

Effect of tectonic setting on the fit and performance of a long-range earthquake forecasting model

David Alan Rhoades,¹
Paul G. Somerville,^{2,3}
Felipe Dimer de Oliveira,²
Hong Kie Thio³

¹GNS Science, Lower Hutt, New Zealand;

²Risk Frontiers, Sydney, Australia;

³URS Corp, Pasadena, CA, USA

Abstract

The Every Earthquake a Precursor According to Scale (EEPAS) long-range earthquake forecasting model has been shown to be informative in several seismically active regions, including New Zealand, California and Japan. In previous applications of the model, the tectonic setting of earthquakes has been ignored. Here we distinguish crustal, plate interface, and slab earthquakes and apply the model to earthquakes with magnitude $M \geq 4$ in the Japan region from 1926 onwards. The target magnitude range is $M \geq 6$; the fitting period is 1966-1995; and the testing period is 1996-2005. In forecasting major slab earthquakes, it is optimal to use only slab and interface events as precursors. In forecasting major interface events, it is optimal to use only interface events as precursors. In forecasting major crustal events, it is optimal to use only crustal events as precursors. For the smoothed-seismicity component of the EEPAS model, it is optimal to use slab and interface events for earthquakes in the slab, interface events only for earthquakes on the interface, and crustal and interface events for crustal earthquakes. The optimal model parameters indicate that the precursor areas for slab earthquakes are relatively small compared to those for earthquakes in other tectonic categories, and that the precursor times and precursory earthquake magnitudes for crustal earthquakes are relatively large. The optimal models fit the learning data sets better than the raw EEPAS model, with an average information gain per earthquake of about 0.4. The average information gain is similar in the testing period, although it is higher for crustal earthquakes and lower for slab and interface earthquakes than in the learning period. These results show that earthquake interactions are stronger between earthquakes of similar tectonic types and that distinguishing tectonic types improves forecasts by enhancing the depth resolution where tectonic categories of earthquakes are vertically separated. However,

when depth resolution is ignored, the model formed by aggregating the optimal forecasts for each tectonic category performs no better than the raw EEPAS model.

Introduction

The Every Earthquake a Precursor According to Scale (EEPAS) model is showing promise as a long-range earthquake forecasting method. It was originally fitted to the New Zealand earthquake catalogue to optimize its performance in forecasting earthquakes of magnitude $M > 5.75$.¹ With unchanged parameters, it was shown to outperform a spatially varying smoothed seismicity model based on proximity to the locations of past earthquakes (PPE) in California for magnitude $M > 5.75$ ¹ and in Japan for $M > 6.75$.² It was applied at lower magnitudes ($M > 4.75$ and $M > 4.95$, respectively) to high-quality catalogues of the Kanto region, central Japan, and southern California,^{3,4} and at $M > 5.95$ in Greece,⁵ all with similarly successful results. Some elaborations of the model to take account of aftershocks have been proposed.^{6,7} Versions of the model are being tested in the Collaboratory for the Study of Earthquake Predictability (CSEP) experiments in California,⁴ New Zealand⁸ and Japan.^{9,10} Useful background information on the CSEP experiments and testing methods is given by Jordan,¹¹ Zechar *et al.*,^{12,13} and Schorlemmer *et al.*^{14,15} The EEPAS model has been shown to fit well to a physics-based synthetic earthquake catalogue involving a dense fault network.¹⁶

The EEPAS model, as applied in previous studies, treats all earthquakes the same, regardless of their tectonic setting. It does not discriminate between three distinct tectonic categories of earthquakes: those occurring within the continental crust, those on plate boundaries, and those within subducted slabs. In regions such as Japan and New Zealand, earthquakes in two or more of these categories occur in close proximity because of the subduction processes that are an important feature of the seismotectonics of these regions. This proximity causes earthquakes from different tectonic categories to interact with each other in the forecast process. However, earthquakes within a given tectonic category are expected to interact more strongly than earthquakes from different tectonic categories, because the background stress is generally discontinuous across a material interface.¹⁷

Notwithstanding this expectation, there is anecdotal evidence of apparent short-term interactions between large earthquakes across earthquake categories. Examples include the 1942 earthquakes in Wairarapa, New Zealand, in which a large earthquake in the crust was followed by one in the slab within a few months,¹⁸

Correspondence: David Alan Rhoades, GNS Science, 1 Fairway Drive, Avalon, Lower Hutt 5010, P O Box 30-368, Lower Hutt 5040, New Zealand.

Tel: +64.4.5704680 - Fax: +64.4.5674600.

E-mail: d.rhoades@gns.cri.nz

Key words: statistical seismology, earthquake forecasting, tectonic setting, Japan.

Acknowledgments: this work was funded by the Earthquake Commission Research Foundation under project n. TVH 08/567, the Foundation for Research, Science and Technology under contract n. C05X0804, and GNS Science Capability Funding. R. Buxton, M. Stirling and J. Zechar provided helpful reviews of the manuscript.

Contributions: DAR, conception and design, analysis and interpretation of data, statistical modelling, main drafting of the article; PGS, conception and design, contributions to drafting the article, critical revision of the article; FDdO, acquisition and analysis of data, critical revision of the article; HKT, analysis and interpretation of data, critical revision of the article. All authors approved the final version of the manuscript.

Conflicts of interests: the authors' employing institutions received funding through a research contract, in the case of GNS Science, or indirectly through a subcontracting process, from the Earthquake Commission Research Foundation, for the express purpose of carrying out the research described here. The research was also supported in part by other institutional research funding not specifically dedicated to this work. No author was paid any money personally for carrying out this research.

Received for publication: 10 November 2011.

Revision received: 29 January 2012.

Accepted for publication: 6 February 2012.

This work is licensed under a Creative Commons Attribution NonCommercial 3.0 License (CC BY-NC 3.0).

©Copyright D.A. Rhoades *et al.*, 2012
Licensee PAGEPress srl, Italy
Research in Geophysics 2012; 2:e3
doi:10.4081/rg.2012.e3

and the 1990 earthquakes in Weber, New Zealand, in which a large earthquake in the slab was followed by one on the plate interface within a few months.¹⁹ There may also be longer term interactions. Shimazaki²⁰ suggested that large interplate earthquakes on subduction zones along the Pacific coast in Japan are preceded (over a period of decades) by an increase in the level of activity of large intraplate earthquakes within the crust of the overriding plates, due to the high intraplate stress caused by the loading of the plate interface.

How the long-term seismogenic process is affected by tectonic setting should be revealed

by different optimal EEPAS parameters for earthquakes in different tectonic categories and different interactions between earthquakes within a given tectonic category than between earthquakes in different categories. Therefore, in this study we examine the hypothesis that the long-term seismogenic process for a major earthquake is affected by its tectonic setting, and that consequently the EEPAS model can be improved by taking account of the tectonic category to which precursory earthquakes belong. Accordingly, we separate earthquakes into the three tectonic categories and develop the EEPAS parameters for different restrictions, by tectonic category, of precursor and target earthquakes.

Classification of earthquakes by tectonic type

The main catalogue of Japan used in this study is that provided by the SEIS-PC earthquake analysis package²¹ for the period 1926-2005. This catalogue was the more complete of the two detailed catalogues of Japan that were available. The recently revised Japan Meteorological Agency catalogue was not available to us for this study. The well-known and readily available model of the plate boundaries by Gudmundsson and Sambridge²² is used to define the plate interfaces in Japan. We validated this model against the locations of events having known accurate depths and earthquake categories, and also against the larger set of events based on the International Seismological Centre (ISC) catalogue having accurate focal depths.²³ We were able to determine a depth range centred on the slab interface where events are most likely to be actual interface events. Since the latter catalogue is by no means complete, we then calibrated the SEIS-PC catalogue against the ISC catalogue by determining an average shift in depth between the two. This way, we eliminated systematic differences in hypocenter depth between the SEIS-PC catalogue and the slab model, and were able to produce a complete catalogue that has event type assigned.

The whole earthquake catalogue is shown in a map view in Figure 1, which distinguishes the three different tectonic categories. Maps of the individual earthquake categories are shown in Figure 2A-C, with depth indicated by color coding. A longitude-depth cross section across Tohoku is shown in Figure 3 for earthquakes within a one-degree latitude band around 39°N, again distinguishing earthquakes in the three tectonic categories.

