We would like to thank Dr. Athanassios Ganas Athens and the anonymous referee for the thoughtful review of our manuscript and for giving these constructive comments and suggestions, which substantially helped us improve the quality of the paper. We are very grateful to the editors for their serious and responsible attitude towards the manuscript, which is very helpful to the timely revision of the manuscript.

Almost all the points that were raised have been adopted in the revised manuscript. We believe the new version of the manuscript has been significantly improved. We have revised the paper considering all the comments, which are shown below point-by-point. In addition, some minor errors have also been corrected in the text.

A marked-up manuscript version has been attached, and the contents revised has been marked in red.

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.

Reply: In the new version of the manuscript, we have added section 3.2 to present the data of two earthquake cases (line278-290). The added content is as follows:

"In order to verify the method and model in this article, we selected two typical historical earthquake cases, i.e., the Wenchuan earthquake and the Tohoku earthquake. These two earthquake cases are not included in the data used for model construction. They are characterized by large magnitude and a large number of aftershocks.

In the Tohoku earthquake case, there were 15,062 aftershocks in the study area within one year after the mainshock (Table 1). In the finite fault model used in this article, the focal depth is 20-25 km, and according to the depth distribution of aftershocks at multiple time scales, the number of aftershocks is the largest at the depth of 35-40 km (Figure 4a). In the Wenchuan earthquake case, there were 1,455 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 10-15 km. According to the depth distribution of aftershocks at multiple time scales, the number of aftershocks is the largest at 10-15 km depth. Aftershocks are not necessarily distributed the most on the focal depth surface (Figure 4b)."

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

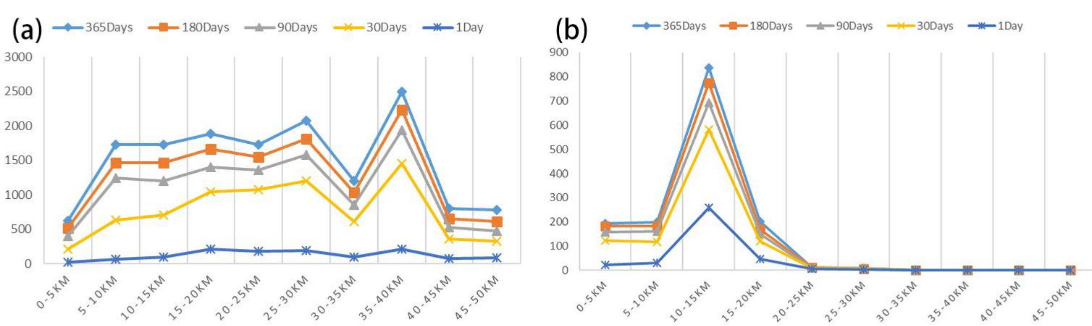

Figure 4. Multi-time scale aftershock depth distribution curves of (a) the Tohoku earthquake and (a) the Wenchuan earthquake.

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.

Reply: The following content has been added to lines 117-119 in the text:

"More precisely, all aftershock data 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."

The following content has been added to the section 4.2 in the text:

"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. To further prove the validity of the model, the actual location of the aftershock event is further used instead of the sub-cell location, and the threshold is set to 0.5. 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 17 and Figure 18). 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."

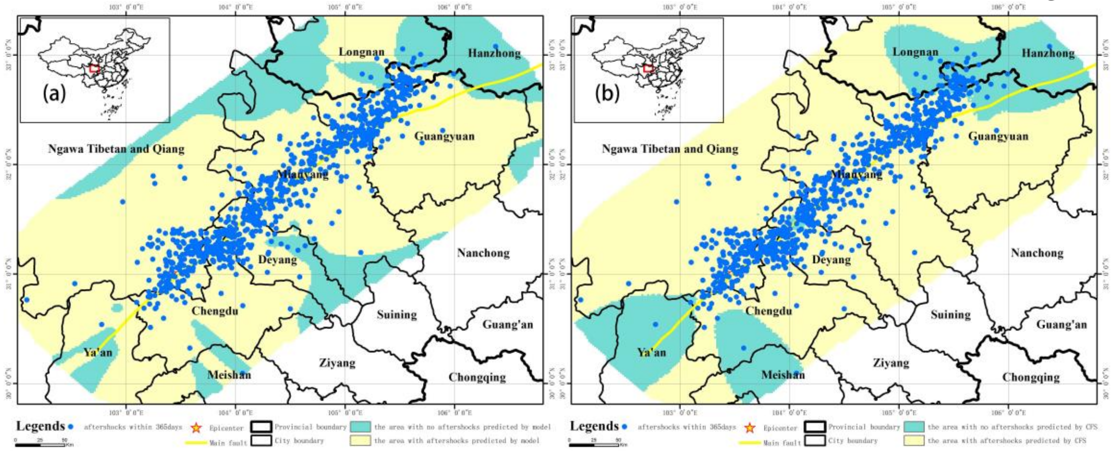

Figure 17. The hysteresis model prediction result and the △CFS prediction result of the Wenchuan earthquake. Figure (a) shows the hysteresis model prediction result and figure (b) shows the △CFS prediction result.

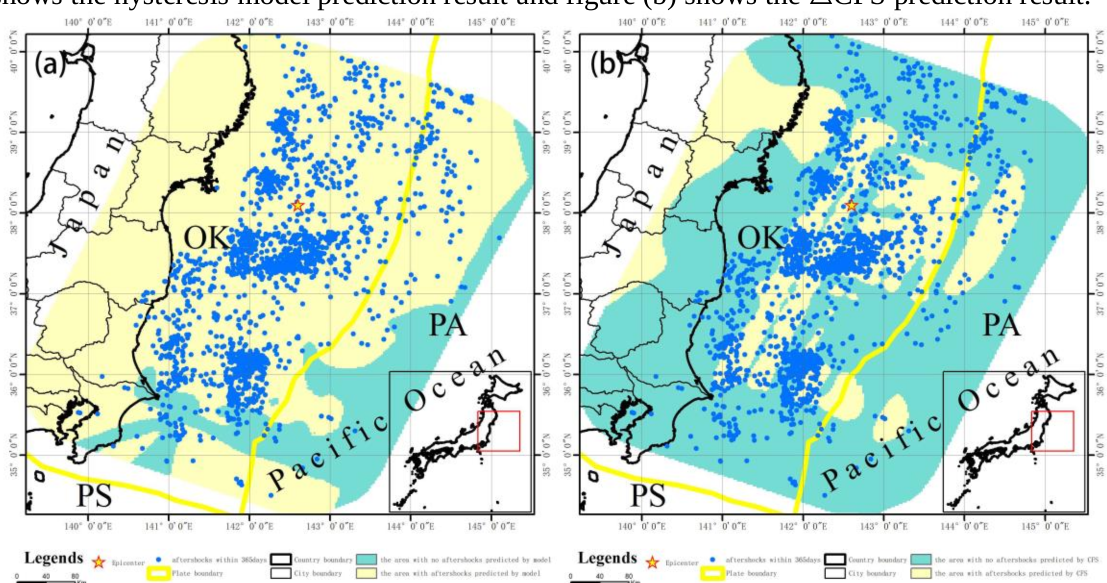

Figure 18. The hysteresis model prediction result and the △CFS prediction result of the Tohoku earthquake. Figure (a) shows the hysteresis model prediction result and figure (b) shows the △CFS prediction result.

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).

Reply: The following content has been added to lines 379-389 in the text:

"In order to analyze the model results better, we use the modified Omori formula to analyze the above two earthquake cases. 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 13 and Figure 14). 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	$N(t)=67.08(t+0.21)^{-1.07}$	$N'(t)=-71.78(t+0.21)^{-2.07}$

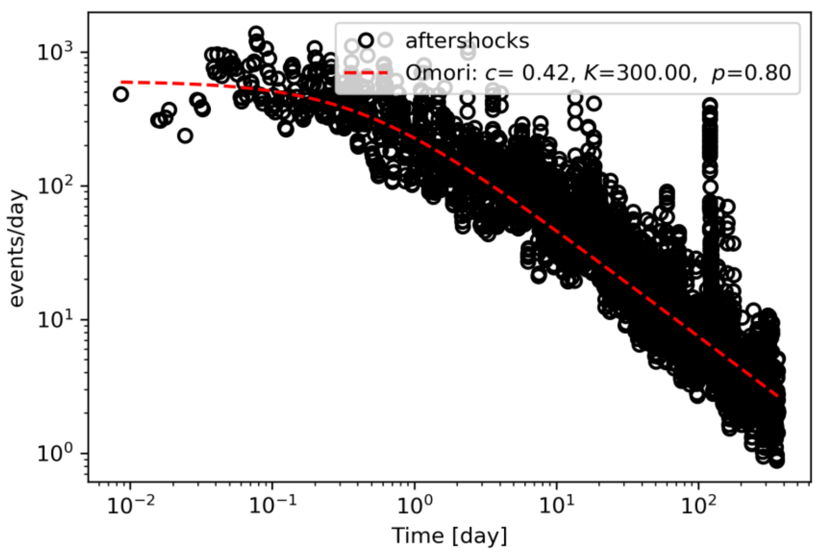

Figure 13. The modified Omori formula aftershock attenuation curve of Tohoku earthquake

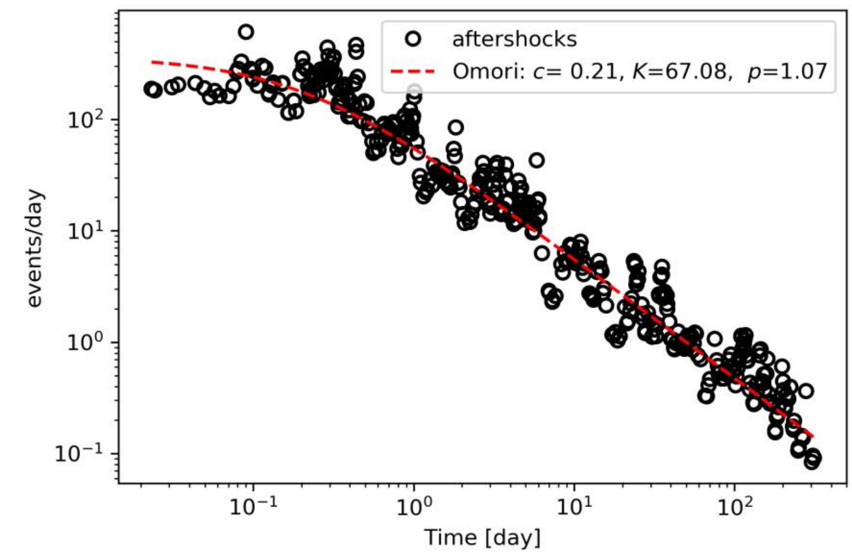

Figure 14. The modified Omori formula aftershock attenuation curve of Wenchuan earthquake

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.

Reply: The following content has been added to lines 347-351 in the text:

"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 9). 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."

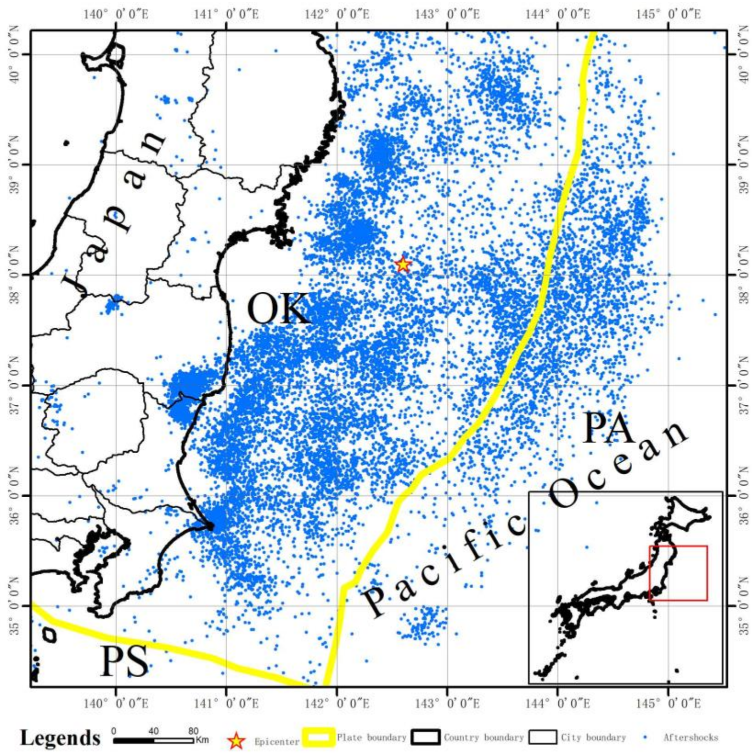

Figure 9. Aftershocks distribution of the Tohoku earthquake. The blue points in the figure are the projection positions of the aftershocks within a depth of 50 km.

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.

Reply: The following content has been added to lines 238-241 in the text:

"The prediction result given by the model is the approximate range of aftershocks, that is, the position of 5 km^3 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."

The following content has been added to lines 244-245 in the text:

"At locations close to the mainshock, the probability value predicted by the model is more likely to be greater than the threshold 0.5 set in the article."

The following content has been added to lines 246-249 in the text:

"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."

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

Reply: The following content has been added to section 4.1 (lines 413-425) in the text:

"Finally, a supplementary explanation is given to the phenomenon that the area predicted by the model is larger than the actual aftershock location. The prediction results of the hysteresis model are the likely locations of aftershocks at different time scales after the mainshock. At each location, the predicted value is a number between 0-1, which represents the probability of aftershocks that may occur at that location. 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 5 km^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 the prediction model, if the threshold increases, the predicted coverage area of aftershocks is gradually reduced, but as the increase of threshold, the local area prediction will also produce more errors and deviations. 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."

