

Statistical models for temporal variations of seismicity parameters to forecast seismicity rates in Japan

Christine Smyth and Jim Mori

Disaster Prevention Research Institute, Kyoto University, Japan

(Received March 30, 2010; Revised October 5, 2010; Accepted October 5, 2010; Online published March 4, 2011)

This paper introduces a model to forecast the rate of earthquakes for a specified period and area. The model explicitly predicts the number of earthquakes and b -value of the Gutenberg-Richter distribution for the period of interest with an autoregressive process. The model also incorporates a time dependency adjustment for higher magnitude ranges, assuming that as time passes since the last large earthquake within the area, the probability of another larger earthquake increases. These predictions are overlaid on a spatial density map obtained with a multivariate normal mixture model of the historical earthquakes that have occurred in the area. This forecast model differs from currently proposed models by its density estimation and its assumption of temporal changes. The model has been submitted to the Earthquake Forecasting Testing Experiment for Japan.

Key words: Forecast, autoregressive models, normal mixture models.

1. Introduction

The Collaboratory for the Study of Earthquake Predictability (CSEP) (<http://www.cseptesting.org/home>) is an initiative to test earthquake forecast models in a fair environment. In collaboration with CSEP, the Earthquake Forecasting Testing Experiment for Japan is focused on evaluating models that forecast the seismicity of Japan (Research group “Earthquake Forecast System based on Seismicity of Japan”, 2009). Submissions to the Japanese experiment, identical to the CSEP, require a forecast of the number of earthquakes to occur in specified $0.1^\circ \times 0.1^\circ$ spatial bins over a particular region and time. The forecast rate for a single spatial bin must be divided into rates for specific magnitude bins over a predetermined magnitude range. The model described herein has been submitted to the Japanese initiative.

To create our forecast, we initially take a subregion of the entire forecast area, comprised of many of the aforementioned spatial bins. We consider subregions so that we can pick up local variations in seismicity rate, and force the algorithm to find pockets of seismicity that otherwise might have been ignored if the entire area was studied at once. We consider various facets of the past seismicity in this subregion. We look at the previous overall seismicity rate; the temporal evolution of the proportion of large versus small earthquakes via the b value of the Gutenberg-Richter distribution (Gutenberg and Richter, 1944); and the location of the previous earthquakes. Then, this information is converted into a forecast via an autoregressive process and a multivariate normal mixture model (Everitt, 1993; Chatfield, 2004). The autoregressive process is used to ex-

trapolate the rate of past seismicity to the forecast period. The mixture model is used to convert the locations of past earthquakes into a spatial density map of the subregion. To generate our forecast for a single spatial bin, we multiply the normalised density of the spatial bin obtained with the mixture model by the predicted number of earthquakes obtained with the autoregressive procedure. In this manner, spatial bins within the subregion with high rates of past seismicity will be assigned larger predictions than bins where past seismicity rates are low. We repeat this process for each subregion of the forecast area until we have a forecast for the entire region.

The forecast model differs from currently proposed models by its assumption of non-stationarity and its density estimation technique. We also incorporate an optional time dependency component in our model by assuming that as time passes since the last large earthquake within a subregion, the probability of another large earthquake increases. This requires the knowledge of the repeat times of earthquakes within an area. We cannot estimate the repeat times empirically owing to the long history required, so here we use a simulation approach based on the available historical data. The forecast model can be run with or without this adjustment. We describe the base algorithm (MARFS) and its optional adjustment (MARFSTA) in the Methods section.

2. Methods

The overall algorithm used to create a forecast is as follows:

- 1) Divide the entire forecast region into smaller subregions. For ease, we divide the entire forecast region into rectangular subregions. We endeavour to minimize the total number of subregions for computational reasons, whilst ensuring that each subregion is small enough so that the multivariate normal mixture model

can be applied. We also ensure that each subregion includes enough historical earthquakes to reliably calculate the Gutenberg-Richter parameters.

- 2) Consider one smaller subregion at a time.
 - (a) Calculate a spatial density map of the previous earthquakes in the area using a multivariate normal mixture model.
 - (b) Predict the parameters of the Gutenberg-Richter distribution for the next period.
 - (c) Multiply the predicted rate of earthquakes with the density of each spatial bin.
 - (d) If desired, obtain the time dependent adjusted rates of larger earthquakes for each spatial and magnitude bin.

If the forecast region is small enough, it is not necessary to divide it into subregions. In this situation, we simply commence the algorithm from step 2(a). However, the forecast region specified by the Earthquake Forecasting Testing Experiment for Japan committee is large, and so we choose to divide the region into 25 subregions in order to force the algorithm to find local variations in seismicity rate. We note here we consider earthquakes only within the depth range specified by the Experiment committee. We now explain each of these steps in detail in the context of a yearly forecast, assuming we have already divided the forecast region into smaller subregions. Extensions to other forecast periods are obvious. The algorithm has been coded with the R software package (<http://www.r-project.org/>).

