Stochastic Environmental Research and Risk Assessment
https://doi.org/10.1007/s00477-023-02397-1

ORIGINAL PAPER

A stochastic approach to the characterization of the seismic sources: a potential method for the assessment of sources of historical and paleo tsunami

Rodrigo Cifuentes-Lobos[1] · Ignacia Calisto[1] · Breanyn MacInnes[2] · Marcos Moreno[1] · Jorge Quezada[3] · Javiera San Martín[1,4] · Matías Fernández-Palma[1] · Cristian Saavedra[2]

Accepted: 4 February 2023

Abstract
Inundation, wave arrival and deformation data gathered from measurement campaigns and historical accounts have been used to study past earthquakes and tsunamis, being crucial for magnitude and extension estimations. Nevertheless, the most commonly used methods cannot provide more information about the spatial characteristics of the event's slip distribution. In this work, we aim to create a methodology for obtaining more realistic, heterogeneous estimations of the slip distribution characteristics of past tsunamigenic earthquakes, by obtaining slip patterns of each earthquake of the same seismic segment, thus establishing a history of the seismic cycle in a rupture zone. To obtain the stochastic characterization of the slip distribution of the seismic source, we combine fault defining parameters using a logic tree structure to generate random slip distributions which are subsequently submitted to successive restrictions to assess their compliance with the available deformation and tsunami data, discarding those that do not satisfy the constraints and using those that do to estimate the most probable seismic source in terms of the data. We test this methodology with synthetic heterogeneous slip fields off the coast of South Central Chile to assess its limitations, data dependence, resolution capabilities and uncertainties. Obtaining a good correlation between the synthetic observations and the locations of the main slip features. Finally, the 9.5 Mw, 1960 Valdivia earthquake is used to benchmark the methodology with real deformation and tsunami data from surveys and historical accounts. This result is then compared to slip distributions obtained with tsunami, seismological and deformation data inversions. This shows that the magnitude of the earthquake could be recovered correctly and the slip distribution is very similar to models calculated from different techniques.

Keywords Slip distribution · Stochastic · Estimation · Paleotsunami

1 Introduction

Studies of historical and prehistorical earthquakes and tsunamis are generally done using detailed revisions of historical documents and accounts (Stewart 2019), 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). Measurements from modern techniques, such as diatom and plankton analysis (Dura et al. 2016), are of incredible value and a necessity in characterizing of past earthquakes and tsunamis. However, these studies are commonly hard and expensive to deploy and require a large workforce due to logistic difficulties. Other great efforts in reassessing the location, rupture mechanism, magnitude and tsunami magnitudes have been done using backward raytracing (e.g. Baptista et al. 1998, 2006) for constraining the source of historical earthquakes. The use of macroseismic intensity data points (Baptista et al. 2006; Wronna et al. 2021), the study of plate kinematics in the area of interest (Wronna et al. 2019) and the use of previously proposed sources (Reis et al. 2017) could complement this methodology in constraining the possible sources. On the other hand, several numerical techniques and tools have been developed that can deal with uncertainties (LeVeque et al. 2016; Grezio et al. 2017), random or epistemic, especially within the Probabilistic Tsunami Hazard Assessment

Ignacia Calisto, Breanyn Macinnes, Marcos Moreno and Jorge Quezada contributed equally to this work.

Extended author information available on the last page of the article

Published online: 01 March 2023

(PTHA) discipline (Fukutani et al. 2015; Kulkarni 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. 2015; Goda and Song 2016; Fukutani et al. 2018) and its use in the study of past earthquakes is an opportunity seldom used. This study proposes using the LT approach to characterize the slip distribution and rupture limits of seismic sources of a past tsunami. This can be achieved by replacing branches associated with hazard assessment, such as recurrence interval (Annaka et al. 2007), with branches to mitigate the uncertainty in the rupture geometry, for example, the along-strike limits, northern and southern limits in the chilean subduction zone, and the rupture area's aspect ratio, to achieve more realistic, heterogeneous models. This way, seismic sources are characterized by their moment magnitude, length, width (aspect ratio and along-strike limits), distance to trench and slip distribution. Random slip distributions are generated using a combination of source defining parameters given by the LT. The approach used in this study corresponds to a decision-schema organised in several levels that part ways, separating themselves as "branches" of a tree, with each level representing a source defining parameter. This structure is built upon a mutually exclusive ramification principle, so as for each generation of a random model, one and only one branch combination is active. This vast number of models is necessary to compensate for empirical and random uncertainties (LeVeque et al. 2016; Grezio 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, as the latter tends to underestimate tsunami intensities (Carvajal et al. 2017), leading to an underrepresentation of the importance of the tsunami data, leading to a possible overestimation of the earthquake's moment magnitude to compensate the lower tsunami intensities. The scope of this methodology is to shine a light on the spatial aspects of the slip distributions of past earthquakes, aiming to be a tool for characterizing of seismic segment's slip patterns, deepening our understanding of the seismic cycle.

2 Method

2.1 Generation of random tsunami sources

To produce hypothetical rupture scenarios, we generate random slip distributions 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 and Beroza 2002; LeVeque et al. 2016), representing a linear combination of eigenvectors and eigenvalues of a covariance matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$, which in turn is a function of the distribution of the subfaults that define the fault geometry. This distribution of subfaults for the generation of stochastic earthquake scenarios is achieved with a LT structure as shown in Fig. 1 (Fukutani et al. 2015; Goda and Song 2016; Fukutani et al. 2018). The dip, strike and depth values for each subfault are obtained from the Slab2 model (Hayes 2018), while the rake values are computed from NUVEL-1 model (DeMets et al. 1990).

To construct a LT structure that generates appropriate random distributions, the choice of value ranges for the different branches of the LT is crucial to adequately estimate the most probable seismic source and not to overutilize computational costs. We propose the following steps to define the values Define a number of discrete subfaults to divide the rupture zone. Define a maximum rupture length according to the span of deformation data determining a buffer zone to the north and south, to accommodate for possible variations in range. According to this length, estimate its maximum seismic moment with scaling laws (Abe 1975; Geller 1976; Skarlatoudis et al. 2016; Thingbaijam et al. 2017) defining 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 (Lay et al. 2012). According to empirical relations (Kanamori and Anderson 1975; Purcaru and Berckhemer 1982) select a range of aspect ratios to configure 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, numbers around 20 work well. The amount of these values defines 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 amount of random distributions is generated, a vertical deformation field is computed for each one using ClawPack's implementation of the Okada (1985) model for computing

