LETTER

https://doi.org/10.1038/s41586-018-0438-y

Deep learning of aftershock patterns following large earthquakes

Phoebe M. R. DeVries[1,2]*, Fernanda Viégas[3], Martin Wattenberg[3] & Brendan J. Meade[1]

Aftershocks are a response to changes in stress generated by large earthquakes and represent the most common observations of the triggering of earthquakes. The maximum magnitude of aftershocks and their temporal decay are well described by empirical laws (such as Bath's law[1] and Omori's law[2]), but explaining and forecasting the spatial distribution of aftershocks is more difficult. Coulomb failure stress change[3] is perhaps the most widely used criterion to explain the spatial distributions of aftershocks[4–8], but its applicability has been disputed[9–11]. Here we use a deep-learning approach to identify a static-stress-based criterion that forecasts aftershock locations without prior assumptions about fault orientation. We show that a neural network trained on more than 131,000 mainshock–aftershock pairs can predict the locations of aftershocks in an independent test dataset of more than 30,000 mainshock–aftershock pairs more accurately (area under curve of 0.849) than can classic Coulomb failure stress change (area under curve of 0.583). We find that the learned aftershock pattern is physically interpretable: the maximum change in shear stress, the von Mises yield criterion (a scaled version of the second invariant of the deviatoric stress-change tensor) and the sum of the absolute values of the independent components of the stress-change tensor each explain more than 98 per cent of the variance in the neural-network prediction. This machine-learning-driven insight provides improved forecasts of aftershock locations and identifies physical quantities that may control earthquake triggering during the most active part of the seismic cycle.

The deep-learning aftershock location forecasts that we have developed are trained and tested using co-seismic slip distributions from the SRCMOD online database of finite-fault rupture models (http://equake-rc.info/SRCMOD/). We calculated elastic stress-change tensors for 199 of the SRCMOD slip distributions (118 distinct mainshocks; Supplementary Table 1) at the centroids of 5 km × 5 km × 5 km cells in a volume extending 100 km horizontally from each mainshock rupture plane and 50 km vertically[12]. The aftershocks that occurred between one second and one year after the mainshocks in each grid cell (162,741 aftershocks in total) were compiled from the International Seismological Center (ISC) event catalogue. By discretizing the volume around each mainshock in this way, aftershock forecasting can be formulated as a large-scale binary classification problem, with the goal of accurately classifying each 5 km × 5 km × 5 km grid cell in the volume around each mainshock as either 'containing aftershocks' or 'not containing aftershocks'.

Neural networks are machine-learning algorithms that are well suited and widely used to classify data[13]. The neural networks used here are fully connected and have six hidden layers with 50 neurons each and hyperbolic tangent activation functions (13,451 weights and biases in total). The first layer corresponds to the inputs to the neural network; in this case, these inputs are the magnitudes of the six independent components of the co-seismically generated static elastic stress-change tensor calculated at the centroid of a grid cell and their negative values. In neural networks designed for binary classification problems, the final layer is often a single sigmoid. In our case, the output of this final neuron may be interpreted as the predicted probability that a grid cell generates one or more aftershocks.

The stress changes and aftershock locations associated with about 75% of randomly selected distinct mainshocks were used as training data; the remaining 25% were reserved to test the trained neural networks. The training and testing datasets both consist of the elements of the stress-change tensor as features and the corresponding labels of either 0, for grid cells without aftershocks, or 1, for grid cells with aftershocks.

We assess the accuracy of the neural-network aftershock location forecasts on the test dataset using receiver operating characteristic (ROC) analysis. ROC curves are widely used to assess the efficacy of diagnostic medical tests. To build these curves, the true positive rate of a binary classifier is plotted against the false positive rate for all possible thresholds of the classifier (see Methods for more details). The area under an ROC curve (AUC) then quantifies the overall performance of a test across all thresholds (Fig. 1). The ROC analysis reveals that the neural-network forecast can explain aftershock locations better than can widely used metrics: the merged AUC value across all slip distributions and grid cells in the test dataset for the neural-network forecast is 0.849, which is larger than that of the classic Coulomb failure stress criterion[3] (AUC = 0.583) resolved on receiver planes parallel to the average orientation of the mainshock fault ($\Delta CFS(\mu = 0.4)$, in which μ is the effective coefficient of friction). Neither classifier has particularly high precision, defined as the percentage of grid cells predicted to be positive that actually are positive: the overall precision associated with $\Delta CFS(\mu = 0.4)$ at a cut-off threshold of 0.01 MPa (Methods) is 3% and that of the neural-network classifier at a threshold of 0.5 is 6% (Fig. 2a, d). Permutation tests (Methods) reveal that the neural-network forecast is significantly better than random assignment for most of the slip distributions in the test dataset: the mean empirical P value is 0.026 across all 57 distributions, and only four distributions are associated with empirical P values larger than 0.1. Additional tests, based on realizations of the training and test datasets that incorporate only one slip distribution per mainshock and variable limits on grid-cell depth depending on the depth of each slip distribution are included in Methods.

The spatial pattern of the deep-learning location forecast can be visualized for the idealized synthetic reference case of an earthquake with a uniform 1 m of slip on a 60-km-long right-lateral strike-slip fault (moment magnitude $M_w \approx 7.0$, Fig. 2; see Extended Data Fig. 1 for an idealized dip-slip case). A location forecast based on the Coulomb failure stress criterion[3] for this idealized strike-slip fault would assign a low risk of aftershocks adjacent to the mainshock rupture plane and a heightened risk in lobes extending from the termini of the mainshock rupture plane (Fig. 2e). By contrast, the learned forecast developed here suggests that aftershock risk may be heightened within 10 km of the mainshock fault in all directions (Fig. 2h). This deep-learning forecast is therefore not consistent with the idea of well-defined stress shadows[6] immediately adjacent to the mainshock.

The learned forecast (Fig. 2h) has implications for the physics of aftershock triggering and earthquake generation. Qualitatively, the

[1]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA. [2]Center for Integrative Geosciences and Department of Physics, University of Connecticut, Storrs, CT, USA. [3]Google, Cambridge, MA, USA. *e-mail: phoebe.devries@uconn.edu

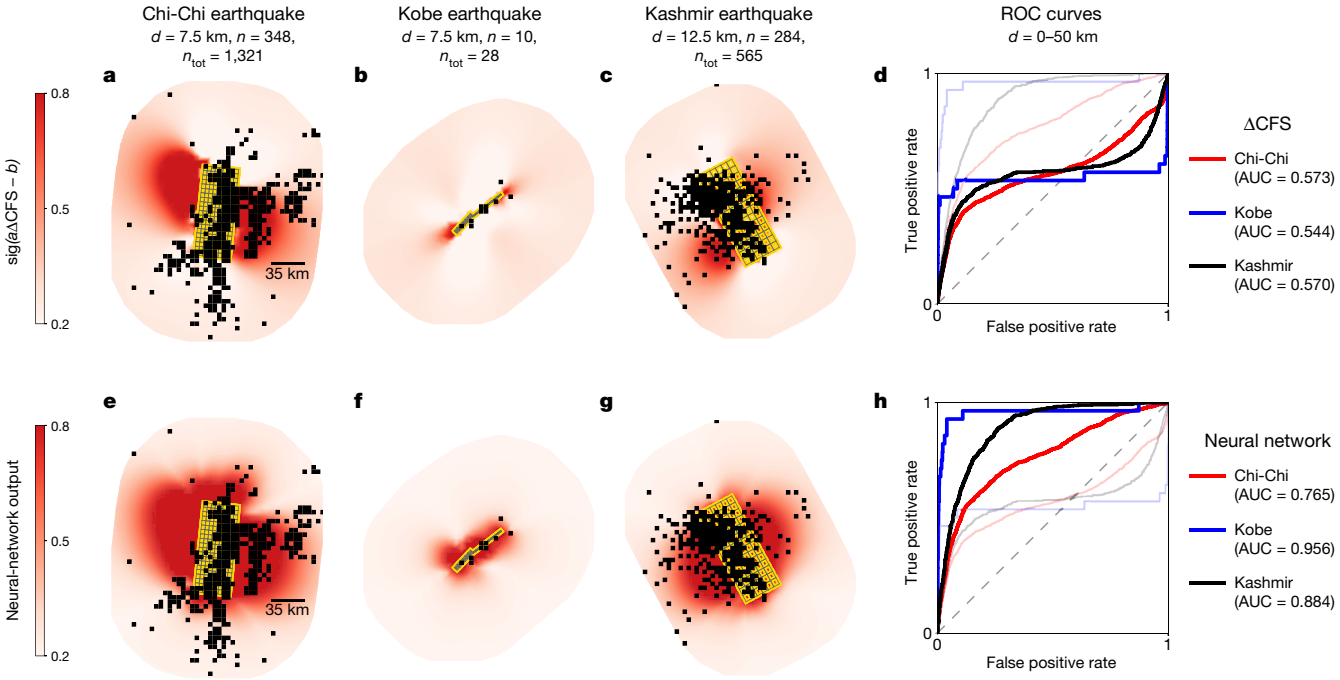


Fig. 1 | Mainshock–aftershock examples. a–d, Spatial patterns of $\Delta\mathrm{CFS}(\mu = 0.4)$ for the 1999 $M_w = 7.7$ Chi-Chi earthquake[17] at a depth of 7.5 km (**a**), the 1995 $M_w = 6.9$ Kobe earthquake[18] at a depth of 7.5 km (**b**) and the 2005 $M_w = 7.6$ Kashmir earthquake[19] at a depth of 12.5 km (**c**), along with ROC curves for all three earthquakes across all depths (**d**). In **a**–**c**, n refers to the number of positive grid cells at the depth shown and n_{tot} is the number of positive grid cells across all depths. A 1:1 grey dashed line is included in **d** for reference. Because of possible sign ambiguities, we calculate four versions of $\Delta\mathrm{CFS}(\mu = 0.4)$ and use the best-performing sign convention for each slip distribution. In **a**–**c**, $\Delta\mathrm{CFS}(\mu = 0.4)$ values (in megapascals) are fed through a sigmoid filter $\mathrm{sig}(x) = 1/(1 + e^{-x})$ ($\mathrm{sig}(a\Delta\mathrm{CFS}(\mu = 0.4) - b)$, with $a = 10$, $b = 1$; colour scale) to facilitate comparison to the neural network; faults are shown in yellow and grid cells that contain aftershocks are shown in black. **e**–**h**, Analogous to **a**–**d** but for the neural network. To facilitate easy comparison, the ROC curves in **d** are plotted as pale lines in **h** and the ROC curves in **h** are plotted as pale lines in **d**.

