2 # Deep Learning of Aftershock Hysteresis Effect Based on Elastic Dislocation Theory

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

5 [1] Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Faculty of
6 Geographical Science, Beijing Normal University, Beijing 100875, China

7 [2] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal
8 University, Beijing 100875, P.R. China

9 [3] Institute of Lanzhou Earthquake Research, China Earthquake Administration, Gansu Lanzhou 730000,
10 China

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

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

26

27 **Key Words:** Okada dislocation theory; FCNN; aftershock hysteresis effect; aftershock pattern; the Omori formula.

28

29

30

31

32

33

34

35

36

37

## 1 Introduction

After the occurrence of strong earthquakes, there are often a large number of aftershocks, which constitute the aftershock sequence. The aftershocks can lead to new damage to the area damaged 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 et al., 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 (Ma et al., 2005; Toda, 2003; Harris, 1998). 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 author to determine the possibility of a spatial distribution of aftershocks (Phoebe 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 change but also to temporal change, 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, the strain energy of the crustal medium will accumulate 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 in the crust, the crust will lose its stability and form the main earthquake fracture surface, which will produce the main earthquake. Simultaneously, the elastic strain energy stored in the crust will be released suddenly. After the occurrence of the main earthquake, the source body and its surrounding medium will return to 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 one time, and it will continue to accumulate 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 temporal-spatial 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). The above temporal-spatial 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.

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

## 2 Data and Methods

101  2.1 Data

102  2.1.1 Raw data

103    Two types of data are used in this paper, SRCMOD finite fault data and the ISC (International
104  Seismological Centre) seismic catalogue.
105    The inversion of finite fault source data facilitates a better understanding of the complexity of the
106  earthquake rupture process. Although the spatial resolution of the model is low, it can provide information
107  on deep seismic slip and fault evolution over time. Therefore, the finite fault model is an important means to
108  further study the mechanics and kinematics of the process of earthquake fracture. The online SRCMOD
109  database provides the inversion results for many typical earthquakes from 1906 to present. These results are
110  uploaded by seismologists globally after the main earthquake through inversion. Because the earth's crust is
111  used as an elastic medium in the calculation of coseismic displacement stress, we don't consider the impact
112  of the background of each earthquake. There are 19 finite fault source models used in this analysis: 15 are
113  used as training data and 4 are used as validation data.
114    The aftershocks following each main earthquake are obtained from the international earthquake center
115  (ISC). For the main earthquake examples in this paper, the aftershocks within 1 d, 30 d, 90 d, 180 d, and 365
116  d and within a volume extending 100km horizontally from each mainshock rupture plane and 50km vertically
117  are used for analysis of the aftershock sequences.

118  2.1.2 Data processing

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

126    2.2 Methods

127    2.2.1 Okada elastic dislocation theory

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

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

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

149    and the lame constant in the linear solid medium is used to obtain the surrounding stress change tensor **σ** as

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

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

153    2.2.2 FCNN

154    To analyze the hysteresis effect of aftershocks, it is necessary to establish a model that can predict the
155    damage modes of aftershocks at multiple time scales. We constructed a fully connected deep neural network
156    to simulate the relationship between the change value of the elastic stress tensor and aftershock and to explain
157    the hysteresis effect of aftershocks. The network established here is a network with six hidden layers. Except
158    for the second hidden layer, which has 100 neurons, the other five hidden layers have 50 neurons. The input
159    layer dimension of the entire network is 12, and its input eigenvalue is the combination of the absolute value
160    of six independent components of the elastic stress at the center of each subunit and the negative number of
161    the absolute value, for a total of 12 inputs.

162    Then we analyze the correlation between aftershocks and stress change, which is closely related to the
163    inputs of FCNN. At present, the research on aftershocks is primarily based on statistical methods, and the
164    research content primarily focuses on the distribution of aftershock strength and time attenuation. The
165    intensity distribution of aftershocks follows the G-R relationship lgN = a − bM, where M is the magnitude,
166    N is the number of aftershocks with magnitude greater than or equal to M, and a and b are the scale
167    coefficients. The value of b generally varies from 0.6 to 1.1 (Utsu, 2002), and its value is related to the

168  regional stress state (Scholz, 1968; Mogi, 1962). A low b value is related to a high stress background, i.e.,
169  the b value is relatively low during strong aftershock activity (Yi et al., 2011).

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

181      The stress change caused by the main earthquake can be calculated by the Coulomb fracture stress
182  change, which is also the most widely used analytical method at present. The change in Coulomb stress
183  produced by the main earthquake will trigger the stress of the following aftershocks (Harris, 1998). Some
184  seismologists believe that if the change in Coulomb fracture stress is positive around the main earthquake, it
185  will promote fault movement and trigger aftershocks; if the change in Coulomb fracture stress is negative, it
186  will inhibit fault movement, and the probability of triggering an aftershock is reduced (Lin, 2004; Harris,
187  1998; Han, 2003). According to the research of DeVries et al., the Coulomb fracture stress change is an
188  inadequate explanation for aftershocks, and the relationship between the positive and negative values of stress
189  change and the triggering of aftershocks requires further exploration. DeVries et al. modeled the relationship
190  between stress change and aftershock triggering by training a neural network (DeVries et al., 2018). The
191  variation in Coulomb fracture stress depends on the geometric properties and coseismic dislocations of the
192  source fault (King et al., 1994; Zhu et al., 2009). Therefore, the change value of the stress tensor, which is
193  closely related to the dislocation of the same earthquake, can be used as the aftershock variable to build the
194  model.

