Performance of national scale smoothed seismicity estimates of earthquake activity rates

Jonathan Griffin[1], Graeme Weatherill[2] and Trevor Allen[3]

1. Corresponding Author, Geoscience Australia, Symonston, ACT, 2601
   Email: jonathan.griffin@ga.gov.au
2. Helmholtz Centre Potsdam, German Research Centre for Geosciences (GFZ)
3. Geoscience Australia, Symonston, ACT, 2601

Abstract

Four smoothed seismicity models were submitted for consideration to the 2018 National Seismic Hazard Assessment (NSHA18). Three of the models used fixed smoothing kernels, each with slightly different implementations, while the fourth used an adaptive smoothing kernel that varies spatially based on earthquake density. In this paper we assess the performance of the adaptive kernel model in forecasting seismic activity rates over decadal timescales. We analyse part of the earthquake catalogue used for NSHA18 to develop the models, and calculate the log-likelihood of the model activity rates against the remainder of the catalogue. The performance of the models at forecasting rates of earthquakes with different magnitudes over different lengths of training and forecast periods is tested against a null hypothesis of a uniform spatial distribution. The results are used to evaluate and inform the final implementation of the adaptive kernel smoothed seismicity model for calculating ground motions with a 10% in 50 years probability exceedance for the NSHA18.

Keywords: national seismic hazard assessment, smoothed seismicity, earthquake forecast testing

1) INTRODUCTION

Smoothed seismicity models provide spatially-varying earthquake occurrence rates that are derived by smoothing the observed rates of earthquake occurrence across a given smoothing kernel. This gives these models the advantage of being purely data-driven, with less dependence on the interpretation of the model developer. While smoothed seismicity models give forecasts of the spatial distribution of earthquake occurrence rates above a given minimum magnitude, they do not provide information on the full magnitude-frequency distribution of earthquakes.

The use of smoothed seismicity models is predicated on the assumption that earthquakes observed in the historical catalogue will be representative of future seismicity. With palaeoseismological evidence showing that faults can be quiescent for tens or even hundreds of thousands of years (Clark *et al.*, 2012; Clark *et al.* 2014; Clark *et al.* 2015), this assumption clearly does not hold for very long return periods, and a smoothed seismicity model does not account for future activation of a presently quiescent fault. However, an episodic model of seismicity in Australia does imply that currently seismically active regions may continue to be active over the geologically short timescales (e.g. 50 years) considered for seismic hazard assessments. In this paper we develop smoothed seismicity forecasts based on decadal scale subsets of the Australian earthquake catalogue and test their performance against remaining subsets of the catalogue. This analysis seeks to test the extent to which short records of seismicity are predictive of future seismicity in a stable continental region such as Australia.

Although four smoothed seismicity models (Cuthbertson, 2016; Griffin *et al.*, 2016; Hall *et al.*, 2007) have been submitted for inclusion in the NSHA18 seismic source model, this paper focuses on one of the two models developed at Geoscience Australia, which uses an adaptive smoothing kernel (Griffin *et al.*, 2016). Due to continuing work on the earthquake catalogue for NSHA18, we use the 2012 NSHA earthquake catalogue, with adjusted magnitudes following work by Ghasemi *et al.* (2016) and Ghasemi and Allen (2017). This catalogue contains earthquakes up to the end of 2010, and we use the completeness model of Burbidge (2012) and Leonard *et al.* (2014; Table 1). The catalogue has been declustered using the Leonard (2008) time and distance windows, and based on the recommendations of Burbidge (2012), the Leonard (2008) declustering algorithm was modified to account for the long aftershock sequences proposed by Stein and Liu (2009) for large continental earthquakes.

Table 1: Completeness model for the earthquake catalogue.

Year	Completeness magnitude
1990	3.0
1980	3.5
1965	4.0

2) METHODOLOGY

In developing our smoothed seismicity model for forecasting future seismicity rates, we use both constant (Frankel, 1995) and adaptive (Helmstetter *et al.*, 2007) smoothing kernel methods. Both methods use a Gaussian smoothing kernel to spatially smooth the rates of earthquake occurrence. The Frankel (1995) model uses a fixed 50 km kernel bandwidth and smooths over three neighbouring cells. Large variations in the density of observed seismicity in Australia (e.g. high density in the Flinders ranges, low density in central Queensland) substantiate the use of an adaptive kernel approach, where the size of the smoothing kernel is dependent on the distance to the *K*th nearest earthquake (Helmstetter *et al.*, 2007). Helmstetter and Werner (2012) presented an extension of this method that uses an adaptive smoothing kernel in time and space, which avoids declustering the earthquake catalogue by taking the median value of the time history of each grid cell. However, we are cautious about applying this method to a low seismicity region such as in Australia where all, or almost all, the earthquakes in a particular region (e.g. Tennant Creek) may consist of a single mainshock and its aftershock sequence. Therefore, this study uses an earthquake catalogue declustered using traditional techniques and only smooths the spatial distribution following Helmstetter *et al.* (2007).

