Project No 6OPR1B
Towards a New Zealand model for short-term earthquake probability: Aftershock productivity and parameters from global catalogue analysis ("Catalogue Analysis")

June 2006

Annemarie Christophersen

Technical Abstract

The project 'Towards a New Zealand model for short-term earthquake probability: Aftershock productivity and parameters from global catalogue analysis ("Catalogue Analysis")' had two main objectives: First to study the occurrence of foreshocks to extend an existing aftershock model into a model for short-term earthquake probabilities and second to include the new model into the Californian 'Regional Likelihood model' (RELM) testing.

The question whether foreshocks behave like mainshocks that happen to have larger aftershocks is important for the understanding of earthquake nucleation and the modelling of earthquake clustering. A commonly used model for short-term earthquake occurrence, the Reasenberg-Jones model, combines empirical observations of aftershock occurrence and extends them to potential foreshocks assuming a single triggering mechanism for earthquake clusters. However, observed foreshock probabilities have been reported significantly lower than predicted by this kind of model.

To investigate the discrepancy between the model prediction and the observations we reviewed an earlier description of aftershock rates. We derived a variation of the Reasenberg-Jones model which was based on the mean abundance, i.e. the average number of aftershocks per sequence in a chosen time and mainshock magnitude interval. We calculated model parameters from the Californian catalogue by first defining earthquake sequences and secondly analyzing aftershock occurrence. The new model formulation predicted foreshock numbers that were consistent with the data. We concluded that no separate triggering mechanism was required to derive foreshock occurrence from aftershock models and thus foreshocks could be regarded as mainshocks with larger aftershocks. Finally, we investigated the spatial and temporal separation between foreshocks and their mainshocks and found that the distribution of foreshock-mainshock pairs had similar spatial and temporal distribution as mainshock-aftershocks.

We set up a model for the occurrence of aftershocks in a chosen time and magnitude interval that can easily be incorporated into the existing model for short-term earthquake probabilities (STEP) in California. The model was based on the mean number of aftershocks, also referred to as mean abundance in a time interval for which aftershocks have been completely recorded. For California, the abundance model was derived for mainshocks in the magnitude interval 2.95 – 5.95 and included aftershocks above magnitude 2.45 in the time interval 0.1 – 10 days. While the mean abundance was well behaved as a function of mainshock magnitude, the abundance for any given earthquake sequence was found to vary significantly. Therefore we proposed to up-date model parameters in the STEP model hourly and not just when more than 100 earthquakes have been observed.

Our model has not been incorporated in the RELM testing because the testing of short-term earthquake models has been delayed. However, efforts in respect to earthquake model development and testing have gained momentum over the last few years, with a predecessor to RELM in the US, the Collaboratory for the Study of Earthquake Predictability (CSEP), a couple of European projects that include earthquake hazard modelling and testing and an EQC funded project with similar aims.

Overall the objectives of the project have been achieved.

Layman's Abstract

The project 'Towards a New Zealand model for short-term earthquake probability: Aftershock productivity and parameters from global catalogue analysis ("Catalogue Analysis")' had two main objectives: First to study the occurrence of foreshocks to extend an existing aftershock model into a model for short-term earthquake probabilities and second to include the new model into the Californian 'Regional Likelihood model' (RELM) testing.

The question whether foreshocks behave like mainshocks that happen to have larger aftershocks is important for the understanding of earthquake nucleation and the modelling of earthquake clustering. A commonly used model for short-term earthquake occurrence, the Reasenberg-Jones model, combines empirical observations of aftershock occurrence and extends them to potential foreshocks assuming a single triggering mechanism for earthquake clusters. However, observed foreshock probabilities have been reported significantly lower than predicted by this kind of model.

To investigate the discrepancy between the model prediction and the observations we reviewed an earlier description of aftershock rates. We derived a variation of the Reasenberg-Jones model which was based on the mean abundance, i.e. the average number of aftershocks per sequence in a chosen time and mainshock magnitude interval. We calculated model parameters from the Californian catalogue by first defining earthquake sequences and secondly analyzing aftershock occurrence. The new model formulation predicted foreshock numbers that were consistent with the data. We concluded that no separate triggering mechanism was required to derive foreshock occurrence from aftershock models and thus foreshocks could be regarded as mainshocks with larger aftershocks. Finally, we investigated the spatial and temporal separation between foreshocks and their mainshocks and found that the distribution of foreshock-mainshock pairs had similar spatial and temporal distribution as mainshock-aftershocks.

We set up a model for the occurrence of aftershocks in a chosen time and magnitude interval that can easily be incorporated into the existing model for short-term earthquake probabilities (STEP) in California. The model is ready to be incorporated in a model testing centre that will soon start for California.

Overall the objectives of the project have been achieved.

Contents

1 Introduction

This is the final report for the project No 6OPR1B 'Towards a New Zealand model for short-term earthquake probability: Aftershock productivity and parameters from global catalogue analysis ("Catalogue Analysis")'. The work follows on from my PhD study and has been closely linked to other EQC funded projects as further outlined below. The introduction reviews the context of the work, including the RELM testing and further international developments on earthquake hazard models and testing. I provide an overview about the original objectives and discuss in how far they have been achieved. The results are reported in form of the two included manuscripts 'Foreshock probabilities from aftershock characteristics' and 'Aftershock abundance: Forecasting aftershock occurrence when catalog completeness is low'. The first manuscript has been submitted to the Bulletin of the Seismological Society of America. It was completed with funding from the EQC project 'Testing and development of earthquake forecasting models' as an extension of the earlier project. The second manuscript is still in a draft stage. However, the results that relate to project 6OPR1B are complete.

1.1 Background

For my PhD, I studied 'the probability of a damaging earthquake following a damaging earthquake (Christophersen, 2000). I established a global earthquake catalogue, which I divided into tectonic regions. I defined earthquake sequences and analysed aftershock behaviour. I then set up a model for the probability of aftershocks which was based on the number of aftershocks, from here on called abundance. Due to time constraints I had no chance to analyse the foreshock occurrence or implement my model during my thesis. The purpose of project No 6OPR1B was to study foreshock occurrence in the global database, extend the aftershock model to allow for larger earthquakes to happen later in an on-going sequence and to collaborate with other New Zealand researchers with view of implementing a short-term hazard model in New Zealand.

1.2 Linkages to other projects

Project 6OPR1B has had a companion project, project No 6OPR1A, 'Evaluating New Zealand Models for Sort-term Earthquake Probability'. At the time of proposal writing Thessa Tormann was a diploma student working with Dr Martha Savage at Victoria University on incorporating foreshock probabilities into the New Zealand earthquake hazard maps and looking to undertake a Master of Science project. Thessa and Martha worked closely with Dr Mark Stirling. We started a little working group, also including Professor Euan Smith and Dr Matt Gerstenberger who had just returned to New Zealand around that time in November 2003. We put our proposals forward with view to carry out the work in parallel. However, there were some initial delays with the funding agreement and contract negotiations and Thessa's project shifted slightly towards the development of a methodology to compare predicted earthquake probabilities with strong motion recordings (Tormann, 2005). However, we still have kept close communication throughout the process of both projects.

As outlined above, project No 6OPR1B is an extension of my PhD thesis. A large component of the work is about the development of the earthquake catalogue and defining aftershock sequences.

Therefore the results have also fed into the EQC funded project Euan Smith 'Triggered earthquake probability forecasting', which was officially completely in November 2004 (Christophersen and Smith, 2004). However, my co-worker Professor Euan Smith has continued work on the modelling of all inter event times between earthquakes in a catalogue during his sabbatical in 2005. Recently he has also started to use the Californian data that I prepared during this project. We have drafted a substantial manuscript for publication (Smith and Christophersen, 2006) and have started on a physical model to explain the power law decay of aftershock activity with time.

Our initial results of the foreshock analysis as reported in June last year, led us to believe that there was a significant difference in foreshock and aftershock behaviour (Christophersen, 2005a). Therefore we put forward a proposal on a more detailed study of foreshock occurrence. The proposal was funded as component of the collaborative project 'Testing and development of earthquake forecasting models'. The work on foreshocks has already been completed and as already mentioned above, the results are included in this report.

1.3 The RELM testing

The RELM (Regional Earthquake Likelihood Models) project was the first comprehensive attempt to test earthquake forecasts against data and different forecasts against one another. The project had three main goals :

1. Develop a variety of viable, geophysically based earthquake-rupture forecast models for California
2. Test these models for consistency with existing geophysical data and design and document conclusive tests with respect to future observations.
3. Examine and compare the implications of each model with respect to seismic hazard and loss estimate.

A number of models have been developed and a special issue of Seismological Research Letters has been dedicated to them but is not yet published. Two types of models are distinguished: Short-term models with earthquake rates varying from day to day and quasi-stationary models assuming earthquake rates to be relatively stable over about a year.

A testing methodology has also been developed (Schorlemmer et al, 2005, and Schorlemmer Gerstenberger, 2005) and was meant to be implemented during 2005. The testing for the quasi-stationary models began in January 2005. However, the testing centre for the short-term models still has to be implemented.

1.4 International developments

The past few years have seen an increase in interest for short-term earthquake hazard models and model testing. The United States Geological Survey (USGS) began publishing daily online maps of short term earthquake probabilities (STEP) for California in May 2005 (Gerstenberger, et al., 2005 and http://pasadena.wr.usgs.gov/step/). Also in the US, the RELM project has expanded into the 'Collaboratory for the Study of Earthquake Predictability' (CSEP) with significant funding for the development of earthquake models and their testing (Schorlemmer et al., 2006). In the European Union, two new projects have started, NERIES and SAFER, which both include components for the development of tools for seismic hazard assessment. Testing of models will also be addressed. In New Zealand, EQC has funded the collaborative project 'Testing and development of earthquake forecasting models' which will replicate the RELM procedure with some New Zealand specific variations. Quests for further funding in New Zealand are underway to strengthen the collaboration with international partners on this topic. Results from project No 6OPR1B will be applied internationally.

1.5 Overview of proposal and work completed

Originally I had identified four objectives with a number of sub-objectives each. I list them below and review the progress on each of them.

1. Global foreshock occurrence

1.1 Establish retrospective global foreshock occurrence by tectonic region, i.e. what is the probability that a mainshock was preceded by a foreshock

1.2 Establish prospective foreshock occurrence, i.e. what is the probability that an earthquake will be followed by a larger one (outside an aftershock sequence)

1.3 Compare foreshock probabilities of different regions, in particular California and New Zealand

1.4 Analyse time between foreshock and mainshock

2. Global model for short-term conditional probabilities of earthquake occurrence

2.1 Test whether global data are consistent with the model that foreshock are part of the same phenomenon as aftershocks

2.2 Extend model from PhD study to include foreshocks

2.3 Compare aftershock productivity in new model to Californian model

3. Testing of abundance component in RELM

3.1 Determine abundance for California

3.2 Incorporate abundance into Californian model

3.3 Test new Californian model within RELM

4. New Zealand model for short-term conditional probabilities of earthquake occurrence

4.1 Provide baseline parameters for Gerstenberger's model

4.2 Provide alternative description of aftershock activity

4.3 Assist with model testing

Results on point 1, the global foreshock occurrence were reported in June 2005. However, as realised later, the original search windows in time were too big and included significant background seismicity. The background seismicity gave the appearance that foreshocks occurred randomly in space and thus behaved quite differently to aftershocks. Results have been up-dated. For California, they are reported in the included manuscript 'Foreshock probabilities from aftershock characteristics'.

The foreshock manuscript also reports on our results for point 2 that foreshock occurrence is consistent with mainshocks having larger aftershocks. As a consequence no specific addition to the model from the PhD study is required to include foreshock behaviour. However, we had reported on a model framework for foreshock occurrence in October 2005 when we still thought that foreshocks showed different behaviour to aftershocks and thus needed separate treatment in short-term earthquake modelling. A comprehensive comparison between aftershock productivity in the new model and the Californian STEP model is undertaken in the second manuscript 'Aftershock abundance: Forecasting aftershock occurrence when catalog completeness is low'.

We determined the abundance model for California. Results from using the catalogue of the Advanced National Seismic Systems (ANSS) are included in the foreshock manuscript and results from our work with the Southern California Earthquake Data Center (SCEDC) are reported in the abundance manuscript. We are ready to incorporate the abundance model into the STEP model. However, as the RELM testing has been delayed, work on the models is on-going.

Point 4 was originally intended to occur parallel with project No 6OPR1A. In particular point 4.3 was allowing time to assist with the development of Thessa's model testing procedure. Point 4.1 and 4.2 were both completed for the Californian data and are reported on in the included manuscript.

1.6 Conclusions

In summary, the project had two main objectives: First, to study foreshock occurrence to test whether foreshocks behaved like mainshocks that happen to have larger aftershock and second to work towards the implementation of a New Zealand short-term earthquake model by way of setting up a model for California and include it in a framework for model testing. The project also aimed at fostering the collaboration between Victoria University and GNS Science and building international relationships in this exciting area of research.

The foreshock analysis took a brief detour by including background seismicity in the analysis and thus concluding that foreshock occurred more randomly in space and over longer time periods than aftershocks. Adjusting the search parameters for setting up the initial earthquake sequences led to the finding that mainshocks relate to their foreshocks in space and time as aftershocks to their mainshocks. In collaboration with Euan Smith and Matt Gerstenberger I set up a model for the number of earthquake following a mainshock in a chosen time and magnitude interval. The model can easily be incorporated into the STEP model to predict ground shaking. As the RELM model testing has been delayed, we have not yet set up the final model(s) for testing.

During the course of the project I had the opportunity to travel to the US twice, including attending the American Geophysical Union conference in December 2004. I also had the chance to visit Zurich and I was invited and funded by Professor Yosi Ogata to attend the fourth workshop on Statistical Seismology in Japan in January this year. The interactions throughout my travels have ultimately led to my appointment with Dr Stefan Wiemer at the Swiss Federal Institute for Technology. I am grateful to EQC for opening up such wonderful opportunities.

2 Foreshock probabilities from aftershock characteristics

Manuscript by Annemarie Christophersen and Euan Smith, submitted to the Bulletin of the Seismological Society of America

2.1 Abstract

The question whether foreshocks behave like mainshocks that happen to have larger aftershocks is important for the understanding of earthquake nucleation and the modeling of earthquake clustering. A commonly used model for short-term earthquake occurrence, the Reasenberg-Jones model, combines empirical observations of aftershock occurrence and extends them to potential foreshocks assuming a single triggering mechanism for earthquake clusters. However, observed foreshock probabilities have been reported significantly lower than predicted by this kind of model. To investigate the discrepancy between the model prediction and the observations we review an earlier description of aftershock rates. We also present a variation of the Reasenberg-Jones model which is based on the mean abundance, i.e. the average number of aftershocks per sequence in a chosen time and mainshock magnitude interval. We derive model parameters from the Californian catalog by first defining earthquake sequences and secondly analyzing aftershock occurrence. The new model formulation predicts foreshock numbers that are consistent with the data. We conclude that no separate triggering mechanism is required to derive foreshock occurrence from aftershock models and thus foreshocks can be regarded as mainshocks with larger aftershocks. Finally, we investigated the spatial and temporal separation between foreshocks and their mainshocks and found that the distribution of foreshock-mainshock pairs had the same spatial and temporal distribution as mainshock-aftershocks.

2.2 Introduction

Earthquakes cluster in space and time. In retrospect, the largest earthquake in a cluster is called the mainshock. The earthquakes preceding a mainshock are named foreshocks, and the earthquakes following a mainshock aftershocks. Two earthquakes of similar size are sometimes referred to as doublets. No physical differences between any of these earthquakes are known. Some researchers argue that a single triggering mechanism is responsible for foreshock-mainshock-aftershock occurrence (e.g. Felzer et al., 2004). In that sense a foreshock could be regarded as a mainshock that happened to have a larger aftershock (Jones et al., 1999). However, applying empirical relations derived from aftershock occurrence to the prediction of foreshocks overestimates the occurrence of foreshock by nearly a factor of two (Jones et al., 1999).

The epidemic type aftershock model (ETAS) describes complete earthquake catalogs as one family of earthquakes (e.g. Console et al. 2003, Zhuang et al., 2004). The model distinguishes between triggered earthquakes and background seismicity. Foreshocks are modeled as part of the background, which has a lower triggering capability than triggered earthquakes (Zhuang et al., 2006). This modeling approach is consistent with the observation that empirical relations from aftershock statistics overpredict foreshock rates but contradicts the idea of a single consistent triggering mechanism for all clustered earthquakes.

Understanding possible differences between foreshocks and mainshocks would help understanding the process of earthquake nucleation and allow better modeling of short-term earthquake hazard.

To investigate the discrepancy between model prediction and foreshock observations we first review the most commonly used empirical models for earthquake clustering. We study a model for aftershock numbers and adjust it to predict foreshock occurrence. We then define earthquake sequences for a Californian earthquake catalog. We analyze some cluster characteristics and show that the prediction of the new model for an earthquake to be a foreshock is consistent with the observation of foreshock numbers. We also investigate the differences in time and space between foreshocks and their mainshocks and find them to be consistent with mainshock-aftershock behavior.

2.3 Empirical relations for aftershock occurrence

Aftershocks are generally far more numerous than foreshocks and therefore most relationships for short term earthquake occurrence have been derived for aftershocks. Utsu and Seki (1955) studied 40 Japanese mainshock-aftershock sequences in the mainshock magnitude range $6 \leq M \leq 8.5$ and found that the aftershock area increased with mainshock magnitude as

$$A(M) = 10^{1.02M-4.02}, \qquad (1)$$

which Utsu later simplified to $A(M) = 10^{M-3.9}$ (Utsu, 2002). To define aftershocks, researchers often draw a box around the mainshock fault ruptures using scaling relations (e.g. Wells and Coppersmith, 1994). Unfortunately not all researchers clearly define what they mean by foreshocks or aftershocks.

The Omori-Utsu law describes the decay of aftershock activity with time t.

$$dN/dt = K/(t_c + t)^p, \qquad (2)$$

where dN is the number of earthquakes in the time interval dt; K is a parameter that is proportional to the aftershock productivity; p describes the decay and takes values around 1.0; and t_c stands for a small time interval just after the mainshock (e.g. Utsu et al., 1995).

The Gutenberg-Richter relation describes the magnitude-frequency distribution

$$n(M) = 10^{a-bM}, \qquad (3)$$

where $n(M)$ is the number of earthquakes of magnitude M and a and b are parameters (e.g. Gutenberg and Richter, 1949).

The average number of aftershocks $N(M)$ for mainshock magnitude M can be described as

$$N(M) = 10^{c(M-M_1)}, \quad (4)$$

where the parameter M_1 corresponds to the mainshock magnitude that on average has one aftershock above the completeness magnitude of the catalog (Christophersen, 2000). The value *c*, in many studies referred to as *α*, corresponds to the triggering capability of differently sized mainshocks. In a global study, the value *c* was found to be 1.0 for most tectonic settings (Christophersen, 2000). For California it was observed to be 0.8 (Helmstetter, 2003), for Italy 0.5 (Console et al., 2003) and for Japan 0.8 on average with values varying between 0.2 and 1.9 (Guo and Ogata, 1997). The *c*-value is sometimes assumed to be equal to the *b*-value in the Gutenberg-Richter relation (e.g. Felzer et al., 2004). This assumption is consistent with another empirical law, Båth's law, which states that the average magnitude difference between a mainshock and its largest aftershock is independent of mainshock magnitude and takes values of about 1.2 (e.g. Richter, 1958). It has remained controversial whether Båth's law has a physical origin or simply is a consequence of other statistical features of aftershock occurrence. However, Vere-Jones, et al. (2006) showed that Båth's law can be derived from the scaling laws for the magnitude distribution and the average number of aftershocks.

Reasenberg and Jones (1989, 1994) combined equations (2) and (3) to determine the rate of aftershocks of magnitude M and above, at time t following a mainshock of magnitude Mm:

$$R(t, M_m) = 10^{a' + b(Mm - M)} / (t + t_c)^p \quad (5)$$

Reasenberg and Jones analyzed 62 Californian earthquake sequences and derived the parameters a' = -1.67, p = 1.08, b = 0.91 and t_c = 0.05. These parameters have become known as the generic Californian aftershock model parameters (e.g. Gerstenberger et al., 2004). The rate can be used to calculate the probability that at least one earthquake of magnitude M or above occurs in the time interval $[t_1, t_2]$ as follows:

$$P(M) = 1 - \exp \int_{t_1}^{t_2} R(t, M)\, dt \quad (6)$$

Equation (6) yields a probability of 10.5% that an earthquake is followed by one of the same magnitude or larger within one week but only 6.0 ± 0.5% of earthquakes in southern California are foreshocks, i.e. they were of magnitude $M \geq 3.0$ and followed by a larger earthquake within 5 days and 10 km (Jones, 1985).

