Natural Hazards and Earth System Sciences (2003) 3: 95–101

Natural Hazards and Earth System Sciences

A stochastic marked point process model for earthquakes

L. Holden[1], **S. Sannan**[1], **and H. Bungum**[2]

[1]Norwegian Computing Center, P.O. Box 114 Blindern, N-0314 Oslo, Norway
[2]NORSAR, P.O. Box 51, N-2027 Kjeller, Norway

Received: 20 July 2001 – Revised: 28 January 2002 – Accepted: 27 February 2002

Abstract. A simplified stochastic model for earthquake occurrence focusing on the spatio-temporal interactions between earthquakes is presented. The model is a marked point process model in which each earthquake is represented by its magnitude and coordinates in space and time. The model incorporates the occurrence of aftershocks as well as the build-up and subsequent release of strain. The parameters of the model are estimated from a maximum likelihood calculation.

1 Introduction

Earthquake forecasting in the strict sense with the exact prediction of the time, the location, and the magnitude of an earthquake has been a difficult area of research for several decades. One outcome of this research, however, is that we today know much more about why earthquake prediction is difficult (e.g. Kagan, 1997; Sykes et al., 1999). This difficulty is in part tied to concepts such as self-similarity, criticality and nucleation processes: All earthquakes start small, and while we know much about the limits to growth, we do not know in sufficient details when and why a rupture stops before that.

In this paper we outline a stochastic model for earthquake occurrence which is focusing on the spatio-temporal interactions between earthquakes, including the effects of aftershocks as well as the build-up and release of strain. The model is a marked point process model (e.g. Cressie, 1993; Ripley, 1987) in which each earthquake is represented by its magnitude and coordinates in space and time. Hence, this is a simplified earthquake model which does not include physical quantities such as the dimensions of the faults, rupture characteristics, etc. Because each earthquake is represented separately, however, it is feasible to include known physical quantities connected to the individual earthquakes in an extended version of the model.

Correspondence to: L. Holden (lars.holden@nr.no)

The model is based upon a Bayesian approach with user specified prior distributions for all parameters, while empirical data are used for deriving posterior distributions. There are many ways to parameterize the model, however, and in this paper we present only one of these. The basic principles behind the algorithms we have used and the following calculations remain independent of this particular parameterization. Another freedom of the model is the choice of prior distributions, where, when faced with an unresolved situation, one may revert to flat priors. This, however, gives less information and subsequently may lead to less precision in the estimates.

2 Marked point process model

Marked point processes are commonly used stochastic models for representing a finite number of events located in space and time. Earthquake occurrence can very well be described by a marked point process model. Each earthquake has, in addition to a location in space and time, parameters representing the magnitude and quite often also information about the earthquake fault lines. Point process models for earthquakes have previously been discussed by Vere-Jones (1995) and Ogata (1998). The model presented in this paper treats aftershocks in a similar fashion to the work by Ogata. In addition, the model takes into account the effect of strain build-up. The ultimate goal is to include as much information as possible of known physical processes into the model.

2.1 The model

In our notation an earthquake is represented by $E = (x, M)$ and t, where $x = (x^1, x^2, x^3)$ is the hypocenter coordinates (x^1 = longitude, x^2 = latitude, x^3 = depth), M is the moment magnitude, and t is the time. An earthquake catalog $H_T = \{(E_i, t_i)\}_{t_i<T} = \{(x_i, M_i, t_i)\}_{t_i<T}$ consists of all observed earthquakes above a certain magnitude in a specified region, and in a given time period (T_0, T).

The two major assumptions made in the proposed model are:

- The intensity $\lambda_1(E, t|H_t, \beta)$ of earthquakes is a function of the parameters $(E, t) = (x, M, t)$, all previous earthquakes H_t in the region, and some parameters β to be determined by Bayesian updating. If additional data or physical information is available, this should be included in this intensity.

- The time averaged intensity $\lambda(E) = \lambda(x)\lambda(M|x)$ as a function of magnitude and location is known.

