Supplementary Information: Empirical Relationships Between Ground Motion Intensities and JMA Seismic Intensity Scale

Patrick W. Ball *[1], **Toby Sansom**[1], **Karina Loviknes**[2]

[1]Inigo Insurance Ltd, 25 Fenchurch Avenue, London, United Kingdom, EC3M 5AD, [2]Njord Centre, Department of Physics and Geosciences, University of Oslo, Oslo, Norway

S1 Supplementary Table 1

Supplementary Table 1 contains 21,354 strong motion recordings across 112 Japanese earthquakes from the National Research Institute for Earth Science and Disaster Resilience maintain two strong-motion seismographic networks: KiK-Net (Kiban-Kyoshin network) and K-NET (Kyoshin network; National Research Institute for Earth Science and Disaster Resilience, 2019). These earthquakes were recorded between 1/1/1997 and 31/12/2025, are $\geq$ M_w 5.0, and have a recording of ≥ 5.0 I_{JMA} for at least one station. We only use K-Net observations from the surface seismometer and all recordings are filtered, processed, and acceleration metrics are calculated by Loviknes et al. (2026a,b). The following columns are included in the table:

- Origin_Time - the date and time the earthquake occurred.
- EQ_Longitude, EQ_Latitude, EQ_Depth_km - location of earthquake hypocenter.
- Magnitude - earthquake moment magnitude.
- Network - network of recording station (kik or knet).
- Station_Code - station identifying code.
- Station_Longitude, Station_Latitude, Station_Height_m - Location of recording station.
- Record_Time - the date and time the recording was taken.
- energy_ratioSignal, energy_ratioNoise, freq_range - parameters assessing accelerogram processing quality (Loviknes et al., 2026a).
- VS30 - shear wave velocity within the upper 30m of the subsurface (m/s) (Zhu et al., 2021).
- Rhypo - hypocentral distance (km), calculated as the shortest distance between hypocenter and station(Loviknes et al., 2026a).
- Max_Acc_gal, Max_h_Acc_gal - maximum peak ground acceleration recorded as the resultant in 3-component and horizontal directions.
- PGA_GM, PGA_RotD50 - Geometric mean (Geom) and RotD50 (Rot50) horizontal peak ground acceleration.
- Sa_GM_X, Sa_RotD50_X - Geometric and RotD50 mean horizontal spectral acceleration (Sa) with 5% damping where X is 0.2, 0.3, 0.6, 1.0, 2.0 or 3.0 s.
- Shindo_Intensity - Japanese Meteorological Agency seismic intensity value.
- Shindo Class - Conversion from Shindo Intensity using Table 1 in the main text.

This table was constructed using Strong_Motion_Data_Processing.ipynb which leverages the PyRotD and shindo.py python libraries (Kottke et al., 2023; RR-Inyo, 2021). All PGA and S_a values are reported in gal.

*Corresponding author: patrick.ball@inigoinsurance.com

S2 Supplementary Table 2

Supplementary Table 2 contains the best-fitting regression model coefficients for all possible combinations of the input parameters (see Regression Modelling, Methodology in main text). Column headers include the variable associated with each coefficient. These are:

- intercept - the Y-axis intercept.
- M_w - moment magnitude of the earthquake.
- log10_D_{hyp} - the log of the distance between the hypocenter and the recording station.
- log10_PGA - the log of peak ground acceleration.
- log10_Sa_X - the log of spectral acceleration at X period, where X is 0.2, 0.3, 0.6, 1.0, 2.0, or 3.0 s.

Hypocentral distances and strong motion observations are presented in km and gal, respectively. All strong motion parameters can be presented as the geometric means of their horizontal components (GM) or as the median response when rotated over all horizontal orientations (RotD50). Metadata provided includes:

- Orientation - the orientation of strong motion parameters, GM or RotD50.
- Degrees_of_Freedom - the number of regression parameters considered.
- lamda - the regularisation parameter.
- wRSME - weighted root-mean-squared error.
- BIC - Bayesian Information Criterion.
- Effective_Degrees_of_Freedom - effective degrees of freedom.