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.

Reply: 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?

Reply: According to the modified Omori formula $n(t) = K(t + c)^{-p}$ 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 1 d, 30 d, 90 d, 180 d, and 365 d, and the N(t) is obtained as the Table A, 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 d. When the time is 29 d or 31 d, the decay rate is not much different from 30 d. 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.

This comment and the third comment are combined together to be revised in Section 4.1. The following content has been added to section 4.1 (lines 389-391) in the text to explain the above:

"If the aftershock attenuation rate of each earthquake case is calculated on all time scales, it can be found that it gradually decreases as the time scale increases."

Table A 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.
Reply: Stein et al. at line 43 of the original 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.

Reply: The references at line 50 and line 421 of original manuscript have been corrected. The revised content is at line 52 and line 497 of the new version of the manuscript.

The revised references are as follows:

DeVries, P.M.R., Viégas, F., Wattenberg, M. and Meade, B.J.: Deep learning of aftershock patterns following large earthquakes, Nature, 560(7720), 632-634, https://doi.org/10.1038/s41586-018-0438-y, 2018.

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.

Reply: 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 original 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". The revised content is at lines 66-70 of the new version of the manuscript.

4. Line 169: The paper by Yi et al. (2011) deals with the period before the Wenchuan mainshock.

Reply: This reference is prone to ambiguity, and the citation of the Chinese reference has been deleted at lines 168-169 of the original manuscript.

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

Reply: The "Chengdu", "Mianyang", "Deyang", "Guanyuan" and "Ngawa" mentioned at line 282 of the text are all Chinese city names, which are located around the epicenter of the Wenchuan earthquake. The revised content is at line 314 of the new version of the manuscript.

Minor comments（RC1）：

1. Add the references Kapetanidis et al. (2015) and Papadimitriou et al. (2018) regarding the spatial and temporal evolution of earthquake sequences.

Reply: Corresponding references have been added at line 57 in the text.

Kapetanidis, V., Deschamps, A., Papadimitriou, P., Matrullo, E., Karakonstantis, A., Bozionelos, G., Kaviris, G., Serpetsidaki, A., Lyon-Caen, H., Voulgaris, N., Bernard, P., Sokos, E. and Makropoulos, K.: The 2013 earthquake swarm in Helike, Greece: Seismic activity at the root of old normal faults, Geophys. Journ. Int., 202, 2044–2073, https://doi.org/10.1093/gji/ggv249, 2015.

Papadimitriou, P., Kassaras, I., Kaviris, G., Tselentis, G.A., Voulgaris, N., Lekkas, E., Chouliaras, G., Evangelidis, C., Pavlou, K., Kapetanidis, V., Karakonstantis, A., Kazantzidou-Firtinidou, D., Fountoulakis, I., Millas, C., Spingos, I., Aspiotis, T., Moumoulidou, A., Skourtsos, E., Antoniou, V., Andreadakis, E., Mavroulis, S., and Kleanthi, M.: The 12th June 2017 Mw=6.3 Lesvos earthquake from detailed seismological observations, Journal of Geodynamics, 115, 23–42, 2018.

2. Page 2, Line 64: Add the references Kaviris et al. (2017) and Kaviris et al. (2018) regarding stress accumulation

Reply: Two articles on stress accumulation have been added to line 65 of the text.

Kaviris, G., Spingos, I., Kapetanidis, V., Papadimitriou, P., Voulgaris, N. and Makropoulos, K.: Upper crust seismic anisotropy study and temporal variations of shear-wave splitting parameters in the Western Gulf of Corinth (Greece) during 2013, Physics of the Earth and Planetary Interiors, 269, 148-164, https://doi.org/10.1016/j.pepi.2017.06.006, 2017.

Kaviris, G., Millas, C., Spingos, I., Kapetanidis, V., Fountoulakis, I., Papadimitriou, P., Voulgaris, N. and Makropoulos, K.: Observations of shear-wave splitting parameters in the Western Gulf of Corinth focusing on the 2014 Mw=5.0 earthquake, Physics of the Earth and Planetary Interiors, 282, 60-76, https://doi.org/10.1016/j.pepi.2017.06.006, 2018.

3. Page 3, Line 89: Add the reference Mai and Thingbaijam (2014) regarding SRCMOD.

Reply: A reference to SRCMOD has been added at line 89 in the text.

Mai, P.M., and Thingbaijam, K.K.S.: SRCMOD: An online database of finite-fault rupture models, Seismological Research Letters, 85(6), 1348-1357, https://doi.org/10.1785/0220140077, 2014.

4. Page 3, Line 104: Add the reference Bondár and Storchak (2011) regarding the ISC catalogue.
Reply: A reference to the ISC earthquake catalog has been added to line 106 of the text.
Bondár, I. and Storchak, D.A.: Improved location procedures at the International Seismological Centre, Geophys. J. Int., 186, 1220-1244, https://doi.org/10.1111/j.1365-246X.2011.05107.x, 2011.

5. What is FCNN? The authors must explain in detail what this acronym is.
Reply: The full name of FCNN is "fully convolutional neural network". This is an error in expression, and I am grateful to the reviewer for his comment here. The convolution operation is not used in the model structure in the article, and all "FCNN" have been replaced by "DNN".
The explanation of the abbreviation "DNN" has been mentioned in the 8th reply.

6. Page 4, Line 167: Add the reference Gutenberg and Richter (1944) regarding the G-R law.
Reply: A reference to the G-R law has been added at line 167 in the text.
Gutenberg, B. and Richter, C.F.: Frequency of earthquakes in California, Bull. Seism. Soc. Am., 4, 185–188, https://doi.org/10.1038/156371a0, 1944.

7. Page 5, Lines 168-169: "A low b value is related to a high stress background, i.e., 168 the b value is relatively low during strong aftershock activity (Yi et al., 2011)." The b value of aftershocks is widely known to have high values. The work you cite is in Chinese and obviously I am not able to read it. I suggest that the authors delete this sentence.
Reply: This reference is prone to ambiguity, and the citation of the Chinese reference has been deleted at lines 168-169 in the article.

8. Figure 1: What is DNN? The authors must explain in detail what this acronym is.
Reply: The following content has been added to lines 161-164 in the text:

The abbreviation DNN mentioned in Figure 1 has been explained in the text. The full name of DNN is "deep neural network". Neural network is based on the extension of the perceptron, and DNN can be understood as a neural network with many hidden layers. Multi-layer neural network and deep neural network actually refer to one thing. DNN is sometimes called Multi-Layer perceptron (MLP).

9. Page 5, Line 211: What is ROC? The authors must explain in detail what this acronym is.
Reply: The abbreviation ROC mentioned in line 211 has been explained in the text. The following content has been added to lines 217-219 in the text:

The full name of ROC is "receiver operating characteristic curve", which considers the results obtained under a variety of different criteria. In this article, the ROC curve can reflect the prediction results of the model under multiple thresholds.

Deep Learning of Aftershock Hysteresis Effect Based on Elastic Dislocation Theory

Jin Chen[1,2], Hong Tang[1,2], and Wenkai Chen[3]

[1] Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

[2]State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and In stitute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100875, China

[3] Institute of Lanzhou Earthquake Research, China Earthquake Administration, Gansu Lanzhou 730000, China

Correspondence to: Hong Tang (tanghong@bnu.edu.cn)

Abstract: This paper selects fault source models of typical earthquakes across the globe and uses a volume extending 100 km horizontally from each mainshock rupture plane and 50 km vertically as the primary area of earthquake influence for calculation and analysis. A deep neural network is constructed to model the relationship between elastic stress tensor components and aftershock state at multiple time scales, and the model is evaluated. Finally, based on the aftershock hysteresis model, the aftershock hysteresis effect of the Wenchuan earthquake in 2008 and Tohoku earthquake in 2011 is analyzed, and the aftershock hysteresis effect at different depths is compared and analyzed. The correlation between the aftershock hysteresis effect and the Omori formula is also discussed and analyzed. The constructed aftershock hysteresis model has a good fit to the data and can predict the aftershock pattern at multiple time scales after a large earthquake. Compared with the traditional aftershock spatial analysis method, the model is more effective and fully considers the distribution of actual faults, instead of treating the earthquake as a point source. The expansion rate of the aftershock pattern is negatively correlated with time, and the aftershock patterns at all time scales are roughly similar and anisotropic.

Key Words: Okada dislocation theory; DNN; aftershock hysteresis effect; aftershock pattern; the Omori formula.

1 Introduction

After the occurrence of strong earthquakes, there is often a large number of aftershocks, which constitute the aftershock sequence. The aftershocks can lead to new damage to the area affected by the main earthquake. Therefore, it is necessary to study aftershocks and stimulate further discussion. Stein et al. systematically discussed the influence of the static stress of the main earthquake on the spatial distribution of aftershocks (Stein and Lisowski, 1983). A large number of earthquake examples show that the change in Coulomb stress produced by the main earthquake is greater than 0.01 Mpa, which readily triggers aftershocks (Harris, 1998; Toda, 2003; Ma et al., 2005). In addition to the Coulomb failure stress change method, the deep learning method is a new emerging method that can address some questions of physical mechanism. The prediction of the aftershock sequence based on the stress state of the crustal medium is also problematic and is a focus of source physics. The neural network has the characteristics of a black box, which can avoid the complicated physical mechanisms when predicting the aftershock pattern (Brodi, 2001; Moustra et al., 2011). In 2018, Phoebe et al. proposed a deep neural network to study the spatial distribution of aftershocks following the main earthquake. A neural network classifier based on stress variation was designed by the authors to determine the possibility of a spatial distribution of aftershocks (DeVries et al., 2018). This idea combines traditional physical analysis mechanisms with data-driven machine learning mechanisms, which can improve our understanding of the complex physical mechanism of earthquakes. Kong et al. also analyzed its necessity (Kong et al., 2019).

The distribution of aftershocks is not only related to spatial changes but also to temporal changes (Kapetanidis et al. 2015; Papadimitriou et al. 2018), which may be related to the actual properties of the medium, i.e., the viscoelastic medium and the porous two-phase medium are closer to the actual geological medium than the elastic medium. The hysteresis effect of the viscoelastic medium on stress change, the effect of readjustment of pore fluid on stress change and other time-dependent medium properties are equally important to post-earthquake stress change, which is an issue that is receiving increasing attention in post-earthquake effects research. In the study of the propagation of a seismic wave and its focal mechanism, the earth medium is assumed to be a completely elastic body. Prior to the main earthquake, the crustal medium will be continuously deformed due to the long-term and slow action of tectonic stress. In the process of stress accumulation (Kaviris et al. 2017; Kaviris et al. 2018), the strain energy of the crustal medium will be accumulated continuously and be stored in the crust in the form of elastic strain energy. 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. Simultaneously, the elastic strain energy stored in the earth's crust will be released in this process. After the occurrence of the main earthquake, the source body and its surrounding medium will return to the steady state. However, because the main earthquake causes a sudden change in the stress state of the medium, the accumulated elastic strain energy in the entire stress field cannot be released completely at once, but it will continue to be accumulated in other areas, and it will ultimately be released in the form of an aftershock sequence. Therefore, there is a hysteresis effect between the aftershock and the main earthquake (Gu et al., 1979). Omori and Utsu proposed the time distribution formulas of aftershocks. However, the formulas are based on statistical significance, which cannot reflect the underlying reason for the change in aftershock distribution over time, and cannot spatialize the temporal change in aftershocks (Omori, 1894; Utsu, 1961). Many scholars also analyzed the spatial-temporal distribution characteristics of aftershocks by building a model. For example, the ETAS model proposed by Ogata (Ogata, 1988), the KJ model proposed by Kagan et al. (Kagan et al., 1994), and the model improved by Ogata based on ETAS (Ogata, 1998). In 2009, Wong et al. proposed a joint distribution model that parameterized the aftershock location based on the distance and relative angle between aftershocks and mainshocks (Wong et al., 2009). All the above spatial-temporal models of aftershocks are all based on point-source earthquakes, while the actual earthquake sources are faults. So the distribution of the main fault zone should be considered when predicting the aftershock pattern. Some spatial models also ignore the relative angle or distance between mainshock and aftershocks. These deficiencies are taken into account when building the new prediction model.

In this paper, a method based on deep neural networks is proposed to analyze the probability distribution of aftershocks following the main earthquake on multiple time scales, which indirectly reflects the hysteresis effect of aftershocks at different positions under the stress field of the main earthquake. The SRCMOD fault source model database and earthquake events are used as raw data (Mai and Thingbaijam, 2014). First, the analysis area of each main earthquake is gridded, and then the aftershocks of each main earthquake are entered into the grids. The DC3D displacement model is used to calculate the components of stress change tensor for each cell. Based on this grid, the results of the calculation are used as the input to train the neural network, and the aftershock hysteresis model is then obtained. As the application analysis cases for the model, the Wenchuan and the Tohoku earthquakes are not included in the training set or the validation set. Finally, the spatial distribution and expansion characteristics of the aftershock hysteresis model are obtained for both the horizontal and vertical directions. In addition, we focus on two important concepts, namely the "hysteresis effect" and the "aftershock pattern". The "hysteresis effect" refers to the change in spatial distribution of aftershocks with the change of time scale. The "aftershock pattern" refers to the spatial distribution of aftershocks at a certain time.

2 Data and Methods

2.1 Data

2.1.1 Raw data

Two types of data are used in this paper, SRCMOD finite fault data and the ISC (International Seismological Centre) earthquake catalogue (Bondár and Storchak, 2011).