The Reasenberg-Jones model was applied to 17 earthquake sequences in New Zealand where a wide variation in model parameters were found (Eberhart-Phillips, 1998). From this model, Savage and Rupp (2000) determined foreshock probabilities and compared them to observed foreshock rates. The results are given in Table 1. As Jones et al. (1999) reported for California and Italy, the observed occurrence of foreshocks in New Zealand was lower than the mean values predicted by the model. However, the range of model probabilities was very large and spanned the observations. Thus model and observations were possibly consistent.

Utsu (1969) had used the same rate equation for Japanese earthquakes with the parameters a' = -1.83, p = 1.3, b = 0.85 and t_c = 0.3. Using these rates in equation (5), the probability that an earthquake is followed by one of the same magnitude or larger within one week can be calculated to be 4.2%. This is significantly smaller than the result from the Californian generic model and agrees reasonably well with observations from other studies already discussed. To investigate the potential difference between foreshock occurrence and predictions made from aftershock models, we studied a model for aftershock numbers, applied it to the Californian catalog and compared foreshock predictions to observations.

2.4 Foreshock and aftershock numbers

Our approach is similar to that of Felzer et al. (2004). They proposed a single triggering mechanism for fore-, main-, and aftershocks and thus derived that numbers of aftershocks, foreshocks and doublets should vary linearly with each other. For example, an area with fewer aftershocks should also have fewer doublets and fewer foreshocks. In addition they asserted the validity of equation (4) for the growth of aftershocks with mainshock magnitude and combined this with the Gutenberg-Richter relation (equation 3) to predict the number of earthquakes above a chosen magnitude M_2 following an earthquake of magnitude M_1. Felzer et al. (2004) assumed that $b = c = 1.0$. In contrast we determine b and c independently from aftershock data and then allow that one of the triggered earthquakes has a larger magnitude than the initiating one. For a more detailed description of the model see Vere-Jones et al. (2006).

Assuming independence of abundance and the magnitude-frequency relation, the expected number of aftershocks of magnitude M and larger, following a mainshock of magnitude M_m in a chosen time interval $[t_1, t_2]$ can be written as

$$N(M) = 10^{\,c(Mm-M1)}\, 10^{\,-b\,(M-Mo)} \quad (7)$$

Assuming that the initial earthquake Mm was to be followed by a larger earthquake, we set $M = M_m$ and get the expected number of aftershocks of magnitude equal or larger than the magnitude of the initiating earthquake:

$$N(M_m) = 10^{\,(c-b)Mm}\, 10^{\,b\,Mo-\,cM1} \quad (8)$$

For foreshocks N will be considerably smaller than 1, e.g. 0.06 for the rate of foreshock occurrence in California (Jones, 1985). For $c = b$, there is no dependence on the magnitude of the initiating earthquake, i.e. the expected number of larger earthquakes to follow an initiating event is the same for all initiating earthquake magnitudes. For $c < b$, larger earthquakes are less likely to be followed by even larger ones. For $c > b$, smaller earthquakes are less likely to be followed by larger ones. Once we have derived a model for the number of triggered earthquakes in one time interval $[t_1, t_2]$, we can use Omori's law to derive the expected number N_t of triggered earthquakes in another time interval $[t_3, t_4]$. For this purpose we define the fraction γ of the total number of triggered earthquakes by $N_t = \gamma\, N_m$. The Omori-Utsu law (equation 2) can be integrated to determine γ:

$$\gamma = \ln\,[(t_4 + t_c)/\,(t_3 + t_c)]\,/\,\ln\,[(t_2 + t_c)/\,(t_1 + t_c)] \qquad \text{for } p = 1 \quad (9)$$

$$\gamma = [(t_4 + t_c)^{\,1-p} - (t_3 + t_c)^{1-p}]\,/\,[(t_2 + t_c)^{\,1-p} - (t_1 + t_c)^{1-p}] \qquad \text{for } p \neq 1 \quad (10)$$

In the following section we outline what data we used to derive the model parameters c, M_l, b, and p.

2.5 The earthquake catalog

For our analysis we used the Californian catalog of the ANSS (Advanced National Seismic Systems) for 1984-2004 between latitude 31-43 North and longitude 127-112 East (ANSS, 2005). Data were restricted to events with depths shallower than 40 km. The completeness magnitude outside an on-going earthquake sequence can be as low as 1.2 (Woessner and Wiemer, 2005). However, as we are mainly interested in earthquake clusters which can have a much higher completeness magnitude, especially immediately following a large earthquake, we initially restricted the data to events above magnitude 2.45. The catalog included about 51,000 earthquakes. Figure 1 shows the magnitude-frequency distribution of all 813,000 earthquakes from magnitude 0 and above. As one way to investigate the completeness magnitude and to find potential problems with the magnitude scale, we calculated the b-value in the Gutenberg-Richter relation (equation 3) using Aki's (1965) maximum likelihood method and also applying a correction for the limitation of the magnitude interval at the top and the magnitude bin width of 0.1.

$$b = 1/(\ln(10) *(M_{ave} - cor) \qquad (12)$$

$$cor = [(M_u - 0.05)\, e^{\beta (Mu - 0.05)} - (M_l - 0.05)\, e^{\beta (Ml - 0.05)}] / [e^{\beta (Mu - 0.05)} - e^{\beta (Ml - 0.05)}] \qquad (13)$$

where $\beta = \ln(10)\, b$ and M_u and M_l are the upper and lower limits of the magnitude interval (Christophersen, 2000). The results including 95% confidence intervals are shown in figure 2.

2.6 Defining earthquake clusters

As already stated in the introduction, no physical differences are known to exist between foreshocks, mainshocks and aftershocks and therefore fore- and aftershocks are usually defined by their closeness in space and time to a larger earthquake, the mainshock. To study fore- and aftershock occurrence we first looked for an objective method to define earthquake sequences that allows for differences between sequences, e.g. in epicenter distribution or sequence duration. We applied a two step process that was developed to define earthquake sequences for a global earthquake catalog (Christophersen, 2000). First we set up a data set of initial clusters: Every earthquake in the catalog above a chosen magnitude qualified as potential mainshock. We searched for earthquakes close in time and space in a magnitude dependent search radius and a sliding time window. If a larger earthquake was found in the specified time window the distance in space was compared to the search radius of the larger earthquake. If the larger earthquake was part of the on-going cluster, its epicenter coordinates and search radius were used to find subsequent earthquakes. For each associated event, the search window in time was extended by 10 days for completeness magnitude 2.5. Different time windows that depended on the completeness magnitude were also trialed. For each initial cluster with at least three earthquakes, the scatter of epicenters was fitted by a bivariate normal distribution with center, orientation of principal axes and lengths of axes determined by least squares. Earthquakes within the 90 percentile ellipse were counted as foreshocks or aftershocks.

Events outside the 90 percentile ellipse were regarded as background (Christophersen, 2000). However, in some cases the largest earthquake occurred outside the ellipse. In those instances the largest earthquake was still counted as mainshock rather than selecting the largest earthquake from within the ellipse as mainshock.

Search windows in space

The initial search radius r is important and can depend on the catalog. We trialed various search radii and compare some of them in figure 3. The Utsu and Seki formula (equation 1) describes how aftershock area increases with mainshock magnitude M. Assuming a circular aftershock area, we derived a search radius by taking four times the circle radius. We also tried half and double this search radius. The Gardner and Knopoff (1974) and Uhrhammer (2005) search windows were both developed to strip aftershocks from earthquake catalogs. Uhrhammer used windows of 20, 30, 45, 67 and 100 km to remove aftershocks following earthquake of magnitude 5, 5.5, 6, 6.5 and 7, respectively. Gardner and Knopoff's windows are 22.5, 30, 40, 54, 70, and 94 km for magnitude 3, 4, 5, 6, 7, and 8, respectively. We fitted those windows by smooth functions and extrapolated the Uhrhammer windows to smaller magnitudes:

$$\text{Gardner-Knopoff:} \quad r = 10^{\,0.12\,M + 0.98} \tag{14}$$

$$\text{Uhrhammer:} \quad r = 10^{\,0.35\,M - 0.44} \tag{15}$$

$$\text{Utsu \& Seki:} \quad r = 10^{\,0.51\,M - 1.65} \tag{16}$$

Evaluating search windows

We evaluated the different search windows by plotting the earthquake epicenters of each cluster on a map centered on the fitted ellipse and with a half-width of the search radius. Figure 4 gives an example of one earthquake of magnitude 4.1, marked by a star, with search results from three different search radii. Figure 4a used the Gardner-Knopoff window which is 31 km for magnitude 4.1. The search radius is too large. Seismicity outside the dense cluster was picked up in the search and therefore the shape of the ellipse is not a good fit to the overall distribution of epicenters. The ellipse takes up a relative small part of the overall map. This is another indication that the search radius was too big. Figure 4b used the Uhrhammer extrapolated window size of 10 km for magnitude 4.1. The ellipse is a good fit to the main cluster and the ellipse occupies a good portion of the map area. In this case the three events outside the fitted ellipse would not have been counted in the aftershock analysis. Figure 4c used the Utsu-Seki search radius of 2.8 km for magnitude 4.1. The epicenters scatter relatively widely and the ellipse takes up a significant proportion of the map area. These observations are an indication that the search radius was too small to find the complete cluster. We visually inspected all fitted ellipses for various search windows and found that Uhrhammer's modified windows worked best as initial search radius.

Search windows in time

In time we used sliding windows that extended the search period by 5, 10, 15 and 30 days following each newly associated earthquake. For the completeness magnitude of 2.5 and an extending time window of 10 days, the resulting sequences had durations varying from less than one day to just under 1000 days. For sliding windows of 30 days, sequence durations of more than 5 years were obtained. If the window length was reduced from 30 to 10 days, the number of mainshocks larger than magnitude 5.0 increased from 116 to 128. The 8 'missing' earthquakes occurred with their own aftershock sequences in the tail of other larger earthquake sequences. It seemed more sensible to use a smaller window length that excluded them and analyze them separately. Thus we chose 10 days to extend the sequence for every newly associated earthquake.

2.7 Analysing earthquake clusters

Our aim was to use equation 8 to determine the number of aftershocks in a cluster, calculate the expected number of foreshocks and compare this to the observed foreshock occurrence.

Completeness considerations

We first investigated the completeness of aftershocks in the early part of the sequence. As thoroughly discussed by Kagan (2004), the completeness magnitude sharply increases following a large earthquake. Earthquakes of smaller events cannot be identified in the coda of the mainshock for a period that depends on the size of the earthquake. Limited resources can also lead to a delay in processing abundant earthquake sequences. Figure 5 is a scatter plot of aftershock magnitudes and times of occurrence in days after the Landers magnitude 7.3 earthquake. There are hardly any earthquakes of magnitude 4.5 or below in the first 0.1 days and there are significant gaps between magnitude 2.5 and 3 for all of the first day. Figure 6 shows the percentage of all aftershocks in the first 10 days following a mainshock that occurred in the first 0.1 and 1 day. For Landers, of the aftershocks in the first 10 days only 2% and 23% were recorded in the first 0.1 and 1 day, respectively. The corresponding average percentages over all mainshock magnitudes are 24% and 53%, respectively. We would have expected a clear trend of the number of aftershocks recorded in the early part of the sequence diminishing with increasing magnitude. However, only for mainshock magnitudes above 7.0 were the observations consistently below the average. The middle magnitude range between 4.5 and 6.0 have more than average numbers of aftershocks in the early part of the sequence, 29% and 61% in the first 0.1 and 1 day, respectively.

The mean abundance and other model parameters

Figure 7 compares the mean abundance as a function of mainshock magnitude for the time intervals [0, 10], [0.1, 10] and [1, 10] days. The mean abundance increases consistently in the mainshock interval 3.0 – 4.1. For magnitude 4.2 – 4.4, the data dip, and then increase again from magnitude 4.5. The break at 4.1 in the mean abundance coincides with an anomaly in the magnitude-frequency distribution which we discuss below. The increase of mean abundance is steeper for the higher magnitudes. However, the scatter is larger as there were fewer mainshock observations in that magnitude range. As there was no clear trend showing the loss of data between 0 and 1 days as a function of mainshock magnitude, we decided to determine the mean abundance in the time interval [0, 10] days. Figure 8 shows the best least squares fit for the data in the magnitude ranges 3.0 – 4.1 and 4.5 and above. We analyzed the aftershocks separately for the mainshocks in the two intervals. We determined the b-value and its 95% confidence interval by equation (12) and found $b = 1.10 \pm 0.04$ for the mainshock magnitude range 3.0 – 4.1 and 0.88 ± 0.02 for 4.5 and above, respectively. We used the IASPEI 1996 software to fit the p-value to the two sets and found $p = 0.69 \pm 0.10$ and

0.97 ± 0.03, respectively. The results were sensitive to starting values and therefore the standard deviation could be larger. Figure 9 shows the best fitting models.

Foreshock occurrence and the foreshock model

To calculate the rate of foreshock occurrence we focused on the initiating earthquake in each cluster. We regarded the first earthquake within the fitted ellipse as the initiating earthquake, unless the mainshock occurred outside the cluster and before the first event in the ellipse. We defined a foreshock as an initiating earthquake that was followed by a larger earthquake within the ellipse. Figure 10 shows the observed foreshock percentages. 95% confidence intervals were determined assuming a binomial distribution.

Model for mainshocks in the magnitude 4.5 and above

There were only six foreshocks of magnitude 4.5 and larger and none above 5.5. For magnitude 4.5, the observed foreshock occurrence was 2.1% and the 95% confidence interval 0 – 10%. The new aftershock model applied to foreshocks predicted 5.0% and between 2.5% and 5.7% when using the 95% limits for each model parameter. The confidence limits of the model are smaller than for the observations because they are derived from the more numerous aftershock data. Consistent with observations in California and Italy (Jones et al., 1999) as well as New Zealand (Savage and Rupp, 2000), the foreshock occurrence is smaller than predicted by the aftershock model. However, data and model agree within the confidence limits.

Model for mainshocks in the magnitude range 3.0 – 4.1

In the mainshock magnitude interval 3.0 – 4.1 there were hundreds of foreshocks and thousands of mainshocks and thus a comparison between model prediction and observations was much more meaningful. Figure 11 is a blow-up of figure 10 for the magnitude range 3.0 – 4.1. The solid line is the model derived from aftershocks according to equation (8) with $c = 0.70$, $b = 1.10$, $M_1 = 3.69$, $M_0 = 2.5$, and $\gamma = 1$. The rates increase for decreasing b-value, M_1 and c and decrease for larger b-value, M_1 and c. We calculated 95% confidence bounds c = 0.63-0.78, b = 1.06-1.14, M_1 = 3.66-3.72 displayed as dotted lines in Figure 11. The data generally agree well with the model demonstrating that foreshock prediction can be made from aftershock characteristics.

The model is consistent with the data in the two magnitude ranges considered. This is consistent with the hypothesis that foreshocks behave like mainshocks that happen to have larger aftershocks. We next look at the spatial distribution between foreshocks and mainshocks and finally at the distribution of time differences between foreshocks and mainshocks.

Distribution of foreshock mainshock pairs in space

If foreshocks were just like mainshocks that happen to have larger aftershocks, then we would expect that the mainshocks would occur within the aftershock area of the foreshock. We take the search area as proxy for the aftershock area. In our original search for clusters we used the magnitude of the larger earthquake to calculate a magnitude dependent search radius and compare this with the distance between two events. We are therefore certain to find mainshocks beyond the search radius of the foreshock. We had 3,804 foreshocks in our data set and for about 60% of them the mainshock occurred within their search radius. Many of the sequences contained multiple foreshocks. If we only look at the distance between the largest foreshock and the mainshock in each sequence, there were 1,056 pairs and 85.5% the mainshocks occurred within the search radius of their respective foreshock. To compare the distribution of foreshock-mainshock distances with the distribution of mainshock-aftershock distances we ran a new catalog search in which spatial comparisons were done using the first earthquake in a cluster, not the largest. We found 598 sequences for which the initiating earthquake was smaller than the largest. We scaled each distance between the largest foreshock and the mainshock with the search radius of the foreshock and plotted the cumulative frequency of all pairs that occurred within 10 days in figure 12 (plus signs, $N = 539$) and within 2 days ($N = 362$). Also shown are the cumulative frequency of distances between the largest aftershock and its mainshock, scaled by the mainshock search radius, that occurred within 10 days (diagonal crosses, $N = 1,653$) and two days (dots, $N = 1,142$).

The curves show that over time the distances of the mainshocks from the foreshocks and the aftershocks from the mainshocks both become larger, implying that the areas occupied by the aftershocks following the mainshocks and the mainshocks following the foreshocks increase with time in a similar way. Aftershocks following mainshocks are more tightly clustered than mainshocks following foreshocks. The curve of two-day distances of foreshocks and mainshocks coincides with the 10-day distances of mainshocks and aftershocks. By definition, foreshocks have smaller magnitudes than mainshocks. We suspect that the greater size of foreshock-mainshock distances arises from an overly large search radius for initiating earthquakes with small magnitudes. We conclude that a significant difference between the distance distributions cannot be resolved, and that therefore there is nothing in the distance data to negate the hypothesis that foreshocks are initiating events that happened to have a larger aftershock.

Distribution of foreshock mainshock pairs in time

Our new model for aftershock (or foreshock) numbers includes a parameter γ that can be used to calculate expected number of events in a time interval different from the one for which parameters were determined. As the data are already scattered significantly in the 10 day period we did not predict any fore- or aftershocks in another time interval. Here, we quickly compare the distribution of time differences in days between the initiating earthquake and its mainshock ($N = 359$) with the distribution of times between the largest aftershock and its mainshock ($N = 1,653$) during the first 10 days. Figure 13 shows the results. Between 0.02 and 0.2 days the foreshock distribution dips slightly below the aftershock one. Otherwise the graphs agree very well indicating that foreshock-mainshock pairs follow the same time decay as mainshock- aftershock pairs.

Break-down in scaling relation relations at mainshock magnitude 4.2?

We found a break down in the mean abundance as function of mainshock magnitude around magnitude 4.2. We observed an anomaly in the magnitude-frequency distribution of the complete catalog at the same magnitude. There were more observations of magnitude 4.2 (275) than magnitude 4.1 (237) rather than the 79% (about 187) expected with a Gutenberg-Richter relation and a b-value of 1.0. In the b-value plot (figure 2), this corresponds to a higher b-value because the average magnitude is smaller for that point. Rather than a break down in scaling we suspect a bias in magnitude determination. Assigning a magnitude of 4.2 to smaller earthquakes would lead to a decrease in the mean abundance as observed. Investigation of this possible bias is outside the scope of this study.

2.8 Discussion and conclusion

We have derived model parameters for the expected number of aftershocks in a chosen magnitude and time interval. The model is based on the mean abundance and uses the Gutenberg-Richter relation for the magnitude distribution and Omori's law for the temporal distribution. We described an objective method to define earthquake sequences and fine tuned it for the ANSS catalog in California. Studying aftershock characteristics we found two branches in the model for mean abundance. The growth exponent was 0.70 ± 0.08 in the magnitude range $3.0 - 4.1$ and 0.93 ± 0.16 for mainshock magnitude larger than 4.5. The break down in abundance scaling at magnitude 4.1 coincided with an anomaly in the magnitude-frequency distribution of all magnitudes and we suspect a bias in the catalogue magnitudes. We fitted the Gutenberg-Richter relation and Omori's law to aftershocks of mainshocks in each of the two branches of the mean abundance model and calculated expected foreshock numbers from the aftershock model. In the mainshock magnitude interval $3.0 - 4.1$ data and model agreed well. For mainshock magnitudes 4.5 and above data are few; the data and model agreed within the confidence intervals.

We have shown that for the Californian data foreshock occurrence is consistent with predictions made from aftershock characteristics. Thus foreshocks can be seen as mainshocks that happen to have larger aftershocks. Both the distributions of distances and time differences between foreshocks and mainshocks were also consistent with those between mainshocks and aftershocks. These findings are consistent with a single triggering mechanism being responsible for short-term earthquake occurrence.