The time independent intensity $\lambda(x)$ represents the average number of earthquakes per unit time and unit volume. We have here estimated $\lambda(x)$ from a catalog. It may, however, be possible to estimate $\lambda(x)$ on the basis of additional geological data of earthquakes in combination with a catalog. This is believed to give more stable estimates because a catalog may have completeness problems. In particular, it may not cover a sufficient number of large earthquakes in each region to give the stability that is desired. There has also been a great deal of debate recently concerning the frequency-magnitude distribution for very large earthquakes (e.g. Kagan, 1999; Main, 2000). For the simple model presented here, however, any particular choice of the high-end cutoff is of no fundamental importance for the results. To simplify matters to this end we have let $\lambda(M|x)$ be determined by the well-known Gutenberg-Richter frequency-magnitude law such that $\lambda(M|x) \propto 10^{-bM}$, where the value of the scaling parameter b usually is in the interval (0.7, 1.2) (e.g. Vere-Jones, 1995).

We will assume that the intensity λ_1 is given by

$$\lambda_1(E, t|H_t, \beta) = \lambda_2(E|\beta)s(x, t|H_t, \beta) + \lambda_3(E, t|H_t, \beta), \quad (1)$$

where λ_2 is the background intensity, s represents the effect of strain build-up, and λ_3 represents the increase in the intensity after an earthquake and is used for modeling the aftershocks. The background intensity $\lambda_2(E|\beta)$ depends implicitly on the parameters β through the requirement that $\lambda_1(E, t|H_t, \beta)$, averaged over time, equals $\lambda(E)$. If the strain build-up is omitted, $s = 1$, while if the aftershock treatment is omitted, $\lambda_3 = 0$. In the case that both $s = 1$ and $\lambda_3 = 0$ the model is just reduced to a simple Poisson model.

The intensity λ_3 is used to model the aftershocks. Let M_i be the magnitude of a shock in the catalog at the time t_i, and M the magnitude of a subsequent shock. Aftershocks M are then modeled by $\lambda_3(E, t|H_t, \beta) > 0$ for earthquakes $M < M_i$ for $t > t_i$. We will assume that λ_3 has the form

$$\lambda_3(E, t|H_t, \beta) = \sum_{(E_i, t_i) \in H_t} g(E, t, E_i, t_i, \beta), \quad (2)$$

where

$$g(E, t, E_i, t_i, \beta) = \beta_1 g_1(M, M_i, \beta_2)\, g_2(t, t_i, \beta_3, \beta_4)\, g_3(x, x_i, \beta_5). \quad (3)$$

The functions g_1, g_2, and g_3 represent magnitudial, temporal, and spatial effects, respectively. Note that the summation implies that if there is a large earthquake followed by a series of smaller earthquakes, all of these earthquakes contribute to the intensity. A typical form of g_1 is $g_1(M, M_i, \beta_2) = \exp(\beta_2 M_i)\, 10^{-bM}$, in accordance with the Gutenberg-Richter law. For the temporal effect we assume that $g_2(t, t_i, \beta_3, \beta_4) = 1/(t - t_i + \beta_4)^{\beta_3}$, which is essentially Omori's law for aftershocks (see e.g. Lay and Wallace, 1995). The spatial effect can be represented by a function based on the distance between the hypocenters, i.e. $g_3(x, x_i, \beta_5) = \exp(-\beta_5 ||x - x_i||^2)$.

It seems to be generally accepted that there is more regularity in the occurrence of earthquakes than can be accounted for in a Poisson model (Working Group on California Earthquakes Probabilities, 1995). The assumption is that in any particular region, strain is slowly building up and then released due to earthquakes. This effect can be incorporated into a point process model. We first define a state variable S that can be connected to strain, or to stress, for that matter. The interpretation of S may be different than the standard definition of strain (or stress), but its general nature will be such that it reflects the spatio-temporal release of strain/stress energy. For simplicity, we refer to S as strain in the following. We define S by

$$S(x, t|H_t, \beta) = S_0(x, T_0) + \phi(x, \beta)(t - T_0) - \sum_{(E_i, t_i) \in H_t} h(x, E_i, \beta), \quad (4)$$

with

$$\phi(x, \beta) = \int h(x, E', \beta)\lambda(E')dE', \quad (5)$$

and

$$h(x, E', \beta) = \exp(\beta_7 M') \exp(-\beta_8 ||x - x'||^2), \quad (6)$$

where the integral in Eq. (5) means an integration $\int_R dx$ over the particular region R we are considering, plus a sum over magnitudes $\sum_{M_{\min}}^{M_{\max}}$, where $M_{\min}$ is the smallest M in the catalog H_T and $M_{\max}$ is taken to be the largest M in H_T. The function ϕ represents the average strain build-up per unit time and h is the release of strain for each earthquake. We have made the assumption that the mean strain release equals the mean strain build-up, which of course is not quite correct for catalogs being shorter than the recurrence time for larger earthquakes. At the current stage of testing the model, however, this is not so crucial. A mismatch between the assumption and reality is not likely to affect the balance between the estimated parameters significantly.

