Testing and development of earthquake forecasting models

David Rhoades  Matthew Gerstenberger
Annemarie Christophersen  Martha Savage
Jiancang Zhuang

GNS Science Consultancy Report 2008/70
May 2008
CONTENTS

EXECUTIVE SUMMARY ........................................................................................................ III
NON-TECHNICAL SUMMARY ............................................................................................ V
1.0  INTRODUCTION ............................................................................................................. 1
2.0  NEW ZEALAND EARTHQUAKE FORECAST TESTING CENTRE ......................... 2
   2.1  Test region and grid specifications .......................................................................... 3
         2.1.1  Catalogue completeness issues .................................................................... 9
         2.1.2  Other catalogue quality issues .................................................................. 14
   2.2  Requirements for model submission ...................................................................... 15
         2.2.1  Five-year models ......................................................................................... 18
         2.2.2  Three-month models .................................................................................... 18
         2.2.3  24-hour models ......................................................................................... 19
   2.3  Tests of model performance ...................................................................................... 19
         2.3.1  N-test ............................................................................................................ 19
         2.3.2  L-test ............................................................................................................. 19
         2.3.3  R-test ............................................................................................................. 19
         2.3.4  Catalogue uncertainties ............................................................................... 20
         2.3.5  Declustering ................................................................................................ 20
         2.3.6  Other considerations .................................................................................... 20
3.0  MODELS INSTALLED IN THE TESTING CENTRE ................................................. 21
   3.1  Five-year models ..................................................................................................... 21
         3.1.1  SUP ............................................................................................................ 21
         3.1.2  PPE ........................................................................................................... 22
         3.1.3  NZNSHM ................................................................................................. 23
   3.2  Three-month models ............................................................................................... 23
         3.2.1  PPE ........................................................................................................... 23
         3.2.2  EEPAS .................................................................................................... 24
   3.3  24-hour models ...................................................................................................... 26
         3.3.1  STEP ....................................................................................................... 26
         3.3.2  ETAS ...................................................................................................... 26
4.0  MODIFIED MERCALLI INTENSITY HAZARD FORECASTS .............................. 32
   4.1  Synthetic catalogue generation .............................................................................. 32
   4.2  Attenuation model and uncertainties .................................................................... 32
   4.3  MM intensity probability maps ............................................................................. 35
5.0  MODEL DEVELOPMENT ............................................................................................. 38
   5.1  Overview ................................................................................................................ 38
   5.2  Time, distance and magnitude dependent foreshock probability model for New Zealand ......................................................................................................................... 39
   5.3  Foreshock rates from aftershock abundance .......................................................... 40
   5.4  Differences between spontaneous and triggered earthquakes: their influence on foreshock probabilities .......................................................... 40
   5.5  Discussion ............................................................................................................... 41
6.0  CONCLUSION .............................................................................................................. 42
7.0  ACKNOWLEDGEMENTS ......................................................................................... 43
8.0  REFERENCES ............................................................................................................... 43

GNS Science Consultancy Report 2008/70
FIGURES

Figure 1
Map showing test region (darkerly shaded) and buffer zone (lightly shaded). The whole shaded region is the data collection region.

Figure 2
Epicentres of earthquakes in the New Zealand catalogue, 1951–2006, with magnitudes M > 2.95 and hypocentral depths h ≤ 45 km.

Figure 3
Cumulative frequency versus magnitude for time-period subsets of the New Zealand Earthquake Catalogue (with depths h ≤ 45 km) within the test region (a) 1951–1963, (b) 1964–1976, (c) 1977–1999, (d) 2000–2006. The Gutenberg-Richter b-value listed for each subset corresponds to the line plotted and is calculated using the minimum magnitude threshold given in parentheses.

Figure 4
Cumulative frequency versus magnitude for time-period subsets of the New Zealand Earthquake Catalogue (with depths h ≤ 45 km) within the buffer region (a) 1951–1963, (b) 1964–1976, (c) 1977–1999, (d) 2000–2006. The Gutenberg-Richter b-value listed for each subset corresponds to the line plotted and is calculated using the minimum magnitude threshold given in parentheses.

Figure 5
(a) Cumulative numbers of earthquakes (depths h ≤ 45 km) within the buffer region exceeding certain magnitude thresholds as a function of time; (b) Ratio of numbers of earthquakes N(M > m + 0.5) / N(M > m) in three year intervals and smooth local fits. The solid line shows the expected value of the ratio under catalogue completeness and a Gutenberg-Richter b-value of 1.

Figure 6
(a) Cumulative numbers of earthquakes (depths h ≤ 45 km) within the buffer region exceeding certain magnitude thresholds as a function of time; (b) Ratio of numbers of earthquakes N(M > m + 0.5) / N(M > m) in three year intervals and smooth local fits. The solid line shows the expected value of the ratio under catalogue completeness and a Gutenberg-Richter b-value of 1.

Figure 7
Frequency versus magnitude plot for earthquakes in the buffer region for hypocentral depths h ≤ 45 km from 2004–2006. The Gutenberg-Richter b-value listed corresponds to the line plotted and is calculated using the minimum magnitude threshold given in parentheses.

Figure 8
Hypocentral depth h versus (a) latitude; (b) longitude; (c) time; for earthquakes with magnitude M > 2.95 in the New Zealand earthquake catalogue 1951–September 2006. (d) Hypocentral depth versus time for M > 3.95.

Figure 9
Histograms of hypocentral depth h for earthquakes with magnitude M > 2.95 in the New Zealand earthquake catalogue, 1951–September 2006. (a) all depths; (b) h ≤ 100 km.

Figure 10
Log_{10} of expected number of earthquakes with magnitude M ≥ 5.0 in a five year period under the NZNSHM model.

Figure 11
Log_{10} of expected number of earthquakes with magnitude M ≥ 5.0 in a three month period (July-September 2006) under the PPE model.

Figure 12
Log_{10} of expected number of earthquakes with magnitude M ≥ 5.0 in a three month period (July-September 2006) under the EEPAS model.

Figure 13
Log_{10} of expected number of earthquakes with magnitude M ≥ 5.0 in a 24-hour period under the STEP model.

Figure 14
Rate of occurrence of MM7 in the 90 day period of a forecast according to the PPE model.

Figure 15
Rate of occurrence of MM7 in the 90 day period of a forecast according to the EEPAS model.

TABLES

Table 1
Polygon defining test region.

Table 2
Polygon defining data-collection region.

Table 3
PPE model parameters used for tests against the whole and declustered catalogues.

Table 4
EEPAS model parameters.

Table 5
ETAS model parameters.

Table 6
Parameter estimates and standard errors from along-strike attenuation model (18) of Dowrick and Rhoades (2005).

Table 7
Correlation matrix for parameters of along-strike attenuation model (18).

Table 8
Parameter estimates and standard errors from strike-normal attenuation model (17) of Dowrick and Rhoades (2005).

Table 9
Correlation matrix for parameters of strike-normal attenuation model.

APPENDICES

Appendix 1
Time, distance and magnitude dependent foreshock probability model for New Zealand.

Appendix 2
Foreshock rates from aftershock abundance.

Appendix 3
Differences between spontaneous and triggered earthquakes: their influences on foreshock probabilities.
EXECUTIVE SUMMARY

The New Zealand Earthquake Forecast Testing Centre, based at GNS Science, is being established as one of several similar regional testing centres under the umbrella of the Collaboratory for the Study of Earthquake Predictability (CSEP). The Centre aims to encourage the development of testable models of time-varying earthquake occurrence in the New Zealand region, and to conduct verifiable prospective tests of their performance over a period of five or more years.

The test region, data-collection region and requirements for testing are described in this report. Models must specify in advance the expected number of earthquakes with epicentral depths $h \leq 40$ km in bins of time, magnitude and location within the test region. Short-term models will be tested using 24-hour time bins at magnitude $M \geq 4$. Intermediate-term models and long-term models will be tested at $M \geq 5$ using three-month and five-year bins, respectively. The tests applied will be the same as at other CSEP testing centres: the so-called N-test of the total number of earthquakes expected over the test period; the L-test of the likelihood of the earthquake catalogue under the model; and the R-test of the ratio of the likelihoods under alternative models.

Three long-term, two intermediate-term and two short-term models have so far been installed in the testing centre, with tests of these models commencing on the New Zealand earthquake catalogue from the beginning of 2008. Submission of models is open to researchers worldwide. New models can be submitted at any time.

The New Zealand testing centre makes extensive use of software produced by the CSEP testing centre in California. It is envisaged that, in time, the scope of the testing centre will be expanded to include new testing methods and differently-specified models, but that the New Zealand testing centre will develop in parallel with other regional testing centres through the CSEP international collaborative process.

Modified Mercalli intensity probability maps can be produced for any model submitted to the New Zealand testing centre, using synthetic catalogues derived from the model, and simulations of a standard MM intensity attenuation model. This allows the hazard indicated by time-varying models to be compared with accepted time-invariant models, such as the national seismic hazard model.

Development of new and modified forecasting models is an important activity. Here we report on several investigations into aspects of short-term models of earthquake clustering, providing insights for future model development.

The first investigation was an analysis of the decay of foreshock activity in time and space, and of the difference in magnitude between foreshocks and mainshocks in New Zealand. It showed that foreshock probabilities decrease with increasing inter-event time according to a modified-Omori law with an exponent $p$ somewhat less than one; that foreshock probabilities decrease with increasing epicentral distance, also following a modified-Omori-type decay; and that the magnitude distribution of mainshocks follows the Gutenberg-Richter relation. Differences in the Omori and Gutenberg-Richter parameters between the TVZ and the rest of New Zealand are consistent with higher attenuation in the TVZ than elsewhere.
The second investigation concerned the relative triggering capability of foreshocks and mainshocks in the southern California earthquake catalogue and the global centroid moment tensor (CMT) catalogue. Several different clustering methods were used, and foreshocks were defined as initiating events of a cluster that were followed by a larger earthquake within a certain time period. It was found that foreshock rates estimated under the assumption that foreshocks have the same triggering capability as mainshocks are in good agreement with the observed rates. For southern California, the probability of an initiating event in an earthquake cluster in the magnitude range 2.0 – 4.5 being followed by a larger event within one day is about 4%. In the global CMT catalogue (with magnitudes ≥ 5), the probability that the initiating earthquake of a cluster is followed by a larger event within 30 days is about 3%.

The third investigation was an analysis of differences between spontaneous and triggered earthquakes, and in particular their influence on foreshock probabilities, using earthquake catalogues from Japan, southern California and New Zealand. A stochastic declustering method based on a space-time epidemic-type aftershock (ETAS) model was used to classify events as "triggered" or "spontaneous". A foreshock here was defined as a spontaneous event that triggered one or more larger descendants. The triggering capability as well as the proportion of foreshocks in the spontaneous earthquakes was found to be lower than the triggering capability and the proportion of foreshocks in triggered events. This appears to be a potential difference between the ETAS model and the real data.

Each of these investigations has been described in detail in a separate manuscript which has been submitted or accepted for publication in a scientific journal. The three manuscripts are appended to this report.
NON-TECHNICAL SUMMARY

The New Zealand Earthquake Forecast Testing Centre, based at GNS Science, is one of several regional testing centres now being established. The aim of the Centre is to encourage the development of forecasting models for New Zealand earthquakes and to conduct tests of their performance over a period of five or more years. It is intended that the New Zealand testing centre will develop in parallel to other regional testing centres through ongoing international collaboration, with sharing of methods and computer software.

The test region, data-collection region and requirements for testing are described in this report. Statistical tests will measure how closely the expected earthquake occurrence under each model conforms to the actual earthquake catalogue. Short-term forecasting models will be tested over a series of 24-hour time-intervals. Intermediate-term models and long-term models will be tested over a series of three-month and five-year time-intervals, respectively. Short-term and intermediate-term models may use only the catalogue of past earthquakes to make their forecasts.

Three long-term, two intermediate-term and two short-term forecasting models have so far been installed in the testing centre, with tests commencing at the beginning of 2008. Submission of models is open to researchers worldwide. New models may be submitted at any time.

For any model submitted to the testing centre, we can calculate the probability of a given level of ground shaking occurring at any location as a consequence of earthquakes expected under the model. This allows the time-varying hazard indicated by each forecasting model to be compared with existing standard models in which the hazard is constant over time.

This project included several studies which may help us to refine short-term forecasting models. These were concerned with different aspects of earthquake clustering and, in particular, foreshock occurrence. A foreshock is an earthquake which is followed soon after by a larger one.

The first study examined the decay of foreshock activity in time and space, and the difference in magnitude between foreshocks and main shocks in New Zealand. It identified clear differences in foreshock occurrence between the Taupo Volcanic Zone and the rest of New Zealand. The second study showed that there is little difference in the tendency of foreshocks and main shocks to trigger other earthquakes in both the Californian earthquake catalogue and a world-wide catalogue. The third study examined differences between the "spontaneous" earthquakes that initiate clusters and the "triggered" earthquakes that follow them. It found that the triggered earthquakes themselves have a greater tendency to trigger other earthquakes, and are more likely to be foreshocks, than the spontaneous earthquakes.
1.0 INTRODUCTION

Learning how to forecast earthquakes is one of the most important problems in seismology. It is important for two reasons. From a scientific perspective, our ability to forecast earthquakes is a measure of our understanding of how earthquakes are generated. From a practical perspective, foreknowledge of an increased hazard of earthquake occurrence in a particular location would be useful for decision-making on the timing of mitigation measures, such as protection and upgrading of building stocks and lifelines networks.

After some years of relative neglect, earthquake forecasting is again becoming a target of geophysicists worldwide. It is now widely recognised that, in order to make progress in this field, there is a need both to develop testable earthquake forecasting models and to conduct verifiable tests of their practical forecasting performance. Internationally, efforts to develop models, agree on testing procedures, and establish testing centres to undertake the performance tests, are gaining momentum (Jordan, 2006; Field, 2007). Broadly speaking, the requirements for a model to be testable are that it must be well-defined, i.e., the forecasts are derived in an unequivocal way from the available data, and capable of generating synoptic estimates of the time-varying rate of earthquake occurrence for any source location and magnitude level within a substantial region of surveillance. Models meeting these requirements are called Regional Earthquake Likelihood Models (RELMs).

A major objective of this study is to establish an earthquake forecast testing centre in the New Zealand region. This includes the specification of the detailed requirements for models to be tested in this centre, including the spatial extent of the test region, the magnitude levels and time periods that will be used, and the grid cells within which forecasts will be made and evaluated. Decisions on such specifications depend on the quality and extent of the New Zealand earthquake catalogue, and the data requirements for models that are presently envisaged for installation in the testing centre. Also to be borne in mind is the maintenance of consistent practices with other similar testing centres, especially the California testing centre of the Collaboratory for the Study of Earthquake Predictibility (CSEP). There are many benefits to be derived from maintaining such consistency across the testing centres in the area of software development costs, which are considerable, especially in light of the level of automation that is needed.

A second objective is to install certain existing models into the testing centre. The authors of this report include developers of some of the existing models, namely the STEP — “Short-Term Earthquake Probability” (Gerstenberger, 2003; Gerstenberger et al., 2005) – and EEPAS — “Every Earthquake a Precursor According to Scale” (Rhoades and Evison, 2005, 2006; Rhoades, 2007) – models. Another existing model is the New Zealand National Seismic Hazard model – NZNSHM – (Stirling et al., 2002), which is already widely used for drawing up earthquake engineering design codes, as well as for many other practical purposes. Although this model is in principle static, rather than time-varying, it is an important reference model to compare models of time-varying earthquake occurrence against. The Epidemic-Type Aftershock Sequence (ETAS) model (Ogata, 1989; 1998) is probably the most widely used short-term earthquake clustering model, and it is desirable to have one or more versions of the space-time ETAS model in the testing centre.
Because it is envisaged that one or more of the models will eventually be used for practical earthquake mitigation measures, it is necessary that the results can be expressed not only in the raw form, as expected numbers of earthquakes in cells of the testing centre grid, but also in a form which is comparable with the existing NZNSHM, i.e., as the probability of a given intensity of shaking being experienced at any location in a future time-period. Therefore in this study, the software to provide the forecasts in this way has been developed.

The development of new forecasting models also is an important objective. A considerable research effort is required to develop a single new model, or even a significant refinement of an existing model. Therefore, a research project such as this can only provide steps-on-the-way towards such development. Several ideas for model improvements are examined in this study, and attempts are made to check them out, as far as it is possible to do so, by retrospective analysis of earthquake catalogues, and simulation of synthetic catalogues from existing models – the ETAS model in particular. The ideas discussed here raise questions as to the detailed nature of short-term earthquake clustering, including foreshocks and aftershocks, and on what scales of time and space earthquake triggering occurs. Knowing the answers will assist in development of the next generation of short-term earthquake forecasting models.

The present studies are limited to using the information in earthquake catalogues to forecast future earthquakes. Other data-bases, such as observations of earth deformation by GPS networks and active fault datasets, are not considered here, although they will undoubtedly be used in future earthquake-forecasting studies, as the length of time covered by the data bases increases.

2.0 NEW ZEALAND EARTHQUAKE FORECAST TESTING CENTRE

The present objective to establish a New Zealand Earthquake Forecast Testing Centre is being carried out with the following goals in mind:

- **To encourage modellers to develop testable time-dependent seismicity forecasting models for New Zealand.** Many studies carry out retrospective analyses of seismicity, but the results and ideas emanating from such studies need to be verified by tests against future seismicity, and in order for that to occur they must first be incorporated into testable models.

- **To establish a testing framework appropriate to New Zealand.** There are similarities and differences between the New Zealand region and other regions where testing centres are being established, notably, at present, California and parts of Europe. The differences have to do with the style of seismicity, and the extent of coverage and history of the earthquake catalogue

- **To re-evaluate the RELM/CSEP likelihood-based testing procedure.** This is a long-term goal. Initially the New Zealand centre is being set up with the same testing procedures as other CSEP testing centres. It is envisaged that re-evaluation of the present procedures will take place through a collaborative process, and that when changed procedures are agreed to, they will be made available to all regional testing centres using common software.
• **To investigate other testing methodologies including ground-motion-based testing.** The first generation of testing is for regional earthquake likelihood models, which estimate the expected number of earthquakes in any given window of time, magnitude and location. The expected number of exceedances of a given level of ground motion at any location in a given window of time is also a quantity of interest, and indeed is the primary quantity of interest in the national seismic hazard model. A long-term goal is to extend the testing to ground-motion models.

• **To test multiple forecast models developed for New Zealand in a 5+ year prospective test.** Robust tests require a large number of earthquakes. To obtain a large-enough number of significant earthquakes to test the models against, a period of at least five years will be necessary. The number of test earthquakes can be increased by lowering the magnitude threshold for targeted events, but in practice any magnitude less than about 5 has not much impact on the ground-motion hazard. It is therefore much more important to forecast the larger earthquakes than small-to-moderate sized events. Also, since it is not clear that the earthquake process is entirely self-similar, an ability to forecast small earthquakes is not equivalent to an ability to forecast large ones. Therefore it may be unhelpful to lower the magnitude threshold too much. In any case, testing of forecasting models is likely to remain an important ongoing activity.

• **To test the impact of individual assumptions within models.** The effect of individual assumptions on model performance is not always easy to determine from retrospective studies. For example, a more complex model will always fit better to existing data than a simpler one, but this does not mean that it will perform better against future earthquakes. Also, the performance of models on a discrete test grid of time, magnitude and location cells is not the same as its performance measured on continuous scales of time, magnitude and location. For model development, it is often more computationally efficient to measure performance on continuous scales. The impact of individual assumptions is not necessarily the same when assessed on a discrete grid. It is desirable to make the testing centre software available to researchers developing models, so that they can anticipate the effects of the test grid on model performance, and if necessary, adjust their models accordingly before submitting them for testing against future earthquakes.

• **To maintain a strong relationship with CSEP.** A strong international research community with an interest in evaluating the predictability of earthquakes is now developing within the CSEP framework. It is important that the New Zealand centre can benefit from, and contribute to, the combined knowledge of this research community, as well as the specific software products developed by the CSEP community to facilitate testing.

### 2.1 Test region and grid specifications

Following extensive consultations between the participants and potential participants, including informal meetings of the Wellington-based statistical seismology group, specifications for the test region, and the spatial and magnitude grids were drawn up. Important considerations were that while the quality and completeness of the earthquake catalogue is generally good at for earthquake locations inside or close to the edges of the New Zealand Seismograph Network, i.e., for onshore locations, this quality and completeness can be expected to deteriorate quite rapidly for offshore locations.
Boundaries of the test region are shown in Figure 1, and vertices of the polygon defining the test region are listed in Table 1. The test region covers the New Zealand land area plus a region extending about 50 km offshore. Figure 1 also shows the data-collection region, and the vertices of the polygon defining this region are listed in Table 2. The data-collection region extends about 50 km in all directions beyond the edge of the test region.

The location grid consists of cells of area 0.1 degree squared centred on 1/10th degree coordinates of latitude and longitude which have their centres within the test region, e.g. (-41.5±0.05, 174.5±0.05).

Figure 2 is a map of shallow earthquake epicentres in the New Zealand region. By comparing Figures 1 and 2, it can be seen that many earthquakes occur outside the test and data-collection regions. However, these regions were chosen for reasons of catalogue completeness and quality as mentioned above and discussed in more detail below.
Figure 1  Map showing test region (darker shaded) and buffer zone (lighter shaded). The whole shaded region is the data collection region.
Table 1  Polygon defining test region.

<table>
<thead>
<tr>
<th>Latitude</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>-46.1520</td>
<td>165.8468</td>
<td>-39.7427</td>
<td>177.8053</td>
</tr>
<tr>
<td>-45.4971</td>
<td>165.7431</td>
<td>-40.1637</td>
<td>177.5288</td>
</tr>
<tr>
<td>-44.0936</td>
<td>167.3329</td>
<td>-41.1462</td>
<td>176.8721</td>
</tr>
<tr>
<td>-38.8070</td>
<td>173.4505</td>
<td>-41.5673</td>
<td>176.4574</td>
</tr>
<tr>
<td>-37.6374</td>
<td>174.1071</td>
<td>-41.9415</td>
<td>175.8698</td>
</tr>
<tr>
<td>-36.9825</td>
<td>173.8652</td>
<td>-42.0819</td>
<td>175.3514</td>
</tr>
<tr>
<td>-35.4386</td>
<td>172.5173</td>
<td>-42.0351</td>
<td>174.9712</td>
</tr>
<tr>
<td>-35.1579</td>
<td>172.5173</td>
<td>-42.5029</td>
<td>174.4873</td>
</tr>
<tr>
<td>-34.6901</td>
<td>172.2062</td>
<td>-43.1579</td>
<td>173.8652</td>
</tr>
<tr>
<td>-34.5029</td>
<td>172.1371</td>
<td>-43.3918</td>
<td>173.7270</td>
</tr>
<tr>
<td>-34.0819</td>
<td>172.3790</td>
<td>-43.7661</td>
<td>173.7615</td>
</tr>
<tr>
<td>-33.8947</td>
<td>172.3790</td>
<td>-44.2807</td>
<td>173.4159</td>
</tr>
<tr>
<td>-33.8947</td>
<td>173.1048</td>
<td>-44.4211</td>
<td>172.9320</td>
</tr>
<tr>
<td>-34.2222</td>
<td>173.5887</td>
<td>-44.3275</td>
<td>172.6210</td>
</tr>
<tr>
<td>-34.5965</td>
<td>174.2454</td>
<td>-44.5146</td>
<td>171.9297</td>
</tr>
<tr>
<td>-35.5322</td>
<td>175.2823</td>
<td>-44.8421</td>
<td>171.9297</td>
</tr>
<tr>
<td>-35.6257</td>
<td>175.6624</td>
<td>-45.1696</td>
<td>171.7569</td>
</tr>
<tr>
<td>-36.0936</td>
<td>176.1118</td>
<td>-45.3567</td>
<td>171.6532</td>
</tr>
<tr>
<td>-36.5146</td>
<td>176.3537</td>
<td>-45.7310</td>
<td>171.5495</td>
</tr>
<tr>
<td>-37.1228</td>
<td>176.4919</td>
<td>-46.0117</td>
<td>171.4804</td>
</tr>
<tr>
<td>-37.2632</td>
<td>176.6302</td>
<td>-46.3392</td>
<td>171.1002</td>
</tr>
<tr>
<td>-37.4035</td>
<td>177.0449</td>
<td>-46.8070</td>
<td>170.3399</td>
</tr>
<tr>
<td>-37.0292</td>
<td>178.1509</td>
<td>-47.0877</td>
<td>169.6486</td>
</tr>
<tr>
<td>-37.3567</td>
<td>178.9804</td>
<td>-47.1345</td>
<td>168.9574</td>
</tr>
<tr>
<td>-37.5906</td>
<td>179.1878</td>
<td>-47.5088</td>
<td>168.5081</td>
</tr>
<tr>
<td>-38.2456</td>
<td>179.0150</td>
<td>-47.7427</td>
<td>167.9896</td>
</tr>
<tr>
<td>-38.7602</td>
<td>178.8767</td>
<td>-47.7895</td>
<td>167.4366</td>
</tr>
<tr>
<td>-38.9942</td>
<td>178.6002</td>
<td>-47.6023</td>
<td>166.9873</td>
</tr>
<tr>
<td>-39.3684</td>
<td>178.5657</td>
<td>-46.1520</td>
<td>165.7431</td>
</tr>
<tr>
<td>-39.6959</td>
<td>178.3237</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2  