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

210  2.2.3 Model evaluation metric

211      The ROC curve and the AUC are used to evaluate the model. The AUC (area under curve) is defined as
212  the area enclosed by the coordinate axis under the ROC curve, and the value of the area cannot be greater
213  than 1. Because the ROC curve is generally located above the straight line y = x, the AUC value range is
214  between 0.5 and 1. Based on the AUC value, we can interpret the accuracy of the classifier. When AUC = 1,
215  the classifier is essentially a perfect classifier, whereas when AUC = 0.5, the classifier is making a random

216  assessment and the obtained model is nonsensical. For the training samples in this paper, there is a class
217  imbalance between positive and negative samples. A characteristic the ROC curve is that when the
218  distribution of positive and negative samples in the test set changes, the ROC curve can remain unchanged.
219  The closer the ROC curve is to the Y-axis and y = 1, i.e., the higher the AUC value of the classifier, the
220  greater the classification accuracy. Generally, when the AUC is less than 0.6, the accuracy of the classifier is
221  poor; when the AUC is less than 0.75, the accuracy of the classifier is moderate; and when the AUC is greater
222  than 0.75, the accuracy of the classifier is good. The output of the model in this paper is a probability value
223  between 0 and 1. When the ROC curve is used to evaluate the model, it is conducted at five time scales, and
224  the model under each time scale is evaluated as a two-classification problem.

225  **3 Results**

226  3.1 Evaluation of the aftershock hysteresis model

227      The aftershock hysteresis model under multiple time scales is obtained by using the neural network to
228  train the constructed training dataset. In this paper, five submodels are trained, and the final hysteresis model
229  is composed of five submodels. The output value of the neural network in each cell is binarized with a
230  threshold value of 0.5. A cell with a predicted value greater than 0.5 is assigned as 1, and a cell with a
231  predicted value less than 0.5 is assigned as 0. Based on the trained aftershock hysteresis model, the aftershock
232  patterns are predicted for the Wenchuan earthquake at multiple time scales, and the ROC curves are obtained
233  for the different time scales. The AUC values of the five time scales are all above 0.8 in both the training and
234  validation sets, and some are close to 0.9. The AUC values of the training set are all higher than those of the
235  validation set for the different time scales. The neural network designed by Phoebe et al. is used for aftershock
236  prediction. The AUC value of the training model on the validation set is 0.849 (Figure 2). In this paper, the
237  AUC value of each submodel on the validation set is similar to the research results of Phoebe et al. Therefore,
238  the model achieves good prediction results at different time scales.

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

$$\Delta CFS = \Delta\tau + \mu\Delta\sigma$$

243  where $\mu$ is apparent friction coefficient, $\Delta\sigma$ is the normal stress on the fault plane and $\Delta\tau$ is shear stress in the
244  direction of fault slip. Based on previous studies, the friction coefficient in this paper is 0.4 (King er al., 1994;
245  Wan et al., 2004). Numerous studies have shown that aftershocks will occur when ΔCFS is greater than
246  +0.01Mpa. In order to compare and analyze the output of FCNN, we need to transform the ΔCFS to 0~1.
247  Similar to the last layer of FCNN, the variation function adopts a variant of the sigmoid function as follows

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

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

### 3.2 Application of the model to the Wenchuan earthquake

#### 3.2.1 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 4, the strong aftershocks following the Wenchuan earthquake are primarily reverse faults 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 12,500 m underground (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 5). 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 6). 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, which is consistent with the actual location of the aftershocks. Over time, the area of aftershocks expands outwards, and the rate slows 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.

#### 3.2.2 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 2,500 m, 7,500 m, 12,500 m, 17,500 m, 22,500 m, 27,500 m, 32,500 m, 37,500 m and 42,500 m. 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 7). 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 12500 m, 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

306 changed. Although the overall expansion direction is consistent with the main fault, the pattern is less regular
307 and more random.

3.3 Application of the model to the Tohoku earthquake in 2011

### 3.3.1 Aftershock hysteresis failure mode

310 Japan is located in the Circum-Pacific seismic belt at the intersection of the Eurasian plate and the Pacific
311 plate, which is an area with a frequent occurrence of global earthquakes. Due to the collision between the
312 Pacific plate and the Eurasian plate, the Pacific plate is subducted under the Eurasian plate, thus forming the
313 Japan Trench and the Japanese island arc. "OK" represents the Okhotsk plate, which is part of the Eurasian
314 plate, "PA" refers to the Pacific plate, and "PS" refers to the Philippine Sea plate, which is also part of the
315 Eurasian plate (Bird, 2003) (Figure 8). The epicenter of the earthquake was located in the subduction zone
316 of the Japanese trench. Based on the aftershock hysteresis model, the aftershock patterns within 1 d, 30 d, 90
317 d, 180 d and 365 d after the main earthquake are predicted, and the section (22500 m) at the focal depth of
318 the main earthquake is selected for analysis. Using the Tohoku earthquake in Japan as an example, the
319 greatest expansion of the aftershock distribution area is observed within 30 d. The shape of the aftershock
320 patterns are similar at all time scales. The aftershock and the predicted aftershock patterns are distributed in
321 an approximately North-South direction along the Japan Trench and plate boundary.

322 The aftershock hysteresis model of the Tohoku earthquake in is obtained by synthesizing the aftershock
323 patterns at different time scales (Figure 9). Over time, the expansion rate of the aftershock pattern gradually
324 slows, and the expansion direction is basically perpendicular to the aftershock pattern at the previous scale.
325 Most of the aftershocks of this earthquake occurred in the eastern Sea of Japan, and the concentrated area of
326 terrestrial aftershocks was located in Fukushima.

### 3.3.2 Aftershock hysteresis patterns at different depths

328 Similar to the Wenchuan earthquake, the aftershock hysteresis pattern of the Tohoku earthquake changes
329 with the change in depth. The magnitude of the earthquake was very large, reaching over $M_w9$. The main
330 earthquake has a great impact on the surrounding area and the crust, which stores considerable energy, then
331 releases it in the form of aftershocks. The predicted expansion direction of the aftershock model is generally
332 consistent with that of the plate boundary and the Japan Trench. In this study, the maximum analysis depth
333 is 50 km. Using the depth section of the main shock source as the center, the actual aftershock pattern does
334 not change significantly when the depth change is small. This may be due to the large magnitude of the
335 earthquake. The area of the actual aftershock pattern is reduced at a depth of 42,500 m. However, the location
336 of the aftershock is still widely distributed. The expansion of the aftershock pattern also changes beginning
337 at a depth of 27500 m. The general direction of distribution is along the trench, and some areas begin to
338 expand vertically along the trench (Figure 10).

