The Canterbury sequence in the context of global earthquake statistics

A. Christophersen
S. Hainzl
M. C. Gerstenberger
D. A. Rhoades
E. G. C. Smith

GNS Science Consultancy Report 2013/196
December 2013

DISCLAIMER

This report has been prepared by the Institute of Geological and Nuclear Sciences Limited (GNS Science) exclusively for and under contract to New Zealand Earthquake Commission (EQC). Unless otherwise agreed in writing by GNS Science, GNS Science accepts no responsibility for any use of, or reliance on any contents of this Report by any person other than EQC and shall not be liable to any person other than EQC, on any ground, for any loss, damage or expense arising from such use or reliance.

The data presented in this Report are available to GNS Science for other use from October 2013.

BIBLIOGRAPHIC REFERENCE

Christophersen, A.; Rhoades D. A.; Hainzl, S.; Smith, E. G. C; Gerstenberger, M. C. 2013. The Canterbury sequence in the context of global earthquake statistics, *GNS Science Consultancy Report* 2013/196. 28p.

CONTENTS

NON-TECHNICAL ABSTRACT ..... **III**

TECHNICAL ABSTRACT ..... **IV**

1.0 INTRODUCTION **1**

2.0 AFTERSHOCK BEHAVIOUR AND MODELLING **4**

2.1 The Omori-Utsu law for aftershock decay in time 4
2.2 The Gutenberg-Richter magnitude-frequency relation 4
2.3 Båth's law for the magnitude difference between mainshock and largest aftershock 5
2.4 The New Zealand generic aftershock model 5
2.5 The STEP model 5
2.6 The Epidemic Type Aftershock Sequence model 6

3.0 STATISTICS OF THE CANTERBURY EARTHQUAKE SEQUENCE **7**

3.1 The GeoNet earthquake catalogue 7
3.2 Effects of data incompleteness on aftershock parameters 9
3.3 Modelling the Canterbury sequence with the ETAS Model 11
3.4 How many more M≥6.0 aftershocks to expect? 15

4.0 WHAT WE CAN LEARN FROM GLOBAL EARTHQUAKE CATALOGUES **16**

4.1 The earthquake catalogues 17
4.2 Defining earthquake sequences 18
4.3 Results 19

5.0 CONCLUSIONS AND OUTLOOK **25**

6.0 ACKNOWLEDMENTS **27**

7.0 REFERENCES **27**

FIGURES

Figure 1.1 Map showing the evolution of the Canterbury sequence with earthquakes colour coded according to the occurrence in time; and fault ruptures indicated 1

Figure 1.2 The cumulative number of a) M≥4.0 and b) M≥5.0 aftershocks with time 2

Figure 3.1 The magnitude of earthquakes in the GeoNet catalogue downloaded at different times versus time in days relative to the Darfield mainshock 8

Figure 3.2 The magnitude of earthquakes in the GeoNet catalogue downloaded at different times versus time in days relative to the Darfield mainshock for different time intervals following the major events: a) 50 hours following Darfield; b) a few days around the February Christchurch earthquake; c) a few days in June 2011; and d) a few days in December 2011. 8

Figure 3.3 The maximum likelihood fit of the Omori-Utsu law for the first 100 days following the Darfield, Christchurch and June 2011 earthquakes for the two GeoNet catalogues. 10

Figure 3.4 Maximum likelihood estimate of the *b*-value as a function of time following the Darfield, the Christchurch, and the June 2011 earthquakes for the two GeoNet catalogues. Each value refers to the 200 successive *M*≥3.0 events that occurred after the indicated time. 11

Figure 3.5 ETAS-model rate for the estimated parameters set in comparison with observed $M \geq 3.0$ earthquakes....12
Figure 3.6 Fit of the ETAS model based on generic and estimated parameters for the $M \geq 3$ sequence....12
Figure 3.7 Retrospective ETAS-model forecasts of the number of $M \geq 6$ events following two years after the Darfield earthquake, i.e. covering the period of the catalogue....13
Figure 3.8 Prospective ETAS-model forecasts of the number of $M \geq 6$ events following 50 years after the end of the catalogue....14
Figure 4.1 A magnitude-time plot comparing the USGS Centennial catalogue with the ISC-GEM catalogue....18
Figure 4.2 Histogram of the number of large aftershocks per main shock for the different search criteria in the ISC-GEM catalogue....21
Figure 4.3 Histogram of the number of large aftershocks per main shock for the different search criteria in the Centennial catalogue....21
Figure 4.4 The distribution of relative time of occurrence of large aftershocks from the main shock for four different search criteria for the ISC-GEM catalogue....22
Figure 4.5 The distribution of relative time of occurrence of large aftershocks from the main shock for four different search criteria for the Centennial catalogue....22
Figure 4.6 The distribution of relative distance of large aftershocks from the main shock for four different search criteria for the ISC-GEM catalogue....23
Figure 4.7 The distribution of relative distance of large aftershocks from the main shock for four different search criteria for the Centennial catalogue....23

TABLES

Table 3.1 The number of aftershocks in the first 24h following the four main events of the Canterbury sequence in three magnitude bands: 3: $2.95 \leq M < 3.95$; 4: $3.95 \leq M < 4.95$; and 5: $M \geq 4.95$ for the GeoNet catalogues downloaded in January and September 2012....9
Table 3.2 The parameters of the Omori-Utsu law (Equation 1) for two different GeoNet catalogues and the three largest earthquakes in the Canterbury sequence....10
Table 3.3 The ETAS model parameters used in the simulations and for forecasting....12
Table 3.4 Results from the retrospective ETAS simulations....13
Table 3.5 Results from the prospective ETAS simulations....14
Table 4.1 Time period and available number of shallow earthquakes for subsets of the Centennial and ISC-GEM catalogue that are assumed to be complete....17
Table 4.2 The search radius $r(M)$ in km for different main shock magnitude M, according to scaling relations by Wells & Coppersmith (1994)....19
Table 4.3 The number of main shocks and total number of large aftershocks for the different clustering parameters where 1R means the aftershock occurred within one radius of the main shock search radius according to Equation 10 and Table 4.2, and 2R is double the radius....19
Table 4.4 The percentage sequences that have three or more large aftershocks for the larger subset of each catalogue....20
Table 4.5 The percentage of large aftershocks at a smaller relative distance D from the main shock, and an earlier time of occurrence than the Christchurch earthquake to the Darfield earthquake, for different search criteria and the four different data subsets....24

NON-TECHNICAL ABSTRACT

The Canterbury earthquake sequence is an on-going earthquake sequence, which started with the M7.1 Darfield earthquake on 4 September 2010. The three M≥6.0 aftershocks included the devastating M6.3 Christchurch earthquake on 22 February 2011, which resulted in 185 deaths and extensive damage. As a consequence of the earthquake sequence the New Zealand National Seismic Hazard model was expected to grossly underestimate the level of ground shaking for the Canterbury region for the coming decades due to on-going aftershock activity and the possibility of other triggered moderate to large earthquakes. Therefore a new time varying earthquake hazard model, the EE model, was developed.

Our project set out to investigate the Canterbury sequence in the context of global earthquake statistics. We had two main objectives:

1. Use aftershock parameters derived from global earthquake catalogues to provide bounds for earthquake rates in the Canterbury region for the coming years and decades.
2. Investigate how unusual late and large aftershocks are as observed in the Canterbury sequence.

During our project we realised that many earthquakes, including 14 in the magnitude band 5-5.9 in the first 24h following Darfield, had not been reported in the initial GeoNet catalogue. We slightly changed focus from deriving aftershock parameters from the global catalogue, to modelling aftershocks in the Canterbury sequence to provide bounds for earthquake rates in the Canterbury region for the coming years and decades. The probability for one or more M≥ 6.0 earthquakes in the 50 years starting in September 2012 ranges from 50% to 93% depending on the model. We identified further work required to understand the time frame for which aftershock models can be usefully applied. We also identified further work required to better understand what effect the deficiencies of the real-time earthquake data have on the real-time forecasting models.

To investigate how unusual late and large aftershocks are (such as those observed in the Canterbury sequence) we used two global earthquake catalogues. For each catalogue we identified magnitude thresholds and time periods for which the data can be expected to be complete. We then searched for earthquake sequences using different search criteria. Depending on the search criteria, between 6 and 13 % of main shocks in the more complete catalogue had 3 or more aftershocks within 1.1 magnitude units as observed in the Canterbury sequence. Of these large earthquakes, around 20% had a larger relative distance than Christchurch to Darfield, and around 15% occurred later than 171 days. This confirmed that the M≥6.0 aftershocks in the Canterbury sequence occurred relatively far away from the main shock in space and time but this is not too unusual in global aftershock statistics.

We also searched for large aftershocks within 1.1 magnitude units of the main shock that occurred at least as late and far away as the Christchurch earthquake relative to the Darfield earthquake. In our most complete dataset and with the preferred search criteria, 2% of large earthquake fulfil these criteria. In the other dataset there were another couple of examples. This confirms that while the Christchurch earthquake occurred with a time delay of 171 days and 42 km from the Darfield epicentre, this is not too unusual in global earthquake occurrences.

TECHNICAL ABSTRACT

The Canterbury earthquake sequence has been a long-lived complex earthquake sequence, which started with the M7.1 Darfield earthquake on 4 September 2010. The three M≥6.0 aftershocks included the devastating M6.3 Christchurch earthquake on 22 February 2011, which resulted in 185 deaths and extensive damage. Due to on-going aftershock activity and the possibility of other triggered moderate to large earthquakes, the New Zealand National Seismic Hazard model was expected to grossly underestimate the level of ground shaking for the Canterbury region for the coming decades. Therefore a new time varying earthquake hazard model, called the EE model, was developed.

Our project set out to investigate the Canterbury sequence in the context of global earthquake statistics. We had two main objectives:

1. Use aftershock parameters derived from global earthquake catalogues to provide bounds for earthquake rates in the Canterbury region for the coming years and decades.
2. Investigate how unusual late and large aftershocks are, as observed in the Canterbury sequence.

During our project we realised that many earthquakes in the first 24h following Darfield, including 14 in the magnitude band 5-5.9, had not been reported in the initial GeoNet catalogue. We slightly changed focus from deriving aftershock parameters from the global catalogue, to aftershock modelling in the Canterbury sequence to provide bounds for earthquake rates in the Canterbury region for the coming years and decades. We first illustrated the deficiency of the initial catalogue by comparing the GeoNet data downloaded in January 2012 with the mainly finalised catalogue in September 2012. We identified further work required to better understand what effect the deficiencies of the real-time earthquake data have on the real-time forecasting models. We fitted the two relationships used in aftershock forecasting, the Omori-Utsu law for aftershock decay and the Gutenberg-Richter relation for the magnitude-frequency distribution of earthquakes, to three subsets of the catalogue following M≥6.0 earthquakes. The range of the fitted parameters exceeded what we would have expected from averaging global earthquake sequences. We used the fitted parameters to calculate the number of M≥6.0 earthquakes expected to occur in the 50 years starting in September 2012 for the different combinations of parameters. The expected number ranged from 0.6 to 1.1, and totalled 2.3 for all sub sequences, which is an upper bound estimate due to the dependence of the events. In comparison, the EE model expects about 1.3 M≥6.0 earthquakes in about the same area and time.

We also fitted the Epidemic Type Aftershock Sequence (ETAS) model to the Canterbury sequence. We used the estimated parameters, as well as generic aftershock parameters for New Zealand to simulate ETAS earthquake sequences for the two-year duration of the catalogue, as well as for 50 years starting in September 2012. The retrospective simulations are consistent with the observations. For the prospective simulations the expected number of M≥ 6.0 earthquakes in 50 years is up to six times higher than for the EE model. We identified further work required to investigate the applicability of aftershock models on the time-scale of decades, which is due to begin soon.

In summary, the probability for one or more M≥ 6.0 earthquakes in the 50 years starting in September 2012 ranges from 50% to 93% depending on the model and model parameters.

To investigate how unusual late and large aftershocks are as observed in the Canterbury sequence, we used two global earthquake catalogues. For each catalogue, we identified magnitude thresholds and time periods for which the data can be expected to be complete. We then searched for earthquake sequences using different search criteria. Depending on the search criteria, between 6 and 13 % of main shocks in the more complete catalogue had 3 or more aftershocks within 1.1 magnitude units, as observed in the Canterbury sequence. In space, the relative distance of Christchurch to the Darfield earthquakes was around the 80^{th} percentile, and in time around the 85^{th} percentile of all large aftershocks respective to their main shocks for the preferred search criteria. This confirmed that although the M≥6.0 aftershocks in the Canterbury sequence occurred relatively far away from the main shock in space and time these differences were not unusual in global aftershock statistics. We also searched for large aftershocks within 1.1 magnitude units of the main shock that occurred at least as late and far away as the Christchurch earthquake relative to the Darfield earthquake. In our best dataset, 2% of large earthquake fulfil these criteria. In the other dataset there were another couple of examples. This confirms that while the Christchurch earthquake occurred with a time delay of 171 days and 42 km from the Darfield epicentre, this is not unique in the context of global earthquake occurrences.

1.0 INTRODUCTION

The Canterbury earthquake sequence has been a long-lived complex aftershock sequence, involving a number of different faults. It started with the M7.1 Darfield earthquake on 4 September 2010, which was followed by a series of aftershocks into early 2011. On 22 February 2011, the devastating M6.3 Christchurch earthquake occurred, resulting in 185 deaths and extensive damage. Further significant earthquakes occurred on 13 June and 23 December 2011. Figure 1.1 shows a map of the Canterbury earthquake sequence indicating the Greendale fault rupture as well as the sub-surface fault ruptures of the Darfield earthquake that involved several fault planes (Beavan, et al., 2012), and the sub-surface fault ruptures of the subsequent major earthquakes. The earthquake activity moved eastward from the Canterbury Plains, through the city of Christchurch and eventually offshore. The epicentre of the Darfield earthquake is 42 km from the epicentre of the Christchurch earthquake while the subsequent major earthquakes were more closely spaced.