<table>
<thead>
<tr>
<th>Latitude</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>-46.2368</td>
<td>165.1210</td>
<td>-39.8158</td>
<td>179.0841</td>
</tr>
<tr>
<td>-45.9211</td>
<td>165.0173</td>
<td>-40.0789</td>
<td>178.6694</td>
</tr>
<tr>
<td>-45.3947</td>
<td>165.1901</td>
<td>-40.2368</td>
<td>178.2546</td>
</tr>
<tr>
<td>-44.9737</td>
<td>165.5012</td>
<td>-40.8684</td>
<td>177.8399</td>
</tr>
<tr>
<td>-41.7632</td>
<td>169.1993</td>
<td>-41.6053</td>
<td>177.3214</td>
</tr>
<tr>
<td>-38.4474</td>
<td>172.9320</td>
<td>-42.0263</td>
<td>176.8721</td>
</tr>
<tr>
<td>-37.5000</td>
<td>173.3468</td>
<td>-42.5000</td>
<td>175.9389</td>
</tr>
<tr>
<td>-37.1842</td>
<td>173.3122</td>
<td>-42.6579</td>
<td>175.1440</td>
</tr>
<tr>
<td>-35.7105</td>
<td>172.0680</td>
<td>-43.2368</td>
<td>174.6601</td>
</tr>
<tr>
<td>-35.1842</td>
<td>171.8952</td>
<td>-43.5526</td>
<td>174.4528</td>
</tr>
<tr>
<td>-35.1316</td>
<td>171.7224</td>
<td>-44.0789</td>
<td>174.4528</td>
</tr>
<tr>
<td>-34.5000</td>
<td>171.4804</td>
<td>-44.7105</td>
<td>173.9889</td>
</tr>
<tr>
<td>-33.9211</td>
<td>171.8878</td>
<td>-44.8684</td>
<td>173.1048</td>
</tr>
<tr>
<td>-33.5526</td>
<td>172.1717</td>
<td>-44.9211</td>
<td>172.6210</td>
</tr>
<tr>
<td>-33.4474</td>
<td>172.6210</td>
<td>-45.3421</td>
<td>172.4827</td>
</tr>
<tr>
<td>-33.4474</td>
<td>173.2776</td>
<td>-45.6579</td>
<td>172.2408</td>
</tr>
<tr>
<td>-33.8158</td>
<td>173.9343</td>
<td>-45.9211</td>
<td>172.2062</td>
</tr>
<tr>
<td>-34.4474</td>
<td>174.9712</td>
<td>-46.4474</td>
<td>172.0334</td>
</tr>
<tr>
<td>-35.1842</td>
<td>175.5933</td>
<td>-46.7632</td>
<td>171.5150</td>
</tr>
<tr>
<td>-35.5000</td>
<td>176.3882</td>
<td>-46.8158</td>
<td>171.3767</td>
</tr>
<tr>
<td>-36.0263</td>
<td>176.7339</td>
<td>-47.3947</td>
<td>170.5127</td>
</tr>
<tr>
<td>-36.2895</td>
<td>176.8721</td>
<td>-47.5526</td>
<td>169.8560</td>
</tr>
<tr>
<td>-36.8684</td>
<td>177.1141</td>
<td>-47.6053</td>
<td>169.5449</td>
</tr>
<tr>
<td>-36.8053</td>
<td>177.9781</td>
<td>-47.9737</td>
<td>168.9228</td>
</tr>
<tr>
<td>-36.7632</td>
<td>179.0150</td>
<td>-48.1842</td>
<td>168.0933</td>
</tr>
<tr>
<td>-37.1842</td>
<td>179.6717</td>
<td>-48.2895</td>
<td>167.3675</td>
</tr>
<tr>
<td>-37.8684</td>
<td>179.7753</td>
<td>-48.1316</td>
<td>166.6417</td>
</tr>
<tr>
<td>-38.9737</td>
<td>179.4988</td>
<td>-47.5000</td>
<td>165.9850</td>
</tr>
<tr>
<td>-39.1316</td>
<td>179.3606</td>
<td>-46.2368</td>
<td>165.1210</td>
</tr>
<tr>
<td>-39.6053</td>
<td>179.1532</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2  Epicentres of earthquakes in the New Zealand catalogue, 1951 – 2006, with magnitudes $M > 2.95$ and hypocentral depths $h \leq 45$ km
2.1.1 Catalogue completeness issues

Catalogue completeness is an important issue to consider when specifying the test and data-collection regions. Broadly speaking we can have regard to the following approximate eras of the New Zealand earthquake catalogue when assessing the change of catalogue completeness with time. (a) Pre 1964 (b) 1964 to 1985 (c) 1986 to 1999 (d) 2000 on. Although the changes to the seismograph network have taken place gradually over periods of time rather than instantaneously, the years 1964, 1987 and 2000 are the approximate dates of major upgrades of the New Zealand Seismograph Network, the most recent being the transition to the present GeoNet broadband network. We examine the frequency-magnitude distribution of the earthquakes with local magnitude $M_L \geq 3.0$ in each of these eras within the test region and within the buffer region (Figure 1).

Figure 3 shows frequency-magnitude plots within the test region. For the 1951-1963 period (Figure 3a), the plot is approximately linear for magnitudes above about 4.2. The deviation from linearity at lower magnitudes is clear evidence of incompleteness up to magnitude 4.0. For the 1964-1976 period (Figure 3b), the deviation of the plot from linearity suggests a threshold magnitude of completeness slightly below 4.0. On the other hand, the linearity of the plots in Figure 3c and 3d suggest that the catalogue is complete, or near complete, for all magnitudes above 3.0 since 1987.

Figure 4 shows corresponding frequency-magnitude plots within the buffer region. For the 1951-1963 period (Figure 4a), the magnitude threshold of completeness appears to be about 4.8; for the 1964-1976 period (Figure 4b), it is about 4.2; for the 1987-1989 period (Figure 4c), it is about 3.9; and for 2000-2006 (Figure 4d), it is about 3.4. Therefore, in all time periods the catalogue is not as complete in the buffer region as in the test region.

We further examine the change in completeness of the catalogue with time in Figures 5 and 6.

The numbers of earthquakes in the test region exceeding certain magnitude thresholds, accumulated with time, are plotted in Figure 5a. This shows that there has been a gradual increase in the rate of accumulation for $M > 2.95$ between 1951 and 2006. The step-like increases are most likely associated with large multiple-earthquake sequences such as the Inangahua aftershock sequence in 1968. Figure 5b can be used to judge the variation of magnitude completeness with time. In this figure, ratios of the number of the earthquakes exceeding certain magnitude thresholds have been plotted. Let $N(M > m)$ be the number of earthquakes exceeding magnitude $m$ in a time interval. Under the assumption of catalogue completeness and a Gutenberg-Richter b-value of about 1, the expected value of the ratio $N(M > m + 0.5) / N(M > m)$ is 0.32, shown by the horizontal line in Figure 5b. The time when this ratio drops to this level, as shown by the points plotted for three-year intervals and the associated smooth trend lines, is an indication of the approximate time when the catalogue became complete for magnitude $M > m$.

Based on Figure 5b, it appears that the catalogue in the test region has been complete at $M > 4.45$ since at least 1951; at $M > 3.95$ since about 1960; at $M > 3.45$ since about 1980; and at $M > 2.95$ since the late 1980s.
Figure 3  Cumulative frequency versus magnitude for time-period subsets of the New Zealand Earthquake Catalogue (with depths h ≤ 45 km) within the test region (a) 1951 - 1983, (b) 1984 - 1976, (c) 1987 - 1999, (d) 2000 - 2006. The Gutenberg-Richter b-value listed in each plot corresponds to the line plotted and is calculated using the minimum magnitude threshold given in parentheses.
Figure 4  Cumulative frequency versus magnitude for time-period subsets of the New Zealand Earthquake Catalogue (with depths $h \leq 45$ km) within the buffer region (a) 1951 - 1963, (b) 1964 - 1976, (c) 1987 - 1999, (d) 2000 - 2006. The Gutenberg-Richter b-value listed in each plot corresponds to the line plotted and is calculated using the minimum magnitude threshold given in parentheses.
Figure 5  (a) Cumulative numbers of earthquakes (depths $h \leq 45$ km) within the buffer region exceeding certain magnitude thresholds as a function of time; (b) Ratio of numbers of earthquakes $N(M > m + 0.5) / N(M > m)$ in three year intervals and smooth local fits. The solid line shows the expected value of the ratio under catalogue completeness and a Gutenberg-Richter b-value of 1.
Figure 6  (a) Cumulative numbers of earthquakes (depths $h \leq 45$ km) within the buffer region exceeding certain magnitude thresholds as a function of time; (b) Ratio of numbers of earthquakes $N(M > m + 0.5) / N(M > m)$ in three year intervals and smooth local fits. The solid line shows the expected value of the ratio under catalogue completeness and a Gutenberg-Richter $b$-value of 1.
Figure 6 shows the corresponding analysis for the buffer region. Compared to the test region, the increase over time in the rate of accumulation of earthquakes in the lower magnitude bands is much stronger. An extraordinary feature of Figure 6a is the huge step-up in the cumulative number of earthquakes in all magnitude classes in 1995. This corresponds to the time of the aftershocks of the 1995 Feb 5 M7.0 East Cape earthquake, many of which occurred in the buffer region.

Based on Figure 6b, it appears that the catalogue in the buffer region has been nearly complete at M > 4.45 since 1951; at M > 3.95 since about 1960; at M > 3.45 since about 1995; and perhaps at M > 2.95 since about 2003. However Figure 7 displays the frequency magnitude relation in the buffer zone for earthquakes from 2004-2006, and it shows that the threshold of completeness is still no lower than 3.4 for this period. There appears to be a deficit of at least 250 earthquakes, or about 25% of all earthquakes with M > 2.95 in this period. Unless further improvements to the network occur during the testing period, a similar deficit is likely to apply in that period also. This will have an effect on the performance of models which depend on small earthquakes in the buffer region to estimate earthquake occurrence in the test region. This is an effect that needs to be considered in the preparation of models for testing.

2.1.2 Other catalogue quality issues

Figure 8 displays scatter plots of hypocentral depth h against latitude, longitude and time. It shows that New Zealand earthquakes occur over a wide range of depths. Hypocentral depths up to 300 km are common in the catalogue, and the deepest recorded earthquakes are at about 600 km. The deep earthquakes are mostly associated with the Hikurangi and Fiordland subduction zones. For the shallower earthquakes (h < ∼ 40 km), the depth is not actually estimated, but rather a depth-restricted solution for the earthquake source is given in the catalogue. Common depth restrictions are h = 5, 12 and 33 km. Before the 1980s, most shallow earthquake solutions were depth-restricted to either 12 or 33 km. Figure 8c shows that the number on non depth-restricted solutions increased dramatically in about 1987, but Figure 6d shows that the depth-restricted solutions are a significant proportion of the shallow earthquakes right up to the most recent recordings in the catalogue used here, i.e. September 2008.

Figure 9 displays histograms of hypocentral depths. Figure 9a shows that more than half of all earthquakes have h ≤ 50 km. Actually about 52% have h ≤ 40 km, and only 2% have 40 < h ≤ 50. Figure 9b shows a sharp change in rate of earthquake occurrence versus depth at about h = 35 km. Note that the three apparent peaks in frequency versus depth correspond to the three common depth restrictions applied to shallow earthquakes, at 5, 12, and 33 km, and should not be taken as evidence that earthquakes occur more frequently in the 0 – 5, 10 – 15, and 30 – 35 km classes than in other shallow depth classes.

It has been agreed by the participants of the testing centre that the test region will include earthquakes with hypocentral depths up to 40 km, and that no differentiation of depths will be attempted in the 0 – 40 km range. This decision has been made in the light of the present quality of the catalogue and the present state of modelling of earthquake occurrence. It does not rule out the possibility that at some future time a wider depth range could be included and models which discriminate depth within the 0 – 40 km range could be tested.
2004 - 2006

![Graph showing frequency versus magnitude for earthquakes in the buffer region for hypocentral depths h ≤ 45 km from 2004 - September 2006. The Gutenberg-Richter b-value listed corresponds to the line plotted and is calculated using the minimum magnitude threshold given in parentheses.]

Figure 7  Frequency versus magnitude plot for earthquakes in the buffer region for hypocentral depths h ≤ 45 km from 2004 – September 2006. The Gutenberg-Richter b-value listed corresponds to the line plotted and is calculated using the minimum magnitude threshold given in parentheses.

2.2 Requirements for model submission

Testing for five-year, three-month and 24-hour models will be carried out on the earthquake catalogue beginning from 1 January 2008, based on the expected number of earthquakes with hypocentral depths h ≤ 40 km in bins defined by a grid of magnitude values at 0.1 interval spacing, and a grid of latitude and longitude coordinates at 0.1 degree spacing, with centres inside the testing-region polygon (Table 1 and Figure 1). Models in all classes are invited from the international scientific community. New models may be submitted at any time. The installation of new 3-month and 24-hour models requires substantial work to be done by the testing centre staff.
Figure 8  Hypocentral depth $h$ versus (a) latitude; (b) longitude; (c) time; for earthquakes with magnitude $M > 2.95$ in the New Zealand earthquake catalogue 1951 – September 2006. (d) Hypocentral depth versus time for $M > 3.95$. 
Figure 9  Histograms of hypocentral depth $h$ for earthquakes with magnitude $M > 2.95$ in the New Zealand earthquake catalogue, 1951 – September 2006. (a) all depths; (b) $h \leq 100$ km.
2.2.1 Five-year models

The five-year test is designed for time-invariant or quasi time-invariant models. Modellers must supply a file specifying the expected number of earthquakes over a 5-year period in each location and magnitude bin. Magnitude bins are centred on values from 5.0 up to 9.0 in steps of 0.1. The format of a single line of data in the file is as follows:

\[ \text{<minimum longitude> <maximum longitude> <minimum latitude> <maximum latitude>} \]
\[ \text{<minimum depth> <maximum depth> <minimum magnitude> <maximum magnitude>} \]
\[ \text{<expected number of earthquakes> <bit mask>} \]

An example line is:

174.55 174.65 -41.55 -41.45 0 40 4.95 5.05 0.00412 1

Although only one depth bin is being used at present, the format allows for extension to multiple depth bins in the future.

The purpose of the bit mask is to accommodate models that do not apply to the whole test region. The value of <bit mask> is either 1 or 0. A value of 1 indicates that the model's estimate for the present bin should be used. A value of 0 indicates that the model's estimate for the present bin should be ignored. A modeller may specify that values from another model be used for bins with bit mask 0.

2.2.2 Three-month models

The three-month test is for intermediate-term forecasting models which use the past earthquake catalogue to forecast the earthquake occurrence over the next three-month period. Modellers must supply a computer program which accepts the past earthquake catalogue in the data-collection region (in a format supplied by the testing centre) as input, and outputs a file specifying the expected number of earthquakes over a 3-month period in each location and magnitude bin. The format of the file should be as described in the section above. The program must be written in such a way that the testing centre can control the input files and specify the period for which the forecast is being made.

The catalogue supplied by the testing centre will be a search, within the data collection region, of the online New Zealand earthquake catalogue produced by GeoNet (http://www.geonet.org.nz/earthquake/resources/index.html), and will have one line of free format real numbers for each earthquake. The first ten numbers will give the source parameters of the earthquake in the order: longitude, latitude, time (year and decimal date), month, day, magnitude, depth, hour, minute, second. An example line is:

1.1745500e+02 -4.1550000e+01 1.9910075e+03 1.00000000e+00 3.00000000e+00 3.00000000e+00 6.00000000e+00 1.70000000e+01 5.80000000e+01 1.00100000e+01 .....  

Three month models may be supplied with the finalised catalogue from the beginning of 1951 to 50 days before the start of each three-month test period.
The three month models will be tested over a succession of three-month intervals up to five
years or the total period of the tests.

2.2.3 24-hour models

The 24-hour test is for short-term forecasting models which use the past earthquake
catalogue to forecast the earthquake occurrence over the next 24-hour period. The minimum
magnitude bin for 24-hour models is centred on magnitude 4.0. Otherwise, the test region is
the same as for the five-year and three-month models. Modellers must supply a computer
program which accepts the past earthquake catalogue in the data collection region (in the
format supplied by the testing centre) as input, and outputs a file specifying the expected
number of earthquakes over a 24-hour period in each location and magnitude bin. The
format of the file should be as described in the section above. The program must be written
in such a way that the testing centre can control the input files and specify the period for
which the forecast is being made.

The 24-hour models may be supplied with the finalised catalogue from the beginning of 1951
up to just before the start of each 24-hour test period.

The 24-hour models will be tested over a succession of daily intervals up to five years or the
total period of the tests.

2.3 Tests of model performance

The tests of model performance will initially be the same as those carried out in the CSEP
testing centre in California. These tests have been described in detail by Schorlemmer et al.
(2007), and so are only briefly reviewed here. The tests treat the cell expected values as
means of independent Poisson-valued random variables.

2.3.1 N-test

The N-test compares the total number of earthquakes expected under the model with the
actual number. The N-test will reject a model if the total number of earthquakes occurring
during the test period is inconsistent with a Poisson random variable with mean $N$, where $N$
is the total expected number of earthquakes under the model.

2.3.2 L-test

The L-test compares the likelihood of the actual earthquake catalogue, i.e., the number of
earthquakes occurring in each bin, with the distribution of likelihoods of synthetic catalogues
conforming to the model. The model is rejected if the likelihood of the actual catalogue lies
outside the distribution of likelihoods of the synthetic catalogues conforming to the model.

2.3.3 R-test

The R-test compares the likelihoods of alternative models on the actual data. It tests the
statistical significance of any differences by comparing the observed difference with what
would be expected if each model, in turn, were the correct one. In order to do this, it
generates synthetic earthquake catalogues consistent with each model in turn, and evaluates
the likelihood for each model using its own and each other model's set of synthetic catalogues.

In the R-test, each model is regarded in turn as the null hypothesis $H_0$ to be compared against alternatives $H_A$. Suppose $L(\Omega \mid H)$ denotes the log likelihood of the actual earthquake catalogue under hypothesis $H$. The R-statistic is $L(\Omega \mid H_A) - L(\Omega \mid H_0)$. A high value of $R$ is favourable to $H_A$. Model $H_0$ can be rejected by $H_A$ if $R$ is statistically large compared to the distribution of R-statistics computed using synthetic catalogues consistent with $H_0$.

2.3.4 Catalogue uncertainties

The methodology described in Schorlemmer et al. (2007) allows for uncertainties in the published catalogue to be allowed for in the evaluation of model performance. This includes, in principal, both magnitude uncertainties and epicentral location uncertainties. Initially, it is not proposed to specifically allow for such uncertainties in the New Zealand centre testing. However, this is a refinement that could be included later, with the agreement of the participants. Any allowance for uncertainties in the catalogue values has to be made with care, because treatment of uncertainties inevitably involves using a model to generate these uncertainties. Such a model might not be agreed to by all participants. At the very least, any model that is being used to generate catalogue uncertainties must be well understood by all participants.

2.3.5 Declustering

Some long-term models are designed to forecast only main shocks, and not aftershocks. For such models, tests will be run against a declustered catalogue. The method of Reasenberg (1985), with the default parameter values, will be used to decluster the earthquake catalogue. No declustering will be carried out on the catalogues used for the three-month and 24-hour hour tests.

2.3.6 Other considerations

The test calculations will be carried out using software developed at the California CSEP testing centre run by the Southern California Earthquake Center. Updates of this software are expected to be released from time to time, and these will normally be incorporated in the New Zealand testing Centre also.

If we were to try to define the best performing model, this model would, in the words of Schorlemmer et al. (2007), "never be rejected by an R-test and would show data consistency in the N- and L-tests". For technical reasons, if for no other reason, it is unlikely that any model will perform to this ideal standard. Even if a model provides a perfect description of earthquake occurrence, the limited updating of forecasts allowed for under the CSEP testing will prevent it from making perfect forecasts. For example, when a major earthquake occurs, the short-term forecasting models should in theory be instantly updated to estimate the subsequent aftershocks. In practice, no update of expected seismicity can be made until the end of the next 24-hour forecasting period. The practical effect of this restriction on updating is that the number of earthquakes in a cell will not conform to either of the assumptions
underlying the tests, i.e., deviations from the expected values will be neither Poisson-distributed nor independent between cells.

Therefore, it will be necessary to carry out tests other than those already incorporated in the testing centre software. In particular, we would like to develop tests that emphasize measurement of the information value of the different models, and identify sub-regions where particular models are more or less informative than others. The likely process for acceptance of new tests is that they will be promoted through the CSEP collaborative process, and incorporated in software made available to all regional testing centres.

A web page http://nz.cseptesting.org/ has been established for the New Zealand testing centre. This will be used to disseminate and update information about the centre.

Participants submitting models to the New Zealand earthquake forecasting testing centre retain all existing rights to their own models. The centre claims no right to use the models except for testing purposes.

3.0 MODELS INSTALLED IN THE TESTING CENTRE

Models installed in the testing centre to date are described below.

3.1 Five-year models

3.1.1 SUP

The SUP (Stationary Uniform Poisson) model is included as a reference model of least information in the five-year tests. This time-invariant model is also available as a reference model for testing over any other time period. It is not a realistic model of seismicity because it does not incorporate either temporal or spatial variation of the rate of earthquake occurrence. The assumption of this model is that seismicity is stationary and spatially homogeneous over the whole test region. The rate of earthquake occurrence for $M > 4.95$ is equal to the average rate of occurrence of such earthquakes over the time period 1965 – 2006. Earthquakes are assumed to follow the Gutenberg-Richter frequency magnitude relation

$$N(m) = N(m_o)10^{-(m-m_o)}$$  \hspace{1cm} (1)

where $N(m)$ is the expected number of earthquakes exceeding magnitude $m_o$, and $b$ is the Gutenberg-Richter $b$-value estimated from earthquakes exceeding magnitude $m_o = 4.95$ in the test region, with an area of about 578,760 km$^2$ over the period from January 1965 to September 2006. There are 253 such earthquakes, and the maximum likelihood estimate of their $b$-value is 1.16.

Thus the rate of occurrence of earthquakes with magnitude $m > 4.95$ in the SUP model is $2.867 \times 10^{-8}$ earthquakes per day per km$^2$. The expected number of earthquakes in cells corresponding to the same magnitude bin differs only because the area of such cells varies with latitude.
3.1.2 PPE

The PPE (Proximity to Past Earthquakes) model is closely based on a model proposed by Jackson and Kagan (1999). It is a type of smoothed seismicity model, in which earthquakes are forecast to occur close to the epicentres of past earthquakes above the magnitude threshold \( m_c \) for the test region. The PPE model was described in detail by Rhoades and Evison (2004).

In the PPE model, the rate density of earthquake occurrence \( \lambda_0 \) at time \( t \), magnitude \( m \), and epicentral location \( (x, y) \) is of the form

\[
\lambda_0(t, m, x, y) = f_0(t)g_0(m)h_0(t, x, y),
\]

where

\[
f_0(t) = \frac{1}{t - t_0},
\]

\( g_0(m) \) is a magnitude density derived from the Gutenberg-Richter relation (1), and \( h_0(t,x,y) \) is the sum, over earthquakes with magnitudes \( m_i \geq m_c \) from the start of the catalogue at time \( t_0 \) up to, but not including, time \( t \), of smoothing kernels of the form

\[
h_{0i}(r_i) = a(m - m_c)\frac{1}{\pi}\left(\frac{1}{d^2 + r_i^2}\right) + s.
\]

In equation (4), \( r_i \) is the distance in km between \( (x,y) \) and the epicentre \( (x_i, y_i) \) of the \( i \)th earthquake in the catalogue with \( m_i \geq m_c \) and \( a, d, \) and \( s \) are constant parameters. The optimal values of the parameters \( a, d, \) and \( s \) over the period 1987-2006, using earthquakes in the data-collection region from 1951 – 2006 are given in Table 3.

Table 3 PPE model parameters used for tests against the whole and declustered catalogues.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Whole catalogue</th>
<th>Declustered catalogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>0.346</td>
<td>0.286</td>
</tr>
<tr>
<td>( d )</td>
<td>3.02</td>
<td>7.31</td>
</tr>
<tr>
<td>( s )</td>
<td>( 8.095 \times 10^{-12} )</td>
<td>( 1.394 \times 10^{-11} )</td>
</tr>
</tbody>
</table>
3.1.3 NZNSHM

The NZNSHM (New Zealand National Seismic Hazard Model) is derived from the work of Stirling et al. (2002). It is a model developed using modern probabilistic seismic hazard analysis techniques and consists of earthquake sources of two types: 1) fault sources; and 2) distributed seismicity sources.

The fault source model consists of more than 300 faults where the following parameters are specified for each fault: fault type (eg., normal or strike-slip), maximum and minimum depth, single event displacement, maximum magnitude, and recurrence interval. These parameters are defined through a combination of field work, modelling and expert judgement.

The distributed seismicity sources are based on a smoothed representation of the historical catalogue of earthquakes in New Zealand from 1840 to the present. The average number of earthquakes is defined at point sources that are spaced at 0.1 degrees across New Zealand and distributed at depths of 10km, 30km, 50km, 70km and 90km. In order to smooth the sources, a Gaussian smoothing kernel with a smoothing distance of 50km is applied to the raw earthquake data. A minimum magnitude of 5.25 is used across all point sources, while the maximum magnitude is defined based on seismo-tectonic zones that are defined for the country. The Gutenberg-Richter (Gutenberg & Richter, 1954) b-value is calculated for each zone, based on the earthquakes occurring within it.

The NZNSHM installed in the testing centre differs from the above model in that it is an entirely gridded representation of the model; in other words, the rates of earthquakes applied to a single fault in the original model have been applied to one or more grid cells. This is done to meet the grid based testing requirements of the tests applied within the centre. The faults are transformed to grid cells by projecting the faults to the surface and evenly distributing the fault based event rate to all grid cells through which the fault passes. As each grid cell represents a 0.1 degree by 0.1 degree area, the rates can be appropriately proportioned based upon this.

Figure 10 shows a five-year forecast of earthquakes with magnitude \( M \geq 5.0 \) under the NZNSHM model.

