Short-term probabilistic aftershock hazard mapping

M.C. Gerstenberger
Institute of Geological & Nuclear Sciences, Lower Hutt, New Zealand.

S. Wiemer
Eidgenössische Technische Hochschule Zurich, Switzerland.

L.M. Jones
US Geological Survey, Pasadena, California, USA.

ABSTRACT: We estimate the short-term hazard associated with strong earthquake shaking by combining the existing time-independent New Zealand hazard model with models describing the time-dependent hazard associated with earthquake clustering. We start with a generic New Zealand time-dependent kernel based on the Reasenberg and Jones aftershock model and automatically incorporate increasingly complex spatial models. By using the corrected Akaike Information Criterion, we combine all models and create a single forecast. An analogous model is currently operating in California; it produces in real-time forecasts of the probability of exceeding Modified Mercalli Intensity VI within the next 24 hours; the maps are automatically updated and available via the web. We have tested the California model using likelihood based statistical methods and determined that the model cannot be rejected and that the forecasts are improved by adding spatial complexity.

1 INTRODUCTION

1.1 Probabilistic seismic hazard analysis

Probabilistic seismic hazard analysis (PSHA) is the effort to quantify the potential for ground shaking due to earthquakes at a certain site and for the time period of interest (Cornell, 1968; Giardini, 1999). Traditional PSHA focuses on mapping the probability of exceeding a certain amount of ground shaking for a relatively long period (e.g., 50 years). While the probability at any given site may appear low the probabilities are nonetheless critical as it is these numbers that guide such things as earthquake insurance rates and building codes.

One reasonably well understood aspect of earthquake behaviour that has been slow to be developed in past PSHA efforts is the short-term hazard due to earthquake clusters such as aftershocks and foreshocks. Hazard due to these phenomena can produce drastic daily fluctuations but does typically not remain for the time period of interest in traditional PSHA mapping. For this reason, short term hazard is not of particular use in the development of building codes; however it can assist in recovery efforts from large mainshocks in developed areas. Additionally, it has implications for the insurance industry as it is not uncommon for insurance policies to exclude damage from aftershocks; depending on how aftershocks are defined in a particular earthquake clustering PSHA scheme, these definitions could be affected.

1.2 The Reasenberg & Jones model: Short term hazard

A first attempt to quantify the short term probabilities due to future aftershocks was made by Reasenberg and Jones (1989; 1990; 1994) with the development of a simple earthquake clustering model. Using two established laws in seismology, they calculate the increase in probability of future events following any other event. Their model provides a forecast of numbers of expected events for periods ranging from days to weeks; however, lacking in spatial information, no detailed location information is provided.
To model the distribution of magnitudes Reasenberg and Jones use the Gutenberg-Richter relationship (Gutenberg and Richter, 1944):
\[
\log_{10} N(M_e) = a - bM_e
\]  
where \(a\) is the productivity constant, \(b\) is related to the ratio of small events to large events, \(M_c\) is the magnitude above which all events have been recorded (magnitude of completeness), and \(N\) is the number of events greater than \(M_c\). Time dependence is incorporated using the modified Omori law (Utsu, et al., 1995) which describes the decay in aftershock productivity:
\[
N(t) = \frac{k}{(t - c)^p}
\]  
where \(t\) is the time since the mainshock, \(k\) is a function of the number of events, and \(c\) and \(p\) are empirically determined constants with \(p\) related to the decay in production of aftershocks. A simple schematic of how probabilities due to both foreshocks and aftershocks develop in their model is illustrated in Figure 1.

Reasenberg and Jones applied this model to California and, following the occurrence of a large mainshock, they were able to issue short-term probabilities of exceeding a given magnitude within the aftershock zone.

In order to design the next generation of time-dependent hazard assessment, we build upon this success by overcoming the following shortcomings of the technique:

- No spatial information is computed, only bulk seismicity parameters for an assumed point source are calculated; thus, no maps are calculated.
- No true hazard assessment is performed; forecasts are expressed as probabilities of an event above an arbitrarily assigned magnitude for an entire aftershock sequence region.
- No real-time automatic computation or distribution of the hazard information.