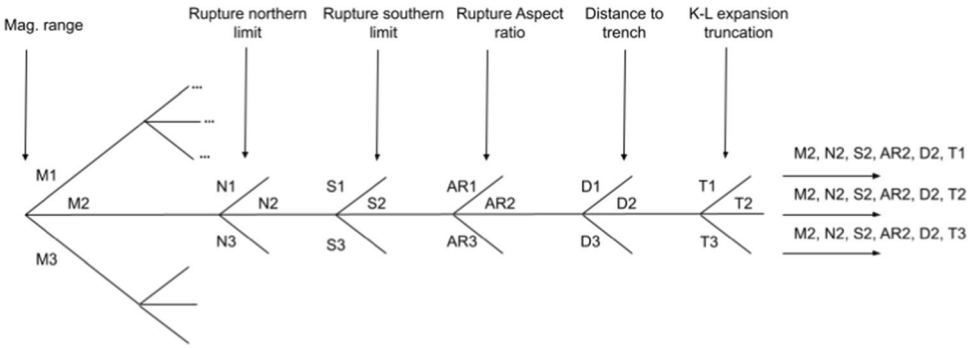

Fig. 1 Example of a logic tree structure used in the creation of random source models. The LT consists of the branches for the combinations of source parameters used to define the random models, magnitudes, northern and southern limits, for the case of the Chilean subduction zone, of the rupture, aspect ratio of the rupture, distance to trench of the updip limit and truncation level of the K–L expansion

seafloor deformation package on Python (Clawpack Development Team. 2018; Mandli et al. 2016).

2.2 Restrictions

LT structures can generate a vast number of random models. However, due to their random nature, not necessarily all of them will be able to represent a natural earthquake. To compensate this, it is necessary to impose restrictions on them to dismiss those that are not plausible, that is, the distributions that do not follow empirical nor source-scaling laws (e.g. Kanamori and 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 the physical characteristics of subduction earthquakes.

Every random model 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 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. Those that were modelled were subsequently filtered with tsunami observations. 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 crucial and affects the final result.

2.2.1 Deformation data restrictions

Deformation restrictions are subdivided into three steps. The first is to divide the maximum length of the possible rupture zone into n bands (see Fig. 5 left 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 on each band is then computed, and the band with more points is selected as reference, from this reference value, the relative density of the remaining bands is computed. If two bands have the same data density, the one with higher tsunami data density is preferred.

The first instance of restrictions considers a comparison between the absolute values of deformation and the slip.

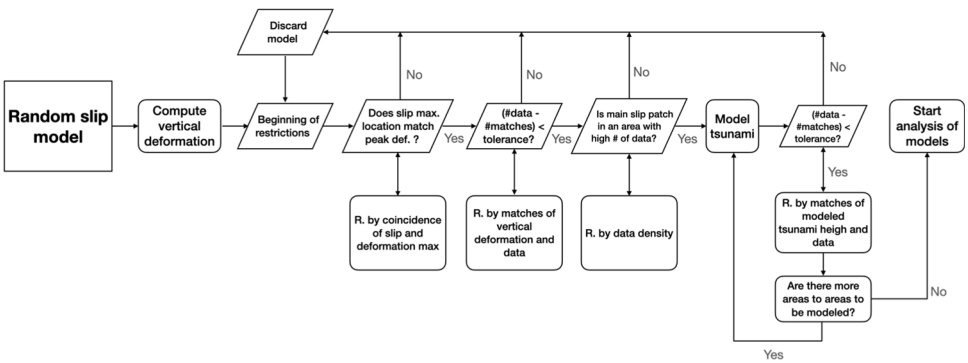

Fig. 2 Schematic flowchart of the restriction process

Along dip, latitudinal slip profiles of the random models are computed by integrating the slip fields along rows of equal latitude and then is compared to the profile of absolute values of deformation (see right panels of Fig. 5 (left)). 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 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 instance of restriction is done 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 these areas may appear as these can comply with the rest of the observations.

This step can be subdivided into several chained restrictions, 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, the model passes the restriction. If not, it is discarded. 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 amounts of uncertainties, the best quality data should be used in a first batch with a low tolerance value, and then continue restricting models with lesser quality data setting a higher threshold, thus the models will have to have more points compliant with the data.

2.2.2 Tsunami data restrictions

There are two steps before the tsunami modelling process. As the number of data points affects 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 slip distributions whose maximum slip value is located in the i-th band is discarded. The methodology discards 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. This ensures that most high valued slip features on the remaining models are located in areas where the data allows for resolving power.

The next step is to define the areas to model tsunamis and the order to model and restrict them. This depends on the availability of high resolution topobathymetric grids and the location of tsunami observations. The methodology is highly sensitive to the order of tsunami restrictions, as for each area modelled, the data has to comply locally and not globally, possibly skewing the estimations. For this reason, modelling should start in areas where tsunami heights values are higher overall and there is higher density of data. Values have priority over density. For example, in the case study, higher tsunami height values and densities are located near Corral ($\sim 39.8^\circ$ S) (Fig. 5b) 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.

Although inundation maps can be computed for on land 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 tsunami models are done, the maximum values of the virtual tide gauges are compared to the wave height, with a tolerance level to account for uncertainties 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.

2.3 Characterization and estimation of the seismic source

Once the random source models have been through every step of the restriction process, the remaining slip distributions are subject to two types of analysis to estimate the most probable source. First is a frequentist analysis of the branches of the LT that generated the remaining models, and second is the characterization of the slip distribution.

A frequentist analysis of the branches shows the most probable moment magnitude, geographic limits and aspect ratio. These values are used to define the span of the distribution. For this characterization, 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 obtained in the frequentist analysis. 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 characterization is Mo_c, the final slip distribution $\mathbf{S_f}$ is obtained by multiplying the characterized slip vector $\mathbf{S_c}$ by the ratio of the seismic moments of the most probable (Mo_{MP}) magnitude and the characterized 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,

$$\mathbf{S_f} = \left(\frac{Mo_{MP}}{Mo_c}\right)\mathbf{S_c} \tag{1}$$

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

3 Testing the methodology

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

3.1 Synthetic earthquakes

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. This synthetic distributions are done to test the importance of the data and the different steps in the restriction process. 120 deformation data points from the synthetic models are sampled with a uniform latitudinal distance along the coast of the rupture zone, for each case. For tsunami height data, tsunami propagation is computed solving linear, for parental grids, and nonlinear, for the higher resolution grid, shallow water equations using numerical model COMCOTv1.7 (Wang et al. 2011) 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', if 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 and Kobrick 2000). Bottom friction was added to the nested grid with higher resolution, with a Manning's coefficient of 0.025, also, for these higher resolution grids. 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 synthtetic model and 22 for the second case. Data points can be seen in Fig. 3.

3.2 Synthetic model estimations

Following the steps described in Sect. 2.2, 200,000 random models are generated for both cases. Then, both are subject to the methodology, clearing all restrictions, respectively, 14 and 9 models. Estimation and characterization of the distribution 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, as shown in Fig. 4. Estimated models tend to spread the total seismic moment onto more subfaults, instead of concentrating it on the displacement peaks as in the synthetic models. Thus, for the same moment magnitude, estimated slip distributions tend to underrepresent the maximum displacement values. This means that estimated magnitudes are often overestimated. This effect is attenuated when tsunami data for the restrictions are taken into account.

