UNIVERSITÀ DEGLI STUDI FIRENZE

PhD in
Earth Sciences
XXXV CICLO – *35th Cycle*

Spatial and temporal analysis of interferometric data for ground deformation detection at regional scale

Doctoral Candidate

Dr.

Davide Festa

Supervisor	**Co-Supervisor**
Prof.	*Prof.*
Nicola Casagli	Federico Raspini

Coordinator

Prof. Sandro Moretti

2019-2022

Table of Contents

Abstract

Aerial mapping and Earth Observation has undergone a great revolution with the advent of Synthetic Aperture Radar (SAR) instruments since the early 90's. SAR acquires images of the Earth in the microwave spectrum, which is able to penetrate clouds, thus making it an all-weather remote sensing system operating day and night. Especially, SAR interferometry (InSAR) is a powerful technique based on radar imagery targeted for surface deformation modelling, from natural hazard mapping to monitoring subsidence and stability in structural engineering, with a precision of millimetre-level scale.
Nowadays, the use of InSAR-derived products for hydrogeological events monitoring and mapping is well established in the scientific community and began to be recognized as a valuable tool for risk management and reduction by administrative entities and Civil Protection authorities.

Within this Thesis work, three complementary approaches for maximizing the interpretation and management of radar-based information in the framework of multi-hazard mapping and analysis are proposed. Large multi-temporal interferometric datasets resulting from different processing algorithms are implemented into post-processing and value-adding routines for extracting readily usable information from Earth observation analytics. This Thesis work represents the main outcome of a three years-long activity at the Department of Earth Sciences of the University of Florence and its goal is to test and evaluate the potential and applicability of operational tools for the characterization of natural and man-made hazards by intensively exploit large amounts of space-borne SAR products.

The core of the Thesis hinges on illustrating different unsupervised and supervised procedures based on feature engineering, Data Mining and Machine Intelligence whose tasks can be merged in a complementary workflow aiming at detecting areas of active motions, spatially and temporally characterize them in detail and cross-correlate the satellite data with other available source of information, making the obtained products and results useful for the end-user.

The first of the three approaches rely on a semi-automatic procedure to quickly identify, map and inventory ground and infrastructures displacements by means of spatial clustering. The case of study regarded ascending and descending deformation maps covering the whole Italian territory and processed through the Parallel Small BAseline Subset (P-SBAS) technique. The detection of deforming areas is followed by an evaluation against the horizontal and vertical components of motion retrieved from a Line of Sight (LOS) vector decomposition and by a cross-correlation with accessible ancillary layers for a preliminary classification of the triggering factors. By setting a displacement rate threshold of ± 1 cm/year, a total number of 14,638 areas resulting from both geometries are found to suffer from instability phenomena in the period spanning from March 2015 to December 2018. The origin of active areas of deformation are tentatively inferred and preliminary sorted in 11 classes split between natural causes and man-made activities such as landslides, volcanic-related deformations, subsidence, and mining-related displacements.

The second procedure focuses on effectively detect high-displacement areas and classify ground motion sources, by exploring the feasibility of a Machine Learning-based approach. This is achieved by applying the Random Forest (RF) technique to deformation maps covering the northern part of Italy and by training the RF model to automatically identify landslide, subsidence, and mining-related ground motion on the

bases of an expertly designed training dataset. The presence of noisy signals and other sources of deformation is also tackled within the model construction. A set of common predictors and explanatory variables extracted from the A-DInSAR datasets and from freely accessible informative layers such as Digital Elevation Model (DEM), land cover maps, and geohazard inventories are exploited for testing the potentialities of prediction of such model. The model performances could achieve promising results given an overall accuracy of 0.97, a true positive rate of 0.94 and an F1-Score of 0.93, thus demonstrating the viability of the proposed procedure.

The third methodology focused on the temporal dimension of A-DInSAR products. It is proposed an automated and purely statical-based rationale for the InSAR Time Series clustering based on Principal Component Analysis (PCA) dimensionality reduction. The popular unsupervised Machine Learning approach called K-Means, the least squares Linear Regression and the Basic Fast Fourier Transform (FFT) are also included in the pipeline of the proposed workflow. The above-mentioned is applied prior a reprojection of the LOS displacement series along the vertical and eastward directions performed for ascending and descending A-DInSAR datasets covering Valle d'Aosta region (North-western Italy), where the applied data fusion rationale is robust about the spatial and temporal dimensions. Ultimately, the retrieved cluster series are further broken down into their trend and seasonality components to enhance the readability of the classified features.

Riassunto esteso

A partire dai primi anni '90, la mappatura e l'osservazione della Terra hanno subito una grande rivoluzione con l'avvento degli strumenti di Synthetic Aperture Radar (SAR). Il SAR acquisisce immagini della Terra nello spettro delle microonde, le quali sono in grado di penetrare le nuvole, rendendolo così un sistema di telerilevamento attivo per tutte le stagioni, giorno e notte. In particolare, l'interferometria SAR (InSAR) è una potente tecnica basata su immagini radar mirata alla modellazione delle deformazioni superficiali, dalla mappatura dei rischi naturali al monitoraggio della subsidenza e della stabilità nell'ingegneria strutturale, con una precisione di scala millimetrica.
Oggi, l'uso di prodotti derivati dall'InSAR per il monitoraggio e la mappatura di eventi idrogeologici è ben consolidato nella comunità scientifica ed è ormai riconosciuto come un valido strumento per la gestione e la riduzione del rischio da parte di enti amministrativi e autorità di Protezione Civile.

In questo lavoro di tesi, vengono proposti tre approcci complementari per massimizzare l'interpretazione e la gestione delle informazioni di origine radar nel quadro della mappatura e dell'analisi multirischio. Grandi moli di dati interferometrici multi-temporali risultanti da diversi algoritmi di elaborazione sono qui implementati in routine di post-elaborazione per estrarre informazioni prontamente utilizzabili e relative all'osservazione della Terra. Questo lavoro di tesi rappresenta il risultato principale di un'attività triennale presso il Dipartimento di Scienze della Terra dell'Università di Firenze e il suo obiettivo è quello di testare e valutare il potenziale e l'applicabilità di strumenti operativi per la caratterizzazione dei rischi naturali e antropici attraverso lo sfruttamento intensivo di grandi quantità di prodotti SAR spaziali.

Il nucleo della tesi è incentrato sull'illustrazione di diverse procedure non supervisionate e supervisionate basate su Data Mining e Machine Intelligence, i cui compiti possono essere fusi in un flusso di lavoro complementare volto a individuare aree di movimenti attivi, a caratterizzarle spazialmente e temporalmente in dettaglio e a correlare i dati satellitari con altre fonti di informazione disponibili, rendendo i prodotti e i risultati ottenuti utili per l'utente finale.

Il primo dei tre approcci si basa su una procedura semi-automatica per identificare, mappare e inventariare rapidamente gli spostamenti del terreno e delle infrastrutture mediante clustering spaziale. Il caso di studio ha riguardato mappe di deformazione ascendente e discendente che coprono l'intero territorio italiano e sono state elaborate attraverso la tecnica Parallel Small BAseline Subset (P-SBAS). L'individuazione delle aree deformanti è seguita da una valutazione rispetto alle componenti orizzontali e verticali del movimento recuperate da una decomposizione dei vettori di spostamento direzionati lungo la linea di vista satellitare e da una cross-correlazione con strati ancillari accessibili per una classificazione preliminare dei fattori scatenanti. Impostando una soglia di tasso di spostamento di ± 1 cm/anno, si è riscontrato un numero totale di 14.638 aree risultanti da entrambe le geometrie interessate da fenomeni di instabilità nel periodo che va da marzo 2015 a dicembre 2018. L'origine delle aree attive di deformazione è stata dedotta in via provvisoria e ordinata in 11 classi, suddivise tra cause naturali e attività antropiche, quali frane, deformazioni di origine vulcanica, subsidenza e spostamenti di origine mineraria.

La seconda procedura si concentra sull'individuazione efficace di aree in deformazione e sulla classificazione delle sorgenti del moto del suolo, esplorando la fattibilità di un approccio basato sul

Machine Learning. Questo obiettivo è stato raggiunto applicando la tecnica Random Forest (RF) a mappe di deformazione che coprono la parte settentrionale dell'Italia e addestrando il modello RF a identificare automaticamente il movimento del suolo legato a frane, subsidenza e miniere sulla base di un dataset di addestramento progettato da esperti. La presenza di segnali rumorosi e di altre fonti di deformazione viene affrontata nella costruzione del modello. Un insieme di predittori comuni e di variabili esplicative estratte dai dataset A-DINSAR e da strati informativi liberamente accessibili, come il modello digitale di elevazione (DEM), le mappe di copertura del suolo e gli inventari di rischio geologico, sono sfruttati per testare le potenzialità di previsione di questo modello. Le prestazioni del modello hanno permesso di ottenere risultati promettenti, con un'accuracy complessiva di 0.97, un true positive rate di 0.94 e un F1-Score di 0.93, dimostrando così la validità della procedura proposta.

La terza metodologia si è concentrata sulla dimensione temporale dei prodotti A-DInSAR. In particolare, viene proposta una routine automatizzata e puramente statica per il raggruppamento delle serie temporali InSAR basata sulla riduzione della dimensionalità tramite analisi delle Componenti Principali (PCA). Il popolare approccio di apprendimento automatico non supervisionato chiamato K-Means, la regressione lineare ai minimi quadrati e la trasformata veloce di Fourier (FFT) sono inclusi nella pipeline del flusso di lavoro proposto. Il procedimento è applicato prima di una riproiezione delle serie di spostamenti LOS lungo le direzioni verticali e orizzontali, eseguita per set di dati A-DISAR ascendenti e discendenti che coprono la regione della Valle d'Aosta (Italia nord-occidentale), dove la logica di fusione dei dati applicata è robusta rispetto alle dimensioni spaziali e temporali. Infine, le serie di spostamento relative ai cluster ottenuti sono ulteriormente suddivise nelle loro componenti di trend e stagionalità per migliorare la leggibilità degli oggetti classificati.

Glossary of frequently used acronyms

Acronym	Meaning
A-DInSAR	Advanced Differential Interferometric Synthetic Aperture Radar
CPA	Civil Protection Authorities
DInSAR	Differential Interferometric Synthetic Aperture Radar
DM	Data Mining
EO	Earth Observation
InSAR	Interferometric Synthetic Aperture Radar
LOS	Line of Sight
MAC	Moving Area Cluster
ML	Machine Learning
MP	Measurement Point
PS	Persistent Scatterer
P-SBAS	Parallel Small Baseline Subset
PSI	Persistent Scatterer Interferometry
RF	Random Forest
S1	Sentinel 1
SAR	Synthetic Aperture Radar
TS-InSAR	Time Series Interferometric Synthetic Aperture Radar

1. Introduction

1.1. Background

Over the past few decades, the flourishing of remote sensing technologies usually involving spaceborne sensors have enshrined the rise of a breakthrough change in the way natural and manmade environment is observed and monitored. Earth observation (EO) data, due to the constant development of space-based technologies, to the increased awareness of climate change impacts and to the challenge of consistently disseminate timely geospatial information, are now at the forefront in the gathering of information about planet Earth's physical, chemical, and biological systems (Sabeur et al., 2018). The ability of satellites to deliver up-to-date and repeat-coverage imagery to the public coupled with the scientific knowledge of turning large amounts of EO data into valuable global information layers have determined an unprecedent contribution to a comprehensive risk mitigation planning and improved disaster response. The current scenario is well underway: the expertise in remote sensing technologies have witnessed the feasibility of a near-real time assessment and global sharing of the state and evolution of our environment during or after natural disaster events.

Prior to the advent of consistent implementation of Earth observation technology, the scientific community collected information about Earth processes predominantly through ground-based sensors, resulting in an insufficient density of networks and a very limited coverage and resolution of data.
Recently, the massive amount of remotely sensed data has brought a shift in the approach to data curation, being ingest processing and information sharing a great challenge considering both temporal and spatial resolutions (Probst et al., 2017). Therefore, Big Data in EO has constituted an emerging paradigm to describe the range of new techniques for the development and application of new image processing techniques for an adequate and computationally efficient exploitation of remotely sensed scenes from a Big Data point of view. Moreover, EO Big Data is concerned with quantitatively analysing complex geoscience processes and spatial-temporal systems characterizing socioeconomic processes (Guo et al., 2015). Therefore, EO research is expected to boom over the next decades while already involving different branches of expertise converging in the newly defined role of the EO data scientist.

Nowadays, the operating space-borne sensors cover a broad portion of the electromagnetic spectrum from visible light and infrared to microwave, where increasingly refined spatial and spectral resolution are integrated with data sampling frequencies that allows the monitoring of the cyclic changes of terrestrial, oceanic, and atmospheric system processes. In 2008, more than 150 EO satellites were in orbit, recording data with both passive (i.e., measuring reflected solar radiation or emitted thermal energy) and active sensors (i.e., emitting energy and recording the reflected or backscattered response) and acquiring more than 10 terabits of data daily (Tatem et al., 2008). By April 2021, the total number of satellites had grown to over 950, according to UCS (Union of Concerned Scientists; ucsusa.org). Most of the EO satellite platforms operates at Low Earth orbit (LEO), namely at an altitude of less than 1000 km from the Earth's surface, thus guaranteeing to take images of higher resolution. Concerning LEO satellites devoted to Earth mapping and monitoring, it is used a specific polar orbit called Sun-Synchronous Orbits (SSO) whose high inclination (ca. 98°) makes the satellite's revisit time longer for equatorial areas than for polar regions. The spatial and temporal coverage is often increased by means of constellations, which consist of equal or similar satellites that are launched together to create a net around Earth. For instance, between the

operational satellites mounting passive sensors, the Copernicus Sentinel-2 mission comprises of a constellation of 2 twin satellites carrying an optical instrument that samples 13 spectral bands (from 10 to 60 m spatial resolution) at a revisit frequency of 5 days at the equator and provides multispectral observations in continuity with QuickBird, SPOT and LANDSAT program. On the other hand, one of the most common active sensors used for EO are those emitting a radar (Radio Detecting and Ranging) signal that can penetrate cloud covers and that can operate both day and night. Those instruments, known as Synthetic Aperture Radar (SAR) sensors, transmits electromagnetic pulses which are reflected or scattered back by the Earth's surface objects. The instrument is provided with an antenna that can detect and record the return pulses within discrete wavelengths in the microwave domain (such as X, C and L bands, from the shortest to the longest wavelength), allowing for the generation of SAR imagery.

SAR acquisition contains a twofold level of information: amplitude and phase. Amplitude is the strength of the radar backscattered response, while phase is the fraction of a single sine wave cycle. The phase of the SAR image primarily related to the-sensor-to-target distance. Through the comparison of the phase of successive radar images acquired on the same target area it is possible to extract distance information about the Earth's terrain: this approach is called Interferometric SAR (InSAR; Bamler & Hartl, 1998b). InSAR constitutes the most used technique to retrieve highly accurate and spatially dense measurements regarding the Earth's surface deformation. InSAR can be used in the context of terrain motion measurement in case of natural hazard such as in the event of subsidence, landslide, earthquake, or it can also be applied for infrastructure, civil construction, mining, or gas reservoirs monitoring.

Starting from the early 2000, many processing algorithms classed as Persistent Scatterer Interferometry (PSI; (Crosetto et al., 2010, 2016; Gernhardt et al., 2010) were developed to obtain the spatial distribution and the temporal evolution of land displacements from local to large scale. The PSI represents a branch of interferometry whose main outcome is to provide a map of ground-based targets, typically with a strong and constant radar backscatter signal over time, for each of which it is retrievable a time series providing the deformation history over the observed period. The capability to sample deformation phenomena over time depends on the characteristics of the acquisition process of operating SAR sensors, determining the availability of SAR data. The active SAR missions nowadays operating are marked by uneven strategies regarding the acquisition plan, the data policy and the sensor acquisition band: for instance, the four-satellite CosmoSkyMed constellation, TerraSAR-X operates in X-band with a revisit time of 16 and 11 days over the same area, while Sentinel-1 (C-band), Radarsat-2 (C-band) and ALOS-2 (L-band) having a revisit time of 12, 24 and 14 days, respectively. In general, it has been widely proved that PSI-derived products are highly sensitive, precise, and accurate, therefore they demonstrated to be a primary tool for the extensive mapping and monitoring of deformation processes affecting Earth's surface (Capes & Teeuw, 2017).

Regularly updated multi-temporal interferometric data covering large area of the planet have represented a great revolution in the EO scenario. In this framework, in this thesis it is introduced the new paradigm "Big SAR data". Relevant efforts towards the extensive exploitation of massive SAR archives have already been carried by research teams and commercial companies in terms of analysis and monitoring services at a various scale of analysis. In this context, a cutting-edge Copernicus initiative named the European Ground Motion Service (EGMS, european-ground-motion-service) will provide, starting from September 2022, multi-temporal interferometric products over Europe with a one year-update plan, thus constituting

a baseline for ground motion applications at continental, national and local level. The latest technological developments in space infrastructures (Upstream) coupled with performing data processing techniques and valuable data distribution approaches (Midstream) have resulted in unprecedented quantities of monitoring data. The need for implementing efficient value-adding tools for the optimization of data usage (Downstream), compels us to find innovative solutions for the automated management of enormous streams of data which can no longer be investigated manually. Therefore, the above-mentioned circumstances must be actively intercepted by analysts to bring an evolution of EO towards the designing of approaches benefitting from Artificial Intelligence (AI) to support entities in charge of natural risk management and global environmental change mitigation.

1.2. Scope of the research

This Thesis work represents the culmination of a three years-long research activity at the Department of Earth Sciences of the University of Florence, with a 6-month period of collaboration with the Remote Sensing group of the British Geological Survey (BGS) based in Keyworth (Nottinghamshire, United Kingdom). The research period with BGS was conducted remotely in first instance, then it consisted of a visiting fellowship of 3 months at the Keyworth offices. The project was partially funded by "Presidenza del Consiglio dei Ministri–Dipartimento della Protezione Civile" (Presidency of the Council of Ministers–Department of Civil Protection), through the IREA-CNR/DPC and CPC-UNIFI agreements and also by the IREA CNR/Italian Ministry of Economic Development DGS-UNMIG agreement and H2020 EPOS-SP (GA 871121).

The project aims at providing Civil Protection Authorities (CPA) and involved stakeholders with the capability of periodically evaluating and assessing the potential impact of geohazards and human-induced deformations of the Earth's surface. This objective will be based on the performances of the spaceborne SAR sensor Sentinel 1, both in terms of wide area coverage and high temporal repeatability.

The goal of this Thesis work is to test and evaluate the feasibility of automated procedures for the management and for the analysis of large datasets of space-borne SAR data, processed by means of different PSI approaches, as operational tools for the spatial and temporal characterization of geohazards in different geological and geomorphological environments. In particular, three different and complementary approaches are proposed for the rapid detection, classification and characterization of the historical motion trend of areas of the territory suffering from significant surface displacements. Landslides, subsidence, mining-related displacements and many other natural or anthropogenic phenomena are evaluated with reference to their impact in terms of ground deformation intensity, deformation pattern and spatial occurrence.

In this framework, the main tasks and implementations of the proposed analysis methodologies are:

- Rapid mapping from large to local scale of areas affected by active ground motions for the construction of continuously updated geohazard inventories ideally suited for territorial planning activities;

- Exploitation of multiple sources of informative and ancillary layers for the pre-screening and assessment of the deformation cause of active areas;

- Statical-based methods for the comprehensive analysis of the motion trend of the Earth's surface over time to detect both seasonal and multiannual surface displacements and their driving processes, in the framework of geohazard mitigation procedures within civil protection activities.

1.3. Thesis structure

Besides the Introduction, Discussions, Conclusions and References, this Theis work is subdivided into different chapters which are organized as follows:

CHAPTER 2. Description of the basic principle of spaceborne radar remote sensing, status of past and recent SAR missions, fundamentals of SAR interferometry, Differential Interferometry and reviews of the actual interferometric processing techniques (both single pairs and multi-interferograms).

CHAPTER 3. Delineation of the main state-of-the-art systems and platforms for EO data consultation with major focus on InSAR-based infrastructures for geohazard monitoring.

CHAPTER 4. Brief description of the potentialities deriving from the consistent post-processing analysis of large A-DInSAR data stream. Here are outlined the proposed approaches for the downstream conversion of interferometric data into value-added products.

CHAPTER 5. Description of post-processing methodology based on clustering radar-targets for the unsupervised hotspot mapping and screening of deforming areas from a nation-wide interferometric dataset covering the Italian Peninsula.

CHAPTER 6. Presentation of an original supervised Machine Learning approach testing the feasibility of statistically combine different morphometric variables and ancillary layers for the automatic recognition of the triggering factor responsible for the displacements detected from very large InSAR deformation maps.

CHAPTER 7. Description of an automated unsupervised procedure designed for the efficient Data Mining (DM) of InSAR Time Series of displacements by means of Python scripting. The main temporal motion trends are statistically extracted, broken down into their main temporal components and deployed for the clustering of radar features.

2. Radar-based remote sensing

Radar is an acronym that stands for "RAdio Detection And Ranging". A radar is an active sensor that transmits microwave radio pulses and receives the backscattered radiation to detect the illuminated targets. The returning signals is processed to quantify the parameters of the backscattered signals produced by the targets (Emery & Camps, 2017). Synthetic aperture radar (SAR) is a form of radar mounted on a moving platform, such as an aircraft or spacecraft. The distance the SAR device travels over a target during the period when the target scene is illuminated creates the large synthetic antenna aperture (a moving antenna). By combining a sequence of acquisitions, SAR allows to create high-resolution images by simulating a much larger antenna than the actual physical one deployed on the space-borne platform.

SAR products are primarily used to monitor surface displacements over large areas and thus assist hazard analysis and spatial planning. A satellite-based SAR sensor scans the Earth's surface by means of electromagnetic microwave pulses. The SAR antenna scans the terrain below the platform through sending microwave pulses and by receiving and. Subsequent signal processing (i.e., the backscattered echoes from the below terrain) allows for a reconstruction of a high-resolution image of the observed scene. SAR imaging is of particular importance for geoscientific applications given the capability of microwaves to penetrate the atmosphere under virtually all conditions (all-weather and day or night) compared to other bands of the electromagnetic spectrum.

Wavelength and length of the sensor's antenna constitute important parameters in determining the signal interaction with the surface, how far a signal can penetrate a medium, and the resolution of an acquiring radar system. In particular, the longer the antenna, the higher the spatial resolution of a radar sensor in the azimuth direction (R_{azimu}), parallel to its orbit(1):

$$R_{azimu} = R \times b$$

(1)

where R is the-sensor-to-target distance and b (2) is defined as:

$$b = \frac{\lambda}{L}$$

(2)

where b is defined as the width of the radar beam, λ is the wavelength of the signal and L is the length of the sensor's antenna. Therefore, from a satellite in space operating at a wavelength of about 5 cm (C-band radar), it would be required a radar antenna of about 4,250 m long in order to get a spatial resolution of 10 m. Concerning the aperture of the antenna, it relates to how long the target area is illuminated, which in turns reflects the achievable resolution (the longer the antenna, the more resolution on the ground). During the forward movement of the satellite, the target scene is repeatedly illuminated from different positions by a single beam antenna, few meters long (Figure 1). The several acquisitions are recorded, stored, combined, and then processed together, simulating a "synthetic aperture", to provide the required azimuth resolution (Hooper, 2006).
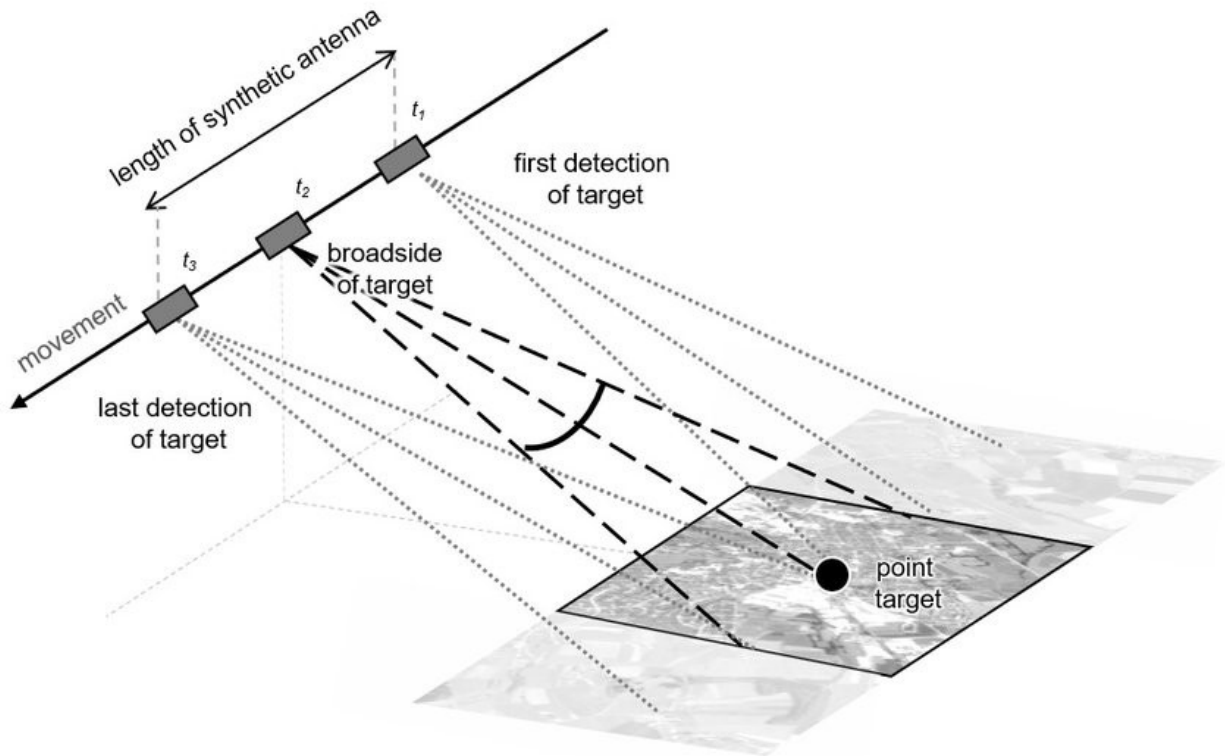
Figure 1. Principle of synthetic aperture radar (source: Braun, 2019).

Radar sensors utilize wavelengths at the centimetre to meter scale (K_a to P band respectively; Figure *2*) which makes the sensor suitable for different purposes (Table 1). Every sensor emits its signal at a specific central frequency (operating frequency) that defines the signal propagation and penetration features. For satellite applications, the three most commonly used bands are: C-band (4-8 GHz, 7.5–3.8 cm wavelength), X-band (8-12 GHz, 3.8–2.4 cm wavelength) and L-band (1-2 GHz, 30–15 cm wavelength). C-band is the SAR workhorse, and its pulses are capable of imaging through tropical clouds and rain showers. However, it holds a limited ability to penetrate vegetation canopies or soils. L-band has a lower resolution than C- and X-band images but are more coherent over time, especially in vegetated regions. X-band, despite being more prone to coherence decay in vegetated areas, has the highest imaging resolution with direct applications for ice and snow monitoring (Table 1).
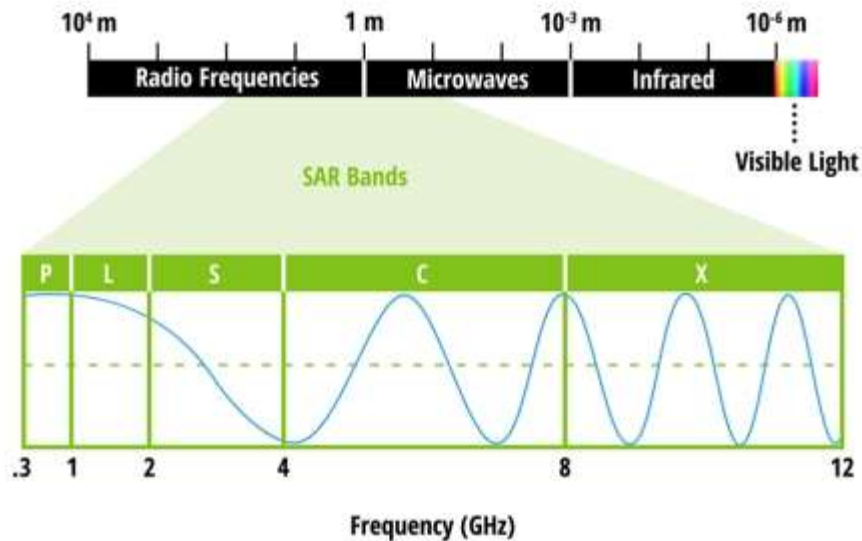
Figure 2. Representation of the different SAR bands and the associated frequency and wavelengths (source earthdata.nasa.gov).

Table 1. The table below denotes the SAR deployed bands with associated frequency, wavelength, and the application typical for that band.

Band	Frequency	Wavelength	Typical application
Ka	27–40 GHz	1.1–0.8 cm	Rarely used for SAR (airport surveillance)
K	18–27 GHz	1.7–1.1 cm	rarely used (H_2O absorption)
Ku	12–18 GHz	2.4–1.7 cm	rarely used for SAR (satellite altimetry)
X	8–12 GHz	3.8–2.4 cm	High resolution SAR (urban monitoring; ice and snow, little penetration into vegetation cover; fast coherence decay in vegetated areas; InSAR)
C	4–8 GHz	7.5–3.8 cm	SAR most used band (global mapping; change detection; monitoring of areas with low to moderate penetration; higher coherence; InSAR); ice, ocean maritime navigation
S	2–4 GHz	15–7.5 cm	Little but increasing use for SAR-based Earth observation; agriculture monitoring (expends C-band applications to higher vegetation density)
L	1–2 GHz	30–15 cm	Medium resolution SAR (geophysical monitoring; biomass and vegetation mapping; high penetration, InSAR)
P	0.3–1 GHz	100–30 cm	Biomass is a P-Band satellite mission expecting to be launched in August 2023; vegetation mapping and assessment. Experimental SAR.

A satellite SAR sensor travels along its path through a predefined polar orbit with the transmitting antenna towards to the ground surface (Hooper, 2006)thus observing the same scene from the East in the descending orbit and from the West when capturing images in ascending orbit. The radar signal hits the assumed flat Earth's surface with an angle which is equal to the inclination of the antenna with respect to the nadir direction (Off-nadir angle; Figure 3). The Off-nadir angle it is also referred to as look angle. The angle between the radar beam centre and the normal to the local topography is the incidence angle, which

can be considered equivalent to the Off-nadir angle by neglecting the Earth's curvature and the local topography. Usually SAR sensors are defined as "side-looking", since radar antenna beam illuminates the ground (antenna footprint) to the right side of the satellite along the flight direction, which is called azimuth direction. The sensor-to-target distance is called slant range, which, if projected on the ground represents the ground range. The direction of the main lobe of the transmitting antenna is called range direction. The width of the illuminated area is called swath (Figure *3*). Polarization can be vertical or horizontal and this refers to the orientation of the plane in which the SAR sensors typically transmit the electromagnetic waves.

An imaging SAR system has different standard acquisition scheme which are designed for different applications but also on the desired coverage and data resolution. The different acquisition systems differ by their geometrical configuration of the SAR antenna or by their temporal properties. Regarding the Stripmap mode (SM; Figure *3*), the antenna usually looks sideways down in a direction perpendicular to the flight direction, with a multi-beam antenna that can illuminate different sub-swath through different beams (Guarnieri & Prati, 1996). Spotlight-SAR is a mode for obtaining high-resolution by steering the radar beam to keep the target within the beam for a longer time implying the formation of a longer synthetic aperture, thus providing a great enhancement in imaging resolution. However, this mode of operation is usually at the expense of spatial coverage. Another acquisition scheme is the ScanSAR, consisting in obtaining wide-swath coverage by periodically switching the antenna elevation beam to points in several range sub swaths, which are acquired quasi simultaneously. The drawbacks associated with that acquisition mode are overcame by the Interferometric Wide swath (IW) mode, which implements a new type of ScanSAR mode called Terrain Observation with Progressive Scan (TOPS) SAR (De Zan & Monti Guarnieri, 2006). Several sub swaths are acquired by TOPS mode, with the increased coverage that is obtained through a reduced azimuth resolution and a nearly uniform signal-to-noise ratio. Like IW mode, Extra Wide Swath Mode (EW) mode uses the TOPSAR technique to cover a very wide area at medium resolution (Figure *3*). A less common acquisition mode is Wave Mode (WM).
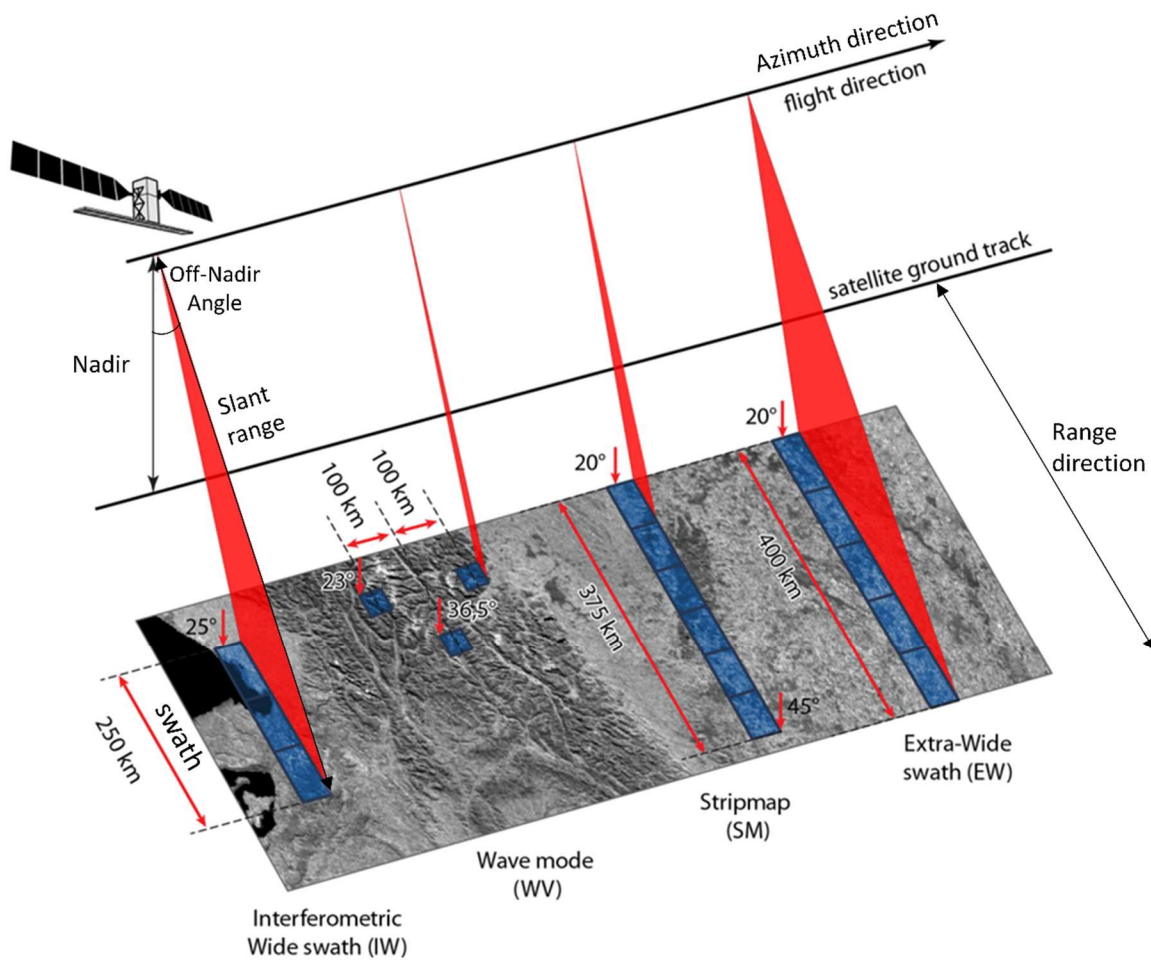
Figure 3. SAR instrument acquisition modes (source: ESA 2016, modified).

The acquisition of SAR images is done in slant range coordinates, corresponding to the radar coordinate system. Consequently, ground elevation (i.e., mountain slopes with high relief) will cause geometric distortions in the radar imagery. The results are clearly visible as effects on radar images (Figure 4) with the compression of features in the near-range relative to features in the far range. These distortions, whose appearance and information can significantly vary in function of the incidence angle respect to the surface morphology, can be distinguished into Foreshortening, Shadowing and Layover. Foreshortening is caused by the shorter travel time of the radar pulse, as the slant range distance to an object on the inclined slope is shorter than to an object on flat areas. The fore-shortened areas appear brighter in the SAR image, due to the compression into a smaller image part of the backscattered slope (Figure 5). On the other hand, when the terrain slope is close to the radar off-nadir angle, the cell dimension becomes very large, and all the details are lost. This happens when topographic surfaces on the opposite side of incoming beam generally are not illuminated, and they suffer of Shadowing (dark-toned). The steeper the sensor incidence angles and the terrain slope, the more pronounced is the effect. Lastly, it can happen that the terrain slope exceeds the radar off-nadir angle, and so the back-scattered signals are imaged in reverse order and superimposed on the contribution coming from other areas. For example, layover can be observed when, in very steep slopes, the radar pulse first hits the top of a mountain and then its foot. Consequently, the top is imaged in front of the mountain's foot, leading to great geometric distortions in the radar image.

9

Figure 4. Geometric distortions effects visible on SAR images acquired over the same area but from different geometries of acquisition (i.e., ascending, and descending orbits).

Figure 5. Type of distortions effects connected to the SAR system incidence angle in relation with the surface morphology (source: Qin et al., 2020. Modified)

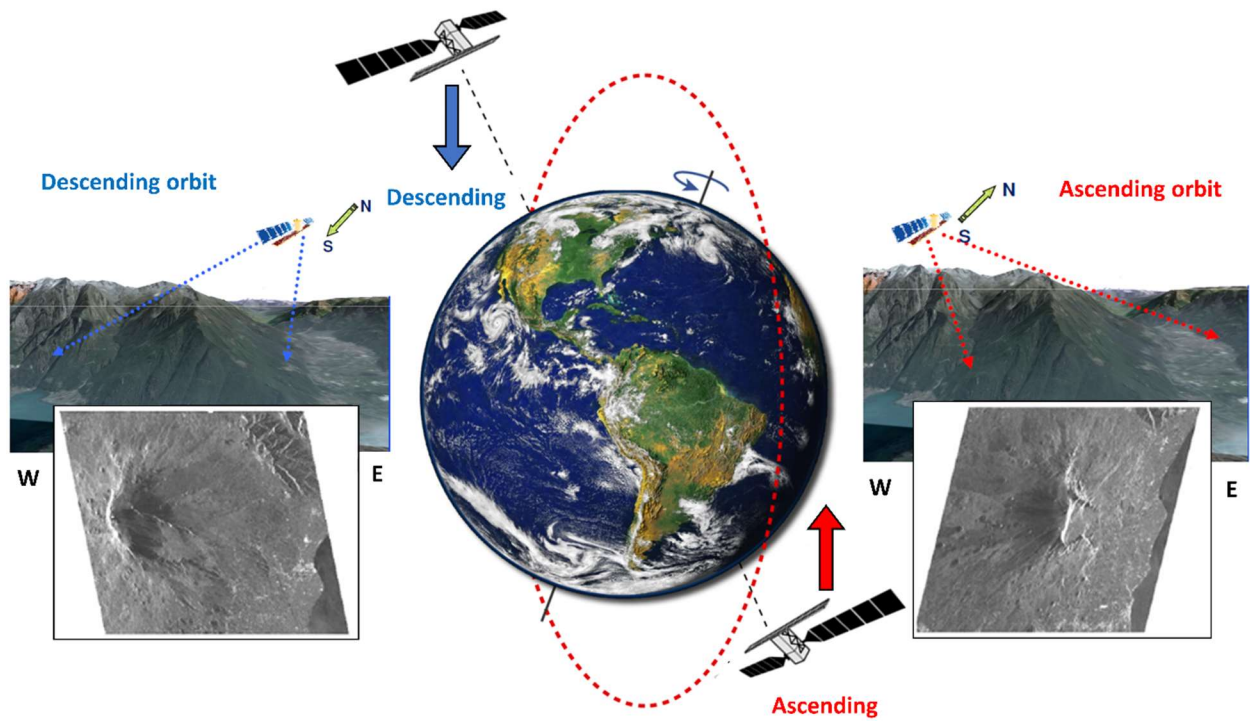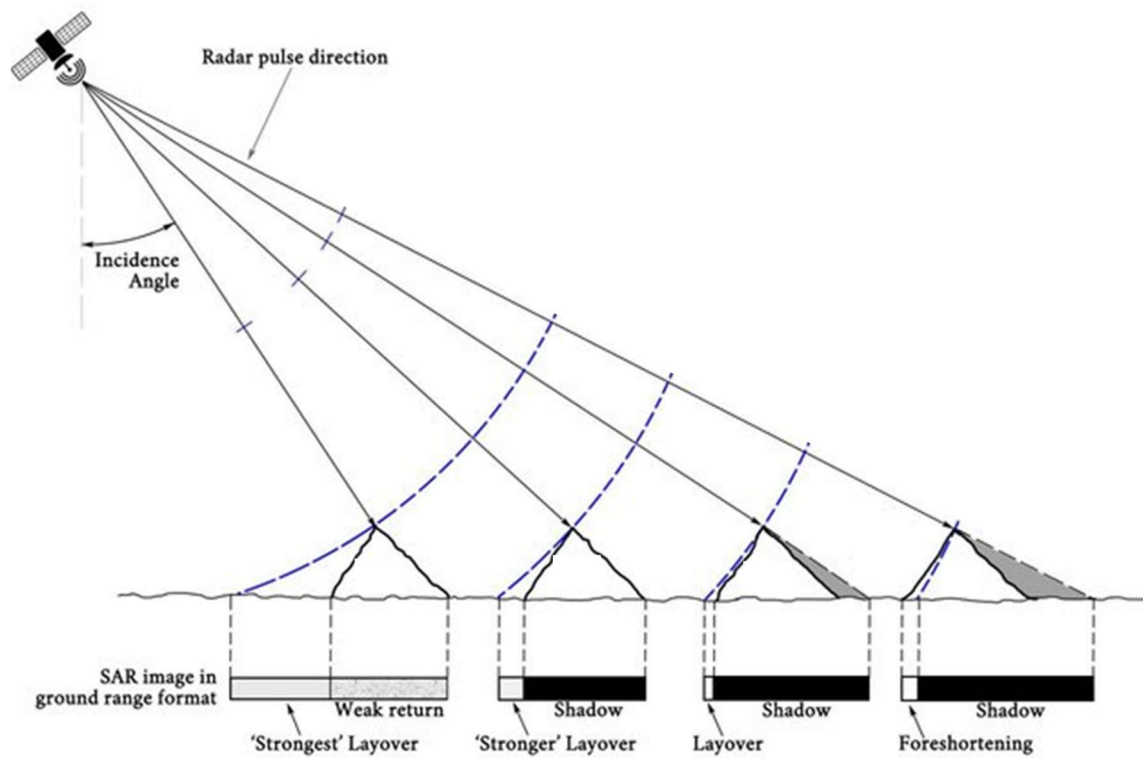It is worth mentioning also the so-called radiometric distortions. It is associated with the combination of the back-scattered signals of distinct objects into a single pixel, which is compressed in a smaller area than the original one. This effect, known as Speckle, is because registered echoes can add both constructively and destructively, creating a noise effect attributed to random interference between the coherent returns issued from the numerous scatterers present on the surface (Gagnon & Jouan, 1997).

2.1. Spaceborne SAR missions

The first satellite designed for the observation and the study of oceans implemented with SAR instruments was SeaSat-1, launched in 1978. Starting from the 1992, several satellite-based SAR sensors have been launched, providing images acquired at different wavelengths and that can be used for ground deformation monitoring. Indeed, the first ESA (European Space Agency) Earth observation mission begun with the ERS 1/2 experience. ESA launched ERS-1 in July 1991 and ERS-2 in April 1995. These two twin satellites were provided with a range of instruments such as a SAR sensor, a radar altimeter and ocean surface temperature and sea winds instruments, with ERS-2 implemented with an extra instrument (GOME) designed to monitor ozone levels in the atmosphere. ERS-1 and ERS-2 mounted a radar C-band sensor and with an incidence angle of 23° and a resolution of 30m x 26m. ERS-1, after acquiring more than 1.5 million individual SAR scenes, was deactivated in March 2000, while ERS-2 was deactivated in 2011 after 16 years of successful operations.

The JERS-1 mission, instead, started in February 1992, led by the National Space Development Agency of Japan (NASDA). ERS-1, which ceased operations in October 1998, was an environmental Japanese satellite with a radar/optical sensor whose task was to cover the global land area for national land survey, agriculture, forestry and fishery, and also for the environmental and disaster protection. It acquired images in L-band and with a westward repeat cycle of 44 days.

The RADARSAT mission consists of a pair of remote sensing satellites from the Canadian Space Agency (CSA) meant for operational monitoring services on a global basis ever since. RADARSAT-1 and RADARSAT-2 data were launched in in November 1995 and December 2007, respectively, where the first one was decommissioned in 2013 and the second still being operational. The two satellites were equipped with a C-band SAR. RADARSAT-1 peculiarity was the ability to shape and steer its radar beam over a 500-km range, while RADARSAT-2 have a superior resolution imaging, flexibility in se-lection of polarization, left and right-looking imaging options, and more precise measurements of spacecraft position and attitude.

EnviSAT, which stands for Environmental Satellite, was another ESA mission successor to the ERS missions, and it was launched in 2002 with 10 instruments aboard, such as radar altimeter, temperature-measuring radiometer instruments, medium-resolution spectrometer. The main objective of EnviSAT was the analysis and monitoring of the Earth's environments on various scales, from local to global, with application in the management of Earth's resources and to the field of meteorology. By May 2012, ESA declared the end for the EnviSAT mission, due to a loss of contact with the satellite.

The Advanced Land Observing Satellite (ALOS) was a Japanese satellite from Japan Aerospace Exploration Agency (JAXA) that was launched on 24 January 2006. The instruments with which it was equipped were: L-band SAR, Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM), and an Advanced Visible and Near-Infrared Radiometer-2 (AVNIR-2). ALOS completed its operational phase on 12 May 2011 after failing due to a power anomaly. In May 2014, the ALOS-2 mission started, where the new satellite is provided with a more advanced radar sensor.

The German X-band sensor TerraSAR-X, launched in June 2007 (Meta et al., 2010), was the first SAR sensor acquiring in TOPS mode. TerraSAR-X was joined by TanDEM-X in 2010 and now the twin satellites offer an unmatched offer in terms of geometric accuracy: a multi-mode X-band sensor permits to acquire with a resolution up to 0.25 m (spotlight mode), and moreover they guarantee an excellent radiometric accuracy and rapid switches between imaging modes and polarizations.

Starting from 2007 to 2010, another mission made up of 4 identical satellites mounting X-band SAR sensors have taken place: COSMO-SkyMed is an Italian Earth-imaging constellation owned and operated by ASI (Agenzia Spaziale Italiana) with a programmable system providing different performance characteristics in terms of swath size, spatial resolution, and polarisation configurations. COSMO-SkyMed is intended for the management of environmental and anthropogenic risks, as well as for civil and military use.

Sentinel missions are the last generation of Earth observation satellites of the ESA and the Copernicus group (program of the European Commission). Sentinel satellites carry a range of technologies, starting from radar to multi-spectral imaging instruments for land, ocean, and atmospheric monitoring. TOPS became the standard acquisition scheme of the Copernicus C-band sensor Sentinel-1A (S1A), launched in October 2014 and Sentinel-1B (S1B), launched in April 2016 (Yagüe-Martínez et al., 2016). The mission ended for S1B in 2022 and plans are in force to launch Sentinel-1C as soon as possible. S1A, and the data records of S1B, allowed to create a long-term consistent data archive to be used in the application based on long time series. The consistency in data collection, the short revisiting time (12 days over the same area) and the free data policy have qualified the S1 mission as a primary SAR system for land and ocean services, such as: monitoring and routine sea-ice mapping, surveillance of the marine environment, monitoring land surface for motion risks, mapping of forests, and support to humanitarian aid and crisis situations.

The SAOCOM (SAtélite Argentino de Observación COn Microondas) satellite series are polarimetric L-band SAR (Synthetic Aperture Radar) constellation of two spacecraft. The relative program is defined, managed and operated by CONAE (Comisión Nacional de Actividades Espaciales), Argentina's Space Agency, Buenos Aires. The SAOCOM-1 mission is composed of two satellites, SAOCOM-1A and -1B, which were launched consecutively in October 2018 and August 2020. The overall objective of SAOCOM is to conduct monitoring services for natural and anthropogenic disaster management support and agriculture, mining and ocean applications - including monitoring surveys of Antarctica (study of continental glacier evolution, global change indicators, etc.).

In Figure 6 the main SAR missions are ordered in chronological order, by the type of sensor and by the revisiting time capability of the related satellites.
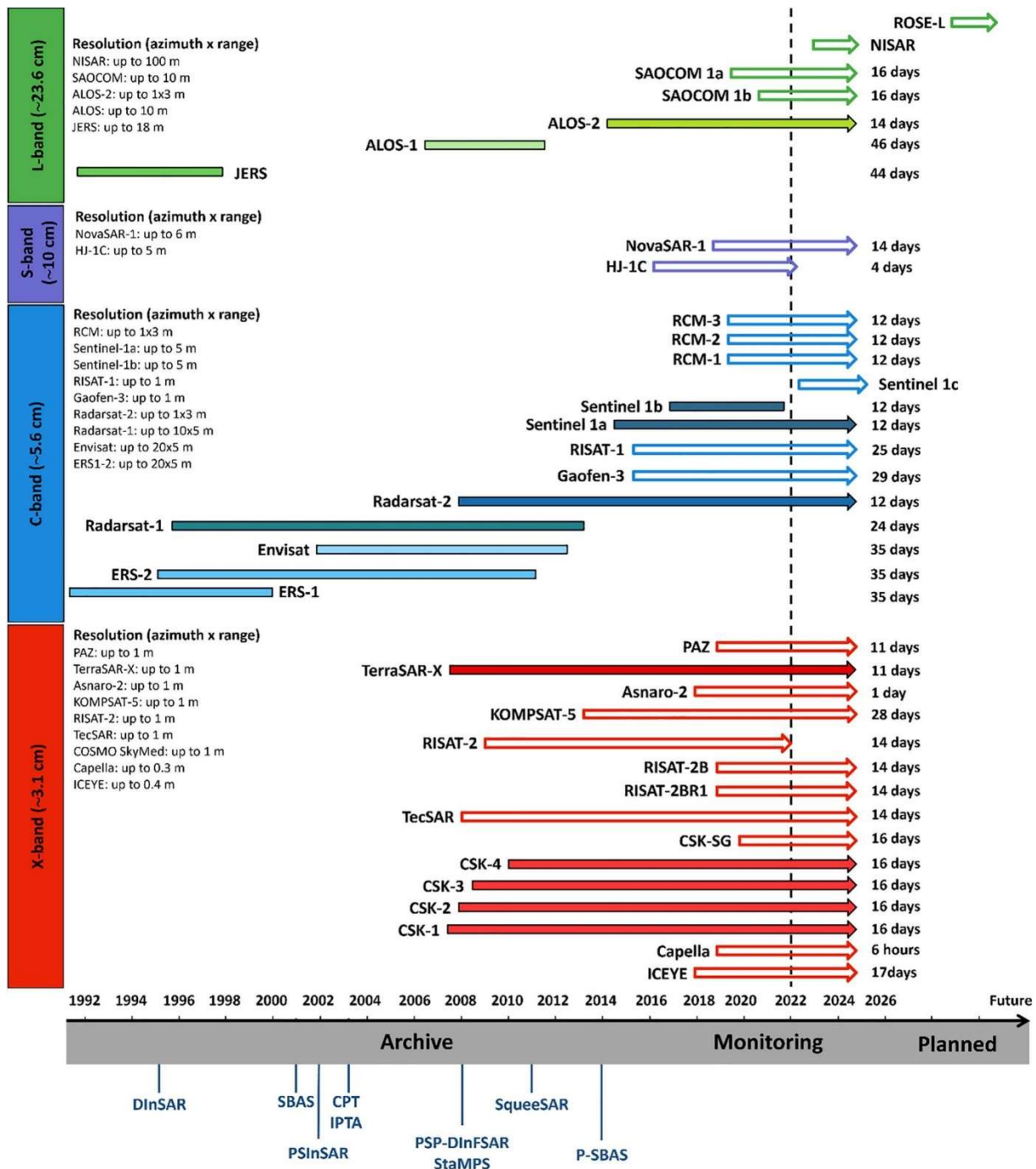
Figure 6. Image depicting the main SAR missions features (Raspini et al., 2022).

