Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Corso di Laurea Magistrale in Fisica del Sistema Terra

Independent Component Analysis of GPS time series in the Altotiberina fault region in the Northern Apennines (Italy)

Relatore:

Prof. Maria Elina Belardinelli

Correlatori:

Dott. Enrico Serpelloni
Dott. Adriano Gualandi

Presentata da:

Cristina Nichele

Sessione II
Anno Accademico 2015/2016

Abstract

L'analisi delle componenti indipendenti (Independent Component Analysis, ICA) è una tecnica di statistica multivariata che consente di scomporre un segnale complesso in un certo numero di componenti, tra loro statisticamente indipendenti, che rappresentano le principali sorgenti di quel segnale. La tecnica ICA è stata applicata a serie temporali di spostamento GPS relative a 30 stazioni localizzate nell'Alta Valle del Tevere, nell'Appennino settentrionale. In quest'area, una faglia normale a basso angolo di immersione (circa 15°), faglia Altotiberina (Altotiberina fault, ATF), risulta attiva e responsabile di microsismicità ($M_L < 3$), nonostante la teoria di Anderson sulla fagliazione affermi che non dovrebbe esserci scorrimento su faglie normali di questo tipo. L'interesse per l'ATF è inoltre dovuto al suo potenziale sismogenetico: un evento che dovesse attivare l'intera faglia (lunga circa 70 km) avrebbe infatti magnitudo intorno a 7. Per questo motivo la zona è monitorata da reti multiparametriche (sismiche, geodetiche, geochimiche) che registrano dati in maniera continuativa, rendendo possibile l'individuazione anche di piccoli segnali di deformazione transiente.

Applicando la ICA alle serie temporali GPS si ottiene una scomposizione del segnale in 4 componenti indipendenti. L'analisi di queste componenti ha portato all'individuazione di correlazioni con serie temporali di piovosità, temperatura e sismicità. Una delle componenti, che presenta un segnale transiente di origine tettonica, è stata poi invertita per determinare la distribuzione dello scorrimento su piani di faglia noti. Il momento sismico associato allo scorrimento sulle faglie risulta maggiore del momento sismico associato ai terremoti registrati nell'area, suggerendo quindi che una parte dello scorrimento sia dovuto a movimenti asismici.

Contents

Introduction

The Altotiberina fault (ATF) is a low angle (15°-25° east-dipping) normal fault (LANF) located in the Umbria-Marche sector of the northern Apennines of Italy. Based on the Andersonian [*Anderson*, 1951] theory of faulting, as typically applied to upper continental crust, normal faults should be precluded to slip at low (<30°) dip angles; nevertheless many field-based studies [*Lister and Davis*, 1989, *Axen*, 1999] and the interpretation of seismic reflection profiles [e.g. *Floyd et al*., 2001] indicate that LANFs can be tectonically active and generate earthquakes [*Rigo et al*., 1996, *Abers et al*., 1997, *Axen*, 1999, *Chiaraluce et al*., 2007]. Despite the fact that the ATF is thought to be the main structure accommodating extension in the region, no large earthquakes have been recorded along the ATF plane since 2 Ka [*Rovida* et al., 2016]. However Chiaraluce et al. [2014] pointed out that based on the size of the fault, an event activating the entire ATF should be around M_W 7. Recently, using a dense network of Global Positioning System (GPS) stations, Anderlini et al. [2016] studied the present-day crustal deformation of this area, finding that the ATF plays an important role in accommodating the SW-NE extension in the Northern Apennines being the largest part of the fault surface creeping at a rate of (1.7±0.3 mm/yr). In order to monitor this region, a research infrastructure (TABOO, *The Altotiberina Near Fault Observatory*), recording high-resolution multiparametric data (seismic, geochemical, geodetic), has been built in 2010. The presence of such a dense network makes it possible to detect even weak transient signals such as slow slip events.

In this thesis I analyze surface displacement time series of 30 GPS stations located around the Altotiberina fault, by means of a multivariate statistical technique called Independent Component Analysis (ICA). ICA provides a way to solve Blind Source

Separation (BSS) problems, finding the sources that are superimposed in a mixed signal. When applied to GPS time series, ICA permits to decompose them into a proper number of components, statistically independent one from the others, which represent the major kinds of variation of displacement as a function of time in the investigated region. The aim of this thesis is to find the best decomposition for the dataset and propose a physical explanation for each IC. Further analyses, aimed at a better description of the physical processes at the base of the highlighted deformation signals, are also pursued using the Independent Component Analysis-based Inversion Method (ICAIM) code developed at the Istituto Nazionale di Geofisica e Vulcanologia (INGV) and University of Bologna [*Gualandi*, 2015] as an extension of the PCAIM (Principal Component Analysis-based Inversion Method) code [*Kositsky* and *Avouac*, 2010], developed at the California Institute of Technology (CalTech). The code allows us also to invert one or more IC in order to model the slip distribution on a predetermined fault geometry.

Chapter 1 contains a description of the Altotiberina fault (ATF) and the geological and tectonic framework of the northern Apennines, where the fault is located.

In Chapter 2, after briefly describing how GPS time series are obtained, I show the results of a strain analysis performed in the ATF region to determine the evolution of the dilatation with time.

The basic principles of the Independent Component Analysis (ICA) technique are exposed in Chapter 3, where the ICAIM code is described. I also show the final decomposition for the 30 GPS stations that compose the dataset. Each IC is then described in Chapter 4, where hypothesis on the sources of those signals are proposed based on the observed correlations between the temporal part of 3 ICs (one is thought to mainly reproduce noise) and time series of cumulated rainfall, mean air temperature and cumulative number of earthquakes.

Finally, in Chapter 5 I present the results of an inversion of the fourth IC, performed in order to determine a model for the slip distribution on rectangular fault planes.

1. The Altotiberina fault

The Altotiberina fault (ATF) is a ~60 km long, NNW-trending, east-dipping (15°-25°) low angle normal fault (LANF). It reaches a depth of approximately 12 km, which suggests a downdip extent of 35 km or more [*Finocchio et al.*, 2013]. It is clearly imaged in the CROP03 deep seismic reflection profile [*Barchi et al*., 1998] and located in the Upper Tiber Valley at the Tuscany–Umbria–Marche regions boundary within the northern Apennines, as shown in Figure 1.1.

The existence of this structure is documented in geological [e.g., *Brozzetti* 1995, *Brozzetti et al*., 2000] and geophysical data [e.g., *Mirabella et al*., 2011], as well as rather well imaged by the associated microseismicity, even though the Andersonian [*Anderson*, 1951] theory of faulting, as typically applied to the upper crust, precludes normal faults from slipping at low (<30°) dip angles. GPS data also documented contemporary extension that is likely accommodated by creeping on the ATF [*Hreinsdottir* and *Bennett*, 2009].

The geophysical interest for this area, mainly due to the fact that, based on the size of the fault, medium to large earthquakes could nucleate along the fault plane, moved the scientific community to build a research infrastructure (TABOO, *The Altotiberina Near Fault Observatory*) to monitor with multiparametric instruments the ATF.

1.1 Geological and tectonic framework

The northern Apennines consist of a NE-verging thrust and fold belt, formed as a consequence of the collision between the Adriatic microplate and the European continental margin (Sardinia–Corsica block) [*Collettini and Barchi*, 2002].

Geologically, the region comprises a cover sequence of continental margin sedimentary rocks deposited upon a Paleozoic metamorphic basement [*Chiaraluce et al.*, 2007].

The study area represents a transitional zone, being located just on the border between the inner extensional Tyrrhenian domain and the outer compressional Adriatic

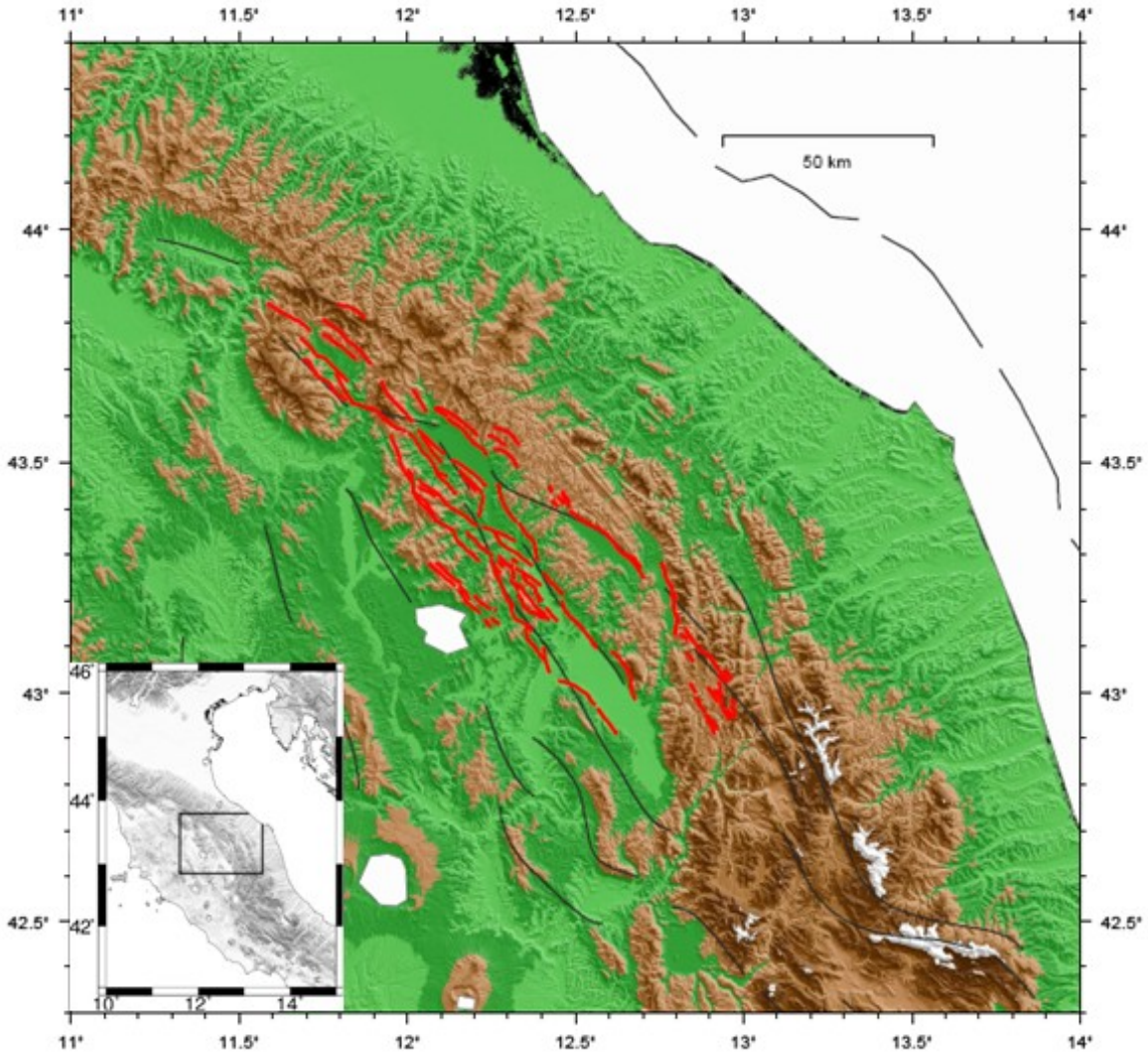


Figure 1.1: Maps of the Northern Apennines region. In red the trace of the Altotiberina fault and of synthetic and antithetic faults obtained from geological survey.

domain [*Collettini and Barchi*, 2002].

The extensional tectonics of the Northern Apennines started in the eastern part of Corsica in the Lower Miocene, then moved eastward, reaching the Umbria region in the Upper Pliocene [*Ambrosetti et al*., 1978; *Jolivet et al*., 1998]: the Altotiberina fault has accumulated 2 km of displacement over the past 2 Ma [*Vadacca et al*., 2014]. This region is still extending at a rate of about 3 mm/yr [*Serpelloni et al*., 2005, *D'Agostino et al.*, 2009]. Recent studies provides more detailed pictures of both present-day strain rates and fault kinematics. Anderlini et al. [2016] showed that ~1 mm/yr of this SW-NE extension is accommodated in the inner (Tyrrhenian) sector of the study area, while ~2 mm/yr are accommodated by activity of the Altotiberina fault and of its antithetic fault system (see Figure 1.2a). The area of maximum extension, concentrated across a 30-40 km wide zone, coincides with the area of the strongest seismic moment release [*D'Agostino et al*., 2009]. As shown in Anderlini et al. [2016] (Figure 1.2b) the strain-rate value in this sector is ~0.67×10^{-7} yr^{-1} of NE-SW oriented extension and it slows down going eastward and westward. The strain-rate calculated in this work (Chapter 2) for the ATF region is consistent with the value estimated in Anderlini et al. [2016].

In the past 20 years, three main seismic sequences have occurred within the hanging-wall of the ATF: the 1984 Gubbio sequence (M_W 5.1), the 1997 Colfiorito multiple-main-shock sequence (M_W 6.0, 5.7 and 5.6), and the 1998 Gualdo Tadino (M_W 5.1) sequence [*Barba and Basili*, 2000, *Boncio and Lavecchia*, 2000, *Boncio et al*., 2000]. Up to now, instead, only micro-earthquakes (M_L<3.0) have been observed to continuously nucleate along the ATF plane [*Chiaraluce et al*., 2007, *De Luca et al*., 2009], mostly in its southern portion [*Piccinini et al*., 2003], at high and constant rate ($r = 7.3 \times 10^{-4}$ earthquakes day^{-1} km^{-2}) [*Chiaraluce et al*., 2009]. The lack of microseismicity in the 20 km long northern section of the ATF suggests the existence of a locked portion along the ATF plane, which may be considered a seismic gap and thus may be hazardous [*Piccinini et al*., 2003]. Anderlini et al. [2016] found that the ATF is mainly locked down to a depth of 4-5 km, while the largest part of the fault surface is creeping at a rate of ~1.7 mm/yr except for a dipper asperity at 7-10 km of depth. The creeping portion correlates with the relocated microseismicity.

The deformation associated with the short and long term slip rate, inferred by geological [*Collettini and Holdsworth*, 2004] and geodetic studies [*D'Agostino et al.*, 2009], can't be fully explained by the microseismicity nucleating along the ATF [*Chiaraluce et al.*, 2014]: Hreinsdóttir and Bennett [2009], investigating regional GPS data, suggest the presence of aseismic deformation and creeping fault behavior.

Based on the size of the fault and considering Wells and Coppersmith [1994] relationship between magnitude and rupture length, the average size of an event activating the entire ATF should be around M_W 7 [*Chiaraluce et al.*, 2014]. However, historical catalogs do not contemplate the occurrence of such an event in the past 2000 years, which could mean that: a) the recurrence time of ATF earthquakes could be larger than 1000 years, or b) the fault is not storing stress, or c) the ATF is sliding aseismically [*Chiaraluce et al.*, 2014].

Another important characteristic of this area is the presence of very high fluids pressure (85% of the lithostatic pressure), mostly due to CO_2, at depths of 4.8 km and 3.7 km, which is considered one of the primary triggering mechanisms of Apennines earthquakes [*Chiodini et al.*, 2004].

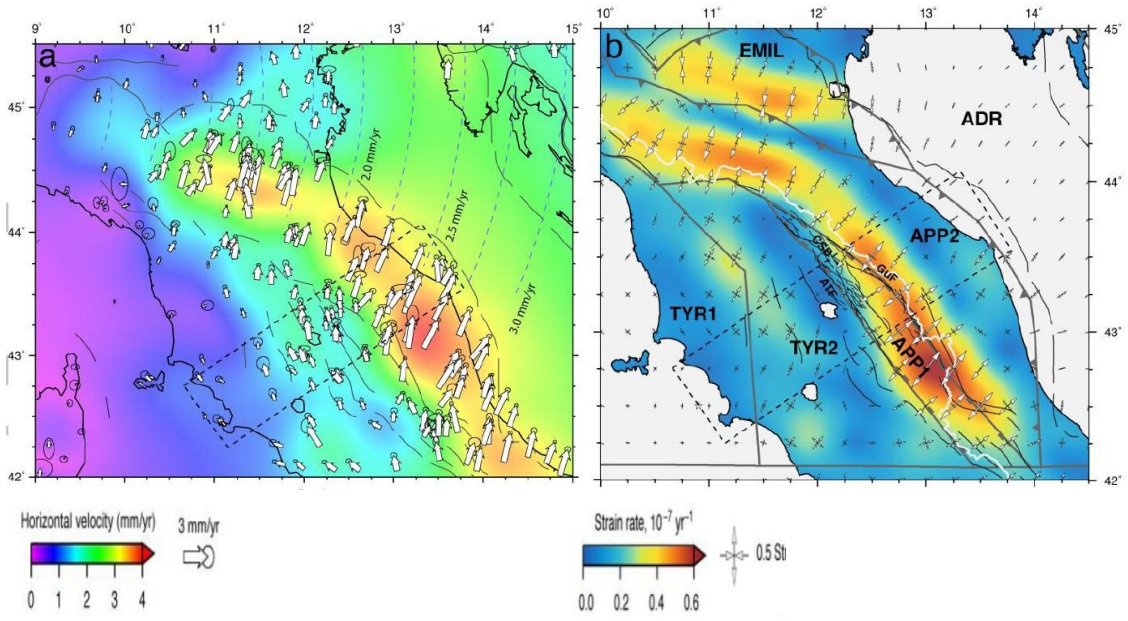


Figure 1.2: a) GPS horizontal velocities (with 95% error ellipses) in a fixed Eurasian frame, superimposed on a map of the continuous velocity field. Dashed blue lines show the motion of Adria relative to Eurasia, while black lines indicate major faults. b) Principal strain-rate axes superimposed on a map of the total (scalar) strain-rate field (Anderlini et al. 2016).

1.1.1 Low angle normal faults, faults mechanics and seismicity

Low angle normal faults (LANFs), characterized by very low dipping (less than 30°), have been largely documented by geological mapping studies [*Lister and Davis*, 1989] and by seismic reflection profiles [*Wernicke*, 1995; *Morley*, 1999; *Taylor et al*., 1999], being firstly recognized in the Basin and Range province, US [*Longwell*, 1945]; however the fact that these structures could generate earthquakes is in contrast with classical fault mechanics. Anderson-Byerlee [1951] theory in fact, predicts that normal faults should form at dips greater than 45°; they may rotate to shallower dips and continue to slip, but for reasonable values of the coefficient of friction, they will lock up before reaching 30°. Thus some studies have suggested that low angle normal faults were initially formed at higher angles, but with time became inactive low angle structures by means of processes like block rotation [*Proffett*, 1977] and/or isostatic adjustment [*Wernicke and Axen*, 1988]. Even though this has been shown to be a possible explanation for low angle normal faults, many field-based studies [*Lister and Davis*, 1989, *Axen*, 1999] and the interpretation of seismic reflection profiles [e.g. *Floyd et al*., 2001] indicate that LANFs can be tectonically active and generate earthquakes [*Rigo et al*., 1996, *Abers et al*., 1997, *Axen*, 1999, *Chiaraluce et al*., 2007]. However, up to now, no moderate-to-large earthquake ruptures have been documented on LANFs [*Jackson and White*, 1989]. This could mean that, mechanically, it is easier to form a new fault instead of activating a severely misoriented low angle structure [*Sibson*, 1985].

However a possible explanation for the lack of large events on LANFs in the contemporary seismic records can be attributable to unusually long recurrence intervals [*Wernicke*, 1995].

If LANF are often believed to be unimportant structures in terms of seismic hazard and in the accommodation of regionally significant amounts of crustal extension [*Chiaraluce* et al., 2007], Anderlini et al. [2016] shows that aseismic sliding on the ATF can partially explain the observed extension.

1.2 The Altotiberina Near Fault Observatory (TABOO)

The Altotiberina Near Fault Observatory (TABOO) is a multidisciplinary research infrastructure devoted to studying tectonic processes tacking place along this normal fault system. Such methodological approach, based on an extremely high spatial resolution, can be more easily applied at the local scale [*Chiaraluce et al*., 2014]. The area monitored by The Altotiberina Near Fault Observatory (TABOO) is identified as the projection at the surface of the ATF plane reconstructed by the use of seismic reflection profiles and surface geology [*Mirabella et al*., 2011].

TABOO was included in 2010 as one of the Near Fault Observatories (NFO) of the European Plate Observing System (EPOS), a European infrastructural project aiming at integrate European research infrastructures aimed to monitor and understand the dynamic of the solid-Earth System.

It consists of 50 permanent seismic stations covering an area of 120 × 120 km^2, 24 of which are equipped with continuous GPS stations, forming two transects across the fault system. Geochemical and electromagnetic stations have been also deployed in the study

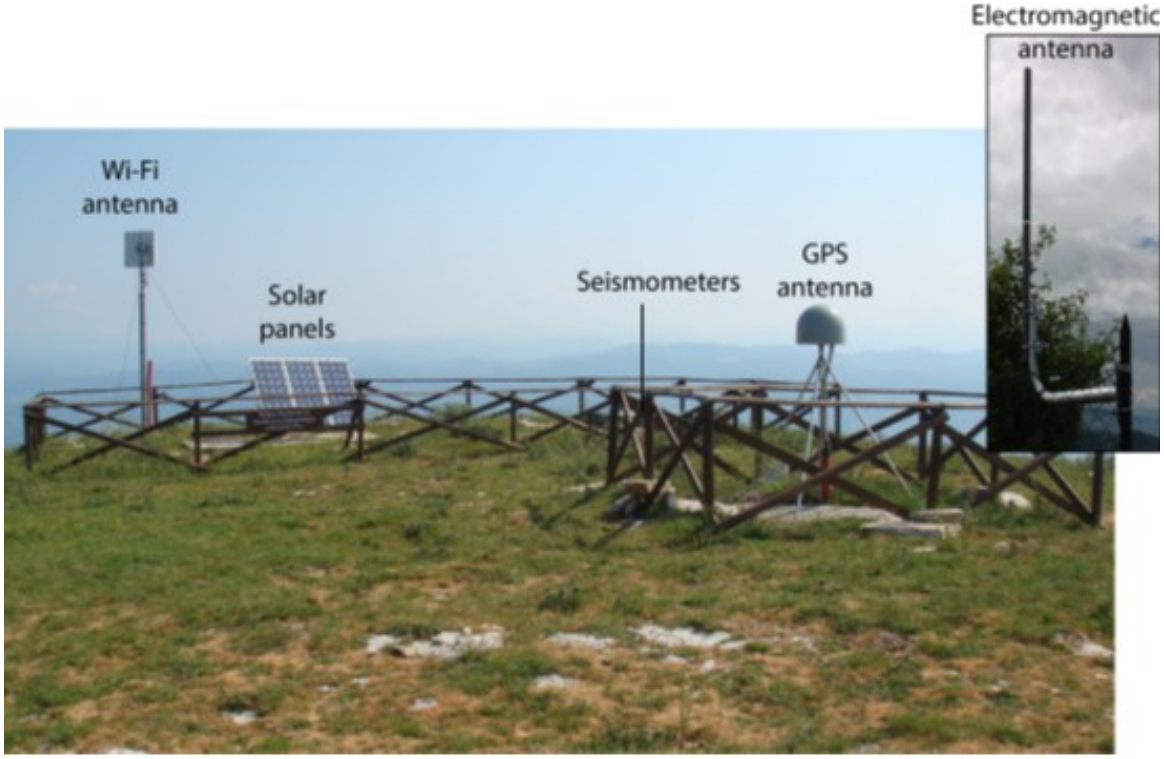


Figure 1.3: Equipment of a TABOO station (ingvterremoti.wordpress.com).

area [*Chiaraluce et al*., 2014]. Figure 1.3 shows the equipment of a TABOO station. The presence of such a dense network, recording continuously high-resolution data, makes this region one of the best monitored in the world and will contribute to understand the mechanics of this LANF and to evaluate the seismic potential associated to this geological structure. The ATF is also an ideal site to study the relationship between fluids, seismicity patterns and faulting because of the presence of deep fluid circulation [*Chiaraluce et al*., 2014].

2. GPS data and surface displacements

The GPS technique, introduced in 1973 in USA for military purposes, is now widely used to detect crustal deformations at different temporal and spatial scales. Due to the increasing number of middle to long GPS time series and the improvement of measurement accuracy and spatial density, a description of the kinematics of complex tectonic frameworks all over the world is becoming possible, including the detection of even small transient tectonic deformations.

