Supporting information for

Seismic Swarm and Surface Deformation in Kamanjab Inlier, Northern Namibia

Text S1. In configuring the parameters for EqTransformer and Siamese-EqTransformer, we first selected a larger overlap parameter in the preprocessor to ensure that larger magnitude events would not be missed (for more details, refer to EqTransformer GitHub repository). Second, using the original model, we set the higher detection threshold for both the P and S phases to 0.4 and the earthquake threshold to 0.7. Finally, we determined that for potential earthquake events, the time difference between the P and S phases must be less than 30 seconds, and we only retained both phases if the other was detected for the same event. All other parameters were set to their default values. We ran the process using an older GPU card, the GTX Super 1660; processing one month's data from 10 stations required approximately two hours for phase pick detection.

For the Siamese-EqTransformer, we set the *P_skip_threshold* and *S_skip_threshold* to 0.6. This indicates that the neural network initially labels picks with a confidence value greater than 0.6 (on a scale from 0 to 1, similar to the EqTransformer detection threshold). Subsequently, the process identifies additional stations to enhance the signal-to-noise ratio within time windows of 6 seconds for the P phase and 12 seconds for the S phase. Then, EqTransformer is run again with the same detection threshold of 0.4. For more detailed information, please refer to the Siamese-EqTransformer GitHub repository.

EqT:

```
preprocessor(preproc_dir="data_processed_hdfs",
             mseed_dir='data',
             stations_json=json_basepath,
             overlap=0.7,
             n_processor=16)

predictor(input_dir='data_processed_hdfs',
          input_model='your/model/location/EqT_original_model.h5',
          output_dir='detection_results',
          estimate_uncertainty=False,
          output_probabilities=False,
          number_of_sampling=5,
          loss_weights=[0.02, 0.40, 0.58],
```

detection_threshold=0.7,
P_threshold=0.4,
S_threshold=0.4,
number_of_plots=0,
plot_mode='time',
batch_size=500,
number_of_cpus=16,
gpuid=0,
gpu_limit=90,
keepPS=True,
spLimit=30)

S_EqT:
txt_folder: ./SEqTRes/
P_branch_config: '/your/SEQT/package/location/SiameseEarthquakeTransformer/S_EqT_codes/configs/P_branch.yaml'
P_branch_model: '/your/SEQT/package/location/SiameseEarthquakeTransformer/models/S_EqT/S_EqT_P_branch.hdf5'
P_threshold: 0.4
P_skip_threshold: 0.6
keep_time_range_P: 6.0

S_branch_config: '/your/SEQT/package/location/SiameseEarthquakeTransformer/S_EqT_codes/configs/S_branch.yaml'
S_branch_model: '/your/SEQT/package/location/SiameseEarthquakeTransformer/models/S_EqT/S_EqT_S_branch.hdf5'
S_threshold: 0.4
S_skip_threshold: 0.6
keep_time_range_S: 12.0

Text S2. The magnitude of completeness, Mc, was first estimated using the maximum‐curvature method: magnitudes were binned in 0.1-unit intervals and the bin with the greatest event count was taken as the initial Mc, yielding Mc = -0.20. We then refined this estimate by applying the goodness-of-fit test of Wiemer and Wyss (2000) to trial values in a range around the initial Mc = -0.2 estimate, specifically here from -0.35 to -0.05 at 0.01 unit increments. The optimal Mc was chosen as the value for which the

observed and modeled Gutenberg–Richter distributions agree at the 95 % confidence level and the coefficient of determination R^2 is maximized. This procedure produced a final completeness magnitude of Mc = -0.29. An example of the calculation output is shown below:

Initial MC from Maximum Curvature Method: -0.200000000000000
Candidate MC: -0.35, R-squared: 0.99323
Candidate MC: -0.34, R-squared: 0.99348
Candidate MC: -0.33, R-squared: 0.99361
Candidate MC: -0.32, R-squared: 0.99391
Candidate MC: -0.31, R-squared: 0.99380
Candidate MC: -0.30, R-squared: 0.99317
Candidate MC: -0.29, R-squared: 0.99422
Candidate MC: -0.28, R-squared: 0.99254
Candidate MC: -0.27, R-squared: 0.99268
Candidate MC: -0.26, R-squared: 0.99254
Candidate MC: -0.25, R-squared: 0.99274
Candidate MC: -0.24, R-squared: 0.99305
Candidate MC: -0.23, R-squared: 0.99323
Candidate MC: -0.22, R-squared: 0.99360
Candidate MC: -0.21, R-squared: 0.99352
Candidate MC: -0.20, R-squared: 0.99285
Candidate MC: -0.19, R-squared: 0.99400
Candidate MC: -0.18, R-squared: 0.99213
Candidate MC: -0.17, R-squared: 0.99232
Candidate MC: -0.16, R-squared: 0.99221
Candidate MC: -0.15, R-squared: 0.99241
Candidate MC: -0.14, R-squared: 0.99278
Candidate MC: -0.13, R-squared: 0.99302
Candidate MC: -0.12, R-squared: 0.99348
...
Candidate MC: -0.06, R-squared: 0.99195
Candidate MC: -0.05, R-squared: 0.99216

Best MC from Goodness-of-Fit Test: -0.29 with R-squared = 0.99422

Depth (km)	Vp(km/s)	Vs (km/s)
0	6.25	3.592
4	6.31	3.626
6	6.74	3.874
20	7.58	4.356
25	8.23	4.73
30	8.48	4.874
38	8.85	5.086

Table S1. The 1D velocity model used in this study, adopted from Midiz, 2010.

Event Time (UTC)	Strike	Dip	Rake	Mw
2018 0714 14:08:11	325.3	62.7	-83.8	2.642
2018 0718 14:40:43	314	61.6	-105.3	2.929
2018 0730 13:48:47	335.7	64.5	-88.7	3.528
2018 0730 18:17:05	333.2	67.5	-92.2	2.661
2018 0809 16:24:31	351	65.0	-70.1	2.524
2018 0812 16:02:13	344.3	63.1	-77.7	3.413
2018 0812 16:38:22	327.9	61.3	-82.3	2.683
2018 0814	10	69.1	-63.2	2.504

16:32:35				
2018 0820 13:43:36	336.4	66.8	-92.0	2.663
2018 0824 00:57:36	339.4	59.7	-82.1	3.290
2018 0902 10:06:13	348.6	60.8	-74.1	3.051
2018 0905 04:56:50	351	60	-98.9	2.627
2018 0908 20:02:30	360	56.2	-77.8	3.07
2018 0910 07:48:04	345	63.5	-89.4	2.609
2018 0910 11:13:47	328	62.4	-102.5	2.628
2018 0917 19:43:22	338.5	62.8	-80.5	2.607
2018 0918 23:15:34	334	66	-108	2.61

Table S2. Focal Mechanism Solutions (FMSs) for 17 events with local magnitude (ML) greater than 2.0. The nodal plane determination is based on InSAR deformation data and FMS data from the 2021 Mw 5.4 events provided by the USGS.

Cluster No.	Daily mean number	Standard deviation
Cluster 1	26.18	20.59
Cluster 2	15.21	9.99
Cluster 3	14.27	16.71
Cluster 4	2.52	4.39

Cluster 5	2.38	1.70

Table S3. Average daily relocated event and standard deviation for different cluster.

Date	Cluster 1 Event Number	Cluster 2 Event Number	Cluster 3 Event Number	Cluster 1 Largest Magnitude (M_L)	Cluster 2 Largest Magnitude (M_L)	Cluster 3 Largest Magnitude (M_L)
2018/07/14	μ+2.13σ	μ+0.68σ	μ-0.38σ	2.17	0.62	0.69
2018/07/19	μ+2.81σ	μ+0.88σ	μ-0.56σ	1.86	1.62	1.07
2018/07/21	μ+1.64σ	μ-0.02σ	μ-0.38σ	1.57	0.62	0.93
2018/07/29	μ+0.04σ	μ+1.28σ	μ+2.20σ	0.96	1.24	2.32
2018/07/30	μ+0.53σ	μ-0.62σ	μ+5.43σ	0.57	1.40	3.34
2018/07/31	μ+0.23σ	μ+3.88σ	μ+5.01σ	1.54	1.87	1.94
2018/08/05	μ+1.84σ	μ-0.42σ	μ-0.48σ	1.72	0.41	0.48
2018/08/12	μ+3.10σ	μ-0.32σ	μ-0.50σ	2.90	0.82	0.71
2018/08/13	μ+2.37σ	μ-0.72σ	μ+0.58σ	1.36	-0.07	0.63
2018/08/14	μ+2.32σ	μ-0.92σ	μ-0.44σ	2.00	0.21	0.51
2018/08/20	μ-0.15σ	μ+1.88σ	μ+0.16σ	0.55	1.63	2.25
2018/08/24	μ-0.25σ	μ-0.22σ	μ-0.26σ	0.91	1.06	2.91
2018/09/10	μ-1.08σ	μ+2.38σ	μ-0.14σ	-0.11	2.10	0.86
2018/09/15	μ-0.64σ	μ-0.52σ	μ+1.54σ	1.18	1.25	1.36
2018/09/17	μ+1.93σ	μ-1.12σ	μ-0.56σ	1.44	0.74	0.90
2018/09/19	μ-0.40σ	μ+4.48σ	μ-0.73σ	0.26	1.46	0.19

Table S4. Days on which a relocated cluster's daily event count exceeds the overall mean (μ) plus 1.5 standard deviations (σ) for the entire study period, along with the event counts (highlighted as μ ± σ) and the largest magnitudes of the three main

clusters (Clusters 1, 2, and 3). Blue text indicates which cluster's event count surpasses the threshold, and highlighted dates denote anomalies in Cluster 4 rather than in the three main clusters. Note that the largest magnitude in this table on a given date only focus on the events have been relocated using the HypoDD method.

