



Calculation of the Rate of $M \geq 6.5$ Earthquakes for California and Adjacent Portions of Nevada and Mexico

By Arthur Frankel and Charles Mueller

Open-File Report 2008–1112

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia 2008

For product and ordering information:
World Wide Web: <http://www.usgs.gov/pubprod>
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth,
its natural and living resources, natural hazards, and the environment:
World Wide Web: <http://www.usgs.gov>
Telephone: 1-888-ASK-USGS

Suggested citation:
Frankel, Arthur, and Mueller, Charles, 2008, Calculation of the rate of $M \geq 6.5$ earthquakes for California and adjacent portions of Nevada and Mexico, U.S. Geological Survey Open-File Report 2008-1112, 14 p.

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted material contained within this report.

Contents

Introduction	1
Procedure	1
Results	3
References Cited	5

Figures

1. $M \geq 6.0$ earthquakes since 1850 in the Working Group area of interest, with zones selected for this study..	7
2. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in extensional zone plotted against time.	8
3. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in San Francisco Bay zone plotted against time	9
4. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in Los Angeles zone plotted against time.	10
5. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in Mendocino zone plotted against time	11
6. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in remainder area plotted against time	12
7. Cumulative number of earthquakes $M \geq 6.5$ in remainder area plotted against time.....	13
8. Cumulative number of earthquakes $M \geq 7.0$ in remainder area plotted against time.....	14

Tables

1. Tests of procedure using synthetic catalogs	6
2. Results of Bayesian procedure for CGS/WG catalog	6

Calculation of the Rate of $M \geq 6.5$ Earthquakes for California and Adjacent Portions of Nevada and Mexico

By Arthur Frankel and Charles Mueller

Introduction

One of the key issues in the development of an earthquake recurrence model for California and adjacent portions of Nevada and Mexico is the comparison of the predicted rates of earthquakes with the observed rates (see, for example, Petersen and others, 2000). Therefore, it is important to make an accurate determination of the observed rate of $M \geq 6.5$ earthquakes in California and the adjacent region.

As part of the Working Group on California Earthquake Probabilities (WGCEP) effort, Felzer (2007) pointed out various issues with the earthquake catalog that must be addressed in order to accurately assess the rates of earthquakes. She noted that the uncertainty in the magnitude and the rounding procedures significantly affect the apparent rate of earthquakes in the catalog. Felzer (2007) assigned magnitude uncertainties and rounding parameters to the catalog compiled by the California Geological Survey (CGS). We declustered this catalog (which we denote as CGS/WG) using the procedure of Gardner and Knopoff (1974). The resulting catalog was used in the analysis described below, including the magnitude uncertainty and rounding determined for each earthquake by Felzer (2007). The region of study is identical to that used in WGCEP and consists of California with a border region including parts of Nevada, Oregon, Arizona, and Mexico.

Procedure

We have developed a procedure to calculate observed earthquake rates from an earthquake catalog, accounting for magnitude uncertainty and magnitude rounding. We present a Bayesian method that corrects for the effect of the magnitude uncertainty in calculating the observed rates. Using Bayes theorem, the probability distribution $p(M | M_{obs})$ of the true magnitude M of each event given an observed (calculated) magnitude M_{obs} can be determined from

$$p(M | M_{obs}) = p(M) p(M_{obs} | M) / P(M_{obs}) . \quad (1)$$

Here $p(M)$ is the *a priori* distribution of M , which is the Gutenberg-Richter recurrence relation, so that $p(M) = 10^{-bM}$. $p(M_{obs} | M)$ is the probability distribution of the observed magnitude

given the true magnitude M , assuming the magnitude uncertainty is normally distributed with a standard deviation of σ ,

$$p(M_{obs} | M) = e^{-(M_{obs}-M)^2/2\sigma^2}.$$

$P(M_{obs})$ in equation (1) is the total probability of having M_{obs} . Substituting into equation (1), we find

$$p(M | M_{obs}) = c10^{-bM} e^{-(M_{obs}-M)^2/2\sigma^2}. \quad (2)$$

Here c is a constant that is calculated so that the probability density for each event sums to unity.

In our procedure to find observed rates from a catalog with magnitude rounding and magnitude uncertainty, we first account for magnitude rounding. We create a probability distribution for each event, which has a magnitude $M_{reported}$ in the catalog. The distribution ranges between $M_{reported} - \text{round}/2$ and $M_{reported} + \text{round}/2$, where round is the rounding increment for that event (for example, 0.5 or 0.1) and the amplitude of the distribution has the form of 10^{-bM} . The distribution is normalized to give unit area.

Then, for each magnitude bin in the rounding distribution we apply equation (2) to account for magnitude uncertainty. Each bin in the rounding distribution represents a possible measured magnitude of the earthquake M_{obs} , before rounding was applied.

The probability density in each bin is the product of the density from the rounding distribution and the density from equation (2) for that magnitude bin. We apply our Bayesian procedure for $M_{obs} - M$ within $\pm 4\sigma$. We found that using even larger limits resulted in no significant change to the results.

The probability densities in each magnitude bin are added for all the events and then cumulative counts and rates are calculated above various magnitude thresholds. The total number of events in all the magnitude bins equals the original number of earthquakes in the catalog.

We tested this procedure using synthetic catalogs of earthquakes. We first produced synthetic catalogs from a Gutenberg-Richter distribution of earthquakes with respect to magnitude. Then, for each earthquake, a Gaussian random-number generator was used to assign an “observed” magnitude adjustment, given a specified sigma for the magnitude uncertainty. Finally, the resulting magnitude is rounded to the specified precision.