4 Valdivia case study

The biggest earthquake in recorded history 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 1000 km, from the Gulf of Arauco in the North ($\sim 37^\circ$ S) and the Ofqui Isthmus, in the South ($\sim 46^\circ$ S) (Benioff et al. 1961; Cifuentes 1989; Cifuentes and Silver 1989; Barrientos and 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 in a span of approximately 15 min (Cifuentes 1989; Cifuentes and Silver 1989). Coseismic and geodetic data from this event in the form of triangulations and relative sea level changes (Plafker and Savage 1970) (shown in Fig. 5a) are used as deformation input in the restriction process, accounting for 150 points. Historical accounts and sedimentological measurements were used to determine tsunami heights (Takahashi and Hatori

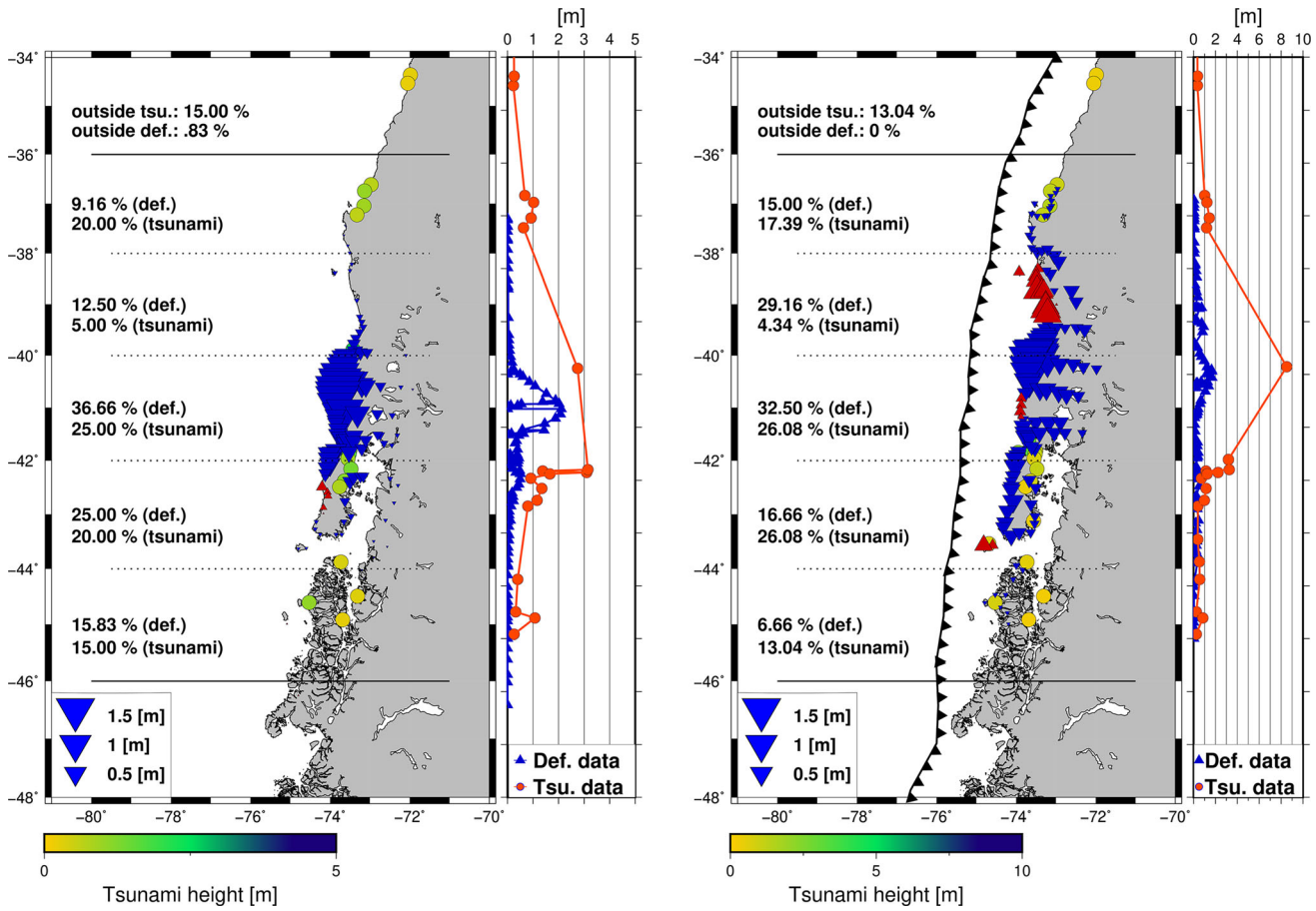

Fig. 3 Distribution of data sampled from synthetic deformation (shown in red upward triangles for uplift and blue downward triangles for subsidence) and tsunami models (shown in colored circles) for synthetic models I (left) and II (right)

1961; Weischet 1963; Sievers et al. 1963; Lockridge 1985; Fritz et al. 2011; Hernandez 2011; Starin et al. 2013 right and Data Set S5 in Supporting Information). As all 103 points available of tsunami data are located on land, offshore 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). We chose to reassess the 1960 Valdivia earthquake as a benchmark to test our methodology, as this event's slip distribution have been deeply studied with multidisciplinary inversion approaches using different datasets (e.g. Barrientos and Ward 1990; Moreno et al. 2009; Fujii and Satake 2013; Ho et al. 2019) thus we have a robust baseline to compare our results. This real event, in addition to the synthetic event, were used to validate our methodology. Logic Tree value ranges for the Valdivia case study and synthetic tests are available in Supporting Information S1. An alternative estimation using qualitative (uplift, subsidence or no change) data instead of quantitative is shown in Fig. S1 in Supporting Information.

4.1 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 Fig. 6. These 20 models satisfied the data available with the tolerances shown in Table S3 in Supporting Information. The variation of slip values on the main slip features is low, with the maximum slip values ranging from $\sim$30 to $\sim$45 ms. The estimations of the seismic source determined by our methodology are compared to previous slip models computed from inversions of seismic, deformation and tsunami data (Barrientos and Ward 1990; Moreno et al. 2009; Fujii and Satake 2013; Ho et al. 2019). Our slip distribution of the Valdivia 9.5 Mw earthquake shows that the most probable rupture spans from the Gulf of Arauco ($\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 ms,

