

An Updated Recurrence Model for Chilean Subduction Seismicity and Statistical Validation of its Poisson Nature

Alan Poulos¹, Mauricio Monsalve¹, Natalia Zamora¹, and Juan Carlos de la Llera^{1,2}

¹National Research Center for Integrated Natural Disaster Management (CIGIDEN), Santiago, Chile

²Department of Structural and Geotechnical Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile.

Abstract

Earthquake recurrence models are the basis of seismic hazard analysis and seismic risk evaluation of physical infrastructure. They are based on statistical analysis of earthquake occurrence data available in a specific geographical region. This work proposes a new earthquake recurrence model for the interface and intraslab seismicity of the subduction margin along Chile. The model improves some of the shortcomings of previous available models in the region, such as the lack of earthquake declustering or the use of magnitude scales inconsistent with modern ground motion prediction equations. Significant differences in seismic rates are found with some previous models. Indeed, the resulting frequencies from the Gutenberg-Richter relations are similar to some of the previous works, but also one order of magnitude higher and lower than two of the previously reported models. Because one of the strongest assumptions in earthquake occurrence models is that they follow a homogeneous Poisson process, this hypothesis is statistically tested herein, finding that the declustered catalog only partially complies with this assumption, showing for instance that the inter-event times follow approximately an exponential distribution.

Introduction

A crucial step in evaluating seismic risk of a given infrastructure is to assess the seismic hazard using Probabilistic Seismic Hazard Analysis, PSHA (Cornell, 1968; McGuire, 2004). This requires recurrence models of earthquakes for all possible seismic sources and ground motion prediction equations (GMPEs) to relate variables of the seismic source with the seismic intensity at a site. Both of these inputs to PSHA are calibrated using historical earthquake data. In this work the main focus is given to recurrence models for subduction seismicity in Chile since recent studies have already provided new GMPEs for subduction earthquakes in the region (e.g., Idini et al., 2017; Montalva et al., 2017).

A common assumption in seismic hazard and risk analyses is that mainshock occurrences follow a homogeneous Poisson process. This process is mathematically simple, stationary (statistical properties are constant over time), and memoryless (events occur independently of each other in time). The constant frequency rate and memoryless assumptions of this random process facilitate the use of recurrence models by researchers, practitioners, and seismic codes, and enable seismic hazard to be estimated as annual rates of exceedance for particular intensity thresholds (Cornell, 1968). However, the memoryless assumption requires the elimination of all temporal correlations in the earthquake database, which in turn demands the careful identification of all foreshocks and aftershocks that could be associated with mainshock events. Hence, the database is usually declustered to remove foreshocks and aftershocks before calibrating the recurrence model. Although there is some debate on the use of declustered recurrence rates (e.g., Marzocchi and Taroni, 2014), it is common practice and some applications even require them in order to be consistent with their underlying assumptions, as is the case for most probabilistic seismic risk assessment frameworks (e.g., Moehle and Deierlein, 2004).

The most commonly used declustering techniques remove events in a space-time window around mainshocks (e.g., Gardner and Knopoff, 1974) or use link-based algorithms (e.g., Reasenber, 1985). More recent models have considered stochastic declustering (e.g., Zhuang et al., 2002; Marsan and Lengliné, 2010), where the classification of earthquakes as mainshocks is done probabilistically. These methods are usually used without validating that the

Table 1: Linear relations used to transform different magnitudes into M_w .

| Magnitude scale | c^* | d^* | s_c^\dagger | s_d^\dagger | n^\ddagger |
|----------------------------------|--------|-------|---------------|---------------|--------------|
| Surface wave magnitude (M_s) | 0.585 | 0.935 | 0.073 | 0.013 | 1115 |
| Local magnitude (M_L) | 0.676 | 0.899 | 0.200 | 0.037 | 891 |
| Body wave magnitude (m_b) | -1.005 | 1.227 | 0.092 | 0.017 | 1716 |

* c and d are the coefficients of the regression.

† s_c and s_d are the standard errors of the estimated coefficients.

‡ n is the number of events with both types of magnitude reported.

resulting catalog of mainshocks resembles a homogeneous Poisson process, which is a key aspect in the construction of a defensible earthquake recurrence model.

After Gardner and Knopoff (1974) reported that the declustered Southern California seismicity was Poissonian, several works have dealt with the problem of evaluating whether declustered catalogs exhibited properties of homogeneous Poisson processes (e.g., Michael, 1997; Luen and Stark, 2012; Beauval et al., 2006, 2008). Other works have compared the adequacy of alternative statistical distributions for describing seismicity, especially in relation to the time between successive earthquakes. In them, the exponential distribution, which derives from a homogeneous Poisson process, is compared with other distributions (Todorovska, 1994; Wang and Kuo, 1998; Utsu, 1984; Chen et al., 2013). In other cases, the Weibull and the gamma distributions, which generalize the exponential distribution, are considered instead (Yakovlev et al., 2006; Zöller and Hainzl, 2007; Touati et al., 2009; Abaimov et al., 2008). Despite these efforts, all these works are concerned with a single statistical aspect of seismicity, either the distribution of times between events or the number of events occurring within a time window.

In the case of Chile, several recurrence models for subduction earthquakes have been previously proposed (e.g., Barrientos, 1980; Martin, 1990; Leyton et al., 2009; Reyes and Cárdenas, 2010; Núñez, 2014). All of these studies, with the exception of Núñez (2014), use surface-wave magnitudes (M_s). However, most GMPEs used with subduction earthquakes use moment magnitudes (e.g., Abrahamson et al., 2016; Zhao et al., 2006), and hence their use requires a magnitude transformation relation that can bias PSHA results. Besides, only in the work of Leyton et al. (2009), the earthquake catalog was declustered by removing foreshocks and aftershocks, which is critical in maintaining mainshock earthquake events distributed in time according to a Poisson process.

This paper proposes a new earthquake recurrence model for subduction interface and intraslab earthquakes in Chile using updated geophysical and seismological data. All model parameters are derived for moment magnitudes, and the earthquake catalog is declustered to consider mainshock events only. The geometry of the seismic sources is defined using geological and geophysical data with seismicity patterns assumed to be homogeneous. The sensitivity of two modeling assumptions in the final seismic parameters are studied: the used relationships for magnitude homogenization and the declustering algorithm. This study also provides a procedure to test the hypothesis if mainshock earthquake occurrences follow a homogeneous Poisson process. Furthermore, the presence of temporal fluctuations that may violate the Poissonian assumption is also checked.