We used various combinations of magnitude uncertainty and rounding to test our method. The selected parameters span the values found in the actual catalog. For each combination, we produced five catalogs each starting with 10,000 earthquakes with $M \geq 4.0$. The b-value chosen was 0.8, as is found in our analysis of the actual catalog and by Felzer (2007). Table 1 shows the average number of events with $M \geq 6.5$ in the catalogs before the magnitude uncertainty and rounding were applied (“actual number of earthquakes”).

Table 1 also shows the average number of events with $M \geq 6.5$ calculated from the procedure described above. Note the similarity of the actual and calculated numbers of events, indicating that the methodology does an excellent job recovering the proper rate of earthquakes.

Results

In our estimation of rates in the actual catalog, we divide the region into zones to determine the proper observation time to consider. Figure 1 is a map of $M \geq 6.0$ earthquakes since 1850, showing the zones we selected. Two of these zones (San Francisco Bay and Los Angeles) were chosen because they encompass large population centers where one might expect longer periods of catalog completeness. The extensional zone contains Reno and towns such as Bishop and Independence which were populated in the mid-1800s (Felzer, 2007). We also outlined a zone around earthquakes in the Mendocino area, because the catalog there seems to effectively start in 1870.

Plots of cumulative number of events with $6.0 \leq M \leq 6.5$ versus time were made for each zone. We do need to include events with catalog magnitudes less than 6.5, because some of those earthquakes will contribute to the rate of $M \geq 6.5$ earthquakes because of magnitude uncertainty. In our rate calculation for $M \geq 6.5$ we use earthquakes in the catalog with magnitudes $M \geq 5.0$. In practice there is no significant difference in the rates found when using a minimum magnitude of 5.0 and 6.0. Obviously, we do not expect the $M5$ earthquakes to be complete for the time periods we considered.

For the extensional zone, it is clear that there is a high rate of $6.0 \leq M \leq 6.5$ earthquakes from 1850 to about 1890, with a lower rate from 1890 to 2000 (see fig. 2). In order to include this higher rate in our calculations we extended the observation time for this zone back to 1850. This zone encompasses the area around Reno, as well as the Mammoth Lakes and Owens Valley areas.

It is well documented (Ellsworth and others, 1981) that the San Francisco Bay area experienced a surge of activity in the 50 years prior to the 1906 $M7.8$ earthquake. This is apparent in the cumulative plot (fig. 3) of events in our San Francisco Bay zone with $6.0 \leq M \leq 6.5$. This area extends from the San Francisco Bay area southeast to the Parkfield area. The slope is much steeper from 1850-1900 than from 1900 to 2006. For the rate calculation for this zone, we start at 1850.

We also chose a zone that encompassed the greater Los Angeles, San Diego, and Santa Barbara areas. The plot of events with $6.0 \leq M \leq 6.5$ (fig. 4) shows a burst of activity in the 1850s with a variable rate after that. We chose an observation time of 1850 for this zone.

The Mendocino area exhibits a higher rate of $6.0 \leq M \leq 6.5$ earthquakes from 1870 to 1920 and a lower rate from 1920 to the present (fig. 5). There are only three earthquakes with $M \geq 4.0$ in the catalog for this region prior to 1870. These earthquakes were assigned to the same location and had a catalog magnitude of 5.5. Therefore, we start our calculation at 1870 for this region.

A key issue is what year to start the observation period for the events in the rest of California. The plot of $6.0 \leq M \leq 6.5$ events (fig. 6) is not conclusive. The rate from 1850 to 1890 is

relatively low, but similarly low rates are observed during other periods. Figure 7 shows the cumulative number of $M \geq 6.5$ for the region. From this plot it is clear that the low rate of $M \geq 6.5$ events prior to 1890 is anomalous compared to the period since 1890. Therefore, we chose 1890 as the start of our observation period for this zone. Most of the earthquakes in the remainder area occur in the Mojave zone specified by Felzer (2007). She derives a completeness time of 1910 for $M=6.6$ earthquakes in this zone. Thus our choice of 1890 as the observation time for $M \geq 6.5$ events in the remainder zone is reasonably consistent with her completeness time for this magnitude range.

We also implemented an alternative approach to calculate the rate of $M \geq 6.5$ events in the remainder area based on completeness times assigned for each magnitude bin. When determining the rate of $M \geq 6.5$ events, one can use a longer completeness time for $M7$ earthquakes than $M6.5$ earthquakes. The disadvantage of this method is that it uses different periods of time for different magnitude ranges above $M6.5$. In Figure 8, we show the cumulative number of $M \geq 7.0$ earthquakes in the remainder area. There are no $M \geq 7.0$ events in the remainder area from 1850 to 1891. It is not clear if this is due to incompleteness of the catalog or to a real lull of production. There is another period of similar duration, 1953-1991, with no $M \geq 7.0$ earthquakes in the remainder area.

We implement this alternative procedure by using the bins in our Bayesian calculation (equation 1). For each magnitude bin, the rate is determined from the effective number of events in that bin divided by the completeness period for that bin. Note that these bins also include the distribution from the rounding. The rates in each magnitude bin are then accumulated to get the total rate above a certain magnitude threshold.

