UNIVERSIDAD DE CONCEPCIÓN

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE GEOFÍSICA

IMPLEMENTATION OF A NUMERICAL METHODOLOGY FOR THE STOCHASTIC CHARACTERIZATION OF THE VALDIVIA 1960 9.5 MW TSUNAMI SOURCE.

BY RODRIGO IGNACIO CIFUENTES LOBOS

Thesis presented to Facultad de Ciencias Físicas y Matemáticas to be eligible for the degree of

Master of Sciences in Geophysics.

Thesis advisor:

Dr. María Ignacia Calisto Burgos

Committee members:

Dr. Breanyn MacInnes - Dr. Marcos Moreno Switt - Dr. Jorge Quezada Flory

July 2022

Concepción, Chile.

IMPLEMENTATION OF A NUMERICAL METHODOLOGY FOR THE STOCHASTIC CHARACTERIZATION OF THE VALDIVIA 1960 9.5 MW TSUNAMI SOURCE.

BY RODRIGO IGNACIO CIFUENTES LOBOS

Thesis presented to Facultad de Ciencias Físicas y Matemáticas to be eligible for the degree of Master of Sciences in Geophysics.

Thesis advisor:

Dr. Ignacia Calisto Burgos

Committee members:

Dr. Breanyn MacInnes - Dr. Marcos Moreno Switt - Dr. Jorge Quezada Flory

July 2022

Concepción, Chile.

Aknowledgements

This thesis would not have been possible without my family's unconditional and wholearted support, especially from my father, who taught me to love learning and to pursue my goals. My gratitude to Sandra, my brethren and to Isadora, whose love and words of encouragement helped me.

To Ignacia, for her endless patience, trust and her outstandig ability to teach and to bring out the best of one.

To my friends, colleagues and collaborators, especially to Franchesca, Javiera, Matías and Chris. The present study is funded by FONDECYT project 11180854 "Source characterization for historical tsunamis of Central-Southern Chile", directed by Ignacia Calisto, PhD. Additional support and funding by Iniciativa Científica Milenio (ICM) through Grant Number NC160025 "Millenium Nucleus CYCLO: The seismic Cycle Along Subduction Zones" was provided. We wanted to extend our gratitude to Dr. Daniel Stewart and Dr. Patricio Winckler for providing insightful comments and tsunami data. All figures were produced using Generic Mapping Tools, GMT (Wessel *et al.,* 1998).

To two generations of Dagoberto Cifuentes.

Contents

List of Figures

List of Tables

Abstract

Probabilistic Tsunami Hazard Assessment (PTHA) brings a variety of mathematical and numerical tools for evaluating long-term exposure to tsunami related hazards in coastal communities, within which the logic tree method stands out for its usefulness and versatility in generating random slip models and dealing with epistemic and aleatory uncertainties, key items for the stochastic study of future tsunami scenarios. This method, by combining parameters that define a source model (such as magnitude, and rupture limits), allows for the creation of a vast number of random source models that can be used for assessing future and long-term hazard. They can also be used in conjunction with data and observations obtained from past tsunamis and earthquakes to open new possibilities for studying past tsunami, and their seismic source models.

This study proposes a numerical methodology for the generation of random tsunami source models, based on the aforementioned logic tree method, for studying past tsunamis and historical tsunamis. In this case this methodology will be tested with compiled data from the great Valdivia 1960 9.5 Mw earthquake and tsunami. This methodology works by filtering the random source models produced by using the logic tree methodology in a staggered fashion. Firstly, they are filtered with empirical relations between magnitudes and rupture dimensions or rupture aspect ratios. The remaining models are then used to compute vertical seafloor deformation using the Okada (1985) solution. These deformation fields are then compared with geodetic data and observations associated with the event of interest, in this case the Valdivia 1960 earthquake, eliminating all

models that do not satisfy these observations. In contrast, all models that do pass this filter, are used as inputs to model tsunami using a staggered scheme, first modelling with low resolution topobathymetry grids, in order to assess if tsunami waves are registered in locations that are known to have been inundated and eliminate the models that do not show this behavior. For those that fulfil the low-resolution modeling, high resolution grids are used to model tsunami and appraise the estimated run up of inundations and compare them with reliable historical accounts and sedimentological observations. The models that pass all filters mentioned above, will be subjected to statistical analysis, such as conditional probability analysis of slip amounts according to its location or the analysis of the cumulative density functions of the slip, to compare them with existent, published, models of the Valdivia 1960 earthquake.

In order to appraise the convergence of the random models generated using a logic tree approach that pass every filter to the existent source models, the Valdivia 1960 9.5 Mw event will be used as a benchmark to test this methodology, due to the number of published studies, data available, reliable historical accounts and source models computed with different techniques and from different data sets, such as geodetic, seismic or tsunami recordings. It is of the utmost importance to further specify that this methodology was designed, and is intended to be used, to study historical tsunami, and will only be tested with modern tsunamis because of the availability of data and studies, such is the case of the Valdivia 1960 earthquake.

The hypothesis proposed in this study is that the estimation of the most likely tsunami source (slip distribution and rupture geometry and limits) of the Valdivia 1960, procured through the analysis of random displacement models obtained using a logic tree structure will be a solution that satisfy the available geodetic, deformation and tsunami data, observations and historical accounts.

This work is subdivided into two parts, an initial resolution test with synthetic deformation, wave arrival and inundation data to test the capabilities and the response of the method to different types of data availability, such as differences in data density (a large number of data points versus sparse points), distribution (uniform distribution along the territory versus clusters of data or

just one cluster) or differences in availability of deformation or tsunami data. Once this test is performed, the methodology will be applied to the case of the Valdivia earthquake.

Chapter 1

Introduction

The subduction zone between the Nazca and South American plates, margin along which Chile is located from its northernmost point in Visviri to the Ofqui Isthmus in Aysén, is a very seismogenic border. It is estimated that every century there are at least five earthquakes with a moment magnitude greater than 8 Mw and that giant events with magnitudes higher than 9 Mw occur every 300 to 400 years (Lomnitz, 2004, Ruiz & Madariaga, 2018). There exists written historical record of earthquakes and tsunamis that took place in Chile since at least the XIII century by way of Japanese records of a tsunami that affected their coast and that can be traced to Chile (Tsuji *et al.,* 2013), and starting in the XVI century via Spanish monks and conquistadors' records located in the "Archivo de Las Indias", in Sevilla, Spain (Cisternas *et al.,* 2012). Lomnitz (2004) counts 14 ($>$ 8 Mw) events from 1570 until 1960, a list to which five new entries have beend added recently, earthquakes with magnitudes higher than 8 Mw starting with the 1985 Algarrobo event (8.0 Mw) and ending with the Coquimbo 2015 (8.5 Mw). On the other hand, Ruiz & Madariaga (2018) estimated the occurrence of at least 3 giant ($>$ 9 Mw) earthquakes (i.e., 1730, 1868 and 1960, also included in the list of 14) in the past 300 years, on segments ranging in locations from the south of Perú and northern Chile (1868 9.0 Mw earthquake) to the Valdivia segment (1960 9.5

Mw). Figure 1.1 shows a historical record of the large earthquakes since 1730 until 2017 along the Chilean coast, plotting with solid lines the estimated rupture length, and with points the NEIC catalog of earthquake epicenters with magnitudes higher than 4 Mw from 1900 to 2017. From the point of view of the rate of tsunamis that hit Chilean coasts, only in the past decade were there 12 events with high enough intensities to be at least measured by offshore instruments (i.e. tide gauges, buoys). Among this 12, a quarter of them had the intensity to destroy infrastructure and cause the loss of life (Global Historical Tsunami Database, NCEI. Last accessed in February 2021).

Despite Chilean's centuries old tradition and a relationship with earthquakes and tsunamis, there are still long ways to go in order to continue advancing towards the construction of more resilient coastal communities and to develop safer means of inhabiting the vast coastal area of Chile it is necessary to further investigate the dynamics of tsunamigenic earthquakes and their consequences. One way help inhabitants to be better prepared for future events is to study the behavior of past earthquakes and tsunami because it will the authorities to take more informed decisions regarding coastal development projects, and hazard assessment (DeRisi *et al.,* 2017).

The biggest earthquake in recorded history, known as the great Valdivia earthquake, took place on May 22, 1960 registering a magnitude of 9.5 Mw. This event ruptured an area on the interface of the Nazca and South American Plates with a length of over 1,000 km, from the Gulf of Arauco in the North (∼37° S) to the Ofqui Isthmus, in the South (∼46° S)(Cifuentes, 1989, Cifuentes & Silver, 1989, Barrientos & Ward, 1990, Lomnitz, 2004, Moreno *et al.,* 2004). This earthquake follows a nearly 33 hours long sequence of earthquakes that started the day before with the 8.1 Mw Concepción event (Cifuentes, 1989). It is hypothesized that this earthquakeis comprised of three events, that ruptured the entire segment in a rapid succession that started in the northern part of the segment and continued southward in a span of approximately 15 minutes (Cifuentes, 1989, Cifuentes & Silver, 1989). The earthquake and consequent tsunami affected an area inhabited by nearly 2.5 million people causing the death of more than 2,000. It affected coasts as far away as Hawaii and Japan, where nearly 100 people lost their life. The earthquake and tsunami caused

damage valued between 500 and 700 million USD (in 1960 value, not adjusted to inflation), besides inundating more than 15,000 hectares of farmland on the outskirts of Valdivia (Barrientos & Ward, 1990). The huge consequences left by this event, the vast areas affected by the tsunami and the loss of infrastructure, farmland and above all life, elicited a great scientific interest, from which emerged a vast number of studies, especially regarding its source and slip distribution (Barrientos & Ward, 1990, Moreno *et al.,* 2009, Fujii & Satake, 2013). On top of that, there are vast amounts of data, biological, sedimentological or geodetic that grant an opportunity to further deepen our understanding of tsunamigenic earthquakes on the Chilean margin, and to help to reconstruct similar past events. Thus, due to the aforementioned data compiled from this event, this earthquake is a perfect candidate to serve as a benchmark in the testing of new methodologies for studying past earthquakes and tsunami.

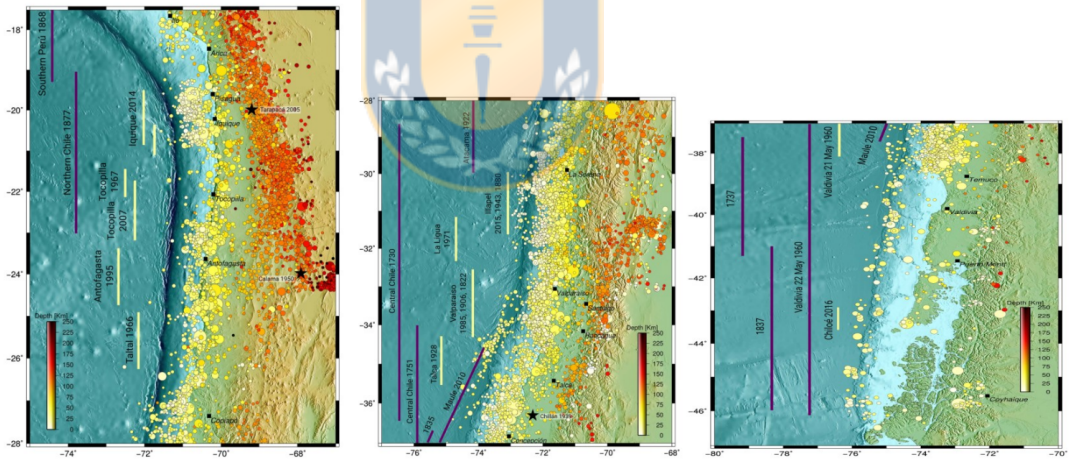


Figure 1.1: Representation of the historical seismicity in the Chilean margin of the subduction between the Nazca and South American plates. The north-south lengths of the estimated rupture zones of $>$ 8.0 Mw earthquakes are shown with lines. Purple lines show the extension of earthquakes that ruptured the entire segment, while yellow lines show those events that partially ruptured their corresponding segments. Colored circles show the epicenters of $>$ 4.0 Mw events from the NEIC catalog ranging from 1900 to 2017. Figure modified from Ruiz & Madariaga (2018).

Historically (and also recently), studies of past and ancient earthquakes and tsunamis are done using sedimentological and deformation measurements of the sites affected by these events (Garrett *et al.,* 2015, Hong *et al.,* 2017, Cisternas *et al.,* 2017, Hocking *et al.,* 2017). However, these

studies are hard and expensive to deploy and require a larger work force than other studies, with complex logistics and schedules. Furthermore, this type of measurements usually has inherently large uncertainties, making them hard to interpret. Nevertheless, sedimentological and deformation measurements taken using modern techniques and approaches, such as diatom and plankton description (Dura, 2016), are of incredible value and a necessity in the study of past tsunami.

On the other hand, several numerical techniques and tools have been developed to deal with uncertainty (LeVeque *et al.,* 2016, Grezio *et al.,* 2017), random or epistemic, especially within the probabilistic tsunami hazard assessment (PTHA) discipline (Fukutani *et al.,* 2014, Kulkarini *et al.,* 2016, Fukutani *et al.,* 2018, Becerra *et al.,* 2020). Among these tools, the Logic Tree (LT) approach excels in dealing with uncertainties, however, this technique is mostly used in studying hazard scenarios of future earthquakes and tsunamis (Fukutani *et al.,* 2014, Goda *et al.,* 2016, Fukutani *et al.,* 2018), and its use in the study of past earthquakes and tsunami is an opportunity seldom used.

This study proposes the conjunction of the logic tree approach and the importance of *in situ* measured data to assess the source models of ancient tsunami. This is achieved by replacing branches in a LT structure associated with hazard assessment, such as recurrence interval (Annaka *et al.,* 2007), for other branches to mitigate the uncertainty in the rupture geometry, for example the limits and aspect ratio of this area, in order to achieve more realistic heterogeneous models. This way, source models are characterized by an earthquake's magnitude, length, width (aspect ratio and north and south limits), location (distance to trench) fand slip distribution. In order to attain a source model, random slip distributions are generated using combination of source defining parameters given by the LT. Figure 1.2 shows an example of the structure of the logic tree proposed to be used in this methodology. These parameters are then used to create a random slip distribution using a Karhunen-Loève (K-L) expansion (see 6.2 for more details) and the Slab2.0 (Hayes, 2018) model for defining the plate interface geometry. This K-L expansion uses a log-normal distribution function in order to avoid negative slip values (LeVeque *et al.,* 2016).

The LT approach used in this study corresponds to a decision-schema organized in several levels

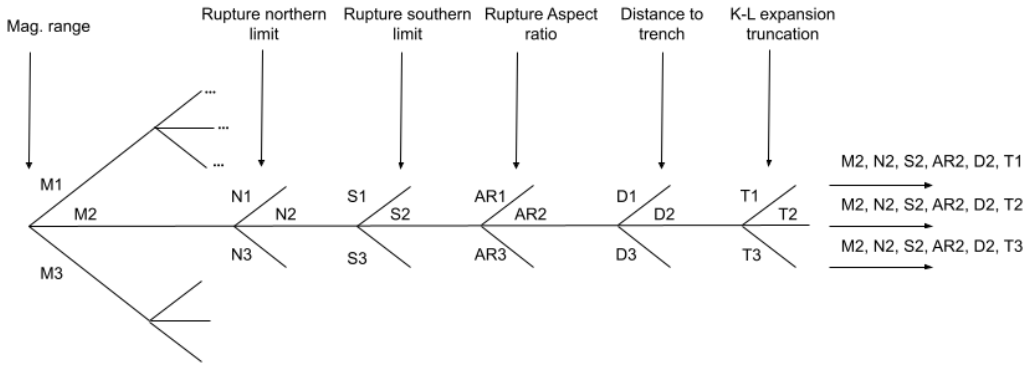


Figure 1.2: Logic tree structure example used in the creation of random source models. The LT consists of the branches for the combinations of source parameters used to define the random models, magnitudes, northern and southern limits of the rupture, aspect ratio of the rupture, distance to trench of the updip limit and truncation level of the K-L expansion.

that part ways, separating themselves as "branches" of a tree, so as if one follows a branch, one could obtain all of the different input parameters needed to create a random slip distribution. If all of the branches are followed one by one, all of the combinations of all of the parameters are obtained. Each level of the tree represents a source defining parameter, in this case (and as is showed in figure 1.2) the tree is comprised of six levels, corresponding to magnitude of the event, northern and southern limit of the rupture, complexity of the K-L expansion, aspect ratio of the rupture and the distance to the trench, based on Lay's domains (Lay *et al.,* 2012). This structure is built upon a mutually exclusive ramification principle, in other words, for each generation of a random model one and only one branch combination is active.