3.2 Three-month models

3.2.1 PPE

The PPE model is submitted for testing as a three-month model as well as a five-year model. Its role in the three-month tests is mainly as a reference model which is spatially varying but quasi time invariant. The only time varying element in the PPE model is due to the augmentation, at three-month intervals, of the earthquake data-base to include the most recent earthquakes. Because of this updating, and because of the general tendency of earthquakes to cluster in both time and location, it is expected that the PPE model may perform slightly better in the three-month testing than in the five-year testing.

The parameters adopted for three-month testing are the same as for five-year testing on the whole catalogue (Table 3).

Figure 11 shows a three-month forecast of earthquakes with magnitude \( M \geq 5.0 \) under the PPE model.
3.2.2 EEPAS

The EEPAS (Every Earthquake a Precursor According to Scale) model is a method of forecasting based on the notion that the precursory scale increase phenomenon (Evison and Rhoades, 2002; 2004) occurs at all scales in the seismogenic process. In past studies the EEPAS model has been applied to New Zealand and California at a magnitude threshold $m_c = 5.75$, to Japan with $m_c = 6.75$ and 6.25, to the Kanto region in central Japan with $m_c = 4.75$, to Greece with $m_c = 5.95$; and to southern California with $m_c = 4.95$ (Rhoades and Evison, 2004, 2005, 2006; Console et al., 2006; Rhoades, 2007).

In the EEPAS model, each earthquake $(t_i, m_c, x_i, y_i)$ contributes a transient increment $\lambda_i(t, m, x, y)$ to the future rate density in its vicinity, given by

$$\lambda_i(t, m, x, y) = w_i f_{i}(t) g_{i}(m) h_{i}(x, y) \quad (5)$$

where $w_i$ is a weighting factor that may depend on other earthquakes in the vicinity, and $f_{i}$, $g_{i}$ and $h_{i}$ are densities of the probability distributions for time, magnitude and location, respectively. The magnitude density $g_{i}$ is of the form

$$g_{i}(m) = \frac{1}{\sigma_m \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{m - a_m - b_m m_i}{\sigma_m} \right)^2 \right] \quad (6)$$

where $a_m$, $b_m$ and $\sigma_m$ are parameters. The time density $f_{i}$ is of the form

$$f_{i}(t) = \frac{H(t-t_i)}{(t-t_i) \sigma_T \ln(10) \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\log(t-t_i) - a_T - b_T m_i}{\sigma_T} \right)^2 \right] \quad (7)$$

where $H(s) = 1$ if $s > 0$ and 0 otherwise, and $a_T$, $b_T$ and $\sigma_T$ are parameters. The location density $h_{i}$ is of the form

$$h_{i}(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sigma_m^{10}} \exp \left[ -\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma_x^2 \sigma_y^2} \right] \quad (8)$$

where $\sigma_x$ and $\sigma_y$ are parameters. The total rate density is obtained by summing over all past occurrences, including earthquakes outside the test region that could affect the rate density within it:

$$\lambda(t, m, x, y) = \mu \lambda_0(t, m, x, y) + \sum_{t_i \geq t \geq m_c \geq m_i} \eta(m_i) \lambda_i(t, m, x, y), \quad (9)$$

where $\mu$ is a parameter representing the proportion of earthquakes with no long-term precursors, $\lambda_0$ is the PPE rate density defined above, $t_0$ is the time of the beginning of the catalogue and $\eta$ is a normalizing function.
Previous applications of EEPAS have included both equal-weighting ($w_i = 1$, for all $i$) and unequal-weighting strategies, and optimization of a varying number of parameters. Applications at relatively low values of $m_c$ have tended to favour equal weighting, and those at relatively high values of $m_c$ have tended to favour down-weighting of aftershocks. The number of parameters to optimize is a difficult choice. Big apparent improvements in the fit to past data can sometimes be obtained by fitting many parameters, but approximate values of most parameters are suggested by the predictive scaling relations derived from many examples of the precursory scale increase phenomenon (Evison and Rhoades, 2004), and there is reason to be concerned about over-fitting, especially where a number of the fitted parameter values differ markedly from the suggested ones.

In consideration of these issues, four different versions of EEPAS have been submitted to the New Zealand testing centre. These are EEPAS_0r (a version with equal weighting and restricted parameter optimization), EEPAS_1r (with down-weighting of aftershocks and restricted parameter optimization), EEPAS_0f (with equal weighting and full parameter optimization); and EEPAS_1f (with down-weighting of aftershocks and full parameter optimization). EEPAS_0f is the best performing model in retrospective tests on the past catalogue, but whether the same is true in prospective testing remains to be seen. The parameter values adopted in each of the models are listed in Table 4. All four versions have been optimized for the time period 1987-2006 using earthquakes since 1951 with magnitudes exceeding the threshold $m_0 = 2.95$ and hypocentral depth $h \leq 45$ km in the data-collection region to forecast earthquakes in the test region with magnitudes exceeding $m_c = 4.95$ and $h \leq 40$ km.

Table 4  EEPAS model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EEPAS_0r</td>
</tr>
<tr>
<td>$a_M$</td>
<td>1.00</td>
</tr>
<tr>
<td>$b_M$</td>
<td>1.0*</td>
</tr>
<tr>
<td>$\sigma_M$</td>
<td>0.32*</td>
</tr>
<tr>
<td>$a_T$</td>
<td>1.85</td>
</tr>
<tr>
<td>$b_T$</td>
<td>0.40*</td>
</tr>
<tr>
<td>$\sigma_T$</td>
<td>0.23*</td>
</tr>
<tr>
<td>$b_A$</td>
<td>0.35*</td>
</tr>
<tr>
<td>$\sigma_A$</td>
<td>1.72</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.0604</td>
</tr>
</tbody>
</table>

* Fixed parameter
Figure 12 shows a three-month forecast of earthquakes with magnitude $M \geq 5.0$ under the EEPAS_0r model.

### 3.3 24-hour models

#### 3.3.1 STEP

The STEP (Short-Term Earthquake Probability; Gerstenberger, 2003; Gerstenberger et al., 2005) model is an aftershock model based on the idea of superimposed Omori (Ogata, 1988, 1998) type sequences. The model comprises 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; for the testing centre the NZNSTM is applied as the background model. The clustering model is based on the work of Reasenberg and Jones (1989) which defines aftershock forecasts based on the $a$- and the $b$- value from the Gutenberg-Richter relationship (Gutenberg & Richter, 1954) and the $p$-value from the modified Omori law (Ogata, 1988, 1998).

The STEP model defines these aftershock forecasts based on three model-components. The first component is based on the average behaviour of aftershocks in New Zealand and uses the median Reasenberg and Jones parameter values for New Zealand aftershocks, and uses parameter estimates from Pollock (2007). The second component uses the development of the ongoing 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. 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 Akaike Information Criterion (AICc), described by 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.

Figure 13 shows a 24-hour forecast of earthquakes with magnitude $M \geq 5.0$ under the STEP model.

#### 3.3.2 ETAS

The ETAS (Epidemic-Type After-Shock) model is a widely used model of earthquake clustering (Ogata, 1988, 1989, 1998; Console and Murr, 2001; Console et al., 2003, 2006). The version of the ETAS model installed in the testing centre is different in several details from most published versions. It is based on the aftershock model used to down-weight aftershocks in the EEPAS model (Rhoades and Evison, 2004).
The rate density \( \lambda' \) is of the form

\[
\lambda'(t, m, x, y) = \nu \lambda_0(t, m, x, y) + \kappa \sum_{i \geq 0} \lambda_i'(t, m, x, y)
\]  

(10)

where \( \lambda_0 \) is the PPE rate density defined above; \( \nu \) and \( \kappa \) are constant parameters, called the "failure rate" and "productivity", respectively; and

\[
\lambda_i'(t, m, x, y) = f_{2i}(t) g_{2i}(m) h_{2i}(x, y).
\]  

(11)

Here, \( f_{2i}, g_{2i} \) and \( h_{2i} \) are functions for the time, magnitude and location of the aftershocks of the \( i \)th earthquake. The time distribution follows the modified Omori-Utsu law (e.g. Ogata, 1983); the magnitude distribution has regard to the Gutenberg-Richter relation; and the location distribution is bivariate normal with circular symmetry and has regard to Utsu's areal relation (Utsu, 1961). Thus

\[
f_{2i}(t) = H(t-t_i) \frac{p-1}{(t-t_i+c)^p},
\]  

(12)

\[
g_{2i}(m) = \beta \exp[-\beta(m-m_i)],
\]  

(13)

and

\[
h_{2i}(x, y) = \frac{1}{2\pi \sigma_U^2 10^m} \exp\left[ -\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma_U^2 10^m} \right],
\]  

(15)

where \( c, p, \) and \( \sigma_U \) are constant parameters, and \( \beta = \text{bin}(10) \), where \( b \) is the Gutenberg-Richter b-value. The Heaviside function \( H \) is as in equation (7).

Parameter values were either fixed or optimised on the New Zealand earthquake catalogue in the test region over the period 1987-2006, using data from the collection region with magnitudes \( m > 2.95 \), and hypocentral depth \( h \leq 40 \) km. The adopted parameter values for the test are given in Table 5.

Table 5 ETAS model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>0.32</td>
</tr>
<tr>
<td>( K )</td>
<td>0.072</td>
</tr>
<tr>
<td>( c ) (days)</td>
<td>0.03*</td>
</tr>
<tr>
<td>( P )</td>
<td>1.11</td>
</tr>
<tr>
<td>( \sigma_U )</td>
<td>0.0020</td>
</tr>
<tr>
<td>( B )</td>
<td>1.16</td>
</tr>
</tbody>
</table>

* Fixed parameter
Figure 10  $\log_{10}$ of expected number of earthquakes with magnitude $M \geq 5.0$ in a five year period under the NZNSHM model.
Figure 11  $\log_{10}$ of expected number of earthquakes with magnitude $M \geq 5.0$ in a three month period (July-September 2006) under the PPE model.
Figure 12. \( \log_{10} \) of expected number of earthquakes with magnitude \( M \geq 5.0 \) in a three month period (July-September 2006) under the EEPAS_0r model.
Figure 13  \( \log_{10} \) of expected number of earthquakes with magnitude \( M \geq 5.0 \) in a 24-hour period under the STEP model.
4.0 MODIFIED MERCALLI INTENSITY HAZARD FORECASTS

For any model installed in the testing centre, the forecast of earthquake occurrence can be translated into a forecast of the probability of a given Modified Mercalli (MM) intensity of shaking being felt at any location in New Zealand over the same period (five years, three months or 24 hours), using software added to the testing centre under this project. This is achieved through a three-stage process:

1. Numerous synthetic Monte Carlo catalogues conforming to the forecast are generated.

2. Using an attenuation model for MM intensity, a number of isoseismal maps are generated for each earthquake in the synthetic catalogues taking into account the epistemic and aleatory uncertainties in the attenuation model.

3. The number of times a given level of MM intensity is exceeded is tabulated at points on a latitude-longitude grid, converted to the probability of exceedance per unit of time. The grid of probabilities is smoothed to produce a contour map.

The methods involved in stages 1 and 2 are not entirely straightforward, and of necessity involve a number of approximations and subjective choices, which are not necessarily optimal. Therefore, we attempt to spell out the methods used in some detail, so that the choices made can be revisited, if necessary, in the future.

4.1 Synthetic catalogue generation

In generating synthetic catalogues, it is necessary to generate a magnitude, epicentre and hypocentral depth for each earthquake. Epicentral locations are generated randomly within a cell of latitude and longitude. Similarly hypocentral depths are generated randomly over 0 – 40 km, because the models do not forecast the distribution of depths within this interval. In reality, the depth distribution of shallow earthquakes is neither uniform on this interval, nor the same at all locations. This can be seen from Figures 8 and 9. However, the distribution of depths at any location is not well known because of uncertainties in hypocentral depths in the catalogue.

4.2 Attenuation model and uncertainties

An attenuation model proposed by Dowrick and Rhoades (2005) is used. The particular model adopted is their "Main Seismic Region" Model 2, which does not require specification of the focal mechanism of the earthquake. This model requires, for each earthquake, specification of the moment magnitude, centroid depth, minimum depth to the top of the source, and whether or not the earthquake occurs in the crust. The following approximations are made in applying it here:

1. The moment magnitude is assumed to be the same as the local magnitude forecast by the model, and for simulated earthquakes is assumed to be known exactly;

2. The isoseismal distance is the distance to the epicentre, i.e., that the epicentre is the centre of the isoseismal pattern;
3. The centroid depth is taken to be the same as the hypocentral depth \( h \);

4. The minimum depth to the top of the source is the maximum of 0 and \((h - w / 2)\), where \( w \) is determined from the scaling relation for fault width as a function of magnitude, given for
New Zealand earthquakes in table 5 of Dowrick and Rhoades (2004), i.e.,
\[ w = -2.19 + 0.5M _w \quad (\text{if} \ M _w < 6.0) \quad \text{and} \quad w = -1.02 + 0.31M _w \quad (\text{if} \ M _w \geq 6.0). \]

5. All earthquakes in the synthetic catalogues are taken to be crustal.

The attenuation model is not isotropic. It assumes elliptical isoseismals, and estimates the
major axis (along-strike) and minor axis (strike-normal) distances separately. The strike-
angle is another parameter which is required as input to the model. For simplicity it is taken
as 33 degrees East of North for every simulated earthquake, a typical, but by no means
universal, angle for fault orientation in the New Zealand region.

The attenuation in the along-strike (a) direction is given by a random-effects regression
model of the form
\[
I _a = A _1 + A _2 M _w + A _3 \log _{10} (r _a ^2 + d _a ^2)^{1/3} + A _4 h _c + A _5 \delta _c + \tau \delta + \sigma \varepsilon _a
\]
where \( I _a \) is the MM intensity on the \( i \)th isoseismal, \( r _a = (a _i ^2 + h _c ^2)^{1/2} \), where \( a _i \) is the along-
strike isoseismal distance, \( h _c \) is the minimum depth of the earthquake source, \( h _c \) is the
centroid depth, \( \delta _c \) is 1 for crustal earthquakes and 0 otherwise, and \( \delta \) and \( \varepsilon _a \) are standard
normal random variables. The fitted parameters are \( A _1, ..., A _5, \tau \) and \( \sigma \). Their estimates and
standard errors (from Dowrick and Rhoades, 2005) are given in Table 6, and their correlation
matrix, not previously published, is given in Table 7.

The attenuation in the strike-normal (b) direction is given by
\[
\log _{10} \left( \frac{b _i}{a _i} \right) = B _1 + B _2 M _w + B _3 I _a + B _4 \ln (a _i) + \tau _g \delta + \sigma _g \varepsilon _b
\]
where \( b _i \) is the strike-normal isoseismal distance, \( \delta \) and \( \varepsilon _b \) are standard normal random
variables and \( B _1, ..., B _4, \tau _g \) and \( \sigma _g \) are fitted parameters. The parameter values and standard
errors are given in Table 8, and their correlation matrix is given in Table 9. Again, these
values come from the Dowrick and Rhoades (2005) study, but not all of them have been
published previously.
Table 6  Parameter estimates and standard errors from along-strike attenuation model (16) of Dowrick and Rhoades (2005).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>4.402</td>
<td>0.401</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1.257</td>
<td>0.0589</td>
</tr>
<tr>
<td>$A_3$</td>
<td>-3.669</td>
<td>0.0798</td>
</tr>
<tr>
<td>$A_4$</td>
<td>0.012</td>
<td>0.0045</td>
</tr>
<tr>
<td>$A_5$</td>
<td>0.409</td>
<td>0.1131</td>
</tr>
<tr>
<td>$d$</td>
<td>11.775</td>
<td>0.9675</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.192</td>
<td>0.0274</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.387</td>
<td>0.0150</td>
</tr>
</tbody>
</table>

Table 7  Correlation matrix for parameters of along-strike attenuation model (16).

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
<th>$d$</th>
<th>$\tau$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>1.00</td>
<td>-0.84</td>
<td>-0.26</td>
<td>-0.36</td>
<td>-0.26</td>
<td>0.22</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1.00</td>
<td></td>
<td>-0.14</td>
<td>-0.02</td>
<td>-0.05</td>
<td>0.11</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$A_3$</td>
<td>-0.26</td>
<td>1.00</td>
<td></td>
<td>-0.71</td>
<td>-0.04</td>
<td>-0.11</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$A_4$</td>
<td>-0.36</td>
<td>-0.14</td>
<td>-0.71</td>
<td>1.00</td>
<td></td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$A_5$</td>
<td>-0.26</td>
<td>-0.02</td>
<td>-0.04</td>
<td>0.61</td>
<td>1.00</td>
<td></td>
<td>0.00</td>
<td>-0.29</td>
</tr>
<tr>
<td>$d$</td>
<td>0.22</td>
<td>0.11</td>
<td>-0.71</td>
<td>-0.11</td>
<td>0.05</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.29</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 8  Parameter estimates and standard errors from strike-normal attenuation model (17) of Dowrick and Rhoades (2005).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_1$</td>
<td>3.62</td>
<td>0.36</td>
</tr>
<tr>
<td>$B_2$</td>
<td>0.45</td>
<td>0.12</td>
</tr>
<tr>
<td>$B_3$</td>
<td>-0.56</td>
<td>0.07</td>
</tr>
<tr>
<td>$B_4$</td>
<td>-0.53</td>
<td>0.10</td>
</tr>
<tr>
<td>$\tau_B$</td>
<td>0.48</td>
<td>0.07</td>
</tr>
<tr>
<td>$\sigma_B$</td>
<td>0.44</td>
<td>0.03</td>
</tr>
</tbody>
</table>
Table 9  Correlation matrix for parameters of strike-normal attenuation model (17).

<table>
<thead>
<tr>
<th></th>
<th>$B_1$</th>
<th>$B_2$</th>
<th>$B_3$</th>
<th>$B_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_1$</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B_2$</td>
<td>-0.11</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B_3$</td>
<td>-0.37</td>
<td>-0.84</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>$B_4$</td>
<td>-0.51</td>
<td>-0.75</td>
<td>0.84</td>
<td>1.00</td>
</tr>
<tr>
<td>$t_B$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\sigma_B$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

In applying the attenuation models (16) and (17), we allow for both epistemic (parameter) uncertainties and aleatory uncertainties (random variation from the mean relation). Allowing for the epistemic uncertainty requires the parameter standard deviations and the correlation matrix of the estimated parameters, from which we derived the variance-covariance matrix of the estimated parameters, which we denote by $\Lambda$. As a positive semi-definite symmetric matrix, $\Lambda$ has a Choleski decomposition of the form $\Lambda = LL'$ where $L$ is a square matrix in which the upper off-diagonal elements are zero. Then if $w$ is a vector of independent standard normal random variables, the vector $u = Lw$ has the same distribution as the vector of parameter uncertainties, and can be used to simulate the epistemic uncertainty, by generating a different realisation of $w$ is used for each synthetic catalogue. For more details, see Rhoades and McVerry (2001).

Allowing for the aleatory uncertainty requires the proper treatment of the between-earthquake and within-earthquake variation. The simulation of isoseismal patterns requires the generation of a value of the random variable $\delta$ for each earthquake and of the random variable $c_j$ for each isoseismal. Because of the physical requirement that (in the case of the along-strike isoseismal distances) $a_i < c_j$ if $I_i > I_j$, any combinations of the $c_j$ values which would violate this assumption have to be discarded. The same applies to simulation of the strike-normal isoseismal distances $b_i$, which can be done once the $a_i$ values are determined, using equation (17) and the parameters, uncertainties and correlation matrix given in Tables 8 and 9.

4.3  MM intensity probability maps

Having calculated a set of $a_i$ and $b_i$ distances for each value of MM intensity $I_i$, the elliptical isoseismals are determined. For any point in a grid of latitude and longitude, the number of exceedances of a given level of MM intensity is accumulated over simulated isoseismal patterns for a given earthquake and then over simulated earthquakes. From the total number of exceedances, the probability of exceeding a given MM intensity over the period of the forecast is estimated for each model at every point in the grid. A smooth interpolation of the grid probabilities is produced using the Generic Mapping Tools software (Wessel and Smith, 1995).

In Figure 14 and 15, we give examples of intensity probability maps for two of the three-month forecasting models, namely, PPE and EEPAS_0r, respectively. They show the expected rate exceedance of MM7 in a particular 90-day forecast period under each of these models.
Figure 14  Rate of occurrence of MM7 in the 90 day period of a forecast according to the PPE model.
Figure 15  Rate of occurrence of MM7 in the 90 day period of a forecast according to the EEPAS_0r model.
5.0 MODEL DEVELOPMENT

This section was contributed by Annemarie Christophersen

5.1 Overview

Earthquakes cluster in space and time. In retrospect, the largest earthquake in a cluster is called the mainshock. The earthquakes preceding a mainshock are named foreshocks, and the earthquakes following a mainshock aftershocks. Clusters with no discernible mainshock are often referred to as swarms. No physical differences between fore-, main- or aftershocks are known. Short-term earthquake clustering models, such as the short term earthquake probability model (STEP) (Gerstenberger, et al., 2005) or the epidemic type aftershock sequence (ETAS) (Console, et al., 2003; Ogata, 1988; Zhuang, et al., 2002), assume that all earthquakes trigger subsequent earthquakes in a self-similar manner, only dependent on the size of the triggering event. However, Tormann found in her EQC funded master thesis that foreshock activity in New Zealand decayed faster in time, space and magnitude than aftershock activity (Tormann, 2005). Further inconsistencies between foreshock rates and aftershock occurrence have been observed in other studies and are outlined in (Christophersen and Smith, 2008). Quantifying any potential differences between foreshocks and mainshocks is an important input for the modelling of short-term earthquake occurrence.

This project and associated work set out to either confirm the assumptions of models like ETAS and STEP that foreshocks behave like mainshocks that happen to have larger aftershocks, or to describe the different characteristics as input for new short-term earthquake models that could be tested within the new testing centre. We proceeded in three different ways, and have three manuscripts for publications associated with our work:

1. Analyse the decay of foreshock activity in time, space and with difference in magnitude between foreshock and mainshock (Tormann et al., 2008; Appendix 1).

2. Analyze the triggering capability (abundance) of mainshocks, and combine the abundance model with the empirical relationship for magnitude to derive expected foreshock rates and compare with observations (Christophersen and Smith, 2008; Appendix 2)

3. Create synthetic earthquake catalogues with known characteristics, to test that results from 1 and 2 are reproducible and not an effect of the data selection (Zhuang et al., 2008; Appendix 3).

To study the characteristics or foreshocks, mainshocks and aftershocks, the earthquakes first need to be classified into these different events. The identification of earthquake clusters from an earthquake catalogue is often referred to as declustering. In principle, two ways of declustering exist: deterministic and stochastic methods. In the deterministic methods, an earthquake is assigned to an earthquake cluster either by spatial and temporal windows or by some linking algorithm. Our analysis 1 applied a window method to the New Zealand earthquake catalogue to first remove smaller earthquakes subsequent to an earthquake and then study the remaining earthquakes in the catalogue (Tormann, et al., 2008). In the second analysis, a range of window methods and one linking algorithm with varying parameter
settings were used to assign each earthquake in the catalogue to a cluster, possibly consisting of only one earthquake (Christophersen and Smith, 2008). In the third analysis, the earthquake catalogue was declustered by a stochastic method based on the ETAS model (Zhuang, et al., 2008).

The parameters derived from the stochastic method were used to simulate synthetic earthquake catalogues to test the methods from analyses 1 and 2. Interestingly, the deterministic methods were not able to reproduce the results from the real data in the synthetic catalogues. The ETAS model produced many late aftershocks (for the New Zealand parameters about 90% of first generation aftershocks had occurred after 1000 days but for Californian parameters only 68%). The traditional window methods were not able to identify late aftershocks and therefore most likely split clusters into several smaller parts. This way a smaller aftershock that was followed by a larger one could have been counted as a foreshock thus artificially increasing the foreshock rate. For New Zealand and California, the observed foreshock rate in the real data was only about half the foreshock rate in the ETAS simulated data.

Analysis 1 found that the foreshock characteristics were consistent with aftershocks characteristics. Analysis 2 showed that the observed foreshock rate in different catalogues and with different clustering algorithm was consistent with the triggering capability of mainshocks and the magnitude-frequency distribution of the complete catalogues. Analysis 3 found a smaller triggering capability and foreshock rate in the background earthquakes than in the triggered earthquakes.

Below we briefly summarise the content of each manuscript and discuss the apparent discrepancy between the different results.

\subsection*{5.2 Time, distance and magnitude dependent foreshock probability model for New Zealand}

We calculated the probability of an initial earthquake (a foreshock) being followed by a mainshock in New Zealand, considering the elapsed time and distance and magnitude differences between foreshock and mainshock. We used non-aftershock events between 1964 and 2007, with magnitude 4.0 and shallower than 40 km. In our definition, aftershocks are earthquakes smaller in magnitude than some previous earthquake. Aftershocks were removed from the catalogue by a New Zealand-specific mainshock-magnitude dependent window in time and space (Savage and Rupp, 2000). We separated the catalogue into events within and outside of the Taupo Volcanic Zone (TVZ). We derived a model for the probability $P(t, r, \delta M)$ that at time $t$ after a potential foreshock of magnitude $M_p$, and at distance $r$, a mainshock with magnitude $M_p + \delta M$ will occur:

$$P(t, r, \delta M) = P_0 10^{(-B \delta M)} (t + c_1)^{p_1} (r + c_2)^{p_2}, \quad (18)$$

where $P_0, B, p_n, c_n, p_r$ and $c_r$ are constants that we determined from the data. Our study had four main results. First we found that binning data using fixed intervals of time or space before fitting the parameters returned different values than a more robust approach of fitting directly the entire range. Secondly, foreshock probabilities decreased with increasing inter-
event time as described by a modified Omori law with an exponent $p_i$ close to one (0.9±0.2 (TVZ) and 0.8±0.1 elsewhere, where uncertainty estimates are 95% confidence intervals. Thirdly, foreshock probabilities decreased with increasing epicentral distance, also following a modified-Omori-type decay with exponent $p_i = 0.9 \pm 0.2$ (non-TVZ) and $1.7 \pm 0.6$ (TVZ). Finally, the magnitude distribution of mainshocks followed the Gutenberg-Richter relation with $B = 1.0 \pm 0.17$ (Non-TVZ) and $1.5 \pm 0.5$ (TVZ). The differences between the TVZ and the rest of New Zealand were consistent with higher attenuation in the TVZ, deduced from previous studies.