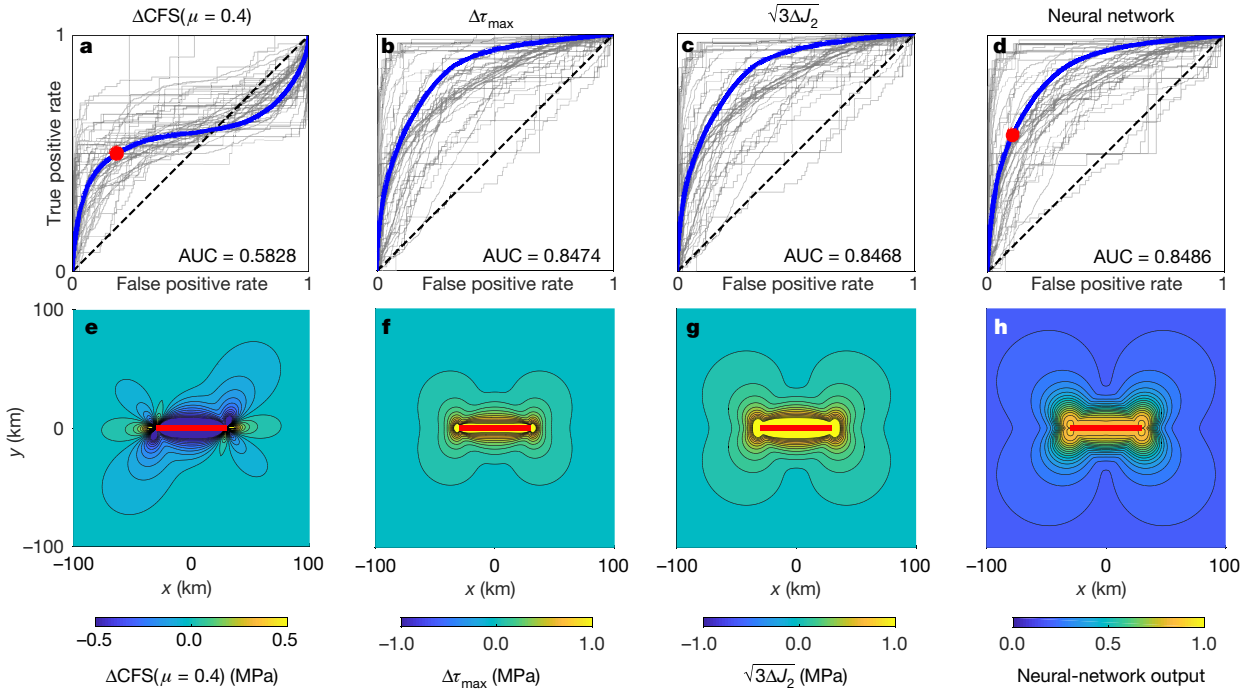


Fig. 2 | Comparison of performance. a–d, ROC curves for every slip distribution in the test dataset (grey curves) for $\Delta\mathrm{CFS}(\mu = 0.4)$ (**a**), $\Delta\tau_{max}$ (**b**), $\sqrt{3\Delta J_2}$ (**c**) and the neural network (**d**). Merged ROC curves are shown in blue and the associated AUC values are listed. The red circles in **a** and **d** highlight the thresholds of 0.01 MPa and 0.5, respectively. **e**–**h**, For a synthetic case of a 60-km-long, right-lateral strike-slip fault (red lines) at a depth of 10 km, we show a comparison of the spatial patterns of $\Delta\mathrm{CFS}(\mu = 0.4)$ (**e**), $\Delta\tau_{max}$ (**f**), $\sqrt{3\Delta J_2}$ (**g**) and the neural network (**h**), averaged over the fault strike.

strike-averaged neural-network forecast appears to closely resemble the spatial patterns of the maximum change in shear stress ($\Delta\tau_{max}$) and the von Mises yield criterion ($\sqrt{3\Delta J_2}$, in which ΔJ_2 is the second invariant of the deviatoric stress-change tensor; Fig. 2). To examine these potential links to physical quantities quantitatively, we compare a suite of scalar static-stress metrics (including the invariants of the stress-change tensor, Coulomb failure stress change and maximum shear stress change; see Methods)—after they are scaled, shifted and normalized with a sigmoid filter—to the neural-network forecast (Fig. 2h). In addition to Coulomb failure stress change, several of the quantities, including shear stress changes and the invariants of the stress change tensor, have been proposed and used successfully in previous studies of aftershock patterns[3,14–16]. Of the metrics considered, the maximum change in shear stress, the von Mises yield criterion and the sum of the absolute values of the six independent components of the stress change tensor can explain the largest percentages (more than 98%, or $R^2 > 0.98$ for both the idealized strike-slip and dip-slip cases) of the variance in the strike-averaged learned forecast (see, for example, Fig. 2; Extended Data Fig. 1) within a 300 km × 300 km area centred on the fault. The last quantity is particularly interesting because it is not invariant under rotation and to our knowledge has not previously been proposed as a quantity that could explain aftershock location patterns. A few other quantities, such as the sum of the absolute values of the shear stress changes on fault-parallel receiver planes can also explain more than 90% of the variance in the strike-slip neural-network forecast (Extended Data Table 1). In other words, without any assumptions about receiver plane orientation or geometry, the neural network identified an aftershock location forecast that is strongly correlated with a small number of physical quantities, most notably the sum of the absolute values of the six independent components of the stress-change tensor, the von Mises yield criterion and the maximum change in shear stress. These results highlight how deep-learning approaches can lead to improved aftershock forecasts and provide physical insights into the mechanisms of earthquake triggering.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0438-y.

Received: 11 November 2017; Accepted: 10 July 2018;
Published online 29 August 2018.

1. Båth, M. Lateral inhomogeneities of the upper mantle. *Tectonophysics* **2**, 483–514 (1965).
2. Utsu, T. A statistical study on the occurrence of aftershocks. *Geophys. Mag.* **30**, 521–605 (1961).
3. King, G. C., Stein, R. S. & Lin, J. Static stress changes and the triggering of earthquakes. *Bull. Seismol. Soc. Am.* **84**, 935–953 (1994).
4. Toda, S., Stein, R. S., Reasenberg, P. A., Dieterich, J. H. & Yoshida, A. Stress transferred by the 1995 $M_w = 6.9$ Kobe, Japan, shock: effect on aftershocks and future earthquake probabilities. *J. Geophys. Res.* **103**, 24543–24565 (1998).
5. Parsons, T., Stein, R. S., Simpson, R. W. & Reasenberg, P. A. Stress sensitivity of fault seismicity: a comparison between limited-offset oblique and major strike-slip faults. *J. Geophys. Res.* **104**, 20183–20202 (1999).
6. Reasenberg, P. A. & Simpson, R. W. Response of regional seismicity to the static stress change produced by the Loma Prieta earthquake. *Science* **255**, 1687–1690 (1992).
7. Jacques, E., King, G. C. P., Tapponnier, P., Ruegg, J. C. & Manighetti, I. Seismic activity triggered by stress changes after the 1978 events in the Asal Rift, Djibouti. *Geophys. Res. Lett.* **23**, 2481–2484 (1996).
8. Nostro, C., Cocco, M. & Belardinelli, M. E. Static stress changes in extensional regimes: an application to southern Apennines (Italy). *Bull. Seismol. Soc. Am.* **87**, 234–248 (1997).
9. Hardebeck, J. L., Nazareth, J. J. & Hauksson, E. The static stress change triggering model: constraints from two southern California aftershock sequences. *J. Geophys. Res.* **103**, 24427–24437 (1998).
10. Mallman, E. P. & Zoback, M. D. Assessing elastic Coulomb stress transfer models using seismicity rates in southern California and southwestern Japan. *J. Geophys. Res.* **112**, B03304 (2007).
11. Felzer, K. R. & Brodsky, E. E. Testing the stress shadow hypothesis. *J. Geophys. Res.* **110**, B05S09 (2005).
12. Okada, Y. Internal deformation due to shear and tensile faults in a half-space. *Bull. Seismol. Soc. Am.* **82**, 1018–1040 (1992).
13. Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
14. Das, S. & Scholz, C. H. Off-fault aftershock clusters caused by shear stress increase? *Bull. Seismol. Soc. Am.* **71**, 1669–1675 (1981).
15. Kagan, Y. Y. & Jackson, D. D. Spatial aftershock distribution: effect of normal stress. *J. Geophys. Res.* **103**, 24453–24467 (1998).
16. Meade, B. J., DeVries, P., Faller, J., Viegas, F. & Wattenberg, M. What is better than Coulomb failure stress? A ranking of scalar static stress triggering mechanisms from 105 mainshock–aftershock pairs. *Geophys. Res. Lett.* **44**, 11409–11416 (2017).
17. Ma, K. F., Song, T. R. A., Lee, S. J. & Wu, H. I. Spatial slip distribution of the September 20, 1999, Chi-Chi, Taiwan, earthquake (M_w7.6)—inverted from teleseismic data. *Geophys. Res. Lett.* **27**, 3417–3420 (2000).
18. Yoshida, S. et al. Joint inversion of near- and far-field waveforms and geodetic data for the rupture process of the 1995 Kobe earthquake. *J. Phys. Earth* **44**, 437–454 (1996).
19. Shao, G. & Ji, C. Preliminary result of the Oct 8, 2005 Mw 7.64 Pakistan earthquake. *UCSB* http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2005/10/smooth/2005pakistan.html (accessed 2 June 2018).

Acknowledgements This work was supported by Harvard University and Google. The computations in this paper were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

Reviewer information *Nature* thanks D. Trugman and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions All authors conceived the idea for this paper; P.M.R.D. and B.J.M. implemented the analysis and wrote the paper.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0438-y.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0438-y.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to P.M.R.D.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Neural-network structure and development. To build and train the neural networks, we use the Python toolkit Keras (https://keras.io), which provides a high-level application programming interface to access the Theano[20] (http://deeplearning.net/software/theano) and TensorFlow[21] (https://www.tensorflow.org) deep-learning libraries. We train the networks using Theano, an adaptive learning rate (Adadelta) optimization method[22], and a binary cross-entropy cost function. Negative training data (grid cells without aftershocks) are downsampled during training. During training, 10% of the positive training data samples were used for validation, along with an equal number of randomly selected negative samples.

Aftershock catalogue. The aftershocks that occurred between one second and one year after the mainshocks in each grid cell were compiled from the reviewed ISC event catalogue. Note that for seven earthquakes in 2012, the time period included is shorter (as short as about one month for the $M_w = 7.7$ Masset, Canada, earthquake) because the catalogue ends on 30 November 2012.

ROC curve analysis. ROC curves are widely used to evaluate the efficacy of binary classifiers and diagnostic tests in medicine, machine learning and many other fields. To construct an ROC curve, the true positive rate (the ratio of the number of true positive classifications to the total number of positives, or in this case the ratio of the number of grid cells correctly identified as containing aftershocks to the total number of grid cells containing aftershocks) is plotted against the false positive rate (the ratio of the number of false positive classifications to the total number of negatives, or in this case the ratio of the number of grid cells incorrectly identified as containing aftershocks to the number of grid cells that do not contain aftershocks) for the range of possible test thresholds. In this way, an ROC curve represents the performance of a test across all possible thresholds. A binary classification method that is no better than random assignment would plot near the 1:1 line, whereas a test that is more effective than random assignment would plot above the 1:1 line. The AUC quantifies the overall performance of a test; a test that is no better than random assignment would correspond to AUC = 0.5, whereas more accurate tests would have AUC values approaching 1. Note that the 2010 Darfield, New Zealand, slip distribution[23] in the test dataset is excluded from the merged AUC value for Coulomb failure stress change (Fig. 2a); the geometric complexity of this rupture precludes meaningful definitions of average strike, rake and dip.

