16TH EUROPEAN CONFERENCE ON
EARTHQUAKE ENGINEERING THESSALONIKI
18 - 21 JUNE 2018

STATISTICAL ASSUMPTIONS OF MAINSHOCK SEQUENCES AND THEIR VALIDITY UNDER DIFFERENT MAGNITUDE RANGES

Alan POULOS[1], Mauricio MONSALVE[2], Natalia ZAMORA[3], Juan Carlos DE LA LLERA[4]

ABSTRACT

Probabilistic seismic hazard and risk assessment normally rely on the assumption that mainshock earthquakes of a given seismic source follow a homogenous random Poisson process. This work aims at evaluating the validity of this assumption in the subduction margin of Chile by identifying and testing necessary statistical consequences from these assumptions against the catalog of Chilean mainshocks. These consequences manifest through specific distributions, i.e., exponential distributions for inter-event times, and Poisson distributions for time series counts. The validity of the assumption was statistically evaluated using chi-squared tests for the goodness-of-fit. The tests were performed using minimum magnitudes to filter out the seismic catalog ranging from magnitudes 5.0 to 6.0. The results indicated that the statistical assumption characterized earthquakes of higher magnitudes with greater accuracy, especially those above M_w 5.25. However, as the catalog was restricted to earthquakes of greater magnitudes, the data supplied to the statistical models became scarce, degrading the quality of the statistics. On the other hand, while considering earthquakes of smaller magnitudes may greatly increase the number of earthquakes considered in these statistics, this comes at the expense of considering data of lower quality, since lower magnitude earthquakes are harder to measure instrumentally.

Keywords: Seismic hazard; Statistical seismology; Poisson process; Statistical hypothesis testing.

1. INTRODUCTION

Earthquake recurrence models are a primary input for probabilistic seismic hazard analysis (PSHA), which has become a standard procedure for the selection of seismic design levels for a wide range of engineered infrastructure (McGuire 2008). These models are calibrated using historical catalogs of earthquake occurrences from a specific region. The catalogs have two types of events: independent mainshock earthquakes and dependent earthquakes (i.e., foreshocks and aftershocks). Since PSHA conventionally considers only the former type of earthquakes, catalogs are first declustered to extract independent seismicity using available algorithms (e.g., Gardner and Knopoff 1974; Reasenberg 1985).

Seismic hazard and risk assessments usually rely on the statistical assumption that mainshock earthquake occurrences are described by a temporal homogeneous random Poisson process. In this process, the times between events follow an exponential distribution, while time series counts, i.e., the number of events that occur within a time window, follow a Poisson distribution. Several studies have used these properties to test whether seismic catalogs are consistent with a homogeneous random

[1]National Research Center for Integrated Natural Disaster Management (CIGIDEN) CONICYT/FONDAP/15110017, Santiago, Chile, alan.poulos@cigiden.cl
[2]National Research Center for Integrated Natural Disaster Management (CIGIDEN) CONICYT/FONDAP/15110017, Santiago, Chile, mauricio.monsalve@cigiden.cl
[3]National Research Center for Integrated Natural Disaster Management (CIGIDEN) CONICYT/FONDAP/15110017, Santiago, Chile, natalia.zamora@cigiden.cl
[4]Department of Structural and Geotechnical Engineering, Pontificia Universidad Católica de Chile and National Research Center for Integrated Natural Disaster Management (CIGIDEN) CONICYT/FONDAP/15110017, Santiago, Chile, jcllera@ing.puc.cl

Poisson process (e.g., Gardner and Knopoff 1974; Michael 1997; Luen and Stark 2012). Other works compared the adequacy of alternative statistical distributions for describing seismicity, especially in relation to the time between successive earthquakes (e.g., Wang and Kuo 1998; Chen et al. 2013).

