Tests of Seismic Hazard Models

by D A Rhoades, W D Smith and M J Stirling

Tests of Seismic Hazard Models

by D A Rhoades, W D Smith and M J Stirling

Prepared for

Earthquake Commission Research Foundation Project No. 01/460

CONFIDENTIAL

Institute of Geological & Nuclear Sciences client report 2011/99
Project Number: 410W3110

The data presented in this Report are
available to GNS for other use from
September 2002

August 2002

CONTENTS

EXECUTIVE SUMMARY

Tests are described for checking the adequacy of a seismic hazard model against independent data. They have been applied to the national seismic hazard model (NSHM) for New Zealand, developed by Stirling *et al*. (2000; 2002). Significant discrepancies revealed by these tests direct attention towards the most unsatisfactory aspects of the NSHM, including data and assumptions.

Frequency of surface fault ruptures

The Poisson process assumption provides a basis for statistical tests comparing the frequency of surface fault rupture in the seismic hazard model with the historical record of surface fault rupture.

Application of such tests to the NSHM has revealed a statistically significant discrepancy between the model and the historical record. The NSHM predicts a frequency of surface fault rupture in New Zealand earthquakes more than five times the historical rate, with more than 50 fault ruptures expected since 1840 and only 10 observed. The discrepancy is most marked in the Taupo Volcanic Zone, but is also appreciable in northern Canterbury and southern Fiordland. It is likely to be due to multiple causes, including underestimation of the mean recurrence interval between ruptures of some fault sources in the NSHM, a possible tendency for fault ruptures to occur in clusters, and probable incompleteness in the record of historical surface fault ruptures.

Times of last rupture and mean recurrence intervals

Under the stationary Poisson model, the elapsed time since the last rupture on any fault has an exponential distribution. If the recurrence interval is known, this allows the construction of a statistic, from the elapsed time, which has a common uniform distribution for all faults. Using the data from all faults together, goodness of fit tests can be used to check the overall consistency of the elapsed time and recurrence interval data.

Application of such tests to the NSHM has not revealed any significant discrepancies. The statistical power of the tests is weak because of the small proportion of fault sources for which any information on the time of last rupture is available. Future paleoseismic studies should aim to increase the available information constraining the times of last rupture of faults.

Seismic moment release and strain accumulation

Geodetic data provide estimates of strain accumulation in any area of the crust and these can be converted into estimates of moment accumulation rate. Similarly seismic hazard models

provide estimates of rates of occurrence of earthquakes of any given magnitude in any area, and these can be converted into estimates of the rate of moment release. If most of the moment rate is released in earthquakes, the ratio of the seismic moment rate to the geodetic moment rate should be not much less than one.

A comparison of the seismic moment rate in the crustal seismotectonic zones of the NSHM with that expected from geodetic data has shown that overall the strain released under the model accounts for 0.6-0.8 of the geodetic moment rate. Within the individual seismotectonic zones, the ratio of NSHM moment to geodetic moment varies between 0.05 and 2. Further research is needed to understand these discrepancies. The present analysis does not include the moment release from the subduction interface and slab sources of the NSHM, because it is not known how much of the associated strain accumulation would be reflected in strain measurements at the surface.

Magnitude-frequency relation

Catalogues of earthquakes covering long time intervals and large areas are found nearly always to conform to the Gutenberg-Richter magnitude-frequency relation with modifications to accommodate the limit on the maximum size of an earthquake in the Earth's crust. Seismic hazard models of a region should therefore conform both with this relation and with the magnitude-frequency distribution in the historical record.

Comparison of the NSHM magnitude-frequency distribution with the historical record and with a plate motion-balanced frequency-magnitude distribution has revealed some discrepancies between the model and the data, and between the model and the Gutenberg-Richter relation. While the magnitude-frequency relation in the NSHM shows wide scatter above M7.0, there is a surplus of earthquakes in the range M7.0 - M7.3 in the NSHM when compared with either historical data or the Gutenberg-Richter relation. Within the same magnitude range there is a deficit of distributed-source earthquakes and a more-than-compensating surplus of fault-source earthquakes under the NSHM. Modifications to the NSHM should be considered in order to rectify these discrepancies. In particular:

- the characteristic magnitude and/or mean recurrence interval probably needs to be adjusted for some fault sources;
- the upper magnitude limit of distributed earthquakes may need to be increased for some seismotectonic zones and a roll-off rather than a sharp magnitude cut-off adopted;
- variability in the magnitude of earthquakes on a given fault source should be allowed for.

The distribution of the difference in magnitude between largest and the next few largest earthquakes under the NSHM was found to be consistent with the historical record.

1.0   INTRODUCTION

In probabilistic seismic hazard analysis (PSHA), information on the frequency and size of rupture of active faults and on historical seismicity from a study region is used to derive a seismicity model.   Then models of strong-motion attenuation are applied to calculate the hazard at any site (Cornell, 1968; Kulkarni *et al*. 1979). The construction of the seismicity model involves various assumptions and standard procedures. For example, it is central to the use of fault data to assume that the earthquakes on a fault have a particular characteristic magnitude and average recurrence interval that can be derived from paleoseismic studies through measurements of fault dimensions and slip rate (Wells and Coppersmith 1994; Somerville *et al*. 1999).   Also a large fault is sometimes divided into segments which are assumed to rupture independently (Schwartz and Coppersmith, 1984; Wald and Heaton, 1994). Again in using seismicity data, a large region is usually divided into a number of smaller regions and for each sub-region the seismicity parameters are estimated and a maximum magnitude is set. Moreover, it is the usual practice to decluster a seismicity catalogue according to some standard method (e.g. Reasenberg 1985) before computing the seismicity parameters.

The resultant hazard model is often thought to be reliable if it incorporates all the current-best-practice standard procedures, and to be beyond any further testing. In any case, testing of such models against subsequent earthquakes is such a long-term process that it is impracticable.

However, a model can in principle be checked against other information that has not been used directly in deriving it. Dowrick and Cousins (2002) have recently compared the historical incidence of Modified Mercalli intensity with that expected under the New Zealand national seismic hazard model (NSHM) developed by Stirling *et al*. (2000; 2002). In this study, we focus on the seismicity model component of a PSHA, and set out to define summary statistics and formal tests by which its adequacy can be checked.   These include tests of consistency of the predicted frequency of surface fault rupture with the historical record, the distribution of estimated times of last rupture of surface faults with the distribution of recurrence times, the seismic moment rate of the seismicity model with that expected from plate motion, and the magnitude and frequency distribution of earthquakes with the historical record and the plate motion. Where practicable, the tests are applied to the NSHM. The purpose is to identify any serious shortcomings in the current model, and thereby to pinpoint areas where improvements can be made to the databases, assumptions and procedures applied in PSHA.

The long-term and large-scale statistics of the earthquakes generated by a seismicity model should conform with long-term and large-scale statistics of real earthquakes and with the long-term plate motion. Because of the way in which seismicity models are constructed, there

is no guarantee that this will be so.

The seismicity model used in the NSHM has the same general features as other PSHA models described above. The recurrence intervals of the surface rupturing faults have been estimated independently. It is therefore of interest to know how well the overall rate of occurrence of surface rupturing earthquakes in the model compares with the historical record of such events over the past c.160 years.

Data on the estimated time elapsed since the last rupture of individual surface faults are not used directly in constructing the seismicity model, and therefore provide another possible consistency check. The expected distribution of such times can be derived from the distribution of average recurrence times and the Poisson assumption. A comparison of the expected and actual distributions will throw further light on the adequacy of the active fault database.

Data on surface strain accumulation, as revealed by geodetic data, provide another means of checking the adequacy of a seismic hazard model. The strain accumulation rates can be used to derive estimates of the moment accumulation rate in seismogenic zones. The overall rate of seismic moment release predicted by the seismic hazard model should be consistent with the overall rate of moment accumulation.

The fault-source contributions to seismic hazard are derived from fault dimensions and slip rates, and are not constrained by consideration of the overall magnitude-frequency distribution in the model. Nevertheless, it is desirable that the model should produce a distribution of magnitudes that is consistent overall with both the historical earthquake catalogue and the ubiquitous Gutenberg-Richter magnitude-frequency relation for earthquake populations.

2.0   FREQUENCY OF SURFACE FAULT RUPTURES

2.1   Comparing the predicted frequency of surface fault ruptures with the historical record

2.1.1   Stationary Poisson Process

In seismic hazard modelling, it is usual to assume that the major earthquakes follow a stationary Poisson process. Although it has long been recognised that earthquakes are clustered in the short term, and that such clustering represents a significant deviation from a stationary Poisson process, there is little evidence for deviation from Poisson behaviour for the larger earthquakes that contribute most to earthquake hazard in the long run. For example, in the case of New Zealand earthquakes, McGinty (2001) has shown that after de-clustering to remove aftershocks, the New Zealand earthquakes since 1964 can be approximated well by a

stationary Poisson process. Hence probabilities of a given level of shaking over long-periods of time are calculated using the stationary Poisson assumption. Thus, if λ is the modelled average rate of occurrence of events per unit time, the number N of events expected in a time-period of length t is given by:

$$E(N) = \lambda t \tag{1}$$

and the probability P of n such events occurring is given by:

$$P(N = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \tag{2}$$

If, in a time-period of length t, the observed number of events is k, then the consistency of that observation with the model can be assessed by summing tail-probabilities, as follows.

1. If $k > \lambda t$, calculate

$$P(N \geq k) = \sum_{n=k}^{\infty} P(N = n) \tag{3}$$

2. If $k < \lambda t$, calculate

$$P(N \leq k) = \sum_{n=0}^{k} P(N = n) \tag{4}$$

The statistical significance of any deviation from the model can then be quantified by the smallness of the probability $P(N \geq k)$ or $P(N \leq k)$ respectively.

The Poisson distribution has the property that the variance is equal to the mean, i.e.

$$Var(N) = \lambda t \tag{5}$$

For large values of λt, the probability of N taking its value in some interval can thus be approximated by the probability of a normal distribution with mean λt and standard deviation $\sqrt{\lambda t}$ taking its value in the same interval. This approximation is justified by the central limit theorem. However, it must be applied with care, since the normal distribution is continuous and N is discrete (integer-valued). For a good approximation, the end-points of the interval should be chosen midway between integer values.

2.1.2 Cluster Processes

Alternatives to the stationary Poisson process model have been proposed, but have not been widely adopted in seismic hazard modelling. For example, Jackson and Kagan (1999) showed that the annual number of earthquakes exceeding magnitude 5.75 in large regions such as the NW Pacific and SW Pacific are better fitted by the negative binomial distribution than the Poisson distribution. This distribution has two parameters, τ and ρ, and the average seismicity rate is $\tau(1-\rho)/\rho$. The probability distribution for the number of events in time t is given by

$$P(N = n) = \frac{\Gamma(t\tau + n)}{\Gamma(t\tau)n!}\rho^{t\tau}(1-\rho)^n \quad \tau > 0,\ 0 < \rho < 1 \quad (6)$$

The expected value and variance are given by

$$E(N) = t\tau(1-\rho)/\rho\,; \qquad\qquad\qquad (7)$$

$$Var(N) = E(N)/\rho \qquad\qquad\qquad (8)$$

It is equation (8) that gives the clue as to why such a model might fit better than the stationary Poisson model. It allows for *Var(N)* to be greater than *E(N)*. Jackson and Kagan (1999) estimated values of ρ ranging from 0.32 to 0.57, corresponding to the ratio *Var(N)/E(N)* ranging from 1.8 to 3.1. This ratio may be interpreted as a mean cluster size, as is shown below.

Suppose that events occur in clusters of fixed size k, and that the clusters themselves follow a stationary Poisson process with rate parameter μ. Then the number of events occurring in time t has expected value given by

$$E(N) = kE(M) = k\mu t\,, \qquad\qquad\qquad (9)$$

where M is the number of clusters, and variance given by

$$Var(N) = k^2 E(M) = k^2 \mu t \qquad\qquad\qquad (10)$$

Hence, for such a process

$$Var(N)/E(N) = k \qquad\qquad\qquad (11)$$

Now suppose that instead of all being the same, cluster sizes K are independently and identically distributed according to a distribution with expected value *E(K)* and variance

var(K). Then the number of events in time *t* has expected value

$$E(N) = E(K)\mu t \tag{12}$$

and variance

$$Var(N) = Var(\overline{K}M) = E(\overline{K})^2 Var(M) + E(M)^2 Var(\overline{K}) + o(t) \tag{13}$$

where $\overline{K}$ is the sample mean cluster size over *M* clusters, the last step is by the approximate formula for the variance of a function of random variables (Kendall and Stuart, 1977, p247) and $o(t)/t \rightarrow 0$ as $t \rightarrow \infty$. Noting that $Var(\overline{K}) \rightarrow 0$ as $t \rightarrow \infty$, we have

$$Var(N) = E(K)^2 \mu t + O(t) \tag{14}$$

where $O(t) \rightarrow 0$ as $t \rightarrow \infty$. Hence, for large t, $Var(N)/E(N) \cong E(K)$, the mean cluster size. For further discussion of cluster processes, see Vere-Jones (1970).

In seismic hazard models, fault-rupturing events are usually assumed to occur independently. Actually, it is possible that more than one fault might rupture in a single earthquake, and that that the rupture of a particular fault might trigger the rupture of a neighbouring fault a short time later. Whereas seismicity data are usually de-clustered before being used in a seismic hazard model, it is not possible to similarly de-cluster the active fault data, since it is usually impossible to determine from paleoseismic studies whether faults have ruptured close together in time in the past. Insofar as there is dependence between faults, we may expect to see discrepancies between the historical distribution of fault rupture and that predicted by a seismic hazard model.