The strain release h is factored into two terms related to the magnitude of the earthquake and a spatial effect, respectively. Thus, S represents the strain at any point (x, t) in space and time, given all the previous earthquakes contained in the catalog H_t. It is assumed that S builds up linearly and then decreases instantaneously with each earthquake. The

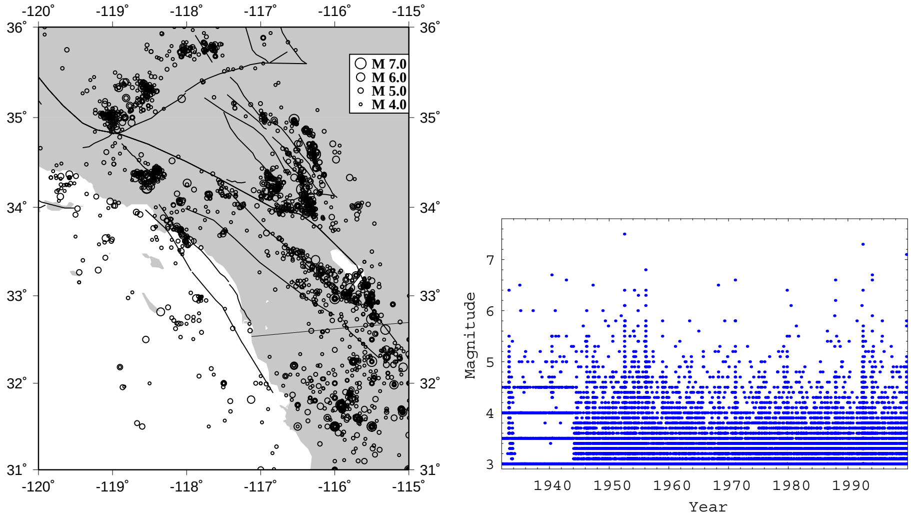

Fig. 1. The location of earthquakes with $M \geq 4.0$ plotted together with the major faults (left), and the magnitude vs. time for earthquakes with $M \geq 3.0$ (right) in Southern California in the period 1932–1999.

strain has a variability that is independent of time, and s is an increasing function of S given by

$$s(x, t|H_t, \beta) = \exp(\beta_6 S). \qquad (7)$$

By this parametrization it is implicitly given that the time-averaged intensity of earthquakes equals $\lambda(E)$. The effect of s is to reduce the variability of strain in time periods between very large earthquakes compared to the simple Poisson model. The variability in time periods between large earthquakes becomes smaller and smaller with increasingly large values of β_6 and β_7. The parameter β_8 specifies the surrounding region of an earthquake in which strain is released.

2.2 Posterior distributions for the parameters

From the real catalog H_T of the period (T_0, T) it is possible to find the posterior distributions of the parameters β. These posterior distributions represent the best guesses for the parameters and should be used in all predictions. The posterior distributions for the parameters β, given the data in the catalog H_T, are defined by the equation

$$f(\beta|H_T) \propto f(\beta) f(H_T|\beta). \qquad (8)$$

The likelihood $f(H_T|\beta)$ can be calculated from

$$f(H_T|\beta) =$$
$$\exp\left(-\int_{T_0}^{T}\int \lambda_1(E, t|H_t, \beta) dE dt\right) \prod_{i=1}^{n} \lambda_1(E_i, t_i|H_{t_i}, \beta)$$
$$\approx c(\beta) \prod_{i=1}^{n} \lambda_1(E_i, t_i|H_{t_i}, \beta), \qquad (9)$$

where the factor $\exp\left(-\int_{T_0}^{T}\int \lambda_1(E, t|H_t, \beta) dE dt\right)$ is approximated by a constant $c(\beta)$. This factor is due to periods (t_{i-1}, t_i) without earthquakes, while the factor $\prod_{i=1}^{n} \lambda_1(E_i, t_i|H_{t_i}, \beta)$ represents the intensities for the actual earthquakes. The maximum likelihood estimates for the parameters are the parameters that maximizes the expression (9).