Previous discrepancies between aftershock models and foreshock data

One motivation of our study was to investigate the inconsistency between claims for a single triggering mechanism for short-term earthquake occurrence (Reasenberg and Jones, 1989 and Felzer et al., 2004) and the over prediction of foreshock numbers from aftershock models (Jones et al., 1999 and Savage and Rupp, 2000). We used the same data for the calculation of foreshock numbers that had been used for deriving parameters for our aftershock model and found good agreement between model predictions and observations. In previous studies different subsets of the catalog were used. For foreshock studies the catalog was first stripped of aftershocks (e.g. Jones, 1985, and Savage and Rupp, 2000). To study aftershock parameters, large earthquake sequences were investigated. Consequently the productivity was larger than average and thus higher foreshock occurrence was predicted.

Are foreshocks self-similar?

For the magnitude interval $3.0 - 4.1$ we found that the probability for an earthquake to be a foreshock decreased from 8.8% for $M = 3.0$ to 3.3% for $M = 4.1$ because the growth parameter in the abundance model $c = 0.70 \pm 0.08$ was smaller than the b-value in the magnitude-frequency relation, $b = 1.10 \pm 0.04$. We found good agreement with the data. However, due to some bias in data selection foreshocks could still be self-similar, i.e. $c = b$: According to Utsu and Seki (equation 1) aftershock area increases 10-fold per mainshock magnitude unit. The Uhrhammer search area only increases 5-fold per magnitude unit. Thus for smaller mainshocks the search area is proportionally larger compared to the aftershock area. If we assume that background seismicity is randomly distributed in space, then the search area of smaller mainshocks is proportionally more likely to include background seismicity. Consequently the mean abundance would be artificially increased for smaller mainshocks and the growth exponent for abundance would be reduced. With increasing mainshock magnitude, the difference between the Uhrhammer and the Utsu & Seki search radius gets smaller (see figure 2).

For mainshock magnitudes of 4.5 and above, the values for $c = 0.93 \pm 0.16$ and $b = 0.88 \pm 0.02$ agreed within their 95% confidence interval. For the mean values, the probability for an initiating earthquake to be a foreshock increased from 5% at magnitude 4.5 to 6.5% at magnitude 7.0. If we assume $c = b = 1.0$, then a value of $M_l = 3.72$ yields a foreshock probability of 6% as observed by Jones (1985). This is not too different from our observations. The problem is that there are fewer mainshocks with large magnitudes and therefore the uncertainty in the model for these events is fairly large.

In summary we found that observed foreshock occurrence is consistent with aftershock characteristics in number, space and time. Thus a foreshock can be regarded as a mainshock that happened to have a larger aftershock.

2.9 Acknowledgements

An initial version of the paper was presented as a poster at the Fourth International Workshop on Statistical Seismology in Tokyo. The first author gratefully acknowledges travel support from the workshop organizer Professor Ogata and the Workshop sponsors. Workshop participants provided useful comments. We thank David Vere-Jones and Martha Savage for useful feedback on the manuscript. Thessa Tormann assisted with the initial analysis of the ANSS catalog. The study has been partly funded by a grant from the New Zealand Earthquake Commission Research Foundation.

2.10 References

Aki, K. (1965) *Maximum likelihood estimate of b in the formula log N = a –bM and its confidence limits*, Bull. Earthquake Res. Inst. 43, 237-239.

ANSS (Advanced National Seismic System) (2005), http://www.ncedc.org/anss/catalog-search.html

Christophersen, A. (2000), *The probability of a damaging earthquake following a damaging earthquake*, PhD Thesis, Victoria University of Wellington.

Console, R., M. Murru, and A. M. Lombardi (2003), *Refining earthquake clustering models*, J. Geophys. Res. 108, doi:10.1029/2002JB002130.

(Dodge, D. A., G. C. Beroza, and W. L. Ellsworth (1996). *Detailed observations of California foreshock sequences: Implications for the earthquake nucleation process*. J. Geophys. Res. 101, 22371-22392.)

Eberhart-Phillips, D. (1998) *Aftershock sequence parameters in New Zealand*, Bull. Seism. Soc. Am. 88, 1095-1097.

Felzer, K. R., R.E Abercrombie, and G Ekstroem (2004), *A common origin for aftershocks, foreshocks, and multiplets*, Bull. Seism. Soc. Am. 94, 88-98.

Gardner, J.K., and L. Knopoff (1974), *Is the sequence of earthquakes in Southern California, with aftershocks removed Poisson*, Bull. Seism. Soc. Am. 64, 1363-1367.

Gerstenberger, M.C., Wiemer, S., L.M. Jones, and P.A. Reasenberg (2005), *Real-time forecasts of tomorrow's earthquakes in California*, Nature 435, doi:10.1038/ nature03622.

Guo, Z. and Y. Ogata (1997), *Statistical relations between the parameters of aftershocks in time, space and magnitude*, J. Geophys. Res., 102, 2857-2873.

Gutenberg, B., and C. F. Richter (1949), *Seismicity of the earth and associated phenomena*, Princeton University Press.

Helmstetter, A. (2003), *Is earthquake triggering driven by small earthquakes?*, Phys. Rev. Lett. 91, DOI: 10.1103/PhysRevLett.91.058501.

Jones, L.M. (1985), *Foreshocks and time-dependent earthquake hazard assessment in Southern California*, Bull. Seism. Soc. Am. 75, 1669-1679.

Jones, L. M., R. Console, F. Di Luccio, and M. Murru (1999). *Are foreshocks mainshocks whose aftershocks happen to be big?* in press.

Kagan, Y.Y. (2004), *Short-term properties of earthquake catalogs and models of earthquake source*, Bull. Seism. Soc. Am. 94, 1207-1228.

Reasenberg, P.A., and L.M. Jones (1989), *Earthquake hazard after a mainshock in California*, Science 243, 1173-1176.

Reasenberg, P.A., and L.M. Jones (1994), *Earthquake Aftershocks: Update*, Science 265, 1251-1251.

Richter, C.F. (1958), *Elementary Seismology*, Freeman.

Savage, M.K., and S.H. Rupp (2000), *Foreshock probabilities in New Zealand*, New Zealand Journal of Geology & Geophysics 43, 461-469.

Uhrhammer (2005), *personal communication.*

Utsu, T. (1969), *Aftershocks and earthquake statistics (I) Source parameters which characterize an aftershock sequence and their interrelations,* J.Fac. Sci., Hokkaido Univ., Ser. VII 3, 129-195.

Utsu, T, (2002), *Statistical features of seismicity*, in International Handbook of Earthquake & Engineering Seismology Part A, Eds. W. H. K. Lee, H. Kanamori, P.C.Jennings, and C. Kisslinger (Academic Press, San Diego), 719-732.

Utsu, T., Y. Ogata, and R. S. Matsu'ura (1995), *The Centenary of The Omori Formula for a decay law of aftershock activity*, J. Phys. Earth 43, 1–33.

Utsu, T., Y., and A. Seki (1955), *A relation between the area of aftershock region and the energy of the mainshock*, Zisin, 7, 233-240.

Vere-Jones, D., J. Murakami, and A. Christophersen, (2006), *A Further Note on Båth's Law*, manuscript submitted to Bull. Seism. Soc. Am.

Wells, D. L. and K. L. Coppersmith, (1994), *New empirical relationships among magnitude, rupture width, rupture area, and surface displacement*, Bull. Seism. Soc. Am., 84, 974-1002.

Woessner, J., and S. Wiemer (2005), *Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty,* Bull. Seism. Soc. Am. 95, 684-698, doi: 10.1785/0120040007.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2004), *Analyzing earthquake clustering features by using stochastic reconstruction*, J. Geophys. Res. 109 B05301, 4755-4768.

Zhuang, J., and Y. Ogata, (2006), *Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks,* manuscript.

2.11 Tables

Table 1: Foreshock probabilities for New Zealand earthquakes. Values from Savage and Rupp (2000) using the model for 17 earthquake sequences in New Zealand analyzed by Eberhart-Phillips (1998).

	Magnitude difference between first event, M_1 and second, M_2.		
	$M_2 > M_1$	$M_2 > M_1+0.5$	$M_2 > M_1+1.0$
Observed foreshock probability within 5 days and 30 km	4.5 ± 0.7%	1.5 ± 0.4%	0.8 ± 0.3%
Mean model foreshock probabilities	10.7%	3.4%	1.1%
Range of model foreshock probabilities	0.21-52%	0.054-23%	0.058-10%

2.12 Figure captions

Figure 1: The magnitude frequency distribution of the complete catalog.

Figure 2: The b-value for the complete catalog as a function of magnitude and 95% confidence interval.

Figure 3: Example of three different search radii in space. Models are adapted from Utsu and Seki (1955) solid, Gardner and Knopoff (1974) dashed, and Uhrhammer (2005) dot-dash.

Figure 4: Example of the effect of different search radius. The same earthquake of magnitude 4.1, with search results from (a) the Gardner-Konpoff window, (b) the extrapolated Uhrhammer window, and (c) the Utsu-Seki window.

Figure 5: A magnitude-time scatter plot of aftershocks for the Landers $M = 7.3$ mainshock.

Figure 6: Percentage of all aftershocks in the first 10 days that occurred in the first 0.1 (open diamonds) and 1 day (solid diamonds).

Figure 7: Mean abundance in time intervals [0,10] (solid diamonds), [0.1,10] (open triangle), and [1,10] (dash) days.

Figure 8: The mean abundance and models for the magnitude intervals 3 – 4.1 (dashed line) and 4.5 – 7.3 (solid line).

Figure 9: The decay of aftershock activity with time and the best fitted model for the two mainshock magnitude ranges 3.0 – 4.1 (triangles) and 4.5 and above (diamonds).

Figure 10: The percentage of initiating earthquakes that were followed by larger events with 95% confidence intervals for the observations.

Figure 11: Foreshock occurrence in the mainshock magnitude interval 3.0 – 4.1 and the model prediction from aftershock analysis.

Figure 12: The comparison of the distribution of foreshock-mainshock distances scaled by the foreshock search radius and mainshock-largest aftershock distances scaled by the mainshock magnitude. Diamonds: foreshocks within two days; + all foreshocks; x aftershocks.

Figure 13: The cumulative distribution of time differences between foreshocks and their mainshocks (diamonds) and mainshocks and their largest aftershock (x).

2.13 Figures

Figure 1:

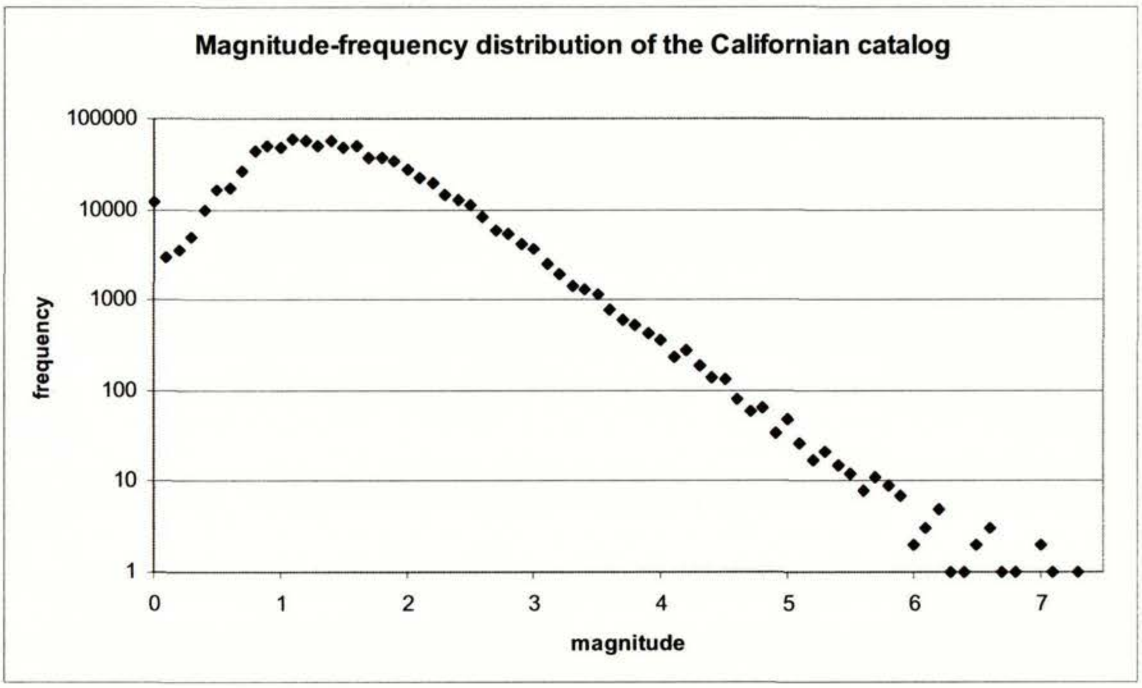


Figure 2:

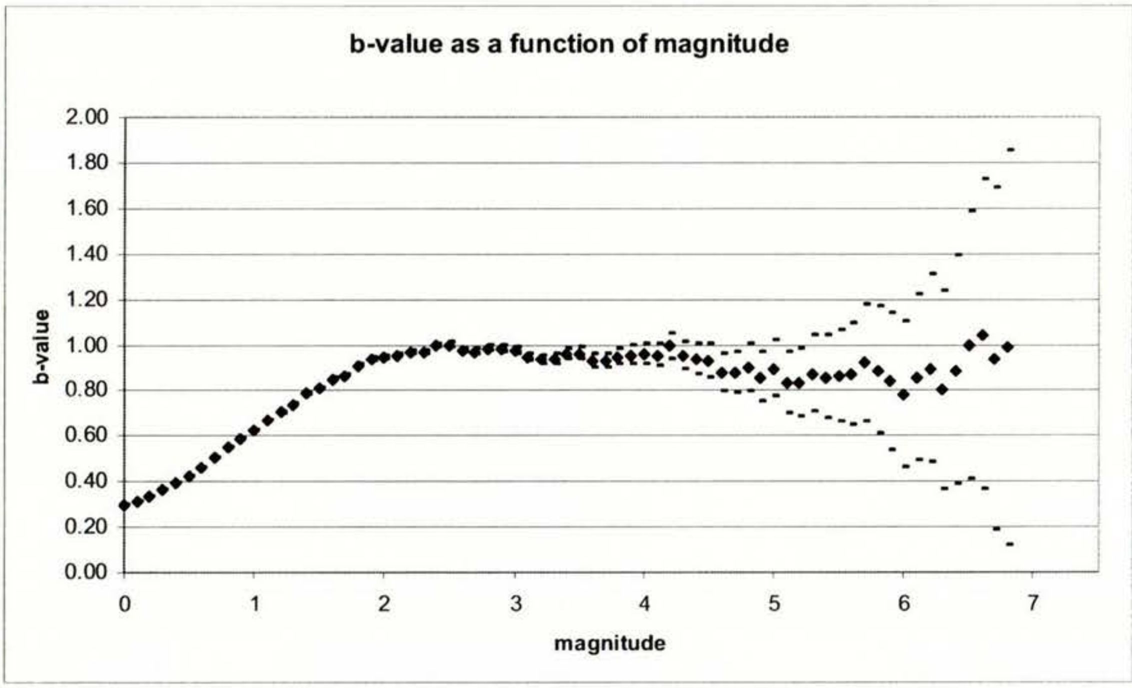


Figure 3

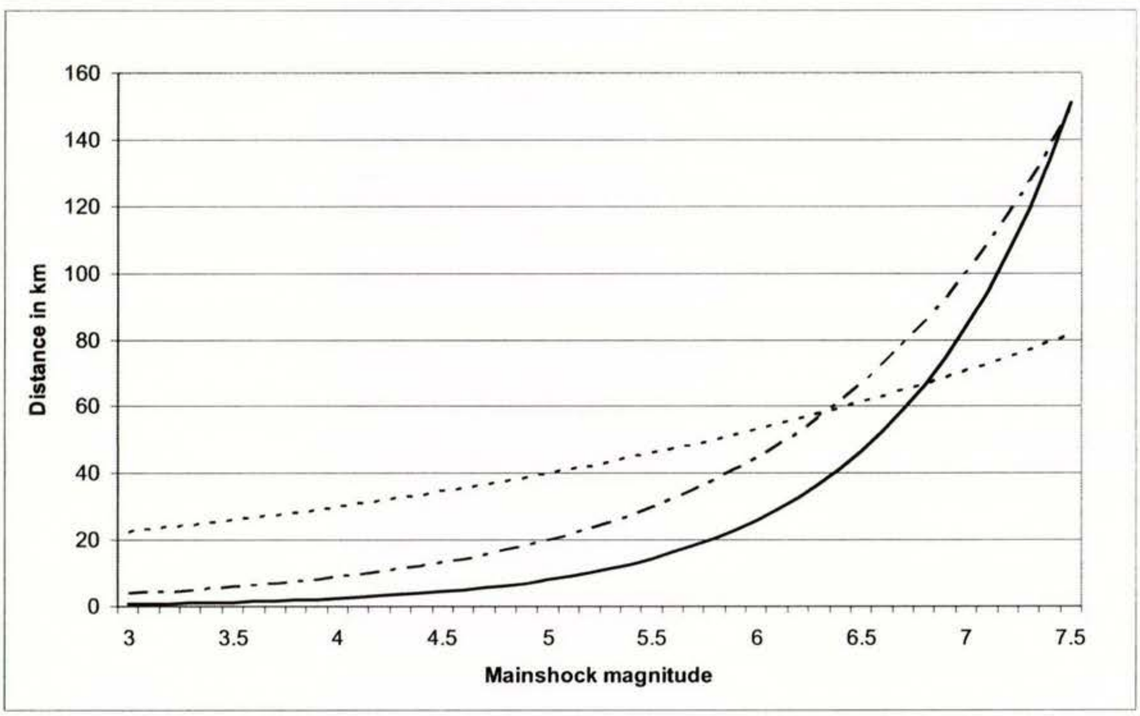
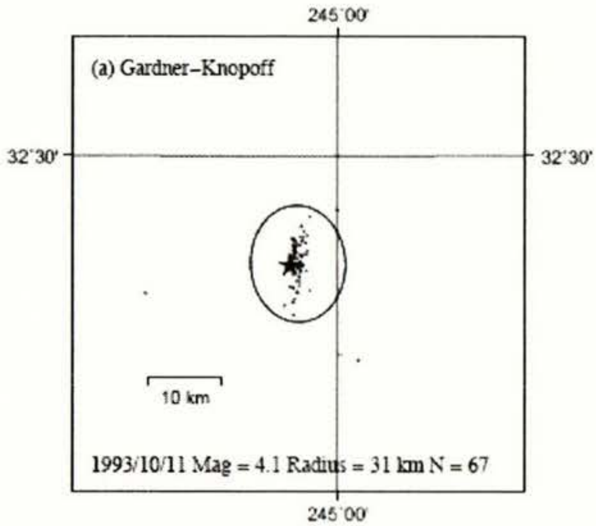
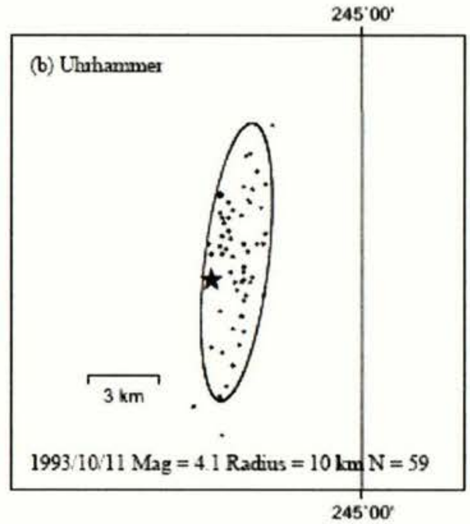


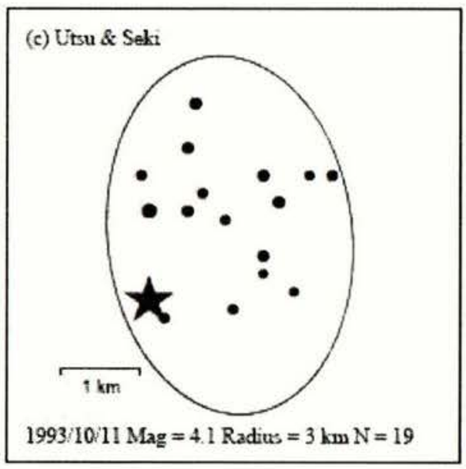
Figure 4



245°00'
(a) Gardner–Knopoff
32°30'
32°30'
10 km
1993/10/11 Mag = 4.1 Radius = 31 km N = 67
245°00'