In the alternative procedure, we use the completeness times developed by Felzer (2007) based on the distribution of population and newspapers across California as a function of time. All but one of the $M \geq 6.5$ events in the remainder area are in either Felzer's Mojave or "rest of state" zones (exception is the 1927 Hosgri earthquake). Because most of the events are in the Mojave zone, we use Felzer's completeness times for that zone: $M < 6.9$ 1890, $M6.9$ to 7.0 1880, $M7.0$ to 7.1 1875, $M7.1$ to 7.3 1870, $M7.3$ to 7.4 1865, $M7.4$ to 8.0 1855, $M \geq 8.0$ 1850. We found that this produces an 8 percent decrease in the calculated rate of $M \geq 6.5$ events in the remainder area. Because there were no $M \geq 7.0$ events in the catalog from 1850 to 1890, but we are applying completeness times prior to 1890 for these magnitudes, this lowers the calculated rate of $M \geq 6.5$ events compared to using an 1890 observation time for all magnitudes. The central question here is whether the lack of $M \geq 7.0$ events in the catalog for the remainder area from 1850 to 1890 is real or an artifact of catalog incompleteness.

Table 2 shows the observation time, number of $M \geq 6.5$ events, and rate we calculated for each zone. For the remainder zone and total, we cite the values determined from both methods described above.

To estimate uncertainty bounds, we apply the standard estimate of uncertainty for a Poissonian process that the standard deviation (sigma) of the observed count of earthquakes equals the square root of the number of earthquakes (see Weichert, 1980). We calculated the standard deviation in rate in each zone from the square root of the effective number of events with $M \geq 6.5$

divided by the observation period in each region (table 2). Then we added the variances of the five regions and took the square root of the sum to get the standard deviation for the total rate. We found a two-sigma value of 0.086 for the total rate. For the alternative method, the variance was determined for each magnitude bin based on the effective count and completeness time for that bin. These variances were summed for bins with $M \geq 6.5$ events and the square root was taken to find the standard deviation for the remainder area. We think this is a minimum estimate of the uncertainty, because we have not considered other sources of uncertainty.

It is reasonable to average the rates determined by the two methods, so our recommended determination of the observed rate of $M \geq 6.5$ is 0.246 ± 0.085 (for two sigma) per year. We stress that this rate is likely an underestimate of the true rate. For example, it is unlikely that the catalog for the extensional zone is complete to $M 6.5$ back to 1850 and it is certainly incomplete for $M 6.0$ - 6.4 earthquakes back to 1850. So the surge of the events in the catalog from 1850 to 1900, may still under-represent the actual increase in events during this time.

Acknowledgments

David Perkins suggested summing the variances of the rates in each region to get the total variance. We thank David Perkins and Mark Petersen for their reviews and valuable advice.

References Cited

Ellsworth, W.L., Lindh, A.G., Prescott, W.H. and Herd, D.G., 1981, The 1906 San Francisco earthquake and the seismic cycle, in *Earthquake Prediction*: D.W. Simpson and P.G. Richards, eds. American Geophysical Union, Washington, D.C., p. 126-140.

Felzer, K. (2007). Appendix I: Calculating California seismicity rates, in *The Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2): 2007 Working Group on California Earthquake Probabilities and the USGS National Seismic Hazard Mapping Program (NSHMP)*, U.S. Geological Survey Open-File Report 2007-1437-I.

Gardner, J.K. and Knopoff, L., 1974, Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America*, v. 64, p. 1363-1367.

Petersen, M.D., Cramer, C.H., Reichle, M.S., Frankel, A.D., and Hanks, T.C., 2000, Discrepancy between earthquake rates implied by historic earthquakes and a consensus geologic source model for California, *Bulletin of the Seismological Society of America*, v. 90, p. 1117-1132.

Weichert, D.H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, *Bulletin of the Seismological Society of America*, v. 70, p. 1337-1346.

Table 1. Tests of procedure using synthetic catalogs

Magnitude Uncertainty (?)	Rounding	Actual Number of $M \geq 6.5$ Earthquakes	Calculated Number of $M \geq 6.5$ Earthquakes from Bayesian Procedure
0.333	0.1	99.6	99.2
0.4	0.5	100.2	98.4
0.1	0.1	97.2	96.7
0.2	0.01	93.0	93.8

Table 2. Results of Bayesian procedure for CGS/WG catalog

zone	observation start	calculated number of $M \geq 6.5$ events	calculated rate of $M \geq 6.5$ events per year	estimated standard deviation of rate
extensional	1850	3.91	0.0251	0.0127
San Francisco Bay	1850	7.09	0.0454	0.0171
Los Angeles	1850	5.05	0.0324	0.0144
Mendocino	1870	6.39	0.0470	0.0186
remainder of region	1890	11.55	0.0996*, 0.0918**	0.0293*, 0.0271**
total			0.250*, 0.242**	0.043*, 0.042**