Event Time (UTC)	M_L	M_W	Cluster
2018 0714 14:08:11	2.17	2.642	1
2018 0718 14:40:43	2.60	2.929	1
2018 0730 13:48:47	3.34	3.528	3
2018 0730 18:17:05	2.28	2.661	N/A
2018 0809 16:24:31	2.07	2.524	1
2018 0812 16:02:13	2.90	3.413	1
2018 0812 16:38:22	2.28	2.683	1
2018 0814 16:32:35	2.01	2.504	1
2018 0820 13:43:36	2.18	2.663	3
2018 0824 00:57:36	2.91	3.290	3
2018 0902 10:06:13	2.51	3.051	N/A
2018 0905 04:56:50	2.10	2.627	2
2018 0908 20:02:30	2.51	3.07	3
2018 0910 07:48:04	2.02	2.609	2
2018 0910 11:13:47	2.10	2.628	2
2018 0917 19:43:22	2.11	2.607	N/A
2018 0918 23:15:34	2.07	2.61	2

Table S5. M_L over 2.0 event from the machine-learning enhanced catalog, showing the corresponding moment magnitude (Mw) derived from moment tensor inversion and the clustering results after relocation. Note that three events are not assigned to any cluster because they were not relocated through HypoDD.

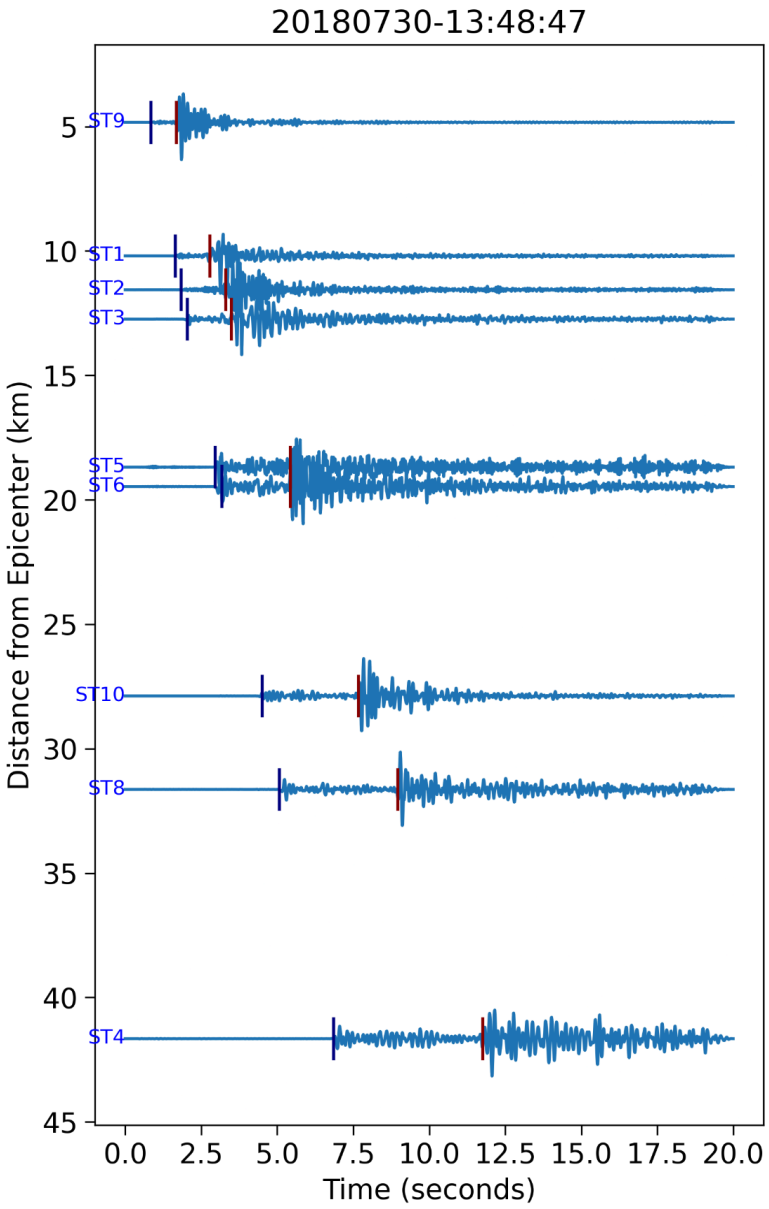

Figure S1. An example of an event and its corresponding detected phase picks. All events that occurred on July 30, 2018, along with their detected phase picks, are showcased in Movie S1.

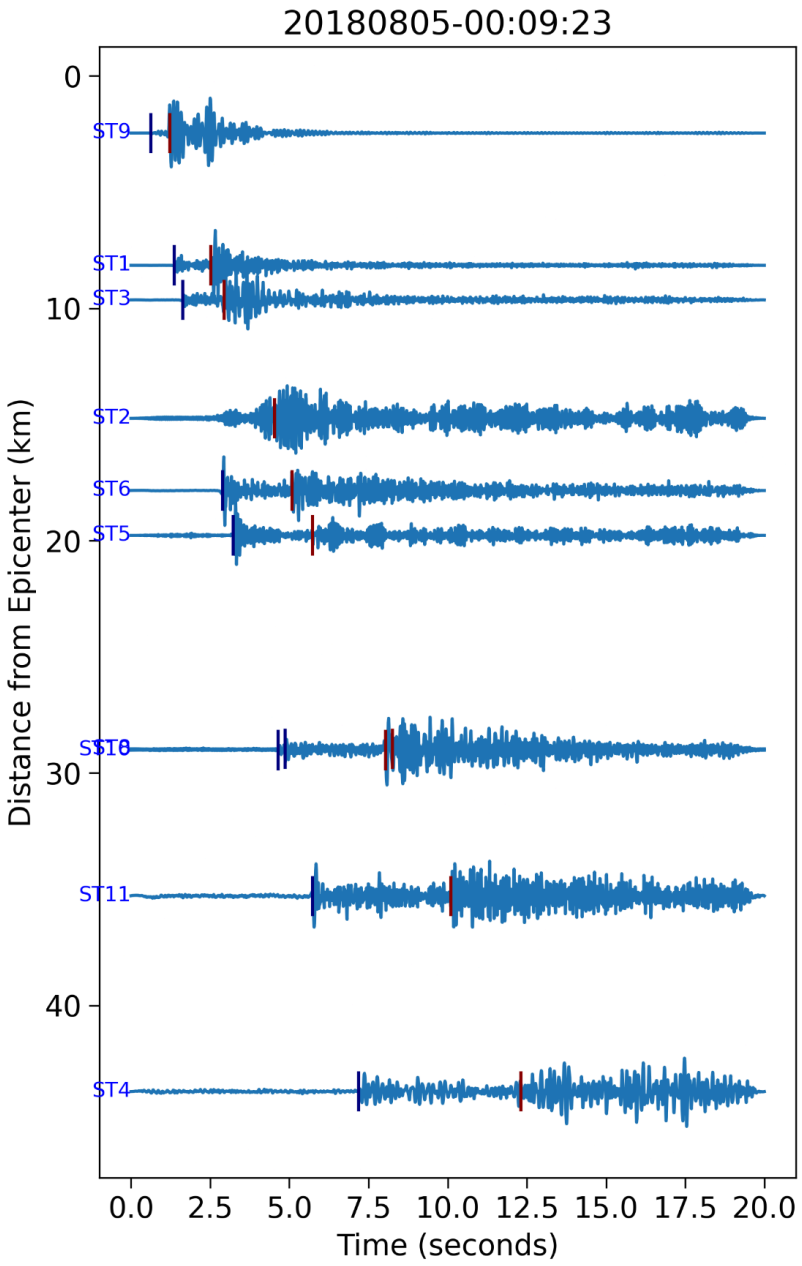

Figure S2. An example of an event and its corresponding detected phase picks. All events that occurred on Aug 5, 2018, along with their detected phase picks, are showcased in Movie S2.

