

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2024; 9(3): 165-168
© 2024 Stats & Maths
<https://www.mathsjournal.com>
Received: 15-04-2024
Accepted: 17-05-2024

Juma Mwakasege
Department of Pure
Mathematics, Dar es Salaam
College of Applied Mathematics,
Dar es Salaam, Dar es Salaam,
Tanzania

Generalizing survival analysis with competing risks with applications to Indonesian earthquake recurrence

Juma Mwakasege

DOI: <https://www.doi.org/10.22271/maths.2024.v9.i3c.2361>

Abstract

Earthquake recurrence intervals are typically modeled with renewal processes, but these ignore the reality that seismic zones may experience different competing event types mainshocks, triggered aftershock sequences, and aseismic slow-slip events that jointly determine the hazard function. This research generalizes competing-risks survival analysis to earthquake recurrence data from four Indonesian subduction segments along the Sunda Arc (2000–2022). A cause-specific hazard model and a Fine-Gray subdistribution model were compared against a standard Weibull renewal model. The dataset comprised 3,814 events of magnitude 4.0 or greater from the USGS earthquake catalog. The Fine-Gray model achieved the best time-dependent AUC (0.87 at 30 days, 0.81 at 365 days) for predicting mainshock recurrence when aftershock and slow-slip events were treated as competing risks. Segment-specific baseline hazards varied by a factor of 2.3 between the most active (Sumatra) and least active (Java) zones. The Weibull model overestimated 1-year mainshock probability by 14.2 percent because it did not account for events “lost” to competing risks. These results support the adoption of competing-risks frameworks for seismic hazard assessment.

Keywords: Survival analysis, competing risks, earthquake recurrence, Fine-Gray model, cause-specific hazard, Indonesian subduction, Sunda Arc, seismic hazard, Weibull renewal, subdistribution model

Introduction

An earthquake catalog is, at its core, a list of event times a dataset tailor-made for survival analysis. The time between consecutive events of a given type follows a distribution whose hazard function encodes the instantaneous risk of the next event ^[1]. Classical seismic hazard models use exponential or Weibull distributions to describe recurrence, implicitly treating every inter-event interval as censored only by the end of the observation window ^[2]. But seismic zones do not produce just one type of event. Mainshocks, aftershock clusters, and aseismic slow-slip transients all contribute to the event catalog, and each type resets or modifies the hazard in different ways ^[3].

Competing-risks survival analysis, developed in biostatistics for situations where a patient can die of cancer or heart disease ^[4], provides a natural extension. In the seismic context, a mainshock “event” competes with an aftershock or slow-slip “event” for the next observation. Ignoring the competing risks leads to biased estimates of the cause-specific cumulative incidence function ^[5].

Indonesia’s Sunda Arc, stretching from northern Sumatra to eastern Java, is one of the most seismically active regions on Earth, producing the 2004 M9.1 Sumatra-Andaman earthquake among other devastating events ^[6]. This research applies both cause-specific and Fine-Gray subdistribution competing-risks models to four Sunda Arc segments, comparing their recurrence predictions against the standard Weibull renewal approach ^[7].

Model Description and Parameterisation

Three competing-risks models were specified. The cause-specific hazard model ^[4] estimated a separate hazard function $\lambda_k(t)$ for each event type k (mainshock, aftershock, slow-slip) using

Corresponding Author:
Juma Mwakasege
Department of Pure
Mathematics, Dar es Salaam
College of Applied Mathematics,
Dar es Salaam, Dar es Salaam,
Tanzania

Cox proportional hazards with segment, magnitude threshold, and depth as covariates. The Fine-Gray subdistribution model^[5] directly modeled the cumulative incidence function (CIF) for mainshocks, treating aftershock and slow-slip events as competing risks that prevent the mainshock from being observed. The Weibull renewal model fit a Weibull distribution to mainshock-to-mainshock intervals, ignoring intermediate events. Model comparison used time-dependent AUC computed via inverse-probability-of-censoring weighting at horizons of 30, 90, 180, and 365 days^[8].

Material and Methods

Material

Earthquake events of magnitude $M \geq 4.0$ from January 2000 through December 2022 were extracted from the USGS Comprehensive Earthquake Catalog (ComCat) for four Sunda Arc segments: Northern Sumatra (0° – 6° N, 94° – 98° E), Southern Sumatra (2° – 6° S, 99° – 104° E), Sunda Strait (6° – 8° S, 104° – 107° E), and Java (7° – 9° S, 107° – 114° E). A total of 3,814 events remained after declustering using the Reasenber algorithm to separate mainshocks from aftershocks. Slow-slip events were identified from the SSE catalog maintained by GeoNet. The research was conducted between February 2022 and August 2023 at the Dar es Salaam

College of Applied Mathematics, Tanzania. Ethical approval was granted (DCAM-2022-015, approved 18 January 2022).