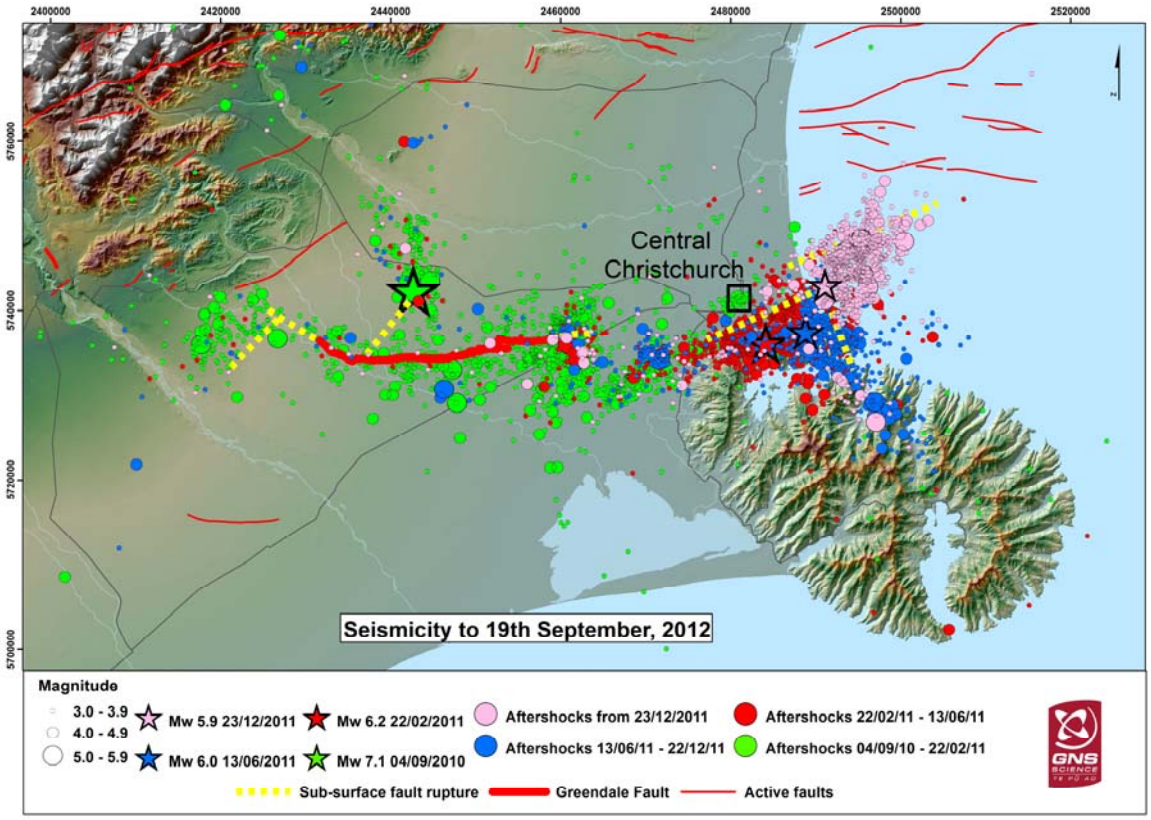

Figure 1.1 Map showing the evolution of the Canterbury sequence with earthquakes colour coded according to the occurrence in time; and fault ruptures indicated (Figure prepared by Rob Langridge).

Figure 1.2 shows the cumulative number of earthquakes as a function of time for two magnitude cut-offs and compares the data with the New Zealand generic aftershock model (Pollock, 2007) applied to a M7.1 main shock (blue line), as well as for each M≥6.0 earthquake (black line). Both plots show that the larger earthquakes were followed by their own bursts of aftershock activity. The generic model provides a reference and allows comparison to the average New Zealand aftershock behaviour. For example, the Darfield earthquake had fewer than average M≥4 aftershocks; the Christchurch earthquake had about twice as many immediate aftershocks as the New Zealand generic aftershock model (Pollock, 2007); the June 2011 earthquake had an average number of aftershocks, and

December 2011 earthquake had several times more aftershocks than the New Zealand generic aftershock model. The catalogue data stop in September 2012 when GeoNet changed to a new data acquisition system. The general rule of thumb says that the largest aftershock is about one magnitude unit smaller than the main shock (Bath, 1965). The Canterbury sequence has three aftershocks within 1.1 magnitude units of the main shock. Also, the large aftershocks occurred over a period of 15 months, and were located between 42 and 47km away from the initial epicentre. This raises the question of how usual or unusual the Canterbury sequence is. There are only limited numbers of large aftershock sequences in the New Zealand earthquake catalogue; we therefore proposed to use global earthquake catalogues to investigate this question.

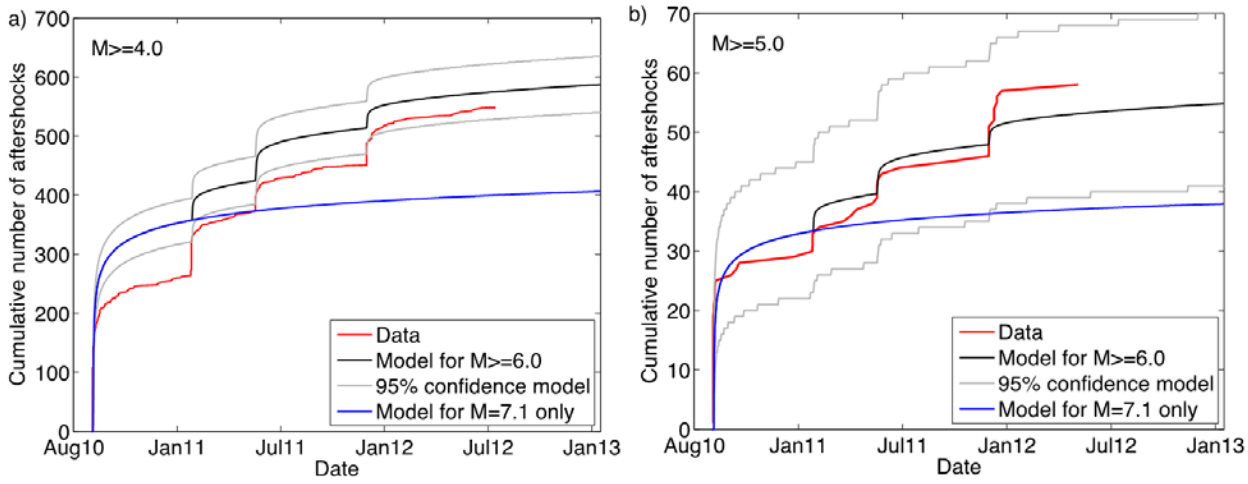

Figure 1.2 The cumulative number of a) M≥4.0 and b) M≥5.0 aftershocks with time. The blue line shows the expected number of aftershocks for a M7.1 main shock for the New Zealand generic aftershock model (Pollock, 2007); the black line shows the generic aftershock model applied to each earthquake of magnitude 6 and larger, and the grey lines represent the 95% confidence bounds of this model.

Prior to the Darfield earthquake, the Canterbury Plains region was seismically quiet when compared to most other regions of New Zealand. Only 10 M≥4 earthquakes occurred in the area shown in Figure 1.1 in the 50 years prior to the Darfield earthquake, with none greater than M5.0. Due to on-going aftershock activity and the possibility of other triggered moderate to large earthquakes, the National Seismic Hazard Model (NSHM) (Stirling et al., 2012) was expected to grossly underestimate the level of ground shaking for the Canterbury region for the coming decades. As a consequence, new time-dependent seismicity models for the Christchurch region have been developed (Gerstenberger, et al., 2013). A number of short-term, medium-term and long-term forecasting models have been combined to provide an earthquake forecast on a 0.05 * 0.05 degree grid in 0.1 magnitude bins for the next 50 years. An international panel of experts provided feedback on how to combine the models. The expert elicitation ('EE') model, is now the basis for seismic hazard calculations in the Canterbury region. The model covers a much wider area than shown in Figure 1, i.e. a 3 by 4 degree box from latitude 45°S to 42°S and longitude 170°E to 174°E compared to the smaller area of about 0.6 by 1.5 degrees. The EE model forecasts 79 M≥5 and 15 M≥6 earthquakes in the larger grid, and 15 M≥5 and 1.3 M≥6 earthquakes in the smaller grid in the 50 years starting in July 2012.

We were interested to learn from global earthquake catalogues about the duration of aftershock sequences and whether using their aftershock parameters would confirm the rates derived in the EE model. Thus we set out to investigate the Canterbury sequence in the context of global earthquake statistics with two main objectives:

1. Use global earthquake statistics to provide bounds for earthquake rates in the Canterbury region for the coming years and decades, and
2. Investigate how unusual late and large aftershocks are as observed in the Canterbury sequence.

We originally proposed to systematically define earthquake clusters in global earthquake catalogues, fit the parameters of relationships for aftershock behaviour and calculate long-term earthquake rates from parameter variation observed in the global data to put boundaries on future aftershock occurrence in Canterbury.

We also fitted the parameters of relationships for aftershock behaviour to the Canterbury sequence: We found that many earthquakes had been missed in the first 24h after Darfield, and also after the other large earthquakes in the sequence, and had been added to the GeoNet data up to 18 months after their occurrence. We also found that, during the sequence, aftershock parameters varied more than we expected from averaged global parameters (Christophersen, 2000). Thus we slightly shifted focus and put more emphasis on the statistical analysis of the Canterbury sequence.

In this report, we first provide an overview of general aftershock behaviour and introduce the relationships involved in statistical modelling of aftershocks. We further introduce the Short-Term Earthquake Probability (STEP) and Epidemic Type Aftershock Sequence (ETAS) models. Section 3 shows results of aftershock modelling for the Canterbury sequence, starting with a comparison of the GeoNet catalogue from January 2012 with the GeoNet catalogue available since September 2012, by which time the processing of the Canterbury sequence was mostly finalised. We illustrate the effect of the initial data incompleteness on the common aftershock parameters. We then use the ETAS model to calculate the probability of future M≥ 6.0 earthquakes in the Canterbury sequence and compare the results with the EE model. Section 4 looks at what we can learn from global earthquake catalogue and helps us answer the question of how usual or unusual the Canterbury sequence has been. We close with conclusions and an outlook for further work.

2.0 AFTERSHOCK BEHAVIOUR AND MODELLING

In this section we briefly introduce the empirical relationships that underpin most aftershock modelling, as well as two aftershock models: The Short-Term Earthquake Probability (STEP) model (Gerstenberger, et al., 2004; Gerstenberger, et al., 2005) model and the Epidemic Type Aftershock Sequence (ETAS) model (Helmstetter & Sornette, 2002; Ogata, 1988).

2.1 THE OMORI-UTSU LAW FOR AFTERSHOCK DECAY IN TIME

The Omori law is the oldest empirical relationship in seismology and originally described how the number of felt aftershocks per day decayed with time t as $1/t$ for the 1891 Nobi, Japan earthquake (Omori, 1894). This relationship for the Nobi sequence was found to still be ongoing 100 years after the Nobi main shock (Utsu, et al., 1995), which can be partly attributed to Nobi being located in an otherwise relatively seismically quiet area in Japan. In areas with higher seismicity rate, the aftershock rate might drop below the long-term average rate faster. As the seismicity in Canterbury was fairly low before the Darfield earthquake, we can expect to observe the aftershock decay for many decades.

Equation 1 shows the modified Omori-Utsu law, i.e. the frequency of aftershocks per unit time interval $\dot{n}(t)$ at time t. The parameter p controls the temporal decay. The parameter K relates to the productivity, and while the meaning of the constant c is still being debated, its presence is necessary to avoid a singularity in the modelled number of aftershocks at time $t = 0$.

$$\dot{n}(t) = \frac{K}{(t+c)^p} \qquad \textbf{Equation 1}$$

Integrating equation 1 over the time interval $[S,T]$ results in the number N of earthquakes

$$N = \int_S^T \frac{K}{(t+c)^p}\,dt = KI_{OU}(S,T), \text{ where} \qquad \textbf{Equation 2}$$

$$I_{OU}(S,T) = \ln\left(\frac{T+c}{S+c}\right) \text{ for } p=1, \text{ and} \qquad \textbf{Equation 3}$$

$$I_{OU}(S,T) = \frac{(T+c)^{1-p} - (S+c)^{1-p}}{1-p} \text{ for } p\neq 1. \qquad \textbf{Equation 4}$$

These equations are used for the models in Figure 1.2.

2.2 THE GUTENBERG-RICHTER MAGNITUDE-FREQUENCY RELATION

The Gutenberg-Richter relation describes the magnitude-frequency distribution of earthquakes, where the number of earthquakes of magnitude M decreases exponentially with increasing magnitude $N(M) \sim 10^{-bM}$ (Gutenberg & Richter, 1944; Ishimoto & Iida, 1939). The relationship also holds for the distribution of aftershock magnitudes (Utsu, 1969). The b-value describes the relative frequency of small and large earthquakes. The b-value for a set of earthquake data, with magnitudes above M_C with sample mean magnitude M_{ave}, can be estimated by the maximum likelihood method (Aki, 1965).

$$b = \frac{1}{\ln(10)(M_{ave} - M_C)}. \qquad \textbf{Equation 5}$$

The Gutenberg-Richter relation is often used to estimate the completeness of earthquake catalogue data. Data start to fall below the straight line at the lower magnitude end at M_c.

2.3 BÅTH'S LAW FOR THE MAGNITUDE DIFFERENCE BETWEEN MAINSHOCK AND LARGEST AFTERSHOCK

Båth's law is the simple empirical observation that the average difference between the magnitudes of the main shock and its largest aftershock is about 1.2 regardless of the magnitude of the main shock (Båth, 1965).