For a large majority of the events, *i.e.* the smaller ones, this assignment is *statistical* in the sense that it is based on the estimate of focal depth alone, without knowing the actual

event type. The assignment is based on the distance of the hypocenter from the plane defining the plate boundary. Events within 5 km of the plate boundary are assumed to be interface events. Events more than 5 km above or below it are assumed to be crustal or slab events, respectively. One could in principle experiment with different threshold values. However, this is not done here because of the high computational burden of such experimentation. It is inevitable that some earthquakes are misclassified because of uncertainties in the plate boundary model and earthquake locations. However, the tectonic types are regarded as fixed for the purposes of the analyses carried out here. This is reasonable, because the statistical models of seismicity fitted here do not depend upon a perfect classification.

Overview of models

The EEPAS model is based on the Precursory Scale Increase (Ψ) phenomenon²⁴⁻²⁹ of an increase in the rate and magnitude of minor earthquakes, which has been found to precede most major shallow earthquakes in the long term in several well-catalogued regions of the world. The precursory swarm phenomenon³⁰⁻³⁶ is a special case of the Ψ -phenomenon. Associated with the Ψ -phenomenon are pre-

dictive scaling relations: regressions of mainshock-magnitude, precursor-time, and precursor area on the precursory magnitude level (Figure 4). These relations allow the magnitude, time of occurrence and source area of a major earthquake to be predicted from the precursory magnitude level. The EEPAS model treats each earthquake as a possible long-term precursor and uses the predictive scaling relations, together with the earthquake's magnitude, to estimate the distribution in time, magnitude and location of the earthquake's contribution to the future earthquake occurrence rate density.¹

The parameters of the EEPAS model are linked to the predictive scaling relations: a_M , b_M and σ_M to the intercept, slope and standard deviation, respectively, of the regression of mainshock magnitude on precursor magnitude; a_T , b_T and σ_T to the intercept, slope and standard deviation, respectively, of the regression of the logarithm of precursor time on precursor magnitude; and b_A and σ_A to the slope and intercept, respectively, of the regression of the logarithm of precursor area on precursor magnitude.

The EEPAS model is a mixture of a time-varying component based on the Ψ -phenomenon, as described above, and the Proximity to Past earthquakes (PPE) models were given by smoothed seismicity model). The *failure rate*

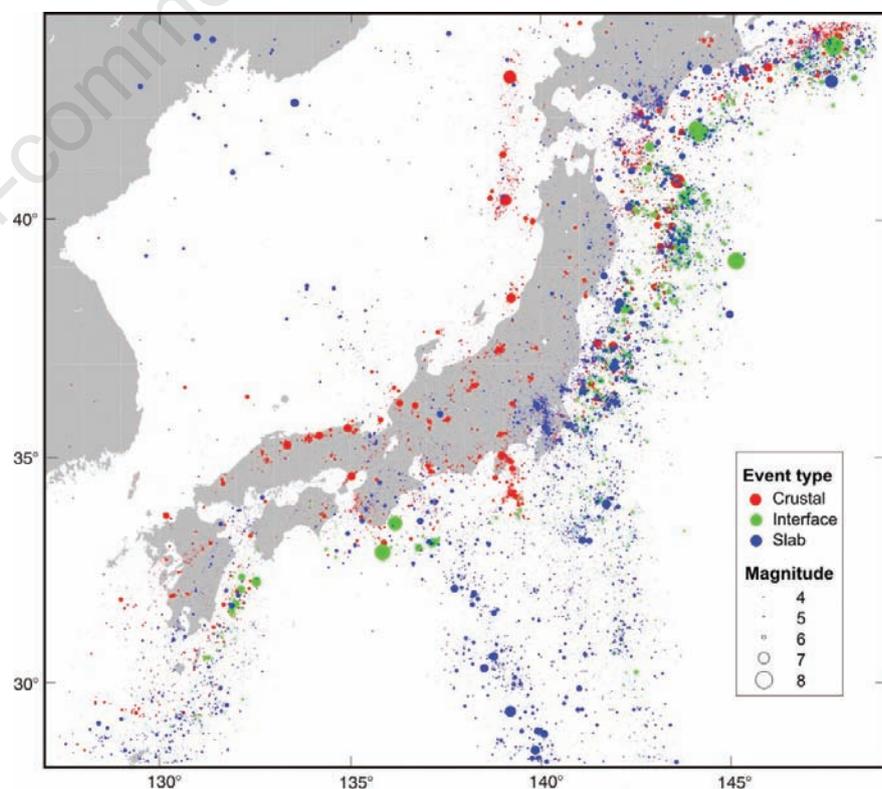


Figure 1. Earthquakes in the Japan catalogue, 1926-2005, showing different event types.

parameter μ ($0 \leq \mu \leq 1$) controls the proportional contributions of the PPE component to the mixture. It represents the proportion of targeted earthquakes that do not display the Ψ -phenomenon. If $\mu=0$, there is no contribution from the PPE model, and if $\mu=1$, there is no contribution from the time-varying component. The PPE model is based on the notion that large earthquakes occur close to where they have occurred in the past, and is also a useful reference model against which to measure the performance of the EEPAS model.

Another important reference model in the analyses presented here is the Stationary Uniform Poisson (SUP) model, in which the magnitude distribution conforms to the Gutenberg-Richter frequency-magnitude relation, but the rate of earthquake occurrence is

both temporally and spatially invariant. The SUP model is useful as a model of *least information*, and is used as a baseline for information scores defined below. The PPE model is a more elaborate form of Poisson model in which the magnitude distribution conforms to the Gutenberg-Richter frequency-magnitude relation and the rate of earthquake occurrence is in principle time-invariant (although continually updated as new earthquakes occur) but spatially varying. On the other hand, in the EEPAS model, the rate of earthquake occurrence is both temporally and spatially varying, and in small space-time volumes does not necessarily conform to the Gutenberg-Richter relation.

Technical details of the SUP, PPE and EEPAS models were by Rhoades and Evison.¹⁻³

Fitting and testing

Fitting is carried out by maximum likelihood, and quantification of improvements is assessed using the likelihood method, which is widely used for comparison of probabilistic forecasting models and has been adopted by the CSEP earthquake forecast testing centers.^{14,37-40}

Comparison of alternative models is rigor-

ously achieved by splitting the catalogue into two sets: a learning set and a testing set. The model selection and fitting is carried out on the learning set, and the performance assessment and significance testing on the testing set. In order to avoid the well-known problems associated with over-fitting,⁴¹ a goodness of fit statistic which includes a penalty for fitting extra free parameters is used. At the testing stage, the number of fitted parameters is irrelevant because there are no free parameters. By comparing goodness of fit or performance statistics of the PPE and SUP models, we can measure the information value of spatial variation of earthquake occurrence under the PPE model, and by comparing similar statistics of the EEPAS and PPE models, we can measure the information value of time-variation of earthquake occurrence under the EEPAS model. If over-fitting is successfully avoided, the information value of models tends to be similar whether assessed at the fitting stage or the testing stage.

The goodness of fit is assessed using the Akaike Information Criterion (AIC) statistic,⁴² defined for a particular model M as

$$AIC_M = -2\ln L_M + 2p_M \quad (1)$$

where L_M is the optimized log likelihood of the

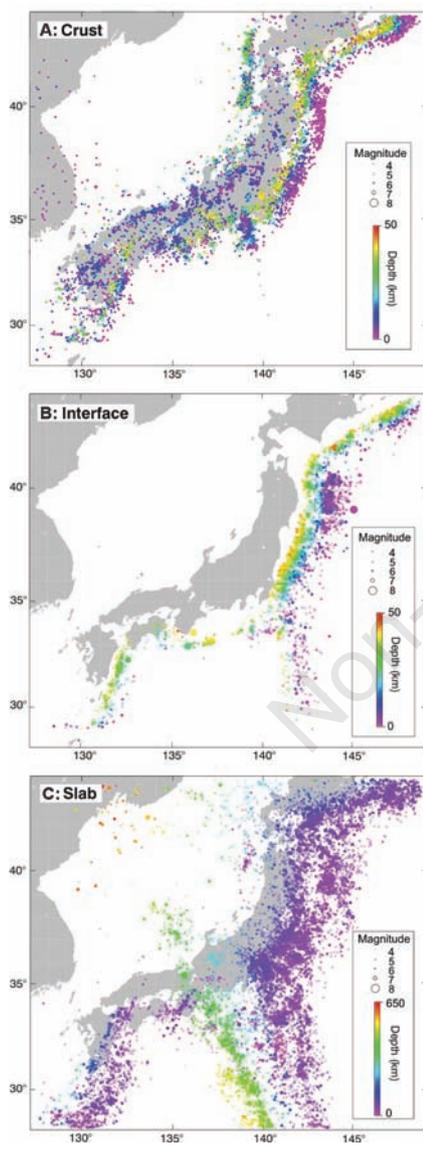


Figure 2. Location, magnitude and depth of (A) crustal earthquakes, (B) interface earthquakes, and (C) slab earthquakes in the Japan catalogue, 1926-2005.

