

1 In these supplementary information pages one figure and three tables are made available.

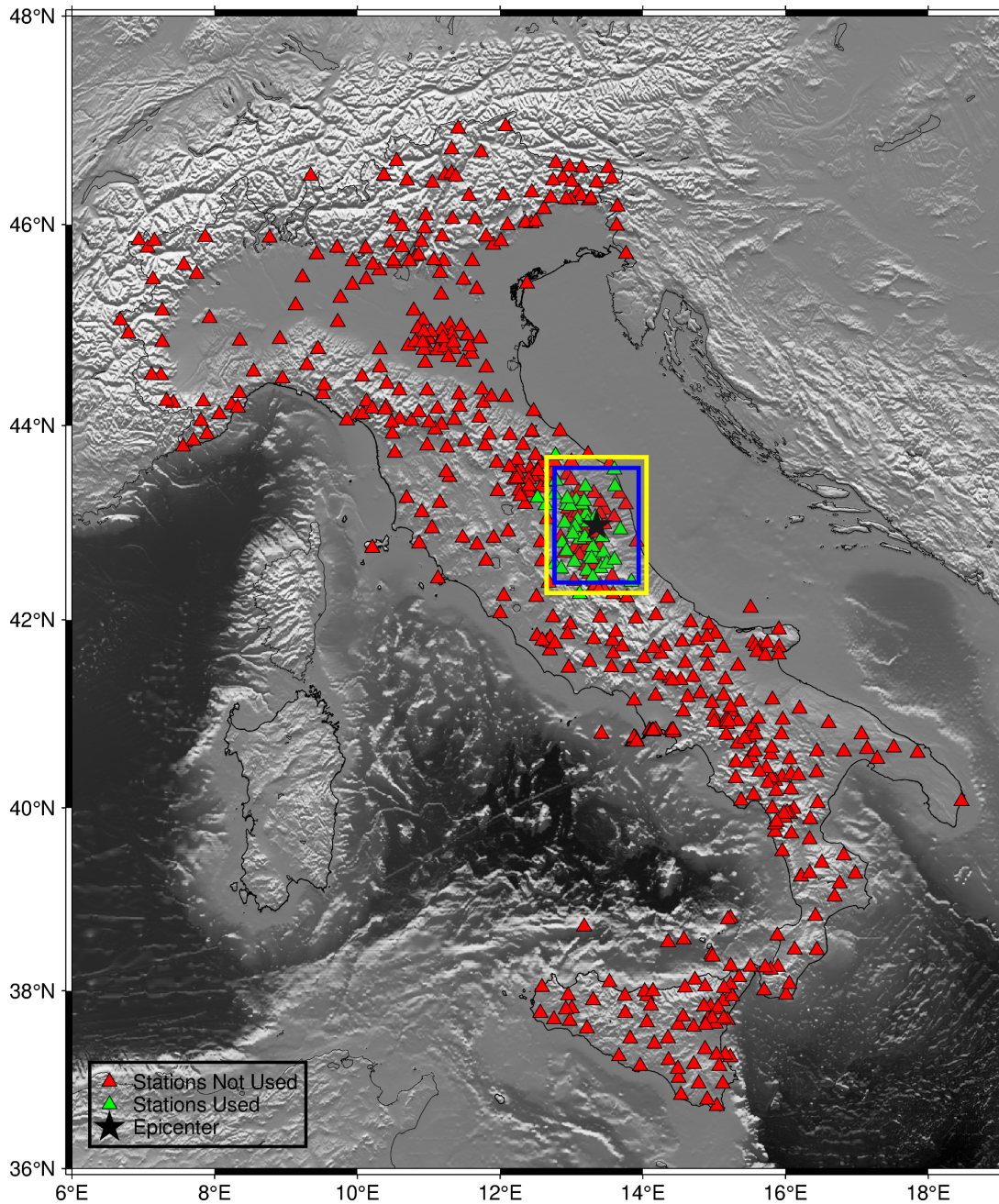


Figure S1 This map illustrates the application of the mask on the nearly 600 seismic stations from the INSTANCE benchmark dataset (Michellini et al., 2021). A sample earthquake from those used in this study was examined: 9068331. All green triangles denote stations that detected the P-wave arrival within 10 seconds for this event according to the INSTANCE metadata, and for this reason, they are all included by the mask during the training phase. Conversely, all red triangles indicate stations where the P-wave does not arrive within the first 10 seconds starting from the P arrival at the first recording station, and for this reason, they are excluded by the mask in the training phase. In the test phase, the stations included by the mask are represented exclusively by the green triangles within the blue bounding box, located approximately ± 0.58 degrees from the first recording station. This setup replicates an operational scenario in which the model estimates travel times assuming a P-wave velocity of 6.5 km/s based on the P arrival at the closest station provided in the INSTANCE metadata. The yellow bounding box defines the area used for metric calculations, set at approximately ± 0.7 degrees from the first recording station.

Experiment	A	B	C	PGA	PGV	SA(0.3)	SA(1.0)	SA(3.0)
1	no	no	yes	0.188	0.199	0.182	0.230	0.267
2	yes	yes	yes	0.205	0.210	0.195	0.243	0.280
3	yes	yes	no	0.202	0.209	0.194	0.241	0.276
4	no	yes	no	0.205	0.219	0.198	0.257	0.300
5	yes	no	no	0.197	0.209	0.192	0.243	0.280
6	no	yes	yes	0.189	0.200	0.181	0.229	0.268
7	no	no	no	0.198	0.207	0.190	0.238	0.274
8	yes	no	yes	0.208	0.218	0.201	0.254	0.290

Table S1 MSE results for the eight experiments on our data set. The letters A, B, and C denote the coordinates of the first station that records the P-wave, the interstation distances and the maximum amplitude information, respectively. These values are inserted as additional metadata into the flattened layer.

IM	ML (Masked GCN)			GMM (Bindi et al., 2011)		
	Median	Mean	STD	Median	Mean	STD
PGA	0.008	0.006	0.318	-0.037	-0.027	0.310
PGV	0.004	0.008	0.310	-0.240	-0.214	0.289
SA(0.3)	-0.019	-0.007	0.323	-0.292	-0.265	0.319
SA(1.0)	-0.024	-0.005	0.346	-0.259	-0.240	0.313
SA(3.0)	-0.034	-0.012	0.362	-0.037	-0.027	0.310

Table S2 IM residual statistics $\log_{10} \frac{IM_{obs}}{IM_{pred}}$ for the Masked GCN predictions and for the estimates obtained using the Bindi et al. (2011) GMM. The comparison was performed solely for the observed IMs at stations with available recorded data. The statistics are calculated after the outlier residuals (those with a residual $> |1|$) have been removed. The number of outliers of the GCN is 5.8% and for the GMM is 2.4%

Model	PGA		PGV		SA(0.3)		SA(1.0)		SA(3.0)	
	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
CNN	0.34	0.22	0.35	0.26	0.36	0.24	0.35	0.26	0.36	0.25
TISER-GCN	0.31	0.20	0.32	0.21	0.31	0.19	0.31	0.20	0.32	0.21
Masked TISER-GCN	0.32	0.19	0.32	0.20	0.32	0.18	0.35	0.23	0.35	0.27

Table S3 MAE and MSE results of each IM for the CNN model from Jozinovic et al. (2020), the TISER-GCN model from Bloemheugel et al. (2023) and our proposed model (Masked TISER-GCN).