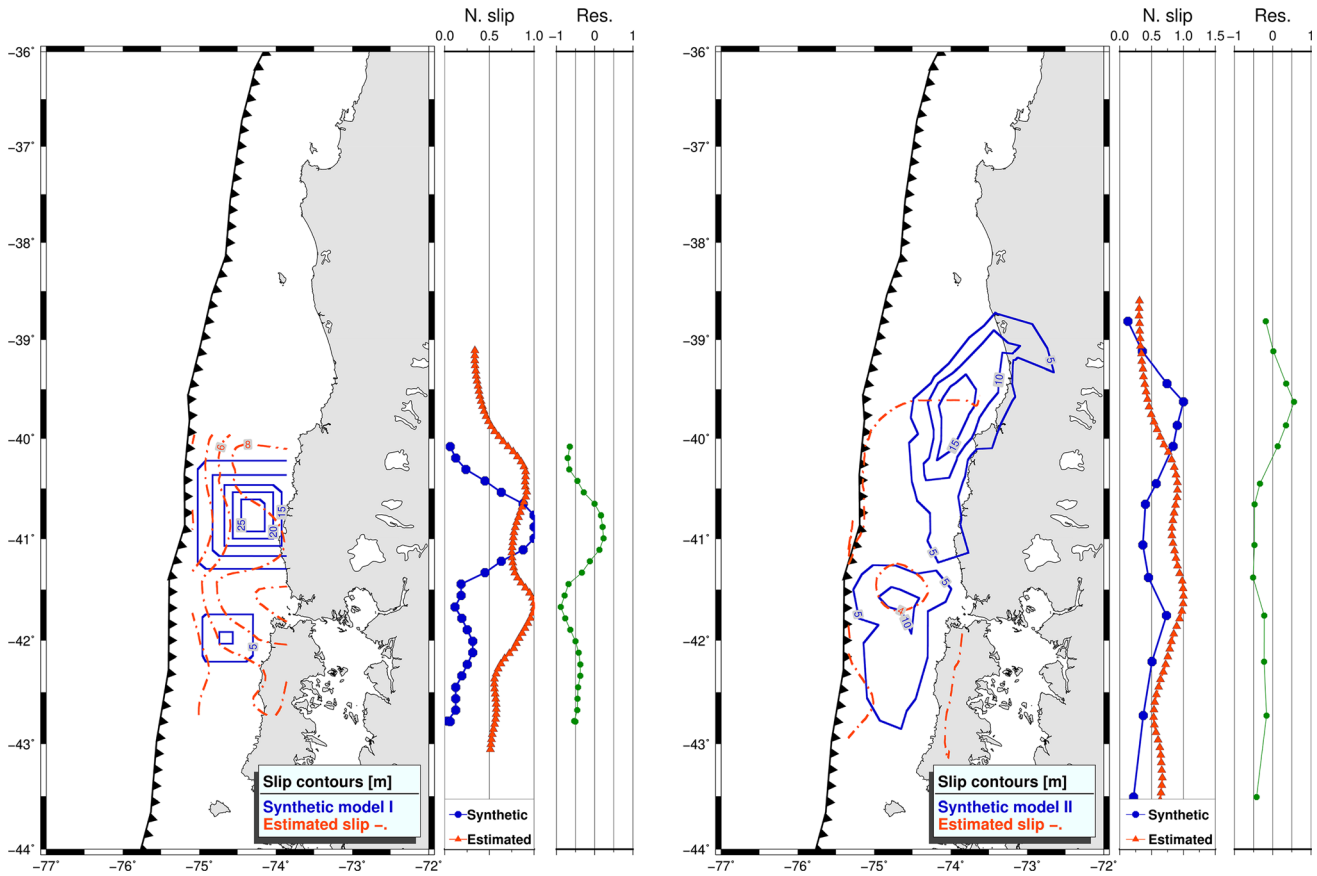

Fig. 4 Final estimations of the most probable synthetic models. Synthetic model I (left) and II (right) show concordance in the location of the main slip features, as it can be seen in the slip profiles

consistent with displacement concentrated in the northernmost part of the rupture in slip models from Ho et al. (2019), Fujii and Satake (2013) and Moreno et al. (2009), albeit not as large (Fig. 6a). In contrast with Moreno et al. (2009), Fujii and Satake (2013) and Ho et al. (2019) 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 ms, larger than showed by Barrientos and Ward (1990) and Moreno et al. (2009) (inversions that share most of the input data), but in line with the values obtained with Fujii and Satake (2013). Compared to the estimation by Ho et al. (2019), our results underestimate the slip south of Corral. Normalized latitudinal-integrated slip profiles (Fig. 6 b) show that both our estimated slip distribution and the three previous inversion results share the latitude of the maximum values, however our results do not resolve a secondary peak at $\sim 41°$ S shown in Moreno et al. (2009). Both profiles in Fig. 6b show a correlation value of $\rho = 0.50$ and the residuals a RMSE of 0.36.

The deformation computed from the estimated slip distribution (see Fig. 7) shows a larger uplift feature northward of the maximum deformation obtained by Barrientos and Ward (1990) or Moreno et al. (2009), in line with slip results. Our model tends to show similar or slightly lower uplift values in the southern area compared with other models, resolving the uplift measured in Guafo and Guamblin Islands. We obtain larger subsidence values in the coast of Chile than the aforementioned results. Nevertheless, due to the lack of deformation data off-shore, it is not possible to delve much deeper into these results.

Although the comparison between tsunami data points and maximum wave heights is done in absolute terms, with tolerance values in meters, to better illustrate the results we computed the percent error (see Fig. 8) for the data points near Corral ($\sim 39.8°$ S). Except four points, all modelled values are within 20% from the observed wave height. Modelled values tend to underestimate wave height with respect to measured values moved to the closest offshore grid point.

