Alma Mater Studiorum — Università di Bologna

Dipartimento di Fisica e Astronomia "Augusto Righi"
Laurea Magistrale in Fisica del Sistema Terra

Characterization of ground deformation related with hydrological processes in the northwestern Po Plain by InSAR data

Presentata da:
Daniele Guidi

Relatrice:
Prof.ssa Eleonora Rivalta

Correlatrice:
Dott.ssa Francesca Silverii

Anno Accademico 2024 - 2025
Appello II

Contents

Chapter 1

Introduction

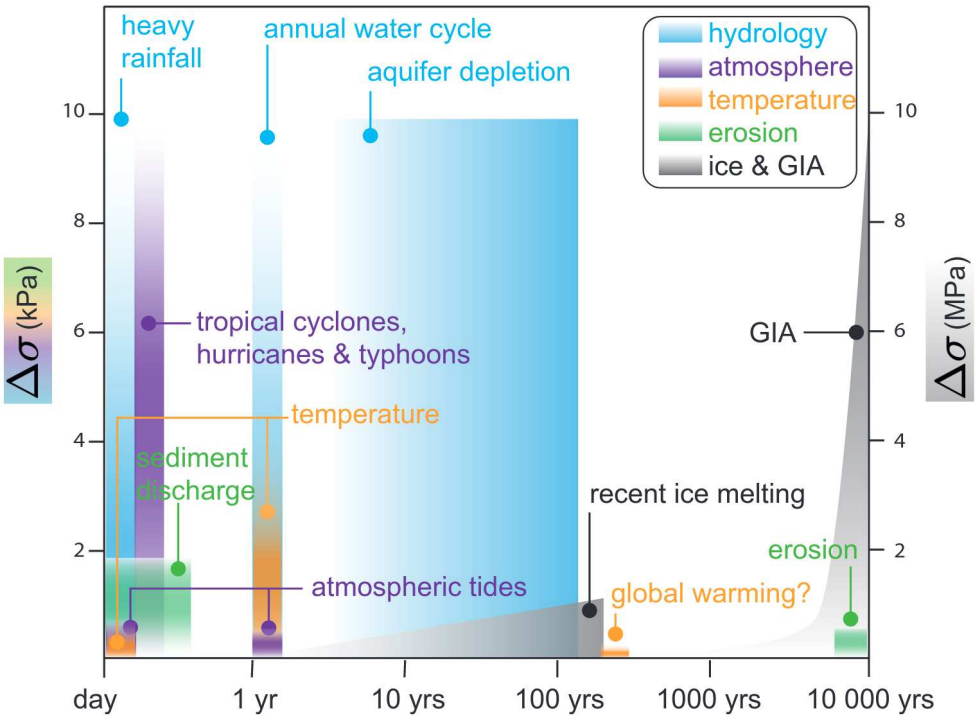
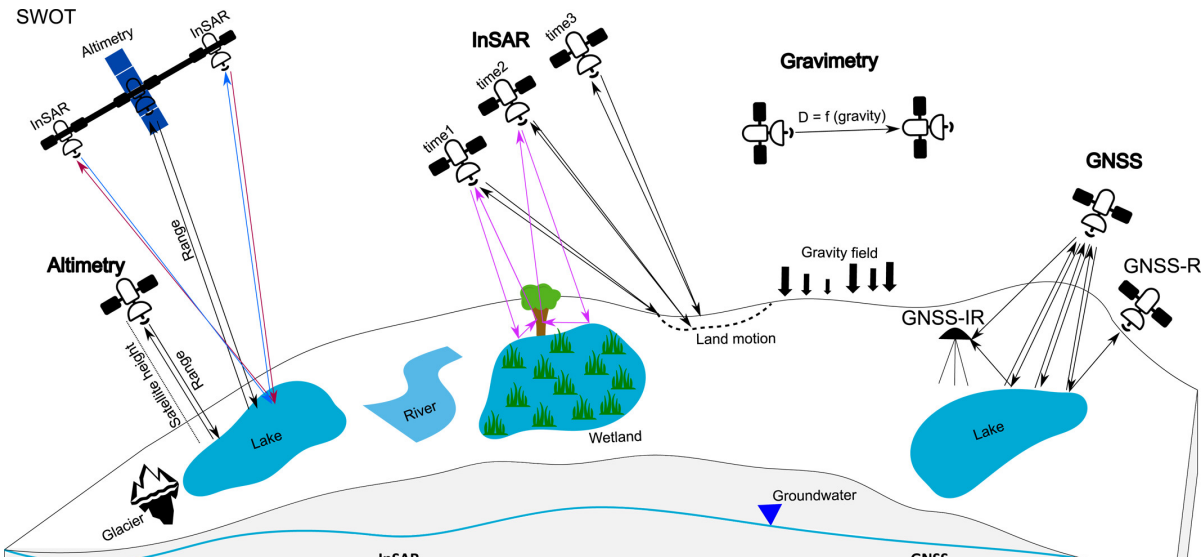


Figure 1.1: Timescales of climate-driven surface processes and associated stress variations. From Burgmann et al. 2024

The Earth's surface represents the interface of the lithosphere with the atmosphere, the hydrosphere and the cryosphere and, as such, its dynamics reflects the interaction processes among these components [Burgmann et al. 2024]. These processes can produce measurable (up to several cm) deformation of the Earth surface at a wide range of



	Altimetry	InSAR		Gravimetry	GNSS		
		Repeat pass	Single pass (SWOT)		GNSS	GNSS-IR	GNSS-R
Hydrologic parameter	Surface water level	Surface water level and water storage change	Surface water level	Water storage change	Water storage change	Surface water level change	Surface elevation, soil moisture
Measuring principle	2-way travel time of radar/laser	Difference in 2-way radar travel time between multiple acquisitions	Difference in 2-way radar travel time between simultaneous acquisitions	Precise orbit determination	Correlation of satellite and receiver signals	Difference in signal-to-noise ratio from 2 GNSS signal paths	Reflected GNSS's signal travel time and backscatter
Temporal, spatial Resolution	~10- 35 days, 50-300 m (along-track)	≥6 days, 10-100 m	≤21 days, 95% of all water bodies	monthly, >300 km)	Daily, Point measurements	Seconds, 100m-100km	Seconds, 100m-100km

Figure 1.2: Summary of geodetic techniques and their use in hydrological monitoring. The table reports the key hydrological parameters, the physical principles exploited for their detection, and the typical spatial and temporal scales of the measurements. [Jaramillo et al. 2024]

spatial (local to global) and temporal (seasonal to multi-year) time scales (Figure 1.1). Examples are surface deformation resulting from climate-driven variations in surface mass load (e.g. terrestrial water storage, snow) and groundwater storage (e.g, aquifer withdrawal/recharge). Since the past century, advances in measurement systems—particularly the development of satellite geodesy techniques—have significantly improved the spatio-temporal resolution and accuracy of surface deformation measurements. In parallel, our ability to analyze the increasing volume of deformation data has improved, both through advances in the physical modeling of deformation processes and in the development of statistical approaches aimed at identifying relevant deformation signals within the data.

An increasing number of studies focus on deformation processes associated with the redistribution of water associated within the hydrological cycle. The use of geodetic techniques—primarily Global Navigation Satellite System (GNSS), Gravity Recovery and Climate Experiment (GRACE), and Interferometric Synthetic Aperture Radar (InSAR)—to characterize water distribution and storage is referred to as hydrogeodesy

[Jaramillo et al. 2024](Figure 1.2).

These processes are of interest for several reasons. On one hand, we aim to isolate and model the hydrological contribution in geodetic time series as a noise component, in order to better identify other geophysical signals. On the other hand, we aim to extract information on geological and hydrological variables by studying the associated surface deformation. In fact, by measuring variations in the Earth's surface and gravity field, it is possible to infer changes in hydrological quantities independently of traditional hydrological methods such as groundwater level measurements from piezometers or river discharge data from streamflow gauges. When the hydrological forcing is known, the study of the resulting deformation processes can provide estimates of the terrain mechanical parameters [Jaramillo et al. 2024].

The Po Plain is a lowland region in northern Italy that hosts abundant freshwater resources stored within sediment layers. The variable geological structure and topography give rise to a complex hydrological system that includes both surface and groundwater reservoirs [De Donatis et al. 2007,De Luca et al. 2020]. Measurable surface deformation has been associated with water resource variations within these reservoirs, resulting from both natural processes and human extraction activities [Carminati et al. 2002, Farías et al. 2024, Pintori et al. 2024].

The strong socio-economic relevance of the region further increases the importance of characterizing these deformation processes. This is essential for assessing related risks — such as damage to infrastructure and increased flood hazard due to subsidence — as well as for complementing traditional investigation methods in understanding the variability of water dynamics and distribution, ultimately supporting informed water management decisions.

In this work, I analyzed InSAR-derived surface deformation data provided by the European Ground Motion Service (EGMS) over the Po Plain to investigate possible effects of hydrological processes. The EGMS service offers freely available deformation time series with a spatial resolution of $100\,\mathrm{m} \times 100\,\mathrm{m}$ and a temporal sampling of 6 days.

I analyzed the dataset relative to the interval 2018-2022. This period was marked by significant climatic variability, including a severe drought event. This is reflected in below-average precipitation across the region and record-low water levels in the Po River — indicative of a major decline in terrestrial water storage [Baronetti et al. 2020, Montanari et al. 2023]. This loss was also accompanied by a regional-scale elastic uplift, observed in GNSS time series by Pintori et al. 2024. The uplifting response was attributed to the lithosphere's elastic deformation due to the decrease in surface load associated with Terrestrial Water Storage (TWS). Terrestrial water storage refers to the total amount of water on land, encompassing all form of water stored on the surface and subsurface, including surface water, soil moisture, groundwater, snow, ice and water stored in vegetation.

My intent is to employ high spatial resolution data in the region to investigate the effect of variable climate conditions on deformation processes occurring at regional-to-

local (tens of kilometers to hundreds of meters) spatial scales.

In addition to gravitational loading water masses can exert other type of interaction with the solid earth. At local scales, groundwater storage (GWS) variation due to fluids recharge/withdraw in porous terrain layers (aquifers) can cause measurable deformation. This is due to poroelastic deformation processes and is typically investigated through InSAR technique, leveraging its high spatial resolution (tens of meters). This effect has been recognized and studied in detail in alluvial plains — in particular Central Valley in California where it has been found to contribute to a relevant part of the deformation signal [Kang et al. 2023].

Poroelastic deformation results from fluid pressure increase/decrease in saturated porous terrain layers, and manifests as, respectively, uplift and subsidence. The response in term of vertical deformation is therefore the opposite of the one produced by increase in hydrosphere loading. Additionally, the two types of deformation commonly occur together and may overlap.

To analyze hydrologically-driven deformation at the local scale, it is first necessary to detect and isolate the signals of interest. Statistical analysis techniques have been widely applied to detect areas characterized by similar deformation patterns [Kang et al. 2023], or to separate signals associated with independent sources [Gualandi et al. 2016]. In this work, I applied different data-analysis techniques and compared the identified signal with precipitation trends to identify processes likely influenced by hydrological dynamics.

Subsequently, I aimed to characterize the detected deformation processes and relate them to their respective forcing mechanisms. Since various coexisting processes through which water influences deformation can overlap and produce contrasting effects—as in the case of loading and pore pressure variations—it is necessary to quantitatively characterize their individual contributions in order to properly describe the dynamics of each individual process. To do this, I used the result of the Total Water Storage variation estimated by Pintori et al. (2024) and employed models of elastic loading response to estimate the elastic loading effect in the study region and separate poroelastic and elastic loading contributions.

Several areas across the study region exhibited deformation signals potentially associated with hydrological forcing at both seasonal and multi-year temporal scales. Among these, the Brescia area showed particularly clear multi-year trends and was selected for a more detailed analysis

I found that this area, in which the subsurface is characterized by a 100 m-thick aquifer body corresponding to the apical sector of an alluvial fan, undergoes significant poroelastic deformation in response to multi-year water variability. By comparing the poroelastic deformation contribution with groundwater level data, I then estimated the elastic parameters of the aquifer layers in correspondence with six groundwater measurement points.

Chapter 2

Study Area

2.1 Po plain

The region analyzed in this thesis is part of the broader Po Plain (Figure 2.1), the largest plain in Italy and one of the largest alluvial plains in Europe. The Po Plain is located in the north of Italy, between latitudes 44°N and 46°N. It covers an area of approximately 46 000 km^2 and it is enclosed between the mountain belts of Apenninens to the south and the Alps to the north and west. The terrain gently slopes from the mountain belt towards the center of the plain and from the west to the east, towards the Adriatic Sea.

The climate in the Po Plain is classified as humid subtropical (Köppen Cfa), characterized by hot and humid summers, with mean temperatures ranging from 22 °C to 25 °C, and cool winters with mean temperature from 1 °C to 4 °C. Annual precipitation ranges from 700 mm to 1200 mm with peaks in spring and autumn [Baronetti et al. 2020].

The region is crossed by numerous rivers. The largest is the Po River, which is Italy's longest river (691 km). It flows centrally approximately from West to East, crossing the entire plain before it reaches the Adriatic sea. Several tributaries from the surrounding mountain ranges flow towards the center of the plain and feed into the Po. The region also hosts some of the largest lakes in Italy which are located on the border between the mountainous area and the alluvial plain, over the course of Po tributataries. The abundance of surface water resources is associated with significant groundwater reserves, as the terrain in this region is composed of sedimentary deposits containing porous layers capable of storing large amounts of water.

Water availability, in combination with the favorable terrain conformation, makes this region very suitable for farming and for human settlements, which have existed in the area for millennia. Nowadays this region is extremely relevant from a socio-economic point of view, hosting approximately one-third of the Italian population and a significant portion of the country's agricultural and industrial production [Baronetti et al. 2020].

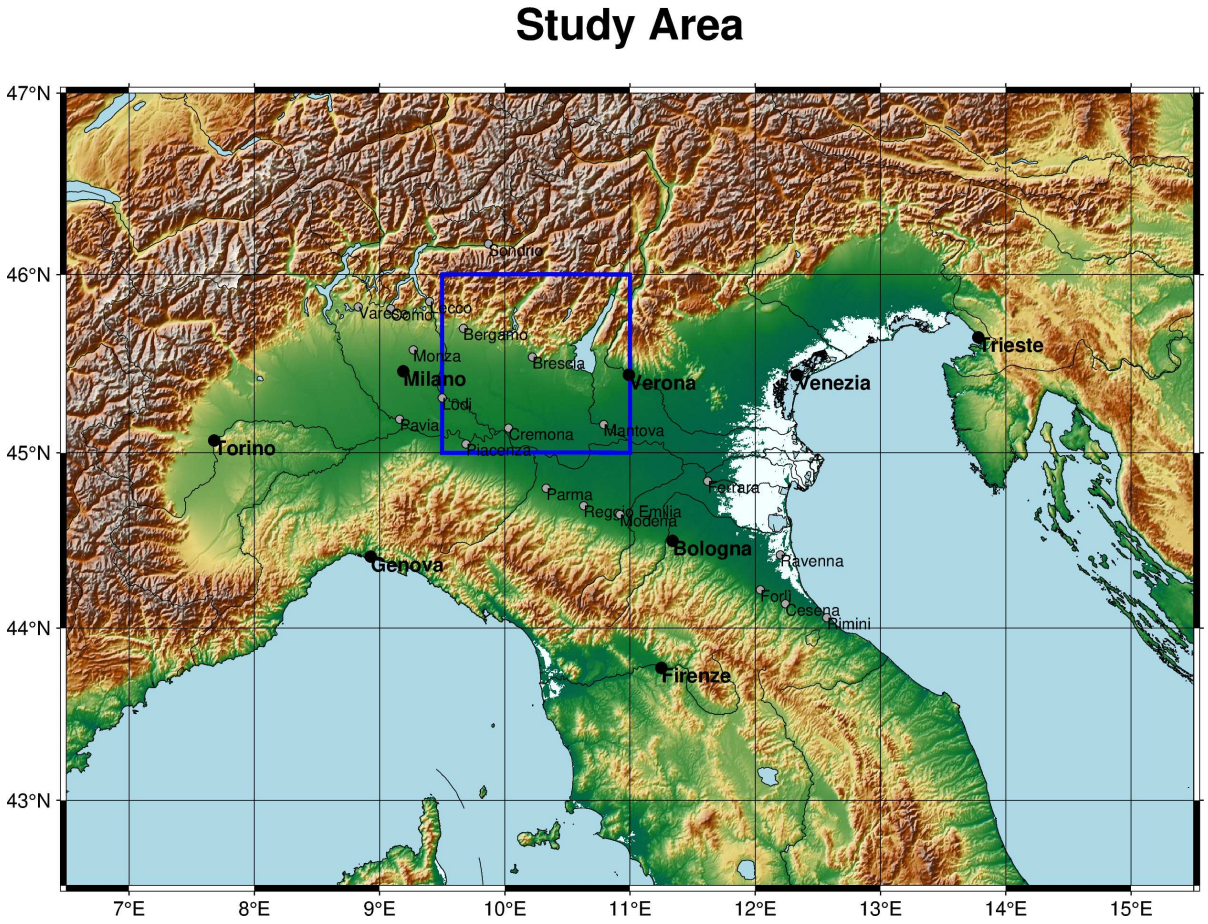


Figure 2.1: Map of northern Italy. The blue rectangle represent the study area.

2.2 Hydrogelogy

The sediments that nowadays constitute the Po Plain have been produced by erosion of the surrounding mountain belts and have progressively accumulated, transforming the basin, which 1.7 million years ago was an oceanic basin [De Donatis et al. 2007]. The rocky substrate is part of the Adriatic Plate, which is subsiding under the Apennine mountain belt. The tectonic subsidence is the primary cause of sediment accumulation [Carminati et al. 2002]. Materials transported by rivers were deposited in the lowest parts of the basin, forming an accretionary wedge prograding from west to east. Climatic and eustatic oscillations further influenced stratification dynamics [De Donatis et al. 2007]. The study by Regione Lombardia (2002), using borehole and seismic profile data, reconstructed in detail the evolution of the basin and terrain stratigraphy, identifying four hydrostratigraphic units, i.e. aquifer groups separated by permeability barriers of regional extension. These have been denoted, from the surface downward, A, B, C, D

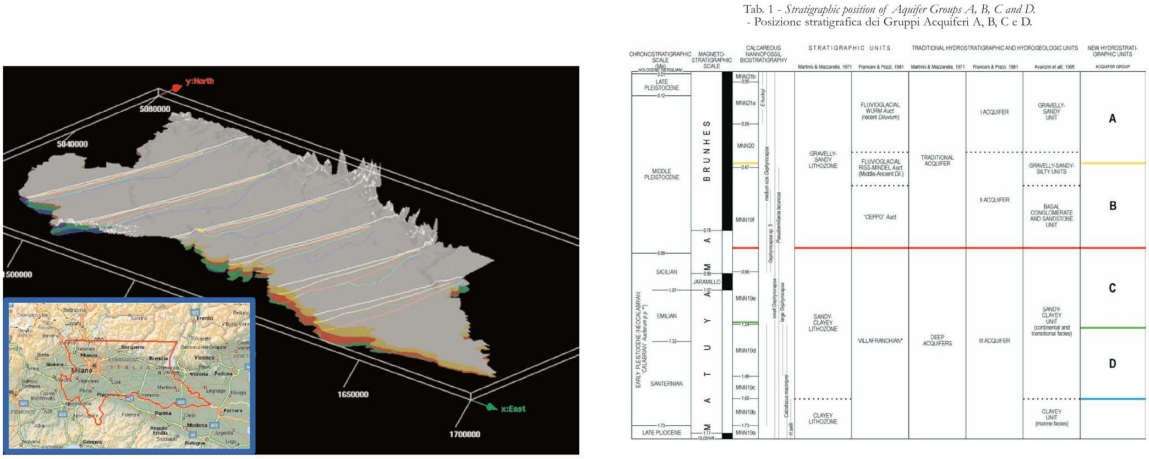


Figure 2.2: 3D reconstruction of the shape of the four hydrostratigraphic units in the Lombardian Plain (left). Stratigraphy of the Po Plain (right) [De Donatis et al. 2007]

(Figure 2.2).

The deepest groups (C,D) correspond to marine and transitional depositional phases. The alternation of marine, deltaic, and continental alluvial plain sediments results in a variety of permeable strata (aquifers), where sandy deposits dominate, and impermeable strata (aquitards), where clay prevails. Groups A and B are composed of continental deposits. Sediments are composed of sand and gravel with high permeability, with grain size decreasing toward the center of the plain, reflecting the loss of energy of the depositional systems. Impermeable layers of fine sediments are less relevant in continental layers and are found mainly in the deepest part of the sequence. Their thickness and frequency increase towards the center of the plain, producing a progressive confinement of the aquifers [Regione Lombardia 2002].

The ground structure in the plain can host substantial water resources, both in continental and oceanic sediments. The continental sediments host unconfined aquifers, which are directly recharged by precipitation in the whole plain area due to the permeable characteristics of the terrain, while the oceanic sediments host confined aquifers which are vertically limited by permeability barriers. These aquifers are recharged near the mountain fronts, where coarser material allows water to infiltrate and reach the deepest layers.

2.3 Water stress in the Po Plain

The complex hydrogeological setting of the Po Plain, shaped by a variety of depositional environments and hosting both confined and unconfined aquifers, results in a highly dynamic water system. This system is sensitive to natural climatic variability but

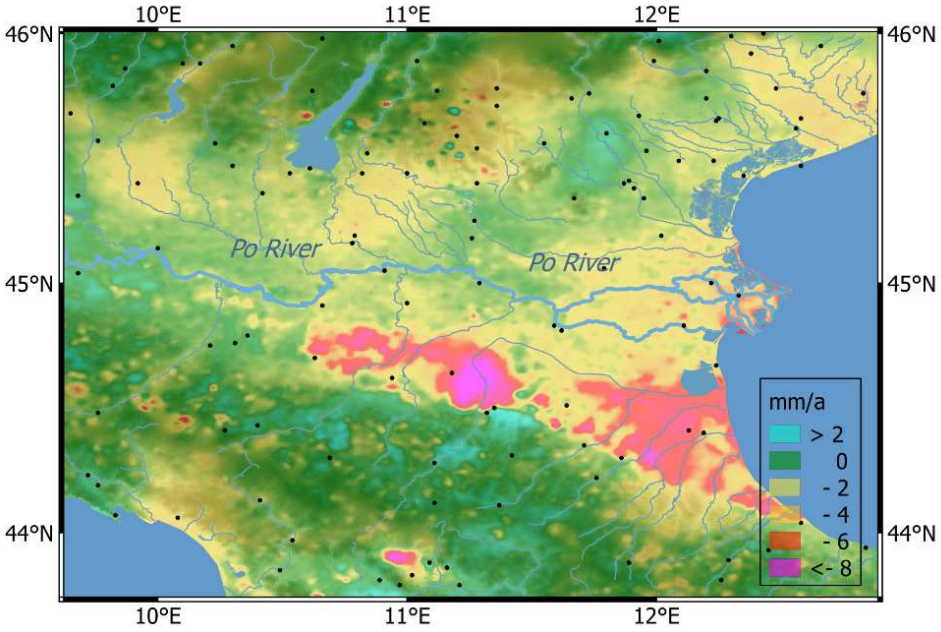


Figure 2.3: Vertical movement of Earth's surface in the Po Plain estimated by combination of GNSS and InSAR data showing subsidence in the northeastern sector of the Plain [Farolfi et al. 2019]

also increasingly exposed to anthropogenic pressures and climatic extremes, which have recently intensified the region's water stress.

Among the most prominent geophysical consequences of this increasing water stress is land subsidence, primarily driven by groundwater overexploitation. In particular, excessive pumping from the shallow unconfined aquifer systems has been recognized as the main driver of the high subsidence rates observed in the eastern sector of the Plain, where values exceed $10\,\mathrm{mm\,y^{-1}}$ (Figure 2.3) [Farías et al. 2024, Farolfi et al. 2019, Carminati et al. 2002]. This phenomenon has been extensively documented and is largely attributed to the compaction of sediments following prolonged aquifer depletion.

While subsidence due to groundwater over-exploitation represents a long-recognized anthropogenic impact on the region's geodynamics, more recent climatic shifts have introduced additional sources of stress, leading to distinct but equally significant deformation signals. In particular, the evolving climate regime of the 21st century has led to an increase in both the frequency and intensity of drought events, with notable episodes occurring in 2003, 2007, and 2011, linked to rising mean temperatures and shifts in seasonal precipitation patterns [Baronetti et al. 2020]. In 2022, the region experienced the worst drought in the last two centuries (Figure 2.4), with the Po River reaching record low levels [Montanari et al. 2023]. This decline was caused by a combination of reduced snowfall, earlier snowmelt, and increased evapotranspiration. Additionally, the expan-

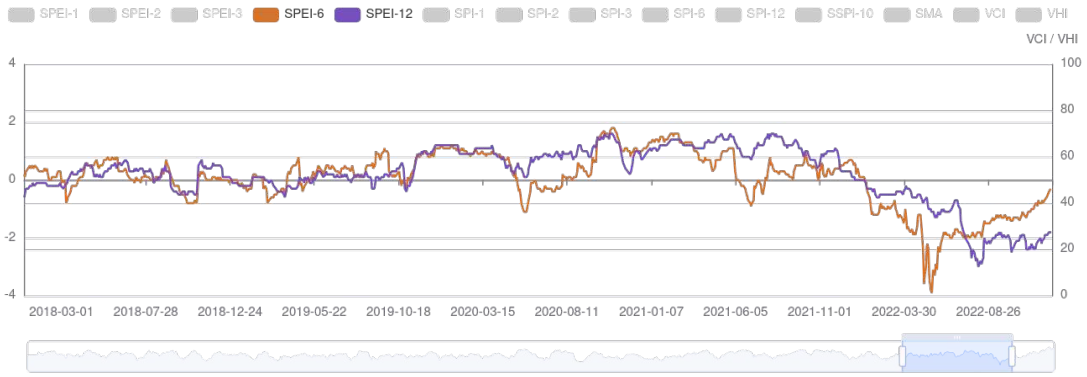


Figure 2.4: Standardized Precipitation and Evapotranspiration Index [Vicente-Serrano et al. n.d.] at 6-months (orange line) and 12-months (purple line) aggregation timescales. The drought period of 2022 is evident in both indices [Slovenian Environment Agency et al. 2022]

sion of irrigated land, which doubled between 1900 and 2015, likely contributed to the progressive depletion of water resources [Montanari et al. 2023].

Beyond its hydrological implications, this extreme drought also triggered measurable geophysical responses. The associated loss of water mass caused an elastic uplift of the ground, detected through GNSS (Global Navigation Satellite System) observations. Based on a regional-scale analysis of GNSS data across the Po Plain, Pintori et al. (2024) identified a multi-year deformation signal strongly correlated with both the Po River discharge and the SPEI-12 index. This signal was interpreted as the solid Earth's elastic response to drought-induced depletion of water storage. From the observed uplift — reaching values of up to approximately 7 mm — they estimated a total terrestrial water loss of around 70 Gt between January 2021 and August 2022 (Figure 2.5).

The critical socio-economic role of the Po Plain and the importance of water resources for its development make the study of groundwater dynamics in this region a topic of high interest. Furthermore, as demonstrated by the cases of long-term anthropogenic subsidence and drought-related elastic uplift, water dynamics in the region is responsible for the occurrence of relevant geophysical processes that could pose significant risks to human settlements and activities. The characterization of hydrological processes and their interaction with the solid earth can thus be very useful for the quantification and mitigation of these risks, as well as providing useful information to help an informed management of water resources.

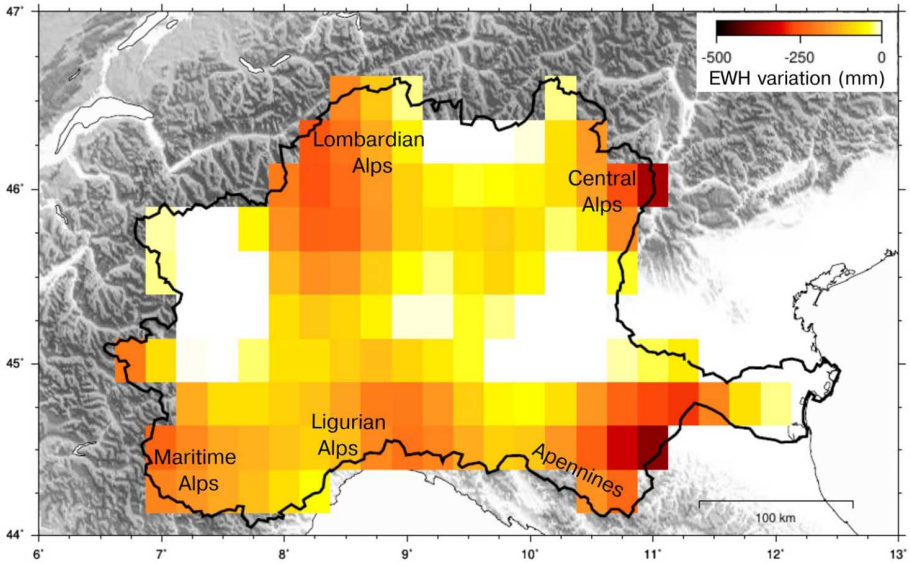


Figure 2.5: Estimated distribution of water loss in the Po Plain by inversion of GNSS data [Pintori et al. 2024]

2.4 Brescia area

In this work I conducted further analyses focusing on a sub-region of the Lombardy plain (Figure 2.6) which revealed to exhibit a clear hydrology-related deformation (Section 7). The selected area is located in the high plain sector, between the foothills of the Alps to the north and the central Lombard plain to the south, and approximately corresponds to the province of Brescia. The area lies between two major Po tributaries: the Oglio river and the Chiese river. In this area the terrain is mainly made of the result of deposition from the Mella river, a minor stream originating in the Trompia Valley and flowing into the Oglio river. This stream is responsible for the creation of an alluvial fan [Fontana et al. 2014]. In this kind of structure coarse, highly permeable sediments accumulate in their upper sector creating favorable recharge conditions and favoring significant groundwater storage [Zhu et al. 2017]. The regional environmental agency ARPALombardia collected and analyzed data on the subsurface geology and hydrology obtained through a network of boreholes, and summarized the results in technical reports [ARPALombardia 2018a, ARPALombardia 2018b, ARPALombardia 2019]. The data allowed for a 3D reconstruction of the aquifer body, represented in cross-sections and 2D maps (Figure 2.8,2.7).

In the northern sector, the uppermost layer is composed of coarse gravel sediments, about 50 m thick, associated with aquifer group A. This overlies layers of compacted gravel and sand conglomerates of variable thickness, associated with group B. Both

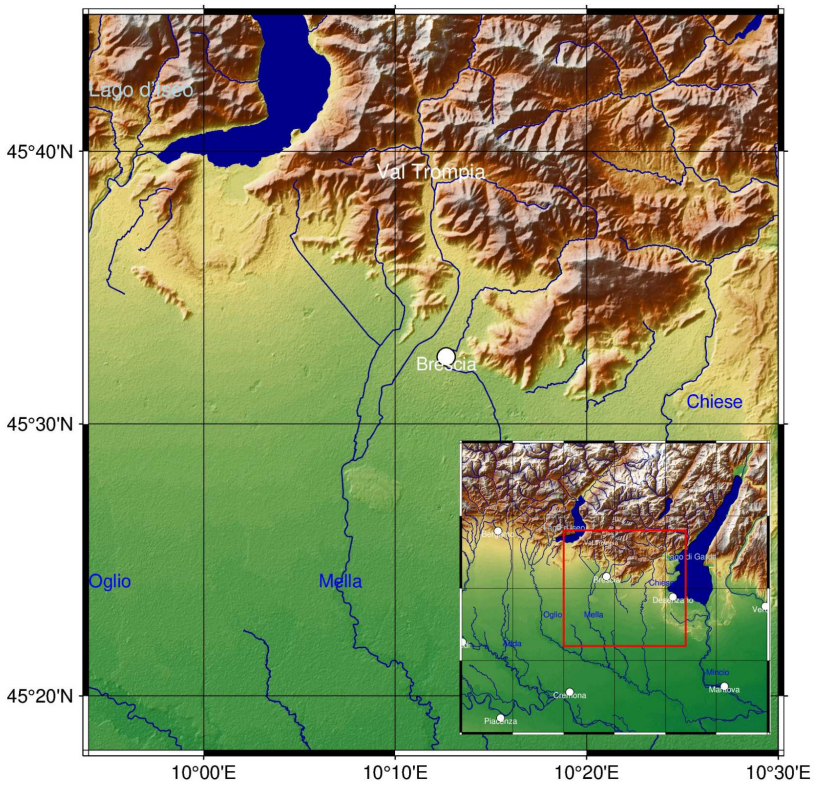


Figure 2.6: Location of Brescia area in the region of interest.

deposits are of continental origin. Marine and transitional deposits of group C are absent in this section but appear progressively southward at shallower depths. In the northern section, the two layers of continental deposits are only locally separated by a layer of clay-rich material which does not present sufficient extension to produce complete hydraulic separation. This results in a mostly unconfined aquifer directly overlying the carbonate substrate. In this section, the aquifer reaches a thickness of over 100 meters. This has been attributed to local morphological characteristics, in particular a buried fault, which in the northern section could have displaced the substrate towards lower depth creating space for sediment accumulation [ARPALombardia 2018a]. Southward, the combined thickness of the two units decreases due to the thinning of conglomerates and the increasing height of group C deposits.

The aquifer, hosted in gravel and conglomerates, acts as a high-capacity reservoir recharged by groundwater inflow from the Trompia Valley and by local precipitation. From this reservoir, groundwater flows southward under the influence of gravity, contributing to the recharge of deeper aquifer systems in the Po Plain.

Previous studies conducted a detailed reconstruction of the spatial and temporal

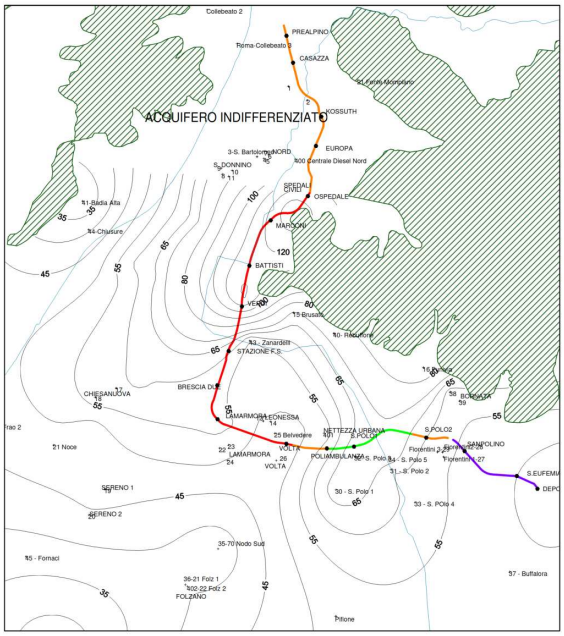


Figura 11 - Andamento della base dell'acquifero monostrato ghiaioso-sabbioso e conglomeratico (m s.l.m.) nella città di Brescia

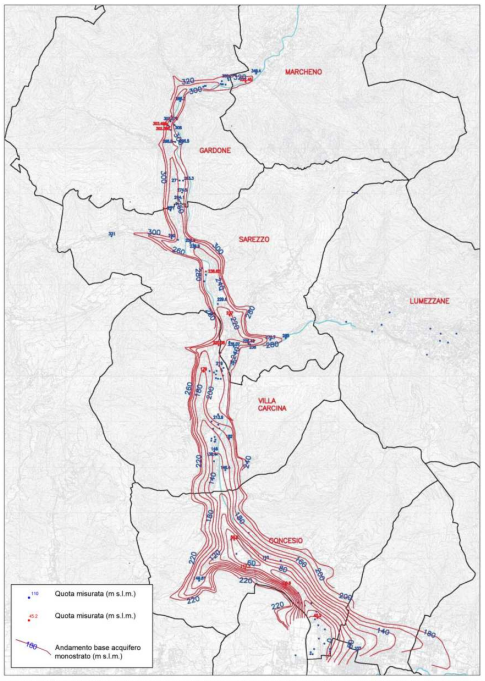


Figura 14 - Andamento della base dell'acquifero monostrato ghiaioso-sabbioso e conglomeratico (m s.l.m.) lungo il fondovalle della Val Trompia

Figure 2.7: Reconstruction of the basal surface of the aquifer in the Brescia area (left)[Provinca di Brescia 2016] and in the Trompia Valley (right)[Provinca di Brescia 2015]. A deepening of the substrate from approximately 140 m a.s.l. to 60 m a.s.l. is observed in the southern section of the Trompia Valley.

variations of the water table in the area. Combined with an in-depth characterization of aquifer geometry and lithology, this has supported the development of a groundwater flow model [ARPALombardia 2018a]. The main boundary conditions and hydrological inputs of the model are schematized in Figure 2.9. Groundwater flow is directed from the mountainous area towards the plain, following the natural topographic gradient.

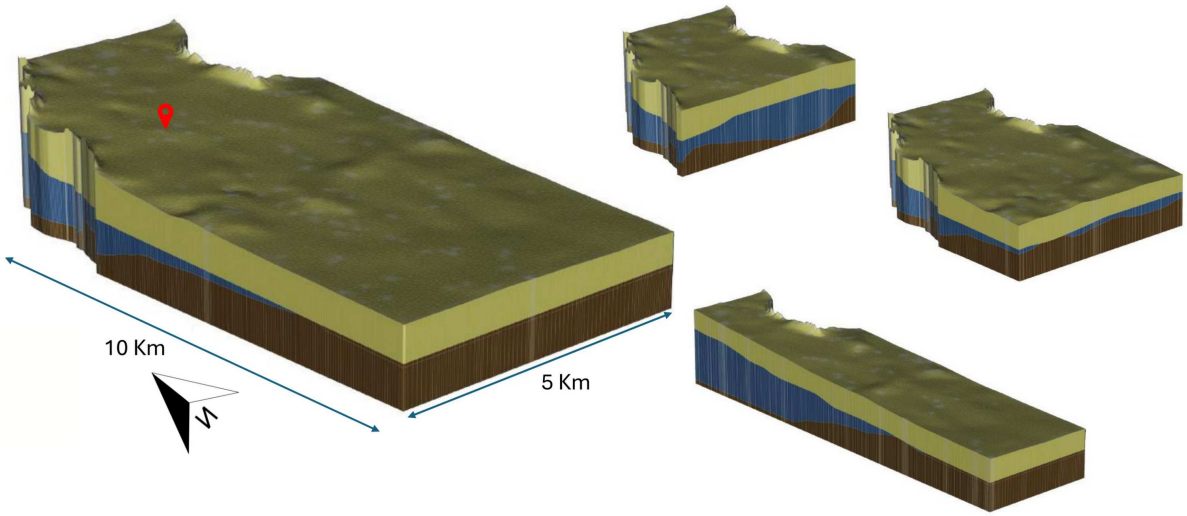


Figure 2.8: 3D qualitative sketch of the terrain structure in the Brescia area (left). Different colors represent respectively gravel (G, yellow), conglomerates (C, blue) and clay (A, brown) [Pastore et al. 2024]

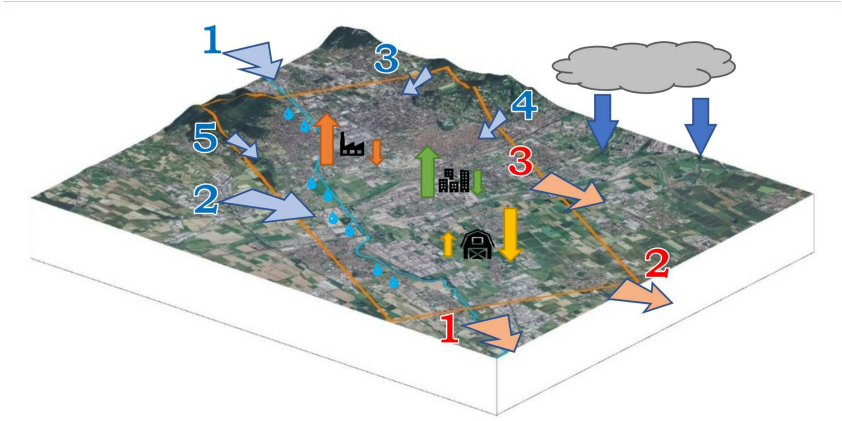


Figure 2.9: Conceptual scheme of the main hydrological inputs and outputs governing the groundwater balance in the Brescia area. Blue arrows indicate inflows from the Trompia Valley (1), western aquifer bodies (2), and surrounding slopes (3, 4, 5). Red arrows represent groundwater outflows directed towards the Po Plain. Orange, green, and yellow arrows correspond to anthropogenic components—industrial, residential, and agricultural activities—that contribute to both recharge and extraction processes [ARPALombardia 2018a].

Chapter 3

Data

3.1 Deformation Data

The deformation data used in this study were retrieved by Synthetic Aperture Radar Interferometry (InSAR) . In particular, I used a standardized and scientifically validated dataset processed and freely available within the framework of the European Ground Motion Service (EGMS) [Crosetto et al. 2021], an initiative of the Copernicus Programme.

Thanks to increasing availability of SAR observations, and improvements in processing algorithms and computational capabilities, InSAR data offer nowadays the possibility to monitor ground deformation at large spatial scales and sub-centimeter accuracy. These features allow for both large-scale characterization of ground deformation and focused analysis on specific areas of interest. The following section introduces the basic principles of the InSAR technique and describes the processing workflow adopted by the EGMS. Further details can be found in [Ferretti, Passera, et al. 2023].

3.1.1 InSAR

SAR

InSAR is based on Synthetic Aperture Radar (SAR), a remote sensing technique based on the use of satellite Earth's surface images acquired by microwave SAR sensors. SAR sensors emit an electromagnetic pulse and recover both amplitude and phase of the signal backscattered from the ground to the satellite. Typical frequency bands used by SAR sensors are L-band, C-band and X-band, with wavelengths respectively of about 24 cm, 5 cm and 3 cm. By using microwaves in an active configuration, radar sensors present some key benefits with respect to optical ones: they can work day and night and, employing long wavelengths, they can see through canopy and clouds (Figure 3.1) [Ferretti, Passera, et al. 2023].

The acquisition geometry of a SAR sensor is sketched in Figure 3.2. The sensor

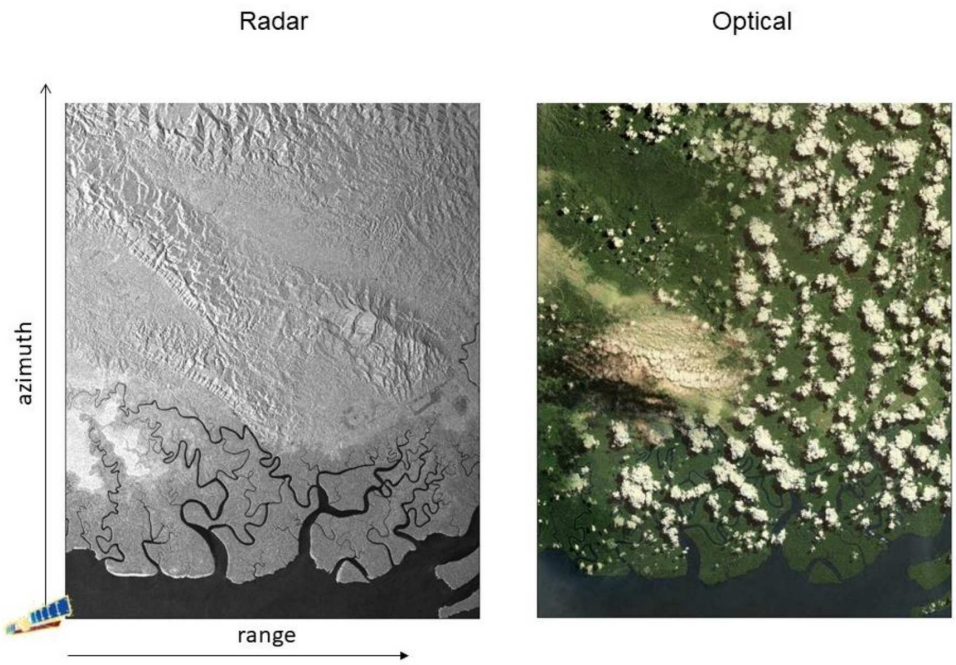


Figure 3.1: Comparison between SAR and optical image of the same area [Ferretti, Passera, et al. 2023]

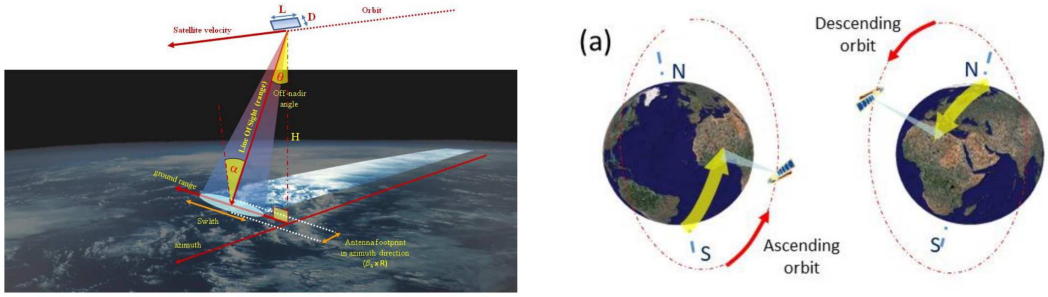


Figure 3.2: Geometry of SAR sensor acquisition and satellite orbit [Ferretti, Passera, et al. 2023]

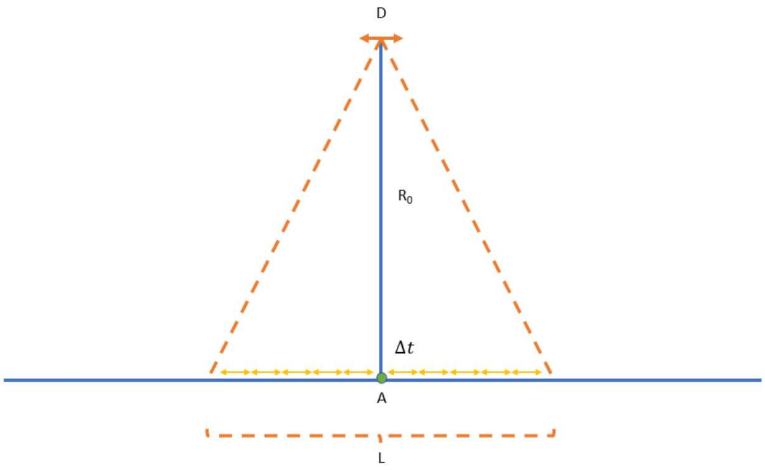


Figure 3.3: Representation of a synthetic aperture [Ferretti, Passera, et al. 2023]

is mounted on the satellite facing ground and is side-looking at an angle α called the off-nadir angle. It acquires images of the Earth's surface in two dimensions, cross-track (range) and along-track (azimuth). The resolution of the image in the two directions is a key characteristic of the system and depends on different factors. In the cross-track direction, it depends on the pulse length and bandwidth, and resolutions of the order of meters can be achieved using a frequency-modulated pulse of high bandwidth. In the along-track direction, the resolution is instead limited by the dimension of the satellite antenna:

$$L \approx \frac{R_0 \lambda}{D} \tag{3.1}$$

where L is the ground resolution, R_0 the distance of the sensor from the target, λ the wavelength of the signal, and D the antenna length in the along-range direction. For a sensor imaging the Earth's surface from a satellite orbiting at an altitude of 900 km with a wavelength of 5.5 cm (C-band), an antenna length of 10 m would yield a ground resolution in the along-track direction of 4950 km, which is too low for detailed surface monitoring. To overcome this limitation, SAR employs the synthetic aperture technique. This consists of combining acquisitions along the satellite motion to simulate a larger antenna (Figure 3.3) and increase resolution at the cost of additional processing. In this way ground resolution of the order of meters can be achieved in both imaging directions.