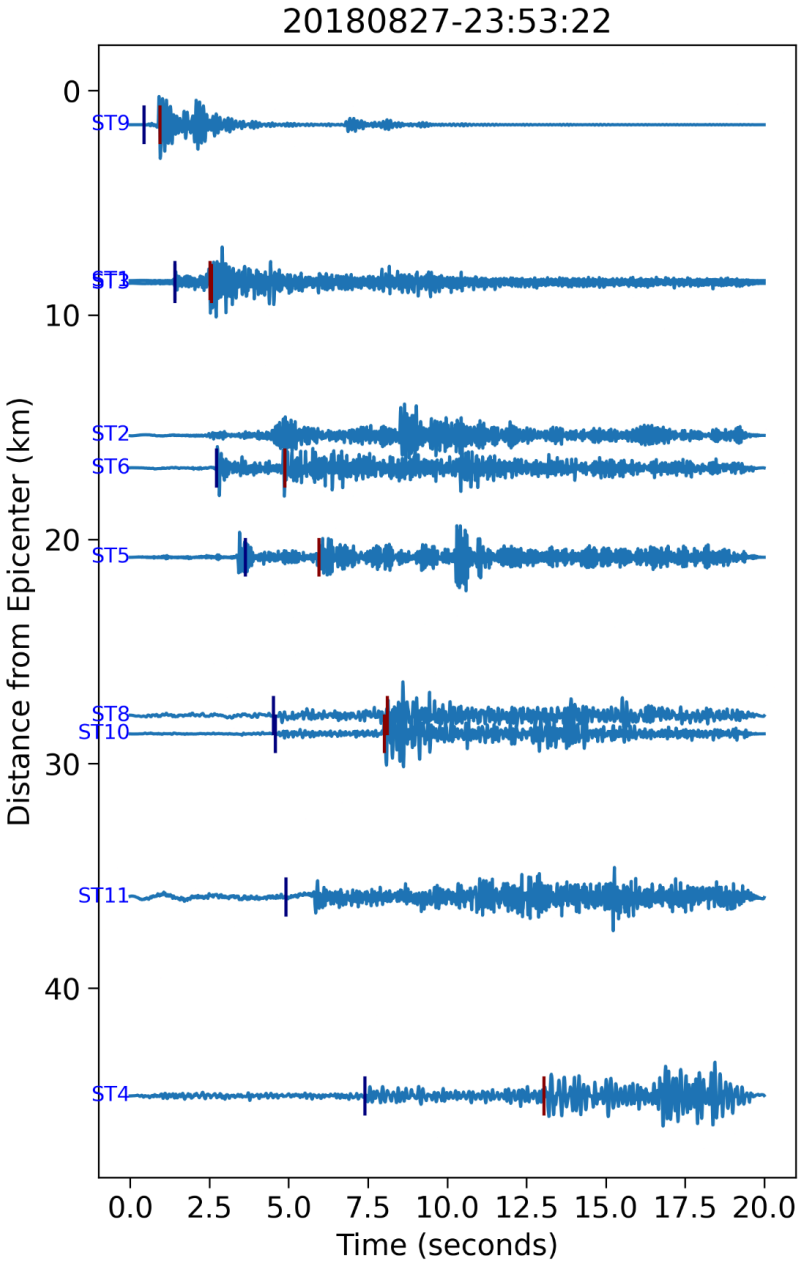

Figure S3. An example of an event and its corresponding detected phase picks. All events that occurred on Aug 27, 2018, along with their detected phase picks, are showcased in Movie S3.

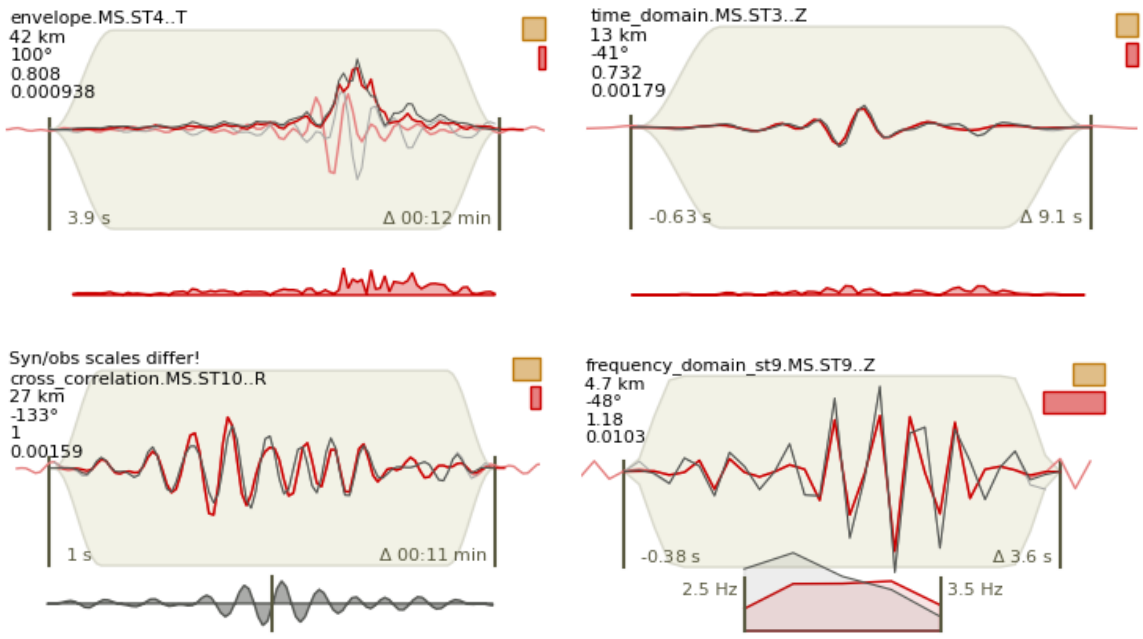

Figure S4. Four examples of fits across the envelope, time, frequency, and waveform cross-correlation domains. The synthetic waveforms are depicted by the red lines, while the original waveforms are represented by the black lines.

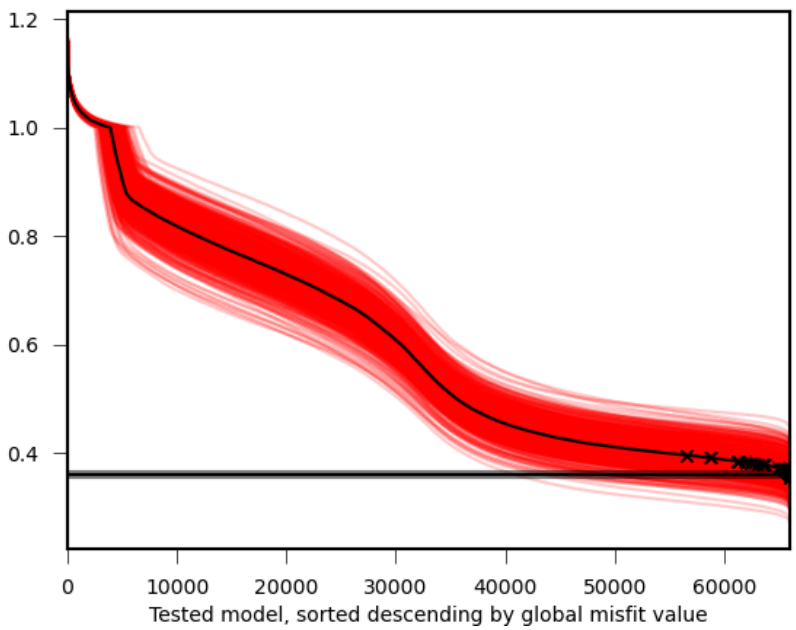

Figure S5. Demonstration of the bootstrapping misfit value decreases to approximately 0.37 and stabilizes after 60,000 iterations.

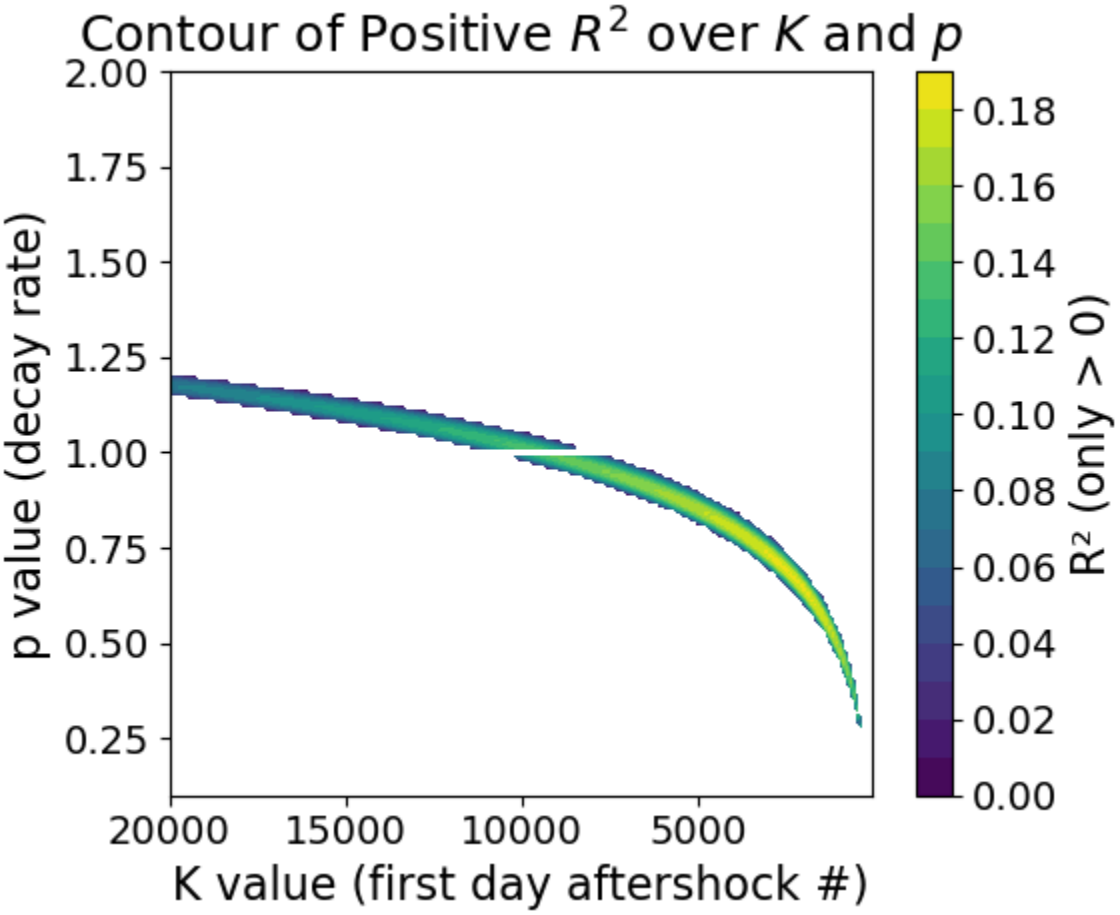

Figure S6. Distribution of R square value between daily event number and Omori fitting with different decay rates p and assumed first-day aftershock numbers k.

