Seismic Source Quantitative Parameters Retrieval from InSAR Data and Neural Networks

XVIII Doctorate in GEOINFORMAZIONE

University "Tor Vergata" - Rome

Doctoral Student:
Salvatore Stramondo

Supervisor:
Prof. Fabio Del Frate

Introduction

The basic idea of this PhD thesis is to exploit the capabilities of neural networks in a very new framework: the quantitative modelling of the seismic source and the interferogram inversion for retrieving its geometric parameters.

In this study an innovative approach for the seismic source classification and the fault parameter retrieval has been investigated. Furthermore it concerns the exploitation of neural networks capabilities and the use of InSAR measurements.

The problem can be sum up as follows. When a moderato-to-strong earthquake occurs we can apply SAR Interferometry (InSAR) technique to compute a differential interferogram. The latter is used to detect and measure the surface displacement field. The earthquake is generated by an active, seismogenic, fault having its own specific geometry. Therefore each differential interferogram contains the information concerning the geometry of the seismic source responsible for the earthquake; its shape and size, the number of fringes, the lobe orientation and number are the main features characterizing the surface displacement field. To retrieve the geometric parameters of the fault plane we compute a huge number of synthetic interferogram, also referred to as models. Each of them is generated by a different combination of such geometric parameters.

The inversion procedure stems from an artificial neural network properly generated and trained to recognize the synthetic interferograms and to classify them as belonging to a certain subset. Furthermore the already trained neural network is used to work with the coseismic InSAR measures. First it singles out the synthetic interferogram best fitting the measured one, and then it is used to retrieve the fault plane parameters by this latter.

More in particular the solution of the inverse problem means to recover the source parameters from the knowledge of InSAR surface displacement field. To this aim, some significant results have already been achieved by means of the simulated annealing technique (Billings, 1994; Kirkpatrick, 1983; Lundgren and Stramondo, 2002). Inversion algorithms using geodetic data and based on the maximum likelihood criteria (Fukahata et al., 2004) or a quasi Montecarlo (Gareth et al., 2005) iterative approach have been also proposed. However, due to the intrinsic ill-posedeness of the problem, some issues remain still open and more investigation is needed.

Neural networks have already been recognized as being a powerful tool for inversion procedure in remote sensing applications (Bishop, 1995; Del Frate et al., 2003). Among the application fields the urban changes (Del Frate F. and Solimini C., 2005), environmental studies concerning oil spill (F. Del Frate, L. Salvatori, 2004), atmospheric and meteorological models (F. Del Frate and G. Schiavon, 1999), ozone retrieval (Iapaolo M., 2003).

They are composed of many nonlinear computational elements (called neurons) operating in parallel and connected by the so called synapses. The use of neural networks is often effective because they can simultaneously address nonlinear dependencies and complex physical behaviour. Two problems have been analysed in this work. The first is the identification of the seismic source mechanism. The second is a typical inversion exercise concerning the fault plane parameter estimation.

1 Tectonic elements

Scope of ***Tectonics*** (from the Greek τεκτον = constructor) is the observation and the description of the main structures of the land crust to study as these structures are formed.
Thanks to the theory of the Plate Tectonics, the sources of the deformations of the land crust can be searched inside the mantle, where with its convective movements inside tears, dislocates, transports and makes colluder between the *litospheric plates*.
To a much smaller scale, these plate movements are generated by the forces responsible for the deformation of the land crust.

1.1 The inside of the Earth

Thanks to the study of the *seismograms,* i.e. the registration of the seismic waves, the inside of the Earth called also *geosphere* can be divided in some concentric shells. The analysis of the seismic waves allows pointing out phenomena of refraction while crossing the inner planet. The refraction consists in the modification of speed and trajectory of a wave when this is transmitted to mediums with different density. Therefore deep surfaces are present where an abrupt acceleration and shunting line of the waves can be verified. Based on such discontinuities, four concentric spherical zones have been identified: *crust*, *mantle*, *external nucleus* and *inner nucleus*.
External nucleus is mainly composed by iron, with approximately 10% of sulfurs (S). Due to its pressure and temperature it is considered "liquid". Instead the inner nucleus reaches the pressure conditions of the "solid" state. The greater part of the mass of the Earth is composed by the *mantle* (Figure 1), which essentially is a mixture of Iron (Fe), Magnesium (Mg), Aluminum (To), Silice (Si) and composed of silicates of Oxygen (O). With temperatures beyond the 1000 centigrade degrees the *mantle* turns out solid, and it can be affected by plastic deformations. The crust is much thinner than the other "shells" (or layers) and it is composed by less dense materials and mainly from Calcium (Ca), and minerals alum-silicates of Sodium (Na). Being relatively cold it turns out fragile and can therefore fracture itself thus generating earthquakes.
The earthquakes however take place also at high depths in particular regarding the crustal seismic events, in the zones of *subduction* for instance.

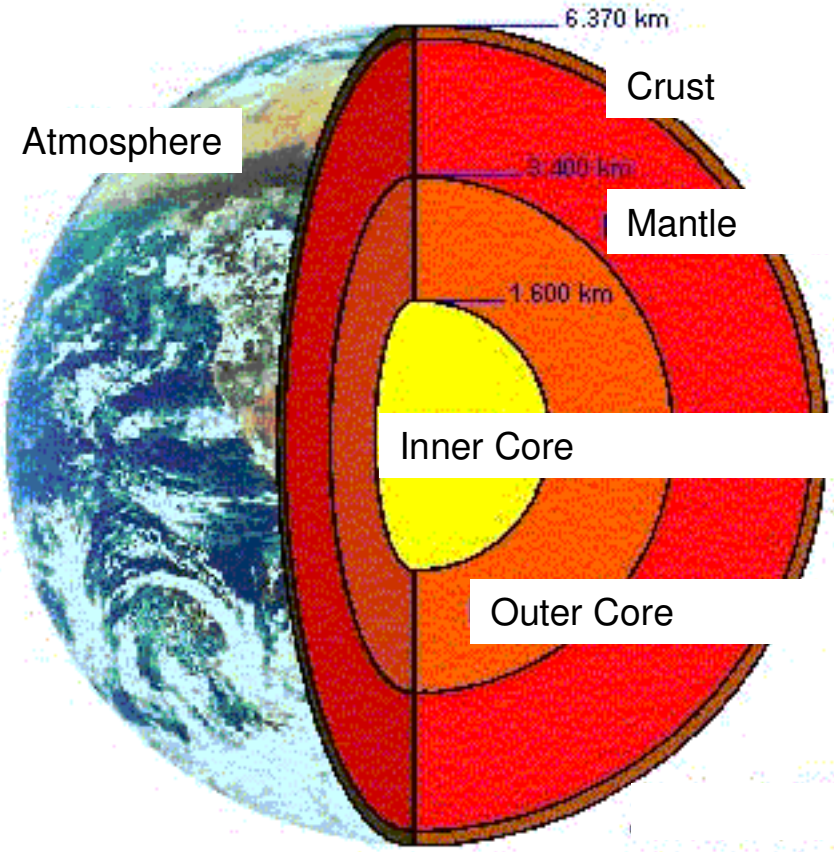


Fig.1: Section of the Earth interior.

1.2 Chemical-physical properties of the geosphere

The composition of the Earth ordered according to the mass of the constituents is:

- 34.6 % Iron
- 29.5% Oxygen
- 15.2% Silicon
- 12.7% Magnesium
- 2.4% Nickel
- 1.9% Sulfur
- 0.05% Titanium

The temperature inside the Earth progressively increases until reaching the 5270°K to its center. Heat inside has been partially generated during the formation of the planet and from then on it has been continuously generated from radioactive decay of isotopes of uranium, thorium and potassium.

The rocks are bad conductor of heat; therefore the heat transmitted from the inside towards outside of the planet is only 1/20000 of the heat that the planet receives from the Sun.
The density average of the Earth is 5.515 g/cm^3, rendering it the denser planet of the Solar System. Since the density of rocks in the land crust is from 2.6 to 3.5 g/cm^3, the material of the nucleus must be the densest one: approximately 5.7 g/cm^3.

1.3 Lithosphere, Astenosphere, Mesosphere

From the point of view of the mechanical behavior the layers of *geosphere* can be divided in a various order.
The more external, rather rigid layer, called *lithosphere* is often approximately 100 km under the continents and 50 km under the oceans and is composed by the *crust* and the rigid upper mantle. The layer beneath, the *Astenosphere,* is extended until 700 km of depth and is characterized by less fragile rocks, comprising a zone of mantle much warmer and in one that we could define plastic. Such plasticity offers a discrete modularity, while the more rigid lithosphere is fallen apart, with abrupt fractures, if sped up from tractions or jams of one sure entity. The inferior mantle, rigid because of the elevated pressures caused from the advanced layers, is the *mesosphere* (Figure 2)

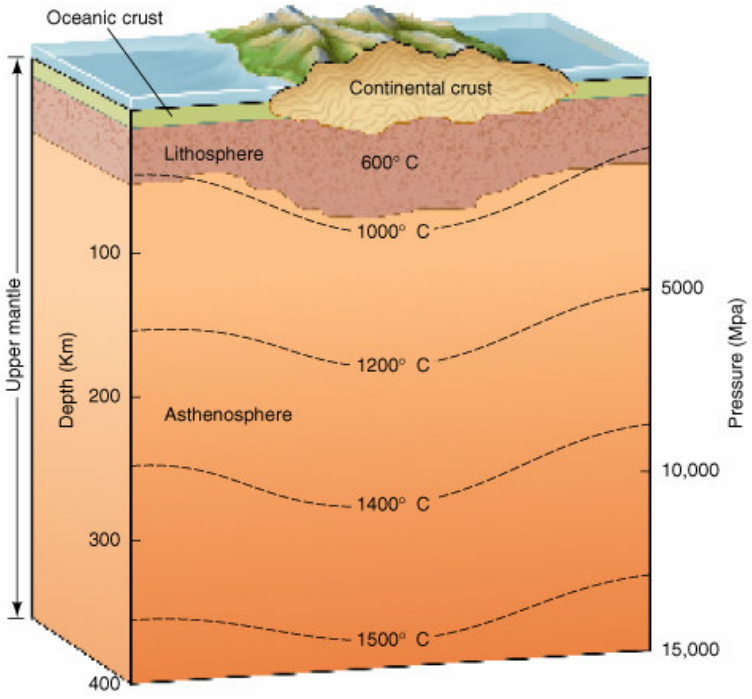


Fig.2: the upper mantle and the crust: lithosphere e astenosphere

1.4 The plate tectonics

The plate tectonics theory explains the causes of the earthquakes. According to this never universally accepted theory, the upper part of the Earth is composed by two layers with different deformative properties, the lithosphere and the astenosphere. The former is composed by a huge amount of plates of oceanic litosphere (thinner) and/or continental lithosphere moving one each other, dragged by the slow convective movements of the upper mantle (astenosphere).

The plate interactions are responsible for the main structural features of the Earth, due to volcanic, seismic and orogenic activities. The dynamic model of the lithosphere, that is the plate tectonics, explains coherently a huge amount of independent data, like the magnetic anomalies of the oceanic floors, the age changes of the oceanic sediments and the underlying oceanic lithosphere, the ipocentral earthquake distribution, the oceanic floors morphology, etc..

The idea of moving plates was originally proposed in 1912 by the German scientist A.Wegener. According to him, about 200 My ago it existed a unique supercontinent, the Pangea, surrounded by the ocean Pantalassa. Later on the Pangea was separated in some continental edges, called sialic (composed by aluminium silicates). They were like rafts floating on the denser and viscous Sima (silicium and magnesyum), moving away one each other from the Poles and toward W, according to the mechanism called continental drift.

Such mechanism justify the Westward movement of the Sial blocks on the Sima layer as the consequence of the delay of such blocks respect to the Eastward Earth rotation. Moreover the inertia of the sialic blocks nowadays causes their slow drift. The sialic blocks should have also been affected by the centrifugal force due to the Earth rotation. The continental drift explains the complementary shapes between continental margins, as the Atlantic coasts of the Latina America and Africa, but also the analogous rock types, fossil flora and fauna. In Figure 3 below is a simplified sketch of the plate tectonics where the main plates are showed.

The earthquakes can only take place in the fragile lithosphere. The border areas between plates are the most affected by earthquakes.

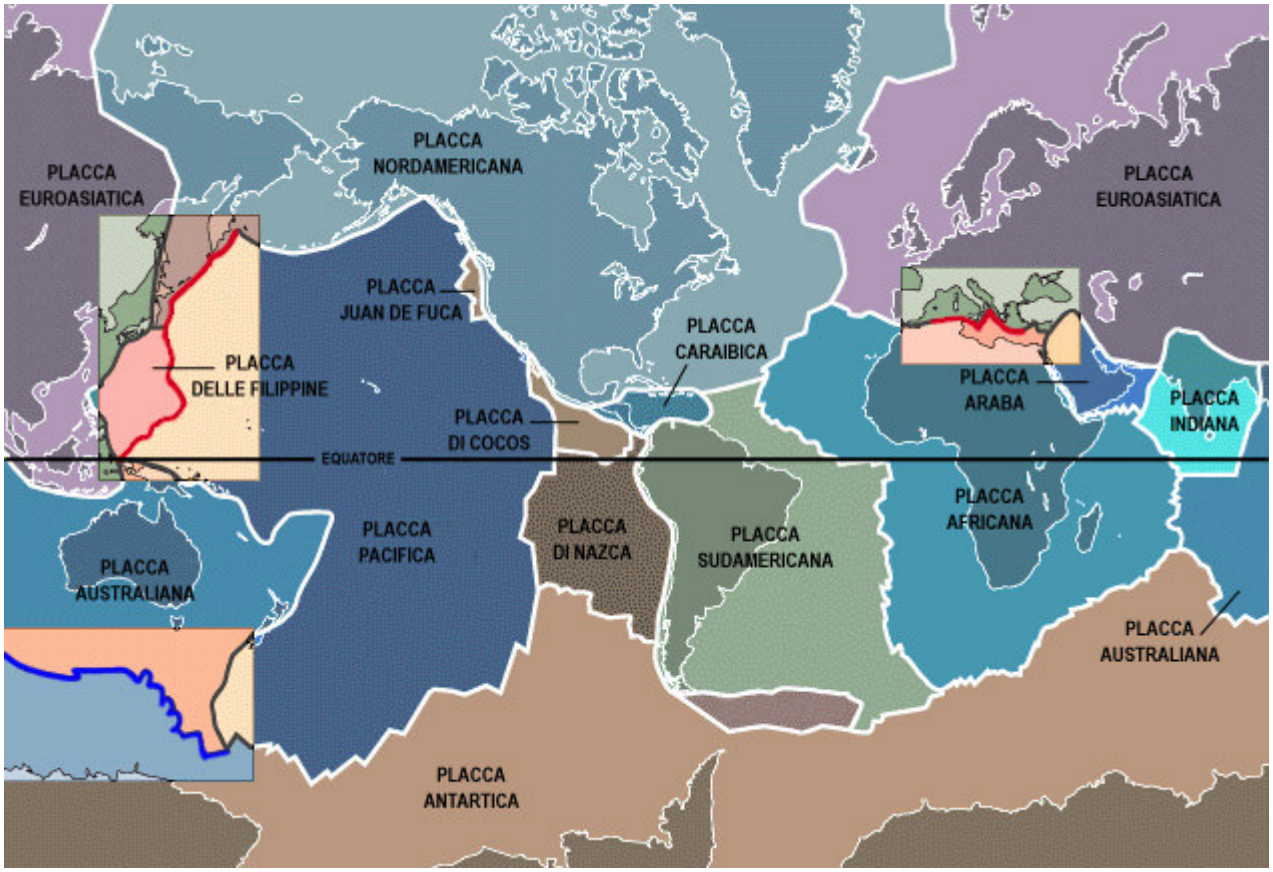


Fig.3: Plate tectonics

The plate movements can be classified in three different types. The converging plate boundaries are called *trenches*; the diverging boundaries are called ridges. Whenever the plates move horizontally one each other they are called *transforms*. The trenches are destructive boundaries, meaning that one plate is subducting under the other. The diverging *ridges*, on the contrary, are continuously filled by mantle materials.

1.4.1 The Earth crust deformations

The average velocity of one plate relatively to the nearby one has been estimated about 5 cm/y. Such reciprocal movements generate forces exploiting the Earth crust deformation. Be S the stress and F the Force applied to a Surface A:

$S = F/A$

The Stress is a function of the gravity force and of the specific geodynamic behaviour (distensive, compressive or thrust). The Stress may generate a permanent deformation of the crust called strain, defined as the shape and volume deformation of an object affected by a certain stress.

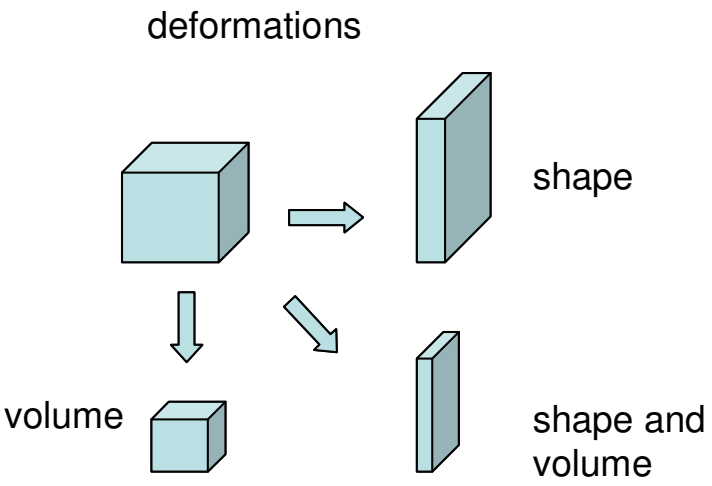


Fig. 4: The strain is the amount of volume and shape deformation of an object a stress regime has been applied to.

Therefore the stress is the cause of the deformation while the strain is the quantification of such effect. Each geodynamic environment is characterized by a specific stress regime.
The Earth crust deformations are mainly functions of:

1. the temperature;
2. the pressure;
3. the velocity at which they take place;
4. the time materials are affected by the stress;
5. the mechanical properties of materials;
6. the presence of water and fluids or not;
7. the variation along time of the previous properties.

1.5 The faults

The term fault indicates all plane discontinuities along which a relative movement of two mass rocks takes place. The faults are a fragile response to a stress operating on the Earth crust. The latter

justify why the faults are mainly seismogenic. These fractures may reach hundred kilometers, as in the San Andreas fault or the North Anatolic fault.
The main observations to locate a fault are:

1. a structural discontinuity
2. a lithologic discontinuity
3. an area of deformated rocks
4. surface deformations
5. sedimentation due to the fault.

However not all the structural or lithologic discontinuities are faults. At depth the geothermic gradient increases and the movement on the fault plane turns to a ductile regime. The depth of such transitions changes due to the material properties. For instance, faults in salt and gypsum are ductile also in the upper crust, while silicates may have a plastic behaviour from minimum 10 km depth.
The analysis of phenomenon concerning faults also involves the elements composing the faults themselves. Therefore the characteristic parameters have to be defined.
The fault plane is the irregular surface of the fracture; it can be vertical, horizontal or sloping. It usually seems striated or smoothed due to the friction of the walls of the fault.
The intersection between the fault and the surface is named trace. The geometrical parameters are the Dip angle, i.e. the angle between the fault plane and the surface; the Strike angle, i.e. the angle between the trace of the fault and the N-S direction. In non-vertical planes, the upper block is called hanging wall whilst the other below is the footwall.
The Slip is the displacement measured along the fault plane. The stratigraphic throw is the displacement of the two blocks along the fault plane.
The fault may assume any slope angle in the range 0° - 90° and the movement may happen along any direction. Based on the characteristics of the displacement the faults can be schematically divided in three classes:

- ***normal faults***
- ***strike slip faults***
- ***thrust faults***

1.5.1 Normal faults

In the normal fault the displacement is perpendicular to the separation surface. Therefore the hangingwall, i.e. the portion of the fault moving downward respect to the footwall. From a physical point of view this fault is the result of an extensional force moving the blocks away on each other. This fault can be observed as a clear plain, a fault zone or a ductile shear zone. Usually the normal faults have an inclination between 55°-70° but nearby the surface may reach 90°.

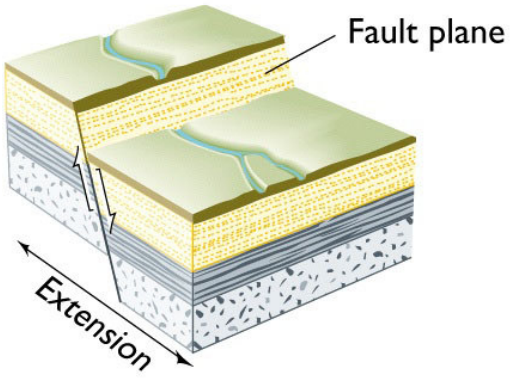


Fig. 5: normal fault mechanism. The simplified sketch shows the movement of the hangingwall relative to the footwall.

1.5.2 Strike slip faults

In the strike slip faults the relative movement of the two blocks is along a horizontal line. Concerning the direction the strike faults are right and left. This latter can be when an observer on one block side looks at the other moving towards left, while the opposite takes place in the former case. The strike slip faults usually do not affect the horography the displacements being parallel to the Earth surface. These faults can be usually recognized by aerial photos or satellite images, able to clearly point out such features.

Concerning the relative movement of a block in some cases it works as a simple slip. Anyway more complex geometries are also possible. For instance, the lateral displacement can be associated to a relaxation component moving away the blocks (transtensive regime), while a compression deals with blocks approaching (transpressive regime).

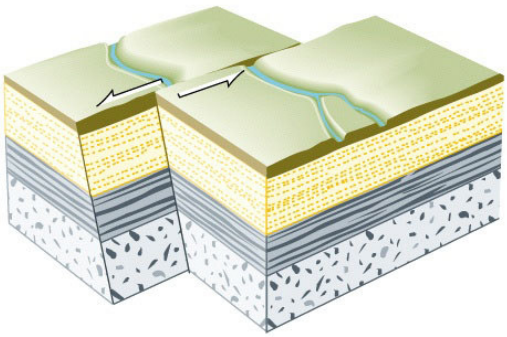


Fig. 6: strike slip fault mechanism

1.5.3 Thrust or reverse faults

In the thrust faults the displacement is perpendicular respect to the separation surface. It is an uplift of the hangingwall referred to the footwall.

These faults are generally due to compressive regime. The thrust mechanism takes place with inclinations between 0° and 45°. Despite this high angle thrust faults exist but they usually are strike slip faults, thrust faults inclined or high angle normal faults reactivated in thrust faults. This latter case takes place when the area is affected by a compressive regime.

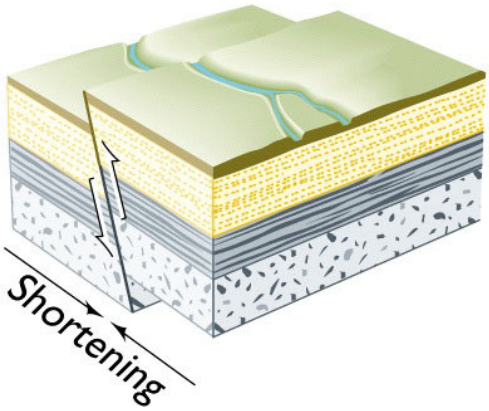


Fig. 7: thrust (or reverse) fault mechanism

1.6 The stress field

The proposed classification (normal, strike and thrust faults) is a simply way to characterize each fault according to its own prevalent behaviour and geometry. But usually faults have an intermediate behaviour. Therefore dealing with predominantly normal faults and vertical displacements one may talk about normal and/or thrust faults with a right or left lateral strike slip. On the contrary, whenever the fault is predominantly strike with horizontal displacements it would be right or left lateral strike with a normal and/or a thrust component.

The deformation takes place in a homogeneous and isotropic medium along symmetric planes. Therefore it can be considered also a quantitative classification criterion dealing with the geometry of the stress field affecting the fault. To this aim one may talk about:

Normal Faults: The stress field components have the σ_1 vertical, the σ_3 horizontal and perpendicular to σ_1 while the σ_2 is on the horizontal plane and parallel to the fault plane (Fig. 8).

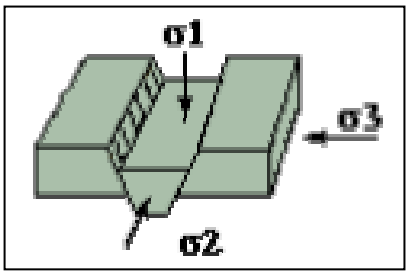


Fig. 8

Strike Faults: The stress field has a maximum σ_1 and a minimum σ_3 on the horizontal plane and oriented according with the bisector of the acute angle and the obtuse angle generated by two conjugate planes, while the σ_2 is vertical (Fig. 9).

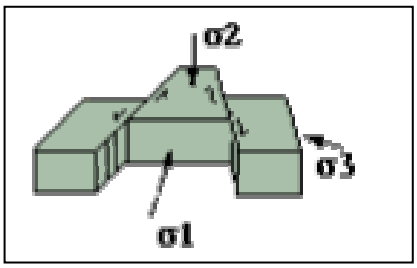


Fig. 9

Thrust faults: They are characterized by the presence of the maximum component σ_1 on the horizontal plane, the minimum component σ_3 vertical and the σ_2, on the horizontal plane and parallel to it (Fig. 10).

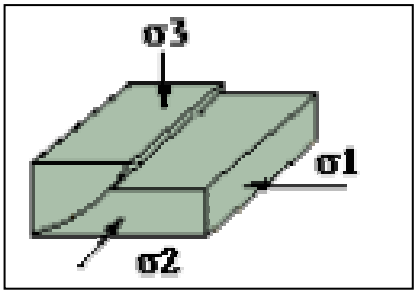


Fig. 10

The previous classification refers to the relations between the three stress components and the horizontal line. Such relations are meaningful only for superficial levels being deeper the direction of the vectors affected by other factors. Therefore the fault planes will be oriented according to the geometry of the stress field.

1.7 The earthquakes

The stress field affecting the fault plane may generate the earthquake. In fact a medium compressed and stretched tends to be deformed according to the physical and chemical properties of the material itself, as the temperature, the pressure, the presence of fluids.

In 1906 the American seismologist Harry F. Reid proposed a model concerning the mechanism of earthquakes. Reid stated that rocks, subjected to compression and stretching, have an elastic behaviour and progressively are deformed up to the break threshold. Whenever the latter limit is exceeded a new fault has been created. The two margins of the previous mass rock are now able to elastically react leaving in a short time the energy accumulated in tens or hundreds years. The energy is released as heat or seismic waves moving from the ipocenter towards all directions. Therefore the mass rock may achieve its original volume and equilibrium condition.