The possibility to measure surface movements from SAR images derives from their active characteristic, i.e. the capability to measure both amplitude and phase of the backscattered signal. In fact, while amplitude values are related to the backscattered energy, and can yield information on the physical characteristics of the surface, the phase value is related to the time delay between the emission and the registration of the signal, which contains information on the travel path and thus the distance between the satellite

and a target on the surface.

In the next section I describe how this information is exploited in Interferometric SAR to measure ground movement.

SAR Interferometry (InSAR)

SAR Interferometry (InSAR) consists of using SAR images to create an interferogram, i.e. a map of the difference of the measured phase values for each pixel between two images acquired from the satellite in two passes over the area of interest.

Typically, the satellites carrying the sensor orbit the Earth in sun-synchronous polar orbit with a revisit time of some days, acquiring images in two geometries: ascending, when the satellites moves from south to north and descending, in the opposite case (Figure 3.2). If two or more images of the same area, acquired in the same geometry, are available and can be aligned with a sub-pixel precision, then the phase difference between the two images can be exploited to obtain information on topography and surface deformation.

When images are taken from two slightly different angles, the relative phase shift in each pixel is produced by the different lengths of the path traveled by the pulse to reach that pixel, from the two slightly different viewpoints of the satellite. With precise knowledge of the satellite position and orientation in the two passes, the phase shift can be related to the distance of the satellite from each reflecting point on the surface, and thus yields a measure of surface topography.

If instead the topographic contribution is removed from the phase shift values using a Digital Elevation Model (DEM), then the residual phase shift reflects the variation of the travel path of the signal is related to the movement of the target that ocurred between the two acquisitions. Thus, phase shift measure can yield a measure of the cumulative surface deformation occurred in the line-of-sight direction of the satellite, in that interval, e.g. a phase shift of 2π corresponds to a displacement of the target towards the sensor by half a wavelength (Figure 3.4).

The process by which the displacement-related phase shift is isolated from the other contributions is referred to as Differential InSAR (DInSAR). Different factors besides deformation and topography however contribute to the measured phase shift. The total phase shift can be therefore expressed as:

$$\Delta\phi = \Delta\phi_d + \Delta\phi_{TOPO} + \Delta\phi_{ATM} + \Delta\phi_{NOISE} \tag{3.2}$$

where:

- $\Delta\phi_d = \frac{4\pi}{\lambda} d_{LOS}$ is the phase shift due to displacement in the line-of-sight direction between the two satellite passes.
- $\Delta\phi_{TOPO}$ is the phase shift produced by different viewing angles over varying topography

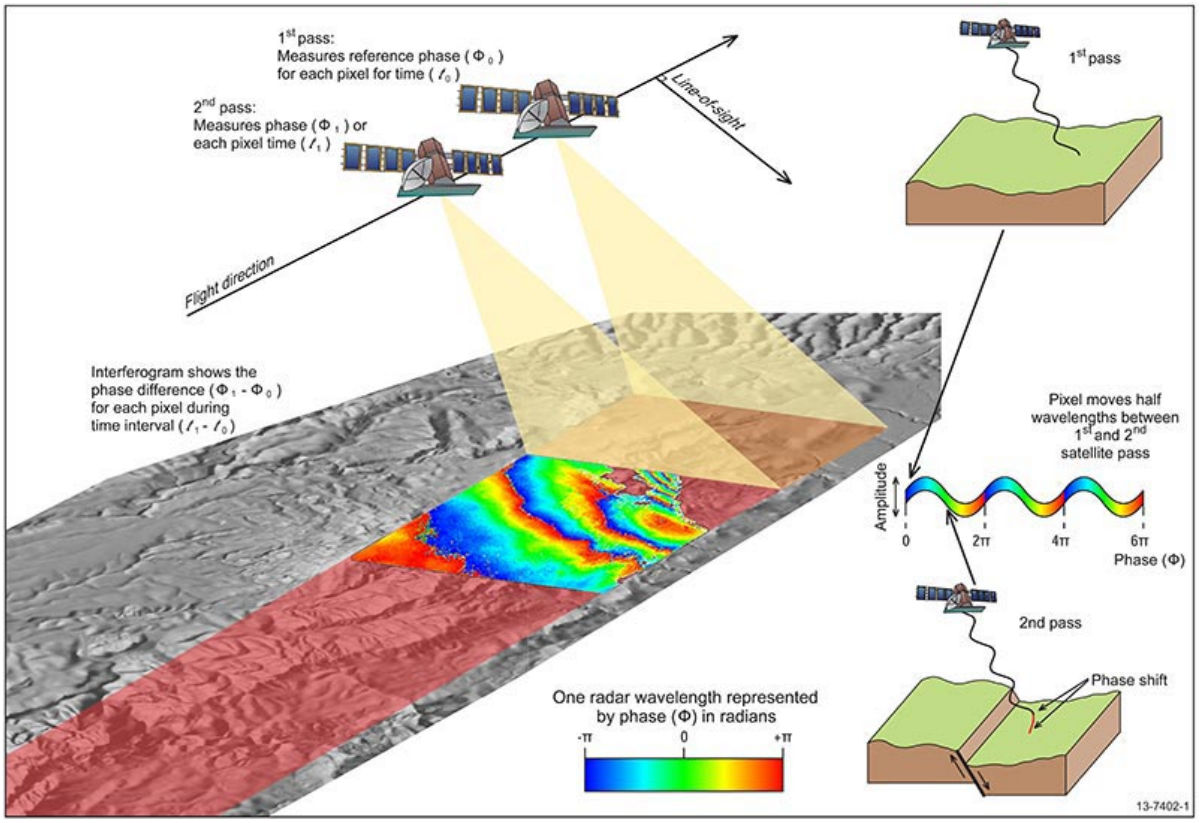


Figure 3.4: Representation of InSAR measurement of topography ad displacement [Australia 2017]

- $\Delta\phi_{ATM}$ is related to changes in ionospheric and tropospheric delay effect between the two acquisitions.

- $\Delta\phi_{NOISE}$ includes all remaining noise sources.

From Equation 3.2 it is evident that not all pixels provide useful information on surface displacement. In fact, in areas where vegetation changes, erosion or rapid movement of the surface cause inconsistent interaction of the electromagnetic signal with the surface in the two images, the noise component $\Delta\phi_{NOISE}$ can overwhelm the signal. Therefore, the phase signal can be used to obtain information on displacement only in areas characterized by a high coherence of the reflected signal, typically associated with the presence of highly reflective and stable surfaces such as buildings or rock outcrops [Ferretti, Passera, et al. 2023].

In these coherent areas, once topographic and atmospheric contributions have been removed, the remaining phase signal reveals fringe patterns in the interferogram (Figure 3.4). Each fringe is associated wit a displacement of half wavelength between the two images, therefore, to obtain surface deformation, an additionital 'unwrapping' operation must be performed. This consists in estimating, for each pixel, the correct number of 2π phase cycles to be added to the phase signal to retrieve the full phase shift.

From the above discussion two key points related to InSAR measurement are highlighted:

- InSAR, being obtained from differences in the travel path of a signal between two acquisition, are relative measurements of surface displacement both in time and in space, and are not therefore immediately comparable measurement.

- In order to obtain surface deformation, SAR data have to undergo several processing steps: alignment of the SAR images, topographic correction, atmospheric denoising, and phase unwrapping. These operations, given also the high spatial resolution of SAR images, can be computationaly intensive.

In the next section I describe the procedure by which the EGMS chain performs these processing operations and the calibration procedure applied to the measurements to provide surface deformation in an absolute reference frame.

3.1.2 EGMS InSAR processing

This section outlines the main steps of the EGMS processing workflow, including interferogram generation, correction of topographic and atmospheric effects, identification of measurement points, and geocoding.

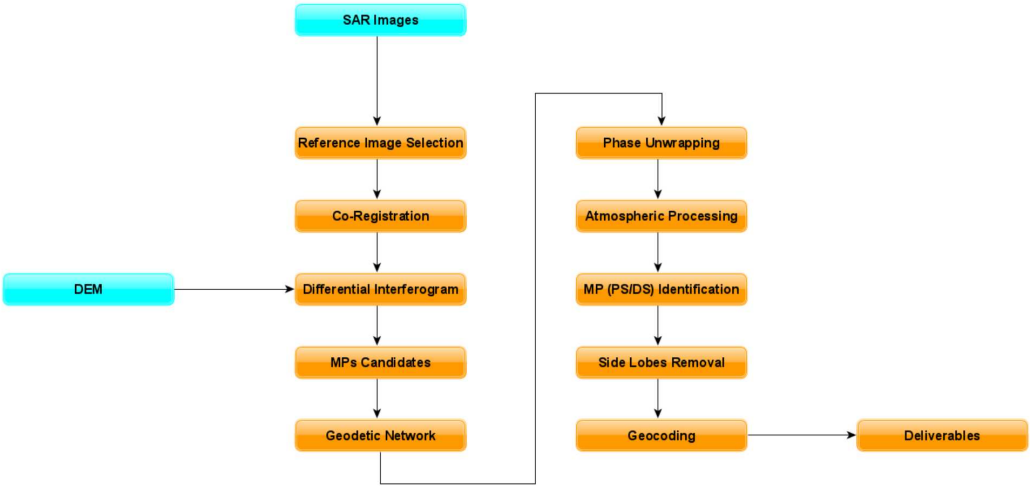


Figure 3.5: EGMS processing workflow [Ferretti, Passera, et al. 2023]

Acquisition

Data processed by the EGMS consortium are derived from SAR images acquired by the Sentinel-1 satellites. These satellites are equipped with twin SAR sensors operating in the C-band and are put in orbits with a minimum revisiting time of 6 days for each acquisition geometry covering most of the Earth's surface. Images are acquired in the Interferometric Wide (IW) swath mode, with a resolution of 20 m (azimuth) by 3.5 m (range). The acquisition mode combines different burst to obtain an overall 250 km by 180 km image.

The acquired images are saved in Single Look Complex (SLC) format in which each pixel has a complex value representing the amplitude and phase of the radar signal. Multiple SLC images of the same area are then processed togheter through several steps (Figure 3.5) which are syntethically described hereafter.

Reference image selection

The first step is the selection of the master image of the interferogram, that is the image used as reference to perform the interferometry. The interferograms produced by SAR images are characterized by the parameters:

- Temporal baseline B_t: time interval between two acquisition. This parameter influences the noise values of the interferogram. The smaller is the temporal baseline the higher is the signal to noise ratio.
- Normal baseline B_n: difference in position of the sensor between the two acquisitions. The larger the normal baseline, the more sensitive is the system to topogra-

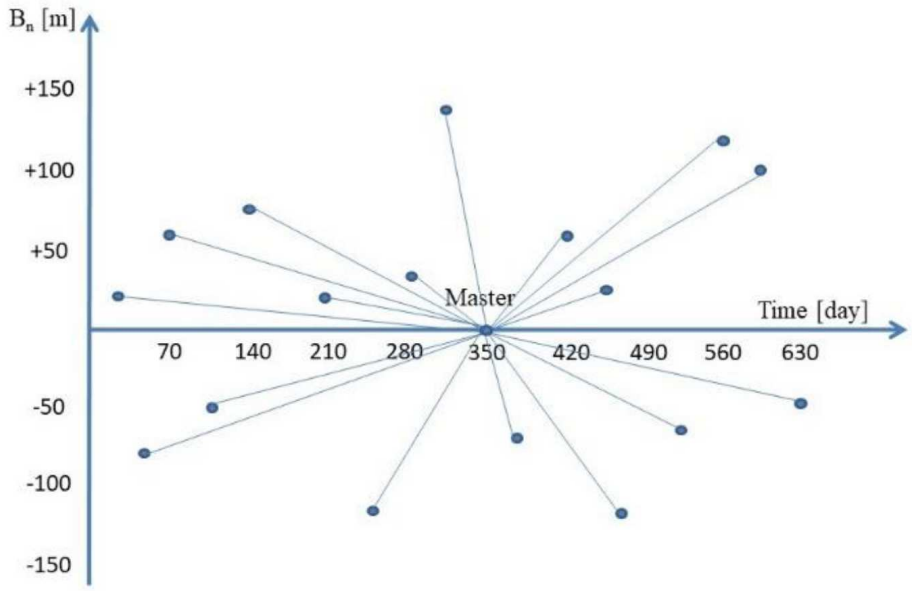


Figure 3.6: Master image is selected to minimize dispersion of baseline [Ferretti, Passera, et al. 2023]

phy. Values of the normal baseline above a critical value, called critical baseline B_c, which depends on satellite acquisition geometry and bandwidth, produce temporal decorrelation of the phase values of the signal.

In EGMS processing, the master images of the stack is selected as the one minimizing the dispersion of the temporal and normal baseline values (Figure 3.6) and maximizing the expected coherence of the interferogram:

$$\max_i \left(\sum_j^N \left(1 - \frac{|B_n^{ij}|}{B_c} \right) e^{-|B_t^{ij}|/\tau} \right) \tag{3.3}$$

where B_n^{ij} B_t^{ij} are the values of the normal and temporal baseline between images i and j, N is the number of images, B_c the critical baseline of the system and τ a temporal decorrelation constant, usually set to 30 days.

Co-registration

Co-registration consists in the resampling of all secondary images to the master images. This procedure is necessary since the same target does not correspond to the same pixel in different images due to difference in the acquisition geometry and in the

sensor characteristics. The second effect is limited in Sentinel-1 data since the two satellites carry twin sensors. The first effect is corrected using an approach based on cross-correlation computation to find analogous image patches, combined with a geometric approach that corrects for different viewpoints using precise satellite orbits information. Due to the high precision required (up to 0.001 pixel), this step is one of the most computationally intensive of the whole process.

Differential interferogram creation

From co-registered images the interferogram is then created by multiplying the complex value of each pixel in the master image by the conjugate of the corresponding pixel in the other image. The resulting phase shift is composed of the terms in equation 3.2. Using an high-resolution digital elevation model (DEM) and information on satellite track, a synthetic interferogram containing only the topographic component $\Delta\phi_{TOPO}$ is computed and then subtracted pixel by pixel to the original interferogram (Figure 3.7). The phase shift in the topographically-corrected interferogram, $\Delta\phi_c$ then contains the displacement-related term, the difference in the contribution of phase delay, a term proportional to the normal baseline B_n and the error ϵ associated with the used DEM, plus noise:

$$\Delta\phi_c = \Delta\phi_d + \epsilon B_n + \Delta\phi_{ATM} + \Delta\phi_{NOISE} \tag{3.4}$$

Selection of candidate measurement points

The core of EGMS processing is represented by the Permanent Scatters (PS) technique [Ferretti, Prati, et al. 2001]. This technique involves the search in multiple SAR images of the same area for points exhibiting stability of the reflected signal. These points are referred to as Permanent Scatters (PS). In EGMS, PS are searched by approaches based on both phase and amplitude information:

- In phase-based approaches PS are identified as pixels with high values of the temporal coherence:

$$\gamma = \left| \frac{1}{N} \sum_{j}^{N} e^{i\Delta\phi_{cj}} \right| \tag{3.5}$$

  where $\Delta\phi_{cj}$ is the topographic-corrected phase shift of a given pixel in the j-th image.

- Amplitude based methods instead relies on the fact that pixels with low values of the amplitude dispersion index - the ratio between mean and standard deviation of aplitude values - also exhibit low values of phase dispersion.

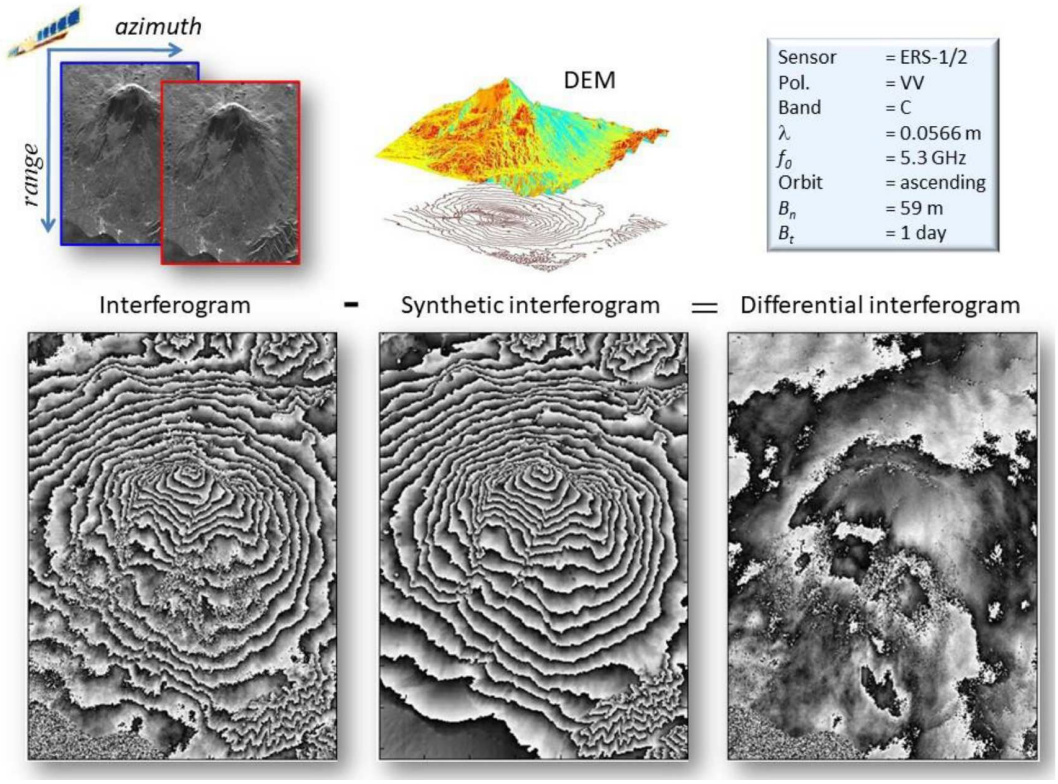


Figure 3.7: Topographic correction of interferogram [Ferretti, Passera, et al. 2023]

Measurement point selection is performed both before and after the estimation and removal of the atmospheric effect since additional points could be identified after the operation.

Phase unwrapping

The set of selected PS constitutes a sparse grid of points with relative phase shift $\Delta\phi_{ij}$ between them (Figure 3.8). The unwrapping operation in EGMS is performed modeling each displacement induced phase shift between neighbouring pixels P_1 and P_2 as produced by a relative linear motion with rate a v, assuming atmospheric effect to be very similar and thus canceling out:

$$\Delta\phi_{12}^{j} \approx \frac{4\pi}{\lambda} B_t^j v_{12} + C_z^j \epsilon_z = \hat{\Delta\phi}_{12}^{j} \quad j = 1, ..., N \tag{3.6}$$

where C_z is a constant and ϵ' is the error in the DEM. The values of v_{12} and ϵ' are estimated maximizing the temporal coherence:

$$\gamma = \max_{v_{12}, \epsilon_z} \left(\frac{1}{N} \sum_{k}^{N} e^{i\Delta\phi_{12}^{j}} e^{-i\hat{\Delta\phi}_{12}^{j}} \right) \tag{3.7}$$

The parameters provide an estimate of the phase shift between neighbouring points, from which the unwrapped phase values are computed by integrating the estimated phase difference across the PS network. Phase unwrapping is one of the most challenging steps in the processing chain. Various algorithms have been developed to perform this task, but they are not discussed here. For further details, the reader is referred to the EGMS Algorithm Theoretical Basis Document [Ferretti, Passera, et al. 2023].

Atmospheric delay correction

A C-band radar signal passing through the atmosphere experiences a phase delay mainly due to the interaction of radiation with tropospheric water vapor and to effects related to variations of the index of refraction with heigth under stratified conditions [Ferretti, Passera, et al. 2023]. The first effect has a characteristic length scale similar to turbulence phenomena -which varies from 1 km to 10 km- while the second is strongly correlated with topography. An additional, but marginal, effect is due to ionospheric disturbances, which occur on a larger spatial scale and are well modeled by low-order polynomials.

In EGMS these effects are estimated and corrected with a precedure that starts by assuming a parametric model for the phase shift produced by the displacement of the measurement point. Then, after fitting the unwrapped, topographically corrected, phase shift $\Delta\phi_c$ to the model, the atmospheric effect is isolated in the residuals with a filtering

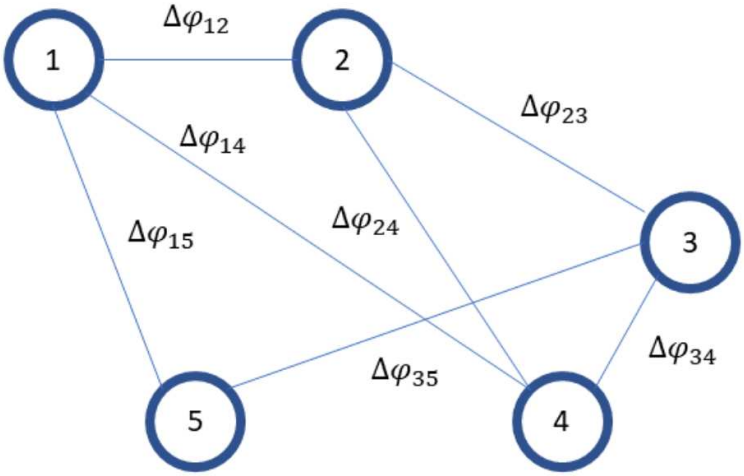


Figure 3.8: Sparse grid of candidate measurement points [Ferretti, Passera, et al. 2023]

procedure based on its statistical properties. The chosen model for the displacement-induced phase shift is a low-order polynomial plus a seasonal component with a period of one year:

$$\Delta\phi_{di} = \sum_{p}^{q} C_{tp} B_{ti}^{p} + S_s \sin(2\pi B_{ti}) + S_c \cos(2\pi B_{ti}) \quad p = 1, .., q \quad i = 1, .., N \tag{3.8}$$

wehere B_{tj} are values of the temporal baseline of the j-th image of the interferogram, expressed in years, q the order of the polynomial. Both the parameter of the model, $C_{t0}, .., C_{tq}$, S_c, C_c, and the error on the DEM (ϵ in Equation 3.4) are estimated from the fit. The fitted curve is and estimate of the displacement-induced phase shift $\Delta\phi_d$, while the residuals of the fit contain the contributions of $\Delta\phi_{ATM}$ and $\Delta\phi_{NOISE}$. Since atmospheric delay effects are expected to be correlated in space but uncorrelated in time, the residuals of the fit are low-pass filtered in space and high-pass filtered in time. The phase shift resulting from this filtering operation is assumed to derive only from atmospheric effects. The estimated atmospheric phase shift are then interpolated onto the grid of the master image using a model for its expected spatial dependency:

$$\Delta\phi_{ATM}(r, x) = \tau(x, r) + C_h h(x, r) + C_r r + C_x x + C_0 \tag{3.9}$$

where x, r are the azimuth and range coordinates, τ the tropospheric delay, $h(x, r)$ is the elevation at pixel x, r and C_h, C_x, C_r, C_0 are regressed constants. Obtained phase shift values at each pixel are then removed from the differential interferogram.

Selection of PS

After the estimation and correction of the atmospheric phase delay, another pixel-by-pixel analysis is performed on the interferogram to identify potentially additional measurement points. In this step, PS are identifed analyzing the dispersion of phase values with respect to the one expected from a dispalcement model [Ferretti, Passera, et al. 2023]:

$$\gamma(P) = \max_{P} \left| \frac{1}{N} \sum_{j}^{N} e^{i\Delta\phi_{aj}} e^{-i\Delta\phi_{mj}(P)} \right| \tag{3.10}$$

where $\Delta\phi_{aj}$ is the phase shift in the j-th image, corrected for both topography and atmospheric delay, and $\Delta\phi_{mj}(P)$ is the expected phase shift from the displacement model employed. P is the parameter vector of the model, estimated as the ones maximizing γ. The displacement model used in EGMS for coherence estimation is a third-order polynomial plus a seasonal component of period of one year:

$$d(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + A\cos(2\pi t + \phi) \tag{3.11}$$

where time is expressed in years and $c_0, c_1, c_2, c_3, A, \phi$ are the estimated model parameters. Each PS characterized by value of γ above a given treshold is selected as a measurement point.

Geocoding and time series computation

Phase values of selected PS are then unwrapped, and all displacement time series obtained are referenced in time to the start of the displacement model in equation 3.11 used for coherence estimation (the value t_0 at which $d(t_0) = 0$). For each PS, the image coordinates (range, azimuth) are converted into northing and easting coordinates in the European Terrestrial Reference System 1989 (ETRS89) using satellite position information at the time of acquisition.

The result of these processing steps is a set of single frames, each containing the time dependent deformation of measurement points belonging to the same interferogram and each having its own reference point. These data are the first type of product delivered by EGMS with the name of baseline products.

EGMS displacement products

EGMS aims to provide a standardized information on surface displacement consistent over a wide area. To achieve this goal, EMGS performs additional processing steps and provides three layers of product, differing for the aim and the type of information provided [Ferretti, Passera, et al. 2023].

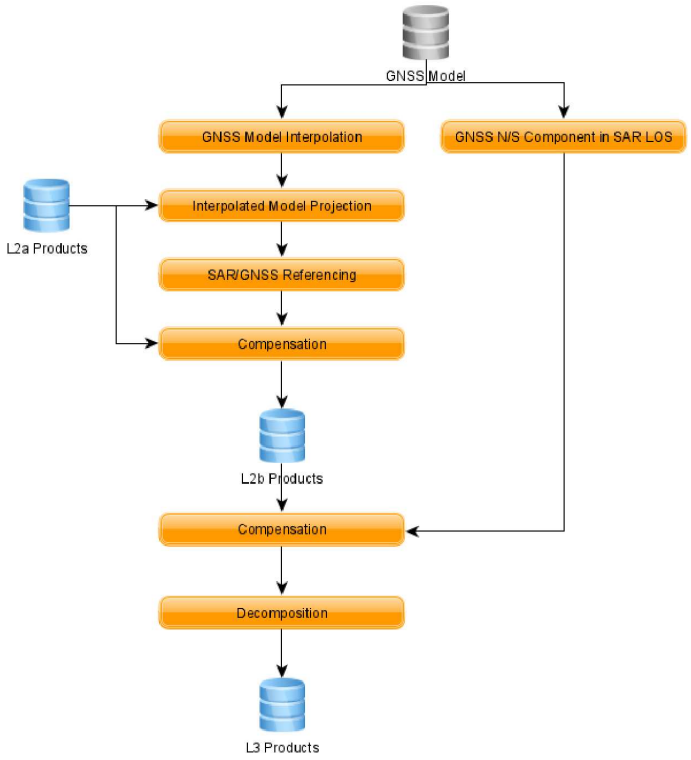


Figure 3.9: EGMS step for the creation of L2 and L3 products [Ferretti, Passera, et al. 2023]

- EGMS Basic product: This is the starting point for the generation of the other two products. Displacement time series belonging to the various frames are referenced spatially to a virtual point whose movement is estimated as the common mode of the dataset. All the images are then mosaicked together, and the data are resampled to a regular 100 m x 100 m grid. For all grid cells, all time series belonging to the same acquisition geometry are averaged and referred to the center of the cell. This dataset, while being standardized over the whole continental area, still contains the relative displacement in the LOS direction.

- EGMS Calibrated product. In this product the reference point of the displacement dataset is anchored to the network of Global Navigation Satellite System (GNSS) stations from the European Reference Frame (EUREF) Permanent Network Densification (EPND). This procedure has the double goal of making the deformation measurement absolute and to validate, with GNSS data, low frequency displacements, such as those caused by tectonic deformation, which in InSAR processing could be misinterpreted as phase artifact.

The procedure by which the dataset is obtained consists in interpolating EPND data onto PS positions and projecting them onto the LOS direction. The GNSS derived displacement are subtracted from the InSAR displacements. The difference is low-pass filtered and interpolated in space to create a smooth surface. This surface does not contain the real high-wavelength displacement (assumed to have been removed with the GNSS data), and it is removed from InSAR data. The virtual reference point is estimated from the obtained dataset. Its movement is estimated with GNSS then added back to the InSAR data to reference them to an absolute reference frame.

Obtained displacement are still one dimensiomal measurements, i.e. they are in the satellite LOS direction, thus the product is delivered in two datasets, one containing the displacement measured in the ascending geometry, the other in descending geometry.

- The EGMS Ortho product, the more advanced of the three, is intended to overcome this limitation combining data from different acquisition geometries from the Calibrated datasets to obtain 2D deformation components. This is done by combining ascending and descending LOS displacement to retreive vertical and horizontal (east-west) deformation components. It is relevant to note that, even when combining different acquisition geometries, a SAR sensor is almost blind to displacement components in the north–south direction, since this is approximately the direction along which it moves. Therefore, this component of displacement is entirely removed from InSAR data before performing the projection. The procedure consists in interpolating northward displacement from GNSS data onto the InSAR measurement points. Then, the northward displacement is projected onto the LOS direction for each acquisition geometry, and the result is removed from the Calibrated data. The resulting dataset should contain only the vertical and eastward components of displacement.

  All grid cells with at least one time series for both descending and ascending acquisition geometry are then selected and for each grid cell all the time series are resampled to a common 6-day time interval. At each time step the displacement in vertical and eastward direction can then be determined from the ascending and descending components using the following relation (Figure 3.10):

$$\begin{pmatrix} U_{east} \\ U_{vert} \end{pmatrix} = \frac{1}{2\sin\theta\cos\theta} \begin{pmatrix} -\cos\theta & \cos\theta \\ \sin\theta & \sin\theta \end{pmatrix} \begin{pmatrix} U_{asc} \\ U_{desc} \end{pmatrix} \tag{3.12}$$

  where θ is the mean incidence angle. The resulting deformation data is delivered in a product containing two datasets, one for the vertical component of displacement, one for eastward component.

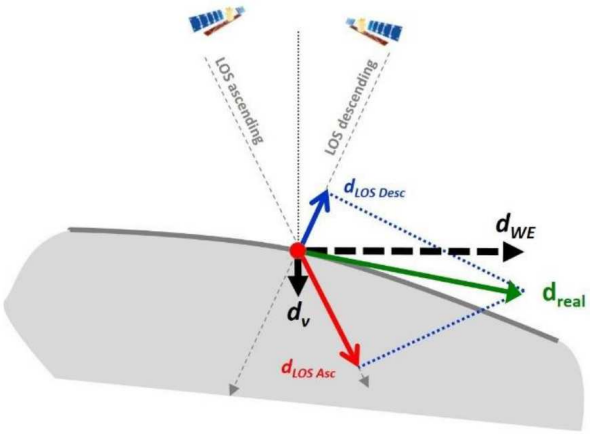


Figure 3.10: Decomposition of ascending and descending LOS displacement in vertical and eastward direction [Ferretti, Passera, et al. 2023]

The obtained displacement products are updated by EGMS on an annal basis and each update contain 5 years of data processed togheter. Data are available for consultation and download from the EGMS official website [https://egms.land.copernicus.eu/].

3.1.3 Data used in this study

For this work, I used the Level 3 EGMS Ortho dataset produced during the second update of EGMS dataset [European Environment Agency 2023]. This dataset spans the time interval from 06-01-2018 to 17-12-2022 with a temporal resolution of 6 days.I downloaded a total of four 100 km x 100 km tiles from the EGMS website. I mosaicked them together and clipped them to the region of interest: a rectangular region bounded by longitudes 9.5°E and 11°E and latitudes 45°N 46°N. I filtered the points by the value of the root mean square error (RMSE) provided with the dataset. This gives and indications of the sparsity of the time series as compared to a displacement model composed by a quadratic and polynomial function. I retained gridpoints with RMSE lower than 1.2 mm, for a total of 183.972 gridpoints and 302 time steps for each deformation time series (Figure 3.11).

3.2 Ancillary data

To explore the relationship between ground deformation and hydrological conditions, I analyzed data describing the hydrological variability across the study area. These include precipitation records collected by rain gauges during the time span covered by the EGMS dataset.

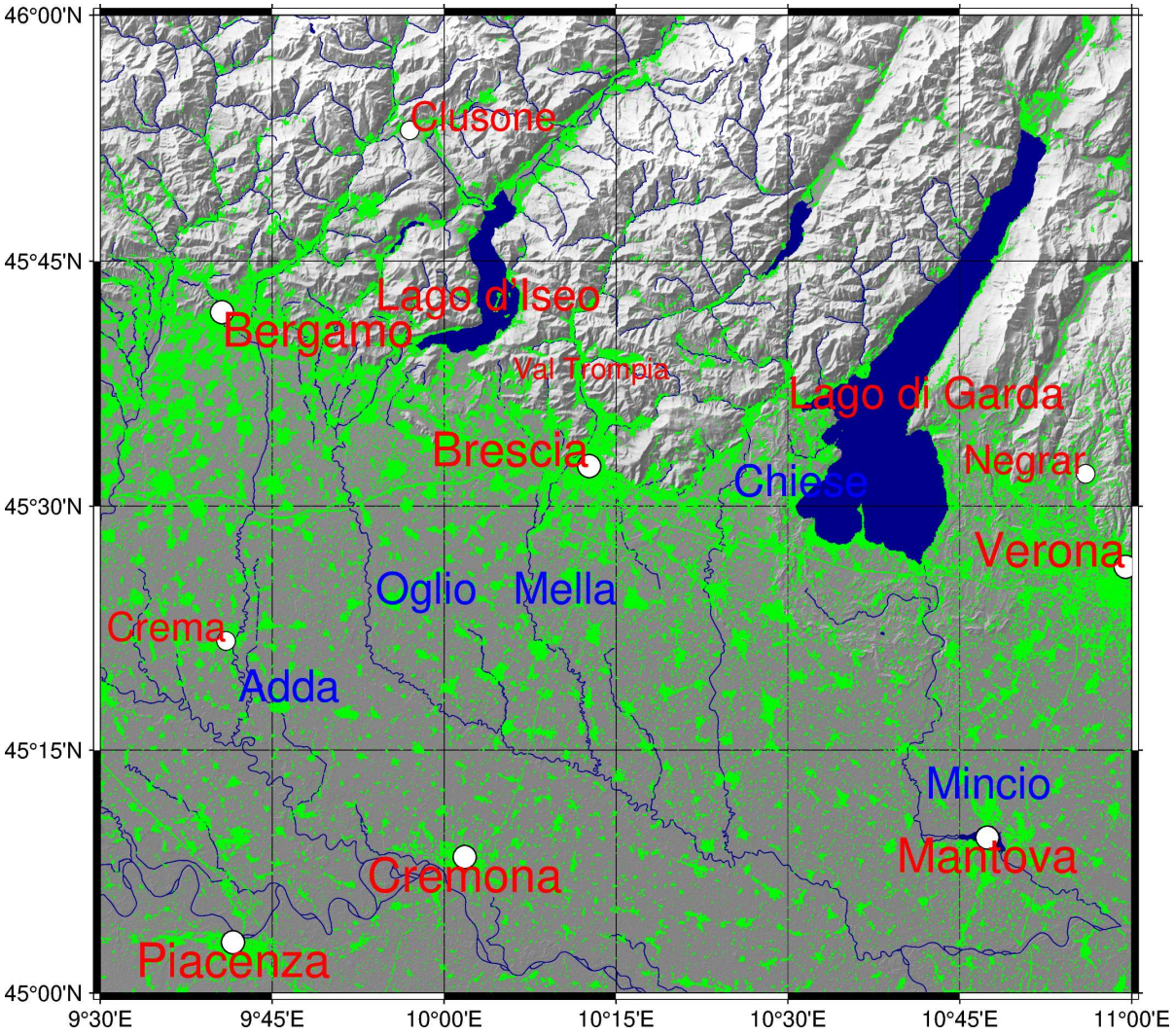


Figure 3.11: EGMS gridpoints (in green) analyzed

To better understand the processes underlying a deformation signal detected near the city of Brescia—likely associated with groundwater dynamics—I also analyzed groundwater level data from piezometers and stratigraphic information of the subsurface. The datasets used in this analysis are described in the following section.

3.2.1 Precipitation data

Precipitation data were recorded by pluviometers of the ARPALombardia network and are available for download from the Open Data Regione Lombardia portal [ARPA Lombardia 2020a, ARPA Lombardia 2023]. Among available sensors in the region of

interest, only sensors with a complete record in the considered period were selected, resulting in a total of 53 sensors retained (Figure 3.12). Data are provided at a variable sampling frequency, up to 10 min.

Time series of precipitation were visually inspected to remove unrealistically high values (example in Figure 3.13), which may be due to transcription errors. Afterwards, I computed 6-day cumulative precipitation time series at the same time steps of EGMS. Then, with the aim of highlighting multi-year variations in precipitation trends, I computed a detrended cumulated precipitation time series for each sensor, removing a linear trend from the series of cumulated precipitation (Figure 3.14).

This figure shows all the time series of cumulated detrended precipitation in gray, with their average overlaid in red. The average shows seasonal oscillations with peaks approximately in the months of December-January, reflecting the precipitation pattern typical of the climate of the area. The seasonal oscillation is superimposed to a clear multi-year trend showing an increase in cumulated detrended precipitation starting from 2018, which reaches the maximum values in 2021, followed by a sharp decrease starting in January 2022 and continuing until July of the same year. This trend reflects the drought conditions in the area in the year 2022 (Section 2.3) as cumulative detrend precipitation values sharply declined starting in January.

I considering two separate intervals, T_1 and T_2, in which to compare deformation and hydrological variability such that T_1 is representative of the rainy period and T_2 of the period associated with the 2022 drought. I defined T_1 as going from the beginning of the EGMS time series, i.e. the 1st of August 2018 to the 1st of January of 2021, and T_2 from 1st of January of 2021 to the 1st August 2022. These two intervals are shown as blue and yellow stripes in FIgure 3.14.

3.2.2 Hydrogeological data

A particular region in the center of the study area, showed a deformation signal correlated with precipitation. The area is located in the Brescia province (2.6) and it extends from the outlet of the Trompia Valley in the north to the transition to the lower plain in the south. For this region, I collected information on the underground structure and hydrogeological features to characterize the distribution and dynamics of ground water resources. Then I analyzed the variation of the groundwater level measured by piezometric sensors to search for correlation between groundwater level variation and surface deformation.

Piezometric data

Data of piezometric sensors in the Brescia area were not available in the ARPALombardia dataset [*Dati quantitativi | Brescia | 2023 - ARPA Lombardia* 2025] for the period of interest. Nevertheless I have been able to gather values of recorded piezometric

levels from hydrogeological reports produced by ARPALombardia from 2018 to 2022 [ARPALombardia 2018b, ARPALombardia 2019, ARPALombardia 2020b, ARPALombardia 2021, ARPALombardia 2022]. The referred documents contain water head level measurement in meters, with associated the date and hour of measurment, location and depth of piezometers. Measurements are contained both in text and in the plots. I gathered measurements for seven piezometers in the area of interest.

Data and location of piezometers are shown in Figure 3.15. All but one of the sensors (Vantini) are located in the southern part of the area and record lower absolute values of groundwater levels (above sea level) with respect to the Vantini piezometer. The time variation of the water level reflects the general multi-year hydrological trend of the area. A general increase in groundwater level is recorded from January 2018. This increase reaches the maximum in the year 2021, with water table increase of up to 15 m (Vantini piezometer). Then, a decrease of the level is recorded by the piezometers. The multiyear variation of the water table level is different among the piezometers, generally piezometers located in the north register greater absolute variation of groundwater level both in the first and in the second period.

I computed the multiyear water table variation at each well for the intervals previously defined. Since time series of data end before the 1st August 2022 I computed the water table variation for the drought period as the variation of water table between the 1s January 2021 and the last available data for each piezometer, which is in January 2022. The spatial distribution of the groundwater level variation between the start and the end of these two interval is shown respectively in the left and right panel of Figure 3.16, from which the positive values in the rainy period, the negative values in the drought period, as well as the latitudinal dependency, can be appreciated.

Geological data

Employed geological data include estimates for the geometric and physical characteristics of the terrain layers in the area of interest. This information is contained in reports produced by ARPALombardia [ARPALombardia 2018a, Provinca di Brescia 2016, Provinca di Brescia 2024]. In particular, I analyzed multiple geological cross sections, realized on the basis of stratigraphic data obtained from a network of sampling wells. Two exmples are shown in Figure 3.17.

I used these cross sections to assess the presence of an unconfined aquifer layer and to make a first-order estimate of its geometrical parameters. In particular I am interested in the thickness of the permeable layers of the ground. I made an estimate of these quantity by summing the thicnkess of gravel and conglomerate deposit measured in the cross section in correspondence of the wells (Figure 3.18). To account for errors produced in this estimate of geological parameters I attribuited an uncertainty of 5 m to the estimates obtained in this way (Figure 3.19)

Furthermore, from these reports I collected information about aquifers hydrogeolog-

Well	Surface (m a.s.l.)	Gravel base (m a.s.l.)	Congl. base (m a.s.l.)	Gravel thickness (m)	Congl. thickness (m)
vantini	183.6	130.0	80.0	53.6	50.0
caffaro	146.0	105.0	50.0	41.0	55.0
emporio	131.7	100.0	70.0	31.7	30.0
vivaio	127.8	80.0	50.0	47.8	30.0
caprera	121.5	70.0	50.0	51.5	20.0

Table 3.1: Elevation and thickness of main geological units at selected. All elevations are expressed in meters above sea level (m a.s.l.). Data also shown in Figure 3.18 and 3.19

ical parameters, such as porosity and permeability (table 3.1)

The reconstruction of the terrain structure shows a multilayer aquifer composed of two permeable rock layers differing for their geological characteristics but in hydraulic contact with each other. The terrain altitude shows variations, decreasing towards the southern part from approximately 200 m.s.l. in the valleys outlet to approximately 100 m.s.l. in the plain area. The aquifer layers underground also show a sloped structure deepening towards the south. The thickness of the porous layers generally also decreases towards the south while minor variations are present in the longitudinal direction (Figure 3.17).

