Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

GEOFISICA

Ciclo XXXII

Settore Concorsuale di afferenza: 04/A4 – GEOFISICA

Settore Scientifico Disciplinare: GEO/10 – GEOFISICA DELLA TERRA SOLIDA

Analysis of slip distribution of large earthquakes oriented to tsunamigenesis characterisation

Presentata da: **Enrico Baglione**

Coordinatore Dottorato:

Prof. Nadia Pinardi

Supervisore:

Prof. Stefano Tinti

Esame finale anno 2020

Abstract

The present thesis focuses on the on-fault slip distribution of large earthquakes in the framework of tsunami hazard assessment and tsunami warning improvement. It is widely known that ruptures on seismic faults are strongly heterogeneous. In the case of tsunamigenic earthquakes, the slip heterogeneity strongly influences the spatial distribution of the largest tsunami effects along the nearest coastlines. Unfortunately, after an earthquake occurs, the so-called finite-fault models (FFM) describing the coseismic on-fault slip pattern becomes available over time scales that are incompatible with early tsunami warning purposes, especially in the near-field.

Our work aims to characterize the slip heterogeneity in a fast, but still suitable way.

Using finite-fault models to build a starting dataset of seismic events, the characteristics of the fault planes are studied with respect to the magnitude. The patterns of the slip distribution on the rupture plane, analysed with a cluster identification algorithm, reveal a preferential single-asperity representation that can be approximated by a two-dimensional Gaussian slip distribution (2D GD). The goodness of the 2D GD model is compared to other distributions used in literature and its ability to represent the slip heterogeneity in the form of the main asperity is proven. The magnitude dependence of the 2D GD parameters is investigated and turns out to be of primary importance from an early warning perspective.

The Gaussian model is applied to the 16 September 2015 Illapel, Chile, earthquake and used to compute early tsunami predictions that are satisfactorily compared with the available observations. The fast computation of the 2D GD and its suitability in representing the slip complexity of the seismic source make it a useful tool for the tsunami early warning assessments, especially for what concerns the near field.

Table of contents

1 Introduction

This PhD thesis investigates the heterogeneity of the coseismic slip on the fault and its modelling in the context of tsunamigenesis. It is known that the slip distribution on the earthquake fault plane is not homogeneous. Nevertheless, a homogeneous slip is adopted in several applications including the modelling of tsunamis induced by earthquakes.

The objective of this study is to provide a strategy for obtaining a seismic source model that:

- can be derived in a very short time after the earthquake occurs;
- considers the slip heterogeneity on the fault surface;
- is not overly complicated, but is sophisticated enough to allow for a satisfactory approximation of the generated tsunami.

To achieve this goal, the thesis is structured in the following main steps:

1) to build a dataset of finite fault models (FFM), a procedure outlined in Chapter 2;
2) to derive the rupture dimensions and the main features of the FFMs (Chapter 3);
3) to find the slip distribution characterizing the on-fault slip heterogeneity using the FFMs (Chapter 4);
4) to apply the characteristic distribution model derived from reference FFMs to a real tsunamigenic earthquake (the 2015 Illapel, Chile earthquake) and to study its performance on tsunami calculations (Chapter 5)
5) to obtain the characteristic distribution model for the same earthquake in an alternative way, more adequate for tsunami early warning applications, that is to derive it from suitable scaling laws by using information such as earthquake location and magnitude that are available soon after the event (Chapter 6).

1.1 The collection of finite fault models

The data used for this study are the finite-fault models (FFM) of recent (after 1990) large earthquakes (magnitude larger than 6), present in the SRCMOD online database (http://equake-rc.info/SRCMOD/).

These slip models were obtained by using different data sources (geodetic, strong motion, teleseismic, local P waves, inSAR, tsunami, and, in the best cases, a combination of two or more of them), and

different crustal models, and by different inversion techniques and stabilization methods, and different spatial sampling (Manighetti et al., 2005).

1.2 The scaling laws

Starting from the chosen FFM dataset, scaling laws between the geometrical properties of the fault and the magnitude are obtained. Several studies have investigated earthquake source-scaling properties, differing as regards the amount of data, the classifications (fault slip type, tectonic regions, tsunamigenic and non-tsunamigenic, etc.), the regression techniques. Nevertheless, the epistemic uncertainty in regression formulation is very large, and preferred methods have not yet been defined in the international seismological community (Stirling et al. 2013). In this thesis, the scaling laws derived from the FFMs are treated only for the sake of completeness, but they will not play any role in the development of our strategy.

Parallel to the FFM regressions of fault length, width, area, aspect ratio, and average and maximum slip against the magnitude, also analogous laws concerning the relative positions of the hypocentre and the maximum displacement on the fault are investigated.

1.3 The on-fault slip distribution

The distribution of the slip over the fault has been the object of numerous studies especially in the last decades, thanks to the source-inversion methods and the rapid growth of the available amount of data.

The study of the slip heterogeneity spans different fields of investigation, from a more theoretical to a more empirical one. The starting point of the modern fault slip theory can be attributed to Haskell (1964), who introduced a fault model as a rupture that propagates over a finite section of a fault, providing analytical formulas for the radiated elastic waves. The Haskell model presents a uniform release of displacement and stress on the fault surface. The author identified this as a deficiency, suggesting that some form of heterogeneity in the slip distribution was required either in space or time.

Thatcher (1990), with the aim of relating long wavelength features of the slip space distribution to observations and to models of earthquake recurrence, suggested that earthquake recurrence is more likely to be controlled by the maximum rather than the average fault slip. Further, he affirmed that the pattern of earthquake slip is typically very irregular, as indicated by slip seismologic determinations or by seismic moment release distribution.

Mai and Beroza (2000) identified a stochastic characterisation of earthquake slip complexity, modelling the slip as a Von Karman autocorrelation function (ACF) model, with parameters depending on the source dimensions. They suggested the use of their stochastic slip model to generate realizations of scenario earthquakes for near-source ground motion simulations.

Somerville et al. (1999), examined systematic features of slip models summarised by Mendoza and Hartzell (1988b) and Heaton (1990). From a rectangular representation of the fault rupture, they delineated the concept of asperity as a region on the fault rupture surface having large slip relative to the average slip. They proposed slip heterogeneity characterisation through rectangular asperities.

The importance of asperities, as zones of large and concentrated seismic displacement, is one of the key features to describe slip heterogeneities, and represents our starting point to develop a simple but not uniform characterization of the slip on the fault.

The concept of asperity is often associated with the stress on the fault and widely studied for the fault dynamical processes. Although our study concentrates more on static displacement fields, nevertheless we will mention some examples.

Myatake (1992) reconstructed the dynamic rupture process of four large Japanese earthquakes exploiting the concept of "weak" and "strong" asperities, considering the strength excess characterising the stress drop regions on the fault planes.

Johnson and Nadeau (2002) developed an earthquake asperity model, based on the concept of small strong asperities that resist the motion on the fault. They justify the accumulation of significant tectonic displacement before failure occurs in terms of interaction between close asperities. One of the main characteristics of their asperity model is that the occurrence of an earthquake is driven by stress concentrations that can be several orders of magnitude larger than the average stress on the fault. Hence, they pointed out how interpretations founded on this feature of a heterogeneous stress field may help explain certain aspects difficult to be explained in terms of homogeneous average stress.

Dragoni and Lorenzano (2017) modelled the fault as a discrete dynamical system, whose state is described by two variables expressing the slip deficits of two asperities, describing in a unique frame both seismic and aseismic slip on a fault. They applied their model to the fault of the 2011 Tohoku-Oki earthquake. They found that the amount of after-slip is found to be proportional to the seismic slip of the asperity, with a proportionality factor depending on the geometry of the fault and on the velocity of tectonic motion.

Our interest in the slip distribution leads us to privilege the definition of asperity as a region with intense value of slip, in particular with respect to its maximum value on the fault plane. We have therefore developed a single asperity model which identifies the main asperity slip pattern with a bivariate Gaussian distribution (hereafter called 2D GD).

1.4 Source slip distribution and tsunami hazard assessment

The interest in the modelling of slip on the fault is mainly motivated, in this research work, in terms of its direct influence on the tsunami field produced by a strong earthquake.

Very often, especially in tsunami warning operations, the assumption is made that the slip is uniform over the entire rupture plane, but, as modern teleseismic and geodetic inversion techniques have confirmed, this assumption is invalid and may lead to unsatisfactory estimations of the tsunami waves height at the coast. Indeed, the ground motion and tsunami heights associated with large earthquakes are controlled by the size and distribution of asperities. Therefore, an understanding of the scaling relations between the parameters characterising big earthquakes is essential for an accurate assessment of future earthquake hazards (Murotani et al. 2013).

Despite scaling relationships evaluating length and width of the fault plane as a function of moment magnitude are sufficient to determine the basic geometry of an earthquake rupture, the randomness and heterogeneity of spatial source slip distribution have major influence on tsunami hazard assessment. This fact is now documented by several studies.

Variations in the local tsunami wave field in relation to heterogeneous slip distributions were the object of study by Geist and Dmowska (1999). They reported that, in most cases, the assumption of uniform slip in dip direction leads to a significant underestimation of the maximum amplitude and leading wave steepness of the local tsunami. Geist (2002), investigating the effect of rupture complexity on the local tsunami wave field, pointed out that the variability in local tsunami runup scaling can be ascribed to tsunami source parameters that are independent of seismic moment. Among the latter, the complexity of the heterogeneous slip distribution patterns results of considerable importance.

Løvholt et al. (2012), with the aim of quantifying the effects of the non-uniform coseismic slip in the dip direction on stochastic tsunami runup variability, investigated how heterogeneous coseismic slip affects the initial water surface elevation and the subsequent tsunami runup occurring along a coastline in the near field for a hundreds of stochastic slip realizations. Exploring the correlations between the maximum runup with the different seabed displacements induced by the slip

distributions, they found, as most relevant parameters, the scaled seabed volume displaced per unit length and the maximum peak-to-peak vertical seabed displacement.

These studies received considerable attention after the great Sumatra-Andaman earthquake of the 26 December 2004. The 2004 earthquake ($M_W > 9$), indeed, was the first event to be recorded by a global network of broadband seismic stations and regional GPS networks and plays a role of historical importance in the field of tsunami hazard and forecasting.

Alongside the factors affecting the distribution of tsunami heights, great importance was also given to time needed to provide a possible tsunami scenario. Geist et al. (2006) compared different measures of tsunami generation for the Sumatra earthquakes of December 2004 and March 2005 ($M_W = 8.7$) with the aim of determining which information about the earthquake is needed to rapidly assess tsunami generation and local tsunami severity. The authors affirmed that many of the differences between the two tsunamis relate to slip distribution patterns for the two earthquakes. Indeed, while point- and line-source parameterization of the earthquake can accurately predict far-field tsunami amplitudes, local tsunami runup is dependent on propagation paths from localized regions of seafloor displacement, arising from earthquake rupture complexity: this fact introduces significant uncertainty in local tsunami warning systems (Geist et al. 2006). The study underlined also that determining reliable finite fault solutions within the time tsunamis take to arrive at shore is extremely difficult with the current state of knowledge and technology. The accuracy that can derive from seismic inversions of slip distributions must inevitably clash with the time required to obtain them.

Always referring to the 2004 Sumatra-Andaman earthquake, Sobolev et al. (2007) demonstrated that two earthquakes with the same magnitude, location and fault geometry but different distribution of slip may generate tsunami waves with drastically different impacts on the coast (Padang coast, West Sumatra, in particular). They found that the presence of islands between the trench and the Sumatran coast makes earthquake-induced tsunamis especially sensitive to the slip distribution on the rupture plane. The authors highlighted also the fact that for local tsunamis the near-field tsunami heights are controlled by the slip variability rather than by the seismic moment.

However, events like the 2004 Sumatra earthquake should be treated cautiously in terms of source and tsunami modelling, being these models at the upper bound of the known scaling relationships and due to the fact that mega-thrust subduction earthquakes may have very long fault rupture zones, exceeding 1000 km (Goda et al., 2016).

The 11 March 2011 Tohoku earthquake, differently from the Sumatra one, was characterised by a small rupture area compared to its magnitude, while the estimated maximum on fault displacement reached extremely high values (about 50 metres). Also in this case, a heterogeneous source model is essential to simulate the observed distribution of the runup correctly, as reported by Løhvolt et al. (2012).

Using the aforementioned Von Karman ACF model, Goda et al. (2014) developed stochastic earthquake slip models for the 2011 Tohoku and evaluated the impact of earthquake slip and fault geometry on tsunami simulation results in terms of near-shore sea surface profiles and inundation height. Their results highlighted strong sensitivity of tsunami wave heights to site location and slip characteristics, and also to variations in dip. The authors, moreover, affirmed that future investigations should assess the influence of earthquake slip distributions and fault geometry when quantifying the variability of impact assessments for tsunami hazards.

Other tsunamigenic earthquakes object of numerous studies are those characterizing the Pacific coast of Chile, where the Nazca plate subducts beneath the South American plate. The largest modern event was the 22 May 1960 (M_w=9.5) earthquake. Considering the last decade, the 27 February 2010 Maule earthquake (Mw 8.8), the 1 April 2014 Iquique earthquake (Mw 8.1) and the 16 September 2015 Illapel earthquake (Mw 8.3) are noteworthy. In particular, the 16 September 2015 Illapel earthquake is the one selected to exemplify our idea of simplified and fast model of slip distribution to compute early tsunami predictions.

Overcoming the temporal obstacle that affects seismic source inversion, a 2D GD slip pattern is easily derived directly from standard earthquake parameters like magnitude and hypocentre location. The impact of our model on tsunami simulations is evaluated in terms of tsunami waveforms and run-ups. We will show that the 2D GD is a reliable representation of the seismic source that can be obtained in a time interval compatible with timely tsunami early warning assessments.

2 The dataset

In this chapter we give a brief description of the database we used in the analyses described in the following chapters. In particular, we illustrate the procedure adopted to build a unique and homogeneous database of earthquake events and show some simple statistics on the earthquakes with magnitude larger than 6 occurred over the last 30 years. We restrict statistics to those events for which Finite-Fault Models (FFMs) are available, which turns out to be a rather small, but still significant, percentage of the total number of worldwide earthquakes.

2.1 Data collection

The first step of the present thesis consisted in the collection of the available FFMs. These can be found in published papers or in specific databases, the latter alternative being preferable for a PhD thesis work since the former one requires a long time for data mining, i.e. for finding and homogenizing a sufficient amount of data. The FFMs considered here were taken from the SRCMOD database (http://equake-rc.info/SRCMOD/).

Finite-fault inversions have become a topic of increasing interest in seismological research, since they allow a better understanding of the rupture mechanism and rupture evolution on the seismic fault. The data used in inversions can be of different types. They may be geodetic data of final deformation, in which case source inversions put constraints on the fault geometry and on the static slip distribution (i.e. the final displacements over the fault surface) (Mai and Thingbaijam, 2014).

Often, joint inversions that combine available geodetic, seismic, tectonic (and when it is the case also tsunami) data are conducted to match all, or most of, the observations and provide a more detailed representation of the rupture process. Some joint inversions use all data simultaneously. Others, instead, follow an iterative approach where one set of observations is used to build an initial model to be used in following inversions where other available data are considered. The field of finite-fault inversion was pioneered in the early 1980s (see Olson and Apsel, 1982; Hartzell and Heaton, 1983). Subsequently, the method has been applied to numerous earthquakes (e.g., Hartzell, 1989; Hartzell et al., 1991; Wald et al., 1991; Hartzell and Langer, 1993;Wald et al., 1993; Wald and Somerville, 1995), while additional source-inversion strategies were developed and applied (e.g., Beroza and Spudich, 1988; Beroza, 1991; Hartzell and Lui, 1995; Hartzell et al., 1996; Zeng and Anderson, 1996).

Finite-fault source inversions highlight the complexity of the earthquake rupture process. The source images obtained give useful information, although with a restricted spatial resolution, on the earthquake slip at depth, and potentially also on the temporal rupture evolution. Therefore, they are of paramount importance for works on the mechanics and kinematics of earthquake rupture processes and play an important role in our comprehension of earthquake source dynamics.
It is beyond the scope of this thesis to treat and discuss source-inversion methods that are the basis of the FFMs we use. In the next sub-sections we delineate the data source from which these models were taken.

2.1.1 SRCMOD

The SRCMOD website is an online database of finite-fault rupture models of past earthquakes. These earthquake source models are obtained from inversion or modelling of seismic, geodetic and other geophysical data, and characterize the space-time distribution of kinematic rupture parameters (from *http://equake-rc.info/SRCMOD/*).
The current version of the SRCMOD dataset provides earthquake scientists, source modellers, and any interested users with open access to more than 300 earthquake rupture models corresponding to about 100 earthquakes, in a unified representation, published over the last 30 years. The website is built on a three-tier architecture, which comprises client-side software (data presentation), server-side coding (data processing), and the back-end data storage.
Three are the file formats used to store the primary data for the source models: 1) MATLAB (http://it.mathworks.com/products/matlab) structures (.mat files), 2) ASCII files containing finite-source parameters (.fsp files), and 3) ASCII files containing a comprehensive slip model (.slp files). We actually focus our attention on the last data-file type.
The entire database represents an inhomogeneous global collection of earthquake rupture models. Inhomogeneous in the sense of:
- faulting type
- location of the earthquake and, consequently
- tectonic province (interplate, intraplate, subduction)
- data and observations used in the source inversion
- inversion techniques applied
- available rupture-model information provided by the authors
- model parameterizations selected and modelling choices made by the modellers.

A finite-fault (also known as kinematic) rupture model typically comprises several parameters, which include the seismic slip, rise time (local duration of the slipping process), rupture-onset time, and rake (angle of slip direction) (Mai and Thingbaijam, 2014). The source studies do not always invert for all these parameters. Some of them could have been fixed/assumed in advance, depending on data and the used inversion technique. The parameters may vary spatially: they are defined at node points or subfaults that constitute the rupture surface. In case of inversions using seismic data, the source time function describes the temporal slip evolution on each point of the fault and is typically chosen using either a simple parametric shape or a linear combination of a number of elementary slip functions (so-called multi time-window inversions) (Mai and Thingbaijam, 2014).

The spatial resolution of the model is defined by the size of the subfaults (or spacing of node points); typically, the details of the rupture process are resolved at a larger scale as a result of the chosen smoothing constraints or of the regularization method (to handle ill-posed inversion problems) and of the trade-off between parameters (Mai et al., 2007; Monelli et al., 2009).

Thanks to the increased availability of seismic and geodetic networks, the FFMs of recent earthquakes are more detailed comparing to those of the previous decades. The number of available source models is also affected by the contribution of fast FFMs, generated from the so-called fast finite-fault inversions in a semi-automatic way within days of a sizeable earthquake and then published online on institutional web pages. But more accurate models are available with some delay, typically after several months.

The variability in rupture models is due to the variety of source-inversion methods and available data. In the current SRCMOD database several earthquakes have more than two rupture models. There can be a significant difference between different source models for the same event. Multiple models are available, for example, to name a few, for the 1992 Landers, the 1995 Kobe, the 1999 İzmit, the 2004 Sumatra, the 2008 Wenchuan and the 2011 Tohoku earthquakes. However, the nominal uncertainties of each of the source inversions are not well known. This has been pointed out previously (e.g., Beresnev, 2003) and can be understood in the context of the data used, the model parameterizations chosen, and the inversion techniques applied in such studies (Cohee and Beroza, 1994; Das and Suhadolc, 1996; Henry et al., 2000; Graves and Wald, 2001; Yokota et al., 2001; Delouis et al., 2002).

Regarding the characteristics reported for every event, the earthquake source dimensions (i.e. length, width or area) are generally estimated prior to the inversion from the spatial distribution of aftershocks. Regarding the rupture width, the thickness of the seismogenic crust is often used to constrain it. In other cases, the fault plane size is estimated using source-scaling relationships.

However, the inversion procedures commonly assume conservatively large source dimensions, so that the entire rupture can be accommodated, which leads to an overestimation of the rupture sizes with small (even zero) displacements at the fault boundaries. Thus, it is necessary to trim the rupture model faults by eliminating superfluous small slips at the fault edges (Mai and Thingbaijam, 2014). In the following sections we will discuss about this procedure.

As already pointed out in this section the main analyses carried out in this thesis work are based on FFMs of the SRCMOD database. The main reason is that SRCMOD is a collector of models also from other databases, that are incorporated after accurate quality check. Moreover, it is a user-friendly online platform, easy to use.

2.2 Fault size

Establishing the rupture size of a fault (fault length or fault width) may be problematic and represents a sensitive issue. Different criteria can be used to determine the involved parameters.
In case of surface rupture occurrence, the fault length can be derived from the visible surface-rupture length LS. The uncertainties in the measurement of LS can derive from:

- incomplete field surveys and studies of the broken fault area;
- outcrops deficiency at some localities;
- uncertain interpretation of properties and numbers of surface ruptures by different scientists;
- different reported values of observable values by different field workers.

Alternatively, it can be estimated from the spatial extent of early aftershocks. But these two approaches can give substantially different results (Wells and Coppersmith, 1994).
In addition, since many aftershocks occur on sub-faults rather than on the main fault as a result of stress diffusion, the fault area estimated from the aftershock area could be biased and larger than the real one.

As for the width, the determination of the fault width relies on aftershocks only, or, which is the case for large strike-slip earthquakes, on the thickness of the seismogenic layer (Scholz, 1994).
Hence, the main errors in the evaluation of the fault dimension parameters can be found in the temporal evolution of the aftershock zone, in the accuracy of aftershock locations in 3D, in the interpretation of the initial extent of the aftershock sequence, in the reliability of the model.

Typically, the overall dimension of the fault area to be discretized is pre-assigned based on some accepted criterion, such as general scaling laws relating earthquake magnitude and total fault area. It may happen that the predefined fault area is larger than needed.

Hence, depending on the inversion algorithm accuracy and on the quantity, quality and spatial distribution of the data to be inverted, the resulting slip matrix, that is the variable slip over the discretized rupture plane, may exhibit rows and/or columns full of zeroes along the edges. In those cases, the rupture length and width reported by the FFMs overestimate the true dimensions.

2.3 Fault size correction and events removal

In order to get an image of the fault area affected by the seismic dislocation, we operated directly on the slip matrix with the criterion that, if there is an empty (i.e. full of zero slip values) row and/or column along the edges of the rupture plane, this is removed from the original source model.

The removal of part of the fault plane from the FFM implies a corresponding update of the fault dimensions and a change of the aspect ratio (defined as the ratio between length and width of the fault). The strike-slip events are those with the lowest average width (24 km), and consequently the largest average aspect ratio (4.60), while dip-slip events present the lowest values (1.81). The dip-slip events (normal and reverse) are the most affected by the trimming procedure with the down-dip dimension more altered than the along-strike one. Indeed, the recalculated aspect ratio for dip-slip events reveals to be higher than the original one. For the strike-slip earthquakes the opposite happens: the final average aspect ratio is smaller than the original one.

Table 2.1: Earthquakes' source average geometric characteristics after the trimming procedure (schematized in Figure 2.1) for different slip-types.

	Strike-slip	Normal	Reverse	Oblique	All
Minimum aspect ratio	1.07	0.89	0.74	0.95	0.74
Average aspect ratio	4.60	2.15	1.81	2.00	2.5
Minimum Area (km^2)	150.0	150.0	411.25	90.0	90.0
Average Area (km^2)	4205	3370	34500	13670	16365
Minimum Length (km)	14	15	17	10	10
Average Length (km)	124	72	211	127	143
Minimum Width (km)	10	10	18	9	9
Average Width (km)	24	33	104	66	64
Minimum mean slip (m)	0.144	0.136	0.120	0.068	0.068

Average mean slip (m)	1.45	1.03	1.07	1.35	1.24

The procedure of deleting the empty rows and columns on the edges does not exclude the possibility for some solutions to have many empty cells on the fault's plane or to present some very small values on the edges that do not permit to remove the edge itself. This led us to an additional control. For each model the percentage of zero (PZ) slip values in the slip matrix has been computed and, when higher than 50%, has been taken as an index of a possible anomaly that required an individual checking. Usually, these anomalous events are characterised by some isolated cells with small slip values that are on the edge of the fault. Killing these cells has the effect that a full row or a full column turns out to be empty and can be removed. If this "blanking" procedure produces a fault model with an empty cell percentage that is still too high (larger than 50 %), the fault is removed from the database, otherwise it is kept.

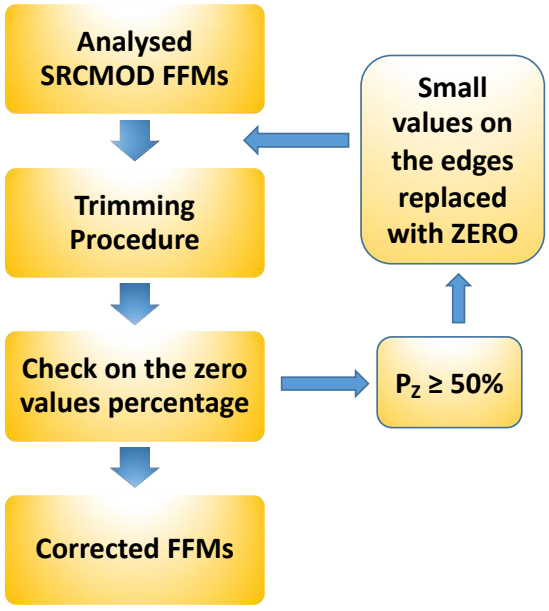

Figure 2.1 Scheme followed to check the SRCMOD FFMs.

2.4 Events with more than one solution

Some of the events considered have more than just one FFM, which poses a problem of selection. Generally, we have selected the most recent solution, but sometimes, after cross-checking different fault representations, we have made a different choice. With events with a high number of solutions, such as the 11 March 2011 Tohoku, Japan, those models that presented solutions very different from the average have been rejected.

Often, one of the discriminating criteria was the degree of fullness of the FFM slip matrix. In particular, for the event of 15 August 2007 in Peru, we noticed the lack of slip values different from zero in a significant portion of the matrix (Figure 2.2, left image) only after we made the visual control. So the FFM was removed from our database and an alternative model was included. The two models are shown, one next to the other, in Figure 2.2.

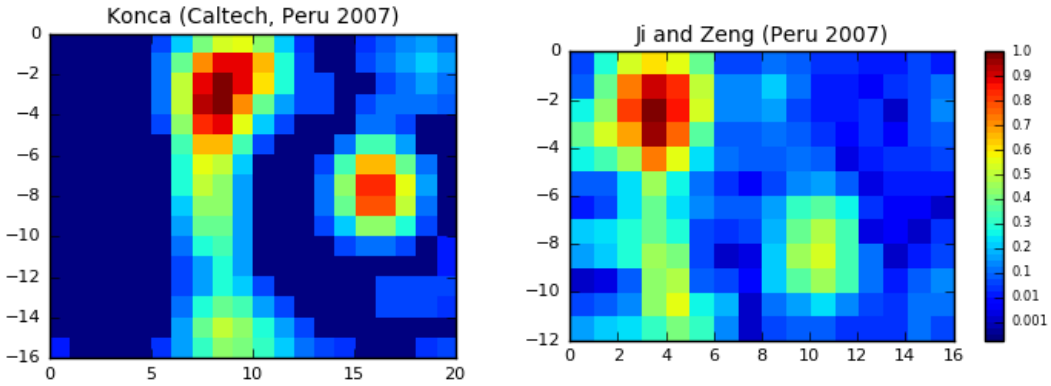

Figure 2.2 FFMs of the 15 August 2007 Peru earthquake by two different authors. On the left, the discarded model; on the right, the chosen model.

2.5 The selected FFMs models

The set of the earthquakes we select for the present analysis is summarised in Table 2.1. Their total number resulted to be 105. An Identification Number (I.N.) is reported for each model as well as: date, slip type, magnitude, seismic moment, hypocentre depth, rupture dimensions, fault area, maximum and average displacement. In the last column of the Table the author/authors of the FFM are also given.

The next two graphs (Figure 2.3, Figure 2.4) show the distribution of the selected 105 FFMs per magnitude and slip type. The majority of the earthquakes fall in the range $7 \leq M_W \leq 8$ with a peak of 16 events centred in *7.3*, while 17 events have magnitude larger than *8*.

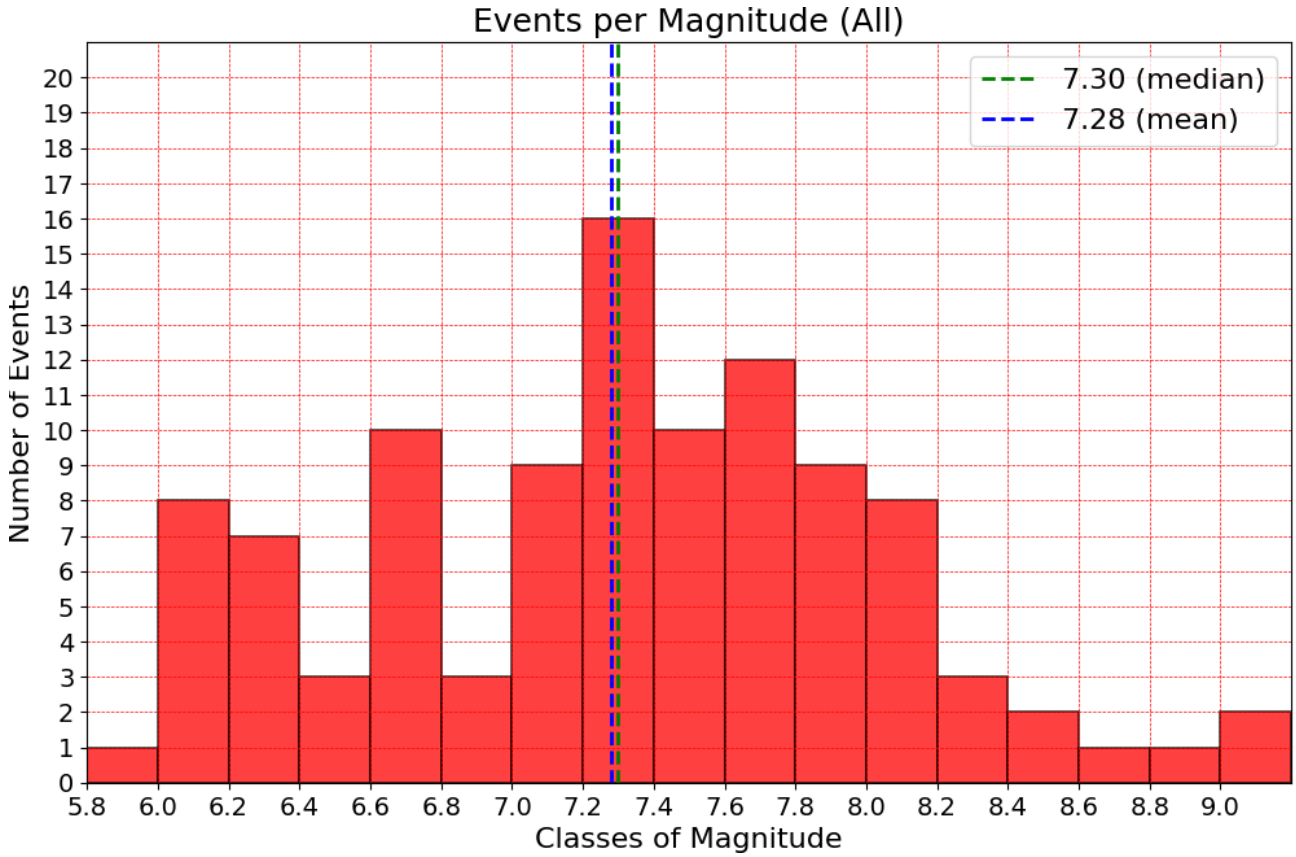

Figure 2.3 Number of FFMs per classes of Magnitude (step = 0.2 Mw). The dashed blue line represents the mean value and the green one the median value.

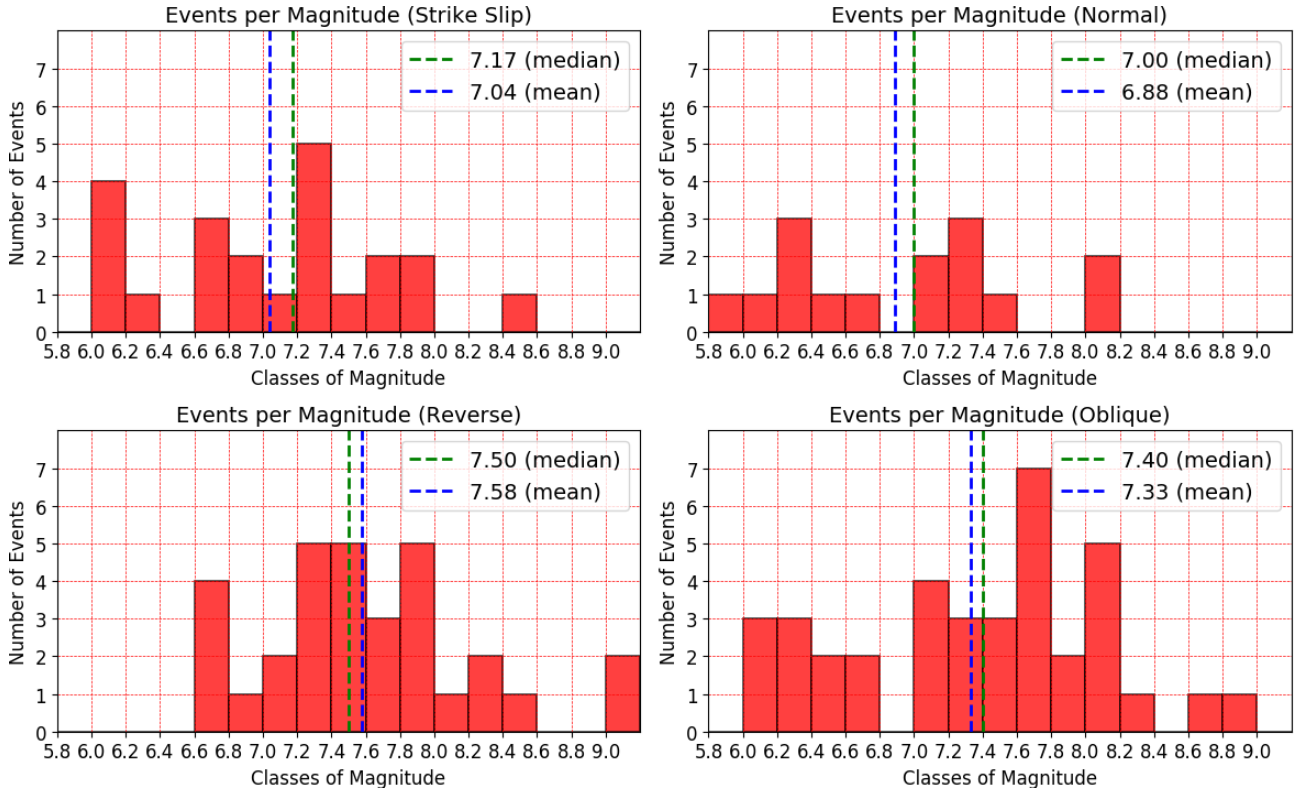

Figure 2.4 Magnitude distribution of FFMs (step = 0.2 Mw) distinguished per focal mechanism type. The dashed blue line represents the mean value and the green one the median value.

Figure 2.4 shows how many models fall in the defined slip-type categories: 15 normal, 22 strike-slip, 31 reverse and 37 oblique events. The categories are defined here according to the slip rake. An earthquake with slip rake falling in a 20°-interval centred on 90° (-90°) is considered a reverse (normal) event. Further, an earthquake with slip rake falling in a 20°-interval centred either on 0° or on 180° is considered a strike-slip event. The “oblique” category includes all the events that do not belong to the other categories.

Table 2.1 . Finite-Fault Models used in this study (ST: Slip-Type, MD: Maximum Displacement, AD: Average Displacement).

I. N.	Location	Date (m/d/y)	ST	Mw	Mo (Nm)	Depth (km)	Length (km)	Width (km)	Area (km^2)	MD (m)	AD (m)	References
1	Joshua Tree (Calif.)	04/23/1992	SS	6.2	2.70E+18	12.5	28	20	560	0.84	0.14	Bennet et al. (1995)
2	Landers (Calif.)	06/28/1992	SS	7.2	7.20E+19	7	76	15	1140	6.77	1.91	Zeng and Anderson (2000)
3	Tibet, Pumqu-Xainza	03/20/1993	O	6.3	2.97E+18	8.25	30	22	660	0.52	0.14	Wang et al. (2014)
4	Hokkaido-nansei-oki (Japan)	07/12/1993	R	7.6	2.85E+20	20	200	70	14000	4.36	0.62	Mendoza and Fukuyama (1996)
5	Northridge (Calif.)	01/17/1994	R	6.7	1.30E+19	17.5	17.5	23.5	412	4.14	0.76	Zeng and Anderson (2000)
6	Sanrikuki (Japan)	12/28/1994	R	7.7	3.99E+20	10	110	140	15400	4.03	0.71	Nagai et al. (2001)
7	Kobe (Japan)	01/16/1995	SS	6.8	1.76E+19	14	52	20	1040	2.75	0.50	Cho and Nakanishi (2000)
8	Colima (Mexico)	10/09/1995	R	8.0	9.67E+20	16.55	200	100	20000	4.77	1.18	Mendoza and Hartzell (1999)
9	Tibet, Pumqu-Xainza	07/03/1996	O	6.1	1.49E+18	8.25	25	18	450	0.45	0.10	Wang et al. (2014)
10	Hyuga-nada1 (Japan)	10/19/1996	R	6.8	1.84E+19	11.6	32.12	32.12	1032	2.92	0.54	Yagi et al. (1999)
11	Nazca Ridge (Peru)	11/12/1996	O	7.8	6.57E+20	21	180	120	21600	4.37	0.49	Salichon et al. (2003)
12	Hyuga-nada2 (Japan)	12/02/1996	R	6.7	1.19E+19	20.4	29.2	29.2	853	1.65	0.42	Yagi et al. (1999)
13	Kagoshimaen-hoku-seibu (Japan)	03/26/1997	SS	6.1	1.50E+18	7.6	15	10	150	0.87	0.34	Horikawa (2001)
14	Kagoshimaen-hoku-seibu (Japan)	05/13/1997	SS	6.0	1.16E+18	7.7	17	10	170	0.41	0.21	Horikawa (2001)
15	Antarctica (Strike-Slip Segment)	03/25/1998	SS	8.0	1.07E+21	12	290	35	10150	35.16	3.14	Antolik et al. (2000)
16	Antarctica	03/25/1998	O	7.8	4.85E+20	12	90	60	5400	21.10	2.83	Antolik et al. (2000)
17	Tibet, Pumqu-Xainza	08/25/1998	O	6.2	1.91E+18	8.25	38	23	874	0.20	0.07	Wang et al. (2014)
18	Iwate (Japan)	09/03/1998	O	6.3	3.20E+18	3	10	9	90	1.40	0.44	Nakahara et al. (2002)
19	Izmit (Turkey)	08/17/1999	SS	7.5	1.77E+20	16	160	28	4480	5.51	1.30	Cakir et al. (2004)
20	ChiChi (Taiwan)	09/20/1999	O	7.6	3.11E+20	7	78	39	3042	11.90	3.75	Sekiguchi et al. (2002)
21	Oaxaca (Mexico)	09/30/1999	N	7.5	1.82E+20	39.7	90	45	4050	2.46	0.64	Hernandez et al. (2001)
22	Hector Mine (Calif.)	10/16/1999	SS	7.1	5.82E+19	7.5	54	18	972	9.46	1.81	Salichon et al. (2004)
23	Duzce (Turkey)	11/12/1999	SS	6.7	1.28E+19	10	40.95	12.6	516	5.09	0.93	Birgoren et al. (2004)
24	Tottori (Japan)	10/06/2000	SS	6.7	1.40E+19	14.5	32	20	640	3.21	0.62	Semmane et al. (2005)a
25	Bhuj (India)	01/26/2001	R	7.4	1.33E+20	20	60	35	2100	12.44	1.51	Antolik and Dreger (2003)
26	Geiyo (Japan)	03/24/2001	R	6.7	1.19E+19	46.46	30	18	540	2.40	0.67	Kakehi (2004)

27	Denali (Alaska)	11/03/2002	SS	7.9	7.08E+20	6	292.5	18	5265	10.57	4.25	Asano et al. (2005)
28	Colima (Mexico)	01/22/2003	R	7.5	2.30E+20	20	70	85	5950	3.14	0.61	Yagi et al. (2004
29	Boumerdes (Algeria)	05/21/2003	R	7.3	8.40E+19	16	64	32	2048	3.52	1.24	Semmane et al. (2005)
30	Carlsberg Ridge	07/15/2003	SS	7.6	2.82E+20	11.32	320	36	11520	3.16	0.55	Wei (Caltech, Carlsberg 2003)
31	Miyagi-hokubu (Japan)	07/25/2003	SS	6.1	1.80E+18	6.5	18	10	180	1.03	0.31	Hikima and Koketsu (2004)
32	Tokachi-oki (Japan)	09/25/2003	R	8.2	2.36E+21	25	120	100	12000	7.06	3.11	Koketsu et al. (2004)
33	Bam, Iran	12/26/2003	O	6.5	7.30E+18	8	25	20	500	1.62	0.48	Poiata et al. (2012a)
34	Irian-Jaya, indonesia	02/07/2004	SS	7.2	7.08E+19	11.23	100	28	2800	3.37	1.03	Wei (Caltech, Irian-Jaya 2004)
35	Zhongba, Tibet	07/11/2004	N	6.2	2.24E+18	10	20	22.27	445	0.69	0.16	Elliott et al. (2010)
36	Parkfield (Calif.)	09/28/2004	O	6.1	1.36E+18	8.26	36.1	11.9	430	0.52	0.10	Custodio et al. (2005)
37	Niigata-Ken Chuetsu, Japan	10/23/2004	R	6.6	1.07E+19	10.6	28	18	504	3.08	0.67	Asano and Iwata (2009)
38	Sumatra	12/26/2004	R	9.1	6.50E+22	35	1480	224	331520	11.43	2.94	Ammon et al. (2005)
39	Fukuoka (Japan)	03/20/2005	SS	6.6	1.15E+19	14	26	18	468	2.67	0.68	Asano and Iwata (2006)
40	Sumatra	03/28/2005	O	8.7	1.17E+22	25.69	380	260	98800	12.50	2.56	Shao and Ji (UCSB, Sumatra 2005)
41	Zhongba, Tibet	04/07/2005	O	6.2	2.24E+18	5.98	28	18.72	524	1.29	0.19	Elliott et al. (2010)
42	Northern California	06/15/2005	SS	7.2	7.08E+19	9.003	102	35	3570	2.96	0.67	Shao and Ji (UCSB, Northern California 2005)
43	Honshu, Japan	08/16/2005	R	7.5	2.00E+20	34.49	96	56	5376	1.32	0.22	Shao and Ji (UCSB, Honshu 2005)
44	Kashmir, Pakistan	10/08/2005	O	7.6	2.82E+20	10.51	126	54	6804	6.37	1.75	Shao and Ji (UCSB,Kashmir 2005)
45	Kuril Islands	11/15/2006	R	8.3	3.16E+21	25.85	400	137.5	55000	8.93	1.69	Ji (UCSB, Kuril 2006)
46	Kuril Islands	01/13/2007	O	8.1	1.58E+21	18.15	200	35	7000	20.25	7.02	Ji (UCSB, Kuril 2007)
47	Noto Hanto, Japan	03/25/2007	O	6.7	1.57E+19	9.62	30	16	480	5.07	1.09	Asano and Iwata (2011)
48	Solomon islands	04/01/2007	R	8.1	1.58E+21	11.6	300	80	24000	3.73	1.47	Ji (UCSB, Solomon Islands 2007)
49	Pisco, peru	08/15/2007	O	8.0	1.12E+21	29.41	192	108	20736	8.21	1.63	Ji and Zeng (Peru 2007)
50	Niigata-ken Chuetsu-oki	08/17/2007	N	6.6	1.60E+19	8.9	33.25	29.75	990	2.58	0.32	Cirella et al. (2008)
51	Bengkulu, indonesia	09/12/2007	R	8.5	6.70E+21	21.23	400	250	100000	5.22	1.21	Gusman et al. (2010)
52	Tocopilla, Chile	11/14/2007	R	7.8	5.82E+20	36.95	375	200	75000	2.99	0.22	Ji (UCSB, Tocopilla 2007)
53	Gerze, Tibet	01/09/2008	O	6.4	4.47E+18	7.5	20	19.65	393	1.96	0.37	Elliott et al. (2010)
54	Gerze, Tibet	01/16/2008	N	5.9	7.94E+17	4	15	10	150	0.88	0.20	Elliott et al. (2010)
55	Simeulue, Indonesia	02/20/2008	R	7.4	1.41E+20	24.8	152	112	17024	1.08	0.15	Sladen (Caltech, Simeulue 2008)
56	Yutian, Tibet	03/20/2008	N	7.1	5.01E+19	4.104	54	19.05	1029	5.14	1.50	Elliott et al. (2010)
57	Wenchuan, China	05/12/2008	O	8.0	1.41E+21	16	320	60	19200	8.01	3.21	Yagi et al. (2012)
58	Iwate Miyagi Nairiku	06/13/2008	N	7.0	3.65E+19	6.5	42.66	17.38	741	6.36	1.82	Cultrera et al. (2013)
59	Zhongba, Tibet	08/25/2008	O	6.7	1.26E+19	7.626	30	30.4	912	1.54	0.25	Elliott et al. (2010)