Reid called such model the “theory of the elastic rebound” particularly useful to explain the genesis of the surface earthquakes close to the ridges or the diverging plate boundaries.

The main energy releases are usually followed by smaller ones, generating smaller earthquakes called aftershocks.

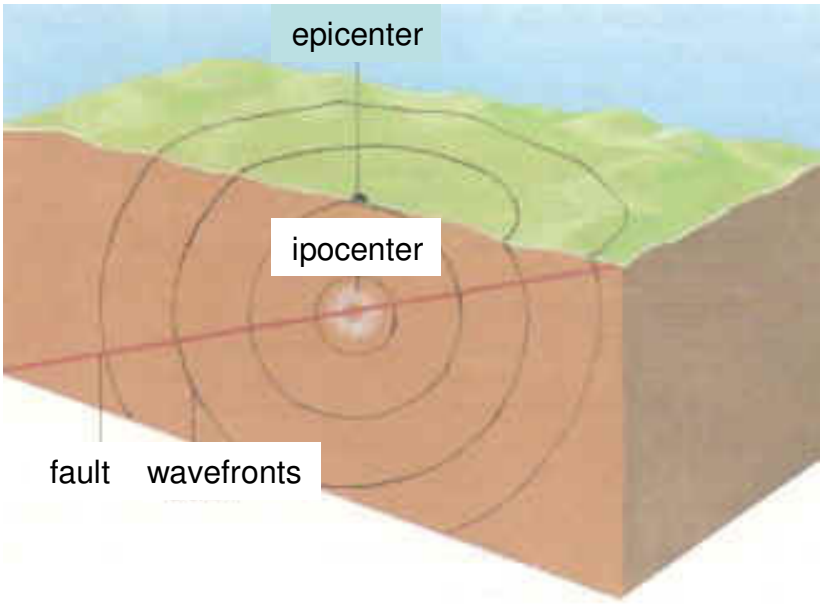


Fig. 11: The seismic waves propagating from the ipocenter through the Earth globe reach also the surface. The wavefronts are refracted by the different layers at variable densities.

1.8 The fault plane from seismological data and ground survey

The definition of the geometric parameters of a seismic source is based on the availability of some auxiliary data that have been parallely analysed.

The fault plane solutions are basically obtained by the seismological data. Besides these data a multisource dataset, ranging from the latter data (seismic waveforms, focal mechanism) to the GPS measurements, from InSAR results to the geological settings and the tectonic regime can be available and contribute to better define the seismic source. Without providing an in-depth description of the previous issues that is beyond the scope of this thesis, some points are noteworthy.

Once an earthquake takes place the seismic waves propagate from the seismic source at depth towards the surface. As more carefully described below the seismic source fault plane can be oriented according to the direction of the motion vectors. The measured movements due to the P-waves and S-waves allow drawing the focal mechanism. The latter is a 3-D diagram showing the two complementary fault geometries based on the compression and dilatation of the seismic waves (see figure below). The lower half of the focal sphere is the interception of two (primary and auxiliary) fault planes, each oriented according to the Strike and the Dip angles. These two are among the source parameters retrieved in chapter 4.

The focal mechanisms can be schematically sum up by three main classes: Normal, Strike slip, Thrust or Reverse. The seismological computing of the focal mechanism, i.e. based on the seismic waves, is the official way to provide a reliable source plane solution. The focal mechanism is usually depicted as a sphere containing the P-wave first motion and the S-wave first motion (Fig. 12). Two complementary fault geometries can be found, the primary and the auxiliary one.

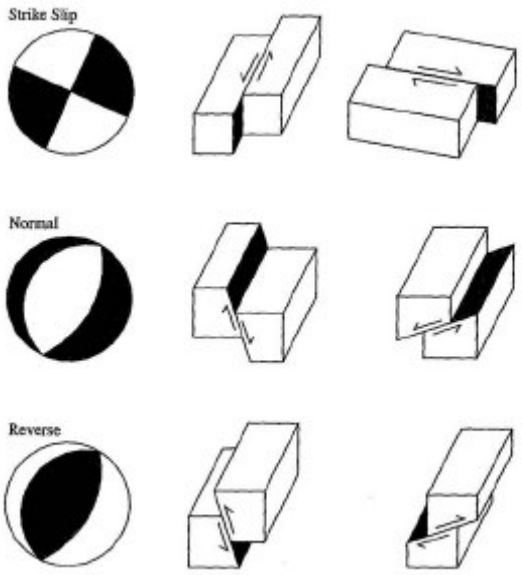


Fig. 12: Example of focal spheres and their corresponding fault geometries. The lower half of the focal sphere is plotted to the left, with compressional quadrants shaded. The block diagrams show the two fault geometries (the primary and auxiliary fault planes) that could have produced the observed radiation pattern.

From a physical point of view nine different force couples are applied to the fault plane thus generating the Seismic Moment Tensor.

$$M_{jk} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}$$

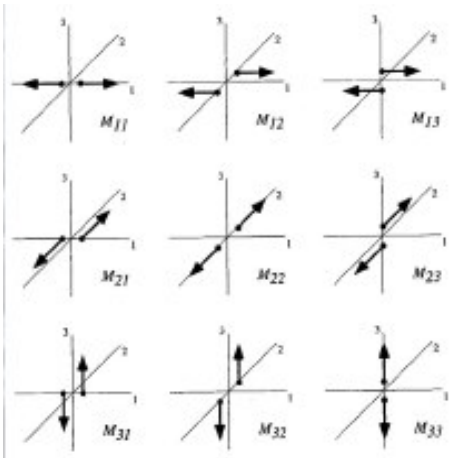


Fig. 13: The nine different force couples make up the components of the moment tensor.

Each fault plane geometry is the result of a specific combination of the applied forces. For instance a right lateral movement on a strike slip fault corresponds to the following one:

$$M = \begin{pmatrix} 0 & M_0 & 0 \\ M_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

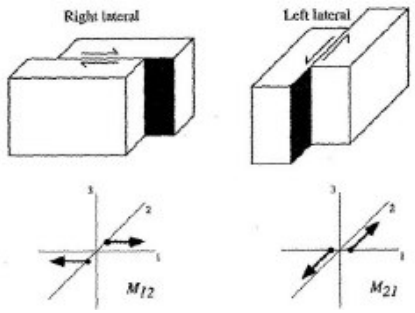


Fig. 14: Example of right lateral movement on a strike-slip fault in the x_1 direction.

where M_0 is the Scalar Seismic Moment. This latter is proportional to the Length and Width of the fault plane according to the simple following equation:

$$M_0 = \mu A s$$

where μ is the rigidity or shear modulus ($\approx 3.2x10^{11} dyne/cm^2$ in the crust), A the fault area (cm^2), s the fault slip (cm) (Fig. 15).

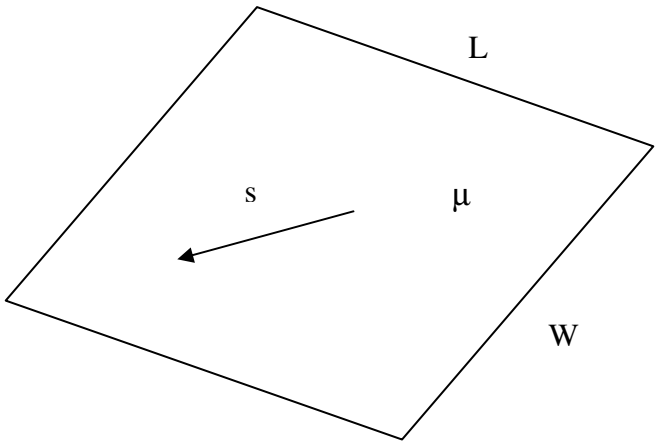


Fig. 15: Schematic fault plane geometry; the fault area is A=LW.

The seismic moment can be determined by:

- The geological measurements of the fault offset
- The estimates of the fault area, from the aftershock distribution, for instance, and from the surface ruptures
- The modelling of the waveforms from broadband data to estimate the fault slip, the fault area and the earthquake mechanism.

The seismic moment is measured in Nm or Dyne/cm; therefore it provides an estimation of the energy of the earthquake. This means the seismic moment is proportional to the magnitude of the earthquake, according to empirical relations. For instance the Moment Magnitude can be computed as follows:

$$M_w = \frac{2}{3}\log_{10}(M_0) - 10.7 \qquad 1.1$$

International seismological institutions systematically compute and provide the magnitude for each significant earthquake and the relative focal mechanism (see for instance: http://web.ics.purdue.edu/~braile/edumod/MagCalc/AS1Results.htm). Despite the seismic moment is the measurement of the energy release, i.e. a physical quantity, its estimation is affected by an uncertainty of the correct estimation of the geometric source parameters. The latter, the fault area and the fault slip, can be estimated from the aftershocks and the surface offsets. As the earthquake localization is computed by measuring the P- and S-waves travel times, the latter parameters are

very accurate at local scale wherever the seismic network is dense, while at global level the epicentral location has tens up to hundred km errors being very sparse the overall seismic network.
In the recent years the growing diffusion of the InSAR technique and the GPS data provided a further approach to the estimation of the fault area and slip. Dedicated softwares (see chapter 6) to compute such parameters from SAR measurements have been developed and are currently used. However the seismological institutions in charge of the Seismic Moment and Magnitude computation still base their analysis on the broadband records.

2. SAR Interferometry

2.1 Introduction

The term Remote Sensing generally indicates a number of observation techniques where the sensor and the observed object (the target) have not a physical contact. The Remote Sensing of the Earth is based on the interaction of the electromagnetic energy at different wavelengths and the observed surface. The Remote Sensing systems can be basically classified as active and passive. The former are equipped with a transmitting system and receive the signal backscattered from the illuminated surface, the latter make use of the radiation naturally emitted or reflected by the Earth's surface. The active sensors are mostly RADAR (RAdio Detection And Ranging) systems, operating in the microwave region (between $\lambda \approx 1\mu m$ and $\lambda \approx 1m$) of the electromagnetic spectrum, while the passive are called Radiometers. Passive sensors record the portion of the electromagnetic spectrum ranging from Visible ($4.0 < \lambda < 7.0\ \mu m$) to Near Infrared (NIR, from $\lambda = 7.0$ to $\lambda \approx 2\ \mu m$) and Thermal Infrared (TIR, from $\lambda \approx 4\ \mu m$ to $\lambda = 20\ \mu m$).

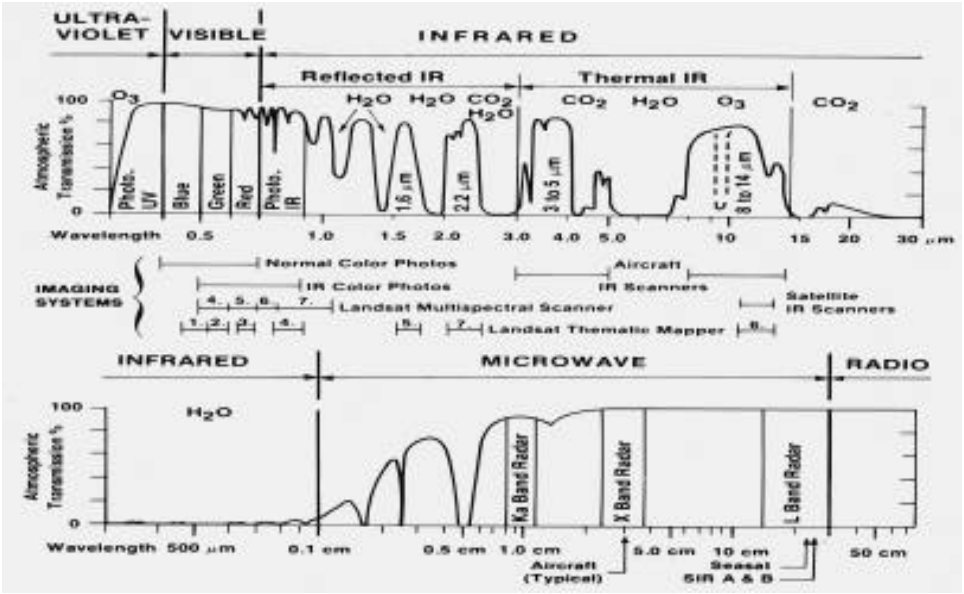


Figure 16: The electromagnetic spectrum.

The active sensors avoid some of the limitations of the passive ones. In particular, they are independent on sunlight; their frequency bands drastically reduce the impact of clouds, fog and rain. Therefore they can be considered time and weather independent.

An important characteristic of active and passive sensors is the spatial resolution, i.e. their capability to distinguish two points at a certain distance. The main limitation of Real Aperture Radar (RAR) is the poor spatial resolution they achieve with the working wavelengths. In fact the spatial resolution is the sensor-to-surface distance times the angular resolution. Moreover it is proportional to the ratio between the operating wavelength and the sensor antenna dimension.
Taking into account the previous relations it easily emerges how microwave sensors on satellite platforms, operating at hundred kilometres far from the target, should have antennas from hundred meters to some kilometres to reach reliable spatial resolution on the order of some meters.
The way for solving such strong limitation is to introduce the concept of synthetic antenna. The latter is obtained by exploiting the orbital movement of the antenna along the orbital path and then processing the received signal. The processing operations on both amplitude and phase components, let obtain the along-path resolution independent on the sensor altitude.

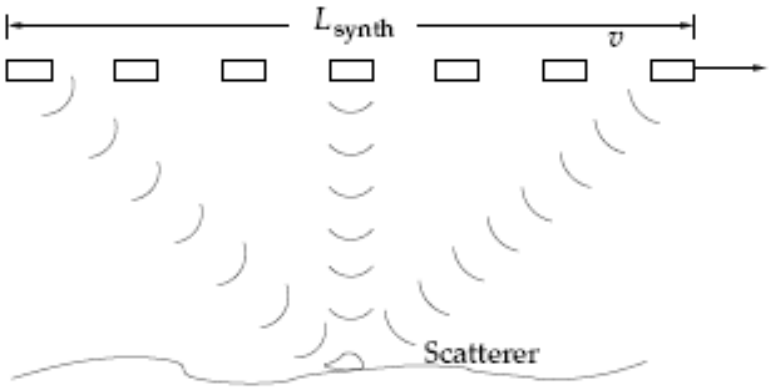


Figure 17: by exploiting the orbital path the synthetic antenna can be created. The scatterer is then observed by different angles. Therefore the spatial resolution is strongly increased.

The Synthetic Aperture Radar (SAR) is an active sensor based on this approach (Wiley, 1965). It maintains the same properties of the RAR systems, in particular the all-weather and all-time capabilities, but its high spatial resolution makes it a basic instrument for Earth Observation. The first SAR sensor was the SEASAT. In 1989 some researchers developed the SAR data processing technique known as SAR Interferometry. From 1992 (Landers Earthquake) this technique has been frequently applied to study the surface displacement field. The availability of new SAR satellites as the European ERS1 and ERS2 and recently the ENVISAT allowed a much wider availability of SAR images.

2.2 The SAR Interferometry (InSAR) technique

The acronym InSAR refers to a SAR data processing technique. It basically calculates the pixel-to-pixel phase differences between two SAR scenes of the same area acquired from analogue geometric conditions. The radar signal is transmitted with its own phase recorded by a local oscillator onboard. While the backscattered signal reaches the receiver it has been modified both in amplitude and in phase. This latter contains the data relative to the two way distance. Be S_1 and S_2 two satellite SAR sensors. Two acquisition geometries can be considered. The dual pass geometry means the satellites overpass the same area at different instants. On the contrary the single pass indicates the contemporaneous acquisitions. In both cases the InSAR technique can be applied to the received signals. In the following we refer to the dual pass mode. Let express the complex backscattered signal recorded at S_1 and S_2 as follows:

$$\begin{aligned} S_1 &= A_1 \cdot \varphi_{r_1}(r) \cdot e^{-j\frac{4\pi}{\lambda}r_1} \\ S_2 &= A_2 \cdot \varphi_{r_2}(r) \cdot e^{-j\frac{4\pi}{\lambda}r_2} \end{aligned} \qquad (2.2.1)$$

The previous equations, referred to Figure 3, contain two phase terms. The former is responsible for the contribution due to the surface backscattering properties (backscattering phase: $\varphi_r(r)$). The latter component is proportional to the sensor to target distance (propagation phase: $\frac{4\pi}{\lambda}r$). The backscattering phase component results from the sum of the elementary contributions of each scatterers of the resolution cell computed onto the plain perpendicularly oriented respect to the satellite Line Of Sight (LOS). To this aim the hypothesis is that the difference between the two sight angles $\hat{S_1 S_2}$ is very small. Therefore the acquisition geometry is substantially the same while the backscattered phase for each elementary scatterer is stable in S_1 and S_2.

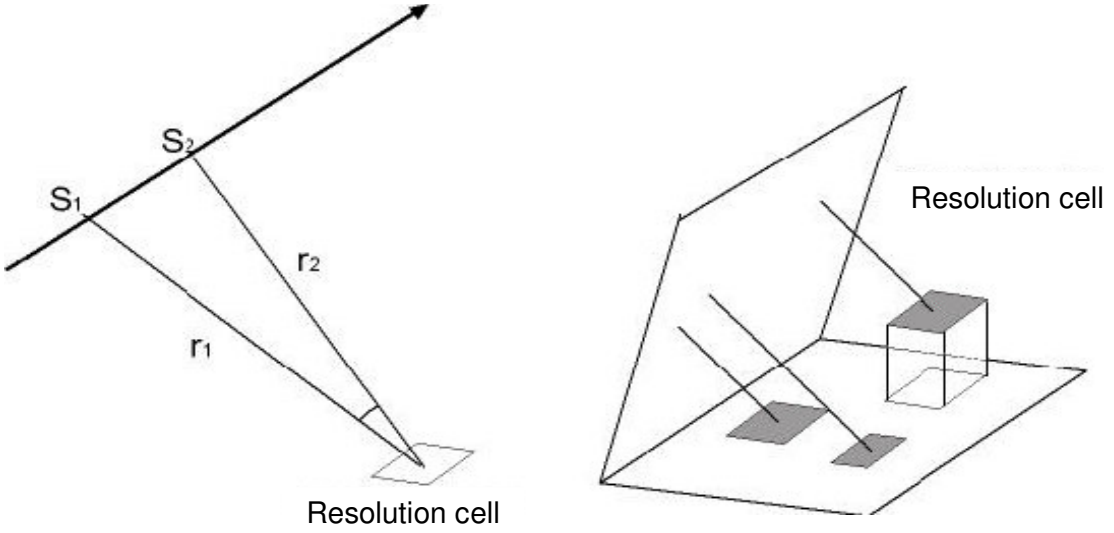


Figure 18: The backscattered phase is the sum of the elementary contributions from the scatterers of each resolution cell.

Based on the previous properties the interferometric phase image, or interferogram, is the map of the pixel-by-pixel differences between S_1 and S_2:

$$S_1 S_2^{*} = A_1 A_2 \cdot \left|\varphi_{r_1}\right\|\varphi_{r_2}\right| \cdot e^{j\left(\angle\varphi_{r_1} - \angle\varphi_{r_2}\right)} e^{-j\frac{4\pi}{\lambda}(r_1 - r_2)} \qquad (2.2.2)$$

The expression $A_1 A_2 \cdot \left|\varphi_{r_1}\right\|\varphi_{r_2}\right|$ is the signal amplitude, or modulus, of the received signal.

The information content of the interferometric phase is function of how much S_1 and S_2 remain similar. To this aim a quality factor, the complex correlation coefficient, can be computed. It provides a measurement of changes in the surface backscattering properties affecting two consecutive SAR images relative to the same area. This quality factor has the following expression:

$$\psi = \frac{E\left[S_1 S_2^{*}\right]}{\sqrt{E\left[S_1 S_1^{*}\right] \cdot E\left[S_2 S_2^{*}\right]}} \qquad (2.2.3)$$

being $E[\ ]$ the expected value.

The $|\psi|$, known as *coherence*, is a basic parameter for InSAR technique being used to provide a change index for comparing two or more SAR scenes.

Let forget the amplitude term and the backscattering phase component in (2.2.2). Thus (2.2.2) becomes a function of the propagation phase that can be assumed as the interferometric phase φ_{int}.

The latter schematically may contain some different contributions:

$$\varphi_{\text{int}} = \varphi_f + \varphi_{topo} + \varphi_{displ} + \varphi_{atm} + \varphi_{err} \qquad (2.2.4)$$

φ_f is the flat Earth interferometric phase, φ_{topo} the topographic component, φ_{displ} the displacement phase (accounting for surface movements), φ_{atm} is the atmospheric phase and finally φ_{err} accounts for possible error contributions. Each phase change corresponding to 2π is called interferometric fringe, or simply fringe.

See the simplified scheme in Figure 4 below summarizing the SAR interferometry acquisition geometry. Let consider two SAR satellite sensors S_1 and S_2 "looking at" the same area from quite similar geometries. The baseline B is the distance between S_1 and S_2 whilst over passing the same target.

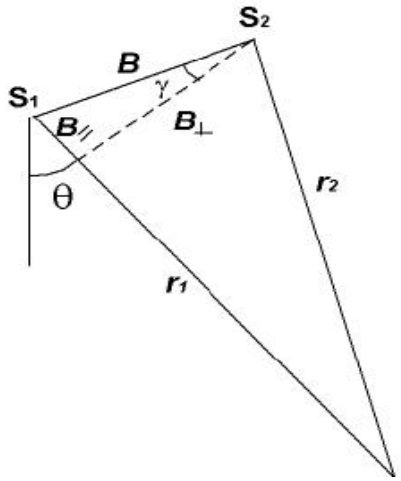


Figure 19: SAR sensors acquisition geometry. The baseline b is the distance between S_1 and S_2.

Let us compute the term φ_f from expression (4). The r_2 can be written as follows:

$$r_2 = \sqrt{r^2 + B^2 - 2r_1 B \cdot \cos\left(\frac{\pi}{2} - \gamma\right)} \cong r_1 - B\sin\gamma \qquad \text{being } B \text{ the } \textit{baseline} \quad (2.2.5)$$

Therefore the φ_f is:

$$\varphi_f = \frac{4\pi}{\lambda} \cdot B \cdot \sin\gamma \qquad (2.2.6)$$

The sensor-to-surface range distance varies for each pixel due to the SAR geometry along range direction. Consequently the flat Earth interferometric phase is not constant also where the Earth surface is completely flat. It varies according to the sensor-to-surface distance and generates one or more orbital fringes. The latter strictly depends on the relative positions of S_1 and S_2, i.e. on their baseline. This means that the sharper the knowledge of the orbital parameters for both satellites the more complete the orbital fringes removal. The flat Earth phase is directly proportional to the spatial baseline, therefore the orbital fringes will increase according to it.

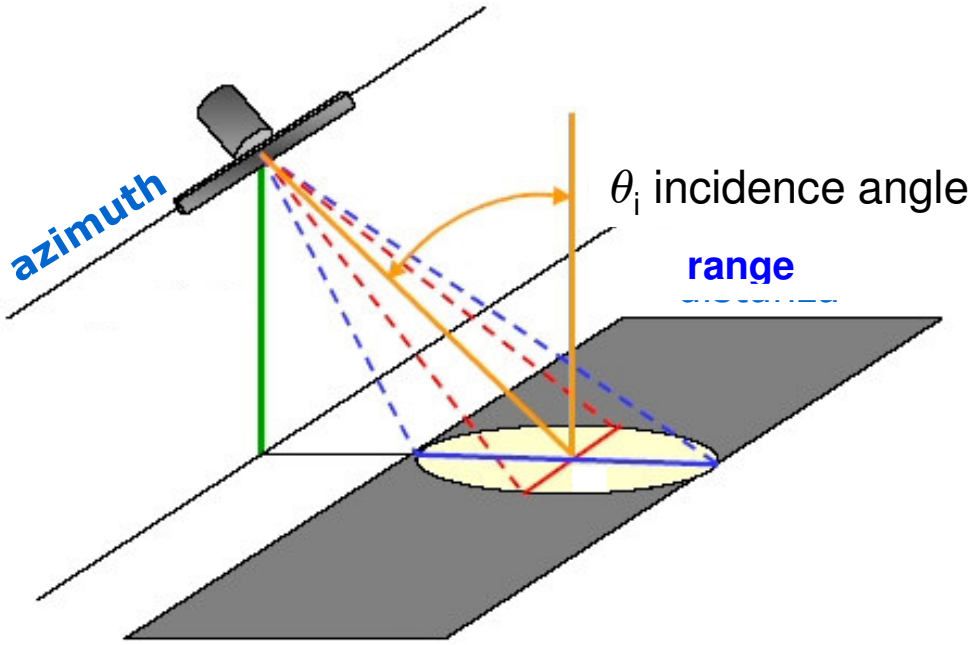


Figure 20: 3D SAR sensor geometry. θ is the incidence angle along the range direction

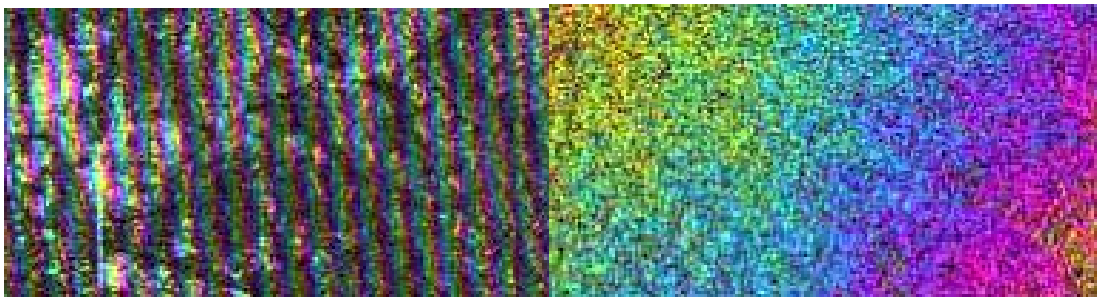


Figure 21: a) interferogram with flat earth component (orbital fringes) and b) after their have been removed

The term φ_{topo} is the phase component containing the topographic contribution to the interferometric phase. Be H the satellite altitude and z the target elevation. The other terms are the same in Figure 19.