2.2. SAR Interferometry

A SAR satellite can acquire over the same area at different times by exploiting repeated orbits of the same satellite or by simultaneously acquire from the same platform with two radars. In this sense, it is possible to "interfere" two SAR images by cross-multiplying, pixel by pixel, the first SAR image with the complex conjugate of the second (Bamler & Hartl, 1998a; Franceschetti et al., 1992; Massonnet & Feigl, 1998): this technique is called SAR interferometry. By combining pixel-by-pixel the phase of these two images

gathered at different times and acquired with slightly different looking angles, an interferogram can be generated where phase is highly correlated to the terrain topography and deformation patterns can be mapped. Finally, the interferogram is built from the complex product of two SAR images (*master* and the *slave* images) and its phase.

Considering a single ground pixel, the variation of the travel path difference (Δr)(3) between the two satellites in the slant-range direction, can be defined with the following equation:

$$\Delta r = -2\frac{B_n q_s}{R}$$

(3)

where: B_n is the perpendicular baseline, R is the distance in slant range between target and sensor and q_s is the distance between two points in perpendicular direction to the slant range.
The interferometric phase variation $\Delta\varphi$ (4) between the two observations is then proportional to Δr divided by the transmitted wavelength λ:

$$\Delta\varphi = \frac{2\pi\Delta r}{\lambda} = \frac{4\pi}{\lambda}\frac{B_n q_s}{R}$$

(4)

The interferometric phase variation $\Delta\varphi$ can be imagined as the result of two contributions (5), namely the (i) phase variation proportional due to altitude difference (q) between the on-ground targets and the (ii) phase variation proportional to the displacement (s) of the point targets along the slant range direction:

$$\Delta\varphi = \frac{-4\pi}{\lambda}\frac{B_n q}{R\sin\theta} - \frac{4\pi}{\lambda}\frac{B_n s}{R\tan\theta}$$

(5)

where θ is the incidence angle and the 4π is referred to the two-way path sensor-target-sensor. From (5) it is computed the so called interferogram flattening, where the perpendicular baseline (B_n) is known from precise orbital data, and the second phase term can be computed and subtracted from the interferometric phase. The flattened interferogram provides an ambiguous measurement of the relative terrain altitude due to the 2π cyclic nature of the interferometric phase. After the flattening, it is performed the reconstruction of the signal's original phase. This is called phase unwrapping and consists in adding appropriate multiples of 2π to each phase input to restore original phase values. Once the interferometric phases are unwrapped, an elevation map in SAR coordinates is obtained (InSAR DEM generation).
Assuming that the targets on the ground change their relative position during two successive SAR acquisition, then a new additive phase term (6) is present within the interferometric phase:

$$\Delta\varphi_d = \frac{4\pi}{\lambda}d$$

(6)

with d representing the point scatterer displacement along the slant range direction. Therefore, the interferometric phase (after the flattening operation) would contain the topographic and motion contributions:

$$\Delta\varphi = \frac{-4\pi}{\lambda}\frac{B_n q}{R\sin\theta} + \frac{4\pi}{\lambda}d$$

<div align="right">(7)</div>

At this point, it is possible to retrieve the terrain displacement component through the so-called differential interferogram, which consists in subtracting the altitude contribution from the interferometric phase by using an already available DEM (Digital Elevation Model).

More precisely, the interferometric phase is the results of the sum of several contributions that may affect the computation of the interferogram. The main purpose of an interferometric process is to separate the different components, extracting only the one related to the displacements by removing or minimizing the others. Having in mind the already mentioned topographic and motion components, the other several phase contributions are illustrated as follows:

- The atmospheric phase component plays an important role for SAR images acquired at different times. Humidity, temperature and pressure affects the radiation travel path producing delays in the signals and reduction of the velocity of the microwaves that contaminate the observations.

- Orbital phase errors are due to small variation of the cross-track and along track components of the orbit during the SAR acquisition: this can result in artifacts in the derived interferograms which are uncorrelated to the real ground deformation (Shirzaei & Walter, 2011).

- Another source of error is the residual topographic error (RTE), which is tied to the local topography and the normal baseline (distance between the two antenna phase centres that illuminate the same area). It should be noted that there exists a critical baseline over which the interferometric phase is pure noise.

- The noise component of the phase, instead, is related to speckle effect or can be associated to the temporal and geometric decorrelation of illuminated scatterers (e.g., vegetated areas or water basins).

- k is an integer value called "phase ambiguity" resulting from the wrapped nature of the interferometric phase, whose estimation is concerned with the correct phase unwrapping.

2.3. Differential Interferometry (DInSAR)

Differential Interferometric Synthetic Aperture Radar (DInSAR) allows to measure terrain deformation from an imaged scene to be derived from two complex SAR images, where the topographic contribution has been removed (Rosen et al., 2000). The term "Differential interferometry" can be misleading because interferometry is a differential technique (i.e., involving the subtraction of the altitude contribution) right from the beginning, however other phases subtraction process can be further executed for surface displacement modelling. DInSAR is recognized to be a great technique for land deformation applications and for the monitoring, the investigation of several types of natural hazards (Gabriel et al., 1989) and for the reconstruction of the surface topography (i.e., DEM creation) as a result of multiple SAR acquisitions. DInSAR is effective with wide areas displacement analysis, the global coverage, and the repeat observation capabilities of SAR satellites.

In the time interval between the acquisition of the two images, the target on the ground changes its relative position and the phase component related to the displacement varies. For every image pixel, the displacement occurred, ranging between $-\pi$ and $+\pi$, is then calculated and the interferogram derived. This "fringe" of coloured pixels cycles represents units of the radar wavelength and represent the difference in signal phase received by the radar instrument over time. It is possible to calculate deformation by since every fringe correspond to a deformation of half the wavelength of the band of the SAR instrument. Taking into account the number of consecutive fringes it is possible to retrieve the cumulated displacement over the observed area. An abrupt colour change in the fringe can be measured relative to the fringe's colour pattern. In Figure 7 it is briefly represented the generation process of an interferogram, over on area encompassing the San Andreas Fault in California. It is possible to detect a colour change in the fringe (pink and yellow meets in correspondence of a fault line) which is caused by a tectonic surface movement that occurred between the observation dates of the two images.
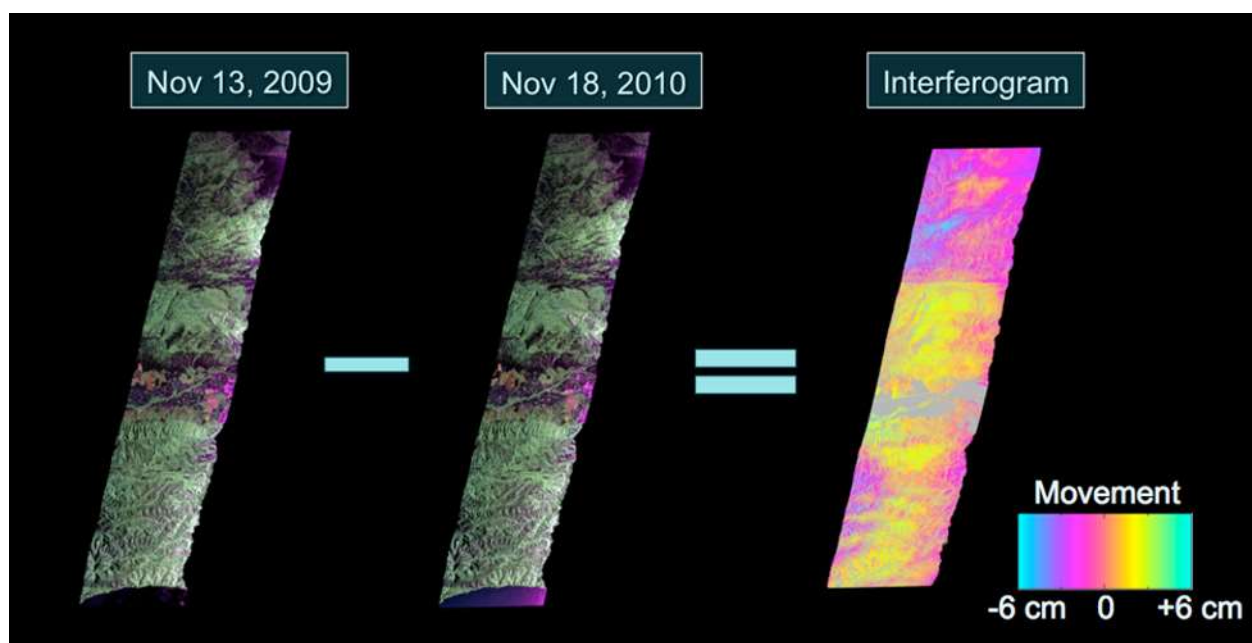
Figure 7. Schematic picture showing basic principle of DInSAR (source: NASA/JPL-Caltech)

By looking at the spread of the fringes it is possible to ascertain the magnitude of the observed ground motion. Close, compact fringes show greater surface movement and wider fringes show gentle to no change as seen from the radar's line of sight. An abrupt change in fringe may be indicative of an earthquake fault ruptured and violently offset, or it can be the consequence of a mountain slope collapse phenomena.

The correct visualization of the displacement fringes can be hindered by a loss of coherence. This can happen when the surface electromagnetic characteristics between the two images used to derive the interferogram are not stable through time (Hanssen, 2001). The coherence decreases when the noise phase component is high (e.g., resulting in radar images affected by speckle). In general, the DInSAR coherence is affected by temporal baseline (longer time intervals between two successive SAR acquisition would decrease the level of coherence), normal baseline (higher distance between orbits of the two images would decrease the level of coherence), the local topography (the presence of steep slopes, oriented parallel to the sensor orbit, reduce the coherence) and the terrain properties (humidity reduce

the coherence). Coherence is measured as an index with values comprised between 0 and 1, with the latter indicating the maximum value of coherence. Despite the application of different filters and various image correction procedure, all the sources of noise cannot be totally removed, thus resulting in a coherence value which is often lower than 1.

DInSAR represents an ideal solution for monitoring surface deformations along the sensor's line-of-sight (LOS) direction with millimetric precision over large areas. In particular, the technique has several applications related to the following natural hazards:

- tectonic motions and earthquake-induced deformations in the co-seismic, post-seismic and inter-seismic events (Klinger et al., 2000; Lavecchia et al., 2016; Peltzer & Rosen, 1995; Wright et al., 1999);

- volcanic-related displacements (Figure 8) due to volcanic deflation and uplift (Hooper et al., 2004a; Lanari et al., 1998; Lundgren & Rosen, 2003);

- landslides, where only slow and very slow movements gravitational phenomena are detectable due to intrinsic DInSAR limitations (Cascini et al., 2010; Colesanti & Wasowski, 2006; Novellino et al., 2021; Peyret et al., 2008; Reyes-Carmona et al., 2020; Squarzoni et al., 2003);

- subsidence due to natural or anthropic causes (Aly et al., 2009; Cascini et al., 2010; Crosetto et al., 2015; Ferretti et al., 2000);

- mining-induced deformation (Fan et al., 2021; Herrera et al., 2007; Ng et al., 2011; Samsonov et al., 2013; Spreckels et al., 2001).
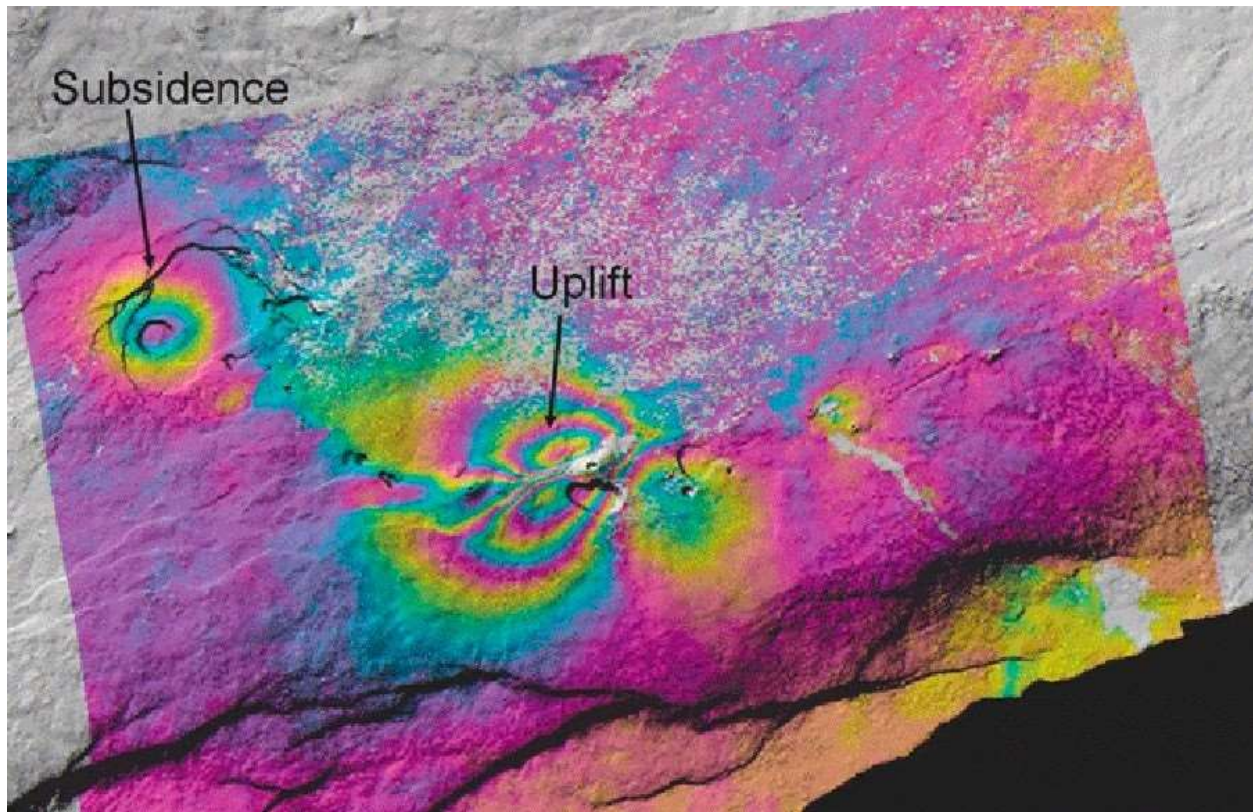
Figure 8. Differential interferogram depicting uplift and subsidence movements associated with a June 2007 earthquake swam on Kilauea Volcano, located on the southeast portion of the island of Hawai (source: Zhong Lu, USGS).

2.4. A-DInSAR processing techniques

Since DInSAR has become a consolidated and well-established measurement tool, many software and web platforms (i.e., SNAP software, developed by ESA) have been ideated to allow new users to apply the DInSAR approach, expanding the scientific community. Moreover, many scientific approaches and techniques have been set up to overcome the main limitations of the single-pair interferometry (atmospheric and decorrelation errors) by exploiting long stack of interferograms. Starting from the early 2000s, a number of algorithms known as Advanced DInSAR (A-DInSAR) techniques were developed to accurately characterize ground deformation by providing spatially dense deformation measurements coupled with a history of movement over time. The term A-DInSAR is used in this work with the same significance of Persistent Scatterers Interferometry (PSI) which represents a specific class of DInSAR techniques exploiting multiple SAR images over the same area to retrieve the displacement phase component, after proper data processing and analysis (Crosetto et al., 2016).

A-DInSAR techniques generally require a minimum number of SAR images (i.e., data stack) (Figure 9) to perform a reliable temporal analysis. Based on the type of sensor, this number may vary: a minimum of ca. 20 images is needed to perform a reliable analysis on C-band data, while a shorter data stack is sufficient for X-band data (Bovenga et al., 2013). This is due to the fact that shorter wavelength allows for higher resolution images and more consistent results. In general, the reliability of the results would greatly improve in relation with the increase of the number of images deployed for the analysis (Colesanti et al., 2003).

A-DInSAR algorithms main outcomes include the deformation time series and the deformation velocity estimated over the analyzed pixels, which are characterized by a small noise component of the phase. These kinds of pixels are typically related to two types of reflectors: those where the response to the radar is dominated by a strong reflecting object and is constant over time within a data stack (Permanent Scatterer, PS) and those where the response is constant over time but is due to different small scattering objects (Distributed Scatterers, DS) (Crosetto et al., 2016; Ferretti et al., 2011). In this thesis, PS, DS and MP (Measurement Point) are used as synonyms. PS are mainly represented by targets exhibiting reduced temporal decorrelation due to their high electromagnetic reflectivity and their coherent and stable scattering behavior (Figure 9). This kind of reflectors are abundant over urban areas, where buildings, man-made metallic structures and roads are present, or also in rocky outcrops. The PS density is very low in vegetated, forested, agricultural or low-reflectivity areas and in steep slopes facing the radar sensor (shadowing). Snow covered surfaces, as well as constantly changing surfaces such as construction sites can lead to the complete or partial loss of PS. Some efforts have been spent towards the maximization of the quality and number of PS selected as reliable a priori, through polarimetric A-DInSAR (Samsonov & Tiampo, 2011). Also, notwithstanding the remarkable increase of measurement points on ground, the position of a PS cannot be known before the processing of the data: this can be regarded as a criticality when coping with small-scales analysis of surface deformation, since some targets of interest may be not sampled at all. The use of X-band-derived high-resolution data can mitigate this deficiency (Crosetto et al., 2010). Moreover, since A-DInSAR results and artifacts removal depends on a stable reference point to estimate relative velocity in respect to that point, a good knowledge of the site characteristics is required.
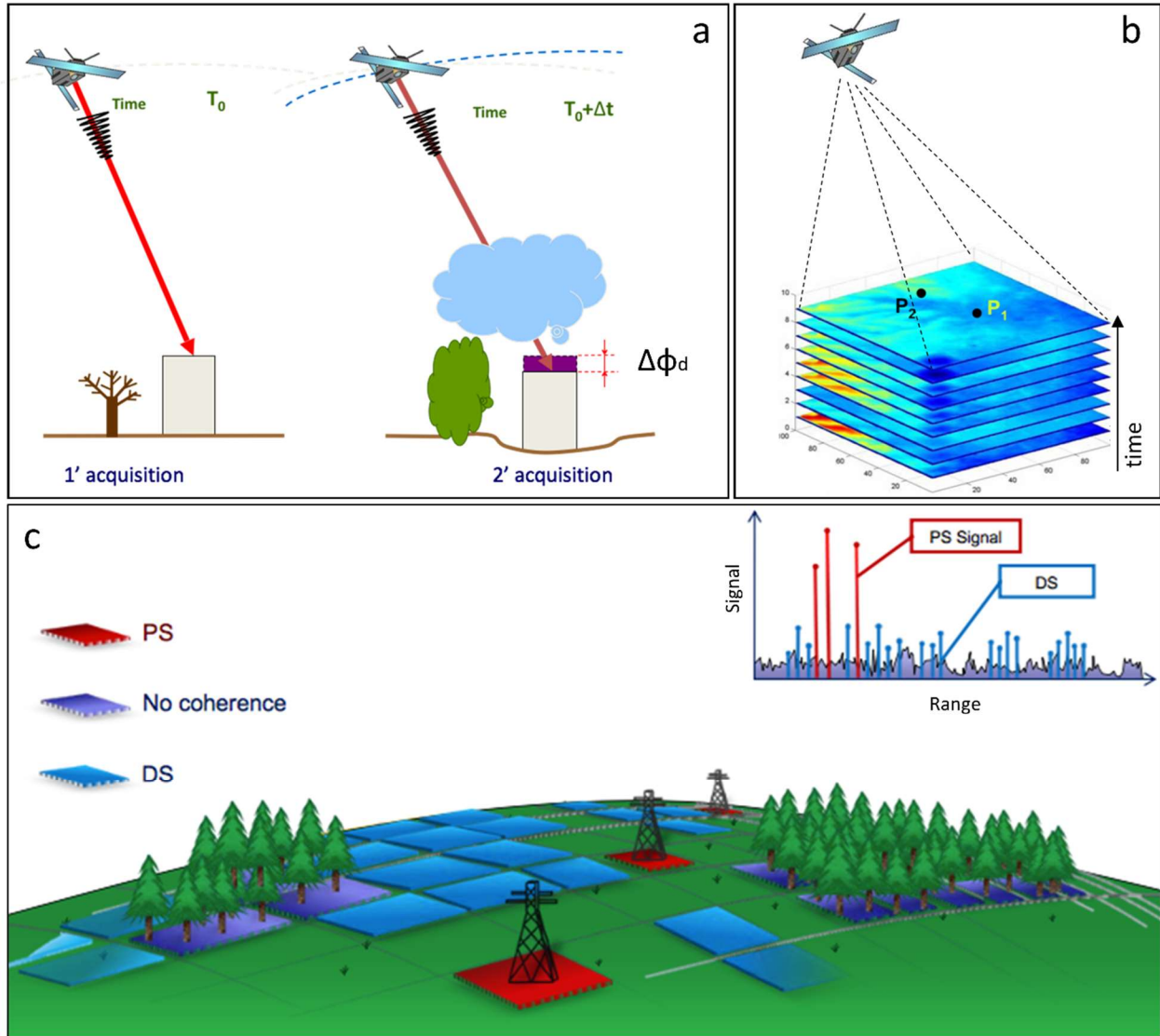
Figure 9. A-DInSAR basic principles: (a) SAR data acquisition over time with phase component variations due to surface displacements, changing atmospheric and ground conditions; (b) collection of a stack of SAR images over the area of interest constituting the starting point of any A-DInSAR processing; (c) different types of reflectors exhibiting different scattered responses.

The deformation time series history is a major information resource and one of the most important products resulting from A-DInSAR processing. The cumulated displacement over the observed period is available for each PS and it is fundamental for many applications (e.g., studying the kinematics of a given phenomenon, quiescence, activation, acceleration, hyperbolic trends etc.) and for correlate deformation trend changes with external and causative factors (Cigna et al., 2011). Every time series is composed of several deformation estimate corresponding to every SAR acquisition. For this reason, they are particularly sensitive to the phase noise. This is evident especially for C-band data, while X-band data show a remarkable quality improvement of the time series (Crosetto et al., 2010). The deformation model deployed by most of the existing A-DInSAR techniques is linear, which is an important assumption that needs to be evaluated, especially when analyzing site-specific displacements. For the same reason, the A-DInSAR techniques deploying a linear model would result in a loss of PS in those areas where the observed

non-linear deformation it is significant, because there is a misfit between the linear model and the real ongoing displacement behavior. Moreover, the deformation time series can be affected also by anomalous trend not related to real ground motions, such as long-term geological processes, errors tied to removal of the atmosphere contribution, uncompensated phase ramps, unstable reference point, thermal effects on the target (Notti et al., 2015).

In addition to the above-mentioned limitations, another severe drawback tied to A-DInSAR capabilities lies within the detectable maximum deformation rate. In particular, "fast" deformation phenomena cannot be derived due to the ambiguous nature of its observations (i.e., the wrapped interferometric phases): if the observed deformation phases overcome π, the real displacement cannot be estimated. The limit is set to a maximum differential deformation between two acquisitions of $\lambda/4$ over the revisit interval. Depending on the orbital characteristics of the different SAR missions and on the deployed sensors capabilities, the maximum differential deformation rate measurable may vary. For instance, for TerraSAR-X mission the maximum velocity rate is 25.7 cm/yr, for S1A it is estimated to be 42.6 cm/yr, while, regarding ALOS, the limit reaches 46.8 cm/yr (Crosetto et al., 2016). It must be stressed that the capability to sample deformation phenomena over time depends on the availability of SAR data (i.e., data acquisition policy, the orbital tube, the revisiting time, the allowed perpendicular baselines), which heavily influences the outcomes of the A-DInSAR processing chains.

The last two decades have seen the flourishing of several A-DInSAR techniques. The so-called Permanent Scatterers approach (Ferretti et al., 1999, 2000) led to the creation of the PSInSAR algorithm by TRE-Telerilevamento Europa, and it constitutes the first A-DInSAR approach to be developed. Another common approach is Small Baselines Subset (SBAS) (Berardino et al., 2002) which exploit distributed targets. Other approaches are: IPTA, developed by Gamma Remote Sensing (Werner et al., 2003), the CPT (Blanco-Sànchez et al., 2008), and the STaMPS method (Hooper et al., 2004b), which is an open-source software for A-DInSAR processing developed at the Stanford University. In Table 2 the main A-DInSAR approaches are reported.

Table 2. Main A-DInSAR approaches.

A-DInSAR approach	Main references
PSInSAR (Permanent Scatterers Interferometry)	(Ferretti et al., 1999, 2000)
SBAS (Small Baseline Subset)	(Berardino et al., 2002)
Modified SBAS	(Schmidt & Bürgmann, 2003)
CPT (Coherent Pixels Technique)	(Mora et al., 2003)
IPTA (Interferometric Point Target Analysis)	(Werner et al., 2003)
SPN (Stable Point Network)	(Duro et al., 2004)
StaMPS (Stanford Method for Persistent Scatterers)	(Hooper et al., 2004b)
STUN (Spatio-Temporal Unwrapping Network)	(Kampes, 2006)
PSP (Persistent Scatterers Pairs)	(Costantini et al., 2009)
DePSI (Delft Persistent Scatterer Interferometry)	(Ketelaar, 2009)
SqueeSAR	(Ferretti et al., 2011)
WAP (Wide Area Product)	(Adam et al., 2011)
SARproZ	(Perissin & Wang, 2012)
MInTS (Multiscale InSAR Time Series)	(Hetland et al., 2012)
Distributed Scatterers approach	(Goel & Adam, 2014)
Modified SqueeSAR	(Lv et al., 2014)
CPS (Cousin PS)	(Devanthéry et al., 2014)
P-SBAS	(Casu et al., 2014)

In this Thesis work, datasets obtained from SqueeSAR and P-SBAS approaches are deployed. The characteristics of the two A-DInSAR techniques are briefly exposed:

- SqueeSAR represents an evolution of the PSInSAR technique. Phase and amplitude are the main parameters exploited by the PSInSAR method, where amplitude allows to individuate the coherent PS with a stable reflectivity while phase allows to estimate the ground movement of the PS. The temporal evolution of the movement is retrieved based on a linear model of displacement, and the temporal unwrapping strategy necessitates to know a priori the characteristics of the motions to be analysed. The result of the PSInSAR is an accurate measurement of the movements along the LOS of each PS, respect to an assumed reference point, in the time interval considered. The SqueeSAR algorithm is able to include and see those measurements points referred to partially coherent and homogeneous areas with low reflectivity (i.e., DS), enhancing the measurement accuracy of surface displacements. Deformation velocity maps and time series of the displacements are the main final products.

- The Parallel Small BAseline Subset (P-SBAS) advanced DInSAR processing bases its principle on the SBAS technique (Berardino et al., 2002), which retrieves time series of deformation exploiting interferograms characterized by small temporal and spatial baseline. The deployment of correlated interferograms would results in the high coherence and the high spatial density of the final product. Independent acquisition datasets, separated by large baselines, are properly linked by applying the Singular Value Decomposition (SVD) method, to retrieve deformation time series and velocity maps. To get beyond the high computational requests of the SBAS technique, the P-SBAS method was put in place. P-SBAS has been adapted to efficiently run on high performance distributed-computing infrastructure (e.g., grid and cloud computing) (Zinno et al., 2015), and tailored for S1 IW TOPS data processing.

3. Big SAR Data management for multi-hazard assessment and monitoring

In recent years, massive amounts of spaceborne geospatial data have been generated thanks to the flourishing of the EO systems. From the technological development of satellites platforms to the increasing interest in investing in the global space segments, from the strategic role played in assessing the socio-economic impact of climatic crises to the management of emergencies, EO data played a crucial role in understanding processes on our planet, enabling the scientific community and CPA to make great strides in monitoring environmental change, risk detection and land cover analysis (Gomes et al., 2020). In this context, managing large EO data represents a recent major challenge considering limited storage and processing capacities of personal computers. The present limitations can prevent EO data users to be fully deployed in their entirety, causing a limited usage of the available data and thus resulting in a limited impact for scientific research and operational applications (Stromann et al., 2020). Thus, there is a need for changing traditional spatial data infrastructures (SDI) to effectively store, ingest, prepare, process, analyze and disseminate these big data sets (Figure 10).
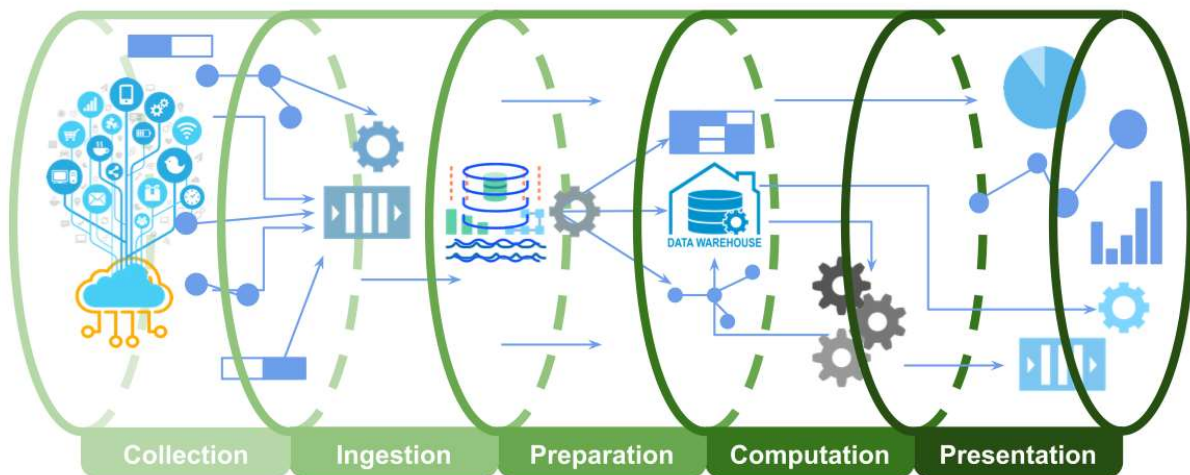
Figure 10. Building a robust data pipeline for processing and managing satellite Imagery (source: medium.com)

A-DInSAR techniques constitute a multi-scale tool offering the ability to detect ground displacements over wide areas with relatively high spatial density of measures and with millimetric sensitivity precision. A-DInSAR technique has seen its popularity dramatically increased with both the availability of X-band data (since 2007) and with the launch of Copernicus S1A and S1B satellites in April 2014 and April 2016, respectively. S1A&B represent a big revolution in the Earth Observation scenario thanks to the unprecedent revisiting time, the main acquisition mode on land (IWS which implements the TOPS), the worldwide coverage, the systematic acquisition plan, the rapid product delivery, and a free data policy. S1 constellation acquired over Europe around 300 passes considering ascending and descending geometries, where an initial repeat pass of 12 days of S1A only (~180 scenes in 6 years) was improved to 6 days in late 2016 thanks to the launch of S1B (~120 scenes in 4 years). However, it worth reporting that

S1B mission ended prematurely due to an anomaly dated back to December 2021. Apart from S1, the number of satellite data sources is continuously increasing, guaranteeing to design redundant acquisition schemes in which the possible failure of one of the systems chosen cannot compromises the monitoring strategy.

The volume of SAR data collected, archived, and processed by A-DInSAR techniques by far falls within the "big data" remit for such type of imagery, which will be referred to as Big SAR Data in this thesis. Therefore, multi-temporal InSAR datasets (and the related studies and elaborations) face some major challenges:

- the need to find efficient combinations of computing infrastructure, algorithms, and software capable to run and optimize the big SAR data processing workflow (Cigna & Tapete, 2021).

- the need for value-adding tools able to manage, organize and analyse the outputs of A-DInSAR processing routines, which often consists of hundreds of thousands or millions of PS (and DS) holding several sources information such as geolocation, coherence, average displacement velocity, and a time-dependent series of displacement values based on the number of SAR scenes deployed for the multi-temporal processing approach.

EO data and multi-temporal InSAR datasets require interpretation to some degree. This is due to the fact that this kind of data are non-repeatable (i.e., unique in space and time), uncertain (i.e., there are different ways of observing, capturing and recording data), multi-dimensional and complex to analyse and to understand (Sudmanns et al., 2020). Due to the above-mentioned reasons the recent developments in big data might be halted. Therefore, the traditional manual post-processing and data integration methodologies must be necessarily re-adapted to pass-through data-driven, machine-based solutions, which is purely based on statistics.

Scientific EO satellites capture data regularly from the Earth's surface allowing for higher-level geospatial information to characterize geological hazards of potential concern of local CPA. This can be extremely useful in characterizing and addressing hazardous areas over hundreds or thousands of square kilometres, allowing CPA to evaluate different scenarios and to plan specific actions based on homogeneous and reliable measurements. Different practical systems and portals are available for processing, storage, and analysis of big interferometric data (Cigna & Tapete, 2021). In this framework, several applications have been already carried out by the scientific community, governmental agencies and by private companies to map and monitor different sources of deformations connected to different scenario of environmental hazard.

3.1. State-of-the-art systems and platforms for EO data

A range of different EO portals dedicated to consult, download, programmatically process and integrate satellite data are nowadays available. Here we focus on the main technical solutions that include SAR imagery and interferometric finite products. Starting from the public initiatives provided by space agencies to the efforts pronounced by private corporates, the most used EO systems are the followings (Figure 11):

- Copernicus Open Access Hub (https://scihub.copernicus.eu) is a web-based system designed to manage the online dissemination of the EO data collected by the Sentinels satellites. It provides *synchronous* access to the latest data, and *asynchronous* access to the historic data.

- Google Earth Engine (GEE) represents the most common planetary-scale platform for Earth science data & analysis. Remote sensing data are available and can be processed without requiring download. Users can also upload their own private data. A Web-based integrated programming environment (IDE), as well as an application programming interface (API), can be used to perform pixel-based image processing (Sudmanns et al., 2020).

- USGS (United States Geological Survey) Earth Explorer provides online search, browse display, metadata export, and data download of satellite, aircraft, and other remote sensing inventories (https://earthexplorer.usgs.gov/).

- The DIAS (Data and Information access services), launched in June 2018, are cloud-based platforms developed to facilitate and standardize the access to Copernicus data and information. The platforms provide access to Sentinel data and to Copernicus operational services, while hosting processing applications and tools (https://www.copernicus.eu/en/access-data/dias).

- The Sentinel Hub (https://www.sentinel-hub.com/) is a cloud API (application programming interface) for satellite imagery, providing service-oriented satellite imagery infrastructure. It represents an interface for handling satellite imagery archive and makes it available for end-users via easy-to-integrate web services.

- PANGEA information system (https://www.pangaea.de/) which functions as an Open Access library for a large variability of data products (including data sets based on satellite imagery).

- Earth Observation Data Centre for Water Resources Monitoring (EODC) is a EO data repository and cloud computing platform. EODC offers access to a variety of data, most notably Copernicus data, with S1 data included (https://eodc.eu/data/).

- Registry of Open Data on AWS (Amazon Web Services) makes available datasets via AWS resources from a variety of satellites data.

- The Copernicus Data and Exploitation platform – Germany (CODE-DE) is a german national access tool, geared towards national and local authorities, dedicated to geoinformation, which provides access to remote sensing data, a virtual working environment for data processing and information material (https://code-de.org/en/).

- The Plateforme d'Exploitation des Produits Sentinel (PEPS) is a platform, constituting the national french counterpart to the CODE-DE platform, providing free access to data sets collected by the Sentinel satellites. Users are allowed to search, select and download EO data (https://peps.cnes.fr/).
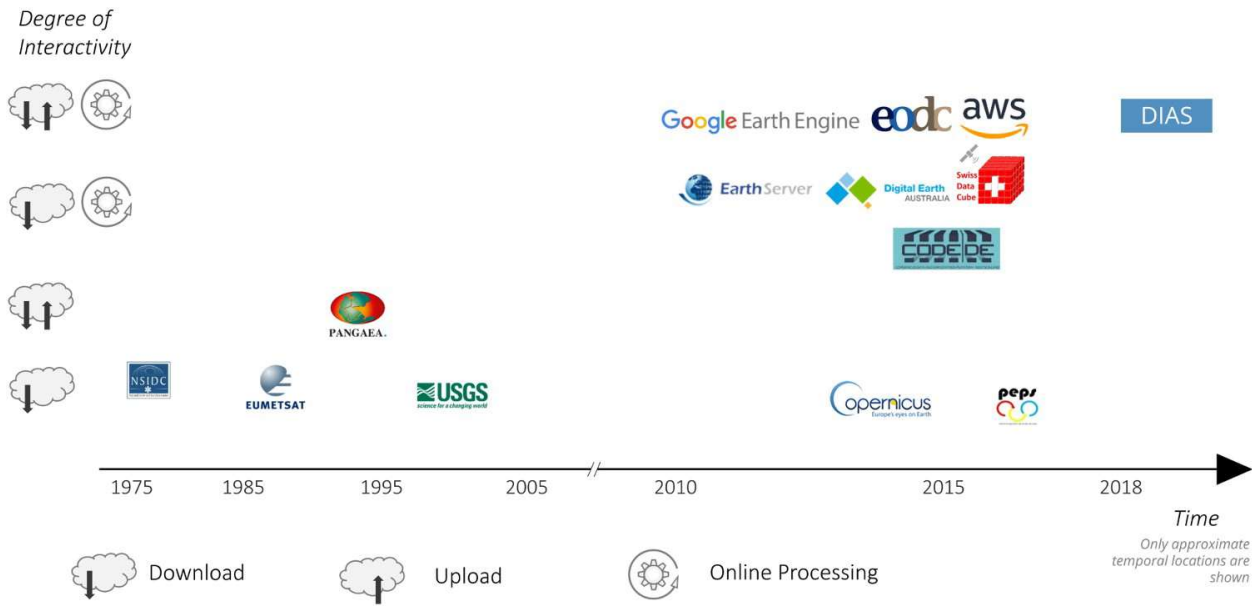
Figure 11. The most popular EO systems and platforms (source: Sudmanns et al., 2020).

3.2. InSAR-based monitoring infrastructures

The feasibility and development of remote sensing platforms and services have led to a big revolution in the EO scenario, in terms of operational capability and amount of radar involved. In this context, some relevant efforts in the direction of extensive exploitation of large SAR archives has already been carried out at a various scale of analysis. In particular, radar images are nowadays used within at least two different lines (Raspini et al., 2018):

- Ground deformation mapping, realized in deferred time, with a yearly frequency plan of update for objectives like territorial planning or geohazard inventories update.

- Ground deformation monitoring, which is a service provided in near-real-time and it is based on the anomaly detection of InSAR time series and can be thought as an operational tool for the support in emergency situations by CPA.

So far, satellite radar data have been mostly used for assessing post-event damages or retrieve data about the occurred event (i.e., a seismic event, a landslide phenomenon or landscape morphological changes due to extreme floods) to assess damages within the impacted area and possibly look retrospectively for possible precursory signals (Raspini et al., 2018). However, given the maturity of the interferometric technique and the recent developments of the underlying space-infrastructure, new initiatives have taken place at exploiting radar data at regional, national, and even continental scale, raising SAR technique to a new level in which it can be used not only for single sites characterization, but for repeated evaluations within short time frames.

Italy is at the forefront for what concerns the realization and the exploitation of continuous monitoring systems. Back in 2007, the Italian Government began supporting the largest ever InSAR analysis to be

26

founded by a national government. In the first instance, the project involved processing, by means of the PSInSAR and the PSP techniques, the entire ERS 1/2 and Envisat archives over Italy (more than 15,000 SAR images covering more than 330,000 km^2), which called for the creation of a dedicated geospatial database to manage millions of radar points (www.pcn.minambiente.it/mattm/progetto-piano-straordinario-di-telerilevamento/). This interferometric database, spanning from November 1992 to September 2010, was then extended in a second phase (di Martire et al., 2017) by adding Envisat images acquired between 2008-2010. During the third phase, the project was foreseen with ingestion of 100 frames (40 x 40 km) from COSMO SkyMed images acquired by ASI (Agenzia Spaziale Italiana). The project was designed to map unstable areas over the entire national territory.

Other important operative regional, national, and continental scale motion services that have been successfully implemented are briefly illustrated as follows:

- Ground Motion Service Germany (BBD, Kalia et al., 2017) constitutes the first S1 based nationwide ground motion dataset. To ensure the combination of remote ground motion data with other independent data sets, the PSI data are calibrated with time series of continuous GNSS stations. In addition, plausibility checks are carried out through site inspections, optical remote sensing data and other geodata.

- InSAR Norway (Bredal et al., 2019) is the Norwegian public national web-based ground motion mapping service, based on Copernicus S1 data. It was launched in November 2018 by the Geological Survey of Norway (NGU), The Norwegian Water Resources and Energy Directorate (NVE) and the Norwegian Space Centre.

- GMS for Denmark (copernicus-sentinel-1-helps-denmark-to-monitor-ground-motion-and-infrastructure), where the nationwide dataset is produced by TRE Altamira and provided by the Agency for Data Supply and Efficiency (SDFE), after processing the entire Copernicus S1 archive over Denmark, from both ascending and descending orbits.

- The Dutch Ground Motion Service (bodemdalingskaart.nl/en-us/) has been finalized by Rijkswaterstaat (RWS—Ministry of infrastructure and water management). It is currently operative and it will be updated twice per year and will be a free and open service. 1467 images from radar satellites were used to build up the dataset composed of ca. 165 million of points, with a time frame comprised between January 2015 and June 2020.

- S1 Monitoring Services of Tuscany, Valle d'Aosta and Veneto Italian regions launched in 2016, 2018 and 2019 respectively (Confuorto et al., 2021). They are defined as near-real time monitoring systems based on S1-derived A-DInSAR products, where time series of each measurement point are systematically analyzed to detect any change in the deformation pattern.

- EGMS initiative (Crosetto et al., 2020a), implemented under the responsibility of the European Environment Agency, is the largest wide-area A-DInSAR service ever conceived. This initiative was defined in the period 2016–2017 and since May 2022 the products are available online for visualization and soon, they will be downloadable (Figure 12). The EGMS include three main

products: deformation maps and deformation time series with measurements along the radar LOS; A-DInSAR deformation map combined with a reference GNSS network; the horizontal east–west and up–down vertical deformation maps (resolution 100 x 100m) obtained by combining the ascending and descending A-DInSAR results.
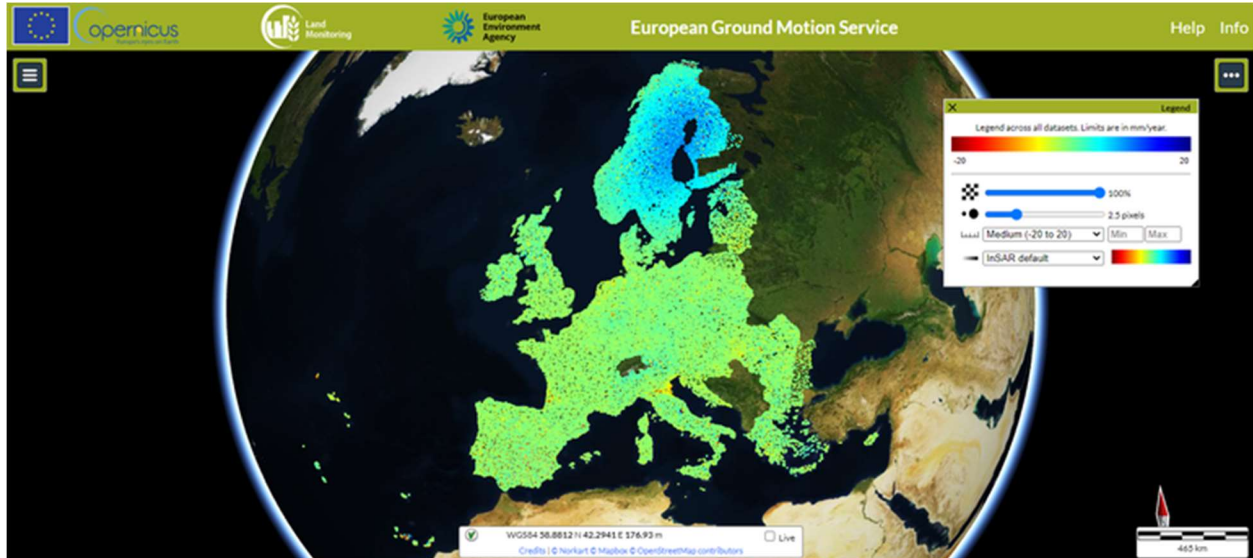

Figure 12. The EGMS online viewer (source: land.copernicus.eu).

4. Value-adding approaches for enhanced interpretation of interferometric products

In this Thesis work, different and complementary approaches for enhancing the interpretation and management of radar-based information are proposed. Large multi-temporal interferometric datasets resulting from different processing algorithms are implemented into post-processing and value-adding routines in the framework of capturing and optimally share values from earth observation analytics.

With the EO business model shifting from selling raw data to foster services for higher-value information to extract, EO data are set to be the main actor in the next years to provide stakeholders with decision-making tools and operational processes. Given the growing volumes of available data and the increase in variety of EO data and the contemporary shift towards cloud-based storage and processing infrastructure, new value-adding solutions have started to develop. EO specialists will be focusing on the analytics for value creation, while the market-agnostic "APIfied" players will further reduce the barriers to EO adoption in high-volume markets, which will fuel demand for data along the entire value chain (capturing-value-earth-observation-analytics).

Along with the SAR specialists in charge of the data processing, A-DInSAR data requires analysis experts able to deliver and interpret the data for different application purposes. Therefore, the final step of any radar-based analysis is conducted on the basis of ancillary information (e.g., geological, hydrogeological, topographic, land cover,…) and in conjunction with any other in-situ or remote sensing sensors available on the area of interest. Based on the scale of the study, different aims and objectives are feasible: at the detailed scale of analysis (>1: 5,000), radar data can provide the deformation history of single buildings affected by intense ground motions, also making it possible to apply techniques for the evaluation of the damage probability (Del Soldato et al., 2019). At medium scale (1:100,000 – 1:25,000), a possible applicability of interferometric products would be to the define the ranking of the municipalities affected by active deformation. At large scale (1:25,000 – 1:5,000) the analysis is carried out for zoning the built-up areas within the most affected territories where damage occurred or is likely to occur. Moreover, A-DInSAR data can be employed to improve the prediction reliability within landslide susceptibility maps (Ciampalini et al., 2016). A multiple scale approach can be combined and used to release the full potential of the interferometric technique. A cascade multiscale approach (Peduto et al., 2015) allows recognizing the most hazardous zones in an area of interest, while the working scale can be progressively increased and ancillary data with different resolution can be adapted consequently.

Different workflows have been proposed to deploy interferometric datasets within post-processing strategies for the detection of areas of active motions, characterize them in detail and cross-correlate the satellite data with relatable geospatial and temporal information. The purpose of implementing value-added products is to evaluate the impact of ground motions (intended as high deformation rates or abrupt changes in the trend of deformation) and therefore to provide the CPA in charge of the hydrogeological risk reduction with solid decisional tools.

Cigna et al. (2011) and Meisina et al. (2008) were among the first to realize that the need for automatic or semi-automatic strategies is becoming urgent and critical topics of discussion for the scientific

community. They showed that through the expert interpretation and analysis of A-DInSAR time series is possible to characterize the ground instabilities of a territory spatially and temporally. In particular, the integration of radar satellite data together with in situ available ancillary data such as geological, geomorphologic, climatic would be an important practice for hazard mapping and for the identification of the appropriate mitigation measures.

Berti et al. (2013) propose an approach for the automated interpretation of A-DInSAR time series of displacement which is based on conditional sequence of common and standard statistical tests. The authors point out that the manual classification and visualization of large quantities of multi-temporal radar data cannot be performed systematically. Therefore, the authors developed a standalone application with a graphical user interface (called PS-Time), which requires A-DInSAR time series as its input and performs their classification into predefined trends of deformation, such as uncorrelated, linear, quadratic, bilinear and discontinuous. Mirmazloumi et al. (2022) proposes an updated procedure to the workflow of Berti et al. (2013), where the objective is to extend the classification rationale of to investigate the misclassified time series and a critical noise source such as vertical jumps caused by PUE (Phase Unwrapping Errors).

Bakon et al. (2017) propose to detect outliers in the postprocessing phase of multi temporal interferometric results, while maintaining a good density of MPs in the final results. By applying clustering techniques, PCA (Principal Component Analysis), Voronoi diagrams, median absolute deviation and Jaccard index the resulting products are cleared of statistically insignificant observations.

Tomás et al. (2019) developed a semi-automatic tool for the identification and pre-screening of geological and geotechnical deformational processes from A-DInSAR processed results. The proposed procedure aims at the detection of Active Deformation Areas (ADA) from deformation map, allow the decomposition of Line of Sight (LOS) displacements from ascending and descending MPs into vertical and east-west components and categorizes each ADA into potential geological hazards. The methods objective is to discriminate landslides, subsidence, sinkholes, and consolidation processes from radar data.

Hussain et al. (2021) presented a statistical approach for detecting offsets and gradient changes in A-DInSAR time series by statistically identifying significant peaks in the first, second and third difference series. The gradient change detector can be used for both online and offline outliers' identification by finding statistically significant movements in the second derivative series. The methodology demonstrates to work well in case of availability of long time series data given that its computation is parallelised and vectorised for major efficiency in order to deal with the large data volumes.