2.1 Spatial density map

The mixture model is a well-known clustering technique which assumes that the data (here the latitude and longitude of earthquakes) come from a probability density function taken to be a mixture of G component density functions in unknown proportions (Everitt, 1993). The mixture model finds G mean vectors that best describe the data, and by estimating G covariance matrices, the entire data space can be represented as a density:

$$f(\mathbf{y}, \boldsymbol{\theta}) = \sum_{g=1}^G \delta_g \Phi(\mathbf{y}, \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g), \quad 0 \leq \delta_g \leq 1, \quad \sum_{g=1}^G \delta_g = 1 \quad (1)$$

where δ_g are the mixing proportions; and $\boldsymbol{\theta}$ is the vector of unknown parameters $\delta_g, \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g, g = 1, \dots, G$. The component density functions are multivariate normal so:

$$\Phi(\mathbf{y}, \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g) = \frac{1}{2\pi^{d/2} |\boldsymbol{\Sigma}_g|^{-1/2}} \exp(-1/2(\mathbf{y} - \boldsymbol{\mu}_g)^T \cdot \boldsymbol{\Sigma}_g^{-1} (\mathbf{y} - \boldsymbol{\mu}_g)) \quad (2)$$

where d is the number of variables; $\boldsymbol{\mu}_g$ is the g^{th} component mean; and $\boldsymbol{\Sigma}_g$ is the g^{th} component covariance matrix. The parameters are calculated using the EM algorithm (Dempster *et al.*, 1977) and the reader is directed to McLachlan and Ng (2009) for more information about the EM algorithm.

For illustrative purposes, we consider the area defined by the vertices shown in Table 1. We refer to this area as the

Table 1. Vertices of Tamba area.

Longitude	Latitude
134.7391E	34.3691N
136.2609E	34.3691N
134.7272E	35.6310N
136.2728E	35.6310N

Tamba area. We randomly sample 500 earthquakes from the Tamba area and cluster these. Figure 1(a) shows the six clusters that the model has found. Each earthquake has been assigned to the cluster for which it has the highest probability of belonging. Although the cluster shapes were determined with 500 earthquakes, we show only 100 earthquakes in the figure for visual clarity. The ellipsoids show the standard deviation of each cluster. The log transformed density of the space, as obtained by the mixture model in Eq. (1) is depicted in Fig. 1(b). Areas of high density correspond to the tight and compact clusters of earthquakes shown in Fig. 1(a).

The mixture model ensures that the density for all points in the space is non-zero. It therefore differs from the average approach sometimes used, which simply obtains a density of any bin in the area by dividing the number of earthquakes in that bin over time by the total number of recorded earthquakes. The mixture model also allows for a smooth transition amongst neighbouring bins. We normalize the density map, so that the sum of the density of each spatial bin of interest is one.

2.2 Gutenberg-Richter parameters

The Gutenberg-Richter distribution (Gutenberg and Richter, 1944) plays a major role in earthquake forecasting and hazards analysis and quantifies the linear relationship between the frequencies of earthquakes, N , and their sizes, M as:

$$\log_{10} N = a - bM. \quad (3)$$

The Gutenberg-Richter distribution indicates how many earthquakes can be expected in some time period for a given region after estimation of its parameters. Maximum likelihood implies that the parameter b is given by:

$$b = \frac{\log_{10} e}{(\bar{M} - M_L)} \quad (4)$$

where \bar{M} is the mean magnitude of the sample of earthquakes; and M_L is a function of M_c (the magnitude of completeness of the data) (Guo and Ogata, 1997). The spatial variability of a and b has been well studied and it is generally agreed that the parameters vary spatially (Wiemer and Wyss, 2002; Schorlemmer *et al.*, 2004). The temporal variability of these parameters has not been studied as rigorously as it is suggested that the variation of these parameters will average out over the long term (Schorlemmer *et al.*, 2004). However, for short term studies, these parameters do vary, and models with temporally variant parameters fit the data better than temporally invariant models (Smyth and Mori, 2009). We exploit the temporal variations of these parameters in this forecast model using an autoregressive process.