5.3 Foreshock rates from aftershock abundance

We investigated whether foreshocks and mainshocks have the same triggering capability. For this purpose we defined abundance as a new formulation for the triggering capability, or productivity of mainshocks. It has two parameters: The growth exponent $\alpha$ and the magnitude $M_f$. In this formulation $M_f$ has a physical meaning, the magnitude that on average has one aftershock in time, space and magnitude windows that are defined by the search algorithm for earthquake clusters. $M_f$ was a useful tool to compare different search algorithms. We used four search algorithms; the Reasenberg linking algorithm (Reasenberg, 1985) and a hybrid method of magnitude dependent search radii and sliding time windows $\Delta T$ with three different radii derived from (i) Gardner-Knopp's Southern Californian declustering algorithm (Gardner and Knopoff, 1974); (ii) Uhrhammer's Northern Californian declustering algorithm (Uhrhammer, 1986); (iii) Utsu-Seki's formula for the growth of aftershock area with mainshock magnitude (Utsu and Seki, 1955). Each algorithm was stretched beyond its limits in the sense that we applied it outside the magnitude range and the geographical area for which it was developed.

We analyzed two subsets of the South Californian catalog and one of the CMT global catalog and calculated the $b_{cat}$-values of the magnitude-frequency distribution, $\alpha$ and $M_f$. We defined foreshocks as initiating events of a cluster that were followed by a larger earthquake within the time period $\Delta T$. Assuming that foreshocks had the same triggering capability as mainshocks, we combined mean abundance with the magnitude-frequency distribution and derived foreshock rates. These compared well with the observations. For Southern California we found that the probability of an initiating event in an earthquake cluster in the magnitude range $2.0 - 4.5$ being followed by at least one larger event within one day is around 4%. This agreed well with corresponding data from the global CMT catalogue where the probability that the initiating earthquake of a cluster was followed by at least one larger event within 30 days is around 3%.

5.4 Differences between spontaneous and triggered earthquakes: their influence on foreshock probabilities

In this study we investigated the foreshock probabilities calculated from earthquake catalogues from Japan, Southern California and New Zealand. Unlike conventional studies on foreshocks, we used a probability-based declustering method to separate each catalogue into stochastic versions of family trees, such that each event was classified as either having been triggered by a preceding event, or being a spontaneous event (Zhuang, et al., 2002). The probabilities were determined from parameters that provided the best fit of the real catalogue to a synthetic catalogue determined using a space-time epidemic-type aftershock
sequence (ETAS) model, which assumed that every earthquake had a magnitude, space and time dependent probability of triggering another event. A foreshock here was defined as a spontaneous event that had one or more larger descendants. The triggering capability as well as the proportion of foreshocks in the spontaneous earthquakes was found to be lower than the triggering capability and the proportion of foreshocks in triggered events. Our results imply that there is a difference between the ETAS model that assumes all earthquakes triggering subsequent earthquakes in the same manner, and the data that when analysed with the ETAS model contradicted these assumptions. Work is on-going to simulate catalogues that feature potential differences the ETAS model and the real data, like spatial anisotropy of triggered earthquakes, missing earthquakes subsequent to a large earthquake and changes in parameters in the rate equation with time.

5.5 Discussion

No differences between foreshock, mainshock and aftershock occurrence could be found using deterministic window and linking methods to identify earthquake clusters. In contrast, the ETAS model that treats the earthquake catalogue as a family of earthquakes finds a lower triggering capability and foreshock probability for spontaneous events compared to triggered earthquakes. These discrepancies are not eliminated when the ETAS declustering is applied to a synthetic ETAS catalogue, even if this is modified by modelling a spatial anisotropic aftershock distribution, removing early aftershocks and varying parameters as a function of time. The window and linking method cannot identify any family structure in an earthquake cluster and therefore cannot distinguish between an immediate aftershock and one that was triggered by another aftershock. Thus these traditional methods cannot be used to study the difference between spontaneous and triggered events, because the triggered events will always include later generation aftershocks (in the terminology of the ETAS model). Therefore mainshocks that occur first in a sequence will, on average, have fewer aftershocks than mainshocks that occur later in the sequence, because we will include in the aftershock count all aftershocks from previous earthquake in the cluster. There is a new stochastic declustering method (Marsan and Lengliné, 2008) that could be applied to shed further light on potential differences between spontaneous and triggered earthquakes.

We close with two conclusions:

1. No separate modelling of foreshock hazard is required for models that include the magnitude-frequency relation and therefore allow for larger earthquakes to follow smaller ones, and that have their parameters derived from window or linking algorithms.

2. The ETAS model does not capture all features of the real earthquake catalogue, and the differences still need to be fully understood.
6.0 CONCLUSION

The establishment of the New Zealand Earthquake Forecast Testing Centre is an important milestone towards the development of usable scientific earthquake forecasts for the New Zealand region. The software and testing methods used in the New Zealand centre are consistent with other testing centres in California and Europe. The New Zealand testing centre, along with the other regional testing centres, provides for more rigorous and transparent testing of a wide range of proposed forecasting models than has occurred in the past. Researchers from all over the world will be encouraged to submit their models for testing in the New Zealand centre, as well as the other regional centres.

The models already submitted include representatives of some of the best established and most studied classes of model already in existence. However, every model is capable of further development; whether it is a long-term model such as NZNSHM, a medium-term forecasting model such as EEPAS, or a short-term model, such as STEP and ETAS. Moreover, there is scope for improving forecasting capability by attempting to combine the information from essentially different models into hybrid forecasting models.

The investigations carried out in the model-development component of the present study have highlighted several useful directions for short-term modelling of earthquake occurrence. In particular, the difference that appears to exist between the ETAS model – which assumes that spontaneous and triggered earthquakes have the same triggering capability and foreshock probability – and the declustered catalogues, which seem to indicate that this assumption is not borne out in practice – demand further investigation. Understanding of this difference is likely to lead to improved modelling of earthquake clustering.

The collaboration that is developing under the umbrella of CSEP is likely to spawn a more rapid development of new testable models of earthquake occurrence, including models based on observations which make use of other data-bases than the past earthquake catalogue, and some models that could not be accommodated in the present testing framework. It will therefore be necessary for the testing centre activities to be expanded to respond to the challenge of testing new types of models.

The centre's activities could also be usefully expanded to make the testing software available to researchers when preparing their models for submission. The process of developing a forecasting model is an arduous one which involves extensive computer code development. Errors in large computer codes are easy to create and difficult to find. It is better that coding errors be found before the models are submitted, rather than after years of formal testing. Error detection could be assisted if the researchers were able to retrospectively test their models using the same software that will be used in the prospective tests, before submitting their models for testing.

The eventual goal of earthquake forecasting research is to be able to provide usable earthquake hazard forecasts based on rigorously tested methods to the hazard management sector. The conversion of earthquake-occurrence probability models to MM intensity probability maps, enabled by this project, is a necessary step towards that goal. However, the work here has highlighted that this conversion is not entirely straightforward, and involves several approximations and choices, which may need to be revisited in the future.
ACKNOWLEDGEMENTS

This work was funded by the EQC Research Foundation under project number 510. Software developed by the California CSEP Earthquake Forecast Testing Center has been extensively used in setting up the New Zealand testing centre. J. Cousins and M. Stirling have provided helpful reviews of the report.

REFERENCES


Field, E. H. (2007), Overview of the working group for the development of regional earthquake likelihood models (RELM), *Seismological Research Letters* 78(1), 7-16


APPENDICES
APPENDIX 1  TIME, DISTANCE AND MAGNITUDE DEPENDENT FORESHOCK PROBABILITY MODEL FOR NEW ZEALAND (21 PAGES)

Time, distance and magnitude dependent foreshock probability model for New Zealand

by

Thessa Tormann, ETH Zurich, thessa.tormann@sed.ethz.ch
Martha K. Savage, VUW, Martha.Savage@vuw.ac.nz
Euan G.C. Smith, VUW, Euan.Smith@vuw.ac.nz
Mark W. Stirling, GNS, m.stirling@gns.cri.nz
Stefan Wiemer, ETH Zurich, wiemer@sed.ethz.ch

Abstract

The possibility that a moderate earthquake may be followed by an equal or larger one (foreshock probability) increases the hazard in its immediate vicinity for a short time by an order of magnitude or more. Thus, foreshock probabilities are of interest for time dependent seismic hazard forecasts. We calculate the probability of an initial earthquake (a foreshock) being followed by a mainshock in New Zealand, considering the parameters of elapsed time and distance and magnitude differences between foreshock and mainshock. We use non-aftershock events between 1964 and 2007, with magnitude ≥ 4.0 and shallower than 40 km, separating the catalogue into events within and outside the Taupo Volcanic Zone (TVZ). We provide a model for the probability $P(t, r, \delta M)$ that at time $t$ after a potential foreshock of magnitude $M_f$, and at distance $r$, a mainshock with magnitude $M_f + \delta M$ will occur: $P(t, r, \delta M) = P_0 * 10^{(-a_{\delta M})} * (t + c_1)^{-b_1} * (r + c_2)^{-b_2}$, where $P_0, B, p, c_0, p, c_1$ are constants to be determined.

We find (1) binning data using fixed intervals of time or space before fitting the parameters returns different values than a more robust approach of fitting directly the entire range, (2) foreshock probabilities decrease with increasing inter-event time as described by a modified Omori law with an exponent $p$, close to one (0.9±0.2 (TVZ) and 0.8±0.1 elsewhere – uncertainty estimates are 95% confidence intervals throughout this study), (3) foreshock probabilities decrease with increasing epicentral distance also following a modified Omori type decay with exponent $p$, of 0.9±0.2 (non-TVZ) and 1.7±0.6 (TVZ) and (4) the mainshock magnitude distribution follows the Gutenberg-Richter relationship ($B=1.0±0.17$ (Non-TVZ) and 1.5±0.5 (TVZ)). The differences between the TVZ and the rest of New Zealand are consistent with higher attenuation in the region, deduced from previous studies.

Introduction

The analysis of “linked events” (e.g. foreshocks-mainshocks-aftershocks, stress triggering) is of increasing interest for the development of time dependent seismic hazard forecasting methodologies (e.g. models submitted to RELM/CSEP, see Seismological
Research Letters, Jan/Feb 2007). Such analysis has been carried out mainly in strike-slip regions such as California, but is less developed in other settings. Many authors consider that foreshocks follow the same decay laws as aftershocks and use these parameters in such models (e.g. Gerstenberger et al. (2005)). Yet theoretical arguments suggest that the damage in large earthquakes might alter the decay parameters, or that the earthquake preparation process before large events could result in different decay laws. Therefore it is worthwhile to examine foreshock decay laws in different tectonic regions.

Located on a plate boundary with complex tectonic setting, New Zealand experiences frequent moderate sized earthquakes (Fig. 1a) and is thus well set up to test some of these hypotheses. New Zealand includes two subduction zones dipping in opposite directions and a transpressional region linking the two. The northernmost subduction zone includes a region of high heat flow, high seismic attenuation, low seismic wave speeds, and volcanism. It is interpreted as a back-arc extensional basin that represents the southern end of the Lau-Havre Trough (e.g. Stern et al. (2006)). Within this extensional zone, New Zealand geologists sometimes differentiate between a western and eastern region. The entire region is sometimes called the Central Volcanic Zone. However, for the purposes of this study we refer to the entire region as the Taupō Volcanic Zone (TVZ) to distinguish it from volcanic zones in other regions of the world (Fig. 1a).

An effort has begun to calculate the likelihood of damaging events occurring anywhere in New Zealand at any given time (Tormann (2005), Gerstenberger (2005) now being applied to NZ). We have developed methodologies to compare such calculations to the strong ground motion record to test if the models are working (Tormann (2005)). However, the methods rely on having adequate models of event occurrence. Therefore, this paper addresses the questions of how the probabilities for future larger events in New Zealand vary with increasing time, distance and magnitude from a potential foreshock. The goal is to later use such results with known aftershock occurrence rates to calculate time varying hazard models for New Zealand.

A foreshock in our definition is an earthquake that is followed by an event of equal or larger magnitude within 5 days and 50 km from the initial event. We develop equations that allow calculation of the probability for a mainshock of a certain magnitude at a certain time after, and distance from, a potential foreshock. We show that the rates of decay of foreshock probability with time and space in New Zealand are comparable to other regions of the world. We also demonstrate that using linear bins of distance or time before determining the parameters distorts the results and leads to different parameters than using unbinned data.

**Previous Studies**

The first investigations of prospective foreshock probabilities, i.e. the probability that an event might be followed by an equal or larger event within a certain time and space window were done in California. Jones (1985) studied the Southern Californian catalogue and found that 6±0.5% of \( M \geq 3 \) events are followed by an equal or larger event within 5 days and 30 km, i.e. qualify as foreshocks. With slightly changed windows of 5 days and
10 km search radius Savage and DePol (1993) found the same 6% chance for an $M \geq 3$ event to be followed by an equal or larger one in the Nevada region. For the volcanic Mammoth/Mono region the probabilities reach 10%. Global foreshock probabilities have been found to be twice as high as the southern Californian values from Jones' 1985 model (Reasenberg (1999)). Reasenberg also found a significant dependency of foreshock probabilities on the mainshock's focal mechanism, thrust earthquakes being the most frequent to have foreshocks.

The hypothesis that foreshocks might be mainshocks whose aftershocks happen to be big and therefore could be modeled and predicted using aftershock relationships has been suggested and tested by several authors (e.g. Jones et al. (1999), Felzer et al. (2004)). Felzer et al. (2004) analysed the global CMT and NEIC catalogues as well as the local Californian CNSS catalogue to argue that one single physical triggering mechanism can explain the occurrence of foreshocks, aftershocks and multiples, i.e. clusters of earthquakes are the result of an initial earthquake triggering subsequent ones. They found that the magnitude of the triggering earthquake is independent of the magnitude of the triggered events and that the number of foreshock and multiples observations is a fixed fraction of the number of aftershock observations as estimated by the Gutenberg-Richter relationship with $b=1$. The Gutenberg-Richter relation (1944) is given by

$$\log 10(N) = a - bM$$

(1)

where $N$ is the cumulative number of earthquakes equal to or greater than magnitude $M$ and the productivity rate $a$ and slope $b$ are constants. Zhuang et al. (2006) examine the Japanese JMA catalogue through a stochastic declustering method, and suggest that triggered events have a higher probability of causing subsequent events than background events.

To clarify aftershock triggering processes, Felzer and Brodsky (2006) studied the distance decay in aftershock sequences within the first few minutes after a mainshock. They observe a good fit of the data by a single inverse power law with an exponent of -1.35, valid over distances of 0.2 to 50 km.

In order to characterize the chances that a sequence of small to moderate earthquakes is a precursory effect of a large mainshock Ogata et al. (1995) and Maeda (1996) considered relationships between foreshock clusters and significantly larger mainshocks as a function of magnitude, temporal and spatial distribution in the cluster. They find that swarms in which the inter-event times and spatial separations between the epicenters are very small (several days to few hours and less than 10 km, respectively) and the magnitudes are increasing, are very likely to be a precursor of a forthcoming mainshock.

The two most recent foreshock probability studies of New Zealand data were carried out by (1) Savage and Rupp (2000) who applied Jones' and Savage and DePol's methodology (with windows of 5 days and 30 km) and obtained average foreshock probabilities for $M \geq 5$ earthquakes of 4.5±0.4% and (2) Merrifield et al. (2004) who used the same windows and analyzed regional differences in foreshock probability for $M \geq 4$ events. Merrifield et al. (2004) found that the Taupo Volcanic Zone (TVZ) shows a significantly higher foreshock probability (9.6±1.7%) than the rest of New Zealand (6.3±0.47%).
Data & Methodology

We use the New Zealand catalogue as available from the GeoNet website (www.geonet.org.nz). We select the period from 01 Jan 1964 to 31 May 2007, the region between longitudes 165° and 180° and latitudes -34° and -49°, and magnitude range $M \geq 4.0$, which is the completeness level suggested by Smith (1981) (see also Fig. 1b).

The time series of the cumulative number of shallow events equal or larger than $M$ 4.0 shows a clear break in slope, with higher rates before 1987, and lower rates thereafter (Fig. 1c). In 1987 the EARSS earthquake detection and recording system was installed (Anderson et al. (1994)), which may have resulted in an unintended change in the magnitude calculations. To verify whether or not this shift affects the foreshock parameter analysis, we carry out calculations for the two periods 1964-1987 and 1987-2007 in addition to considering the whole period.

Aftershock sequences and all events deeper than 40km are eliminated from the catalogue. Aftershock removal is a necessary step as otherwise pairs from within aftershock sequences could match the searching criteria and artificially increase the foreshock probability values considerably (e.g. Jones (1985)). We use the Gardner & Knopoff (1974) algorithm with time and distance windows modified for New Zealand, as published by Savage and Rupp (2000), (see also Table 1). The algorithm applies simple, magnitude-dependent windows in time and space to remove aftershocks from the catalogue. The catalogue is examined in chronological order, and after each event, any further event of smaller magnitude within the appropriate distance and time window is removed from the catalogue. Equal or larger events will be left in the catalogue, and will begin their own search windows.

The effect on foreshock probabilities caused by different time and distance windows in the aftershock removal process is discussed in Savage and Rupp (2000). In that study the authors try to be conservative, in the sense that they prefer to leave out possible background events rather than remove too few aftershocks. The windows are chosen by examining plots of aftershock area and aftershock time period as a function of magnitude. The data sources for the aftershock areas and time periods come from Gibowicz (1973) and Evison and Rhoades (1993). An envelope curve is picked by eye that includes all the events within the sources. Further tests are carried out by doubling the aftershock removal windows. The larger windows have little effect on the foreshock probabilities, but increase the error bars because fewer events are included. However, when shorter windows are chosen, such as those published by Gardner and Knopoff (1974), there is an increase in apparent foreshocks with decreasing magnitude. The Reasenberg (1989) declustering algorithm likewise yielded an apparent increase in foreshocks with decreasing magnitude. Savage and Rupp (2000) attribute that effect to having left too many aftershocks in the catalogue. Among those events, small aftershocks that are followed by larger aftershocks enhance the observed probabilities. Because of this we used only the windows determined for New Zealand (Savage and Rupp (2000)).
Foreshock Probability Model and Maximum Likelihood Fitting

Examining the declustered catalogue in chronological order, we search for all earthquakes that have been followed by an equal or bigger one within 5 days and 50km epicentral distance, i.e. we treat all events as point sources and ignore depth information within the upper 40km as that is problematic in the New Zealand catalogue. We do not treat multiple foreshock sequences separately, so we include a maximum of one foreshock per mainshock.

We follow the probability calculations first proposed by Jones (1985). If \( N \) is the number of earthquakes tested (potential foreshocks) and \( n \) is the number of events that qualified as foreshocks, we calculate the foreshock probability \( P \) as:

\[
P = \frac{n}{N}
\]

(2)

As we assume the probability distribution for an earthquake to be a foreshock to follow a binomial distribution, the 95% confidence interval \( c_{\text{int}95} \) for the foreshock probabilities is given by

\[
c_{\text{int}95} = 1.96/N \times (n(N-n)/N)^{0.5}
\]

(3)

We analyze the foreshock probabilities as a function of (1) mainshock magnitude – what are the chances for same sized mainshocks, 0.1 magnitude units larger, 0.2, etc. up to 2.0 units of magnitudes between foreshock and mainshock, (2) time between foreshock and mainshock – testing how the probabilities decrease with increasing time and (3) epicentral distance between foreshock and mainshock – testing how much more likely events are to follow in the direct vicinity of the foreshock compared to greater distances.

Because Merrifield et al. (2004) find differences in foreshock probabilities between seismicity in the TVZ and outside, we treat the subcataologue of potential foreshocks located in the TVZ separately. We define earthquakes to be in the TVZ if their epicenters lie in the triangle between (longitude/latitude) points (175.85 / -37.0), (175.55 / -39.29), and (177.4 / -37.5) (Fig. 1a).

We fit the data using maximum likelihood and optimize the parameters of the Gutenberg-Richter relation (1944) for the mainshock magnitude dependence, the modified Omori law (Utsu (1995)) for temporal decay, and a power law for the spatial decay. Two sets of parameters are determined, applicable for TVZ and non-TVZ seismicity.

Mainshock magnitude dependence

We assume that the magnitude differences between foreshocks and mainshocks follow the Gutenberg-Richter relation (1944), which is supported by a linear decay in a logarithmic plot of cumulative probabilities for different magnitude differences over the range of magnitudes with sufficient data (up to magnitude differences of 1.6 units (Non-TVZ) and 1 unit (TVZ)) (Fig. 2). Accordingly, the data can be fitted through a function of the type:

\[
P(\Delta M) = 10^{A \cdot B^{M_0+\Delta M}}
\]

(4)
where $\delta M$ is the difference in magnitude between the foreshock and the subsequent mainshock, and $P(\delta M)$ is the probability that an event will be followed by a second one of $\delta M$ magnitude units larger. $A'$ can be calculated by taking the observed probability for $\delta M=0$, i.e. the probability for same sized mainshocks within 5 days and 50km radius, which has been calculated from the catalogue (0.018 for Non-TVZ and 0.092 for TVZ). For simplicity it is assumed to have no errors. Substitution gives the model equation:

$$P(\delta M) = A'10^{B'\delta M}$$ (5)

where $A = 10^{A'\cdot B'Mfs}$. The fit function therefore depends on the parameter $B$ only, which is related to the $b$-value from the Gutenberg-Richter relationship but is not identical. To fit the function, we follow the method of Aki (1965):

$$B = \frac{\log_{10}(e)}{(Mav - Mc)}$$ (6)

where $e$ is the base of the natural logarithm, $Mav$ is the average magnitude in the set, and $Mc$ is the cutoff magnitude for the catalog, here 3.95.

**Distance dependence**

We assume a probability density function for the distance decay of

$$P(r) = \frac{a}{(r + c_r)^p}, \quad 0 \leq r \leq r_{max}$$ (7)

where $r$ is the epicentral distance between foreshock and mainshock, i.e. all events are treated as point sources and depth ignored, $r_{max}$ is the maximum distance window in the foreshock definition, in this study $r_{max}=50km$, $a$ is a normalizing factor and a function of $c_r$ and $p_r$ to be determined by solving

$$\int_0^{r_{max}} P(r) \, dr = 1$$ (8)

and $c_r$ and $p_r$ are parameters of the distribution to be determined by the method of maximum likelihood. For this we perform a 2d grid search for the values of $c_r$ and $p_r$ that give a maximum of the log likelihood function:

$$L_{max} = \max(L = \log \prod_{i=1}^{n} P(r_i))$$ (9)

where the product is taken over all $n$ data points of $r$, i.e. the distance values for all foreshock-mainshock pairs. The maximum is determined over all values of the
parameters \( p, c_r \). In a second analysis we fix the parameter \( c_r \), which can be interpreted as the location accuracy and optimize for the decay exponent \( p \), only.

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

\[
W = 2 \left\{ L_{\text{max}} - L(\text{other specified parameters}) \right\} \quad (10)
\]

Then \( W \) is asymptotically distributed as \( \chi^2 \) with \( k \) degrees of freedom, where \( k \) is the number of parameters being fitted – here two or one. For two parameters, the 95 percentile of \( \chi^2 (2) \) yields \( W=5.99 \), and so the 95\% confidence region is bounded by the contour for \( L_{95}=L_{\text{max}} \cdot W/2 \), or \( L_{95}=L_{\text{max}} \cdot 3 \) (Fig. 3). The corresponding reduction for one degree of freedom is 1.92.

**Time dependence**

We fit the data to a probability density function in time, which describes the modified Omori law (Utsu et al. (1995)) and in type is identical to the pdf assumed for the distance decay.

\[
P(t) = \frac{d}{(t+c_i)^{p_i}}, \quad 0 \leq t \leq t_{\text{max}} \quad (11)
\]

\( d \) is the scaling factor equivalent to \( a \) in equation (7), \( t \) is the time since the potential foreshock in days, \( t_{\text{max}}=5 \) days, and \( c_i \) is a constant that avoids a singularity for \( t=0 \). We follow the same fitting procedure as described above.

**Fitting decay parameters after linear binning**

In an earlier approach (Tormann (2005)), we counted all earthquakes that had been followed by an equal or bigger one on the first day, on the second, on the third, etc., to the 5th day, and within 10km epicentral distance, 10-20km, and so on to 40-50km. The decay parameters were determined on the binned dataset.

To allow for enough events for statistical analysis, we binned the data in several ways. In time, we considered integer days after the foreshock time. We started the count at the exact foreshock time (e.g., if the event occurred at 09:13 UT on day 247, the first time interval ends at 09:12:59 on day 248). In space, we considered rings of 10 km radius. To account for the fact that the area increases with the radius of the ring, we considered spatial probabilities in terms of the number of events per unit area. We displayed the foreshock probabilities as a function of time and epicentral distance in the middle of each bin.