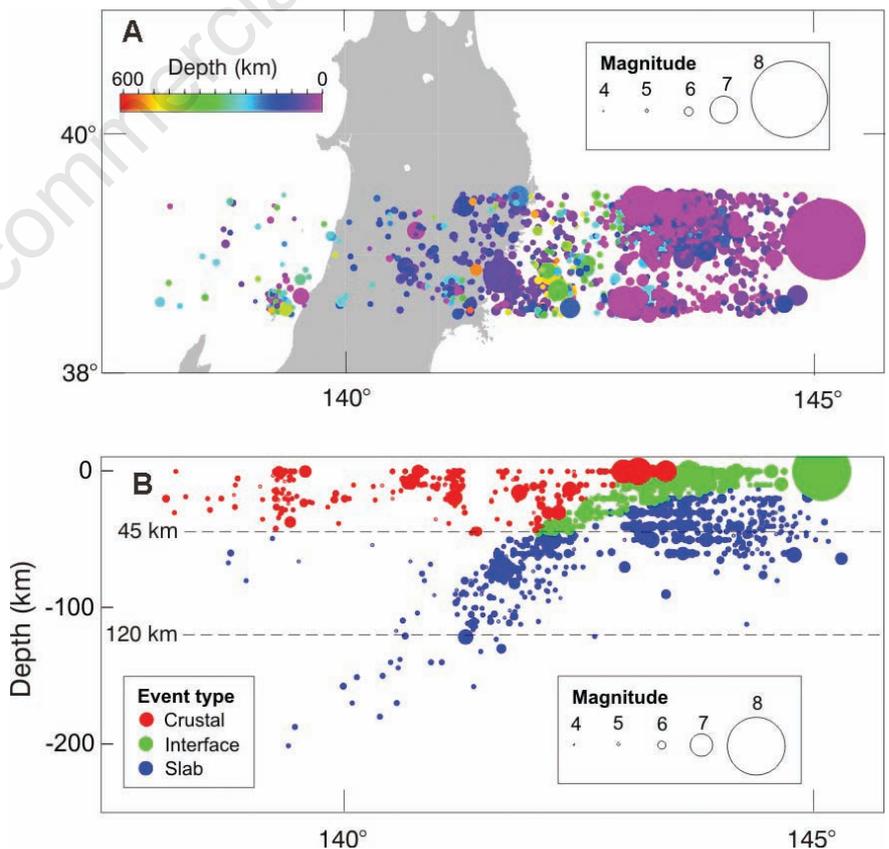


Figure 3. (A) Location of cross-section of earthquakes, selected by latitude range, through Tohoku, Japan, 1926-2005. (B) Longitude-depth cross-section of selected earthquakes, showing tectonic type of each event.

model, and p_M is the number of fitted parameters. A relatively low value of AIC indicates a relatively high information value, *i.e.* a model that explains the data relatively well. Formally, the information value of a fitted model is expressed as the information rate per earthquake, I_M , defined by

$$I_M = (AIC_{SUP} - AIC_M) / (2N) \quad (2)$$

where N is the number of earthquakes in the target set.

The fitting period is 1965 January 1 to 1995 December 31. This is the period over which the models are optimized, although the earthquakes from 1926 onwards are used in the analysis, with those before 1965 serving to warm up the models. Only earthquakes with hypocentral depth $h \leq 120$ km are used, because the precursory scale increase is held to be a phenomenon mainly associated with shallow earthquakes.²⁶ The targeted earthquakes are those inside the polygonal region of surveillance shown in Figure 5A. The region of surveillance was chosen to be as large as possible in order to obtain the maximum possible number of earthquakes in each tectonic subset, subject to adequate catalogue completeness for precursory earthquakes. The crustal, interface and shallow slab earthquakes occupy smaller regions, as shown in Figure 5B-D. The magnitude threshold m_0 for earthquakes contributing to the analysis is 3.95 and the magnitude threshold m_c for targeted earthquakes is 5.95.

Fitting without regard to tectonic type

Initially the SUP, PPE and EEPAS models were fitted to earthquakes in the Japan catalogue without regard to tectonic type. The fit was also examined in the case that only earthquakes with $h \leq 45$ km are used. Note that all earthquakes with $h \leq 45$ km are classified as in the slab. The information rate is not usually improved greatly by fitting many parameters. Here we fit only four parameters of the EEPAS model (a_M , a_T , σ_A and μ), with other parameters being fixed at typical values from previous studies. The equal-weights version of the EEPAS model¹ is used. Here, and in all fits to subsets below, the fitted parameters are optimized subject to constraints. The parameters a_M , a_T , σ_A and μ are constrained to lie in the intervals 1.0-1.8, 0.9-2.5, 0.5-10.0 and 0.0-0.5, respectively.

The optimal value of 0.5 for μ (Table 1) indicates that the fitted EEPAS model is a 50:50 mixture of the time-varying component and the PPE smoothed-seismicity component. This is a high value of μ compared to estimates in previous studies,^{1,3} which are typically much closer to 0.

It can be seen from Table 1 that the fitted parameter values are similar for the two cases $h \leq 120$ km and $h \leq 45$ km. Also the information rates are similar, indicating that the EEPAS model forecasts the deeper earthquakes in the range $45 < h \leq 120$ km about as well as it does the shallower earthquakes.

The difference ΔI between I_{EEPAS} and I_{PPE} , here referred to as the information gain, is an indicator of the quantity of time-varying information provided by the EEPAS model. From Table 1, the difference is 0.16 for $h \leq 120$ km and 0.13 for $h \leq 45$ km. These values correspond to an average probability gain per earthquake of 1.15 and 1.19, respectively. The observed information gains are similar to the value of 0.13 obtained by Rhoades and Evison² in fitting the same models (but with $\mu=0$) to a larger region of surveillance using the Japan Meteorological Agency (JMA) catalogue with $m_c=6.25$. They are, however, much lower than the information gains obtained from studies of the EEPAS model in California, which range from 0.72 for the whole of California¹ with $m_c=5.75$ to 0.82 for Southern California⁴ with $m_c=4.95$. The high information gain of the EEPAS model in California has been ascribed to

the simplicity of the plate boundary tectonics in California with most of the seismicity being shallower than 15 km.³

Fitting to earthquakes in separate tectonic classes

Next we describe the fitting of models with the target set of earthquakes restricted to individual tectonic-type subsets. We consider different predictor sets in order to determine whether the precursors of major earthquakes in a particular tectonic class come from the same class or different classes of earthquakes.

Slab earthquakes

There are 42 slab earthquakes in the target set, including 29 with $h \leq 45$ km. The predictability of the shallow slab events with $h \leq 45$ km is of interest in its own right. Therefore, we