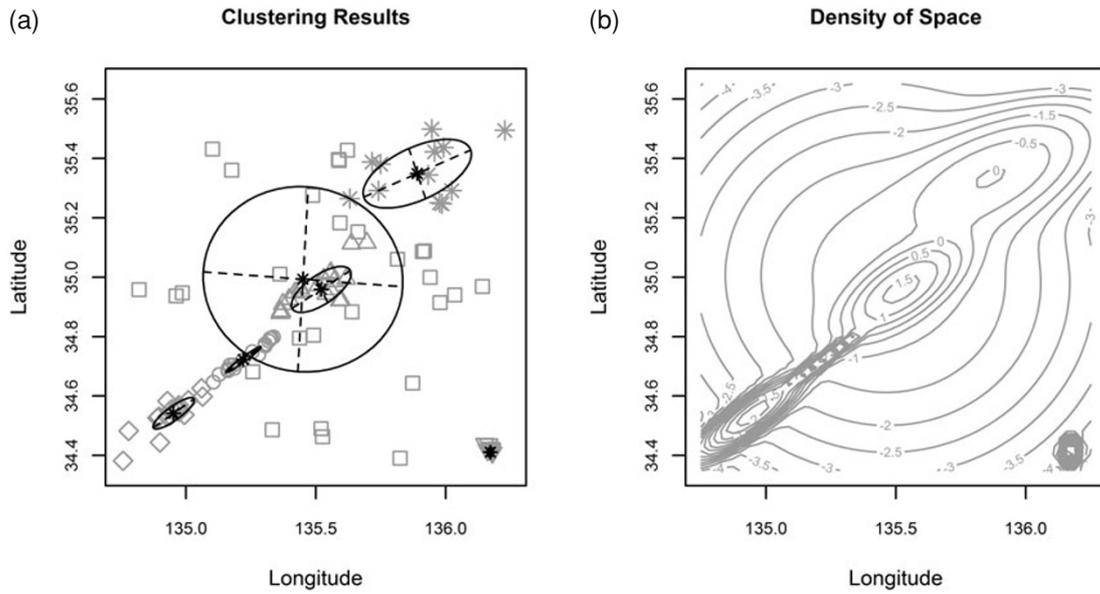


Fig. 1. Multivariate normal mixture model clustering results of the Tamba area. Figure 1(a): The left picture shows the locations of historical earthquakes. The different symbols, squares, circles, triangles, inverted triangles, asterisks and diamonds show the assignment of each earthquake to one of the six clusters. The standard deviation of each cluster is shown by an ellipsoid. The mean of each cluster is shown as a cross. Figure 1(b): The right picture shows the log transformed density of the area as estimated by Eq. (1). The contours represent increasing density.

We obtain the b value for each year, b^t , and count the number of earthquakes above the magnitude of completeness of the data for each year, N^t . The parameter N^t refers now to a simple count of the number of earthquakes in year t . To predict the rate of earthquakes next year, we apply a autoregressive process to these b^t and N^t values. Autoregressive models are used frequently with discrete time series data, when we have measurements of a random variable at equispaced points in time. The variable is regressed on its own previous values, elucidating the prefix ‘auto’ and thereby an autoregressive model assumes that the value of the variable at any time is linearly dependent on its most recent values (Chatfield, 2004):

$$b^T = \sum_{k=1}^p \alpha_k b^{T-k} + \epsilon^T \quad (5)$$

where α is the vector of predictor coefficients; and b^T represents the b value of the year prior to that which we are trying to forecast (the most recently observed value). We use the Akaike Information Criterion to obtain a reliable estimate of p (Akaike, 1974). The vector α is estimated with the Yule-Walker equations:

$$\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ \vdots \\ r_p \end{pmatrix} = \begin{pmatrix} 1 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & 1 & r_1 & \cdots & r_{p-2} \\ r_2 & r_1 & 1 & \cdots & r_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & \cdots & \cdots & 1 \end{pmatrix} \times \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \vdots \\ \alpha_p \end{pmatrix} \quad (6)$$

where r_i is the correlation between observations that are i steps apart (Chatfield, 2004).

After estimating α , we use Eq. (5) directly to obtain the one-ahead forecast by substituting T with $T + 1$. The

autoregressive model is thereby used to predict the next year’s \hat{b}^{T+1} and \hat{N}^{T+1} values. We believe this approach should enable us to model changes in seismicity rates or magnitude distributions.

2.3 Unadjusted earthquakes rates

We obtain a prediction for each spatial bin (indexed by i), \hat{N}_i^{T+1} , within the subregion by multiplying \hat{N}^{T+1} by the density of each bin:

$$\hat{N}_i^{T+1} = f(\mathbf{y}_i, \theta) \times \hat{N}^{T+1} \quad (7)$$

where $f(\mathbf{y}_i, \theta)$ is the density at the midpoint of the i^{th} spatial bin, \mathbf{y}_i , given by the multivariate normal mixture model. To discretize the \hat{N}_i^{T+1} (the total number of earthquakes expected in the bin), into rates for each magnitude bin, we scale by \hat{b}^{T+1} . The predictions obtained to this point are referred to as the MARFS (Multivariate AutoRegressive Forecast of Seismicity) predictions.

2.4 Adjusted earthquake rates

We assume that as time passes since the last large earthquake, the probability of another large earthquake increases. Here, we define a large earthquake for illustrative purposes as $M \geq 5$, although this definition is not stringent. The model proposes a rate adjustment for higher magnitude $M \geq 5$ ranges. Such an adjustment usually requires the knowledge of the repeat time of earthquakes. As the history available to calculate this statistic is quite short, we propose a simulation approach. Firstly, the mean \bar{b} , \bar{a} , and \bar{N} are calculated over all years:

$$\bar{b} = \frac{\sum_{t=1}^T b^t}{T} \quad (8)$$

where b^t represents the b value of the general t^{th} year and b^T represents the b value of the year prior to that which we

are trying to predict. We obtain the Poisson probability for having 0, 1, 2, 3, ... earthquakes greater than $M \geq 5$ using these average numbers. The probability estimate is robust as we are using all data greater than the magnitude of completeness for its calculation. If we were to only consider the previous $M \geq 5$ earthquakes which had occurred within the subregion, this value would not be certain. We then simulate 1000 years of data, where the number of $M \geq 5$ earthquakes in any year is determined by these probabilities. Using the simulated data, we obtain simulated recurrence times of earthquakes, and fit a logistic distribution to these times. We fit a logistic distribution, rather than a normal distribution as the logistic curve has more kurtosis than the normal curve and therefore there is slightly more probability of earthquakes occurring further from the mean repeat time.