We fitted these data points using nonlinear regression (software from www.graphpad.com) for a least squares fit to the above equations: Gutenberg-Richter, modified Omori law and inverse power law decay with distance. We show here that such an approach biases the latter part of the time series and results in more rapid apparent decays with distance and time than using unbinned time series. Therefore, in our discussion we present the results from the maximum likelihood fitting to the unbinned dataset.
Results

The initial catalogue had 7368 events, reduced to 2354 when the aftershocks were removed. In the remaining, de-clustered catalogue, 120 are located in the TVZ, 29 of which are identified as foreshocks. Out of the 2234 events outside the TVZ, 139 are followed by bigger events close in time and space. This translates into an overall 6.2±1.0% foreshock probability for non-TVZ events, which is a close reproduction of Merrifield et al.'s (2004) results, and a 24.2±7.7% probability for TVZ events, which is about twice the value that Merrifield et al. (2004) determined. The difference can be explained by the different definitions of TVZ seismicity used in the two studies: we cut the New Zealand catalogue into two parts, the TVZ subcatalogue and the rest of the country. To reduce boundary effects, we only require the foreshock to be situated within the borders of the TVZ. We then search for all events in the New Zealand earthquake catalogue that fulfill the foreshock-mainshock search criteria. Merrifield et al. (2004) analyze the New Zealand earthquake catalogue by using a lattice of 0.1 degree spacing and for each grid node searching for foreshock-mainshock pairs within a circle of 100 km radius. As the TVZ is only about 50 km wide and 250 km long, nearly all lattice points within that area are assigned probability values dominated by activity outside the TVZ, smearing the high values from within the zone.

Mainshock magnitude dependence

The maximum likelihood fit calculates $B_{TVZ}=1.5±0.5$ and $B_{Non-TVZ}=1.0±0.2$.

Distance dependence

The middle panel of Fig. 3 gives the results of fitting the distance function parameters, $p_r$ and $c_r$, by maximum likelihood. The 95% confidence contour does not close. Moreover, the figure shows a strong correlation between $p_r$ and $c_r$. The problem arises from $c_r$, which is effectively indeterminate for the small datasets we have. Accordingly we elected to fix $c_r$ to a reasonable value of 5km, for both TVZ and non-TVZ. This value for $c_r$ gives values for $p_r$ of 1.7±0.6 (TVZ) and 0.9±0.2 (non-TVZ). These are shown in the lower panel of Fig. 3, which plots log-likelihood v. $p_r$, and in the lower row of Fig. 4. The horizontal line at a level of (maximum log-likelihood-1.92) gives the 95% confidence interval for $p_r$ as the intersection with the curve (Fig. 3). The normalization factor $a$ is 2.66 for TVZ and 0.31 for the rest of the country.

Time dependence

The data can be modeled for $c_t(TVZ)=0.0003$, $p_t(TVZ)=0.9±0.2$, $d(TVZ)=0.14$, $c_t(NonTVZ)=0.0016$, $p_t(NonTVZ)=0.83±0.09$, and $d(NonTVZ)=0.17$ as shown in Table 2 and Fig. 4, upper row.

Stationarity over time

In order to make sure that the parameters derived for the foreshock model do not depend on the time period of the catalogue chosen to calculate them, we repeat the optimization for different subsets of the catalogue. We cut the data in 1987 (a break in the earthquake rate is recorded in that year; Fig. 1c) and use both parts of the catalogue separately to confirm our results. The time and distance parameters including their 95%
confidence limits are compared in Fig. 5. For both the TVZ and the Non-TVZ, the confidence intervals of each parameter overlap for the three different time periods. Although the estimated mean values seem to be quite different, they cannot be distinguished taking their uncertainties into account. So the choice of time period is not significant, and within the errors of our measurements, the foreshock parameters are stationary over time. The different earthquake occurrence rate in the second half of the catalogue does not influence the results at this level of accuracy.

Linear Binning

Results from linear binning of the exponentially distributed data were substantially different from the results when the exact time intervals were used (Fig. 6; Table 2: Tormann, 2005). The difference can be explained as follows: About half the data is put into the first bin, and therefore the fitting procedure of the binned data is limited to the tail of the real data, which biases the shape of the curve greatly (Fig.6).

Discussion & evaluation

Mainshock magnitude dependence

The B-values of 1.5 ± 0.5 for the TVZ and 1.0 ±0.2 for the rest of New Zealand are close to the b-value from the Gutenberg-Richter relationship for New Zealand, which is about 1.2 (Figure 1b). The B-values found in this study are larger than the values found in previous studies in California and Nevada though, where B of 0.75 ± 0.1 and 0.72 ± 0.04 were determined, respectively (Jones (1985), Savage & DePolo (1993)). Although the differences are not significant at 95% confidence, we find a higher B-value for the Taupo Volcanic Zone than for the rest of New Zealand, similar to the catalogue b-value of 1.28 ± 0.13 within the TVZ reported by Smith and Webb (1986). This was also seen for the volcanic Mammoth/Mono region (0.89) compared to the non-volcanic Nevada region (0.72) (Savage and DePolo (1993)). In the TVZ a high occurrence of earthquake swarms of small to moderate magnitudes is observed while the region does not feature faults capable of producing very large events. This explains the steeper slope of the magnitude frequency distribution, i.e. a higher B-value in the TVZ.

Distance dependence

Decay of foreshock probability with distance was not determined directly in the studies of Jones (1985) and Savage & DePolo (1993), but examining their plots suggests the vast majority of foreshocks occur within 1 or 5 km of the mainshock location, respectively. In Jones' plot, the decay with distance appears to be slightly slower than the temporal decay, (i.e., the exponent in the distance dependence appears smaller than the exponent in the temporal dependence) while in Savage & DePolo's plots the spatial decay is faster than the temporal one, as we see for New Zealand data in this study.

Abercrombie and Mori (1996), suggested that foreshocks in California occur in a larger volume around mainshocks than expected from the size of the mainshock “preparation zone”, i.e. the critical slip distance over which laboratory studies of frictional sliding observe pre-seismic slip before unstable dynamic failure.
Pfizer et al. (2004) used earthquakes $M \geq 2.2$ in California that have been followed by a larger one within two days and found that, up to about 10 km distance from the mainshock, foreshock incidences decrease roughly as a power-law decay. At greater distances the foreshock incidences become erratic, and observed seismicity is dominated by background seismicity.

Analyzing foreshock sequences in Japan, Ogata et al. (1995), found an approximate inverse power law of foreshock intensity (i.e. expected number of events per unit area) decreasing with distance from the mainshock with exponent of -1.32 for $1 \leq r \leq 35$ km.

Thus, different regions may yield different distance decay parameters. The stronger localization in the TVZ compared to the rest of New Zealand may be caused by the high seismic attenuation in the TVZ (e.g. Salmon et al. (2002), Eberhart-Phillips and McVerry (2003)). If the mainshocks are caused by dynamic triggering from the foreshocks, which then continue to populate new sequences via the aftershock triggering mechanisms (e.g., Anderson et al. (1994), Jaume (1999)), then the higher attenuation would likely limit the triggering to shorter distances than in the rest of the country.

**Time dependence**

The exponents of 0.9±0.2 for the TVZ and 0.83±0.09 for non-TVZ as determined for the time decay in this study are similar to previous studies that looked into foreshock probability decay over time after an event. The probability behavior is similar to the frequency decay of an aftershock sequence, modeled by a modified Omori’s law $1/(t+c)^p$ (Utsu et al. (1995)) where $c$ is small and $p$ is close to one. Jones (1985) and Savage and DePollo (1993) determined $p$ for Southern Californian and Western Nevada’s foreshock probabilities, and obtained values of 0.9 and 1.0, respectively.

According to Eberhart-Phillips (1998), New Zealand aftershock probabilities do not show a different falloff behavior compared to other regions in the world. She analyzed 17 New Zealand aftershock sequences and although the parameter values varied greatly between the individual sequences, the median values turned out to be very similar to the generic California values as determined by Reasenberg and Jones (1989). Of special interest for comparison with this foreshock probability analysis are the parameters $b$ from the Gutenberg-Richter equation, found to be 0.91 in California (Reasenberg and Jones (1989) and 0.98 in New Zealand (Eberhart-Phillips, 1998), and the temporal decay parameter $p$ from the modified Omori’s law, found to be 1.08 in California (Reasenberg and Jones, 1989) and 1.05 in New Zealand (Eberhart-Phillips, 1998).

New Zealand’s foreshock probabilities follow this generic law with slightly smaller decay exponents compared to Californian foreshocks and compared to New Zealand aftershocks.

**Model equations**

Jones (1985) analyzed the Southern Californian foreshock data for mainshock magnitude distribution and time differences between foreshock and mainshock and described her findings in an equation which calculates the probability $P(M_{ms}, t)$ that a mainshock of $M \geq M_{ms}$ will occur within one hour after time $t$ in hours after an earthquake of $M=M_{0}$ as

Tormann et al. – Time, distance and magnitude dependent foreshock probability model for New Zealand
\[ P(M_{ms}, t) = A * 10^{B(M_{ms} - M_0)} * (t + I)^{p} \]  \hspace{1cm} (12) \\

She determined \( A = 0.016, B = -0.75 \pm 0.1, \) and \( p = 0.9. \)

Savage and DePolo (1993) modeled the data with the same equation as introduced by Jones (1985) and obtained the best fit for \( A = 0.012 \) (Nevada), \( A = 0.019 \) (Mammoth/Mono), \( B = -0.72 \pm 0.04 \) (Nevada), \( B = -0.89 \pm 0.04 \) (Mammoth/Mono) and \( p = 1.0 \) for both regions.

We analyze New Zealand foreshock probabilities and find equations to model their decay behaviour with magnitude difference, time and distance. Multiplying the three separate equations gives a model that calculates the probability \( P(t, r, \delta M) \) that at time \( t \) after an event and distance \( r \) from the potential foreshock a mainshock with magnitude \( M_{fs} + \delta M \) will occur as:

\[ P(t, r, \delta M) = a * d * P(\delta M = 0) * 10^{(-B*\delta M)} * (t + c_i)^{-p_i} * (r + c_r)^{-p_r} \]  \hspace{1cm} (13)

or

\[ P(t, r, \delta M) = P_0 * 10^{(-B*\delta M)} * (t + c_i)^{-p_i} * (r + c_r)^{-p_r} \]  \hspace{1cm} (14)

where \( P_0 = a * d * P(t=5, r=50, \delta M=0) \), with two parameter sets for the TVZ and the rest of New Zealand (Table 2).

Conclusion

Temporal, spatial and magnitude dependent analysis of New Zealand foreshock probability data shows similar decay properties as observed in other regions of the world. Linear binning of data that is exponentially distributed biases the results greatly and leads to completely different parameter values.

Acknowledgements

We thank the New Zealand Earthquake Commission for funding this study. Extensive discussions with Annemarie Christophersen as well as David Vere-Jones and David Rhoades have helped the progress on this work substantially. John Townend provided comments on the thesis on which this work is based. Reviews from M. Gerstenberger and an anonymous reviewer were very helpful. Some plots were produced using ZMAP and GraphPad Prism version 4.03 for Windows, GraphPad Software, San Diego California, USA, www.graphpad.com.
References


Jones, L.M., Console, R., Di Lucio, F. et al. (1999). Are foreshocks mainshocks whose aftershocks happen to be big? In press *Bulletin of the Seismological Society of America*


### Tables

<table>
<thead>
<tr>
<th>Mag</th>
<th>4.0</th>
<th>4.5</th>
<th>5.0</th>
<th>5.5</th>
<th>6.0</th>
<th>6.5</th>
<th>7.0</th>
<th>7.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>R(km)</td>
<td>35.5</td>
<td>44.5</td>
<td>52.5</td>
<td>63</td>
<td>79.4</td>
<td>100</td>
<td>125.9</td>
<td>151.4</td>
</tr>
<tr>
<td>T(days)</td>
<td>22</td>
<td>42</td>
<td>83</td>
<td>155</td>
<td>510</td>
<td>790</td>
<td>1737</td>
<td>3981</td>
</tr>
</tbody>
</table>

Table 1: For New Zealand modified declustering windows used in the Gardner & Knopoff (1974) algorithm. The wide windows are designed to capture as many aftershocks as possible, rather removing non-related events than leaving aftershocks in the catalogue.

<table>
<thead>
<tr>
<th></th>
<th>LS, binned data</th>
<th></th>
<th>ML, cumulative data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TVZ</td>
<td>NonTVZ</td>
<td>TVZ</td>
</tr>
<tr>
<td>$P_0$</td>
<td>0.05</td>
<td>0.0063</td>
<td>0.033</td>
</tr>
<tr>
<td>$B$</td>
<td>1.5±0.05</td>
<td>1.32±0.03</td>
<td>1.5±0.5</td>
</tr>
<tr>
<td>$c_L$</td>
<td>0.0009±0.03</td>
<td>0.003±0.03</td>
<td>0.0003±0.004</td>
</tr>
<tr>
<td>$c_R$</td>
<td>5.0±0.003</td>
<td>5.003±0.09</td>
<td>5 (fixed)</td>
</tr>
<tr>
<td>$p_L$</td>
<td>2.8±0.8</td>
<td>1.8±0.2</td>
<td>0.9±0.2</td>
</tr>
<tr>
<td>$p_R$</td>
<td>3.1±0.1</td>
<td>2.8±0.2</td>
<td>1.7±0.6</td>
</tr>
</tbody>
</table>

Table 2: Foreshock probability model parameters as determined in this study for the TVZ seismicity and the rest of New Zealand, contrasting the different methods of parameter calculation, LS = Least Squares fitting of binned data, ML = Maximum Likelihood fitting of cumulative data. The column in bold letters is the trustworthy and final parameter set describing New Zealand foreshock probability behaviour.
Figure Captions

Fig.1a: New Zealand plate tectonic setting:
Subducting Pacific Plate under the Australian Plate (Hikurangi Subduction Zone) in the North, subducting Australian Plate under the Pacific Plate (Fiordland Subduction Zone) in the South.

The Alpine Fault (South Island) accommodates most of the strike slip motion and is capable of producing large earthquakes. The Taupo Volcanic Zone (TVZ) on the North Island is an extensional geothermally active volcanic area with frequent small to moderate earthquake swarms.

The dots represent the recorded New Zealand seismicity (GeoNet catalogue) between 01 Jan 1964 and 31 May 2007 for magnitudes equal to or greater than M=4.0.

Fig.1b: Cumulative and non-cumulative frequency magnitude distribution of New Zealand seismicity data between 01 Jan 1964 and 31 May 2007, depth ≤ 40km, M≥2.0. The slope has been manually fit to a b-value of 1.2, Mc= 4.0.

Fig.1c: Cumulative number of earthquakes with time, showing a break in seismicity rate in 1987.

Fig. 2: Cumulative foreshock probabilities with increasing magnitude differences between foreshock and mainshock, for TVZ and Non-TVZ. The dotted lines represent the model fit by the method of maximum likelihood.

Fig. 3: Maximum likelihood fitting for distance and time for the TVZ (left panels) and non-TVZ (right panels). Upper panels: log-likelihood contours for time behaviour; middle panels: log-likelihood contours for distance with c, free; bottom panels: log-likelihood v. p, for c, fixed to 5km. The heavy contour is the 95% confidence interval for the two-parameter plots. The horizontal line in the bottom panels gives the 95% confidence interval for p, (for c, = 5km).

Fig. 4: Foreshock probability dependence on time (upper row) and epicentral distance (lower row) between foreshock and mainshock for TVZ (left) and Non-TVZ (right). The stepped lines are the data, the smooth lines the models with 95% confidence bands on the parameters fitted by maximum likelihood.

Fig. 5: Time and distance exponents calculated for different periods of the catalogue for Non-TVZ (left) and TVZ (right). The stars show the maximum likelihood results, the rectangles visualize the 95% confidence limits on the parameters. They all overlap, so the parameters do not vary significantly over different subsets of the catalogue.

Fig. 6: Comparison between least square fit of linearly binned data (left) and maximum likelihood fit of non-binned cumulative data (right). The inset on the plot on the right shows how much data is put into the first one data point in the case of linear binning.
Figures

Fig. 1a:

New Zealand tectonic setting and seismicity

Fig. 1b:

NZ catalogue, 1964-2007, Mmin=2.0, depth=40km

Cumulative / Non-cumulative Number

Magnitude
Fig 1c:

Fig. 2:

Cumulative Foreshock Probabilities

```
Observed Probability

0.5 1 1.5 2

TVZ
NonTVZ

B=1.5
B=1.1
```

```
Magnitude difference (units of M)
```
Fig. 3:

ML fit TVZ foreshock times

ML fit non-TVZ foreshock times

ML fit TVZ foreshock distances

ML fit non-TVZ foreshock distances

Log Likelihood $\nu, p$ for $c$ fixed = 5.00

Log Likelihood $\nu, p$ for $c$ fixed = 5.00
Fig. 4:

Distribution of times between TVZ foreshock-mainshock pairs

- $p = 0.99 \leq 0.10 \div 0.0033$ 

Distribution of distances between TVZ foreshock-mainshock pairs

- $p = 0.73 \leq 0.50 \div 0.05$

Distribution of times between non-TVZ foreshock-mainshock pairs

- $p = 0.95 \leq 0.10 \div 0.0016$

Distribution of distances between non-TVZ foreshock-mainshock pairs

- $p = 0.88 \leq 0.24 \div 0.03$

Fig. 5:

Time and distance decay exponents for different periods of the catalogue (Non-TVZ)

Time and distance decay exponents for different periods of the catalogue (TVZ)
Affiliations and addresses of the authors

Thessa Tormann
Swiss Seismological Service
Institute of Geophysics
ETH Zurich, HPP P3
Schaffmattstrasse 30
CH-8093 Zurich
Switzerland
e-mail: thessa.tormann@sed.ethz.ch

Martha Kane Savage
Institute of Geophysics
Victoria University of Wellington
Box 600, Wellington
New Zealand
e-mail: Martha.Savage@vuw.ac.nz

Euan C.G. Smith
Institute of Geophysics
Victoria University of Wellington Box 600, Wellington
New Zealand
e-mail: Euan.Smith@vuw.ac.nz

Mark Williamson Stirling
GNS Science
PO Box 30368, Lower Hutt
New Zealand
e-mail: m.stirling@gns.cri.nz

Stefan Wiemer
Swiss Seismological Service
Institute of Geophysics
ETH Zurich, HPP P5
Schaffmattstrasse 30
CH-8093 Zurich
Switzerland
e-mail: wiemer@sed.ethz.ch
APPENDIX 2  FORESHOCK RATES FROM AFTERSHOCK ABUNDANCE
(26 PAGES)

Foreshock rates from aftershock abundance

Annemarie Christophersen and Euan G.C. Smith

Abstract

The question whether foreshocks behave like mainshocks that happen to have larger aftershocks is important for the understanding of earthquake nucleation and the modeling of earthquake clustering. Many models, including the STEP model for forecasting short-term earthquake hazard in California, and the ETAS model for the description of earthquake clustering, imply that all earthquakes of magnitude $M$, including foreshocks, can trigger earthquakes according to a magnitude dependent triggering ability $10^{0.2M}$. The magnitudes of the triggered earthquakes are proportional to $10^{0.2M}$, according to the Gutenberg-Richter relation for the magnitude-frequency distribution, and therefore can be larger than the magnitude of the initiating event. However, numerous studies observe a lower foreshock occurrence than predicted from aftershock models. In this paper we show that for homogeneous earthquake catalogues, and a consistent definition of foreshocks, foreshock rates can be predicted from aftershock characteristics. For Southern California the probability that the initiating event of an earthquake cluster in the magnitude range $2.0 - 4.5$ is followed by at least one larger event within one day is around 4%. This agrees well with corresponding data from the global CMT catalogue where the probability that the initiating earthquake of a cluster is followed by at least one larger event within 30 days is around 3%. These foreshock rates are significantly lower than previously reported and we discuss possible reasons for this.

Introduction

Earthquakes cluster in space and time. In retrospect, the largest earthquake in a cluster is called the mainshock. The earthquakes preceding a mainshock are named foreshocks, and the earthquakes following a mainshock aftershocks. Clusters with no discernible mainshock are often referred to as swarms. No physical differences between fore-, main- or aftershocks are known. Therefore spatial and temporal vicinity define whether an earthquake is part of a cluster, or occurred independently. Most earthquake models assume that foreshocks and mainshocks have the same triggering capability scaling with earthquake size. However, there are contradicting observations. Here we investigate whether the occurrence of foreshock-mainshock pairs is consistent with mainshock-aftershock triggering. Understanding possible differences between foreshocks and mainshocks would help understanding the process of earthquake nucleation and allow better modeling of short-term earthquake hazard.

Two empirical relationships are very important for the description of earthquake clusters: The Omori-Utsu law and the Gutenberg-Richter relationship. The Omori-Utsu law describes the decay of aftershock activity with time $t$.

$$\frac{dN}{dt} = K (t_c + t)^p,$$

(1)

Christophersen & Smith, Foreshock rates from aftershock abundance
where $dN$ is the number of earthquakes in the time interval $dt$; $K$ is a parameter that is proportional to the aftershock productivity; $p$ describes the decay and takes values around 1.0; and $t_c$ stands for a small time interval just after the mainshock and avoids a singularity for $t=0$ [e.g. Utsu, et al., 1995]. The Gutenberg-Richter relation describes the magnitude-frequency distribution

$$N(M) = 10^{a-bM},$$ (2)

where $N(M)$ is the number of earthquakes of magnitude $M$ and $a$ and $b$ are parameters [Gutenberg and Richter, 1944; Ishimoto and Iida, 1939]. The equations can be combined into an empirical rate function $\lambda(t,M)$ for aftershocks of magnitude $M$ at time $t$ after the mainshock of magnitude $M_m$ [Utsu, 1969], now using the cumulative magnitude frequency distribution and therefore changing the notation to $a'$,

$$\lambda(t,M) = 10^{a'+b(M_{m}-M)}/(t+c)^p \quad (a', b, c, p; \text{constants}).$$ (3)

The rate can be used to calculate the probability that at least one earthquake of magnitude $M$ or above occurs in the time interval $[t_1, t_2]$ ([Reasenberg and Jones, 1989; Reasenberg and Jones, 1994])

$$P(M) = 1 - e^{-\int_{t_1}^{t_2} \lambda(t,M) \, dt}.$$ (4)

For California, the parameters for equations (3) and (4) were derived by analyzing 62 Californian earthquake sequences and found to be $a' = -1.57$, $p = 1.08$, $b = 0.91$ and $c = 0.05$ [Reasenberg and Jones, 1989]. These parameters have become known as the generic Californian aftershock model parameters [Gerstenberger, et al., 2005]. Equation (4) yields a probability of 10.5% that an earthquake is followed by one of the same magnitude or larger within one week. However, the observed probability for an earthquake of magnitude $M \geq 3.0$ to be followed by a larger earthquake within 5 days and 10 km is 6.0 ± 0.5% in Southern California [Jones, 1985], which is significantly lower than predicted from the aftershock model. To determine foreshock rates, Jones first declustered the catalog by removing all smaller earthquakes within magnitude dependent space and time windows. Each earthquake qualified as a foreshock if it was followed by a larger earthquake within a chosen time and space window. Thus some clusters would have multiple foreshocks increasing the foreshock probability. Jones provides a table with numbers of foreshocks [Jones, 1985]. If every sequence is only allowed to have one (initiating) foreshock, the foreshock probability drops from 6% to about 4%.

The discrepancy between the aftershock model predicting a 10% probability of an earthquake at the beginning off a cluster to be followed by at least one larger [Reasenberg and Jones, 1989] and the observation of about 4% of at least one larger one [Jones, 1985] could indicate that foreshocks might have a lower triggering capability than mainshocks. We note that the studies did not select clusters in a consistent way. Furthermore, the Reasenberg and Jones parameter were determined from 62 well recorded and thus productive aftershock sequences without taking into single events or small clusters to determine the triggering capability. In contrast the foreshock
probabilities were determined from the complete catalog.

A more detailed study to assess whether an earthquake in the vicinity of a fault structure is a foreshock to a characteristic earthquake found a larger range of foreshock probabilities [Michael and Jones, 1998], which are consistent with the Reasenberg-Jones aftershock model. However, specific case studies cannot be sensibly compared to more general studies of the catalog.

In an attempt to evaluate whether the Californian generic model accurately predicted the probability for large events, a global study was undertaken [Reasenberg, 1999a; Reasenberg, 1999b]. The probabilities of an earthquake of magnitude \( M \geq 5 \) being followed within 75 km and 10 days by an earthquake of \( M \geq 5, M \geq 6, \) and \( M \geq 7, \) were 7.5\%, 2.3\% and 0.4\%, respectively. The first value is about 75\% of the Californian model while the latter two are about double the Californian model. Reasenberg’s conclusion was that the actual global probability for an earthquake to be followed by a larger event was 15\% within 10 days [Reasenberg, 1999b].

Other regions in which the Reasenberg-Jones model was applied include Italy and New Zealand. In Italy, the model parameters were similar to California [Jones, et al., 1995; Lolli and Gasperini, 2003]. The foreshock probabilities were determined with yet another method. A foreshock was an earthquake larger than a threshold magnitude \( M_f \) that occurred after a period of quiescence in a temporal and spatial window of length \( T_i \) and \( R_i \) and was followed by a larger earthquake within a spatial and temporal window of length \( T_2 \) and \( R_2 \) [Console and Murr, 1993]. A statistical analysis found best values for \( T_i, R_i, T_2 \) and \( R_2 \) of 80 days, 140 km, 2 days and 30 km, respectively. 63 earthquakes qualified as potential foreshocks of magnitude 3.0 and larger, of which 6 were followed by mainshocks of magnitude 4.0 or greater, resulting in a probability of 9.5\%. In another analysis for central Italy, the rate of magnitude \( M \geq 3.3 \) being followed by \( M \geq 3.8 \) earthquakes within 2 days and 30 km was 7\% [Di Luccio, et al., 1997]. As these rates are for a magnitude difference of 1 and 0.5 magnitude units, the foreshock rates in Italy are higher than predicted by the Reasenberg-Jones model.