Machine learning (ML) technologies have been successfully applied to interferometric products for feature detection and classification. In this sense, the screening of volcanic deformation, coal mining and tunneling subsidence is accomplished by Anantrasirichai et al. (2021) by spatially interpolate the deformation signals, apply overwrapping techniques and by deploying synthetic training datasets. Brengman & Barnhart (2021) use CNN (Convolutional Neural Networks) to detect, locate, and classify temporally and spatially impulsive surface deformation events in interferograms with the main objective being the identification of co-seismic or co-seismic-like surface deformations. Ansari et al. (2021) takes

advantage of a deep Long Short-Term Memory (LSTM) autoencoder model to find temporal relations in deformation time series and cluster similar features.

In the following chapters, three new different approaches for spatial and temporal analysis of large interferometric datasets are outlined. The scale study of this research is from regional to nationwide, while maintaining a major focus on the retrieval of precise and accurate ground deformation information at a very localized scale. The common and underlying objective of the various presented approaches is the development of new A-DInSAR post-processing tools addressed to the management of big and large A-DInSAR datasets covering whole regions or nation. The spatiotemporal characteristics constituting the multi-temporal interferometric dataset are here retrieved from automated routines of supervised and unsupervised approaches for the automated recognition, detection, and characterization of ground deformation areas. In this way, surface displacements occurring on the monitored territories are intended to be periodically seized, characterized, and uploaded in updated inventories concerning multi-level hazards which can be further exploited by the CPA for risk mitigation and remediation practices.

In the following chapters, novelties, advantages, and constraints of the proposed Data Analytics (DA) tools are outlined. The post-processing operations are conducted and evaluated considering A-DInSAR datasets ranging from regional/small to large/detailed scale ("scaling up" approach). Feature engineering, data integration, DM and ML techniques are adopted for a comprehensive selection and transformation of the most relevant variables and information retrievable from already processed interferometric datasets deployed as test cases (Figure 13):

CHAPTER 5 – MOVING AREA CLUSTERS (MAC) DETECTION AT NATIONAL SCALE. Here is presented a GIS geoprocessing tool (Python scripting) chaining together sequences of processes for the spatial clustering of radar targets moving with relatively high values of displacement trend. The detected deforming areas (MAC) are then evaluated against the Line of Sight (LOS) velocity vector decomposition and the accessible ancillary layers for a preliminary classification of the triggering factors. The proposed methodology is here applied to the mean ascending and descending deformation maps covering the whole Italian territory resulting from 3294 and 2868 S1 acquisitions respectively, spanning from March 2015 to December 2018 and processed through the Parallel Small BAseline Subset (P-SBAS) technique. By setting a displacement rate threshold of ± 1 cm/year, a total number of 14,638 areas resulting from both geometries are found to suffer from instability phenomena, the origin of which are in turn preliminary sorted in 11 classes split between natural causes and man-made activities. With 2 degrees of confidence, landslide and subsidence events are found to be the main causes of deformation within the Italian territory, constituting respectively 31% and 27% of the total unstable areas, followed by volcanic-related processes (22%). Lastly, it is displayed a complete overview of the deformation phenomena which have recently occurred on the Italian Peninsula starting from national scale statistical analysis and ending up with local scale investigations according to the deformation patterns visible through the vertical and East-West components of motion.

The proposed DM procedure can be applied to A-DInSAR deformation maps for the static mapping of hotspots of ground deformation due to geological or anthropogenic phenomena.

CHAPTER 6 – MACHINE LEARNING-BASED ASSESSMENT OF GROUND DISPLACEMENT SOURCE. Here is presented an example of automated solution aimed at performing integrated analysis of large amounts of interferometric data. To effectively detect high-displacement areas and classify ground motion sources,

it has been explored the feasibility of a machine learning-based approach. This is achieved by applying the Random Forest (RF) technique to large-scale deformation maps spanning 2015–2018. Focusing on the northern part of Italy, the model is trained to identify landslide, subsidence, and mining-related ground motion with which to construct a balanced training dataset. The presence of noisy signals and other sources of deformation is also tackled within the model construction. The proposed approach relies on the use of explanatory variables extracted from the A-DInSAR datasets and from freely accessible informative layers such as Digital Elevation Model (DEM), land cover maps, and geohazard inventories. In general, the model performance is very promising as we achieved an overall accuracy of 0.97, a true positive rate of 0.94 and an F1-Score of 0.93. The obtained outcomes demonstrate that such transferable and automated approach may constitute an asset for stakeholders in the framework of geohazards risk management.

CHAPTER 7 – <u>AUTOMATED CLUSTERING OF TIME SERIES OF SURFACE DEFORMATION</u>. A statistical-based approach based on the time-dependent displacement behaviour of radar MPs is implemented with the purpose of deriving the most widespread deforming trends within the observed A-DInSAR dataset. The approach copes with the applicability of an automated unsupervised learning approach by applying a statistical dimensionality-reduction method which includes a Principal Component Analysis (PCA) followed by K-means clustering to detect non-stationary deforming trends at a regional scale. The added benefit of combining ascending/descending InSAR data and interpolating displacements in time at different time steps is here explored prior to data dimensionality reduction and feature extraction through PCA. The signal of the ground deformation clusters is then deconstructed into the underlying trend and seasonality components to enhance the interpretability of the classified satellite InSAR features. Using InSAR Time series data spanning 2014-2020 and covering the Valle d'Aosta region (North-western Italy), the proposed approach detects several slope movements and anthropogenic deformations with both linear and seasonal displacement behaviours.
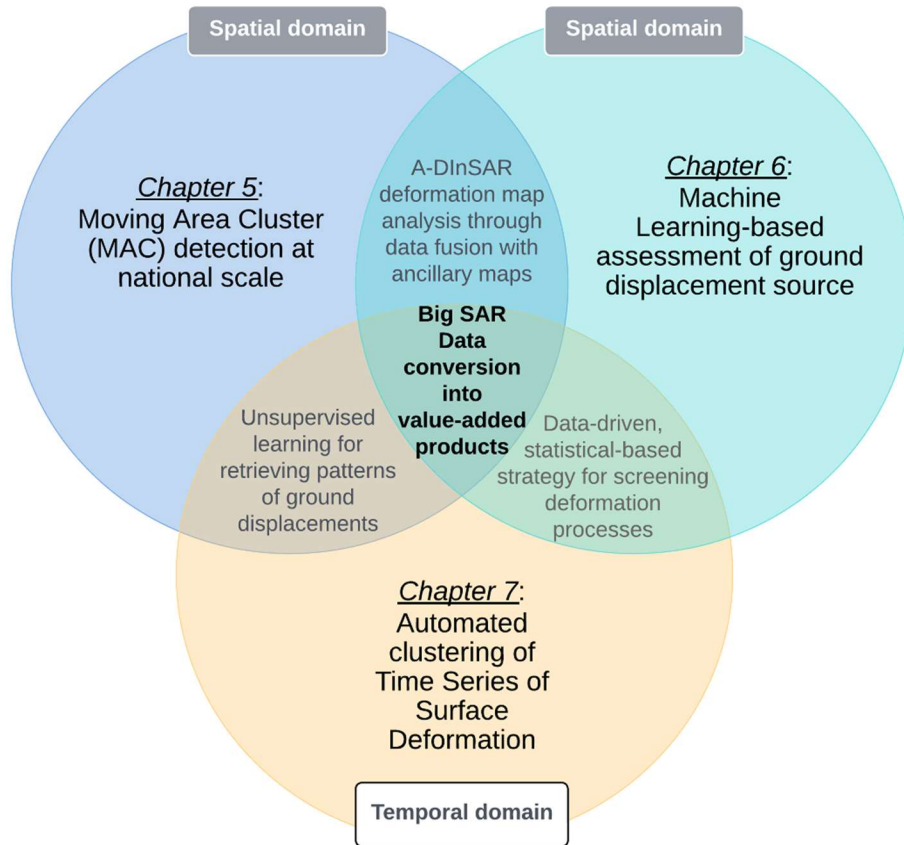
Figure 13. The post-processing tools proposed in this thesis cover Data Mining (DM) concepts and methodologies, while dealing with data integration for the implementation of Machine Learning (ML) techniques and statistical approaches in order to extract useful information from the analysed interferometric datasets.

5. Moving Area Cluster (MAC) detection at national scale

This section of the Thesis work will illustrate an approach for the detection of Moving Area Cluster (MAC) from A-DInSAR deformation maps for the preliminary screening of the high-velocity displacement phenomena concerning a territory of interest. In the presented case, a nation-wide interferometric dataset is deployed, thus allowing for testing a post-processing approach with an unprecedent scale for this kind of analysis. The enhanced temporal repetitiveness and coherence of the S1 constellation is here deployed for a reliable mapping of ground deformation punctuating the Italian Peninsula. Thus, the proposed methodology is based on the semi-automated analysis and interpretation of S1 interferometric data of ground displacements above the whole Italian territory referred to a time span comprised between 2015 and 2018.

In this chapter is presented the result of the research activity which was partially funded by the "Presidenza del Consiglio dei Ministri–Dipartimento della Protezione Civile" (Presidency of the Council of Ministers–Department of Civil Protection), through the IREA-CNR/DPC and CPC-UNIFI agreements. This research was also partially funded by the IREA CNR/Italian Ministry of Economic Development DGS-UNMIG (now Italian Ministry of Ecologic Transition DGISSEG) 2018-2019 agreement.

The rising availability of satellite-based multi-temporal interferometric datasets covering large areas of the Earth surface constitutes a huge asset in the context of operational workflows aimed at improving land risk assessment and management. In order to cost-effectively handle huge amount of data, here it is proposed a semi-automatic procedure to quickly identify, map and inventory ground and infrastructures displacements by means of spatial clustering performed over very large-scale A-DInSAR datasets. The detected deforming areas are then routinely evaluated against the Line of Sight (LOS) velocity vector decomposition and the accessible ancillary layers. This process is devoted to the rapid and effective identification of hotspots of displacements and for a preliminary classification of the triggering factors.

The presented geoprocessing tool is called *MAClustering* and it can be fed with A-DInSAR datasets counting millions of MPs while allowing the user to set up and tune the parameter values, thus allowing to retrieve outputs that match the user specific needs. Therefore, this DM approach is to be used for a large-scale clustering operation on radar MPS allowing for an overview of the main displacement phenomena concerning a large territory and for a detailed scale analysis of local deformations. The retrieved unstable areal deformations are here referred to as *Moving Area Cluster* (MAC). The multiscale approach is to use the interferometric data for mapping ground motions at regional scale, providing reliable information on the spatial location, on the average intensity, on the areal extension and on the possible origin of the detected phenomenon. The results obtained can be implemented in an operative procedure that involves Regional Authorities and CPA in order to create a chain of command based on the A-DInSAR data where priority actions are needed.

5.1. Input data

The presented methodology relies on a complete multi-scale approach starting from a country-wide mosaicked mean deformation velocity map and ending up with a classified active motion areas (MAC) database. Two different kinds of data have been exploited for the above-mentioned task:

- S1 interferometric dataset;

- A set of pre-existing inventories and ancillary data, constituting an ensemble of generally freely accessible informative layers (in order to make the proposed approach transferable).

5.1.1. Nation-wide P-SBAS interferometric data

For the proof of concept of the proposed post-processing tool, an interferometric deformation map covering the whole Italian territory represents the starting point of the MAC generation.

Such a A-DInSAR deformation map covers an area of more than 300,000 km^2 and it has been computed within a project lead by the Italian Ministry of Economic Development aimed at investigating the deformation phenomena affecting the Italian territory (Manzo et al., 2019). The realization of the mean deformation velocity map relies on interferometric analysis at national scale obtained through the P-SBAS algorithm (Casu et al., 2014), which simultaneously exploits a fast access to the S1 data archives, High-Performance Computing resources and external geodetic data (de Luca et al., 2017; Manunta et al., 2019; Zinno et al., 2015). In particular, the P-SBAS processing chain takes full advantage from the peculiarities of the S1 IWS acquisitions mode, where every acquisition strip is composed of several swaths and bursts. This data partitioning promotes the distribution of the S1 interferometric processing of the independent bursts on a large number of different computing nodes or CPUs. Moreover, the S1 constellation, being characterized by a very small orbital tube (with a nominal diameter of about 200 m), makes the derived raw data particularly suitable to be processed through the P-SBAS approach resulting in a minimized spatial decorrelation. Having in mind fast processing of huge volumes of S1 data, the Cloud Computing environments are perfectly appropriate for fast processing of huge volumes of S1 data, leveraging on storage capability, large variety of instances for the computation and resources scalability with respect to the amount of data to be processed.

In this work, Sentinel-1A and 1B radar images acquired in C-band (wavelength 5.5 cm and incidence angle ranging from 33° to 43°) are exploited in ascending and descending orbits covering the whole Italian territory (Figure 14). The overall investigated area has been divided into 19 and 17 frames, from 5 and 6 ascending and descending orbit tracks, respectively, with an average size of almost 200 km × 250 km (azimuth and range). An overlap between frames has been preserved, covering almost 20% of the frame size for two subsequent ones in the azimuth direction, thus guaranteeing an additional consistency check for the continuity preservation of the achieved results.

Frame	Number of S-1 slices	Number of S-1 acquisistions
T88_N1	530	175
T88_N3	354	177
T88_N4	347	170
T15_N2	460	164
T15_N3	474	175
T15_N1	445	164
T117_N2	485	159
T117_N1	506	166
T117_N3	519	180
T117_N4	529	179
T117_N5	439	181
T44_N2	493	163
T44_N1	529	177
T44_N4	531	178
T44_N3	530	178
T44_N5	406	160
T146_N3	538	183
T146_N1	539	183
T146_N2	540	182
TOTAL	9194	3294

Frame	Number of S-1 slices	Number of S-1 acquisistions
T66_N1	494	166
T66_N2	459	180
T168_N1	464	175
T168_N2	438	151
T168_N3	351	175
T168_N5	321	146
T95_N1	472	171
T95_N2	474	172
T95_N3	466	166
T22_N1	398	164
T22_N2	510	171
T22_N3	511	163
T124_N1	502	178
T124_N2	464	178
T124_N3	412	172
T51_N1	494	170
T51_N2	488	170
TOTAL	7718	2868

Figure 14. Representation of the processed S1 frames within the S1 IWS P-SBAS processing chain. (a)-(b) represent the processed frames over ascending and descending orbits, respectively. On the right side of the picture details about the processed slices and the corresponding acquisitions for each frame are shown. Note that each colour in (a) and (b) identifies a different track.

The retrieved A-DInSAR products (Figure 15) have been obtained by using data acquired from March 2015 to December 2018 with a resulting pixel dimension of about 80 m x 80 m (Table 3). The A-DInSAR processing has been carried out following the S1 P-SBAS approach (Manunta et al., 2019) which includes

appropriate interferogram phase filtering (Pepe et al., 2015) and accurate phase unwrapping procedures (Pepe & Lanari, 2006). Atmospheric phase screen and possible residual phase artifacts are filtered out by also exploiting global navigation satellite system (GNSS) position time series, which in turn are properly selected for the sufficient extension in time and for not being affected by local displacement effects, as detailed in (Lanari et al., 2020). The final average A-DInSAR measurement accuracy is of about 5 mm on the single retrieved displacement time series (Manunta et al., 2019). To produce high-precision A-DInSAR measurements, only pixels holding a temporal coherence (Pepe & Lanari, 2006) value greater than 0.9 were retained for this work (Figure 15). Taking into account both ascending and descending A-DInSAR nation-wide datasets, a total of 15,247,842 MPs have been made available for this research (Table 3).

Table 3. Features of the input ascending and descending A-DInSAR datasets, including the MPs (Measurement Points) count, the smallest and largest values of velocities encountered, the standard deviation of the velocity distribution and the number of MPs per km^2.

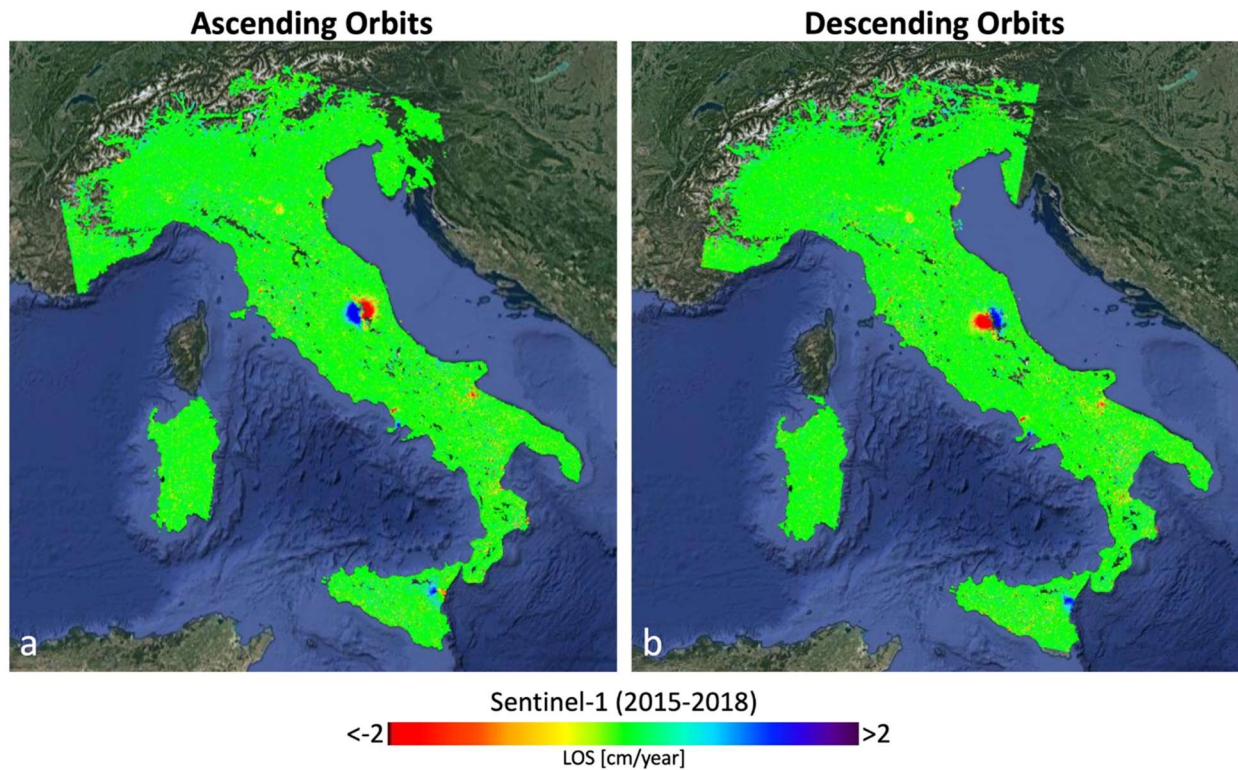| Satellite (Orbit)	Time Interval (Number Of Images)	MP Count	Vel. Min – Max (cm/yr)	Vel. Std. Deviation
Sentinel-1 (Ascending)	23 Mar 2015 – 26 Dec 2018 (*3,294*)	8,066,243	-10.1 – 6.8	0.3
Sentinel-1 (Descending)	26 Mar 2015 – 29 Dec 2018 (*2,868*)	7,181,599	-11.0 – 10.9	0.3

Figure 15. Ascending and descending nation-wide P-SBAS deformation velocity maps obtained from the exploitation of the whole S1 IWS archive spanning the March 2015 – December 2018 time interval.

5.1.2. Pre-existing inventories and ancillary data layers

The joint use of geohazard-related inventories and ancillary maps is crucial for the classification and characterization of the A-DInSAR-based MAC database. In this work, both pre-existing freely accessible informative layers and re-elaborated products were employed. They can be summarized in three categories, pre-existing geohazard inventories, ancillary maps, DEM, here resumed:

- The selected pre-existing inventories are related to different geohazards, namely landslides, earthquake-induced landslides, and subsidence. In the first case, the Italian Landslide Inventory (IFFI project) at 1:10 000 scale (Trigila et al., 2007) has been chosen since it depicts the spatial distribution of past and present mass movements within Italy (represented as polygons or points). It must be pointed out that the data are updated to 2017 for the Region of Umbria; to 2016 for Emilia Romagna, Friuli Venezia Giulia, Liguria, Piedmont, Sicily, Valle d'Aosta and the Autonomous Province of Bolzano; to 2015 for Tuscany; to 2014 for Basilicata and Lombardy. For the remaining regions, the data are updated to 2007. Moreover, the Italian Catalogue of Earthquake-induced ground failures in Italy (Fortunato et al., 2012) supplies a detailed picture of the ground effects due to the seismic sequence occurred on 2016-2017 in the Central Apennines (Improta et al., 2019a) and it is therefore used as a proxy to extrapolate spatially coincident earthquake-related moving areas. Finally, a subsidence inventory map has been derived based on the analysis

conducted by (Solari et al., 2018) who spatially defined subsidence areas in Italy monitored by means of interferometric techniques in the last 20 years;

- Ancillary maps consist of thematic layers such as land use, physiographic and lithological maps. In particular, the following layers have been deployed: CORINE Land Cover map at 1:100 000 scale from the Copernicus Land Monitoring Service and updated for the 2018 reference year, where up to 44 classes of land covers are distinguished; *Carta della Natura* map at 1:250 000 scale ([sistema-carta-della-natura](#)) provided by ISPRA (*Istituto superiore per la protezione e la ricerca ambientale*) where the Italian peninsula is sub-divided into 37 different physiographic units associated to geographically defined sections of the territory; 1:500,000 lithological map of Italy with 16 associated different classes (https://www.isprambiente.gov.it/images/progetti/progetto-1250-ita.jpg);

- The slope map (raster format) of Italy has been computed from the TINITALY/01 DEM, a digital elevation model with 10 m-cell sizes covering the whole Italy territory (Tarquini & Nannipieri, 2016).
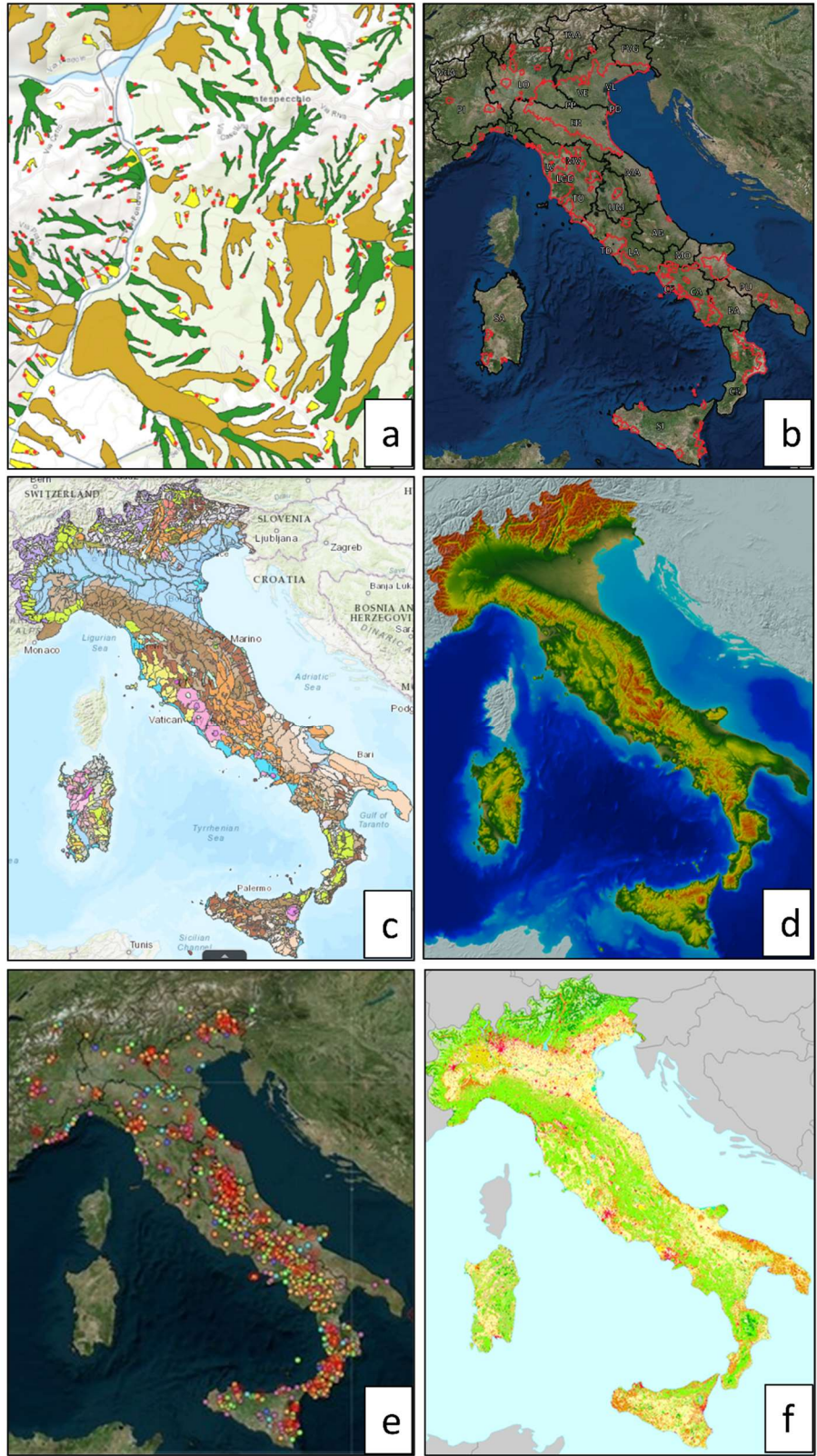
Figure 16. Visual representation of the deployed inventories and ancillary data layers: (a) the Italian Landslide Inventory (IFFI project) at 1:10 000 scale (Trigila et al., 2007); (b) subsidence inventory map derived from (Solari et al., 2018); (c) Carta della Natura map at 1:250 000 scale from ISPRA; (d) TINITALY/01 DEM *(Tarquini & Nannipieri, 2016)*; (e) Italian Catalogue of

Earthquake-induced ground failures in Italy (Fortunato et al., 2012); (f) CORINE Land Cover map at 1:100 000 scale (updated till 2018) from the Copernicus Land Monitoring Service.

5.2. Methodology

The methodological approach consists of three post-processing phases performed over the P-SBAS outcomes resulting from S1 IWS archives: (i) semi-automatic spatial clustering of active ground deformation areas affecting the Italian territory, thus identifying the MACs; (ii) subsequent automatic classification by means of superimposition on pre-existing inventories, thematic maps and by LOS velocity vector decomposition for every MAC; (iii) local scale analysis and characterization of the detected ground deformation processes through the investigation of the deformation pattern.
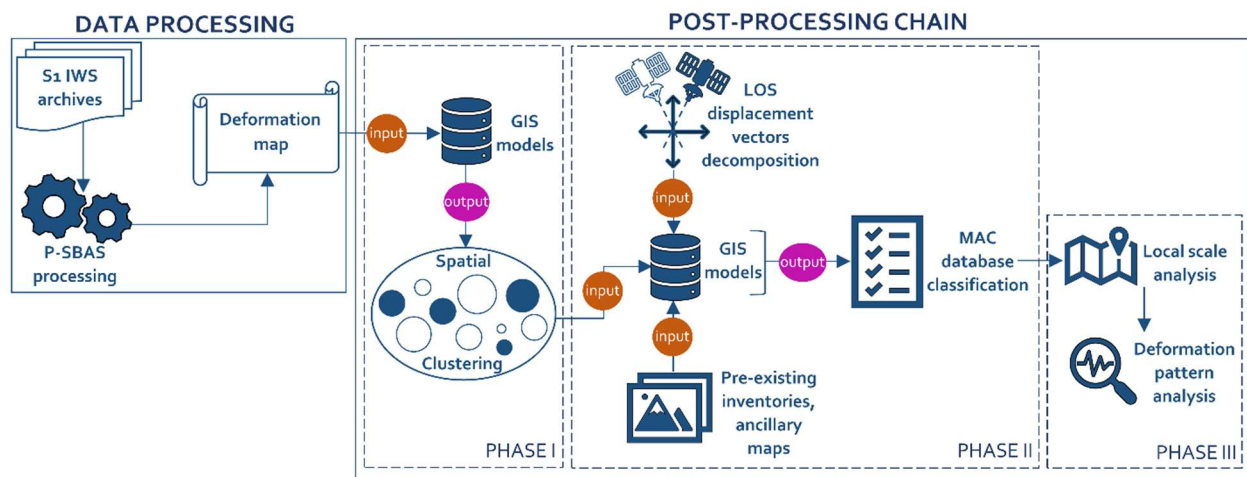

Figure 17. Flow chart of the adopted methodology.

5.2.1. Spatial clustering of radar targets

Groups of spatial objects with similar features can be merged into clusters. Clustering constitutes an unsupervised learning method used as a process to find meaningful structure and explanatory underlying processes in the intrinsic grouping among the unlabelled data present. Spatial clustering analysis is particularly useful when it is applied to filter nation-wide A-DInSAR databases and retain MPs representative of relatively large average displacement behaviour, in order to detect the most potentially active hazardous motions.

MAClustering is a semi-automated GIS model consisting of a sequence of geoprocessing tools that allows the filtering, the extraction and the spatial aggregations of MPs into clusters of moving areas with a significant displacement rate. *MAClustering* can be shared or can be used in Python scripting and other models. The model workflow consists of a chain of tasks connected where the output of every process is fed into another process (Figure 18a). The presented GIS geoprocessing model is provided with an interface dialog box from which the user can supply different values parameters to the model. In particular, 10 different built-in GIS tools were exploited while, besides the need for specifying the input

data file and the working folder, 3 parameters are requested to be set (Figure 18b). The characteristics of the final output strongly depend on the model parameters settings which can be described as follows:

- Displacement rate threshold, whose function is to filter out stable MPs (i.e., ground pixels holding absolute velocity values lower than the chosen threshold);

- Buffer size, which is the MP radius of "influence" that should be set in accordance with the spatial resolution of the A-DInSAR dataset (e.g., if the distance between two MPs is greater than double the radius, they are not aggregated in the same cluster);

- Minimum cluster size, which is referred to the statistically relevant minimum number of MPs to perform spatial clustering.
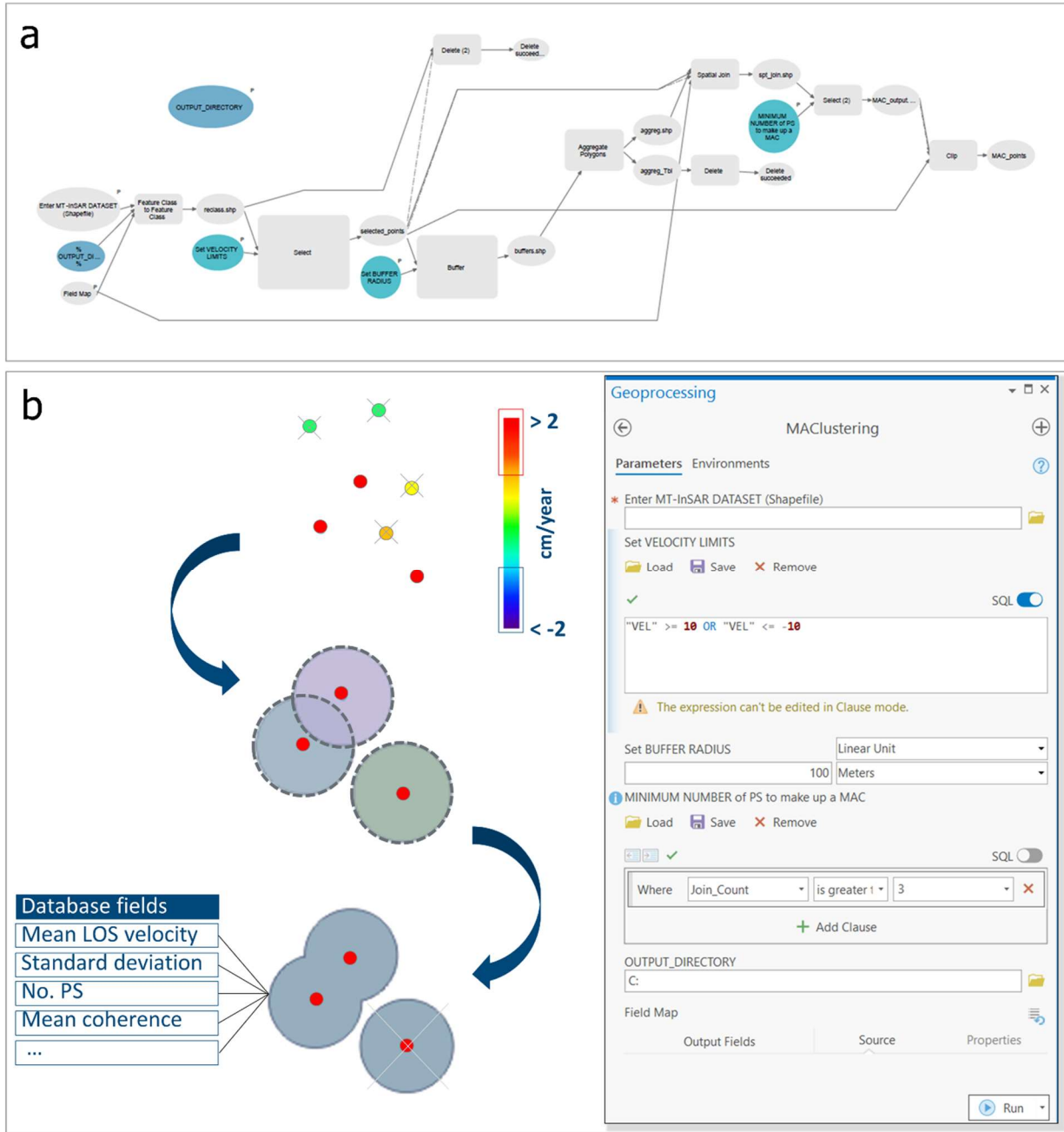
Figure 18. Overview of *MAClustering*: (a) diagram showing the dense network of interactions between the built-in tools deployed within the geoprocessing model; (b) an interface dialog box coming with the model allows to interactively set up 3 parameters for the extraction of MACs.

The proposed model and underlying approach can be assimilated to the Density-based spatial clustering of applications with noise (DBSCAN) algorithm (Ester et al., 1996). DBSCAN identifies a dense region by grouping together data points that are closed to each other based on distance measurement. Compared to this approach, *MAClustering* provide a tailored and more personalized approach for the grouping of radar MPs data by leveraging on the user specific needs.

Taking into account the scale of analysis, it has been implemented an iterative procedure aimed at fine-tuning the displacement rate threshold based on the standard deviation (σ) of velocity, following already existing literature approaches (Barra et al., 2017; Tomás et al., 2019). A final stability threshold of ± 1 cm/year, corresponding to slightly more than 3σ, was here set to distinguish active MPs from stable ground reflectors. Keeping in mind the spatial resolution of the P-SBAS results (about 80 m × 80 m), every ground pixel is here inscribed within a 100 m radius buffer, a minimum number of 3 contiguous MPs was chosen to execute spatial clustering, implying that MACs with an extent lower than 3 times the dataset spatial resolution are not detected (about 0.02 km^2). Thirdly, the average displacement value and coherence of the MPs included in each cluster were automatically derived (as part of the model geoprocessing chain) in order to retrieve relevant statistical features for every deformation area.

The abovementioned parameters are not subjected to fixed rules; any sensitivity analysis regarding the impact of the specific parameter selection is neglected, considering that they should be tailored to the features of the given interferometric dataset and to the scope of the analysis.

It can be deduced that the expected output, generated in a shapefile format, consists of polygonal clusters which in turn might represent an approximation of the detected deformational processes in terms of planar extension. However, this kind of analysis constitutes nonetheless an accurate and quick method to systematically identify and locate sectors of the Italian territory where ongoing deformations due to geological processes are taking place.

5.2.2. LOS displacement vectors decomposition

P-SBAS provides deformation values measured along the radar line of sight, thus representing only a one-dimension estimation of the real motion. Reprojecting the velocity deformation vectors along the vertical (Up-Down direction) and horizontal (East-West) direction is a useful practice to enhance the deformation estimates and the classification of the observed phenomena (de Luca et al., 2017; Manzo et al., 2006; Notti et al., 2014), which can be done when ascending and descending datasets are available and when the contributions of N-S horizontal motions are assumed negligible (due to the intrinsic limitation of the satellite acquisition geometry). The computation of the vertical and horizontal components is performed by solving two simple equations (8), as pointed out by (Notti et al., 2014):

$$V_V = \frac{((V_D/\mathrm{elos}_D) - (V_A/\mathrm{elos}_A))}{(\mathrm{hlos_D/elos_D} - \mathrm{hlos_A/elos_A})}; V_H = \frac{((V_D/\mathrm{hlos}_D) - (V_A/\mathrm{hlos}_A))}{(\mathrm{elos_D/hlos_D} - \mathrm{elos_A/hlos_A})}$$

(8)

where V_V and V_H are the vertical and the horizontal velocity components, V_A and V_D are the ascending and descending (LOS) velocity vectors; while hlos=cos(α) and elos=cos(1.571 – α)*cos(4.712 – θ), where θ and α are the azimuth and the incidence angle respectively.
In order to retrieve V_V and V_H, we consider each pixel on the ground for which we have both ascending and descending LOS velocity values (Figure 19a). The vertical and horizontal components of the velocity can be thought as coordinates depicting the magnitude and the direction of the movement for each MP. Therefore, by representing the horizontal velocity on the x-axis and the vertical velocity on the y-axis it is possible to retrieve the "real" velocity vector, which can be plotted on the Up-Down and East-West plane.

In order to help through an accurate visualization of the direction of the "actual" observed motion, it is here introduced a colormap (Figure 19b) which is based on the ratio between the vertical and horizontal components of deformation. In this way, from (i) Up-East to (ii) East-Down, from (iii) Down-West to (iv) West-Up, 12 patterns of movement are highlighted.

For this research, the ratio between the vertical and horizontal motion components is retrieved and associated to every MAC for classification purposes. On the other hand, detailed analyses at local scale are carried out by investigating the deformation pattern in depth and by fully exploiting the availability of spatially coincident LOS measures from both geometries.
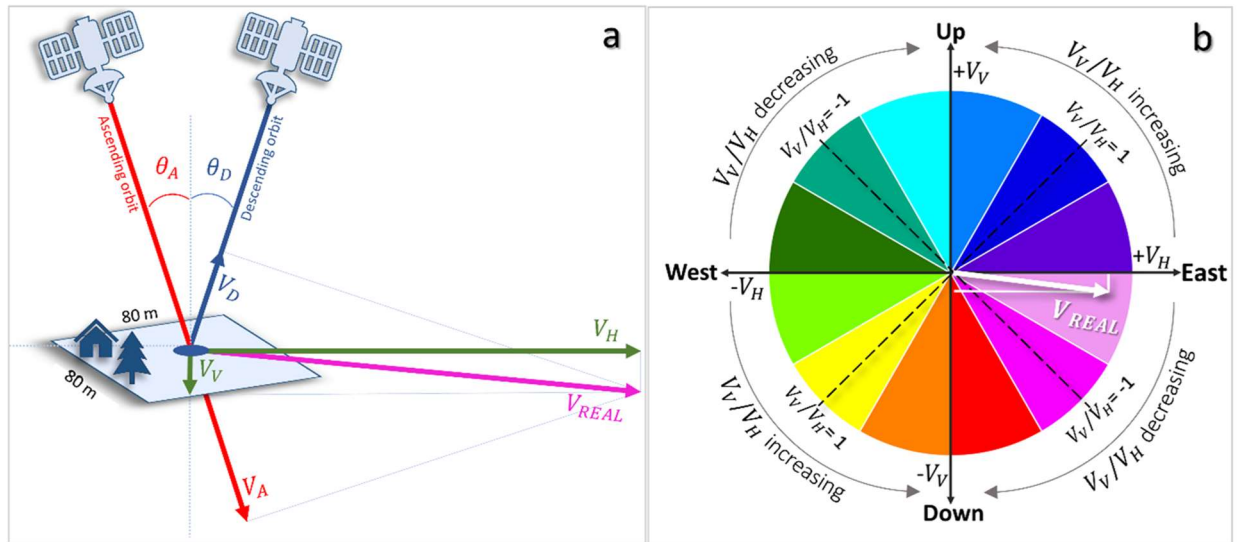

Figure 19. (a) Resolving the E-W (horizontal) and the Up-Down (vertical) components of the LOS velocity within a MP; (b) polar plot of the "real" velocity vector projected on the Up-Down and East-West plane, where the different colours are associated to different directions, thus enabling a better comprehension of the pattern of movement.

5.2.3. MAC classification procedure

The desirable outcome of the proposed post-processing chain is the identification of the cause which leads to the ground deformations outlined by the spatial clustering results. The deforming area can be thought as an effect of natural processes or man-made activities (Matano, 2019). The model developed for the classification task consists of a chain of 8 different built-in GIS tools which focus on the quantification of the spatial relationship between features and on conditional statements dealing with threshold values. In order to make the model run, it is necessary to supply a mandatory polygon shapefile in input (i.e., the MAC dataset), the ancillary maps when available and the slope map of the area of interest. Moreover, the model requires to specify the working folder and to set 4 parameters, namely a series of threshold values which can be summarized as follow: (i) percentage of spatial overlap; (ii) the mean vertical and horizontal components ratio values; (iii) the mean vertical component value; (iv) the mean slope value. The decision tree and the specific parameter setting adopted in our case study is depicted in Figure 20. Based on the working scale and the A-DInSAR data resolution, a spatial overlap between S1 data and ancillary dataset of at least 50% is here considered valid for a correct classification. According to the spatial correlation existing between the active motion areas dataset and the pre-existing inventories or ancillary maps, each

MAC could be initially discriminated into 8 categories (Figure 20): earthquake-induced deformations, volcanic-related processes, landslides, land subsidence, mining activities, dump site activities, construction site activities (where the latter three can be summarized as man-made activities) and unclassified. The last class is retained for the elements falling outside the classification criterion (i.e., concomitant unavailability of the ancillary information and lack of interferometric data from the two geometries of acquisition). Any further distinctions or interpretation regarding the underlying natural or anthropic process (i.e., settlements, sinkholes, natural or anthropogenic subsidence, etc.) which led to ground deformations are not considered because of the lack of comprehensive ancillary information and because out of the scope of this work.

For the deforming areas that lack of spatial correlation with the available ancillary data, the information retrieved from the 2D decomposition of LOS displacements is exploited. In order to interpret the underlying process which has led the ground to deform, it's here introduced the variable K_{VH} (9) expressing the prevailing direction of the areal movement that has been derived for every MAC, that is the mean ratio between the absolute vertical (Up-Down) and the absolute horizontal (East-West) component values:

$$K_{VH} = \frac{|V_V|}{|V_H|}$$

(9)

where the different ratio values of spatially coincident MPs are averaged and related to every MAC.
The outcome of this computation, in combination with the computed average slope value, strengthens the association of MAC to 3 other possible categories (Figure 20): potential landslide, potential subsidence, potential uplift. In particular, assuming that mass wasting processes are generally developed on slopes higher than 5° (Iwahashi et al., 2003; Lee et al., 2004) and land subsidence likely occurs on plains or valleys floor (slopes less than 5°), moving areas showing a prevailing mean horizontal direction ($|V_V|/|V_H| < 1$) are here considered potential landslides, while moving area showing a prevailing mean negative vertical movement ($V_V < 0$ and $|V_V|/|V_H| > 1$) are considered potential subsidence. Upward vertical displacements can take place in different environments as a result of several factors (Bonì et al., 2018; Coda et al., 2019), thus MAC with a prevailing mean positive vertical movement ($V_V > 0$ and $|V_V|/|V_H| > 1$) are here associated to potential uplifts. The adopted classification method operated through $|V_V|/|V_H|$ ratio thresholds (Figure 20), arises from literature examples and is supported by observations made on the portion of the MAC dataset which have been safely classified according to the spatial correlation with the ancillary data (where available).
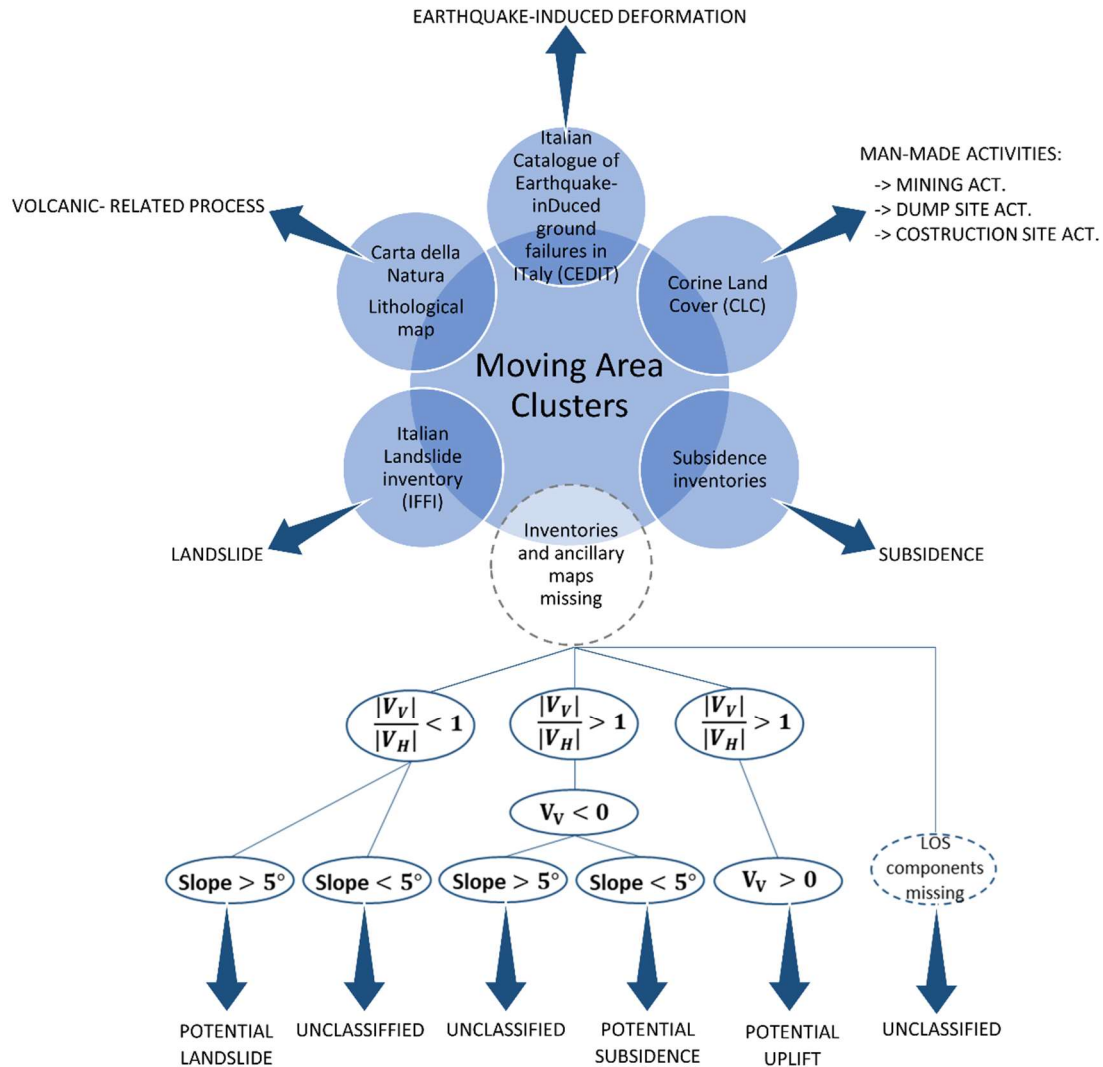
Figure 20. Procedural scheme for the classification of Moving Areas Clusters (MAC).

5.3. Outcomes of the post-processing procedure

The results obtained by means of spatial clustering analysis and through the cross-correlation between the generated MAC and different collected maps covering the whole Italian territory are here described in three different sub-sections: (i) general characteristics of the computed country-wide active motion areas database; (ii) outcome of the performed classification; (iii) presentation of seven case studies, representative of the adopted classification, wherein the validation of the approach is also shown.

5.3.1. The MAC database of Italy

Following the criteria explained in section 5.2.1, a MAC database covering the whole Italian territory has been retrieved for both the ascending and descending input A-DInSAR datasets. The attributes related to the computed active motion areas are summarized in Table *4*. The table fields include the average LOS

velocities for both geometries, the number of MPs falling within every deformation area, the E-W and the vertical component of the LOS velocities, and the classification of the cause from which every MAC originated from.

Table 4. The Moving Areas Clusters (MAC) database is provided with non-quantitative details and with statistical features related to the areal extent of the deforming area, the LOS, vertical and horizontal velocities, and the temporal coherence *(Pepe & Lanari, 2006)*.