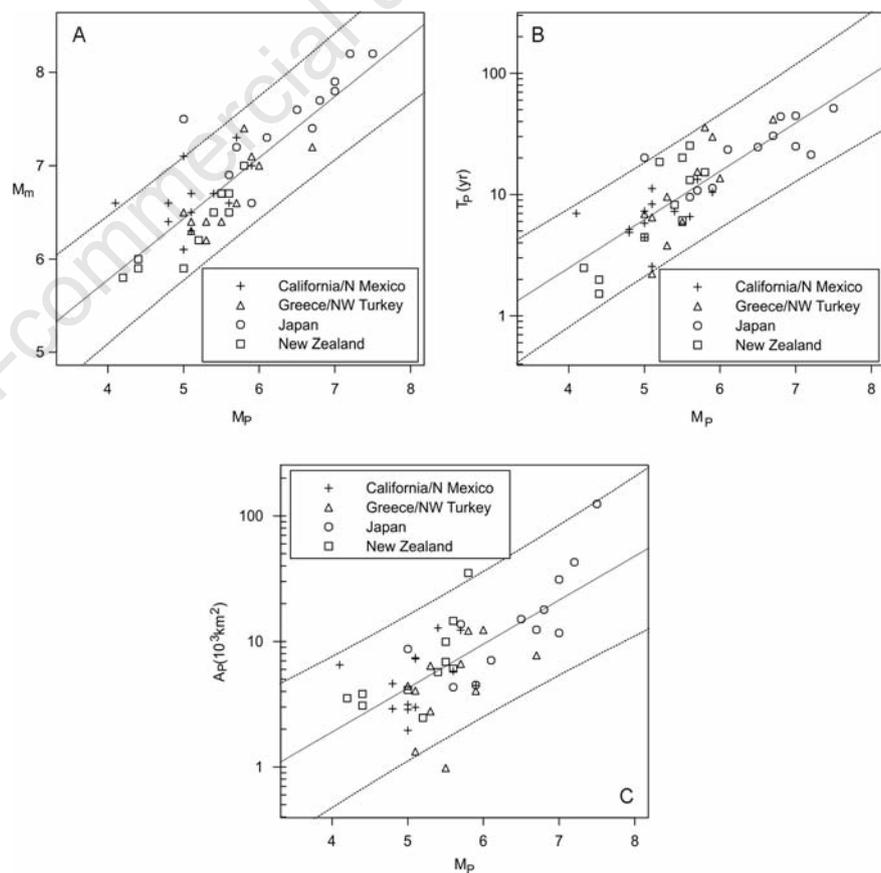


Figure 4. Predictive scaling relations of long-term seismogenesis derived from examples of the precursory scale increase phenomenon in four well-catalogued regions. (A) Mainshock magnitude M_m versus precursor magnitude M_p ; (B) precursor time T_p versus M_p ; (C) precursor area A_p versus M_p . Solid lines are fitted regressions and dotted lines are 95% tolerance limits.²⁶

consider the predictability both of all slab earthquakes within the region of surveillance shown in Figure 5A, and the shallow events only within the smaller region shown in Figure 5D.

Table 2 shows the result of fitting the model to all slab events using different predictor sets of earthquakes (slab, crust and slab, interface and slab, and all earthquakes). Note that the largest values of I_{EEPAS} , I_{PPE} and ΔI are all obtained when both interface and slab events are included in the predictor set. Since the PPE model embodies the hypothesis that large earthquakes tend to occur near to where they have occurred before, the result for I_{PPE} suggests that large earthquakes in the slab tend to occur near to where large earthquakes have previously occurred either in the slab or on the interface. The results for I_{EEPAS} and ΔI together imply that the earthquakes precursory to large earthquakes in the slab may occur either in the slab or on the interface, but hardly ever in the crust. This analysis does not show whether or not it is only relatively shallow slab earthquakes close to the interface that have precursors on the interface.

Note that the values of I_{EEPAS} , I_{PPE} and ΔI in Table 2 are all higher than those in Table 1. It follows that earthquakes in the slab as a separate class are on average more predictable than earthquakes in general in the target region, using any of the predictor sets in Table 2. The optimum value of 0.59 for ΔI indicates that the information gain for slab earthquakes is much greater than that for earthquakes in general in Table 1, and not much less than that found in the studies of California.

Comparing the fitted parameter values for the optimal model in Table 2 with those in Table 1, we note that the values of a_M and σ_A are rather similar. The value of a_T (1.15) for slab earthquakes is less than that in Table 1 (1.43). If all other parameter values were equal, this difference would indicate that earthquakes in the slab have a precursor time which is shorter on average by a factor of 0.52 than earthquakes in general in the target region. Also, we note that the parameter μ of the optimal model, at 0.34, is less than that in Table 1. A lower value of μ indicates that a higher proportion of the target earthquakes have precursory sequences conforming well to the fitted EEPAS model parameters. Again, this analysis does not show whether or not it is the relatively shallow slab earthquakes close to the interface that conform best to the fitted EEPAS model parameters.

It is instructive to compare the parameter values obtained with slab earthquakes only in the predictor set with those of the optimal model when slab and interface events are both included. The high value of 1.7 for the parameter a_M indicates that the precursory earthquakes occurring in the slab have relatively

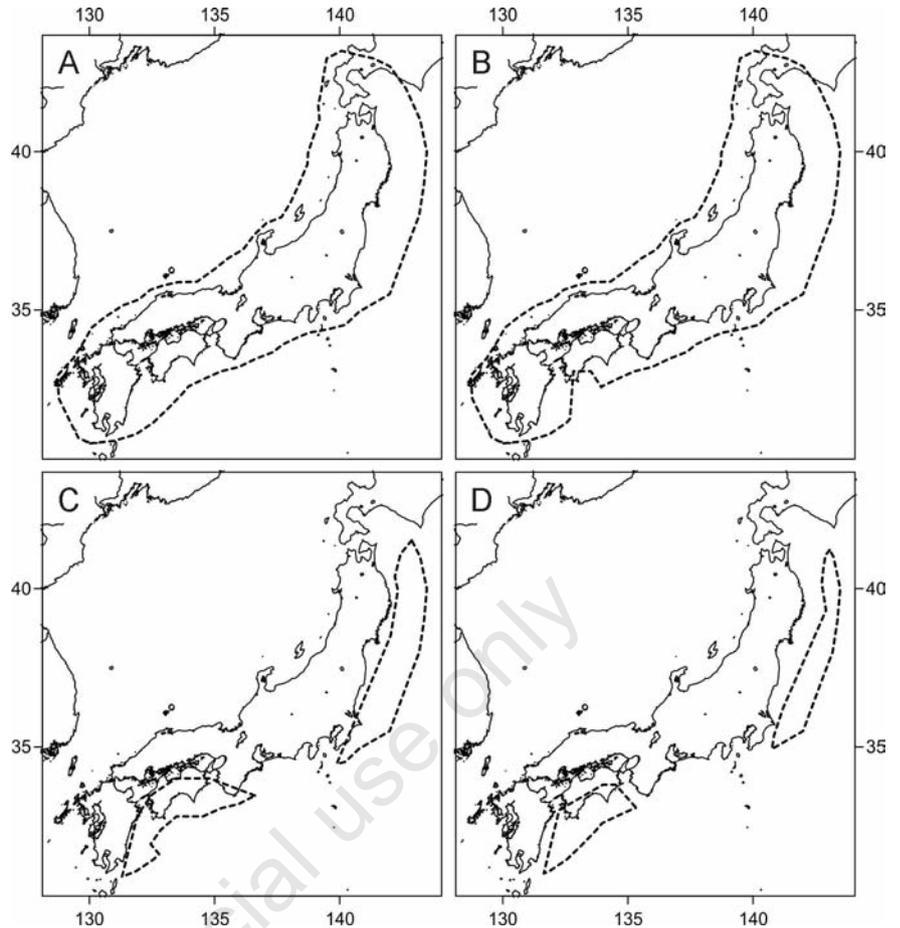


Figure 5. (A) Region of surveillance for fitting EEPAS model (dashed polygon). (B) Sub-region occupied by crustal earthquakes. (C) Sub-region occupied by interface earthquakes. (D) Sub-region occupied by shallow ($h < 45$ km) slab earthquakes.

Table 1. Parameters and information scores for EEPAS model fitted to Japan catalogue within testing region of Figure 5A, without regard to tectonic type.

Parameter	Depth restriction	
	$h \leq 120$ km (n=105)	$h \leq 45$ km (n=76)
a_M	1.47	1.48
b_M	1.0*	1.0*
σ_M	0.32*	0.32*
a_T	1.43	1.42
b_T	0.4*	0.4*
σ_T	0.23*	0.23*
b_A	1.06	1.07
σ_A	0.35*	0.35*
μ	0.50	0.50
Information score	Value	
I_{PPE}	1.42	1.44
I_{EEPAS}	1.58	1.57
$\Delta I = I_{EEPAS} - I_{PPE}$	0.16	0.13

*Fixed parameter value.

low magnitudes. Because of this high value of a_M , the relatively large values of a_T and σ_A should not be taken to indicate relatively long precursor times and large precursory areas, respectively. The relatively high value of 0.49 for μ indicates that a relatively high proportion of target earthquakes do not have precursory sequences in the slab conforming to the fitted EEPAS parameters.

