Major comments（RC2）:

1. On earthquake data presentation: for each case study how many aftershocks are contained in the ISC database vs. hysteresis model predictions. Make a table corresponding to the time-intervals considered in this study, i.e, with 1-d/no, 1-week/no, 30-days/no etc. Also, show a depth distribution of events per time-interval according to the ISC catalogue.

Response: In the Tohoku earthquake case, there were 15062 aftershocks in the study area within one year after the main shock (Table 1). In the finite fault model used in this article, the focal depth is 20-25km, and according to the depth distribution of aftershocks at multiple time scales, the number of aftershocks is the largest at the depth of 35-40km (Figure 1).
In the Wenchuan earthquake case, there were 1455 aftershocks in the study area within one year after the main shock (Table 1). In the finite fault model used in this paper, the focal depth is 10-15km. According to the depth distribution of aftershocks at multiple time scales, the number of aftershocks is the largest at 10-15km depth. Aftershocks are not necessarily distributed the most on the focal depth surface (Figure 2).

Table 1 The number of aftershocks of typical historical earthquakes on multiple time scales

	1Day	30Days	90Days	180Days	365Days
Tohoku earthquake	1241	7642	10984	13002	15062
Wenchuan earthquake	369	957	1180	1327	1455

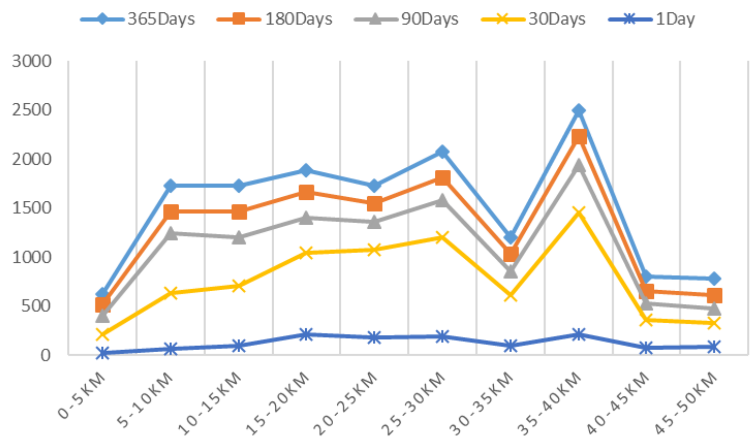

Figure 1 Multi-time scale aftershock depth distribution curve of Tohoku earthquake

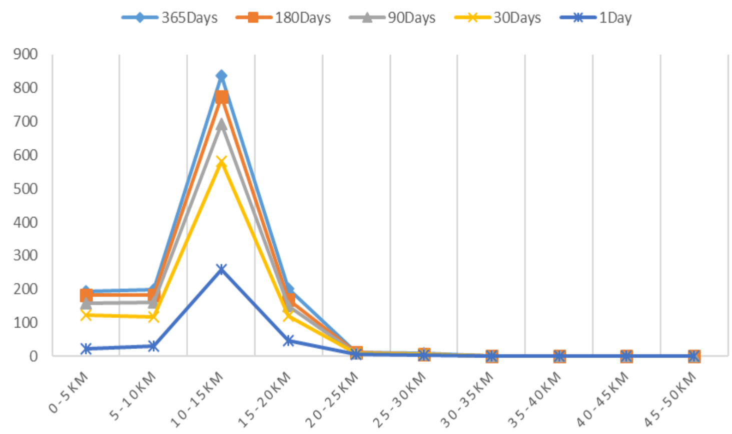

Figure 2 Multi-time scale aftershock depth distribution curve of Wenchuan earthquake

2. To obtain a more accurate comparison of aftershock locations with hysteresis model predictions, I recommend to the authors to use relocated earthquake catalogues. It is obvious that the ISC catalog contains artifacts regarding the location of events (see cross-alignments in Fig. 6) but the purpose of the ISC catalog is not to compare aftershock prediction models with physics-based (Coulomb) predictions.