To aid model comparison, we calculate each parameterisation's Bayesian Information Criterion (BIC) to evaluate their respective performance whilst penalising unnecessary model complexity (Neath and Cavanaugh, 2012). In a regression analysis, the goodness of fit will always improve when an additional variable is introduced. Since Ridge regression regularises coefficients toward zero, the model is less complex than the number of variables implies. The BIC equation is therefore adapted to use the effective degrees of freedom, d_{eff}, rather than the raw parameter count,

$$\text{BIC} = n \ln\left(\frac{\sum w_i (y_i - \hat{y}_i)^2}{n}\right) + d_{eff} \log n. \tag{1}$$

and

$$d_{eff} = \sum \frac{d_i^2}{d_i^2 + \lambda}, \tag{2}$$

where d_i are singular values of the predictor matrix - tabulated input variables used to fit the model - which summarise how much natural variation exists along each independent direction of the data. As λ increases, the regularisation more strongly constrains coefficients towards zero and the effective degrees of freedom, and therefore model complexity, reduces. This table is created using Regression_Methodology.ipynb.

S3 Supplementary Table 3

Supplementary Table 3 details seismograms in the Next Generation Attenuation Subduction (NGA-sub) database that occur outside of Japan, with a hypocentral depth ≤ 30 km, and yield a $\text{PGA}_{\text{RotD50}} \geq 0.1$ g (Mazzoni, 2022). We have also calculated I_{JMA} for each recording:

- All the event and station metadata is taken directly from the NGA-Sub database (Mazzoni, 2022).
- Seismic intensities in RotD50 are taken directly from the database, GM values are calculated from horizontal components using Equation 3 in main text. All values are displayed in gal.

S4 Supplementary Table 4

Supplementary Table 4 details a synthetic flat portfolio of assets and includes the following information:

- Latitude, Longitude - Location of an asset.
- osm_id - unique identifier for each location.
- Type - details whether an asset is a hospital or energy facility.
- Limit Flat - total limit of the asset if 100% loss is achieved.

Locations of these assets are taken from Byers et al. (2018) and Healthsites.io (2026).

S5 Supplementary Figures

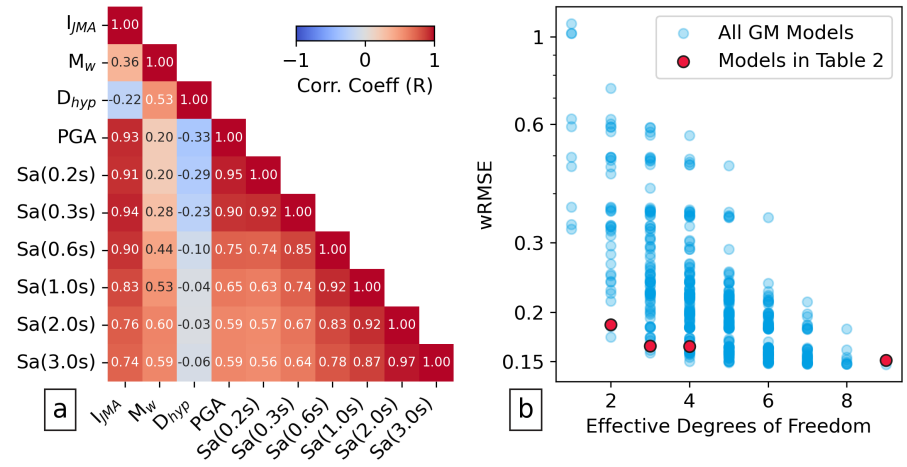


Figure S1 (a) Cross correlation matrix of earthquake observations with $I_{\text{JMA}} \geq 2$ from our earthquake catalogue (Supplementary Table 1). All ground motion intensities are presented as the GM of their horizontal components. D_{hyp} and ground motion parameters are presented as the $\log_{10}()$ of their respective values. Each combination of variables is coloured and annotated with its respective Pearson correlation coefficient calculated using their respective absolute values. (b) Weighted root mean squared error (wRMSE) as a function of the number of variables in 511 possible regression parameterisations constructed from variables shown in Figure S1a. Blue circles = all parameterisations, red circle = best-fitting models detailed in Table 2 in the main text.