60	Kermedac Islands, new Zealand	09/29/2008	R	7.0	3.55E+19	39.5	70	70	4900	0.98	0.14	Hayes (NEIC, New Zealand 2008)
61	Sulawesi, Indonesia	11/16/2008	R	7.3	1.00E+20	25.5	120	56	6720	2.33	0.45	Sladen (Caltech, Sulawesi 2008)
62	Papua	01/03/2009	O	7.6	2.82E+20	34.59	120	96	11520	4.96	0.59	Hayes (NEIC, Papua 2009)
63	L'Aquila, Italy	04/06/2009	N	6.3	3.13E+18	8.639	30	22	660	1.36	0.19	Gualandi et al. (2013)
64	Offshore Honduras	05/28/2009	O	7.3	1.00E+20	13.51	180	36	6480	3.09	0.58	Hayes and Ji (Offshore Honduras 2009)
65	Fiordland, New Zealand	07/15/2009	O	7.6	2.82E+20	24.14	160	96	15360	5.57	0.63	Hayes (NEIC, New Zealand 2009)
66	Java, Indonesia	07/17/2006	R	7.8	6.77E+20	15	250	140	35000	2.12	0.66	Yagi and Fukahata (2011)
67	Gulf of California	08/03/2009	SS	6.9	2.51E+19	9.163	108	20.8	2246	2.32	0.31	Hayes (NEIC, Gulf of California 2009)
68	Samoa	09/29/2009	N	8.0	1.12E+21	16.85	180	49.08	8834	14.92	3.33	Hayes (NEIC, Samoa 2009)
69	Padang, Indonesia	09/30/2009	O	7.6	2.82E+20	80	54	45	2430	5.60	1.78	Sladen (Caltech, Padang 2009)
70	Vanuatu	10/07/2009	O	7.6	2.82E+20	35	91	60	5460	2.93	0.87	Sladen (Caltech, Vanuatu 2009)
71	Haiti	01/12/2010	O	7.0	3.55E+19	11	45	22.5	1013	3.72	1.45	Sladen (Caltech, Haiti 2010)
72	Maule, Chile	02/27/2010	O	8.9	2.51E+22	37	600	187	112200	12.90	4.05	Shao et al. (UCSB, Maule 2010)
73	El Mayor-Cucapah, Mexico	04/04/2010	SS	7.4	1.20E+20	10	120	16	1920	9.25	1.89	Mendoza and Hartzell (2013)
74	Northern Sumatra	04/06/2010	R	7.8	5.62E+20	30.64	240	216	51840	3.17	0.22	Hayes (USGS, Northern Sumatra 2010)
75	Northern Sumatra	05/09/2010	R	7.2	7.08E+19	44.63	90	90	8100	1.13	0.18	Hayes (NEIC, Northern Sumatra 2010)
76	Darfield, South Island New Zealand	09/03/2010	O	7.0	3.80E+19	10.83	80	26	2080	3.51	0.60	Hayes (NEIC, Darfield 2010)
77	Sumatra	10/25/2010	R	7.7	3.98E+20	17.43	375	196	73500	1.20	0.12	Hayes (NEIC, Southern Sumatra 2010)
78	Bonin Islands	12/21/2010	O	7.4	1.41E+20	17.77	110	42	4620	3.71	0.53	Hayes (NEIC, Bonin Islands 2010)
79	Vanuatu	12/25/2010	N	7.3	1.00E+20	15.17	90	42	3780	2.82	0.44	Hayes (NEIC, Vanuatu 2010)
80	Pakistan	01/18/2011	N	7.2	7.08E+19	66.79	60	60	3600	3.49	0.33	Hayes (NEIC, Pakistan 2011)a
81	Offshore Honshu, Japan	03/09/2011	R	7.3	1.00E+20	20.72	126	126	15876	1.35	0.19	Hayes (NEIC, Offshore Honshu 2011)
82	Tohoku-Oki, japan	03/11/2011	R	9.1	5.50E+22	21	525	260	136500	48.00	9.55	Wei et al. (2012)
83	Kermadec Islands	07/06/2011	N	7.3	1.00E+20	19.04	216	72	15552	4.62	0.34	Hayes (USGS, Kermadec Islands 2011)
84	Vanuatu	08/20/2011	R	7.3	1.00E+20	31.42	102	90	9180	0.93	0.12	Hayes (NEIC, Vanuatu 2011)
85	Kermadec Islands	10/21/2011	O	7.4	1.41E+20	32.1	90	90	8100	3.19	0.28	Hayes (NEIC, Kermadec Islands 2011)
86	Van,Turkey	10/23/2011	R	7.1	5.01E+19	19	95	40	3800	4.65	0.50	Konca (2015)
87	Sumatra	01/10/2012	SS	7.2	7.08E+19	18.37	90	21	1890	6.71	1.29	Hayes (NEIC, Sumatra 2012)
88	Oaxaca, Mexico	03/20/2012	R	7.4	1.41E+20	19.74	126	108	13608	4.53	0.28	Hayes (NEIC, Oaxaca 2012)
89	Sumatra	04/11/2012	SS	8.6	8.90E+21	22	384	60	23040	34.00	8.72	Wei (Caltech, Sumatra 2012)
90	Offshore El Salvador	08/27/2012	O	7.3	1.00E+20	19.85	210	128	26880	1.06	0.09	Hayes (NEIC, Offshore El Salvador 2012)
91	East of Sulangan, Philippines	08/31/2012	O	7.6	2.72E+20	34.04	128	90	11520	3.14	0.42	Hayes (USGS, Philippines 2012)

92	Costa Rica	09/05/2012	R	7.6	2.54E+20	39.32	150	120	18000	3.05	0.29	Hayes (NEIC, Costa Rica 2012)
93	Masset,Canada	10/28/2012	R	7.8	7.00E+20	17	210	90	18900	3.16	0.60	Wei (Caltech, Masset 2012)
94	Santa Cruz islands	06/02/2013	O	8.1	1.54E+21	12.7	144	90	12960	12.70	2.86	Lay et al. (2013)
95	Scotia Sea	11/17/2013	SS	7.7	8.94E+20	10.72	392	50	19600	4.40	0.83	Hayes (USGS, Scotia Sea 2013)
96	Iquique, Chile	04/01/2014	O	8.1	1.58E+21	21.54	285	160	45600	4.21	0.67	Wei (Caltech, Iquique 2014)
97	Gorkha, Nepal	04/25/2015	O	7.9	9.09E+20	15	160	88	14080	7.53	2.48	Yagi and Okuwaki (2015)
98	Alaska	01/24/2016	O	7.1	6.22e+19	15	50	52.5	2625	3.02	0.35	Hayes (NEIC, Alaska South 2014)
99	Kumamoto, Japan	04/14/2016	SS	6.1	2.04e+18	15	14	13	182	1.16	0.36	Asano and Iwata (2016)
100	Kumamoto, Japan	04/15/2016	O	7.1	5.12e+19	15	56	20	1120	5.66	1.64	Yagi et al. (2016)
101	Amatrice	08/24/2016	N	6.2	2.6e+18	15	36	12	432	0.50	0.19	Pizzi et al. (2017)
102	Ussita	10/26/2016	N	6.1	1.76e+18	15	30	13	390	0.40	0.14	Pizzi et al. (2017)
103	Norcia	10/30/2016	N	6.5	7.1e+18	15	36	13	468	2.85	0.45	Pizzi et al. (2017)
104	Chiapas, Mexico	09/08/2017	N	8.1	1.88e+21	15	145	65	9425	18.64	5.39	Okuwaki and Yagi (2017)
105	Illapel, Chile	09/16/2015	O	8.3	3.33e+21	15	190	130	24700	10.43	3.65	Okuwaki et al. (2016)

3 Regression laws and hypocentre location

Analysing the set of the selected FFMs we have obtained a number of scaling relations linking the geometrical properties of the seismic fault with the magnitude. Such relations, often used by seismologists and engineers in seismic hazard and risk assessments, could be useful in the frame of tsunami hazard studies, and even in the common tsunami early warning practice, where typically from earthquake location and magnitude (available in a few minutes after the quake) one makes real-time assessment on the generated tsunami size. More precisely, in this section we analyse the scaling laws relating the magnitude (M_W) on one side, to some of the main fault properties, such as rupture length (*Length*), rupture width (*Width*), rupture area (*Area*), maximum displacement (*MD*) and average displacement (*AD*).

As earthquake source-scaling properties are found to depend on the seismotectonic regime and faulting style (Stirling et al., 2013), rupture models have been grouped according to the fault mechanism.

In addition, we have explored the position of the hypocentre with respect to the location of slip peak and the areas of maximum displacement, since this can be one more key to understand, roughly but quickly, how the rupture evolved on the fault. Knowing in a short time where the slip was more or less intense allows one to get a better characterization of the potential tsunami.

Indeed, this is not a secondary feature, because in the very near field of large earthquakes, ground motions are strongly dependent on "local" directivity effects, that is, on the relative position of the hypocentre with respect to regions of high slip (Mai, 2001; Guatteri et al., 2003).

These regions are known as asperities. The criterion to define an asperity is not unique. Somerville et al. (1999) define an asperity as a connected region whose average slip is 1.5 or more times larger than the average displacement over the entire fault. Alternatively, Mai et al. (2005) define asperities with respect to the maximum displacement (MD), in order to distinguish between events with rather smooth slip distributions and ruptures with complex slip distribution presenting locally very high slip values. Hence, they define

as large asperities the regions where slip takes values between 1/3 and 2/3 of MD; and very large asperities the regions where the slip takes values larger than 2/3 of MD.

The rupture nucleation point is of great importance in the definition of the finite-fault inverse problem, because it is very important in spontaneous dynamic rupture modelling. The starting region of an earthquake strongly influences whether the rupture will actually propagate spontaneously or not, for some given initial stress distribution. Indeed, knowing in advance the probable rupture nucleation point for a set of initial conditions will permit to better compute the dynamic rupture process, avoiding, or at least reducing, the inconvenient and frequent trouble of rupture calculations that abort before rupturing the desired earthquake size. In finite source inversions, the rupture nucleation point is determined by waveform modelling and, hence, may differ from the actual location owing to differences between the dynamic range and the frequency passbands of the accelerometers and of the typical high-gain seismic network stations (Mai et al. 2005).

3.1 Source scaling relationships

In the SCRMOD database every model provides the geographical coordinates of the epicentre, the hypocentre depth, the rupture width and length, the seismic moment, the magnitude. The spatial variations of the on-fault slip along the strike and down-dip directions are also provided. The slip values are assigned over a matrix of sub-faults, with assigned number of rows (*Nx*) and columns (*Nz*). The dimensions (*inDx,* i*nDz*) of the rectangle covering the entire fault are also given.

In Table 3.1 the regression formulas computed in this study for maximum slip (MD), average slip (AD), fault length (Length), fault width (Width) and fault area (Area) vs. moment magnitude (M_W) are listed, divided into 5 categories: strike-slip events (SS), reverse (R) events, normal (N) events, oblique events, all events (2nd column). As reported by Blaser et al., 2010, the scaling relations can differ significantly for different slip types. Differently from other studies, such as Thinggbaijam et. al., 2017, we do not distinguish between shallow crustal and subduction-interface events, but we group them together in the reverse-faulting category. The scaling relationships are log-linear in the form:

$$\text{Log } D = q + m \cdot M_W$$

The number of data points used in each case is shown in the 3rd column of Table 3.1. The best estimates for the coefficients *m* and *q* (calculated by least-squares method) are listed in the 4th and 5th columns, followed by the correlation coefficients *r*.

For the calculations and the regressions, the magnitude values are taken with two decimal figures. However, when taken individually as estimates of the moment magnitude, they are considered significant only to one decimal figure.

Table 3.1 Regressions between rupture dimensions (Length, Width), rupture area (Area), maximum displacement (MD), average displacement (AD), and moment magnitude (M_W); r is the correlation coefficient.

Equation	Slip Type	Number of events	q	m	r
$Log(L) = q + m M_W$	SS	22	-2.82	0.67	0.95
	N	15	-1.40	0.45	0.91
	R	31	-2.56	0.62	0.90
	Oblique	37	-1.70	0.50	0.90
	All	105	-2.05	0.55	0.91
$Log(W) = q + m M_W$	SS	22	-0.51	0.26	0.81
	N	15	-0.98	0.35	0.83
	R	31	-1.36	0.43	0.81
	Oblique	37	-1.42	0.42	0.86
	All	105	-1.55	0.43	0.82
$Log(A) = q + m M_W$	SS	22	-3.33	0.92	0.94
	N	15	-2.38	0.80	0.91
	R	31	-3.92	1.05	0.89
	Oblique	37	-3.12	0.92	0.91
	All	105	-3.61	0.99	0.91
$Log(MD) = q + m M_W$	SS	22	-3.90	0.64	0.89
	N	15	-4.01	0.64	0.89
	R	31	-2.11	0.35	0.62
	Oblique	37	-3.39	0.54	0.83
	All	105	-2.90	0.47	0.76

Equation	Slip Type	Number of events	q	m	r
$Log(AD) = q + m\ M_W$	SS	22	-3.94	0.55	0.86
	N	15	-4.46	0.61	0.84
	R	31	-3.24	0.40	0.57
	Oblique	37	-4.15	0.55	0.75
	All	105	-3.46	0.45	0.68

Let us consider the relationships between the fault dimensions and the magnitude. The length of the fault plane (see Figure 3.1) is highly correlated with the magnitude, and the highest value of the correlation coefficient is found for the strike-slip events. They present the largest length for a given magnitude, which is in agreement with the rupture mechanism. The same behaviour is reported by Thingbaijam et al., 2017, with a slope of *0.7*, really close to our $m = 0.67$.

The fault width-magnitude laws present instead a lower slope. The largest values of width are seen for the reverse events. Contrary to the length-magnitude laws, strike-slip events have slope lower than all other categories, confirming an easier rupture propagation along the strike direction for these events. In addition to that, as reported by Thingbaijam et al., 2017, strike-slip events on quasi-vertical faults are strongly affected by the finite width of the seismogenic layer. From Figure 3.2 it is possible to notice the dominant behaviour of the reverse-faulting events, and further that there are no strike-slip and normal-faulting events for which W exceeds the value of 10^2 *km*.

The fact that length L grows more rapidly with magnitude compared to W is highlighted also by Thingbaijam et al., 2017. They also confirm the different behaviour of W for the different focal mechanisms, stating that subduction-interface earthquakes have significantly higher W compared to other faulting regimes.

Indeed, strike-slip events are the only ones showing a correlation between the aspect ratio and the earthquake magnitude: this relationship results instead meaningless for the other rupture mechanisms (see Figure 3.3).

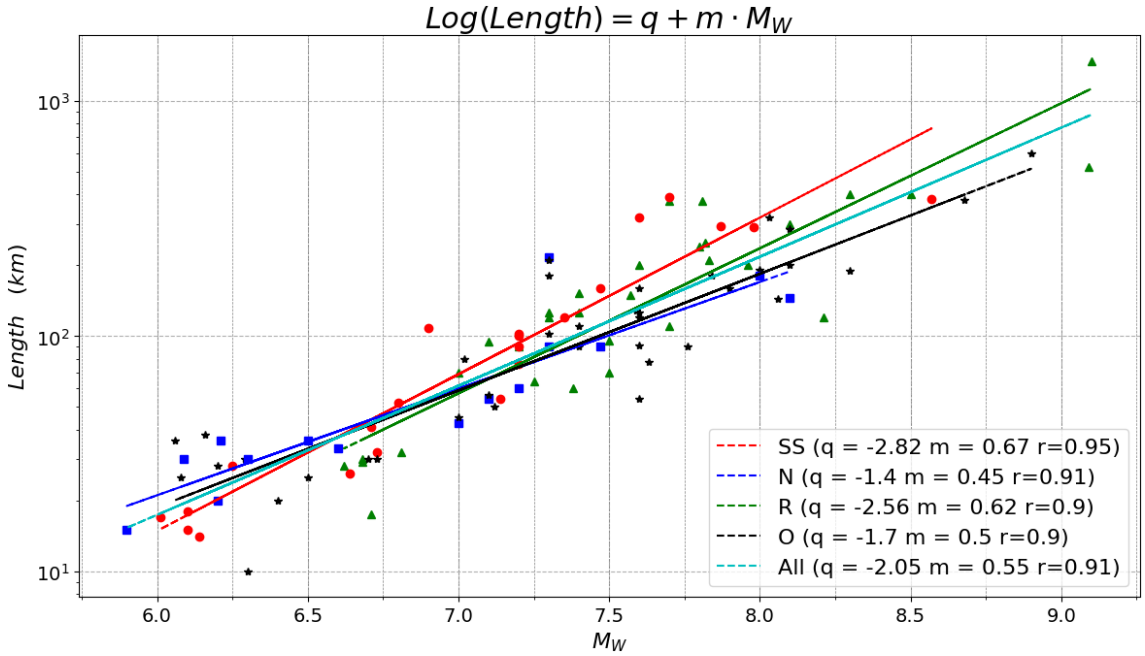

Figure 3.1 Regression law of rupture length vs. magnitude. The different colours and data symbols are related to different focal mechanisms: red circles represent the strike-slip events (SS), blue squares the normal events (N), green triangles the reverse events (R), black stars the oblique-type events (O).

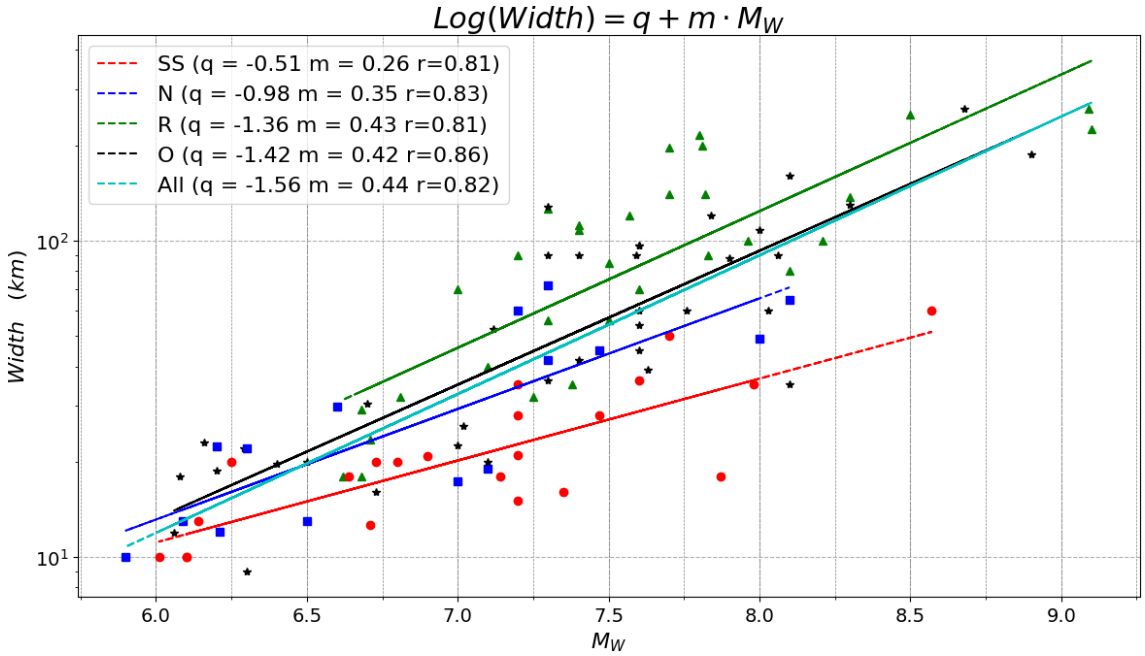

Figure 3.2 Regression law of rupture width vs. magnitude. See caption of Figure 3.1 for details.

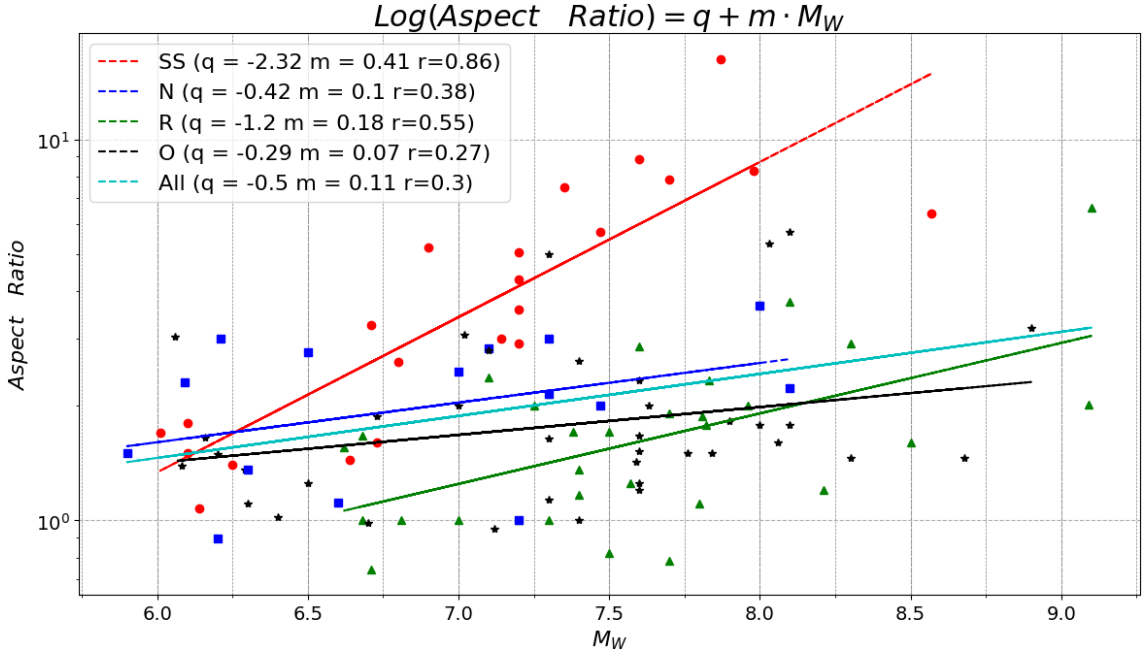

Figure 3.3 Regression law of aspect ratio vs. magnitude. See caption of Figure 3.1 for details.

The area of the fault (see Figure 3.4) also increases with the magnitude, with higher values and slopes for reverse earthquakes. Together with the length, this parameter is the one showing the best correlation with the magnitude.

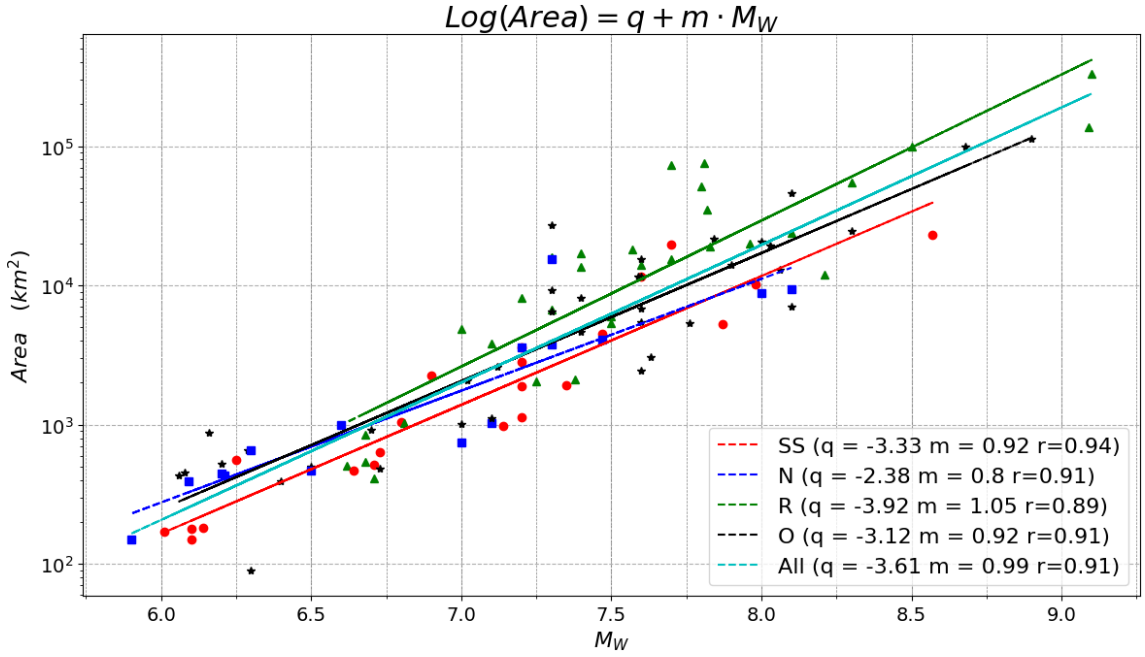

Figure 3.4 Regression law of fault area vs. magnitude. See caption of Figure 3.1 for details.

Considering the maximum displacement MD one observes (see Figure 3.5) that it increases with the magnitude, but the correlation coefficient is not that high. The strike-slip events, for which the correlation coefficient has the highest value (comparable with the one computed for normal events), generally exhibit a larger peak of slip, for a given magnitude, with respect to the other mechanisms. Reverse earthquakes, instead, present the lowest value of slip peak. The same consideration holds when one considers the average displacement AD as visible from Figure 3.6.

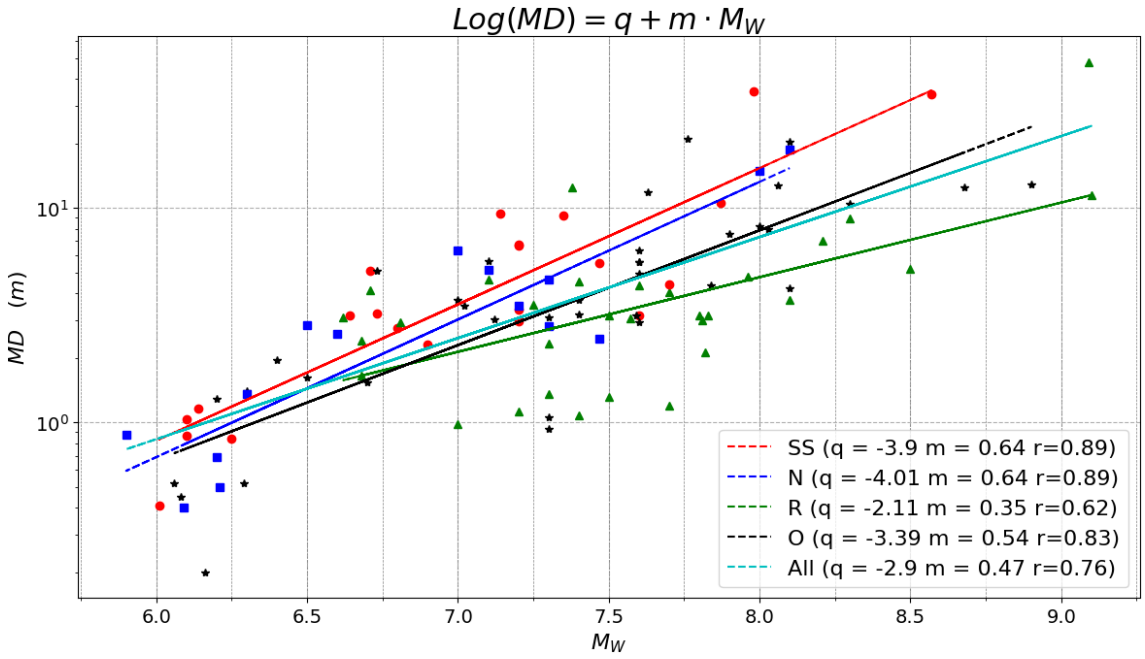

Figure 3.5 Regression law of maximum displacement vs. magnitude. See caption of Figure 3.1 for details.

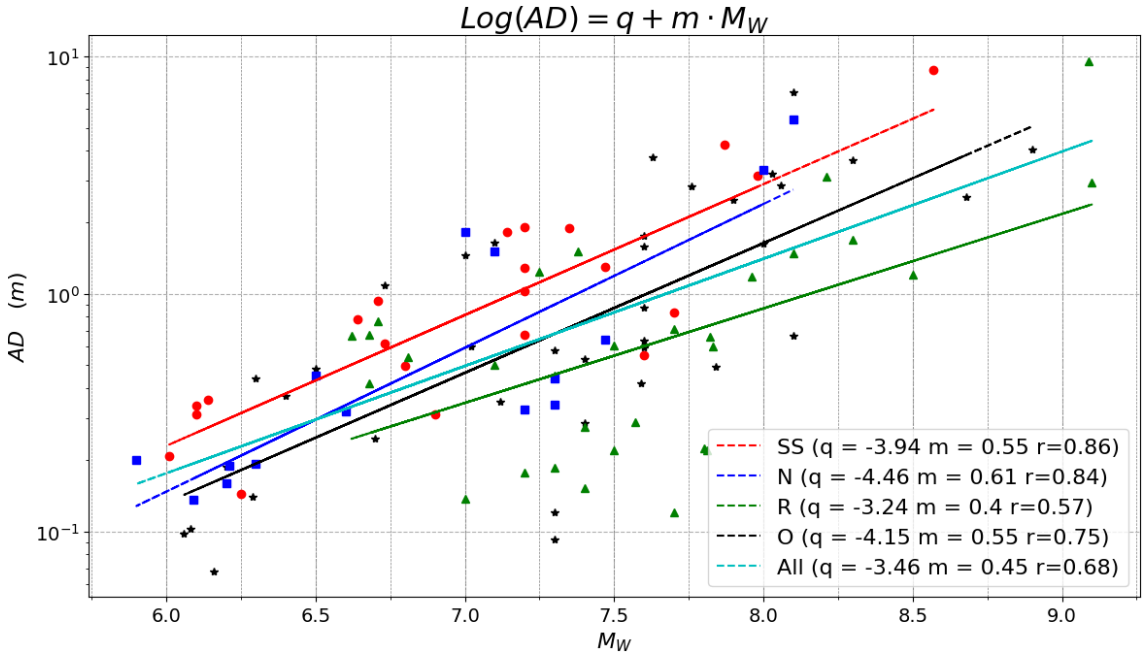

Figure 3.6 Regression law of average displacement vs. magnitude. See caption of Figure 3.1 for details.

Taking into account that the most widely used scaling relations are those derived by Wells and Coppersmith (1994), it is interesting to make some comparisons with their results. They considered a larger number of events, with magnitude also smaller than *5* and distinguished between surface and subsurface rupture length. The surface lengths are obtained thanks to ground-surface outcrops, whereas subsurface estimates derive generally from aftershocks analyses. They showed that relations differ for these two categories, with the latter being more appropriate for a comparison, since we worked considering subsurface lengths.

Table 3.2 Regression laws by Wells and Coppersmith (1994)

Equation	Slip Type	Number of events	a	b	r
Log(L) = a + b M_W	SS	93	-2.57	0.62	0.96
	R	50	-2.42	0.58	0.93
	N	24	-1.88	0.50	0.88
	All	167	-2.44	0.59	0.94

Equation	Slip Type	Number of events	a	b	r
$Log(W) = a + b\ M_W$	SS	87	-0.76	0.27	0.84
	R	43	-1.61	0.41	0.90
	N	23	-1.14	0.35	0.86
	All	153	-1.01	0.32	0.84
$Log(A) = a + b\ M_W$	SS	83	-3.42	0.90	0.96
	R	43	-3.99	0.98	0.94
	N	22	-2.87	0.82	0.92
	All	148	-3.49	0.91	0.95

Comparing Table 3.1 and Table 3.2 we notice that our laws are somewhat different from Wells and Coppersmith's, and generally, our relations lead to higher values of length and width for the same magnitude. An example is given in Table 3.3 where we show size estimates obtained by means of our and Wells and Coppersmith (1994) laws for a strong earthquake ($M_W = 7.5$), that is however on the lower side of the tsunamigenic shock category.

Table 3.3 Comparison on the length and width values predicted for an $M_W = 7.5$ earthquake

Calculated value	Slip Type	Wells and Coppersmith	This study
L (km, $M_W = 7.5$)	SS	120	160
	R	85	123
	N	74	94
	All	97	119
W (km, $M_W = 7.5$)	SS	18	28
	R	29	73
	N	31	44
	All	25	55

Notice that our regression laws that hold for earthquakes with $M_W \geq 6$ give larger rupture length for every slip type, especially for those characterized by strike-slip mechanism.

Figure 3.7 shows the relationship between seismic moment and magnitude of the earthquakes of our database, which is a subset of the SRCMOD database. The two are perfectly correlated, which suggests that in building SRCMOD the two quantities have been analytically derived one from the other according to the Hanks and Kanamori's law (1979). If one wants to link the source dimensions with the seismic moment, rather than the moment magnitude, one can simply proceed following the reasoning explained below. Calling *D* the generic source dimension one has:

$$\log_{10} D = a_d + b_d \cdot M_W$$

and accepting that

$$M_W = a_M + b_M \cdot \log_{10} M_0$$

where the coefficient values are those given by the Hanks and Kanamori's law (1979), one can conclude that:

$$\log_{10} D = (a_d + a_M b_d) + b_M b_d \log_{10} M_0$$

$$D \propto M_0^{b_d b_M} \implies M_0 \propto D^{\frac{1}{b_d b_M}}$$

Following this elementary algebraic strategy, we obtain from our regression laws:

$$L \propto {M_0}^{0.36} \implies M_0 \propto L^{2.75}$$
$$W \propto {M_0}^{0.29} \implies M_0 \propto W^{3.44}$$
$$A \propto {M_0}^{0.65} \implies M_0 \propto A^{1.53}$$
$$MD \propto {M_0}^{0.31} \implies M_0 \propto (MD)^{3.22}$$
$$AD \propto {M_0}^{0.30} \implies M_0 \propto (AD)^{3.37}$$

The relationships obtained here are in good agreement with those calculated by Mai and Beroza (2000), i.e.:

$$L \propto {M_0}^{0.39}$$
$$W \propto {M_0}^{0.32}$$
$$A \propto {M_0}^{0.72}$$
$$AD \propto M_0^{0.29}$$

We remark further that these relationships can be seen also as an indicative support of the principle of self-similarity. This latter states that any change in M_0 implies appropriate changes in *L*, *W*, and *AD* (Kanamori and Anderson, 1975), and if one assumes a constant fault aspect-ratio (*L/W*), the above laws take on the form:

$$L \propto {M_0}^{\frac{1}{3}} \quad W \propto {M_0}^{\frac{1}{3}} \quad A \propto {M_0}^{\frac{2}{3}} \quad AD \propto {M_0}^{\frac{1}{3}}$$

This scaling behaviour is associated with scale-invariant stress-drop (Thingbaijam et al., 2017). However, our relations do not fit perfectly the fault dimensions *L* and *W*, and hence highlight how the aspect ratio is not independent from magnitude (or seismic moment) changes. Indeed, in agreement with our results, several studies reported that *L* grows faster with increasing magnitude ($M_0 > 6$) compared to the growth of *W* (e.g., Mai and Beroza, 2000; Blaser et al., 2010; Leonard, 2010). Regarding particularly the aspect ratio, Blaser et al., 2010, underlined that it changes with magnitude, and moreover, it changes differently for different slip types.

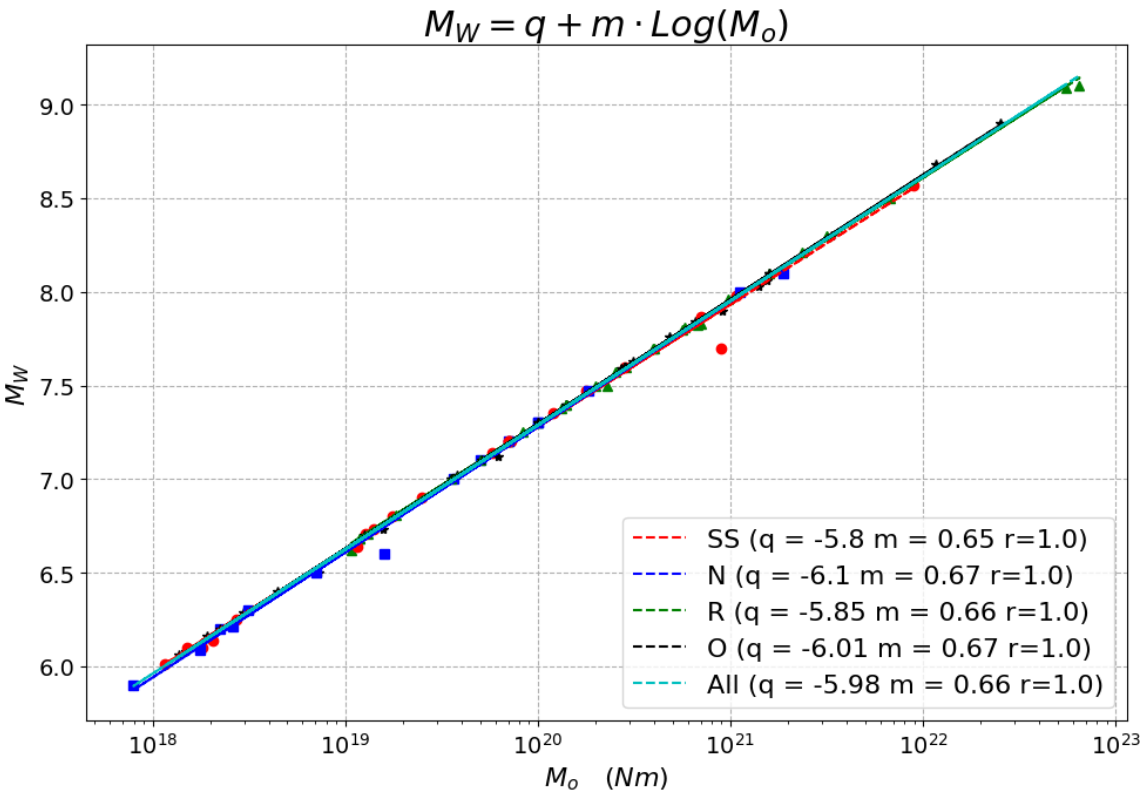

Figure 3.7 Regression law of magnitude (Mw) vs. seismic moment (Mo). See caption of Figure 3.1 for details.

3.2 Hypocentre and maximum slip

We have explored the relation between the positions of the hypocentre and of the slip peak and have summarised our results in the graphs of Figure 3.8.

The nucleation point is almost never coincident with the maximum displacement point. The distance between the two points tends to increase along with the magnitude, which suggests us to try to derive scaling laws of the same type of the ones obtained in the previous section. Figure 3.8 shows the correlation between the distance and the moment

magnitude, where correlation coefficients vary between 0.64 (for reverse-faulting events) and 0.8 (for strike-slip and oblique events).

Strike-slip earthquakes exhibit the largest distance between hypocentre and maximum displacement for a fixed magnitude and the highest slope coefficient. Normal-faulting earthquakes are characterized by the lowest slope. This behaviour follows what was seen for the length vs. the magnitude.

This immediate conclusion is in agreement with previous studies regarding the hypocentre position. Manighetti et al. (2005) found that earthquakes nucleate at a finite distance from the zone of maximum slip. They found that this distance falls on average between 20% and 30% of the total largest asperity length. The results of Mai et al. (2005) confirmed this finding. Distinguishing between large asperities ($1/3\ MD \leq D \leq 2/3$) and very large asperities ($D \geq 2/3\ MD$), they found that ruptures start close to large-asperities and have to encounter a very-large-slip asperity within the first half of the rupture distance to be able to grow larger than the nucleation zone. They justify this behaviour with an energy balance reasoning. An earthquake can only grow in size, potentially becoming a large earthquake, if the energy absorbed to create new crack surface (fracture energy) balances the available elastostatic energy and the energy radiated by seismic waves (Husseini et al. 1975; Husseini, 1977; Madariaga and Olsen, 2000; Favreau and Archuleta, 2003). According to this view, if the rupture nucleation point is located in regions of small stress drop (meaning in a low-slip area), and far from any point of significant stress drop, the fracture energy will soon be too large to maintain further crack propagation, and the rupture will stop prematurely (Mai et al., 2005).

Instead of the geometrical on-fault distance between the hypocentre and the MD position, one can consider the horizontal and the vertical distances (Figure 3.9, Figure 3.10). From the latter Figure one can see that the hypocentre is almost always deeper than point of slip peak. For only 24 cases over 105 the opposite is true. Contrary to the down-dip direction, there is no such a significant difference in the along-strike direction (Figure 3.9), underlining a symmetry in the horizontal direction. The down-dip distance is correlated also vs. the aspect ratio (Figure 3.11), and from this last relation it is possible to notice that most of the events for which the maximum slip is deeper than the hypocentre stay in the window marked by a low aspect ratio and are reverse-mechanism earthquakes. For strike slip events, the maximum displacement never lies deeper than the hypocentre.

All these observations confirm the conclusion by Mai et al. (2005) that hypocentres are not randomly located on a fault.

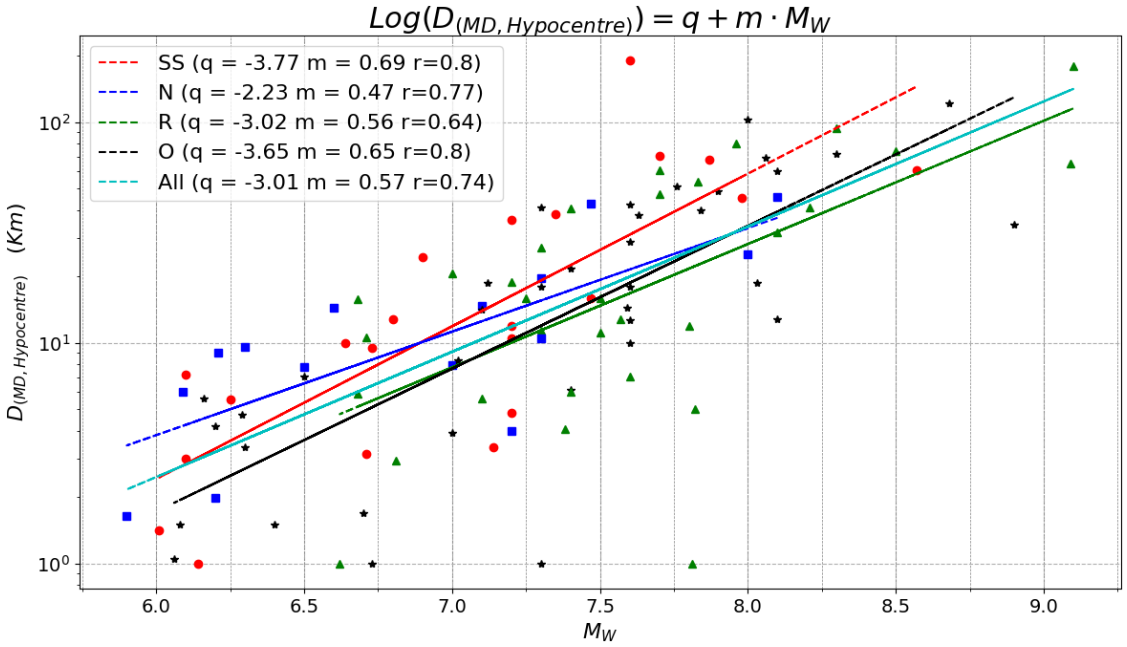

Figure 3.8 Regression law of the MD-Hypocentre distance vs. the magnitude Mw.

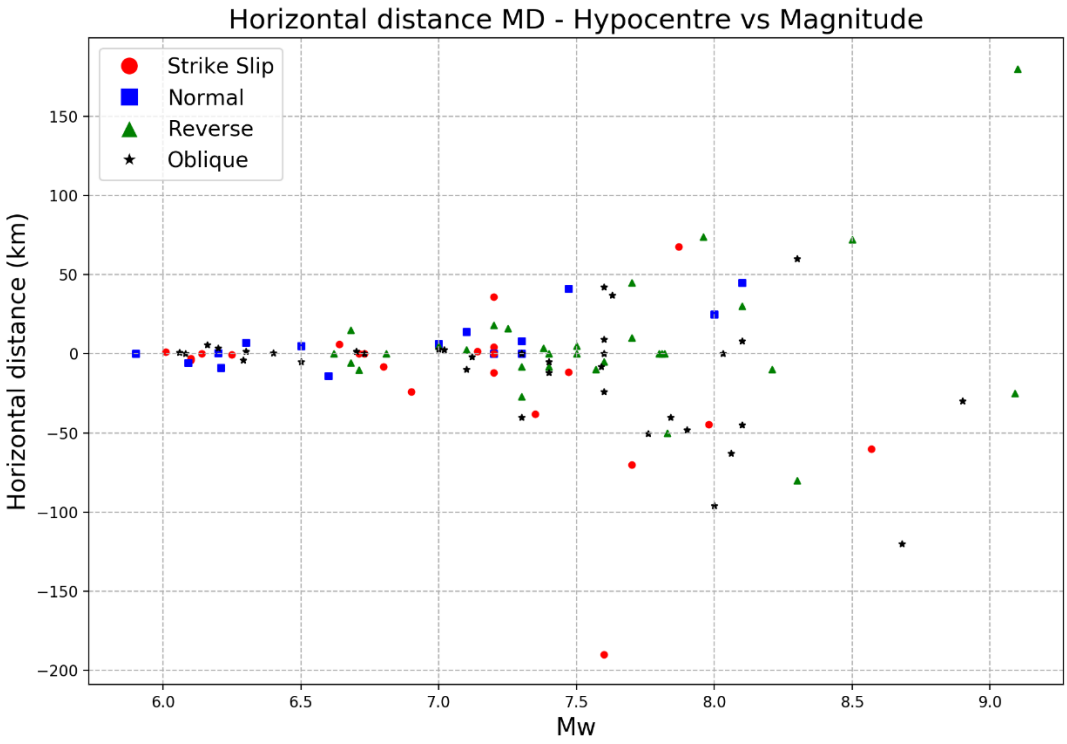

Figure 3.9 Along-strike distance between the point of maximum displacement (MD) and the hypocentre vs. the magnitude.