This work seeks to evaluate the validity of the Poissonian assumption under different magnitude intervals, especially with respect to the minimum magnitude chosen. The geographical region where this study is carried out corresponds to the subduction margin along the Pacific coast line of Chile. A catalog of mainshocks events was built in the region by integrating information from several different sources. An earthquake recurrence model was developed using the declustered catalog, and the tests were used to verify if the distribution of inter-event times and the distribution of the number of events occurring in a time window are Poissonian.

2. DEVELOPMENT OF THE RECURRENCE MODEL

The area of study is the Chilean subduction margin between the Nazca and South American plates, which converge at a rate of about 68 mm/year. Figure 1 shows a cross section of the subduction zone with the different types of earthquakes that affect the country. Interface events occur in the boundary between the oceanic (Nazca) plate and the continental (South American) plate, and can reach very large magnitudes, as evidenced by the 1960 M_w 9.5 Valdivia earthquake and the more recent 2010 M_w 8.8 Maule earthquake. Intraslab events occur in the oceanic plate at depths that are usually greater than 60 km. These earthquakes can generate significant damage and casualties, as was the case for the 1939 M_w 7.8 Chillan earthquake, the deadliest earthquake in Chilean history. Other earthquakes occur in geological faults located on the continental plate. These Crustal earthquakes are less numerous than interface and intraslab events, and affect smaller geographical regions. Finally, outer-rise events occur in the oceanic plate to the west of the oceanic trench with little effect on the continent, but with the potential to generate tsunamis.

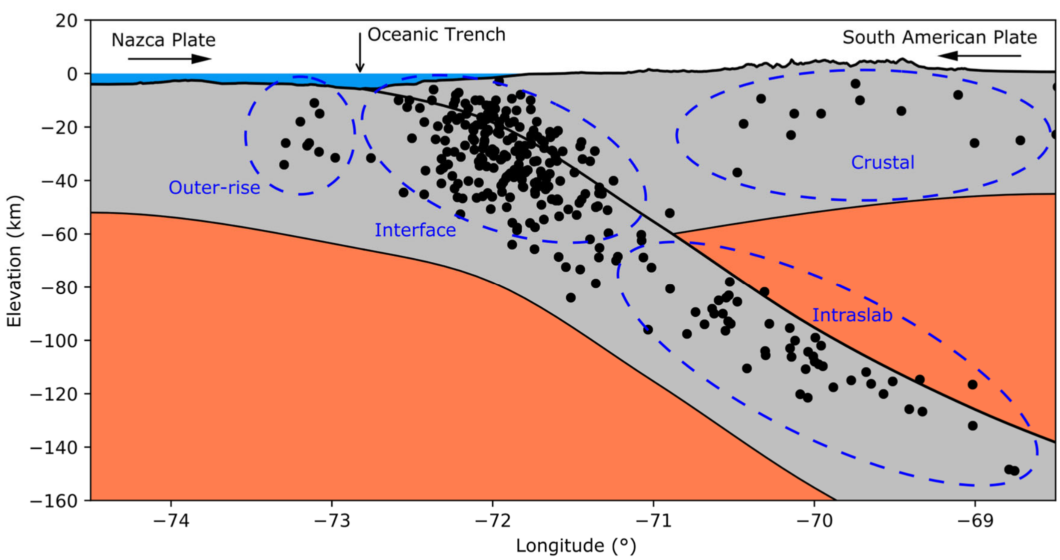


Figure 1. Cross section of the Chilean subduction zone at latitude -33° showing the different types of earthquakes. Black dots represent the hypocenters of earthquakes from the USGS ComCat catalog that occurred between latitude -34° and -32°.

This study only considers interface and intraslab events since they are by far the most numerous, and hence, enable better statistical testing. Moreover, they are the most relevant for engineering purposes. Figure 2a shows the geographical areas (zones) where these earthquakes occur, which were obtained using the slab geometry proposed by Hayes et al. (2012). The interface zone extends from the oceanic trench to the contour line of the slab corresponding to a depth of 60 km, while the intraslab zone ranges from the contour lines of depths 60 and 160 km.