Estimation of Earthquake Recurrence Parameters

Seismic Catalog

Earthquake events were retrieved from two catalogs, the ISC Bulletin and the USGS ComCat catalog, which cover from 1900 to the present. Non-instrumental catalogs, such as the SISRA catalog (Askew and Algermissen, 1985), were discarded in this study due to the inherent high uncertainty of their estimations. The information from the USGS ComCat catalog was preferred for events present in both catalogs. Magnitudes were homogenized to moment magnitude (M_w) using two different approaches: available relations from the literature (i.e., Scordilis, 2006; Núñez, 2014), and linear regressions fitted using events where M_w and at least one other magnitude (M_s , M_L , or m_b) were reported. The transformations considered have the linear form $M_w = c + dM$, with parameters c and d defined in Table 1. The regression parameters were computed using orthogonal least squares regressions since they usually result in better estimations than standard least squares regressions for most magnitudes and measurement errors in practice (Castellaro and Bormann, 2007).

The time window for which earthquakes are observed in a region depends on their magnitude; the higher the magnitude of earthquakes, the longer the observation time windows and vice versa. Such is the case because the increase in instrumentation density over time has improved the ability to sense smaller earthquakes. The years from which the catalog was assumed to be complete were calculated using the methodology proposed by Stepp (1972).

Only events with moment magnitude greater than 5 after homogenization were considered in the analysis, since this is the magnitude threshold for which the seismic catalog is complete since year 1974, and represents a meaningful threshold for engineering purposes.

Declustering

The declustering method used herein was proposed by Gardner and Knopoff (1974) and uses a generalized form of the space-time windows from Knopoff et al. (1982), which enables the independent scaling of the windows for more flexibility. The method was implemented as follows. Let us consider two events i, j of magnitudes m_i, m_j , with $m_i > m_j$, that occurred at locations x_i, x_j on dates t_i, t_j , respectively. Then, if

$$d(x_i, x_j) \leq C_{\text{dist}} \cdot 3^{m_i-7}, \quad (1)$$

and

$$|t_i - t_j| \leq C_{\text{time}} \cdot 3^{m_i-7}, \quad (2)$$

event j is removed from the catalog. In the conditions above, $d(x_i, x_j)$ is the geodesic distance between locations x_i, x_j in kilometers, while $C_{\text{dist}} = 100$ km and $C_{\text{time}} = 130$ days are declustering parameters. The rule is applied in order from the greater to the smaller events (magnitude-wise). Parameters $C_{\text{dist}}, C_{\text{time}}$ were selected so that the times between earthquakes in the declustered catalog follow an exponential distribution and the number of earthquakes in a time window a Poisson distribution. These criteria are consistent with the homogeneous Poisson process assumption as explained later in the section with statistical tests.

Zonation

After declustering the catalog, the remaining events (mainshocks) were partitioned into the seven geographic zones shown in Figure 1, which are composed of three subduction interface zones (Zones 1-3), and four subduction intraslab zones (Zones 4-7). Interface zones extend from the oceanic trench to a depth of 60 km, while intraslab zones extend from a depth of 60 km to approximately 160 km. The geometry of the trench and the depth contours were obtained from the real geometry of the slab proposed by Hayes et al. (2012). While seismotectonic zonation is a fundamental aspect in hazard assessments, the assumed zonation in this work is modified from other models to produce comparable results. Thus, the north-south slab divisions were selected to be the same as the one selected in previous studies (Martin, 1990; Leyton et al., 2009).

It is important to remark that the studied region also features earthquakes that rupture crustal faults in the South American plate, outer rise events, and deep seismicity, as shown in Figure S1, available in the electronic supplement to this article. However, to be consistent with the scope of this study, the aim was to focus separately on the interface and intraslab subduction seismicity, which is much more prominent.

Gutenberg-Richter Relation

The expected annual number of seismic events with M_w exceeding a certain value m , $\lambda_M(m)$, is assumed to follow the well-known Gutenberg-Richter relation (Gutenberg and Richter, 1944):

$$\lambda_M(m) = 10^{a-bm}, \quad (3)$$

where parameters a and b must be calibrated for each zone. This relationship can be rewritten in terms of the expected number of events with magnitude greater than a minimum magnitude (chosen as M_w 5 in this study):

$$\lambda_M(m) = \lambda e^{-\beta(m-M_{\text{min}})}, \quad (4)$$

where $\lambda = \lambda_M(M_{\text{min}})$, and $\beta = b \ln(10)$. The Gutenberg-Richter relation normally considers an upper magnitude bound, M_{max} , since the rupture area of earthquakes and their slip are physically constrained. Further details of the estimation of maximum magnitudes are available in the electronic supplement to this article.