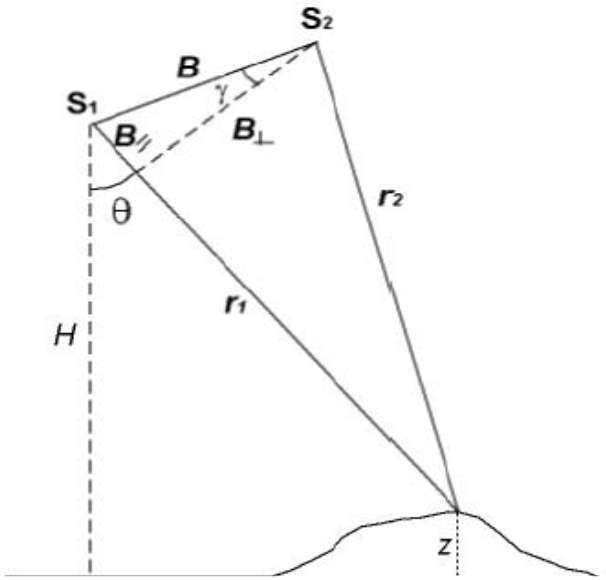


Figure 22: Topographic phase component relative to a target with elevation z.

The sensor-to-target distance is proportional to the propagation phase. Therefore the following expression:

$$\varphi = \frac{4}{\lambda}(r_2 - r_1) \cong -\frac{4\pi}{\lambda} \cdot B \cdot \sin\gamma - \frac{4\pi}{\lambda} B_{\perp} \frac{z}{r_1 \sin\vartheta} = \varphi_f + \varphi_{topo} \qquad (2.2.7)$$

is the sum of the flat Earth and the topographic phase.
As clearly stated in (2.2.7) the topographic term depends on the orthogonal component of the baseline. This means the capability of InSAR technique to detect and measure elevation changes is as much higher as the distance between S_1 and S_2 increases. To this issue refer to Figure 23. The same topography is represented by a different fringe number according to the orthogonal baseline that has a lower value.

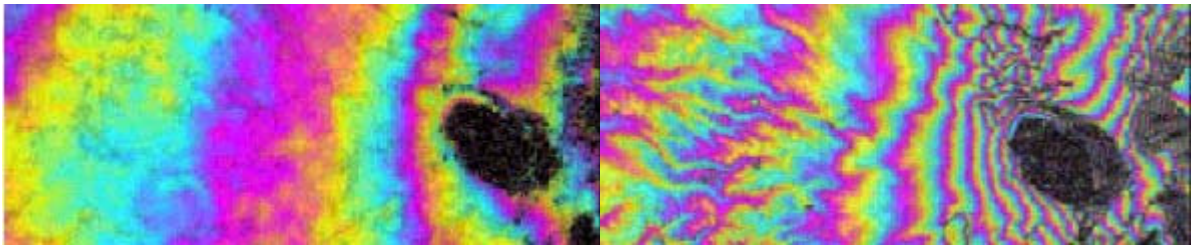


Figure 23: InSAR interferograms at different spatial baselines, 174m in (a) and 40m in (b). Each color cycle corresponds to an interferometric fringe.

As far as the ambiguity height h_a is concerned it corresponds with the topographic variation for each phase cycle 2π . Be:

$$\varphi_{topo} = \frac{4\pi}{\lambda} B_{\perp} \frac{z}{r_1 \sin\vartheta} \tag{2.2.8}$$

with $z = H - r_1 \cos\vartheta = H - \cos\vartheta \cdot \frac{\left(\rho^2 - B^2\right)}{2 \cdot \left(\rho - B\sin\gamma\right)}$

and $\rho = r_2 - r_1 = \frac{\lambda\varphi}{4\pi}$

Deriving both members of (8) respect to the phase:

$$\left.\frac{\partial z}{\partial \varphi}\right|_{2\pi} = \frac{\lambda \cdot r_1 \sin\vartheta}{4\pi \cdot B_{\perp}} \cdot \left.\frac{\partial \varphi_{topo}}{\partial \varphi}\right|_{2\pi} \tag{2.2.9}$$

$$h_a = \frac{\lambda \cdot r_1 \sin\vartheta}{2 \cdot B_{\perp}}$$

It clearly appears that how higher the *ambiguity height* h_a as lower the $B_{\perp}$. This means that increasing the distance between the two satellites while acquiring the images, the sensitivity of the interferometric phase to the topographic changes is also improved.

2.3 The Differential SAR Interferometry (DInSAR)

When the target of InSAR use is the measurement of surface displacements we talk about DInSAR technique. It has been showed in (2.2.4) that the interferometric phase is schematically described as the sum of some contributions, some of which are considered noise ($\varphi_f + \varphi_{atm} + \varphi_{err}$), the two others ($\varphi_{topo} + \varphi_{displ}$) contain information. However, while the main issue is the measurement of the surface displacement, the corresponding phase component (φ_{displ}) has to be exactly computed thus accurately cleaning the topographic phase. To this aim a Digital Elevation Model (DEM) can be used.

Suppose that a natural disaster (earthquake, volcanic eruption) producing a surface movement field affected a certain area and be S_1 and S_2 two SAR satellites overpassing such area before and after this event, respectively. The interferometric phase component φ_{displ} contains such displacement. In particular, it measures the projection $\delta r'$ of the effective displacement δr onto the satellite LOS (see Figure 24) as shown in (2.3.1).

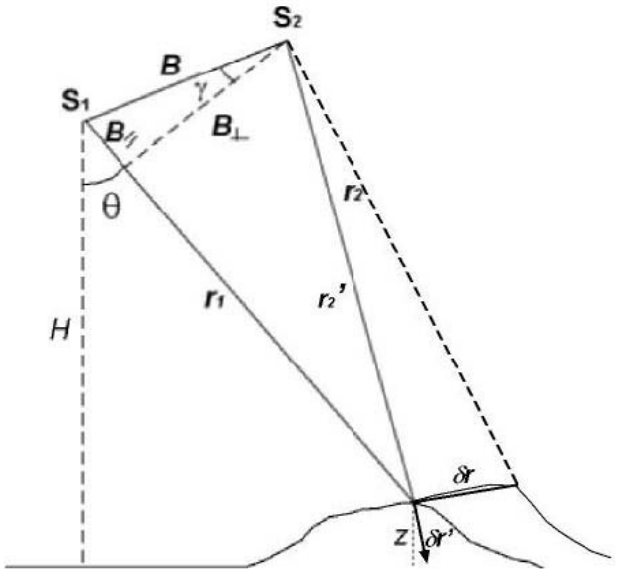


Figure 24: Interferometric phase computing: displacement component

$$\varphi_{displ} = \frac{4\pi}{\lambda} \cdot \delta r' \qquad (2.3.1)$$

2.3.1 Atmospheric effects

The interferometric phase φ_{atm} is the contribution due to the variation of the optical path of the electromagnetic wave due to the atmosphere. More precisely, as the interferometric phase is computed by considering two SAR images overpassing at different times the same area, changes in the atmospheric refraction index may occur due to different temperature, humidity and pressure conditions. Therefore the sensor-target propagation path may be affected thus resulting different for each image (see Figure 19). The quality of the computed interferometric phase image, the interferogram, decreases and parallely the accuracy of displacement measures worsens.

.

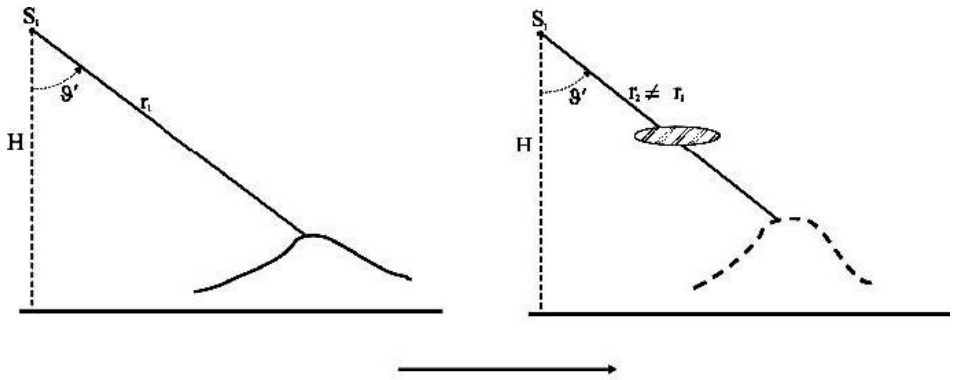
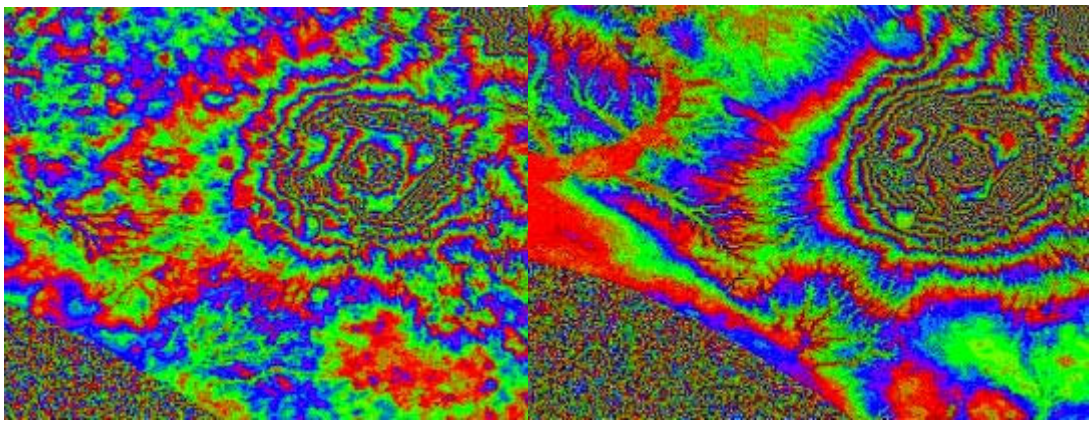


Figure 25: propagation path changes due to different atmospheric conditions of the two SAR images.

The spatial variability of such phenomenon is usually slow and uniform. The coherence does not allow any further consideration being the estimation windows too small for letting appreciable any refraction index variation.



a) b)

Figure 26: interferograms with (a) and without (b) the atmospheric phase contribution.

2.3.2 Phase ambiguity

The interferometric phase in (2.2.4) is the wrapped phase and is known except for multiples of 2π. Such ambiguity avoids the complete knowledge of the phase in InSAR or DInSAR applications. Let express the problem in analytical terms. Be:

$$\varphi_{unwrapp} = \varphi_{wrap} + 2N\pi \qquad (2.3.2.1)$$

The wrapped phase (φ_{wrap}) is between $\pm\pi$. Phase unwrapping algorithms have been developed for computing the exact value of the wrapped one. This means to calculate the exact number N of multiples of 2π. This is a very basic step requiring the accurate knowledge of orbital and topographic data. Some phase unwrapping algorithms are currently used, from the branch cut to the sparse phase unwrapping.

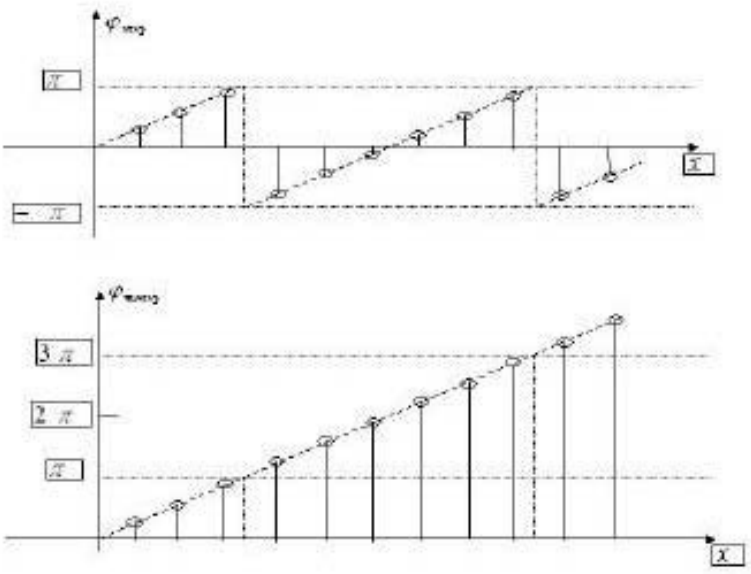


Figure 27: interferometric wrapped phase (above, known between $\pm\pi$) and unwrapped (figure below).

2.3.3 Geometric distortions

The acquisition geometry of the SAR sensor is usually right side looking oriented. Therefore geometric distortions are basically present in the collected data. For instance, ERS1 and ERS2 satellites are right side looking, having their Line Of Sight 23° far from the vertical.

Be τ the duration of the transmitted signal and θ the sight angle, the spatial resolution along the slant range direction in the hypothesis of flat terrain can be as follows (see Figure 20):

$$r_d = \frac{c\tau}{2} \qquad c = 3 \cdot 10^8 \text{ m/s.} \qquad (2.3.3.1)$$

By projecting the range resolution onto the ground the result is:

$$r_p = \frac{c\tau}{2\sin\theta} \qquad (2.3.3.2)$$

It is noteworthy that the point elevation is negligible because of the capability of the SAR sensor of measuring distances with high accuracy, while, on the contrary, it is unable to understand if a target point is at sea level or not. As a matter of fact the SAR sensor registers the two way time (sensor-target-sensor) thus measuring the path distance.

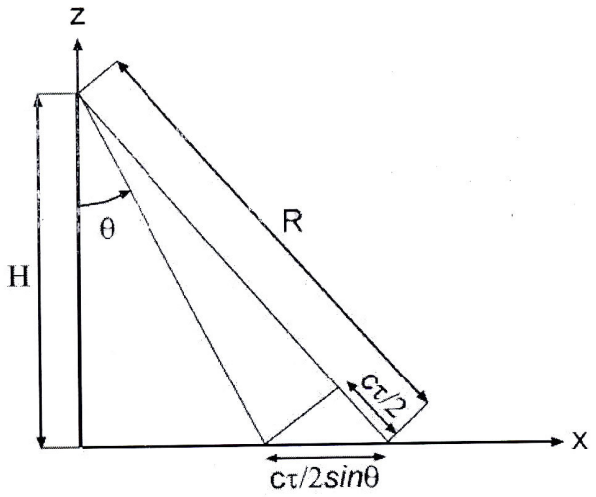


Figure 28: SAR range geometry and spatial resolution

The (3.3.3.1) and (3.3.3.2) have been computed under the simplifying assumption that the locale slope is null. In a real case however be α the slope angle; the range resolution becomes:

$$r_d = \frac{c\tau}{2 \cdot \sin(\vartheta - \alpha)} \qquad (2.3.3.3)$$

The previous equation is a function of the relation between the sight angle θ and the local slope α. Accounting for such angles the geometric distortions can be classified as follows (see Figure 21):

- Foreshortening. Due to the presence of a low slope relief along the slant range direction the resolution cell. The relation between the sight angle and the local slope is $(-\vartheta \le \alpha \le \vartheta)$;
- Layover. Due to the presence of a high slope relief. The backscattered signal from the top is then registered at the receiver before the one from the vally floor. The relation between the sight angle and the local slope is $(\alpha \ge \vartheta)$;
- Shadowing. Part of the relief is in shadow. The relation between the sight angle and the local slope is $\left(\alpha \le \vartheta - \frac{\pi}{2}\right)$.

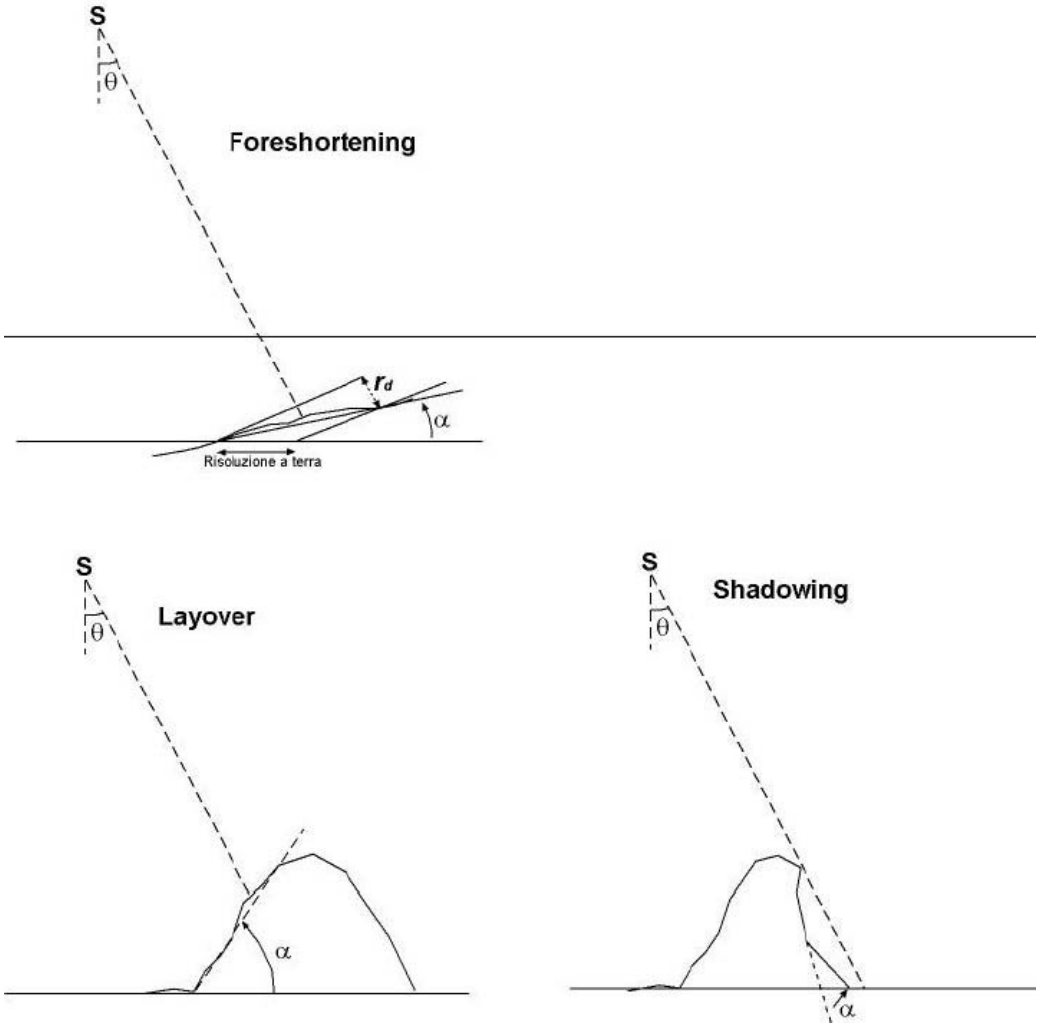
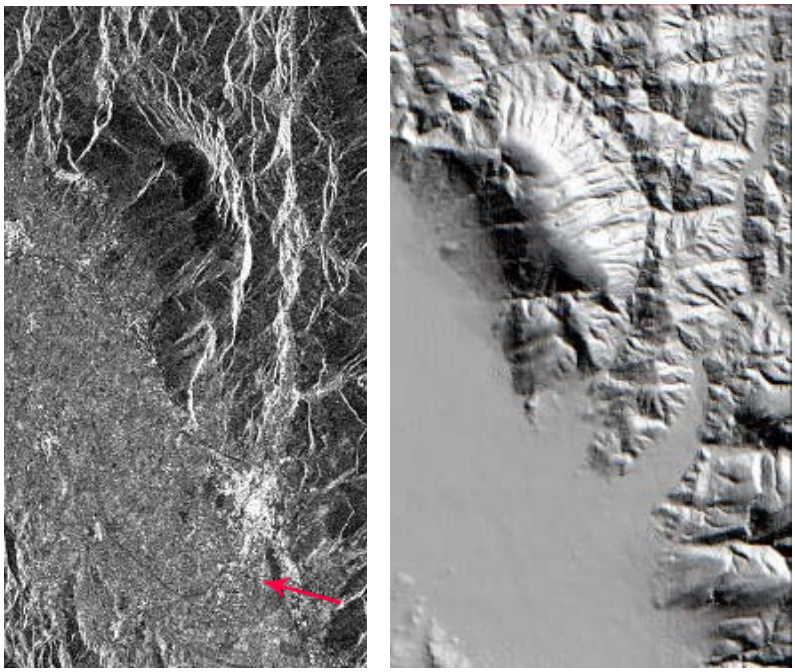


Figure 29: Geometric distorsions due to the combination between the SAR acquisition geometry and the topographic relief. Due to the foreshortening effect the distances heavily change. The layover distorsions are mainly present wherever the SAR sensor is looking at a steep relief. Finally the shadowing is caused by the shadows due to the geometry of the target combined with the satellite position.

In figure below is the comparison between a SAR intensity image and a DEM from 1:25.000 scale digital cartography of the Monte Subasio (Umbria, Central Italy). The geometric distortions of the SAR image are apparent.



a) b)

Figure 30: a) ERS1 image and b) DEM. The red arrow indicates the SAR Line Of Sight (LOS) for the descending orbit. The comparison with the DEM clearly highlights the geometric deformations.

2.4 Other factors influencing the interferometric phase

The SAR signal and then the interferometric phase can be affected by problems schematically summarized in two groups:

- problems due to the interferometric technique;
- instrumental problems.

Among the factor deling with the interferometric technique one may cite:

- Geometric decorrelation. It is due to the different sight angles of the two SAR acquisitions for each interferometric pair. The spatial correlation can be expressed as follows:

$$\left|\gamma_{spatial}\right| = 1 - \frac{2\cos\vartheta\left|\delta\theta\right|R_y}{\lambda} = 1 - \frac{2\left|B\right|R_y\cos^2\theta}{\lambda r} \qquad (2.4.1)$$

where R_y is the ground range resolution and θ is the sight angle.

The equation (26) shows that the spatial correlation linearly grows with the spatial baseline. In fact as larger is the difference between the two satellite points of view as greater is the difference between the electromagnetic answers of the SAR scenes. Therefore the *critical baseline* B_c is defined as the maximum spatial distance between two SAR acquisitions to generate a reliable interferometric phase.

$$B_c = \frac{\lambda r}{2R_y \cos^2 \theta} \qquad (2.4.2)$$

The previous equation indicates that the $|\gamma_{spatial}|$ is linearly variable with the baseline from 1 (baseline 0) to 0 (for baseline equal to B_c). It is noteworthy how the critical baseline is depending on almost all the parameters: B_c is directly proportional to the wavelength λ and the slantrange distance *r* whilst is inversely proportional to the range resolution R_y and the sight angle θ.
The geometric decorrelation can be also explained based on a frequency approach. Due to the different angles the spectra of the received signals S_1 and S_2 have a frequency shift Δf. As larger becomes the angle ε (i.e. the distance of S_1 and S_2 increases) the overlapping between the bands of the two received radar signals tends to reduce thus leading to a complete decorrelation.

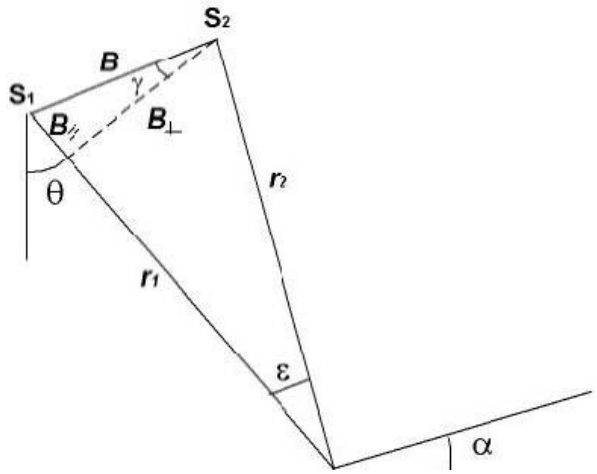


Figure 31: S_1 and S_2 SAR acquisition geometry. As the difference ε between the two sight angles increases the measured backscattered signals S_1 and S_2 decorrelate.

The phase shift can be expressed as follows:

$$\Delta f = -\frac{f \cdot \varepsilon}{tg(\vartheta - \alpha)} \qquad (2.4.3)$$

where α is the local slope angle.
It is also:

$$tg\varepsilon \cong \frac{\cos(\vartheta)}{\frac{r_1}{B} - \sin(\vartheta)} \qquad (2.4.4)$$

- Temporal decorrelation. The backscattered electromagnetic signal from the surface can be variable due to atmospheric events, or to soil exploitation for farming, or simply due to the vegetation growth.

- DEM generation errors due to possible topographic residual phases.

- Orbital data inaccurate knowledge. The latter can have a relevant impact on the interferometric processing.

Concerning the instrumental problems let's consider the following:

- Thermal instrumental noice. A thermal decorrelation index can be defined:

$$\psi_{termico} = \frac{1}{1 + SNR^{-1}}$$

- Radar frequency instability.

- Pulse Repetition Frequency (PRF) variation.

3. Biologic and artificial neurons

The biologic neuron is composed by a cellular body containing the core. Moreover it has one or more dendrites reaching the signals coming from outside, i.e. from other neurons. It contains also an axon transmitting the signal from the neuron outside. Finally, the synapse is the contact point between neurons.

The neuron may have two conditions, active while transmitting a signal towards outside, or not active. The signal exchange between two neurons takes place by means of an electrochemical mechanism based on releasing and absorbing the neurotransmitters. The signals received from outside may have two distinct effects: exciting and inhibitory. According to which dendrite has received the signal a different weight is associated to it. Then, all received and weighted signals are added. Whenever the exciting level exceeds a fixed threshold (activation threshold) the neuron becomes active and let flow through the axon the outgoing signal.

The biologic neuron is the basic component of the nervous system. The capability of a nervous system derives from the complexity of the nervous system itself, in terms of number of neurons and connections. The capabilities of the neurons are developed by the learning that allow modifying the weights associated to each signal and the threshold of each neuron by means of the interaction with outside.

The artificial neural network is planned to simulate some properties of the behaviour of the biologic neuron. The artificial neuron is composed by an activation function, one or more input connections (each of them with his weight) receiving the incoming numerical signals (Fig. 32 Moreover it has one or more output connections driving the signal towards outside.

The weights may have positive values for input exciting connections, negative values for inhibitory connections, null for null connections.