The inversion of finite fault source data facilitates a better understanding of the complexity of the earthquake rupture process. Although the spatial resolution of the model is low, it can provide information on deep seismic slip and fault evolution over time. Therefore, the finite fault model is an important means to further study the mechanics and kinematics of the process of earthquake fracture. The online SRCMOD database provides the inversion results for many typical earthquakes from 1906 to present. These results are uploaded by seismologists globally after the main earthquake through inversion. Because the earth's crust is used as an elastic medium in the calculation of coseismic displacement stress, we don't consider the impact of the background of each earthquake. There are 19 finite fault source models used in this analysis: 15 are used as training data and 4 are used as validation data.

The aftershocks following each main earthquake are obtained from the International Earthquake Center (ISC). More precisely, all aftershock data 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. For the mainshock cases in this paper, the aftershocks within 1 d, 30 d, 90 d, 180 d, and 365 d and within a volume extending 100 km horizontally from each mainshock rupture plane and 50 km vertically are used for analysis of the aftershock sequences.

2.1.2 Data processing

After acquiring the limited fault source data and aftershock sequence data, it is necessary to process them to create the final data for analysis. First, the volume extending 100 km horizontally from each mainshock rupture plane and 50 km vertically is divided into a grid of 5 km^3 cubes. Five time scales of aftershock sequence data are then entered into each cube. The aftershock state of a cell with an aftershock is defined as 1 and that of a cell without an aftershock is defined as 0. The final training data has 15 aftershock sequences containing 318,210 subcells, and the validation data has 4 aftershock sequences containing 89,900 subcells.

2.2 Methods

2.2.1 Okada elastic dislocation theory

The inversion analysis of seismogenic faults after earthquakes is a popular topic in seismology, while in the process of inversion, the application of dislocation theory and models is essential. The dislocation model was first used to analyze fault movement in 1958 (Steketee, 1958). Steketee introduced the dislocation theory into the study of seismic deformation fields, and described the relationship between discontinuous displacement on the dislocation plane and the displacement field in an isotropic medium. Okada summarized the existing research in 1985 and proposed a formula for the calculation of displacement in an isotropic, uniform elastic half space. This formula can be used to calculate the coseismic deformation caused by any fault in the elastic half space (Okada, 1985, 1992). Okada dislocation theory systematically summarizes the relationship between point source dislocation and surface deformation caused by rectangular dislocation. The crustal movement is typically slow, and the crustal medium generally shows viscosity and plasticity over a long time scale. At present, the Okada dislocation theory is the most widely used dislocation theory and is often used in combination with InSAR technology. InSAR is used to monitor the surface coseismic deformation field, and the Okada theory is then used to conduct fault slip inversion (Shan et al., 2017; Wang et al., 2018; Cheng et al., 2019; Zhao, 2019).

Therefore, the Okada elastic dislocation theory is used to calculate the coseismic strain stress field of the main earthquake in the paper. The Okada elastic dislocation model, which ignores the influence of stratification in the earth's medium, is widely used in the study of coseismic deformation of the seismic signal source. Okada gives the analytical expression of the partial derivative $\frac{\partial u_i}{\partial x_j}(\mathrm{i,j}=1,2,3)$ of the displacement $\mathbf{u}$ of the finite fault plane in the elastic half space (Okada, 1992). This expression is used to obtain the strain tensor $\boldsymbol{\varepsilon}$ as

$$\varepsilon_{ij}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)$$

and the lamé constant in the linear solid medium is used to obtain the surrounding stress change tensor $\boldsymbol{\sigma}$ as

$$\sigma=\begin{bmatrix}\lambda tr[\varepsilon]+2\mu\varepsilon_{xx} & 2\mu\varepsilon_{xy} & 2\mu\varepsilon_{xz}\\ 2\mu\varepsilon_{yx} & \lambda tr[\varepsilon]+2\mu\varepsilon_{yy} & 2\mu\varepsilon_{yz}\\ 2\mu\varepsilon_{zx} & 2\mu\varepsilon_{zy} & \lambda tr[\varepsilon]+2\mu\varepsilon_{zz}\end{bmatrix}$$

where λ and μ are lamé constants. In this paper, the crustal medium is regarded as a Poisson body, and the two lamé coefficients are both 3.0×10^{10} Pa. The parameter $tr[\varepsilon]$ is the trace of strain tensor $\boldsymbol{\varepsilon}$.

2.2.2 DNN

To analyze the hysteresis effect of aftershocks, it is necessary to establish a model that can predict the damage modes of aftershocks at multiple time scales. We constructed a fully connected DNN to simulate the relationship between the change value of the elastic stress tensor and aftershock and to explain the hysteresis effect of aftershocks. The full name of DNN mentioned above is “deep neural network”. Neural network is based on the extension of the perceptron, and DNN can be understood as a neural network with many hidden layers. Multi-layer neural network and deep neural network actually refer to one thing. DNN is sometimes called Multi-Layer perceptron (MLP). The network established here is a network with six hidden layers. Except for the second hidden layer, which has 100 neurons, the other five hidden layers have 50 neurons. The input layer dimension of the entire network is 12. Its input eigenvalue is the combination of the absolute value of six independent components of the elastic stress at the center of each subunit and the negative number of the absolute value, for a total of 12 inputs.

Then we analyze the correlation between aftershocks and stress change, which is closely related to the inputs of DNN. At present, the research on aftershocks is primarily based on statistical methods, and the research content primarily focuses on the distribution of aftershock strength and time attenuation. The intensity distribution of aftershocks follows the G-R relationship $\log N = a - bM$, where M is the magnitude, N is the number of aftershocks with magnitude greater than or equal to M, and a and b are the scale coefficients (Gutenberg and Richter, 1944). The value of b generally varies from 0.6 to 1.1 (Utsu, 2002), and its value is related to the regional stress state (Mogi, 1962; Scholz, 1968).

The study of time attenuation of aftershocks begins with the statistical description of frequency attenuation characteristics of the aftershock sequence using the Omori formula (Omori, 1894). In 1961, Utsu et al. proposed that the frequency attenuation rate of the actual aftershock sequence is faster than that calculated by the Omori formula (Utsu, 1961) and proposed the modified Omori formula $n(t) = K(t + c)^{-p}$, where n(t) is the aftershock frequency per unit time, c is a constant, and p is the attenuation coefficient of the aftershock sequence. For a large number of aftershock sequences, the modified Omori formula accurately describes the time attenuation of aftershocks. In the modified Omori formula, the c value is related to the incomplete recording time after the main earthquake (Kagan et al., 2005), which can provide a physical explanation for the aftershock attenuation after the main earthquake (Lindman et al., 2005). This value is also related to the rupture mode of the main earthquake (Narteau et al., 2009), i.e., the aftershock attenuation is affected by the stress state and related to the stress change.

The stress change caused by the main earthquake can be calculated by the Coulomb fracture stress change, which is also the most widely used analytical method at present. The change in Coulomb stress produced by the main earthquake will trigger the stress of the following aftershocks (Harris, 1998). Some seismologists believe that if the change in Coulomb fracture stress is positive around the main earthquake, it will promote fault movement and trigger aftershocks; if the change in Coulomb fracture stress is negative, it will inhibit fault movement, and the probability of triggering an aftershock is reduced (Lin, 2004; Harris, 1998; Han, 2003). According to the research of DeVries et al., the Coulomb fracture stress change is an inadequate explanation for aftershocks, and the relationship between the positive and negative values of stress change and the triggering of aftershocks requires further exploration. DeVries et al. modeled the relationship between stress change and aftershock triggering by training a neural network (DeVries et al., 2018). The variation in Coulomb fracture stress depends on the geometric properties and coseismic dislocations of the source fault (King et al., 1994; Zhu et al., 2009). Therefore, the change value of the stress tensor, which is closely related to the dislocation of the same earthquake, can be used as the aftershock variable to build the model.

In addition, Meade et al. tested many stress-related indicators in 2017 to explain the influence of the coseismic stress field of the main earthquake on the location of aftershocks. Their results show that the sum of the absolute values of the six independent components of the stress tensor, the von Mises yield criterion and the maximum shear stress produce the best interpretation. These variables can be obtained by the combination of the absolute values of the six independent components of the stress tensor and the negative values of the absolute values. Therefore, these variables are also used as the network input (Meade et al., 2017; Mignan et al., 2019). The input components are expressed as $|\sigma_{xx}|$, $|\sigma_{xy}|$, $|\sigma_{xz}|$, $|\sigma_{yy}|$, $|\sigma_{yz}|$, $|\sigma_{zz}|$, $-|\sigma_{xx}|$, $-|\sigma_{xy}|$, $-|\sigma_{xz}|$, $-|\sigma_{yy}|$, $-|\sigma_{yz}|$ and $-|\sigma_{zz}|$. The dimension of the network output layer is 1 and the output value is the relative probability of aftershocks in each cell, which is between 0 and 1. The dropout layer is also set after each hidden layer. The dropout layer is set to alleviate the occurrence of over-fitting in the model training process, which can have a regularization effect. In addition, the activation function of each hidden layer in the network is a Relu function, and the optimizer is Adadelta. The activation function of the output layer is a sigmoid function, which maps variables between 0 and 1 (Figure 1). Five scales are analyzed in this paper. Five neural networks are constructed to train five submodels. Each submodel is independent from the others and does not affect the others.

2.2.3 Model evaluation metric

The ROC curve and the AUC are used to evaluate the model. The full name of ROC is "receiver operating characteristic curve", which considers the results obtained under a variety of different criteria. In this article, the ROC curve can reflect the prediction results of the model under multiple thresholds. The AUC (area under curve) is defined as the area enclosed by the coordinate axis under the ROC curve, and the value of the area cannot be greater than 1. Because the ROC curve is generally located above the straight line y = x, the AUC value range is between 0.5 and 1. Based on the AUC value, we can interpret the accuracy of the classifier. When AUC = 1, the classifier is essentially a perfect classifier, whereas when AUC = 0.5, the classifier is making a random assessment and the obtained model is nonsensical. For the training samples in this paper, there is a class imbalance between positive and negative samples. A characteristic of the ROC curve is that when the distribution of positive and negative samples in the test set changes, the ROC curve can remain unchanged. The closer the ROC curve is to the Y-axis and y = 1, i.e., the higher the AUC value of the classifier, the greater the classification accuracy. Generally, when the AUC is less than 0.6, the accuracy of the classifier is poor; when the AUC is less than 0.75, the accuracy of the classifier is moderate; and when the AUC is greater than 0.75, the accuracy of the classifier is good. The output of the model in this paper is a probability value between 0 and 1. When the ROC curve is used to evaluate the model, it is conducted at five time scales, and the model under each time scale is evaluated as a two-classification problem.

3 Results

3.1 Evaluation of the aftershock hysteresis model

The aftershock hysteresis model under multiple time scales is obtained by using the neural network to train the constructed training dataset. In this paper, five submodels are trained, and the final hysteresis model is composed of five submodels. The prediction result given by the model is the approximate range of aftershocks, that is, the position of 5 km^3 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. The output value of the neural network in each cell is binarized with a threshold value of 0.5. A cell with a predicted value greater than 0.5 is assigned as 1, and a cell with a predicted value less than 0.5 is assigned as 0. At locations close to the mainshock, the probability value predicted by the model is more likely to be greater than the threshold 0.5 set in the article.

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. Based on the trained aftershock hysteresis model, the aftershock patterns are predicted for the Wenchuan earthquake at multiple time scales, and the ROC curves are obtained for the different time scales. The AUC values of the five time scales are all above 0.8, in both the training and validation sets, and some are close to 0.9. The AUC values of the training set are all higher than those of the validation set for the different time scales. The neural network designed by DeVries et al (2018) is used for aftershock prediction. The AUC value of the training model on the validation set is 0.849 (Figure 2). In this paper, the AUC value of each submodel on the validation set is similar to the research results of DeVries et al (2018). Therefore, the model achieves good prediction results at different time scales.

For comparison, we forecast the aftershock location based on the static Coulomb failure stress change. Considering the influence of shear stress, normal stress and friction coefficient on the active fault plane, Coulomb failure stress change (ΔCFS) can be expressed as

$$\Delta\mathrm{CFS} = \Delta\tau + \mu\Delta\sigma$$

where μ is the apparent friction coefficient, $\Delta\sigma$ is the normal stress on the fault plane and $\Delta\tau$ is the shear stress in the direction of fault slip. Based on previous studies, the friction coefficient in this paper is 0.4 (King et al., 1994; Wan et al., 2004). Numerous studies have shown that aftershocks will occur when ΔCFS is greater than +0.01Mpa. In order to compare and analyze the output of DNN, we need to transform the ΔCFS to 0-1. Similar to the last layer of DNN, the variation function adopts a variant of the sigmoid function as follows

$$\Delta\mathrm{CFS}' = \frac{1}{(1 + e^{-10(\Delta CFS - 0.01)})}\ .$$

where $\Delta\mathrm{CFS}'$ represents the Coulomb failure stress change after sigmoid transformation. We know that the traditional sigmoid function is similar to the jump function. In the analysis process of this paper, 0.01Mpa is the threshold value to determine whether aftershocks are generated, so the parameter 0.01 in the formula is the translation coefficient, that is, the traditional sigmoid function shifts 0.01Mpa to the right. Parameter 10 is the zoom coefficient, which compresses the sigmoid function horizontally to make its shape approach the jump function as much as possible. When $\Delta\mathrm{CFS}$ is greater than 0.01Mpa, $\Delta\mathrm{CFS}'$ approaches 1 as much as possible, and when $\Delta\mathrm{CFS}$ is less than 0.01Mpa, $\Delta\mathrm{CFS}'$ approaches 0 as much as possible. Then we evaluated the results and calculated the AUC value on each time scale by the ROC curve. Compared with the results of the previous model, the AUC results obtained by the method based on static Coulomb failure stress change are generally poor, which are no more than 0.6 (Figure 3).