2.4 THE NEW ZEALAND GENERIC AFTERSHOCK MODEL

The New Zealand generic aftershock model results from fitting the large New Zealand aftershock sequence to the Omori-Utsu law (Equation 1), where the productivity parameter K is replaced by

$$K = 10^{a+b(M_m - M_C)}, \qquad \textbf{Equation 6}$$

so that the productivity increases exponentially with the main shock magnitude M_m; the parameters a and b are from the Gutenberg-Richter relation (Reasenberg & Jones, 1989, 1990). The generic parameters are p=1.07, c=0.04 days, a=-1.59, and b=1.03 (Pollock, 2007).

2.5 THE STEP MODEL

The STEP (Short-Term Earthquake Probability; Gerstenberger et al., 2004, 2005) model is an aftershock model based on the idea of superimposed Omori sequences (Equation 1). The model is comprised of two components: 1) a background model; and 2) a time-dependent clustering model. The background model can consist of any model that is able to forecast a rate of events for the entire region of interest at all times. The clustering model uses three model-components. The first component is based on the average behaviour of aftershocks in New Zealand and uses the generic parameter values for New Zealand aftershocks (Pollock, 2007). The second component uses the development of the on-going aftershock sequence to attempt to improve the information in the forecast. In this component the Reasenberg and Jones parameters are estimated for each individual aftershock sequence as it develops. The third component allows for spatial heterogeneities within the sequence by calculating the Reasenberg and Jones parameters on a 0.1-degree by 0.1-degree grid within the aftershock sequence. The second and third components are only practical for large aftershock sequences with more than 100 aftershocks above the threshold of completeness; therefore, in practice, the STEP forecasts are typically dominated by the first component. In order to produce a single forecast based on the three component forecasts, each forecast is weighted by a score from the corrected Aikaike Information Criterion (Burnham & Anderson, 2002) and the three weighted forecasts are summed. The AICc is based on a model's likelihood score, the amount of data and the number of free parameters that must be estimated, thus ensuring a smooth transition from the generic component to the spatially-heterogeneous component as more data becomes available. Finally, for each grid node, all aftershock forecasts are compared to the background forecast for the node and the highest forecast is retained.

The spatial distribution of the forecast aftershocks is partially decoupled from the rate calculation. The rate forecast is isotropically distributed using a radius based on the subsurface rupture length from magnitude-scaling relationships (Wells & Coppersmith, 1994) and using a $1/r^2$ taper. If the radius is less than the grid spacing, a point source is assumed and the rates are smoothed in a circular fashion using the $1/r^2$ taper; otherwise a two segment fault is estimated based on the density of the aftershock distribution and the rates are smoothed using the $1/r^2$ taper from the fault. Additionally, because each aftershock can generate its own rate forecast as well as contributing to the overall rate forecast of the main shock, spatial heterogeneities are added by superimposing individual aftershock forecasts on the main shock forecast. When more than one forecast is calculated for a single grid node, only the largest forecast is retained.

2.6 THE EPIDEMIC TYPE AFTERSHOCK SEQUENCE MODEL

The ETAS model is a stochastic point-process model that is based on overlapping Omori sequences according to Equation 1 (Ogata, 1988). Different implementations of the ETAS model are in use. Here we use the Hainzl implementation (Hainzl, et al., 2008), where each earthquake has a magnitude M dependent ability to trigger aftershocks above magnitude M_C according to

$$k_0 10^{\alpha(M-M_C)}, \qquad \textbf{Equation 7}$$

where k_0 rather than K as in Equation 1 reflects the fact that each earthquake triggers its own aftershock. The parameter α is a constant. The total occurrence rate of earthquakes is given by

$$\lambda(t,x,y) = \mu + \sum_{i:t_i<t} \frac{k_0 10^{\alpha(M_i-M_C)}}{(t-t_i+c)^p} f_i(x-x_i, y-y_i), \qquad \textbf{Equation 8}$$

where μ is the background rate and $f(x,y)$ is a normalised function, which describes the spatial probability distribution of triggered aftershocks. One challenge in applying the ETAS model can be the estimation of the parameters (Harte, 2013).

Recently we applied the ETAS-Hainzl model to aftershock simulations for a Wellington Fault earthquake (Christophersen, et al., 2013). We set the ETAS model parameters equivalent to the generic New Zealand parameters and calculated k_0 to be equivalent to 10^a in Equation 6. In section 3.3 we use those values for ETAS simulations of the Canterbury sequence.

3.0 STATISTICS OF THE CANTERBURY EARTHQUAKE SEQUENCE

Statistical analyses of earthquake catalogues rely on having a consistent and complete set of earthquake records, i.e. on having a set of data that includes all earthquakes above a certain magnitude, the completeness magnitude M_C. We found that many aftershocks within the first 30 hours after the Darfield earthquake were only added to the GeoNet catalogue nearly two years after the event. As a further challenge to the study of the Canterbury sequence, the GeoNet data acquisition system changed in September 2012 as further explained below. A thorough study of the catalogue changes and the effect on aftershock models is beyond the scope of the current project and another proposal to EQC is underway. Here we briefly compare the catalogues downloaded in January 2012 and September 2012 for the time intervals of the major earthquakes in the Canterbury sequence to illustrate the extent of the missing data. Section 3.2 investigates the effect of the incomplete data on the key aftershock parameters. We then apply the ETAS model to the best available data to forecast the rate of large aftershocks in the next 50 years. We compare the results with the EE model.

3.1 THE GEONET EARTHQUAKE CATALOGUE

Over the last couple of years GeoNet has been in the process of changing its earthquake location system from CUSP to the new GeoNet Rapid, which uses the software SeisComP3. The bulk of the Canterbury sequence was located with the CUSP system, which was discontinued in September 2012. Currently the data prior to December 2011 have been finalised, while January – August 2012 are still under review. The long-term plan is to relocate the historic earthquake catalogue, or at least part thereof, with SeisComP3 but until then it is not clear how consistent the old and the new data are in location and magnitude. For this project we used the CUSP data only and therefore the catalogue ends in September 2012.

The many aftershocks that occur immediately following a large earthquake are difficult to detect because of many overlying waveforms. The more earthquakes occur within a short period of time, the harder it is to process them. GeoNet prioritises processing and reviewing of data according to competing workloads. In Figure 3.1 and Figure 3.2 we show the differences in the downloaded catalogues from January and September 2012. Figure 3.1 shows the complete period of the catalogue on a linear and logarithmic time scale, and the differences between the newer and later data are particularly apparent right after the Darfield earthquake on the logarithmic scale and at the end of the catalogue due to the longer duration at the end of the catalogue. Figure 3.2 zooms into a limited time period following the major events in the sequence. The first couple of days are most affected by missing data. When no major event occurs there is not much difference in magnitude and times between the earlier and later catalogue. Table 3.1 compares the aftershocks in the catalogue for the first 24 hours after each $M \geq 6.0$ earthquake in both catalogues. The later catalogue has about three times as many earthquakes in the magnitude bands M4.0-4.9 and M5.0-5.9, and more than seven times as many earthquakes in the magnitude band M3.0-3.9 for the Darfield earthquake. For Christchurch, there is little difference in the catalogues in the magnitude bands M4.0-4.9 and M5.0-5.9, but about 2.5 as many earthquakes in the magnitude band M3.0-3.9 in the later catalogue. The difference in the catalogues is less pronounced for the June and December earthquakes. It is not clear whether the data had already been partly processed at the time of downloading the catalogue. The most important realisation is that many earthquakes were not initially recorded in the GeoNet catalogue, and we will discuss this further in Section 5.

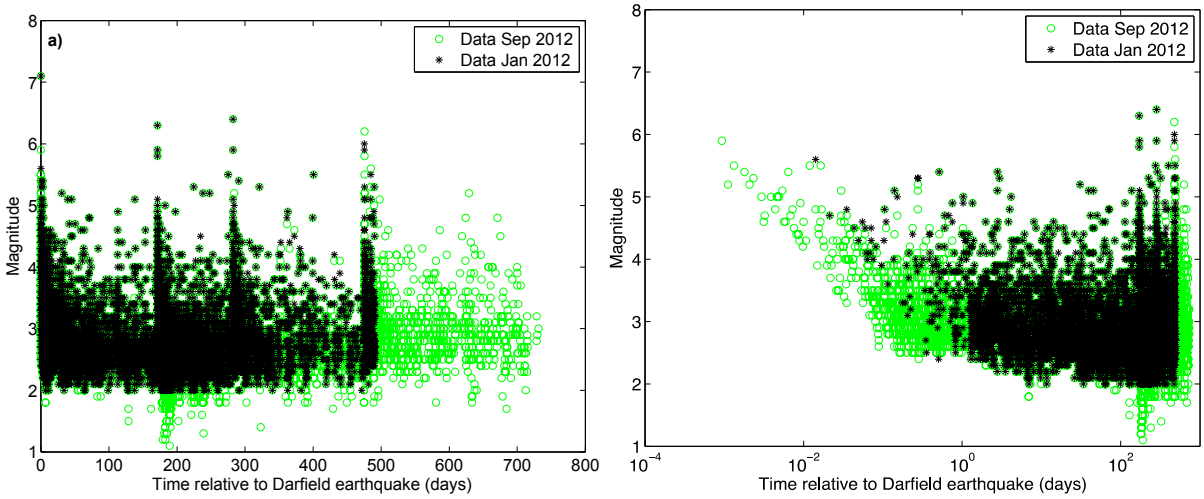

Figure 3.1 The magnitude of earthquakes in the GeoNet catalogue downloaded at different times versus time in days relative to the Darfield mainshock. The logarithmic timescale in plot b) shows how many, even large events were added to the catalogue nearly two years after the M7.1 event.

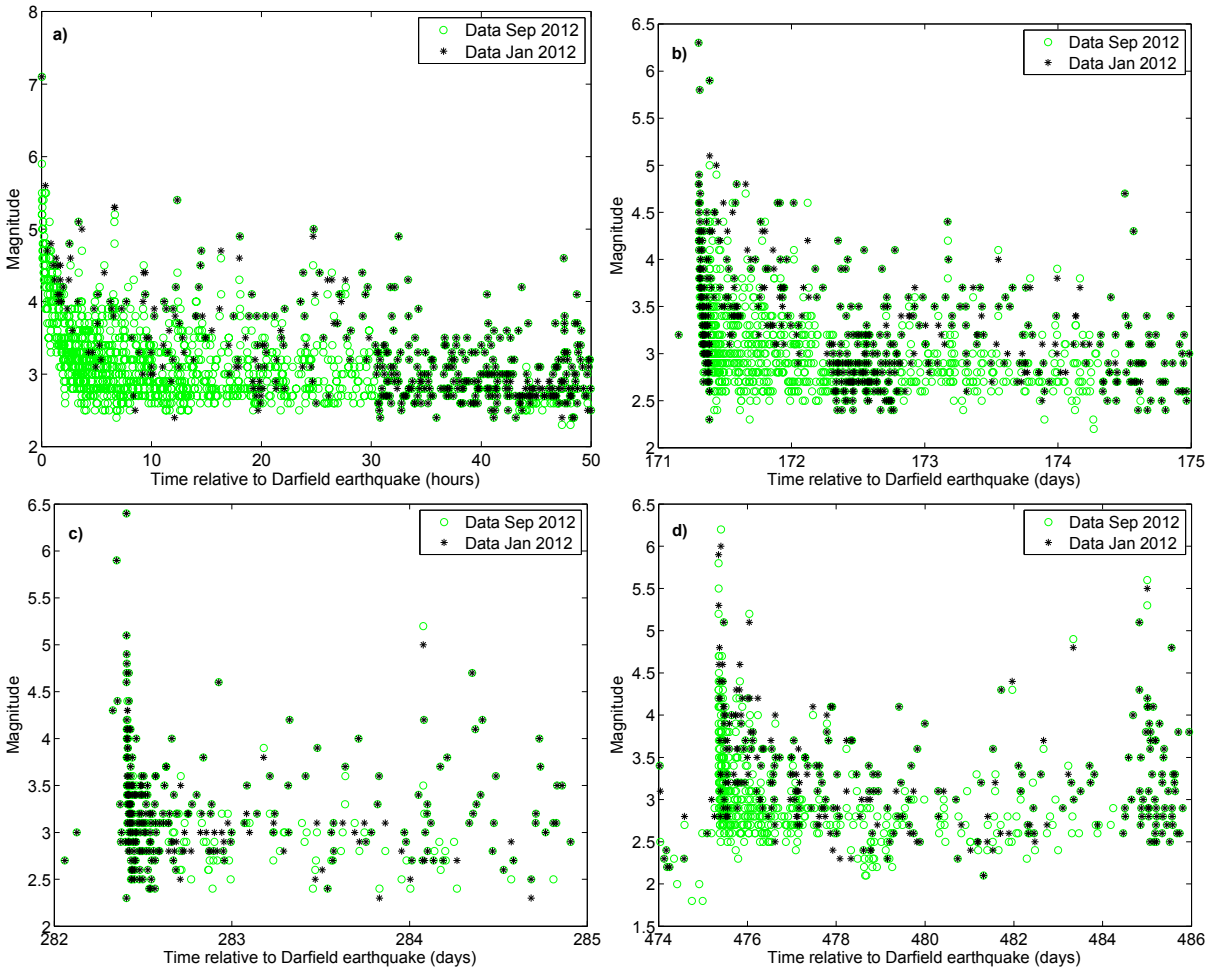

Figure 3.2 The magnitude of earthquakes in the GeoNet catalogue downloaded at different times versus time in days relative to the Darfield mainshock for different time intervals following the major events: a) 50 hours following Darfield; b) a few days around the February Christchurch earthquake; c) a few days in June 2011; and d) a few days in December 2011.