## 4 Discussion

### 4.1 Aftershock hysteresis effect

341 The modified Omori formula is $n(t) = K(t + c)^{-p}$, where n(t) is the aftershock frequency per unit time,
342 and as t increases, n(t) will decrease correspondingly to describe the time attenuation characteristics of
343 aftershocks. Compared with the Omori formula, the aftershock hysteresis effect analyzed in this paper can
344 be reflected by the correlation between the change of time scale and the region of aftershocks. Based on the
345 discussion of focal depth sections of the main earthquake, within one day after the Wenchuan earthquake,
346 the number of subunits with aftershocks is 213; within 30 d, it is 386, representing an increase of 81.6%;
347 within 90 d, it is 432, representing an increase of 11.9%; within 180 d, it is 466, representing an increase of
348 7.9%; and within 365 d, it is 488, representing an increase of 4.7%. Within one day after the Tohoku
349 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 slows. The rate of expansion is most rapid 30d after the earthquake. After 30d, 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 11 and Figure 12, 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.

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 learned. Compared with the static Coulomb failure stress change method, the aftershock hysteresis model obtained has a better prediction effect.

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 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 SRCMOD finite fault data and the ISC seismic

catalogue and is used to construct an aftershock hysteresis model. Using the main aftershock sequences of the Wenchuan earthquake and the Tohoku earthquake in Japan 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 Phoebe et al. In this study, the figures and the subsequent processing of the results are all performed using 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.

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.

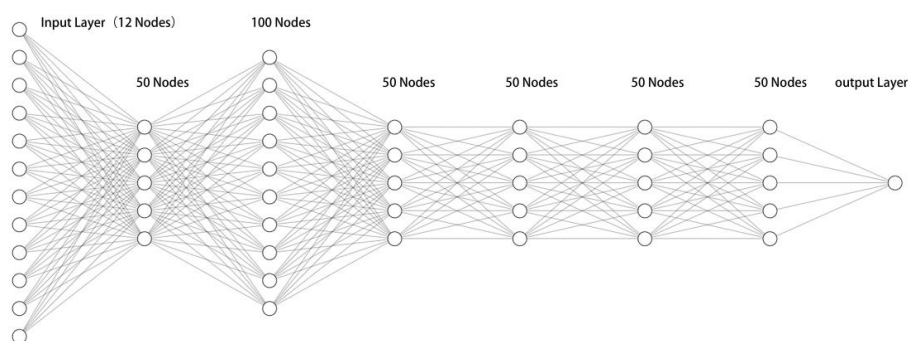
Dieterich, J.: A constitutive law for rate of earthquake production and its application to earthquake clustering, Journal of Geophysical Research, 99(B2), 2601, 1994.

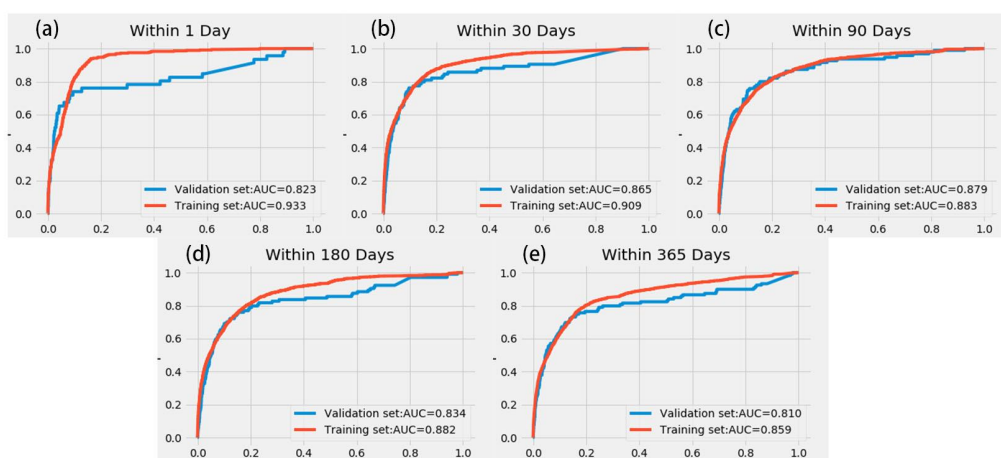
440    Gu, J., Xie, X., and Zhao, L.: On temporal distribution of large aftershocks of the sequence of a major
441        earthquake and preliminary theoretical explanation, Acta Geophysica Sinica, 22(1), 32-46, 1979.
442    Han, Z.: Possible reduction of earthquake hazard on the Wellington Fault, New Zealand, after the nearby
443        1855, M8.2 Wairarapa earthquake and implication for interpreting paleoearthquake intervals, Annals
444        of Geophysics, 46(5), 1141-1154, https://doi.org/10.4401/ag-3450, 2003.
445    Harris, R. A.: Introduction to special section: stress triggers, stress shadows, and implications for seismic
446        hazard. Journal of Geophysical Research, 103(B10), 24347, https://doi.org/10.1029/98JB01576, 1998.
447    Jordan, MI., and Mitchell, T.M.: Machine learning: Trends, perspectives, and prospects, Science,
448        349(6245), 255-260, https://doi.org/10.1126/science.aaa8415, 2015.
449    Kagan, Y.Y., and Heidi, H.: Relation between mainshock rupture process and Omori's law for aftershock
450        moment release rate, Geophysical Journal International, 163(3), 1039-1048,
451        https://doi.org/10.1111/j.1365-246X.2005.02772.x, 2005.
452    Kagan, Y.Y., and Jackson, D.D.: Long-term probabilistic forecasting of earthquakes, Journal of
453        Geophysical Research, 99(B7), 13685-13700, 1994.
454    King, G.C.P., Stein, R.S., and Lin, J.: Static stress changes and the triggering of earthquakes, Bulletin of the
455        Seismological Society of America, 84(3), https://doi.org/10.1029/94JB00611, 1994.
456    Kong, Q., Trugman, D. T., and Ross, Z. E.: Machine Learning in Seismology: Turning Data into Insights,
457        Seismological Research Letters, 90(1), 3-14, https://doi.org/10.1785/0220180259, 2019.
458    Lecun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521(7553), 436,
459        https://doi.org/10.1038/nature14539, 2015.
460    Lin, J.: Stress triggering in thrust and subduction earthquakes and stress interaction between the southern
461        San Andreas and nearby thrust and strike-slip faults, Journal of Geophysical Research,
462        109(B2),B02303, https://doi.org/10.1029/2003jb002607, 2004.
463    Lindman, M., Jonsdottir, K., and Roberts, R.: Earthquakes descaled: on waiting time distributions and
464        scaling laws, Physical Review Letters, 94(10), 108501,
465        https://doi.org/10.1103/PhysRevLett.94.108501, 2005.
466    Ma, K. F., Chan, C.H., and Stein, R.S.: Response of seismicity to Coulomb stress triggers and shadows of
467        the 1999 Mw=7.6 Chi-Chi, Taiwan, earthquake, Journal of Geophysical Research Solid Earth,
468        https://doi.org/10.1029/2004JB003389, 2005.
469    Meade, B. J., Devries, P. M. R., and Faller, J.: What Is Better Than Coulomb Failure Stress? A Ranking of
470        Scalar Static Stress Triggering Mechanisms from 105 Mainshock-Aftershock Pairs, Geophysical
471        Research Letters, 44, https://doi.org/10.1002/2017GL075875, 2017.
472    Mignan, A., and Broccardo, M.: A Deeper Look into 'Deep Learning of Aftershock Patterns Following
473        Large Earthquakes': Illustrating First Principles in Neural Network Physical Interpretability, IWANN
474        2019: Advances in Computational Intelligence, 3-14, https://doi.org/10.1007/978-3-030-20521-8_1,
475        2019.
476    Mogi, K.: On the time distribution of aftershocks accompanying the recent major earthquake in and near
477        Japan, Bull Earthquake Res Inst, 40:107-124, 1962.
478    Moustra, M., Avraamides, M., and Christodoulou, C.: Artificial neural networks for earthquake prediction
479        using time series magnitude data or seismic electric signals, Expert Systems with Applications,
480        38(12), 15032-15039, https://doi.org/10.1016/j.eswa.2011.05.043, 2011.
481    Nanjo, K. Z., Enescu, B., and Shcherbakov, R.: Decay of aftershock activity for Japanese earthquakes,
482        Journal of Geophysical Research Solid Earth, 112(B8), https://doi.org/10.1029/2006JB004754, 2007.
483    Narteau, C., Byrdina, S., and Shebalin, P.: Common dependence on stress for the two fundamental laws of
484        statistical seismology, Nature, 462(7273), 642-645, https://doi.org/10.1038/nature08553, 2009.
485    Narteau, C., Shebalin, P., and Holschneider, M.: Onset of power law aftershock decay rates in southern
486        California, Geophysical Research Letters, 32(22), L22312, https://doi.org/10.1029/2005gl023951,
487        2005.
488    Ogata, Y.: Statistical models for earthquake occurrences and residual analysis for point processes, JASA,
489        83(401), 9-27, 1988.