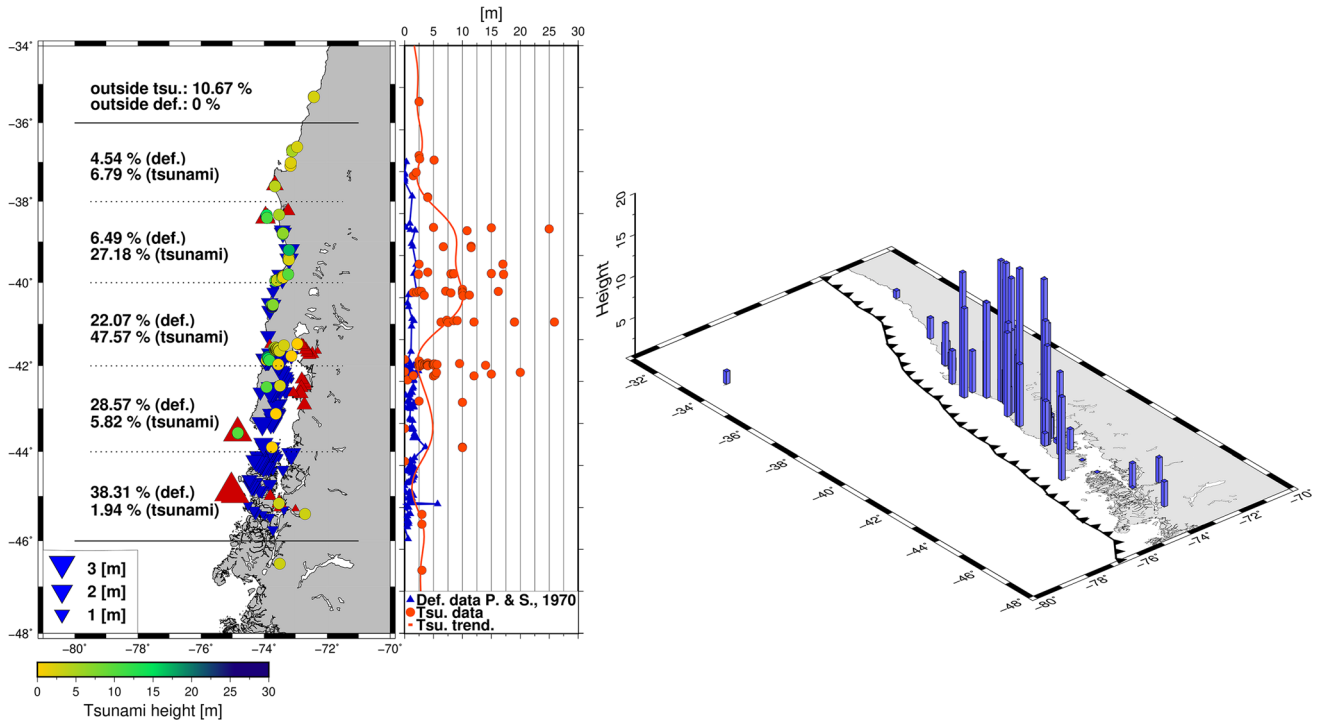

Fig. 5 (Left) Deformation and tsunami data available for the estimation of the seismic source of the 9.5 Mw earthquake of Validivia, 1960. Deformation data (blue downward triangles show areas of subsidence and red upward triangles show uplift) is from Plafker and Savage (1970) and tsunami data from Takahashi and Hatori (1961), Weischet (1963), Sievers et al. (1963), Lockridge (1985), Fritz et al. (2011), Hernandez (2011), Starin et al. (2013). Profiles to the right show the behaviour of data amplitude along latitude. (Right) Tsunami heights distribution

5 Discussion

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.

Choosing the right restriction parameters (thresholds/tolerances) influences the result. As setting higher thresholds allow 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. When choosing threshold values for checking point-to-point matches it is important to note that they 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 methodology may not be able to estimate with high resolution the main slip features, and would underestimate the maximum slip values. However, this estimation would show the location of this features.

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

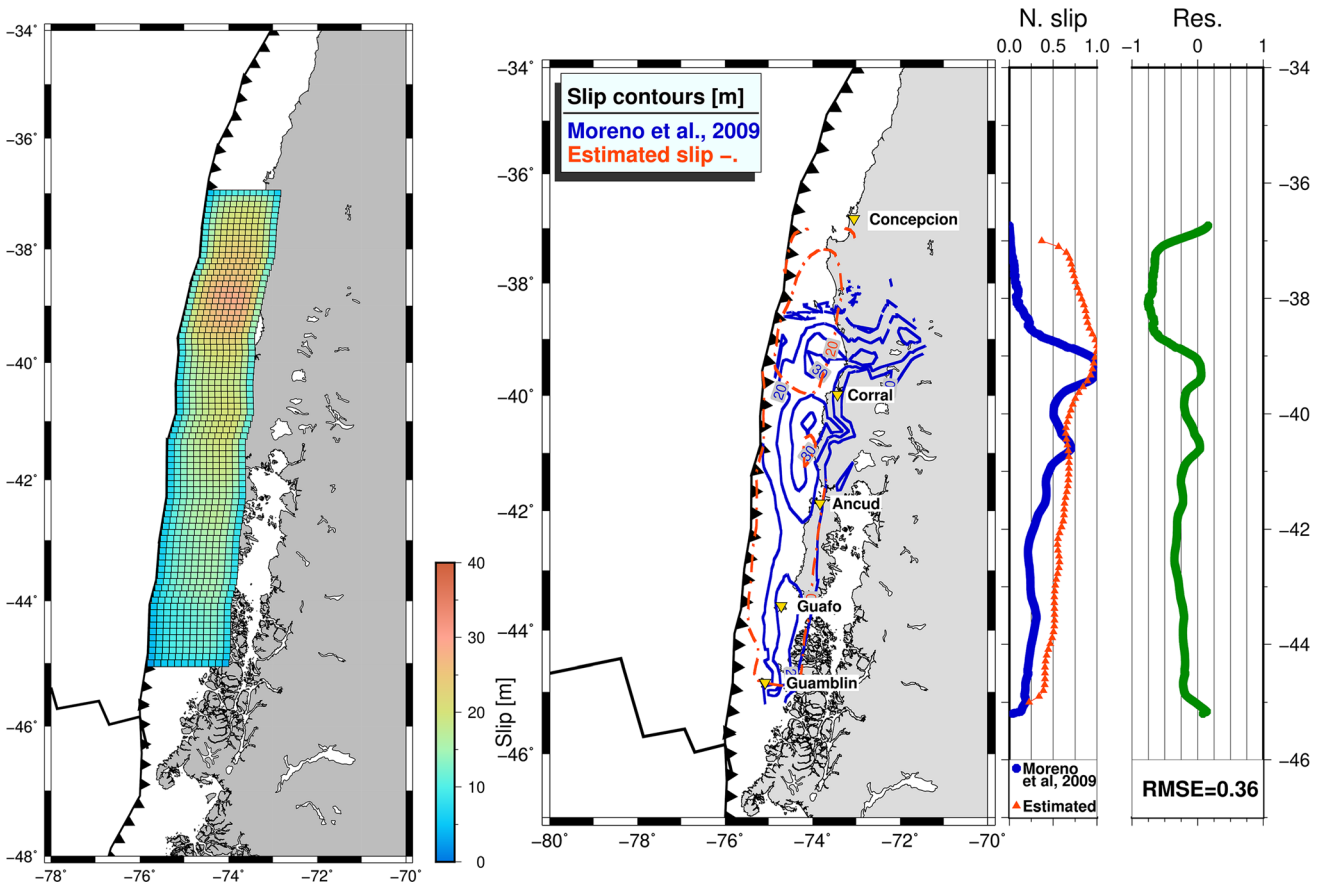

Fig. 6 Estimation of the most probable seismic source model for the Valdivia 1960 9.5 Mw earthquake in terms of the deformation and tsunami data available. Left figure shows the slip distribution of the smoothed subfault-wise maximisation of the PDF. (Right) shows a comparison between the estimation (dotted orange contours) with the inversion result from Moreno et al. (2009) (solid blue contours). Normalized along dip 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)

addition in tsunami height points may skew the estimation, based solely on seismic sources, to the specific area affected by it Bardet et al. (2003), Ulrich et al. (2019) and Nakata et al. (2020), among others, mention the importance of possible underwater landslide sources in tsunami generation, as some run-up measurements cannot be explained only by vertical deformation of the seafloor, such as the case of the tsunami in Palu Bay after the 2018 Sulawesi earthquake. The same can be said to amplification phenomena caused by site effects in bays or fjords, where this addition in height may cause that the estimated model present a high slip feature in this area.