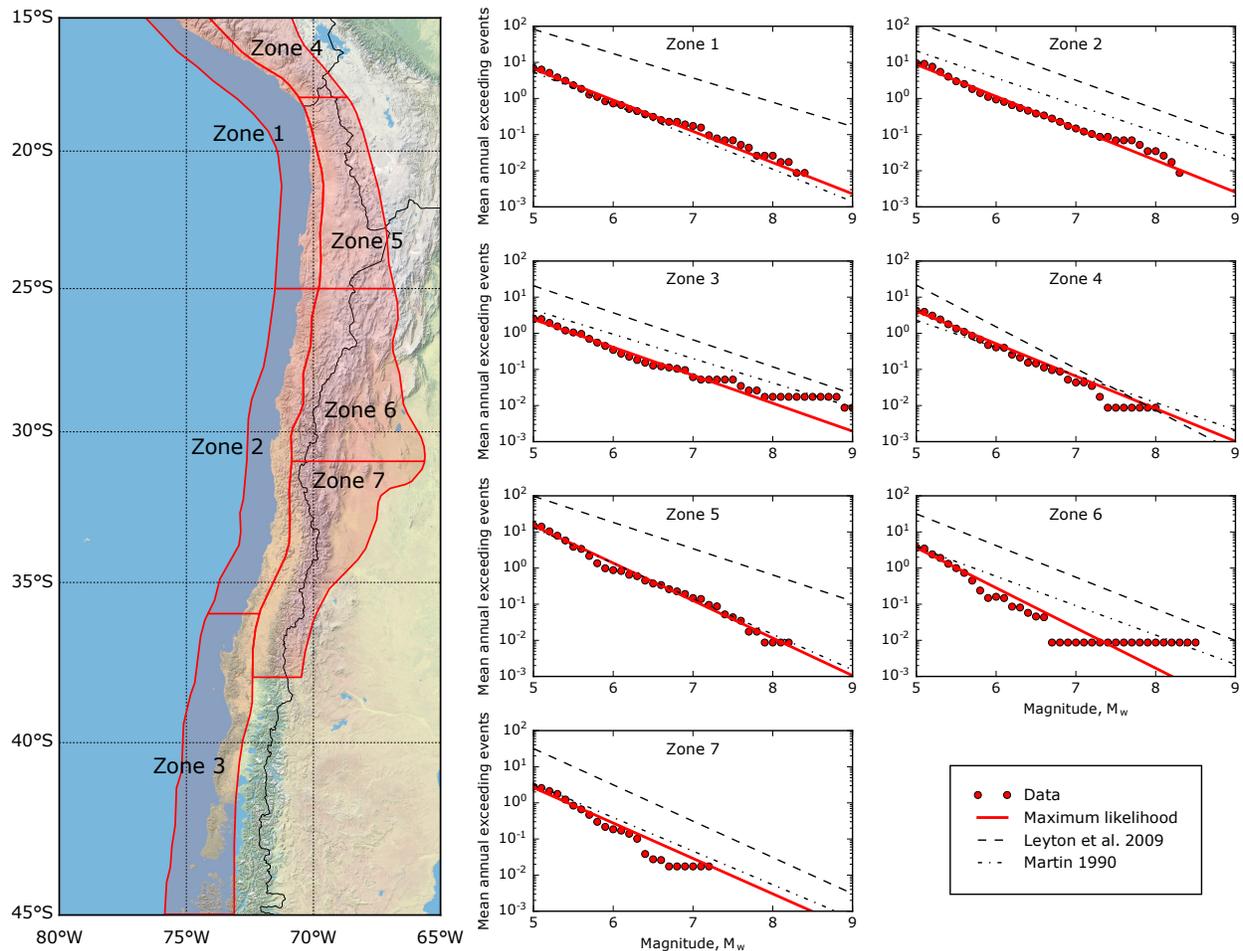


Figure 1: Geometry of the zones in the earthquake recurrence model and their calibrated Gutenberg-Richter relations.

The calibration of the Gutenberg-Richter parameters was performed using the maximum likelihood method proposed by Weichert (1980), which accounts for unequal observation time windows at different magnitudes. The asymptotic standard deviation of the estimators was computed with the expressions derived by Keller et al. (2014), who found that Weichert (1980) overestimated the standard deviation of parameter β .

Table 2 summarizes the values of the resulting parameters for each zone using the homogenization regressions of this study and the declustering method of Knopoff et al. (1982). The table also shows the number of events with $M_w \geq 5$ considered in the catalogs, the estimated maximum magnitudes (M_{max}) used to truncate the Gutenberg-Richter distribution, and the λ parameter normalized by an area of $100,000 \text{ km}^2$ (λ_{norm}). This last parameter shows that Zone 5 has significantly higher seismic rates than the other intraslab zones.

The calibrated Gutenberg-Richter relations are also shown in the plots of the right of Figure 1 and compared with the catalog data. The fit with the data is better for Zones 1, 2, and 5 since there are more events in them than in the other four zones. Figure 1 also shows the Gutenberg-Richter relations of two previous studies that used very similar zones (Martin, 1990; Leyton et al., 2009). Both studies used surface-wave magnitude (M_s) instead of M_w as used herein; therefore the relations have been transformed to M_w using the linear relationship derived presented in Table 1. The figure shows that the differences between the two previous studies are substantial.

The results are also compared with two other studies (Barrientos, 1980; Núñez, 2014) in Figure 2. Since these studies used different zonations, the comparison is performed separately for interface and intraslab seismicity, i.e. Zones 1-3 and Zones 4-7, respectively. Because all of these works roughly study the same region, they should result in frequencies of the same order of magnitude. However, as shown in Figure 2, results from this study are quite

Table 2: Gutenberg-Richter parameters for each seismic zone, and for interface and intraslab zones together.

| Zone | Number of events | a^* | b^* | λ^\dagger | β^\dagger | σ_λ^\ddagger | σ_β^\ddagger | λ_{norm}^\S | M_{max}^\parallel |
|-----------|------------------|-------|-------|-------------------|-----------------|---------------------------|-------------------------|----------------------------|----------------------------|
| Zone 1 | 390 | 5.15 | 0.87 | 6.63 | 1.99 | 0.34 | 0.08 | 2.86 | 9.2 |
| Zone 2 | 508 | 5.36 | 0.88 | 8.71 | 2.03 | 0.39 | 0.07 | 4.07 | 9.3 |
| Zone 3 | 144 | 4.26 | 0.77 | 2.43 | 1.78 | 0.20 | 0.12 | 1.22 | 9.6 |
| Zone 4 | 242 | 5.11 | 0.90 | 4.10 | 2.07 | 0.27 | 0.11 | 4.64 | 8.6 |
| Zone 5 | 843 | 6.37 | 1.04 | 15.04 | 2.39 | 0.53 | 0.07 | 8.15 | 8.4 |
| Zone 6 | 195 | 6.12 | 1.11 | 3.65 | 2.56 | 0.26 | 0.15 | 1.43 | 8.4 |
| Zone 7 | 149 | 5.32 | 0.98 | 2.64 | 2.26 | 0.22 | 0.15 | 1.24 | 8.5 |
| Zones 1-3 | 1042 | 5.55 | 0.86 | 17.78 | 1.98 | 0.56 | 0.05 | 2.75 | – |
| Zones 4-7 | 1429 | 6.48 | 1.02 | 25.43 | 2.34 | 0.69 | 0.05 | 3.43 | – |

* a and b are the Gutenberg-Richter parameters.

† λ and β are alternative Gutenberg-Richter parameters.

‡ σ_λ and σ_β are the asymptotic standard deviations of the alternative Gutenberg-Richter parameters.