* determined from $M \geq 6.0$ events since 1890

** determined from alternative method using completeness times as a function of magnitude above $M6.9$

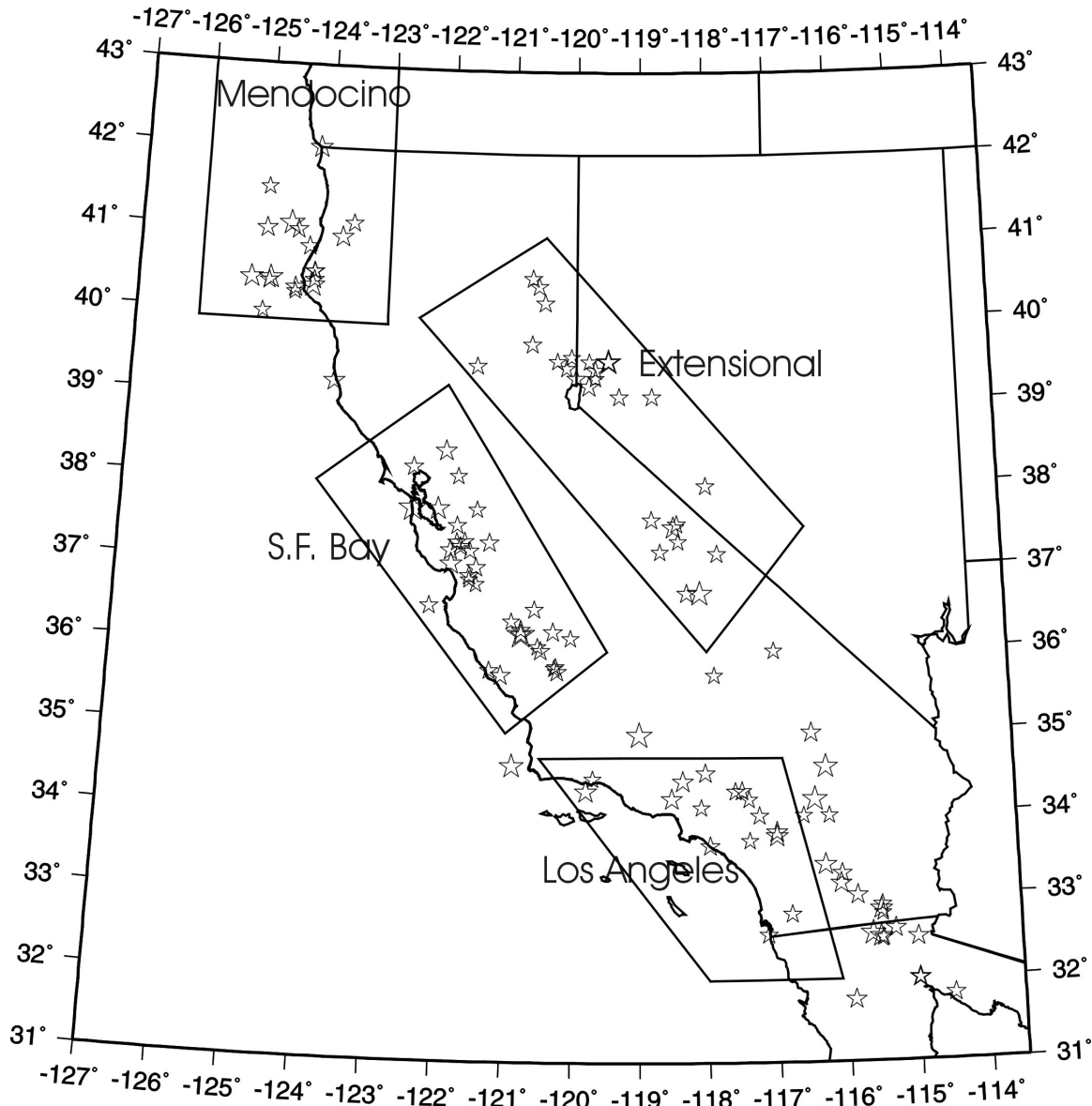


Figure 1. $M \geq 6.0$ earthquakes since 1850 in the Working Group area of interest, with zones selected for this study. Earthquakes from the California Geological Survey/Working Group for California Earthquake Probabilities catalog.