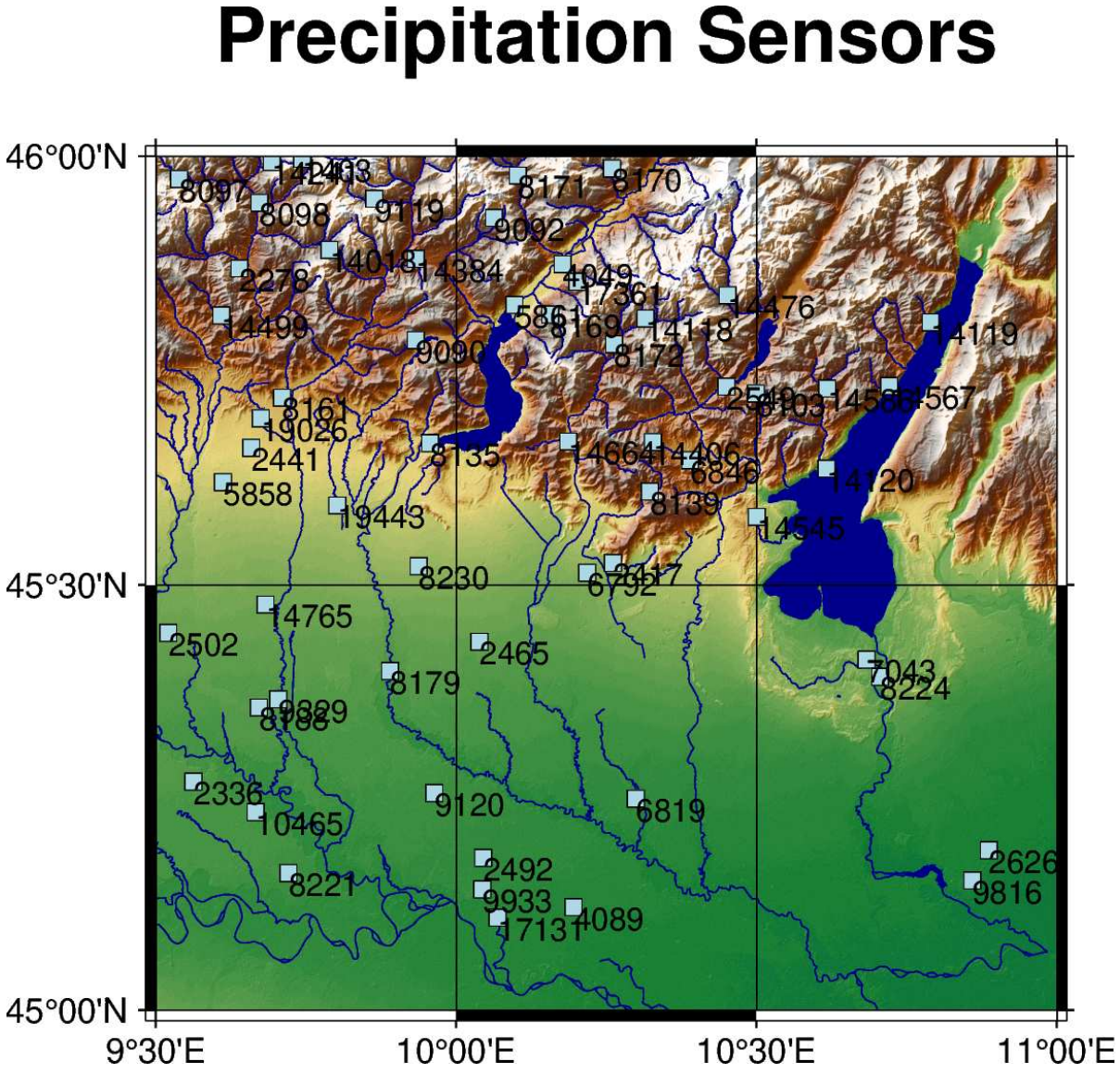


Figure 3.12: Locations of pluviometers of the ARPA monitoring network

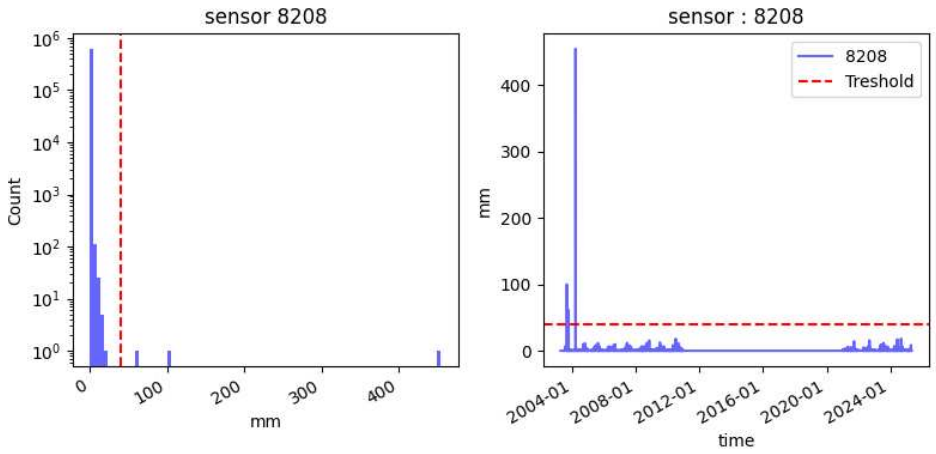


Figure 3.13: Example of unrealistically high values of 10 min cumulated precipitation of sensor 8202.

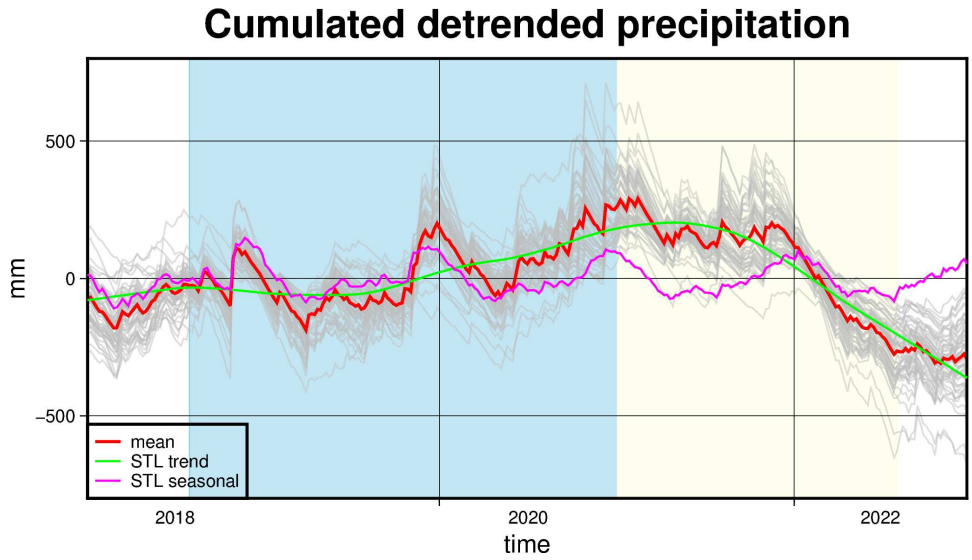


Figure 3.14: Time series of cumulated detrended precipitation. The grey lines represent time series at each sensor,and the red line is the mean time series. To better highlight the contribution of seasonal and multiyear patterns, an STL decomposition (Section 4.2) is performed on the mean time series. The blue and yellow stripe indicate the two intervals T_1 and T_2 respectively associated with rainy and drought conditions

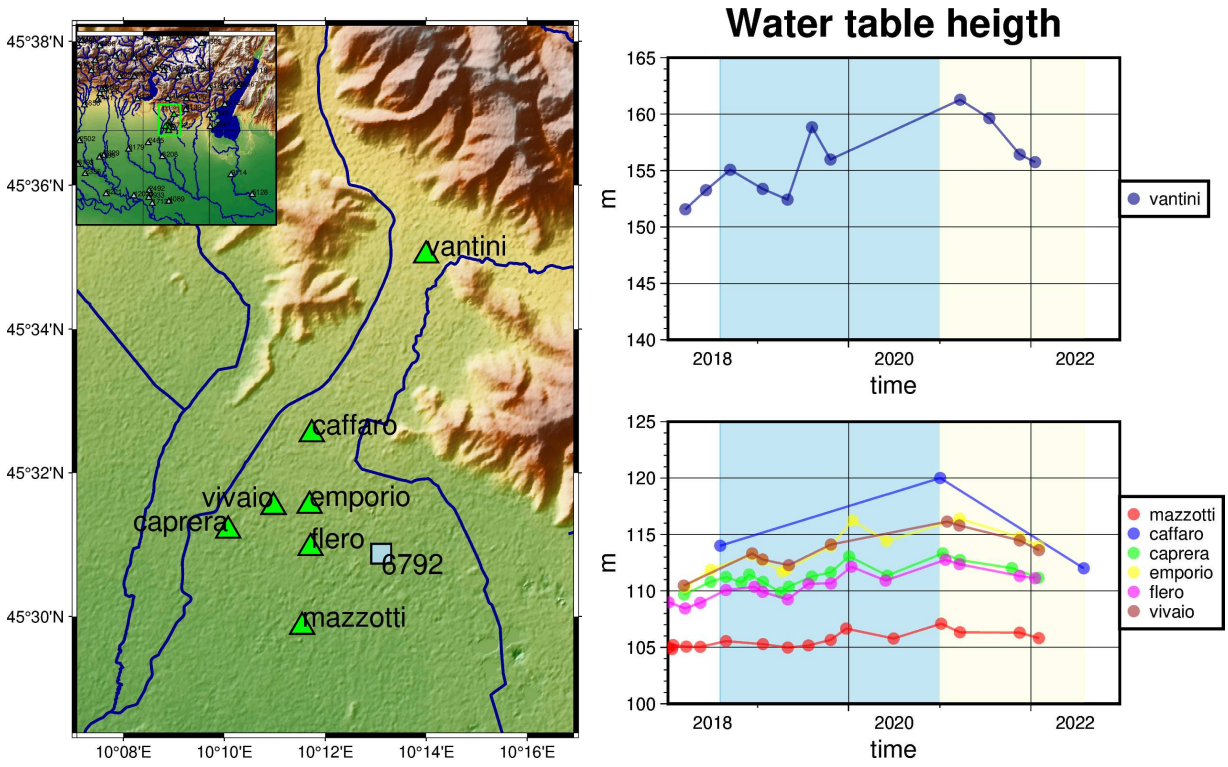
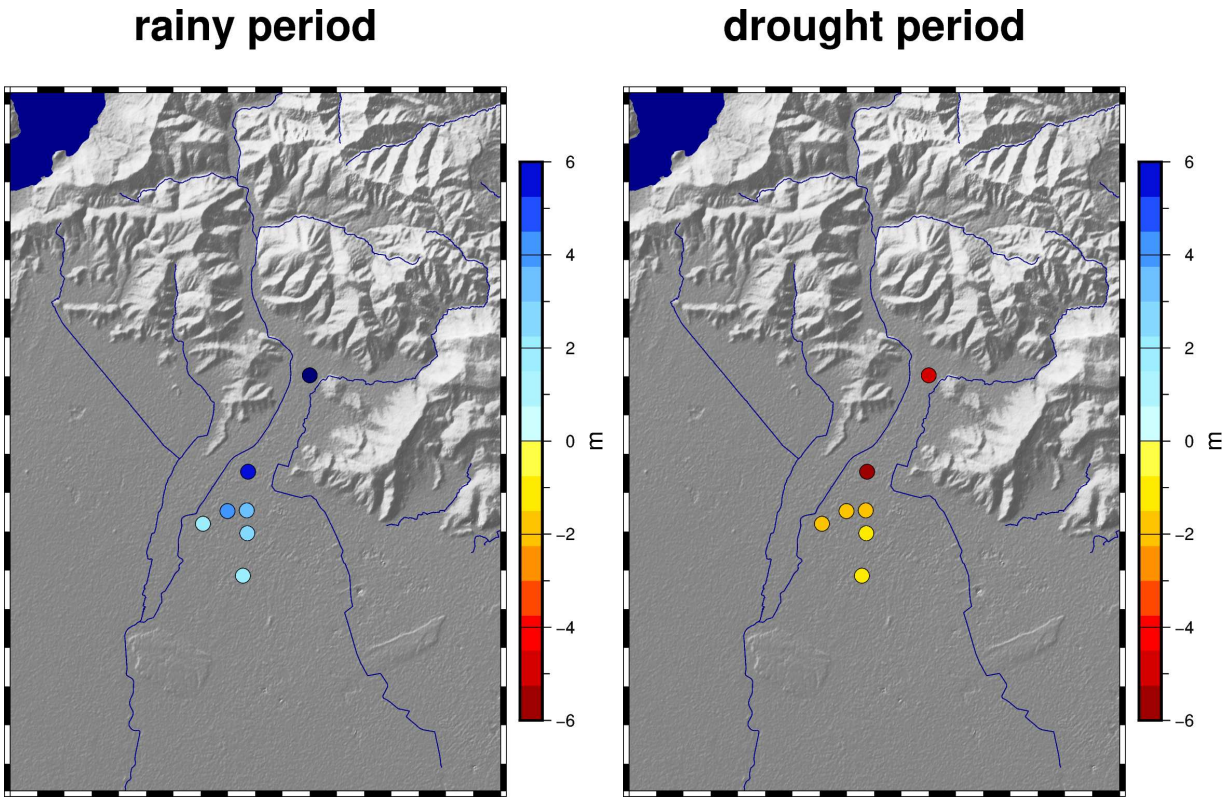
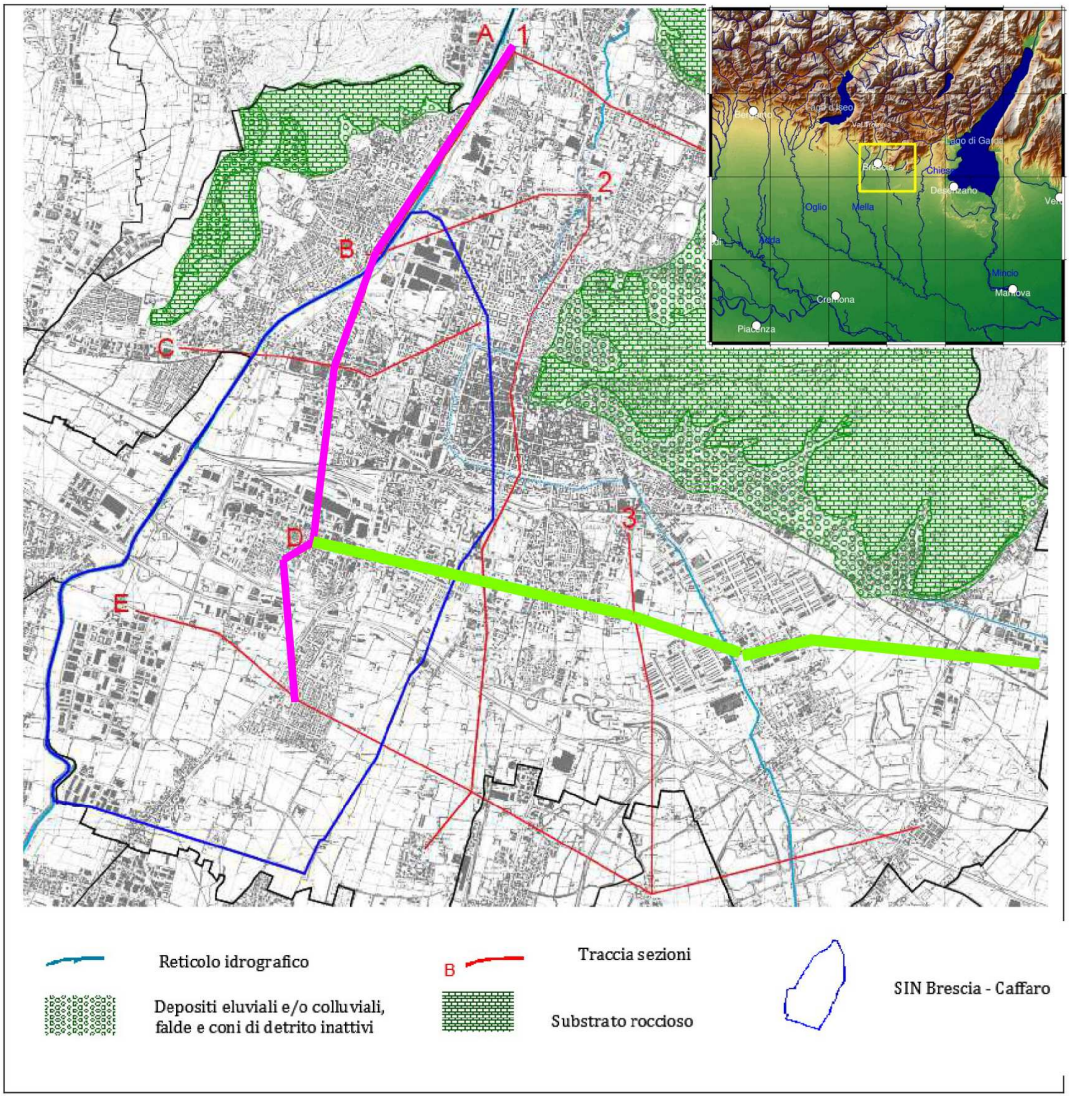


Figure 3.15: Piezometric data time series (right) and piezometric sensor location (left). The blue and yellow stripe indicate respectively the rainy and drought intervals.



(a) (b)

Figure 3.16: Variation of piezometric levels in the rainy (left) and drought (right) interval



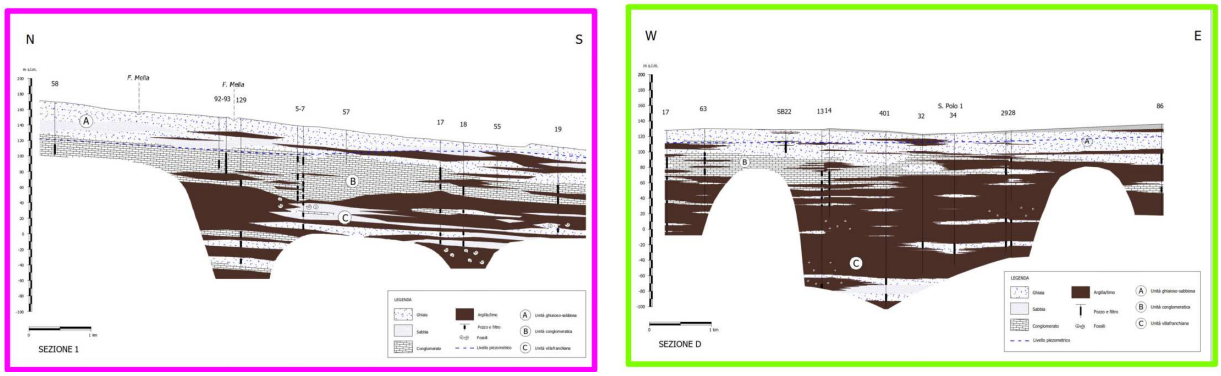


Figure 3.17: Cross sections tracks in the study area (up) Examples of analyzed cross sections (down) [Provinca di Brescia 2015]

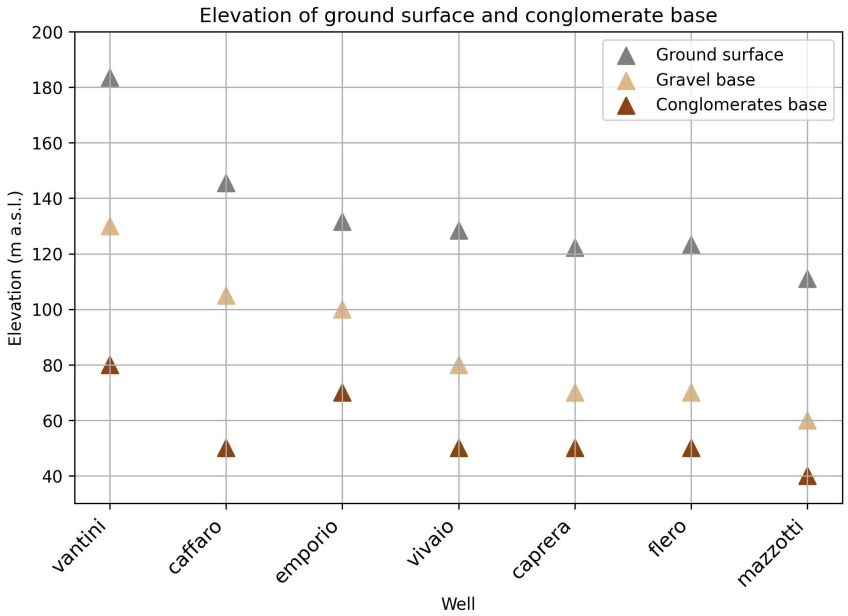


Figure 3.18: Ground surface elevation and depth of gravel and conglomerate layers at selected wells. All elevations are expressed in meters above sea level (m a.s.l.).

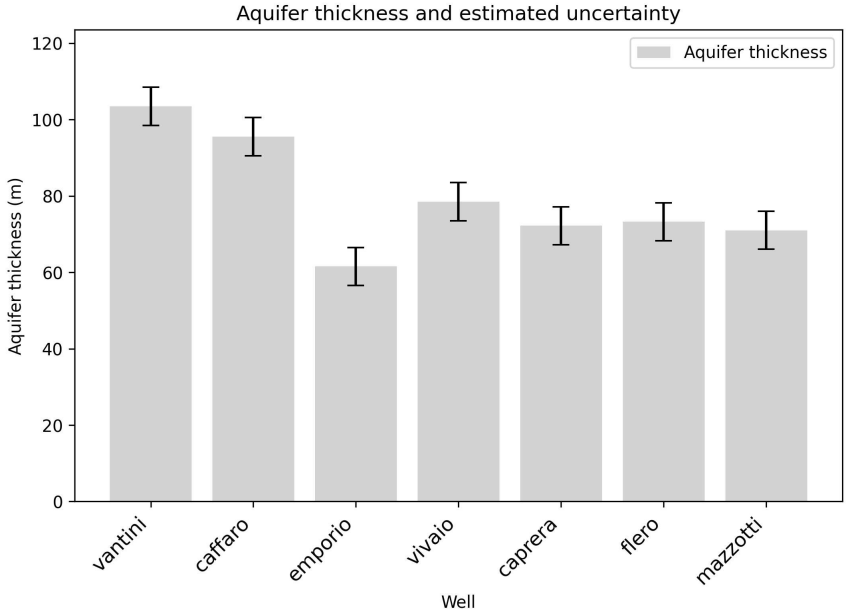


Figure 3.19: Thickness of gravel and conglomerate aquifer units at selected wells.

Chapter 4

Methods for data analysis

Given the high spatial resolution and precision of the InSAR deformation measurement, we expect it to detect the effects of a variety of physical processes. This is particularly true in regions such as the Po Valley, where several geophysical mechanisms influencing surface displacements, such as tectonic movements, natural and anthropogenic subsidence [Farías et al. 2024, Carminati et al. 2002] and elastic rebound due to varying hydrosphere load [Pintori et al. 2024] are known to coexist. Since our goal is to investigate the deformation associated with a specific class of processes — namely those related to continental water dynamics — it is necessary to disentangle the deformation signals associated with different processes.

In this thesis I approach the problem by applying a set of data analysis techniques to characterize the spatiotemporal variability of the EGMS dataset. The final aim is to identify deformation patterns that, through comparison with independent hydro-geological and geographical features or ongoing hydrological forcings, could be related to hydrological processes.

Approaches commonly employed to highlight different deformative processes are generally classifiable in two groups:

- **Parametric approaches**: These approaches aim to find the parameters of a predetermined model of the surface deformation time series. The model is composed by a superposition of analytic functions defined with free parameters. Typically used function are linear trends, sinusoids and exponentials. In these kind of approaches, each time series is individually analyzed to find the best value for these parameters in order to represent the observed data. Parametric approaches have been used to characterize specific components of deformation, as linear trend and seasonal oscillations. They have the advantage of facilitating the association of each contribution to a determined physical process.

- **Data-driven approaches** These approaches do not assume any predefined functional form for the fundamental signals. They instead exploit the statistical properties of

the dataset to identify dominant spatiotemporal patterns of variability [Gualandi et al. 2016]. These approaches are used when we are interested in characterizing the deformation field resulting from a superposition of unknown processes for which we do not have a predetermined model. In this case the shape of fundamental signals results from the analysis and are then interpreted by comparison with the deformation expected from physical processes.

In this work we apply methods belonging to both approaches to obtain o a robust characterization of the ongoing deformation processes detected by the data:

- Identification of long term deformation trend has been performed through a parametric modeling as a degree-one polynomial, whose coefficients, for each time series, have been estimated with least square regression. The estimated linear component has been removed from each time series before the application of the following analysis.

- The multivariate data-driven linear decomposition techniques Principal Component Analysis (PCA) and Independent Component Analysis (ICA) have been applied to the dataset. These two techniques decompose the observed time dependent deformation signal in a linear combination of a discrete number of spatio-temporal components. Both techniques fall in the data-driven approaches as they do not assume any pre-determined form for the spatio-temporal dependence of the deformation components, which are instead characterized by their statistical properties. The two techniques differ in the statistical properties which the components are sought to satisfy. PCA finds a decomposition in terms of components which are mutually *uncorrelated* while ICA enforces the more restrictive condition of statistical *independence*. Both decomposition techniques have been previously used in the study area to isolate components of surface deformation linked to individual physical processes such as natural and anthropogenic subsidence and hydrologically-related deformation [Farías et al. 2024, Nespoli et al. 2021].

- The non-parametric filtering procedure Seasonal Trend Decomposition based on LOESS (STL) [Cleveland et al. 1990] has been applied to each time-series. This techniques is used to isolate a seasonal component and a multi-year trend component with flexibility in selecting the amount of variation contained in each of them [Borsa et al. 2014]. While STL operates on single time series it does not rely on predefined basis functions. Instead, it employs a procedure based on LOESS (Locally Estimated Scatterplot Smoothing) technique to estimate the seasonal and multi-year signals in a data-driven way. The application of STL to the detrended EGMS time series yields two new datasets: one capturing the seasonal variation and one describing the multiyear signal.

- The unsupervised learning technique K-means clustering has been used to classify deformation time series into distinct groups based on the similarity of their temporal evolution. This technique has been proven useful to separate areas dominated by elastic deformation from areas dominated by poroelastic deformation in Central Valley [Kang et al. 2023]. We applied the clustering to the original, detrended data as well as on the output of the STL decomposition in order to identify spatially coherent zones characterized by similar deformation behavior at both seasonal and multiyear temporal scales.
- Computation of the cross-correlation between deformation and precipitation time series has been performed to identify areas were deformation is possibly related to hydrology-related processes.

The employed techniques are described in detail in the following sections.

4.1 Multivariate data-driven decomposition

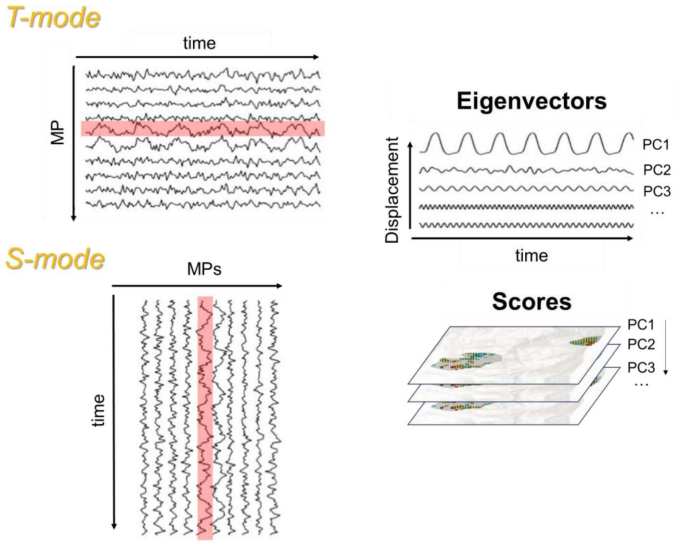


Figure 4.1: Illustration of T-mode and S-mode decompositions highlighting the orientation of the input data matrix. [Rigamonti 2025]

The techniques described in this section belong to the class of multivariate linear decomposition techniques, i.e., they decompose a time- and space-dependent deformation dataset into a set of spatiotemporal components using multivariate statistics. In

this framework, time series of deformation at multiple locations are interpreted as a collection of samples of a multivariate signal, i.e., a vector of random variables whose components – called signal features – are the values of deformation at each time step. This signal reflects the superposition of different processes (e.g., tectonic trends, seasonal deformation, coseismic deformation), which are variably mixed in each of the time series.

The objective of these techniques is to find a set of signals – often referred to as latent components or sources – whose linear combinations can reconstruct the observed dataset. Depending on the specific decomposition used, these components are assumed to be statistically independent or uncorrelated signal. Each of the obsered time series is represented as a weighted sum of these components, with the weights indicating the contribution of each component. If we organise the data in a $N \times M$ matrix where N is the number of observed samples and M the number of related features:

$$\mathbf{X} = \begin{pmatrix} x_1^1 & \dots & x_M^1 \\ \vdots & \ddots & \vdots \\ x_1^N & \dots & x_M^N \end{pmatrix} \tag{4.1}$$

then a linear decomposition of the dataset in term of L components can be written as:

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \tag{4.2}$$

where:

- $S \in \mathbb{R}^{L \times M}$ is the source matrix, and its rows represent the source signals
- $A \in \mathbb{R}^{N \times L}$ is the mixing matrix, containing on the columns the contribution (called score) of each component to the various samples
- $N \in \mathbb{R}^{N \times M}$ collects the data noise

In a dataset containing both temporal and spatial dependency of the deformation, the observed time series can be arranged in two different ways in the data matrix $\mathbf{X}$ (Figure 4.1).

- In the first case (temporal "T-" mode), N represents the number of observations (i.e., of the InSAR points), and M is the time series length. Therefore, each row of $\mathbf{X}$ represents the full time series at a certain point. In this case, we treat every time series as a sample, the sources are temporal functions, and the mixing matrix is given by the associated spatial distributions.
- In the second case (spatial "S-" mode), N represents the time series length and M the number of observations. Therefore, each row of $\mathbf{X}$ represents an InSAR image (deformation map) at a certain time. In this case, we treat every image as a sample, and the sources are images, and the mixing matrix contains the related temporal functions.

The two modes therefore differ in whether we are looking for independent sources in terms of temporal functions (T-mode) or spatial patterns (S-mode). Therefore, we expect T-mode to capture the dominant modes of variation in the temporal evolution of deformation across locations, whereas S-mode to capture the most significant spatial patterns of deformation that change over time across multiple spatial locations. Whether is better to use T- or S- mode mostly depends on the data. It is important to consider that the ability to distinguish between two probability density functions (pdfs) depends on both the number of available samples and the degree to which the pdfs differ across space and time [Adriano Gualandi et al. 2021]. Since we are dealing with InSAR data, characterized by a higher number of spatial samples (InSAR points) than temporal ones (time series length), we applied the S-mode decomposition.

In this work we applied to the dataset two linear decomposition techniques sharing the decomposition form 4.2: Principal Component Analysis (PCA) and Independent Component Analysis (ICA).

In PCA the decomposition is performed by reprojecting the data - i.e. each of the observed signal - onto a new reference frame in which the variables are uncorrelated. This is achieved choosing the axes of the reference frame oriented in the direction of the eigenvectors of the covariance matrix of the data. Thus the reprojection is done on an orthogonal reference frame in which each of the basis vector is associated with a value of the variance and it is uncorrelated with the others. The new axes represent the latent components of the decomposition and each sample will be explained by a linear combination of these component with weights given by the projections of that sample.

Using just the order-two statistical moment of the data (the covariance), PCA implies the assumption that the data projected onto the extracted components are normally distributed. However, the probability density function distributions for some of the sources most commonly present in deformation signals recorded through geodetic techniques are often not Gaussian (such as a linear trend, annual signals, and post-seismic signals, Figure (4.2)). Therefore, in the case of a mix of such sources, the components extracted through PCA would represent only a combination of the true physical sources beneath the observed data, not having any physical meaning if individually taken.

A more stringent constraint, and potentially more suitable for the research of physical signals, is the independence among the components, that is the condition imposed by ICA [Gualandi et al. 2016]. This multivariate technique still remains in the field of linear decompositions (i.e. assumes that the signals from the different sources are linearly combined). However, the data are projected onto a system of coordinates where each component is no longer constrained to be orthogonal to another one, and more statistical moments are used to represent the underlying distribution of the observations. This allows for a higher flexibility in the representation of source signals.

Different ICA algorithms exist based on how the statistical independence between the source signals is achieved. The widely used FastICA algorithm [Hyvärinen et al. 2000] is based on the central limit theorem, which states that the superposition of independent

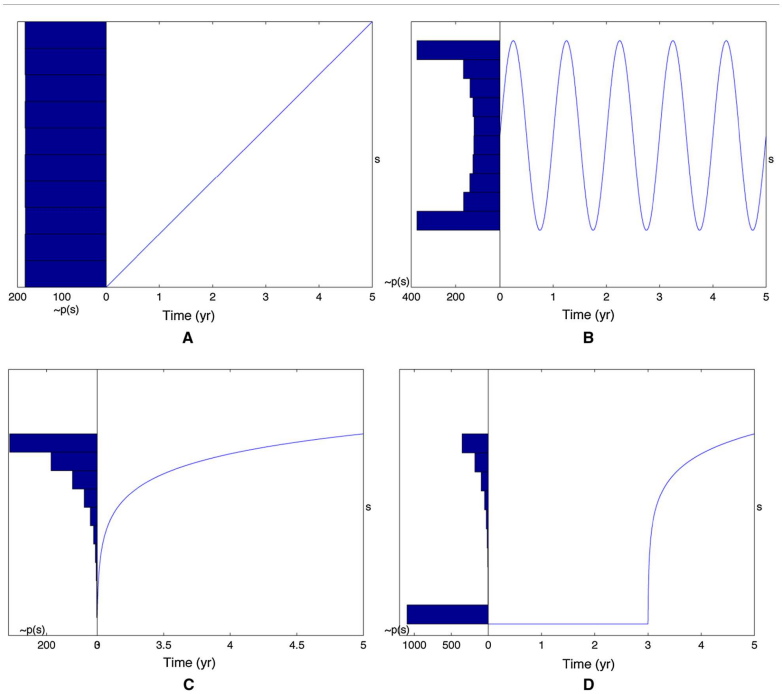


Figure 4.2: Probability density function associated with typical geodetic signals show significant deviations from Gaussianity [Gualandi et al. 2016]

variables follows a distribution which is more similar to a Gaussian distribution than the original ones. FastICA therefore extracts the independent components by maximizing the non-Gaussianity of the extracted sources.

An alternative algorithm, the variational bayesian ICA (vbICA, Gualandi et al. (2016)) employs a generative approach and it explicitly models the pdf of the reconstructed source signals as a mix of Gaussians to better encompass multimodal behaviours. The advantage of the vbICA approach consists in having a natural framework to describe multimodal pdfs for the sources and to deal with missing data (a common problem in geodetic time series). In this thesis, we employed the vbICA approach, which has proven to generally achieve a better performance in similar contexts compared to FastICA algorithm (e.g., Gualandi et al. (2016), Figure 4.3).

In the following two sections I describe the theory behind the two employed techniques: PCA and vbICA.

4.1.1 Principal Component Analysis (PCA)

The goal of Principal Component Analysis is to find a new coordinate system in which the data can be represented as uncorrelated components, ordered by the amount of variance they explain. This is achieved by reprojecting the original dataset onto an

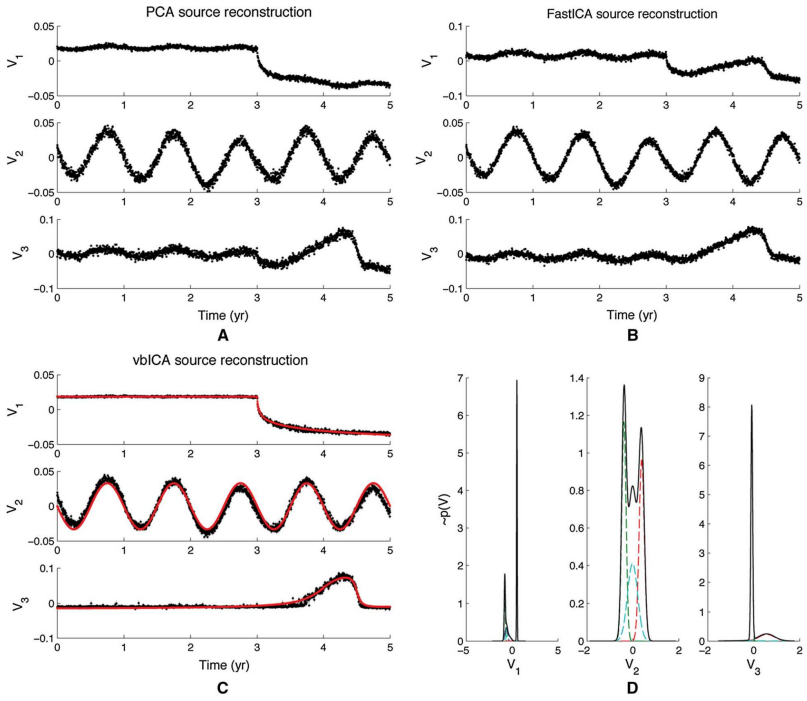


Figure 4.3: Comparison of the results of a synthetic test performing a decomposition with PCA, FastICA and vbICA. The red lines represent the sourcesignals, whose corresponding pdf are displayed in the plot on the bottom right [Gualandi et al. 2016]

orthogonal basis. In this transformed space, the directions (or components) are such that the first one accounts for the maximum variance, the second is orthogonal to the first and accounts for the largest remaining variance, and so on.

Let $\mathbf{X} \in \mathbb{R}^{N \times M}$ be the data matrix, arranged in S-mode: each of the N rows corresponds to a deformation map (i.e., one observation in time), and each of the M columns corresponds to a spatial point (i.e., a time series).

PCA is based on the analysis of the covariance matrix $C_{\mathbf{X}} \in \mathbb{R}^{N \times M}$, which quantifies how the signal features vary together. $\mathbf{C_X}$ is defined as:

$$\mathbf{C_X} = E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^T] \tag{4.3}$$

and can be estimated as the sample covariance, which, assuming to have centered the dataset so that each column has zero mean, has the form:

$$\mathbf{C_X} = \frac{1}{1-N}\mathbf{X}^T\mathbf{X} \tag{4.4}$$

This matrix contains on the diagonal the variance of each feature and off diagonal the pairwise covariances between spatial points across time. If the off-diagonal terms are non-zero, it means the corresponding features (spatial locations) are statistically correlated.

The new basis in which features are uncorrelated is found by computing the eigenvalue decomposition of $\mathbf{C_X}$

$$\mathbf{C_X} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top \tag{4.5}$$

where:

- $U \in \mathbb{R}^{M \times M}$ is the matrix containing on the columns the eigenvectors of C (orthonormal),
- $\Lambda \in \mathbb{R}^{M \times M}$ is a diagonal matrix collecting the eigenvalues λ_i $i = 1, ..., M$ which represent the variance explained by each component.

We then express the transformed variables as:

$$\mathbf{Y} = \mathbf{X}\mathbf{U} \tag{4.6}$$

Each column of $\mathbf{Y}$ contains the projection of a sample of the signal onto one principal direction. The set of projection is a new multivariate signal with N samples characterized by a diagonal covariance matrix:

$$\mathbf{X_Y} = \frac{1}{N-1}\mathbf{Y}^T\mathbf{Y} = \frac{1}{N-1}\mathbf{U}^T\mathbf{X}^T\mathbf{X}\mathbf{U} = \mathbf{U}^T\mathbf{C}\mathbf{U} = \mathbf{\Lambda} \tag{4.7}$$

representing the fact that the new variables are mutually uncorrelated and to each is associated a value of variance given by the corresponding eigenvector.

A decomposition of the form 4.2 in terms of the defined component is found from the Singular Value Decomposition (SVD) of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^T \tag{4.8}$$

In which:

- $\mathbf{U} \in \mathbb{R}^{M \times M}$ contains the right singular vectors, which are the eigenvectors of the matrix $\mathbf{X}^T\mathbf{X}$.
- $\mathbf{V} \in \mathbb{R}^{N \times N}$ contains the left singular vectors, which are the eigenvectors of the matrix $\mathbf{X}\mathbf{X}^T$.
- $\Sigma \in \mathbb{R}^{N \times M}$ is a diagonal matrix containing the square root of the non-zero eigenvalues of the covariance matrix λ_i, called singular values:

From comparison between equation 4.6 and equation 4.8 in fact we can see that:

$$\mathbf{Y} = \mathbf{X}\mathbf{U} = \mathbf{V}\mathbf{\Sigma} \tag{4.9}$$

the principal components are the left singular vector $\mathbf{V}$ scaled by the singular values. In the obtained form (Equation 4.8) the decomposition can be related to the generic form of a linear decomposition. In this case the source matrix is the matrix $\mathbf{U}^T$, while the mixing matrix is the $\mathbf{Y}$ matrix. In s-mode decomposition the first contain on each row the spatial functions associated to the principal components, the second contains on each column the associated temporal functions.

4.1.2 Variational Bayesian Independent Component Analysis

In this work, ICA was performed using the Variational Bayesian ICA algorithm (vbICA) [Gualandi et al. 2016]. This algorithm employs a generative framework in which the mixing process of L independent sources is simulated with a parametric model in which the probability density functions (pdfs) of each source signal are modeled as a Mix of Gaussians (MoG).

The model is characterized by a set of observable variables $\mathbf{X}$ and hidden variables $\mathbf{H}$. Observable variables are the data while hidden variables characterize the sources and the mixing processes. The pdfs of both observable and hidden variables are defined in a parametric way with a set of model parameters $\mathbf{\Theta}$. Both model parameters and hidden variables are unknown and they are indicated as model 'weights' $\mathbf{W} = \{\mathbf{H}, \mathbf{\Theta}\}$. The goal of the generative approach to ICA is to find the best values for the weights that can explain the data $\mathbf{X}$ under the framework of linear combination of independent source signals. At this scope, the independence of the sources is enforced in the parametrization of their pdfs, then the algorithm makes an estimate of the whole set of weights through the computation of their posterior probability distribution, given the observed data $p(\mathbf{W}|\mathbf{X})$.

In a Bayesian framework, given a generative model $\mathcal{M}$ and the observed data $\mathbf{X}$, maximising the posterior pdf over weights $\mathbf{W}$ given the data $\mathbf{X}$ is the best choice for $\mathbf{W}$:

$$p(\mathbf{W}|\mathbf{X},\mathcal{M}) = \frac{p(\mathbf{X}|\mathbf{W},\mathcal{M})\, p(\mathbf{W}|\mathcal{M})}{p(\mathbf{X}|\mathcal{M})} \tag{4.10}$$

where $p(\mathbf{X}|\mathbf{W},\mathcal{M})$ is the likelihood of the data under the generative model $\mathcal{M}$, $p(\mathbf{W}|\mathcal{M})$ is the prior distribution of the weights in the model, and the denominator is a normalizing factor called the *model evidence*:

$$p(\mathbf{X}|\mathcal{M}) = \int p(\mathbf{X}|\mathbf{W},\mathcal{M})\, p(\mathbf{W}|\mathcal{M})\, d\mathbf{W} \tag{4.11}$$

Variational Approximation

Since the expression for the model evidence (equation 4.11) contains an integral which is generally intractable, as it spans the entire weight space, in the vbICA algorithm the computation of the posterior is performed using the variational approach proposed by Choudrey (2002). In this approach instead of directly computing the posterior probability, we introduce an approximating pdf $p'(\mathbf{W})$ whose parameters are refined in order to minimize its distance from the true posterior [Gualandi et al. 2016].

Given an arbitrary pdf over the weight space $p'(\mathbf{W})$, a measure of the distance between $p'(\mathbf{W})$ and the posterior $p(\mathbf{W}|\mathbf{X})$ is given by the the Kullback-Leibler (KL)-divergence [Kullback et al. 1951]:

$$\mathrm{KL}(p'(\mathbf{W})\|p(\mathbf{W}|\mathbf{X})) = \int p'(\mathbf{W}) \ln \frac{p'(\mathbf{W})}{p(\mathbf{W}|\mathbf{X})}\, d\mathbf{W}. \tag{4.12}$$

If we combine equation 4.12 and equation 4.10 we obtain the expression for the KL-divergence.

$$\begin{aligned}
&\mathrm{KL}(p'(\mathbf{W})\|p(\mathbf{W}|\mathbf{X})) = \\
&\quad = \int p'(\mathbf{W}) \ln \frac{p'(\mathbf{W})}{p(\mathbf{W}|\mathbf{X})} d\mathbf{W} = \int p'(\mathbf{W}) \ln \frac{p'(\mathbf{W})p(\mathbf{X})}{p(\mathbf{W},\mathbf{X})}\, d\mathbf{W} = \\
&\quad = \int p'(\mathbf{W}) \ln \frac{p'(\mathbf{W})}{p(\mathbf{X},\mathbf{W})} d\mathbf{W} + \int p'(\mathbf{W}) \ln p(\mathbf{X}) d\mathbf{W} = \\
&\quad = -\int p'(\mathbf{W}) \ln \frac{p(\mathbf{X},\mathbf{W})}{p'(\mathbf{W})} d\mathbf{W} + \ln p(\mathbf{X})
\end{aligned}$$

where the dependency from the model $\mathcal{M}$ has been dropped for conciseness.

The integral quantity in the last expression is called the Negative Free Energy (NFE).

$$\mathrm{NFE}[\mathbf{X}] = \int p'(\mathbf{W}) \ln \frac{p(\mathbf{X},\mathbf{W})}{p'(\mathbf{W})} d\mathbf{W} \tag{4.13}$$

We see that, since $\ln p(\mathbf{X})$ does not depend on the weights, maximizing the Negative Free Energy with respect to the distribution $p'(\mathbf{W})$ is equivalent to minimize the KL-divergence between $p'(\mathbf{W})$ and the true posterior $p'(\mathbf{W}|\mathbf{X})$.

To apply the variational approximation, it is necessary to choose a form for the approximating pdf $p'(\mathbf{W})$ which allows to make the maximization problem tractable. The most common choice for $p'(\mathbf{W})$ is that it factorizes in:

$$p'(\mathbf{W}) = \prod_{i}^{k} p'_i(\mathbf{w}_i) \tag{4.14}$$

for some partitions $\{\mathbf{w}_i, ..., \mathbf{w}_k\}$ of the set of weights [Gualandi et al. 2016, Ormerod et al. 2010].

With this form for $p'(\mathbf{W})$, Choudrey (2002) showed that the maximization of the Negative Free Energy results in explicit expressions for each marginal posterior $p'_i(\mathbf{w}_i)$ in terms of the others. This in turns leads to an iterative procedure to find the joint distribution $p'(\mathbf{W})$ solution of the maximization problem [Gualandi et al. 2016].

Model specification

In the vbICA model it is necessary to define the distribution of the variables used in the decomposition of equation (4.2), $\{\mathbf{A}, \mathbf{S}, \mathbf{N}\}$ and of the other hidden variables required in the definition of their pdf. The definition depends on a set of parameters, and, for each of these parameters, as well as for the hidden variables, it is necessary to define a prior distribution.

Additionally, the variational approach requires to choose the form for the approximating marginal densities $p'_i(\mathbf{w_i})$ to use in equation 4.14. However, Choudrey (2002) showed that there is no need to specify the functional form for the marginal densities if the corresponding prior distribution of the weights $p(\mathbf{w_i})$ and data likelihood $p(\mathbf{X}|\mathbf{w_i})$ are modeled as distribution belonging to conjugate families, i.e. classes of distribution which multiplied give distribution belonging to the same class. In this case, the expressions for the marginal posteriors $p'(\mathbf{w}_i)$ which maximize the NFE are analytical and the distributions have a defined functional form. The parameters characterizing the distributions $p'(\mathbf{w}_i)$ can in general be expressed as functions of the parameters of the corresponding prior $p(\mathbf{w}_i)$, of the data, and of the expected values of hidden variables under the other posterior [Choudrey 2002]. Then, after the specification of the model, a set of updating equations can be derived that allow us to compute the parameters that govern the posterior densities [Gualandi et al. 2016].