Response: The database of aftershocks used in original text is from Reviewed ISC Bulletin, which is a subset of the ISC Bulletin that has been manually reviewed by ISC analysts. This includes all events that have been relocated by the ISC. In the previous comparison of the two methods in this article, the sub-cell location where the aftershock was located was used for evaluation, and the sub-cells with aftershocks were marked, which is also the reason for the cross-alignments in Figure 6 of original text. Therefore, the actual location of the aftershock event is further used instead of the sub-cell location, and the threshold is set to 0.5. Later, the prediction results were verified on the focal depth of the two earthquake cases to compare the effects of the aftershock hysteresis model and the Coulomb failure stress change method. The evaluation results of the aftershock hysteresis model are as follows: 97.6% of the Wenchuan earthquake aftershocks fall in the area with the predicted value greater than 0.5, and 96% of the Tohoku earthquake aftershocks fall within the area with the predicted value greater than 0.5. The evaluation results based on the Coulomb failure stress change method are as follows: 87.3% of the Wenchuan earthquake aftershocks fall in the area with a predicted value greater than 0.5, and 45.3% of the Tohoku earthquake aftershocks fall within the area with a predicted value greater than 0.5 (Figure 3 and Figure 4). Therefore, if the evaluation is made from the specific location of the aftershock event, the prediction result of the constructed model is still better than the result based on the Coulomb failure stress change method.

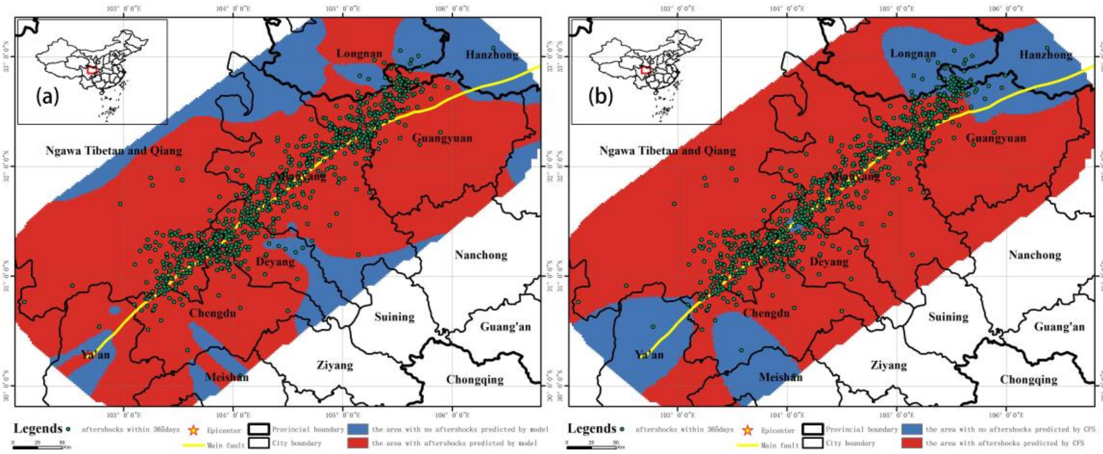

Figure 3 The hysteresis model prediction result and the △CFS prediction result of the Wenchuan earthquake

