S&G3716

VICTORIA UNIVERSITY OF WELLINGTON

Te Whare Wānanga o te Ūpoko o te Ika a Māui

SCHOOL OF EARTH SCIENCES

Te Kura Tātai Aro Whenua

• Geology • Institute of Geography • Institute of Geophysics

in association with

Antarctic Research Centre

Report to the New Zealand Earthquake Commission

EQC Project 03/485

Triggered earthquake probability forecasting: The development of a new model for earthquake recurrence.

Annemarie Christophersen & Euan Smith

School of Earth Sciences

Victoria University of Wellington

Contents

Executive Summary

This project set out to determine the probability of occurrence, on a global scale, of large earthquakes being triggered by large earthquakes; 'triggered earthquakes' being those occurring close in time and space to the parent earthquake but not close enough to be considered aftershocks.

The occurrence of such triggered events was investigated by removing aftershocks from a global catalogue of $M \geq 5$ events recorded between 1964-2000. We built two databases of earthquakes close enough together in space that the assumption that they were physically connected was reasonable. We called the two datasets *Superclustres without aftershocks* and *Superclusters with aftershocks*, the former having been constructed from the aftershock-less catalogue.

It was found that the distinction between 'aftershocks' and 'triggered events' was arbitrary. Consequently we make no distinction between these two types of post-mainshock earthquake.

The collection of events in all the superclusters without aftershocks is equivalent to a catalogue of mainshocks. This was examined for differences in rate of earthquake occurrence before and after large earthquakes. None were found. The times of occurrence of the mainshocks in the catalogue could be modelled well by a random (Poisson) process.

The global database also offers an opportunity to investigate the phenomenon reported by Evison and Rhoades, which they term Precursory Scale Increase (of earthquake activity before a large earthquake) - PSI. A preliminary investigation of some of their examples of PSI suggests that the phenomenon should be systematically investigated using the global catalogue. The resources of the current project were insufficient to do this.

To investigate the distribution of times between large earthquakes, we used the dataset of superclusters with aftershocks. We examined the distribution of interevent times, Δt, between all the earthquakes with $M \geq 6.5$ in any supercluster with aftershocks. It consisted of two parts. The early part corresponds to the well-established Omori law for aftershocks. The later part is well modelled by a random process. Accordingly, we propose that the probability of the time to the next earthquake in a region can be calculated with a mixture of an Omori law, suitably modified to deal with very short and very long times, and a random process. This model can be fitted to any region's data by the standard method of maximum likelihood.

The form of the model allows calculation of the mean time between large earthquakes, and probabilities for times to the next earthquake given the time to the last earthquake. Examples of the latter are given. These derivative results have direct applications in earthquake hazard and risk assessment.

Application of the model to the global database produced the following results and conclusions:

- After an earthquake, about 40-50% of the subsequent earthquakes can be regarded as being aftershocks, and the rest as 'new' earthquakes. This is a collective property, and one cannot say with certainty whether a particular subsequent earthquake is an aftershock or a new earthquake.
- Models calculated for different ranges of magnitudes were consistent with each other.

- The parameters that described the particular features of the model for a particular region have large uncertainties. This means that it will be difficult to obtain a good model for a region that has a limited record of large earthquakes.

An attempt to apply the model to the occurrence of $M \geq 7$ earthquakes in central New Zealand since 1840 had the following consequences:

- With only 15 events, producing 14 interevent times, the model parameters had very large uncertainties.
- The model had rather different parameters than the global models. We speculate that the New Zealand catalogue of $M \geq 7$ earthquakes *is not complete*, as normally assumed. In particular we speculate that large aftershocks of pre-instrumental (pre-1900) earthquakes may not have been identified from the historical record. The 1848 Marlborough earthquakes are known to be a series in which several aftershocks are large, and it is quite possible that the 1855 earthquake had one or more aftershocks of $M > 7$. If true, this absence will have distorted the model we calculated.

The greatest weakness in our modelling is the assumption of physical connectedness between all events in a supercluster. The models will not be able to be used in practice until this weakness is overcome. There are several suggestions in the literature as to how to proceed with the development of an appropriate spatial connection model.

In light of the results of this project, the following additional work should be undertaken:

1. Development of a correction to the spatial-connection assumption, as noted immediately above.
2. After completing 1: determine a model for New Zealand using a suitable dataset or datasets e.g. $M \geq 6$ earthquakes since 1940. The model can then be tested as an alternative to the current model in seismic hazard and risk assessments.
3. Systematically investigate the occurrence of the PSI phenomenon in the global database, to determine what its characteristic manifestations are in this database.
4. Elucidation of what is the underlying physical process and connection between successive earthquakes in a region, and why 'new' earthquakes appear to occur randomly in time.

Technical Abstract

This project set out to determine the probability, on a global scale, of large earthquakes being triggered by large earthquakes; 'triggered earthquakes' being those occurring close in time and space to the parent earthquake but not close enough to be considered aftershocks.

The occurrence of such triggered events was investigated by removing aftershocks – 'declustering' – from a global catalogue of $M \geq 5$ events recorded between 1964-2000. We term the set of spatially proximal events from which aftershocks and triggered events have been removed a *Supercluster without aftershocks*. Subsets of the global catalogue with the same spatial limits as superclusters without aftershocks but from which aftershocks and triggered events have not been removed are termed *Superclusters with aftershocks*.

It was found that the distinction between 'aftershocks' and 'triggered events' was arbitrary. Consequently we make no distinction between these two types of post-mainshock earthquake.

The collection of events in all the superclusters without aftershocks is equivalent to a catalogue of mainshocks. This was examined for differences in rate of earthquake occurrence before and after large earthquakes. None were found. The times of occurrence of the mainshocks in the catalogue could be modelled well by a random (Poisson) process.

The distribution of magnitudes in the whole catalogue followed the Gutenberg and Richter law with a *b*-value of 1.0, as did the distribution of mainshock magnitudes. However, the set comprising the largest earthquake in each of the superclusters is biased in favour of large events. Consequently the b-value for this dataset should be expected to be < 1.0, and was found to be 0.84 ± 0.02.

The global database also offers an opportunity to investigate the phenomenon reported by Evison and Rhoades, which they term Precursory Scale Increase (of earthquake activity before a large earthquake) - PSI. A preliminary investigation of some of their examples of PSI suggests that the phenomenon should be systematically investigated using the global catalogue. The resources of the current project were insufficient to do this.

To investigate the distribution of times between large earthquakes, we used the dataset of superclusters with aftershocks. We assume, for the purpose of subsequent modelling, that all earthquakes within a supercluster so defined are sufficiently close in space for there to be an (unspecified) physical connection between the earthquakes.

We examined the cumulative distribution of interevent times, Δt, between all the earthquakes with $M \geq$ a minimum magnitude (M_r) of 6.5 in any supercluster with aftershocks. It consisted of two parts discernible as straight-line segments when Δt was plotted on a logarithmic scale. The early part corresponds to the Omori law for aftershocks, i.e. the probability density function for Δt, $f(\Delta t) \sim 1/\Delta t$. After making a correction for the finite duration of the catalogue, the later part could not be modelled by a power law $1/t^p$, but is well modelled by a Poisson process. Accordingly, we propose that the time interval between successive earthquakes in a region with $M \geq M_r$ can be modelled with a probability distribution that is a mixture of an Omori law, suitably modified to deal with very short and very long times, and a Poisson process.

The model has four parameters: a short and a long-time constant for the Omori law, the time constant for the Poisson process and the relative weights of the two parts. The short-time constant is likely to be poorly resolvable in practice because of known difficulties with data, but the model is not sensitive to this parameter. The long time constant for the Omori law appears to be less than or equal to the time constant for the Poisson process. It is poorly resolved but the model is not sensitive to its value. This leaves the two controlling parameters beings the relative weight and the Poisson time constant. These can be readily fitted to any dataset of interevent times by maximum likelihood.

The quality of the fit of the model to the long ineterevent times, i.e. those with a low probability of being aftershocks, strongly suggests that the Poisson part of the model is more than just a default for an alternative process, and that the occurrence of 'new' events is truly Poisson (random). If this is so, the challenge is to explain how steady processes such as strain accumulation at plate boundaries can give rise to random 'new' earthquakes. A clue may be given by the long time constant of the Omori part. If it turns out that, as it appears, this has the same value as the Poisson time constant, then this suggests a single physical process controls both aftershocks and 'new' earthquakes.

The form of the model is fairly tractable analytically, allowing calculation of e.g. a mean interevent time, and conditional probabilities for interevent times given the time to the last earthquake. These derivative results have direct applications in earthquake hazard and risk assessment.

Application of the model to the clusters with aftershocks from $M_r = 6.5$ for 1985-2000 and $M_r = 7$, for 1985-2000 and 1977-2000, produced the following results and conclusions:

- The weights for the two parts of the mixture are approximately the same. This means that any subsequent earthquake $M \geq M_r$ has a probability of about 0.4-0.5 of being an aftershock.
- The mean interevent times calculated from the models for $M_r = 6.5$ and $M_r = 7$ are consistent with each other and with the mean interevent times calculated directly from the data.
- The weight and time constant parameters have appreciably large 95% confidence regions as inferred from log-likelihood values. This implies that it will be difficult to obtain a good model for any particular region with a limited record of large earthquakes. However, if M_r is reduced to produce more data, the assumption about physical connection between events is likely to be violated.

An attempt to apply the model to the occurrence of $M \geq 7$ earthquakes in central New Zealand since 1840 had the following consequences:

- With only 15 events, producing 14 interevent times, the model parameters had very large 95% confidence regions.
- The weight for the Omori part of the distribution was low compared to the global models: 0.1-0.15 rather than 0.4. We speculate that in fact the New Zealand catalogue of $M \geq 7$ earthquakes *is not complete*, as normally assumed. In particular we speculate that large aftershocks of pre-instrumental (pre-1900) earthquakes may not have been identified from the historical record. The 1848 Marlborough earthquakes are known to be a series in which several aftershocks are large, and it is quite possible that the 1855 earthquake

had one or more aftershocks of $M > 7$. Absence of any such data will significantly bias the weight parameter.

The greatest weakness in our modelling is the assumption of physical connectedness between all events in a supercluster, whose size is defined by its largest event. The model will not be able to be used in practice until a combined spatial/temporal model is developed. There are several suggestions in the literature as to how to proceed with the development of a spatial model.

In light of the results of this project, the following additional work should be undertaken:

1. Development of a spatio-temporal model, as noted immediately above.
2. After completing 1: determine a model for New Zealand using a suitable dataset or datasets e.g. $M \geq 6$ earthquakes since 1940. The model can then be tested as an alternative to the current (Poisson only) model in seismic hazard and risk assessments.
3. Systematically investigate the occurrence of the PSI phenomenon in the global database, to determine what its characteristic manifestations are in this database.
4. Investigations of the implications of the functional form of the model, to elucidate what is the underlying physical process and connection between successive earthquakes in a region, and why 'new' earthquakes appear to occur randomly in time.

Introduction

Randomness of large mainshocks

In earthquake occurrence studies, a distinction is typically made between 'mainshocks', and 'foreshocks' and 'aftershocks', the latter being causally related to the former. Foreshocks are earthquakes smaller than a subsequent event which are close by in time and space; aftershocks are nearby smaller events that follow. In some earthquake sequences there is little or no difference between the magnitudes of the largest few earthquakes. Such sequences are often called 'swarms'. A third class of earthquakes is now considered to exist by some people: 'triggered earthquakes', which are events causally related to mainshocks but which occur outside the limits of space and time that define aftershocks (e.g King et al.1994). Aftershocks are considered to occur as a result of stress redistribution around the fault zone (e.g. Scholz, 1990). Triggered earthquakes are thought to occur due to static stress changes in the vicinity of the ruptured fault (e.g. Stein, 1999, Parsons, 2002), or due to the transient seismic waves (e.g. Voisin et al., 2004).

In seismic hazard models, 'mainshocks' are typically assumed to be independent of each other. Their independent occurrence is often referred to as the 'background seismicity', in contrast to the non-independent foreshocks, aftershocks and triggered earthquakes. This means that the occurrence of mainshocks is necessarily modelled by a Poisson process, in which the time between successive mainshocks is described by an exponential distribution; i.e. the probability that the time to the next earthquake is less than or equal to time t is given by:

$$F(t) = 1 - e^{-t/t_\mu} \qquad (1)$$

where t_μ is the mean time between events. One feature of the Poisson process is its lack of memory. The occurrence of the next large earthquake does not depend on the timing of the previous one. The Poisson distribution is thus the standard function to model background seismicity in static seismic hazard models. i.e. models in which the hazard does not vary conditionally with earthquake occurrence, except perhaps for the occurrence of foreshocks and aftershocks. The New Zealand hazard model (Stirling et al, 2002) is an example of such a model.

We report here a single, simple model for the occurrence of shallow (depth $\leq$ 70km) earthquakes that describes the collective occurrence of mainshocks, aftershocks and triggered earthquakes (and, implicitly, foreshocks) all together. Among the applications for this model will be the replacement of static seismic hazard models with ones that are conditional on the occurrence of prior seismicity.

Declustering of earthquake catalogues

The parameter for the Poisson occurrence model for mainshocks is derived from catalogues that have had aftershocks removed (e.g. Gardner and Knopoff, 1974). The process of stripping aftershocks and other causally related events from a catalogue is termed 'declustering'. Several methods of aftershock identification have been described in the literature (e.g. Gardner and Knophoff, 1974, Reasenberg, 1985, Davis and Frohlich, 1991, Molchan and Dmitrieva, 1992).

In the simplest method, all earthquakes in a magnitude dependent window in time and space are removed following a large earthquake. Gardner and Knopoff used distances of 70 km and 94 km and time windows of 915 days and 985 days after earthquakes of magnitude 7.0 and 8.0 respectively. There is a problem intrinsic to this approach because it relies on the assumption that the 'background' process that gives rise to the new, independent events is switched off during the aftershock process. Whether the background process derives from stress accumulation due to plate movement, or to some other process, it is implausible that it could switch off and on in this way on a timescale of a few months to years. A further problem with this simple window method is the introduction of hard boundaries for regions in space which are usually centred on the largest earthquake.

Reasenberg (1985) developed an algorithm in which he defined interaction zones to link earthquakes to clusters, and clusters with one another. The interaction zones are dynamically modelled with one spatial and one temporal parameter. The spatial parameter depends on redistribution of stress around the earthquake fault. It is determined from the largest and the last associated earthquakes. Examples for distances are 70 km for an M = 7.0 earthquake and 181 km for M = 8. The temporal parameter is derived from the empirical law for aftershock decay which is further discussed below. For simplicity Reasenberg (1985) used a window of 1 day to link new earthquakes to one another and 10 days to link earthquakes within an existing cluster. As Savage and DePolo (1993) pointed out, Reasenberg's algorithm depends strongly on the magnitude at which the catalogue is complete. However, when adjusting parameters in the Reasenberg algorithm to suit their catalogue, Savage and DePolo found no significant difference between Reasenberg and Gardner & Knophoff declustering. In New Zealand Reasenberg's declustering method was applied by Stirling et al. (2002) to the New Zealand catalogue before deriving their hazard model.

In a different approach, Console et al. (2003) model the complete catalogue as one family of earthquakes which interact through an epidemic model (ETAS - see below), which is defined by the magnitudes and times of previous events and the distance from each catalogue earthquake to each potential source of a new earthquake. They also allow for the occurrence of new, random events. Their approach makes redundant the distinction between foreshocks, mainshocks and aftershocks.

The decay of aftershock activity

Some declustering mechanisms take into account the decay of aftershock occurrence with time. This decay is well known to follow a power law, the modified Omori law (Utsu et al. 1995):

$$\mathrm{dN/dt} = K / (c + \mathrm{t})^{p} \qquad (2)$$

where dN is the number of aftershocks occurring in a time interval dt; K is a parameter that is proportional to the aftershock productivity and varies from sequence to sequence; p describes the decay and takes values around 1.0 (Omori originally proposed $p = 1$); and c stands for a small time interval just after the mainshock. Mathematically and statistically c is important to make the equation non-singular at t = 0. It is identified with a time period after the mainshock in which the aftershocks cannot be fully detected. (Utsu et al., 1995; Kagan, 2004). An alternative rationalisation for c is that it represents a transient time during which aftershock productivity grows to full strength as the mainshock rupture progresses and is completed (Kagan, 2004).

A global model for aftershock probabilities

In a previous study (Christophersen 2000; Christophersen and Smith 2000) we used a global earthquake catalogue to derive the probability of a damaging earthquake following a damaging earthquake. We focussed on short-term earthquake occurrence by extending the time window for spatially proximal earthquakes by 30 days for each new earthquake above a threshold magnitude of M = 5.0. Our longest earthquake sequence lasted under two years. In space, we searched for earthquakes in a magnitude dependent circular area centred on the largest earthquake in a particular sequence. We fitted the distribution of epicentres by ellipses and defined the boundaries of the ellipse so that the area included 90% of the earthquakes. The aftershock area A scaled with mainshock magnitude M in the following manner:

$$\log_{10} A = M - 3.39 \tag{3}$$

We derived parameters for aftershock behaviour by stacking earthquakes from different sequences within a region. We found the *p*-value of Omori's Law to be 1.0 for most tectonic environments, except for mid-ocean ridges which have a faster decay of aftershocks and consequently have a larger *p*-value of about 1.2, and intracontinental zones such as Australia, which have slower than average aftershock decay, reflected in a *p*-value of about 0.9.

Triggered earthquakes

Triggered earthquakes have been defined as earthquakes occurring shortly after a mainshock but outside the aftershock area (Yeats et al., 1997). A well studied example of a triggered earthquake is the June 1992 M = 6.5 Big Bear, California earthquake that followed the M = 7.3 Landers earthquake within 3 hours (King et al., 1994). The second earthquake occurred within 40 km of the epicentre of the first but on a different fault. This occurrence initiated interest in triggered seismicity, and much research has been undertaken on this topic in the last decade. In another example of triggering, the North Anatolian fault has progressively slipped since 1939 in a domino-like chain reaction (Stein et al. 1997). In the latest episode, the Mw 7.4 Izmit earthquake on 17 August 1999 was followed on 12 November by an Mw 7.1 earthquake along-strike (e.g Aydin and Kalafat, 2002). A third type of triggering may occur when stresses due to the passage of seismic waves initiate an increase in seismicity. Such triggering was reported after the Landers earthquake – see e.g. Anderson et al. (1994) and other papers in the same special volume on the Landers earthquake. In our study we necessarily incorporate such transient triggering only when the triggered events occur within the region we designate as being potentially affected by the occurrence of an earthquake (see Superclusters, below).

Stein (1999), Parsons (2002) and numerous other workers have shown that aftershocks and triggered earthquakes are more prevalent in areas where shear stresses are enhanced following a major shear dislocation (mainshock) than in areas where shear stress is reduced. The changes in shear stress are associated with potential earthquake occurrence through a modified Coulomb stress criterion for slip on a fault (Stein 1999).

In a comprehensive analysis of $M_w \geq 7$ events from the Harvard catalogue 1977-2000, Parsons (2002) showed that the 61% of earthquakes following an $M_w \geq 7$ event within a region of shear stress increase of ≥ 0.01 MPa represented 8% of all $M_w \geq 7$ events, and that the rate of occurrence of these decayed according to the Omori law. An important conclusion from this result is that the distinction between aftershocks and triggered events is arbitrary, and that the aftershock region could be redefined to incorporate triggered events. We will show that this conclusion is valid, and subsequently make no distinction between aftershocks and triggered events. We also note Parson's result that 39% of post M 7 proximal events occurred in regions of shear stress *decrease*, representing a further 5% of the catalogue. This suggests that processes other than Coulomb stress redistribution induce subsequent earthquakes. Indeed, it could be argued that if 39% of post-mainshock events are caused by some unknown process, then this process would be occurring in the (equally sized) areas of elevated shear stress as well, and that accordingly only (61-39)% = 22% of post-mainshock events are likely to have resulted from Coulomb shear stress enhancement.

We therefore do not limit inclusion of triggered events (and aftershocks) to those occurring in fault-proximal areas of shear stress increase.

ETAS models

A number of models for earthquake recurrence have been based on the Omori law. One of the most successful types of model has been the Epidemic-Type Aftershock Sequence model (ETAS – see Ogata, 1988, 1998). Ogata (1988) proposed that the rate of occurrence of earthquakes per unit area at any location $\lambda(t)$ could be written as

$$\lambda(t) = \lambda_a(t) + \lambda_c(t) \tag{4}$$

where $\lambda_a(t)$ is the contribution from aftershocks and triggered earthquakes, and $\lambda_c(t)$ is contribution from 'new' independent events. $\lambda_a(t)$ is the sum of a product of terms that allow for the magnitudes and distances of all the earthquakes in the catalogue to the location where the calculation is being made, and the sum is taken over all earthquakes in the catalogue.

In a recent extension of these ideas, Console et al. (2003) proposed that $\lambda(t, x, y, m)$ be written as

$$\lambda(t, x, y, m) = p_r \lambda_0(x, y, m) + \Sigma\, H(t - t_j)\, K\, h(t - t_j)\, f(x - x_j, y - y_j)\, \beta\, e^{\alpha m_j - \beta m} \tag{5}$$

where x, y, m are the location and magnitude of the future event; j denotes the jth event in the catalogue, with the sum being taken over all j; H is the Heaviside function (unit step function at $t - t_j = 0$), h is a modified Omori law normalised so that its integral over time = 1:

$$h(t) = (p-1)c^{p-1} (c + t)^{-p} \qquad (p \neq 1) \tag{6}$$

and f is a spatial kernel; either a Gaussian Kernel:

$$f(r) = (1 / 2\pi\, \sigma) \exp (- r^2 /2\sigma^2) \tag{7}$$

or a (modified) power law:

$$f(r) = (1/\pi) (q-1)\, d^{\,2(q-1)} \; / \; (r^2 + d^2)^{q} \tag{8}$$

and p_r (proportion of new events), K, α, $\beta = \log_e 10\, b$, are (constant) parameters to be determined, along with c, p, and either σ or d and q. The preferred method of determining the parameters is maximum likelihood (e.g. Severini 2000, Dempster et al. 1977).