Field	Description	Units
ID	Identification number of the active motion area	-
Satellite-orbit	Geometry of acquisition of the MPs from which the active motion area has been generated	-
Areal extent	Computed polygonal surface area	km^2
Mean coherence	Computed mean temporal coherence value of the MPs falling within the active motion area	-
Count MP asc	Number of ascending MPs falling within the active motion area	-
Count MP desc	Number of descending MPs falling within the active motion area	-
Mean VLOS asc	Computed mean ascending LOS velocity	cm/yr
Mean VLOS desc	Computed mean descending LOS velocity	cm/yr
Mean V_H velocity	Computed mean E-W component of velocity	cm/yr
Mean V_V velocity	Computed mean Up-Down component of velocity	cm/yr
Mean V_V/ V_H	Computed mean ratio between Up-Down and E-W component	-
Classification	Classification of the underlying ground deforming process	-

The overall detected MACs are 14,638 which are differently distributed among the belonging A-DInSAR geometry: 8,278 clusters for the ascending geometry and 6,360 clusters for the descending one (Figure 21a and Figure 21b, respectively). The first set of MAC covers around 1,647 km2 while the second covers around 1,223 km2 which correspond respectively to ca. 3% and 2% of the ascending and descending deformation maps (where the areal extent of each MP is 80 m × 80 m). The difference between the two subsets is mainly due by the relatively smaller MP total count and MP density of the input descending A-DInSAR dataset. The mean LOS velocity distribution of the ascending and descending MAC has been computed and it is displayed in Figure 21c and Figure 21d: 85.2% and 87.4% of both datasets, respectively, show mean velocity values ≤ −1.0 cm/year, thus demonstrating a large predominance of deformation areas with a mean negative LOS velocity (movement away from the satellite sensor). Only 14.1% of the ascending and 12.2% of the descending areas indicate mean LOS velocities ≥ +1.0 cm/year (movement towards the satellite sensor). Very few areas with values between −1.0 cm/year and + 1.0 cm/year results from the spatial clustering operation considering that possibly MPs with opposite signs coexist in the same clustered area (due to an abrupt change of movement direction recorded within a very small distance).
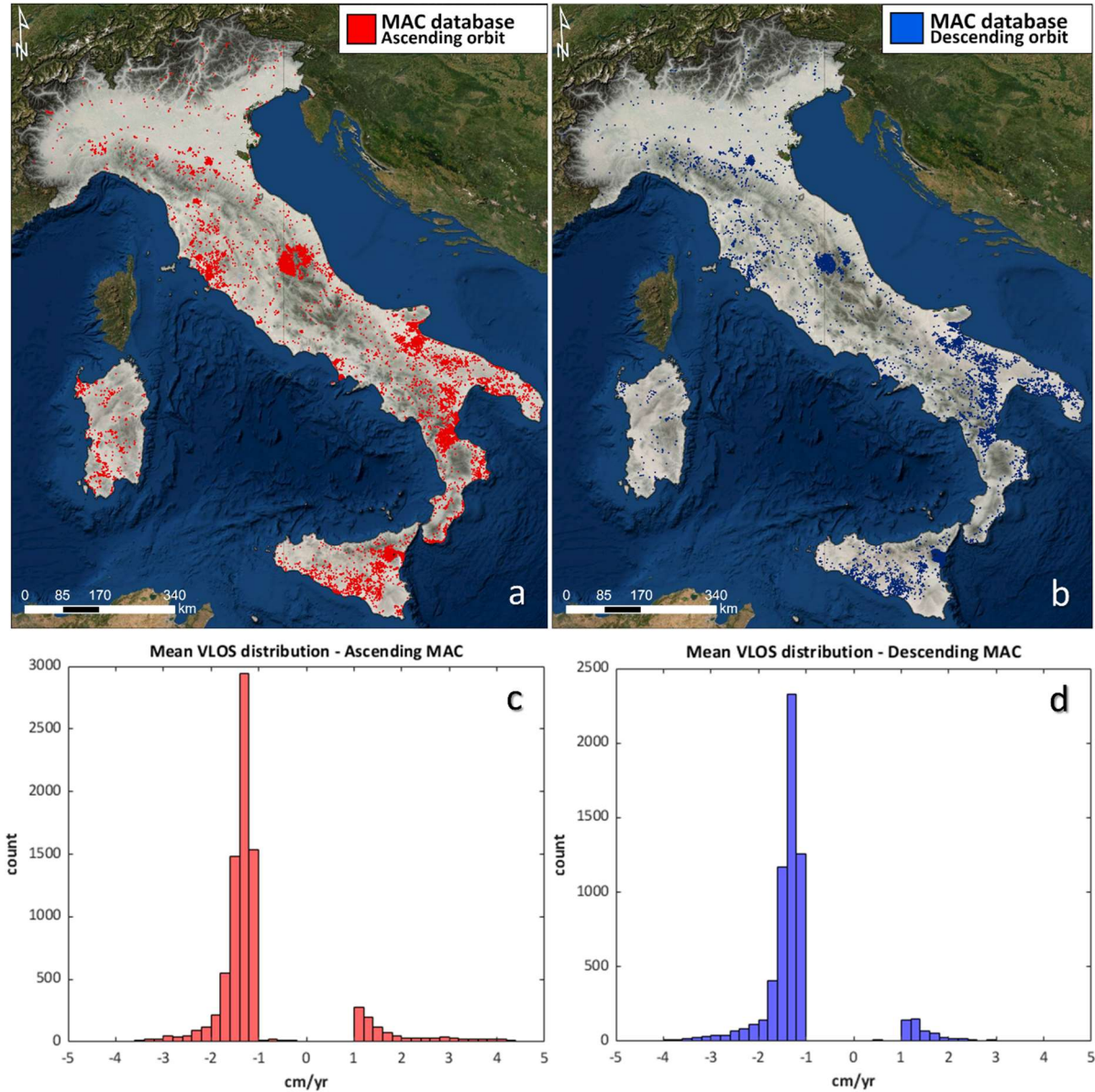
Figure 21. (a) The ascending and (b) the descending Moving Areas Clusters (MAC) database for the reference period March 2015- December 2018 resulting from the spatial clustering analysis performed over the whole Italian territory and (c, d) the mean VLOS value distribution computed for the ascending and the descending MAC, respectively.

The ascending and the descending MAC databases are here compared in terms of areal extent with the lithological map of Italy (1:500 000 scale) and the II level CLC (Corine Land Cover) map. By comparing the MAC coverage (MAC area within class/total class area) with respect to the total MP coverage (MPs area within class/total class area), it is possible to obtain a weighted measure (i.e., the percentage of MAC area within class/MPs area within class) of the extent to which every lithological and land cover class has been affected by surface deformations during the monitored period (Figure 22). Considering MAC belonging to both geometries, it is found that they are mainly located on effusive igneous rocks (7.7% for the ascending and 7.8% for the descending) and pelitic flysch (8.4% and 6.3%), followed by clay and marls (6.8% and

6.4%). Very few deforming areas are found to be spatially coincident with glacial deposits (from 0.1% to 0.2%) and with intrusive igneous rocks.

Concerning the distribution of MAC between land use classes, it is reported a clear disparity between the agricultural areas and the artificial surfaces, with the latter generally not much interested by deformational phenomena. Except for the mine, dump and construction sites (7.8% for the ascending MAC and 7.7% for the descending one), the majority of ground-moving zones is positioned on inland waters, open space with little or no vegetation and maritime wetlands. Compared to the coverage of the MPs, very few MACs are found on artificial, non-agricultural vegetated areas (0.3% for the ascending and 0.2% for the descending dataset) and on urban fabric (0.7% for both datasets). The significant presence of MAC within inland waters is due to the proximity of deformation areas to water bodies, where the clustered MPs mainly refer to dynamic environments such as riverbanks, shores and cliffs.
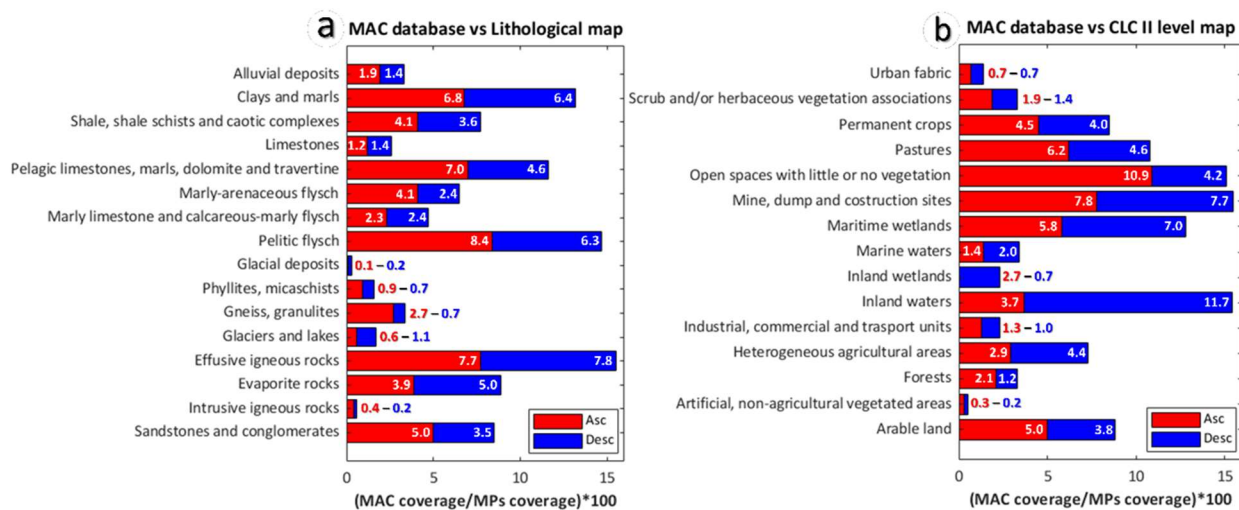
Figure 22. Comparing the coverage of the ascending and the descending Moving Areas Clusters (MACs) with respect to the coverage of the Measurement Points (MPs, constituting the ascending and descending velocity deformation maps) for each categorical class of (a) the 1:500 000 lithological map of Italy and of (b) the second level of the Corine Land Cover map.

5.3.2. MAC classification

The active motion areas have been classified according to the spatial association existing with the geohazard-related inventories and ancillary maps as outlined in Figure 20. In terms of covered area, the overall partitioning of the ascending and descending active motion areas between the various categories is comparable (Figure 23) with some discrepancies induced by the different satellite view. The descending dataset has recorded a relative majority of volcanic-related processes (24.8%) with respect to the ascending dataset (19.5%). An opposite trend arises for the earthquake-induced deformations, which constitute 11.9% of the descending clusters and 17.1% of the ascending ones. Other minor mismatches regard the percentage of subsidence-related events (6.7% of the ascending areas and 7.5% of the descending areas), the fraction of deforming areas positioned on previously mapped landslides (12.0% and 13.8% respectively) and the percentage of not classified phenomena (4.8% of the total area covered

by the ascending dataset vs. 2.7% of the descending one). The remnant categories show differences in the range of ca. one percentage point.
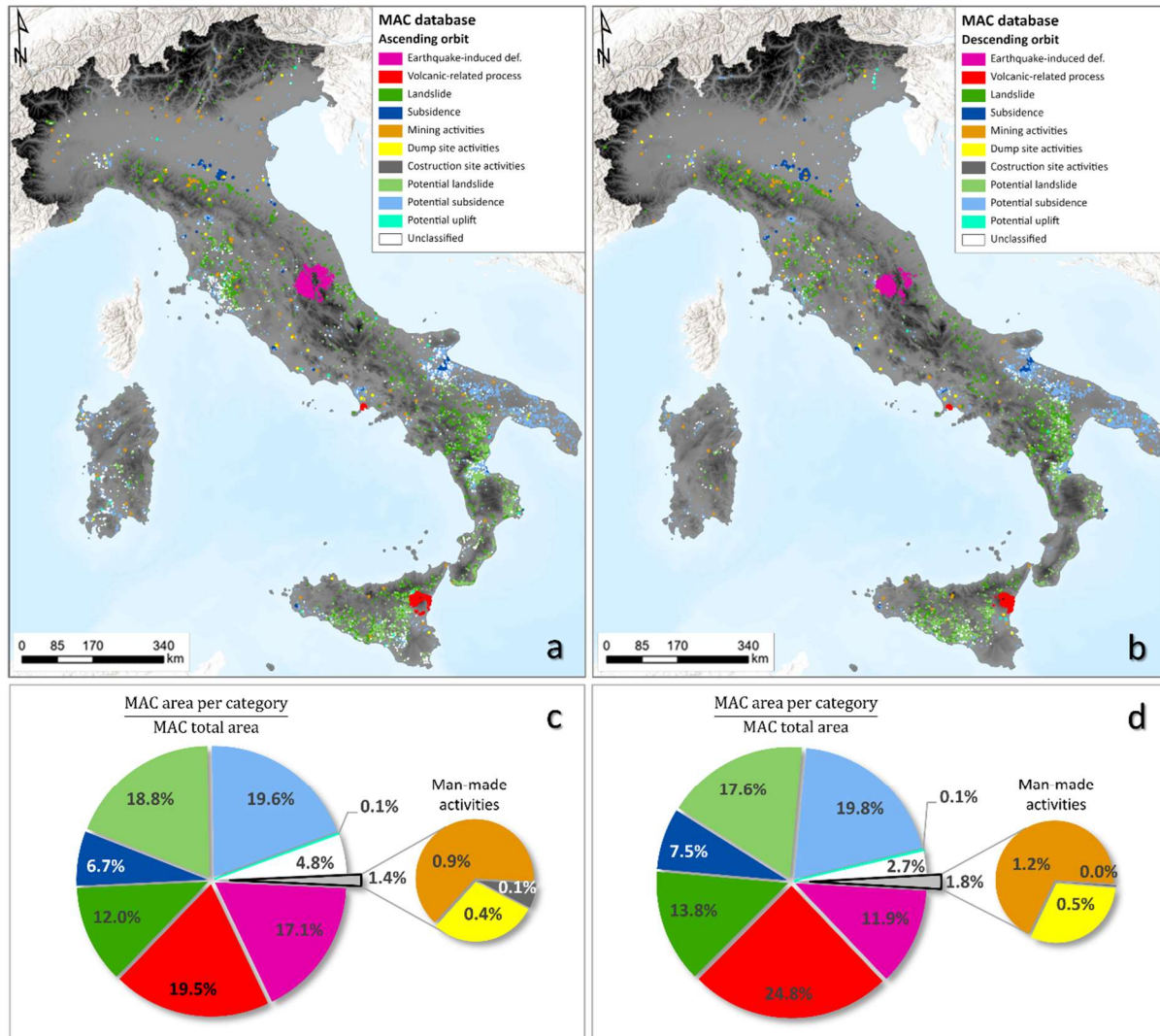
Figure 23. Classification of the (a) ascending and (b) descending moving area clusters (MAC) of Italy retrieved for the reference period March 2015-December 2018. Pie charts showing the (c) ascending and (d) descending MAC distribution within the different categories, in terms of covered area respect to the total area of each MAC database.

If active motion areas categorized as landslide and potential landslide are considered together, it appears that slope instabilities processes could be considered as the most representative phenomenon in the ascending and descending datasets (30.8% and 31.4% respectively). On the contrary, both the potential uplifts and the man-made activities (i.e., dump site activities, construction sites activities and mining activities) constitutes together the minority of the two datasets (Figure 23).

Moreover, an average value of slope and elevation has been computed on the deforming areas belonging to the various categories (Table 5). Nearly identical values (around 19° of slope and elevation comprised between 813 and 833 m.a.s.l.) are shared between the two datasets for the earthquake-induced

deformations, due to the geographical coincidence of the underlying motion areas which correspond to the seismic sequence occurred in 2016–2017 in the Central Apennines. The active volcanic-related processes are likewise reported in the same locations in both ascending and descending MAC (Figure 23): they describe the Mount Etna (an active stratovolcano on the East coast of Sicily) and the Phlegraean Fields (a vast area of volcanic origin located North-West of the city of Naples which has been historically affected by bradyseism) ongoing deformations. In this case the exact geographical position of the MAC is changing, as well as the average slope and the average elevation, according to each orbit.

The distribution of the MAC classified as landslide and subsidence totally reflect the physiographic pattern of the Italian peninsula, with the first category to be found throughout the Apennine chain and the Alpine chain while the second to be placed on the main Italian plains, such as the Po Plain (northern Italy) or the *Tavoliere delle Puglie* (south-eastern Italy). Accordingly, an average slope value of 14.3° – 15.0° (respectively in the ascending and descending dataset) for the deforming areas marked as landslides while an average slope of 0.4° is attributable to the subsidence-related clusters. A similar trend is confirmed when looking at the average slope and elevation values of MAC classified as potential landslide and potential subsidence (Table 5). However, the difference appears to be less pronounced, with a mean slope of 1.6° – 1.5° for potential subsidence phenomena and a mean slope of 12.5° – 13.0° for potential landslides (respectively in the ascending and descending dataset).

Among the remaining categories (i.e., potential uplift and the man-made activities), the mining activities account for the highest percentage of covered area together with the highest average slope and elevation value.

Table 5. Statistics depicting the number and the area covered (both in absolute figures and as a percentage of the total), the average slope and the average elevation computed per every category into which MAC are classified.

Categories	Ascending orbit				Descending orbit			
	Count (%)	Area (%) km²	Avg. Slope deg (°)	Avg. Elev. m.a.s.l	Count (%)	Area (%) km²	Avg. Slope deg (°)	Avg. Elev. m.a.s.l
Earthquake-induced def.	1004 (12.1%)	282.3 (17.1%)	19.3	813.1	629 (9.9%)	145.8 (11.9%)	18.9	833.4
Volcanic-related process	150 (1.8%)	320.0 (19.5%)	8.1	779.7	148 (2.3%)	302.5 (24.8%)	10.6	876.5
Landslide	1349 (16.3%)	197.5 (12.0%)	14.3	417.9	1184 (18.6%)	169.2 (13.8%)	15.0	420.8
Subsidence	279 (3.4%)	109.7 (6.7%)	0.4	34.9	307 (4.8%)	91.9 (7.5%)	0.4	32.5
Mining act.	116 (1.4%)	14.9 (0.9%)	10.7	231.3	115 (1.8%)	14.9 (1.2%)	11.2	273.0
Dump site act.	40 (0.5%)	6.8 (0.4%)	4.2	117.9	36 (0.6%)	6.7 (0.5%)	4.5	133.6
Construction site act.	5 (0.1%)	1.6 (0.1%)	2.9	183.0	1 (0%)	0.1 (0.0%)	6.1	178.7
Potential landslide	2349 (28.4%)	310.0 (18.8%)	12.5	327.6	1765 (27.8%)	215.1 (17.6%)	13.0	335.4
Potential subsidence	2227 (26.9%)	323.5 (19.6%)	1.6	143.0	1856 (29.2%)	242.0 (19.8%)	1.5	146.0
Potential uplift	22 (0.3%)	1.7 (0.1%)	1.4	144.4	15 (0.2%)	1.3 (0.1%)	1.0	133.1
Unclassified	737 (8.9%)	79.0 (4.8%)	3.1	172.9	307 (4.8%)	33.6 (2.7%)	4.1	210.5

5.3.3. Local scale analysis – Case studies

In this section, 7 specific cases, classified according to the different triggering mechanism, are illustrated, showing the large applicability in different scenarios of the MAC algorithm on interferometric results.

5.3.3.1. *Landslide – the case of Maratea (Basilicata Region)*

Figure 24 illustrates the case study of a moving area located at Maratea, in the Potenza province (Basilicata Region, Italy). The Maratea Valley is known for being historically affected by large complex deformational phenomenon which are thought to be firstly controlled by the geological setting and, secondly, by the active role played by tectonics in that area (Berardino et al., 2003; Rizzo, 2002). As a result of several studies, it is reported that the village of Maratea rises on a composite DSGSD (deep-seated gravitational slope deformation) which can be described as a spreading evolving lower-down into a large and deep flow, due to the superimposition of permeable and rigid masses (carbonate units, breccia and detritus) upon plastic ones (clayey flysch). The movements monitored on the surface and in depth, referred to the period 1992-1996 and reported by (Rizzo & Limongi, 1997), are both of the order of 2-4 cm/year.

The slope deformation affecting the study area has remained active for the reference period March 2015-December 2018 as it is recorded by both the ascending and descending MAC database; an overall mean LOS velocity of 1.5 cm/year and -1.6 cm/year respectively for both geometry of acquisition is reported, where the positive value indicates movement towards the satellite while the negative sign implies movement away from the satellite. Within the boundaries of the detected moving areas, a peak velocity of -2.6 cm/year on the eastern sector is reached (Figure 24a; Figure 24b). The areal extent of the detected deformation ranges between 2.4 km^2 (descending MAC) and 2.6 km^2 (ascending MAC) and it covers almost all the residential settlements located on the northern side of the Mt. San Biagio.

Focusing on the deformation pattern, the entire ground displacement follows a general westward horizontal direction ($|V_V|/|V_H| < 1$). However, while the steepest and easternmost sector of the MAC (represented with light-green MPs) has moved slightly downward, the flat sector located on the valley floor has undergone a relative upward deformation (Figure 24c). By observing the ground deformation as seen by satellites from two points of view (ascending and descending geometry of acquisition), the kinematic behaviour of the DSGSD can be inferred: a progressive downslope shift from negative to positive vertical displacements can be linked to a semi-circular or compound failure plane (Frattini et al., 2018). The real velocity vectors (Figure 24d) have recorded a maximum rate of deformation of 3.8 cm/year, thus falling inside the range of the GPS measures observed by Rizzo, 2002) in the inhabited area of the valley.
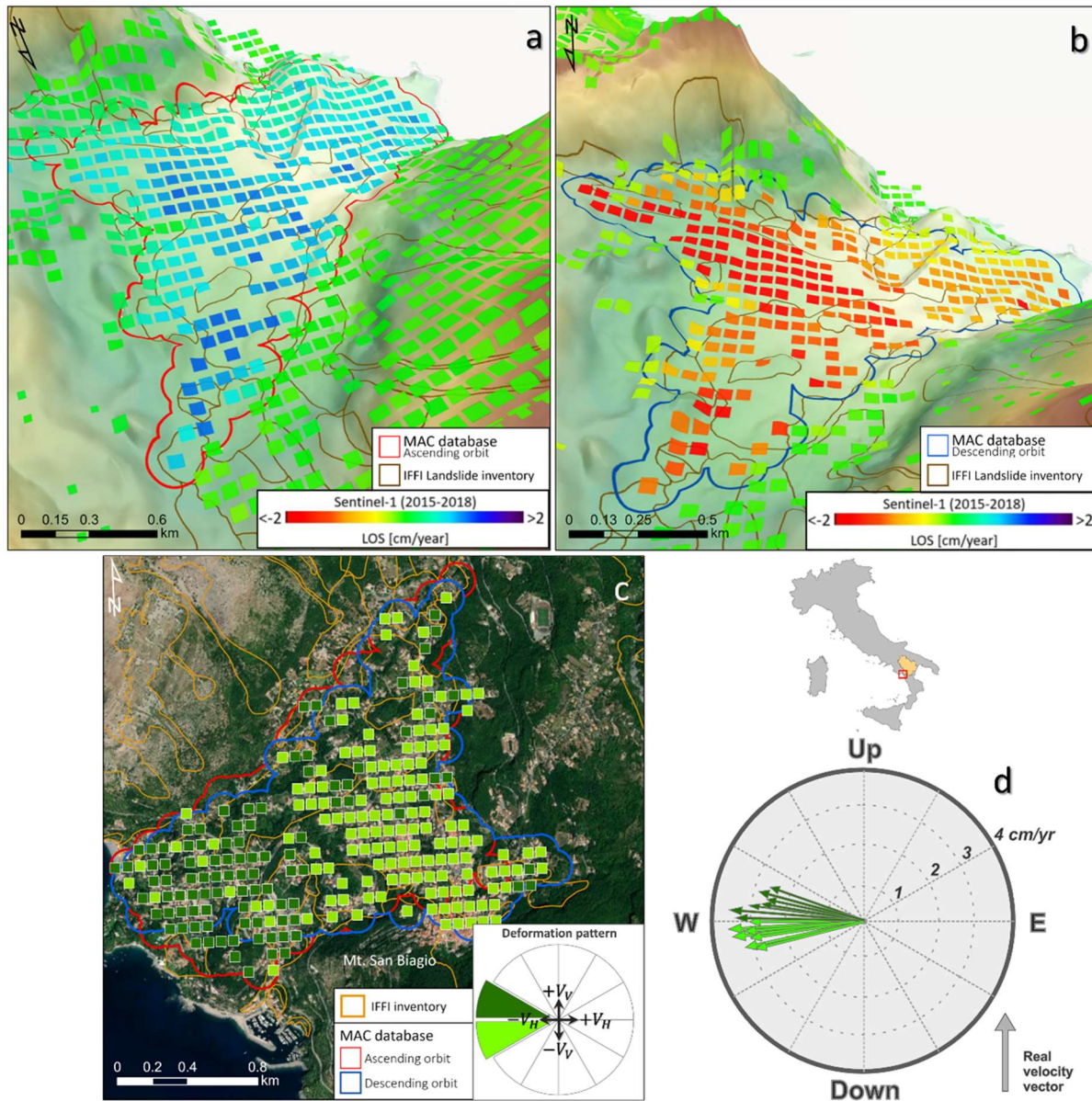
Figure 24. Maratea municipality (Basilicata Region) results. a,b) 3D view of the ascending (a) and descending (b) deformation velocity maps; (c) vertical and horizontal component of the deformation velocity obtained by combining LOS ascending and descending MACs; (d) polar plot depicting a representative sample of the "real" velocity vectors computed for each MP within the two MAC of Maratea showing the magnitude and the direction of the movement on the Up-Down and East-West plane.

5.3.3.2. Volcanic related process – the case of the Phlegraean Fields (Campania Region)

In Figure 25, the case of one of the most studied volcanic areas in the world, the Phlegraean Fields caldera, is presented. This area is located along the western coast of the Campania Region (southern Italy) in a largely urbanized territory, including the western sector of the city of Naples. The present physical aspect of the Phlegraean Fields can be considered as the result of two recent major eruptions, i.e., the Campanian Ignimbrite (~39 ka) and the Neapolitan Yellow Tuff (~15 ka), which produced two nested calderas covering together a 16 km wide area (Acocella, 2008; Bevilacqua et al., 2020; Fedele et al., 2011). At least

since 2nd Century BC the Phlegrean Fields have undergone a bradyseism phenomenon consisting of several uplift and subsidence episodes. Since 2005, after a 20 years-long phase of subsidence, the caldera has been experiencing a phase of uplift movement, registering a slow acceleration in the last years (Bevilacqua et al., 2020; Chiodini et al., 2017; D'Auria et al., 2015; Patra et al., 2019) and a total maximum vertical displacement in the Pozzuoli area of ca. 83 cm (recorded up to June 2021) (Osservatorio Vesuviano, 2021).

The recent uplift movement is confirmed by the A-DInSAR dataset and the consequent MAC database, where both ascending and descending deformation clusters fall East and West of the Pozzuoli Bay covering an area of 33.5 km2 and of 33.2 km2, respectively for a total area of 58.2 km2 (without considering the overlapping parts of the two clusters). For the reference period March 2015-December 2018, the recorded peak velocities of 6.2 cm/year (ascending geometry) and of 6.5 cm/year (descending geometry) are referred to MPs falling exactly inside the Pozzuoli district (Figure 25a; Figure 25b).
Considering the LOS velocity projected on the Up-Down and East-West plane, it emerges that the spot with the highest deformation rate (i.e., the Pozzuoli district) shows MPs holding vertical component of displacement considerably greater than the horizontal ($|V_V|/|V_H| > 1$); that zone seems to acts as a ridge zone dividing the western sector and the eastern sector into areas uplifting with a westward and eastward component respectively (Figure 25c). The polar plot represented in Figure 25d displays that the maximum displacements have acted vertically close to the centre of Phlegraean Fields caldera (located in the town of Pozzuoli) with the magnitude of the "real" velocity reaching a peak velocity of 8.6 cm/year during the monitored period (Osservatorio Vesuviano, 2018). On the other hand, on the outermost parts of the clustered areas, the entity of the displacements decreases in intensity with the horizontal component of motion increasing in parallel. The highlighted results confirm that the ground deformations are located over a semi-circular annular Region and follow a radial pattern both in terms of intensity and orientation (Bevilacqua et al., 2020; Manconi et al., 2010).
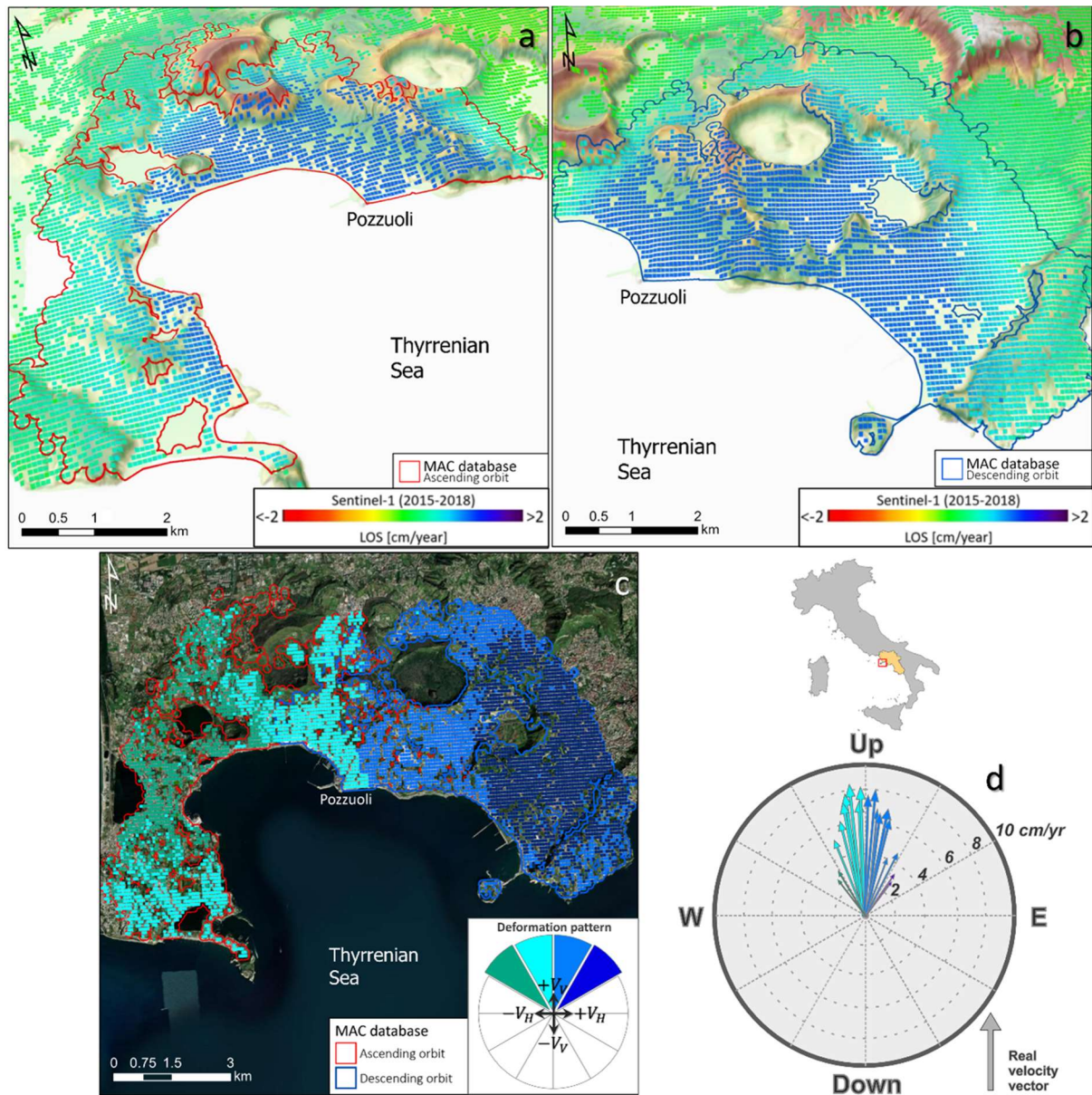
Figure 25. Phlegrean Fields (Campania Region) results. (a, b) 3D view of the ascending (a) and descending (b) deformation velocity maps; c) vertical and horizontal component of the deformation velocity obtained by combining LOS ascending and descending MACs; (d) polar plot depicting a representative sample of the "real" velocity vectors for each MP within the two MACs, showing the magnitude and the direction of the motion on the Up-Down and East-West plane.

5.3.3.3. Earthquake-induced deformations – the case of the 2016 Amatrice-Visso-Norcia seismic sequence (central Italy)

In Figure 26 the mapped coseismic surface deformations related to the *Amatrice-Visso-Norcia* seismic sequence that struck central Apennines between August and October 2016, as a result of the S1 P-SBAS processing are shown. The sequence initiated on 24 August 2016 with the Moment magnitude (Mw) 6.0 *Amatrice* earthquake claiming at least 290 lives and causing widespread damages. On 26 October 2016,

another Mw 5.9 earthquake hit the village of *Visso*, which is ca. 25 km to the North of *Amatrice*. The largest seismic event, the Mw 6.5 *Norcia* earthquake, occurred on 30 October and nucleated in an area comprised between the two previous events. The following months an intense aftershock activity involved the activated zone, which runs NNW-SSE parallel to the axis of the central-northern Apennines (Improta et al., 2019b).

The detected moving areas (which occupy a total area of ca. 70 km × 70 km) correspond to coseismic ground motions and failures related to the *Amatrice-Visso-Norcia* seismic sequence. Surface deformations due to different phenomena (e.g., slope failures, anthropic activities…) that possibly affected the same area during the monitored period March 2015 - December 2018 could not be inferred solely on the basis of the mean deformation velocity.

The ascending and descending MAC database predominantly recorded displacements in the areas surrounding the activated zone, which is roughly located along the three mainshock epicentres (Figure 26a and Figure 26b, respectively). The absence of interferometric data in the vicinity of the hypocentre area has to be ascribed to the loss of coherence because of temporal changes in the scattering conditions on the ground (Kobayashi et al., 2013). The observed displacements are characterized by two NNW-SSE striking deformation lobes with positive and negative velocities depending on their position respect to the epicentres and to the geometry of acquisition of satellites (Figure 26a ;Figure 26b) with values reaching a peak of 5.8 cm/year on the left of the activated zone, where it is also recorded a higher density of MAC. This trend clearly confirms the widely accepted hypothesis of a normal fault mechanism for the Mw 5.9– 6.5 mainshock events, where the coseismic collapse of the hanging wall accounted for the main visible displacement.

Through the analysis of the displacement on the Up-Down and East-West plane, the area hit by the 2016 seismic sequence reveals a complex pattern of deformation (Figure 26c). Despite the large number of MPs throughout the site of interest holding very different values of ratio between the vertical and the horizontal velocity components, it is here highlighted the E-W component of deformation concerning the two deformation lobes, with the eastern one (corresponding to the foot wall area) shifted eastward and the western one (corresponding to the hanging wall area) shifted westward in agreement with previous studies (Cheloni et al., 2017; Huang et al., 2017). Moreover, a clear subsidence pattern is partially visible in correspondence to the Amatrice village (marked with red and orange colours in Figure 26c). The polar plot represented in Figure 26d displays the "real" vectors with computed velocities considerably higher than the LOS ones: a maximum horizontal displacement of 8.4 cm/year has affected the central part of the left lobe (~7 km NW from Norcia village).
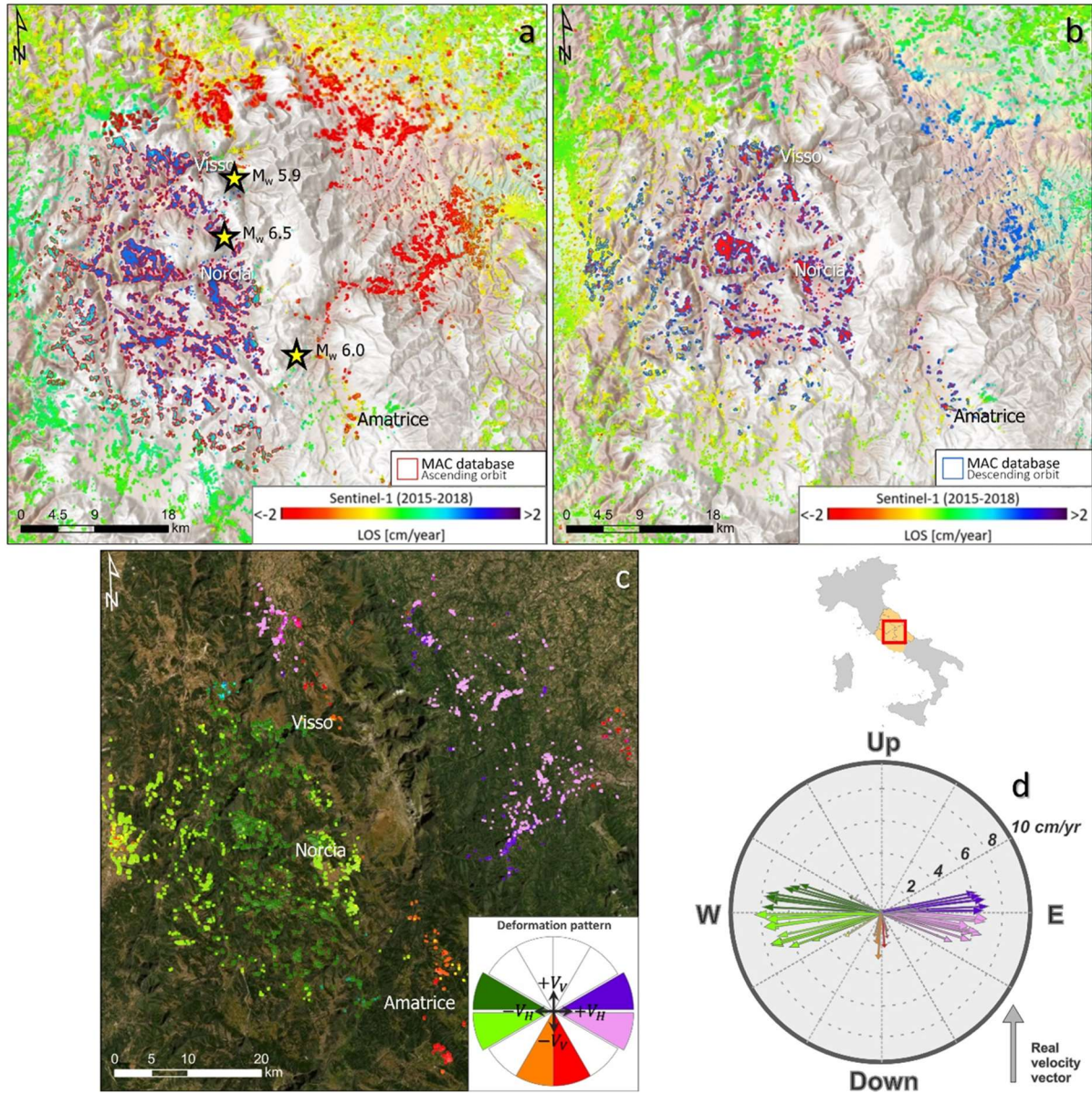
Figure 26. 2016 *Amatrice-Visso-Norcia* seismic sequence results. (a, b) Ascending (a) and descending (b) deformation velocity maps of the area hit by the seismic sequence; (c) vertical and horizontal component of the deformation velocity obtained by combining LOS ascending and descending MACs; (d) polar plot depicting a representative sample of the "real" velocity vectors showing the main direction and magnitude of the observed movement. The yellow stars show the epicentres of the three mainshocks as reported by Istituto Nazionale di Geofisica e Vulcanologia (INGV).

5.3.3.4.   Subsidence – the case of Montemurlo (Tuscany Region)

The municipality of *Montemurlo*, in the Prato province (Tuscany Region), along with the Firenze-Prato-Pistoia basin, is an area historically known for being affected by subsidence phenomena. In particular, the southwestern portion of the municipality hosts a large industrial area, whose presence could be connected to the recent seasonal variation of the water table due to groundwater withdrawal (del Soldato al., 2019).

The MAC database shows two cluster of MPs in the ascending geometry with a total areal extent of 0.4 km^2 while the two equivalent detected descending deforming area are slightly larger (0.5 km^2). For the reference period March 2015-December 2018 the area has registered a peak velocity of -3.9 cm/year and of -4.2 cm/year in the cluster centres of the ascending and descending geometry, respectively (Figure 27a; Figure 27b).

In Figure 27c, the computed components of displacement are visible, highlighting an expected pattern of deformation: the entire area suffered land subsidence resulting in a high value of mean absolute ratio between the vertical and horizontal components ($|V_V|/|V_H| > 1$), with the western sectors moving relatively towards East and the eastern sectors to the opposite direction (i.e., towards West). This trend implies that the sinking of the central zone of the deforming areas (where land subsidence effects are more evident) may drive the surrounding movements. Hence, the interferometric results suggest the existence of a cone-shape depression radiating away from the centre of the clustered areas due to groundwater overexploitation. The polar plot represented in Figure 27d shows a maximum "real" velocity (i.e., computed on the Up-Down and East-West plane) of ca. 4.8 cm/year. Del Soldato et al. (2019) report average interferometric measures of ground velocities reaching approximately −10 cm/year in July 2017 and June 2018 which nevertheless relate to rates of ground subsidence of particular intensity occurred in specific time spans.
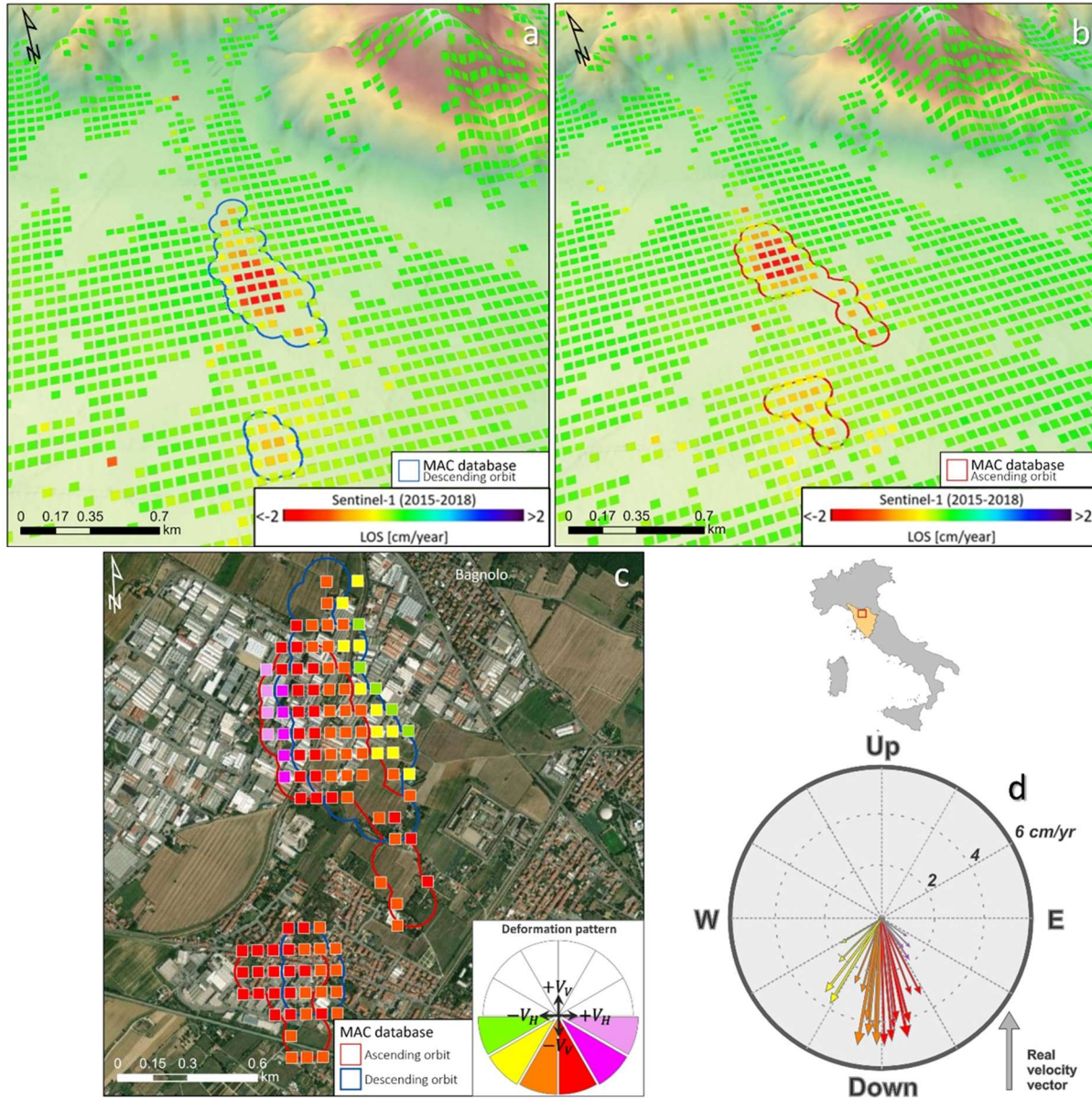
Figure 27. *Montemurlo* municipality (Tuscany Region) results. (a, b) 3D view of the ascending (a) and descending (b) velocity deformation maps; (c) vertical and horizontal component of the deformation velocity obtained by combining LOS ascending and descending MACs; (d) polar plot depicting a representative sample of the "real" velocity vectors showing the main direction and magnitude of the observed movement.

5.3.3.5.   Dump site activities – the case of Masseria del Re (Campania Region)

Masseria del Re, a locality that belongs to the municipality of Giugliano in Campania, gave its name to the storage site located in the province of Naples (southern Italy). As a result of the recent crisis related to the disposal of waste in the Campania Region, the site has been designated since January 2006 as a storage space for tons of fuel pellets produced from waste of different origins. The accumulation of material continued until August 2007 up to the complete saturation of the site, where more than 5 million of fuel

pellets have been amassed on a surface area of about 460,000 m^2. The regional authorities announced the start of work concerning the removal of waste from the site in May 2016.

The S1 derived MAC database referred to the monitored period March 2015-December 2018 shows ground deformation both in ascending and descending geometry covering the whole storage site (Figure 28a; Figure 28b). All the clustered MPs inside the two moving areas are characterized by negative LOS values, thus testifying the overall lowering of the zone where the fuel pellets are set. The estimated areal mean velocity ranges between -2.1 cm/year (as seen from the ascending orbit) and -2.5 cm/year (as seen from the descending orbit), while locally it is reached a maximum velocity of -5.4 cm/year.

While looking at the deformation pattern on the Up-Down and East-West plane (Figure 28c), it emerges the clear predominancy of surface deformations acting vertically throughout the storage site ($|V_V|/|V_H| > 1$). The polar plot represented in Figure 28d shows a maximum "real" velocity of ca. 6.9 cm/year, thus highlighting once again the importance of combining data from different geometries of acquisition for a better estimation of the entity of the displacement. Two main causes can be reasonably connected to the observed lowering deformation over the area of interest: (i) the fuel pellets load imposition with the subsequent ground lowering and (ii) the leachate production, collection and extraction which are regularly performed on site by the facility operators. As a result, it can be concluded that the observed motion is mainly related to the waste material load, rather than the local ground surface.

Figure 28. *Masseria del Re* storage site (Campania Region) results. (a, b) 3D view of the ascending (a) and descending (b) deformation velocity maps; (c) vertical and horizontal component of the deformation velocity obtained by combining LOS ascending and descending MACs; (d) polar plot depicting a representative sample of the "real" velocity vectors showing the main direction and magnitude of the observed movement.

5.3.3.6. Mining activities – the case of Santa Barbara (Tuscany Region)

The former Santa Barbara brown coal mining area is settled in the municipality of *Cavriglia*, in the Arezzo province (Tuscany Region). The mining activities started in 1860 and ended in 1994 due to exhaustion of the lignite basin. The topography of the site has drastically evolved with the site exploitation, with the formation of hills and ground excavations, where the latter have become today visible lakes (Bozzuto & Geroldi, 2021). In particular, the slope close to the *Allori* Lake, due to his orientation eastwards and its

artificial bare terrain, turns out to be greatly monitored from S1 satellite systems. The former mining site is currently object of an Environmental Rehabilitation Plan (Bozzuto & Geroldi, 2021).

Two ascending clusters (Figure 29a) are located on the above-mentioned slope, where the largest (0.5 km^2) is just in front of the *Allori* lake and the *Bomba* area (which is the seat of many warehouses) while the smallest (~0.1 km^2) is ca. 700 m north of the first cluster; one descending MAC (Figure 29b), with an area extent of 0.6 km^2, occupies roughly the same position of the largest ascending cluster. A peak of LOS velocity of about -9.8 cm/year was recorded in the ascending geometry, during the monitored period March 2015 - December 2018, as well as an average displacement rate of -4.1 cm/year and -2.7 cm/year in the two ascending deforming areas. For what concerns the descending geometry, an average LOS velocity of 2.1 cm/year has been measured for the same area. On the contrary, the *Bomba* area appears to have slightly moved toward the satellite sensor (uplift movement) without generating any MAC, due to an estimated average ascending LOS velocity of 0.6 cm/year and an average descending LOS velocity of 0.3 cm/year in that zone (Figure 29a; Figure 29b). All the reported observations are consistent with the latest interferometric measurements available on the area of interest (geoportale.lamma.rete.toscana.it).

In Figure 29c the combination of ascending and descending LOS velocities has allowed a 2D reconstruction of the velocity vectors in the Up-Down and East-West plane: the overlapping clustered areas show an evident East-Down direction of movement, thus behaving similarly to a landslide process. This evidence is supported by the presence of several slope instabilities spread throughout the site of interest (Figure 14). The polar plot represented in Figure 29d shows a maximum "real" velocity of 9.2 cm/year within the perimeter of the outlined clusters, which is comparable to the peak LOS velocity recorded in the ascending A-DInSAR dataset. The "real" velocity vectors related to the Bomba area exhibit upward movement which has reached 1.5 cm/year at most during the reference period. In the context of abandoned coal mine sites which have undergone flooding, the uplift movement could be ascribed to the long-term complex interaction between the mine goaf, the rock mass and the fluctuation of the groundwater level (Zhao & Konietzky, 2020). The development of different lakes in the former mining area over the past century (e.g., the *Allori* Lake) could have played a significant role in triggering the observed surface deformations.
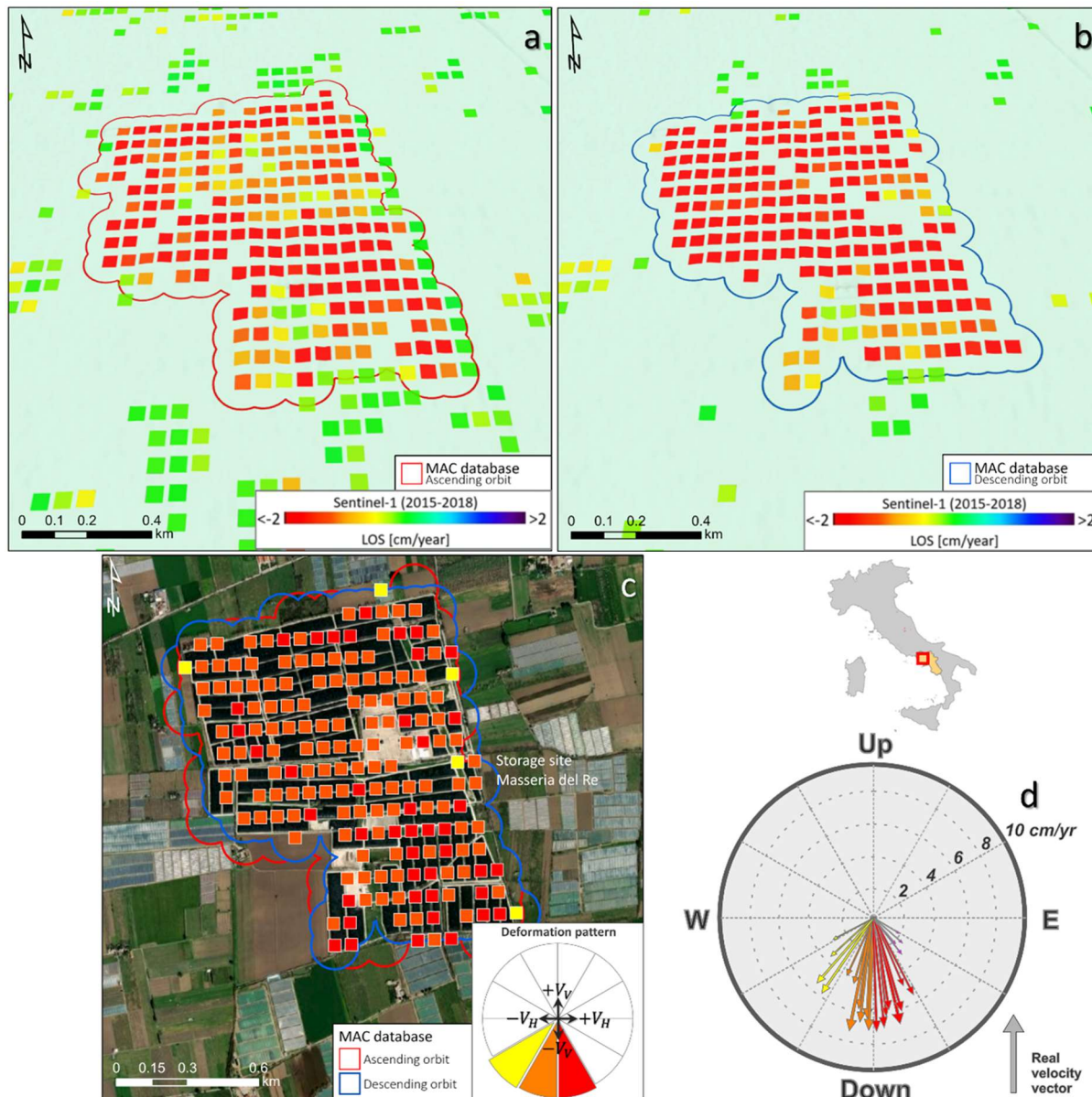
Figure 29. Santa Barbara (Tuscany Region) results. (a, b) 3D view of the ascending (a) and descending (b) deformation velocity maps; (c) vertical and horizontal component of the deformation velocity obtained by combining LOS ascending and descending MACs; (d) polar plot depicting a representative sample of the "real" velocity vectors showing the main direction and magnitude of the observed movement.