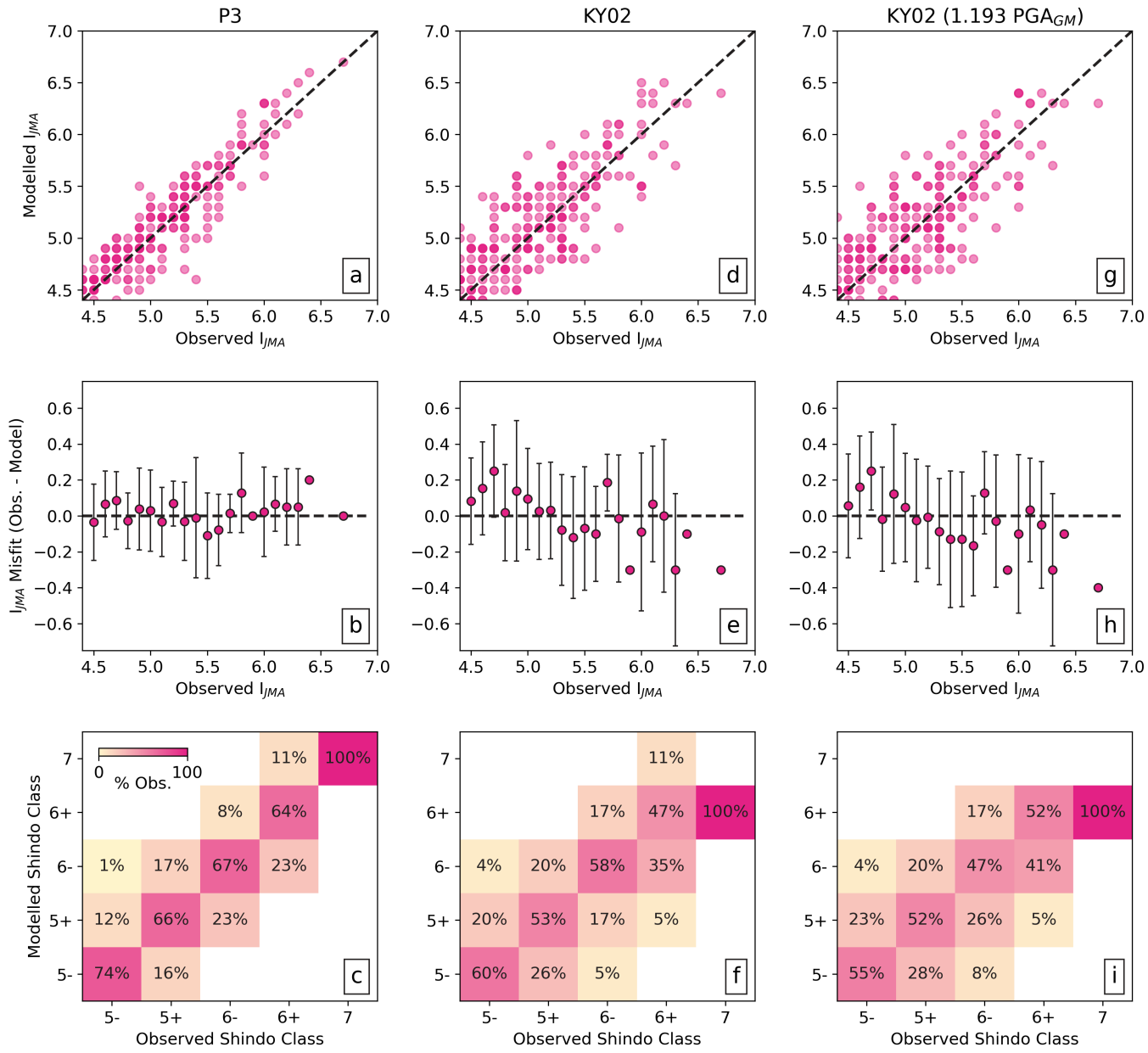


Figure S2 Comparison between observed and modelled I_{JMA} values using all non-Japanese accelerograms within the NGA-Subduction database with $\mathrm{PGA}_{\mathrm{RotD50}} \geq 0.1$ g (Mazzoni et al., 2023). (a) Modelled I_{JMA} as a function of observed I_{JMA} for all ground motions recordings with observed and modelled $I_{\mathrm{JMA}} \geq 4.5$. Dashed line = 1:1 for comparison. (b) Symbols and error bars = mean and standard deviation residual misfit between observed and modelled I_{JMA} for all samples in each 0.1 I_{JMA} interval. (c) A comparison between observed and modelled Shindo classes using conversion outlined in Table 1 in the main text. Number and colour of each bin denotes percentage of observed samples of the class denoted by the X axis in the bin. (d-f) Same as S2a-c, respectively with I_{JMA} calculated using Equation 14 in Karim and Yamazaki (2002). (g-i) Same as S2d-f, respectively, using $\mathrm{PGA}_{\mathrm{GM}} \times 1.193$ in place of $\mathrm{PGA}_{\mathrm{R}}$.

