Real Time Estimation of Earthquake Location and Magnitude Using Large Language Models

Aurora Bassani[1], Daniele Trappolini[1], Giulio Poggiali[1], Elisa Tinti[1], Fabio Galasso[1], Alberto Michelini[2], Chris Marone[1,3]
[1]Sapienza University of Rome [2]Istituto Nazionale di Geofisica e Vulcanologia [3]Pennsylvania State University

SAPIENZA UNIVERSITÀ DI ROMA
ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA
This presentation participates in OSPP — EGU — Outstanding Student & PhD candidate Presentation contest
EGU — Sharing is encouraged

CONTACTS

Code:https://github.com/AuroraBassani/LLMforGEOSCIENCE (private for now)
Email:aurora.bassani@uniroma1.it

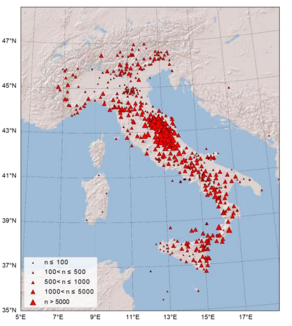
ABSTRACT PINLAB GITHUB

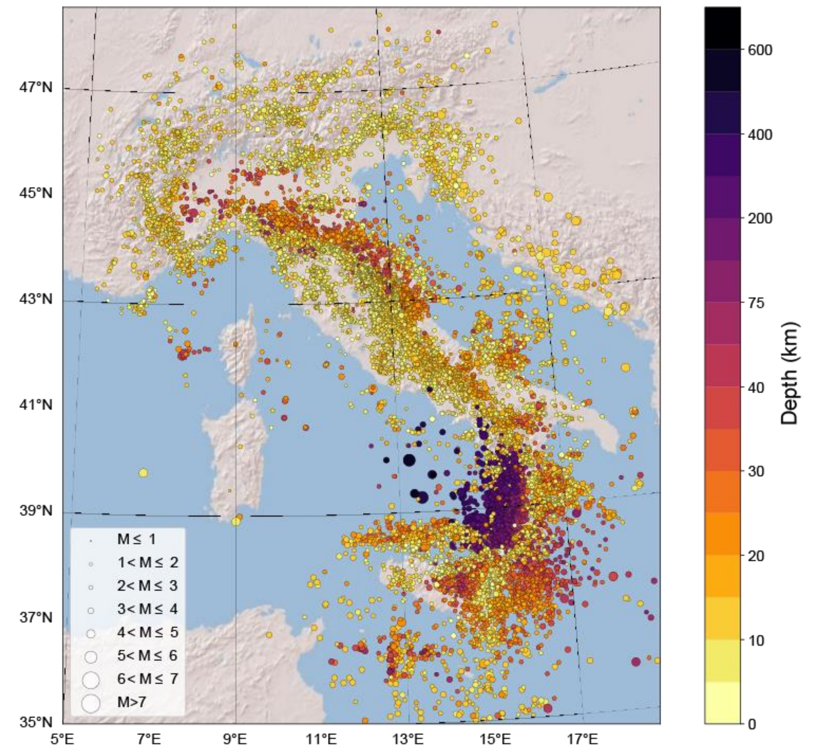
1. INTRODUCTION

Accurately estimating earthquake magnitude and location is crucial for seismic hazard assessment and real-time Earthquake Early Warning Systems (EEWS). While traditional and deep learning methods have advanced, they often lack real-time efficiency and generalization. Recently, studies like PromptCast, Time-LLM, and Chronos have shown promise in applying Large Language Models (LLM) to time series tasks; thus, we investigated whether an LLM could be used for earthquake parameter estimation.