5.3.3.7. Construction site activities – the case of the freight village Amerigo Vespucci (Tuscany Region)

The freight village *Amerigo Vespucci* is located in the municipality of *Collesalvetti* and represents the main container station for the Livorno province (Tuscany Region), due to its strategic vicinity to the developed infrastructural area where it sits. Today, the freight terminal covers a surface area of ca. 3,000,000 m², while the creation of the hub dates to late '80 when the area was designed as a building site to be used

for commercial and logistical-industrial use. Since the construction of the first warehouses, the area has been experiencing ground lowering due to the consolidation processes as a consequence of the buildings load imposition (Ciampalini et al., 2019), which resulted in structural damages to the sheds and consequent economic losses.

Two overlapping ascending and descending MAC with comparable extension concern the area corresponding to the freight village Amerigo Vespucci during the monitored period March 2015-December 2018 (Figure 30a; Figure 30b). It is reported an area mean velocity of -1.8 cm/year for both geometries, while locally observed ground displacements have reached a maximum of -3.7 cm/year. The freight station site appears to have been generally subducted, however few MPs within the clustered areas show to be within the range of stability (velocity ranging between -0.3 – 0.3 cm/year), thus highlighting that the area has possibly undergone a process of differential settlement, as also testified by Ciampalini et al. (2019).

In Figure 30c the components of the LOS velocities on the Up-Down and East-West plane are shown, as a result of the combination of the ascending and descending interferometric data. The ground has displaced in downward direction (V_V <0 and $|V_V|/|V_H|$ > 1) with a mean "real" velocity equal to 2.1 cm/year considering the entire area, while it is recorded a local peak of 7.2 cm/year (Figure 30d) in the north-eastern sector of the freight village. The observed deformation pattern confirms the ongoing process of settlement caused by the consolidation of a highly compressible clay layer in response to building loads, while the observed measures are in good agreement with the rate of velocity derived for the same period from past studies conducted over this area (Ciampalini et al., 2019).
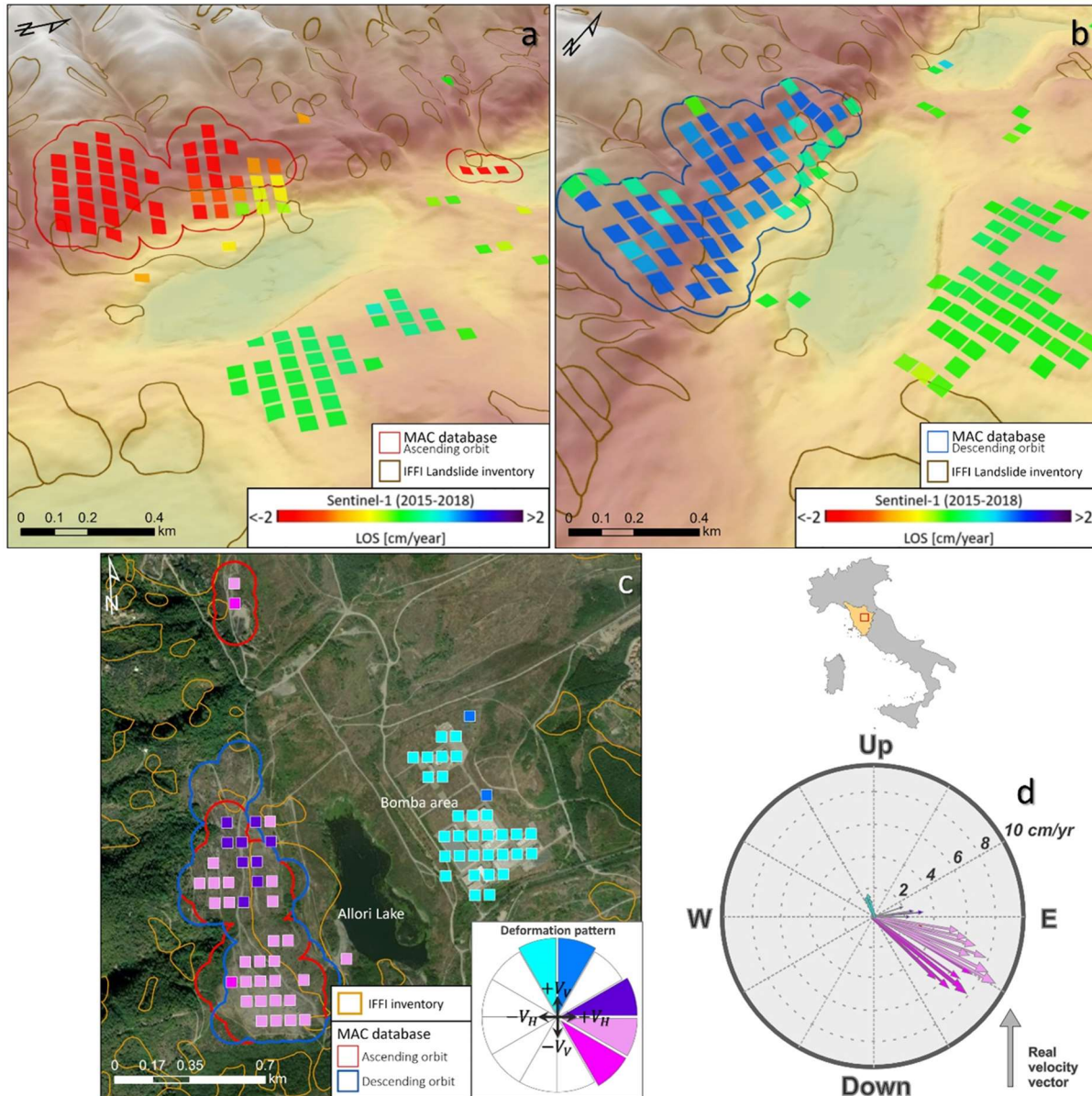
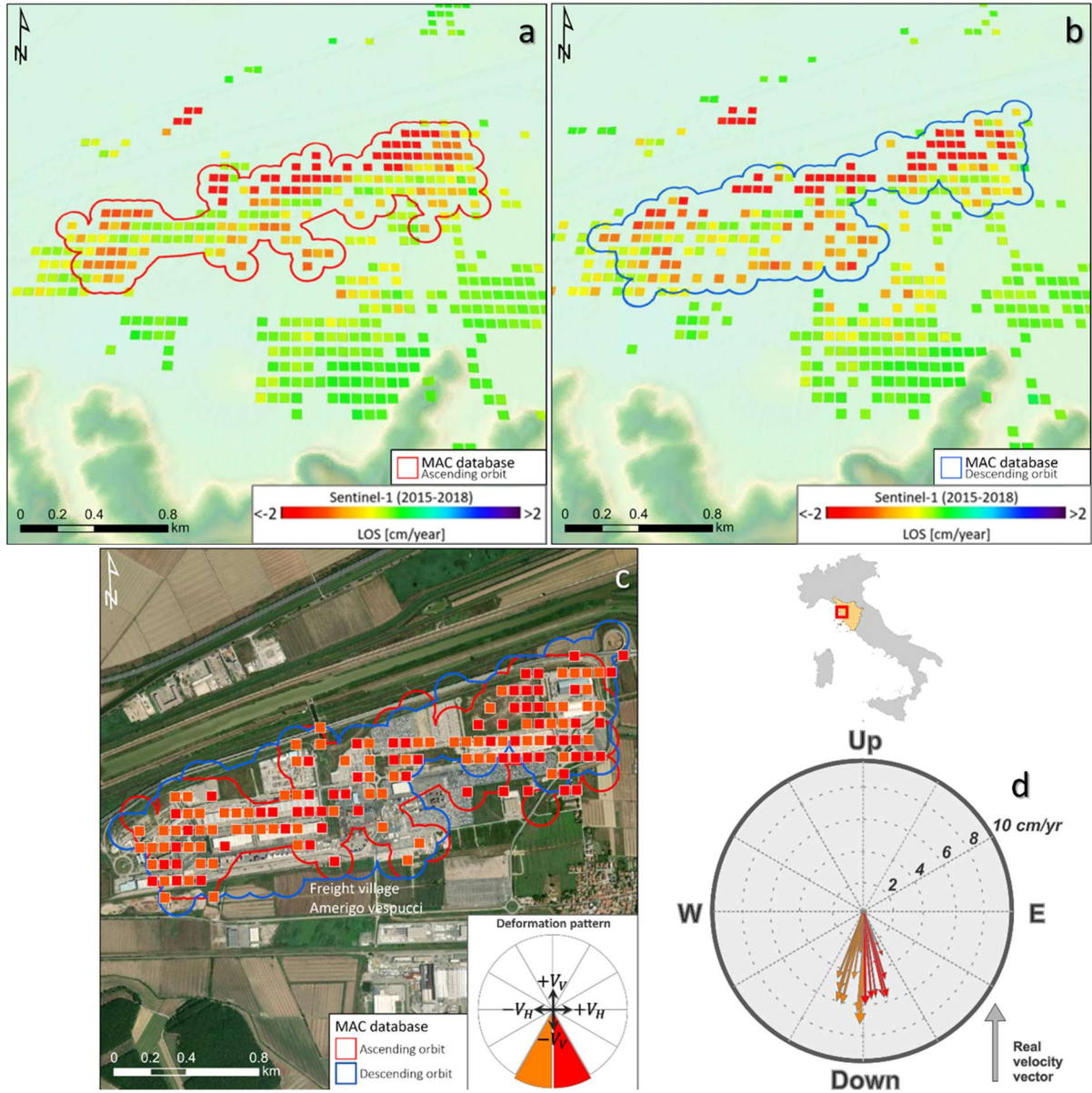Figure 30. Freight village *Amerigo Vespucci* (Tuscany Region) results. (a, b) 3D view of the ascending (a) and descending (b) deformation velocity maps; (c) vertical and horizontal component of the deformation velocity obtained by combining LOS ascending and descending MACs; (d) polar plot depicting a representative sample of the "real" velocity vectors showing the main direction and magnitude of the observed movement.

6. Machine Learning-based assessment of ground displacement source

In this section of this Thesis work it will be illustrated a Machine Learning (ML) based method for the automated recognition and classification of the cause of ground displacements detected through the analysis of A-DInSAR deformation maps. The methodology is aimed at exploiting large-scale C-band S1 data to assess the triggering factor of potential active displacements, based on the quantification of the statistical association between radar-based features and a set of explanatory variables retrieved from the most common morphometric and geohazard inventories.

The test site covers Tuscany and all the Northern Italian regions (Valle d'Aosta, Piemonte, Lombardia, Liguria, Emilia Romagna, Veneto, Trentino-Alto Adige and Friuli-Venezia Giulia).
The interferometric data selected for the presented methodology are made available thanks to the partnership with the Presidency of the Council of Ministers–Department of Civil Protection, through the IREA-CNR/DPC and CPC-UNIFI agreements. This work was also partially funded by the IREA CNR/Italian Ministry of Economic Development DGS-UNMIG agreement and H2020 EPOS-SP (GA 871121).

The proposed original approach regards the implementation of the Random Forest (RF, Breiman, 2001), a well-known and widely implemented ML technique, to automatically detect and classify ground deformations sources across Northern Italy using the spatial characteristics of A-DInSAR data. The significance of such a work relies on the necessity of periodically updating and assessing the geohazard activities detectable from continuously updated space-borne radar data. In this approach, deformation maps resulting from the processing performed via P-SBAS (de Luca et al., 2017; Zinno et al., 2015) were used to retrieve MPs with outlying mean velocity values and spatially associate them to a set of predictors, while the time series of displacements were not considered in this case. Morphometric features of the landscape, land cover, and geohazard inventories were taken into consideration as predictors in order to create a model for the purpose of classifying the origin of ground deformations in urban, non-urban, and forested environments. RF approach creates an ensemble of decision trees, established from randomly generated portions of the training data, whose votes are considered together to finalize a data-driven forest model. The presented innovative idea is to adapt and apply RF at local scale to A-DInSAR MPs. A major breakthrough regards the introduction of new explanatory variables which can be obtained by the post-processing of P-SBAS data: an index expressing the motion trend in the East-West and vertical planes. Besides these, the method relies on other easily accessible input parameters extracted from a Digital Elevation Model (DEM) and from open access catalogs (e.g., landslide inventories). The procedure can also be tested in other geographical areas where DEM, land cover map, and geohazard inventories are available. Therefore, it is demonstrated the feasibility of a simple ML-based classification scheme for assessing the source of active motions, which is a primary task in the management of large A-DInSAR datasets in the framework of supporting stakeholders in land management and civil protection purposes.

6.1. Materials and methods

The flowchart of the presented methodology is synthesized in Figure 31. Input data and rationale of the procedure are explained in more detail in the following sections.
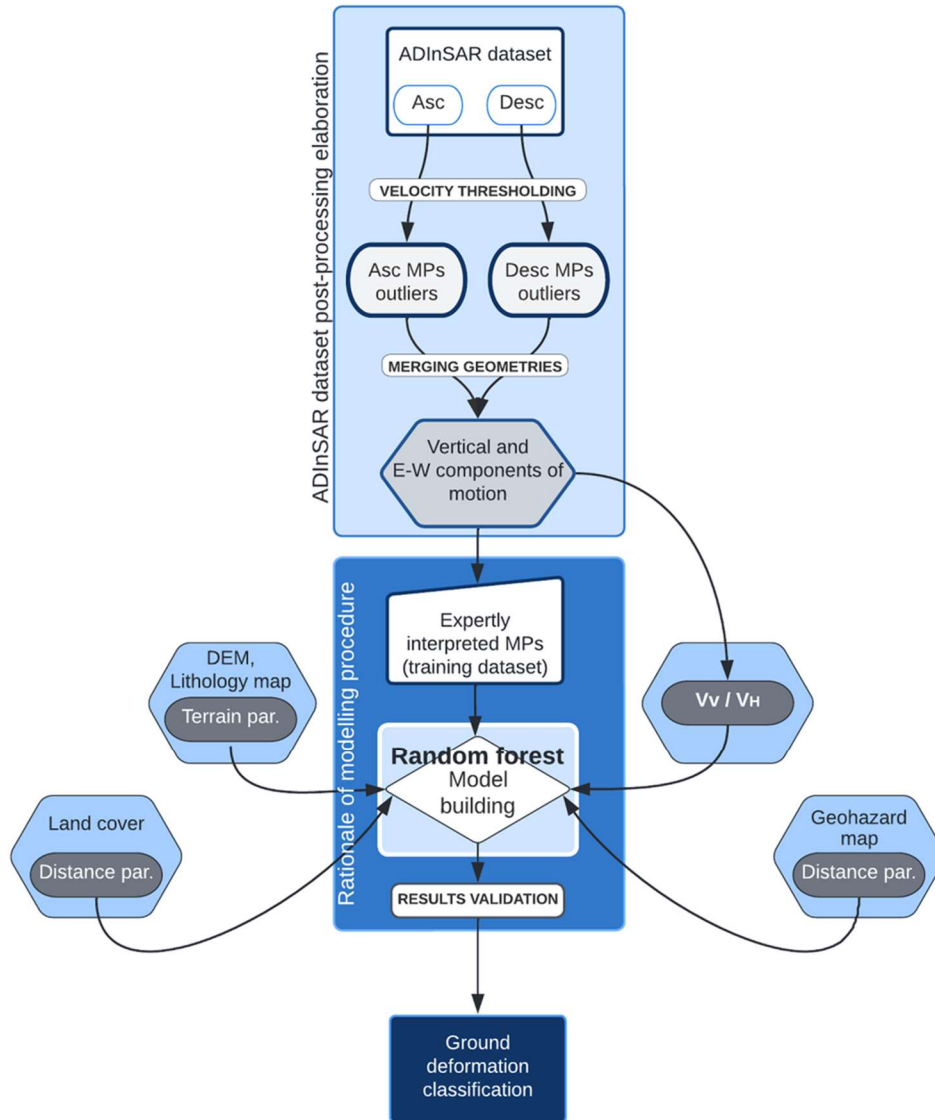
Figure 31. Workflow of the ML-based methodology for the classification of ground deformations detected from the post-processing analysis of A-DInSAR data.

6.1.1.   A-DInSAR dataset post-processing elaboration

S1 ascending and descending datasets covering Northern Italy regions (with Tuscany region included), acquired from 23 March 2015 to 26 December 2018 were employed and processed by means of P-SBAS data, with a resulting pixel resolution of about 80 m x 80 m. Time series of displacement, despite being available for each MPs, are not employed for this work proof of concept since the presented approach is based on the mean velocity for each pixel. To exclude incoherent points, only pixels holding a temporal coherence value greater than 0.9 were retained for this analysis. 2,749,737 velocity MPs were counted in the ascending orbit, while 2,594,041 MPs compose the descending velocity map. The overall MPs density ranges from 19.2 to 18.1 points per km^2 for the ascending and descending pass, respectively. The higher density and significantly cluster are in urban areas compared to forested and agricultural

areas. The mean velocity value (μV) of both datasets is around zero (Figure 32), while the standard deviation (σ) of the velocity distribution stands at ±0.2 cm/year which usually corresponds to the displacement interval related to stable ground conditions (Catani et al., 2012; Raspini et al., 2018).

Given that the velocity values in the two datasets share approximately bell-shaped distributions (Figure 32), the standard deviation of the sample was used as a cut-off threshold for identifying MPs with relatively high deformation rates. In particular, any value out of the range defined by μV±3σ (i.e., ± 0.6 cm/year) can be considered as an outlier (Santos, 2020) and therefore is retained. This criterion allowed to restrict the analysis to a total of ~145,000 unstable ground measurements (Figure 32; 76,464 MPs and 68,193 MPs from the ascending and descending passes, respectively), finally selected as targets of our work.
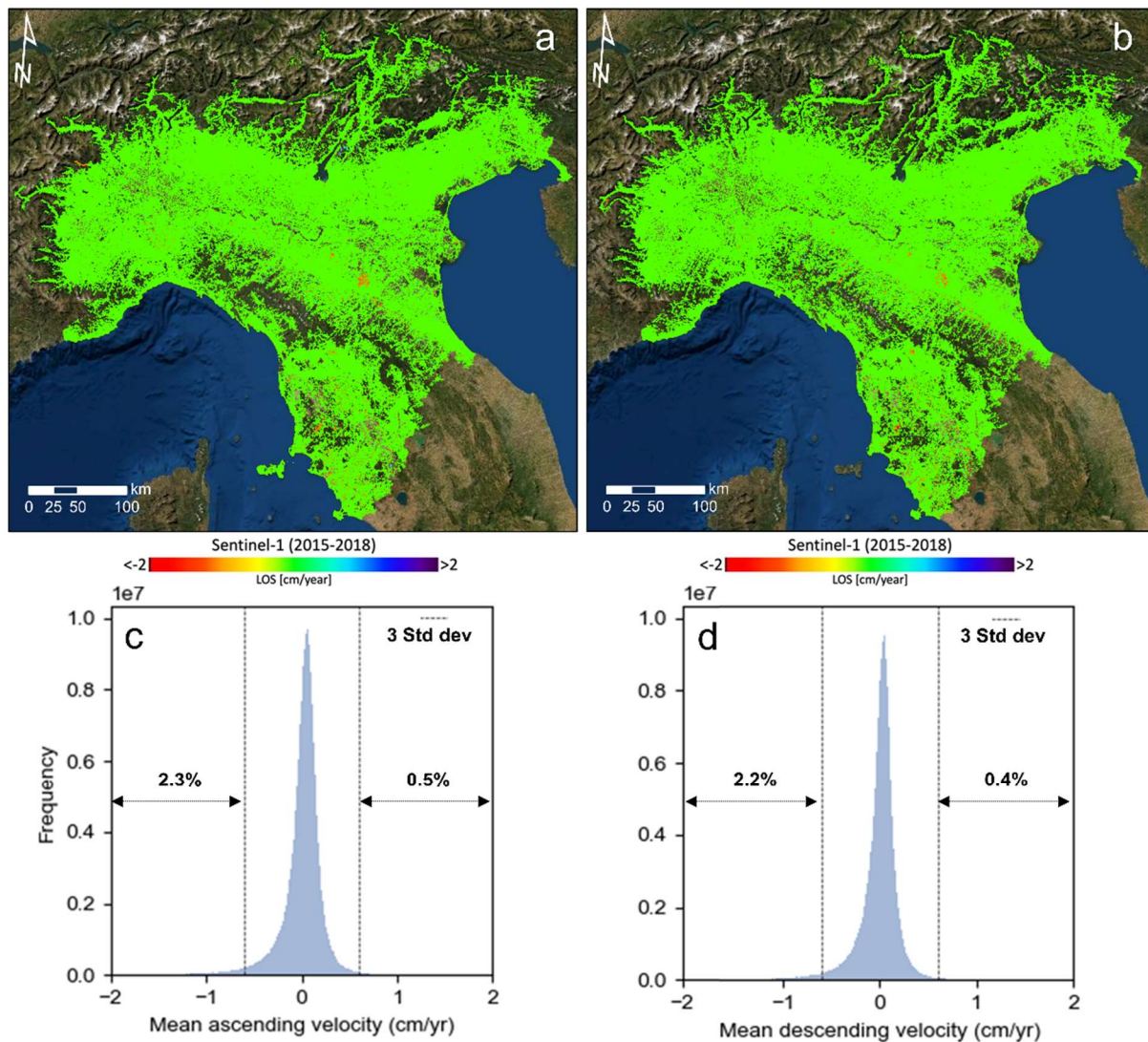

Figure 32. (a) Ascending and (b) descending velocity maps and related frequency distribution of the (c) ascending and (d) descending mean velocity (μV). The cut-off threshold, namely 3 times the standard deviation (3σ), is represented with the dashed line.

Furthermore, being a key-step for the analysis, the information from the two different LOS geometries were combined to resolve the E–W (V_H) and vertical (V_V) velocity components (Notti et al., 2014) by exploiting the S1 P-SBAS data architecture. Since the P-SBAS mean deformation velocity maps are divided following an 80 m x 80 m geographical regular grid (De Luca et al., 2017), the displacement vector decomposition was computed for those pixels accounting for coincident LOS ascending and descending measurements. This can be done by solving the equations (8) reported in Notti et al. (2014).

Despite not being available for every MPs, the reprojection of the displacement rate on the E–W and vertical plane will constitute an important explanatory information in the proposed model-building framework. As a result of the post-processing phase, our univocal final dataset counts 20,212 MPs.

6.1.2. Model parameters
6.1.2.1. Terrain parameters

In this work a 10 m cell-size grid DEM, derived from TINITALY/01 project (Tarquini & Nannipieri, 2016) released by INGV was used. Some of the main DEM-derived indices largely employed in landslide susceptibility (Catani et al., 2013) are here used to infer a correlation between the spatial occurrence of MPs with outlying velocities and the motion trigger. In particular, it is here extracted the exact value to every MPs location for the following maps:

- Elevation, which may be indicative of different altitudes tied to various environmental settings (Ayalew et al., 2004) and different types of surface deformations.

- Aspect identifies the downslope direction. The raster pixels are expressed in terms of compass direction that the surface faces at that location with respect to the cardinal direction.

- Slope describes the slope for each raster cell in degrees based on the elevation at each point.

- Curvature, which expressed the relative change in slope and is computed as the second derivative of the surface topography.

- Profile curvature, indicative of the rate at which the slope gradient changes towards the direction of maximum slope.

- Planar curvature, respect to the latter, is calculated orthogonally to the direction of the maximum slope.

- Topographic Wetness Index (TWI) (10), which can be though as a steady state wetness index, where A is the upslope contributing area of each raster pixel and β is the slope angle.

$$ln\left(\frac{A}{\tan\beta}\right)$$

(10)

The characteristics of the soil coverage may be considered as an important driving property to discriminate different causes of ground deformation. The diverse geomechanical and hydraulic properties reflected by the different units were considered. Their spatial distribution is well depicted in the 1:500,000 lithological map of Italy (www.isprambiente.gov.it/progetto-1250-ita). We grouped all the geological formations into eleven classes: (i) granular soils; (ii) debris and alluvial deposits; (iii) clays and compact marls; (iv) massive calcareous rocks; (v) igneous intrusive rocks; (vi) igneous effusive rocks; (vii) schistous rocks; (viii) metamorphic rocks; (ix) evaporite rocks; (x) complex units with mainly pelitic layers; and (xi) complex units with mainly calcareous and arenaceous layers.

6.1.2.2. Distance parameters

Proximity distances to important land motion predisposing factors can be of great help to calculate and to determine the kind of surface displacement. For our analysis we used the areal features contained in the 1:10,000 scale Italian Landslide Inventory (*Inventario dei Fenomeni Franosi in Italia*, IFFI project; Trigila et al., 2007), since it depicts the spatial distribution of past mass movements for the Italian territory. The spatial diffusion of open-pit extraction sites and flooded mining pits were extracted from the III level classification of 2018 CORINE Land Cover map (CLC) at 1:100,000 scale provided by the Copernicus Land Monitoring Service. Here is computed the distance between every MPs and the closest line segment to the polygon features in an equidistance projected coordinate system, where a distance of zero is considered when a MP lies within a polygon.

6.1.2.3. K_{VH} parameter

In order to find out the deformation pattern which may be linked to a particular geological or anthropogenic phenomenon causing ground displacements as seen by SAR satellites, here is deployed novel index based on the interferometric measures reprojected on the E–W and vertical plane. Here is deployed the variable K_{VH} (9) that is the ratio between V_H and V_V, namely the vertical (Up-Down) and the absolute horizontal (East-West) component values (for more references see section 5.2.3)
The use of K_{VH} is meant to introduce a strong attribute regarding the prevailing direction of the movement evidenced by the MP for which the index is calculated. A prevailing negative value of K_{VH} would suggest a downward motion, while a prevailing positive value of K_{VH} would indicate an upward motion. By definition, a value comprised between -1 and +1 would suggest a predominant horizontal movement (i.e., toward West or East).

Table 6. Explanatory parameters used for the definition of the model.

Parameters	Type	Data format	Scale/resolution	Source
Elevation	Cont.	Grid	10m	TINITALY/01
Aspect	Cont.	Grid	10m	derived from TINITALY/01
Slope	Cont.	Grid	10m	derived from TINITALY/01
Curvature	Cont.	Grid	10m	derived from TINITALY/01
Profile curvature	Cont.	Grid	10m	derived from TINITALY/01
Planar curvature	Cont.	Grid	10m	derived from TINITALY/01
TWI	Cont.	Grid	10m	derived from TINITALY/01

Lithology	Categ.	Shapefile (polygon)	1:500,000	ISPRA	
Landslide distance	Cont.	Shapefile (polygon)	1:10,000	IFFI	
Mining sites distance	Cont.	Shapefile (polygon)	1:100,000	CLC2018	
Kvh	Cont.	Grid	80m	derived from A-DInSAR dataset	

Cont. = Continuous   Categ. = Categorical

6.1.3. Random Forest

Random forest (RF) is an ensemble ML method which have become very popular in the past decade due to its excellent performance in handling non-parametric multidimensional classification and regression problems and in showing low chances of overfitting (Breiman, 2001). RF is used to predict an outcome by involving the bootstrapped aggregation of several regression trees. A model is created based on a random selection of the training data from the input dataset, while the data not included are referred to as "out-of-bag" (OOB). The generation of a random subset of features ensures low correlation among decision trees, whose classification outcomes are aggregated to identify the most popular result. Therefore, the majority vote will yield the predicted class and the OOB sample is then used for cross-validation, finalizing the class prediction. Here, the common characteristics retrievable from the velocity maps plus other ancillary maps (i.e., numerical, categorical, and distance-based variables) were investigated as well as combined for explaining different causes of deformation tied to ground moving MPs. In other words, the proposed procedure is concerned with predicting a categorical variable to point-like objects.

The method itself proves to be largely applicable in many scenarios. Several advantages are associated to RF: i) massive datasets and data heterogeneity can be dealt (Genuer et al., 2017), also allowing for the mixed use of categorical and numerical variables; ii) complex data fabric can be managed, and itis capable of accounting for nonlinear high-dimensional features which may interact with each other (Scornet et al., 2015); iii) it performs well in model interpretation and is fast to train. A further key benefit is the ability of RF to statistically evaluate variable importance to the model. Importance is calculated using Gini coefficients, that is how much the model's accuracy decreases when a given variable is excluded. This capability allows for studying the relative contribution of the different explanatory variables on the overall result.

A way to estimate the ensemble error is by computing the "out-of-bag error" (OOBE), which measures the prediction errors of RF through the comparison of the OOB predicted responses against the true responses. The overall performance of the model can be evaluated through diagnostic metrics aimed at assessing how well features tied to a particular category are correctly predicted or miscategorized, by using TP, TN, FP, FN which are true and false positive and negatives, respectively. In this work, following the proper setting of the forest parameters, a balanced dataset for model training was tested and the classification performances are measured by means of Accuracy (11)**Error! Reference source not found.**, True Positive Rate (TPR) (12) and F1-Score (13), which is the harmonic mean of TPR and Precision (14)**Error! Reference source not found.**. These parameters are mathematically detailed below:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

(11)

$$TPR = \frac{TP}{TP + FN}$$

$$(12)$$

$$Precision = \frac{TP}{(TP + FP)}$$

$$(13)$$

$$F1 - Score = \frac{2 \cdot TPR \cdot Precision}{TPR + Precision} \cdot$$

$$(14)$$

6.1.4. Rationale of modelling procedure

To automatically classify the phenomena responsible for triggering the outlying MPs, we need to build the RF model by feeding it with ground truth examples of different classes of expertly interpreted MPs (Farina et al., 2006). In this approach, the focus is on slope deformation caused by landslides, subsidence (due to both natural or anthropogenic sources, such as groundwater pumping, natural compaction, or loading) and mining-associated activities (which might be linked to complex deformation patterns) as they are the most common cause of surface displacements in the test area. They are also considered the outlying MPs arising from other unspecified triggering factors or inherent noisy signals which are still included in the filtered A-DInSAR dataset after the velocity thresholding. Real sites across Northern Italy which have been analysed remotely within the framework of monitoring services (Confuorto et al., 2021) or through field investigations were exploited to build up the training dataset, while uncorrelated signals were included in the model by randomly picking up sparse MPs to reproduce noise and any other undefined ground deformation. As expected, the manual classification of MPs resulted in an imbalanced dataset. Since ML methods for supervised classification tend to provide different performances for the samples in the most populated classes compared with the samples belonging to the least numerous classes (Johnson & Khoshgoftaar, 2019), we adopted a simple data-levelling method to obtain a balanced training dataset. We performed random under-sampling for all the classes based on the number of elements of the least populated class, thus retaining a total of 2,184 features equally divided into four groups (Figure 33): (i) landslide (L); (ii) subsidence (S); (iii) mining-related displacement (M); (iv) noise or other phenomena (ND).
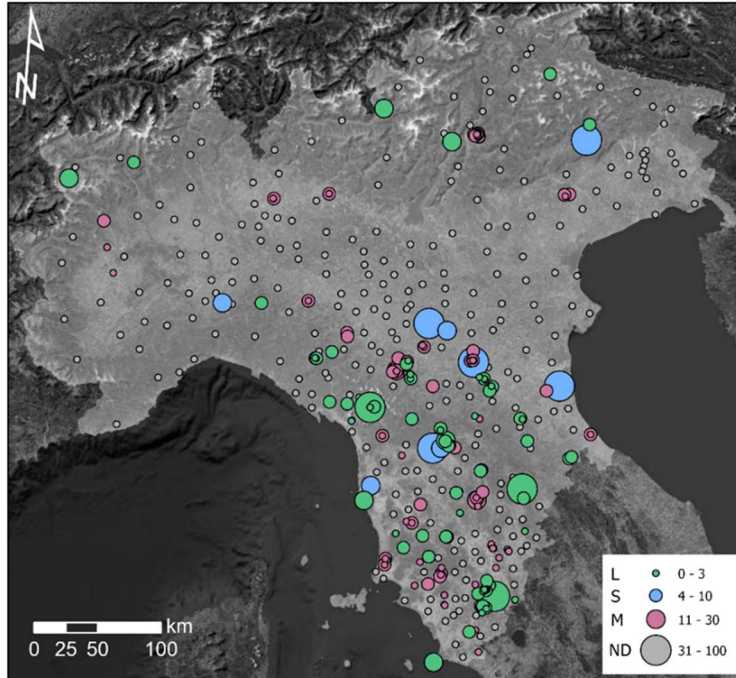
Figure 33. Map depicting the location of expertly interpreted outlying MPs constituting the balanced training dataset. MPs related to landslide (L), subsidence (S), mining-related displacements (M) and noise or other (ND) are shown by means of bubbles whose dimension is proportional to the number of aggregated features within the same discrete phenomenon.

The validation dataset accounts for 15% of the training dataset (i.e., the OOB sample), while the test dataset is constituted by all the remaining MPs which could not be safely interpreted, namely 18,028 MPs.

The RF model settings are tuned regarding the generation of the most accurate model. Different number of trees were here tested trying to minimize OOBE and building a robust model, where every decision tree is created using a random subset of the training data available. Decreasing the number of trees, the computation time of the model is reduced, though not constituting a major concern in our case given the dimensions of our training dataset. To make the ensemble of decision tree an effective predictor, a maximum number of randomly sampled variables for each tree is required, which is here considered empirically by looking at the square root of the total number of explanatory variables. On the other hand, the maximum tree depth (i.e., the maximum number of splits that will be made down a tree) is a data-driven parameter which is here automatically calculated based on the number of trees created and the number of variables included.

6.2. Results of the ML-based methodology
6.2.1. Model tests

Once fixed the RF model settings, they were executed 50 runs on randomized portions of the training dataset, starting from a number of 250 trees and then moving on to 500 and 1000 trees. Considering the full parameters set, the mean and standard deviation of model accuracy obtained through the several

runs were estimated and then the OOBE evaluated as the percentage of incorrect classifications for every predicting class by looking at the most accurate model (Table 7).

Table 7. Mean and standard deviation (Std dev) of accuracy computed for 250, 300 and 1000 trees. Out-of-bag errors (OOBE) displayed for every predicting class and related to the most accurate model.

Number of trees	250	500	1000
Mean accuracy (50 runs)	9.2e-1	9.3e-1	9.3e-1
Std dev accuracy (50 runs)	1.5e-2	1.6e-2	1.6e-2
Best model accuracy	9.5e-1	9.6e-1	9.6e-1
	Best model OOBE	Best model OOBE	Best model OOBE
Landslide (L)	5.2%	4.7%	4.3%
Subsidence (S)	5.9%	5.1%	4.5%
Mining-related disp. (M)	2.0%	2.0%	1.9%
Noise or other (ND)	21.3%	20.5%	20.0%

Results shows that the accuracy values obtained with different number of trees are very similar to each other, with mean values ranging from 9.2e-1 to 9.3e-1 and stable standard deviation values spanning from 1.5e-2 to 1.6e-2. Comparing the OOBE from the best models, a slight decrease of misclassification in the OOB sample was noted, probably correlated with the rise of the number of trees. The latter parameter does not seem to significantly affect the computation time, which is comparable in all cases. It can be asserted that the performances obtained by applying the classification to the training dataset do not considerably enhance across several runs and by varying the number of trees, thus evidencing the practical implementation of a RF method in this kind of approach.

6.2.2. Variable importance and model validation

The best model achieved by setting 1000 trees has taken as a representative to explore its outcomes and characteristics through diagnostic metrics. We evaluated the importance of model variables by looking at the Gini coefficients calculated for every deployed explanatory training variable.

The distance parameter related to CLC mining sites is the main driving variable holding a median value of importance equal to 6.3, followed by K_{VH} with a value of 6.0 (Figure 34). These parameters show similarly higher values if compared to the remnants: In fact, together they account for a total 27% of overall importance in the model, thus remarking their breakthrough role for catching the statistical evidence for discriminating high-velocity MPs in different classes. It can be inferred that the first parameter can be safely correlated with the classification of true mining-related displacement, while K_{VH} may be relevant to distinguish prevailing vertical surface movements (subsidence) from phenomena responsible for mainly horizontal motions (landslide). Slope, aspect, elevation, and distance to IFFI landslides constitute a set of fairly relevant parameters with median importance values comprised between 4.7 and 4.3. Apart from TWI, the last set of parameters composed by profile curvature, total curvature, plan curvature and lithology accounts for a minor responsibility in the splitting process of a decision tree (Figure 34), with the local differences in soil/rock composition described by the lithological map being marginally useful to classify surface displacements (2.6 median importance).
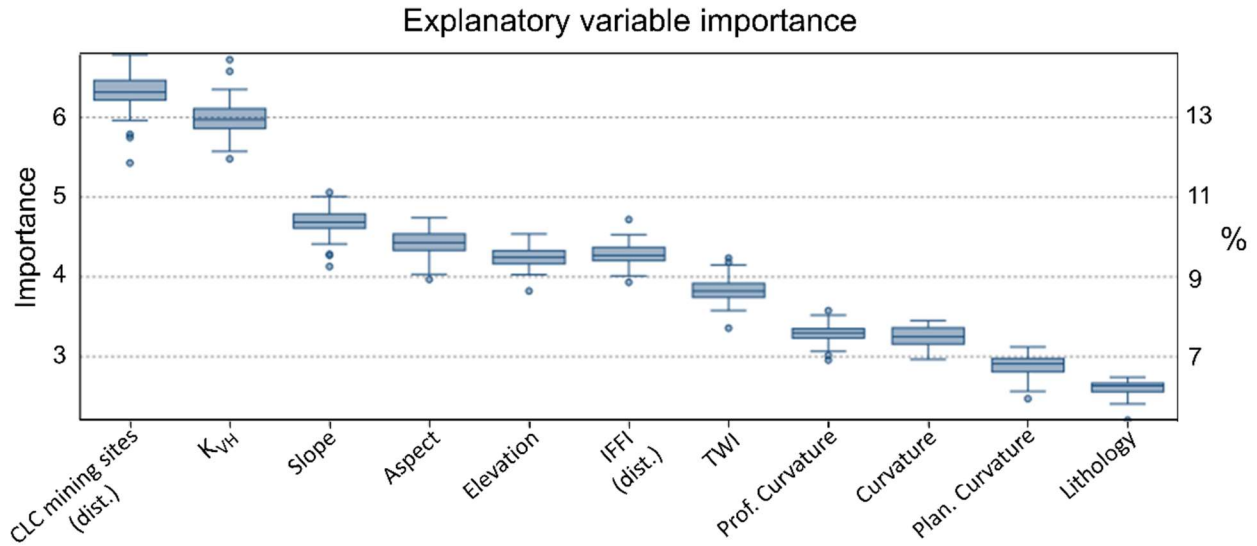
Explanatory variable importance

Figure 34. Boxplot related to the importance of every explanatory variable computed as the sum of the Gini coefficients from all the trees. The values are also reported as the percentage of the total sum of Gini coefficients.

The validation dataset mean confusion matrix indicated as the percentage of predicted samples with respect to the ground truth labelled samples, and a list of performance metrics achieved for every class are reported in Table 3. In general, classification indices are very good considering all classes (arithmetic mean of indices values), with an overall accuracy of 9.7e-1, a TPR of 9.4e-1 and a value of F1-score equal to 9.3e-1. Class M (mining-related displacement) has the highest indices among all the other classes, while classes L (landslide) and S (subsidence) show comparable high performances with both recording relatively smaller F1-score values (9.5e-1 and 9.2e-1, respectively). On the other hand, MPs are more likely to be incorrectly labelled as ND (Noise or other) as witnessed by the relatively lower performance rates achieved (Table 8). In particular, the percentage of actual positives ND which are correctly identified (i.e., the TPR) is equal to 8.1e-1. This can occur considering that class ND gathers together high-velocity MPs with multiple deformation sources, including noisy and sparse signals which could not be related to a discrete and clearly discernible phenomena. Given that, the lack of peculiar features makes their classification more difficult to the RF according to the parameters considered. This is confirmed also by the results reported in the confusion matrix (Table 8). Class ND is more often confused with class S and secondly with class L. Among the other classes, S is never confused with L and vice versa, thus demonstrating that the deployed parameters provide strong statistical attributes to discern subsidence from landslides. Mining-related displacements are confused in 2.4% of the cases with ND and only in 1.2% of the cases with both L and S.

Table 8. Validation dataset mean confusion matrix and performance metrics reported for every class.

	Validation dataset: classification			
Class	L	S	M	ND
Landslide (L)	95.7%	0.0%	1.2%	3.1%
Subsidence (S)	0.0%	95.5%	1.2%	3.3%
Mining-related disp. (M)	0.7%	0.0%	98.1%	1.2%
Noise or other (ND)	5.4%	12.2%	2.4%	80.0%

Diagnostic metrics	Per class				Overall
Accuracy	9.7e-1	9.6e-1	1.0	9.4e-1	9.7e-1
TPR (True Positive Rate)	9.5e-1	9.8e-1	1.0	8.1e-1	9.4e-1
F1-score	9.5e-1	9.2e-1	9.9e-1	8.7e-1	9.3e-1

6.2.3. Outcomes of prediction

After the training and validation phases, we used the best RF model to predict the source of displacements recorded through the A-DInSAR deformation map of Northern Italy. The classification concerned a total number of 18,028 high deformation MPs obtained from the training of 2,184 features expertly classified according to well-known cases of landslide, subsidence, and mining deformation phenomena. The geographical distribution of both outcomes and training data are depicted in Figure 35.
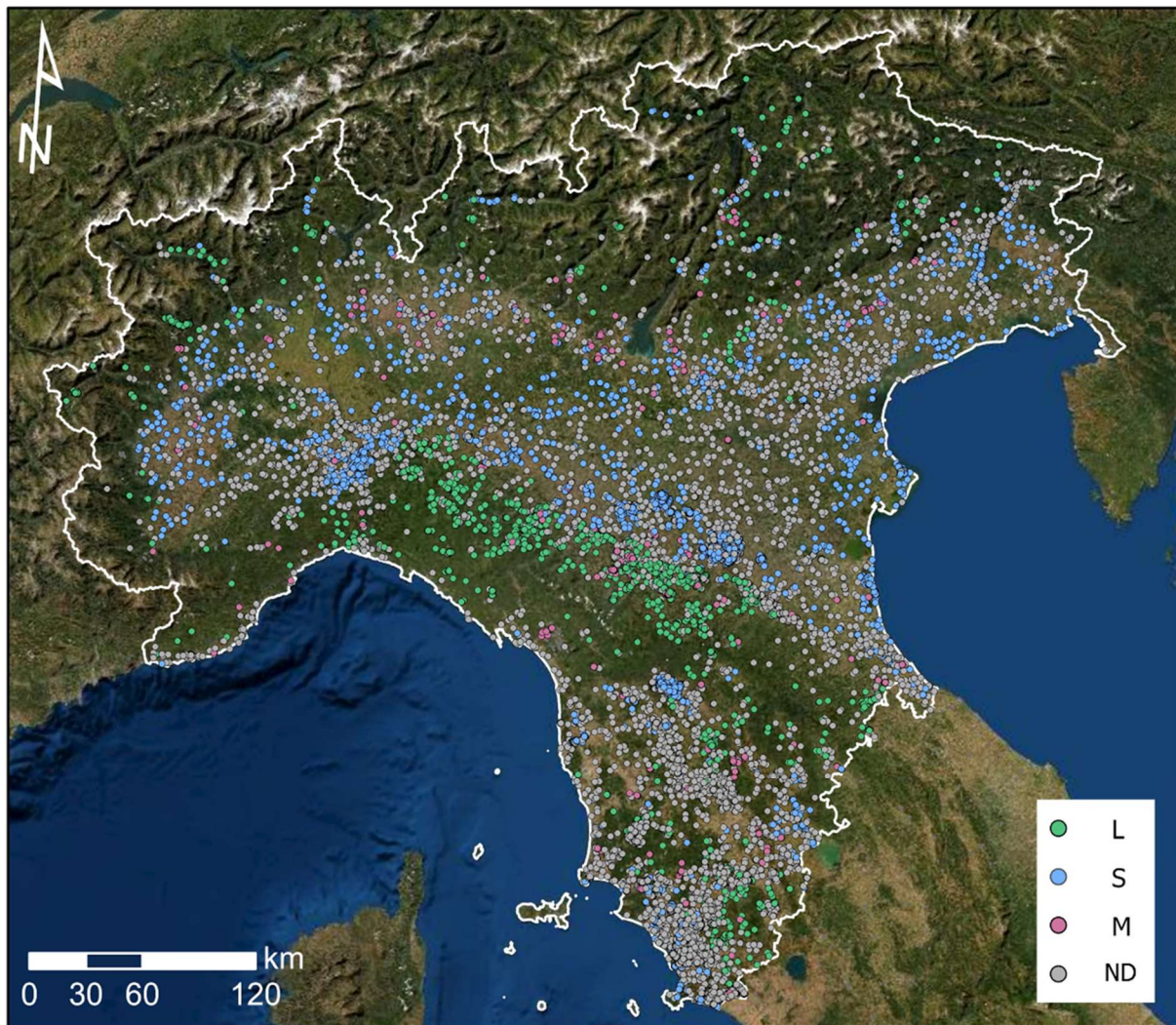

Figure 35. Northern Italy classification results of A-DInSAR MPs with high velocities according to 4 classes of deformation source: landslide, subsidence, mining-related displacement and noise or other phenomena.

The proposed method detected that L and S outlying MPs are settled in different morphometric and physic environments, reflecting the different geological processes. 2,712 MPs, which were labelled as L, are in correspondence of mountainous areas (mainly across the Northern Apennines), while 8,255 S points group themselves in the main plains (e.g., the Po Plain) and the main valley floors (e.g., alpine valleys) of the study area. We estimate a mean slope value of 16.1° for L features (with a standard deviation equal to ±9.4°), while for S-classified MPs we retrieve a mean slope of 0.5° and ±1.3° standard deviation thus confirming that slope constitutes an important parameter in our model. Another strong discriminant between L and S relates to elevation: we find mean values of 55.8±66.6 m a.s.l. for S points and 528.4±320.6 m a.s.l. for L points.

On the other hand, 773 MPs are labelled as mining-related displacements (M) while 8,472 are classified as noise or other phenomena (ND). Most of the M features (i.e., 81% of the totality) appears to cluster within the boundaries of the CLC mining sites, while the remnants are situated in direct proximity (ranging from 0 to 576 meters) since the main driving parameter of our model is the distance from the mining locations. The other morphometric parameters such as slope, elevation, aspect, and TWI do not show meaningful statistical distributions for M predicted features since mining sites are not spatially restricted to a particular physiographic environment. Concerning MPs labelled as ND, they hold, by definition, largely dispersed values among all the parameters.

The classification outcomes at local scale for several sites across the study area are shown in Figure 36. The method detects that the majority of outlying MPs in the metropolitan area of Bologna (Emilia Romagna Region) refer to a well-known subsidence phenomenon (Figure 36a) with a long-term evolution due to local anthropogenic activities (i.e., overexploitation of the aquifers) and natural causes, such as tectonics (Stramondo et al., 2007). In the same area, several MPs linked to mining-related deformations are classified in correspondence of 3 open-pit extraction sites of inert material which run stretches along the Reno River. In general, we find good agreement between the subjective assessment and the predicted classification, even if in few MPs in the buffer zone around mining areas appears misclassified and were marked as ND. This can be ascribed to a misleading combination of explanatory variable values that RF fail to interpret. On the other hand, in Figure 36b case near Oriano hamlet (Emilia Romagna Region), is depicted, where outlying signals within a small area are correctly correlated to landslide and to mining areas which were previously catalogued as such. Also, our method proves to be able to detect L classified MPs which do not fall inside the IFFI landslide inventory. As evidence of this, in Figure 36c we show the case of the Sauze d'Ouilx hamlet (Piedmont Region) which is recognized to be affected by a large deep-seated gravitational slope deformation (DGSD; Fioraso et al., 2020). In this context, the proposed classification method could provide an aid in updating geohazard inventories.

Another interesting capability shown by our model concerns the indirect classification of source of deformations which were not considered in the training dataset. The evidence is provided by the case of Larderello geothermal field, which is settled in southern Tuscany Region (Figure 36d). Almost every MPs is being correctly detected as ND (i.e., noise or other source of deformations) since the origin of the deformation is not related to any L, S or M phenomena. Therefore, here the combined statistical instances of the explanatory parameters have proven to be effective in not confusing ND with S, despite the area suffering from severe vertical movements which have persisted in the last decades (Botteghi et al., 2012).
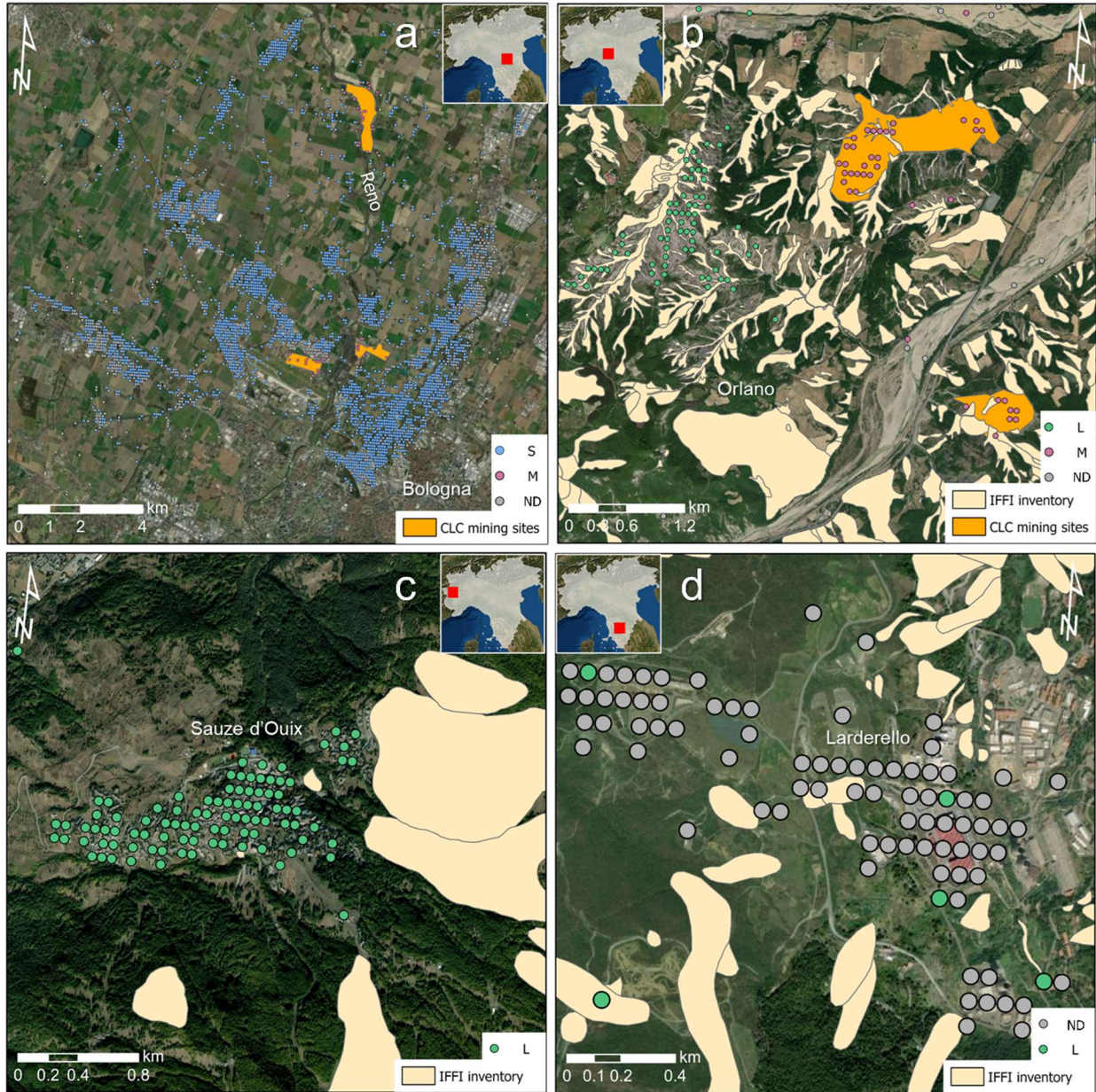
Figure 36. Classification results displayed for different sites across the study area. (a) Subsidence phenomena and extraction sites in Bologna city and surroundings (Emilia Romagna Region); (b) landslide phenomena and mining-related displacements in the Oriano hamlet (Emilia Romagna Region) in correspondence of IFFI and CLC inventories; (c) correctly interpreted landslide phenomena in Sauze d'Oulx (Piedmont Region) which has not been previously reported in the IFFI inventory; (d) surface deformations related to the geothermal activity in Larderello (Tuscany Region) where most of the outlying MPs have been correctly classified as noise or other sources of displacement.