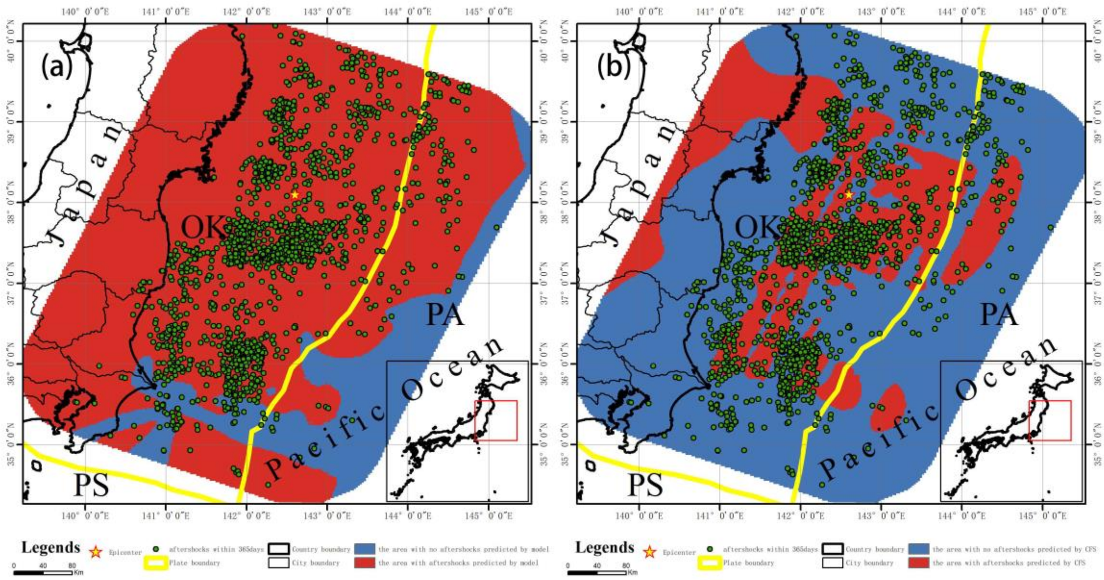

Figure 4 The hysteresis model prediction result and the △CFS prediction result of the Tohoku earthquake

3. The authors compare aftershock decay rates results (hysteresis model) with Omori (Utsu) empirical laws but do not present the actual results of their application in this manuscript. So I would like to see an aftershock decay diagram (Utsu vs hysteresis) for both case studies (Wenchuan and Tohokou).

Response: According to the modified Omori formula, the aftershock attenuation of the two earthquake cases of Tohoku earthquake and Wenchuan earthquake are analyzed, and the three coefficients in the attenuation formula of the two earthquake cases are determined, namely c, K and p. Based on the modified Omori formula, aftershock attenuation maps of two earthquake cases can be obtained (Figure 5 and Figure 6). The modified Omori formula reflects the attenuation trend of the occurrence rate of aftershocks over time. The attenuation equations and derivative functions of the two earthquake cases are shown in Table 2. The revised Omori formula can reflect the attenuation of the aftershock event rate over time. In addition, only the quantitative attenuation formula cannot give a good visualization of the attenuation process in space. From the derivative functions of the attenuation formulas of the two earthquake cases, as time increases, the absolute values of the slope of the derivative functions become smaller and smaller.

Table 2 Modified Omori formula and derived function of typical earthquake cases

Earthquake	Modified Omori formula	Derived function
Tohoku earthquake	$N(t)=300(t+0.42)^{-0.8}$	$N'(t)=-240(t+0.42)^{-1.8}$
Wenchuan earthquake	$N(t)=67.08(t+0.21)^{-1.07}$	$N'(t)=-71.78(t+0.21)^{-2.07}$

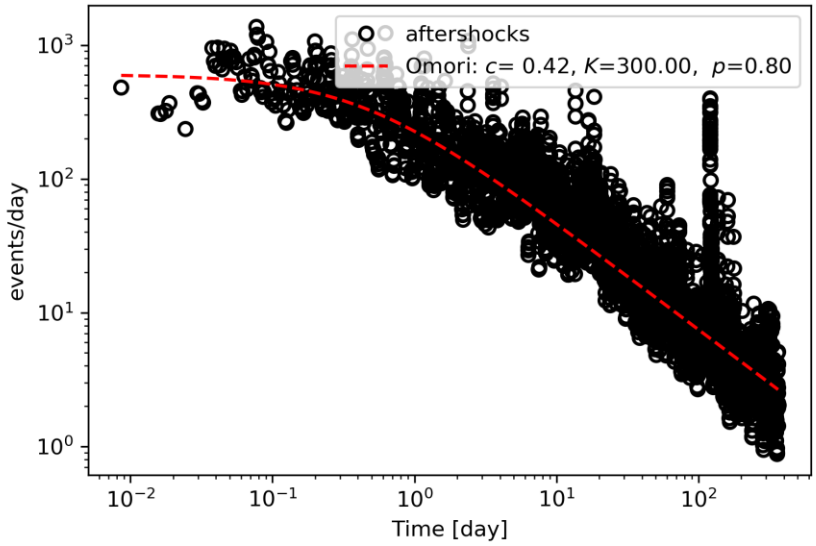

Figure 5 The modified Omori formula aftershock attenuation curve of Tohoku earthquake