Table 3.1 The number of aftershocks in the first 24h following the four main events of the Canterbury sequence in three magnitude bands: 3: 2.95≤M<3.95; 4: 3.95≤M<4.95; and 5: M≥4.95 for the GeoNet catalogues downloaded in January and September 2012.

		GeoNet January 2012			**GeoNet September 2012**		
Major earthquake	*M*	*M*5	*M*4	*M*3	*M*5	*M*4	*M*3
Darfield	7.1	6	36	69	20	111	523
Christchurch	6.3	4	55	147	3	56	350
June 2011	6.4	1	22	130	1	22	133
December 2011	6.2	3	17	34	2	21	116

3.2 EFFECTS OF DATA INCOMPLETENESS ON AFTERSHOCK PARAMETERS

Two of the key empirical relationships in aftershock modelling are the Omori-Utsu law for the decay of aftershock activity with time and the Gutenberg-Richter relation for the magnitude-frequency distribution. Here we fit both relationships to the two different GeoNet catalogues for three *M*≥6.0 earthquakes in the Canterbury sequence. We cut the catalogue in space in about the area shown in Figure 1.1 and included all earthquakes of magnitude 3 and larger. Figure 3.3 shows the Omori-Utsu fit to the number of aftershocks per day against time on a log-log scale. Table 3.2 lists the fitted model parameters. The difference between the catalogues is most striking for the Darfield earthquakes where the *c*-value, i.e. the off-set in time before the onset of the power law decay, reduced from 2.5 days to 0.1 days for the more complete data. The three Omori-Utsu parameters are correlated and for the Darfield earthquake they all decrease with the more complete catalogue. The difference for the other two earthquakes are less obvious as can be expected from the smaller difference in the catalogues. One point to note is the significant decrease of the *p*-value with time, which leads to higher probabilities for large events at later times as further discussed below. The decrease is likely to be caused by aftershocks from the earlier larger earthquakes, which do not follow the Omori-Utsu law of the fitted later earthquakes and therefore seem to slow the decay and result in smaller p-values. This effect could be lessened when fitting the parameters of all three sequences simultaneously or when selecting earthquakes in the region around the fault zone of each earthquake separately. We also note that the higher *p*-value estimate for the incomplete January 2012 data set for Darfield is strongly biased by the strong positive correlations between *c*- and *p*-value estimations (Hohlschneider, et al., 2012) and the large period of incompleteness directly after the mainshock.

Having fitted the parameters of the Omori-Utsu law, we can use Equation 2 to calculate the expected number of earthquakes in a selected time interval. We are interested in the expected number of M≥6.0 earthquakes in the 50 years time period from September 2012. As the parameters are derived for M≥3.0 we use the magnitude-frequency relation (Section 2.2) with a *b*-value of 1 to calculate the values that are also listed in Table 3.2. The expected number of M≥6.0 earthquakes in 50 years from September 2012 ranges from 0.4 to 1.3. Using the simple Omori-Utsu decay, the Christchurch and the June earthquakes have larger expected numbers than the Darfield earthquake that is more than a magnitude unit larger and has much larger *K*-value. The explanation for this is the smaller *p*-value leads to more aftershocks occurring later in the sequence than early on. The 10% difference in expected number for the June 2011 earthquake from the two different catalogues lies in a 0.01 difference in the *p*-value. If we make the simple assumption that the three fitted M≥6.0 earthquakes have their own aftershock sequence with the parameters as estimated, then we

can add the expected numbers in Table 3.2. For the better catalogue, we get 2.3 M≥6.0 earthquakes in the 50 year period starting in September 2012.

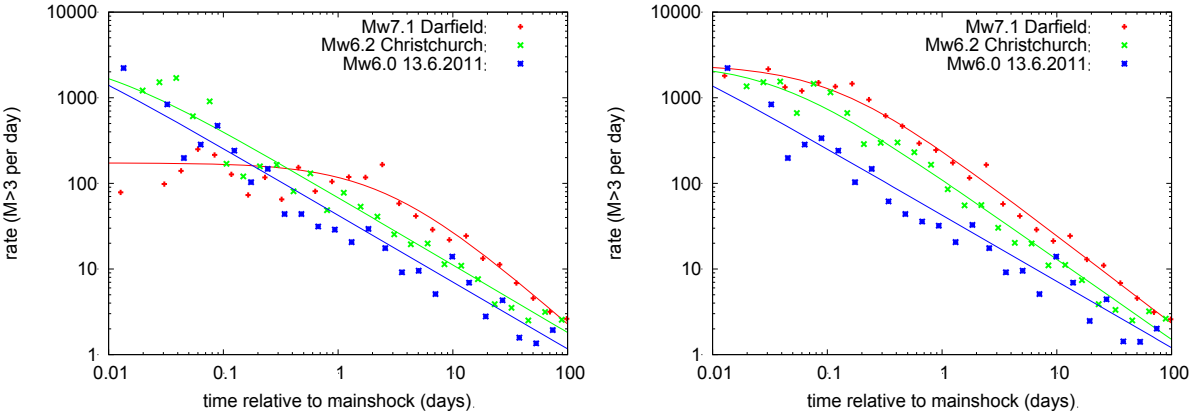

Figure 3.3 The maximum likelihood fit of the Omori-Utsu law for the first 100 days following the Darfield (red), Christchurch (green) and June 2011 earthquakes for the two GeoNet catalogues (January 2012 on left and September 2012 on right). The parameters of the fit are shown in Table 3.2.

Table 3.2 The parameters of the Omori-Utsu law (Equation 1) for two different GeoNet catalogues and the three largest earthquakes in the Canterbury sequence. The last columns shows the expected number of M≥6.0 earthquakes in the 50 years starting in September 2012 for each parameter combination.

	Catalogue	***K*-value**	***c*-value [days]**	***p*-value**	***50 year M≥6.0 rate from Sept. 12***
Darfield	Jan. 2012	511.2	2.514	1.17	0.41
	Sept. 2012	261.4	0.112	1.03	0.55
Christchurch	Jan. 2012	68.3	0.007	0.79	1.34
	Sept. 2012	112.7	0.035	0.94	0.64
June	Jan. 2012	42.5	0.001	0.78	1.01
	Sept. 2012	42.6	0.001	0.77	1.10

Figure 3.4 shows the *b*-value of the Gutenberg-Richter relationship as a function of time relative to the main shock on a logarithmic scale. Again, the differences are most striking for the Darfield earthquake. The June earthquake has the most stable *b*-value with time and the least difference between the catalogues. For later time periods the changes in *b*-value with time are repeated between the two catalogues. We have not tried to match the increase in b-value at 1 day and 10 day after the Darfield earthquake, and the generally higher b-value following the Christchurch earthquake with any other physical observations.

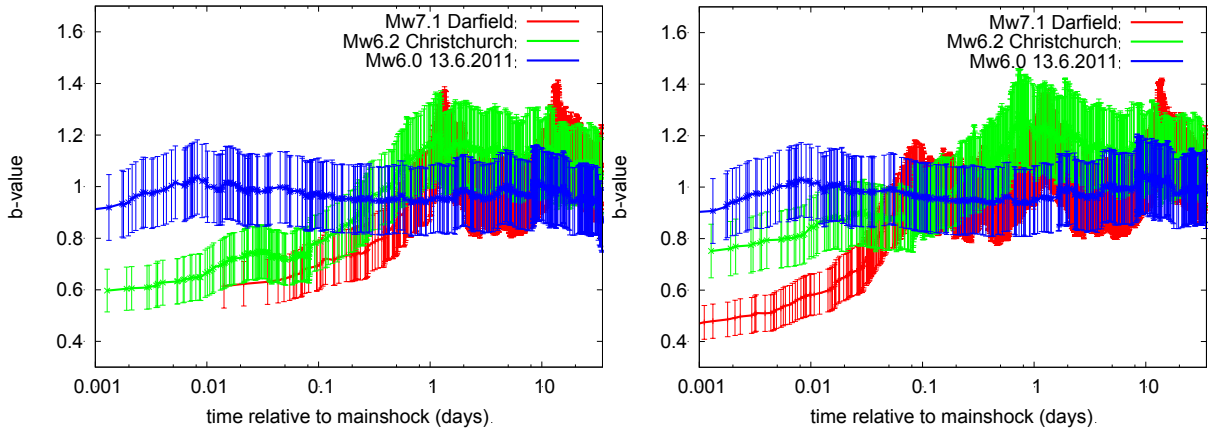

Figure 3.4 Maximum likelihood estimate of the *b*-value as a function of time following the Darfield (red), the Christchurch (green), and the June 2011 (blue) earthquakes for the two GeoNet catalogues (January 2012 on left and September 2012 on right). Each value refers to the 200 successive $M \geq 3.0$ events that occurred after the indicated time. The light coloured lines show 95% confidence intervals.

3.3 MODELLING THE CANTERBURY SEQUENCE WITH THE ETAS MODEL

The ETAS model is a stochastic point process model as explained in more detail in Section 2.6, which allows each earthquake in a cluster to have its own aftershock decay. We fitted the ETAS model to the best data available for the Canterbury earthquake sequence and the results are shown in Table 3.3. Figure 3.5 shows the daily rate of $M \geq 3.0$ earthquakes in red and the magnitude and timing of the observed earthquakes relative to the Darfield event. We also used the New Zealand generic aftershock parameters in the ETAS model as explained in Section 2.4 and shown in Table 3.3. Figure 3.6 compares the cumulative number of $M \geq 3.0$ earthquakes with time relative to Darfield for the two ETAS models. The results are consistent with Figure 1.2 which compares the observed data with the generic model applied only to Darfield and to all $M \geq 6.0$ earthquakes. Overall the sequence has been more productive than the average New Zealand aftershock sequence. Also, there have been fluctuations in the productivity with time. Using the maximum likelihood method to estimate parameters for the sequence ensures that the total number of observed earthquakes matches the modelled number of earthquakes. The right hand side of the plot shows the number of observed earthquakes versus their transformed time

$$T = \int_0^t \lambda(t)\,dt, \qquad \textbf{Equation 9}$$

where $\lambda(t)$ is the ETAS rate function (according to Equation 8). We did not use the spatial terms in the fitting. For a perfect fit of the ETAS model there would be a random fluctuation around the diagonal. The plot shows that more earthquakes than expected occurred from about earthquake 700 and even more so the Christchurch earthquake (around 1500). Around the June earthquake, there was a change in slope for the estimated parameter curve, which indicates that in relation to the ETAS model the sequence got quieter. This result reflects the fact that the Christchurch event was more effective and the June earthquake less so, compared to the expected magnitude-productivity scaling law.

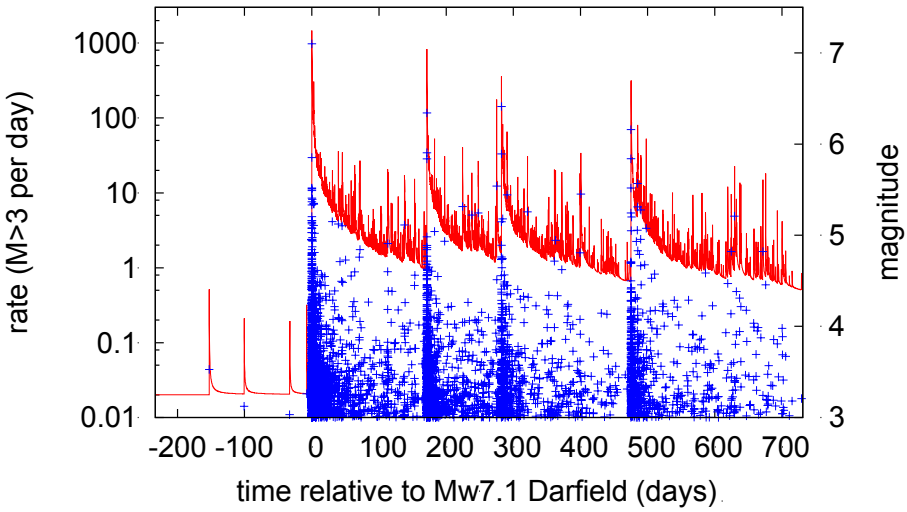

Figure 3.5 ETAS-model rate for the estimated parameters set (red lines) in comparison with observed $M \geq 3.0$ earthquakes (blue points).

Table 3.3 The ETAS model parameters used in the simulations and for forecasting.

Model	M_{max}	*b*-value	Background μ [1/days]	Productivity k_0	*c*-value [days]	α-value	*p*-value
Best fit	8.0	1.03	0.020	0.0137	0.019	0.90	1.11
Generic 7.5	7.5	1.03	0.020	0.0065	0.040	1.03	1.07
Generic 8.0	8.0	1.03	0.020	0.0065	0.040	1.03	1.07

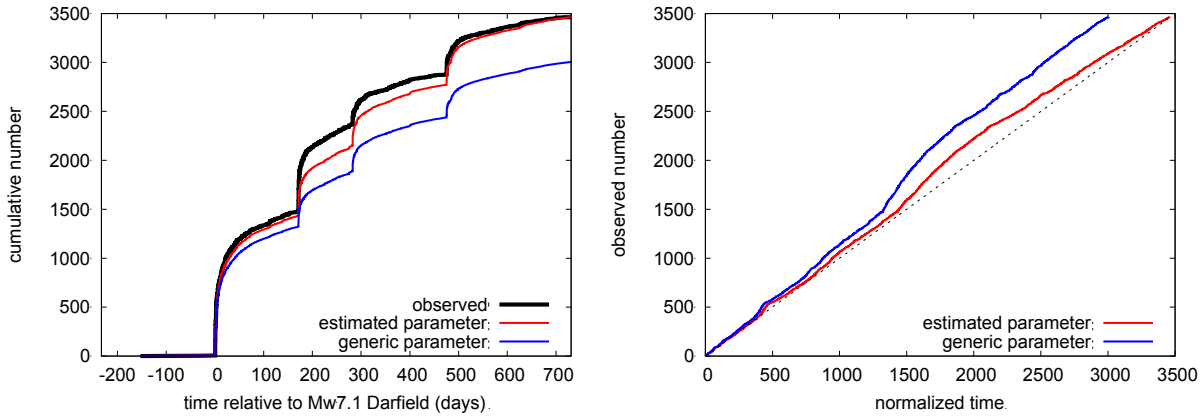

Figure 3.6 Fit of the ETAS model based on generic and estimated parameters for the $M \geq 3$ sequence. Note that the results for generic-7.5 and generic-8.0 are the same because magnitudes are directly taken from the catalogue. Furthermore, as a result of maximum likelihood fits, the total number predicted by the model is the same as the number of observed events and thus the red line intersects with the diagonal at the end point. The right-hand figure shows the transformed time as explained in the text.