§ λ_{norm} is the λ parameter normalized by an area of 100,000 km².

|| M_{max} is the maximum magnitude.

similar to those of Martin (1990) and Núñez (2014). The frequencies reported by Barrientos (1980) are one order of magnitude smaller, which could be the case due to the use of considerable less data. Contrarily, the results of Leyton et al. (2009) are one order of magnitude higher using essentially the same data. It is important to note that most of the studies (Barrientos, 1980; Martin, 1990; Núñez, 2014) did not decluster their catalogs; therefore, we can expect their declustered frequencies to be roughly 22 to 44% lower, based on the number of events extracted in this study by the two declustering algorithms. However, accounting for declustering would make the differences with Barrientos (1980) even more significant and would not affect the differences with Leyton et al. (2009).

Sensitivity to Declustering and Magnitude Homogenization

The analysis was repeated for the complete study region using the clustered data and the declustering methods proposed by Knopoff et al. (1982) and Reasenber (1985). For the latter method, earthquakes with magnitude greater than 3.5 were used to identify clusters, and the parameters were set to the default values (Reasenber, 1985) with the exception of the maximum look-ahead time that was increased to 20 days, which limited testing showed to be more consistent with the seismicity of the region. Furthermore, the analyses considered two alternative sets of magnitude homogenization relationships, the ones proposed by Scordilis (2006) for M_s and m_b , and by Núñez (2014) for M_L ; and the ones derived in this study (Table 1). The results for all combinations of declustering methodologies, including the original clustered data, and magnitude homogenization relationships are shown in Table 3. This table shows the total number of events with $M_w \geq 5$, and the Pearson and Spearman correlation coefficients of successive inter-event times, which in a Poisson stochastic process, should be theoretically zero. The 95% confidence interval of the correlation coefficient associated with a Poisson process is also presented in this table.

Several important observations can be drawn from these global results. First, both declustering algorithms reduce the correlations with respect to the clustered data; however, the reduction obtained by the methodology proposed by Knopoff et al. (1982) is clearly more significant. Second, the selection of a declustering algorithm affects the parameters that represent the total number of events (a and λ) significantly, while the value of parameter b remains almost unchanged. Third, the selection of the relationships for magnitude homogenization has a significant effect on all final parameters, especially on λ (the rate of events with $M_w \geq 5$), which is mostly due to events with no M_w value reported and m_b values between 4.67 and 4.90, the minimum values included when using the relation derived by Scordilis (2006) and this work, respectively. Consequently, we chose to use the Knopoff et al. (1982) declustering algorithm for its significantly lower correlation, and the magnitude homogenization relationships derived by this research since they are specific to the region considered.

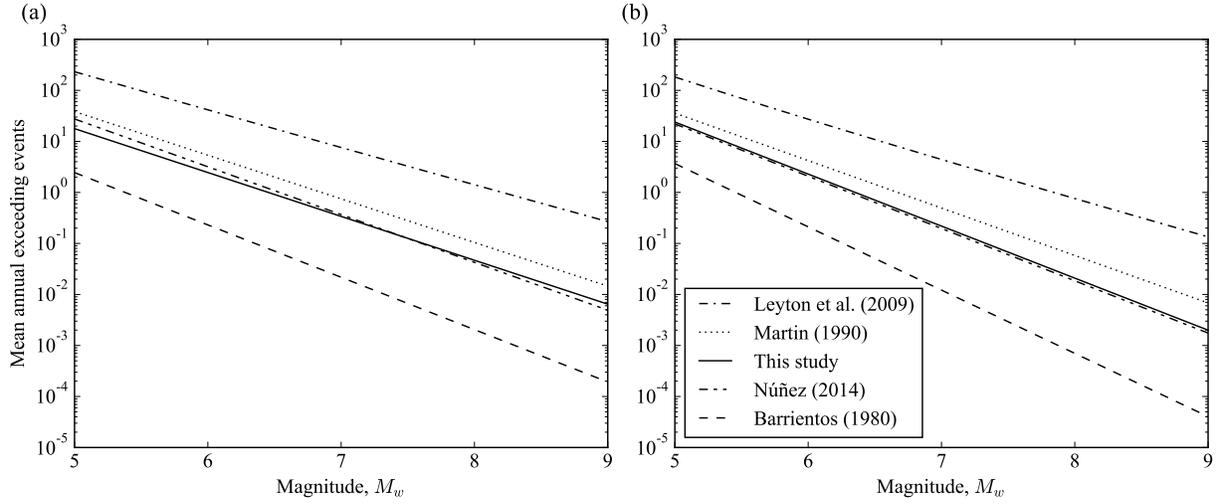


Figure 2: Comparison of recurrence relations for Chilean (a) interface and (b) intraslab seismicity relative to existing previous models.

Table 3: Results for the complete study zone using different assumptions.

| Declustering | None | | Reasenberg (1985) | | Knopoff et al. (1982) | |
|--------------------------|-------------|-------------|-------------------|-------------|-----------------------|-------------|
| | Scordilis | This study | Scordilis | This study | Scordilis | This study |
| a^* | 8.16 | 6.62 | 8.15 | 6.82 | 8.03 | 6.35 |
| b^* | 1.21 | 0.95 | 1.23 | 1.01 | 1.23 | 0.94 |
| λ^\dagger | 124.16 | 77.92 | 95.93 | 59.86 | 73.79 | 43.18 |
| β^\dagger | 2.79 | 2.18 | 2.84 | 2.32 | 2.84 | 2.17 |
| Events with $M_w \geq 5$ | 6757 | 4394 | 5281 | 3441 | 4073 | 2471 |
| Pearson correlation | 0.201 | 0.243 | 0.084 | 0.058 | 0.093 | 0.079 |
| Spearman correlation | 0.465 | 0.517 | 0.105 | 0.107 | 0.059 | 0.039 |
| 95% CI of correlation | ± 0.025 | ± 0.031 | ± 0.028 | ± 0.035 | ± 0.032 | ± 0.041 |

* a and b are the Gutenberg-Richter parameters.

† λ and β are alternative Gutenberg-Richter parameters.

Testing the Poisson Assumption

Test design

Declustered earthquake catalogs are supposed to be well-modeled by a homogeneous Poisson process (Gardner and Knopoff, 1974; Michael, 1997; Luen and Stark, 2012). Therefore, the occurrence relations introduced above are built on this assumption.