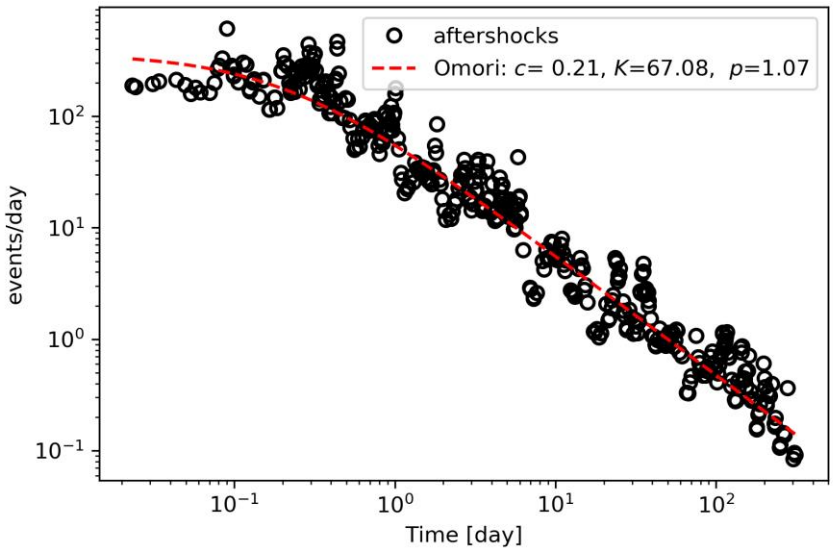

Figure 6 The modified Omori formula aftershock attenuation curve of Wenchuan earthquake

4. The 2011 (Tohoku) hysteresis modeling did not take into account the change in the properties of the elastic medium, i.e the 2011 earthquake ruptured the subduction interface between continental and oceanic crust. Given the configuration of the plate geometry offshore Japan, it is easy to count (within the ISC location uncertainties) how many ISC aftershocks occurred on the subduction interface, inside the upper (Eurasia plate) or inside the slab.

Response: The Tohoku earthquake occurred due to the subduction of the Pacific plate to the Eurasian plate. The aftershocks of the Tohoku earthquake mainly occurred near the junction of the Eurasian plate and the Pacific plate (Figure 7). They all belong to the earthquake between the plates. The Japanese offshore plate is mainly the Okhotsk plate, which is part of the Eurasian plate. A total of 12,462 (about 82.7%) aftershocks occurred in the Okhotsk plate, and 2,576 (about 17.1%) aftershocks occurred in the Pacific plate. This part of the content is intended as the tectonic background of the Tohoku earthquake.