For the Helmstetter *et al.* (2007) method model parameters (*K*, and a minimum rate r_{min} to ensure there exists a non-zero rate of earthquake occurrence everywhere) are chosen through maximum likelihood optimisation. To more consistently evaluate the performance of the alternative models, we initially optimise *K* and r_{min} using a training catalogue (1964 – 2003) and a target catalogue (2004-2010). The same optimised values are then used for all subsequent earthquake rate forecasts. Following (Helmstetter *et al.*, 2007) we test values of *K* from 3-50 and find $K = 3$ and $r_{min} = 1.e\text{-}7$ to provide the optimal parameterisation.

To test the performance of smoothed seismicity models, we need a training catalogue to derive the model and a target catalogue to test the model. Again following Helmstetter *et al.* (2007) we calculate the Poisson log-likelihood (l) of the model as:

$$l = \sum_{i_x=1}^{N_x} \sum_{i_y=1}^{N_y} ln\left(P\left[\lambda\left(i_x, i_y\right), n\left(i_x, i_y\right)\right]\right) \tag{1}$$

where $N_x$and N_y are the number of cells in the *x* and *y* directions respectively, $\lambda\left(i_x, i_y\right)$ is the earthquake occurrence rate in each grid cell $\left(i_x, i_y\right)$ derived from the model,, and $n\left(i_x, i_y\right)$ is the observed rate of earthquakes within the cell.

One measure of the model's performance is the probability gain (*G*) per earthquake relative to a uniform spatial distribution (Helmstetter *et al.*, 2007):

$$G = exp\left(\frac{l - l_{unif}}{N_t}\right) \tag{2}$$

where l_{unif} is the log likelihood of a uniform spatial distribution of seismicity (still Poisson distributed in time, with the same rate everywhere), and $N_t$is the number of earthquakes within the target catalogue. The probability gain is a measure of how well the spatial distribution of seismicity within the training period predicts the spatial distribution of future seismicity (i.e. within the target period), compared with a uniform spatial distribution. If it is no better than a uniform distribution, the probability gain will be 1.0. For large stable continental regions such as Australia, the probability gain can indicate whether the spatial distribution of earthquakes over the short term can be considered stationary over timescales that are of interest to society's earthquake risk reduction efforts.

The results presented in the following section simply use the raw number of events within the catalogue, and do not account for catalogue incompleteness. This has the advantage of providing additional insights into the catalogue's completeness, and also means the analysis presented following does not depend on the Gutenberg-Richter *b*-value.

3) RESULTS

Beginning with first decade of the catalogue, 1965-1974, we calculate the probability gain for each subsequent decadal subset of the catalogue (Figure 1) using values of K = 3, 4 and 5 and for minimum magnitudes (M_{Min}) of 3.5 and 4.0. These results show higher probability gains in the short term (i.e. 1975-1984), particularly for a minimum magnitude of 3.5, compared with subsequent time periods. This trend may be attributed to one or both of the following two possibilities:

a) Incompleteness of the catalogue prior to around 1980 resulting in different spatial distributions of observed seismicity before and after this time; and/or
b) Short-term time-dependence at the decadal scale within the declustered earthquake catalogue.

Figure 2 shows smoothed rates for each of the decadal periods for a minimum magnitude of 3.5, along with the complete catalogue. In Figures 2a and b lower rates are forecast in central and north-eastern Australia, compared with later periods (Figures 2c, d, and e). This suggests that catalogue incompleteness is important here, consistent with the completeness model of Burbidge (2012). When the full catalogue is used (Figure 2f), the spatial distribution of seismicity rates becomes more strongly heterogeneous, since a greater number of total earthquakes in the catalogue reduces the distance to the Kth nearest earthquake and hence the size of the smoothing kernel.