Permutation tests. We perform permutation tests to assess the statistical significance of the performance of the neural-network forecast on the test dataset. For each mainshock slip distribution in the test set, we generated 5,000 random realizations of the locations of positive grid cells. Each of the resulting random ROC curve realizations yields an AUC value. The observed AUC value is compared with the empirical distribution of random AUC values from the permutation tests to obtain a one-sided empirical P value for every slip distribution in the test set.

Quantitative comparison to existing stress metrics for an idealized case. The 40 stress metrics that we consider include the nearest distance to the mainshock rupture, the maximum change in shear stress, the three invariants of the stress-change tensor, the Coulomb failure stress change on receiver planes parallel to the mainshock fault plane (with coefficient of friction $\mu = 0.0$, 0.2, 0.4, 0.6 or 0.8), the total Coulomb failure stress change, the total shear stress change on fault-parallel receiver planes, and the normal-only component of Coulomb failure stress change (see Extended Data Table 1 for mathematical definitions of these quantities). Note that in Figs. 1 and 2a, owing to possible sign ambiguities, we calculate four versions of classic Coulomb failure stress change and use the best-performing sign convention for each slip distribution. In Extended Data Table 1, all of the metrics and their magnitudes are considered for both the deviatoric and full stress tensors in a 300 km × 300 km area centred on the idealized fault. To enable a quantitative and meaningful comparison of the spatial patterns of these stress metrics to the deep-learning forecast, we first scale and shift each stress metric, then feed these values through a sigmoid filter to normalize between 0 and 1.

Constraining grid-cell depths to 5 km beyond the maximum depth of each slip distribution. The results presented in the main text are based on training and test datasets (Supplementary Table 1; earthquake source data taken from refs [17–19,23–189]) that incorporate grid cells down to 50 km for each slip distribution. We originally chose this fixed depth cut-off to keep the analysis as clean, clear and consistent as possible.

In addition, we consider an alternative approach, in which the maximum depth of grid cells considered varies with the maximum depth of each slip distribution. Here we present analogous results to those in the main text, but using training and test datasets that incorporate only grid cells shallower than 5 km below the deepest depth of the slip distribution, although no deeper than 50 km. As in the original training and test datasets, we define these alternative datasets by the random assignments listed in Supplementary Table 1. In other words, for a slip distribution that extends to a depth of 17.5 km, we incorporate all grid cells with centroids at depths shallower than 22.5 km in the test and training datasets specified in Supplementary Table 1.

The size of the training and test datasets are greatly reduced when variable depth limits are incorporated: for the training dataset, the number of total grid cells is reduced by roughly 20% (from 4,743,090 to 3,779,070), while the total number of aftershocks is also reduced by around 3% (from 131,804 to 127,735). For some individual slip distributions, such as for the Landers earthquake, the number of grid cells is reduced by about 50%–60% because the slip distributions extend to depths of 15–18 km, whereas for others—such as for the Tohoku and 2004 Sumatra earthquakes—the number of grid cells incorporated changes by 10% or less because these slip distributions extend to depths of greater than 40 km. We exclude two slip distributions from the analysis when variable depth limits are incorporated: the 2 December 1996 $M_w = 6.7$ Hyuga-nada earthquake (see Supplementary Table 1) in the test dataset and the 1968 $M_w = 7.5$ Hyuga-nada earthquake (see Supplementary Table 1) in the training dataset. When variable depth limits are incorporated for these events, the ISC catalogue does not contain any aftershocks in the volume surrounding the faults.

With these variable-depth-limit training and test datasets, and using a neural network with the same structure as before, we obtain markedly similar results (Extended Data Figs. 2, 3) to those obtained using the original versions of the datasets (Figs. 1, 2). A neural network trained on the variable-depth-limit version of the training dataset yields a merged AUC value across all slip distributions and grid cells in the variable-depth-limit version of the test dataset of 0.8333, which is larger than the classic Coulomb failure stress-change criterion (AUC = 0.5804). For this variable-depth-limit case, the neural-network classifier performs similarly well as, although not better than, the maximum change in shear stress (AUC = 0.8383) and the von Mises yield criterion (AUC = 0.8378). As in the fixed-depth-limit example, more than 98% of the trained-neural-network forward prediction for an idealized strike-slip fault (Extended Data Fig. 3h) can be explained by the maximum change in shear stress, the von Mises yield criterion and the sum of the absolute values of the independent components of the stress-change tensor.

Realizations of the test and training datasets with one slip distribution per mainshock. In the main text, we present results using training and test datasets that are defined by distinct mainshocks, rather than distinct slip distributions. In other words, in each dataset, there may be multiple slip distributions from the same event (see Supplementary Table 1). We took this approach to include as much data as possible; however, certain effects—and perhaps biases—could have been introduced as a result.

To address this concern, here we include results from an ensemble approach to training and testing. In this approach, we generate ten realizations of pairs of training and test datasets. Each realization of the training and test datasets includes only one randomly selected slip distribution per mainshock (Supplementary Table 2). For each of the ten realizations, we used this ensemble approach for both an analysis that incorporated grid cells down to 50 km for each slip distribution and an analysis that incorporated grid cells down to only 5 km beyond the maximum depth of each slip distribution. Thus, a total of 20 networks were trained and tested using the ensemble approach (Supplementary Table 2), with the structure of each the same as that of the original neural network.

Taking this approach substantially reduces the size of the training and test datasets. With grid cells down to 50 km incorporated, the mean number of positive grid cells included in each realization of the training dataset is 33,895.7 (with a minimum number of positive grid cells of 33,670 in realization 7 and a maximum of 34,197 in realization 1). With only grid cells down to 5 km beyond the deepest depth of the slip distributions incorporated, the mean number of positive grid cells included in each realization of the training dataset is further reduced to 31,894.9 (with a minimum number of positive grid cells of 31,162 in realization 5 and a maximum of 32,318 in realization 3). By comparison, in the original training dataset (which incorporates grid cells down to 50 km uniformly), the total number of positive grid cells is 85,850.

Rather than including the ROC curves from all of the realizations of training and test dataset pairs, we instead summarize the relative performances of $\Delta\mathrm{CFS}(\mu = 0.4)$, $\Delta\tau_{\max}$, $\sqrt{3\Delta J_2}$ and the associated trained neural networks for each realization (Extended Data Table 2). ROC curves for realization 6—for which the trained neural network performs the worst in terms of the AUC—are shown in Extended Data Fig. 4.

Overall, we obtain similar results to those presented in the main text when using realizations of training and test datasets with one randomly selected slip distribution per event. The AUC values listed in Extended Data Table 2 are comparable to, and in 18 of 20 cases larger than, the AUC values associated with $\Delta\tau_{\max}$ and $\sqrt{3\Delta J_2}$ evaluated on the same test dataset realizations. The smallest AUC value associated with the trained neural networks is 0.78 (realization 6, incorporating only grid cells down to 5 km beyond the deepest depth of each slip distribution; Extended Data Fig. 4) and the largest AUC value associated with $\Delta\mathrm{CFS}(\mu = 0.4)$ is 0.626 (realization 6, but incorporating grid cells down to 50 km; Extended Data Fig. 4). Furthermore, the forward predictions of all trained neural networks (Extended Data Figs. 5, 6) are qualitatively similar to those displayed in Fig. 2h and Extended Data Fig. 3h.

Data and code availability. This project is based on the freely available Keras (https://keras.io) and Theano[21] (http://deeplearning.net/software/theano) libraries

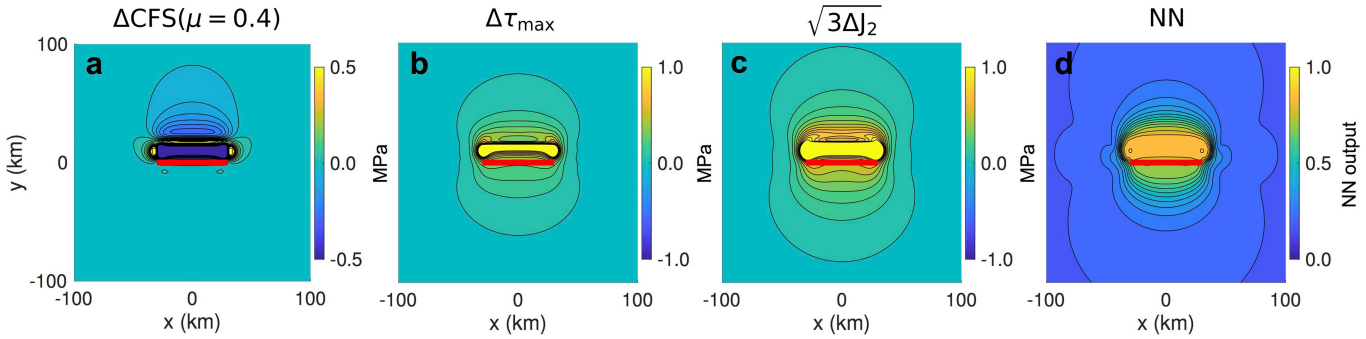
as well as T. B. Thompson's Okada wrapper (https://github.com/tbenthompson/okada_wrapper). All code is available at https://www.github.com/phoebemrdevries. All data are freely available from the SRCMOD catalogue (http://equake-rc.info/SRCMOD) and the ISC event catalogue (http://www.isc.ac.uk/iscgem).