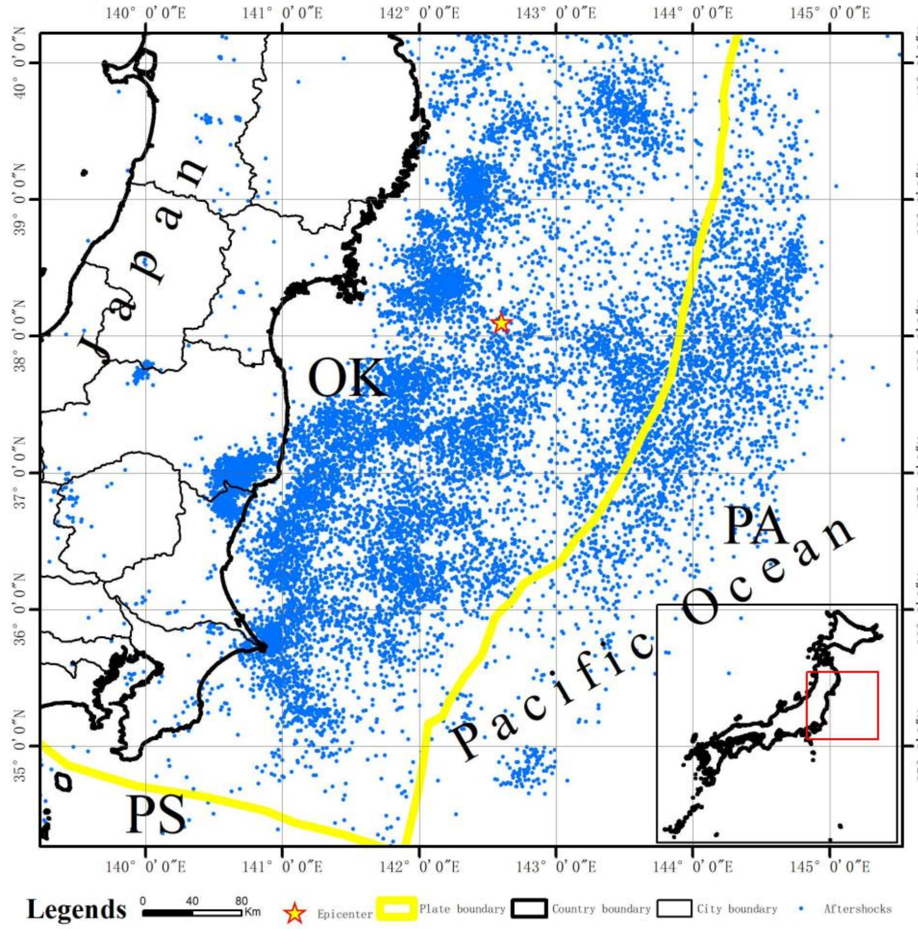

Figure 7 Aftershocks distribution of the Tohoku earthquake

5. The hysteresis model fails to predict no aftershock zones close to the mainshock as it occurred in the Tohoku case, i.e in Fig. 8 the area surrounding the mainshock (a few km at all directions; in fact the main asperity and towards the east; see Hayes 2011, https://link.springer.com/article/10.5047/eps.2011.05.012) is coloured the same as the areas where aftershocks are predicted.

Response: The prediction result given by the model is the approximate range of aftershocks, that is, the position of 5km3 sub-cells where aftershocks may occur. Each cell will have a relative probability of aftershocks, which is between 0-1. Since this probability value is less than 1 and greater than 0, it does not necessarily mean that aftershocks will occur or that aftershocks will definitely not occur. At locations close to the main shock, the probability value predicted by the model is more likely to be greater than the threshold 0.5 set in the article. Therefore, the sub-cells close to the main shock tend to have a greater probability of aftershocks than the cells far away from the main shock. . The precise location of aftershocks cannot be accurately predicted. For some physical models such as the Coulomb failure stress change, it can also reflect the distribution of aftershocks relative to the main shock. In this paper, the evaluation method based ROC curve is used, and all possible thresholds are taken into account to evaluate the model and physical model in the text. According to the ROC curves of the two methods, the effect of the hysteresis model in the article may be poor under some thresholds, but its AUC value is much greater than that of the physical model. The comprehensive results show that its prediction of the likely location of aftershocks is better than that based on the Coulomb failure stress change.

6. The hysteresis model overpredicts aftershocks, see 1-day etc. spatial distribution for the Wenchuan case (Fig. 5) where over 50% of the predicted area to be covered by afteshocks is in fact void of aftershocks. Indeed, aftershocks are distributed parallel to the main rupture and most of them occur on the hangingwall of the thrust and below the cos-seismic slip surface (Tong et al. 2010; https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2009JB006625) which is in good agreement with physics-based models (see for example the Athens 1999 earthquake case https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2007JB005504 or the L’Aquila 2009 earthquake case https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-246X.2011.05279.x