We obtain the initial adjustment factor as:

$$\text{AF} = \frac{P(t^*) - P(t^* - 1)}{1 - P(t^* - 1)} \quad (9)$$

where t^* is the number of years since the last $M \geq 5$ earthquake; and $P(t^*)$ is the value of the cumulative distribution function at t^* of the aforementioned logistic distribution (Stein and Wysession, 2003). This initial adjustment factor is then divided by the Poissonian probability of having one or more earthquakes in the same period. The rates of all bins predicting $M \geq 5$ earthquakes are multiplied by this adjustment factor. If there has been no $M \geq 5$ earthquake within the subregion, no adjustment is made. Forecasts which include this adjustment are called the MARFSTA (Multivariate AutoRegressive Forecasts of Seismicity with a Time Adjustment) predictions.

3. Assessing Model Predictions

An important and necessary step in the introduction of any new model is to illustrate its ability. As this model will be compared to other forecast techniques within the Earthquake Forecasting Testing Experiment for Japan initiative, we do not perform benchmark comparisons here. However, we show that the model is indeed valid, by first presenting the predictions for the Tamba area, and then presenting plots of the entire Japanese forecast region. We show the location of observed earthquakes during the forecast period.

We also present the Receiver Operating Characteristic (ROC) curves of the predictions. The ROC curves plot the true positive rate versus the false positive rate. The true positive rate is given by:

$$\frac{\gamma}{\gamma + \vartheta} \quad (10)$$

and the false positive rate is given by:

$$\frac{\zeta}{\zeta + \omega} \quad (11)$$

where the values of γ , ϑ , ζ , ω are explained in Table 2. We assume a very small alarm rate, then all bins with predicted rates less than this alarm rate where an earthquake occurs are counted and give the value ϑ in Table 2. Similarly,

Table 2. Counts of forecast and observed earthquakes over all $0.1^\circ \times 0.1^\circ$ bins.

		Observed	
		Yes	No
Forecast	Yes	γ	ζ
	No	ϑ	ω

all bins with predicted rates greater than this alarm rate where an earthquake occurs are counted and give the value γ in Table 2. Similar reasoning holds for ζ and ω . Then we increase the alarm rate and repeat this process. Each alarm rate gives a point on the ROC curve. This graphical technique compares to a random prediction. The reader is directed to Fawcett (2006) and Murru *et al.* (2009) for a detailed explanation of the ROC curve validation technique. We report the results of these tests in the following Results section.

4. Results

4.1 Tamba area of Japan

We illustrate the method on data obtained from the small Tamba area defined by the vertices in Table 1. We use this area for illustrative purposes only: it does not constitute one of the 25 rectangular subregions we use to create our entire Japan forecast. If we were to use a subregion as small as the Tamba area, the algorithm would be prohibitively long to run. We have high quality data for this area from January 1976 to December 2007 inclusive (Hiroshi Katao, 2008, personal communication). The dataset is compiled by the Disaster Prevention Research Institute (DPRI), Kyoto University. The hypocenters have been determined by DPRI using combined data from the Japan Meteorological Agency, High Sensitivity Seismograph Network Japan (Hinet) (<http://www.hinet.bosai.go.jp/>), and DPRI stations. We have considered the Tamba area in other work, where it was shown that the Gutenberg-Richter parameters of the Tamba area are temporally variant (Smyth and Mori, 2009). Therefore, we use this area to illustrate the forecast method for the year January through December 1995 and stress that the predictions were obtained using only information available up to the 31st of December, 1994. They are retrospective predictions only in the sense that we are not waiting for validation of our results: we already have the data with which to test the model. We removed all data less than $M 2.5$. We trialled various cut off values with the all Japan forecasts and found the best cut off values (with retrospective testing) for the different forecast classes. The smallest successful value for mainland Japan was 2.5. As the Tamba area is mainland Japan, we chose to use 2.5 as our cut off magnitude.