In the region of interest for this work, a dense network of continuous GPS stations is now operating: many of them were built when the TABOO project started, in 2010, while others, belonging to INGV RING (*Rete Integrata Nazionale GPS*) network, local and European (EUREF) networks, are active from several years.

Section 2.1 describes the type of data that I have used in this thesis and how they have been obtained (section 2.1.1), while in section 2.2 I will describe strain time series obtained from GPS position time series.

2.1 Data: GPS time series

In this thesis I used high-resolution data obtained from continuous GPS stations operating in a region of ~7800 km^2 surrounding the Altotiberina fault. Daily position time series, recorded by each station, are obtained through a three-step approach as described in section 2.1.1. These time series have been used to compute the evolution of strain with time (section 2.2). Furthermore, an Independent Component Analysis (ICA) has been

performed on raw displacement time series to find the decomposition that best describes the observations (Chapter 3).

2.1.1 GPS data processing and position time series

GPS data processing consists in a three-step approach as in Serpelloni et al. [2006, 2010], including 1) raw phase data reduction, 2) combination of loosely-constrained

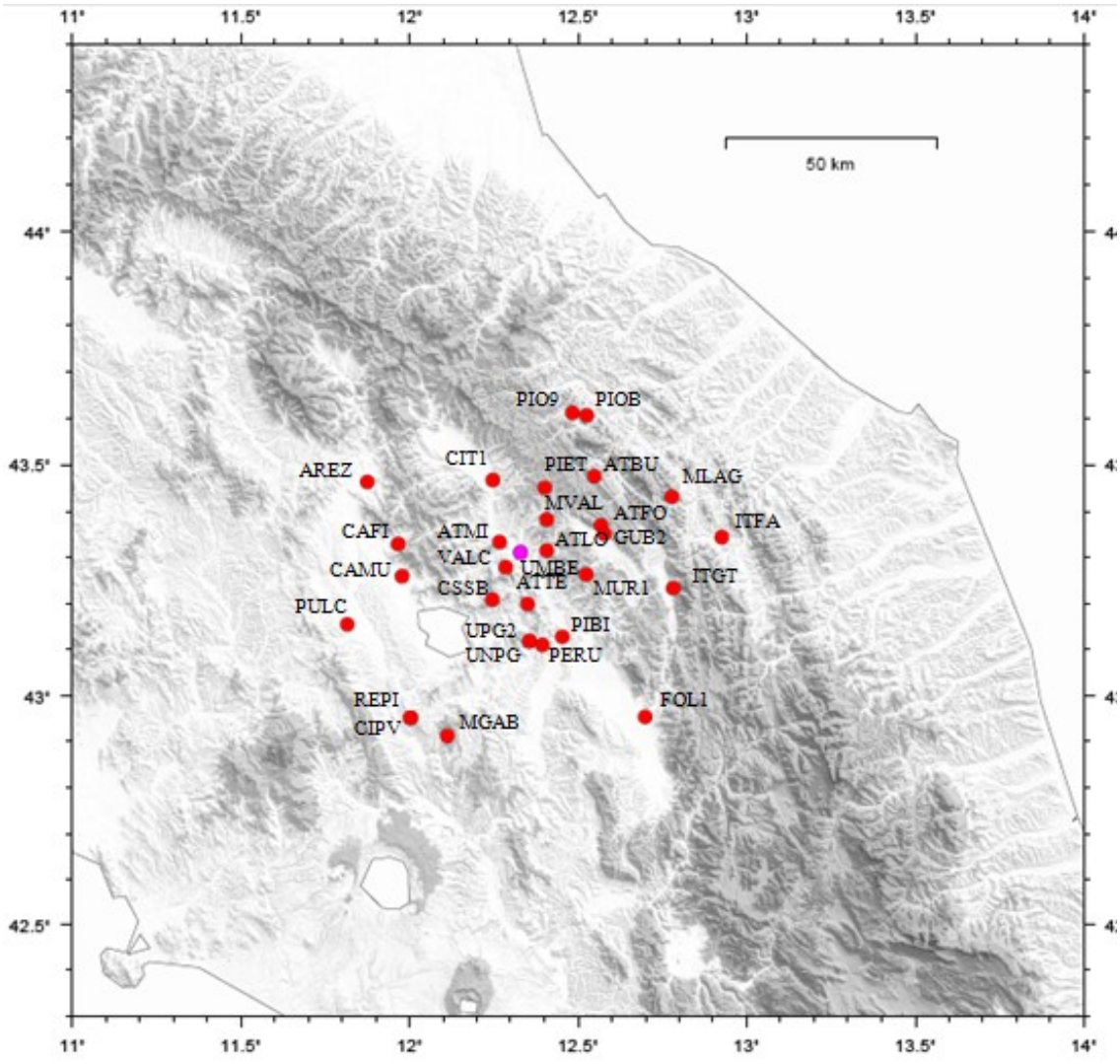


Figure 2.1: Distribution of the GPS stations used on this work. The pink dot indicates the station UMBE, placed almost at the center of the study area.

solutions and reference frame definition, and 3) time series analysis.

In the first step, daily GPS phase observations are used to estimate site position, adjustments to satellite orbital parameters (EOPs) and time-variable piecewise linear zenith and horizontal gradient tropospheric delay parameters by means of the GAMIT (V10.4) software [*Herring et al*., 2010], applying loose constraints to geodetic parameters. The ocean-loading and a pole-tide correction model FES2004 [*Lyard et al*., 2006], the Global Mapping Function (GMF) [*Boehm et al*., 2006] for both hydrostatic and non-hydrostatic components of the tropospheric delay model, the IGS absolute antenna phase center model for both satellite and ground-based antennas and the Global Pressure and Temperature (GPT) model are applied in this step. Continuous GPS (cGPS) and Survey-mode GPS networks (sGPS) are processed separately: following a simple geographic criterion, they are divided into sub-networks sharing a set of high quality IGS stations, which are used as tie-stations in the second step. Loosely constrained solutions are stored in form of ASCII GAMIT H-files and SINEX files, and contribute to the combined solution provided by Istituto Nazionale di Geofisica e Vulcanologia (INGV) [*Avallone et al*., 2010; *Devoti et al*., 2010].

In the second step the ST_FILTER program of the QOCA software [*Dong et al*., 2002] allows us to combine daily loosely constrained solutions with the global solutions made available by SOPAC [*http://sopac.ucsd.edu*] and to simultaneously realize a global reference frame by applying generalized constraints [*Dong et al*., 1998]. The reference frame is obtained minimizing coordinates and velocities of the IGS global core stations [*http://igscb.jpl.nasa.gov*], while estimating a seven-parameter transformation with respect to the IGS realization of the ITRF2008 NNR frame [*Altamimi et al*., 2011].

In the third step, tridimensional velocities and related uncertainties are estimated through the analysis of the position time series using the *analyze_tseri* program of the QOCA software. For cGPS stations, annual and semi-annual seasonal components are estimated together with a secular component while for sGPS sites only the latter component is estimated. Outliers, if present, are cleaned adopting a post-fit RMS criterion discarding values larger than 3 times the post-fit Weighted Root Mean Square (WRMS). Only stations having a minimum length of 2.5 years are retained in this analysis, as shorter intervals may result in biased estimates of linear velocities [*Blewitt and Lavallee*, 2002]. Residual time series contain various systematic and random errors, as well as unmodeled

signals; one of the major spatially correlated error sources in GPS solutions is the so-called common mode error (CME); CME can be removed from the residual time-series applying a principal component analysis (PCA) [*Dong et al*., 2006], improving significantly the signal to noise ratio. Velocity uncertainties are estimated adopting the maximum likelihood estimation (MLE) technique, implemented in the CATS software [*Williams*, 2008], considering a combination of white noise (WN) and power law noise (PL) models.

The dataset used in this work consists of 30 cGPS stations (see Figure 2.1) located within 50 km from the station UMBE, which is placed almost at the center of the investigated area. Time series for these stations are analyzed in a time span of 8 years, from 2008 to the end of 2015.

Figure 2.2 shows the occupation history of the GPS stations used in this work (the percentage of missing data is 34,6%). As described in the following chapters, reducing the amount of missing data could improve the results of an independent component analysis (ICA); for this reason stations with too shorts time series have been removed from the dataset.

The independent component analysis has been performed on raw time series, de-trended for a constant velocity term and cleaned for CME, instrumental offsets and outliers. While both raw time series cleaned for CME, instrumental offsets and outliers and

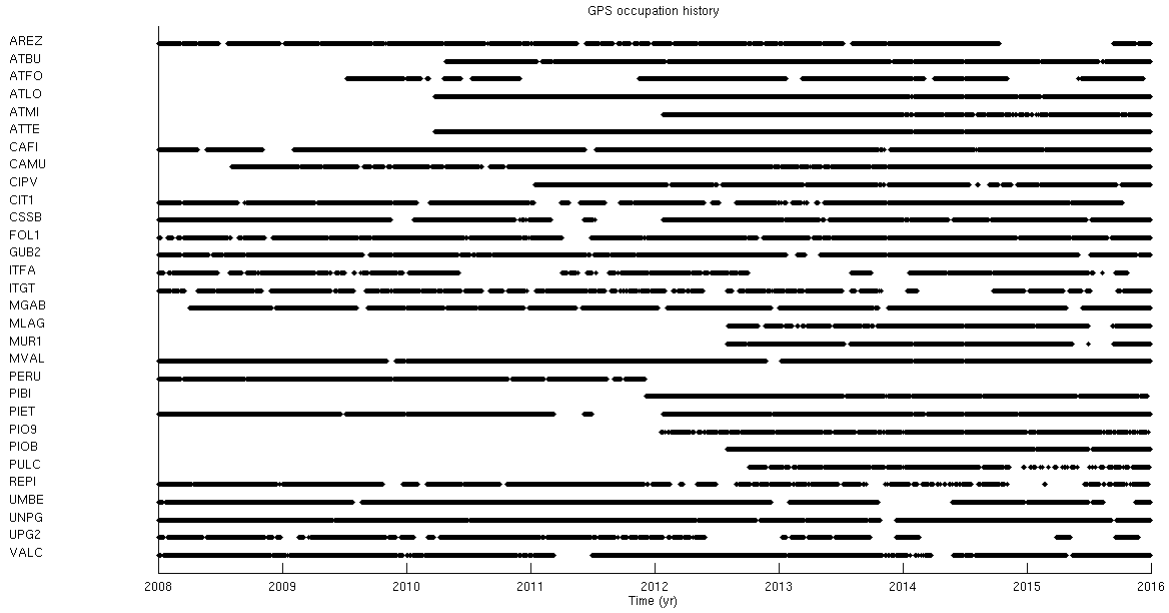


Figure 2.2: Occupation table for the dataset. Horizontal axis: time span analyzed. Vertical axis: station names.

raw time series cleaned for CME, instrumental offsets, outliers and for a seasonal velocity term (annual and semi-annual) have been used in the strain analysis.

2.2 Strain-rates and time-series of strain

The strain-rate field computed from geodetic velocity data measures the contribution of both seismic and aseismic deformation of the crust and is a fundamental quantity that helps in quantifying the seismic potential of active faults. Commonly, strain-rate analyses are performed starting from horizontal velocity fields, providing information on the kinematics and tectonics of a deforming region. For the study area, this analysis has been performed by Anderlini et al. [2016] showing that the area of highest deformation rate (~0.67×10^{-7} yr^{-1} of NE-SW oriented extension) is located near the crest of the Apennines; the highest compressional rate instead is located at the eastern foothills of the Northern Apennines and continues to the SE of the Adriatic coast at lower rates. The transition between the compressional and extensional regimes takes place around longitude 13°E. The Tyrrhenian sector instead exhibits relatively higher strain rates ($0.3\text{-}0.4 \times 10^{-7}$ yr^{-1} NE-SW oriented extension).

Alternatively, one can analyze the evolution of strain with time starting from the position time-series rather than from the stationary velocities. The cumulative strain time series, obtained from the analysis of position time series of a continuous GPS network, can be a useful tool to detect anomalous deformation signals and transients of different origin.

The strain estimation in a three-dimensional medium necessitates of the knowledge of the continuous displacement field, i.e. at each point of the medium [*Pietrantonio and Riguzzi*, 2004]. Here I will describe the strain tensor and how it can be derived from geodetic displacements and velocity measurements. Without loss of generality, for simplicity, one can consider the two-dimensional (i.e., horizontal) case [*Feigl* et al., 1990].

It is useful to first treat a single epoch survey, performed at time *t*, then the full time-dependent problem will simply follow. The calculation of the two-dimensional displacement gradient tensor from GPS data, allows the computation of the strain tensor [*Allmendinger* et al., 2007]:

$$ {}^{k}u_i = t_i + \frac{\partial {}^{k}u_i}{\partial {}^{k}x_j} {}^{k}x_j = t_i + e_{ij} {}^{k}x_j \tag{2.1} $$

where ${}^{k}u_i$ is the i horizontal component of displacement at the station k (in east and north directions), t_i is a constant of integration, ${}^{k}x_j$ is the position of the station and e_{ij} is the displacement gradient tensor. Following Allmendinger et al. [2007], the equations for n stations can be rewritten as:

$$ \begin{bmatrix} {}^{1}u_1 \\ {}^{1}u_2 \\ {}^{2}u_1 \\ {}^{2}u_2 \\ \dots \\ \dots \\ {}^{n}u_1 \\ {}^{n}u_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & {}^{1}x_1 & {}^{1}x_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & {}^{1}x_1 & {}^{1}x_2 \\ 1 & 0 & {}^{2}x_1 & {}^{2}x_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & {}^{2}x_1 & {}^{2}x_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & {}^{n}x_1 & {}^{n}x_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & {}^{n}x_1 & {}^{n}x_2 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \end{bmatrix} \tag{2.2} $$

Data from at least 3 stations ($n \geq 3$) are needed because in two dimensions there are six unknowns. The solutions can be found using a least squares method. For infinitesimal strains, tensor e_{ij} can be additively decomposed into a strain and a rotation tensor:

$$ e_{ij} = \varepsilon_{ij} + \omega_{ij} = \left(\frac{e_{ij} + e_{ji}}{2}\right) + \left(\frac{e_{ij} - e_{ji}}{2}\right) \tag{2.3} $$

where ε_{ij} is the infinitesimal strain tensor and ω_{ij} is the rotation tensor.

The diagonal elements of ε_{ij} represent compression or dilatation according to the sign, while the extra-diagonal ones are the shear terms. The principal deformations λ_i, are obtained by solving the classical eigenvalue problem for ε_{ij}:

$$ \det\left(\varepsilon_{ij} - \lambda I\right) = 0 \tag{2.4} $$

The maximum shear (γ_{max}) is:

$$\gamma_{max} = \frac{1}{2}(\lambda_3 - \lambda_1) \tag{2.5}$$

where λ_3 and λ_1 are the highest and the smallest values, respectively; the directions of the maximum shear are the orthogonal bisectors of the four angles defined by the directions of the principal deformations [*Pietrantonio and Riguzzi*, 2004].

A strain analysis in the investigated area has been performed through the utility *analyze_strain* of the software QOCA, using as input the daily displacements in different reference frames. The underlying assumptions of the software are:

- elastic medium;
- linear deformations;
- 2-D (horizontal) strain model.

From coordinate adjustments one can obtain the two-dimensional displacement gradient tensor for each station at each time which can be decomposed into 7 linear combinations of the displacement gradient components [*Feigl et al.*, 1990]:

$$\begin{bmatrix} E_{11} \\ E_{12} \\ E_{22} \\ \omega \\ \gamma_1 \\ \gamma_2 \\ \Delta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0.5 & -0.5 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{bmatrix} \tag{2.6}$$

where u (v) is the x (y) component of displacement, E_{11}, E_{12}, E_{22} can be used to form the strain tensor ellipsoidal, γ_1 and $\gamma_2 = 2E_{12}$ are shear strains and Δ is the dilatation.

Figure 2.3 shows time-series of dilatation obtained using GPS displacements as input of *analyze-strain*. The slope of the interpolating lines represents the dilatation rate and has the same value (~0.21×10^{-7} yr^{-1}, NE-SW oriented extension) in both Figure 2.3a and 2.3b. It is coherent with the strain-rate field estimated in Anderlni et al. [2016] (see Figure 1.2b) for the region surrounding the Altotiberina fault. The latter was obtained

adopting spherical wavelets in order to estimate a spatially continuous velocity field from which the strain-rate field is evaluated as in Tape et al [2009].

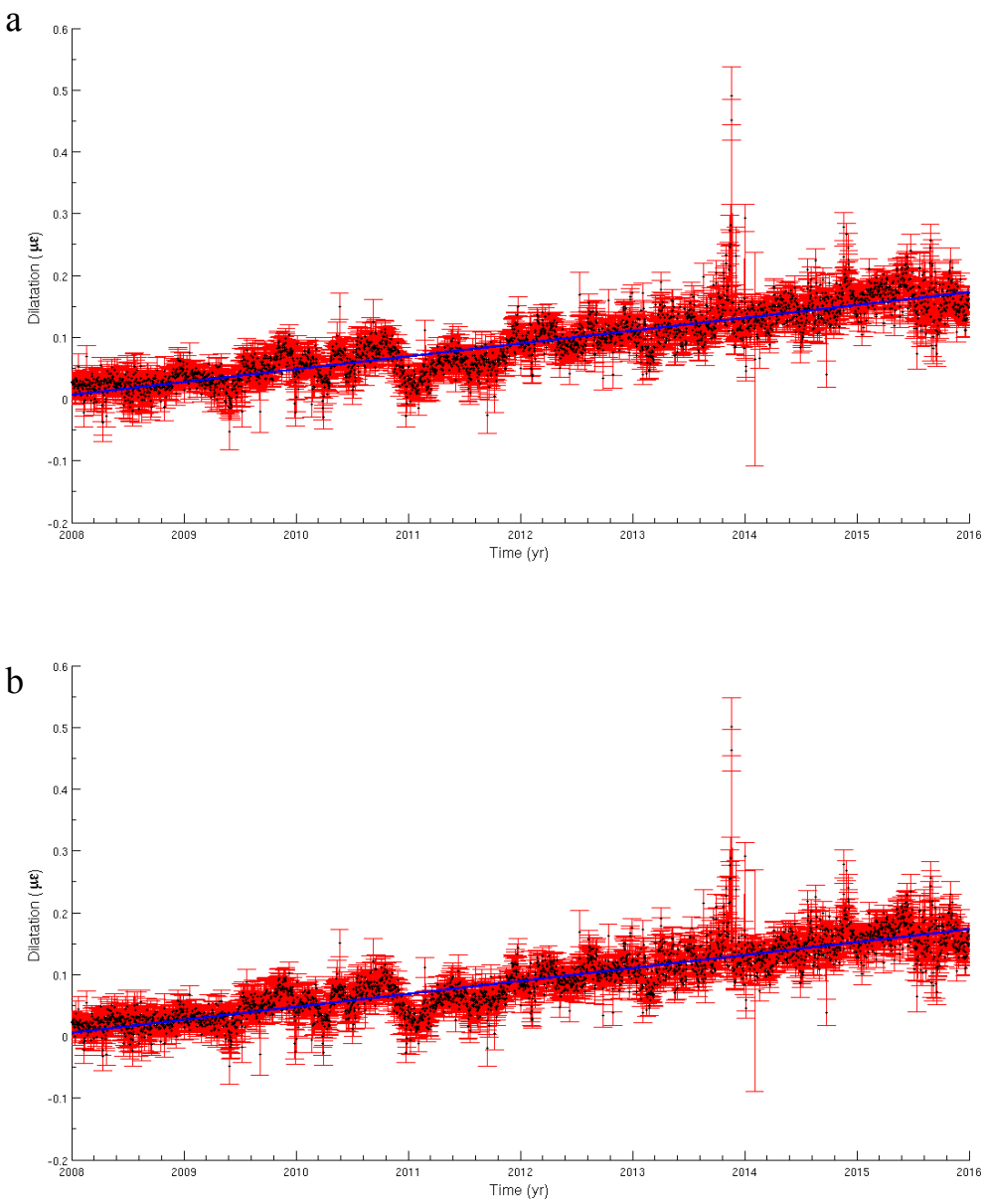


Figure 2.3: Dilatation time series obtained using as input a) raw time series cleaned for CME, instrumental offsets and outliers, b) raw time series cleaned for CME, instrumental offsets, outliers and a seasonal velocity term (annual and semi-annual). Black dots indicate the dilatation value at each epoch with error bars (in red) obtained from the time series of the stations showed in Figure 2.1. The blue line interpolates the curve (its slope is the dilatation rate).

3. Independent Component Analysis (ICA) on GPS time series

As discussed in Chapter 2, a way to analyze GPS time series is to model them as a superposition of functions (linear, seasonal, logarithmic, exponential, etc.) described by a proper number of parameters. Another way to perform time series analysis is to use multivariate statistics techniques such as Principal Component Analysis (PCA) and Independent Component Analysis (ICA). PCA allows to decompose time series in a certain number of components, ordered according to the amount of variance explained in the initial geodetic dataset. Similarly ICA looks for a decomposition of the original data matrix, but with some more assumptions on the source signal.

In this work the analysis of GPS time series has been performed using the ICAIM (ICA-based Inversion Method) software developed at the INGV (Istituto Nazionale di Geofisica e Vulcanologia) and University of Bologna by Adriano Gualandi [*Gualandi*, 2015] as an extension of the PCAIM (Principal Component Analysis-based Inversion Method) code [*Kositsky* and *Avouac*, 2010], developed at the California Institute of Technology (CalTech). After decomposing original data into the appropriate number of linear components that describe the observations, the code allows us to invert the spatial response of one or more components for slip on a predetermined fault geometry. The basic principles of ICA are exposed in this chapter as well as the fundamental steps of the ICAIM code. At last the final decomposition for stations surrounding the Altotiberina fault (ATF) is shown and discussed.

3.1 Independent Component Analysis (ICA): basic principles

The Independent Component Analysis (ICA) is a multivariate statistical method searching a linear transformation of the original data that minimizes the statistical dependence between its components [*Comon*, 1994]. Herault and Jutten were the first to propose the problem of independent component analysis in 1983, while in 1994 Comon gave the first rigorous definition of ICA.

This technique was originally developed to deal with problems of Blind Source Separation (BSS) like the so called *cocktail-party problem*, describing a situation in which two people are speaking simultaneously in a room and their voices are recorded by two microphones placed in different locations. As in Hyvarinen and Oja [2000], one can express the two recorded time signals ($x_1(t)$ and $x_2(t)$) as weighted sums of the speech signals emitted by the two speakers ($s_1(t)$ and $s_1(t)$):

$$x_1(t) = a_{11}s_1 + a_{12}s_2 \tag{3.1}$$

$$x_2(t) = a_{21}s_1 + a_{22}s_2 \tag{3.2}$$

where a_{11}, a_{12}, a_{21}, and a_{22} are some parameters that depend on the distances of the microphones from the speakers. This is a simplified model where time delays, noise and other extra factors are omitted. To solve this equation by classical methods one needs to know the parameters a_{ij}, otherwise it would be very difficult to find a solution. It turns out, however, that assuming $s_1(t)$ and $s_2(t)$ statistically independent at each time instant allows us to separate those source signals from their mixtures $x_1(t)$ and $x_2(t)$ [*Hyvarinen and Oja*, 2000].

A precise definition of ICA could be provided assuming the following linear statistical model [*Comon*, 1994, *Gualandi*, 2015]:

$$X_{M\times T} = A_{M\times L}S_{L\times T} + N_{M\times T} \tag{3.3}$$

where X is the data matrix, A is the so called mixing matrix, S is the source signal matrix and N is some additive noise. X is an M × T matrix, where M is the number of sensors and T the number of epochs; S is a L × T matrix, where L is the number of sources. Independent Component Analysis is the practice of transforming these M statistically dependent sequences into L ≤ M statistically independent sequences [*Choudrey*, 2002].

It is useful to treat noise as one of the sources, obtaining:

$$X_{M\times T} = A_{M\times L} S_{L\times T} \tag{3.4}$$

Each M-dimensional column vector of X can be plotted as a point in the sensor space (M-dimensional) obtaining a distribution of T points. To perform an ICA means to project the sensor space onto an L-dimensional source space, described by an independent coordinate frame [*Choudrey*, 2002].

The underlying assumptions of the independent component analysis are:

- statistical independence of the source signals;
- the values in each source signal have non-Gaussian distributions.