For New Zealand, the analysis of 17 earthquake sequences found a wide variation in model parameters with the averages being similar to California and Italy [Eberhart-Phillips, 1998]. From the range of model parameters, Savage and Rupp determined foreshock probabilities and compared them to observed foreshock rates as determined in the same way as in California [Savage and Rupp, 2000]. The observed foreshock probability within 5 days and 30 km was 4.5±0.7%, 1.5±0.4%, and 0.8±0.3% for the magnitude of the second earthquake larger, larger by half a magnitude unit and larger by one magnitude unit, respectively. This compared to 10.7%, 3.4% and 1.1% prediction from the Reasenberg and Jones model with New Zealand parameters. Applying the large range of observed parameters gave a wide range of possible foreshock rates, e.g. 0.21-52% for the foreshock probability within 5 days and 30 km.

Despite the discrepancies between model predictions and foreshock observations the Reasenberg-Jones model still is the core of the Short-Term-Earthquake-Probability model (STEP) that forecasts real-time earthquake shaking in California [Gerstenberger, et al.,]
2005]. If the hypothesis of a single triggering mechanism for all earthquakes was not correct, it would be advisable to separate the forecasting of larger and smaller events than those that had already occurred.

Discrepancies in triggering capability for different types of earthquakes have also been observed in a different model context. The Epidemic Type Aftershock Sequence (ETAS) model is a stochastic point process model that accounts for stationary seismicity and secondary aftershocks and describes complete earthquake catalogs as one family of earthquakes [Console, et al., 2003; Ogata, 1988; Zhuang, et al., 2002]. In the ETAS model all earthquakes have the same magnitude dependent triggering capability. The ETAS model distinguishes between triggered earthquakes and background seismicity. The first earthquake of a cluster is part of the background because it was not triggered by another event. Applying the ETAS model for the de-clustering of earthquake catalogs reveals a lower triggering capability for the background than for triggered events [Zhuang, et al., 2008]. This observation is consistent with foreshock numbers being lower than predicted from the Reasenberg and Jones model. It might not necessarily contradict the idea of a single triggering mechanism for all earthquakes because the mainshock may change the stress regime causing more earthquakes to happen and resulting in an apparent higher triggering capability. On the other hand, the discrepancies that the ETAS model finds between simulated and real data could hint at features in the real catalog that the ETAS model does not capture, e.g. the spatial distribution of aftershocks and the spatially and temporally varying magnitude of completeness, in particular following a large earthquake.

The hypothesis of the single triggering mechanism was recently validated for global and Californian data [Felzer, et al., 2004]. In particular, foreshock rates for California could be predicted from aftershock characteristics by assuming that an initial earthquake of magnitude $M$ triggers $N$ subsequent events according to $N = c10^{ad}$, where $c$ is a productivity constant and $a$ the growth parameter. The magnitude-frequency distribution (equation 2) determines the size of the triggered earthquakes. For the analysis $a$ and $b$ were both set equal to 1 [Felzer, et al., 2004]. However, various values for $a$ and $b$ have previously been reported. In California the $b$-value is usually found to be somewhat smaller than 1.0 for the complete catalog, e.g. 0.91 in the generic model [Reasenberg and Jones, 1989]. In a global study, the value $a$ was found to be 1.0 for most tectonic settings [Christophersen, 2000]. For California it was observed to be 0.8 [Helmstetter, 2003], for Italy 0.5 [Console, et al., 2003] and for Japan 0.8 on average with values varying between 0.2 and 1.9 [Guo and Ogata, 1997]. Hainzl, et al. (2008) showed that some of the smaller $a$ values could be explained by assuming spatial isotropy in the space dependent ETAS model.

In summary, there are contradicting observations on whether the triggering capability of foreshocks differs from the one of later earthquakes in a cluster. The clarification of this issue has been hampered by differences in data selection for foreshocks and aftershocks, the heterogeneity of earthquake catalogs, in particular the varying completeness magnitude in time and space, and simplified assumptions, e.g. that $a$ and $b$ take the same value. Here we want to investigate whether the hypothesis of the same triggering capability of fore-and mainshocks holds when $a$ and $b$ are determined from the
data. For this purpose we first select two high quality earthquake catalogs and identify homogenous sub-sets of the catalogs. We then define earthquake clusters in time and space by various methods and a range of parameters. We introduce a new formulation for the number of aftershocks, which we call abundance. We assume that the first earthquake in an earthquake cluster has the same triggering capability as a mainshock. By combining mean abundance with the magnitude-frequency relation we can calculate how often larger earthquakes can be expected. We compare the model with the observed foreshock rate and conclude that foreshock rates are consistent with mainshock-aftershock triggering but depend on the clustering algorithm. Finally we discuss possible reasons for observed discrepancies between past studies.

Data selection

Earthquake catalogs are generally heterogeneous in space and time due to changes in the seismic station network and processing procedures ([Woessner and Wiemer, 2005] and references therein). For our purpose of counting aftershocks and determining foreshock rates, magnitude consistency and completeness of data are of particular importance. The completeness magnitude $M_C$ is the magnitude for which all events are reported in the catalog. Immediately following a large earthquake the completeness magnitude generally increases (e.g. [Kagan, 2004]), which we need to take into account in our data analysis.

We selected two earthquake catalogs for their high quality and data consistency: The regional Southern Californian Earthquake Data Center (SCEDC) with local magnitude $M_L$ and the global central moment tensor (CMT) catalog with moment magnitude $M_w$. For Southern California we selected earthquakes in the period from 1984 – 2006 inclusive, during which all magnitudes have been calculated and determined consistently [SCEDC, 2007]. Figure 1 shows a completeness map using the entire magnitude range method [Woessner and Wiemer, 2005] calculated within the software package zmap [Earthquake Statistics Group, 2007; Wiemer, 2001], using 100 bootstraps on a 0.1x0.1 degree grid with 100 earthquakes per grid node and 30 earthquakes above $M_C$. Outside a major aftershock sequence, the completeness magnitude for the catalog is 3.0, and 2.0 within the polygon displayed in figure 1. Table 1 describes four sub-sets of the SCEDC catalog and the global data. Data sets 1 and 2 cover the complete region of the catalog, and 3 and 4 the data within the polygon. To minimize the problem of missing aftershocks following a large earthquake, we removed from data set 1 and 3 all clusters that have at least one earthquake of magnitude 6 and above, using the Reasenberg declustering code [Reasenberg, 1985] as further outlined below. The resulting data sets are 2 and 4. We calculated the $b$-value by maximum likelihood method [Aki, 1965], where $b$ is inverse to the average magnitude above the completeness magnitude. The average magnitude is increased by the missing small aftershocks, and consequently the $b$-value is higher for data sets 2 and 4 than data sets 1 and 3. Our $b$-value of around 1.12 for Southern California is larger than generally reported (e.g. the generic values of 0.91, [Reasenberg and Jones, 1989]), and we suspect that other data sets also suffer from missing small events, and consequently have higher average magnitudes and lower $b$-values.

The CMT catalog started in 1976 and has been consistent in data processing ever since. However, the increase in station coverage with time has led to lowering of the
completeness magnitude. It has been shown that the first 10 years or more significantly under sampled earthquakes smaller than magnitude 5.6 [Kagan, 1997]. A spatial completeness analysis of the data from 1983 to 2002 showed spatial variation in completeness from magnitude 5.3 along the Aleutian arc to 6.0 close to Antarctica [Woessner and Wiemer, 2005]. Figure 2 shows the completeness magnitude with time for the whole CMT catalog calculated by the entire magnitude range method [Woessner and Wiemer, 2005], using 100 bootstraps with 500 earthquakes per sampling windows, an overlap of 10 earthquakes between windows and magnitudes in 0.1 bins. We selected earthquakes from 1980 of magnitude 5.7 and above. With the large completeness magnitude of 5.7, we are not concerned about missing aftershocks except within the first few hours of a larger earthquake, and thus only start counting aftershocks after 0.1 days.

Defining earthquake clusters

No physical differences are known to exist between foreshocks, mainshocks and aftershocks, and therefore earthquake clusters are usually defined by their closeness in space and time to one another. Most algorithms for identifying clusters have been developed for removing aftershocks, and thus ‘declustering’ of earthquake catalogs. In principle, two approaches for declustering exist: stochastic and deterministic algorithms. The stochastic algorithms like the stochastic declustering based on the ETAS model [Zhuang, et al., 2002] and a new ‘model independent’ clustering algorithm [Marsan and Lengliné, 2008] calculate probabilities for each earthquake to be triggered by a preceding one. 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. Here we focus on the deterministic algorithms where again 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. The most commonly used linking algorithm was developed for California [Reasenberg, 1985] and is available as FORTRAN code on the USGS web-site [http://earthquake.usgs.gov/research/software/#CLUSTER]. The Reasenberg code is based on the Omori-Utsu decay of aftershock activity [Utsu, et al., 1995] and requires a few parameters to be set. The most important parameters are the minimum and maximum look ahead time $r_{\text{min}}$ and $r_{\text{max}}$, the effective completeness magnitude $M_{\text{eff}}$, and the factor for the spatial interaction distance $r_{\text{int}}$. Table 2 shows the values that we used for the Californian and global data respectively as adapted from previous studies (e.g. [Helmstetter, et al., 2007]). We did not study the effect of changing individual parameters in the Reasenberg code.

To investigate the effect of different spatial windows, we developed a hybrid method where we centered the search in space on the largest earthquake within a cluster and used sliding windows of duration $\Delta T$ to link a cluster in time. Figure 3 illustrates the search algorithm: each earthquake in the catalog equal to or larger than the completeness magnitude qualified as a potential initiating earthquake for a cluster. We searched within a magnitude dependent spatial radius as defined in equations 5-7 and time $\Delta T$ for related events. If a larger earthquake occurred within the time $\Delta T$ we used the distance calculated
from the larger magnitude to decide whether the earthquakes were part of the same cluster. If they were, the larger event became the new center for finding related earthquakes. For California we used $\Delta T = 3$ days and 1 day for completeness magnitudes of 3.0 and 2.0 respectively; and for the global data $\Delta T = 30$ days with a completeness magnitude of 5.7. Halving or doubling of $\Delta T$ had no significant effect on the results as shown for the mean abundance in the CMT catalog in table 3.

We applied three different search windows in space: (i) Gardner and Knopoff [Gardner and Knopoff, 1974], (ii) Uhrhammer [Uhrhammer, 1986], which were developed for declustering the Southern and Northern Californian earthquake catalogs respectively, and (iii) Utsu-Seki, a variation on the relation for the growth of aftershock area with mainshock magnitude [Utsu and Seki, 1955]. Gardner and Knopoff’s windows are 22.5, 30, 40, 54, 70, and 94 km for magnitude 3, 4, 5, 6, 7, and 8, respectively. Uhrhammer used windows of 20, 30, 45, 67 and 100 km to remove aftershocks following earthquake of magnitude 5, 5.5, 6, 6.5 and 7, respectively. We fitted those windows by smooth functions and extrapolated the Uhrhammer windows to smaller magnitudes as shown in equations 5 and 6. Utsu and Seki [Utsu and Seki, 1955] studied 40 Japanese mainshock-aftershock sequences in the mainshock magnitude range $6 \leq M \leq 8.5$ and found that the aftershock area increased with mainshock magnitude as $A(M) = 10^{0.12M - 4.02}$, which Utsu later simplified to $A(M) = 10^{M - 3.9}$ [Utsu, 2002]. Assuming a circular aftershock area, we derived a search radius $r(M)$ by taking four times the radius as shown in equation 7.

\[
\begin{align*}
\text{Gardner-Knopoff:} & \quad r(M) = 10^{0.12M + 0.08} \\
\text{Uhrhammer:} & \quad r(M) = 10^{0.35M - 0.44} \\
\text{Utsu & Seki:} & \quad r(M) = 10^{0.50M - 1.60}
\end{align*}
\]

Figure 4 visually compares the different search radii. Gardner-Knopoff is largest for the small magnitudes, with 17 km for magnitude 2, compared to 1.8 km for Uhrhammer and 230 m for Utsu-Seki. The latter two search radii were extrapolated to low magnitudes from magnitude 5 for Uhrhammer and 6.4 for Utsu-Seki, and are likely to be smaller than the location uncertainty in the catalog for small magnitudes. The Gardner-Knopoff search radius increases most gradually and is 125 km at magnitude 9.0 small in comparison to Uhrhammer (513km) and Utsu-Seki (800km). In figure 4 the stars represent the largest distance to the furthest cluster member of the Reasenberg clustered data for mainshocks of each magnitude bin. The data fall mostly between the Uhrhammer and Utsu-Seki window in the magnitude range 2 to 6. In the magnitude range 5.7 to 7.5, in which most the global data lie, all algorithms are fairly similar in space.

**Aftershock abundance and expected foreshock rate**

The purpose of our analysis is to investigate whether foreshocks have the same triggering capability as mainshocks. Thus we model the mean number of aftershocks as a function of mainshock magnitude in a new formulation, which we call abundance. We then assume that the initiating event of a cluster has the same triggering capability as the mainshocks and use the magnitude-frequency distribution to determine how many
earthquakes larger than the initiating one can be expected.

The number of aftershocks and its variation between sequences and regions is not as extensively studied as the Omori law for the decay of aftershock activity with time, even though the integral of aftershock rate over time gives the number of aftershocks. The number of aftershocks depends on the definition of aftershocks in space, time and magnitude. While it is generally accepted that the number of aftershocks grows exponentially with mainshock magnitude as $10^{aM}$, the definition of aftershocks is not always specified. Here, we use different algorithms as described above to define earthquake clusters. We call the largest earthquake in the cluster the mainshock, including single events, and count aftershocks in the time interval $\Delta T$ used in the clustering algorithm. For the Southern California data, we are fairly confident that aftershocks would be detected above the completeness magnitude for earthquakes smaller than magnitude 6.0. In case of the CMT data we exclude the first 0.1 days following the mainshock for counting aftershocks to avoid the issue of missing aftershocks.

We call the number of aftershocks abundance to reflect that we study the triggering capability or the abundance of aftershock sequences. We model the mean abundance as

$$N(M) = 10^{a(M-M_I)}.$$  \hspace{1cm} (8)

This can be derived from the more common formulation $N(M) = c10^{aM}$ [Felzer, et al., 2004] by including the constant in the exponent and shifting the reference for $N(M)$ to 1.0. For $N(M) = 1$, the exponent has to be 0 and thus $M = M_I$ corresponds to the magnitude that has on average one aftershock above the completeness magnitude in the time interval $\Delta T$ and in the space defined by the clustering algorithm. Assuming independence of abundance and the magnitude-frequency relation, the expected number of aftershocks of magnitude $M$ and larger, following a mainshock of magnitude $M$ can be written as

$$N(M) = 10^{a(M-M_I)}10^{-b(M-M_c)} \hspace{1cm} \text{or}$$

$$N(M) = 10^{(\alpha - b)M - \alpha M_I + \beta M_c}$$  \hspace{1cm} (9)

For foreshocks, $N(M)$ will be considerably smaller than 1, e.g. 0.06 for the rate of foreshock occurrence in California [Jones, 1985]. For $\alpha = b$, there is no dependence on the magnitude of the initiating earthquake, i.e. the expected number of larger earthquakes to follow an initiating event is the same for all initiating earthquake magnitudes. For $\alpha < b$, the number of larger earthquakes decreases with magnitude of the initiating earthquake and thus large earthquakes have smaller triggering capability, while for $\alpha > b$, the number of larger earthquakes increases with magnitude of the initiating earthquake and thus smaller earthquakes are less likely to be followed by larger ones.

The mean abundance with different search algorithms

To determine mean abundance from the data, we counted all mainshocks, including all events that had no foreshock or aftershock. Next we counted for each mainshock magnitude all aftershocks in the time window $\Delta T$ used in the clustering algorithm and added up the
aftershocks for all clusters of one mainshock magnitude, using bins of 0.1 magnitude units centered on magnitudes rounded to one decimal place. The mean abundance was calculated by dividing the number of aftershocks by the number of mainshocks.

Figure 5a compares the mean abundance for all four search algorithms for the Southern Californian polygon data (data set 4). The left-hand column of plots shows the magnitude-frequency distribution of mainshocks and the total numbers of aftershocks per mainshock magnitude. We included a reference line for the mean total number of aftershocks $N_{ave}$ for mainshocks of all magnitudes. The most striking difference between the clustering algorithms are the differences in $N_{ave}$. $N_{ave}$ increases from 178 for Utsu-Seki, to 197 for Reasenberg, 240 for Uhrhammer and 289 for Gardner-Knopoff. For Reasenberg and Utsu-Seki windows, the total number of aftershocks per mainshock magnitude bin remains relatively constant with increasing mainshock magnitude, except for a roll-off close to the completeness magnitude. The roll-off can be explained by magnitude uncertainty and the closeness to the completeness magnitude. A magnitude 2.0 earthquake has a smaller chance of having an aftershock, because to be above the completeness magnitude an aftershock would have to have the same magnitude 2.0. Uhrhammer and Gardner-Knopoff also feature the roll-off. In addition, the Uhrhammer and Gardner-Knopoff plots have a bulge above the reference line for mainshocks between magnitude 2.1 and 3.2, which causes the higher $N_{ave}$. The large search radius for small earthquakes in the Gardner-Knopoff algorithm, somewhat less for Uhrhammer, leads to many small earthquakes being linked in clusters. As a consequence significantly fewer magnitude 2 earthquakes are mainshocks in the Gardner-Knopoff algorithm than for example Reasenberg (5243 compared to 8557).

The right-hand column of plots in figure 5a shows the mean abundance. Due to the nature of the magnitude-frequency distribution with few observations at the high magnitude end, error bars would be increasing with mainshock magnitude. The input data in the left-hand column are shown to illustrate how many observations were available for each magnitude bin. The mean abundance falls on a straight line in a log-linear plot. We therefore used a linear regression on the logarithm to the base of 10 of the data to determine the parameters. We tested that the residuals of the model fit were normally distributed, which justified the use of linear regression. We fitted the data in an interval of half a magnitude unit above the completeness magnitude to exclude the roll-off near the completeness magnitude, and we included in the fitting all magnitudes with at least 10 mainshock observations. The 95% confidence interval for the model and the data, as well as for the parameter $\alpha$ can be determined by standard methods for linear regression.

However, our parameter $M_f$ is not fitted directly in the linear regression but is calculated by dividing the intercept from the linear regression by the slope $\alpha$. As these two parameters are dependent on one another, the calculation of a confidence interval is not straightforward. We derived an approximate 95% confidence interval for $M_f$ from the 95% confidence models. $M_f$ is the value for which $\log_{10}(N(M))$ takes the value 0. We found confidence intervals for $M_f$ by solving for $M$ at the 95% model values $\log_{10}(N(M)) = 0$. Figure 5a includes the results for $\alpha$ and $M_f$ and their 95% confidence intervals. Both $\alpha$ and $M_f$ are smallest for Gardner-Knopoff and increase for Uhrhammer, Utsu-Seki and Reasenberg. The latter two agree within the 95% confidence interval of the parameters. The increase in $\alpha$ and $M_f$ for the different search algorithms is linked to the bulge of total

Christopher & Smith, Foreshock rates from aftershock abundance
number of aftershocks for small earthquakes in the left-hand column of figure 5a. Smaller earthquakes in the Gardner-Knopoff and Uhrhammer search algorithm are more likely to have aftershocks. Thus $M_t$, the magnitude that has on average one aftershock above $M_C$, is smaller than for the other methods. As the mean abundance is larger for smaller mainshocks for the Gardner-Knopoff and Uhrhammer search radii, the slope $\alpha$ in the mean abundance fit comes out smaller.

Table 3 summarizes the results of the mean abundance fitting for all data sets. The same trend of increasing $\alpha$ and $M_t$ for decreasing search radii can be seen for the complete Southern California catalog with $M_C=3.0$. However, as the search radii for the larger magnitudes are closer together (see figure 4), and fewer data points are available for the fitting leading to larger confidence intervals, the confidence intervals for $\alpha$ and $M_t$ for the different search algorithms overlap and thus the parameters are not significantly different.

For the global data, the fitting range starts from magnitude 6.2 and goes to magnitude 7.7 – 7.9, depending on search algorithm. In that magnitude range the search radii of the window methods are all relatively similar and cross over as can be seen in figure 4. An effect of the search radii can be seen on the results for $\alpha$ and $M_t$. The parameter $\alpha$ grows for the window methods in the same order in which the growth exponent for the search radii increases in equations 5-7: Gardner-Knopoff, Uhrhammer and Utsu-Seki (see table 3 for results). The parameter $M_t$ is smallest for the larger search radii in the same magnitude range, because more earthquakes were clustered and thus the abundance increased, decreasing $M_t$. For Utsu-Seki and for Uhrhammer $M_t$ is about the same and a little less for Gardner-Knopoff. As a consequence of the overlapping search radii in the magnitude range of data fitting, and relatively large confidence intervals, the differences in $\alpha$ or $M_t$ are not significant for the window methods. In contrast, the Reasenberg clustering algorithm gives different results. Figure 6a shows the comparison of the mean abundance fitting for Utsu-Seki and Reasenberg for the global data, which are the methods that have the largest difference in $\alpha - b_{cut}$ in contrast to having the smallest for the Southern Californian data. $N_{ave}$ of 13 for Reasenberg is only about half of $N_{ave} = 21$ for Utsu-Seki. The total numbers of aftershocks in the Reasenberg algorithm shows a bulge for the smaller magnitudes and consequently $\alpha$ comes out smaller. However, as $N_{ave}$ is so much lower, $M_t$ is significantly higher for Reasenberg than for the Utsu-Seki search algorithm.

The last two rows in table 3 show the results for halving and doubling $\Delta T$ for the Utsu-Seki search algorithm and the global data. The slope $\alpha$ does not change with varying $\Delta T$ but $M_t$ decreases with increasing $\Delta T$, as more time was available for aftershocks to occur. However, as the confidence intervals of the $M_t$ for different $\Delta T$ overlap, halving or doubling does not have a significant effect on the outcome of the mean abundance fitting. The same observation holds for the other window methods and the different data sets.

In summary, the size of the search radius affects the magnitude $M_t$. The larger the search radius, the more earthquakes can become part of the same cluster, increasing the
number of aftershocks, and decreasing the magnitude $M_i$ that has on average one aftershock. Thus the difference in $M_i$ for the spatial window method is largest for the data set 4, where the fitting started at magnitude 2.5 and the differences between the search radii was largest with 19.6 km, 2.7 km and 450 m for Gardner-Knopoff, Uhrhammer and Utsu-Seki, respectively. For dataset 2 with the fitting range $3.5 - 4.9$, the increase in $M_i$ from Gardner-Knopoff to Uhrhammer to Utsu-Seki was much smaller and as fewer data were available for the fitting, the uncertainty was larger and the confidence intervals overlapped. The difference in search radii decreased to 56 km, 53 km and 32 km at magnitude 6.2, at which the fitting of mean abundance started for data set 5. Between magnitude 6.2 and 6.3 the Uhrhammer search radius crossed over with the Gardner-Knopoff and was largest up to magnitude 8.0 from where Utsu-Seki became largest (see figure 4). The difference in $M_i$ was negligible for Uhrhammer and Utsu-Seki and $M_i$ was now larger for Gardner-Knopoff.

The crossovers of search radii for the different algorithms were caused by the different growth exponents of the search radii. The growth exponent affects the parameter $\alpha$ of the mean abundance, which increases with increasing growth parameter. For the window-based methods there is a trend for $\alpha$ to increase with the magnitude of the fitting range. However, the confidence intervals overlapped for each window method.

The Reasenberg linking algorithm produced very similar results for the Southern Californian datasets as the Utsu-Seki algorithm. In contrast, the Reasenberg algorithm did not work well for the CMT data with $M_c = 5.7$, presumably because there were not enough earthquakes above the completeness magnitude to link together the clusters.

**Foreshock occurrence and the foreshock model**

To determine the foreshock rate we counted all earthquakes that initiated a cluster as possible initiating earthquakes, again including the single events that were not followed by other earthquakes. We defined foreshocks to be those initiating earthquakes that were followed by a larger earthquake within $\Delta T$. The foreshock rate was calculated by dividing the number of foreshocks in each magnitude bin by the total number of initiating earthquakes in the same magnitude bin. Figures 5b and 6b show the foreshock data corresponding to the mean abundance in figures 5a and 6a. The magnitude-frequency distribution of the initiating events and foreshocks are shown in the left-hand columns as well as their $b$-values, $b_{tot}$ and $b_f$. A difference in these two $b$-values reflects a magnitude dependence of the foreshock rate. The magnitude-frequency distributions show that the amount of data available to calculate foreshock rates decays exponentially with increasing magnitude due to the Gutenberg-Richter relation. There is often only one or two foreshock observations one magnitude unit from where we started fitting mean abundance. Thus error bars for the data would be very large. To reduce the scatter in the foreshock rate data, we used a moving average of three magnitude bins. The data are shown in the right hand columns of figures 5b and 6b. Included in the plots are the mean models for foreshock rates, which were calculated according to equation 10, using the parameters from the mean abundance for each declustering algorithm and the $b$-value of the complete catalog, $b_{cat}$. Also shown are the 95% confidence intervals for the models. These were derived from 1000 simulations of the model for each magnitude bin, where
all parameters were assumed to be normally distributed with variance given from the data. The models generally fit the observed foreshock rate well. Close to the completeness magnitude the data tend to fall below the model for the Californian data. We explain this in the same manner as the roll-off of mean abundance close to the completeness magnitude. Due to magnitude uncertainty the events close to the cut-off had a lower probability of being observed as a foreshock. The exception is the Gardner-Knopoﬀ method, where the roll-off is off-set by the large search radius for small events. Where the foreshock data get sparse the data scatter about the model.