Figure 2(a) shows the predicted rate of all $M \geq 2.5$ earthquakes for each spatial bin within the Tamba area. The scale is given on the right. We can see the predictions clustering along a diagonal, and the maximum prediction of any bin is approximately 3.5 earthquakes. The rate of observed earthquakes is given in the figure directly to the right. We can see the earthquakes cluster along the diagonal, however slightly further south west than predicted. The number of earthquakes is also by far larger, with over 200 earthquakes

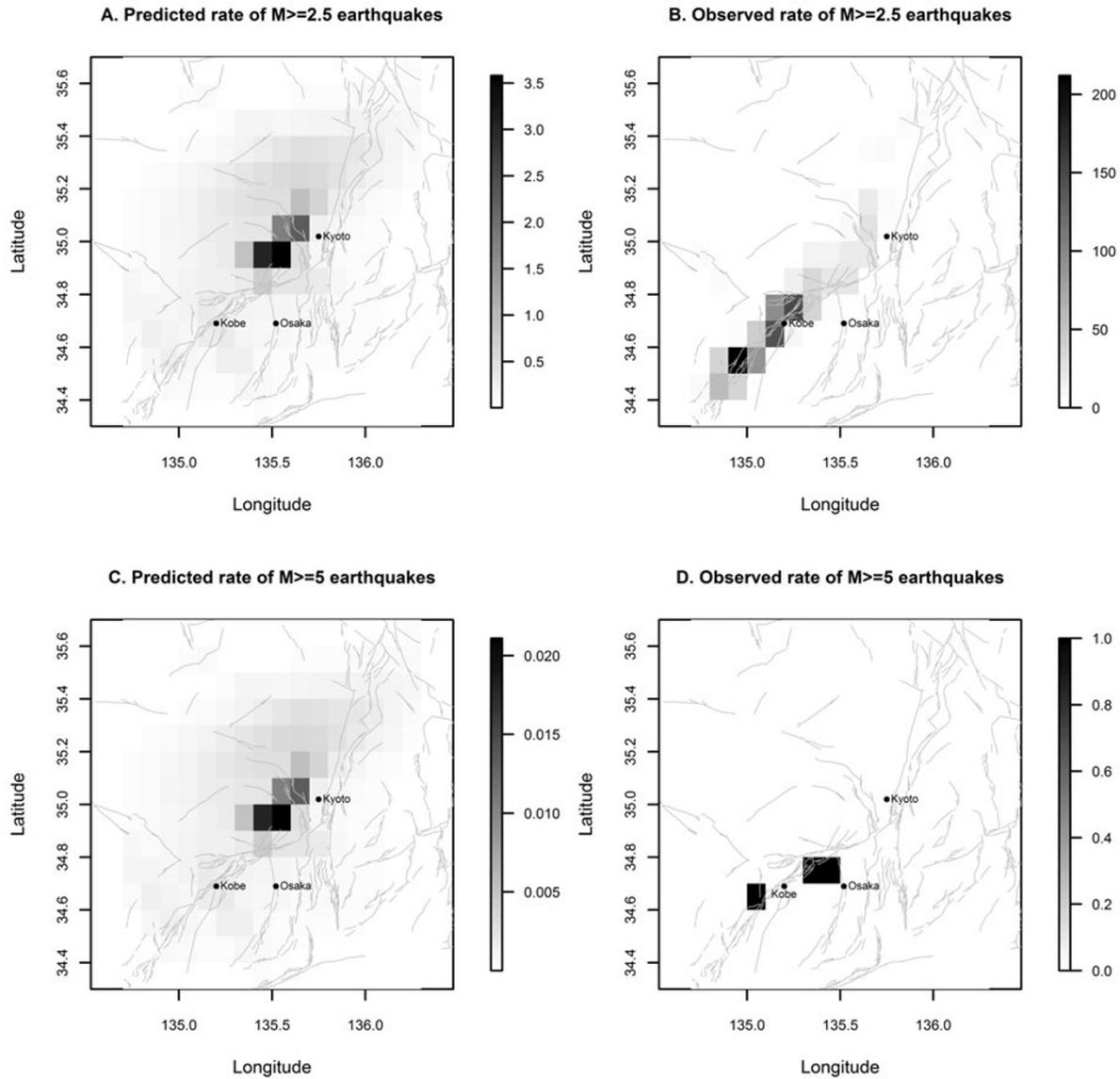


Fig. 2. Comparison of predictions and observed seismicity for the Tamba area in 1995. The scale shows the number of expected or observed earthquakes at each point. We have also included the fault segments within the Tamba area obtained from the active fault database of Japan (available at <http://riodb02.ibase.aist.go.jp/activefault/>).

occurring in one bin. Those familiar with the history of Japanese earthquakes will remember 1995 as the year of the Hyogo-ken Nambu earthquake (Kobe earthquake). The Kobe earthquake, $M_{JMA} = 7.3$, (magnitude as determined by the Japan Meteorological Agency) was located within the Tamba area, and was followed by many aftershocks also located within the area. Obviously, the forecast rate of earthquakes was far less than the observed number of earthquakes for this year. We remind the readers that the forecast is finalized with data until December 31st 1994. The Kobe earthquake occurred on January 17th 1995. If we were to update the forecast on January 31st, obviously we would have predicted a higher seismicity rate. In this regard, these algorithms can be used with seismicity data that include aftershocks; however the algorithm must be updated after the main shock. Figure 2(c) shows the number of $M \geq 5$ earthquakes predicted. The spatial pattern of earthquakes in Fig. 2(c) is the same as Fig. 2(a): only the forecast rate

decreased. Figure 2(d) shows the actual location of those observed earthquakes, including the Kobe earthquake main shock.