This methodology of slip distribution estimation may allow for the characterization of sequences of earthquakes of the same seismic segment, providing information about possible spatial patterns in the seismic cycle. Thus, by knowing the behaviour of slip distributions across different earthquakes and the location of potentially recurrent slip maxima, it could prove to be a useful tool in seismic hazard assessment studies.

Due to the rectangular construction of the fault, deeper slip values than the downdip limit of the constructed fault cannot be directly resolved, however, its possible effect on measured deformation values may be accounted for by overestimating slip values in the same latitudes. For example in Valdivia case study, deep slip values obtained in the inversions at $\sim 39^\circ$ S cannot be resolved with our method, however, the excess of slip near the down dip rupture boundary near $\sim 39^\circ$ and in the northern limit can be related to the restrictions compensating for this deep values. 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

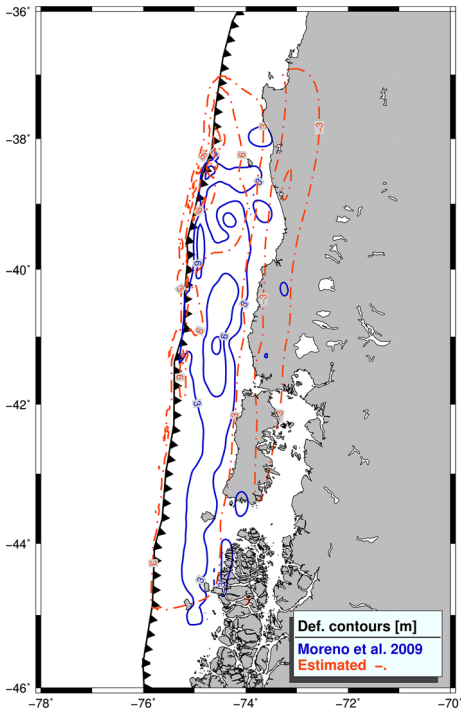

Fig. 7 Comparison of the deformation obtained computing the Okada (1985) solution to the estimated slip distribution (orange dotted contours) and the deformation from inversion results from Moreno et al. (2009) (blue solid contours)

distribution for Valdivia match the $Mw = 9.5$ value (Cifuentes and Silver 1989). This underestimation tends to underrepresent tsunami heights and inundation (see Fig. 8), but the estimation results shine a bright light on the areas that ruptured with greater amplitude in past earthquakes.

6 Conclusion

This methodology is designed for studying past earthquakes, and not to replace nor blank inversions. When modern seismic data is available, inversion methodologies are superior to this for computing slip distributions. However, the necessity of a tool that allows the leap from homogeneous, feature-lacking distributions to more realistic heterogeneous models of historic earthquakes; prior to the invention of seismic instruments, to better understand the behaviour of a seismic segment, makes this methodology a great opportunity to evolve from earlier estimations to newer ones that take advantage of the availability of historic observations. Estimated seismic sources results from this methodology are of good quality in terms of locating slip patches and estimating moment magnitude, albeit underestimating slip maxima from these patches. The main slip features of the 1960 Valdivia earthquake and synthetic models are resolved, constraining the location of the area where most of the seismic moment is concentrated. The rupture length and magnitudes are also resolved, obtaining exactly 9.5 for the Valdivia benchmark and overestimating by 0.1 each synthetic model. Overall, given the availability of deformation and tsunami data, be it quantitative or qualitative, characterizations of heterogeneous slip distributions is possible and of good quality. These results can give valuable insight in characterizing 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 characterization of areas with repeated high slip patches or areas where high slip values are not common. Heterogeneous, realistic slip

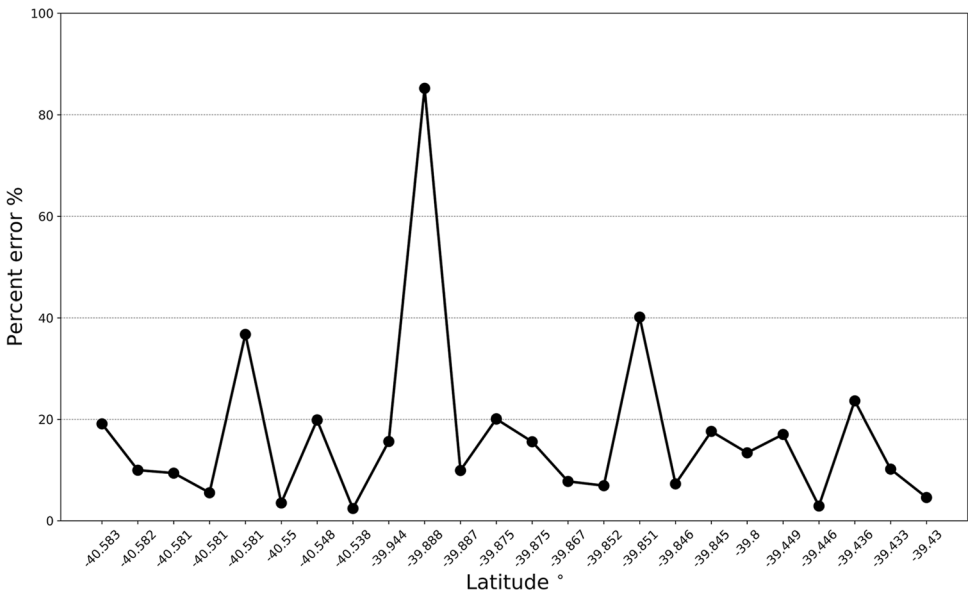

Fig. 8 Percent error between tsunami data points near Corral ($\sim 39.8°$ S) translated to water with the methodology proposed by Smart et al. (2016) and the maximum wave height modelled with our estimation. Modelled values tend to underestimate wave height

distributions may lead to a more effective hazard assessment and more prepared and resilient communities.

This methodology was originally designed to estimate slip distributions from subduction earthquakes. Specifically the codes were developed for the subduction zone of the Nazca plate beneath the South American plate. However, by changing the area of the Slab2 model to another and recomputing rake values, we can estimate seismic sources for any subduction zone. Furthermore, by having geometric information of a normal fault and deformation and/or tsunami data, this methodology can be extended to estimate a slip distribution on different tectonic settings.

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 and/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 (Clawpack Development Team. 2018; Mandli et al. 2016) that offer Adaptative Mesh Refinement 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.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00477-023-02397-1.

Acknowledgements 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 and Smith 1998). Please refer to Journal-level guidance for any specific requirements. We want to thank the reviewers for their insightful corrections.

Funding The present study is funded by FONDECYT project 11180854 "Source characterization for historical tsunamis of Central-Southern Chile", directed by Ignacia Calisto, Ph.D. Additional support and funding by Iniciativa Científica Milenio (ICM) through Grant Number NC160025 "Millenium Nucleus CYCLO: The seismic Cycle Along Subduction Zones" was provided.