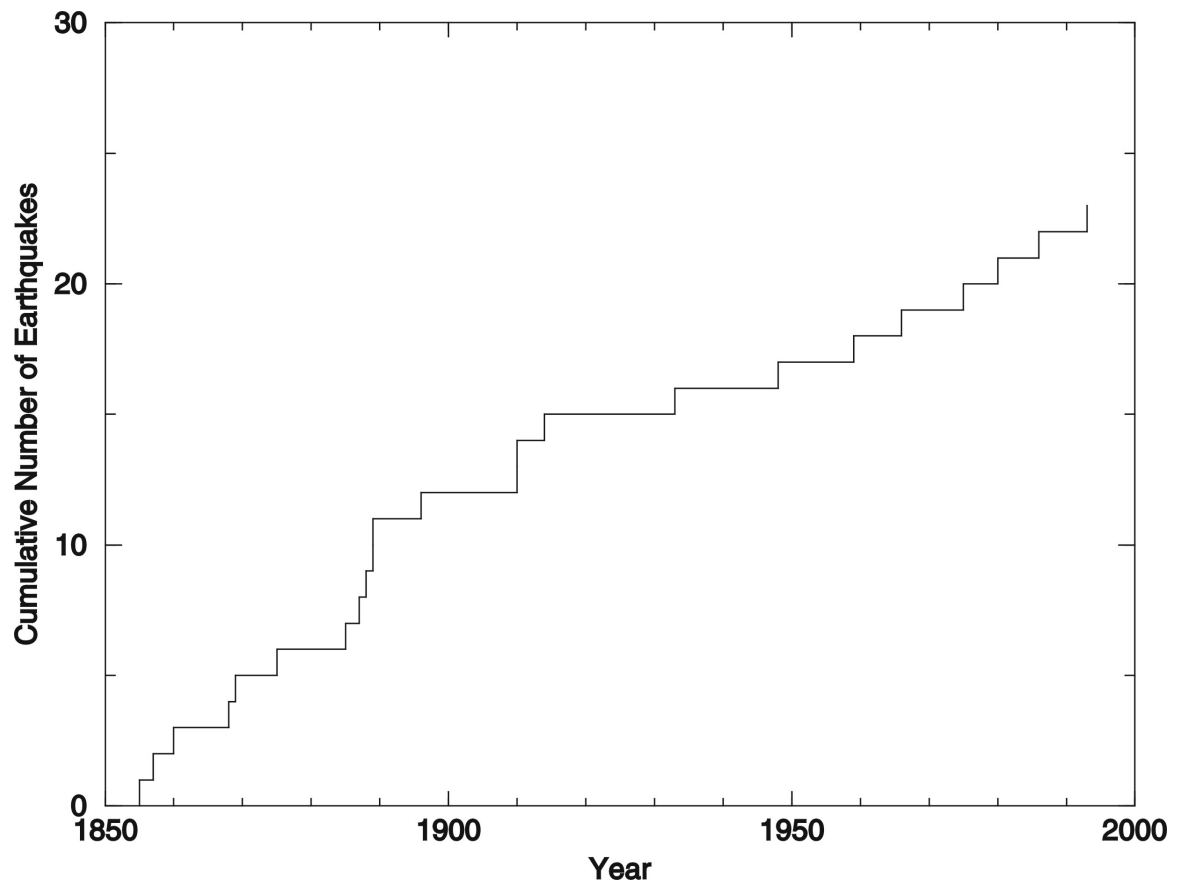


Figure 2. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in extensional zone plotted against time.

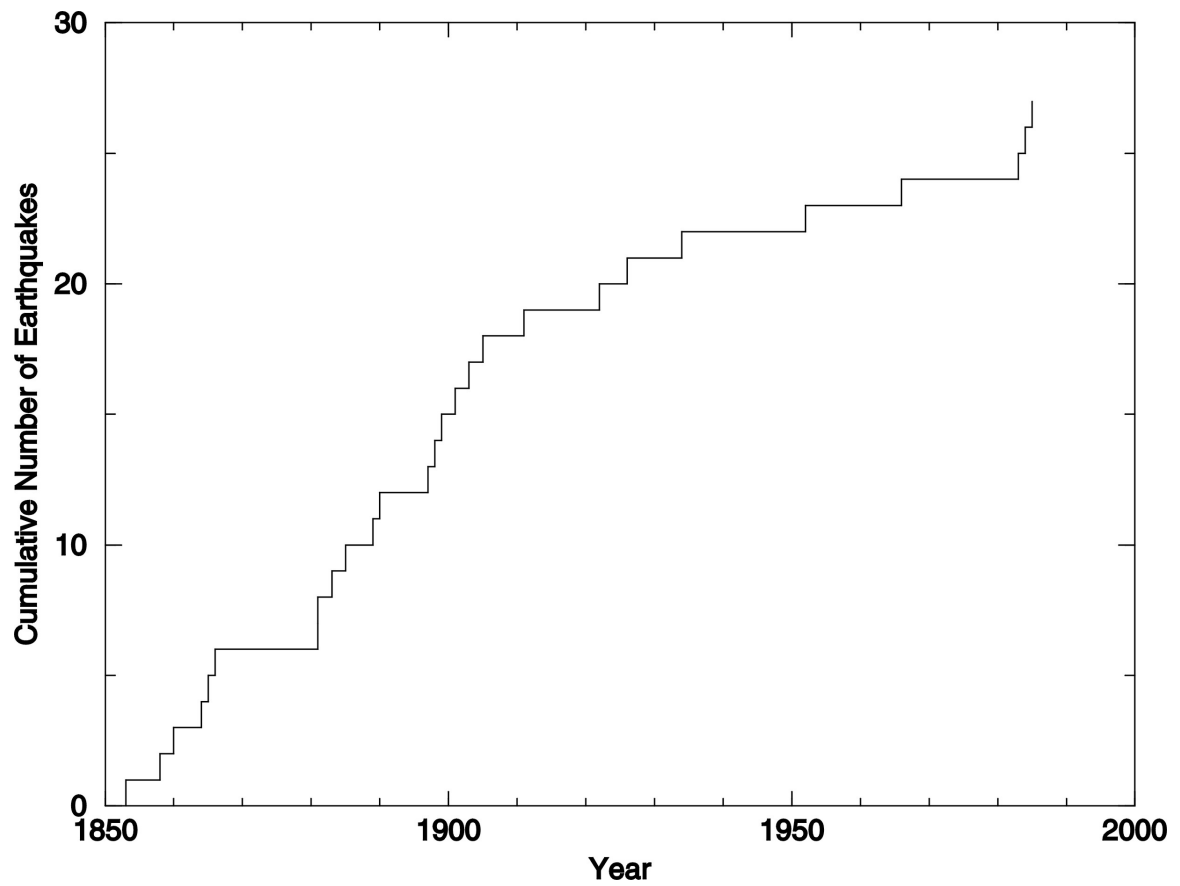


Figure 3. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in San Francisco Bay zone plotted against time.