In a homogeneous Poisson process, events are generated independently of each other following solely an occurrence rate $\lambda > 0$. For a time window $[t, t+\tau]$, with τ small, an event occurs with probability $\lambda\tau + o(\tau^2)$, regardless of the time passed since the previous event or the history of events to date (memoryless property or Markovianity). Elaborating on this probability, it follows that:

(P1) The times between consecutive events follow an exponential distribution of rate λ .

(P2) The number of events occurring in a time window of size T follows a Poisson distribution of rate λT .

Furthermore, since events occur independently of each other in time, from a time series perspective, it follows that:

(P3) The frequency of events does not consistently increase nor decrease over time, nor exhibits seasonality or other periodic behavior, except for erratic fluctuations due to randomness.

Table 4: Results of the χ^2 goodness-of-fit test for properties (P1) and (P2).

| Zone | (P1) Exponential distribution | | | | (P2) Poisson distribution | | | |
|--------|-------------------------------|---------------------|------------|---------------------|---------------------------|---------------------|------------|-------------------|
| | Bins | χ^2 -statistic | p -value | Assessment | Bins | χ^2 -statistic | p -value | Assessment |
| Zone 1 | 11 | 11.459 | 0.323 | Consistent | 7 | 5.339 | 0.376 | Consistent |
| Zone 2 | 11 | 14.581 | 0.148 | Weakly consistent | 9 | 11.655 | 0.112 | Weakly consistent |
| Zone 3 | 11 | 11.397 | 0.327 | Consistent | 5 | 1.917 | 0.590 | Consistent |
| Zone 4 | 11 | 9.483 | 0.487 | Consistent | 5 | 5.460 | 0.141 | Weakly consistent |
| Zone 5 | 11 | 15.150 | 0.127 | Weakly consistent | 12 | 25.032 | 0.005 | Inconsistent |
| Zone 6 | 11 | 17.767 | 0.059 | Weakly inconsistent | 5 | 6.177 | 0.103 | Weakly consistent |
| Zone 7 | 11 | 18.573 | 0.046 | Inconsistent | 5 | 2.021 | 0.568 | Consistent |

To assess the adequacy of the homogeneous Poisson process assumption, properties (P1), (P2), and (P3) were tested statistically. Properties (P1) and (P2) were tested using the χ^2 goodness-of-fit test as done in the literature (e.g., Hoel, 1945; Luen and Stark, 2012). For (P1), the bins in the χ^2 test were selected to have equal probability with respect to the maximum likelihood estimate of the exponential distribution. For (P2), the bins were naturally defined by the natural numbers, so low probability bins were combined with adjacent bins until all had probability above 0.1, according to the maximum likelihood estimate of the Poisson distribution. The counts of each zone were computed over quarter years and their time series are shown in Figure S2, available in the electronic supplement to this article.

Property (P3) was tested by measuring four correlations and testing if they are different from zero: the correlation between local inter-event times is zero, and their global trend is zero; and the correlation between local counts in the time series is zero, and their global trend is zero. The statistical significance of these correlation was assessed using the t -test as discussed in Artusi et al. (2002). The t -transform is computed as $t = |r|(\frac{1-r^2}{N-2})^{-0.5}$, where r is the correlation tested and N is the number of observations used to calculate r .

Local and global correlations were measured as follows. Local correlations were defined as the correlations of observations that are consecutive in time, i.e., as $r(X_i, X_{i+1})$, where X_1, X_2, \dots, X_m are the inter-event times (or time series counts) ordered chronologically. On the other hand, global correlations were defined over the entire 43-year observation period being tested (from 1974 to 2018) and were calculated as $r(X_i, i)$, i.e., by comparing each observation to its index i . Local and global correlations were computed using both Pearson’s and Spearman’s formulas.

The test results were interpreted according to the following rubric: the hypothesis is labeled “consistent” if $p \in [0.2, 1]$; “weakly consistent” if $p \in [0.1, 0.2]$; “weakly inconsistent” if $p \in [0.05, 0.1]$; and “inconsistent” if $p \in [0, 0.05]$. This labeling scheme is an extension of the traditional yet simplistic p -value criterion of rejecting a hypothesis (i.e., deeming it inconsistent) if $p < 0.1$ or another threshold such as 0.01 or 0.001.

Test results

The statistical test results of properties (P1) and (P2) are summarized in Table 4 and the distributions are shown in Figure 3. As the table and the figure show, the catalog is largely consistent with properties (P1) and (P2). Inconsistencies occur in Zone 5, with the Poisson distribution, and in Zone 7, with the exponential distribution. Figure 3 shows that these inconsistencies are slight, although the distribution of the number of events in Zone 5 reveals a non-Poissonian peak. This corresponds to activity following the 2010 M_w 8.8 Maule earthquake that failed to be declustered.

Table 5 presents the results of property (P3), i.e., the local and global correlations on both the inter-event times and the time series. Since several dependent tests were performed in each zone, each subtable contains the minimum (worst) p -value obtained for each zone. In general, the tests lead to mixed results in terms of the temporal independence of the mainshocks. Despite statistical significance in some cases, correlations seem small in general, and the correlations that are statistically significantly different from zero are concentrated over the short term. Correlations above 0.2 are only present in the short term (local correlations) in the time series counts of Zone 5 and the inter-event times of Zone 6.

From these results, it may be concluded that the declustered catalog is not strictly consistent with a homogeneous Poisson process. However, if we analyze the stochasticity from the perspective of the inter-event times, then the catalog does resemble a homogeneous Poisson process, i.e., the inter-event times distribute according to an exponential distribution, although some local correlations are significant. On the contrary, if the stochasticity is analyzed from the perspective of the time series, it is apparent that the counts follow a Poisson distribution, but data set correlations are significant, especially in the short-term.

