July 25, 2016

To: Natural Hazards and Earth System Sciences Editorial Office

Dear Sir/Madam,

I appreciate your constructive comments. I quoted those comments, which follow with my responses. All of the changes are marked with underlines in the revised manuscript.

1) Formula (1) gives a rate lambda that has the following dimensions: number of earthquakes/(time x area). This is consistent with the dimensions of the Kernel in formula (2): number of earthquakes/area, but it is not with units of Figure 2 (number of earth- quakes/time/Volume). Is it by purpose, or is it a mistake? Please explain.

The unit for $\lambda(M,x)$ can be in per area or volume, depending on the dimensions of forecasting models, i.e. number of earthquake/year/km^2 for two-dimensional models while number of earthquake/year/km^3 for three-dimensional ones. The explanation is described in Lines 70-71.

2) Further, explain clearly what lambda is in formula (1). Is it the seismicity rate for earthquake of magnitude M or with magnitude exceeding M, as it seems?

In Lines 71-73, I stated that '*Although* $\lambda(M,x)$ *is a function of magnitude, for simply representation, forecasting seismicity rates in different magnitude ranges will be shown in this study.*' For example, forecasted seismicity rates for M ≥ 3.0 are presented in Figure 2; rates for different magnitude intervals are presented in Figures 6(a)-(d).

3) For reader's better understanding, it would be useful to provide tables where all parameters used in the estimation of the two models are gathered. For example, provide estimates of TM, NM, PL, c and d.

It is rather difficult to provide a table to specify the assumed parameters, which varies case by case. For example, in the case of the Ryukyu region, T_M are 37 and 14 years while N_M are the total number of earthquakes with magnitudes larger than 4.0 and 3.0 for TTSN and CWBSN cases, respectively; in the Kanto case, T_M are 89, 32 and 22 years while N_M are the total number of earthquakes in the catalog with magnitudes larger than 4.5, 3.5 and 2.5, respectively.

4) Moreover, explain: did you try with different power laws?

A fixed value of 1.75 is assumed for the power law index. I explained in Lines 81-84 that *'Molina et al. (2001) recommend PL in between 1.5 and 2.0 and Chan et al. (2010) concluded insignificant differences between the results when PL is assumed to be between the recommended values. I thus assumed an intermediate value of 1.75 in this study.'*

5) how did you estimate the coefficients c and d?

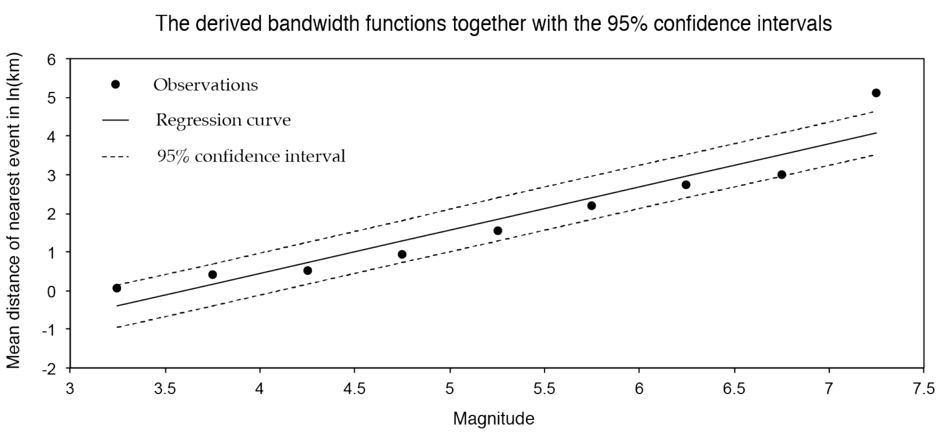
The *c* and *d* values of the bandwidth function for the two regions were determined through the linear regression. I stated that in Lines 144-146 and 239-240 for the Ryukyu and Kanto cases, respectively.

6) From your data, it seems that the bandwidth H(M) is very concentrated about the hypocenter/epicenter xi since c is at most 1 km, and d is less than M. The consequence is that H(M) goes rapidly to 0, and decays more rapidly for larger magnitudes. When the distance x-xi is much larger than the bandwidth, the Kernel becomes constant (not depending from the position). A discussion of the kernel properties could be useful.