The variable pdfs are specified as follows:

- Each source component $\mathbf{s}_i$ with $i = 1, ..., L$ is modeled as a Mix of Gaussians (MoG). The parameters defining the mix are the number of participating Gaussians m_i and for each of them a value of the mean $\mu_{i,j}$, the variance $\beta_{i,j}$ and the

weight given in the mix $\pi_{i,j}$ called mixing proportions with $j = 1, .., m_i$. For each source, the set of parameters is indicated as $\boldsymbol{\theta}_i = \{\boldsymbol{\mu}_i, \boldsymbol{\beta}_i, \boldsymbol{\pi}_i\}$ for $i = 1, ..., L$ where $\boldsymbol{\mu}_i = \{\mu_{i,1}, ..., \mu_{i,m_i}\}$, and similarly $\boldsymbol{\beta}_i$, $\boldsymbol{\pi}_i$ are vectors of size m_i. The complete set of source parameters is indicated as $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\pi}\}$ where $\boldsymbol{\mu} = \{\boldsymbol{\mu}_i, ..., \boldsymbol{\mu}_L\}$ and similarly $\boldsymbol{\beta}$, $\boldsymbol{\pi}$.

To complete the source model, an indicator variable called source state $\mathbf{q}_i$ is used to select which Gaussian of the i-th source is used to generate $\mathbf{s}_i$. For each i, $\mathbf{q}_i$ is a vector of the same size of the sources vector and whose components q_i^k for $k = 1, ..., M$, take the values $\{1, .., m_i\}$ with probability given by the mixing proportions π_{i,q_i^k}. The complete collection of source states for all components is indicated with $\mathbf{q} = \{\mathbf{q}_1, ..., \mathbf{q}_L\}$.

- The mixing matrix elements are assumed to follow a zero-mean Gaussian distribution with column-dependent precision α_i. The set of variances for all component is indicated with $\boldsymbol{\alpha} = \{\alpha_i, ..., \alpha_L\}$

- The noise variables are modeled as a Gaussian with zero mean and precisions: $\boldsymbol{\Lambda} = \{\Lambda_1, ..., \Lambda_M\}$

The chosen factorization of the variational distribution $p'(\mathbf{W})$ over the model weights $\{\mathbf{A}, \mathbf{S}, \boldsymbol{\Lambda}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{q}\}$ is therefore [Gualandi et al. 2016]:

$$p'(\mathbf{W}) = p(\mathbf{A})\, p(\boldsymbol{\Lambda})\, p(\mathbf{S}|\mathbf{q})\, p(\boldsymbol{q})\, p(\boldsymbol{\mu})\, p(\boldsymbol{\beta})\, p(\boldsymbol{\pi}) \tag{4.15}$$

Statistical independence between the sources is enforced by further factorizing the variational posterior over $\mathbf{S}$ as:

$$p(\mathbf{S}) = \prod_{i=1}^{L} p(\mathbf{s}_i) \tag{4.16}$$

where L is the number of independent components used in the decomposition.

Finally it is necessary to define prior distributions over all model parameters $\{\Lambda, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\pi}\}$. On the basis of the consideration in the previous section the priors of the model parameter are chosen from conjugate families of the corresponding variable pdfs. In particular:

- The **component means** $\boldsymbol{\mu}$ of the MoG are modeled as independent Gaussian variables of given mean m_{i0} and variance τ_{i0} for $i = 1, ..L$

$$p(\boldsymbol{\mu}) = \prod_{i=1}^{L} \prod_{q_i=1}^{m_i} \mathcal{N}(\mu_{i,q_i}; m_{i0}, \tau_{i0})$$

- The **component precisions** $\boldsymbol{\beta}$ are modeled using Gamma distributions with parameters $b_{i0}, c_i 0$, $i = 1, ..., L$:

$$p(\boldsymbol{\beta}) = \prod_{i=1}^{L} \prod_{q_i=1}^{m_i} \Gamma(\beta_{i,q_i}; b_{i0}, c_{i0})$$

- The **mixture proportions $\boldsymbol{\pi}$** associated with the MoG for each source component are assumed to follow symmetric Dirichlet distribution with parameters λ_i, $i = 1, .., L$

$$p(\boldsymbol{\pi}) = \prod_{i=1}^{L} \mathrm{Dir}(\pi_i; \lambda_{i0})$$

- The **mixing matrix precisions $\boldsymbol{\alpha}$** are modeled with a product of Gamma distributions with parameters $b_{\alpha i}, c_{\alpha i}$, $i = 1, ..., L$:

$$p(\boldsymbol{\alpha}) = \prod_{i=1}^{L} \prod_{q_i=1}^{m_i} \Gamma(\alpha_i; b_{\alpha_i}, c_{\alpha_i})$$

- The **noise precision $\boldsymbol{\Lambda}$** is modeled as a Gamma distribution with parameters $b_{\Lambda i}, c_{\Lambda i}$, $i = 1, ..., L$:

$$p(\boldsymbol{\Lambda}) = \prod_{j=1}^{M} \Gamma(\Lambda_j; b_{\Lambda_j}, c_{\Lambda_j})$$

These distributions depends on the values of a set of hyperparameters which have to been choosen by the user. In this work we folllow the default settings which correspond to weakly informative priors, allowing the data to reveal their internal structure [Gualandi et al. 2016].

Selection of number of components

In the ICA technique the number L of source components is a parameter specified by the user. The selection of the correct value for L is thus a crucial step in the application of the technique since a suboptimal choice could yield a decomposition in which component are not well separated, with signal leaking from one component to another, rendering the interpretation of the results difficult.

Even thought ICA does not assign a relevance to each component, like PCA does by ranking the components by explained variance, the employed variational Bayesian approach allows to adopt a method known as Automatic Relevance Determination (ARD) [Gualandi et al. 2016]. This method characterizes the relevancy of the components through the inferred values of the mixing matrix precision parameters α_i. Each α_i in fact defines how strong is the assumption that the mean value of the corresponding column of the mixing matrix is zero. A large value of α_i corresponds to a posterior over the mixing matrix column i dominated by the prior density, effectively setting the elements of the column i to zero. Since each column of the mixing matrix is associated with a source components, the α_i values inferred by the model characterize the relevance of the corresponding component in the decomposition. Then, in each decomposition,

we can directly compare the α_i values associated with each component to assess if some are more relevant than the other, and eventually identify the less relevant as noise and discard it.

In this work, we selected the number of components by repeating the decomposition for various number of components. For each case, we computed and plotted the ratio between the maximum and the minimum α_i values as a function of the number of components. This ratio generally decreases as more components are introduced, reflecting the progressive inclusion of less relevant components into the model. We selected the decomposition in which this ratio becomes smaller than one tenth, identifying the component characterized by the greatest value of α_i as noise and discarding it.

4.2 STL decomposition

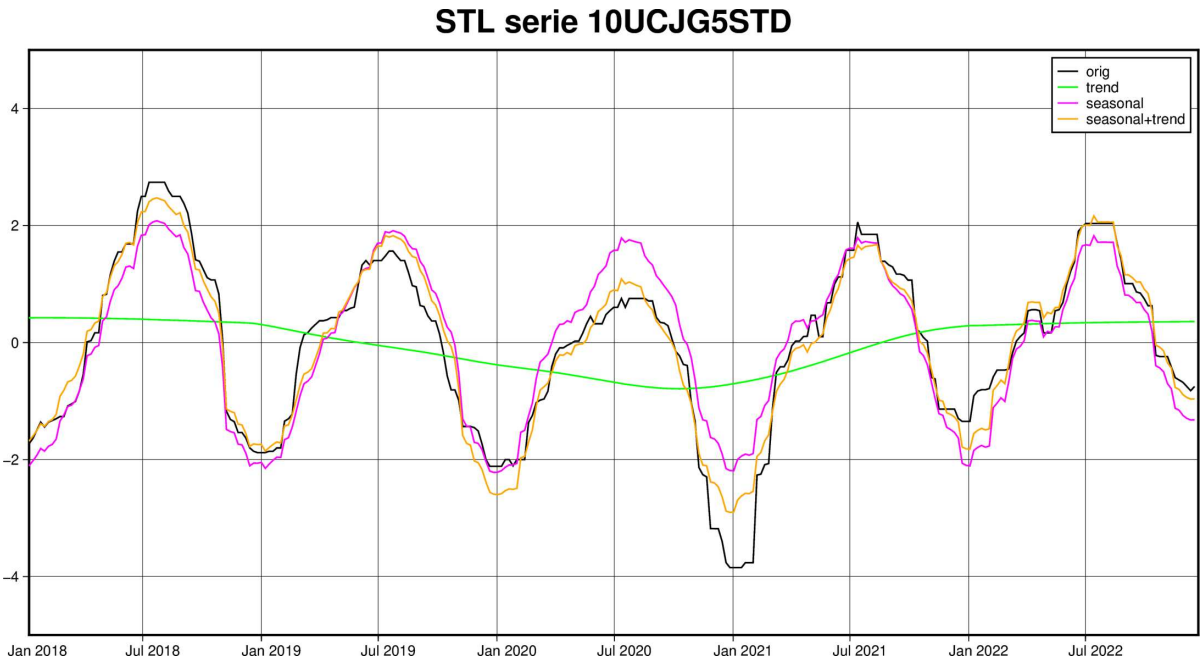


Figure 4.4: Example the decomposition result for a selected displacement time series.

STL is a decomposition technique which is applied to a single time series. This procedure decomposes the input time series in a *seasonal* component, a *trend* component, and a *residual* component (Figure 4.4) [Cleveland et al. 1990]. In this work, I employ STL to characterize the deformation field in terms of seasonal and multi-year deformation trends, which could be potentially linked to variations of the hydrological forcings at these two timescales.

Supposing we denote a time series defined with N epochs t_i as Y_i with $i = 1, .., N$, the STL technique obtains an additive decomposition of Y_i of the form:

$$Y_i = T_i + S_i + R_i \tag{4.17}$$

In which the T_i, S_i, R_i for $i = 1, .., N$ indicates the time series respectively representing the seasonal, trend, and residual components. The components time series are retrieved applying a sequence of filtering operations to the original time series Y_i, all but two of them procedure employ the Locally Estimated Scatterplot Smoothing (LOESS) smoother [Cleveland et al. 1990].

In this section I describe the theory behind the LOESS smoother and the steps performed by the STL algorithm.

4.2.1 LOESS

LOESS (locally estimated scatterplot smoothing) is a procedure which obtains a regression curve from a time series by fitting a linear or quadratic function of the independent variable locally, in a moving fashion, analogous to the way in which a moving average is computed .

La procedure takes as input a time series Y_i with values relative to N epochs t_i, $i = 1, .., N$ and [restituisce] a smoothed time series Y'_k defined in an arbitrary set of epochs t'_k, $k = 1, .., N'$.

Each Y'_k is obtained performing a weighted polynomial regression of the time series Y_i and computing the result at the time t'_k. In this regression, the values Y_i are assigned weights that depend on their temporal distance from t'_k so that observations corresponding to epochs t_i closer in time to t'_k, have greater influence on the fitted value. Repeating the regression for each t'_k we obtained a curve which locally represent a regression of the time series Y_i.

To regulate the effective smoothing window an integer parameter q, called 'smoothing span', is introduced. If $q < N$ we define the quantity $\lambda_q(t'_k)$ as the time distance between t'_k and the q-th nearest epoch of the time series Y_i. If $q > N$ the definition is scaled as:

$$\lambda_q(t'_k) = \frac{q}{N}\lambda_n(t'_k) \tag{4.18}$$

Then, when computing the regression curve for the value Y'_k the weights $w_i^{(k)}$ given to each Y_i in the fit are:

$$w_i^{(k)} = W\left(\frac{|t_i - t'_k|}{\lambda_q(t'_k)}\right) \tag{4.19}$$

in which the function $W(u)$ is the tricube weight function:

$$W(u) = \begin{cases} (1-u^3)^3 & \text{if } |u| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.20}$$

The LOESS procedure can be easily extended in the presence of a reliability measure, or robustness weights, ρ_k associated with the observations. In this case, the weights $w_i^{(k)}$ are modified as:

$$\hat{w}_i^{(k)} = \rho_k w_i^{(k)} \tag{4.21}$$

4.2.2 STL

I employ the implementation of STL contained in the `statsmodel.tsa` python package [Perktold et al. 2024], derived from the original fortran implementation developed by Cleveland et al. (1990). The result of the decomposition for an EGMS time

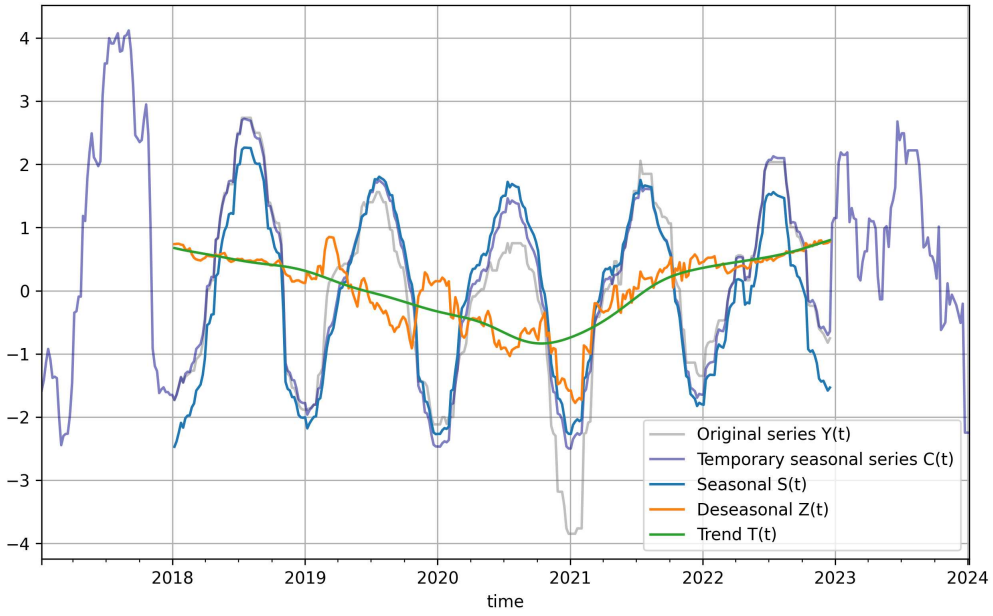


Figure 4.5: Intermediate time series computed in an iteration of the inner loop

series is shown in Figure 4.4 and the intermediate series produced in the decomposition procedure are shown in Figure 4.4.

The decomposition procedure depends on a set of parametrs which regulate the size of the filtering windows employed, effectively determing the amount of variation to be attributed respective to the seasonal and trend component. These parameters are synthetized in the table 4.1 along with their meaning and effect on the decomposition. Also reported are the values employed in this work. The most important one is the number of observation in a seasonal cycle, which determines the frequency of the seasonal component to isolate.

In the next section I synthetize the fundamental steps performed by the algorithm to decompose an example time series Y_i, $i = 1, ..., N$ while in the following section I discuss the employed parameters and their effects.

Algorithm overview

The STL algorithm structure is iterative, performing a loop which refines the estimates of the components in each iteration. Additionally, the algorithm can perform and outer loop which repeats the decomposition refining the weights given to each value in the series. This is done in order to obtain a result which is robust in case of data showing

an outlier behaviour.

A complete run contains the following steps:

1. **Initialization** Performed only at the beginning. The trend component is initialized as $T_i = 0$ and the robustness weigths as $\rho_i = 0$ for each $i = 1, ..., N$.

2. **Inner loop**

   (a) A detrended time series $D_i = Y_i - T_i$, for $i = 1, ..., N$, is computed using the trend T_i of the previous iteration.

   (b) Given an integer parameter n_p representing the number of observation contained in a seasonal cycle, n_p seasonal subseries C_j^p are extracted, with $p = 1, ..., n_p$ and j going from 1 to the number of seasonal cycle spanned by the time series. These are composed by grouping all observations with the same position in the cycle (E.g. all January values for annual observations) and ordering them by the index of the seasonal cycle - E.g. the year to which they belong. They are referred as *cycle subseries*.

   (c) Each cycle subseries is smoothed using LOESS with smoothing span defined by an integer parameter n_s. The smoothed time series are computed at all times and are additionally extended one period of the seasonal cycle before the start and after the end.

   (d) The smoothed cycle subseries are recomposed and the resulting series is indicated with C_i (purple line in Figure 4.5). C_i is a temporary seasonal series and contains $N + 2n_p$ times due to the additional values computed in the previous step.

   (e) To prevent low-frequencies to enter the seasonal component then an high pass filter is applied to the series C_i. To do this, two running averages of length n_p, a running average of length 3, and a LOESS smoothing with span defined by a parameter n_l, are applied to C_i. Each running average removes n_p points from the series so the resulting series is defined at the same times of the original one. The result is indicated with L_i and it contains a low frequencies signal.

   (f) The L_i series is subtracted from C_i obtaining the estimation for the seasonal component S_i (blue curve in Figure 4.5).

   (g) A deseasonalized time series is computed $Z_i = Y_i - S_i$, $i = 1, ..., N$.

   (h) The trend component T_i of the current iteration is found smoothing of the deseasonalized Z_i series with a loess span define by the parameter n_t. This component is used in the following iteration of the loop.

   After a number of iteration sufficient to reach convergence of the component estimates the decomposition residuals are computed as:

$$R_i = Y_i - S_i - T_i \quad i = 1, .., N \tag{4.22}$$

Parameters	meaning	influences:	value
$\mathbf{n_p}$	number of step in a seasonal cycle	frequency of the seasonal component	61
$\mathbf{n_s}$	seasonal smoothing span	amount of interannual variation of seasonal component	7
$\mathbf{n_l}$	low-pass smoothing span	low frequencies power in the seasonal component	63
$\mathbf{n_t}$	trend smoothing span	time scale of variation in the trend component	117

Table 4.1: Parameters used in the decomposition along with selected values and effect on the decomposition

3. **Outer loop (robustness iterations)**

   By setting on the optional robust parameter, the algorithm performs a correction of the decomposition after the estimation of the residuals. As explained by Cleveland et al. (1990) this has the purpose to reduce the sensibility of the decomposition to points showing abherrant behaviours (outliers). In fact, expecting that these points will be characterized by higher values of the residuals, the robustness weight are updated in this step as:

   $$\rho_k = B(|R(t_k)|/h) \tag{4.23}$$

   where $h = 6 \cdot \text{median}(|R(t_i)|)$ and B is the bisquare function:

   $$W(u) = \begin{cases} (1-u^2)^2 & \text{if } |u| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.24}$$

   Then the inner loop is repeated but in the LOESS smoothing steps (c) and (g) the values of the weigths are modified as in 4.21 giving less importance to outlier points.

The number of iteration of both the inner and outer loop are selectable parameters. In this work I left the parameters to their default values by Cleveland et al. (1990) to ensure convergence.

Parameter selection

In this section I describe the parameters of which the decomposition relies. They are summarized in table 4.1 along their selected values. For some of these parameters a data-independent optimal defaults is proposed by Cleveland et al. (1990) on the basis of a frequency response analysis of the operators associated with the smoothing operations,

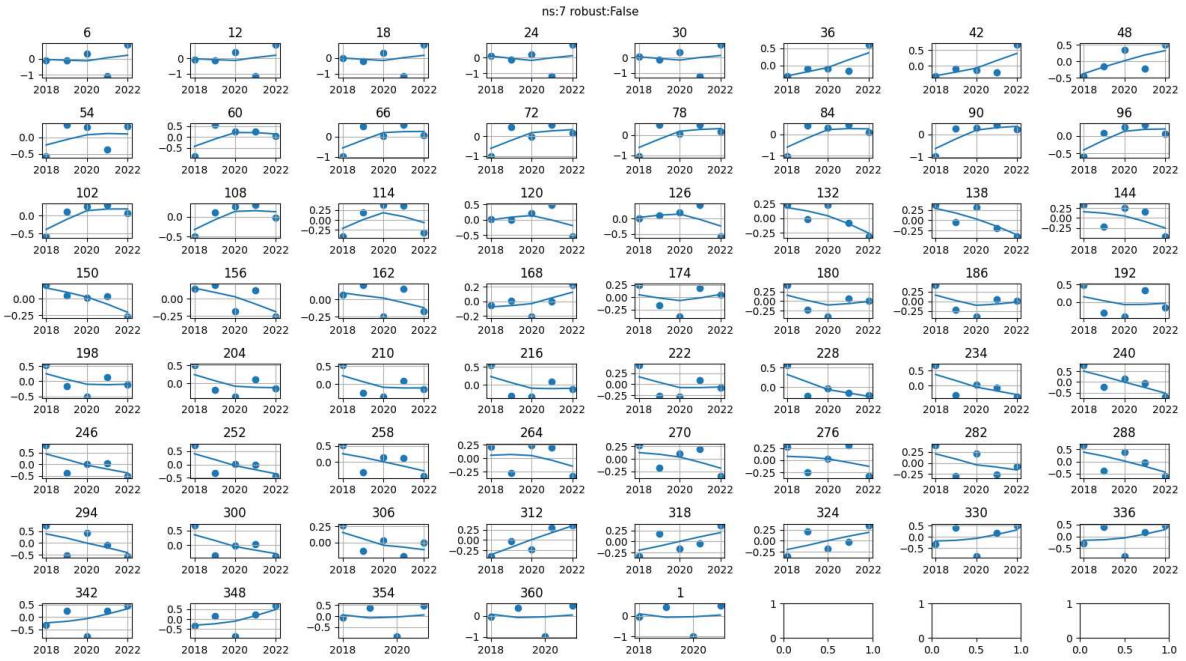


Figure 4.6: Example of a seasonal diagnostic plot used for the selection of the $\mathbf{n_s}$ parameter. Each panel is realtive to a day-of-year. Points represent the cycle subseries from the data and solid lines the corresponding points in the seasonal component as estimated from STL. Robustness iteration not performed

and adopted by `statsmodels`. In each of these cases I have followed this indication leaving the parameter value to its default.

- **Number of data in a seasonal cycle** n_p : It is used to determine the seasonal periodicity. Its value naturally derives from the temporal sampling of the timeseries. In our case, for a six day sampling rate and a periodicity of one year $n_p = 365/6 \approx 61$

- **Seasonal smoothing span** n_s : It is used in the smoothing of the cycle subseries, its value determines how much variation should go into the seasonal component by controlling the entity of the seasonal smoothing performed in step (c). It has no default value but has to be choosen on the basis of the data. A smaller n_s allows the seasonal component to capture more interannual variation while with larger n_s the seasonal component becames more periodic. To help an informed choice, a seasonal diagnostic plot, as in Figure 4.6, can be made, in which each panel contains a subseries (circles) and the corresponding seasonal cycle (solid line). The amount in which the solid line follows the points gives an indication of how much of the variability is captured by the seasonal smoothing.

In our case, with just 5 years of data available, we chose for n_s the value 7 to capture the interannual fluctuations without incurring in the risk of overfitting.

- **Low-pass span** n_l : Controls the cutoff for the low frequency which should be preserved by the LOESS smoothing of step. and thus the frequencies which should be removed from the seasonal component. Cleveland et al. (1990) shows that an optimal value is obtained setting it equal to the smallest odd integer greater or equal to n_p.

- **Trend smoothing span** n_t : Determines the amount of smoothing of the trend component performed the step (g) of the inner loop. It should be as small as possible to closely follow long term changes. At the same time, it should be large enough to filter out the high frequencies that are selected by the seasonal smoother at step (c) so that the trend and seasonal component do not compete for the same frequencies. In Cleveland et al. (1990) it is shown that a lower bound for n_t is the smallest odd integer n_t such that:

$$n_t \geq \frac{1.5 n_p}{1 - 1.5 n_s^{-1}} \tag{4.25}$$

- **Robustness iterations** In this work we have repeated the decomposition both with and without robustness iterations and compared results.

4.3 K-mean clustering

K-means clustering is an unsupervised learning algorithm that classify time series in a given number of groups (clusters) minimizing the intra-cluster variance. Given a set of time series of N epochs indicated with $\mathbf{X} = \{X_i^{(1)}, ..., X_i^{(M)}, \}$ with $i = 1, \ldots, N$ we define the K-means cost function, given a parameter K indicating the number of clusters:

$$J = \sum_{k=1}^{K} \sum_{X_i \in \mathbf{C}_k} \|X_i - \mu_i^{(k)}\|^2 = \sum_{k=1}^{K} var(\mathbf{C}_k) \tag{4.26}$$

where $\mathbf{C}_k$ the set of time series assigned to cluster k ($k = 1, ..., K$), $\mu_i^{(k)}$ the average time series of cluster k ($i = 1, ..., N$) and $|X_i - \mu_i^{(k)}\|^2$ is the square of the Euclidean distance between time series X_i and $\mu_i^{(k)}$.

As can be seen from equation 4.26 the cost function correspond to the sum of the intra-cluster variances $var(\mathbf{C}_k)$ of each cluster.

The algorithm used to minimize J follows an iterative procedure:

- The cluster centroid are initialized choosing K random time series among the group in the M time series.

- For each timeseries the euclidean distance from each of the centroids is computed and the series is assigned to the cluster whose centroid has the less distance.
- Each cluster centroid is updated as the mean of the time series assigned to its cluster.

The procedure is interrupted when either the cluster assignments do not change or the cost function, as defined by Equation 4.26, changes by less than a threshold between two iterations.

4.3.1 Silhouette score

Since the number of cluster is an input to the algorithm we need a method to assest the quality of a given choice for K. A commonly used metric for this purpose is the *silhouette score* [Kang et al. 2023]. Given a time series X_i, classified in the cluster k, the silhouette score $s(X_i, k)$ is defined as:

$$s(X_i, k) = \frac{b(X_i, k) - a(X_i, k)}{\max\left(b(X_i, k), a(X_i, k)\right)} \tag{4.27}$$

where:

- $a(X_i, k)$ is the average distance of X_i to all other time series in the same cluster k (i.e., the intra-cluster distance). With N_k being the number of series classified in cluster k:
$$a(X_i, k) = \frac{1}{N_k} \sum_{Y_i \in \mathbf{C}_k} \|X_i - Y_i\|^2 \tag{4.28}$$
- $b(X_i, k)$ is the minimum average distance of X_i to all points in a different cluster, computed over all other clusters (i.e., the nearest-cluster distance):
$$b(X_i, k) = \min_{k' \neq k} \frac{1}{N_{k'}} \sum_{Y_i \in \mathbf{C}_{k'}} \|X_i - Y_i\|^2 \tag{4.29}$$

For each series X_i, $s(X_i, k)$ ranges from -1 to 1. Values close to 1 indicate that the time series is well matched to its own cluster and poorly matched to neighboring clusters. Values close to 0 indicate that the time series lies between clusters, and negative values suggest incorrect cluster assignment. Therefore the average silhouette score over all time series in the dataset can be used as a measure of the goodness of the clustering.

We compute this metric for several values of K and we apply the elbow method (or knee method) to the silhouette score curve. This technique identifies the value of K beyond which the improvement in clustering quality starts to diminish, marked by a change in slope. This value has been selected as the cluster number.

4.4 Cross-correlation

To investigate the possible drivers of the signals extracted by the former analysis and identify areas where the observed deformation is possibly related to hydrological-related processes, we employed a lagged correlation analysis using precipitation data (Section 4.2.1) as a proxy of the hydrological trend. This consists in computing the Pearson correlation coefficient between the time series of the deformation signal and the precipitation time series, while varying the time lag applied to the latter. For each measurement point, the lag that maximizes the absolute correlation was selected, providing both an estimate of the strength of the correlation and an indication of the possible temporal relation between the two variables.

The cross-correlation ρ at lag τ, between two time series of N data X_i and Y_i, $i = 1, ..., N$ is defined as:

$$\rho(\tau) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_{i+\tau} - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_{i+\tau} - \bar{Y})^2}} \tag{4.30}$$

with $\bar{X}$ and $\bar{Y}$ being the average values of the two time series. In this equation τ represent the lag measured in number of time steps of the time series. To obtain the value of the time lag τ is multiplied by the sampling period of the time series, in our case 6 d. For $\tau = 0$ the Pearson correlation coefficient is recovered.

$$\rho(\tau = 0) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{4.31}$$

Chapter 5

Data analysis results

5.1 Linear deformation trend

The first post-processing operation performed on the EGMS data consisted in fitting each time series wit a first order polynomial using least-squares regression. The coefficient of the first-order term represents the mean velocity at each location, the obtained values are shown in Figure 5.1. The mean velocities are mostly negative in the study area, indicating subsidence. The subsidence rates are higher in the center of the plain, with average values of $5\,\mathrm{mmy}^{-1}$ and localized maxima reaching $10\,\mathrm{mmy}^{-1}$. These rates decrease towards the Alps, reaching values slightly lower than $3\,\mathrm{mmy}^{-1}$ in the foothills regions.

Two localized areas exhibit an average uplift over the considered time period. One is located at the center of the domain, corresponding to the city of Brescia, partially extending northward into the Trompia Valley and broadening southward in latitude, forming an inverted "T" shape. The second area is smaller and located in the easternmost part of the domain, around latitude 45°30′N and it corresponds to the Municipality of Negrar di Valpolicella in the Verona Province.

The linear trend component is typically associated with long-term deformation processes that are approximately uniform over the period of observation, such as tectonic movements or sediment compaction [Farolfi et al. 2019]. Since the focus of this work is on shorter-term deformation effects related to fluctuations in meteoclimatic conditions, the estimated linear component was removed from each time series.

The subsequent analyses were conducted on the resulting detrended dataset. These are STL decomposition, which I applied to each detrended time series to separate the annual seasonal variability from the non linear multiyear variations, K-Means clustering which I applied to both the detrended series an their STL components and the linear decomposition techniques PCA and ICA which I applied to the detrended dataset.

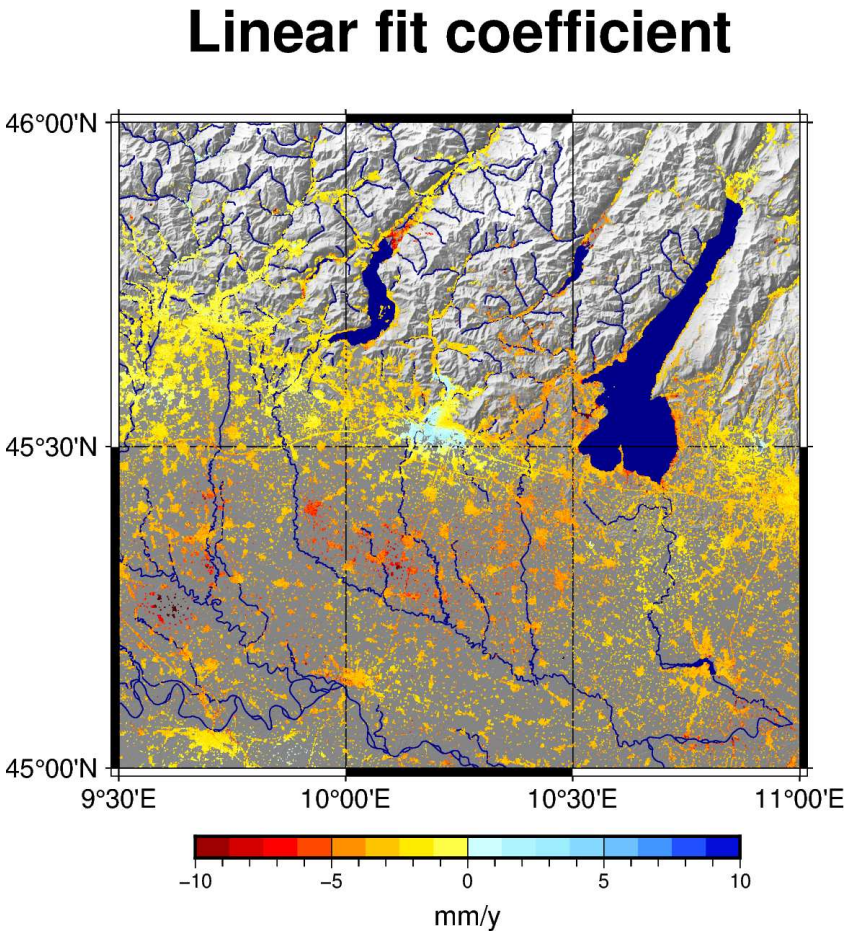


Figure 5.1: Linear fit coefficient values for vertical velocity dataset

5.2 STL

This section presents the results of the STL decomposition performed on each detrended time series. The decomposition separates each series into three components: seasonal, multi-year trend, and residual. The spatial variability of each component's amplitude was examined by mapping the difference between maximum and minimum values of the corresponding time series at each location (Figure 5.2). These amplitude values were compared to the amplitude of the original time series to assess the relative contribution of each component.

From the map of maximum deformation, it can be observed that in several locations the maximum deformation during the study period reached values around 10 mm. In particular, the previously highlighted "T" shape corresponding to the city of Brescia shows up, where deformation reaches up to 15 mm. Other locations with significant deformation include a valley in the northwest (corresponding to the municipality of

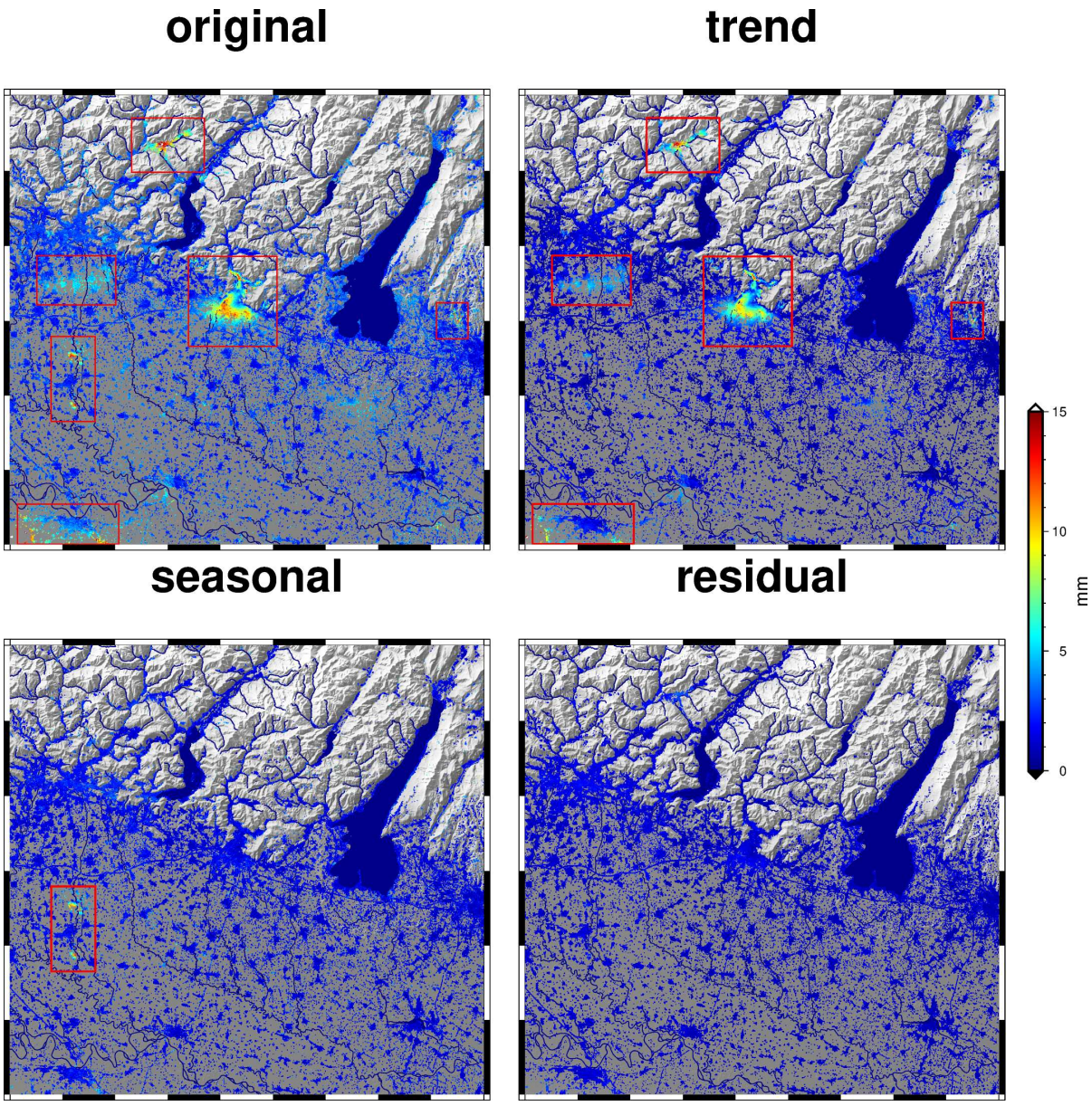


Figure 5.2: Maps of deformation amplitude for the original time series (a), the STL-extracted trend (b), seasonal (c), and residual (d) components. Red boxes highlight areas showing significant deformation signals.

Clusone), an area on the southwestern edge of the domain corresponding to the city of Piacenza, and some smaller localized areas in the western part of the plain. Among these, the areas of Brescia, Piacenza and Clusone are also clearly visible in the amplitude map of the multi-year trend component, indicating significant variability at the multi-year timescale. An additional area southwest of the city of Bergamo is also identified, with lower deformation values but still exceeding 5 mm.

The seasonal component amplitude highlights a seasonal signal of up to 15 mm in two small areas located north and south of the city of Crema, anywhere else the amplitude values do not exceed 5 mm.

The residual component map also shows deformation values below 5 mm throughout the entire study area, including the locations highlighted in the other components. Residual deformation is associated with high-frequency variability filtered out during the STL smoothing process, e.g., short-term (weekly-monthly) signals removed during trend smoothing.

5.3 K-means clustering

In this section I describe the results of the application of the K-means clustering algorithm. If the procedure identifies areas characterized by similar deformation trends, then the points of each cluster aggregate in localized areas. The average deformation time series of these areas is represented by the cluster centroid.

The centroid time series are compared to the cumulative detrended precipitation time series (see Section 3.2.1). The clustering procedure was applied to the detrended dataset as well as to the datasets containing the STL-derived seasonal and multi-year trend components. In the latter, the objective is to identify areas characterized by similar seasonal or multi-year deformation patterns, analyzing each timescale separately. Accordingly, the precipitation time series used for comparison was also decomposed into two components—one representing seasonal variability and the other representing multi-year variability—using the STL method.

Also, for each clustering result, the values of the silhouette coefficient were computed for each time series. These coefficients quantify how much more similar is a series to the members of its own cluster than to the members of other clusters. Higher values indicate well-classified time series, while negative values suggest possible misclassification. The average silhouette coefficient provides a general indication of the quality of the clustering.

In this work, K-means clustering was repeated for cluster numbers ranging from 2 to 8, and the average silhouette coefficient was plotted against the number of clusters. From these plots, the optimal number of clusters was selected using the elbow method (Section 4.3.1)

5.3.1 Original dataset

For the original dataset the silouhette coefficient methods selected a number of three clusters (Figure 5.3b). Figure 5.3a and 5.3c shows the clustering results while Figure 5.3b shows the values of the silouhette coefficient for the different clusters.

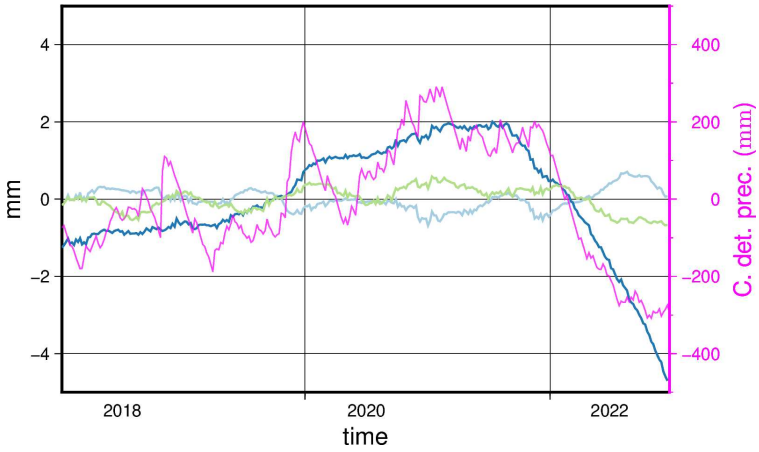
The majority of the time series were classified into clusters 1 and 3, with a smaller fraction in cluster 2. The centroid of cluster 2 is characterized by a deformation pattern positively correlated with the precipitation trend. The average deformation in this cluster shows an uplift starting in 2018 and peaking in 2021. During this period, the average displacement reaches approximately 3 mm. After 2021, a subsidence trend is observed, with a total displacement exceeding 6 mm by 2022. Time series belonging to cluster 2 are aggregated in five spatially distinct areas. These correspond to areas of Brescia, Clusone, Bergamo and Negrar previously highlighted wthrought STL decomposition.

Outside these zones, points are divided between clusters 1 and 3. In most of the study area, the spatial distribution of these clusters is scattered, although points classified as cluster 3 generally surround those of cluster 2. Clusters 1 and 3 show small average deformations - below 2 mm - and no significant multi-year trend. Both corresponding centroids exhibit seasonal oscillations of small amplitude - approximately 1 mm - which are in antiphase with seasonal oscillation of precipitation for cluster 1 and in phase for cluster 3.

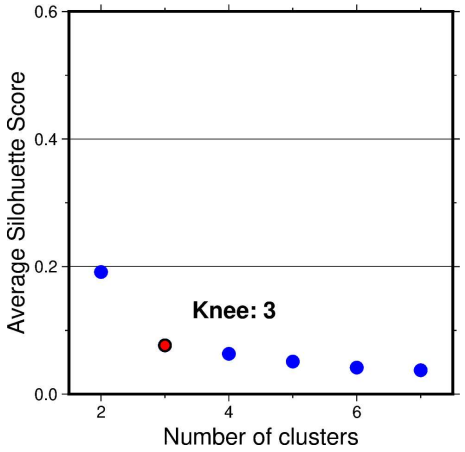
5.3.2 Multi year trend STL component

For the STL multi-year component, the silhouette method selected four clusters (Figure 5.4b). Results are shown in Figure 5.4a and 5.4c. The spatial distribution highlights the same areas previously identified, although they are now separated into two groups. Cluster 4 includes the time series from the Brescia area and the Clusone valley. Its average deformation is strongly correlated with the detrended cumulative precipitation. Cluster 3 has a similar temporal pattern, but a lower amplitude. It includes points in from the southern area of Bergamo, Clusone, Negrar and in the Brescia.

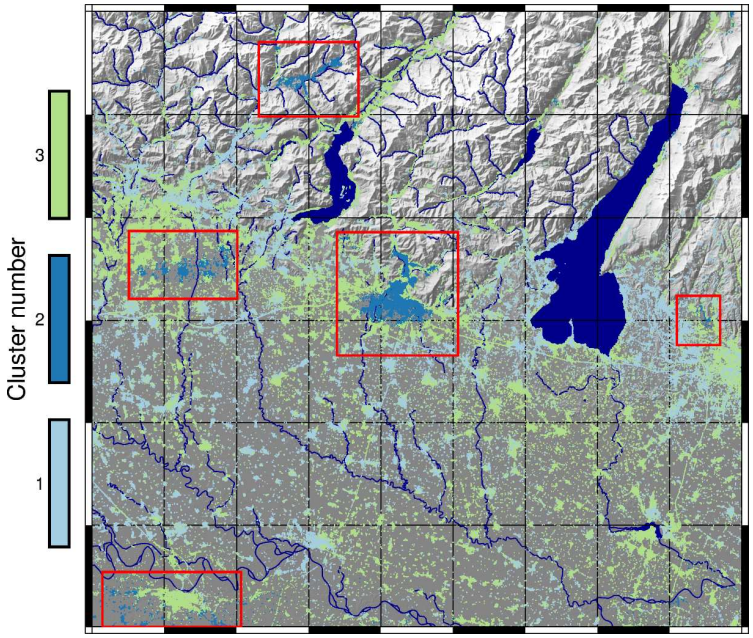
The remaining series are divided between cluster 2 - characterized by a stable temporal trend - and cluster 1, which exhibits low-amplitude trends that are negatively correlated with precipitation (Figure 5.4a (up)). Comparison of the silhouette distribution plot for this case (Figure 5.4d), with the one relative to the clustering of the original dataset (Figure 5.3d) shows an improvement of the classification, when using only the multi-year deformation component, reflected also in an higher value of the average silouhette coefficient.