2.2 Application to the national seismic hazard model

2.2.1 Historical fault ruptures

The active faults that form the database for the NSHM (Appendix 1 of Stirling *et al*. (2000)) include faults characterised entirely by paleoseismic studies, faults characterised partly from the magnitudes of historical earthquakes, and faults which have no onshore surface expression. The latter category includes, for example, 15 notional fault segments on the Hikurangi subduction interface, with assessed characteristic magnitudes ranging from 7.5 to 8.4. In Table 1 are listed all the fault sources of the NSHM which are considered to have ruptured to the surface in historical earthquakes (1840-2002). The total number of such faults is 10. It should be noted that two of the faults in this list, Napier (1931 rupture) and Inangahua, are essentially "buried" faults, the extent of which could not have been identified from the surface rupture alone, i.e., without knowledge of the corresponding earthquakes.

It has to be acknowledged that not every historical surface fault rupture would necessarily have been noticed at the time of the corresponding earthquake occurrence. This is particularly true in the nineteenth century, when even large earthquakes distant from populated areas might not have been noticed, or at least their large size might not have been appreciated from their felt effects. For example, a large earthquake on an offshore fault might not have been recognised. The instrumental record of earthquakes in the 20th century is such that New Zealand earthquakes of magnitude 7.0 and above since 1901, and those of magnitude 5.5 since 1940, would likely have been recorded on numerous seismographs either in or outside of New Zealand (Dowrick and Rhoades, 1998). Nevertheless, association of offshore earthquakes with faults, notional or otherwise, is problematical, and is not attempted here.

Table 1   Faults from NSHM database which are considered to have ruptured in historical earthquakes (1840-2002)

Fault index	Fault name	Date of rupture (AD)	NSHM Earthquake magnitude	Mean recurrence interval (yr)
4	Awatere NE	1848	7.5	1000
58	Wairarapa (1855 rupture)	1855	8.1	1500
11	Hope (1888 rupture)	1888	7.2	120
140	Kaiapo	1922[*]	6.2	731
52	White Ck	1929	7.6	34000
151	Arthur's Pass (1929 rupture)	1929	7.0	3500
170	Napier (1931 rupture)	1931[#]	7.8	2500
274	Waipukaka	1934	7.6	1900
55	Inangahua	1968[#]	7.4	4400
110	Edgecumbe (1987 rupture)	1987	6.5	1362

[#] Surface rupture insufficient to indicate size of earthquake

[*] Rupture resulted from earthquake swarm of lower magnitude events

2.2.2   Discrepancy between data and model

We now compare the above observations with the number of expected fault ruptures over the same period under the NSHM. For this purpose a synthetic catalogue of 10,000 years of earthquakes was generated, using the assumptions of the NSHM. This synthetic catalogue was found to include a total of 3260 events rupturing faults in the database (Figure 1). The number of such events expected in 162 years is thus 52.8. According to the normal approximation described above, the probability of a value less than 10.5 occurring is 2.9×10^{-9}, i.e., extremely small. On this basis, it appears that the historical data on fault ruptures differs so significantly from the NSHM that the NSHM can be confidently rejected.

There are several possible causes of the discrepancy between the data and the model.

1.   Incompleteness of the data in Table 1

If the data in Table 1 are anywhere near complete, the rejection of the model still stands. For example, if we suppose that Table 1 contains only half of the faults that have ruptured in historic time, i.e. that there have actually between 20 fault ruptures in historic times, then the above probability would still be less than 0.002. About a third of the fault ruptures in the synthetic catalogue occurred on offshore faults. It is to be expected that there have been ruptures of offshore faults in the historical period also and that the list of historical fault ruptures is incomplete. It is therefore probable that incompleteness of the data in Table 1 is a contributing factor, but it seems unlikely that it is the only cause of the discrepancy between the data and the model.

2.   Underestimation of average recurrence intervals of fault rupture

Large systematic errors in the average recurrence intervals for rupture of faults in the database would needed to account for the discrepancy. To accommodate the observed data within the 95% tolerance limits, a decrease in the expected number of events to somewhere in the range 5.7 to 19.3 is required. This represents a decrease of the fault rupture occurrence rate by a factor in the range 0.11 to 0.37. Such a decrease could be achieved by, for example, increasing each average recurrence interval in the database by a common factor in the range 2.7 to 9.0. If such underestimation is the cause of the discrepancy, it would suggest that there is a flaw in the methodology of estimating mean recurrence intervals for fault rupture or in its application to New Zealand faults. It should be noted that if there exist active surface-rupturing faults that are not in the present database, the discrepancy would be exacerbated. It seems probable that some undiscovered active faults exist, because some faults in the existing database, such as those at Napier and Inangahua, could not have been characterised without the occurrence of historical earthquakes on them.

3.   Clustering of fault-rupturing events

We apply the theory developed in section 2.1.2 above. In order to account for the discrepancy by clustering, we need an increase in the standard deviation of the number of faults in 162 years by a factor of at least 2.9, i.e. an increase in the variance by a factor of at least 8.4. Thus, according to the analysis in section 2.1.2, a mean cluster size of at least 8.4 is required. In view of the scant evidence for clustering of historical fault ruptures, such a mean cluster size seems to be extremely unlikely.

None of these possible causes provides, by itself, a satisfactory explanation of the

discrepancy. However, two or more factors may be interacting to produce the discrepancy.

2.2.3 Spatial distribution of fault ruptures

It is instructive to compare the locations of historical fault ruptures with their expected distribution under the NSHM. Figure 2 shows the number of historical fault ruptures considered to have occurred in one-degree windows over the period 1840-2002. Figure 3 shows the number expected under the NSHM in any period of 162 years in the same windows. The biggest differences between these two figures are found in the central North Island (especially the Bay of Plenty region), the North Canterbury region, and south Fiordland. These areas are highlighted in Figure 3. In all of these three regions the modelled number of fault ruptures exceeds the observed number by a wide margin.

In the Fiordland region, no fault ruptures have been observed where about 11 would be expected under the NSHM. Such a deviation from the expected number would occur once in about 60,000 such periods due to random fluctuations under the stationary Poisson model. But this mostly offshore region is one area where the historical record could well be incomplete.

In North Canterbury two fault ruptures have been observed where about seven would be expected under the model. Such a deviation from the expected number would occur once in about 20 such periods due to random fluctuations under the stationary Poisson model.

In the central North Island region one surface fault rupture has been observed where about 15 would be expected under the model. (The one surface fault rupture that did occur, on the Kaiapo fault (Table 1), resulted from a swarm of earthquakes with no individual magnitude as large as the NSHM characteristic magnitude.) Such a deviation from the model would occur only once in about 400,000 such periods due to random fluctuations under the stationary Poisson model. Therefore the NSHM, in its present formulation, may confidently be rejected in this region. It appears that either the average rate of occurrence of surface fault rupture has been greatly overestimated (especially in the Taupo Volcanic Zone) or the stationary Poisson model is a poor model for fault ruptures in this region. If the latter is true, then one possible explanation could be that the many separate faults identified in this region do not rupture independently.

Research to produce an improved fault-source model for the Taupo Volcanic Region is now being undertaken.

3.0 DATES OF LAST RUPTURE OF SURFACE FAULTS

3.1 Comparing the distribution of estimated times of last rupture of surface faults with the distribution of mean recurrence intervals

Under the stationary Poisson model, the hazard is constant in time, and events are distributed uniformly in time. The inter-event-time distribution that produces a constant hazard rate is the exponential distribution. For an individual fault with mean recurrence interval *T*, the cumulative distribution for the time between successive events is thus given by:

$$F(t) = 1 - \exp(-t/T) \tag{15}$$

Because the stationary Poisson process has no memory, the time to the next event from any randomly chosen point in time has the same distribution as the time between successive events. The same is true if time is reversed, i.e. at any randomly chosen point in time, the time elapsed since the last event has the same exponential distribution indicated by equation (15). It follows that if fault ruptures follow the stationary Poisson model, and if *t* is the time elapsed since the last rupture of a fault and *T* the mean recurrence interval for events on the fault, then $F(t)$ calculated as in equation (15) is distributed uniformly on the interval (0,1).

This property can in principle be used as a test of seismic hazard models that are based on the stationary Poisson process. If the time elapsed since the last rupture is known for a sufficiently large sample of active faults, and if the sample can reasonably be considered as a random (or representative) sample of all faults in the database, then the distribution of $F(t)$ over faults in the sample can be formally compared with a uniform distribution by any one of a number of standard statistical tests, some of which are described briefly below.

3.1.1   Binomial test of number of values in a class interval

Suppose that n_i is the number of observations, from a sample of *n* values of $F(t)$, that fall in some class interval $I_i = (a_i, b_i)$. If $F(t)$ follows a uniform distribution, then the probability p_i of any individual value falling in I_i is $b_i - a_i$. Thus the value n_i can be compared with the outcome of a *n* binomial trials with probability p_i of success. For large enough *n*, the normal approximation to the binomial distribution can applied to give symmetric tolerance limits for n_i about the expected value np_i.

3.1.2   Chi-square test of goodness of fit

The chi-square test (e.g. Fisz, 1963, p426) is constructed from the counts in all *k* class intervals in a histogram. The test statistic is defined by

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \tag{16}$$

This statistic can be compared to a chi-square distribution with *k*-1 degrees of freedom, to see if the sample conforms to the theoretical distribution. It is important that numbers in each class interval should not be too small, since the test is based on the asymptotic distribution of

the statistic as $n \to \infty$. The minimum expected number in a class interval should not be less than about 10.

3.1.3 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test (e.g. Fisz, 1963, p445) is a test based on the maximum deviation of the empirical distribution from the theoretical distribution. Let $S(x)$ is the theoretical cumulative distribution (uniform in this case) and $S_n(x)$ the empirical distribution of the n sample values, i.e.,

$$S_n(x) = n(x)/n \qquad (17)$$

where $n(x)$ is the number of sample values $\leq x$. The test statistic is

$$\lambda = \sqrt{n} \sup_x | S(x) - S_n(x) | \qquad (18)$$

The limiting distribution of λ as $n \to \infty$ when the sample conforms to the theoretical distribution is the so-called Kolmogorov-Smirnov distribution. From tabulations of this distribution, it is found that values of λ greater than 1.36 and 1.63 indicate a deviation of the data from the theoretical distribution at the 0.05 and 0.01 levels of significance, respectively.

3.2 Application to the national seismic hazard model

3.2.1 Synthetic catalogue

Let us first illustrate the application of these tests to a synthetic catalogue that conforms exactly to the NSHM. In a 10,000-year synthetic catalogue, earthquakes occurred on 249 of the faults in NSHM database. For each of these faults, the time t between the time of last rupture and the end of the catalogue was found and $F(t)$ computed by equation (15), with T set to the average recurrence interval for ruptures of that fault under NSHM. A histogram of the resulting distribution values of $F(t)$ is given in Figure 4. This plot also shows the expected number in each class interval as a solid line, and, as dashed lines, the 95% tolerance limits for the number in any class interval, calculated according according to the normal approximation to the binomial distribution (as discussed in 3.1.1 above). It can be seen that numbers in each class interval fall within the limits, indicating no deviation from the theoretical uniform distribution according to the binomial test. However, the numbers in the class intervals have the appearance of a downward trend. Could this trend be real? The chi-square test seems to indicate that it is; for this data χ^2, calculated as in equation (16), is 36.0, much greater than the critical value of 21.7 at the 0.01 level of significance. The Kolmogorov-Smirnov test confirms that the data do not conform to the theoretical uniform distribution; here λ (equation 16) is 2.49, much greater than the critical value of 1.63 at the 0.01 level of significance.

How can it be that a catalogue generated to conform exactly to the theoretical model does not produce times of last rupture consistent with that model? The answer lies in the non-randomness of the sample produced by a 10,000-year catalogue. All the faults that did not rupture in the 10,000-year catalogue are excluded from the data in Figure 4. To balance the sample, we need to add information from these faults to the sample. All that we know, for these faults, is that the elapsed time since the last rupture is greater than 10,000-years, i.e., $F(t)$ lies somewhere between $F(10,000)$ and 1. We assume that anywhere in this interval is equally likely and re-compute the histogram (Figure 5). The heights of the bars now represent expected numbers, rather than actual counts, in the class intervals, because we have averaged over the possible values of $F(t)$ for the unruptured faults.

It can be seen in Figure 5 that the downward trend apparent in Figure 4 is no longer present. The chi-square test confirms this; now $\chi^2 = 3.3$, a value that is much less than the critical value of 16.9 at the 0.05 level of significance. Also the Kolmogorov-Smirnov λ statistic is now 0.50, much below the critical value at the 0.05 level. Hence it can be seen that the augmented data conform well to the theoretical distribution.

3.2.2   Prehistoric fault-ruptures

The above analysis is a warning of how easy it is to inadvertently construct samples that are not random, or not properly balanced across all of the faults in the database. In particular, it is apparent that when we consider the actual data on past fault ruptures, the data on historical fault ruptures constitute a hopelessly biased sample of elapsed times, that is weighted strongly towards short elapsed times. Such a sample is of no use when applying the present tests. We therefore need to assemble the information, such as it is, on pre-historic fault ruptures.

Only limited information on the date of last rupture is available for faults that are not known to have ruptured in historical time. In Table 2, we have brought together information on the last rupture of active faults in the NSHM database. These data are taken from three sources: the GNS active fault database, the Mouslopoulou *et al*. (2001) compilation for the International Lithosphere Project World-wide Earthquake Recurrence Database and a study of earthquake sources in the Canterbury region by Pettinga *et al* (2001). The index number is from Appendix 1 of Stirling *et al*. (2000). The last rupture dates or intervals are mostly from C^{14} dates expressed as years B.P., i.e. prior to 1950, and are only approximate for calendar years. The mean recurrence intervals are taken from Appendix 1 of Stirling *et al*. (2000). F(elapsed time) is calculated using equation (15), and treating the dates of last rupture as if they were exact for calendar years.