Technically, independence can be defined by means of the probability densities considering two scalar-valued random variables y_1 and y_2. The marginal probability density function (pdf) of y_1 and y_2 is defined as:

$$p_1(y_1) = \int p(y_1, y_2) dy_2 \tag{3.5}$$

$$p_2(y_2) = \int p(y_1, y_2) dy_1 \tag{3.6}$$

where $p(y_1, y_2)$ is the joint pdf of y_1 and y_2. The two variables are said to be independent if and only if the joint pdf is factorizable as:

$$p(y_1, y_2) = p(y_1)p(y_2) \tag{3.7}$$

This definition extends naturally for any number n of random variables, in which case the joint density must be a product of n terms [*Hyvarinen and Oja*, 2000].

As regard the second assumption, that is that the independent components must be non-Gaussian for ICA to be possible, it can be shown that the joint probability density function of Gaussian variables is completely symmetric and does not contain any information on the directions of the columns of the mixing matrix A. Therefore A cannot be estimated for Gaussian variables [*Hyvarinen and Oja*, 2000].

Following the Central Limit Theorem, which tells that the distribution of a sum of independent random variables tends toward a Gaussian distribution, means that the observed data density is more Gaussian than the hypothetical source [*Choudrey*, 2002].

Hyvarinen and Oja [2000] suggest that a way to find independent variables is to measure the deviation from Gaussianity: the classical method to estimate it is *kurtosis*; other methods require the calculation of *negentropy*, *mutual information* (MI) or the use of generative models [*Hyvarinen and Oja*, 2000]. The one adopted in this work is a variational Bayesian algorithm (vbICA), developed by Choudrey [2002] and implemented in the ICAIM code (Independent Component Analysis-based Inversion Method) [*Gualandi*, 2015; *Gualandi* et al., 2016]. A short description of the basic principles of the code is given in next subsection.

3.1.1 The ICAIM software

The ICAIM code, completely written in MATLAB (The MathWorks, Inc., Natick, MA), determines the Independent Components (ICs) using a generative model and applying a variational Bayesian algorithm (vbICA) as described in Choudrey [2002]. The goal of a generative model is both to explain the observations and to match the a priori knowledge; however, this requires the estimation of different parameters, which is not always possible, and approximations, like the variational Bayesian one, are needed [*Gualandi*, 2015]. This approximation allows us to model each independent component i = 1,…,L by a mixture of m_i Gaussians (MoG). Each Gaussian q_i = 1,…,m_i is fully described by a mean μ_{i,q_i} and a precision β_{i,q_i}. The vbICA algorithm imposes that these hidden

variables $\mu_{i,q_i}$and β_{i,q_i}, follow a Gaussian and a Gamma distribution, respectively. In other words, the MoG describing the *i*-th IC is characterized by m_i Gaussians pdf, described by m_i normally distributed means μ_{i,q_i} and Gamma distributed precisions β_{i,q_i}. These $2 \times m_i$-hidden variables are thus described by some hyper-parameters that control the named distributions. In particular, the normally distributed variables μ_{i,q_i} are described by a mean m_{i,q_i} and a precision τ_{i,q_i}; while the Gamma distributed variables β_{i,q_i} are described by a scale parameter b_{i,q_i} and a shape parameter c_{i,q_i}. Being vbICA a generative model it is necessary to specify the a priori starting values of the hyper-parameters: in this study I specify the same a priori hyper-parameters for every IC. In particular I use $m_{i,q_i} = m_{i0} = 0,\ \tau_{i,q_i} = \tau_{i0} = 1,\ b_{i,q_i} = b_{i0} = 10^3$ and $c_{i,q_i} = c_{i0} = 10^{-3},\ q_i = 1,2,3,4$, as a standard configuration assuming that each IC is described by $m_i = 4$ Gaussians as in Gualandi [2015].

The first step of an ICA is to center the data matrix X by subtracting to each time series its mean value. Then the vbICA is performed starting from a Principal Component Analysis (PCA) or a random initialization. In the first case, the starting points for the mixing matrix and for the sources are set using the U and the V values obtained through a Singular Value Decomposition (SVD) of the data matrix:

$$X_{M\times T} \approx U_{M\times L} S_{L\times L} V^T_{T\times L} \tag{3.8}$$

where the diagonal values of S are the first L singular values of X, which explain most of the variance of the data matrix, and the column of U and V are respectively the eigenvectors of the spatial covariance matrix XX^T and of the temporal covariance matrix X^TX. Such decomposition can be also obtained using the weighted low-rank approximation of Srebro and Jaakkola [2003], which permits to easily account for missing data and uncertainties [*Kositsky* and *Avouac*, 2010].

After having chosen the number of components that best describe the observations, as explained in 3.2.2, one could invert the spatial response of one or more components assuming a predetermined fault geometry. The inversion step will be described in details in Chapter 5.

3.2.2 Determining the number of useful components

In order to obtain the optimal ICA decomposition it is necessary to identify the number of components that best describe the observations, without over fitting.

The reduced $\chi 2$ test and the F-test are two commonly used procedures to get the number of components needed to explain the dataset without reproducing the noise. These tests are based on the measure of the residuals in a L_2 norm space. A more advanced way to select the number of components for the vbICA is based on the Negative Free Energy (NFE) quantity (see *Choudrey*, 2002). In order to narrow down the search among several potentially good models an Automatic Relevance Determination (ARD) procedure is also introduced and combined with the NFE estimation (see *Choudrey*, 2002 and *Gualandi*, 2015).

The reduced $\chi 2$ test allows to determine the number of components such that the misfits between the original data and data reconstructed through ICA are, on average, of the order of magnitude of the measurement uncertainties [*Kositsky* and *Avouac*, 2010].

The F-test is based on the ratio of the difference between the $\chi 2$ obtained before and after adding a new component, divided by the $\chi 2$ obtained without the additional component and multiplied by a ratio of degrees of freedom [*Kositsky* and *Avouac*, 2010].

The ARD criterion allows the determination of the number of useful components by comparing the variance value α_i (with i = 1,…,L) associated to each component; this values describe how relevant is the source for the explanation of the data. Considering a decomposition in n components, if the maximum variance is more than 10 times bigger than the smallest one, one can discard the component with the minimum variance [*Gualandi*, 2015; *Gualandi* et al., 2016].

The procedure adopted in this thesis, follows that of Choudrey [2002] and Gualandi [2015], and consists in first selecting the number of components using the ARD criterion, then testing several combinations of priors and choosing the one that minimizes the absolute value of the NFE (see also *Gualandi* et al., 2016).

3.2 ICA decomposition of GPS signals

The methods described so far have been used to analyze daily position time series of GPS stations, obtained following the procedures described in Chapter 2.

Considering that three time-series are associated to each station (one describing the motion along the East-West direction, one along the North-South and one along the vertical), the data matrix X is now a M × T matrix, where M is 3 times the number of stations and T equals the number of recorded epochs.

Following Gualandi [2015], the resulting decomposition can be written as the product of three matrices (U, S, V): the sizes of U and V are M × L and T × L respectively, where T is the number of epochs and L is the number of independent components, while the size of S is L × L. Each column of V describes the temporal evolution of one component; U, instead contains the spatial information relative to the dataset. Both U and V are non dimensional, since the unit of measurement is carried by S.

3.3 Final decomposition for the ATF

As described in section 2.1.1, the dataset used in this work consists of time series belonging to 30 cGPS stations located within 50 km from the center of the studied fault system, assumed to be in the station UMBE (Figure 2.1). The analysis is carried out on time series de-trended for a constant velocity term and cleaned for CME, instrumental offsets and outliers. The reason why time series have been de-trended before starting the ICA is that the linear signal is the most powerful signal and keeping it could cover lower signals. Furthermore, it is difficult to reproduce a linear signal with a mixing of Gaussians [*Gualandi* et al., 2016]. To reduce the amount of missing data, time series have been analyzed in a time span of 8 years, from 2008 to the end of 2015. The ICA has been performed using the ICAIM code, choosing a PCA initialization under the algorithm of Srebro and Jaakkola [2003] (see section 3.1.1).

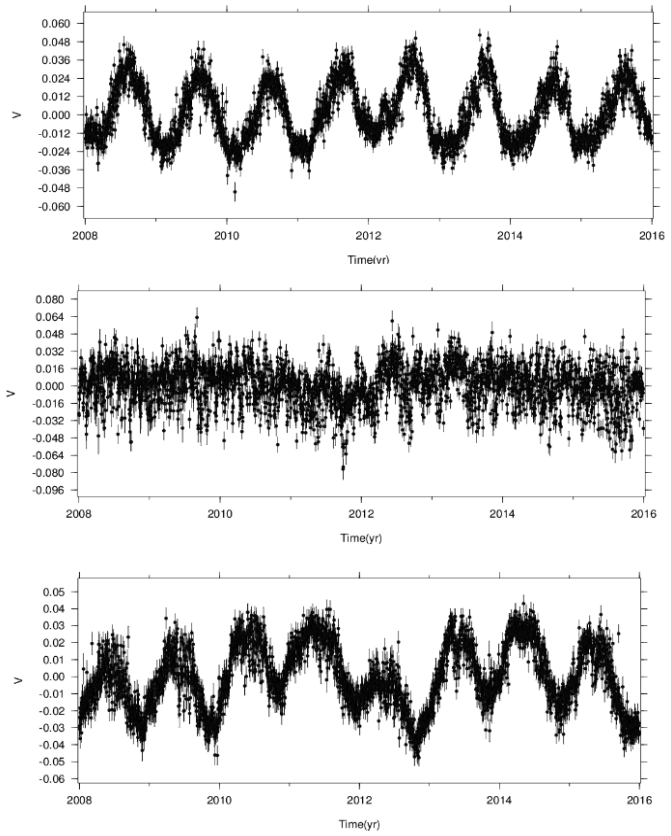


Figure 3.2: Independent components on raw time series de-trended and cleaned for CME, instrumental offsets and outliers.

The ARD criterion suggests considering 3 components in the decomposition (Figure 3.1). Since the Signal to Noise Ratio (SNR), that is a measurement of signal strength relative to background noise, of the third component is greater than 1 only in the horizontal components I decided to analyze separately horizontal and vertical time series. In the first case, the ARD criterion indicates 4 ICs as the proper number of components, while a 3 components decomposition is suggested in the latter case. Figure 3.2 shows the ICs obtained after a separate analysis of horizontal and vertical time series, randomly changing the initialization point and choosing the decomposition with the highest NFE value among 10 random initializations.

When horizontal components are used as input for ICAIM, one more IC, showing a transient signal in 2014, appears; while using vertical time series, the third component of the original decomposition disappears.

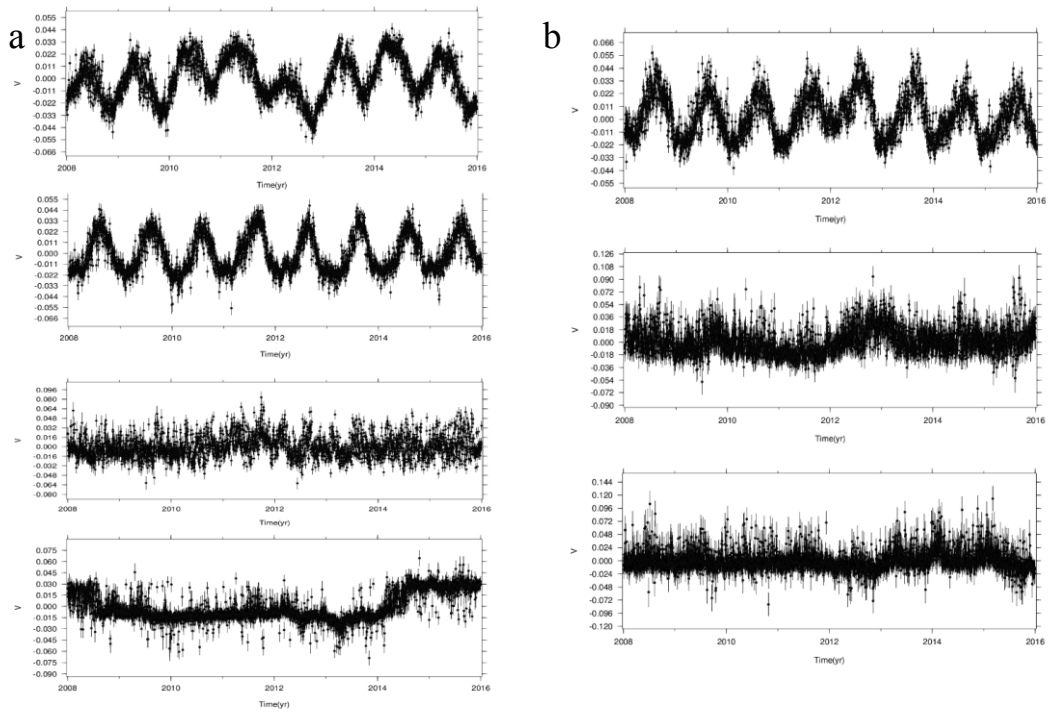


Figure 3.3: ICs obtained for horizontal (a) and vertical (b) time series.

For this reason I decided to pursue the analysis of the ICs obtained from the horizontal time series analysis, searching for a physical interpretation of all the components.

A frequency analysis on the ICs shows a dominant annual period in the first and second IC (Figure 3.3). IC1 and IC2 annual periods are out of phase of 90 days: the first component follows the second by 90 days (Figure 3.4).

IC1, IC2 and IC4 will be described in Chapter 4, where hypothesis on the interpretation of those signals will be formulated. IC3 instead seems to reproduce noise (Figure 3.2a).

The spatial response U is plotted in Figure 3.5 with vectors indicating the horizontal components and S values, which give the unit of measurement, placed on top right of each panel.

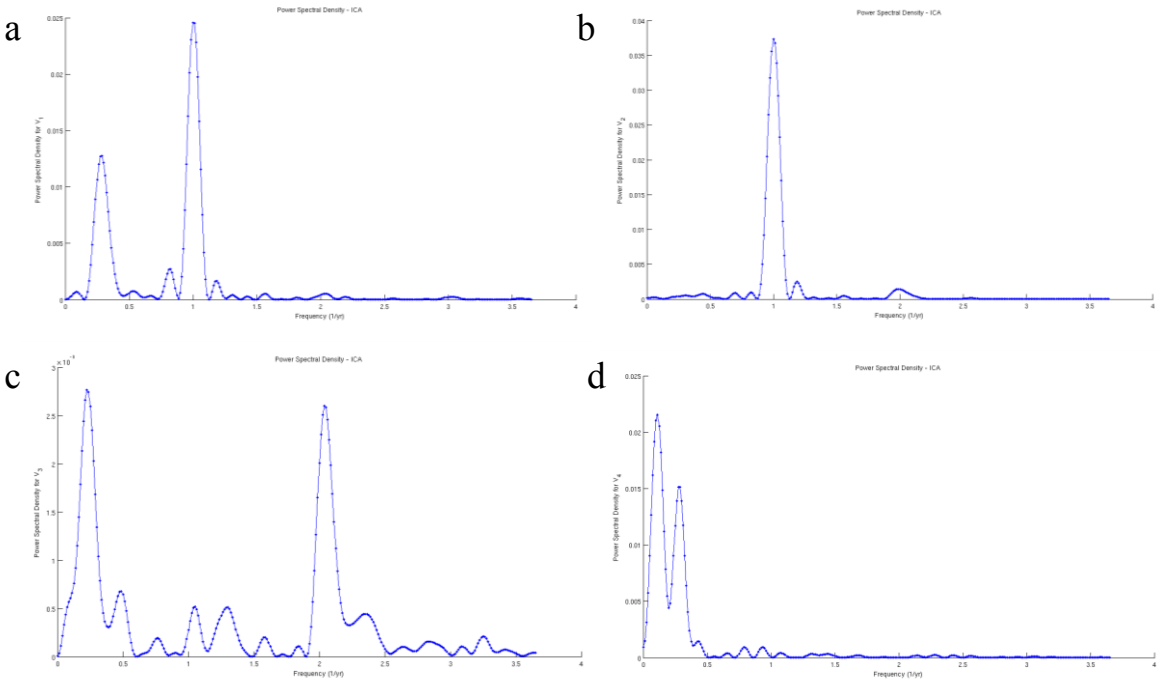


Figure 3.4: Power spectral density of the ICs (IC1 (a), IC2 (b), IC3 (c), IC4 (d)) of Figure 3.2a

IC4 exhibits a spatial response which seems to reproduce stretching, typical of normal faults. For this reason, an inversion has been performed over this component on a fault geometry placed at the center of the opening, as it will be described in Chapter 5.

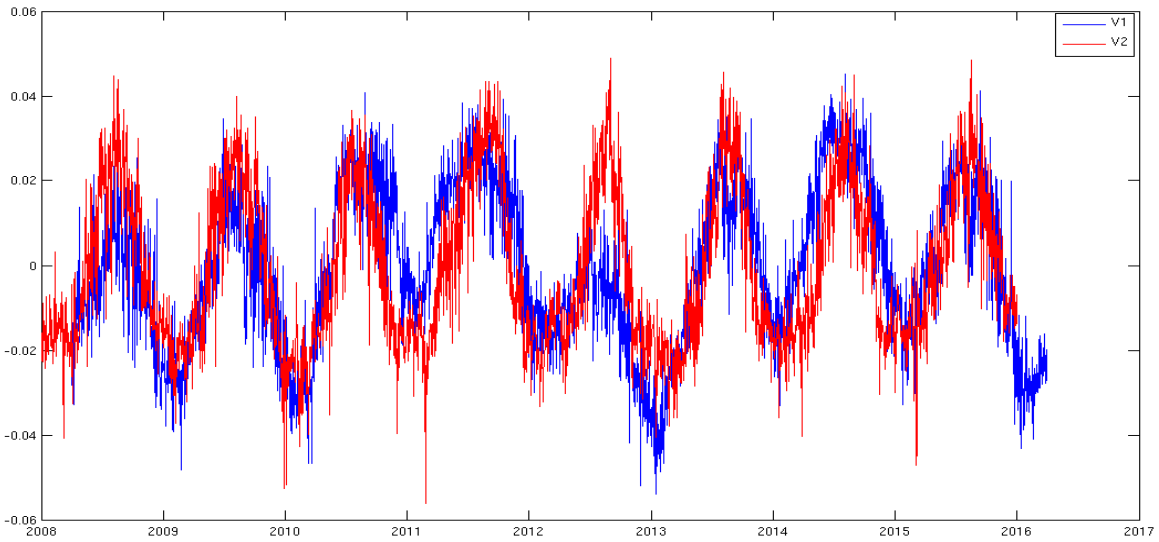


Figure 3.4: Superposition of IC1 (blue) shifted beyond of 90 days and IC2 (red).

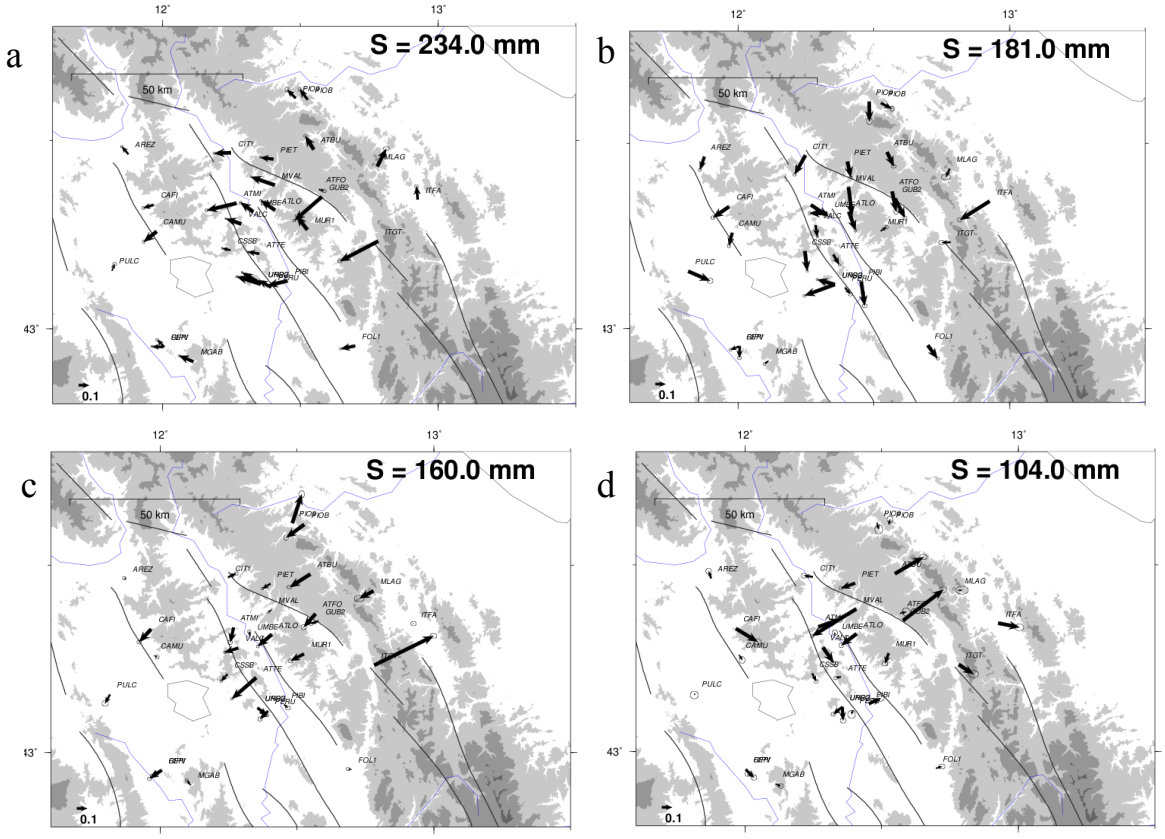


Figure 3.5: Spatial response for IC1 (a), IC2 (b), IC3 (c), IC4 (d).

4. Analysis of the independent components

In Chapter 3 I have described how to apply ICA to GPS time series, ending up with a decomposition in four independent components, obtained using only horizontal GPS time series as input of the ICAIM code. The next step consists in finding a physical interpretation of the components, trying to identify the sources of those signals.

In this chapter I will analyze each IC of the final decomposition, except for the third one which seems to reproduce mainly noise. Some correlations with other signals will be presented, which could explain the presence of those deformation signals (or fingerprints) in GPS records.

4.1 First component

As seen in Chapter 3, the first IC of the final decomposition exhibits a seasonal periodicity. Figure 4.1 shows that in a region around 2011-2013, the temporal component of IC1 (black) differs from a sinusoidal curve (red). This is the most powerful signal in the decomposition, being the Signal to Noise Ratio (SNR) for this component bigger than 1 for more than $\frac{2}{3}$ of the stations. The time series of the station GUB2 (Figure 4.2) is reported as an example of a time series in which the IC1 signal presence is clear at visual inspection.

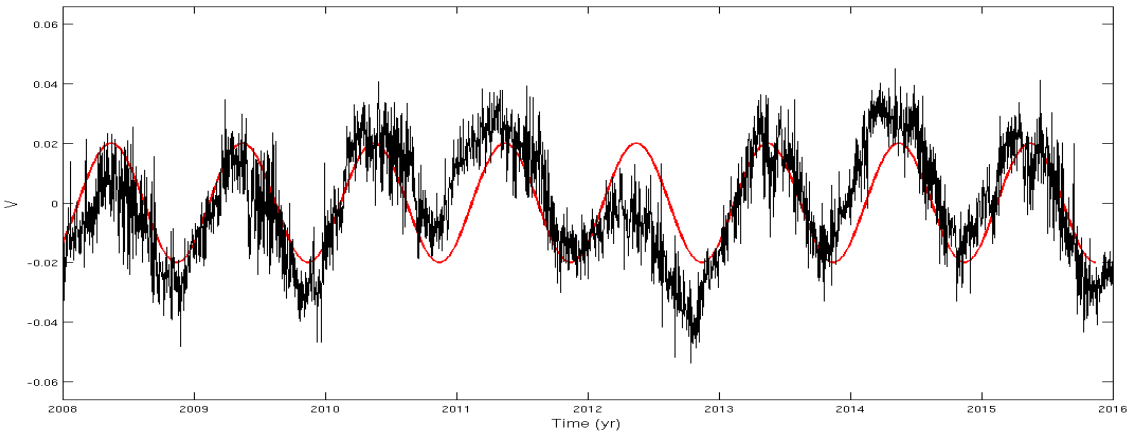


Figure 4.1: First independent component of the final decomposition (black). A sinusoid with annual period (red) has been superimposed to the IC1 in order to show that the IC1 signal differs from a sinusoid in the time interval 2011-2013.

In section 4.1.1 the correlation between the first component of the IC decomposition and the cumulative precipitations is described.