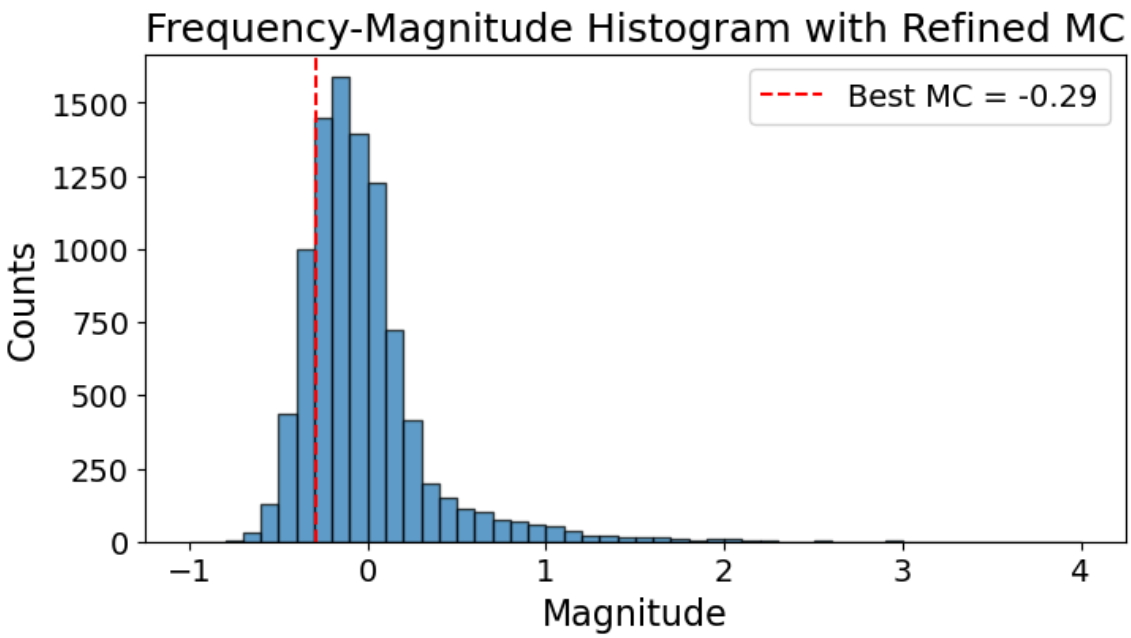

Figure S7. Histogram of magnitude distribution of initial machine-learning catalog and highlighted Mc (see Text S2 for detail calculation process).

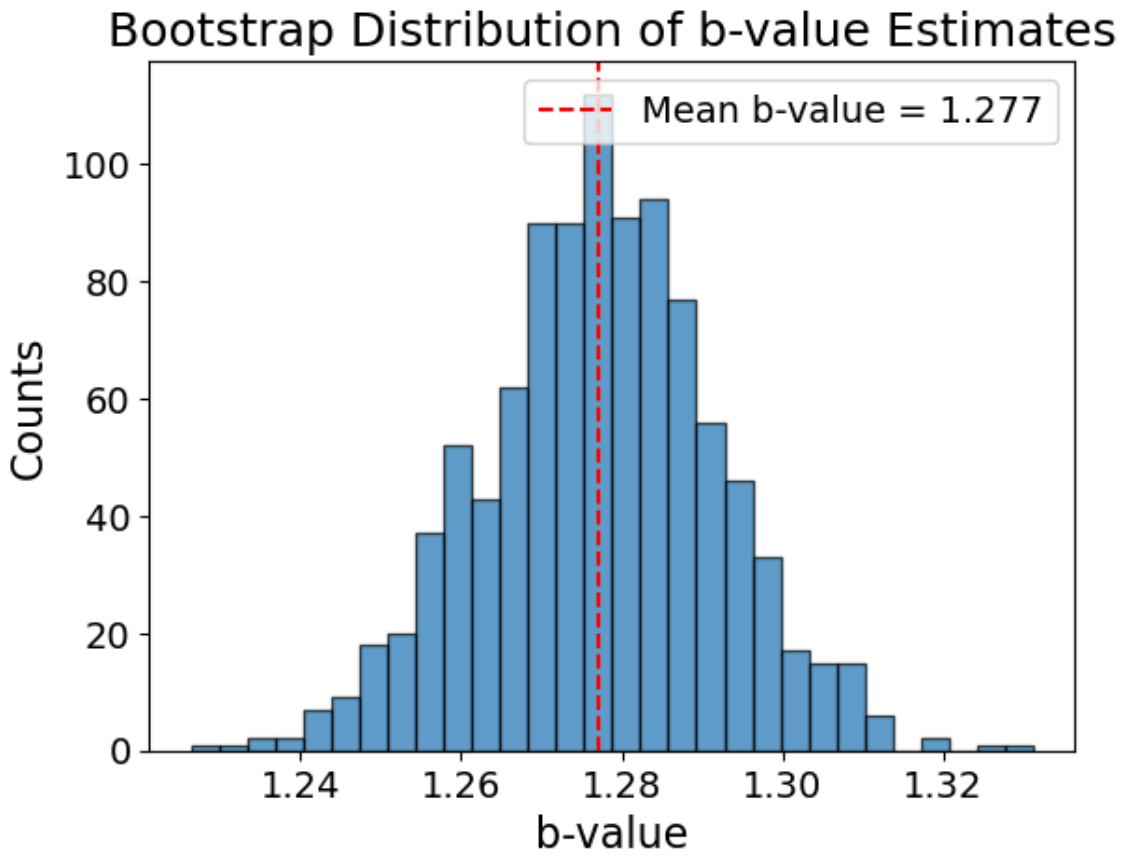

Figure S8. b-value estimation of 1000 times bootstrapping distribution.

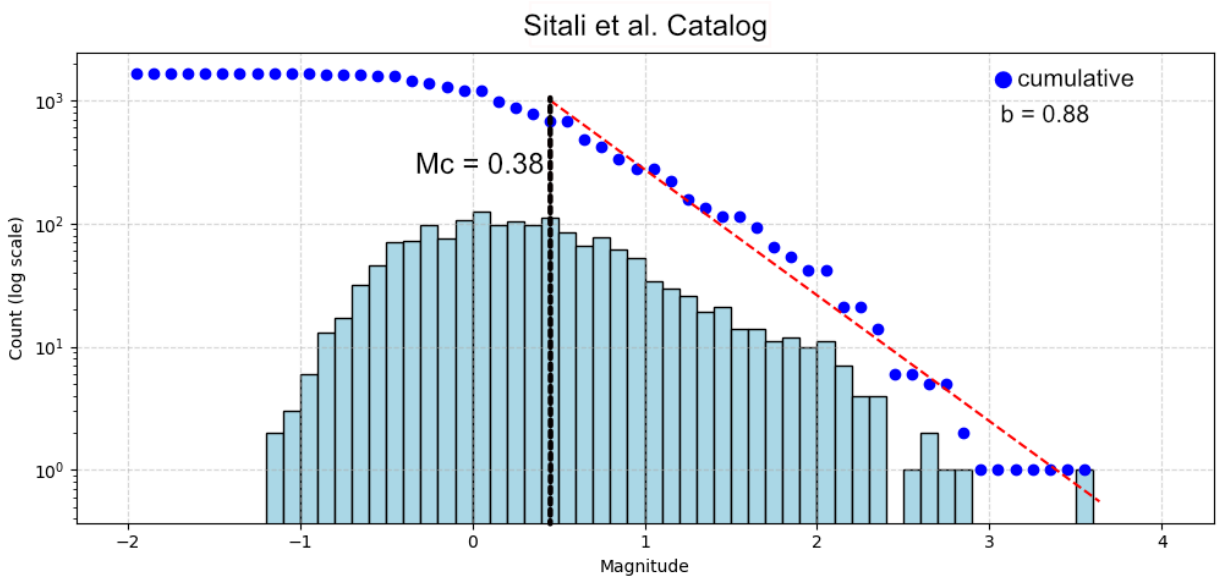

Figure S9. Magnitude-frequency distribution, fitted using the maximum likelihood method (equation 1 in main text), yields a b-value of 0.88 and a completeness magnitude Mc of 0.38 (black dashed line) in Sitali et al. catalog. Note that the magnitude distribution may not be correct because of the station offset (see main text in Section 3.1.2).