References

Byers, L., Friedrich, J., Hennig, R., Kressig, A., Li, X., Malaguzzi Valeri, L., and McCormick, C. A Global Database of Power Plants. Technical note, World Resources Institute, 2018. https://www.wri.org/research/global-database-power-plants.

Healthsites.io. The Global Healthsites Mapping Project. https://healthsites.io/, 2026. Open healthcare facility dataset and API. Accessed: 2026-05-29.

Heath, D. C., Wald, D. J., Worden, C. B., Thompson, E. M., and Smoczyk, G. M. A global hybrid V_{S30} map with a topographic slope-based default and regional map insets. *Earthquake Spectra*, 36(3):1570–1584, 2020. doi: 10.1177/8755293020911137.

Karim, K. R. and Yamazaki, F. Correlation of JMA instrumental seismic intensity with strong motion parameters. *Earthquake Engineering & Structural Dynamics*, 31(5):1191–1212, 2002. doi: 10.1002/eqe.158.

Kottke, A., Vanneste, K., and Bot, S. arkottke/pyrotd: v0.6.1, 2023. doi: 10.5281/zenodo.8101955.

Loviknes, K., Von Specht, S., Lilienkamp, H., Händel, A., Zhu, C., and Cotton, F. KiK-net and K-NET flatfile with automatically pro-