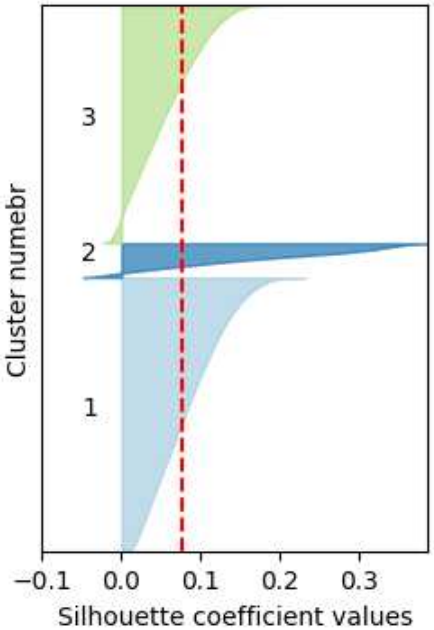
(a) Time dependence of cluster centroids



(b) Average silhouette coefficient for each clustering

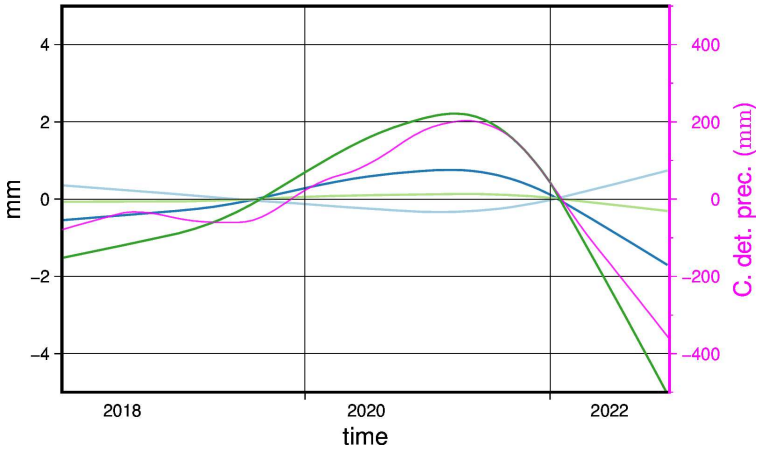


(c) Spatial distribution of clusters

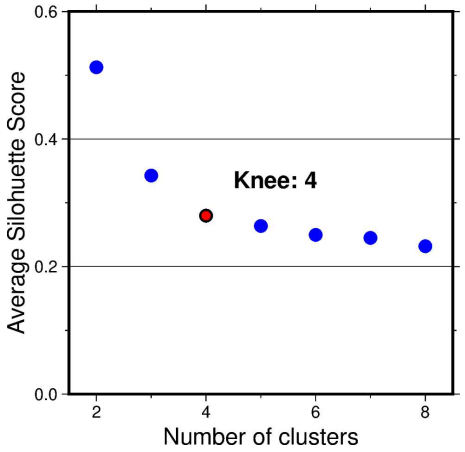


(d) Distribution of values of the silhouette coefficient

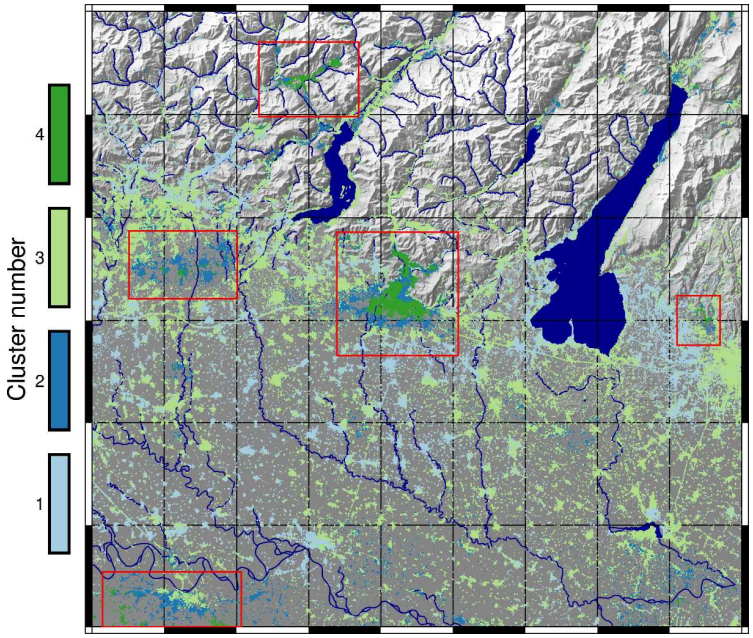
Figure 5.3: K-Means clustering of the original detrended dataset: (a) Time dependence of cluster centroids, (b) spatial distribution of clusters, (c) average silhouette coefficient for each clustering, (d) distribution of values of the silhouette coefficient.



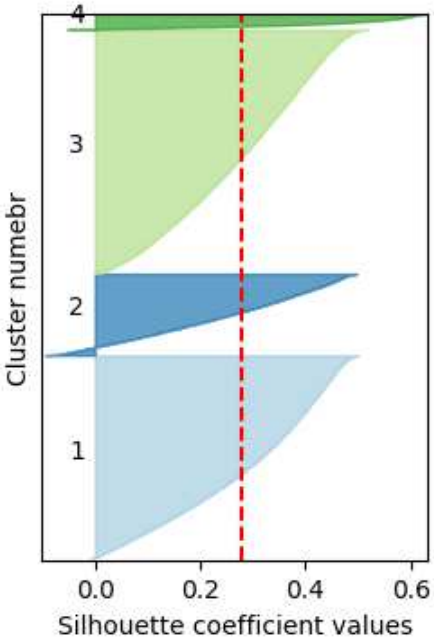
(a) Time dependence of cluster centroids



(b) Average silhouette coefficient for each clustering

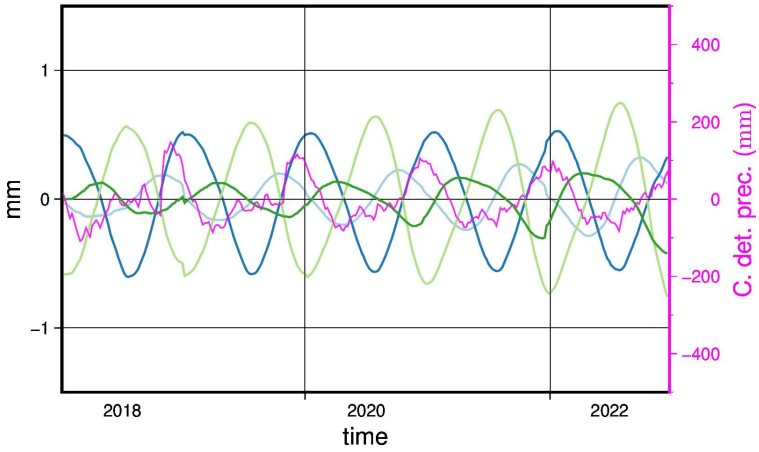


(c) Spatial distribution of clusters

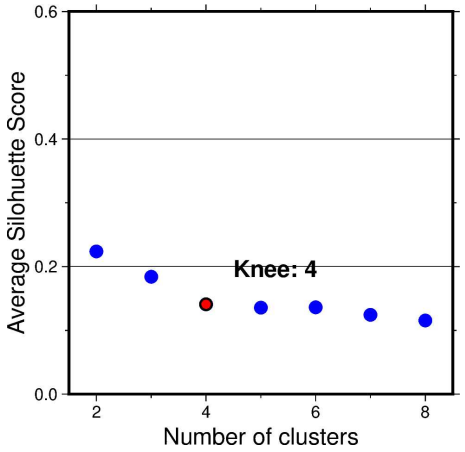


(d) Distribution of values of the silhouette coefficient

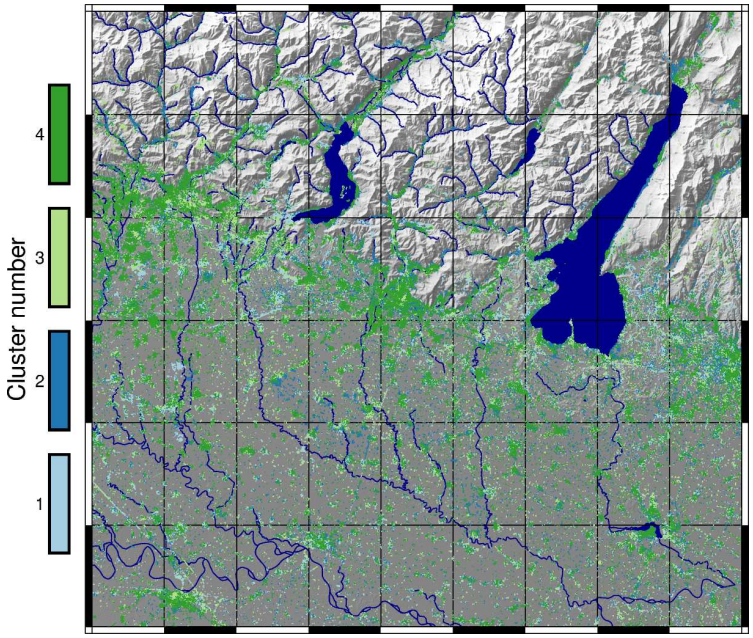
Figure 5.4: K-Means clustering of the STL multi-year component: (a) Time dependence of cluster centroids, (b) spatial distribution of clusters, (c) average silhouette coefficient for each clustering, (d) distribution of values of the silhouette coefficient.



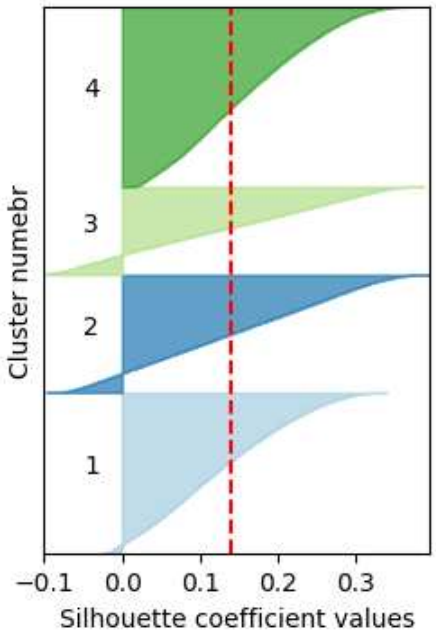
(a) Time dependence of cluster centroids



(b) Average silhouette coefficient for each clustering



(c) Spatial distribution of clusters



(d) Distribution of values of the silhouette coefficient

Figure 5.5: K-Means clustering of the STL seasonal component: (a) Time dependence of cluster centroids, (b) spatial distribution of clusters, (c) average silhouette coefficient for each clustering, (d) distribution of values of the silhouette coefficient.

5.3.3 Seasonal STL component

For the clustering of the STL seasonal component, the silhouette method selected four clusters. The results are shown in Figure 5.5b, and in this case the seasonal component of the detrended precipitation was compared to the cluster centroids.

From Figure 5.5b, it is evident that clustering on the seasonal component does not provide as meaningful a classification as in the multi-year case. Cluster distributions are spatially scattered across the domain, and the silhouette plot (Figure 5.5c) reflects low clustering quality, with many time series likely misclassified, as indicated by negative values of the siouhette coefficient.

5.4 Summary of clustering results

The clustering analysis reveals that the multi-year deformation component plays a significant role in several areas across the study domain. In contrast, the seasonal component appears to be of limited relevance and is strongly expressed only in two small, localized regions.

When clustering is applied to the original detrended dataset, three main groups of time series are identified: one characterized by a pronounced multi-year trend positively correlated with precipitation, and two exhibiting small-amplitude seasonal oscillations. These seasonal trends show different behaviors in relation to precipitation—one group oscillates in phase, while the other in antiphase with the seasonal precipitation pattern.

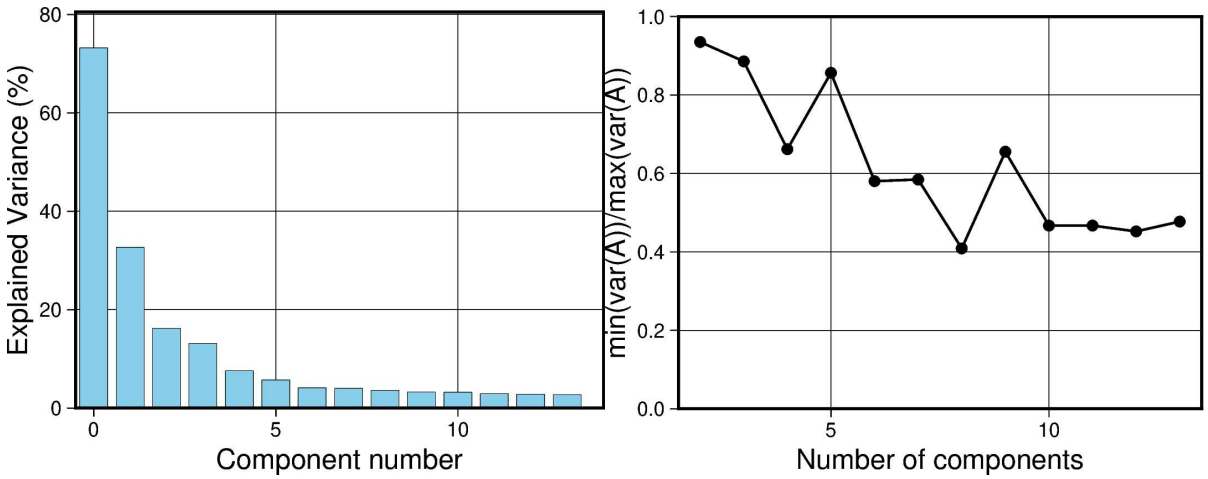
A more detailed classification emerges when clustering is performed exclusively on the STL-derived multi-year trend component. In this case, four distinct types of deformation behaviors are observed: (1) multi-year trends positively correlated with precipitation and high amplitude, (2) multi-year trends positively correlated with precipitation but with lower amplitude, (3) trends negatively correlated with precipitation, and (4) time series that remain nearly stable throughout the observation period.

5.5 Data driven decomposition

The detrended dataset was analyzed using the linear decomposition techniques PCA and ICA, in order to isolate spatiotemporal signals using the whole dataset.

In both cases, it was necessary to select the number of components to retain. To this end, both decompositions were computed with a number of components ranging from 2 to 14. Two diagnostic plots were produced: one showing the variance explained by each PCA component (Figure 5.6a), and the other showing the ratio between the maximum and minimum ARD values obtained from ICA (Figure 5.6b).

The first plot allows for the evaluation of the relative importance of each PCA component by examining how the explained variance is distributed. Since PCA components



(a) (b)

Figure 5.6: (a) Variance explained by each PCA component, showing a decreasing trend and plateauing after the fifth component. (b) Ratio between the maximum and minimum Automatic Relevance Determination (ARD) values for each ICA component,.

are ordered by decreasing variance, the values decline as the number of components increases. In this case, the curve reaches a plateau after the fifth component.

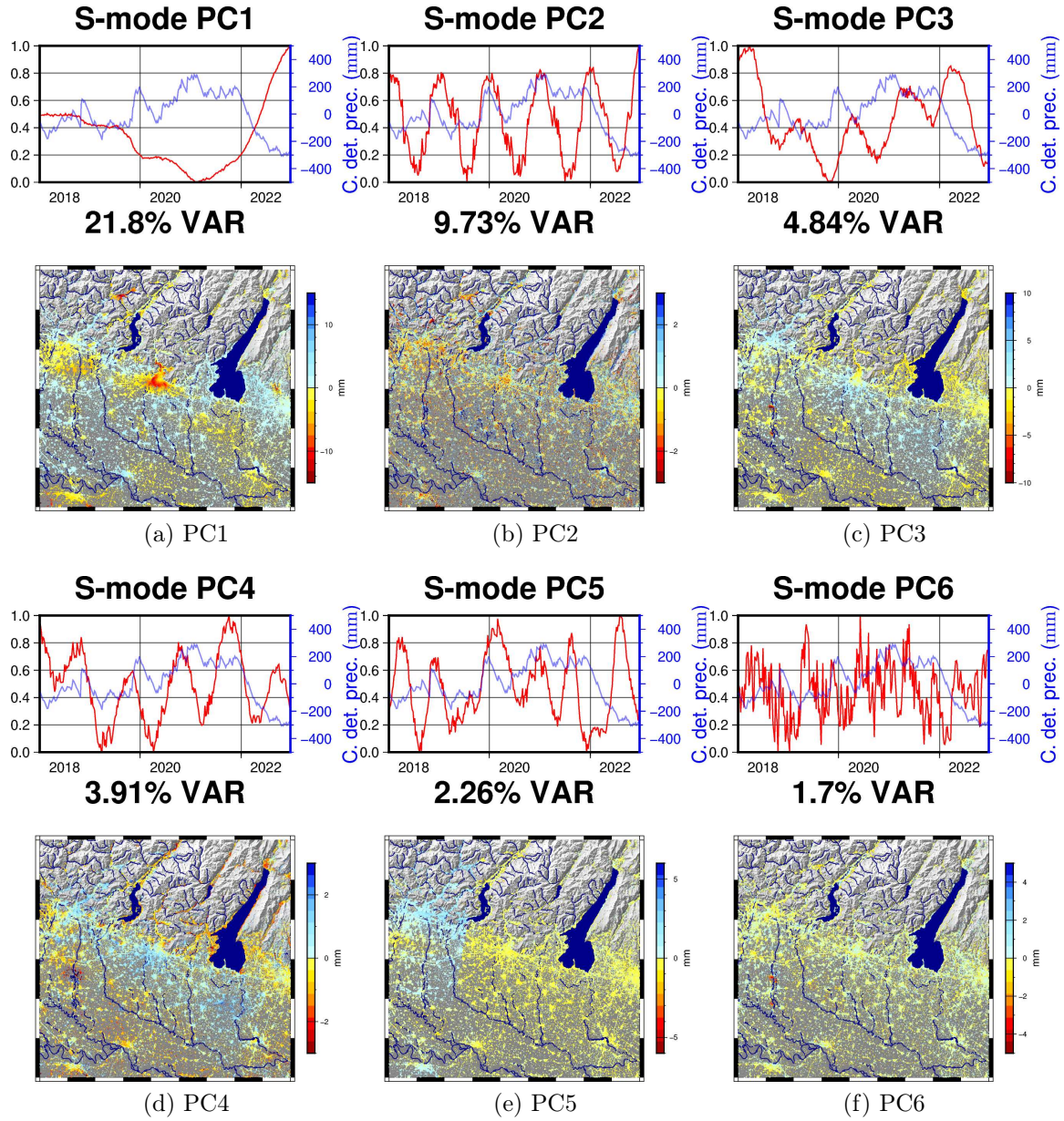
Even if showing some fluctuations the maximum-over-minimum ARD ratio associated with vbICA decreases with the components number and approximately stabilizes starting from the sixth component. For this reason, the decomposition with six components was selected as a reference and compared with those obtained using fewer or more components.

The results of the PCA and ICA are presented graphically in Figures 5.7 and 5.8. For each component, the temporal function is plotted above the corresponding spatial function. For visualization purposes, the temporal signals were normalized to the [0, 1] interval using their amplitude range, and the same scaling factor was applied to the spatial components. The displacement at a given pixel can be retrieved by multiplying the spatial value at the pixel and the corresponding temporal function.

The results show that including the sixth component consistently led to the isolation of a high-frequency temporal signal with a significantly lower deformation amplitude than the others (Figure 5.7f, Figure 5.8f). Adding more components produced additional high-frequency patterns, interpreted as noise. This behavior was observed for both PCA and ICA, supporting the identification of the sixth component as noise. Consequently, only the first five components were considered physically meaningful.

The results of the two decompositions show noticeable similarities. In both cases, the most relevant signal is represented by the first component, which exhibits a similar spatiotemporal pattern in both methods. This component reflects a multi-year trend with

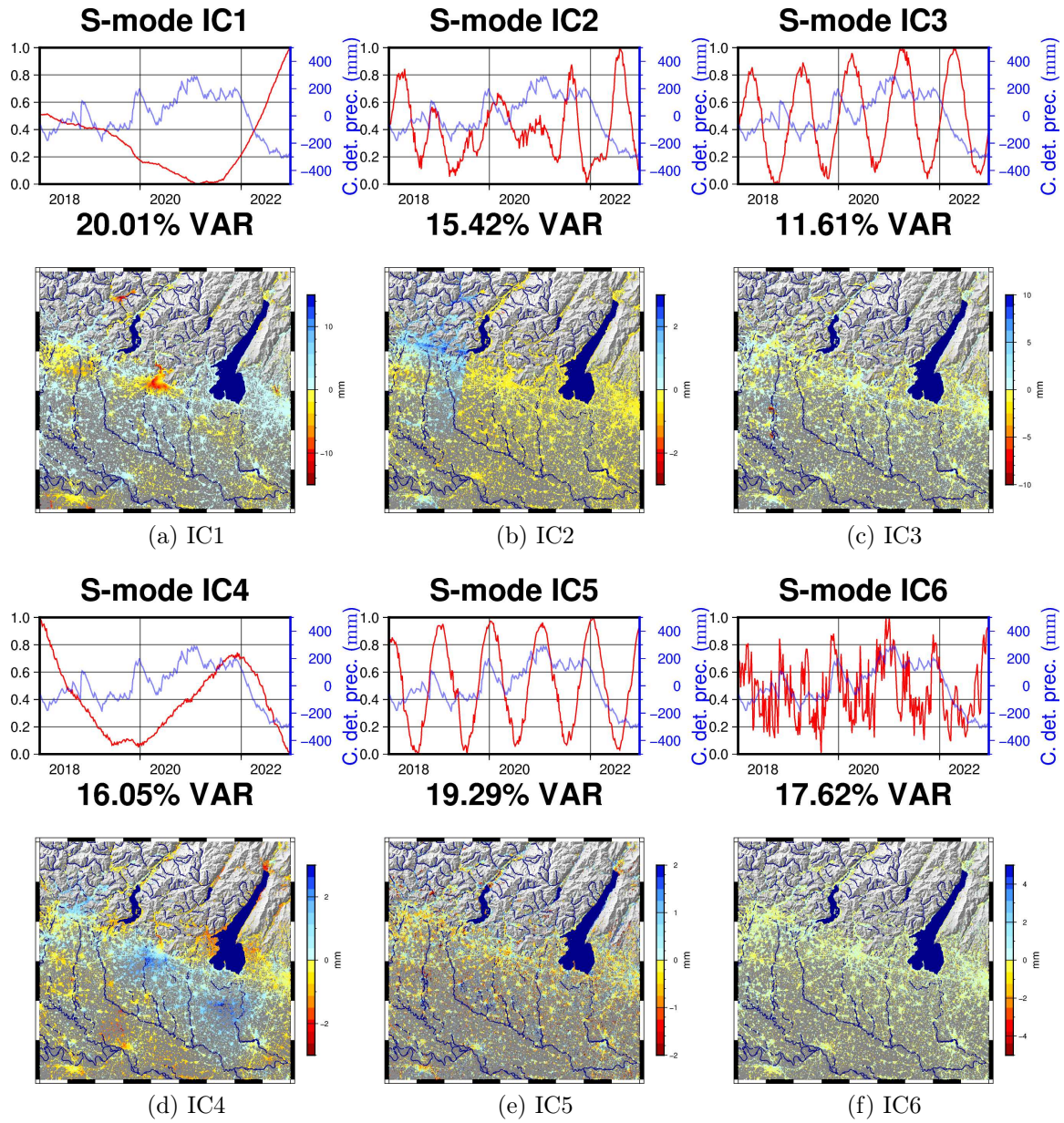
no superimposed oscillations, characterized by a decreasing behavior until 2021, followed by an increase that continues until the end of the observation period. The associated spatial function again highlights the areas of Brescia, Clusone, and Negrar, whit related negative values of the spatial function exdeeding 10 mm in absolute magnitude. In other



(a) PC1 (b) PC2 (c) PC3

(d) PC4 (e) PC5 (f) PC6

Figure 5.7: PCA decomposition results: temporal (top) and spatial (bottom) components.



(a) IC1 (b) IC2 (c) IC3

(d) IC4 (e) IC5 (f) IC6

Figure 5.8: ICA decomposition results: temporal (top) and spatial (bottom) components.

regions where a multi-year signal had previously been identified, i.e. the area south of Bergamo and near the south-west border of the domain, the spatial function remains negative but with lower absolute values. Outside these areas, the values of the spatial function are positive and below 3 mm in absolute terms. The deformation associated with this component therefore matches what was previously observed: a multiyear uplift of

localized region between 2018 and 2021, followed by subsidence thereafter, both occurred with spatially variable amplitude, and an opposite multi-year deformation trend outside these regions. Comparison with the detrended cumulative precipitation series, displayed alongside the temporal functions, reveals a clear correlation between this deformation signal and the multi-year evolution of precipitation.

The two techniques also produce similar results for components IC2 and PC5, which exhibit comparable behavior in both their spatial and temporal functions (Figures 5.7e and 5.8b). Both are characterized by small positive values in the eastern part of the domain and negative values in the west. These two regions are separated by a roughly north–south-oriented boundary located near longitude . The associated temporal functions display oscillations of variable amplitude and period, with no evident correlation to precipitation. The lack of correlation with known physical forcings, combined with the sharp spatial transition that does not correspond to any notable geographic feature, suggests that this component is unlikely to represent a physically meaningful process and is possibly related with processing artefacts.

Differences between the two decomposition techniques are more relevant in the remaining components. ICA identifies two periodic components (IC3 and IC5) and one additional multi-year component (IC4), whereas PCA extracts only one periodic (PC2) component and two components with more complex temporal patterns (PC3 and PC4) (see Figures 5.8c, 5.8e, 5.8d, and Figures 5.7b, 5.7c, 5.7d).

In the case of ICA, both periodic components exhibit annual seasonality and are temporally shifted, with IC5 lagging behind IC3. In their spatial pattern, the most notable feature of IC3 is the presence of two deformation hotspots in the western part of the domain, located north and south of the city of Crema, where the seasonal oscillations exceed 10 mm in amplitude (Figure 5.8c). Outside this zone, the spatial function of IC3 remains close to zero. The strong spatial localization of the signal is attributed to the presence of a natural gas storage facility in this area, whose operations are responsible for the observed deformation [Rigamonti 2025]. In such facilities, natural gas is injected under pressure into a confined aquifer, replacing the pore water with gas. The gas is typically stored during the summer and extracted in winter for consumption. The resulting variations in pore pressure can induce seasonal ground deformation.

A deformation hotspot is also visible in component IC5 (Figure 5.8e), although the associated amplitude values are lower and less concentrated around the Crema area. IC5 can be associated with the main seasonal component identified by PCA (PC2; Figure 5.7b); both components likely capture, to varying degrees, the periodic localized signal related to gas storage activities along with broader natural seasonal variations present throughout the region.

The final component identified by ICA, IC4 (Figure 5.8d), displays a multi-year character, with a temporal minimum and maximum occurring in the winters of 2019 and 2022, respectively. Its spatial pattern is defined by positive amplitudes in the central part of the plain and negative values toward the surrounding mountainous areas, in both cases

not exceeding 2 mm amplitueds. The low correlation between its temporal evolution and precipitation makes it difficult to directly associate this component with a hydrological signal, as was done for IC1. This component might be driven by a low amplitude and temporarily shifted response in some areas, but further investigation would be required to confirm this hypothesis.

The PCA results for components PC3 and PC4 differ significantly from the corresponding components obtained through ICA. As shown in the figures (Figures 5.7c and 5.7d), PCA does not successfully isolate the deformation observed in the Crema area and the multiyear signal in IC4 into distinct components. Instead, both PC3 and PC4 exhibit seasonal oscillations superimposed on a multi-year trend, resembling that of IC4. The overlap is also evident in their spatial patterns, showing both a broad signal similar to the spatial function of IC4 and the two deformation hotspots in the area of Crema.

These results reflect the difference in the effectiveness of the two techniques in addressing the blind source separation problem. ICA successfully identifies and disentangles deformation signals arising from distinct physical processes—namely, hydrological forcing and anthropogenic deformation associated with gas storage—whereas PCA only partially separates these contributions, yielding components that reflect a mixture of different sources.

5.6 Cross-Correlation

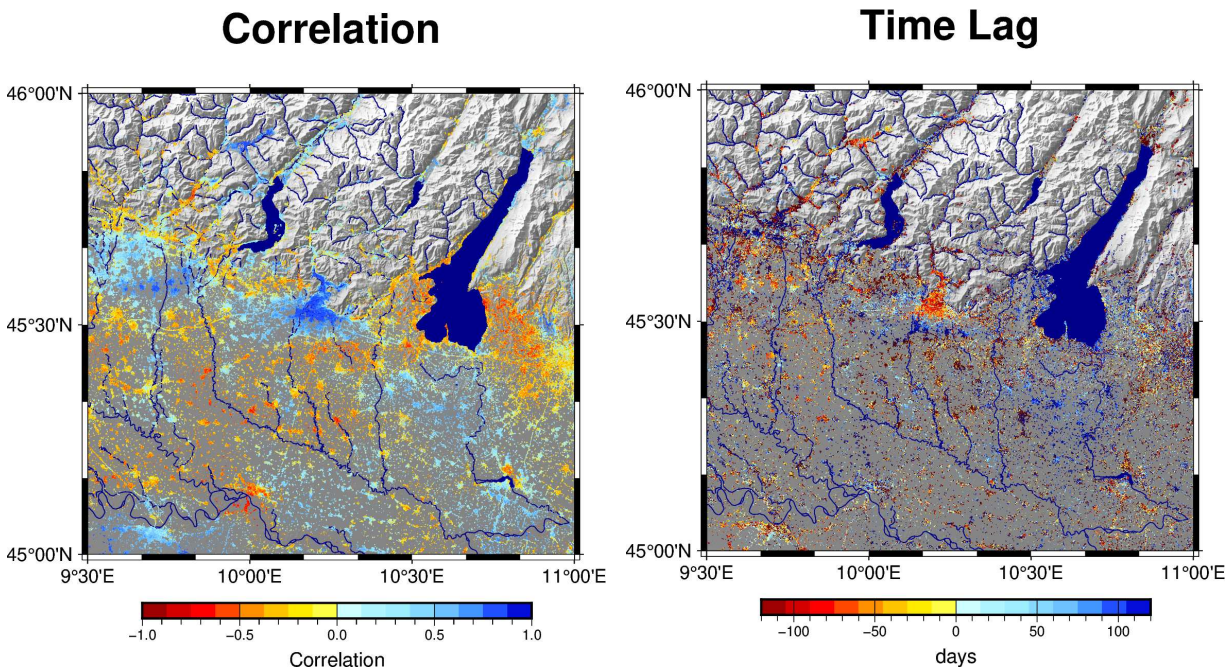
The results of the cross-correlation between the detrended deformation dataset and the cumulative detrended precipitation time series are shown in Figure 5.9. The left panel displays the spatial distribution of the maximum correlation coefficient, while the right panel maps the temporal lag (in days) at which the maximum correlation occurs.

The map of the correlation values confirms previous observations: the areas of Brescia, Clusone, southern Bergamo, Negrar, and the southwestern sector of the study area exhibit positive correlation. The magnitude of the correlation varies across the domain, with the highest values observed in the Brescia and Clusone regions, where the correlation coefficient exceeds $\rho > 0.9$ and shows minor time lag vaues ($<$ 2-3 months)

In contrast, the correlation becomes negative across the central portion of the plain but with lower absolute values, indicating a weaker inverse relationship between precipitation and deformation in these areas.

5.7 Summary

The data-driven decompositions confirmed the previously observed relevance of multiyear deformation signals, which were detected with varying amplitude in almost all areas and showed a temporal pattern strongly resembling that of precipitation. These signals



(a) (b)

Figure 5.9: Maximum correlation coefficient between detrended deformation and cumulative detrended precipitation time series (a) Time lag (in days) at which the maximum correlation occurs (b)

exhibit significant amplitude in localized regions, typically situated in valleys located in the mountains or at the boundary between the plain and the mountainous area. In these zones, deformation is positively correlated with hydrological trends ($\rho > 0.9$), uplifting in rainy periods and subsiding in drought periods, and reaches amplitudes up to 15 mm. A similar, though less pronounced, signal is also present in areas adjacent to these, whereas in the plain, deformation shows smaller amplitude and a negative correlation with precipitation (e.g., original cluster, multi-year cluster, PC1, IC1). The contrasting correlation patterns observed suggest that the deformation field results from the superposition of multiple processes driven by long-term hydrological variability, such as elastic loading response and poroelastic deformation, each occurring with variable magnitude across the study area. In the next chapter I explore the possible physical mecahnism which could be associated to these deformation signals.

Chapter 6

Models

In this work I studied transient deformation processes occurring on yearly to multi-year time scales (up to 10 years). Both my results (5.7) and previous studies in the region [Pintori et al. 2024], support the hypothesis that deformation in the region at these time scale has a relevant contribution associated with hydrological processes. These processes involve the redistribution of water masses, which can exert stress on the solid Earth through different mechanisms.

One of the principal mechanism is due to the flexure of the lithosphere under the water masses load. This effect is sketched in Figure 6.1 for an increase in the water storage. The added weight produces surface deformation characterized by subsidence under the load application point and horizontal movement towards it. Another mechanism occurrs in presence of materials in which the solid matrix has void space (pores) that can be saturated with water. Terrain layers with these characteristics are referred to as aqufiers. In this case the deformation is due to the variations of the fluid pore-pressure whithin these layers which produces stresses on the solid matrix. Processes of this type are termed poroelastic processes, the effect is sketeched in Figure 6.1 for an increase in water storage. The increased storage is reflected in an increase of the pore pressure trought the whole saturated layers, which produce a local expansion of the medium [Larochelle, Chanard, et al. 2022]. The resulting deformation is in this case of uplift and outward movement.

These two processes thus, altought both forced by variations and redistribution of water masses, present some key difference, due to different physical mechanism interveneing in each case. In the first process the responsible force is the weight of all water masses above the lithosphere in a given point. The loading force is associated in fact with the total Terrestrial Water Storage (TWS), which includes both water stored in surface reservoir, i.e the Surface Water Storage (SWS) and in the underground, i.e the Ground Water Storage (GWS). In the second process, instead, the forcing is associated with the GWS term. Another key difference is the direction of the response which is opposite between the two processes, i.e. elastic upflit and poroelastic contraction for a decrease in water storage and elastic subsidence and poroelastic uplift for an increase. It

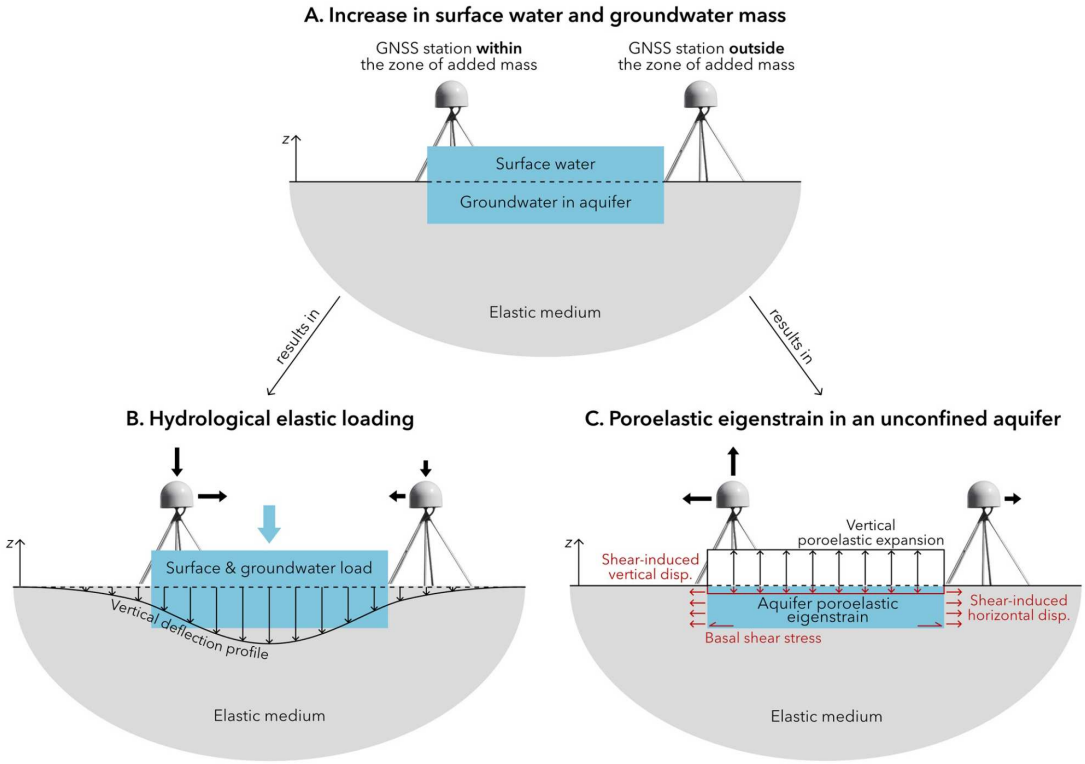


Figure 6.1: Deformation due to change in water storage (A). The added weight of water mass result in subsidence and movement towards the load (B). The increment in pore pressure produce by an increase in groundwater storage produce uplift and outward horizontal movement (C). From Larochelle, Chanard, et al. 2022

is therefore necessary to individually model both processes in order to disentangle their effect on the observed deformation.

6.0.1 Elastic deformation

Governing equations

The elastic response to surface loading was computed employing the LoadDef software [Martens et al. 2019b], which computes the elastic response of a stratified, spherically symmetric, non-rotating, self-gravitating Earth. The exposition of the model follows the formulation presented in Martens et al. 2019a with partially modified notation

The governing equations for the deformation of a spherical Earth of radius a due to a surface mass distribution γ which exhert a gravitative load on the Earth are the

momentum-conservation equation and Poisson equation.

$$\begin{aligned} \rho \frac{\partial^2 \vec{u}}{\partial t^2} &= \nabla \cdot \hat{\sigma} + \vec{F} \\ \nabla^2 \psi &= 4\pi G \rho \end{aligned} \tag{6.1}$$

Where $\vec{u}$ the displacement vector, t is the time, $\hat{\sigma}$ is the stress tensor, ρ the mass density, ψ is the gravitational potential and G the Universal gravitational constant. The term $\vec{F}$ represents the volume forces which in this case is the gravity:

$$\vec{F} = \rho \vec{g} \qquad \text{where} \qquad \vec{g} = \vec{\nabla} \psi \tag{6.2}$$

Equations are valid inside the Earth. Outside the Earth the potential satisfies Laplace's equation:

$$\nabla^2 \psi = 0 \tag{6.3}$$

The equations (6.1) are linearized by expressing each variable as the sum of an unperturbed term, representing a state of spherically symmetric hydrostatic equilibrium, and a small perturbation due to surface loading and Earth deformation. Using subscripts 0 and 1 for unperturbed and perturbation quantities respectively, we write:

$$\begin{aligned} &\hat{\sigma} = \hat{\sigma_0} + \hat{\sigma_1} \\ &\rho = \rho_0 + \rho_1 \\ &\psi = \psi_0 + \psi_1 \\ &\vec{g} = \vec{g_0} + \vec{g_1} \qquad \text{where} \qquad \vec{g_0} = \vec{\nabla} \psi_0 \quad \vec{g_1} = \vec{\nabla} \psi_1 \end{aligned} \tag{6.4}$$

The stress tensor associated with a state of hydrostatic equilibrium can be written as [Martens et al. 2019a]:

$$\hat{T_0} = -p_0 \hat{I} \tag{6.5}$$

with p_0 satisfying:

$$\vec{\nabla} p_0 = \rho_0 \vec{g_0} \tag{6.6}$$

The umperturbed density distribution ρ_0 is associated to the umperturbed potential ψ_0 with:

$$\nabla^2 \psi_0 = 4\pi G \rho_0 \tag{6.7}$$

In the umperturbed state, the displacement is assumed to be zero. When the surface load is applied a deformation $\vec{u}$ is produced. The stress in the deformed body is composed by the pre-stress, the advection of the pre-existing hydrostatic stress field by the displacement field and the perturbation stress associated with the deformation:

$$\hat{\sigma} = \hat{T_0} - (\vec{u} \cdot \vec{\nabla}) \hat{T_0} + \hat{\sigma_1} = -p_0 \hat{I} + (\vec{u} \cdot \vec{\nabla})(p_0 \hat{I}) + \hat{\sigma_1} \tag{6.8}$$

The perturbation density associated with the deformed earth density has to satisfy the continuity equation for mass conservation:

$$\rho_1 + \vec{\nabla} \cdot (\rho_0 \vec{u}) = 0 \tag{6.9}$$

Substituting the expressions in (6.4) into the base equations (6.1), and applying the definitions in (6.8), (6.9), and (6.7), yields the following equations system:

$$\begin{aligned} \rho_0 \frac{\partial^2 \vec{u}}{\partial t^2} &= \nabla \cdot \hat{\sigma}_1 + \rho_0 \nabla \psi_1 + \nabla(\rho_0 \, \vec{u} \cdot \vec{g}_0) - \vec{g}_0 \, \nabla \cdot (\rho_0 \, \vec{u}) \\ \nabla^2 \psi_1 &= -4\pi G \, \nabla \cdot (\rho_0 \, \vec{u}) \end{aligned} \tag{6.10}$$

where all terms involving products of perturbation quantities have been omitted, since they can be considered negligible compared to the others.

From this point onward, the subscript 1 on perturbation quantity is dropped.

The perturbation potential ψ is the sum of two terms: ψ^E, generated by the Earth's internal mass redistribution, and ψ^M, generated by the external surface load:

$$\psi = \psi^E + \psi^M \tag{6.11}$$

Since the loading mass lies entirely outside the Earth, the associated potential ψ^M satisfies Laplace's equation within the Earth's interior:

$$\nabla^2 \psi^M = 0 \tag{6.12}$$

As a result, the surface mass density γ which generates it does not appear explicitly in the equations of motion within the Earth 6.1 but only in the boundary conditions.

These require that ψ^M, ψ^E are continous at the surface ($r = a$ where a is the radius of the Earth)

$$\psi(r = a_-) = \psi(r = a_+) \tag{6.13}$$

that the quantity $\vec{e_r} \cdot (\vec{\nabla}\psi^E + 4\pi\rho_0 \vec{u})$, where $\vec{e_r}$ is a unit-norm vector normal to the Earth's surface, is continous as well, thus

$$\vec{e_r} \cdot (\vec{\nabla}\psi^E|_{r=a_-} + 4\pi G \rho_0 \vec{u}(r = a_-)) = \vec{e_r} \cdot \vec{\nabla}\psi^E|_{r=a_+} \tag{6.14}$$

since $\vec{u}$ is only defined inside the Earth, and that $\vec{e_r} \cdot \vec{\nabla}\psi^M$ changes by $4\pi G\gamma$:

$$\vec{e_r} \cdot \vec{\nabla}\psi^M|_{r=a_-} + 4\pi G\gamma = \vec{e_r} \cdot \vec{\nabla}\psi^M|_{r=a_+} \tag{6.15}$$

which yields for $\vec{\nabla}\psi = \vec{\nabla}\psi^E + \vec{\nabla}\psi^M$:

$$\vec{e_r} \cdot \nabla\psi|_{r=a_-} + 4\pi G \rho_0 \vec{e_r} \cdot \vec{u}(r = a_-) = \vec{e_r} \cdot \nabla\psi|_{r=a_+} - 4\pi G\gamma \tag{6.16}$$

where γ is the surface mass density.

Boundary conditions for the surface tractions are:

$$\begin{aligned} \hat{\sigma} \cdot \vec{e_\theta} &= 0 \\ \hat{\sigma} \cdot \vec{e_r} &= \gamma \vec{g_0} \end{aligned} \tag{6.17}$$

Numerical solution

The constitutive relation for the elastic rehology is given by Hooke's Law,

$$\sigma_{ij} = 2\mu\epsilon_{ij} + \lambda\epsilon_{kk}\delta_{ij} \tag{6.18}$$

with λ and μ the (radius dependent) Lamé constant. Substituting 6.18 into the perturbed equations (6.10) eliminates the stress variable $\hat{\sigma}$, yielding a system expressed only in terms of displacement and gravitational potential [Martens et al. 2019a].

IN spherical coordinates the displacement component are the radial u_r, u_θ and azimuthal component u_ϕ. Assuming an axis symmetric load distribution, the azimuthal component u_ϕ vanishes. The remaining three variables u_r, u_θ, and ψ can be separated in a radial and an angular part by expanding them in spherical harmonics:

$$\begin{aligned} u_r(r,\theta,t) &= \sum_n U_n(r)\, P_n(\cos\theta)\, e^{i\omega t} \\ u_\theta(r,\theta,t) &= \sum_n V_n(r)\, \frac{\partial P_n(\cos\theta)}{\partial\theta}\, e^{i\omega t} \\ \psi(r,\theta,t) &= \sum_n \Psi_n(r)\, P_n(\cos\theta)\, e^{i\omega t} \end{aligned} \tag{6.19}$$

Subsituting this expression, and the consitutive relation 6.18 into the equation of motion 6.10 leads to a system of second-order ODEs for $U_n(r), V_n(r), \Psi_n(r)$.

The obtained system can be further reduce to a first order system introducing the additional variables T^n_{rr} and $T^n_{r\theta}$ which are the harmonic coefficient of the radial and tangential stress, and the variable Q_n is defined as

$$Q_n = \frac{d\Psi_n}{dr} - 4\pi G\rho_0 U_n \tag{6.20}$$

The radial and tangential stress coefficients:

$$T^n_{rr} = \lambda\left(\frac{dU_n}{dr} + 2\frac{U_n}{r} - \frac{n(n+1)}{r}V_n\right) + 2\mu\frac{dU_n}{dr} \tag{6.21}$$

$$T^n_{r\theta} = \mu\left(\frac{dV_n}{dr} - \frac{V_n}{r} + \frac{U_n}{r}\right) \tag{6.22}$$

are derived from the constitutive law (6.18), and the strain tensor expressions in spherical coordinates:

$$\begin{aligned} \epsilon_{rr} &= \frac{\partial u_r}{\partial r} \\ \epsilon_{r\theta} &= \frac{1}{2}\left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} + \frac{1}{r}\frac{\partial u_r}{\partial\theta}\right) \end{aligned} \tag{6.23}$$

As shown in Martens et al. (2019a), using the variables U_n, T^n_{rr}, V_n, $T^n_{r\theta}$, Ψ_n, $Q_n$the system of three second-order ODEs for U_n, V_n, and Ψ_n is reduced to a form:

$$\frac{d\vec{y}}{dr} = \hat{A}(r)\,\vec{y} \tag{6.24}$$

where the state vector $\vec{y}$ is defined as:

$$\vec{y}(r) = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} U_n \\ T^n_{rr} \\ V_n \\ T^n_{r\theta} \\ \Psi_n \\ Q_n \end{pmatrix} \tag{6.25}$$

and $\hat{A}$ is a six by six matrix depending from r, μ, λ, ρ_0, ω and n.