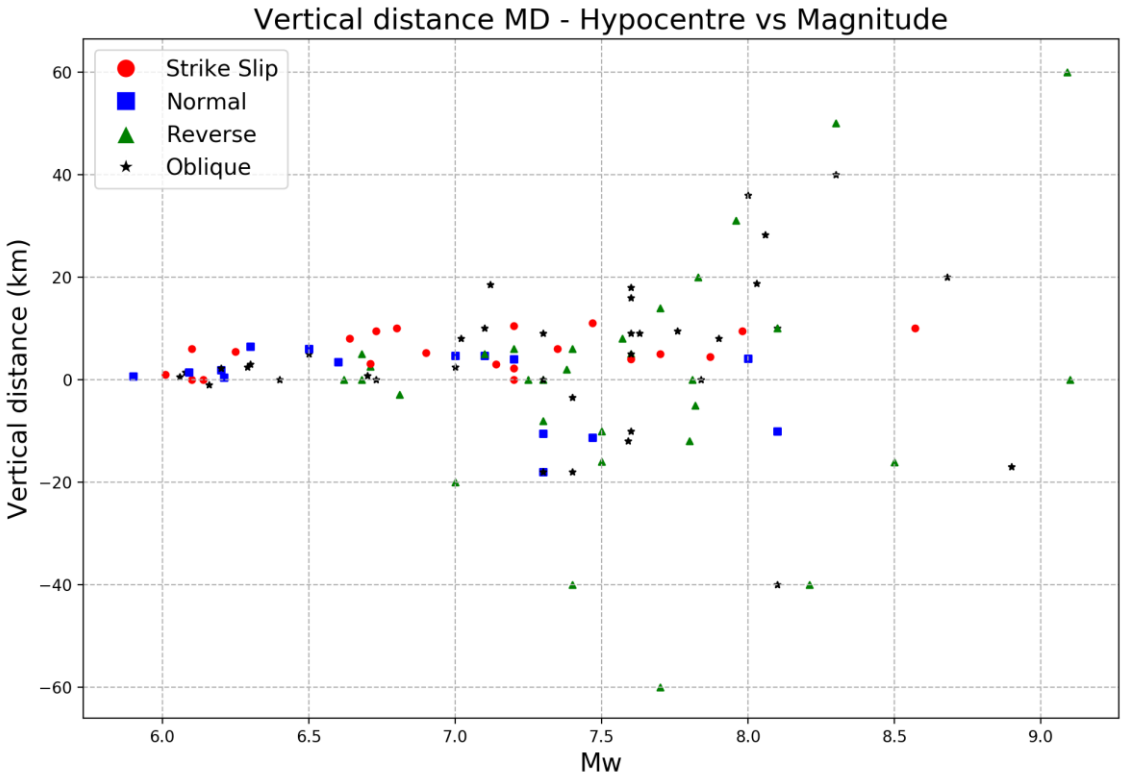

Figure 3.10 Along-dip between the point of maximum displacement (MD) and the hypocentre vs. the magnitude.

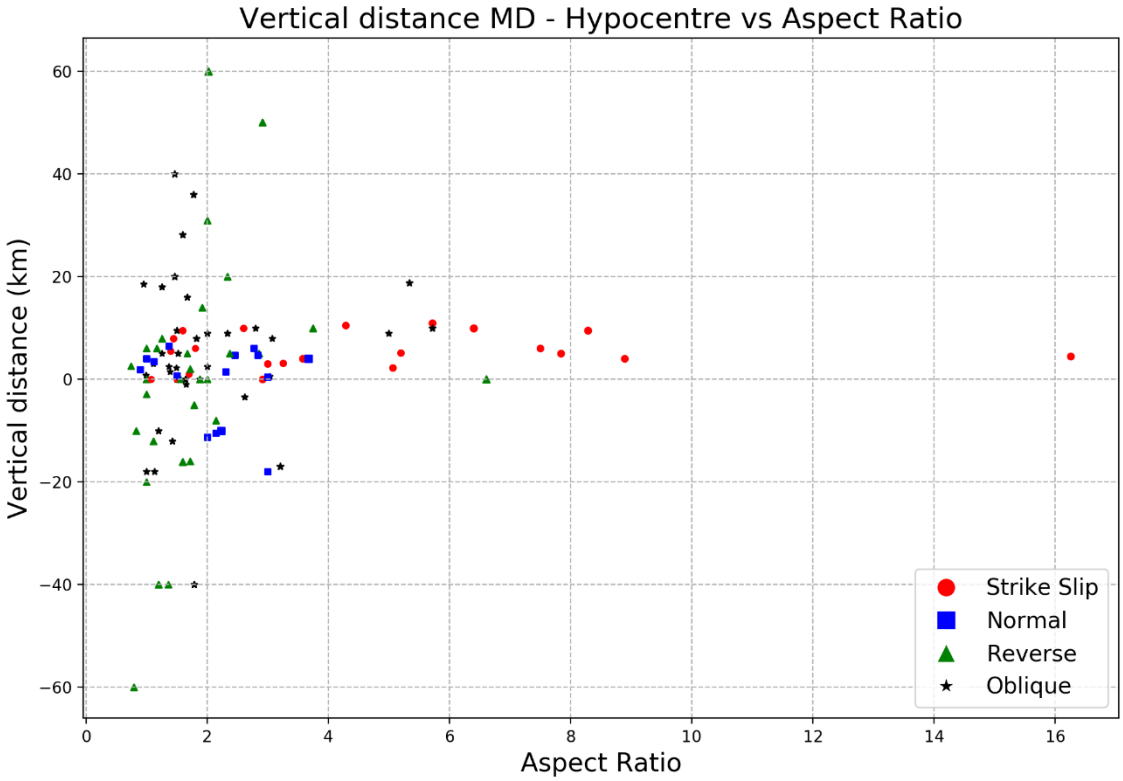

Figure 3.11 Along-dip distance between the point of maximum displacement (MD) and the hypocentre *vs. the aspect ratio.*

We have also considered the ratios d_{HM}/D, d_{HM}/L, X_{HM}/L, Y_{HM}/W, $|Y_{HM}/W|$ where:

- X_{HM} and Y_{HM} are the along-strike and along-dip distance of the maximum slip from the hypocentre;
- L and W are the length and the width of the fault plane;
- d_{HM} and D are:

$$d_{HM} = \sqrt{X_{HM}^2 + Y_{HM}^2} \; ; \qquad D = \sqrt{L^2 + W^2}$$

Tables 3.4 -*3.6* summarize some interesting results regarding these parameters, taking into account respectively the average values (Table 3.4), the 90th percentiles (Table 3.5) and the largest values (Table 3.6).

Table 3.4 shows that strike-slip earthquakes exhibit the largest average values for the normalised distance between hypocentre and maximum displacement, both in strike and dip directions. They are the only ones presenting the position of the maximum displacement always shallower than the hypocentre, and consequently the value Y_{HM}/W practically coincident with $|Y_{HM}/W|$. Normal and oblique events are those immediately

following. Reverse earthquakes are the only ones showing a negative value of Y_{HM}/W: for most of these then the maximum displacement position tends to be deeper than the hypocentre.

Strike-slip events are also those presenting the highest maximum ratio (about 0.6) for both strike and dip directions, indicating that the distance d_{HM} exceeds the half of the main rupture dimension along the related axis.

Table 3.4 Average values of X_{HM}/L, Y_{HM}/W, $|Y_{HM}/W|$, d_{HM}/D, d_{HM}/L for the different focal mechanisms

	X_{HM}/L	Y_{HM}/W	$\|Y_{HM}/W\|$	d_{HM}/D	d_{HM}/L
Strike-slip	0.153	0.251	0.251	0.194	0.208
Normal	0.174	0.063	0.184	0.188	0.215
Reverse	0.127	-0.004	0.143	0.149	0.190
Oblique	0.154	0.108	0.174	0.173	0.207
All	0.149	0.099	0.182	0.172	0.203

Table 3.5 Thresholds of X_{HM}/L, Y_{HM}/W, $|Y_{HM}/W|$, d_{HM}/D, d_{HM}/L under which one finds the 90% (the 90th percentile) of the values for the different focal mechanisms

	X_{HM}/L	Y_{HM}/W	$\|Y_{HM}/W\|$	d_{HM}/D	d_{HM}/L
Strike-slip	0.308	0.498	0.498	0.346	0.382
Normal	0.377	0.286	0.286	0.310	0.388
Reverse	0.250	0.231	0.357	0.342	0.401
Oblique	0.447	0.321	0.321	0.395	0.471
All	0.366	0.345	0.370	0.350	0.421

Table 3.6 Maximum values of X_{HM}/L, Y_{HM}/W, $|Y_{HM}/W|$, d_{HM}/D, d_{HM}/L for the different focal mechanisms

	X_{HM}/L	Y_{HM}/W	$\|Y_{HM}/W\|$	d_{HM}/D	d_{HM}/L
Strike-slip	0.594	0.600	0.600	0.590	0.594
Normal	0.458	0.462	0.462	0.425	0.475
Reverse	0.586	0.364	0.429	0.452	0.604
Oblique	0.561	0.500	0.500	0.475	0.571
All	0.594	0.600	0.600	0.590	0.604

The 90th percentile ratios given in *Table 3.5* show that strike-slip and reverse-faulting earthquakes present the highest values in the down-dip direction, with the distinction however that for the strike-slip events the slip peak is found always above the hypocentre, whereas in reverse ruptures the cases in which the maximum is below the hypocentre are preponderant. For the along-strike direction the behaviour is however different: strike-slip and reverse-faulting mechanisms present low ratios, suggesting that along the horizontal direction the maximum slip tends to remain closer to the hypocentre. Considering the ratio d_{HM}/D, the 90th percentile falls in the interval [0.3, 0.4] for all the different focal mechanisms, which can also be rephrased with the statement that hypocentres are located in the window [0.0, 0.4] for the 90% of all ruptures, or equivalently that in the 90% of the rupture models the nucleation point "encounters" the maximum slip within the first half of the main rupture distance D.
If one considers the maximum values for all the different focal mechanisms (see Table 3.6), d_{HM}/D stays in the window [0.4, 0.6].

As mentioned above, Mai et al. 2005 have investigated the relationship between the hypocentre and the closest asperity, having this issue important implications for the generation of near-source ground motion. In agreement with our results, they found no dependence between moment magnitude M_W and the normalized distances, where the normalisation value was the distance to the farthest point of the fault plane (R_{max} in their notation). However, they pointed out that the ranges of normalized distances to the closest large-slip and to the very-large-slip asperities are rather limited. They stated that, in the 48% of all the source models they investigated, the earthquake nucleated outside an asperity, 35% started within a large-slip asperity, and 16% within a very-large-slip asperity. From their analysis, hypocentres are located in the interval $0.0 < D \leq 0.4R_{max}$ for about 95% of all ruptures that did not start within a very-large-slip region, and never within a distance lower than $0.04R_{max}$. For none of the ruptures they found $R > 0.6\ R_{max}$. They concluded that ruptures generally nucleate close to large-slip asperities (about 0.15 R_{max}) and encounter a very-large-slip asperity within the first half of the total rupture distance. Similarly, Manighetti et al. (2005) found that ruptures start at a finite distance from the zone of maximum displacement, quantifying their statement in terms of

asperities dimension d_A. They asserted that distance between hypocentre and the zone of maximum slip averages between 20% and 30% of d_A.

3.2.1 Position of the hypocentre with respect to the fault plane

The analysis of the position of the nucleation point on the rupture plane could reveal features relevant especially when a fast determination of the fault geometry is important: the higher the magnitude, the larger is expected to be the variability of the hypocentre with respect to the fault plane, or equivalently of the fault placement with reference to the hypocentre. Mai et al. (2005) found that small earthquakes ($M_W < 6$) tend to rupture in the centre of the fault plane. Our database does not include events with moment magnitude lower than 6. We investigated the position of the nucleation point in the fault plane. The results are well represented in *Figure 3.12 Figure 3.13, Figure 3.14,* where the distance of the hypocentre from the fault plane centre ($d_{H\text{-}FC}$) and its along-strike and along-dip components are plotted in a logarithmic scale versus the moment magnitude. Even if the correlation coefficients are not that high (varying in the window [0.64, 0.8]), the increasing of the distance $d_{H\text{-}FC}$ with the magnitude is evident. The events most exhibiting this trend are the strike-slip events with a slope coefficient of 0.69 and a correlation coefficient of 0.8. On the other hand, normal-faulting events are the ones for which growth is more inhibited. This is a further indicator of the different evolution of the rupture for different fault mechanisms.

The regressions reported in *Figure 3.13* and *Figure 3.14* refer respectively to the along-strike $X_{H\text{-}FC}$ and along dip $Z_{H\text{-}FC}$ components of the distance $d_{H\text{-}FC}$, and are plotted to test the existence of a dominant direction. $X_{H\text{-}FC}$ presents almost the same value as above for the correlation coefficient (varying in [0.66, 0.7]) and a slope coefficient comprised between 0.6 and 0.7 with the exception of the reverse-faulting earthquakes characterized by a value of *0.76*.

Differently, the along-dip component $Z_{H\text{-}FC}$ shows a significant correlation coefficient only in the case of normal and reverse faulting events, with a slope lower than the one characterizing the along-strike direction.

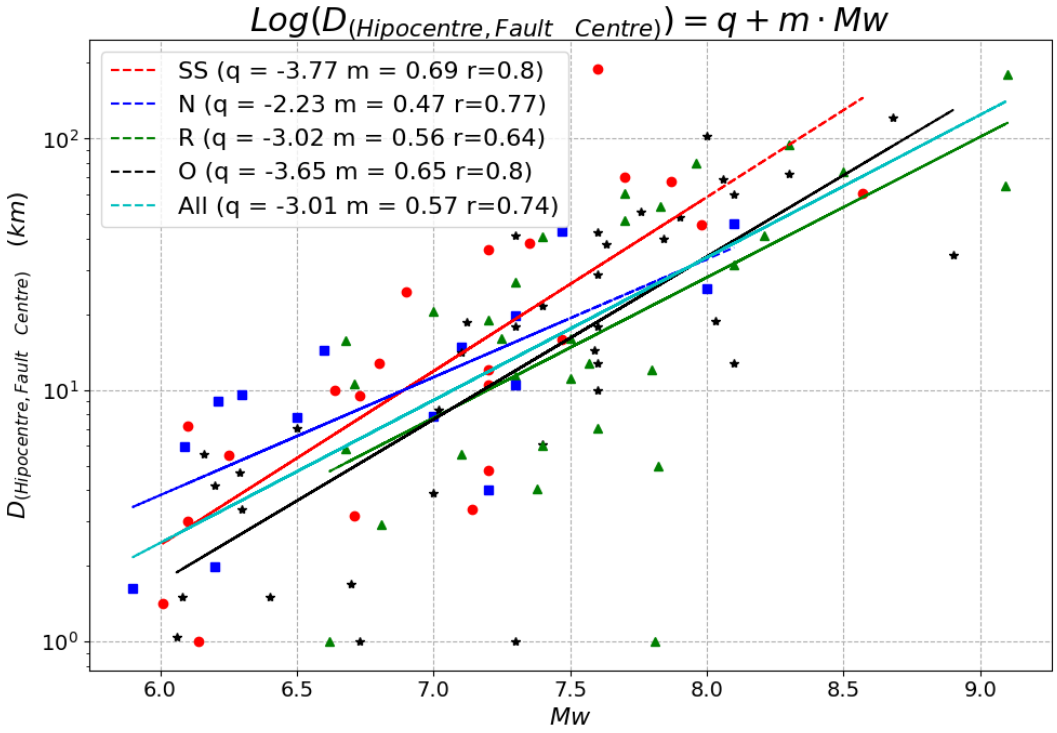

Figure 3.12: Regression of the distance between the hypocentre and the fault centre vs. magnitude.

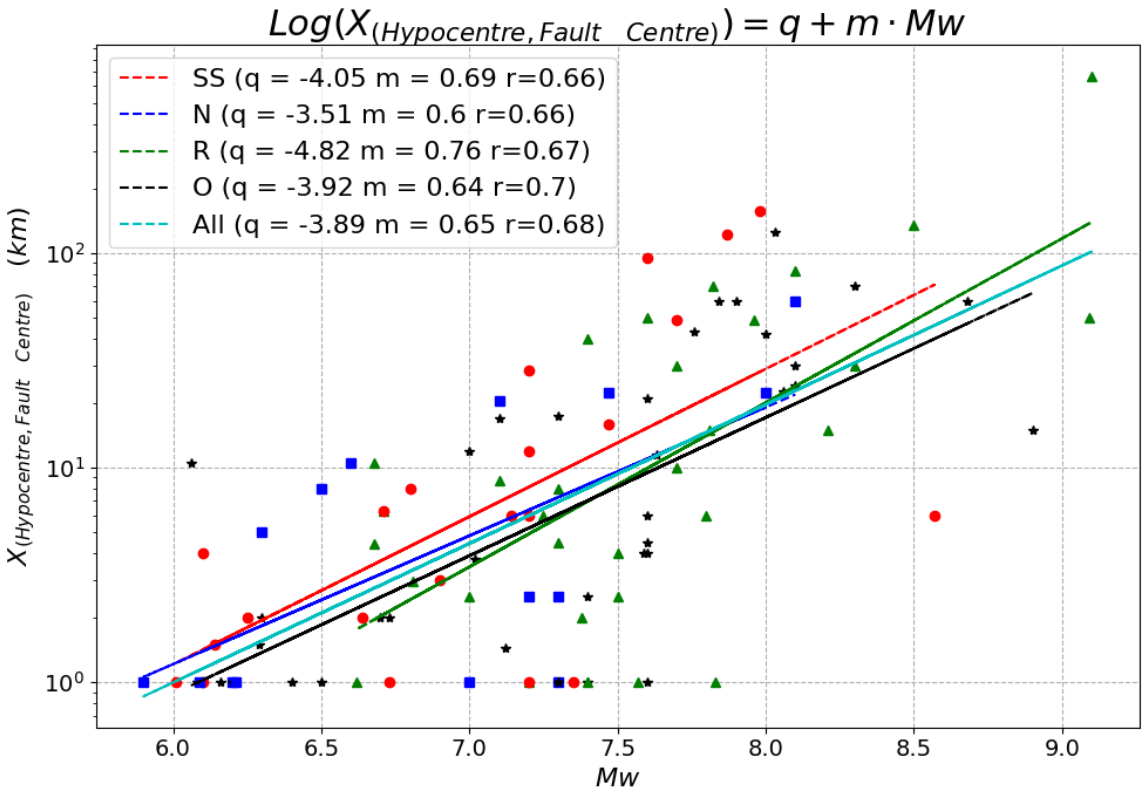

Figure 3.13 Regression of the along-strike distance between the hypocentre and the fault centre vs. magnitude.

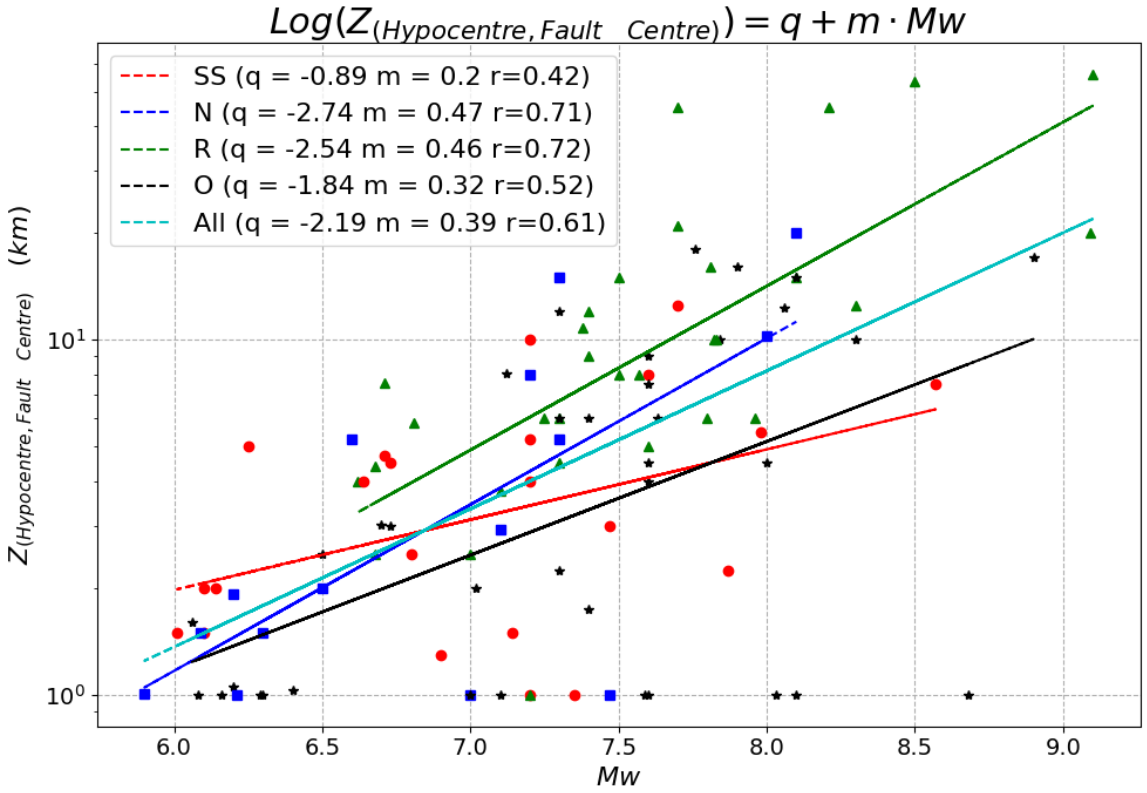

Figure 3.14 Regression of the along-dip distance between the hypocentre and the fault centre vs. magnitude.

3.2.2 Position of the slip peak

We analysed also the position of the maximum displacement on the fault plane. Looking at the regression law in Figure 3.15 it is noticeable that the peak of slip tends to move away from the fault centre as the magnitude increases. This distance growth seems to be larger, for the same magnitude values, for strike-slip earthquakes.

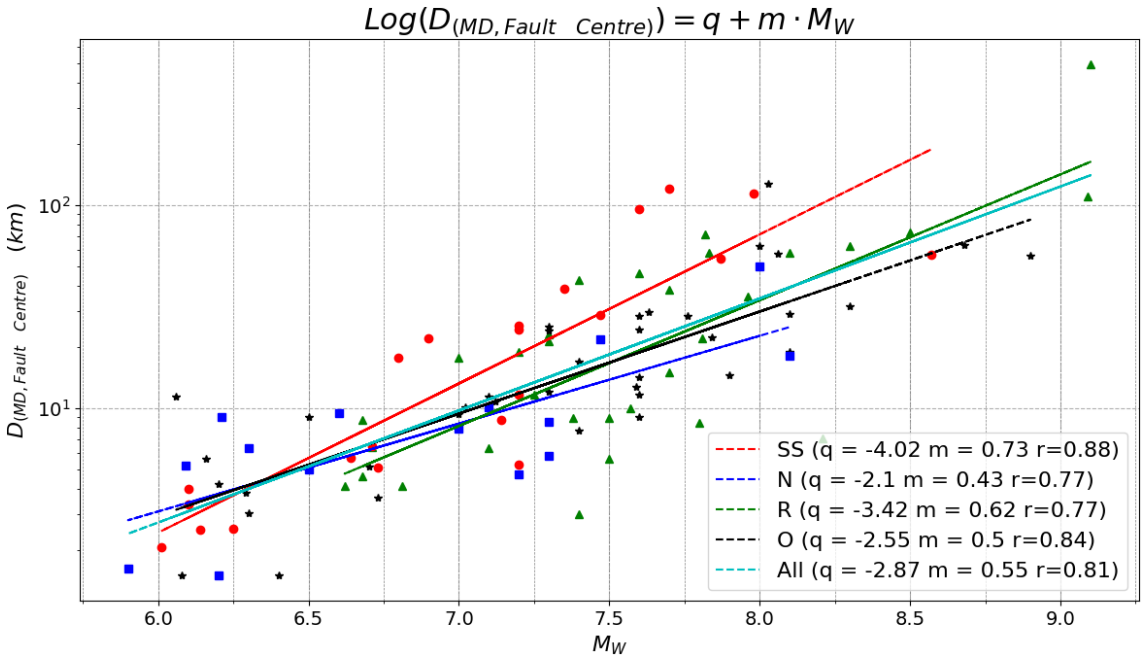

Figure 3.15 Regression of the distance between the point of maximum displacement and the fault centre vs. magnitude.

The question may arise whether this correlation is attributable to the horizontal or to the vertical components of the distance vector or to both. To provide an answer, the single components (along-strike and down-dip) have been plotted vs. magnitude. From Figure 3.16 and Figure 3.17 one can observe that the correlation between the vertical component and magnitude is higher than the one involving the horizontal component. Strike-slip events are those that, reasonably, exhibit the highest slope in the along-strike component plot: this agrees with the natural evolution of the rupture for such events.

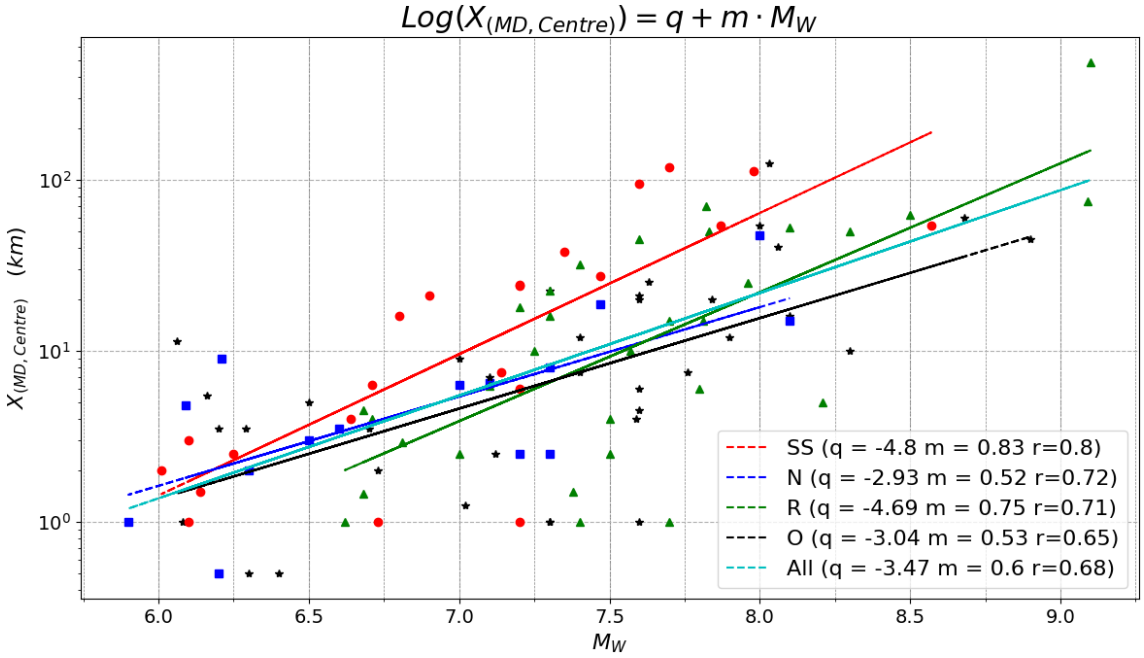

Figure 3.16 Regression of the along-strike component of the distance between the position of the maximum slip and the fault centre vs. magnitude.

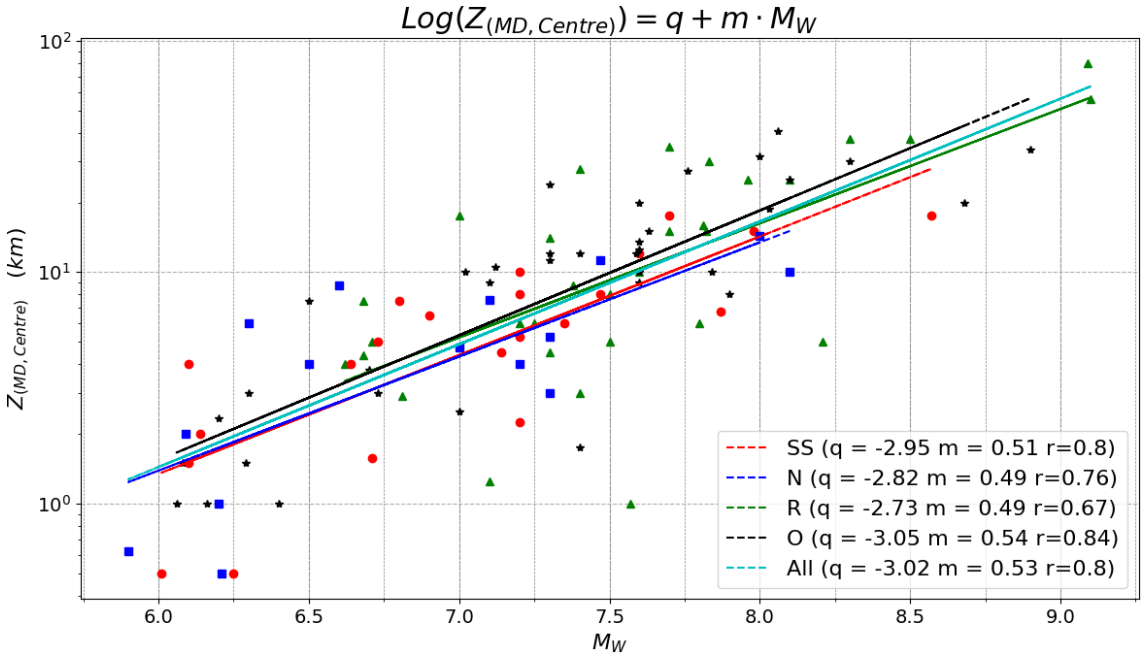

Figure 3.17 Regression of the along-dip component of the distance between the position of the maximum slip and the fault centre vs. magnitude.

3.3 Conclusions

From the above analysis one can draw some interesting considerations, synthesised in the following.

The quantities that best correlate with magnitude are the rupture length ($0.90 \leq r \leq 0.95$) and the rupture area ($0.89 \leq r \leq 0.94$), followed by the rupture width ($0.81 \leq r \leq 0.86$). The slip presents a weaker correlation than the fault dimensions with the magnitude. The strike-slip events are those with the best correlation coefficients among the derived scaling laws ($0.83 \leq r \leq 0.95$). They are also the only ones that show a good correlation between the aspect ratio and the magnitude (see Figure 3.3), with $r = 0.86$.

The reverse events are characterized by the highest values of rupture width and area for an assigned magnitude.

The hypocentre rarely coincides with the position of the peak of slip, and often lies at larger depth.

Regarding the position of the slip peak, this generally does not fall near the fault centre, with distance increasing with the earthquake magnitude, and more quickly for strike-slip events.

4 The 2D Gaussian Distribution

In this chapter we focus our attention on the distribution of the slip over the fault plane, which plays an important role in seismic studies. It is known that the propagation pattern and the frequency content of the seismic waves radiating from the source, as well as the permanent deformation field produced by the dislocation in the surrounding medium, heavily depend on slip heterogeneities. Moreover, keeping an eye on tsunami generation by earthquakes, it can be pointed out that the fault plane slip distribution affects remarkably, together with other factors, the geographical pattern of the maximum heights of the tsunami waves on the coast. Hence, setting up tools allowing to obtain reliable on-fault slip distributions from a few seismic parameters (such as the earthquake magnitude) would ensure improvements in the response of tsunami warning systems especially in areas where tsunamigenic sources are often located a few kilometres or a few tens of kilometres away from the coast. One could state that a better knowledge of the co-seismic slip distribution could improve the so-called "decision support chain" as part of the tsunami early warning systems.

Considering the four elements of systematic people-centred early warning systems (Basher et al., 2006; UN/ISDR, 2006):

- the risk knowledge,
- the monitoring and warning service,
- the dissemination and communication of the warnings,
- the response capability,

a better definition of the source could be of significant importance in the first two elements of the system.

As a matter of fact, a relevant problem for the tsunami warning centres in all cases where the tsunami leading time is short is that the decision whether issuing or not an alert after a potentially tsunamigenic earthquake, has to be made in real-time operations with a lack of information on the earthquake source. The focal mechanism is one of the main discriminants concerning the earthquake's ability to generate tsunami waves. Nowadays, the focal mechanism information is available relatively late (namely after some tens of minutes), being based on the inversion of seismic data, and even later (from hours to days) one gets a full model (FFM) of the slip distribution.

Getting a fast and acceptable estimate of the slip distribution based on earthquake magnitude could improve the performance of the tsunami warning centres procedures, especially for what concerns the near field region.

4.1 Identification of asperities

In this study, as a starting point, we have restricted our attention to those events that exhibit a clear single asperity. Indeed, as reported by Manighetti et al. (2005), typically slip distributions are dominated by only one zone of large slip, whose rupture accounts for most of the moment release.

As already outlined in the previous chapter, there are different definitions of an asperity. Here, we privilege the one adopted by Mai et al. (2005), who consider high-level asperity/asperities as region/regions where the slip takes on values larger than 2/3 of the peak value, and low-level asperity/asperities as region/regions where the slip is comprised between 1/3 and 2/3. Low-level asperities could be either connected to the high-level ones, in which case they form a peripheral extension, or disconnected, therefore forming a separate large slip area.

In this study we concentrate on high-level asperities only and the procedure used to identify the number of such asperities on a given rupture plane is a clustering method based on the so-called mean shift algorithm (Comaniciu and Meer, 2002) implemented in a Python framework.

More specifically, we have considered the two-dimensional space represented by the centres of the subfaults on the rupture plane. Over this domain, the data points selected for the cluster recognition are those exceeding a slip value threshold:

$$x: \qquad u(x) \geq \frac{2}{3} u_{max} \tag{4.1}$$

where x denotes the generic point on the fault plane. Let's call with *S* the set of subfaults fulfilling the above inequality. The mean shift method is an iterative procedure that starts with an initial estimate x_0 and with its neighbourhood $N(x_0)$ defined through a kernel function *K(x)* and that computes the density as well as the weighted mean (or centroid) $m(x)$ of the points within $N(x_0)$ according to the expression:

$$m(x) = \frac{\sum_{x_i \in N(x_0)} K(x_i - x_0) x_i}{\sum_{x_i \in N(x_0)} K(x_i - x_0)} \tag{4.2}$$

The algorithm involves shifting this kernel iteratively to a higher density region until convergence. Every shift is defined by the mean shift vector, that always points toward the direction of the maximum increase in the density. At every iteration the kernel is shifted to the centroid or the mean of the points within it. At convergence, there will be no direction at which a shift can accommodate more points inside the kernel. Practically, it is a centroid based algorithm, which works by updating candidates for centroids to be the mean of the points within a given region. At the end, one finds a candidate cluster for each selected starting point x_0. These candidates have to be filtered in a post-processing stage to eliminate near-duplicates to form the final set of centroids.

In our case, we have selected a procedure using a flat kernel involving a fraction of the characteristic rupture distance $\sqrt{(L^2 + W^2)}$, where L and W are the length and width of the fault respectively, and have taken as x_0 all the subfaults belonging to S. The final stage has recognised that the obtained clusters are mostly duplicates of one or two independent clusters or asperities. Figure 4.1 illustrates the initial set S and the final asperity for the FFM of the 16 September 2015 Illapel Chile earthquake (I.N. 105 of Table 2.1) obtained by applying the mean shift algorithm inclusive of the final filtering selection.

The application of the clustering algorithm to the dataset of 105 earthquakes recognises that 72 events have a single asperity slip distribution. In the following we will restrict our analysis to this subset. The list of their identification numbers is given in the first column of Table 4.1 later on in this chapter.

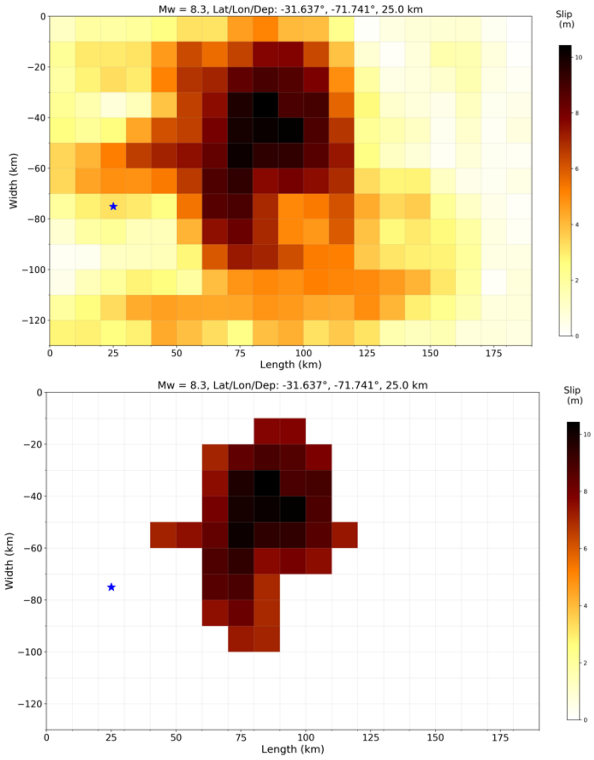

Figure 4.1 Example of filtered fault plane resulting from the mean shift clustering procedure. Above, the original SRCMOD FFM for the 16 September 2015 Illapel Chile earthquake (by Okuwaki et al., 2016). Below, the asperity identified by the method.

Parallel to the Gaussian modelling of single asperity events based on using the definition of asperities mentioned above, we have also investigated the relationship between the slip and area that characterise the asperities, following the work by Lee et al. (2016). The scaling of heterogeneous slips with the fault surface can provide a basis for ground motion simulation for earthquake scenarios, particularly in the near-fault region and is relevant also for tsunami generation and hazard assessment. We observe however that this topic does not change the general strategy of this thesis and is treated only for the sake of completeness. Details are reported in Appendix E.

4.2 The 2D Gaussian Distribution

Focusing on the new subset of events characterized by a single asperity, we can fit the on-fault slip through a distribution with a bi-dimensional Gaussian shape, hereafter denoted with 2D GD. The 2D GD, centred on the FFM subfault with maximum slip, has an elliptical distribution field, depending on four parameters, namely the standard deviations (σ_1, σ_2) along two perpendicular axes over the fault plane, the angle θ between

the along-strike axis and the major axis (σ_1), and the amplitude factor characterising the peak value (u_{max}) of the distribution (see Appendix A for graphic examples). In such a field the magnitude of the slip at the generic point (x, y) of the fault plane has the expression given below:

$$u(x,y) = u_{max} \cdot e^{-[a(x-x_0)^2+2b(x-x_0)(y-y_0)+c(y-y_0)^2]} \quad (4.3)$$

where u_{max} is the magnitude of the slip peak, and (x_0, y_0) are the coordinates of the peak position. The coefficients *a*, *b* and *c* are given by the formulas:

$$a = \frac{\cos^2\theta}{2\sigma_1^2} + \frac{\sin^2\theta}{2\sigma_2^2}\ ; \qquad b = \frac{\sin 2\theta}{4\sigma_1^2} - \frac{\sin 2\theta}{4\sigma_2^2}; \qquad c = \frac{\sin^2\theta}{2\sigma_1^2} + \frac{\cos^2\theta}{2\sigma_2^2}$$

The optimal parameters have been determined by a least-squares fitting procedure applied over the parameter space scanned at regular steps.

Different optimizations have been conducted, varying the number of free parameters or their variability domain. The tested configurations are:

1) σ_1, σ_2, θ, u_{max} free parameters;
2) σ_1, σ_2, θ free parameters (u_{max} = MD);
3) σ_1, σ_2, u_{max} free parameters ($\theta = 0$);
4) σ_1, σ_2, free parameters ($u_{max} = MD$, $\theta = 0$)
5) σ_1, σ_2, free parameters ($\theta = 0$, u_{max} is fixed to the value obtained from the regression law)

In the following we will denote with 2D GD*i* (*i*=1,2,…,5) the distributions obtained through the above optimization procedures.

4.3 Optimal 2D GD for tsunamigenic earthquakes

Fitting least-squares procedures can be targeted in different ways depending on the objective of fitting. In this study we have special interest for tsunamigenic earthquakes. We point out that for earthquakes occurring off-shore or near to the coast and having the potential to be tsunamigenic, what matters is the seafloor vertical-displacement field pattern since it constitutes the real tsunami hydrodynamic source. As a consequence, the free surface displacement field induced by a prescribed slip distribution on the fault at depth will be used as the natural metric to compare different slip patterns. More precisely, for any given earthquake, we compare the co-seismic vertical-component displacements

corresponding to the SRCMOD FFM, taken as the reference case, with the same displacement fields computed from the following alternative slip distributions:

- a homogeneous fault model;
- a "Smooth Closure Condition" distribution (hereafter SCC), in which the heterogeneity depends only on depth; it was originally introduced by Freund and Barnett (1976) (see e.g. Geist and Dmowska, 1999 for details, see also Appendix B);
- the best-fitting 2D GDs, i.e. 2D GD*i* (*i*=1,2,…5).

The choice of the first type of models is dictated by the observation that it is a widely used option in many studies dealing, for example, with the deformation fields induced by earthquakes, with earthquake hazard assessment, with the simulation of earthquake-induced tsunamis, with tsunami hazard assessment and tsunami early warning, while the second option is sometimes taken into account in the literature (see e.g. Geist and Dmowska, 1999; Tinti et al., 2005; Gutscher et al. 2006; Babeyko et al., 2010; Tonini et al., 2011, and many others).

For a given FFM, the surface displacement field, and in particular its vertical component, is computed here by means of the analytical formulas by Okada (1992), in which a rectangular fault is buried in a perfectly elastic, homogeneous and isotropic elastic half-space. A heterogeneous slip distribution is obtained by linearly superposing all contributions coming from the homogeneous-slip sub-faults.

For each of the 72 events of the single-asperity database, we started from the SRCMOD FFM reference case (hereafter FFM_{REF}). We adopted the same Cartesian reference frame and the same sub-fault tessellation of the fault plane as proposed in FFM_{REF}. More precisely, the basic geometric and focal properties of the fault (total and subfaults' length, total and subfaults' width, average strike, average dip, average rake, subfault positions) have been taken from FFM_{REF} and used for all other cases. The only varying parameter is the slip amount on each subfault, computed as follows:

- for FFM_{REF}, it is simply retrieved from the database;
- for the homogeneous slip distribution case, it is the average slip computed over FFM_{REF} and assigned to all subfaults;
- for the SSC case, see Appendix A;

- for the five 2D GDs, it is the slip matrix obtained by means of the least-squares algorithm.

The vertical surface displacements were computed over a flat domain corresponding to the altitude z = 0. The computational domain is centred in the mid-point of the surface projection of the FFM_{REF} fault plane, and its extension is twice the total length of the fault in both the strike and normal directions. The domain has been discretized with a grid of 100 x 100 nodes. As the domain extension varies for the different considered events, the same happens also for the grid steps.

4.4 Misfit

For each of the comparison mentioned above, the goodness of the fitting to the reference SRCMOD case is measured by computing the following misfit value:

$$msf = \sqrt{\frac{\sum_{i=1}^{n}(U_z^i - u_z^i)^2}{\sum_{i=1}^{n}(U_z^i)^2}} \tag{4.4}$$

where U_z is the free-surface vertical displacement of the reference field and u_z is the corresponding value of one of the considered fields, and the sum is extended over all nodes of the computational domain.

The misfit values resulting from the different comparisons are reported in Table 4.1, where they are calculated for all the single-asperity events identified through the mean shift clustering procedure.

Table 4.2 summarizes the minimum, maximum and average values for all the slip distributions. For all the 2D GD distributions the mean value of the misfit is lower than both the Uniform and SCC mean values. Among all the Gaussian optimizations, the best behaviour is given by the 2D GD_1 (mean misfit = 0.350) and by the 2D GD_3 distributions (mean misfit = 0.352).

Table 4.1 Misfit between the vertical-displacement fields calculated for the reference FFM events and the corresponding fields for each of the tested slip distributions. The total number of events in the Table is 72. Identification numbers in the first column (I.N.) are the ones given in Table 2.1.