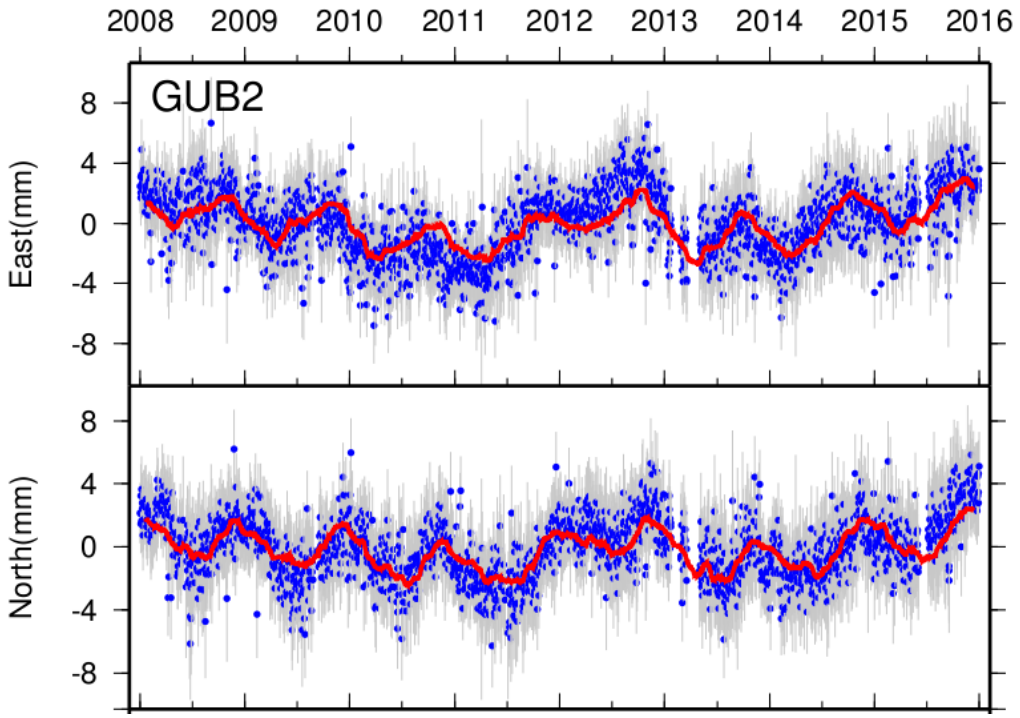


Figure 4.2: Horizontal components of the time series of station GUB2. Blue dots indicate surface displacement at each epoch, while the red curve is the reconstruction based on the ICA.

4.1.1 Correlation with rainfall

In order to analyze the precipitation trend, rainfall values are taken from repositories of the European Commission's 6th Framework Programme (EU-FP6) project ENSEMBLES (*http://ensembles-eu.metoffice.com*). They are high resolution daily gridded data obtained through a three-step process of interpolation of precipitation data, recorded by meteorological stations located all around Europe [*Haylock* et al., 2008]. The catalog contains the millimeters of rainfall per day for each cell of a regular 0.25°×0.25° grid.

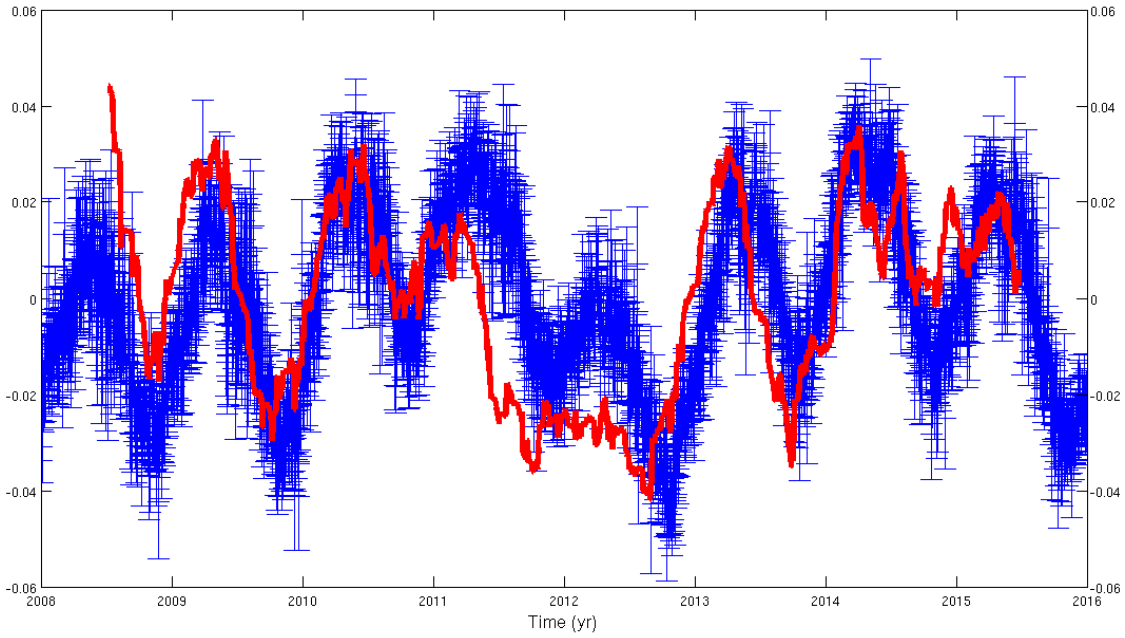


Figure 4.3: Superposition of the first IC (blue) and of the 190-days rainfall accumulation curve (red). Both curves have been rescaled.

In order to obtain the daily amount of rainfall (mm/day) in the same region analyzed with the ICAIM software (within 50 km from the station UMBE) and in the same time span (2008-2015), a stacking has been performed over each cell of the grid belonging to the analyzed region. After that, curves of cumulated precipitation, obtained varying the number of accumulation days, have been compared with the temporal component of the first IC finding that a 190-days accumulation curve best fits the IC1 trend, with a

correlation of 0.6. The precipitation curve has been centered along the y axis and both curves have been rescaled in order to be compared. Figure 4.3 shows the 190-days rainfall accumulation curve (red) superimposed on the temporal component of the first IC (blue). It can be seen that the precipitation curve departs from a sinusoidal behavior in 2012; indeed, that year was characterized by a very slow rainfall rate that seems to have influenced the horizontal component of surface displacement in the region of interest. The departure from a sinusoidal signal is partially present also in the vertical component decomposition (Figure 3.2b), but the first IC of that analysis is out of phase with respect to the cumulated rainfall. Indeed it has the same phase of the IC2 resulting from the analysis of the horizontal time series that is described in next section. This mixture of signals may be caused by the higher level of noise in the vertical component, typical of GPS measurements.

This analysis suggests that seasonal geodetic displacements seem to reflect a lithospheric response to the seasonal variation of hydrological surface loads. The hydrological cycle in fact, has been shown to be capable of producing measurable deformations of the lithosphere [e.g. *Christiansen* et al., 2007; *Bettinelli* et al., 2008]. Modeling of these effects requires knowledge of the hydrological properties of the crust and is beyond the scope of this thesis.

4.2 Second component

As seen in Chapter 3 (Figure 3.2a), the second IC of the final decomposition describes a sinusoidal signal with annual period (Figure 4.4). The analysis of the local seismic catalog shows an annual variation of the number of earthquakes that correlates with this IC, as detailed in the following of this section.

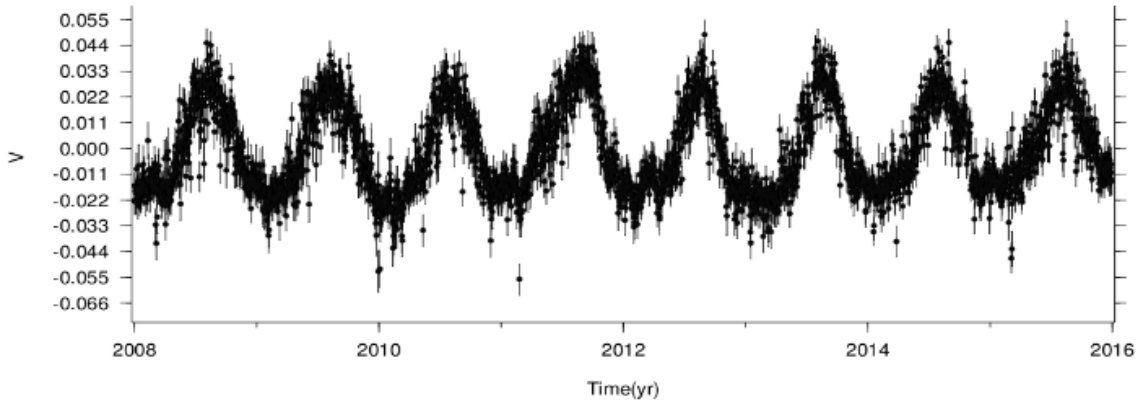


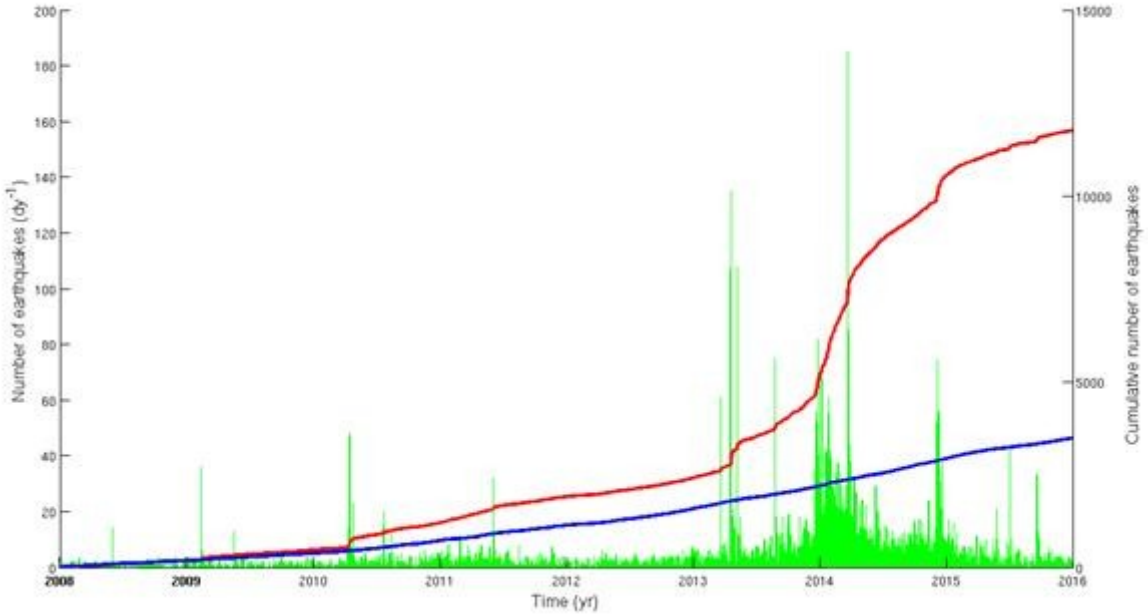
Figure 4.4: Second IC of the final decomposition.

4.2.1 Correlation with the seismic catalog

I have analyzed the seismic catalog of the study region in order to search for an annual periodicity in the seismicity rate. This analysis has been performed on the Italian Seismological Instrumental and parametric Data-basE (ISIDE) catalog of INGV (iside.rm.ingv.it), selecting events occurred in the same region analyzed with ICAIM (within 50 km from the station UMBE) and in the same time span (2008-2015), as well as on events of another catalog of relocated seismicity (Luisa Valoroso, pers. comm.) which covers a time span of almost 4 years (April 2010 – May 2014). Histograms (green) indicating the number of events per day for each catalog are reported in Figure 4.5.

Next section contains a description of the Schuster spectrum code used to search periodicities in the seismic catalogs of the ATF area.

a



b

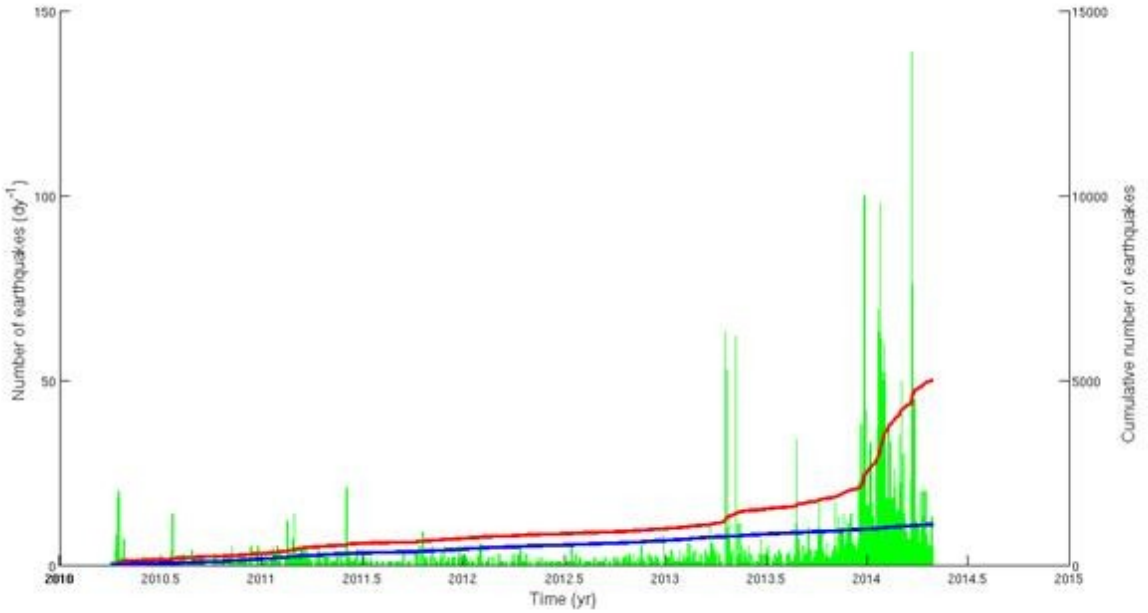


Figure 4.5: Events for the ISIDE catalog (a) and for the catalog of relocated seismicity (b). The green histograms represent the number of events per day. The continuous lines indicate the cumulative number of earthquakes with magnitudes above the detection thresholds (M_l = 1.2 in panel a and M_l = 1 in panel b) of the original catalog (blue line) and of the de-clustered catalog (red line).

4.2.1.1 Schuster spectrum software

The Schuster spectrum software was developed by Thomas J. Ader and Jean-Philippe Avouac [*Ader* and *Avouac*, 2013] as a generalization of the Schuster test [*Schuster*, 1897]. While the Schuster test allows to determine whether a given periodicity is present or not in a seismic catalog, the Schuster Spectrum looks for periodicities in a given period range allowing finding more than one periodicity, if present.

The Schuster test is a statistical test that associates a phase to each event of the seismic catalog:

$$\vartheta_k = 2\pi \frac{t_k}{T} \tag{4.1}$$

where t_k is the time of event number *k* and *T* is the tested period. The phase of each event is represented in the complex plane as:

$$d_k = e^{i\vartheta_k} = \cos\vartheta_k + i\sin\vartheta_k \tag{4.2}$$

Starting at the center of the complex plane, after *K* events the total distance from the origin will be:

$$D_K = \sum_{k=1}^{K} d_k = |D_K| e^{i\theta_K} \tag{4.3}$$

where:

$$\theta_K = \frac{1}{K}\sum_{k=1}^{K} \vartheta_k \tag{4.4}$$

$$|D_K| = \sqrt{A_K^2 + B_K^2} \tag{4.5}$$

and where:

$$A_K = \sum_{k=1}^{K} \cos\vartheta_k \tag{4.6}$$

$$B_K = \sum_{k=1}^{K} \sin \vartheta_k \tag{4.7}$$

Calling D the final distance obtained when K equals the total number of events in the catalog, Schuster [1897] defines the probability (Schuster p-value) that a distance equal to or greater than D can be reached by a random walk as:

$$p = e^{-\frac{D^2}{N}} \tag{4.8}$$

where N is the number of events in the catalog. Therefore, 1-p represents the significance level to reject the hypothesis that earthquakes occur randomly. This means that a lower p-value corresponds to a higher probability to have a periodicity at the period T in the catalog.

It can be seen [*Ader* and *Avouac*, 2013] that the logarithm of the Schuster p-value is independent of the tested period T, depending only on the number N of events in the catalog and on the amplitude α of the seismicity rate variations:

$$\langle -\ln p \rangle = \langle \frac{D^2}{N} \rangle = 1 + \frac{N\alpha^2}{4} \tag{4.9}$$

Ader and Avouac [2013] generalize the Schuster test to overcome its main problem: while it is sure that a periodicity in the catalog gives a low p-value, a low p-value itself cannot indicate unambiguously a periodicity in the catalog. Calculating a spectrum of Schuster p-values allows to resolve this issue and to search for any unknown periodicities, as explained hereafter.

In order to determine the subset of periods to test suitable to obtain a complete spectrum of Schuster p-values for the catalog, it is important to notice that when testing a period T, the catalog gets stacked over that period, so the continuous range of periodicities around T that remain coherent throughout the stacking process gets tested. For this reason the first step is to define a proper period sampling in order to avoid both long computation time due to oversampled testing and skipping to test any relevant period. Two consecutive tested periods have to verify the condition:

$$\Delta T_i = \frac{\varepsilon T_i^2}{t} \tag{4.10}$$

where ΔT_i is the period increment, T_i is the i-th tested period, t is the length of the time series and ε determines the period sampling. The optimal value of ε is $\varepsilon_0 \approx 1$ (see Ader and Avouac [2012] for the proof).

The second step is to determine the probability threshold below which a Schuster p-value can be regarded as indicating with confidence a periodicity of the catalog. Following Ader and Avouac [2012] the expected value of the maximum Schuster probability for a given tested period T_i is:

$$\langle \delta_i \rangle = \frac{T_i}{t} \tag{4.11}$$

which means that the expected value is higher for higher periods.

If a p-value is significantly lower than the threshold value, the probability to have a periodicity in the catalog is high. In other words a low p-value means that the null hypothesis that earthquakes take place randomly is rejected.

The Schuster spectrum code needs as input the epochs of the events in the catalog and a range of periods to be tested. The proper range should be between $t \times 10^{-4}$ and t, where t is the length of the catalog [*Ader* and *Avouac*, 2012].

Sometimes declustering the catalog is necessary because keeping events not related to the background activity (e.g. aftershocks or swarms) might conceal some periodic variations in the background seismicity rate [*Ader* and *Avouac*, 2012]. This misinterpretation of clusters is due to the stacking process and it results in low p-values for periods above a given period. Since a not properly declustered catalog shows a cloud of low p-values at large periods, performing a Schuster spectrum would help to understand whether in a catalog a given periodicity is real or an artifact produced by the presence of clusters, while a simple Schuster's test cannot achieve this target.

4.2.1.2 Results

A first test to search periodicities has been performed on the ISIDE catalog in the range between 8×10^{-4} and 8 years. 11768 events, with magnitudes above $M_l = 1.2$, have been analyzed. I have estimated the magnitude detection threshold from the magnitude-frequency distribution of earthquakes. Results are shown in Figure 4.6, where Schuster *p*-values are plotted as a function of the period tested, in a log-log graph (note the reverse scale in the y-axis). A periodicity of one year is evident but the cloud of points, with low *p*-values (periods larger than 10^{-2} yr) suggests the presence of clusters in the catalog. Therefore, a declustering seems to be necessary in order to verify if the low *p*-value obtained at 1 year describes a real periodicity of the catalog. I have performed the declustering using the Reasenberg algorithm [*Reasenberg*, 1985], implemented in the ZMAP software [*Wiemer*, 2001], using default parameters. The code looks for temporally and spatially dependent events in the catalog (aftershocks occurring after some foreshock) and removes them returning a catalog composed of independent events. Reasenberg [1985] defines an interaction zone as a spatiotemporal window and assumes that if an earthquake occurs within the interaction zone of a previous event it is classified as an aftershock and thus it is removed from the catalog.

After the declustering, only 3473 independent events remain in the catalog. The cumulative number of earthquakes before and after the declustering process is plotted in Figure 4.5a (red and blue lines respectively). Figure 4.7a, instead shows the Schuster spectrum for the declustered catalog. The periodicity at one year seems to disappear in this case, while low *p*-values at large periods indicate that some clusters are still present in the catalog.

The analysis of the relocated seismicity gives different results. A first test on the catalog reveals the presence of clusters, that have to be removed. In this case declustering reduces the number of events from 5005 to 1098 events. The cumulative number of earthquakes before and after the declustering process is plotted in Figure 4.5b (red and blue lines respectively). Figure 4.7b, instead shows the Schuster spectrum for this catalog. The declustering process seems to be well fulfilled, since no low *p*-values are present at high periods except for that at 1 year. This means that an annual periodicity for this catalog is

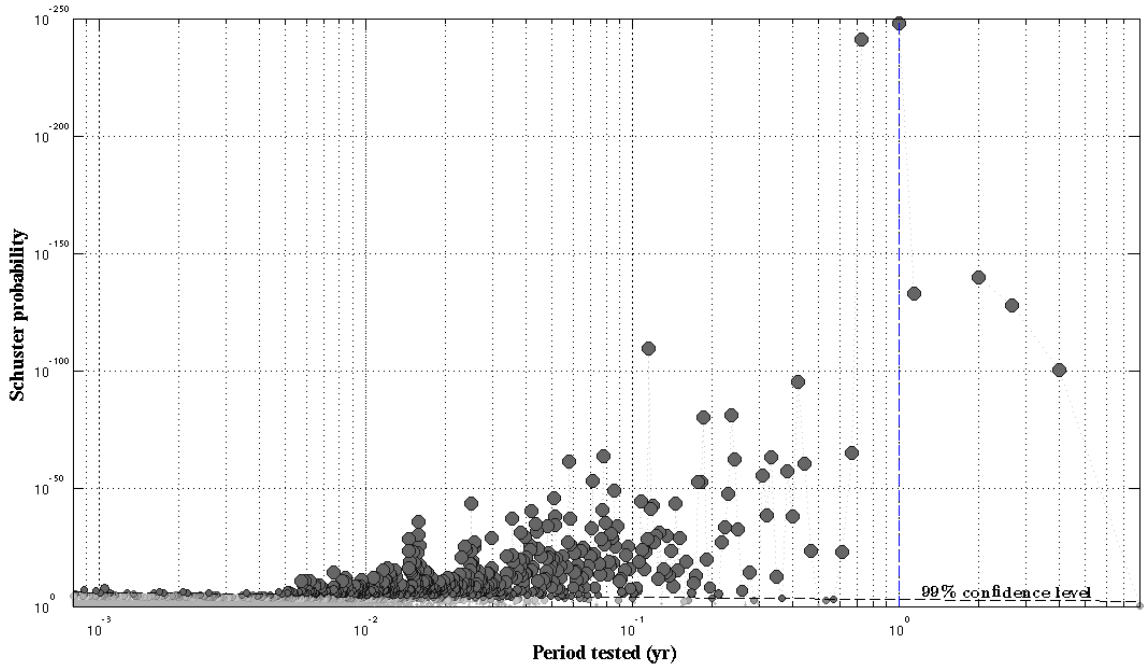


Figure 4.6: Schuster spectrum for the ISIDE catalog. Grey dots indicate the probability values corresponding to each of the period tested. The black dashed line in the bottom identifies the 99% confidence level (0.01 x $\langle \delta_i \rangle$). Dots among this level are darker. The dashed blue line indicates the 1year period.

highly probable. This analysis results in extremely low probability (less than 1%) that seasonality in this catalog is due to chance.

The reason why the analysis of the two catalogs produces different results could be the imprecise location of hypocenters in the ISIDE catalog that might affect the results of the declustering. It could be that earthquakes occurring closely in space are instead localized at a distance greater than the selected interaction zone and thus they are considered as independent. Indeed the percentage of residual events after the declustering is 30% for the ISIDE catalog, while for the catalog of relocated seismicity it is 22%. Figure 4.8 shows epicenters location for both declustered catalogs.