The difference between the diﬀerent declustering algorithms is striking, particularly for the small magnitudes in the Southern Californian catalog. The difference between \( \alpha \) and \( b_{cat} \) determines the magnitude dependence of the foreshock rate. For the Southern California data inside the polygon this difference is largest for Gardner-Knopoﬀ, followed by Uhrhammer with similar results for Reasenberg and Utsu-Seki (see table 3 for the results and 95\% conﬁdence limits). All differences between \( \alpha \) and \( b_{cat} \) are < 0, indicating that smaller earthquakes are more likely to be followed by larger earthquakes than large ones. For Gardner-Knopoﬀ the foreshock rate decreases from 13.2\% at magnitude 2.0 to 5.3\% at magnitude 3.0, and 2.1\% at magnitude 4.0. There are hardly any foreshock data above magnitude 4.0 in the selected data, and therefore the comparison of model results for larger foreshocks is not useful. The 95\% conﬁdence intervals of the foreshock model are not symmetrical as can be seen on ﬁgure 5a. The reason for the strong magnitude dependence of the foreshock rate for the Gardner-Knopoﬀ search windows lies in the very large search radius for small earthquakes. The magnitude dependence is less pronounced for the Uhrhammer search radii with 8.3\% at magnitude 2, 4.0\% at magnitude 3.0 and 1.9\% at magnitude 4.0. For Reasenberg and Utsu-Seki the results are very similar with 3.1\% and 3.5\% respectively, at magnitude 2, around 2.2\% at magnitude 3.0 and 1.5\% at magnitude 4.0. Note that for the Reasenberg and the Utsu-Seki algorithms the magnitude-frequency distributions of the foreshocks fall below the straight line at the lower magnitude end, leading to relatively low \( b_{s} \) values 1.15 and 1.09, respectively. The most likely reason for this lies in the search radii, which are very small for small earthquakes, possibly smaller than the location uncertainty of the catalog. Therefore not all related earthquakes may be counted in the same cluster and the foreshock rate drops.

The results for the global data reﬂect the results for the abundance ﬁtting and are similar to the results from the Californian data. Generally the foreshock rates agree well with the model derived from ﬁtting the mean abundance. Figure 6b shows two examples for the Reasenberg and the Utsu-Seki algorithms, which had the largest diﬀerence in \( \alpha - b_{cat} \) values. The magnitude dependence of the mean model for Reasenberg is less pronounced than for Gardner-Knopoﬀ method in California, even though the diﬀerence of \( \alpha - b_{cat} \) is larger for the global data. For the Utsu-Seki algorithm the diﬀerence of \( \alpha - b_{cat} \) is close to being 0 and thus the magnitude dependence of the foreshock rate is negligible.

In summary, the probability of an initiating event being followed by a larger earthquake within the speciﬁed \( \Delta T \) is around 3\%, except for the Gardner-Knopoﬀ and Uhrhammer algorithms in Southern California. There the probability for an earthquake to
be a foreshock is close to 13% and 6%, respectively, for earthquakes around magnitude 2.0. The probability drops within one magnitude to 4%. This is smaller than reported in previous studies and predicted by the Reasenberg-Jones model as discussed in the introduction.

Discussion and conclusion

We investigated whether foreshocks and mainshocks have the same triggering capability. For this purpose we defined abundance as a new formulation for the triggering capability, or productivity of mainshocks. It has two parameters: The growth exponent $\alpha$ and the magnitude $M_f$. In this formulation $M_f$ has a physical meaning, the magnitude that on average has one aftershock in time, space and magnitude windows that are defined by the search algorithm for earthquake clusters. $M_f$ is a useful tool to compare different search algorithms. We used four search algorithms; the Reasenberg linking algorithm and a hybrid method of magnitude dependent search radii and sliding time windows $\Delta T$ with three different radii derived from (i) Gardner-Knopoff's Southern Californian declustering algorithm; (ii) Uhrhammer's Northern Californian declustering algorithm; (iii) Utsu-Seki's formula for the growth of aftershock area with mainshock magnitude. Each algorithm was stretched beyond its limits in the sense that we applied it outside the magnitude range and the geographical area for which it was developed.

We analyzed two subsets of the South Californian catalog and one of the CMT global catalog and calculated the $b_{cat}$-values of the magnitude-frequency distribution, $\alpha$ and $M_f$. We defined foreshocks as initiating events of a cluster that were followed by a larger earthquake within the time period $\Delta T$. Assuming that foreshocks had the same triggering capability as mainshocks, we combined mean abundance with the magnitude-frequency distribution and derived foreshock rates. These compared well with the observations.

Below we discuss the effects of the different search algorithms on the parameters $\alpha$ and $M_f$ and on the difference ($\alpha - b_{cat}$), which controls the magnitude dependence of the foreshock rate. We stress the importance of consistency in data selection and compare our foreshock rate results with foreshock probabilities in previous studies. We close with four key messages.

The differences in $\alpha$ and $M_f$ with search algorithm

The search algorithms affected the fitting of mean abundance and consequently the foreshock models. For the window methods, the larger the search radius, the more earthquakes were part of the same cluster, increasing the number of aftershocks, and decreasing the magnitude $M_f$ that has on average one aftershock. The growth exponent of the search radii, which were 0.12, 0.35, and 0.50 for Gardner-Knopoff, Uhrhammer and Utsu-Seki, respectively, affected the parameter $\alpha$, increasing it with increasing growth exponent. Aftershocks are generally associated with the rupture area of the mainshock, which scales with mainshock magnitude [Kanamori and Anderson, 1975; Utsu and Seki, 1955]. For earthquakes of magnitude $M \geq 6.0$, Kanamori and Anderson found the rupture length $L$ to scale like $M = 2\log L$. This corresponds to a growth exponent for the search radius of 0.5, like Utsu-Seki. For smaller earthquakes, Kanamori and Anderson proposed
the rupture length \( L \) to scale like \( M = 3 \log L \), which corresponds to a growth exponent in search radius of 0.33, very similar to Uhrhammer’s 0.35. In comparison, Gardner-Knopoff’s growth exponent for the search radius is only 0.12. However, the Gardner-Knopoff search radii start large with 22.5 km for magnitude 3.0 [Gardner and Knopoff, 1974]. We associated the ‘bulge’ in the total number of aftershocks in the top left-hand plot of figure 5a with the overly large search radius, interpreting it as being caused by unrelated earthquakes that were associated with clusters that occurred relatively far away. The same phenomenon occurs with the Uhrhammer method in the magnitude range 2.2 to 3.2, in which the search radius increases from 2.1 km to 4.8 km. The Utsu-Seki radius increases from 320 m to 1 km in the same magnitude range and shows signs of missing aftershocks for magnitude 2.3 and smaller. We can partly explain that with magnitude uncertainty close to the completeness magnitude, which also seems to affect the other algorithms. However, it is not clear what the effect is of the location uncertainty in the catalog.

The Reasenberg algorithm had similar results for \( \alpha \) and \( M_l \) to the Utsu-Seki algorithm for Southern California, but failed to effectively cluster the global data with \( M_c = 5.7 \).

In summary, the search algorithms affect the mean abundance and we can infer some issues with each algorithm, which we will not further discuss here because they are not the focus of this paper.

The difference \( (\alpha - b_{cat}) \) and its effect on foreshock rates

The difference \( (\alpha - b_{cat}) \) determines the magnitude dependence of the foreshock rate (see equation 10). For \( \alpha = b_{cat} \), there is no magnitude dependency of foreshock rates, for \( \alpha < b_{cat} \) capability to trigger larger earthquakes decreases with magnitude, and for \( \alpha > b_{cat} \) aftershock activity would explode because larger earthquakes are more likely to trigger even larger ones. In the ETAS model, \( \alpha < b_{cat} \) is strictly required for the model not to be critical [Zhuang and Ogata, 2006] if the maximum magnitude is not constrained. When applying the ETAS model for simulating earthquakes in a limited magnitude range, it is possible to have \( \alpha > b_{cat} \). In our analysis, the difference \( (\alpha - b_{cat}) \) was always negative, ranging from -0.20 ± 0.16 to -0.29 ± 0.17 in dataset 2, -0.15 ± 0.08 to -0.40 ± 0.05 in dataset 4, and -0.03 ± 0.16 to -0.40 ± 0.19 in dataset 5 (see table 3). However, for the Uhrhammer and Utsu-Seki algorithms applied to the global data \( \alpha \) and \( b_{cat} \) were close enough to being equal as previously assumed [Felzer, et al., 2004]. As discussed in the previous section, the parameter \( \alpha \) depends on the search algorithm and therefore the calculated magnitude dependence of the foreshock rate could be an artifact of the data selection. However, as already pointed out, our purpose is not to evaluate the search algorithms.

In the context of discussing the difference between \( \alpha \) and \( b \), we note that the empirical relationship Båth’s law that states that the magnitude between mainshock and the largest aftershock in about 1.2, regardless of mainshock magnitude [Richter, 1956], implies that \( \alpha = b_{cat} \). The distribution of magnitude difference between mainshock and
largest aftershock can be derived from the triggering capability of mainshocks and the magnitude-frequency distribution, which feature $\alpha$ and $b_{cat}$ as parameters, and thus the difference between $\alpha = b_{cat}$ is an important input [Vere-Jones, 2008; Vere-Jones, et al., 2006].

**Consistency in data selection**

Comparing the foreshock rates for the Gardner-Knoppoff algorithm with the model predictions from abundance fitting for the Reasenberg algorithm in figure 5b, and vice versa, shows that defining aftershocks with one search algorithm and foreshock data with another does not give a good match between foreshock model and data. The issue is subtler for using different sliding time windows $\Delta T$. Figure 7 gives an example for Uhrhammer search algorithm with $\Delta T=1$. However, mean abundance parameters and foreshock rates were determined within 3 days. The data clearly fall below the model, in particular for the smaller foreshock magnitudes. Aftershock sequences are extended by $\Delta T$ for each subsequent earthquake within the current search parameters. Thus the abundant sequences would continue beyond three days in which we determine abundance. However; the less abundant sequences would be cut after $\Delta T=1$ and some later larger earthquakes would be in a different cluster and therefore the foreshock rate comes out smaller. This example stresses the importance of consistency in data selection when defining fore- and aftershocks.

**Foreshock probabilities – comparison with previous studies**

We find that the probability of the initiating earthquake in a pre-defined cluster to be followed by a larger one within one day is around 4% and less for California, for the Reasenberg and Utsu-Seki algorithms. For the Gardner-Knoppoff and Uhrhammer algorithms the probability is up to 13% and 8% respectively for the small magnitudes and decays quickly with increasing magnitude (see figure 5b). These probabilities agree reasonably well with Jones’ result of 6% which reduced to 4% when only one foreshock per sequence is counted [Jones, 1985]. Foreshock probabilities for Italy were larger with 7-9% [Console and Murr, 1993; Di Luccio, et al., 1997]. However, mainshock were allowed to occur within a radius of 30 km for foreshocks of magnitude 3 and larger. This radius is similar to 22.5 km and 30 km at magnitude 3.0 and 4.0, respectively, for the Gardner-Knoppoff algorithm, and the respective foreshock rates are comparable too. To investigate whether there is a real difference in foreshock rates between Italy and California both earthquake catalog would have to be analyzed with the same methodology.

For the global data, we find that the probability of the initiating earthquake in a pre-defined cluster to be followed by a larger one within 30 days is around 3%. This is much lower than 15% as previously reported for the same data set [Reasenberg, 1999b]. The previous study had some issues with magnitude completeness. If the data are not complete, then once one earthquake from a cluster is detected, it is likely to detect other events from the same cluster. By missing individual events the denominator for calculating foreshock probabilities would be reduced and consequently the probability increased. Furthermore the study seemed to include smaller aftershocks being followed.
by larger ones and thus increasing the foreshock probability.

We close with four conclusions: First, the ability of foreshocks to trigger larger earthquakes is consistent with the triggering capability of mainshocks. Thus there is not evidence for a separate mechanism for foreshock triggering and thus foreshocks have no more predictive power than other earthquakes. Secondly, data need to be selected consistently and care has to be taken about the completeness magnitude. Combining the first two points leads to the third conclusion: Using aftershock models for earthquake forecasting including the larger events is sensible, as long as the cluster definitions in the forecasting are consistent with the cluster definitions used for deriving model parameters. Finally, foreshock probabilities are only around 3-4%, with the value depending on the spatial, temporal and magnitude window.

Acknowledgements

We gratefully acknowledge the use of the SCEDC and the CMT data. Figure 1 was done with the mapping tool gmt and the help of Jochen Woessner. An early version of the paper was presented as a poster at the Fourth International Workshop on Statistical Seismology in Tokyo in January 2006. The first author gratefully acknowledges travel support from the workshop organizer Professor Ogata and the Workshop sponsors. We thank David Vere-Jones, Martha Savage, Yan Kagan, Sebastian Hainzl, an anonymous reviewer and the associate editor Stefan Wiemer for useful feedback on the manuscript. The study has been partly funded by a grant from the New Zealand Earthquake Commission Research Foundation. The investigation of the effect of different clustering algorithms on aftershock abundance and foreshock rates was funded by the European Union projects NERIES (contract number 026130) and SAFER (contract number 036935).

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


SCEDC (2007), SCEDC holdings, edited.


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


Zhuang, J., et al. (2008), Differences between background and triggered earthquakes: their influences on foreshock probabilities submitted to *JGR*.

ETH Zurich, Institute of Geophysics, Schafmattstrasse 30, 8093 Zurich, Switzerland (A.C.)
The Institute of Geophysics, Victoria University of Wellington, PO Box 600, Wellington, New Zealand (E.G.C.S.)

Christopher & Smith, Foreshock rates from aftershock abundance
### Tables

**Table 1: Overview of data sets used in the analysis**

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Region</th>
<th>Time period</th>
<th>$M_C$</th>
<th>Comment</th>
<th>Number of events</th>
<th>$b_{cat}$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Southern California</td>
<td>1984-2006</td>
<td>3.0</td>
<td>All SCEDC data</td>
<td>10,171</td>
<td>1.05±0.01</td>
</tr>
<tr>
<td>2</td>
<td>Southern California</td>
<td>1984-2006</td>
<td>3.0</td>
<td>Clusters with M≥6.0 removed</td>
<td>7,201</td>
<td>1.13±0.01</td>
</tr>
<tr>
<td>3</td>
<td>Southern California</td>
<td>1984-2006</td>
<td>2.0</td>
<td>Data within polygon in Fig.1</td>
<td>75,689</td>
<td>1.03±0.004</td>
</tr>
<tr>
<td>4</td>
<td>Southern California</td>
<td>1984-2006</td>
<td>2.0</td>
<td>As above but clusters with M≥6.0 removed</td>
<td>53,030</td>
<td>1.11±0.005</td>
</tr>
<tr>
<td>5</td>
<td>Global</td>
<td>1980-2004</td>
<td>5.7</td>
<td>All CMT data</td>
<td>4,842</td>
<td>1.01±0.01</td>
</tr>
</tbody>
</table>

**Table 2: Parameters used for the Reasenberg declustering code.**

<table>
<thead>
<tr>
<th>Data subset Parameter</th>
<th>SCEDC All data</th>
<th>SCEDC Polygon</th>
<th>CMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{min}$ [days]</td>
<td>3</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>$\tau_{max}$ [days]</td>
<td>10</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>$\alpha_{eff}$</td>
<td>3.0</td>
<td>2.0</td>
<td>5.7</td>
</tr>
<tr>
<td>$\tau_{bet}$</td>
<td>10</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 3: Results of fitting mean abundance for different data sets and different clustering algorithms**

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Clustering</th>
<th>$\Delta T$ [days]</th>
<th>$M_C$</th>
<th>Magnitude bins</th>
<th>$\alpha$</th>
<th>$M_1$</th>
<th>$\alpha - b_{cat}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Gardner-Knoff</td>
<td>3</td>
<td>3.0</td>
<td>15</td>
<td>$0.84 \pm 0.17$</td>
<td>$4.21 \pm 0.09$</td>
<td>$-0.29 \pm 0.17$</td>
</tr>
<tr>
<td>2</td>
<td>Uhrhammer</td>
<td>3</td>
<td>3.0</td>
<td>15</td>
<td>$0.88 \pm 0.18$</td>
<td>$4.28 \pm 0.09$</td>
<td>$-0.25 \pm 0.18$</td>
</tr>
<tr>
<td>2</td>
<td>Utsu-Seki</td>
<td>3</td>
<td>3.0</td>
<td>15</td>
<td>$0.93 \pm 0.16$</td>
<td>$4.44 \pm 0.08$</td>
<td>$-0.20 \pm 0.16$</td>
</tr>
<tr>
<td>2</td>
<td>Reasenberg</td>
<td>3</td>
<td>3.0</td>
<td>15</td>
<td>$0.92 \pm 0.24$</td>
<td>$4.39 \pm 0.13$</td>
<td>$-0.21 \pm 0.24$</td>
</tr>
<tr>
<td>4</td>
<td>Gardner-Knoff</td>
<td>1</td>
<td>2.0</td>
<td>22</td>
<td>$0.71 \pm 0.05$</td>
<td>$3.16 \pm 0.06$</td>
<td>$-0.40 \pm 0.05$</td>
</tr>
<tr>
<td>4</td>
<td>Uhrhammer</td>
<td>1</td>
<td>2.0</td>
<td>22</td>
<td>$0.79 \pm 0.06$</td>
<td>$3.30 \pm 0.06$</td>
<td>$-0.32 \pm 0.06$</td>
</tr>
<tr>
<td>4</td>
<td>Utsu-Seki</td>
<td>1</td>
<td>2.0</td>
<td>22</td>
<td>$0.93 \pm 0.08$</td>
<td>$3.51 \pm 0.06$</td>
<td>$-0.18 \pm 0.08$</td>
</tr>
<tr>
<td>4</td>
<td>Reasenberg</td>
<td>1</td>
<td>2.0</td>
<td>22</td>
<td>$0.96 \pm 0.08$</td>
<td>$3.52 \pm 0.06$</td>
<td>$-0.15 \pm 0.08$</td>
</tr>
<tr>
<td>5</td>
<td>Reasenberg</td>
<td>30</td>
<td>5.7</td>
<td>17</td>
<td>$0.61 \pm 0.19$</td>
<td>$7.76 \pm 0.31$</td>
<td>$-0.40 \pm 0.19$</td>
</tr>
<tr>
<td>5</td>
<td>Gardner-Knoff</td>
<td>30</td>
<td>5.7</td>
<td>16</td>
<td>$0.83 \pm 0.13$</td>
<td>$7.30 \pm 0.09$</td>
<td>$-0.18 \pm 0.13$</td>
</tr>
<tr>
<td>5</td>
<td>Uhrhammer</td>
<td>30</td>
<td>5.7</td>
<td>15</td>
<td>$0.91 \pm 0.15$</td>
<td>$7.23 \pm 0.09$</td>
<td>$-0.10 \pm 0.15$</td>
</tr>
<tr>
<td>5</td>
<td>Utsu-Seki</td>
<td>30</td>
<td>5.7</td>
<td>15</td>
<td>$0.97 \pm 0.15$</td>
<td>$7.24 \pm 0.09$</td>
<td>$-0.04 \pm 0.15$</td>
</tr>
<tr>
<td>5</td>
<td>Utsu-Seki</td>
<td>15</td>
<td>5.7</td>
<td>15</td>
<td>$0.98 \pm 0.18$</td>
<td>$7.29 \pm 0.11$</td>
<td>$-0.03 \pm 0.18$</td>
</tr>
<tr>
<td>5</td>
<td>Utsu-Seki</td>
<td>60</td>
<td>5.7</td>
<td>15</td>
<td>$0.98 \pm 0.16$</td>
<td>$7.19 \pm 0.09$</td>
<td>$-0.03 \pm 0.16$</td>
</tr>
</tbody>
</table>
Figure captions

Figure 1: $b$-value map of Southern California as calculated by the entire magnitude range method [Woessner and Wiemer, 2005] with the software package zmap [Earthquake Statistics Group, 2007; Wiemer, 2001], using 100 bootstraps on a 0.1x0.1 degree grid with 100 earthquakes per grid node and 30 earthquakes above the completeness magnitude. The grey line shows a polygon in which the completeness magnitude is 2.0, except immediately following an earthquake of magnitude 6.0 and larger.

Figure 2: The completeness magnitude with time for the CMT catalog. The solid line is a moving average of 500 earthquakes per sampling windows with an overlap of 10 earthquakes between windows, in magnitudes in 0.1 bin width. The dotted lines are the 95% confidence intervals of the data. The dashed line presents the chosen completeness magnitude of 5.7 since 1980.

Figure 3: Illustration of cluster linking in time. Each cluster is extended by a time window $\Delta T$ following an associated event. The first earthquake in the cluster is called the initiating event. It is a foreshock when a larger earthquake follows in the same cluster.

Figure 4: The different search radii in space as a function of magnitude for the window methods Gardner-Knopoff (triangles), Uhrhammer (solid line) and Utsu-Seki (squares). The stars show the largest distance between mainshock and aftershock for any mainshock magnitude for the Reasenberg linking algorithm with the parameter setting in table 1.

Figure 5a: SCEDC – polygon data: Comparison of different clustering algorithms. The left-hand column shows the magnitude-frequency distribution of the mainshocks (triangles) and the maximum likelihood fit. The stars represent all aftershocks within $\Delta T = 1$ day of the mainshock for each mainshock magnitude. The dashed reference line shows the value $N_{ave}$ of the mean number of aftershocks per mainshock magnitude. The right-hand column shows the mean abundance as calculated from the data on the left. The solid line is the model fit including all data at least half a magnitude unit from the completeness and 10 observations. The fitted $\alpha$ and $M_l$ are shown and their 95% confidence intervals. The dashed line and the thin solid lines are the 95% confidence intervals of the model and the data, respectively.

Figure 5b: SCEDC – polygon data: Comparison of different clustering algorithms. The left-hand column shows the magnitude-frequency distribution for initiating earthquakes (squares) and foreshocks (triangles) and their $b$-values as fitted by maximum likelihood. The difference of $b_{im}$ and $b_{PS}$ has an effect on the magnitude dependence of the foreshock rate. The right-hand column shows the foreshock rate calculated with a moving average of three adjacent magnitude bins. The solid line presents the model according to equation 10 with the parameters derived from the mean abundance fit. The 95% confidence interval (dashed lines) is derived from 1000 simulations of equation 10 with normally distributed parameters, $\alpha$, $M_l$ and $b$.

Figure 6a: The same as in figure 5a for the CMT data with $M_C = 5.7$. Comparison of Reasenberg and Utsu-Seki algorithms.

Figure 6b: The same as in figure 5b for the CMT global data.

Figure 7: Example of inconsistent data selection in time: Sliding time windows of $\Delta T = 1$ day and Uhrhammer spatial windows were used to search for clusters but mean abundance and foreshock rates were determined for three days. (see text for further explanation)
Figure 3:

Figure 4:
Figure 5a:

SCED polygon

Gardner Knopoff

Mean Abundance
- Model
- 95% model
- 95% data

Mα = 0.71 ± 0.05
M_l = 3.16 ± 0.06

Uhrammer

N_{ave} = 240

Mα = 0.79 ± 0.06
M_l = 3.50 ± 0.06

Reasenberg

N_{ave} = 187

Mα = 0.98 ± 0.08
M_l = 3.52 ± 0.06

Utsu-Seki

N_{ave} = 178

Mα = 0.93 ± 0.08
M_l = 3.50 ± 0.06

Christophersen & Smith, Foreshock rates from aftershock abundance
Figure 5b:

- **SCED polygon**
  - Gardner Knopoff
    - \( b_{\text{sh}} = 1.19 \)
    - \( b_{\text{ts}} = 1.72 \)
  - Uhrhammer
    - \( b_{\text{sh}} = 1.20 \)
    - \( b_{\text{ts}} = 1.49 \)
  - Reasenberg
    - \( b_{\text{sh}} = 1.20 \)
    - \( b_{\text{ts}} = 1.15 \)
  - Utsu-Seki
    - \( b_{\text{sh}} = 1.19 \)
    - \( b_{\text{ts}} = 1.09 \)

- **Foreshock Rate**
  - Model
  - 95% model

Christopher & Smith, Foreshock rates from aftershock abundance
Figure 6a:
Figure 6b:

![Graphs showing foreshock rate and fitted lines for Reasenberg and Utsu-Seki](image)

Figure 7:

![Graph showing foreshock rate over magnitude](image)
APPENDIX 3  DIFFERENCES BETWEEN SPONTANEOUS AND TRIGGERED EARTHQUAKES: THEIR INFLUENCES ON FORESHOCK PROBABILITIES (12 PAGES)

Differences between spontaneous and triggered earthquakes: their influences on foreshock probabilities

Jiancang Zhuang\textsuperscript{1}, Martha K. Savage\textsuperscript{2}, Annemarie Christophersen\textsuperscript{3}
David Vere-Jones\textsuperscript{4}, Yoshihiko Ogata\textsuperscript{1} and David D. Jackson\textsuperscript{5}

\textsuperscript{1}Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-Ku, Tokyo 106-8569, Japan
\textsuperscript{2}Institute of Geophysics, School of Earth Sciences, Victoria University of Wellington
P.O.Box 600, Wellington, New Zealand
\textsuperscript{3}Swiss Seismological Service
ETH Zürich, Schafmattstr. 30, CH-8093 Zürich, Switzerland
\textsuperscript{4}School of Mathematics, Statistics and Computer Sciences, Victoria University of Wellington
P.O.Box 600, Wellington, New Zealand
\textsuperscript{5}Department of Earth and Space Sciences, University of California
Los Angeles, CA 90095-1567, USA

Abstract

In this study we investigate the foreshock probabilities calculated from earthquake catalogs from Japan, Southern California and New Zealand. Unlike conventional studies on foreshocks, we use a probability-based de-clustering method to separate each catalog into stochastic versions of family trees, such that each event is classified as either having been triggered by a preceding event, or being a spontaneous event. The probabilities are determined from parameters that provide the best fit of the real catalogue to a synthetic catalogue determined using a space-time epidemic-type aftershock sequence (ETAS) model, which assumes that every earthquake has a certain probability to trigger another event. A foreshock here is defined as a spontaneous event that has one or more larger descendants. The proportion of foreshocks in spontaneous events of each catalog is found to be lower than the proportion of the foreshocks in the triggered events. We infer that this is caused by different triggering behaviors in spontaneous versus triggered events. This implies that, if we evaluate the risk of foreshocks on the assumption that foreshocks trigger mainshocks in the same manner as mainshocks trigger aftershocks, then we over-predict the risk of foreshocks.