3.2 Case selection and data presentation

In order to verify the method and model in this article, we selected two typical historical earthquake cases, i.e., the Wenchuan earthquake and the Tohoku earthquake. These two earthquake cases are not included in the data used for model construction. They are characterized by large magnitude and a large number of aftershocks.

In the Tohoku earthquake case, there were 15,062 aftershocks in the study area within one year after the mainshock (Table 1). In the finite fault model used in this article, the focal depth is 20-25 km, and according to the depth distribution of aftershocks at multiple time scales, the number of aftershocks is the largest at the depth of 35-40 km (Figure 4a). In the Wenchuan earthquake case, there were 1,455 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 10-15 km. According to the depth distribution of aftershocks at multiple time scales, the number of aftershocks is the largest at 10-15 km depth. Aftershocks are not necessarily distributed the most on the focal depth surface (Figure 4b).

3.3 Application of the model to the Wenchuan earthquake

3.3.1 Aftershock hysteresis failure mode

According to the tectonic stress figure of the Wenchuan earthquake, the Wenchuan earthquake was located in the Longmenshan area in the border mountains east of the Qinghai Tibetan Plateau. The geological structure in this area is complex. The main Longmenshan fault zone is composed of a series of roughly parallel thrust faults. It is divided into a front mountain zone and a back mountain zone with the Yingxiu-Beichuan central fault as the boundary. From Northwest to Southeast, the main fault zone consists of the back mountain fault, the central fault and the front mountain fault. The main fault forming the Wenchuan earthquake is the Yingxiu-Beichuan central fault. According to the beachball plot of the focal mechanism solution in Figure 5, the strong aftershocks following the Wenchuan earthquake are mainly related to reverse or thrust faults under the action of compressive stress.

Based on the aftershock hysteresis model, the failure patterns of aftershocks are predicted at different time scales, and the section observation is conducted at a depth of 12.5 km (essentially at the same depth as the source). Combined with the focal mechanism solution analysis of strong aftershocks around the main fault zone, the aftershocks in this area are mainly caused by the NW-trending and SE-trending crustal compressive stress (Figure 6). The expansion of the aftershock hysteresis pattern is observed, which is generally distributed along the fault strike and extends along the trend line of the main fault. Within one day after the main earthquake at Wenchuan, there were aftershocks over a wide area. The location of the aftershocks is distributed along the fault zone, and the location of the aftershocks is basically distributed in the geographical space predicted by the model.

Finally, the spatial results of the hysteresis effect of the Wenchuan earthquake are obtained by synthesizing the damage modes of the aftershocks at multiple time scales (Figure 7). The location of the aftershocks is basically along the main fault, i.e., the Yingxiu-Beichuan central fault. The model predicts that aftershocks are mainly distributed in Chengdu, Mianyang, Deyang, Guangyuan and Ngawa cities, which is consistent with the actual location of the aftershocks. Over time, the area of aftershocks expands outwards, and the rate decreases gradually. Using the main aftershock sequence from the Wenchuan earthquake as an example, the aftershock hysteresis patterns at different time scales are similar, and the direction of outward expansion is basically perpendicular to the distribution direction of the previous time scale. Compared with the attenuation map of earthquake intensity, the spatial distribution map of aftershock attenuation can provide some reference for follow-up disaster prevention and mitigation work after a large earthquake. We can further understand the attenuation law of aftershocks, and attempt to extend its time attenuation from a statistical perspective to a spatial perspective.

3.3.2 Aftershock hysteresis patterns at different depths

At different focal depths, the aftershock hysteresis patterns will also change. The focal depth range of the aftershocks analyzed in this paper is 0-50 km. The aftershock hysteresis effect is analyzed by selecting sections with depths of 2.5 km, 7.5 km, 12.5 km, 17.5 km, 22.5 km, 27.5 km, 32.5 km, 37.5 km and 42.5 km. Many previous studies have shown that the seismogenic layers in Central and western China are located in the middle and upper layers of the crust at a depth of no more than 20 km (Zhao et al., 1995; Yang et al., 2003). The aftershocks with a focal depth within 20 km are widely distributed (Figure 8). When the focal depth exceeds 20 km, the area where the aftershocks are generated suddenly decreases with increasing depth until no aftershocks are observed. The focal depth of the largest aftershock distribution range is 12.5 km, which is in the same range as the focal depth of the main earthquake. In the middle and upper layers of the earth's crust, the shapes of the aftershock hysteresis patterns are generally similar at different time scales. Over time, the shape of the aftershock hysteresis pattern generally expands outward in a similar pattern as the previous timescale. However, when the focal depth exceeds a certain value, the hysteresis pattern of the aftershocks substantially changes. In this case, when the focal depth is greater than 20 km, the area predicted for aftershocks significantly decreases, and the evolution of the hysteresis pattern is also changed. Although the overall expansion direction is consistent with the main fault, the pattern is less regular and more random.

3.4 Application of the model to the Tohoku earthquake in 2011

3.4.1 Aftershock hysteresis failure mode

Japan is located in the Circum-Pacific seismic belt at the intersection of the Eurasian plate and the Pacific plate, which is an area with a frequent occurrence of global earthquakes. Due to the collision between the Pacific plate and the Eurasian plate, the Pacific plate is subducted under the Eurasian plate, thus forming the Japan Trench and the Japanese island arc. “OK” represents the Okhotsk plate, which is part of the Eurasian plate, “PA” refers to the Pacific plate, and “PS” refers to the Philippine Sea plate, which is also part of the Eurasian plate (Bird, 2003) (Figure 9). The epicenter of the earthquake was located in the subduction zone of the Japanese trench. 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. 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. Based on the aftershock hysteresis model, the aftershock patterns within 1 d, 30 d, 90 d, 180 d and 365 d after the main earthquake are predicted, and the section (22.5 km) at the focal depth of the main earthquake is selected for analysis (Figure 10). Using the Tohoku earthquake in Japan as an example, the greatest expansion of the aftershock distribution area is observed within 30 d. The shape of the aftershock patterns are similar at all time scales. The aftershock and the predicted aftershock patterns are distributed in an approximately North-South direction along the Japan Trench and plate boundary.

The aftershock hysteresis model of the Tohoku earthquake in is obtained by synthesizing the aftershock patterns at different time scales (Figure 11). Over time, the expansion rate of the aftershock pattern gradually decreases, and the expansion direction is basically perpendicular to the aftershock pattern at the previous scale. Most of the aftershocks of this earthquake occurred in the eastern Sea of Japan and the area of concentrated terrestrial aftershocks was located in Fukushima.

3.4.2 Aftershock hysteresis patterns at different depths

Similar to the Wenchuan earthquake, the aftershock hysteresis pattern of the Tohoku earthquake changes with the change in depth. The magnitude of the earthquake was very large, reaching over M_w9. The main earthquake has a great impact on the surrounding area and the crust, which stores considerable energy, then releases it in the form of aftershocks. The predicted expansion direction of the aftershock model is generally consistent with that of the plate boundary and the Japan Trench. In this study, the maximum analysis depth is 50 km. Using the depth section of the main shock source as the center, the actual aftershock pattern does not change significantly when the depth change is small. This may be due to the large magnitude of the earthquake. The area of the actual aftershock pattern is reduced at a depth of 42.5 km. However, the location of the aftershocks is still widely distributed. The expansion of the aftershock pattern also changes beginning at a depth of 27.5 km. The general direction of distribution is along the trench, and some areas begin to expand vertically along the trench (Figure 12).

4 Discussion

4.1 Aftershock hysteresis effect

The modified Omori formula is $\mathrm{n(t)} = \mathrm{K}(t+c)^{-p}$, where n(t) is the aftershock frequency per unit time, and as *t* increases, n(t) will decrease correspondingly to describe the time attenuation characteristics of aftershocks. In order to analyze the model results better, we use the modified Omori formula to analyze the above two earthquake cases. 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 13 and Figure 14). 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. If the aftershock attenuation rate of each earthquake case is calculated on all time scales, it can be found that it gradually decreases as the time scale increases.

Compared with the Omori formula, the aftershock hysteresis effect analyzed in this paper can be reflected by the correlation between the change of time scale and the region of aftershocks. Based on the discussion of focal depth sections of the main earthquake, within one day after the Wenchuan earthquake, the number of subunits with aftershocks is 213; within 30 d, it is 386, representing an increase of 81.6%; within 90 d, it is 432, representing an increase of 11.9%; within 180 d, it is 466, representing an increase of 7.9%; and within 365 d, it is 488, representing an increase of 4.7%. Within one day after the Tohoku earthquake, the number of subunits with aftershocks was 137; within 30 d, it was 595, representing an increase of 334%; within 90 d, it was 724, representing an increase of 21.7%; within 180 d, it was 799, representing an increase of 10.4%; and within 365 d, it was 856, representing an increase of 7.1%. The aftershock pattern predicted by the model expands over time, but the expansion speed of the aftershock pattern also gradually decreases. The rate of expansion is most rapid 30 d after the earthquake. After 30 d, the speed decreases significantly from 30 to 90 d. The aftershock pattern of the Wenchuan earthquake expanded at a speed of 28.7 units/day within 30 d after the earthquake and then rapidly dropped to 7.8 units/day. The aftershock pattern of the Tohoku earthquake in Japan expanded at a rate of 38.6 units/day within 30 d after the earthquake and then dropped rapidly to 7.3 units/day. According to the correlation curve in Figure 15 and Figure 16, the aftershock hysteresis effect is reflected by the expansion pattern of the aftershocks. Combined with the comprehensive analysis of the previous two earthquake cases, the expansion rate of the aftershock hysteresis effect is $v \propto n(t) \propto \frac{1}{t}$. Unlike previous research on the attenuation law of aftershocks based on statistics (Narteau et al., 2005; Nanjo et al., 2007), this paper starts from another perspective, namely, spatial distribution, and returns to the discussion of the attenuation law of aftershock spatial distribution.

Finally, a supplementary explanation is given to the phenomenon that the area predicted by the model is larger than the actual aftershock location. The prediction results of the hysteresis model are the likely locations of aftershocks at different time scales after the mainshock. At each location, the predicted value is a number between 0-1, which represents the probability of aftershocks that may occur at that location. 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 5 km^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 the prediction model, if the threshold increases, the predicted coverage area of aftershocks is gradually reduced, but as the increase of threshold, the local area prediction will also produce more errors and deviations. 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.

4.2 Comparative analysis of prediction models

The widely used temporal-magnitude earthquake generation model (ETAS) was proposed by Ogata (Ogata, 1988). Later on, he observed that the distribution of aftershock sequences tended to be elliptic rather than circular. He established the anisotropic aftershock attenuation function and took the normal distribution as the spatial distribution model of aftershock (Ogata, 1998). It is a widely observed fact that aftershocks usually occur on or near the fault of mainshock. However, the normal distribution model does not include the source mechanism information of the mainshock when predicting the aftershock mode. Kagan et al. introduced the anisotropy function of the spatial smooth core into the long-term earthquake prediction, and established the spatial smooth core model, including the source mechanism information of the mainshock (Kagan et al., 1994). However, the above models ignore the internal relationship of the relative distance or direction between the mainshock and the aftershocks. Based on this, Wong et al. proposed a joint distribution model to parameterize the aftershock location according to the distance and relative angle between the mainshock and aftershocks (Wong et al., 2009). In the prediction process of the above models, the epicentre of the mainshock is used as a point source for analysis. Actually, the distribution of the fault plane of the mainshock should be fully considered. Based on the finite fault model, the distribution information of the main fault is considered in the paper. At the same time, the relative position and direction between the mainshock and aftershocks have been considered in the process of calculating the variation of stress tensor by using the Okada dislocation theory. Therefore, in the process of model training and learning, the relative position relation is also identified. Compared with the static Coulomb failure stress change method, the aftershock hysteresis model has a better prediction effect.

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. To further prove the validity of the model, the actual location of the aftershock event is further used instead of the sub-cell location, and the threshold is set to 0.5. 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 17 and Figure 18). 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.

In addition, the model is a six layer neural network, which is a black box model. Compared with the traditional statistical model or physical model, is the deep learning model more complex? We think this complexity is relative. In fact, the starting point of the traditional model and of the model in this paper are similar. They are all based on data, trying to find a relationship between some basic physical quantities and aftershocks. The complexity of traditional models lies in the process of finding such a connection. The complexity of deep learning model lies in its seemingly complex structure. The complex structure will lead to the increase of the number of internal variables to be learned, and the rapid computing ability of today's computers can solve this problem, so as to reduce manpower and time-consuming. In addition, the deep learning model is a data-driven method. It will be more convenient than the traditional model when the data set or the amount of data changes greatly or the model needs to be adjusted.

5 Conclusions

In this paper, based on the criterion of correlation between aftershocks and stress changes caused by the main earthquakes, a deep neural network is trained using the SRCMOD finite fault data and the ISC earthquake catalogue and is used to construct an aftershock hysteresis model. Using the main aftershock sequences of the Wenchuan and the Tohoku earthquakes as examples, the characteristics of the aftershock hysteresis effect in plane space and at different depths are then analyzed. The main contributions are as follows:

(1) The trained model of aftershock hysteresis is accurate. It can predict the aftershock patterns at multiple time scales after a large earthquake and produce a spatial distribution map of the aftershock hysteresis effect. Compared with static Coulomb failure stress change, this model is more effective.

(2) Compared with the traditional aftershock spatial analysis method, the model fully considers the distribution of actual faults in the prediction of aftershock pattern, instead of treating the earthquake as a point source. In the analysis of the model, the relative position information between the mainshock and aftershocks has been included.

(3) The expansion rate of the aftershock patterns changes over time, i.e., $v \propto n(t) \propto \frac{1}{t}$. In the middle and upper layers of the crust, the shape of the aftershock pattern is generally consistent, and the expansion direction is typically perpendicular to the direction of distribution of the previous time scale.