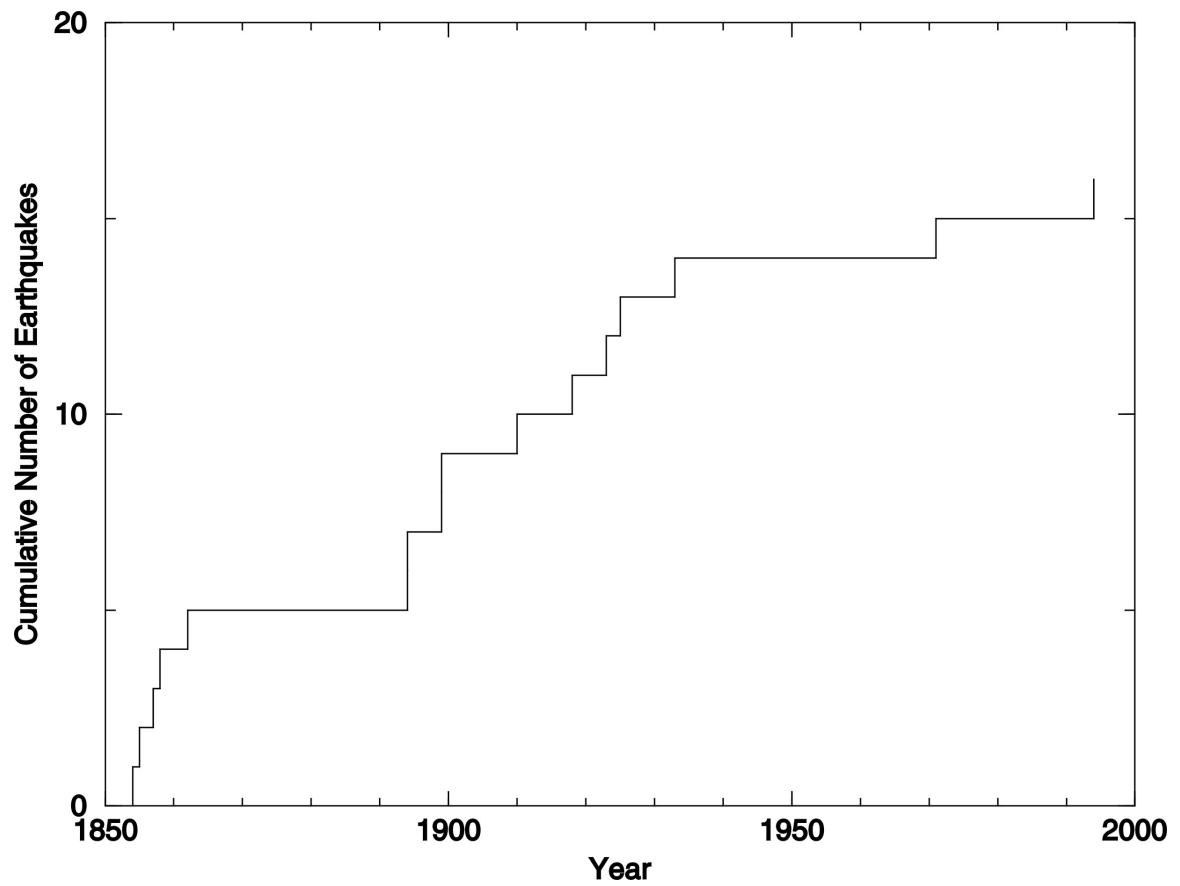


Figure 4. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in Los Angeles zone plotted against time.

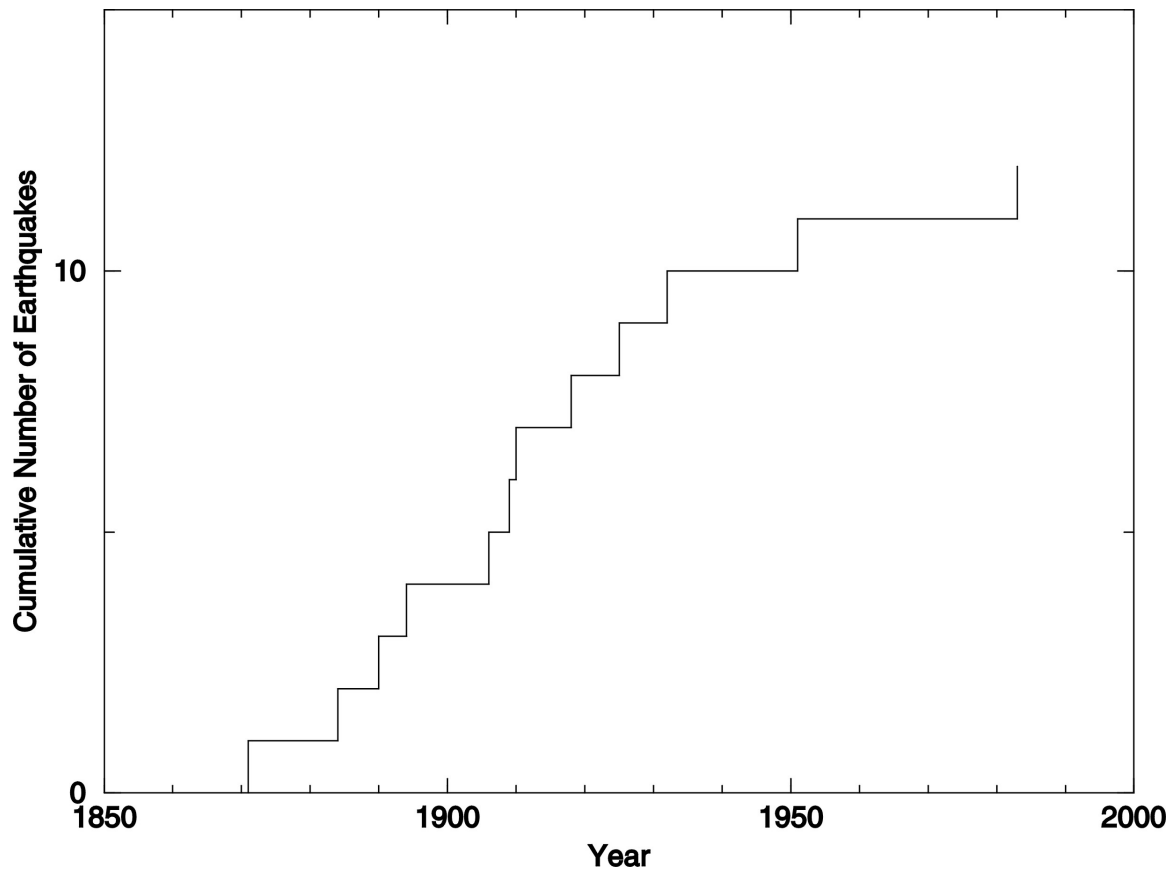


Figure 5. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in Mendocino zone plotted against time.