245°00'
(b) Uhrhammer
3 km
1993/10/11 Mag = 4.1 Radius = 10 km N = 59
245°00'



(c) Utsu & Seki
1 km
1993/10/11 Mag = 4.1 Radius = 3 km N = 19

Figure 5

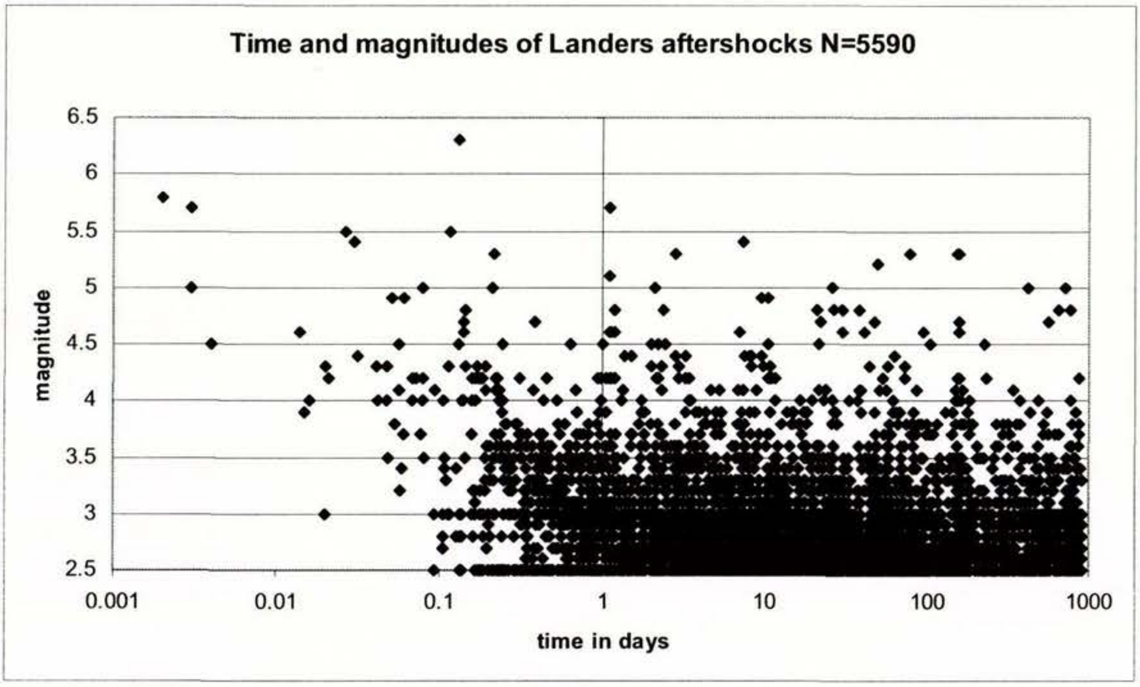


Figure 6

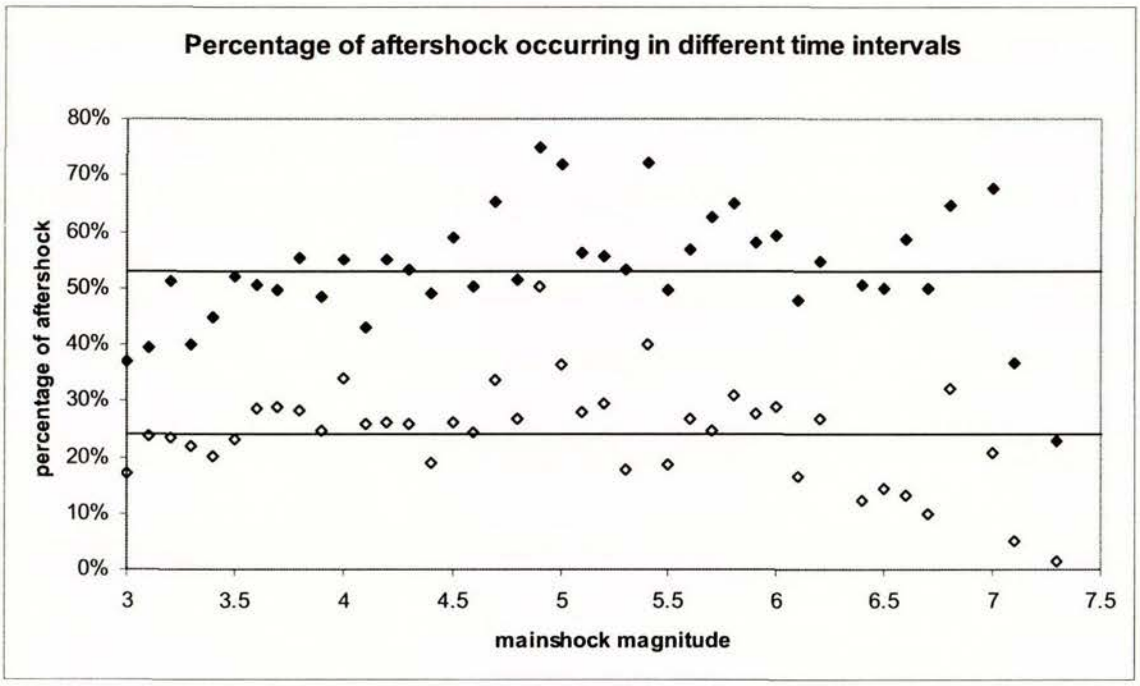


Figure 7

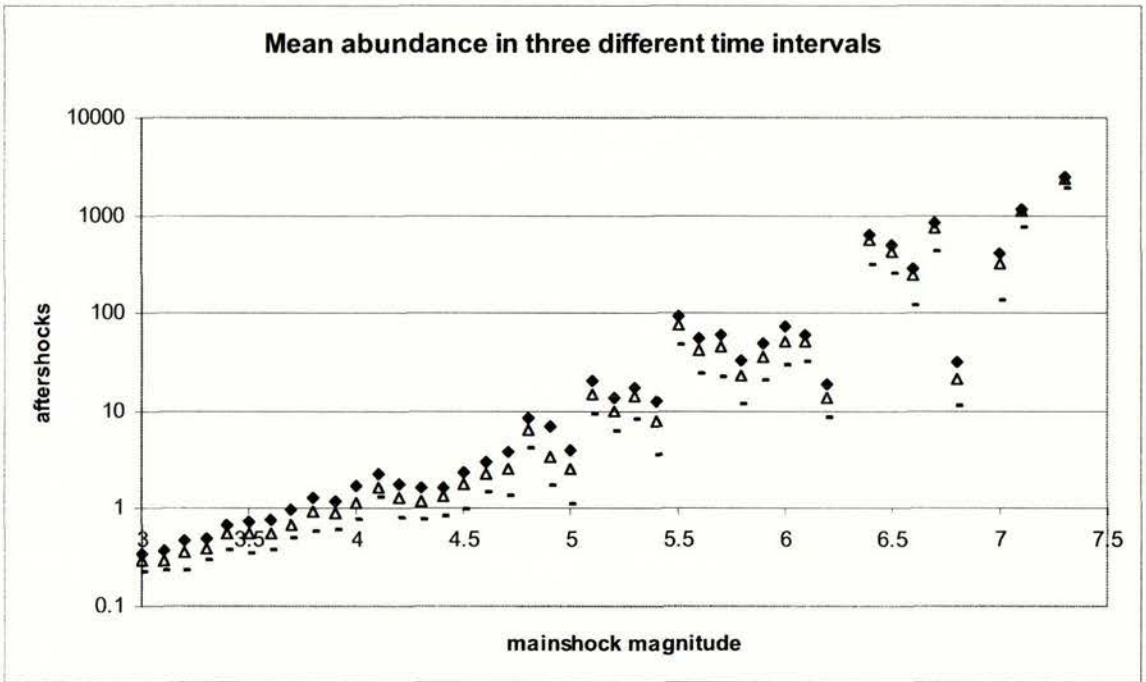


Figure 8

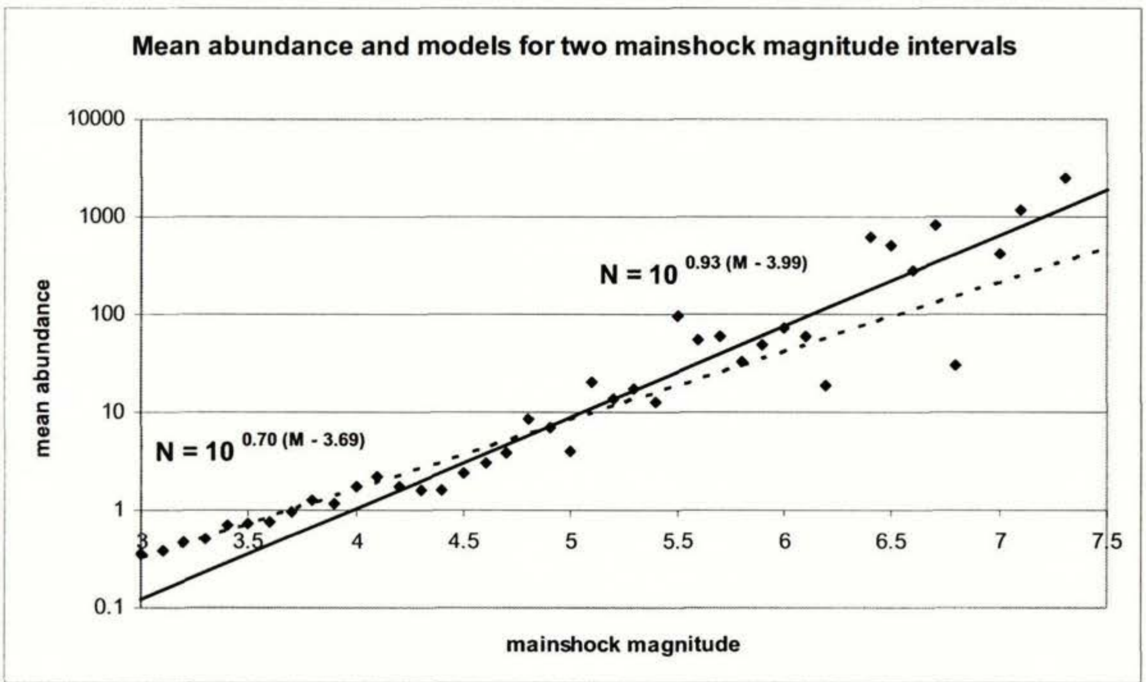


Figure 9

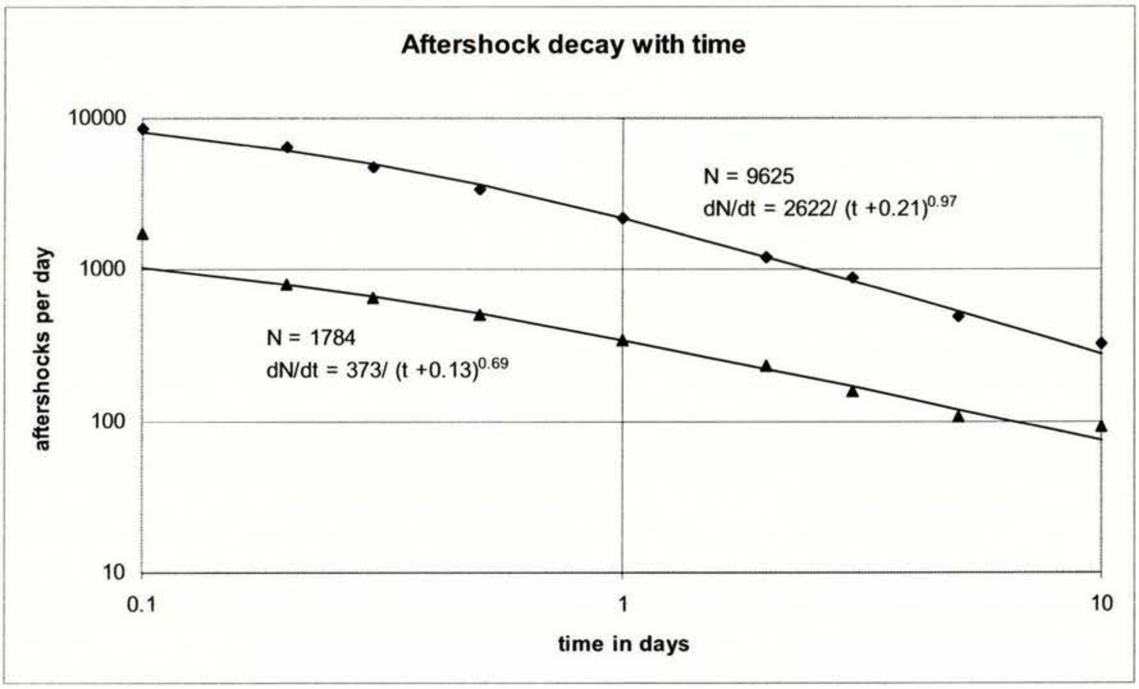


Figure 10

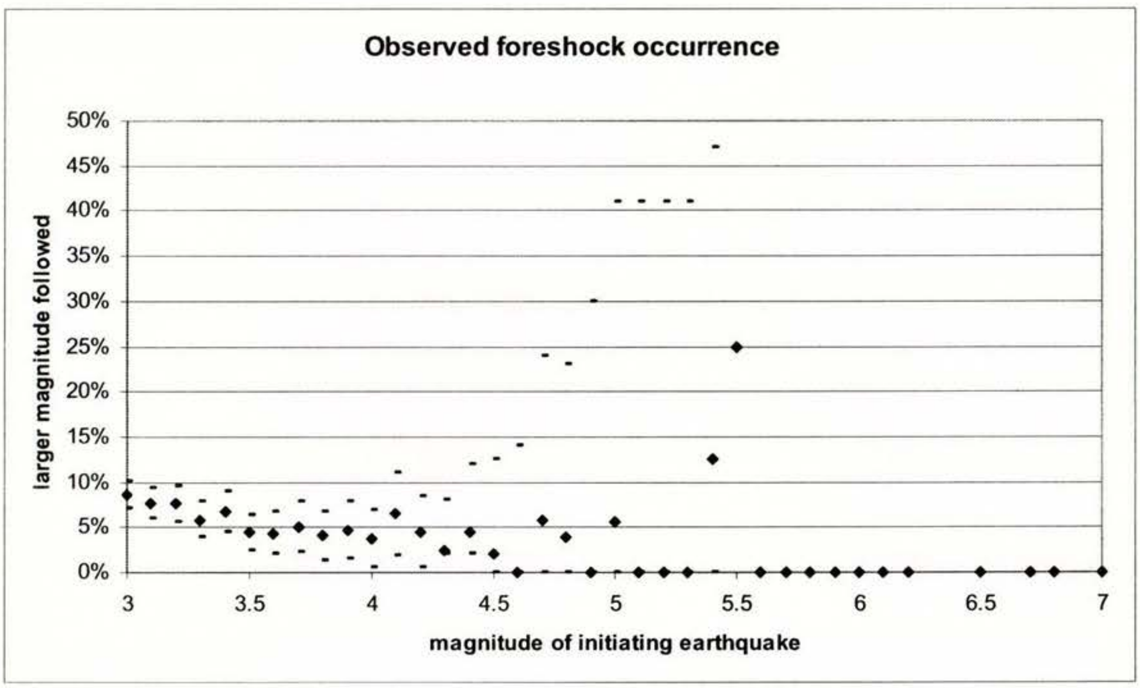


Figure 11

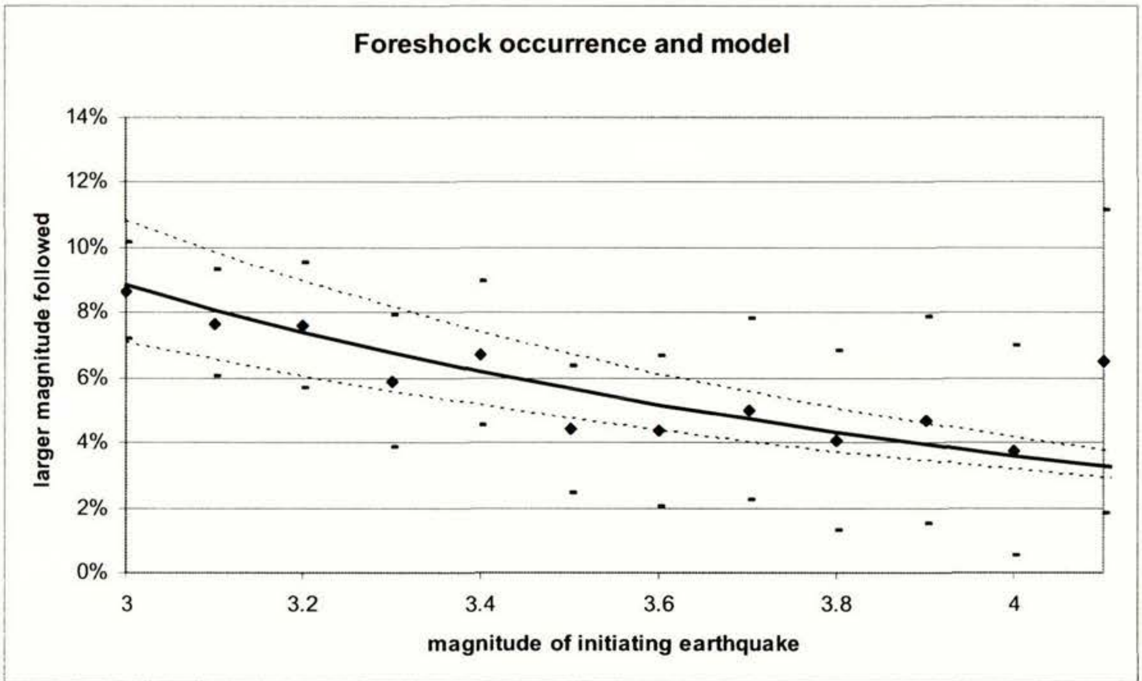


Figure 12

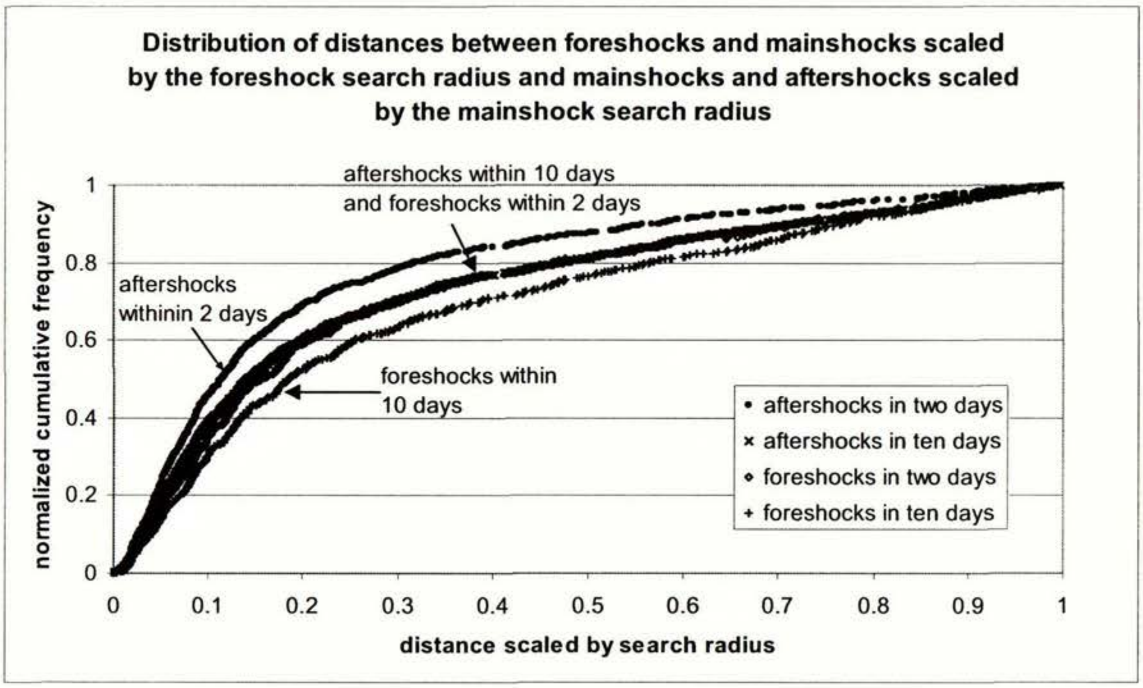
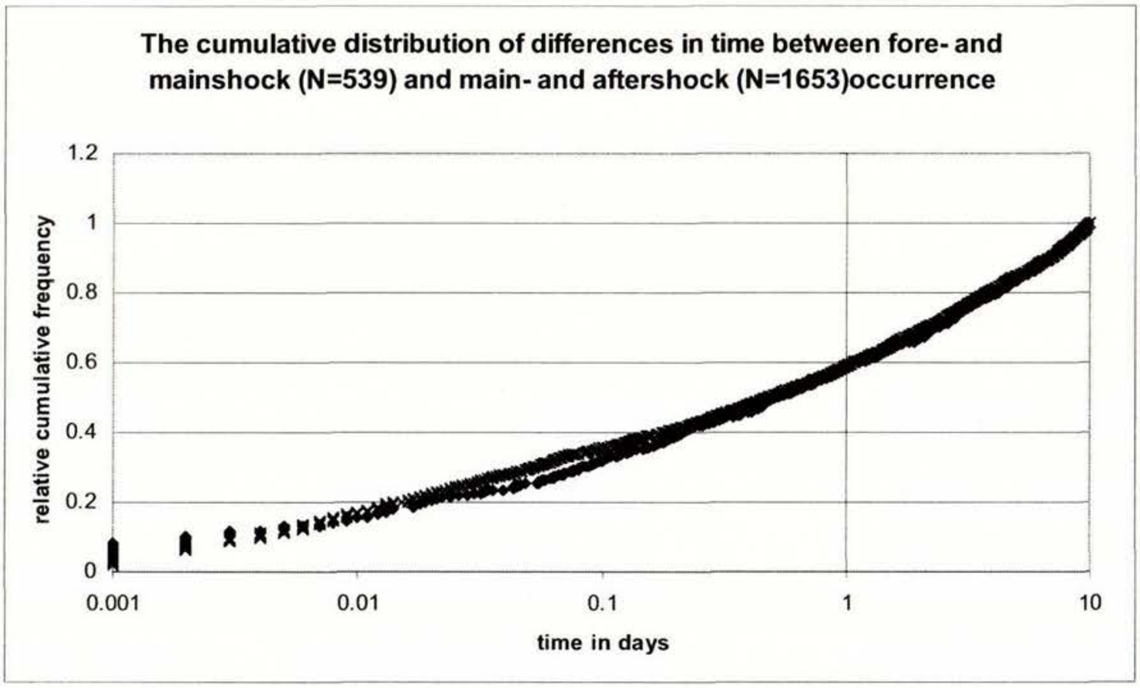