2.3 Simulations and predictions

Simulation from the model is fairly straightforward. The simulated earthquakes are generated one at a time in chronological order, and the intensity for each new earthquake is given by Eq. (1). An example of a simulation is given in the next section.

Predictions can be done by first sampling n values of the parameter set β from the distributions (8), using the most recent earthquake catalog H_t for the region. These parameter sets can be found by simulating β by Markov chain Monte Carlo methods (e.g. Cressie, 1993; Ripley, 1987). In a Markov chain Monte Carlo simulation one defines a chain of parameter sets β that satisfies the specified distributions. For each of these sets of parameters β new earthquakes are simulated based upon the intensity $\lambda_1(E, t|H_t, \beta)$, where H_t is continuously updated during simulation. Probabilistic predictions are then obtained simply by counting the number of earthquakes in the different simulations. For short-term predictions and when the intensity is low it is also possible to use the intensities directly.

3 Results

3.1 Parameter estimation

The initial body of work for testing our model has been to implement an optimization algorithm in C++ that maximizes the likelihood (9), i.e. an algorithm that calculates the maximum likelihood estimator for the parameters β. The empirical data we have used are based on an earthquake catalog over the time span 1932–1999 compiled by the Southern California Earthquake Center, SCEC (2000), and limited to the region 31–36° N, 115–120° W. We have also limited the data to include only the epicenter coordinates x^1 = longitude and x^2 = latitude, thus neglecting the hypocenter coordinate x^3 = depth. The location of all earthquakes of magnitudes

Table 1. The absolute values of the relative variations of $\log f(H_T, \beta)$ (given in %) when the parameters β are varied $\pm 1\%$ relative to the values giving maximum likelihood

	Relative variations of $\log f(H_T, \beta)$ (in %)	
Parameter	No strain	Strain
β_1	0.00438	0.00446
β_2	0.101	0.103
β_3	0.0470	0.0477
β_4	0.000442	0.000447
β_5	0.00416	0.00422
β_6	–	0.0000870

$M \geq 4.0$ for this region and the given time span are shown in Fig. 1 (left). We have included all earthquakes of magnitude $M \geq 3.0$ in our data set, which comprises 15 804 earthquakes over the time-span of 24837 days, giving an average intensity of 0.636 earthquakes per day. A simple time-magnitude plot in Fig. 1 (right) shows that the completeness seems to be reasonable back to 1944. The 5×5 degree area is divided into 1600 grid cells, with each cell corresponding to a size of about 14×12 km. The intensity $\lambda(E)$ for each grid cell is calculated from the empirical data. An average b-value of 0.93 for the Gutenberg-Richter relation is estimated from the data.

To calculate the maximum likelihood estimator for the parameters β, we have used the 58-year period 1942–1999 of the SCEC catalog. When the strain build-up is omitted, and hence $s = 1$, the estimation yields $\beta_1 = 0.478$, $\beta_2 = 1.20$, $\beta_3 = 1.24$, $\beta_4 = 0.0191$, and $\beta_5 = 2.07 \times 10^3$. In the estimation we have assumed that the contributions to $\lambda_3(E, t|H_t, \beta)$ for each earthquake go no more than 10 years back in time. Hence, the 10-year period 1932–1941 of the empirical data is used indirectly in the calculation, giving contributions to λ_3 for quakes in the period 1942–1951. A detailed analysis of the parameters indicates that about 60% of the earthquakes in the period 1942–1999 (according to the model) were aftershocks, while the remaining 40% of the earthquakes were related to background activity. An interpretation of the parameter β_2 is that the increased intensity due to an earthquake of magnitude $M = 5.6$ is twice that of a $M = 5.0$ earthquake. The values of β_3 and β_4 indicate that the increase of intensity is halved 20 minutes after an earthquake. Finally, the value of β_5 indicates that the increased intensity is halved at a distance of 1.9 km from the epicenter of the earthquake.