The vast number of models created using the LT structure is necessary in order to compensate for empirical and random uncertainties (LeVeque *et al.,* 2016, Grexio *et al.,* 2017), as well as to ensure the heterogeneity of slip distributions. The use of heterogeneous slip distribution models is preferable compared to the use of homogeneous ones, because the latter tend to underestimate tsunami intensities (Carvajal *et al.,* 2017), making more difficult to satisfy the observations with the modelling, and tending to overestimate the magnitude of earthquakes that comply with the data.

As stated before, this type of structures is capable of generating a vast number of random models, however, due to the random nature of the generated models, not necessarily all of them will be able to represent a natural or possible (in a physical sense) earthquake. To compensate this, it is necessary to include filters and restrictions to the models in order to dismiss those that are not plausible, and to limit the computational costs and complexity using only models that comply with physical restrictions and data, both geological and historical records.

In this stage is where the conjunction of the studies of ancient earthquakes and tsunami using sedimentological and deformation data, and the mathematical and numerical tools given by PTHA come together. Every random model created with the LT is subjected to two types of filters (or restrictions), one group based on physical properties of earthquakes and another based on the available deformation and/or tsunami data. This tests are performed in a staggered fashion, applying first those which require less computational costs first, and then increasing in complexity in each subsequent filter. By the nature of this method, each successive filter is applied to a smaller number of models, thus decreasing the toll on the computer, and decreasing the run time of the algorithm. Initially, the models' compliance to the observed physical dynamics of subduction earthquakes is tested, checking if they meet the specifications given by the empirical relations and scaling laws (Kanamori & Anderson, 1975, Abe *et al.,* 1975, Geller *et al.,* 1976, Purcaru *et al.,* 1982). Those models that pass this first filtering instance are then subject to a second group of filters, comprised of the data available, this is done in two, a first stage with geodetic and deformation data, and a second and third with tsunami data, divided in first filter based on wave arrival and a second one on wave height and inundation.

1.1 State of the art

One of the main aims of this study is to deepen our understanding of the seismic recurrent characteristics and earthquake dynamics in Chile, proposing a methodology for estimating and studying

slip distributions of historic earthquakes and a case study of the Valdivia 1960 9.5 Mw earthquake, a well-studied earthquake (e. gr. Barrientos & Ward (1990), Moreno *et al.,* (2009), Fujii & Satake (2012)), albeit the lack of available data on the southern part of the rupture makes for low resolution in this area. The study of the recurrence and seismic history in the Chilean margin is highly advanced, with publications such as Ruiz & Madariaga (2018) showing a recurrence analysis of the Chilean subduction zone. The behaviour and dynamics of subduction earthquakes are also well understood, with studies laying the foundations for understanding the role that mechanisms and physical parameters such as sediment dehydration, pore pressure (e. gr. Moreno *et al.,* 2014), coupling degree, gravity anomalies or friction (Molina *et al.,* 2020) play in subduction earthquakes. Thus, some of these parameters may be used in the physical restriction of random slip models.
The methodology developed in this instance is based on mathematical tools and algorithms designed for disciplines that study tsunami hazards, that have been modified to suit our needs.

The main tool used and redesigned for this study is the Logic Tree structure, a relatively recent development in the Probabilistic Tsunami Hazard Assessment field, starting in 2007 with Annaka *et al.,* 2007, who set the groundwork for future tsunami hazard work. However, the use of this structure for the study of past earthquake is far more contemporary, with Goda & Song (2016) and Fukutani *et al.* (2018), whose studies centered in the Tohoku-Oki earthquake and its related uncertainties from a probabilistic point of view, based on the Logic Tree. On the other hand, from the standpoint of historical tsunami, sedimentology, stratigraphic and coastal deformation data, its usage in reconstructing and estimating seismic scenarios for the case of central Chile is recent, with examples found in Dura (2016), Hocking *et al.* (2017), whose work in the role in diatom and other biological indicator data has been of the utmost importance in defining tsunami restrictions, among others.

This study, being based on computationally costly algorithms and computations (e. gr. LT structure, random slip distribution generation, computation of vertical deformation) and on tsunami modelling, needs to be as streamlined as possible, making every step as efficient as it can be, so as

to not waste time and/or computational resources. Great efforts were put into creating an efficient management of the large number of output created by the methodology, that can reach easily orders of tens of thousands or even hundreds of thousands of files, tallying up to hundreds of Gb of data. However, more work can be done, especially in modelling tsunami in a more cost-effective way. An option is to use state of the art models that parallelize and/or use the GPU instead of the CPU to compute the propagation of the tsunami, accelerating considerably the computations. Qin *et al.* (2019) and Galaz *et al.* (2019) show the advantages of using GPUs in tsunami modelling and the acceleration of said processes. Other advances in this regard are the inclusion of ray-tracing methods in wave propagation (Schambach *et al.,* 2019) and using A.I. in the study of tsunami hazards (Song *et al.,* 2019). These advances can be proposed as a continuation of this study.

1.2 Research question, goals and objectives

Research questions

- Is it possible to characterize the source of a historical tsunami, such as Valdivia 1960, in the Chilean margin based on restrictions given by paleo tsunami records?
- Does a logic tree structure approach to generating random source models allows to converge to the source model of a historical tsunami?

Goals and objectives

- Main goal:
  - To estimate and characterize the source model of the Valdivia 1960 earthquake using a logic tree approach to generate random slip distributions and restrictions arisen from tsunami and geodetic records.
- Specific goals:

- To implement a numerical methodology for the generation of random slip distribution models using a logic tree structure.
- To implement a methodology to test the compliance of random slip distribution to deformation and tsunami data, taking into account their geographical distribution and density.
- To assess the resolution, capabilities and limitations of the methodology depending on the data availability.
- To generate random slip models and to restrict them using data compiled from the Valdivia earthquake.
- To model tsunami propagation and estimate wave heights and run up in areas where data is available, finally filter the models using the data.
- To analyze the remaining models and to study the convergence of them with the known and published Valdivia 1960 source models.

1.3 Hypothesis

As a hypothesis, it is proposed that:

• The estimation of the most probable source model of the Valdivia 1960 tsunami (slip distribution on the rupture and rupture limits), obtained via analysis of random source models generated using a logic tree structure approach will converge on a solution that satisfies the available paleotsunami data, geodetic and tide gauge derived.

Chapters 2 to 4, are from the publication titled "A Stochastic Approach to the Characterization of the Seismic Source of Historical and Paleotsunami"

Chapter 2

Methodology

2.1 Generation of random tsunami sources

2.1.1 The logic tree approach

In order to produce hypothetical rupture scenarios, random slip distributions are generated based on a combination of fault defining parameters. This fault is subdivided into a matrix of $n \times m$ rectangular subfaults (whose dimensions depend on the number of subfaults along-dip and strike, the length of the fault and its aspect ratio) and, following the methodology proposed by LeVeque *et al.,* 2016, a random slip value s_i is assigned to the ith subfault, defining a $\mathbf{s} \in \mathbb{R}^n$ slip vector. These slip values have a joint lognormal distribution given by the exponentiation of a Karhunen–Loève (K-L) sum (Mai *et al.,* 2002, Leveque *et al.,* 2016), representing a linear combination of eigenvectors and eigenvalues of a covariance matrix $\hat{C} \in \mathbb{R}^{n \times n}$, which in turn is a function of the distribution of the subfaults that define the fault geometry (more details can be found in section 6.2.1).

Said distribution of subfaults for the generation of stochastic earthquake scenarios is achieved with a LT structure (fig 2.1), most commonly used in PTHA (Fukutani *et al.,* 2014, Goda *et al.,* 2016, Fukutani *et al.,* 2018), but in this study it is used as a tool for generating random past

scenarios.

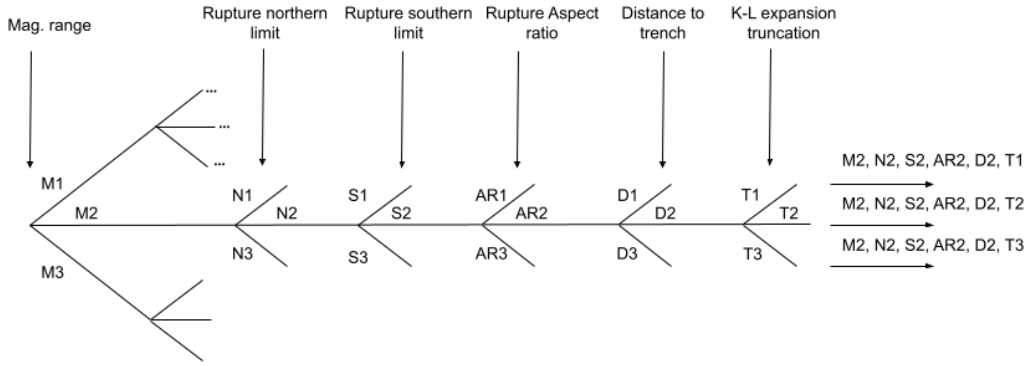


Figure 2.1: Logic tree structure example used in the creation of random source models. The LT consists of the branches for the combinations of source parameters used to define the random models, magnitudes, northern and southern limits of the rupture, aspect ratio of the rupture, distance to trench of the updip limit and truncation level of the K-L expansion.

2.1.2 Choosing branch values

In order to construct a LT structure that generate appropriate random slip distributions for the event of interest, the election of a correct range of values for the different branches of the LT structure is crucial to be able to adequately estimate the most probable seismic source and not to over utilize computational costs to do so. The steps proposed to define the values are as follows:

- Define a number of discrete subfaults to divide the rupture zone (Note that the number of subfaults increase drastically the computational costs)
- Define a maximum rupture length according to the span of deformation data including a buffer zone to the north and south to accommodate for possible variations in range
- Estimate the earthquake´s seismic moment with scaling laws (e.g. Abe *et al.*, 1975, Geller *et al.*, 1976, Skarlatoudis *et al.*, 2016, Thingbaijan *et al.*, 2017) and definine a range according to the uncertainties of the scaling laws.
- Estimate a mean slip value according to each moment magnitude of the range.

- Define a range of distances to the trench according to local geology and seismic properties (e.g. Lay *et al.*, 2012)
- According to empirical relations (e.g. Kanamori & Anderson, 1975, Purcaru *et al.*, 1982) select a range of aspect ratios for the configuration of the number of subfaults defined previously
- Finally, select a level of truncation for the K-L expansion sum, the larger the number, the larger the deviation from a homogeneous slip distribution (and the larger computational costs), numbers around 20 work well

The amount of these values define the number of correlation matrices to be computed, albeit not necessarily the total number of random distributions to be generated. It is possible to choose, for each combination of parameters of the LT, an independent number of times that the random distribution in the K-L expansion is drawn, thus generating a larger amount of distributions for each combination. Once the random distributions are generated, a vertical deformation field is computed for each one using Clawpack's implementation of the Okada (1985) model for computing seafloor deformation (dtopotools v5.5.0) package on Python (Clawpack Development Team, 2018, Mandli *et al.*, 2016).

2.1.3 Restrictions

A vast number of models is necessary in order to compensate for empirical and random uncertainties, as well as to ensure the heterogeneity of slip distributions. This type of slip distribution is preferable compared to homogeneous ones, as the latter tend to underestimate tsunami intensities (Carvajal *et al.,* 2017).

LT structures are capable of generating a vast number of random models, however, due to their random nature, not necessarily all of them will be able to represent a natural or possible earthquakes. To compensate for this, it is necessary to include and impose restrictions to them in order to dismiss those that are not plausible (do not follow empirical nor source-scaling laws, e.g.

Kanamori & Anderson, 1975, Thingbaijam *et al.,* 2017), and to limit the computational costs and complexity using only models that comply with physical restrictions and data for the estimation, both geological and historical records and accounts.

As discussed in the previous section, empirical relations and source scaling laws are used in the first place to choose the branches of the LT, to ensure that the earthquake scenarios generated agree with physical characteristics of subduction earthquakes.

After the generation of physically possible earthquake scenarios, the next stage is combining sedimentological and deformation data from studies of past earthquakes and tsunami, and the slip distributions given by PTHA. Every random model created with the LT is subjected to two types of restrictions, one group based on deformation data and another based on tsunami inundation and/or run-up data. These tests are performed in a staggered fashion (see fig. 2.2 for a schematic of the restriction process), assessing the compliance of the models to deformation data first, reducing the number of source models needed to be modelled with tsunami modelling software for filtering with tsunami observations. By the nature of this method, each successive filter is applied to a smaller number of models, thus decreasing the toll on the computer, and decreasing the run time of the methodology. Each type of restriction is further subdivided into different steps to be described in the next section. It's important to note that the order of the restriction is important and affects the final result

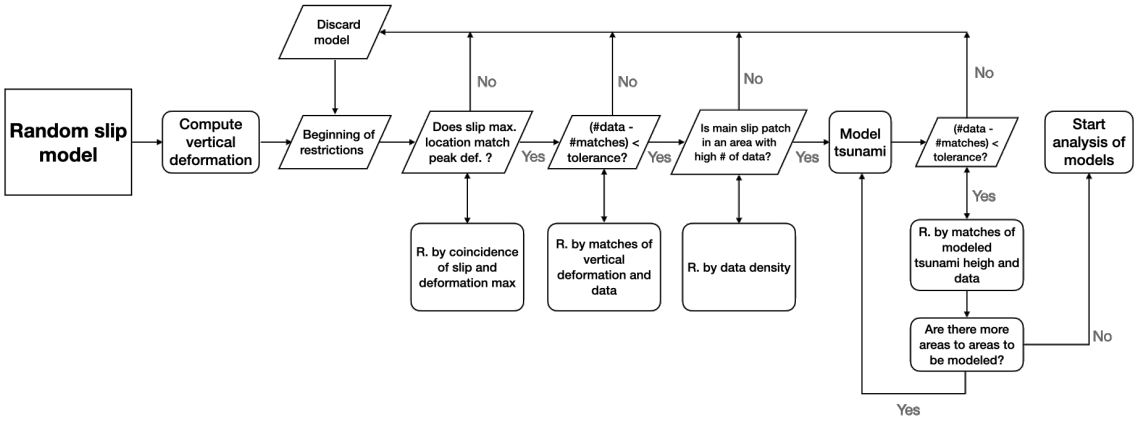


Figure 2.2: Simplified flowchart of the restriction process.

Deformation data restrictions

This first type of restriction is subdivided into three steps, with the possibility of further dividing the last one into as many steps as the user would want, depending on the data availability. The first is to divide the maximum length of the possible rupture zone into n bands (see fig. 2.3 for example) depending on the length and quantity of data points available. For most earthquakes with rupture lengths higher than 500 km, 5 bands is recommended. The relative number of data points in each band is then computed, and the one with the densest deformation data is selected as a reference, from which data density value the relative density values of the other bands will be computed. If two bands have the same data density, the one with higher tsunami data density is preferred over the other. From north to south, the third and second bands are selected for each synthetic case, respectively, since they are the ones with the higher number of deformation data points (32.66 % of the data for first synthetic model and 32.50 % for the second. See figure 2.3).

The first instance of restrictions considers a comparison between the absolute values of deformation and the slip. The latitudinal slip profile (slip fields integrated along rows of equal latitude) of the random models is compared with the profile of absolute values of deformation (see right panels of fig. 2.3 (a) and (b)). As larger deformation corresponds to larger slip (Okada, 1985), random models whose maximum slip values are located in the same band as the maximum deformation data pass this filter. Models whose maximum values are located m bands away from the band of the maximum deformation data are penalised, discarding *f(m)* models, where *f(m)* is a user defined function. By default, a linear function is applied in the form of $f(m) = m \times 0.2$, i.e. discarding the 20 % of models whose maximum slip values are located one band away from the maximum deformation value. This previous (to comparing point-to-point data and random models) instance of restriction is done in order to compensate for possible deficiencies and/or gaps in data coverage. An area with a data gap will not be able to discard models in the next restriction step, as it does not have any means to compare data to the vertical deformation of a given model. This means that

if this filter is not applied, possible models with high slip values in this area may appear as these can comply with the rest of observations.

The last step is the most important and can be subdivided into several chained restrictions levels, according to data quality. Vertical deformation fields computed from the random slip distributions are compared point-to-point with deformation data, testing if the computed vertical deformation values are within a tolerance level to the data. The number of matches is compared to the total number of points and if the amount is higher than a threshold (tolerance defined *a priori* by the user), the model passes the restriction. If not, it is discarded. If the user chooses to, this step can be performed in a staggered fashion, dividing the data into groups and forcing the models to comply with the data in batches. If the data has different uncertainties, the best quality data should be used in a first batch with a low tolerance, and then continue restricting models with lesser quality setting a higher threshold, thus the models will have to have more points comply with the data.
It is important to note that if the user chooses to lower or increase the threshold value, for the number of matches compared with the number of data or the tolerance to consider a point of the random model to be a match to the data, respectively, the number of models that will clear the restriction will be higher. There is no *a priori* relation to estimate the quantity of models that will comply in excess given a certain increase in the threshold.