20. The Theano Development Team. Theano: a python framework for fast computation of mathematical expressions. Preprint at http://arxiv.org/abs/1605.02688 (2016).
21. Abadi, M. et al. TensorFlow: large-scale machine learning on heterogeneous systems. In *Proc. 12th USENIX Symposium on Operating Systems Design and Implementation* 265–283 (USENIX Association, 2016).
22. Zeiler, M. D. ADADELTA: an adaptive learning rate method. Preprint at https://arxiv.org/abs/1212.5701 (2012).
23. Atzori, S. et al. The 2010–2011 Canterbury, New Zealand, seismic sequence: multiple source analysis from InSAR data and modeling. *J. Geophys. Res.* **117**, B08305 (2012).
24. Yagi, Y., Kikuchi, M., Yoshida, S. & Yamanaka, Y. Source process of the Hyuga-nada Earthquake of April 1, 1968 (M_{JMA}7.5), and its relationship to the subsequent seismicity. *Zishin J. Seis. Soc. Japan*. **51**, 139–148 (1998) [in Japanese].
25. Nagai, R., Kikuchi, M. & Yamanaka, Y. Comparative study on the source processes of recurrent large earthquakes in Sariku-oki Region: the 1968 Tokachi-oki earthquake and the 1994 Sanriku-oki earthquake. *Zishin J. Seis. Soc. Japan* **54**, 267–280 (2001) [in Japanese].
26. Takeo, M. Fault heterogeneity of inland earthquakes in Japan. *Bull. Earthq. Res. Inst. Univ. Tokyo* **65**, 541–569 (1990).
27. Heaton, T. H. The 1971 San Fernando earthquake: a double event? *Bull. Seismol. Soc. Am.* **72**, 2037–2062 (1982).
28. Hartzell, S. & Langer, C. Importance of model parameterization in finite fault inversions; application to the 1974 M_w 8.0 Peru earthquake. *J. Geophys. Res.* **98**, 22123–22134 (1993).
29. Liu, H. & Helmberger, D. V. The near-source ground motion of the 6 August 1979 Coyote Lake, California, earthquake. *Bull. Seismol. Soc. Am.* **73**, 201–218 (1983).
30. Mendoza, C. Finite-fault analysis of the 1979 March 14 Petatlan, Mexico, earthquake using teleseismic P-wave-forms. *Geophys. J. Int.* **121**, 675–683 (1995).
31. Takeo, M. Rupture process of the 1980 Izu-Hanto-Toho-Oki earthquake deduced from strong motion seismograms. *Bull. Seismol. Soc. Am.* **78**, 1074–1091 (1988).
32. Hartzell, S., Langer, C. & Mendoza, C. Rupture histories of eastern North American earthquakes. *Bull. Seismol. Soc. Am.* **84**, 1703–1724 (1994).
33. Mendoza, C. & Hartzell, S. H. Inversion for slip distribution using teleseismic P waveforms: North Palm Springs, Borah Peak, and Michoacan earthquakes. *Bull. Seismol. Soc. Am.* **78**, 1092–1111 (1988).
34. Fukuyama, E. & Irikura, K. Rupture process of the 1983 Japan Sea (Akita-Oki) earthquake using a waveform inversion method. *Bull. Seismol. Soc. Am.* **76**, 1623–1640 (1986).
35. Hartzell, S. H. & Heaton, T. H. Rupture history of the 1984 Morgan Hill, California, earthquake from the inversion of strong motion records. *Bull. Seismol. Soc. Am.* **76**, 649–674 (1986).
36. Takeo, M. & Mikami, N. Inversion of strong motion seismograms for the source process of the Naganoken-Seibu earthquake of 1984. *Tectonophysics* **144**, 271–285 (1987).
37. Mendoza, C., Hartzell, S. & Monfret, T. Wide-band analysis of the 3 March 1985 Central Chile earthquake: overall source process and rupture history. *Bull. Seismol. Soc. Am.* **84**, 269–283 (1994).
38. Mendoza, C. Coseismic slip of two large Mexican earthquakes from teleseismic body wave-forms: implications for asperity interaction in the Michoacan Plate Boundary Segment. *J. Geophys. Res.* **98**, 8197–8210 (1993).
39. Hartzell, S. Comparison of seismic waveform inversion results for the rupture history of a finite fault: application to the 1986 North Palm-Springs, California, earthquake. *J. Geophys. Res.* **94**, 7515–7534 (1989).
40. Larsen, S., Reilinger, R., Neugebauer, H. & Strange, W. Global positioning system measurements of deformations associated with the 1987 Superstition Hills earthquake: evidence for conjugate faulting. *J. Geophys. Res.* **97**, 4885–4902 (1992).
41. Wald, D. J., Helmberger, D. V. & Hartzell, S. H. Rupture process of the 1987 Superstition Hills earthquake from the inversion of strong-motion data. *Bull. Seismol. Soc. Am.* **80**, 1079–1098 (1990).
42. Hartzell, S. H. & Iida, M. Source complexity of the 1987 Whittier Narrows, California, earthquake from the inversion of strong motion records. *J. Geophys. Res.* **95**, 12475–12485 (1990).
43. Emolo, A. & Zollo, A. Kinematic source parameters for the 1989 Loma Prieta earthquake from the nonlinear inversion of accelerograms. *Bull. Seismol. Soc. Am.* **95**, 981–994 (2005).
44. Steidl, J. H., Archuleta, R. J. & Hartzell, S. H. Rupture history of the 1989 Loma Prieta, California, earthquake. *Bull. Seismol. Soc. Am.* **81**, 1573–1602 (1991).
45. Wald, D. J., Helmberger, D. V. & Heaton, T. H. Rupture model of the 1989 Loma Prieta earthquake from the inversion of strong-motion and broad-band teleseismic data. *Bull. Seismol. Soc. Am.* **81**, 1540–1572 (1991).
46. Hough, S. E. & Dreger, D. S. Source parameters of the 23 April 1992 *M* 6.1 Joshua Tree, California, earthquake and its aftershocks: empirical Green's function analysis of GEOS and TERRAscope data. *Bull. Seismol. Soc. Am.* **85**, 1576–1590 (1995).
47. Cohee, B. P. & Beroza, G. C. Slip distribution of the 1992 Landers earthquake and its implications for earthquake source mechanics. *Bull. Seismol. Soc. Am.* **84**, 692–712 (1994).
48. Cotton, F. & Campillo, M. Frequency-domain inversion of strong motions: application to the 1992 Landers Earthquake. *J. Geophys. Res.* **100**, 3961–3975 (1995).
49. Hernandez, B., Cotton, F. & Campillo, M. Contribution of radar interferometry to a two-step inversion of the kinematic process of the 1992 Landers earthquake. *J. Geophys. Res.* **104**, 13083–13099 (1999).
50. Wald, D. J. & Heaton, T. H. Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bull. Seismol. Soc. Am.* **84**, 668–691 (1994).
51. Zeng, Y. & Anderson, J. *Evaluation of Numerical Procedures for Simulating Near-Fault Long-Period Ground Motions using Zeng Method.* Report No. 2000/01 to the PEER Utilities Program (Pacific Earthquake Engineering Research Center, UC Berkeley, 2000).
52. Mendoza, C. & Fukuyama, E. The July 12, 1993, Hokkaido-Nansei-Oki, Japan, earthquake: coseismic slip pattern from strong-motion and teleseismic recordings. *J. Geophys. Res.* **101**, 791–801 (1996).
53. Hartzell, S., Liu, P. C. & Mendoza, C. The 1994 Northridge, California, earthquake; investigation of rupture velocity, risetime, and high-frequency radiation. *J. Geophys. Res.* **101**, 20091–20108 (1996).
54. Hudnut, K. W. et al. Co-seismic displacements of the 1994 Northridge, California, earthquake. *Bull. Seismol. Soc. Am.* **86**, S19–S36 (1996).
55. Shen, Z.-K. et al. Northridge earthquake rupture models based on the global positioning system measurements. *Bull. Seismol. Soc. Am.* **86**, S37–S48 (1996).
56. Wald, D. J., Heaton, T. H. & Hudnut, K. W. The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS, and leveling data. *Bull. Seismol. Soc. Am.* **86**, S49–S70 (1996).
57. Nakayama, W. & Takeo, M. Slip history of the 1994 Sanriku-Haruka-Oki, Japan, earthquake deduced from strong-motion data. *Bull. Seismol. Soc. Am.* **87**, 918–931 (1997).
58. Mendoza, C. & Hartzell, S. Fault-slip distribution of the 1995 Colima-Jalisco, Mexico, earthquake. *Bull. Seismol. Soc. Am.* **89**, 1338–1344 (1999).
59. Courboulex, F., Santoyo, M. A., Pacheco, J. F. & Singh, S. K. The 14 September 1995 ($M = 7.3$) Copala, Mexico, earthquake: a source study using teleseismic, regional, and local data. *Bull. Seismol. Soc. Am.* **87**, 999–1010 (1997).
60. Yagi, Y., Kikuchi, M., Yoshida, S. & Sagiya, T. Comparison of the coseismic rupture with the aftershock distribution in the Hyuga-nada earthquakes of 1996. *Geophys. Res. Lett.* **26**, 3161–3164 (1999).
61. Salichon, J. et al. Joint inversion of broadband teleseismic and interferometric synthetic aperture radar (InSAR) data for the slip history of the $M_w = 7.7$, Nazca ridge (Peru) earthquake of 12 November 1996. *J. Geophys. Res.* **108**, 2085 (2003).
62. Hernandez, B. et al. Rupture history of the 1997 Umbria-Marche (central Italy) main shocks from the inversion of GPS, DInSAR and near field strong motion data. *Ann. Geophys.* **47**, 1355–1376 (2004).
63. Horikawa, H. Earthquake doublet in Kagoshima, Japan: rupture of asperities in a stress shadow. *Bull. Seismol. Soc. Am.* **91**, 112–127 (2001).
64. Sudhaus, H. & Jònsson, S. Source model for the 1997 Zirkuh earthquake ($M_w = 7.2$) in Iran derived from JERS and ERS InSAR observations. *Geophys. J. Int.* **185**, 676–692 (2011).
65. Ide, S. Complex source processes and the interaction of moderate earthquakes during the earthquake swarm in the Hida-Mountains, Japan, 1998. *Tectonophysics* **334**, 35–54 (2001).
66. Miyakoshi, K., Kagawa, T., Sekiguchi, H., Iwata, T. & Irikura, K. Source characterization of inland earthquakes in Japan using source inversion results. In *Proc. 12th World Conference on Earthquake Engineering abstr.* 1850 (New Zealand Society for Earthquake Engineering, 2000).
67. Nakahara, H. et al. Broadband source process of the 1998 Iwate prefecture, Japan, earthquake as revealed from inversion analyses of seismic waveforms and envelopes. *Bull. Seismol. Soc. Am.* **92**, 1708–1720 (2002).
68. Ji, C., Wald, D. J. & Helmberger, D. V. Source description of the 1999 Hector Mine, California, earthquake, part II: complexity of slip history. *Bull. Seismol. Soc. Am.* **92**, 1208–1226 (2002).
69. Salichon, J., Lundgren, P., Delouis, B. & Giardini, D. Slip history of the 16 October 1999 M_w 7.1 Hector Mine earthquake (California) from the inversion of InSAR, GPS, and teleseismic data. *Bull. Seismol. Soc. Am.* **94**, 2015–2027 (2004).
70. Bouchon, M. et al. Space and time evolution of rupture and faulting during the 1999 Izmit (Turkey) earthquake. *Bull. Seismol. Soc. Am.* **92**, 256–266 (2002).
71. Çakir, Z. et al. Coseismic and early post-seismic slip associated with the 1999 Izmit earthquake (Turkey), from SAR interferometry and tectonic field observations. *Geophys. J. Int.* **155**, 93–110 (2003).
72. Delouis, B., Giardini, D., Lundgren, P. & Salichon, J. Joint inversion of InSAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip: application to the 1999 Izmit mainshock. *Bull. Seismol. Soc. Am.* **92**, 278–299 (2002).
73. Reilinger, R. E. et al. Coseismic and postseismic fault slip for the 17 August 1999, $M = 7.5$, Izmit, Turkey earthquake. *Science* **289**, 1519–1524 (2000).
74. Yagi, Y. & Kikuchi, M. Source rupture process of the Kocaeli, Turkey, earthquake of August 17, 1999, obtained by joint inversion of near-field data and teleseismic data. *Geophys. Res. Lett.* **27**, 1969–1972 (2000).
75. Copley, A., Avouac, J. P., Hollingsworth, J. & Leprince, S. The 2001 Mw 7.6 Bhuj earthquake, low fault friction, and the crustal support of plate driving forces in India. *J. Geophys. Res.* **116**, B08405 (2011).