6.2.4. Impact of K_{vh} variable within the RF model

It was also analysed the impact of the newly introduced KVH variable in the modelling procedure. Being K_{vh} a proxy for estimating the deformation pattern on the vertical and East-West directions, it was expect that it would constitute a solid explanatory index in the RF classification procedure. To this end, it was

performed the two-tailed Mann Whitney U test (Fay & Proschan, 2010) for checking a significant difference between the K_{vh} distributions of the predicted MPs (Figure 37). We tested the four classes non-normal distributions pairwise for a level of significance (α) equal to 0.05, conducting 6 tests for every combination. The U statistic with normal approximation is then used to compute the probability p-value and evaluated against the selected level of significance: if p is less than α, the null hypothesis of no correlation (H0: class1 = class2) is rejected and we accept that the difference between class1 and class2 populations is big enough to be statistically significant. Tests outcome confirms that each class distribution (Figure 37) substantially differs from the others since we obtain a p-value equal to 0.5e-2 when testing M predictions against ND predictions, while we obtain a p-value of zero for all the other test combinations. In the end, it is demonstrated that in the proposed methodology the use of K_{vh} is fundamental for discerning between different sources of deformations detected through A-DInSAR data.
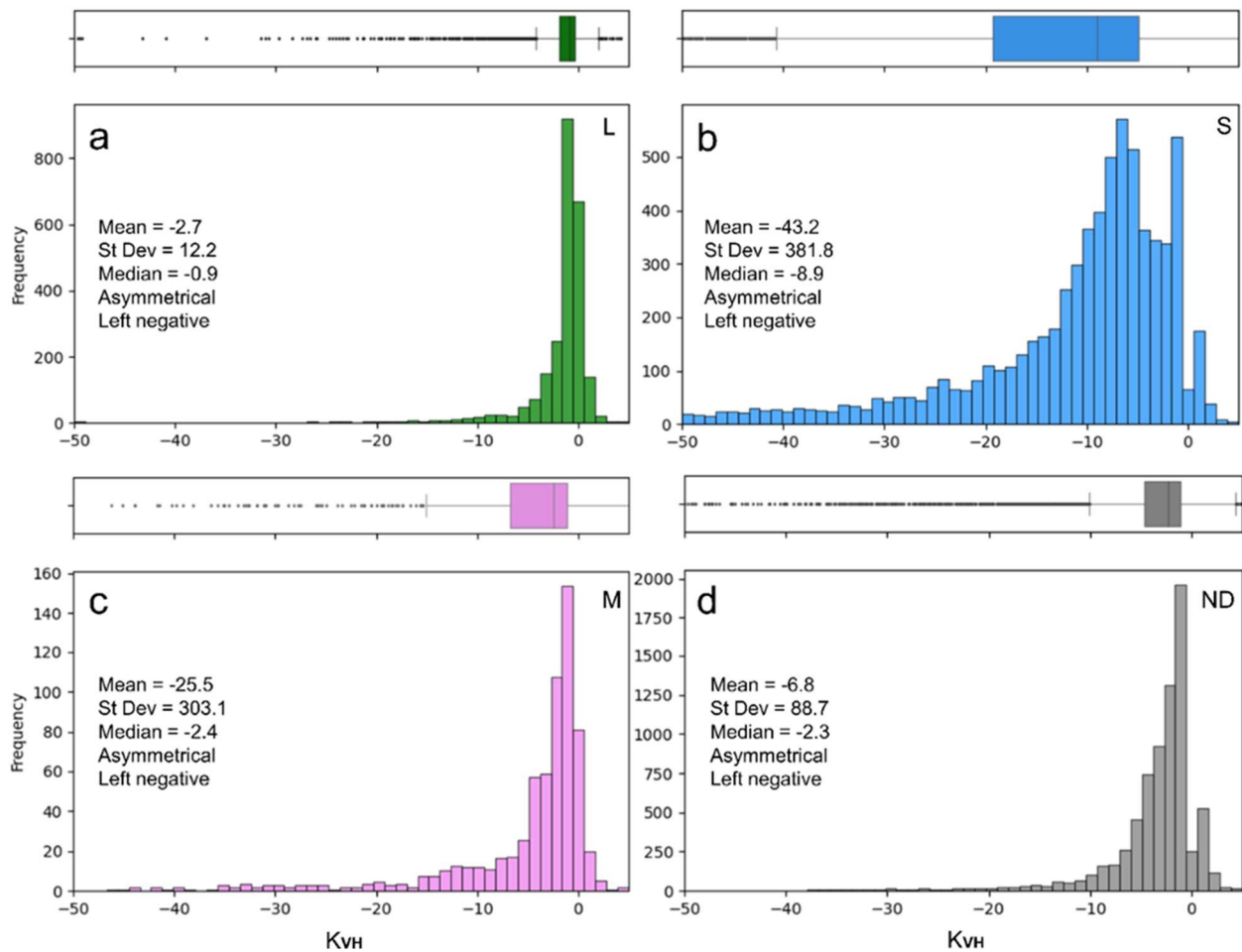
Figure 37. Histogram and plot box of the K_{VH} distributions related to predicted (a) landslide MPs, (b) subsidence MPs, (c) mining-related MPs and (d) noise or other MPs.

6.2.5. Spatial pattern analysis of the predicted features

The spatial distribution and aggregation of the high velocity MPs partly reflect the parent dataset distribution. Moreover, our data should be inherently grouped in space being this a constraint of A-DInSAR

technique (MPs usually aggregate around reflecting elements that satellites can easily see; Lu et al., 2012). For this reason, we evaluated the Average Nearest Neighbour (ANN) for every predicted class, and we compared them to the overall distribution. In our case, ANN is based on the averaging of all the distances found between each MP of the same class and its nearest neighbour. The ANN ratio is the observed average distance divided by the expected average distance of a hypothetical random distribution with an equal number of points covering the same total area. If the observed average distance is less than the expected for a random distribution, the MPs class being analysed is considered clustered, or conversely, it is considered dispersed. As can be seen in Table 9, all the classes exhibit ANN ratios largely greater than 1, denoting a trend toward dispersion with likelihood that the dispersed pattern could be result of random chance (z-scores largely greater than the critical value) less than 1%. This could be ascribed to the MPs density which has been greatly lowered during the post-processing operation (see section 6.1.1) while preserving the underlying surface area. On the other hand, if we compare the ratio between predicted classes ANN and the overall ANN related to whole outlying MPs we can retrieve interesting remarks: M, L and S classes appear to be more clustered in space than the overall distribution, with ANN ratios being around half of the overall ANN (Table 9). Instead, MPs belonging to ND class appear to be slightly more dispersed (value of ANN equal to the overall ANN multiplied by 1.1). As a result, we can deduce that the proposed method confirms to be helpful in detecting spatially aggregated features associated to discrete deforming phenomena, and besides, it proves reliable in identifying noisy deformation signals.

Table 9. Computed Average Nearest Neighbor ratio (ANN) for every predicted class and for the overall spatial distribution of outlying MPs, where also every Class ANN is evaluated against the overall ANN.

	Observed mean dist. (m)	Expected mean dist. (m)	ANN ratio	Class ANN/Overall ANN
Overall	512.1	1.3	394.0	-
L	752.2	3.6	209.0	0.5
S	491.5	2.1	234.0	0.6
M	981.0	6.9	142.2	0.4
ND	937.8	2.1	446.6	1.1

7. Automated clustering of Time Series of Surface Deformation

In this section of this Thesis work it will be presented the third and last approach aimed at developing an automated pipeline for the extraction of the most meaningful and recurrent trends of deformation described by the temporal evolution of displacement values of every MPs constituting an A-DInSAR dataset. The final goal of the proposed approach is the effective management and automated post-processing of large EO data to make ground motion data readily understandable and interpretable by a broad range of nonexpert stakeholders and end-users.

Any supervised manual classification of large A-DInSAR datasets can no longer be performed, being a time-consuming activity and the results affected by subjectivity. For systematic large-scale mapping, the large volume of Time Series Interferometric Synthetic Aperture Radar (TS-InSAR) information enclosed within millions of Measurement Points (MPs) calls for efficient DM approaches to disclose the underlying displacement patterns. Therefore, here is explored the applicability of an automated unsupervised machine learning approach by applying a statistical dimensionality-reduction method which includes a Principal Component Analysis (PCA; Jolliffe & Cadima, 2016) followed by K-means clustering (Wu et al., 2008) in order to detect non-stationary deforming trends at a regional scale.

The test site is the Valle d'Aosta region (VdA) which is among the regions with the highest mortality rate from landslides in Italy (polaris.irpi.cnr.it) and where other identified sources of ground motion include subsidence and mining-related displacements (Confuorto et al., 2021; Solari et al., 2019). The deployed A-DInSAR dataset is processed by means of SqueeSAR™ algorithm (Ferretti et al., 2011) and for the proof of concept of the proposed methodology only TS-InSAR data are deployed.

The novelties brought by this work, are the following:

- The proposed DM approach is applied to the decomposed vertical and eastward components displacement TS-InSAR rather than to the LOS values.

- The robustness of the clustering method is evaluated against different TS-InSAR interpolation techniques.

- Every retrieved cluster is further decomposed into trend and seasonality components for a complete mapping of the temporal displacement behaviour.

To the best of our knowledge, this method has not been tested on PSI TS-InSAR before.
The proposed work addresses the following literature gaps:

- most of studies aiming to map and characterise ground motion mainly rely on the average displacement rate obtained from fitting a linear regression model to the entire TS-InSAR data (Festa et al., 2022; Lu et al., 2012; Novellino et al., 2019; Tomás et al., 2019), rather than deploying the whole temporal dimension.

- Mapping TS-InSAR temporal displacement signals to a priori pattern models (Berti et al. 2013; Chang & Hanssen, 2016) would inhibit the scalability in the classification of unknown temporal patterns in large InSAR datasets.

7.1. Site characterization and interferometric data used
7.1.1. Study area

VdA is a region located in the north-western part of the Italian Alpine arc and is bordered by Switzerland to the north and by France to the west (Figure 38). The 3,262 km^2 area encompasses a mostly mountainous landscape where almost 50% of its territory has an elevation higher than 2,000 m.a.s.l.

As a result of the Alpine orogenesis and the Quaternary glacial modelling, VdA constitutes a complex structural-geomorphological context (Dal Piaz et al., 2001) where a main central valley east-west oriented is framed by several sub-valleys north-south oriented, constituting an ideal scenario for the application of the InSAR technique (i.e., ground deformation detection along N-S direction it is poorly constrained by the satellite geometry). The post-glacial action (Holocene to present) has been influencing the reliefs setting and the ongoing slope dynamics (Carraro & Giardino, 2004). Indeed, VdA is extensively characterized by gravitational mass movements such as shallow landslides, rockfalls, rock glaciers and deep-seated gravitational slope deformation (DSGSD) which sometimes involve entire mountain flanks (Trigila et al., 2010). The study of such phenomena represents a primary necessity for a mountain territory, and it is functional to more effective risk management for such a region.

Considering the historical causes of motion in VdA, the validation of our procedure included:

- Landslide inventory from the Italian Landslide Inventory project (IFFI; Trigila et al., 2007), where 2,188 landslides and 141 DSGSDs have been mapped until 2016 throughout the VdA territory (Figure 38), covering more than 590 km^2 (ca. 18% of the entire region).

- Glaciers database on more than 900 rock glaciers marked as active or relict landforms (catastoghiacciai.it/GeoGhiacciai). Besides the well know geohazards, other sources of land motion have been linked to human-induced surface displacements, namely subsidence due to groundwater overexploitation and mining-related activities (Confuorto et al., 2021).

- Finally, a database on the localization of possibly moving areas subject to past and recent quarrying and dump activities (Figure 38) available online at geoportale.regione.vda.it.
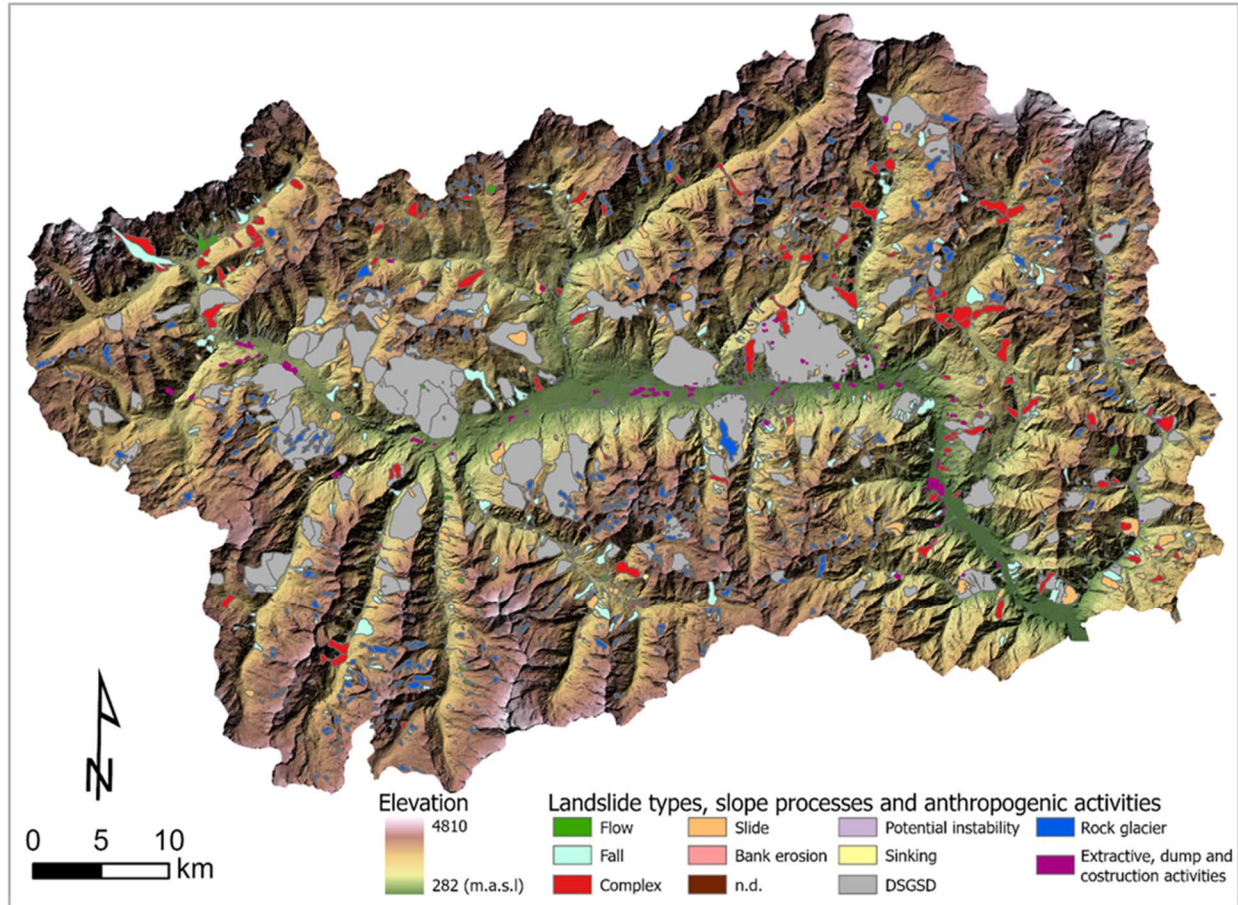
Figure 38. Geomorphologic setting of the VdA, with landslide types, slope processes and man-induced moving areas outlined according to the IFFI catalog and to the regional databases available at /geoportale.regione.vda.it/.

7.1.2. Interferometric dataset covering VdA

The launch of the C-band S1 constellation has allowed to acquire huge volumes of freely accessible radar images with an unprecedented temporal sampling over Europe. In this study, 256 and 261 S1 scenes covering VdA were exploited from ascending and descending acquisition geometries Figure 39a, respectively. The stacks of images, acquired with the Interferometric Wide (IW) swath mode by both S1A and S1B and spanning the time interval from 19 October 2014 to 12 May 2020, were processed by means of the SqueeSAR™ algorithm (Ferretti et al., 2011) which allows a high density of interferometric data, especially in environments such as the alpine area (Notti et al., 2012). Images have been captured along satellite track n.88 in ascending orbit, with an incidence angle of 38.59°, and along-track n.66 in descending orbit, with an incidence angle of 43.12° (Figure 39b). The SqueeSAR processing chain was able to provide more than 730,000 ascending and descending displacement TS-InSAR subdivided between point-wise coherent scatterers (PS) and partially coherent distributed scatterers (DS), allowing a rather high MP density (ca. 56 MP/km^2).

b	**S1 parameters and PSI results**	
SAR satellite	S1A&B	S1A&B
Satellite geometry	Ascending	Descending
Track	88	66
Sensor mode	IW	IW
LOS angle	38.59°	43.12°
Azimuth angle	8.87°	10.49°
No. of scenes	256	261
No. of MPs	365,384	367,263
Time spans	04-11-2014 to 12-05-2020	10-10-2014 to 11-05-2020
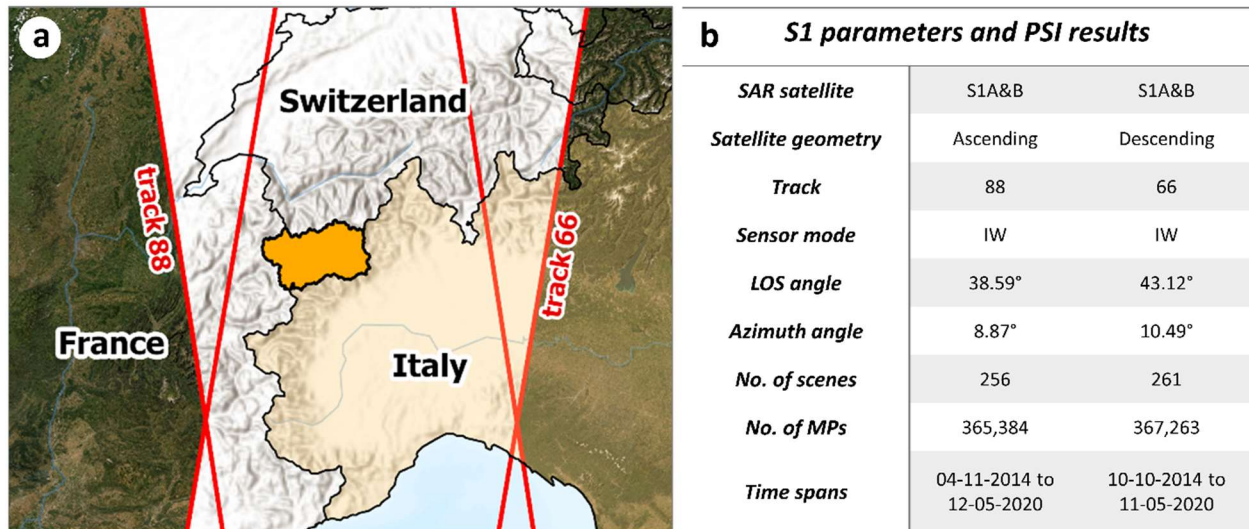
Figure 39. (a) S1 datasets location and (b) PSI SqueeSAR™ results.

7.2. Rationale of the proposed methodology

In this section, it is briefly described the pipeline of the proposed TS-InSAR DM approach (Figure 40) which is entirely performed automatically through a code assembled in Python 3.8.10 (https://github.com/maybedave/InSAR-Time-Series-Clustering). The code consists of three main steps: 1) Spatial and temporal post-processing of the PSI dataset to retrieve newly interpolated vertical and eastward displacement TS-InSAR; 2) PCA-based dimensionality reduction and features retrieval; 3) unsupervised K-Means learning for TS-InSAR automated clustering and decomposition of the cluster centroids.

Figure 40. Workflow of the proposed approach.

7.2.1. TS-InSAR post-processing

Multi-geometry InSAR data fusion aims to reproject and estimate the LOS observed displacement signal along the vertical and eastward directions based on the combination of measurements acquired from two or more different SAR satellite orbit acquisitions. Independent InSAR datasets can be merged on the (i) spatial and (ii) temporal level if LOS measurements are referred to the same location and period.
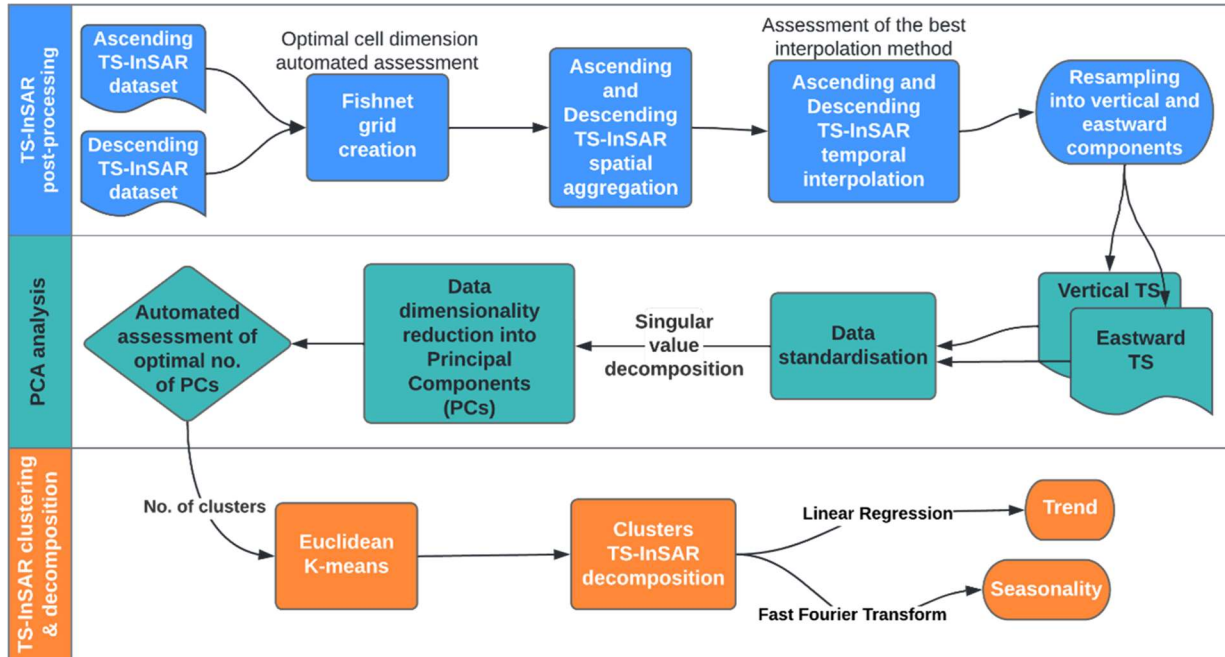
The accomplishment of the (i) spatial interpolation is performed by merging sparse MPs from overlapping geometries into spatially coherent domains. We propose a squared gridding of point-like data consisting of an adaptive scaling process able to enclose at least one MP from ascending and descending geometries. The process starts by assessing the extent of the InSAR datasets to produce each individual grid cell as a geometric polygon. The optimal size of the grid cell is automatically assessed as follows: considering an interval $n = [10 \; .. \; 1000]$ the totality of cells C_n produced at the n grid size is evaluated against $Cad_n = Ca_n \cup Cd_n$, where Ca_n and Cd_n are the number of cells with at least one MP from the ascending and descending geometry, respectively. The optimal size of the grid cell corresponds to the smallest grid size (so to capture motion at higher resolution using more cells – C_n) but large enough to contain at least one ascending and descending MP. This size is evaluated by plotting the results of C_n/Cad_n against the n range of size values and by identifying the point of maximum curvature (Figure *41*) through the Kneed algorithm (Satopaa et al., 2011). Empty grid cells resulting from the absence of simultaneous data from both geometries are automatically deleted.
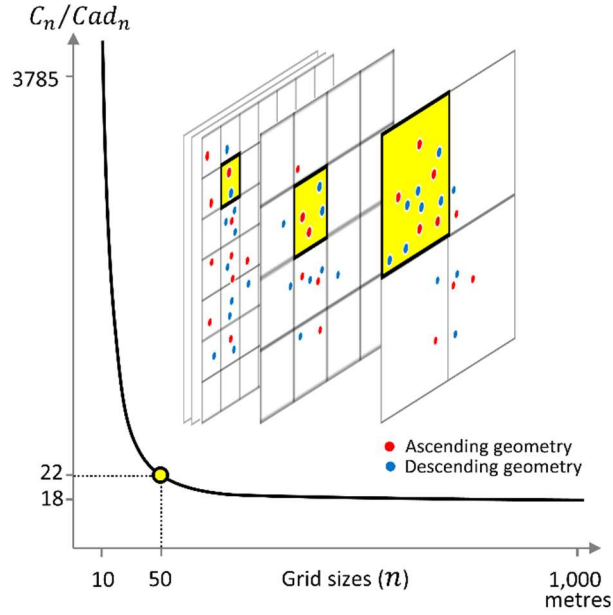
Figure 41. Procedure for the identification of the optimal size of the grid for the spatial interpolation of MPs from the two LOS satellite geometries of acquisition.

Also, (ii) interpolation in time is required to retrieve vertical and eastward TS-InSAR since the image acquisition dates are usually different in each viewing geometry. In this framework, the proposed methodology performs the fusion of geometry-wise LOS velocities on a cell-by-cell basis on the interpolated grid, thus resulting in unique ascending and descending TS-InSAR for each grid squared polygon. After aggregating and averaging the displacement values per time step, the approach pipeline follows with the selection of the ascending and descending overlapping time window (oTw) as the new temporal frame for TS-InSAR resampling. Then, the new frequency conversion nf (15) is automatically computed as:

$$nf = oTw/(S_{asc} + S_{desc})$$

(15)

where oTw unit is days, while S_{asc} and S_{desc} are the total numbers of ascending and descending geometry scenes. The interpolation of LOS velocities (V_{asc}; V_{desc}) to a new frequency enables the computation of vertical (V_v) and eastward (V_E) (16) TS-InSAR on each time step by solving the following set of equations:

$$\begin{pmatrix} V_{asc} \\ V_{desc} \end{pmatrix} = \begin{pmatrix} -\sin\theta_{asc}\cos\alpha_{asc} & \cos\theta_{asc} \\ -\sin\theta_{desc}\cos\alpha_{desc} & \cos\theta_{desc} \end{pmatrix} \begin{pmatrix} V_E \\ V_v \end{pmatrix}$$

(16)

where θ and α are the satellite incidence and azimuth angle, respectively.

The vertical and horizontal components of motion are here retrieved assuming the N component of motion negligible. The TS-InSAR temporal interpolation is performed by testing linear, quadratic and cubic methods in order to find out the most accurate model able to fill data gaps in the new frequency. By trying to minimize the difference between observed data and interpolated value, the aim is to preserve as much as possible of the original non-stationary TS-InSAR data. To achieve that, we randomly select 100 time

series from the dataset, eliminate every 2nd displacement measurement, interpolate the missing values, and evaluate the Root Mean Square Error (RMSE) (17) between the observed and forecasted values:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - o_i)^2}$$

(17)

where f is the forecasted value and o is the observed value and n is the total sample size. RMSE can be defined as the standard deviation of the residuals, therefore lower values are associated to a better average model prediction. The lower mean RMSE value retrieved for 100 time series interpolated via linear, quadratic, and cubic methods is here considered as a proxy for choosing an individual fitting model.

7.2.2. Principal Component Analysis (PCA) applied to TS-InSAR

Given the large quantity of temporal displacement data embedded within TS-InSAR datasets, here we use PCA, a statistical technique which consists of data dimensionality reduction by balancing between information loss and optimal number of dimensions to retain (Lischeid, 2009; Prentice, 1982), in the framework of interpretability enhancement of the dataset (Jolliffe & Cadima, 2016a). The main principle behind the data dimensionality reduction is to preserve as much statistical variability as possible. This relates to finding new variables (linear functions of those in the original dataset), namely, the principal components (PCs), which explain a maximal amount of variance and that are totally uncorrelated. PCs reduce to solving an eigenvector/eigenvalue problem, where the eigenvector and the associated eigenvalue constitute the variance accounted by each of the components. Retrieving the PCs aims, in our case, at characterizing temporally and spatially changing deformation patterns from any TS-InSAR dataset without a priori constraints (Chaussard & Farr, 2019). Therefore, prior assuming that the sources are uncorrelated, the decomposition of a mixed deformation signal through the PCA results in a unique solution (Ebmeier, 2016).

The here performed PCA is applied to multiple univariate time series (i.e., the TS-InSAR dataset) and can be broken down into (i) data standardization, (ii) data dimensionality reduction into PCs, (iii) automated assessment of the optimal number of PCs to retain. The aim of (i) is to standardize all the TS-InSAR to a comparable scale by removing the mean and scaling to unit variance, so that each feature is equally considered and that PCA will not result in biased estimations. The second step (ii) involves linear dimensionality reduction using Singular Value Decomposition (Halko et al., 2011) of the data to project it to a lower dimensional space, namely orthogonal components that explain a maximum amount of the variance. To achieve (iii), we adopt the Scree-plot criterion, where the eigenvalues are plotted against the number of components, and the optimal number of PCs is retrieved by looking for the "elbow" in the curve (see section 7.2.1) thus selecting all the meaningful components just before the curve flattens. This is done to automatically select the optimal number of PCs able to explain much of the variability of the TS-InSAR dataset.

7.2.3. TS-InSAR clustering and decomposition

The K-means algorithm is an inductive and unsupervised clustering method that is here adopted in the form of the standard Euclidean version to cluster unlabelled sparse TS-InSAR data into groups with similar characteristics. Given two time series, $Q = q_1 \dots q_n$ and $C = c_1 \dots c_n$, their Euclidean distance D (18) is defined as:

$$D(Q, C) = \sqrt{\sum_{i=1}^{n} (q_i - c_i)^2}$$

(18)

K-means aims to separate n samples x in k clusters (c) of equal variances, each described by the mean μ_i (i.e., the centroid) of the TS samples in the cluster. The final clustering algorithm objective is to perform the minimisation of centroids inertia (19):

$$\sum_{i=0 \,\mu_i \in c}^{n} \min(||x_i - \mu_i||^2)$$

(19)

Therefore, inertia can be defined as a metric expressing how internally coherent clusters are, where the lower the values the better the model convergence.

K-means is a scalable approach towards very large samples and needs only a parameter to be initialized, namely the number of clusters, which is in our case automatically defined by the optimal number of PCs. Running PCA before k-means clustering has multiple advantages: it is the continuous solution of the cluster membership indicators in K-means clustering (Ding & He, 2004), it can alleviate the problem of inflation of Euclidean spaces and speed up the computations.

Every MP is grouped into a cluster of similar patterns. These patterns are retrieved by considering the barycentre (i.e., centroid) of the average shape of the grouped time series. To enhance the readability of the TS-InSAR clusters, the cluster centroid is further analysed to decompose the signal into trend and seasonality, where the first can be defined as the general direction of the overall data, while the seasonality is the repeating short-term cycle component in the series. Finally:

- The least squares Linear Regression technique is applied to fit a linear model by minimizing the residual sum of squares between the observed targets in the time series and the targets predicted by the linear approximation in order to retrieve the trend.

- The basic Fast Fourier Transform (FFT) is here applied within our code to compute the power spectrum (Priestley, 1981) in order to retrieve the seasonality. By looking at the frequencies corresponding to the peaks of spectral power with the highest magnitude, it is possible to assess the periodicity of a time-related series.

7.3. Outcomes of the TS-InSAR analysis procedure

7.3.1. Parameters calibration

The results of the code looping through the extension of the ascending and descending TS-InSAR datasets covering the VdA generated different fishnet configurations based on a range of selected grid sizes. The best trade-off grid size is automatically found at a value of 48 m, corresponding to a totality of 2,480,665 cells subdividing a rectangular area of ca. 80 x 40 km enclosing the study area; on the other hand, Cad, namely the number of cells retaining both ascending and descending TS-InSAR, results in 93,385 cells, which constitutes the final number of features subject to the proposed DM procedure.

Due to the inherent temporal mismatch between the ascending and descending TS-InSAR datasets (256 and 261 scenes respectively), the proposed method resamples the LOS measurements from the two different geometries into a new shared temporal frequency (nf), thus enabling the LOS reprojection into the vertical and eastward components. The nf, computed as shown in section 7.2.1, is equal to 4 days considering the deployed test datasets, bringing the number of synthetic scenes to 504. Interpolation methods were tested by evaluating the prediction performance obtained with 100 random time series interpolated with linear, quadratic and cubic methods. In our case (Figure 42) the linear interpolation method is preferred since it carries a mean RMSE value of 3.32, which is lower compared to the mean RMSE obtained by both quadratic and cubic interpolations (3.76 and 3.83 respectively). This result is indicative of the deployed data used within this work and it may differ based on different TS-InSAR datasets used as inputs.
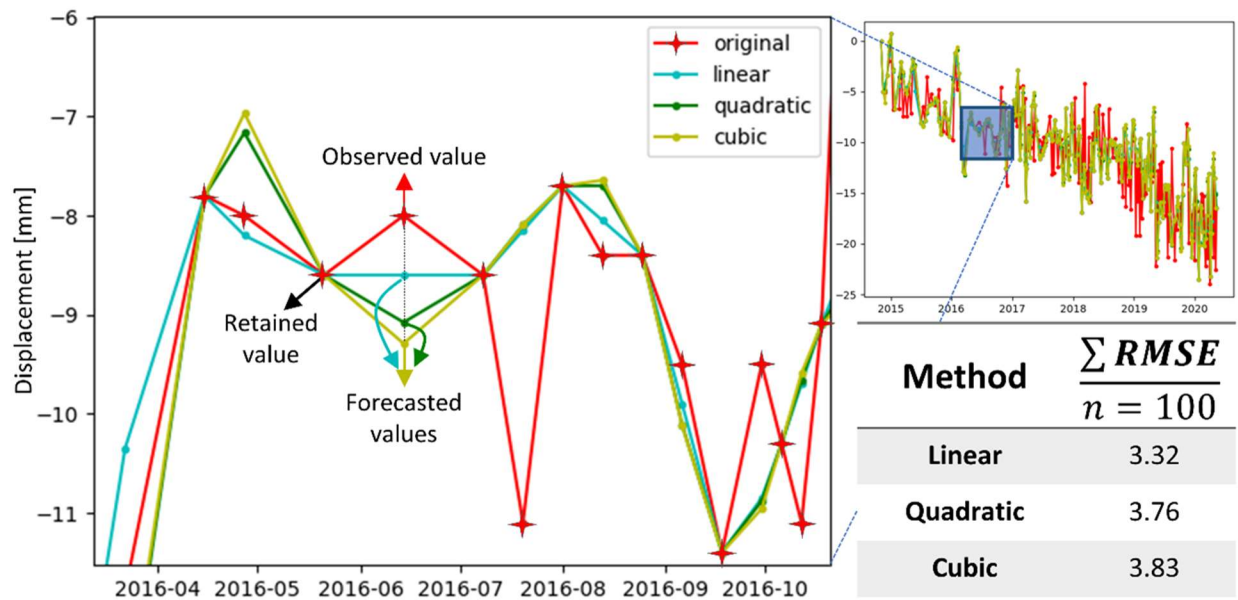
Figure 42. Adopted procedure for the assessment of the interpolation method which best approximates the observed values, in the framework of TS-InSAR frequency up-sampling. The original time series (i.e., from the SqueeSAR dataset) is reprojected into synthetic time series via linear, quadratic and cubic interpolation methods. We evaluate the mean RMSE of every method by interpolating 100 random time series from the ascending and descending datasets.

The optimal number of PCs to retain is automatically selected through the Scree plot criterion, by plotting the eigenvalues on the y-axis and the number of factors on the x-axis. It always displays a downward curve, since the first component usually explains much of the variability, while the next components gradually explain lower amount fractions of the overall variability. Finding the "elbow" in that curve relates to the number of factors retaining much of the TS-InSAR dataset information since they will constitute the cluster membership indicators in the K-means clustering. In our case, after performing the PCA analysis over the vertical and eastward TS-InSAR datasets, the Kneedle algorithm selected a number of 4 components to retain.

7.3.2.   Automated Data Mining of TS-InSAR data
7.3.2.1.   PCA-based K-means clustering

The standardized PCA scores for the reprocessed vertical and eastward TS-InSAR dataset (Figure 43) depict the underlying data structure. The two successive sets of the selected 4 orthogonal components explaining the maximum amount of the two datasets variance exhibit similar characteristics. The eigenvectors of the first PC (PC1) show that the direction of maximum variance (i.e., the eigenvector) is stretched parallel to the x-axis direction (representing the satellites acquisitions) showing a mostly stable and steady behaviour with displacement values slightly fluctuating around zero. PC1 demonstrates that most of the TS-InSAR signals covering the VdA (explained amount of variance stands at 76.55% and 74.14% for the eastward and vertical dataset, respectively) are related to stable targets, which corresponds to an expected scenario when looking at a regional PSI dataset. PC2, PC3 and PC4 show a strong periodic fluctuation of the displacement signals and, in particular, the observed ground motion phenomena tend to follow a clear yearly seasonal cycle. Regarding the eastward dataset, PC2 accounts for 8.22% of the total variance and displays a seasonal deformation with a positive trend (i.e., the relative movement towards east). Concerning the vertical dataset, PC2 accounts for 7.91% of the total variance while displaying an overall negative trend, which translates into a downward displacement. PC3 and PC4 are indicative of displacement signals with similar amplitudes and similar trends oscillating around zero, but also, they show to be out of phase over time in both analysed TS-InSAR datasets.
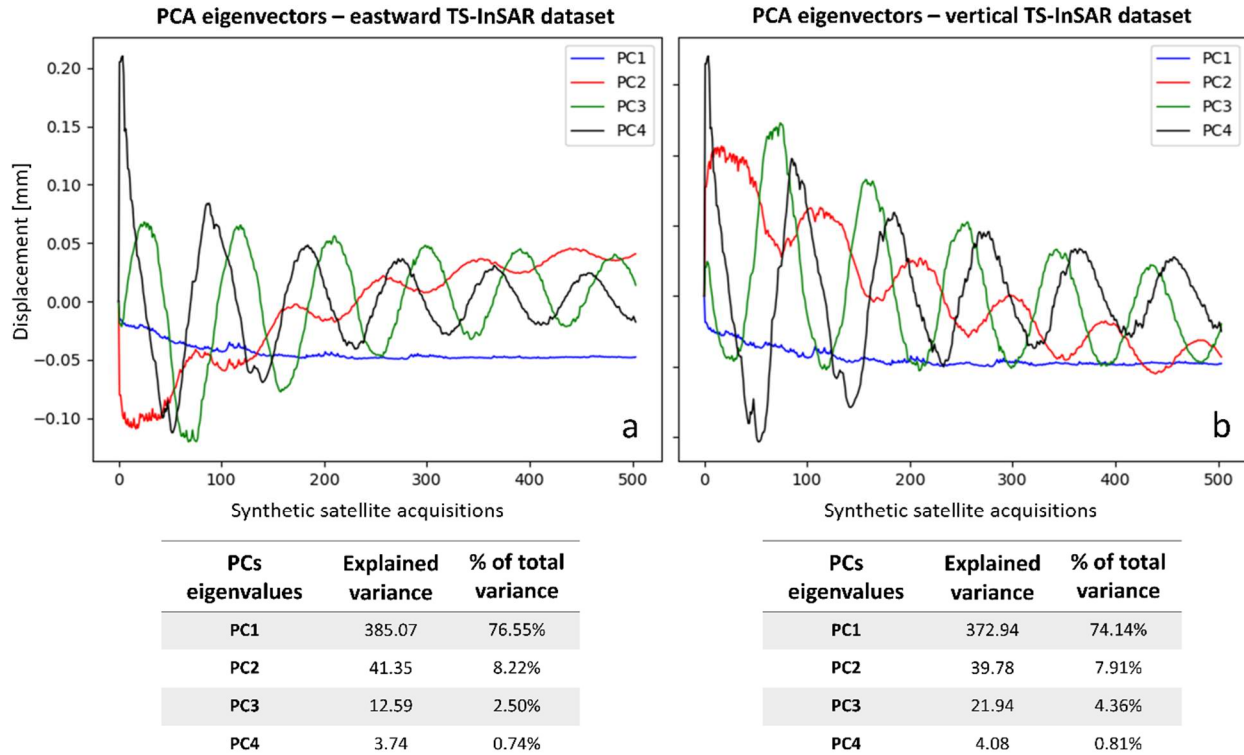
Figure 43. Plots of the eigenvectors and corresponding eigenvalues scores resulting from the standardized PCA conducted for the reprocessed a) eastward and b) vertical TS-InSAR dataset.
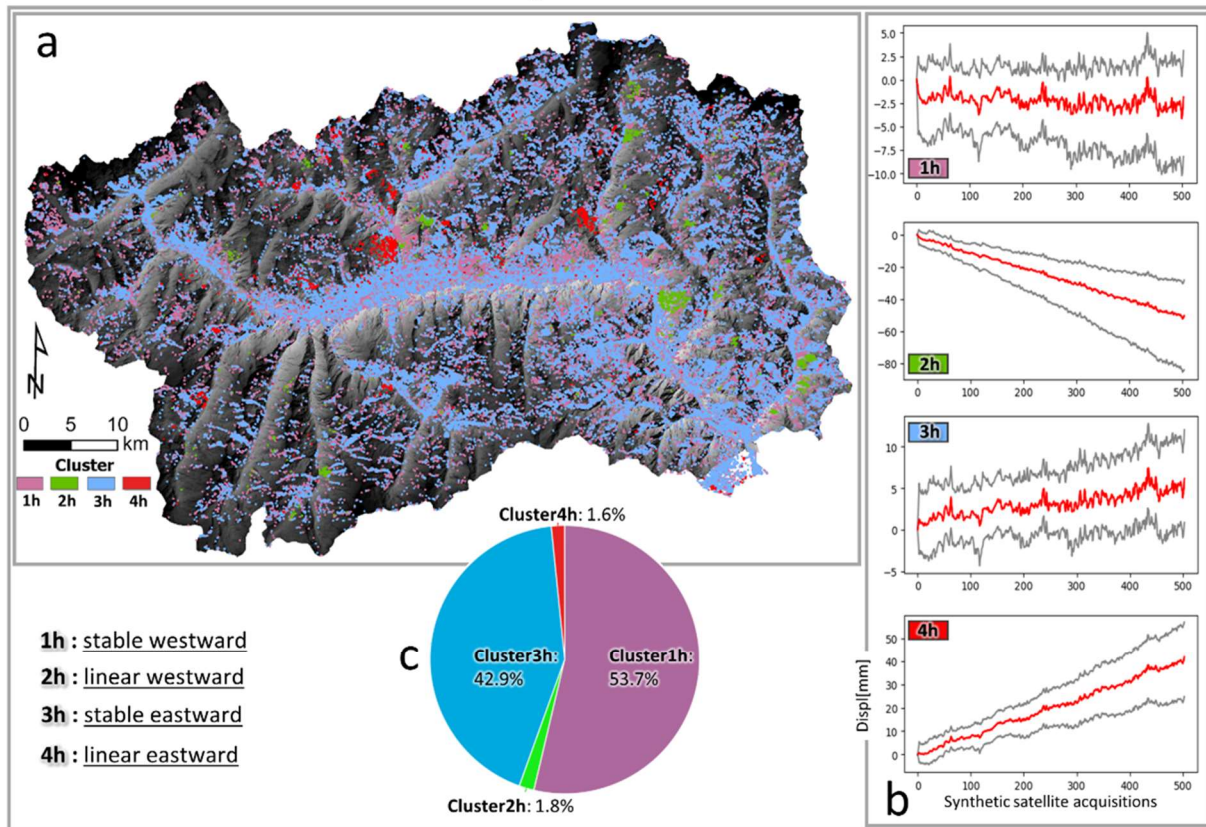
Accounting for the variance, the K-means clustering resulted in the classification of the analysed TS-InSAR dataset into 4 different types of clusters. The spatial distribution of the classified features, along with the representation of the centroid series (within the 10th and 90th percentiles) of each (i) eastward and (ii) vertical TS-InSAR cluster and the relative percentage distribution are shown in Figure 44.

Concerning (i) clusters, around 54% of the classified features belong to Cluster1e (Figure 44b), which shows an overall uncorrelated behaviour, with random fluctuations of displacements around zero indicating stable conditions and no prevailing movements towards east or west. On the other hand, Cluster3e represents around 43% of features, denoting again a noisy series of displacement which however show a general slightly positive trend, indicative of a relative eastward movement. Cluster2e and Cluster4e (only 1.8% and 1.6% of the overall features, respectively) can be assimilated if looking at the linear pattern and at the high rate of displacement (Figure 44b), even though they show opposite motion directions: while Cluster2e exhibit a decisive westward motion (negative trend), Cluster4e exhibit an eastward motion (positive trend). The geographical distribution of both Cluster1e and Cluster3e appears uncorrelated, while Cluster 2e and Cluster 4e tends to be spatially aggregated on slopes with specific aspect, hence denoting a common process influencing the temporal behaviour of displacement of the ground, such as an active landslide.

By looking at (ii) clustering results, 50.5% and 39.8% of features are classified as Cluster1v and Cluster3v, respectively (Figure 44f). These two show several similarities with the previously exposed Cluster1e and Cluster3e, both in terms of series shape (Figure 44e) and geographical distribution (Figure 44d): they show mostly uncorrelated and noisy behaviour, with almost null displacement rates (within ±5mm), while they

tend to be spatially scattered throughout the test area. On the other hand, Cluster2v and Cluster 3v are both less frequent within the classified features (0.8% and 8.9%, respectively), they exhibit a decisive negative trend (with Cluster2v holding higher rates of displacement) and tend to be mostly present on slopes. What is noticeable about the latter, is the repeating seasonal cycle markedly affecting the series. Trend and seasonality component analysis of the retrieved cluster series are addressed more specifically in the next section (7.3.2.2).

K-means clustering horizontal TS-InSAR dataset

a

Cluster
1h 2h 3h 4h

1h : stable westward
2h : linear westward
3h : stable eastward
4h : linear eastward

c

Cluster4h: 1.6%
Cluster3h: 42.9%
Cluster1h: 53.7%
Cluster2h: 1.8%

1h
2h
3h
4h
Displ[mm]
Synthetic satellite acquisitions
b

K-means clustering vertical TS-InSAR dataset

d

Cluster
1v 2v 3v 4v

1v : stable downward
2v : subsiding primary
3v : stable upward
4v : subsiding secondary

f

Cluster4v: 8.9%
Cluster3v: 39.8%
Cluster1v: 50.5%
Cluster2v: 0.8%

1v
2v
3v
4v
Displ[mm]
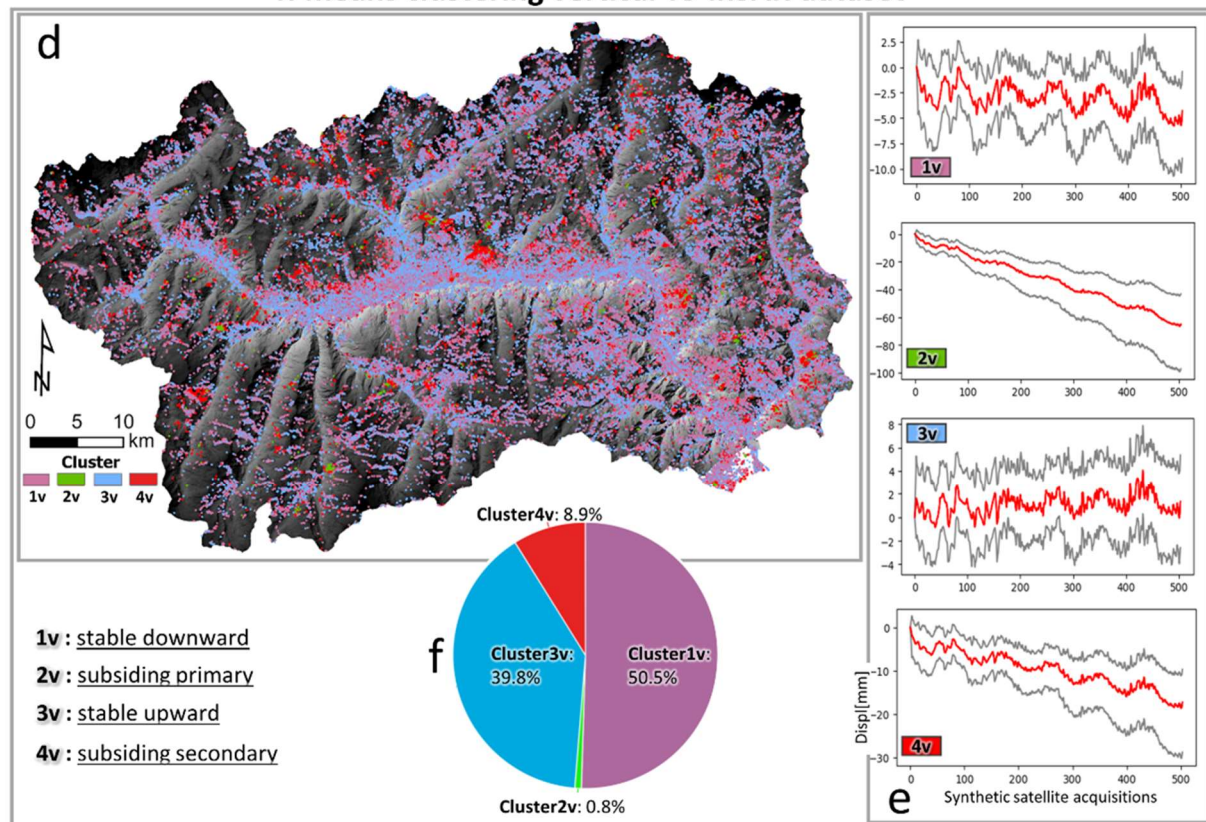Synthetic satellite acquisitions
e

Figure 44. Map of the geographical distribution of the a) eastward and d) vertical features classified according to the K-mean clustering analysis, where the relative (b, e) clusters series are depicted by plotting the 10th, 90th percentile (grey series) and the centroid (red series). The cluster percentage distributions related to the eastward and vertical TS-InSAR datasets are shown as (c, f) pie charts.