When we incorporate a time dependent adjustment for greater than $M \geq 5$, the forecast rate of $M \geq 5$ earthquakes slightly changes, however the overall spatial pattern does not. The adjusted rate is less than half that of the unadjusted rate. At this point there had been four $M \geq 5$ earthquakes in the dataset. Two earthquakes had occurred in 1985, ten years from the start of the catalogue, one earthquake occurred in 1987 and one earthquake occurred in 1992. The repeat times are therefore 10, 0, 2 and 5 years. The mean repeat time is being simulated as larger than three years. Hence, the chance of a greater than $M \geq 5$ earthquake was reduced by the adjustment factor as the last $M \geq 5$ earthquake had occurred only three years previously in 1992. This highlights the potential pitfalls of this approach. We only have a handful of events greater than $M \geq 5$ within

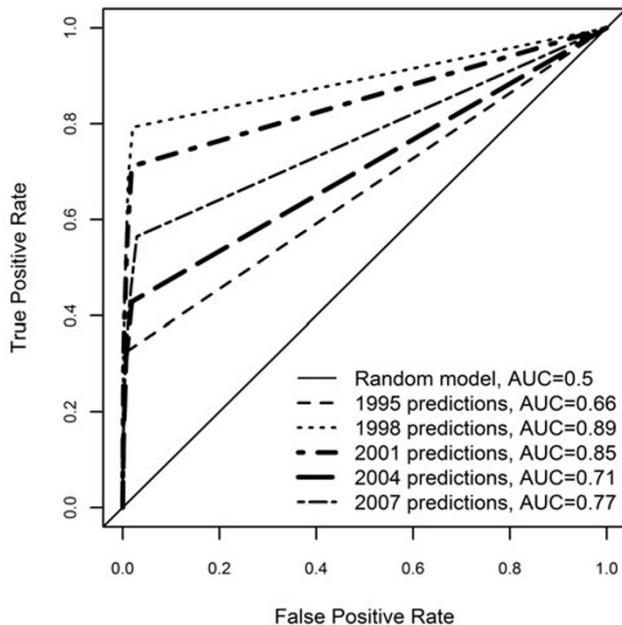


Fig. 3. Receiver Operating Characteristic (ROC) curves showing fit of the predicted seismicity in the Tamba area for various years.

this area. It is difficult to calculate reliable repeat times and increasing and decreasing trends until we have enough events in the data. When we forecast rates for all Japan, our data history is longer and our subregions are larger, and thereby we can increase our history of $M \geq 5$ events.

The ROC curves are presented in Fig. 3. The true positive rate corresponds to the fraction of cells with earthquakes that are correctly preceded by an alarm. The false positive rate give the fraction of cells without earthquakes that are incorrectly preceded by an alarm. So, the larger the area under the curve (AUC), the better the predictions. Here, we present results for 5 years; 1995, 1998, 2001, 2004 and 2007. The graphs are obtained using all the earthquakes $M \geq 2.5$ available in the catalog. As expected, the worst performing year is 1995. After this year, the AUC is greater than 0.7, and obtains a maximum value of 0.89 in 1998. This graph shows that the technique is doing far better than a random guess (diagonal line), particularly as more data are made available.

The observed events in 1995 directly influence predictions for the following years. For the immediate years, the density moves south west along the diagonal and the adjustment factor for a $M \geq 5$ earthquake is less than 1, implying that as we have recently had an earthquake greater than $M \geq 5$, the probability of another should be scaled down. Slowly, activity along the off-diagonal forces the density into a more circular shape and the adjustment factor for the probability of an $M \geq 5$ earthquake edges above 1.

In this section we looked at the small Tamba area of Japan to illustrate the model. However, in order to submit our model we must forecast rates for all Japan. To apply this model to all Japan, we repeatedly take subregions, treat each subregion individually, run the algorithm, and append the forecasts together. We illustrate an entire Japan forecast in the following section.

4.2 Entire Japanese forecast area

We present results of the model applied to the entire Japan forecast region. We generate a forecast of the number of $M \geq 5$ earthquakes for February 2009 to January 2010 inclusive. The area has been specified by the Earthquake Forecasting Testing Experiment for Japan committee. We have taken the log transform of the predicted total rate of $M \geq 5$ earthquakes per bin and plotted these rates on the forecast map shown in Fig. 4. For clarity, we do not plot any rates less than 0.001 earthquakes per year. The scale is shown on the right hand side of the plot. The darker the color, the higher the forecast rate of earthquakes. Squares show the location of $M \geq 5$ earthquakes that were observed during this period. The figure shows that recent large earthquakes in Japan usually occur in regions with high forecast rate. We forecast a total 77.52 $M \geq 5$ (MARFS) and 77.17 (MAFSTA) earthquakes for the period. We observed approximately 64 earthquakes during this time.