I.N.	Uniform	SCC	2D GD_1	2D GD_2	2D GD_3	2D GD_4	2D GD_5
1	0.770	0.671	0.181	0.183	0.184	0.171	0.277
3	0.500	0.431	0.356	0.388	0.344	0.374	0.495
4	0.603	0.751	0.659	0.760	0.675	0.760	0.729

6	0.855	0.623	0.380	0.381	0.436	0.429	0.500
7	0.821	0.802	0.492	0.419	0.419	0.419	0.428
9	0.469	0.371	0.260	0.275	0.223	0.248	0.308
10	0.668	0.658	0.424	0.525	0.422	0.532	0.462
11	0.621	0.463	0.445	0.865	0.445	0.865	1.636
12	0.550	0.537	0.422	0.679	0.409	0.679	0.700
13	0.397	0.344	0.142	0.178	0.142	0.178	0.243
15	0.832	0.819	0.851	0.877	0.767	0.776	0.765
16	0.835	0.748	0.381	0.470	0.383	0.490	0.610
17	0.420	0.328	0.213	0.229	0.288	0.333	0.896
18	0.555	0.503	0.195	0.209	0.201	0.209	0.237
20	0.485	0.525	0.551	0.734	0.551	0.688	0.600
22	0.791	0.789	0.508	0.509	0.508	0.509	0.515
25	0.446	0.551	0.641	0.659	0.641	0.659	0.598
30	0.771	0.787	0.421	0.421	0.489	0.489	0.760
31	0.424	0.482	0.296	0.533	0.296	0.518	0.494
32	0.184	0.342	0.141	0.170	0.141	0.170	0.139
33	0.605	0.428	0.169	0.267	0.169	0.267	0.173
35	0.507	0.324	0.163	0.223	0.146	0.128	0.145
37	0.315	0.292	0.286	0.593	0.286	0.593	0.237
40	0.908	0.835	0.692	0.707	0.692	0.738	0.870
41	0.696	0.549	0.448	0.449	0.455	0.438	0.444
42	0.959	0.953	0.352	0.342	0.352	0.342	1.001
43	0.236	0.207	0.326	0.326	0.244	0.218	0.378
44	0.730	0.702	0.304	0.304	0.304	0.304	0.320
45	0.828	0.782	0.360	0.400	0.360	0.400	0.391
46	0.421	0.444	0.204	0.263	0.255	0.322	0.409
47	0.311	0.375	0.349	0.587	0.349	0.581	0.374
49	0.689	0.781	0.382	0.418	0.376	0.436	0.447
51	0.883	0.642	0.434	0.471	0.434	0.471	0.655
52	0.811	0.727	0.434	0.566	0.453	0.597	0.601
53	0.576	0.420	0.138	0.149	0.138	0.149	0.265
54	0.554	0.421	0.196	0.184	0.159	0.159	0.245
55	0.774	0.657	0.238	0.265	0.238	0.265	0.466
59	0.637	0.490	0.214	0.246	0.260	0.246	0.282
60	0.376	0.163	0.264	0.217	0.199	0.247	0.544
61	0.697	0.540	0.177	0.237	0.182	0.167	0.209
62	0.548	0.470	0.333	0.421	0.333	0.447	0.440
63	0.617	0.543	0.407	0.553	0.428	0.545	0.519
65	0.863	0.824	0.267	0.277	0.267	0.277	0.265
66	0.552	0.662	0.500	0.888	0.492	0.883	0.869

67	0.839	0.833	0.493	0.492	0.493	0.492	0.520
68	0.655	0.649	0.457	0.479	0.458	0.481	0.466
69	0.131	0.091	0.037	0.077	0.042	0.080	0.028
71	0.401	0.442	0.282	0.262	0.282	0.262	0.301
74	0.889	0.844	0.695	0.845	0.697	0.838	0.928
77	0.960	0.833	0.473	0.570	0.483	0.570	1.150
78	0.614	0.653	0.464	0.449	0.464	0.449	0.509
80	0.129	0.138	0.262	0.262	0.260	0.260	0.168
81	0.827	0.558	0.397	0.509	0.397	0.509	0.621
82	0.958	0.707	0.446	0.485	0.446	0.485	0.708
84	0.590	0.334	0.385	0.470	0.295	0.470	0.505
85	0.602	0.419	0.148	0.148	0.148	0.148	0.169
86	0.625	0.543	0.289	0.361	0.289	0.361	0.296
87	0.872	0.869	0.305	0.353	0.305	0.353	0.351
88	0.868	0.754	0.446	0.500	0.446	0.500	0.436
89	0.731	0.705	0.495	0.538	0.495	0.538	0.543
90	0.860	0.882	0.601	0.725	0.602	0.725	1.631
91	0.642	0.566	0.365	0.365	0.333	0.317	0.436
92	0.625	0.511	0.315	0.315	0.377	0.377	0.344
93	0.700	0.563	0.557	0.701	0.569	0.701	0.728
96	0.692	0.673	0.472	0.496	0.506	0.575	0.673
97	0.650	0.602	0.262	0.328	0.262	0.328	0.327
98	0.065	0.022	0.066	0.044	0.066	0.044	0.044
100	0.625	0.626	0.230	0.265	0.243	0.280	0.569
102	0.909	0.402	0.092	0.092	0.076	0.076	0.268
103	0.831	0.492	0.159	0.156	0.213	0.238	0.282
104	0.411	0.401	0.096	0.110	0.110	0.110	0.097
105	0.595	0.759	0.288	0.330	0.478	0.478	0.498

Table 4.2 Minimum, maximum and average values of the misfit for each type of slip distribution.

Slip distribution	Min Misfit	Max Misfit	Mean Misfit
Uniform	0.065	0.960	0.630
SCC	0.022	0.953	0.564
2D GD_1	0.037	0.851	0.350
2D GD_2	0.044	0.888	0.410
2D GD_3	0.042	0.767	0.352
2D GD_4	0.044	0.883	0.413

2D GD_5	0.028	1.636	0.494

We notice that the different 2D GD optimizations are obtained by letting varying a different number of parameters. A way to judge the significance of the resulting fitting is provided by the Bayesan Information Criterion (BIC) based on the formula:

$$BIC = n \cdot \ln(R^2) + k \cdot \ln(n) \tag{4.5}$$

where:

- *n* is the samples number;
- *R* is the misfit value;
- *k* is the number of free parameters to be estimated.

The BIC is an increasing function of *R* and an increasing function of *k*. The lowest BIC value indicates the preferred model, penalising models with larger number of free parameters.

We adopt the BIC formulation in (4.5), even if it is often reported as:

$$BIC_0 = n \cdot \ln(\sigma_e^2) + k \cdot \ln(n)$$

with: $\sigma_e^2 = \frac{1}{n}\sum_{i=1}^{n}\left(U_z^i - u_z^i\right)^2$

Now, the expression (4.5) can be rewritten using the misfit definition (4.4):

$$BIC = n \cdot \ln\left(\frac{\sum_{i=1}^{n}\left(U_z^i - u_z^i\right)^2}{\sum_{i=1}^{n}(U_z^i)^2}\right) + k \cdot \ln(n)$$

With some basic algebra the expression assumes the form:

$$BIC = n \cdot \ln\left(\frac{\sum_{i=1}^{n}\left(U_z^i - u_z^i\right)^2}{n}\right) + k \cdot \ln(n) + n \cdot \ln(n) - n \cdot \ln\left(\sum_{i=1}^{n}\left(U_z^i\right)^2\right)$$

and can be rewritten as:

$$BIC = n \cdot \ln(\sigma_e^2) + k \cdot \ln(n) + C \qquad \text{with } C = n \cdot \ln(n) - n \cdot \ln\left(\sum_{i=1}^{n}\left(U_z^i\right)^2\right)$$

Hence:

$$BIC = BIC_0 + C$$

The two expressions for BIC and BIC_0 differ just for the constant term *C*.

Table 4.3 includes the BIC values for all the events and all the slip distributions, whereas Table 4.4 summarizes the minimum, maximum and average values of BIC. One can observe that for all the 2D GDs the mean value of BIC is lower than the Uniform and SCC mean values. Among all the 2D Gaussian optimizations, the best behaviour is given by the 2D GD_1 (mean BIC = -23.6×10^3) and by the 2D GD_3 (mean BIC = -23.5×10^3). Taking advantage from this result and observing that the difference between the two mean values is very small, so that the two approximation methods may be considered equivalent, in the following we will privilege the 2D GD_3 fitting, that differs from 2D GD_1 in that the angle of the major axis of the Gaussian ellipse is assumed to be fixed and equal to zero (see Section 4.2).

Table 4.3 BIC values of the 7 tested slip distributions. The total number of events in the Table is 72. Identification numbers in the first column (I.N.) are the ones given in Table 2.1.

I.N.	Uniform ($\times 10^3$)	SCC ($\times 10^3$)	2D GD_1 ($\times 10^3$)	2D GD_2 ($\times 10^3$)	2D GD_3 ($\times 10^3$)	2D GD_4 ($\times 10^3$)	2D GD_5 ($\times 10^3$)
1	-5.2	-8.0	-34.1	-34.0	-33.9	-35.3	-25.7
3	-13.9	-16.8	-20.6	-18.9	-21.3	-19.6	-14.0
4	-10.1	-5.7	-8.3	-5.5	-7.8	-5.5	-6.3
6	-3.1	-9.4	-19.3	-19.3	-16.6	-16.9	-13.9
7	-3.9	-4.4	-14.1	-17.4	-17.4	-17.4	-16.9
9	-15.1	-19.8	-26.9	-25.8	-30.0	-27.9	-23.6
10	-8.0	-8.4	-17.1	-12.9	-17.2	-12.6	-15.4
11	-9.5	-15.4	-16.2	-2.9	-16.2	-2.9	9.9
12	-11.9	-12.4	-17.2	-7.7	-17.9	-7.7	-7.1
13	-18.5	-21.3	-39.0	-34.5	-39.0	-34.6	-28.3
15	-3.7	-4.0	-3.2	-2.6	-5.3	-5.1	-5.3
16	-3.6	-5.8	-19.3	-15.1	-19.2	-14.2	-9.9
17	-17.3	-22.3	-30.9	-29.5	-24.8	-22.0	-2.2
18	-11.8	-13.7	-32.7	-31.3	-32.0	-31.3	-28.8
20	-14.5	-12.9	-11.9	-6.1	-11.9	-7.5	-10.2
22	-4.7	-4.7	-13.5	-13.5	-13.5	-13.5	-13.2
25	-16.1	-11.9	-8.9	-8.3	-8.9	-8.3	-10.3
30	-5.2	-4.8	-17.3	-17.3	-14.3	-14.3	-5.5
31	-17.2	-14.6	-24.3	-12.5	-24.3	-13.2	-14.1
32	-33.9	-21.5	-39.1	-35.5	-39.1	-35.5	-39.4
33	-10.0	-17.0	-35.5	-26.4	-35.5	-26.4	-35.0

35	-13.6	-22.5	-36.3	-30.0	-38.5	-41.2	-38.5
37	-23.1	-24.6	-25.0	-10.4	-25.0	-10.4	-28.8
40	-1.9	-3.6	-7.3	-6.9	-7.3	-6.1	-2.8
41	-7.2	-12.0	-16.0	-16.0	-15.7	-16.5	-16.2
42	-0.8	-0.9	-20.8	-21.4	-20.8	-21.5	0.0
43	-28.9	-31.5	-22.4	-22.4	-28.2	-30.4	-19.5
44	-6.3	-7.1	-23.8	-23.8	-23.8	-23.8	-22.7
45	-3.8	-4.9	-20.4	-18.3	-20.4	-18.3	-18.8
46	-17.3	-16.2	-31.8	-26.7	-27.3	-22.6	-17.8
47	-23.4	-19.6	-21.0	-10.6	-21.0	-10.8	-19.6
49	-7.4	-4.9	-19.2	-17.4	-19.5	-16.6	-16.1
51	-2.5	-8.8	-16.6	-15.0	-16.7	-15.0	-8.4
52	-4.2	-6.4	-16.7	-11.4	-15.8	-10.3	-10.2
53	-11.0	-17.3	-39.5	-38.0	-39.5	-38.0	-26.6
54	-11.8	-17.3	-32.6	-33.9	-36.7	-36.7	-28.1
55	-5.1	-8.4	-28.7	-26.5	-28.7	-26.5	-15.3
59	-9.0	-14.3	-30.8	-28.1	-26.9	-28.1	-25.3
60	-19.5	-36.2	-26.6	-30.6	-32.2	-28.0	-12.2
61	-7.2	-12.3	-34.6	-28.8	-34.0	-35.8	-31.3
62	-12.0	-15.1	-22.0	-17.3	-22.0	-16.1	-16.4
63	-9.6	-12.2	-17.9	-11.8	-16.9	-12.1	-13.1
65	-2.9	-3.9	-26.4	-25.6	-26.4	-25.7	-26.5
66	-11.9	-8.2	-13.8	-2.3	-14.2	-2.5	-2.8
67	-3.5	-3.6	-14.1	-14.2	-14.1	-14.2	-13.1
68	-8.4	-8.6	-15.6	-14.7	-15.6	-14.6	-15.2
69	-40.6	-47.9	-66.0	-51.3	-63.6	-50.6	-71.4
71	-18.3	-16.3	-25.3	-26.7	-25.3	-26.8	-24.0
74	-2.4	-3.4	-7.2	-3.3	-7.2	-3.5	-1.5
77	-0.8	-3.6	-14.9	-11.2	-14.5	-11.2	2.8
78	-9.8	-8.5	-15.3	-16.0	-15.3	-16.0	-13.5
80	-41.0	-39.6	-26.8	-26.8	-26.9	-26.9	-35.7
81	-3.8	-11.7	-18.4	-13.5	-18.5	-13.5	-9.5
82	-0.9	-6.9	-16.1	-14.4	-16.1	-14.4	-6.9
84	-10.5	-21.9	-19.1	-15.1	-24.4	-15.1	-13.7
85	-10.1	-17.4	-38.2	-38.2	-38.2	-38.2	-35.5
86	-9.4	-12.2	-24.8	-20.3	-24.8	-20.4	-24.4
87	-2.7	-2.8	-23.7	-20.8	-23.7	-20.8	-20.9
88	-2.8	-5.6	-16.1	-13.8	-16.1	-13.8	-16.6
89	-6.2	-7.0	-14.0	-12.4	-14.0	-12.4	-12.2
90	-3.0	-2.5	-10.1	-6.4	-10.1	-6.4	9.8
91	-8.9	-11.4	-20.1	-20.1	-22.0	-23.0	-16.6

92	-9.4	-13.4	-23.1	-23.1	-19.5	-19.5	-21.3
93	-7.1	-11.5	-11.7	-7.1	-11.3	-7.1	-6.3
96	-7.4	-7.9	-15.0	-14.0	-13.6	-11.0	-7.9
97	-8.6	-10.1	-26.7	-22.3	-26.8	-22.3	-22.4
98	-54.7	-76.0	-54.3	-62.4	-54.3	-62.5	-62.3
100	-9.4	-9.4	-29.4	-26.6	-28.3	-25.5	-11.2
102	-1.9	-18.2	-47.7	-47.7	-51.4	-51.4	-26.3
103	-3.7	-14.2	-36.7	-37.1	-30.9	-28.7	-25.3
104	-17.8	-18.2	-46.8	-44.2	-44.2	-44.2	-46.6
105	-10.4	-5.5	-24.9	-22.1	-14.7	-14.8	-13.9

Table 4.4 Minimum, maximum and average values of the BIC for each type of slip distribution.

Slip distribution	Min BIC ($\times 10^3$)	Max BIC ($\times 10^3$)	Mean BIC ($\times 10^3$)
Uniform	-54.7	-0.8	-11.0
SCC	-76.0	-0.9	-13.6
2D GD_1	-66.0	-3.2	-23.6
2D GD_2	-62.4	-2.3	-20.8
2D GD_3	-63.6	-5.3	-23.5
2D GD_4	-62.5	-2.5	-20.7
2D GD_5	-71.4	9.9	-18.0

4.5 Regression laws of the 2D GD parameters

In analogy with regression laws linking focal and slip parameters with magnitude, one can establish scaling laws including the Gaussian Distribution parameters. For the reason explained in the previous section, we restrict our analysis to the 2D GD_3 fitting optimizations. Accordingly, the parameters to be taken into account are the standard deviations along horizontal and down-dip axis σ_1 and σ_2 and the maximum displacement MD.

One finding is that both the along-strike and down-dip standard deviations seem to correlate with the magnitude (see Figure 4.2 and Figure 4.3). In both cases, reverse

earthquakes present the largest slope coefficient and the highest values of σ_1 and σ_2 for an assigned moment magnitude M_W. Concerning σ_1 (along-strike), the lowest values of σ_1 and the lowest slope coefficients are found for earthquakes characterized by normal focal mechanism. Instead, σ_2 (down-dip) has its lowest values for strike-slip earthquakes, which underlines once more the natural evolution of the rupture process in the along-strike direction. Plots of σ_2 show a larger spread in the slope values of the different focal mechanisms laws than plots of σ_1. This fact can suggest a higher sensitivity of the down-dip standard deviation to rupture mechanism.

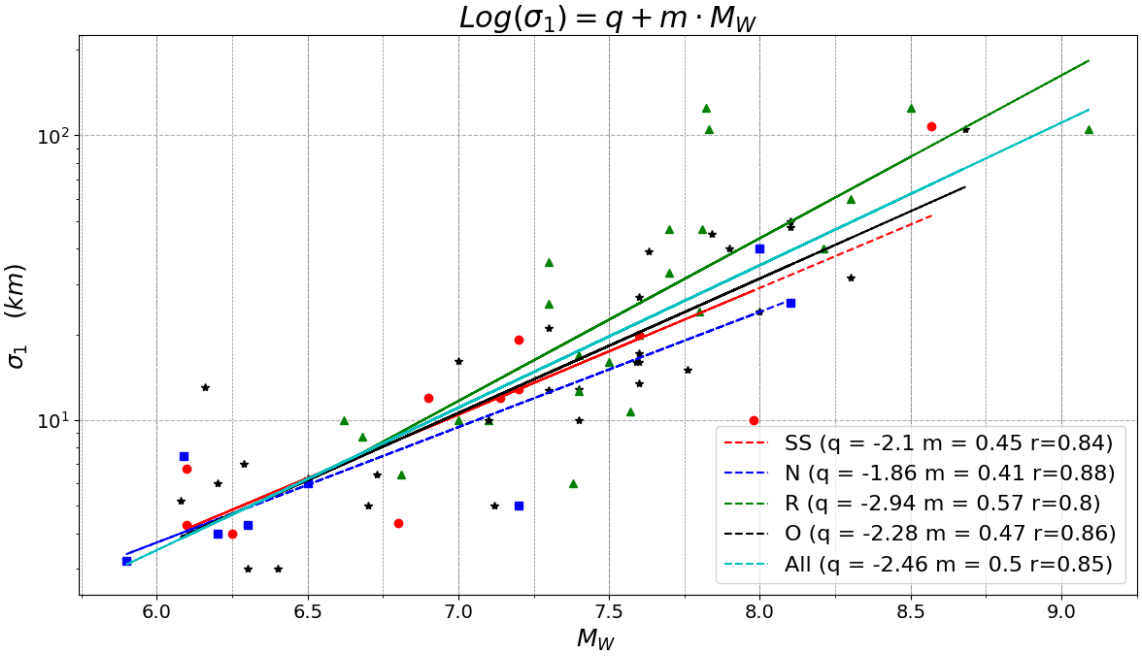

Figure 4.2 Regression law of along-strike standard deviation σ_1 vs. magnitude.

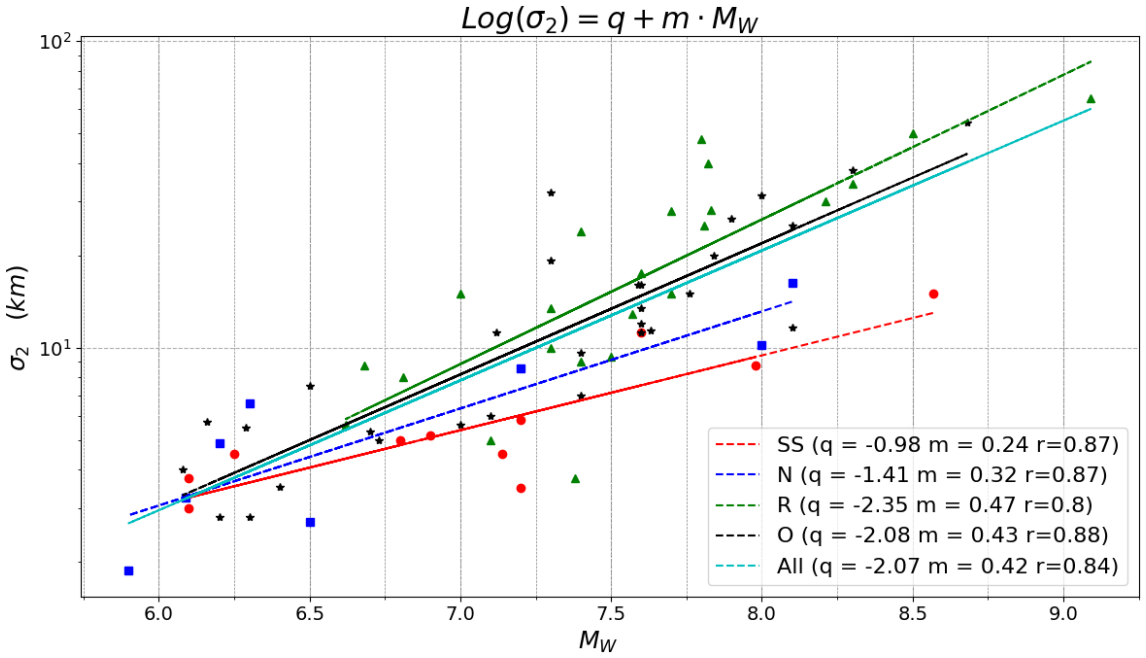

Figure 4.3 Regression law of down-dip standard deviation σ_2 vs. magnitude.

It is interesting to analyse the trend of the standard deviations in relation to the rupture dimensions as portrayed in Figure 4.4 and Figure 4.5. In this case, a linear regression is effective in describing the correlation between the Gaussian parameters σ_1 and σ_2 on one side, and the rupture length and width on the other. Remarkably, with the exception of the strike-slip events, the slope coefficient is almost the same for all the earthquakes, which indicates a homogeneous regularity in the increasing of the asperity size with rupture dimensions. From our results, we can conclude that the standard deviations of the slip distribution increase with the rate of 1/5 of the source size along the relative direction. The correlation coefficient is particularly good for the down-dip direction.

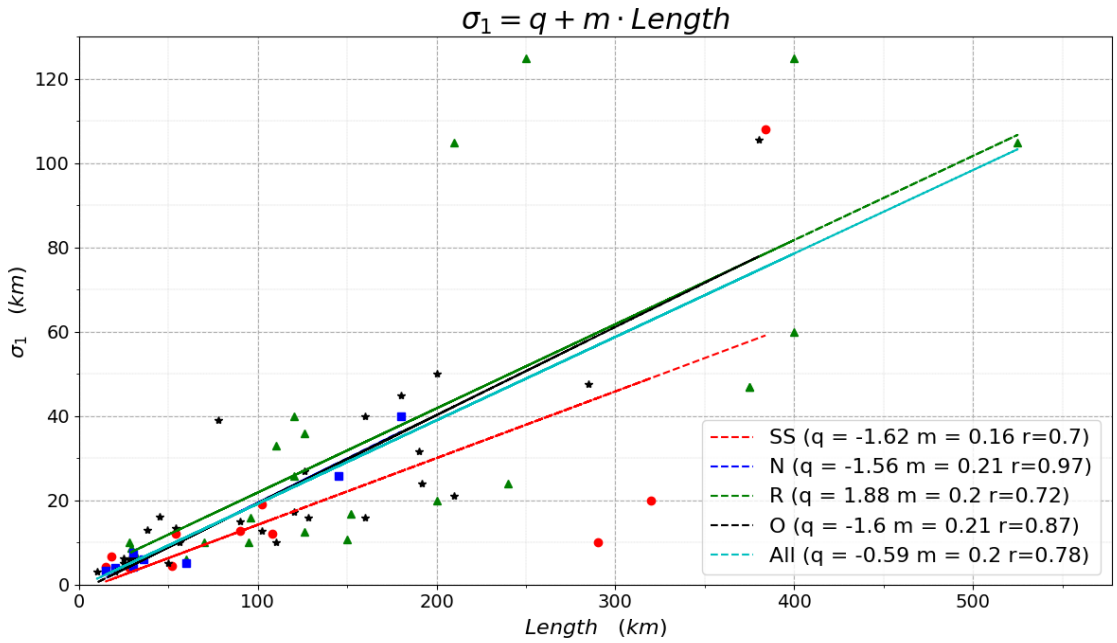

Figure 4.4 Linear regression law of the along-strike standard deviation σ_1 vs. rupture length.

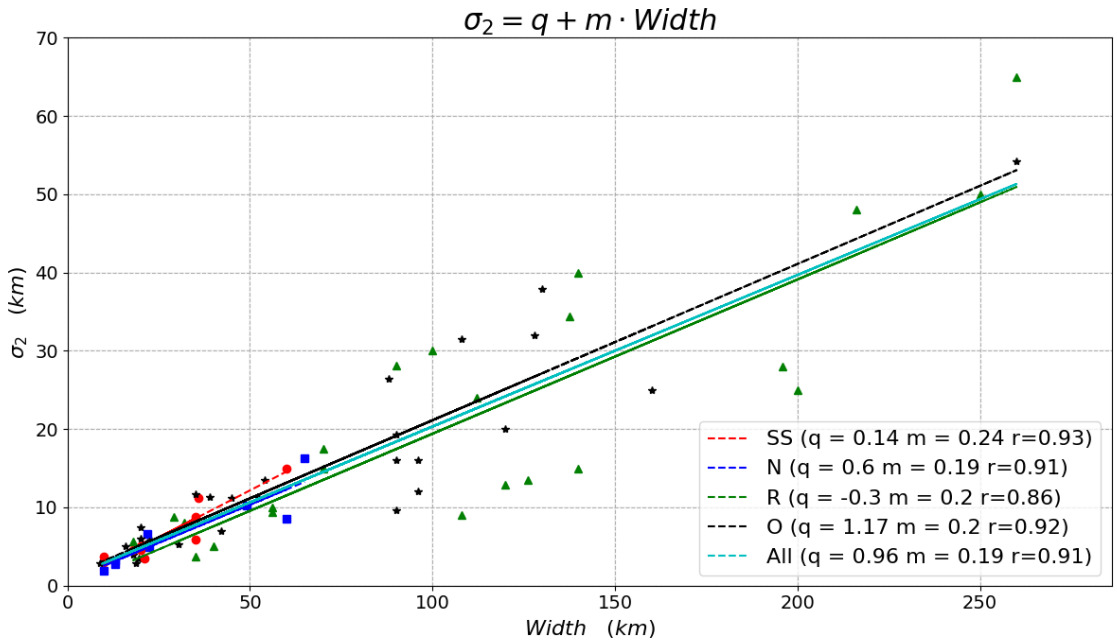

Figure 4.5 Linear regression law of the down-dip standard deviation σ_2 vs. rupture width.

The other parameter taken into account is the maximum displacement of the 2D GD centre value (see Figure 4.6). It is not surprising that the correlations we find here are not so different from the ones involving the MD values of the FFM, since the Gaussian peak MD_{2DGD} is found through a minimization process exploring a space window around the

FFM MD. In both cases the best correlation is obtained for the strike-slip and normal earthquakes, and for the Gaussian parameter the correlation coefficient is even higher (r = 0.93, r = 0.96 for strike-slip and normal shocks, respectively).

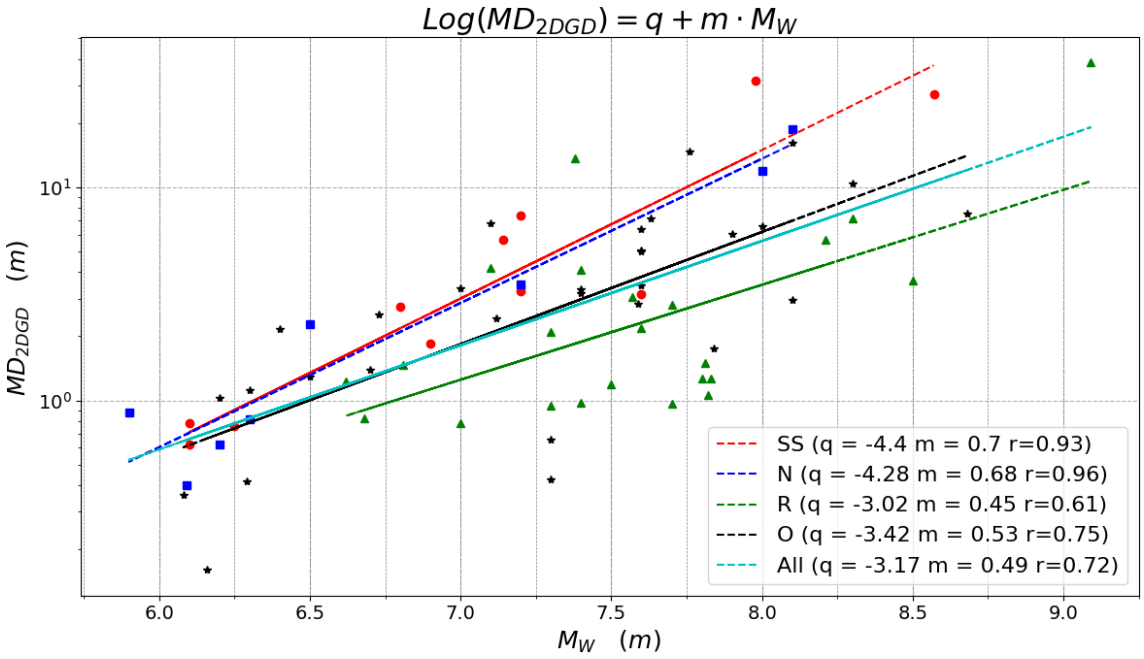

Figure 4.6 Regression law of the 2D GD maximum displacement MD vs. magnitude.

4.6 The slip in the hypocentre for distributions of the type 2D GD_3

Considering the relation between the magnitude of the slip in the hypocentre and the peak value for the Gaussian distributions is an additional way to analyse the relation between the hypocentre and the main fault asperity.

Let us consider the 2D GD slip expression given by the formula (4.3) in Section 4.2. If we deal with the 2D GD_3, then $\theta = 0$, hence:

$$a = \frac{1}{2\sigma_1^2}\ ; \qquad b = 0; \qquad c = \frac{1}{2\sigma_2^2}$$

If we call the coordinates of the hypocentre with (x_H, y_H) over the fault plane, the value assumed by the slip in correspondence with the hypocentre is:

$$u_H = u_{max} \cdot e^{-\left[\frac{(x_H - x_0)^2}{2\sigma_1^2} + \frac{(y_H - y_0)^2}{2\sigma_2^2}\right]}$$

The histogram in Figure 4.7 shows the number of events for different intervals of the ratio u_H/u_{max}. The average value of this ratio is 0.555.

There are 16 cases in which the ratio is smaller than 0.2, which indicates that the hypocentre lies far from the main asperity.

Considering the histograms for different focal mechanisms it is possible to observe that the strike-slip events are those with the smallest average value (0.355) while other mechanisms show an average value larger than 0.5. Reverse earthquakes are those with the highest one (0.642).

Table 4.5 shows, together with the averages, different percentile values for the ratio u_H/u_{max}. Looking at the medians, with the exception of the strike-slip mechanisms, one sees that its upper limit is larger than 0.62 and that the maximum is reached by normal events, presenting a median of 0.713. Taking into account the 10^{th} percentile, one can observe the peculiar behaviour of the strike-slip events.

Table 4.5 Mean value and five different percentile values (10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}) of the ratio u_H/u_{max} for different focal mechanisms.

	Mean	10^{th}	25^{th}	Median (50^{th})	75^{th}	90^{th}
Strike-slip	0.355	0.0	0.018	0.233	0.734	0.795
Normal	0.574	0.132	0.178	0.713	0.902	0.926
Reverse	0.642	0.224	0.466	0.693	0.868	0.988
Oblique	0.557	0.111	0.259	0.627	0.776	0.950
All	0.555	0.063	0.228	0.657	0.802	0.945

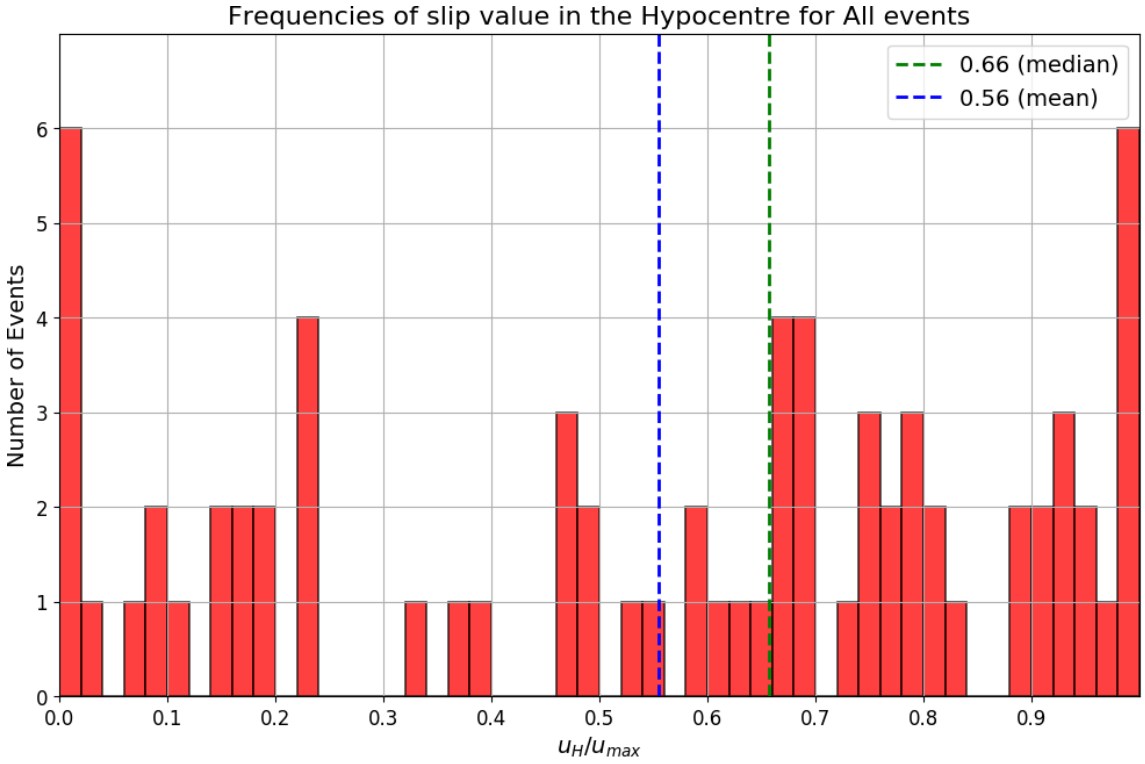

Figure 4.7 Histogram of the hypocentre slip value (u_H) normalized to the 2D GD$_3$ maximum displacement (u_{max}).

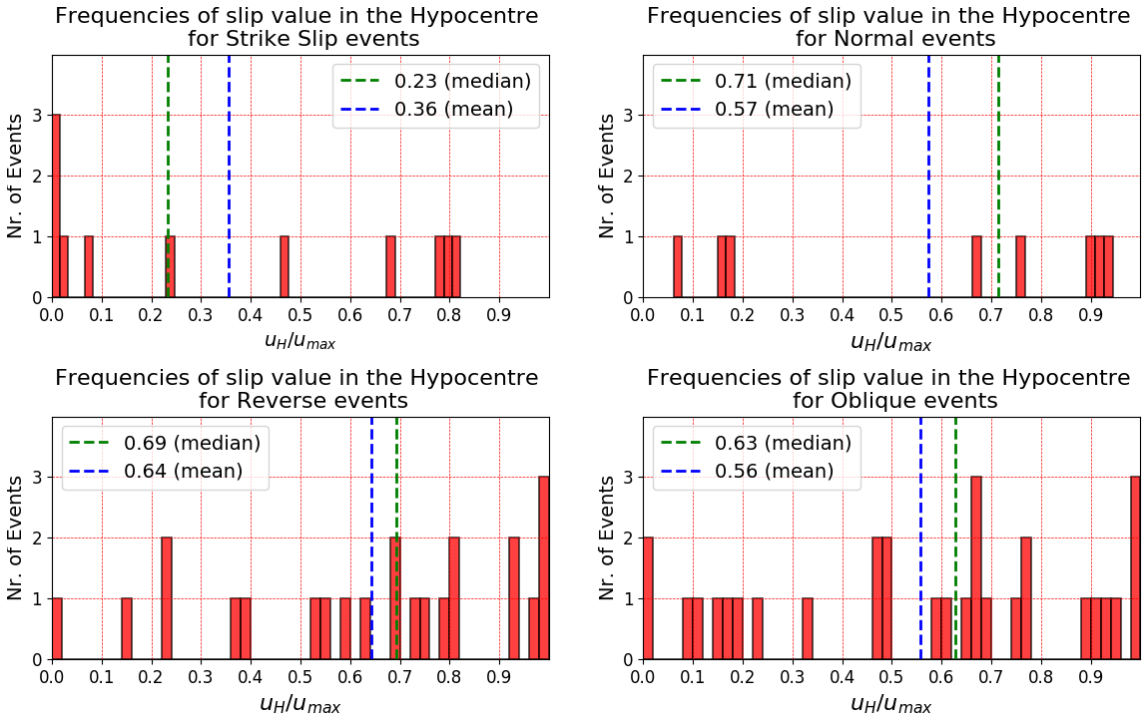

Figure 4.8 Histograms of the u_H/u_{max} ratios for the different focal mechanisms.

The slip value at the hypocentre u_H has been analyzed also in relation to the average slip u_{mean} characterizing the slip distribution.

Table 4.6 summarizes some characteristic values of the ratio u_H/u_{mean} in a way similar to what was done above for u_H/u_{max}. The Table suggests some observations. Reverse-faulting earthquakes are those showing the highest u_H/u_{mean} average ratio (4.509), followed by normal, oblique, and finally strike-slip events (1.416). For reverse, normal and oblique ruptures, moreover, 90% of cases present the hypocentral slip value definitely larger than the 50% of the average slip over fault plane. For strike-slip events the 75% of the ruptures are characterized by u_H larger than 17.5% of u_{mean}.

Table 4.6 Mean value and five different percentile values (10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}) of the ratio u_H/u_{mean} for the different focal mechanisms.

	Mean	10^{th}	25^{th}	Median (50^{th})	75^{th}	90^{th}
Strike-slip	1.416	0.001	0.175	1.416	2.294	2.847
Normal	3.352	0.608	0.767	3.352	3.677	6.296
Reverse	4.509	0.919	1.758	4.509	6.777	10.747
Oblique	2.619	0.766	1.332	2.619	3.099	5.359
All	3.120	0.345	1.181	2.146	3.677	7.077

The analysis reported in the above tables and figures highlights some peculiarities of the relative position between the main asperity and the nucleation point. Indeed, while in the previous chapter we have considered the relation between hypocentre and maximum slip locations in terms of distances, here we have directly taken into account the slip value assumed in the hypocentre, following our 2D GD distribution. Generally, it is possible to conclude that hypocentres are preferentially located in regions with significant amount of slip. The only faulting mechanisms for which this trend is not true are the strike-slip ruptures.

These results agree with findings by Manighetti et al. (2005) and Mai et al. (2005), but are better highlighted and quantified here.

5 The 2015 Illapel earthquake, Chile

By taking advantage from the analysis on the slip distributions carried out in the previous chapters, here we study the 16 September 2015 Illapel earthquake, Chile.

There are several reasons inducing us to consider this event, in particular:

- it generated a tsunami;
- there is an FFM in the database, more precisely the event with I.N.105 in Table 2.1;
- being a recent event, a lot of data from observations and studies are available.

The Illapel earthquake occurred in a seismic region that was the scene of other major earthquakes. As the consequence of the subduction of the Nazca plate beneath the South American plate many great earthquakes repeatedly occur offshore Chile. We cite, as recent example, the Maule earthquake of 27 February 2010, M_W = 8.9 (I.N. = 72 in Table 2.1), the Iquique earthquake on April 1st 2014, M_W = 8.1 (I.N. = 96 in Table 2.1).

The Illapel event in particular, is considered to be a re-rupture of the 1943 earthquake (Mw 7.9, Beck et al. 1998), whose related tsunami was recorded in Japan with height of 10 cm in Hanasaki and 25 cm in Kushimoto (Hatori 1968; Watanabe 1998).

As a matter of fact, many of the offshore Chilean earthquakes generated tsunamis, causing damage not only on the Chilean coast but also across the Pacific Ocean. As regards for instance the famous earthquake of 1960 (22 May, M_w=9.5, Barrientos and Ward, 1990), the largest in the South America seismic history, its tsunami caused the death of about 2000 people on the Chilean coast (Atwater et al. 1999), plus 61 and 142 fatalities in Hawaii and Japan, respectively (Atwater et al. 1999; Watanabe 1998). Regarding the 2010 Maule earthquake, its tsunami had a run-up exceeding 15 m on the Chilean coast, and the total fatalities were 156; tsunami waves reached 2 m height on the Japanese coast, still causing some property damage (Fujii and Satake, 2013).

The 2015 Illapel earthquake triggered a tsunami that reached the nearest coastal areas very fast, within few minutes of the earthquake initiation. According to some eyewitness accounts the tsunami attacked one coastal village immediately after the earthquake. A tsunami threat message from the Pacific Tsunami Warning Center (PTWC) was issued 7 min after the main shock and the National Hydrographic and Oceanic Service (SHOA),

the organization in charge of the Chile's National Tsunami Warning System, issued a tsunami alarm message 8 min after the earthquake (SHOA, 2015a). A preventive evacuation status for the entire Chilean Coast was declared by the National Emergency Office (ONEMI) 11 min after the earthquake (ONEMI, 2015). Unfortunately, despite the quick evacuation, the tsunami caused the death of 8 people (ONEMI, 2015).

A number of post-tsunami field surveys measured tsunami heights on the Chilean coast (Aranguiz et al. 2016; Contreras Lopez et al. 2016). The maximum runup heights were reported as 10.8 m at the Totoral fishing village (30.37°S) by Aranguiz et al. (2016), and as 13.6 m at La Cebada (30.97°S), by Contreras-Lopez et al. (2016). Aranguiz et al. (2016) attributes the high runup at Totoral to the deep offshore bathymetry and the pocket beach morphology that funneled tsunami waves ashore. Except for the two anomalous locations mentioned above, tsunami heights were up to 9 m on the coast between 29°S and 32°S, and smaller farther south and north (Satake and Heidarzadeh 2017).

The tsunami was also recorded on coastal tide gauges (Aranguiz et al. 2016; Heidarzadeh et al. 2016) located in places shown in Figure 5.1(a). The earliest measured tsunami arrival of ≈ 15 min with zero-to-peak amplitude of ≈ 2 m was recorded at the Pichidangui tide-gauge station, just south of the epicentre. To the north, at the Coquimbo tide-gauge station, the first arrival was at 23 min with ≈1 m amplitude, but the largest tsunami amplitude of 4.7 m was recorded ≈1.5 h after the earthquake.

The tide-gauge tsunami amplitudes vary in the range 1–2 m with a maximum value of 4.7 m. Instead, runup heights vary in the range 3–6 m with a peak of ≈ 11–13.6 m (Satake and Heidarzadeh, 2017). One can state that tsunami runup heights were approximately up to three times the tide-gauge amplitudes along the Chilean coast.

Thanks to the NOAA Global Historical Tsunami Database one can get a picture of the tsunami heights across the Pacific Ocean: 1.37 m in Marquesas Islands, 0.83 m in Hawaii (Hilo), 0.52 m in New Zealand (Chatham), and 0.10 m in Australia (Port Kembla). In Japan, the Japan Meteorological Agency issued tsunami advisory approximately 19 h after the earthquake and 3 h before the first tsunami arrival at Japan, with expected tsunami heights of ≈ 1 m: the largest observed tsunami amplitude was 0.78 m on the Kuji GPS buoy.

5.1 The slip models

The slip distribution of the 2015 Illapel earthquake was studied by several researchers, using various observation data, including geodetic data, teleseismic waveforms, near-field seismic data, and tsunami waveforms.

The model reported in the SRCMOD database is the one by Okuwaki et al. (2016) who made a hybrid inversion of teleseismic waveforms and backprojection data.

They adopted the Yagi and Fukahata's (2011) inversion method, that can mitigate the effect of uncertainties in the Green's function, that is known to be a major source of modeling errors in waveform inversion procedures and that resulted in the non-uniqueness of seismic source models for the same earthquake by different researchers (e.g. Beresnev, 2003). The fault geometry was built with the constant strike and dip angles being 2.7° and 15.0°, respectively.

The initial rupture point (assumed hypocentre) was the CSN (Centro Sismologico Nacional, Universidad de Chile) determined location (31.637°S, 71.741°W, 25 km depth). The rake angle on each source node was assumed a pure thrust motion relative to the plate motion direction, in agreement with the MORVEL model (Demets et al. 2010).

Large slip was found in the shallow up-dip portion of the fault plane where a large-slip patch was centered 72 km northwest of the epicentre. The total seismic moment release was calculated to be 3.3×10^{21} Nm (MW 8.3). Okuwaki et al. (2016) assert that the slight difference in the seismic moment wrt. other studies (e.g. 2.67×10^{21} Nm, Ye et al., 2016) may derive from differences in fault geometry and slip locations along the dip direction since the seismic moment depends on assumed rigidity and this latter increases with depth.

The final slip map obtained by Okuwaki et al. (2016) is portrayed in Figure 5.1(b). It is the result of two distinct episodes of rupture propagation near the hypocentre and near the north of it, characterized by variable rupture front velocities.

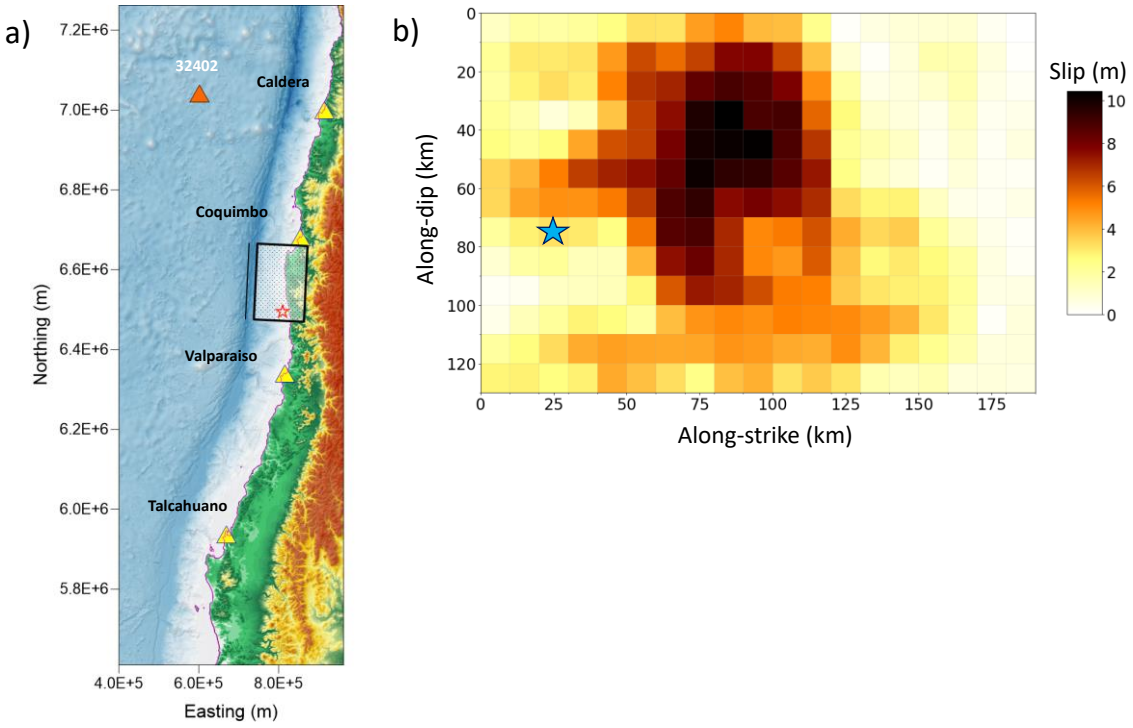

Figure 5.1 Illapel earthquake. Fault model by Okuwaki et al., 2016. (a) Fault plane along the Chilean coast. Star: epicentre, Yellow triangles: tide-gauge stations, Red Triangle: offshore DART buoy. The reference system used for this map and for the following ones is the WGS84, with UTM projection (18th zone, southern hemisphere). (b) Slip map obtained by Okuwaki et al., included in the SRCMOD database (Star: hypocentre).

5.2 Slip distributions and tsunami simulations

Starting from the reference SRCMOD slip map we have calculated the seven on-fault slip distributions corresponding to the types illustrated in the previous chapter, namely:

- the homogeneous distribution;
- the SCC distribution;
- the five best fitting 2D GDs.

In this study we will compare the tsunami resulting from the various models We abstain from considering how well the calculated tsunamis match the real data here, since minimizing the discrepancy wrt the experimental data, though a relevant factor, is not the goal of the analysis carried out in this chapter. Indeed, expectedly the performance of all the derived models is linked to the performance of the reference model.

Tsunami simulations have been carried out by solving linear shallow-water (LSW) equations by means of the UBO-TSUFD model (Tinti and Tonini, 2013), that is a numerical code developed and maintained by the Tsunami Research Team (TMT) of the University of Bologna. It has been run on a coarse resolution bathymetric grid (900 m)

built on GEBCO_2014 data. Since coseismic deformations occur on the fault in a very short time, i.e. in a time so short that tsunami waves can travel only a very short distance before the rupture process ends, we assume the following initial conditions for the tsunami simulations that are quite typical in tsunami modelling:

- the initial vertical displacement of the sea surface is equal to the vertical displacement of the sea bottom:
- the initial horizontal velocity of the sea water is zero everywhere.

The coastal boundary has been treated as a vertical wall where pure wave reflection occurs. We have therefore not considered inundation effects, and technically we have not calculated runup heights. Nevertheless, since usually runup data are well correlated with coastal tsunami heights, we consider that a tsunami linear model with fixed coastal boundary is able to provide results sufficiently good for our goal, that is to compare tsunamis from different slip distributions.

The simulated tsunami data of interest for our analysis are:

- extreme (maxima and minima) water elevations along the 10-m isobaths.
- synthetic water-level time series at one offshore DART station and at four coastal tide-gauge stations.

5.2.1 Observed data

The instrumental data of the DART buoy (Station 32402) have been retrieved from the site https://www.ndbc.noaa.gov/station_page managed by the NOAA National Centers for Environmental Information, while tide-gauge data have been downloaded from the site http://www.ioc-sealevelmonitoring.org/ developed and maintained by the VLIZ (Vlaams Instituut voor de Zee, Ostenda, Belgium) under IOC mandate. The Chilean stations from North to South are Caldera (station code: cald2), Coquimbo (station code: coqu2), Valparaiso (station code: valp and valp2) and Talcahuano (station code: valp and valp2).

These time series have been cleaned from the ocean tide, filtered from high frequencies, and temporarily shifted in such a way that their origin time coincides with the origin time of the earthquake (namely 22:54:33 UTC, 16 September 2015). After processing, the time series rate is 1 sample per minute.

Runup height is defined as the maximum ground elevation reached by the tsunami on a sloping shoreline. Runup measurements have been obtained from the Global Historical Tsunami Database maintained by the NGDC/WDS National Geophysical Data Center / World Data Service, that is part of the NOAA structure of the National Centers for Environmental Information, doi:10.7289/V5PN93H7 [last access date 26 June 2019].

5.2.2 Synthetic data

Figure 5.2 and Figure 5.3 show the vertical co-seismic displacement fields induced by the considered slip distributions. The corresponding misfit data are all reported in the last row of Table 4.1. Because the fields of 2D GD_3 and 2D GD_4 happen to be quite similar for this particular case we display the outputs only for the 2D GD_3. All fields have to be compared against the Okuwaki et al. model, taken as reference and portrayed in Figures 5.2- 5.3. Looking at fields of Figure 5.2, it is immediate to notice that the one computed through the 2D GD_1 (d) reproduces the shape of the displacement more faithfully than the uniform (b) and SCC (c) source distributions.

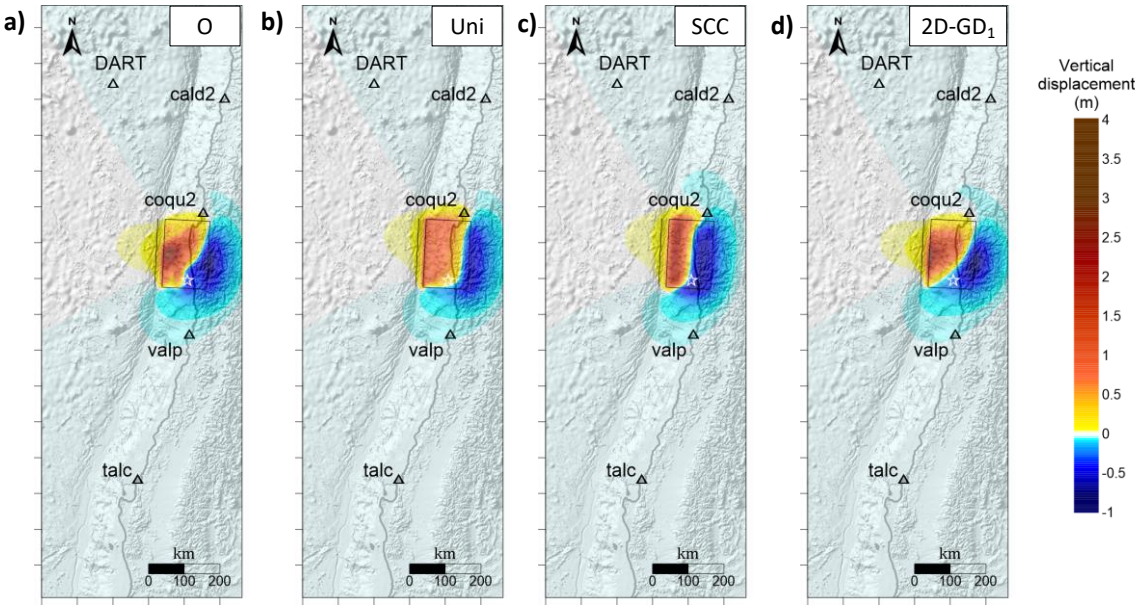

Figure 5.2 Vertical co-seismic displacement fields obtained by applying Okada's formulas to the slip distributions under study. (a): The reference SRCMOD model by Okuwaki et al. (O). (b), (c), (d) The uniform (Uni), SCC and

2D GD_1 slip distributions, respectively. The five triangles in each map identify the four coastal stations and the DART buoy. The rectangle highlights the seismic fault. The white star is the epicentre.

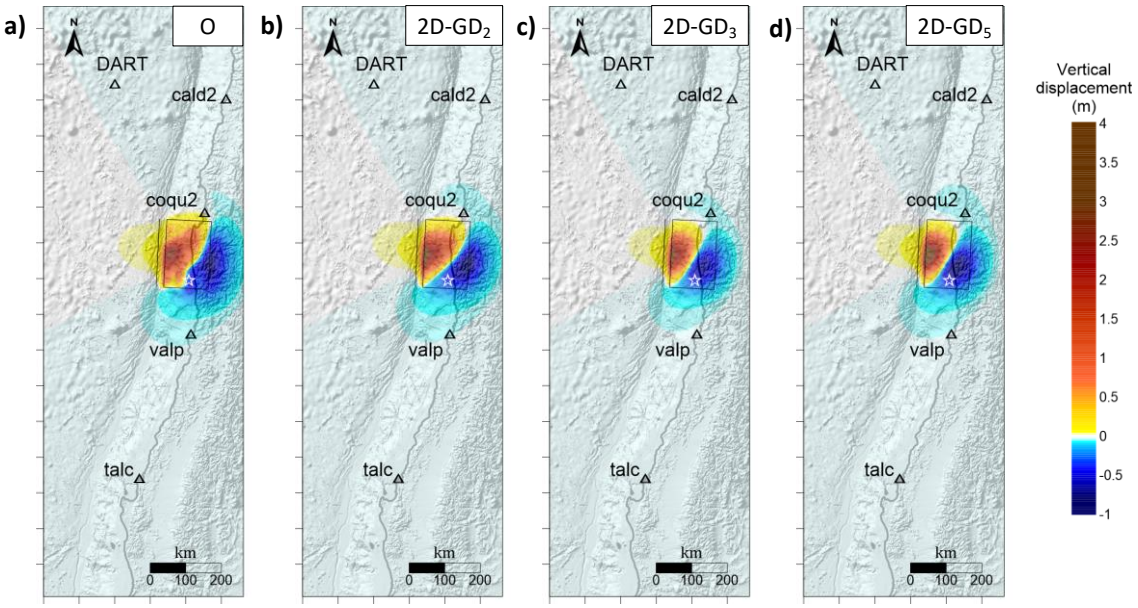

Figure 5.3 Vertical co-seismic displacement field from the reference model and from the other Gaussian slip distributions, namely Okuwaki et al.'s reference model (a), 2D GD_2 (b), 2D GD_3 (c), 2D GD_5 (d). For further details, see Figure 5.2.

5.2.3 Maxima and minima water elevations

The good behaviour of the Gaussian distributions is also confirmed by the maximum (Figures 5.4 - 5.5) and minimum (Figures 5.6 - 5.7) water elevation fields obtained from the tsunami simulations, that have been conducted for the first 10 hours after the earthquake started. The shape of the field is better reproduced by the 2D GDs and do not underestimate the analogue results from SRCMOD model, which could be an important feature from the tsunami alert point of view.

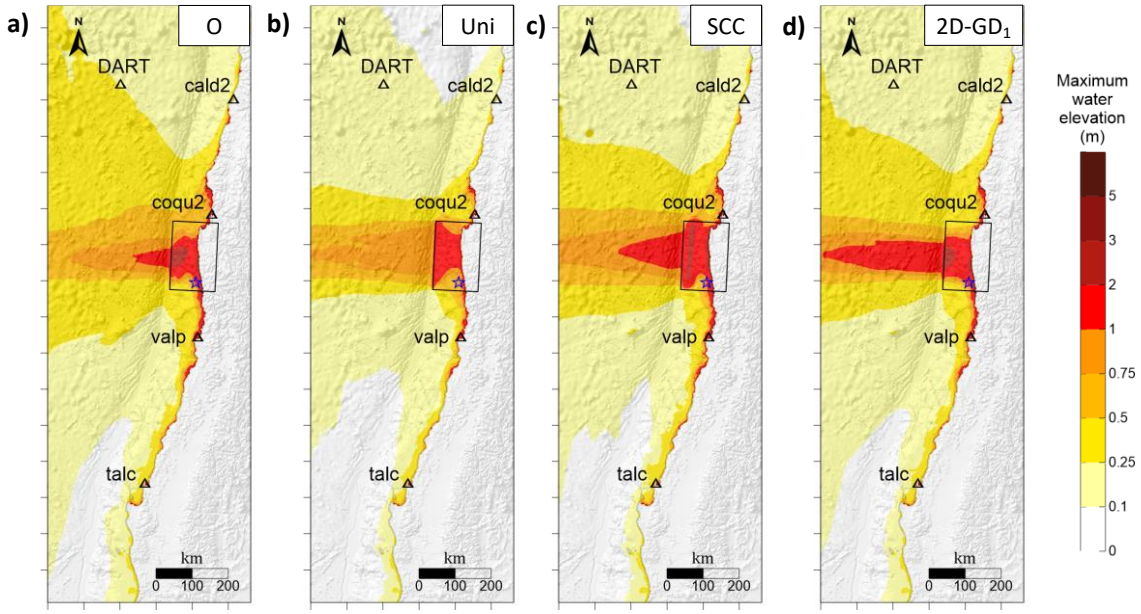

Figure 5.4 Maximum water elevation fields produced by the different sea-bottom displacements. O (a), uniform (b), SCC (c), 2D GD_1 (d). For further details, see Figure 5.2.

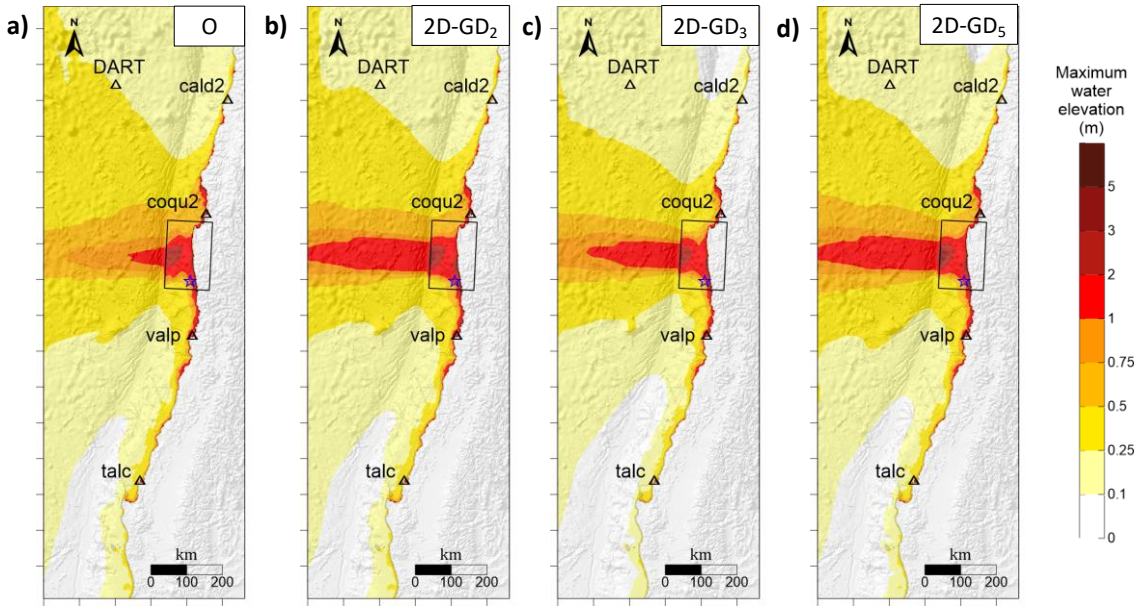

Figure 5.5 Maximum water elevation fields produced by the different sea-bottom displacements. O (a), 2D GD_2 (b), 2D GD_3 (c), 2D GD_5 (d). For further details, see Figure 5.2.

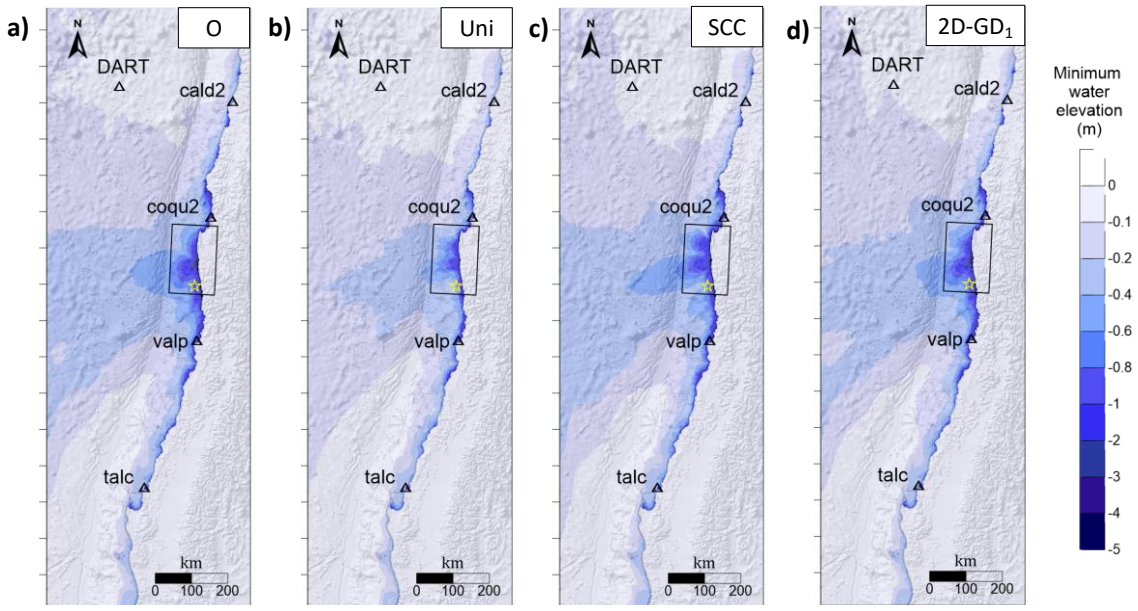

Figure 5.6 Minimum water elevation fields produced by the different sea-bottom displacements. O (a), uniform (b), SCC (c), 2D GD_1 (d). For further details, see Figure 5.2.

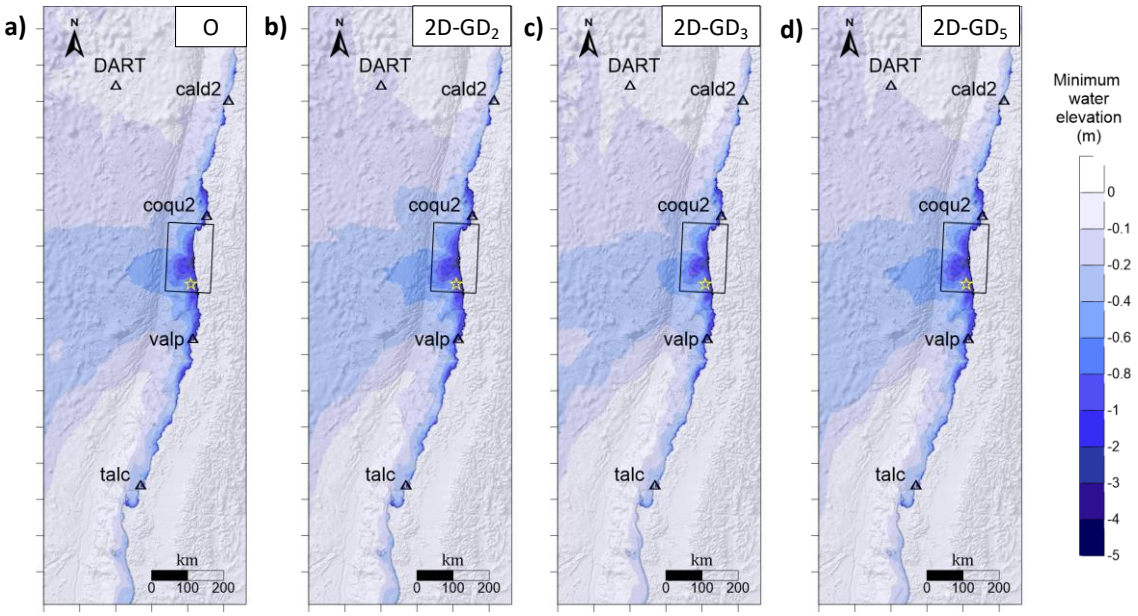

Figure 5.7 Minimum water elevation fields produced by the different sea-bottom displacements. O (a), 2D GD_2 (b), 2D GD_3 (c), 2D GD_5 (d). For further details, see Figure 5.2.

5.1 Analysis of the waveform signals

In this subsection we focus on the marigrams obtained in the four coastal station of Caldera (cald2), Coquimbo (coqu2), Valparaiso (valp/valp2), Talcahuano (talc/talc2) and in the offshore DART buoy station 32402 (180 nautical miles west of Caldera, Chile).
We present the results of the quantitative analyses performed over the waveforms signals obtained. These analyses aim at quantifying the degree of similarity between the signals produced by the different slip distributions (Uni, SCC and the 2D-GDs) and the signal produced by the reference slip model (the O model). The signals are taken in the time window [0, 36000] s (the first 10 hours after the rupture occurred), with a time step of one minute between one measure and the adjacent one.

Figures 5.8 - 5.12 report, for each station, the waveforms calculated through the simulations, overlapped to the observed signal (dashed line). For each station, it is reported the time from which the comparison analyses have been started (the blue dotted line in the figures). The starting time t_i differs from station to station due to the different tsunami arrival time at the stations, that is clearly related to the distance of the stations from the source. Removing the zero-part of the signals is important to perform a better comparison between the major oscillations of the waveforms.
The starting time is taken coincident with the earthquake rupture time (hence t_i=0 s) for the closer stations of Coquimbo and Valparaiso. For the station of Caldera and Talcahuano are respectively 20 min and 60 min. For the DART buoy, the starting time is settled at 25 min to exclude the part of the record due to the seismic signal, since we are interested in the analysis of the signal that refers to the tsunami evolution.

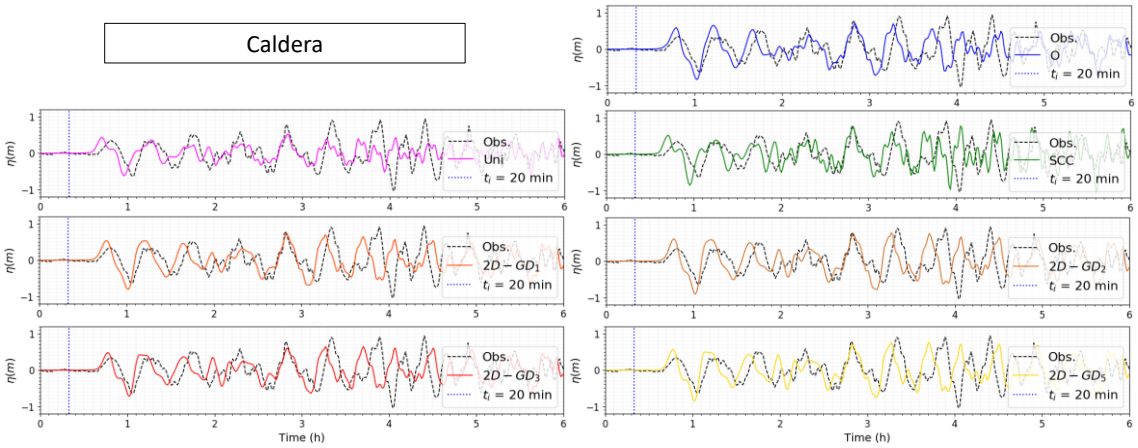

Figure 5.8 Marigrams obtained at the station of Caldera for (from top-right to bottom-right) the O, Uni, SCC, 2D GD_1, 2D GD_2, 2D GD_3, 2D GD_5 distributions. The black dashed signal refers to the observed one. The vertical blue dotted line indicates the starting time related to the station.

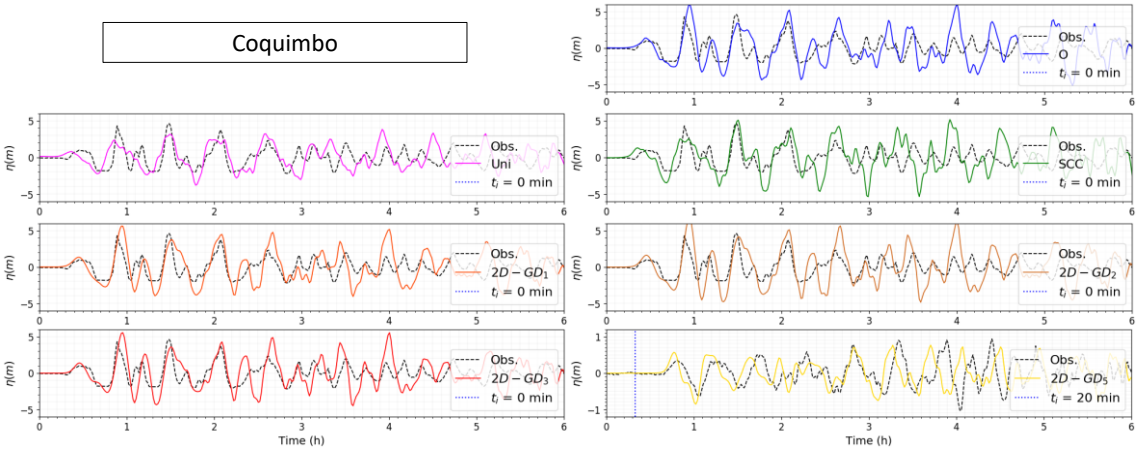

Figure 5.9 Marigrams obtained at the station of Coquimbo. For further details, see Figure 5.8.

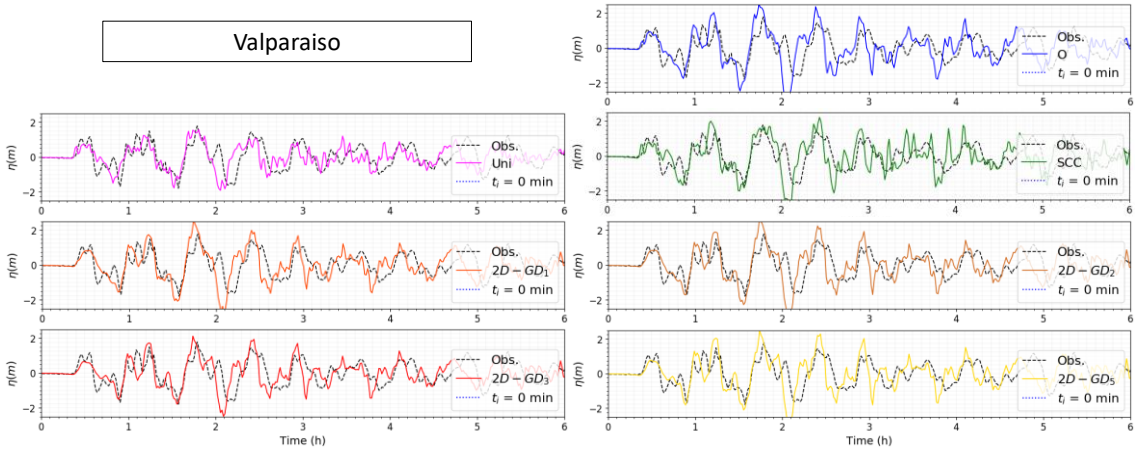

Figure 5.10 Marigrams obtained at the station of Valparaiso. For further details, see Figure 5.8.

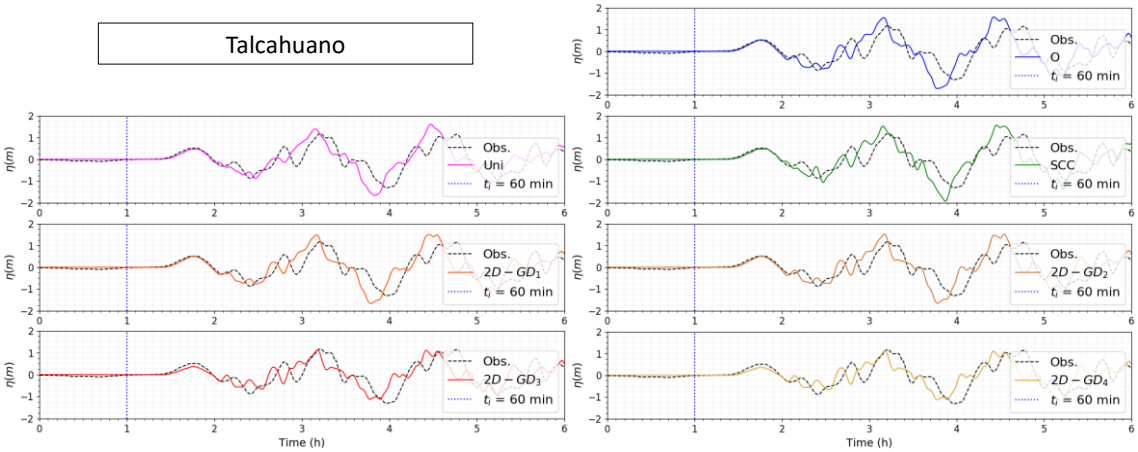

Figure 5.11 Marigrams obtained at the station of Talcahuano. For further details, see Figure 5.8.

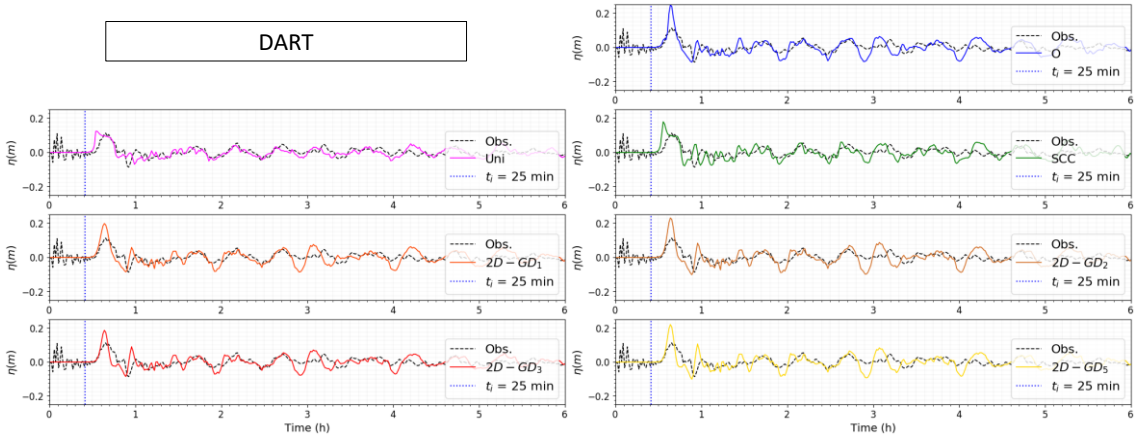

Figure 5.12 Marigrams obtained at the buoy DART station 32402. For further details, see Figure 5.8.

5.1.1 Time-shifting procedure

Considering the reported marigrams for each station, a time mismatch is often found between the observed and modelled tsunami waveforms, with the latter arriving generally earlier.

This fact is in agreement with other studies (e.g. Heidarzadeh et al., 2016). As reported by Romano et al., 2017, these early arrivals of the synthetic waveforms are mainly due to the inaccurate or too coarse bathymetric models, or to unknown or unaccounted for instrumental responses, and possible unknown clock errors of old instruments. In order to model these effects properly, a high-resolution bathymetric grid around the instrument location would be necessary, but such data are not always publicly available. In the source inversion optics, this time mismatch may affect the tsunami source model inferred by data

inversion, for example the earthquake slip distribution. But this is not our case, because we are directly deriving the tsunami waveforms from a fault slip model, and not the opposite. Moreover, we are interested in reproducing the main waveforms rather than finding the exact time of the first arrival.

Hence, the synthetic signal has been time shifted by a time τ such that the cross-correlation:

$$(f * g)[\tau] = \sum_{t} f[t]g[t+\tau] \tag{5.1}$$

assumes its maximum value within the first 1 hour of signal from the starting time t_i.

The cross-correlation of two continuous functions *f* and *g* is defined as:

$$(f * g)(\tau) = \int_{-\infty}^{\infty} f^*(t)g(t+\tau)dt \tag{5.2}$$

where *f*(t)* denotes the complex conjugate of *f(t)*.

For two real discrete functions (which is our case) the cross-correlation is defined by the equation (5.1), where *t* is the discrete index varying through the signals and τ represents the discrete shift between the two signals. In our computations, the temporal shift applied to the synthetic waveforms corresponds to the value of τ that maximizes the value expressed by the (5.1).

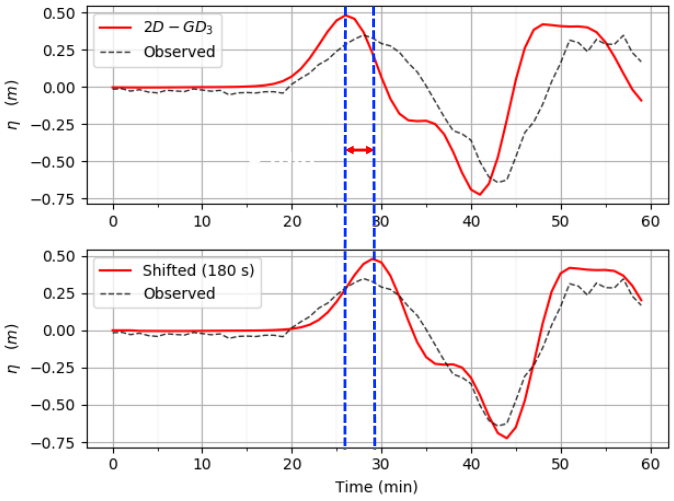

Figure 5.13 Example of waveforms taken from the set used for the analyses. Top graph: the 2D GD_3 signal at the Caldera station for the first 1 hour. Bottom graph: the signal shifted by the time maximizing the cross-correlation (180 s) The blue vertical lines highlight the temporal shift between the top and bottom signals.

Figure 5.13 permits to visualize the procedure, applied to the 2D GD_3 signals at the tide-gage station of Caldera: it is evident how the shifted signal is better superimposed over the observed signal.

Table 5.1 reports the time shifts that maximize the cross-correlation between synthetic and the observed signals for all the distributions during the first hour of recording. For all the stations the shift is never higher than 5 minutes.

The Coquimbo station is the only one for which the synthetic signals should be back shifted (by 1 minute), meaning that the synthetic waveforms tend to lag behind the observed signal. For all the other stations the modeled tsunami waveforms tend to arrive earlier than the observed ones. However, this does not represent a significant difference, since the delay times are much less than the average periods of the oscillations.

Table 5.1 Time-shift values that maximize the cross-correlation between the synthetic signals and the observed one in the first record hour for all the considered stations.

	Caldera (min)	Coquimbo (min)	Valparaiso (min)	Talcahuano (min)	DART (min)
O	2	-1	0	0	2
Uni	5	4	0	0	3
SCC	5	3	1	0	4
2D-GD_1	3	-1	0	0	2
2D-GD_2	3	-1	0	0	2
2D-GD_3	3	-1	0	0	3
2D-GD_5	3	-1	0	0	3

After the above mentioned preliminary procedures, we compare each synthetic signal *w(t)* with the observed one $w_0(t)$. The measures assumed to quantify the goodness of this comparison are:

- the misfit:

$$misfit = \sqrt{\frac{\sum_n (w[n] - w_0[n])^2}{\sum_n (w_0[n])^2}} \tag{5.3}$$

- the Pearson correlation coefficient:

$$R = \frac{\sum_n (w[n] - \overline{w})\,(w_0[n] - \overline{w_0})}{\sqrt{\sum_n (w[n] - \overline{w})^2 \sum_n (w_0[n] - \overline{w_0})^2}} \tag{5.4}$$

The two indexes have been calculated for the shifted signals. The results are reported in Figure 5.14. For each station, the two values are shown graphically, with different color referring to different slip distribution signal.

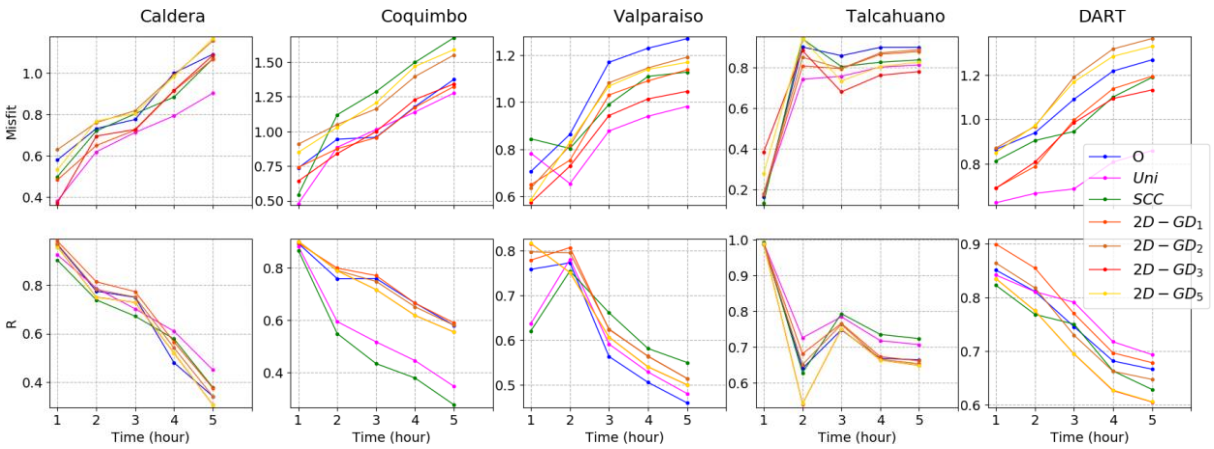

Figure 5.14 Misfits (top graphs) and correlation coefficients (bottom graphs) obtained for the different distributions. The values refer to the shifted waveforms.

The misfit values grow fast after the first 1-2 hours. The lower limits are still quite high (going below 0.20 only for the Talcahuano station).

The same applies to the correlation coefficient R, whose value decreases rapidly after the first main tsunami oscillations.

It is interesting to notice how the misfit of the reference model O (marked by the blue color) is never the lowest one. This peculiarity is confirmed by the trends of the correlation coefficient R: the reference model value is never the best one. The highest R values, for the first 2 hours of signal analysis, are always attained by the 2D GDs, with the exception of the Talcahuano station, for which all the slip distributions reproduce the first arrival well, as can be easily seen from Figure 5.11. With greater prominence from the R graphs, it can be seen how the values relative to the 2D GDs follow the trend of the reference model.

From these considerations, we can deduce that:

- all the models reproduce the first major oscillations of the tsunami;
- the reference model O is not so suitable to reproduce the observed signal;
- the 2D GDs are good substitutes for a heterogeneous slip model.

This last point suggests us to evaluate the comparison between the reference model O and the other slip distributions to better quantify the fit of the 2D Gaussian function.

Hence, we repeat the same procedures (time-shifting and computation of the indexes chosen for the comparison) but with respect to the signal produced by the reference model O, and not to the observed one. Table 5.2 summarizes the time shifts characterizing each slip distribution for every station. It is immediate to notice how, for the 2D Gaussian distributions the cross-correlation is already at its maximum value without the need for a temporal translation, with the exception of the Caldera station, whose shift is however only one minute (hence one time-step). Uniform and SCC distributions present, on the other hand, a zero-time-shift value just for the station of Valparaiso. The uniform source signals reached a 5-minute time shift value for the Coquimbo station, while the SCC waveforms present a 3-minute shift for the Caldera and Coquimbo stations.
Figure 5.15 shows the misfits and correlation coefficients for the shifted waveforms. These graphs show very well how the fit of the 2D Gaussians signals is far better than those of the uniform and of the SCC distributions. Regarding the 2D GDs, for the first three hours, the misfit is always lower than 0.40 and the correlation coefficient R is always larger than 0.92. Taking into consideration only the first hour of the signals, the 2D GDs misfit goes under 0.20 for the Caldera (min = 0.15), the Coquimbo (min = 0.11), the Valparaiso (min = 0.19) and the Talcahuano (min = 0.07) stations, while the misfit is always very close to 1. The difference shown vs. the uniform and SCC distributions is high: they are not as suitable to replace the heterogeneous slip reference model.

Table 5.2 Time shift values that maximize the cross-correlation between the synthetic signals produced by the slip distributions and the signal produced by the reference in the first hour of records for all the considered stations.

	Caldera (min)	Coquimbo (min)	Valparaiso (min)	Talcahuano (min)	DART (min)
Uni	2	5	0	-1	1
SCC	3	3	0	-1	1
2D-GD_1	1	0	0	0	0
2D-GD_2	1	0	0	0	0
2D-GD_3	1	0	0	0	0
2D-GD_5	1	0	0	0	0

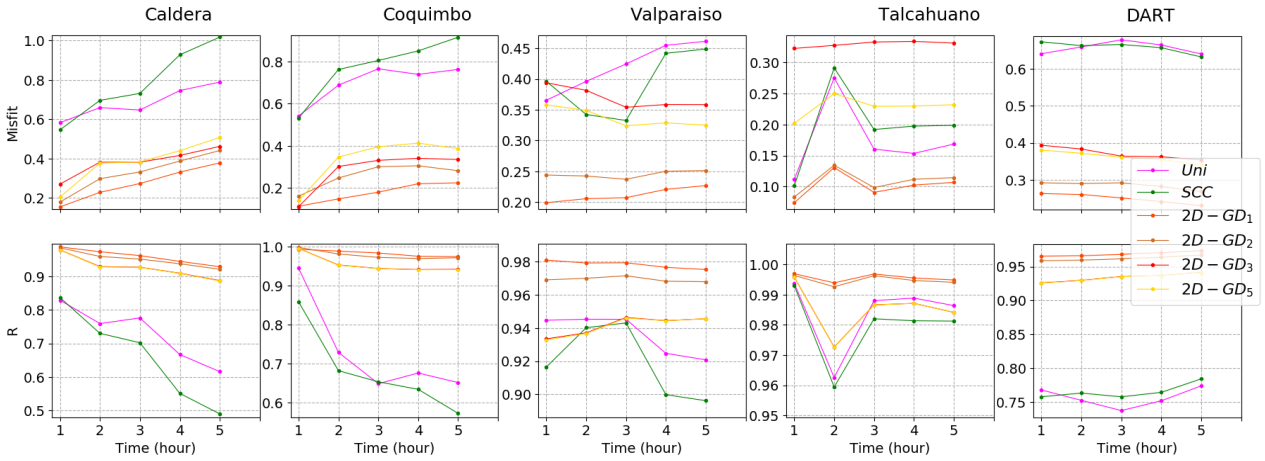

Figure 5.15 Misfits (top graphs) and correlation coefficients (bottom graphs) obtained for the different distributions with respect to the reference model O. The values refer to the shifted waveforms.

5.1.2 Water elevations and runup

For each distribution, the maximum water elevation along the 10-meter isobath has been also extracted. The results are plotted together with the experimental data, that refer to the maximum heights along the coast (taken from the NOAA website). The comparison is pure qualitative, due to the difference existing between the two measures. All the plots (Figures 5.16 - 5.17) present the water heights obtained for a particular distribution next to the water heights related to the reference case. The water height plots are placed next to the map with the northing aligned and with the same scale of the map. The observed runup heights are also reported for a qualitative comparison.

The 2D GD distribution (see Figure 5.16) fit better the maximum runup at the epicentre latitude, immediately followed by the reference SRCMOD model. SCC and uniform water heights tend, on the other hand, to underestimate the maxim runup measured in the coastal areas close to the epicentre. The focused profile behaviour of the Gaussian distribution could represent an important feature in the optics of tsunami early warning and of the estimation of the worst scenarios.

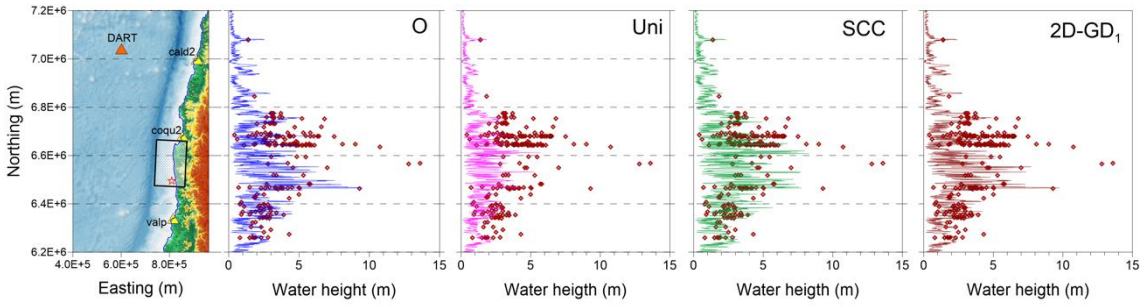

Figure 5.16 Maximum water heights along the 10-m isobath. The four plots at the right of the map refer to the reference case O (blue line), the uniform distribution (magenta line), the SCC distribution (green line) and the 2D GD$_1$ (red line). The observed runup heights are reported with red diamonds.

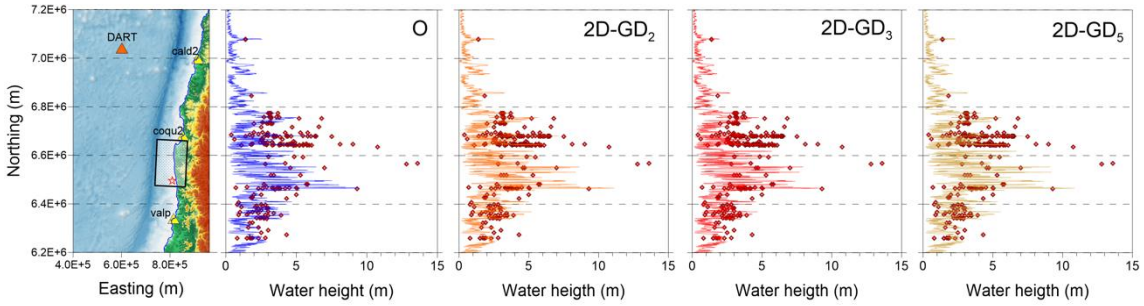

Figure 5.17 Maximum water heights along the 10-m isobath corresponding to the different sources: O, 2D GD$_2$, 2D GD$_3$, 2D GD$_5$.

All these results indicate the potential of the 2D GDs in reproducing the observations if the starting model is a faithful representation of the real seismic source. This is the reason that led us to repeat the above strategy,y but applied to another finite fault model.

5.2 The model by Heidarzadeh et al. (2016)

As already mentioned in the previous sections, the Illapel earthquake is well documented and several fault models have been obtained by different authors by means of different techniques. In inversion procedures, tsunami waveforms and tide gauge records can be used to study earthquake source processes, as reported by Satake and Kanamori (1991). Indeed, being the bathymetry better known than seismic velocity structure in the Earth, it is possible to accurately evaluate the tsunami propagation. Tsunami data can be used alone to determine the tsunami source or jointly with other kind of geophysical data, which allows one to enrich the knowledge concerning the original cause.

A model obtained by a joint inversion technique is the one by Heidarzadeh et al. (2016). They proposed a source model obtained using teleseismic and tsunami data.

In this section we are going to present our Gaussian distributions applied to their finite fault model, following the same procedure used for the Okuwaki et al. model.

The work by Heidarzadeh et al. was driven by two main reasons:

- furnishing a source model consistent with both seismic and tsunami data;
- investigate potential relationship between the 2015 Illapel and 2010 Maule earthquakes.

We are interested in the first item. The combination of teleseismc inversions and forward tsunami simulations permits one to get stable results in the time- (thanks to seismic data) and in the space-domain (thanks to tsunami data). This occurs because seismic waves travel much faster than tsunami waves (Satake, 1987). The data used for the inversions by Heidarzadeh et al. (2016) consisted of 62 teleseismic records and 33 tsunami records (the latter ones with a sampling interval of 1 min). They used a total number of 96 subfaults (12 along-strike, 8 along-dip) with dimension 20 km × 20 km. They conducted numerical simulations of tsunami propagation applying the numerical model by Satake (2005), based on linear shallow-water equations, using a single uniform grid. They also calculated the initial seafloor deformation through the Okada's formulas (1992).

The final slip model presents a large-slip area of 80 km (along strike) × 100 km (along dip) located 70 km to the northwest of the epicentre. This large asperity represents the 20% of the total rupture area (considering the non-zero slip subfaults) and the slip patch to be substituted by our 2D GDs. The slip map by Heidarzadeh et al. 2016 ("H" model hereafter) is reported in Figure 5.18(b). The rectangular plane reveals to be larger than the fault plane by Okuwaki et al., with a lower maximum slip value.

The H slip map presents different dip and rake values from subfault to subfault, and it has been used for the following analyses in its original form. In order to model it with our 2D GD distributions, however, a simplified model has been obtained, with the same number of subfaults, each presenting the same values of dip and rake (taken as the average value). Hence, the slip distributions considered for the analyses are the reference H model, the H simplified model (H_S), the uniform distribution, the SCC distribution and the five 2D GDs. As the modeling for the different 2D GDs provides similar outputs, in this section

we are going to present the graphs related only to the distribution 2D GD_3, which is fully sufficient to emphasize the difference against the other slip distributions.

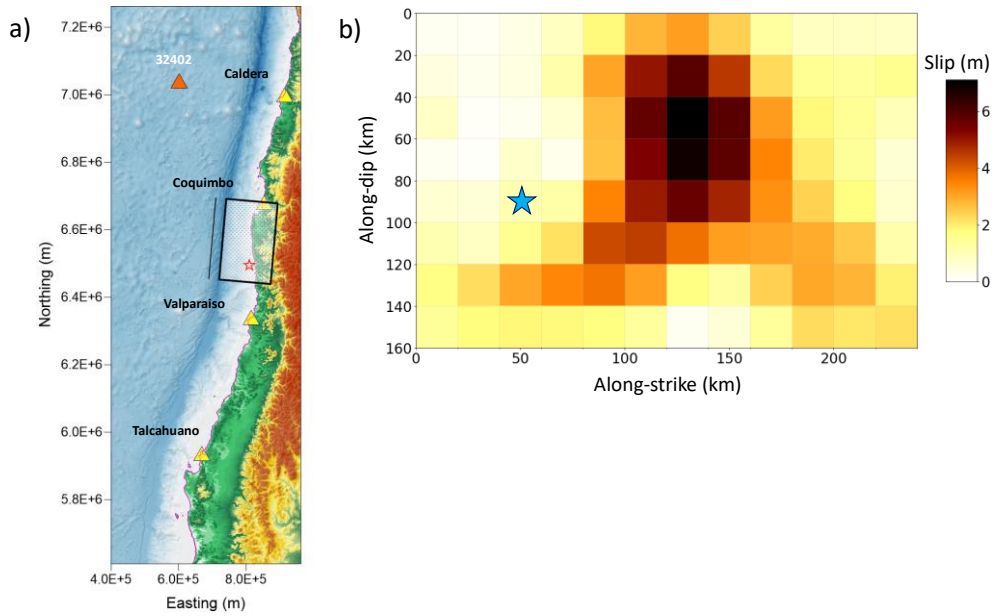

Figure 5.18 Illapel earthquake. Fault model by Heidarzadeh et al., 2016. (a) Fault plane along the Chilean coast. Star: epicentre, Yellow triangles: tide-gauge stations, Red triangle: offshore DART buoy (b) Slip map obtained by Heidarzadeh et al. (Star: hypocentre).

We calculated the vertical coseismic displacement fields induced by the cited slip distributions. Figure 5.19 shows these fields for the H, H_S, Uniform, SCC, 2D GD_3 distributions. The displacement patterns appear to be a little bit more extended than that produced by the Okuwaki et al. model, which is in agreement with what just said for the on-fault slip characterizing the two different models.

Considering instead the similarity between the vertical displacement induced by the 2D GD_3 and the one induced by the original model H, the shapes appear more different than what was found for the Okuwaki et al. model. This also holds for the Uniform and the SCC slip distributions. A more due judgement can be however given after the analysis of the waveforms.

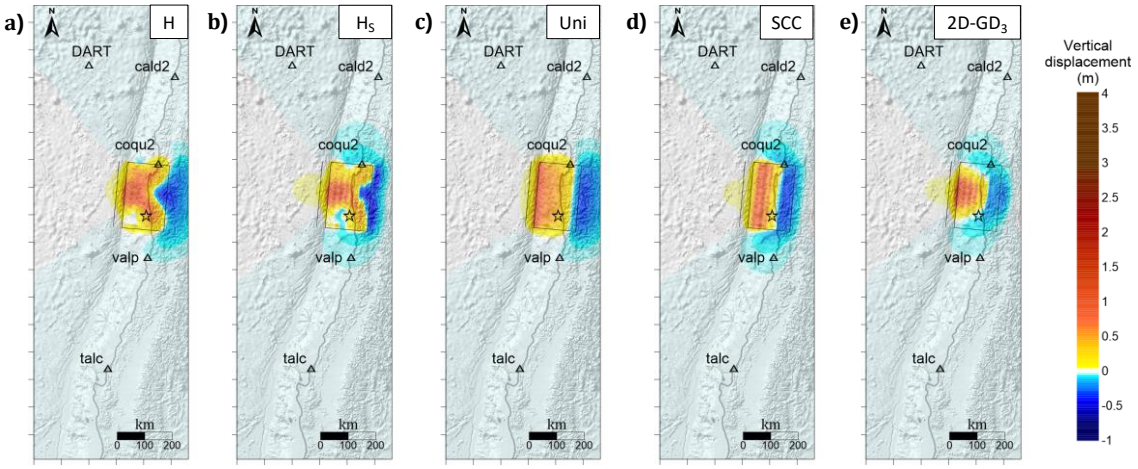

Figure 5.19 Vertical co-seismic displacement fields obtained by applying Okada's formulas to the slip distributions under study. (a): The reference model by Heidarzadeh et al. (H). (b), (c), (d), (e) the simplified model (H_S), Uniform (Uni), SCC and 2D GD_3 slip distributions, respectively. The five triangles in each map identify the four coastal stations and the DART buoy. The rectangle highlights the seismic fault. The white star is the epicentre.

The maximum and minimum water elevation fields have been extracted over the 10 hours of tsunami evolution. Figures 5.20 - 5.21 represent respectively the maxima and minima for the five slip distributions considered. Looking at the maxima the better similarity between the 2D GD and the Heidarzadeh models with respect to the Uniform and SCC distributions is evident. The graphs point out how the water elevations reach their highest value in the region north-west with to the epicentre, in accordance with the sea-surface vertical displacement behaviour, and along the coasts.

The minima do not highlight this difference probably due to the choice of the color scale, made to be consistent over all the graphs reported in this study (including models discussed in this and in the following Chapter).

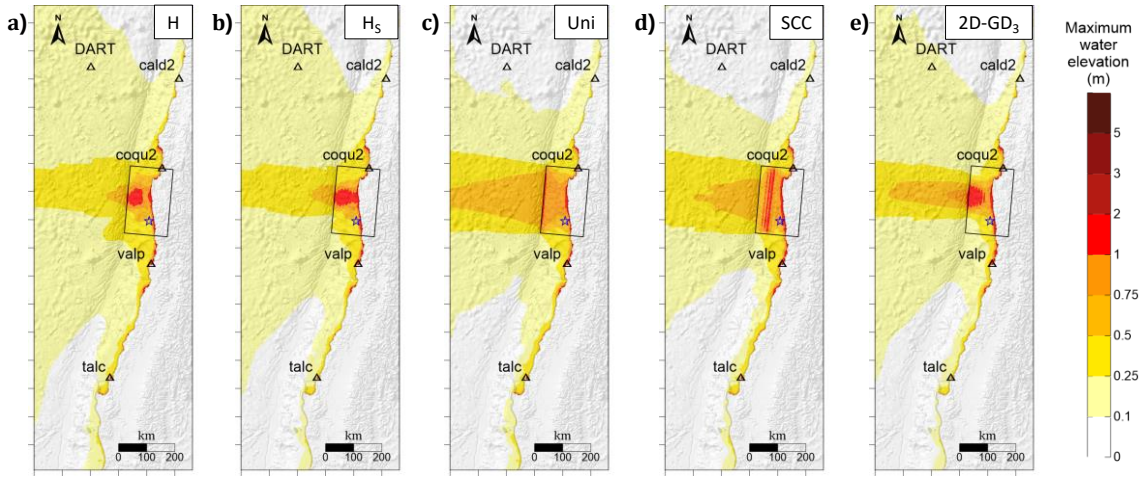

Figure 5.20 Maximum water elevation fields produced by the different sea-bottom displacements. H (a), H_S (b), Uniform (c), SCC (d), 2D GD_3 (e). For further details, see Figure 5.19.

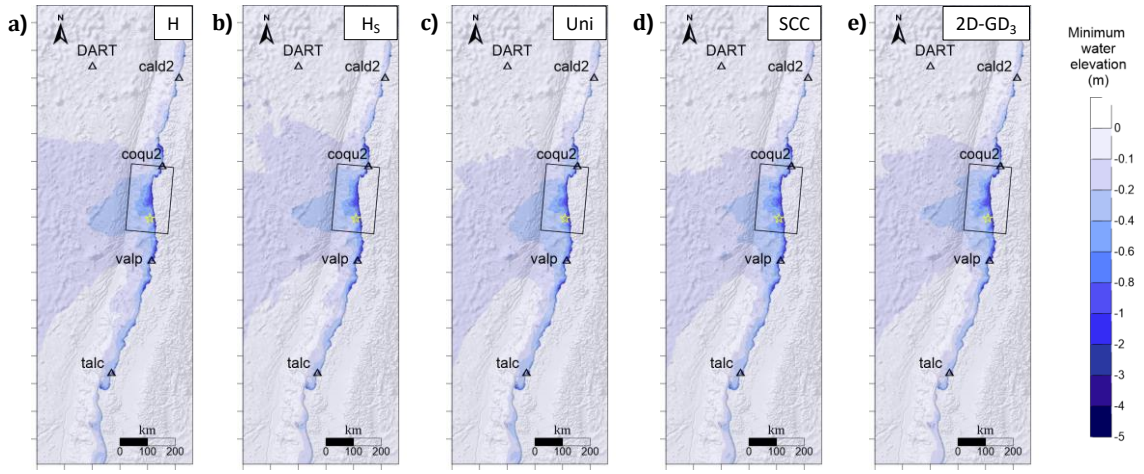

Figure 5.21 Minimum water elevation fields produced by the different sea-bottom displacements. H (a), H_S (b), Uniform (c), SCC (d), 2D-GD_3 (e). For further details, see Figure 5.19.

5.2.1 Analysis of the waveform signals (second FFM)

Tsunami waveforms have been calculated for the four stations of Caldera, Coquimbo, Valparaiso and Talcahuano and for the DART buoy station. Synthetic and observed (dashed line) records of the first 10 hours after the earthquake are reported in Figures 5.22 - 5.26. The waveforms show how the H, H_S and 2D GD_3 reproduce better the observed signal, especially for the first oscillations of the records. The discrepancy between the

Uniform and the SCC waveforms with the observed signals regarding the first tsunami arrival is particularly evident at the station of Caldera, Coquimbo and the DART buoy. While the Okuwaki et al. waveforms tended to overestimate the observed signals, especially for the stations of Caldera, Coquimbo and for the offshore buoy station, the signals derived from the Heidarzadeh et al. model turn out to be a little weaker and more faithful to the observed marigrams. This is not surprising considering that the H model is obtained also using tsunami waveform data.

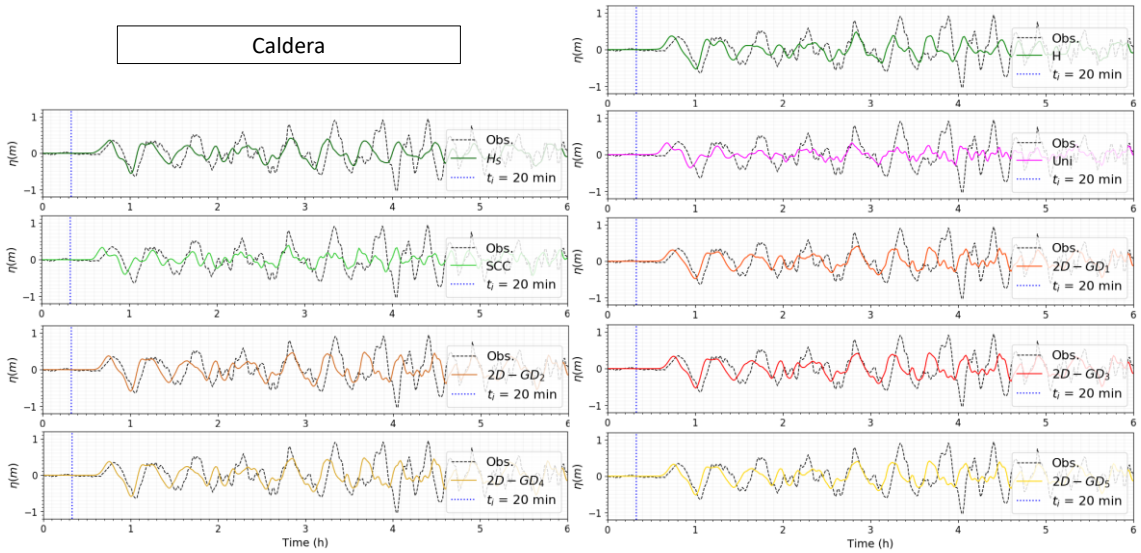

Figure 5.22 Marigrams calculated at the station of Caldera for (from top-right to bottom-right) the H, H_S, Uni, SCC, 2D GD_1, 2D GD_2, 2D GD_3, 2D GD_4, 2D GD_5 distributions. The black dashed signal refers to the observed one. The vertical blue dotted line indicates the starting time related to the station.

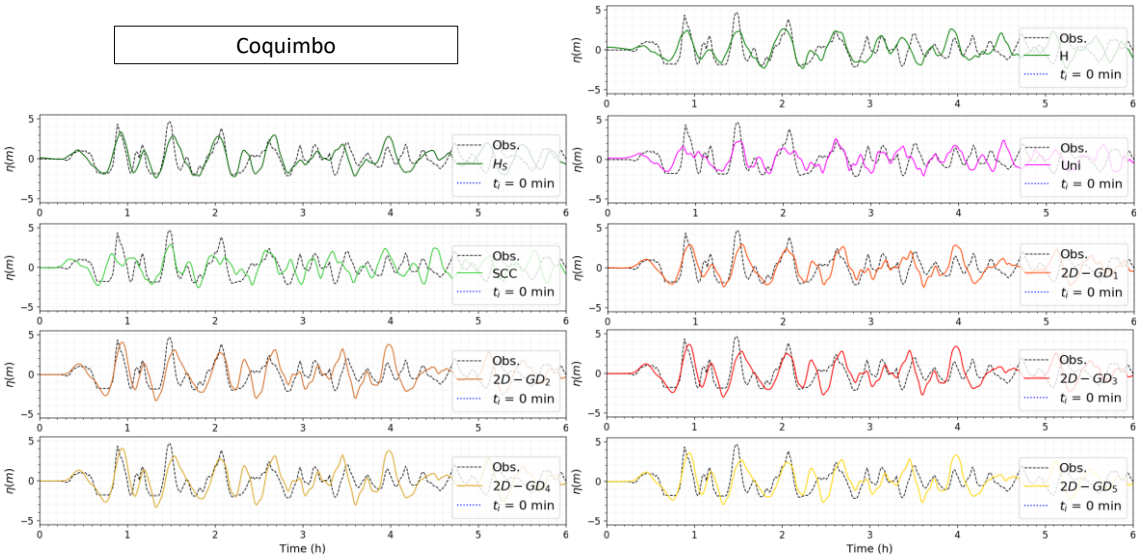

Figure 5.23 Marigrams obtained at the station of Coquimbo. For further details, see Figure 5.22.

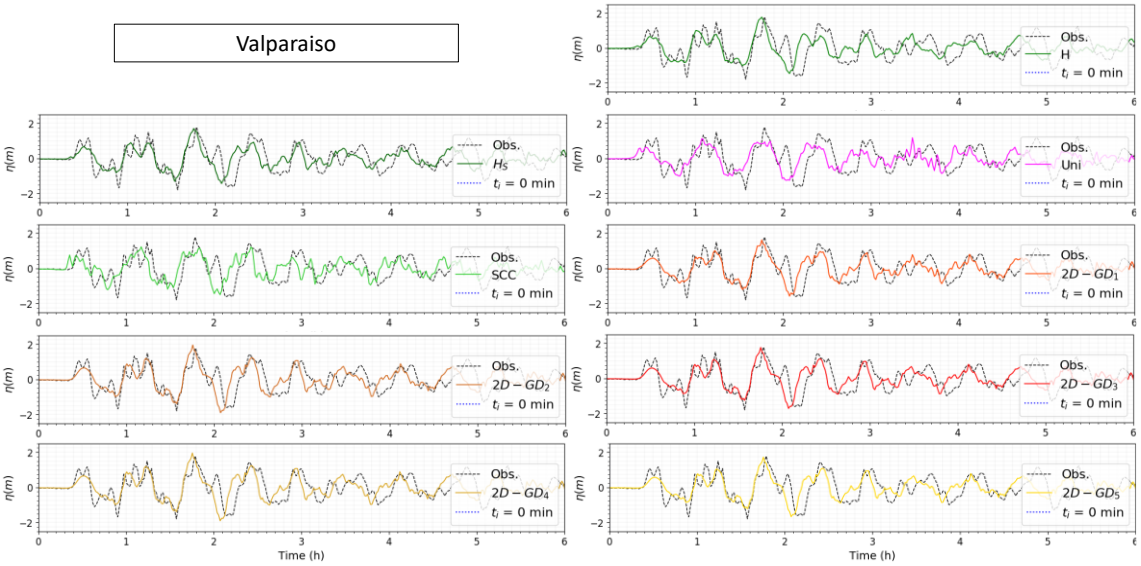

Figure 5.24 Marigrams obtained at the station of Valparaiso. For further details, see Figure 5.22.

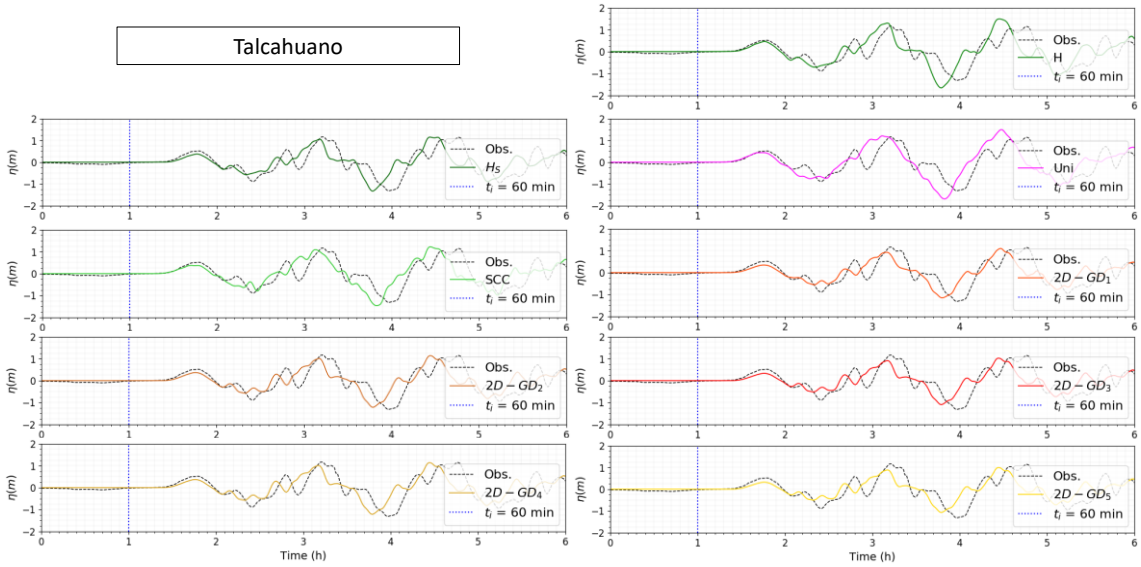

Figure 5.25 Marigrams obtained at the station of Talcahuano. For further details, see Figure 5.22.

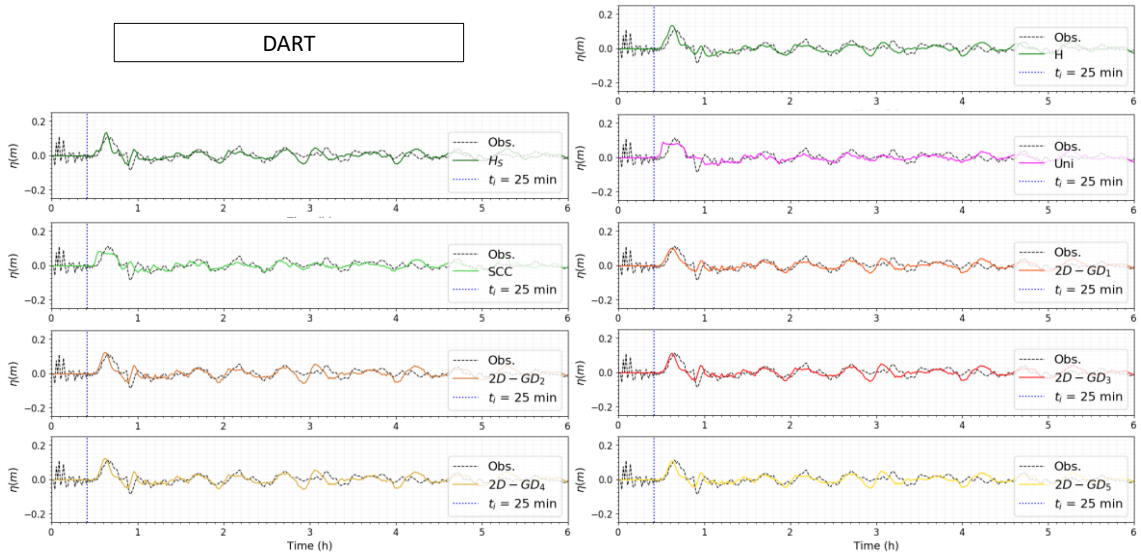

Figure 5.26 Marigrams obtained at the buoy DART station 32402. For further details, see Figure 5.22.

Synthetic signals, as in the previous case, do not exactly match the first arrival of the tsunami. This fact agrees with what noted by Heidarzadeh et al. 2016, according to whom a tsunami travel time difference of up to 5 minutes was reported for coastal tide gauges between simulations using coarse and using high-resolution bathymetry. After applying the the time shifting procedure on the first recording hour illustrated in the first sections of this Chapter, we obtain the results summarized in Table 5.3.

It is immediate to notice that the 2D GDs present the lowest values of time-shift, especially for the four coastal stations: the 2D GD distributions take the first oscillations in a much better way. For the Gaussian distributions, values of time shift never exceed the 3 minutes. The SCC and the Uniform slip distribution results to be the ones with the highest time shift values.

The three stations of Caldera, Coquimbo and Valparaiso present at least a 2-minute difference between the time shifts characterizing the Gaussians and those calculated for the Uniform and SCC. The situation is different for the Talcahuano station, for which the first hour is dominated by the first main oscillation, taken quite well by all the distributions.

For all the stations, except Coquimbo, the synthetic waveforms tend to anticipate the observed signal, in agreement with what said in the previous section. As already pointed out, we are not interested in tsunami source inversions, but in reproducing, as faithfully as possible, the observed waveforms.

Table 5.3 Time shift values that maximize the cross-correlation between the synthetic signals and the observed ones in the first hour of records for all the considered stations.

	Caldera (min)	Coquimbo (min)	Valparaiso (min)	Talcahuano (min)	DART (min)
H	4	2	0	1	3
H_S	3	0	0	1	3
Uni	5	4	5	2	3
SCC	5	5	5	2	4
2D GD_1	3	0	0	1	2
2D GD_2	3	-1	0	1	3
2D GD_3	3	-1	0	1	3
2D GD_4	3	-1	0	1	3
2D GD_5	3	-1	0	1	3

The following graphs (Figure 5.27) display the behaviour of the indices for the model by Heidarzadeh et al.. It is immediate to observe how the indices show values that are by far better than those obtained for the Okuwaki et al. model simulations.

The misfits emphasize the best behaviour of the H, H_S and 2D GDs in reproducing the first hours of the tsunami evolution. The same observation can be made for the correlation coefficient, for which the 2D GDs values are always larger than those characterizing the Uniform and the SCC distributions.

We stress once more that the better behaviour of the H model with respect to the O model is not surprising, since the former was obtained by inverting also the tsunami waveforms. Hence, reproducing the H source model with a heterogeneous slip distribution like the 2D GDs can lead us closer to reproducing better the observed data.

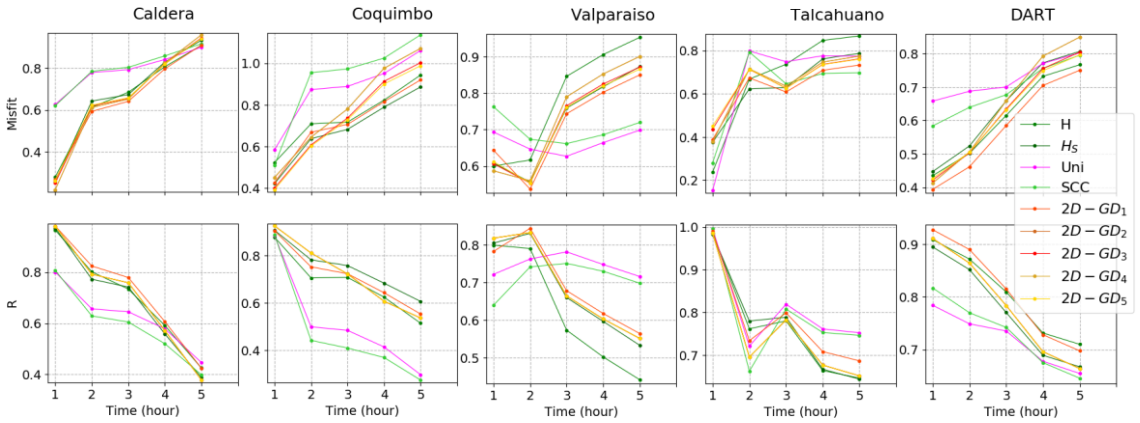

Figure 5.27 Misfits (top graphs) and correlation coefficients (bottom graphs) obtained for the different distributions. The values refer to the shifted waveforms.

We evaluate the misfits and correlation coefficients with respect to the reference model H. Table 5.4 displays the time shift values maximizing the cross-correlation between the signals produced by the different slip distributions and the signal produced by the reference model H. With the exception of the Coquimbo station, there is an accurate time alignment of the H_S and of 2D GDs waveforms with the reference H signal. Instead, the Uniform and the SCC waveforms always tend to anticipate the reference one (only for the SCC distribution at the Valparaiso station the time shift is zero).

Misfits and correlation coefficients of the shifted signals are reported in Figure 5.28. One can notice that the 2D GDs (and the H_S) misfit values are always lower than those obtained from the Uniform and SCC source models, with the exception of the first hour of the Coquimbo station (where the lowest misfit is the Uniform source one), and of the Talcahuano station, for which both the Uniform and the SCC distributions present the best misfits, but they lose the primacy already at the second hour of records. If one

considers the first 2-3 hours of records, the best behaviour of the 2D GDs waveforms is unquestionable. Therefore, we can conclude that the 2D Gaussian distributions are better substitutes for a heterogeneous slip pattern compared to the Uniform and the SCC slip distributions.

Table 5.4 Time shift values that maximize the cross-correlation between the synthetic signals produced by the different slip distributions and the signals produced by the reference one (H) in the first hour of records.

	Caldera (min)	Coquimbo (min)	Valparaiso (min)	Talcahuano (min)	DART (min)
H_S	0	1	0	0	0
Uni	2	2	2	1	1
SCC	2	3	0	1	1
2D GD_1	0	2	0	0	0
2D GD_2	0	2	0	0	0
2D GD_3	0	2	0	0	0
2D GD_4	0	2	0	0	0
2D GD_5	0	2	0	0	0

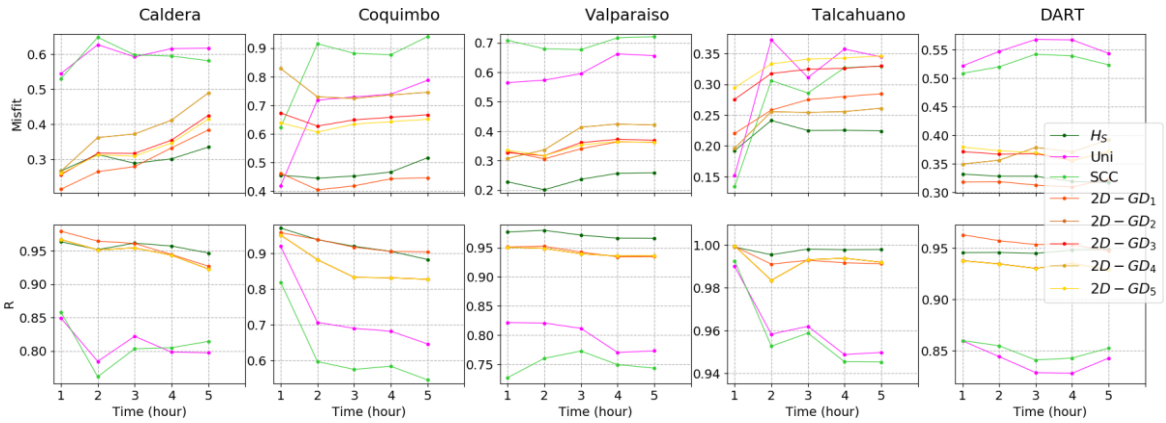

Figure 5.28 Misfits (top graphs) and correlation coefficients (bottom graphs) obtained for the different distributions with respect to the reference model H. The values refer to the shifted waveforms.

5.2.2 Water elevations and runup (second FFM)

In this sub-section we report the maximum water elevation along the 10-meter isobath. The results are plotted together with the experimental run-up data taken from the NOAA

website. For the Gaussian distributions, like for the coseismic and water elevation fields, only the 2D GD_3 results are shown.

One may observe that the maximum elevations derived from the H slip models are lower than those obtained from the O model. We note also the gap between the calculated elevations and the observed run-ups. However, the Gaussian function is the one that gives the highest results, especially in the epicentral area, marking a further advantage point in using it in terms of early warnings, a topic that will be addressed in the next Chapter.

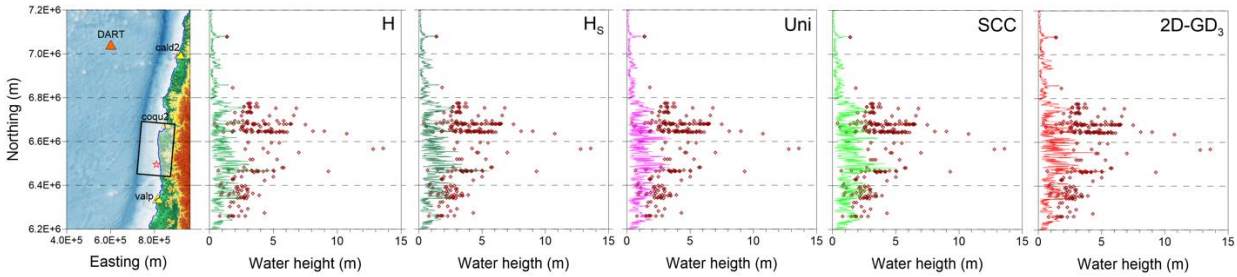

Figure 5.29 Maximum water heights along the 10-m isobath. The five plots at the right of the map refer to the reference case H (green line), the reference simplified H_S (dark green line), the Uniform distribution (magenta line), the SCC distribution (light green line) and the 2D GD_3 (red line). The observed runup heights are reported with red diamonds.

5.1 Conclusions

Starting from an application to a real case (the earthquake of Illapel, 16 September 2015) we derived the Uniform, the SCC, and the 2D GDs source slip models referring to two different finite fault model: the Okuwaki et al. model (2016), denoted as O model, and the Heidarzadeh et al. model (2016), designated by H model. Though it is not the focus of this analysis to find which one is better than the other, it is clear that since the earthquake caused a tsunami and the H model was obtained by inverting also near-field and far-field tsunami data, this has to be preferred especially if one is interested in tsunami modelling. Our point however was different. We have compared tsunamis produced by different slip distributions derived from the finite-fault models, and we have been able to show that for both cases (that is for the O and H models) the tsunamis corresponding to the heterogeneous, though simple, sources like the 2D GDs are preferable than the tsunamis computed from the Uniform and from the SCC sources in terms of relevant goodness indicators of the waveforms, such as the misfit and the correlation coefficients.

We will exploit this advantage to envisage a strategy for the tsunami early warning where 2D GD sources can be fast derived and utilised for quick tsunami predictions. This will be the topic of Chapter 6.

As a final consideration of this Chapter, we can add that we have found no significant difference in the waveforms resulting from the 5 2D Gaussian optimizations we used. This fact is reported in Appendix B.

6 The case of the Illapel earthquake from an early warning perspective

In the previous chapters, our analysis moved from an earthquake, for which one FFM (model O) was taken from the SRCMOD and one FFM (model H) was taken from the literature. From this information, we derived the Gaussian distributions best fitting the FFM slip heterogeneity where the best fit was measured in terms of the Earth's surface vertical co-seismic displacement field.

In this chapter, we change our perspective focusing on the early warning problem. When an earthquake occurs, the only information that becomes available after a few minutes concerns the location of the earthquake and its magnitude. The first finite-fault models, based on seismic/geodetic data inversion, become available much later, say several hours/days after the earthquake origin time. And in the case of tsunamigenic earthquakes, tsunami waveforms useful for inversion become available after the tsunami passage at the recording stations. From the warning perspective, usually, the time to get FFM representations is therefore too long for the near-source coastal areas, and sometimes even for the most distance coasts. We already know that slip heterogeneity influences significantly the distribution of tsunami run-ups, especially for near-field areas. Hence, if we knew the on-fault slip-distribution in quasi-real-time, we could build a more realistic tsunami scenario and activate a more accurate and focussed warning procedure.

We can take advantage of what we saw in the previous chapters to devise a strategy that allows us, once the hypocentre position and magnitude are known, to immediately derive an earthquake fault model where the heterogeneity is in the form of a Gaussian slip distribution and hence to provide very quickly a forecast of the tsunami maximum heights distribution pattern.

Throughout this chapter, we are going to present the main steps of the method and quantify its reliability using the indexes described in the previous chapter. In this thesis we test the strategy on the Illapel earthquake case in order to allow a direct comparison with the previously obtained results.

6.1 The strategy

Let us imagine we have to provide a tsunami warning and a real-time forecast of the expected tsunami effects immediately after the outbreak of the 16 September 2015 earthquake.

Realistically, the hypocentre location and the magnitude of the event with a good degree of approximation can be available within two-three minutes. Knowing the hypocentre location permits us to place the fault plane in a definite geographical reference, while the knowledge of magnitude allows us to derive the parameters characterizing the fault dimension, and those related to the slip distribution adopted (homogeneous, Gaussian, or other).

Actually, the mere knowledge of the hypocentre is not enough to correctly arrange the fault in the 3D space, since in principle the rupture can propagate in any direction, allowing infinite possible choices of the final extension of the fault rectangle with respect to the hypocentre position. Moreover, the information on magnitude alone cannot tell anything about the preferential slip mechanism, since this information is known only once the focal parameters (strike, dip, rake) have been obtained.

However, usually, focal parameters can be safely derived from seismological databases, knowing the tectonic region and the seismic history characterizing the interested area. In our case, considering the seismotectonic features of the source area, we assume that the Illapel shock was a reverse-type earthquake.

Here, we have taken, as starting experiment, the focal parameters provided by the USGS catalogue (https://earthquake.usgs.gov/earthquakes/eventpage/us20003k7a/moment-tensor): they are summarized in *Table 6.1*.

Table 6.1 Seismic parameters characterizing the Illapel earthquake, 16/09/2015.

Lat (°S)	Lon (°W)	Depth (m)	Mw	Strike (°)	Dip (°)	Rake (°)
31.573	71.674	22400	8.3	353	19	83

The on-fault slip distributions adopted in this chapter are:

1) A uniform slip distribution whose length L, width W and mean slip $\hat{u}$ are obtained through the empirical source-scaling laws by Thingbaijam et al. 2017 (hereafter called T model);

2) A uniform slip distribution whose length L, width W and mean slip $\hat{u}$ are obtained through the regression laws by Goda et al. 2016 (hereafter called G model);
3) A uniform slip distribution with L, W, and $\hat{u}$ obtained through our regression laws characterising "All" the events, reported in section 3.1 (model Uni_A);
4) A uniform slip distribution with L, W, and $\hat{u}$ obtained through our regression laws characterising the "Reverse" events, reported in section 3.1 (model Uni_R);
5) A 2D-GD_3 slip distribution, where L, W are obtained through the regression laws characterising "All" the events, reported in section 3.1 and the maximum slip u_{max}, the two standard deviations σ_1, σ_2 are derived from the regression laws characterising "All" the events, reported in section 4.5 (2D-GD_A);
6) A 2D-GD_3 slip distribution, where L, W are obtained through the regression laws characterising the "Reverse" events, reported in section 3.1 and the maximum slip u_{max}, the two standard deviations σ_1, σ_2 are derived from the regression laws characterising the "Reverse" events, reported in section 4.5 (2D-GD_R).

Notice that the Gaussian distributions of cases 5) and 6) can be derived without making recourse to any specific FFM, but using the regression laws that have been established from the SRCMOD dataset. The choice of the 2D GD_3 among the possible 2D GD*i* (*i*=1,2…5) has been made because it was proven to be a good compromised between the fitting performance and the number of free parameters.

Length and width have been rounded to get an integer number of sub-faults of size 20 km × 20 km. The characteristics of each on-fault slip distributions are reported in Table 6.2 (for uniform slip distributions) and Table 6.3 (for 2D GD_3 slip distributions).

Table 6.2 Number of subfaults, rupture dimensions and mean slip values for uniform slip distributions.

	N_{strike}	N_{dip}	Length (km)	Width (km)	û (m)
T	13	4	260	80	3.87
G	12	6	240	120	2.32
UNI_A	16	6	320	120	1.88
UNI_R	19	8	380	160	1.20

Table 6.3 Number of subfaults, rupture dimensions and Gaussian parameters for the 2D GD distributions.

	N_{strike}	N_{dip}	Length (km)	Width (km)	u_{max} (m)	σ_1 (km)	σ_2 (km)
2D-GD_A	19	6	320	120	7.89	48.97	26.06
2D-GD_R	12	6	380	160	5.19	61.80	35.56

From Table 6.2 it is evident how our scaling laws tend to overestimate the fault dimensions, mostly the length, especially in the case of reverse-faulting regression. On the other hand, they provide a lower value of average slip. This result is probably due to the way the fault patterns have been cleaned up (as reported in section 2.3), where only zero-slip sub-faults rows/columns on the edge have been deleted. However, this fact does not represent a particular problem for our purposes, due to the peaked profile of the function characterizing the 2D GD with respect to a homogeneous model. Indeed, the inclusion of our uniform models is more for completeness than an actual novelty, since our attention is mainly focused on Gaussian distribution models and their direct application.

6.1.1 The relative position between the fault plane and the hypocentre

As observed in section 3.2.1, the fault plane is not necessarily centered at the earthquake hypocentre. To take this uncertainty into account, we decided, as a first approach, to consider not one but three faults for each of the distributions mentioned above.

The three fault planes are:

1- a plane centered on the hypocentre, hereafter called “mid” source;
2- a fault shifted northwards, having the hypocentre in the central point of the southern edge of the rectangular plane, hereafter called “north” source;
3- a fault shifted southwards, having the hypocentre in the central point of the northern edge of the rectangular plane, hereafter called “south” source.

To provide a more complete scenario it might be useful to consider other faults of intermediate position and, perhaps, of different depth. In this chapter we have just considered the faults with the maximum possible shift where the hypocentre is moved to the maximum distance from the fault centre.

Figure 6.1 shows the arrangement of the four mid faults derived through the different regressions in the geographical reference system. North and South faults are reported only for the source derived from the Thingbaijam et al. (2017) scaling laws, to illustrate the distinction between North, Mid and South faults.

The shown source rectangles are only four because for the Uniform and Gaussian distributions obtained from our regression laws that is the values of Lengths and Width are the same: more specifically cases 3) and 5) provide the values inferred from the regression of type "All" and cases 4) and 6) provide the values obtained from the regressions of type "Reverse".

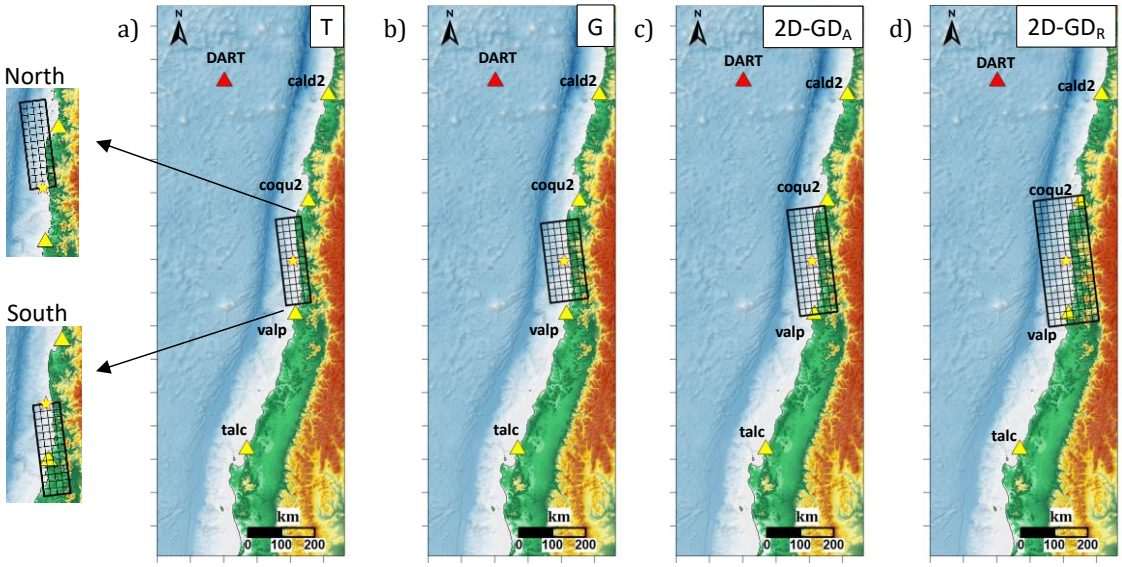

Figure 6.1 Mid fault planes obtained from regressions by Thingbaijam et al. 2017 (a), Goda et al. 2016 (b), our "All" case (c), our "Reverse" case (d). North (top left) and South (bottom left) fault planes for the seismic source obtained from regressions by Thingbaijam et al. (2017) are reported in the leftmost panels. The yellow star represents the epicentre location. Yellow triangles indicate the four coastal stations and the red triangle the DART buoy station.

6.2 The tsunami simulations

Once defined the seismic sources, we can calculate the initial conditions for the tsunami propagation, which is the vertical co-seismic displacement fields induced by the different slip distributions.

The co-seismic fields are obtained thanks to the Okada's formulas (1992) and their representations for all of the Mid faults are reported in Figure 6.2.

Fields obtained for the North and South faults are not shown because they have, more or less, the same structure as those obtained from the Mid faults, simply translated along the strike direction. Among the uniform models, the one showing the most concentrated and intense vertical displacement is the T model, followed by the G model and then our Uni_A. The Uni_R, instead, shows a more extended and less peaked pattern. The fields derived from the Gaussian distributions are, as expected, the most focused.

The differences with the fields obtained from the two FFMs in the previous Chapter is unavoidable. The FFMs were both characterized by fault plane extended primarily north to the hypocentre, while the Mid faults are centered within the nucleation point. Furthermore, the difference in strike values inevitably alters the directivity of the field.

To be consistent with the previous chapter, the maxima and minima water elevation fields on the first 10 hours of tsunami simulation were also extracted (Figures 6.3- 6.4). These are more similar to the Heidarzadeh et al. model than to the Okuwaki et al. one as far as the magnitude of the water levels is concerned. The Gaussian models exhibit a more focused profile, while uniform models show the tendency to over-flatten the elevation field.

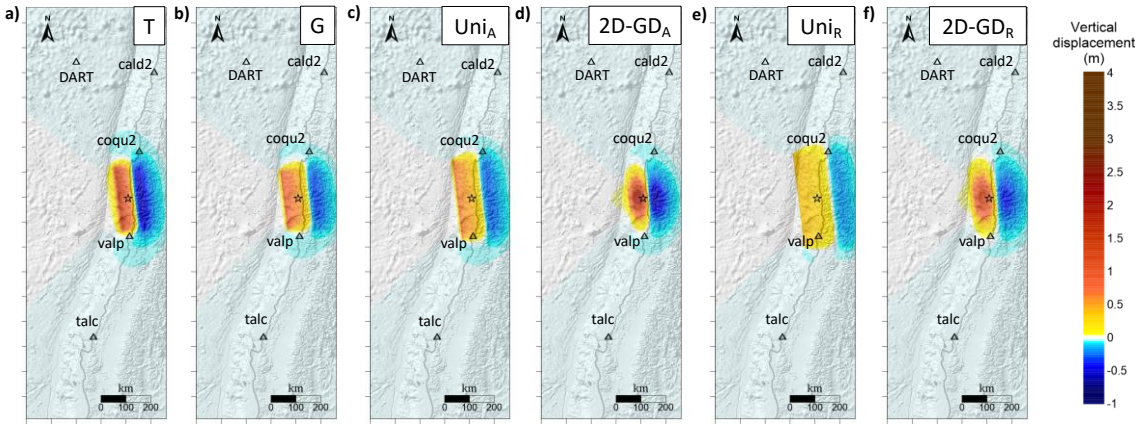

Figure 6.2 Vertical co-seismic displacement fields induced by the Mid faults of the slip distributions considered. (a) Thingbaijam et al. (2017) regression, (b) Goda et al. (2016) regression, (c) UNI_A, (d) 2D-GD_A, (e) UNI_R, (f) 2D-GD_R.

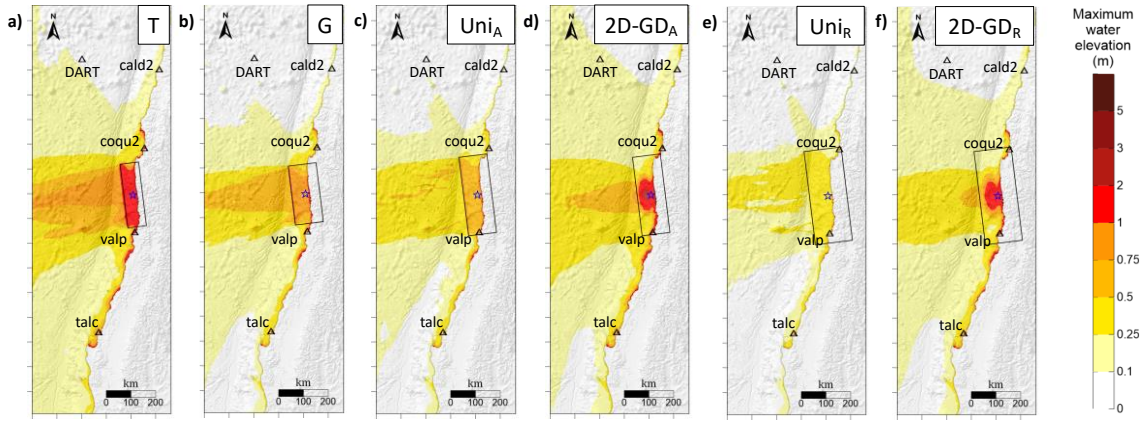

Figure 6.3 Maxima water elevation fields induced by the Mid faults of the slip distributions considered. See Figure 6.2 for further details.

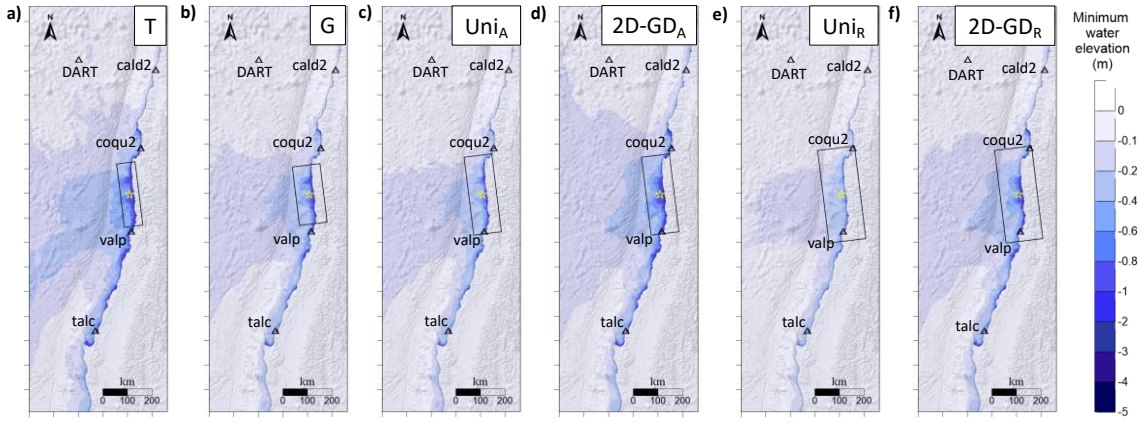

Figure 6.4 Minima water elevation fields induced by the Mid faults of the slip distributions considered. See Figure 6.2 for the further details.

6.2.1 Analysis of the waveform signals

To quantify the goodness of these models, we obtained from magnitude and hypocentre coordinates, we proceed according to the method illustrated in Chapter 5. Tables 6.4, 6.5, 6.6 summarize the time shifts for the mid, north and south faults respectively.

The mid faults present time-shift values within 8 minutes, with the exception of the Valparaiso station, for which there is a significant time difference between the calculated first arrival and the registered one. Regarding the uniform slip distribution, the calculated signals anticipate the observed one by a time interval that goes from 14 (G model) to 25

(Uni_R) minutes. The Gaussian distributions, on the other hand, maintain a time anticipation not exceeding 7 minutes.

Quite different are the time-shift values for the faults translated with respect to the hypocentre. For the north faults, the synthetic arrivals at the Caldera station anticipate the observed signal, a fact which is not surprising since this station is the northernmost of those considered. Also for the station of Coquimbo and the DART 32402 all synthetic arrivals have to be shifted forward, but by a smaller amount than those calculated at Caldera. For the station of Valparaiso, the synthetic waveforms have to be anticipated to better correlate with the first observed tsunami oscillations, with larger shifts for the Gaussian functions.

The south faults, as expected, present the opposite behaviour. The synthetic waveforms arrive with delay at the northern coastal station of Caldera and Coquimbo and at the DART buoy station. Instead, they arrive earlier at the southern stations of Valparaiso and Talcahuano.

Table 6.4 Time shift values that maximize the cross-correlation between the synthetic signal produced by the mid faults slip distributions and the observed signal on 1 hour of records for all the considered stations.

Mid faults	Caldera (min)	Coquimbo (min)	Valparaiso (min)	Talcahuano (min)	DART (min)
T	1	-2	16	6	-4
G	-1	-3	14	7	-5
Uni_A	4	1	18	7	2
Uni_R	7	4	25	8	2
$2D\text{-}GD_A$	-5	-6	6	6	-5
$2D\text{-}GD_R$	-4	-5	7	6	-5

Table 6.5 Time shift values that maximize the cross-correlation between the synthetic signal produced by the north faults slip distributions and the observed signal on 1 hour of records for all the considered stations.

North faults	Caldera (min)	Coquimbo (min)	Valparaiso (min)	Talcahuano (min)	DART (min)
T	13	4	-1	1	4
G	11	5	0	1	7
Uni_A	17	4	0	-3	4
Uni_R	19	3	0	-3	4
$2D\text{-}GD_A$	11	4	-9	-4	7
$2D\text{-}GD_R$	14	4	-10	-5	11

Table 6.6 Time shift values that maximize the cross-correlation between the synthetic signal produced by the south faults slip distributions and the observed signal on 1 hour of records for all the considered stations.

South faults	Caldera (min)	Coquimbo (min)	Valparaiso (min)	Talcahuano (min)	DART (min)
T	-12	-13	21	12	-15
G	-12	-13	20	13	-15
Uni_A	-12	-13	20	13	-15
Uni_R	-12	-13	20	15	-15
$2D\text{-}GD_A$	-19	-20	24	13	-23
$2D\text{-}GD_R$	-20	-21	24	15	-25

As we have considered the extreme cases of fault translation, with the intention of underlining the problem deriving from the uncertainty in the relative position between the hypocentre and the fault plane, it is reasonable that the results in terms of misfits and correlation coefficients are not significant. Moreover, in the configuration adopted for this particular case, the south faults result to be, for a significant portion, in-land. The same reasoning holds for the co-seismic field induced by the latter faults: and this alters the propagation of the tsunami, since its initial condition (given by the off-shore/coastal vertical displacement field) is reduced.

For this reason, regarding the comparison between the synthetic waveforms and the observed ones, we focus on the results obtained only for the mid faults, that are shown in

Figure 6.5. The misfits of the first hour, with the exception of Coquimbo and Valparaiso stations, show the lowest values for the Gaussian distributions. However, for the Valparaiso station it is necessary to consider the fact that uniform distributions have been translated by a time that is not at all negligible, at least twice the shift value of the Gaussians: hence, the behaviour of the uniform indexes for the Valparaiso station should be treated carefully, since they derive from signals that fail to reach the first arrival by more than a quarter of an hour.

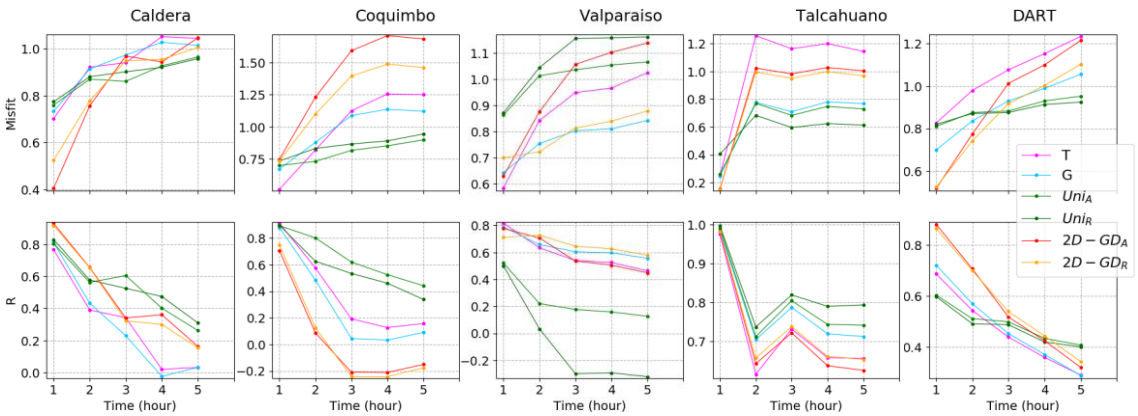

Figure 6.5 Misfits (top graphs) and correlation coefficients (bottom graphs) obtained for the different slip distributions on the mid faults. The values refer to the shifted waveforms.

6.2.1 Water elevations and runup

Concerning the maximum water elevation along the 10-meter isobath, we decided to show the results for all three source configurations examined. Figures 6.6 - 6.7 - 6.8 show the water elevations obtained for the north, mid, south faults respectively. Even if the comparison with the runup values (small red diamonds in figures) is purely qualitative, it is possible to draw interesting inferences.

Starting from the mid faults, the best behaviour in reproducing the experimental data in the epicentral zone is given by the 2D GD_A. This is due to the peculiarity of the Gaussian distributions that naturally represent the main asperity rather than the entire fault plane. And the main asperity, as we have pointed out several times, has the greatest influence on the co-seismic field and on the distribution of heights along the coast.

Among the homogeneous slip model, the T model is the one that produces the greatest heights, due to the fact that it is characterized by the greater average on-fault slip. For the same reason the Uni_A and the Uni_R are those exhibiting the lowest heights.

The same comments can be repeated for north faults. In this case, the privileged area is the one at the Coquimbo latitude, where, once again, the Gaussian functions reproduce better the trend of the run-ups.

For the south faults, all the results appear smaller. This because, in such configuration, the non-negligible onshore components of the co-seismic deformations greatly reduce the initial tsunami source area.

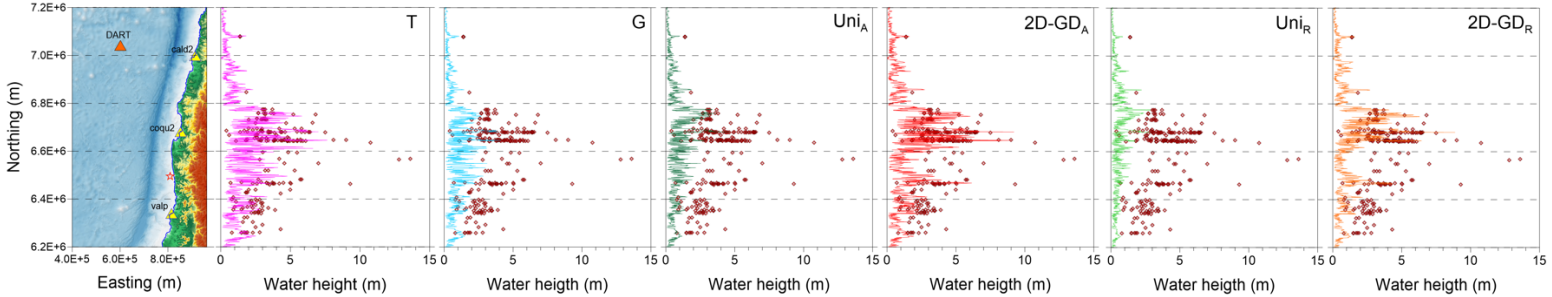

*Figure 6.6 Maximum water heights along the 10-m isobath for the **north faults**.*

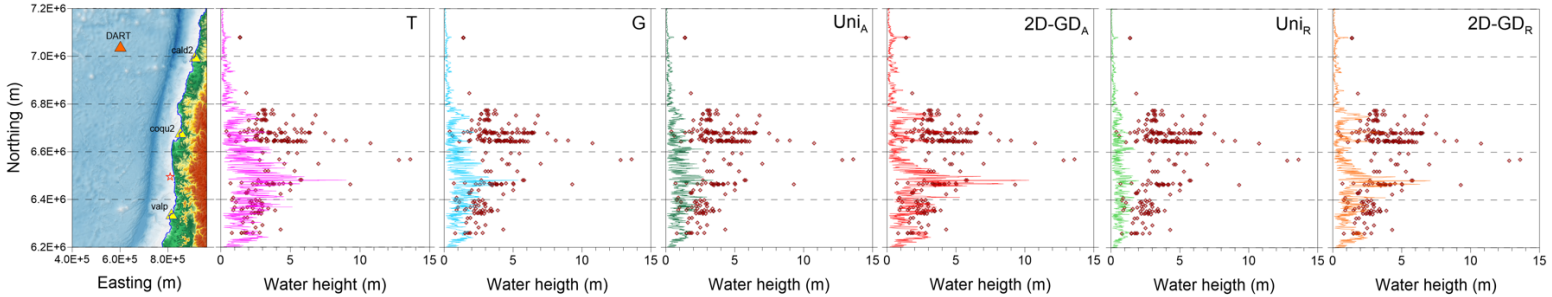

*Figure 6.7 Maximum water heights along the 10-m isobath for the **mid faults**.*

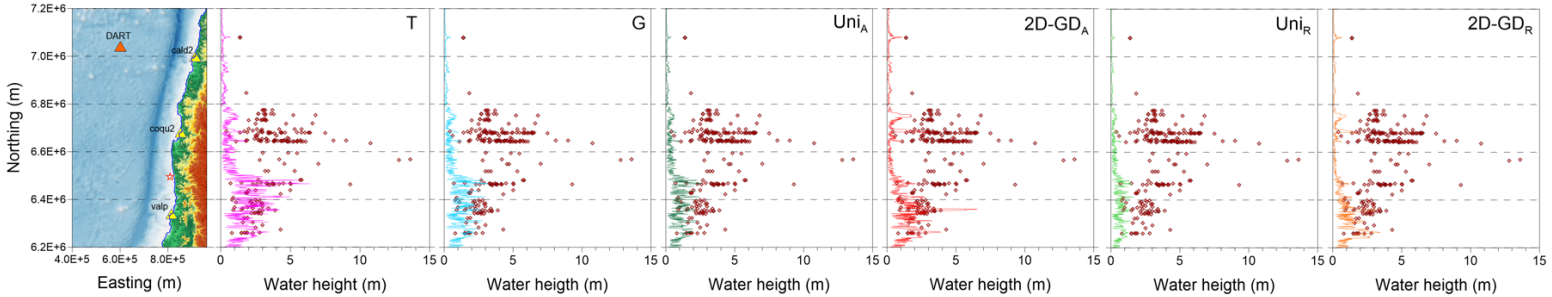

*Figure 6.8 Maximum water heights along the 10-m isobath for the **south faults**.*

6.1 The final comparison

As a final step in this chapter, it is our intention to compare the Gaussian distribution slip patterns obtained directly from magnitude and the finite fault models that derive from inversions of seismic and tsunami data.

We summarize in this final section the waveforms analysis results for:

- The Okuwaki et al. model: [O].
- The 2D GD_3 obtained from [O]: [2D GD_O].
- The Heidarzadeh et al. model: [H].
- The 2D GD_3 obtained from [H]: [2D GD_H].
- The 2D GD_A.
- The 2D GD_R.

The next list of figures (6.9 - 6.13) displays, for each station the trend of the misfits and correlation coefficients (a) and the tsunami waveforms calculated for the first 5 hours (b). All the graphs refer to the shifted synthetic signals. The misfits and the correlation coefficients are calculated with respect to the observed waveforms.

It is possible to deduce several interesting conclusions from the graphs.

First of all, the direct comparison between the two FFMs (O and H) highlights the best behaviour of the Heidarzadeh model and, consequently, of the Gaussian function obtained from it.

Moreover, and this is the most important point for our results, the 2D GDs obtained from magnitude regressions present misfit values that are always lower or equal (in Coquimbo station), at least for the first recording hour, than those referring to the O model.

The comparison over several hours of recording becomes in any case less significant, since the values of misfit rise dramatically for all distributions. Furthermore, with a view to early warning, it is precisely the first series of oscillations that interest most. With reference to this, it is possible to draw some observations, looking directly at the tsunami waveforms and focusing on the trends of the H and 2D $GD_{A/R}$ models. Indeed, despite the indices prefer the FFM, the synthetic waveforms (b) of the 2D $GD_{A/R}$ turn out to reproduce the first oscillation sufficiently well. Furthermore, considering as an example the station of Talcahuano (Figure 6.12 b), even if the time of the second main wave is not

perfectly coincident with the observed one (which also applies to the other models), the shape and magnitude of the oscillation is very satisfactorily respected.

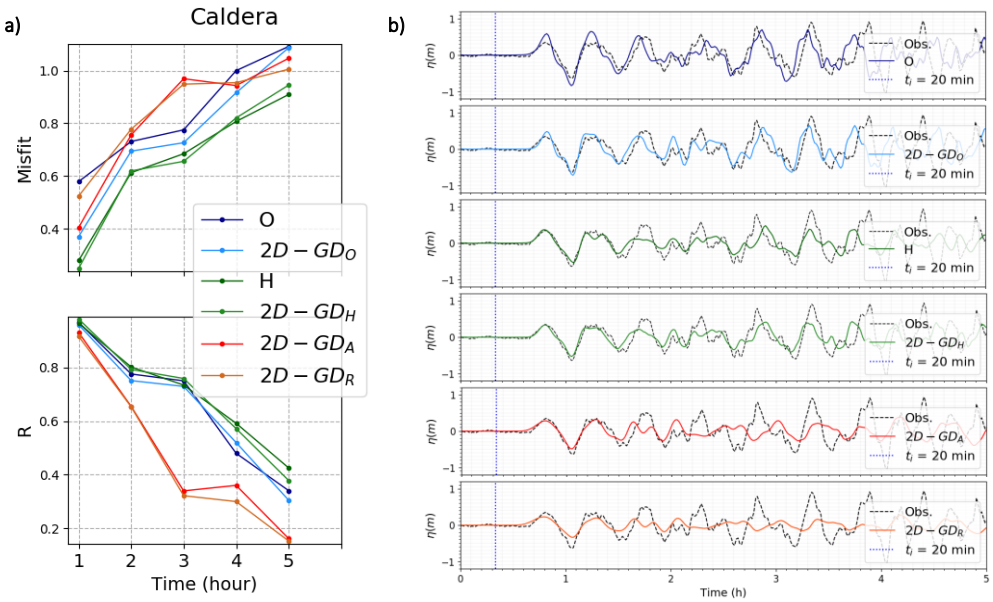

Figure 6.9 a):Misfits (top graph) and correlation coefficients (bottom graph) calculated for the different shifted waveforms; b): Marigrams (with shifted synthetic waveforms) obtained at the station of Caldera. The black dashed line refers to the observed signal. The vertical blue dotted line indicates the starting time related to the station.

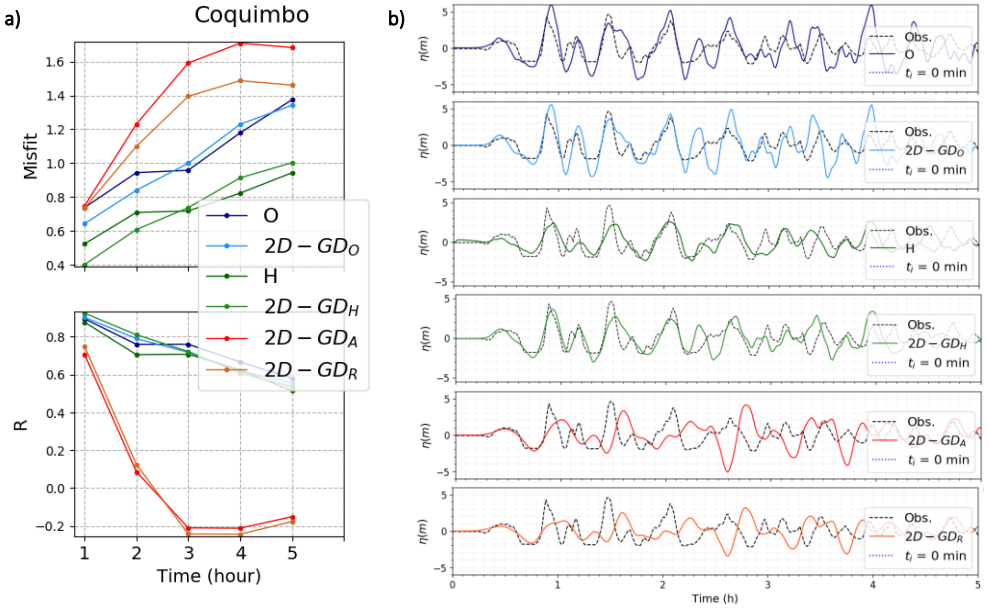

Figure 6.10 a):Misfits (top graph) and correlation coefficients (bottom graph); b): Marigrams obtained at the station of Coquimbo. For further details, see Figure 6.9.

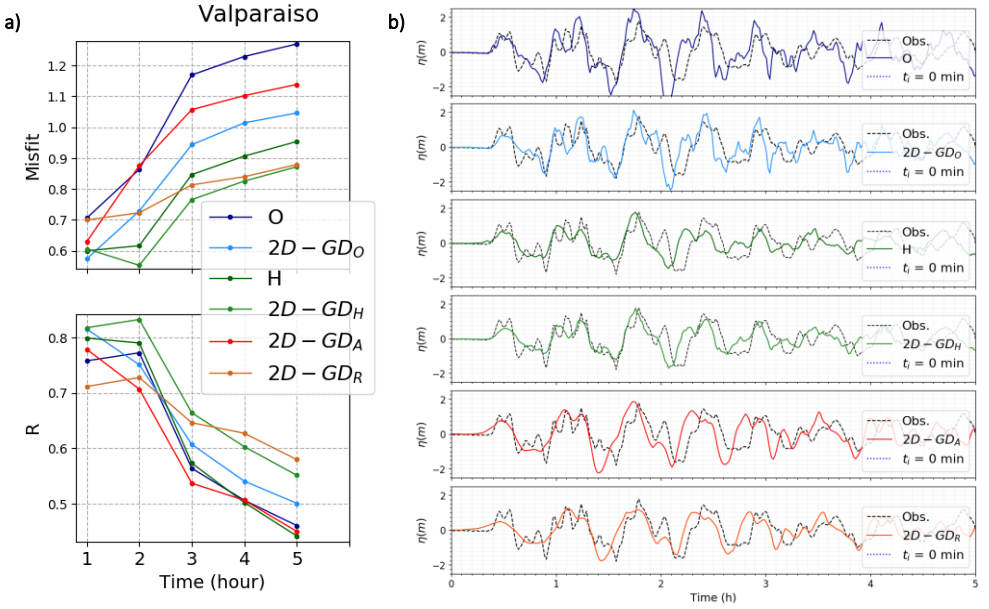

Figure 6.11 a):Misfits (top graph) and correlation coefficients (bottom graph); b): Marigrams obtained at the station of Valparaiso. For further details, see Figure 6.9.

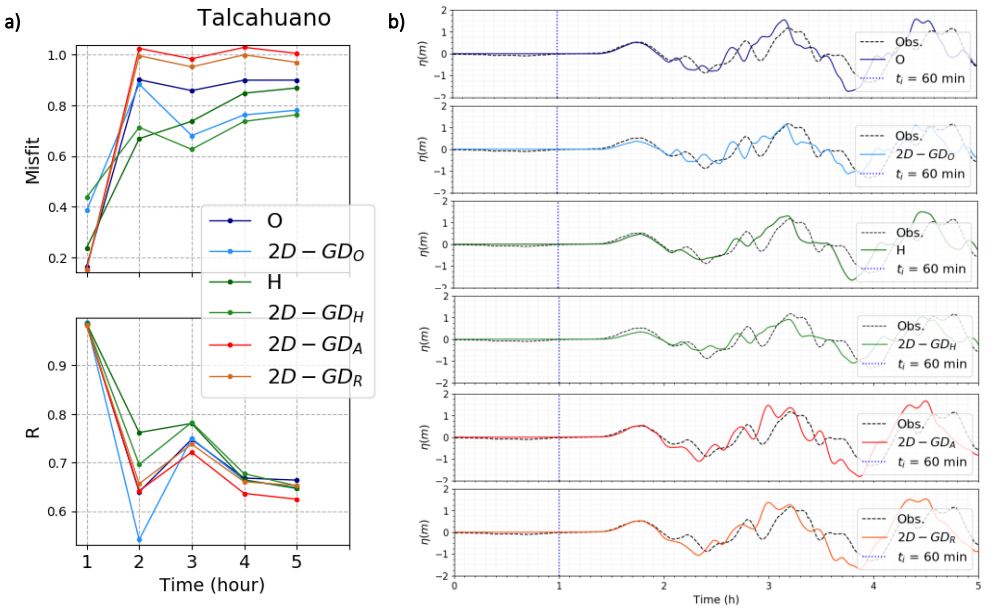

Figure 6.12 a):Misfits (top graph) and correlation coefficients (bottom graph); b): Marigrams obtained at the station of Talcahuano. For further details, see Figure 6.9.

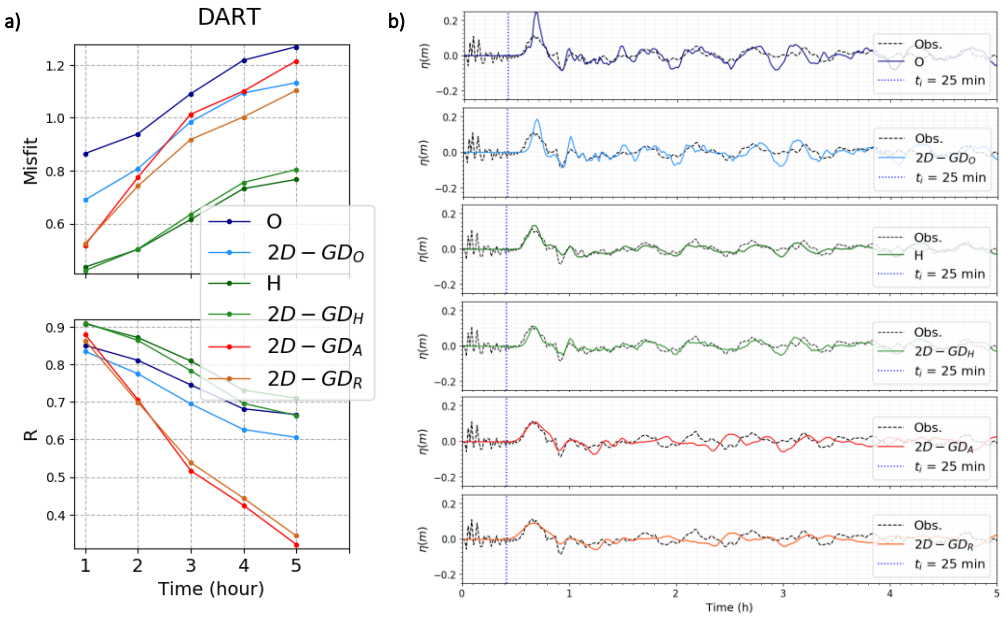

Figure 6.13 a):Misfits (top graph) and correlation coefficients (bottom graph); b): Marigrams obtained at the buoy DART station 32402. For further details, see Figure 6.9.

6.1.1 Timing considerations and early warning implications

The marigrams listed above show a good agreement both by the FFM H and by the Gaussians derived from the hypocentre location and magnitude. However, as quantified by the misfit and the correlation coefficients, the faithfulness of the model H model to the observed data is greater.

Nevertheless, there is an important and fundamental distinction between the two distributions. The H model was obtained a posteriori, by inversion of seismic data and tsunami waveforms. The 2D $GD_{A/R}$ are obtained directly from the magnitude and location of the earthquake, which is a much simpler and quicker procedure.

Let us try to make a rough estimate of the time-line involved in these processes, starting from the computation of the source model H by Heidarzadeh et al. 2016. The model is contained in a manuscript received by the Geophysical Research Letters on the 4 December 2015, and then published on the 23 January 2016. The time elapsed since the outbreak of the earthquake to the day on which the work was presented is therefore approximately 3 months. Potentially, a first version of this FFM obtained by also exploiting tsunami data could be obtained several hours after the tsunami has impacted the coasts (that is as soon as the tide-gauge data were available).

If instead we consider a source model that can be derived directly from the magnitude (like the 2D GDs), then the time to obtain them is comparatively extremely lower. Taking the case under study, since it was a great earthquake there were several warnings launched by the warning operational centers. For example, let's consider the bulletins launched by the Pacific Tsunami Warning Centre (PTWC, https://ptwc.weather.gov/ptwc).

The time the earthquake originated was 22:54:32 (UTC). At 23.01 (7 min after the main shock) the first tsunami threat message was issued from the PTWC. The message contained the following earthquake information:

- MAGNITUDE 7.9;
- ORIGIN TIME 22:55 UTC SEP 16 2015;
- COORDINATES 31.5 SOUTH, 71.9 WEST;
- DEPTH 33 KM / 20 MILES;
- LOCATION NEAR THE COAST OF CENTRAL CHILE.

The message contained also the estimated times of arrival (eta) of the initial tsunami wave for places with a potential tsunami threat as well as for other places having an eta within the following six hours.

At 23.23 (29 minutes after the main shock), a second threat message was issued, with some updates and a new value of the magnitude (8.3).

The issuing time of the first alert bulletin can be taken as the upper limit for the time necessary to know magnitude and hypocentre position, that are thus potentially estimated, in the first instance, only a few minutes after the event.

This is a topic that makes the 2D Gaussian models extremely interesting for early warning. It means that potentially, a few minutes after the main shock of a dangerous earthquake it is possible to get a first asperity model of the event.

The first arrival of the tsunami at the considered stations was of about 40, 20, 24, 90 minutes for Caldera, Coquimbo, Valparaiso, Talcahuano respectively and about 30 minutes for the buoy DART 32402.

This means that the availability in a few minutes of the finite heterogeneous slip model can be sufficiently prompt. To make tsunami predictions even faster one can pre-compute tsunami Green's functions, which can result in launching alerts in a sufficiently short time and with details better than a less realistic uniform model. As reported by Ohta et al., 2012, using pre-computed tsunami Green's functions, tsunami simulations can be accomplished in approximately 1 min. Realistically, adding the tsunami computation time to the time needed to get the relevant earthquake information on the magnitude and location of the event, it takes about 5 minutes and certainly less than 10 minutes for the total computation of the on-fault slip and of the tsunami heights. Moreover, the multi-faults (north, mid, south) solution, presented here only as seminal method, can be a useful tool to provide conservative values of coastal heights.

7 Discussion and Conclusions

This thesis dealt with the characterization of the slip distribution on the faults responsible for the largest magnitude earthquakes oriented to tsunamigenesis characterisation.

Starting from a dataset of 105 events extracted from the SRCMOD database (http://equake-rc.info/SRCMOD/) occurred after 1990 and with moment magnitude equal or larger than 6, a first step consisted in the computation of regression laws between the geometrical properties of the fault and the moment magnitude. Although there are several studies of this kind existing in the literature, we wanted to build our own set of support laws for the next step, that is, to find a simple, yet realistic, representation of the heterogeneous distribution of slip on the fault.

The analysis has been performed considering the entire dataset of finite fault models (FFM) as well as distinct types of focal mechanisms.

The studied parameters included the rupture area, length, width aspect ratio, maximum and average slip. The best correlation with magnitude is shown by the area and the length of the fault. Strike-slip events are those showing the best correlations, and furthermore the only category for which the aspect ratio shows a satisfactory level of correlation.

We also investigated relative positions of the hypocentre and of the maximum displacement MD and also their positions with respect to the fault plane centre. The distance to the fault centre of the hypocentre and of MD tends to increase with magnitude. Moreover, for smaller magnitudes, it is often close to the centre of the fault. The hypocentre rarely lies close to the fault centre or to the peak of slip.

As second main step, a method to characterize the slip heterogeneity on the fault has been devised. We consider high-level asperity/asperities as region/regions where the slip takes on values larger than 2/3 of the peak value. For each event of the database, the number of asperities has been identified through a clustering algorithm called mean shift method. This latter has been applied to the two-dimensional space represented by the centres of the subfaults on the rupture plane. The application of the clustering algorithm to the dataset of 105 earthquakes recognises that as many as 72 events exhibit a single asperity slip distribution. Then, we have best-fitted these 72 FFMs by means of 2D Gaussian distributions (2D GD). The optimal parameters characterising the 2D GD have been

determined by a least-squares procedure applied over the entire multi-parameter space scanned by regular steps. Five different optimizations have been conducted, varying the number of free parameters and/or their variability domain.

To quantify how well these distributions are able to mimic the original slip heterogeneity, they have been compared to a uniform-slip model (Uni) and a depth-dependent slip model (SCC), that is to models assumed in many tsunami hazard analyses found in the literature and in Tsunami Warning System forecasts. As a performance matrix, the co-seismic vertical surface displacements have been taken. This choice has been driven by the fact that, in the perspective of tsunami generation, the vertical seafloor deformation is used as the initial condition for the tsunami propagation. The vertical displacements at the Earth's surface field induced by the FFM slip have been assumed to be the real ones and taken as reference cases in the comparisons. The comparisons have been quantified by a misfit value and a BIC value. Results show that the best results are obtained with 2D GDs, in particular the 2D GD_1 (σ_1, σ_2, θ, u_{max} free parameters) and the 2D GD_3 (σ_1, σ_2, u_{max} free parameters).

For the three main parameters characterising the 2D DGs, i.e. σ_1, σ_2, u_{max}, scaling laws with moment magnitude have been derived, in analogy with what is commonly done for the fault geometry parameters. The two standard deviations σ_1, σ_2 also correlate well to the length and width of the fault plane respectively.

Furthermore, considering the position of the hypocentre, this has been studied on the fault planes characterized by the 2D Gaussian slip distributions. It results that hypocentres are preferentially located in regions with significant amount of slip. The only faulting mechanisms for which this trend is not true are the strike-slip ruptures. Focusing our attention to our "one asperity" model, it is possible to conclude that hypocentre often lies on the edge of the asperity.

The third main step of the study consisted in the application of the 2D GD slip models to the real case represented by the 16 September 2015 Illapel (Chile) tsunamigenic earthquake (Mw=8.3).

Two different FFMs have been considered:

- the Okuwaki et al. (2016) model (O model) included in the SRCMOD database, obtained by seismic data inversion;

- the Heidarzadeh et al. (2016) model (H model), obtained by seismic and tsunami data inversion.

The comparison among the different on-fault slip distributions highlighted that the 2D GDs are found to mimic the maximum elevations and waveforms obtained starting from the reference FFM slip models much more closely than the SCC and uniform distributions. We may draw the conclusion that the 2D GD solution is a very good representation of the "true" slip model, while being much "easier" and "faster" to be computed.

Considering again the Illapel earthquake, we derive slip models only knowing the magnitude and the location of the hypocentre, which is the most important result of the thesis. Among these models, the best behaviour, in terms of tsunami waveforms and maximum elevations is represented by the 2D GDs, whose tsunami waveforms are even better than the FFM model obtained by seismic inversion (Okuwaki et al. 2016).
The FFM by Heidarzadeh et al. 2016 is yet the best, but only slightly better than the considered 2D GD. However, it has been obtained by tsunami waveform inversions, and hence it can be potentially computed only several hours after the tsunami hits the coasts. Moreover, the idea of multi-faults solution presented in the previous Chapter, derived from the investigated uncertainty on the mutual relation between the hypocentre location and the fault plane, may represent an interesting aspect to optimize the maximization of tsunami heights to the coast.

We can conclude that the 2D Gaussian distribution is a simple representation of the seismic source, that however takes into account the slip heterogeneity, effectively replacing the main asperity, and takes a very short time to be derived. One of the winning features of the Gaussian modelling consists precisely in reproducing the asperity (main cause of the effects characterizing the tsunami) rather than the fault plane, property that also makes the 2D GD less dependent on the size of the fault plane on which it is superimposed. We believe the presented 2D GD method permits to produce reliable real-time tsunami simulations very quickly and can be used as an experimental procedure in the frame of operational tsunami warning systems. Hence, its possible application in the tsunami warning context deserves further attention and research.

In this thesis we have restricted our attention to the single-asperity earthquakes and shown that the 2D Gaussian distribution models provide good approximations of the on-fault seismic slip and can be used for tsunami modelling. Our choice neglects all those earthquakes that do not fit a single-asperity representation. To evaluate how critical is this restriction that excludes 33 events out of the 105 earthquakes of the SRCMOD database we decided to extend the 2D-GD_3 parametrization to all the 105 FFMs and calculated how much the misfit values relative to the induced coseismic vertical displacement vary. The results are reported in Appendix D, summarized in Table D.1 and provide some hint on the uncertainty introduced by applying the single-asperity method.

Another source of uncertainty is related to the focal characterisation of the seismic source, which is known to depend heavily on the seismic region. Some considerations can help to elucidate this issue. The estimation of the focal parameters characterizing a seismic event in quasi-real-time inevitably implies a certain degree of uncertainty, that can be quantified if one adopts a probabilistic approach. As reported by several studies that investigate tsunami hazard (Gonzalez et al. 2009, Burbidge et al. 2009, Grezio et al. 2017) in regions where long-term tsunami hazard is dominated by subduction sources, one can establish sound seismotectonic models that restrict the variability of focal mechanisms as well as rupture location and orientation. Consider further that subduction zones account for the vast majority of tsunamigenic earthquakes worldwide (Grezio et al. 2017).
Differently, in some specific tectonic contexts (e.g. Caribbean and Mediterranean) dominated by a large number of crustal faults with different orientations and focal mechanisms, earthquake parameters can be characterized by a higher degree of variability. Selva et al. (2016) proposed a method to include the "background" seismicity (a term used by the authors to indicate seismicity not occurring on a subduction slab) and its variability in hazard quantification. Using a logic tree structure, they define several parameters for the background seismicity (such as fault geometry and earthquake mechanism, that is strike, dip and average rake (s, d, r)) and establish a corresponding joint discrete Probability Density Function (dPDF). For the Mediterranean region, for instance, this joint dPDF can be constrained by focal mechanism catalogues, such as AllCMT, which merges data from the Global Centroid Moment Tensor (GCMT) catalogue (Dziewonski et al, 1981; Ekström et al, 2012) and the Regional Centroid

Moment Tensor (RCMT) catalogue (Pondrelli et al, 2011) and the Earthquakes Mechanisms of the Mediterranean Area (EMMA) database (Vannucci and Gasperini, 2004), which collects the focal solutions in published literature over the period 1905-2014. Besides, also data on sufficiently well-known local faults can be used. It is worthy of mention that the method proposed by Selva et al. (2016) is adopted in the framework of the Probabilistic TSUnami Hazard MAPS for the NEAM Region (TSUMAPS-NEAM), http://www.tsumaps-am.eu/wpcontent/uploads/2019/09/NEAMTHM18.

Acknowledgements

I thank Dr. Mohammad Heidarzadeh for providing us their finite fault model (Heidarzaedh et al., 2016) and the related data.

I am grateful to Professor Martin Mai and Dr. Kiran Kumar Thingbaijam for maintaining the SRCMOD database.

Special thanks to Professor Alberto Armigliato for the assistance in tsunami numerical modelling and to my supervisor Professor Stefano Tinti for the continuous and helpful suggestions during the thesis period.

References

Abe, K. (1975). Reliable estimation of the seismic moment of large earthquakes. J. Phys. Earth, 23(4), 381-390.

Aránguiz, R., González, G., Cienfuegos, R., Yagi, Y., Okuwaki, R., Urra, L., ... and Rojas, C. (2017). The 16 September 2015 Chile tsunami from the post-tsunami survey and numerical modeling perspectives. Pure Appl. Geophys. 173, 333–348. doi:10.1007/s00024-015-1225-4.

Atwater, B.F., Cisternas, M.V., Bourgeois, J., Dudley, W.C., Hendley II J.W. and Stauffer, P.H. (1999). Surviving a tsunami — lessons from Chile, Hawaii, and Japan. U.S. Geological Survey Circular No. 1187, pp. 20.

Babeyko, A. Y., Hoechner A. and Sobolev, S.V. (2010). Source modeling and inversion with near real-time GPS: a GITEWS perspective for Indonesia. Nat. Hazards Earth Syst. Sci., 10(7), 1617-1627.

Barrientos, S. E. and Ward, S. N. (1990). The 1960 Chile earthquake: inversion for slip distribution from surface deformation. Geophys. J. Intern., 103(3), 589-598.

Basher, R. (2006). Global early warning systems for natural hazards: systematic and people-centred. Philos. Trans. Roy. Soc. A: Mathematical, Physical and Engineering Sciences, 364(1845), 2167-2182.

Beresnev, I.A. (2003), Uncertainties in finite-fault slip inversions: To what extent to believe? (A critical review), Bull. Seism. Soc. Am., 93(6), 2445-2458.

Beroza, G. C. (1991). Near-source modeling of the Loma Prieta earthquake; evidence for heterogeneous slip and implications for earthquake hazard, Bull. Seism. Soc. Am., Beroza, G. C. (1991). Near-source modeling of the Loma Prieta earthquake; evidence for heterogeneous slip and implications for earthquake hazard, Bull. Seism. Soc. Am., 81, 1603–1621.