This system is integrated numerically using the Runge–Kutta method, initializing three linearly independent solutions inside the Earth and propagating them to the surface. Refer to [Martens et al. 2019b] for details on how these independent solutions are initialized, the choice of starting radius for the integration, and the treatment of fluid layers.

The solution has to satisfy the boundary conditions at the surface. These are obtained by inserting the expressions in (6.19) into (6.16):

$$\left.\frac{\mathrm{d}\Psi_n}{\mathrm{d}r}\right|_{r=a_-} - 4\pi G\rho_0 U_n(a_-) = \left.\frac{\mathrm{d}\Psi_n}{\mathrm{d}r}\right|_{r=a_+} + 4\pi G K_n \tag{6.26}$$

where K_n are the harmonic coefficients of the mass distribution:

$$\gamma(\theta) = \sum_n K_n P_n(\cos\theta) \tag{6.27}$$

Outside the Earth ($r > a$), the potential ψ satisfies Laplace's equation and can be written as:

$$\psi(r,\theta,t) = \sum_n \Psi_n \left(\frac{a}{r}\right)^{n+1} P_n(\cos\theta) e^{i\omega t} \tag{6.28}$$

thus:

$$\left.\frac{\mathrm{d}\Psi_n}{\mathrm{d}r}\right|_{r=a_+} = -\frac{n+1}{a}\Psi_n(a_+) \tag{6.29}$$

and equation (6.26) becomes:

$$-4\pi G U_n(a_+) + \frac{n+1}{a}\Psi_n(a_+) = 4\pi G K_n \tag{6.30}$$

In terms of the y-variables:

$$y_6(a_+) + \frac{n+1}{a} y_5(a_+) = 4\pi G K_n \tag{6.31}$$

and for the stress-related variables:

$$\begin{aligned} y_2(a_+) &= -g_0 K_n \\ y_4(a_+) &= 0 \end{aligned} \tag{6.32}$$

A solution satisfying the boundary conditions is obtained by forming a linear combination of the three linearly independent solutions propagated to the surface. Indicating them with y^{I}, y^{II}, y^{III} and letting m_1, m_2, m_3 be the coefficients of the linear combination, the following system must be satisfied:

$$\begin{bmatrix} y_2^{\mathrm{I}} & y_2^{\mathrm{II}} & y_2^{\mathrm{III}} \\ y_4^{\mathrm{I}} & y_4^{\mathrm{II}} & y_4^{\mathrm{III}} \\ y_6^{\mathrm{I}} + \dfrac{n+1}{a} y_5^{\mathrm{I}} & y_6^{\mathrm{II}} + \dfrac{n+1}{a} y_5^{\mathrm{II}} & y_6^{\mathrm{III}} + \dfrac{n+1}{a} y_5^{\mathrm{III}} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} -g_0 K_n \\ 0 \\ 4\pi G K_n \end{bmatrix} \tag{6.33}$$

Once the coefficients m_1, m_2, m_3 are determined, the displacement and potential harmonic coefficient at the surface can be reconstructed as:

$$U_n(a) = m_1\, y_1^{\mathrm{I}} + m_2\, y_1^{\mathrm{II}} + m_3\, y_1^{\mathrm{III}} \tag{6.34}$$

$$V_n(a) = m_1\, y_3^{\mathrm{I}} + m_2\, y_3^{\mathrm{II}} + m_3\, y_3^{\mathrm{III}} \tag{6.35}$$

$$\Psi_n(a) = m_1\, y_5^{\mathrm{I}} + m_2\, y_5^{\mathrm{II}} + m_3\, y_5^{\mathrm{III}} \tag{6.36}$$

from which the full solution for u_r, u_θ, ψ is found using (6.19).

Green Functions

Since the mass distribution enters linearly in the boundary conditions, without loss of generality the response can be computed for surface point mass and then integrated over a finite mass distribution [Farrell 1972]. When γ is a point mass load, the functions decribing the solutions u_r,u_θ,ψ are termed Green Functions for the surface load and indicated respectively with:

$$\Gamma_{u_r}(r, \theta, \phi; r', \theta', \phi') \qquad \Gamma_{u_\theta}(r, \theta, \phi; r', \theta', \phi') \qquad \Gamma_{\psi}(r, \theta, \phi; r', \theta', \phi') \tag{6.37}$$

where the primed coordinates are the source coordinates and the unprimed are the coordinates in which the solution is computed.

The surface potential due to a point mass is [Farrell 1972]:

$$\psi^M = \sum_n \frac{a g_0}{m_e} P_n(\cos\theta) = \sum_n \Psi_n^M P_n(\cos\theta) \tag{6.38}$$

where Ψ^M are the harmonic coefficient of the surface potential associated with the loading mass γ, and m_e is the mass of the Earth. Introducing the set of coefficient called loading Love Numbers k'_l, h'_l, l'_n, and defined as [Farrell 1972]:

$$\begin{aligned} U_n(r) &= h'_n(r)\frac{\Psi_n^M(r)}{g_0} \\ V_n(r) &= l'_n(r)\frac{\Psi_n^M(r)}{g_0} \\ \Psi_n^E(r) &= k'_n(r)\Psi_n^M(r) \end{aligned} \tag{6.39}$$

the Green functions (6.37) are written as a sum of harmonic terms as:

$$\begin{aligned} \Gamma_u(r,\theta) &= \frac{a}{m_e}\sum_{n=1}^{\infty} h'_n(r)\, P_n(\cos\theta) \\ \Gamma_v(r,\theta) &= \frac{a}{m_e}\sum_{n=1}^{\infty} l'_n(r)\, \frac{\partial P_n(\cos\theta)}{\partial\theta} \\ \Gamma_\psi(r,\theta) &= \frac{ag_0}{m_e}\sum_{n=1}^{\infty}(1+k'_n(r))\, P_n(\cos\theta) \end{aligned} \tag{6.40}$$

The employed definition of the Love Numbers follows Farrell (1972) which express them as function of r. When the radial argument is not specified the value at surface of the Earth ($r = a$) is to be taken.

By the superposition principle, once we obtained the Green Function the displacement and the potential for a finite load is computed convolving the Green Functions with the surface mass distribution.

$$\begin{aligned} u_r(r,\theta,\phi) &= \int_0^{2\pi}\int_0^{\pi} \Gamma_{u_r}(r,\theta,\phi;r'=a,\theta',\phi')\gamma(\theta',\phi')a^2\sin(\theta')d\theta'd\phi' \\ u_\theta(r,\theta,\phi) &= \int_0^{2\pi}\int_0^{\pi} \Gamma_{u_\theta}(r,\theta,\phi;r'=a,\theta',\phi')\gamma(\theta',\phi')a^2\sin(\theta')d\theta'd\phi' \\ \psi(r,\theta,\phi) &= \int_0^{2\pi}\int_0^{\pi} \Gamma_{\psi}(r,\theta,\phi;r'=a,\theta',\phi')\gamma(\theta',\phi')a^2\sin(\theta')d\theta'd\phi' \end{aligned} \tag{6.41}$$

where the integral span the whole extension of the surface mass load on the surface ($r = a$).

LoadDef

The required inputs of the LoadDef model are a model of the interior properties of the Earth, a load distribution model and a set of observation points in which to compute the response.

I employed the Preliminary Reference Earth Model (PREM). The load model can be time dependent or static. If time dependent, the load model has to be provided as an amplitude and phase value associated to a location. The phase value is used to model harmonic load, such as the one of tidal origin. In our application this value was set to zero, which corresponds to istantaneous load application [Martens et al. 2019b]. The computation procedure employs threes steps:

1. Computation of the loading Love Numbers for the provided model of Earth structure and rheology. This is done by numerical integrating the equation of motion with surface boundary conditions associated to a point mass. This yield the three solution U_n, V_n, Ψ_n from which the Love Numbers at the surface are computed as:

$$h'_n = \frac{U_n(a)}{a} \tag{6.42}$$

$$l'_n = \frac{V_n(a)}{a} \tag{6.43}$$

$$k'_n = \frac{\Psi_n(a)}{a\, g_s} - 1 \tag{6.44}$$

2. Green Functions computation from the values of the Love Numbers. This consists in numerically performing the sum of Legendre polynomials in equation 6.40. See Martens et al. 2019b for information on how this is done to ensure convergence.

3. Performing the convolution 6.41 on a numerical grid. The grid is defined for each observation point dividing concentrical circles around the point location. Both the Green Function and the load model are interpolated on the grid. Then the convolution is performed as an arithmetical sum.

6.1 Poroelastic model

The description of poroelastic dynamics requires variables for both the solid matrix and the pore fluid. The former is described with the strain and stress tensors, while the latter is described with the fluid mass content in a control volume, Δm, and the pore pressure p. The constitutive relation for linear poroelasticity can be written with any pair of a tensor variable and a scalar variable in dependence on the other two. Using strain and mass content as dependent variables the constitutive relation results in [Segall 2010]:

$$\begin{aligned} E\epsilon_{ij} &= (1+\nu)\sigma_{ij} - \nu\sigma_{kk}\delta_{ij} - (1-2\nu)\alpha p\delta_{ij} \\ \Delta m &= \frac{(1-2\nu)\alpha\rho_w}{2\mu(1+\nu)}\left[\sigma_{kk} + \frac{3}{B}p\right]. \end{aligned} \tag{6.45}$$

where E is the Young's Modulus, ν the Poisson's ratio, α is the Biot-Willis coefficient, and B the Skempton coefficient and ρ_w the denisty of water.

The description of the complete dynamics involves the Navier-Cauchy quasi-static equilibrium equation:

$$\nabla \cdot \hat{\sigma} + \vec{F} = 0 \tag{6.46}$$

the strain compatibility equations:

$$\varepsilon_{ij,kl} + \varepsilon_{kl,ij} - \varepsilon_{ik,jl} - \varepsilon_{jl,ik} = 0 \tag{6.47}$$

the continuity equation for the fluid mass.

$$\frac{\partial q_i}{\partial x_i} + \frac{\partial m}{\partial t} = 0 \tag{6.48}$$

where $\vec{q}$ is the mass flux vector, and the Darcy's law:

$$\vec{q} = -\rho_w \frac{\kappa}{\eta} (\nabla p - \rho_w \vec{g}) \tag{6.49}$$

where η is the fluid viscosity and the permeability k is assumed to be isotropic.

Combining equations 6.47,6.48,6.49, neglecting the term related to gravity in the Darcy's law, yields to two coupled equation for the three components of the displacement field and the pore pressure [Segall 2010]:

$$\mu \nabla^2 u_i + \frac{\mu}{(1-2\nu)} \frac{\partial^2 u_j}{\partial x_i \partial x_j} - \alpha \frac{\partial p}{\partial x_i} + f_i = 0 \tag{6.50}$$

$$S_\alpha \frac{\partial p}{\partial t} - \frac{\kappa}{\eta} \nabla^2 p = -\alpha \frac{\partial}{\partial t} (\nabla \cdot \mathbf{u}) \tag{6.51}$$

with:

$$S_\alpha = \frac{3\alpha(1-2\nu)(1-\alpha B)}{2\mu B(1+\nu)}. \tag{6.52}$$

The first equation is analogous to the equation of motion in traditional elasticity with the additional term $-\alpha \nabla p$ representing the effect of the pore pressure gradient which enters as a body force distribution. The second is a diffusion equation for the pore pressure. The term $\alpha \frac{\partial}{\partial t} \nabla \cdot u$ on the right-hand side of the equation represents the coupling with the displacement field, entering this equation as a source term. The equations are uncoupled if this term is neglected. In this case, we consider a one-way coupling, in which variations of pore pressure produce displacement, but the displacement does not influence pore pressure. We can then first solve for the equilibrium pore pressure distribution using equation 6.51 and then insert the result in equation 6.50 to solve for the displacements [Wang 2001]. The poroelastic problem in this case is analogous to the thermoelastic

problem, with α analogous to the thermal expansion coefficient and the pore pressure to the temperature.

The poroelastic deformation in this work has been described using the model developed by Larochelle, Chanard, et al. (2022). This model employs the eigenstrain theory to obtain the vertical deformation of a disk of porous material embedded in an elastic layer. Eigenstrain generally refers to deformation due to intrinsic material properties, i.e., not associated with external stress, which, in the absence of a restraining medium, would produce uniform deformation. In the case of poroelastic deformation, eigenstrain arises from the variation of the pore pressure, associated with the variation of the water table inside the aquifer [Mura 1982].

Assume that a uniform disk of radius a and of porous material, which is uniformly saturated by water for a thickness b, undergoes an increase in water storage. This is reflected in an increase in the water table h of a quantity Δh. This produce an increase in the pore pressure of :

$$\Delta p = \rho_w g \Delta h_w \tag{6.53}$$

over all the saturated thickness.

In the absence of the restraining medium, for an increase in pore pressure, the material would undergo an uniform expansion given by Larochelle, Chanard, et al. (2022):

$$\epsilon_{eig} = \frac{\alpha \Delta p (1-2\nu)}{E_{aq}} = \frac{\alpha \rho_w g \Delta h (1-2\nu)}{E_{aq}} \tag{6.54}$$

Employing a system of cyclindrical coordinate the first equation in 6.45 becomes:

$$\begin{aligned}
\varepsilon_{zz} &= \frac{1}{E_{\text{aq}}} \left[(1+\nu)\, \sigma_{zz} - \nu \left(\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz} \right) \right] + \varepsilon_{\text{eig}} \\
\varepsilon_{rr} &= \frac{1}{E_{\text{aq}}} \left[(1+\nu)\, \sigma_{rr} - \nu \left(\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz} \right) \right] + \varepsilon_{\text{eig}} \\
\varepsilon_{\theta\theta} &= \frac{1}{E_{\text{aq}}} \left[(1+\nu)\, \sigma_{\theta\theta} - \nu \left(\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz} \right) \right] + \varepsilon_{\text{eig}}
\end{aligned} \tag{6.55}$$

Larochelle, Chanard, et al. (2022) assume that the horizontal strain $\varepsilon_{\theta\theta}$ and ε_{rr} are much smaller than the eigenstrain due to the effect of the restraining medium, thus neglect them in the equation 6.55, obtaining:

$$\sigma_{rr} = \sigma_{\theta\theta} = \frac{-E_{\text{aq}}\, \varepsilon_{\text{eig}} + \nu\, \sigma_{zz}}{1-\nu} \tag{6.56}$$

where σ_{zz} is the change in total vertical stress associated with a change in groundwater level:

$$\sigma_{zz} = -\phi\, \rho_w g\, \Delta h \tag{6.57}$$

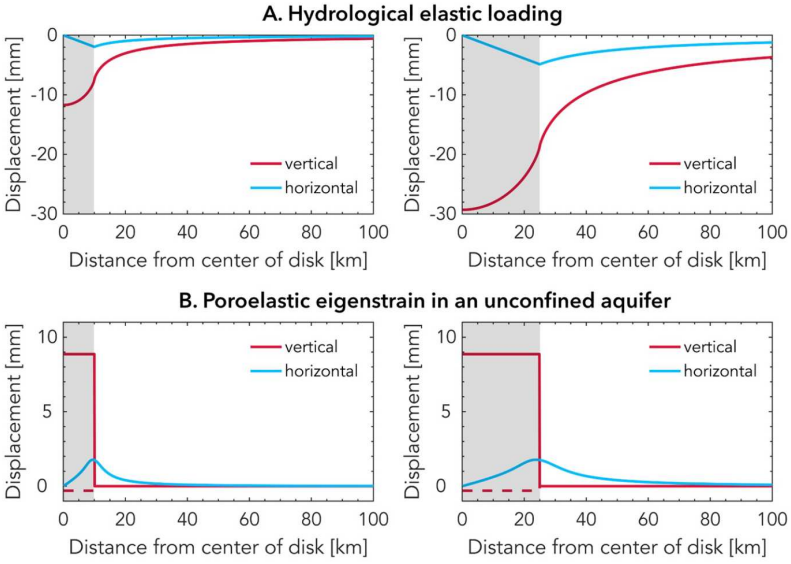


Figure 6.2: The figure shows an example of the radial dependency of the solution for a disk-shaped aquifer. The function is compared with the elastic response associated with the load variation inside the same disk. From the comparison of vertical deformation responses, we see that while the vertical elastic deformation is negative and affects points outside the load application, the vertical poroelastic deformation is positive above the aquifer and zero outside, reflecting the local character of the deformation.

Inserting 6.57 and 6.56 in the first of the Equations 6.55, a differential equation for the vertical displacement is obtained, which can be integrated from the aquifer base through all the saturated thickness of the aquifer (b), yielding [Larochelle, Chanard, et al. 2022]:

$$u_z(r) = \begin{cases} \dfrac{(1+\nu)(1-2\nu)}{(1-\nu)} \dfrac{(\alpha-\phi)\,\rho_w g\,\Delta h\, b}{E_{\text{aq}}}, & r \leq a \\ 0, & r > a \end{cases} \tag{6.58}$$

The equation therefore represents the vertical displacement, at the surface, associated with a water level variation Δh within a porous disk of radius a with porosity ϕ and Biot-Wills coefficient α. The deformation is proportional to the thicnkess of the aquifer b and to the water table variation Δh and it is zero outside the aquifer.

6.2 Application

In the time interval considered in this study, the trend of hydrological variables experienced significant multiannual fluctuations in the study area, as described in Section 2.3. Using GNSS data over the whole Po Plain, Pintori et al. (2024) characterized

the hydrologically-related ground deformation at the regional scale during the period 2010–2023. They identified a multi-year deformation signal strongly correlated with Po River level time series and the SPEI-12 (Standardized Precipitation and Evapotranspiration index) (Figure 6.3). By modeling this deformation signal as an elastic response they estimated the spatial distribution of the total water storage variation (Figure 2.5).

The analysis of EGMS data confirmed the presence of deformation related to multi-year hydrological variability. However, the spatial resolution of this dataset is higher than that of GNSS data, allowing for the identification of hydrologically-related deformation processes at smaller spatial scales. In Chapter 5, I showed that the application of different data analysis methods to the dataset identifies a set of areas, with spatial extents on the order of 1 to 10 km, uplifting in rainy periods and subsiding in drought periods. Here therefore poroelastic effects are likely the predominant deformation mechanism.

The area characterized by the most evident poroelastic signal is located in the surroundings of Brescia. Geologically, the terrain here consists of a layer of gravel and sand deposits, over 100 m thick, hosting an unconfined aquifer (Section 2.4).

I assumed that the observed deformation in this area ΔU is the superposition of an elastic contribution ΔU^e — associated with the regional-scale, multi-year variation in total water storage (Figure 6.5) — and a contribution due to poroelastic processes occurring at smaller spatial scales within the sediment layers ΔU^p:

$$\Delta U = \Delta U^e + \Delta U^p \tag{6.59}$$

Under this assumption I analyzed the cumulative deformation in two periods characterized by different hydrological conditions: a 'rainy' period, indicated as T_1, spanning from August 2018 to January 2021, and a 'drought' period, indicated as T_2 from Janury 2021 to August 2022 (Figure 6.5). To speed up computations, the modeling was performed on a regular 500 m × 500 m rectangular grid, shown in Figure 6.4.

The EGMS deformation data were resampled onto this grid by averaging all time series located within each grid cell and assigning the result to the corresponding grid center. Let $p \in \{1, \ldots, M\}$ index the grid points containing at least one EGMS time series, and let t_i with $i \in \{1, \ldots, N\}$ denote the time steps. Then the resampled deformation value at time t_i for grid point x_p is denoted as:

$$\Delta U(t_i, x_p) \tag{6.60}$$

Let t_1^s and t_1^f be the starting and final times of the first interval T_1 and t_2^s, t_2^f be the starting and final time for the second period T_2, the cumulative deformation during each period is computed as:

$$\begin{aligned} \Delta U_{T_1}(x_p) &= \Delta U(t_1^f, x_p) - \Delta U(t_1^s, x_p) \\ \Delta U_{T_2}(x_p) &= \Delta U(t_2^f, x_p) - \Delta U(t_2^s, x_p) \end{aligned} \tag{6.61}$$

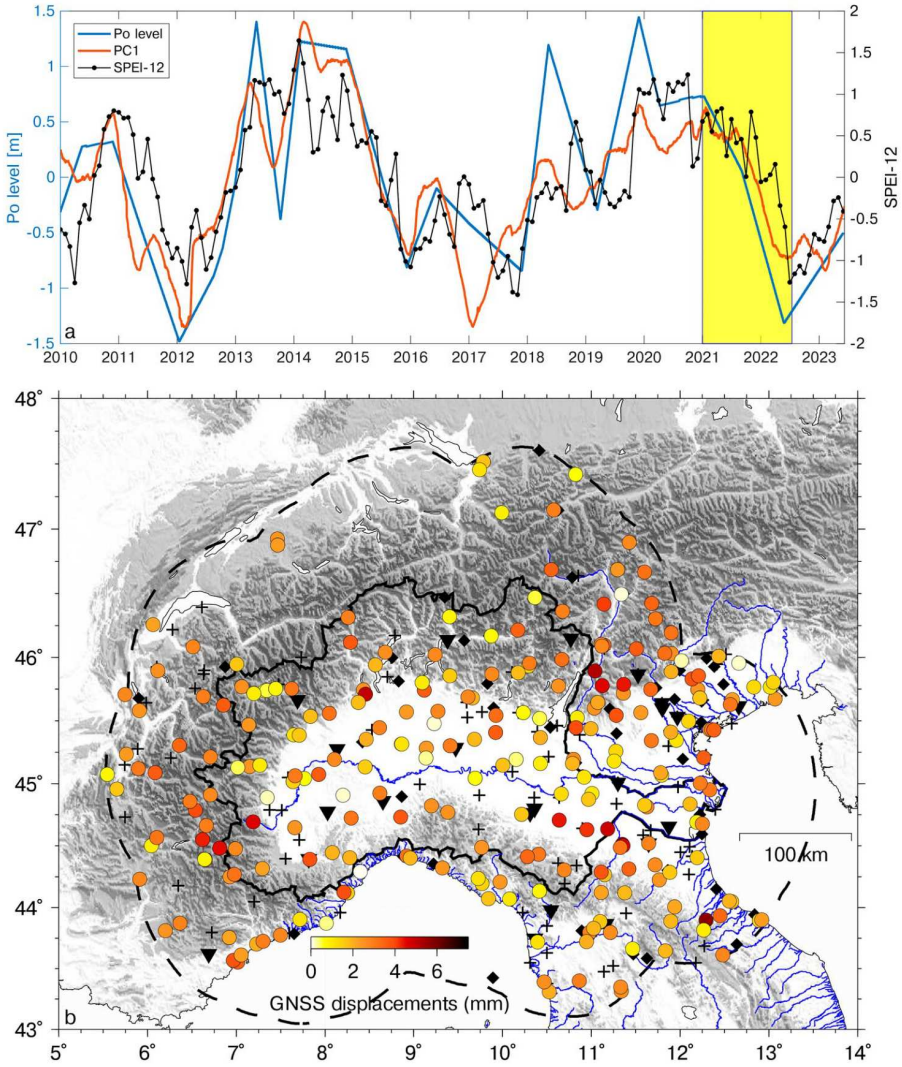


Figure 6.3: Temporal trend (top) and spatial amplitud (bottom) of hydrologically-related deformation, from PCA applied to GNSS vertical displacement. Temporal trend is compare with Po River level, and SPEI-12 index. [Pintori et al. 2024]

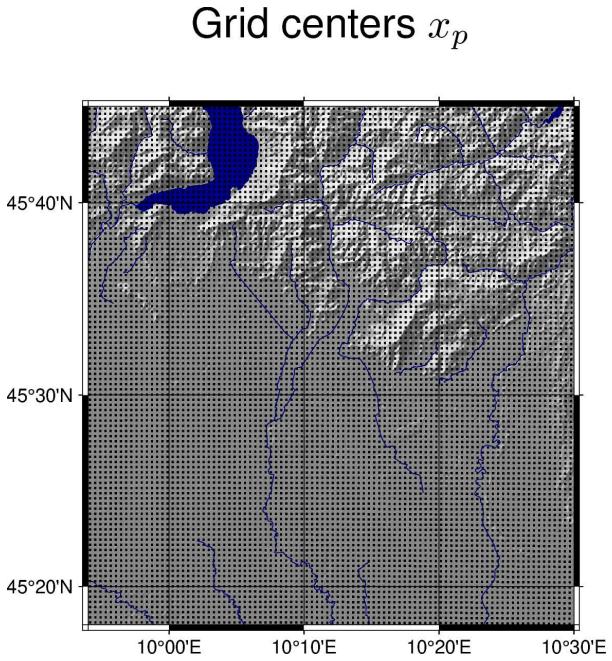


Figure 6.4: Regular grid used for elastic deformation computation with LoadDef.

I employed the models described in the previous section to quantify the two contributions to the deformation in each period:

$$\begin{aligned} \Delta U_{T_1}(x_p) &= \Delta U^e_{T_1}(x_p) + \Delta U^p_{T_1}(x_p) \\ \Delta U_{T_2}(x_p) &= \Delta U^e_{T_2}(x_p) + \Delta U^p_{T_2}(x_p) \end{aligned} \tag{6.62}$$

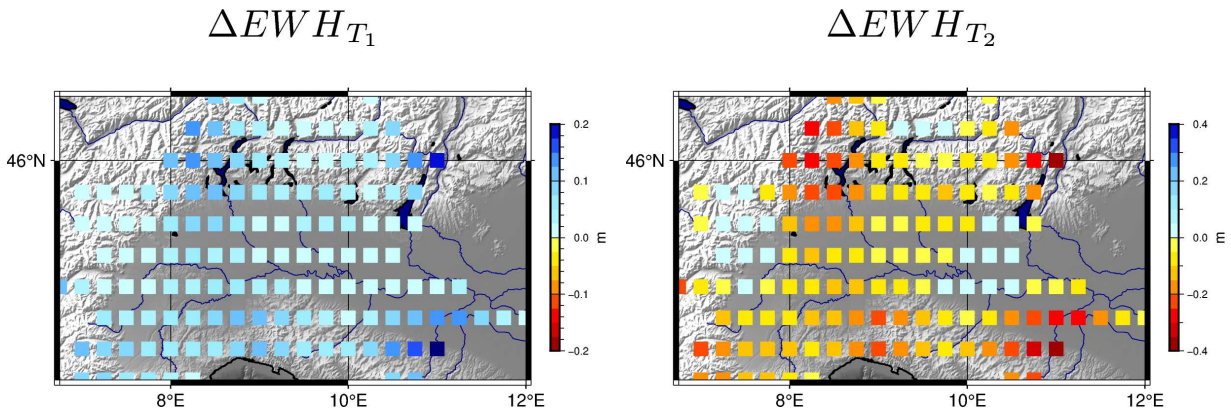
To compute the elastic component of deformation, I used the Equivalent Water Height (EWH) variations shown in Figure 6.5, for T_1 and T_2, as input to the LoadDef software. The output of the two integrations are the elastic contribution to deformation at each gridpoint in the two periods:

$$\Delta U^e_{T_1}(x_p) \quad \text{and} \quad \Delta U^e_{T_2}(x_p) \tag{6.63}$$

The poroelastic contribution is then obtained by subtracting the modeled elastic deformation from the observed cumulative deformation for each gridpoint.

$$\begin{aligned} \Delta U^p_{T_1}(x_p) &= \Delta U_{T_1}(x_p) - \Delta U^e_{T_1}(x_p) \\ \Delta U^p_{T_2}(x_p) &= \Delta U_{T_2}(x_p) - \Delta U^e_{T_2}(x_p) \end{aligned} \tag{6.64}$$

To compare poroelastic deformation with water level data, I averaged $\Delta U^p(x_p)$ within a radius of 600 m around each piezometer, obtaining a cumulative deformation value representative of the local behaviour at each site. The standard deviation of the values



(a) (b)

Figure 6.5: Equivalent Water Height (EWH) variation during the rainy period T_1 (a) and drought period T_2 (b), from [Pintori et al. 2024], used as input for elastic deformation modeling.

within each buffer area was used as the uncertainty associated with the deformation estimate. I then used Equation(6.58) to estimate the Young's modulus of the aquifer at each well location.

$$E_{\text{aq}} = \frac{(1+\nu)(1-2\nu)}{(1-\nu)} \cdot \frac{(\beta-\phi)\,\rho_w\,g\,\Delta h\,b}{\Delta U^p} \tag{6.65}$$

where Δh is the water table variation at the considered well and ΔU is the poroelastic deformation measured in the same time interval in correspondence of the same well.

Porosity values of sediment retrived from geological reports [ARPALombardia 2018a] are 0.3 for the gravel layer and 0.2 for the underlying conglomerate. I therefore used the average value of 0.25 along the whole sediment thickness. The Biot–Willis coefficient was set to 0.8, a value typically adopted in literature in the absence of site-specific information for the study area [Carlson et al. 2024, Larochelle, Chanard, et al. 2022] . The Poisson ratio was set to 0.25.

The computational uncertainty associated with the Young's modulus was then estimated using standard error propagation:

$$\delta E_{aq} = \frac{(1+\nu)(1-2\nu)}{(1-\nu)} \cdot (\beta-\phi) \cdot \rho_w \cdot g \cdot \sqrt{\left(\frac{hb\,\mathrm{d}\Delta U}{\Delta U^2}\right)^2 + \left(\frac{h\,\mathrm{d}b}{\Delta U}\right)^2 + \left(\frac{\mathrm{d}h\,b}{\Delta U}\right)^2} \tag{6.66}$$

Additional sources of error, which have not been modeled, could be associated with the model assumptions. These are discussed in Section 7.

6.2.1 Local elastic response

Up to this point, I assumed that the elastic component of the deformation was solely due to regional-scale water storage variations (Figure 6.5) from GNSS. These have a

spatial resolution of 0.25° [Pintori et al. 2024], thus are to be considered average values over a gridcell of approximately 30 km side. Assuming that the elastic response is entirely due these forcing I neglected the effect of heterogeneities of the surface load distribution at the smaller scale, such as the ones that are associated to water mass changes in lakes and groundwater reservoirs. These could be responsible for small-scale features in the observed deformation signal which are due to elastic effects not considered in our model. Neglecting these contributions could lead to an inaccurate estimate of the amplitude of the poroelastic signal and distort the inferred value of the Young's modulus E_{aq}.

To assess this effect, I additionally computed the elastic deformation associated with the change in water storage inside the considered aquifer.

At this aim, I considered a simplified model of the piezometric level variation inside the aquifer in the period T_1 (Figure 6.6). The aquifer is assumed to be bounded by shown perimeter, which was derived by tracing the results of the K-means clustering (Section 5.3). I performed a two-dimensional linear interpolation of the water table variation data at the water wells. I then assigned the resulting values to grid points falling whitin the perimeter. The equivalent water height variation (ΔEWH) was then obtained multiplying the water table variation by the porosity ϕ.

$$\Delta\text{EWH}(x'_p) = \phi \Delta h(x'_p) \tag{6.67}$$

where x'_p indicates gridpoints falling whitin the perimeter.

The value of $\Delta\text{EWH}(x'_p)$ where used as a load model to force LoadDef. The response was computed on the same 500 m × 500 m grid shown in Figure 6.4. The resulting deformation is indicated with $\Delta U^{e,l}_{T_1}(x_p)$.

The total deformation is now assumed to be composed of three components:

$$\Delta U_{T_1}(x_p) = \Delta U^{p}_{T_1}(x_p) + \Delta U^{e,r}_{T_1}(x_p) + \Delta U^{e,l}_{T1}(x_p) \tag{6.68}$$

where $\Delta U^{e,r}_{T_1}(x_p)$ is the regional elastic response previously defined as $\Delta U^{e}_{T_1}(x_p)$. The estimate of the poroelastic component was then updated by subtracting both the regional and local elastic contributions from the observed deformation.

$$\Delta U^{p}_{T_1}(x_p) = \Delta U_{T_1}(x_p) - \Delta U^{e,r}_{T_1}(x_p) - \Delta U^{e,l}_{T1}(x_p) \tag{6.69}$$

The Young's modulus was recomputed accordingly comparing the new estimate for the poroelastic deformation with groundwater level data.

6.3 Results

In Figure 6.7, the cumulative total deformation $\Delta U_{T_1}(x_p)$ and $\Delta U_{T_2}(x_p)$ for the two periods are shown. It can be seen that, in correspondence with the aquifer perimeter, there is uplift during the rainy period T_1, with average value of 3 mm above the aquifer

$$\Delta EWH$$

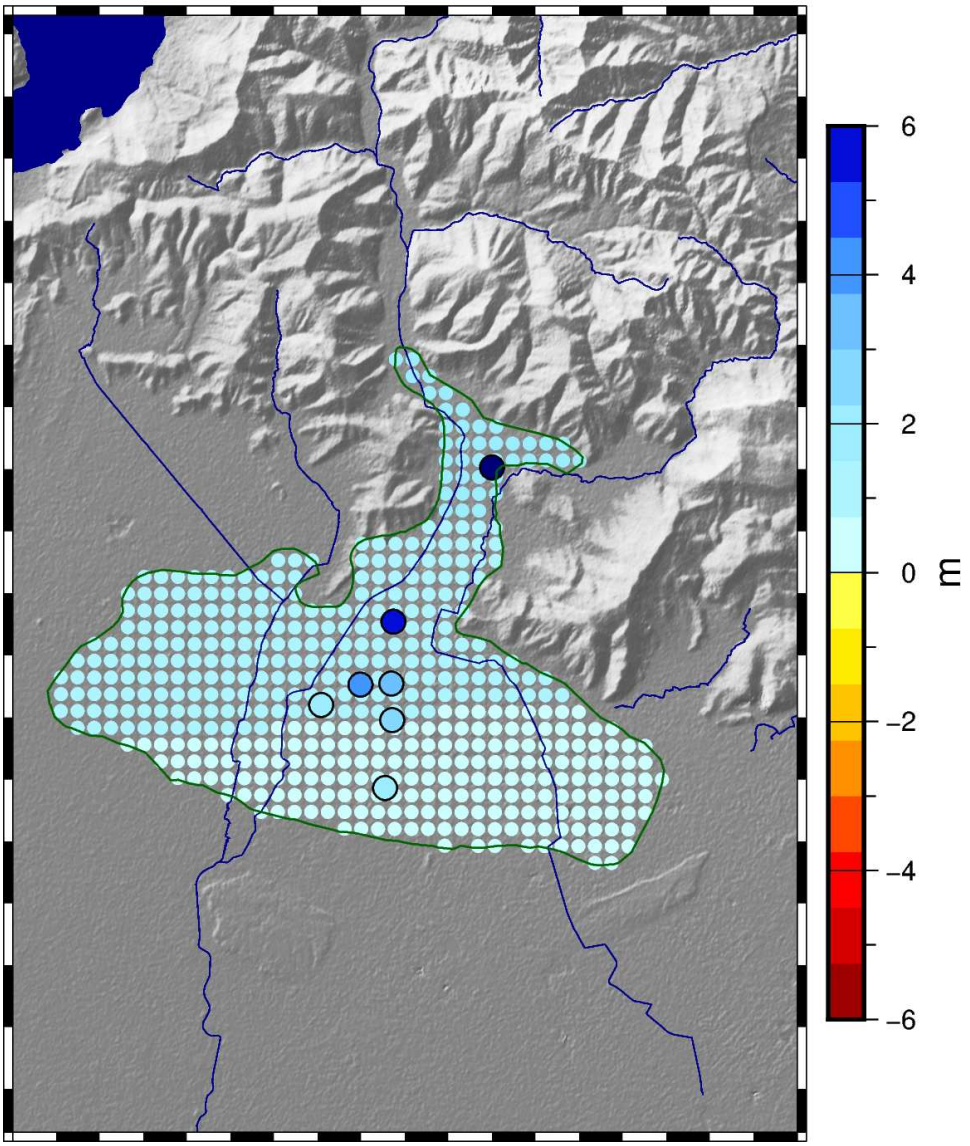
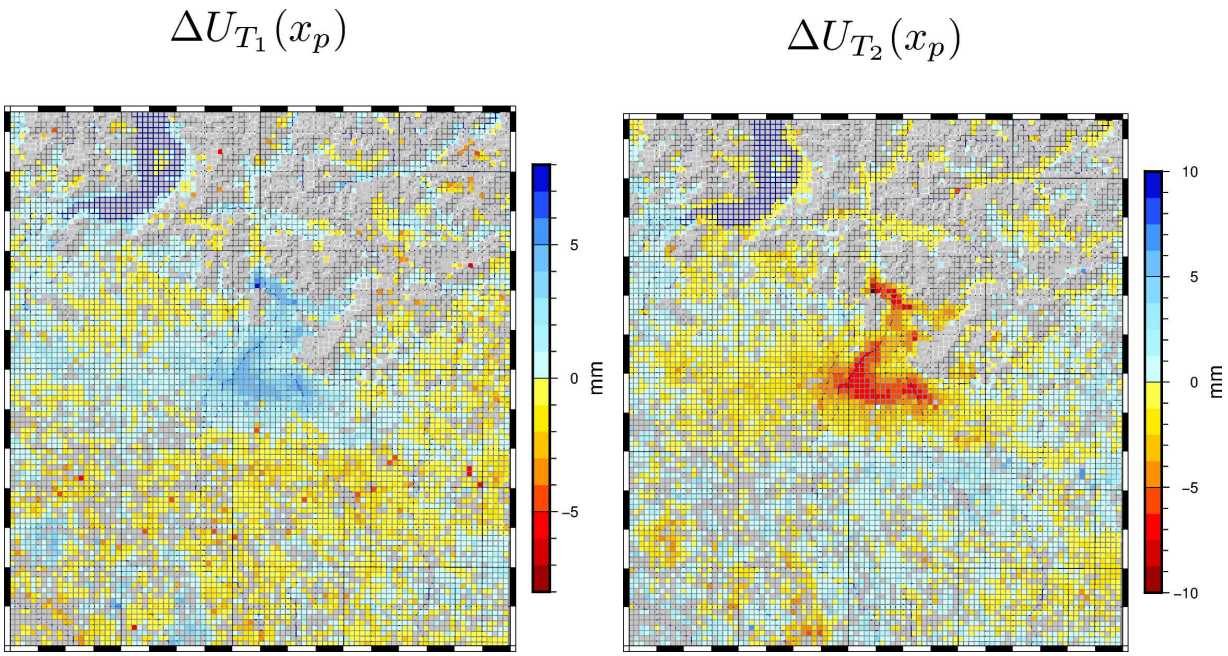


Figure 6.6: Model of the water table variation within the aquifer, used to estimate the local elastic response. Bigger circles are the measured values of water table variation. Smaller circles represent the reconstructed equivalent water heigth distribution



(a) (b)

Figure 6.7: Cumulative deformation computed on the grid

and maxima above 5 mm, and subsidence associated with the drought period T_2, with average values of 5 mm. Maximum values of both uplift and subsidence are reached in the northern part of the area, at the outlet of the Trompia Valley. In the plain to the south, the values become much smaller in absolute magnitude and changes sign, being negative during the rainy period and positive during the drought period, with absolute displacements remaining below two millimeters.

In Figure 6.8, the elastic response $\Delta U^e_{T_2}(x_p)$ and $\Delta U^e_{T_2}(x_p)$ computed with LoadDef are shown. The values of the elastic response are approximately uniform over the study area, around -0.5 mm for T_1 and 1 mm for T_2.

By removing the elastic response to the cumulative deformation I obtained an estimate of the poroelastic contribution (Figure 6.9). Given the uniformity of the elastic response over the studied area, the poroelastic deformation results in a version of the original cumulative deformation shifted towards more negative values in the first period — when the elastic response is of subsidence — and towards positive value in the second — when the elstic response is of uplift. The maximum poroelastic deformation reached is of about 7 mm for the rainy period and -10 mm in the drought period.

The available groundwater time series (Figure 3.15) end before the conclusion of the defined drought period T_2. Therefore, the quantitative comparison between the poroelastic deformations water table variations was only possible for the first period T_1

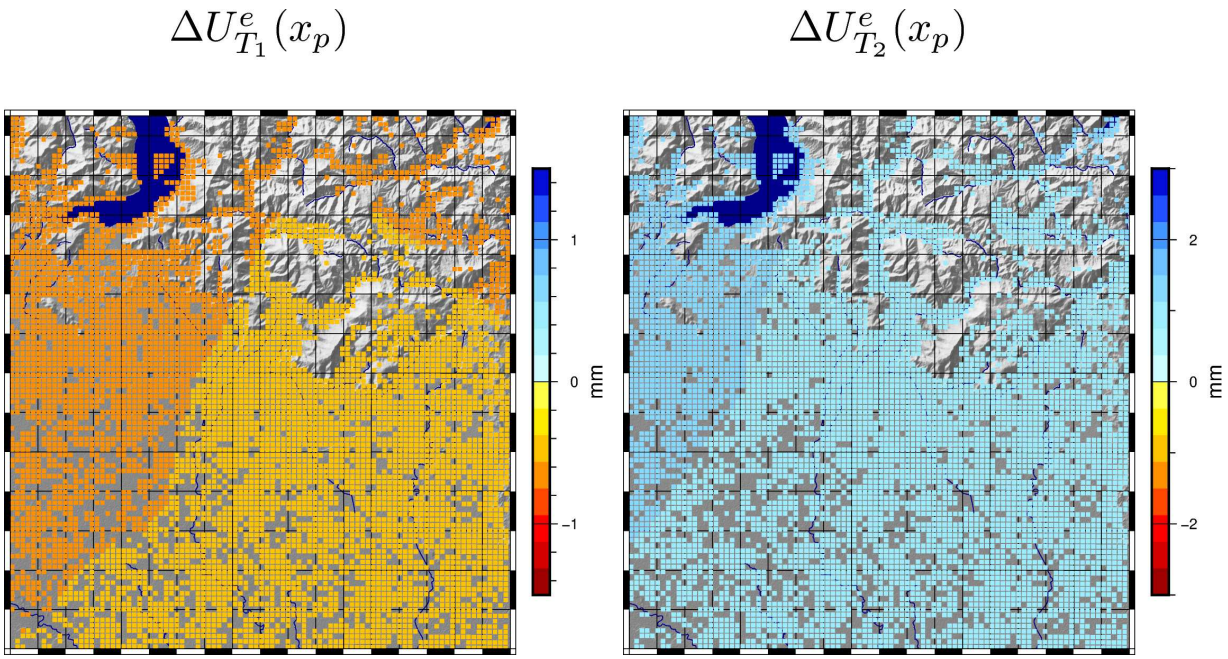


Figure 6.8: Elastic response to the regional elastic load distribution.

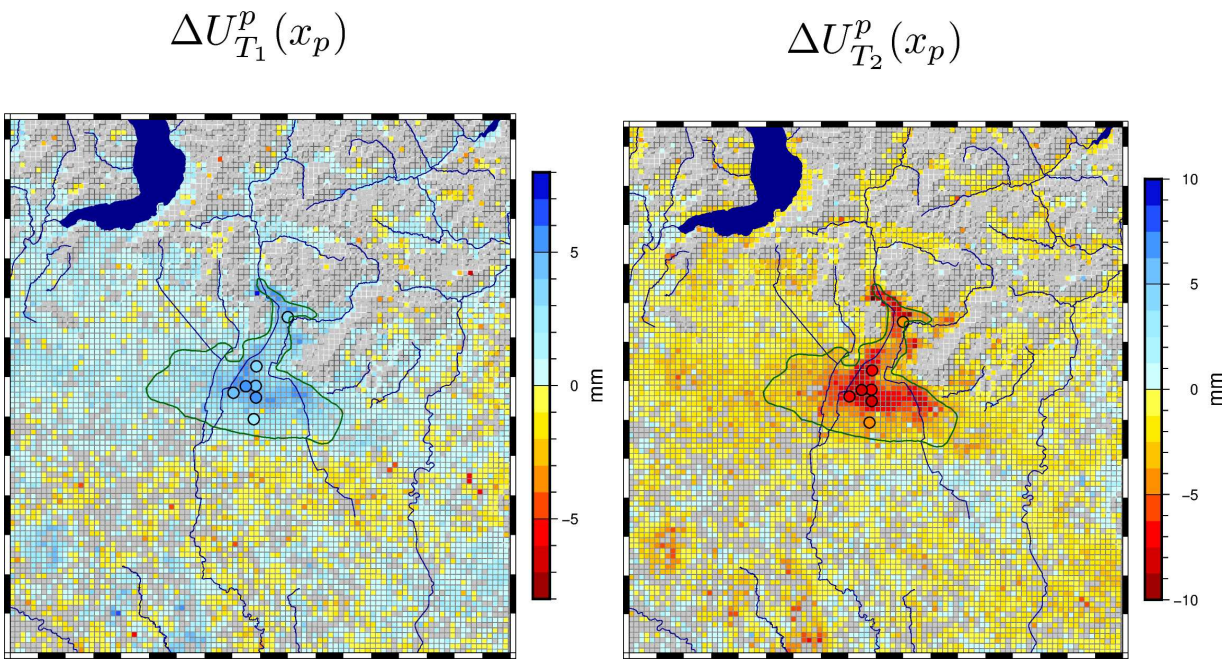
associated with rainy condition. For this period I applied equation 6.66 using the average values of the poroelastic deformation around each well (big circles in Figure 6.9) and the value of the water table variation (Figure 3.16). Figure 6.10 shows the estimated values of the Young's modulus of the aquifer for each piezometer location and their spatial distribution.

The aquifer Young's modulus values span from 200 MPa to 600 MPa. The values are higher in the northern part of the area and lower in the south, closer to the plain. The uncertainty associated with the value also is higher for the piezometer located in the north.

6.3.1 Local elastic deformation correction

In this section, the results obtained by considering the local elastic response due to the aquifer load are shown. Due to the availability of groundwater data only in the first period, this computation has been performed only for this interval.

The deformation (Figure 6.11) is characterized by small values of subsidence, with maximum values around 1 mm. The distribution shows the typical pattern of the elastic response: the maximum subsidence value is found in the central part of the aquifer, and the deformation drops to zero outside the aquifer area.