In 2008, the Iwate-Miyagi Nairiku earthquake struck Iwate prefecture, northern Honshu, $M_{JMA} = 7.2$ (epicentre 39.0283N,140.88E). The aftershock sequence of this earthquake induces a high forecast rate within the immediate area, visible in Fig. 4. Figures 2 and 4 show us that if there is a very large main shock, its aftershock sequence will completely dominate the spatial density of the area. Although this may be a realistic scenario, where earthquakes are predicted in the same area for many years to come, it may be necessary to move the density estimation away from this position. This could be achieved by taking only the last 10 years of data to obtain the density, or by using some form of random jitter. Furthermore, a large main shock and its associated aftershock sequence will induce an elevated forecast rate for the future years in the immediate vicinity. If we were trying to forecast only independent events, our model would severely over-predict in this situation. However, the forecast experiment requires a forecast of the total number of earthquakes during a time period, and does not distinguish between main shocks and aftershocks. Therefore, we do not try and alter the resulting spatial density or forecast rate following a large main shock.

5. Discussion and Conclusions

In conclusion, this forecast model will consider earthquakes more likely in areas where they have already occurred via the mixture model. There will be a gradual slope in density across neighbouring bins. The mixture model should produce areas of high density coincident with previous seismicity. The autoregressive model will pick up any changes in rate and magnitude distributions. As the data increase over time, it may be appropriate to include trend or seasonality analysis, or even some more complicated time series modelling. At this point in time, we use the simplest model possible, owing to lack of data. Overlaying the mixture model with the autoregressive model gives a spatially and temporally variant forecast of seismicity.

We also introduced a time dependency adjustment factor for large magnitude ranges. Time dependency models have been advocated in the literature, for example see Petersen *et al.* (2007). It is only natural to assume that large earthquakes become more likely as time passes. Here

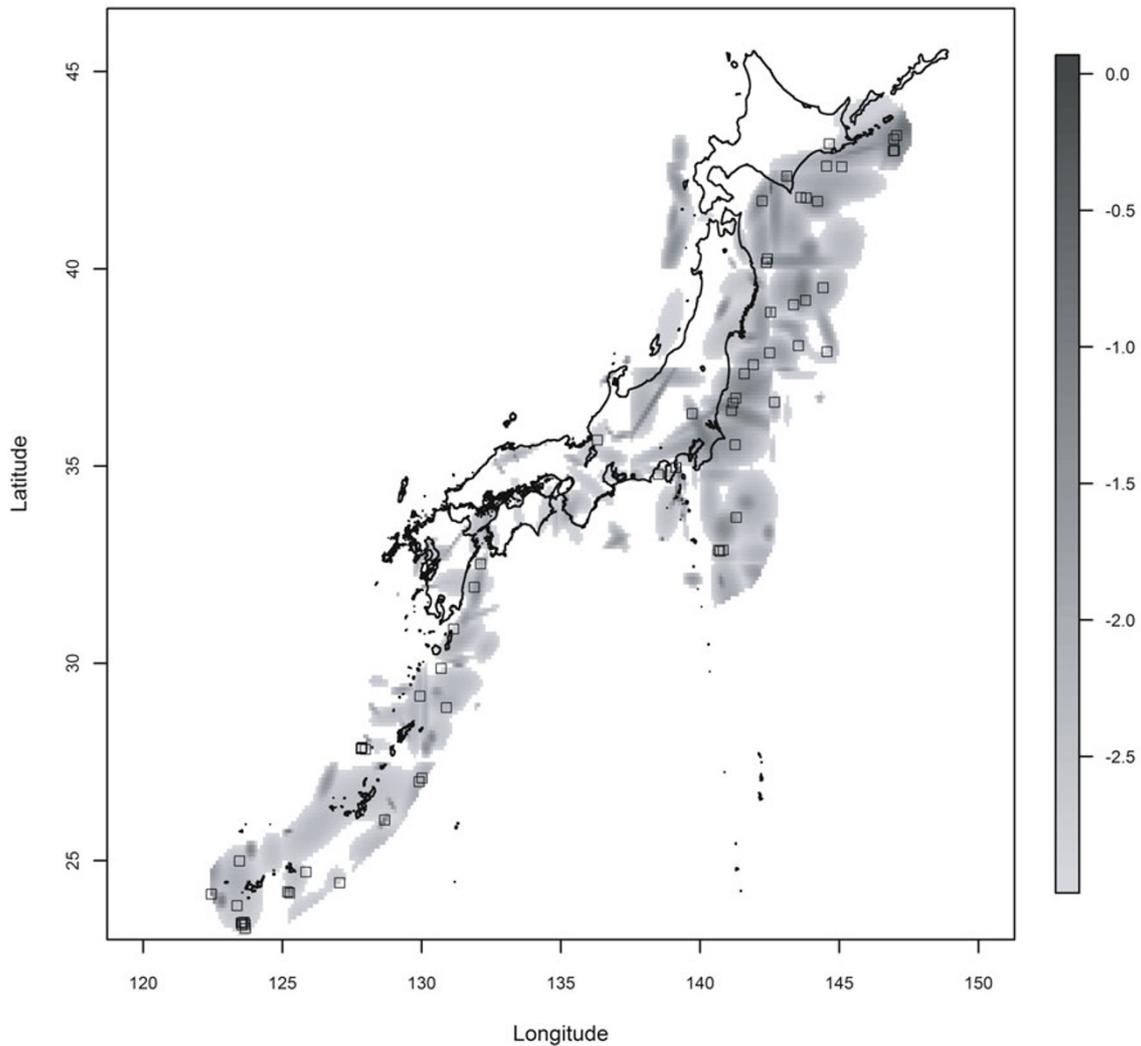


Fig. 4. Predicted (log) number of $M \geq 5$ earthquakes for the one year period from February 2009 through January 2010. Locations of $M \geq 5$ earthquakes that occurred during this period are marked with a square.