Response: The prediction results of the hysteresis model are the likely locations of aftershocks at different time scales after the main shock. At each location, the predicted value is a number between 0-1, which represents the probability of aftershocks that may occur at that location. In this article, we take the prediction threshold as 0.5, and think that when the prediction value is greater than 0.5, an earthquake is more likely to occur in the sub-cell with a volume of $5km^3$. In fact, when the predicted value is less than 0.5, there is also the possibility of aftershocks, but this possibility is relatively small. For this prediction model, if the threshold is increased from 0.5 to 0.6, 0.7, 0.8, it can be seen that in the focal depth section of the two earthquake cases, the predicted coverage area of aftershocks is gradually reduced, but while increasing the threshold, the local area prediction will also produce more errors and deviations (Figure 8 and Figure 9). In addition, if we focus on some aftershocks far away from the fault, we will find that aftershocks are also likely to occur at locations far away from the fault on multiple time scales, but the density of aftershocks is relatively small at these locations. Therefore, if these sparsely distributed aftershocks are taken into account, the predicted aftershock coverage area is wider than the area where the aftershocks are concentrated along the fault.

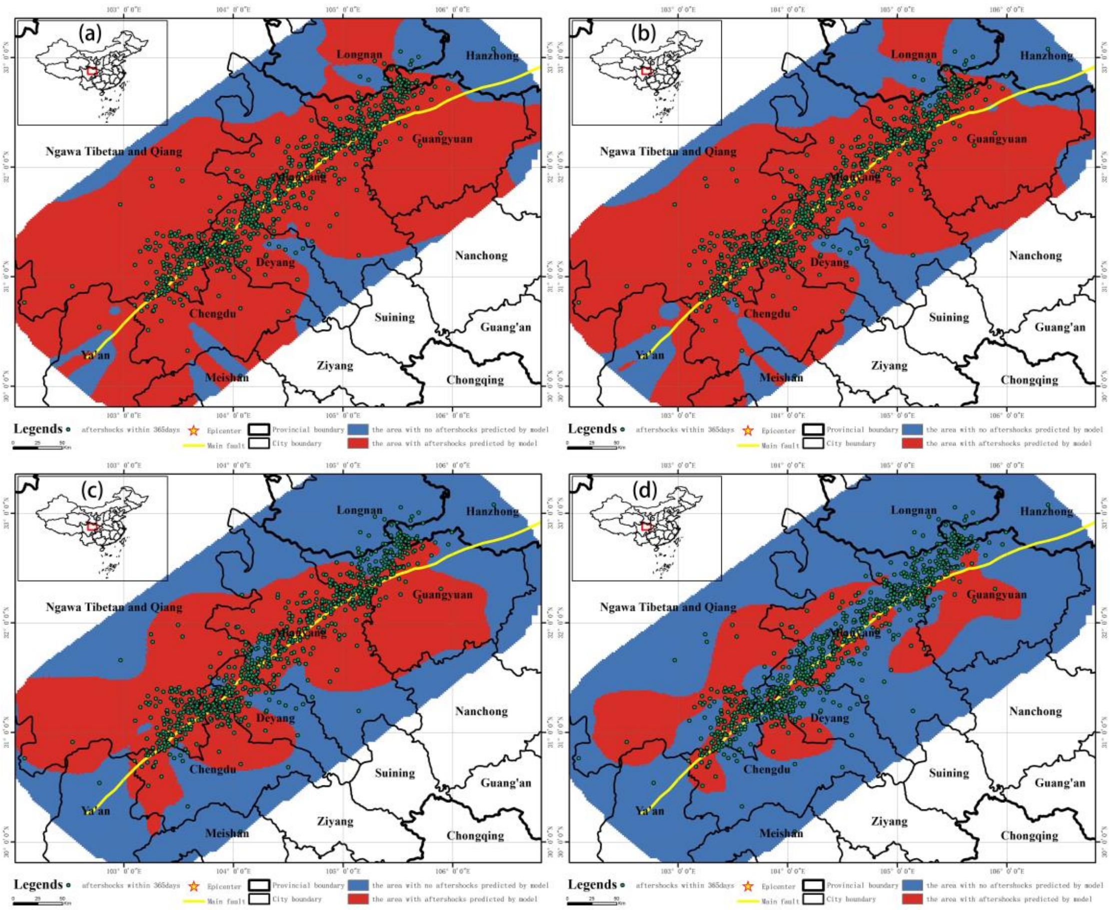

Figure 8 Wenchuan earthquake hysteresis model prediction results under different thresholds (Figure a to d show the model prediction results when the threshold is 0.5, 0.6, 0.7, 0.8, respectively)