If we include the build-up of strain, the calculation of the maximum likelihoods yields $\beta_1 = 0.480$, $\beta_2 = 1.20$, $\beta_3 = 1.24$, $\beta_4 = 0.0193$, $\beta_5 = 2.07 \times 10^3$, $\beta_6 = 0.000470$, $\beta_7 = 0$, and $\beta_8 = 0$. Therefore, turning on the strain build-up yields minor corrections only to β_1, β_2, β_3, β_4, and β_5, while β_7 and β_8 both vanish. The direct interpretation of $\beta_7 = 0$ is that the build-up and release of strain (note that this is not strain energy) is independent of the magnitudes of the individual earthquakes, and that it is the number of earthquakes (the cumulative effect) that matters in this context. Indirectly, however, a large earthquake has a greater impact than a small one because it triggers a larger number of aftershocks.

The result $\beta_8 = 0$ means that each earthquake in the catalog has a global effect on the build-up and release of strain. This is difficult to interpret physically but a plausible explanation of the result is that we so far in the model have used epicentral distances instead of the smallest distance to the causative fault (the rupture plane). Another significant source of error may be related to our crude assumption that $S = S_0$ at the beginning and the end of the catalog. This assumption was made from simplicity and from the lack of knowledge about the actual state of strain in the various regions and different points of time. It is also clear that the earthquake catalog we have used in estimating the parameters of the model covers a too short time span (1932–1999) to carry sufficient information on the build-up and release of strain. The vanishing of β_7 suggests that this parameter, as given by the expression (6), is superfluous in the model. A better representation of the strain release might therefore be given by

$$h(x, E', \beta) = \exp(-\beta_7 ||x - x'||^2 / 10^{M'}), \tag{10}$$

in accordance with the common observations that the spatial extension L of small earthquakes behaves like $L \propto 10^{M/2}$. With $M \propto \frac{2}{3} \log M_0$, where M_0 is the seismic moment, this corresponds to $M_0 \propto L^3$, implying self-similarity. In the expression (10) the parameter β_7 corresponds to β_8 in (6), i.e., there is advantageously one less parameter in the model to be estimated. For large earthquakes the scaling law question is a lot more controversial, even though the original suggestion by Scholz (1982) of $L \propto 10^{3M/4}$ (or $M_0 \propto L^2$) still seems to be the most viable one. In that case a slightly different expression for $h(x, E', \beta)$ would have to be used. The incorporation of such scaling relationships, however, lies in future work of our model.

To test the sensitivity of the likelihood (9) with respect to the parameters β we calculated the values of $\log f(H_T, \beta)$ when the parameters in turn were varied 1% away from the values giving maximum likelihood. In Table 1 are shown the absolute values of the relative variations of $\log f(H_T, \beta)$ when the parameters are varied for both the case of no strain and the case of including the build-up of strain. Since $\log f(H_T, \beta)$ is fairly symmetric about its maximum in all directions, the given variations are the means of the relative variations of $\log f(H_T, \beta)$ when the parameters are varied $\pm 1\%$ relative to the values corresponding to maximum likelihood. We notice that $\log f(H_T, \beta)$ is very stable with respect to variations in the β's. Furthermore, the sensitivity of $\log f(H_T, \beta)$ to the parameters $\beta_1, \ldots, \beta_5$ is quite unaffected by the turn-on of strain.

3.2 Simulations of earthquakes

The second part of our work has been to implement an algorithm for simulating earthquakes in the region of interest,

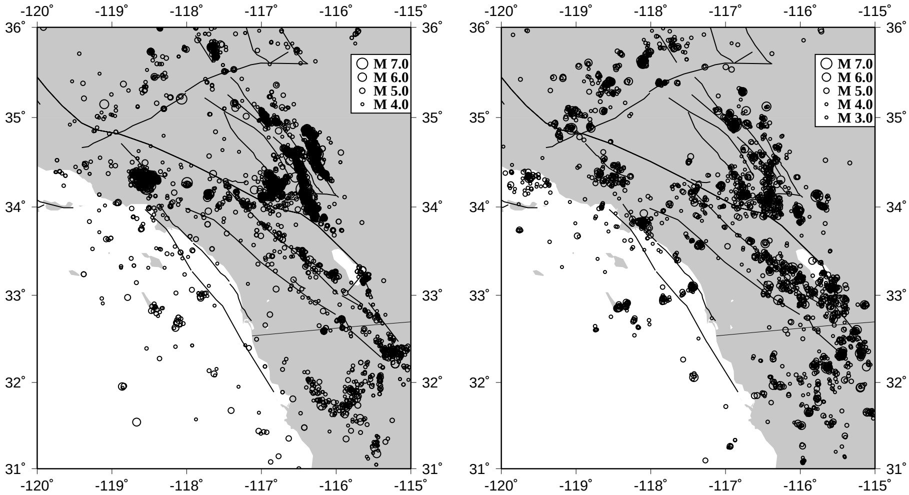

Fig. 2. Earthquakes of $M \geq 4.0$ in the period 1990–1999 in Southern California (left), and simulated earthquakes over the same 10-year period using the marked point process model (right). The simulation is based on data from the time period 1932–1999.