The artifical neuron

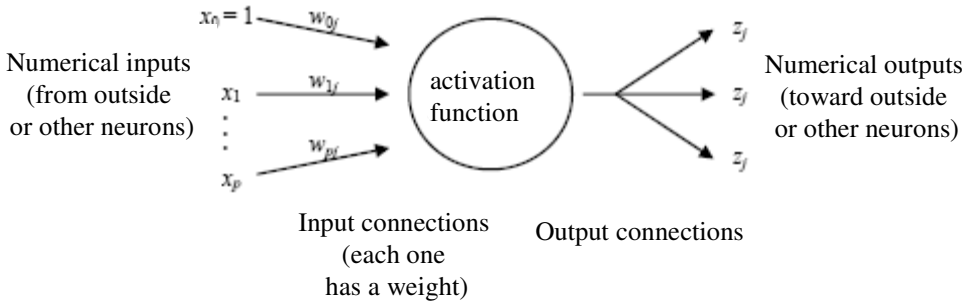


Figure 32: schematic structure of the artificial neuron.

3.1 Application fields of artificial neural networks

The neural networks are able to identify complex input-output connections, whose structure is unknown. However some examples concerning such connections are available.
Therefore the neural networks are typically used when a huge amount of data is available. Based on such data the neural network will allow extracting synthetic conclusions. The following applications are among the most suitable ones:

- Classification: in such problem solving a supervisor is present. According to its own experience it decides which class the input belongs to. The first step is a training phase using a dataset of examples. Afterwards the neural network shall be able to execute a classification as much as possible like the supervisor's one.
  The network implements a function F(X,q) splitting the input space X in N regions each belonging to a different class. This function depends on the parameters fixing the neural network and present in q. The performance of the neural network is measured by statistical quantities as the classification error or the root mean square error.
  Be *N* the number of examples the classification error is:

$$e_i = \begin{cases} 1 & F(X_i, \theta) \neq y_i \\ \\ 0 & F(X_i, \theta) = y_i \end{cases} \qquad (3.1.1)$$

being $E = \sum_{i=1}^{N} e_i$

The root mean square error is: $E = \frac{1}{N} \sum_{i=1}^{N} [F(X_i, \theta) - y_i]^2$ 3.1.2)

- Functional approximation: the *N* examples refer to an unknown process having *y* as output while the input exciting process is *X*. The function $F(X_i, \theta)$ implemented by the neural network is a modelling function of such process. The modelled *y* should be as much close to the value of *y* obtained if the input *X* is different from the available examples. The root mean square error provides an estimate of the effective error. The classification can be considered a particular case of approximation.

- Regression: the set of *N* examples is relative to an unknown process generating an output *y* whenever the input is *X*. The relationship is defined by the unknown conditioned distribution $\eta_X(y)$. Then the following regression is defined:

$$r(X) = \int y \cdot d\eta_X(y) \quad (3.1.3)$$

- Density estimation: the dataset is the space *X;* the supervisor is not present. The examples refer to an unknown density *P(X)* to be modelled by means of a neural network having a suitable structure. The parameters are contained in the vector *q*. The performance of the network has to be evaluated comparing the *P(X)* and the estimated density *p(X, q).*

- Prediction: it concerns the prevision of $f(t_0+d)$ concerning an unknown temporal process *f(t),* beginning from a certain number of samples at times $t \leq t_0$. "*d*" is the prediction distance.

The definition of the characteristics of a neural network deals with the choice of the best solution for three basic issues:

- The structure of the network: the architecture of the network should be customized to the work
- The learning set: it contains information concerning what the network should do. Its role is let the neural network improve its experience from the available examples belonging to it. The main issue is to understand if the latter set contains all the necessary information or if, on the contrary, such information is redundant.
- The learning algorithm: it transfers the information of the learning set to the neural network architecture for opportunely fixing the parameters.

The previous issues are strictly connected with the choice of the activation function. This latter defines the output signal function of the numerical signals collected by the input connections, based on the weights and on the activation threshold of the neuron.
The activation function connects the linear combination of the inputs and the output. This function can be split in the two components below:

1- the net potential (combination function)
2- the transfer function

The net potential P_j allows defining how the signals coming from the connections are compared and then combined with the activation level of the neuron. Among the possible combination functions the most used is the sum:

$\sum_{h=1}^{p} w_{hj} x_h$. The function product, $\prod_{h=1}^{p} w_{kj} x_h$, the function $\max_h w_{hj} x_h$ and $\min_h w_{hj} x_h$ are further reliable choices.

Now, the input signals are weighted and combined. Afterwards they are compared with the activation level of the neuron $\vartheta_j : P_j = \sum_{h=1}^{p} w_{hj} x_h - \vartheta_j$. Therefore if an input connection transmitting a constant signal whose weight is $w_{0j} = -\vartheta_j$ is added, the net potential results: $P \sum_{h=1}^{p} w_{hj} x_h$.

The transfer function $z_j = f(P_j)$ is able to transform the net potential and to generate the output signal. It is usually chosen among one of the following:

1- the step function: it may reach two values depending on the net potential higher or minor than zero. The step function can be binary, $f(P_j)=\begin{cases}0 & P_j<0\\ 1 & P_j\geq 0\end{cases}$ (3.1.4)

or bipolar, $f(P_j)=\begin{cases}-1 & P_j<0\\ 1 & P_j\geq 0\end{cases}$ (3.1.5)

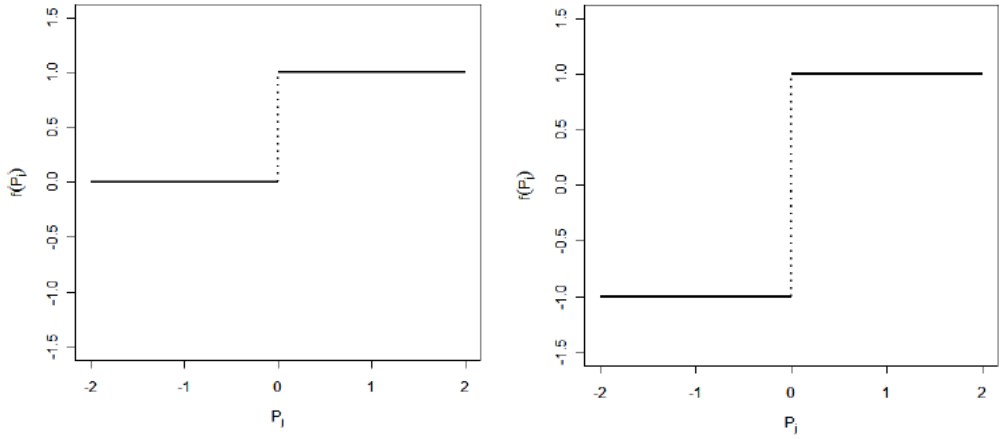


Figure 33: a) binary function; b) bipolar function.

2- the ramp function: this is a combination of a step function and a linear one. It is also known as linear saturated function:

$$f(P_j)=\begin{cases}-\gamma & P_j<\gamma\\ P_j & |P_j|<\gamma\\ y & P_j\geq\gamma\end{cases} \quad (3.1.6)$$

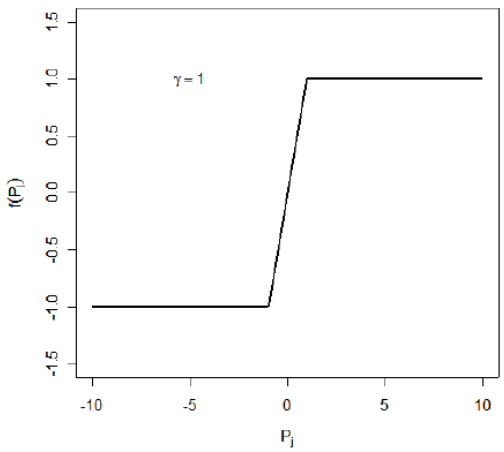


Figure 34: the ramp function

3- the sigmoid function: it is a monotone growing function, it is a continuous and differential function and it asymptotically tends to an upper and a lower limit. It may assume the following expression and behaviours:

$$f(P_j) = \frac{e^{P_j}}{1 + e^{-P_j}} \qquad (4.1.7) \qquad\qquad f(P_j) = \tanh(P_j) = \frac{e^{P_j} - e^{-P_j}}{e^{P_j} + e^{-P_j}} \qquad (3.1.7)$$

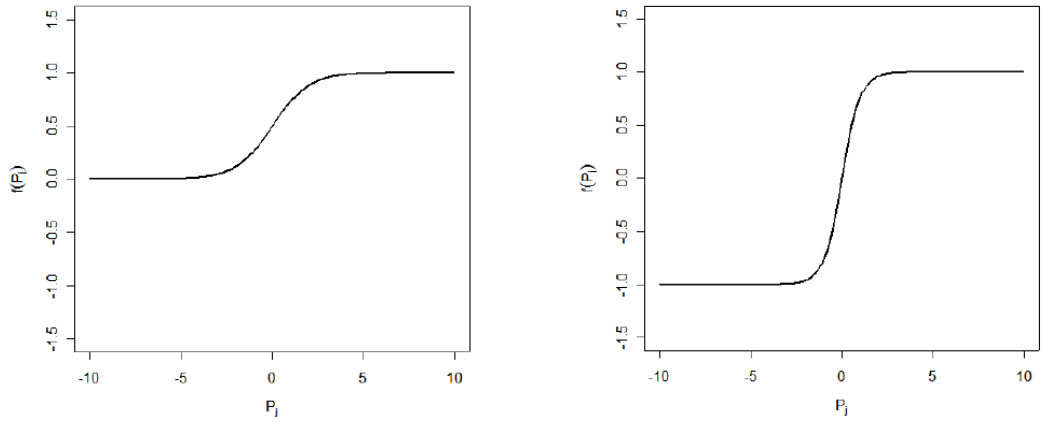


Figure 35: a) and b) sigmoid functions

3.2 Architecture of an artificial neural network

The organized structure of the neurons belonging to an artificial neural network is called topology or architecture of the net. The neurons are structured in layers, input and output layer for instance. These latter are the contact points with outside. Between the input and output one or more hidden layers are also present.

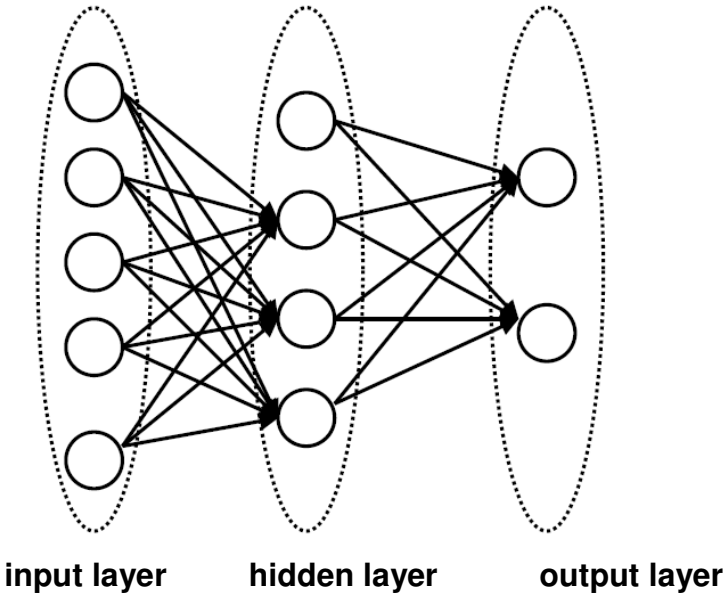


Figure 36: schematic architecture of an artificial neural network

The input layer is composed by a number of artificial neurons equal to the number of input signals. Each neuron belonging to this layer receives only one external signal and does not generate any kind of transformation. The output layer let transmit outside the signals from the net. The input and the output layers of a self associative net coincide whilst it does not happen for a hetero associative one.

Concerning the connections between neurons, whenever are between neurons of different layers are called interlayer connections whilts are intralayer connections when refer to neurons of the same layer.

A neural network is totally connected if each neuron is connected to all the others. A neural network is interconnected if each neuron is connected to all the others belonging to the previous and to the following layers.

Concerning the direction of the flux of the signals, a partially or totally connected net is feed-forward if the connections are one-way, i.e. from a layer to the following one. The net is feed-back if the signals are two-way.

3.3 Learning

The main property of artificial neural networks is their capability to learn. In other words they are able to improve their experience. This property is connected to the algorithm used for the learning process.

The quality of the learning algorithm is evaluated and tested by means of unknown examples. This allows defining the generalization capability of the net.

Two different learning approaches can be defined: supervised and unsupervised. Regarding the supervised approach the network answer is compared with the desired one (supervision). Consequently the parameters are modified to reduce the differences. The way for measuring the shift and the learning strategy are basic aspects for this approach.

The following scheme summarizes the logic of a supervised net.

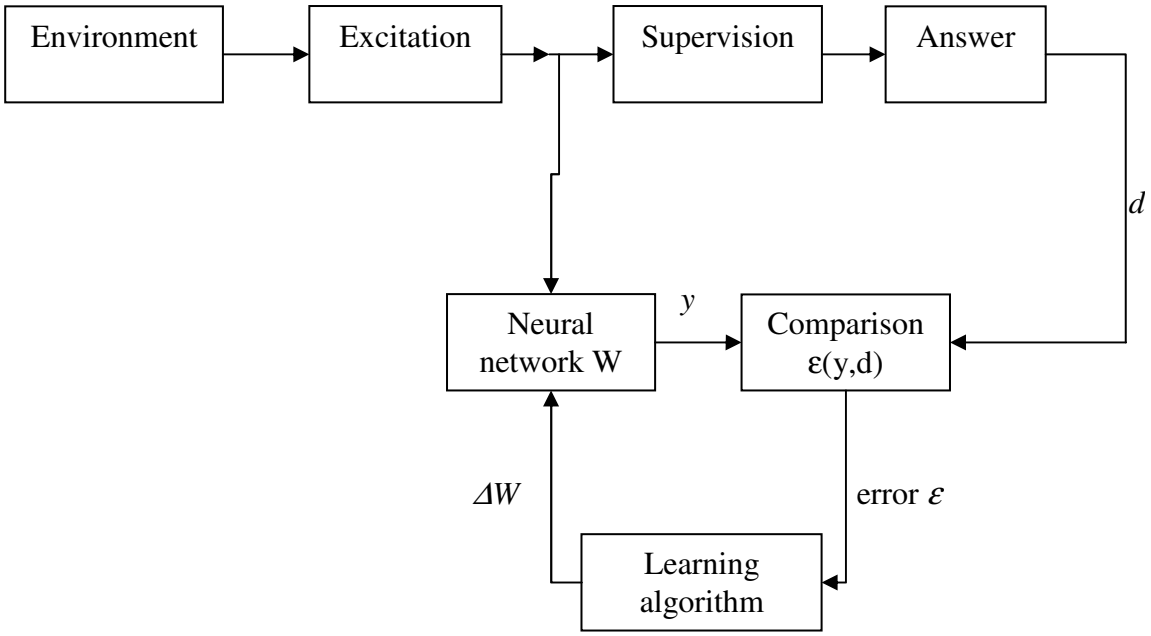


Figure 37: block diagram of the supervised neural network. Be *X* the input, *y* the output of the network, *d* the desired (supervision process) output, *W* is the parameter vector and *ΔW* the correction of the learning algorithm.

On the other side the unsupervised neural networks are based on a more general laws dealing with the natural selection and addressing the output. Among such laws the Hebbian mechanism and the competitive learning

3.3.1 Hebbian learning

In the Hebbian learning approach the synaptic link between two neurons is enforced whenever they are simultaneously activated whilst it weakens otherwise. The derived algorithm corrects the weight of the connection between two artificial neurons according to the correlation of their activities: as more positive the correlation so greater the weight.
Be *x(t)* and *y(t)* the neuron activities, *w* is the connection weight, the correction to the network is:

$\Delta w = \eta \cdot \mathrm{cov}\{y(t), x(t)\}$ with positive η. The simplified form $\Delta w = \eta \cdot x(t) \cdot y(t)$ can be also considered.
The correction *w* can generate an exponential growth up to the saturation. Let's take *y(t)* just depending on the neuron *x(t)*. If x_0 is constant and being *w* small: $dw = \eta \cdot y(t) \cdot x_0 = \eta \cdot w \cdot {x_0}^2$.
Therefore is: $w(t) = w(0) \cdot e^{\eta \cdot {x_0}^2 \cdot t}$. (3.3.1.1)

The saturation can be avoided by accounting for a corrective factor, a forgetfulness factor, controlling the weight growth:
$\Delta w = \eta \cdot x(t) \cdot y(t) - \alpha \cdot y(t) \cdot w(t)$. (3.3.1.2)
Therefore starting from the previously cited hypothesis:
$dw = \eta \cdot w \cdot {x_0}^2 - \alpha \cdot w^2 \cdot x_0$, (3.3.1.3)
with: $w(t) = \frac{\eta \cdot x_0}{\alpha} \cdot \frac{1}{1 - A \cdot e^{-\eta \cdot {x_0}^2 \cdot t}}$ and $A = 1 - \frac{\eta \cdot x_0}{\alpha \cdot w(0)}$.

Then, the weight *w(t)* asymptotically tends to a positive value $\frac{\eta \cdot x_0}{\alpha}$, starting from a positive *w(0)*.

3.3.2 Competitive learning

The competitive learning implements a natural selection favouring just one neuron. The learning logic in this case is favourable to the more active neuron.
Consider a neural network having *N* neurons connected with *M* inputs, both with binary values

$\{0,1\}$. Each neuron is connected to all inputs. Be $A_j = \sum_{i=1}^{M} w_{ij} u_i$ the excitation value of the *i-th* neuron, u_i is the activity of the *i-th* input and w_{ij} the weight of the connection between the input *i* and the neuron *j*. The following bonds characterize the network: $\sum_{i=1}^{M} w_{ij} = 1$ and $w_{ij} \geq 0$.

The competitive mechanism can be summarized as follows:

1- The *m-th* excitation is applied as input

2- The neuron excitations ($A_j^{(m)}$) are defined and the maximum among them (corresponding with the *k-th* neuron) is selected

3- The weights relative to the *k-th* neuron are corrected using the following expression:

$$\Delta w_{ik} = g \cdot \frac{u_i^{(m)}}{n^{(m)}} - g \cdot w_{ik} \ . \qquad (3.3.2.1)$$

In the latter 0<g<1, $u_i^{(m)}$ is the activity of the *i-th* input beginning with the *m-th* excitation, $n^{(m)} = \sum_{i=1}^{M} u_i^{(m)}$ is the number of active inputs referred to the *m-th* excitation.

4- To avoid the atrophy for some neurons corrective weights are applied also to them.

The competitive learning algorithm has some properties:

a- The weights are non negative:

$$w_{ik} + \Delta w_{ik} = (1-g) \cdot w_{ik} + g \cdot \frac{u_i^{(m)}}{n^{(m)}} \geq 0 \qquad (3.3.2.2)$$

b- The sum of the weights of the connections for a neuron is equal to 1:

$$\sum_{i=1}^{M} (w_{ij} + \Delta w_{ij}) = \sum_{i=1}^{M} w_{ij} + \sum_{i=1}^{M} \Delta w_{ij} = 1 \qquad (3.3.2.3)$$

3.3.3 The Occam principle

The base for learning is the capability of acquiring experience from examples. This is exactly what happens in the nervous system that, moreover, has also to improve the efficiency by optimizing the available resources.

This general principle, called the Occam principle, can be expressed as follows if applied to the neural networks: let consider two neural networks satisfying the learning set; the less complex of them will behave better when working on unknown case studies, i.e. it has greater generalization capabilities.

This principles applied to supervised neural networks suggests to use an objective function and to minimize it. The function is composed by two terms: the first E_s concerns the adequacy of the networks relatively to the objective; the second E_c its complexity. Be m the network parameters the objective function is as follows: $E(m)=E_s(m)+E_c(m)$.

4 The methodology

The idea to reproduce the behaviour of a seismogenic structure and to allow geophysicists to study its properties and understand its dynamics was first proposed by formulating the dislocation theory. Steketee (1958) applied such theory to the seismological fieldwork. Further pionieristic manuscripts (Rongvad and Frasier, 1958) developed a theoric formulation of the deformation in a homogeneous, isotropic and semi-infinite medium.

Such formulation takes its origin from the hypothesis of point source, with a vertical strike slip component, in a Poisson solid (Steketee, 1958). At depth, it corresponds with a stress field generated by a finite dimension fault, sloping and shear zone, in an elastic medium (Iwasaki and Sato, 1979). In the following years more efforts have been made allowing the model to be closer to the truth. To this aim the Earth curvature and the topography have been taken into account, while the crust has been considered inhomogeneous and multilayer. Furthermore the effects of the lateral disomogeneities together with the presence of sloping layers have been considered too. Anyway these complexities highlighted that the Earth curvature effect is usually negligible mainly for superficial sources. On the contrary the inhomogeneous (either vertically or horizontally) crust may strongly affect the deformation field.

Notwithstanding a huge number of studies and published manuscripts, the analysis of the answer of a medium to the seismic source is mainly based on the simplified hypothesis of an elastic and isotropic half space, while the source itself is geometrically very simple. Three are the reasons. First, the model needs simple boundary conditions to provide a comprehensive scenario and reproduce the behaviour of the medium. Secondly, the model is one of the possible, simplified copies of the reality. Finally, the data concerning the crustal movements are not enough to provide an accurate knowledge of the Earth upper layer. It is noteworthy that in the last years the availability of much more instruments allowed reaching an improved and more detailed knowledge of the Earth crust.

As far as the seismic source modelling is concerned a forward and an inverse approaches can be used. In the next paragraphs both methods are going to be described in details.

4.1 The forward problem.

With his article published in 1985 on Bulletin of the Seismological Society of America, Yoshimitsu Okada leaves from a wide review of the previous jobs, some of which cited above, in order to

propose an analytical formulation describing the superficial deformation due to a zone of cut (shear fault) in a half space. Okada focused on the thin seismogenic superficial layer, being his formulation relevant to practical applications. Moreover previous analytical descriptions are too long and complex in order to adequately answer to the requested target. For instance, Savage and Hastie (1966) proposed a not usable algorithm for vertical and horizontal faults, while Sato and Matsu' ura (1974) found a singularity in case the fault intercepts the surface. Moreover being these studies mainly relevant to the analysis of the shear deformation this latter carried to neglect the studies relative to the tensile component.

Anyway the formulation of Okada supplies a compact analytical instrument in order to calculate the superficial movements, the effort and the inclination due to the shear zone in the considered half space. Therefore it is widely adopted by the geophysicists for the seismic source modeling. Moreover for sake of simplicity it naturally introduces geometric limitations and requires the supply of the boundary conditions.

4.2 The model of Okada

The analytical models offer a description of the *coseismic static deformations*. Okada developed a set of equations for computing the surface deformation field due to a point or finite size, i.e. rectangular, source, for an isotropic and homogenous semi-infinite space, using the laws of the classic linear elasticity.

Steketee (1958) demonstrated that the displacement field $u_i\,(x_1, x_2, x_3)$ due to a dislocation $\Delta u_j\,(\zeta_1, \zeta_2, \zeta_3,)$ through a surface Σ in an isotropic medium is:

$$u_i = \frac{1}{F} \iint_{\Sigma} \Delta u_j \left(\lambda \delta_{jk} \frac{\partial u_i^n}{\partial \xi_k} + \mu \left(\frac{\partial u_i^j}{\partial \xi_k} + \frac{\partial u_i^k}{\partial \xi_j} \right) v_k \right) d\Sigma \qquad 4.1.1$$

with δ_{jk} delta of Kronecker, λ and μ the Lamé coefficients, v_k the cosine of the normal to the surface element $d\Sigma$, u_i^j the i-th displacement component of the point (x_1, x_2, x_3) due to the j-th direction of the force F applied in $(\zeta_1, \zeta_2, \zeta_3,)$ on the surface.

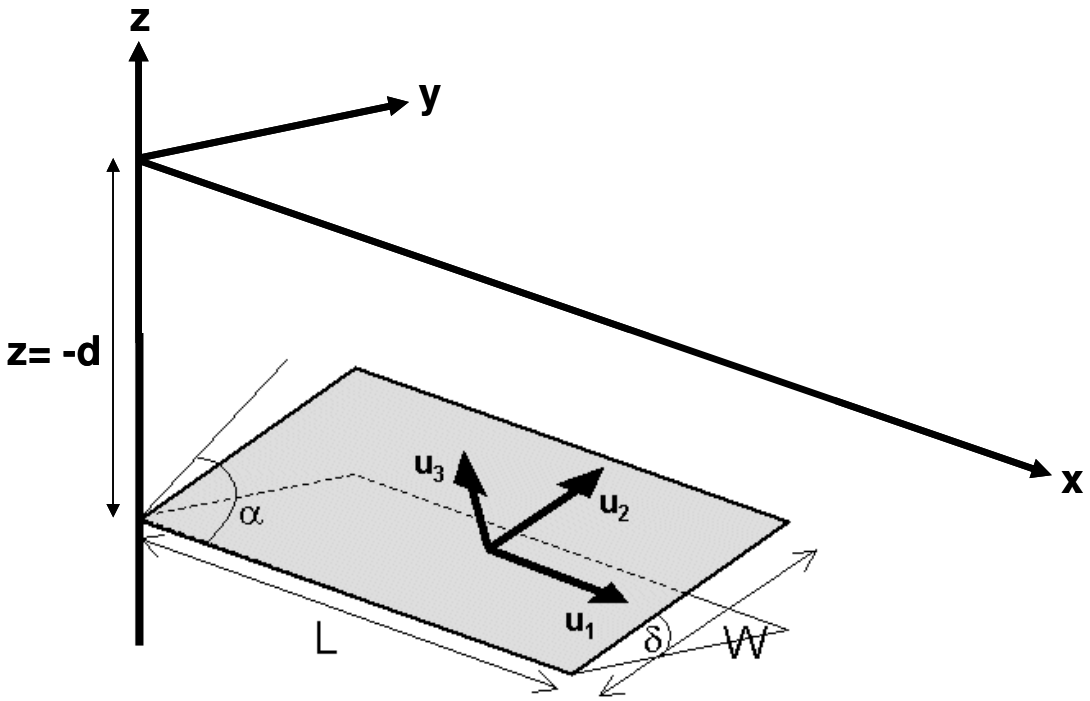


Fig. 38: Fault geometry and symbols.