Figure 13


The cumulative distribution of differences in time between fore- and mainshock (N=539) and main- and aftershock (N=1653)occurrence
relative cumulative frequency
1.2
1
0.8
0.6
0.4
0.2
0
0.001
0.01
0.1
1
10
time in days

3 Aftershock abundance: Forecasting aftershock occurrence when catalog completeness is low

Manuscript drafted by Annemarie Christophersen to be submitted
to the Bulletin of the Seismological Society of America
Co-authors: Matt Gerstenberger and Euan Smith.

3.1 Abstract

We set up a model for making 24 hours earthquake forecasts which is based on the number of observed earthquakes and can be fitted to an on-going earthquake sequence even when the number of recorded earthquakes is low. We derived model parameters from the catalog of the Southern California Earthquake Data Center. We first selected time periods, for which earthquake locations and magnitude had been determined consistently and then applied an objective method to define aftershocks. We analyzed aftershocks in time, magnitude and space. We found that on average our model predicted only about half as many aftershocks than the generic Reasenberg and Jones aftershock model for California. We incorporated our aftershock model into the Short Term Earthquake Probability Model (STEP) to probabilities of exceeding a level of ground shaking in a 24 hours period. Our model it ready to be incorporated into the Regional Earthquake Likelihood Model Testing (RELM).

3.2 Introduction

In May 2005, the United States Geological Survey began publishing daily online maps of short term earthquake probabilities (STEP) for California (Gerstenberger, et al., 2005 and http://pasadena.wr.usgs.gov/step/). The inputs to the ground shaking are earthquake rates with mean aftershock parameters for California (Reasenberg and Jones, 1989). As real time data for an on-going earthquake sequence becomes available, the model parameters are recalculated and the model allows for more complexity as described below.

Immediately following a large earthquake, smaller earthquakes may not be detected in the coda of the mainshock, and the numerous aftershocks may overwhelm the processing capability. Thus, when operating in real time, the STEP model relies on generic aftershock behavior in the initial time after a mainshock when calculating forecasts.

We present a variation on the STEP model that allows for fitting of the data even when the amount of data is limited. The model is based on the abundance, i.e. the number of aftershocks per mainshock in a time interval which in retrospective analysis aftershocks are assumed to be recorded completely. We derive mean model parameters for California and show that the current generic parameters in STEP overpredict the number of aftershocks by about a factor of two. We incorporate the abundance model into STEP and illustrate the difference in predicted ground motion following a moderate mainshock with few aftershocks and a large mainshocks with many aftershocks that are not fully recorded in the early stages of the sequence.

In recent years efforts have increased to set up testing procedures for earthquake models, with the RELM (Regional Earthquake Likelihood Models) project at the forefront (Schorlemmer et al., 2005, and Schorlemmer and Gerstenberger, 2005). Similar test centers are being developed in Europe and New Zealand. Our model is ready to be incorporated in the RELM testing and we are looking forward to test our models in other settings.

3.3 The STEP model

The STEP model is based on probabilistic seismic hazard analysis and produces maps that show the probability of exceeding a level of ground shaking in a 24 hours period (Gerstenberger et al., 2004 and 2005). The time-varying component of the STEP model uses a simple model for aftershock occurrence as introduced by Reasenberg and Jones (1989, 1990 & 1994) where the rate $\lambda(t, M_m)$ of aftershocks of magnitude M and larger above a magnitude threshold M_0 is given by

$$\lambda\,(t,\, M_m) = 10^{\,a' + b\,(Mm - M)} \,/\, (t + t_c)^p. \tag{1}$$

The magnitude threshold M_0 is usually the smallest magnitude for which a set of earthquake data is completely recorded and thus is also referred to as completeness magnitude.

Equation (1) is a variation on the modified Omori's law for aftershock decay with time (e.g. Utsu, 1969 or Utsu et al., 1995)

$$\lambda(t) = K/(t+t_c)^p\,, \tag{2}$$

where K is a parameter that is proportional to the aftershock productivity; p describes the decay and takes values around 1.0; and t_c stands for a small time interval just after the mainshock. Reasenberg and Jones (1989) introduced dependency on the mainshock magnitude by replacing K with $10^{\,a' + b\,(Mm - M)}$. They analyzed 62 Californian earthquake sequences and derived the parameters $a' = -1.67$, $p = 1.08$, $b = 0.91$ and $t_c = 0.05$ which have become known as the generic Californian aftershock model parameters (e.g. Gerstenberger et al., 2004).

To calculate hazard maps, the aftershock rates need to be distributed in space. Gerstenberger (2003) developed a three step method, starting out with a circle around the mainshock with a radius $r(M)$ of at least 5 km and a $1/r^2$ decay from the center. The radius $r(M)$ is derived from scaling relations for the subsurface rupture length according to Wells and Coppersmith (1994)

$$r(M) = 10^{0.59M - 2.44}\,. \tag{3}$$

As aftershock data become available a fault trace is estimated and the aftershock area refined. For details see Gerstenberger et al. (2004). Once at least 100 aftershocks above the completeness magnitude are recorded, sequence specific model parameters are calculated. For very productive sequence the parameters are allowed to vary spatially (Gerstenberger et al., 2005).

3.4 The abundance model

The average number of aftershocks, or mean abundance N (M) in a chosen time interval [t_1, t_2] increases exponentially with mainshock magnitude M and can be written in the form

$$N(M) = 10^{c(M-M_1)}, \tag{4}$$

where the parameter M_1 corresponds to the mainshock magnitude that on average has one aftershock above the completeness magnitude of the catalog (Christophersen, 2000, and Christophersen and Smith, 2006). The value c, in many studies referred to as α, corresponds to the triggering capability of differently sized mainshocks. In a global study, we found the value c to be 1.0 for most tectonic settings (Christophersen, 2000) and for California it was 0.70 ± 0.08 in the magnitude range 3.0 – 4.1 and 0.93 ± 0.16 for mainshock magnitude larger than 4.5 (Christophersen and Smith, 2006). Other observations for c include 0.8 for California (Helmstetter, 2003), 0.5 for Italy (Console et al., 2003) and 0.8 on average for Japan with values varying between 0.2 and 1.9 (Guo and Ogata, 1997).

Christophersen (2000) proposed that the abundance for a given mainshock magnitude follows a geometric distribution where the probability P_n of observing n aftershocks in given by

$$P_n = (1-q)q^n. \tag{5}$$

The parameter q can be calculated from the average abundance N_{ave}

$$q = N_{ave}/(1 + N_{ave}) \tag{6}$$

The frequency size distribution of earthquakes, also know as Gutenberg-Richter relation, can be used to distribute the total number of aftershocks with magnitude.

$$N(M) = 10^{a-bM}, \tag{7}$$

where $N(M)$ is the of number of earthquakes of magnitude M and larger, and a and b are parameters (Gutenberg and Richter, 1944 and 1949).

Assuming independence of abundance and the magnitude-frequency relation, the expected number of aftershocks of magnitude M and larger, following a mainshock of magnitude M_m in a chosen time interval [t_1, t_2] can be written as

$$N = 10^{c(Mm-M1)}\, 10^{-b(M-Mo)}, \tag{8}$$

where M_0 is the completeness magnitude for which abundance was determined.

Once we have derived a model for the number of triggered earthquake in one time interval $[t_1, t_2]$, we can derive the expected number of triggered earthquakes in another time interval $[t_3, t_4]$. For this purpose we define a parameter γ that relates the number of aftershocks in one time interval to the number in another: $N_t = \gamma \; N_m$. The Omori' law (equation 2) can be integrated to determine γ:

$$\gamma = \ln [(t_4 + t_c)/ (t_3 + t_c)] / \ln [(t_2 + t_c)/ (t_1 + t_c)] \qquad \text{for } p = 1 \qquad (9)$$

$$\gamma = [(t_4 + t_c)^{1-p} - (t_3 + t_c)^{1-p}] / [(t_2 + t_c)^{1-p} - (t_1 + t_c)^{1-p}] \qquad \text{for } p \neq 1 \qquad (10)$$

To be able to establish the abundance model according to equation (8) we first need to identify aftershocks. We used a method developed by Christophersen (2000) and described in Christophersen and Smith (2006) and Vere-Jones et al., (2006). In brief, we first searched for related earthquakes in a magnitude dependent radius and space and a time window that depended on the completeness magnitude of the catalog and the tectonic region. This way we established a data set of earthquake sequences. Each sequence with at least three events was fitted by an ellipse using a least squares method. Events within the 90^{th} percentile ellipse were regarded as foreshocks or aftershocks.

3.5 The earthquake catalog

For the study of earthquake occurrence it is important to have a set of data that had magnitude and location determined consistently. For our analysis we used data from the Southern California Earthquake Data Center (SCEDC). The catalog includes earthquakes between latitude 32-37 North and longitude 122-114 East. Data were restricted to events with depths shallower than 40 km. SCEDC (2006) identifies two time period with consistent magnitude and location determinations: 1932 – 1976 and 1981 – 1999. We used those subsets of the SCEDC catalog for our analysis. Figure 1 shows a map of the data and figure 2 presents the magnitude-frequency distribution for each subset.

3.6 Analyzing earthquake clusters

Our first aim was to set up the mean abundance model according to equation (4). Abundance depends on the time interval, the completeness magnitude and the mainshock magnitude. The completeness magnitude varies with mainshock magnitude and time from the mainshock. We first investigated the completeness magnitude, and calculate the b-value of the magnitude-frequency relation at the same time. We then set up a model for the mean abundance. We calculated the p-value of Omori's law and then compare the abundance model with the Reasenberg & Jones model. We then investigated the variation of abundance and finally close with some spatial analysis.

Completeness considerations

Outside an on-going earthquake sequence the completeness magnitude can be as low as 1.2 in some regions of California (Woessner and Wiemer, 2005). However, subsequent to a large earthquake the completeness magnitude sharply increases because smaller events cannot be detected in the coda of larger events (Kagan, 2004). Limited resources can also lead to a delay in analysts processing earthquake sequences. Thus the average recorded aftershock magnitude is much larger in the early part of the sequence than later on. We assume that all earthquakes above a chosen cut-off are recorded, once the average aftershock magnitude does not change with time. For the time period 1981-1999, we calculated the average magnitude for all aftershocks of magnitude 2.45 and larger in logarithmic time bins 0.001-0.01, 0.01-0.1, 0.1-1, 1-10 and 10-100 days for different mainshock magnitude intervals. The results are shown in figure 4. For the mainshock magnitude range 3 – 6, the average aftershock magnitude does not change significantly from 0.1 days onward.

b-value calculations

The *b*-value in the Gutenberg-Richter relation (equation 7) is proportional to the inverse of the average magnitude. Thus the *b*-value as function of cut-off magnitude can be used as another tool to investigate the completeness magnitude. If the *b*-value is calculated in a limited magnitude range a correction is required (Christophersen, 2000).

$$b = [(\ln(10)*(M_{ave} - cor)]^{-1} \qquad (11)$$

$$cor = (M_u - 0.05)\, e^{\beta\,(Mu - 0.05)} - (M_l - 0.05)\, e^{\beta(Ml - 0.05)} / [e^{\beta(Mu - 0.05)} - e^{\beta(Ml - 0.05)}] \qquad (12)$$

where $\beta = \ln(10)b$ and M_u and M_l are the upper and lower limits of the magnitude interval (Christophersen, 2000). The correction also includes an adjustment for the magnitude bin width of 0.1.

Figure 5 (a) shows the *b*-value and 95% confidence intervals as function of cut-off magnitude for aftershocks in the time interval 0.1-10 days for mainshocks in the magnitude range 2.95 – 5.95 for the data from 1981 - 1999. The *b*-value was 1.00 ± 0.02 at magnitude 2.5 and dropped off for smaller magnitudes. Figure 5 (b) shows the data for aftershocks in the time period 1932 – 1976. The *b*-value was 0.98 ± 0.09 at magnitude 4.0 and dropped off for smaller magnitudes. We also calculated *b*-values for the not declustered catalogs. In the time period 1932 – 1976 we found the completeness magnitude to be 4.0 with a *b*-value of 0.99 ± 0.06. For the time period 1981 – 1999, we found that the *b*-value increased with cut-off magnitude and then decreased in the magnitude interval 2.1 – 3.1. Felzer (2006) ascribed such a pattern to increasing magnitude error with decreasing magnitude combined with catalog incompleteness. The *b*-value at magnitude 3.1 was 1.02 ± 0.02.

The mean abundance model

We derived the mean abundance model by counting all the aftershocks above magnitude 2.45 in the time interval 0.1 – 10 days as a function of mainshock magnitude and then divided the total number of aftershocks per mainshock by the number of respective mainshocks. Figure 6 shows the data and the model in the mainshock magnitude interval 3-5.9. The model was fitted by a least squares method to the logarithm of abundance. We excluded the four largest mainshocks from the analysis because the aftershock data would have not been complete in the chosen time and magnitude intervals. However, the data still fitted the model reasonably well, indicating that those sequences had more than average number of aftershocks. The best fitting slope *c* was 0.89 ± 0.02 and the parameter M_l was 3.75 ± 0.02.

For the time period 1932 – 1976, we counted aftershocks above magnitude 3.95, in the time interval 0.1 -1 10 days. We used the Gutenberg-Richter law to calculate the mean abundance from the 1981 – 1999 model if the completeness magnitude was increased to 3.95. The data fitted the model well. Figure 7 shows the data and the best fitting model for the data as well as the model from the later time period. The models agreed reasonably well.

The variation of aftershock abundance

Figure 6 and 7 show that the mean abundance per mainshock magnitude is relatively well behaved. However, Christophersen (2000) and Christophersen and Smith (2000) showed that for any given mainshock magnitude in a similar tectonic setting, the abundance can vary significantly. This is also the case in California. Figures 8 (a) – (d) show the smallest and largest abundance observed for each mainshock magnitude in the range 2.95 – 5.95 as well as the mean abundance model for different completeness magnitudes M_0. For example, the number of observed aftershocks of $M_0 \geq 2.45$ varied from 7 – 262 for $M_m = 5.5$ and the model mean was 40.

The dotted line in figure 8 corresponds to the first n that reached the 95% confidence interval of the geometric distribution for the mean abundance model (equation 5). The steps in the model are caused by the discrete nature of the distribution. More than 5% of the data are above the model, indicating that a distribution with a longer tail was needed to describe the abundance for a given mainshock magnitude.

The decay of aftershock activity

To find the p-value for the aftershocks in the abundance model, we used the IASPEI 1996 software (IASPEI, 1996) that fits the parameters of the Omori-Utsu law by maximum likelihood method. For the mainshock magnitude range 2.95 – 5.95 we calculated the days between mainshock and aftershocks for each sequence, sorted the times chronologically and fitted the p-value in the time interval 0.1 – 10 days. We used different cut-off magnitude for completeness. For $M_0 = 2.5$ there were 2,995 aftershocks. The p-value and one standard deviation was 0.79 ± 0.03. A relatively small p-value could indicate that background seismicity was included in the sequences and thus the aftershock decay appeared to be slower. We limited the data to earthquake sequences that had at least 10 events. For $M_0 = 2.5$ there were 2,038 aftershocks. The p-value and one standard deviation was 0.89 ± 0.03. We observed a tendency for the p-value to increase with increasing cut-off magnitude M_0 but the interval of the estimate plus and minus one standard deviation overlapped. We decided not to pursue further a possible change in p-value with completeness magnitude.

Statistically, a p-value of 1.0 or smaller can only apply in a limited time period because the integral of the Omori-Utsu law only converges for $p < 1.0$. However, in a global study Smith and Christophersen (2006) found $p = 1.0$ until the background seismicity started to dominate after about 200 days. For simplicity we assigned $p = 1.0$ and $t_c = 0.1$ day. In our model the p-value and t_c are used to calculate the parameter γ (equation 9 and 10) that relates the observed abundance in one time interval to abundance in another time interval. Table 1 shows how the predicted number of events varied in percentage in respect to $p = 1.0$ for $p = 0.89$ and $p = 1.08$ in different time-intervals.

Some spatial considerations

To include our model into RELM or any other testing center, we need to spread our expected number of aftershocks in space. With limited data available, the spatial extent will be circular around the mainshock. To estimate an appropriate radius, we calculated the average length of the longer ellipse axis and plotted the data in figure 9. Also shown in the figure are the Uhrhammer search radius which was used to search for the initial set of earthquake sequences, a trend line from the data, which was calculated using a least squares method on the logarithm of the distances, and the Wells-Coppersmith radius used in the STEP model. The smallest earthquake used to derive the scaling relationship between mainshock and fault length was 4.8 (Wells and Coppersmith, 1994). The data agreed reasonably well with the model for mainshocks larger than magnitude 5.0. For smaller mainshocks the increase of ellipse length with mainshock magnitude was much flatter. For mainshocks of magnitude 4.0, the trend line had a value of 3.4 km. The location uncertainty was in the order of a few kilometers and acted as a lower boundary for the average distance. Thus our results could not provide an improvement on the current STEP model procedure for limited data: To choose a minimum circular radius of 5 km and apply scaling relations for the subsurface rupture with magnitude according to Wells and Coppersmith (1994) for larger mainshocks.

Comparison of observed and predicted abundance

Figure 10 shows some examples of how the predicted abundance from earlier time intervals compared to the observed abundance on day 2, 3 and 4. Prediction and observation agreed reasonably well, but less so when the maximum observed abundance dropped below 10. Overall, the data available for this comparison were relatively limited in number and we only used this comparison as an illustration that the method should work. More formal testing will be applied by incorporating the model into the RELM testing.