490    Ogata, Y.: Space-time point-process models for earthquake occurrences, Annals of the Institute of
491        Statistical Mathematics, 50(2), 379-402, 1998.
492    Okada, Y.: Surface deformation due to shear and tensile faults in a half-space, Bull. Seismol. Soc. Amer.,
493        75, 1135-1154, https://doi.org/10.1016/0148-9062(86)90674-1, 1985.
494    Okada, Y.: Internal deformation due to shear and tensile faults in a half-space, Bull. Seismol. Soc. Amer.,
495        82(2), 1018-1040, 1992.
496    Omori, F.: On aftershocks of earthquakes, J. Coll. Sci. Imp. Univ. Tokyo. 7, 11-200, 1894.
497    Phoebe, M. R. D., Fernanda, V., Martin, W. and Brendan, J. M.: Deep learning of aftershock patterns
498        following large earthquakes, Nature, 560(7720), 632-634, https://doi.org/10.1038/s41586-018-0438-y,
499        2018.
500    Scholz, C.H.: Microfracturing and the inelastic deformation of rock in compression, Journal of Geophysical
501        Research, 73(4), 1417-1432, https://doi.org/10.1029/jb073i004p01417, 1968.
502    Shan, X., Qu, C., and Gong, W.: Coseismic deformation field of the Jiuzhaigou Ms7.0 earthquake from
503        Sentinel-1A InSAR data and fault slip inversion, Chinese Journal of Geophysics, 60(12), 4527-4536,
504        https://doi.org/10.6038/cjg20171201, 2017.
505    Stein, R. S., and Lisowski, M.: The 1979 homestead valley earthquake sequence, California: control of
506        aftershocks and postseismic deformation, Journal of Geophysical Research Solid Earth, 88(B8),
507        https://doi.org/10.1029/JB088iB08p06477, 1983.
508    Steketee, J. A.: Some geophysical applications of the elasticity theory of dislocations. Canadian Journal of
509        Physics, 36(9), 1168-1198, https://doi.org/10.1139/p58-123, 1958.
510    Toda, S.: Toggling of seismicity by the 1997 Kagoshima earthquake couplet: A demonstration of time-
511        dependent stress transfer, Journal of Geophysical Research, 108(B12), 2567,
512        https://doi.org/10.1029/2003jb002527, 2003.
513    Utsu, T.: A statistic study on the occurrence of aftershocks, Geophys, 30, 521-545, 1961.
514    Utsu, T.: 43 Statistical features of seismicity, International Geophysics, 81, 719-732,
515        https://doi.org/10.1016/S0074-6142(02)80246-7, 2002.
516    Wan, Y., Wu Z., and Zhou G.: Focal mechanism dependence of static stress triggering of earthquakes,
517        Tectonophysics, 390(1-4), 235-243, 2004.
518    Wang, Z., Zhang, R., and Wang, X.: InSAR coseismic deformation monitoring and fault inversion of 2016
519        Qinghai Menyuan Earthquake, Remote Sensing Information, 33(006), 103-108,
520        https://doi.org/10.3969/j.issn.1000-3177.2018.06.015, 2018.
521    Wong, K., and Schoenberg, F. P.: On mainshock focal mechanisms and the spatial distribution of
522        aftershocks, Bulletin of the Seismological Society of America, 99(6), 3402-3412, 2009.
523    Yang, Z., Chen, Y., and Zheng, Y.: The application of double differential seismic location method in the
524        accurate location of earthquakes in central and western China, Science in China(Series D), 33(z1),
525        https://doi.org/10.1360/03dz0014, 2003.
526    Yi, G., Wen, X., and Xin, H.: Distributions of seismicity parameters and seismic apparent stresses on the
527        Longmenshan-Minshan tectonic zone before the 2008 Ms8.0 Wenchuan earthquake, Acta Geophysica
528        Sinica, 54(006), 1490-1500, https://doi.org/10.3969/j.issn.0001-5733.2011.06.008, 2011.
529    Zhu, H, and Wen, X.: Stress triggering process if the 1973 to 1976 Songpan, Sichuan, sequence of strong
530        earthquake, Chinese J. Geophys, 52(4), 994-1003, https://doi.org/10.3969/j.issn.0001-
531        5733.2009.04.016, 2009.
532    Zhao, J.: Application of InSAR technology in surface deformation monitoring, M.S. thesis, Nanjing
533        University, China, 2019.
534    Zhao, Z., and Chen, N.: Accuracy amendment of focal location along the Longmenshan fault in the north of
535        Sichuan, Earthquake Research in Sichuan, 4, 19-30, 1995.
536
537
538
539