1 Introduction

The foreshock is one of the most popular issues in research on seismicity and earthquake prediction, and even of public attention. Immediately after the occurrence of a moderate earthquake, the public and the researchers want to know whether a larger, destructive earthquake will follow. The reported forecast of the Haicheng earthquake (1976-7-26, $M_{7.6}$) in China was made mainly based on foreshocks. However, the answer is not straightforward.

The first problem in foreshock studies is how to distinguish foreshocks from other shocks. To do this, an explicit definition of foreshocks is required. In the past, foreshocks have usually been defined as earthquakes that are followed in a short time by at least one larger earthquake nearby. The size of the space-time window is subjectively chosen. But aftershock sequences contain many events, of which many small ones are followed by bigger aftershocks, even though they might all be considered to have been triggered by the same, large mainshock. Thus, aftershocks will “contaminate” the foreshock statistics by making it appear that many events are followed by larger ones. So, much effort has been focussed on trying to remove aftershocks from catalogues (often called de-aftershocking or declustering), which is a complicated problem. There are many declustering methods, classified into three types: a) window-based (e.g., Gardner and Knopoff, 1974), b) link-based (e.g., Frohlich and Davis, 1990) and c) probability based (e.g., Kagan and Knopoff 1976, Zhuang et al 2002). The first two give deterministic separations of
earthquake clusters. Type c) gives the probabilities that each earthquake is spontaneous one or triggered by another previous event.

The second problem is whether a foreshock is a mainshock whose aftershocks happen to be large. Jones (1985) tries to find the difference in several geophysical variables between foreshocks and mainshocks in southern California; her results do not show any clear differences. Jones et al. (2005, unpublished manuscript) use the Omori law as the baseline model to describe the manners in which mainshocks produce aftershocks and find no evidence to show that foreshocks trigger mainshocks in a different way. Helmstetter et al. (2003, 2004) and Felzer et al. (2004) compare simulation data from the ETAS model and conclude that foreshocks are no different from other shocks in triggering seismicity. But they used a window based method for the definitions of foreshocks and earthquake clusters, which includes arbitrary artificial parameters. Simulations can be influenced or controlled by several parameters, whose different values may change the conclusion completely. To make a more strict test, it is necessary to use more objective methods to formulate the foreshocks and earthquake clusters.

The third problem is how to recognize a foreshock before the occurrence of the larger mainshock and how to make use of foreshocks in prediction. The answer to this question depends mainly on the answers to the second question. If foreshocks are merely mainshocks that have larger aftershocks, we cannot find any difference between foreshocks and other types of earthquakes. We can then rely on a clustering model for earthquake clusters, such as the ETAS model, for the purpose of probabilistic prediction; otherwise, we have to formulate a new model purely for foreshocks, as done by Ogata (1996).

To understand the above problems, we need a good model of earthquake clustering. For Problem 1 (foreshock definition), we need to remove the aftershocks that occurred in earthquake clusters, because there are many events in an aftershock sequence satisfying the conditions of foreshocks. For example, Hainzl et al. (2005) showed that Reseanberg (1985)'s declustering methods yield incorrect estimation of the proportion of spontaneous events in the catalogue. For Problem 2, if we have a good baseline (null) model for seismicity clustering, we can make extensive studies on whether there are distinguishing features between the patterns of foreshocks and mainshocks in triggering seismicity. The hypothesis of the same mechanisms for both foreshocks and mainshocks can be rejected if distinctive differences can be found. For Problem 3, if foreshocks are just mainshocks that have larger aftershocks, a clustering model is good enough for the purposes of prediction.

There have been many studies on the modelling of earthquake clustering (see, e.g., Console et al, 2003; Console and Murrn, 2001; Helmstetter and Sornette, 2003; Helmstetter et al, 2004; Kagan, 1991; Ogata, 1998, 2004; Zhuang et al 2002, 2004). Their models assume that the seismicity can be divided into a spontaneous component and a clustering component, and that each event, whether it is spontaneous or it is directly triggered by another event, triggers its own offspring according to some general rules.

According to the Omori-Utsu formula and the Gutenberg-Richter law, Vere-Jones et al (2006) give the probability distribution of the magnitude of the largest earthquake that one earthquake can trigger. Zhuang and Ogata (2004, 2006) and Salchev and Sornette (2005) also calculate the probability distributions associated with the magnitude of the largest event in an earthquake cluster. These studies are closely related to foreshocks probability, because the probability that a spontaneous event is a foreshock is the same probability that the largest event in its cluster is greater than it. Our study will follow such ideas together with sufficient data analysis.

In this article, we are going to use the ETAS model to study the above problems. We will firstly give a brief description of necessary concepts and methodologies associated with the ETAS model. We secondly concentrate on the discrepancies of the ETAS model and earthquake data, by which we proposed an improved version of the model. Then, different from conventional studies, we re-define foreshocks as spontaneous events that have at least one larger descendant in all the generation. Making use of the ETAS model and the updated version, we will investigate the features of foreshocks.
2 Concepts and methodologies

The ETAS model The ETAS model has the time-varying seismicity rate (termed as conditional intensity function, see, e.g., Delay and Vere-Jones, 2003),

\[
\lambda(t, x, y, m) = \lim_{\delta_1 \to 0^+, \delta_2 \to 0^+, \delta_3 \to 0^+} \frac{\mathbb{E} \{ N((t, t + \delta_1] \times (x, x + \delta_2] \times (y, y + \delta_3]) \mid \mathcal{H}_t \} }{\delta_1 \delta_2 \delta_3}
\]

\[
= \mu(x, y, m) + \sum_{i, t_i < t} \xi(t, x, y, m; t_i, x_i, y_i, m_i),
\]

i.e., the expected number of earthquake in the unit space-time-magnitude at \((t, x, y, m)\) given the observations before \(t\), where \(\mathcal{H}_t\) represents the history of observation up to time \(t\), but not including \(t\), \(\mu(x, y, m)\) represents the stationary spontaneous seismicity rate and \(\xi(t, x, y, m; t_i, x_i, y_i, m_i)\) is the contribution to seismicity rate by the \(i\)th event occurring previously. In practice, it is usual to make the following assumptions: (1) The whole process is magnitude separable, i.e.,

\[
\lambda(t, x, y, m) = \lambda(t, x, y) s(m),
\]

where

\[
\lambda(t, x, y) = \mu(x, y) + \sum_{i, t_i < t} \xi(t, x, y; t_i, x_i, y_i, m_i)
\]

and

\[
s(m) = \beta e^{-\beta(m-m_c)}, \; m \geq m_c
\]

is the probability density form of the Gutenberg-Richter law (exponential distribution); (2) the response function \(\xi(t, x, y; t_i, x_i, y_i, m_i)\) is separable and depends on the difference in time and spatial locations, in the form of

\[
\xi(t, x, y; t_i, x_i, y_i, m_i) = \kappa(m_i) g(t - t_i) f(x - x_i, y - y_i; m_i),
\]

where

\[
\kappa(m) = A e^{-a(m-m_c)}, \; m \geq m_c
\]

is the the expectation of the number of children (productivity), which is a Poisson random variable, from an event of magnitude \(m\),

\[
g(t) = \frac{p - 1}{c} \left( 1 + \frac{t}{c} \right)^{-p}, \; t > 0,
\]

is probability density function (p.d.f.) of the length of the time interval between a child and its parent, and

\[
f(x, y; m) = \frac{q - 1}{\pi D e^{q(m-m_c)}} \left( 1 + \frac{x^2 + y^2}{D e^{q(m-m_c)}} \right)^{-q}
\]

is the p.d.f. of the relative locations between the parent and children, \(m_c\) being the magnitude threshold. The above formulations are according to Zhuang et al (2004, 2005) and Ogata and Zhuang (2005), which is an improved version of the one in Ogata (1998).

Stochastic declustering and stochastic reconstruction Once the conditional intensity function is estimated, it provides us a good method to estimate the probability that an event is a spontaneous event or is instead triggered by others (Kagan and Knopoff, 1980; Zhuang et al 2002). Consider the contribution of the spontaneous seismicity rate at the occurrence of the \(i\)th event,

\[
\varphi_i = \frac{\mu(x_i, y_i)}{\lambda(x_i, x_i, y_i)}.
\]

If we remove the \(i\)th event with probability \(1 - \varphi_i\), for all the events in the process, we can realize a process with the occurrence rate of \(\mu(x, y)\) (see Ogata, 1984, or Zhuang 2006, for justification). Thus it
is natural to regard $\varphi_i$ as the probability that the $i$th event is a spontaneous events, given the realization and the fitted model. Similarly,

$$\rho_{ij} = \frac{\kappa(m_i)g(t_j - t_i)f(x_j - x_i, y_j - y_i; m_i)}{\lambda(t_j, x_j, y_j)}$$  \hspace{1cm} (10)

can be regarded as the probability that the $j$th event is directly produced by $i$. Alternatively, one can consider $\sum_j \rho_{ij}$ as the expected number of direct aftershocks (children) produced by event $i$. Here we refer to Zhuang et al. (2002, 2004) for the declustering and cluster separation algorithms.

Estimating parameters and spontaneous seismicity rate  For a realization of the process, $\{(t_i, x_i, y_i, m_i) : i = 1, \cdots, N\}$, in a spatial region $S$ and a time interval $[0, T]$, the likelihood has the standard form

$$\log L = \sum_{i=1}^{N} \log \lambda(t_i, x_i, y_i, m_i) - \int_{0}^{T} \int_{S} \lambda(t, x, y) dx dy dt + \sum_{i=1}^{N} \log s(m_i).$$  \hspace{1cm} (11)

The model parameters, $\theta = (\mu, A, \alpha, c, p, D, q, \gamma)$, can be estimated through maximizing the likelihood function. For the version of a model with nonhomogeneous spontaneous rate, the procedure for estimating the spontaneous rate and the model parameters is given by Zhuang et al. (2002, 2004, 2005) as an iterative algorithm.

3 Data

Three datasets are used in this study: the Japan Meteorological Agency (JMA) catalogue, the Southern California earthquake catalogue and the New Zealand catalogue.

We use the (JMA) catalog in the range of longitude 121° - 155°E, latitude 21° - 48°N, depth 0 - 100 km, time 1926/January/1 - 1999/December/31 and magnitude \( > M_{J}4.2 \) (Figure 1). For an earthquake catalogue covering records of a long history, completeness and homogeneity are always problems causing troubles for statistical analysis. To tackle these problems, we choose a target space-time range, in which the seismicity seems to be relatively and visually complete and homogeneous. We choose a polygon as shown in Figure 1a as the target region with the same depth and magnitude ranges to fit the model and to build empirical analysis. The events outside of this study space-time range are used as complementary events for calculating the boundary effect.

The second dataset is the New Zealand earthquake catalog compiled by GNS Science, which covers the seismicity in the years 1964 to 2003. We choose the earthquakes of $M_L \geq 4.0$ with depths less than 40 km in the years 1985 to 2003 and with locations inside the polygon as shown in Figure 1.

The third catalog is the southern California catalogue, which covers the years 1932 to 2006 compiled by the South California Earthquake Data Center (SCEDC). As almost all of the earthquake in Southern California are shallow ones, all of the earthquakes in the polygon shown in Figure 1(c) and during the years 1980 and 2003 are selected for the analysis. The magnitude threshold is chosen as $M_{3.4}$.

4 Analysis

Foreshocks probabilities under the ETAS model and obtained by stochastic reconstruction

We define foreshocks as spontaneous events that are followed by at least one larger offspring. Thus,
spontaneous events are divided into two classes: foreshocks and non-foreshock spontaneous events, respectively. Similarly, we classify triggered events into events that have larger offspring and events that have no larger offspring, namely, triggered foreshocks and triggered non-foreshocks, respectively.

A mainshock in the conventional sense can be a spontaneous non-foreshock or triggered non-foreshock, depending on whether it is the initiating event of the cluster. To understand whether foreshocks are mainshocks that have larger aftershocks or whether the whole catalogue has only a unique triggering mechanism, a meaningful and direct test is to find statistical differences between the features of foreshocks and triggered foreshocks.

The ETAS model assumes that there is no difference between spontaneous events and triggered events. In this case, as shown by Zhuang and Ogata (2006), the probability $H(m)$ that an event of magnitude $m$ has at least 1 larger descendant is

$$H(m) = 1 - \exp[-\kappa(m) F(m)]$$

where $F(m)$ is the solution of the equation

$$F(m) = 1 - \int_{m_b}^m s(m') \exp[-\kappa(m') F(m')] \, dm'$$

and is interpreted as the probability that the largest earthquake in an arbitrary cluster, including the initial event and all its descendants, is greater than $m$.

Using the stochastic declustering method (Zhuang et al. 2002, 2004), we separate the events in each of the three catalogues into different family trees: Each event is randomly assigned to belong to the spontaneous or to a child of another event according to the probabilities obtained with the declustering parameters. Within each family of triggered events, we keep track of whether it has a larger descendant. Such a declustering procedure is repeated many times in order to get enough copies of the stochastic realizations. We then count the proportion of foreshocks in the spontaneous events and the proportion of triggered in the triggered events. These are considered the foreshock probability rates, as plotted in Figure 6. It is clear that for all the cases, the proportion of foreshocks in spontaneous events is much lower than the proportion of triggered foreshocks in triggered events.

**Difference between spontaneous seismicity and clustering seismicity** What causes the difference between the foreshocks and triggered foreshocks? Zhuang and Ogata (2006) suggest such differences may be mainly caused by the difference between spontaneous events and triggered events in the ability of triggering children and in the magnitude distribution. In the ETAS model, there is no difference between the spontaneous events and the triggered events, i.e., once an event occurs, no matter whether it is spontaneous or triggered, it triggers direct children in the same way as other events of the same magnitudes. Yet the earthquakes in the real catalogue that we have separated into families show differences between the spontaneous and triggered events. For example, Zhuang et al. (2004) introduce the stochastic reconstruction method for the JMA catalog, and find that there are differences between spontaneous events.
and triggered events in the triggering productivity and the magnitude distribution. We applied the same procedures to the New Zealand and the SCEDC catalogs to verify whether this is true for all the three datasets used in this study. We first reconstructed the following empirical productivity functions for the spontaneous events and the triggered events, respectively,

\[ \hat{\kappa}(m) = \frac{\sum_{i} \varphi_i I(||m_i - m|| < \Delta)}{2\Delta \sum_{i} \varphi_i I(||m_i - m|| < \Delta)} \]  

(14)

and

\[ \hat{\kappa}(m) = \frac{\sum_{i} (1 - \varphi_i) I(||m_i - m|| < \Delta)}{2\Delta \sum_{i} (1 - \varphi_i) I(||m_i - m|| < \Delta)} \]  

(15)

where \( \Delta \) is a small positive number, \( \sum_{i} \rho_{ij} \) can be regarded as the total number of children produced by event \( i \), and \( \sum_{i} \varphi_i I(||m - m_i|| < \Delta) \) can be regarded as the total number of spontaneous event with magnitudes in the interval \( (m - \Delta, m + \Delta) \). Thus, \( \hat{\kappa}(m) \) and \( \hat{\kappa}(m) \) are the estimates of the average number of children produced by a event of magnitude \( m \).

In Figure 4, we found that, for all three datasets, a spontaneous event produces less children than a triggered event of the same magnitude. Such differences get smaller when the magnitudes get larger.

Another interesting feature to investigate is the difference in the magnitude distribution of spontaneous seismicity and triggered seismicity. Using similar reconstruction method, we have

\[ \hat{\delta}(m) = \frac{\sum_{i} \varphi_i I(||m_i - m|| < \Delta)}{2\Delta \sum_{i} \varphi_i} \]  

(16)

and

\[ \hat{\delta}(m) = \frac{\sum_{i} (1 - \varphi_i) I(||m_i - m|| < \Delta)}{2\Delta \sum_{i} (1 - \varphi_i)}. \]  

(17)

The reconstructed results are also shown in Figure 4. Again the differences are clear, spontaneous seismicity has a larger proportion of smaller earthquakes than triggered seismicity.

**Explanation of the difference between foreshocks and triggered foreshocks**  If the differences between spontaneous events and triggered events exist, the probabilities that an triggered event and a spontaneous event of magnitudes \( m \) have one or more larger descendant are,

\[ H_c(m) = 1 - \exp[-\kappa_c(m)F_c(m)] \]  

(18)

and

\[ H_b(m) = 1 - \exp[-\kappa_b(m)F_b(m)], \]  

(19)

respectively, where \( F_c(m) \) satisfies the equation

\[ F_c(m) = 1 - \int_{m_e}^{m} s_c(m') \exp[-\kappa_c(m')F_c(m')] dm'. \]  

(20)

After we reconstruct the empirical functions, \( \hat{\kappa}_b, \hat{\kappa}_c, \) and \( \hat{\delta}_c, \) we substitute them into (18) and (19) to calculate \( H_b \) and \( H_c \), denoted by \( \hat{H}_b \) and \( \hat{H}_c \). In this way, we have two sets of estimates, \( \{\hat{H}_b, \hat{H}_c\} \) and \( \{\hat{H}_b, \hat{H}_c\} \). For all the three cases in this study, as shown in Figure 6, \( \hat{H}_b \) and \( \hat{H}_c \) are close enough to \( H_b \) and \( H_c \), respectively.

**Changing magnitude threshold**  The magnitude threshold plays a critical role in the determination of foreshocks. A foreshock might not be a foreshock any longer when the magnitude threshold is lowered down, if it is an offspring event smaller than the previous magnitude threshold. Thus, the foreshock is a catalog-related concept. To see whether the above results hold for a different magnitude threshold, we set alternative magnitude thresholds for the JMA, New Zealand, SCEDC catalogs as 4.5, 4.3 and 3.8, respectively, and repeat the above analysis. The results are shown in Figures 4 and 5. Figure 4 shows that the differences in spontaneous and triggered event still exist even if the magnitude threshold changes. Such results show that a more reasonable model than the ETAS model is to formulate spontaneous events and triggered event with difference triggering behaviors and magnitude distributions. Of course,
as shown in previous sections, the most influence of this feature to the catalog is that the proportions of foreshocks in real catalogs are over predicted by the ETAS model. Figure 2 shows differences between the proportions of foreshocks in spontaneous events and of triggered foreshocks in triggered events, which can be also well explained by the differences between spontaneous events and triggered events in productivity functions and magnitude distributions.

**Study on a simulated catalog** To make sure that the existence of the above difference between spontaneous and triggered events are not caused by the numerical procedures, we applied the same procedure to a synthetic catalogue simulated by using the ETAS model with parameters and spontaneous rate fitted from the JMA catalog. Please see Zhuang et al. (2004) for the simulation algorithm. In this synthetic catalogue, as indicated by the ETAS model, there is no difference between spontaneous events and triggered events. Thus, if the reconstruction works, the reconstructed \( k_b \) and \( k_c \) should show no difference between each other as well as \( \hat{s}_b(m) \) and \( \hat{s}_c(m) \). Similarly, \( \hat{H}_b(m) \) and \( \hat{H}_c(m) \) should also be the same.

We applied the same procedures, i.e., fitting the model to the simulated data to estimate the spontaneous rate and obtain maximum likelihood estimate of model parameters, calculate \( \varphi_j \) and \( \rho_j \) with the estimated spontaneous rate and model parameters, reconstructing \( k_b, k_c, \hat{s}_b \) and \( \hat{s}_c \), evaluating \( \hat{H}_b(m) \) and \( \hat{H}_c(m) \) by repeating stochastic separation of clusters for many times, and calculating \( \hat{H}_b(m) \) and \( \hat{H}_c(m) \) using \( k_b, k_c, \hat{s}_b \) and \( \hat{s}_c \). The results are shown in Figure 5. It clear to see that the difference between the reconstructed characteristics of the spontaneous events and triggered event are ignorably small. Consequently, the difference between the proportions of foreshocks in spontaneous events and of triggered foreshocks in triggered events can be ignorable.

5 Discussion

The magnitude distributions of spontaneous events and triggered events are different. Spontaneous events have a higher Gutenberg-Richter \( b \)-value than triggered events for all three dataset. This seems contrary to our knowledge that mainshocks usually have a smaller \( b \) value than aftershocks. The spontaneous events are initial events of each cluster, but not the largest events (mainshocks), and thus have a lower mean magnitude than mainshocks and a higher \( b \) value. This low \( b \) value indicates that clusters tend to be initiated by small events.

A triggered event triggers more direct children on average than a spontaneous event of the same magnitude. This can be explained by the fact that the number of children events triggered by a parent event depends not only on the parent’s magnitude, but also on the heterogeneity of the stress field. The more heterogeneous the stress field, the more children each parent triggers. At the beginning of an earthquake cluster, the stress field is maybe at a relatively homogenous state. The stress field becomes more heterogeneous as the cluster evolves until it is adjusted by the occurrence of most events in the cluster. The stress field recovers to a state being relatively homogeneous at the final stage of the cluster.

Are foreshocks mainshocks whose aftershocks happen to be bigger is of interest? In our opinions, the classification of spontaneous events and triggered events is more essential than the classification of foreshocks, mainshocks, and aftershocks. Foreshocks and mainshocks are not easily comparable in triggering other events, because a foreshock is always a spontaneous event while a mainshock may be a spontaneous event or an event triggered by a previous event. The events among the triggered events corresponding to foreshocks in the spontaneous events are the triggered events which have one or more larger descendants. From our analysis, the later ones have higher ability in producing offspring, larger or not, than foreshocks, which is due to the main differences between triggered events and spontaneous events. Thus, the spontaneous seismicity consists of foreshocks, isolated spontaneous events, and mainshocks that has no foreshocks.

Another important issue is how to make use of the above results for the purpose of earthquake forecast. The results of this paper and other related ones prove these facts: (1) the occurrence of big earthquakes following the so-called “foreshocks” is not due to coincidence; (2) given an event, the probability that it is a foreshock is overestimated if it is estimated by using the ETAS model or the Omori-Utsu formula with the Gutenberg-Richter law for the magnitude (STEP model, Gerstenberger et al., 2005); (3) prediction based on foreshocks can be implemented based on a clustering model more complicated than the ETAS model; (4) Even though a spontaneous has a lower probability that be followed

*Zhuang et al., Differences between spontaneous and triggered earthquakes*
by a larger earthquake, according to our analysis, a swarm of events with similar magnitude increase the risk of the occurrence of a larger earthquake dramatically.

6 Conclusion

Based on the theory of the ETAS model and the stochastic reconstruction method, we obtained the probabilities associated with the spontaneous events which have larger descendants (defined as foreshocks here) in the regions of Japan, New Zealand and Southern California. We found that the proportion of foreshocks in spontaneous events is much lower than the proportion of triggered foreshocks in triggered events. Such a feature can be well explained with the different clustering characteristics between spontaneous event and clustering events. Moreover, we may over-predict earthquake risks with foreshocks if the prediction is based on the assumption that foreshocks trigger mainshocks in the same way as mainshocks trigger aftershocks.

Acknowledgement

This study has partially been funded by a grant of the New Zealand Earthquake Commission and US NSF/USGS SSEC awards for Projects 07168 and 07175. Helpful discussions with Yan Kagan, David Rhoads and Euan Smith are gratefully acknowledged.

References


Figure 2: Empirical functions of the proportions of foreshocks (blue triangles) in the spontaneous seismicity and triggered foreshocks in the clustering seismicity (orange pluses) in the JMA, Southern California and New Zealand catalogs. Solid lines: theoretical $H_b(m)$ (blue), $H_c(m)$ (orange) and $\bar{H}(m)$ (black) in Equations (18), (19), and (12), respectively.

Figure 3: Empirical productivity functions of spontaneous events (blue triangles), triggered events (orange pluses) and all the events (black circles) from the JMA catalog, the New Zealand catalog, and the Southern California catalog are shown in (a), (b), and (c), respectively. Solid lines representing $\hat{A}e^{\hat{\alpha}(m-m_0)}$ where $\hat{A}$ and $\hat{\alpha}$ are the maximizing likelihood estimates of $A$ and $\alpha$ respectively. Panels (d), (e), and (f) show the empirical magnitude probability densities of spontaneous events (blue triangles), triggered events (orange pluses) and all the events (black circles) from the JMA catalog, the New Zealand catalog, and the Southern California catalog, respectively.
Figure 4: Same as in Figure 3, expect that the results from the analysis on the datasets with a magnitude threshold 0.3 higher.

Figure 5: Same as in Figure 2, except that the results from the analysis on the datasets with a magnitude threshold 0.3 higher.
Figure 6: (a) Empirical productivity functions of spontaneous events (blue triangles), triggered events (orange pluses) and all the events (black circles) for a simulated catalog. (b) Empirical magnitude probability densities of spontaneous events (blue triangles), triggered events (orange pluses) and all the events (black circles) from the simulated catalog. (c) Empirical functions of the proportions of foreshocks (blue triangles) in the spontaneous seismicity and triggered foreshocks in the clustering seismicity (orange pluses) in the simulated catalog. Panels (d) to (f) are the same as (a) to (c) except that the analysis is on the same simulated data with a higher magnitude threshold. See Figures 2 to 5 for more explanations.