Tsunami data restrictions

Two previous steps are to be taken before the tsunami modelling process. As the number of data points affect the resolving power of the methodology, bands with higher deformation data density are given more weight in the restriction process, defining a penalty for those that have less with respect to the reference band. This way, for band i, an amount corresponding to a function of the ratio between the i-th band density and the reference density of random slips distributions whose maximum slip value is located in the i-th band is discarded. By default, the methodology discard the $\frac{\rho_i}{\rho_{RB}}\%$ of the models (where ρ_i and ρ_{RB} are the data densities of the i-th band and of the

reference band, respectively), but this value can be changed according to the user. This ensures that most high valued slip features on the remaining models are located in areas where the data allows for resolving power.

After every step of the deformation restriction, the next step is to define the areas to model tsunami and the order in which they are to be modelled and restricted. This step depends on the availability of high-resolution topobathymetric grids (COMCOT accepts only up to 12 nested grids) and the location of tsunami observations. The methodology is highly sensitive to the order of tsunami restrictions, and for each area modelled, the data has to comply locally (only in the modelled area) and not globally (the whole are where there are data points), possibly skewing the estimations. For this reason, the modelling order starts with areas where tsunami data values are higher and there is higher density of data. Values have priority over density. For example, in the case study, higher values and densities are located near Corral ($\sim 39.8°$ S) (fig 3.1 (b)) and thus this area is selected as the starting point for tsunami modelling and restriction. Other areas are modelled in decreasing order of values and density, with each step requiring fewer models to be computed. One possible alternative, albeit a discouraged one due to its extremely high computational cost, is to model the entire area covered with tsunami data in as many batches as the grid availability requires, and then restricting checks the compliance of the models to the data. Computational times and the amount of disk space required increase steeply with each extra tsunami model, so this alternative can prove to be unwieldy if the amount of models surpass a hundred or two hundred.

Although inundations map can be computed for onland tsunami data, this is discouraged as the computational costs are much higher than estimating off-shore wave heights and computing virtual tide gauge time series. In all of the examples shown in this work, the latter option is used. Once the models are done, the maximum values of the virtual tide gauges are compared to the wave height, with a tolerance level defined by the user and the possibility to account for local tidal changes. If the modelled value is higher than the data point $\pm$ the tolerance (or if the inland location of the point is inundated) the model passes, else, it is discarded.

2.2 Analysis of models

Once the random source models have been through every step of the restriction process, all the remaining slip distributions that cleared them are subject to two types of analysis in order to estimate the most probable source. First is a frequentist analysis of the branches that generated the remaining models, and second is the characterisation of the slip distribution and vertical deformation fields based on the slip values of the subfaults that compose the remaining random models.

2.2.1 Characterisation and estimation of the seismic source

A frequentist analysis of the branches show the most probable source defining parameters, in terms of the most repeated value, indicating the moment magnitude, limits and aspect ratio. These values are used to define the span of the rupture zone and the distribution of the subfaults that compose the fault. For the characterisation of the slip distribution, the most probable value of the i-th subfault is estimated computing a probability density function (PDF) of the values of the i-th subfault of every model that passed the restrictions, and computing the slip value that maximises this PDF, ensuring a subfault-wise maximisation of the probability. The resulting slip distribution will not, necessarily, have the most repeated moment magnitude computed in the first estimation step. To remedy this, the field is scaled by a multiplicative factor so that the sum of the contribution of each subfault gives the desired moment magnitude Mo_{MP}, given by the most probable (MP) moment magnitude. Thus, if the moment resulting from the characterisation is Mo_c, the final slip distribution $\vec{S}_f$ is obtained by multiplying the characterised slip vector $\vec{S}_c$ by the ratio of the seismic moments of the most probable (Mo_{MP}) magnitude and the characterised slip vector moment, so as if the new moment of each subfault is computed and the contribution of each of them is summed, the desired moment magnitude is obtained. This is computed with,

$$\vec{S}_f = \left(\frac{Mo_{MP}}{Mo_c} \right) \vec{S}_c \tag{2.1}$$

The resulting distribution is finally smoothed using a Gaussian filter and subsequently scaled to obtain the desired moment magnitude.

2.3 Testing the methodology

Several resolution tests are performed with varying distributions of data in order to test the capabilities and limitations of the methodology to characterise the seismic source and to assess the dependence to deformation and tsunami data availability, distribution, and density.

2.3.1 Synthetic earthquake

Two random synthetic slip distributions with different rupture lengths and spatial characteristics were generated, from which vertical deformation field are computed with the Okada (1985) solution, then tsunami and deformation data are sampled from them and used as input for the restrictions in the estimation process. The synthetic distributions are done to test the importance of the data and the different steps in the restriction process.

Deformation data from the synthetic models is sampled with a uniform latitudinal distance along the coast of the rupture zone, 120 points where sampled for each case. For tsunami height data, tsunami propagation is computed solving linear and nonlinear shallow water equations using numerical model COMCOTv1.7 with a nested grid system of four levels with increasing topobathymetric resolution. Areas historically inhabited on the Chilean coast along the rupture zone of the synthetic models are modelled using fourth-level grids with resolutions of 0.09' and 0.04', where available. Grids were obtained from Global Multi-Resolution Topography (GMRT) (Ryan *et al.,* 2009), National Center for Hidrographic and Oceanographic Data (CENDHOC in Spanish) of the Chilean Navy Hydrographic and Oceanographic Service (SHOA in Spanish) and SRTM 1 Arc-Second Global (Farr & Kobrick, 2007). Tsunami heights are sampled computing virtual time series of ocean surface heights near locations historically affected by tsunami and where real tsunami data

has been measured. Both sampling approaches were chosen to simulate possible *in situ* measurement campaigns. 19 points were sampled for the first synthetic model and 22 for the second case. Figure 2.3 show the distribution of synthetic data.

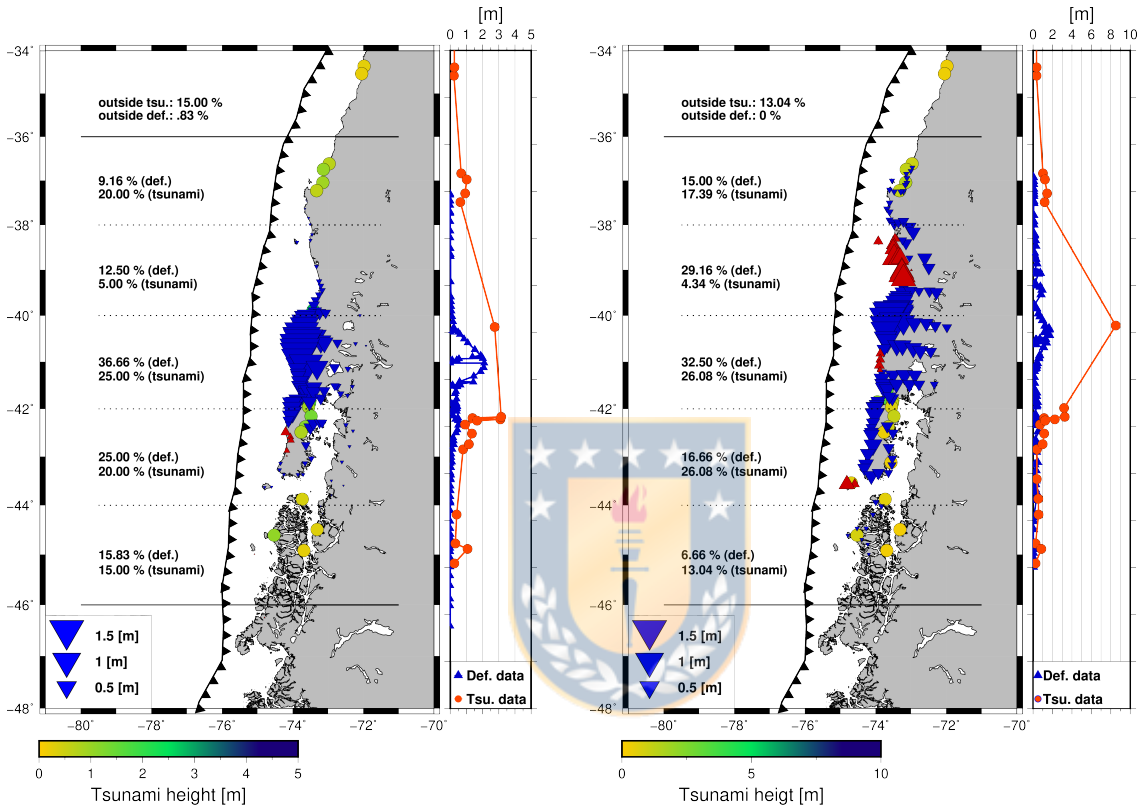


Figure 2.3: Deformation and tsunami data sampled from two synthetic models used to test the methodology. Possible rupture zone is divided into 5 strips, where data density is estimated for the restriction process. Profiles to the right of the maps show the absolute value in meters of the deformation sampled (green dots) and the values of tsunami height on virtual tide gauges sampled near inhabited coastal areas (blue dots). Left figure shows data sampled from Synthetic Model I and right shows data from Synthetic Model II.

2.3.2 Synthetic model estimations

Following the steps described in section 2.2, value ranges for the six levels of the LT for both synthetic models are chosen, and the number of iterations for random distributions is set so as to have 200,000 random models generated for both cases. Then, both are subject to the methodology, clearing all restrictions (see table 6.3 for an overview of the user defined restriction values) respectively

14 and 9 models, obtaining the results shown in figure 2.4. The normalised latitudinal slip profiles show correlations of $\rho_1 = 0.34$ and $\rho_2 = 0.55$ with root mean square errors of $\text{RMSE}_1 = 0.50$ and $\text{RMSE}_2 = 0.35$ for synthetic models I and II, respectively.

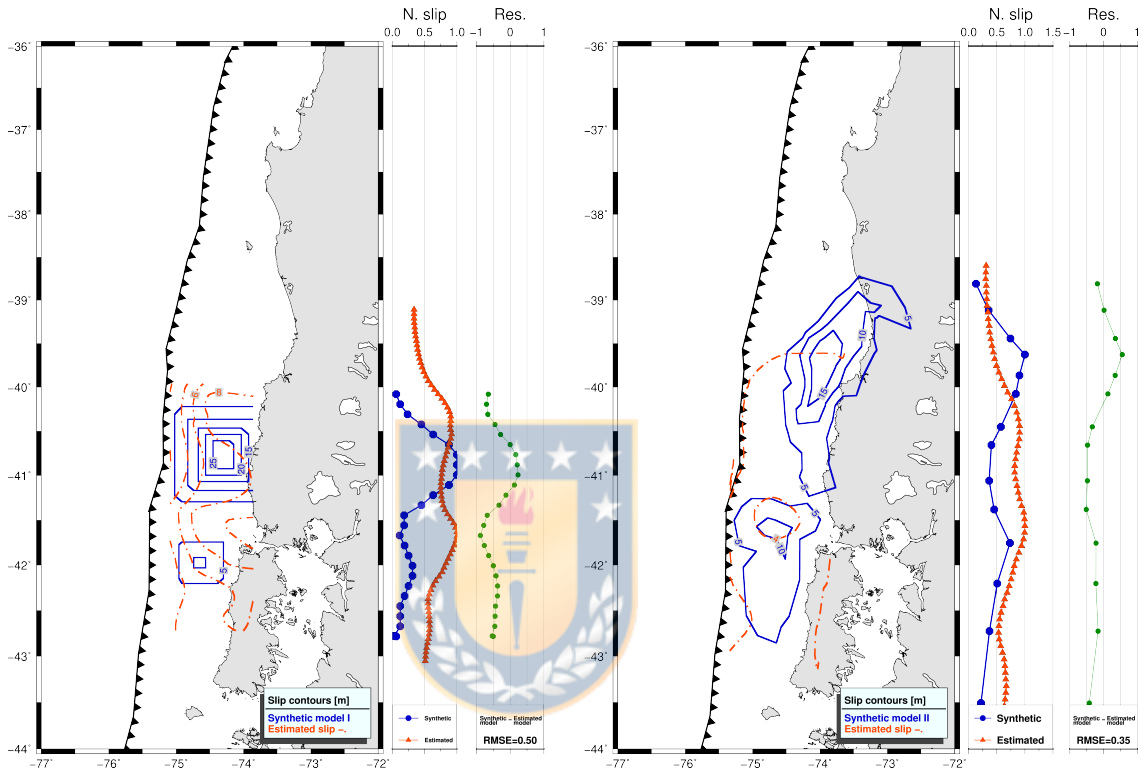


Figure 2.4: Comparison of the synthetic slip distributions (solid blue contours) and the estimated results (dotted orange contours). Latitudinal profiles of normalised slip for the estimation results (orange triangles and lines) and synthetic slip distributions (blue dots and lines) show the along-strike distribution of slip. The plot to the right show the latitudinal profile of residuals between estimations. Left figure shows the results of Synthetic Model I and right shows the results of Synthetic Model II.

Table 2.1: Restriction tolerances used in the synthetics test estimation process. The first restriction penalty function and the number of bands were set as defined in earlier sections.

	Deformation Tol.	Match Tol.	Tsunami Tol.	Match Tol.
Synthetic Model 1	0.3 [m]	10	0.5 [m]	0
Synthetic Model 2	0.3 [m]	10	0.8 [m]	0

Estimation and characterisation of the slip distribution by the proposed methodology resolve the location of the main patches of slip for both synthetic models, although presenting more spread on the distribution of these values and not reaching the same slip maximum values. This can be explained by the fact that estimated models tend to spread the total seismic moment onto more subfaults, instead of concentrating it in the peaks of displacement as is the case of the synthetic models used in this test. Thus, for the same moment magnitude, estimated slip distributions tend to under represent the maximum values of displacement. This means that estimated magnitudes are often overestimated. This effect is attenuated when tsunami data for the restrictions is taken into account.

Regarding the quality of the estimated seismic sources, the obtained slip distributions are able to locate the number and location of the slip maxima features (with some leeway), however it tends to underestimate the values of these maximum slips. For example, this results are good for studies of recurrence of ruptured areas, as one could characterise the zones that are more prone to present higher slip values in a given rupture zone, if one could characterise the historical earthquakes that took place there.

Chapter 3

Valdivia case study

3.1 Introduction

The biggest earthquake in recorded history, known as the great Valdivia earthquake, took place on May 22, 1960 registering a magnitude of 9.5 Mw. This event ruptured an area in the interface of the Nazca and South American Plates with a length of over 1,000 km, from the Gulf of Arauco in the North (∼37° S) and the Ofqui Isthmus, in the South (∼46° S)(Cifuentes, 1989, Cifuentes & Silver, 1989, Barrientos & Ward, 1990, Lomnitz, 2004, Moreno *et al.,* 2009). It is hypothesised that this earthquake is comprised of three subevents, that ruptured the entire segment in a rapid succession, starting in the northern part of the segment and continuing southward over a span of approximately 15 minutes (Cifuentes, 1989, Cifuentes & Silver, 1989).

Coseismic and geodetic data from this event in the form of triangulations and relative sea level changes (Plafker & Savage, 1970) (shown in figure 3.1 (a)) is used as deformation input in the restriction process, accounting for 150 points. Historical accounts and sedimentological measurements were used to determine tsunami heights (courtesy of Dr. Winckler, original data) (fig. 3.1 (b)). As all 103 points available of tsunami data is located on land, off-shore wave height

values (located on the closest node to the data point on the highest resolution topobathymetric grid available for the area) are estimated using the inverse of the equations proposed in Smart *et al.,* 2016. The value ranges used as input for the LT in the generation process are shown in Table 3.1. The number of iterations was set as to have 200,000 random slip distribution at the start of the restriction process. The values for the restriction parameters are shown in Table 3.2

Table 3.1: Value ranges of the Logic Tree for the generation of random slip distributions for estimating the seismic source of the Valdivia 1960 earthquake. Moment magnitudes range from 9.3 to 9.7 in 0.1 increments, complexity values (N) range from 20 to 24 in unit increments, aspect ratio values (AR) range from 2 to 5 also in unit increments.

Branch	Mw	N	AR [L/W]	D. [km]	N. l. [°]	S. l. [°]
Range	[9.3,9.7]	[20,24]	[2,4]	[0,60]	[9.3,9.7]	[9.3,9.7]

Table 3.2: Restriction tolerances used in the estimation process for the Valdivia 1960 earthquake slip distribution. The first restriction penalty function and the number of bands were set as defined in earlier sections.