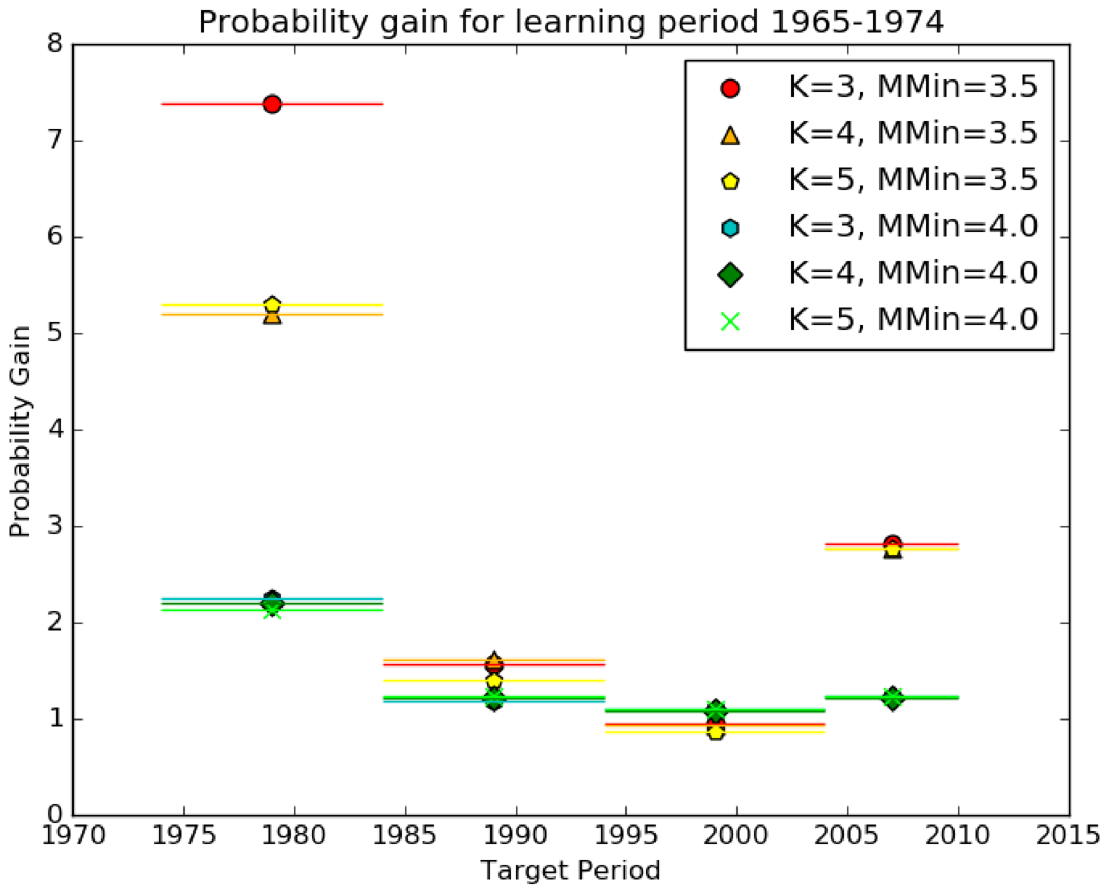

Figure 1: Probability gain for adaptive kernel smoothed seismicity models trained on data from 1965-1974 and tested decadal subsets of the catalogue: 1975-1984; 1985-1994; 1995-2004; and 2005-2010. Horizontal bars show the temporal extent of the target period and points show the middle of the period. Variations in the smoothing parameter *K* and the minimum magnitude threshold are shown.