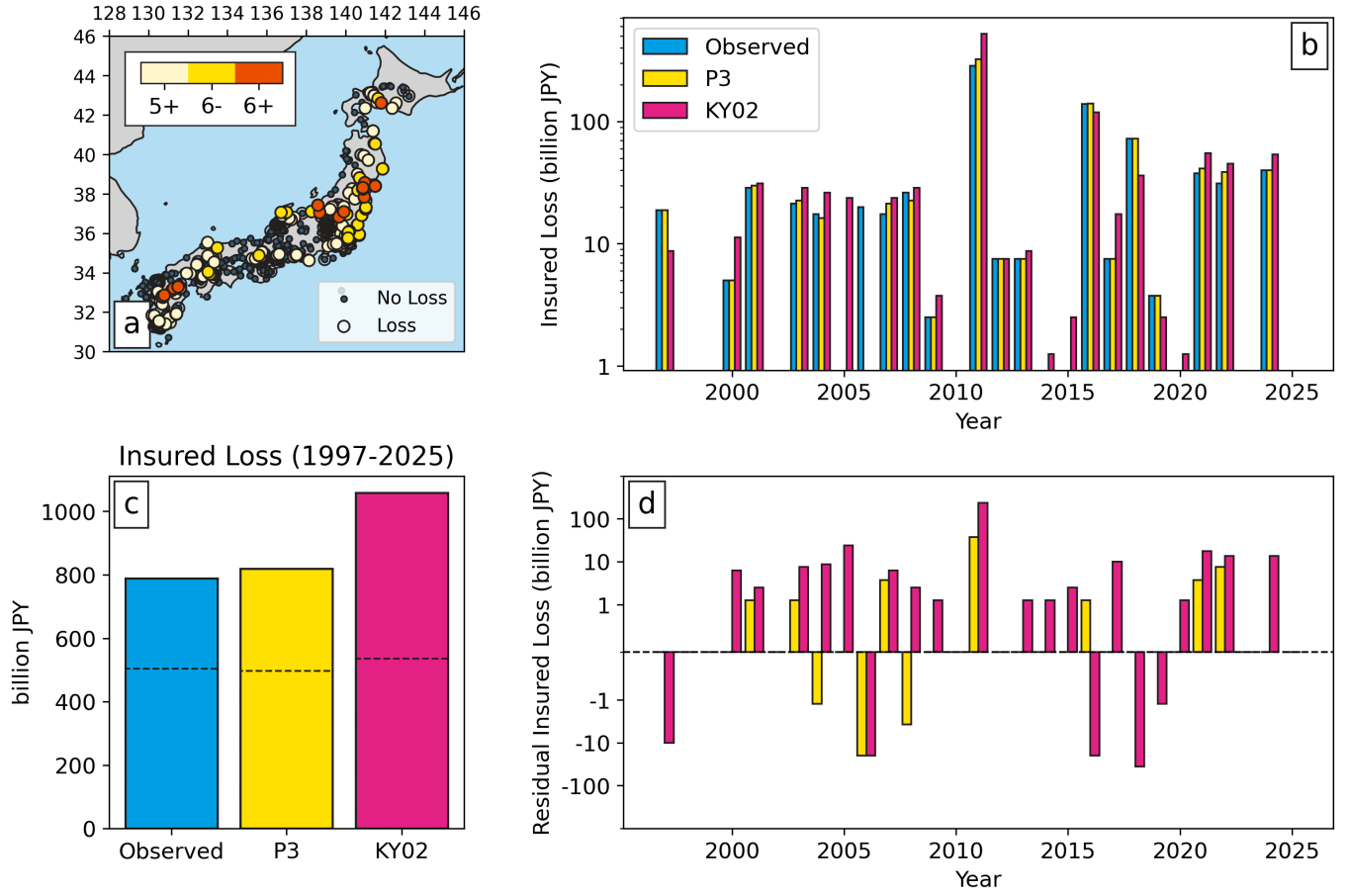


Figure S3 Application of I_{JMA} parameterisations to a synthetic parametric insurance portfolio using real events between 1997 and 2026 (Supplementary Table 1). (a) Locations within the synthetic portfolio; small and large circles are untriggered and triggered locations, respectively. Triggered locations coloured by the largest Shindo Scale experienced. (b) Loss experience time series; blue, yellow and pink bars show annual insured losses calculated using Shindo Class observations, those estimated using P3 and those calculated using KY02, respectively. (c) Aggregated losses for all events between 1997 and 2025; colours are the same as Figure S3b. Dashed line shows losses for all years except 2011. (d) Annual loss differences using Shindo Scale observations and parameterisation, positive values are parameterisation overestimates.

cessed ground motions and associated metadata V. 0002. Technical report, GFZ Data Service, 2026a. https://doi.org/10.5880/GFZ.LKUT.2026.003.

Loviknes, K., Von Specht, S., Lilienkamp, H., Händel, A., Zhu, C., and Cotton, F. Harmonization of KiK-net and K-NET Ground-Motion Parameters and Metadata: A Flatfile of 950,000 Three-Component Records Spanning 29 Years. *Seismica*, 5(1), 2026b.

Mazzoni, S. NGA-subduction portal: Ground-motion record selection and download. Technical report, The B. John Garrick Institute for the Risk Sciences, 2022. doi: 10.34948/N3D59V.

Mazzoni, S., Kishida, T., and Bozorgnia, Y. NGA-Subduction Component PSA IM Tables: H1, H2, V, RotD50, rotD100 5 Technical report, The B. John Garrick Institute for the Risk Sciences, 2023. doi: 10.34948/N3MS3N.

National Research Institute for Earth Science and Disaster Resilience. K-NET, KiK-net. Dataset, 2019. doi: 10.17598/NIED.0004.

Neath, A. A. and Cavanaugh, J. E. The Bayesian information criterion: background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2):199–203, 2012.

RR-Inyo. shindo: Calculates JMA (Japan Meteorological Agency) Seismic Intensity (Shindo) Scale from Acceleration Data Recorded in NumPy Array. https://github.com/RR-Inyo/shindo, 2021. https://github.com/RR-Inyo/shindo. Accessed: 2026-06-01.

Zhu, C., Weatherill, G., Cotton, F., Pilz, M., Kwak, D. Y., and Kawase, H. An open-source site database of strong-motion stations in Japan: K-NET and KiK-net (v1. 0.0). *Earthquake Spectra*, 37(3):2126–2149, 2021.

[ht!]