7.3.2.2.   Cluster series decomposition

The barycentre of the cluster's series is the subject of decomposition analysis, which constitutes the terminal step of the pipeline automatically operated through the deployed code. Each centroid is fitted to a linear model to better evaluate the underlying trend that causes increasing or decreasing patterns in the cluster series, while the potential seasonal component is obtained by transforming the centroid series into the frequency domain. The outcomes of the twofold analysis are presented as follows: the visual plots of the eastward and vertical centroid decomposition are depicted in Figure 45 and Figure 46 respectively, while the components are addressed quantitatively in Table 10. The spectral peaks obtained through the FFT are shown by plotting the x-axis covering the frequencies from zero to the Nyquist frequency (half the sampling rate), while the y-axis indicates the spectral power. Peaks corresponding to frequencies holding the largest power (amplitude) are shown in the tabular form, while with P_0 we indicate the peak with the lowest frequency (f_0) and with P_1 the peak carrying a greater frequency (f_1). The existence of an annual periodicity (i.e., seasonality) is eventually witnessed by the main peak approximating a frequency of $f \sim 365$ days^{-1}, while a peak associated to lower frequencies typically corresponds to the fundamental frequency (no predominant seasonality).

Among the eastward TS-InSAR clusters, the centroid marked as 1e (Figure 45) is the one depicting an overall stable trend (with the linear segment with the best fit holding a slope of approximately zero). 3e and 4e carry a similarly positive trend, however, 4e has a higher slope (i.e., higher deformation rate). Instead, 2e carries a negative linear trend but with a slope magnitude comprised between the latter. On the other hand, the periodicity analysis highlighted that only 1e and 3e have a small trace of seasonality within their signals, since they both resulted in having P_1 corresponding to a frequency of 336 days^{-1}, however P_0 recording lower frequencies.

Figure 45. Eastward TS-InSAR cluster centroid decomposition into a trend (Linear Regression) and seasonality (power spectrum) components.

Regarding the vertical TS-InSAR clusters, centroids 1v and 3v depict trend with very low deformation rates (Figure 46), with the first being slightly negative and the second being slightly positive (see slope values in Table 10), while 2v holds a decisive negative linear trend, with slope value of two orders greater than 1v and 3v. Centroid 4v has still a negative trend but less pronounced than 2v. The periodicity analysis of the vertical clusters' centroids allowed to retrieve interesting results: 1v and 3v centroids have their main spectral peaks set to an annual cycle (P_0 corresponding to a frequency of 336 days^{-1}). 2v shows no seasonality at all (both P_0 and P_1 relate to lower frequencies), instead, centroid 4v holds a trace of seasonality having the peak P_1 set to an annual frequency (Table 10).
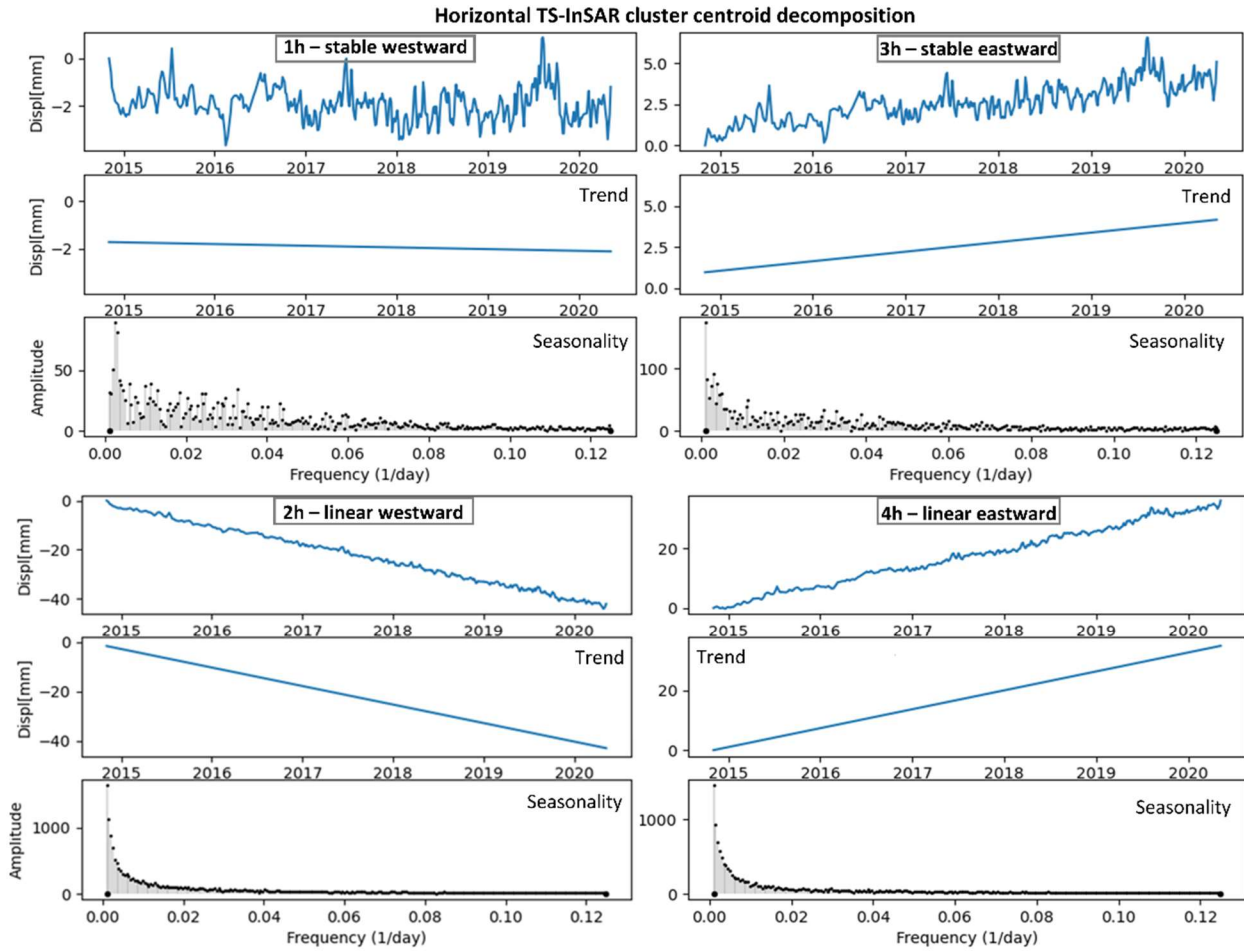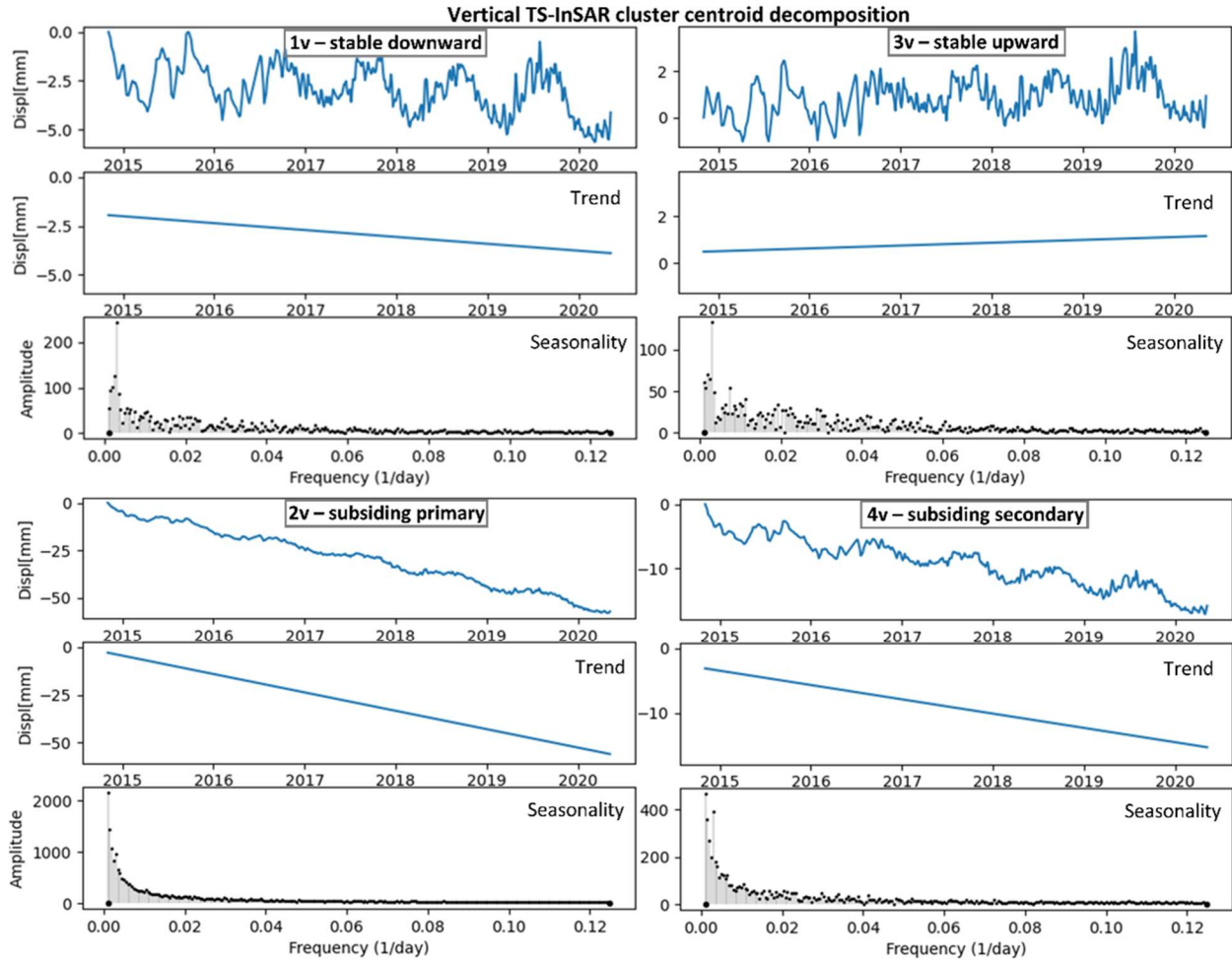
Figure 46. Vertical TS-InSAR cluster centroid decomposition into trend (Linear Regression) and seasonality (power spectrum) components.

Table 10. Clusters centroids components obtained from the Ordinary least squares Linear Regression and Discrete Fourier Transform analysis. RMSE indicates the Root Mean Square Error estimated for the linear model. P_0 and P_1 denote the two main spectral peaks encountered, with P_0 being the peak with the highest spectral power.

Cluster centroids	Linear Regression			Discrete Fourier Transform			
	Slope	Intercept	RMSE	Spectral peaks	Amplitude	Frequency	Period (days)
1e	-0.08e-2	-1.72	0.67	P_0	88.92	0.24e-2	403
				P_1	81.18	0.30e-2	336
2e	-0.80e-1	-1.52	0.61	P_0	1637.68	0.99e-3	1008
				P_1	1126.94	0.15e-2	672
3e	0.64e-2	0.97	0.64	P_0	173.15	0.99e-3	1008
				P_1	91.13	0.30e-2	336
4e	0.69e-1	-0.8e-1	0.84	P_0	1458.41	0.99e-3	1008

				P_1	940.21	0.15e-2	672
1v	-0.39e-2	-1.94	1.03	P_0	242.59	0.30e-2	336
				P_1	126.21	0.24e-2	403
2v	-0.11	-2.81	1.61	P_0	2143.65	0.99e-3	1008
				P_1	1443.22	0.15e-2	672
3v	0.13e-2	0.49	0.76	P_0	132.62	0.30e-2	336
				P_1	69.78	0.20e-2	504
4v	-0.02	-3.08	1.18	P_0	465.70	0.99e-3	1008
				P_1	391.62	0.30e-2	336

7.3.2.3. Comparison between clustering results and regional inventories

The spatial distribution of the retrieved clusters is explored with respect to the pre-existing regional inventories listed in Section 7.1.1. By looking at the degree of spatial aggregation of classified target feature trends within mapped unstable areas, it is possible to both evaluate the meaningfulness of the proposed DM method and to summarize the average temporal displacement behaviour of possibly active sources of ground motion. In this regard, Figure 47 depicts the spatial partition of the clustered features according to the available inventories. According to what highlighted in section 7.3.2.1 and 7.3.2.2, Cluster2e and Cluster4e are indicative of temporally linear movements with high rates of displacement: despite being the least frequent throughout the study area, a very high percentage of 2e and 4e features tends to aggregate within mapped landslides (38.2% and 24.6% of Cluster2e and Cluster4e features, respectively) and within mapped DSGSDs (55% and 70.2% of Cluster2e and Cluster4e features, respectively). Comparable numbers are obtained for the same inventories if taking into consideration Cluster2v and Cluster4v: they show to have the highest rates of displacement between the retrieved vertical TS-InSAR clusters. In general, it can be stated that DSGSDs do not show to displace by following a predominant component; indeed, both vertical and horizontal motions can be equally recognized within DSGSDs.

The mining/quarrying extraction sites alongside the rock glacier inventory have a generally low grade of spatial intersection with the classified features. This can be linked to their relatively low areal extension, if compared to landslides and DSGSDs, which are higher in number (2,329 landslides and DSGSDs vs 86 extraction sites and 941 rock glaciers) and underpin a greater areal extension throughout VdA (591 km^2 vs 71 km^2). For the same reason, the density of features is generally low for landslides and DSGSDs and higher for mining and quarrying sites. What can be worth mentioning is the relative higher intersection of Cluster2v (compared to the other clusters) with the rock glacier inventory: this can be regarded as evidence of the active state of movement of some of the phenomena mapped throughout VdA.
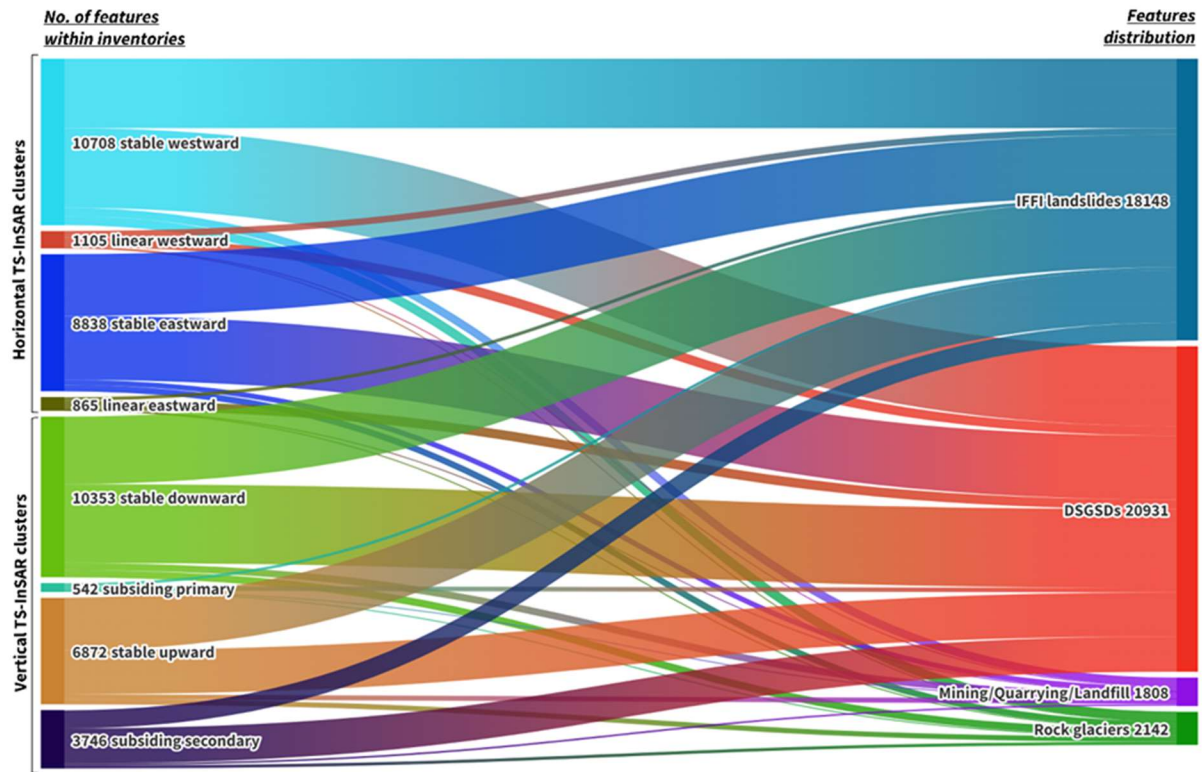
Figure 47. Sankey diagram for the spatial comparison between the classified features of the horizontal and vertical TS-InSAR datasets and the pre-existing regional inventories relatable to possibly active source of ground motion.

The presented method shows its full potential of detecting and characterizing different patterns of motion when looking at local unstable areas, such as the DSGSD depicted in Figure 48. In particular, the clustered TS-InSAR features demonstrate to be aggregated in space suggesting the existence of phenomena influencing contiguous targets (Figure 48). Moreover, the spatial pattern is coherent with respect to the downhill slope direction, the displacement direction, and the main morphological units. Indeed, Cluster2e features, which denote a mostly linear displacement towards West, stand out on a slope facing West, while, since Cluster3e and Cluster4e reside on a mostly opposite-oriented slope, they both hold a trend of movement towards the east (Figure 48a). By looking at vertical clustered TS-InSAR features (Figure 48b), it is possible to extend the degree of interpretability of the deformation framework. The spatial distribution of the classified features, when analysed according to the displacement reprojected on the vertical direction, is rather different and this enables us to make specific considerations on how differently the various sectors of the DSGSD (e.g., the head scarp, the main body) have been deforming. For instance, Cluster2v features are in the scarp areas and highlight the cores of the DSGSD moving downslope with the highest rates of displacement and a mostly linear trend (as shown in Figure 46). The detected most active zones fade into the surrounding Cluster4v features punctuating the main body of DSGSD (Figure 48b), which are indicative of a motion with a less severe downward displacement trend and with a trace of seasonality. In particular, it should be mentioned that several studies conducted over VdA and the Alpine chain (Bajni et al., 2021; Frayssines & Hantz, 2006; Paranunzio et al., 2019) report that the frequency of not-ordinary climatic conditions, which follow a seasonal trend, is majorly related to slope failures events. What is worth noting is that spatially coincident features retrieved from the analysed eastward and

vertical TS-InSAR datasets depict dissimilar deformation patterns in terms of both temporal trends and periodicity (Figure 48c; Figure 48d): the proposed method allows to temporally extract the undergoing displacement into the eastward and vertical components of motion, enabling a deep understanding of how differently the motion increase and/or fluctuation has distributed over time along the two orthogonal directions.
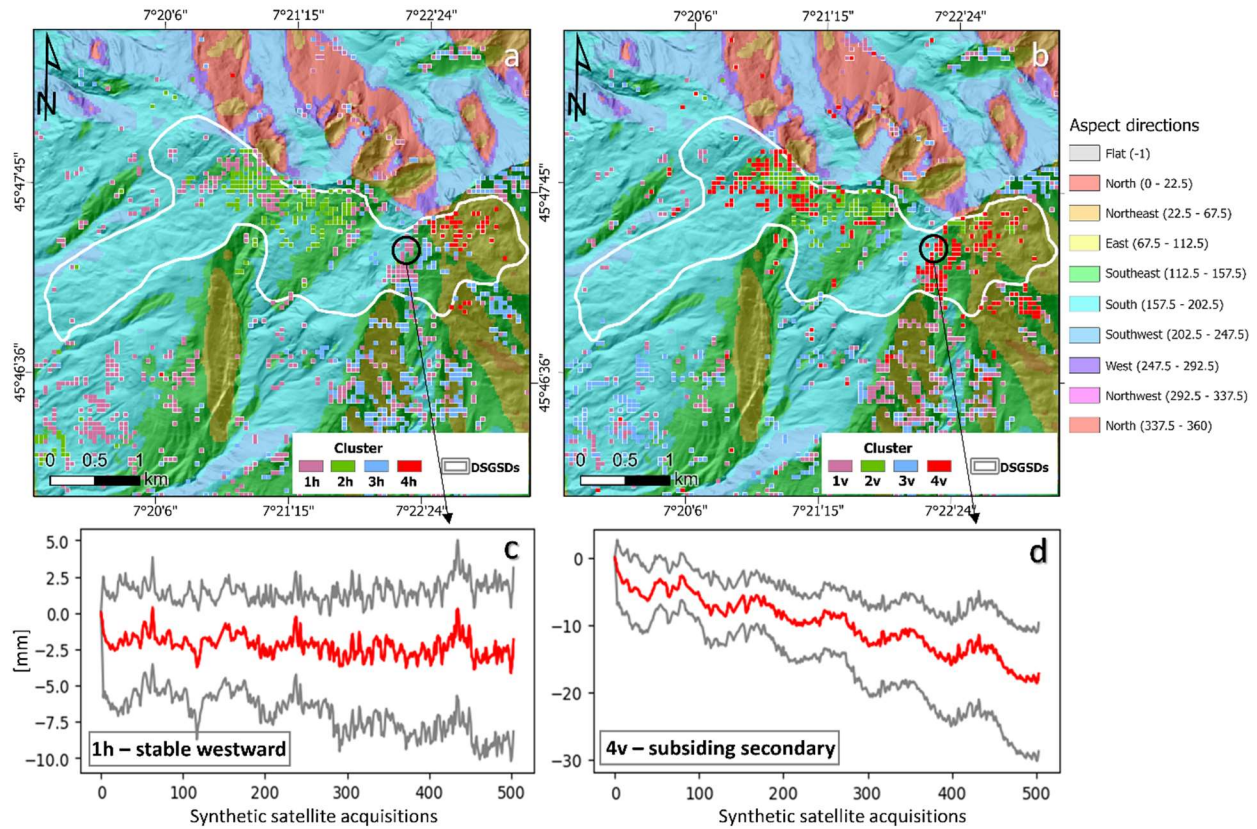

Figure 48. Sample result of the applied procedure regarding a DSGSD phenomenon. Different temporal displacement behaviors are noticeable as witnessed by the features classified according to different types of (a)eastward and (b)vertical TS-InSAR clusters. The relative (c; d) cluster plots of some spatially coincident features denote that the ongoing displacement can be unevenly subdivided between the eastward and vertical components of motion, hence enabling an accurate interpretation of the phenomenon.

The strength of the method is further evidenced by the example shown in Figure 49, depicting unstable conditions within a landfill site, in the municipality of Aosta. The deformation map containing both ascending and descending TS-InSAR S1 MPs (Figure 49a) sets the starting point of our analysis, while the clustered features retrieved for the vertical component of motion (Figure 49b) represent one part of a twofold outcome of the DM approach.

A well-known limitation of conventional InSAR techniques is the low sensitivity to North-South (N-S) deformations. Since modern spaceborne radar systems move along near-polar orbits, there is no diversity in the viewing geometries to accurately estimate N-S displacements. Therefore, N-S displacement time series are here neglected. Aside from the intrinsic limitations of InSAR, some other constraints may arise from the proposed approach. For instance, the MPs aggregation into grid cells might cause a significant

drop in terms of spatial resolution (Figure 49b) hindering the correct detection of very small-scale surface movements. Also, the method does not take into account any weight related to the quantity of MPs from the two geometries and contained within the cell. Moreover, it should be considered that the absence of regular ascending and descending acquisitions constitutes a major obstacle for the applicability of the proposed method.

Anyway, the procedure ensures an unprecedent improved and quick readability of big amounts of radar data in the following terms: 1) motion is directly interpreted into purely vertical and eastward displacements, allowing an analysis of the earth surface independent of the radar look geometry; 2) the large quantity of targets is squeezed into few data-driven clusters based on the similar temporal displacement behaviour, hence enabling to rapidly pinpoint common deformation processes; 3) the detected process is further broken down into the main temporal trend of deformation and into the eventual seasonality component.

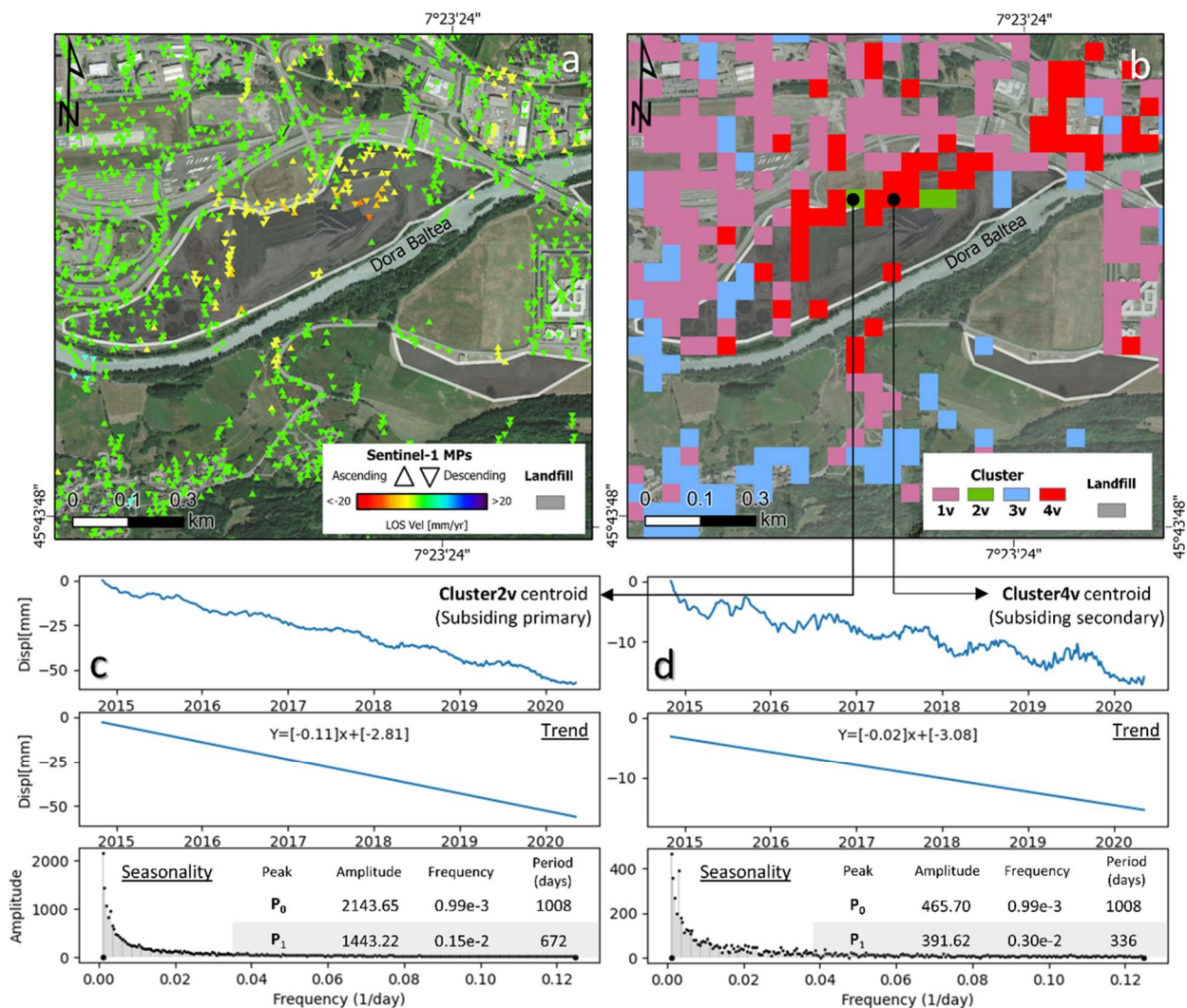Figure 49. (a) mean LOS velocity of ascending and descending MPs on a landfill site located in the Aosta municipality. (b) Clustering results of the reprocessed vertical TS-InSAR dataset where the clustered features point out an ongoing subsiding process with a mostly linear trend and high rate of displacement (c) and with a more pronounced seasonal pattern while having a lower displacement rate (d).

8. Discussions

Within this Thesis work, unsupervised DM techniques and supervised machine learning techniques are tested and critically analysed within the framework of Big SAR Data management for the efficient extraction and characterization of the Earth's surface deformations due to natural hazards or anthropogenic activities.

The data produced from Earth Observation satellites has recently become very abundant. As a direct expression of that, the Upstream and Midstream of Earth Observation data have registered a great technological development, translating into both very powerful satellites infrastructures carrying accurate sensors (Tatem et al., 2008) and consolidated data processing techniques (Crosetto et al., 2016) exploiting web cloud services to reduce data storage and analysis burden (Zinno et al., 2015). Nowadays one of the main interests of this market is the Downstream of the Earth observation Big Data (Probst et al., 2017). This means finding optimal ways the convert this data into value-added products to allow a greater exploitation of the available information and to make them more understandable and easier for a practical use. The manual analysis is no longer an option, however the main challenges for studying and analyse this data are their size, the complex nature, and the availability of datasets used for training data. Because of this, there has been a prominent trend in techniques used to automate these processes and host the data processing in online cloud servers. These processes include Data Mining and Machine Learning, and they are concerned with clustering procedures and other statical-based techniques.

Starting from the 90's, several studies all around the world have demonstrated the usefulness of single and multi-interferogram approaches for detecting ground motions at different scales. Thus, the monitoring and mapping capabilities of the interferometric techniques, and especially one above all the Persistent Scatterer Interferometry, are widely used thanks to the unmatched capability of covering large areas coupled with a good temporal repeatability and a great cost-benefits ratio.
Nowadays, updated interferometric products at regional, national and continental scale (Crosetto et al., 2020b) are already available and deployed to obtain information on ground deformation phenomena. The continuous stream of Big SAR Data and thus consequently processed interferometric datasets accounts for billions of point-wise targets that are characterized by a time series of information that includes the "deformation history" of every measurement point. Machine intelligence is therefore a helpful tool in such cases, since data-driven algorithms and statistical models allows to analyse datasets and then draw inferences from identified patterns to detect the most useful information for studying and preventing geohazard-related ground deformations in the framework of civil protection activities.

It is also important to remark that the satellite interferometry intrinsic limits must be taken into account. In fact, even if data acquisition is made possible in inaccessible areas which could not be investigated with the same measurement points density without time and money consuming in situ surveys, on the other side, not all types of land covers are suited for radar remote sensing. For this reason, SAR satellites can acquire reliable data only where electromagnetic-stable targets can be found. This is not the case of highly vegetated areas. The spatial density of points is also dependent on the acquisition band of the sensor; in fact, in an urban environment the X-band satellite can achieve a density of points hundreds to thousand times higher than C-band data because of their higher spatial and temporal resolution (Crosetto et al.,

2010). On the other hand, the longer L-band sensors will perform better in rural and forested environments than C- and X-band satellites because of their higher penetration capacity (Aoki et al., 2021). Another big drawback is the impossibility of detecting fast movements. The velocity limits for SAR satellites can be calculated based on the wavelength and on the satellite revisiting time. When multi-interferogram approaches are used the displacement rate threshold varies not only for the band length but also for the revisiting time. For Sentinel1 C-band constellation it is estimated to be around 85 cm/year (considering a 6-day revisiting time, when Sentinel1B was still operating along with Sentinel1A; mission-status-report-415-11-24-october-2022). Considering that, longer wavelength (L-band) can record greater displacements between two acquisitions of the sensor. On the other hand, ERS 1/2 satellite have a detection threshold of 15 cm/year, showing the importance of the revisiting time to limit the decorrelation of the signal between two consecutives interferograms

In this Thesis work the potentialities of the PSI technique for detecting, classifying and analysing hydrological phenomena have been fully exploited, always considering the intrinsic drawbacks of the satellite interferometry. Through the presentation of three post-processing approaches and value-adding tools, the conducted spatial and temporal analysis of very large A-DInSAR datasets has allowed to take a step forward on making radar data more usable and accessible to non-expert users. The proposed approaches, notwithstanding realized for large scale analysis, still preserve a high level of detail with respect to the local scale characterization.

The first proposed approach (see chapter 5) lies within the framework of rapid detection and characterization of the most active ground motion phenomena affecting the Earth's surface. The activity of filtering and mapping hotspots of ground displacements from nation-wide A-DInSAR datasets proves to be a particularly useful technique in order to retain meaningful information suited for risk mitigation strategies of the territory.

The improvements that have been made in the field of the analysis and exploitation of the interferometric data move in parallel with the rising quality and quantity of SAR images (especially considering the temporally coherent acquisition plan of Sentinel 1), along with the continuous development of more precise and efficient A-DInSAR algorithms. The P-SBAS approach can meet the present need of boosting the efficiency in computing precise deformation maps out of extremely large Sentinel 1 SAR archives with a worldwide coverage and unprecedent temporal sampling. The present research demonstrates that, starting from continuous SAR-derived information, it is nowadays possible to obtain a periodical long-term mapping plan of the sectors of territories affected by relevant surface deformations, thus guaranteeing a support in the assessment of the location, the extent, the probable underlying cause, the magnitude and direction of movement of the detected phenomena.

The performance of the GIS-models adopted for this research demonstrates the applicability of our procedure to very large interferometric datasets, even with not high-performance machine. The equipment used for the implementation of the methodology was an Intel Core i7-4790QM at 3.60 GHz, 4 cores, 8 threads, 8 MB cache, 16 GB RAM, 250 GB SSD disk and Windows 10, 64 bits. The performance testing showed that less than 21 min were needed to perform spatial clustering (Phase I of the post-processing chain) over the ascending geometry A-DInSAR dataset covering whole Italy (8,066,243 MPs), while it needed less than 7 min to perform the classification of the Moving Area Clusters (MAC) ascending database (Phase II of the post-processing chain), which accounts for 8,278 polygonal clusters. Moreover,

it should be pointed out that the manual preparation of the input data (e.g., pre-existing inventories, ancillary information) has required additional time.

Spatial clustering analysis, aside from suffering the same inherent limitations of interferometric techniques, strongly depends on the input parameters that are subjectively set in the semi-automatic GIS model: the estimation of both magnitude and areal extent of the observed ground deformation phenomena are tight to the setting of the displacement rate threshold, the buffer size and the minimum cluster size. Another important drawback is the classification of the underlying process which has triggered the diffuse ground deformation, thus resulting in the generation of MAC. The ancillary information and the pre-existing inventories are sometimes unavailable, incomplete, or not updated, hindering a correct pre-screening of the deforming areas. Hence, the combination of other available source of information, such as slope maps and the mean value of vertical and horizontal components of deformation, when this information is available, may become necessary (see Section 5.2.3). In particular, the lack of adequate ancillary information and the insufficient interferometric data resulted in 1,044 unclassified MAC, which corresponds to 7.1% of the total clustered areas considering both the ascending and descending database. On the other hand, a total of 8,234 out of 14,638 MAC (considering both geometries of acquisition) have been classified as potential landslides, potential subsidence or potential uplift, solely on the basis of the ratio between mean vertical and horizontal components of displacement and local slope value. In terms of number of units, either potential landslide and potential subsidence classes are two of the main representative categories within the classified MAC database (i.e., 28.1% and 27.9% respectively, considering the totality of MAC), thus suggesting that many sectors of the Italian territory are suffering from ground deformation phenomena which have not still been recorded by the present inventories. The significance of the semi-automated classification method inherent with the two potential phenomena is tied to some considerations carried out on MAC classified as landslide and subsidence according to the pre-existing geohazard inventories, which therefore can be considered as "ground truth".

Considering the LOS components mean ratio values ($|V_V|/|V_H|$), the 100% of the subsidence-MAC (534 units) are associated to values greater than 1, while the landslide-MAC (2083 units) accounts for 1879 units with values less than 1 and 204 units with values greater than 1. Having also considered a 5° slope angle threshold to distinguish between landslide and subsidence events (Iwahashi et al., 2003; Lee et al., 2004), it is reported that, again, the totality of subsidence-MAC stands below that limit (average slope value less than 5°), while 2058 out of 2083 landslide-MAC stand above that limit (average slope value greater than 5°).

It is evident that the vertical component of the movement is predominant in subsidence-related moving areas, while the landslide-related ones does not show a sharp tendence in this sense. The scattered values associated to slope deformations could be ascribed to a great variability in terms of kinematic behaviours (since all type of slope movements are here considered) and to the relative spatial distribution of the MPs within each landslide, where the deformation pattern may be partially hidden to satellite sensors (Frattini et al., 2018). Despite the limitation that could arise in the pre-screening of potential landslides, it emerges the relative high confidence in correctly interpreting MAC as potential subsidence phenomena.

The presented work is to be considered one of the first example of semi-automated mapping and pre-assessment of the Earth surface deformations at national scale. Other approaches with a similar task have been previously applied only for local scale (Barra et al., 2017), basin scale (Bianchini et al., 2012) or regional scale (Montalti et al., 2019; Solari et al., 2019) analysis. The presented site-specific pattern deformation analysis (i.e., the third phase of the post-processing chain) constitutes an innovative scheme for data interpretation and visualization. Indeed, the colormap and the polar plot, proposed throughout the various case studies (see section 5.3.3), enable a more precise and comprehensive view of the magnitude and direction of the observed motion occurred on the Up-Down and East-West plane.

Leveraging on thematic information already available, such as landslide databases (Trigila et al., 2007), outputs of this mapping systems may be fully integrated along with the upcoming deliveries of continental-scale ground motion data (EGMS), as they represent the key step for the optimal management of billions of MPs, for the benefit of expert and non-expert users, for public and private scenarios. Moreover, the creation of updated databases regarding the ongoing ground deforming areas could serve for: (i) reference for Civil Protection purposes and policy makers devoted to territorial planning; (ii) prioritization of the available funds in the framework of development of new structural plans; (iii) updating of geohazard inventories; (iv) creation of more advanced products such as landslide intensity maps (Catani et al., 2005; Lu et al., 2014), given the lack of information on the expected velocities for a large number of landslides (Cigna et al., 2013); (v) validation data for landslide and subsidence susceptibility maps.

The second proposed procedure (see chapter 6) can be regarded as a further evolution in the methodological approach respect to the one proposed in chapter 5. In both approaches A-DInSAR deformation map are analysed through data fusion with ancillary informative layers and open-source inventories, however, while the first procedure relies on an unsupervised learning strategy, the second procedure is based on a supervised learning manner.

In this work it is address once again the need to find a transferable approach to safely analyse and exploit large volumes of interferometric data in the framework of screening surface displacements related to both natural and anthropogenic sources. The presented methodology aims at the automated detection of pixel scale ground deformations, departing from A-DInSAR deformation maps covering very wide areas and composed of millions of MPs. In the presented case, the feasibility of a Machine Learning model for the prediction of the cause of deformation tied to radar targets with outlying velocity values is evaluated by using an interferometric dataset covering Northern Italy as a proof of concept. Based on the time spent on building the training dataset used within this work, we estimate that the expert interpretation of high velocity MPs covering Northern Italy might take up to several weeks.

This data-driven approach relies on the transferable use of Random Forest (RF) applied to sparse MPs derived from P-SBAS A-DInSAR and on input parameters which can be easily extracted from land cover maps, geohazard inventories and DEMs. The results show great prediction performances for a four-class overall classification with a F1-score equal to 9.3e-1, a True Positive Rate (TPR) equal to 9.4e-1 and an Accuracy value of 9.7e-1. Visually inspected and interpreted MPs are used as geostatistical instances for modelling and exploring the spatial distribution of different ground deformation processes.

Some limitations are linked with the usage of the presented methodology. Besides the largely acknowledged drawbacks of interferometry techniques (e.g., impossibility to capture fast movements,

difficulties in retrieving coherent data from vegetated areas), the quality of the input A-DInSAR dataset and of the ancillary information is crucial for the correct extrapolation of statistically relevant value intervals with regards to the deployed parameters used for the RF model. In the case study, the coexistence of different kind of deforming phenomena within the same area occasionally leads to a mismatch in the screening process of the source of displacement. In particular, mining areas might show typical complex displacement patterns which can replicate subsiding or slope movements, especially if restricting our analysis to satellite radar data only. This translates into the very high importance value of the distance parameter related to Corine Land Cover (CLC) mining sites, which demonstrates to be a real defining factor in the proposed model. Moreover, since ND (i.e., noise or other phenomena) is a broad ensemble containing different sources of unstable radar signals, the method understandably registers weaker predicting performances for that class. In this respect, it should be pointed out that the present approach is liable to misclassify MPs if the built model is not trained with a proper number of features or an adequate selection of explanatory variables. Moreover, the method requires the decomposition of the LOS velocities into the vertical and horizontal components, which can be accomplished only where both ascending and descending measurements are available. This implies a decreasing of the scene coverage, thus resulting in the decrease of the number of classifiable targets. Nevertheless, the method proves to be robust in overall correctly classifying the retrieved classes, demonstrating that by only using satellite data, it is possible to achieve a quick, automated, and reliable screening of the source of deformation of unstable radar targets.

In the third proposed procedure (chapter 7) it is introduced an unsupervised data mining for the automatic recognition of complex surface motion trends in order to optimally interpret large-scale displacement InSAR Time Series (TS-InSAR) datasets, providing entirely data-driven results to be employed in the management and mitigation of the geohazards impact on both vulnerable population and infrastructures. The deployed assembled code provides processing performances resilient to the ever-increasing volumes of SAR data and an automatically balanced cost-benefit strategy to produce accurate outputs for stakeholders in need of handy tools for territorial planning operations and civil protection purposes.

One of the main novelties of the proposed unsupervised data mining approach is the integration of an automated pipeline for the extraction of TS-InSAR components of motion into the vertical and eastward directions, departing from overlapping targets belonging to both ascending and descending LOS geometry. It is also introduced a simple method to automatically tailor the grid size for the spatial join of the targets into the squared grid cell features. The eastward and vertical TS-InSAR datasets decomposition is then followed by a well-established procedure of data-dimensionality reduction and clustering (PCA and K-Means), where the obtained clusters are ascribable to a defined range of temporal displacement patterns. To enhance the interpretability of the data-driven results, the designed original code automatically addresses the main underlying temporal trend and periodicity of each cluster centroid by applying the Ordinary least squares Linear Regression and Discrete Fourier Transform analysis. As a result of this novel pipeline, it is possible to greatly improve the InSAR interpretation of movements such as mass wasting processes, subsidence, or man-made activities. The procedure can be therefore regarded as a key contribution to the optimal management of big InSAR data, for the benefit of expert and non-expert users, for public and private scenarios.

Also, some limitations may arise from the proposed method, hence influencing the usability of the obtained results. For instance, the MPs aggregation into grid cells might cause a significant drop in terms of spatial resolution (see section 7.2.1) hindering the correct detection of very small-scale surface movements. Also, the method does not take into account any weight related to the quantity of MPs from the two geometries and contained within the cell. Moreover, it should be considered that the absence of regular ascending and descending acquisitions constitutes a major obstacle for the applicability of the proposed method. Anyway, whether only a single geometry of acquisition is available, the approach can be thought as still valid to reduce a TS-InSAR dataset into its PCs, to evaluate the temporal clusters and the related temporal components.

The series averaging within each grid cell and the rationale of temporal interpolation can possibly make an impact in correctly reproducing the motion trend along the vertical and eastward components. Since many sources of errors due to radar acquisitions are responsible for the generally high noise component in the TS-InSAR signals, a low resampling rate (down sampling) might result in a general positive effect of smoothing out errors from the series data. On the other hand, an excessive loss of data might cause an oversimplification that can be detrimental to the accuracy of the analysis. In this regard, the TS-InSAR dataset quality also plays a role: discontinuity and unbalance in radar data acquisitions might bring to a misleading reconstruction of the TS-InSAR components of motion, regardless of the adopted interpolation method (see section 7.3.1). This, rather than affecting the clustering performances, might have pronounced effects on the trend and seasonality components estimation. In view of the above-mentioned, Sentinel 1 archives coupled with a balanced resampling rate have shown to be appropriate for time-dependent series data mining, given the temporal consistency in data collection.

The restriction of the trend and seasonality analysis to the cluster centroid only, despite partially defeaturing the cluster series characterization, requires a much lower computational effort while still constituting a valuable operation for the readability of the data-driven TS-InSAR clusters. Although the here presented method is thought to extract, enrich, and summarize meaningful information from large scale InSAR data in an optimized way, more detailed results can be obtained by either using a reduced InSAR dataset as input (for small-scale analysis) or by deliberately increasing the number of Principal Components (i.e., number of clusters) to retain. Moreover, the presented method can be further developed by integrating Independent Component Analysis (ICA) to retrieve the additive subcomponents of PCA-based results (i.e., to characterize the temporal behaviours of vertical and horizontal TS-InSAR clusters).

The three proposed approaches illustrate different but complementary ways of post-processing that could be merged in one single workflow to detect areas of active motions, automatically classify the origin of their deformation through cross-correlation of the satellite data with ancillary maps and, finally determine the principal components describing the historical trends of motion. Therefore, all the described procedures constitute ready-to-use tools for risk assessment and management activities conducted by entitled governmental departments and specialized bodies.

9. Conclusions

In the framework of the increasing availability of space borne radar data with unprecedent coverage (e.g., the European Ground Motion Service, https://egms.land.copernicus.eu/) the downstream of InSAR-derived products can be considered a developing market sector which is an essential asset for multi-hazard events monitoring and mapping linked to ground surface and single asset displacements and for risk management and reduction. Consequently, this thesis work serves as an example of the primary importance of implementing automated solutions aimed at performing integrated analysis of large amounts of interferometric data.

In this PhD Thesis the potentialities and drawbacks of the interferometric technique have been presented and then exploited to define three different value-adding procedures, applied in different environment and at different scales, for the enhanced interpretation of interferometric products. This Thesis work represents the main outcome of a three years-long activity at the Department of Earth Sciences of the University of Florence and its goal is to test and evaluate the potential applicability of spatial and temporal post-processing analysis routines of space-borne SAR data, processed by means of different PSI approaches, as operational tools for the characterization of the Earth's surface motion tied to natural or anthropogenic processes.

Three study cases regarding interferometric datasets retrieved from Sentinel 1 multi-temporal radar images acquired over Italy with varying extensions (from regional to a nation-wide extent) were deployed as proof of concepts for three semi-automated and automated InSAR post-processing approaches.

In the first instance, the analysis of nation-wide P-SBAS results through unsupervised clustering has enhanced the detection of newly activated surface deformation events with a millimetric accuracy, thus demonstrating that such activities may constitute an asset for future geohazards risk management practices. In particular, a semi-automated GIS model consisting of a sequence of geoprocessing tools was designed and tested for double geometry A-DInSAR deformation maps spanning 2015–2018 and covering whole Italy. A total number of 14,638 Moving Area Clusters (MAC) constituted by at least 3 neighbouring Measurement Points (MPs) with a displacement threshold of ± 1 cm/year has been retrieved from both the ascending and descending DInSAR datasets which were then pre-screened in 11 categories: earthquake-induced deformations, volcanic-related processes, landslides, potential landslides, subsidence, potential subsidence, potential uplift, dump site activities, mining activities, construction site activities and unclassified.

The second methodology regarded the use of a Machine Learning technique, called Random Forest (RF), has been applied to large-scale deformation maps covering the northern part of Italy. After a training phase realized through a dataset based on expert and manual classification, the model was able to identify landslide, subsidence, mining-related ground motion and noisy signals or other sources of deformation. The proposed approach relies on the use of explanatory variables extracted from the A-DInSAR dataset to be analysed and from freely accessible informative layers such as Digital Elevation Model (DEM), land cover maps, and geohazard inventories. In general, the model performance achieved promising results with an overall accuracy of 0.97, a true positive rate of 0.94 and an F1-Score of 0.93.

The third approach concerned the automated retrieval of clusters from InSAR Time Series (TS-InSAR) of displacement, therefore primarily focusing on the temporal dimension of interferometric results. The focus was placed on SqueeSAR™ results covering Valle d'Aosta region (North-western Italy). The core of the designed algorithm relies on a statistical data dimensionality reduction technique prior to a K-Means-based method for the automated spatial clustering of TS-InSAR data based on similar temporal displacement behaviour of the analysed radar targets. The latter are fed to the algorithm prior a reprojection of the LOS displacement series along the vertical and eastward directions, where the applied data fusion rationale is robust about the spatial and temporal dimensions. Ultimately, the retrieved cluster series were further broken down into their trend and seasonality components to enhance the readability of the classified features. 4 Principal components were retained for both Vertical and Horizontal Time Series corresponding to 4 different clusters for each of the two analysed direction of movements, which were able to highlight the coexistence of linear and seasonal deformation patterns tied to both slope mass wasting processes and quarrying activities present in the Valle d'Aosta region.

The proposed value-adding tools case demonstrated that starting from A-DInSAR datasets composed of billions of MPs is possible to largely improve the readability of remotely sensed data and to make the most out of them by using automatic post-processing procedures. Thus, it is demonstrated their suitability for mapping and analysing multi-hazards in the framework of territorial planning operations and civil protection purposes.

Many potential improvements can be developed in the next future to enhance the presented approaches. To advance the detection, the classification and the analysis of newly activated surface deformations new data sources can be involved in the data fusion process. In particular, by enriching the amount and the diversity of data and by deploying complex Deep Learning algorithms more accurate results can be produced even. In this sense, ground deformation forecasting approaches will be tested in the near future by focusing on the prediction of precursors to rapid collapses or anomalous behaviors via TS-InSAR analysis. Indeed, new, original, and transferable frameworks for routine environmental monitoring should be developed to increase the potentiality of interferometric products and to fill the gap between the scientific community and the end user, for rapid disaster information and response. Being natural and anthropogenic multi-hazards a serious threat to human life, properties and environment, the evaluation of their impact and the development of adequate mitigation strategies requires a comprehensive analysis. This can be accomplished through decision-making tools which enables to fully identify the source of multi-hazard and multi-scale events and the expected related effects with the aim of designing effective resilience strategies.

Publications derived from this PhD Thesis

Davide Festa, Nicola Casagli, Francesco Casu, Pierluigi Confuorto, Claudio De Luca, Matteo Del Soldato, Riccardo Lanari, Michele Manunta, Mariarosaria Manzo & Federico Raspini (2022). Automated classification of A-DInSAR-based ground deformation by using random forest, GIScience & Remote Sensing, 59:1, 1749-1766, DOI: 10.1080/15481603.2022.2134561.