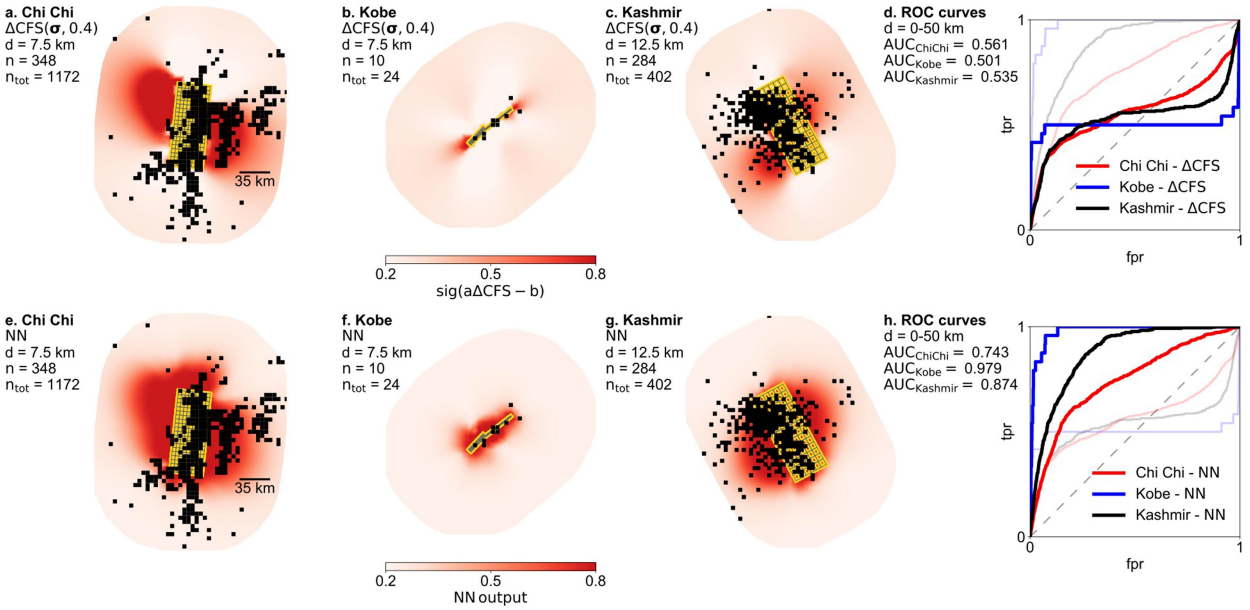
76. Copley, A. Source models of large earthquakes: Jan/26/2001 (Mw 7.6), Bhuj, India. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2001_Bhuj/index.html (accessed 1 July 2013).
77. Yagi, Y. A slip model for the Jan 26, 2001 Bhuj (India) earthquake using teleseismic recordings. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2001BHUJIN01YAGI (accessed 18 May 2003).
78. Asano, K., Iwata, T. & Irikura, K. Estimation of source rupture process and strong ground motion simulation of the 2002 Denali, Alaska, earthquake. *Bull. Seismol. Soc. Am.* **95**, 1701–1715 (2005).
79. Poiata, N., Miyake, H., Koketsu, K. & Hikima, K. Strong motion and teleseismic waveform inversions for the source process of the 2003 Bam, Iran, earthquake. *Bull. Seismol. Soc. Am.* **102**, 1477–1496 (2012).
80. Semmane, F., Campillo, M. & Cotton, F. Fault location and source process of the Boumerdes, Algeria, earthquake inferred from geodetic and strong motion data. *Geophys. Res. Lett.* **32**, L01305 (2005).
81. Wei, S. Source models of large earthquakes: July/15/2003, Carlsberg Ridge, Mw7.6. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2004_carlsberg-ridge/index.html (accessed 1 July 2013).
82. Koketsu, K., Hikima, K., Miyazaki, S. & Ide, S. Joint inversion of strong motion and geodetic data for the source process of the 2003 Tokachi-oki, Hokkaido, earthquake. *Earth Planets Space* **56**, 329–334 (2004).
83. Tanioka, Y., Hirata, K., Hino, R. & Kanazawa, T. Slip distribution of the 2003 Tokachi-oki earthquake estimated from tsunami waveform inversion. *Earth Planets Space* **56**, 373–376 (2004).
84. Yagi, Y. Source rupture process of the 2003 Tokachi-oki earthquake determined by joint inversion of teleseismic body wave and strong ground motion data. *Earth Planets Space* **56**, 311–316 (2004).
85. Yamanaka, Y. & Kikuchi, M. Source process of the recurrent Tokachi-oki earthquake on September 26, 2003, inferred from teleseismic body waves. *Earth Planets Space* **55**, e21–e24 (2003).
86. Wei, S. Source models of large earthquakes: Feb/07/2004 (Mw 7.2), Irian Jaya, Indonesia. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2004_indo-irian_jaya/index.html (accessed 1 July 2013).
87. Ammon, C. J. et al. Rupture process of the great 2004 Sumatra-Andaman earthquake. *Science* **308**, 1133–1139 (2005).
88. Rhie, J., Dreger, D., Burgmann, R. & Romanowicz, B. Slip of the 2004 Sumatra-Andaman earthquake from joint inversion of long-period global seismic waveforms and GPS static offsets. *Bull. Seismol. Soc. Am.* **97**, S115–S127 (2007).
89. Shao, G. & Ji, C. Preliminary result of the Aug 16, 2005 Mw 7.19 Honshu earthquake. *UCSB* http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2005/08/smooth/honshu.html (accessed 22 August 2013).
90. Shao, G. & Ji, C. Preliminary result of the Jun 15, 2005 Mw 7.2 northern California earthquake. *UCSB* http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2005/06/smooth/northernca.html (accessed 25 September 2013).
91. Lay, T. et al. The 2006-2007 Kuril Islands great earthquake sequence. *J. Geophys. Res.* **114**, B11308 (2009).
92. Sladen, A. Source models of large earthquakes: preliminary result, 11/15/2006 (Mw 8.3), Kuril Islands. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2006_kuril/kuril.html (accessed 1 July 2013).
93. Yen, Y.-T., Ma, K.-F. & Wen, Y.-Y. Slip partition of the 26 December 2006 Pingtung, Taiwan (*M*6.9, *M*6.8) earthquake doublet determined from teleseismic waveforms. *Diqiu Kexue Jikan* **19**, 567–578 (2008).
94. Ji, C. Rupture process of the 2007 Jan 13 magnitude 8.1 - KURIL Island earthquake (revised). *UCSB* http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2007/01/13/kuril.html (accessed 22 August 2013).
95. Sladen, A. Source models of large earthquakes: preliminary result, 01/13/2007 (Mw 8.1), Kuril Islands. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2007_kuril/kuril.html (accessed 1 July 2013).
96. Ji, C. & Zeng, Y. Preliminary result of the Sep 12, 2007 Mw 7.9 Kepulauan earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2007PAGAII01JIxx (accessed 14 June 2018).
97. Sladen, A. & Ozgun Konca, A. Source models of large earthquakes: preliminary result, 09/12/2007 (Mw 7.9), Central Sumatra earthquake. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2007_c_sumatra/c-sumatra.html (accessed 1 July 2013).
98. Ji, C. Rupture process of the 2007 April 1, magnitude 8.1, Solomon Islands earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2007SOLOMO01JIxx (accessed 14 June 2018).
99. Béjar-Pizzaro, M. et al. Asperities and barriers on the seismogenic zone in North Chile: state-of-the-art after the 2007 M_w 7.7 Tocopilla earthquake inferred by GPS and InSAR data. *Geophys. J. Int.* **183**, 390–406 (2010).
100. Sladen, A. Source models of large earthquakes: preliminary result, 11/14/2007 (Mw 7.7), Tocopilla earthquake, Chile. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2007_tocopilla/tocopilla.html (accessed 1 July 2013).
101. Zeng, Y., Hayes, G. & Ji, C. Preliminary result of the Nov 14, 2007 Mw 7.7 Antofagasto, Chile earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2007TOCOPI01ZENG (accessed 14 June 2018).
102. Sladen, A. Source models of large earthquakes: preliminary result, 11/16/2008 (Mw 7.3), Sulawesi. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2008_sulawesi/sulawesi.html (accessed 1 July 2013).
103. Ji, C. & Hayes, G. Preliminary result of the May 12, 2008 Mw 7.9 eastern Sichuan, China earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2008WENCHU01JIxx (accessed 18 June 2018).
104. Sladen, A. Source models of large earthquakes: preliminary result, 05/12/2008 (Mw 7.9), East Sichuan. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2008_e_sichuan/e_sichuan.html (accessed 1 July 2013).
105. Yagi, Y., Nishimura, N. & Kasahara, A. Source process of the 12 May 2008 Wenchuan, China, earthquake determined by waveform inversion of teleseismic body waves with a data covariance matrix. *Earth Planets Space* **64**, e13–e16 (2012).
106. Fielding, E. J. et al. Kinematic fault slip evolution source models of the 2008 *M*7.9 Wenchuan earthquake in China from SAR interferometry, GPS and teleseismic analysis and implications for Longmen Shan tectonics. *Geophys. J. Int.* **194**, 1138–1166 (2013).
107. Hayes, G. Preliminary result of the July 15, 2009 Mw 7.6 Fiordland earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2009FIORDL01HAYE (accessed 14 June 2018).
108. Hayes, G. Preliminary result of the August 3, 2009 Mw 6.9 Gulf of California earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2009GULFOF01HAYE (accessed 14 June 2018).
109. Cirella, A., Piatanesi, A., Tinti, E., Chini, M. & Cocco, M. Complexity of the rupture process during the 2009 L'Aquila, Italy, earthquake. *Geophys. J. Int.* **190**, 607–621 (2012).
110. Gualandi, A., Serpelloni, E. & Belardinelli, M. E. s2009LAQUIL01GUAL. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2009LAQUIL01GUAL (accessed June 2018).
111. Hayes, G. & Ji, C. Preliminary result of the May 28, 2009 Mw 7.3 earthquake offshore Honduras. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2009OFFSHO01HAYE (accessed 14 June 2018).
112. Hayes, G. A preliminary result of the Sep 30, 2009 Mw 7.6 southern Sumatra earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2009PADANG01HAYE (accessed 19 June 2018).
113. Sladen, A. Source models of large earthquakes: preliminary result, 09/30/2009 (Mw 7.6), Padang, Indonesia. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2009_padang/padang.html (accessed 1 July 2013).
114. Hayes, G. Preliminary result of the Sep 29, 2009 Mw 8.0 Samoa earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2009SAMOAx01HAYE (accessed 19 June 2018).
115. Sladen, A. Source models of large earthquakes: preliminary result, 10/07/2009 (Mw 7.6), Vanuatu. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2009_vanuatu/index.html (accessed 1 July 2013).
116. Wei, S. et al. Superficial simplicity of the 2010 El Mayor–Cucapah earthquake of Baja California in Mexico. *Nat. Geosci.* **4**, 615–618 (2011).
117. Calais, E. et al. Transpressional rupture of an unmapped fault during the 2010 Haiti earthquake. *Nat. Geosci.* **3**, 794–799 (2010).
118. Hayes, G. P. et al. Complex rupture during the 12 January 2010 Haiti earthquake. *Nat. Geosci.* **3**, 800–805 (2010).
119. Sladen, A. Source models of large earthquakes: preliminary result, 01/12/2010 (Mw 7.0), Haiti. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2010_haiti/index.html (accessed 1 July 2013).
120. Hayes, G. Updated result of the Jan 12, 2010 Mw 7.0 Haiti earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2010HAITIx02HAYE (accessed 20 June 2018).
121. Hayes, G. Updated result of the Feb 27, 2010 Mw 8.8 Maule, Chile earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2010MAULEC01HAYE (accessed 19 June 2018).
122. Hayes, G. Updated result of the Apr 6, 2010 northern Sumatra earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2010NORTHE01HAYE (accessed 19 June 2018).
123. Hayes, G. Preliminary result of the Dec 25, 2010 Mw 7.3 Vanuatu region earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2010VANUAT01HAYE (accessed 19 June 2018).
124. Hayes, G. Preliminary result of the Oct 21, 2011 Mw 7.4 Kermadec Islands region earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2011KERMAD02HAYE (accessed 3 June 2018).
125. Hayes, G. Preliminary result of the July 6, 2011 Mw 7.6 Kermadec Islands region earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2011KERMAD01HAYE (accessed 1 June 2018).
126. Hayes, G. Updated result of the Mar 9, 2011 Mw 7.3 earthquake offshore Honshu, Japan (Tohoku EQ foreshock). *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2011OFFSHO01HAYE (accessed 1 June 2018).