In the Cartesian space in Fig. 34 the elastic medium is in the region with $z \leq 0$ and the x axis is parallel to the fault strike direction. The elementary dislocation components U_1, U_2 and U_3 are the stress fields along the strike-slip, the dip-slip and the tensile directions, respectively. In Fig. 34 each vector shows as the hanging-wall behaves respect to the foot-wall. Therefore whenever U_2 is oriented as in Fig. 34 the fault is a reverse one, or thrust fault, whilst it is a normal fault if $sen2\delta < 0$. In this coordinate system the abovementioned vectors u_i^j can be expressed as follows:

$$u_1^{\ 1} = \frac{F}{4\pi\mu}\left\{\frac{1}{R} + \frac{(x_1-\xi_1)^2}{R^3} + \frac{\mu}{\lambda+\mu}\left[\frac{1}{R-\xi_3} - \frac{(x_1-\xi_1)^2}{R(R-\xi_3)^2}\right]\right\}$$

$$u_2^{\ 1} = \frac{F}{4\pi\mu}(x_1-\xi_1)(x_2-\xi_2)\left\{\frac{1}{R^3} - \frac{\mu}{\lambda+\mu}\left[\frac{1}{R(R-\xi_3)^2}\right]\right\}$$

$$u_3^{\ 1} = \frac{F}{4\pi\mu}(x_1-\xi_1)\left\{-\frac{\xi_3}{R^3} - \frac{\mu}{\lambda+\mu}\left[\frac{1}{R(R-\xi_3)}\right]\right\}$$

$$u_1^{\ 2} = \frac{F}{4\pi\mu}(x_1-\xi_1)(x_2-\xi_2)\left\{\frac{1}{R^3}-\frac{\mu}{\lambda+\mu}\left[\frac{1}{R(R-\xi_3)^2}\right]\right\}$$

$$u_2^{\ 2} = \frac{F}{4\pi\mu}\left\{\frac{1}{R}+\frac{(x_2-\xi_2)^2}{R^3}+\frac{\mu}{\lambda+\mu}\left[\frac{1}{R-\xi_3}-\frac{(x_2-\xi_2)^2}{R(R-\xi_3)^2}\right]\right\} \qquad 4.2.1$$

$$u_3^{\ 2} = \frac{F}{4\pi\mu}(x_2-\xi_2)\left\{-\frac{\xi_3}{R^3}-\frac{\mu}{\lambda+\mu}\left[\frac{1}{R(R-\xi_3)}\right]\right\}$$

$$u_1^{\ 3} = \frac{F}{4\pi\mu}(x_1-\xi_1)\left\{-\frac{\xi_3}{R^3}+\frac{\mu}{\lambda+\mu}\left[\frac{1}{R(R-\xi_3)}\right]\right\}$$

$$u_2^{\ 3} = \frac{F}{4\pi\mu}(x_2-\xi_2)\left\{-\frac{\xi_3}{R^3}+\frac{\mu}{\lambda+\mu}\left[\frac{1}{R(R-\xi_3)}\right]\right\}$$

$$u_3^{\ 3} = \frac{F}{4\pi\mu}\left\{\frac{1}{R}+\frac{\xi_3^{\ 2}}{R^3}+\frac{\mu}{\lambda+\mu}\frac{1}{R}\right\}$$

In the previous equations it is $R^2 = (x_1-\xi_1)^2+(x_2-\xi_2)^2+\xi_3^{\ 2}$.

Taking into account the previous expressions in (4.1.1) the elementary dislocation for each surface element $\Delta\Sigma$ is provided by:

Displacements: strike-slip

$$\frac{1}{F}\mu U_1\Delta\Sigma\left[-\left(\frac{\partial u_i^{\ 1}}{\partial\xi_2}+\frac{\partial u_i^{\ 2}}{\partial\xi_1}\right)\sin\delta+\left(\frac{\partial u_i^{\ 1}}{\partial\xi_3}+\frac{\partial u_i^{\ 3}}{\partial\xi_1}\right)\cos\delta\right] \qquad 4.2.2$$

Dip slip

$$\frac{1}{F}\mu U_2\Delta\Sigma\left[\left(\frac{\partial u_i^{\ 2}}{\partial\xi_3}+\frac{\partial u_i^{\ 3}}{\partial\xi_2}\right)\cos 2\delta+\left(\frac{\partial u_i^{\ 3}}{\partial\xi_3}-\frac{\partial u_i^{\ 2}}{\partial\xi_2}\right)\sin 2\delta\right] \qquad 4.2.3$$

Tensile fault

$$\frac{1}{F}\mu U_3\Delta\Sigma\left[\lambda\frac{\partial u_i^{\ n}}{\partial\xi_n}2\mu\left(\frac{\partial u_i^{\ 2}}{\partial\xi_2}\sin^2\delta+\frac{\partial u_i^{\ 3}}{\partial\xi_3}\cos^2\delta\right)-\mu\left(\frac{\partial u_i^{\ 2}}{\partial\xi_3}+\frac{\partial u_i^{\ 3}}{\partial\xi_2}\right)\sin 2\delta\right] \qquad 4.2.4$$

The previous expressions state the body force equivalents of a double-couple with moment $\mu U_1\Delta\Sigma$ or $\mu U_2\Delta\Sigma$ in case of a shear fault. Moreover for tensile fault they represent the centre of dilatation with intensity $\lambda U_3\Delta\Sigma$, combined with a double-couple without moment and intensity $2\mu U_3\Delta\Sigma$.

4.2.1 Displacement vector components

The displacement vector along the strike-slip, the dip-slip and the tensile directions are provided by computing such components for a point like source in (0, 0, -d), by replacing the 4.2.2 in 4.2.3, 4.2.4 in the 4.2.1, being $\xi_1 = \xi_2 = 0$ and $\xi_3 = -d$:

Strike-slip direction

$$u_x^{\,0} = -\frac{U_1}{2\pi}\left[\frac{3x^2 q}{R^5} + I_1^{\,0}\sin\delta\right]\Delta\Sigma$$
$$u_y^{\,0} = -\frac{U_1}{2\pi}\left[\frac{3xyq}{R^5} + I_2^{\,0}\sin\delta\right]\Delta\Sigma \qquad 4.2.1.1$$
$$u_z^{\,0} = -\frac{U_1}{2\pi}\left[\frac{3xdq}{R^5} + I_4^{\,0}\sin\delta\right]\Delta\Sigma$$

Dip-slip direction

$$u_x^{\,0} = -\frac{U_2}{2\pi}\left[\frac{3xpq}{R^5} - I_3^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma$$
$$u_y^{\,0} = -\frac{U_2}{2\pi}\left[\frac{3ypq}{R^5} - I_1^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma \qquad 4.2.1.2$$
$$u_z^{\,0} = -\frac{U_2}{2\pi}\left[\frac{3dpq}{R^5} - I_5^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma$$

Tensile direction

$$u_x^{\,0} = \frac{U_3}{2\pi}\left[\frac{3xq^2}{R^5} - I_3^{\,0}\sin^2\delta\right]\Delta\Sigma$$
$$u_y^{\,0} = \frac{U_3}{2\pi}\left[\frac{3yq^2}{R^5} - I_1^{\,0}\sin^2\delta\right]\Delta\Sigma \qquad 4.2.1.3$$
$$u_z^{\,0} = \frac{U_3}{2\pi}\left[\frac{3dq^2}{R^5} - I_5^{\,0}\sin^2\delta\right]\Delta\Sigma$$

Be considered in the previous equations:

$$I_1^{\,0} = \frac{\mu}{\lambda+\mu} y\left[\frac{1}{R(R+d)^2} - x^2\frac{3R+d}{R^3(R+d)^3}\right]$$

$$I_2^{\,0} = \frac{\mu}{\lambda+\mu} x\left[\frac{1}{R(R+d)^2} - y^2\frac{3R+d}{R^3(R+d)^3}\right]$$

$$I_3^{\,0} = \frac{\mu}{\lambda+\mu}\left[\frac{x}{R^3}\right] - I_2^{\,0} \qquad 4.2.1.4$$

$$I_4^{\,0} = \frac{\mu}{\lambda+\mu}\left[-xy\frac{2R+d}{R^3(R+d)^2}\right]$$

$$I_5^{\,0} = \frac{\mu}{\lambda+\mu}\left[\frac{1}{R(R+d)} - x^2\frac{2R+d}{R^3(R+d)^2}\right]$$

$$p = y\cos\delta + d\sin\delta$$

$$q = y\sin\delta - d\cos\delta \qquad 4.2.1.5$$

$$R^2 = x^2 + y^2 + d^2 = x^2 + y^2 + q^2$$

4.2.2 Strain vector components

The displacement vector has to be derived along each axis for computing the three components of the strain vector.

For instance along the Strike slip direction:

$$\frac{\partial u_x^{\,0}}{\partial x} = -\frac{U_1}{2\pi}\left[\frac{3xq}{R^5}(2-\frac{5x^2}{R^2}) + J_1^{\,0}\sin\delta\right]\Delta\Sigma$$

$$\frac{\partial u_x^{\,0}}{\partial y} = -\frac{U_1}{2\pi}\left[-\frac{5x^2yq}{R^7} + (\frac{3x^2}{R^5} + J_2^{\,0})\sin\delta\right]\Delta\Sigma$$

$$\frac{\partial u_y^{\,0}}{\partial x} = -\frac{U_1}{2\pi}\left[\frac{3yq}{R^5} - (1-\frac{5x^2}{R^2}) + J_2^{\,0}\sin\delta\right]\Delta\Sigma \qquad 4.2.2.1$$

$$\frac{\partial u_y^{\,0}}{\partial y} = -\frac{U_1}{2\pi}\left[\frac{3yq}{R^5}(1-\frac{5y^2}{R^2}) + (\frac{3xy}{R^5} + J_4^{\,0})\sin\delta\right]\Delta\Sigma$$

Along the Dip-slip direction:

$$\frac{\partial u_x^{\,0}}{\partial x} = -\frac{U_2}{2\pi}\left[\frac{3pq}{R^5}(1-\frac{5x^2}{R^2}) - J_3^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma$$
$$\frac{\partial u_x^{\,0}}{\partial y} = -\frac{U_2}{2\pi}\left[\frac{3x}{R^5}(s-\frac{5ypq}{R^2}) - J_1^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma$$
$$\frac{\partial u_y^{\,0}}{\partial x} = -\frac{U_2}{2\pi}\left[-\frac{15xypq}{R^7} - J_1^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma$$
$$\frac{\partial u_y^{\,0}}{\partial y} = -\frac{U_2}{2\pi}\left[\frac{3pq}{R^5}(1-\frac{5y^2}{R^2}) + \frac{3ys}{R^5} - J_2^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma$$

4.2.2.2

Along the tensile direction:

$$\frac{\partial u_x^{\,0}}{\partial x} = \frac{U_3}{2\pi}\left[\frac{3q^2}{R^5}(1-\frac{5x^2}{R^2}) - J_3^{\,0}\sin^2\delta\right]\Delta\Sigma$$
$$\frac{\partial u_x^{\,0}}{\partial y} = \frac{U_3}{2\pi}\left[\frac{3xq}{R^5}(2\sin\delta-\frac{5yq}{R^2}) - J_1^{\,0}\sin^2\delta\right]\Delta\Sigma$$
$$\frac{\partial u_y^{\,0}}{\partial x} = \frac{U_3}{2\pi}\left[-\frac{15xyq^2}{R^7} - J_1^{\,0}\sin^2\delta\right]\Delta\Sigma$$
$$\frac{\partial u_y^{\,0}}{\partial y} = \frac{U_3}{2\pi}\left[\frac{3q}{R^5}(q+2y\sin\delta-\frac{5y^2q}{R^2}) - J_2^{\,0}\sin^2\delta\right]\Delta\Sigma$$

4.2.2.3

In the previous 4.2.2.1, 4.2.2.2 and 4.2.2.3 be: $s = p\sin\delta + q\cos\delta$

$$J_1^{\,0} = \frac{\mu}{\lambda+\mu}\left[-3xy\frac{3R+d}{R^3(R+d)^3} + 3x^3y\frac{5R^2+4Rd+d^2}{R^5(R+d)^4}\right]$$
$$J_2^{\,0} = \frac{\mu}{\lambda+\mu}\left[\frac{1}{R^3} - \frac{3}{R(R+d)^2} + 3x^2y^2\frac{5R^2+4Rd+d^2}{R^5(R+d)^4}\right]$$
$$J_3^{\,0} = \frac{\mu}{\lambda+\mu}\left[\frac{1}{R^3} - \frac{3x^2}{R^5}\right] - J_2^{\,0}$$
$$J_4^{\,0} = \frac{\mu}{\lambda+\mu}\left[-\frac{3xy}{R^5}\right] - J_1^{\,0}$$

4.2.2.4

Finally compute the tilt components along the axis.

Tilt along the Strike-slip:

$$\frac{\partial u_z^{\,0}}{\partial x} = -\frac{U_1}{2\pi}\left[\frac{3dq}{R^5}\left(1-\frac{5x^2}{R^2}\right)+K_1^{\,0}\sin\delta\right]\Delta\Sigma$$
$$\frac{\partial u_z^{\,0}}{\partial y} = -\frac{U_1}{2\pi}\left[-\frac{15xydq}{R^7}+\left(\frac{3xd}{R^5}+K_2^{\,0}\right)\sin\delta\right]\Delta\Sigma \qquad 4.2.2.5$$

Tilt along the Dip-slip:

$$\frac{\partial u_z^{\,0}}{\partial x} = -\frac{U_2}{2\pi}\left[-\frac{15xdpq}{R^7}-K_3^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma$$
$$\frac{\partial u_z^{\,0}}{\partial y} = -\frac{U_2}{2\pi}\left[\frac{3d}{R^5}\left(s-\frac{5ypq}{R^2}\right)-K_1^{\,0}\sin\delta\cos\delta\right]\Delta\Sigma \qquad 4.2.2.6$$

Tilt tensile:

$$\frac{\partial u_z^{\,0}}{\partial x} = \frac{U_3}{2\pi}\left[-\frac{15xdq^2}{R^7}-K_3^{\,0}\sin^2\delta\right]\Delta\Sigma$$
$$\frac{\partial u_z^{\,0}}{\partial y} = \frac{U_3}{2\pi}\left[\frac{3dq}{R^5}\left(2\sin\delta-\frac{5yq}{R^2}\right)-K_1^{\,0}\sin^2\delta\right]\Delta\Sigma \qquad 4.2.2.7$$

4.2.3 The Okada equations for a finite size source

The previous formulation summarizes the parameterization of the point like seismic source. Even though a simplified geometry a more realistic case is the rectangular, finite size, source with length L and width W (see figure below). The Displacement, Strain and Tilt can be modified changing equations from 4.2.1.1 onward by replacing x, y and d, with $x-\xi^{'}$, $y-\eta^{'}\cos\delta, d-\eta^{'}\sin\delta$. Moreover the definite integral $\int_0^L d\xi^{'}\int_0^W d\eta^{'}$ be considered. Being: $x-\xi^{'}=\xi, p-\eta^{'}=\eta$, with $p=y\cos\delta+d\sin\delta$, it becomes:

$$\int_x^{x-L} d\xi \int_p^{p-W} d\eta \qquad 4.2.3.1$$

The result of previous replacements in compact form is:
$f(\xi,\eta)=f(x,p)-f(x,p-W)-f(x-L,p)+f(x-L,p-W))$
The displacement vector components become the following:

Displacement Strike-slip

$$u_x = -\frac{U_1}{2\pi}\left[\frac{\xi q}{R(R+\eta)} + \tan^{-1}\frac{\xi\eta}{qR} + I_1 \sin\delta\right]$$
$$u_y = -\frac{U_1}{2\pi}\left[\frac{\tilde{y} q}{R(R+\eta)} + \frac{q\cos\delta}{R+\eta} + I_2 \sin\delta\right] \qquad 4.2.3.2$$
$$u_z = -\frac{U_1}{2\pi}\left[\frac{\tilde{d} q}{R(R+\eta)} + \frac{q\sin\delta}{R+\eta} I_4 \sin\delta\right]$$

Displacement Dip-slip

$$u_x = -\frac{U_2}{2\pi}\left[\frac{q}{R} - I_3 \sin\delta\cos\delta\right]$$
$$u_y = -\frac{U_2}{2\pi}\left[\frac{\tilde{y} q}{R(R+\xi)} + \cos\delta\tan^{-1}\frac{\xi\eta}{qR} - I_1 \sin\delta\cos\delta\right] \qquad 4.2.3.3$$
$$u_z = -\frac{U_2}{2\pi}\left[\frac{\tilde{d} q}{R(R+\xi)} + \sin\delta\tan^{-1}\frac{\xi\eta}{qR} - I_5 \sin\delta\cos\delta\right]$$

Displacement Tensile

$$u_x = \frac{U_3}{2\pi}\left[\frac{q^2}{R(R+\eta)} - I_3 \sin^2\delta\right]$$
$$u_y = \frac{U_3}{2\pi}\left[\frac{-\tilde{d} q}{R(R+\xi)} - \sin\delta\left\{\frac{\xi q}{R(R+\eta)} - \tan^{-1}\frac{\xi\eta}{qR}\right\} - I_1 \sin^2\delta\right] \qquad 4.2.3.4$$
$$u_z = \frac{U_3}{2\pi}\left[\frac{\tilde{y} q^2}{R^5} - I_5{}^0 \sin^2\delta\right]\Delta\Sigma$$

4.3 The generation of the synthetic interferograms

4.3.1 The model

The surface expression of a seismic event is the deformation field. It can be simply explained by a dislocation model in an elastic half space. To detect and measure such effect the InSAR technique is

available. The RNGCHN (RaNGe CHaNge) program (Feigl et al., 1999) deals with the parametrization of a dislocation model for computing the displacement vector. The program handles the differential interferograms relative to the coseismic measurements and uses the dislocation formulation provided by Okada (1985) in closed analytic form (refer to the paragraph 4.2).

Being the medium the boundary plane of an elastic half space the hypothesis of flat Earth is considered. Furthermore the elastic medium is specified by the Lamé coefficients λ and μ. In order to simplify the model $\lambda=\mu$ therefore these parameters drop out of the expressions for surface displacement. The Okada notation and its parameters are summarized in the following Table 1:

Quantity	**Symbol**	**Units**
Range change	$\Delta\rho$	Mm
Slip vector on fault	**U**	Mm
Dip angle	δ	Deg
Strike angle	α	Deg
Fault Width	W	Km
Fault Length	L	Km
Lower left corner of fault	[X, Y, d]	Km
Displacement vector at surface	**u**	Mm
Point on ground surface	[E, N]	Km
Satellite Unit vector	$\hat{s}$	
Lamé coefficients	λ, μ	Pa s
Fault centroid	$[X_c, Y_c, d_c]$	Km
Locking depth	D	Km
Seismic moment	M_0	Nm
Rake angle	**r**	Deg

Table 1: fault plane parameters notation in the reference system used in RNGCHN software

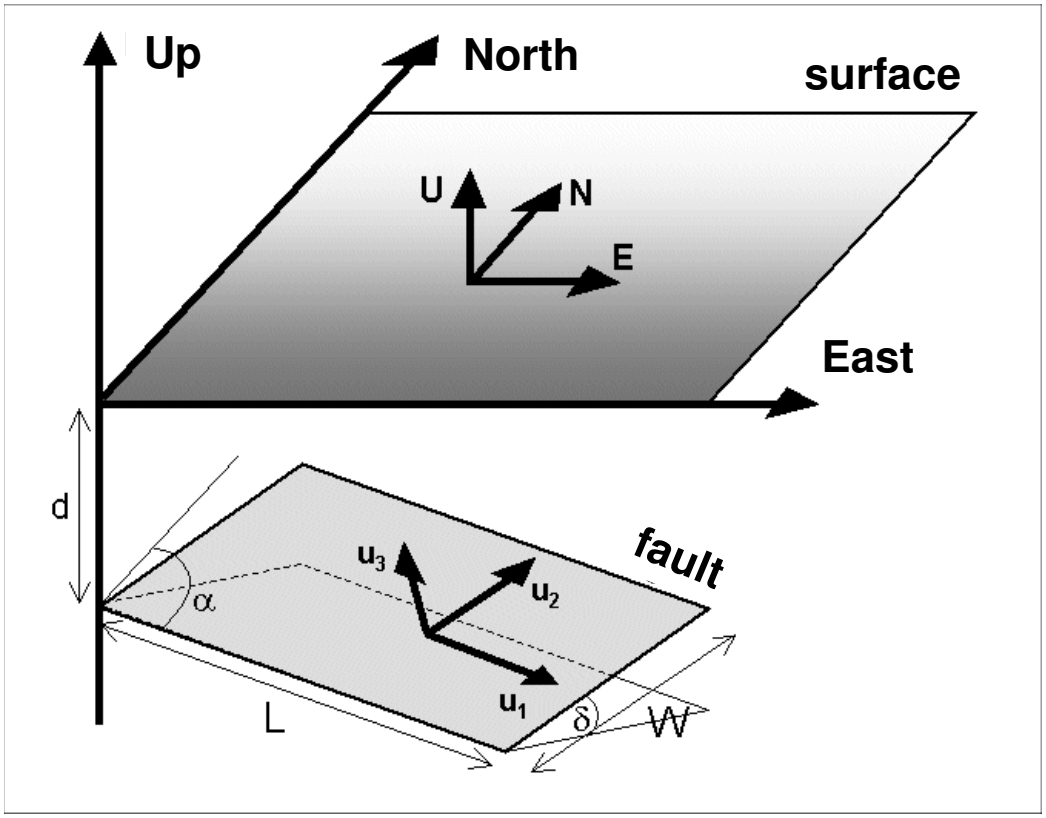


Fig. 39 The fault plane is in described within its own reference system (d, α, δ).

The Okada formulation is foresees the point source and the rectangular, finite dimension, shape. As far as the slip vector is concerned it is divided in the strike, dip, and tensile components are U_1, U_2 and U_3. Whenever conceiving the tensile component U_3=0, the slip vector lies on the fault plane. The strike (α) is the clockwise angle between the fault trace line and the NS direction. The dip (δ) is the angle between the fault plane and the surface. Finally the rake angle is computed as $\tan r = \frac{U_2}{U_1}$.

Exploiting the fault type classification provided above, the normal fault has U_2<0 and r<0; the reverse or thrust fault has U_2>0 and r>0; the left lateral strike slip implies U_1>0 and $|r| \le 90°$ while the right lateral one U_1<0 and $|r| \ge 90°$. As far as the seismic moment is concerned it is computed as $M_0 = \mu ULW$. Common values used for μ in the Earth crust range 30-36 GPa are considered constant in the investigated, elastic half space.

The input parameters for the RNGCHN program mainly concern the fault geometry (W, L, D, α, δ) and the displacement (**u**). The (East, North, Up) RNGCHN coordinate system stems from the (X, Y, d) Okada's one by rotating for the fault strike α clockwise from the North-South direction.

The origin of the fault in the Okada's reference system does not coincide with the centroid of the rectangle. Therefore be $[E_c, N_c, d_c]$ the coordinates of the centroid, the Okada's origin is as follows:

$$\begin{bmatrix} E \\ N \\ d \end{bmatrix} = \begin{bmatrix} E_c \\ N_c \\ d_c \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \sin\alpha & -\cos\alpha & 0 \\ \cos\alpha & \sin\alpha & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} L \\ W\cos d \\ W\sin d \end{bmatrix} \qquad (4.3.1.1)$$

4.3.2 Geometric parameters

The unit vector **ŝ** (normalized to a unit vector) stands for the target-to-satellite direction. It is determined by two angles, the SAR incidence from vertical and the azimuth from North of the satellite ground track. Therefore the range change $\Delta\rho$ measured along the satellite Line Of Sight (LOS), i.e. the line joining the SAR sensor and the target ground point [*E, N*] is:

$$-\Delta\rho = \mathbf{u}\cdot\hat{\mathbf{s}} \qquad (4.3.2.1)$$

whereas the convention is $\Delta\rho$ positive whenever the point target moving towards the satellite, while it is negative otherwise.

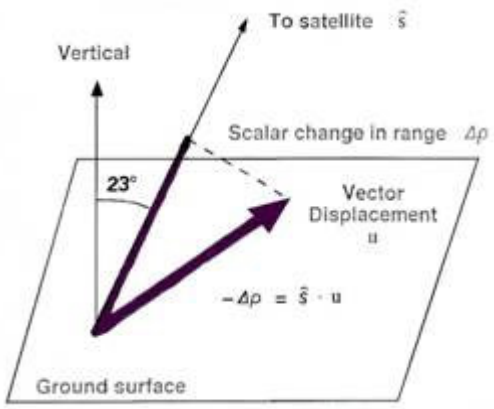


Fig. 40: Geometric relations between u, ŝ e $\Delta\rho$.

4.3.3 Synthetic interferograms: the dataset

The RNGCHN software has been then used to generate the synthetic interferograms and the proper dataset used for training the artificial neural network and for validating its capabilities during the classification and the parameter retrieval. As stated above, the input parameters used for the generation of the synthetic interferograms, or models, either concerns the geometry of the fault or the displacement vector. Anyway in this study the fault slip has been assumed a constant value for each fault class. As far as the dataset of models is concerned it has been generated by moving the parameters all along a certain range. To be exhaustive this latter has been defined accounting for the effective variation range of each parameter within the frame of the three fault classes (normal, thrust and strike slip). In the table below (Table 2) are the guidelines for the model generation, i.e. the range interval for each parameter and the sampling rate.

Parameter Variation range / Fault class	Length (km)	Width (km)	Dip angle (deg)	Strike angle (deg)	Bottom Depth (km)
Normal	9-21 (3)	6-8 (2)	45-85 (10)	90-180 (30)	8-12 (4)
Thrust	10-50 (10)	10-20 (10)	30-70 (10)	90-180 (30)	10-20 (10)
Strike slip	25-125 (25)	15-20 (5)	70-110 (10)	90-180 (30)	20-25 (5)

Table 2: the fault plane parameters are made varying within a certain range. The data combination univocal identify each fault class

The fault slip vectors are U_1= 0 mm U_2= -500 mm U_3= 0 mm for the Normal fault; U_1= 1600 mm U_2= 0 mm U_3= 0 mm for the Strike slip fault; finally U_1= 0 mm U_2= 500 mm U_3= 0 mm for the Thrust or Reverse fault.

The seismological and geological knowledge have been exploited to select a reliable data range for each parameter. Indeed the strike and the dip angles stem from the measurements provided by the geologists during field surveys, or are inferred from the earthquake focal mechanism showing the effective fault orientation. To this aim the upper and lower boundaries in Table 2 are stated referring to case studies and fault catalogues. Based on such ranges and sampling rates the number of synthetic interferograms obtained is 400 for each fault class, thus leading to 1200 synthetic interferograms in all (Fig. 41 and 42). Each wrapped synthetic interferogram has been visualized with a colour scale. The colour cycle corresponds to a complete phase cycle, $\pm\pi$.

The interferometric phase measured by InSAR accounts for the satellite-to-target LOS range change. Therefore the colour cycle is also equal to half wavelength $\frac{\lambda}{2}$. In this work C-Band images have been used meaning $\frac{\lambda}{2} \cong 2.8cm$.