Table 2   Information on times of last rupture, mean recurrence interval and cumulative exponential distribution of elapsed time for fault sources in the NSHM database. See Table 1 for historical ruptures.

Index	Fault name	Last rupture (yr B.P. unless otherwise stated)	Mean recurrence interval (yr)	*F*(elapsed time)
1	Wairau	>1400[2]	1650	0.57 – 1
3	Awatere SW	522-597[3]	2930	0.16 – 0.18
5	Alpine (Milford–Haupiri)	1717 AD[2]	300	0.61
9	Clarence NE	1700-1880[1]	1500	0.68 – 0.71
10	Clarence SW	1700-1880[1]	1080	0.79 – 0.82
12	Hope (Conway-Offshore)	1838 AD[3]	120	0.75
18	Omihi	<10000	474	0 – 1.00
19	Lowry	>10000	5000	0.86 – 1
20	Culverdon	1495-1925[3]	7500	0.18 – 0.23
22	Mt Grey	300-450[3]	3300	0.09 – 0.13
27	Harper	>10000[3]	10000	0.63 – 1
28	Porters Pass	500-700[3]	2900	0.16 – 0.21
32	Pegasus 1	>10000	10000	0.63 – 1
33	Pegasus 2	>10000	10000	0.63 – 1
34	Pegasus 3	>10000	10000	0.63 – 1
37	Lake Heron	<10000	5000	0 – 0.86
39	Mt Hutt-Mt Peel	<10000	7500	0 – 0.74
40	Fox Peak	<10000	7000	0 – 0.76
41	Hunter Hills Nth	>10000	15000	0.49 – 1
42	Hunter Hills Sth	>10000	15000	0.49 – 1
43	Dryburgh SE	>10000	22000	0.37 – 1
44	Dryburgh NW	>10000	22000	0.37 – 1
46	Wharakuri	<20000	10000	0 – 0.86
48	Waitangi	<20000	50000	0 – 0.33
74	Ostler Nth	2850-4410	3000	0.61 – 0.77
75	Ostler Central	>439	3000	0.14 – 1
78	Irishman Creek	>10000	15000	0.49 – 1
96	Kerepehi Central	450-900[1]	5000	0.09 – 0.16
97	Kerepehi South	1800-4800[1]	5000	0.30 – 0.62
138	Ngangiho	before 1922 AD[1]	698	0.11 – 1
142	Waiohau North	1800-4800[1]	843	0.88 – 1.00
143	Waiohau South	1800-4800[1]	533	0.97 – 1.00
134	Paeroa North	110-1800[1]	303	0.41 – 1.00
135	Paeroa Central	110-1800[1]	269	0.45 – 1.00
136	Paeroa South	110-1800[1]	322	0.41 – 1.00

Index	Fault name	Last rupture (yr B.P. unless otherwise stated)	Mean recurrence interval (yr)	F(elapsed time)
172	Whakatane	620+/-120 – 627+/-94[1]	3500	0.15 – 0.20
171	Waimana	150-1800[1]	3500	0.06 – 0.40
120	Ngakuru NE	110-1800[1]	1100	0.14 – 0.81
121	Ngakuru SW	110-1800[1]	983	0.15 – 0.85
141	Aratiatia	1800-4000[1]	300	0.93 – 1.00
105	Braemar	800-4800[1]	797	0.63 – 1.00
104	Matata	150-700[1]	374	0.41 – 0.87
106	Rotoiti	100-800[1]	130	0.68 – 1.00
107	Te Teko	1987 AD[1]	339	0.04
165	Mohaka South	350-1800[1], <600[2]	1000	0.33 – 0.48
166	Mohaka North	350-1800[1], <600[2]	1000	0.33 – 0.48
193	Poukawa North	7000-12000[1]	9500	0.52 – 0.72
194	Waipukurau-Poukawa	3000-75001[1]	5300	0.43 – 0.76
157	Wellington	1350-1450 AD[2]	600	0.60 – 0.66
197	Rangiora	100-1800[2]	962	0.14 – 0.85

[1] From GNS active fault database
[2] From Mouslopoulou *et al.* (2001)
[3] From Pettinga *et al.* (2001)

For nearly all faults in Table 2 there is a range of F(time elapsed) values, and in many cases the range is rather wide. Out of the 305 active faults in the NSHM database, we have information on prehistoric ruptures for only 49 faults. It would perhaps be unwise to assume that these are a random or representative sample. Nevertheless, the histogram of expected numbers has been assembled in Figure 6. Wider class intervals have been used than before because of the need for the expected number in each class interval to be no less than about 10.

It can be seen from the tolerance limits in Figure 6 that there is no deviation of the data from the model according to the binomial test. Nevertheless, visually there appears to be some hint of an increasing trend in the histogram. The apparent trend is consistent with a tendency to underestimate the mean recurrence intervals of active faults. The chi-squared and Kolmogorov-Smirnov tests show that, in this case, there is no statistical support for the existence of such a trend. The χ^2 statistic is 4.84 on 4 degrees of freedom, much less than the critical value of 9.5 at the 0.05 level. And the λ statistic is 1.04, much less than the critical value 0f 1.36 at the 0.05 level.

Hence the data, such as they are, on the time elapsed since prehistoric ruptures of faults in the NSHM database, are consistent with the estimated mean recurrence intervals for these faults. However, with such a small and poorly constrained data set, the statistical tests are not very powerful for detecting deviations from the theoretical distribution. It is to be hoped that a

larger sample will become available as a result of future paleoseismic investigations. It should be noted from Table 2 that sometimes $F(t)$ can be quite well constrained even when the date of last rupture is not. This is evident, for example, in the case of the Lowry and Aratiatia faults.

4.0   SEISMIC MOMENT RATE

4.1   Comparing the seismic moment rate of the seismicity model (strain release) with that expected from geodetic data (strain accumulation)

Seismic moment rates are a convenient measure of earthquake potential that can be derived from geological and seismological data sets. Seismic moments may also be estimated for individual earthquakes from geodetic data. For this report we concentrate on the estimation of moment accumulation rates from geodetic data in the intervals between major earthquakes. We wish to compare the pattern and magnitude of moment accumulation estimated from geodetic data with the pattern and magnitude of moment release estimated from geological and seismological data. Our eventual aim is to be able to develop the geodetically-measured strain accumulation as a useful forecasting tool for strain release in earthquakes.

For geological data, the seismic moment rate $\dot{M}_0$ can be deduced from the length L, width W and slip rate $\dot{s}$ of the fault source, as follows:

$$\dot{M}_0 = \mu L W \dot{s} \tag{19}$$

where μ is the average rigidity in the seismogenic layer. For seismological data, the moment magnitude M_w of an earthquake is of course closely related to the seismic moment M_0 by

$$M_w = \frac{2}{3}\log M_0 - 6.03 \tag{20}$$

where M_0 is expressed in Newton metres (Hanks and Kanamori, 1979). Where the moment magnitude has not been measured, empirical relations (e.g. Dowrick and Rhoades, 1998) can be used to convert other magnitude measures, such as local magnitude M_L and surface-wave magnitude M_s to an approximate value of M_w.

Geodetic observations can give estimates of the surface strain accumulation rate over wide areas. If these areas are subdivided into smaller areas corresponding, say, to seismotectonic zones, then the moment accumulation rate can be calculated by first deriving the principal axes of the strain rate tensor for each zone, and then using the method of Savage and Simpson (1997) to calculate the moment accumulation rate in each zone:

$$\dot{M}_0 = 2\mu HA \max(|\varepsilon_1|, |\varepsilon_2|, |\varepsilon_1 + \varepsilon_2|) \tag{21}$$

where μ is the rigidity, A is the area of the zone, H is the depth of the seismogenic zone, and ε_1 and ε_2 are the principal surface strain rates acting over the area A. One assumption involved here is that the average strain rate measured at the surface is representative of the average strain rate over the seismogenic volume beneath.

The seismic moment rates predicted within any zone by a seismic hazard model can be obtained by summing the rates contributed by all sources in the model. Thus we can compare the model seismic moment rate with the minimum moment accumulation rate indicated by geodetic observations within each zone separately, and over the wider area. Because some aseismic deformation may occur, it is not expected that these two rates should be equal, but rather that the seismic moment rate should not generally exceed the geodetic moment accumulation rate (provided that the seismic moment rate is not biased by recent large earthquakes).

4.2   Application to the national seismic hazard model

New Zealand sits astride the boundary of the Australian and Pacific plates. The relative motion across the plate boundary, as revealed by geodetic studies, is c. 50mm/yr in the northeast, and c. 30mm/yr in the Southwest. The sum total of slip achieved by the occurrence of earthquakes in the NSHM over time should be proportional to the rate of plate motion, but there is no constraint in the NSHM to ensure that it is.

Here we compare seismic moment rates calculated for the 14 crustal seismotectonic zones defined in the NSHM (Stirling *et al*., 2000; 2002) against moment accumulation rates calculated for these zones from geodetic data. The NSHM uses geological and seismological data as input to calculating the location, magnitude and frequency of occurrence of earthquakes across New Zealand. Geodetic data are not used as input to the NSHM because of the presently-unresolved difficulties encountered in using these data for probabilistic seismic hazard assessment (PSHA) in the past (e.g. Working Group of California Earthquake Probabilities, 1995). The c. 10 year long geodetic strain record in New Zealand therefore represents a data set independent of the NSHM that can be used to evaluate the performance of the NSHM for predicting contemporary moment rate accumulation in New Zealand.

Geodetically derived surface velocities at points throughout New Zealand are first combined into a smooth velocity field throughout New Zealand (Beavan and Haines, 2001). Moment accumulation rates are calculated from the geodetic velocities by first deriving the principal axes of the strain rate tensor for each of the 14 crustal NSHM seismotectonic zones, and then using the method of Savage and Simpson (1997) to calculate the moment accumulation rates. Two sets of moment rates are calculated, according to whether or not the Pacific-Australia

relative plate motion rate is imposed as a boundary condition when the individual velocities are combined in to the smoothed velocity field. In general, the estimated moment accumulation rates are larger in the case where the relative plate motion rates are applied at the boundary. This is especially true for regions that are partially offshore, where the surface velocities are controlled more by the boundary conditions than by observed geodetic data.

For the NSHM, we simply sum up the seismic moment rate of all earthquake sources inside each crustal seismotectonic zone. The moment rates are either calculated for an earthquake source by multiplying the seismic moment of the earthquake magnitudes predicted for the source by the rate of occurrence of those earthquakes, or from the length, width and slip rate of the fault source. The latter method is used for the cases where faults cross boundaries of seismotectonic zones.

We show the results of our comparisons in Figures 7 and 8. Figure 7 shows the ratio of the seismic moment rates predicted from the NSHM to the geodetic moment accumulation rates for each crustal seismotectonic zone (numbered 1 to 14 from left to right on the *x*-axis in Figure 7). Figure 8 is a map of the seismotectonic zones colour-coded according to ratio. The ratio is at about 0.6 - 0.8 averaged over the whole of New Zealand. This means that on average the geodetic data predict slightly higher moment accumulation rates than the seismic moment release rates from the combined geologic and seismological datasheet of the NSHM. The largest differences are in the zones located away from the major plate boundary faults and seismicity of New Zealand (i.e. zones 1, 5, 7, 8, 9, 12, and 13). This is most probably because moment accumulation must occur over a wider area than moment release, particularly when large faults are involved in the strain release mechanism. (In an extreme case where all the moment release occurs on one fault, the moment release is localised on this one fault. By contrast, the strain accumulation leading up to the earthquake has occurred over a region more than three times the locking depth of the fault.)

The discrepancy between the geodetic moment rate and the moment rate of the NSHM may also arise because we have not considered the entire NSHM source model in this study. The subduction interface and slab sources have not been included in the summation of total moment rates for the country, because it is not clear how much of the strain accumulation in these sources would be reflected in strain measurements at the surface. It is to be expected that a component of the contemporary strain measured across New Zealand will be released on the subduction zones.

With the qualifications mentioned above, the agreement between the geodetic and NSHM-derived seismic moment rates across the country as a whole is reasonably good. Future research will be focused on understanding the observed discrepancies, through critical examination of the data used to define the moment rates in the NSHM. It is expected that improved understanding will be obtained by eventually incorporating the subduction zone and slab sources into the comparison, and perhaps using the geodetic data to modify the NSHM

seismotectonic zones to be more consistent with the spatial distribution of geodetic strain across New Zealand.

5.0   MAGNITUDE-FREQUENCY DISTRIBUTION

5.1   Comparing the magnitude and frequency distribution of the seismicity model with the historical record and the plate motion

5.1.1   Universality of the earthquake magnitude-frequency distribution

The Gutenberg-Richter magnitude-frequency relation (Gutenberg and Richter, 1944)

$$\log N = a - bM \ , \tag{21}$$

where *N* is the number of earthquakes exceeding any magnitude *M*, and *a* and *b* are parameters, is almost universally (e.g. Kagan, 1999) observed to hold for populations of earthquakes in large areas and over long time-periods. Where adequate data are available, they usually conform well to the relation (21) over several orders of magnitude, with the number of earthquakes tapering off at high magnitudes, as is necessary for the rate of seismic energy release to be finite. From an analysis of world-wide catalogues of seismic moments using the Flinn-Engdahl regionalisation of global seismicity, Kagan (1999) found that the b-value of 1 is universally observed, and that the maximum moment magnitude does not vary between regions for shallow earthquakes at or near continental boundaries. He also found that although the fluctuations in earthquake size distribution increase with depth, the b-values for earthquakes in the depth range 70-500 km do not exhibit statistically significant regional variations.