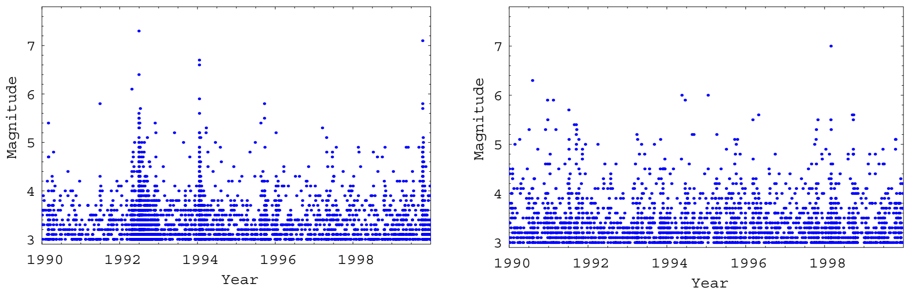

Fig. 3. Magnitude vs. time of actual earthquakes in the period 1990–1999 (left), and of simulated earthquakes over the same 10-year period (right) in Southern California.

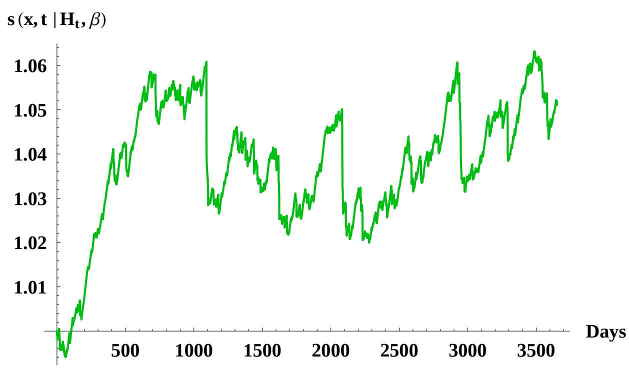

Fig. 4. The strain function $s(x, t|H_t, \beta)$ for a typical simulation taking into account the build-up and release of strain. $s(x, t|H_t, \beta)$ increases when the activity is below the average background intensity and drops suddenly due to the aftershock activity following a large earthquake. The abrupt increase of $s(x, t|H_t, \beta)$ over the first 500 days is due to a low S_0 value.

given the estimated values of the β's. Simulated earthquakes are generated one at a time in chronological order. For each new earthquake there is an increase or a decrease in the intensity (1), due to a sudden increase in the term λ_3 and a sudden decrease in the strain function s. The algorithm may use a longer catalog to establish the intensity of the background activity, while it may use a 10-year catalog to establish an intensity for the aftershock activity. As the simulation proceeds, the latter catalog is gradually replaced with the simulated earthquakes to regulate the intensity of aftershocks. An example of a simulation over the 10-year period 1990–1999 in the case of $s = 1$ is shown to the right in Figs. 2 and 3. This particular simulation thus does not take the build-up and release of strain into account. The background activity is due to the 1932–1999 catalog and the aftershock activity is initially due to the 1980–1989 data. The simulation resulted in 2315 earthquakes over 3652 days, corresponding to 0.634 earthquakes per day. This agrees very well with the average intensity of 0.636 earthquakes per day found for the entire period 1932–1999 in the same region. Furthermore, of the 2315 earthquakes, 1419 (or 61.3%) were aftershocks, which is in good agreement with the 60% of aftershocks estimated from the model for the period 1942–1999. For a comparison with observed data, the corresponding data from the period 1990–1999 are shown to the left in Figs. 2 and 3. Except for the fact that the period 1990–1999 is a period of higher than average intensity, we see from Fig. 2 that the simulation produces earthquakes with a similar spatial distribution as the observed data. In addition it can be seen that the simulated catalog reflects features in the 1932–1999 data that are not present in the 1990–1999 data. The simulation results also display aftershock activity indicated by event sequences