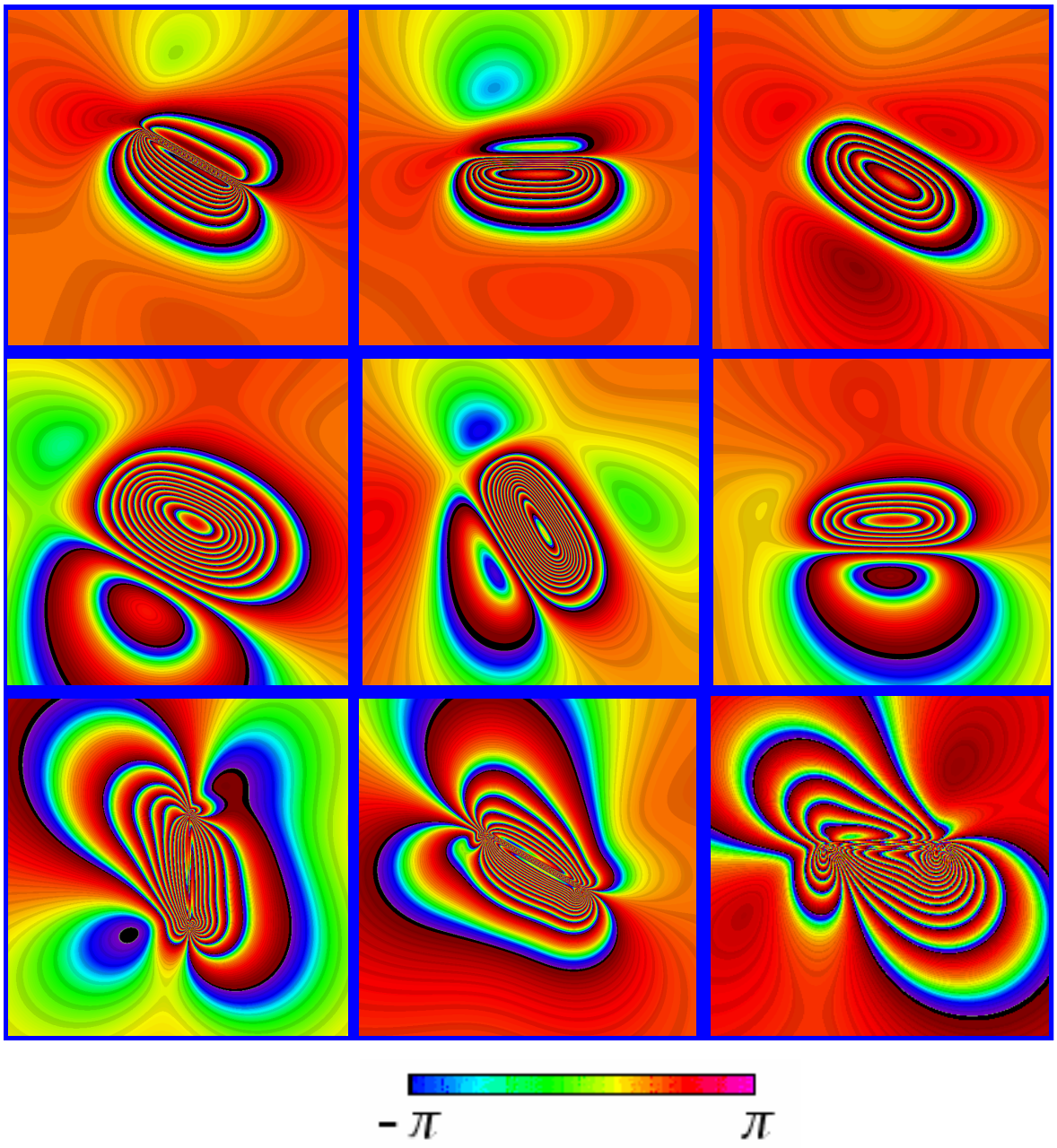


Fig.41: Example of wrapped synthetic interferograms. From top to down: the first three along the upper line belong to the Forward fault class; the other three along the central line are Thrust or Reverse fault; the lower three concern the Strike slip fault. The color cycle and the phase scale is also shown

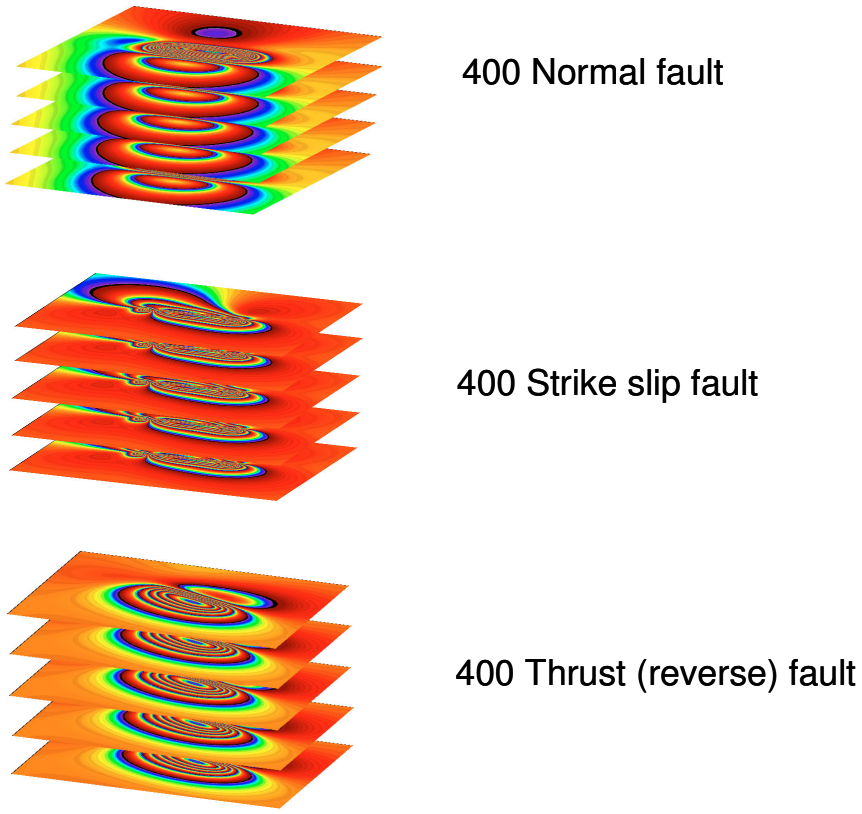


Fig. 42: The dataset of synthetic interferograms is composed by 1200 models, 400 for each source fault mechanism (Normal, Strike slip and Thrust).

4.3.4 The inverse problem

The goal of this thesis is exploiting the performance of a neural network for geophysical applications, with particular attention to the seismic source model. The latter topic is an unprecedented fieldwork for neural networks being previous studies not available in literature yet. Define the inverse problem means choosing the opportune topology for the neural network besides the characteristics of the input and output layer.

Two inverse problems have been conceived thus resulting in two different solutions. Therefore the work can be divided in two phases:

- the seismic source classification
- the fault parameter quantitative retrieval

Each of the proposed exercises has been split in Training, Validation and Test.
The main advantage of the neural networks is the capability to generalize, meaning they are able to recognize and correctly classify data from a dataset it has never seen before. Leading to effectively exploit the potential of the net the learning phase has to be correctly performed. To this aim the input dataset is split in three subsets:

1- the training set, used during the for running the whose error is minimized during training
2- the validation set, to evaluate the capability of the trained network to work with data new data
3- the test set, for checking the overall performance of the neural network using a real, non-synthetic, dataset.

The neural network reaches its best performance when the learning phase has been properly stopped when the validation error is low. Whenever the training carries on after the error has reached its minimum, the overtraining reduces the capability to provide a more general solution.
The Stuttgart Neural Network Simulator (SNNS) is the freeware software used to generate the neural networks in this work. SNNS allows planning and testing the preferred topology.
The performance evaluation of SNNS learning is measured by means of statistic parameters. The more common is the Sum Squared Error (SSE) of the learning function. It can be written as follows:

$$SSE = \sum_{p \in patterns} \sum_{j \in output} \left(t_{pj} - o_{pj}\right)^2 \qquad (4.3.4.1)$$

where t_{pj} is the desired output of neuron j for the pattern p, o_{pj} is the actual output.
The Mean Squared Error (MSE) is:

$$MSE = \frac{SSE}{n - p} \qquad (4.3.4.2)$$

where n is the number of observations and p the number of parameters, or weights, to be estimated.
The SSE/o-units is the SSE divided by the number of outputs.

It is straightforward that the displayed error is computed on the Training or on the Test set according to what indicated in the first column. Usually the learning phase is stopped while the MSE of the Test tends to be constant (see example in Fig. 43).

```
Learning all patterns:
  epochs   : 20
  parameter:  0.20000
  #o-units : 5
  #patterns: 50 (total: 50)

     epoch:         SSE           MSE          SSE/o-units

Train   20:      21.31374        0.42627          4.26275
Test    20:       9.92320        0.33077          1.98464
Train   18:      12.96179        0.25924          2.59236
Train   16:       9.19350        0.18387          1.83870
Test    15:       7.13944        0.23798          1.42789
Train   14:       6.35968        0.12719          1.27194
Train   12:       5.13774        0.10275          1.02755
Train   10:       4.42500        0.08850          0.88500
Test    10:       6.79930        0.22664          1.35986
Train    8:       3.74046        0.07481          0.74809
Train    6:       3.13978        0.06280          0.62796
Test     5:       6.74240        0.22475          1.34848
Train    4:       2.72080        0.05442          0.54416
Train    2:       2.38300        0.04766          0.47660
Train    1:       2.21536        0.04431          0.44307
```

Fig. 43: Neural network learning process: the statistic parameters SSE and MSE are computed for Train and Test. The learning is generally stopped whenever such parameters tend to a constant value.

4.3.4.1 Training and Validation

The overall dataset of synthetic interferograms is composed by 1200 examples each one of 1500 x 1500 pixels. The very high range for the geometric parameters Length (9-125 km) and Width (6-20 km) leads to generate synthetic interferograms with very different sizes. For sake of simplicity and to work with a homogeneous dataset the interferograms are resized to the dimension above. Therefore being the effective extension of the main part of them minor than the maximum size, the synthetic interferograms are filled with a square crown of zeroes. Reasonably it can be conceived that the inner, central portion of 750 x 750 pixels contains the main part of the significant information, at least. Therefore this subset has been extracted from each model thus providing a

further dataset of 1200 synthetic interferograms of 750 x 750 pixels. Despite such processing the input layer of the neural network would remain too large and a further approach is needed. The provided solution is a 75 x 75, 10-pixel spaced grid to perform a data sampling. Therefore the neural network receives in input a set of 5625 values for each interferogram. In other words the input layer of the neural network is composed by 5625 neurons (Fig. 44).

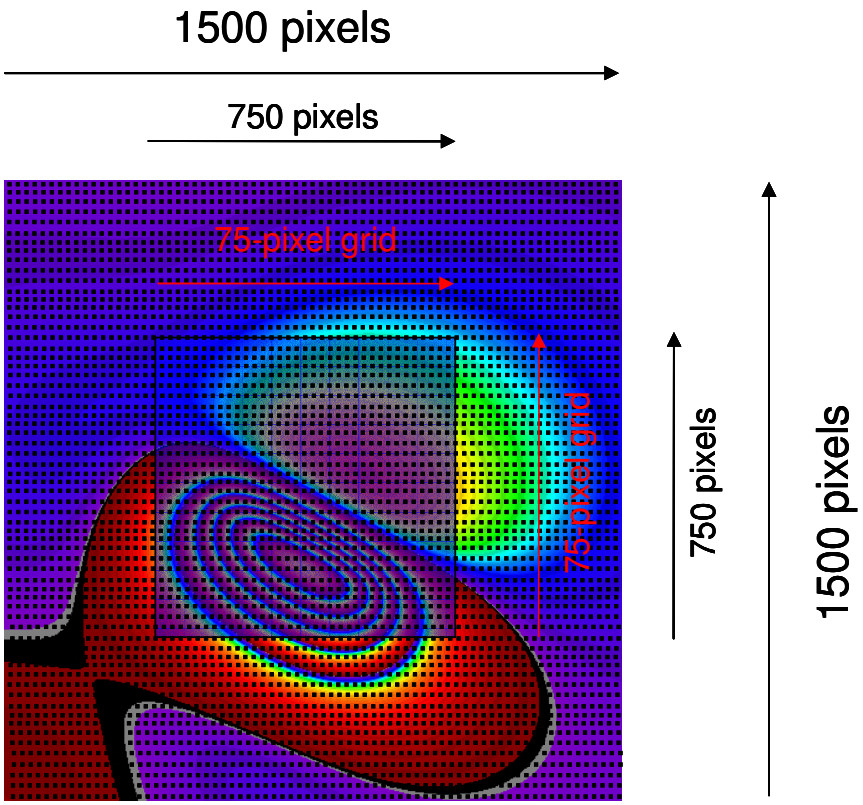


Fig. 44: Each synthetic interferogram is 1500x1500 pixels. The 750x750 pixels central part is then extracted. Furthermore a sampling procedure is applied to select a grid of 75x75 pixels with a constant rate of 10 pixels.

4.3.4.2 The seismic source classification

The dataset of 1200 synthetic interferograms is split in two subsets, the training set and the validation set. A random process is applied for data selection for avoiding any possible subjective criteria. The resulting Training set is thus composed by 750 randomly selected elements whilst the Validation set contains 450 ones. The Learning phase is the most critical one being addressed to assess the capabilities of the network. The leading criteria are basically a competitive learning where, at each layer, the winner neurons are then capable to drive towards the output. It is roughly

composed by a Training phase and a Validation one. During the Training phase the neural network receives as input the 750 synthetic interferograms of the Training set to opportunely set up the weights for each connection. The scope of the Validation phase is verifying the effective capability of the trained network to recognize new data. Therefore the 450 synthetic interferograms of the Validation set are the new input to the network. Statistical parameters can be computed to evaluate the performance of the neural network. In Table 3 the confusion matrix referred to the Validation set is showed.

As far as the input data are concerned to better handling them and to reduce the range values each interferogram is scaled between -1 and 1.

The preferred topology of the neural network is the result of a heuristic approach aimed to select the better suited one for this exercise. SNNS software allows creating the preferred neural network. Therefore after some attempts addressed to this choice the most effective one is a MLP with two hidden layers, characterized by the topology [5625] [200] [60] [3] (see Figure 45). The SCG (Scaled Conjugate Gradient) function has been used to perform the learning phase. The empirical criteria addressing the learning phase to provide a neural network able to maintain a significative capability to generalize accounts for the Training and Validation curves. In fact while the Learning phase is in progress the Training curve tends to lower its error so as the Validation one. In particular whenever this latter error decreases up to a certain threshold the Learning procedure should be stopped. On the contrary if the Learning phase is in progress after reaching this lower the Validation error increases thus limiting the potential of the network (see Fig. 46).

The competitive learning is the capability of the neural network to decide the winner neuron after a huge number of connections. In the classification exercise the goal is the correct association to the input of one of the three output nodes.

Accounting for the selected test case the output layer can simply provide three combinations, 1 0 0 for Normal fault, 0 1 0 for Strike slip fault, 0 0 1 for Thrust or Inverse fault.

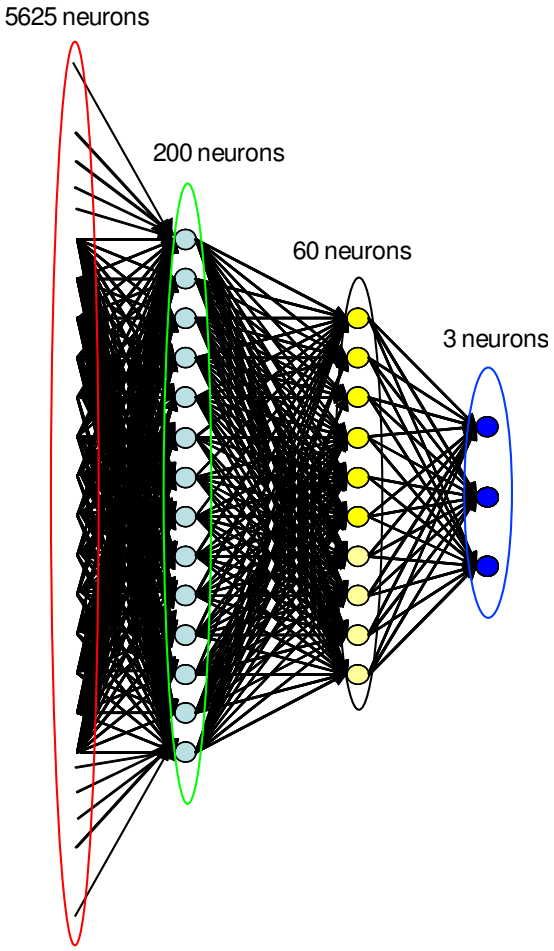


Fig. 45: topology of the neural network used for the classification exercise. The input layer is composed by a huge number of nodes, 5625, extracted from each synthetic interferogram by applying a grid of 75 x 75 pixels.

	Normal (1 0 0) 135	**Strike slip** (0 1 0) 162	**Thrust** (0 0 1) 153
Normal (1 0 0) 135	131	1	3
Strike slip (0 1 0) 162	4	158	0
Thrust (0 0 1) 153	0	3	150

Table 3: Confusion matrix computed for the Validation set. The high number of synthetic interferograms concentrated along the diagonal means the network is able to correctly recognize them.

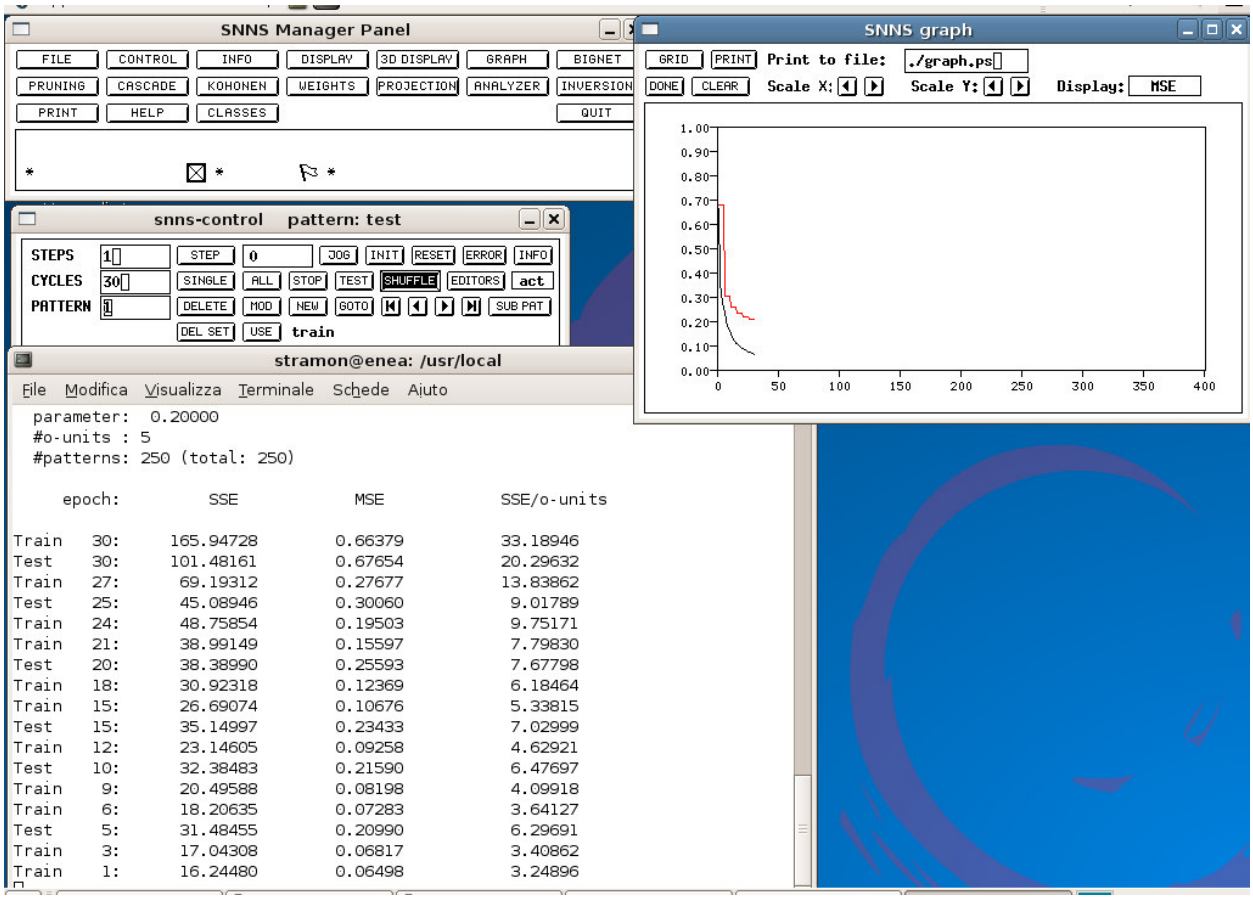


Fig. 46: Output view from SNNS Training and Validation phase. The graph plot shows the behaviour of the Validation (in red) and the Training (in black) curves after 30 cycles. Both have reached the minimum error level.

4.3.4.3 The fault parameter quantitative retrieval

The second exercise concerns the inversion problem, i.e. the retrieval of the geometric fault parameters from the dataset of synthetic interferograms. The parameters involved in such exercise are: Length, Width, Dip angle, Strike angle, Bottom Depth. For sake of simplicity the inversion problem has been referred to the Strike slip case study only. Anyway further works might be focused on completing this exercise accounting for the Normal and Reverse fault classes.

The procedure is quite similar to the previous one. First, stemming from the Strike slip dataset of 400 synthetic interferograms the Training and Validation subsets are defined. As for the case above the selection criteria adopted is a random process for avoiding any personal choice. To handle this case study as for the classification one, the Training set is composed by 250 synthetic interferograms whilst the remaining 150 belong to the Validation set. The further processing steps

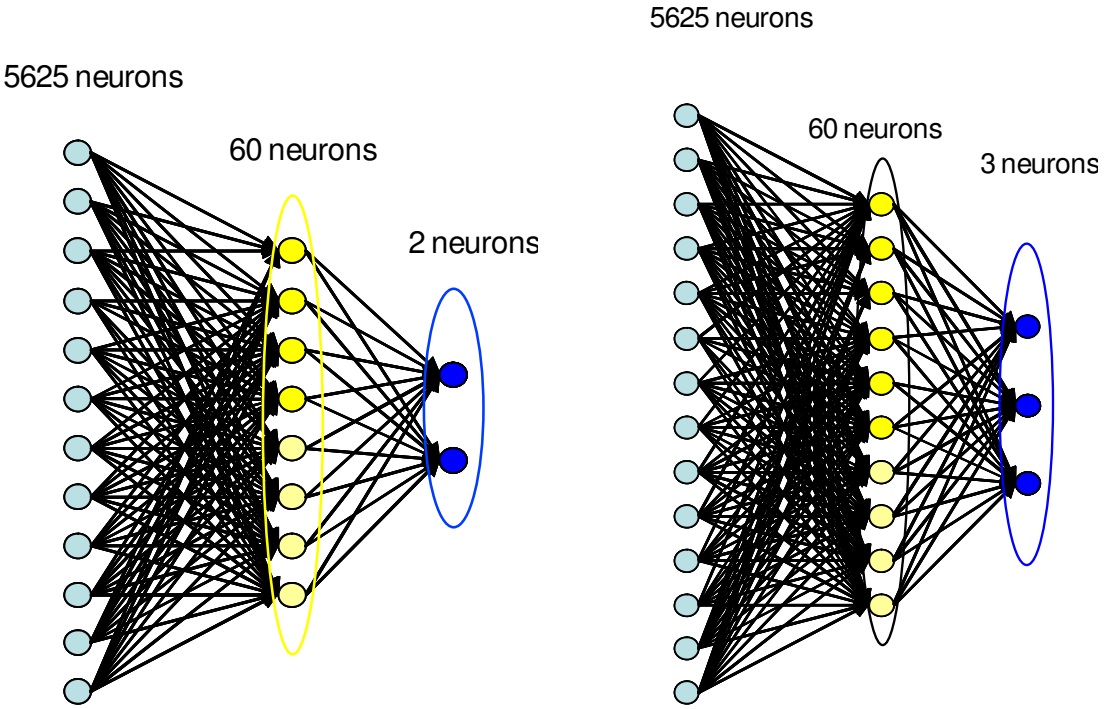
necessary to opportunely format the data have been already described in the previous classification exercise and are applied here. The goal is limiting the number of neurons for the input layer of the neural network without reducing the information content. Therefore each synthetic interferogram has been resized and sampled according to the abovementioned procedure (see previous paragraphs).

The parameter retrieval exercise has been split in three exercises each one with growing complexity:

- 2 parameter: Strike and Dip angles;
- 3 parameters: Strike and Dip angles, Bottom Depth;
- 5 parameters

The artificial neural network used for each of them is not the same. The first layer, i.e. the input layer, has exactly the same number of nodes (5625) used for the classification. Concerning the output layer it depends on the number of geometric parameters to be estimated. In the first two cases, two and three output parameters, the topology of the neural network has just one hidden layer of 60 neurons.

The heuristic approach allowed inferring this topology whereas it seemed the simplest one capable to provide effective parameter retrieval. Therefore the topology is [5625][60][2] (2-parameters) and [5625][60][3] (3-parameters) (Figure 46). As far as the more complex exercise is concerned, the five parameter retrieval, based on the same heuristic approach it has been assumed that two hidden layers are necessary. This leads to the following larger topology: [5625][200][60][5] (Figure 47).



(a) (b)

Fig. 46: topology of the neural networks used for the 2-parameters (a) and 3-parameters (b) inversion. The input layer contains a huge number of nodes, 5625, extracted from each synthetic interferogram by means of a grid of 75 x 75 pixels.