Davide Festa, Manuela Bonano, Nicola Casagli, Pierluigi Confuorto, Claudio De Luca, Matteo Del Soldato, Riccardo Lanari, Ping Lu, Michele Manunta, Mariarosaria Manzo, Giovanni Onorato, Federico Raspini, Ivana Zinno & Francesco Casu (2022). Nation-wide mapping and classification of ground deformation phenomena through the spatial clustering of P-SBAS InSAR measurements: Italy case study, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 189, Pages 1-22, ISSN 0924-2716, DOI: https://doi.org/10.1016/j.isprsjprs.2022.04.022.

References

Acocella, V. (2008). Activating and reactivating pairs of nested collapses during caldera-forming eruptions: Campi Flegrei (Italy). *Geophysical Research Letters*, *35*(17). https://doi.org/https://doi.org/10.1029/2008GL035078

Adam, N., Rodriguez Gonzalez, F., Parizzi, A., & Liebhart, W. (2011). Wide area persistent scatterer interferometry. *2011 IEEE International Geoscience and Remote Sensing Symposium*, 1481–1484. https://doi.org/10.1109/IGARSS.2011.6049347

Aly, M. H., Zebker, H. A., Giardino, J. R., & Klein, A. G. (2009). Permanent Scatterer investigation of land subsidence in Greater Cairo, Egypt. *Geophysical Journal International*, *178*(3), 1238–1245. https://doi.org/10.1111/j.1365-246X.2009.04250.x

Anantrasirichai, N., Biggs, J., Kelevitz, K., Sadeghi, Z., Wright, T., Thompson, J., Achim, A. M., & Bull, D. (2021). Detecting Ground Deformation in the Built Environment Using Sparse Satellite InSAR Data With a Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(4), 2940–2950. https://doi.org/10.1109/TGRS.2020.3018315

Ansari, H., Ruβwurm, M., Ali, M., Montazeri, S., Parizzi, A., & Zhu, X. X. (2021). InSAR Displacement Time Series Mining: A Machine Learning Approach. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 3301–3304. https://doi.org/10.1109/IGARSS47720.2021.9553465

Aoki, Y., Furuya, M., de Zan, F., Doin, M.-P., Eineder, M., Ohki, M., & Wright, T. J. (2021). L-band Synthetic Aperture Radar: Current and future applications to Earth sciences. *Earth, Planets and Space*, *73*(1), 56. https://doi.org/10.1186/s40623-021-01363-x

Ayalew, L., Yamagishi, H., & Ugawa, N. (2004). Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. *Landslides*, *1*(1), 73–81. https://doi.org/10.1007/s10346-003-0006-9

Bajni, G., Camera, C. A. S., & Apuani, T. (2021). Deciphering meteorological influencing factors for Alpine rockfalls: a case study in Aosta Valley. *Landslides*, *18*(10), 3279–3298. https://doi.org/10.1007/s10346-021-01697-3

Bakon, M., Oliveira, I., Perissin, D., Sousa, J. J., & Papco, J. (2017). A Data Mining Approach for Multivariate Outlier Detection in Postprocessing of Multitemporal InSAR Results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(6), 2791–2798. https://doi.org/10.1109/JSTARS.2017.2686646

Bamler, R., & Hartl, P. (1998a). Synthetic aperture radar interferometry. *Inverse Problems*, *14*(4), R1–R54. https://doi.org/10.1088/0266-5611/14/4/001

Bamler, R., & Hartl, P. (1998b). Synthetic aperture radar interferometry. *Inverse Problems*, *14*(4), R1–R54. https://doi.org/10.1088/0266-5611/14/4/001

Barra, A., Solari, L., Béjar-Pizarro, M., Monserrat, O., Bianchini, S., Herrera, G., Crosetto, M., Sarro, R., González-Alonso, E., Mateos, R. M., Ligüerzana, S., López, C., & Moretti, S. (2017). A Methodology to Detect and Update Active Deformation Areas Based on Sentinel-1 SAR Images. *Remote Sensing*, *9*(10). https://doi.org/10.3390/rs9101002

Berardino, P., Costantini, M., Franceschetti, G., Iodice, A., Pietranera, L., & Rizzo, V. (2003). Use of differential SAR interferometry in monitoring and modelling large slope instability at Maratea (Basilicata, Italy). *Engineering Geology*, *68*(1), 31–51. https://doi.org/https://doi.org/10.1016/S0013-7952(02)00197-7

Berardino, P., Fornaro, G., Lanari, R., & Sansosti, E. (2002). A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(11), 2375–2383. https://doi.org/10.1109/TGRS.2002.803792

Berti, M., Corsini, A., Franceschini, S., & Iannacone, J. P. (2013). Automated classification of Persistent Scatterers Interferometry time series. *Natural Hazards and Earth System Sciences*, *13*(8), 1945–1958. https://doi.org/10.5194/nhess-13-1945-2013

Bevilacqua, A., Neri, A., de Martino, P., Isaia, R., Novellino, A., Tramparulo, F. D., & Vitale, S. (2020). Radial interpolation of GPS and leveling data of ground deformation in a resurgent caldera: application to Campi Flegrei (Italy). *Journal of Geodesy*, *94*(2), 24. https://doi.org/10.1007/s00190-020-01355-x

Bianchini, S., Cigna, F., Righini, G., Proietti, C., & Casagli, N. (2012). Landslide HotSpot Mapping by means of Persistent Scatterer Interferometry. *Environmental Earth Sciences*, *67*(4), 1155–1172. https://doi.org/10.1007/s12665-012-1559-5

Blanco-Sànchez, P., Mallorquí, J. J., Duque, S., & Monells, D. (2008). The Coherent Pixels Technique (CPT): An Advanced DInSAR Technique for Nonlinear Deformation Monitoring. In *Earth Sciences and Mathematics* (pp. 1167–1193). Birkhäuser Basel. https://doi.org/10.1007/978-3-7643-8907-9_10

Bonì, R., Bosino, A., Meisina, C., Novellino, A., Bateson, L., & McCormack, H. (2018). A Methodology to Detect and Characterize Uplift Phenomena in Urban Areas Using Sentinel-1 Data. *Remote Sensing*, *10*(4), 607. https://doi.org/10.3390/rs10040607

Botteghi, S., Ventisette, C., Montanari, D., Manzella, A., Moretti, S., Dipartimento, D., Scienze, D., & Terra. (2012, May). *CONTRIBUTION OF PERSISTENT SCATTERER INTERFEROMETRY (PSI) TO MAP SURFACE DISPLACEMENT IN THE TRAVALE – RADICONDOLI GEOTHERMAL AREA (TUSCANY, ITALY) Institute of Geosciences and Earth Resources*.

Bovenga, F., Nitti, D. O., Fornaro, G., Radicioni, F., Stoppini, A., & Brigante, R. (2013). Using C/X-band SAR interferometry and GNSS measurements for the Assisi landslide analysis. *International Journal of Remote Sensing*, *34*(11), 4083–4104. https://doi.org/10.1080/01431161.2013.772310

Bozzuto, P., & Geroldi, C. (2021). The former mining area of Santa Barbara in Tuscany and a spatial strategy for its regeneration. *The Extractive Industries and Society*, *8*(1), 147–158. https://doi.org/https://doi.org/10.1016/j.exis.2020.09.007

Braun, A. (2019). *Radar satellite imagery for humanitarian response. Bridging the gap between technology and application*. 7 Mathematisch-Naturwissenschaftliche Fakultät.

Bredal, M., Dehls, J., Larsen, Y., Marinkovic, P., Lauknes, T. ~R., Stødle, D., & Moldestad, D. ~A. (2019). The Norwegian National Ground Motion Service (InSAR.no): Service Evolution. *AGU Fall Meeting Abstracts*, *2019*, G13C-0559.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brengman, C. M. J., & Barnhart, W. D. (2021). Identification of Surface Deformation in InSAR Using Machine Learning. *Geochemistry, Geophysics, Geosystems*, *22*(3), e2020GC009204. https://doi.org/https://doi.org/10.1029/2020GC009204

Capes, R., & Teeuw, R. (2017). On safe ground? Analysis of European urban geohazards using satellite radar interferometry. *International Journal of Applied Earth Observation and Geoinformation*, *58*, 74–85. https://doi.org/10.1016/j.jag.2017.01.010

Carraro, F., & Giardino, M. (2004). Quaternary glaciations in the western Italian Alps – a review. In J. Ehlers & P. L. Gibbard (Eds.), *Quaternary Glaciations Extent and Chronology* (Vol. 2, pp. 201–208). Elsevier. https://doi.org/https://doi.org/10.1016/S1571-0866(04)80071-X

Cascini, L., Fornaro, G., & Peduto, D. (2010). Advanced low- and full-resolution DInSAR map generation for slow-moving landslide analysis at different scales. *Engineering Geology*, *112*(1), 29–42. https://doi.org/https://doi.org/10.1016/j.enggeo.2010.01.003

Casu, F., Elefante, S., Imperatore, P., Zinno, I., Manunta, M., de Luca, C., & Lanari, R. (2014). SBAS-DInSAR Parallel Processing for Deformation Time-Series Computation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(8), 3285–3296. https://doi.org/10.1109/JSTARS.2014.2322671

Catani, F., Canuti, P., & Casagli, N. (2012). The Use of Radar Interferometry in Landslide Monitoring. *Paper Presented at 1st Meeting OfCold Region Landslides Network and 1st Symposium on Land-Slides in Cold Region*.

Catani, F., Casagli, N., Ermini, L., Righini, G., & Menduni, G. (2005). Landslide hazard and risk mapping at catchment scale in the Arno River basin. *Landslides*, *2*(4), 329–342. https://doi.org/10.1007/s10346-005-0021-0

Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. *Natural Hazards and Earth System Sciences*, *13*(11), 2815–2831. https://doi.org/10.5194/nhess-13-2815-2013

Chang, L., & Hanssen, R. F. (2016). A Probabilistic Approach for InSAR Time-Series Postprocessing. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(1), 421–430. https://doi.org/10.1109/TGRS.2015.2459037

Chaussard, E., & Farr, T. G. (2019). A New Method for Isolating Elastic From Inelastic Deformation in Aquifer Systems: Application to the San Joaquin Valley, CA. *Geophysical Research Letters*, *46*(19), 10800–10809. https://doi.org/https://doi.org/10.1029/2019GL084418

Cheloni, D., de Novellis, V., Albano, M., Antonioli, A., Anzidei, M., Atzori, S., Avallone, A., Bignami, C., Bonano, M., Calcaterra, S., Castaldo, R., Casu, F., Cecere, G., de Luca, C., Devoti, R., di Bucci, D., Esposito, A., Galvani, A., Gambino, P., … Doglioni, C. (2017). Geodetic model of the 2016 Central Italy earthquake sequence inferred from InSAR and GPS data. *Geophysical Research Letters*, *44*(13), 6778–6787. https://doi.org/https://doi.org/10.1002/2017GL073580

Chiodini, G., Selva, J., del Pezzo, E., Marsan, D., de Siena, L., D'Auria, L., Bianco, F., Caliro, S., de Martino, P., Ricciolino, P., & Petrillo, Z. (2017). Clues on the origin of post-2000 earthquakes at Campi Flegrei caldera (Italy). *Scientific Reports*, *7*(1), 4472. https://doi.org/10.1038/s41598-017-04845-9

Ciampalini, A., Raspini, F., Lagomarsino, D., Catani, F., & Casagli, N. (2016). Landslide susceptibility map refinement using PSInSAR data. *Remote Sensing of Environment*, *184*, 302–315. https://doi.org/https://doi.org/10.1016/j.rse.2016.07.018

Ciampalini, A., Solari, L., Giannecchini, R., Galanti, Y., & Moretti, S. (2019). Evaluation of subsidence induced by long-lasting buildings load using InSAR technique and geotechnical data: The case study of a Freight Terminal (Tuscany, Italy). *International Journal of Applied Earth Observation and Geoinformation*, *82*, 101925. https://doi.org/https://doi.org/10.1016/j.jag.2019.101925

Cigna, F., Bianchini, S., & Casagli, N. (2013). How to assess landslide activity and intensity with Persistent Scatterer Interferometry (PSI): the PSI-based matrix approach. *Landslides*, *10*(3), 267–283. https://doi.org/10.1007/s10346-012-0335-7

Cigna, F., del Ventisette, C., Liguori, V., & Casagli, N. (2011). Advanced radar-interpretation of InSAR time series for mapping and characterization of geological processes. *Natural Hazards and Earth System Sciences*, *11*(3), 865–881. https://doi.org/10.5194/nhess-11-865-2011

Cigna, F., & Tapete, D. (2021). Sentinel-1 Big Data Processing with P-SBAS InSAR in the Geohazards Exploitation Platform: An Experiment on Coastal Land Subsidence and Landslides in Italy. *Remote Sensing*, *13*(5). https://doi.org/10.3390/rs13050885

Coda, S., Confuorto, P., de Vita, P., di Martire, D., & Allocca, V. (2019). Uplift Evidences Related to the Recession of Groundwater Abstraction in a Pyroclastic-Alluvial Aquifer of Southern Italy. *Geosciences*, *9*(5). https://doi.org/10.3390/geosciences9050215

Colesanti, C., Ferretti, A., Prati, C., & Rocca, F. (2003). Monitoring landslides and tectonic motions with the Permanent Scatterers Technique. *Engineering Geology*, *68*(1), 3–14. https://doi.org/https://doi.org/10.1016/S0013-7952(02)00195-3

Colesanti, C., & Wasowski, J. (2006). Investigating landslides with space-borne Synthetic Aperture Radar (SAR) interferometry. *Engineering Geology*, *88*(3), 173–199. https://doi.org/https://doi.org/10.1016/j.enggeo.2006.09.013

Confuorto, P., del Soldato, M., Solari, L., Festa, D., Bianchini, S., Raspini, F., & Casagli, N. (2021). Sentinel-1-based monitoring services at regional scale in Italy: State of the art and main findings. *International Journal of Applied Earth Observation and Geoinformation*, *102*, 102448. https://doi.org/https://doi.org/10.1016/j.jag.2021.102448

Costantini, M., Falco, S., Malvarosa, F., Minati, F., & Trillo, F. (2009). Method of persistent scatterer pairs (PSP) and high resolution SAR interferometry. *2009 IEEE International Geoscience and Remote Sensing Symposium*, III-904-III–907. https://doi.org/10.1109/IGARSS.2009.5417918

Crosetto, M., Devanthéry, N., Cuevas-González, M., Monserrat, O., & Crippa, B. (2015). Exploitation of the full potential of PSI data for subsidence monitoring. *Proceedings of the International Association of Hydrological Sciences*, *372*, 311–314. https://doi.org/10.5194/piahs-372-311-2015

Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., & Crippa, B. (2016). Persistent Scatterer Interferometry: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*, 78–89. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.10.011

Crosetto, M., Monserrat, O., Iglesias, R., & Crippa, B. (2010). Persistent Scatterer Interferometry: Potential, Limits and Initial C- and X-band Comparison. *Photogrammetric Engineering and Remote Sensing*, *76*, 1061–1069. https://doi.org/10.14358/PERS.76.9.1061

Crosetto, M., Solari, L., Balasis-Levinsen, J., Casagli, N., Frei, M., Oyen, A., & Moldestad, D. A. (2020a). GROUND DEFORMATION MONITORING AT CONTINENTAL SCALE: THE EUROPEAN GROUND MOTION SERVICE. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B3-2020*, 293–298. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-293-2020

Crosetto, M., Solari, L., Balasis-Levinsen, J., Casagli, N., Frei, M., Oyen, A., & Moldestad, D. A. (2020b). GROUND DEFORMATION MONITORING AT CONTINENTAL SCALE: THE EUROPEAN GROUND MOTION SERVICE. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B3-2020*, 293–298. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-293-2020

Dal Piaz, G., Cortiana, G., del Moro, A., Martin, S., Pennacchioni, G., & Tartarotti, P. (2001). Tertiary age and paleostructural inferences of the eclogitic imprint in the Austroalpine outliers and Zermatt–Saas ophiolite, western Alps. *International Journal of Earth Sciences*, *90*(3), 668–684. https://doi.org/10.1007/s005310000177

D'Auria, L., Pepe, S., Castaldo, R., Giudicepietro, F., Macedonio, G., Ricciolino, P., Tizzani, P., Casu, F., Lanari, R., Manzo, M., Martini, M., Sansosti, E., & Zinno, I. (2015). Magma injection beneath the urban area of Naples: a new mechanism for the 2012–2013 volcanic unrest at Campi Flegrei caldera. *Scientific Reports*, *5*(1), 13100. https://doi.org/10.1038/srep13100

de Luca, C., Zinno, I., Manunta, M., Lanari, R., & Casu, F. (2017). Large areas surface deformation analysis through a cloud computing P-SBAS approach for massive processing of DInSAR time series. *Remote Sensing of Environment*, *202*, 3–17. https://doi.org/https://doi.org/10.1016/j.rse.2017.05.022

de Zan, F., & Monti Guarnieri, A. (2006). TOPSAR: Terrain Observation by Progressive Scans. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(9), 2352–2360. https://doi.org/10.1109/TGRS.2006.873853

del Soldato, M., Solari, L., Poggi, F., Raspini, F., Tomás, R., Fanti, R., & Casagli, N. (2019). Landslide-Induced Damage Probability Estimation Coupling InSAR and Field Survey Data by Fragility Curves. *Remote Sensing*, *11*(12). https://doi.org/10.3390/rs11121486

del Soldato, M., Solari, L., Raspini, F., Bianchini, S., Ciampalini, A., Montalti, R., Ferretti, A., Pellegrineschi, V., & Casagli, N. (2019). Monitoring Ground Instabilities Using SAR Satellite Data: A Practical Approach. *ISPRS International Journal of Geo-Information*, *8*(7). https://doi.org/10.3390/ijgi8070307

Devanthéry, N., Crosetto, M., Monserrat, O., Cuevas-González, M., & Crippa, B. (2014). An Approach to Persistent Scatterer Interferometry. *Remote Sensing*, *6*(7), 6662–6679. https://doi.org/10.3390/rs6076662

di Martire, D., Paci, M., Confuorto, P., Costabile, S., Guastaferro, F., Verta, A., & Calcaterra, D. (2017). A nation-wide system for landslide mapping and risk management in Italy: The second Not-ordinary Plan of Environmental Remote Sensing. *International Journal of Applied Earth Observation and Geoinformation*, *63*, 143–157. https://doi.org/https://doi.org/10.1016/j.jag.2017.07.018

Ding, C., & He, X. (2004). K-means clustering via principal component analysis. *Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004*, 225 – 232. https://www.scopus.com/inward/record.uri?eid=2-s2.0-14344257496&partnerID=40&md5=09310cc192cf973a0d4a7c2fcbc9e39f

Duro, J., Inglada, J., Closa, J., Adam, N., & Arnaud, A. (2004). High Resolution Differential Interferometry using Time Series of ERS and ENVISAT SAR Data. In H. Lacoste (Ed.), *FRINGE 2003 Workshop* (Vol. 550, p. 72).

Ebmeier, S. K. (2016). Application of independent component analysis to multitemporal InSAR data with volcanic case studies. *Journal of Geophysical Research: Solid Earth*, *121*(12), 8970–8986. https://doi.org/https://doi.org/10.1002/2016JB013765

Emery, W., & Camps, A. (2017). Radar. In *Introduction to Satellite Remote Sensing* (pp. 291–453). Elsevier. https://doi.org/10.1016/B978-0-12-809254-5.00005-1

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 226–231.

Fan, H., Wang, L., Wen, B., & Du, S. (2021). A New Model for three-dimensional Deformation Extraction with Single-track InSAR Based on Mining Subsidence Characteristics. *International Journal of Applied Earth Observation and Geoinformation*, *94*, 102223. https://doi.org/https://doi.org/10.1016/j.jag.2020.102223

Farina, P., Colombo, D., Fumagalli, A., Marks, F., & Moretti, S. (2006). Permanent Scatterers for landslide investigations: outcomes from the ESA-SLAM project. *Engineering Geology*, *88*(3), 200–217. https://doi.org/https://doi.org/10.1016/j.enggeo.2006.09.007

Fay, M. P., & Proschan, M. A. (2010). Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys*, *4*(none). https://doi.org/10.1214/09-SS051

Fedele, L., Insinga, D. D., Calvert, A. T., Morra, V., Perrotta, A., & Scarpati, C. (2011). 40Ar/39Ar dating of tuff vents in the Campi Flegrei caldera (southern Italy): toward a new chronostratigraphic reconstruction of the Holocene volcanic activity. *Bulletin of Volcanology*, *73*(9), 1323–1336. https://doi.org/10.1007/s00445-011-0478-8

Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., & Rucci, A. (2011). A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(9), 3460–3470. https://doi.org/10.1109/TGRS.2011.2124465

Ferretti, A., Prati, C., & Rocca, F. (1999). Permanent scatterers in SAR interferometry. *IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No.99CH36293)*, *3*, 1528–1530 vol.3. https://doi.org/10.1109/IGARSS.1999.772008

Ferretti, A., Prati, C., & Rocca, F. (2000). Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(5), 2202–2212. https://doi.org/10.1109/36.868878

Festa, D., Bonano, M., Casagli, N., Confuorto, P., de Luca, C., del Soldato, M., Lanari, R., Lu, P., Manunta, M., Manzo, M., Onorato, G., Raspini, F., Zinno, I., & Casu, F. (2022). Nation-wide mapping and classification of ground deformation phenomena through the spatial clustering of P-SBAS InSAR measurements: Italy case study. *ISPRS Journal of Photogrammetry and Remote Sensing*, *189*, 1–22. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2022.04.022

Fioraso, G., Tararbra, M., Negro, N., & Quaternario, I. (2020). *La deformazione gravitativa profonda di Sauze d'Oulx (Alpi Occidentali): analisi dei meccanismi defromativi in relazione all'evoluzione trado quaternaria dell'alta Valle di Susa*. *23*, 37–54.

Fortunato, C., Martino, S., Prestininzi, A., & Romeo, R. W. (2012). New release of the Italian catalogue of earthquake-induced ground failures (CEDIT). *Italian Journal of Engineering Geology and Environment*, *2*, 63–74. https://doi.org/10.4408/IJEGE.2012-02.O-05

Franceschetti, G., Migliaccio, M., Riccio, D., & Schirinzi, G. (1992). SARAS: a synthetic aperture radar (SAR) raw signal simulator. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(1), 110–123. https://doi.org/10.1109/36.124221

Frattini, P., Crosta, G. B., Rossini, M., & Allievi, J. (2018). Activity and kinematic behaviour of deep-seated landslides from PS-InSAR displacement rate measurements. *Landslides*, *15*(6), 1053–1070. https://doi.org/10.1007/s10346-017-0940-6

Frayssines, M., & Hantz, D. (2006). Failure mechanisms and triggering factors in calcareous cliffs of the Subalpine Ranges (French Alps). *Engineering Geology*, *86*(4), 256–270. https://doi.org/https://doi.org/10.1016/j.enggeo.2006.05.009

Gabriel, A. K., Goldstein, R. M., & Zebker, H. A. (1989). Mapping small elevation changes over large areas: Differential radar interferometry. *Journal of Geophysical Research: Solid Earth*, *94*(B7), 9183–9191. https://doi.org/https://doi.org/10.1029/JB094iB07p09183

Gagnon, L., & Jouan, A. (1997). Speckle Filtering of SAR Images - A Comparative Study Between Complex-Wavelet-Based and Standard Filters. *Proceedings of SPIE - The International Society for Optical Engineering*, *3169*. https://doi.org/10.1117/12.279681

Genuer, R., Poggi, J.-M., Tuleau-Malot, C., & Villa-Vialaneix, N. (2017). Random Forests for Big Data. *Big Data Research*, *9*, 28–46. https://doi.org/https://doi.org/10.1016/j.bdr.2017.07.003

Gernhardt, S., Adam, N., Eineder, M., & Bamler, R. (2010). Potential of very high resolution SAR for persistent scatterer interferometry in urban areas. *Annals of GIS*, *16*(2), 103–111. https://doi.org/10.1080/19475683.2010.492126

(Gini) Ketelaar, V. B. H. (2009). *Satellite Radar Interferometry* (Vol. 14). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9428-6

Goel, K., & Adam, N. (2014). A Distributed Scatterer Interferometry Approach for Precision Monitoring of Known Surface Deformation Phenomena. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(9), 5454–5468. https://doi.org/10.1109/TGRS.2013.2289370

Gomes, V. C. F., Queiroz, G. R., & Ferreira, K. R. (2020). An Overview of Platforms for Big Earth Observation Data Management and Analysis. *Remote Sensing*, *12*(8). https://doi.org/10.3390/rs12081253

Guarnieri, A. M., & Prati, C. (1996). ScanSAR focusing and interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *34*(4), 1029–1038. https://doi.org/10.1109/36.508420

Guo, H.-D., Zhang, L., & Zhu, L.-W. (2015). Earth observation big data for climate change research. *Advances in Climate Change Research*, *6*(2), 108–117. https://doi.org/https://doi.org/10.1016/j.accre.2015.09.007

Halko, N., Martinsson, P. G., & Tropp, J. A. (2011). Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions. *SIAM Review*, *53*(2), 217–288. https://doi.org/10.1137/090771806

Hanssen, R. F. (2001). *Radar Interferometry* (Vol. 2). Springer Netherlands. https://doi.org/10.1007/0-306-47633-9

Herrera, G., Tomás, R., Lopez-Sanchez, J. M., Delgado, J., Mallorqui, J. J., Duque, S., & Mulas, J. (2007). Advanced DInSAR analysis on mining areas: La Union case study (Murcia, SE Spain). *Engineering Geology*, *90*(3), 148–159. https://doi.org/https://doi.org/10.1016/j.enggeo.2007.01.001

Hetland, E. A., Musé, P., Simons, M., Lin, Y. N., Agram, P. S., & DiCaprio, C. J. (2012). Multiscale InSAR Time Series (MInTS) analysis of surface deformation. *Journal of Geophysical Research: Solid Earth*, *117*(B2), n/a-n/a. https://doi.org/10.1029/2011JB008731

Hooper, A. (2006). *Persistent scatter radar interferometry for crustal deformation studies and modeling of volcanic deformation*.

Hooper, A., Zebker, H., Segall, P., & Kampes, B. (2004a). A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers. *Geophysical Research Letters*, *31*(23). https://doi.org/https://doi.org/10.1029/2004GL021737

Hooper, A., Zebker, H., Segall, P., & Kampes, B. (2004b). A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers. *Geophysical Research Letters*, *31*(23). https://doi.org/https://doi.org/10.1029/2004GL021737

Huang, M.-H., Fielding, E. J., Liang, C., Milillo, P., Bekaert, D., Dreger, D., & Salzer, J. (2017). Coseismic deformation and triggered landslides of the 2016 Mw 6.2 Amatrice earthquake in Italy. *Geophysical Research Letters*, *44*(3), 1266–1274. https://doi.org/https://doi.org/10.1002/2016GL071687

Hussain, E., Novellino, A., Jordan, C., & Bateson, L. (2021). Offline-Online Change Detection for Sentinel-1 InSAR Time Series. *Remote Sensing*, *13*(9). https://doi.org/10.3390/rs13091656

Improta, L., Latorre, D., Margheriti, L., Nardi, A., Marchetti, A., Lombardi, A. M., Castello, B., Villani, F., Ciaccio, M. G., Mele, F. M., Moretti, M., Battelli, P., Berardi, M., Castellano, C., Melorio, C., Modica, G., Pirro, M., Rossi, A., Thermes, C., … Group, T. B. S. I. W. (2019a). Multi-segment rupture of the 2016 Amatrice-Visso-Norcia seismic sequence (central Italy) constrained by the first high-quality catalog of Early Aftershocks. *Scientific Reports*, *9*(1), 6921. https://doi.org/10.1038/s41598-019-43393-2

Improta, L., Latorre, D., Margheriti, L., Nardi, A., Marchetti, A., Lombardi, A. M., Castello, B., Villani, F., Ciaccio, M. G., Mele, F. M., Moretti, M., Battelli, P., Berardi, M., Castellano, C., Melorio, C., Modica, G., Pirro, M., Rossi, A., Thermes, C., … Group, T. B. S. I. W. (2019b). Multi-segment rupture of the 2016 Amatrice-Visso-Norcia seismic sequence (central Italy) constrained by the first high-quality catalog of Early Aftershocks. *Scientific Reports*, *9*(1), 6921. https://doi.org/10.1038/s41598-019-43393-2

Iwahashi, J., Watanabe, S., & Furuya, T. (2003). Mean slope-angle frequency distribution and size frequency distribution of landslide masses in Higashikubiki area, Japan. *Geomorphology*, *50*(4), 349–364. https://doi.org/https://doi.org/10.1016/S0169-555X(02)00222-2

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, *6*(1), 27. https://doi.org/10.1186/s40537-019-0192-5

Jolliffe, I. T., & Cadima, J. (2016a). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Jolliffe, I. T., & Cadima, J. (2016b). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Kalia, A. C., Frei, M., & Lege, T. (2017). A Copernicus downstream-service for the nationwide monitoring of surface displacements in Germany. *Remote Sensing of Environment*, *202*, 234–249. https://doi.org/https://doi.org/10.1016/j.rse.2017.05.015

Kampes, B. (2006). *Radar interferometry*. Springer Netherlands. https://doi.org/10.1007/978-1-4020-4723-7

Klinger, Y., Michel, R., & Avouac, J.-P. (2000). Co-seismic deformation during the Mw7.3 Aqaba Earthquake (1995) from ERS-SAR interferometry. *Geophysical Research Letters*, *27*(22), 3651–3654. https://doi.org/https://doi.org/10.1029/1999GL008463

Kobayashi, T., Tobita, M., Koarai, M., Okatani, T., Suzuki, A., Noguchi, Y., Yamanaka, M., & Miyahara, B. (2013). InSAR-derived crustal deformation and fault models of normal faulting earthquake (Mj 7.0) in the Fukushima-Hamadori area. *Earth, Planets and Space*, *64*(12), 15. https://doi.org/10.5047/eps.2012.08.015

Lanari, R., Bonano, M., Casu, F., Luca, C. de, Manunta, M., Manzo, M., Onorato, G., & Zinno, I. (2020). Automatic Generation of Sentinel-1 Continental Scale DInSAR Deformation Time Series through an Extended P-SBAS Processing Pipeline in a Cloud Computing Environment. *Remote Sensing*, *12*(18). https://doi.org/10.3390/rs12182961

Lanari, R., Lundgren, P., & Sansosti, E. (1998). Dynamic deformation of Etna Volcano observed by satellite radar interferometry. *Geophysical Research Letters*, *25*(10), 1541–1544. https://doi.org/https://doi.org/10.1029/98GL00642

Lavecchia, G., Castaldo, R., Nardis, R., de Novellis, V., Ferrarini, F., Pepe, S., Brozzetti, F., Solaro, G., Cirillo, D., Bonano, M., Boncio, P., Casu, F., de Luca, C., Lanari, R., Manunta, M., Manzo, M., Pepe, A., Zinno, I., & Tizzani, P. (2016). Ground deformation and source geometry of the 24 August 2016 Amatrice earthquake (Central Italy) investigated through analytical and numerical modeling of DInSAR measurements and structural-geological data. *Geophysical Research Letters*, *43*(24). https://doi.org/10.1002/2016GL071723

Lee, S., Ryu, J.-H., Won, J.-S., & Park, H.-J. (2004). Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Engineering Geology*, *71*(3), 289–302. https://doi.org/https://doi.org/10.1016/S0013-7952(03)00142-X

Lischeid, G. (2009). Non-linear visualization and analysis of large water quality data sets: A model-free basis for efficient monitoring and risk assessment. *Stochastic Environmental Research and Risk Assessment*, *23*(7), 977 – 990. https://doi.org/10.1007/s00477-008-0266-y

Lu, P., Casagli, N., Catani, F., & Tofani, V. (2012). Persistent Scatterers Interferometry Hotspot and Cluster Analysis (PSI-HCA) for detection of extremely slow-moving landslides. *International Journal of Remote Sensing*, *33*(2), 466–489. https://doi.org/10.1080/01431161.2010.536185

Lu, P., Catani, F., Tofani, V., & Casagli, N. (2014). Quantitative hazard and risk assessment for slow-moving landslides from Persistent Scatterer Interferometry. *Landslides*, *11*(4), 685–696. https://doi.org/10.1007/s10346-013-0432-2

Lundgren, P., & Rosen, P. A. (2003). Source model for the 2001 flank eruption of Mt. Etna volcano. *Geophysical Research Letters*, *30*(7). https://doi.org/https://doi.org/10.1029/2002GL016774

Lv, X., Yazıcı, B., Zeghal, M., Bennett, V., & Abdoun, T. (2014). Joint-Scatterer Processing for Time-Series InSAR. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(11), 7205–7221. https://doi.org/10.1109/TGRS.2014.2309346

Manconi, A., Walter, T. R., Manzo, M., Zeni, G., Tizzani, P., Sansosti, E., & Lanari, R. (2010). On the effects of 3-D mechanical heterogeneities at Campi Flegrei caldera, southern Italy. *Journal of Geophysical Research: Solid Earth*, *115*(B8). https://doi.org/https://doi.org/10.1029/2009JB007099

Manunta, M., de Luca, C., Zinno, I., Casu, F., Manzo, M., Bonano, M., Fusco, A., Pepe, A., Onorato, G., Berardino, P., de Martino, P., & Lanari, R. (2019). The Parallel SBAS Approach for Sentinel-1 Interferometric Wide Swath Deformation Time-Series Generation: Algorithm Description and Products Quality Assessment. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(9), 6259–6281. https://doi.org/10.1109/TGRS.2019.2904912

Manzo, M., Riccardo, L., Antoncecchi, I., Solaro, G., Bonano, M., Buonanno, S., Castaldo, R., Casu, F., Ciccone, F., Luca, C., de Novellis, V., Manunta, M., Onorato, G., Pepe, S., Rossi, G., Tizzani, P., Zeni, G., Zinno, I., & Franco, T. (2019, October). *Ground deformation analysis of the Italian Peninsula through space-borne SAR interferometry and geophysical modelling: the IREA-CNR/MiSE-DGS-UNMIG agreement*.

Manzo, M., Ricciardi, G. P., Casu, F., Ventura, G., Zeni, G., Borgström, S., Berardino, P., del Gaudio, C., & Lanari, R. (2006). Surface deformation analysis in the Ischia Island (Italy) based on spaceborne radar interferometry. *Journal of Volcanology and Geothermal Research*, *151*(4), 399–416. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2005.09.010

Massonnet, D., & Feigl, K. L. (1998). Radar interferometry and its application to changes in the Earth's surface. *Reviews of Geophysics*, *36*(4), 441–500. https://doi.org/https://doi.org/10.1029/97RG03139

Matano, F. (2019). Analysis and Classification of Natural and Human-Induced Ground Deformations at Regional Scale (Campania, Italy) Detected by Satellite Synthetic-Aperture Radar Interferometry Archive Datasets. *Remote Sensing*, *11*(23). https://doi.org/10.3390/rs11232822

Meisina, C., Zucca, F., Notti, D., Colombo, A., Cucchi, A., Savio, G., Giannico, C., & Bianchi, M. (2008). Geological Interpretation of PSInSAR Data at Regional Scale. *Sensors*, *8*(11), 7469–7492. https://doi.org/10.3390/s8117469

Meta, A., Mittermayer, J., Prats-Iraola, P., Scheiber, R., & Steinbrecher, U. (2010). TOPS imaging with TerraSAR-X: Mode design and performance analysis. *Geoscience and Remote Sensing, IEEE Transactions On*, *48*, 759–769. https://doi.org/10.1109/TGRS.2009.2026743

Mirmazloumi, S. M., Wassie, Y., Navarro, J. A., Palamà, R., Krishnakumar, V., Barra, A., Cuevas-González, M., Crosetto, M., & Monserrat, O. (2022). Classification of ground deformation using sentinel-1 persistent scatterer interferometry time series. *GIScience & Remote Sensing*, *59*(1), 374–392. https://doi.org/10.1080/15481603.2022.2030535

Montalti, R., Solari, L., Bianchini, S., Soldato, M. del, Raspini, F., & Casagli, N. (2019). A Sentinel-1-based clustering analysis for geo-hazards mitigation at regional scale: a case study in Central Italy. *Geomatics, Natural Hazards and Risk*, *10*(1), 2257–2275. https://doi.org/10.1080/19475705.2019.1690058

Mora, O., Mallorqui, J. J., & Broquetas, A. (2003). Linear and nonlinear terrain deformation maps from a reduced set of interferometric sar images. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(10), 2243–2253. https://doi.org/10.1109/TGRS.2003.814657

Ng, A. H.-M., Ge, L., Zhang, K., Chang, H.-C., Li, X., Rizos, C., & Omura, M. (2011). Deformation mapping in three dimensions for underground mining using InSAR – Southern highland coalfield in New South Wales, Australia. *International Journal of Remote Sensing*, *32*(22), 7227–7256. https://doi.org/10.1080/01431161.2010.519741

Notti, D., Calò, F., Cigna, F., Manunta, M., Herrera, G., Berti, M., Meisina, C., Tapete, D., & Zucca, F. (2015). A User-Oriented Methodology for DInSAR Time Series Analysis and Interpretation: Landslides and Subsidence Case Studies. *Pure and Applied Geophysics*, *172*(11), 3081–3105. https://doi.org/10.1007/s00024-015-1071-4

Notti, D., Herrera, G., Bianchini, S., Meisina, C., García-Davalillo, J. C., & Zucca, F. (2014). A methodology for improving landslide PSI data analysis. *International Journal of Remote Sensing*, *35*(6), 2186–2214. https://doi.org/10.1080/01431161.2014.889864

Notti, D., Meisina, C., Zucca, F., Alessio, C., Anselmo, C., & Poggi, F. (2012). *Pro and Cons of non-linear SqueeSAR Time Series compared with other monitoring systems in landslides study*. 8599.

Novellino, A., Cesarano, M., Cappelletti, P., di Martire, D., di Napoli, M., Ramondini, M., Sowter, A., & Calcaterra, D. (2021). Slow-moving landslide risk assessment combining Machine Learning and InSAR techniques. *CATENA*, *203*, 105317. https://doi.org/https://doi.org/10.1016/j.catena.2021.105317

Novellino, A., Terrington, R., Christodoulou, V., Smith, H., & Bateson, L. (2019). *Ground motion and stratum thickness comparison in Tower Hamlets, London*.

Paranunzio, R., Chiarle, M., Laio, F., Nigrelli, G., Turconi, L., & Luino, F. (2019). New insights in the relation between climate and slope failures at high-elevation sites. *Theoretical and Applied Climatology*, *137*(3), 1765–1784. https://doi.org/10.1007/s00704-018-2673-4

Patra, A., Bevilacqua, A., Pitman, E. B., Bursik, M., Voight, B., Neri, A., Macedonio, G., Flandoli, F., de Martino, P., Giudicepietro, F., & Vitale, S. (2019, October). *A statistical approach for spatial mapping and temporal forecasts of volcanic eruptions using monitoring data*.

Peduto, D., Cascini, L., Arena, L., Ferlisi, S., Fornaro, G., & Reale, D. (2015). A general framework and related procedures for multiscale analyses of DInSAR data in subsiding urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 186–210. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.04.001

Peltzer, G., & Rosen, P. (1995). Surface Displacement of the 17 May 1993 Eureka Valley, California, Earthquake Observed by SAR Interferometry. *Science*, *268*(5215), 1333–1336. https://doi.org/10.1126/science.268.5215.1333

Pepe, A., & Lanari, R. (2006). On the Extension of the Minimum Cost Flow Algorithm for Phase Unwrapping of Multitemporal Differential SAR Interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(9), 2374–2383. https://doi.org/10.1109/TGRS.2006.873207

Pepe, A., Yang, Y., Manzo, M., & Lanari, R. (2015). Improved EMCF-SBAS Processing Chain Based on Advanced Techniques for the Noise-Filtering and Selection of Small Baseline Multi-Look DInSAR Interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(8), 4394–4417. https://doi.org/10.1109/TGRS.2015.2396875

Perissin, D., & Wang, T. (2012). Repeat-Pass SAR Interferometry With Partially Coherent Targets. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(1), 271–280. https://doi.org/10.1109/TGRS.2011.2160644

Peyret, M., Djamour, Y., Rizza, M., Ritz, J.-F., Hurtrez, J.-E., Goudarzi, M. A., Nankali, H., Chéry, J., le Dortz, K., & Uri, F. (2008). Monitoring of the large slow Kahrod landslide in Alborz mountain range (Iran) by GPS and SAR interferometry. *Engineering Geology*, *100*(3), 131–141. https://doi.org/https://doi.org/10.1016/j.enggeo.2008.02.013

Prentice, I. C. (1982). Multivariate methods for the presentation and analysis of data. In *Palaeohydrological changes in the temperate zone in the last 15 000 years. IGCP 158 B. Lake and mire environments. Project guide III,*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0020342606&partnerID=40&md5=362cd61065f78ec10cf1e852b5f4ae3b

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press Inc.

Probst, L., Pedersen, B., & Dakkak-Arnoux, L. (2017). *Big Data in Earth Observation*.

Qin, Y., Hoppe, E., & Perissin, D. (2020). Slope Hazard Monitoring Using High-Resolution Satellite Remote Sensing: Lessons Learned From a Case Study. *ISPRS International Journal of Geo-Information*, *9*, 131. https://doi.org/10.3390/ijgi9020131

Raspini, F., Bianchini, S., Ciampalini, A., del Soldato, M., Solari, L., Novali, F., del Conte, S., Rucci, A., Ferretti, A., & Casagli, N. (2018). Continuous, semi-automatic monitoring of ground deformation using Sentinel-1 satellites. *Scientific Reports*, *8*(1), 7253. https://doi.org/10.1038/s41598-018-25369-w

Raspini, F., Caleca, F., del Soldato, M., Festa, D., Confuorto, P., & Bianchini, S. (2022). Review of satellite radar interferometry for subsidence analysis. *Earth-Science Reviews*, *235*, 104239. https://doi.org/10.1016/j.earscirev.2022.104239

Reyes-Carmona, C., Barra, A., Galve, J. P., Monserrat, O., Pérez-Peña, J. V., Mateos, R. M., Notti, D., Ruano, P., Millares, A., López-Vinielles, J., & Azañón, J. M. (2020). Sentinel-1 DInSAR for Monitoring Active Landslides in Critical Infrastructures: The Case of the Rules Reservoir (Southern Spain). *Remote Sensing*, *12*(5). https://doi.org/10.3390/rs12050809

Rizzo, V. (2002). GPS monitoring and new data on slope movements in the Maratea Valley (Potenza, Basilicata). *Physics and Chemistry of the Earth, Parts A/B/C*, *27*(36), 1535–1544. https://doi.org/https://doi.org/10.1016/S1474-7065(02)00174-2

Rizzo, V., & Limongi, P. (1997). Risultati inclinometrici ed indagini geologico-stratigrafiche nel Centro Storico di Maratea (Lucania, Italia). *Geografia Fisica e Dinamica Quaternaria*, *20*, 137–144.

Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodriguez, E., & Goldstein, R. M. (2000). Synthetic aperture radar interferometry. *Proceedings of the IEEE*, *88*(3), 333–382. https://doi.org/10.1109/5.838084

Sabeur, Z., TABASCO, A., Zissis, D., Stevenot, B., SERBAN, F., Rodríguez-Aubó, N., GONZÁLEZ, A., PISCITELLI, R., Ilie, C., GUTIERRES, F., GONZALO-MARTÍN, C., PEDRERO, A., GALE, L., COSTEA, A., Chatzikokolakis, K., Berre, A., & ÁLVAREZ-CID, M. (2018). *Whitepaper Big Data in Earth Observation Big Data Value Association TF7-SG5: Earth Observation and Geospatial*. https://doi.org/10.13140/RG.2.2.28253.84969

Samsonov, S., d'Oreye, N., & Smets, B. (2013). Ground deformation associated with post-mining activity at the French–German border revealed by novel InSAR time series method. *International Journal of Applied Earth Observation and Geoinformation*, *23*, 142–154. https://doi.org/https://doi.org/10.1016/j.jag.2012.12.008

Samsonov, S., & Tiampo, K. (2011). Polarization Phase Difference Analysis for Selection of Persistent Scatterers in SAR Interferometry. *IEEE Geoscience and Remote Sensing Letters*, *8*(2), 331–335. https://doi.org/10.1109/LGRS.2010.2072904

Santos, F. (2020). Modern methods for old data: An overview of some robust methods for outliers detection with applications in osteology. *Journal of Archaeological Science: Reports*, *32*, 102423. https://doi.org/https://doi.org/10.1016/j.jasrep.2020.102423

Satopaa, V., Albrecht, J., Irwin, D., & Raghavan, B. (2011). Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior. *2011 31st International Conference on Distributed Computing Systems Workshops*, 166–171. https://doi.org/10.1109/ICDCSW.2011.20

Schmidt, D. A., & Bürgmann, R. (2003). Time-dependent land uplift and subsidence in the Santa Clara valley, California, from a large interferometric synthetic aperture radar data set. *Journal of Geophysical Research: Solid Earth*, *108*(B9). https://doi.org/10.1029/2002JB002267

Scornet, E., Biau, G., & Vert, J.-P. (2015). Consistency of random forests. *The Annals of Statistics*, *43*(4), 1716 – 1741. https://doi.org/10.1214/15-AOS1321

Shirzaei, M., & Walter, T. R. (2011). Estimating the Effect of Satellite Orbital Error Using Wavelet-Based Robust Regression Applied to InSAR Deformation Data. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(11), 4600–4605. https://doi.org/10.1109/TGRS.2011.2143419

Solari, L., del Soldato, M., Bianchini, S., Ciampalini, A., Ezquerro, P., Montalti, R., Raspini, F., & Moretti, S. (2018). From ERS 1/2 to Sentinel-1: Subsidence Monitoring in Italy in the Last Two Decades. *Frontiers in Earth Science*, *6*. https://doi.org/10.3389/feart.2018.00149

Solari, L., Soldato, M. del, Montalti, R., Bianchini, S., Raspini, F., Thuegaz, P., Bertolo, D., Tofani, V., & Casagli, N. (2019). A Sentinel-1 based hot-spot analysis: landslide mapping in north-western Italy. *International Journal of Remote Sensing*, *40*(20), 7898–7921. https://doi.org/10.1080/01431161.2019.1607612

Spreckels, V., Wegmüller, U., Strozzi, T., Musiedlak, J., & Wichlacz, H.-C. (2001). *DETECTION AND OBSERVATION OF UNDERGROUND COAL MINING-INDUCED SURFACE DEFORMATION WITH DIFFERENTIAL SAR INTERFEROMETRY*.

Squarzoni, C., Delacourt, C., & Allemand, P. (2003). Nine years of spatial and temporal evolution of the La Valette landslide observed by SAR interferometry. *Engineering Geology*, *68*(1), 53–66. https://doi.org/https://doi.org/10.1016/S0013-7952(02)00198-9

Stramondo, S., Saroli, M., Tolomei, C., Moro, M., Doumaz, F., Pesci, A., Loddo, F., Baldi, P., & Boschi, E. (2007). Surface movements in Bologna (Po Plain — Italy) detected by multitemporal DInSAR. *Remote Sensing of Environment*, *110*(3), 304–316. https://doi.org/https://doi.org/10.1016/j.rse.2007.02.023

Stromann, O., Nascetti, A., Yousif, O., & Ban, Y. (2020). Dimensionality Reduction and Feature Selection for Object-Based Land Cover Classification based on Sentinel-1 and Sentinel-2 Time Series Using Google Earth Engine. *Remote Sensing*, *12*(1). https://doi.org/10.3390/rs12010076

Sudmanns, M., Tiede, D., Lang, S., Bergstedt, H., Trost, G., Augustin, H., Baraldi, A., & Blaschke, T. (2020). Big Earth data: disruptive changes in Earth observation data management and analysis? *International Journal of Digital Earth*, *13*(7), 832–850. https://doi.org/10.1080/17538947.2019.1585976

Tarquini, S., & Nannipieri, L. (2016). The 10m-resolution TINITALY DEM as a trans-disciplinary basis for the analysis of the Italian territory: Current trends and new perspectives. *Geomorphology*, *281*. https://doi.org/10.1016/j.geomorph.2016.12.022

Tatem, A. J., Goetz, S. J., & Hay, S. I. (2008). Fifty Years of Earth-observation Satellites. *American Scientist*, 390–398.

Tomás, R., Pagán, J. I., Navarro, J. A., Cano, M., Pastor, J. L., Riquelme, A., Cuevas-González, M., Crosetto, M., Barra, A., Monserrat, O., Lopez-Sanchez, J. M., Ramón, A., Ivorra, S., del Soldato, M., Solari, L., Bianchini, S., Raspini, F., Novali, F., Ferretti, A., … Casagli, N. (2019). Semi-Automatic Identification and Pre-Screening of Geological–Geotechnical Deformational Processes Using Persistent Scatterer Interferometry Datasets. *Remote Sensing*, *11*(14). https://doi.org/10.3390/rs11141675

Trigila, A., Iadanza, C., & Guerrieri, L. (2007). The IFFI project (Italian landslide inventory): Methodology and results. In *Guidelines for Mapping Areas at Risk of Landslides in Europe, edited by: Hervás, J., ISPRA, Rome, Italy*.

Trigila, A., Iadanza, C., & Spizzichino, D. (2010). Quality assessment of the Italian Landslide Inventory using GIS processing. *Landslides*, *7*(4), 455–470. https://doi.org/10.1007/s10346-010-0213-0

Werner, C., Wegmuller, U., Strozzi, T., & Wiesmann, A. (2003). Interferometric point target analysis for deformation mapping. *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*, *7*, 4362–4364 vol.7. https://doi.org/10.1109/IGARSS.2003.1295516

Wright, T. J., Parsons, B. E., Jackson, J. A., Haynes, M., Fielding, E. J., England, P. C., & Clarke, P. J. (1999). Source parameters of the 1 October 1995 Dinar (Turkey) earthquake from SAR interferometry and seismic bodywave modelling. *Earth and Planetary Science Letters*, *172*(1), 23–37. https://doi.org/https://doi.org/10.1016/S0012-821X(99)00186-7

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z.-H., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37. https://doi.org/10.1007/s10115-007-0114-2

Yagüe-Martínez, N., Prats-Iraola, P., Rodríguez González, F., Brcic, R., Shau, R., Geudtner, D., Eineder, M., & Bamler, R. (2016). Interferometric Processing of Sentinel-1 TOPS Data. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(4), 2220–2234. https://doi.org/10.1109/TGRS.2015.2497902

Zhao, J., & Konietzky, H. (2020). Numerical analysis and prediction of ground surface movement induced by coal mining and subsequent groundwater flooding. *International Journal of Coal Geology*, *229*, 103565. https://doi.org/https://doi.org/10.1016/j.coal.2020.103565

Zinno, I., Elefante, S., Mossucca, L., de Luca, C., Manunta, M., Terzo, O., Lanari, R., & Casu, F. (2015). A First Assessment of the P-SBAS DInSAR Algorithm Performances Within a Cloud Computing Environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(10), 4675–4686. https://doi.org/10.1109/JSTARS.2015.2426054