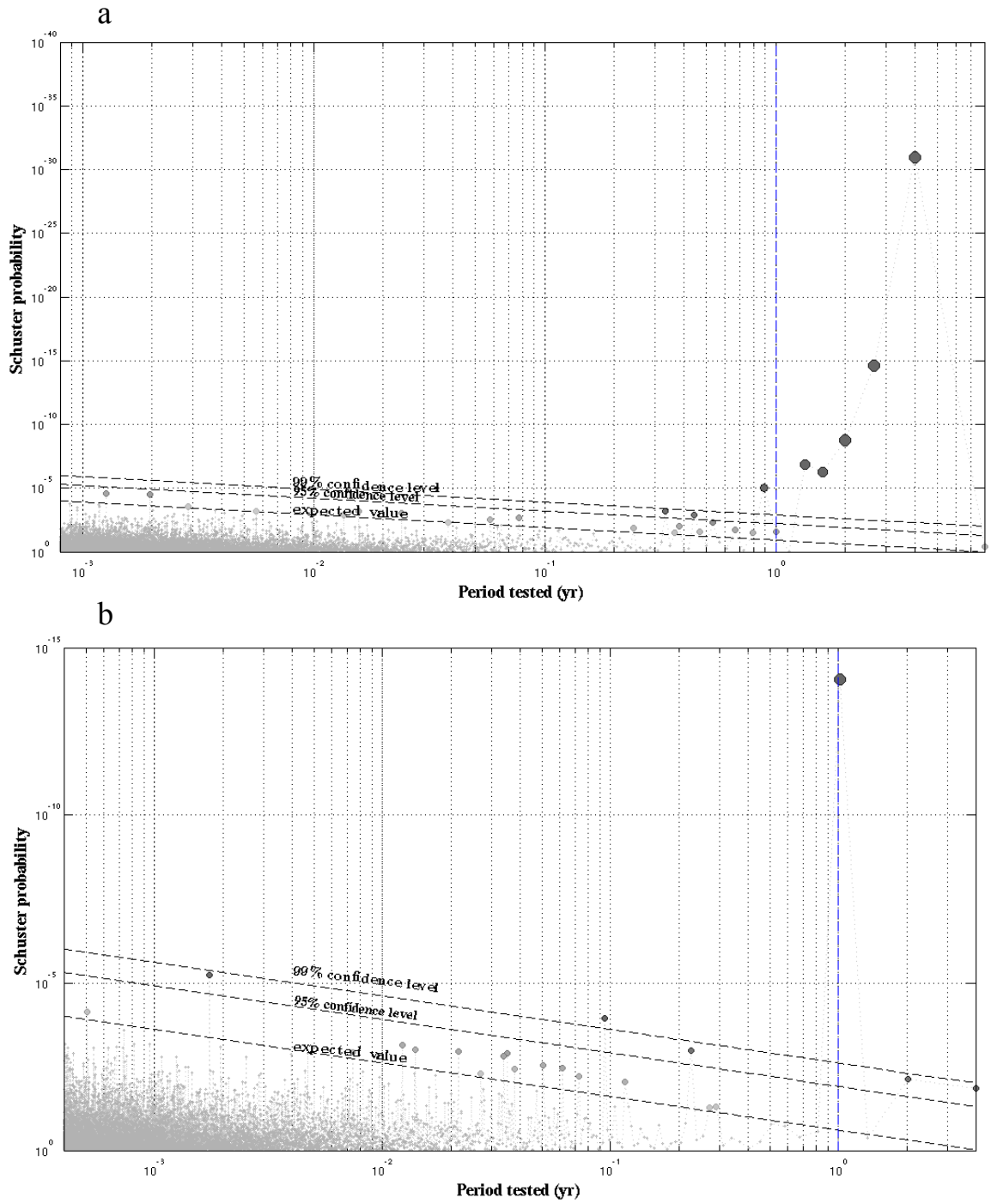


Figure 4.7: Schuster spectrum for the ISIDE catalog (a) and for the catalog of relocated seismicity (b) both declustered. Black dashed lines in the bottom identify the expected value and the 95% and 99% confidence levels, while the dashed blue line indicates the 1year period. Only the catalog of relocated seismicity shows a periodicity at 1 year.

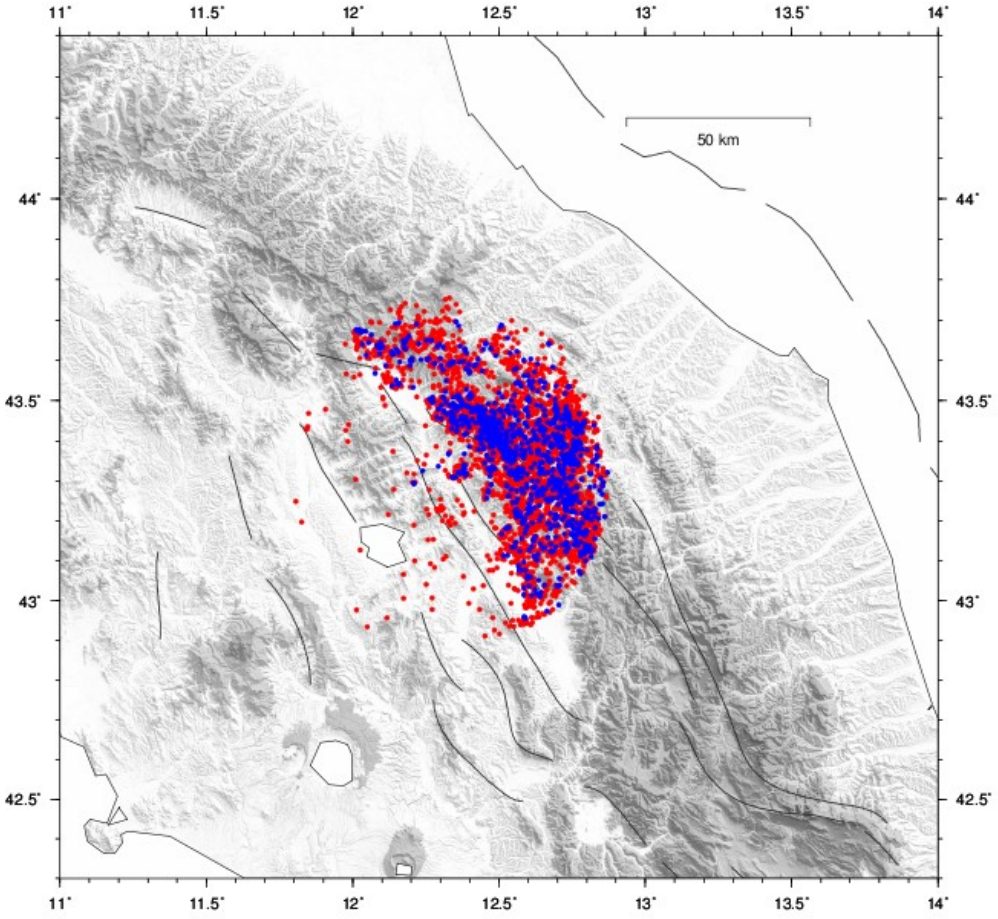


Figure 4.8: Epicenters location of the background events of the ISIDE (red dots) and the relocated (blue dots) catalogs.

Another way to represent the results consists in plotting the 2D walk of the phase angle ϑ_k and verify if the probability that a random walk reaches the distance D is sufficiently low. To do this it is useful to first translate the time series so that it starts from time $t_1 = 0$. The 2D walk (red curve) described by the phases ϑ_k is shown in Figure 4.9 for the period T=1 yr. Since the red curve ends out of the circle of 99% confidence level one can conclude that the probability to have an annual periodicity in the catalog of relocated seismicity exceed the 99%.

Figure 4.10 shows a stacking of the number of earthquakes per week for the catalog of relocated seismicity with superimposed a sinusoidal curve (red) with annual period. The same red curve has been superimposed on the plot of the second independent component of the ICA switched in sign (Figure 4.11) in order to visualize the correlation between the two

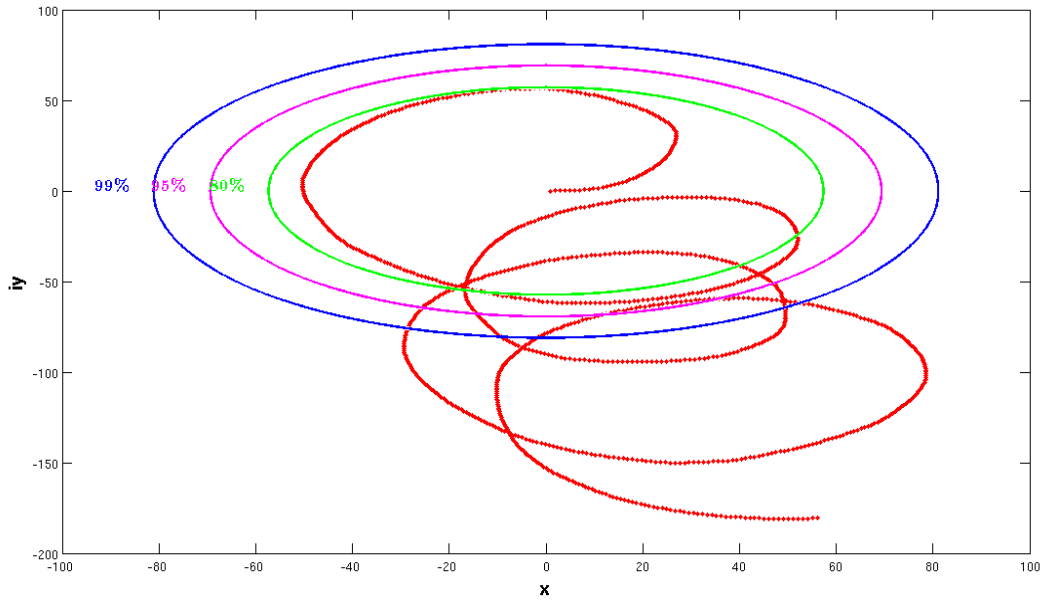


Figure 4.9: 2D walk of the phase angle θ_K, with K=1,…,N (N red dots), calculated for the period of 1 year, represented in the complex plane. The x component of each point of the curve is $D_K \cos\theta_K$, while the y component is $D_K \sin\theta_K$. Green, pink and blue circles represent the 80%, 95% and 99% confidence levels, respectively.

signals. This correlation can be seen also in Figure 4.12 where a stacking of seismicity and of the IC2 over 1 year are superimposed. Figure 4.13 shows a curve (red) of daily mean temperature of the area superimposed on the temporal component of the second IC (blue). Temperature data have been downloaded from the dataset of the ENSEMBLE project (*http://ensembles-eu.metoffice.com*) and processed as described in section 4.1.1 for precipitation data, then daily mean temperature has been centered around y = 0 and rescaled. The correlation between the two signals of Figure 4.13 is 0.82. Therefore, considering the anticorrelation between the temporal component of the IC2 and seismicity (Figures 4.11 and 4.12), it turns out that the highest earthquakes rates are observed during winter, when temperature is lower, while the lowest earthquakes rates take place during summer months, when temperature is higher. In other words, temperature variations seem to modulate earthquakes occurrence in this area. Therefore seismicity seems not to be modulated by hydrological cycle as it is proposed for Nepal by Bettinelli et al. [2007], since the IC1, that correlates with rainfall and the IC2, that correlates with seismic periodicity are out of phase of 90 days.

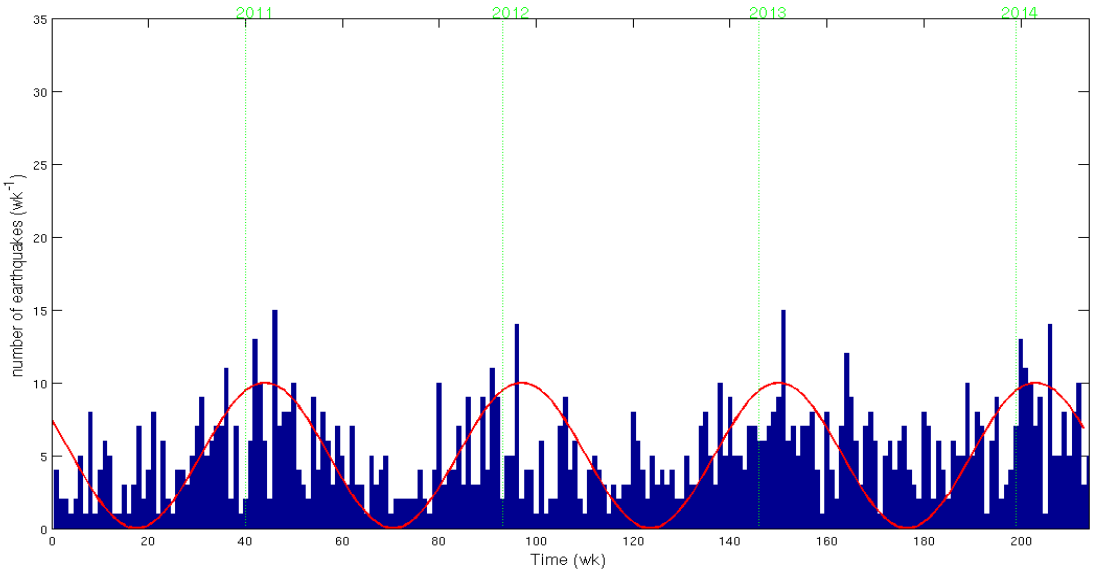


Figure 4.10: Number of earthquakes per week (blue bars) of the catalog of relocated seismicity. A sinusoid with annual period has been superimposed to the histogram in order to emphasize the annual periodicity of seismicity.

As described in Chapter 1, seismicity in this region is mostly due to microearthquakes (M_L<3.0) occurring at high rate. Many studies [e.g. *Gao* et al., 2000; *van Dam* et al., 2001; *Ebel* and *Ben-Zion*, 2014] have demonstrated that microearthquakes

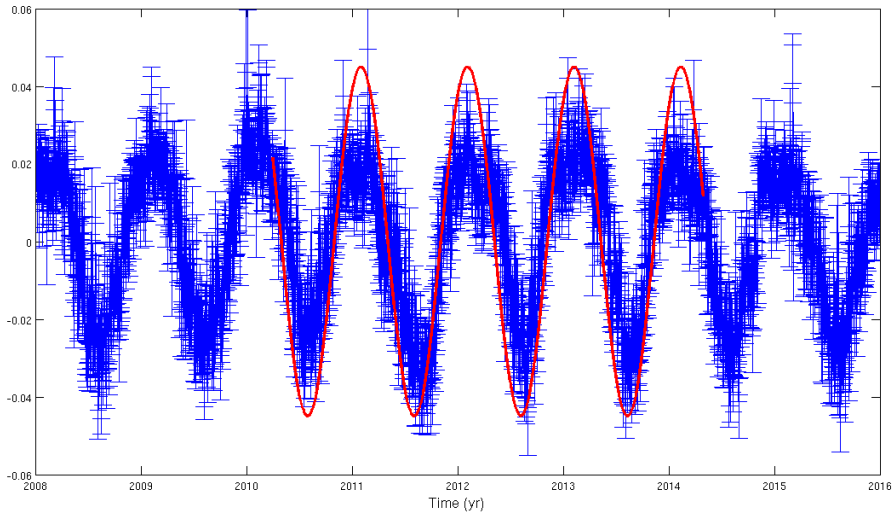


Figure 4.11: Temporal component of the IC2 (blue) switched in sign and sinusoidal curve of seismicity of figure 4.9 (red). This indicates an anticorrelation between the seismic activity and the geodetic displacement associated to the second IC.

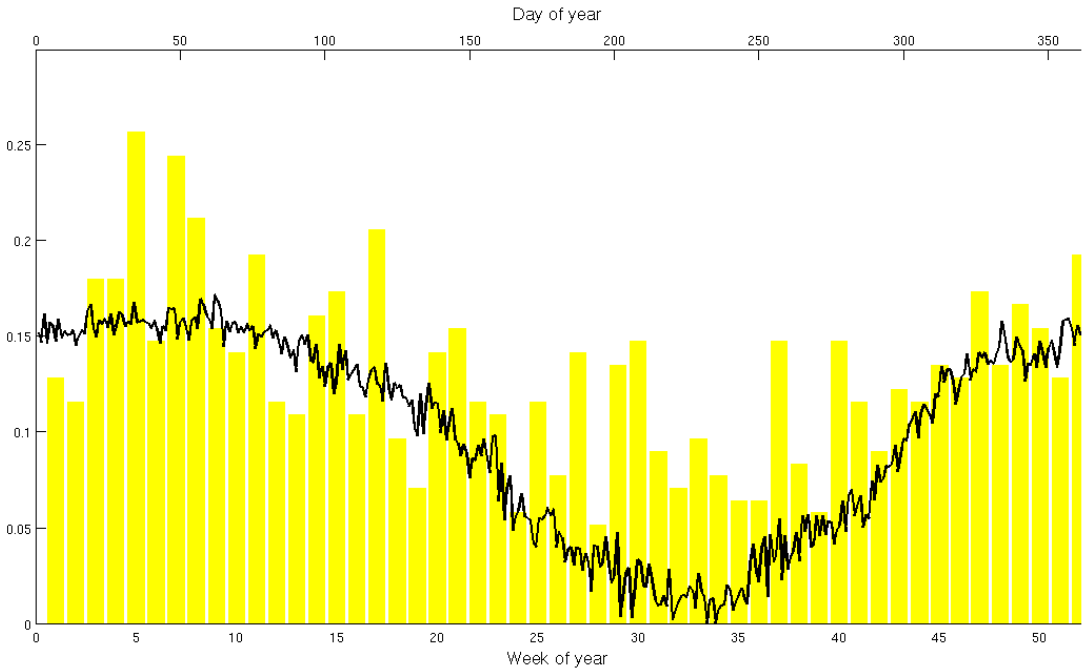


Figure 4.12: Superposition of a stacking of the number of earthquakes per week over 1 year (yellow bars) and of a stacking of the IC2 over 1 year switched in sign and translated along the y axis in order to have only positive values. Both the histogram and the black curve are rescaled.

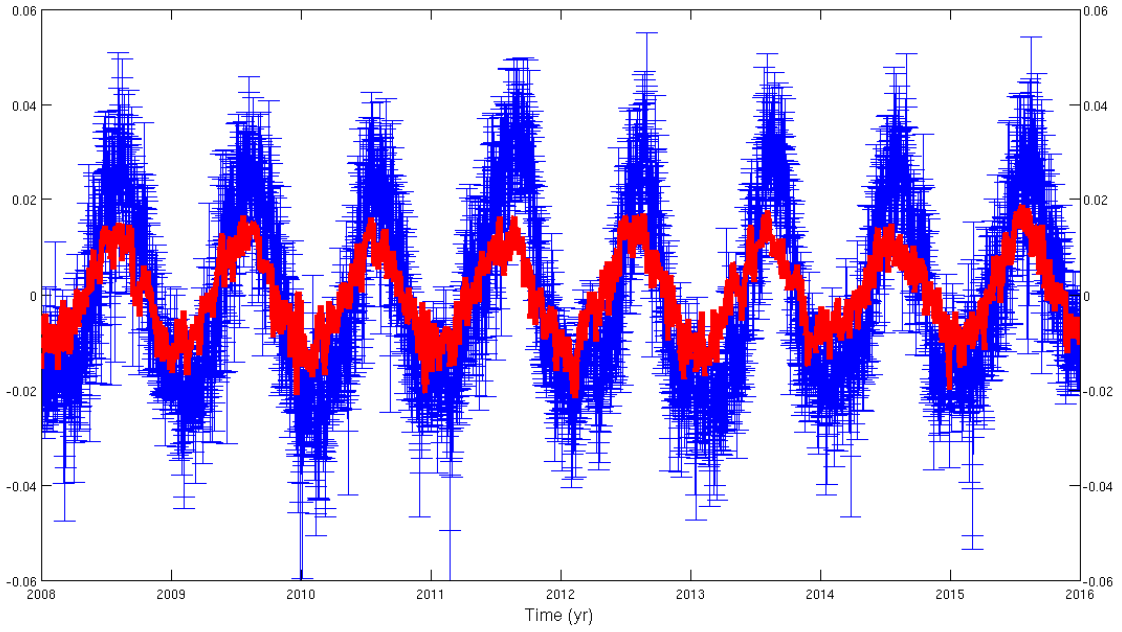


Figure 4.13:Mean temperature curve (red) superimposed on the temporal component of the second IC (blue). Temperature values have been centered along the y axis and rescaled.

may be triggered or modulated by climatic forces with annual period such as atmospheric pressure, hydrology and temperature. This means that stresses induced by those forces are sufficient to fracture near-critically stressed rocks [*Christiansen* et al., 2007]. In particular, Ebel and Ben-Zion (2014) proposed temperature variations to be the cause of annual modulation of seismicity in California and Nevada. In order to evaluate the effects of temperature variations on deformation in the study region, the thermoelastic strain should be calculated, but this is out of the scope of this thesis.

4.3 Fourth component

The IC4 signal is the lowest (in term of S value) among all the components: the SNR is bigger than 1 for only 5 stations (ATBU, ATMI, GUB2, ITFA, MVAL). From figure 4.13 it can be seen that this component exhibits a transient signal in 2014. Transient deformation signals are defined as non periodic, non secular accumulation of strain in the crust [*Riel* et al., 2014]; they can be the result of post-seismic deformations, slow slip events, volcanic deformations, hydrologic deformations or other processes. Large transient motions, such as those due to big earthquakes, can be detected directly from data, while weaker deformations, usually covered by noise, are difficult to be measured.

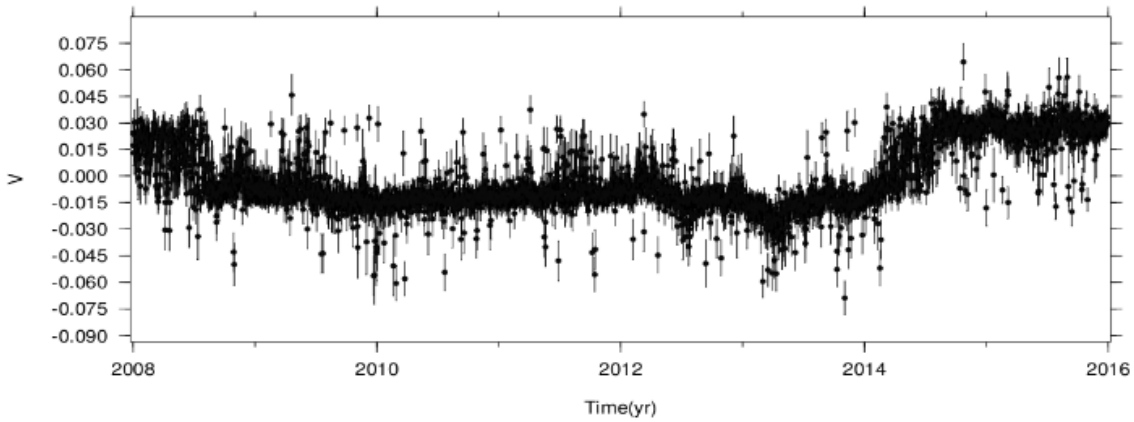


Figure 4.14: Fourth component of the final decomposition.

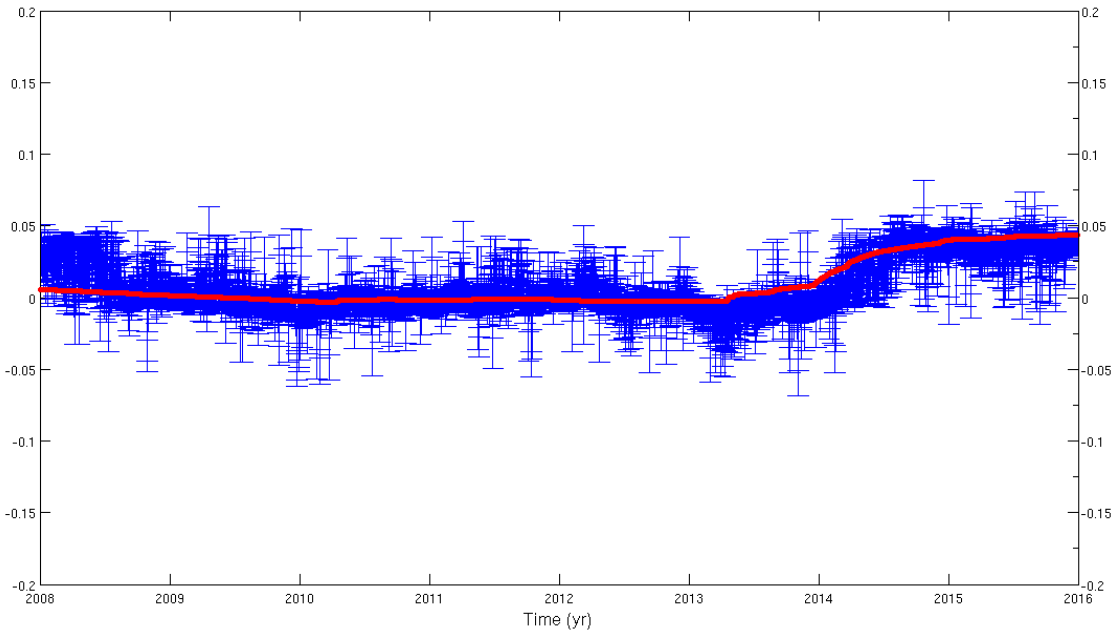


Figure 4.15: Superposition of the temporal component of the fourth IC (blue), plotted with uncertainties bars, and a curve of seismicity obtained subtracting the cumulative number of earthquakes of the declustered ISIDE catalog from the cumulated seismicity of the original catalog. Both curves are rescaled.