(4) According to the prediction results of the model, the aftershock patterns at all time scales are roughly similar and anisotropic. The distribution law of aftershock hysteresis effect will change with the increase of the depth.

In the analysis of each aftershock sequence, we only consider the influence of the main earthquake fault zone. If we comprehensively consider the stress field superposition of multiple or all faults in the analysis area of each earthquake case, the prediction of the aftershock pattern will be more accurate. In addition, we focus on the location of the aftershocks and will further explore and study aftershocks from the perspectives of magnitude and energy in the future.

Acknowledgments

The SRCMOD finite fault model data can be obtained from http://equake-rc.info/srcmod/. The ISC earthquake events can be obtained from http://www.isc.ac.uk/iscgem/. The results of the aftershock hysteresis effect are obtained by programming in Python, and some codes refer to previous research by DeVries et al. (2018). In this study, the figures and the subsequent processing of the results are all performed using the ArcGIS software. This work was supported by the National Natural Science Foundation of China (No. 41971280) and the National Key R&D Program of China (No. 2017YFB0504104).

Author Contributions

Conceptualization, H.T.; Funding acquisition, H.T.; Investigation, J.C.; Methodology, J.C. and H.T.; Software and Code, J.C.; Manuscript editing, J.C.; Manuscript revision, W K. C. and H.T. Supervision, H.T.

Competing Interests

The authors declare no conflict of interest.

References

Bird, P.: An updated digital model of plate boundaries, Geochemistry, Geophysics, Geosystems, 4(3), https://doi.org/10.1029/2001GC000252, 2003.

Bodri, B.: A neural-network model for earthquake occurrence, Journal of Geodynamics, 32(3), 289-310, https://doi.org/10.1016/S0264-3707(01)00039-4, 2001.

Bondár, I. and Storchak, D.A.: Improved location procedures at the International Seismological Centre, Geophys. J. Int., 186, 1220-1244, https://doi.org/10.1111/j.1365-246X.2011.05107.x, 2011.

Cheng, D., Zhang, Y., and Wang, X.: Coseismic deformation and fault slip inversion of the 2017 Mw7.3 Halabjah, Iraq, earthquake based on Sentinel-1A data, Acta Seismologica Sinica, 41(4), 484-493, https://doi.org/10.11939/jass.20180113, 2019.

DeVries, P.M.R., Viégas, F., Wattenberg, M. and Meade, B.J.: Deep learning of aftershock patterns following large earthquakes, Nature, 560(7720), 632-634, https://doi.org/10.1038/s41586-018-0438-y, 2018.

Dieterich, J.: A constitutive law for rate of earthquake production and its application to earthquake clustering, Journal of Geophysical Research, 99(B2), 2601, 1994.

Gu, J., Xie, X., and Zhao, L.: On temporal distribution of large aftershocks of the sequence of a major earthquake and preliminary theoretical explanation, Acta Geophysica Sinica, 22(1), 32-46, 1979.

Gutenberg, B. and Richter, C.F.: Frequency of earthquakes in California, Bull. Seism. Soc. Am., 4, 185–188, https://doi.org/10.1038/156371a0, 1944.

Han, Z.: Possible reduction of earthquake hazard on the Wellington Fault, New Zealand, after the nearby 1855, M8.2 Wairarapa earthquake and implication for interpreting paleoearthquake intervals, Annals of Geophysics, 46(5), 1141-1154, https://doi.org/10.4401/ag-3450, 2003.

Harris, R. A.: Introduction to special section: stress triggers, stress shadows, and implications for seismic hazard. Journal of Geophysical Research, 103(B10), 24347, https://doi.org/10.1029/98JB01576, 1998.

Jordan, MI., and Mitchell, T.M.: Machine learning: Trends, perspectives, and prospects, Science, 349(6245), 255-260, https://doi.org/10.1126/science.aaa8415, 2015.

Kagan, Y.Y., and Heidi, H.: Relation between mainshock rupture process and Omori's law for aftershock moment release rate, Geophysical Journal International, 163(3), 1039-1048, https://doi.org/10.1111/j.1365-246X.2005.02772.x, 2005.

Kagan, Y.Y., and Jackson, D.D.: Long-term probabilistic forecasting of earthquakes, Journal of Geophysical Research, 99(B7), 13685-13700, 1994.

Kapetanidis, V., Deschamps, A., Papadimitriou, P., Matrullo, E., Karakonstantis, A., Bozionelos, G., Kaviris, G., Serpetsidaki, A., Lyon-Caen, H., Voulgaris, N., Bernard, P., Sokos, E. and Makropoulos, K.: The 2013 earthquake swarm in Helike, Greece: Seismic activity at the root of old normal faults, Geophys. Journ. Int., 202, 2044–2073, https://doi.org/10.1093/gji/ggv249, 2015.

Kaviris, G., Spingos, I., Kapetanidis, V., Papadimitriou, P., Voulgaris, N. and Makropoulos, K.: Upper crust seismic anisotropy study and temporal variations of shear-wave splitting parameters in the Western Gulf of Corinth (Greece) during 2013, Physics of the Earth and Planetary Interiors, 269, 148-164, https://doi.org/10.1016/j.pepi.2017.06.006, 2017.

Kaviris, G., Millas, C., Spingos, I., Kapetanidis, V., Fountoulakis, I., Papadimitriou, P., Voulgaris, N. and Makropoulos, K.: Observations of shear-wave splitting parameters in the Western Gulf of Corinth focusing on the 2014 Mw=5.0 earthquake, Physics of the Earth and Planetary Interiors, 282, 60-76, https://doi.org/10.1016/j.pepi.2017.06.006, 2018.

King, G.C.P., Stein, R.S., and Lin, J.: Static stress changes and the triggering of earthquakes, Bulletin of the Seismological Society of America, 84(3), https://doi.org/10.1029/94JB00611, 1994.

Kong, Q., Trugman, D. T., and Ross, Z. E.: Machine Learning in Seismology: Turning Data into Insights, Seismological Research Letters, 90(1), 3-14, https://doi.org/10.1785/0220180259, 2019.

Lecun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521(7553), 436, https://doi.org/10.1038/nature14539, 2015.

Lin, J.: Stress triggering in thrust and subduction earthquakes and stress interaction between the southern San Andreas and nearby thrust and strike-slip faults, Journal of Geophysical Research, 109(B2),B02303, https://doi.org/10.1029/2003jb002607, 2004.

Lindman, M., Jonsdottir, K., and Roberts, R.: Earthquakes descaled: on waiting time distributions and scaling laws, Physical Review Letters, 94(10), 108501, https://doi.org/10.1103/PhysRevLett.94.108501, 2005.

Ma, K. F., Chan, C.H., and Stein, R.S.: Response of seismicity to Coulomb stress triggers and shadows of the 1999 Mw=7.6 Chi-Chi, Taiwan, earthquake, Journal of Geophysical Research Solid Earth, https://doi.org/10.1029/2004JB003389, 2005.

Mai, P.M., and Thingbaijam, K.K.S.: SRCMOD: An online database of finite-fault rupture models, Seismological Research Letters, 85(6), 1348-1357, https://doi.org/10.1785/0220140077, 2014.

Meade, B. J., Devries, P. M. R., and Faller, J.: What Is Better Than Coulomb Failure Stress? A Ranking of Scalar Static Stress Triggering Mechanisms from 105 Mainshock-Aftershock Pairs, Geophysical Research Letters, 44, https://doi.org/10.1002/2017GL075875, 2017.

Mignan, A., and Broccardo, M.: A Deeper Look into 'Deep Learning of Aftershock Patterns Following Large Earthquakes': Illustrating First Principles in Neural Network Physical Interpretability, IWANN 2019: Advances in Computational Intelligence, 3-14, https://doi.org/10.1007/978-3-030-20521-8_1, 2019.

Mogi, K.: On the time distribution of aftershocks accompanying the recent major earthquake in and near Japan, Bull Earthquake Res Inst, 40:107-124, 1962.

Moustra, M., Avraamides, M., and Christodoulou, C.: Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals, Expert Systems with Applications, 38(12), 15032-15039, https://doi.org/10.1016/j.eswa.2011.05.043, 2011.

Nanjo, K. Z., Enescu, B., and Shcherbakov, R.: Decay of aftershock activity for Japanese earthquakes, Journal of Geophysical Research Solid Earth, 112(B8), https://doi.org/10.1029/2006JB004754, 2007.

Narteau, C., Byrdina, S., and Shebalin, P.: Common dependence on stress for the two fundamental laws of statistical seismology, Nature, 462(7273), 642-645, https://doi.org/10.1038/nature08553, 2009.

Narteau, C., Shebalin, P., and Holschneider, M.: Onset of power law aftershock decay rates in southern California, Geophysical Research Letters, 32(22), L22312, https://doi.org/10.1029/2005gl023951, 2005.

Ogata, Y.: Statistical models for earthquake occurrences and residual analysis for point processes, JASA, 83(401), 9-27, 1988.

Ogata, Y.: Space-time point-process models for earthquake occurrences, Annals of the Institute of Statistical Mathematics, 50(2), 379-402, 1998.

Okada, Y.: Surface deformation due to shear and tensile faults in a half-space, Bull. Seismol. Soc. Amer., 75, 1135-1154, https://doi.org/10.1016/0148-9062(86)90674-1, 1985.
Okada, Y.: Internal deformation due to shear and tensile faults in a half-space, Bull. Seismol. Soc. Amer., 82(2), 1018-1040, 1992.
Omori, F.: On aftershocks of earthquakes, J. Coll. Sci. Imp. Univ. Tokyo. 7, 11-200, 1894.
Papadimitriou, P., Kassaras, I., Kaviris, G., Tselentis, G.A., Voulgaris, N., Lekkas, E., Chouliaras, G., Evangelidis, C., Pavlou, K., Kapetanidis, V., Karakonstantis, A., Kazantzidou-Firtinidou, D., Fountoulakis, I., Millas, C., Spingos, I., Aspiotis, T., Moumoulidou, A., Skourtsos, E., Antoniou, V., Andreadakis, E., Mavroulis, S., and Kleanthi, M.: The 12th June 2017 Mw=6.3 Lesvos earthquake from detailed seismological observations, Journal of Geodynamics, 115, 23–42, 2018.
Scholz, C.H.: Microfracturing and the inelastic deformation of rock in compression, Journal of Geophysical Research, 73(4), 1417-1432, https://doi.org/10.1029/jb073i004p01417, 1968.
Shan, X., Qu, C., and Gong, W.: Coseismic deformation field of the Jiuzhaigou Ms7.0 earthquake from Sentinel-1A InSAR data and fault slip inversion, Chinese Journal of Geophysics, 60(12), 4527-4536, https://doi.org/10.6038/cjg20171201, 2017.
Stein, R. S., and Lisowski, M.: The 1979 homestead valley earthquake sequence, California: control of aftershocks and postseismic deformation, Journal of Geophysical Research Solid Earth, 88(B8), https://doi.org/10.1029/JB088iB08p06477, 1983.
Steketee, J. A.: Some geophysical applications of the elasticity theory of dislocations. Canadian Journal of Physics, 36(9), 1168-1198, https://doi.org/10.1139/p58-123, 1958.
Toda, S.: Toggling of seismicity by the 1997 Kagoshima earthquake couplet: A demonstration of time-dependent stress transfer, Journal of Geophysical Research, 108(B12), 2567, https://doi.org/10.1029/2003jb002527, 2003.
Utsu, T.: A statistic study on the occurrence of aftershocks, Geophysics, 30, 521-545, 1961.
Utsu, T.: 43 Statistical features of seismicity, International Geophysics, 81, 719-732, https://doi.org/10.1016/S0074-6142(02)80246-7, 2002.
Wan, Y., Wu Z., and Zhou G.: Focal mechanism dependence of static stress triggering of earthquakes, Tectonophysics, 390(1-4), 235-243, 2004.
Wang, Z., Zhang, R., and Wang, X.: InSAR coseismic deformation monitoring and fault inversion of 2016 Qinghai Menyuan Earthquake, Remote Sensing Information, 33(006), 103-108, https://doi.org/10.3969/j.issn.1000-3177.2018.06.015, 2018.
Wong, K., and Schoenberg, F. P.: On mainshock focal mechanisms and the spatial distribution of aftershocks, Bulletin of the Seismological Society of America, 99(6), 3402-3412, 2009.
Yang, Z., Chen, Y., and Zheng, Y.: The application of double differential seismic location method in the accurate location of earthquakes in central and western China, Science in China(Series D), 33(z1), https://doi.org/10.1360/03dz0014, 2003.
Zhu, H., and Wen, X.: Stress triggering process if the 1973 to 1976 Songpan, Sichuan, sequence of strong earthquake, Chinese J. Geophys, 52(4), 994-1003, https://doi.org/10.3969/j.issn.0001-5733.2009.04.016, 2009.
Zhao, J.: Application of InSAR technology in surface deformation monitoring, M.S. thesis, Nanjing University, China, 2019.
Zhao, Z., and Chen, N.: Accuracy amendment of focal location along the Longmenshan fault in the north of Sichuan, Earthquake Research in Sichuan, 4, 19-30, 1995.