	Deformation Tol.	Match Tol.	Tsunami Tol.	Match Tol.
Valdivia earthquake	0.3 [m]	15	0.4 [m]	0

3.2 Results

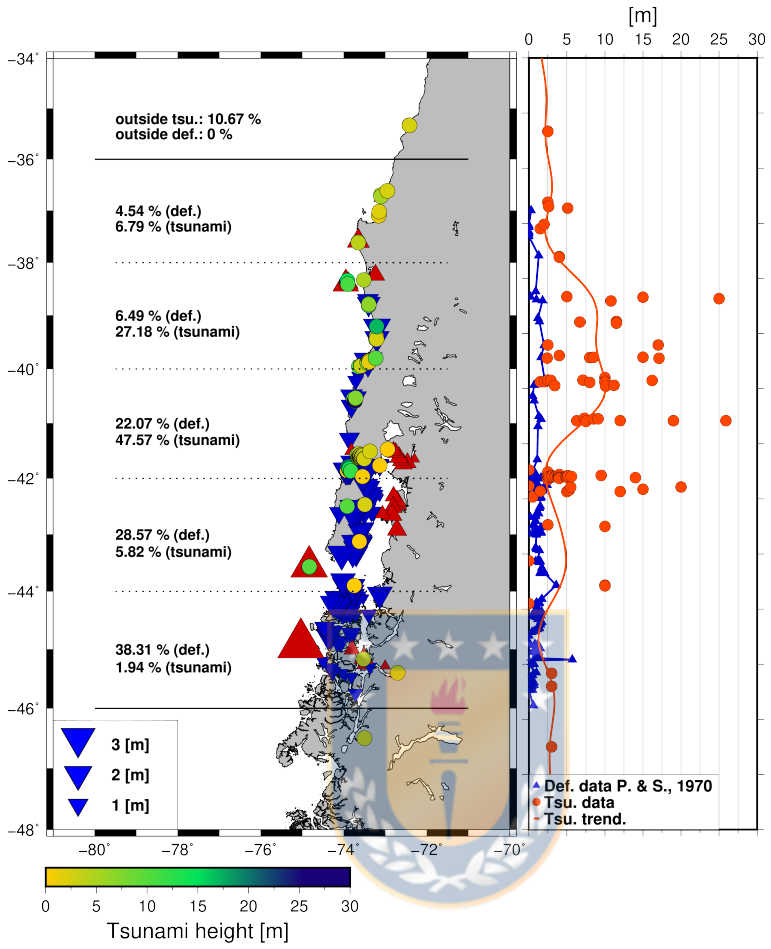
After the restriction process, from the starting 200,000 models, 20 cleared every step. Subjecting the set of remaining models to the estimation process yield the results shown in figure 3.2.

The estimations of the seismic source determined by our methodology are compared to previous slip models computed from inversions of seismic, deformation and tsunami data (Barrientos & Ward, 1990, Moreno *et al.*, 2009, Fujii & Satake, 2013). Our slip distribution of the Valdivia 9.5 Mw earthquake show that the most probable rupture spans from the Gulf of Arauco ($\sim 37°$ S) to the south of the Guaitecas Archipelago ($\sim 45°$ S), reaching farther north with larger slip values than previous slip models, while sharing the southern limit. Results show a large primary slip patch in the northern part of the rupture zone, between 38° and 40°, with maximum slip values up to 35 meters, consistent with displacement concentrated in the northernmost part of the rupture

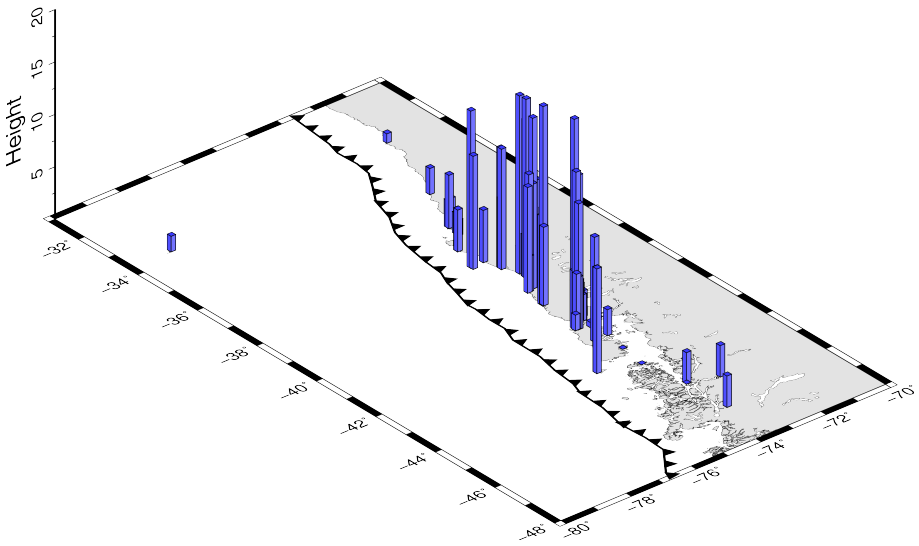
in other slip models (albeit not as large) (fig 3.2 (a)). In contrast with Moreno *et al.,* 2009 and Fujii & Satake, 2013 results, our large slip values do not extend as far south, while they do reach farther north than the reference slip models. South of Corral ($\sim 40°$), our estimated values are homogeneous with values in excess of 10 meters, larger than Barrientos & Ward, 1990 and Moreno *et al.,* 2009 (inversions that share most of the input data), but in line with the values obtained with Fujii & Satake, 2013. Normalized latitudinal-integrated slip profiles (fig 3.2 (b)) show that both our stochastically estimated slip distribution and the three previous inversion results share the latitude of the maximum slip values, however our results do not resolve a secondary peak at $\sim 41°$ S shown in Moreno *et al.,* 2009. Both profiles in fig 3.2 (b) show a correlation value of $\rho = 0.50$ and the residuals a RMSE of RMSE=0.36. Due to the construction of the rectangular fault, deep slip values obtained in the inversions at $\sim 39°$ S cannot be resolved with our method, however, the excess of slip near the down dip rupture boundary near $\sim 39°$ and in the northern limit can be related to the restrictions accounting for this deep values.

Deformation values tend to show larger uplift in the southern area compared with other models, resolving the uplift measured in Guafo and Guamblin Islands, nevertheless, due to the lack of deformation data off-shore, it is not possible to delve much deeper into these results.

Overall, slip distribution estimations results obtained from the synthetic tests and the Valdivia case study are good, with the three of them being able to resolve, with some tolerance, the location and quantity of the main slip features. Although maximum slip values are underestimated, the magnitude of the estimated slip distribution for Valdivia match the Mw=9.5 value (Cifuentes & Silver, 1989). This underestimation tends to under represent tsunami heights and inundation, but the estimation results shine a bright light on the areas that ruptured with greater amplitude in past earthquakes.



(a) Deformation and tsunami data distribution



(b) Tsunami height data distribution

Figure 3.1: (a) Deformation and tsunami data available for the estimation of the seismic source of the 9.5 Mw earthquake of Validivia, 1960. Deformation data (blue downward triangles show areas of subsidence and red upward triangles show uplift) is from Plafker & Savage, 1975 and tsunami data was provided by Dr. Winckler at Universidad Católica de Valparaíso. Profiles to the right show the behaviour of data amplitude along latitude. (b) Tsunami heights distribution.

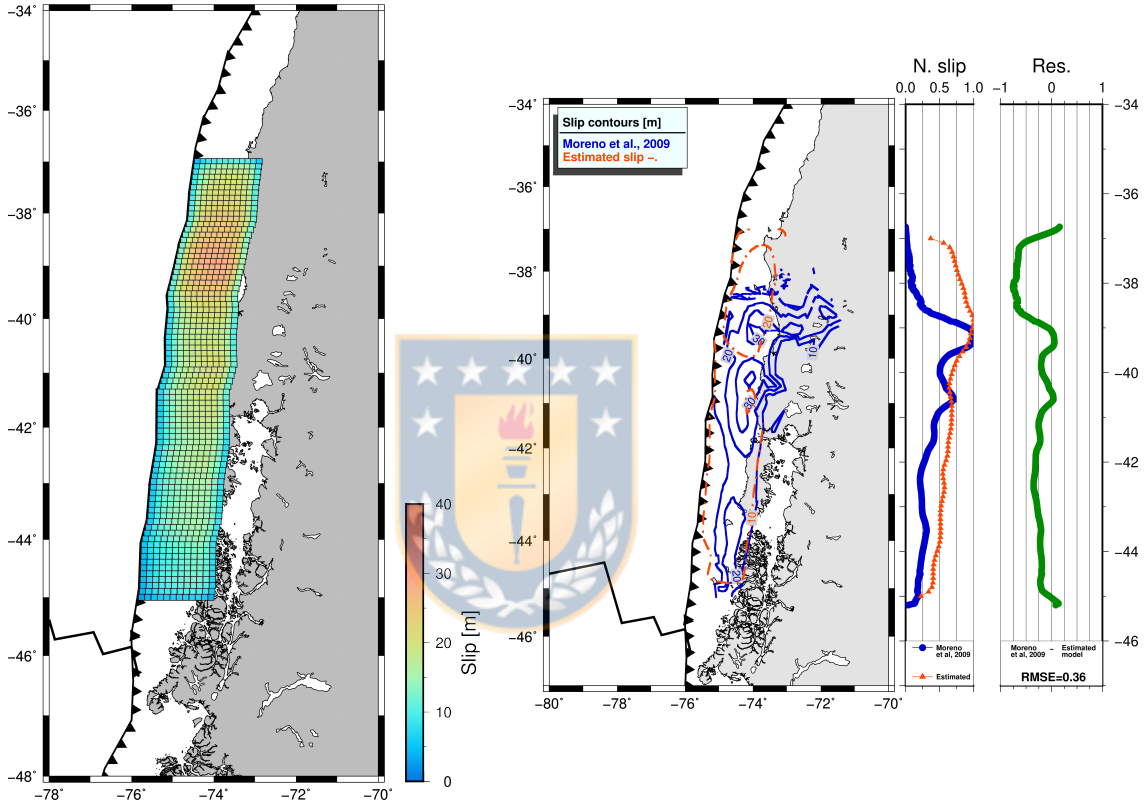


Figure 3.2: Estimation of the most probable seismic source model for the Valdivia 1960 9.5 Mw earthquake in terms of the deformation and tsunami data available. Left figure shows the slip distribution of the smoothed subfault-wise maximisation of the PDF. Right shows a comparison between the estimation (dotted orange contours) with the inversion result from Moreno *et al.*, 2009 (solid blue contours). Normalized latitudinal slip profiles show the compared behaviour of the slip distributions, showing the main slip patch in the northern section of the rupture and a decrease in values to the south. Residuals show that the main differences in estimated displacement are in the northern limit of the fault)

Chapter 4

Discussion

The most critical part of the estimation process is the definition of the LT branches, which will ultimately decide the number of random models generated. Although convergence to an estimation will always happen, an inadequate definition of branch parameters could severely hinder the time that takes to arrive at the estimation, increasing the numbers of models to be generated and thus the number of vertical deformation models that need to be computed, leading to more models that need to be restricted and computed for their respective tsunami propagation. The estimation of the slip distribution, from which is derived the moment M_0, the magnitude Mw, and the northern and southern limits of the fault, depending on the aspect ratio desired for the fault, is heavily dependent on the distribution and span of deformation data and/or *a priori* knowledge, thus large uncertainties or gaps in data can lead to flawed estimations, that may satisfy the deformation data and, only locally, the areas used to model tsunami, however a global convergence may not happen. Depending on the dimensions of the earthquake, different scaling laws (e.g. Murotani, *et al.,* 2008, Murotani *et al.,* 2013, Skarlatoudis *et al.,* 2016, Thingbaijam *et al.,* 2017) should be used to estimate the magnitude, defining a plausible range for Mw depending on the scaling law used. A safe way to choose the limits for giant earthquakes (Mw∼9) is to define them as the limits of

the seismic segment, assuming that they ruptured along its entirety. For smaller earthquakes, the limits can be chosen by the range of deformation data available, if the last available data location is south to northern limit of the segment, and vice versa.

The role of the data for the estimation is, as expected, of the utmost importance. Evenly distributed deformation data, especially if it is not located only on the coast, gives more resolving power to slip features, leading to more concentrated slip patches and a larger deviation from homogeneous distributions. The lack of data, deformation and/or tsunami related, in an area of the rupture zone leads to possible appearance of artefacts in the estimation, due to the inability to discard random models containing large slip values in such areas.

The order of the deformation data used as input in the restriction process does not impact in the results, however, the order in which tsunami data is used to restrict heavily affects the results, possibly skewing the final estimations. This is due to the necessity of dividing the area into subareas to model tsunami (due to limitations in the number of grids that can be nested) and restricting, a process that leads to random distributions complying locally to tsunami data and not necessarily to the global distribution of data. This brings the imperative requirement of choosing carefully the order in which said divisions of the area are to be modelled.

Aligned with the importance of data for the estimation process, choosing the right restriction parameters (thresholds/tolerances) does influence the result. As setting higher thresholds allows for more models to clear the filtering process, the final estimation will include a larger variability of models, resulting in more homogeneous slip distributions, not being able to resolve in a finer way slip patches. This will also mean that, for the same magnitude, maximum slip values will be lower, as the seismic moment will be more spread out through a large number of subfaults. From the point of view of computational times, higher thresholds and higher number of models clearing restriction steps mean a higher number of models that have to be subject to subsequent filters and tsunami modelling, increasing the toll on the computer.

Thresholds values for checking point-to-point matches should never be equal or higher than the

highest uncertainty value of the data points. In the case of the tolerance for the number of matches in the comparison, this value should not be higher than the minimum number of data points in a given band, in the case of deformation restrictions, or the quantity in a modelling area, for tsunami.

In parallel with the effect that tolerances have regarding the number of models that clear the restrictions, the number of data points has a similar effect. Fewer data points allow more random models to pass the filtering process, inducing a more homogeneous slip distribution estimation. Thus, for the study of past earthquakes from more than a century ago, where the retrieval and measurement of deformation or tsunami inundation is hard, the user could expect the solution to not show with high resolution the main slip features, and would underestimate the maximum slip values. However, this estimation would still show the location of this features.

The stochastic nature of the estimation methodology requires the generation of a vast number of random rupture scenarios, in order to deal with uncertainties, creating an important toll to the computers that perform the calculations, generating in some cases up to hundreds of Gb of data. The restriction process is far more efficient, and its run time is many times faster than the generation process. In the case of the tsunami modelling, the run time is constrained in part by the modelling software of choice; COMCOTv1.7 is a fast and lightweight software. However, depending on the tsunami data available (on-shore and/or off-shore), one could generate only virtual tide gauge time series and/or wave height maps from the tsunami models to perform the tsunami restrictions, further increasing the amount of output data generated and the toll on the computers and drives. It is recommended using of a method such as the inverse of the proposed by Smart *et al.*, 2016 to transport an on-shore tsunami measurement to a point on the ocean corresponding to a node of a high resolution topobathymetry grid, thus removing the need of computing an inundation map, and leaving the necessity of computing virtual tide gauges time series in the forementioned off-shore point.

If the earthquake caused landslides, underwater or not, that in turn may have caused a tsunami, it can affect also negatively the estimation methodology, as this possible addition in tsunami height

points may skew the estimation to the specific area affected by it, forcing additional slip patches to compensate for this extra wave height (e.g. Gusman *et al.,* 2019). The same can be said to amplification phenomena caused by site effects in bays or fjords, where this addition in observed height may cause that the estimated model present a high slip feature in this area.

Future improvements for the methodology may gravitate to the inclusion of time dependence of the rupture, including a branch to the LT structure that accounts for a dynamic rupture model instead of a static model for the tsunami propagation computation process. However, due to the large uncertainties inherent to past data, the resolution and fitting improvements are expected to be marginal.

Chapter 5

Conclusion

Estimated seismic sources results from this methodology are of good quality in terms of locating slip patches, albeit they underestimate overall slip maxima from these patches. Given the availability of quantitative and qualitative deformation and tsunami data, be it quantitative or qualitative, characterisations of heterogeneous slip distributions is therefore possible. These results can give valuable insight into characterising the behaviour of a specific past earthquake, or, if used for every earthquake recorded in a given rupture zone, this methodology can shine light into possible slip distribution patterns in the rupture zone, such as the characterisation of areas with recurrent high slip patches.