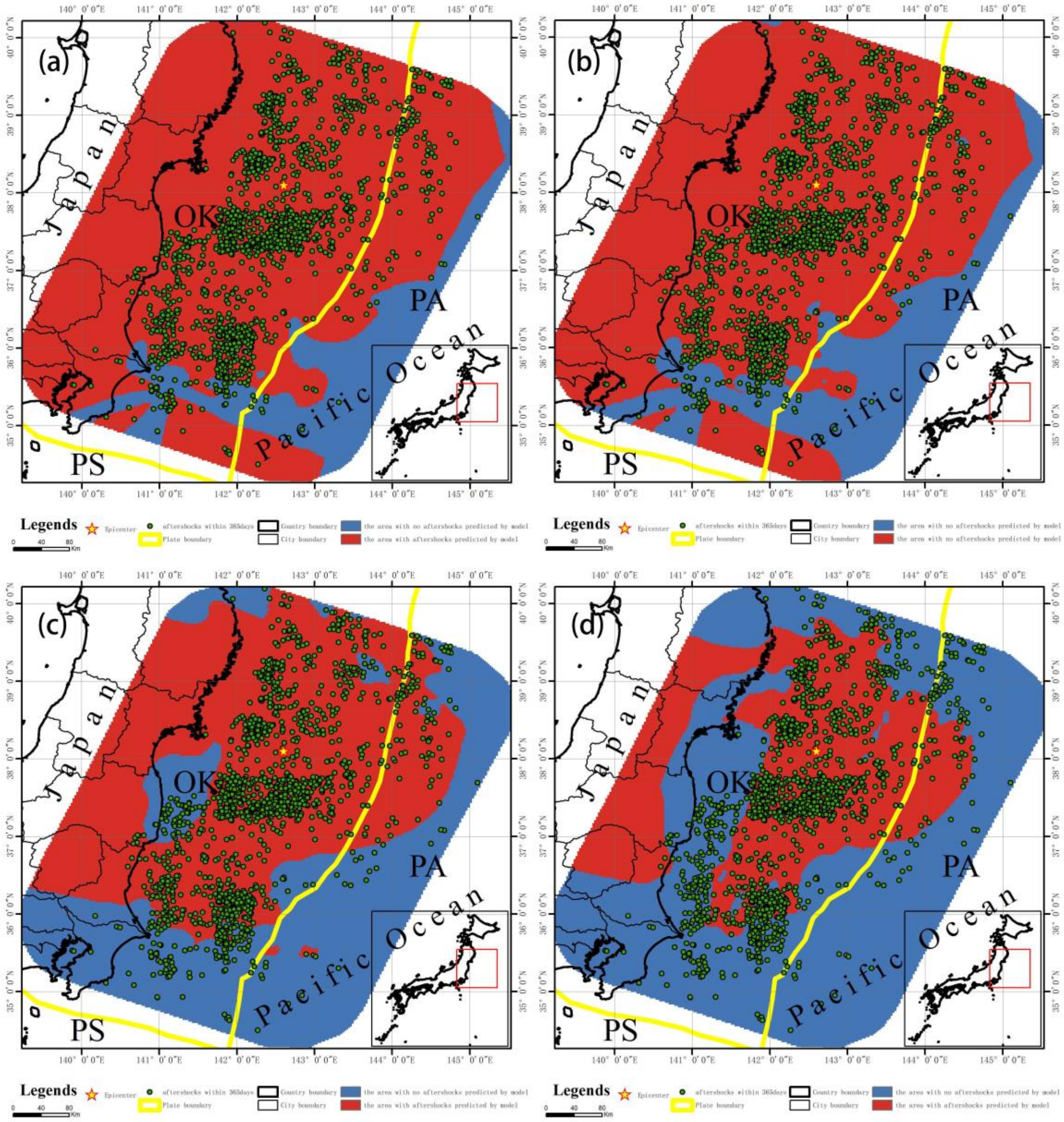

Figure 9 Tohoku earthquake hysteresis model prediction results under different thresholds (Figure a to d show the model prediction results when the threshold is 0.5, 0.6, 0.7, 0.8, respectively)

7. A complication regarding the Tohoku case is the Mw=7.9 aftershock https://web.archive.org/web/20110412003136/http://neic.usgs.gov/neis/bulletin/neic_iiav.html and its own aftershocks. That dataset should be added in Fig. 9 and compared with hysteresis model predictions.