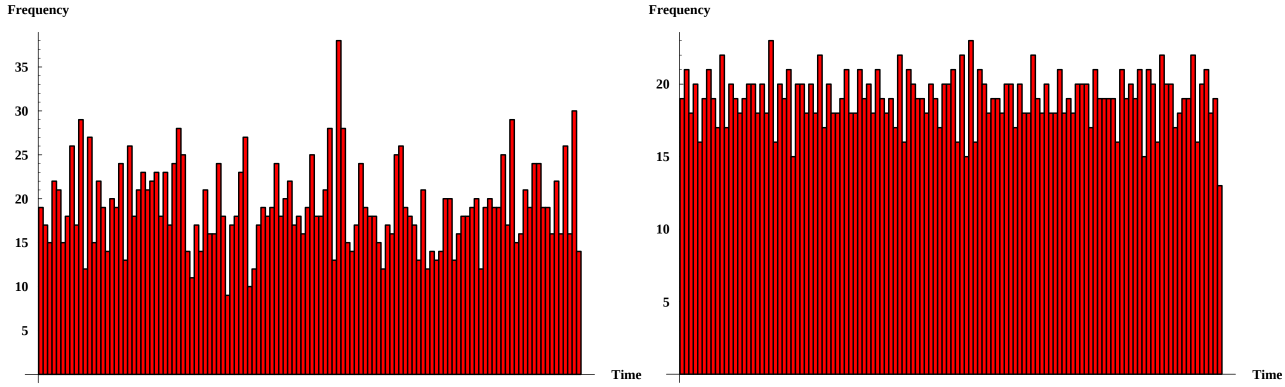

Fig. 5. The frequency of earthquakes in intervals of 30 days for a given simulation. The left figure shows the frequency when the parameter $\beta_6 = 0.000470$, while to the right is shown the frequency when $\beta_6 = 0.470$.

following the larger earthquakes as shown in Fig. 3 (right), although this effect is not as pronounced as for the observed data in Fig. 3 (left).

Simulations where the build-up and release of strain are taken into account produce results that are similar to those displayed in Figs. 2 and 3. This is not so surprising since we do not expect to see any real periodicity pattern on the occurrence of large earthquakes over the time-span of only 10 years. The return period for the largest earthquakes in this region is much longer than this. In Fig. 4, however, we have shown the typical behavior of the strain function $s(x, t|H_t, \beta)$ during a given simulation. The sudden drops in the graph correspond to the release of strain due to the aftershocks following a large earthquake. To demonstrate the pure effect of the strain function $s(x, t|H_t, \beta)$ in our model, we have done simulations where the aftershocks treatment have been omitted by setting $\beta_1 = 0$, effectively giving $\lambda_3 = 0$. The other parameters are left unchanged, except that in the first simulation we set $\beta_6 = 0.000470$ (unchanged), while in the second simulation we take β_6 to be three orders of magnitude higher. The effect of varying β_6 in this manner is shown in Fig. 5 where the graphs show the frequency of earthquakes for each 30-day period of the simulation period. A greater value of β_6 yields a more even release of strain in each time interval, and hence a more even frequency of earthquakes in each 30-day period. The model is not particularly sensitive to the value of β_6 since it was necessary to multiply β_6 with a factor 10^3 to achieve a significant difference. However, an improved estimate of the parameter β_8 (or β_7 in the expression (10)) may imply a greater sensitivity to β_6.