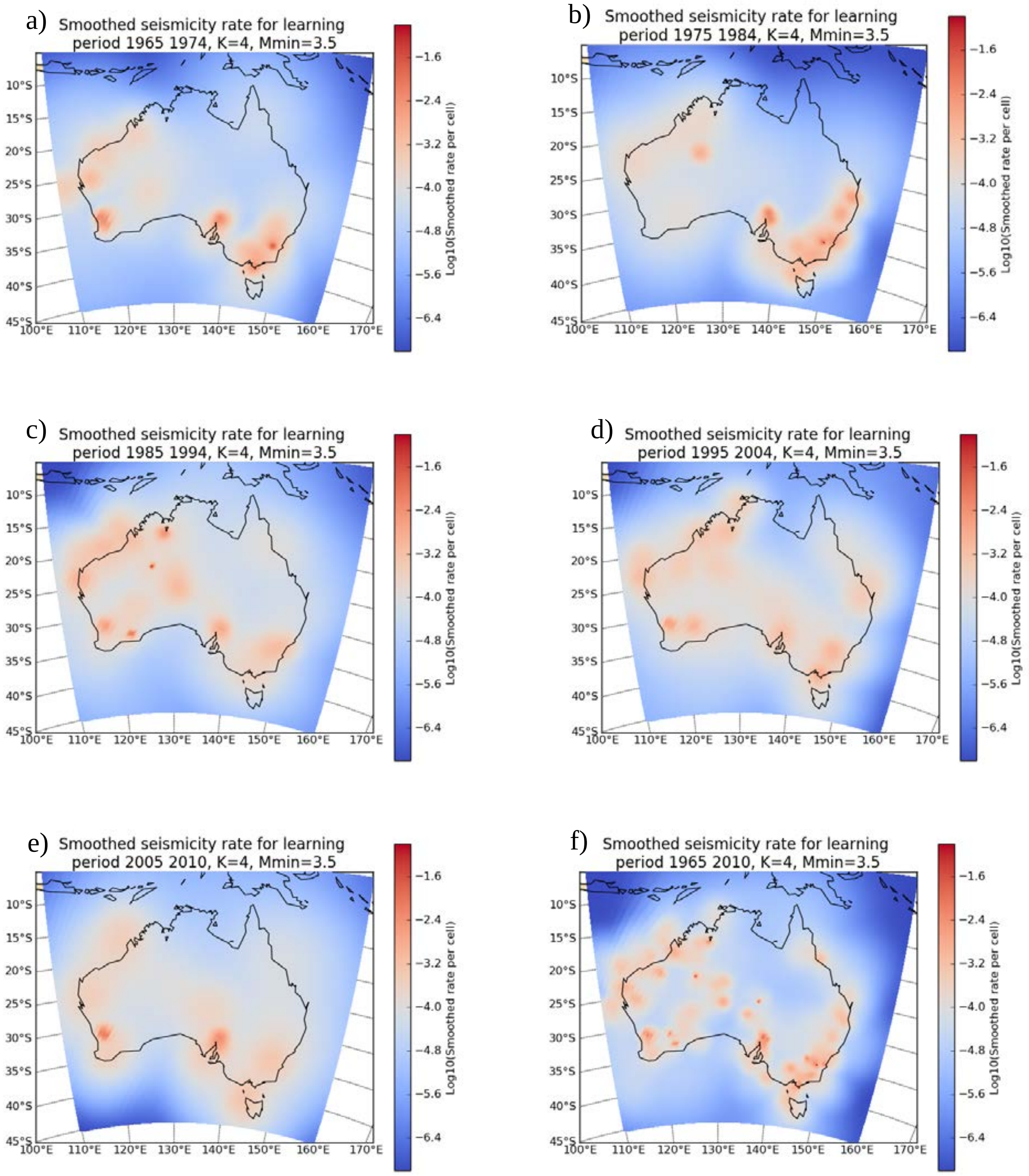

Figure 2 Adaptive kernel (*K*=4) smoothed seismicity rates of earthquakes with magnitude >= 3.5 based on decadal learning periods of a) 1965-1974; b) 1975-1984; c) 1985-1994; d) 1995-2004 and e) 2005-2010. Also shown in f) is the smoothed rate for the entire catalogue from 1965-2010. Note that earthquakes outside the Australian continental area (e.g. New Guinea and the Banda Arc) are not included in this analysis.

Figure 3 shows the probability gain for alternative decadal learning periods when tested on the catalogue from 2005-2010. Probability gains calculated from smoothed seismicity models developed with a minimum magnitude of 4.0 increase with time. There is no clear trend for probability gains calculated for models developed with a minimum magnitude of 3.5, although *G* is larger when using the latter training catalogues and the optimal value of *K*=3. This is suggestive, although by no means definitive, that the spatial distribution of larger earthquakes is influenced by the recent (decadal scale) time history of larger earthquakes. Figure 4 shows smoothed seismicity rates for a minimum magnitude of 4.0 for each decadal period, and the whole catalogue. Again, an increase in model resolution is seen when the full catalogue is used (Figure 4f).

By including a longer record of earthquakes in our learning period we see an increase in probability gain (Figure 5), with the highest probability gain made by including all earthquakes above magnitude 3.5, the full learning period from 1965-2004 and using *K*=3. The probability gain of 4.5 is similar to the value of 4.8 that Helmstetter *et al.* (2007) calculated for their best learning catalogue of earthquakes in California.