Tables

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

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	$N(t)=67.08(t+0.21)^{-1.07}$	$N'(t)=-71.78(t+0.21)^{-2.07}$

Figures

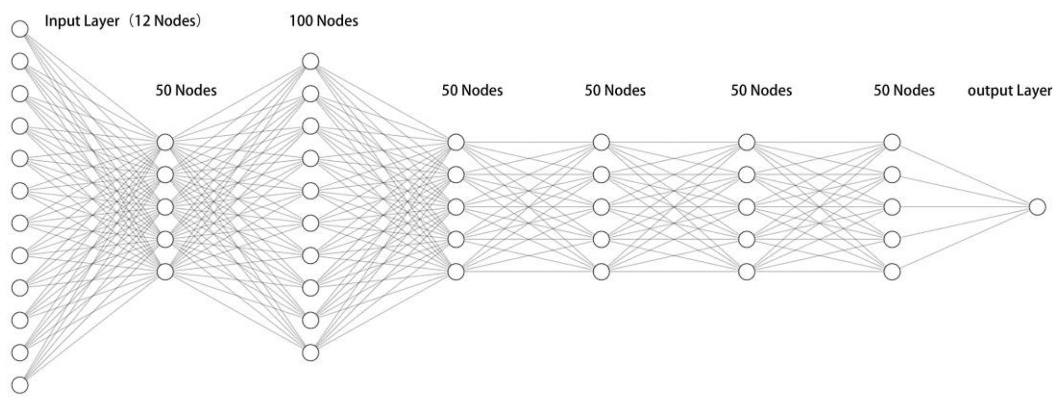

Figure 1. The structure of DNN. The neural network is composed of an input layer, hidden layers, output layers and the connections between each layer. The function of each hidden layer is to transform the features of the network input.

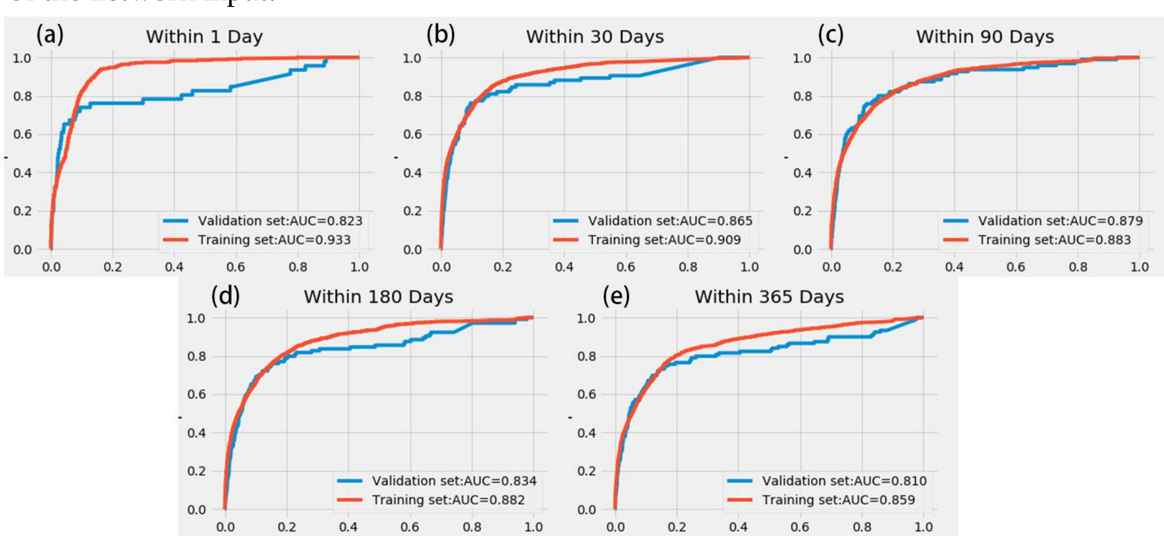

Figure 2. ROC curve for multiple time scales. Figures (a) through (e) show the ROC curve of the model within 1 d, 30 d, 90 d, 180 d and 365 d, respectively. The horizontal axis represents the FPR (false positive rate) and the vertical axis represents the TPR (true positive rate).

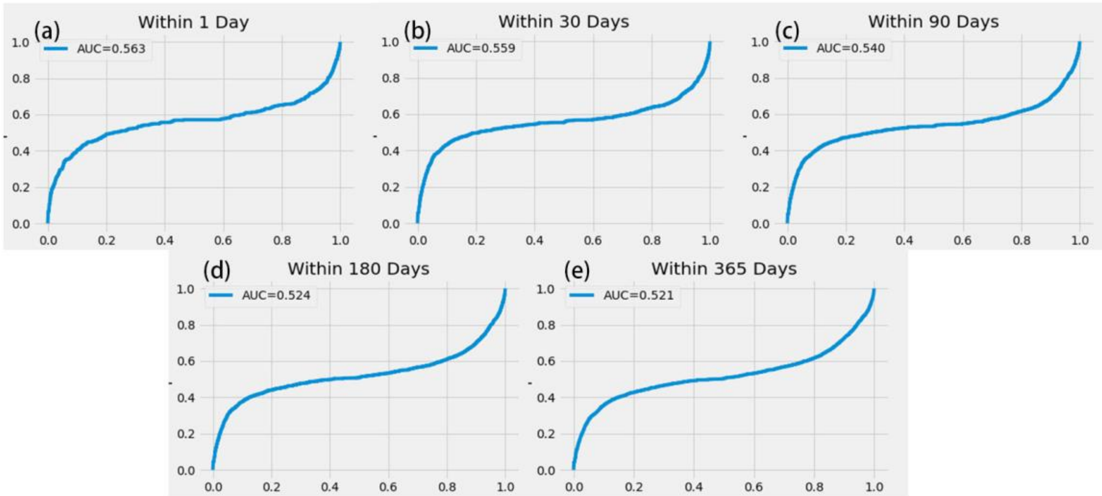

Figure 3. ROC curve of △CFS for multiple time scales. Figures (a) through (e) show the ROC curve of the model within 1 d, 30 d, 90 d, 180 d and 365 d, respectively. The horizontal axis represents the FPR (false positive rate) and the vertical axis represents the TPR (true positive rate).

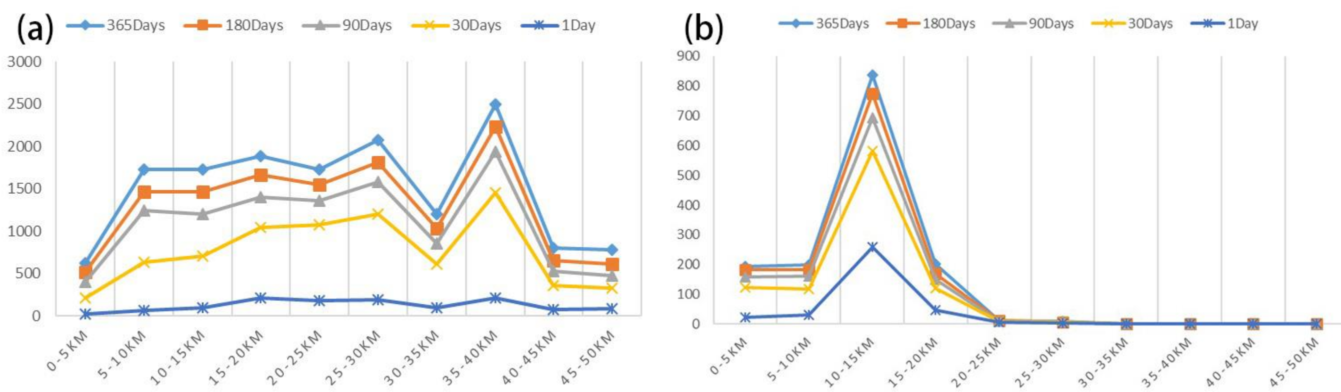

Figure 4. Multi-time scale aftershock depth distribution curves of (a) the Tohoku earthquake and (a) the Wenchuan earthquake.

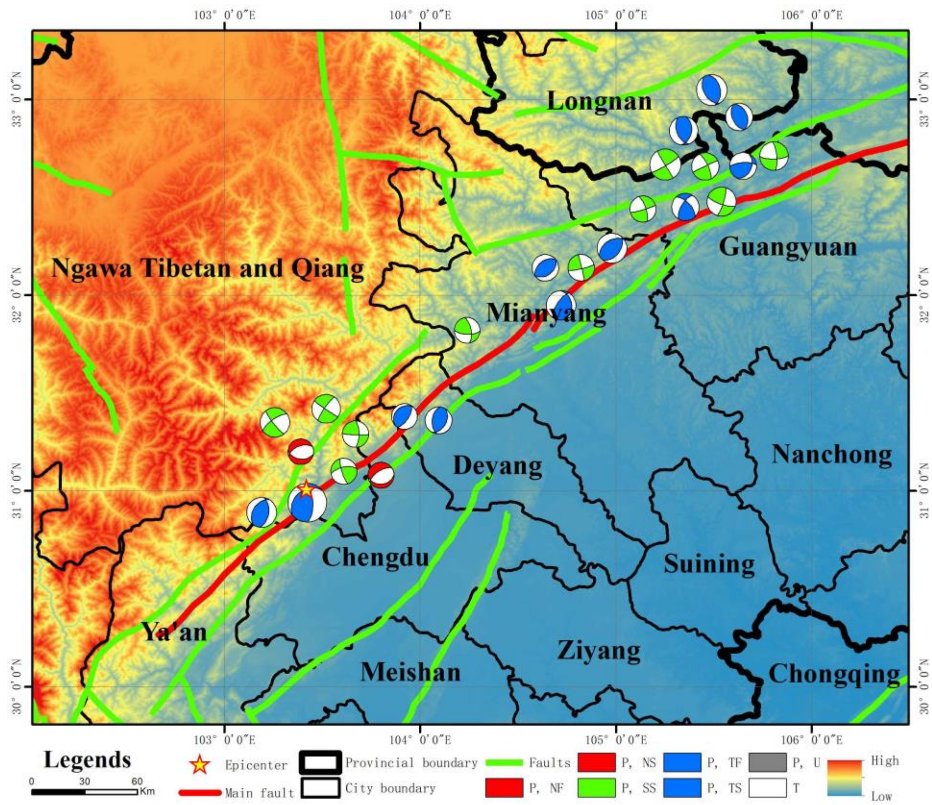

Figure 5. Structural background map of the Wenchuan earthquake. The red and green lines represent the fault structures in this area. The red line is the main fault zone of the Wenchuan earthquake, and the green lines represent other fault zones. The focal mechanism of the main aftershocks are also shown.

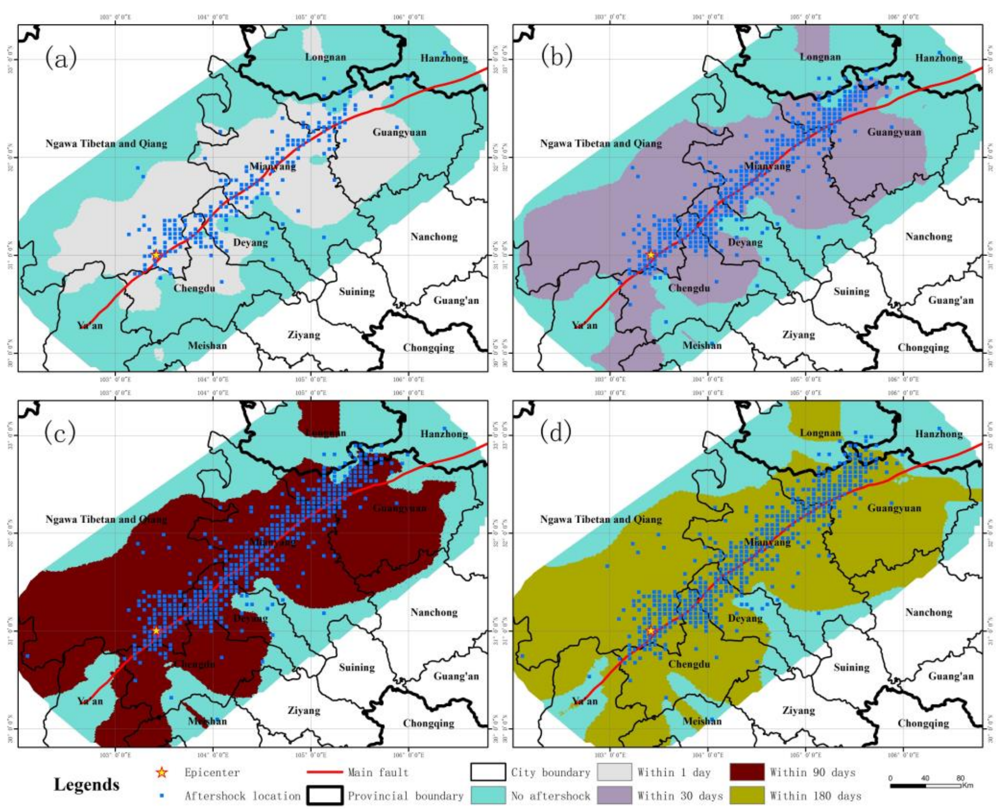

Figure 6. Aftershock damage patterns of the Wenchuan earthquake at multiple time scales. Figures (a) through (d) show the aftershock damage pattern within 1 d, 30 d, 90 d and 180 d, respectively. The blue dots indicate the actual location of aftershocks at the corresponding time scale.