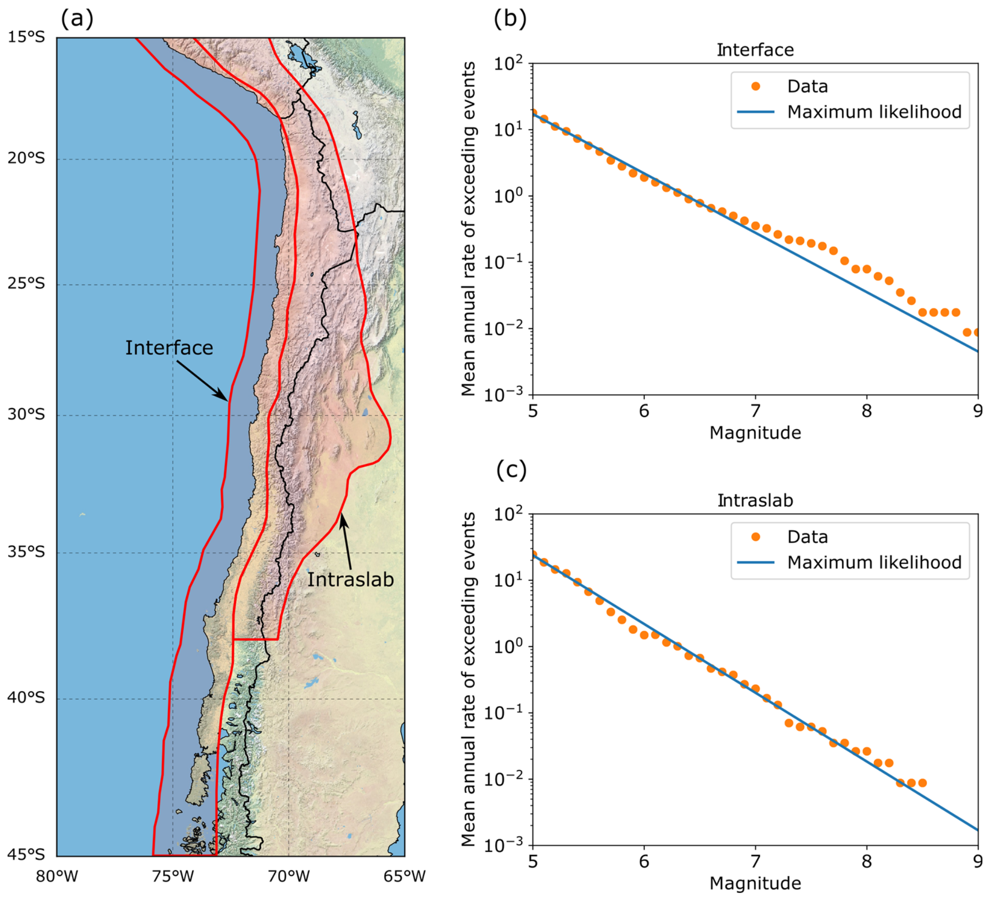


Figure 2. (a) Geometry the interplate and intraslab subduction zone of Chile. Calibrated Gutenberg-Richter laws for: (b) interplate and (c) intraslab earthquakes.

The input earthquake catalog was constructed by merging two publicly available catalogs, the USGS ComCat catalog (USGS 2017) and the ISC Bulletin (International Seismological Centre 2017). The events cover from 1902 to the end of 2016. Duplicate events were identified when two conditions were met simultaneously, the difference between event times was less than 16 s and the difference in epicentral distances was less than 100 km. The information from the ComCat catalog was used in these cases. The moment magnitude scale (M_w) was used in this study since it is widely used in ground motion prediction models, which are also required for PSHA. When M_w for an event was unavailable, it was estimated from a linear regression that relates M_w with another magnitude scale (M_s, M_L, or M_b), also derived from the same data, but using events were M_w and another magnitude scale was reported. Events with magnitude less than 5 after homogenization were removed from the analysis since this is the magnitude threshold for which the seismic catalog is complete since the year 1974, and also represents a meaningful threshold for most engineering purposes.