When applying this model to an Italian earthquake catalogue, Console et al. (2003) preferred a Gaussian kernel for the new earthquakes and the power law function for the aftershocks. We note that Stock and Smith (2002a, b) demonstrate that an adaptive Gaussian kernel represents spatial seismicity features better than the fixed Gaussian kernel (constant σ) above. Also, Console et al.'s parameter *d* should, *a priori*, be magnitude dependent. Thus the Colsole et al. model is complex, and possibly needs to be made more so to model seismicity in a region where seismicity is very heterogeneous, such as New Zealand. It, and other ETAS models, have some other undesirable features, which we discuss while describing our modelling, below.

A power law model for large earthquake recurrence

Kagan and Jackson (1999) showed in a global earthquake catalogue study, including aftershocks, that time intervals between $M \geq 7.5$ earthquakes with intersecting fault areas are better described by a power law $1/t^p$, with $p \sim 0.75$, than a Poisson distribution over post-earthquake intervals up to ten years. This finding may suggest that a power law is a better descriptor of earthquake probabilities in the medium term than the Poisson distribution.

Precursory Scale Increase (of seismicity - PSI)

Evison and Rhoades (2004) propose that before at least some large ($M \geq 5.8$) earthquakes in several different regions world-wide, there is an increase in seismicity, characterised by an increase in the average magnitude occurring per unit time, starting at a time T_p which scales directly with the magnitude M_m of the subsequent large earthquake (mainshock):

$$\log_{10} T_p \propto M_m$$

and also that the average magnitude of the largest three earthquakes in the precursory set also scales with M_m. Evison and Rhoades do not necessarily propose that the PSI phenomenon is universal. We note that the occurrence of local precursors to large earthquakes may not be inconsistent with their occurrence appearing random on a large scale.

The datasets we developed in the course of this project offer an opportunity to investigate PSI on a global scale, and we discuss some very preliminary results below.

New Zealand aftershock sequences

Analysis of 17 New Zealand aftershock sequences with mainshock magnitudes ≥ 5.5 that occurred between 1987 and 1995 showed a wide variation of aftershock parameters (Eberhart-Phillips, 1999). The observed variation could be due to a number of reasons: The range of different tectonic settings within New Zealand, the limited number of aftershocks in some of the sequences, or the intrinsic variability of the aftershock process (Christophersen 2000).

New Zealand's large earthquakes

Since 1840, New Zealand has experienced 18 shallow earthquakes of $M \geq 7$, including 15 in a central region between Arthur's Pass and Gisborne, two in Fiordland and one off East Cape (Smith, 1994; see Table 4 below).

Of these, eight occurred between 1929 and 1942, about six times the average rate (see Table 4, which does not include one in Fiordland in 1938). In addition, there have been long intervals with no large events. For example, there was no magnitude 7 or greater event between 1897 and 1929 or between 1968 and 1995. However, because of the large variance for the mean interevent time, the departure of the distribution of interevent times of large New Zealand earthquakes is only marginally different from a Poisson distribution (Smith, 1994).

Aims and objectives of this project

We want to develop probabilistic forecasts for earthquake occurrence in the medium term after a large earthquake. The ultimate goal of our project is to use the new probabilistic forecasts in conjunction with the existing aftershock model to develop estimates of time-varying earthquake hazard.

Our research has been explorative in the sense that we wanted to investigate the distinction between aftershocks and other triggered earthquakes using a global earthquake catalogue. The catalogue was established during a previous project and we have extended the duration of the catalogue using new data now available. We began with the hypothesis that, after removing aftershocks (declustering) the rate of seismicity following a large earthquake would be greater than before it due to the occurrence of triggered events outside the defined aftershock space-time window, and we sought to quantify this increase.

We can now identify four broad objectives that have been addressed:

1. Establish a database of mainshocks by removing immediate foreshocks and aftershocks as previously defined, and analyse the rate of earthquake occurrence before and after a large earthquake.
2. Apply Kagan and Jackson's methodology to further investigate the power law behaviour of large earthquakes for up to 10 years after a major earthquake.
3. Test our database for any evidence of the Precursory Scale Increase phenomenon
4. Apply our results to an appropriate New Zealand earthquake catalogue.

Objectives 3 and 4 were not in the original proposal but evolved during the project. While the overall aim to develop a model for time varying hazard has remained unchanged, the objectives listed in the proposal had to be adapted when objective 1 did not lead to the expected results.

Development of Superclusters

We wanted to analyse the temporal occurrence of spatially proximal earthquakes. For this purpose we searched an earthquake catalogue, or subsets of it, for earthquakes that were spatially close over the duration of the catalogue. We called the resulting data sets of clusters of earthquakes 'superclusters'. We distinguish between superclusters without aftershocks and superclusters with aftershocks. Superclusters without aftershocks have immediate aftershocks and foreshocks removed from the dataset by a declustering algorithm (see above) before running the superclustering algorithm. In this section we describe the earthquake catalogue, the search radius, the removal of immediate fore- and aftershocks and the search algorithm for superclusters. We then investigate some of the properties of superclusters. We close with an overview of our subsequent analyses.

The Earthquake Catalogue

The catalogue is based on the earthquakes reports by the International Seismological Centre (ISC) (ISC, 2002). It includes body wave magnitude mb and surface wave magnitude Ms as calculated by ISC and the American National Earthquake Information Center, NEIC; a re-calculated body wave magnitude mbcor; and moment magnitude Mw calculated from Harvard moment-tensor reports. Table 1 shows an overview of the magnitudes and the time periods for which they are available.

Table 1: Available time periods of the magnitudes in the earthquake catalogue

Magnitude	Harvard Mw	Corrected mb	ISC mb	ISC Ms	NEIC mb	NEIC Ms
Time period	1977 – 2000	1964 – 1985	1964 – 2000	1978 – 1995	1964 – 2000	1968 – 1995

An extensive magnitude and catalogue completeness study found the catalogue to be complete for magnitude 5.0 and above (Christophersen, 2000). However, the size of large earthquakes often is underreported in the early part of the catalogue because the body wave magnitude starts saturating at about magnitude 6.0, and the moment magnitude and the surface wave magnitudes were not uniformly available until 1977 and 1978 respectively. Some moment magnitudes for earlier earthquakes were added to the catalogue by hand. We estimate that the sizes of earthquakes larger than magnitude 7.0 would have been correctly estimated since 1977 due to the regular reports of the moment magnitude. Since 1985, earthquakes larger than 6.5 would have been correctly estimated because the Harvard moment and the ISC and NEIC surface wave magnitudes were well established by then. However, our earlier study (Christophersen, 2000) shows that between 1977 and 1985 the catalogue was not consistently complete for magnitudes in the range $6.5 \leq M < 7$.

Our catalogue includes 44,568 earthquakes.

The search radius

The size of the search radius for related earthquakes is derived from the study of aftershock areas (Christophersen, 2000), for which the areas occupied by epicentres of earthquakes within a sequence of aftershocks were modelled by ellipses centred on the mean of the epicentre locations. The areas A of ellipses were set to include 90% of the epicentres (see Introduction). For a given magnitude M, the resultant areas on average satisfied the scaling equation (see Introduction)

$$\log_{10} A = M - 3.39 \qquad (3)$$

For the purpose of defining a search radius for spatially related earthquakes, we assume a ratio of ellipse major axis a to ellipse minor axis b of up to 4:

$$\max a / b = 4 \qquad (9)$$

The area A of an ellipse is given by

$$A = \pi\, a\, b \qquad (10)$$

Combining equations (3), (9) and (10), we derive a maximum distance d between earthquakes within an the ellipse of

$$d = 2\,\mathrm{a} = 4 \sqrt{(10^{(M-3.39)} / \pi)} \qquad (11)$$

d becomes the search radius. The search for related events is centred on the largest earthquake within a cluster. The search area is 16 times the size of the corresponding 90% aftershock ellipse and therefore may include more earthquakes than immediate aftershocks. Table 1 shows the search radius for some selected magnitudes.

Table 2: Search radius in kilometres for some selected magnitudes

Magnitude	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.2
Search Radius [km]	14	26	46	81	144	256	455	810	1813

The removal of immediate aftershocks

We applied a simple window method in time and space to remove immediate fore- and aftershocks from the earthquake catalogue. In our declustering, every event in the catalogue qualified as a potential mainshock. We searched for related earthquakes within 30 days in the magnitude dependent radius in space described above. If a larger earthquake occurred within the time window we used the larger search radius to decide whether the earthquakes were close enough to be related. If they were, the longitude, latitude and magnitude of the larger event became the new parameters for searching. For every related earthquake found, we extended the search window in time by 30 days. From each cluster of related earthquakes we kept the largest event and thus created a database of mainshocks. The database of mainshocks included 29,224 earthquakes.

The magnitude-frequency distribution of the catalogue and the mainshock database

Gutenberg and Richter first described the magnitude-frequency distribution of earthquakes in the following form

$$\log_{10} N(M) = a - b M \tag{12}$$

where N (M) is the number of earthquakes of magnitude M and *a* and *b* are parameters (e.g. Gutenberg and Richter, 1949). The *b*-value describes the ratio between numbers of small and large earthquakes and takes values of about 1.0. The relationship applies to complete catalogues (e.g. Kagan, 1999), aftershock sequences (e.g. Utsu, 1969) and mainshocks (Christophersen, 2000).

The *b*-value can be calculated by maximum likelihood method to be

$$b = 1/\{\log_e 10 * (M_{ave} - M_c)\} \tag{13}$$

where M_{ave} is the average magnitude of the data set for all earthquakes larger than M_c and M_c is the magnitude above which the data are assumed to be complete (Aki, 1965).

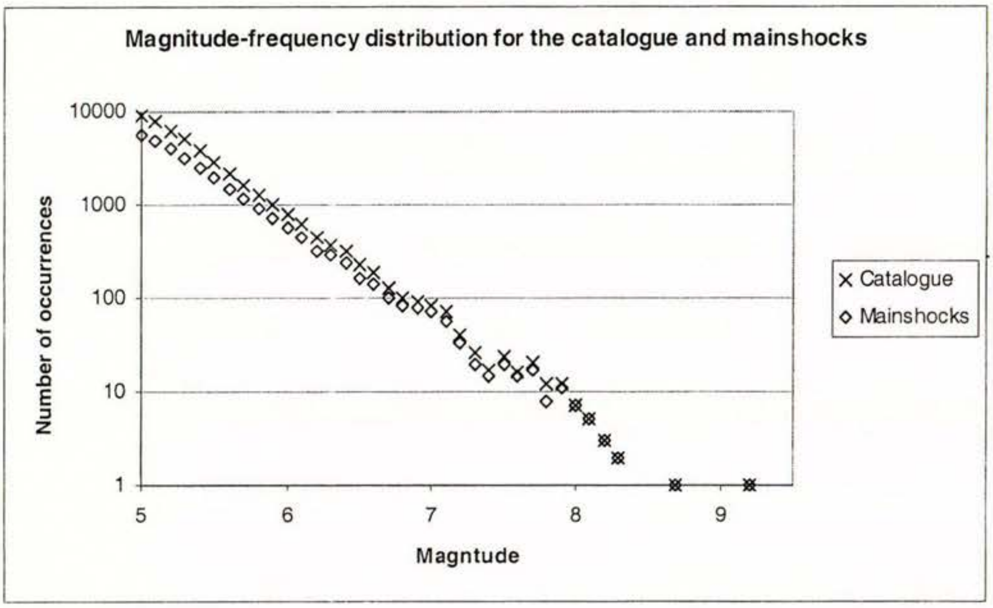


Figure 1: Magnitude-frequency distribution for the whole catalogue and mainshocks only, for the period1964-2000.

The 95% confidence interval for the b-value is given approximately by

$$\pm 1.96\, b / \sqrt{N_{tot}} \tag{14}$$

where N_{tot} is the sample size and *b* is the estimated *b*-value (Aki, 1965).

Figure 1 shows the magnitude-frequency distribution for our complete catalogue and for mainshocks from 1964 to 2000. The estimated b-values are 1.09 ± 0.01 and 1.03 ± 0.01. We know that some magnitudes of earthquakes in the earlier part of the catalogue are underestimated in size. Therefore the average magnitude in eqn (13) is too small and the *b*-value too large. Calculating the *b*-value for the period 1985 – 2000 we find $b = 1.03 \pm 0.01$ and $b = 0.99 \pm 0.03$ for the catalogue and the mainshocks respectively. We expect a smaller *b*-value for the mainshocks, because the declustered catalogue removes relatively more small earthquakes than large ones.

Search algorithm for superclusters

The supercluster datasets were built by searching a given catalogue (i.e. possibly declustered) for spatially proximal events. The search was undertaken in chronological order, one earthquake at a time. So for example, the magnitude of the first earthquake was compared to the magnitude of each subsequent earthquake. The larger of the two magnitudes was used to calculate a magnitude dependent search radius (*d*, eqn 11) to decide whether the events were proximal. Now the location and the magnitude of the larger event were used to find further related events. We start again at the beginning of the catalogue to find earthquakes that would have not have been included using the first, smaller radius. All related events were extracted to a new database. Once the end date of the catalogue is reached, the supercluster is complete. Each earthquake that was written into the supercluster database was flagged in the catalogue and excluded from further searches, which started at the beginning of the catalogue at the first unflagged earthquake. As a consequence, in our algorithm each earthquake can only be part of one supercluster, even if the search areas for superclusters overlap.

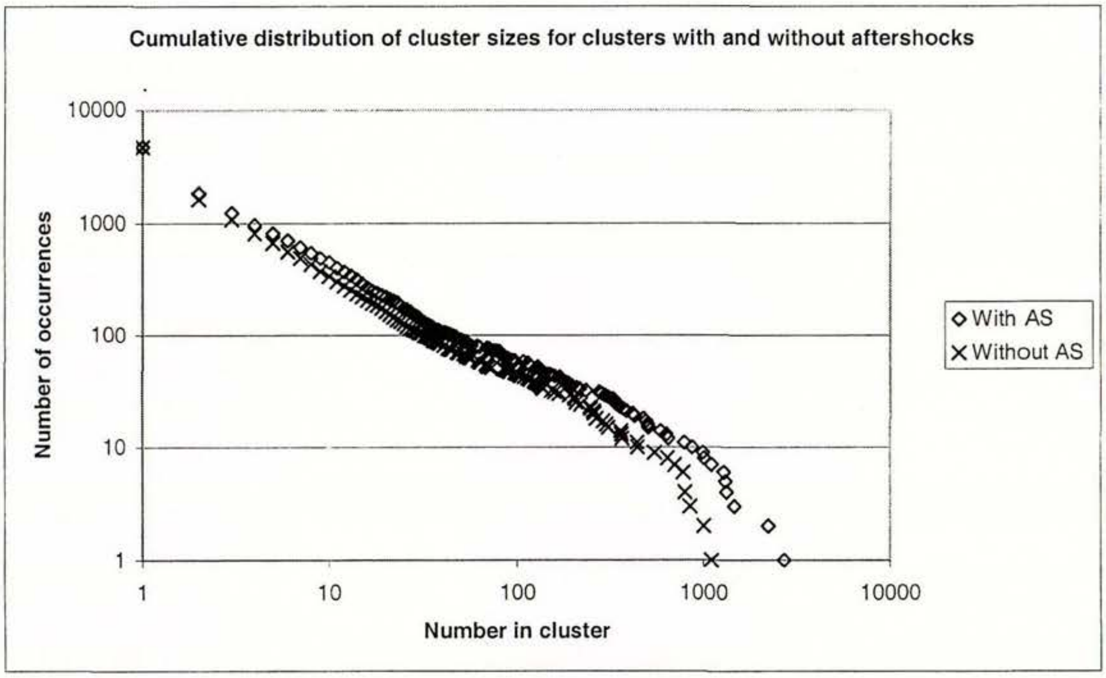


Figure 2: Size distribution of superclusters with and without aftershocks, for the period 1964-2000.

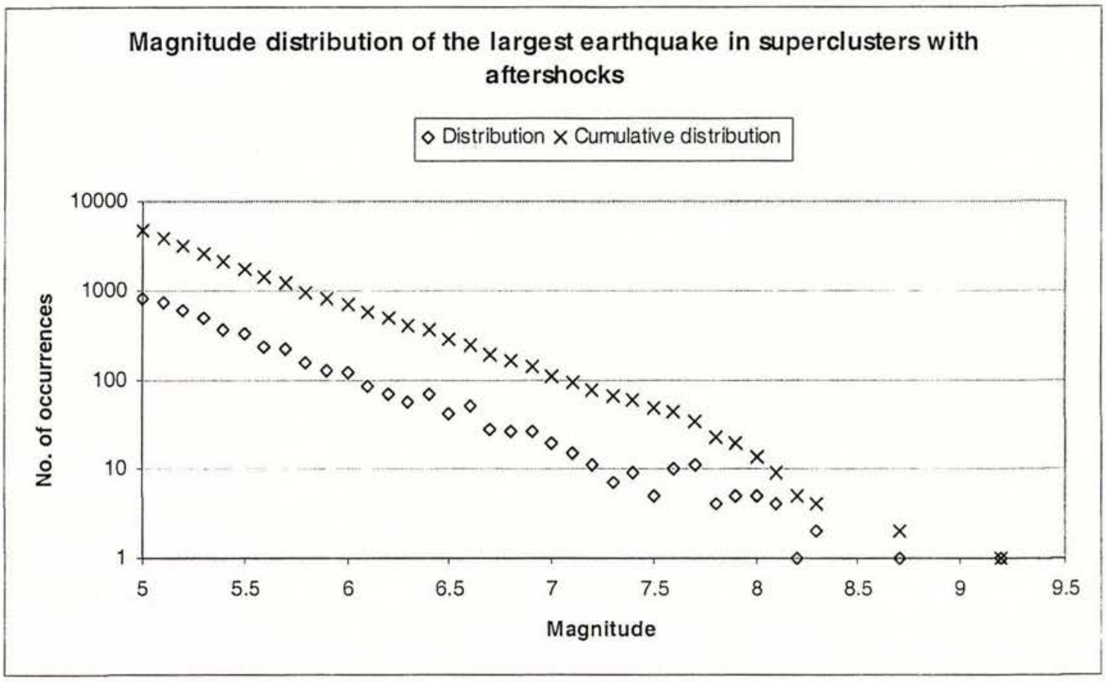


Figure 3: Magnitude distribution, and cumulative distribution for the largest earthquakes in each supercluster with aftershocks.

Numbers of superclusters and their frequency-size distribution

Superclusters with aftershocks and without aftershocks were searched from catalogues with 44,569 and 29,224 earthquakes respectively. Applying the search algorithm described above, we find just over 4,700 superclusters. If there were no overlap between search areas at all, the declustering of fore- and aftershocks would not affect the size of the search area and the total number of superclusters would be the same with or without aftershocks. However, we find a slight mismatch, with 0.6% fewer superclusters with than without aftershocks.

Figure 2 illustrates the size of cluster – frequency distribution of superclusters. About 60% of the clusters consist of only one earthquake. However, more than 60% of the earthquakes are in superclusters larger than 100 earthquakes and about 75% of earthquakes are in superclusters with 10 or more events. The power-law behaviour in Figure 2 results from the combination of magnitude scaling of the search radius and the Gutenberg and Richter magnitude-frequency law. The departure from a power law at the right, for high numbers, probably reflects the similar fall-off of the magnitude-frequency law at high magnitudes.

The magnitude-frequency distribution of superclusters

Figure 3 shows the magnitude frequency relation for the largest earthquake from each supercluster with aftershocks. The data for superclusters without aftershocks are insignificantly different. The Gutenberg-Richter relation applies to the superclusters. However, the b-value is only 0.84 ± 0.02, which is significantly smaller than for the complete catalogue or the catalogue of mainshocks. The difference in b-value is caused by a bias of magnitude selection towards large earthquakes, as we have chosen the largest event in each supercluster.

Worldwide distribution of superclusters

Figure 4 a-c shows the locations of those superclusters worldwide that have a mainshock of M 7 or greater. To give a feeling for their relationship to the catalogue, they are superimposed on epicentres of (a) the whole catalogue; (b) M 6.5 and greater and (c) M 7 and greater. The significance of these sets is discussed below.

Random occurrence of earthquakes in superclusters?

In the introduction we discussed the assumption that large earthquakes have a Poisson distribution in time. The number of large earthquakes per year in a region therefore has a Poisson distribution, with parameter $t_{\mu} = 1/$ (mean number of earthquakes per year in that region). For a Poisson distribution the mean equals the variance. In figure 5 a and 5 b we plot variance versus mean for 1,844 superclusters with aftershocks that had 2 and more earthquakes and for 1,642 superclusters without aftershocks that had at least 2 events per cluster. Note the difference in the variance axis scale. We used a logarithmic scale to better see all data. The model line represents unity. As expected due to the clustering of aftershocks, the variance for superclusters with aftershocks is significantly higher than the mean. For superclusters without aftershocks the variance and mean are fairly close to each other. However, the data still sit more above the line than below, indicating slightly greater variation in seismicity rate in fact than predicted by the Poisson model. This is consistent with specific examples; e.g. the larger than predicted numbers of both short and long interevent times for New Zealand earthquakes already discussed.

