Mapping the Magnitude of Completeness of the Australian Earthquake Catalogue

Stephen Sagar and Mark Leonard
Geoscience Australia

ABSTRACT

The minimum magnitude of completeness (Mc) of a compiled earthquake catalogue is considered to be an important factor in accurate hazard parameter estimation, and one which can vary on both a temporal and spatial basis.

This paper utilises a Mc grid-mapping procedure developed by Wiemer & Wyss (2000) and tested on highly seismic regions across the world, based on a statistical analysis of the frequency magnitude distribution of events sampled by a regular-grid. Variations of the underpinning regression and sampling techniques in this procedure have been tested on the most recent Australian catalogue, to assess their suitability in dealing with the episodic and sparsely distributed nature of Australian seismicity.

The process has emphasised the need to consider a number of issues in the Australian catalogue to successfully complete comparative Mc mapping at a national scale. For example, regional variations in minimum event magnitude reporting and the inclusion of event observations from smaller localised networks need to be addressed. Refinements of the tested techniques have been proposed which incorporate well established parameters of Australian seismicity to address some of the limitations encountered with the grid-mapping method, both statistical and spatial, that result from the sparse distribution of events in many areas of the country.

INTRODUCTION

In the accurate estimation of earthquake hazard parameters, one of the most important factors is the completeness of the sample under investigation (Wiemer & Wyss, 2002), and hence the appropriate determination of the minimum magnitude of complete recording (Mc) for the sample is critical. Mc is defined as the lowest magnitude above which all seismic events in a space-time volume are detected (Wiemer & Wyss, 2000).

Determinations of spatial variations in the Mc of the Australian earthquake catalogue have to this point been restricted to a pre-defined regional or seismic zone analysis (Leonard, 2007). This application of discrete Mc values to zones then introduces a high variability at the zone boundaries when developing hazard maps. Alternative techniques to derive detection thresholds based on network configuration, instrumentation and crustal attenuation have been developed by Cuthbertson (2006), independent of event catalogue data.

This paper applies variations of a technique developed by Wiemer & Wyss (2000), based on a uniform grid analysis of the earthquake catalogue. The determination of Mc is based on a linearity assumption of the cumulative frequency-magnitude distribution (FMD), introduced by Gutenberg & Richter (1944):

$$Log(N) = a - bM \tag{1}$$

where N is the number of earthquakes of a magnitude M, and a and b are constants describing the productivity of a volume and relative size distribution of events respectively.

Unlike the highly seismic regions to which this technique has been previously applied (Woessner & Wiemer, 2005; Wiemer & Wyss, 2000), Australia is characterised by a moderate to low level of episodic seismicity (Brown & Gibson, 2004; Leonard, 2007). A number of variations of the Mc mapping technique have been tested to evaluate their effectiveness in dealing with the nature of Australian seismicity and the sparse distribution of events across many areas of the country.

CATALOGUE SOURCE

The compiled catalogue used for analysis in this paper is considered the most current and complete catalogue in use at Geoscience Australia. Events such as swarms and aftershocks were identified by the catalogue provider, Gary Gibson, and were filtered to create a main-shock catalogue of 23,339 events between 1900 and 2007.

ISSUES IN CREATING A CATALOGUE FOR ANALYSIS

Magnitude shifts

Issues involving inadvertent changes in magnitude scale are evident in most catalogues around the world (Zuniga & Wyss, 1995). These changes can be significant in their effect on studies of catalogue completeness, both in the successful application of algorithms which assume a linearity of the FMD and the final estimation of Mc.

The Australian catalogue is subject to a number of magnitude scale variations, both on a temporal and regional basis (Leonard, 2007). The most notable of these in respect to mapping Mc occurred in the late 1980's and early 1990's, with the shift from the use of the general Richter scale to local magnitude scales reflecting local attenuation. The regional variation in this period of magnitude scale changes is illustrated in Figure 1, using regions defined in Leonard (2007). Northern Western Australia shows almost a 0.5 magnitude scale decrease between the period pre and post 1992 most likely due to a magnitude calculation change. In comparison, the South Australian region shows no discernable scale shift over the same period.