540
541
542
543
544
545
546

547 **Figure**

548

549
550
551 **Figure 1.** The structure of DNN. The neural network is composed of an input layer, hidden layers, output
552 layers and the connections between each layer. The function of each hidden layer is to transform the
553 features of the network input.

554

555
556
557 **Figure 2.** ROC curve for multiple time scales. Figures (a) through (e) show the ROC curve of the model
558 within 1 d, 30 d, 90 d, 180 d and 365 d, respectively. The horizontal axis represents the FPR (false positive
559 rate) and the vertical axis represents the TPR (true positive rate).

560

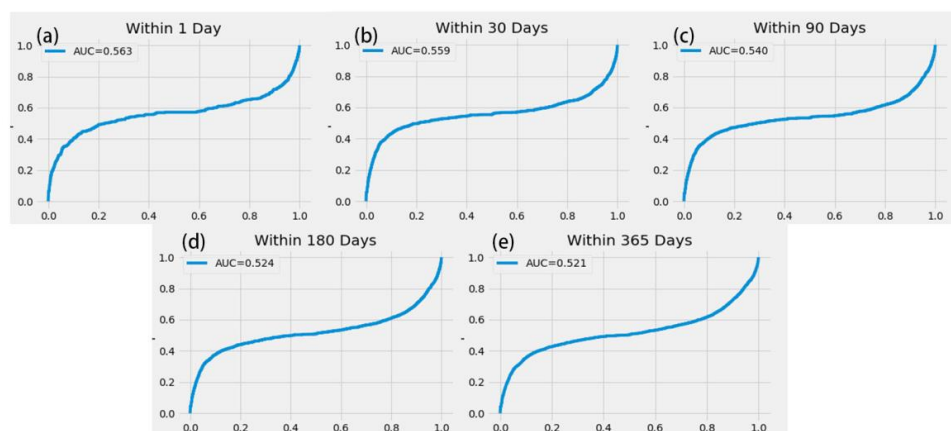



**Figure 3.** ROC curve of △CFS for multiple time scales. Figures (a) through (e) show the ROC curve of the model within 1 d, 30 d, 90 d, 180 d and 365 d, respectively. The horizontal axis represents the FPR (false positive rate) and the vertical axis represents the TPR (true positive rate).

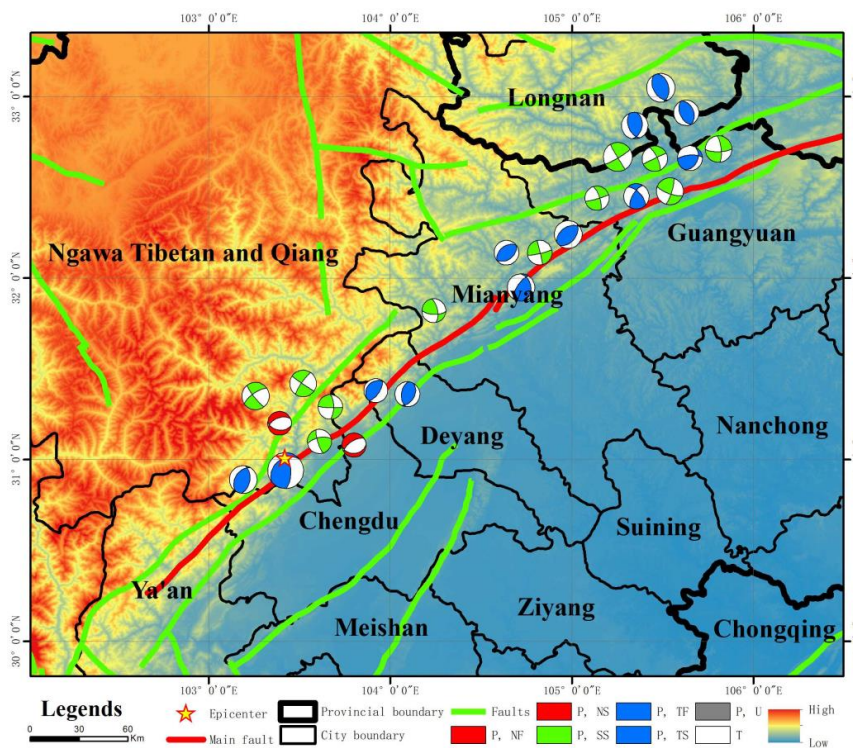



**Figure 4.** Structural background map of the Wenchuan earthquake. The red line and green line represent the fault structures in this area. The red line is the main fault zone of the Wenchuan earthquake, and the green line represents other fault zones. The ellipsoid of the focal mechanism of the main aftershocks is also shown, which represents the types of reverse faults: strike slip reverse fault, normal fault, strike slip normal fault and strike slip fault.