2. INSTANCE DATASET

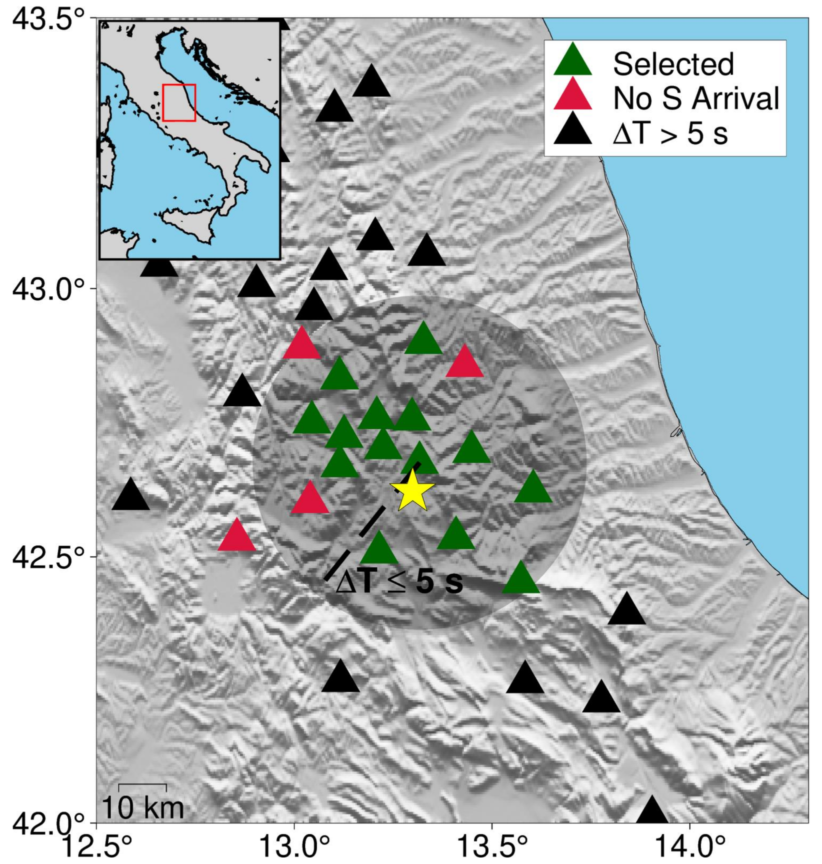
- 54,008 earthquakes
- 1,159,249 3-channel waveforms
- 19 networks
- 620 seismic station