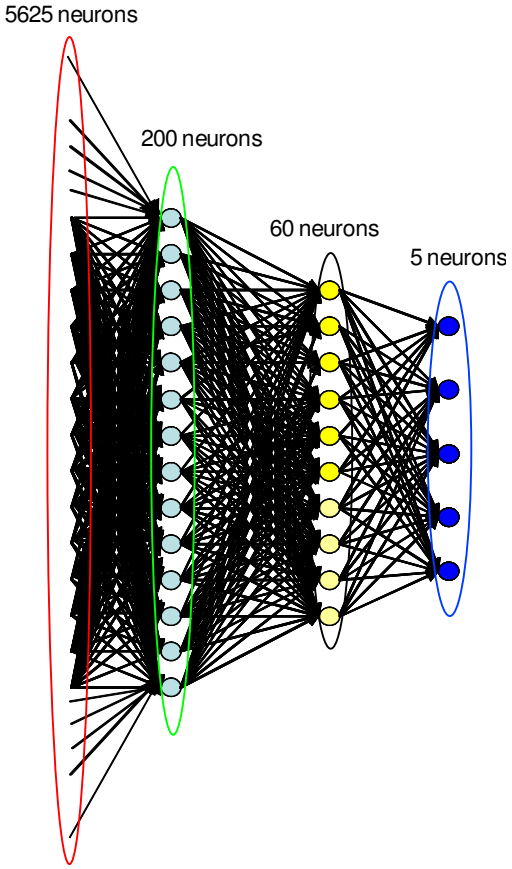


Fig. 47: topology of the neural network created for the inversion problem solution. The input layer contains a huge number of nodes, 5625, extracted from each synthetic interferogram by means of a grid of 75 x 75 pixels.

To fully exploit the capability of the neural network the Learning phase has to be completed accounting for its main properties.

The Training and Validation curves move towards an error threshold and the Learning phase should be stopped before reaching such value. This to avoid the Validation error might increase after achieving its minimum. For this case study the empirical driving criteria suggested to stop the Training phase after 30 cycles. The Validation is thus provided every 5 cycles to verify its behaviour. To infer the performance of the neural network the SSE and the MSE indexes are then computed.

To provide a statistics about the performance of the network learning phase the Root Mean Square Error (RMSE) has been computed for the 2-, 3- and 5-parameter cases. The RMSE provides an estimate of the agreement between the expected output and the computed one. The RMSE for each parameter is roughly of the same order for those parameters common to the exercises (Table 4).

	Length	**Width**	**Dip**	**Strike**	**Depth**
2-parameters	0.014 (1.4%)	0.015 (1.5%)			
3-parameters	0.014 (1.4%)	0.015 (1.5%)	0.05 (5%)		
5-parameters	0.015 (1.5%)	0.017 (1.7%)	0.077 (7.7%)	0.07 (7%)	0.08 (8%)

Table 4: RMSE computed for the three exercises, 2-, 3- and 5-parameters. The RMSE values are of the same order of magnitude.

5 The experimental results

5.1 The test cases and their fault parameters

In order to evaluate the effective capability of the proposed innovative approach some test cases have been studied. The classification exercise and the seismic source parameter retrieval have been analyzed. The four selected test cases concern moderate to strong earthquakes (Mw around 6.0 or higher) dated since 1992 when InSAR technique was first applied to Landers. The leading criteria to choice were first the availability or not of InSAR data (based on ERS1-ERS2, Envisat, Radarsat, … images). Second, the amount and quality of the information concerning each seismic source. Third, the test cases should roughly have fault plane geometries belonging to the three classes above. Finally, the number of manuscripts and papers already available in literature concerning the earthquake.

The 1997 Colfiorito (Italy) earthquake is characterized by a seismogenic fault with a normal mechanism. The 1999 Izmit (Turkey) and the 2003 Bam (Iran) earthquakes refer to the strike slip mechanism. Finally the Northridge earthquake is the example of the thrust or reverse fault geometry.

The proposed test cases have been widely studied to analyze the seismic source geometry. Papers available in literature analysed the issue of forward and inverse modelling and fixed the constraints using several data, seismic, geodetic and InSAR.

Hereafter the test cases are summarized in the Tables below, one for each of them. Here are the estimates of the five geometric parameters of the seismic sources, the same five parameters of the retrieval exercise, as provided by the cited Authors. The comparison among the proposed solutions clearly shows the wide variability range of the provided estimates.

Colfiorito Geometric Parameters	Length (km)	Width (km)	Dip (deg)	Strike (deg)	Bottom Depth (km)	Data used
Stramondo et al., 1999	12.5	7.5	45	144	6.5	SAR+GPS
Salvi et al., 2000	10 6	8 7	46 45	154 138	8 6	SAR+GPS
Cocco et al., 2000	12 7	12 7	42 46	144 152	6.5 6	Seismic data
Hernandez et al., 2004	12.5 7.5	7.5 7.5	47 51	144 152	6.5 8	SAR+GPS+Seismic data

Table 5: geometric parameters relative to the Colfiorito seismic source fault plane as obtained by authors in literature.

Izmit Geometric Parameters	Length (km)	Width (km)	Dip (deg)	Strike (deg)	Bottom Depth (km)	Data used
Stramondo et al., 2002	120 135	20-25 “ “ “ “	87 “ “ “ “	90 “ “ “ “	25 “ “ “ “	SAR
Cakir et al., 2003	150-160	24	85	100-110	25	SAR
Feigl et al., 2002	130-150	21	90	90	25	SAR+GPS+ SPOT
Reilinger et al., 2000	97	16	90	87	20	GPS

Table 6: geometric parameters relative to the Izmit earthquake concerning the modelled seismic source fault plane as obtained by the cited authors.

Bam Geometric Parameters	Length (km)	Width (km)	Dip (deg)	Strike (deg)	Bottom Depth (km)	Data used
Stramondo et al., 2005	14	8	88	177	9	SAR
Jackson et al., 2006	12 (15)	8.6 (1.5)	86 (64)	175 (180)	14 (7)	SAR
Talebian et al., 2004	20	15	88 (180)	177 (30)	15	SAR

Table 7: geometric parameters relative to the Bam seismic source fault plane as obtained by some authors.

Northridge Geometric Parameters	Length (km)	Width (km)	Dip (deg)	Strike (deg)	Bottom Depth (km)	Data used
Wald et al., 1996	15	20	40	122	15	Strong Ground Motion +Teleseismic+GPS+Leveling Data
Thio and Kanamori, 1996			42	130	19	Seismic waveforms

Table 8: geometric parameters relative to the Northridge seismic source fault plane as obtained by some authors.

5.2 The Test phase

The capability of the proposed methodology has to be verified for a test case. After the Training and Validation phases the neural network is now able to infer the geometric characteristics of a generic fault plane from the available interferograms. Therefore the Test phase is based on the use of a coseismic interferogram generated by means of specific softwares leading to compute such surface displacement map.

The four case studies have taken into account the whole possible set of seismic sources, whereas they have been split within the Normal, Thrust and Strike slip classes.

Therefore the abovementioned earthquakes InSAR data have been available for are the following:

- **Colfiorito** (September 26th 1997, Central Italy, Mw= 6.0)
- **Izmit** (August 17th 1999, Turkey, Mw= 7.4)
- **Northridge** (January 25th 1994, South California, US, Mw=6.7)
- **Bam** (December 26th 2003, Iran, Mw=6.5)

5.2.1 Normal mechanism: the 1997 Colfiorito earthquake

The 1997 Umbria-Marche, Central Italy, seismic sequence was originated by a main event on September 26 at 00:33 GMT (Greenwich Medium Time), Mw 5.7. This earthquake caused strong damage to the ancient and historical towns nearby the epicenter. Some hours later, at 09:40 GMT, a

Mw 6.0 hit the already damaged area, thus completing the collapse of the main part of the man made structures.

During the following months a huge amount of small to moderate aftershocks took place, up to another main event on October 14, Mw 5.6, 15 km S of the previous earthquakes.

As far as the fault parameters are concerned they have been inferred from the manuscripts in literature, based on the seismological data, the GPS measurements and the InSAR results. In particular the differential interferogram computed using a 35 days ERS2 image pair on descending pass shows a surface displacement pattern generally characterized by an enough good signal to noise ratio.

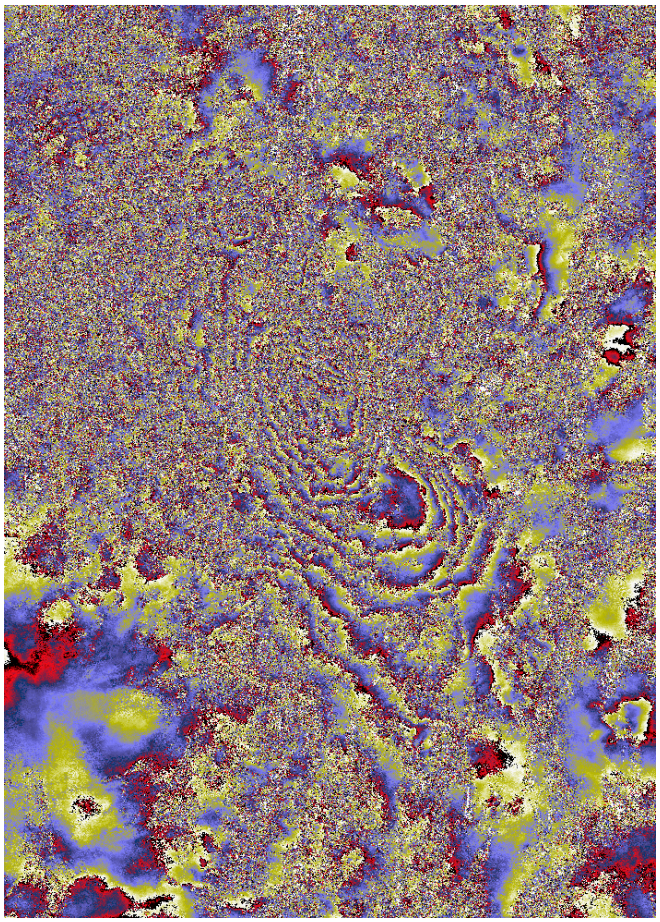


Fig.48: Coseismic Interferogram relative to the Colfiorito (Central Italy) earthquake. The temporal baseline is 70 days.

5.2.2 Strike slip mechanism: the 1999 Izmit earthquake

The Izmit earthquake is one of the main seismic events hitting the Mediterranean during the last Century. The earthquake affected the urban and industrial area surrounding the Gulf of Izmit (Marmara Sea) and to the eastward at Adapazari city (see figure below). The mainshock was located along the northern strand of the North-Anatolian fault system, approximately 9 km southeast of the city of Izmit, at an ipocentral depth of 16-18 km. The focal solution shows a right-lateral strike-slip mechanism (*Barka*, 1999; *Toksöz et al.*, 1999). As far as the surface effects are concerned the Izmit event was characterized by up to 110-120 km of surface ruptures.

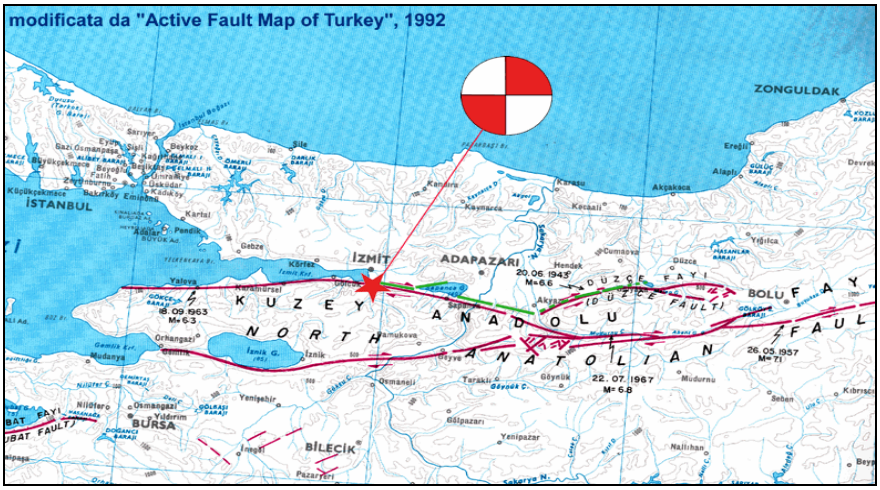


Fig. 49: The Western portion of the North Anatolian Fault System and the epicenter of the Izmit seism. The focal mechanism is a pure strike slip as also shown.

The average horizontal displacement measured along the North-Anatolic Fault System reaches a deformation rate of about 20 mm/y. The coseismic effects have been due to the activation of a fault patch up to 110 km long whereas the horizontal displacement was between 2 and 5m, with a mean value of about 2.5m.

The measurement of the surface displacement field has been provided by applying InSAR processing. Two ERS1-ERS2 SAR image pairs on ascending pass, either pre- and post-seismic ones, have been used to generate the coseismic interferograms.

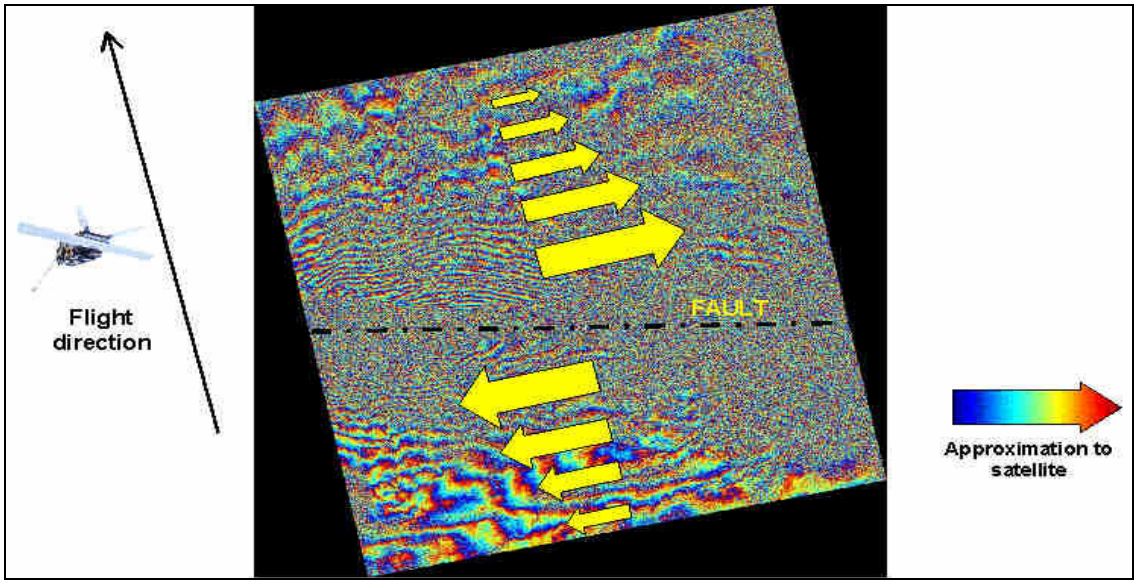


Fig. 50: Coseismic differential interferogram from the image pair August 13th – September 17th. The SAR data where acquired along the ascending pass. Indeed the LOS is from SW to NE and almost parallel to the fault line.

5.2.3 Thrust (reverse) mechanism: the 1994 Northridge earthquake

The Northridge earthquake affected the area nearby the city of Los Angeles in 1994. It has been originated by an inverse fault mechanism, leading to 43 cm vertical displacement and 20 cm horizontal as measured by the GPS.

The whole surface displacement field has been retrieved by applying InSAR technique to an ERS1 coseismic image pair along descending orbit. The surface coherence is very limited mainly towards SW probably due to the topographic relief and the very high temporal baseline (1992-1995).

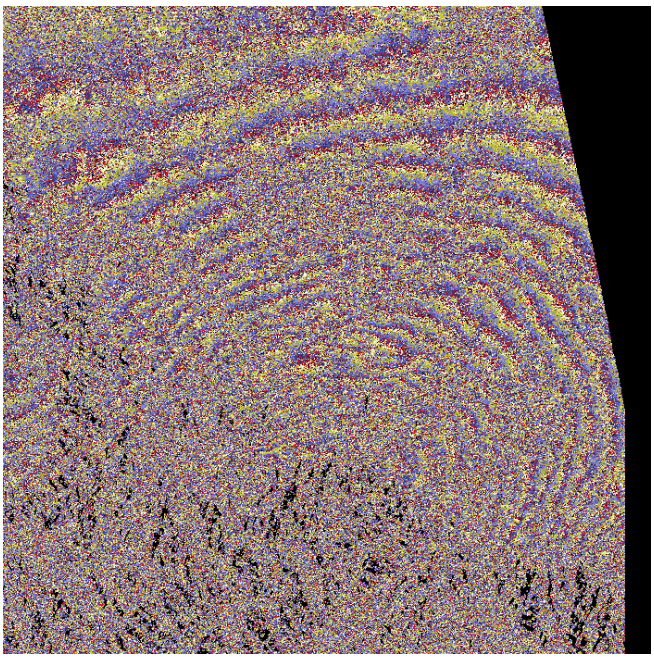


Fig. 51: Differential coseismic interferogram relative to the Northridge earthquake. The very high surface decorrelation is particularly clear in the high relief area toward S.

5.2.4 A further test for the Strike slip mechanism: the 2003 Bam earthquake

The previous three examples are representative for the fault plane geometries, normal, strike, inverse. Despite this to provide a further test case another recent strong earthquake answering to the selection criteria above has been chosen. The Bam earthquake took place close to the historical city of Bam and near the Bavarat village (see figure below). It was originated at a depth of about 10 km (USGS-United States Geological Survey; IIEES-International Institute of Earthquake Engineering and Seismology) in a desert and almost flat area.

The seismic focal mechanism suggests that the slip occurred along a high-angle dipping fault plane, with a prevailing strike-slip sense of motion (USGS). The InSAR technique was applied by Stramondo et al. (2004, 2005) to investigate the Bam earthquake source and the resulting differential interferogram revealed the existence of an un-mapped fault, called the “Arg-e-Bam” fault. The Bam earthquake is considered the first one investigated by the new European SAR Satellite, Envisat Advanced SAR, addressed to replace the ERS1 and ERS2 satellites. The Envisat IS2 acquisition geometry is similar to the ERS1/2 (right-side looking, 23° average incidence angle). The coseismic interferogram has been computed on the descending pass.

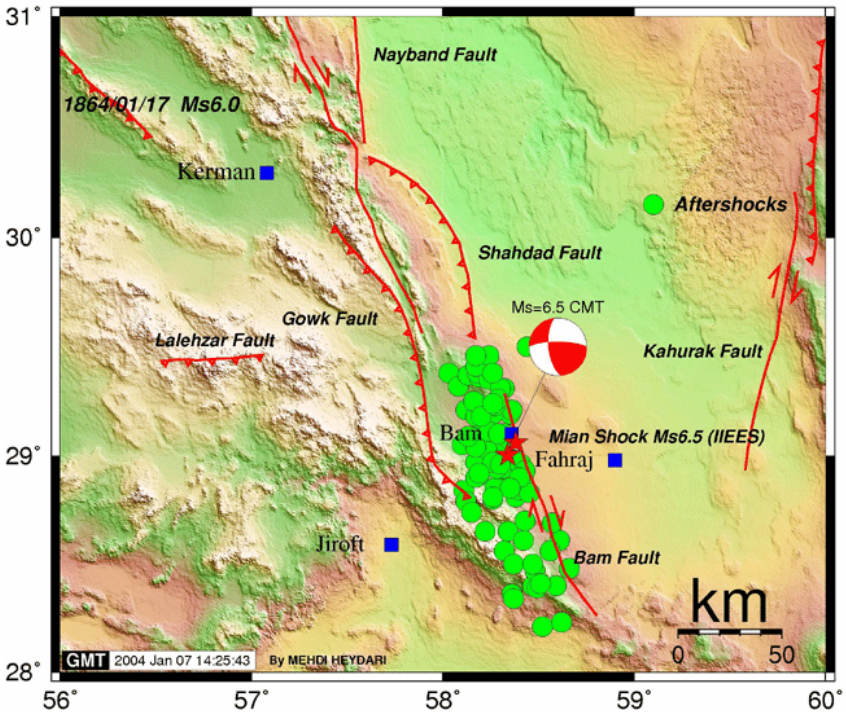


Fig. 52: the December 2003 Bam earthquake. The epicentral area is located in the SE part of Iran close to the historical city of Bam. The fault mechanism is mainly a strike slip geometry.

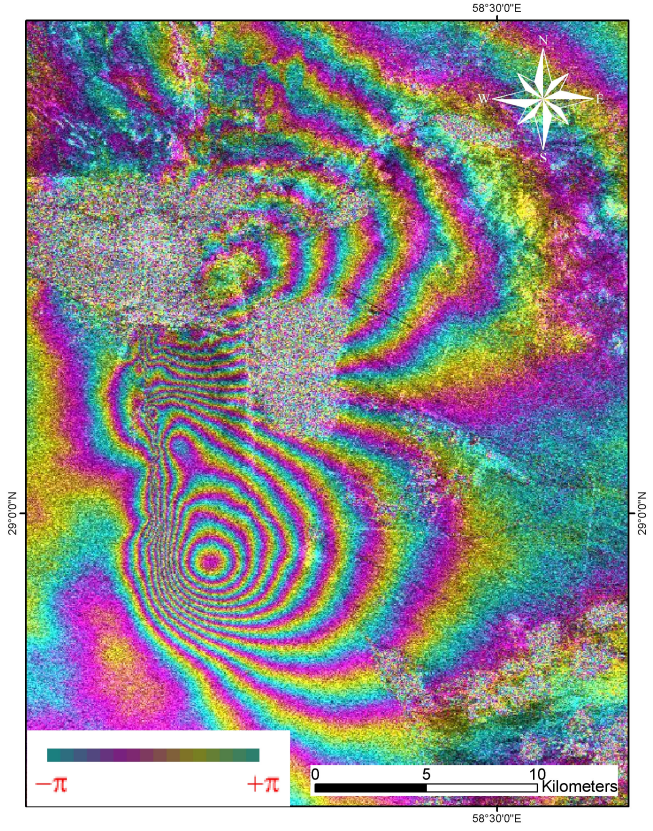


Fig. 53: Coseismic differential interferogram of the Bam earthquake. Each color cycle, or fringe, corresponds to λ/2 displacement along the LOS. This is the first seismological application of InSAR to Envisat ASAR data.

5.3 The classification exercise

The classification exercise has been performed on the surface displacement field measured by InSAR. The available interferograms relative to these four earthquakes have been then used within the Test phase that is based on the properly trained neural network. The coseismic interferogram computed for each test case differs from the others either for the surface extension or for the shape. These differences are due to the inhomogeneities of the seismic source geometry. For avoiding any discrepancy the interferograms relative to the test cases are formatted as the 1200 synthetic interferograms of the Normal, Strike and Inverse datasets. To this aim the already applied approach has been used. The interferograms of the test cases have been first resized to 1500 x 1500 pixels, afterwards the centred spatial subset of 750 x 750 pixels has been extracted. Finally the input layer of the trained neural network composed by 5625 neurons has been generated applying a 10 pixel spaced grid of 75 x 75 nodes. In Figure 54 is a schematic view of the exercise.

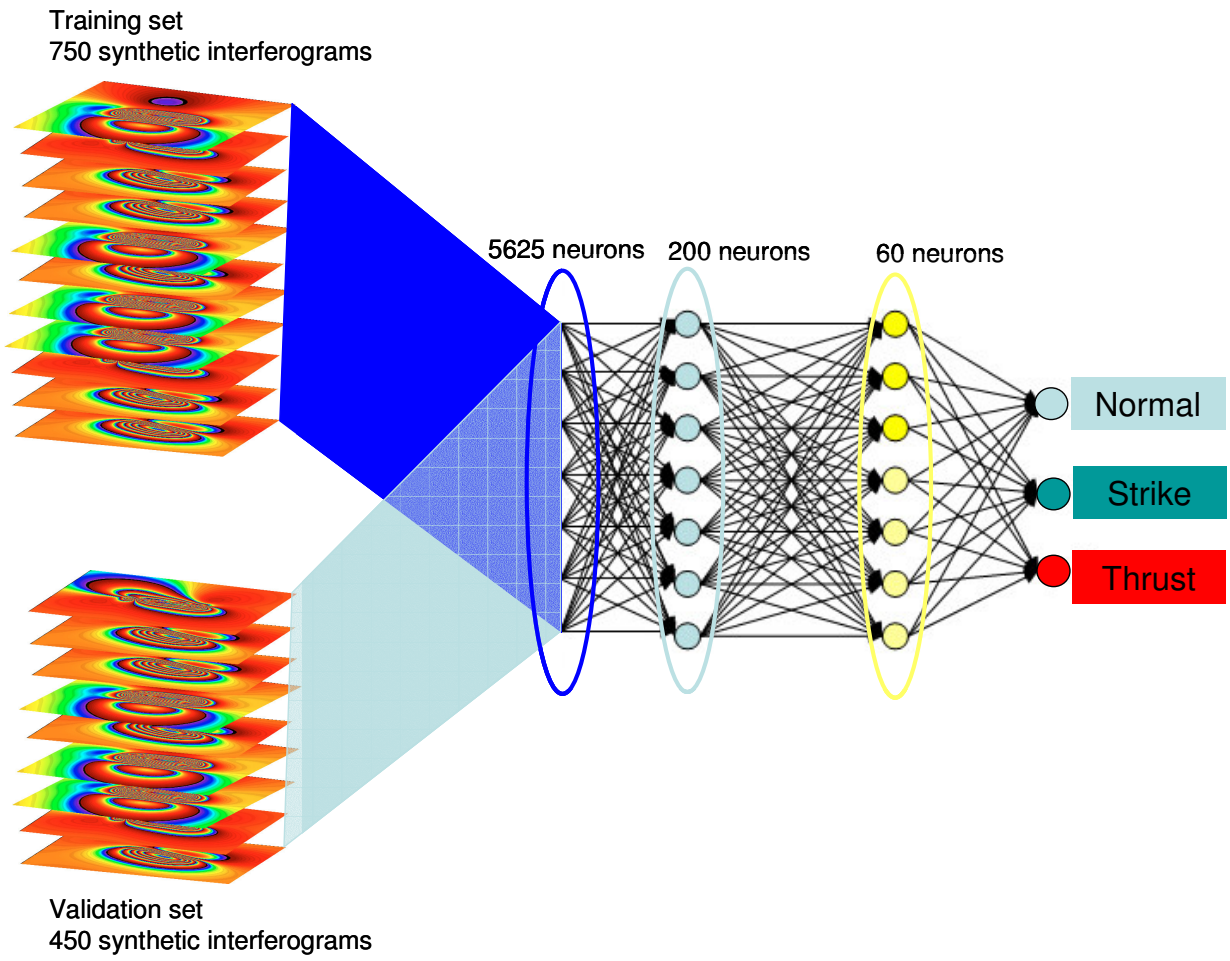


Fig. 54: General scheme of the neural network used for the seismic source classification exercise. The overall set of synthetic interferograms is split into the Training and the Validation test. The net has two hidden layers of 200 and 60 neurons. The three output classes refer to the possibile source mechanism.