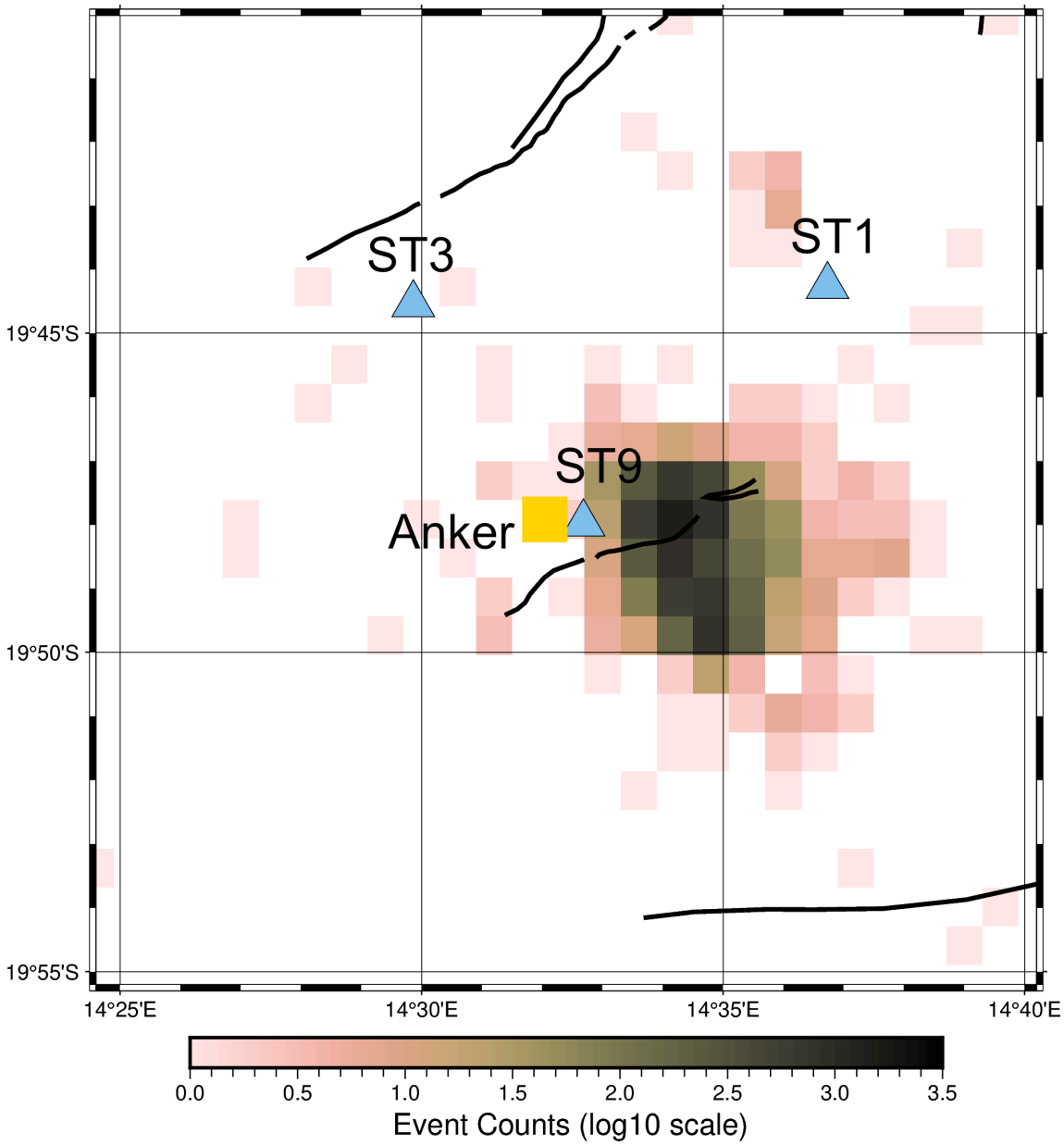

Figure S10. The number of events from the machine learning catalog is located in approximately 1 km* 1 km box area. Most of the events are near ST9 (Anker Village).

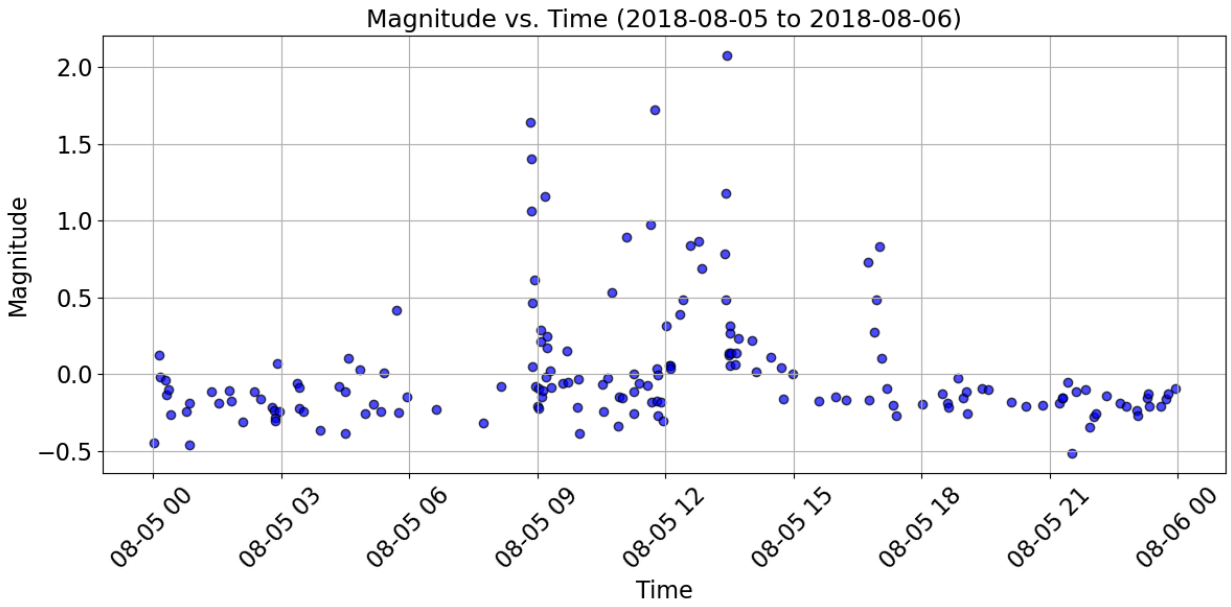

Figure S11. An example of a daily event bursting on 5th August.

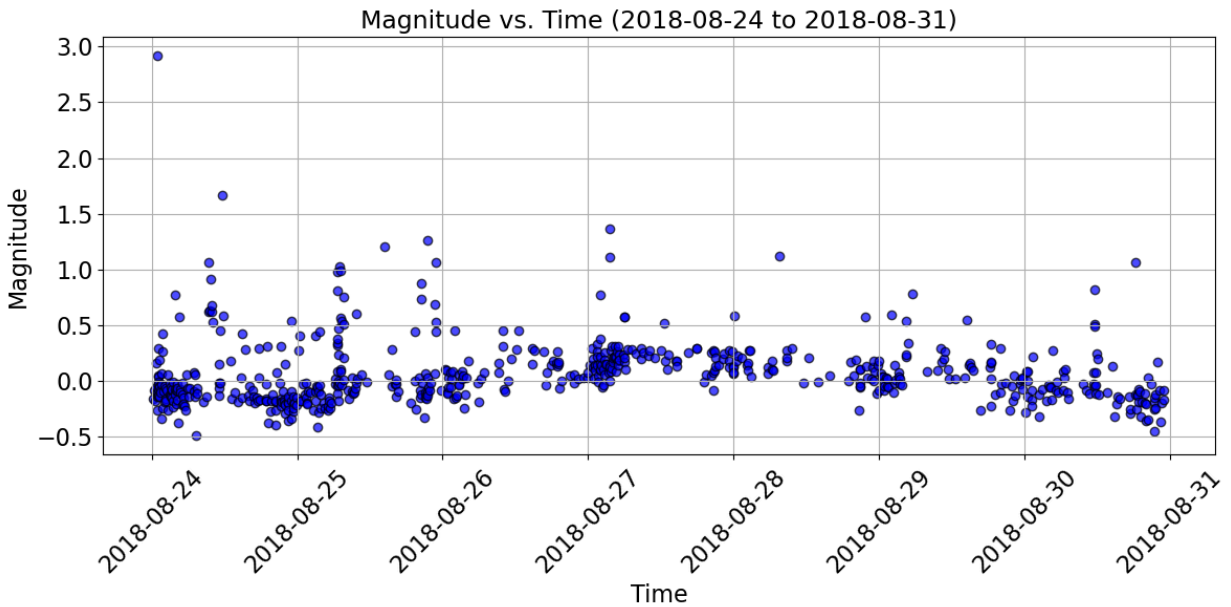

Figure S12. An example of the variation of a minor magnitude event. The minor magnitude events on the date of 27th August have an overall larger magnitude compared to other dates in a week.

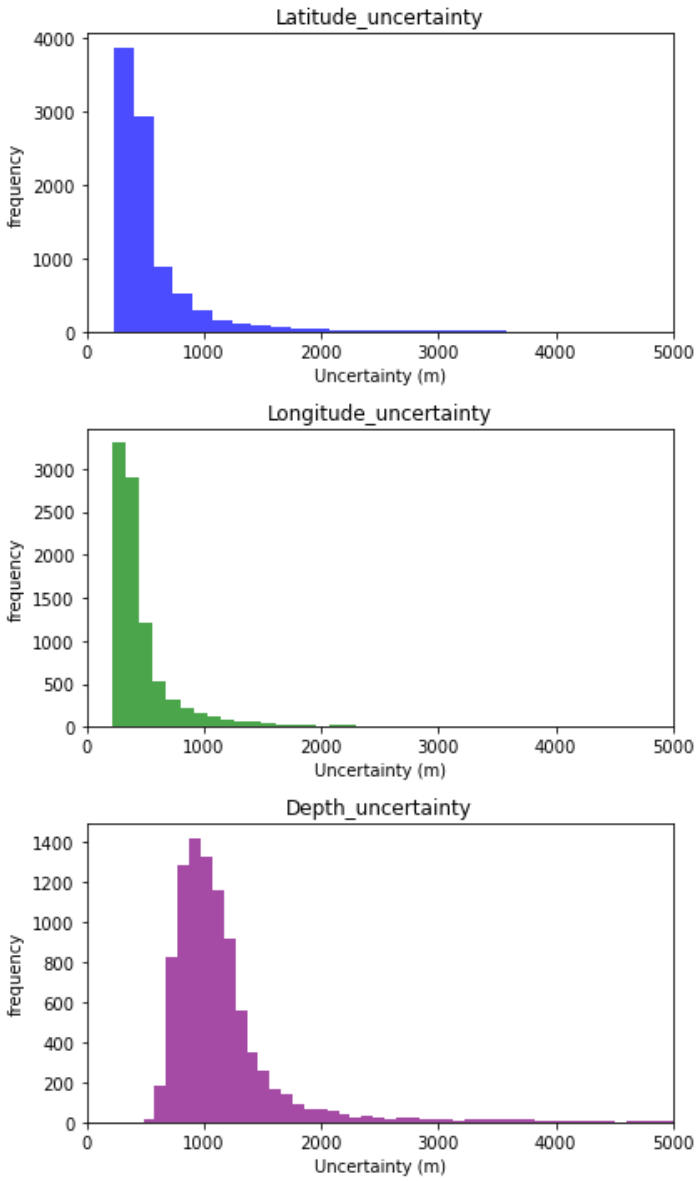

Figure S13. Location uncertainty in 3 directions with a 95% confidence level estimated from NonLinLoc.