We can use the ETAS model for retrospective and prospective forecasting of earthquake occurrence. In retrospective forecasting, we perform 1000 Monte-Carlo simulations of the seismicity following the Darfield M7.1 main shock based only on the information of earthquakes that occurred prior to the main shock and the main shock itself. For each simulation, we count the number of M≥*6.0* events in the first two years after the main shock. The resulting distribution of 1000 values is shown in Figure 3.7, where the results are compared for simulations with the New Zealand generic parameters and the best fitting

parameters as shown in Table 3.3. For the estimated parameters, the probability to observe three or more aftershocks is 36%, while it is 30% for the generic parameters. The simulations of the generic model with maximum magnitude of 8.0 has a higher probability of getting larger numbers of M≥6 aftershocks than the generic model with maximum magnitude of 7.5. This is due to the larger number of aftershocks that the more rare large earthquakes cause, and raises the question of what the optimum M_{max} would be.

In the case of the prospective forecast, we repeat the procedure but start the simulation at the end of the earthquake catalogue, i.e. in September 2012 and simulate for 50 years. In this case, all earthquakes in the catalogue are used as input information for the calculation of the ETAS rate $\lambda(t)$. The results for the expected number of M≥6.0 events within the next 50 years are shown in Figure 3.8.

Table 3.4 Results from the retrospective ETAS simulations

ETAS model	Average number of M≥*6.0*	Observed frequency of 1 or more M≥*6.0*	Observed frequency of 3 or more M≥*6.0*	95th percentile of number of M≥*6.0* events
Generic parameters with M_{max} = 7.5	1.9	72%	30%	13
Generic parameters with M_{max} = 8.0	2.3	73%	29%	19
Estimated parameters with M_{max} = 8.0	2.9	76%	36%	26

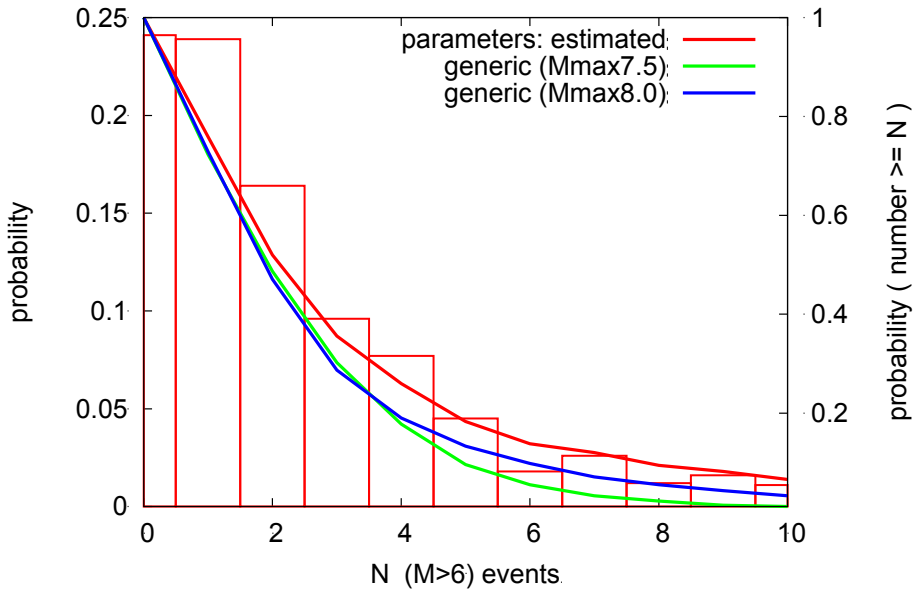

Figure 3.7 Retrospective ETAS-model forecasts of the number of M≥6 events following two years after the Darfield earthquake, i.e. covering the period of the catalogue. The red histograms and left axis of plot show the relative frequency of occurrence of the number of large earthquake for the estimated parameters. The solid curves and right axis of plots gives the probability of observing N or more large aftershocks for the estimated parameters and the generic parameters with different maximum magnitude M_{max}.

Table 3.5 Results from the prospective ETAS simulations.

ETAS model	Average number of *M*≥6.0	Observed frequency of 1 or more *M*≥6.0	Observed frequency of 3 or more *M*≥6.0	95th percentile of number of *M*≥6.0 events
Generic parameters with M_{max} = 7.5	3.4	83%	47%	12
Generic parameters with M_{max} = 8.0	5.1	84%	47%	21
Estimated parameters with Generic parameters with M_{max} = 8.0	7.7	93%	67%	28

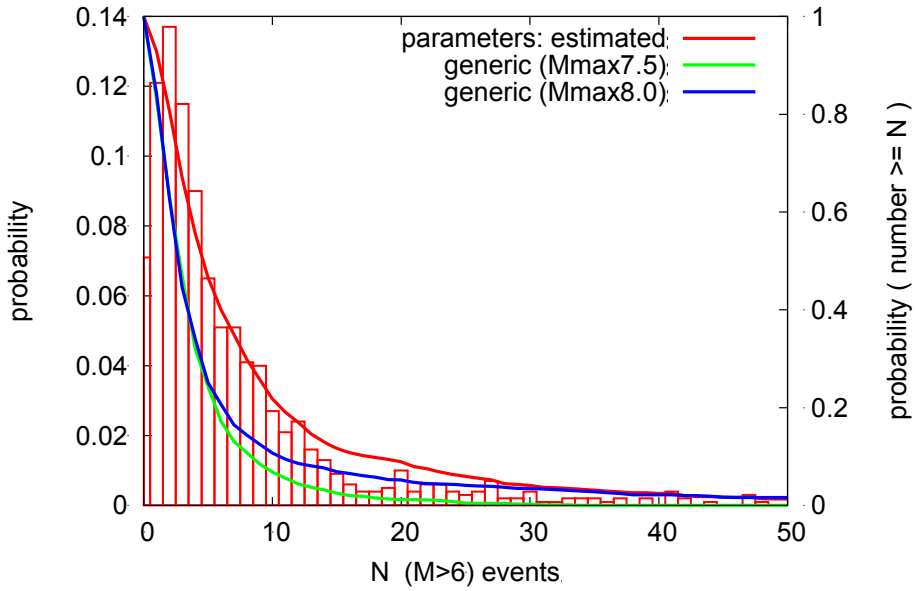

Figure 3.8 Prospective ETAS-model forecasts of the number of *M*≥6 events following 50 years after the end of the catalogue. The red histograms and left axis of plot show the relative frequency of occurrence of the number of large earthquake for the estimated parameters. The solid curves and right axis of plots gives the probability of observing *N* or more large aftershocks for the estimated parameters and the generic parameters with different maximum magnitude M_{max}.

3.4 HOW MANY MORE M≥6.0 AFTERSHOCKS TO EXPECT?

As already noted in the introduction, the EE model expects about 1.3 M≥6 earthquakes in the 50 years starting in July 2012 in the area shown in Figure 1.1, which corresponds roughly to the aftershock area used for the ETAS modelling. The probability of observing one or more M≥6 earthquakes in the 50 years is 73%.

Applying the STEP model to the best available data to forecast the expected number of earthquakes for the 50 years starting on 1 September 2012 results in 0.91 M≥6 earthquakes for the Canterbury sequence and thus a 60% probability of another M≥6 earthquake.

The new generic model expects an average of 3.5 M≥6 earthquakes to occur within two years following a M7.1 main shock, and 4.2 M≥6 earthquakes within 50 years. However, by the time the Christchurch earthquake had occurred, the expected number of M≥6 earthquakes from the generic model had already reduced to 1.1. Applying the generic model for 50 years starting two years after a M7.1, i.e. in September 2012, results in 0.7 M≥6 earthquakes and therefore roughly a 50% probability of observing another M≥6 earthquake.

Originally we had planned to derive aftershock parameters from global earthquake catalogues and apply them to the Canterbury sequence to provide some bounds on the long-term rates. However, we found larger variation in parameters from fitting the Omori-Utsu law to three M≥6 earthquakes in the Canterbury sequence (Table 3.2) than we would have expected from averaging of global sequences (Christophersen, 2000). The number of expected M≥6 earthquakes for 50 years from September 2012 range from 0.6 to 1.1 for those three earthquakes. Making the simple assumption that the three M≥6 earthquakes have their independent Omori-Utsu decay, the combined expected number of M≥6 earthquakes in the 50 years from September 2012 is 2.3 which corresponds to a 90% probability of one or more M≥6 earthquake.

The average number of M≥6 earthquakes for the prospective ETAS simulations are up to six times higher than the EE model and have an observed occurrence of up to 93% for another M≥6 earthquake (Table 3.5). Because the ETAS model simulates aftershocks of aftershocks the average number of large earthquakes can be expected to be larger than from the other models that do no forward simulations. However, we have also seen that the average number of M≥6 earthquakes varies by a factor of two for the different input parameters (Table 3.5). The maximum magnitude of aftershocks M_{max} is an important parameter because large aftershocks trigger their own cluster of aftershocks, and the more earthquakes there are in the cluster, the more likely it is to have a large event. Less obvious is the importance of the aftershock decay parameter p. Harte (2013, Table 1) pointed out the long time it takes for different combinations of c and p different percentiles of events to occur. For the generic New Zealand parameter p=1.07, and a c-value of 0.04 days, it takes nearly 800 days for 50% of aftershocks to occur. Given the long tail of the Omori-Utsu law, the fitting period of two years is small, and the truncation of the data is likely to lead to smaller p-values than realistic. We had similar problems when applying the ETAS model to simulate 50 years of aftershocks for a hypothetical Wellington Fault earthquake (Christophersen, et al., 2013). The observation of bursts of aftershock activity on a time-scale of decades rather than months as observed in the Canterbury sequence led to the EQC proposal ‘Testing aftershock models on a time-scale of decades’ which is about to start. The new project will work towards a better understanding of useful timeframes for the application of aftershock models.

4.0 WHAT WE CAN LEARN FROM GLOBAL EARTHQUAKE CATALOGUES

One of our research questions asks how usual or unusual late and large aftershocks are. The M7.1 Darfield earthquake had three M≥6.0 aftershocks with time delays of months between them and at distances beyond the original fault area. We first ask the question how often do main shocks have three or more aftershocks within 1.1 magnitude units. We then look at the distribution of aftershocks in time relative to the main shock. As aftershocks generally follow the Omori-Utsu law regardless of the magnitude of the main shock, we can directly compare the occurrence of the Christchurch, June and December 2011 earthquake with the time delay of all other global large aftershocks. In space, aftershock area usually increases exponentially with main shock magnitude. Therefore we compare relative distances between the Darfield earthquake and its large aftershocks with other relative distances between main shocks and their large aftershocks, i.e. we use a scaling relation between magnitude and fault rupture length to calculate relative distances between main shock and aftershocks. Finally we want to know whether there are any other aftershocks that were within 1.1 magnitude units of the main shock and occurred as far away in space and time as the Christchurch earthquake from the Darfield event.

To answer all these questions we aim to use the longest available complete earthquake records, i.e. data that have been recorded and processed consistently and are complete to a low magnitude of completeness M_C. The United States Geological Survey provides the Centennial Earthquake Catalogue which covers the time period 1900-2007 (Engdahl & Villaseñor, 2002; USGS, 2013). The International Seismological Centre has recently completed an earthquake catalogue compilation for the Global Earthquake Model (GEM) project, which covers the time period 1900-2009 (Storchak et al, 2013). We describe both catalogues in more detail below. Both catalogues rely on a global seismic network to detect earthquakes. In the very early 20th century, only 20-30 standardised seismometers were installed globally (Ammon et al., 2010) and therefore the global detection completeness is not below magnitude 7 (e.g. Engdahl & Villaseñor, 2002; Storchak et al., 2013). Between 1960 and 1966, the World-Wide Standardized Seismograph Network (WWSSN) was constructed to help detect underground nuclear testing (Ammon et al., 2010). As a consequence M_C decreased significantly. Doubling the length of the earthquake record from, say 50 to 100 years, on average doubles the amount of data available. On the other hand, lowering M_C, by one magnitude unit, e.g. from 7 to 6, on average increases the number of earthquakes by a factor of 10 due to the nature of the magnitude-frequency relation of earthquake occurrence (Section 2.2). Thus there is a trade-off between a long earthquake record in time and a low M_C, and we therefore look at two different subsets for each catalogue, i.e. four data sets in total which range in size from 1,023 to 12,565 earthquakes.

Once we have complete sets of earthquake data, we define earthquake sequences. As no physical differences are known to exist between foreshocks, mainshocks and aftershocks, earthquake clusters are usually defined by their closeness in space and time to one another. We use a simple window method as described in Section 4.2 to define earthquake clusters and vary the size of the window length. Section 4.3 discusses the answers to the questions posed above.