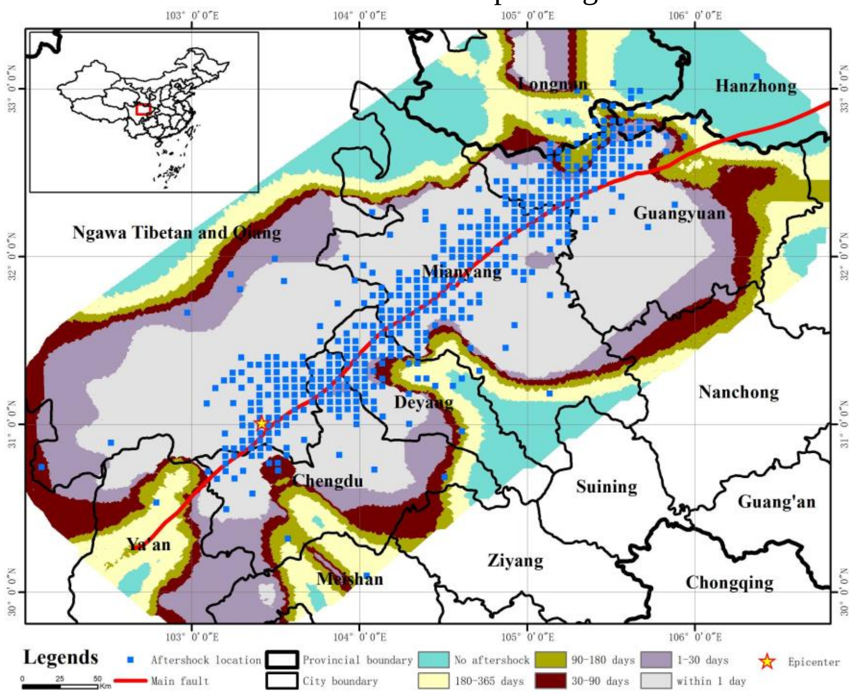

Figure 7. Aftershock hysteresis effect of the Wenchuan earthquake. The aftershock hysteresis effect can be observed by combining the aftershock patterns of the Wenchuan earthquake at different time scales. The blue dots indicate the locations of the actual aftershocks over one year.

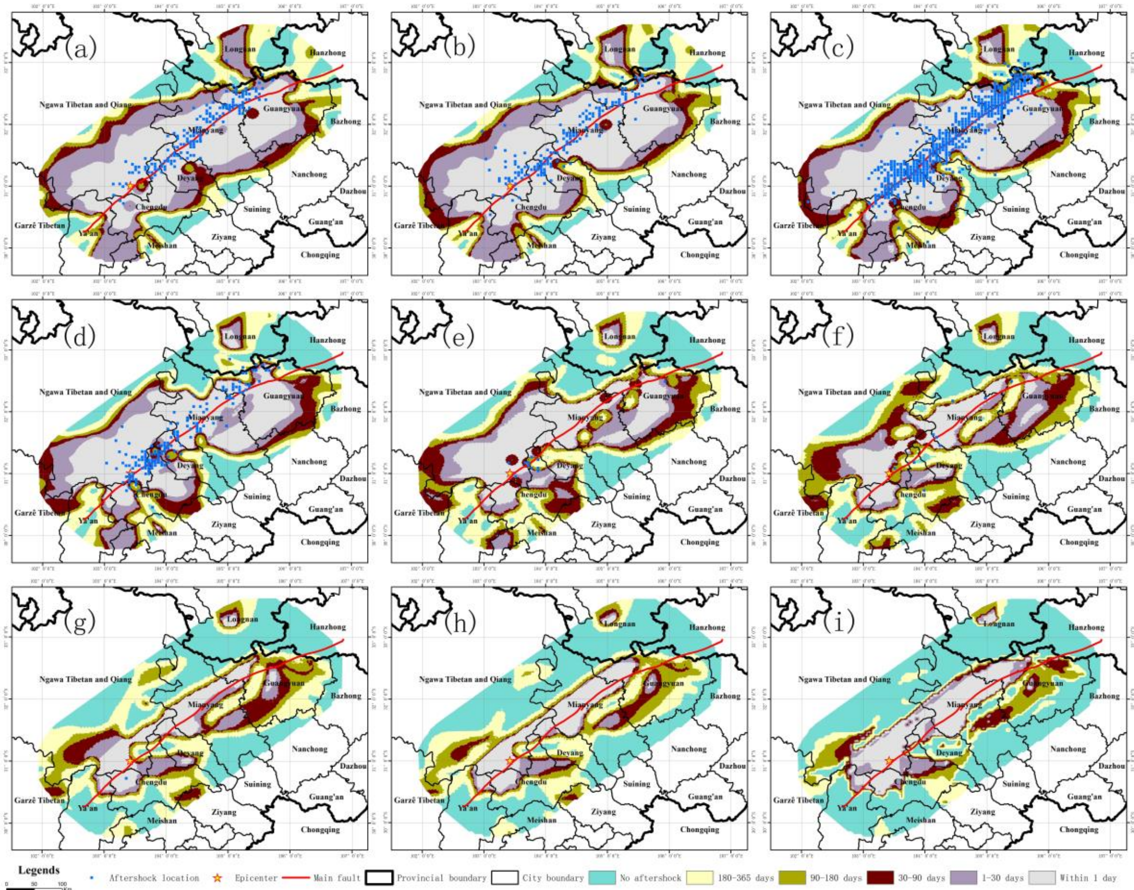

Figure 8. Aftershock hysteresis effect of the Wenchuan earthquake in different depths. Figures (a) through (i) show the aftershock hysteresis effect of the Wenchuan earthquake for depth sections of 2.5 km, 7.5 km, 12.5 km, 17.5 km, 22.5 km, 27.5 km, 32.5 km, 37.5 km, and 42.5 km, respectively.

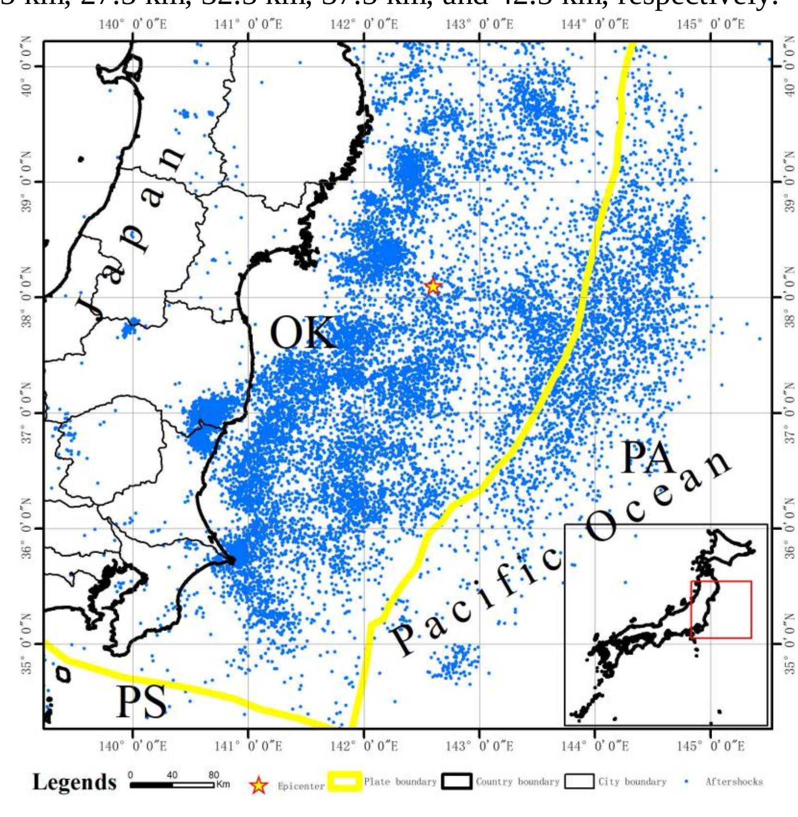

Figure 9. Aftershocks distribution of the Tohoku earthquake. The blue points in the figure are the projection positions of the aftershocks within a depth of 50 km.

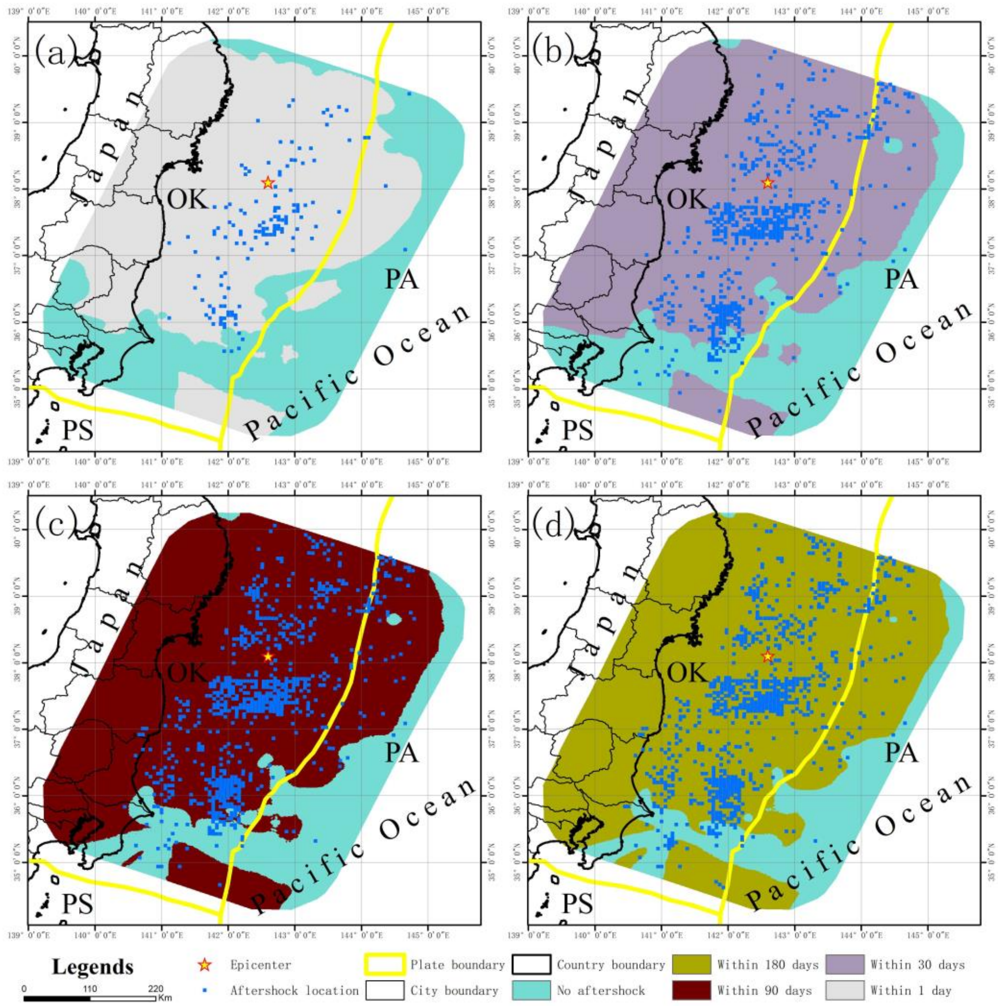

Figure 10. Aftershock damage patterns of the Tohoku earthquake at multiple time scales. Figures (a) through (d) show the aftershock damage pattern of the Tohoku earthquake within 1 d, 30 d, 90 d and 180 d, respectively. The blue dots indicate the actual location of aftershocks at the corresponding time scale.

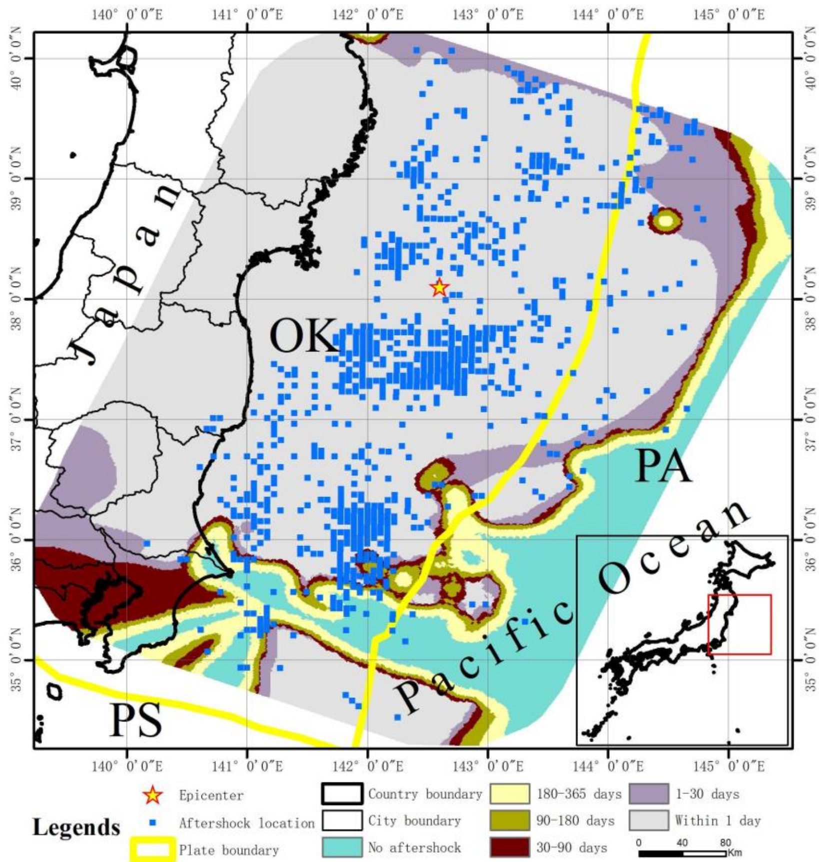

Figure 11. Aftershock hysteresis effect of the Tohoku earthquake in Japan. The aftershock hysteresis effect can be observed by combining the aftershock patterns of the Tohoku earthquake at different time scales. The blue dots indicate the locations of the actual aftershocks over one year.