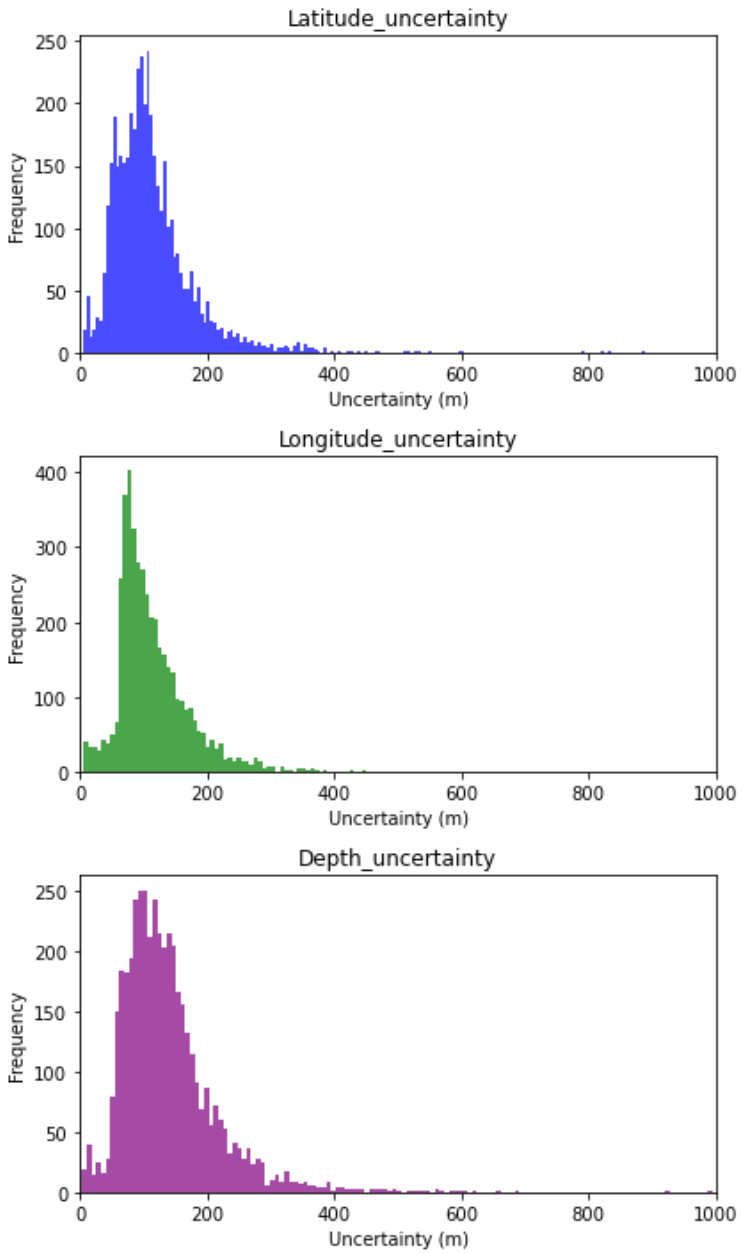

Figure S14. Location uncertainty for HypoDD relocation result, with the majority of events' uncertainty less than 250 m.

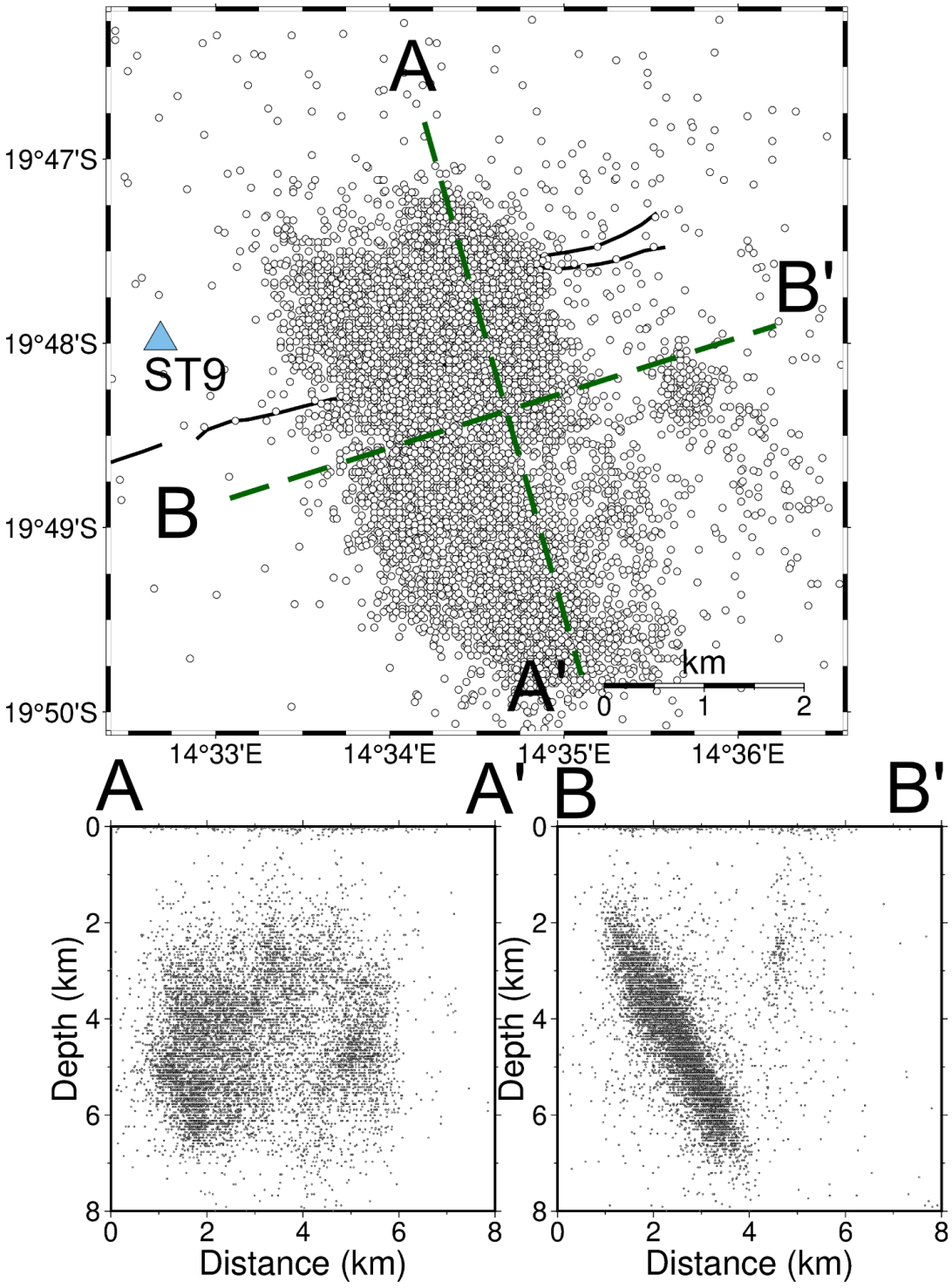

Figure S15. Earthquake location for the initial Machine Learning catalog in the same area as Figure 4a, 4b, and 4c in main text for comparison.