Declarations

Competing interests All authors declare that they do not have any financial interests. Original idea by Ignacia Calisto. All authors contributed to the design of the estimation algorithm. Software design and implementation were made by Rodrigo Cifuentes-Lobos, Ignacia Calisto, Javiera San Martín and Matías Fernández-Palma. Data collection and preparation were done by Cristian Saavedra. This draft was written by Rodrigo Cifuentes-Lobos and all authors have commented and contributed in it. All authors have read and approved this manuscript. All codes (v1.0.0) for the stochastic estimation of slip distributions can be found on https://github.com/RCifuentesLobos/PaleoTsunami/tree/v1.0.0 (Cifuentes-Lobos and Calisto 2022). This code is built upon ClawPack's GeoClaw v5.5.10 (Clawpack Development Team. 2018; Mandli et al. 2016). All tsunami modeling was performed with COMCOTv1.7 (Wang et al. 2011).

References

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

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. In: Tsunami and its hazards in the Indian and Pacific Oceans. Springer, pp 577–592

Baptista MA, Miranda PMA, Miranda JM, Victor LM (1998) Constrains on the source of the 1755 Lisbon tsunami inferred from numerical modelling of historical data on the source of the 1755 Lisbon tsunami. J Geodyn 25(1–2):159–174. https://doi.org/10.1016/s0264-3707(97)00020-3

Baptista MA, Miranda JM, Luis JF (2006) In search of the 31 March 1761 earthquake and tsunami source. Bull Seismol Soc Am 96(2):713–721. https://doi.org/10.1785/0120050111

Bardet JP, Synolakis CE, Davies HL, Imamura F, Okal EA (2003) Landslide tsunamis: recent findings and research directions. In: Bardet JP, Imamura F, Synolakis CE, Okal EA, Davies HL (eds) Landslide tsunamis: recent findings and research directions. Pageoph Topical Volumes. Birkhäuser, Basel. https://doi.org/10.1007/978-3-0348-7995-8_1

Barrientos SE, Ward SN (1990) The 1960 Chile earthquake: inversion for slip distribution from surface deformation. Geophys J Int 103(3):589–598

Becerra I, Aránguiz R, González J, Benavente R (2020) An improvement of tsunami hazard analysis in central Chile based on stochastic rupture scenarios. Coast Eng J 62(4):473–488

Benioff H, Press F, Smith S (1961) Excitation of the free oscillations of the Earth by earthquakes. J Geophys Res 66(2):605–619. https://doi.org/10.1029/JZ066i002p00605

Carvajal M, Cisternas M, Catalán P (2017) Source of the 1730 Chilean earthquake from historical records: implications for the future tsunami hazard on the coast of metropolitan Chile. J Geophys Res Solid Earth 122(5):3648–3660

Cifuentes IL (1989) The 1960 Chilean earthquakes. J Geophys Res Solid Earth 94(B1):665–680

Cifuentes IL, Silver PG (1989) Low-frequency source characteristics of the great 1960 Chilean earthquake. J Geophys Res Solid Earth 94(B1):643–663

Cifuentes-Lobos R, Calisto I (2022) [software]. https://github.com/RCifuentesLobos/PaleoTsunami/tree/v1.0.0. https://doi.org/10.5281/zenodo.6998929

Cisternas M, Carvajal M, Wesson R, Ely LL, Gorigoitia N (2017) Exploring the historical earthquakes preceding the giant 1960 Chile earthquake in a time-dependent seismogenic zoneexploring the historical earthquakes preceding the giant 1960 Chile earthquake. Bull Seismol Soc Am 107(6):2664–2675

Clawpack Development Team (2018) Clawpack software. http://www.clawpack.org (Version 5.5.10). https://doi.org/10.5281/zenodo448.1405834

DeMets C, Gordon RG, Argus DF, Stein S (1990) Current plate motions. Geophys J Int 101(2):425–478. https://doi.org/10.1111/j.1365-246X.1990.tb06579.x

Dura T, Hemphill-Haley E, Sawai Y, Horton BP (2016) The application of diatoms to reconstruct the history of subduction zone earthquakes and tsunamis. Earth Sci Rev 152:181–197

Farr TG, Kobrick M (2000) Shuttle radar topography mission produces a wealth of data. Eos Trans Am Geophys Union 81(48):583–454

Fritz HM, Petroff CM, Catalán PA, Cienfuegos R, Winckler P, Kalligeris N et al (2011) Field survey of the 27 February 2010 Chile tsunami. Pure Appl Geophys 168(11):1989–2010

Fujii Y, Satake K (2013) Slip distribution and seismic moment of the 2010 and 1960 Chilean earthquakes inferred from tsunami

waveforms and coastal geodetic data. Pure Appl Geophys 170(9):1493–1509
Fukutani Y, Suppasri A, Imamura F (2015) Stochastic analysis and uncertainty assessment of tsunami wave height using a random source parameter model that targets a Tohoku-type earthquake fault. Stoch Env Res Risk Assess 29(7):1763–1779
Fukutani Y, Suppasri A, Imamura F (2018) Quantitative assessment of epistemic uncertainties in tsunami hazard effects on building risk assessments. Geosciences 8(1):17
Garrett E, Shennan I, Woodroffe S, Cisternas M, Hocking E, Gulliver P (2015) Reconstructing paleoseismic deformation, 2: 1000 years of great earthquakes at chucalén South Central Chile. Quat Sci Rev 113:112–122
Geller RJ (1976) Scaling relations for earthquake source parameters and magnitudes. Bull Seismol Soc Am 66(5):1501–1523
Goda K, Song J (2016) Uncertainty modeling and visualization for tsunami hazard and risk mapping: a case study for the 2011 Tohoku earthquake. Stoch Env Res Risk Assess 30(8):2271–2285
Grezio A, Babeyko A, Baptista MA, Behrens J, Costa A, Davies G et al (2017) Probabilistic tsunami hazard analysis: multiple sources and global applications. Rev Geophys 55(4):1158–1198
Hayes G (2018) Slab2: a comprehensive subduction zone geometry model: U.S. Geological survey data release. https://doi.org/10.5066/F7PV6JNV
Hernandez J (2011) 1960. memorias de un desastre. Arte Sonoro Austral Ediciones
Ho T-C, Satake K, Watada S, Fujii Y (2019) Source estimate for the 1960 Chile earthquake from joint inversion of geodetic and transoceanic tsunami data. J Geophys Res Solid Earth 124:2812–2828. https://doi.org/10.1029/2018JB016996
Hocking EP, Garrett E, Cisternas M (2017) Modern diatom assemblages from Chilean tidal marshes and their application for quantifying deformation during past great earthquakes. J Quat Sci 32(3):396–415
Hong I, Dura T, Ely LL, Horton BP, Nelson AR, Cisternas M, Wesson RL (2017) A 600-year-long stratigraphic record of tsunamis in South-Central Chile. Holocene 27(1):39–51
Kanamori H, Anderson DL (1975) Theoretical basis of some empirical relations in seismology. Bull Seismol Soc Am 65(5):1073–1095
Kulkarni V, Arcos ME, Alcinov T, Lavine A, Youngs R, Roussel P, Mullin D (2016) Probabilistic tsunami hazard assessment for a site in eastern Canada. Pure Appl Geophys 173(12):3719–3755
Lay T, Kanamori H, Ammon CJ, Koper KD, Hutko AR, Ye L, et al (2012) Depth-varying rupture properties of subduction zone megathrust faults. J Geophys Res Solid Earth 117(B4)
LeVeque RJ, Waagan K, González FI, Rim D, Lin G (2016) Generating random earthquake events for probabilistic tsunami hazard assessment. In: Global tsunami science: past and future, volume I. Springer, pp 3671–3692
Lockridge PA (1985) Tsunamis in peru-chile (Vol 39). The Center
Lomnitz C (2004) Major earthquakes of Chile: a historical survey, 1535–1960. Seismol Res Lett 75(3):368–378
Mai PM, Beroza GC (2002) A spatial random field model to characterizen complexity in earthquake slip. J Geophys Res Solid Earth 107(B11):ESE-10
Mandli KT, Ahmadia AJ, Berger M, Calhoun D, George DL, Hadjimichael Y, LeVeque RJ (2016) Clawpack: building an open source ecosystem for solving hyperbolic PDEs. PeerJ Comput Sci 2:e68
Moreno MS, Bolte J, Klotz J, Melnick D (2009) Impact of megathrust geometry on inversion of coseismic slip from geodetic data: application to the 1960 Chile earthquake. Geophys Res Lett 36(16)
Nakata K, Katsumata A, Muhari A (2020) Submarine landslide source models consistent with multiple tsunami records of the 2018 Palu tsunami, Sulawesi, Indonesia. Earth Planets Space 72:44. https://doi.org/10.1186/s40623-020-01169-3
Okada Y (1985) Surface deformation due to shear and tensile faults in a half-space. Bull Seismol Soc Am 75(4):1135–1154
Plafker G, Savage JC (1970) Mechanism of the Chilean earthquakes of may 21 and 22, 1960. Geol Soc Am Bull 81(4):1001–1030
Purcaru G, Berckhemer H (1982) Quantitative relations of seismic source parameters and a classification of earthquakes. Tectonophysics 84(1):57–128
Reis C, Omira R, Matias L, Baptista MA (2017) On the source of the 8 May 1939 Azores earthquake–tsunami observations and numerical modelling. Geomat Nat Haz Risk 8(2):328–347. https://doi.org/10.1080/19475705.2016.1218944
Ryan WB, Carbotte SM, Coplan JO, O'Hara S, Melkonian A, Arko R et al (2009) Global multi-resolution topography synthesis. Geochem Geophys Geosyst 10(3)
Sievers C, Hellmuth A, Villegas CG, Barros G (1963) The seismic sea wave of 22 May 1960 along the Chilean coast. Bull Seismol Soc Am 53(6):1125–1190
Skarlatoudis A, Somerville P, Thio H (2016) Source-scaling relations of interface subduction earthquakes for strong ground motion and tsunami simulation. Bull Seismol Soc Am 106(4):1652–1662
Smart G, Crowley K, Lane E (2016) Estimating tsunami run-up. Nat Hazards 80(3):1933–1947
Starin F et al (2013). The 1960 tsunami on beach ridge plains near Maullín, Chile: Landward descent, renewed breaches, aggraded fans, multiple predecessors. Andean Geology
Stewart Daniel M (2019) Historical tsunamis in the Concepcion bay, as seen in the reconstructed flood levels from the colonial city of Concepcion (Penco), Chile (1570–1835). Revista de historia (Concepción) 26(2):97–127. https://doi.org/10.4067/S0717-88322019000200097
Takahashi R, Hatori T (1961) A summary report on the Chilean tsunami of may 1960. Report on the Chilean Tsunami 23–34
Thingbaijam KKS, Mai PM, Goda K (2017) New empirical earthquake source-scaling LawsNew empirical earthquake source-scaling laws. Bull Seismol Soc Am 107(5):2225–2246
Ulrich T, Vater S, Madden EH et al (2019) Coupled, Physics-Based Modeling Reveals Earthquake Displacements are Critical to the 2018 Palu, Sulawesi Tsunami. Pure Appl. Geophys. 176:4069–4109. https://doi.org/10.1007/s00024-019-02290-5
Wang X, Power W, Staff GSN (2011). Comcot: A tsunami generation propagation and run-up model. GNS Science. https://books.google.cl/books?id=ioXPoQEACAAJ543
Weischet W (1963) Further observations of geologic and geomorphic changes resulting from the catastrophic earthquake of May 1960 in Chile. Bull Seismol Soc Am 53(6):1237–1257
Wronna M, Baptista MA, Miranda JM (2019) Reanalysis of the 1761 transatlantic tsunami. Nat Hazards Earth Syst Sci 19:337–352. https://doi.org/10.5194/nhess-19-337-2019
Wronna M, Baptista MA, Miranda JM (2021) Reevaluation of the 11 November 1858 earthquake and tsunami in Setúbal: a contribution to the seismic and tsunami hazard assessment in southwest Iberia. Pure Appl Geophys 178:4717–4742. https://doi.org/10.1007/s00024-021-02885-x

Authors and Affiliations

Rodrigo Cifuentes-Lobos[1] · Ignacia Calisto[1] · Breanyn MacInnes[2] · Marcos Moreno[1] · Jorge Quezada[3] · Javiera San Martín[1,4] · Matías Fernández-Palma[1] · Cristian Saavedra[2]

✉ Ignacia Calisto
icalisto@dgeo.udec.cl

Rodrigo Cifuentes-Lobos
rodcifuentes@udec.cl

Breanyn MacInnes
breanyn.macinnes@cwu.edu

Marcos Moreno
marcosmoreno@udec.cl

Jorge Quezada
jquezad@udec.cl

Javiera San Martín
jasanmartin2016@udec.cl

Matías Fernández-Palma
matiafernandez@udec.cl

Cristian Saavedra
crisaavedra@udec.cl

[1] Geophysics Department, University of Concepcion, Avda. Esteban Iturra s/n - Barrio Universitario, Concepción 4081447, Biobio, Chile

[2] Department of Geological Sciences, Central Washington University, 400 E. University Way, Ellensburg 98926, WA, USA

[3] Departamento de Ciencias de la Tierra, Universidad de Concepción, Victor Lamas 1290, Concepción 4081447, Biobio, Chile

[4] CTPA-Putemun, Instituto de Fomento Pesquero (IFOP), Castro - Chiloe, Chile