This methodology was calibrated with two synthetic slip distributions where it was tested the effect of the order of restrictions, the truncation level of K-L expansion and the use of qualitative and/or quantitative data, in order to assess the methodology's sensitivity to these variables. The estimations of these synthetic tests converged into a solution that correctly estimated the number and location of the main slip features, correctly estimating the magnitude of the earthquakes, but underestimating the maximum slip values. After calibration, this estimation process was applied to the 1960 Valdivia earthquake, to test the methodology in a real world scenario. In this case,

the estimation also converged into a solution that shared the number and location of slip patches showed by other authors (Barrientos & Ward, 1990, Moreno *et al.,* 2009, Fujii & Satake, 2013), and also correctly estimating the moment magnitude of the earthquake, albeit underestimating the maximum slip of the models computed with inversions.

Some drawbacks described in previous chapters can be rectified with the advent of new computational technologies and more powerful CPUs and the possible parallelisation of the creation of random slip models anrd/or the computation of tsunami models. In the same page, the possibility of using different software for tsunami modeling such as tools from ClawPack's GeoClaw that offer Adaptative Mesh Refinement (AMR) and the possibility of including more grids may solve the issue of having to divide the rupture zone into different areas to model tsunami locally.

Although this methodology was initially designed for the Chilean subduction zone, the implementation for it to work in any subduction zone was in mind when writing the code, thus by changing only a few parameters, earthquakes from other subduction zones can be estimated.

Chapter 6

Annex material

This chapter delves into more detailed explanations of topics exposed in previous chapters and contains mathematical formulations and intermediate results not show in the publication.

6.1 Tsunami and subduction earthquakes

Tsunami waves are usually defined as gravitational waves with periods ranging from 10^2 [s] to 10^4 [s]. They are the result of rapid transient processes such as earthquakes deforming the seafloor, landslides and collapses (above and underwater), volcanic eruptions, among other causes. These process results in the formation of long gravitational waves, with wavelengths λ within the range 10^4-0^6 [m], usually corresponding to the horizontal dimension L of the source (in the case of seismic motions of the seafloor) such as that $\lambda \sim L$. The dimensions of tsunami waves are such that they significantly exceed the depth H of the ocean ($\lambda \ggg H \implies H/\lambda \lll 1$), meaning the possibility of modelling tsunami using the long-wave or shallow-water theory. This implies that horizontal velocities and accelerations are far superior to the vertical ones ($U_{xy} \ggg U_z$ and $\dot{U}_{xy} \ggg \dot{U}_z$), fact that consequently simplifies the governing equations of tsunami propagation, giving a tsunami phase velocity $c = \sqrt{gH}$ (considering an incompressible ocean with no elasticity) in the limit $\lambda \ggg H$ of

the more general expression for phase velocity $c = \sqrt{\frac{g \tanh(kH)}{k}}$, where $k = \frac{2\pi}{\lambda}$ is the wave number. This dependence on the seafloor depth and shape makes the wave propagation sensitive to the bathymetry, being affected by ridges, elevations, depressions and the shelf, leading to effects such as scattering, diffraction and changing the amplitude of the tsunami. In case of propagation on open ocean, the amplitude is often limited to $\sim 10^{-1}$ [m] and rarely exceeds 1 meter.

As the propagation of the tsunami waves approaches the coast, the sudden decrease of water depth leads to a decrease of the phase velocity, compressing the wave train, diminishing the wavelength and, because of energy conservation, an increase in the amplitude of the tsunami. Phenomena related with the interaction of the waves with the shallower coastal bathymetry and other coastal geographical features (ridges, canyons, bays), such as diffraction, resonance or reverberation, often result on making the first wave not the strongest. The onshore advance of the waves can reach up to several kilometers and the run-up height[1] is capable of attaining tens of meters. It is important to note that the inland maximum water level of the tsunami can be located at the shoreline or the inundation line or anywhere in between, and can be higher than the run-up height. This massive amount of water entering at high speeds of tens of meters per second, can erode and carry vast amounts of sand, sediments, sea animals and plant life inland, depositing it as it inundates the region affected by the tsunami. These strong flows of water can inundate, destroy and wash out buildings, causing great loss of infrastructure and more importantly, life. As such, it is necessary to deepen the understanding of the characteristics of the seismic sources in a defined rupture area, in order to delve deeper in the generation and propagation processes of tsunamis.

Most devastating tsunami are caused by underwater subduction earthquakes; according to historical data (GHTD, NCEI) more than 70 % of tsunamis worldwide have been caused by strong underwater earthquakes. One of the main hurdles of studying tsunami caused by earthquakes is the difficulty of relating the magnitude of an event with the intensity of a tsunami. For the same moment magnitude, tsunami intensities may vary significantly, varying, for example, the observed

[1] Elevation reached by the tsunami relative to the shoreline.

run-up heights for the same moment magnitude 8.0 Mw by a factor of 64 (Levin & Nosov, 2009). These massive uncertainties in terms of tsunami intensities using the moment magnitude may be caused by the depth of the source, the focal mechanism, the water depth on the source area or the tectonic setting of the source (Gusiakov *et al.,* 2011). Thus, one of the most important factors, and often neglected in contrast with the hydrodynamic part of the phenomenon, is the understanding of the role of the seismic source on the tsunami generation process.

From an hydrodynamic standpoint, the generation and propagation processes are somewhat well understood, even though it is an extremely difficult and challenging problem involving complex mathematical formulations and complicated fluid mechanics and dynamics concepts. There are even existing analytical solutions for obtaining information about wave-fields and velocity potentials for some ideal cases. Similarly, subduction earthquakes themselves are relatively well understood, however our understanding og the direct contribution of a heterogeneous slip distribution and vertical deformation field to tsunami generation process has some gaps, for example, tsunami modelling typically makes the assumption that the sea-surface initial condition matches the seafloor deformation. The role of the seismic cycle and segment rupture lengths in tsunami generation needs to be studied in the light of data from paleotsunami as they may define a path to study the behaviour of a rupture zone from the viewpoint of the characteristics of possible historical and recurrent slip distribution patters. From the point of view of hazard assessment, studying and comprehending paleotsunami also, provides an opportunity to be better prepared for future earthquakes.

In order to attain the goals set up in previous paragraphs, this work proposes the study of past earthquakes and tsunami with a stochastic approach, generating random models and testing their compliance with data compiled from the event of interest, in order to estimate the most probable slip distribution that caused the tsunami.

6.2 Random model generation

6.2.1 Mathematical formulation of slip distribution generation

In order to produce the aforementioned hypothetical rupture scenarios, random slip distributions are generated based on a combination of fault geometry defining parameters. This fault is subdivided into n rectangular subfaults and, following the methodology proposed by LeVeque *et al.,* 2016, a random slip value s_i is assigned to the i-th subfault, defining a $\mathbf{s} \in \mathbb{R}^n$ slip vector. These final slip values have a joint lognormal distribution given by the exponentiation of a Karhunen–Loève (K-L) sum (LeVeque *et al.,* 2016, Mai & Beroza, 2002), representing a linear combination of eigenvectors and eigenvalues of a covariance matrix $\hat{C} \in \mathbb{R}^{n \times n}$, which is a function of the distribution of the subfaults that define the fault geometry.

Said distribution of parameters for the generation of stochastic earthquake scenarios is achieved with a LT structure (see figure 1.2 for reference example), most commonly used in PTHA (Fukutani *et al.,* 2014, Goda *et al.,* 2016, Fukutani *et al.,* 2018), but in this study it is used as a tool for generating random scenarios in the form of slip distributions. Once a distribution is obtained, the desired number of subfaults and their locations[2] are defined, and finally with this information as input, the correlation matrix $\hat{C}$ is computed as proposed by LeVeque *et al.* (2016),

$$C_{ij} = \exp\left(-\left(\frac{d_{strike}(i,j)}{r_{strike}}\right) - \left(\frac{d_{dip}(i,j)}{r_{dip}}\right)\right) \tag{6.1}$$

where d_{strike} and d_{dip} are the distances between the centers of the subfaults i and j along-strike and along-dip respectively, meanwhile r_{strike} and r_dip are the correlations lengths in said directions. As proposed in LeVeque *et al.* (2016), the correlation lengths are defined as 40 % of the along-strike and along-dip dimensions. Distances between subfaults are defined differently for the along-strike and along-dip cases, firstly $d_{strike}(i,j) = \sqrt{d_{ij}^2 - d_{dip}(i,j)^2}$, where d_{ij} is the great

[2]It is possible to define the location of each subfault with their corners or their centre point, in this work the latter is used for its definition.

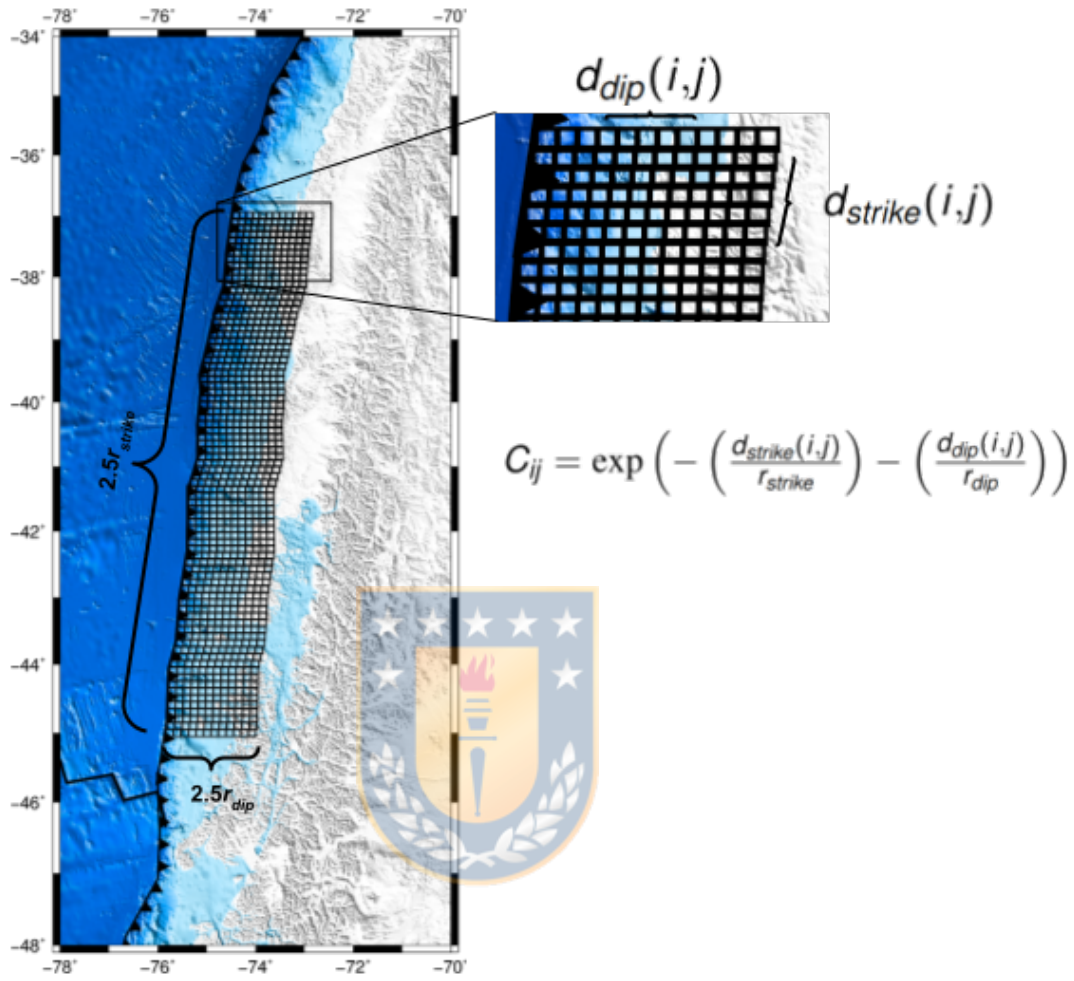


Figure 6.1: Geometric meaning of the terms for computing C.

circle path distance between the i-th and j-th subfault, and d_{dip} is defined using the difference in depth of the subfaults d_{depth} and the dip angle δ as $d_{dip}(i,j) = d_{depth}/\sin(\delta)$. Subsequently, the eigenvalues and eigenvectors of C_{ij} can be computed, and thus an initial Gaussian slip distribution vector $\vec{S}^g$ can be obtained with

$$\vec{S}^g = \vec{\mu}^g + \sum_{k=1}^{m} \sqrt{z_k \lambda_k} \vec{v}_k \tag{6.2}$$

where $\vec{\mu}^g$ is a vector with the Gaussian mean defined as $\mu_i^g = \log(\bar{\mu}\tau_i) - \frac{1}{2}\log\left(\alpha^2 + 1\right)$, τ is a taper function of the depth of each subfault defined by $\tau(d) = 1 - \exp\left(-20\left|d - d_{max}\right|/d_{max}\right)$, d_{max} is

the maximum depth of the fault, $\alpha = 0.5$ is a scalar and z is a set of random independent number with a distribution $z \sim \mathcal{N}(0,1)$. Figure 6.1 shows a schematic example of the computation of C.The number m of eigenvalues and eigenvectors used in the sum determine the complexity, or the degree of heterogeneity, in the spatial distribution of the slip, controlling how different from a homogeneous distribution the field is created. Heterogeneous slip distributions are preferable compared to homogeneous ones, because the latter tend to underestimate tsunami intensities for the same moment magnitude (Carvajal *et al.,* 2017), as well as not give valuable insight on the location of the maximum slip values. Then, in order to obtain the desired final lognormal distribution, each component of $\vec{S}^g$ is exponentiated, as follows,

$$\vec{S} = \exp\left(\vec{\mu}^g + \sum_{k=1}^{m} \sqrt{z_k \lambda_k}\vec{v}_k\right) = \exp\left(\vec{\mu}^g\right) \exp\left(\sum_{k=1}^{m} \sqrt{z_k \lambda_k}\vec{v}_k\right)$$

Lay *et al.* (2012) notes that close to the trench big subduction earthquakes tend not to have comparatively important slip values, so a Hanning window shaped taper is applied to the near-trench end of the distribution (west end in the case of Chilean earthquakes). This may not necessarily be true for some earthquakes, such as 2011 Tohoku-Oki, so thi staper is optional in the generation code. As this affects the magnitude of the event, the field is then scaled by a multiplicative factor so that the sum of the contribution of each subfault gives the desired moment magnitude Mo_{LT}, given by the LT parameter combination. Thus, if the magnitude given by the LT is Mw_{LT}, and the magnitude of the slip vector is Mw_s, the final slip distribution S_f is obtained by multiplying the slip vector by the ratio of the seismic moments of the LT magnitude and the slip vector magnitude, so as if the new moment of each subfault is computed and the contribution of each of them is summed, the desired moment magnitude is obtained. This is computed with,

$$\vec{S}_f = \left(\frac{Mo_{LT}}{Mo_s}\right)\vec{S} \tag{6.3}$$

Where Mo_{LT} is the seismic moment related to the magnitude given by the LT and Mo_s is the moment of the previews slip distribution. With this, the generation of the slip distribution for one of the combinations of the LT is finished and as a result 7 matrices are created, containing slip, latitude, longitude, depth, dip, strike and rake for each subfault.

The vertical deformation that the earthquake causes can be computed with the Okada (1985) solution. For this purpose, Clawpack (a Python package) is used as proposed by LeVeque *et al.,* (2016). This package contains functions to compute vertical deformation, manipulate slip distributions and to plot maps, among others. The final output is a file containing the vertical deformation field, ready for the next restrictions with deformation data and/or for tsunami modelling.

6.2.2 Definition of LT branches values

As it was mentioned in previous paragraphs, $\vec{S}$ is computed with a linear combination of eigenvalues and eigenvectors of a correlation matrix $\hat{C}$, function of the geometry of the array of subfaults. In turn, this geometry is completely controlled by the branches of the LT, so the correct definition of the range of values for every branch is a crucial first step in the estimation process.

6.3 Tsunami modelling

Tsunami modelling in this study is performed using the software COMCOTv1.7 (**Co**rnell **M**ulti-grid **Co**upled **T**sunami model). This software uses explicit staggered leap-frog finite difference schemes to solve Shallow Water Equations. As it name implies, a nested grid system, dynamically coupled up to 12 levels (or up to 1 parent grid and 11 individual grids, or a combination of both cases), each nested level with a lower resolution in an integer ratio with its previous level, can be implemented in the model to allow for high resolution modelling in areas of interests (figure 6.2 shows an example).