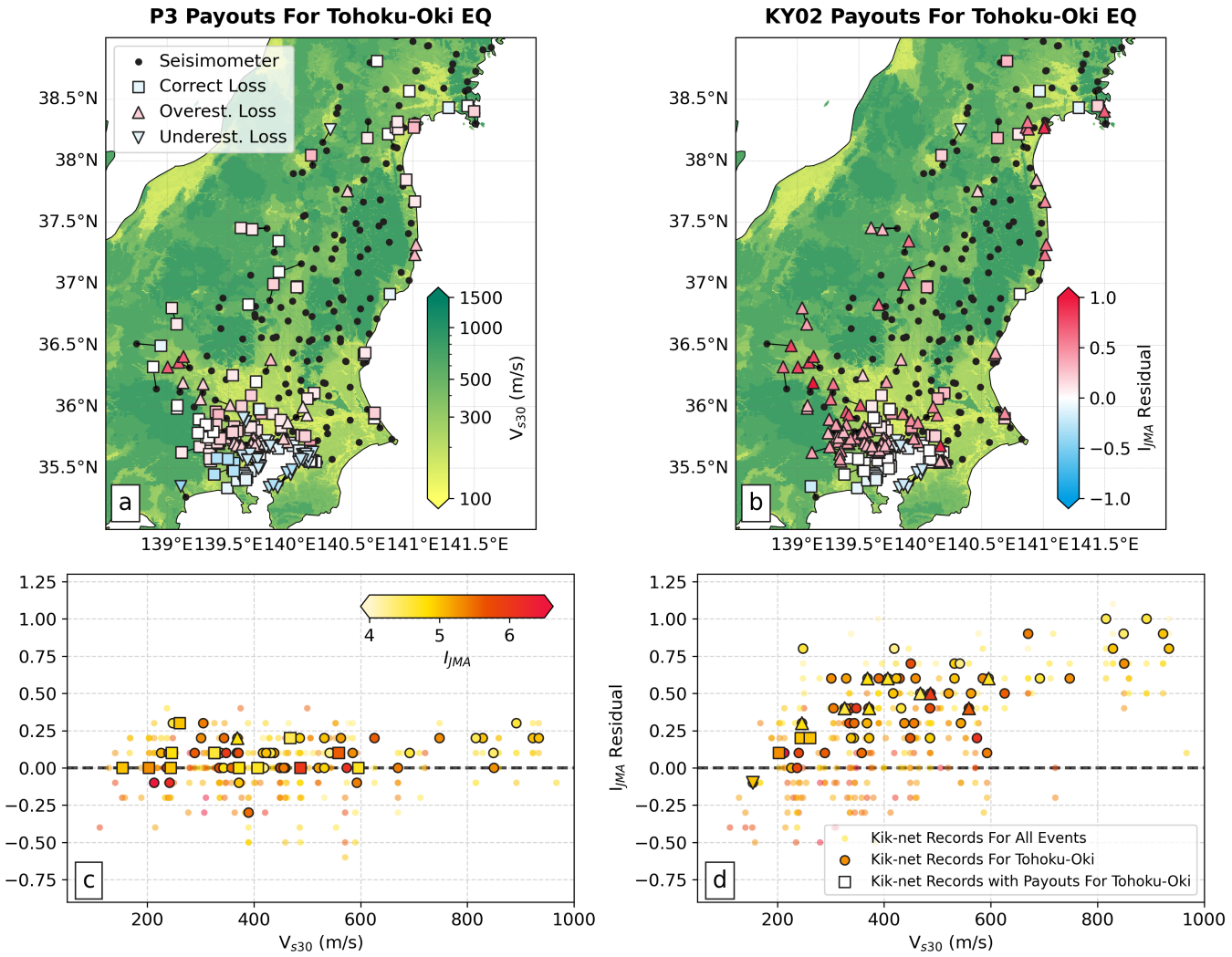


Figure S4 Parameterisation performance for the Tohoku-Oki Earthquake, 2011. (a) Map of Tohoku-Oki-earthquake-affected region with background grid coloured by shear-wave velocity in the top 30 m of the subsurface (V_{s30}; Heath et al., 2020). Inverted-triangles, squares, and triangles are insured locations where triggered losses in the synthetic portfolio are under-, correctly and over-estimated using P3, respectively. Symbols are coloured by the residual misfit between parameterised and observed I_{JMA} and black lines tie insured locations to their nearest seismometer. Black dots are locations of seismometers with Tohoku-Oki earthquake ground motion intensity recordings (Supplementary Table 1). (b) Same as Figure S4a with I_{JMA} values calculated using KY02. (c) Difference between observed and parameterised KiK-net station I_{JMA} values ≥ 4.5 using Equation 14 in Karim and Yamazaki (2002) as a function of V_{s30} values provided by Zhu et al. (2021). Small circles are all seismometer recordings for all earthquakes in Supplementary Table 1; large circles are recordings for Tohoku-Oki earthquake; inverted-triangles/squares/triangles are recordings that trigger losses in the synthetic portfolio that are under-/correctly/over-estimated using P4. All symbols are coloured by their respective observed I_{JMA} values. (d) Same as Figure S4c but parameterised I_{JMA} calculated using KY02 (Equation 10 in main text).