Based on the spatial response of this component, showed in Figure 3.5d, the signal seems to describe a stretching along NE-SW direction, that is kinematically coherent with the long-term extensional deformation of the area (*Anderlini* et al., 2016); nevertheless this deformation is probably too small to be observed in the strain analysis reported in Chapter 2.

In order to improve the SNR of this component a further ICA has been performed focusing on an area of 30 km around the point of coordinate (12.4772, 43.4291) and analyzing data on a time span of 3.5 years from July 2012 to December 2015, reducing the percentage of missing data. After that, I have inverted the signal for slip on a predetermined fault geometry, as it will be explained in Chapter 5.

In the next section, the correlation between the fourth IC and the seismicity rate of the study area is presented.

4.3.1 Correlation with seismicity

In this section data from the ISIDE catalog will be used to present a correlation between the seismicity rate and the IC4.

Figure 4.15 shows the superposition of the temporal component of the IC4 and of a curve of seismicity obtained subtracting the cumulative number of earthquakes of the declustered catalog (blue curve in Figure 4.5) from the cumulated seismicity of the original catalog (red curve in Figure 4.5). Both curves have been rescaled in order to be compared. Seismicity seems to reproduce the IC4 trend, even in correspondence of the transient signal in 2014, suggesting a temporal correlation between the two signals. Furthermore, Figure 4.16 shows that the clustered seismicity of this area is mostly located in the hanging wall of the ATF (NE of its trace) and that the stretching process displayed by the spatial response of the IC4 occurs in the same region.

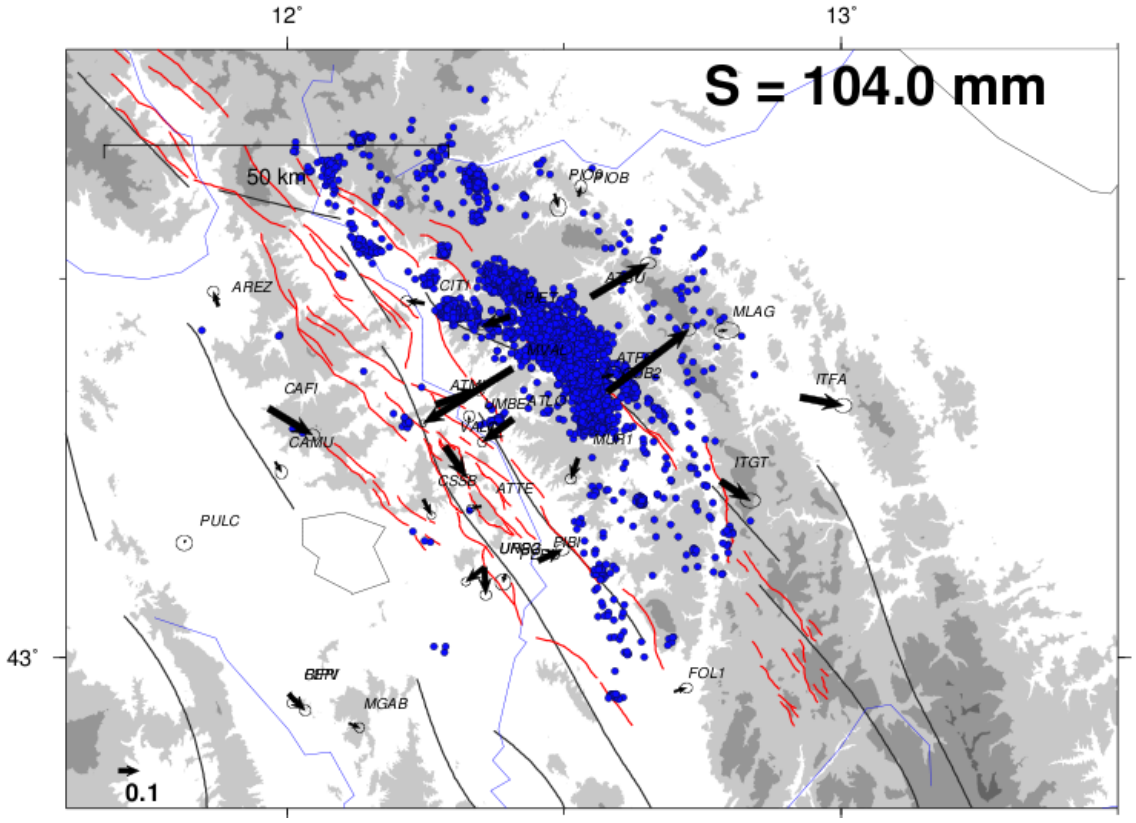


Figure 4.16: Blue dots indicate epicenters location of the clustered earthquakes with magnitude above the detection threshold, while black arrows represent the spatial response of the forth IC as in figure 3.5d. S value is placed on top right of the panel. In red the trace of the Altotiberina fault and of synthetic and antithetic faults obtained from geological survey.

5. Inversion of the fourth independent component

In order to describe a set of observed data, it is useful to implement our physical knowledge into a model that considers the more relevant relations between a finite number of parameters and the data.

The inverse problem theory is a branch of mathematics that is largely applied in geophysics as well as in many other fields (e.g. medical physics, nuclear physics) in order to determine the parameters of a physical model based on a set of observed data and the underling relations existing between the two.

In this chapter I will briefly describe the basic principles of inversion theory and how it is implemented in the ICAIM code. Then I will present the results of an inversion of the spatial part (U) of the fourth IC of the final decomposition in order to find the slip distribution on faults with known geometry. I will finally discuss some of the implications of my findings.

5.1 Inversion theory

The study of a physical system follows a three step procedure [*Tarantola*, 2005]:

- parameterization of the system: consists in finding a minimal set of model parameters that completely describe the system. A parameterization of the system is

represented as a point in the model space *M*;

- forward modeling: discover some physical laws that allow us to predict the observed data when model parameters are known, by means of the equation:

$$d = G(m) \tag{5.1}$$

where *d* is the data vector and represents a point in the data space *D*, *m* is the vector of model parameters and *G* is called the *forward operator* expressing the mathematical model of the physical system under study. If the spaces of the model *M* and data *D* are both linear, equation 5.1 becomes linear and can be written as:

$$d = Gm \tag{5.2}$$

- inverse modeling: consists in predicting model parameters based on a set of data or observations, by inverting equation 5.2.

  The inverse problem is classified as over-determined, well-determined or under-determined, if the number of unknown model parameters to determine is lower, equal or higher than the number of observed data. The inversion of geodetic data is often a mixed-determined problem, where some model parameters can be well constrained, while for others insufficient input information is available.

  The simplest case is the well-determined one: in this case the problem has an exact solution.

  In the over-determined case, the least-squares formulation allows us to find an approximate solution by minimizing the distance between *d* and *Gm*. In other words, the best solution is the one that minimizes a cost function $S(m)$ defined as the L2 norm of the difference between observed and reconstructed data:

$$S(m) = \|d - Gm\|^2 = \frac{1}{2}[(d - Gm)^T C_d^{-1}(d - Gm)] \tag{5.3}$$

where C_d is the covariance matrix for the data. The least square solution of the minimization of $S(m)$ is the best estimated model $\tilde{m}$:

$$\widetilde{m} = (G^T C_d^{-1} G)^{-1} G^T C_d^{-1} d \tag{5.4}$$

The covariance matrix for such a model is:

$$\widetilde{C_m} = (G^T C_d^{-1} G)^{-1} \tag{5.5}$$

In the last case (under-determined), when the number of data is lower than the number of parameters to be estimated, there are infinite possible solutions for the minimization of $S(m)$. Therefore, further (a priori) information should be added in order to infer the values of the model parameters. This process is called regularization of the problem and it allows passing from an under-determined to an over-determined problem, which can be solved with a least-squares method. The regularization approach is used also when the problem is mixed-determined in order to add extra information to infer the under-determined parameters of the problem.

5.2 Inversion in ICAIM

The theory of linear elasticity allows us to predict the theoretical surface displacements produced by a certain slip distribution on a given fault at depth. The cumulative amount of fault slip can be inferred by inverting the displacement occurred in given time interval. In order to compute the temporal evolution of fault slip it is thus necessary to invert the displacement measured between every two consecutive epochs. This procedure, schematically presented in the left-hand track of Figure 5.1, is very intense from a computational point of view. In Figure 5.1 the ICAIM procedure flows clockwise from top to bottom left. Once a proper independent components decomposition of the initial dataset is obtained, the code allows one or more IC to be inverted on predetermined fault geometries in order to find a possible slip distribution, under the assumption that the medium surrounding the faults is elastic. Each fault is subdivided into a certain number of

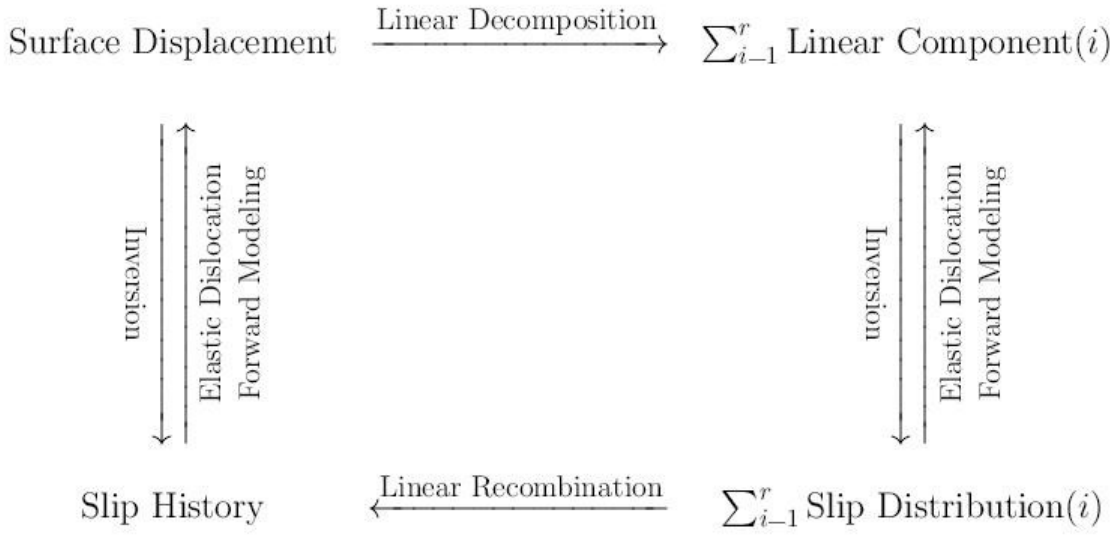


Figure 5.1: Schematic for PCAIM and ICAIM [Kositsky and Avouac, 2010].

rectangular patches identified by the *x, y, z* coordinates of their centers, their strike and dip angles, and their length and width.

The slip inversion is performed on the spatial part (U) of the ICs. The direct problem in this case is:

$$U = GL \tag{5.6}$$

where *G* denotes the Green's functions, computed from the semi-analytical solutions of Okada [1985] (Poisson's ratio ν equal to 0.25), relating surface displacements with fault slip at depth, and the matrix *L* contains the strike-slip and dip-slip for each patch. Since the inversion of equation 5.6 is often an ill-posed problem, i.e. there are many plausible solutions, the introduction of some constrains is necessary to regularize the problem. The a-posteriori slip model and the corresponding covariance matrix are calculated following the approach of Tarantola [2005] for linear inverse problems:

$$m = m_0 + C_{m_0} G^T \left(G C_{m_0} G^T + C_d^{-1}\right)^{-1} (d - G m_0) \tag{5.7}$$

$$C_m = C_{m_0} - C_{m_0} G^T \left(G C_{m_0} G^T + C_d\right)^{-1} G C_{m_0} \tag{5.8}$$

where m_0 is the a priori model and C_{m_0} is the a priori covariance matrix of the model parameters. The elements of the C_{m_0} matrix are chosen following Radiguet et al. [2011]:

$$C_{m_0}(i,j) = \left(\sigma_m \frac{\lambda_0}{\lambda}\right)^2 e^{\left(-\frac{d(i,j)}{\lambda}\right)} \tag{5.9}$$

where σ_m is the a priori standard deviation of the model parameters, λ_0 corresponds to the square root of the area of a patch, λ fixes the distance over which parameters are correlated, and $d(i,j)$ is the distance between the centers of two patches *i* and *j*.

Based on equation 5.6, the spatial component (U) of each GPS station can be reconstructed as the linear sum of the contribution of all patches along the strike and dip directions.

5.3 Results

As seen in Chapter 3 (Figure 3.5d) the spatial response of IC4 reproduces extension across the Gubbio fault system. In order to focus on the signal of interest, a new decomposition has been performed, selecting stations located within 30 km from the point of coordinate (12.4772, 43.4291) (mean point between the stations ATBU and MVAL) and analyzing horizontal GPS time series in a time span of 3.5 years, from July 2012 to December 2015.

The ARD criterion suggests to consider a decomposition in 4 ICs. Figure 5.2 shows the temporal and spatial components of the final decomposition obtained choosing the combination of priors that maximizes the NFE. In this decomposition the IC that reproduces the transient signal at 2014 is the second one. In Figure 5.3 a temporal filtering of the IC2 with a moving window of 30 days is superimposed on the original signal. Figure 5.4 instead, shows the horizontal time series of the MVAL station, as an example of a station in which the transient signal at 2014 is visible.

In this analysis I used 3 different rectangular faults, obtained from geological and seismological evidences [*Pucci* et al., 2003; *Moretti* et al., 2009] that represent the Altotiberina fault (ATF), and a synthetic-antithetic fault system located in the hanging wall

of the ATF. Table 5.1 reports fault dimensions and their orientation in space (strike and dip angles).

I have performed inversions over the three faults separately and over different

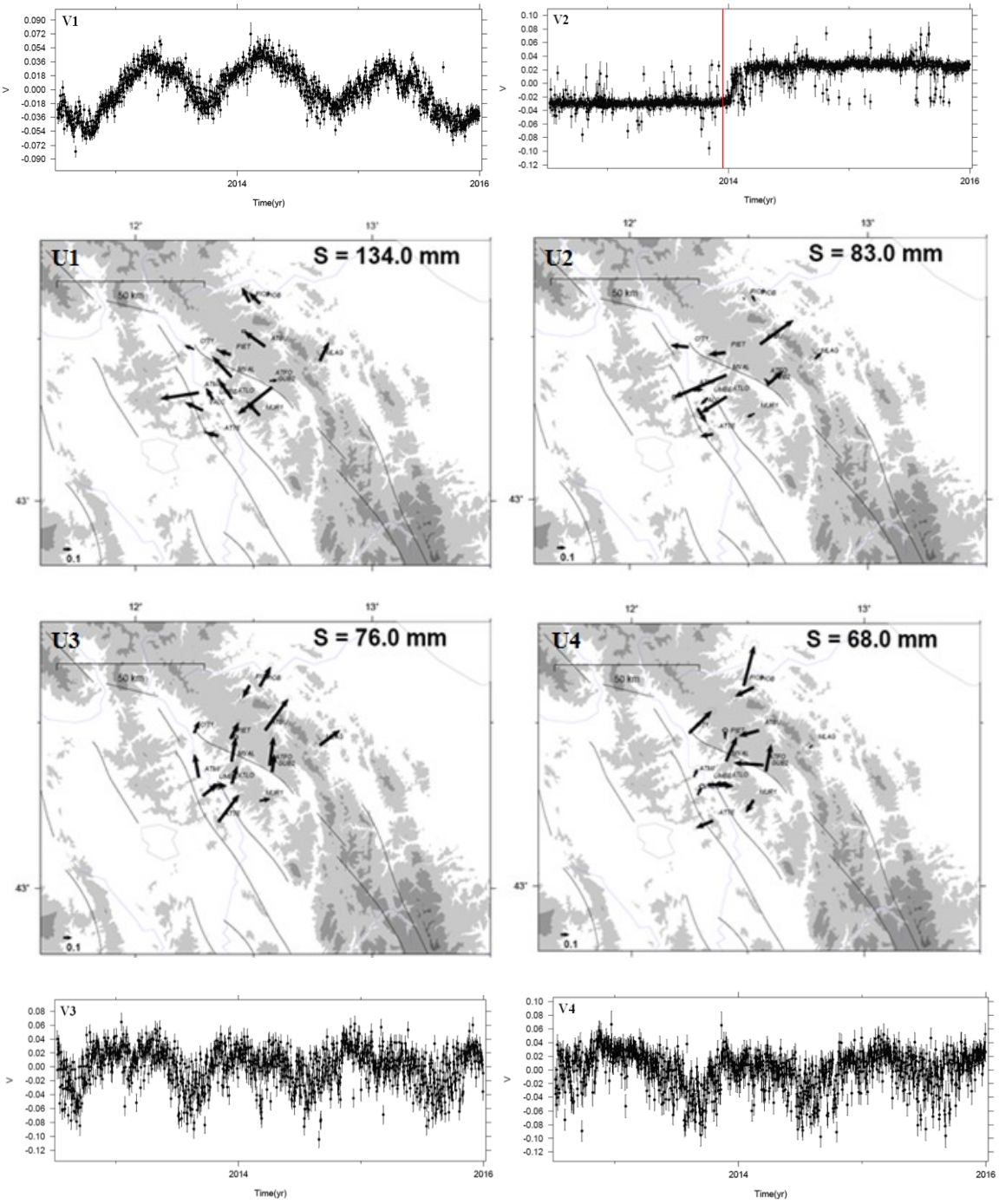


Figure 5.2: Independent components obtained analyzing GPS time series of 15 stations located within 30 km from the point (12.4772, 43.429) from July 2012 to December 2015.The red line in the second panel individuate the epoch (22-12-2013) of the largest event (M_w = 3.9).

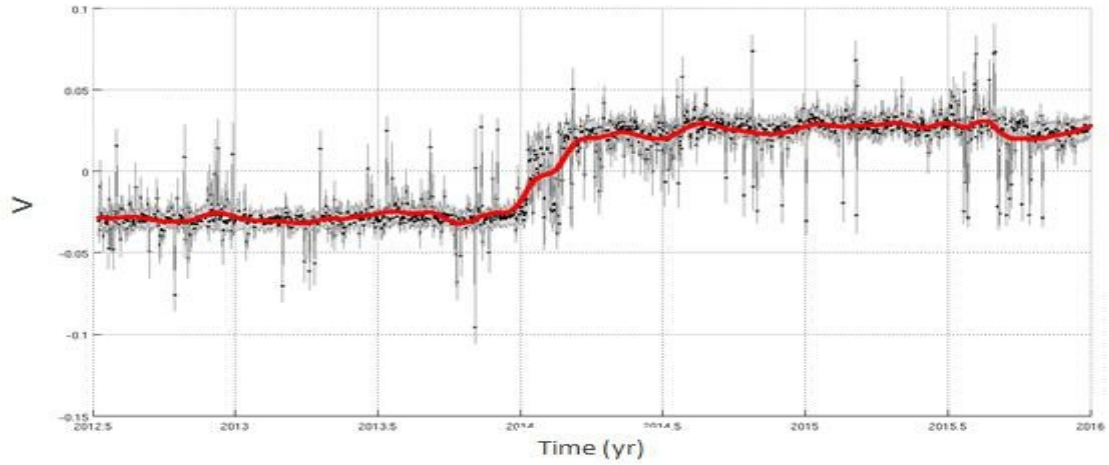


Figure 5.3: Temporal component of the IC2 (black dots) with error bars. The red line is a filtering obtained using a moving window of 30 days.

combinations of them. I have also tried to subdivide the fault planes in a different number of patches.

First of all I have analyzed a simple case in which both the synthetic and antithetic faults are composed by a single patch of dimension 40 x 6 km, while the ATF is

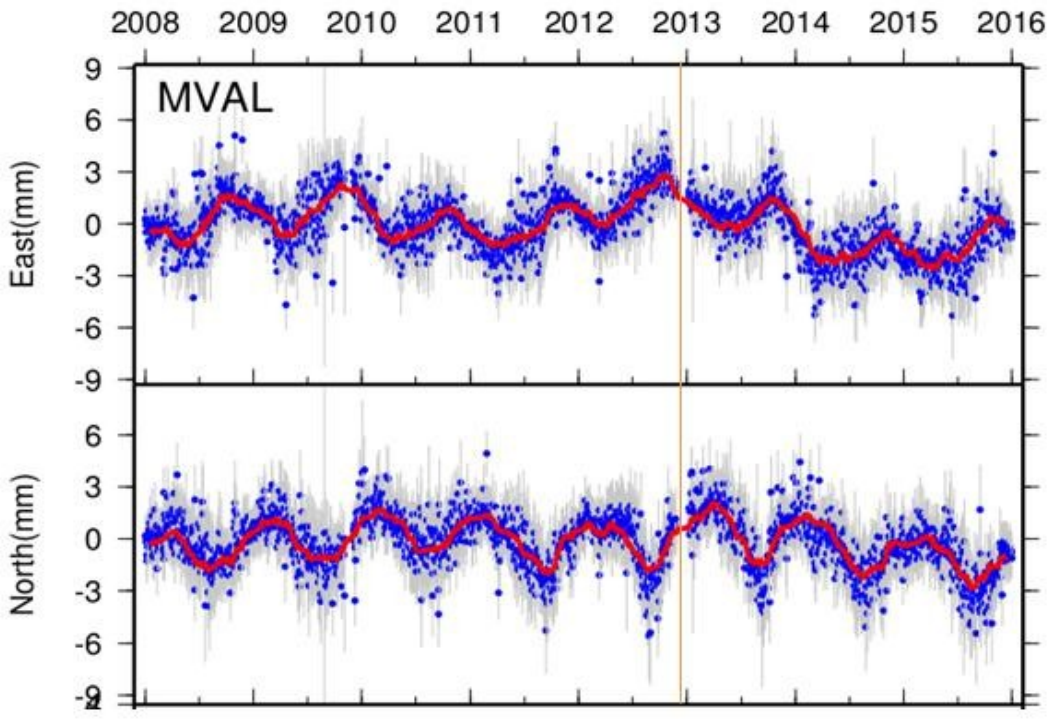


Figure 5.4: Horizontal components of the time series of station MVAL. Blue dots indicate surface displacement at each epoch, while the red curve is the reconstruction based on the ICA. The orange line individuate the occurrence epoch (22-12-2013) of the largest event (M_w = 3.9).

	Length (km)	Width (km)	Top depth (km)	Bottom depth (km)	Lon Lat top northern corner (°)	Lon Lat top southern corner (°)	Strike angle (°)	Dip angle (°)
ATF	72	50	0.06	13	11.9854 43.4946	12.5025 42.9670	324.8	15
Synthetic	40	6	0.40	5	12.3187 43.5269	12.6045 43.2326	324.8	50
Antithetic	40	6	0.40	5	12.3967 43.5670	12.6824 43.2723	144.8	50

Table 5.1: Dimensions and orientation in space of the three faults (ATF, synthetic and antithetic).

subdivided into 2 patches along the strike direction and 8 patches along the dip direction, so that the dimension of each patch (~ 36 x 6 km) is almost equal to the size of the other faults. Knowing that all the faults behave as normal faults, I have constrained the positivity of the solution and the slip to be only along the dip direction (fixed rake). In this case, the number of model parameters (equal to the number of patches) is lower than the number of data (number of time series = 2 x 15) even considering all the faults in the inversion. This means that a least square solution can be obtained without introducing any regularization. The best solution for each fault configuration is the one that best fits the original data. In order to understand which combination of faults gives the best solution I have applied an F-test. This statistical test allows us to determine if a more complex model provides a significantly better fit to the data than a simpler one. The F value is calculated as [e.g. *Press* et al., 2002]:

$$F = \frac{\chi_1^2 - \chi_2^2}{\chi_2^2} \frac{d - p_2}{p_2 - p_1} \tag{5.10}$$

where χ_1^2 and χ_2^2 are the chi squares of the simpler and the more complex model, respectively, and p_1 and p_2 are the number of model parameters for the two models. The null hypothesis that the fit is not significantly improved considering a model with a higher number of parameters is rejected if the value of *F* calculated from the data, i.e. provided by

5.10, is greater than a critical value of the *F*-distribution for a given significance level, here assumed to be 5%.