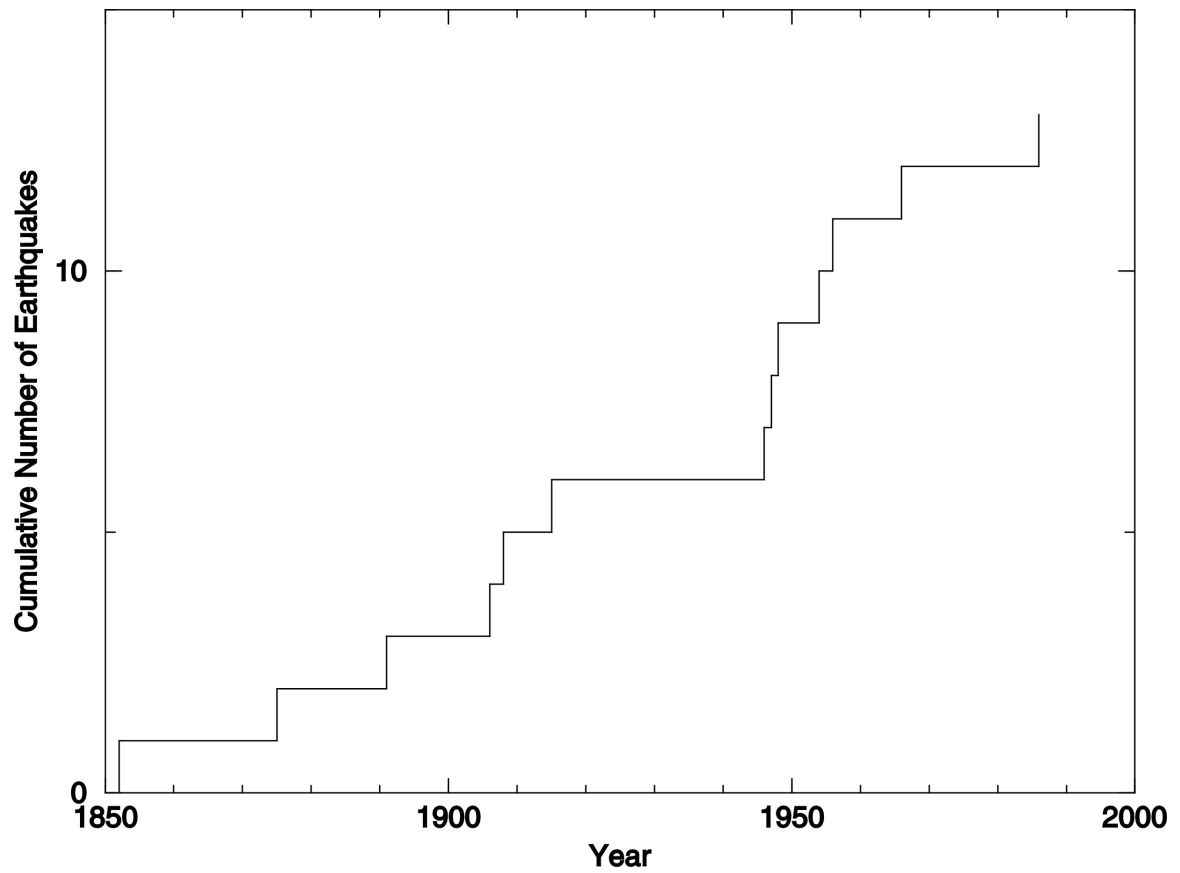


Figure 6. Cumulative number of earthquakes $6.0 \leq M \leq 6.5$ in remainder area plotted against time.