The model accepts as inputs different types of initial conditions (e.g.: Landslides, fault models,

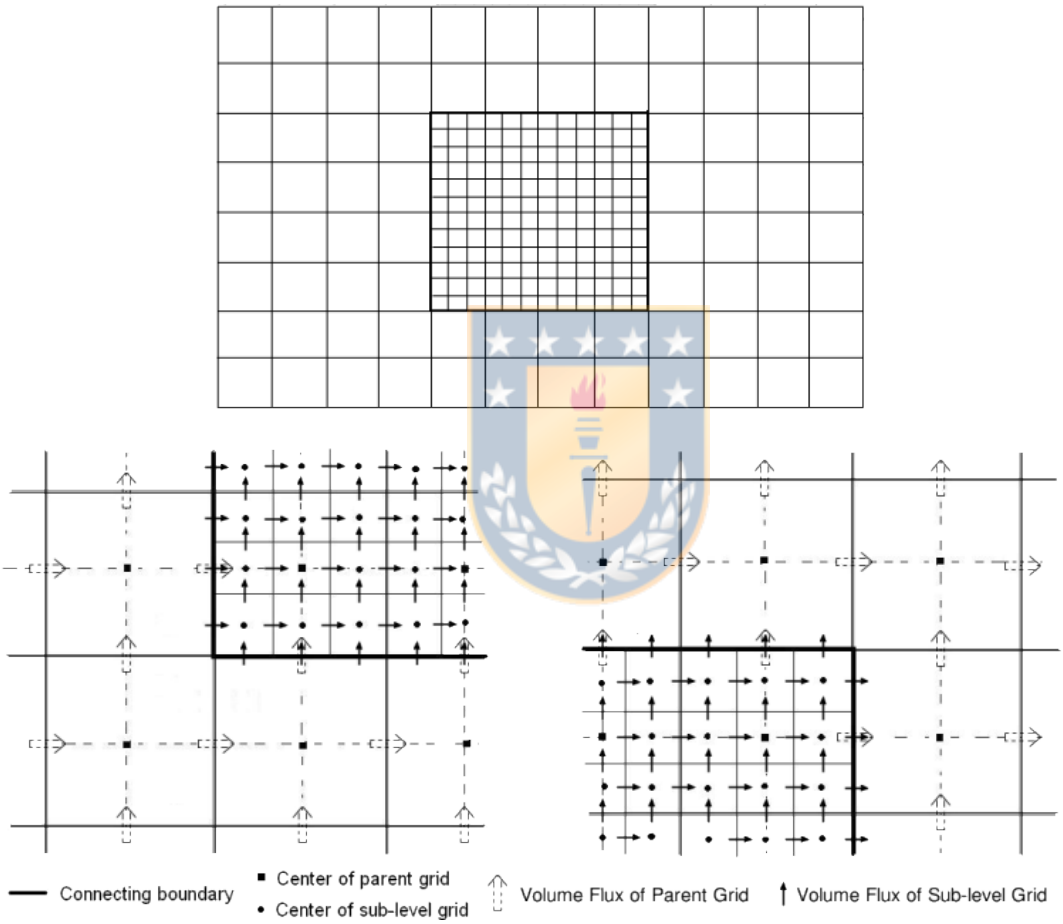


Figure 6.2: Detailed schematic of nested grids in COMCOTv1.7. Upper panel shows a parent grid and a nested grid with a resolution ratio of 3 between them. Lower left and right panel show a detailed schematic of the behaviour of volume fluxes for parent and nested grids (sub-level grids).

solitary waves, seafloor deformation). When seafloor deformation models are available they are used as initial conditions; in this case, water surface displacement is considered the same as the deformation of the seafloor.

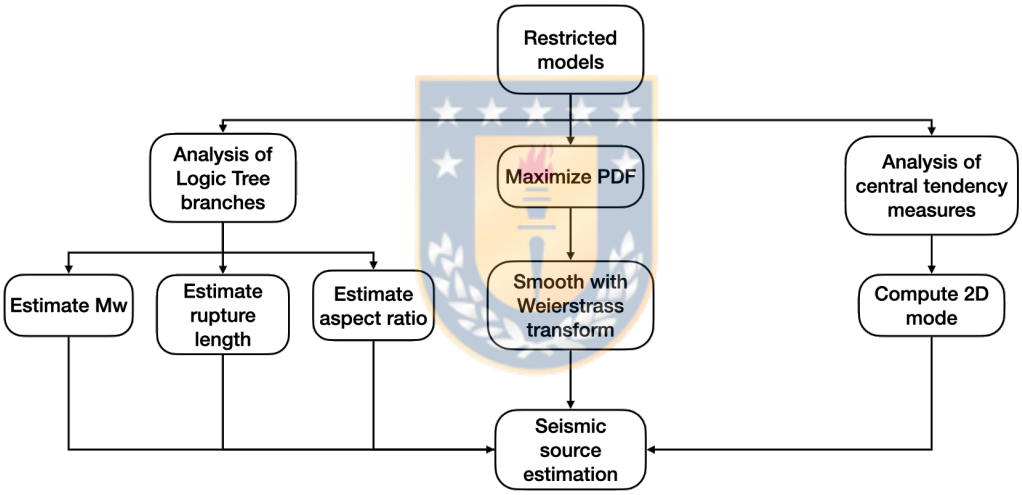
As it is common for tsunami from subduction earthquakes to inundate vast areas of land, the distribution of data available tends to be greatly scattered along the shore. This in tandem with the model's limit of 12 grids, creates a necessity for dividing the area of interest for tsunami modelling in different sections. The division of the entire modelling area into more manageable sub-areas is a crucial part of the estimation process, as the algorithm is highly sensitive to the order of tsunami restrictions. A more detailed discussion can be found in previous chapters. In addition of solving the problem with the limited number of grids, this allows the user to model tsunamis in smaller areas, with smaller grids, lowering computational costs. Also, restrictions with tsunami data can be performed in a staggered fashion, starting with the most important sub-area in terms of data availability (a definition of "most important" can be found in the next section), one can model tsunami using all the deformation models that cleared the first round of restrictions (deformation data restriction), then filter the results with the data corresponding to said geographical sub-area, and continue modelling in the following sections with fewer deformation fields, further cutting computational costs.

6.4 Statistical analysis

The characterization of the seismic source for past earthquakes from which only deformation, tsunami inundation data and historical accounts are available carries the heavy burden of great uncertainties (see chapter 4 for a thorough discussion on the topic). For this, the estimation of the most probable source in terms of its data available, shown in figure 6.3 with a simplified flowchart schematic of the estimation process. The estimation of the slip distribution using the maximisation of the PDF is far more computationally demanding and slower compared with computing a 2D

mode. For exploratory analysis of the estimation in intermediate steps of restriction, the 2D mode is encouraged versus the PDF method, as with more than a 100 models the computation of the latter method can take up to a day (all computations where done using a Intel i7-8750H Six-Core processor with 16 Gb of RAM)

6.4.1 Analysis process summary

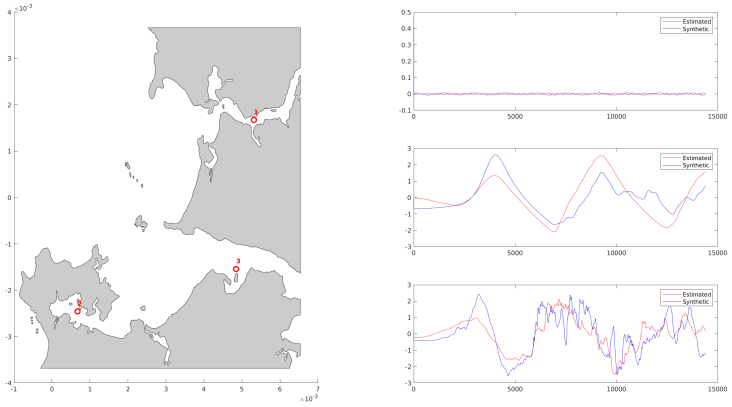


Anal

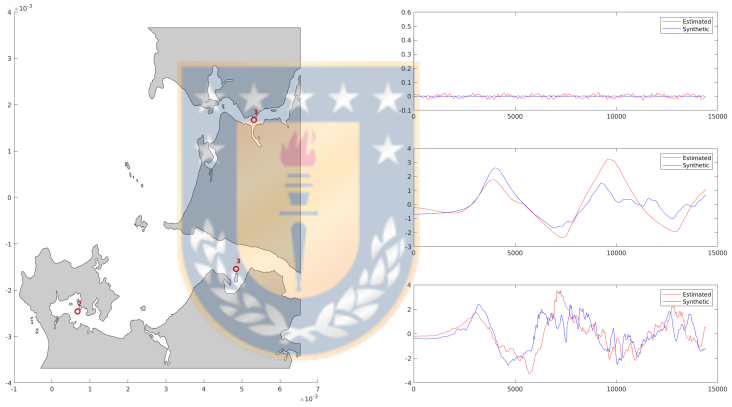
Figure 6.3: Summarized flowchart of the statistical analysis process.

6.4.2 Comparison between mode and PDF maximization

With two possible methods of estimation, it is important to know their performance. Both methods excel in adjusting to sea surface height data. This was tested in the synthetic models. Maximizing the subfault-wise PDF takes up to 100 times longer time to compute in comparison with computing the 2-D mode of the models, however, this methods presents a slight better performance, presenting higher correlations and not under estimating sea surface heights.



(a) Mode



(b) PDF

Figure 6.4: Comparison example between the sea surface height from the synthetic model and from the estimated model, both for the one estimated with the 2-D mode and the one maximizing the PDF.

6.5 Synthetic tests

Two synthetic slip models are created in order to assess the capabilities of the algorithm for resolving features of slip distributions, namely slip patches or slip maxima locations, given certain characteristics of data distribution, size of the ruptured area, depth of the slip distribution among others. Deformation and tsunami data are sampled from them and then are submitted to the algorithm so as to estimate the initial synthetic model to perform a resolution test.

This section shows a summary of the parameters used in the estimation process of the synthetic

models.

6.5.1 Summary

Branch Parameters

Table 6.1: Value ranges of the Logic Tree for the generation of random slip distributions for estimating the seismic source of the first synthetic model. Moment magnitudes range from 8.6 to 9.0 in 0.1 increments, complexity values (N) range from 14 to 20 in two units increments, aspect ratio values (AR) range from 2 to 4 also in unit increments. In the case of southern and northern limits variations, both ranges have 5 values

Branch	Mw	N	AR [L/W]	D. [km]	N. l. [°]	S. l. [°]
Range	[8.6,9.0]	[14,20]	[2,4]	[0,60]	[-39.0,-38.0]	[-44.5,-43.0]

Table 6.2: Value ranges of the Logic Tree for the generation of random slip distributions for estimating the seismic source of the second synthetic model. Moment magnitudes range from 8.6 to 9.0 in 0.1 increments, complexity values (N) range from 20 to 24 in unit increments, aspect ratio values (AR) range from 2 to 4 also in unit increments. In the case of southern and northern limits variations, both ranges have 5 values

Branch	Mw	N	AR [L/W]	D. [km]	N. l. [°]	S. l. [°]
Range	[8.6,9.0]	[20,24]	[2,4]	[0,60]	[-39.0,-38.0]	[-44.5,-43.0]

Restriction parameters

Table 6.3: Restriction tolerances used in the synthetics test estimation process. The first restriction penalty function and the number of bands were set as defined in earlier sections.

	Deformation Tol.	Match Tol.	Tsunami Tol.	Match Tol.
Synthetic Model 1	0.3 [m]	10	0.5 [m]	0
Synthetic Model 2	0.3 [m]	10	0.8 [m]	0

Evolution of the number of models throughout the restrictions

The number of models steeply decline in the first two deformation restrictions, then decrease at a much lower pace. However, both cases show a similar behaviour in the overall decrease in number up to the final amount of the analysis set. With the cardinality of both sets being 0.015 % and 0.055 % of the initial amount.

Table 6.4: Evolution of quantity of random models after every step of the restriction process up to the final set for analysis and estimation.

Case	Initial	Def. 1	Def. 2	Pre Tsunami	Tsunami	Analysis set
Synthetic I	200000	10114	248	221	30	30
Synthetic I	200000	7909	108	102	11	11

6.6 Intermediate results

6.6.1 Valdivia 1960

For the main estimation of the Valdivia 1960 slip distribution shown in earlier chapters, a set of 150 deformation (Plafker & Savage, 1975) and 103 tsunami (Winckler, original data) data points where used. This quantitative data allows for the point-to-point comparison of the amount of deformation of each random model against the measurements, giving the possibility of defining tighter tolerances for the restriction process. However, for earlier earthquakes and tsunamis, the possibility of being able to measure the amount of deformation and/or tsunami height or run up is slim, or if possible, the amount of uncertainty in the measurements may be too high. Alternatively, the methodology allows for the input of categorical data, both for deformation and tsunami to check for compliance in the restriction process, the categories are shown in table 6.5. This way the methodology doesn't check an amount of deformation, but the "direction" of deformation, and for the case of tsunami, it checks if a certain site in a domain was or not inundated. A combination of quantitative and qualitative data is also implemented in the methodology.

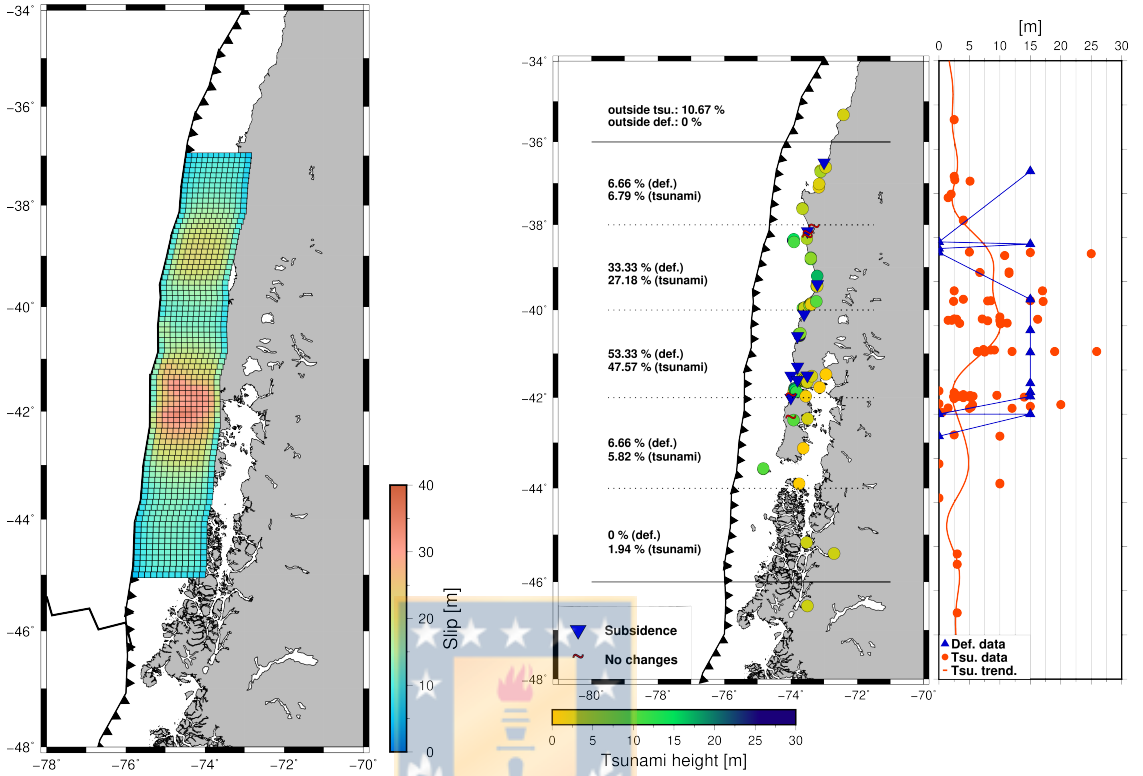
Table 6.5: Categories for the input of categorical data to the restriction process. Deformation allows for three types of data, checking for subsidence, uplift or no change. On the hand, Tsunami restrictions can be made by checking if a site was or not inundated.

Case	Category I	Category II	Category III (Deformation only)
Deformation	Subsidence	Uplift	No change
Tsunami	Does inundate site	Doesn't inundate site	

This case scenario, in terms of quality of the estimations, performs not as well as the one using quantitative data. It is capable of resolving the location of the main slip features with some leeway, however it tends to underestimate or overestimate slip values, depending of the prevalent category of data. The results shown in figure 6.5 (a) were obtained with 14 categorical deformation data points (Cisternas *et al.,* 2000, Garret *et al.,* 2015, Hocking *et al.,* 2017, Hong *et al.,* 2017, Kempf *et al.,* 2017) and the same 103 quantitative tsunami points. This slip distribution show two slip patches, a main one between $\sim$ 41° S and $\sim$ 43° S, with maximum slip values of up to $\sim$40 [m], and a secondary one up north between $\sim$ 38° S and $\sim$ 39.5° S with lower slip maxima, reaching $\sim$ 20 [m]. This secondary slip feature location coincides with features of the slip models from Barrientos *et al.,* 1990, Moreno *et al.,* 2009 and Fujii & Satake, 2013. However, the main slip patch in the southern portion of the rupture presents slip values much larger than the ones in the forementioned models. This difference can be explained by the excess of deformation data (see figure 6.5 (b)) indicating subsidence in this area, in contrast with other latitudes, and the fact that these deformation restrictions where performed with qualitative data that only had information of subsidence and no change in the terrain, lacking uplift data (as no uplift was recorded in this area).