The *F*-test suggests as best model the one that considers a slip distribution only on the synthetic and antithetic faults (see Table 5.2). From now on, I will refer to it as model 1. Figure 5.5 shows the slip distribution over the two faults and the fitting between original (red arrows) and reconstructed (blue arrows) data.

A second cycle of inversions has been performed using the same faults but reducing the patch dimensions (2 x 2 km^2): the ATF has been subdivided into 36 patches along the strike direction and 25 patches along the dip direction, while the synthetic and antithetic faults have been subdivided into 20 and 3 patches along the strike and dip directions, respectively. The aim of this test is to see where the slip is localized along the faults. In this case, since the number of model parameters is higher than the number of data, a regularization is required. I have forced the positivity (normal motion) of the solution and selected as a priori model the one that assigns null slip on the fault, as in Radiguet et al. [2011], in order to avoid large slips in regions with low data-coverage. I have tested two different values of σ_m (0.1, 0.2 m) while λ was considered in the range from 0.1 to 15 km. The best model for each fault configuration was selected on the basis of the L curve criterion [*Hansen*, 1992]. The selected model corresponds to the best compromise between slip roughness and misfit with respect to the data [*Radiguet* et al. 2011]. In this case, an *F*-test cannot be used to choose the best fault configuration, because the number of data is lower than the number of model parameters. Therefore, I have compared models based on

	Antithetic vs Synthetic + Antithetic	Synthetic vs Synthetic + Antithetic	Synthetic + Antithetic vs Synthetic + Antithetic + ATF
F calculated	2.59×10^4	2.59×10^4	-0.02
F-distribution	0.22	0.22	0.38

Table 5.2: Calculated F values and values of the F-distribution obtained when comparing the misfit of an inversion performed with one antithetic or synthetic fault alone and an inversion performed considering both faults (first two columns). In the last column the fit to the data of an inversion over a synthetic-antithetic fault system is compared with the one obtained considering the synthetic, antithetic faults and the ATF together.

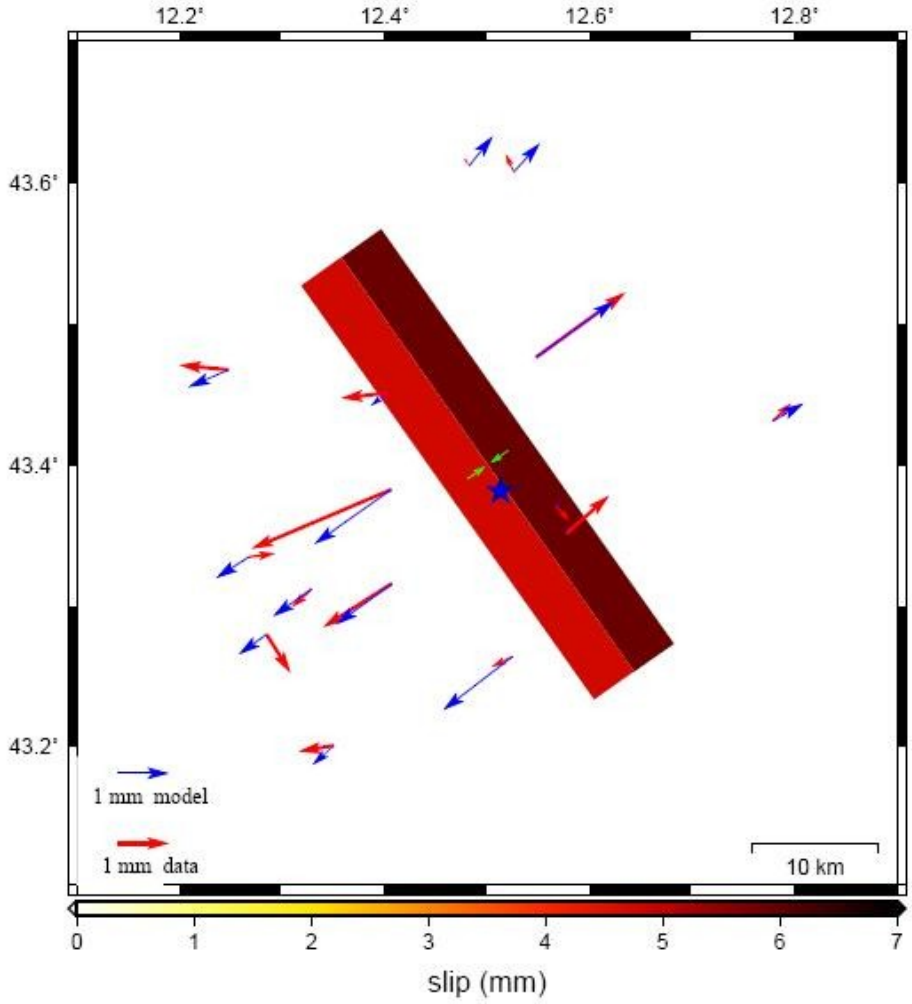


Figure 5.5: Slip distribution over a synthetic and an antithetic fault composed by a single patch each (model 1). Green vectors indicate the fixed rake direction. Blue and red arrows represent reconstructed and original data, respectively, while the blue star individuates the epicenter location of the largest event (Mw = 3.9).

their fit to the data. One synthetic or antithetic fault alone cannot explain data as well as if they are considered together, as we have verified. The χ^2 of the inversion performed over a synthetic or antithetic fault are 3194.55 and 3580.56, respectively, while the χ^2 obtained considering the two faults together is 1676.30.

Figure 5.6 shows the best model (model 2) obtained performing an inversion over both synthetic and antithetic faults. The maximum slip is localized in the shallowest part of the synthetic fault. Figure 5.7 instead shows the result of the inversion performed over the three faults. The fit between original (red arrows) and reconstructed (blue arrows) data is reported in both Figures 5.6 and 5.7. As one can expect, the model obtained considering

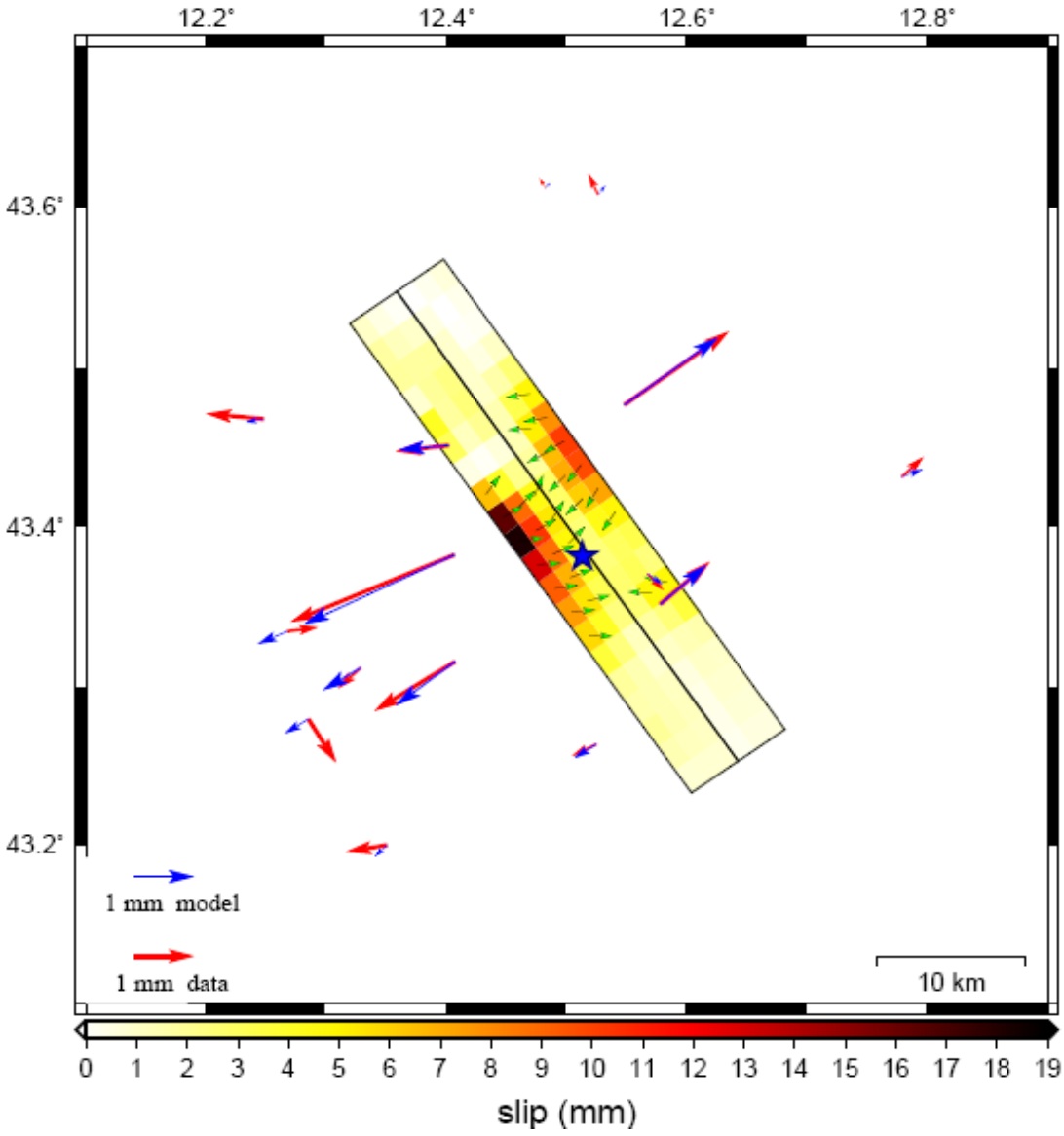


Figure 5.6: Slip distribution for model 2. Green vectors indicate the rake direction. Blue and red arrows represent reconstructed and original data, respectively, while the blue star individuate the epicenter location of the largest event (M_w = 3.9).

three faults in the inversion is the one that best fits the data, being able to reproduce even the displacement of stations ATMI and VALC. However, it can be seen from Figure 5.7 that some strike-slip is introduced in the ATF plane beneath the two stations. This solution is thus discarded being in contrast with the typical behavior of a normal fault and the preferred model is model 2.

I have calculated the seismic moment associated to the slip distribution over the synthetic and antithetic faults in both models 1 and 2 and compared it to the seismic

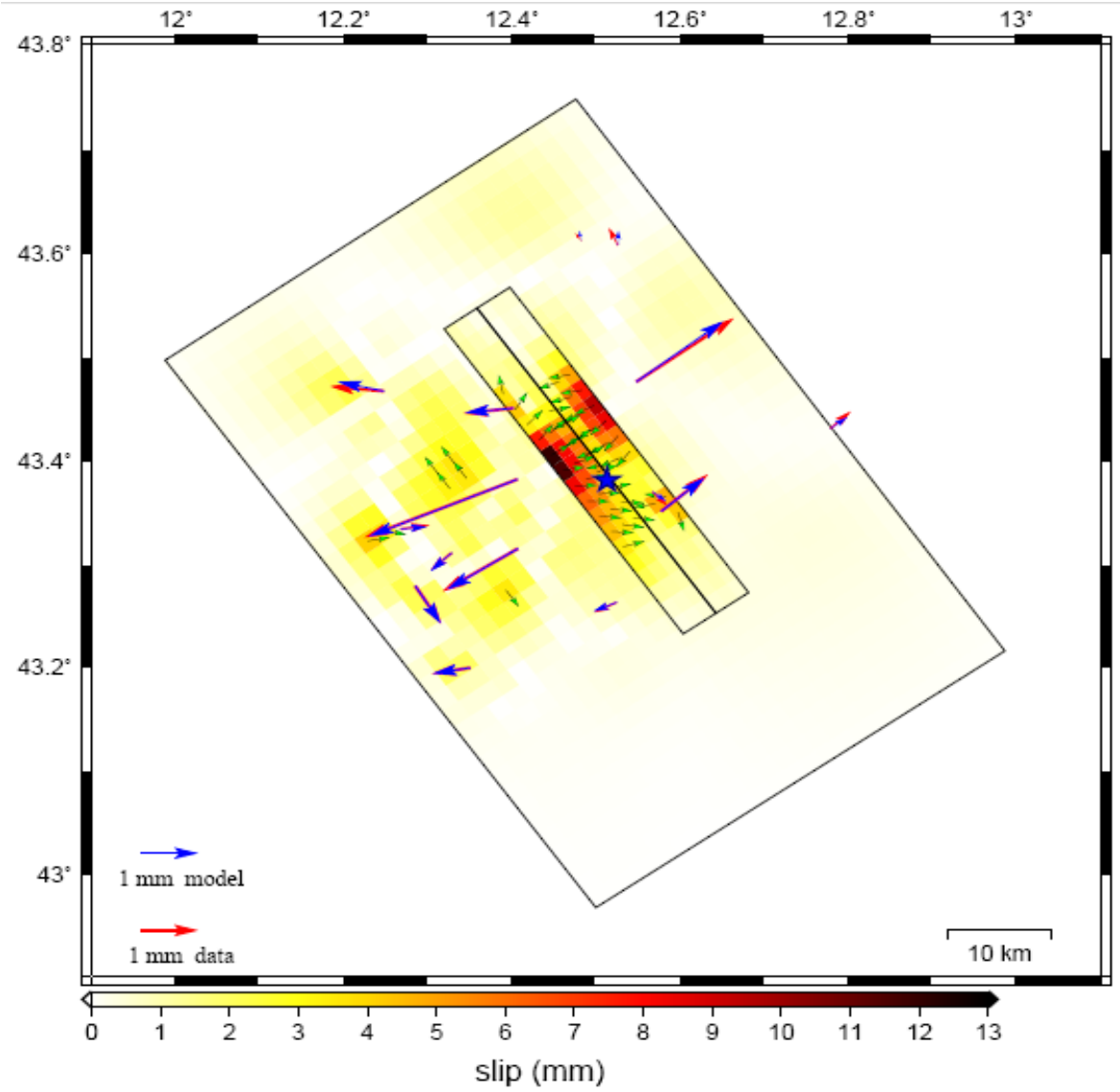


Figure 5.7: Slip distribution over the ATF, a synthetic and an antithetic fault. Green vectors indicate the rake direction. Blue and red arrows represent reconstructed and original data, respectively, while the blue star individuate the epicenter location of the largest event (M_w = 3.9).

moment released by clustered earthquakes occurred in the same time span and in the same region considered in the analysis (see Figure 5.8), including the largest event (M_W = 3.9) of a 2014 sequence. Local magnitudes M_L of events in the ISIDE catalog have been converted into moment magnitudes M_W using the equation (valid for $M_L \leq 4$) deduced for the seismic catalog of the ATF region in Munafò et al. [2016]:

$$M_W = \frac{2}{3} M_L + 1.15 \qquad (5.11)$$

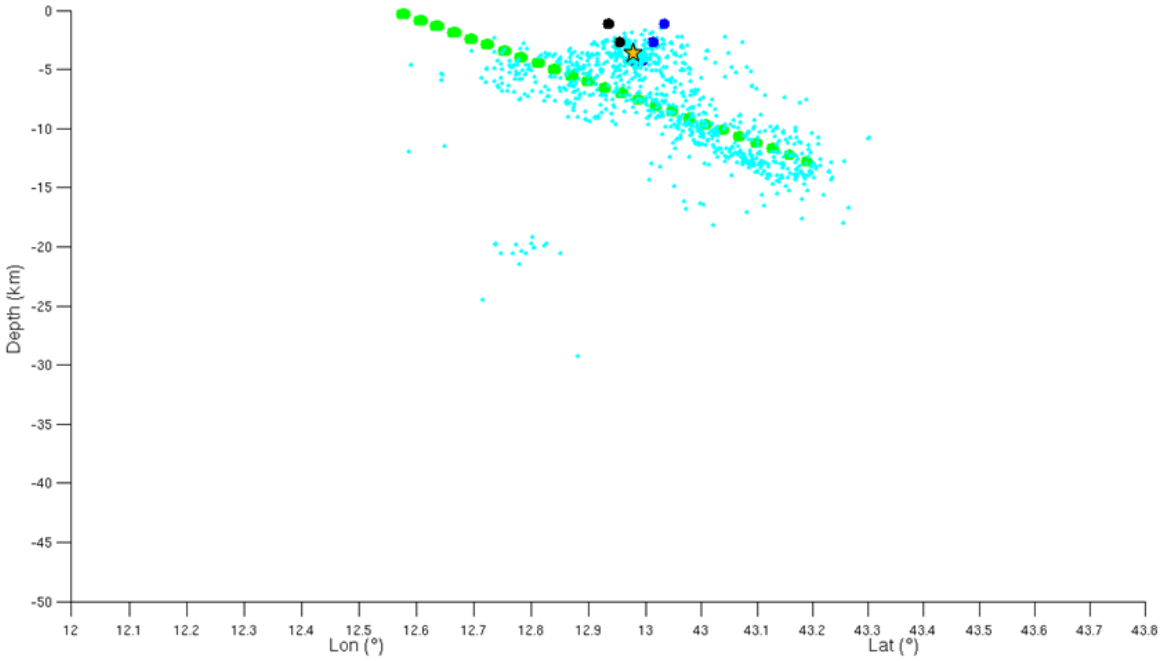


Figure 5.8: Cross section along the NE-SW direction, showing the three faults and the hypocenter location of clustered earthquakes (cyan dots). Green dots indicate the ATF, while black and blue dots indicate the synthetic and antithetic fault, respectively. The yellow star is the hypocenter location of the largest event (Mw = 3.9).

Figure 5.9 shows the trend of the cumulated seismic moment deduced from the inversion (black dots) for models 1 and 2 and the one released by clustered earthquakes (green line), assuming a shear modulus μ = 30 GPa. The red (filtering of the cumulated seismic moment deduced from geodetic measurements) and green curves show the same temporal evolution but differ in their final amplitudes., the equivalent moment magnitude for model 1 and 2 is M_W 5.18 and M_W 5.01, respectively. The sequence of earthquakes instead, has an equivalent magnitude of M_W 4.98. In these calculations, uncertainties in the shear modulus, in patches dimensions and in slip values have not been taken into account when determining the seismic moment related to model 1 and 2. Uncertainties in the magnitudes of earthquakes of the ISIDE catalog have been neglected too.

In conclusion, the inversion of the second IC of the new decomposition results in a slip distribution over the synthetic and antithetic faults that could be both seismic and aseismic. Model 1 suggests that the cumulated seismic moment released by clustered earthquakes can explain only part of the cumulated seismic moment deduced from the

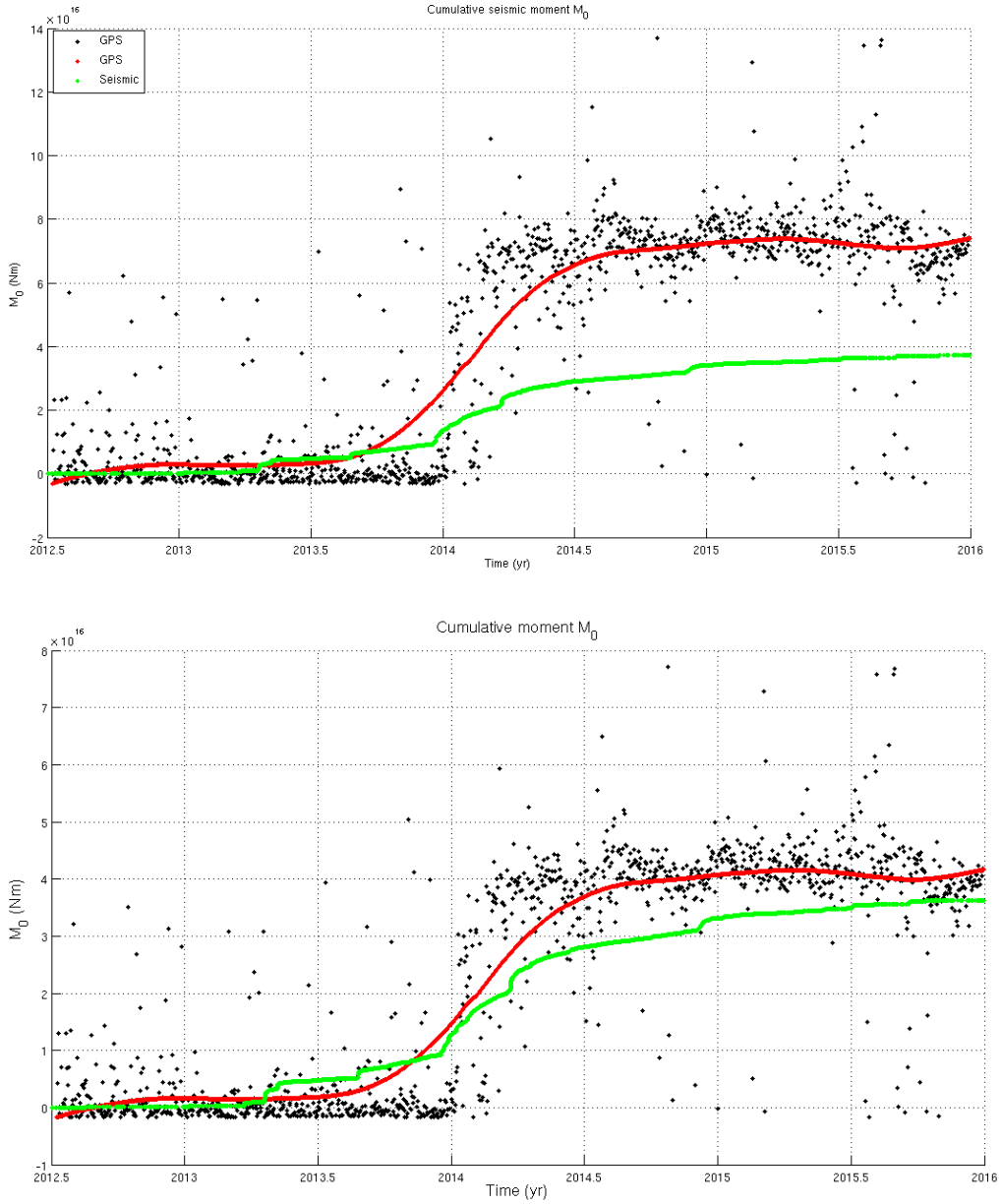


Figure 5.10: Superposition of the cumulated seismic moment deduced from the inversion (black dots) and the cumulated seismic moment calculated from the magnitude of the earthquakes in the catalog (green line) for model 1 (a) and 2 (b). The difference between the red line (filtering of the cumulated seismic moment deduced from the inversion) and the green line is of about 3.7x10^{16} Nm for model 1 and 0.5 x10^{16} Nm for model 2.

inversion, so that in this case, the contribution of earthquakes equals that of the aseismic motion. Model 2 instead suggests that the majority of slip can be related to the seismic sequence of 2014, while the remaining part can be linked to aseismic motion.

Conclusions

In this thesis, an Independent Component Analysis (ICA) technique has been used to study surface displacement time series of 30 Global Positioning System (GPS) stations, located in the Upper Tiber Valley, in central Italy. In this region, the existence of a low angle (15°-25° east-dipping) normal fault (Altotiberina fault, ATF) has been documented in geological [e.g., *Brozzetti* 1995, *Brozzetti et al*., 2000] and geophysical data [e.g., *Mirabella et al*., 2011]. The fault geometry is well imaged by the associated microseismicity. A multidisciplinary research infrastructure (TABOO, *The Altotiberina Near Fault Observatory*) was built in 2010 to monitor the ATF with a multidisciplinary approach and is now one of the Near Fault Observatories (NFO) of the European Plate Observing System (EPOS). The geophysical interest for this area is mainly due to the fact that micro-earthquakes (M_L<3.0) have been observed to continuously nucleate along the ATF plane [*Chiaraluce et al*., 2007, *De Luca et al*., 2009], despite the fact that Anderson-Byerlee [1951] theory precludes normal fault from slipping at low (<30°) dip angles. Furthermore, based on the size of the fault, an earthquake activating the entire ATF should be around M_W 7 [*Chiaraluce et al*., 2014]; however, historical catalogs do not contemplate the occurrence of such an event in the past 2000 years. The detection of time-dependent crustal deformation, transient deformation or deformation induced by non-tectonic processes is of fundamental importance, and is now possible also in region with low tectonic strain rates thanks to high resolution geophysical networks.