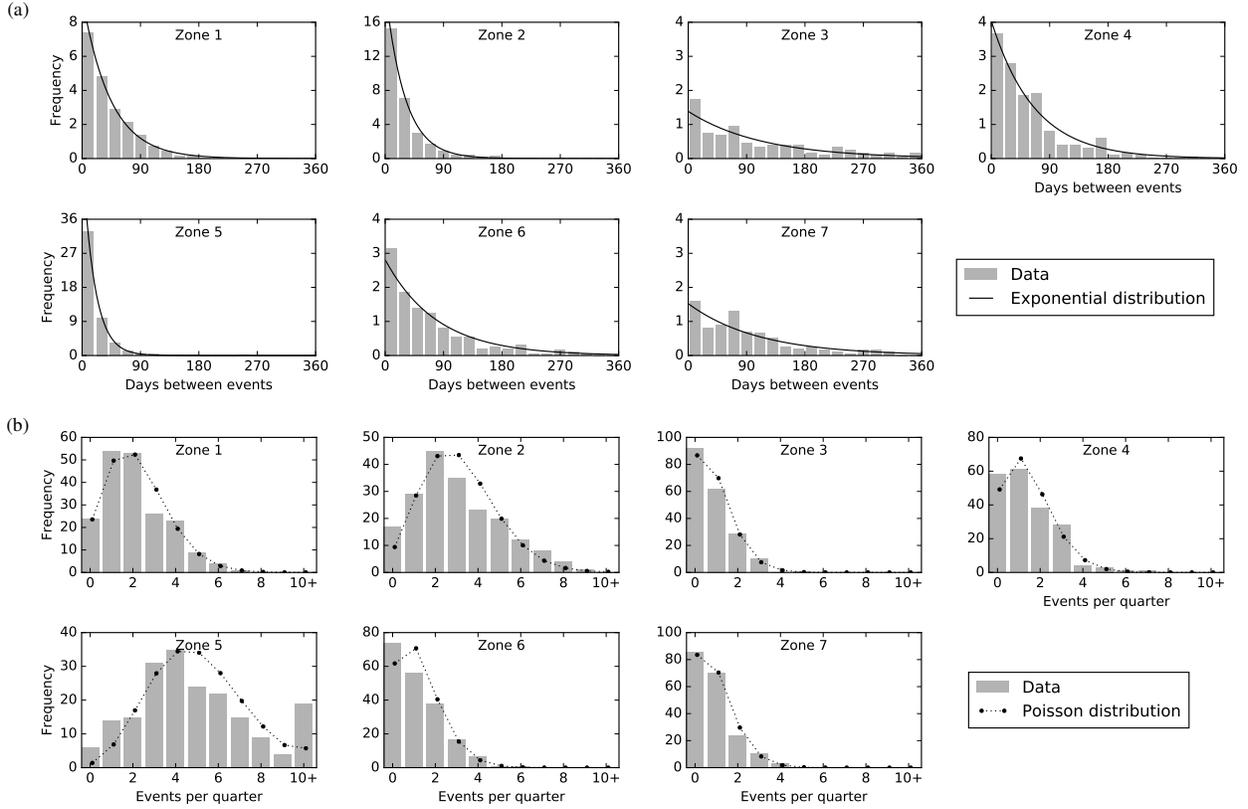


Figure 3: Statistical fit of: (a) the exponential distribution to the times between successive mainshocks (inter-event times), and (b) the Poisson distribution to the number of mainshocks per quarter (counts). Both insets compare the distributions to the data for the seven zones defined.

Table 5: Results of testing property (P3): local and global correlations and their statistical significance for inter-event times and time series counts, of each zone.

| Variable | Zone | r_{local}^* | $p\text{-value}^\dagger$ | s_{local}^* | $p\text{-value}^\dagger$ | r_{global}^* | $p\text{-value}^\dagger$ | s_{global}^* | $p\text{-value}^\dagger$ | min $p\text{-value}^\ddagger$ | Assessment |
|--------------------|------|----------------------|--------------------------|----------------------|--------------------------|-----------------------|--------------------------|-----------------------|--------------------------|-------------------------------|---------------------|
| Inter-event times | 1 | 0.058 | 0.243 | 0.019 | 0.707 | -0.013 | 0.789 | -0.032 | 0.517 | 0.243 | Consistent |
| | 2 | 0.058 | 0.163 | 0.064 | 0.119 | -0.052 | 0.203 | -0.066 | 0.108 | 0.108 | Weakly consistent |
| | 3 | -0.088 | 0.276 | -0.01 | 0.905 | 0.072 | 0.371 | 0.104 | 0.195 | 0.195 | Weakly consistent |
| | 4 | 0.131 | 0.032 | 0.121 | 0.048 | 0.071 | 0.246 | 0.018 | 0.772 | 0.032 | Inconsistent |
| | 5 | 0.102 | 0.002 | 0.051 | 0.117 | -0.048 | 0.133 | -0.097 | 0.003 | 0.002 | Inconsistent |
| | 6 | 0.212 | 0.002 | 0.159 | 0.018 | 0.010 | 0.878 | -0.075 | 0.263 | 0.002 | Inconsistent |
| | 7 | 0.104 | 0.186 | 0.023 | 0.775 | 0.124 | 0.112 | 0.012 | 0.879 | 0.112 | Weakly consistent |
| Time series counts | 1 | 0.065 | 0.371 | 0.017 | 0.814 | 0.009 | 0.900 | -0.007 | 0.921 | 0.371 | Consistent |
| | 2 | 0.179 | 0.014 | 0.190 | 0.009 | 0.090 | 0.214 | 0.070 | 0.330 | 0.009 | Inconsistent |
| | 3 | -0.078 | 0.284 | -0.085 | 0.242 | -0.067 | 0.355 | -0.053 | 0.462 | 0.242 | Consistent |
| | 4 | 0.018 | 0.808 | 0.071 | 0.330 | -0.104 | 0.151 | -0.125 | 0.082 | 0.082 | Weakly inconsistent |
| | 5 | 0.272 | <0.001 | 0.250 | <0.001 | 0.100 | 0.166 | 0.049 | 0.495 | <0.001 | Inconsistent |
| | 6 | 0.076 | 0.296 | 0.082 | 0.263 | -0.008 | 0.913 | -0.064 | 0.372 | 0.263 | Inconsistent |
| | 7 | 0.053 | 0.467 | 0.142 | 0.050 | -0.135 | 0.060 | -0.147 | 0.041 | 0.041 | Inconsistent |

Correlations and p -values labeled inconsistent or weakly inconsistent are in bold.

* Pearson correlations are denoted by r and Spearman correlations are denoted by s . Subscripts local and global denote the time-reach of the correlations.

† The p -value of each correlation is presented to their right.

‡ Minimum p -value of each row.