Seismic hazard models are usually constructed so that the distributed seismicity within a seismotectonic zone conforms to the Gutenberg-Richter relation up to some upper magnitude limit within, with b-value estimated from the earthquake catalogue of the zone. The magnitudes associated with fault sources are determined mostly from the physical dimensions of the faults. The fault sources contribute most to the high end of the magnitude range. There is thus no guarantee that the combined magnitude-frequency distribution will be consistent with that in the earthquake catalogue.

It should be noted that Kagan did not examine the seismicity within small seismotectonic zones such as are used in the construction of many seismic hazard models. However, if the result of applying the PSHA methodology to small zones and faults within a region is a magnitude-frequency distribution that, on the larger scale, is not consistent with the universal relation, it is matter of concern. Any such inconsistency should call into question how well founded are the assumptions and data that went into the model.

5.1.2   Distribution of the largest magnitudes

Consider a set of earthquakes exceeding some magnitude threshold and conforming to the Gutenberg-Richter relation (1). We shall call such a set a Gutenberg-Richter set. The magnitudes in a Gutenberg-Richter set follow the exponential distribution. Now consider the relative magnitudes of the largest few earthquakes in the set. Let D_n denote the difference in magnitude between the largest and the $(n+1)^{th}$ largest earthquakes. It is well known (e.g. Vere-Jones, 1969) that the distribution of D_n is independent of the number of earthquakes in the set. Figure 9 shows the expected value and 95% tolerance limits for D_n for $n=1,\ldots,10$.

As indicated above, populations of earthquakes in large areas and over long periods of time are close to being Gutenberg-Richter sets. In small areas and over short periods of time, the earthquakes often differ from a Gutenberg-Richter set. For example, a typical mainshock-aftershock sequence has one earthquake of outstanding magnitude, with $D_1 \approx 1.2$ as indicated by Båth's law (Richter, 1958; Båth, 1965). The same tends to be for regions and time period in which the seismicity is dominated by one major mainshock-aftershock sequence. Any value of D_1 larger than 1.2 represents a significant deviation from a Gutenberg-Richter set. Over large areas and very long times, the earthquakes again tend to differ from a Gutenberg-Richter set in the smallness of the first few values of D_n, because of the effect of the tail-off in the frequencies for earthquakes at the high end of the range of magnitudes.

Synthetic catalogues generated according to seismic hazard models should have similar properties to catalogues of actual earthquakes, including the distributions of the largest earthquakes over similar time-periods.

5.1.3   Plate motion-balanced magnitude-frequency distribution

In the long run the magnitude-frequency distribution of earthquakes in a seismic hazard model should be consistent both with plate motion and with the magnitude-frequency relations displayed by large populations of earthquakes. Wesnousky *et al*. (1983) showed how to calculate such an ideal distribution for a plate boundary region. In their "B value" method they assumed that earthquakes had the same distribution as a Gutenberg-Richter set up to a maximum magnitude M_{max}. It is useful to compare the overall magnitude-frequency distribution of earthquakes generated by a seismic hazard model with the ideal distribution. Again, any deviations need to be investigated to see if they might be the result of unrealistic assumptions or poor data quality.

5.2   Application to the national seismic hazard model

5.2.1   Magnitude-frequency distribution of synthetic and actual catalogues

The magnitude-frequency distribution of earthquakes in a 10,000-year synthetic catalogue conforming to the NSHM is displayed in Figure 10, with cumulative frequencies in Figure 10(a) and discrete frequencies in Figure 10(b). Separate plots are shown for all earthquakes in the catalogue, and for the distributed earthquakes, i.e. those that do not rupture modelled faults in the NSHM fault database. Also shown in Figure 10(a) are maximum likelihood fits of the Gutenberg-Richter relation (Aki, 1965) to each plot. The *b*-values calculated for *M*>5.25 are 0.93 and 1.16 for all earthquakes and distributed earthquakes, respectively. It can be seen from Figure 10(a) that there is a surplus of earthquakes at around *M*>7 in the NSHM relative to the fitted relation, and then a deficit beginning at M>7.5 becomes more pronounced for higher magnitudes. Figure 10(b) shows that between M7 and 7.3 there are about twice as many earthquakes under the NSHM as expected under the fitted relation, and that these are nearly all fault-rupturing events. For such a surplus to occur in a particular magnitude range over an area as large as New Zealand and a period as long as 10,000 years is a significant deviation from the universal scaling relation.

Figure 11 compares the cumulative rate of earthquake occurrence of the NSHM with that indicated by instrumental earthquake data. The Harvard world-wide catalogue of earthquake centroid moment tensor solutions (Dziewonski *et al*.) provides estimates of M_w for earthquakes of magnitude about 5 and greater over the period 1979-2002. For our analysis we consider just those events in the New Zealand region between 36° and 48° S. The b-value estimated for M>5.35 is 0.91±0.10. This is consistent with the b-value of 0.93 for all earthquakes in the NSHM. For M>6.45, a composite catalogue which is considered to be complete from 1901-2002, has been obtained by using the values for M_w listed by Dowrick and Rhoades (1998) for earthquakes prior to 1979 and the Harvard catalogue values from 1979 on. Some of the Dowrick and Rhoades M_w values are estimated from the measured M_s. It can be seen from Figure 11 that for M>5.35 the rate of earthquake occurrence is about 50% greater in the NSHM than the Harvard catalogue. This may be due to incompleteness of the Harvard catalogue at lower magnitudes, or to differences between the M_w and M_L magnitude scales; the distributed seismicity for the NSHM was primarily based on instrumental M_L values. However further research, outside the scope of this study, is necessary to clarify the reason for this discrepancy. The rates are similar for all three catalogues at M>6.3, but at around M>7 the NSHM rate is about twice the rate of the Harvard and 1901-2002 catalogues.

In Figure 12, the events rupturing modelled faults have been removed from all three catalogues, and the rates recomputed. This result is in strong contrast to Figure 11. Now the NSHM rate is generally below the other two catalogues from about *M*>6.25 on. For *M*>7, the NSHM rate is only about one-tenth of the rate for the Harvard and 1901-2002 catalogues. So the NSHM underestimates the number of large "non-fault-rupturing" earthquakes by a wide

margin and, by an even wider margin, overestimates the number of "fault-rupturing" events.

Figure 13 further elucidates the information in the previous figures. It shows, as a function of magnitude, the proportion of earthquakes that rupture modelled faults in the NSHM synthetic catalogue compared with the Harvard and 1901-2002 catalogues. Only one earthquake in the Harvard catalogue ruptured a modelled fault (the Edgecumbe earthquake of 1987), but six earthquakes in the 1901-2002 catalogue did (see Table 1 above). Under the NSHM, the proportion of earthquakes rupturing modelled faults is close to one for all magnitudes above 7.0. Except at magnitudes 7.8 and 7.9, the corresponding proportions in the Harvard and 1901-2002 catalogues are invariably closer to zero than one, if not actually zero.

The discrepancies evident from the analyses in this section indicate that some revision of the NSHM may be necessary. The discrepancies appear to have arisen from the following elements of PSHA procedure:

1. Too low a value of magnitude cut-off for distributed earthquake in certain zones;
2. Underestimation of the mean recurrence intervals for rupture of some faults;
3. Not allowing for variability in the magnitude of earthquakes on an individual fault.

The magnitude cut-off for distributed earthquakes is as low as 7.0 for quite a lot of the zones in the NSHM. Twentieth century catalogues clearly show that many of the earthquakes with M>7.0 have not ruptured faults in the NSHM database. As well as raising the threshold for certain zones, there could be merit in having a tail-off of the Gutenberg-Richter relation for distributed earthquakes rather than a hard cut-off.

Underestimation of the mean recurrence intervals for some faults has already been discussed in section 2 above, where it was seen that there the problem is most acute in the Taupo Volcanic Zone, but not entirely restricted to that zone.

The wide scatter of points about the Gutenberg-Richter trend at M>7 in Figure 10(b) could be reduced by allowing for variability in the size of a characteristic earthquake on individual faults. To allow for such variation would be consistent with evidence of variability in the amount of slip in successive earthquakes on faults (e.g. Stein *et al*. 1997). Moreover Jackson (2001) has recently reported that the size of earthquakes that have occurred on pre-mapped faults in California have varied substantially from the magnitudes previously estimated from fault dimensions. Actually there is both uncertainty and expected variability in the size of earthquakes that will rupture a given fault. Both of these are ignored when a single value of earthquake magnitude is chosen for each fault, as in the NSHM.

5.2.2   Distribution of the largest magnitudes

In Figure 14, the value of D_n, the difference in magnitude between the largest and $(n+1)$th

largest earthquakes has been plotted as a function of *n*, for the NSHM and the Harvard and 1901-2002 catalogues. In the case of the NSHM, 100 separate synthetic catalogues, each with a 101-year duration, were generated. This was to make the NSHM results comparable with the 1901-2002 catalogue, because for catalogues with an upper magnitude limit the distribution of D_n can be expected to change with the length of the catalogue, as already noted above.

The distribution of D_n across the 100 catalogues has been represented in Figure 14 by the mean, plotted as a point, and error bars extending to the 5th and 95th percentiles. It can be seen that, for each value n, the D_n values of the Harvard and 1901-2002 catalogue lies within the error bars. Hence the distribution of D_n values in the NSHM is consistent with the 1901-2002 catalogue. In this respect the NSHM is quite satisfactory.

On the other hand, for *n*>3 the error bars do not intersect the curve for the mean D_n of a Gutenberg-Richter set. Hence D_n for 3<*n*<20 is significantly lower under the NSHM than would be expected in a catalogue conforming completely to the Gutenberg-Richter relation. Such a result is to be expected for a catalogue covering a wide area and long duration, when there is an upper limit to the range of possible magnitudes. It is therefore not surprising that D_n values for all three catalogues in Figure 14 lie consistently below the expected value of a Gutenberg-Richter set.

5.2.3 Comparison of NSHM with plate motion-balanced magnitude-frequency distribution

Here the magnitude-frequency distribution for all of the earthquake sources of the NSHM is compared with one that is "plate motion balanced". The latter distribution is calculated according to the length (1950 km), relative plate motion rate (average 40 mm/yr), and maximum magnitude (M8.3, 0.1 units larger than the largest historical earthquake in New Zealand) of the New Zealand plate boundary, and according to the global average b-value of 1. The model used to make these calculations is the "Bvalue Model" of Wesnousky *et al.* (1983).

The comparison of the two magnitude-frequency distributions is illustrated in Figure 15. Here the rates are shown as the number of earthquakes per century greater than or equal to a magnitude M. There is generally good agreement between the two distributions, except in the magnitude range about M6.7 - 7.4, where the NSHM rates are greater, by a factor of 1.4 to 1.8, than the plate motion-balanced rates. Since these magnitudes and frequencies mainly come from the fault sources of the NSHM, it seems likely that the parameters for some of these fault sources may be in error.

A possible reason for the discrepancies is that the magnitudes for some of the fault sources are too low, due to underestimation of fault length and use of the Wells and Coppersmith

earthquake scaling relations to derive magnitude from fault parameters (Stirling *et al*. 2002). Work is in progress at GNS to incorporate new scaling relations into the NSHM, and by this and other means to modify fault sources in the model. This work may result in reduced discrepancies between the NSHM and the plate motion-balanced model. It is worth noting that similar discrepancies have been reported in recent years for the Californian component of the US national seismic hazard model, and that these have been reduced through modification of the fault source data set and by the use of newly developed scaling relations in the model.

6.0   CONCLUSION

A number of tests applicable to seismic hazard models have been described and applied to the NSHM. They have in some cases revealed possible deficiencies in the data and assumptions of the NSHM, and areas for future research.

A comparison of the frequency of surface fault rupture in the NSHM with the historical record has revealed a statistically significant discrepancy. The NSHM predicts a much higher frequency of fault rupture in New Zealand earthquakes than has occurred historically. The discrepancy is most marked in the Taupo Volcanic Zone, but is also appreciable in northern Canterbury and southern Fiordland. It is likely to be due to multiple causes, including underestimation of the recurrence intervals of fault rupture for some faults in the NSHM, a possible tendency for fault ruptures to occur in clusters, and probable incompleteness in the historical record of fault ruptures.

A test of the consistency of the distribution of estimated times of last rupture of surface faults with the distribution of mean recurrence intervals in the NSHM has not revealed any significant discrepancies. The statistical power of the test is weak because of the small proportion of faults in the database for which any information on the time of last rupture is available. Future paleoseismic studies should aim to increase the available information constraining the times of last rupture of faults.

A comparison of the seismic moment rate of the NSHM with that expected from geodetic data has shown that overall the strain released under the model accounts for 0.6-0.8 of the geodetic moment rate. Within the individual seismotectonic zones of the NSHM, the agreement is not so good, with the ratio of NSHM moment to geodetic moment varying between 0.05 and 2. Further research is needed to understand these discrepancies. The present comparison does not include the contribution of subduction interface and slab sources to the moment release under the NSHM.

Comparisons of the magnitude-frequency distribution of the NSHM with the historical record and plate motion balanced distribution have revealed discrepancies between the model and the data, and between the model and the universal Gutenberg-Richter magnitude-frequency relation and its extensions. While the magnitude-frequency relation in the NSHM shows wide

scatter above M7.0, there is a surplus of earthquakes in the range M7.0 - M7.3 in the NSHM when compared with either historical data or the Gutenberg-Richter relation. Within the same magnitude range there is a deficit of distributed earthquakes and a more-than-compensating surplus of fault-rupturing earthquakes under the NSHM. Modifications to the NSHM are needed to rectify these discrepancies. In particular the characteristic magnitude and/or mean recurrence interval probably needs to be adjusted for some faults, and the upper magnitude limit of distributed earthquakes increased for some seismotectonic zones and a roll-off rather than a sharp cut-off adopted. Also variability in the magnitude of earthquakes rupturing a given fault source should be allowed for in the model.