(a) (b)

Figure 6.9: Poroelastic deformation obtained removing the regional elastic contribution. Small points are the value at the defined gridpoints. Big circles are deformation values in correspondence of each piezometer

The poroelastic deformation in this case is computed by removing both the contribution due to the regional elastic response (Figure 6.8) and the local response (Figure 6.11). The result is shown in Figure 6.12. Maximum uplift values in this case are up to 8 mm.

Using this result for the poroelastic deformation, I repeated the estimate of the aquifer Young's modulus. The obtained results are shown in Figure 6.13. Obtained result in this case are lower since the removal of the additional contribution yield an higher poroelastic deformation. Values span from around 150 MPa to 400 MPa and are characterized by the same latitude dependency of the previous case.

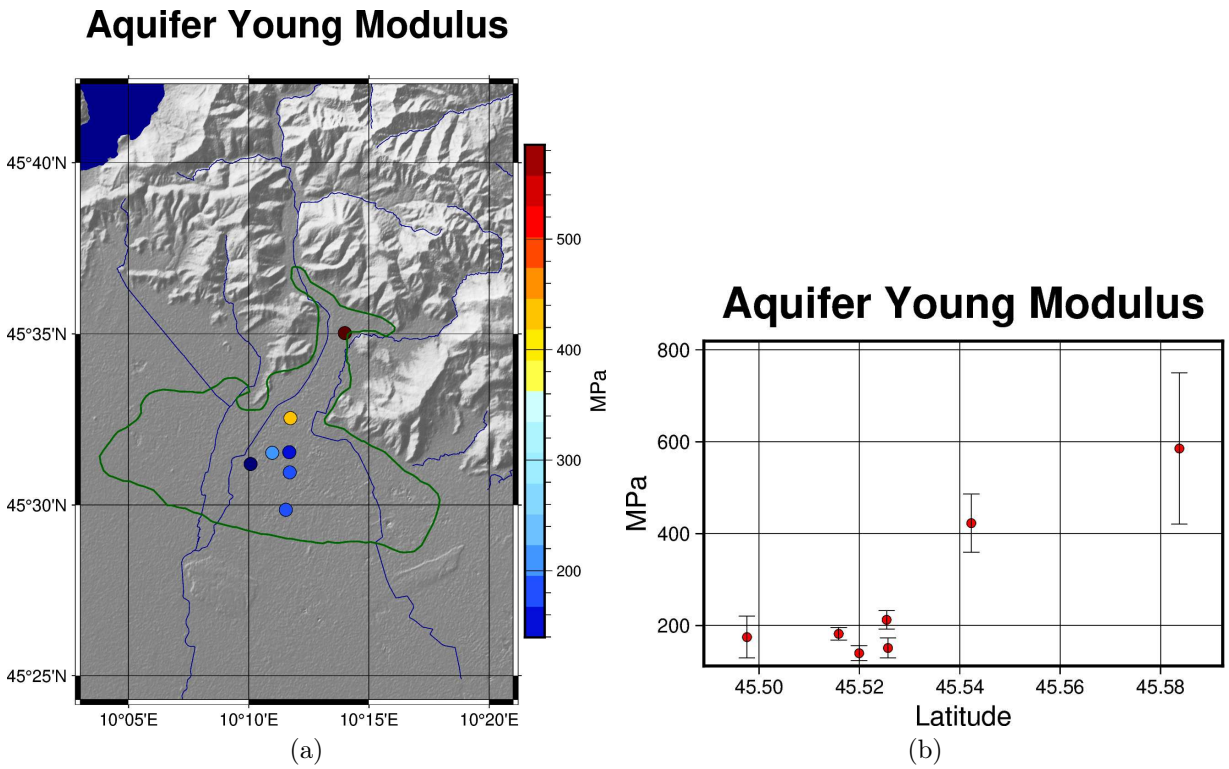


Figure 6.10: Results for the aquifer Young modulus. Spatial distribution of value estimated for each piezometer location (a). Estimated values with uncertainty shown in function of latitude (b)

$$\Delta U_{T_1}^{e,l}$$

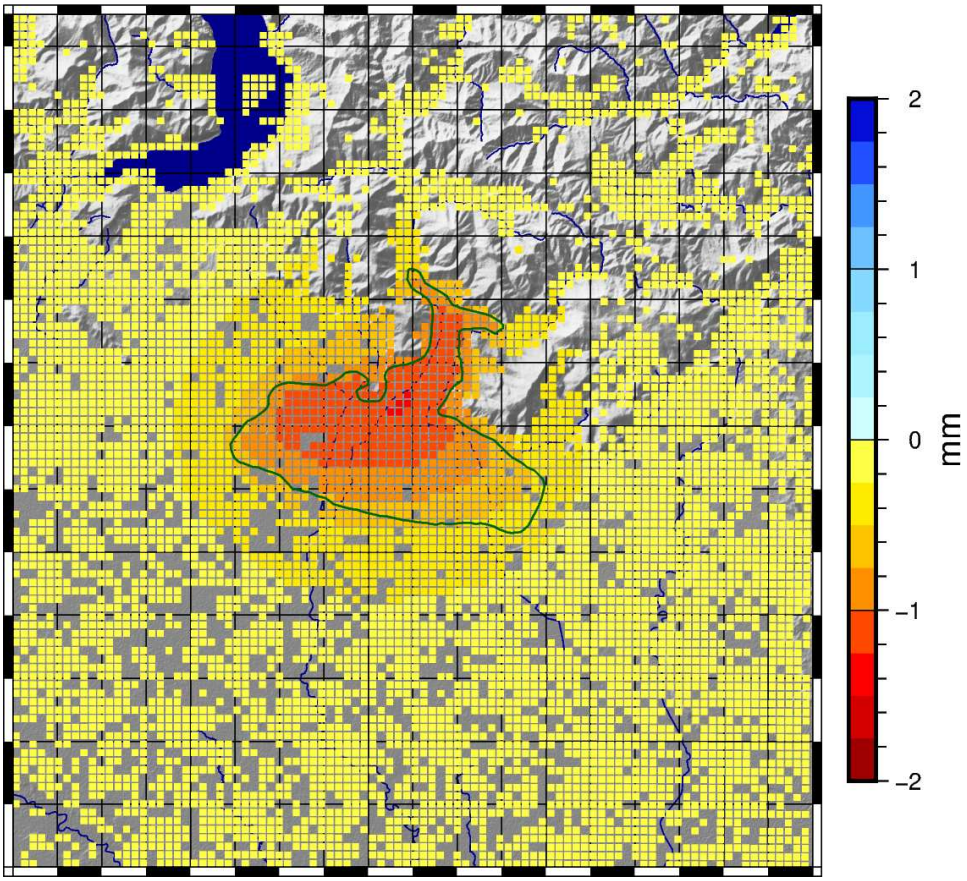


Figure 6.11: Modeled vertical deformation induced by the local elastic response to groundwater level variations within the aquifer during the rainy period T_1. The deformation pattern exhibits maximum subsidence in the central part of the aquifer, gradually decreasing to zero toward the aquifer boundaries.

$$\Delta U^{p}_{T_1}(x_p)$$

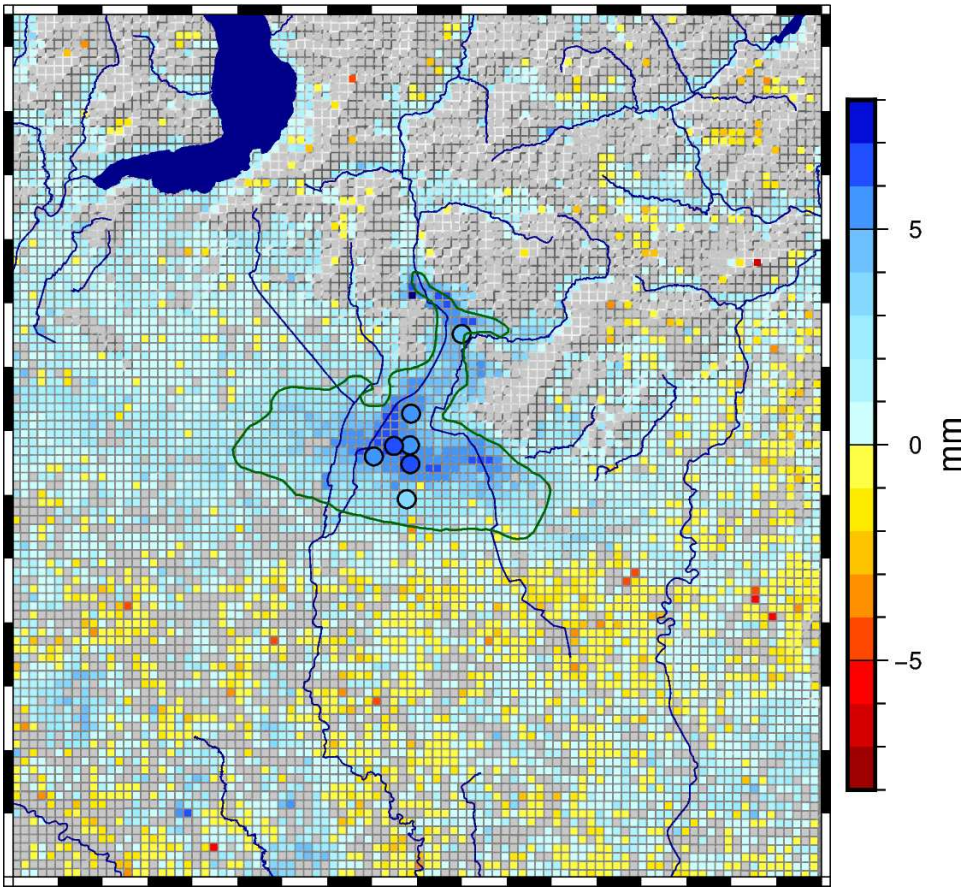


Figure 6.12: Poroelastic deformation obtained removing both regional and local elastic contributions. Small point are the deformation values at the defined gridpoints. Big circles are the average deformation values in correspondence of each piezometer

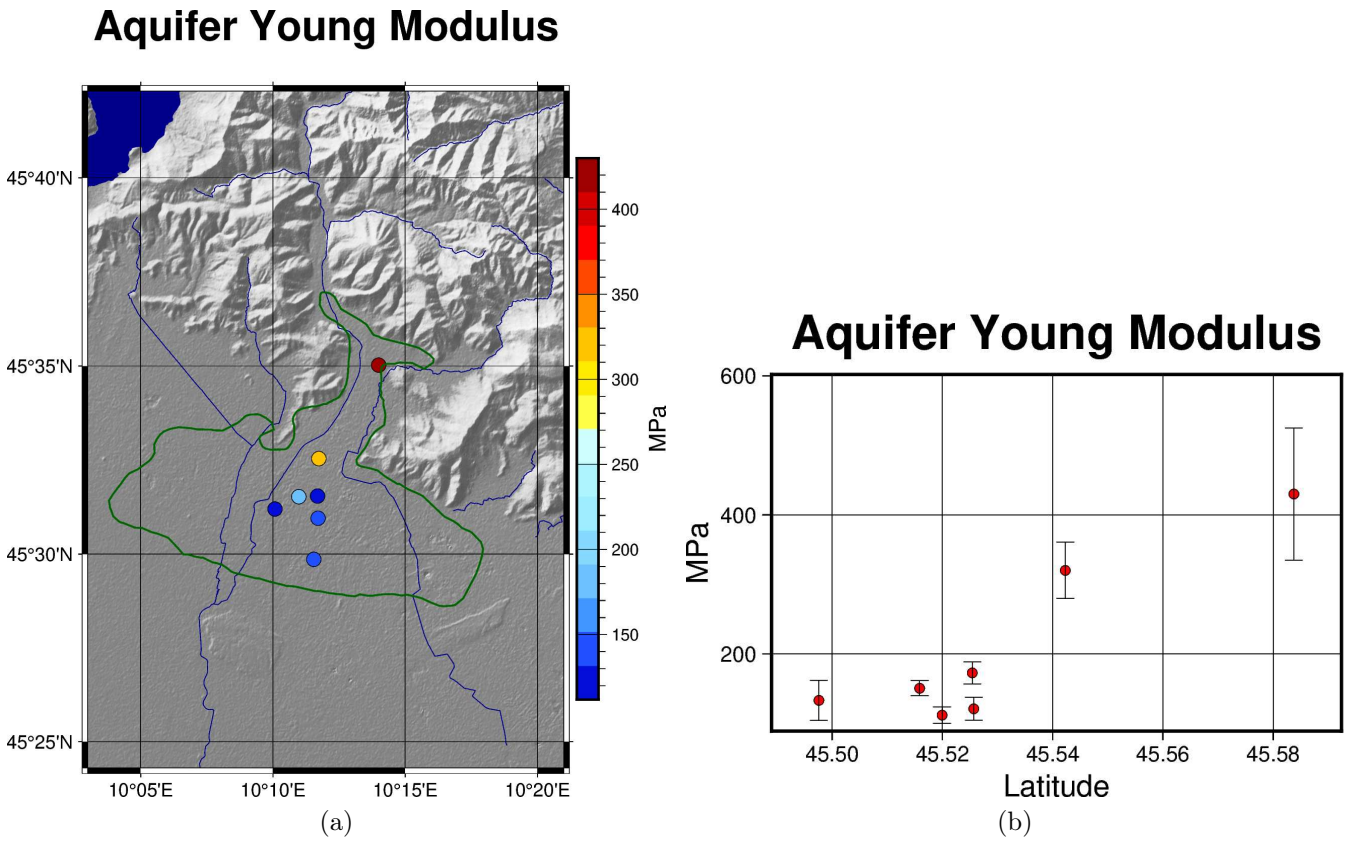


Figure 6.13: Results for the aquifer Young modulus after removing also the load contribution of the local aquifer. Spatial distribution of value estimated for each piezometer location (a). Estimated values with uncertainty shown in function of latitude (b)

Chapter 7

Discussion

I analyzed the EGMS Ortho L3 dataset, containing Sentinel-1 satellites InSAR-derived vertical deformation time series with a 6-day sampling rate from 06/01/2018 to 22/12/2022, on a 100 m × 100 m grid over an area encompassing a section of the Po Plain and part of the Alpine mountain belt. I employed different statistical data analysis techniques to identify relevant deformation processes occurring in the area and to compare their temporal evolution with seasonal and multi-year trends observed in precipitation data measured by the ARPALombardia sensor network.

The majority of the time series in the region are characterized by negative average vertical velocity (Figure 5.1), consistent with the subsidence effects known to occur in the region [Carminati et al. 2002, Farolfi et al. 2019]. The subsidence signal has been investigated in previous studies [Farías et al. 2024] and has been linked to a superposition of natural and anthropogenic processes. In this work I focused on seasonal and multi-year variations; therefore, I removed the estimated linear trend and applied statistical analysis techniques to the detrended time series.

I identified different deformation processes occurring at seasonal and multi-year timescales superimposed to the long-term trend.

The seasonal component is most evident in two small areas nearby the city of Crema, likely associated with the periodic activity of nearby underground gas storage facilities [Rigamonti 2025]. Both the amplitude map of the seasonal STL component and the ICA results highlight these deformation patterns (Figure 5.2 and 5.8). Notably, PCA was not able to isolate this localized seasonal deformation signal when applied to the entire domain, as shown by the comparison between PC3 and PC4 with IC3 and IC4 (Figure 5.8 and 5.7). However, when the analysis was restricted to a smaller area encompassing only the affected region, the two techniques yielded similar results (Figures 7.1 and 7.2) These findings highlight that ICA is more effective than PCA at addressing the blind source separation problem in the presence of multiple deformation sources, while the two techniques yield similar results when a dominant deformation process is associated with most of the variability.

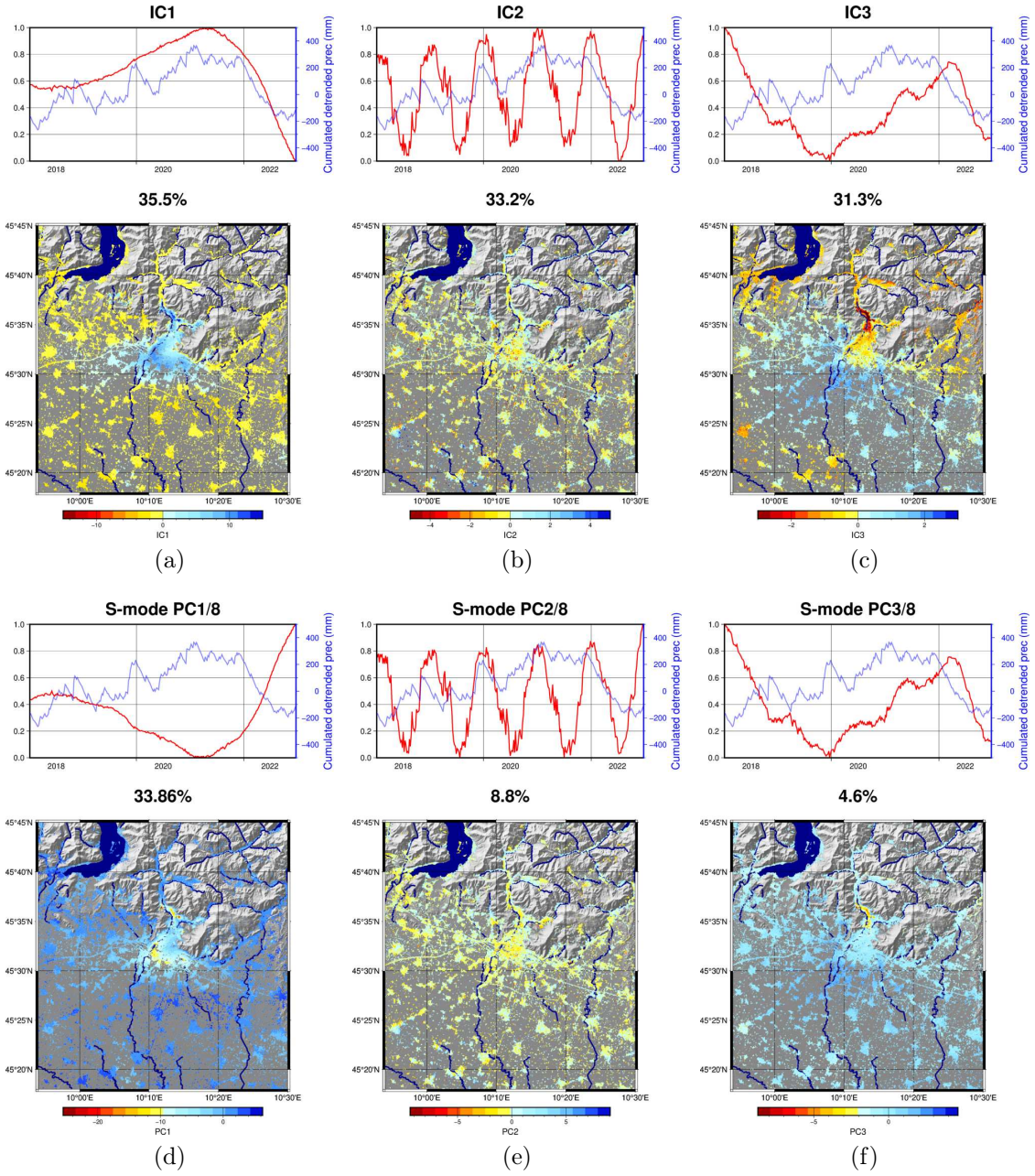


Figure 7.1: Comparison of PCA and ICA performed on the area characterized by the strongest multiyear signal.

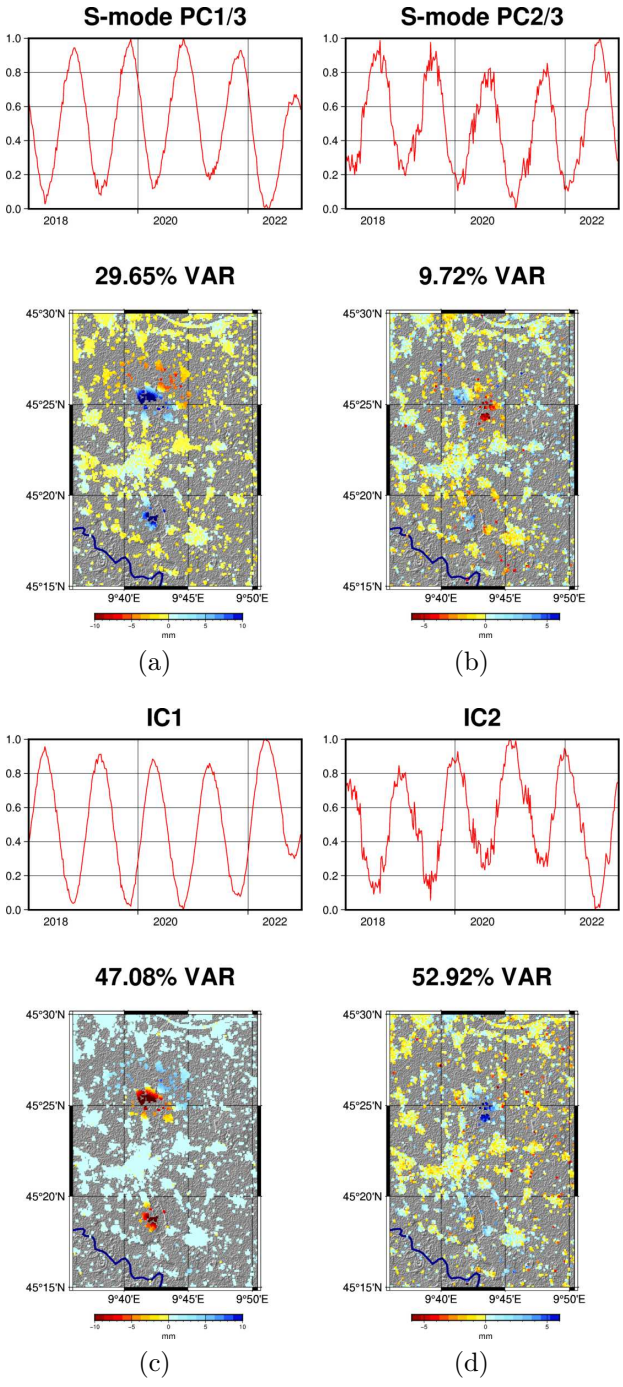


Figure 7.2: Comparison of PCA and ICA performed on the are characterized by the strongest seasonal signal.

From the ICA results (Figures 5.8 and 7.2), it can be seen that the seasonal signal in this region is split into two components. A possible reason is the non-stationarity of the source signal. Larochelle, A. Gualandi, et al. 2018 showed that, since vbICA assumes nonmoving sources, multiple components are necessary to fully recover the effects of a potentially non-static source. For the case of a source moving at a constant speed in a constant direction, the authors found that two components are sufficient to explain the observations and that these two components must be derivatives of one another. Further analysis would be required to verify this hypothesis.

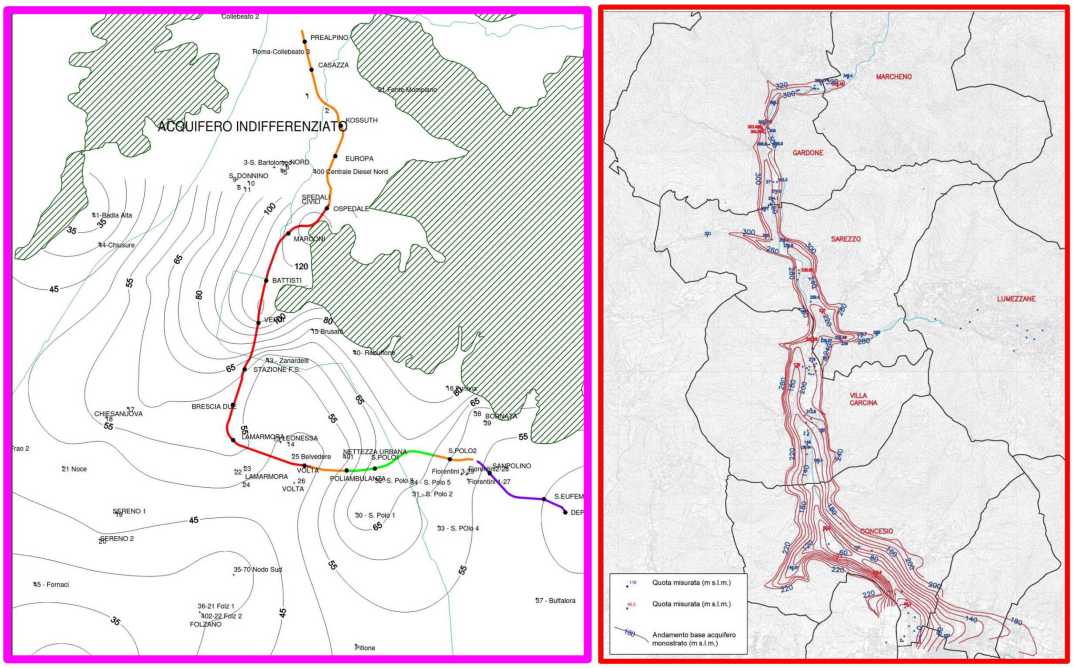
For the multi-year trend, the application of clustering on both the original and STL-filtered time series shows that this deformation component plays a significant role with typical spatial scales ranging from a few kilometers to tens of kilometers (Figure 5.3 and 5.4). Among the main areas showing this deformation pattern, one is located in a mountain valley (Clusone), three lie on the foothills of the Alpine valleys (including Brescia and southern Bergamo), and one is situated on the foothills of the Apennines (Piacenza). The data decomposition techniques confirmed the relevance of the multi-year signal, which was identified as the most significant component by both PCA and ICA (Figure 5.8a and 5.7a). The associated spatial distribution maps show that the highest amplitude of this signal is located in the same areas previously identified through clustering. The multi-year deformation in these areas is strongly correlated in time with cumulative detrended precipitation ($\rho > 0.8$) (Figure 5.9).

In addition to positively correlated signals, both clustering and decomposition techniques identified deformation patterns that are negatively correlated with precipitation, although with smaller amplitude. These are mainly located in the central part of the plain and in the area west of Lake Garda (Figure 5.8a, 5.7a, 5.4).

The magnitude of the observed deformation is relatively low (below 1 cm). At this aim we rejected the pixels associated with highly variable time series (Section 3.1.3). The evident spatio-temporal correlation with independent data (e.g., precipitation and aquifers distribution) provide strong indication that the highlighted deformation patterns are mostly related with hydrological or antropogenic processes, particularly in areas affacted by the highest deformation (i.e, Brescia and Crema).

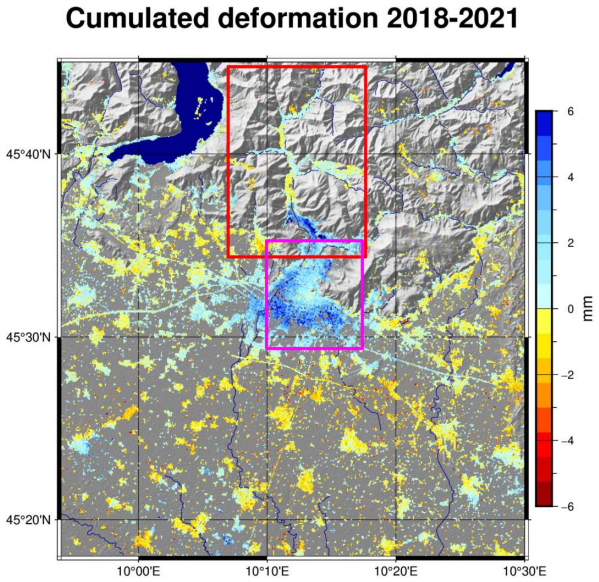
The presence of negatively correlated areas is consistent with what is observed in GNSS measurements and could be attributed to elastic loading effects. In contrast, in the Brescia area — where the positively correlated deformation signal is most prominent — geological and hydrogeological data support the hypothesis that poroelastic deformation is the dominant process. The subsurface in this area consists, in fact, of a sequence of permeable sediments that host in the northern part an unconfined aquifer with thickness up to over 100 m. Water level data in this aquifer, measured by piezometers, show a multi-year evolution that clearly matches the deformation and precipitation multi-year variations (Figure 7.4).

The spatial distribution of deformation amplitude in this area closely resembles geological estimates of aquifer thickness (Figure 2.7 and reproduced in 7.3 for comparison



(a) Saturated aquifer thicnkess b obtained from geological surveys

(b) Basal surface of the aquifer layer in the Trompia Valley



(c) Values of cumulated deformation occurred in the first period

Figure 7.3: Comparison of cumulated deformation and aquifer thicnkess

with deformation). This similarity is evident both in the plain and at the outlet of the Trompia Valley, where the highest deformation values spatially coincide with areas of deepening of the substrate. This observation is consistent with the model proposed by Larochelle, Chanard, et al. 2022, which predicts a direct proportionality between the amplitude of poroelastic deformation and the saturated thickness of the aquifer b.

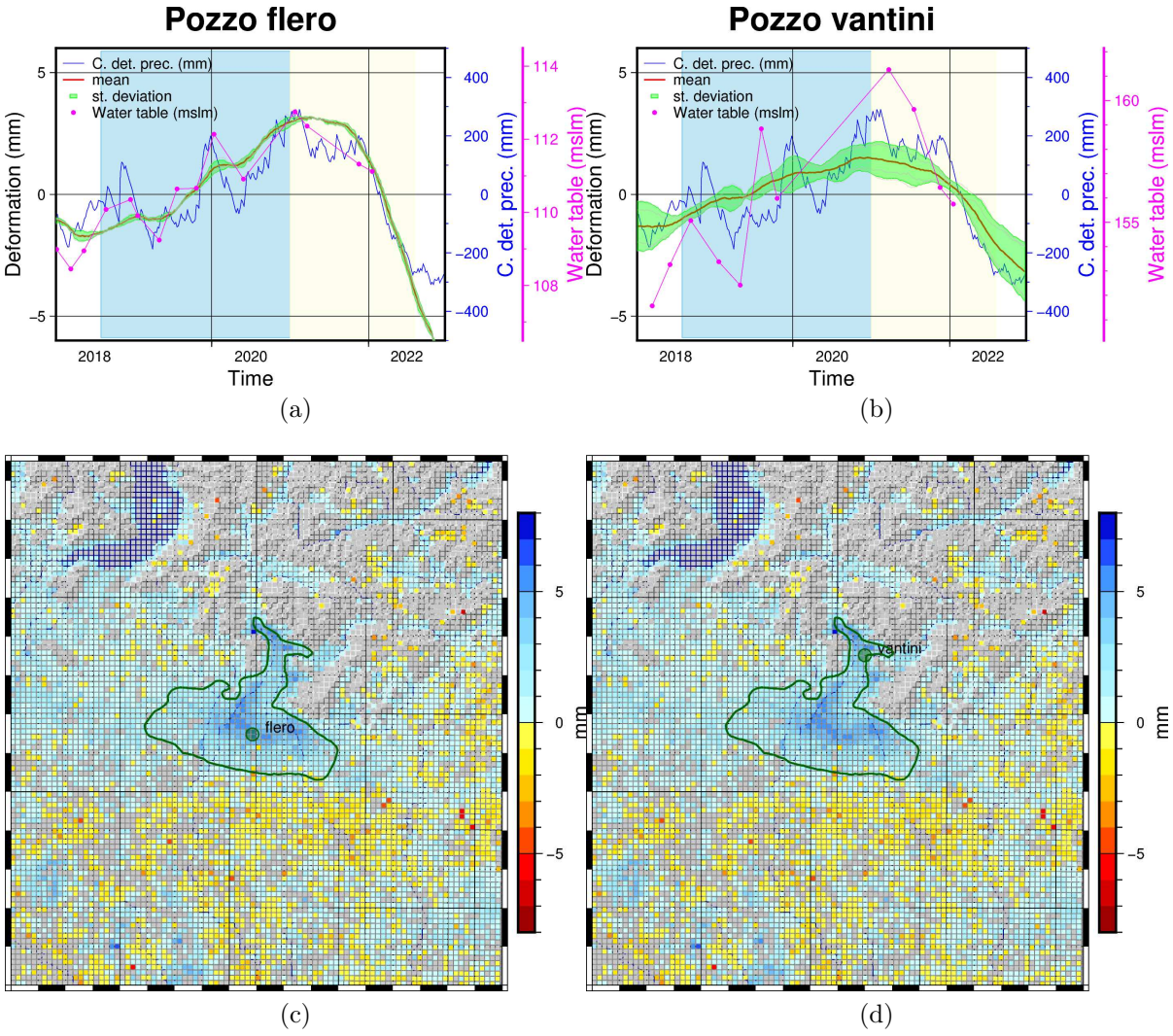


Figure 7.4: Time series and spatial maps of poroelastic deformation at two representative piezometer sites. (a) and (b) show the comparison between vertical deformation (blue) and water table variations (purple) at the Flero and Vantini piezometers, respectively. (c) and (d) spatial location of piezometers and gridpoints used in estimating deformation.

To isolate the poroelastic contribution in this area, I computed the expected elastic response induced by surface water storage variations (Figures 6.5) and removed it from

the observed deformation signal. The modeled amplitude of the elastic contribution (Figure 6.8) is around one millimeter. This value is comparable to the amplitude of the negative/positive values observed in the plain south of of the area characterized by poroelastic signal in respectively the rainy/drought period. With correction of this effect, in both periods the area characterized by signal anticorrelated with precipitation is diminished (Figure 6.9), suggesting that the small-amplitude signals in this region are mainly associated with elastic deformation.

Employing equation 6.66, I performed six point-based estimates of the aquifer's Young modulus, under the assumption that the residual deformation, after the removal of the elastic contribution, is entirely due to poroelastic eigenstrain. The estimation was based on assumed values of the material parameters appearing in equation 7.3—namely, porosity and the Biot–Willis coefficient. These parameters were assumed to be spatially uniform and representative of the sediment types present in the aquifer.

I found values of E_{aq} ranging from 200 MPa and 600 MPa. From the data it appears there is a decrease of the rigidity in the direction of the fan deposition. Quantitative information found in literature for gravel and sand material reports values for the Young Modulus ranging from 30 MPa to 300 MPa depending on the sediment looseness [Kezdi 1974]. This approximately agrees with what was found in the southernmost locations, while higher values in the northern part could be associated with higher sediment cementification which is supported by geological observations [Provinca di Brescia 2016].

The models presented in this study have to be intended as first-order and present inevitable simplifications which might have affected the inferred quantities (e.g. E_{aq}) From Figures 7.5, 7.6 and 7.7, it can be observed that the estimated values of E_{aq} are consistent across the four southernmost piezometers, for water table variations in the range of 1.5 m to 4.5 m and poroelastic deformations between approximately 2.5 mm to 5 mm. In contrast, the two northernmost piezometers exhibit lower deformation values associated with higher water table variation

This discrepancy may be attributed to errors in the water level values measured in these two locations. In particular, the Vantini piezometer may be intercepting a perched aquifer [ARPALombardia 2018a], hydraulically disconnected from the main saturated system, thereby misrepresenting the effective load acting on the deforming medium. Additionally, a replacement of the Vantini well in the month of April 2021 [ARPALombardia 2022] may have introduced inconsistencies in the measurement continuity. Both Vantini and Caffaro are also located in areas affected by intense anthropogenic groundwater extraction, which may induce localized perturbations of the hydraulic head, further complicating the interpretation of the poroelastic response [Pili et al. 2017].

Another possible source of error may arise from the assumed values of the material parameters ϕ (porosity) and α (Biot-Willis coefficient). These parameters are intrinsic properties of the subsurface layers. In particular, the Biot-Willis coefficient is often linked to the degree of confinement of the aquifer. In this analysis, both ϕ and α were considered uniform across the study area. However, the geological characteristics of the

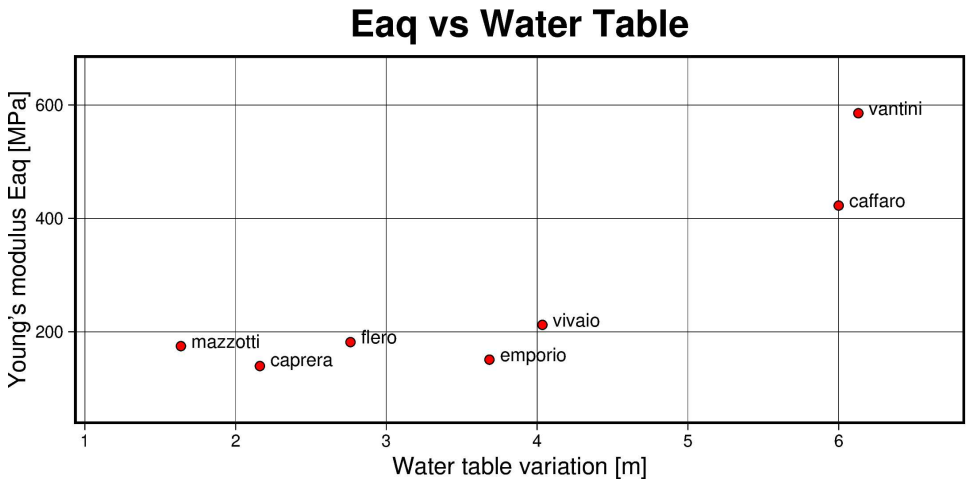


Figure 7.5: Estimated Young's modulus E as a function of water table variation Δh.

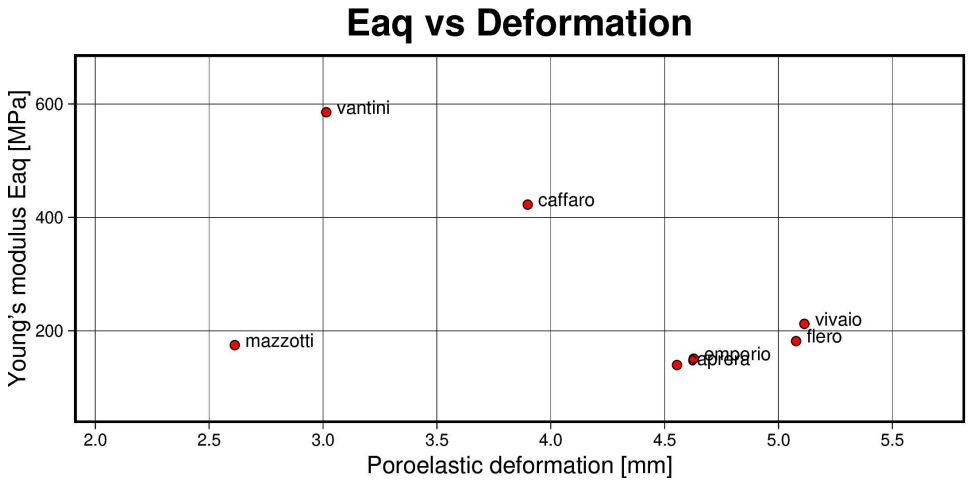


Figure 7.6: Estimated Young's modulus E as a function of vertical poroelastic deformation Δu.

region are known to vary gradually, with transitional features with a decrease in porosity towards the plain and increase in aquifer confinement. As a consequence, the assumed uniform values may lead to biased estimates of the aquifer stiffness.

It is also important to consider potential errors arising from the physical model employed. These may stem from unmodeled contributions to the observed deformation, including both non-hydrological multi-year effects, such as thermoelastic processes, and unresolved hydrological sources. In particular, deformation due to small-scale variations in hydrological loading may not be adequately captured. The elastic deformation component that was removed from the signal was derived from equivalent water height data representing spatial averages over a relatively large area, which may miss localized load

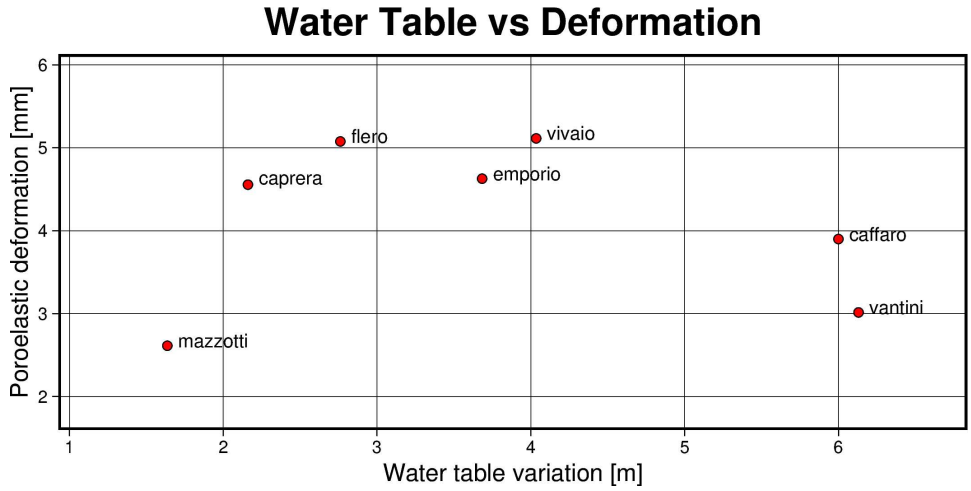


Figure 7.7: Water table variation Δh as a function of vertical poroelastic deformation Δu.

changes that could still produce significant elastic deformation at smaller scales.

To quantify this effect, the influence of small-scale load variations associated with groundwater storage (GWS) changes within the aquifer was explicitly modeled. The maximum predicted elastic deformation was found to be on the order of 1 mm (Figure 6.12). The removal of this additional local elastic response resulted in slightly lower estimates of E_{aq} across all sites, without significantly altering the spatial distribution (Figure 6.10). These results suggest that, if accurately modeled, the inclusion of this component could help reduce biases in the estimation of E_{aq}. However, the adopted modeling approach relies on a simplified reconstruction of the water table signal, which may introduce additional systematic errors due to extrapolation and limited data resolution.

Chapter 8

Conclusions

In this work, I investigated surface deformation in a sector of the Po Plain between 2018 and 2022, using vertical InSAR data from the EGMS Ortho L3 dataset and hydroclimatic observations from regional monitoring networks. The study aimed to investigate the role of hydrological processes in driving surface displacements at seasonal to multiyear scales.

To analyze the large volume of data, I employed several statistical techniques, including STL decomposition, principal component analysis (PCA), variational Bayesian independent component analysis (vbICA), and K-means clustering. These methods proved effective in identifying distinct deformation sources related to natural and anthropogenic processes. In particular, I found a class of deformation signals characterized by multiyear vertical displacements strongly correlated with precipitation trends. These signals are often localized near the Alpine-Po Plain transition zone and are likely associated with poroelastic deformation of permeable sediment layer.

To investigate this mechanism quantitatively, I focused on the area showing the strongest poroelastic signal, located near the city of Brescia. In this region, I modeled the elastic deformation induced by the regional TWS variations, using equivalent water heigth values estimated from GNSS [Pintori et al. 2024]. Removing this deformation component led to a clearer identification of the poroelastic contribution. I applied a poroelastic model, using information on the local geology, to compare the residual deformation to available water table data and estimated the sediment Young's modulus at six locations. The estimated values ranged from 200 MPa to 600 MPa. These estimates were compared with literature values and found to be consistent in magnitude with values relative to gravel and sand deposit, though potentially biased by assumptions in the model and input parameters have been made.

Overall, this study demonstrates the potential of integrating InSAR data, hydroclimatic records, and surface deformation modeling to detect and interpret surface deformations linked to water cycle dynamics and infer aquifer features at unprecedented spatio-temporal resolution. Future work should focus on improving the resolution and

continuity of hydrological input data, and accounting for lithological heterogeneity.

Bibliography

ARPA Lombardia (2020a). *Precipitation from 2011 to 2020*. Dataset published on October 1, 2020. License: CC BY 4.0. Regione Lombardia. URL: https://www.dati.lombardia.it/d/2kar-pnuk.

— (2023). *Precipitation from 2021 onward*. Dataset published on April 4, 2023. License: CC BY 4.0. Updated daily. Regione Lombardia. URL: https://www.dati.lombardia.it/d/pstb-pga6.

— (June 2018a). *Modello idrogeologico di flusso e trasporto della contaminazione*. Tech. rep. Brescia, Italia: Agenzia Regionale per la Protezione dell'Ambiente della Lombardia (ARPA). URL: http://www.arpalombardia.it.

— (Aug. 2018b). *Monitoraggio delle acque sotterranee - Risultati delle indagini geochimiche e piezometriche - Gennaio - Giugno 2018*. Tech. rep. Brescia, Italia: Agenzia Regionale per la Protezione dell'Ambiente della Lombardia (ARPA). URL: http://www.arpalombardia.it.

— (May 2019). *Monitoraggio delle acque sotterranee - Risultati delle indagini geochimiche e piezometriche - Luglio - Dicembre 2018*. Tech. rep. Brescia, Italia: Agenzia Regionale per la Protezione dell'Ambiente della Lombardia (ARPA). URL: http://www.arpalombardia.it.

— (Jan. 2020b). *Monitoraggio delle acque sotterranee - Risultati delle indagini geochimiche e piezometriche - Giugno - Luglio 2019*. Tech. rep. Brescia, Italia: Agenzia Regionale per la Protezione dell'Ambiente della Lombardia (ARPA). URL: http://www.arpalombardia.it.

— (Aug. 2021). *Monitoraggio delle acque sotterranee - Risultati delle indagini geochimiche e piezometriche - Settembre 2020 - Febbraio 2021*. Tech. rep. Brescia, Italia: Agenzia Regionale per la Protezione dell'Ambiente della Lombardia (ARPA). URL: http://www.arpalombardia.it.

— (Mar. 2022). *Monitoraggio delle acque sotterranee - Risultati delle indagini geochimiche e piezometriche - Settembre 2021 - Marzo 2022*. Tech. rep. Brescia, Italia: Agenzia Regionale per la Protezione dell'Ambiente della Lombardia (ARPA). URL: http://www.arpalombardia.it.

Australia, Geoscience (2017). *Interferometric Synthetic Aperture Radar*. URL: https://www.ga.gov.au/scientific-topics/positioning-navigation/

positioning-australia/geodesy/geodetic-techniques/interferometric-synthetic-aperture-radar (visited on 06/27/2025).