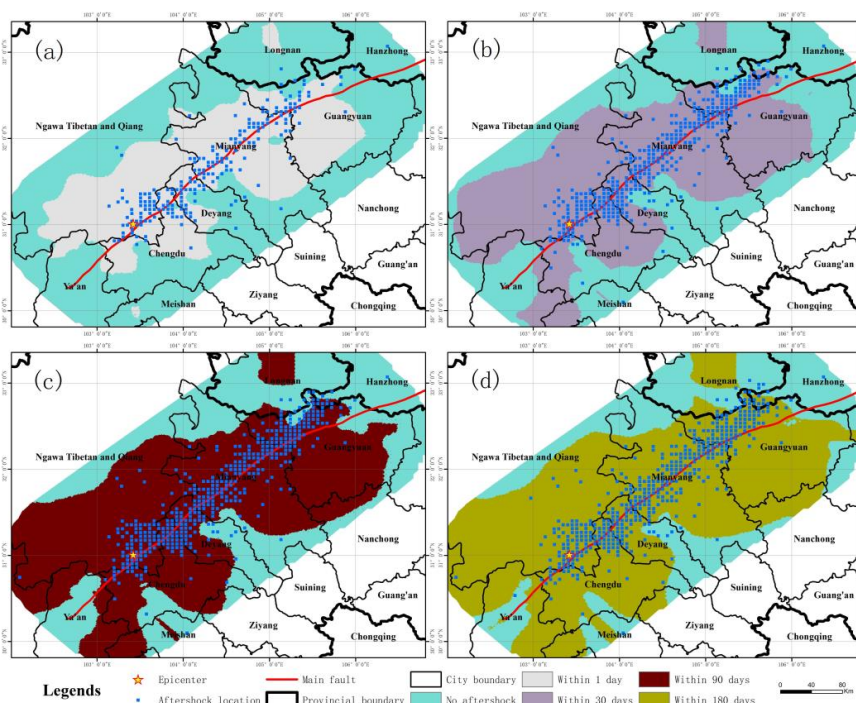
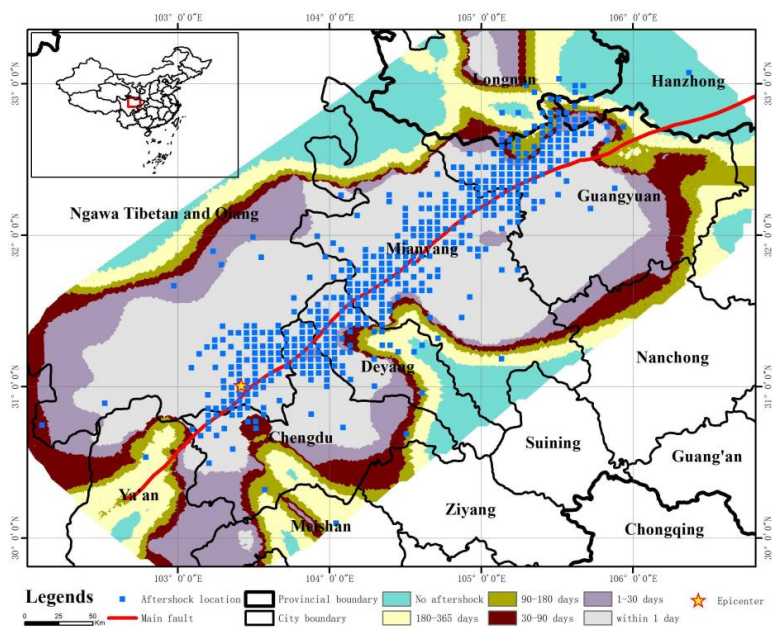
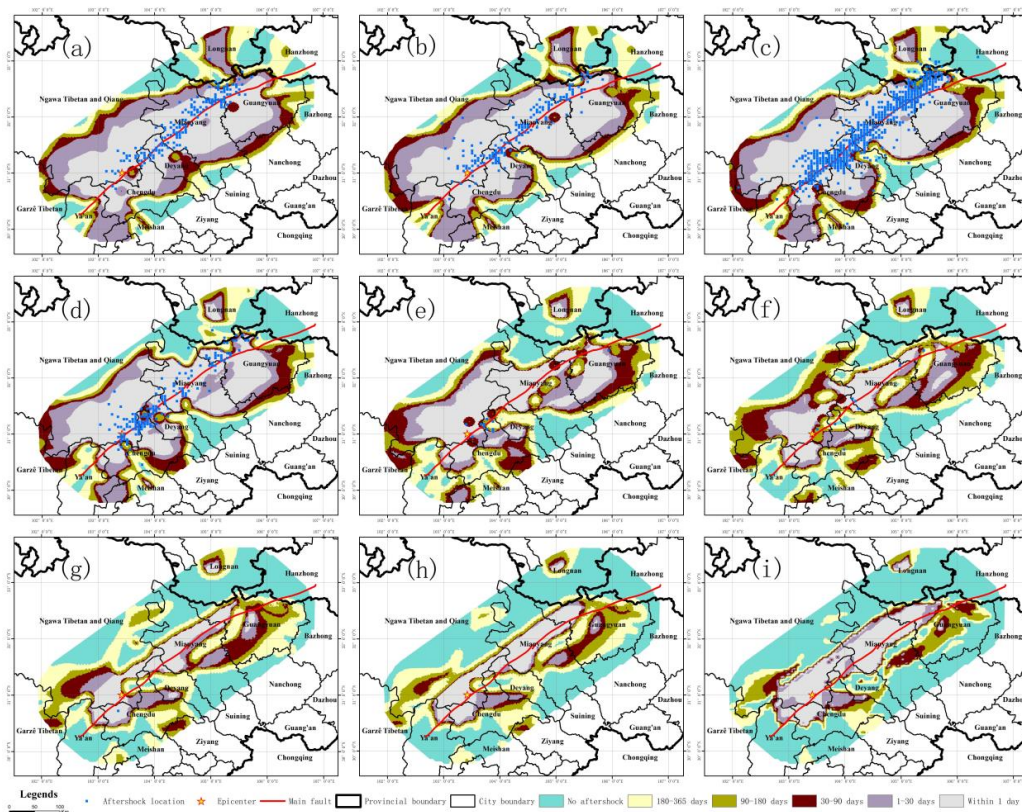



**Figure 5.** Aftershock damage patterns of the Wenchuan earthquake at multiple time scales. Figures (a) through (d) show the aftershock damage pattern within 1 d, 30 d, 90 d and 180 d, respectively. The blue dots indicate the actual location of aftershocks at the corresponding time scale.