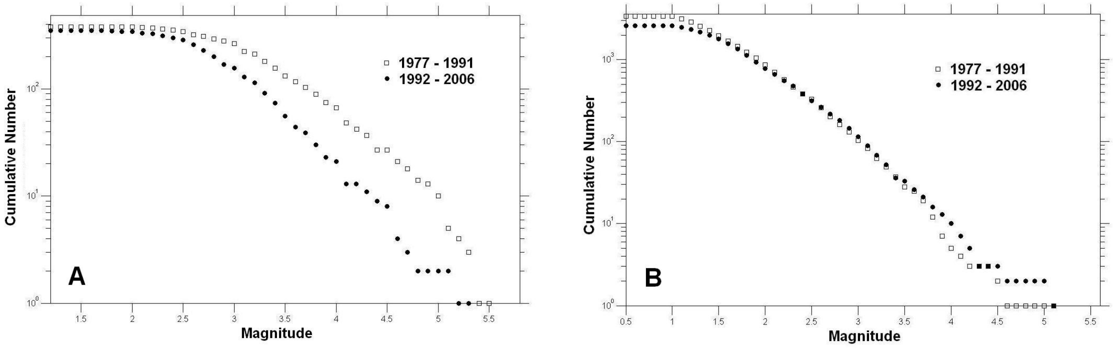

Figure 1 – A) FMD of Northern WA pre and post 1992 showing the magnitude shift evident in the change to a local magnitude scale. **B)** FMD of SA pre and post 1992 showing no magnitude scale shift between the periods.

To maintain a homogenous magnitude scale on a regional basis, and minimise the effect the various regional scale shifts may have on the application of the mapping algorithms, the period for further analysis was restricted to events post 1992 to ensure all regions had adopted the magnitude scale that would remain current to the extent of the catalogue.

Variable reporting and Network Changes

The FMD linearity assumption that is the basis of the 'goodness-of-fit' (GFT) mapping technique requires the space–time volume of events being examined to be generated from a consistent network configuration. Any possible changes in network configuration and/or variations in reporting, such as rounding of event magnitudes, have the potential to effect the successful application of the GFT mapping algorithm. Figure 2 illustrates two areas where these issues may be of concern in the current catalogue.

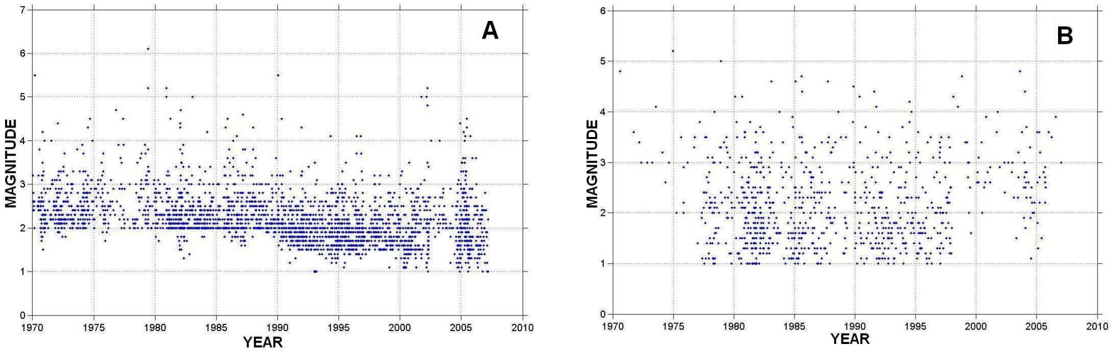

Figure 2 – A) Southern Western Australia Magnitude v Time. Note the increased number of events between 1 and 2 detected from 1990. Prior to 1990 evidence of these smaller events being truncated to magnitude 2 can been seen. **B)** Queensland Magnitude v Time. A marked increase in smaller magnitude activity can be clearly seen between 1977 – 1998, which is probably explained by an increase or change in the reporting network configuration.

MAPPING METHODOLOGY

The spatial mapping of Mc was completed using ZMAP software, an open-source program developed and written in Matlab by Wiemer (2001).

Event sampling in ZMAP was completed using a regularly spaced grid, set at 1 degree intervals, to extract events at each grid node. Two sampling methods were considered for event extraction: Fixed Radius and the Closest 'x' Events.

The Fixed Radius approach was considered to be unsuited to the Australian catalogue due to the sparse spatial distribution of events in many areas of the country. Due to the variability in sample sizes that this method produces, a high variability in the uncertainty of the derived Mc values would also be expected.

The Closest 'x' Events method was used to obtain the nearest 250 events to each grid node, considered a sufficient sample to maximise the potential of generating a Mc value at each node.