4.1 THE EARTHQUAKE CATALOGUES

The Centennial Catalogue (Engdahl & Villaseñor, 2002) is a global catalogue of locations and magnitudes of instrumentally recorded earthquakes from 1900 to 2007, which was put together with the aim of providing a realistic picture of the global seismicity distribution. It is based on a number of pre-existing catalogues and used a hierarchical scheme for integrating hypocentre and magnitude information. Earthquakes for which phase data were available were relocated. The magnitude cut-off for the period prior to the WWSSN is 6.5 and the estimated completeness 7.0; since the introduction of the WWSSN in 1964, the magnitude cut-off is 5.5 and the catalogue is deemed to be complete at this level (Engdahl & Villaseñor, 2002). We investigated two sub-sets of the catalogue. First the complete catalogue with cut-off magnitude 7.0, and second the post WWSSN period with cut-off magnitude 5.5. We also cut the catalogue in depth and only included earthquakes shallower than 70km. Table 4.1 shows a summary of available earthquakes.

The International Seismological Centre–Global Earthquake Model (ISC–GEM) Global Instrumental Earthquake Catalogue covers the time period 1900–2009, and stems from a special project by the Global Earthquake Model (GEM) Project to improve and extend existing global earthquake catalogues for the purpose of seismic hazard and risk assessment (Storchak, et al., 2013). Prior to the start of the project the following cut-off magnitudes were set: M_C ≥7.5 for 1900-1918, M_C ≥6.25 for 1919-1959, and M_C ≥5.5 for 1960-2009. The estimated completeness for the final catalogue is M_C =6.4 for 1900-1963, and M_C =5.6 for 1964-2009 (Storchak et al., 2013). However, there are not many earthquakes with M≥6.25 in the catalogue prior to 1919 (Storchak et al., 2013, Figure 2). We therefore select the following two sub-sets of the catalogue: First 1919-2009 with M_C =6.4, and second, 1964 - 2009 with M_C = 5.6. We also cut the catalogue in depth and only included earthquakes shallower than 70km. A summary of the available earthquakes is included in Table 4.1.

The subset of the catalogues selected for further study cover slightly different time intervals and different time periods. Therefore the number of earthquakes cannot be compared directly. However, Table 4.1 already gives the impression that the ISC-GEM catalogue has more data, which is confirmed in Figure 4.1, which compares the magnitude and time of occurrence of earthquakes in the two global catalogues.

Table 4.1 Time period and available number of shallow earthquakes for subsets of the Centennial and ISC-GEM catalogue that are assumed to be complete.

Catalogue	M_C	Period	Number of earthquakes
Centennial 1	7.0	1st Jan. 1919- 3rd Sept. 2007	1023
Centennial 2	5.5	1st Jan. 1964- 3rd Sept. 2007	8712
ISC-GEM 1	6.4	1st Jan. 1919- 31st Dec. 2009	3701
ISC-GEM	5.6	1st Jan. 1964- 31st Dec. 2009	12565

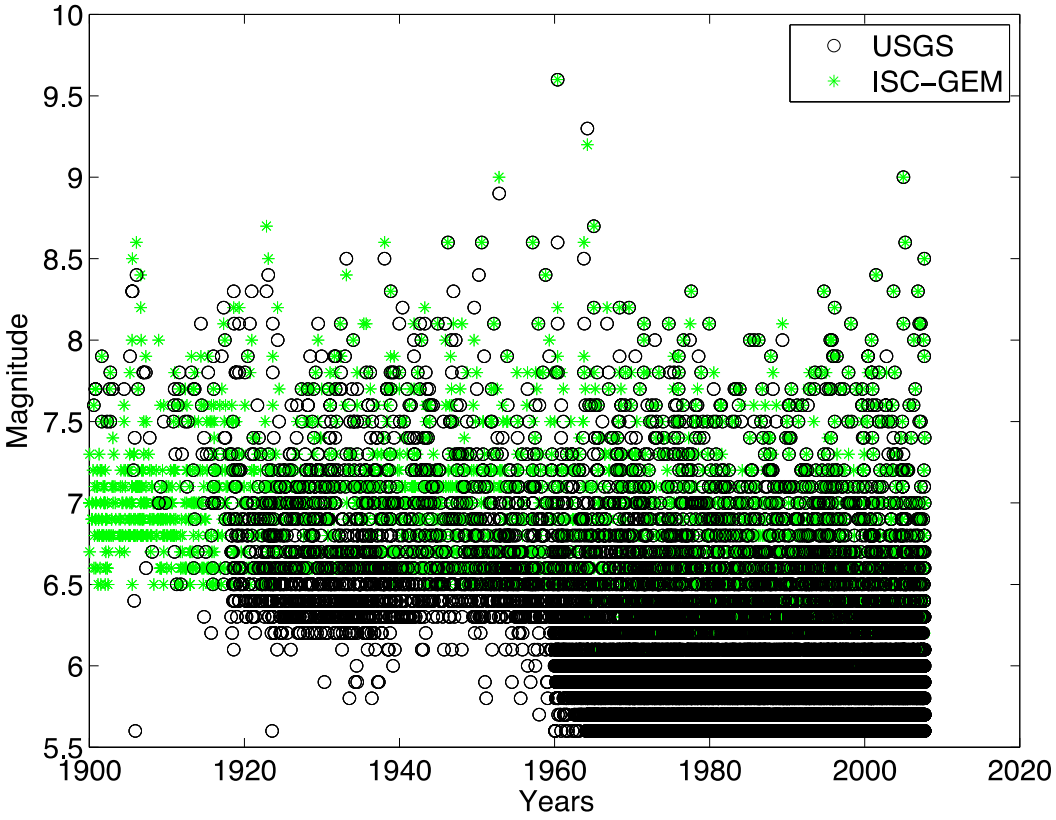

Figure 4.1 A magnitude-time plot comparing the USGS Centennial catalogue with the ISC-GEM catalogue.

4.2 DEFINING EARTHQUAKE SEQUENCES

Most algorithms for identifying earthquake sequences have been developed for removing aftershocks, and thus involve 'declustering' of earthquake catalogues. In principle, two approaches for declustering exist: stochastic and deterministic algorithms. The stochastic algorithms calculate probabilities for each earthquake to be triggered by a preceding one. Examples are the stochastic declustering, based on the ETAS model (Zhuang, et al., 2002) and the model independent clustering algorithm (Marsan & Lengliné, 2008), which requires the users to make informed decisions about spatial and temporal distribution of clusters and therefore implies an underlying model on which to base those decisions. Any earthquake can thus have a probability of being associated with a number of different clusters. For the deterministic algorithms, each earthquake is taken to be part of one cluster. Two different approaches can be distinguished: (i) linking algorithms where clusters are linked by smaller earthquakes and are allowed to grow in time and space as seismicity develops; and (ii) window algorithms where magnitude-dependent windows in space and time are used to identify earthquakes of the same cluster. Both the stochastic method and the linking algorithm usually need data to a low M_C to be able to associate earthquakes with one another. For example, the Reasenberg declustering algorithm does not identify a M6.0 earthquake following 17 days after and within 30 km of the epicentre of the 1855 M8.2 Wairarapa earthquake as an aftershock (Christophersen, et al., 2011). For the purpose of analysing large aftershocks of large main shocks, we use a window method in time and space to define aftershocks. We start the search with the largest earthquake in the catalogue and search within a magnitude-dependent spatial radius, and a selected time window for aftershocks. The time window is restarted each time an earthquake is found within the spatial radius around the main shock. Once there has been no earthquake within the search radius for the duration of the time window, the sequence is terminated, and the search is restarted

from the largest earthquake in the catalogue that has not been associated with a sequence. Thus each earthquake can only be part of one sequence, and the search for related earthquakes is centred around the largest earthquake in a sequence. The search radius in space is based on scaling relation between the magnitude of an earthquake and its rupture length (Wells & Coppersmith, 1994, Table 2.3)

$$r(M) = 10^{0.59M - 2.44}. \qquad \textbf{Equation 10}$$

Table 4.2 shows the search radius $r(M)$ according to Equation 10 for selected main shock magnitudes. We test the sensitivity to our selected windows by doubling them. The search radius is also used to calculate the relative distance between main shocks and their aftershocks.

Table 4.2 The search radius $r(M)$ in km for different main shock magnitude M, according to scaling relations by Wells & Coppersmith (1994).

M	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.3
r(M) [km]	6.4	13	25	49	97	191	376	741	1462

The Canterbury sequence started with the M7.1 Darfield earthquake which was the largest earthquake in the sequence and had three M≥6.0 earthquakes; therefore we define a large earthquake to be within 1.1 magnitude units of the mainshock. Table 4.4 provides an overview of how many main shocks and large aftershocks there are in the different data sets for the different search criteria. The main shock has to be at least 1.1 magnitude units larger than M_C so that each main shock has the same chance to sample large aftershocks which are within 1.1 units of its magnitude. This consideration does not take into account uncertainty in the magnitudes.

Table 4.3 The number of main shocks and total number of large aftershocks for the different clustering parameters where 1R means the aftershock occurred within one radius of the main shock search radius according to Equation 10 and Table 4.2, and 2R is double the radius. 1T refers to rolling time windows of one year, and 2T to rolling time windows of two years. The red numbers highlight the number of aftershocks for the analysis in Figure 4.4 – Figure 4.7.

Catalogue	***Min*** M_m	**1R1T**		**1R2T**		**2R1T**		**2R2T**	
		# M_m	**#** ***AS***	**#** M_m	**#** ***AS***	**#** M_m	**#** ***AS***	**#** M_m	**#** ***AS***
Centennial 1	8.1	47	16	47	19	47	19	46	32
Centennial 2	6.7	724	316	652	343	629	420	499	427
ISC-GEM 1	7.5	284	137	281	159	277	169	254	214
ISC-GEM 2	6.7	768	455	696	483	643	563	529	584

4.3 RESULTS

Figure 4.2 and Figure 4.3 show the histograms of the number of large aftershocks per main shock of the different search criteria for the larger subsets of the two catalogues. Table 4.4 shows the percentage of sequences that have at least three or more large aftershocks. The percentage varies from 3-10% for the Centennial catalogue and 6-13% for the ISC-GEM catalogue. The higher percentages in the latter catalogue reflect that there are more events in the catalogue and probably means that some aftershocks were missed in the Centennial catalogue. As expected, the percentage increases with increasing search windows, more so in time than in space. The results show that it is not particularly unusual for a main shock to have three or more aftershocks within one magnitude unit.

Table 4.4 The percentage sequences that have three or more large aftershocks for the larger subset of each catalogue.

Catalogue	1R1T	1R2T	2R1T	2R2T
Centennial 2	3%	7%	5%	10%
ISC-GEM 2	6%	10%	8%	13%

For Figure 4.4 and Figure 4.5 we calculated the time difference for all large aftershocks to their main shocks and sorted the times and plotted them against number of aftershocks. The red, blue and green vertical lines show the time in days between the Darfield earthquake and the Christchurch (171 days), June (282 days) and December 2011 (475 days) aftershocks, respectively. The time axis is logarithmic; if all aftershocks were to follow the Omori-Utsu law we would expect a straight line. The data fall on a straight line from about 10^{-2} days (14 minutes) until about the occurrence of the Christchurch aftershocks for the search criteria of one fault length (R=1) and a rolling time window of one year (1T). Earlier aftershocks might have been missed in the coda of the main shock, and therefore the data fall below the line for small times after the main shock. Increasing the search parameters in time and/or space leads to a steepening of the slope at later times. This could be an indication that earthquakes are included that are not part of the sequence because they occurred further away in time and space and therefore do not follow the Omori-Utsu law. With larger time delays, and greater distances from the main shock it is more likely to include unrelated earthquakes. Looking at the plots, 1R and 1T are our preferred search criteria. Both catalogues show the same trends. The data for the Centennial catalogue, which has fewer earthquakes, look a bit more scattered.

It is not so clear what relationship the data should follow in space. For Figure 4.6 and Figure 4.7 we calculated the relative distances between main shock and aftershock for all large aftershocks, which were highlighted in red in Table 4.3. Again we sorted them in increasing order and plotted the relative distance against number of aftershocks. The plots include the relative distances of the Christchurch (red), June (blue) and December (green) earthquakes as reference lines, i.e. 42 km, 46km, and 47 km, respectively divided by the search radius of 56 km for a M7.1 main shock. Except for very small relative distances, the data fall on a straight line until the occurrence of the large Canterbury aftershocks, then the data start falling below the line. This observation holds for both catalogues and all search parameter combinations. When doubling the search radius, more large aftershocks are found but not at the same rate as up to relative distances of about 0.8. This is interesting as the search area quadruples when doubling the search radius and could be an indication that the aftershock area of the large main shocks in the database reaches the boundaries of the seismogenic zone available for earthquake rupture.

Table 4.5 shows the percentage of large aftershocks at a smaller relative distance *D* from the main shock, and an earlier time of occurrence than the Christchurch earthquake to the Darfield earthquake for different search criteria and the four different data subsets. For the preferred search criteria of one magnitude dependent fault length in space and rolling time windows of one year in time, we see that the Christchurch earthquake occurred around the 80^{th} percentile in space for all data subsets, and the 85^{th} percentile in time. This confirms, again, that the Christchurch aftershock was distant in space and time but not unusually so in the context of global aftershock statistics. For increasing search windows in space and time, more earthquakes further away in space and time are found, and the Christchurch earthquake occurs around the 50^{th} percentile in space and between the 50^{th} and 60^{th} percentiles in time.