6.6.2 Importance of data coverage

To test the importance of data coverage in the estimation process, several tests were made in order to assess the importance of gaps in data coverage. As stated previously, gaps in coverage lead to areas where there is no possibility to check if a given random model complies or not with data, which in turn leads to unusual areas of large slip. As long as the compliance conditions given by



(a) Slip estimation (b) Data

Figure 6.5: Left. Estimation of the slip distribution for the Valdivia 1960 earthquake using categorical deformation data accounting for a 10 % of the total used in the main estimation (blue downward triangles show subsidence, red tilde sign show no change). Right. Categorical deformation and tsunami data used in the estimation. Latitudinal profiles to the right of the figure show the changes of these values along strike. Categorical deformation is shown as blue triangles with zero [m] amplitude for no changes and 15 [m] for subsidence.

data in areas where there is coverage are met, the restriction process is blind to such high slip features, not being able to restrict them.

As an extreme case, the second synthetic slip distribution is estimated using only the northernmost section of the data, discarding every point south of 41° S. Results can be seen in figure 6.6. High slip features appear south of 41° S, not corresponding to any data point, as well as low slip artefacts in the down dip rupture limit that come from the lack of restriction capabilities. As well as these artefacts the estimated slip distribution presents a much larger aspect ratio, with its down dip limit much deeper than the one in the estimation with all data points.

The lack of data points, be it gaps or low density in the geographical distribution of such points,

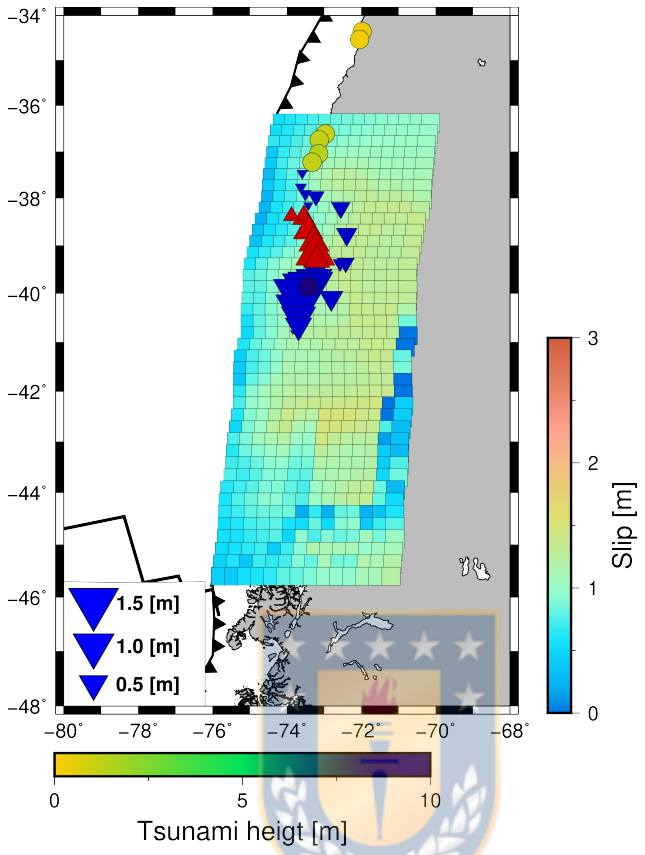


Figure 6.6: Estimated slip distribution for the second synthetic model case. This estimation was made using only data points north of 41°. Along with slip values, deformation (blue triangles indicate subsidence, red indicate uplift) and tsunami (coloured circles) data points are shown in the figure. High slip features appear south of the data limit, as well as low slip artefacts in the down slip limit.

primarily leads to the emergence of the aforementioned high slip patches in poorly covered areas, to more homogeneous slip distribution, and to an overestimation of the along dip dimension of the rupture, to compensate for such lack of data.

In the case where there are large gaps that span up to one of the along strike rupture limits (such as this example), this methodology will most surely estimate flawed slip distributions. In contrast, if gaps in data coverage are constrained within areas with data, although artefacts may arise due to the gaps, the estimated slip structures will present correct slip features, albeit with underestimated slip values due to the amount of seismic moment that is located in these high slip artefacts.

6.7 Valdivia 1960 data

The main deformation dataset used in the estimation of the Valdivia 1960 Mw 9.5 earthquake is taken from Plafker & Savage, 1970. Data points can be seen in tables 6.6, 6.7, 6.8 and 6.9.

- Categorical Deformation data available for the Valdivia 1960 earthquake

Table 6.6: Vertical deformation measured extracted from Plafker & Savage, 1975. Regardless of uncertainties in the measurements, all data points were taken into account using the tolerance shown in table 3.2. Part 1

Latitude °	**Longitude °**	**Vertical deformation [m]**
-36.766700	-73.150000	0.3
-37.067200	-73.151700	0.0
-37.152500	-73.185300	0.0
-37.233600	-73.417800	0.0
-37.185800	-73.551400	0.0
-37.035000	-73.517800	0.0
-37.585800	-73.651400	1.3
-38.235000	-73.235300	0.9
-38.333900	-73.501100	-0.2
-38.384200	-73.884400	1.0
-38.401100	-73.950800	1.8
-38.718300	-73.418600	-1.2
-38.783300	-73.400000	-1.4
-39.216900	-73.202200	-2.0
-39.433300	-73.201100	-1.6
-39.851900	-73.417800	-2.1
-39.933300	-73.583300	-0.7
-40.133600	-73.651900	-0.5
-40.218300	-73.716900	-0.7
-40.501100	-73.817200	-1.3
-40.551100	-73.751700	-1.6
-40.683900	-73.817200	-1.2
-41.285000	-73.868100	-1.1
-41.616700	-73.616700	-1.6
-41.633330	-73.600000	-1.7
-41.750600	-73.716700	-2.4
-41.768600	-73.417200	-1.3
-41.800600	-73.367200	-1.0
-41.768600	-73.234700	-0.7
-41.800000	-73.118600	-0.4
-41.768600	-73.133300	-1.0
-41.750560	-73.118600	-0.6
-41.717200	-73.067200	-0.4
-41.517200	-73.035300	0.0
-41.500000	-73.816700	0.5

- Categorical Deformation data available for the Valdivia 1960 earthquake

Table 6.7: Vertical deformation measured extracted from Plafker & Savage, 1975. Regardless of uncertainties in the measurements, all data points were taken into account using the tolerance shown in table 3.2. Part 2

Latitude °	**Longitude °**	**Vertical deformation [m]**
-41.851400	-73.951900	-1.8
-41.867200	-73.918100	-1.5
-41.868600	-73.883900	-1.5
-41.885000	-73.850000	-1.3
-41.818600	-73.633300	-1.9
-41.985300	-73.518600	-1.5
-41.966700	-73.533900	-1.6
-42.134700	-73.467200	-1.5
-42.167200	-73.401400	-1.4
-42.268100	-73.350000	-1.3
-42.301900	-73.266700	-0.9
-42.334700	-73.133300	-0.2
-42.318100	-72.800600	0.8
-42.401400	-72.718100	0.9
-42.485600	-72.650000	0.5
-42.517200	-72.816700	0.9
-42.651400	-72.833900	1.0
-42.633300	-73.068100	0.6
-42.616700	-73.218600	0.0
-42.601400	-73.283900	-0.8
-42.600600	-73.333300	-0.8
-42.550000	-73.433900	-0.9
-42.616700	-73.500600	-1.1
-42.583300	-73.618100	-1.4
-42.600000	-73.683900	-1.6
-42.616700	-73.733300	-1.5
-42.616700	-73.818600	-1.6
-42.517200	-73.800000	-1.6
-42.618600	-74.116700	-1.0
-42.883600	-73.484200	-0.9
-41.501400	-72.800000	0.7
-41.533900	-72.751100	0.8
-41.617500	-72.667500	0.9
-41.685300	-72.635000	0.7
-41.735300	-72.567200	0.9
-41.718100	-72.467200	0.8
-41.650000	-72.384700	0.0
-41.650600	-72.300000	0.3
-41.850600	-73.985300	-1.0

- Categorical Deformation data available for the Valdivia 1960 earthquake

Table 6.8: Vertical deformation measured extracted from Plafker & Savage, 1975. Regardless of uncertainties in the measurements, all data points were taken into account using the tolerance shown in table 3.2. Part 3

Latitude °	**Longitude °**	**Vertical deformation [m]**
-42.934200	-73.618900	-1.4
-42.916700	-72.718100	1.0
-43.133300	-73.518100	-0.9
-43.116700	-73.568100	-1.2
-43.118100	-73.618600	-1.2
-43.133300	-73.633300	-1.2
-43.150000	-73.685300	-1.4
-43.135300	-73.766700	-1.5
-43.250000	-73.701900	-1.7
-43.333900	-73.700000	-1.5
-43.350000	-73.783900	-1.5
-43.350000	-74.050000	-2.1
-43.566700	-74.833300	3.6
-43.816700	-74.018100	-1.7
-43.884700	-74.001400	-1.4
-43.883300	-73.868100	-1.0
-43.934700	-73.801400	-0.9
-43.967200	-73.801400	-1.1
-43.900000	-73.751900	-1.3
-43.900000	-73.750000	-1.1
-43.983900	-73.735300	-1.0
-44.050600	-73.716700	-0.9
-44.084700	-73.650000	-0.7
-44.133300	-73.651400	-0.7
-44.118100	-73.534700	-0.4
-44.133300	-73.483300	0.0
-44.150600	-73.818100	-1.0
-44.050000	-73.133900	-1.7
-44.634700	-74.733300	0.0
-44.651900	-74.618100	0.0
-44.601900	-73.600600	0.0
-44.783900	-73.601400	0.0
-44.734700	-73.818100	-1.0
-44.701400	-74.251400	-0.9
-44.716700	-74.333300	-1.5

- Categorical Deformation data available for the Valdivia 1960 earthquake

Table 6.9: Vertical deformation measured extracted from Plafker & Savage, 1975. Regardless of uncertainties in the measurements, all data points were taken into account using the tolerance shown in table 3.2. Part 4

Latitude °	**Longitude °**	**Vertical deformation [m]**
-44.634700	-74.467200	-0.8
-44.634700	-74.450600	-1.2
-44.718100	-74.418600	-1.0
-44.783300	-74.384700	-1.3
-44.818600	-74.316700	-1.0
-44.867200	-74.151900	-0.9
-44.884700	-74.066700	-0.8
-44.933300	-73.918600	0.0
-45.000600	-73.801400	0.7
-45.750300	-73.718600	-0.6
-45.016700	-72.500000	0.0
-44.916700	-75.033900	5.7
-45.201900	-74.501900	-0.5
-45.317200	-74.317200	-0.6
-45.334700	-74.316700	-0.7
-45.383300	-74.266700	-0.7
-45.401400	-74.035300	-0.4
-45.466700	-73.968100	-0.2
-45.468100	-73.868100	0.0
-45.466700	-73.816700	-0.5
-45.350000	-73.817200	0.0
-45.283300	-73.001400	0.3
-45.283900	-73.601400	0.0
-45.283900	-73.568600	0.3
-45.234700	-73.551900	0.0
-45.200000	-73.533300	0.0
-45.151900	-73.518100	1.0
-45.300000	-73.367200	-0.5
-45.284700	-73.200600	0.0
-45.351900	-72.084700	0.0
-44.116700	-74.166700	-1.6
-44.184700	-74.318600	-1.2
-44.216700	-74.251900	-1.6
-44.235300	-74.168100	-1.7
-44.250600	-74.118100	-1.5
-44.283900	-74.101900	-1.5
-44.283900	-74.083300	-1.4
-44.350600	-74.017200	-1.4
-44.318100	-74.000600	-1.3
-44.335300	-73.934700	-1.2
-44.351400	-73.867200	-1.1
-44.368100	-73.850000	-1.1
-44.400000	-73.400000	-0.7
-44.418100	-73.718600	-0.4
-44.483300	-73.617200	0.0

In addition with deformation data from Plafker & Savage, 1970, this study has categorical deformation data (see table 6.5) used in subsection 5.6.1.

- Categorical Deformation data available for the Valdivia 1960 earthquake

Table 6.10: Categorical vertical deformation data available for the Valdivia 1960, 9.5 Mw earthquake. Zero stands for no deformation and -1 for subsidence measured by the study.

Location	**Latitude °**	**Longitude °**	**Vertical deformation [m]**
Maullín River (Cisternas *et al.,* 2000)	-41.5	-73.5	-1
Cocotué (Cisternas *et al.,* 2018)	-42.0	-74.0	0
Quidico (Dura *et al.,* 2017)	-38.25	-73.49	0
Quidico (Hocking *et al.,* 2017)	-38.1	-73.3	0
Tirúa (Dura *et al.,* 2017)	-38.34	-73.5	0
Huelde Lake (Kempf *et al.,* 2017)	-42.5	-74.0	0
Chucalén (Garret *et al.,* 2015)	-41.5	-74.0	-1
Andalién River (Hocking *et al.,* 2017)	-36.5	-73.0	-1
Tirúa River (Hocking *et al.,* 2017)	-38.15	-73.5	-1
Lingue River (Hocking *et al.,* 2017)	-39.4	-73.2	-1
Chaihuén River (Hocking *et al.,* 2017)	-40.1	-73.6	-1
Pucatrihue (Hocking *et al.,* 2017)	-40.6	-73.8	-1
Llico (Hocking *et al.,* 2017)	-41.3	-73.8	-1
Maullín River (Hocking *et al.,* 2017)	-41.6	-73.8	-1

Tsunami data, gently provided by Dr. Patricio Winckler, (tables 6.11 and 6.12) was used in the restriction process.

- Tsunami wave height data available for the Valdivia 1960 earthquake

Table 6.11: Tsunami wave height data of the Valdivia 1960 earthquake (part 1).

Location	Latitude °S	Longitude °W	Wave height [m]
Unnamed location	-39.875373	-73.389478	2.1
Niebla	-39.875373	-73.389478	1.5
Unnamed location	-39.852284	-73.391825	16.2
Niebla	-39.850571	-73.391796	2.9
Unnamed location	-39.846332	-73.393392	7.1
Caleta los Molinos	-39.844983	-73.396299	2.5
Caihuín Corral	-39.938885	-73.589891	11.2
Caleta Huito	-39.957281	-73.642152	10.1
Unnamed location	-39.944342	-73.584756	3.4
Mehuín	0	-73.216304	6.9
Mehuín (fishermen's cove)	-39.445972	-73.209361	2.4
Mehuín Mississippi	-39.448626	-73.225873	17.1
Bahía Mansa	-40.582421	-73.734575	6.3
Bahía Mansa (beach)	-40.581348	-73.731758	25.9
Pucatrihue	-40.537638	-73.714546	7.3
Pucatrihue (beach)	-40.547968	-73.718563	9.1
Punta Corona lighthouse	-41.787152	-73.893898	5.5
Cancha Chalcura (Chiloé)	-41.841666	-73.869997	5.3
Bahía Mansa	-40.580888	-73.738015	19
Bahía Mansa	-40.582563	-73.734782	19
Toltén viejo	-39.208302	-73.208308	2.5
Caleta Estaquilla	-39.208302	-73.208308	17
Quenuir	-41.573412	-73.678171	9.5
West of Maullín (mouth of San Pedro Nolasco River)	-41.62001	-73.659215	4
West of Maullín (lower reach of del Rey River)	-41.626199	-73.649483	4
West of Maullín (Los Caules)	-41.616198	-73.64129	5.1
Chanhué (between Quenuir and La Pasada)	-41.558534	-73.637702	4
West of Maullín Los Caules	-41.605255	-73.62952	5
West of Maullín (Huichamilla)	-41.611383	-73.614988	2.6
West of Maullín (Huichamilla)	-41.609568	-73.613827	2.7
West of Maullín (northeast of Huichamilla)	-41.604153	-73.612743	4
West of Maullín (near Los Pinos)	-41.596894	-73.608099	3
La Pasada	-41.58242	-73.602411	5
Maullín (western outskirts)	-41.595021	-73.60214	5.6
Maullín (western part of town)	-41.594977	-73.598541	2.8
Maullín (western part of town)	-41.595413	-73.597332	3.8
La Pasada	-41.581882	-73.595225	5.2
La Pasada	-41.583225	-73.594596	3.1
Maullín (Plaza de Armas)	-41.595804	-73.592524	4.9
Maullín (along Cariquilda River)	-41.603905	-73.592348	2.4
Maullín (south end of Tentén hill)	-41.609306	-73.59223	2.9
La Pasada	-41.582288	-73.591618	3
Chuyaquén	-41.603564	-73.564755	3.9
Chuyaquén	-41.597218	-73.561296	3.5
Chuyaquén	-41.602574	-73.557577	2.5
Chuyaquén	-41.588799	-73.536288	3.5
Cariquilda (upper tidal reaches of Quiriquilda River)	-41.642526	-73.51346	2.5
Cariquilda (upper tidal reaches of Quiriquilda River)	-41.646933	-73.506154	2.5
Misquihué	-41.512721	-73.369041	2.5
Puerto Montt	-41.474841	-72.942598	0
Calbuco	-41.765697	-73.131663	0
Ancud	-41.868311	-73.832198	5

Table 6.12: Tsunami wave height data of the Valdivia 1960 earthquake (part 2).