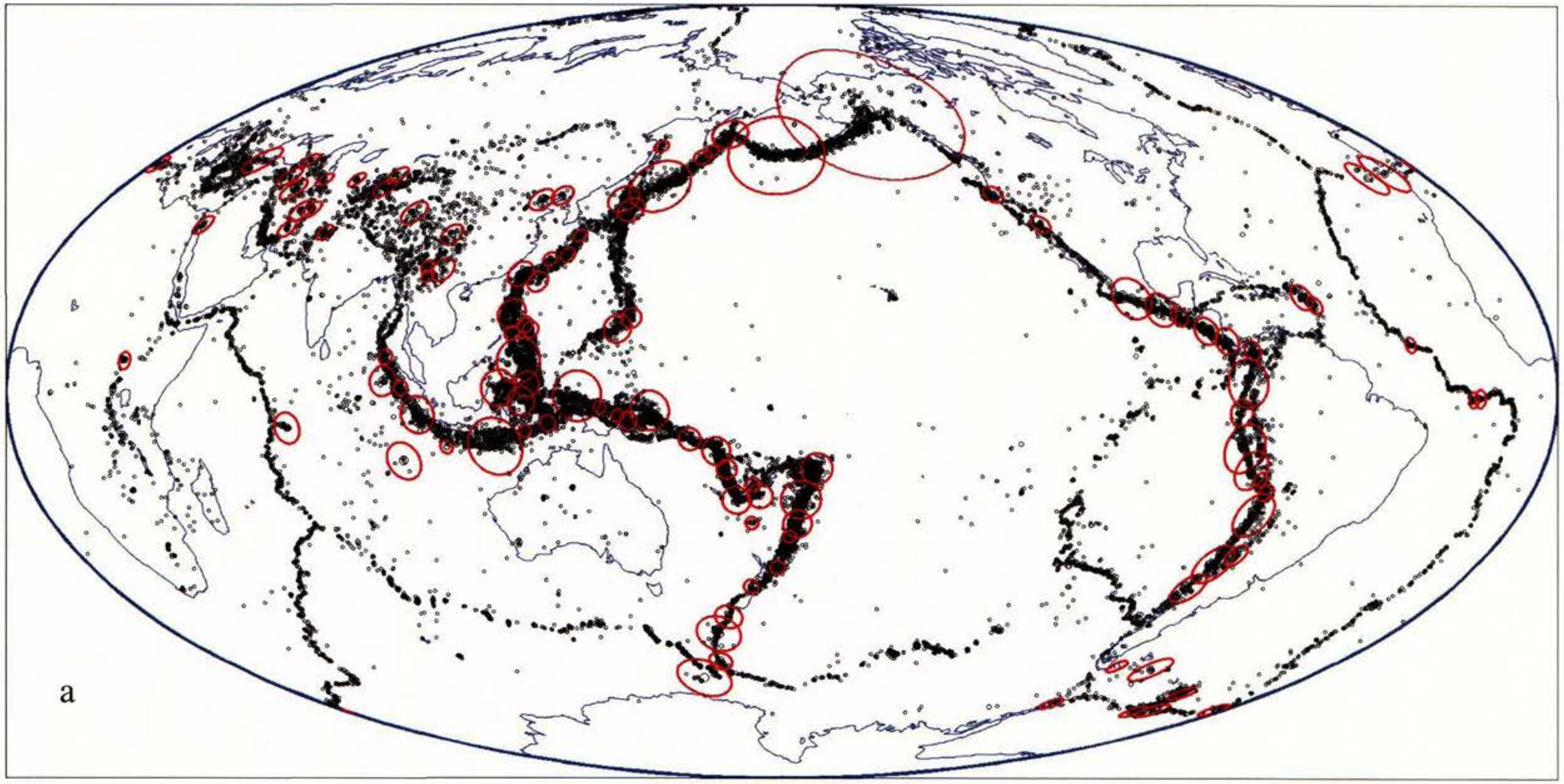


Figure 4: Epicentres of catalogue earthquakes and 110 superclusters (red) with largest earthquake M ≥ 7.
a. All earthquakes, M ≥ 5. b. Earthquakes M ≥ 6.5; c. Earthquakes M ≥ 7. For significance, see text.

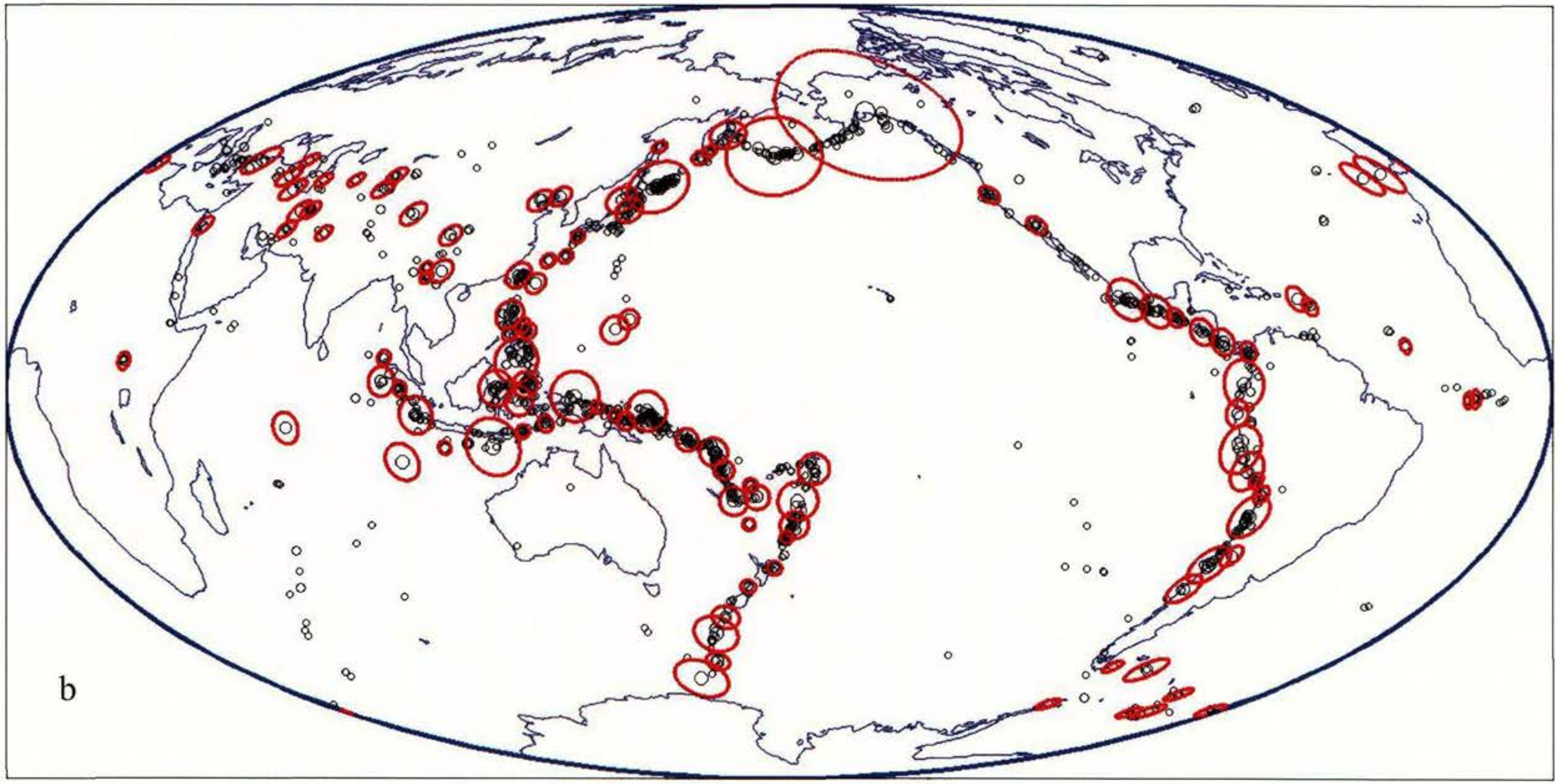


Figure 4: Epicentres of catalogue earthquakes and 110 superclusters (red) with largest earthquake M ≥ 7.
b. Earthquakes M ≥ 6.5.

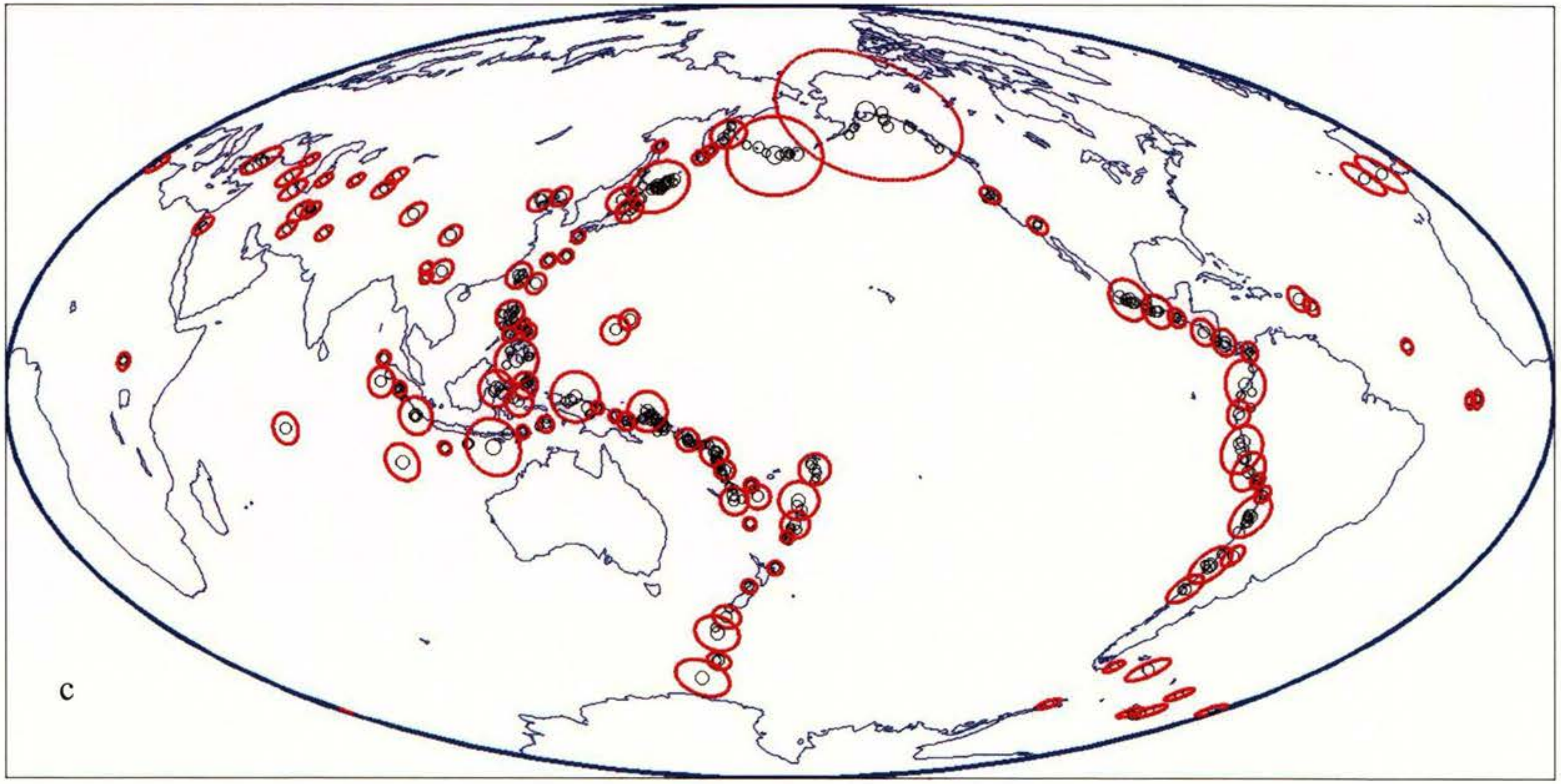
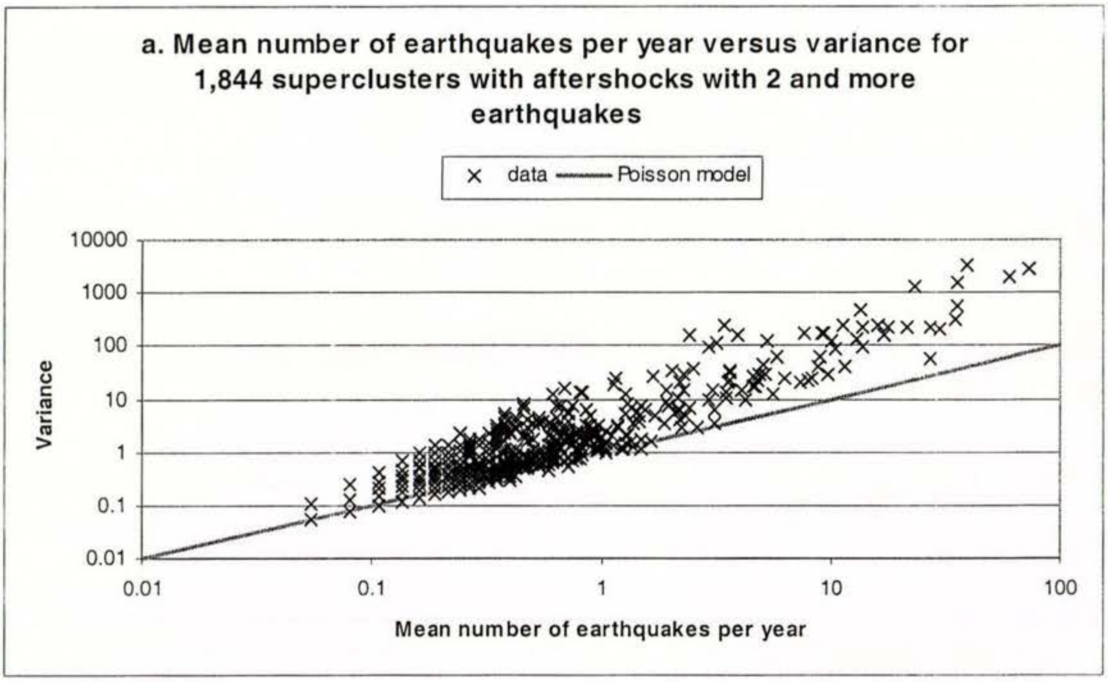


Figure 4: Epicentres of catalogue earthquakes and 110 superclusters (red) with largest earthquake M ≥ 7.
c. Earthquakes M ≥ 7.



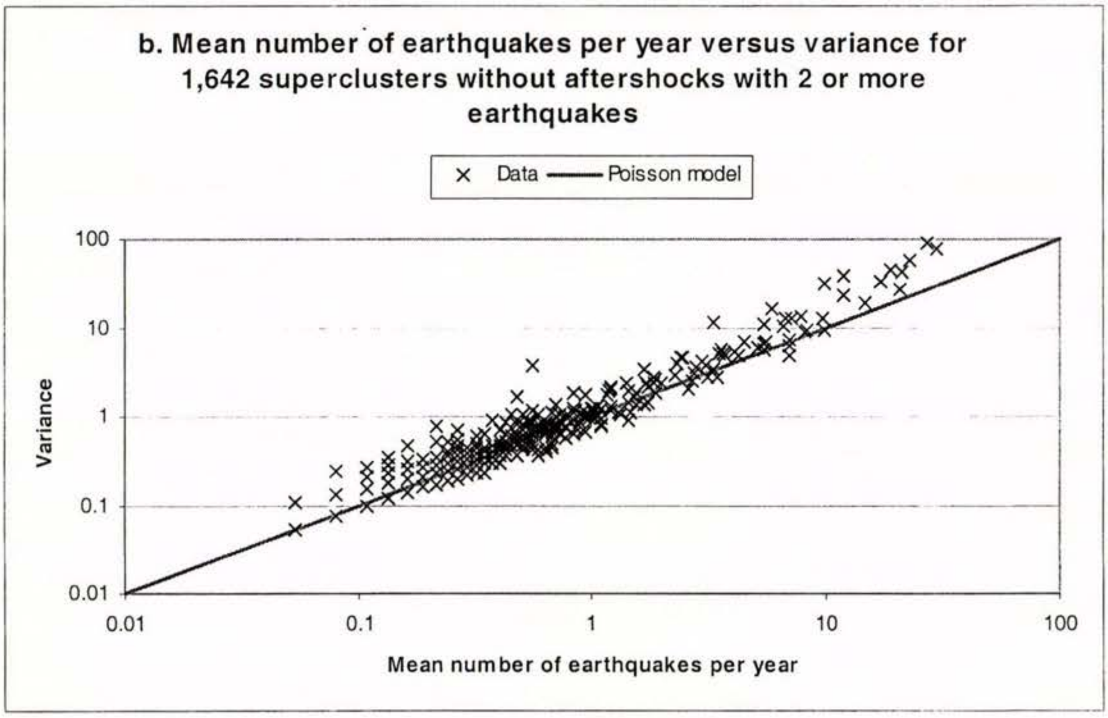


Figure 5: Mean number of earthquakes per year versus variance for superclusters a: with aftershocks and b: without.

Duration of superclusters

The distribution function for time differences between earthquakes that occur randomly in time is exponential with parameter t_μ = mean interevent time (see Introduction). We compared the mean time between earthquakes in a supercluster to the time between the first observed event in a supercluster and the beginning of the catalogue, and between the last observed event and the end of the catalogue to investigate how many clusters could have started and ended during the period of the catalogue (37 years).

We calculated average time t_μ between earthquakes for each of the 1,087 superclusters without aftershocks with three or more earthquakes

$$t_\mu = (37 - t_{start} - t_{end}) / (n - 1) \tag{15}$$

where n is the number of events in the supercluster, t_{start} is the time of the first earthquake in the supercluster and t_{end} the time of the last. We then used the cumulative distribution

$$F(t; t_\mu) = 1 - e^{-t/t_\mu} \tag{16}$$

for $t = t_{start}$ and t_{end} respectively to calculate the probability of observing t_{start} and t_{end} given t_μ. For 1,087 independent clusters, we would expect 2.5% or 27 to lie outside the upper 97.5 percentile of the cumulative exponential distribution function. For the start time we found 171 clusters outside the 97.5 percentile and 130 for the end time, equalling 16% and 11% of all clusters respectively. We interpret this result so that about 13.5% of clusters seem to have started during the duration of the catalogue and 8.5% ended. For the superclusters that appear to be continuous throughout the duration of the catalogue we calculate a background rate of expected number of earthquakes per year by dividing the total number of earthquakes in the supercluster by the duration of the catalogue (37 years).

Limitations of the data

The 37 year duration of the global earthquake catalogue is small in geological time. We are mindful that there will have been large earthquakes that occurred before the start of the catalogue, which may have triggered earthquakes in the catalogue. Similarly, we might not see all the triggered earthquakes that could have followed a large earthquake towards the end of our catalogue.

We are aware that earthquake sizes in the early part of our catalogue are systematically underestimated. For some aspects of the behaviour of large earthquakes, this problem effectively shortens the length of catalogue we can use, as described above.

Finally, we undertake a global analysis by stacking data from various regions. While we may thus be able to derive general behaviour of earthquake occurrence, our models are likely to have limited application in a particular region, as rate of occurrence varies widely, even in relatively active areas. Additional work will be required to determine appropriate regional parameters.

Overview of further analyses

Figure 6 provides an overview of the analyses that followed from the establishment of the supercluster datasets. These are discussed in detail in later sections of the report.

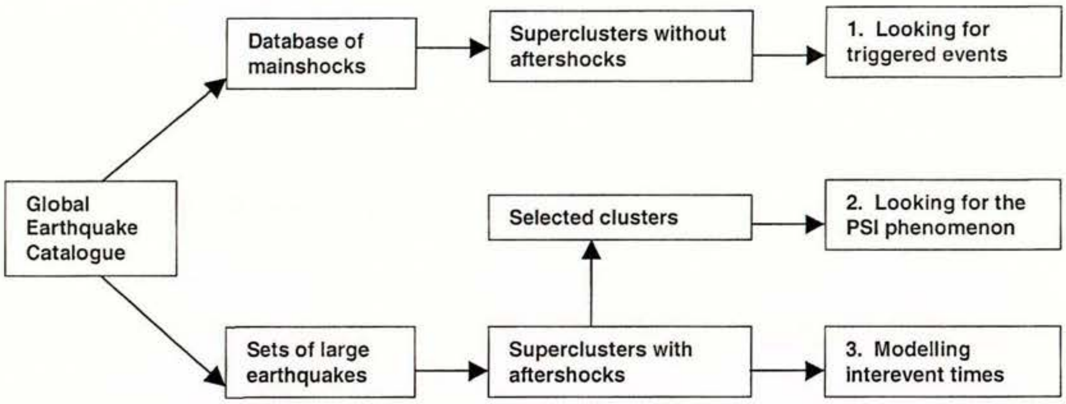


Figure 6: Flow path from establishment of the catalogue of superclusters with and without aftershocks through to the project objectives – probability of triggered events, Precursory Scale Increase (PSI), and the modelling of interevent times.