7.0   ACKNOWLEDGEMENTS

8.0   REFERENCES

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

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

Beavan, R.J. and Haines, J. (2001) Contemporary horizontal velocity and strain rate fields of the Pacific-Australian plate boundary zone through New Zealand *Journal of Geophysical Research, 106(B1)*, 741-770.

Cornell, C.A. (1968). Engineering seismic risk analysis. *Bull. Seismol. Soc. Am., 58*, 1583-1606.

Dowrick, D.J. and Rhoades, D.A. (1998). Magnitudes of New Zealand earthquakes, 1901-1993. *Bull. NZ Nat. Soc. Earthqu. Eng. 31(4)*, 260-280.

Dowrick, D.J. and Cousins W.J. (2002). Historical incidence of Modified Mercalli intensity in New Zealand and comparisons with hazard models. *Bull. NZ Soc. Earthqu. Eng.*, submitted.

Dziewonski A.M. *et al*. Centroid moment tensor solutions, *Physics of the Earth and Planetary Interiors*, various issues.

Fisz, M. (1963). *Probability Theory and Mathematical Statistics,* 3rd Edition. Wiley, New York, 677 pp.

Gutenberg, B and Richter, C.F. (1944). Frequency of earthquakes in California. *Bull. Seismol. Soc. Am. 34*, 185-188.

Hanks, T.C., and Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research B84*, 2348-2350.

Jackson, DD (2001) Does fault length limit earthquake size? Symposium Sa, *IAA/YAPS Joint Scientific Assembly, Hanoi, Vietnam,* Book of abstracts, Paper no. 2345.

Jackson, D.D. and Kagan, Y.Y. (1999). Testable earthquake forecasts for 1999. *Seismological Research Letters 70(4)*, 393-403.

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

Kendall, M. and Stuart, A. (1977). *The Advanced Theory of Statistics*, Vol. 1, 4th Edition. McMillan, New York, 472 pp.

Kulkarni, R.B., Sadigh, K. and Idriss, I.M. (1974). Probabilistic evaluation of seismic exposure, *Proc. 2nd U.S. Nat. Conf. on Earthq. Eng.,* Vol. I, 263-270.

McGinty, P. (2001). Preparation of the New Zealand earthquake catalogue for a probabilistic seismic hazard analysis. *Bull. NZ Soc. Earthqu. Eng. 34(1),* 60-67.

Mouslopoulou, V., Stirling, M. and Berryman, K. (2001) New Zealand's Contribution to the ILP Worldwide Earthquake Recurrence Database. Poster prepared for ILP meeting, Kaikoura, New Zealand, November 2001.

Pettinga, J.R., Yetton, M.D., Van Dissen, R.J. and Downes, G. (2001). Earthquake source identification and characterisation for the Canterbury region, South Island, New Zealand, *Bull. NZ Soc. Earthqu. Eng*, *34(4)*, 282-317.

Reasenberg, P.A. (1985). Second-order moment of California seismicity, 1969-1982. *J. Geophys. Res. 90*, 5479-5490.

Richter, C.F. (1958). *Elementary Seismology*, Freeman, San Francisco, California, 768 pp.

Savage, J.C. and Simpson, R.W. (1997). Surface strain accumulation and the seismic moment tensor. *Bull. Seismol. Soc. Am*. 87, 1345-1353.

Schwarz, D.P. and Coppersmith, K.J. (1984). Fault behaviour and characteristic earthquakes from the Wasatch and San Andreas faults. *J. Geophys. Res. 89,* 5681-5698.

Somerville, P.G., Irikura, K., Graves, R., Sawada, S., Wald, D., Abrahamson, N., Iwasaki, Y., Kagawa, T., Smith, N. and Kowada, A. (1999). Characterising earthquake slip models for the prediction of strong ground motion. *Seismol. Res. Lett. 70*, 59-80.

Stein, R.S., Barka, A.A. and Dieterich, J.H. (1997). Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Journal of Geophysical Research 99(B7)*, 13701-13712, 1994.

Stirling, M.W., McVerry, G.H., Berryman, K.R., McGinty, P., Villamor, P., Van Dissen, R.J., Dowrick, D., Cousins, J., Sutherland, R. (2000). Probabilistic seismic hazard assessment of New Zealand. New active fault data, seismicity data, attenuation relationships and methods. *Institute of Geological and Nuclear Sciences Client report 2000/53.*

Stirling, M.W., McVerry, G.H., and Berryman, K.R. (2002) .A new seismic hazard model for New Zealand. *Bull. Seismol. Soc. Am.,* in press.

Vere-Jones, D. (1969). A note on the statistical interpretation of Båth's law. *Bull. Seismol. Soc. Am., 59*, 1535-1541.

Vere-Jones, D. (1970) Stochastic models for earthquake occurrence (with discussion). *J. Roy.*

Stat. Soc. B 32, 1-62.

Wald, D.J. and Heaton, T.H. (1994). Spatial and temporal distribution of slip of the 1992 Landers, California earthquake. *Bull. Seismol. Soc. Am. 84*, 668-691.

Wells, D.L and Coppersmith, K.J. (1994). Analysis of empirical relationships among magnitude, rupture length, rupture area and surface displacement. *Bull. Seismol. Soc. Am. 84*, 974-1002.

Wesnousky, S.G, Scholz, C.H., Shimazaki, K., and Matsuda, T. (1983). Earthquake frequency distribution and the mechanics of faulting. *J. Geophys. Res. 88(B11),* 9331-9340.

Working Group on California Earthquake Probabilities (1995). Seismic hazards in southern California: probable earthquakes, 1994-2024. *Bull. Seismol. Soc. Am. 85*, 379-439.