1.3 Spatial variations in seismicity

An important aspect in the improvement of the Reasenberg and Jones model is the exploration and understanding of how the behaviour of earthquakes can vary from region to region. Wiemer and Katsumata (1999) and Wiemer et al. (2002) have shown that systematic changes in the magnitude distribution, productivity and decay rate of earthquakes are observed when mapping aftershock sequences in detail. These observations are poorly understood but are likely to reflect regional differences in the stress field. Importantly, these changes impact the seismic hazard for the region. It is
clear that a region that produces more earthquakes has a greater hazard of future earthquakes; regional
differences in the size distribution are much less understood and are often ignored in seismic hazard
calculations. It is a logical investigation to include these changes into hazard calculations. If regions
with the propensity for larger events sustain this pattern, they will exhibit a greater seismic hazard than
regions producing only smaller events (Wiemer, 2000; Wiemer et al., 2002). An example of spatial
variability in seismicity parameters is shown in Figure 2 for the 1992 Landers and Big Bear, as well as
the 1999 Hector Mine aftershocks sequences.

**Figure 2:** Maps showing three seismicity parameter estimates from the Landers and Hector Mine mainshocks.
A) Aftershock productivity; the number of observed events greater than the magnitude of completeness level. B) The
\( b \) value; a high \( b \) value corresponds to a region with relatively more small events and a low \( b \) value indicates
the opposite. C) Temporal change in aftershock productivity; a high \( p \) value indicates areas where the aftershock
productivity slowed relatively quickly while a low \( p \) value indicates that the productivity remained high for
longer periods of time.

1.4 **Short term hazard assessment**

To best make seismic hazard information understandable and useful to potential end users (scientists,
engineers and the public), PSHA is commonly portrayed in a hazard map. By summing up the
contribution from all magnitudes at each location, a map allows for an interpretation of what the
probability is for ground shaking at any location. As this format has become the *de facto* standard in
PSHA, potential users are familiar with and able to efficiently utilize the information.

1.5 **Real time processing**

Because the changes in hazard caused by earthquake clusters can change from day-to-day, or even
hour-to-hour at the beginning of an aftershock sequence, it is necessary for the PSHA to be done in
real-time; the workload that this generates is most easily accomplished in an automated process. Not
only is this an IT task, but it entails a better scientific understanding of processes related to aftershock
occurrence (e.g. aftershock zone size).

2 **A REAL TIME HAZARD MODEL**

We calculate the time-dependent seismic hazard by combining stochastic models: starting from a
background model, which is a long-term, stationary but spatially-varying Poisson hazard model, and
depending on the amount of currently available information, we add clustering model elements of
varying complexity. For both California and New Zealand, we have chosen to adopt the national
hazard maps as the background model. Most of the time a recent earthquake large enough to generate
any appreciable increase in hazard will not have occurred and we then assume that the background
model adequately represents the hazard. After any event of $M3$ or larger, the hazard associated with possible aftershocks is added to the background hazard to complete the hazard setting.

2.1 Model elements and spatial heterogeneity

Our clustering model utilizes the method of Reasenberg and Jones and is therefore based on the four seismicity parameters, $a$, $b$, $c$, and $p$. In order to optimise the use of available information, our model is a composite of three model elements that differ from each other in complexity and input data used: 1) a generic-clustering element; 2) a sequence-specific element; and 3) a spatially heterogeneous element. In the generic clustering element, a previously derived set of seismicity parameters is used and therefore the rate and location of aftershocks depends solely on the mainshock magnitude and location. For both New Zealand and California we use the median parameter values as estimated from aftershock sequences in each region. In the case that the ongoing earthquake sequence produces a sufficient number of aftershocks, a sequence-specific element is additionally calculated that uses the \textit{a posteriori} parameter values estimated for the sequence at a given time. Concurrent with the estimation of the second element, a third, spatially heterogeneous element is derived with seismicity parameters estimated independently at each of the $i^{th}$ nodes of a 5-km grid covering the region and based on the seismicity within a maximum of 15 km of the node.