Calculation of Mc from each sample of 250 events was completed by application of the quantitative criterion developed by Wiemer & Wyss (2000), broadly summarised below. Their technique is based on the generation of a series of synthetic magnitude distributions, using *b* and *a* values generated at each magnitude bin by a maximum likelihood estimate (Aki, 1965). Each synthetic distribution represents a perfect fit to a power law for events above the specific minimum magnitude bin.

- Using the assumption of the linearity of the cumulative FMD above Mc, each synthetic distribution is then evaluated for 'goodness-of-fit' (GFT) to the FMD generated from the sampled 250 events.

- The residuals between the synthetic and observed distributions are calculated and used to produce an R value reflecting the extent to which the synthetic power-law can model the data variability of the observations.

- Mc is then determined at the magnitude for which the power law fits the observed catalogue at the R = 90% level, on the basis of the assumption made by Wiemer & Wyss (2000) that many catalogues will only reach this level of fit due to issues discussed above of magnitude scaling, varied reporting and natural changes in seismicity that cannot be filtered in the generation of a 'high quality' catalogue.

This hypothesis was tested on the Australian catalogue at both the 95% and 90% level. The generation of a valid Mc value was restricted to node samples which contained a sufficient number of events, set at *Nmin* = 50, above the determined Mc magnitude. This is suggested by Schorlemmer et al. (2004) as an appropriate cut-off, below which the synthetic power-law is fitted only to a small proportion of the data, generating a statistically unreliable result.

The results clearly verified the assumption of Wiemer & Wyss (2000), with only a small proportion of the grid nodes able to fit the power law and generate a Mc value at the 95% level. Mapping at the 90% level is more successful, although large sample areas still fall short of the 90% fit. These areas occurred in areas of both low and high seismicity, suggesting that the linearity of the FMD is being corrupted by some irregularity of reporting or clustering that was not identified in the catalogue analysis. It is not immediately apparent as to why the 90% GFT would not fit as expected, especially in areas of relatively high seismicity such as the Flinders region of SA.

Examination of a node located at the NW extent of the active South Australian seismic area (Long 137°, Lat -30°) shows evidence of the breakdown of the linearity assumption. The FMD for this node (Figure 3) shows a bi-modal distribution, with a clear curvature point at magnitude 2.9, which prevents the power-law fit at the 90% level, while still modelling to a sufficient portion of the data.

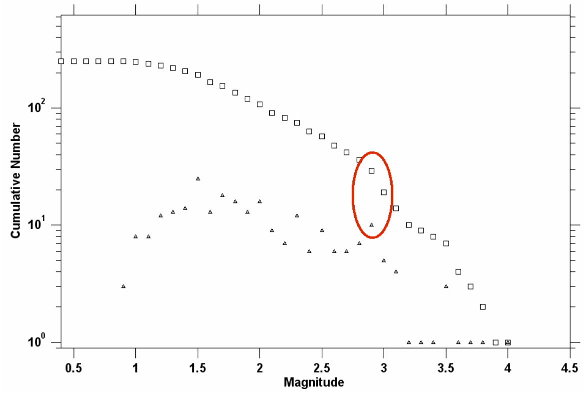

Figure 3 – Illustration of a bi-modal distribution and the breakdown of the linearity assumption of the cumulative FMD

Testing was then completed using an alternative method offered in the ZMAP software, considered a Maximum curvature fit (MaxC), which determines Mc at the highest value of the non-cumulative FMD.

In their assessment of the Mc estimation techniques used in this paper, Woessner & Wiemer (2005) found MaxC to produce a reliable and relatively robust estimation of Mc, yet with a somewhat systematic value underestimation of approximately 0.1-0.2. Figure 4 compares the derived Mc and line fit of both the 90% GFT and MaxC methods on the same subset of 250 events from the Australian catalogue.