The original objective of the project was Figure 6 Box 1 - looking for the frequency and hence probability of occurrence of triggered events. As we discuss below, we reached a similar conclusion to Parsons (2002) that the distinction between aftershocks and triggered events was arbitrary. This led to our giving more weight to the objective of using the datasets to test the proposal of Kagan and Jackson (1999), that interevent times between spatially proximal large earthquakes had a power-law distribution. This led on to Figure 6 Box 3 - the development of a new interevent time model. The datasets also offered the opportunity of a very preliminary exploration of Figure 6 Box 2 - the PSI phenomenon proposed by Evison and Rhoades (2004).

Looking for triggered earthquakes

Objective

If large earthquakes trigger earthquakes outside the space-time window of the aftershocks, we expect to observe a systematic increase of seismicity there after the large earthquake. In this section we investigate whether seismicity systematically increases after the occurrence of a large earthquake after the time of the immediate aftershocks. For this purpose, we first focus on superclusters without aftershocks with only one earthquake of magnitude 7.0 or larger and with the next largest earthquake at least one magnitude unit smaller. We exclude superclusters with multiple large earthquakes or no large earthquakes from our analysis so that each supercluster has an easily identifiable potential trigger.

Method of analysis

We define the background seismicity rate or expected number of earthquakes per year for each supercluster without aftershocks to be the total number of earthquakes in the supercluster divided by the duration of the catalogue of 37 years. We checked that all superclusters used in this analysis are continuous over the duration of the catalogue as discussed in the section 'Duration of superclusters'. We define the ratio of seismicity for each year as:

$$\text{ratio (year)} = \{\text{no. of earthquakes in that year}\} / \{\text{expected no. of earthquakes per year}\} \quad (17)$$

By construction, the ratio is always positive. For values less than 1.0, the number of earthquakes in that year is below average, for values larger than 1.0, the seismicity is above average. For our analysis, we set the time of the largest earthquake in a supercluster to 0 and count the number of earthquakes in each 365.25 days period before and after the largest earthquake. We disregard the earthquakes at the very beginning and end of the catalogue that occur in intervals smaller than 365.25 days.

Individual superclusters

We look first at single superclusters, and then discuss pooling data from them in the next section. There were only four sequences with the largest earthquake $M \geq 7.0$, at least one magnitude unit difference to the second largest event and with at least 100 earthquakes in the sequences. The largest earthquake was the 1964 Alaska earthquake, which had no prior data. The others were 1990 Saipan, Marianas Trench, 1993 Guam, and a 1998 event with little subsequent data. The Saipan event had 123 events in the supercluster and the Guam supercluster had 260.

Figure 7 shows an example of a ratio plot for the 1993, $M_w = 7.7$ Guam earthquake. There is some scatter in ratio between years. However, no clear increase in seismicity after the largest earthquake can be observed. The average of the ratios before and after the largest earthquake are 0.96 and 1.16 respectively. The seismicity ratios for other individual superclusters without aftershocks shows similar scatter. However, no significant increase of seismicity can be observed following the potential trigger event.

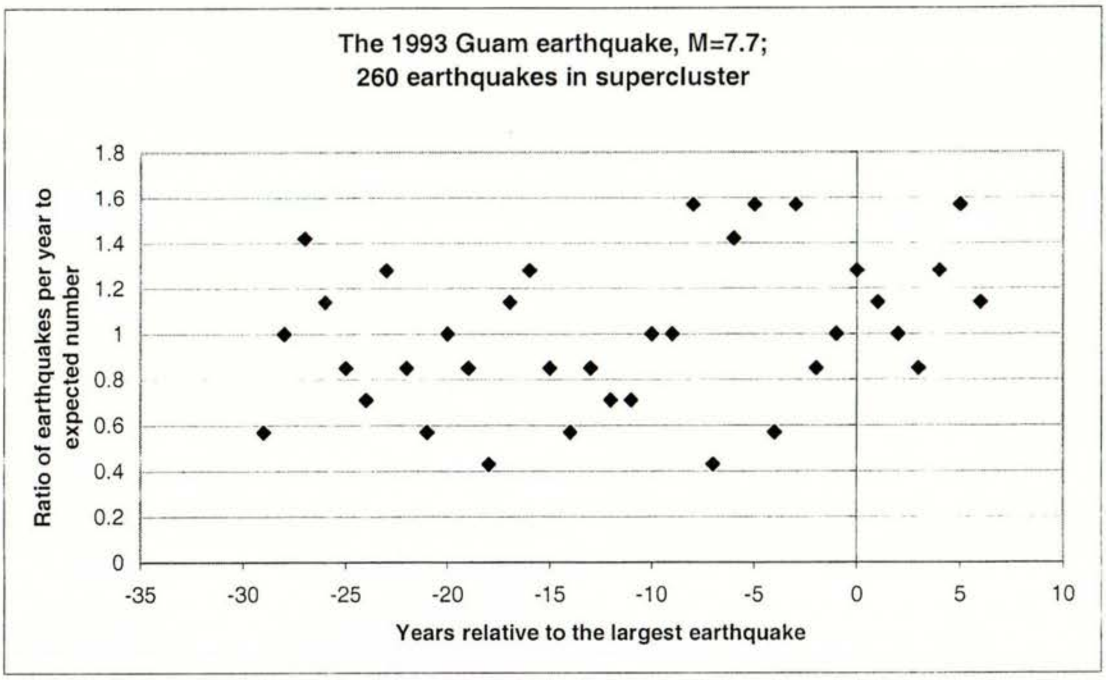


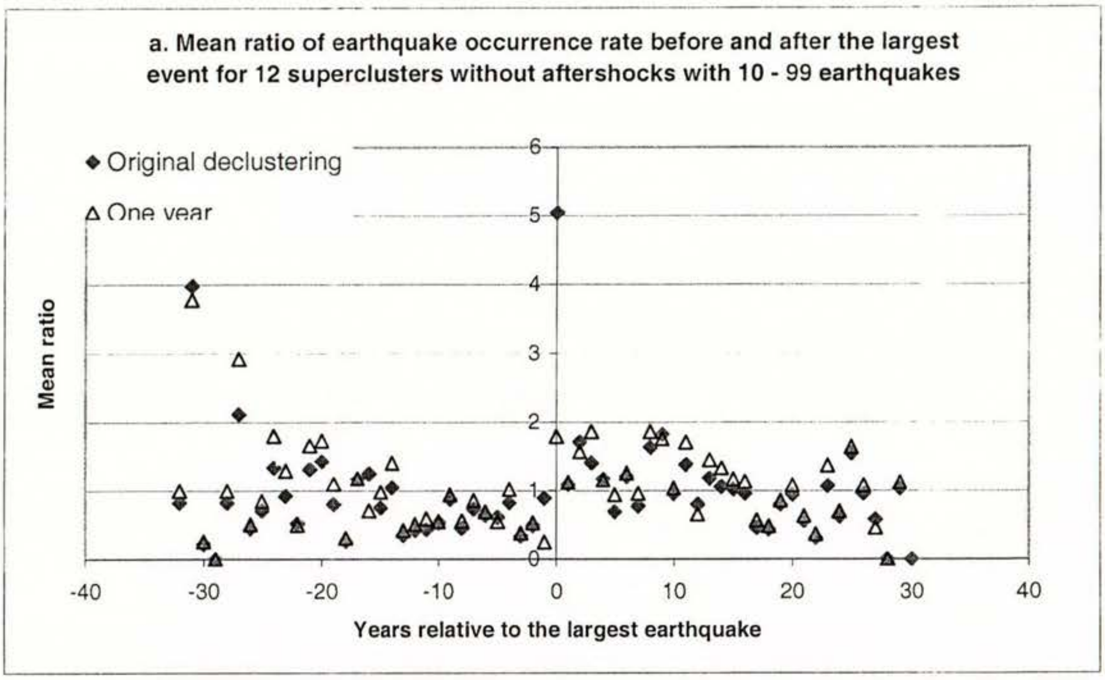
Figure 7: Rate of occurrence of earthquakes per year relative to background in the supercluster containing the 1993 Guam earthquake (year 0).

Combined superclusters

Next we combined data from various superclusters without aftershocks. There were a total of 110 superclusters with $M \geq 7.0$ (see Fig. 4). We selected those 18 clusters that had at least one earthquake larger than magnitude 7.0 and the second largest earthquake one magnitude unit smaller. We excluded two superclusters that had their largest earthquakes within two years of either end of the earthquake catalogue. We split the remaining data into two sets: one with 12 superclusters that had between 10 – 99 earthquakes per supercluster, comprising a total of 449 earthquakes; and the 4 superclusters each with at least 100 earthquakes mentioned in the previous sectuion, comprising a total of 1,262 earthquakes. Seismicity ratios were then averaged for each year before and after the largest potential trigger earthquake. The diamonds in figures 8a and 8b show the mean ratio of seismicity versus time for those two sets.

Differences in ratio before and after the largest earthquake

There is still some scatter in the mean ratios. The average mean ratio before and after the largest earthquake is 0.86 and 0.94 for the smaller superclusters and 0.93 both before and after for the larger superclusters. For the smaller superclusters, the ratio seems to be consistently smaller in the 10 years prior to the largest earthquake compared to the 10 years following: the average ratio for years -10 to -1 is 0.63, compared to 1.25 for years 1 to 10.



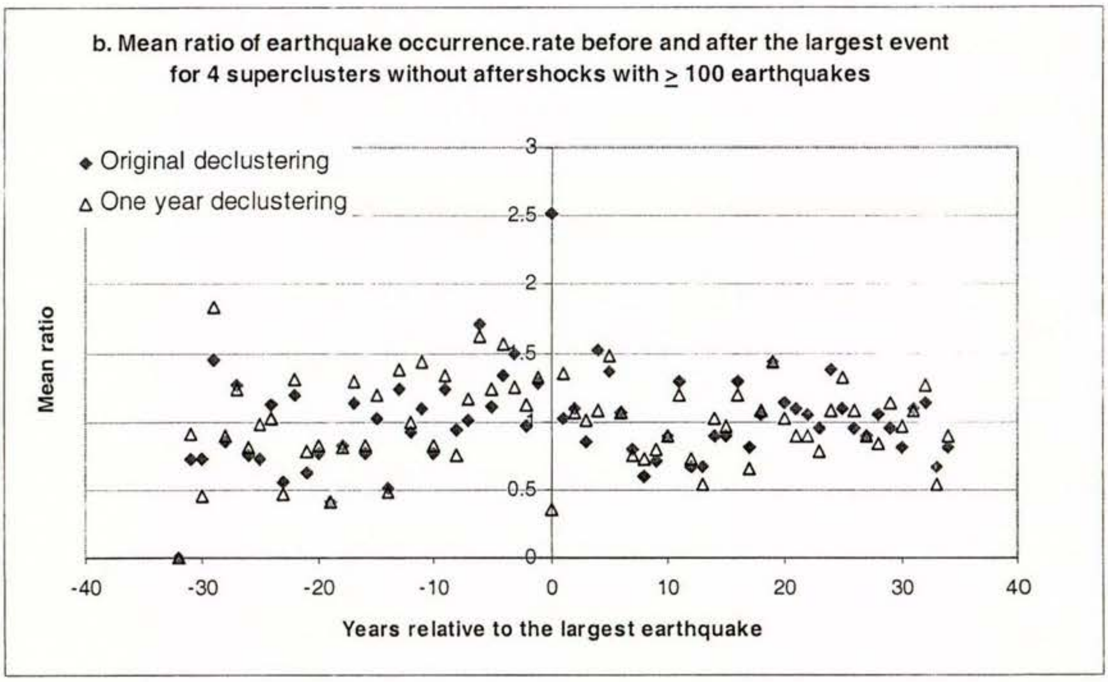


Figure 8 a and b: Mean seismicity ratios per year before and after the largest earthquake in a supercluster with 10-99 events (a) and 100 or more events (b). When the declustering period is extended to one year (triangles), the high value in year 0 disappears.

However, for the larger superclusters, which have significantly more earthquakes, the trend is slightly reversed. The average ratio for years -10 to -1 is 1.18, compared to 0.99 for years 1 to 10. We therefore interpret the difference in average seismicity rate pre- and post- the trigger earthquake for the smaller superclusters to be insignificant.

Increased seismicity in the first year

The feature that stands out most in both graphs is the significant increase in seismicity in the first 365.25 days after the largest earthquake (year 0): 5.0 for the smaller superclusters and 2.5 for the larger superclusters. We suspected this increase to be an artefact arising from inadequate aftershock declustering. Therefore we repeated the analysis with a newly declustered catalogue, using the same spatial window but with a time window which was one year after each mainshock, instead of a sliding window that was extended by 30 days for each spatially related earthquake above the cut-off magnitude of 5.0. The data are displayed in figure 8 with the black triangles. The ratio in year 0 decreases significantly to 1.8 for small superclusters and to 0.4 for large superclusters.

Other observations

In the smaller superclusters, the mean yearly ratio is much larger than average for a couple of data points well before the largest earthquake e.g. in year –31 the ratio is 4 for the original declustering. Two superclusters contribute to this data point: for one of them, the largest earthquakes was the 1997, Mw = 7.1 Pakistan earthquake. This supercluster has a total of 40 earthquakes, 5 of which occurred in 1966, 31 years earlier. We suspect that at least one of the earlier earthquakes was underestimated in size. Therefore, the initial search radius would not have been large enough to properly decluster the aftershocks.

Conclusions

We could not detect a consistent pattern of decreased or increased seismicity before or after a large earthquake, once immediate aftershocks were removed and the post-event time window was extended to one year. In our aftershock study (Christophersen 2000), 90% of post-earthquake seismicity falling within an ellipse defined by eqn (3) was deemed to be 'aftershocks'. It would be possible to define the earthquakes falling outside the aftershock space-time window as 'triggered earthquakes'. But the choice of 90% for the cut-off is arbitrary. We therefore agree with Parsons (2002) that 'triggered events' and aftershocks cannot be distinguished physically, and we therefore reject an arbitrary distinction between them.

Examples of PSI from the global catalogue

Evison and Rhoades (2004) describe a pattern of change in seismicity prior to some earthquakes of $M \geq 5.8$ in several regions worldwide, which they name Precursory Scale Increase (PSI). The change is measured by a change with time in the mean magnitude of earthquakes in a region surrounding the subsequent large earthquake, the region being 'optimised' to demonstrate the effect (Evison and Rhoades, 2004). A method for determining an optimal area algorithmically is not described in their paper. We have used a circular region centred on the mainshock with radius given by d in eqn (11) and calculated CuMag plots, using their methodology for this (Evison & Rhoades eqns 2, 3), for (an arbitrary) five of their examples.

$$\mathrm{CuMag}(t) = \sum_{\text{all events above threshold } M_c \text{ up to time } t} \{M(t) - M_c + 0.1\} - k(t - t_s)$$

where

$$k = \sum_{\text{all events above threshold } M_c} \{M(t) - M_c + 0.1\} - k(t_f - t_s)$$

and M_c is the threshold magnitude, and t_s and t_f are the start and finish times for the catalogue.

Thus the regions from which we calculate CuMag plots are different from Rhoades and Evison's. The results are shown in Fig. 9. Note that since our catalogue is for earthquakes of $M \geq 5$ only, we have very many fewer data per example than Evison and Rhoades.

The objective of this examination was not to verify the PSI phenomenon or otherwise. It was to determine whether its manifestation could be seen in the global catalogue. If it could, then a project to systematically search the global catalogue could be formulated, which would include developing the optimal search region. If PSI could be demonstrated, a predictive algorithm could then be developed.

The phenomenon is evident in the examples for Loma Prieta (Fig. 9a) and Landers (b). As Evison and Rhoades found, the CuMag falls for some years prior to the named earthquake before reversing and increasing, starting 10-11 years before Loma Prieta (1989) and about 14 years before Landers (1993). The S Cascadia example (c) may show the phenomenon too, with a termination of the falling CuMag about 7 years before the mainshock.

The other examples have too few data to demonstrate the phenomenon or deny it. In the Kobe example (e) the large increase in the CuMag in 1974 could mark the start of the process i.e. the initial decrease in CuMag. In both the W Tottori (d) and Kobe examples, no precursory increase appears at the $M \geq 5$ level, with the mainshock representing the return of the CuMag to 'normal'.

These examples suggest two tentative conclusions:

1. In at least some instances, the PSI phenomenon may be discernible using the global database
2. In other cases the global database will have too little seismicity at $M > 5$ to identify the phenomenon, if it occurs.
3. On balance, we think a systematic examination of the global catalogue for the PSI phenomenon would be worthwhile.

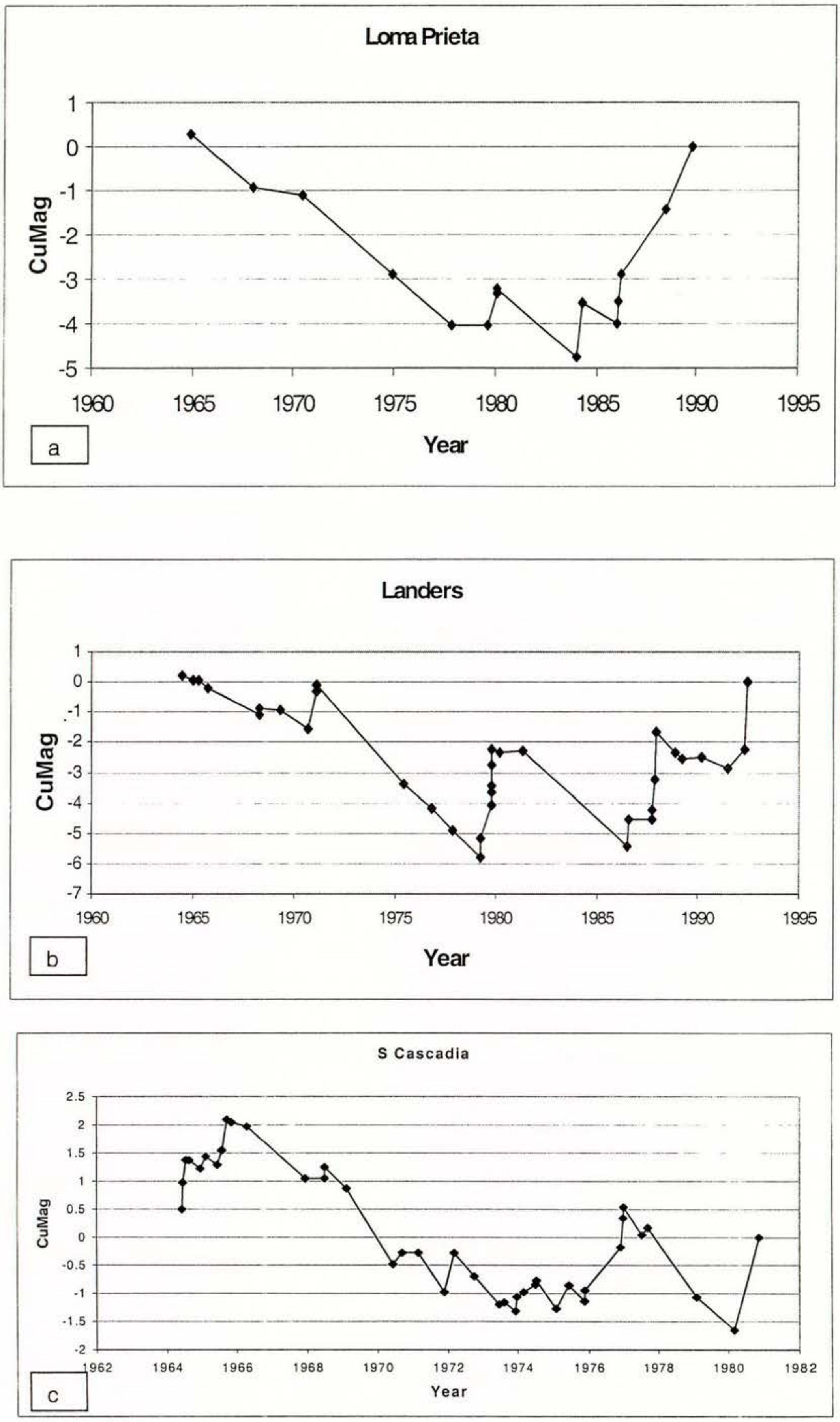
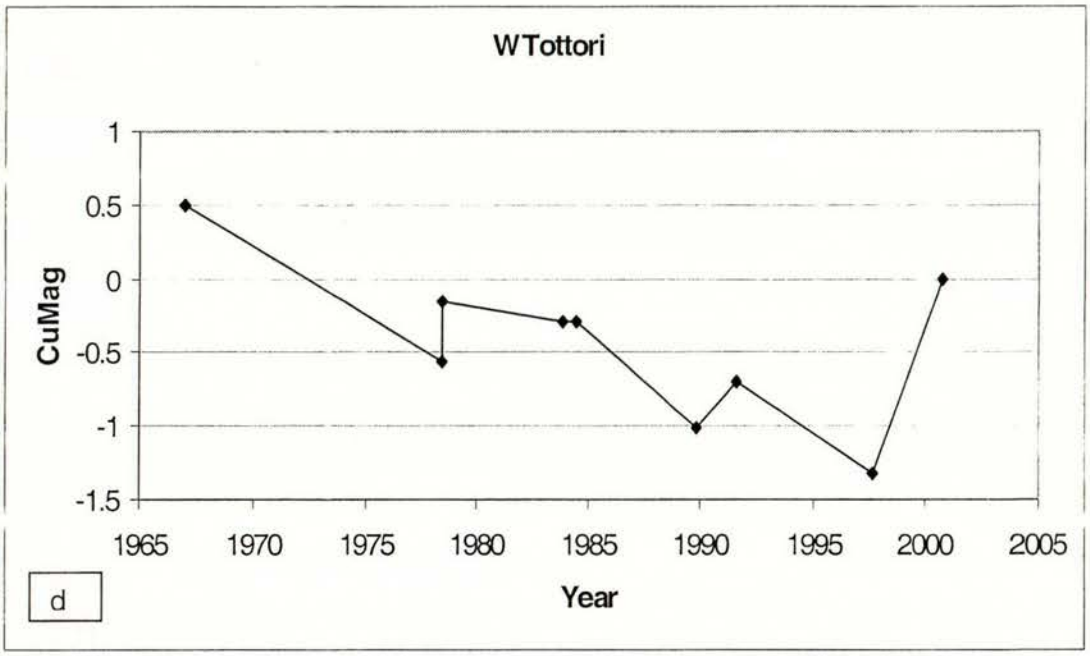


Figure 9 a, b, c. See caption below.