3.7 Comparison between the abundance and the Reasenberg & Jones model

Both the mean abundance model and the Reasenberg and Jones model are based on the modified Omori law (equation 2). We derived our model parameters by studying earthquake sequences, including single events and doublets, from a complete and homogeneous catalog. We stacked data from all sequences to derive model parameters. In contrast, Reasenberg and Jones (1989) studied 62 well recorded aftershock sequences and averaged parameters between sequences. Table 2 provides an overview of the model parameters. Figure 10 compares the mean rates of both models for different mainshock magnitudes and completeness magnitude M_0. The Reasenberg-Jones model has consistently higher rates. The gap between the two models narrows for increasing time due to the difference in p – value. Figure 11 shows the expected number of earthquakes in the previous time interval for different mainshock and completeness magnitudes. The data at point 1 day cover the time interval 0.1 – 1 days. The data subsequent days always cover the previous 24 hours period. We note two key observations: The numbers predicted by the Reasenberg and Jones model are about double the numbers predicted by the mean abundance model. Though the completeness magnitudes are chosen optimistically, i.e. the completeness in the first 1 or 2 days following a magnitude 7.0 earthquake is unlikely to be as low as 3.95 in the real time catalog, the number of expected earthquakes is fairly small. A longer time period or a lower completeness magnitude would be required to observe at least 100 earthquakes so that the STEP model would begin to calculate sequence specific parameters.

Here two examples of the effects of the different aftershock descriptions in the STEP model. Example 1 for a large mainshocks with more than the average number of aftershocks but a very high completeness magnitudes in the first couple of days following the mainshock. Example 2 for a moderate mainshock with less than the average number of aftershocks.

3.8 Implementation of the abundance model into STEP for real-time application

Still to finalize in discussion with Matt

3.9 Discussion

We have proposed a variation to the STEP model for California, firstly to change the generic model parameters and secondly to assess the productivity of an on-going sequence when limited data is available. The key assumption for our second change is that Omori's law for aftershock decays also applies to sequences with few aftershocks and that the number of aftershocks observed in the first day following the mainshocks can be used to derive the parameters *K*. As the model will be up-dated hourly, aftershock activity following a major aftershock can be accounted for. We have not tested our model variations, only illustrated them with some data from the SCEDC catalog. We are planning to introduce our model to the RELM model testing, as well as calculating region specific parameters to include the model in testing centers that are starting in Europe and New Zealand. We expect the new model to perform better than STEP for small sequences and in the early stages of a large earthquake when limited aftershock data is available. For larger sequences, there will be no difference because the model will follow STEP procedures in up-dating sequence specific parameters. If the new model performs successfully in the testing, then it will find useful application in regions where the seismic network is not as advanced as in California. Some of our results warrant some more in depth discussion as follows:

Magnitude completeness

The up-dating of the model in the early stages of an on-going sequence relies heavily on the completeness magnitude. There will be not enough data to estimate the completeness magnitude and thus some assumptions are required. No good data on real time earthquake catalogs exists because catalogs are usually up-dated when earthquakes have been reviewed by seismologists and corrections have been made. Thus the testing will provide some valuable insight into the quality of the real time data. However, seismic networks continually change and the magnitude completeness will generally vary with the location of the earthquakes.

The difference between new and generic model parameters

The generic parameters predict on average about twice as many aftershocks than the abundance model. This is because the generic parameters were determined from 62 well recorded and thus abundant aftershock sequences (Reasenberg and Jones, 1989) while we included all earthquakes in the catalog in our analysis. The generic *b*-value of 0.91 is smaller than the *b*-value of 1.0 that we found. One possible explanation is that the magnitude completeness was actually higher than assumed and thus the *b*-value was underestimated.

Some more discussion on our implementation.

3.10 Conclusion

To be finalized once all sections are completed.

3.11 Acknowledgements

The study has been partly funded by a grant from the New Zealand Earthquake Commission Research Foundation.

3.12 References

Console, R., M. Murru, and A. M. Lombardi (2003), *Refining earthquake clustering models*, J. Geophys. Res. 108, doi:10.1029/2002JB002130.

Christophersen, A. (2000), *The probability of a damaging earthquake following a damaging earthquake*, PhD Thesis, Victoria University of Wellington.

Christophersen, A. and E.G.C. Smith (2006), *Foreshock probabilities from aftershock characteristics*, manuscript submitted to Bull. Seism. Soc. Am..

Gerstenberger, M.C. (2003), *Earthquake clustering and time-dependent probabilistic seismic hazard analysis for California*, Dissertation submitted to the Swiss Federal Institute of Technology for the degree of Doctor of Science.

Gerstenberger, M.C., Wiemer, S., and L.M. Jones (2004), *Real-time forecasts of tomorrow's earthquakes in California: A New Mapping Tool*, U.S. Geological Survey, open-file report 2004-1390.

Gerstenberger, M.C., Wiemer, S., L.M. Jones, and P.A. Reasenberg (2005), *Real-time forecasts of tomorrow's earthquakes in California*, Nature 435, doi:10.1038/ nature03622.

Guo, Z. and Y. Ogata (1997), *Statistical relations between the parameters of aftershocks in time, space and magnitude*, J. Geophys. Res., 102, 2857-2873.

Gutenberg, B., and C. F. Richter (1949), *Seismicity of the earth and associated phenomena*, Princeton University Press.

Helmstetter, A. (2003), *Is earthquake triggering driven by small earthquakes?*, Phys. Rev. Lett. 91, DOI: 10.1103/PhysRevLett.91.058501.

Kagan, Y.Y. (2004), *Short-term properties of earthquake catalogs and models of earthquake source*, Bull. Seism. Soc. Am. 94, 1207-1228.

Reasenberg, P.A., and L.M. Jones (1989), *Earthquake hazard after a mainshock in California*, Science 243, 1173-1176.

Reasenberg, P.A., and L.M. Jones (1990), *California aftershock hazard forecast*, Science 247, 345-346.

Reasenberg, P.A., and L.M. Jones (1994), *Earthquake Aftershocks: Update*, Science 265, 1251-1251.

Richter, C.F. (1958), *Elementary Seismology*, Freeman.

Schorlemmer, D. and M.C. Gerstenberger (2005), *RELM Testing Center*, submitted to Seismological Research Letters.

Schorlemmer, D., M.C. Gerstenberger, S. Wiemer, and D. D. Jackson (2005), *Earthquake Likelihood model Testing*, submitted to Seismological Research Letters.

Smith, E.G.C. and A. Christophersen (2006), *A new formulation of Omori's law for aftershock decay and its application in models for earthquake occurrence conditional on the time since the last earthquake*, manuscript in preparation.

Utsu, T. (1969), *Aftershocks and earthquake statistics (I) Source parameters which characterize an aftershock sequence and their interrelations,* J.Fac. Sci., Hokkaido Univ., Ser. VII 3, 129-195.

Utsu, T., Y. Ogata, and R. S. Matsu'ura (1995), *The Centenary of The Omori Formula for a decay law of aftershock activity*, J. Phys. Earth 43, 1–33.

Vere-Jones, D., J. Murakami, and A. Christophersen, (2006), *A Further Note on Båth's Law*, manuscript submitted to Bull. Seism. Soc. Am.

Wells, D. L. and K. L. Coppersmith, (1994), *New empirical relationships among magnitude, rupture width, rupture area, and surface displacement*, Bull. Seism. Soc. Am., 84, 974-1002.

Woessner, J., and S. Wiemer (2005), *Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty,* Bull. Seism. Soc. Am. 95, 684-698, doi: 10.1785/0120040007.

3.13 Tables

Table 1: Changes in mean abundance in different time intervals for $p = 0.89$ and $p = 1.08$ in respect to $p = 1.0$.

Time interval in days	$p = 0.89$	$p = 1.0$
0.1 – 1	88%	108%
1 – 2	100%	99%
2 – 3	106%	95%
3 – 4	109%	93%
4 – 5	113%	91%

Table 2: Comparison of model parameters for the mean abundance model and the Reasenberg-Jones model (1989)

Model parameter	Abundance model	Reasenberg & Jones
b - value	1.0	0.91
p - value	1.0	1.08
tc	0.1	0.05
K	Mean abundance/ [ln(10.1)-ln(0.2)]	$10^{a'+b(M-Mo)}$

3.14 Figure captions

Figure 1: A Figure 1: Map of the SCEDC earthquake catalog: Green dots: 1932-1976 M>3.5; Blue dots: 1981-1999 M>2.5

Figure 2: Magnitude-frequency distribution for two different time periods (1932 – 1976 in red and 1981 – 1999 in black; crosses are cumulative data).

Figure 3: Magnitude-frequency of mainshocks, SCEDC 1932-76 (diamonds) and 1981-99 (crosses).

Figure 4: The average aftershock magnitude in logarithmic time intervals; (a) for the mainshocks in the magnitude range 5.95 – 6.95 (red) and 6.95 – 7.45 (black) and (b) for mainshocks in the magnitude range 2.95 – 3.95 (green), 3.95 – 4.95 (red) and 4.95 – 5.95 (black). Also shown is are the one standard deviations.

Figure 5: b-value and 95% confidence intervals for aftershocks, within the time interval 0.1 - 10 days of mainshocks and the mainshock magnitude range 2.95 – 5.95, (a) SCEDC 1981-1999, N=14,860 for $M \geq 2.0$ and N=7,076 for $M \geq 2.5$ and (b) SCEDC 1932-1976 N=1,958 for $M \geq 3.0$ and N=417 for $M \geq 4.0$.

Figure 6: Mean abundance in time intervals [0.1, 10] (solid diamonds) for mainshocks in between 1981 and 1999 and aftershock cut-off 2.45, and best fitting model (solid line).

Figure 7: Mean abundance for magnitude completeness 4.0 within 10 days,
SCEDC 1932-76, model from 1981-99 data (solid line) and least squares fit (dashed line).

Figure 8 (a) – (d): The range of observed abundances as a function of mainshock magnitude in the range 2.95 – 5.95 for different completeness magnitudes.

Figure 9: Average length of longer ellipse axis and comparison to search radius.

Figure 10: Comparison of observed and predicted number of aftershocks in different time intervals.

Figure 11: The mean aftershocks rates predicted from the abundance model, dashed lines, in comparison to Reasenberg-Jones (1989), solid lines, for mainshock magnitude. 4.0, 5.0 and 6.0.

Figure 12: The mean number of aftershocks rates predicted from the abundance model, dashed lines, in comparison to Reasenberg-Jones (1989), solid lines, for mainshock magnitude. 4.0, 5.0 and 6.0.

3.15 Figures

Figure 1

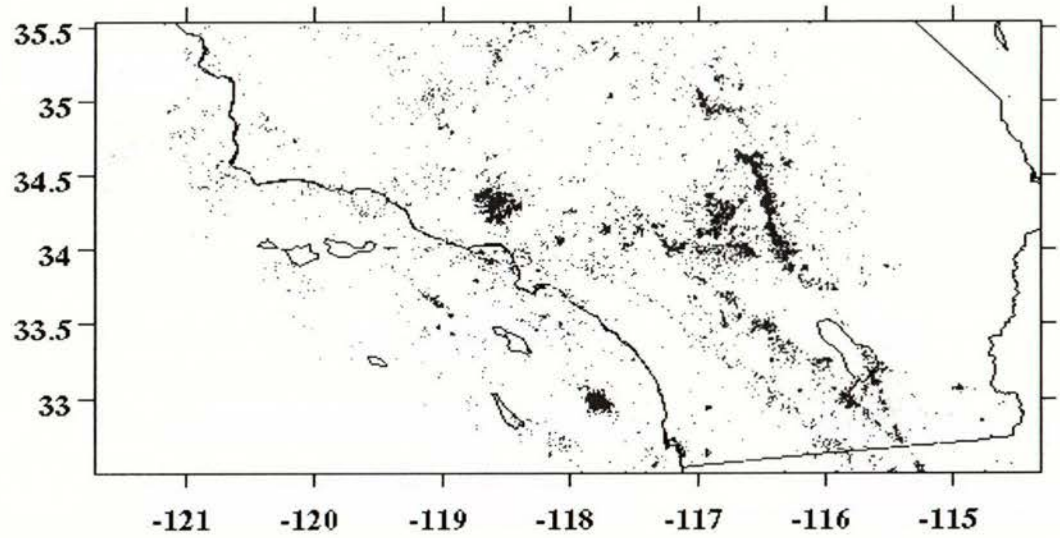


Figure 2

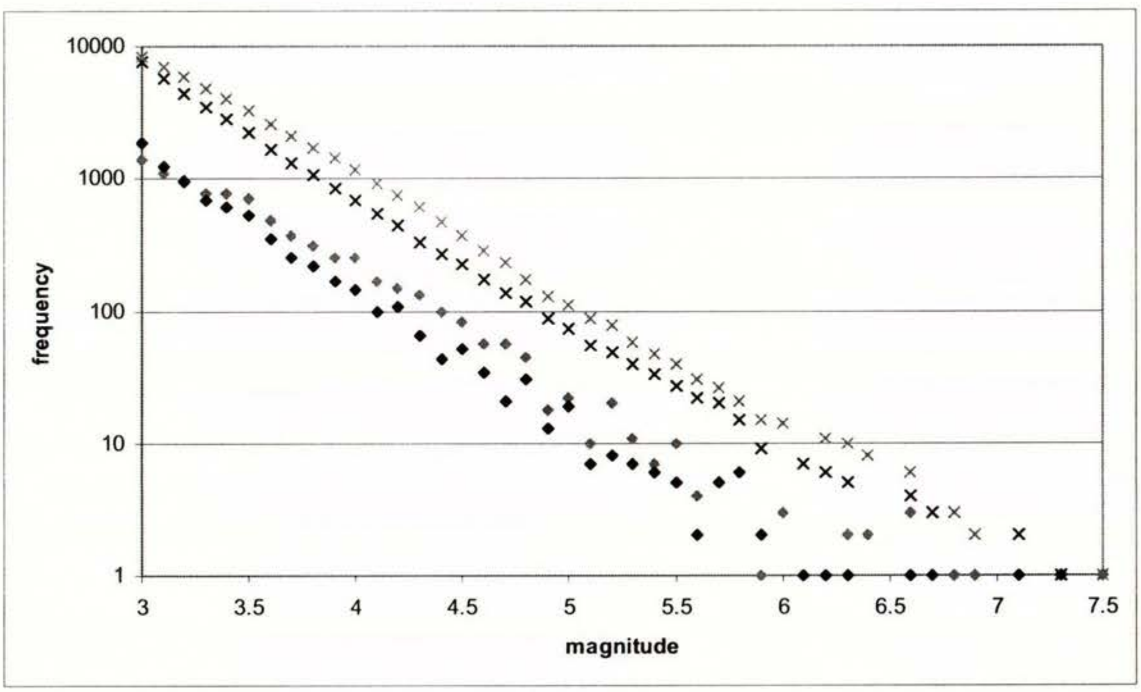


Figure 3

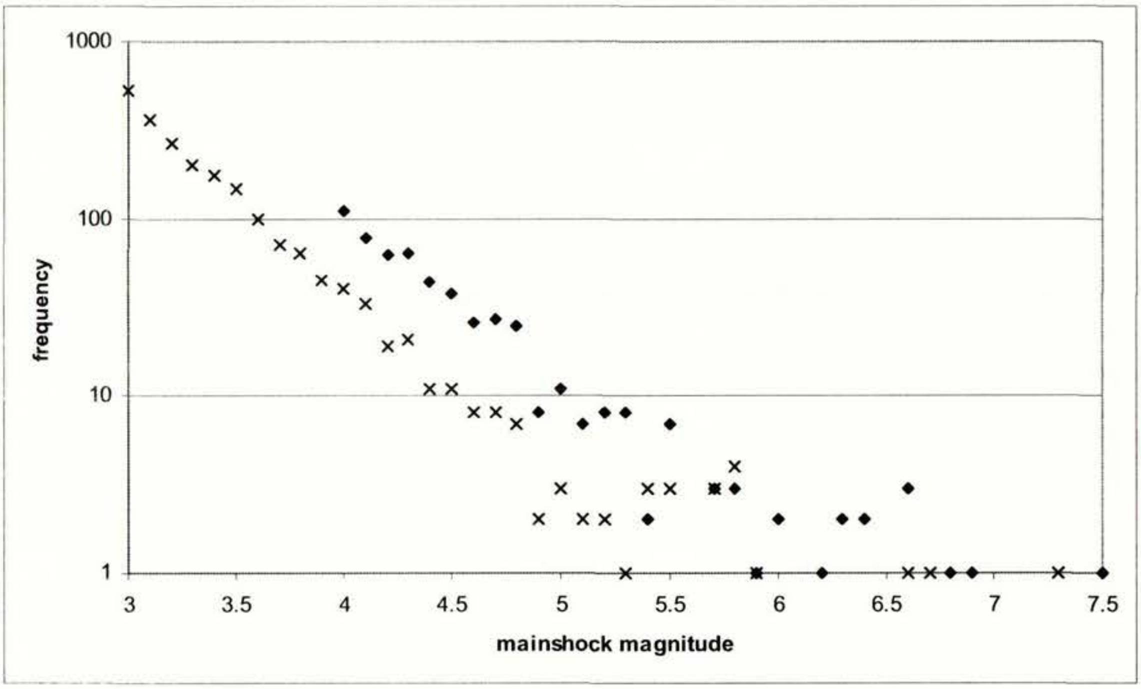


Figure 4 (a)

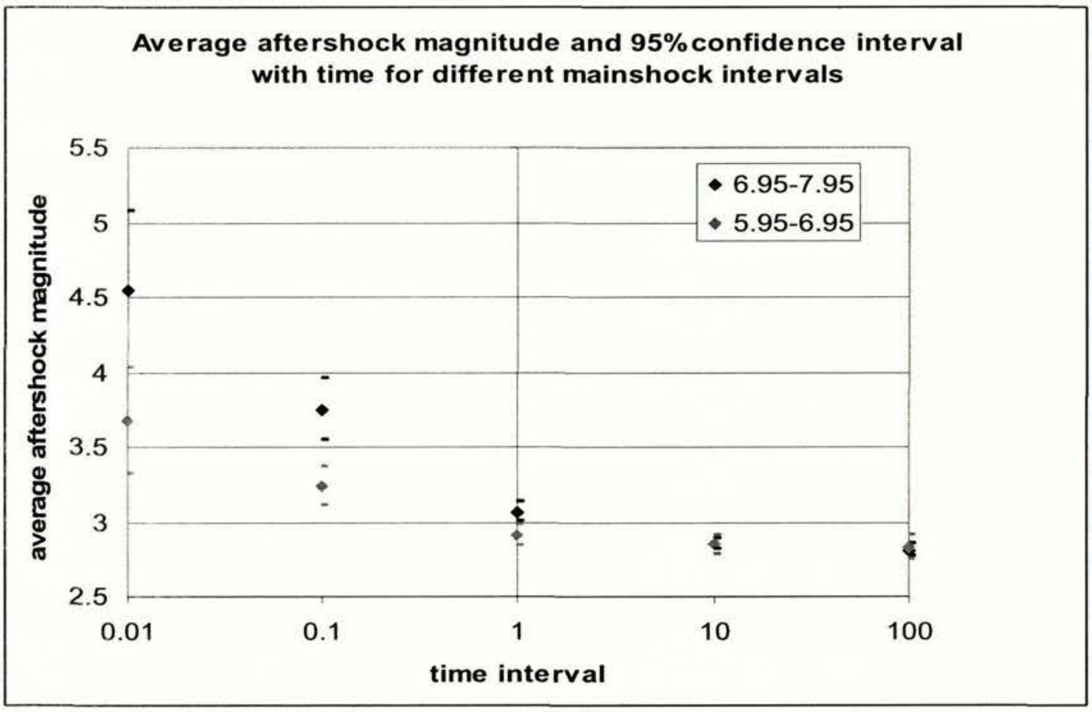


Figure 4 (b)