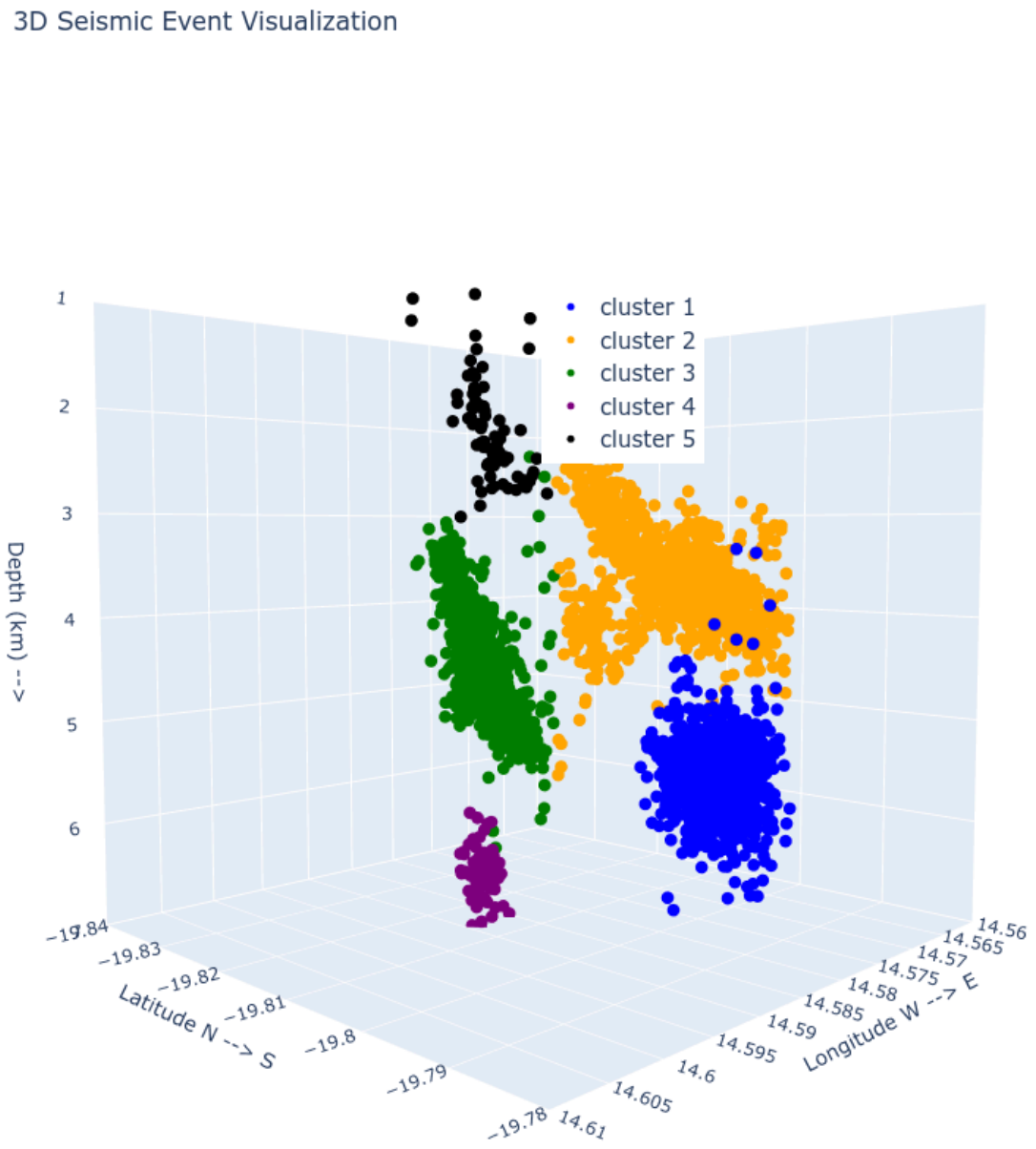

Figure S16. Distribution of relocated events from the initial machine learning catalog with 3D (latitude, longitude, depth) DBSCan clustering. See Movie S.4 for different angles of view.

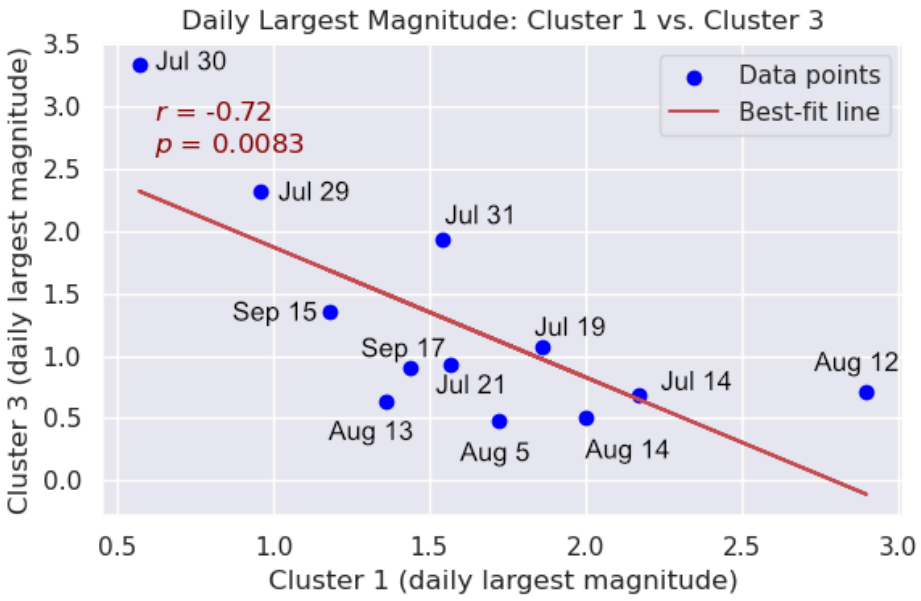

Figure S17. The correlation of daily largest magnitude in anomaly dates of Cluster 1 and 3 highlighted in Table S3 between Cluster 1 and Cluster 3.

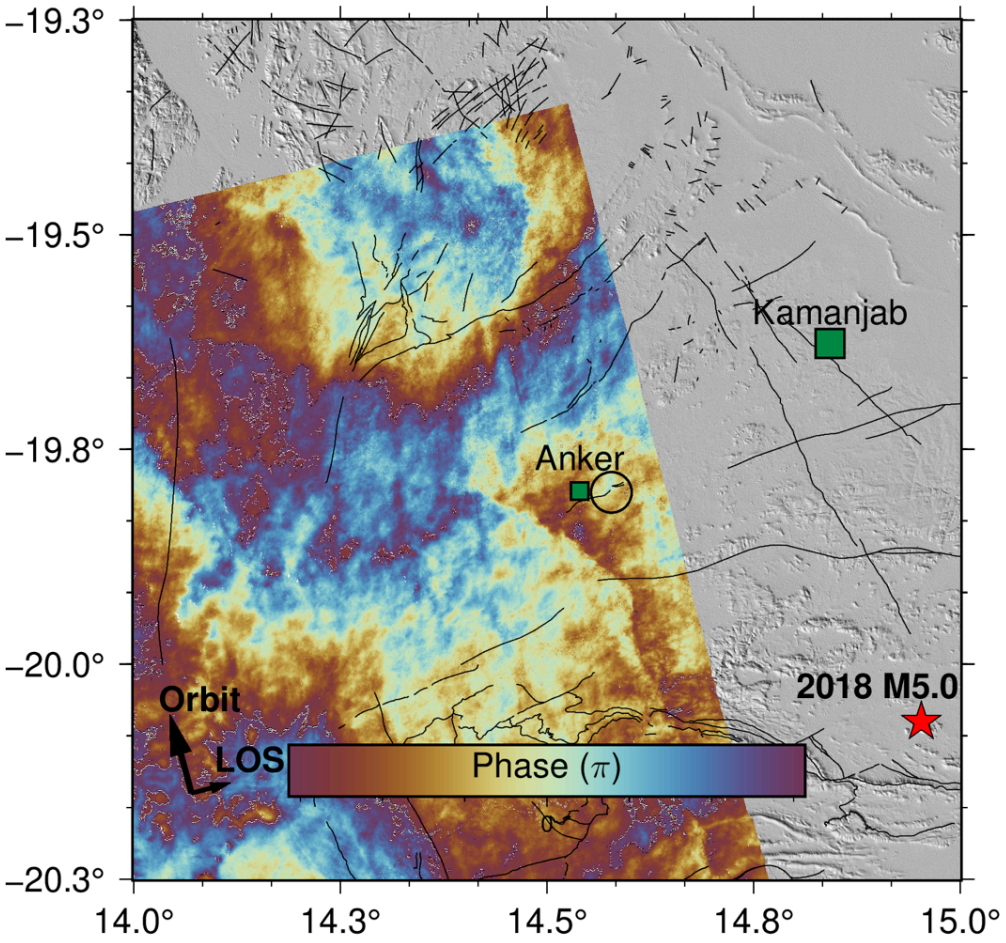

Figure S18. A Coseismic 12-day ascending co-seismic interferograms (20 May—1 June 2018) for the earthquake from Sentinel-1B imagery shows the USGS 24 May 2018 Mw 5.0 event location (red star).

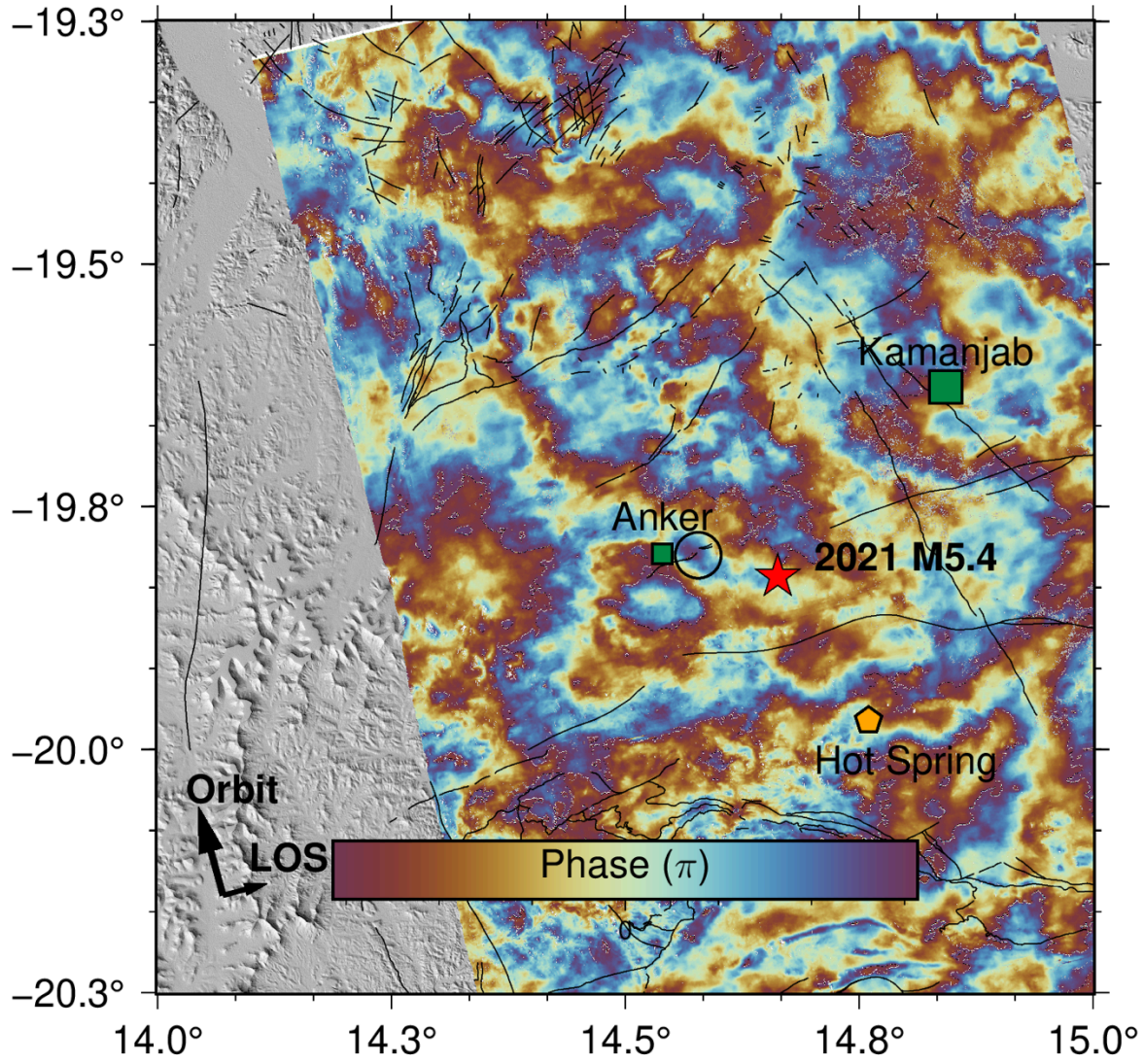

Figure S19. A Coseismic 12-day ascending co-seismic interferograms (29 March—10 April 2021) for the earthquake from Sentinel-1B imagery shows the USGS 04 April 2021 Mw 5.4 event location (red star).