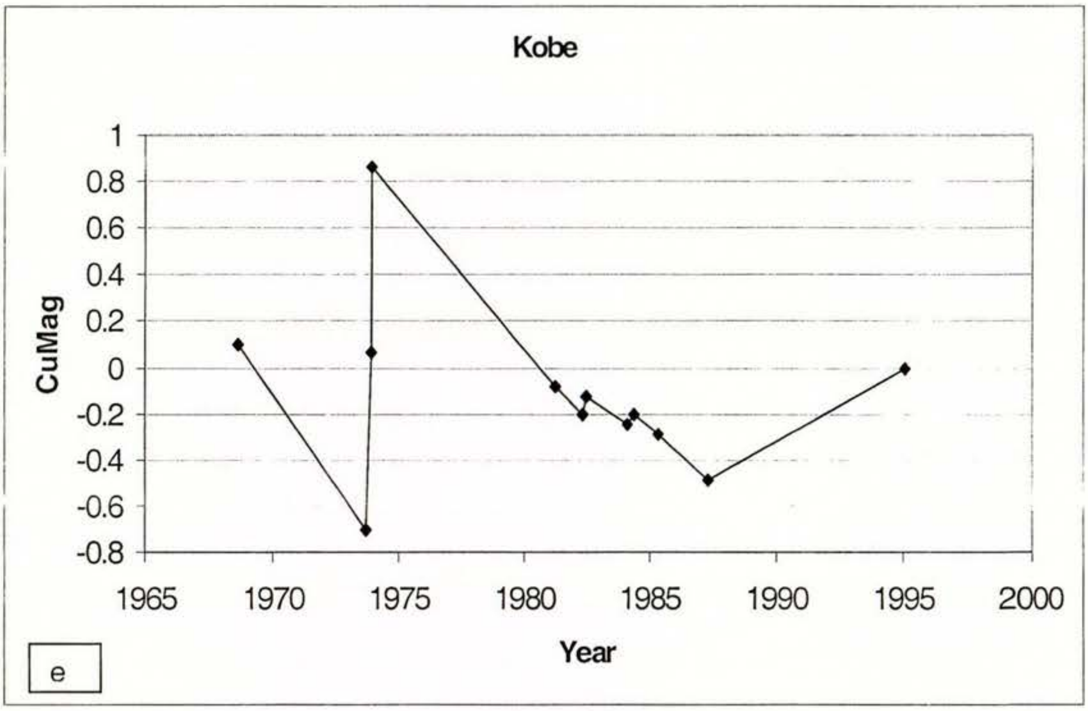


Figure 9. Five examples of CuMag plots, from and following Evison and Rhoades (2004) using the global database ($M \geq 5$) and a search area as described in the text. a. Loma Priet; b. Landers; c. S Cascadia; d. W Tottori; e. Kobe. The named earthquake occurs at the end of the CuMag graph.

A new model for earthquake interevent times

Distribution of times between M 6.5 events in the global catalogue

For the purposes of considering interevent times of related earthquakes, we assume that each Supercluster contains events that could physically influence subsequent events in the Supercluster, but not outside. Limitations of this assumption are discussed below.

From these supercluster data, datasets of interevent times were constructed: in each supercluster with aftershocks the time between successive events of $M \geq M_r$ was taken, and then the results from all superclusters were pooled. When the empirical distribution of interevent times is plotted, two parts to the distribution are evident (Fig 10). Similar results have been reported by others e.g. Musson et al. (2002).

The two parts are very straight on the log(time) scale, with best fitting lines having $R^2 = 0.997$ (R^2 = fraction of variance modeled by the line) and 0.989 for the two parts. A linear plot against log(time) for the cumulative probability implies that the density function is behaving as 1/t. This 1/t behaviour is well known as the Omori law for aftershock decay.

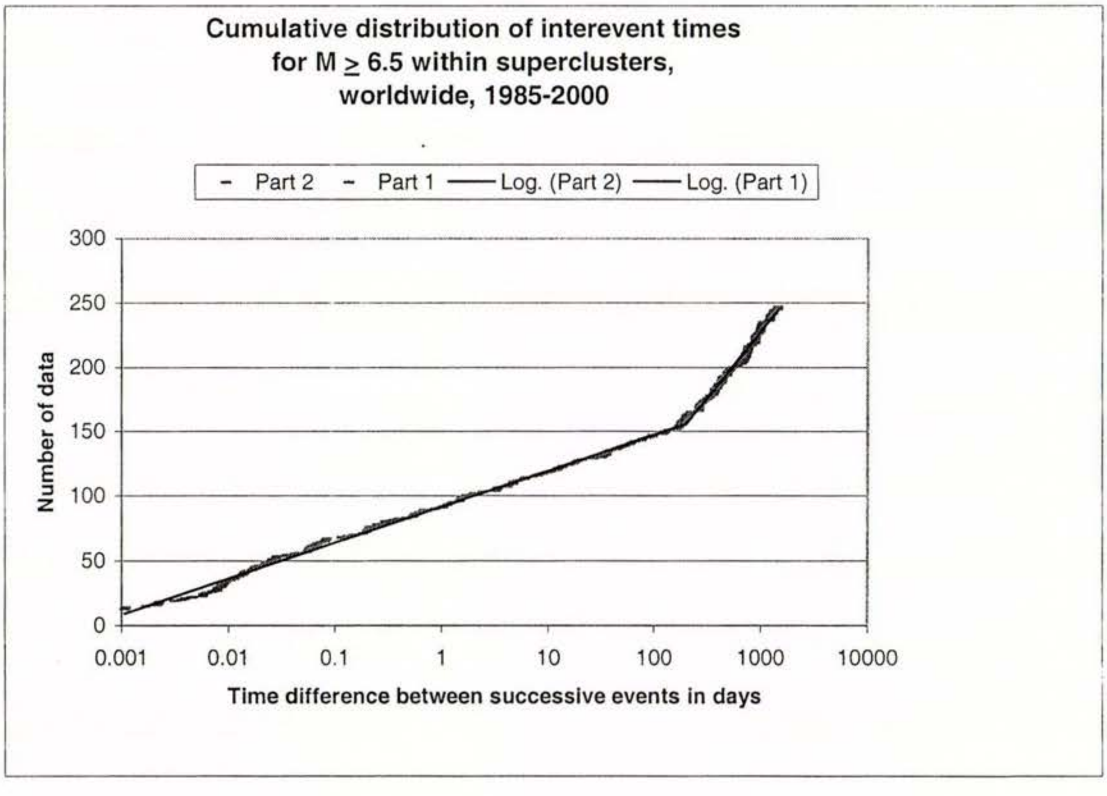


Figure 10: Cumulative distribution of interevent times for $M \geq 6.5$ earthquakes within superclusters worldwide, 1985-2000.

Therefore the first (left hand) log-linear plot corresponds to the Omori law. Note that the duration of this behaviour, about 200 days, is rather greater than the normal duration of aftershock sequences, but not greater than allowed in 'declustering' algorithms that strip aftershocks from a catalogue e.g. (e.g. Gardner and Knopoff, 1974).

The appearance of Fig. 9 suggests, therefore, that we should attempt to fit a model of the general form

$$f(t) = A/t + B/t^p \tag{18}$$

where the exponent p should be < 1 as the right-hand log-linear branch is steeper than the left one, and A and B are constants to be determined by fitting to the data e.g. by maximum likelihood.

However, there are several problems with this approach:

(1) The two parts have non-integrable singularities at both $t=0$ and $t = \infty$, meaning that, as they stand, these forms are unsuitable to be probability density functions. The problem is well known, and usually 'solved' by starting from some arbitrary start time $t_s > 0$ (often designated 'c' in the literature) and terminating at some other arbitrary end time t_e. t_s is interpreted to be a period immediately after a mainshock during which some aftershocks may not have been detected, meaning that the catalogue of aftershocks is incomplete during this time, or a transient time after initiation of the rupture of the mainshock during which the aftershock-causing process builds up. Kagan (2004) gives a comprehensive discussion of t_s, including especially the deficiencies of aftershock catalogues.

(2) t_e has other problems associated with it. It is usually defined to be the time at which rate of occurrence of aftershocks becomes equal to the 'background' i.e. to the pre-mainshock rate of activity in the region. But this approach would imply that prior to t_e all earthquakes were aftershocks i.e. the background rate was, implausibly, switched off between the time of the mainshock and t_e; and, second, that the aftershock occurrence abruptly terminates at t_e . The latter is physically implausible and not observed in practice. Continuous physical processes can only end abruptly as a result of a physical discontinuity e.g the occurrence of a mainshock abruptly disrupts strain accumulation. Some aftershock sequences have been observed to last for decades, a notable example being the aftershocks of the 1891 Nobi, Japan, earthquake, which Omori used to define his law. Utsu et al. (1995) update the data through to 1991: the 1/t decay is still observed after 100 years.

(3) An undesirable effect of introducing such arbitrary constants is that various properties of the distribution depend on them strongly. For example, the expectation of the distribution, i.e. the mean time between earthquakes, depends strongly on t_s and t_e (Sornette and Knopoff 1997). t_s represents an insignificant amount of time, and t_e is effectively arbitrary. That the properties of the distribution should depend strongly on them is not satisfactory.

Therefore, the functional form of eqn (2) was modified by the introduction of exponential modulating functions, namely:

$$f(t) = A\{1 - \exp(-t/t_s)\}\exp(-t/t_1)/t + B\{1 - \exp(-t/t_s)\}\exp(-t/t_0)/t^p \tag{19}$$

This choice of modification was motivated as follows: the term $\{1 - \exp(-t/t_s)\} \sim 1$ for times t greater than a few multiples of t_s, and so does not influence the distribution at times of interest. The term $\exp(-t/t_1)$ similarly ~ 1 for times that are a modest fraction of t_1 and so does not disrupt the 1/t behaviour at shorter times. It also represents, probabilistically, the addition of an independent random process to the 1/t process of aftershock decay and in this sense constitutes, in Bayesian terms, a neutral prior assumption.

This form is now well behaved at zero and infinity. The expectation is not strongly dependent on t_s, which may be set to a small value e.g. 0.001 day; but it does depend somewhat on t_0 and t_1, which must be determined from the data.

Choice of p = 1 in the Omori law

The modified form of the Omori law allows aftershock decay to proceed as $1/t^p$ (e.g. Utsu et al 1995). In an earlier result using a similar global catalogue, Christophersen and Smith (2000) established that $p = 1.0$ was the best average global value, but that there were regional variations in p, with a range of about 0.8-1.2. The good linear fit to the left hand part of Fig. 10 confirms $p = 1.0$ is appropriate for the 'A' term of eqn (18) and the datasets considered here. But an important point is that, since $1/t^p$ is continuous in p, the model should be stable to small variations in p about 1.0. In many formulations of Omori's law, the value $p = 1$ is excluded because of its double singularity. Distributions with $p > 1$ avoid the singularity at infinity, but their expected value is still infinite. The form of eqn (19) presents no such difficulties.

Effect of a finite catalogue

Before attempting to fit the model to a dataset, a correction may have to be made for the finite duration of the catalogue. The catalogue used for the $M \geq 6.5$ events had a duration of 16 years, which is long compared to the mean time between M 6.5 events (~ 500 days; see Table 3), but not necessarily long compared to the mean time for larger events e.g. M 7. In such a case the set of observed time intervals is biased against longer times i.e. times which are a significant fraction of the catalogue duration t_c - and no times longer than t_c can be observed at all. A correction can be approximated as follows. If the probability density of inter-event times is f(t), then the probability of observing a pair of earthquakes $\leq$ t apart in time is given by:

$$\int_0^t (1 - \Delta / t_c) f(\Delta) \, d\Delta \Big/ \int_0^{t_c} (1 - \Delta / t_c) f(\Delta) \, d\Delta \qquad (20)$$

and so the corresponding density function

$$(1 - t / t_c) f(t) \Big/ \int_0^{t_c} (1 - \Delta / t_c) f(\Delta) \, d\Delta \qquad (21)$$

should be used for the maximum likelihood fitting procedure. The function converges to f(t) as $t_c \rightarrow \infty$.

Note that this is only an approximate correction as it applies where there are just two events in a particular supercluster, and allows for the amount of time before the first occurs.

Outcome from trial model fitting

When eqn (21) was used to try to fit the data, it quickly became evident that the best fit was to be found with p = 0, i.e. the second part of the distribution is just

$$B\{1 - \exp(-t/t_s)\}\exp(-t/t_0) \tag{22}$$

and since it comes into effect for times long compared to t_s, and it has no singularity at $t = 0$, the term in the braces can be dropped.

This is an important result. Our assumption was that the behaviour of the right-hand branch of Fig. 10 should be represented by an exponentially modulated power law. With $p = 0$, the power law degenerates, and we are left with a negative exponential function. The straightness of the right hand segment derives mainly from the functional form of (21), when the correction for a finite catalogue is applied to (19). This means that in the family of function forms just described, the *best* fit to the data is given by a random process. The implications of this will be discussed below.

Thus the model form adopted for the fitting was the mixture:

$$f(t) = w_1 f_0\{1 - \exp(-t/t_s)\}\exp(-t/t_1)/t + w_2(1/t_0)\exp(-t/t_0) \tag{23}$$

where normalising constants f_0 and $1/t_0$ have been introduced, so that $w_1 + w_2 = 1$.
f_0 is given, to a good approximation, by:

$$f_0 = 1/\{E_1(a/t_s) + \log(a/t_s) + \gamma + E_1(a/t_1)\} \tag{24}$$

where E_1 is the exponential integral (see Abramowitz and Stegun, 1965, p228, 5.1.1 and p230 5.1.39)

$$E_1(x) = \int_x^\infty \exp(-u)/u \, du \tag{25}$$

γ is Euler's constant 0.57721566, and a is a constant chosen so that $t_s \ll a \ll t_1$. f_0 is then not sensitive to a.

Derived results

Mean interevent time

The form of eqn (23) makes for ready calculation of various quantities of interest. For example, the expected value (mean interevent time) is given by:

$$\mathcal{E}(t) = w_1 f_0 t_1 + w_2 t_0 \tag{26}$$

This is consistent with the standard, 'aftershock stripping' approach to interevent time modeling: If all the aftershocks are removed then $w_1 = 0$, $w_2 = 1$ and $\mathcal{E}(t) = t_0$.

We can also use eqn (23) to calculate the probability distribution for time to the next earthquake given a time t_L since the last, as follows.

Interevent time conditional on the time since the last earthquake

Bayes theorem says that: Probability (time to next event $\leq t \mid t_L$)

= Probability (time to next $\leq t$ and t_L since the last) / Probability (time to the last $\geq t_L$)

= Probability ($t_L \leq$ time to next $\leq t + t_L$) / (1- Probability (time to the last $\leq t_L$))

If F(t) is the distribution function for the density function given by eqn (23), i.e.

$$F(t) = \int_0^t f(\Delta)\, d\,\Delta$$

Then

$$F(t \mid t_L) = \{F(t + t_L) - F(t_L)\}/\{1 - F(t_L)\} \qquad (27)$$

Note that for $t_L = 0$ this reverts to being F(t). Examples of the use of eqn (27) are given in 'Conditional interevent time probabilities: an example'.

Application to the global catalogue

Fig. 11 a-c shows the results of fitting the model eqn (23), modified by eqn (21) to the time intervals between all events of $M \geq 6.5$ in each supercluster with aftershocks, derived from catalogue 2: 1985-2000. This yielded 341 interevent times, ranging from 0.0011 to 4471 days, with interevent times less than the t_s time = 0.001 days being discarded as being, in some cases, the result of duplicate entries of the same event.

Fitting is by the method of maximum likelihood where we seek the maximum, of the log-likelihood:

$$\mathcal{L} = \sum_{i=1}^{N} \log f(t_i) \qquad (28)$$

where there are N data.

An approximate 95% confidence limit is given as follows (Severini 2000, p113). Define W:

$$W = 2\,\{\,\mathcal{L}\,(\text{best fit parameters}) - \mathcal{L}\,(\text{other, specified parameters})\,\} \qquad (29)$$

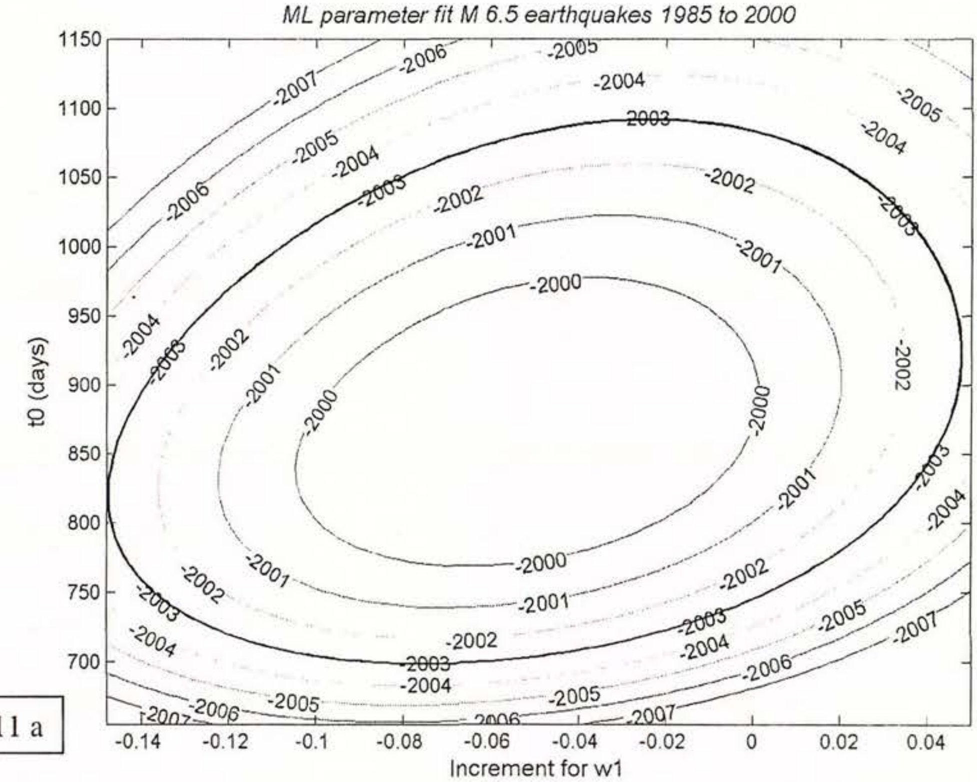
then W is asymptotically distributed as χ^2 with k degrees of freedom, where k is the number of parameters – here 2. The 95 percentile of χ^2 (2) is 5.99, and so the 95% confidence region is given by the value of $\mathcal{L} = 5.99/2 = 3$ smaller than the maximum (Fig. 11a).

The contributions to the model are most evident in c, which is similar to Fig 10. Aftershocks dominate at times up to about 500 days, after which the occurrence of 'new' random events becomes the main contributor. But note that 'new' events contribute from t = 0, and the aftershocks are not artificially truncated. Aftershocks may occur, with decreasing probabilities, at any interevent time.

Resolving t_s and t_1

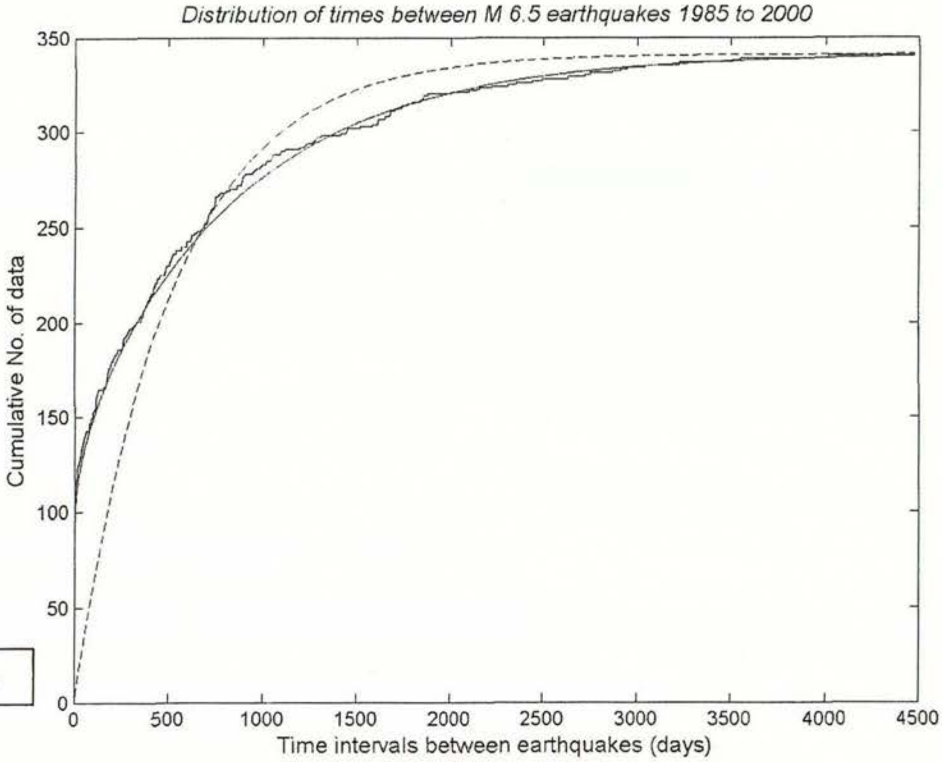
Table 3 gives the results of fitting the model to various datasets, and illustrates the sensitivity of the model, and its expectation, to changes in t_s and t_1.