ICA is a statistical tool used for solving Blind Source Separation (BSS) problems: it allows us to detect source signals from a mixing of recorded signals. The Independent Component Analysis-based Inversion Method (ICAIM) code [*Gualandi*, 2015], has been

used to find the best independent-component decomposition for the dataset of GPS time series. The final decomposition has been obtained analyzing the horizontal components of GPS time series, de-trended for a constant velocity term. The proper number of ICs has been chosen based on the Automatic Relevance Determination (ARD) criterion, comparing the variance associated to each component. In this way we obtained four independent components.

Hypothesis on the sources of each IC (except for the third one that seems to mainly reproduce noise) have been proposed, based on some observed temporal correlations. The first IC (IC1), that exhibits an annual periodicity, is found to correlate with a 190-days rainfall accumulation curve. The second IC (IC2) is characterized by a 1 year periodicity too, but it is out of phase of almost 90 days compared to the IC1. A correlation between IC2 and a time series of mean air temperature of the region has been pointed out. Furthermore, an analysis of the seismic catalog of the area, performed by means of a Schuster test, revealed that the seismicity rate exhibits an annual periodicity, that is in phase opposition with the mean temperature trend: the number of earthquakes is higher during the winter than during the summer. Therefore, it seems that temperature variations are responsible for surface displacement measured by GPS and that they regulate the seismicity rate. Finally, the fourth IC (IC4), consists in a monotonic signal whose non-vanishing rate is limited to a period of 7-8 months (transient signal) from December 2013 to June 2014. A correlation between the IC4 and the cumulative number of clustered earthquakes occurred in the region has been pointed out. The fact that this component could represent a tectonic motion is suggested also by its spatial part, that reproduces extension across the Gubbio fault system. For this reason, an inversion of the IC4 has been performed in order to infer if slip on faults located in the hanging wall of the ATF may explain the observed deformation transient. Three rectangular fault planes, representing the ATF and a synthetic-antithetic fault system, have been used for the inversion. The preferred model is the one that accounts only for slip on the synthetic and antithetic faults, suggesting that the IC4 does not reproduce surface displacement related to slip on the ATF plane. Furthermore, I found that the seismic moment released by the earthquakes is lower compared to the cumulated seismic moment deduced from the IC4 inversion, suggesting that a certain amount of slip should be related to aseismic motion. Despite the low amplitude of this transient signal (lower than 3 mm in the horizontal components), the

approach here used is able to capture it, thanks to the dense spatio-temporal distribution of available geodetic data.

Bibliography

Abers G.A., Mutter C.Z., Fang J., 1997 - Shallow dips of normal faults during rapid extension: earthquakes in the Woodlark-D'Entrecasteaux rift system, Papua New Guinea., *Journal of Geophysical Research*, 102, 15301–15317.

Ader T. J. and Avouac J-P., 2013 - Detecting periodicities and declustering in earthquake catalogs using the Schuster spectrum, application to Himalayan seismicity, *Earth Planet. Sci. Lett*., 377-378, 97–105, doi:10.1016/j.epsl.2013.06.032.

Allmendinger R.W., Reilinger R., Loveless J., 2007 - Strain and rotation rate from GPS in Tibet, Anatolia and the Altiplano., *Tectonics*, 26:TC3013. doi:10.1029/2006TC002030.

Altamimi Z., Collilieux X., Métivier L., 2011 - ITRF2008: An improved solution of the International Terrestrial Reference Frame, *J. Geodesy*, 85(8), 457–473, doi:10.1007/s00190-011-0444-4.

Ambrosetti P., Carboni M.G., Conti M.A., Costantini A., Esu D., Gadin A., Girotti O., Lazzarotti A., Mazzanti R., Nicosia U., Parisi G., Sandrelli F., 1978 - Evoluzione paleogeografica e tettonica dei bacini tosco-umbro-laziali nel Pliocene e nel Pleistocene inferiore. *Mem. Soc. Geol. Ital*. 19, 573–580.

Anderlini L., Serpelloni E., Belardinelli M.E., 2016 - Creep and locking of a low-angle normal fault: Insights from the Altotiberina fault in the Northern Apennines (Italy), *Geophys. Res. Lett.*, 43 (2016), pp. 4321–4329 http://dx.doi.org/10.1002/2016GL068604.

Axen G.J., 1999 - Low-angle normal fault earthquakes and triggering., *Geophys. Res. Lett.*, 26 (24), 3693–3696. doi:10.1029/1999GL005405.

Anderson E.M., 1951 - The Dynamics of Faulting, 2nd edition, *Oliver and Boyd*, Edinburgh, 206.

Avallone A., Selvaggi G., D'Anastasio E., D'Agostino N., Pietrantonio G., Riguzzi F., Serpelloni E., Anzidei M., Casula G., Cecere G., D'Ambrosio C., De Martino P., Devoti R., Falco L., Mattia M., Rossi M., Obrizzo F., Tammaro U., Zarrilli L., 2010 - The RING network: Improvements to a GPS velocity field in the central Mediterranean, *Ann. Geophys.*, 53(2), 39–54, doi:10.4401/ag-4549.

Barba S. and Basili R., 2000 - Analysis of seismological and geological observations for moderate-size earthquakes: the Colfiorito Fault System (Central Apennines, Italy), *Geophys. J. Int.*, 141, 241-252.

Barchi M.R., Minelli R., Pialli G., 1998 - The crop 03 profile: a synthesis of results on deep structures of the northern Apennines, *Memorie della Società Geologica Italiana*, 52, 383-400.

Bettinelli P., Avouac J.P., Flouzat M., Bollinger L., Ramillien G., Rajaure S., Sapkota S., 2008 - Seasonal variations of seismicity and geodetic strain in the Himalaya induced by surface hydrology, *Earth Planet. Sci. Lett.*, 266, 332–344.

Blewitt G. and Lavallee D., 2002 - Effect of annual signals on geodetic velocity, J. *Geophys. Res.*, 107(B7), 2145, doi:10.1029/2001JB000570.

Boehm J., Niell A., Tregoning P., Schuh H., 2006 - Global Mapping Function (GMF): A new empirical mapping function based on numerical weather model data, *Geophys. Res. Lett*., 33, L07304. doi:10.1029/2005GL025546.

Boncio P., Brozzetti F., Lavecchia G., 2000 - Architecture and seismotectonics of a regional low-angle normal fault zone in central Italy, *Tectonics*, 19,1038-1055.
Boncio P., Lavecchia G., 2000 - A geological model for the Cofliorito earthquakes (September–October 1997, central Italy), *J. Seismol*. 4, 345– 356.

Brozzetti F., 1995 - Stile strutturale della tettonica distensiva nell'Umbria occidentale: l'esempio dei Massicci Mesozoici Perugini, *Studi Geologici Camerti*, special vol. 1995/1, 105-119.

Brozzetti F., Luchetti L., Pialli G., 2000 - La successione del Monte Rentella (Umbria Occidentale); biostratigrafia a nannofossili calcarei ed ipotesi per un inquadramento tettonico regionale, *Bollettino della Società Geologica Italiana*, 119, 407-422.

Chiaraluce L., Chiarabba C., Collettini C., Piccinini D., Cocco M., 2007 - Architecture and mechanics of an active low-angle normal fault: Alto Tiberina Fault, Northern Apennines, Italy, *J. Geophys. Res.*, 112, B10310, doi:10.1029/2007JB005015.

Chiaraluce L., Amato A., Carannante S., Castelli V., Cattaneo M., Cocco M., Collettini C., D'Alema E., Di Stefano R., Latorre D., Marzorati S., Mirabella F., Monachesi G., Piccinini D., Nardi A., Piersanti A., Stramondo S., Valoroso L., 2014 - The Alto Tiberina Near Fault Observatory (northern Apennines, Italy), *Ann. Geophys.*, 57(3), S0327, doi:10.4401/ag-6426.

Chiodini G., Cardellini C., Amato A., Boschi E., Caliro S., Frondini F., Ventura G., 2004 - Carbon dioxide Earth degassing and seismogenesis in central and southern Italy, *Geophys. Res. Lett*., 31, L07615; doi:10.1029/2004GL019480.

Choudrey R.A., 2002 - Variational Methods for Bayesian Independent Component Analysis. Pattern analysis and machine learning - robotics research group, *University of Oxford*.

Christiansen L.B., Hurwitz S., Ingebritsen S.E., 2007 - Annual modulation of seismicity along the San Andreas Fault near Parkfield, CA, *Geophysical Research Letters*, p. L04306 http://dx.doi.org/10.1029/2006GL028634.

Collettini C. and Barchi M.R. 2002 - A low angle normal fault in the Umbria region (central Italy): A mechanical model for the related microseismicity, *Tectonophysics*, 359, 97-115.

Collettini C. and Holdsworth R.E., 2004 - Fault zone weakening processes along low-angle normal faults: insights from the Zuccale Fault, Isle of Elba, Italy, *J. Geol. Soc. London*, 161, 1039-1051.

Comon P., 1994 - Independent component analysis, A new concept?, *Signal Processing*, 36:287-314.

D'Agostino N., Mantenuto S., D'Anastasio E., Avallone A., Barchi M., Collettini C., Radicioni F., Stoppini A., Fastellini G., 2009 - Contemporary crustal extension in the Umbria–Marche Apennines from regional CGPS networks and comparison between geodetic and seismic deformation, *Tectonophysics*, 476, 3-12.

De Luca G., Cattaneo M., Monachesi G., Amato A., 2009 - Seismicity in central and northern Apennines integrating the Italian national and regional networks, *Tectonophysics*, 476, 121-135.

Devoti R., Pietrantonio G., Pisani A., Riguzzi F., Serpelloni E., 2010 - Present day kinematics of Italy, in The Geology of Italy: Tectonics and Life Along Plate Margins, *J. Virtual Explorer*, Electronic Ed., vol. 36, edited by M. Beltrando et al., paper 2, doi:10.3809/jvirtex.2009.00237.

Dong D., Herring T., King R., 1998 - Estimating regional deformation from a combination of space and terrestrial geodetic data, *J. Geodesy*, 72(4), 200–214.

Dong D., Fang P., Bock Y., Cheng M., Miyazaki S., 2002 - Anatomy of apparent seasonal variations from GPS-derived site position time series, *J. Geophys. Res*., 107(B4), 2075, doi:10.1029/2001JB000573.

Dong D., Fang P., Bock Y., Webb F., Prawirodirdjo L., Kedar S., Jamason P., 2006 - Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis, *J. Geophys. Res*., 111, B03405, doi:10.1029/2005JB003806.

Ebel J.E., Ben-Zion Y., 2014 - Annual Temperature Vatiations and Seasonal Seismicity in California and Nevada, http://earth.usc.edu/~ybz/pubs_recent/Ebel_BZ_BSSA09/ Ebel_Ben-Zion_BSSA08.pdf.

Feigl K.L., King R.W., Jordan T.H., 1990 - Geodetic measurements of tectonic deformation in the Santa Maria fold and thrust belt, California, *J. Geophys. Res*., 95, 2679–2699.

Finocchio D., Barba S., Santini S., Megna A., 2013 - Interpreting the interseismic deformation of the Altotiberina Fault (central Italy) through 2D modeling, *Annals of Geophysics*, 56 (6), S0673; doi:10.4401/ag-5806.

Floyd J., Mutter J., Goodliffe A.M., Taylor B., 2001 - Evidence for fault weakness and fluid flow within an active low-angle normal fault., *Nature* 411, 779–783.

Gao S.S., Silver P.G., Linde A.T., Sacks I.S., 2000 - Annual modulation of triggered seismicity following the 1992 Landers earthquake in California, *Nature*, 406, 500–504.

Gualandi A., 2015 - Spatial and temporal characterization of ground deformation recorded by geodetic techniques, *University of Bologna*.

Gualandi A., Avouac J-P., Galetzka J., Genrich J.F., Blewitt G., Adhikari L.B., Koirala B.P, Gupta R., Upreti B.N, Pratt-Sitaula B., 2016 - Pre-and post-seismic deformation related to the 2015, Mw7. 8 Gorkha earthquake, Nepal Tectonophysics.

Hansen P. C., 1992 - Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Rev*., 34, 561–580.

Herault J. and Jutten C., 1986 - Space or time adaptive signal processing by neural models. In *Proceedings* AIP *Conference*: *Neural Networks for Computing*, volume 151, pages 206-211. J.S. Denker, Ed. American Institute for Physics.

Herring T., King R.W., McClusky S., 2010 - GAMIT Reference Manual, Release 10.4, *Massachusetts Institute of Technology*, Cambridge, MA.

Hewitt C.D. and Griggs D.J., 2004 - Ensembles-based Predictions of Climate Changes and their Impacts., *Eos*, 85, p566.

Hyvärinen A. and Oja E., 2000 - Independent Component Analysis: Algorithms and Applications., *Neural Networks*, 13(4-5):411-430.

Hreinsdottir S. and Bennett R. A., 2009 - Active aseismic creep on the Alto Tiberina lowangle normal fault, Italy, *Geology*, 37, 683-686.

Jackson J.A., White N.J., 1989 - Normal faulting in the upper continental crust: observations from regions of active extension., *Journal of Structural Geology*, 11, 15–36.

Jolivet L., Faccenna C., Goffe´ B., Mattei M., Rossetti F., Brunet C., Storti F., Funiciello R., Cadet J.P., D'Agostino N., Parra T., 1998 - Midcrustal shear zones in postorogenic

extension: example from the northern Tyrrhenian Sea., *J. Geophys. Res*. 103, 12123–12160.

Kositsky A.P., Avouac J-P., 2010 - Inverting geodetic time series with a principal component analysis-based inversion method., *J. Geophys. Res*., 1150 (B03401). http://dx.doi.org/10.1029/2009JB006535.

Lister G.S. and Davis G.A., 1989 - The origin of metamorphic core complexes and detachment faults formed during Tertiary continental extension in the northern Colorado River region, USA, *J. Struct. Geol*., 11, 65-93.

Longwell C.R., 1945 - Low-angle normal faults in the Basin and Range province., *Trans. Am. geophys*, Un.26,107-118.

Lyard F., Lefevre F., Letellier T., Francis O., 2006 - Modelling the global ocean tides: Modern insights from FES2004, *Ocean Dynam*., 56, 394–415.

Mirabella F., Brozzetti F., Lupattelli A., Barchi M.R., 2011 - Tectonic evolution of a low angle extensional fault system from restored cross sections in the northern Apennines (Italy), *Tectonics*, 30, TC6002; doi:10.1029/2011TC002890.

Moretti M., De Gori P., Chiarabba C., 2009 - Earthquake relocation and three-dimensional Vp and Vp/Vs models along the low angle Alto Tiberina Fault (Central Italy): evidence for fluid overpressure, *Geophys. J. Int*., 176 (3): 833-846

Morley C.K., 1999 - Marked along-strike variations in dip of normal faults - the Lokichar fault, N. Kenya rift: a possible cause for metamorphic core complexes. *J. Struct. Geol*. 21, 479– 492.

Munafò I., Malagnini L., Chiaraluce L., 2016, Short Note On the Relationship between M_w and ML for Small Earthquakes, *Bulletin of the Seismological Society of America*, Vol. 106, No. 5, doi: 10.1785/0120160130.

Okada Y., 1985 - Surface deformation to shear and tensile faults in a half-space, *Bull. Seismol. Soc. Am*., 75, 1135–1154.

Piccinini D., Cattaneo M., Chiarabba C., Chiaraluce L., De Martin M., Di Bona M., Moretti M., Selvaggi G., Augliera P., Spallarossa D., Ferretti G., Michelini A., Govoni A., Di Bartolomeo P., Romanelli M., Fabbri J., 2003 - A microseismic study in a low seismicity area of Italy: the Città di Castello 2000-2001 experiment, *Annals of Geophysics*, 46 (6), 1315-1324.

Pietrantonio G., Riguzzi F., 2004 - Three-dimensional strain tensor estimation by GPS bservations: methodological aspects and geophysical applications, *Journal of Geodynamics*, 38 (2004), pp. 1–18.

Press W., Teukolsky A., Vetterling W., Flannery B., 2002 - Numerical Recipes in C: the Art of Scientific Computing, *Cambridge Univ Press*, Cambridge, U. K.

Proffett J.M., 1977 - Cenozoic geology of the Yerington district, Nevada, and implications for the nature of Basin and Range faulting., *Geological Society of America Bulletin* 88, 247–266.

Pucci S., De Martini P.M., Pantosti D., Valensise G., 2003 - Geomorphology of the Gubbio Basin (Central Italy): understanding the active tectonics and earthquake potential, *Annals of Geophysics*, [S.l.], v. 46, n. 5, dec. 2003, doi:10.4401/ag-3458.

Radiguet M., Cotton F., Vergnolle M., Campillo M., Valette B., Kostoglodov V., Cotte N., 2011 - Spatial and temporal evolution of a long term slow slip event: the 2006 Guerrero slow slip event, *Geophys. J. Int*., 184 (0), 816–828. http://dx.doi.org/10.1111/j. 1365-246X.2010.04866.x.

Reasenberg E. A., 1985 - Second-order moment of central California seismicity, *J. Geophys. Res*., 90,5,479 –5,495.

Riel B., Simons M., Agram P., Zhan Z., 2014 - Detecting transient signals in geodetic time series using sparse estimation techniques, *J Geophys Res Solid Earth*, 119:5140–5160. doi:10.1002/2014JB011077.

Rigo A., Lyon-Caen H., Armijo R., Deschamps A., Hatzfeld D., Makropoulos K., Papadimitriou P., Kassaras I., 1996 - A microseismic study in the western part of the Gulf of Corinth (Greece): implications for large scale normal faulting mechanisms., *Geophys. J. Int.,* 126, 663–688.

Rovida A., Locati M., Camassi R., Lolli B., Gasperini P. (eds), 2016 - CPTI15, the 2015 version of the Parametric Catalogue of Italian Earthquakes, Istituto Nazionale di Geofisica e Vulcanologia, doi:http://doi.org/10.6092/INGV.IT-CPTI15.

Schuster A., 1897 - On lunar and solar periodicities of earthquakes, *Proc. R. Soc. Lond*., 61, 455–465.

Serpelloni E., Anzidei M., Baldi P., Casula G., Galvani A., 2005 - Crustal velocity and strain-rate fields in Italy and surrounding regions: New results from the analysis of permanent and non-permanent GPS networks, *Geophys. J. Int*., 161 (3), 861-880; doi:10.1111/j.1365246X.2005.02618.x.

Serpelloni E., Anzidei M., Baldi P., Casula G., Galvani A., 2006 - GPS measurement of active strains across the Apennines, *Ann. Geophys*., 49(1), 319–329, doi:10.4401/ag-5756.

Serpelloni E., Bürgmann R., Anzidei M., Baldi P., Mastrolembo Ventura B., Boschi E., 2010 - Strain accumulation across the Messina Straits and kinematics of Sicily and Calabria from GPS data and dislocation modeling, *Earth Planet. Sci. Lett*., 298 (2010), pp. 347–360 http://dx.doi.org/10.1016/j.epsl.2010.08.005.

Sibson R.H., 1985 - A note on fault reactivation., *Journal of Structural Geology*, 7, 751–754.

Srebro N. and Jaakkola T., 2003 - Weighted low-rank approximations, *Twentieth International Conference on Machine Learning*, http://ttic.uchicago.edu/nati/ Publications/SrebroJaakkolaICML03.pdf.

Tape C., Musé P., Simons M., Dong D., Webb F., 2009 - Multiscale estimation of GPS velocity fields, *Geophys. J. Int*., 179(2), 945–971, doi:10.1111/j.1365-246X.2009.04337.x.

Tarantola A., 2005 - Inverse Problem Theory and Methods for Model Parameter Estimation, *SIAM*.

Taylor B., Goodliffe A.M., Martinez F., 1999 - How continents break up: insights from Papua New Guinea. *J. Geophys. Res*. 104, 7497– 7512.

Vadacca L., Anderlini L., Casarotti E., Serpelloni E., Chiaraluce L., Polcari M., Albano M., Stramondo S., 2014 - Active tectonic extension across the Alto Tiberina normal fault system from GPS data modeling and InSAR velocity maps: new perspectives within TABOO Near Fault Observatory, *Geophysical Research Abstracts*, 16, EGU2014-6325-1, EGU General Assembly.

van Dam T., Wahr J., Milly P.C.D., Shmakin A.B., Blewitt G.,Lavallee D., Larson K.M., 2001 - Crustal displacements due to continental water loading, *Geophys. Res. Lett*, 28, 651–654.

Wernicke B. and Axen G.J., 1988 - On the role of isostasy in the evolution of normal fault systems., *Geology* 16, 848–851. doi:10.1130/0091-7613(1988) 016.

Wernicke B., 1995 - Low-angle normal faults and seismicity: A review, *J. Geophys. Res*., 100, 20159-20174.

Wiemer S., 2001 - A software package to analyze seismicity: ZMAP, *Seismol. Res. Lett*., 72(3), 373–382.

Williams S.D.P., 2008 - CATS: GPS coordinate time series analysis software, *GPS Solut*., 12(2), 147–153, doi:10.1007/s10291-007-0086-4.

Acknowledgements

Alla fine di questo percorso che comprende i tre anni della laurea triennale e i due della laurea magistrale vorrei spendere qualche parola per ringraziare tutte le persone che mi hanno aiutata e sostenuta.

In primo luogo vorrei ringraziare la Prof. Maria Elina Belardinelli per aver accettato di essere la mia relatrice di tesi per la seconda volta e per avermi sempre guidata con amorevole professionalità. Un grande grazie anche al Dott. Enrico Serpelloni per la sua pazienza e disponibilità e per avermi proposto di fare la bellissima esperienza a Caltech che mi ha dato tanto dal punto di vista umano e professionale.

A special thank goes to Prof. Jean-Philippe Avouac for having welcomed me to his group at Caltech and for his valuable help and support.

Vorrei ringraziare Adriano per tutto il tempo speso ad aiutarmi e a chiarire i miei dubbi, per avermi accolta in casa sua i primi giorni dopo il mio arrivo a Pasadena e per avermi trasmesso la grande passione per il suo lavoro.

Ringrazio Letizia per le informazioni sull'ATF e per i preziosi consigli e Francesco per la gentile disponibilità nel fornirmi i dati di temperatura e piovosità della regione.

Thank you Sylvain and Mortaza for having shared lunch with me since my first day at Caltech.

Thank you Claudia and Alice for having been great roommates and friends during my time in Pasadena.

Grazie ad Anna, Barbara, Davide, Desy, Fra, Linda, Pede, Sere, Simo, Tappo, Vane e Viki per i bei momenti trascorsi insieme, tra chiacchiere e allegria.

Ringrazio Enrico e Manu per la loro sincera amicizia di lunghissima data. Grazie anche per avermi presentato Giulia C. e Giulia S. che stimo e sono sicura diventeranno grandi Sanlorenzesi.

Grazie Vicky per la tua amicizia, gentilezza e disponibilità.

Ringrazio Giorgia per i tre anni trascorsi da coinquiline e i cinque anni da compagne di studio a Bologna. Questo periodo ha segnato una parte importante della mia vita e sono felice di averlo condiviso con te.

Grazie a Bea per la nostra amicizia che dura da tantissimi anni nonostante le nostre diversità. Grazie infinite per essermi stata così vicina, soprattutto in questi mesi e per avermi scelta per essere al tuo fianco nel tuo giorno più importante.

Ringrazio i miei zii Franca e Carlo e i miei cugini Marco, Monica e Manuela per il loro supporto in questi anni.

Grazie zia Maria, zio Giovanni, Dany, Massimo e Luca per l'affetto che sempre mi dimostrate.

Vorrei ricordare Don Vittorio e la Rosanna per avermi sempre sostenuta e incoraggiata nel mio percorso e ringraziare tutta la comunità di San Lorenzo per essermi stata vicina durante il periodo trascorso in California, facendomi sentire sempre a casa anche quando non lo sono.

Un grande grazie a Suor Cinzia: la tua amicizia e il tuo sostegno sono sempre fondamentali.

Ringrazio mia sorella Chiara per avermi insegnato le virtù della calma e della riflessione e mio fratello Andrea per avermi trasmesso l'entusiasmo e la gioia di vivere.

Grazie ai miei genitori, Laura e Bruno, per avermi insegnato con l'esempio, che l'impegno, la fatica e il coraggio sono la chiave per realizzare i propri sogni.