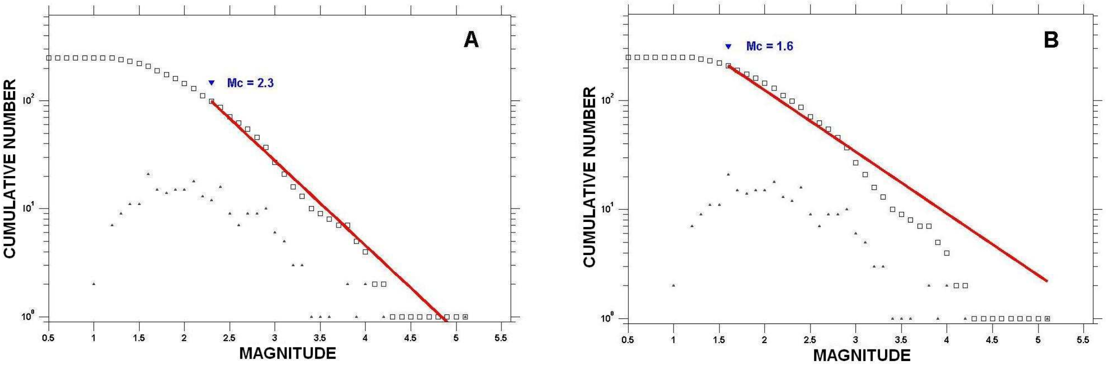

Figure 4 – A) 90% GFT Mc model **B)** MaxC Mc model

RESULTS AND LIMITATIONS

Due to the lower sensitivity of the MaxC method to any linearity deviations of the FMD, Mc values were resolved in areas not mapped by the 90% GFT method to create a complete national coverage for the period 1998-2007 (Figure 5). The period for this preliminary map was chosen to avoid the temporal variation of reporting shown in QLD in Figure 2 from pre-1998 to the current day.

A bootstrap procedure outlined by Woessner & Wiemer (2005), drawing samples with replacement, was used to get an indication of the uncertainty of the MaxC Mc map values. The standard deviation of the mapped values averaged between 0.15 and 0.2, with a maximum of 0.5 in areas marked by the transition from a low (1.2 – 2.0) to high (>2.4) Mc.

There is potential to use the mapping procedure to analyse smaller temporal windows of the catalogue, to identify the varying Mc over time due to these types of reporting/network variability. However, the small volume of event data available over these smaller discrete time periods would have an adverse effect on the statistical strength of the algorithm results.

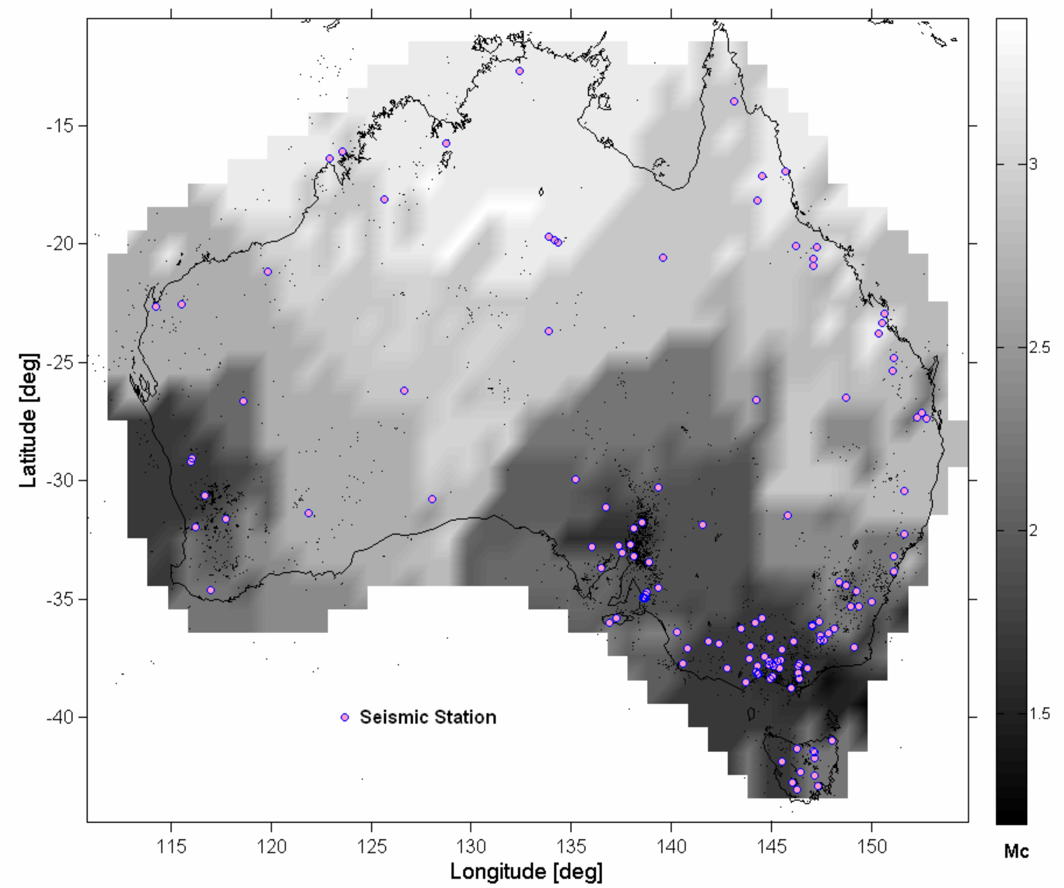

Figure 5 – Magnitude of Completeness (Mc) of the Australian Earthquake Catalogue 1998-2007 determined using the Maximum Curvature Method (MaxC). Seismic stations shown are an indicative representation of the network over this period.