First, the models for different M_r values are consistent with each other and have expected values (means) that are similar to the means of the data. The last three columns of the Table illustrate this: the first two of these compare means calculated from the model and from the data, and the last column gives the mean for $M_r = 6.5$ assuming a Gutenberg and Richter distribution of magnitudes, with $b = 0.84$ (see above) in which case the mean for M_r 6.5 = $10^{-0.84 \times (7.0-6.5)}$ times the mean for M_r 7. Thus for $M_r = 6.5$ for the case $t_1 = t_0$, the expected value from the distribution is 507 days, the mean of the data is 515 days and the mean predicted, by taking the mean of the M_r 7 data = 1300 days, is 494 days. The differences are not significant.

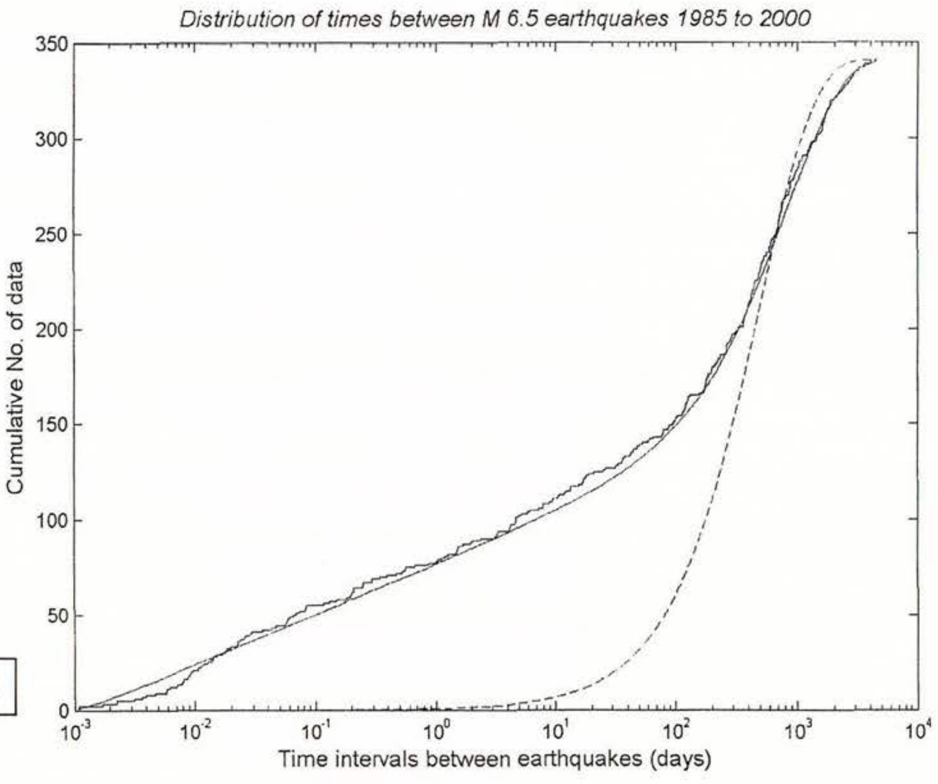


11 a

Figure 11: Interevent times of $M \geq 6.5$ earthquakes from the Harvard Catalogue 1985-2004.
a: Log-likelihood contours for w_1 increments and t_0 in model (23) modified according to (21), with $t_s = 0.001$ day and $t_1 = t_0$. w_1 values are to be added to 0.5. The heavy line gives an approximate 95% joint confidence region for the parameter (Severini, 2000).
b and c: Cumulative distribution functions for interevent times: data (stepped), model (23) (solid) and ML fit Poisson models (dashed) on linear (b) and log(time) scales (c).


Distribution of times between M 6.5 earthquakes 1985 to 2000
Cumulative No. of data
Time intervals between earthquakes (days)
0
50
100
150
200
250
300
350
500
1000
1500
2000
2500
3000
3500
4000
4500

11 b


Distribution of times between M 6.5 earthquakes 1985 to 2000
Cumulative No. of data
Time intervals between earthquakes (days)
0
50
100
150
200
250
300
350
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}

11 c

The effect of increasing t_s to 0.01 day is shown in lines one and two. As might be expected, t_0 is not altered, but w_1 is, decreasing by 0.05 for $t_s = 0.01$. As a result the expected value increases from 1260 to 1350. Neither of these changes is significant.

The model parameter t_1 is not easily resolved. The value of t_1 affects the time during which the Omori law is the greater contributor to the model. However, the Omori decay of 1/t is the primary factor decreasing aftershock occurrence, and so there is not much sensitivity to t_1, in either the fit or in the subsequent properties of the model, such as its expected value. This is illustrated in Fig. 12 which shows log-likelihood contours for values of $t_1 = 5\ t_0$ and $0.2\ t_0$. These should be compared to Fig 11 a where $t_1 = t_0$. It can be seen that $t_1 > t_0$ has the effect of worsening the fit (ML = -2001 cf –2000 for $t_1 = t_0$), increasing w_1 slightly (by about 0.03) and t_0 decreasing slightly (870 to 790). The changes are not significant. For $t_1 < t_0$ the changes are also small: the fit is only marginally worse (smaller ML –2000 contour), with a small decrease in w_1 and t_0.

A possible discriminant is, however, offered by the expected value of the distribution. The final three columns of Table 2 show that $t_1 = 5\ t_0$ and $t_1 = 0.2\ t_0$ produce ML models with expectations matching the data mean of 505 days worse than $t_1 = t_0$ case. In view of this and the rather poorer log-likelihood fits, we adopt $t_1 = t_0$ as a probable best value. If truly $t_1 = t_0$, the obvious question is why this should be so? *A priori* there is no reason for the (long) time constants of the two parts of the model to be the same. If they were truly the same it would suggest that there is a physical linkage between the decay of aftershocks and the generation of new earthquakes.

Table 3: Datasets modelled and resulting parameters and derived results.

Dataset	No. in set	t_0 day	Approx. 95% conf. $\pm$	w_1	Approx. 95% conf. $\pm$	t_s day	t_1/t_0	Mean time (days) from model	Mean time (days) from data	Mean predicted from M 7 mean 1300
$M \geq 7.0$ 1977-2000	140	1750	450	0.3	0.11	0.001	1	1260	1316	
$M \geq 7.0$ 1977-2000	132	1750	450	0.25	0.11	0.01	1	1350	1396	
$M \geq 7.0$ 1985-2000	79	1550	600	0.46	0.19	0.001	1	887	906	
$M \geq 6.5$ 1985-2000	341	870	200	0.45	0.1	0.001	1	507	515	494
$M \geq 6.5$ 1985-2000	341	800	200	0.49	0.1	0.001	5	433	515	494
$M \geq 6.5$ 1985-2000	341	840	200	0.4	0.1	0.001	0.2	532	515	494

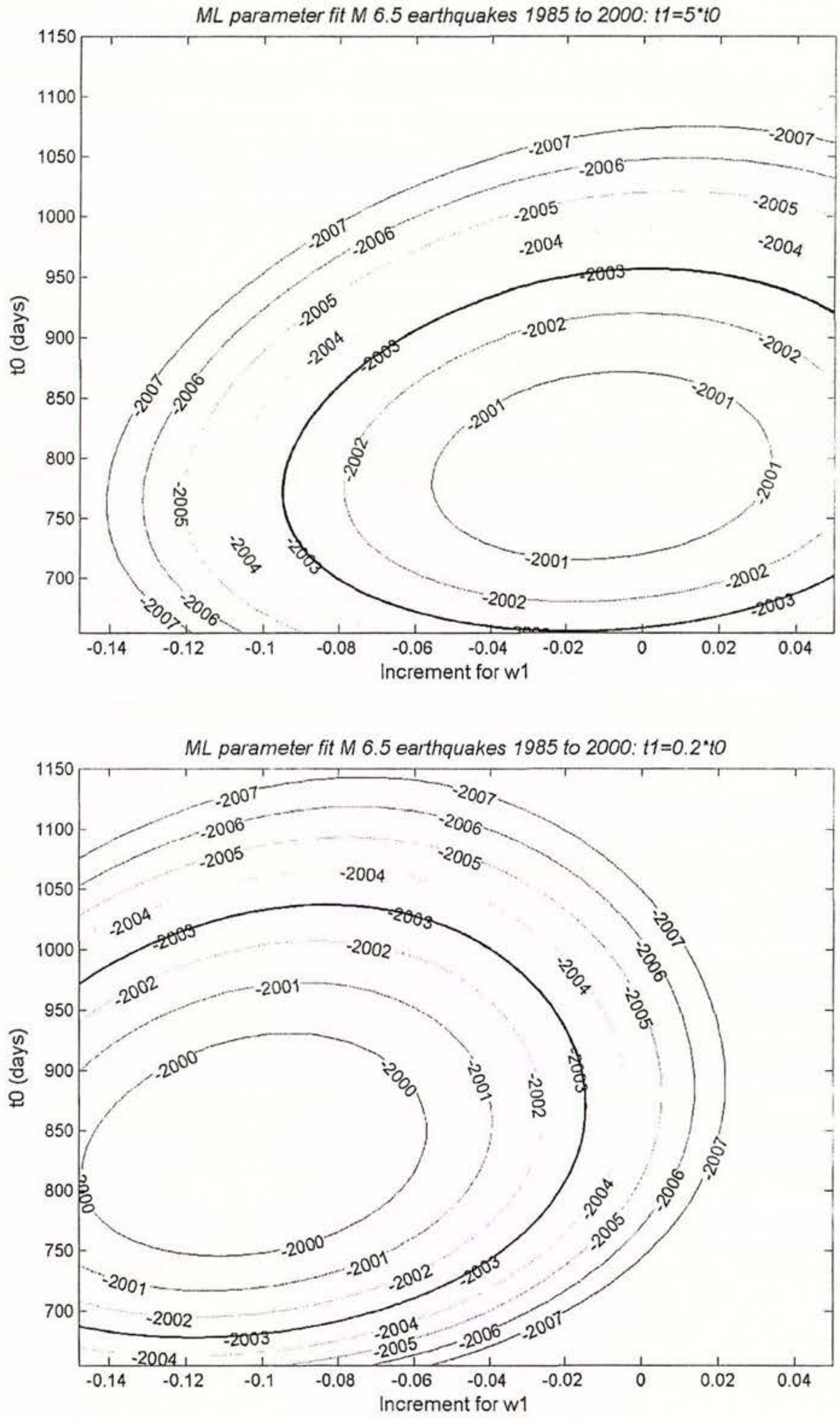


Figure 12: Log-Likelihood contours for models with see eqn (23); cf. Fig. 11 a.
a. $t_1 = 5\ t_0$ *b.* $t_1 = 0.2\ t_0$.

Application to New Zealand

An additional dataset available to provide a test for the model is the set of $M \geq 7$ earthquakes in the region of New Zealand which has been seismically active with large earthquakes since European times, from northern Hawke's Bay to Arthur's pass (Table 4). The largest earthquake in this set is the M ~ 8.2 Wairarapa earthquake of 1855 (Grapes and Downes, 1997). In accordance with the criteria used to construct the superclusters, the region of influence for this event would have a radius of 570 km (c.f. Table 2) centred on the 1855 event, and it would encompass the area just described, but not southwest of about 44 S, 169 E or northeast of about 37.5 S, 179.5E. The 1995 M 7.0 event off East Cape is just at the margin of the supercluster and is not included (no other event in Table 4 is more than 300km away). A list of events of $M \geq 7$ in the supercluster is given in Table 4. There are only 15 such events, giving 14 interevent times. This dataset then constitutes a supercluster of $M \geq 7$ events for the period 1843 to the present.

Table 4: Earthquakes of $M \geq 7$ in central New Zealand (Gisborne to Arthur's Pass) since 1840 (after Smith, 1994).

No.	Name/ Region	Year	Month	Day	Time interval (days)	Latitude	Longitude	Distance from 1855 (km)	Magnitude
1	Wanganui	1843	7	8		-39.9	175.0	166.7	7.5
2	Marlborough	1848	10	15	1923.3	-41.5	173.8	100.5	7.1
3	Wairarapa	1855	1	23	2294.8	-41.4	175.0	0.0	8.2
4	Hawkes Bay	1863	2	22	2951.0	-40.0	176.5	200.4	7.5
5	Cape Farewell	1868	10	18	2062.3	-40.0	173.0	229.3	7
6	N. Canterbury/ Glenn Wye	1888	8	31	7258.0	-42.6	172.3	259.7	7
7	Nelson	1897	12	7	3383.3	-40.0	175.0	155.5	7
8	Arthur's Pass	1929	3	9	11420.0	-42.8	171.9	296.5	7.1
9	Buller	1929	6	16	97.0	-41.7	172.2	235.2	7.8
10	Hawke's Bay	1931	2	2	596.5	-39.3	177.0	288.3	7.8
11	Hawke's Bay (aftershock)	1931	2	13	11.0	-39.5	177.5	298.7	7.3
12	Pahiatua	1934	3	5	1117.8	-40.6	176.3	143.6	7.6
13	Wairarapa I	1942	6	24	3031.0	-40.9	175.9	93.6	7.2
14	Wairarapa II	1942	8	1	37.0	-41.0	175.8	80.3	7.0
15	Inangahua	1968	5	23	9428.5	-41.8	172.0	249.2	7.4

Fitting the model (23) to these 14 intervals by maximum likelihood yielded the parameter values $w_1 = 0.13$ and $t_0 = 3500$ days (9.6 years). Figures 13 a-c show the maximum likelihood contours, and fit of the model against the data in comparison with a Poisson model. The following are evident:

- The parameter uncertainties from the fitting are large. An approximate 95% confidence interval for w_1 and t_0 would be given by the contour –3 smaller than the centre of the ML plot (i.e. value –131.3) – but this is well outside the range of this figure. This arises principally because there are so few data.

- Notwithstanding this, the w_1 parameter is greatly smaller than that found from the global datasets, where $w_1 \sim 0.3$ - 0.45 (see Table 3, lines 1 - 2). Either the lower value is (approximately) correct for New Zealand, which would imply a magnitude-frequency relationship in New Zealand rather different from the global average or, more likely, that the New Zealand catalogue of $M \geq 7$ events is not complete; in particular, some short interevent time M 7 earthquakes have been missed. Since about 1/3 of the catalogue is pre-instrumental, this is not as implausible as it may seem. The 1848 Marlborough earthquake was part of a complex sequence in which some aftershocks were large and had a pattern of felt reports suggesting a spread of locations (Grapes et al 2004). It is possible that at least one additional M 7 earthquake occurred in the 1848 sequence.

  Similarly in the aftermath of the 1855 earthquake a large (M 7) aftershock would not have gone unnoticed but it would have been difficult to assess for size. The occurrence of one or more M 7 aftershocks in 1855 is thus plausible.

Figure 14 (c.f. fig 13) shows the effect of modeling a synthetic dataset in which three additional times of 1, 2 and 4 days have been added to the New Zealand M 7 dataset. The substantial effect is that w_1 changes from 0.13 (increment –0.37) to about 0.36 (increment –0.14). At the same time, t_0 increases (insignificantly) from 3500 to 3800 days because even though there are now more short time intervals, the Omori law part of the distribution gets more weight.

Table 5: Comparison of model parameters and derived results for the best fitting New Zealand model (c.f. Fig. 13) and one fitted to synthetic data that include the 14 actual interevent times and three additional, short.times.

Dataset	No. in set	t_0 day	Approx. 95% conf. $\pm$	w_1	Approx. 95% conf. $\pm$	t_s day	t_1/t_0	Mean time(days) from model	Mean time (days) from data
NZ $M \geq 7$ 1843-2004	14	3500	Large	0.13	Large	0.001	1	3075	3258
NZ $M \geq 7$ plus 3 synthetic events	17	3800	Large	0.36	Large	0.001	1	2520	2683

The results are illustrated in Table 5 (c.f Table 3). In the second row, the results of modeling the synthetic data i.e. real data plus the three imagined events, are shown. Both models under-predict the mean interevent time slightly (last two columns), but not significantly.

The conclusion from these considerations is that the development of a useful model for New Zealand, or any region for that matter, is not straightforward. Doubt about the possible omission of short interevent times from the catalogue suggests using a different dataset for the model e.g. $M \geq 6$ events since 1940, which catalogue is thought to be substantially complete (but we note Kagan's 2004 conclusions about catalogue completeness at short interevent times). Attempting to model the $M \geq 6$ events would run into another difficulty already mentioned: that the spatial relationship between events in the same supercluster is an approximation with obvious shortcomings. These are discussed below.

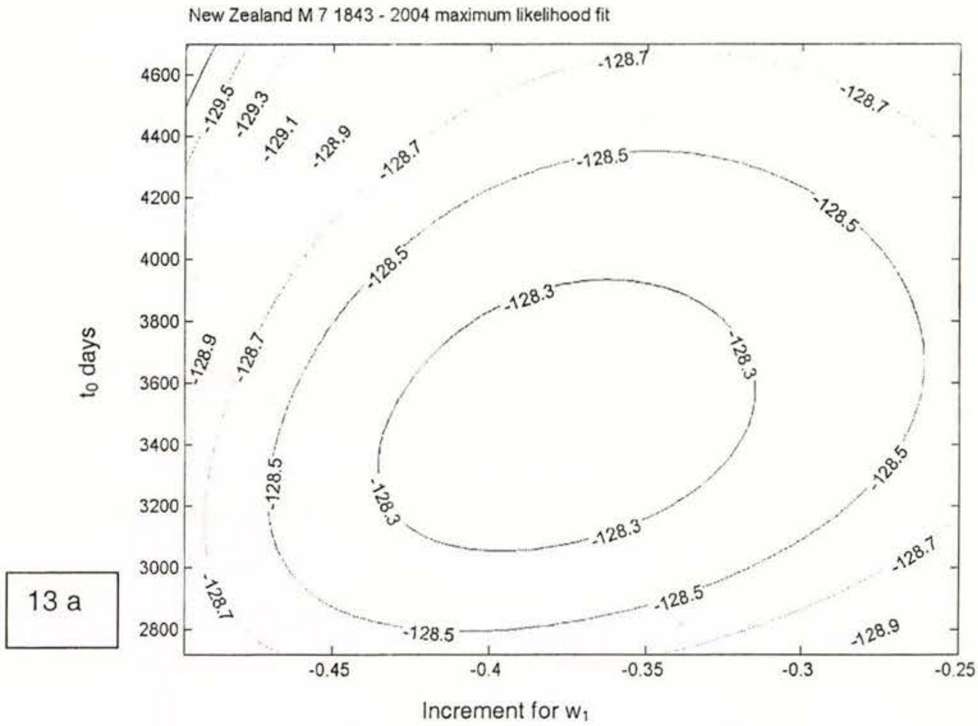
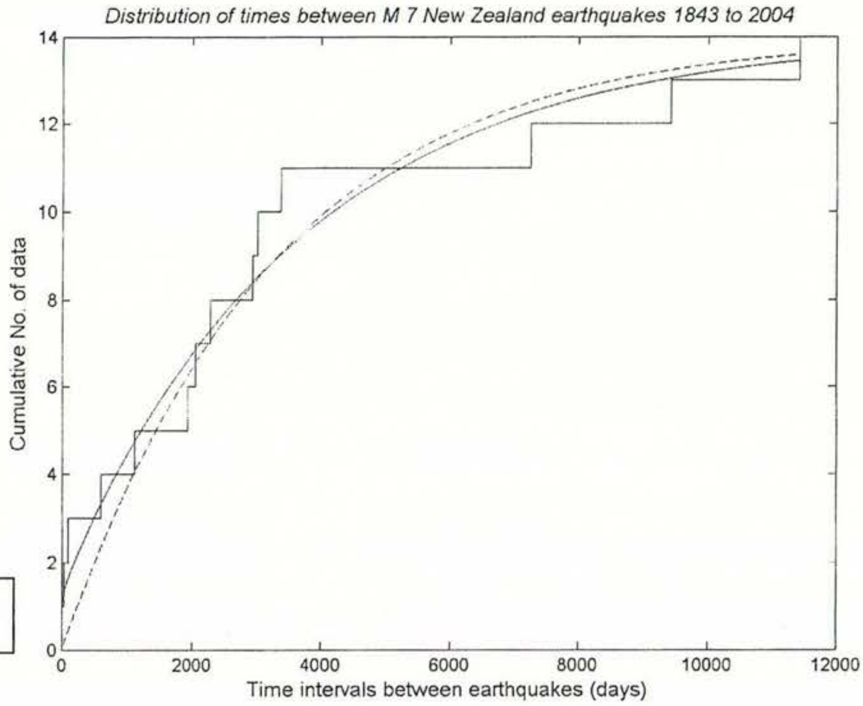
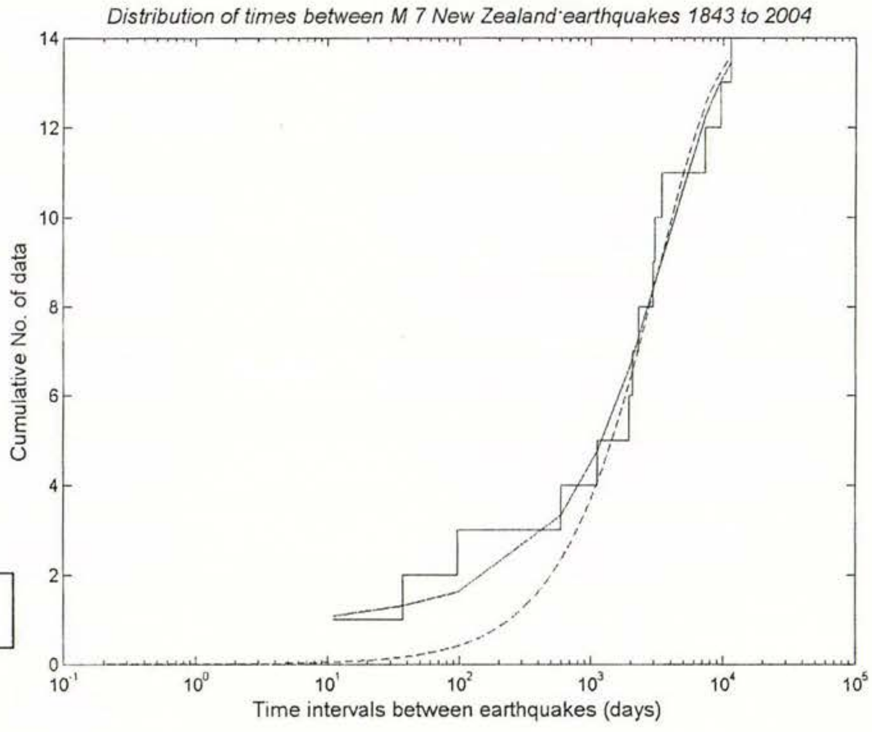


Figure 13: Results of fitting the model eqn (23) with correction eqn (21), with t_s = 0.001 day and $t_1 = t_0$, to the New Zealand interevent times, Table 4. a: Log-likelihood contours for w_1 increments, to be added to 0.5. A 95% joint confidence region for the parameters is beyond the edge of the plot. b and c: Cumulative distribution functions for interevent times: data (stepped), model (23) (solid) and ML fit Poisson models (dashed) on linear (b) and log(time) scales (c).