In Table 9 the result of the classification exercise has been simply summarized. The trained neural network correctly assigned the belonging of each earthquake fault mechanism according to the information provided to the net by the coseismic interferograms. The geometry of the fringe pattern, its spatial extension and the fringe shape has been recognized by the network. Besides the positive result means that despite the stated procedure for extracting the information content reduce to 75 x 75 values each interferogram this is enough to characterize it. Furthermore the exercise demonstrated the effective capability of the neural network to achieve the expected result. In Table 8 are the output values of the network. These numbers between 0 and 1 provide an estimate of how correctly each source mechanism has been assigned to a fault class. Concerning the Colfiorito earthquake it has been undoubtedly recognized belonging to the Normal fault class. The decision for both Izmit and Bam to be assigned to the Strike slip fault is straightforward. Finally the Northridge earthquake is clearly an example of a Thrust or Inverse fault mechanism.

Earthquake / Source mechanism	**Normal fault**	**Strike Slip Fault**	**Thrust (reverse) Fault**
Colfiorito	**0.96629**	0.06156	0.01176
Izmit	0.03653	**0.95714**	0.04217
Northridge	0.05263	0.05109	**0.96284**
Bam	0.04876	**0.95507**	0.00268

Table 9: Schematic summary of the classification exercise. Each test case has been correctly classified according to its own fault mechanism. The numbers have been normalized between 0 and 1. They mean how clearly the network associated each earthquake to a fault class.

5.4 The fault parameter quantitative retrieval

As far as the parameter retrieval exercise is concerned after the Learning phase ended the trained neural network can be used to exploit its capabilities. To this aim the Test cases are the Izmit and the Bam earthquakes both belonging to the Strike slip fault class. The input data of the neural network are extracted by two coseismic interferograms, one for each case, computed to measure the overall displacement field. It is noteworthy that the interferogram of the Izmit earthquake does not contain the overall fringe pattern, being the displacement field beyond the surface extension of one

SAR image frame. The latter is about 100 x 100 km square whereas the length of the surface rupture exceeds 110-120 km. Looking at the fringe pattern it is manifest that the more external ones are tents km outside the interferogram. Concerning the Bam case the surface deformation is fully within the coseismic interferogram being the seismic fault plane quite smaller.
As already pointed out the parameter retrieval has been performed on three exercises with growing number of unknown factors:

- 2 parameter: Strike and Dip angles;
- 3 parameters: Strike and Dip angles, Bottom Depth;
- 5 parameters

In Table 10 are the unscaled parameters retrieved by the trained neural network. The first two exercises, 2- and 3-parameters, are solved using the [5625][60][2] and [5625][60][3] topologies. The estimated values for Strike-Dip and Strike-Dip-Depth are in fully agreement with those present in literature and retrieved by other techniques.

	Strike (deg)	**Dip (deg)**	**Depth (km)**
Izmit 2-parameters	90°	89°	
Izmit 3-parameters	89°	87°	16

Table 10: the table summarizes the geometric parameters retrieved for the Izmit case with the proposed approach for the two and three parameter test cases.

5.5 The fault plane retrieval

The conclusive test case concerns the retrieval of the overall set of five variable parameters as listed above. The parameters are the fault Length and Width, the Dip and Strike angles, the Depth.
To fully exploit the capability of the neural network the topology used for this exercise is the same one set for the classification test. The MLP network has thus the following topology with two hidden layers: [5625][200][60][5] (see Fig. 55).

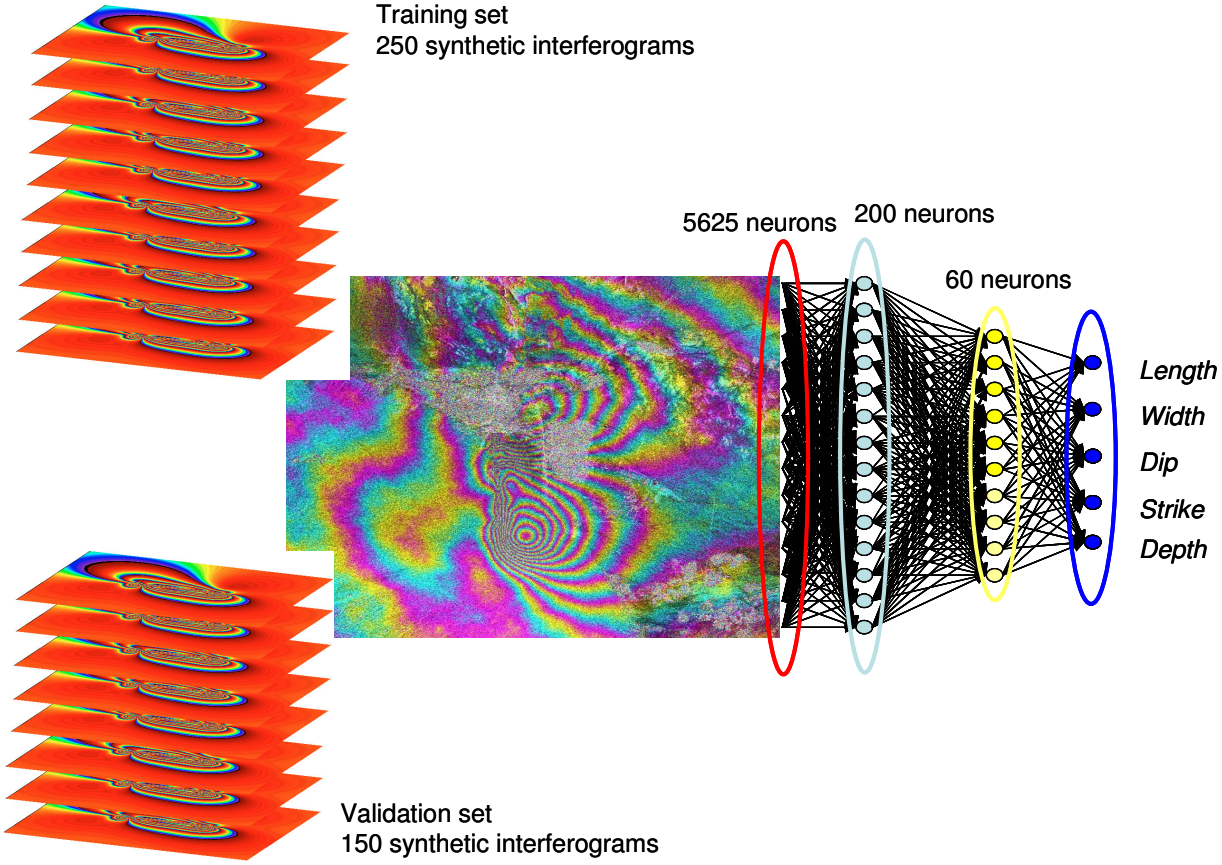


Fig. 55: General scheme of the neural network used for the quantitative retrieval of the geometric parameters. The overall set of synthetic interferograms is split into the Training and the Validation test. The net has two hidden layers of 200 and 60 neurons. The five outputs refer to the Width, Length, Dip, Strike and Depth.

The Test phase has been referred to two earthquakes, Izmit and Bam, both concerning the strike slip fault. The trained neural network received the input layer of 5625 samples extracted from the interferogram relative to the Izmit case. Afterwards the same number of samples relative to the Bam earthquake has been provided to the network. In Table 6 and 7 the five geometric parameters estimated by other approaches and already published in literature are summarized. The large range variability of such parameters means that according to the data used or the modeling approach, either forward or inverse, and finally the algorithm for the interferogram inversion, the results differ within a certain range. Despite this the main part of the retrieved parameters achieves values close to effective ones.

The results of the inversion based on the use of the trained neural network for both case studies are in Table 11. In particular for Izmit the Length, Width, Dip and Strike are in agreement with the corresponding ones in Table 6 within a reliable range.

The only Depth value is too shallow. The disagreement between this latter and the possible depth (see also Table 6) is probably due to the hypothesis of constant slip along the fault plane. Moreover as the coseismic surface displacement exceeds the extension of a single frame interferogram a significative portion of the overall displacement due to the earthquake is outside it. Therefore the lack of part of the deformation probably affected the depth estimation of the neural network.
Furthermore the exercise is fully successful in the Bam case where the retrieved values for the Length, Width, Dip, Strike and Depth (see Table 11) are close to those in Table 7.

	Length (km)	**Width (km)**	**Dip (deg)**	**Strike (deg)**	**Depth (km)**
Izmit	120	16.6	83	81	0.25
Bam	16.25	13.5	102	188	22.5

Table 11: Schematic summary of the network outputs relative to the five parameter Test phase for both Izmit and Bam cases. The results of the inversion of the differential interferogram have been opportunely rescaled.

To complete the parameter retrieval exercise the Okada formulation has been applied to the output values estimated by the network. The best fit model based on the InSAR data inversion has been computed and compared with the measured coseismic interferogram. The very large extension of the affected area avoids to fully comparing the estimated and measured displacement fields. Notwithstanding this the modeled fringe pattern well agrees with the measured interferogram wherever the latter is available (Fig. 56).
The output value provided by the neural network to perform the exercise for the Bam earthquake have been used to generate the best fit Okada model. The comparison with the coseismic interferogram points out the general agreement of the fringe pattern even tough further refinements are possible (Fig. 57). In particular the uniform slip distribution along the fault plane also for Bam seems causing the simplification of the fringe pattern thus avoiding the retrieval of possible local inhomogeneities.

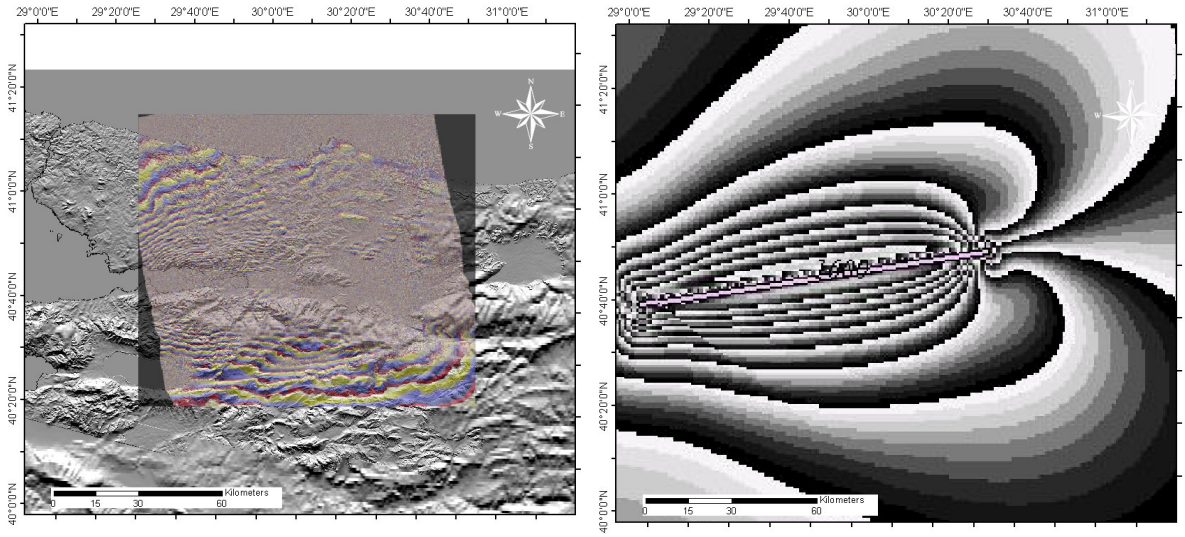


Fig. 56: Differential interferogram relative to the Izmit earthquake (right) and best fit inverse model based on the retrieved geometric parameters (left).

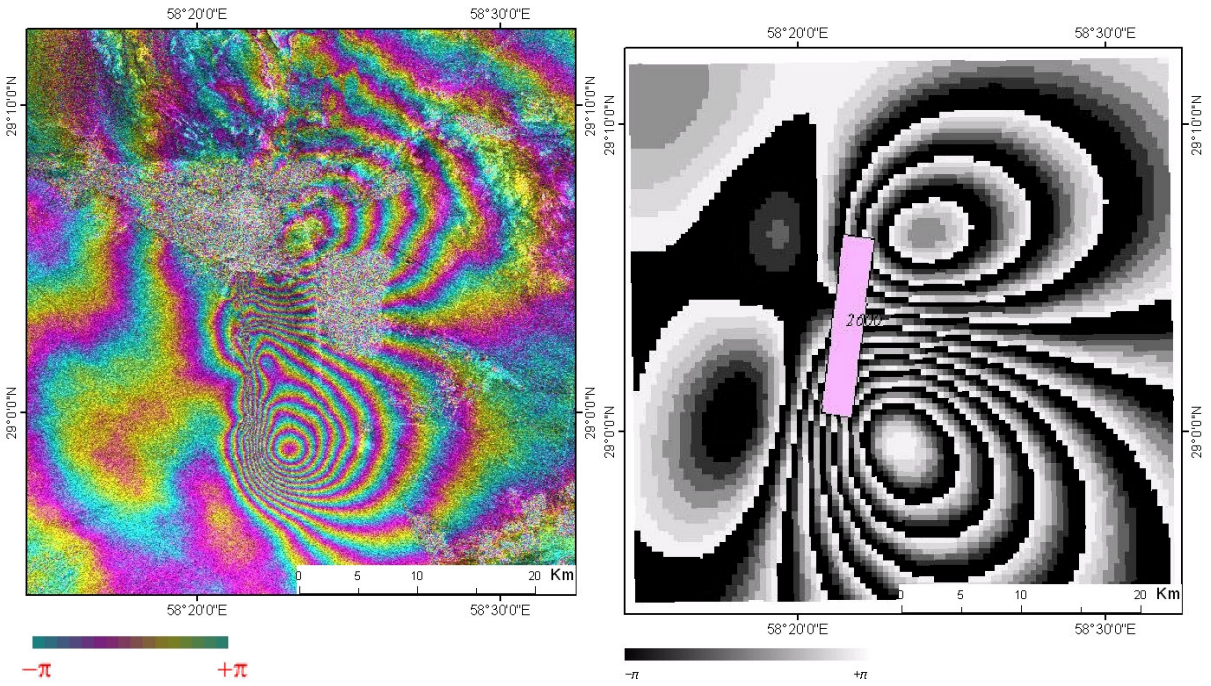


Fig. 57: Differential interferogram relative to the Bam earthquake (right) and best fit inverse model based on the retrieved geometric parameters (left).

6 Conclusions

The capabilities of neural networks have been exploited in a very new framework, i.e. the seismic source parameter retrieval. Despite the artificial neural networks have become useful tools for the solution of classification and inversion problems, anyway they have not been used in seismology until now. This is true for both the specific issues here discussed and analysed. As a matter of fact the detection and classification of the fault mechanism usually deals with the use of seismological measurements. Moreover the geometric parameter retrieval of the fault plane is also performed using still valid approaches involving the simulating annealing technique or other well known inversion methods. In this context the proposal of an innovative approach involving the artificial neural network and the Interferometric SAR technique relies on the purpose to exploit the effective potential of the synergic use of both.

An artificial neural network has been properly designed to perform an inversion procedure and single out the geometric parameters of the fault.

The proposed methodology involves different disciplines ranging from remote sensing to seismology. In fact the Multi Layer Perceptron neural network used synthetic interferograms during the Learning phase and coseismic data within the Test phase later. This confirms the very wide flexibility of the neural networks and its effective role among the available inversion tools.

As far as the schematic chain of this thesis is concerned it has to be noted that the innovative method required a preliminary effort to correctly assign the forward problem and to allow the application of the inversion scheme. To this aim the forward problem has been studied using the method by Okada whilst the inversion process has been planned from the beginning. This means that the procedure for the interferogram formatting and the data extraction procedures have been specifically conceived.

Two problems have been analysed: the seismic source classification according to the fault mechanism and the inversion of those parameters responsible for the determination of the geometry of the fault plane.

Concerning the first problem the neural network correctly identified the fault mechanism of the four test cases taken into account. Despite the coseismic interferograms are rather noisy the neural network clearly classified the Colfiorito case as a Normal fault, Northridge as a Thrust (or reverse), and Izmit and Bam as Strike slip faults. The decision of the network is effective for each case.

As far as the second problem is concerned five parameters among those determining the fault geometry have been simultaneously inverted. The inversion exercise has been performed on two seismic events dealing with the strike slip mechanism, Izmit and Bam earthquakes. The output

values conceived by the neural network for such parameters, Length, Width, Dip angle, Strike angle and Depth, are generally comparable with those already present in literature and estimated by other forward and inverse procedures.

The obtained results allow to consider the approach based on the neural networks and the interferograms a possible option to the already exploited methods. Furthermore a possible advantage of this approach relies on the high level of automation of the procedures. The proposed method is innovative and should be applied to further test cases. The very promising results lead to increasing the studies involving the neural networks and the satellite-based data. Further achievements should concern implementing a variable slip model and the relative inversion procedure of six parameters.

Bibliography

Billings, S. D., Simulated annealing for earthquake location, Geophys. J. Int., 118, 680-692, 1994.

Bishop C.M., “Neural Networks for Pattern Recognition”, Oxford Univ. Press, pp. 374--375, 1995.

Cakir Z., Jean-Bernard de Chabalier, Rolando Armijo, Bertrand Meyer, Aykut Barka and Gilles Peltzer, Coseismic and early post-seismic slip associated with the 1999 Izmit earthquake (Turkey), from SAR interferometry and tectonic field observations, Geophys. J. Int. (2003) 155, 93-110

Cocco M., Nostro C. And G. Ekstrom, Static stress changes and fault interaction during the 1997 Umbria-Marche earthquake sequence, *Journal of Seismology*, Vol.4, n.4, pp. 501-516, October 2000

F. Del Frate, G. Schiavon, "Nonlinear principal component analysis for the radiometric inversion of atmospheric profiles by using neural networks," IEEE Trans. Geosci. Remote Sensing, Vol. 37, pp. 2335--2342, 1999.

Del Frate F., Ferrazzoli P., and Schiavon G., Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks, Remote Sensing of Environment, Vol. 84, n. 2, pp. 174-183, February 2003

M. Iapaolo, "Automatic selection by means of neural networks of GOME optimum spectral channels for the retrieval of Ozone vertical profiles", 2003 Int. Geosci. Remote Sensing Symp., Toulouse, France, July 2003.

F. Del Frate, L. Salvatori, "Oil spill detection by means of neural network algorithms: a sensitivity analysis", IGARSS'04, ANCHORAGE, USA

Del Frate, C. Solimini "Change detection in urban areas with QuickBird imagery and neural networks algorithms", 3rd Int. Symp. Remote Sensing and Data Fusion Over Urban Areas (URBAN 2005) and 5th Int. Symp. Remote Sensing of Urban Areas (URS 2005), Tempe, Arizona, USA, March 2005

Feigl K.L., Francesco Sarti, He´le`ne Vadon, Simon McClusky, Semih Ergintav, Philippe Durand, Roland Burgmann, Alexis Rigo, Didier Massonnet, and Rob Reilinger, Estimating Slip Distribution for the Izmit Mainshock from Coseismic GPS, ERS-1, RADARSAT, and SPOT Measurements, Bulletin of the Seismological Society of America, 92, 1, pp. 138–160, February 2002

Franceschetti Giorgio e Riccardo Lanari, Synthetic Aperture Radar Processing, CRC Press

Fukahata, Y., Nishitani A. and M. Matsu'ura, Geodetic data inversion using ABIC to estimate slip history during one earthquake cycle with viscoelastic slip-reponse functions, GJI, 156, 140-153, 2004.

Gareth J. Funning, Barry Parsons, and Tim J. Wright, James A. Jackson, Eric J. Fielding, Surface displacements and source parameters of the 2003 Bam (Iran) earthquake from Envisat advanced synthetic aperture radar imagery, JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 110, B09406, doi:10.1029/2004JB003338, 2005

Gatelli F., Monti Guarnieri A., Parizzi F., Pasquali P., Prati C. and Rocca F., The Wavenumber Shift in SAR Interferometry, IEEE Trans. On Geosci. And Remote Sensing, vol. 32, n.4, pp. 855-864

Hernandez. B., M. Cocco, F. Cotton, S. Stramondo, O. Scotti, F. Courboulex and M. Campillo , Rupture history of the 1997 Umbria-Marche (central Italy) main shocks from the Inversion of GPS, DInSAR and near field strong motion data, Annals of Geophysics, vol.47, n.4, August 2004, pp. 1355-1376

Iwasaki, T. and R. Sato (1979). Strain field in a semi-infinite medium due to an inclined rectangular fault, J. Phys. Earth 27, 285-314.

J. Jackson, M. Bouchon, E. Fielding, G. Funning, M. Ghorashi, D. Hatzfeld, H. Nazari, B. Parsons, K. Priestley, M. Talebian, M. Tatar, R. Walker and T. Wright, Seismotectonic, rupture process, and earthquake-hazard aspects of the 2003 December 26 Bam, Iran, earthquake, *Geophys. J. Int.* (2006) 166, 1270–1292

Kanamori Hiroo, Quantification of Earthquakes, Nature, vol.271, n.5644, pp.411-414, February

1978, July 1994

Kirkpatrick, S. C., D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, Science, 220, 671-680, 1983

Lundgren P. and S. Stramondo, Slip Distribution of the 1997 Umbria-Marche earthquake sequence through joint inversion of GPS and DInSAR data, Journal of Geophysical Research, VOL. 107, NO. B11, 2316, doi:10.1029/2000JB000103, November 2002

Martinelli Giuseppe, Reti neurali e nurofuzzy, Editrice Universitaria di Roma – La Goliardica

Massonnet D. and Rabaute T., Radar Interferometry:Limits and Potential, IEEE Trans. On Geosci. And Remote Sensing, vol. 31, n.2, pp. 455-464, March 1993

Okada Y., Surface deformation due to shear and tensile faults in a half space, Bull. of the Seism. Soc. of America, vol. 75, n.4, pp. 1135-1154, August 1985.

Picardi Giovanni, Elaborazione del segnale radar, Metodologie ed applicazioni, Franco Angeli Collana scientifica

Prati C. and Rocca F., Focusing SAR Data with Time-Varying Doppler Centroid, IEEE Trans. On Geosci. And Remote Sensing, vol. 30, n.3, pp.550-559, May 1992

Reilinger, R. E., S. Ergintav, R. Bu¨rgmann, S. McClusky, O. Lenk, A. Barka, O. Gurkan, L. Hearn, K. L. Feigl, R. Cakmak, B. Aktug, H. Ozener, and M. N. Tokso¨z (2000). Coseismic and postseismic fault slip for the 17 August 1999, M _ 7.4, I˙zmit, Turkey earthquake, *Science* 289**,** 1519–1524.

Rondved, L. and J.T. Frasier (1958). Displacement discontinuity in the elastic half-space, J. Appl. Mech. 25, 125-128.

P. Rosen H.A Zebker, C. Werner. Accuracy of topografy maps derived from ERS-1 interferometric radar. Ieee transaction on geoscience and remote sensing, 32 (4):19617-19634, *July 1994*

Salvi S, Stramondo S., Cocco M., Tesauro M., Hunstad I., Anzidei M., Briole P., Baldi P., Sansosti E., Lanari R., Doumaz F., Pesci A. and A. Galvani (1999) – "Modeling Coseismic Displacements resulting from SAR Interferometry and GPS measurements during the 1997 Umbria-Marche seismic sequence", *Journal of Seismology*, Vol.4, n.4, pp. 479-499, October 2000

Sato, R. and M. Matsu'ura (1974). Strain and tilts on the surface of a semi-infinite medium, J. Phys. Earth 22, 213-221.

Savage, J.C. and L.M. Hastie (1966). Surface deformation associated with dip-slip faulting, J. Geophys. Res. 71, 4897-4904.

Steketee, J.A. (1958). On Volterra's dislocation in a semi-infinite elastic medium, Can. J. Phys. 36, 192-205

Stramondo S., Tesauro M., Briole P. Sansosti E., Salvi S., Lanari R., Anzidei M., Baldi P., Fornaro G., Avallone A., Buongiorno M.F., Franceschetti G., Boschi E. (1999) – "The September 26, 1997 Colfiorito, Italy, earthquakes: modeled coseismic surface displacement from SAR Interferometry and GPS", *Geophys. Res. Lett., Vol. 26, n.7, pp.883-886*

Stramondo S., F. R. Cinti, M. Dragoni, S. Salvi, S. Santini, The August 17, 1999 Izmit, Turkey, earthquake: new insights on slip distribution from dislocation modeling of DInSAR and surface offset, *Annals of Geophysics*, vol. 45, n.3/4, pp. 527-536, June/August 2002

S. Stramondo, M. Moro, C. Tolomei, F.R. Cinti, F. Doumaz, InSAR surface displacement field and fault modelling for the 2003 Bam earthquake (southeastern Iran), *Journal of Geodynamics*, vol. 40, Nos. 2-3, pp. 347-353, September-October, 2005

Talebian M., Eric J. Fielding, Gareth J. Funning, Manoucher Ghorashi, James Jackson, Hamid Nazari, Barry Parsons, Keith Priestley, Paul A. Rosen, Richard Walker, and Tim J. Wright, The 2003 Bam (Iran) earthquake: Rupture of a blind strike-slip fault, Geophysical Research Letters, VOL. 31, L11611, doi:10.1029/2004GL020058, 2004

Hong Kie Thio and Hiroo Kanamori, Source Complexity of the 1994 Northridge Earthquake and Its Relation to Aftershock Mechanisms, Bulletin of the Seismological Society of America, Vol. 86, No. 1B, pp. 84-92, February 1996

David J. Wald, Thomas H. Heaton and Kenneth W. Hudnut, The Slip History of the 1994 Northridge, California, Earthquake Determined from Strong Ground Motion, Teleseismic, GPS, and Leveling Data, *Bull. Seism. Soc. Am., 86, S49-S70*

H. Zebker and J. Villasenor*, Decorrelation in Interferometric Radar Echoes, IEEE Transaction on Geoscience and Remote Sensing, Vol 30, N° 5, September 1992*

H. A. Zebker, J. Villasenor, "*Decorrelation in Interferometric Radar Echoes",* IEEE Trans. On Geoscience and Remote Sensing, VOL. 30, N° 5, pp. 823-836, 1994

Zell A. et al., "SNNS Stuttgart Neural Network Simulator User Manual", Report N6/95, University of Stuttgart, Institute for Parallel and Distributed High Performance Systems, Stuttgart, Germany, 1995.