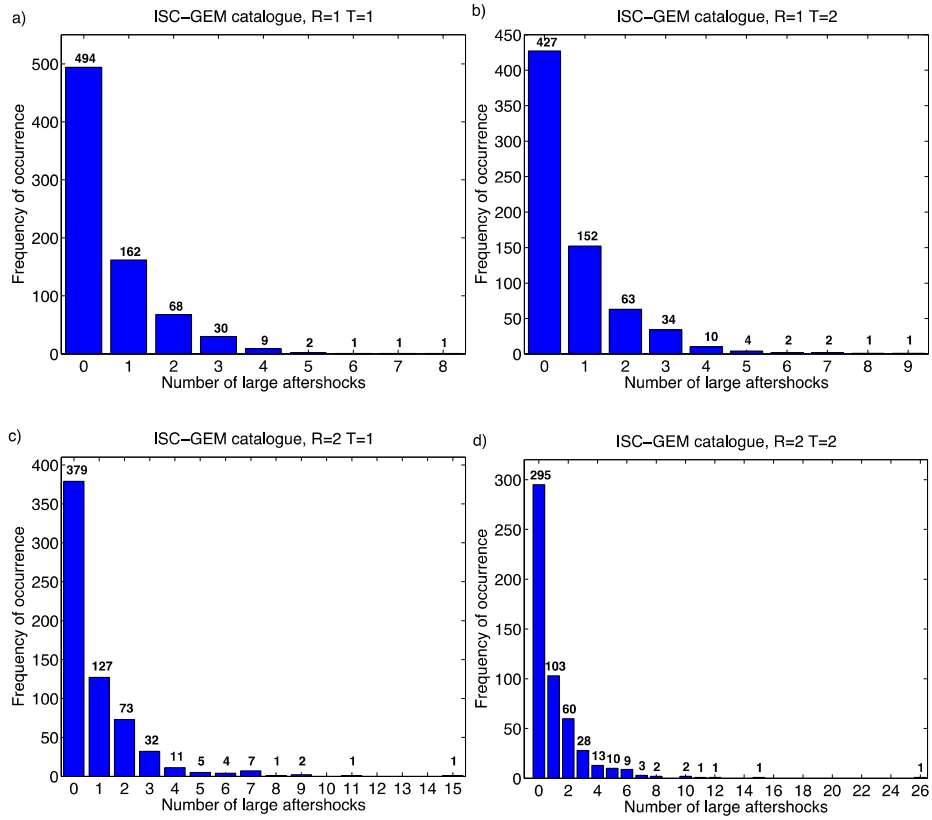

Figure 4.2 Histogram of the number of large aftershocks per main shock for the different search criteria in the ISC-GEM catalogue.

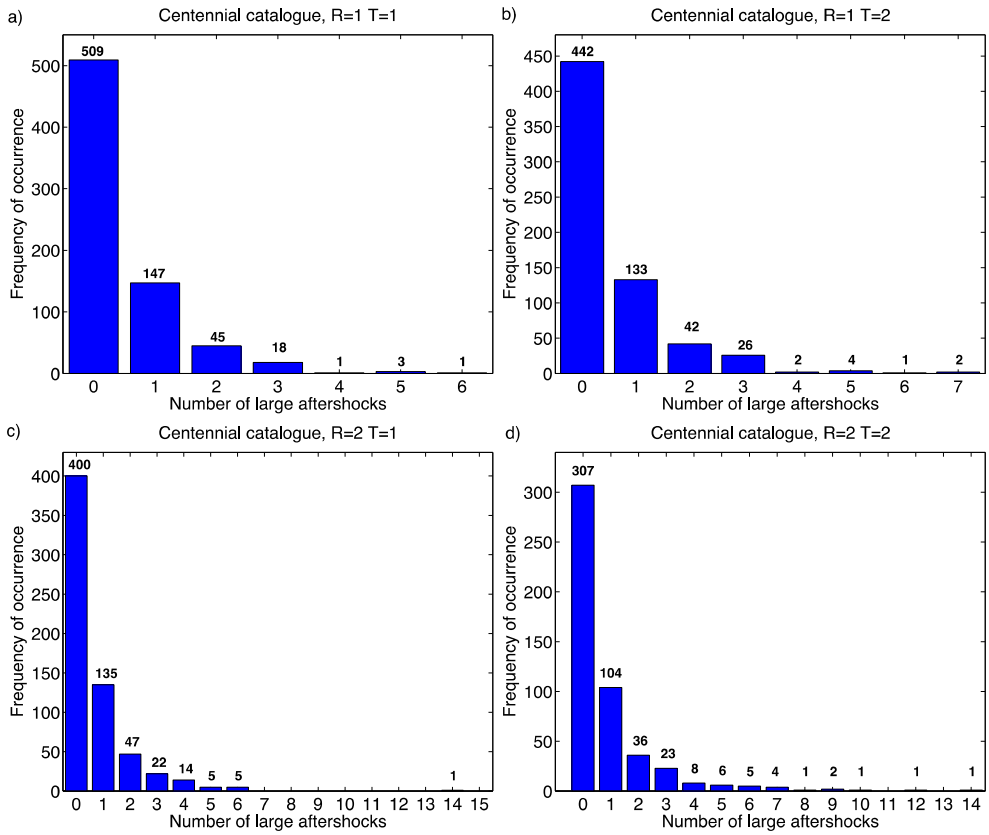

Figure 4.3 Histogram of the number of large aftershocks per main shock for the different search criteria in the Centennial catalogue.

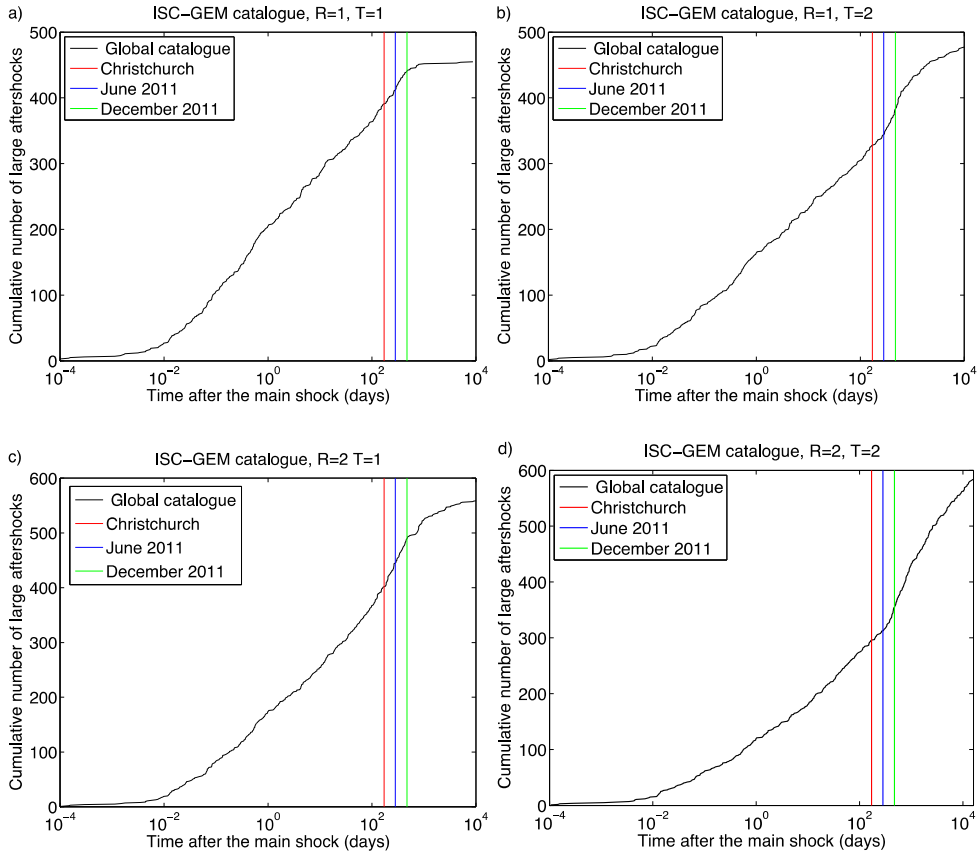

Figure 4.4 The distribution of relative time of occurrence of large aftershocks from the main shock for four different search criteria for the ISC-GEM catalogue.

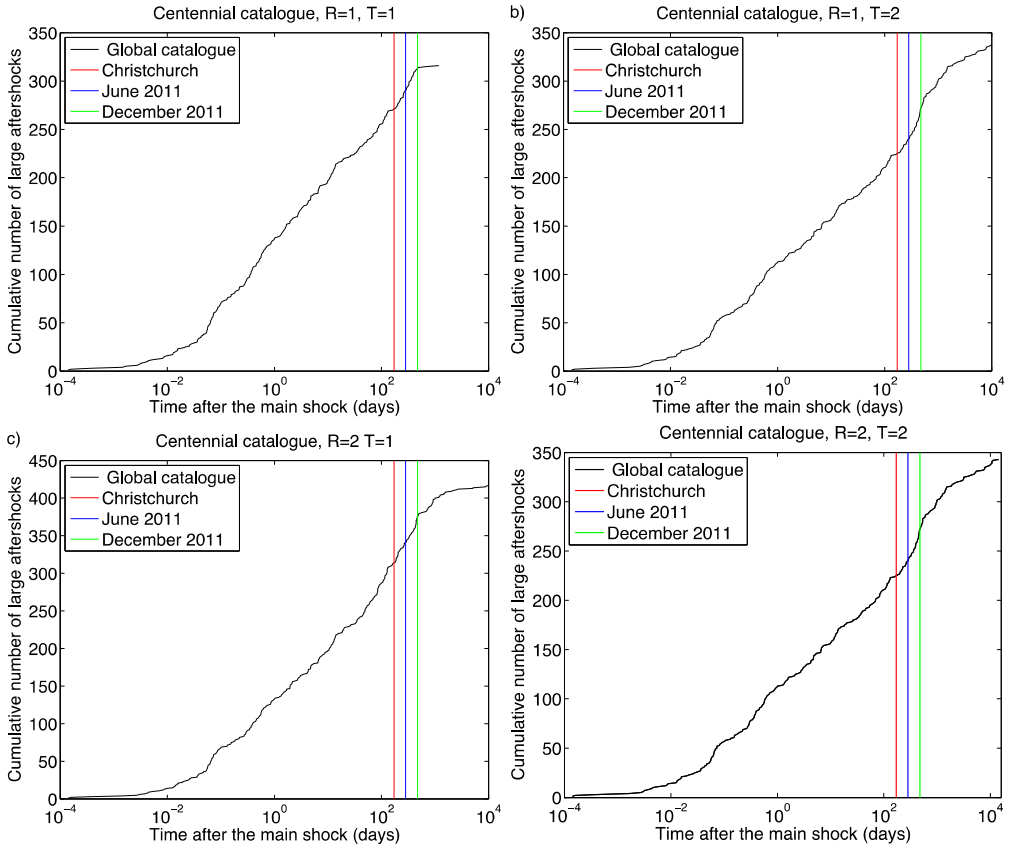

Figure 4.5 The distribution of relative time of occurrence of large aftershocks from the main shock for four different search criteria for the Centennial catalogue.

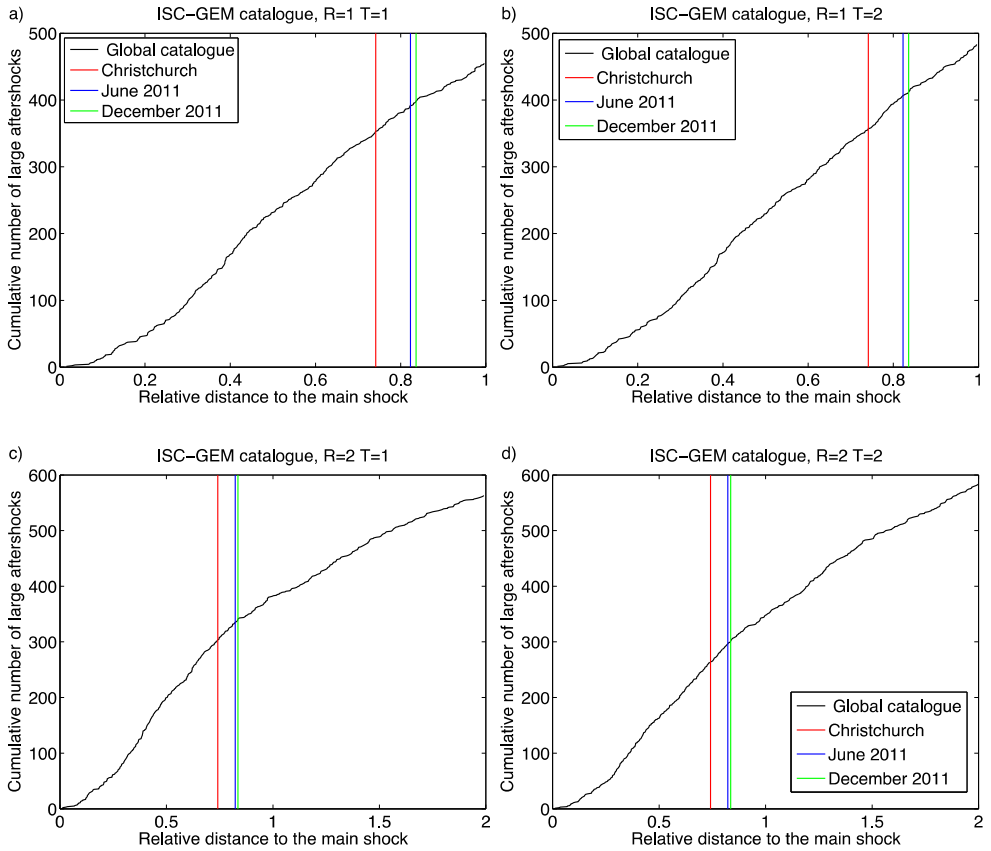

Figure 4.6 The distribution of relative distance of large aftershocks from the main shock for four different search criteria for the ISC-GEM catalogue.