We do not limit an aftershock to occur at the same location as its mainshock. In the first two elements, the aftershock zone is initially assumed to be circular with a radius based on the Wells and Coppersmith (1994) relationship for subsurface rupture length. After 100 aftershocks have been recorded, an aftershock zone is estimated based on the spatial aftershock distribution; this zone is continually updated to account for newly recorded events. The total rate of aftershocks is determined by the model and distributed across an area that extends one-half of a fault length from the source. The spatial density of aftershocks is assumed to be proportional to $1/r^2$ where $r$ is distance from the mainshock's source. In the third element, the spatially heterogeneous element, the spatial variations in the rate are calculated directly.

2.2 Composite model and rate forecasts

We calculate the expected rate of $4 \leq M \leq 8$ aftershocks at each node and for each element using the equation described by Reasenberg and Jones:

$$\lambda(t) = \frac{10^{a' + \beta(M_w - M_0)}}{(t + c)^p}$$

(3)

where $M_w$ is the mainshock magnitude and $t$ is the time since the mainshock. We then combine the rates using Akaike weights, which are derived from the corrected Akaike Information Criterion (AICc) (Burnham and Anderson, 2002). The AICc is based on the model's likelihood, the amount of data and the number of free parameters that must be estimated, thus ensuring a gradual addition of sequence-specific and spatially-heterogeneous information to the initial generic information as sufficient data warranting a more complex model become available. Because the AICc prefers fewer free parameters, the spatially heterogeneous model will be strongly weighted only in active areas.

We treat every earthquake ($M \geq 3$) (including aftershocks) as a potential mainshock and calculate an aftershock model for it. At any time, more than one earthquake may be contributing to the expected rate of aftershocks at a given location, most commonly after a large aftershock. In these cases, the earthquake producing the highest expected rate at a location is adopted, and its rate assigned to the node. Thus, a cascade of aftershock sequences can be modelled, with secondary sequences represented where and when they exceed the primary sequence's rate. A sequence is terminated once its rate of $M \geq 5$ earthquakes falls below that of the background model at that site.

Figure 3 shows the model progression from the generic element to the more complex spatially heterogeneous element for the 1999 $M_w$ 7.1 Hector Mine, California event as seen in the forecasted rates above the background level. Initially, 1 minute before the occurrence of the mainshock, a small amount of increased rates are seen due to some small foreshock activity. Immediately after the event,
an isotropic increase in rates is observed centred on the epicentre of the mainshock. One day after the mainshock, the spatially heterogeneous model has become dominant near the rupture zone (white lines) as can be seen in the heterogeneity of the forecasted rates; this trend continues throughout the next year and beyond.

Figure 3: Forecasted number of earthquake (M>4) above the background level for the 1999 Hector Mine, California aftershock sequence. Time periods correspond to the time periods used for the hazard calculations in Figure 4.

2.3 Probabilistic seismic hazard analysis

The hazard calculations are done in using standard PSHA techniques; the hazard associated with earthquake shaking at site $S$ depends on the expected rate of earthquake sources nearby, their magnitudes and distances from $S$, propagation effects (primarily geometric spreading and seismic attenuation) and local soil conditions at $S$. For $M \geq 5.5$ sources, we use the Boore, Joyner and Fumal (1997) attenuation function, using its 0.0 s period coefficients for rock sites. For $4 \leq M < 5.5$ sources, we use the 0.0 s period coefficients estimated for $4 \leq M < 5.5$ earthquakes in California (Quitoriano, pers. comm., 2003). We currently ignore local site effects. The seismic hazard at a location is calculated for the probability of exceeding $1.126g$ in the next 24 hours, and expressed as the probability that ground shaking will exceed Modified Mercalli Intensity (MMI) VI using the relationship of Wald, et al. (1999). The hazard forecast for the time periods corresponding to the rate forecasts in Figure 3 are shown in Figure 4.