583 **Figure 6.** Aftershock hysteresis effect of the Wenchuan earthquake. The aftershock hysteresis effect can be
584 observed by combining the aftershock patterns of the Wenchuan earthquake at different time scales. The
585 blue dots indicate the locations of the actual aftershocks over one year.

586

587
588
589 **Figure 7.** Aftershock hysteresis effect of the Wenchuan earthquake in different depths. Figures (a) through
590 (i) show the aftershock hysteresis effect of the Wenchuan earthquake for depth sections of 2,500 m, 7,500
591 m, 12,500 m, 17,500 m, 22,500 m, 27,500 m, 32,500 m, 37,500 m, and 42,500 m, respectively.

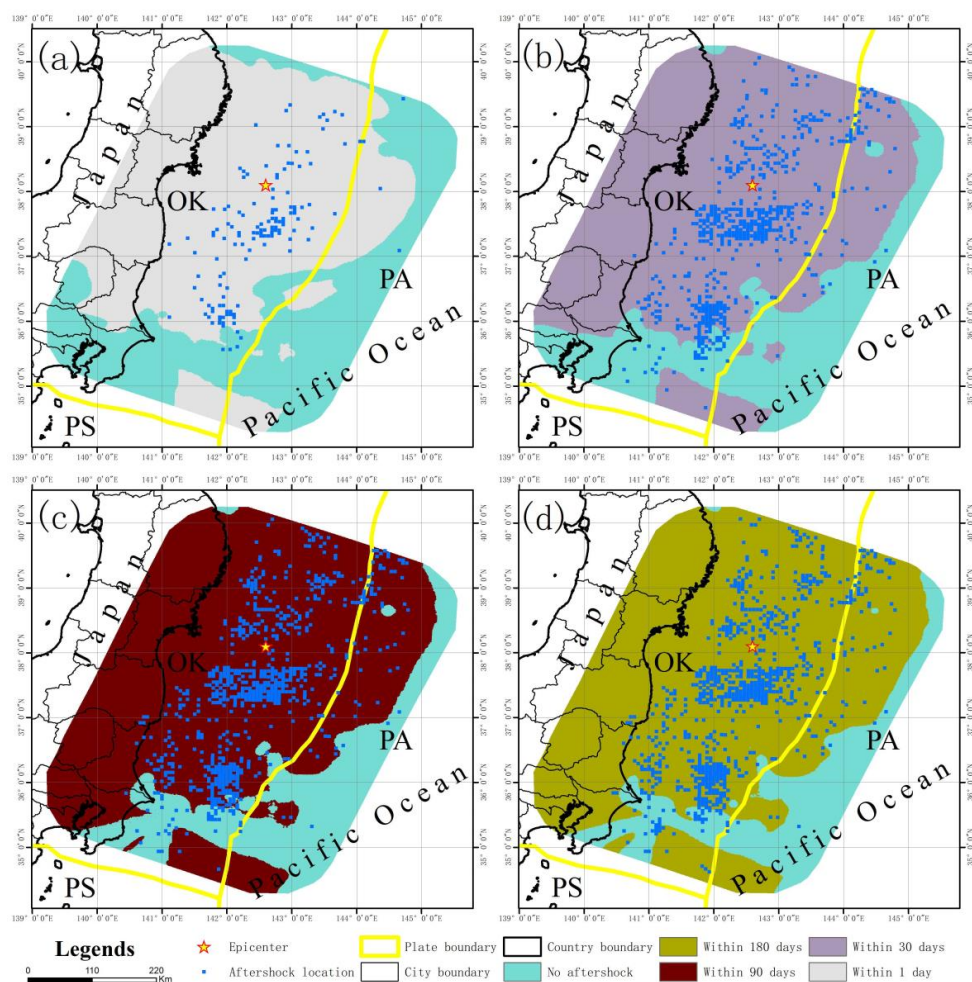



**Figure 8.** Aftershock damage patterns of the Tohoku earthquake at multiple time scales. Figures (a) through (d) show the aftershock damage pattern of the Tohoku earthquake within 1 d, 30 d, 90 d and 180 d, respectively. The blue dots indicate the actual location of aftershocks at the corresponding time scale.

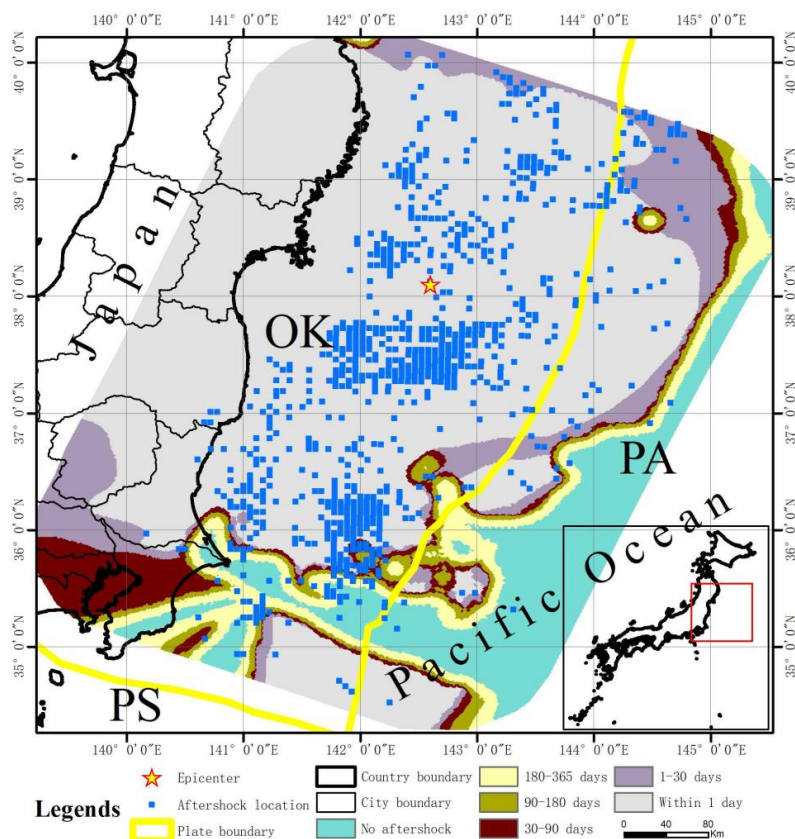



**Figure 9.** Aftershock hysteresis effect of the Tohoku earthquake in Japan. The aftershock hysteresis effect
can be observed by combining the aftershock patterns of the Tohoku earthquake at different time scales.
The blue dots indicate the locations of the actual aftershocks over one year.