Conclusions

This work accounted for the interface and intraslab seismicity of the subduction margin that runs north-south in Chile and resulted in an earthquake recurrence model for subduction earthquakes, which improves some limitations of previous available models and uses up-to-date earthquake catalogs. The resulting frequencies from the Gutenberg-Richter relation were similar in some cases to some of the results reported in previous work (i.e., Martin, 1990; Núñez, 2014), with roughly one order of magnitude larger than the one proposed by Barrientos (1980), and one order of magnitude smaller than one reported by Leyton et al. (2009). These differences in earthquake rates would have large impact on PSHA, seismic risk assessment, and structural response and design.

The declustered catalog was also analyzed under the assumptions of a homogeneous Poisson process, finding that its most relevant features are consistent with the data, i.e. the distribution of times between events followed approximately an exponential distribution and a Poisson distribution for events in a time window defined in quarters. Testing for the presence of different types of temporal associations between events, it was found that such associations between the events considered were significant. This is not surprising since the declustering process itself introduces spurious associations in the catalog. For example, quarters in the temporal vicinity of a severe mainshock could have a lower than normal number of events after the declustering process. This effect could create a positive correlation between the number of mainshock events for successive quarter, which should be zero in a Poisson process. Although these kinds of associations may be statistically significant, they have a small effect on practical PSHA, though they could be included in more sophisticated hazard analyses.

In the calibration process, considerable attention was paid to uncertainties, especially to the impact of the choice of the magnitude homogenization relations and declustering method on the resulting Gutenberg-Richter parameters. The bias observed in these choices suggests that these topics require further research, so that their effects are minimized in future hazard assessments in the region. In particular, the two declustering algorithms that were tested showed a significant difference in their similarity to a homogeneous Poisson process, which is a common assumption in seismic hazard and risk assessments. Thus, testing the Poisson nature of the declustered catalog should be a common step in the construction of recurrence models.

The recurrence model proposed in this study has several limitations. First, the basis of the assessment is a declustered catalog, which only contains mainshocks. However, mainshocks are only a part of the relevant seismicity, since foreshocks and aftershocks are also important in the seismic design of structures. Indeed, characterizing the duration of clusters, as well as the frequency and magnitudes of the dependent events (e.g., Boyd, 2012; Iervolino et al., 2014) may have important applications in seismic risk and resilience analysis. Strong mainshocks may lead to aftershocks that cause damage accumulation or collapse in already damaged structures. In addition, considering clustered seismicity rates in PSHA may solve the problem of underestimating the seismic hazard that has been shown in previous studies (e.g., Marzocchi and Taroni, 2014).

Furthermore, this work only considered a single zonation model for the country, with seven zones that are essentially the same as the ones used in previous works (Martin, 1990; Leyton et al., 2009); however, zonation is a critical contributing factor in getting representative recurrence relations, and hence should be the subject of future research. In theory, one could ignore zonation and provide spatially continuous estimates of the Gutenberg-Richter parameters. However, the use of this type of model would be impractical in seismic hazard and risk assessments. This opens another research question: how could a data-driven zonation be generated that is simple (i.e., with few zones) and that leads to representative parameters?

Moreover, since no crustal faulting was considered in the analysis, a complete PSHA is not proposed for the region. Such result would require accounting for the superposition of the subduction and crustal seismicity for the Chilean case. More research on the design and impact of declustering methods might also contribute greatly to the statistical quality of recurrence relations. Consequently, Bayesian statistics and other uncertainty quantification methods may be necessary for eliciting and including the uncertainties into robust hazard estimations.

Data and Resources

The seismic event catalogs used in this study were downloaded from the USGS webpage (<https://earthquake.usgs.gov/earthquakes/search/>, last accessed September 2018) and from the ISC webpage (<http://www.isc.ac.uk/iscbulletin/search/bulletin/>, last accessed September 2018). The focal mechanisms from the Global Centroid Moment Tensor Project database were used as counterpart of the analysis (<http://www.globalcmt.org/CMTsearch.html>, last accessed September 2018).

Acknowledgments

This study was supported by the National Research Center for Integrated Natural Disaster Management (CIGIDEN) CONICYT/FONDAP/15110017, and by grants CONICYT/FONDECYT/1170836, CONICYT/FONDECYT/3170867, and CONICYT/FONDECYT/3170895 from the Chilean National Commission for Scientific and Technological Research (CONICYT).

References

- Abaimov, S. G., D. L. Turcotte, R. Shcherbakov, J. B. Rundle, G. Yakovlev, C. Goltz, and W. I. Newman (2008). Earthquakes: recurrence and interoccurrence times, *Pure and Applied Geophysics* **165**(3), 777–795.
- Abrahamson, N., N. Gregor, and K. Addo (2016). BC Hydro ground motion prediction equations for subduction earthquakes, *Earthquake Spectra* **32**(1), 23–44.
- Artusi, R., P. Verderio, and E. Marubini (2002). Bravais-Pearson and Spearman correlation coefficients: meaning, test of hypothesis and confidence interval, *The International Journal of Biological Markers* **17**(2), 148–151.
- Askew, B. L. and S. T. Algermissen (1985). Catalog of earthquakes for South America: Hypocenter and intensity data, Chile, Project SISRA. Technical Report volume 5, Centro Regional de Sismología para América del Sur, Lima, Peru.
- Barrientos, S. (1980). Regionalización sísmica de Chile. Master's thesis, Universidad de Chile, Santiago, Chile.
- Beauval, C., P.-Y. Bard, S. Hainzl, and P. Gueguen (2008). Can strong-motion observations be used to constrain probabilistic seismic-hazard estimates?, *Bulletin of the Seismological Society of America* **98**(2), 509–520.
- Beauval, C., S. Hainzl, and F. Scherbaum (2006). Probabilistic seismic hazard estimation in low-seismicity regions considering non-Poissonian seismic occurrence, *Geophysical Journal International* **164**(3), 543–550.
- Boyd, O. S. (2012). Including foreshocks and aftershocks in time-independent probabilistic seismic-hazard analyses, *Bulletin of the Seismological Society of America* **102**(3), 909–917.
- Castellaro, S. and P. Bormann (2007). Performance of different regression procedures on the magnitude conversion problem, *Bulletin of the Seismological Society of America* **97**(4), 1167–1175.
- Chen, C.-H., J.-P. Wang, Y.-M. Wu, C.-H. Chan, and C.-H. Chang (2013). A study of earthquake inter-occurrence times distribution models in Taiwan, *Natural Hazards* **69**(3), 1335–1350.
- Cornell, C. A. (1968). Engineering seismic risk analysis, *Bulletin of the Seismological Society of America* **58**(5), 1583–1606.
- Gardner, J. K. and L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, *Bulletin of the Seismological Society of America* **64**(5), 1363–1367.
- Gutenberg, B. and C. F. Richter (1944). Frequency of earthquakes in California, *Bulletin of the Seismological Society of America* **34**(4), 185–188.
- Hayes, G. P., D. J. Wald, and R. L. Johnson (2012). Slab1. 0: A three-dimensional model of global subduction zone geometries, *Journal of Geophysical Research: Solid Earth* **117**(B1).
- Hoel, P. G. (1945). Testing the homogeneity of Poisson frequencies, *The Annals of Mathematical Statistics* **16**(4), 362–368.
- Idini, B., F. Rojas, S. Ruiz, and C. Pastén (2017). Ground motion prediction equations for the Chilean subduction zone, *Bulletin of Earthquake Engineering* **15**(5), 1853–1880.
- Iervolino, I., M. Giorgio, and B. Polidoro (2014). Sequence-based probabilistic seismic hazard analysis, *Bulletin of the Seismological Society of America* **104**(2), 1006–1012.