Blaser, L., Krüger, F., Ohrnberger, M. and Scherbaum, F. (2010). Scaling Relations of Earthquake Source Parameter Estimates with Special Focus on Subduction Environment., Bull. Seism. Soc. Am., 100(6), 2914-2926.

Burbidge, D., Cummins, P. R., Mleczko, R., and Thio, H. K. (2008). A probabilistic tsunami hazard assessment for Western Australia. Pure Appl. Geophys., 165(11–12), 2059–2088. https://doi.org/10.1007/S00024-008-0421-X

Choi, B. H., Kim, K. O., Yuk, J. H., Kaistrenko, V. and Pelinovsky, E. (2015). Analytical Rapid Prediction of Tsunami Run-up Heights: Application to 2010 Chilean Tsunami. Ocean & Polar Research, 37(1), 1-9.

Choi, B. H., Pelinovsky, E., Kim, K. O. and Min, B. I. (2012). Estimation of Run-up Heights of the 2011 off the Pacific Coast of Tohoku Earthquake Tsunami Based on Numerical Simulations. Open Ocean. J., 6, 5-13.

Cholifah, L. and Prastowo, T. (2017). A simple parameterization for tsunami run-up prediction. Journal of Science & Science Education, 1(2), 7-13.

Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis & Machine Intelligence, 24(5), 603-619.

Dziewonski, A.M., Chou, T.-A., & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, J. Geophys. Res. 86, 2825-2852

Dragoni, M., and Lorenzano, E. (2017). Dynamics of a fault model with two mechanically different regions. Earth Planets Space, 69(1), 145. DOI 10.1186/s40623-017-0731-2.

Ekström, G., Nettles, M. & Dziewonski, A.M., 2012. The global CMT project 2004–2010: centroid-moment tensors for 13,017 earthquakes, Phys. Earth planet. Inter., 200–201, 1–9.

Freund, L. B. and Barnett, D. M. (1976). A two-dimensional analysis of surface deformation due to dip-slip faulting, Bull. Seism. Soc. Am., 66, 667–675.

Fujii, Y. and Satake, K. (2006) Source of the July 2006 West Java tsunami estimated from tide gauge records, Geophys. Res. Lett., 33, L24317, doi:10.1029/2006GL028049.

Fujii, Y. and Satake, K. (2007). Tsunami source of the 2004 Sumatra–Andaman earthquake inferred from tide gauge and satellite data. Bull. Seism. Soc. Am., 97(1A), S192-S207.

Fukunaga, K. and Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on Information Theory, 21(1), 32-40.

Geist, E. L. (2002) Complex earthquake rupture and local tsunamis, J. Geophys. Res., 107(B5), doi:10.1029/2000JB000139.

Geist, E.L. and Dmowska, R. (1999). Local tsunamis and distributed slip at the source. Pure Appl. Geophys., 154, 485-512.

Geist, E. L., S. L. Bilek, D. Arcas and Titov V. V. (2006). Differences in tsunami generation between the 26 December 2004 and 28 March 2005 Sumatra earthquakes, Earth Planets Space, 58, 185–193.

Geist, E., Titov, V.V., Arcas, D., Pollitz, F.F. and Bilek, S.L. (2007). Implications of the 26 December 2004 Sumatra–Andaman earthquake on tsunami forecast and assessment models for great subduction-zone earthquakes, Bull. Seism. Soc. Am., 97(1A), S249–S270, doi:10.1785/0120050619.

Goda, K., Mai, P.M., Yasuda, T. and Mori, N. (2014) Sensitivity of tsunami wave profiles and inundation simulations to earthquake slip and fault geometry for the 2011 Tohoku earthquake. Earth Planets Space, 66(1), 105, doi:10.1186/1880-5981-66-105.

Goda, K., Yasuda, T., Mori, N. and Maruyama, T. (2016). New scaling relationships of earthquake source parameters for stochastic tsunami simulation. Coast. Engin. J., 58(3), 1650010-1.

González, F. I., Geist, E. L., Jaffe, B., KaNogLu, U., Mofjeld, H., Synolakis, C. E., ... Yalciner, A. (2009). Probabilistic tsunami hazard assessment at seaside, Oregon, for near- and far-field seismic sources. J. Geophys. Res., 114, C11023. https://doi.org/10.1029/2008JC005132.

Gutscher, M.A., Roger, J., Baptista, M.A., Miranda, J.M. and Tinti S., 2006. Source of the 1693 Catania earthquake and tsunami (southern Italy): New evidence from tsunami modeling and locked suduction fault plane. Geophys. Res. Lett., 33, L08309, doi:10.1029/2005GL025442.

Haskell N. A. (1964). Total energy and energy spectral density of elastic wave radiation from propagating faults. Bull. Seism. Soc. Am., 54 (6A), 1811–1841.

Hatori, T. (1968). Study on distant tsunamis along the coast of Japan, Part 2, tsunamis of South American origin, Bull. Earthq. Res. Inst., University of Tokyo, 46, 345–359.

Heaton, T. H. (1990). Evidence for and implications of self-healing pulses of slip in earthquake rupture. Phys. Earth Planet. Inter., 64(1), 1-20.

Heidarzadeh, M., Murotani, S., Satake, K., Ishibe, T. and Gusman, A. R. (2016), Source model of the 16 September 2015 Illapel, Chile, Mw 8.4 earthquake based on teleseismic and tsunami data, Geophys. Res. Lett., 43, 643–650, doi:10.1002/2015GL067297.

Iinuma, T., Hino, R., Kido, M., Inazu, D., Osada, I., Ito, Y., Ohzono M., Tsushima, H., Suzuki, S., Fujimoto, H. and Miura, S. (2012), Coseismic slip distribution of the 2011 off the Pacific Coast of Tohoku Earthquake (M9.0) refined by means of seafloor geodetic data, J. Geophys. Res., 117, B07409, doi:10.1029/2012JB009186.

Johnson, L. R., and Nadeau, R. M. (2002). Asperity model of an earthquake: static problem. Bull. Seism. Soc. Am., 92(2), 672-686.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. J. Am. Stat. Ass., 90(430), 773-795.

Lee, Y. T., K. F. Ma, and Y. T. Yen, (2016): Heterogeneous slip distribution self-similarity on a fault surface. Terr. Atmos. Ocean. Sci., 27, 181-193, doi: 10.3319/TAO.2015.11.05.01(T).

Leonard, M. (2010). Earthquake Fault Scaling: Self-Consistent Relating of Rupture Length, Width, Average Displacement, and Moment Release. Bull. Seism. Soc. Am., 100 (5A), 1971-1988, doi:10.1785/0120090189.

Leonard, M. (2014). Self-Consistent Earthquake Fault-Scaling Relations: Update and Extension to Stable Continental Strike-Slip Faults. Bull. Seism. Soc. Am., 104(6), 2953-2965, doi:10.1785/0120140087.

Løvholt, F., Pedersen, G., Bazin, S., Kühn, D., Bredesen, R. E. and Harbitz, C. (2012), Stochastic analysis of tsunami runup due to heterogeneous coseismic slip and dispersion, J. Geophys. Res., 117, C03047, doi:10.1029/2011JC007616.

Løvholt, F., Kaiser, G., Glimsdal, S., Scheele, L., Harbitz, C.B. and Pedersen, G. (2012). Modeling propagation and inundation of the 11 March 2011 Tohoku tsunami, Nat. Hazards Earth Syst. Sci., 12(4):1017-1028

Mai, P. M., Spudich P. and Boatwright J. (2005). Hypocenter locations in finite-source rupture models, Bull. Seism. Soc. Am., 95(3), 965-980.

Mai, P. M., and Beroza, G. C. (2000) Source scaling properties from finite-fault rupture models, Bull. Seism. Soc. Am., 90, 604–615.

Mai, P. M., and Beroza, G. C. (2002), A spatial random field model to characterize complexity in earthquake slip, J. Geophys. Res., 107(B11), 2308, doi:10.1029/2001JB000588.

Mai, P. M. and Thingbaijam, K. K. S. (2014). SRCMOD: An online database of finite source rupture models, Seismol. Res. Lett., 85(6), 1348-1357.

Manighetti, I., Campillo, M., Sammis, C., Mai, P. M. and King, G. (2005), Evidence for self-similar, triangular slip distributions on earthquakes: Implications for earthquake and fault mechanics, J. Geophys. Res., 110, B05302, doi:10.1029/2004JB003174.

Mendoza, C., and Hartzell, S. H. (1988b). Aftershock patterns and main shock faulting. Bull. Seism. Soc. Am., 78(4), 1438-1449.

Miyatake, T. (1992). Dynamic rupture processes of inland earthquakes in Japan weak and strong asperities. Geophys. Res. Lett., 19(10), 1041-1044.

Milliner, C. W. D., Sammis, C., Allam, A. A., Dolan, J. F., Hollingsworth, J., Leprince, S. and Ayoub, F. (2016). Resolving fine-scale heterogeneity of co-seismic slip and the relation to fault structure, Sci. Rep., doi:10.1038/srep27201.

Murotani, S., Satake, K., and Fujii, Y. (2013). Scaling relations of seismic moment, rupture area, average slip, and asperity size for M~9 subduction-zone earthquakes. Geophys. Res. Lett., 40(19), 5070-5074.

Ohta, Y., Kobayashi, T., Tsushima, H., Miura, S., Hino, R., Takasu, T., ... and Sato, T. (2012). Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: Application to the 2011 Tohoku-Oki earthquake (Mw 9.0), J. Geophys. Res., 117, B02311, doi:1029/2011JB008750.

Okada, Y. (1992) Internal deformation due to shear and tensile faults in a half-space, Bull. Seism. Soc. Am., 82(2), 1018–1040.

Okuwaki, R., Yagi, Y., Aránguiz, R., González, J. and González, G. (2016). Rupture process during the 2015 Illapel, Chile earthquake: Zigzag-along-dip rupture episodes. Pure Appl. Geophys. 173, 1011–1020. DOI 10.1007/s00024-016-1271-6.

Pondrelli, S., Salimbeni, S., Morelli, A., Ekström, G., Postpischl, L., Vannucci, G. and Boschi, E., 2011. European–Mediterranean Regional Centroid Moment Tensor catalog: solutions for 2005–2008, Phys. Earth Planet. Inter., 185(3–4), 74–81.

Raghukanth, S.T.G. and Sangeetha, S. (2016) A stochastic model for earthquake slip distribution of large events, Geomatics, Natural Hazards and Risk, 7(2), 493-521, DOI: 10.1080/19475705.2014.941418.

Romano, F., Piatanesi, A., Lorito, S., Tolomei, C., Atzori, S. and Murphy, S. (2016), Optimal time alignment of tide-gauge tsunami waveforms in nonlinear inversions: Application to the 2015 Illapel (Chile) earthquake, Geophys. Res. Lett., 43, 11,226– 11,235, doi:10.1002/2016GL071310.

Satake, K. (1987). Inversion of tsunami waveforms for the estimation of a fault heterogeneity: Method and numerical experiments. J. Phys. Earth, 35(3), 241-254.

Satake, K. and Heidarzadeh, M. (2017). A review of source models of the 2015 Illapel, Chile earthquake and insights from tsunami data. Pure Appl. Geophys. 174, 1–9.

Satake, K. and Kanamori, H. (1991). Use of tsunami waveforms for earthquake source study. Natural Hazards, 4(2-3), 193-208.

Selva, J., Tonini, R., Molinari, I., Tiberti, M. M., Romano, F., Grezio, A., ... Lorito, S. (2016). Quantification of source uncertaintiesin seismicprobabilistic tsunami hazard analysis (SPTHA). Geophys. J. Int, 205, 1780–1803. https://doi.org/10.1093/gji/ggw107.

Sobolev, S. V., Babeyko, A. Y., Wang, R., Hoechner, A., Galas, R., Rothacher, M., Sein, D. V., Schröter, J., Lauterjung, J. and Subarya, C. (2007), Tsunami early warning using GPS-Shield arrays, J. Geophys. Res., 112, B08415, doi:10.1029/2006JB004640.

Somerville, P. G., K. Irikura, R. Graves, S. Sawada, D. J. Wald, N. Abrahm- son, Y. Iwasaki, T. Kagawa, N. Smith and A. Kowada (1999). Characterizing crustal earthquake slip models for the prediction of strong ground motion, Seism. Res. Lett., 70(1), 59–80.

Stirling, M., T. Goded, K. Berryman, and N. Litchfield (2013). Selection of earthquake scaling relationships for seismic-hazard analysis, Bull. Seism. Soc. Am., 103, 2993–3011.

Strasser, F. O., Arango, M. C. and Bommer, J. J. (2010). Scaling of the Source Dimensions of Interface and Intraslab Subduction-zone Earthquakes with Moment Magnitude. Seismol. Res. Lett., 81(6), 941-950.

Thatcher, W. (1990). Order and diversity in the modes of circum-Pacific earthquake recurrence. J. Geophys. Res., 95(B3), 2609-2623.

Thingbaijam, K. K. S., Martin Mai, P. and Goda, K. (2017). New empirical earthquake source-scaling laws. Bull. Seism. Soc. Am., 107(5), 2225-2246.

Tinti, S., Armigliato, A., Pagnoni, G. and Zaniboni, F (2005). Scenarios of giant tsunamis of tectonic origin in the Mediterranean, ISET J. Earthq. Techn., 42(4), 171-188.

Tinti, S., Gavagni, I. and Piatanesi, A. (1994), A finite-element numerical approach for modeling tsunamis, Annals of Geophysics 37, 1009–1026.

Tonini, R., Armigliato, A., Pagnoni, G., Zaniboni, F. and Tinti, S. (2011). Tsunami hazard for the city of Catania, eastern Sicily, Italy, assessed by means of Worst-case Credible Tsunami Scenario Analysis (WCTSA). Nat. Hazards Earth Syst. Sci., 11(5), 1217–1232.

Vannucci, G., and Gasperini, P., 2004. The new release of the database of earthquake mechanisms of the Mediterranean area (EMMA Version 2), Annals of Geophysics 47, 307-334.

Watanabe, H. (1998). Materials for comprehensive list of destructive tsunamis in Japan (‘‘Nihon Higai Tsunami Souran’’), pp. 206. Tokyo: University of Tokyo Press.

Wells, D. L. and Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull. Seism. Soc. Am., 84(4), 974-1002.

Ye, L., Lay, T., Kanamori, H. and Koper, K.D. (2016), Rapidly estimated seismic source parameters for the 16 September 2015 Illapel, Chile Mw 8.3 Earthquake, Pure Appl. Geophys. 173, 321–332. doi:10.1007/s00024-015-1202-y.

http://equake-rc.info/SRCMOD/ (last access on 20 February 2019)

https://earthquake.usgs.gov/earthquakes/ (last access on 26 June 2019)

https://www.ngdc.noaa.gov/hazard/tsu_db.shtml (last access on 26 June 2019)

http://www.tsumaps-neam.eu/wp-content/uploads/2019/09/NEAMTHM18 (last access on 30 October 2019)

Appendix A

Graphic examples of the 2D GD functions

We present the 2D GD function expressed by the Eq. 4.3 in Section 4.2 for four different configurations, letting the angle θ vary. Placing $u_{max} = 0.15$, $\sigma_1 = 1.5$, $\sigma_2 = 0.5$, the four configurations reported in Figure A.1 correspond to:

a) $\theta = 0°$;

b) $\theta = 30°$;

c) $\theta = 90°$;

d) $\theta = 150°$.

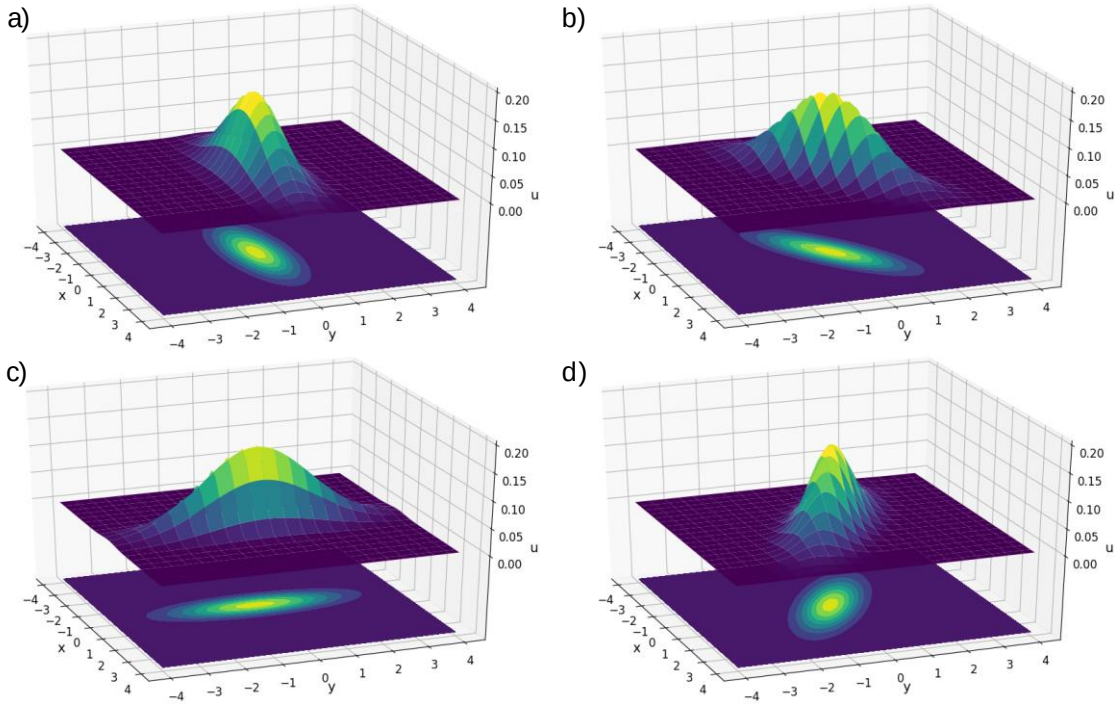

Figure A.1 Plots of the 2D GD functions for $\theta = 0°$ (a), $\theta = 30°$ (b), $\theta = 90°$ (c), $\theta = 150°$ (d).

Appendix B

Tsunami waveforms produced by the different 2D GD functions

Considering the specific case of the Illapel, 16th September 2015 earthquake, there is no great difference in terms of tsunami waveforms produced by the different Gaussian optimizations we have used.

The graphs reported in Figure B.1- Figure B.5 show, in the same plot for each considered stations, the marigrams obtained with the five Gaussian optimizations. The differences are negligible. The largest discrepancies are found for the station of Coquimbo, where the 2D-GD_1 appears to have slightly weaker peaks and the 2D-GD_4 presents the maximum value of water elevations. However, in the context of this study, these subtle differences are of little significance and do not affect the result of the work.

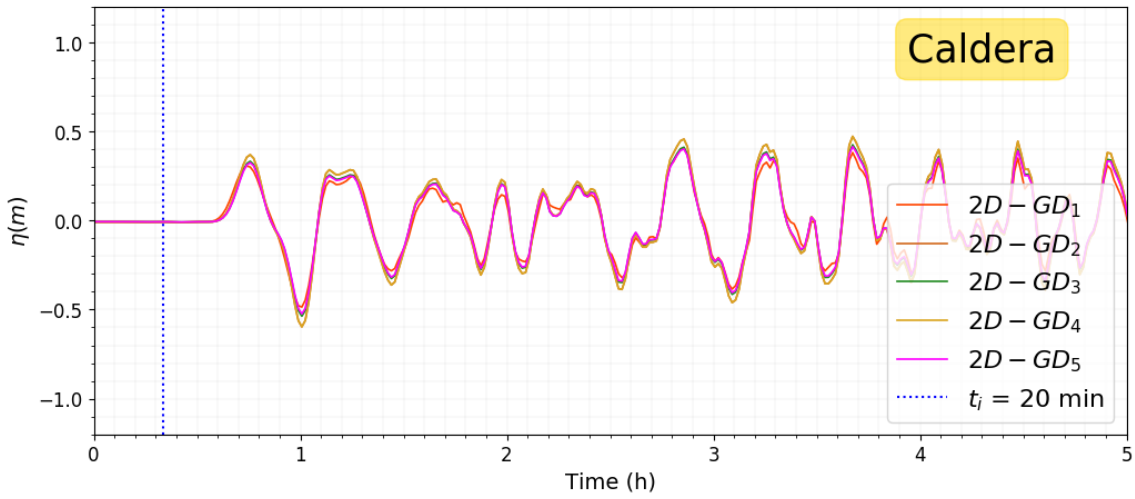

Figure B.1 Marigrams obtained from the 2D-GD optimizations at the Caldera station.

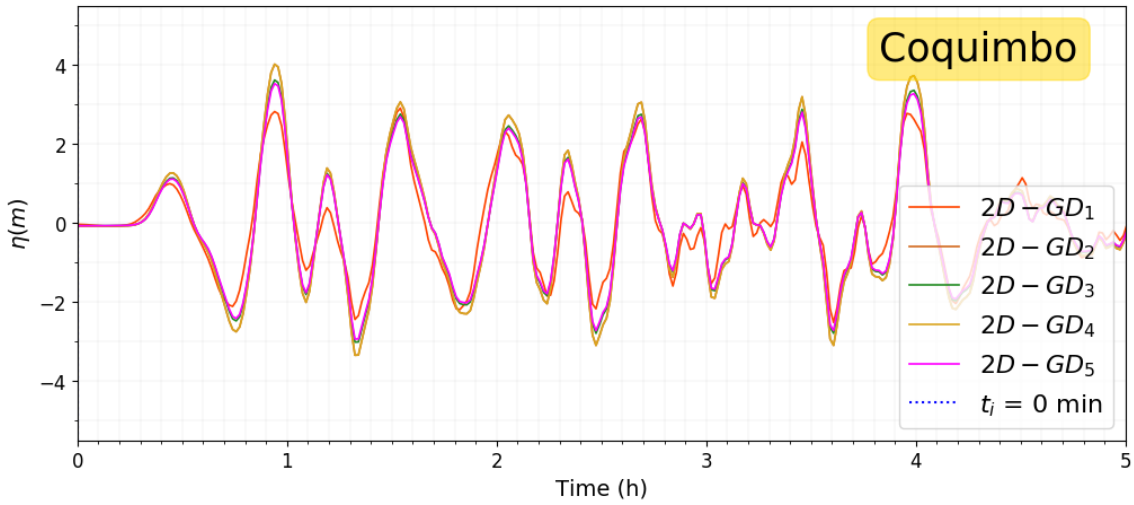

Figure B.2 Marigrams obtained from the 2D-GD optimizations at the Coquimbo station.

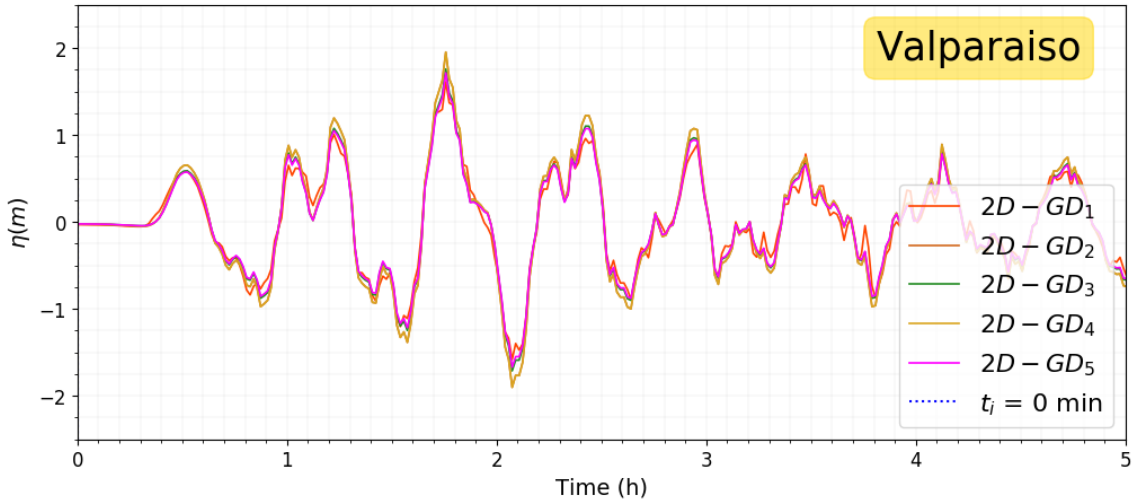

Figure B.3 Marigrams obtained from the 2D-GD optimizations at the Valparaiso station.

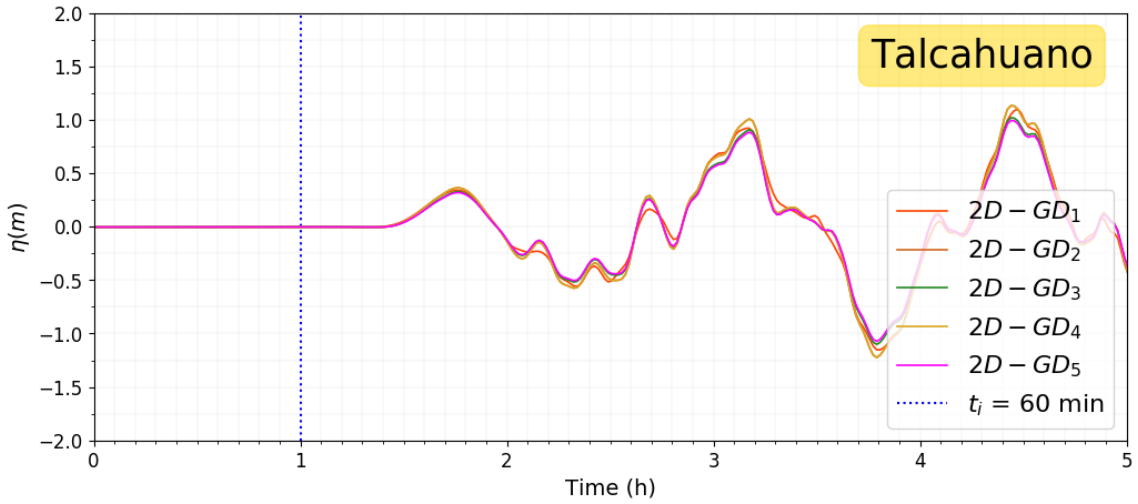

Figure B.4 Marigrams obtained from the 2D-GD optimizations at the Talcahuano station.

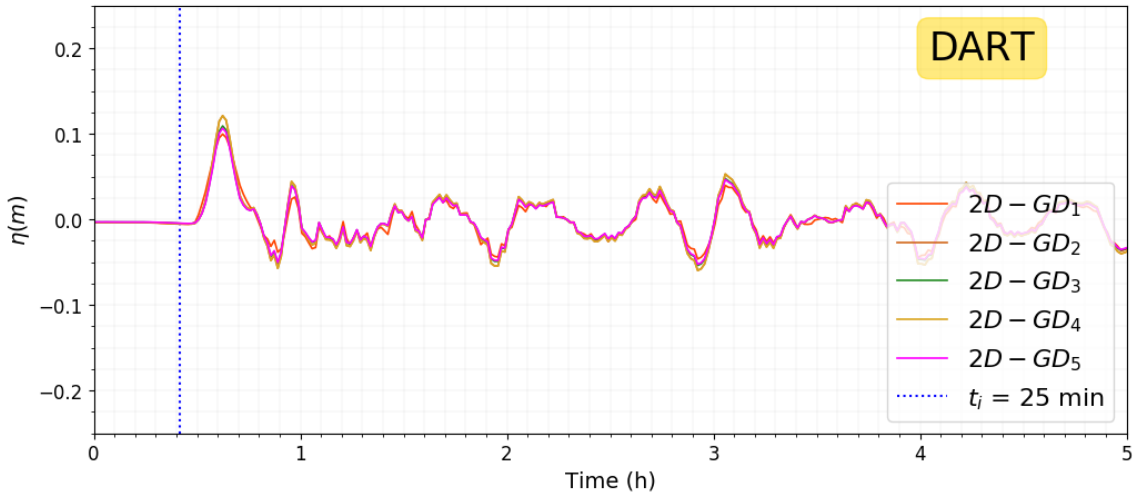

Figure B.5 Marigrams obtained from the 2D-GD optimizations at the DART station 32402.

Appendix C

Computation of the SCC on-fault slip

If we consider a fault with a normalised width W ranging from 0 to 1, we can express the position of the slip peak by a parameter q ranging as well from 0 to 1. In the Smooth Closure Condition (SCC) algorithm the function that expresses the slip at the depth z is the following:

$$f(z,q) = \begin{cases} 2\dfrac{z^2}{q^2}\left(3 - 2\dfrac{z}{q}\right) & , \quad z < q \\ 2 + \dfrac{4(z^3 - q^3) + 12q(z-q) - 6(1+q)(z^2 - q^2)}{(1-q)^3}, & \quad z \geq q \end{cases}$$

where a distinction is made between the slip above and below the peak. Since there is no dependence from the along-strike coordinate, the slip distribution is depth dependent but laterally homogeneous. In this thesis, an SCC slip distribution is obtained by assigning to the generic fault cell (i_x,i_z) the slip given by:

$$s(i_x, i_z) = f(z,q) \cdot \frac{AD}{AD_{SCC}}$$

where AD is the average FFM displacement and AD_{SCC} is the average of the function *f(z,q)* over all the cells.

Figure C.1 shows trends of the function *f(z,q)* for different values of *q*.

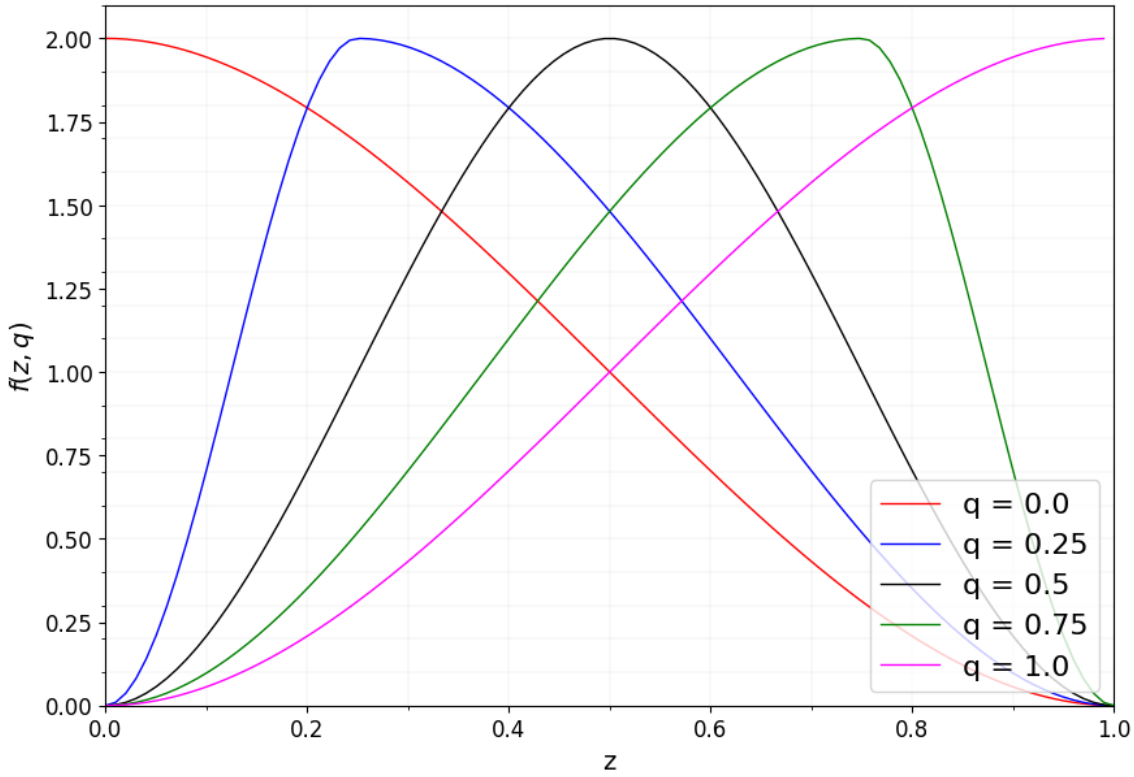

Figure C.1 Plot of the SCC function for different values of q.

Appendix D

Misfit values for all the 105 FFMs of the database

In this Appendix, the misfits of the 2D Gaussian models for the one-asperity subset (72 cases) and the full set (105 cases) of FFMs will be compared, and this can provide a clue to estimate the error brought by assuming a 2D Gaussian distribution a priori.

In Figure D.1 the misfits related to the 72 single-asperity models selected with the mean-shift algorithm are reported in red, while the misfits of the remaining 33 events are reported in blue. It is evident that the multi-asperity events worsen the misfit performance since they are shifted towards the high-misfit region of the frequency plot. The misfit value marking approximately the boundary between single- and multi-asperity events is about 0.77, which is the value of the 95th percentile of the misfit distribution, considering all the events. Looking more carefully at the percentile values, one can notice that, in any case, 75% of all the events present a misfit value lower than 0.5.

The misfit trends for the Uniform (Figure D.2) and SCC (Figure D.3) distributions are definitely worse because the 75th percentile is 0.81 and 0.75 for the former and for the latter. Notice further that in these two cases, the addition of the 33 events does not change the general character of the distribution, which confirms that these distributions are less performant than the 2D Gaussian distributions for all cases (single- and multi-asperity slip fields).

Table D.1 Minimum, maximum and average values of the misfit for each type of slip distribution (105 models).

Slip distribution	Min Misfit	Max Misfit	Mean Misfit
Uniform	0.065	1.066	0.628
SCC	0.022	1.121	0.578
2D GD_3	0.042	1.064	0.413

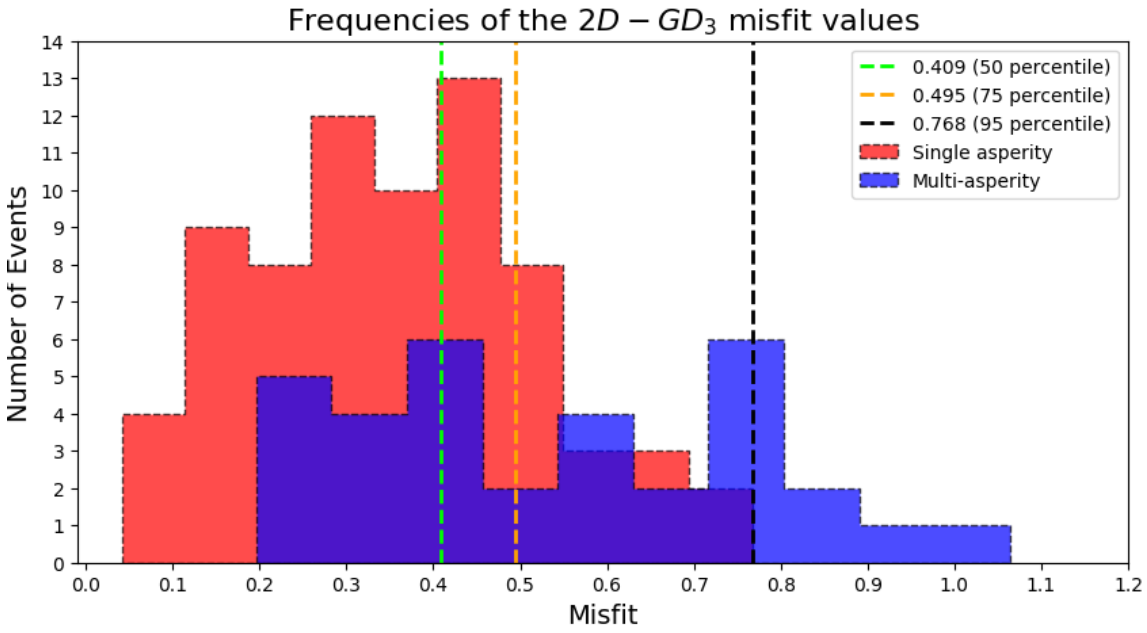

Figure D.1 Distribution of the 2D-GD$_3$ misfit values for all the 105 events of the database. In red: the single asperity models (72) selected by the mean shift algorithm. In blue: the multi-asperity models (33).

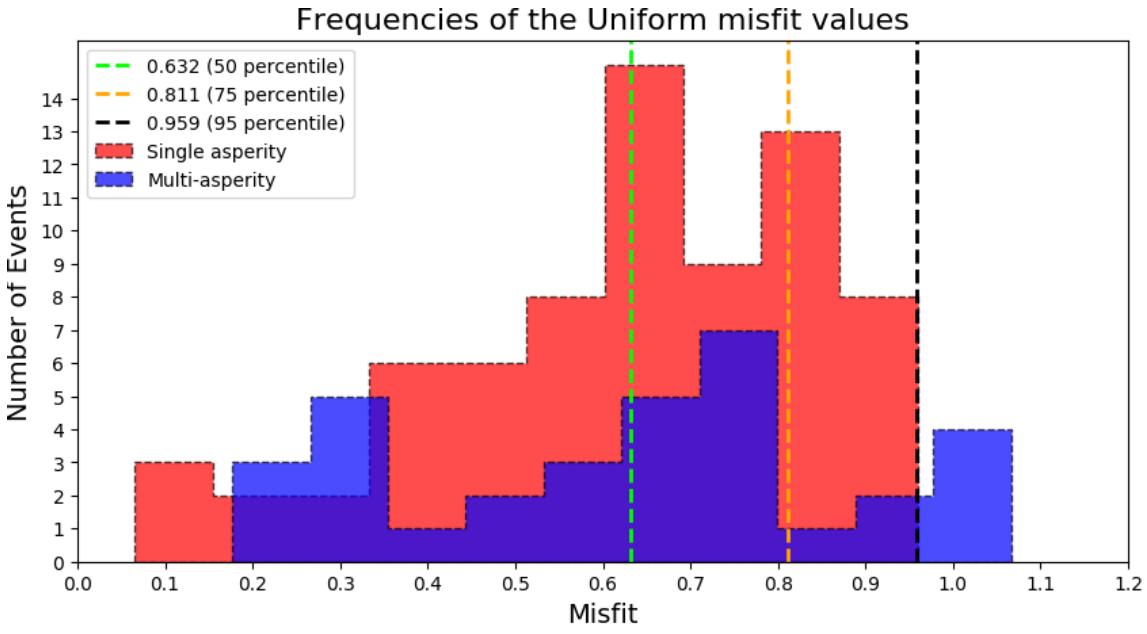

Figure D.2 Distribution of the Uniform misfit values for all the 105 events of the database.

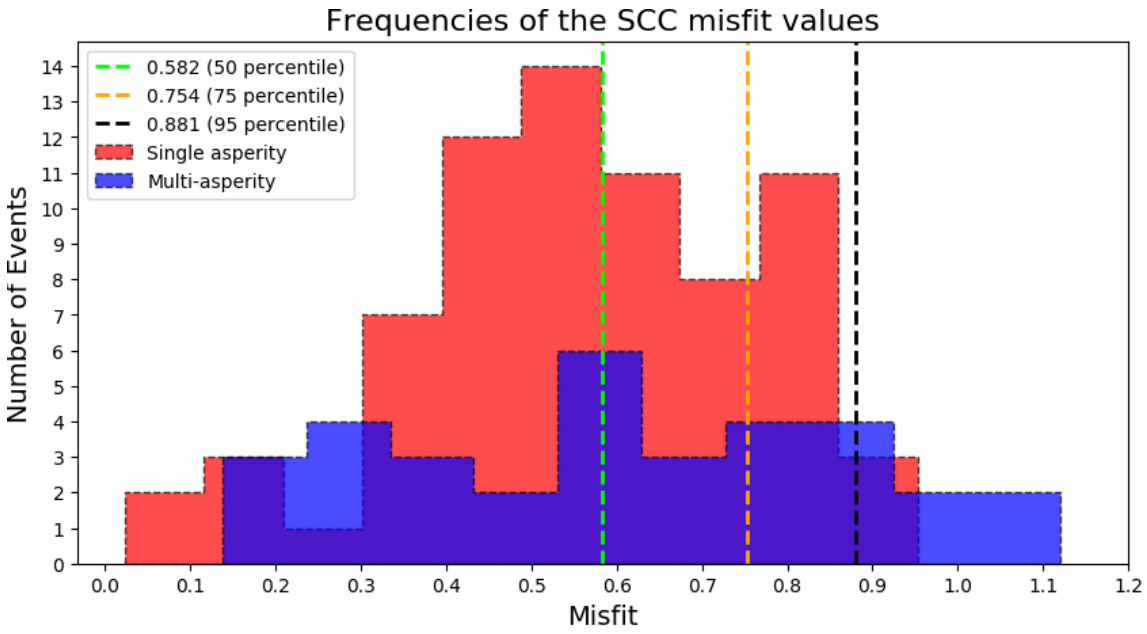

Figure D.3 Distribution of the SCC misfit values for all the 105 events of the database.

Appendix E

Scaling between the slip and area characterising the asperities

The concept of asperities is of great relevance in the framework of seismological studies and tsunami hazard assessment. Among the several studies concerning this topic, we consider the work proposed by Lee et al. (2016). Examining 41 earthquakes in the magnitude range Mw=4.6-8.9, the authors found that the fault slip exhibited fractal scaling between the rupture slip and area. They consider the slip ratio R_d between the slip above a certain threshold *d* and the average slip, and the average area ratio R_s between the area characterised by slip equal to, or larger than, *d* and the total rupture area. The scaling relationship found is of the form:

$$R_s = 10^{a-n(R_d)} \tag{E.1}$$

Examining this relationship for different values of the ratio R_d and different magnitude intervals, the authors found a self-similarity in the heterogeneity in slip distribution over the fault. The self-similar scaling exponent indicates the degree of fractal dimension in the fault slip system according to Lee et al. (2016). They affirm that spatial slip distribution for large earthquakes (M_w > 7) tends to have a more homogeneous slip distribution compared to the moderate events.

Equation (E.1) can be written as:

$$Log(R_s) = a - n \cdot R_d \tag{E.2}$$

Applying the equation to our 105 FFMs, we investigated the relationship between the slip ratio that characterizes our asperities and their area ratio. Having defined our asperities with respect to the maximum slip and not to the average one, the ratio R_d characterizing our events is not fixed. Lee et al., instead, analyse the scaling behaviour for different fixed values of the slip ratio.

The plot in Figure E.1 displays our results: the scaling behaviour appears to be confirmed with a correlation coefficient modulus equal to or larger than 0.7 for all the focal mechanisms, except for reverse events (|r| = 0.56). The parameter values differ from those obtained by Lee et al. (2016), but this difference can easily be attributed to the different subcategories considered, to the different adopted ratio R_d, and to the different magnitude intervals considered.

Despite these factors, the general trend of the area ratio as a function of the slip that characterizes the asperity respects the trend suggested by Lee et al. (2016) and, as reported by them, the scaling relationship for the slip partition in a finite-fault provides an important basis for ground motion prediction, which is particularly crucial in assessing seismic and tsunami hazards and simulated earthquake and tsunami scenarios.

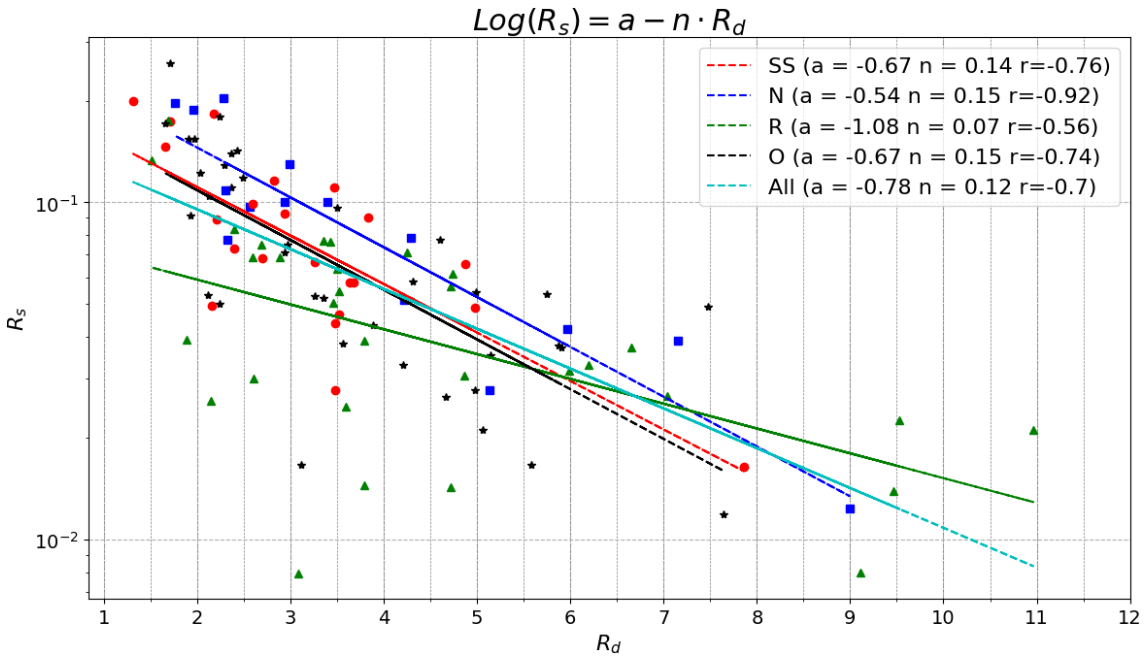

Figure E.1 Scaling law between area and slip ratios (respectively R_s and R_d) that characterise the asperities of this study.