Table 3 shows the results of fitting the shallow slab events only. There are only 13 earthquakes in the target set, making the fitting exercise somewhat suspect. However, the highest value of I_{PPE} is again obtained when slab and interface events are included in the predictor set. In contrast, the highest value of I_{EEPAS} is obtained with only the slab earthquakes in the predictor set. A notable feature is the relatively low value of 0.18 for μ , indicating that most of the targeted shallow slab earthquakes have precursory sequences in the shallow slab conforming to the EEPAS model.

Interface earthquakes

There are 21 interface earthquakes in the target set. Table 4 shows the result of fitting the models to these data for different predictor sets of earthquakes.

The highest value of I_{PPE} (1.18) is obtained when the predictor set consists of the crustal, interface and shallow ($h \leq 45$ km) slab earthquakes. This indicates that large interface earthquakes tend to occur where large earthquakes have previously occurred either in the crust, on the interface or in the shallow part of the slab.

The largest value of I_{EEPAS} (1.26) is obtained when the predictor set consists of interface events only. This indicates that the precursory earthquakes for interface events occur mostly on the interface itself.

Comparing the fitted parameter values for the optimal model in Table 4 with those in Table 1, we note that a_M is larger (1.60 compared to 1.47), a_T is smaller (1.13 compared to 1.43), σ_A is much larger (1.71 compared to 1.06), and μ is smaller (0.30 compared to 0.5). The larger value of a_M indicates that precursory earthquakes tend to be relatively small in magnitude. Taking into account the difference for a_M , the differences for a_T and σ_A indicate that on average precursor times are relatively short and precursor areas relatively large for earthquakes on the interface compared with those for earthquakes in general. Also, the relatively low value of μ indicates that a high proportion of the large earthquakes on the interface have precursory sequences conforming well to the fitted parameters.

The results in Table 4 suggest that an EEPAS

model with the PPE component derived from crustal, interface and shallow slab earthquakes, and the time-varying component derived from only the interface might provide a slightly better fit to the data than any model in Table 4. However, such a model was found to reduce the AIC value by only a negligible amount. Therefore, the model using only interface earthquakes for both components is preferred on account of its simplicity.

Crustal earthquakes

There are 42 crustal earthquakes in the target set. Table 5 shows the result of fitting the models to these data for different predictor sets of earthquakes. We optimize the same four EEPAS parameters as in the previous sec-

tion, with other parameters fixed to the same values listed in Table 1.

It is clear from Table 5 that there is no benefit to the EEPAS model information score from including the slab earthquakes, and particularly the deeper events with $45 < h \leq 120$ km in the predictor set. It follows that not many precursors to crustal earthquakes occur in the slab.

Note that the highest values of both I_{EEPAS} and I_{PPE} in Table 5 are obtained when the crust and interface earthquakes are used as the predictor set. However, the highest value of ΔI is obtained when only the crustal earthquakes are used.

To investigate further whether a significant proportion of precursory earthquakes to major crustal earthquakes occur on the plate interface, we fit an EEPAS model in which the PPE component is estimated using earthquakes from both the crust and interface, and the time-varying component is estimated using

Table 2. Parameters and information scores for EEPAS model fitted to slab earthquakes in the Japan catalogue (n=42) within region of surveillance in Figure 5A. EEPAS parameters not listed are constrained to the same values as in Table 1.

Model component	Predictor set (C: Crust; I: Interface; S: Slab)			
PPE	S	C, S	I, S	C, I, S
Time-varying	S	C, S	I, S	C, I, S
Parameter	Fitted value			
a_M	1.70	1.70	1.52	1.45
a_T	1.31	1.25	1.15	1.09
σ_A	1.10	0.60	0.96	0.94
μ	0.49	0.50	0.34	0.36
Information score	Value			
I_{PPE}	1.67	1.57	1.70	1.58
I_{EEPAS}	2.15	1.86	2.29	2.02
ΔI^*	0.48	0.29	0.59	0.44

$$^*\Delta I = I_{EEPAS} - I_{PPE}$$

Table 3. Parameters and information scores for EEPAS model fitted to shallow ($b \leq 45$ km) slab earthquakes (n=13) within the region of surveillance shown in Figure 5D. EEPAS parameters not listed are constrained to the same values as in Table 1.

Model component	Predictor set (C: Crust; I: Interface; S: Slab)			
PPE	S	I, S	C, S	C, I, S
Time-varying	S	I, S	C, S	C, I, S
Parameter	Fitted value			
a_M	1.66	1.62	1.65	1.65
a_T	1.20	1.19	1.14	1.14
σ_A	0.60	0.65	0.55	0.63
μ	0.18	0.18	0.33	0.29
Information score	Value			
I_{PPE}	1.47	1.27	1.55	1.39
I_{EEPAS}	2.52	2.12	2.08	1.88
ΔI^*	1.05	0.84	0.52	0.49

$$^*\Delta I = I_{EEPAS} - I_{PPE}$$

only the earthquakes from the crust. As well as the equal-weighting strategy adopted above, we also consider the alternative weighting strategy of down-weighting aftershocks. The results are summarized in Table 6.

For both variations on the EEPAS model considered in Table 6, the value of I_{EEPAS} is greater than the maximum value in Table 5. The maximum value of I_{EEPAS} is obtained for the case of down-weighted aftershocks. This is our preferred model for crustal earthquakes, and we infer that the precursory earthquakes to major crustal earthquakes occur predominantly in the crust.

Effect of tectonic category on predictive scaling relations

The fitted parameters of the optimal models for each tectonic category indicate some differences in the predictive scaling relations of long-term seismogenesis depending on the tectonic category of the target earthquakes. These differences are illustrated in Figure 6.

In Figure 6A, the median main shock magnitude $M_m = a_M + b_M m$ for an observed precursory earthquake of magnitude m is plotted for main shocks in each tectonic category, based on the optimal fitted EEPAS model for that category. The differences between tectonic categories are up to a third of a magnitude unit, with the median main shock magnitude for crustal earthquakes being lowest. In other words, the precursor earthquakes for a given main shock magnitude are, on average, higher for crustal earthquakes than for interface and slab earthquakes.

In Figure 6B, the median precursor time is plotted for each tectonic category based on the optimal EEPAS parameters. This figure shows that, for a given precursor magnitude, the precursor time for crustal earthquakes is, on average, about double that for slab and interface events.

In Figure 6C, the 99th percentile of the distribution of precursory area is plotted for each tectonic category, based on the parameters σ_A and b_A of the optimal EEPAS model for that category. This figure shows that precursory areas for earthquakes in the slab are smaller by a factor of two or three than those for earthquakes in the slab or interface, respectively. The 99th percentile was used here rather than the median to make Figure 6C as comparable as possible with Figure 4C, in which the whole precursory area from examples of the precursory scale increase is plotted.

Although the two approaches to deriving the predictive scaling relations, through the precursory scale increase phenomenon (Figure 4) and the EEPAS model parameters (Figure 6) are not entirely equivalent, the similarities

Table 4. Parameters and information scores for EEPAS model fitted to interface earthquakes (n=21) within the region of surveillance shown in Figure 5C. EEPAS parameters not listed are constrained to the same values as in Table 1.

Model component	Predictor set (C: crust; I: interface; S: slab)					
	I	C, I	I, S	I, S ($h \leq 45$)	C, I, S	C, I, S ($h \leq 45$)
PPE	I	C, I	I, S	I, S ($h \leq 45$)	C, I, S	C, I, S ($h \leq 45$)
Time-varying	I	C, I	I, S	I, S ($h \leq 45$)	C, I, S	C, I, S ($h \leq 45$)
Parameter	Fitted values					
a_M	1.60	1.00	1.60	1.60	1.18	1.00
a_T	1.13	1.26	1.00	1.05	0.90	0.90
σ_A	1.71	1.50	2.92	2.32	2.27	2.12
μ	0.30	0.39	0.50	0.50	0.50	0.50
Information score	Value					
I_{PPE}	0.91	1.02	0.99	1.06	1.12	1.18
I_{EEPAS}	1.26	1.13	1.03	1.12	1.08	1.16
ΔI^*	0.36	0.11	0.04	0.06	-0.04	-0.02

* $\Delta I = I_{EEPAS} - I_{PPE}$

Table 5. Parameters and information scores for EEPAS model fitted to crustal earthquakes in the Japan catalogue (n=42) within the region of surveillance shown in Figure 5B. EEPAS parameters not listed are constrained to the same values as in Table 1.