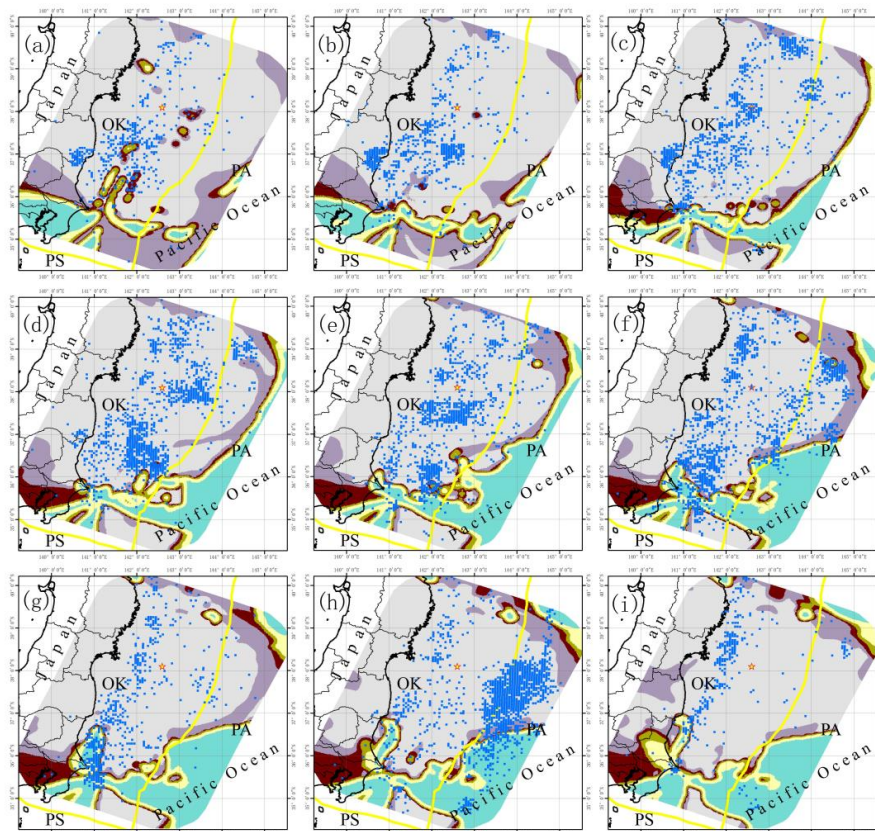



**Figure 10.** Aftershock hysteresis effect of the Tohoku earthquake at different depths. Figures (a) through (i) show the aftershock hysteresis effect of the Tohoku earthquake for depth sections of 2,500 m, 7,500 m, 12,500 m, 17,500 m, 22,500 m, 27,500 m, 32,500 m, 37,500 m, and 42,500 m, respectively.

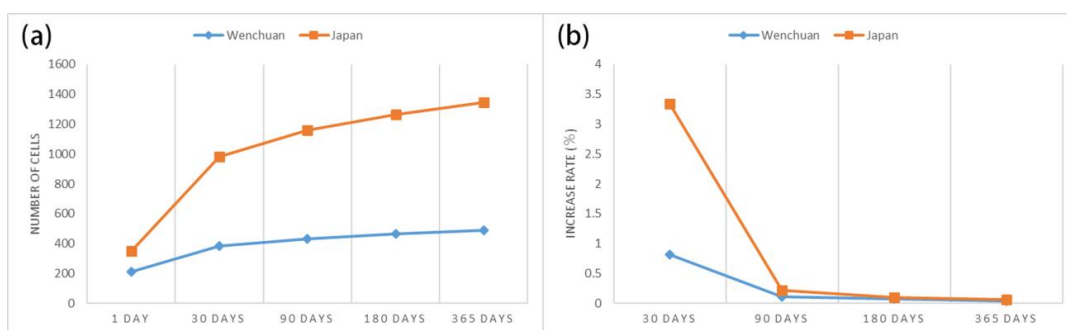



**Figure 11.** The curve of aftershock hysteresis effect (actual aftershocks). Figure (a) shows the change in the number of cells with aftershocks at different time scales, and figure (b) shows the change in the growth rate of the number of cells with aftershocks at different time scales.

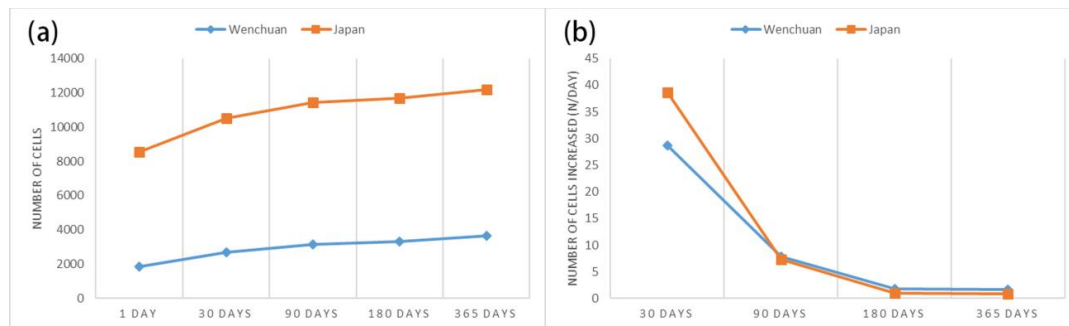



**Figure 12.** The curve of aftershock hysteresis effect (predicted aftershock pattern). Figure (a) shows the number of cells with aftershocks predicted at different time scales, and figure (b) shows the increment of cells with aftershocks at each time scale.