we scaled the probabilities based on the time since the last $M \geq 5$ earthquake. This is not a realistic cut off for areas where $M \geq 5$ earthquakes are a yearly occurrence. For regions that are prone to larger earthquakes we suggest using a higher magnitude cut off for the scaling probability. It is also possible that the scaling probability be adjusted manually. If the researcher had reason to believe (based on more physical models) that there was to be a $M \geq 8$ earthquake imminently, the adjustment factor could be increased to reflect this expert knowledge. The researcher could also, within a particularly seismically active region, have different adjustment factors for each of $M 5$, $M 6$, $M 7$ and $M 8$ bins.

The overall product of our research is the earthquake forecast algorithms, MARFS and MARFSTA, ready for real time testing. Our model differs from currently proposed models by its density estimation technique and by the inclusion of potential temporal changes, often ignored, within the Gutenberg Richter distribution. We also incorporate a further time dependency component in MARFSTA, by assuming that as time passes since the last large earthquake, the probability of another large earthquake increases. The

models described here are submitted to the Earthquake Forecasting Testing Experiment for Japan and are undergoing testing against other well known models in a prospective environment to ascertain which submitted model best forecasts the seismicity of Japan. We look forward with interest to the results of the testing experiment over the coming years, and the subsequent increased understanding of the physics and statistics of earthquake occurrence.

Acknowledgments. Christine Smyth is the recipient of a Japan Society for the Promotion of Science Postdoctoral Fellowship. We gratefully acknowledge the Japanese Meteorological Agency and Hiroshi Katao for providing the necessary data used in this publication, and the Earthquake Research Institute at Tokyo University for hosting the test center. We also thank the National Institute of Advanced Industrial Science and Technology for making the fault database of Japan publicly available.

References

- Akaike, H., New look at statistical-model identification, *IEEE Trans. Automatic Control*, **19**, 716–723, 1974.
- Chatfield, C., *The Analysis of Time Series*, Chapman and Hall, Florida, 2004.
- Dempster, A. P., N. M. Laird, and D. B. Rubin, Maximum likelihood from

- incomplete data via EM algorithm, *J. Roy. Stat. Soc. B. Stat. Meth.*, **1**, 1–38, 1977.
- Everitt, B. S., *Cluster Analysis*, Edward Arnold, London, 1993.
- Fawcett, T., An introduction to ROC analysis, *Pattern Recognit. Lett.*, **27**, 861–874, 2006.
- Guo, Z. and Y. Ogata, Statistical relations between the parameters of aftershocks in time, space and magnitude, *J. Geophys. Res.*, **102**, 2857–2873, 1997.
- Gutenberg, B. and C. F. Richter, Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.*, **34**, 185–188, 1944.
- McLachlan, G. and S. Ng, The EM algorithm, in *The Top-Ten Algorithms in Data Mining*, edited by X. Wu and V. Kumar, 93–115 pp, Chapman and Hall/CRC, Boca Raton, Florida, 2009.
- Murru, M., R. Console, and G. Falcone, Real time earthquake forecasting in Italy, *Tectonophysics*, **470**, 214–223, 2009.
- Petersen, M. D., T. Q. Cao, K. W. Campbell, and A. D. Frankel, Time-independent and time-dependent seismic hazard assessment for the State of California: Uniform California earthquake rupture forecast model 1.0, *Seismol. Res. Lett.*, **78**, 99–109, 2007.
- Research group “Earthquake Forecast System based on Seismicity of Japan” (K. Z. Nanjo, N. Hirata, H. Tsuruoka are responsible for the wording of the article), Earthquake forecast testing experiment for Japan, *Newslett. Seismol. Soc. Jpn.*, **20**, 7–10, 2009 (in Japanese).
- Schorlemmer, D., S. Wiemer, and M. Wyss, Earthquake statistics at Parkfield: 1. Stationarity of b values, *J. Geophys. Res.*, **109**, 2004.
- Smyth, C. and J. Mori, Assessing temporal variations in the Gutenberg-Richter distribution for a short-term forecast model, *Japan Geoscience Union Meeting*, Chiba, Japan, 2009.
- Stein, S. and M. Wyssession, *An Introduction to Seismology, Earthquakes, and Earth Structure*, Blackwell, Malden, 2003.
- Wiemer, S. and M. Wyss, Mapping spatial variability of the frequency-magnitude distribution of earthquakes, *Adv. Geophys.*, **45**, 259–301, 2002.

C. Smyth (e-mail: christine@eqh.dpri.kyoto-u.ac.jp) and J. Mori