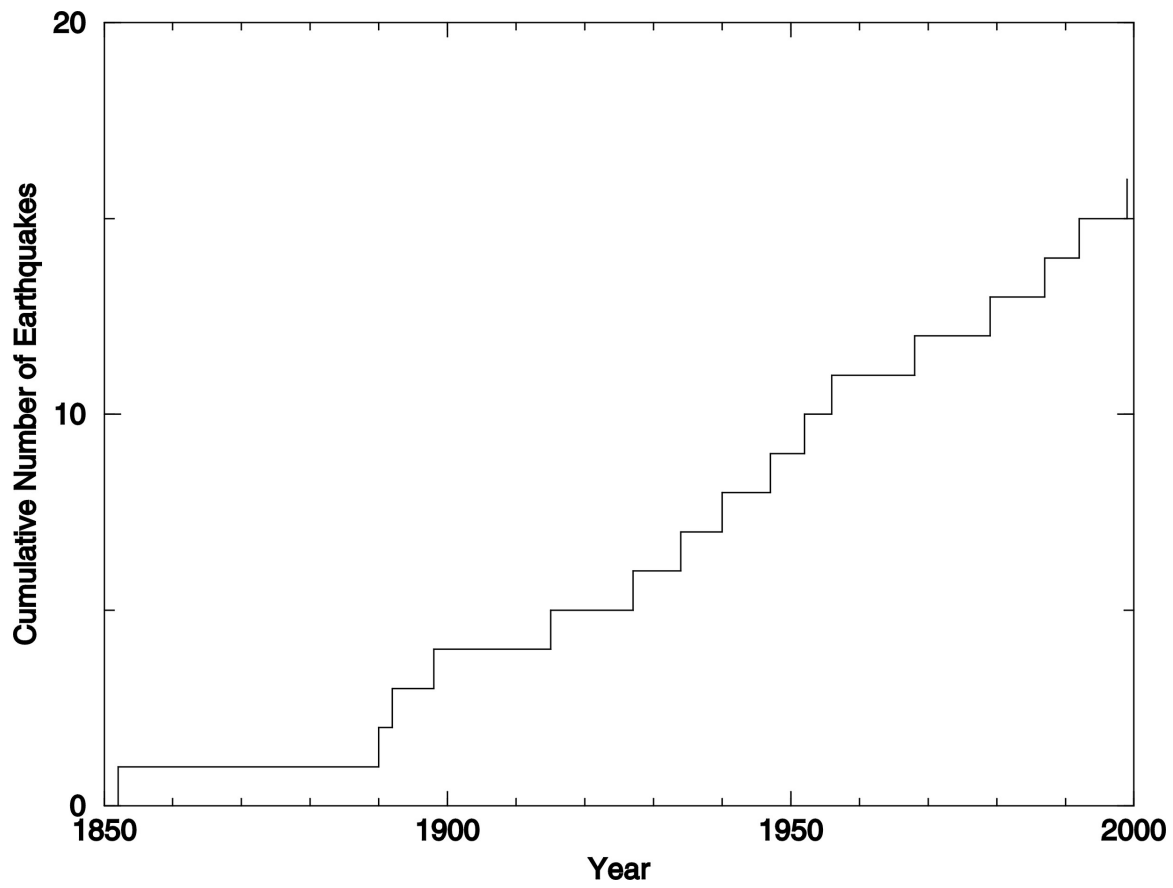


Figure 7. Cumulative number of earthquakes $M \geq 6.5$ in remainder area plotted against time.

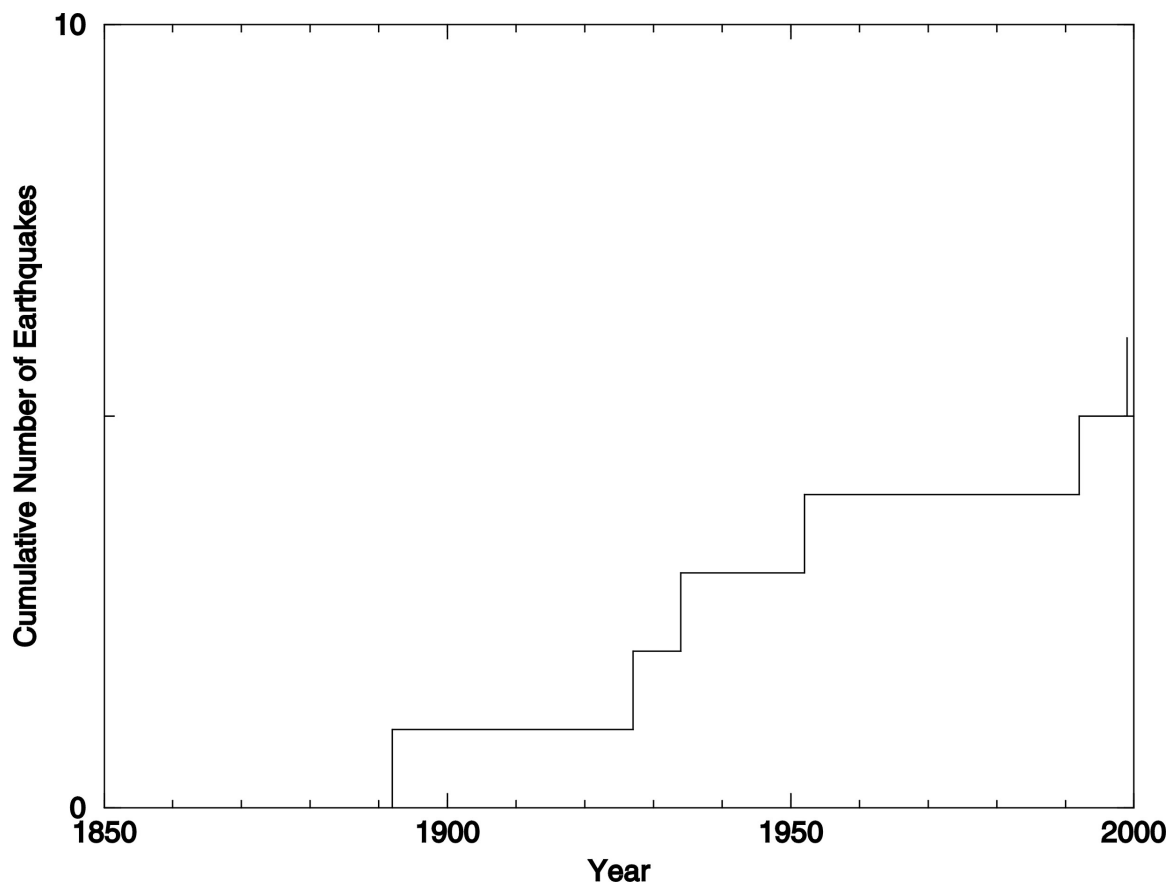


Figure 8. Cumulative number of earthquakes $M \geq 7.0$ in remainder area plotted against time.