- Keller, M., A. Pasanisi, M. Marcilhac, T. Yalamas, R. Secanell, and G. Senfaute (2014). A Bayesian methodology applied to the estimation of earthquake recurrence parameters for seismic hazard assessment, *Quality and Reliability Engineering International* **30**(7), 921–933.
- Knopoff, L., Y. Y. Kagan, and R. Knopoff (1982). b Values for foreshocks and aftershocks in real and simulated earthquake sequences, *Bulletin of the Seismological Society of America* **72**(5), 1663–1676.
- Leyton, F., S. Ruiz, and S. A. Sepúlveda (2009). Preliminary re-evaluation of probabilistic seismic hazard assessment in Chile: From Arica to Taitao Peninsula, *Advances in Geosciences* **22**, 147–153.
- Luen, B. and P. B. Stark (2012). Poisson tests of declustered catalogues, *Geophysical Journal International* **189**(1), 691–700.
- Marsan, D. and O. Lengliné (2010). A new estimation of the decay of aftershock density with distance to the mainshock, *Journal of Geophysical Research: Solid Earth* **115**(9), 1–16.
- Martin, A. (1990). Hacia una nueva regionalización y cálculo del peligro sísmico en Chile. Master's thesis, University of Chile, Santiago, Chile.
- Marzocchi, W. and M. Taroni (2014). Some thoughts on declustering in probabilistic seismic-hazard analysis, *Bulletin of the Seismological Society of America* **104**(4), 1838–1845.
- McGuire, R. K. (2004). Seismic hazard and risk analysis. EERI Monograph MNO-10, Earthquake Engineering Research Institute.
- Michael, A. J. (1997). Testing prediction methods: Earthquake clustering versus the Poisson model, *Geophysical Research Letters* **24**(15), 1891–1894.
- Moehle, J. and G. G. Deierlein (2004). A framework methodology for performance-based earthquake engineering. In *13th World Conference on Earthquake Engineering*, Vancouver, Canada.
- Montalva, G. A., N. Bastías, and A. Rodríguez-Marek (2017). Ground-Motion Prediction Equation for the Chilean Subduction Zone, *Bulletin of the Seismological Society of America*. Advance online publication.
- Núñez, I. (2014). Nuevo peligro sísmico para Chile. Master's thesis, Universidad de Chile, Santiago, Chile.
- Reasenber, P. (1985). Second-order moment of central California seismicity, 1969–1982, *Journal of Geophysical Research: Solid Earth* **90**(B7), 5479–5495.
- Reyes, J. and V. H. Cárdenas (2010). A Chilean seismic regionalization through a Kohonen neural network, *Neural Computing and Applications* **19**(7), 1081–1087.
- Scordilis, E. M. (2006). Empirical global relations converting M_S and m_b to moment magnitude, *Journal of Seismology* **10**(2), 225–236.
- Stepp, J. C. (1972). Analysis of completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard. In *Proc. of the 1st Int. Conf. on Microzonation, Seattle*, Volume 2, pp. 897–910.
- Todorovska, M. I. (1994). Comparison of response spectrum amplitudes from earthquakes with a lognormally and exponentially distributed return period, *Soil Dynamics and Earthquake Engineering* **13**(2), 97–116.
- Touati, S., M. Naylor, and I. G. Main (2009). Origin and nonuniversality of the earthquake interevent time distribution, *Physical Review Letters* **102**(16), 168501.
- Utsu, T. (1984). Estimation of parameters for recurrence models of earthquakes, *Bulletin of the Earthquake Research Institute* **59**, 53–66.
- Wang, J.-H. and C.-H. Kuo (1998). On the frequency distribution of interoccurrence times of earthquakes, *Journal of Seismology* **2**(4), 351–358.

- Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, *Bulletin of the Seismological Society of America* **70**(4), 1337–1346.
- Yakovlev, G., D. L. Turcotte, J. B. Rundle, and P. B. Rundle (2006). Simulation-based distributions of earthquake recurrence times on the San Andreas fault system, *Bulletin of the Seismological Society of America* **96**(6), 1995–2007.
- Zhao, J. X., J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H. K. Thio, P. G. Somerville, Y. Fukushima, and Y. Fukushima (2006). Attenuation relations of strong ground motion in Japan using site classification based on predominant period, *Bulletin of the Seismological Society of America* **96**(3), 898–913.
- Zhuang, J., Y. Ogata, and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrences, *Journal of the American Statistical Association* **97**(458), 369–380.
- Zöller, G. and S. Hainzl (2007). Recurrence time distributions of large earthquakes in a stochastic model for coupled fault systems: the role of fault interaction, *Bulletin of the Seismological Society of America* **97**(5), 1679–1687.

National Research Center for Integrated Natural Disaster Management (CIGIDEN), CONICYT/FONDAP/15110017
Pontificia Universidad Católica de Chile
Av. Vicuña Mackenna 4860
Macul, Santiago 7820436
Chile
alan.poulos@cigiden.cl
mauricio.monsalve@cigiden.cl
natalia.zamora@cigiden.cl
jcllera@ing.puc.cl