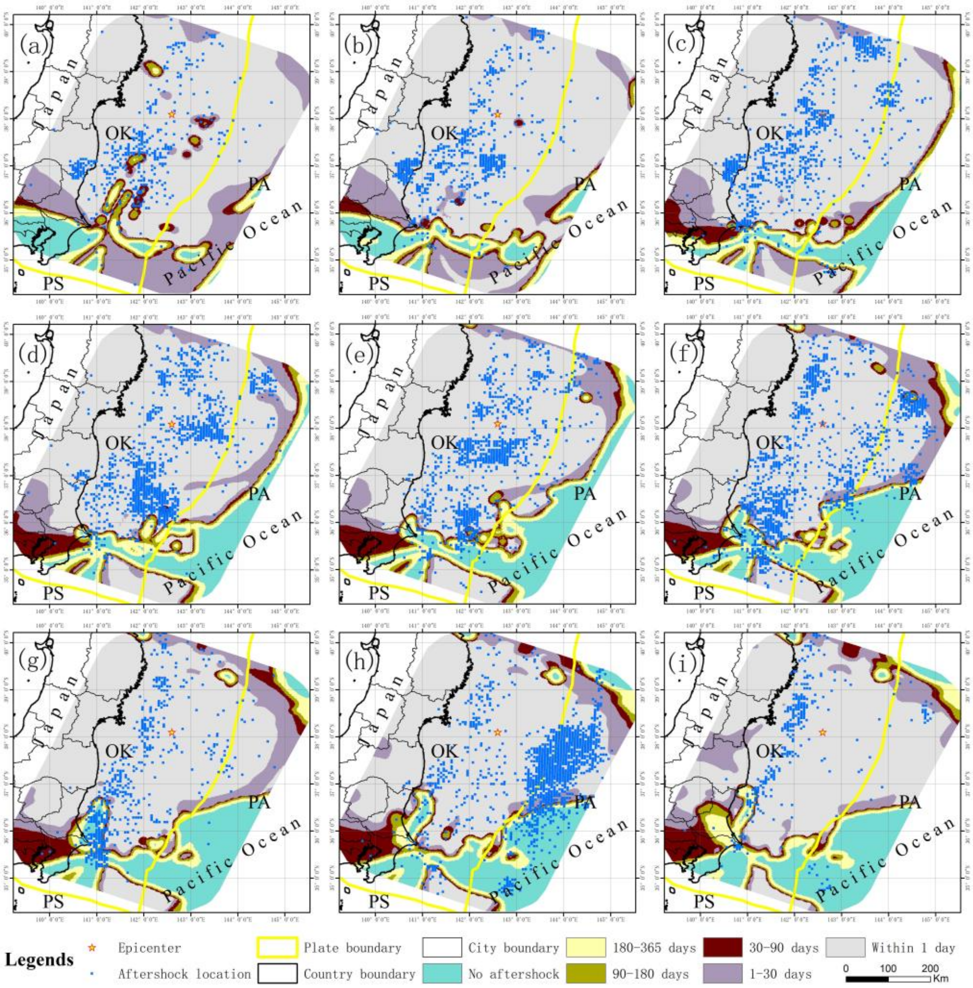

Figure 12. Aftershock hysteresis effect of the Tohoku earthquake at different depths. Figures (a) through (i) show the aftershock hysteresis effect of the Tohoku earthquake for depth sections of 2.5 km, 7.5 km, 12.5 km, 17.5 km, 22.5 km, 27.5 km, 32.5 km, 37.5 km, and 42.5 km, respectively.

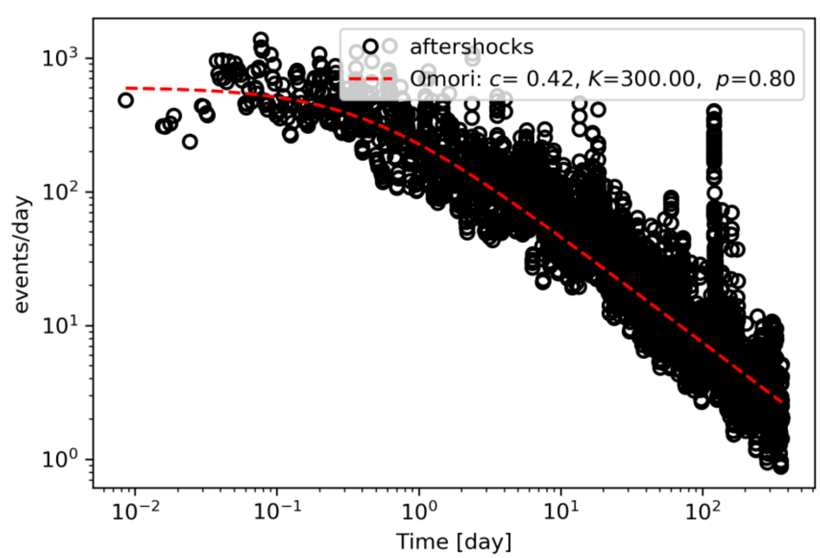

Figure 13. The modified Omori formula aftershock attenuation curve of Tohoku earthquake.

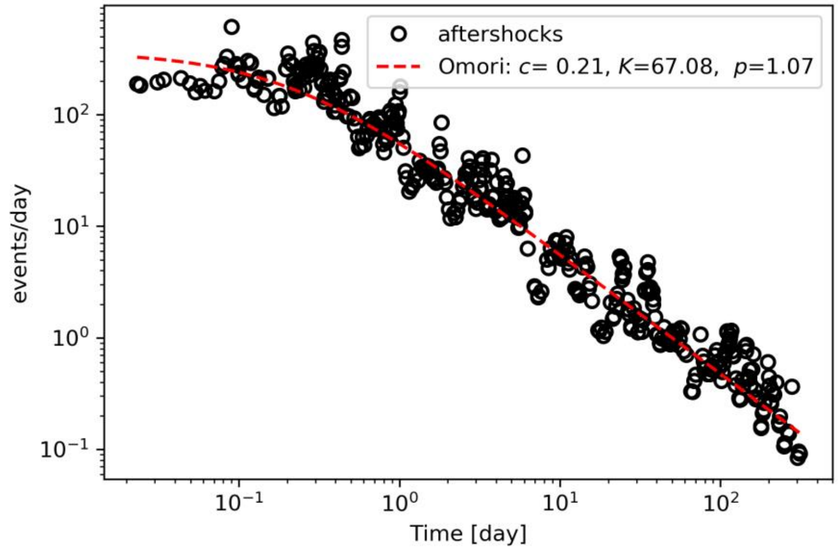

Figure 14. The modified Omori formula aftershock attenuation curve of Wenchuan earthquake.

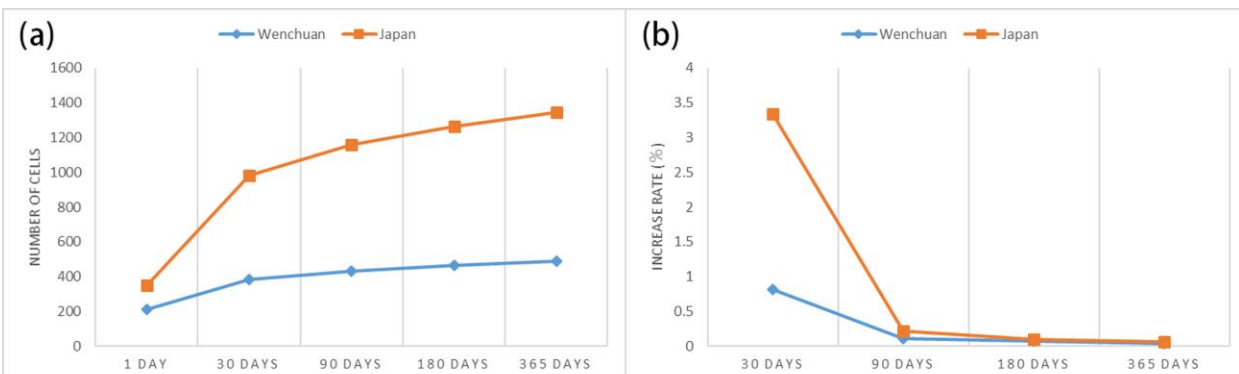

Figure 15. The curve of aftershock hysteresis effect (actual aftershocks). Figure (a) shows the change in the number of cells with aftershocks at different time scales, and figure (b) shows the change in the growth rate of the number of cells with aftershocks at different time scales.

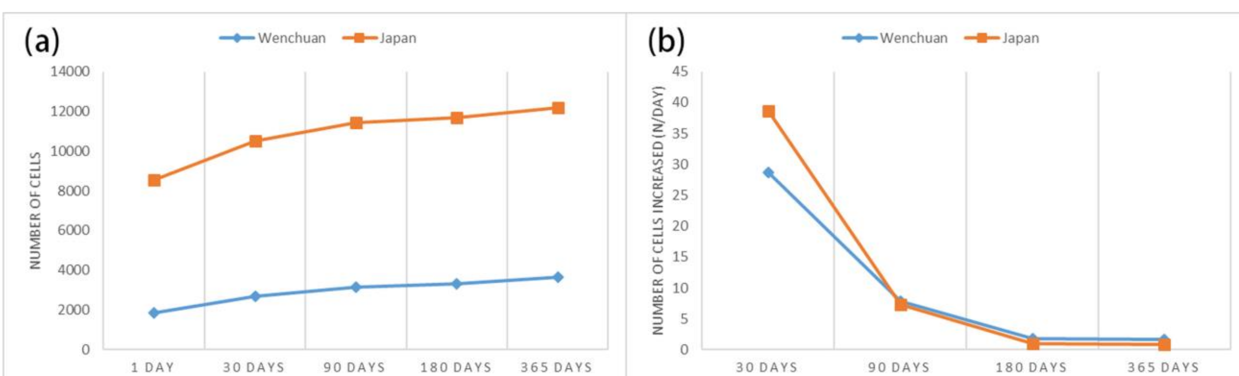

Figure 16. The curve of aftershock hysteresis effect (predicted aftershock pattern). Figure (a) shows the number of cells with aftershocks predicted at different time scales, and figure (b) shows the increment of cells with aftershocks at each time scale.

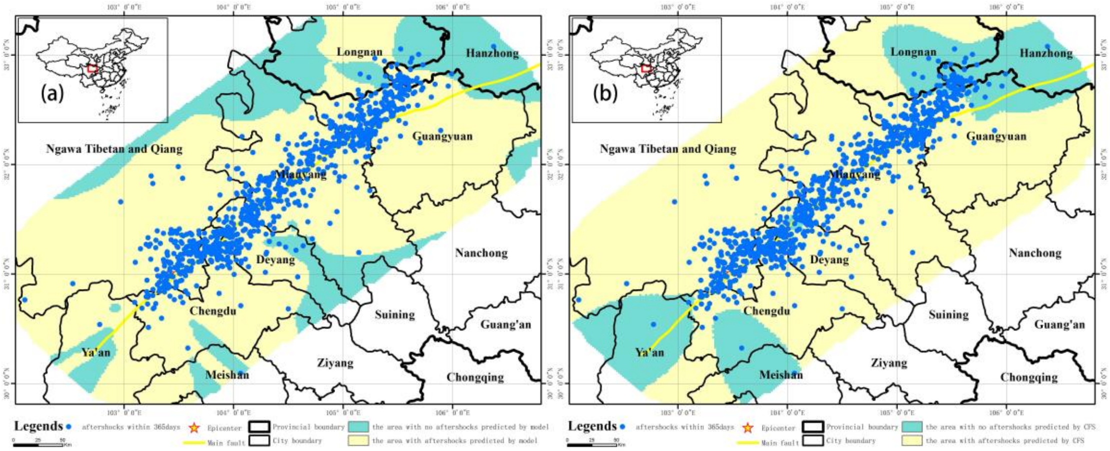

Figure 17. The hysteresis model prediction result and the △CFS prediction result of the Wenchuan earthquake. Figure (a) shows the hysteresis model prediction result and figure (b) shows the △CFS prediction result.

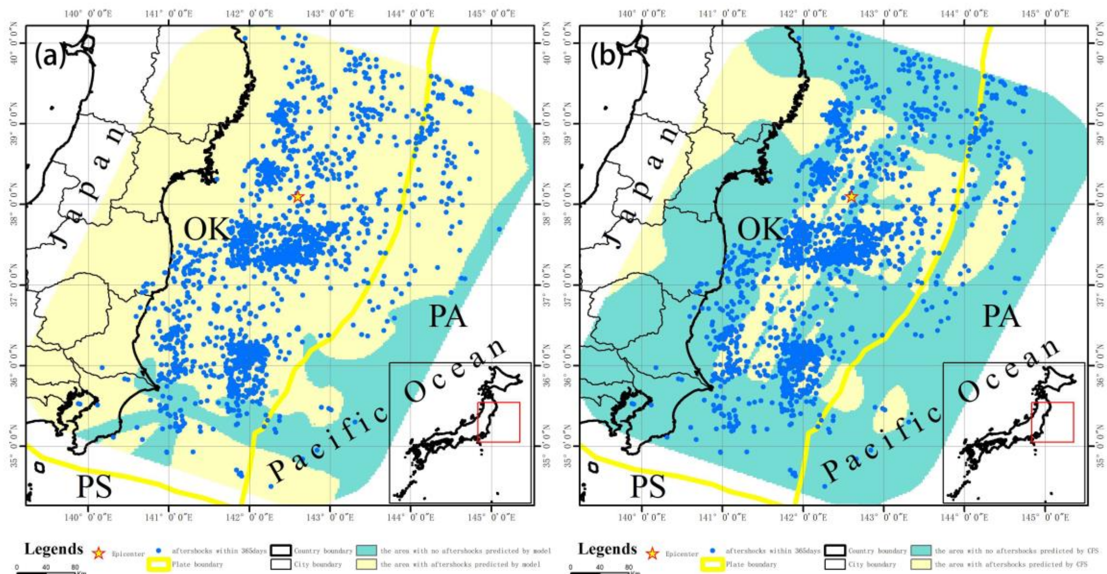

Figure 18. The hysteresis model prediction result and the △CFS prediction result of the Tohoku earthquake. Figure (a) shows the hysteresis model prediction result and figure (b) shows the △CFS prediction result.