Location	Latitude °S	Longitude °W	Wave height [m]
Quellón	-43.124172	-73.623283	0
Isla Guafo	-43.566283	-74.827354	10
Lebu	-37.60564	-73.653048	4
Puerto Saavedra	-38.79247	-73.39726	6.7
Lota	-37.097808	-73.159962	1.5
Queule	-39.398523	-73.213686	4
Mehuin	-39.432953	-73.217398	8.5
Maullín	-41.61502	-73.598995	14
Caleta Mansa	-40.580895	-73.736373	12
Linao	-41.965351	-73.547685	0.5
Achao	-42.469858	-73.492219	2.5
Gulf of Quetalmahue	-41.860033	-73.916794	1.5
Punta Ahuí	-41.825388	-73.864778	15
Corral	-39.88758	-73.430888	10
Isla Mocha	-38.340331	-73.912968	15
Robinson Crusoe Juan Fernandez	-33.635865	-78.83055	1.5
Melinka	-43.898948	-73.747535	0
Puerto Saavedra	-38.812538	-73.401186	11.5
Mehuín	-39.436379	-73.216074	8
Corral	-39.887009	-73.430034	8
Maicolpué	-40.594828	-73.734798	7.5
Ancud	-41.868311	-73.832198	5
Coronel	-37.02	-73.15	2
Guafo	-43.58	-74.83	10
Lebu	-37.617	-73.65	4
Mansa River	-40.55	-73.76	8.5
Mehuín	-39.43	-73.22	15
Isla Mocha	-38.367	-73.933	25
Isla Mocha	-38.41	-73.905	10.8
Puerto Aysén	-45.4	-72.7	3
Punta Arenas	-53.17	-70.93	0.2
Punta Saavedra	-38.78	-73.4	11.5
Punta Tirna (Tirera)	-38.33	-73.52	5
Antofagasta	-23.65	-70.4	0.7
Arica	-18.467	-70.333	1.1
Aysén	-46.5	-73.5	3
Puerto Aguirre	-45.165	-73.524	3
Ancud	-41.867	-73.828	12
Chiloé (undisclosed location)	-42.5	-73.917	10
Talcahuano	-36.695	-73.106	2.6
Talcahuano	-36.725	-73.105	5.1
Tomé	-36.617	-72.95	2.5
Caldera	-27.064	-70.825	1.5
Coquimbo	-29.95	-71.335	1.1
Eastern Island	-27.15	-109.33	6
Constitución	-35.333	-72.417	2.5
Punta Constitución	-35.33	-72.42	2.5
Corral	-39.867	-73.431	10
Valdivia	-39.8	-73.233	10
Valparaíso	-33.033	-71.633	0.9
Punta Corona	-41.784162	-73.879817	20

6.8 Acknowledgements

The present study is funded by FONDECYT project 11180854 “Source characterization for historical tsunamis of Central-Southern Chile”, directed by Ignacia Calisto, PhD. Additional support and funding by Iniciativa Científica Milenio (ICM) through Grant Number NC160025 “Millenium Nucleus CYCLO: The seismic Cycle Along Subduction Zones” was provided. We wanted to extend our gratitude to Patricio Winckler for providing insightful comments and tsunami data. All figures were produced using Generic Mapping Tools, GMT (Wessel & Smith, 1998).

Chapter 7

References

Bibliography

[1] Abe, K. (1975). Reliable estimation of the seismic moment of large earthquakes. *Journal of Physics of the Earth*, 23(4):381–390.

[2] Annaka, T., Satake, K., Sakakiyama, T., Yanagisawa, K., Shuto, N. (2007). Logic-tree Approach for Probabilistic Tsunami Hazard Analysis and its Applications to the Japanese Coasts. *Pure appl. geophys.* 164, 577–592. https://doi.org/10.1007/s00024-006-0174-3

[3] Becerra, I., Aránguiz, R., González, J., & Benavente, R. (2020). An improvement of tsunami hazard analysis in Central Chile based on stochastic rupture scenarios, *Coastal Engineering Journal*, 62:4, 473-488, DOI:10.1080/21664250.2020.1812943

[4] Carvajal, M., & A. Gubler. 2017. "The effects on tsunami hazard assessment in Chile of assuming earthquake scenarios with spatially uniform slip". In Geist, E. L., Fritz, H. M., Rabinovich, A. B., Tanioka, Y. (eds), Global Tsunami Science: Past and Future, Volume I, 3693-3702. Birkhäuser, Cham: Springer International Publishing. doi: 10,1007/978-3-319-55480-8_3.

[5] Carvajal, M., Cisternas, M., & Catalán, P. A. (2017). Source of the 1730 Chilean earthquake from historical records: Implications for the future tsunami hazard on the coast of Metropolitan Chile. Journal of Geophysical Research: Solid Earth, 122(5), 3648-3660.

[6] Cifuentes, I. L. (1989), The 1960 Chilean earthquakes, J. Geophys. Res., 94(B1), 665– 680, doi:10.1029/JB094iB01p00665.

[7] Cifuentes, I. L., & Silver, P. G. (1989). Low-frequency source characteristics of the great 1960 Chilean earthquake. Journal of Geophysical Research, 94(B1), 643. doi:10.1029/jb094ib01p00643

[8] Cisternas, M., Torrejón, F., Gorigoitía, N., (2012). Amending and complicating Chile's seismic catalog with the Santiago earthquake of 7 August 1580. J. S. Am. Earth Sci.33 (1), 102–109.

[9] Cisternas, M., Carvajal, M., Wesson, R., Ely L. L. & Gorigoitia N. (2017) Exploring the Historical Earthquakes Preceding the Giant 1960 Chile Earthquake in a Time-Dependent Seismogenic Zone. *Bulletin of the Seismological Society of America*; 107 (6): 2664–2675. doi: https://doi.org/10.1785/0120170103

[10] Clawpack Development Team (2018), Clawpack Version 5.5.10 http://www.clawpack.org, doi: 10.5281/zenodo.1405834

[11] Dura, T., Hemphill-Haley, E., Sawai, Y., Horton, B. P. (2016). The application of diatoms to reconstruct the history of subduction zone earthquakes and tsunamis. Earth-Science Reviews, 152, 181-197.

[12] Dura, T., Horton, B. P., Cisternas, M., Ely, L. L., Hong, I., Nelson, A. R., ... Nikitina, D. (2017). Subduction zone slip variability during the last millennium, south-central Chile. Quaternary Science Reviews, 175, 112–137. doi:10.1016/j.quascirev.2017.08.023

[13] Farr, T. G., and M. Kobrick, 2000, Shuttle Radar Topography Mission produces a wealth of data. Eos Trans. AGU, 81:583-583.

[14] Fukutani, Y., Suppasri, A., & Imamura, F. (2014). Stochastic analysis and uncertainty assessment of tsunami wave height using a random source parameter model that targets a Tohoku-type earthquake fault. Stochastic Environmental

[15] Fukutani, Y., Suppasri, A., & Imamura, F. (2018). Quantitative Assessment of Epistemic Uncertainties in Tsunami Hazard Effects on Building Risk Assessments. Geosciences, 8(1),17. doi:10.3390/geosciences8010017

[16] Galaz J., Cienfuegos, R., Echeverria A., Pereira S., Bertin C., Prato G., Karich J. (2019) Integrating tsunami simulations in web applications using Nami, an open source client side GPU powered tsunami simulation library

[17] Garrett, E., Shennan, I., Woodroffe, S.A., Cisternas, M., Hocking, E.P. & Gulliver, P. (2015) Reconstructing paleoseismic deformation, 2: 1000 years of great earthquakes at Chucalén, south central Chile. *Quaternary Science Reviews*, 113, pp. 112-122.

[18] Geller, R. J. (1976). Scaling relations for earthquake source parameters and magnitudes. *Bulletin of the Seismological Society of America*, 66(5):1501–1523.

[19] Global Historical Tsunami Database, NCEI. doi:10.7289/V5PN93H7 Last accessed 31st August, 2021

[20] Goda, K., Song, J. (2016). Uncertainty modeling and visualization for tsunami hazard and risk mapping: a case study for the 2011 Tohoku earthquake. *Stoch. Environ. Res. Risk Assess* 30, 2271–2285 (2016). https://doi.org/10.1007/s00477-015-1146-x

[21] Grezio, A., Babeyko, A., Baptista, M. A., Behrens, J., Costa, A., Davies, G.,... Thio, H. K. (2017). Probabilistic Tsunami Hazard Analysis: Multiple sources and global applications. *Reviews of Geophysics*, 55, 1158– 1198. https://doi.org/10.1002/2017RG000579

[22] Gusiakov, V.K.(2011) Relationship of tsunami intensity to source earthquake magnitude as retrieved by historical data. Pure Appl. Geophys. **168**, 2033-2041 (2011)

[23] Gusman, A. R., Supendi, P., Nugraha, A. D., Power, W., Latief, H., Sunendar, H., et al. (2019). Source model for the tsunami inside Palu Bay following the 2018 Palu earthquake, Indonesia. Geophysical Research Letters, 46, 8721– 8730. https://doi.org/10.1029/2019GL082717

[24] Hayes, G., 2018, Slab2 - A Comprehensive Subduction Zone Geometry Model: U.S. Geological Survey data release, https://doi.org/10.5066/F7PV6JNV.

[25] Hocking, E. P., Garrett, E., & Cisternas, M. (2017). Modern diatom assemblages from Chilean tidal marshes and their application for quantifying deformation during past great earthquakes. *Journal of Quaternary Science*, 32(3), 396–415. doi:10.1002/jqs.2933

[26] Hong, I., Dura, T., Ely, L. L., Horton, B. P., Nelson, A. R., Cisternas, M., Nikitina, D., & Wesson, R. L. (2017). A 600-year-long stratigraphic record of tsunamis in south-central Chile. *The Holocene*, 27(1), 39–51. https://doi.org/10.1177/0959683616646191

[27] Kanamori, H., Anderson, D. L. (1975). Theoretical basis of some empirical relations in seismology. Bulletin of the seismological society of America, 65(5), 1073-1095.

[28] Kempf, P., Moernaut, J., Van Daele, M., Vandoorne, W., Pino, M., Urrutia, R., & De Batist, M. (2017). Coastal lake sediments reveal 5500 years of tsunami history in south central Chile. Quaternary Science Reviews, 161, 99-116

[29] Kulkarini, V., Arcos, M.E.M., Alcinov, T. (2016). Probabilistic Tsunami Hazard Assessment for a Site in Eastern Canada. *Pure Appl. Geophys.* 173, 3719–3755 (2016). https://doi.org/10.1007/s00024-016-1414-9

[30] Lay, T., Kanamori, H., Ammon, C. J., Koper, K. D., Hutko, A. R., Ye, L., Yue, H., & Rushing, T. M. (2012), Depth-varying rupture properties of subduction zone megathrust faults, *J. Geophys. Res.*, 117, B04311, doi:10.1029/2011JB009133.

[31] LeVeque R. J., Waagan, K., González F. I., Rim, D., Lin, G. (2016) Generating Random Earthquake Events for Probabilistic Tsunami Hazard Assessment. In: Geist E.L., Fritz H. M., Rabinovich A. B., Tanioka, Y. (eds) Global Tsunami Science: Past and Future, Volume I. Pageoph Topical Volumes. Birkhäuser, Cham.

[32] Lomnitz, C. (2004). Major Earthquakes of Chile: A Historical Survey, 1535-1960. Seismological Research Letters, 75(3), 368-378. doi:10.1785/gssrl.75.3.368

[33] Mai, P. M. & 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.

[34] Mandli, K.T., Ahmadia, A.J., Berger, M.J., Calhoun, D.A., George, D.L., Hadjimichael, Y., Ketcheson, D.I., Lemoine, G.I., LeVeque, R.J.,(2016). Clawpack: building an open source ecosystem for solving hyperbolic PDEs. PeerJ Computer Science. doi:10.7717/peerj-cs.68

[35] Melnick, D., Cisternas, M., Moreno, M., Norambuena, R., 2012. Estimating coseismic coastal uplift with an intertidal mussel: calibration for the 2010 Maule Chile earthquake (Mw = 8.8). Quat. Sci. Rev. 42, 29–42. https://doi.org/10.1016/j.quascirev.2012.03.012.

[36] Molina, D., Tassara, A., Melnick D., Abarca R., Madella, A., (2020) Frictional segmentation of the Chilean megathrust from a multivariate analysis of geophysical, geological and geodetic data. In preparation.

[37] Moreno, M., Haberland, C., Oncken, O., Rietbrock, A., Angiboust, S., & Heidbach, O. (2014).Locking of the Chile subduction zone controlled by fluid pressure before the 2010 earthquake. Nature Geoscience, 7(4), 292–296. doi:10.1038/ngeo2102

[38] Murotani, S., H. Miyake, & K. Koketsu (2008), Scaling of characterized slip models for plate-boundary earthquakes, *Earth Planets Space*, 60, 987–991.

[39] Murotani, S., Satake, K., & Fujii, Y. (2013). Scaling relations of seismic moment, rupture area, average slip, and asperity size for M∼9 subduction-zone earthquakes. Geophysical Research Letters, 40(19), 5070–5074. doi:10.1002/grl.50976

[40] Ormeño, F. (2021). Caracterización de las fuentes sísmicas más probables para terremotos históricos. Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Concepción para optar al título profesional de Geofísica. Universidad de Concepción.

[41] Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the seismological society of America*, 75(4):1135–1154.

[42] Purcaru, G. and Berckhemer, H. (1982). Quantitative relations of seismic source parameters and a classification of earthquakes. *Tectonophysics*, 84(1):57–128.

[43] Qin, X., LeVeque, R., Motley, M., (2019) Efficient Tsunami Modeling on Adaptive Grids with Graphics Processing Units (GPUs) Preprint ArXiv.com

[44] Ryan, W. B. F., et al. (2009), Global Multi-Resolution Topography synthesis, Geochem. Geophys. Geosyst., 10, Q03014, doi:10.1029/2008GC002332.

[45] Schambach, L., Grilli, S. T., Tappin, D. R. (2018), New simulations and understanding of the 1908 Messina tsunami for a dual seismic and deep submarine mass failure source, Marine Geology.

[46] Servicio Hidrográfico y Oceanográfico de la Armada de Chile, SHOA (2000). El maremoto del 22 de mayo de 1960 en las costas de Chile.

[47] Smart, G. M., Crowley, K. H. M., & Lane, E. M. (2016). Estimating tsunami run-up. Natural Hazards, 80(3), 1933-1947.

[48] Skarlatoudis, A., Somerville, P., and Thio, H. (2016). Source-scaling relations of interface subduction earthquakes for strong ground motion and tsunami simulation. *Bulletin of the Seismological Society of America*, 106(4):1652–1662.

[49] Song M., Cho Y. (2020) Probabilistic Tsunami Heights Model using Bayesian Machine Learning. Journal of Coastal Research, 95(sp1), 1291-1296

[50] Synolakis, C. E. (1991). Tsunami runup on steep slopes: How good linear theory really is. In Tsunami Hazard, pp 221-234. Springer

[51] Thingbaijam, K. K. S., Martin Mai, P., and Goda, K. (2017). New empirical earthquake source-scaling laws. *Bulletin of the Seismological Society of America*, 107(5):2225–2246.

[52] Tsuji, Y., (2013). Catalog of distant tsunamis reaching Japan from Chile and Peru. Tsunami Eng. 30, 61–68.

[53] Wessel, P., and W. H. F. Smith (1998), New improved version of Generic Mapping Tools released, EOS *Trans.*, AGU ,79, 579.