Baronetti, Alice et al. (Aug. 2020). "A weekly spatio-temporal distribution of drought events over the Po Plain (North Italy) in the last five decades". In: *International Journal of Climatology* 40.10, pp. 4463–4476. ISSN: 0899-8418, 1097-0088. DOI: 10.1002/joc.6467. (Visited on 01/02/2025).

Borsa, Adrian Antal et al. (2014). "Ongoing drought-induced uplift in the western United States". In: *Science* 345.6204, pp. 1587–1590. DOI: 10.1126/science.1260279.

Burgmann, Roland et al. (2024). "Climate- and weather-driven solid Earth deformation and seismicity". In: *GNSS Monitoring of the Terrestrial Environment*. Ed. by Yosuke Aoki et al. Elsevier, pp. 257–285. DOI: 10.1016/B978-0-323-95507-2.00011-6.

Carlson, G. et al. (Sept. 1, 2024). "A novel hybrid GNSS, GRACE, and InSAR joint inversion approach to constrain water loss during a record-setting drought in California". In: *Remote Sensing of Environment* 311, p. 114303. ISSN: 0034-4257. DOI: 10.1016/j.rse.2024.114303. (Visited on 12/17/2024).

Carminati, E et al. (Nov. 1, 2002). "Subsidence rates in the Po Plain, northern Italy: the relative impact of natural and anthropogenic causation". In: *Engineering Geology* 66.3, pp. 241–255. ISSN: 0013-7952. DOI: 10.1016/S0013-7952(02)00031-5. (Visited on 01/13/2025).

Choudrey, R. A. (2002). *Variational Methods for Bayesian Independent Component Analysis*. URL: https://www.robots.ox.ac.uk/~parg/_projects/ica/riz/Thesis/index.html (visited on 05/21/2025).

Cleveland, Robert B. et al. (Mar. 1990). "STL: A Seasonal-Trend Decomposition Procedure Based on Loess". In: Publisher: Statistics Sweden (SCB). ISSN: 0282423X. (Visited on 05/10/2025).

Crosetto, M. et al. (June 28, 2021). "DEFORMATION MONITORING AT EUROPEAN SCALE: THE COPERNICUS GROUND MOTION SERVICE". In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLIII-B3-2021, pp. 141–146. DOI: 10.5194/isprs-archives-XLIII-B3-2021-141-2021. (Visited on 05/17/2025).

Dati quantitativi | Brescia | 2023 - ARPA Lombardia (2025). URL: https://www.arpalombardia.it/dati/2023/acqua/dati-quantitativi-brescia-2023/ (visited on 05/19/2025).

De Donatis, Mauro et al. (Jan. 2007). "A 3D model of the aquifers of the Lombardian Po Plain". In: *Memorie descrittive della Carta Geologica d'Italia* 76, pp. 193–206.

De Luca, Domenico Antonio et al. (Dec. 9, 2020). "Hydrogeology of the western Po plain (Piedmont, NW Italy)". In: *Journal of Maps* 16.2, pp. 265–273. ISSN: 1744-5647. DOI: 10.1080/17445647.2020.1738280. (Visited on 01/09/2025).

European Environment Agency (2023). *EGMS Ortho Vertical Component (2018-2022)*. https://doi.org/10.2909/943e9cbb-f8ef-4378-966c-63eb761016a9.

Copernicus Land Monitoring Service, European Ground Motion Service (EGMS), v01 r00, release date: 25 October 2023. DOI: 10.2909/943e9cbb-f8ef-4378-966c-63eb761016a9. (Visited on 07/09/2025).

Farías, Celina Anael et al. (Jan. 2024). "Statistical and Independent Component Analysis of Sentinel-1 InSAR Time Series to Assess Land Subsidence Trends". In: *Remote Sensing* 16.21, p. 4066. ISSN: 2072-4292. DOI: 10.3390/rs16214066. (Visited on 01/13/2025).

Farolfi, Gregorio et al. (2019). "Integration of GNSS and Satellite InSAR Data: Derivation of Fine-Scale Vertical Surface Motion Maps of Po Plain, Northern Apennines, and Southern Alps, Italy". In: *IEEE Transactions on Geoscience and Remote Sensing* 57.1, pp. 319–328. DOI: 10.1109/TGRS.2018.2854371.

Farrell, W. E. (Aug. 1972). "Deformation of the Earth by surface loads". In: *Reviews of Geophysics* 10.3, pp. 761–797. DOI: 10.1029/RG010i003p00761. (Visited on 05/29/2025).

Ferretti, A., E. Passera, et al. (Oct. 2023). *End-to-end Implementation and Operation of the European Ground Motion Service (EGMS) - Algorithm Theoretical Basis Document*. Technical Report EGMS-D3-ALG-SC1-2.0-006. European Environment Agency (EEA). URL: https://land.copernicus.eu/en/technical-library/egms-algorithm-theoretical-basis-document/@@download/file.

Ferretti, A., C. Prati, et al. (Jan. 2001). "Permanent scatterers in SAR interferometry". In: *IEEE Transactions on Geoscience and Remote Sensing* 39.1, pp. 8–20. ISSN: 1558-0644. DOI: 10.1109/36.898661. (Visited on 05/25/2025).

Fontana, Alessandro et al. (Mar. 2014). "Alluvial fans and megafans along the southern side of the Alps". In: *Sedimentary Geology* 301, pp. 150–171. ISSN: 00370738. DOI: 10.1016/j.sedgeo.2013.09.003. (Visited on 03/12/2025).

Gualandi, A et al. (2016). "Blind source separation problem in GPS time series". In: *Journal of Geodesy* 90.4, pp. 323–341. DOI: 10.1007/s00190-015-0875-4.

Gualandi, Adriano et al. (2021). "Variational Bayesian Independent Component Analysis for InSAR Displacement Time-Series With Application to Central California, USA". In: *Journal of Geophysical Research: Solid Earth* 126.4, e2020JB020845. DOI: https://doi.org/10.1029/2020JB020845.

Hyvärinen, A. et al. (2000). "Independent component analysis: algorithms and applications". In: *Neural Networks* 13.4, pp. 411–430. ISSN: 0893-6080. DOI: https://doi.org/10.1016/S0893-6080(00)00026-5.

Jaramillo, Fernando et al. (2024). "The Potential of Hydrogeodesy to Address Water-Related and Sustainability Challenges". In: *Water Resources Research* 60.11. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023WR037020, e2023WR037020. ISSN: 1944-7973. DOI: 10.1029/2023WR037020. (Visited on 02/14/2025).

Kang, S. et al. (2023). "Isolating the Poroelastic Response of the Groundwater System in InSAR Data From the Central Valley of California". In: *Geophysical Research Letters*

50.9. e2023GL103222 2023GL103222, e2023GL103222. DOI: https://doi.org/10.1029/2023GL103222.

Kezdi, Arpad (1974). *Handbook of Soil Mechanics*. Amsterdam: Elsevier.

Kullback, S. et al. (Mar. 1951). "On Information and Sufficiency". In: *The Annals of Mathematical Statistics* 22.1. Publisher: Institute of Mathematical Statistics, pp. 79–86. ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177729694. (Visited on 05/14/2025).

Larochelle, S., K. Chanard, et al. (2022). "Understanding the Geodetic Signature of Large Aquifer Systems: Example of the Ozark Plateaus in Central United States". In: *Journal of Geophysical Research: Solid Earth* 127.3. e2021JB023097 2021JB023097, e2021JB023097. DOI: https://doi.org/10.1029/2021JB023097.

Larochelle, S., A. Gualandi, et al. (2018). "Identification and Extraction of Seasonal Geodetic Signals Due to Surface Load Variations". In: *Journal of Geophysical Research: Solid Earth* 123.12. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018JB016607 pp. 11, 031–11, 047. ISSN: 2169-9356. DOI: 10.1029/2018JB016607. (Visited on 05/14/2025).

Martens, Hilary R. et al. (2019a). "LoadDef: A Python-Based Toolkit to Model Elastic Deformation Caused by Surface Mass Loading on Spherically Symmetric Bodies". In: *Earth and Space Science* 6.2, pp. 311–323. DOI: https://doi.org/10.1029/2018EA000462.

— (Feb. 2019b). "LoadDef: A Python-Based Toolkit to Model Elastic Deformation Caused by Surface Mass Loading on Spherically Symmetric Bodies". In: *Earth and Space Science* 6.2, pp. 311–323. ISSN: 2333-5084, 2333-5084. DOI: 10.1029/2018EA000462. (Visited on 05/29/2025).

Montanari, Alberto et al. (Aug. 11, 2023). "Why the 2022 Po River drought is the worst in the past two centuries". In: *Science Advances* 9.32, eadg8304. ISSN: 2375-2548. DOI: 10.1126/sciadv.adg8304. (Visited on 01/11/2025).

Mura, Toshio (1982). "General theory of eigenstrains". In: *Micromechanics of defects in solids*. Dordrecht: Springer Netherlands, pp. 1–62. ISBN: 978-94-011-9306-1. DOI: 10.1007/978-94-011-9306-1_1.

Nespoli, Massimo et al. (May 1, 2021). "The interaction between displacements and water level changes due to natural and anthropogenic effects in the Po Plain (Italy): The different point of view of GNSS and piezometers". In: *Journal of Hydrology* 596, p. 126112. ISSN: 0022-1694. DOI: 10.1016/j.jhydrol.2021.126112. (Visited on 01/13/2025).

Ormerod, J. T. et al. (2010). "Explaining Variational Approximations". In: *The American Statistician* 64.2, pp. 140–153. DOI: 10.1198/tast.2010.09058. eprint: https://doi.org/10.1198/tast.2010.09058. URL: https://doi.org/10.1198/tast.2010.09058.

Pastore, Maria Luisa et al. (2024). *Stato contaminazione acque sotterranee. Monitoraggio Giugno 2024*. ARPA Lombardia - Dipartimento di Brescia, Presentazione convegno

11 October 2024. URL: http://commissario.bresciacaffaro.it/images/ARPA%20Monitoraggio%20giugno%20-%20presentazione%20convegno%2011-10-24.pdf (visited on 06/27/2025).

Perktold, Josef et al. (Apr. 17, 2024). *statsmodels/statsmodels: Release 0.14.2*. Version v0.14.2. DOI: 10.5281/ZENODO.593847. (Visited on 05/05/2025).

Pili, Nicola et al. (2017). "SIN Brescia - Caffaro: risultati monitoraggio acque sotterranee. Indagine geochimica e piezometrica, Gennaio 2015". In: *Geologia tecnica e ambientale*.

Pintori, F. et al. (Mar. 2024). "Drought-Induced Vertical Displacements and Water Loss in the Po River Basin (Northern Italy) From GNSS Measurements". In: *Earth and Space Science* 11.3, e2023EA003326. ISSN: 2333-5084, 2333-5084. DOI: 10.1029/2023EA003326. (Visited on 02/23/2025).

Precipitazioni dal 2011 al 2020 | Open Data Regione Lombardia (2025). URL: https://www.dati.lombardia.it/Ambiente/Precipitazioni-dal-2011-al-2020/2kar-pnuk/about_data (visited on 05/19/2025).

Precipitazioni dal 2021 | Open Data Regione Lombardia (2025). URL: https://www.dati.lombardia.it/Ambiente/Precipitazioni-dal-2021/pstb-pga6/about_data (visited on 05/19/2025).

Provinca di Brescia ARPALombardia, Gruppo di lavoro Engineering Geology (2015). *Attività di affinamento delle conocenze sulla contaminazione delle acque sotterraneee in cinque aree della provincia di Brescia con definizione dei plumes di contaminanti ed individuazione delle potenziali fonti di contaminazione, Lotto A, Area BS001, Fondovalle della Val Trompia e Valle di Lumezzane - Fase 1, Relazione Idrogeologica di inquadramento*. Tech. rep.

— (2016). *Attività di affinamento delle conocenze sulla contaminazione delle acque sotterraneee in cinque aree della provincia di Brescia con definizione dei plumes di contaminanti ed individuazione delle potenziali fonti di contaminazione, Lotto A, Area BS002, Brescia - Caffaro, Fase 1, Relazione Idrogeologica di inquadramento*. Tech. rep.

— (2024). *Attività di affinamento delle conocenze sulla contaminazione delle acque sotterraneee in cinque aree della provincia di Brescia con definizione dei plumes di contaminanti ed individuazione delle potenziali fonti di contaminazione, Lotto A, Area BS002, Brescia - Caffaro, Fase 3, Esiti delle attività svolte - Relazione illustrativa finale*. Tech. rep.

Regione Lombardia, Eni-Divisione Agip (2002). *Geologia degli Acquiferi Padani della Regione Lombardia*. Tech. rep. A cura di Cipriano Carcano e Andrea Piccin. S.EL.CA. (Firenze). (Visited on 03/17/2025).

Rigamonti, S. (Feb. 20, 2025). "Role of triggering and controlling factors on ground deformation processes by time series analyses and numerical modelling. (Tesi di dottorato, 2025)". In: URL: https://boa.unimib.it/handle/10281/543585 (visited on 03/15/2025).

Segall, Paul (2010). *Earthquake and volcano deformation*. Princeton, N. J: Princeton university press. ISBN: 978-0-691-13302-7.

Slovenian Environment Agency, Central Institution for Meteorology et al. (2022). *Standardised Precipitation-Evapotranspiration Index - ERA5_ QM SPEI-1 (Version 1.0)*. DOI: https://doi.org/10.48784/166e51ee-534a-11ec-9143-02000a08f41d.

Vicente-Serrano, Sergio M et al. (n.d.). "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index". In: (). DOI: 10.1175/2009JCLI2909.1.

Wang, Herbert F. (2001). *Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology*. Princeton: Princeton University Press. ISBN: 9781400885688. DOI: doi:10.1515/9781400885688.

Zhu, Lin et al. (Feb. 3, 2017). "Modeling 3-D permeability distribution in alluvial fans using facies architecture and geophysical acquisitions". In: *Hydrology and Earth System Sciences* 21.2. Publisher: Copernicus GmbH, pp. 721–733. ISSN: 1027-5606. DOI: 10.5194/hess-21-721-2017. (Visited on 06/27/2025).

Riassunto

Negli ultimi decenni, lo sviluppo delle tecniche di geodesia satellitare ha portato alla possibilità di misurare la deformazione della superficie terrestre con precisione subcentimetrica ed elevata risoluzione spazio temporale. Tra i fenomeni che possono essere caratterizzati da misure geodetiche, particolare interesse è associato a quelli legati alla redistribuzione dell'acqua nel ciclo idrogeologico [Burgmann et al. 2024]. Lo studio di questi processi risulta di interesse sia per isolare e rimuovere la componente idrologica dai segnali geodetici, facilitando l'identificazione di segnali legati ad altre forzanti, sia per ricavare in modo indiretto informazioni sulle proprietà idrogeologiche e meccaniche del sottosuolo, difficilmente accessibili con metodi di indagine idrologica convenzionali, come campionamenti in situ [Burgmann et al. 2024].

L'area di studio analizzata in questa tesi si colloca nella Pianura Padana, la più estesa pianura italiana e una delle principali pianure alluvionali europee. Situata tra le Alpi a nord e l'Appennino a sud, essa copre circa 46.000 km^2. Il clima è classificato come subtropicale umido (Köppen Cfa), con estati calde e umide (22–25°C), inverni freddi (1–4°C) e precipitazioni annue comprese tra 700 e 1200 mm [Baronetti et al. 2020]. La regione è ricca sia di acque superficiali, rappresentate da numerosi laghi e fiumi – tra cui il Po, il maggiore fiume italiano – sia di acque sotterranee, distribuite in sedimenti porosi di origine alluvionale. Secondo una classificazione idrogeologica della Regione Lombardia [Regione Lombardia 2002], i corpi idrici sotterranei sono suddivisi in quattro gruppi (A, B, C, D), disposti in profondità decrescente e distinti per origine deposizionale e caratteristiche idrauliche. I gruppi più profondi (C e D) sono costituiti da sedimenti marini o transizionali, con alternanza di livelli sabbiosi permeabili (acquiferi) e strati argillosi impermeabili (aquitardi), mentre i gruppi A e B sono composti da depositi continentali, più permeabili, costituiti da sabbie e ghiaie, con granulometria maggiore, specialmente nella fascia pedemontana.

La disponiblità di risorse idriche ha sostenuto per secoli lo sviluppo agricolo e industriale della regione, che oggi ospita circa un terzo della popolazione italiana e include alcuni dei principali poli industriali del paese. Questo ha portato a un considerevole stress sulle risorse idriche, a cui di recente si è aggiunto l'impatto dei cambiamenti climatici, che ne hanno accentuato la vulnerabilità: a partire dagli anni 2000 si è osservato un aumento nella frequenza e nell'intensità degli episodi siccitosi, con eventi rilevanti negli anni 2003,

2007, 2011 [Baronetti et al. 2020] e, più recentemente, nel 2022 [Montanari et al. 2023].

Le dinamiche idrologiche regionali inoltre influenzano direttamente la deformazione della superficie terrestre. Diversi studi hanno documentato infatti segnali deformativi legati sia a variazioni naturali dell'idrosfera [Pintori et al. 2024], sia all'effetto delle attività di estrazione umana [Carminati et al. 2002; Farías et al. 2024], che sopratutto nella parte sud orientale sono responsabili di una subsidenza del terreno con velocità che superano i $10\,\mathrm{mm y^{-}1}$ [Carminati et al. 2002; Farolfi et al. 2019]. Tali osservazioni evidenziano l'importanza di monitorare e comprendere questi processi deformativi, sia per valutare i rischi legati a fenomeni come la subsidenza e l'impatto sulle infrastrutture, sia per sfruttarli come strumenti indiretti di indagine delle dinamiche idroclimatiche.

Nel presente lavoro si analizzano i dati di deformazione superficiale ottenuti con la tecnica InSAR e distribuiti dall'European Ground Motion Service (EGMS) con l'obiettivo di individuare segnali di deformazione associabili a processi idrologici nel periodo 2018–2022.. Particolare attenzione è dedicata alla zona di Brescia, collocata nel settore nord-occidentale della pianura lombarda, dove è stato rilevato un segnale deformativo a lungo termine significativamente correlato con la variabilità multiannuale delle precipitazioni registrata nello stesso periodo.

La tecnica InSAR (Interferometric Synthetic Aperture Radar) è una tecnica di telerilevamento satellitare che permette di misurare spostamenti del suolo, con precisione sub-centimetrica, a partire dalla differenza di fase tra due immagini radar acquisite su una stessa area in due tempi diversi [Ferretti, Passera, et al. 2023].

Il principio alla base della tecnica si basa sul fatto che la differenza tra la fase del segnale emesso e quella dell'eco ricevuto dipende dalla lunghezza del tragitto percorso dal segnale e può quindi essere usata per misurare la distanza tra il satellite e il bersaglio a terra (pixel) o la sua variazione nel tempo [Ferretti, Passera, et al. 2023].

Per misurare lo spostamento della superficie è necessario disporre di almeno due immagini SAR della stessa area. Le due immagini vengono quindi allineate con precisione sub-pixel in modo che i valori di fase possano essere confrontati. La differenza tra i valori di fase misurati nei due passaggi è legata alla variazione della distanza tra il satellite e il pixel. Se la superficie in corrispondenza del pixel ha subito uno spostamento, allora sarà presente nello sfasamento un contributo legato alla proiezione dello spostamento della superficie nella direzione di vista del satellite (Line Of Sight - LOS).

I dati InSAR utilizzati in questo lavoro sono stati elaborati e distribuiti dal servizio europeo EGMS, un'iniziativa del programma Copernicus [European Environment Agency 2023]. I dati derivano da immagini SAR acquisite dai satelliti Sentinel-1 in banda C con un tempo di rivisitazione minimo di 6 giorni. La procedura con cui, a partire da una serie di immagini SAR si ottengono i prodotti di deformazione rilasciati dall'EGMS, prevede una serie di passaggi, esposti in dettaglio in Ferretti, Passera, et al. (2023). Queste operazioni hanno come scopo quello di:

- Selezionare l'immagine da usare come riferimento (master image) per la creazione dell'interferogramma,
- Allineare le altre immagini sulla griglia dell'immagine di riferimento e realizzare l'interferogramm,
- Modellare e rimuovere, utilizzando un Digital Elevation Model, il contributo dello sfasamento prodotto dalla topografia e da differenze tra i punti di vista del satellite tra i due passaggi,
- Modellare e rimuovere gli effetti di sfasamento prodotti dall'atmosfera,
- Escludere pixel in cui caratteristiche variabili della superficie, come ad esempio la presenza di vegetazione, producono segnali riflessi non consistenti tra i vari passaggi del satellite, ed individuare punti, detti permanent scatterers (PS), in cui dal segnale riflesso è possibile ottenere informazioni sullo spostamento
- Ricostruire, a partire da valori di fase noti modulo 2π, la fase assoluta per ciascun pixel, attraverso un processo detto di phase unwrapping.

I dati di spostamento così ottenuti costituiscono il prodotto **BASIC**. Due prodotti più avanzati sono rilasciati. Nel primo, (**CALIBRATED**), le misure InSAR sono ancorate a dati GNSS dalla rete di stazioni EPND (European Permanent Network Densification) dell' European Reference Frame (EUREF). Nel secondo, (**ORTHO**), spostamenti in direzione LOS ottenuti da misure effettuate in più geometrie di acquisizione, vengono combinati per ottenere misure di spostamento verticale e orizzontale [Ferretti, Passera, et al. 2023].

In questo studio sono stati utilizzati i dati di deformazione verticale contenuti nel prodotto **ORTHO**, rilasciato con il secondo aggiornamento EGMS. Il dataset copre il periodo compreso tra il 6 gennaio 2018 e il 17 dicembre 2022, con un campionamento temporale di 6 giorni. Sono stati scaricati quattro quadranti di $100 \times 100\,\text{km}$, successivamente uniti e ritagliati su un'area compresa tra le longitudini 9.5°E – 11°E e le latitudini 45°N – 46°N.

Per indagare la relazione tra deformazioni del suolo e processi idrologici, i dati di deformazione sono stati confrontati con dati di precipitazione misurati dalla rete di monitoraggio di ARPA Lombardia, disponibili tramite il portale Open Data di Regione Lombardia [*Precipitazioni dal 2011 al 2020 | Open Data Regione Lombardia* 2025; *Precipitazioni dal 2021 | Open Data Regione Lombardia* 2025]. Sono stati selezionati 53 sensori con copertura completa nel periodo 2018–2022. I dati grezzi, con un campionamento fino a 10 minuti, sono stati sottoposti a controllo qualità e aggregati in serie cumulative a 6 giorni, in corrispondenza delle date del dataset EGMS. Su ogni serie è stato rimosso un trend lineare per evidenziare la variabilità interannuale. La media spaziale delle serie mostra una ciclicità stagionale sovrapposta a una chiara tendenza pluriannuale: un

aumento progressivo delle precipitazioni fino al 2021, seguito da un forte calo durante la siccità del 2022.

Nell'area di Brescia, analizzata più nel dettaglio, sono stati acquisiti anche dati di livello piezometrico da sette pozzi, estratti da rapporti idrogeologici pubblicati da ARPA Lombardia tra il 2018 e il 2022 [ARPALombardia 2018b; ARPALombardia 2019; ARPALombardia 2020b; ARPALombardia 2021; ARPALombardia 2022]. Le serie piezometriche mostrano un andamento coerente con quello delle precipitazioni: un progressivo innalzamento della falda fino al 2021, seguito da una fase di calo nel 2022.

Per caratterizzare la struttura del sottosuolo in quest'area, sono state analizzate sezioni stratigrafiche presenti nei rapporti ARPA [Provinca di Brescia 2016]. La stratigrafia mostra un acquifero multistrato composto da due unità permeabili, geologicamente distinte ma idraulicamente connesse. La topografia e la profondità degli strati acquiferi mostrano una tendenza decrescente da nord verso sud, con variazioni più contenute lungo l'asse est-ovest.

Ho applicato diverse tecniche di analisi statistica al dataset EGMS per individuare nell'area i principali segnali di deformazione, e, successivamente, interpretare i processi fisici responsabili sulla base di informazioni indipendenti sulle forzanti climatiche e le caratteristiche geologiche dell'area. I metodi applicati sono i seguenti:

- **Stima del trend lineare**: effettuata tramite regressione ai minimi quadrati, per evidenziare e rimuovere la componente di lungo termine della deformazione.
- **Principal Component Analysis (PCA) e Independent Component Analysis (ICA)**: entrambe queste tecniche decompongono l'intero dataset in un numero di componenti spazio-temporali che soddisfano determinate proprietà statistiche. La PCA ricerca componenti non correlate tra loro, diagonalizzando la matrice di covarianza del dataset. L'ICA, invece, ricerca componenti statisticamente indipendenti. Un limite della PCA per le applicazioni geofisiche è l'assunzione che i dati seguando delle distribuzioni di probabilità gaussiane, ipotesi spesso non soddisfatta nel caso dei segnali geodetici [Gualandi et al. 2016]. L'ICA invece modellizza esplicitamente la densità di probabilità delle sorgenti come combinazione di distribuzioni gaussiane descritte da parametri. Nell'approccio utilizzato all'ICA [Gualandi et al. 2016] la stima di questi parametri è effettuata tramite un modello generativo, che ricerca i valori ottimali dei parametri massimizzano la probabilità a posteriori bayesiana in modo variazionale (Variational Bayesian Independent Component Analysis).
- **STL (Seasonal-Trend decomposition using LOESS)**: metodo non parametrico, applicato a singole serie temporali, basato su smoothing locale, utilizzato per separare variazioni stagionali e tendenze pluriannuali [Cleveland et al. 1990]. L'algoritmo STL applica alle serie temporali una sequenza di operazioni di filtraggio, basate su

regressioni locali (LOESS, Locally Estimated Scatterplot Smoothing), con l'obiettivo di isolare una componente stagionale e una di trend a lungo termine. Il metodo dipende da una serie di parametri che determinano la porzione di variabilità assegnata alla componente stagionale e a quella multiannuale.

- **Clustering K-means**: utilizzato per classificare le serie temporali in base alla similitudine della loro evoluzione. L'algoritmo suddivide le serie in un numero predefinito di gruppi, minimizzando la varianza interna ad ogni gruppo. Il clustering è stato applicato sia alle serie originali che ai risultati della decomposizione STL (componente stagionale e multiannuale), con l'obiettivo di identificare similitudini nel comportamento delle serie temporali a diverse scale temporali.

- **Cross-correlazione**: è stata condotta un'analisi di cross-correlazione tra le serie temporali di deformazione e la serie di precipitazione cumulata (detrendata). Per ogni punto di misura è stato determinato il ritardo temporale che massimizza il valore assoluto della correlazione con la precipitazione, al fine di quantificare l'accordo tra deformazione osservata e forzanti climatiche.

L'applicazione delle tecniche sopra descritte alle serie di deformazione detrendate ha permesso di evidenziare diversi processi deformativi che si verificano nell'area di studio.

La decomposizione STL ha mostrato che le componenti multiannuali sono predominanti in diverse aree montane e pedemontane, dove raggiungono ampiezze superiori a 10 mm. La componente stagionale, invece, risulta generalmente debole, ad eccezione di due hotspot locali nell'area di Crema, corrispondenti con la posizione di due siti di stoccaggio di gas naturale.

I risultati del clustering confermano, nelle stesse aree montane e pedemontane, la presenza di serie temporali con comportamento coerente, caratterizzato da un sollevamento nella prima parte del periodo e una subsidenza nella seconda, ovvero con un comportamento positivamente correlato alle precipitazioni. Nelle aree di pianura è stato invece osservato un comportamento opposto, con ampiezze minori (≤ 2 mm). Questa distinzione tra aree è visibile sia nei risultati del clustering applicato alle serie originali, sia a quelli effettuati sulle componenti multiannuali. La componente stagionale ha prodotto invece una classificazione meno informativa, con distribuzioni spaziali frammentate.

Le decomposizioni tramite PCA e ICA hanno confermato il ruolo dominante del segnale multiannuale e permesso di identificare ulteriori segnali di deformazione. Entrambe le tecniche hanno individuato una componente di deformazione associata alla variabilità delle precipitazioni, con distribuzione spaziale coerente ai risultati precedenti: segnale correlato positivamente alla precipitazione nelle aree montane e pedemontane con ampiezze fino a 10 mm, e comportamento opposto nelle zone pianeggianti, con ampiezza minore (meno di 2 mm).

In aggiunta a questo segnale, la decomposizione ICA ha isolato un contributo stagionale in corrispondenza dei siti di stoccaggio di gas naturale. Il segnale, con ampiezze fino a 6 mm, presenta un andamento stagionale evidente che riflette i cicli di iniezione e prelievo. Si osserva che, mentre l'ICA ha permesso una separazione netta di questo contributo dalla variabilità multiannuale, nel caso della PCA esso risulta mescolato ad altre componenti.

Infine, l'analisi di cross-correlazione ha verificato un'elevata coerenza spaziale tra deformazione e precipitazione nelle aree identificate nella fascia pedemontana, con coefficienti di correlazione positivi superiori a 0.9, mentre, prevalentemente in zone pianeggianti, sono state individuate anche serie anticorrelate, con valori assoluti di correlazione più bassi.

La presenza di aree caratterizzate da segnali di deformazione sia positivamente che negativamente correlati con l'andamento delle precipitazioni, evidenziata dalle analisi condotte, suggerisce la coesistenza di diversi meccanismi deformativi indotti dalla variabilità idrologica.

In particolare, le aree che mostrano una chiara correlazione positiva, situate prevalentemente lungo il margine pedemontano, sono compatibili con una risposta poroelastica del sottosuolo. In questi contesti, acquiferi poco confinati e con sedimenti grossolani possono subire un'espansione in seguito all'aumento della pressione di poro durante i periodi piovosi, e quindi produrre un sollevamento superficiale [Larochelle, Chanard, et al. 2022]. Tale interpretazione è coerente con la distribuzione spaziale dei segnali osservati e con le caratteristiche idrogeologiche note.

Al contrario, nelle aree di pianura, la correlazione negativa tra precipitazioni e deformazione è compatibile con una risposta elastica della litosfera a variazioni del carico idrologico. Questo tipo di comportamento è in accordo con quanto osservato da Pintori et al. (2024), che hanno documentato un sollevamento regionale della superficie terrestre durante la siccità del 2022, associato alla diminuzione delle risorse idriche avvenuto in questo periodo.

Queste osservazioni nel loro insieme suggeriscono che la deformazione osservata possa essere il risultato della sovrapposizione di due effetti distinti: una risposta elastica regionale al carico idrologico, dominante nelle aree di pianura, e una risposta poroelastica lprevalente a scale locali in corrispondenza di depositi permeabili. Mentre il primo meccanismo produce sollevamento nei periodi di siccità e subsidenza in quelli piovosi, il secondo induce un comportamento opposto, con sollevamento durante i periodi umidi e subsidenza nei periodi secchi.

I due meccanismi sono stati studiati applicando dei modelli fisici di risposta al carico [Martens et al. 2019a] e di deformazione poroelastica [Larochelle, Chanard, et al. 2022].

La deformazione elastica della Terra in risposta ad un carico superficiale è calcolabile numericamente per una Terra stratificata, radialmente simmetrica, non rotante e autogravitante, a partire dalle equazioni del moto perturbate rispetto ad uno stato di riferimento di equilibrio idrostatico [Martens et al. 2019a].

$$\begin{aligned} \rho_0 \frac{\partial^2 \vec{u}}{\partial t^2} &= \nabla \cdot \hat{\sigma}_1 + \rho_0 \nabla \psi_1 + \nabla(\rho_0 \, \vec{u} \cdot \vec{g}_0) - \vec{g}_0 \, \nabla \cdot (\rho_0 \, \vec{u}) \\ \nabla^2 \psi_1 &= -4\pi G \, \nabla \cdot (\rho_0 \, \vec{u}) \end{aligned} \tag{8.1}$$

dove $\vec{u}$ è il campo di spostamento, $\hat{\sigma}_1$ è la perturbazione del tensore degli sforzi, ρ_0 è la densità allo stato di riferimento, ψ_1 è il potenziale gravitazionale perturbato, $\vec{g}_0$ è il campo gravitazionale imperturbato, G è la costante gravitazionale universale.

In virtù della simmetria sferica del problema, le variabili possono essere separate in una parte radiale e una parte angolare, espressa in termini di armoniche sferiche. La dipendenza temporale è assunta essere sinusoidale di frequenza ω ($\omega = 0$ per carichi indipendenti dal tempo).

$$\begin{aligned} u_r(r, \theta, t) &= \sum_n U_n(r) \, P_n(\cos\theta) \, e^{i\omega t} \\ u_\theta(r, \theta, t) &= \sum_n V_n(r) \, \frac{\partial P_n(\cos\theta)}{\partial \theta} \, e^{i\omega t} \\ \psi(r, \theta, t) &= \sum_n \Psi_n(r) \, P_n(\cos\theta) \, e^{i\omega t} \end{aligned} \tag{8.2}$$

dove: $P_n(\cos\theta)$ sono i polinomi di Legendre, U_n, V_n, Ψ_n sono i coefficienti radiali della decomposizione armonica.

Inserendo nelle equazioni del moto la decomposizione 8.2 e la relazione costitutiva per la reologia elastica data dalla legge di Hooke:

$$\sigma_{ij} = 2\mu\epsilon_{ij} + \lambda\epsilon_{kk}\delta_{ij} \tag{8.3}$$

si ottiene un sistema di equazioni differenziali ordinarie del secondo ordine nelle variabili $U_n(r)$, $V_n(r)$ e $\Psi_n(r)$.

Il sistema è ulteriormente riducibile ad un sistema di primo ordine nella forma [Martens et al. 2019a]:

$$\frac{d\vec{y}}{dr} = \hat{A}(r) \, \vec{y} \tag{8.4}$$

definendo il vettore:

$$\vec{y}(r) = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} U_n \\ T^n_{rr} \\ V_n \\ T^n_{r\theta} \\ \Psi_n \\ Q_n \end{pmatrix} \tag{8.5}$$

dove T^n_{rr} e $T^n_{r\theta}$ sono i coefficienti armonici delle componenti dello sforzo radiale e tangenziale rispettivamente, Ψ_n è il potenziale gravitazionale perturbato e la variabilie Q_n è legata alle derivate radiali del potenziale:

$$Q_n = \frac{d\Psi_n}{dr} - 4\pi G\rho_0 U_n \tag{8.6}$$

Il sistema viene integrato numericamente propagando tre soluzioni linearmente indipendenti per il vettore $\vec{y}$ dall'interno della Terra fino alla superficie. Alla superficie le tre soluzioni vengono poi combinate linearmente per soddisfare le condizioni al contorno associate alla presenza del carico superficiale [Martens et al. 2019a].

Per una massa puntiforme unitaria, le funzioni che descrivono la risposta sono dette Green Functions. I coefficienti della decomposizione armonica sono detti numeri di Love [Farrell 1972]:

$$\begin{aligned}
\Gamma_u(r,\theta) &= \frac{a}{m_e}\sum_{n=1}^{\infty} h'_n(r)\, P_n(\cos\theta) \\
\Gamma_v(r,\theta) &= \frac{a}{m_e}\sum_{n=1}^{\infty} l'_n(r)\, \frac{\partial P_n(\cos\theta)}{\partial\theta} \\
\Gamma_\psi(r,\theta) &= \frac{ag_0}{m_e}\sum_{n=1}^{\infty}(1+k'_n(r))\, P_n(\cos\theta)
\end{aligned} \tag{8.7}$$

Dove:

$$\Gamma_{u_r}(r,\theta,\phi;r',\theta',\phi') \qquad \Gamma_{u_\theta}(r,\theta,\phi;r',\theta',\phi') \qquad \Gamma_\psi(r,\theta,\phi;r',\theta',\phi') \tag{8.8}$$

sono le funzioni di Green. Le variabili primate sono relative alla massa che genera la perturbazione, le variabili non primate sono relative al punto in cui la risposta viene calcolata.

La conoscenza dei numeri di Love, e quindi delle funzioni di Green (equazione 8.7) permette il calcolo della risposta ad un carico di estensione generica $\gamma(\theta',\phi')$ attraverso una convoluzione:

$$\begin{aligned}
u_r(r,\theta,\phi) &= \int_0^{2\pi}\int_0^{\pi} \Gamma_{u_r}(r,\theta,\phi;r'=a,\theta',\phi')\gamma(\theta',\phi')a^2\sin(\theta')d\theta'd\phi' \\
u_\theta(r,\theta,\phi) &= \int_0^{2\pi}\int_0^{\pi} \Gamma_{u_\theta}(r,\theta,\phi;r'=a,\theta',\phi')\gamma(\theta',\phi')a^2\sin(\theta')d\theta'd\phi' \\
\psi(r,\theta,\phi) &= \int_0^{2\pi}\int_0^{\pi} \Gamma_{\psi}(r,\theta,\phi;r'=a,\theta',\phi')\gamma(\theta',\phi')a^2\sin(\theta')d\theta'd\phi'
\end{aligned} \tag{8.9}$$

In questo lavoro la deformazione in risposta ad un carico elastico è stata calcolata con il software LoadDef. Questo esegue tre passaggi:

1. Integrazione numerica delle equazioni del moto 8.1 con condizioni al contorno date da una massa unitaria puntiforme. Da questo passaggio vengono calcolati i numeri di Love per ogni grado armonico.

2. Calcolo delle Green Function attraverso la somma numerica della serie 8.7.

3. Calcolo della convoluzione su una griglia numerica, definita attorno a ciascun punto di osservazione mediante cerchi concentrici. La funzione di Green e il modello di carico fornito in input, definito su griglia numerica, sono interpolati sulla griglia definita dal codice e la convoluzione è calcolata come somma aritmetica.

La deformazione poroelastica è stata descritta tramite il modello di Larochelle, Chanard, et al. (2022), che calcola la deformazione di un disco omogeneo di materiale poroso, saturo d'acqua fino ad un livello b, in conseguenza dell'aumento di Δh del livello della tavola d'acqua. L'effetto dell'aumento del livello dell'acqua è associato ad un *eigenstrain* nel mezzo poroso che, in assenza di vincoli meccanici esterni, produrrebbe un'espansione uniforme data da:

$$\epsilon_{\text{eig}} = \frac{\alpha \rho_w g \Delta h (1 - 2\nu)}{E_{\text{aq}}} \tag{8.10}$$

dove, α è il coefficiente di Biot-Wills, ρ_w è la densità dell'acqua, g è l'accelerazione gravitazionale, Δh è la variazione di carico piezometrico (livello della falda), ν ed E_{aq} sono rispettivamente il rapporto di Poisson e il modulo di Young dell materiale poroso. La relazione costitutiva poroelastica in presenza di eigenstrain è (in coordinate cilindriche) [Larochelle, Chanard, et al. 2022]:

$$\begin{aligned}
\varepsilon_{zz} &= \frac{1}{E_{\text{aq}}} \left[(1+\nu)\, \sigma_{zz} - \nu \left(\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz} \right) \right] + \varepsilon_{\text{eig}} \\
\varepsilon_{rr} &= \frac{1}{E_{\text{aq}}} \left[(1+\nu)\, \sigma_{rr} - \nu \left(\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz} \right) \right] + \varepsilon_{\text{eig}} \\
\varepsilon_{\theta\theta} &= \frac{1}{E_{\text{aq}}} \left[(1+\nu)\, \sigma_{\theta\theta} - \nu \left(\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz} \right) \right] + \varepsilon_{\text{eig}}
\end{aligned} \tag{8.11}$$

dove ε_{ij} sono le componenti del tensore delle deformazioni, σ_{ij} sono le componenti del tensore degli sforzi. Larochelle, Chanard, et al. (2022) assume che le deformazioni orizzontali $\varepsilon_{\theta\theta}$ e ε_{rr} siano molto minori di ε_{eig}, trascurandole nelle equazioni 8.11 e arrivando ad un'equazione per u_z che può essere integrata in verticale dalla base dell'acquifero, risultando in:

$$u_z(r) = \begin{cases} \dfrac{(1+\nu)(1-2\nu)}{(1-\nu)} \dfrac{(\alpha - \phi)\rho_w g \Delta h\, b}{E_{\text{aq}}}, & r \leq a \\ 0, & r > a \end{cases} \tag{8.12}$$

dove $u_z(r)$ è lo spostamento verticale alla superficie alla distanza radiale r, a è il raggio della zona in cui avviene l'aumento del livello di falda, ϕ è la porosità del materiale, b è lo spessore dell'acquifero.Tale soluzione è caratterizzata da un comportamento locale, con deformazione solo in corrispondenza dell'aumento della tavola d'acqua e proporzionale allo spessore dell'acquifero.

Per studiare la deformazione osservata sulla base dei processi fisici responsabili, il periodo di studio è stato suddiviso in due intervalli temporali distinti, caratterizzati da condizioni idrologiche opposte: un periodo umido (T_1) e uno siccitoso (T_2). Le serie EGMS sono state ricampionate su una griglia regolare di $500\,\mathrm{m} \times 500\,\mathrm{m}$. In ciascun punto x_p della griglia, la deformazione cumulata osservata $\Delta U(x_p)$ è stata modellata come sovrapposizione di un contributo associato ai processi elastici e uno associato alla risposta elastica:

$$\begin{aligned} \Delta U_{T_1}(x_p) &= \Delta U^e_{T_1}(x_p) + \Delta U^p_{T_1}(x_p) \\ \Delta U_{T_2}(x_p) &= \Delta U^e_{T_2}(x_p) + \Delta U^p_{T_2}(x_p) \end{aligned} \tag{8.13}$$

La risposta elastica è stata calcolata utilizzando LoadDef nei due periodi considerati, a partire dai dati di variazione di Equivalent Water Height (EWH) ottenuti nell'area tramite inversione di dati GNSS [Pintori et al. 2024]. La componente poroelastica è stata quindi ottenuta per differenza:

$$\begin{aligned} \Delta U^p_{T_1}(x_p) &= \Delta U_{T_1}(x_p) - \Delta U^e_{T_1}(x_p) \\ \Delta U^p_{T_2}(x_p) &= \Delta U_{T_2}(x_p) - \Delta U^e_{T_2}(x_p) \end{aligned} \tag{8.14}$$

La deformazione poroelastica è stata confrontata con le variazioni del livello della falda, Δh, osservate in sette piezometri. Il modulo di Young dell'acquifero E_{aq} è stato stimato utilizzando la seguente formula, derivata dal risultato di Larochelle, Chanard, et al. (2022):

$$E_{\mathrm{aq}} = \frac{(1+\nu)(1-2\nu)}{(1-\nu)} \cdot \frac{(\beta - \phi)\,\rho_w\, g\, \Delta h\, b}{\Delta U^p} \tag{8.15}$$

dove:

- Δh è la variazione piezometrica osservata;
- ΔU^p è la deformazione poroelastica calcolata;
- $\phi = 0.25$ è la porosità tipica dei sedimenti nella zona;
- $\beta = 0.8$ è il coefficiente di Biot–Willis assunto;
- $\nu = 0.25$ è il rapporto di Poisson;
- b è lo spessore dell'acquifero stimato da cross-section geologiche;

- ρ_w è la densità dell'acqua, $\rho_w = 1000\,\mathrm{kg\,m^{-3}}$

L'analisi della deformazione cumulata ha mostrato una risposta coerente con processi poroelastici nella zona di Brescia. Durante il periodo piovoso si osserva un sollevamento dell'area corrispondente all'acquifero, con valori medi superiori a 3 mm e picchi oltre 5 mm, mentre durante il periodo siccitoso si osserva una subsidenza della stessa area, con valori medi di circa 5 mm. I massimi valori si concentrano nella zona settentrionale dell'area di studio, in corrispondenza dello sbocco della Val Trompia, mentre verso sud i segnali si attenuano e cambiano segno.

Sottraendo il contributo elastico regionale stimato mediante il modello LoadDef, i segnali rimanenti sono stati interpretati come deformazione poroelastica. Questa appare amplificata rispetto alla deformazione originaria, con valori fino a 7 mm nel periodo piovoso e -10 mm nel periodo siccitoso.

La stima per il modulo di Young dell'acquifero ha ottenuto valori compresi tra 200 MPa e 600 MPa. I valori più elevati si osservano nella porzione settentrionale dell'area, mentre quelli più bassi si localizzano nella pianura a sud. La distribuzione spaziale del modulo di Young è consistente con le variazioni di granulometria e di compattazione dei depositi caratteristiche della struttura dell'acquifero.

Le stime del modulo di Young dell'acquifero sono simili tra i piezometri della porzione meridionale dell'area studiata, mentre sono maggiori in corripondenza dei pozzi più settentrionali. Queste discrepanze possono derivare da errori nella misura del livello piezometrico, come nel caso del piezometro Vantini, che potrebbe intercettare una falda sospesa non rappresentativa della falda principale [ARPALombardia 2018a]. Altri fattori che possono aver introdotto errori sistematici sono la sostituzione del pozzo durante il periodo di osservazione [ARPALombardia 2022], l'influenza di emungimenti locali [Pili et al. 2017], e l'assunzione di parametri geologici uniformi nello spazio (porosità ϕ e coefficiente di Biot-Willis α). Inoltre, il modello fisico adottato include semplificazioni che potenialmente trascurano contributi deformativi rilevanti (es. effetti termoelastici o effetti idrologici non modellati). Come analisi ulteriore è stata calcolata anche la risposta elastica associata alle variazioni di carico idrico a scala locale, mostrando che, sebbene di entità contenuta (massimo 1 mm), essa può influenzare le stime del modulo di Young. L'inclusione di questo contributo ha portato a una lieve riduzione dei valori stimati, suggerendo che una modellazione più accurata delle variazioni idriche locali potrebbe contribuire a rimuovere fonti di errore.

In conclusione questo studio mostra che, integrando dati InSAR e dati meteoclimatici è possibile caratterizzare i processi defomativi associati alla dinamica del ciclo dell'acqua con una elevata risoluzione spaziale e ricavare informazioni sulla struttura e sulle caratteristiche dei corpi idrici sotterranei