Dependent seismicity (foreshocks and aftershocks) were removed from the catalog using the declustering algorithm proposed by Knopoff et al. (1982), which removes events in a space-time window around mainshocks. The remaining events after declustering were classified as interface and intraslab using these zones. However, events originating from crustal faults were removed from the analysis. These events were identified using the focal mechanism from the global CMT catalog (Ekström et al. 2012) when available; and if not, they were identified by occurring in the upper part of the slab and at a depth shallower than 40 km.

The mean annual rate of earthquakes that surpasses a given magnitude m in a certain geographical region is described by a Gutenberg-Richter (G-R) law (Gutenberg and Richter, 1944):

$$\lambda_M(m) = 10^{a-bm} \quad (1)$$

where a and b are calibration parameters. These G-R parameters were computed using the methodology proposed by Weichert (1980), which uses a maximum likelihood estimation that considers unequal observation time windows for different magnitude ranges. These differences arise from the fact that the instrumentation density has improved the ability to sense lower magnitude earthquakes over time. The observation periods were estimated with the methodology proposed by Stepp (1972). The resulting recurrence parameters are shown in Table 1, and the fit between the calibrated G-R laws and the catalog is shown in Figures 2b and 2c for interplate and intraslab seismic events, respectively.

Table 1. Gutenberg-Richter parameters calibrated for the interface and intraslab zones.

Zone	Number of events	a	b
Subduction interface	1030	5.71	0.90
Subduction intraslab	1398	6.56	1.04

3. STATISTICAL TESTS

The assumption that mainshock sequences follow a random Poisson process was tested statistically. The adequacy of this process was assessed by plotting and testing the temporal distribution of the events, and filtering by different minimum M_w magnitudes. As stated before, in a homogeneous Poisson process, the times between events theoretically follow an exponential distribution, while time series counts follow a Poisson distribution. In addition, the events should occur independently from each other, which means that the data cannot exhibit seasonality, trends, or any other temporal associations. The latter was not tested in this work.

Visual assessments of the distributions of the inter-event times and time series counts were carried out by plotting their probability-probability plots (pp-plots) and histograms. In a pp-plot, the empirical cumulative probability of the data points is compared against their theoretical cumulative probability, which was estimated using maximum likelihood. When the empirical and theoretical probabilities match, the dotted points in the pp-plot form a diagonal line.

The distributions of the subduction interface and intraslab earthquakes are depicted in Figure 3 and Figure 4, respectively. From left to right, the Figures shows the pp-plot for the inter-event times, the histogram of the inter-event times, the pp-plot for the time series counts, and the histogram of the time series counts. From top to bottom, the minimum magnitudes considered range from 5.0 to 6.0. The histogram plots were designed to best portray the probability distributions. For the exponential distribution (inter-event times), the histogram window spanned from zero to three times the theoretical average inter-event time ($1/\lambda$), in which the observed inter-event times were discretized into ten bins of equal width. The values of $1/\lambda$ are provided in the plots in days. For the Poisson distribution (time series counts), the histograms were centered on the average rate of the series (λt), where the interval window *t* was selected so that the most populated bin considered, at least, 8 observations.

The pp-plots show that the exponential distribution closely resembles inter-event times, something which is confirmed by the histograms. The counts, however, do not match well a Poisson distribution if the minimum magnitude is 5.0. Indeed, the corresponding pp-plot shows a sinusoidal line. This is corrected by considering a minimum magnitude of 5.25. However, as the minimum magnitude increases, the adequacy of the Poisson distribution becomes difficult to assess, due to the decrease in the number of data points. For minimum magnitudes above M_W 6.0, it was even difficult to visualize.

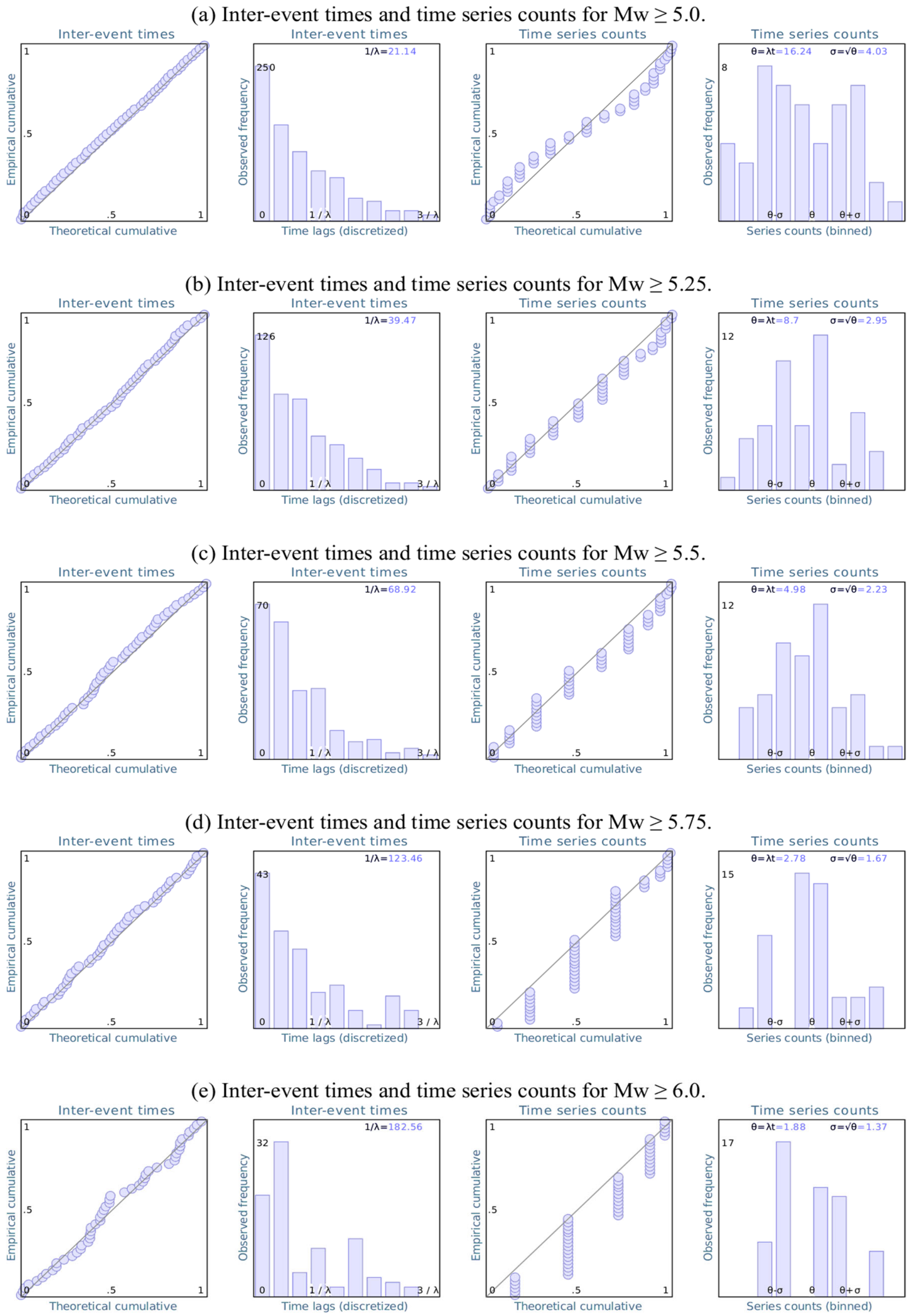
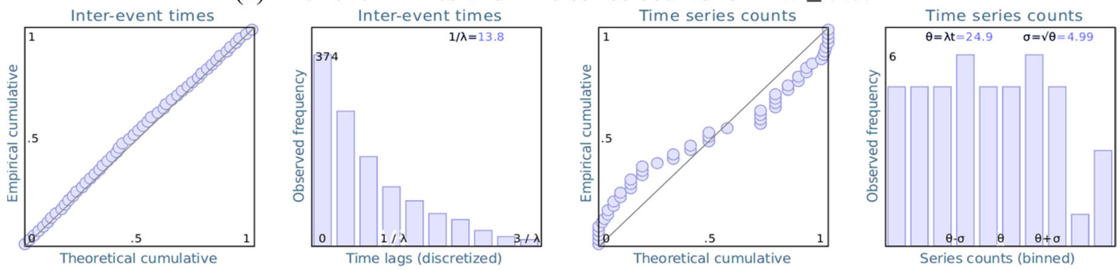
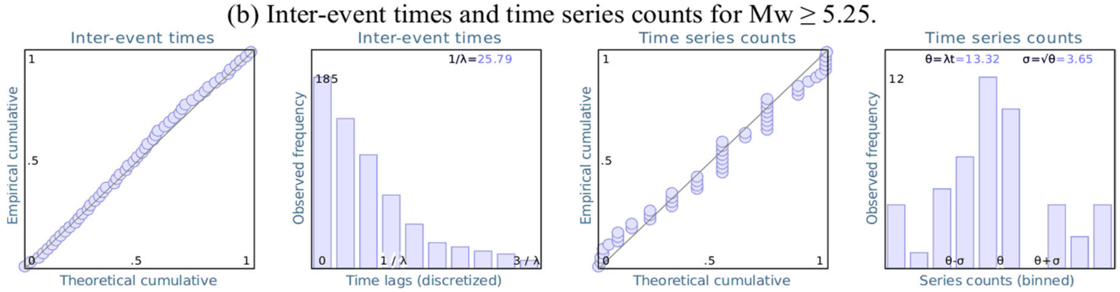


Figure 3. Visualization of the distributions of inter-event times and time series counts for the subduction interface zone, for minimum magnitudes M_w (a) 5.0, (b) 5.25, (c) 5.5, (d) 5.75, and (e) 6.0.

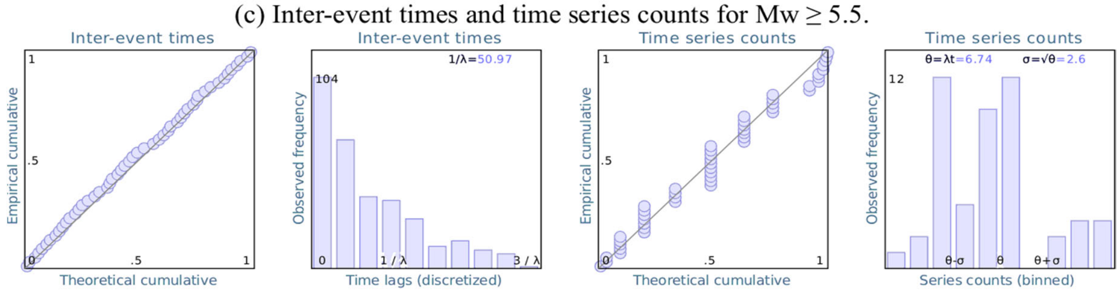
(a) Inter-event times and time series counts for Mw ≥ 5.0.