9.0   FIGURES
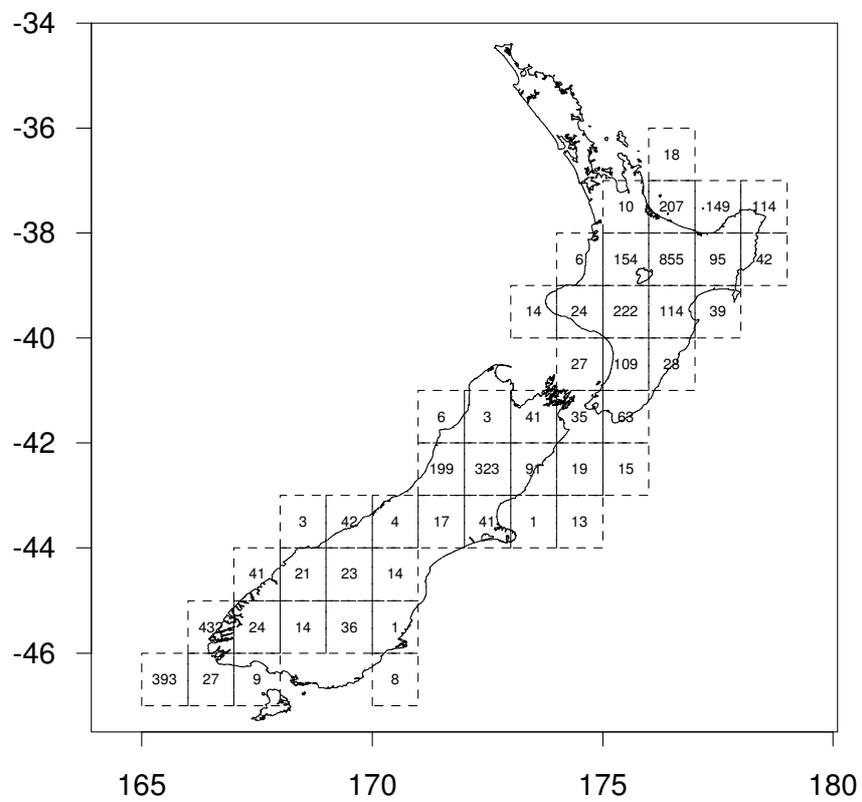
Figure 1   Number of earthquakes, in one-degree windows, rupturing modelled fault sources in a 10,000-year synthetic catalogue conforming to the NSHM.
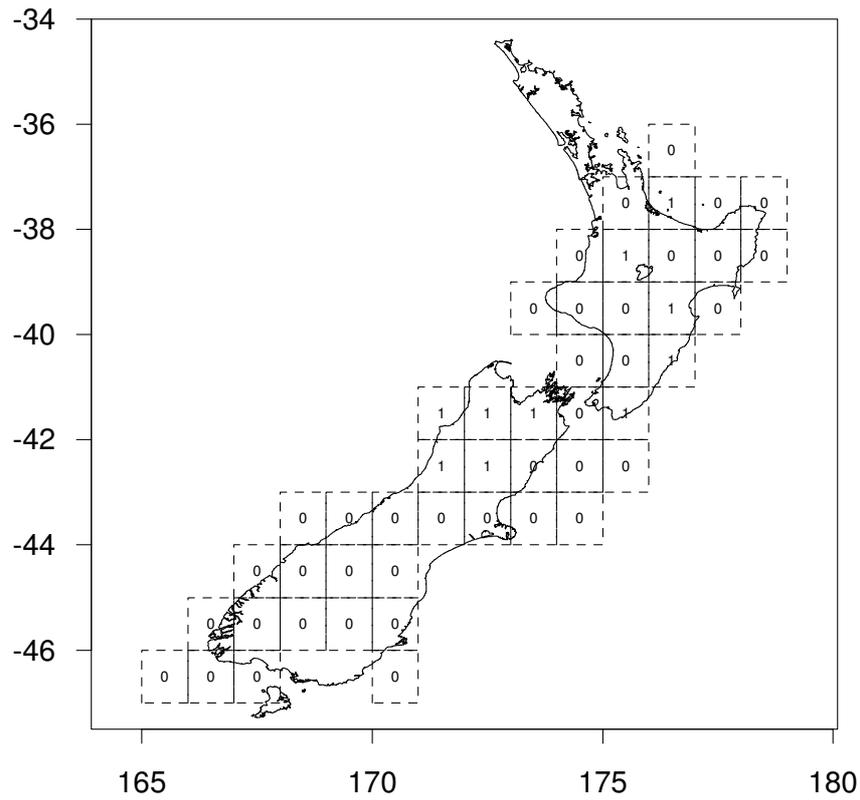
Figure 2   Number of earthquakes, in one-degree windows, considered to have ruptured modelled fault sources of the NSHM during the historical period 1840-2002.
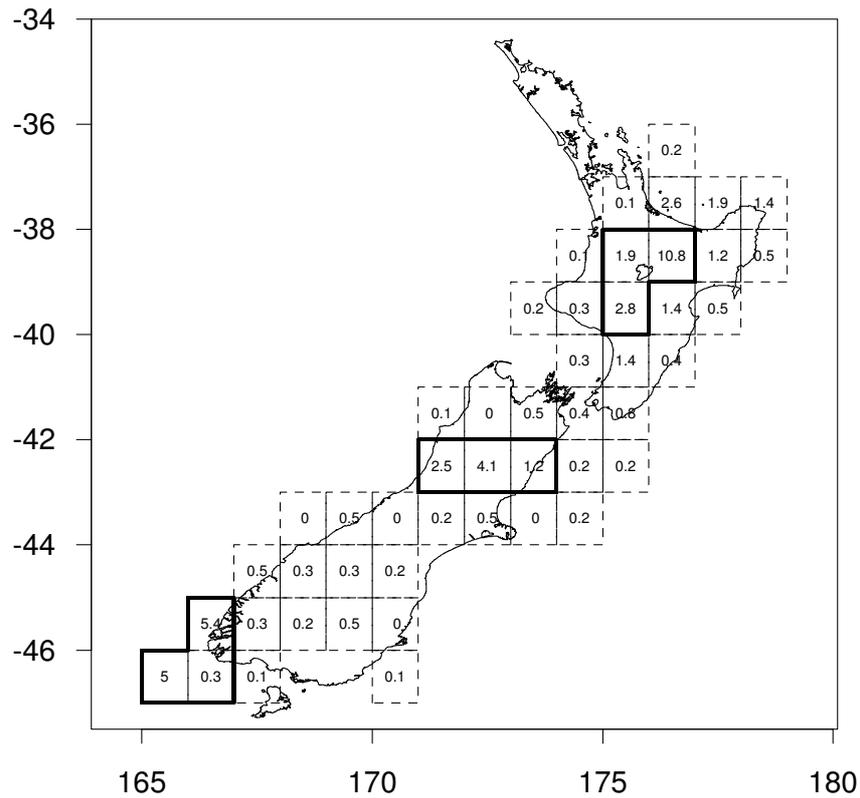
Figure 3   Number of earthquakes, in one-degree windows, expected to rupture modelled fault sources in any period of 162 years under the assumptions of the NSHM. The highlighted regions in the central North Island, northern Canterbury and southern Fiordland are discussed in the text.
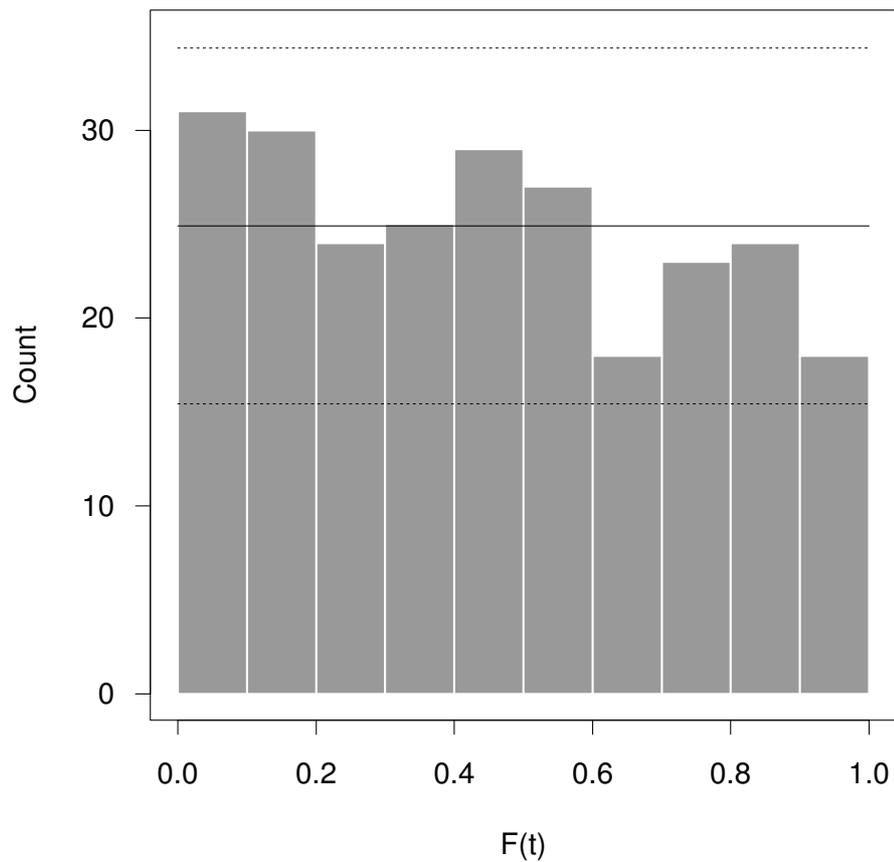
Figure 4   Histogram of *F*(elapsed time) for a 10,000-year synthetic catalogue conforming to the NSHM. Only faults that generated earthquakes in the catalogue are included. The solid line shows the expected number in each class interval and the dashed lines the 95% tolerance limits for the number in each class under the theoretical model.
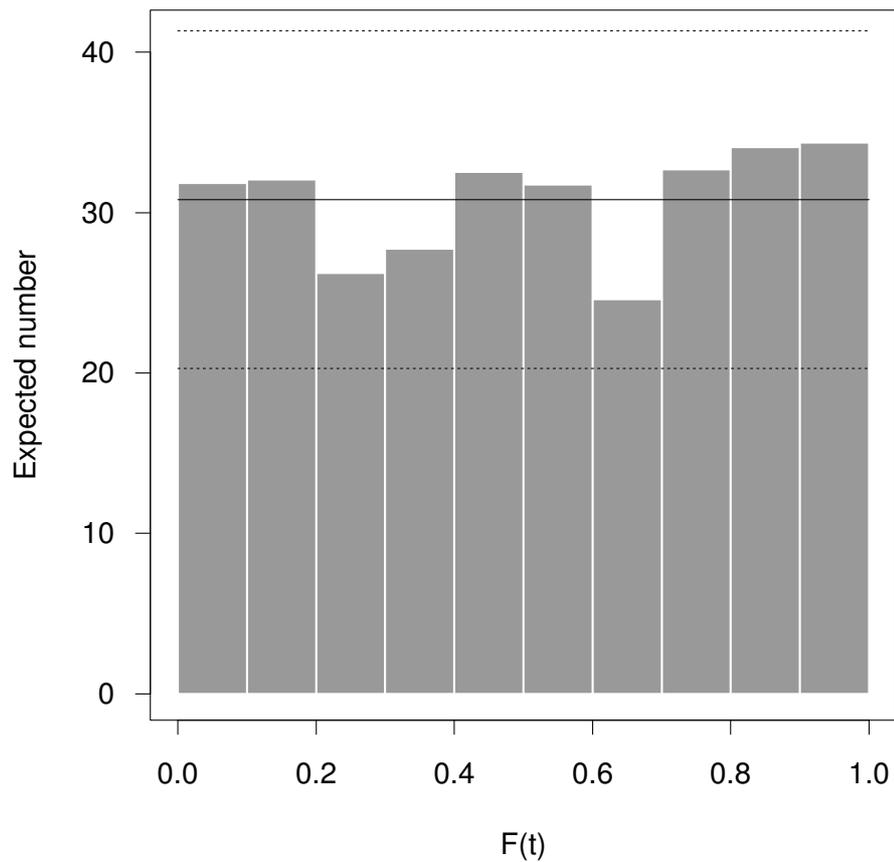
Figure 5   Histogram of *F*(elapsed time) for a 10,000-year synthetic catalogue conforming to the NSHM, including information from faults that did not generate earthquakes in the catalogue. The solid line shows the expected number in each class interval and the dashed lines the 95% tolerance limits for the number in each class under the theoretical model.
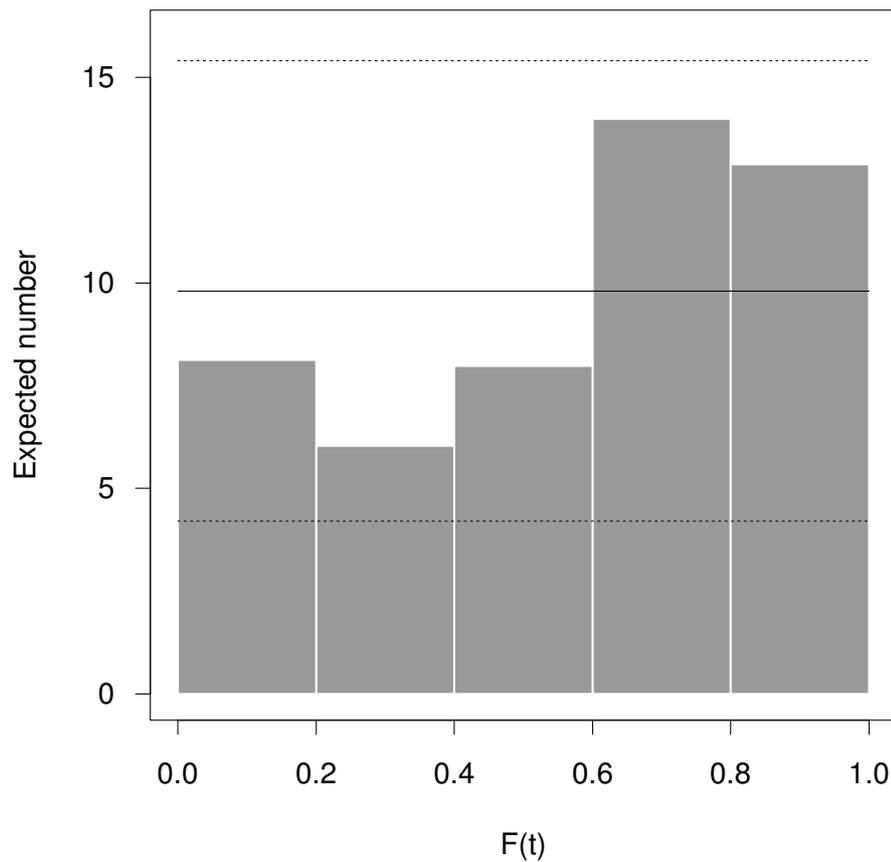
Figure 6   Histogram of *F*(elapsed time) constructed from information on pre-historic ruptures of modelled fault sources in the NSHM (Table 2). The solid line shows the expected number in each class interval and the dashed lines the 95% tolerance limits for the number in each class under the theoretical model.
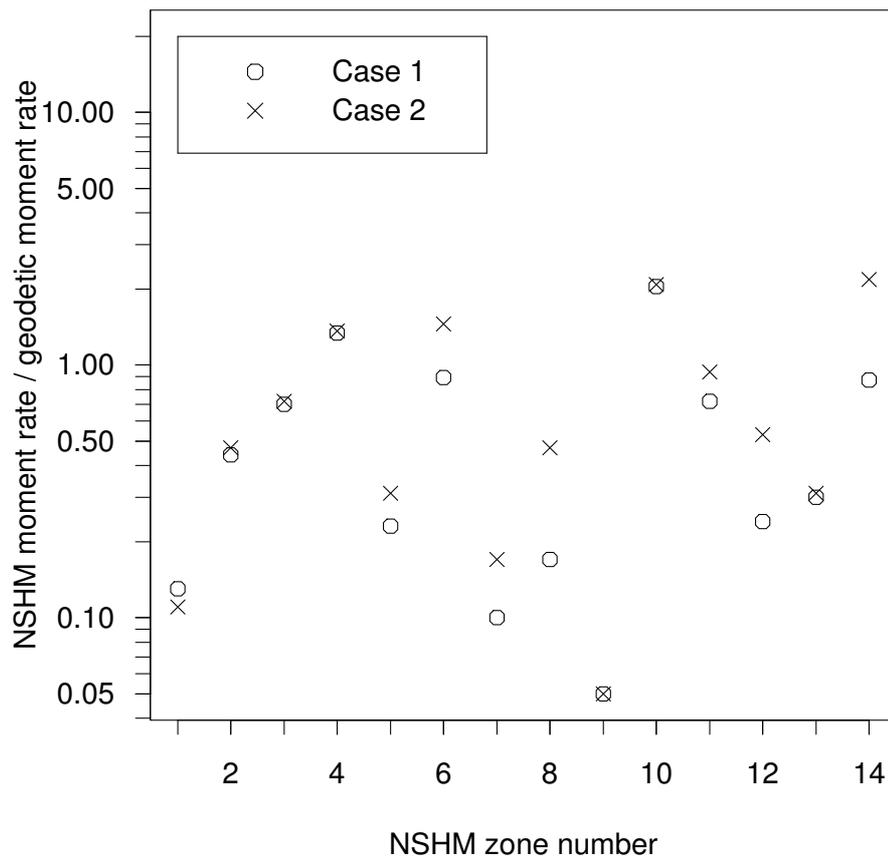
Figure 7 Ratio of the NSHM-derived seismic moment rates to the geodetically-derived seismic moment rates for the 14 crustal seismotectonic zones of the NSHM. In Case 1 and Case 2 the geodetic moment rates are calculated, respectively, with and without the Pacific-Australia relative plate motion rate imposed as a boundary condition.
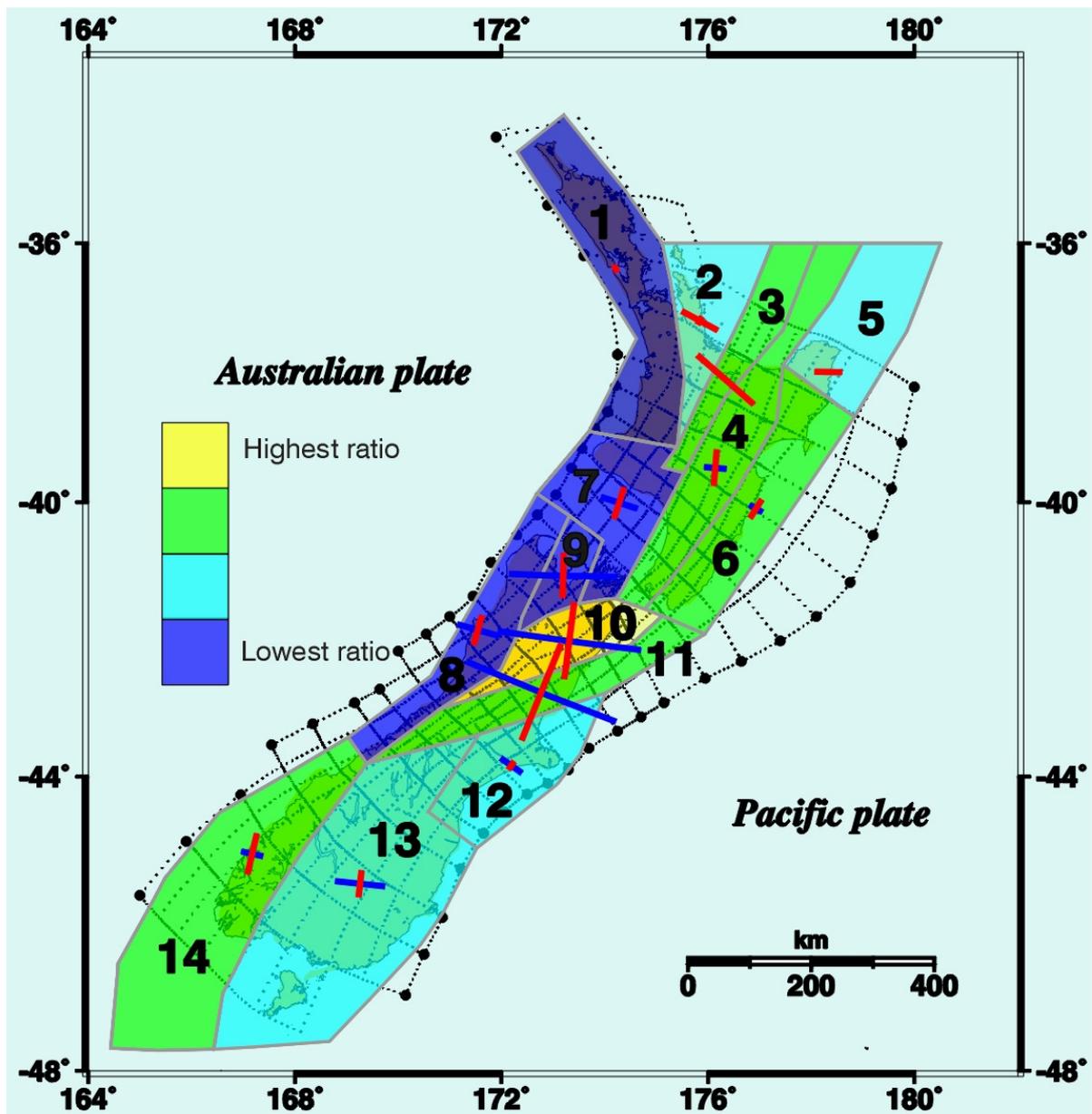
Figure 8   Map of New Zealand showing the 14 crustal seismotectonic zones colour coded according to ratio NZSHM moment rate/Geodetic moment rate. The ratio of 1 falls at about the boundary of the green and yellow, showing that the zones that enclose the major plate boundary faults and seismicity are in generally good agreement with the geodetic data.
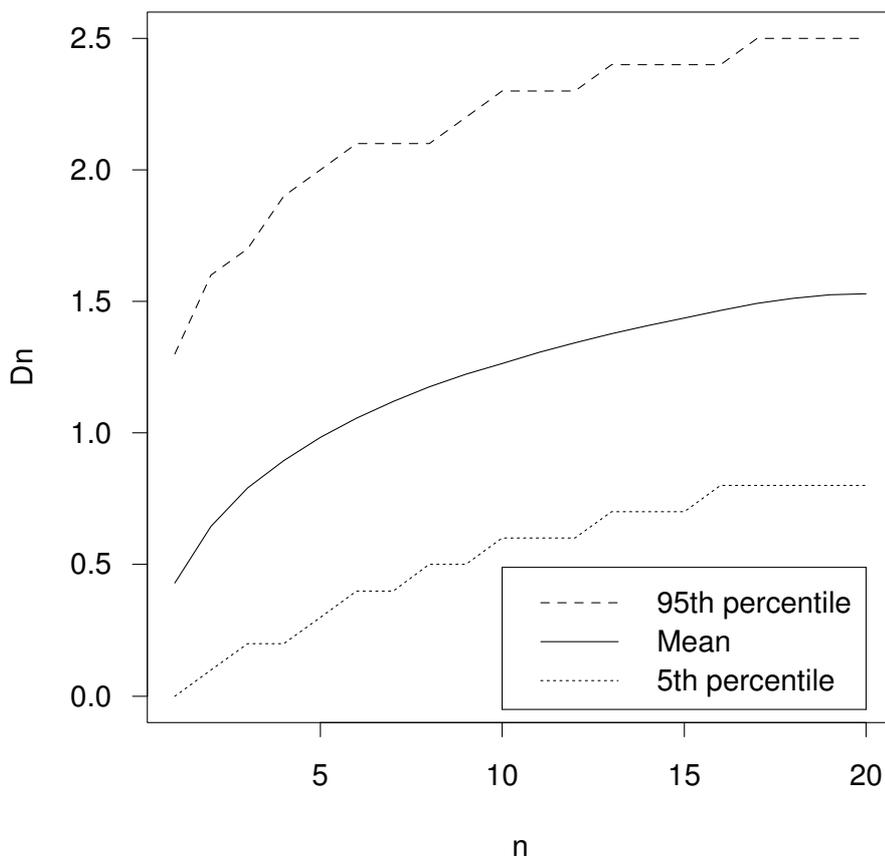
Figure 9   Distribution of D_n for in a Gutenberg-Richter set of earthquakes: expected value and 5th and 95th percentiles. Here D_n is the difference between the largest and $(n+1)$[th] largest magnitudes in the set.