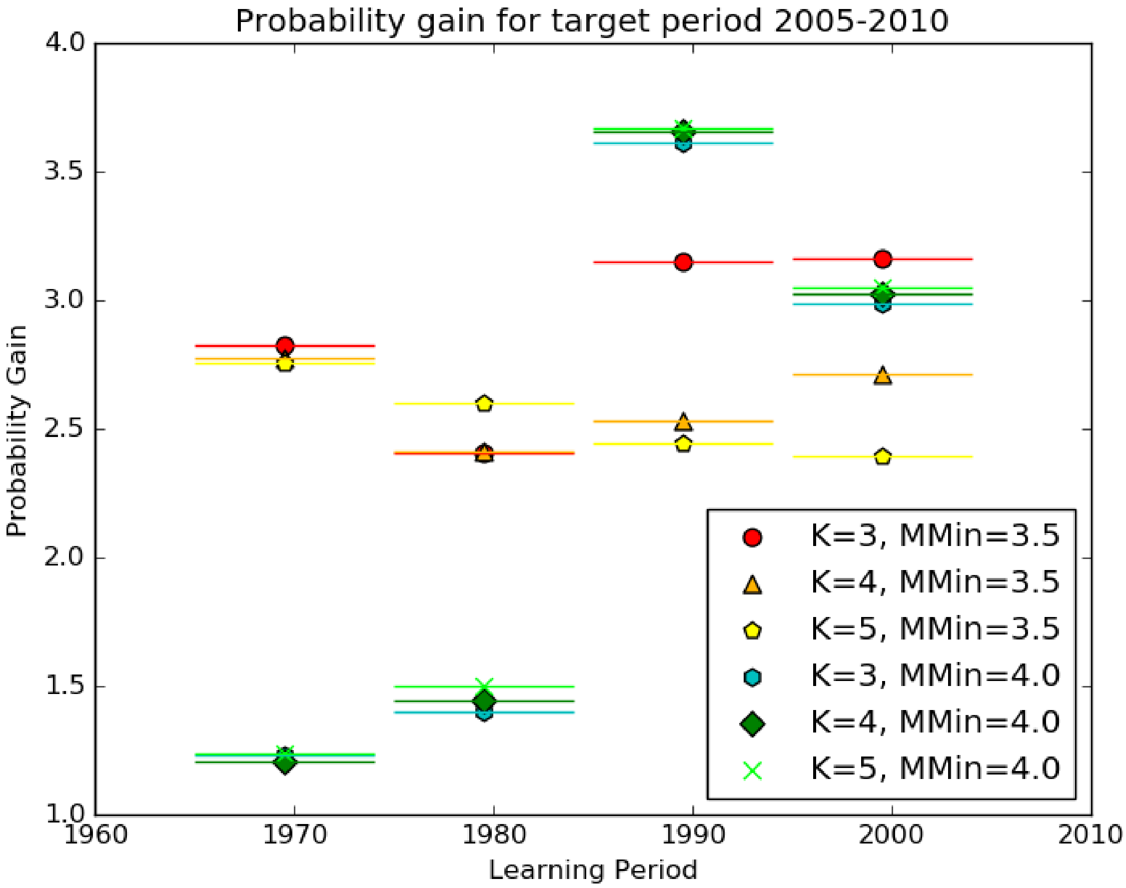

Figure 3: Probability gain for adaptive kernel smoothed seismicity models trained on decadal subsets of the catalogue: 1965-1974; 1975-1984; 1985-1994; 1995-2004 and tested on the period 2005-2010. Horizontal bars show the temporal extent of the learning period and points show the middle of the period. Variations in the smoothing parameter *K* and the minimum magnitude threshold are shown.