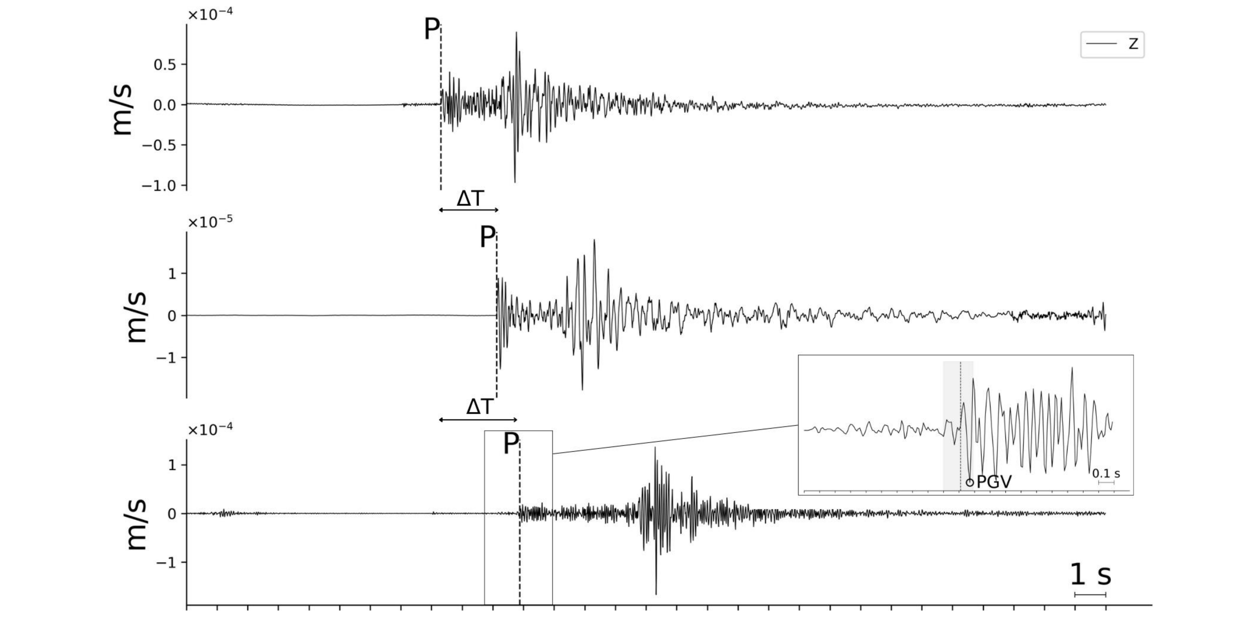




https://www.pi.ingv.it/banche-dati/instance/

3. SELECTION PROCESS





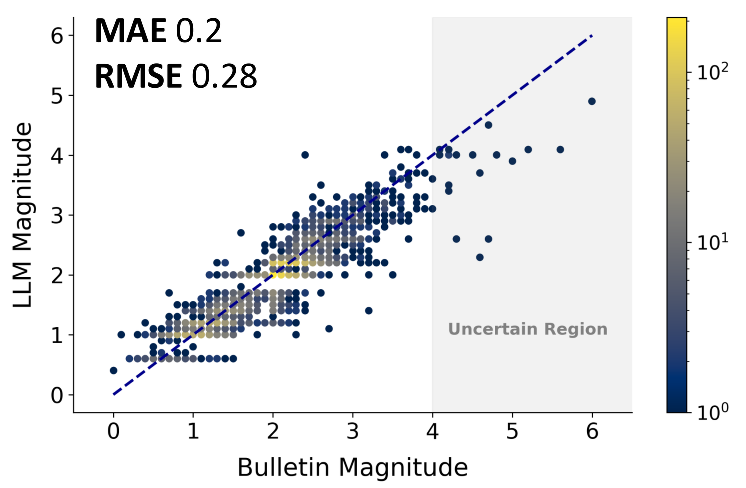
INSTANCE							
Magnitude Range	[0, 1)	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)
Number of Events	4462	15249	30845	3106	315	28	3

Final Dataset							
Magnitude Range	[0, 1)	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)
Number of Events	2562	9654	16941	1537	162	21	2

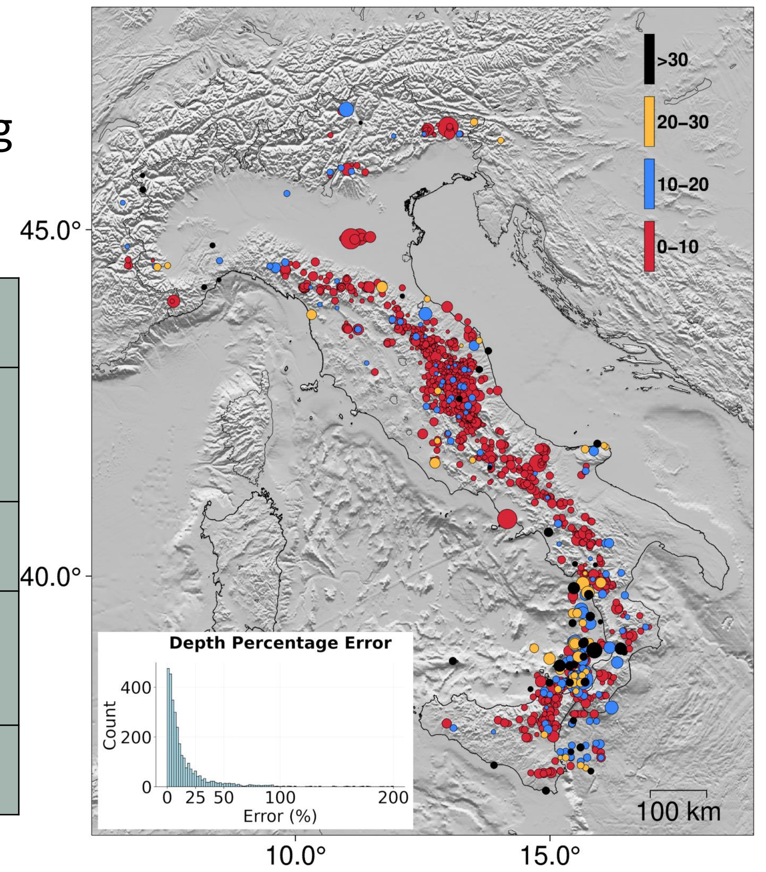
4. PROPOSED MODEL

5. RESULTS

We assessed the model performance on 3,094 events by comparing each parameter with the INGV bulletin values.



Mean Epicentral Error	4 km
Mean Hypocentral Error	6.8 km
MAE Depth	4.3 km
MAE First P Arrival Time	0.6 s
MAE Magnitude	0.2



CONCLUSION

- Our model using an LLM shows promising results for rapid and automated estimation of earthquake parameters.
- The promising outcomes from the comparison with INGV and the model's ability to process seismic data within a short time window pave the way for potential real-time applications.