4 Concluding remarks

In conclusion, we have by maximum likelihood optimization estimated the parameters β of the marked point process model, using earthquake data from Southern California from the period 1932–1999. The first five parameters $\beta_1, \ldots, \beta_5$ are connected to the term λ_3 and represent the aftershock activity. The estimated values for these parameters seem plausible. For instance, we may compare the results $\beta_3 = 1.24$ and $\beta_4 = 0.0191$ with the corresponding estimates found by Ogata (1998) in various extensions of his Epidemic Type Aftershock-Sequences model. Applied to data from two different districts of Japan, Ogata estimated a parameter p (corresponding to β_3 in our model) to be in the range 0.900–1.136 and a parameter c (corresponding to β_4) to be in the range 0.00172–0.0357. The last three parameters $\beta_6, \beta_7, \beta_8$ are connected to the factor s, which represents the build-up and release of strain in the model. Surprisingly, we found both β_7 and β_8 to vanish in the maximum likelihood estimation. The vanishing of β_7 suggests that the release of strain locally is independent of the magnitude of the earthquakes. The release of total strain energy, however, does depend on the size of the earthquake since a large earthquake causes release of strain over a much larger area than a small earthquake. Because of this, the parameter β_7 seems superfluous in the model and could be omitted by a slightly different representation of the strain release, as indicated in the previous section. The vanishing of β_8 in our estimation is difficult to interpret physically, however. This implies that all earthquakes in the catalog have a global effect on the release of strain for the region considered. This is not what we would have expected and suggests that the given parametrization is too crude to handle the spatial dependencies between the release of strain and the individual earthquakes. One explanation for this may simply be that we have used epicentral distances instead of the distances to the rupture plane. We are hoping to incorporate distances to the rupture plane in future work, together with the inclusion of more geological information (such as extended sources) in the model. The present set-up does facilitate such extensions.

An obvious problem in the current estimation, however, is that the catalog time 1932–1999 is very short compared to the recurrence times of larger earthquakes in California. From Fig. 1 (right) we see that the Southern California Catalog is reasonably homogeneous back to the early 1940's. For the initial testing of our model we have therefore chosen to use the period 1942–1999 for calculating the maximum likelihood estimator for the parameters β. When the build-up and release of strain is omitted, this choice does not seem to impose any kind of problems. In fact, the values obtained for $\beta_1, \ldots, \beta_5$ are in agreement both with what we expected and with previous work (e.g. Lay and Wallace, 1995; Ogata, 1998). But when the build-up of strain is included and the parameters $\beta_6, \beta_7, \beta_8$ are estimated, the shortness of the catalog may very well impose difficulties. The next step for testing and further development of the model is therefore clearly to extend the test catalog with data prior to 1932. However, since the quality of data gets poorer the further back we go in time, such an extension is likely to introduce some new problems, albeit most likely not any serious ones. With an assumption that the intensity of the background activity prior to 1932 was the same as for the period 1932–1999, it would be straightforward to simulate a background intensity for earlier times with our simulation algorithm. This background intensity can then be used to complement data sets that originally are incomplete. We believe that this method will lead to improved estimates for the parameters $\beta_6, \beta_7, \beta_8$, or for β_6, β_7 if (10) is used. The key point is that our marked point process model is very flexible to changes, and that the results presented in this paper follow from initial testing of the algorithms involved. More realistic exercises will follow later, including, as already noted, the use of extended sources.

Acknowledgements. We are grateful to B. Natvig and J. Gåsemyr for a number of constructive comments and suggestions.

References

Cressie, N.: Statistics for spatial data, John Wiley & Sons Inc., New York, 1993.

Kagan, Y. Y.: Are earthquakes predictable?, Geophys. J. Int., 131, 505–525, 1997.

Kagan, Y. Y.: Universality of the seismic moment-frequency relation, Pure Appl. Geophys., 155, 537–573, 1999.

Lay, T. and Wallace, T. C.: Modern Global Seismology, Academic Press, 521 pp., San Diego, 1995.

Main, I.: Apparent breaks in scaling in the earthquake cumulative frequency-magnitude distribution: Fact or artifact?, Bull. Seism. Soc. Am., 90, 86–97, 2000.

Ogata, Y.: Space-time point-process models for earthquake occurrences, Ann. Inst. Statist. Math., 50, 379–402, 1998.

Ripley, B. D.: Stochastic simulation, John Wiley & Sons Inc., New York, 1987.

SCEC: Southern California Earthquake Center, Catalog of Southern California earthquakes publicly available at http://www.scecdc.scec.org/ftp/catalogs/SCSN/, 2000.

Scholz, C. H.: Scaling laws for large earthquakes: consequences for physical models, Bull. Seism. Soc. Am., 72, 1–14, 1982.

Sykes, L. R., Shaw, B. E., and Scholz, C. H.: Rethinking earthquake prediction, Pure Appl. Geophys., 155, 207–232, 1999.

Vere-Jones, D.: Forecasting earthquakes and earthquake risk, Int. J. of Forecasting, 11, 503–538, 1995.

Working Group on California Earthquake Probabilities: Seismic Hazard in Southern California: Probable earthquakes, 1994 to 2024, Bull. Seism. Soc. Am., 85, 379–439, 1995.