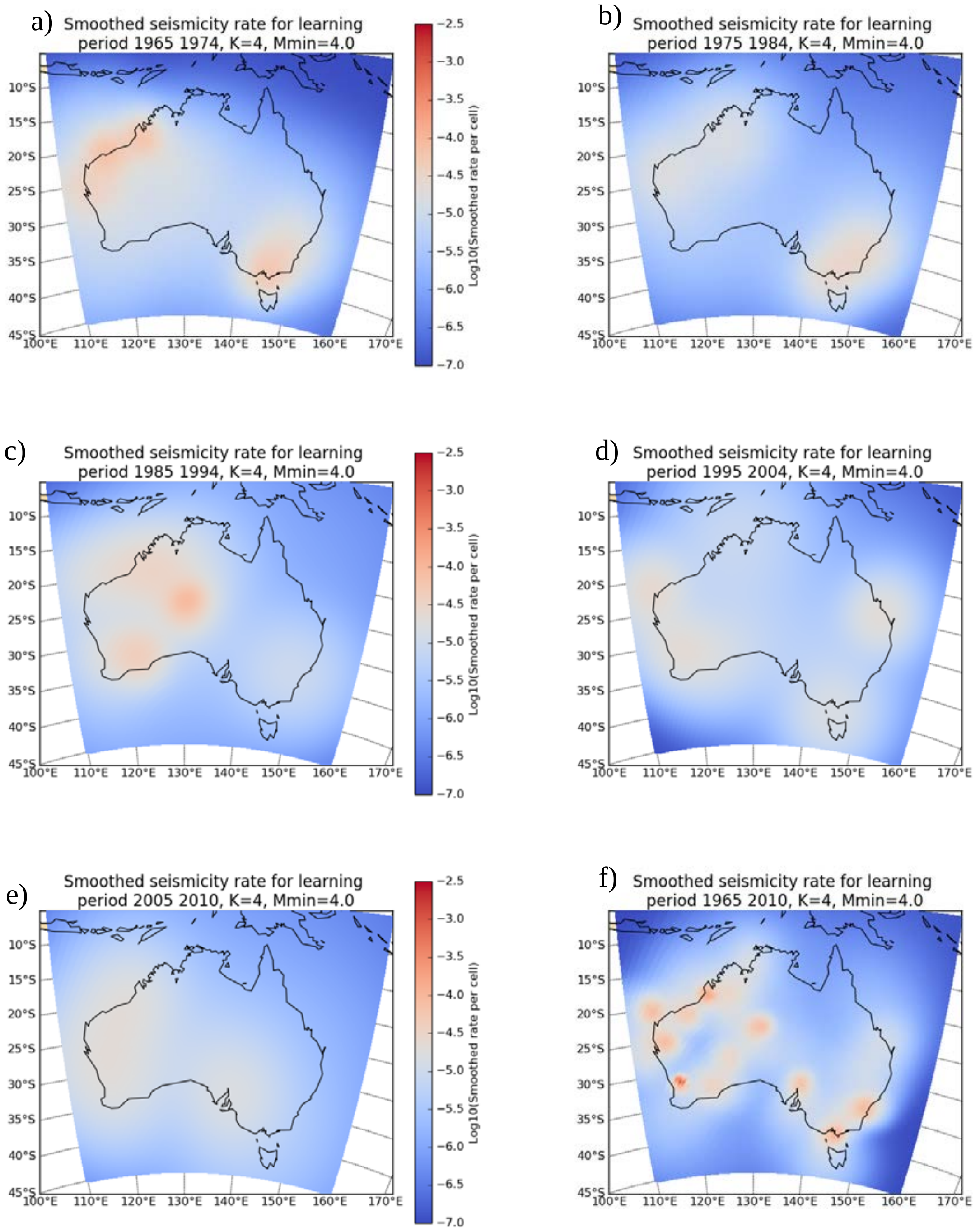

Figure 4: Adaptive kernel (*K*=4) smoothed seismicity rates of earthquakes with magnitude >= 4.0 based on decadal learning periods of a) 1965-1974; b) 1975-1984; c) 1985-1994; d) 1995-2004 and e) 2005-2010. Also shown in f) is the smoothed rate for the entire catalogue from 1965-2010. Note that earthquakes outside the Australian continental area (e.g. New Guinea and the Banda Arc) are not included in this analysis.