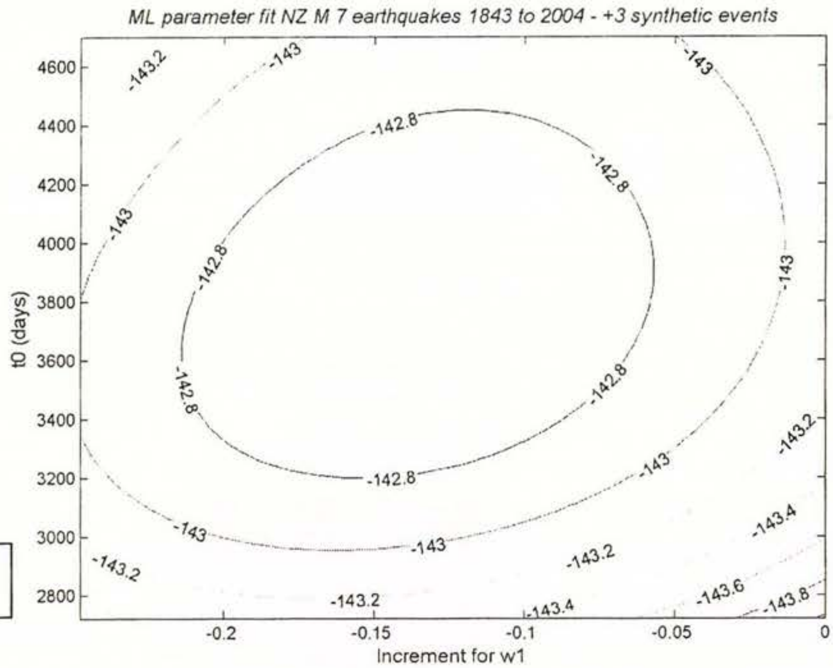

Distribution of times between M 7 New Zealand earthquakes 1843 to 2004
Cumulative No. of data
Time intervals between earthquakes (days)
0
2
4
6
8
10
12
14
0
2000
4000
6000
8000
10000
12000

13 b


Distribution of times between M 7 New Zealand earthquakes 1843 to 2004
Cumulative No. of data
Time intervals between earthquakes (days)
0
2
4
6
8
10
12
14
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}

13 c

14 a



14 b

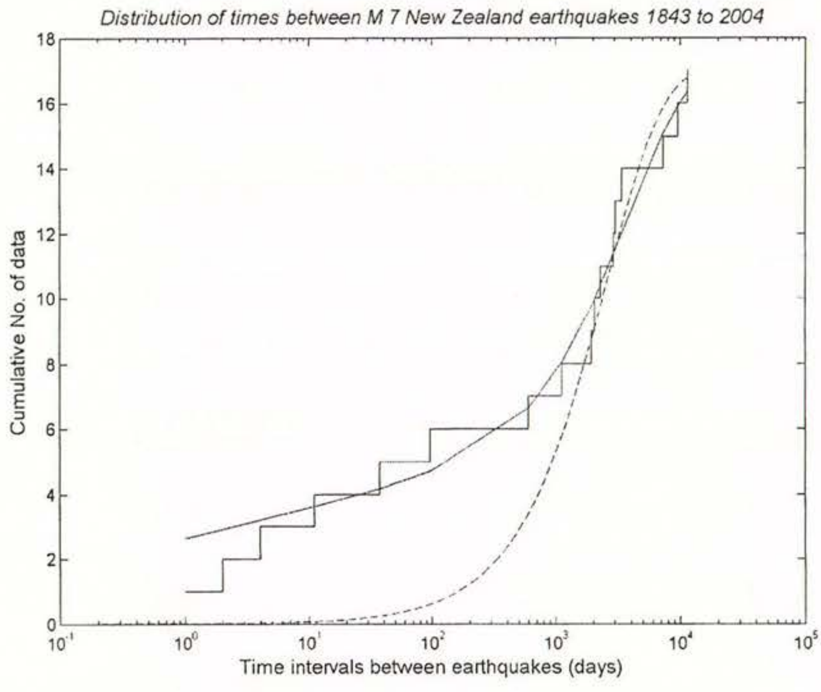


Figure 14: Similar to Fig. 13, except that the dataset is the synthetic New Zealand dataset described in the text. b. of Fig. 13 is omitted.

Spatial relationship between events

We have assumed an unspecified but equal physical connection between all the events in a supercluster with aftershocks. We justify this assumption on the basis of the progress with the interevent modelling that has resulted. But the assumption is clearly poor: smaller events at the extremities of a supercluster for which the largest event has M > 8 cannot be expected to have significant influence on each other.

However, it would be a mistake to prescribe the spatial relationship too tightly. As is evident from Table 4, the whole of central New Zealand became active during the period 1929 to 1942, with New Zealand's second equal largest events (1929 Buller and 1931 Hawkes Bay mainshock; Table 4) occurring within just two years at a separation of more than 450 km. Therefore, a spatial relationship model that allows for possible long range interactions (e.g. Kagan and Jackson 1999) needs to be carefully developed. Possible functional forms for such a relationship have been recently discussed, e.g. Console et al (2003), who prefer a power law kernel to a Gaussian kernel for modelling Italian seismicity - see eqn (8).

Conditional interevent time probabilities: an example

As we have discussed, the models that we have developed so far have limited application because of the deficiencies we have acknowledged. However, in this section we present an example of the recurrence time probability calculations that the models will enable when the difficulties have been overcome.

For this purpose we use the artificial model described by the parameters in the second line of Table 5. We use this rather than the line 1 model because we think the line 2 model, with its additional synthetic data, is closer to global models, such as lines 1-2 of Table 3 and, accordingly, is more likely to represent the actual situation in New Zealand than the line 1 model. Nonetheless, it is a synthetic model. The purpose of the example is to demonstrate the effect of the model, not to predict conditional interevent time probabilities for New Zealand.

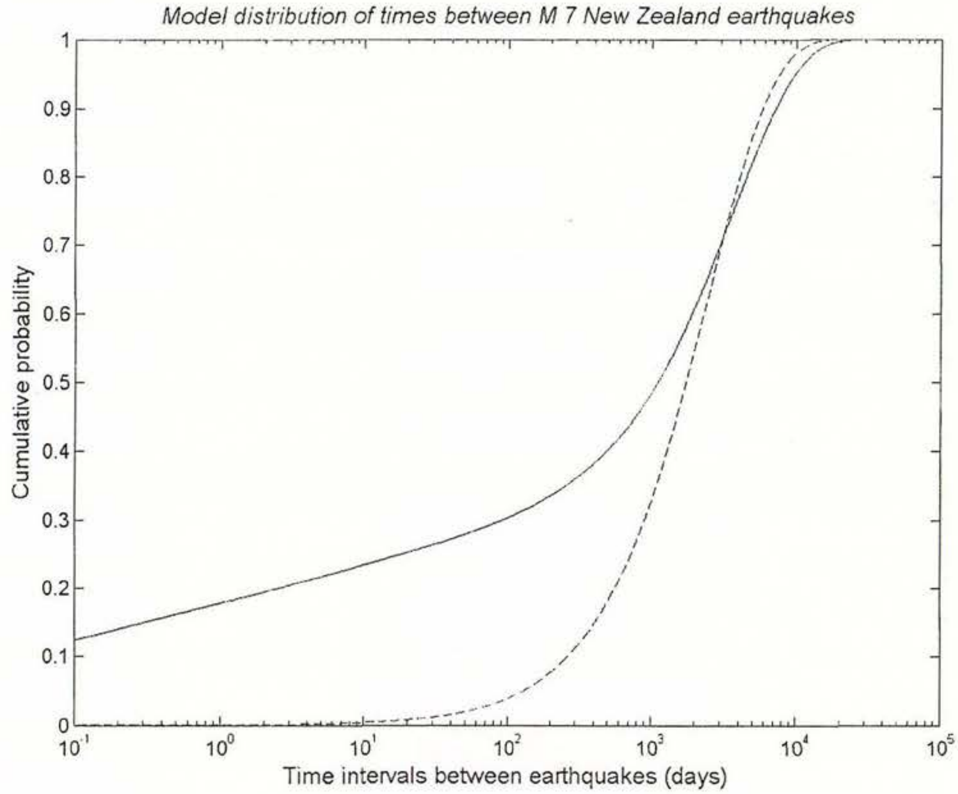


Figure 15: Probability distribution function of the model of line 2 in Table 5 (solid). A best – fit Poisson model for the same data is shown dashed.

Figure 15 compares the model, using Table 5 line 2 parameters, and Poisson probability distributions for interevent times. The condition that applies is than an $M \geq 7$ earthquake has just occurred. Note that model probabilities for interevent times are greater than Poisson until about 3000 days, when model probabilities become, and stay, less than Poisson.

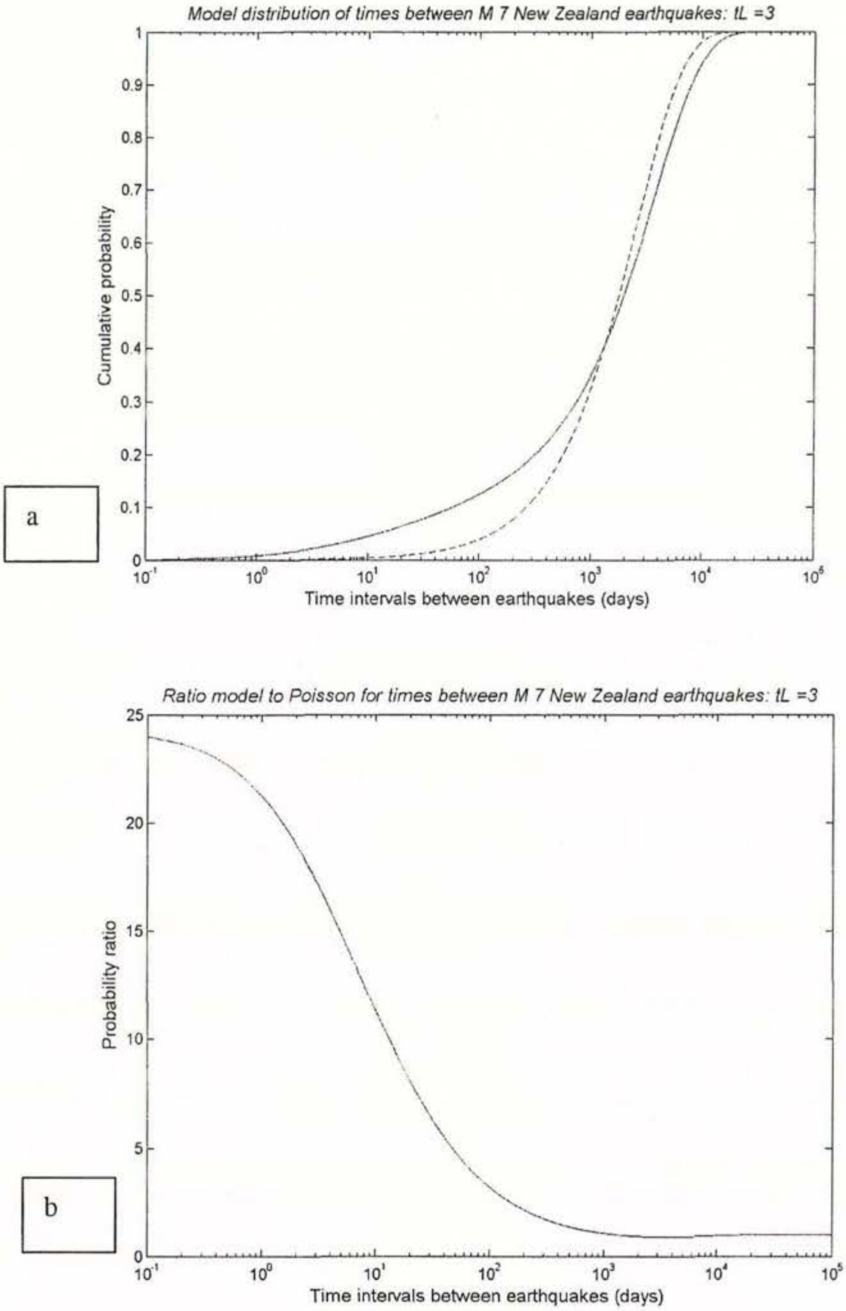
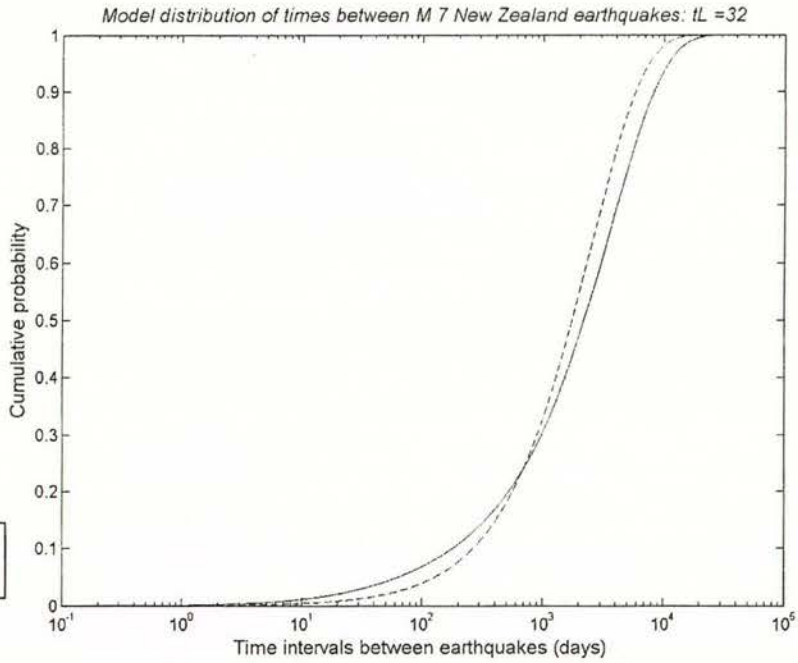


Figure 16-1: Comparison of model, using Table 5 line 2 parameters, and Poisson probability distributions for interevent times, given an elapsed time of $t_L = 3$ days since the last $M \geq 7$ event. a: model distribution (solid) and Poisson distribution (dashed. b: ratio of model to Poisson probabilities.

a



b

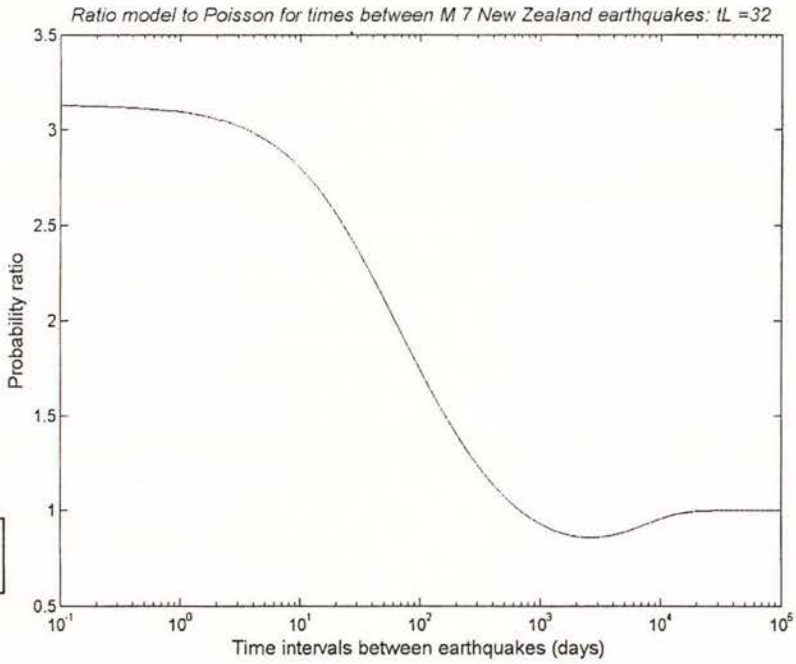
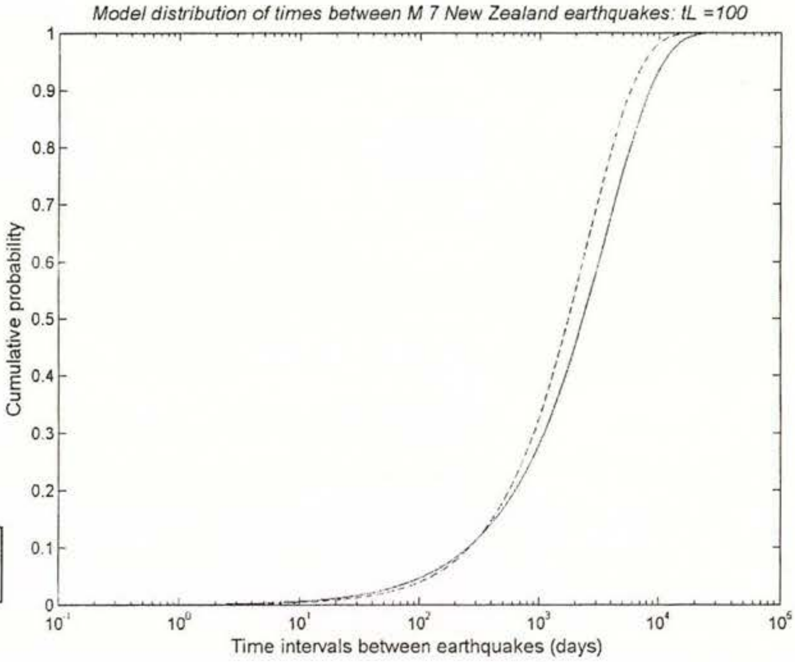


Figure 16-2: As for Fig. 16-1 except that the elapsed time t_L = 32 days.

a



b

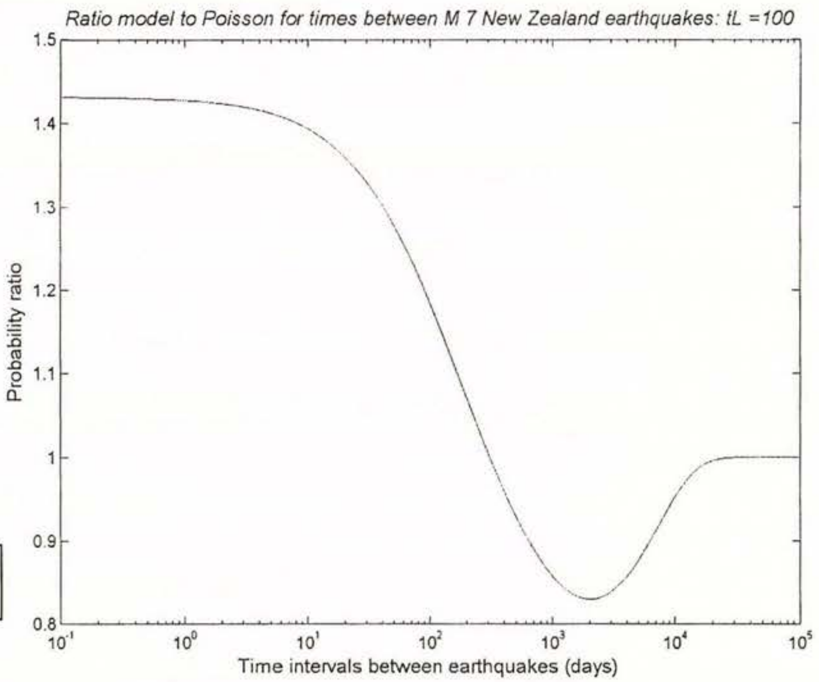
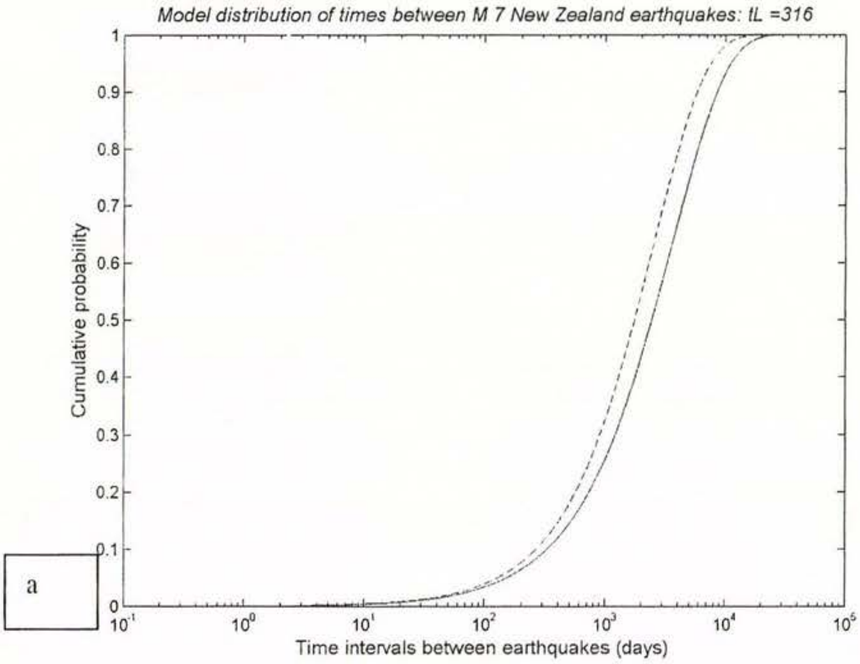


Figure 16-3: As for Fig. 16-1 except that the elapsed time $t_L = 100$ days.