127. Fujii, Y., Satake, K., Sakai, S., Shinohara, M. & Kanazawa, T. Tsunami source of the 2011 off the Pacific coast of Tohoku earthquake. *Earth Planets Space* **63**, 815–820 (2011).
128. Satake, K., Fujii, Y., Harada, T. & Namegaya, Y. Time and space distribution of coseismic slip of the 2011 Tohoku earthquake as inferred from tsunami waveform data. *Bull. Seismol. Soc. Am.* **103**, 1473–1492 (2013).
129. Yue, H. & Lay, T. Source rupture models for the M_w 9.0 2011 Tohoku earthquake from joint inversions of high-rate geodetic and seismic data. *Bull. Seismol. Soc. Am.* **103**, 1242–1255 (2013).
130. Hayes, G. Updated result of the Oct 23, 2011 Mw 7.1 eastern Turkey earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2011VANTUR01HAYE (accessed 20 June 2018).
131. Shao, G. & Ji, C. Preliminary result of the Oct 23, 2011 Mw 7.13 Turkey earthquake. *UCSB* http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2011/10/23/turkey.html (accessed 22 August 2013).
132. Hayes, G. Preliminary result of the Aug 20, 2011 Mw 7.1 Vanuatu earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2011VANUAT01HAYE (accessed 1 June 2018).
133. Wei, S. et al. Complementary slip distributions of the largest earthquakes in the 2012 Brawley swarm, Imperial Valley, California. *Geophys. Res. Lett.* **40**, 847–852 (2013).
134. Hayes, G. Preliminary result of the Aug 31, 2012 Mw 7.6 earthquake east of Sulangan, Philippines. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2012EASTOF01HAYE (accessed 21 June 2018).
135. Lay, T. et al. The October 28, 2012 M_w 7.8 Haida Gwaii underthrusting earthquake and tsunami: slip partitioning along the Queen Charlotte fault transpressional plate boundary. *Earth Planet. Sci. Lett.* **375**, 57–70 (2013).
136. Shao, G. & Ji, C. Preliminary result of the Oct 28, 2012 Mw 7.72 Canada earthquake. *UCSB* http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2012/10/canada.html (accessed 20 August 2013).
137. Wei, S. Source models of large earthquakes: Oct./28/2012 (Mw 7.8), Masset, Canada. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2012_Masset/index.html (accessed 1 July 2013).
138. Hayes, G. Preliminary result of the Mar 20, 2012 Mw 7.4 Oaxaca, Mexico earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2012OAXACA01HAYE (accessed 20 June 2018).
139. Wei, S. Source models of large earthquakes: March/20/2012 (Mw 7.4), OAXACA, Mexico. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2012_Mexico/index.html (accessed 1 July 2013).
140. Hayes, G. Preliminary result of the Jan 10, 2012 Mw 7.2 off the west coast of northern Sumatra, Indonesia earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2012SUMATR03HAYE (accessed 14 June 2018).
141. Shao, G., Li, X. & Ji, C. Preliminary result of the Apr 11, 2012 Mw 8.64 Sumatra earthquake. *UCSB* http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2012/04/10/sumatra.html (accessed 19 August 2013).
142. Hayes, G. Preliminary result of the Apr 11, 2012 Mw 8.6 earthquake off the west coast of northern Sumatra. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2012SUMATR01HAYE (accessed 19 June 2018).
143. Hayes, G. Preliminary result of the Apr 11, 2012 Mw 8.6 earthquake off the west coast of northern Sumatra. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2012SUMATR02HAYE (accessed 19 June 2018).
144. Yamanaka, Y. & Kikuchi, M. Asperity map along the subduction zone in northeastern Japan inferred from regional seismic data. *J. Geophys. Res.* **109**, B07307 (2004).
145. Hartzell, S. & Mendoza, C. Application of an iterative least-squares wave-form inversion of strong-motion and teleseismic records to the 1978 Tabas, Iran, earthquake. *Bull. Seismol. Soc. Am.* **81**, 305–331 (1991).
146. Archuleta, R. J. A faulting model for the 1979 Imperial Valley earthquake. *J. Geophys. Res.* **89**, 4559–4585 (1984).
147. Hartzell, S. H. & Heaton, T. H. Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake. *Bull. Seismol. Soc. Am.* **73**, 1553–1583 (1983).
148. Olson, A. H. & Apsel, R. J. Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake. *Bull. Seismol. Soc. Am.* **72**, 1969–2001 (1982).
149. Mendoza, C. & Hartzell, S. H. Slip distribution of the 19 September 1985 Michoacan, Mexico, earthquake: near-source and teleseismic constraints. *Bull. Seismol. Soc. Am.* **79**, 655–669 (1989).
150. Wald, D. J. Strong motion and broad-band teleseismic analysis of the 1991 Sierra-Madre, California, earthquake. *J. Geophys. Res.* **97**, 11033–11046 (1992).
151. Silva, W. et al. A slip model for the Little Skull Mountain earthquake of June 29, 1992. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s1992LITTLE01SILV (accessed June 2018).
152. Cho, I. & Nakanishi, I. Investigation of the three-dimensional fault geometry ruptured by the 1995 Hyogo-Ken Nanbu earthquake using strong-motion and geodetic data. *Bull. Seismol. Soc. Am.* **90**, 450–467 (2000).
153. Horikawa, H., Hirahara, K., Umeda, Y., Hashimoto, M. & Kusano, F. Simultaneous inversion of geodetic and strong-motion data for the source process of the Hyogo-ken Nanbu, Japan, earthquake. *J. Phys. Earth* **44**, 455–471 (1996).
154. Ide, S., Takeo, M. & Yoshida, Y. Source process of the 1995 Kobe earthquake: determination of spatio-temporal slip distribution by Bayesian modeling. *Bull. Seismol. Soc. Am.* **86**, 547–566 (1996).
155. Koketsu, K., Yoshida, S. & Higashihara, H. A fault model of the 1995 Kobe earthquake derived from the GPS data on the Akashi Kaikyo Bridge and other datasets. *Earth Planets Space* **50**, 803–811 (1998).
156. Sekiguchi, H., Irikura, K. & Iwata, T. Fault geometry at the rupture termination of the 1995 Hyogo-ken Nanbu earthquake. *Bull. Seismol. Soc. Am.* **90**, 117–133 (2000).
157. Wald, D. J. Slip history of the 1995 Kobe, Japan, earthquake determined from strong motion, teleseismic, and geodetic data. *J. Phys. Earth* **44**, 489–503 (1996).
158. Sekiguchi, H., Irikura, K. & Iwata, T. Source inversion for estimating the continuous slip distribution on a fault introduction of Green's functions convolved with a correction function to give moving dislocation effects in subfaults. *Geophys. J. Int.* **150**, 377–391 (2002).
159. Chi, W. C., Dreger, D. & Kaverina, A. Finite-source modeling of the 1999 Taiwan (Chi-Chi) earthquake derived from a dense strong-motion network. *Bull. Seismol. Soc. Am.* **91**, 1144–1157 (2004).
160. Jonsson, S., Zebker, H., Segall, P. & Amelung, F. Fault slip distribution of the 1999 M_w 7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurements. *Bull. Seismol. Soc. Am.* **92**, 1377–1389 (2002).
161. Zhang, W., Iwata, T., Irikura, K., Pitarka, A. & Sekiguchi, H. Dynamic rupture process of the 1999 Chi-Chi, Taiwan, earthquake. *Geophys. Res. Lett.* **31**, L10605 (2004).
162. Wu, C. J., Takeo, M. & Ide, S. Source process of the Chi-Chi earthquake: a joint inversion of strong motion data and global positioning system data with a multifault model. *Bull. Seismol. Soc. Am.* **91**, 1128–1143 (2004).
163. Zeng, Y. H. & Chen, C. H. Fault rupture process of the 20 September 1999 Chi-Chi, Taiwan, earthquake. *Bull. Seismol. Soc. Am.* **91**, 1088–1098 (2004).
164. Ma, K. F., Mori, J., Lee, S. J. & Yu, S. B. Spatial and temporal distribution of slip for the 1999 Chi-Chi, Taiwan, earthquake. *Bull. Seismol. Soc. Am.* **91**, 1069–1087 (2004).
165. Birgören, G., Sekiguchi, H. & Irikura, K. Rupture model of the 1999 Duzce, Turkey, earthquake deduced from high and low frequency strong motion data. *Geophys. Res. Lett.* **31**, L05610 (2004).
166. Delouis, B., Lundgren, P. & Giardini, D. Slip distributions of the 1999 Décze (Mw 7.2) and Izmit (Mw 7.6) earthquakes on the North Anatolian Fault (Turkey): a combined analysis, internal report. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s1999DUZCET01DELO (accessed 1 April 2018).
167. Hernandez, B. et al. Rupture history of September 30, 1999 intraplate earthquake of Oaxaca, Mexico (M_w 7.5) from inversion of strong-motion data. *Geophys. Res. Lett.* **28**, 363–366 (2001).
168. Iwata, T., Sekiguchi, H., Matsumoto, Y., Miyake, H. & Irikura, K. Source process of the 2000 western Tottori Prefecture earthquake and near-source strong ground motion. In *2000 Fall Meeting of the Seismological Society of Japan* (Seismological Society of Japan, 2000).
169. Sekiguchi, H., Iwata, T., Sugiyama, Y., Fusejima, Y. & Horikawa, H. Faulting process and condition for its occurrence of 2000 Tottori-ken Seibu Earthquake. In *2001 Japan Earth and Planetary Science Joint Meeting abstr.* S3-006 (2001).
170. Kakehi, Y. Analysis of the 2001 Geiyo, Japan, earthquake using high-density strong ground motion data: detailed rupture process of a slab earthquake in a medium with a large velocity contrast. *J. Geophys. Res.* **109**, B08306 (2004).
171. Yagi, Y., Mikurno, T., Pacheco, J. & Reyes, G. Source rupture process of the Tecoman, Colima, Mexico earthquake of 22 January 2003, determined by joint inversion of teleseismic body-wave and near-source data. *Bull. Seismol. Soc. Am.* **94**, 1795–1807 (2004).
172. Custódio, S., Liu, P. C. & Archuleta, R. J. The 2004 M_w 6.0 Parkfield, California, earthquake: inversion of near-source ground motion using multiple data sets. *Geophys. Res. Lett.* **32**, L23312 (2005).
173. Dreger, D. S., Gee, L., Lombard, P., Murray, M. H. & Romanowicz, B. Rapid finite-source analysis and near-fault strong ground motions: application to the 2003 M_w 6.5 San Simeon and 2004 M_w 6.0 Parkfield earthquakes. *Seismol. Res. Lett.* **76**, 40–48 (2005).
174. Ji, C. Source models of large earthquakes: slip history the 2004 (Mw 5.9) Parkfield earthquake (single-plane model). *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2004_ca/parkfield2.html (accessed 1 July 2013).
175. Ozgun Konca, A. Source models of large earthquakes: preliminary result, 06/10/08 (Mw 7.6), Kashmir earthquake. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2005_kashmir/kashmir.html (accessed 1 July 2013).
176. Yagi, Y. & Fukahata, Y. Rupture process of the 2011 Tohoku-oki earthquake and absolute elastic strain release. *Geophys. Res. Lett.* **38**, L19307 (2011).
177. Ji, C. Preliminary result of the 2006 July 17 magnitude 7.7 - south of Java, Indonesia earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2006SOUTHE01JIxx (accessed 21 June 2018).
178. Ozgun Konca, A. Source models of large earthquakes: preliminary result, 06/07/17 (Mw 7.9), southern Java earthquake. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2006_s_java/s_java.html (accessed 1 July 2013).