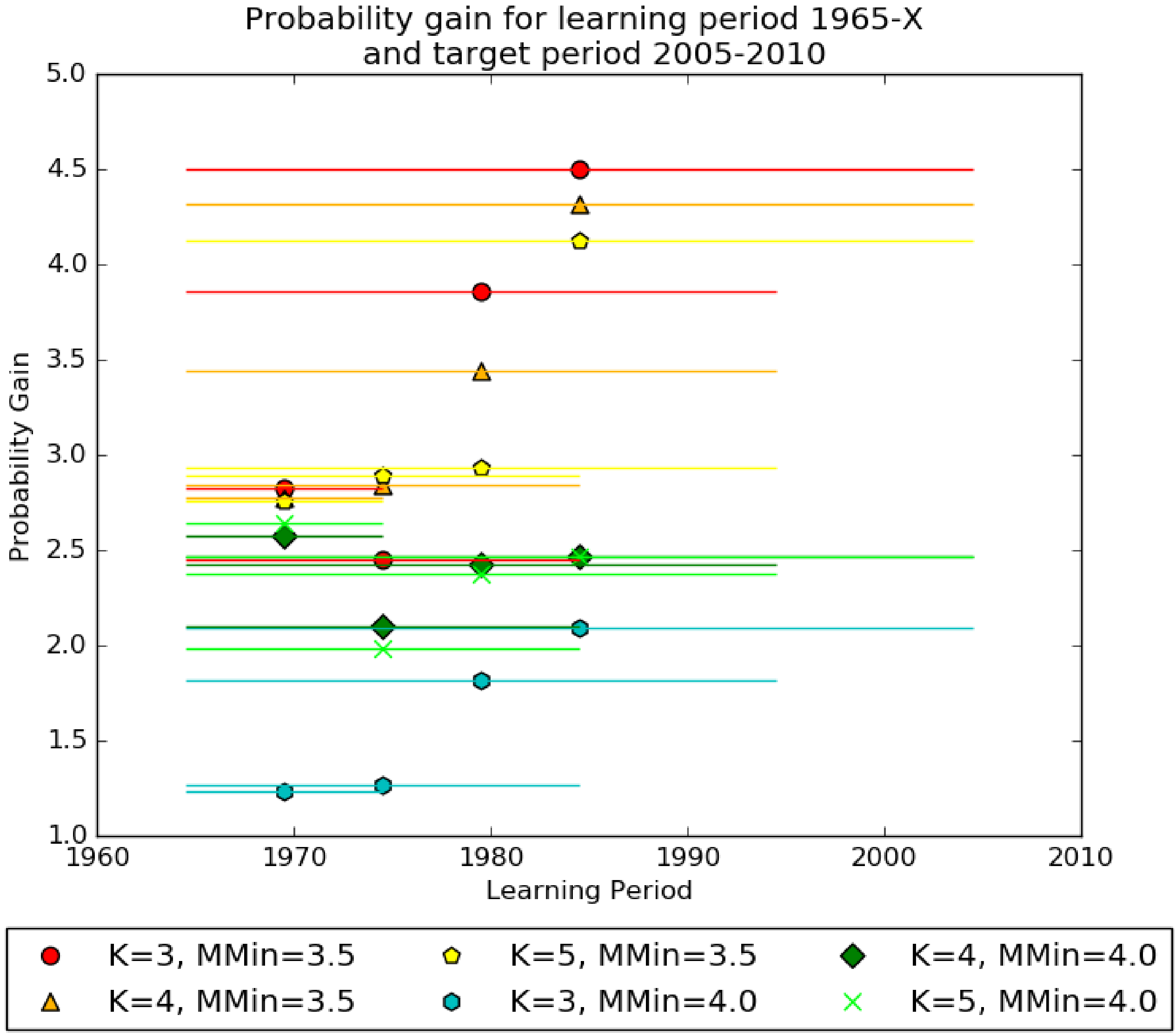

Figure 5: Probability gain for adaptive kernel smoothed seismicity models trained on data from: 1965-1974; 1965-1984; 1965-1994; and 1965-2004 and tested the period 2005-2010. Horizontal bars show the temporal extent of the learning period and points show the middle of the period. Variations in the smoothing parameter *K* and the minimum magnitude threshold are shown.

4) DISCUSSION AND CONCLUSIONS

The results shown here demonstrate that the spatial distribution of past seismicity within stable continental Australia significantly improves forecasts of future seismicity compared with the null hypothesis of a uniform spatial distribution. This result must be limited to temporal time-scales on the order of a few decades based on the mostly complete instrumental catalogue of small earthquakes in Australia (i.e. 1965-present). However, it is a result that is consistent with the non-uniform spatial distribution of topography generating neotectonic features included in the NSHA18 fault source model (Clark *et al.*, 2016).

The results suggest that smoothed seismicity rates are more informative by including as many earthquakes as possible. Including all earthquakes above magnitude 3.5 is more informative than including all earthquakes above magnitude 4.0, even though we do not adjust for catalogue incompleteness of magnitude 3.5 earthquakes pre-1980. Including a greater number of earthquakes increases the spatial resolution of the model by decreasing the distance to the *K*th nearest earthquake, and this increase in resolution in general increases the performance of the model.