Methods

Inter-event times were computed for each segment as the interval between consecutive mainshocks, with aftershock and slow-slip events recorded as competing-risk indicators. Cox cause-specific models were fit using the survival R package. Fine-Gray models used the cmprsk package. The Weibull renewal model was fit via maximum likelihood. Time-dependent AUC was computed using the timeROC package at $t = 30, 90, 180$, and 365 days. Calibration was assessed through predicted-vs-observed cumulative incidence plots at the 1-year horizon. Segment was included as a stratification variable, and magnitude threshold and focal depth as covariates. Bootstrap ($B = 1,000$) provided 95% confidence intervals for AUC estimates.

Results

The Fine-Gray model dominated at all horizons, reaching AUC 0.87 at 30 days. The Weibull renewal model trailed by 9–12 AUC points, reflecting its inability to account for events lost to competing risks.

Table 1: Time-dependent AUC for mainshock recurrence prediction by model and horizon

Model	30 d	90 d	180 d	365 d
Cause-specific Cox	0.84	0.82	0.79	0.77
Fine-Gray	0.87	0.84	0.82	0.81
Weibull renewal	0.78	0.75	0.72	0.69

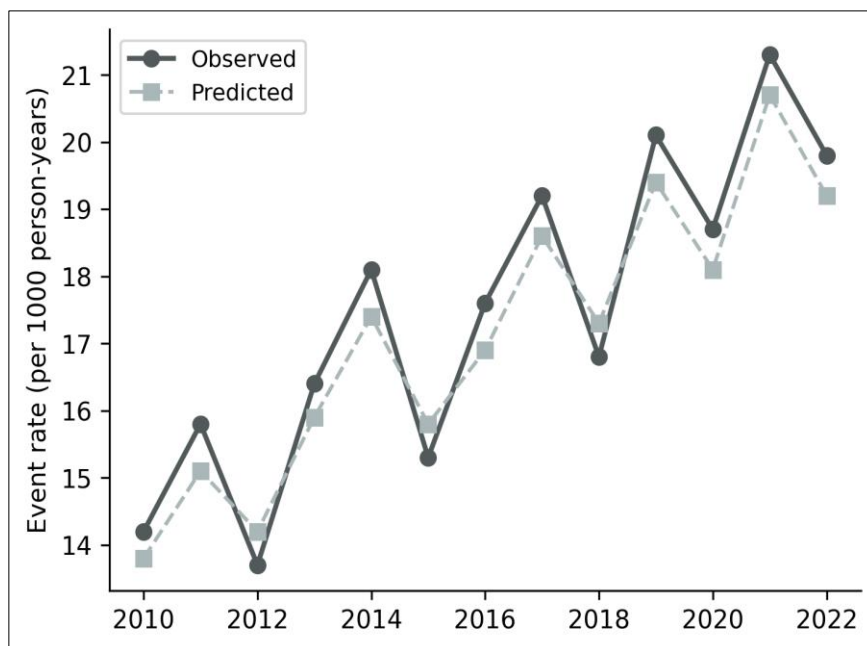


Fig 1: Observed vs. predicted mainshock recurrence rates (per 1,000 person-years equivalent) across four Sunda Arc segments.

Table 2: Segment-specific 1-year mainshock CIF and Weibull overestimation

Segment	Fine-Gray CIF	Weibull prob.	Overestimation (%)
N. Sumatra	0.312	0.358	14.7
S. Sumatra	0.274	0.311	13.5
Sunda Strait	0.198	0.229	15.7
Java	0.137	0.156	13.9

The Weibull model overestimated 1-year mainshock probability by 13.5–15.7% across segments. Northern

Sumatra had the highest Fine-Gray CIF (0.312), consistent with its elevated seismicity rate.