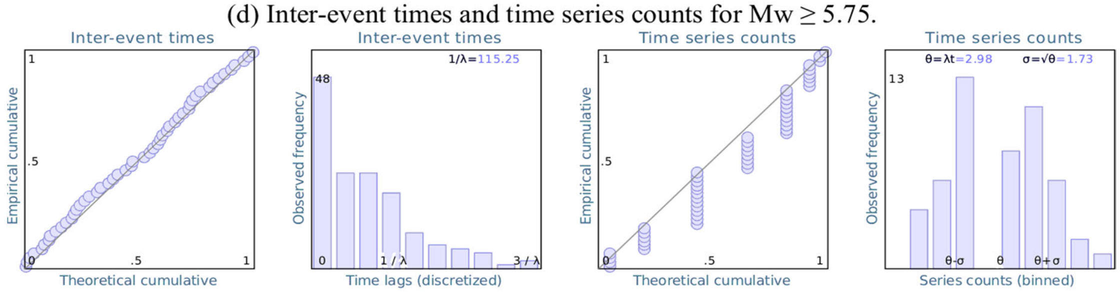
(b) Inter-event times and time series counts for Mw ≥ 5.25.



(c) Inter-event times and time series counts for Mw ≥ 5.5.



(d) Inter-event times and time series counts for Mw ≥ 5.75.



(e) Inter-event times and time series counts for Mw ≥ 6.0.

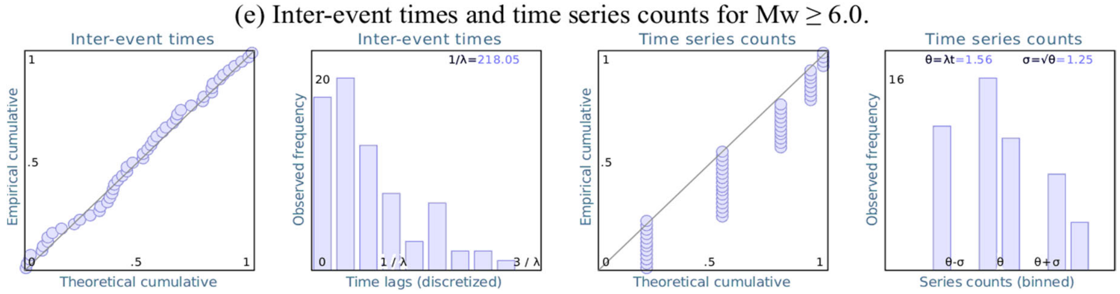


Figure 4. Visualization of the distributions of inter-event times and time series counts for the subduction intraslab zone, for minimum magnitudes Mw (a) 5.0, (b) 5.25, (c) 5.5, (d) 5.75, and (e) 6.0.

While visual assessment of distributions is useful, it must be accompanied by proper tests of the goodness-of-fit. Table 2 shows the p-values of this goodness-of-fit test for the exponential and Poisson distributions, using both interface and intraslab seismicity. The chi-squared test was used to compare the empirical and theoretical distributions. For the exponential distribution, the bins were constructed following the theoretical cumulative distribution, with upper bounds 0.1, 0.2, …, and 1.0. For the Poisson distribution, the events were binned in quarters, with theoretically low-probability (<0.1) adjacent bins being merged together to make the test more robust to tail instabilities.

Table 2. p-values of the goodness-of-fit tests when filtering the earthquakes by different minimum magnitudes.

	Interface		**Intraslab**	
Minimum M_w	**Exponential**	**Poisson**	**Exponential**	**Poisson**
5.00	0.521	0.110	0.808	0.005
5.25	0.778	0.741	0.725	0.481
5.50	0.467	0.425	0.427	0.488
5.75	0.278	0.218	0.627	0.990
6.00	0.329	0.689	0.815	0.876

Generally speaking, the results presented in Table 2 are consistent with the distributions shown in Figures 3 and 4. The Poisson distribution for time series counts is rejected for the intraslab seismicity when the minimum magnitude is 5.0. The fit improves as the minimum magnitude increases. However, while the distributions appear more Poissonian for high minimum magnitudes, fewer observations are available for statistical testing, and hence the results are less conclusive. The exponential distributions for inter-event times were not rejected, regardless of the minimum magnitude.

