Comparing Machine Learning to Manual Earthquake Location procedures: evaluating the performance of LOC-FLOW on a microseismic sequence occurred in Collalto area (NE Italy)

M. Sugan[1], L. Peruzza[1], M.A. Romano[1], M. Guidarelli[1], L. Moratto[1], D. Sandron[1], M.P. Plasencia Linares[1], M. Romanelli[1]

[1] *National Institute of Oceanography and Applied Geophysics - OGS, Italy*

Detecting earthquakes and picking seismic phases are fundamental elements in numerous seismological processes, being crucial in both seismic monitoring and in-depth seismological investigations. The utilization of machine learning (ML) techniques has experienced a notable improvement lately, offering a promising way to face the complexities associated with earthquake detection and localization.

The reliability of ML methods remains an open question in settings with dense and localized seismic networks. In such contexts, fast and accurate detection and localization of earthquakes are essential for decision-making, playing a pivotal role in seismic risk mitigation strategies, even for events of very low magnitude.

In the field of microseismic monitoring, ML applications are similar to those of earthquake monitoring, but have the task of processing weak seismic signals characterized by low signal-to-noise ratios at individual receivers or very short target time signals (Anikiev et al., 2023). Therefore, evaluating the performance of ML models trained on regional datasets in a microseismic sequence is challenging but crucial, especially for applications in the field of induced seismicity (e.g., Mousavi et al., 2016) or for activities of observatories near faults.

This study focuses on evaluating the performance of the PhaseNet algorithm (Zhu and Beroza, 2018), a prominent deep learning model for earthquake phase identification. The evaluation is conducted within the extensive LOC-FLOW workflow for earthquake location proposed by Zhang et al. (2022). The study is performed on the seismic events of the Refrontolo sequence that occurred in August 2021 on an antithetic fault segment of the Montello thrust system in the Pedemontana district of Southeastern Alps (Peruzza et al., 2022).

The seismic sequence displayed remarkable activity despite its low energy release (M_L 2.5 for the main event). The sequence consisted of 374 events occurring at approximately 9 km depth within a confined volume, and was monitored by the permanent Collalto Seismic Network (RSC). This sequence is a significant case study for testing and refining automated techniques to detect and locate microearthquakes using machine learning.

The RSC is composed of 10 stations and has monitored microseismic activity potentially induced by underground gas storage activities since 2012. Rigorous manual processing conducted by the RSC, involving daily and monthly offline procedures, is undertaken to guarantee data accuracy and metadata reliability. Nevertheless, this approach can be time consuming and demanding, particularly in densely populated seismic sequences where very fast analysis is preferable.

Comparing seismic catalogs derived from associations approved by experienced analysts and revised manual picks with those generated using the PhaseNet integrated with LOC-FLOW provides a unique opportunity to evaluate the performance of ML methods in detecting and localizing local microearthquakes. Our particular focus includes:

1. PhaseNet phase picker performance: we examine the effectiveness of the PhaseNet phase picker in comparison to manual phase picks. This evaluation aims to assess the accuracy and reliability of PhaseNet's automated phase detection.

2. LOC-FLOW-generated earthquake catalogs: we analyse the earthquake catalogs produced by LOC-FLOW, evaluating both origin time and absolute locations. The examination involves comparing catalogs formed with PhaseNet picks against those created with the original RSC manual picks. This analysis offers insights into the coherence and efficacy of PhaseNet in creating the overall earthquake catalog.

3. Contribution of template matching: we assess the influence of template matching on the final catalog and compare it to the dataset obtained from the original RSC procedures. This evaluation aims to clarify the degree to which template matching enhances the accuracy and comprehensiveness of the earthquake catalog within the LOC-FLOW workflow.

4. Spatio-temporal characteristics of seismicity: furthermore, we assess the spatio-temporal characteristics of the acquired seismicity. This examination is useful to check the efficiency of the method in discerning tectonic structures activated during the sequence. Gaining insights into the spatial and temporal patterns of seismic activity offers valuable understanding of the underlying geologic processes.

We find that PhaseNet achieved a detection rate of 79% for manual P arrival times and 90% for S arrival times at the same stations. While P picks exhibited satisfactory accuracy, a noticeable delay was observed for S picks. This delay is presumed to be a common feature, even in other datasets, given the high quality of the manual picking used for comparison.

After integrating events identified by the template matching procedure, the final LOC-FLOW catalog is characterized by an increased number of events compared to the initial manual catalog. However, in our specific case study, PhaseNet did not contribute significantly to the augmentation of the earthquake count during the most active days of the sequence (e.g., 2-3 August), where template matching played a crucial role. Despite the observed lower accuracy in S picks, PhaseNet's overall performance is commendable, especially when considering the time of processing, significantly reduced compared to manual picking. The seismicity pattern observed vividly depicted the geometry of the activated fault in both temporal and spatial dimensions (Sugan et al., 2023).

References

Anikiev D., Birnie C., bin Waheed U., Alkhalifah T., Gu C., Verschuur D.J., Eisner L.; 2023. Machine learning in microseismic monitoring. Earth-Science Reviews. 239:104371. https://doi.org/10.1016/j.earscirev.2023.104371.

Mousavi S.M., Horton S.P., Langston C.A., Samei B.; 2016. Seismic Features and Automatic Discrimination of Deep and Shallow Induced-Microearthquakes Using Neural Network and Logistic Regression. Geophys J Int. 207(1):29-46. https://doi.org/10.1093/gji/ggw258.

Peruzza L., Romano M.A., Guidarelli M., Moratto L., Garbin M., Priolo E.; 2022. An unusually productive microearthquake sequence brings new insights to the buried active thrust system of Montello (Southeastern Alps, Northern Italy). Front Earth Sci. 10:1044296. https://doi.org/10.3389/feart.2022.1044296.

Sugan M., Peruzza L., Romano M.A., Guidarelli M., Moratto L., Sandron D., Plasencia Linares M.P., Romanelli M.; 2023. Machine learning versus manual earthquake location workflow: testing LOC-FLOW on an unusually productive microseismic sequence in northeastern Italy. Geomatics, Natural Hazards and Risk, 14:1. https://doi.org/10.1080/19475705.2023.2284120.

Zhang M., Liu M., Feng T., Wang R., Zhu W.; 2022. LOC-FLOW: An End-to-End Machine-Learning-Based High-Precision Earthquake Location Workflow. Seismol Res Lett. 93(5):2426-2438. https://doi.org/10.1785/0220220019.

Zhu W., Beroza G.C.; 2018. PhaseNet: A deep-neural-network based seismic arrival-time picking method. Geophys J Int. 216(1): 261–273. https://doi.org/10.1093/gji/ggy423.

Corresponding author: msugan@ogs.it