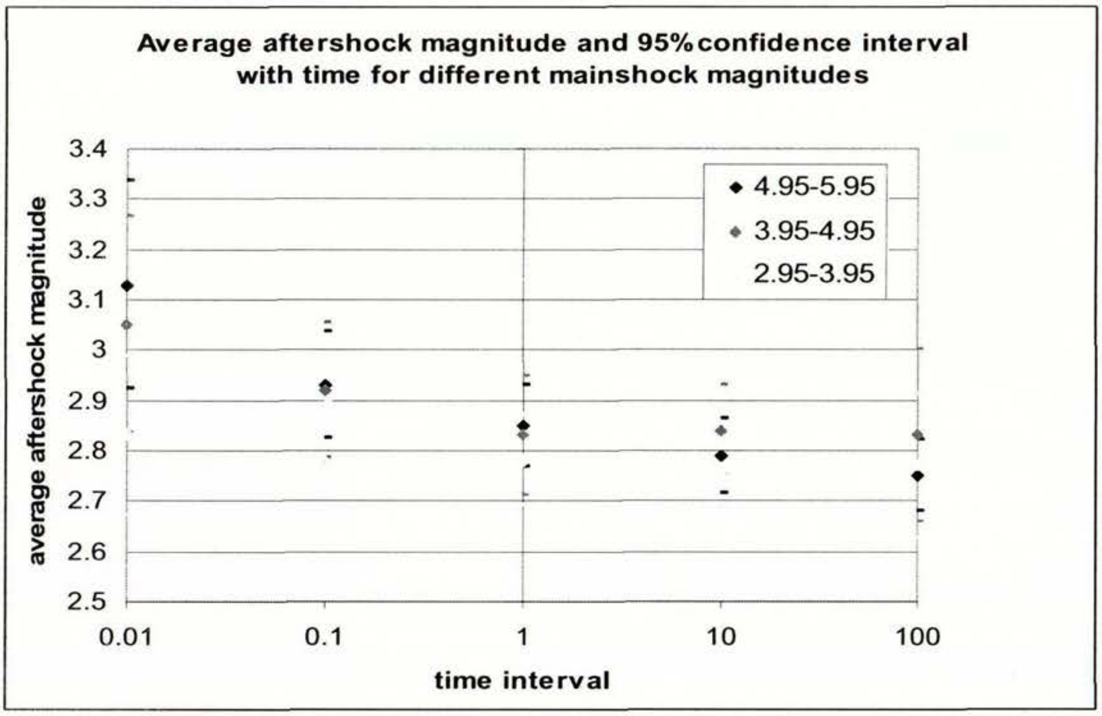


Figure 5(a)

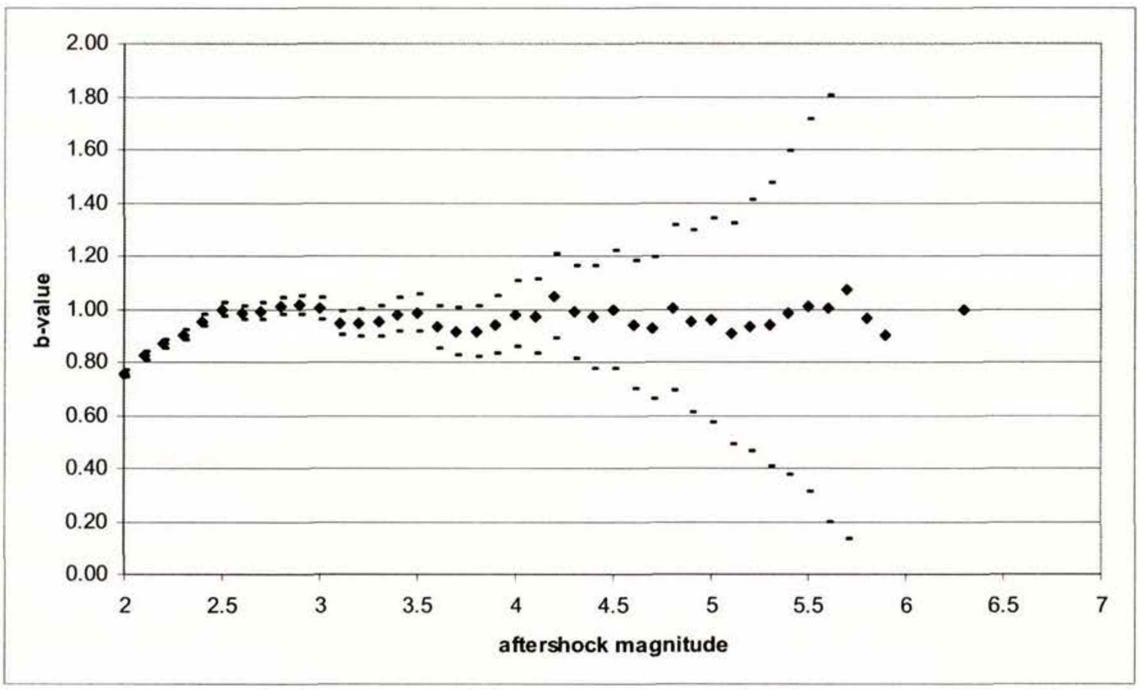


Figure 5(b)

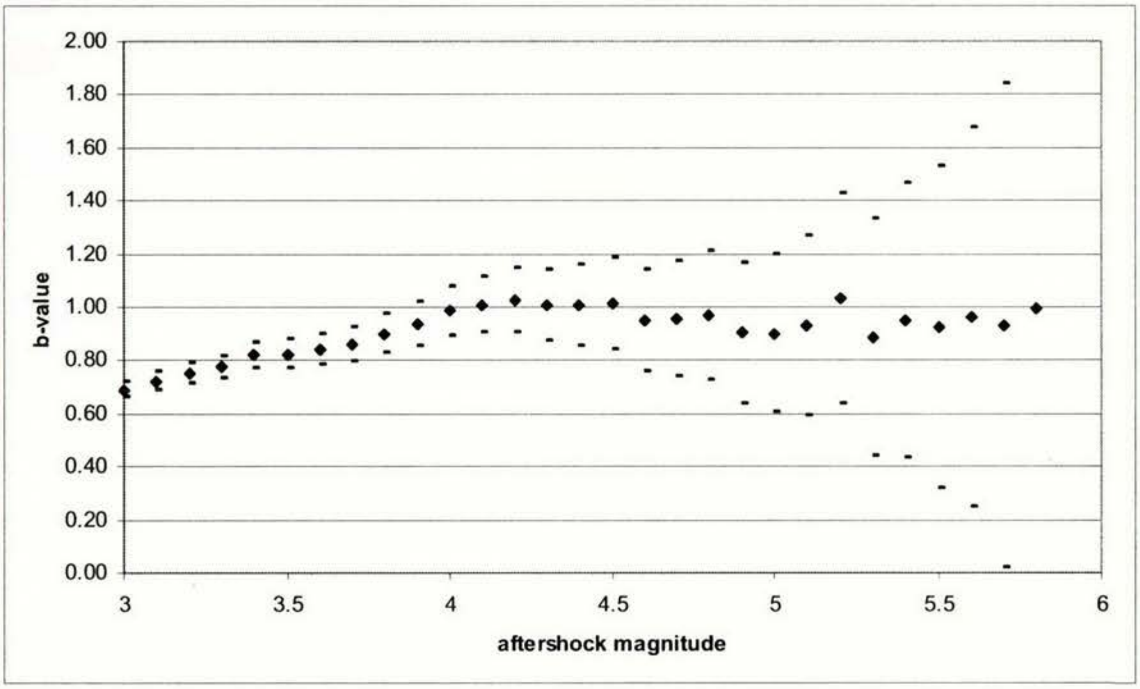


Figure 6

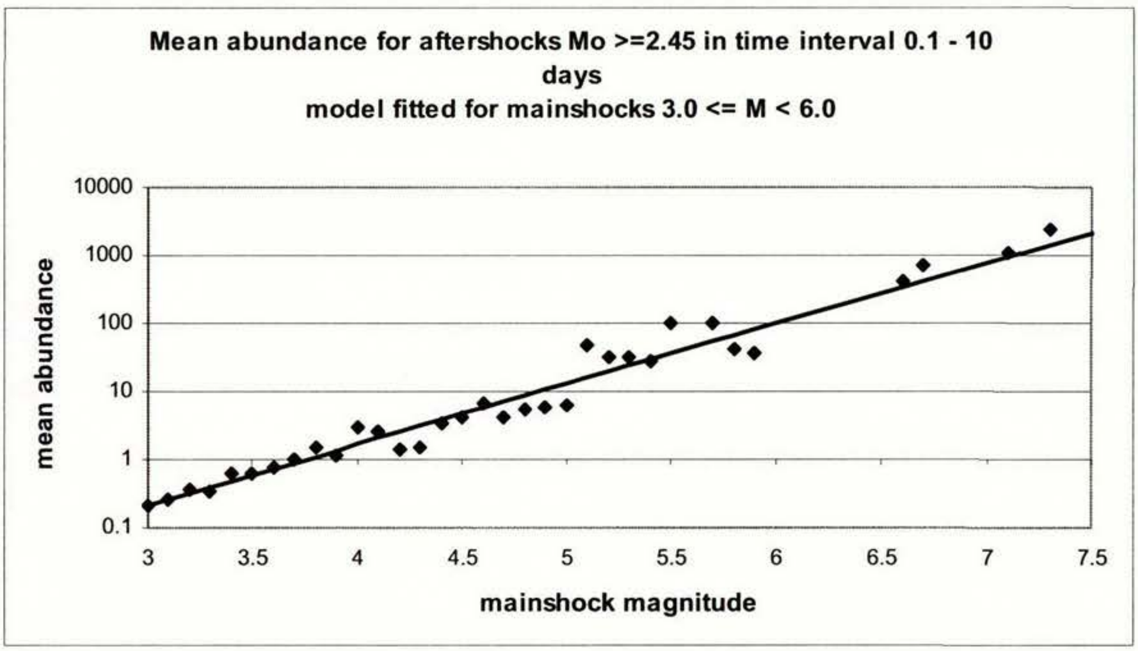


Figure 7

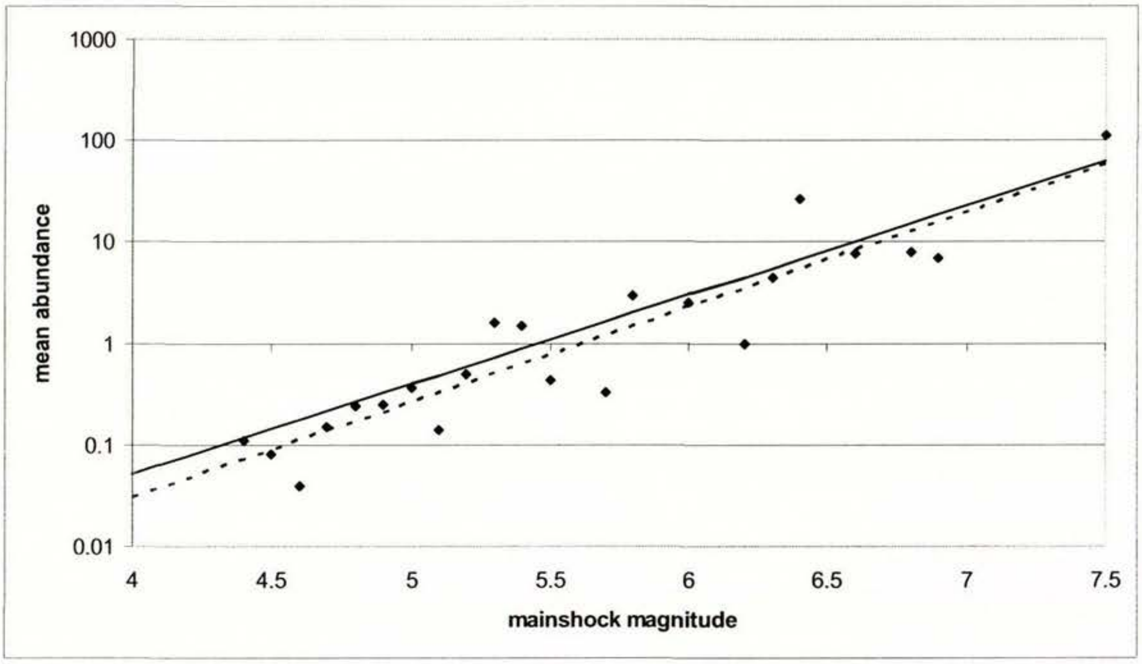


Figure 8 (a)

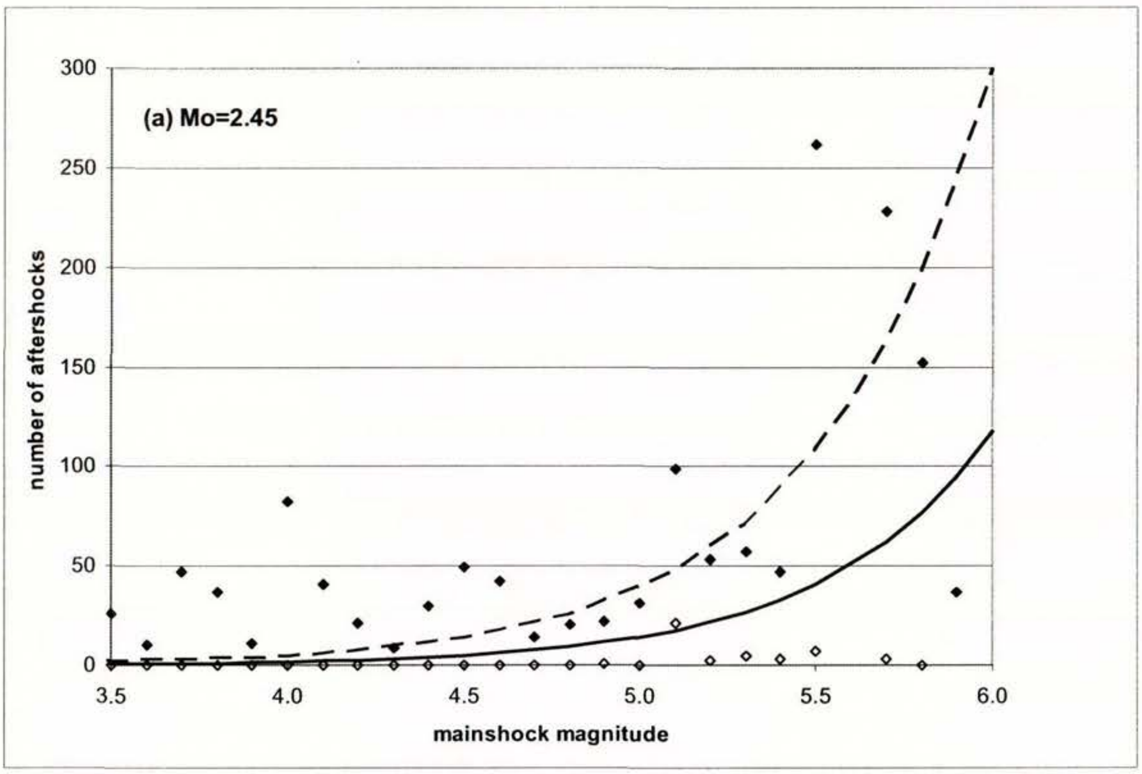


Figure 8 (b)

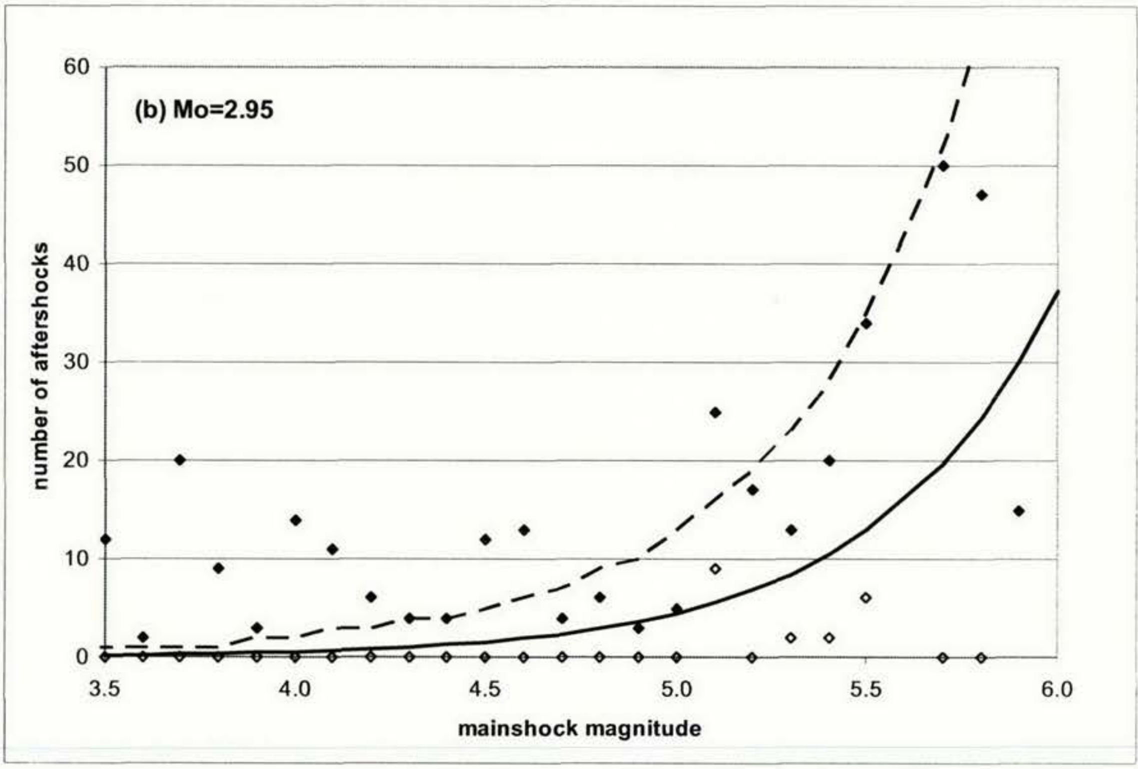


Figure 8 (c)

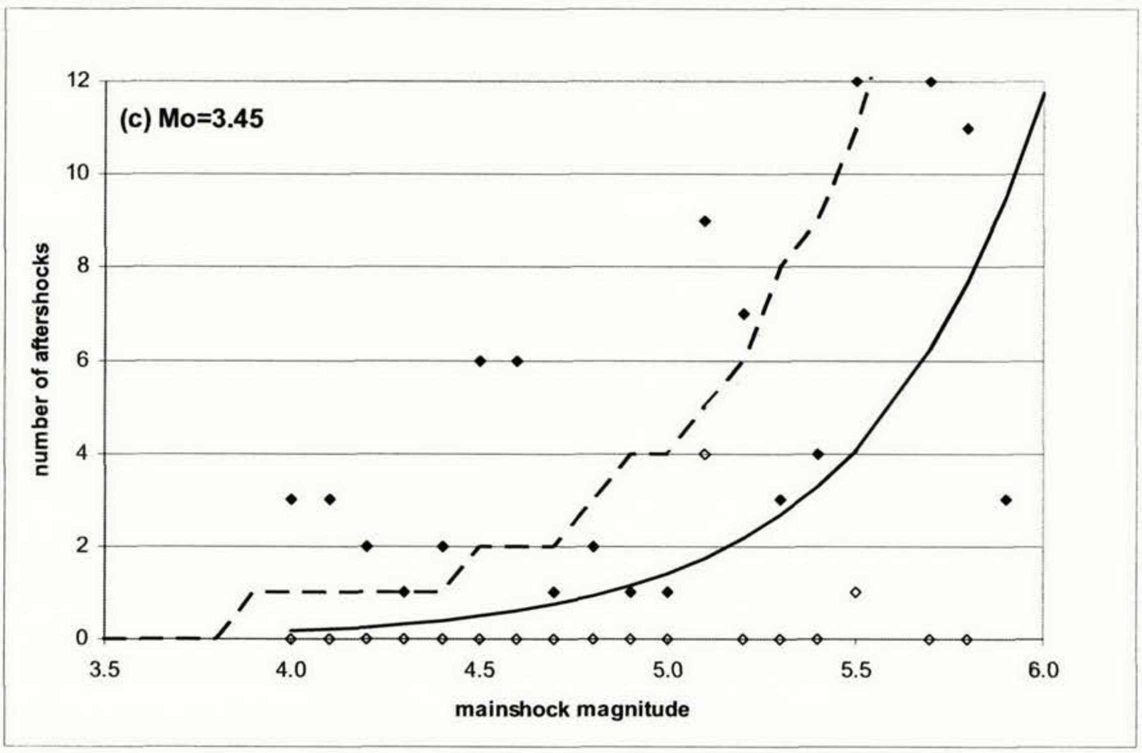


Figure 8 (d)