Some temporal changes in performance are observed for earthquakes with magnitude greater than 4.0 (Figure 3), which may suggest a decadal-scale temporal dependence in the spatial distribution of larger earthquakes. This may also be explained by uncertainties in the declustering method used on the catalogue, or incompleteness of the early catalogue. Further work is needed to understand whether this is a real manifestation of time-dependent spatial distributions of larger earthquakes.

A limitation of the analysis presented here is that it focuses only on the spatial distribution of earthquake occurrence, and does not consider the magnitude distribution. While the past spatial distribution of magnitude 3.5 and greater earthquakes has the highest probability gain for forecasting rates of all future magnitude 3.5 and greater earthquakes, we have not assessed whether such a model performs better at predicting the spatial distribution of, for example, only magnitude 5.0 or greater earthquakes. Future work will be undertaken in this area.

5) ACKNOWLEDGEMENTS

This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government. This paper was improved by reviews by Gregory Brenn and Hadi Ghasemi. This paper is published with the permission of the CEO, Geoscience Australia.

6) REFERENCES

Burbidge, D.R. (ed), 2012. The 2012 Australian Earthquake Hazard Map, Geoscience Australia Record 2012/71.

Clark, D., Griffin, J., Stirling, M., Leonard, M., Volti, T., 2016. Incorporating fault sources into the Australian National Seismic Hazard Assessment (NSHA) 2018. In: Australian Earthquake Engineering Society 2016 Conference, Nov 25-27, Melbourne, Vic.

Clark, D., McPherson, A., Allen, T., 2014. Intraplate earthquakes in Australia. In: Talwani, P. (Ed.), Intraplate Earthquakes. Cambridge University Press, New York, pp. 8–49.

Clark, D., McPherson, A., Cupper, M., Collins, C.D.N., Nelson, G., 2015. The Cadell Fault, southeastern Australia: a record of temporally clustered morphogenic seismicity in a low-strain intraplate region. Geol. Soc. London, Spec. Publ. SP432.2-.

Clark, D., McPherson, A., Van Dissen, R., 2012. Long-term behaviour of Australian stable continental region (SCR) faults. Tectonophysics 566, 1–30.

Cuthbertson, R.J., 2016. Automatic determination of seismicity rates in Australia. In: Earthquake Engineering Society 2016 Conference, Melbourne, VIC.

Frankel, A., 1995. Mapping Seismic Hazard in the Central and Eastern United States. Seismol. Res. Lett. 66, 8–21.

Ghasemi, H., Allen, T., 2017. Testing the sensitivity of new Australian empirical magnitude conversion equations on seismic hazard analyses. In: Australian Earthquake Engineering Society 2017 Conference, Nov 24-26, Canberra, ACT.

Ghasemi, H., Griffin, J., Heimann, S., Leonard, M., Allen, T., 2016. Towards a homogeneous earthquake catalogue for Australia. In: Australian Earthquake Engineering Society 2016 Conference, Nov 25-27, Melbourne, Vic.

Griffin, J., Volti, T., Clark, D., Ghasemi, H., Leonard, M., Allen, T., 2016. Development of the 2018 Australian National Seismic Hazard Assessment (NSHA). Aust. Earthq. Eng. Soc. 2016 Conf. Nov 25-27, Melbourne, Vic.

Hall, L., Dimer, F., Somerville, P., 2007. A Spatially Distributed Earthquake Source Model for Australia. Proc. 2007 Annu. Meet. Aust. Earthq. Eng. Soc.

Helmstetter, A., Kagan, Y., Jackson, D., 2007. High-resolution Time-independent Grid-based Forecast for $M \geq 5$ Earthquakes in California. Seismol. Res. Lett. 78.

Helmstetter, A., Werner, M.J., 2012. Adaptive spatiotemporal smoothing of seismicity for long-term earthquake forecasts in California. Bull. Seismol. Soc. Am.

Leonard, M., 2008. One Hundred Years of Earthquake Recording in Australia. Bull. Seismol. Soc. Am. 98, 1458–1470.

Leonard, M., Burbidge, D.R., Allen, T.I., Robinson, D.J., McPherson, A., Clark, D., Collins, C.D.N., 2014. The Challenges of Probabilistic Seismic-Hazard Assessment in Stable Continental Interiors: An Australian Example. Bull. Seismol. Soc. Am. 104, 3008–3028.

Stein, S., Liu, M., 2009. Long aftershock sequences within continents and implications for earthquake hazard assessment. Nature 462, 87–89.