Response: The M_w7.9 earthquake is considered as an aftershock of the Tohoku earthquake, and subsequent earthquakes are also considered as aftershocks of the Tohoku earthquake. In addition to the aftershock of magnitude 7.9, there are also several earthquakes with magnitude 7 or higher. (Toda S , Lin J , Stein R S . Using the 2011 M w 9.0 off the Pacific coast of Tohoku Earthquake to test the Coulomb stress triggering hypothesis and to calculate faults brought closer to failure[J]. Earth Planets & Space, 2011, 63(7):39.) These earthquakes are also considered as aftershocks of the earthquake.

8. It seems that the DNN shows a marked decay in the expansion rate at about 30-days following the mainshock. I suggest to investigate this pattern further and provide a more quantitative approach, i.e. is it 29, 30 or 31 days etc? is it a function of the magnitude of the mainshock?

Response: According to the modified Omori formula $\mathrm{n(t)} = \mathrm{K}(t + c)^{-p}$ and the K, c, and p values of the two earthquake cases. The specific attenuation formulas for two earthquake cases can be obtained. Take the Tohoku earthquake as an example, its K=300, c=0.42, p=0.8. The attenuation curve is respectively derived at 1day, 30days, 90days, 180days, and 365days, and the N(t) is obtained as the Table 3, and then the attenuation rate is calculated. The conclusion obtained is similar to the hysteresis model, and the decay rate is the fastest at the time scale of 30 days. When the time is 29 days or 31 days, the decay rate is not much different from 30 days. Since the derivative function of the attenuation curve is a decreasing function, its attenuation speed decreases with time. So far, it has not been found to be related to the strength of the main shock.

Table 3 The attenuation of aftershocks of the Tohoku earthquake on multiple time scales

	1Day	30Days	90Days	180Days	365Days
N（t）	127.67	75.80	50.61	40.83	33.58
Change Rate	/	0.406	0.332	0.193	0.178

Minor comments（RC2）：

1. Line 43: Stein et al., 1983 should be corrected to Stein and Lisowski,1983.
Response: Stein et al. in line 43 of the text have been corrected to Stein and Lisowski, 1983.

2. Line 50: the reference to Phoebe et al. should be corrected according to https://www.nature.com/articles/s41586-018-0438-y. The same comment applies to line 421 of the manuscript.
Response: The references at line 50 and line 421 have been corrected.

3. Line 65-67: These views treat the crust as a continuoum body while it is widely know that the crust is cross-cut by discontinuities/dislocations where tectonic strain is accumulating.
Response: The whole crust is indeed a discontinuity, but it can be approximated as a continuum in some local areas. The content of lines 65-67 in the text has been revised as follows. “When the stress intensity is greater than the bearing stress intensity of the crust, the crust will lose its stability. Discontinuous crust will produce displacement at the location of its fracture, forming an earthquake. Sometimes fracture surfaces are produced in some locally continuous areas. At the same time, the elastic strain energy stored in the earth's crust will be released in this process”.

4. Line 169: The paper by Yi et al. (2011) deals with the period before the Wenchuan mainshock.
Response: This reference is prone to ambiguity, and the citation of the Chinese reference has

been deleted at lines 168-169 in the article.

5. Line 282: ...Chengdu, Mianyang, Deyang, Guangyuan and Ngawa... Are these provinces? regions? towns?

Response: The "Chengdu", "Mianyang", "Deyang", "Guanyuan" and "Ngawa" mentioned in line 282 of the text are all Chinese city names, which are located around the epicenter of the Wenchuan earthquake.