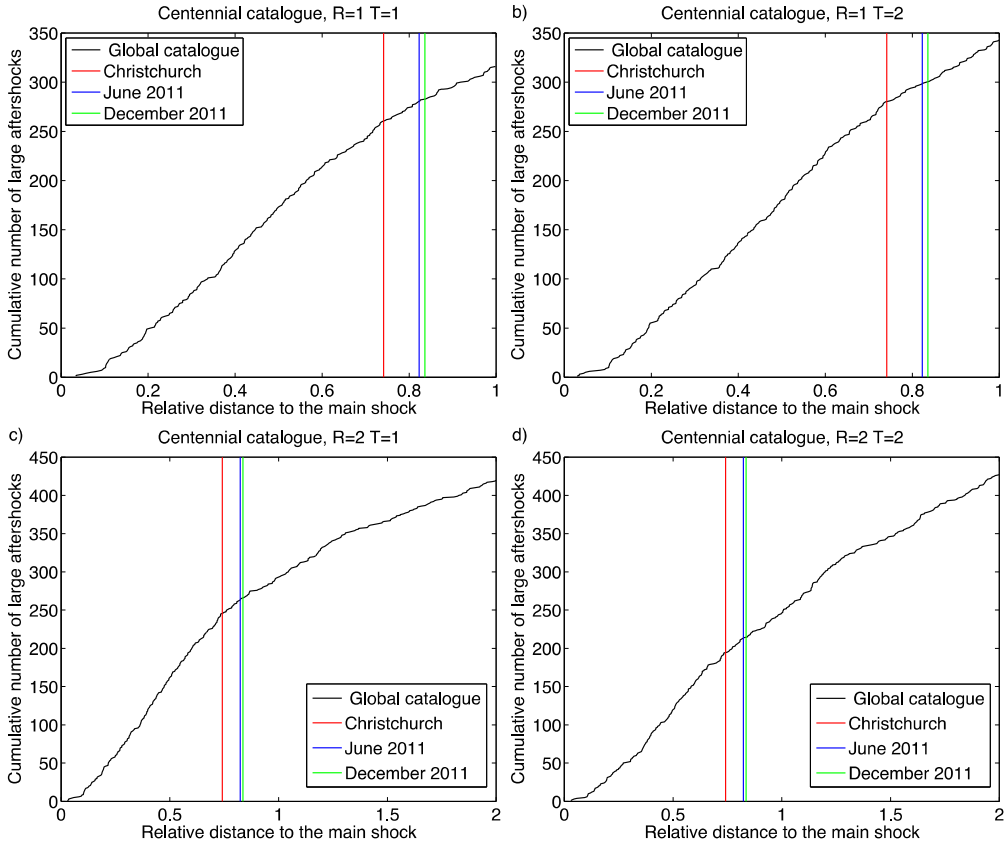

Figure 4.7 The distribution of relative distance of large aftershocks from the main shock for four different search criteria for the Centennial catalogue.

Table 4.5 The percentage of large aftershocks at a smaller relative distance *D* from the main shock, and an earlier time of occurrence than the Christchurch earthquake to the Darfield earthquake, for different search criteria and the four different data subsets.

Catalogue	1R1T		1R2T		2R1T		2R2T	
	D	***T***	***D***	***T***	***D***	***T***	***D***	***T***
Centennial 1	83%	84%	80%	69%	63%	79%	53%	57%
Centennial 2	82%	86%	81%	65%	58%	75%	45%	48%
ISC-GEM 1	80%	90%	78%	78%	57%	81%	51%	60%
ISC-GEM 2	77%	86%	73%	68%	54%	71%	45%	51%

Our final question in respect to late and large aftershocks is whether there are any other aftershocks that were within 1.1 magnitude units of their main shock and occurred as far away in space and time from it as the Christchurch earthquake from the Darfield event. We concentrated our search on the dataset with the preferred search criteria to avoid picking up too many unrelated earthquakes. We first searched the ISC-GEM 2 dataset. Out of 455 large aftershocks, 9 large aftershocks in 7 earthquake sequences were at least as far away in time and space from their main shock as the Christchurch earthquake from the Darfield earthquake. This corresponds to 2% of all large earthquakes. Most of them are part of four aftershock sequences in the Solomon Islands or the Solomon Sea. One is an aftershock of the 1987 Antofagasta, M7.5 Chile earthquake, and another is an aftershock to the 2004 Papua Indonesia earthquake. Searching the longer database of the ISC-GEM catalogue with a minimum main shock magnitude of 7.5 that has 137 aftershocks as listed in Table 4.3, we find no large aftershocks that fulfil the search criteria. Removing the main shock magnitude threshold, we find four large aftershocks that are at least as late and distant as the Christchurch to the Darfield earthquake. Three of those occurred later than 1964 and have already been identified in the ISC-GEM 2 dataset. The fourth is an aftershock to the 1937 Taidong earthquake in Taiwan. Identical searches in the Centennial catalogue find only one additional large aftershock as part of the 1980 M7.5 main shock near New Caledonia. In summary, there are other examples of large aftershocks, i.e. within 1.1 magnitude unit of the main shock, that occur at least as late as the Christchurch to the Darfield earthquake and with at least the same relative distance. For our most complete dataset, there are 9 examples, which corresponds to 2% of all large aftershocks. Another couple of examples can be found in the other datasets but all examples occur in areas with much higher strain rates than the Canterbury region.

5.0 CONCLUSIONS AND OUTLOOK

Our project 'The Canterbury sequence in the context of global earthquake statistics' had the aim to provide bounds for earthquake rates in the Canterbury region for the coming years and decades and to investigate how unusual late and large aftershocks are as observed in the Canterbury sequence. The probability of a further M≥6.0 earthquake in the 50 years starting from September 2012 ranges from 50 to 93% depending on the model. We systematically searched for aftershock sequences in two global earthquake catalogues within two different time periods and for two magnitude cut-offs to satisfy completeness as well as possible, using a range of search criteria. For the more complete catalogue and the preferred search criteria, 6% of main shocks have three or more large aftershocks within 1.1 magnitude units of the main shock. The relative distance, i.e. the epicentral distance between the main shock and the large aftershock divided by a magnitude-dependent length for the rupture length of the main shock, and the time delay between the Christchurch and the Darfield earthquakes were around the 80th and 85th percentiles, respectively, of all the distributions derived from all large aftershocks and their main shocks. There are about a dozen examples in the global earthquake data where a large aftershock occurred with at least as great a relative distance and time delay as the Christchurch to the Darfield earthquake. Thus in the context of global earthquake statistics, the Canterbury sequence is not unique.

During our project we realised that the real-time GeoNet earthquake catalogue misses many earthquakes, especially within the first couple of days following a M≥ 6.0 earthquake. In the case of the Canterbury sequence it took up to 18 months before the data was added to the catalogue. For example, the January 2012 catalogue still had only 6 M≥5.0 aftershocks in the 24 hours after the Darfield earthquake. The New Zealand generic aftershock model expects around 15 magnitude 5 and larger aftershocks within 24 hours of a M7.1 main shock. As a consequence we thought for a long period of time that the Darfield earthquake was less productive than the average New Zealand earthquake sequence. The September 2012 GeoNet catalogue had 20 M≥5.0 aftershocks within 24 hours of the Darfield event, and thus the Darfield earthquake was not 'under-productive' compared to other New Zealand main shocks. We further illustrated the deficiency of the catalogue by comparing the January 2012 and September 2012 catalogues downloaded from GeoNet for a couple of days following each M≥6.0 earthquake in the Canterbury sequence. We showed the effect on the Omori-Utsu law for aftershock decay and the Gutenberg-Richter relation for the magnitude-frequency distribution. These empirical relationships form the basis for the aftershock forecast models, which are used in real-time earthquake forecasting. Thus, the time that is really important for accurate fitting of aftershock models is deficient of data. More work is required to understand this effect on these models.

Since the Canterbury sequence began, GeoNet has changed its data acquisition system from the CUSP system to the new SeisComp3 system. In this project we only used data from the CUSP system, as more work is required to understand how consistent the two systems are. The Cook Strait sequence in July and August 2013 has provided an earthquake sequence that will enable us to learn more about the capability of SeisComp3 to detect aftershocks following a large earthquake.

For this project, as well as another recent project (Christophersen, et al., 2013) we used the ETAS model for simulating aftershock sequences for time periods up to 50 years. We found bursts of activity as observed in the Canterbury sequence on a time scale of months but on a

time scale of decades. In the current project, the average number of M≥6.0 earthquake for the 50-years period from September 2012 was up to a factor of six higher than determined by the EE Model. The maximum magnitude for aftershocks is an important parameter that controls the total number of large earthquakes. We also discussed the importance of the aftershock decay parameter *p*. Using the Omori-law parameter from the New Zealand generic model, less than 50% of aftershocks occur within the first two years after a main shock. Thus from an ETAS model perspective, the Canterbury sequence will continue for thousands of years. However, the two years of data that were available for fitting model parameters was comparatively small, and one must be very cautious about making extrapolations in the future. We have another EQC project underway to investigate the applicability of aftershock models on a time-scale of decades in which we propose to include a parameter for the duration of a sequence.

In summary, we answered our original research questions, not necessarily by the means that we had originally anticipated as we shifted from the analysis of global aftershock parameters to the analysis and simulation of the Canterbury sequence. We have identified a suite of further work required to learn more about the quality of the real-time data and their effect on aftershock models so that we can improve real-time earthquake forecasting in New Zealand.

6.0 ACKNOWLEDMENTS

This research was funded by the Earthquake Commission Research Foundation under grant number 12/634. The report has been internally reviewed by Hannah Brackley and David Harte.

7.0 REFERENCES

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

Bath, M. (1965). Lateral inhomogeneities of the upper mantle. *Tectonophysics,* **2**(6), 483-514.

Beavan, J., Motagh, M., Fielding, E. J., Donnelly, N., & Collett, D. (2012). Fault slip models of the 2010 - 2011 Canterbury, New Zealand, earthquakes from geodetic data and observations of postseismic ground deformation. *New Zealand Journal of Geology and Geophysics,* **55**(3), 207-221.

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel inferences A Practical Information-Theoretic Approach*: Springer, New York.

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

Christophersen, A., Gerstenberger, M., Rhoades, D., & Stirling, M. (2011). Quantifying the effect of declustering on probabilistic seismic hazard, *Ninth Pacific Conference on Earthquake Engineering*. Auckland, New Zealand.

Christophersen, A., Rhoades, D., & Hainzl, S. (2013). *Sensitivity study of aftershock occurrence for a Wellington Fault earthquake.* Paper presented at the New Zealand Society for Earthquake Engineering Technical Conference, Wellington, New Zealand.

Engdahl, E. R., & Villaseñor, A. (2002). 41 Global seismicity: 1900-1999. In H. K. P. C. J. William H.K. Lee & K. Carl (Eds.), *International Geophysics* (pp. 665-XVI): Academic Press.

Gerstenberger, M., Wiemer, S., & Jones, L. (2004). Real-time Forecasts of Tomorrow's Earthquakes in California: a New Mapping Tool. *United States Geological Survey Open-File Report,,* **2004-1390**.

Gerstenberger, M. C., McVerry, G., Rhoades, D. A., & Stirling, M. (2013). Seismic Hazard Modelling for the Recovery of Christchurch, New Zealand. *Earthquake Spectra*(Canterbury Special volume), in press.

Gerstenberger, M. C., Wiemer, S., Jones, L. M., & Reasenberg, P. A. (2005). Real-time forecasts of tomorrow's earthquakes in California. *Nature,* **435**(7040), 328-331.

Gutenberg, R., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America,* **34**, 185-188.

Hainzl, S., Christophersen, A., & Enescu, B. (2008). Impact of earthquake rupture extensions on parameter estimations of point-process models. *Bull. Seism. Soc. Am.,* **98**(4), 2066-2072.

Harte, D. S. (2013). Bias in fitting the ETAS model: a case study based on New Zealand seismicity. *Geophysical Journal International,* **192**(1), 390-412.

Helmstetter, A., & Sornette, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research-Solid Earth,* **107**(B10), art. no.-2237.

Hohlschneider, M., Narteau, C., P., S., Peng, Z., & Schorlemmer, D. (2012). Bayesian analysis of the modified Omori law. *Journal of Geophysical Research,* **117**(B06317).

Ishimoto, M., & Iida, K. (1939). Observations of earthquakes registered with the microseismograph constructed recently. *Bull. Earthq. Res. Inst.,* **17**, 443-478.

Marsan, D., & Lengliné, O. (2008). Extending earthquakes' reach through cascading *Science,* **319**, 1076-1079.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of American Statistical Association,* **83**(401), 9-27.

Omori, F. (1894). On aftershocks. *Report of Imperial Earthquake Investigation Committee,* **2**, 103-109.

Pollock, D. (2007). *Aspects of short-term and long-term seismic hazard assessment in New Zealand.* Unpublished Master Thesis, ETH Zurich.

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

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

Stirling, M., McVerry, G., Gerstenberger, M., Litchfield, N., Van Dissen, R., Berryman, K., Barnes, P., Wallace, L., Villamor, P., Langridge, R., Lamarche, G., Nodder, S., Reyners, M., Bradley, B., Rhoades, D., Smith, W., Nicol, A., Pettinga, J., Clark, K., Jacobs, K. (2012). National seismic hazard model for New Zealand: 2010 update. *Bulletin of the Seismological Society of America,* **102**(4), 1514-1542.

Storchak, D. A., Di Giacomo, D., Bondár, I., Engdahl, E. R., Harris, J., Lee, W. H. K., Villaseñor, A., & Bormann, P. (2013). Public Release of the ISC–GEM Global Instrumental Earthquake Catalogue (1900–2009). *Seismological Research Letters,* **84**, 810-815.

USGS. (2013). *Centennial Earthquake Catalog*, 2013, from http://earthquake.usgs.gov/research/data/centennial.php

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

Utsu, T., Ogata, Y., & Matsu'ura, R. S. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *Journal of the Physics of the Earth,* **43**, 1-33.

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area and surface displacement. *Bulletin of the Seismological Society of America,* **84**, 974-1002.

Zhuang, J., Ogata, Y., & Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association,* **97**(458), 369-380.