4. CONCLUSIONS

This article tests the Poissonian nature of the mainshock earthquakes in the Chilean subduction margin. The mainshock catalog was homogenized to moment magnitude, declustered to remove dependent seismicity, and classified into interface and intraslab events along the subduction margin. Gutenberg-Richter parameters were fitted to both zones considering unequal observation time windows for different magnitude ranges.

The hypothesis that the occurrence of mainshock interface and intraslab earthquakes follow a Poisson process was tested statistically for different minimum magnitudes from 5.0 to 6.0. In general, inter-event times closely follow an exponential distribution as it should theoretically for a Poisson process. However, distributions get distorted as the minimum earthquake magnitude for the analysis increases due to less data availability. Counts are also consistent with a Poisson distribution for minimum earthquake magnitudes in the range of 5.25 and 6.0, and less so for a minimum magnitude of 5.0. No significant changes in the results are observed between interface and intraslab events. In conclusion, results of this study suggest that the hypothesis of a homogeneous Poisson random process is consistent for mainshock earthquake sequences of minimum magnitudes greater than 5.25.

There are several specific modeling assumptions that limit the conclusions of this work. First, other earthquake declustering algorithms are also used in the literature, such as the link-based algorithm proposed by Reasenberg (1985). Previous studies have shown that these methods have an impact on the results of statistical tests (e.g., Luen and Stark 2012). Furthermore, only one interface zone and one intraslab zone were used for the tests; however, these zones are usually partitioned into smaller zones with similar seismicity rates for use in PSHA, which could also change the test results presented.

Finally, the statistical properties of Chilean subduction earthquakes might not be representative of other geographical regions or seismological regimes, and hence specific studies should be carried out to test the local seismicities

5. ACKNOWLEDGMENTS

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

6. REFERENCES

Chen CH, Wang JP, Wu YM, Chan CH, Chang CH (2013). A study of earthquake inter-occurrence times distribution models in Taiwan. *Natural hazards*, 69(3): 1335-1350.

Ekström Nettles GM, Dziewoński A (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200: 1-9.

Gardner JK, Knopoff L (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, Richter CF (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4): 185-188.

Hayes GP, Wald DJ, Johnson RL (2012). Slab1.0: A three-dimensional model of global subduction zone geometries, J. Geophys. *Journal of Geophysical Research: Solid Earth*, 117(B1).

International Seismological Centre (2017). *On-line Bulletin*, http://www.isc.ac.uk/iscbulletin/search/bulletin/, Internatl. Seismol. Cent., Thatcham, United Kingdom. Last accessed 2017-01-20.

Knopoff L, Kagan YY, Knopoff R (1982). b Values for foreshocks and aftershocks in real and simulated earthquake sequences. *Bulletin of the Seismological Society of America*, 72(5): 1663-1676.

Luen B, Stark PB (2012). Poisson tests of declustered catalogues. *Geophysical journal international*, 189(1): 691-700.

McGuire RK (2008). Probabilistic seismic hazard analysis: Early history. *Earthquake Engineering & Structural Dynamics*, 37(3): 329-338.

Michael AJ (1997). Testing prediction methods: earthquake clustering versus the Poisson model. *Geophysical research letters*, 24(15): 1891-1894.

Reasenberg P (1985). Second-order moment of central California seismicity, 1969-1982. *Journal of Geophysical Research: Solid Earth*, 90(B7): 5479-5495.

Stepp JC (1972). Analysis of completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard, *Proceedings of the 1st International Conference on Microzonation*, Vol. 2, 897-910. Seattle, WA.

USGS (2017). *USGS ComCat catalog*, https://earthquake.usgs.gov/earthquakes/search/. United States Geological Survey. Last accessed 2017-01-20.

Wang JH, Kuo CH (1998). On the frequency distribution of interoccurrence times of earthquakes. *Journal of Seismology*, 2(4): 351-358.

Weichert DH (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.