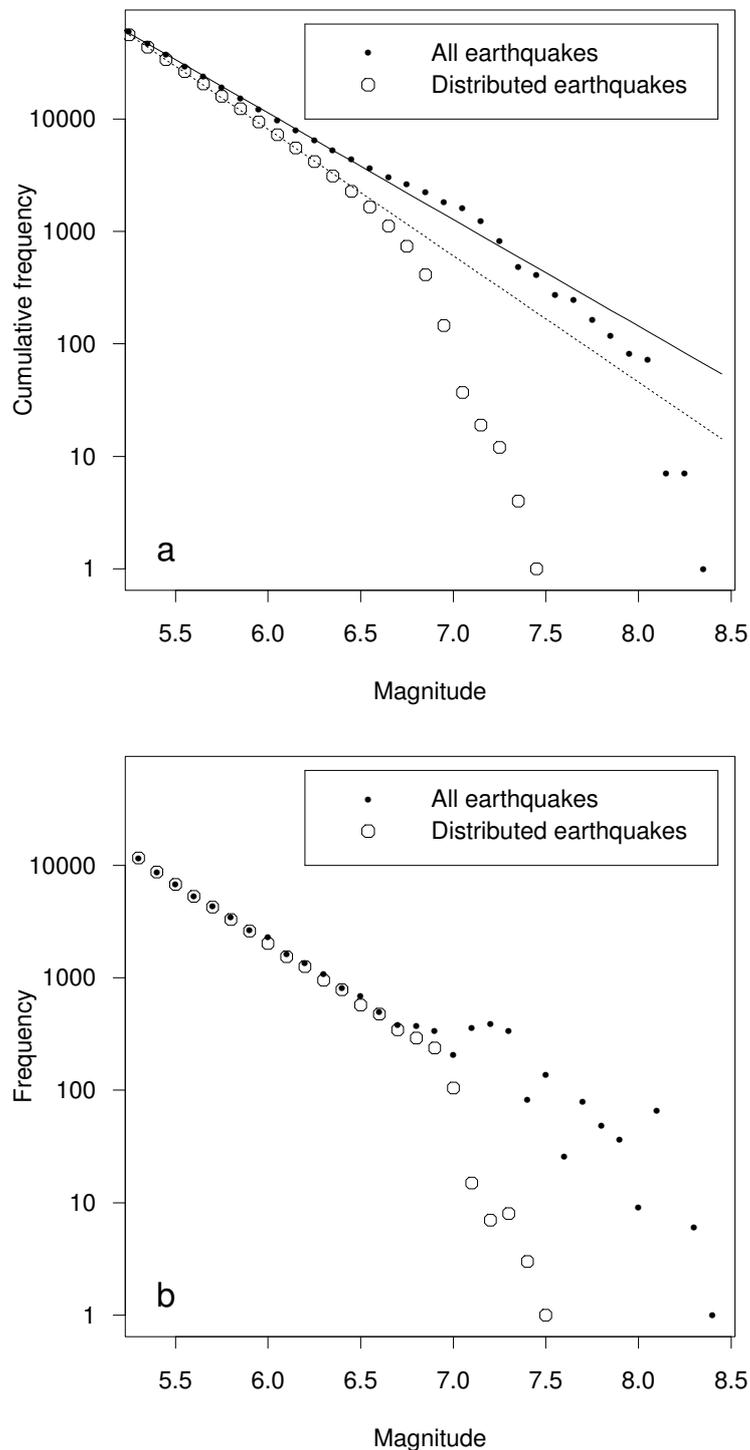## NSHM synthetic catalogue

a

b

Figure 10 (a) Cumulative frequency – magnitude plot and fitted Gutenberg-Richter relation and (b) frequency-magnitude plot, for a 10,000-year synthetic catalogue conforming to the NSHM. Plots are shown for all earthquakes in the catalogue and for just the distributed earthquakes (i.e. those not associated with a modelled fault source in the NSHM).
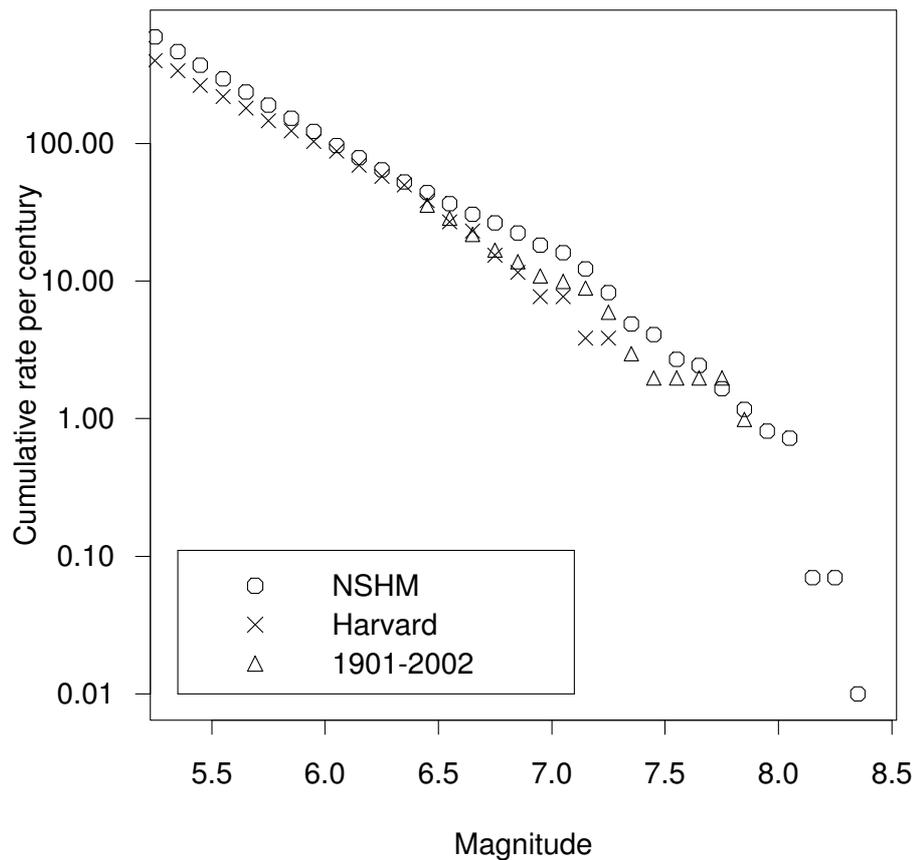
All earthquakes

Figure 11 Cumulative rate of earthquake occurrence as a function of magnitude for a 10,000 year synthetic catalogue conforming to the NSHM, the Harvard catalogue of M_w 1979-2002 for the New Zealand region between 36° and 48° S., and a composite catalogue of $M_w >$ 6.45 from 1901-2002 (see text for details). All earthquakes are included, i.e. both modelled fault-rupturing and distributed events.
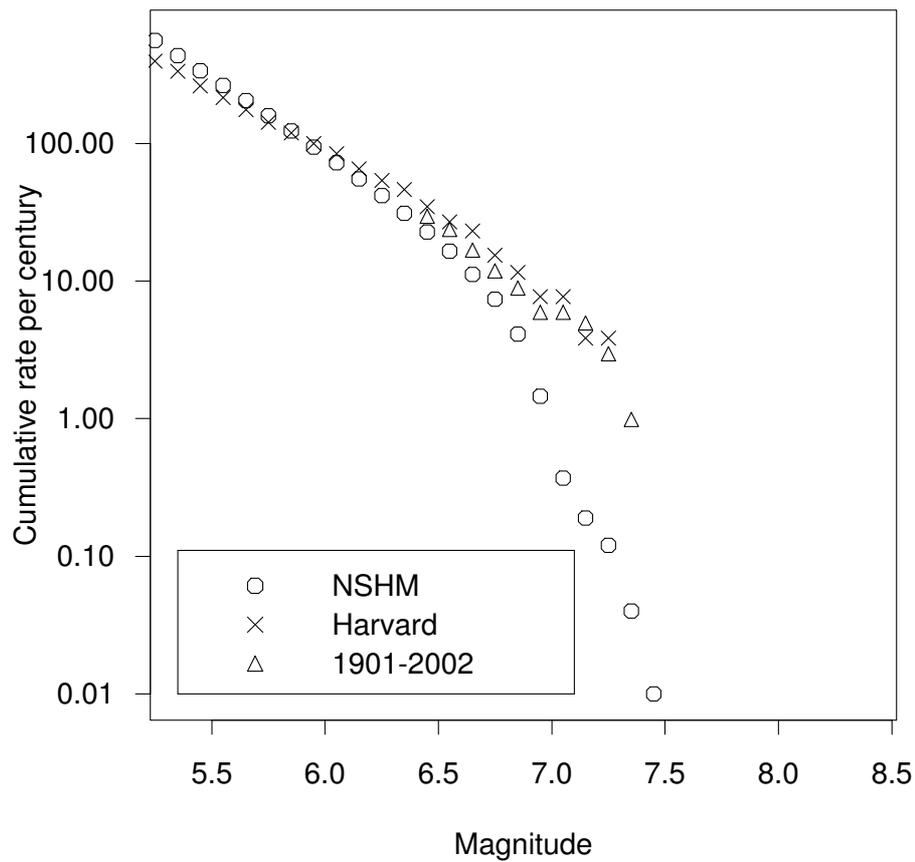
Non fault-rupturing earthquakes