2.4 Statistical testing and model validation

To test the forecasting ability of the model, we initially tested the relative accuracy of the full composite aftershock model described above against other less complex models by applying likelihood ratio tests. These tests compare the ability of a hypothesis to forecast an observed dataset as compared to the ability of a null hypothesis, giving a probability with which the null hypothesis can be rejected. We chose not to test based on the forecasted hazard due to the inherent smoothing involved in the hazard calculation and the limited amount of observed ground motion data. For each location and magnitude bin ($\Delta M = 0.1M$, $M \geq 4$) in southern California, and the period 1992 and 2001, we compared the forecasted rates in each model with the observed numbers of events.

We compared our model against successively more challenging null hypotheses. In the first test, we showed that the background model of the 1996 U.S. National Hazard map could be rejected with a significance of <1%. In the second and third tests, we compared the full composite model to stages of the model, the second using only the generic clustering element, and the third adding the sequence-
specific element where it could be determined. We were able to again reject both partial models with 1% significance. We thus can state that the composite model did a better job of predicting the daily seismicity of 1992-2001 in southern California than the U.S. National Hazard maps, an average clustering model, or a sequence-specific parameter model.

![Image](image_url)

Figure 4: 24 hour probability of exceeding MMI VI for 6 time periods following the Hector Mine, California event. Heterogeneity in the hazard forecast is clearly observable from one week after the mainshock.

As an additional test, to check the consistency of our model with the data, we performed a likelihood test of our composite model against the observed data; we could not reject the model with a significance of <1%. Therefore we cannot reject our model as insufficiently explaining the data at our chosen significance level.

3 SHORT TERM HAZARD IN NEW ZEALAND: ISSUES AND COMPLICATIONS

The adaptation of the real time model to New Zealand involves, but may not be limited to, the following issues:

- Development of generic New Zealand seismicity parameters
- Adaptation of N.Z. national seismic hazard model including depth zones and attenuation relationship (McVerry, et. al. 2000)
- Understanding limitations of real time data quality
- Development of an algorithm to identify and properly handle swarms

Due to the complex tectonic setting of New Zealand, the national hazard model currently consists of five depth layers (Stirling et al. 2002). While these layers are necessary for the implementation of the national hazard map, it is impractical to develop a suite of generic parameters appropriate for this
complexity; this is primarily due to a lack of aftershock data that would allow us to calculate separate parameters for each depth zone. Previous investigations of generic parameters for N.Z have not attempted to handle the depth issues (Eberhart-Phillips 1998, Pancha, et.al. 2004). Therefore we currently use a single set of parameters for all events.

One issue that requires further research is the handling of earthquake swarms. This type of behaviour frequently occurs throughout New Zealand and does not follow a power law relationship as is described by the modified Omori law: the basis for our model. An algorithm must be emplaced that can automatically identify swarm-like behaviour based on seismicity and location.

A limitation in the real-time application is data quality, and specifically catalogue completeness. In the early hours and days after a large event, such as the 2003 Fiordlands event (Figure 5), the network becomes swamped with events; during this time, the catalogue completeness generally remains very high and, due to a lack of data, restricts the hazard model to using only the generic element. In the case of the Fiordlands event, using a catalogue updated 3 weeks after the mainshock, too few events are available to use a more complex model element than the generic.

Figure 5: 24 hour hazard forecast calculated 1 minute following the $M_w$ 7.1 Fiordlands event on August, 22, 2003.

4 CONCLUSIONS

Building upon the simple concepts of the Reasenberg and Jones model, we have developed a short term hazard model with the following results:

- Twenty-four hour hazard forecasts for New Zealand and California;
- Statistically based earthquake clustering model including foreshock and aftershock hazard;
- Combines long-term poissonian hazard information with hazard estimates based on current earthquake data;
• Additional aftershock complexity is conservatively added by using AICc;

• Statistical testing has shown: 1) our model sufficiently fits historical data; 2) the added complexity of spatial heterogeneity better fits the data than less complex models;

• New Zealand provides a more difficult proving ground than California, requiring the model to be able to handle additional complexities such as depth zones and swarms. Additionally, data quality will be a significant issue when operating in real time.

• The model is currently operating in California with real time hourly updates at http://step.gps.caltech.edu.

REFERENCES:


Quitoriano, V. 2003. personal communication.