The bandwidth function can be represented as:

$$H(M) = c \cdot e^{d \cdot M} \ .$$

The function assumes that the distances to other events are longer for larger events and vice versa. For the Ryukyu case as example (Figure below), *c* and *d* values were determined using the linear regression (solid line) based on observations (black dots). Based on the curve, it is assumed that the distance to other events for M3.25 is 0.66 km, while longer distance of 58.86 km for M7.25 events. Using this function, spatial distribution of seismicity can be forecasted.



7) In formula (5) ΔR (M,x,t) could be better denoted as ΔRn (M,x,t).

The formula (5) has been revised based on the suggestion.

8) Since the rate- and-state friction model requires the knowledge of the focal mechanism parameters, it's worth stressing that it can be applied only when these are available, which usually happens only for large earthquakes.

For the application of the rate-and-state friction model, the treatment of receiver faults followed the procedure of Catalli and Chan (2012) and assumed a spatially variable receiver fault plane for each calculation grid. A receiver fault plane for each grid consists with the closest reference focal mechanism. I stated that in Lines 157-160.

9) The author mentions that ΔCFS was calculated in a homogeneous half-space: please provide the assumed P and S wave velocities.

The use of parameters for ΔCFS calculation is included in Lines 168-170. Note that the information on wave velocities is not required for the calculation through the program of COULOMB 3.3 (Toda and Stein, 2002).

10) Specify better what the locution "target depth" means for the rate-and-state friction model computations.

In Lines '154-157', I stated that '*For source parameters of receiver fault planes, both location and focal mechanism should be defined and utilized for ΔCFS calculation. The locations are assumed to be the coordinate and the target depth of each calculation grid.*'

11) The importance of the depth factor is partly surprising. I explain why. In all predictions, the less you predict, the more your predictions are reliable. More details one adds, the more he risks to be wrong. Applying this principle to the seismicity predictions, one would expect that restricting predictions to the epicenter location one would obtain results more accurate than the ones got when predicting the hypocenters (since the depth parameter is added). This is not the case for this paper. This should be explain. My opinion is that in the calculations, the kernel parameters were estimated for a 3D case and applied both for 2D and 3D, which of course is unfair. The 2D case should be treated with a proper 2D kernel.

I am aware that each forecasting factor results in additional misfit. As mentioned by the reviewer, misfit of hypocenter would be larger than that of epicenter (in distance). In this manuscript, however, I validated the forecasting ability through the Molchan diagrams, that normalize forecasting result from best (0 % of alarm-occupied space in Figures 8 and 9) to worst (100 % of alarm-occupied space in Figures 8 and 9). The results of both Ryukyu and Kanto cases suggest relatively better forecasting ability for the three-dimensional models.

For application of the Kernel function, proper space dimensions are implemented, i.e. $\lambda(M,x)$ (in equation 1) is in per area for two-dimensional model whereas in per volume for three-dimensional one. I stated that in Lines 70-71.

12) The Kanto case is treated too quickly. It seems that only results for the rate-and-state friction model are shown. Figures 8 and 9 should be illustrated and discussed better. In particular, the curves "kernel function + rate/state" of Figure 9 should be explained. Are they the result of a combined method?

I revised the manuscript and included additional discussion on the Kanto case in Chapter 5.1-5.2 (e.g. the null hypothesis test).

In Figure 9, the model that combines the smoothing Kernel function and the rate-and-state friction model is validated. I explained in Lines 259-260 and the caption of Figure 9.

13) The section 5.3 on application on seismic and tsunami hazard must be either cancelled or treated more accurately. Forecasting epicenters and hypocenters of large tsunami- genic earthquakes is a very relevant result. However, it is not clear how this can be achieved with the present methods. Results should be shown where seismicity distri- bution is given for band of magnitudes larger than 8 (like in Figure 6, where they are given only for M up to 7.9).

The discussion on application to seismic and tsunami hazard assessments has been removed.

Sincerely yours,

Chung-Han Chan

Earth Observatory of Singapore, Nanyang Technological University
N2-01A-14, 50 Nanyang Avenue, Singapore 639798
Tel: +65 6592-3129
Email: chchan@ntu.edu.sg