Model component	Predictor set (C: crust; I: interface; S: slab)					
	C	C, I	C, S	C, S ($h \leq 45$)	C, I, S	C, I, S ($h \leq 45$)
PPE	C	C, I	C, S	C, S ($h \leq 45$)	C, I, S	C, I, S ($h \leq 45$)
Time-varying	C	C, I	C, S	C, S ($h \leq 45$)	C, I, S	C, I, S ($h \leq 45$)
Parameter	Fitted values					
a_M	1.25	1.27	1.18	1.26	1.27	1.27
a_T	1.51	1.48	1.46	1.48	1.46	1.46
σ_A	1.35	1.38	1.17	1.34	1.21	1.21
μ	0.50	0.50	0.50	0.44	0.50	0.50
Information score	Value					
I_{PPE}	0.89	1.06	0.89	0.88	1.02	0.99
I_{EEPAS}	1.05	1.15	0.97	1.04	1.03	1.08
ΔI^*	0.16	0.09	0.08	0.16	0.01	0.07

* $\Delta I = I_{EEPAS} - I_{PPE}$

Table 6. Parameters and information scores for EEPAS model fitted to crustal earthquakes (n=42) within the region of surveillance shown in Figure 5B. The PPE model is derived using crustal and interface earthquakes. The time-varying component of EEPAS is derived from crustal earthquakes only.

Parameter	Weighting strategy	
	Equal weights	Aftershocks down-weighted
Parameter	Fitted values	
a_M	1.27	1.31
a_T	1.49	1.45
σ_A	1.29	1.41
μ	0.50	0.50
Information score	Value	
I_{PPE}	1.06	
I_{EEPAS}	1.17	1.23
ΔI^*	0.11	0.17

* $\Delta I = I_{EEPAS} - I_{PPE}$

and differences can be noted. First note that all three relations in Figure 6A are close to the fitted relation between mainshock magnitude and precursor magnitude in Figure 4A. Secondly, the slab and interface relations in Figure 6B are close to, although slightly below, the fitted relation for precursor time in Figure 4B, but the crustal relation is appreciably above, although well within the 95% tolerance limits. Finally, the interface relation in Figure 6C is close to the fitted relation for precursory area in Figure 4C, but the slab and interface relations are lower. The slab relation is even outside the 95% tolerance limits of Figure 4C, and so represents the most significant departure from the relations derived from the precursory scale increase phenomenon.

Enhancement of information scores

The overall results of fitting the EEPAS model are summarized in Table 7, in which the information scores for the best model for each targeted subset of earthquakes are shown, together with those for the fit to all earthquakes without regard to tectonic type. Note that for each targeted subset the information gain ΔI is greater than that for all earthquakes; for the slab events by a factor of more than 3 and for the interface events by a factor of more than 2.

Tests on independent data

For testing on independent data the information score I_M for a model M is defined as

$$I_M = (\ln L_M - \ln L_{SUP}) / N \quad (3)$$

where $\ln L$ is the log likelihood statistic and N is the number of target earthquakes in the testing set. The number of fitted parameters does not affect the information score for independent testing because the parameter values are all fixed at the testing stage.

Fitted models described in the previous section were applied to the testing set of earthquakes (1996-2005) and the information scores were computed. These testing results are summarized in Table 8, in which the information scores for the best model for each targeted subset of earthquakes are shown, together with those for the model for all earthquakes without regard to tectonic type.

The information scores in Table 8 can be compared to those in Table 7. It is noticeable that, unlike Table 7, the value of ΔI in Table 8 is higher for the crust than for the interface or

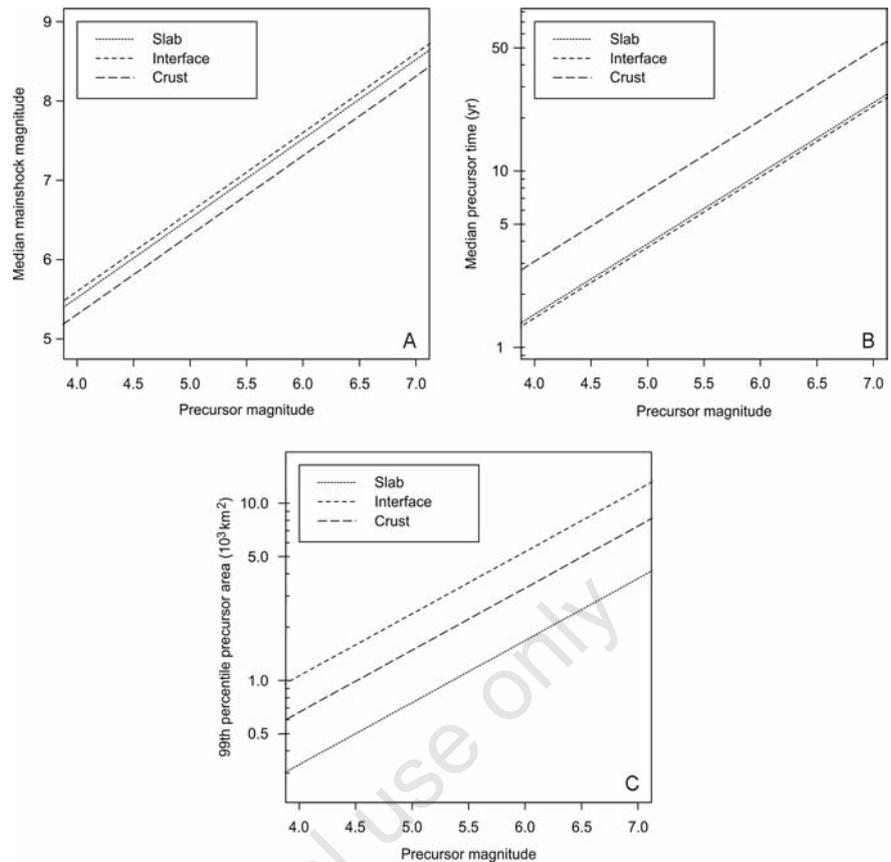


Figure 6. Predictive scaling relations for earthquakes in the slab interface and crust, as indicated by the EEPAS model parameters fitted separately to each tectonic type. (A) Median mainshock magnitude versus precursor magnitude; (B) median precursor time versus precursor magnitude; (C) 99th percentile of precursor area versus precursor magnitude.

Table 7. Information scores for best model in each class in fitting to the learning set (1966-1995).

Information score	Target set			
	All (n=105)	Slab (n=42)	Interface (n=21)	Crust (n=42)
I_{PPE}	1.42	1.70	2.27	1.06
I_{EEPAS}	1.58	2.29	2.62	1.23
ΔI^*	0.16	0.59	0.35	0.17

Table 8. Information scores for best model in each class evaluated on the testing set of earthquakes (1996-2005).

Information score	Target set			
	All (n=37)	Slab (n=13)	Interface (n=7)	Crust (n=17)
I_{PPE}	1.07	1.90	2.42	0.38
I_{EEPAS}	1.29	2.07	2.69	1.05
ΔI^*	0.22	0.17	0.27	0.67

slab subsets. The high value for the crust is due to a particularly low value of I_{PPE} rather than a high value of I_{EEPAS} . Conversely, the relatively low values of ΔI for the slab and interface appear to be due to high values of I_{PPE} rather than to low values of I_{EEPAS} . In any case, not too much reliance should be placed on the results for the individual subsets because of the relatively small number of target earthquakes in each. But it is notable that in Table 8, as in Table 7, the overall information scores of the best models combined are higher than those for the model fitted to all data without regard to tectonic type, and by similar amounts: for I_{PPE} by 0.23, for I_{EEPAS} by 0.43 and for ΔI by 0.20.

In Figure 7, the performance of the best fitting model in each class on the testing set is

compared with that of other models of the same class using other data inputs. The tests used here are the t-test and w-test.⁴⁰ Note that an information gain per earthquake greater than zero indicates that the model named in the header of the plot outperforms the model named in the left side panel.

Figure 7A shows that among models for slab earthquakes, the performance of the best-fitting model (EEPAS using interface and slab earthquakes as predictors) is significantly better at about the 95% confidence level (the w-test being significant and t-test not quite significant) than that of the model using all tectonic categories of earthquake as predictors, but not significantly different from the other models. It also shows that the EEPAS model using only the slab earthquakes as predictors is the best-per-

forming model during the testing period.