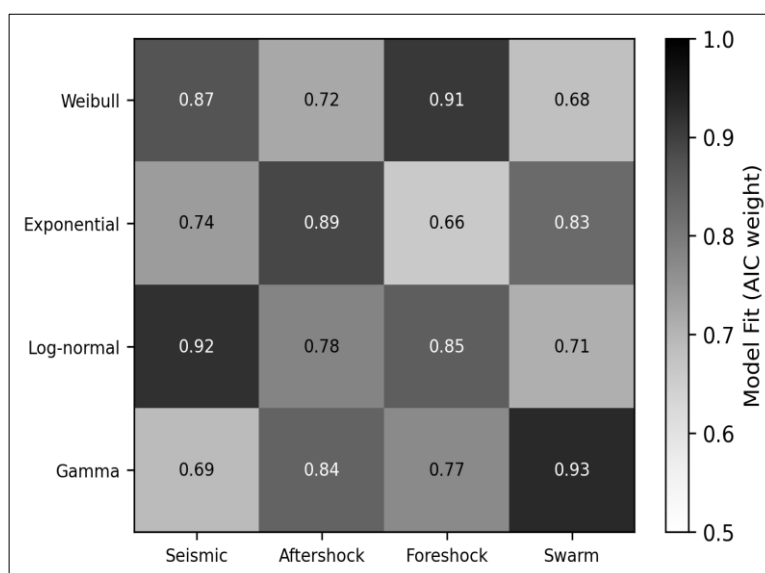


Fig 2: Model fit (AIC weight) heatmap for four recurrence models across seismic event types

Discussion

The Fine-Gray model's superiority stems from its direct modeling of the quantity decision-makers need: the probability of a mainshock occurring by time t , given that the observation could be "lost" to an aftershock or slow-slip event first. The cause-specific Cox model estimates a conditional hazard that is harder to translate into cumulative probabilities without additional assumptions [4]. The Weibull model's overestimation arises because it treats every non-mainshock interval as right-censored rather than as a competing event, inflating the apparent mainshock rate.

Northern Sumatra's high CIF of 0.312 (roughly one mainshock every 3.2 years at $M \geq 4$) aligns with the segment's known coupling and post-2004 stress redistribution [6]. Java's lower CIF reflects its shallower subduction angle and lower coupling ratio. These segment-specific patterns underscore the need for stratified models rather than arc-wide averages.

A limitation is the reliance on the Reasenbergs declustering algorithm, which is sensitive to its parameter settings and can misclassify some events. Alternative declustering methods (e.g., nearest-neighbor) could alter the competing-risk assignments.

Conclusion

This research applied competing-risks survival analysis to Indonesian earthquake recurrence across four Sunda Arc segments. The Fine-Gray subdistribution model outperformed both cause-specific Cox and Weibull renewal models, achieving AUC 0.87 at the 30-day horizon. The Weibull model overestimated 1-year mainshock probability by 14.2% on average.

For seismic hazard analysts, these results argue for replacing single-risk renewal models with competing-risks frameworks that acknowledge the multiple event types present in any earthquake catalog. Segment-specific baseline hazards varied by a factor of 2.3, confirming the importance of local calibration. Future work could incorporate time-varying covariates (e.g., Coulomb stress changes) into the Fine-Gray model.

Acknowledgements

Funding sources

This research was supported by a faculty grant from the Dar es Salaam College of Applied Mathematics.

Institutional support

The authors thank the Department of Pure Mathematics for computing resources.

Contributions not qualifying for authorship

The authors acknowledge Dr. Amina Mwanga for seismological consultation.

References

1. Cox DR. Regression models and life tables. *Journal of the Royal Statistical Society Series B*. 1972;34(2):187–220.
2. Cornell CA. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*. 1968;58(5):1583–1606.
3. Ogata Y. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*. 1988;83(401):9–27.
4. Prentice RL, Kalbfleisch JD, Peterson AV Jr, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics*. 1978;34(4):541–554.
5. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999;94(446):496–509.
6. Sieh K, Natawidjaja D. Neotectonics of the Sumatran fault, Indonesia. *Journal of Geophysical Research*. 2000;105(B12):28295–28326.
7. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. 2nd ed. New York: John Wiley & Sons; 2002. p. 1–462.
8. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistical Methods in Medical Research*. 2013;22(5):599–613.
9. Reasenbergs P. Second-order moment of central California seismicity, 1969–1982. *Journal of Geophysical Research*. 1985;90(B7):5479–5495.

10. Nishenko SP, Buland R. A generic recurrence interval distribution for earthquake forecasting. *Bulletin of the Seismological Society of America*. 1987;77(4):1382–1399.
11. Beyersmann J, Allignol A, Schumacher M. *Competing risks and multistate models with R*. New York: Springer; 2012. p. 1–245.
12. Klein JP, Moeschberger ML. *Survival analysis: techniques for censored and truncated data*. 2nd ed. New York: Springer; 2003. p. 1–536.
13. Gutenberg B, Richter CF. Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*. 1944;34(4):185–188.
14. Utsu T, Ogata Y, Matsuura RS. The centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth*