# Testing the Hypothesis that Earthquake Hazard is Uniform Across Australia

**Cvetan Sinadinovski**[1] and Kevin McCue[2]

[1] - Advanced Research Centre, Saudi Aramco, Saudi Arabia
[2] - Department of Informatics and Computing, Central Queensland University, Australia

## Abstract

Estimating the probability of damaging ground shaking at any site in some specified time period for earthquake hazard assessment, depends on many assumptions and models. In intraplate areas such as Australia there is no causative model akin to Plate Tectonics for interplate regions. On a World scale, seismologists confidently expect that the distribution of seismicity over the next few years will mirror that of the past since most earthquakes occur along plate boundaries.

In Australia, a similar inspection of epicentre maps shows that past earthquakes are not evenly spread and appear to be clustered. Plate tectonics assumes the plates are rigid and without internal deformation. Even so, one of the critical assumptions underlying all past earthquake hazard assessments in Australia is that earthquakes are more likely to continue to happen where they have in the last 100 years than in places without epicentres, ie the clustering pattern will be sustained. This assumption is demonstrably not always true leading some critics to claim that earthquakes can occur anywhere and therefore that the hazard should be uniform across the continent.

We model the past seismicity statistically to show that the pattern is not representative of a uniform distribution at a high level of confidence.

**Keywords:** Seismic hazard assessment, intraplate earthquakes, chi-square test, declustering foreshocks and aftershocks

.

## Introduction

Estimating the probability of damaging ground shaking at any site in some specified time period for earthquake hazard assessment, depends on many assumptions and models. In intraplate areas such as Australia there is no causative model akin to Plate Tectonics that explains earthquake behavior in interplate regions so successfully. The seismicity along plate boundaries is in continuous broad bands but in Australia zones of past earthquakes are not evenly spread and appear to be clustered. Plate tectonics assumes the plates are rigid and without internal deformation without internal earthquakes, which is evidently incorrect.

Epicentre maps of Australian earthquakes since 1901 when seismographs were first deployed show that events are not evenly distributed. Some have argued that there is a pattern in the earthquake epicentres which can be explained as a response to the externally applied tectonic stress at the plate boundaries (McCue *et al.*, 1998). Others claim that earthquakes can occur anywhere because sometimes they occur where least expected such as the Tennant Creek sequence of 1987/88 (Jones *et al,* 1991), and therefore that any hazard rating should be uniform across the continent.

One of the critical assumptions underlying all past earthquake hazard assessments in Australia is that earthquakes are more likely to continue to happen where they have in the last 100 years than in new places, i.e. that the clustering pattern will be sustained. The Australian continent is an assemblage of crustal blocks of vastly different geological age and rock composition in the middle of a thin plate composed mainly of relatively young oceanic basaltic crust. Stress at the plate boundaries is redistributed within the plates and measurements (Lambeck *et al.*, 1984) have shown that within the continent it is neither of equal amount nor constant direction so it would be surprising if earthquakes occurred uniformly across continental Australia. We hypothesise that the distribution of earthquakes is not uniform.

However, such clustering might be observed when a uniform distribution is undersampled and does not necessarily prove this hypothesis. Thus it is necessary to undertake a formal analysis to test whether the perceived spatial clustering could be the consequence of random sampling from a uniform distribution. If it is, then the perceived spatial clustering provides us with no information regarding the location of future earthquakes, for the probability of occurrence would be independent of past epicentre locations.

Our procedure is to characterise the continental crust as a series of adjoining rectangular blocks along lines of latitude and longitude which can then be conveniently gridded into ‘cubes’ of crust of roughly equal surface area for the analysis. Different sized cubes can be tested and the test consists of comparing the number of observed and predicted earthquakes in each cube.

## Formal Analysis

The analysis follows that of McFadden *et al.* (2001) but updated for earthquakes that have occurred in the interim. We chose again the chi-square, $\chi^2$, goodness of fit test for comparison of the observed and expected frequency of earthquakes within each cube. Given $k$ cubes each of area $A_j$ ($A = \Sigma A_j$), then under the null hypothesis of a uniform distribution of earthquakes the expected number of earthquakes $e_j$ in each cube is given by

$$e_j = \frac{A_j}{A} N \tag{1}$$

where $N$ is the total number of earthquakes. If $o_j$ is the observed number of earthquakes in the cube $j$, then under the null hypothesis the statistic

$$X^2 = \sum_{j=1}^{k} \frac{(o_j - e_j)^2}{e_j} = \sum_{j=1}^{k} \frac{o_j^2}{e_j} - N \tag{2}$$

is $\chi_\nu^2$-distributed with ν = ($k$-1) degrees of freedom (Hoel, 1971). If the observations are in fact drawn from the distribution specified by the null hypothesis then one would expect $X^2$ to be about ν, the number of degrees of freedom. Conversely, if the observations are not drawn from the assumed distribution then the deviation of observed values from expected values will be large and so the observed value of $X^2$ will be much larger than expected under the null hypothesis. In that case, if the observed value of $X^2$ is too large (e.g., if the probability of observing such a large or larger value is less than 5%) then the null hypothesis may be rejected.

First an area outlining Australian continental crust was defined and then a uniform (in degrees of latitude and longitude) grid was placed over it. Those grid boxes that lie entirely outside the continental area were ignored for the analysis. Those grid boxes that lie entirely within the continental area were included as individual test regions with their area adjusted for the latitude range.

Two data sets have been used for testing. The first is the complete set of earthquakes in the Australian region as extracted from the Geoscience Australia Earthquake Database. Small earthquakes are naturally more difficult to detect and locate than large earthquakes. Consequently, a poor distribution of seismic stations can introduce a bias in the spatial distribution of small observed earthquakes. To avoid such a bias we have analysed the spatial distribution of the data sets only for those times in which we are confident that the seismic network has been able to identify all earthquakes of the specified magnitude in the Australian continental region. The first seismographs were established in Australia in 1901 and we are confident of having detected all quakes with magnitudes M≥6.0 since then (McFadden *et al.*, 2001). Carnegie Institution instruments were installed in 1959 and consequently all quakes with magnitudes M≥5.0 have been detected since then. Major seismic arrays were commissioned in Northern Territory at Alice Springs and Warramunga and so the data set is complete for magnitudes M≥4.0 since 1965. Subsequently, many dams have been instrumented and the national networks expanded so that the data set is complete for magnitudes M≥3.2 since 1980. In these data sets no distinction or identification is made of the temporal and spatial clustering of quakes as foreshocks or aftershocks from a single main shock.

The second type of data set, referred to as a declustered data set, has the same magnitude ranges but identified foreshocks and aftershocks were removed. A quake was considered to be a foreshock or an aftershock and was removed if it was within a distance $d$ km and time $t$ years of the main shock (McCue,1990)

$$d = 10^{(\mathrm{M}-4.11)/1.65} , \tag{3}$$

where M is the local magnitude of the main shock, and the time is 10 years for magnitude 7, 1 year for magnitude 6, 3 months for magnitude 5, and 10 days for magnitude 4.

The maps are next presented simply as a square grid in latitude and longitude. The continental area is shaded and a grid superposed on the image. Most earthquakes in continental Australia are crustal, shallower than 40 km in depth and no further segregation by focal depth was attempted in this study.

## Data Analysis

Figure 1 shows an analysis of the complete data set with M≥3.2 on a 1°x1° grid. As mentioned earlier the analysis is restricted to events observed since 1980. For this analysis $X^2 \sim 2.5\text{x}10^5$ with 1399 degrees of freedom. Had the earthquakes been drawn from a spatially uniform distribution then we would expect $X^2$ to be of the order of the number of degrees of freedom. We would have been able to reject the null hypothesis of a spatially uniform distribution at the 95% confidence level had $X^2$ exceeded 1475 or at the 99.9% confidence level had $X^2$ exceeded 2100. Clearly, with such a large $X^2$ the possibility that the observed spatial clustering could have been obtained by random sampling from a spatially uniform distribution is effectively zero.

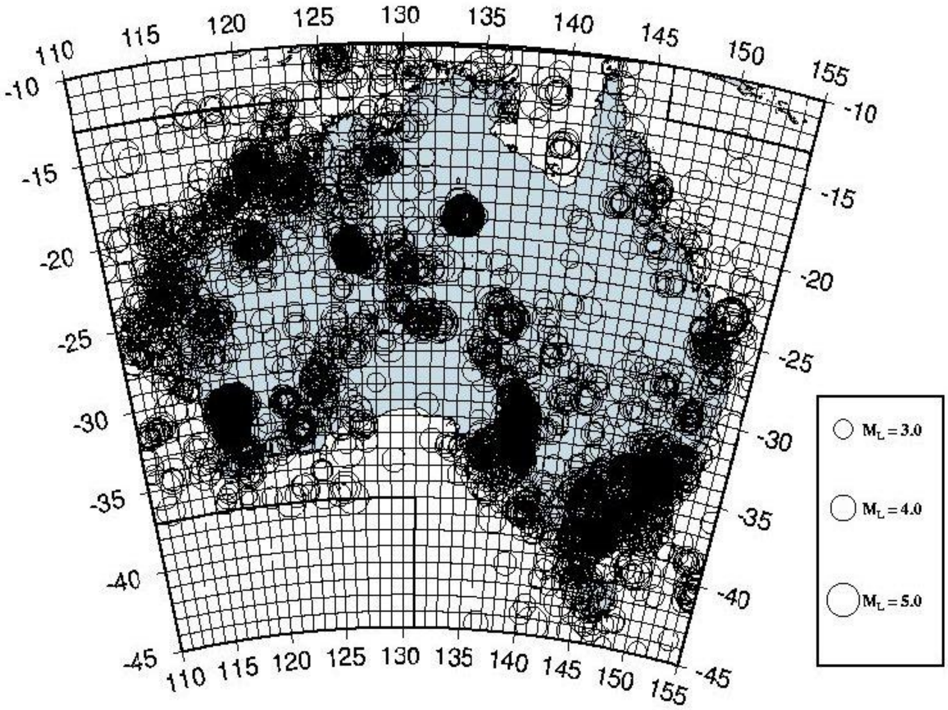


**Figure 1.** The complete data set for earthquakes M≥3.2 since 1980 on a 1°x1° grid. Total number of earthquakes = 5193, total number of boxes=1400, $X^2 \sim 2.5\text{x}10^5$ for 1399 degrees of freedom with a frequency of occurrence 4, probability is ~ $4\text{x}10^{-4}$%.

An analysis of the same data set but with a 5°x5° grid also gives a very large $X^2$ (around $2.5\text{x}10^4$) with 55 degrees of freedom. Again, there is effectively no possibility that the observed spatial clustering could have been obtained by sampling from a uniform distribution (99.9% confidence level is 90). This conclusion appears secure from a scale of 1º to 10º.

Figure 2 shows an analysis of the declustered data set with M≥3.2. For this analysis $X^2 \sim 10^5$ with 1299 degrees of freedom the null hypothesis of a spatially uniform distribution would be rejected at the 99.9% confidence level if $X^2$ exceeded 2000. Obviously, with such a large $X^2$ there is effectively no possibility that the observed spatial clustering could have been obtained by random sampling from a spatially uniform distribution. An analysis of the same data set but with a 2.5ºx2.5º grid gives $X^2$=3.6x10$^4$ with 1399 degrees of freedom, and again, there is effectively no possibility that the observed spatial clustering could have been obtained by sampling from a uniform distribution. This conclusion is consistent from a scale of 1º to 10º.

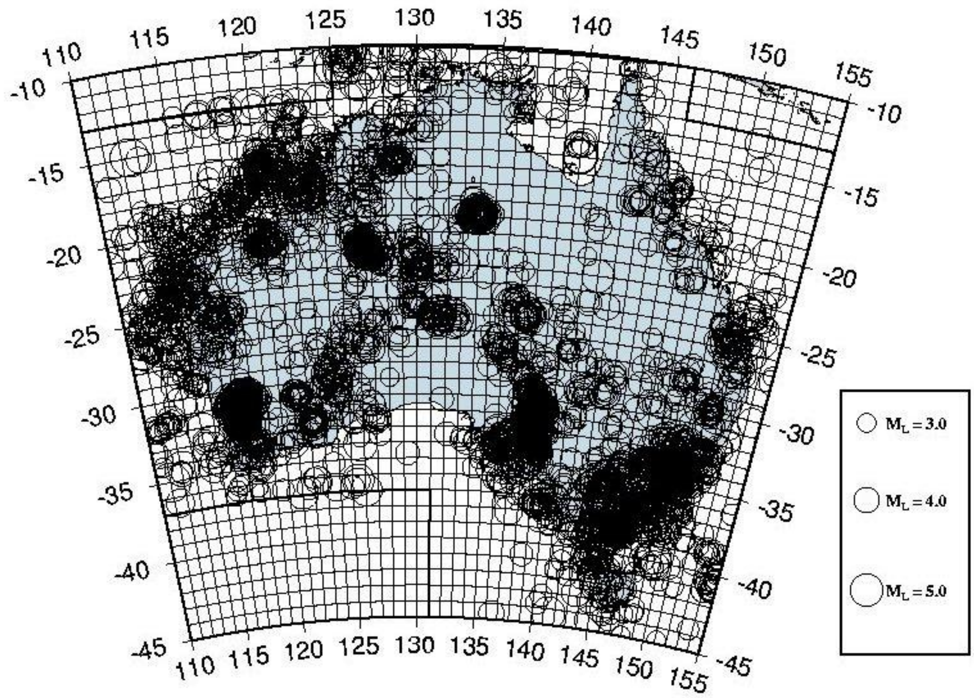


**Figure 2.** The declustered data set for earthquakes M≥3.2 since 1980 on a 1ºx1º grid. Total number of earthquakes=4308, total number of boxes=1300, $X^2$=3.6x10$^4$ for 1299 degrees of freedom with a frequency of occurrence 3, probability is ≥ 10$^{-3}$%.

For considerations of hazard assessment, a more relevant analysis is perhaps that of the declustered data set for magnitudes of earthquakes that can cause damage. Figure 3 shows the analysis of the declustered data set with magnitudes M≥4.0 since 1965 on a 2.5ºx2.5º grid. The slightly larger grid has been chosen because there are too few earthquakes in this analysis to give reliable statistics for a smaller grid. With 224 degrees of freedom, the null hypothesis of a uniform spatial distribution could be rejected at the 99.9% confidence level if $X^2$ exceeded 400: with an observed value of 7.78x10$^3$ it is concluded that there is effectively no possibility that the observed spatial clustering could have been obtained from a uniform distribution. Figure 4 shows the analysis of a data set with magnitudes M≥5.0 on a 5ºx5º grid. The larger grid has been chosen because there are too few earthquakes in this analysis to give reliable statistics for a smaller grid.

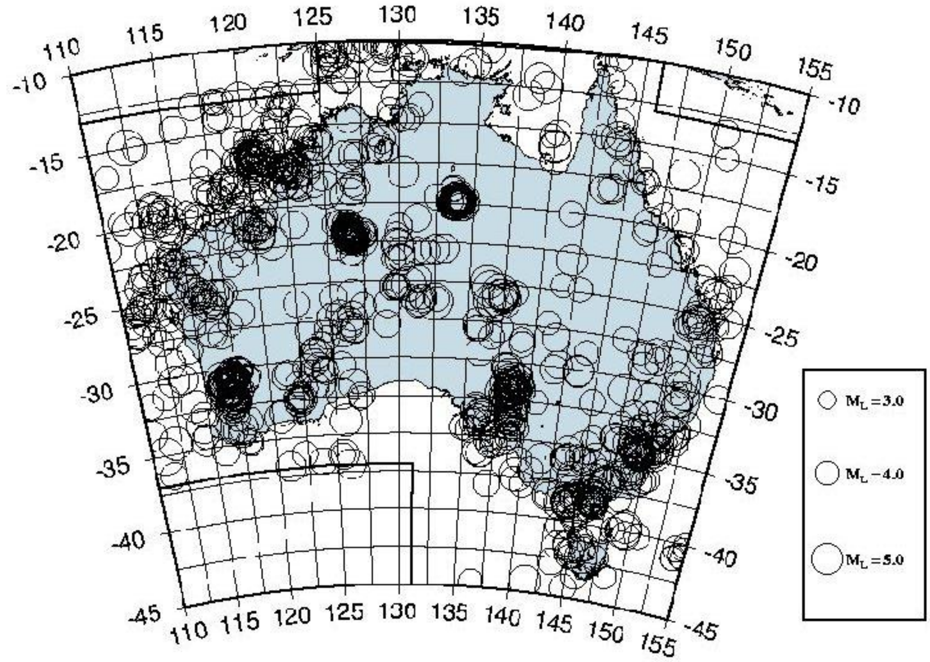


**Figure 3.** Analysis of the declustered data set for earthquakes with M≥4 since 1965 on a 2.5°x2.5° grid. Total number of earthquakes = 1057, total number of complete boxes = 225, $X^2$=7.78x$10^3$ for 224 degrees of freedom with a frequency of occurrence 4, probability is around $10^{-4}$%.

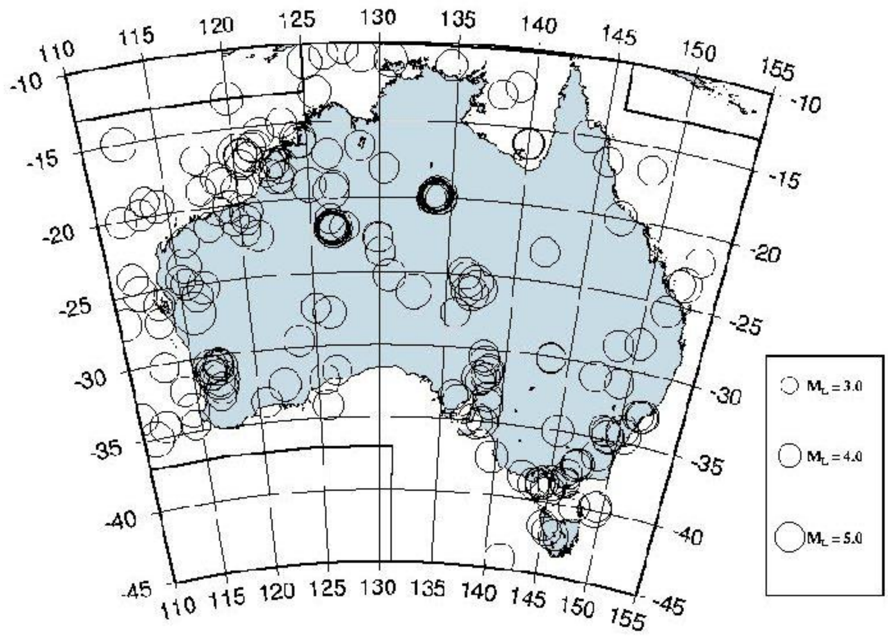


**Figure 4.** Analysis of the declustered data set for earthquakes with M≥5 since 1959 on a 5°x5° grid. Total number of earthquakes = 227, total number of boxes = 55, $X^2$ ~ 500 for 54 degrees of freedom, probability is around 0.01%.

With 54 degrees of freedom, the null hypothesis of a uniform spatial distribution can be rejected at the 99.9% confidence level if $X^2$ exceeded 90: with an observed value of 500 it is evident that there is again insignificant possibility that the observed spatial clustering could have been obtained from a uniform distribution.

On a $10^o$x$10^o$ grid the observed value of $X^2$ is again such that with 12 degrees of freedom leads to the similar conclusion for the larger scale. Since 1901 there have been only 11 large earthquakes with magnitudes 6 and more. These numbers are inadequate for statistical analysis even with large boxes.

## Conclusions

Use of chi-square, $X^2$, goodness of fit test is most appropriate when the expected frequency of occurrence is not too small. Some statisticians suggest that each value should be greater or equal to five. To strictly honour that rule we have to increase the cell size because there are not enough large earthquakes for such statistics, however increasing the area for the cells could lead towards too greater generalization of the geological structures/blocks of interest potentially involved in causing earthquakes.

Small magnitude quakes can still pose a risk if they are very shallow. Thus it might be interesting to look at quakes with magnitude down to 2.0 and depths within the top 5km. Those earthquakes again appear strongly clustered but there are problems in the analysis in that there is an absence of depth information for many earthquakes in the Geoscience Australia catalogue and for small earthquakes the spatial variation of the network sensitivity imposes a bias on the spatial distribution of observed events.

To summarise, this study has produced the following results:

**Full data set**

| M | Grid | Critical value of $X^2$ at the 99.9% confidence level for a uniform distribution | $X^2$ (Calculated) |
|---|---|---|---|
| ≥3.2 | $1^o$x$1^o$ | 2100 | $>2.5\times10^5$ |
| ≥3.2 | $5^o$x$5^o$ | 90 | $>2.5\times10^4$ |

**Declustered data set**

| M | Grid | Critical value of $X^2$ at the 99.9% confidence level for a uniform distribution | $X^2$ (Calculated) |
|---|---|---|---|
| ≥3.2 | $1^o$x$1^o$ | 2000 | $\sim10^5$ |
| ≥3.2 | $5^o$x$5^o$ | 90 | $>1.5\times10^4$ |
| ≥4.0 | $5^o$x$5^o$ | 90 | >2500 |
| ≥5.0 | $5^o$x$5^o$ | 90 | ~500 |

For each of the data selections considered in the figures and for scales ranging from a 1ºx1º grid to a 10ºx10º grid, the null hypothesis of a spatially uniform underlying distribution of earthquakes can be rejected with an exceptionally high degree of confidence.

The sample is, of course, highly restricted in time and might not necessarily provide a good representation of the earthquake distribution over several thousands of years. However, in terms of estimating the hazard across Australia for the time span of a hundred years such restriction is unlikely to be of consequence. Thus it may be safely concluded that the current earthquake hazard across Australia is non-uniform. This is as was expected and the observed distribution must be the consequence of geology and tectonics and demands the assignment of different levels of hazard to different regions of the continent.

## References


Hoel, P.G. 1971. *Introduction to Mathematical Statistics, Fourth Edition.* Wiley International, New York.

Jones, T.D. McCue, K.F. Denham, D. Gregson, P. Bowman, R., and Gibson, G. 1991. Three large intraplate earthquakes near Tennant Creek, Northern Territory on 22 January 1988. *BMR Journal of Australian Geology and Geophysics,* 12 (4), 339 - 344.

Lambeck, K., McQueen, H., W.S, Stephenson, R., A, Denham, D. 1984. The state of stress within the Australian continent. Annales Geophysicae 2(6) p723-741.

McCue, K. 1990. Australia’s Large Earthquakes and Recent Fault Scarps, *Journal of Structural Geology*, **Vol. 12**, No. 5/6, pp 761-766.

McCue, K.F., Somerville, M., and Sinadinovski, C., 1998 — GSHAP and the Proposed Australian Earthquake Hazard Map in Meckering 30 years on – how would we cope today? Proceedings of the Australian Earthquake Engineering Society Conference, Perth WA. Nov 4-5 1998, 18-1 – 18-6.

McFadden, P., Sinadinovski, C., McCue, K.F., and Collins, C. 2001. Is Earthquake Hazard Uniform Across Australia?, (Unpublished review).