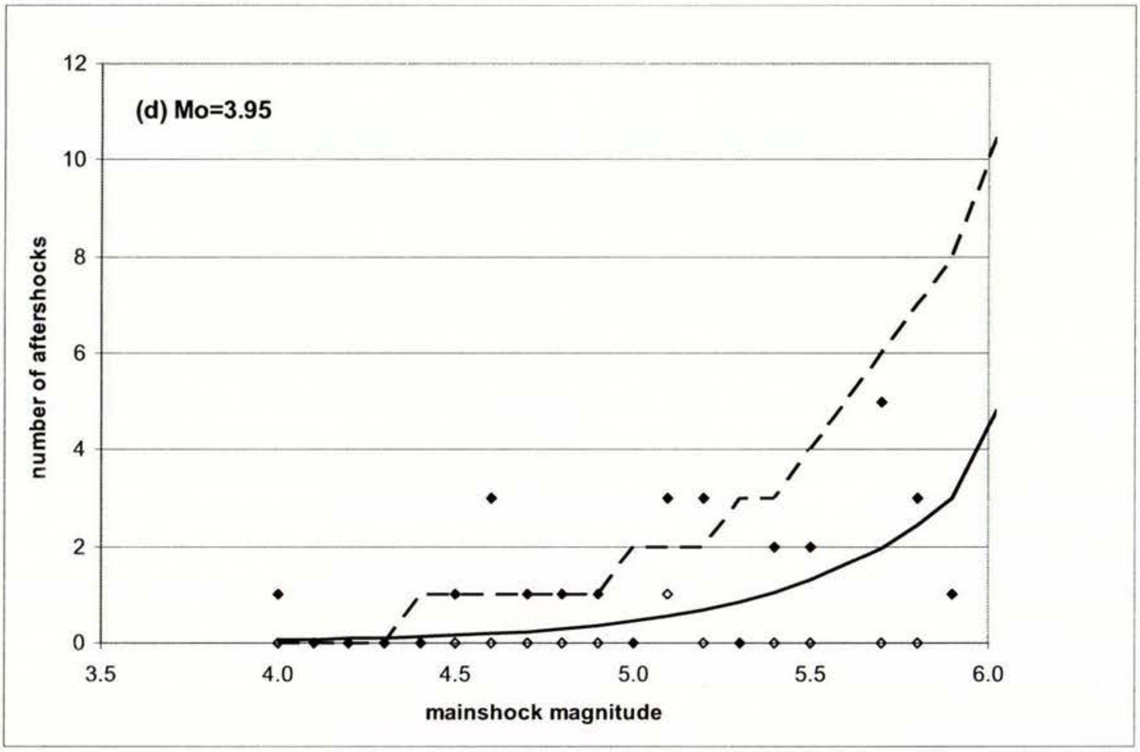


Figure 9

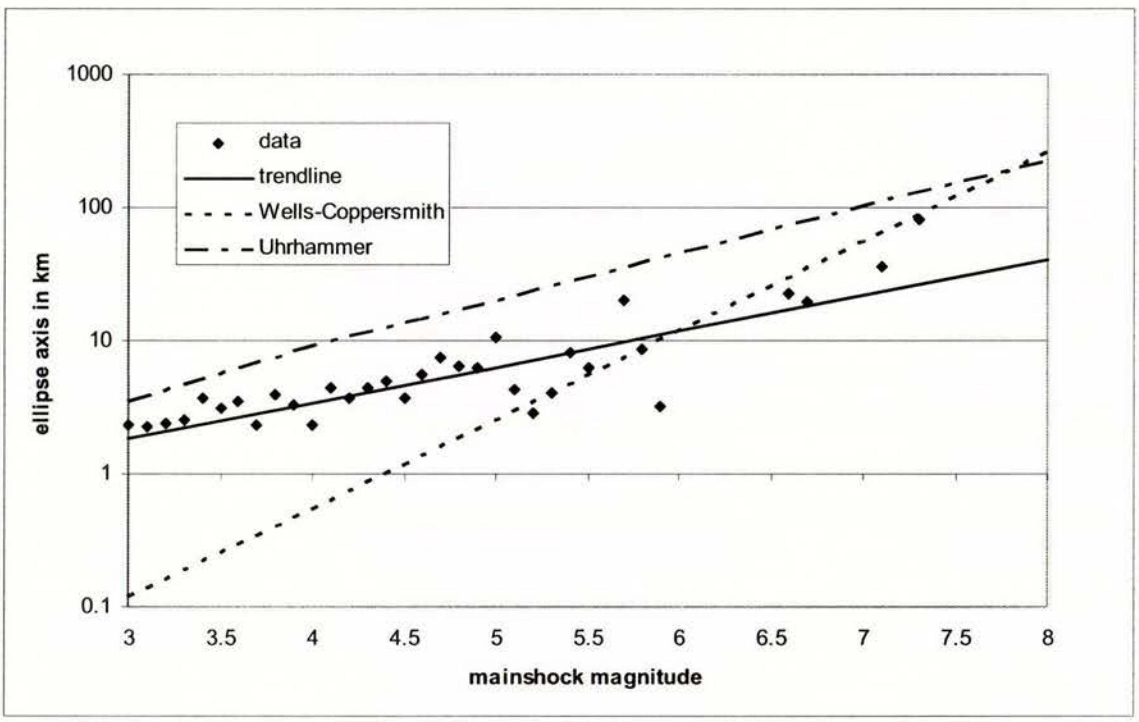


Figure 10(a)

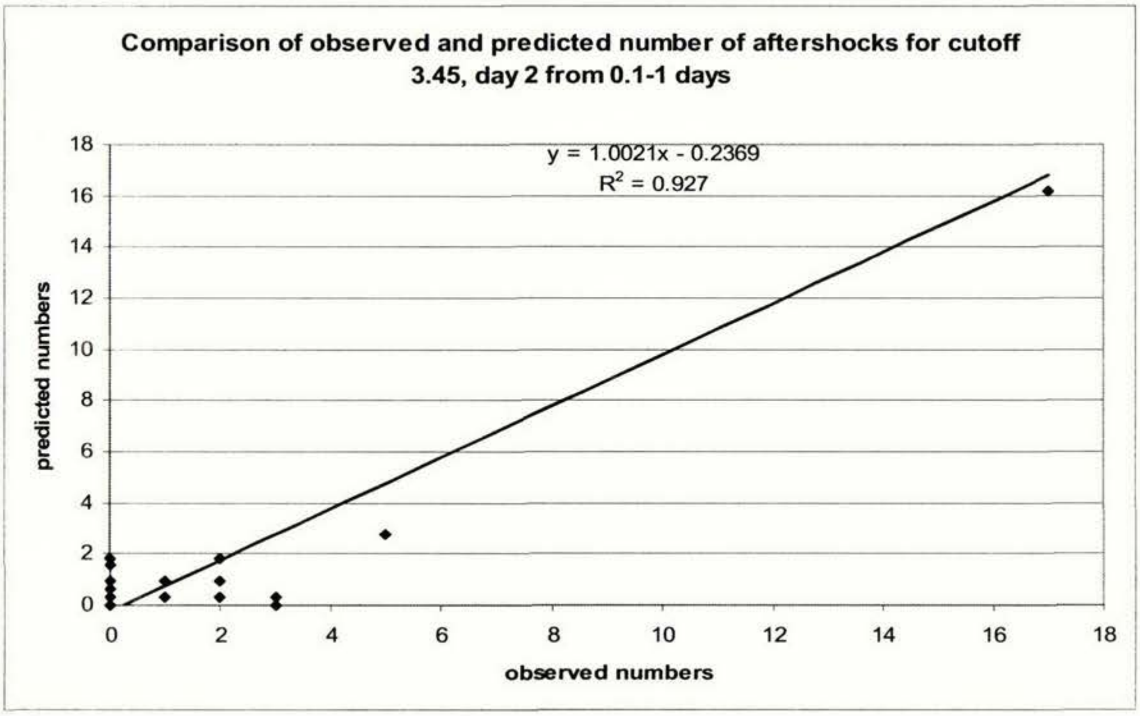


Figure 10(b)

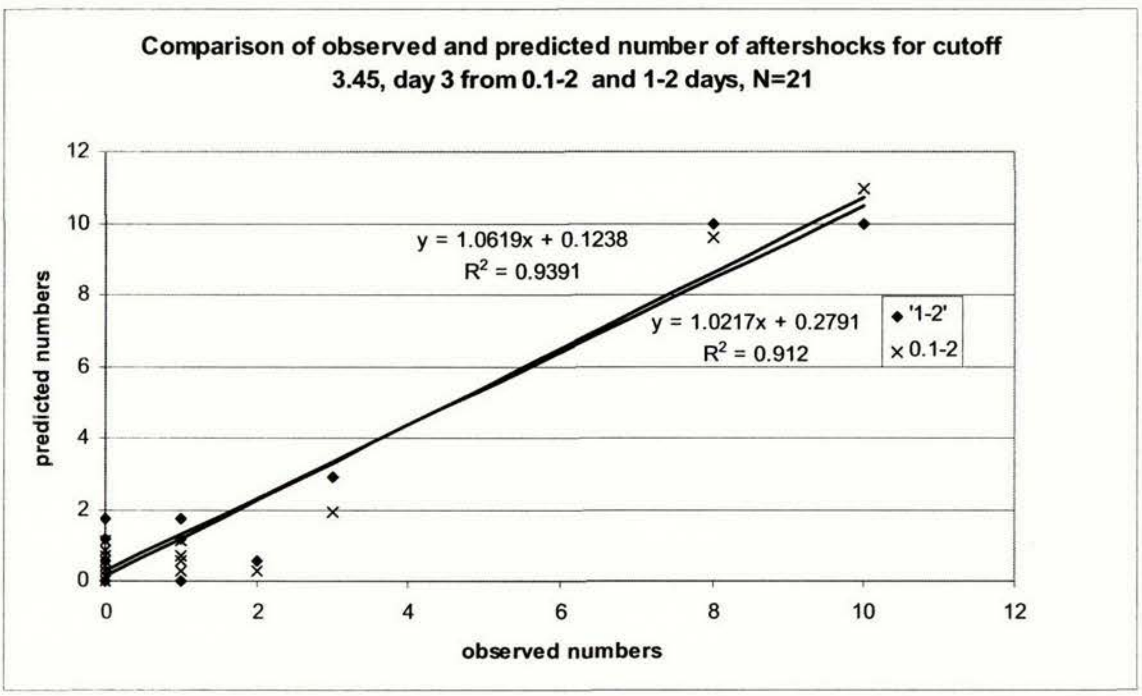


Figure 10(c)

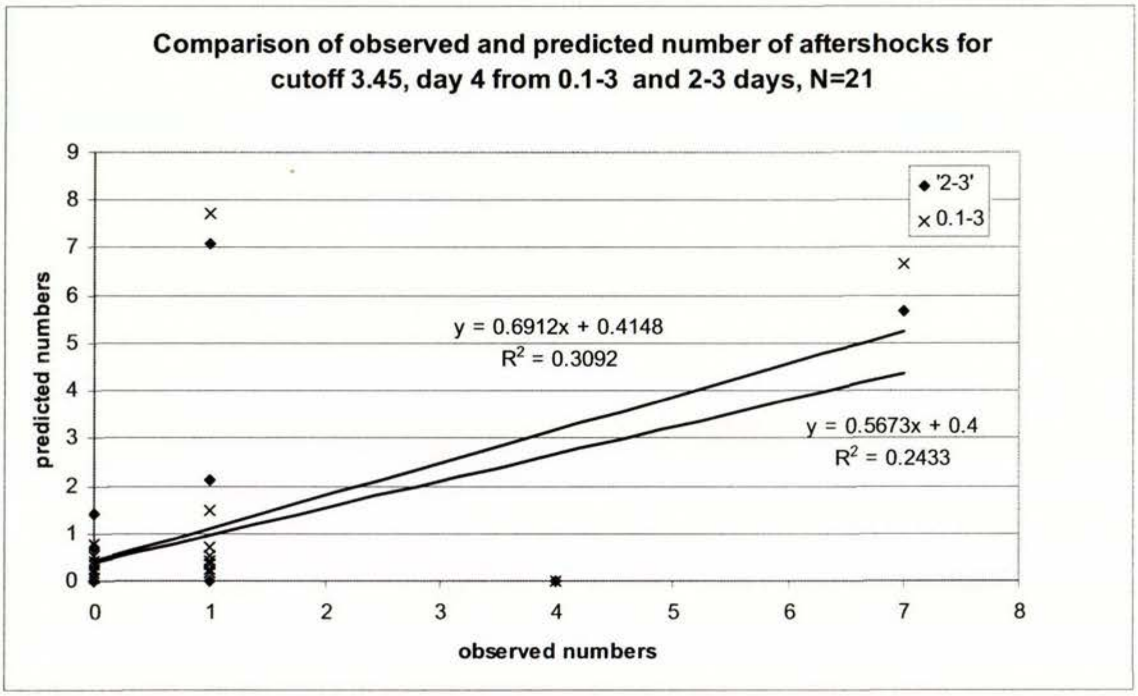


Figure 11

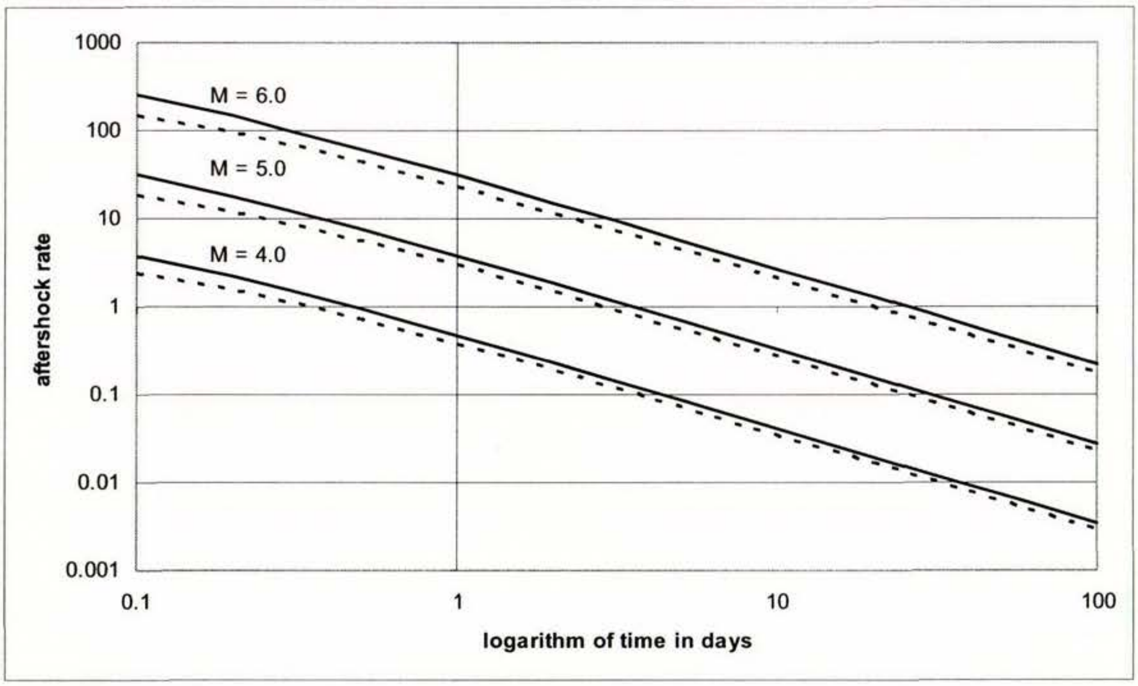
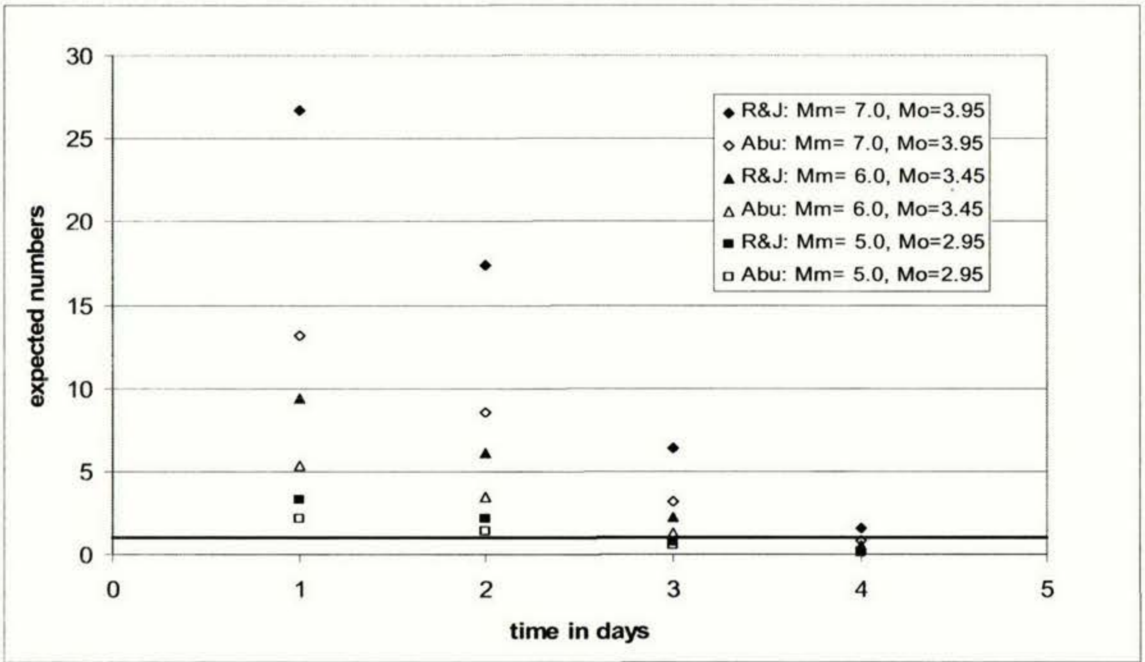


Figure 12



30
25
20
15
10
5
0
expected numbers
0
1
2
3
4
5
time in days
R&J: Mm= 7.0, Mo=3.95
Abu: Mm= 7.0, Mo=3.95
R&J: Mm= 6.0, Mo=3.45
Abu: Mm= 6.0, Mo=3.45
R&J: Mm= 5.0, Mo=2.95
Abu: Mm= 5.0, Mo=2.95

Appendix A: References

Christophersen, A. (2000), *The probability of a damaging earthquake following a damaging earthquake*, PhD Thesis, Victoria University of Wellington.

Christophersen, A. (2005a), *Towards a New Zealand model for short-term earthquake probability: Aftershock productivity and parameters from global catalogue analysis*, Second progress report, EQC Project No 6OPR1B, June 2005.

Christophersen, A. (2005b), *Towards a New Zealand model for short-term earthquake probability: Aftershock productivity and parameters from global catalogue analysis*, Third progress report, EQC Project No 6OPR1B, October 2005.

Christophersen, A. and E. Smith (2004) *Triggered earthquake probability forecasting: The development of a new model for earthquake recurrence*, A report to the New Zealand Earthquake Commission, EQC Project 03/485.

Gerstenberger, M.C., Wiemer, S., L.M. Jones, and P.A. Reasenberg (2005), *Real-time forecasts of tomorrow's earthquakes in California*, Nature 435, doi:10.1038/ nature03622.

Schorlemmer, D. and M.C. Gerstenberger (2005), *RELM Testing Center*, submitted to Seismological Research Letters.

Schorlemmer, D., M.C. Gerstenberger, S. Wiemer, and D. D. Jackson (2005), *Earthquake Likelihood model Testing*, submitted to Seismological Research Letters.

Schorlemmer, D., M. Gerstenberger, T. Jordan, E. Field, S. Wiemer, L. Jones, and D. D. Jackson (2006), *From the Testing Center of Regional Earthquake Likelihood Models to the Collaboratory for the Study of Earthquake Predictability*, presentation at the fourth workshop of Statistical Seismology, Tokyo, Japan.

Smith, E.G.C. and A. Christophersen (2006), *A new formulation of Omori's law for aftershock decay and its application in models for earthquake occurrence conditional on the time since the last earthquake*, manuscript in preparation.

Tormann, T. (2005) *Dynamic seismic hazard model for New Zealand*, Master Thesis, Victoria University of Wellington.