6. TOWARD EEWS: COMPARISON OF LLM AND RAPID AUTOMATED RESULTS

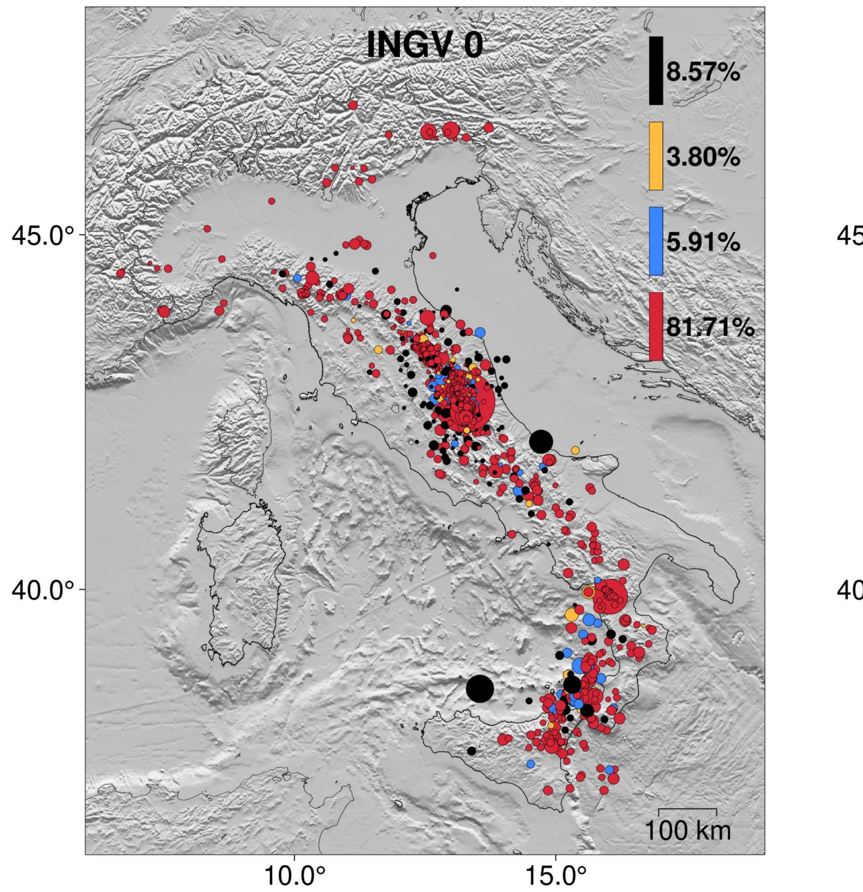
INGV Automatic Solutions

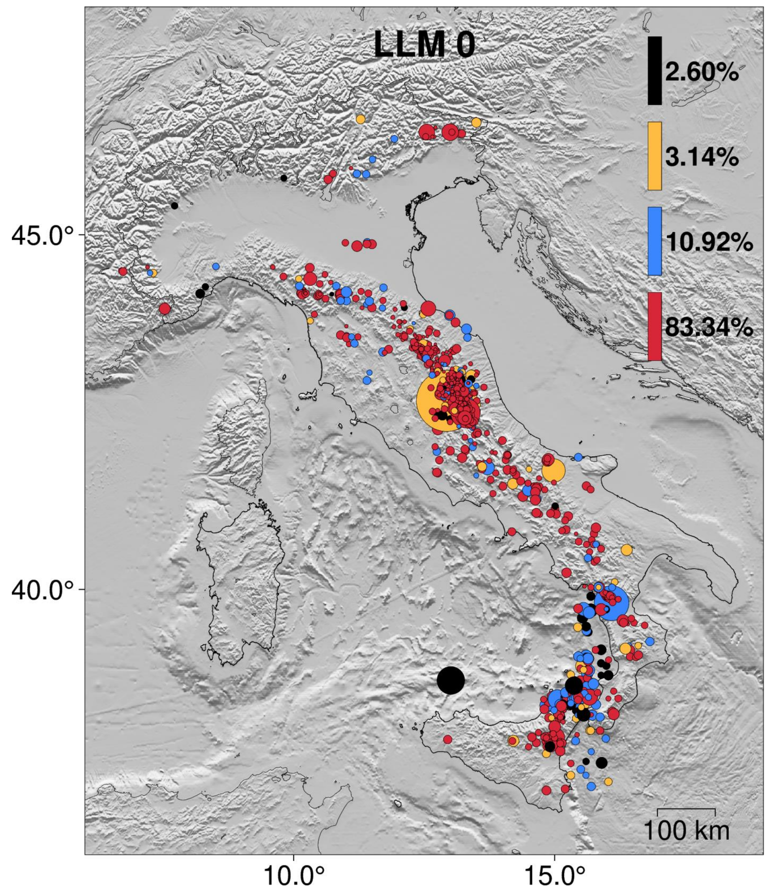
Solution '0' is generated as soon as at least 8 seismic phase arrivals are associated with an event. This means that a solution is typically available about 30 s after the origin time.