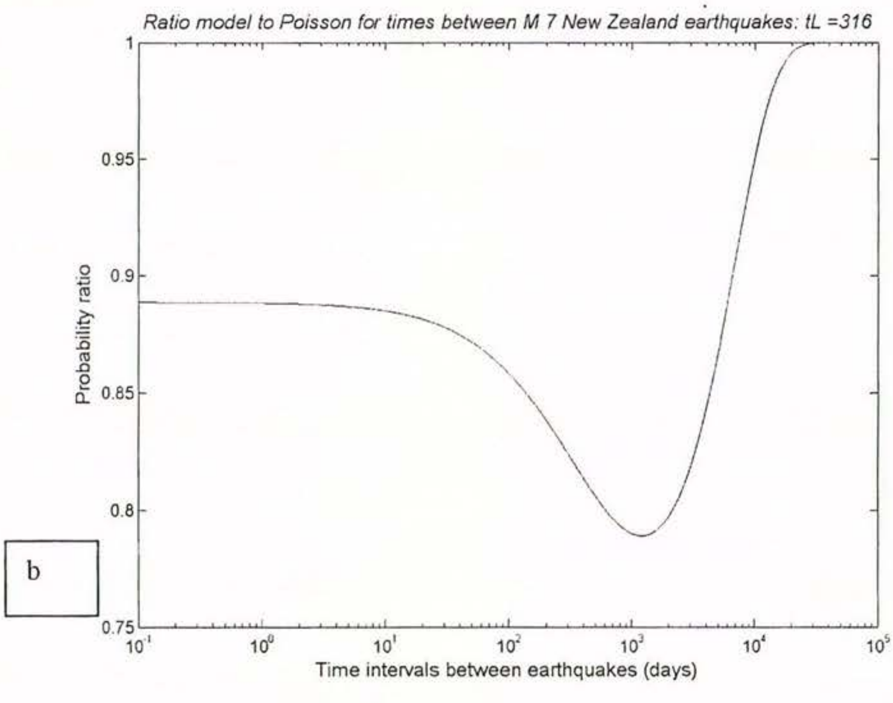


Figure 16-4: As for Fig. 16-1 except that the elapsed time $t_L = 316$ days.

Figs. 16-1 - 4 show how the interevent time probability distribution changes with different elapsed times since the last earthquake. In each case, a time t_L has elapsed since an $M \geq 7$ earthquake, and the distributions give the probabilities for the times to the next $M \geq 7$ event. To make the comparisons easier, Fig. b. in each case gives the ratio of model to Poisson probabilities. The changes to the conditional distributions are continuous and steady with $\log(t_L)$, and in Fig. 16 the elapsed times t_L have been chosen to illustrate the general behaviour.

Notice that in all cases except perhaps $t_L = 3$, the model probabilities are little different from Poisson in absolute terms.

For an elapsed time of about 3 days since an M 7 earthquake, probabilities are initially 24 times higher than Poisson. The ratio falls steadily until it is about 1 at interevent times of about 1000 days. Then probabilities become, and stay, less than Poisson.

For longer elapsed times, the probability ratios are initially lower. For an elapsed time t_L of about 30 days, probabilities are up to 3 times higher than Poisson up to interevent times of about 500 days.

$t_L = 100$ days approximately marks a transition. For longer elapsed times, the probability ratio is less than Poisson for all intervent times i.e. the probability of the next earthquake occurring within any specified time is lower than for Poisson. In Fig 16-4, where t_L is about 300 days, probabilities are only 80-90% of Poisson until interevent times of about 5000 days.

This behaviour can be summed up as follows: Following a large (M 7) earthquake, the probability of getting another large earthquake is very much higher than Poisson for short interevent times. As time passes without an earthquake, probabilities fall, until after about 100 days without a second large event, probabilities become (slightly) less than the Poisson probabilities. This reproduces what we see in the data: rather more both short and long interevent times than the Poisson model would predict.

Summary and conclusions

We investigated the occurrence of triggered events by declustering a global catalogue of $M \geq 5$ events recorded between 1964-2000. This led to the formation of two datasets: *Superclusters without aftershocks* and *Superclusters with aftershocks*. We concluded that, within the space window we defined, the distinction between 'aftershocks' and 'triggered events' was arbitrary.

The collection of events in all the Superclusters without aftershocks is equivalent to a catalogue of mainshocks. This was examined for differences in rate of earthquake occurrence before and after large earthquakes. No significant changes were found. The mainshock catalogue could be modelled well by a random, Poisson process.

We used the superclusters with aftershocks to investigate the distribution of interevent times between large earthquakes. We assumed that all earthquakes within a supercluster are sufficiently close in space for there to be a physical connection between the earthquakes. This is a crude assumption, but is justified by its leading to successful isolation of the temporal behaviour of earthquakes. Improving the spatial relationship is the most important additional work that should be undertaken, and this is discussed below.

The cumulative distribution of interevent times, Δt, between all the earthquakes in any supercluster with aftershocks with $M \geq M_r$ consists of two parts: The early part corresponds to the Omori law for aftershocks, i.e. the probability density function for Δt, $f(\Delta t) \sim 1/\Delta t$. The later part is well modeled by a Poisson process. An attempt to model the second part of the mixture as a power law $1/t^p$ resulted in the conclusion that $p = 0$, meaning that the second process was indistinguishable from a purely random, Poisson process. It follows that attempts to model the times of occurrence of 'new' earthquakes using other functions or assumptions are most unlikely to be successful. We recommend below that there should be further work on determining how such a random pattern can result from a stress accumulation process, via plate movements, that is known to be very steady on geological timescales.

Our resultant model for the time interval between successive earthquakes in a region with $M \geq M_r$ is a probability distribution that is a mixture of an Omori law, suitably modified to deal with very short and very long times, and a Poisson process. This model has four parameters: a short and a long-time constant for the Omori law, the time constant for the Poisson process and the relative weights of the two parts. The short-time constant is likely to be poorly resolved in practice because of known difficulties with data (Kagan, 2004), but the model is not sensitive to this parameter. The long time constant t_1 for the Omori law appears to be less than or equal to the time constant for the Poisson process. t_1 is poorly resolved but the model is not sensitive to its value. This leaves the two controlling parameters beings the relative weight and the Poisson time constant.

The form of the model is easy to use analytically. It allows ready calculation of the mean interevent time, and conditional probabilities for interevent times given the time to the last earthquake. The model is consistent with the Gutenberg and Richter magnitude frequency law: the mean interevent times calculated from the models for $M_r = 6.5$ and $M_r = 7$ were consistent with each other and also with the mean interevent times calculated directly from the data. The derivative results will have direct applications in earthquake hazard and risk assessment when region-specific models are developed.

We fitted the model by the method of maximum likelihood to the superclusters with aftershocks from $M_r = 6.5$ for 1985-2000 and $M_r = 7$, for 1985-2000 and 1977-2000. We concluded that:

- The weights for the two parts of the mixture are of the same order, meaning that any earthquake in a supercluster has a probability of about 0.3 - 0.5 of being an aftershock.
- The weight and time constant parameters have appreciably large 95% confidence regions as inferred from log-likelihood values. This implies that it will be difficult to obtain a good model for any particular region with a limited record of large earthquakes. However, if M_r is reduced to produce more data, the assumptions about physical connection between events is likely to be even less good an approximation. Hence the development of a proper spatial component for the model is necessary before useful models can be developed.
- To model the recurrence of large earthquakes in a region, it is neither necessary nor desirable to decluster the catalogue. Declustering at long time intervals after an event is likely to remove 'new' earthquakes as well as aftershocks, and the proposed mixture model renders declustering unnecessary anyway.

An attempt was made to apply the model to the occurrence of $M \geq 7$ earthquakes in central New Zealand since 1840. With only 14 interevent times, the model parameters had very large 95% confidence regions. The weight for the Omori part of the distribution was low compared to the global models: 0.1-0.15 rather than 0.4-0.5. We speculate that the New Zealand catalogue of $M \geq 7$ earthquakes may be missing $M \geq 7$ aftershocks that occurred shortly after some of the larger, pre-1900 events. Absence of any such data significantly biases the weight parameter and leaves the model being dubiously useful. As discussed below, a useful New Zealand model will require the spatial model improvement and fitting to a numerically larger dataset.

The global database also offered an opportunity to investigate the phenomenon reported by Evison and Rhoades (2004), which they term Precursory Scale Increase (of earthquake activity before a large earthquake) - PSI. A preliminary investigation of some of their examples of PSI suggests that the phenomenon should be systematically investigated using the global catalogue. The resources of the current project were insufficient to do this, but we recommend, below, that further research be done on this topic.

Recommendations for additional research

The results and conclusions described above have alluded to many issues that could not be resolved within the context of the current project. The proposed model for earthquake interevent times offers important possibilities for earthquake hazard and risk assessment. However, to implement and utilise the model the issues raise will have to be addressed. The steps required are as follows.

1. Development of a spatio-temporal model, as noted immediately above.

The most promising starting point for this is a kernel method e.g. Stock and Smith 2002a, b, Console et al. 2002. In Console et al.'s 2002 approach there are in fact two kernels: one for the spatial influence of the Poisson part of the model – the w_2 term in eqn (23) - and a second for the aftershock/triggered event part. Console et al. use a Gaussian kernel for the Poisson part but prefer a power law function for the aftershocks to model Italian seismicity. The temporal distribution is treated independently in the sense that the spatial and temporal parts are multiplied together to give the combined model. Both spatial and temporal components require scaling with earthquake size. Christophersen 2000 has established the appropriate magnitude scaling for both time and space. The result will be similar to an ETAS epidemic seismicity model, except that it will have good distributional properties that enable its ready application to hazard assessment.

The structure of this model should be developed using the global database, because substantial amounts of data are required to keep parameter uncertainties to a minimum during the development phase. Once the appropriate, scaled, spatio-temporal form is established, it can then be applied to produce regional models.

2. Determine a model suitable for New Zealand

This will require a suitable dataset or datasets e.g. $M \geq 6$ earthquakes since 1940. Candidate datasets are readily available from GeoNet. Once developed, the model can be tested as an alternative to the current (Poisson only) model in seismic hazard and risk assessments. Note that the long-term average hazard and risk are not likely to change as a result of such a model. What will change are the conditional assessments of hazard and risk given the occurrence of earthquakes. For example, the model probability of multiple events in a short time interval (few years) will be distinctly greater than the current Poisson probability.

3. Systematically investigate the occurrence of the PSI phenomenon

Our very preliminary consideration of a few examples cited by Rhoades and Evison, using the global database, was inadequate to confirm the PSI phenomenon or otherwise. However, it would not be difficult to make a systematic search through the database to see what were the characteristics of the seismicity prior to large earthquakes. Note that our 'global' model for interevent times is not incompatible with the possibility that precursors to earthquakes, and PSI in particular, exist. We have demonstrated here, however, that there is no large, global increase or decrease in the seismicity prior to, or following a large earthquake, other than aftershocks (and possibly immediate foreshocks) as already noted.

In such an investigation, we would test the proposed PSI against an alternative model, namely that when there is an increase in seismicity the expected time until an earthquake of a specified magnitude (e.g. M 7) will vary with the extent of the seismicty increase consistent with the Gutenberg and Richter magnitude-frequency law.

Finally, there are implications from our model for the physical processes that give rise both to aftershocks/triggered events and non-aftershocks.

4. Investigations of the implications of the functional form of the model.

It is remarkable that, although the Omori law is 110 years old, no-one has proposed a physical model that yields 1/t relaxation of stress, and hence a similar decay of aftershock numbers. Current thinking is that following an earthquake the ductile layers beneath the elastic-brittle seismogenic crust relax in response to the stess change caused by the earthquake. This relaxation in turn induces stresses in the brittle crust, causing aftershocks.

However, current rheological models do not predict 1/t behaviour. If the ductile lower crust and upper mantle behaved as a Newtonian fluid, with stress proportional to rate of strain, the decay would be exponential. It is generally considered that these rocks are not Newtonian, and power law rheologies, where stress is proportional to (rate of strain) ^{n} are preferred. However, these models do not lead to 1/t decay either.

Computer codes now exist to model realistic rheologies. We would use these programs to test the hypothesis that the 1/t behaviour is the integration of a complex, multiple stress feedback loop between the brittle and ductile layers of the crust.

The other issue to be investigated is the random behaviour of the non-aftershocks. In plate boundary regions, it is known that plate boundary processes cause stress to accumulate at a steady rate. How is it then, that 'new' earthquakes should be Poisson (or indistinguishable from Poisson)? In an earlier study Smith and Smith (1994) found that after a short period (~ 5000 years) the state of stress at any location in New Zealand became chaotic as a result of steady stress increase and stress decrease resulting from Gutenberg & Richter-distributed earthquakes. Their model allowed for no stress relaxation however. We would propose to develop a model that permitted interevent relaxation on a New Zealand scale to test whether the stress state at future epochs was likely to be chaotic.

Acknowledgements

This study was funded in part by a grant from the EQC Research Foundation, Project 03/485. We thank Martha Savage for thoughtful and useful comments on the Report.

References

Abramowitz, M., and I A Stegun (eds) (1965) Handbook of Mathematical Functions. *Dover Publications, New York.* 1046pp

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.

Anderson, J.G., J.N. Brune, J.N. Louie, Y. Zeng, M. Savage, G. Yu, Q. Chen, and D. dePolo. (1994) Seismicity in the western Great Basin apparently triggered by the Landers, California, earthquake, 28 June 1992. *Bull. Seism. Sic. Am.* **84** (Landers special issue): 863-91.

Aydin, A. and D Kalafat (2002) Surface ruptures of the 17 August and 12 November Izmet and Duzce earthquakes in northwestern Anatolia, Turkey: their tectonic and kinematic significance and associated damage. *Bull. Seism. Soc. Am* **92** (Izmit Special Issue): 95-106.

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 (2000) A global model for aftershock behaviour. *Proceedings, 12th World Conf. Earthq. Engg.* Auckland, 2000. Paper 0379, 8pp.

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

Davis, S. D., and C. Frohlich (1991) Single-link cluster analysis, synthetic earthquake catalogues, and aftershock identification, *Geophys. J. Int.* **104**, 289-306.

Dempster A. P., N. M. Laird, and D. B. Rubin, 1977, Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* **B 39**, 1-39.

Eberhart-Philips, D. (1999) Aftershock sequence parameters in New Zealand, *Bull. Seism. Soc. Am.* **88**, 1095-1097.

Evison, F. F., and D. A. Rhoades (2004) Demarcation and scaling of long-term seismogenesis, *Pure appl. geophys.* **161**, 10.1007/s00024-003-2435-8.

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. Knophoff (1974) Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, *Bull. Seism. Soc. Am.* **64**, 1363-1367.

Grapes, R. H., and G. Downes. (1997). The 1855 Wairarapa, New Zealand, earthquake - analysis of historical data. *Bull. NZ Soc. Earthq. Engg. (Spec. Issue)* **30**, 271-369.

Grapes, R., G. Downes, and A. Goh. (2004). The 1848 Marlborough, New Zealand, earthquakes – historical data. *Institute of Geological and Nuclear Sciences science report.*

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

International Seismological Centre (2002). ISC Catalogue 1904 – 2000, *cd rom* **11**.

Kagan, Y. Y. (1999) Universality of the seismic moment-frequency relation, *Pure appl. Geophys.* **155**, 537-573.

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

Kagan, Y. Y. and D. D. Jackson (1999) Worldwide doublets of large shallow earthquakes, *Bull. Seism. Soc. Am.* **89**, 1147-1155.

King, G. C. P., R. S. Stein, and J. Lin (1994) Static stress changes and the triggering of earthquakes, *Bull. Seismol. Soc. Amer.* **84**, 935-953.

Molchan, G. M., and O. E. Dmitrieva (1992) Aftershock identification: methods and new approaches, *Geophys. J. Int.* **109**, 501-516.

Musson, R.M.W., T Sapanos, and C.T Nakas (2002). A power-law function for earthquake interarrival time and magnitude. *Bull. Seism. Soc. Am.* **92**: 1783-94.

Ogata, Y. (1988) Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Ass.* **83,** 9-27.

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

Parsons, T. (2002) Global Omori law decay of triggered earthquakes: Large aftershocks outside the classical aftershock zone, *J. Geophys. Res.* **107**, 10.1029/2001JB001051.

Reasenberg, P. (1985) Second-order moment of central California seismicity, 1969–1982, *J. Geophys. Res.* **90**, 5479–5495.

Rhoades, D. A., and F. F. Evison (2004) Long-range earthquake forecasting with every earthquake a precursor according to scale, *Pure appl. geophys.* **161**, 10.1007/s00024-003-2435-9.

Savage, M. K., and D. M. DePolo (1993) Foreshock probabilities in the western great-basin Eastern Sierra Nevada, *Bull. Seism. Soc. Am* **83**, 1910-1938.

Scholz, C.H. (1990) The mechanics of earthquakes and faulting, *Cambridge University Press.*

Severini, T.A. (2000) Likelihood methods in statistics. *Oxford Statistical Sciences Series* **22**. OUP 380pp.

Smith, E and W.D Smith. (1994) Frequency of occurrence of catastrophic earthquakes. Report to the NZ earthquake Commission. *Inst. Geol. Nuc. Sci. Client Report* 32390112B 15pp + figs

Smith, E.G.C. (1994) Temporal pattern of occurrence of New Zealand's largest historical earthquakes. Report to the NZ Earthquake Commission. *IGNS Client Report 322907A*. 6pp.

Sornette, D. and L. Knopoff (1997) The paradox of the expected time until the next earthquake. *Bull. Seism. Soc. Am.* **87:** 789-798.

Stein R. S. (1999) The role of stress transfer in earthquake occurrence, *Nature* **402**, 605-609.

Stein, R. S., A. A. Barka, and J. H. Dieterich (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering, *Geophys. J. Int* **128**, 594-604.

Stirling, M., G. McVerry, K. Berryman, P. McGinty. P. Villamor, R. Van Dissen, D. Dowrick, J. Cousins, and R. Sutherland (2002) Probabilistic seismic hazard assessment of New Zealand: New active fault data attenuation relationships and methods. Report to the NZ Earthquake Commission. *Client report 2000/53 Institue of Geological & Nuclear Sciences Ltd.*

Stock, C. and E.G.C. Smith (2002a) Adaptive kernel estimation and continuous probability representation of historical earthquake catalogues. *Bull. Seism. Soc. Am.* **92**, 904-912.

Stock, C. and E.G.C. Smith (2002b) Comparison between seismicity models generated by different kernel estimations. *Bull. Seism. Soc. Am.* **92**, 913-922.

Utsu, T. (1969) Aftershocks and earthquake statistics (I) – Some parameters which characterize an aftershock sequence and their interrelations, *J. Fac. Sci. Hokkaido Univ. Jap. Ser. VIII* **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.

Vosin, C., F. Cotton, and S. Di Carli (2004) A unified model for dynamic and static stress triggering of aftershocks, antishocks, remote seismicity, creep events, and multisegmented rupture, *J. Geophys. Res.* **109**, B06304, doi:10.1029/2003JB002886.

Yeats, R. S., K. Sieh, and C. R. Allen (1997) The geology of earthquakes, *Oxford University Press*, Glossary, p. 499.