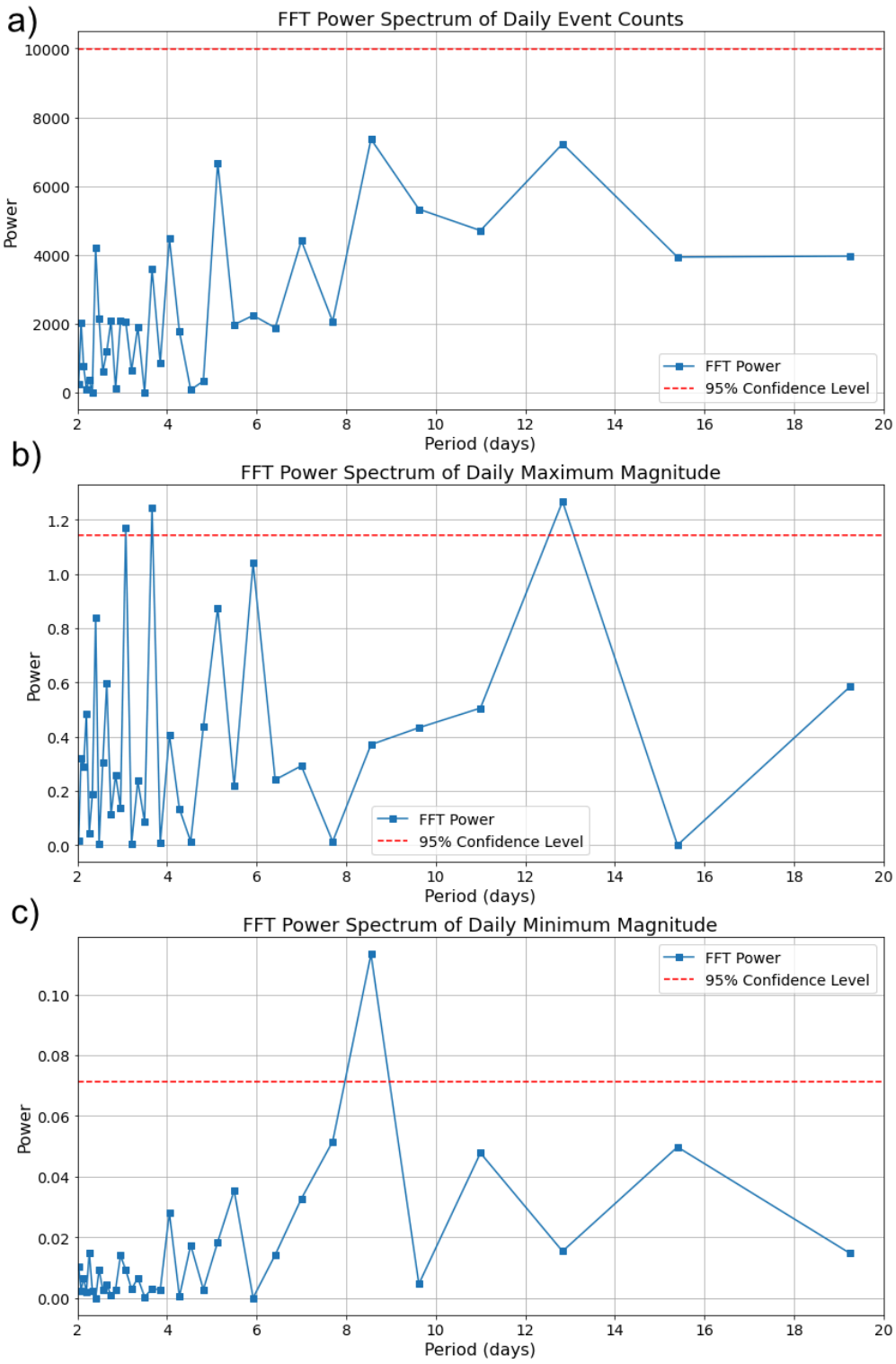

Figure S20. Fast Fourier transform analysis with all machine-learning catalog events with daily event number, maximum magnitude, and minimum magnitude.

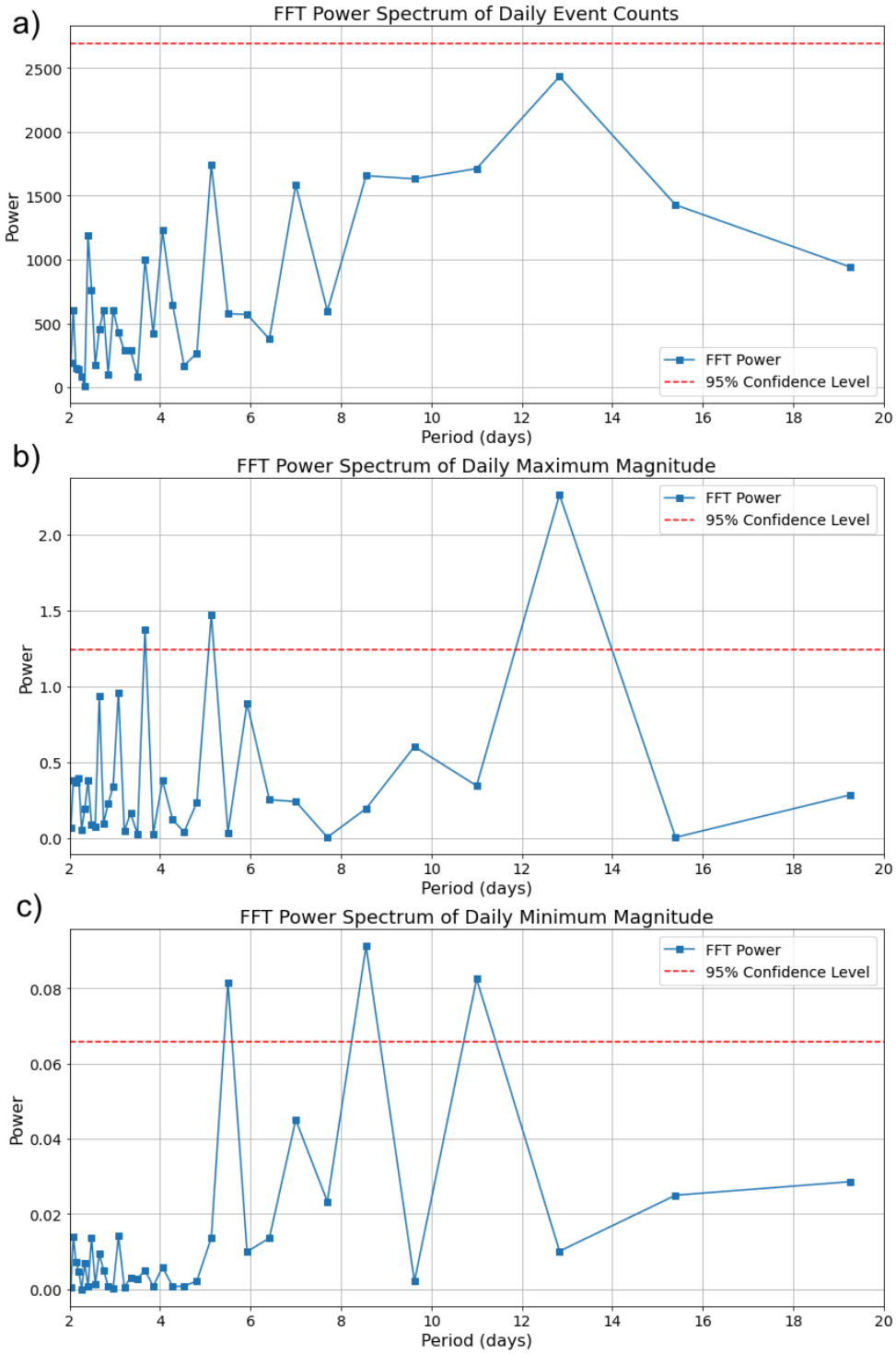

Figure S21. Fast Fourier transform analysis with all relocated catalog events with daily event number, maximum magnitude, and minimum magnitude

Movie S1. All detected events that occurred on July 30, 2018, along with their detected phase picks (blue for P phase, and red for S phase)
Movie S2. All detected events that occurred on August 5, 2018, along with their detected phase picks (blue for P phase, and red for S phase)
Movie S3. All detected events that occurred on August 27, 2018, along with their detected phase picks (blue for P phase, and red for S phase)
Movie S4. Distribution of relocated events from the initial machine learning catalog with 3D (latitude, longitude, depth) DBSCan clustering
Movie S5. Daily relocated event migration pattern with a cross-section of view.

Reference

Wiemer, S., & Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bulletin of the Seismological Society of America*, *90*(4), 859-869.