Figure 7B shows that among models to forecast interface earthquakes, the performance of the best fitting model (EEPAS using only interface earthquakes as predictors) is not significantly different from that of other models in the same class. The lack of any significance differences between any of the models here is likely due to there being only seven earthquakes in the target set during the testing period.

Figure 7C shows that among models to forecast crustal earthquakes, the performance of the best fitting model (the EEPAS model with aftershocks down-weighted using only crustal earthquakes as predictors for the time-varying component and crustal and interface events for the PPE component) is significantly better with 95% confidence than all other models in

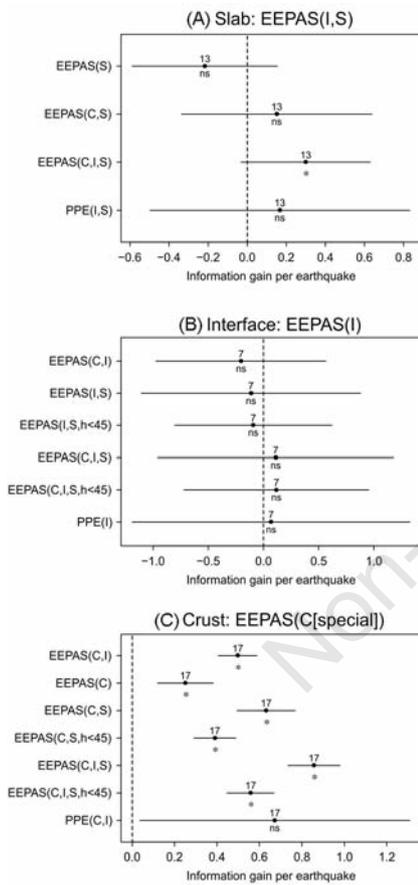


Figure 7. t-test comparison of the information gain per earthquake, during the test period, of the best-fitting EEPAS model targeting (A) slab earthquakes, (B) interface earthquakes, (C) crustal earthquakes over other models. Predictor sets are shown in brackets. (C: Crust; I: Interface; S: Slab). Error bars are 95% confidence intervals. The number of target earthquakes is shown above the mean information gain and the w-test significance below (ns, not significant; *significant at 95% confidence level).

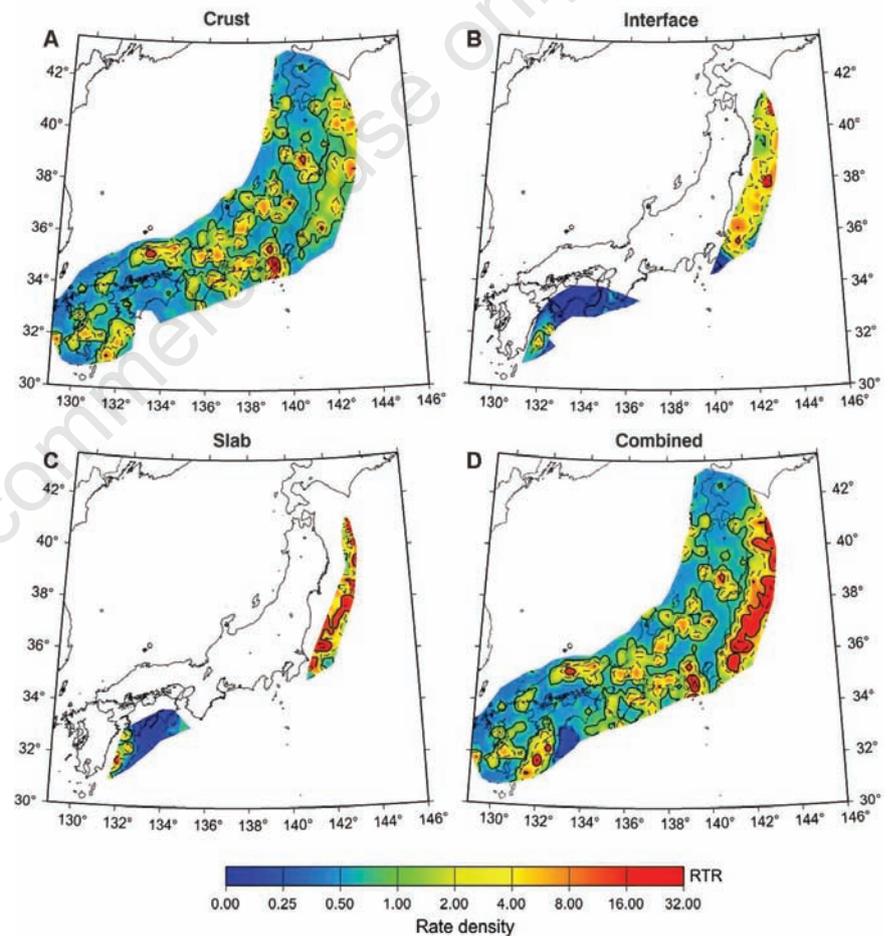


Figure 8. Illustrating the construction of the aggregate model for shallow earthquakes: EEPAS model forecast of rate density of shallow ($h \leq 45$ km) earthquake occurrence on the plate interface for magnitude 7.0 on 2009 06 30, based on catalogue up to end of 2005 within the region of surveillance. (A) Crustal model; (B) Interface model; (C) Slab model; and (D) Aggregate model.

the same class. In this case, the results obtained from the fitting period are strongly confirmed by the performance of the fitted models during the testing period.

Aggregate model

Aggregate EEPAS models were created for each depth range (0-120 km, and 0-45 km) from the optimal models for each tectonic category. For the 0-120 km depth range, slab earthquakes can occur anywhere within the region of surveillance of Figure 5A, crustal earthquakes within the region of surveillance of Figure 5B, but interface events only within the smaller region of Figure 5C. The rate density of the aggregate model is, therefore, obtained by summing the rate densities of the crustal, interface and slab models at locations within the region of Figure 5C, by summing those of the crustal and slab models at other locations within the region of Figure 5B, and by using the slab rate density only at other locations. For the 0-45 km depth range, the difference is that slab earthquakes can occur only within the region of Figure 5D. The rate density of the aggregate model is, therefore, the sum of the rate densities of the crustal, interface and slab models at locations within region of Figure 5D, the sum of the rate densities of the crustal and interface models at other locations within region of Figure 5C, and the rate density of the crustal model alone at other locations within region of Figure 5B. The composition of this aggregate model is illustrated in Figure 8.

The fit and performance of the aggregate EEPAS models were compared with that of the simple EEPAS models fitted to all of the data in the 0-120 km and 0-45 km depth ranges (Table 8). The results are summarized in terms of the information scores for the fitting and testing periods in Table 9. This table shows that in three comparisons out of four, the aggregate model has a lower information score than the corresponding simple model, the exception being the 0-45 km depth range during the fitting period. During the fitting period, the information score of the aggregate model is negatively affected by the high number of fitted parameters (12) compared to the simple model (4). However, it is during the testing period, when the number of fitted parameters has no impact on the information scores, that the deficit of the aggregate model compared to the simple model is greatest.

Depth resolution in earthquake forecast experiments

The analysis of the previous section indicates

Table 9. Information scores of aggregate and simple EEPAS models.

Period	Information score Depth range	Aggregate model	Simple model
Fitting period	0-120 km	1.54	1.59
	0-45 km	1.60	1.58
Testing Period	0-120 km	1.08	1.29
	0-45 km	1.03	1.15

that depth selection is an important issue when constructing earthquake forecasting models based on past seismicity. The EEPAS model is typical of such forecasting models in that it does not discriminate between the depths of earthquakes in the explanatory catalogue or the target set. Ideally such models would be 3-dimensional rather than 2-dimensional, and forecast experiments, like those being conducted in the CSEP testing centers, would measure the ability of models to forecast the depth of earthquakes as well as their latitude, longitude, magnitude and time of occurrence. In practice, the CSEP testing centers have adopted different target depth ranges: 0-30 km in California,³⁹ 0-40 km in New Zealand,⁸ and 0-30 km and 0-100 km in different testing regions in Japan.⁹ Depth resolution is clearly more important in subduction regions like those of Japan and New Zealand, where the seismicity extends to depths of several hundred kilometres, than in transform plate boundary regions like California, where the seismicity is mostly confined to the top 20 km of the crust. In the former regions, dividing the target region into depth layers, as is being proposed for a CSEP experiment in the Kanto region of central Japan (K. Nanjo, personal communication, 2011), might allow 2D models to be used effectively in a 3D experiment. Modellers could take account of tectonic categories by careful selection of depths in the input catalogue for each layer. However, such layering is practicable only in the best-instrumented onshore regions where the seismograph network is able to reliably estimate earthquake depths.