Figure 12 Cumulative rate of earthquake occurrence as a function of magnitude for a 10,000 year synthetic catalogue conforming to the NSHM, the Harvard catalogue of M_w 1979-2002 for the New Zealand region between 36° and 48° S., and a composite catalogue of $M_w >$ 6.45 from 1901-2002 (see text for details). Only earthquakes that did not rupture a modelled fault in the NSHM database are included.
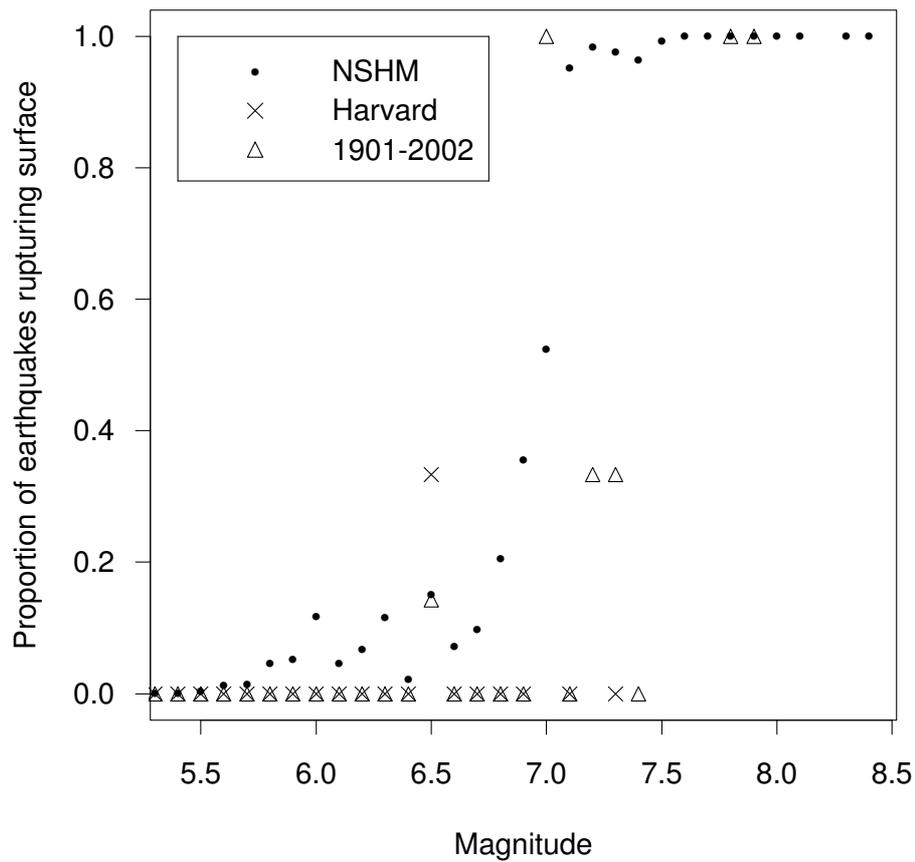
Figure 13   Proportion of earthquakes causing surface fault rupture as a function of magnitude in a 10,000 year synthetic catalogue conforming to the NSHM, the Harvard catalogue of M_w 1979-2002 for the New Zealand region between 36° and 48° S., and a composite catalogue of $M_w > 6.45$ from 1901-2002 (see text for details).
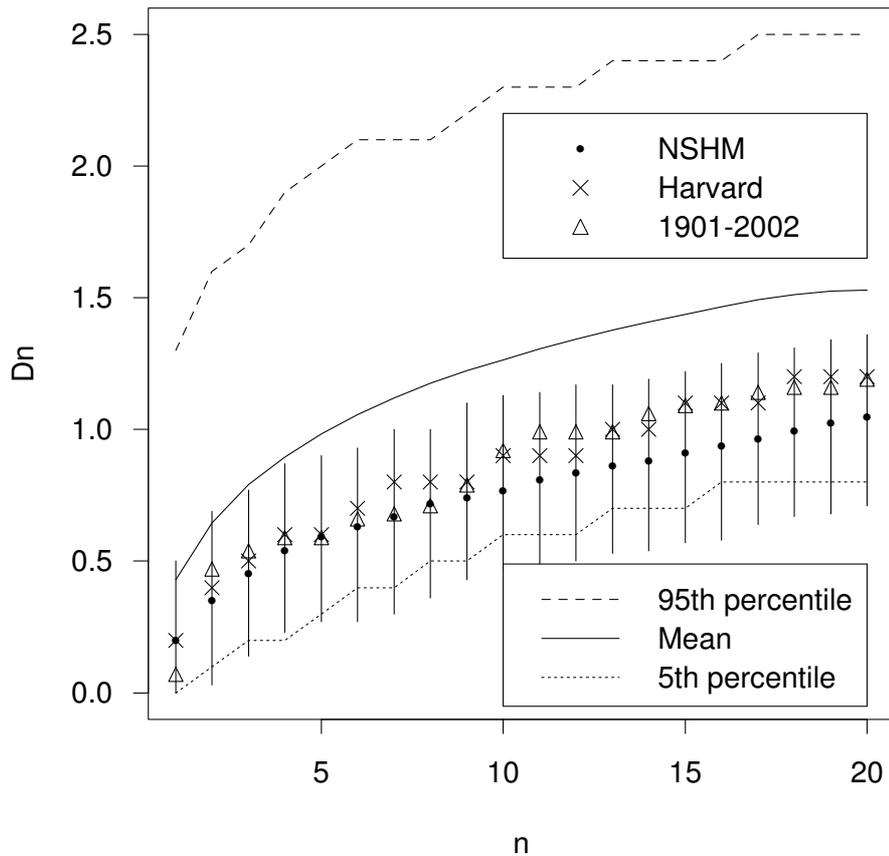
Figure 14 Plot of D_n against n for 100 synthetic catalogues of 101 years duration consistent with the NSHM, the Harvard catalogue of M_w 1979-2002 for the New Zealand region between 36° and 48° S., and a composite catalogue of $M_w > 6.45$ from 1901-2002 (see text for details). The error bars on the NSHM points extend to the 5th and 95th percentiles. Also shown are the expected value and 5th and 95th percentiles of the distribution of D_n for a Gutenberg-Richter set.. Here D_n is the difference between the largest and $(n+1)$[th] largest magnitudes in the catalogue.
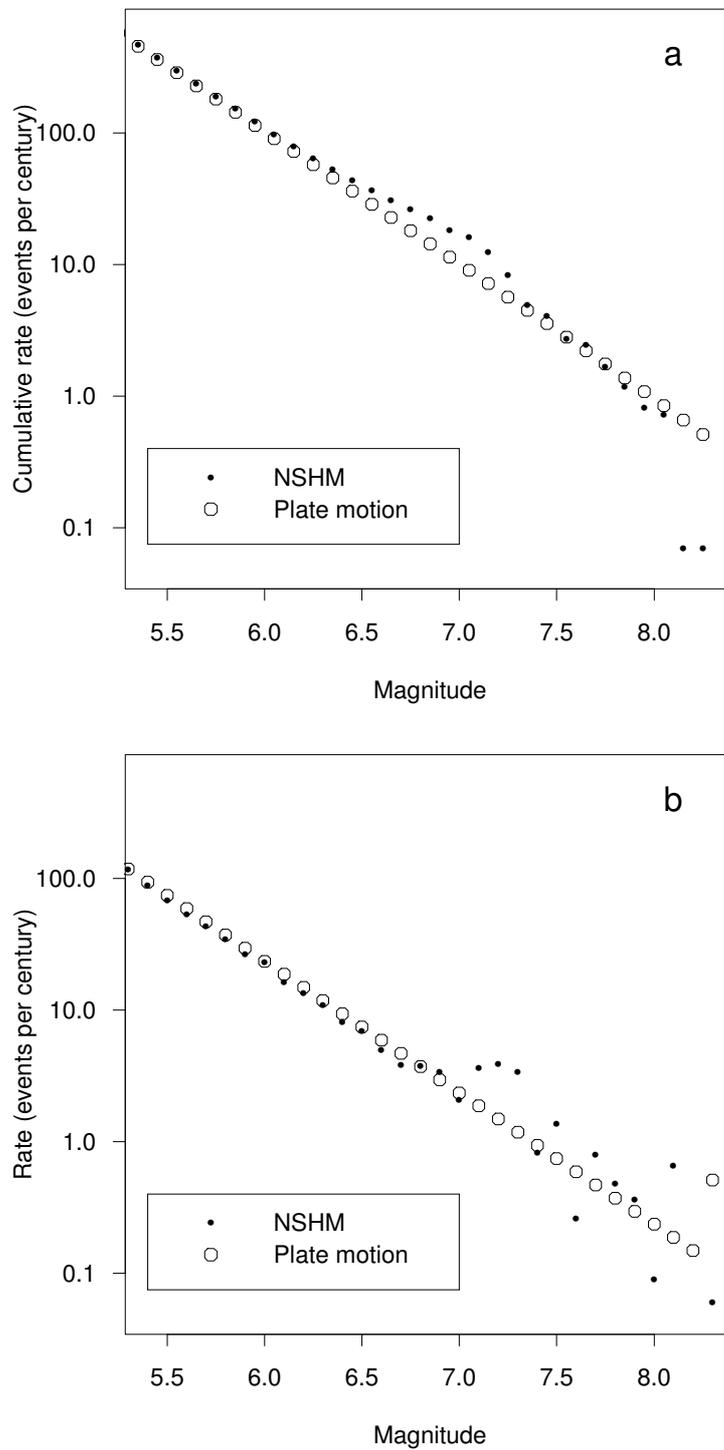
Figure 15   (a) Cumulative rate and (b) rate of earthquake occurrence as a function of magnitude for the total population of earthquakes from the NSHM and a plate motion-balanced magnitude-frequency distribution. See the text for further explanation.

10.0 NON-TECHNICAL ABSTRACT

Tests of Seismic Hazard Models

by D A Rhoades, W D Smith and M W Stirling

Assessment of seismic hazard in New Zealand relies on the national seismic hazard model (NSHM). This model embodies a catalogue of active faults and a so-called background seismicity model. The former is derived mostly from geological studies of active faults throughout New Zealand, and estimates the likely magnitude and frequency of occurrence of earthquakes on each fault. The latter describes the very large number of mostly smaller earthquakes that occur without any apparent relation to known faults. Assessing earthquake hazard at any given location involves examining the likely ground motion that can occur there, caused by nearby earthquakes (a) on faults, and (b) in the background.

The report examines the accuracy with which the NSHM represents the occurrence of earthquakes in New Zealand: where, how often and how large. Some significant discrepancies are found from what has occurred historically. In particular, the NHSM predicts that in a period of 160 years there should be more than 50 earthquakes that are accompanied by surface rupture of faults; only 10 have been observed between 1840 and 2000.

The report also checks the rate at which seismic energy is released throughout New Zealand, as estimated by geodetic data and as predicted by the NHSM. It seems that the NHSM prediction is too low, although there are some complicating factors that are not yet well understood so this apparent discrepancy may not be significant.

A widely accepted model for earthquakes predicts that for every event of magnitude 7 there will be about 10 of magnitude 6, 100 of magnitude 5, and so on. The multiplier of 10 varies a little from place to place, but the rule seems ubiquitous. The Report finds that the NHSM does not follow this model as closely as would be expected.

Each of these tests is done by careful attention to statistical procedure. Significant discrepancies have been found, which suggest that further development work is needed to refine the model so that it represents more accurately the earthquakes that are likely to occur in New Zealand.

11.0   TECHNICAL ABSTRACT

Tests of Seismic Hazard Models

by D A Rhoades, W D Smith and M W Stirling

Tests are described for checking the adequacy of a seismic hazard model against independent data. They have been applied to the national seismic hazard model (NSHM) for New Zealand Significant discrepancies revealed by these tests direct attention towards the most unsatisfactory aspects of the NSHM, including data and assumptions.

The Poisson process assumption provides a basis for statistical tests comparing the frequency of surface fault rupture in the seismic hazard model with the historical record of surface fault rupture. Application of such tests to the NSHM has revealed a statistically significant discrepancy between the model and the historical record. The NSHM predicts a frequency of surface fault rupture in New Zealand earthquakes more than five times the historical rate, with more than 50 fault ruptures expected since 1840 and only 10 observed. The discrepancy is most marked in the Taupo Volcanic Zone, but is also appreciable in northern Canterbury and southern Fiordland. It is likely to be due to multiple causes, including underestimation of the mean recurrence interval between ruptures of some fault sources in the NSHM, a possible tendency for fault ruptures to occur in clusters, and probable incompleteness in the record of historical surface fault ruptures

Under the stationary Poisson model, the elapsed time since the last rupture on any fault has an exponential distribution. If the recurrence interval is known, this allows the construction of a statistic, from the elapsed time, which has a common uniform distribution for all faults. Using the data from all faults together, goodness of fit tests can be used to check the overall consistency of the elapsed time and recurrence interval data. Application of such tests to the NSHM has not revealed any significant discrepancies. The statistical power of the tests is weak because of the small proportion of fault sources for which any information on the time of last rupture is available. Future paleoseismic studies should aim to increase the available information constraining the times of last rupture of faults.

Geodetic data provide estimates of strain accumulation in any area of the crust and these can be converted into estimates of moment accumulation rate. Similarly seismic hazard models provide estimates of rates of occurrence of earthquakes of any given magnitude in any area, and these can be converted into estimates of the rate of moment release. If most of the moment rate is released in earthquakes, the ratio of the seismic moment rate to the geodetic moment rate should be not much less than one. A comparison of the seismic moment rate in

the crustal seismotectonic zones of the NSHM with that expected from geodetic data has shown that overall the strain released under the model accounts for 0.6-0.8 of the geodetic moment rate. Within the individual seismotectonic zones, the ratio of NSHM moment to geodetic moment varies between 0.05 and 2. Further research is needed to understand these discrepancies.

Catalogues of earthquakes covering long time intervals and large areas are found nearly always to conform to the Gutenberg-Richter magnitude-frequency relation with modifications to accommodate the limit on the maximum size of an earthquake in the Earth's crust. Seismic hazard models of a region should therefore conform both with this relation and with the magnitude-frequency distribution in the historical record. Comparison of the NSHM magnitude-frequency distribution with the historical record and with a plate motion-balanced frequency-magnitude distribution has revealed some discrepancies between the model and the data, and between the model and the Gutenberg-Richter relation. While the magnitude-frequency relation in the NSHM shows wide scatter above M7.0, there is a surplus of earthquakes in the range M7.0 - M7.3 in the NSHM when compared with either historical data or the Gutenberg-Richter relation. Within the same magnitude range there is a deficit of distributed-source earthquakes and a more-than-compensating surplus of fault-source earthquakes under the NSHM. Modifications to the NSHM should be considered in order to rectify these discrepancies. The characteristic magnitude and/or mean recurrence interval probably needs to be adjusted for some fault sources. The upper magnitude limit of distributed earthquakes may need to be increased for some seismotectonic zones and a roll-off rather than a sharp magnitude cut-off adopted Also, variability in the magnitude of earthquakes on a given fault source should be allowed for.