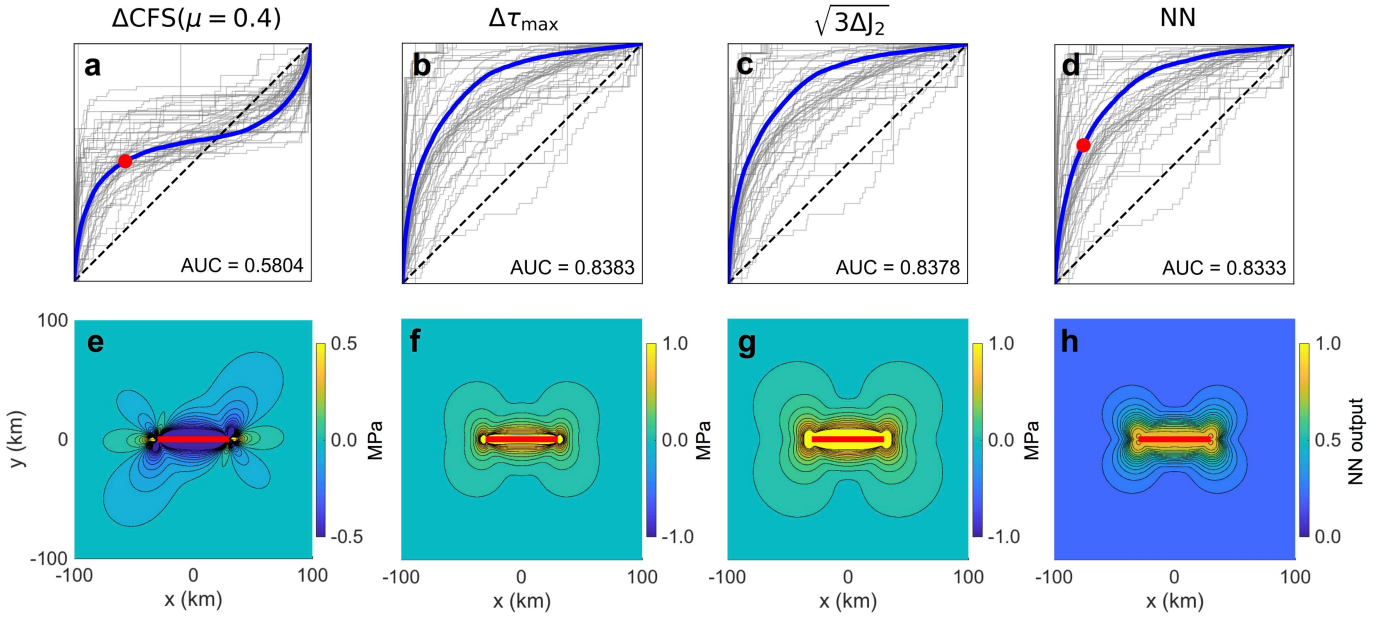
179. Ji, C. & Zeng, Y. Preliminary result of the Aug 15, 2007 Mw 8.0 coast of central Peru earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2007PISCOP01JIxx (accessed 21 June 2018).
180. Ozgun Konca, A. Source models of large earthquakes: preliminary result, 07/08/15 (Mw 8.0), Peru earthquake. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2007_peru/pisco.html (accessed 1 July 2013).
181. Hayes, G. & Ji, C. Preliminary result of the Jun 13, 2008 Mw 6.8 Honshu earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2008IWATEx01HAYE (accessed 28 June 2018).
182. Hayes, G. Preliminary result of the Sep 29, 2008 Mw 7.0 Kermedac Islands earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2008KERMED01HAYE (accessed 19 June 2013).
183. Hayes, G. & Ji, C. Preliminary result of the Feb 20, 2008 Mw 7.4 Simeulue earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2008SIMEUL01HAYE (accessed 20 June 2018).
184. Sladen, A. Source models of large earthquakes: preliminary result 02/20/2008 (Mw 7.4), Simeulue earthquake, Indonesia. *Caltech Tectonics Observatory* http://www.tectonics.caltech.edu/slip_history/2008_n_sumatra/simeulue.html (accessed 1 July 2013).
185. Hayes, G. Preliminary result of the Jan 3, 2009 Mw 7.6 Papua earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2009PAPUAx01HAYE (accessed 19 June 2018).
186. Hayes, G. Preliminary result of the Dec 21, 2010 Mw 7.4 Bonin Islands earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2010BONINI01HAYE (accessed 19 June 2018).
187. Hayes, G. Preliminary result of the May 9, 2010 Mw 7.2 northern Sumatra earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2010NORTHE02HAYE (accessed 19 June 2018).
188. Hayes, G. Preliminary result of the Sep 5, 2012 Mw 7.6 Costa Rica earthquake. *eQuake-RC Finite-Source Rupture Model Database* http://equake-rc.info/SRCMOD/searchmodels/viewmodel/s2012COSTAR01HAYE (accessed 21 June 2018).
189. Yue, H. et al. The 5 September 2012 Nicoya, Costa Rica M_w 7.6 earthquake rupture process from joint inversion of high-rate GPS, strong-motion, and teleseismic P wave data and its relationship to adjacent plate boundary interface properties. *J. Geophys. Res.* **118**, 5453–5466 (2013).



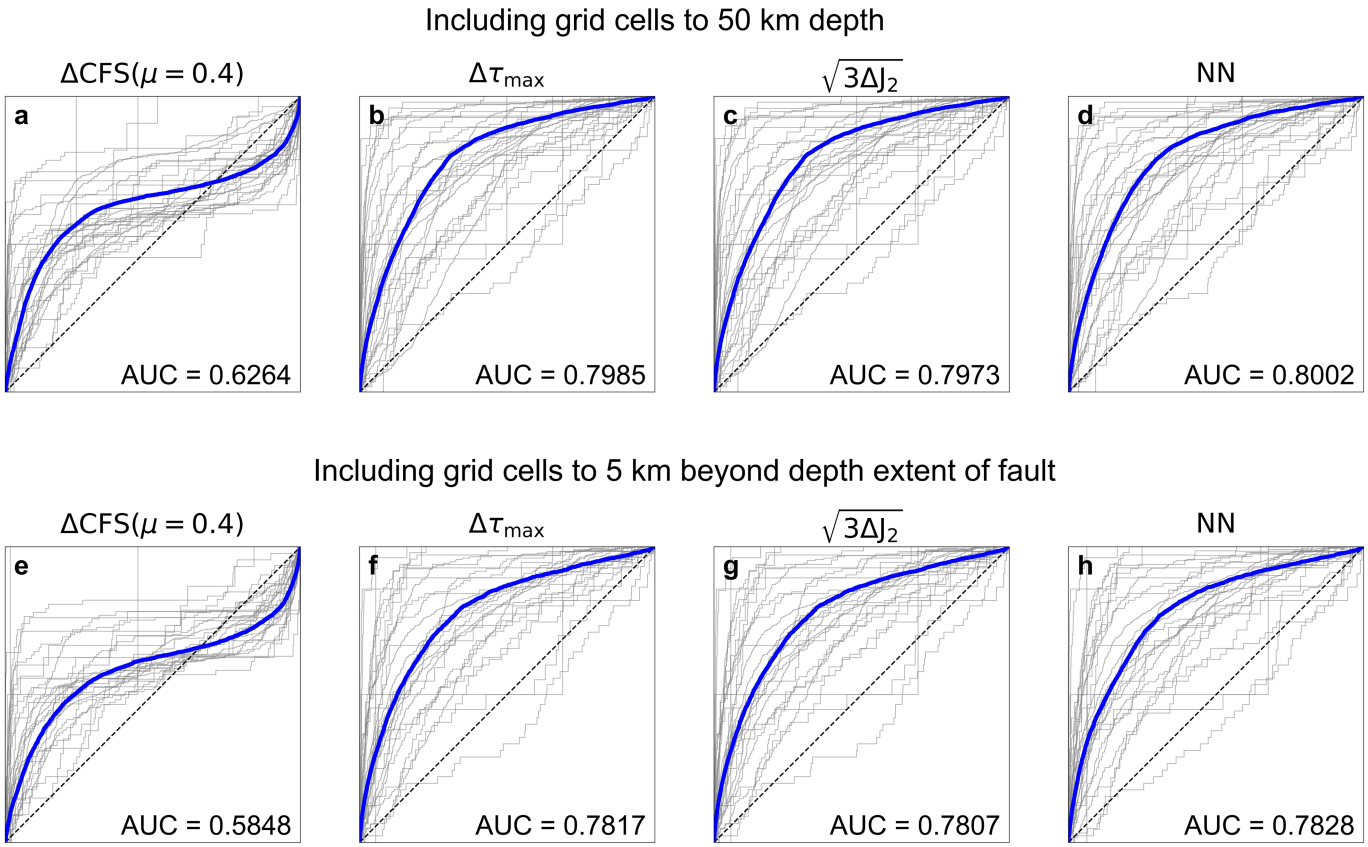
Extended Data Fig. 1 | Comparisons of spatial patterns of stress metrics. a–d, Analogous to Fig. 2e–h, but for an idealized thrust earthquake. The fault plane dips 45° to the north and the red line is the trace of the fault at the surface. Depth shown is 10 km.