Conclusions

Distinguishing the three tectonic categories of earthquake has generally confirmed the hypothesis that earthquake interactions are stronger between earthquakes in similar tectonic categories than between those of different categories. This is an intuitively obvious result, but one that it is important to demonstrate. Our analyses have shown that for forecasting the major earthquakes within one tectonic category, the fit and performance of the EEPAS model is improved when restrictions are applied to the tectonic category of the predictor set of earth-

quakes. The optimal EEPAS model uses slab and interface events as precursors to major slab earthquakes, interface events only as precursors to major interface events, and crustal events only as precursors to major crustal events. For the smoothed-seismicity component of the EEPAS model, it is optimal to use slab and interface events to forecast the location of earthquakes in the slab, interface events only to forecast the location of earthquakes on the interface, and both crustal and interface events to forecast the location of events on the interface. The optimal fits to different categories reveal appreciable differences in the scaling of precursor time and area. In particular, the precursor times for crustal earthquakes are relatively long and the precursory areas for slab earthquakes are relatively small, compared to the other categories and to the scaling relations derived from examples of the precursory scale increase phenomenon.

The results of the fitting exercise have been partially borne out by the performance of the models on the testing set, but further testing on future data would provide a stronger confirmation of the optimal model for each tectonic category.

The robustness of the conclusions obtained here using the Japan earthquake catalogue could be tested by similar studies of other well-catalogued regions of the world where subduction is present. However, there are presently no other regions with catalogues to match that of Japan in quality and quantity of data.

The obvious explanation for the differences between tectonic categories in the EEPAS parameters reflecting the scaling of precursor time and area is that different tectonic regimes behave differently with respect to their crustal properties, such as state of stress and rate of deformation. However, studies of wider regions could clarify whether the observed differences in scaling are actually a characteristic of the tectonic categories themselves or of the particular conditions prevailing in the Japan region. Numerical and physical modelling of the earthquake generating process may in time provide insights into these differences.

References

1. Rhoades DA, Evison FF. Long-range earthquake forecasting with every earthquake a precursor according to scale. *Pure Appl Geophys* 2004;161:47-71.
2. Rhoades DA, Evison FF. Test of the EEPAS forecasting model on the Japan earthquake catalogue. *Pure Appl Geophys* 2005;162:1271-90.
3. Rhoades DA, Evison FF. The EEPAS forecasting model and the probability of moderate-to-large earthquakes in central Japan. *Tectonophysics* 2006;417:119-30.
4. Rhoades DA. Application of the EEPAS model to forecasting earthquakes of moderate magnitude in southern California. *Seismol Res Lett* 2007;78:110-5.
5. Console R, Rhoades DA, Murru M, et al. Comparative performance of time-invariant, long-range and short-range forecasting models on the earthquake catalogue of Greece. *J Geophys Res* 2006;111:B09304.
6. Rhoades DA. Long-range earthquake forecasting allowing for aftershocks. *Geophys J Int* 2009;178:244-56.
7. Rhoades DA, Gerstenberger MC. Mixture models for improved short-term earthquake forecasting. *Bull Seism Soc Am* 2009;99:636-46.
8. Gerstenberger MC, Rhoades DA. New Zealand Earthquake Forecast Testing Centre. *Pure Appl Geophys* 2010;167:877-92.
9. Nanjo KZ, Tsuruoka H, Hirata N, Jordan TH. Overview of the first earthquake forecast testing experiment in Japan. *Earth Planets Space* 2011;63:159-69.
10. Rhoades DA. Application of a long-range forecasting model to earthquakes in the Japan mainland region. *Earth Planets Space* 2011;63:197-206.
11. Jordan TH. Earthquake predictability, brick by brick. *Seismol Res Lett* 2006;77:3-6.
12. Zechar JD, Gerstenberger MC, Rhoades DA. Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts. *Bull Seism Soc Am* 2010;100:1184-95.
13. Zechar JD, Schorlemmer D, Liukis M, et al. The Collaboratory for the Study of Earthquake Predictability perspective on computational earthquake science. *Concurr Comp Pract Exper* 2010;22:1836-47.
14. Schorlemmer D, Gerstenberger MC, Wiemer S, et al. Earthquake likelihood model testing. *Seismol Res Lett* 2007;78:17-29.
15. Schorlemmer D, Zechar JD, Werner MJ, et al. First results of the Regional Earthquake Likelihood Models experiment. *Pure Appl Geophys* 2010;167:5-22.
16. Rhoades DA, Robinson R, Gerstenberger MC. Long-range forecasting in synthetic earthquake catalogues. *Geophys J Int* 2011;185:1037-48.
17. Hirano S, Yamashita T. Analysis of the static stress field around faults lying along and intersecting a biomaterial surface. *Geophys J Int* 2011;187:1460-78.
18. Doser DI, Webb TH. Source parameters of large historical (1917-1961) earthquakes, North Island, New Zealand. *Geophys J Int* 2003;152:795-832.
19. Anderson HJ. Focal mechanisms of some recent large New Zealand earthquakes. *NZ J Geol Geophys* 1991;34:103-9.
20. Shimazaki K. Intra-plate seismicity and inter-plate earthquakes: historical activity in Southwest Japan. *Tectonophysics* 1976;33:33-42.
21. Ishikawa Y. SEIS-PC, new version. *Geol Data Proc* 1986;11:65-74.
22. Gudmundsson O, Sambridge M. A regionalized upper mantle (RUM) seismic model. *J Geophys Res* 1998;103:7121-36.
23. Engdahl ER, van der Hilst R, Buland R. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull Seism Soc Am* 1998;88:722-43.
24. Evison FF, Rhoades DA. Model of long-term seismogenesis. *Ann Geofis* 2001;44:81-93.
25. Evison FF, Rhoades DA. Precursory scale increase and long-term seismogenesis in California and northern Mexico. *Ann Geophys* 2002;45:479-95.
26. Evison FF, Rhoades DA. Demarcation and scaling of long-term seismogenesis. *Pure Appl Geophys* 2004;161:21-45.
27. Evison F, Rhoades DA. Long-term seismogenesis and self-organized criticality. *Earth Planets Space* 2004;56:749-60.
28. Evison FF, Rhoades DA. Multiple-main-shock events and long-term seismogenesis in Italy and New Zealand. *NZ J Geol Geophys* 2005;48:523-36.
29. Papadimitriou EE, Evison FF, Rhoades DA, et al. Long-term seismogenesis in Greece: comparison of the evolving stress field and precursory scale increase approaches. *J Geophys Res* 2006;111:B05318.
30. Evison FF. The precursory earthquake swarm. *Phys Earth Planet Int* 1977;15:19-23.
31. Evison FF. Generalised precursory swarm hypothesis. *J Phys Earth* 1982;30:115-24.
32. Evison FF, Rhoades DA. The precursory earthquake swarm in New Zealand: hypothesis tests. *NZ J Geol Geophys* 1993;36:51-60.
33. Evison FF, Rhoades DA. The precursory earthquake swarm in New Zealand: hypothesis tests II. *NZ J Geol Geophys* 1997;40:537-47.
34. Evison FF, Rhoades DA. The precursory earthquake swarm in Japan: hypothesis test. *Earth Planets Space* 1999;51:1267-77.
35. Evison FF, Rhoades DA. The precursory earthquake swarm in Greece. *Ann Geofis* 2000;43:991-1009.
36. Rhoades DA. Lessons and questions from thirty years of testing the precursory swarm hypothesis. *Pure Appl Geophys* 2010;167:629-44.
37. Rhoades DA, Evison FF. Time-variable factors in earthquake hazard. *Tectonophysics* 1989;167:201-10.
38. Jackson DD. Hypothesis testing and earthquake prediction. *Proc Natl Acad Sci USA* 1996;93:3772-5.
39. Schorlemmer D, Gerstenberger MC. RELM Testing Center. *Seismol Res Lett* 2007;78:1-16.
40. Rhoades DA, Schorlemmer D, Gerstenberger MC, et al. Efficient testing of earthquake forecasting models. *Acta Geophys* 2011;59:728-47.
41. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. New York: Springer; 2001. pp. 533.
42. Akaike H. A new look at the statistical model identification. *IEEE Trans Auto Cont* 1974;19:716-23.