Event Sampling Distances

An issue highlighted in previous studies using these techniques (Wiemer & Wyss, 2000; 2002), and one which is amplified by the sparse event distribution over a large area in the northern regions of Australia, is the search radius used to sample the required events in all of the mapping methods.

In Northern Queensland the search radius to sample 250 events has been extended up to 1500 km, in comparison to a radius of approx 200-600 km for the more active seismic regions of southern Australia. In these areas of large search radius, consideration must be given to whether the sampled subset of events represents the true detection capability of a station at that position. As the algorithm samples events purely based on distance to the node, it is possible that an event of a very low magnitude may be sampled for analysis at a node up to 1500 km away, well beyond where the actual event would be detected.

In their comment on the work of Wiemer & Wyss (2000), Rydelek & Sacks (2003) identify this as shortcoming of the Closest 'x' Events method; that the algorithm can theoretically calculate the same Mc value for a node located above a highly seismic area as one located some distance away in an aseismic region. This effect is not as highly pronounced in the Australian Mc maps as may be expected. An extended transition of low Mc values can be identified moving from the highly seismic SA region towards the area of lower activity in QLD, however this was considered preferable to the high sensitivity and variability of the Mc values derived at the zone edges of the 'Fixed Radius' method.

Constraint of B values

The MaxC method has been shown to systematically underestimate Mc, and has the potential to be sensitive to catalogue issues such as magnitude truncation (Figure 2a) which may artificially increase a particular magnitude to the highest value in the non-cumulative FMD.

To enable the application to the catalogue of the 90% GFT method, which to this point has been statistically limited by the small number of events above the derived Mc, it may be possible to constrain the GFT model slope by the known *b* values in Australia.

In future work, the derived *b*-values known in various zones of Australia (Leonard, 2007), may be coded into the Matlab ZMAP program to serve as constraints on the slope of the GFT line that is fitted at each node. This may have the potential to overcome some of the statistical and underestimation issues currently being encountered.

CONCLUSIONS

The techniques used in this paper have highlighted the need for close catalogue examination before these Mc mapping techniques can be applied. Issues such as the identification of events generated from temporary arrays and the reporting of events only above a particular magnitude threshold need to be considered to produce a more accurate 'national level' catalogue completeness picture. The sparse and episodic nature of Australian seismicity may require that further elements, such as constraining *b*-values, be added to the modelling algorithms to increase the statistical reliability of the results.

REFERENCES

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

Brown, A. and G. Gibson. (2004) A multi-tiered earthquake hazard model for Australia, *Tectonophysics, 390*, 25-43.

Cuthbertson, R. (2006) Automatic Calculation of seismicity rates in eastern Queensland. *Proceedings of the 2006 Australian Earthquake Engineering Society Conference*. 137-144

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

Leonard, M. (2007) One Hundred Years of Earthquake Reporting in Australia. (in press)

Rydelek, P. A., and I. S. Sacks (2003). Comment on "Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the Western United States, and Japan" by Stefan Wiemer and Max Wyss, *Bull. Seism. Soc. Am.* 93**,** 1862–1867.

Schorlemmer, D., S. Wiemer and M. Wyss (2004) Earthquake statistics at Parkfield: 1. Stationarity of b values, *J. Geophys. Res., 109*, B12307.

Wiemer, S. (2001). A software package to analyze seismicity: ZMAP, *Seism. Res. Letts.*, 72, 373-382

Wiemer, S., and M. Wyss. (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western US and Japan, *Bull. Seism. Soc. Am.*, *90*, 859-869.

Wiemer, S., and M. Wyss (2002). Mapping spatial variability of the frequency-magnitude distribution of earthquakes, *Adv. Geophys.* 45**,** 259–302.

Woessner, J., and S. Wiemer. (2005). Assessing the quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and Its Uncertainty, *Bull. Seism. Soc. Am.*, *95*, 684-698.

Zuniga, R., and M. Wyss. (1995). Inadvertent changes in magnitude reported in earthquake catalogs: Their Evaluation through b-value estimates, *Bull. Seism. Soc. Am.*, *85*, 1858-1866.