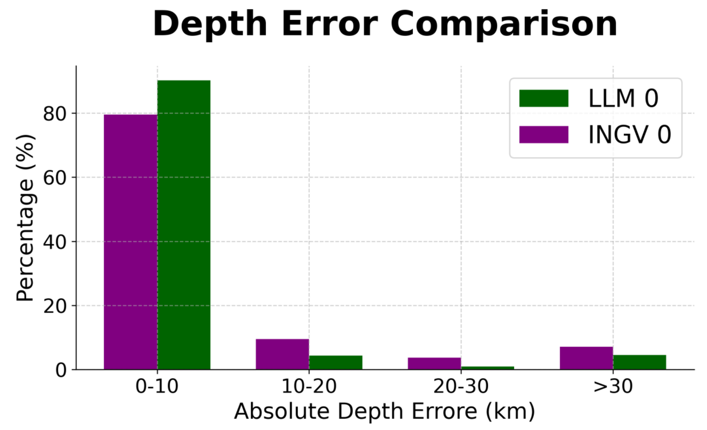
Solution '1' is provided when events have at least 30 phases identified, which typically occurs 1 minute after the earthquake.

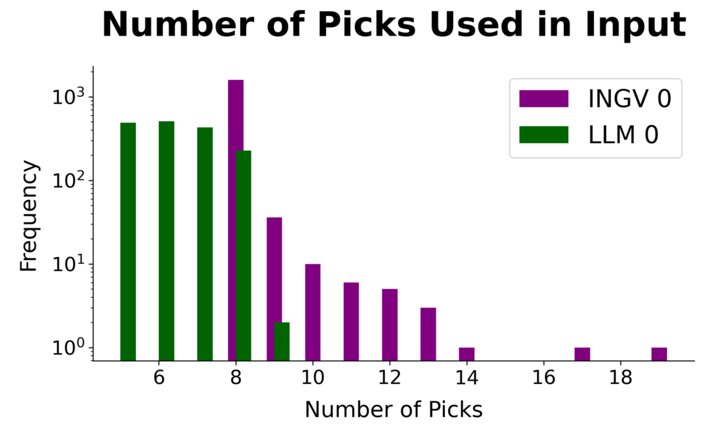
Solution '2' is available when there are at least 4 associated phases identified and the solution has been stable for the last 60 s; that is, when no additional phases have been associated within 60 s. This implies a waiting time of at most 5 to 6 minutes after the initial time.

	LLM			INGV		
	0	1	2	0	1	2
Number of Samples	1657	889	2241	1657	889	2241
Mean Epicentral Error (km)	6.26	5.61	5.29	8.62	9.84	10.68
Mean Hypocentral Error (km)	11.05	9.51	8.62	15.05	15.41	16.43
MAE Magnitude	0.30	0.32	0.25	0.25	0.14	0.17
MSE Magnitude	0.18	0.22	0.13	0.15	0.08	0.15
MAE First P Travel Time (s)	1.10	1.14	0.90	1.76	1.89	1.85









7. TESTING THE INFLUENCE OF INPUT PARAMETERS ON MODEL PERFORMANCE

Here we explore the influence of input parameters on model performance. On the left, a plot compares azimuthal gap, number of input stations, and bulletin depth with epicentral error. On the right, we address the long-standing question of whether earthquake size can be estimated from early rupture characteristics. For this analysis, we trained and tested the same model using the same data, but with PGVs computed on the horizontal component over a wider time window around the S-wave arrival (2 s instead of 0.2 s). The resulting performance difference is not significant, indicating that waiting for the S-wave information does not improve the estimate of magnitude.