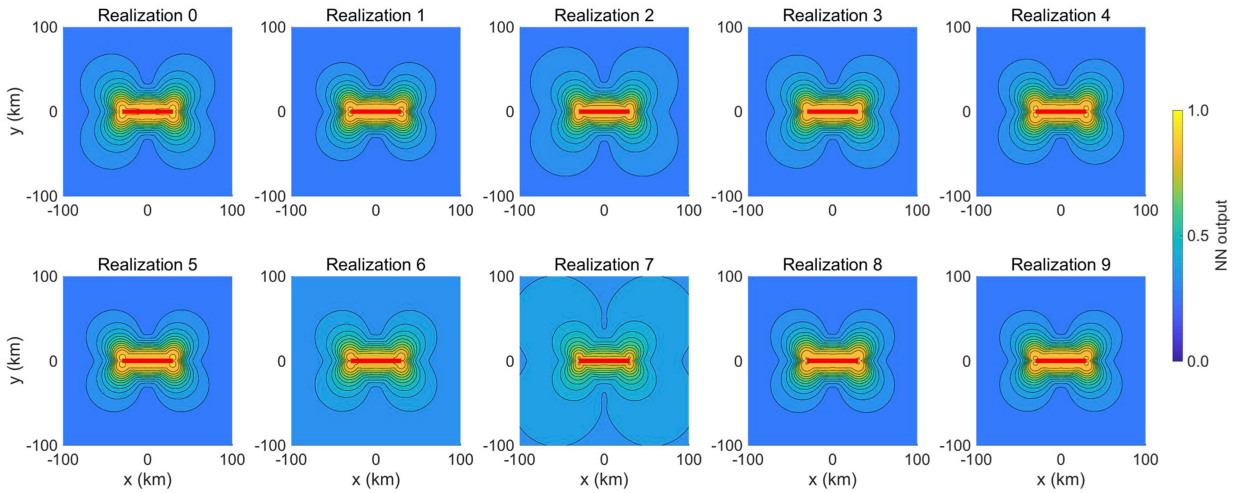
Extended Data Fig. 2 | Mainshock–aftershock examples. a–h, Analogous to Fig. 1a–h, using the same sign conventions for Coulomb failure stress change, but with results based on a training dataset (Supplementary Table 1) that excludes grid cells more than 5 km below the maximum depth of each slip distribution.



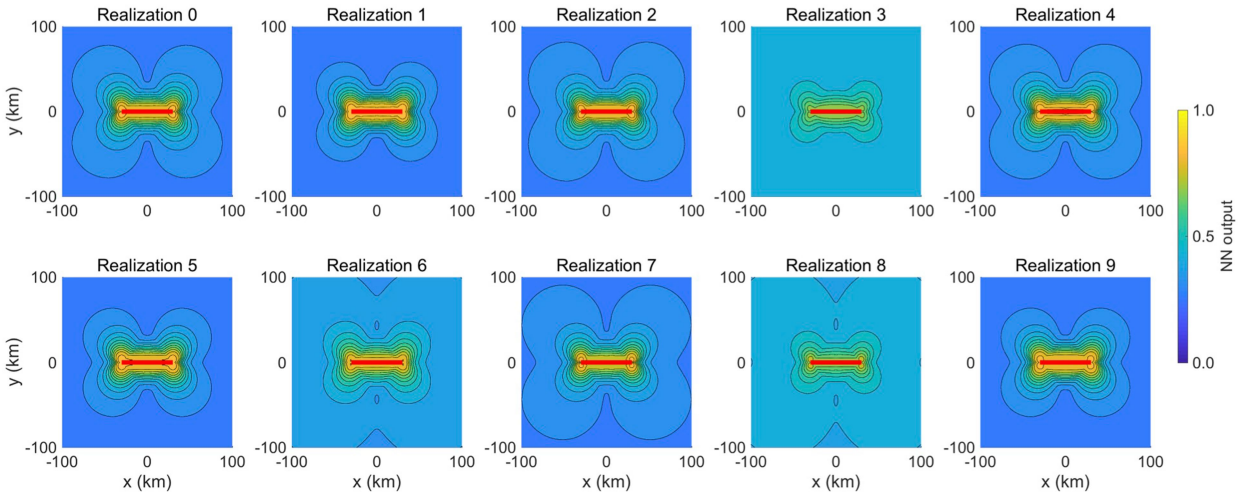
Extended Data Fig. 3 | Comparisons of performance. a–h, Analogous to Fig. 2a–h, using training and test datasets (Supplementary Table 1) that exclude grid cells more than 5 km below the maximum depth of each slip distribution.



Extended Data Fig. 4 | ROC curves associated with realization 6 of the datasets. a–d, Curves incorporate grid cells down to a depth of 50 km. **e–h**, Curves including grid cells down to 5 km beyond the maximum depth of each slip distribution. Thus, the neural network in **d** is trained and evaluated on a version of dataset realization 6 (Supplementary Table 2) that incorporates grid cells down to a depth of 50 km, whereas that in **h** is trained and evaluated on the same realizations of slip distributions, but incorporating only grid cells down to 5 km below each slip distribution.



Extended Data Fig. 5 | Forward predictions of the neural networks from each realization of the training dataset, incorporating all grid cells down to 50 km. Each panel is analogous to Fig. 2h, but uses one of ten distinct neural networks trained on one of ten different realizations of the training dataset (Supplementary Table 2). See Methods for further discussion.



Extended Data Fig. 6 | Forward predictions of the neural networks from each realization of the training dataset, incorporating grid cells down to 5 km beyond the depth of each slip distribution. Each panel is analogous to Fig. 2h, but uses one of ten distinct neural networks trained on one of ten different realizations of the training dataset (Supplementary Table 2). See Methods for further discussion.

Extended Data Table 1 | Comparison of physical metrics to the neural network for an idealized case

Quantity	Symbol	Evaluation	%VE
nearest distance	r	$r = \min\left(\sqrt{(x - x_f)^2 + (y - y_f)^2}\right)$	46%
maximum shear	$\Delta\tau_{max}(\boldsymbol{\chi})$	$\Delta\tau_{max}(\boldsymbol{\chi}) = \lvert\chi_1 - \chi_3\rvert/2$	98%
1st invariant	$\Delta I_1(\boldsymbol{\chi})$	$\Delta I_1(\boldsymbol{\chi}) = \chi_1 + \chi_2 + \chi_3$	66%
2nd invariant	$\Delta I_2(\boldsymbol{\chi})$	$\Delta I_2(\boldsymbol{\chi}) = \chi_1\chi_2 + \chi_2\chi_3 + \chi_1\chi_3$	96%
von-Mises criteria	$\sqrt{3\Delta J_2}$	$\sqrt{3\Delta J_2} = \sqrt{\Delta I_1^2(\boldsymbol{\sigma}) - 3\Delta I_2(\boldsymbol{\sigma})}$	98%
3rd invariant	$\Delta I_3(\boldsymbol{\chi})$	$\Delta I_3(\boldsymbol{\chi}) = \chi_1\chi_2\chi_3$	84%
Coulomb failure criteria μ = 0.0, 0.2, 0.4, 0.6, 0.8	$\Delta CFS(\boldsymbol{\chi}, \mu)$	$\Delta CFS(\boldsymbol{\chi}, \mu) = (\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot \mathbf{n}_\parallel - \mu(\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot \mathbf{n}_\perp$	89%
Coulomb failure criteria normal only, μ = 0.4	$\Delta CFS_n(\boldsymbol{\chi})$	$\Delta CFS_n(\boldsymbol{\chi}) = -\mu(\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot \mathbf{n}_\perp$	59%
Coulomb failure criteria total shear	$\Delta CFS_\tau(\boldsymbol{\chi})$	$\Delta CFS_\tau(\boldsymbol{\chi}) = \lvert(\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot \mathbf{n}_\parallel\rvert + \lvert(\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot (\mathbf{n}_\parallel \times \mathbf{n}_\perp)\rvert$	95%
Coulomb failure criteria total, μ = 0.4	$\Delta CFS_{total}(\boldsymbol{\chi}, \mu)$	$\Delta CFS_{total}(\boldsymbol{\chi}, \mu) = \lvert(\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot \mathbf{n}_\parallel\rvert + \lvert(\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot (\mathbf{n}_\parallel \times \mathbf{n}_\perp)\rvert - \mu(\mathbf{n}_\perp \cdot \boldsymbol{\chi}) \cdot \mathbf{n}_\perp$	93%
Sum of magnitudes of stress components	$m(\Delta\boldsymbol{\chi})$	$m(\Delta\boldsymbol{\chi}) = \lvert\Delta\chi_{xx}\rvert + \lvert\Delta\chi_{yy}\rvert + \lvert\Delta\chi_{zz}\rvert + \lvert\Delta\chi_{xy}\rvert + \lvert\Delta\chi_{xz}\rvert + \lvert\Delta\chi_{yz}\rvert$	>99%

χ represents either the full (σ) or deviatoric (σ') stress-change tensor, χ_i are the corresponding eigenvalues, x_f and y_f are the x and y locations of the fault plane, respectively, and $\boldsymbol{n}_\perp$ and $\boldsymbol{n}_\parallel$ are the unit vectors perpendicular to the average orientation of the mainshock fault plane and parallel to the mean mainshock slip direction, respectively. %VE is the proportion of the variance in the strike-averaged neural-network forecast for the idealized strike-slip case (Fig. 2) that is explained by each strike-averaged physical metric. We include the largest %VE for each metric. For Coulomb failure stress change, the largest %VE corresponds to the magnitude of the Coulomb failure stress change associated with the full stress-change tensor $|\Delta CFS(\sigma, \mu = 0.0)|$. See Methods for details.

Extended Data Table 2 | Summary of results for ten realizations of the training and test datasets

Grid cells down to 50 km depth

	$AUC_{\Delta CFS(\mu=0.4)}$	$AUC_{\Delta\tau_{max}}$	$AUC_{von\text{-}Mises}$	AUC_{NN}
Realization 0	0.5810	0.8147	0.8137	0.8185
Realization 1	0.6037	0.8104	0.8092	0.8127
Realization 2	0.6031	0.8067	0.8056	0.8123
Realization 3	0.6056	0.8174	0.8164	0.8224
Realization 4	0.5818	0.8183	0.8173	0.8200
Realization 5	0.5951	0.8083	0.8073	0.8128
Realization 6	0.6264	0.7985	0.7973	0.8002
Realization 7	0.5774	0.8202	0.8192	0.8242
Realization 8	0.6025	0.8214	0.8205	0.8258
Realization 9	0.5748	0.8161	0.8151	0.8186

Grid cells down to 5 km deeper than the depth of each slip distribution

	$AUC_{\Delta CFS(\mu=0.4)}$	$AUC_{\Delta\tau_{max}}$	$AUC_{von\text{-}Mises}$	AUC_{NN}
Realization 0	0.5759	0.8016	0.8009	0.7982
Realization 1	0.5822	0.7937	0.7927	0.7946
Realization 2	0.5825	0.7946	0.7938	0.7964
Realization 3	0.6053	0.8073	0.8067	0.8095
Realization 4	0.5775	0.8054	0.8047	0.8066
Realization 5	0.5757	0.7937	0.7930	0.7941
Realization 6	0.5848	0.7817	0.7807	0.7828
Realization 7	0.5759	0.8087	0.8079	0.8096
Realization 8	0.6047	0.8110	0.8104	0.8133
Realization 9	0.5721	0.8018	0.8010	0.8000

The top half of the table displays results based on realizations that incorporate grid cells down to 50 km and the bottom half displays results based on realizations that incorporate grid cells down to 5 km below the maximum depth of each slip distribution (see Supplementary Table 2).