

MACHINE-LEARNING-BASED PHASE PICKER: ANALYZING THE TEMPORAL AND SPATIAL CHANGES OF THE OCTOBER 2019 COTABATO AND DECEMBER 2019 DAVAO DEL SUR EARTHQUAKES

Paulo Pangan SAWI¹
MEE22704

Supervisor: Saeko KITA^{2*}

ABSTRACT

A deep-neural-network-based phase picker, PhaseNet, was used to pick the arrival times of the P and S waves during the earthquake sequence in Cotabato and Davao del Sur, Philippines, which occurred from October to December 2019, involving five M~6 inland earthquakes with magnitudes MW 6.4, 6.6, 5.9, 6.5 and 6.7. In this study, we utilized 80 days of seismic data from stations located within 200 km of the event area and input them into PhaseNet for analysis.

The phase picks, the output of PhaseNet, were first associated and initially located using Rapid Earthquake Association and Location (REAL). Subsequently, the earthquakes were relocated using VELEST. The hypocenters were further refined using the relative location method called HypoDD.

Using these methods, we successfully created an earthquake catalog comprising 5,017 earthquakes, which is more than those on the list by the Department of Science and Technology – Philippine Institute of Volcanology and Seismology (DOST-PHIVOLCS) on their website. This catalog reveals the spatial and temporal changes in seismicity following each significant event. It also uncovers detailed patterns in aftershock clustering, which are likely linked to complex fault system structures that may have contributed to the seismic activity.

Keywords: Machine learning, Hypocenter relocation, Inland earthquakes, Philippines, Southeast Asia.

1. INTRODUCTION

Cotabato and Davao del Sur experienced significant earthquakes in October and December 2019. The initial earthquake occurred on 16 October 2019, registering a moment magnitude (M_w) of 6.4 with strike-slip faulting. Approximately 13 days later, a more substantial earthquake struck on 29 October 2019 at 09:04 AM, with a magnitude of M_w 6.6, as per Philippine Standard Time (PST). Following this event, another earthquake occurred less than two hours later at 10:42 AM (PST), with a magnitude of M_w 5.9. Another notable earthquake occurred on 31 October 2019, registering a magnitude of M_w 6.5. All of these earthquakes exhibited strike-slip faulting based on the SWIFT-CMT analysis.

The aftershock distribution and focal mechanisms indicate that the source belongs to the Cotabato fault system (Perez et al., 2019). Also, on 15 December 2019, Davao del Sur was hit by M_w

¹ Department of Science and Technology – Philippine Institute of Volcanology and Seismology (DOST-PHIVOLCS), Philippines.

² International Institute of Seismology and Earthquake Engineering, Building Research Institute.

* Chief examiner

6.7 with strike-slip faulting. According to the report from the DOST-PHIVOLCS Quick Response Team (Perez et al., 2020), the Makilala-Malungon Fault is recognized as the fault associated with this event.

Table 1: Earthquake Parameters for the $M > \sim 6$ in Cotabato and Davao del Sur from October and December 2019 by DOST-PHIVOLCS.

Date- Time (PST)	Magnitude	Depth (km)	Location	Highest Intensity (PEIS)
16 October 2019 – 07:37 PM (EQ1)	6.3(M_S)/6.4(M_W)	09	06.76°N, 125.01°E	VII
29 October 2019 –09:04 AM (EQ2a)	6.6(M_S)/6.6(M_W)	07	06.81°N,125.03°E	VII
29 October 2019 –10:42 AM (EQ2b)	6.1(M_S)/5.9(M_W)	11	06.85°N,125.02°E	VI
31 October 2019 – 09:11 AM (EQ3)	6.5(M_S)/6.5(M_W)	08	06.92°N,125.06°E	VII
15 December 2019–02:11 PM (EQ4)	6.9(M_S)/6.7(M_W)	09	06.76°N,125.13°E	VII

In the last twenty years, there has been a significant advancement in the field of machine learning, which is commonly and widely used in image processing, voice recognition, language understanding, and many other applications (Jordan & Mitchell, 2015). Another application of machine learning in the field of seismology that uses deep neural networks (DNN) is the PhaseNet (Zhu & Beroza, 2019). PhaseNet is an automatic phase picker that picks the arrival times of P and S waves (Zhu & Beroza, 2019). PhaseNet was trained on a dataset encompassing over 30 years of seismic data from the Northern California Earthquake Data Center, which has manually picked P and S wave arrival times by skilled analysts. By leveraging the learned features from this training data, PhaseNet can estimate the arrival times of P and S waves for new data by providing probability-based predictions (Zhu & Beroza, 2019). Many researchers used PhaseNet to make a high-precision earthquake catalog. They used these high-precision catalogs to provide a detailed study and insight into an area's aftershock activity, earthquake sequence, or series.

This research aims to develop a precise catalog that captures the temporal and spatial variations of the October 2019 Cotabato earthquakes and the December 2019 Davao Del Sur earthquake. To achieve this, a machine-learning-based phase picker, PhaseNet, will be used in this study. Additionally, earthquake association and relocation techniques will be utilized to enhance the accuracy and reliability of the created catalog. The catalog generated in this study has the potential to provide valuable insights into the fault structure within the studied area. Moreover, the findings and methodologies presented here can be applied to analyze other earthquake sequences and series that have occurred in the Philippines. This study also paves the way for the utilization of machine learning techniques, specifically in phase picking, within the context of earthquake research in the Philippines.

2. DATA

This study focuses on analyzing the dataset derived from the 2019 Cotabato earthquake sequence, spanning from October 16 to December 31, 2019. The dataset encompasses a total of 80 days of earthquake waveforms. The dataset consists of seismic recordings obtained from the Philippine Seismic Network (PSN) operated by DOST-PHIVOLCS. Only seismic stations within a 200-kilometer (km) radius were selected for this study. This selection process included 18 seismic stations, comprising 14 broadband stations and four short-period stations.

Each station's data consists of three components: a North-South component, an East-West component, and an Up-down or horizontal component. The data were resampled to a uniform sampling rate of 100 Hz and subsequently divided into 30-second time windows to facilitate further analysis using PhaseNet.

3. METHODOLOGY

The methodology of this study consists of the following steps: (1) Phase picking, (2) Earthquake Association, (3) Absolute location, and (4) Relative location.

For phase picking, the P-wave and S-wave arrival times were automatically determined using a deep-neural network-based phase picker called PhaseNet (Zhu & Beroza, 2019). PhaseNet utilizes the three components of a seismogram to determine the arrival times of P and S waves, along with their corresponding probabilities (Park et al., 2020).

After applying PhaseNet, the subsequent step involves associating these phase picks with an earthquake using Rapid Earthquake Association and Location (REAL) (Zhang et al., 2019). REAL associates the P and S arrival times and determines earthquake location by counting the number of observed picks (P and S arrival times) and calculates the travel-time residual, and the grid point with the highest number of picks and the smallest travel-time residual will be selected (Zhang et al., 2019). The IASPEI 91 reference model was employed as the velocity model to generate a theoretical travel timetable using TauP Toolkit (Crotwell et al., 1999). A threshold was established for the number of picks, requiring each event to include a minimum of four P picks and at least one S pick.

Upon associating the phase picks and acquiring the initial earthquake locations using REAL, these earthquakes were relocated using VELEST (Kissling et al., 1994). VELEST was utilized twice in the process. In the first iteration, the earthquake locations obtained from REAL were used to derive a 1D velocity model by utilizing Crust 1.0 (Laske et al., 2013) as the initial velocity model. After obtaining the updated velocity model, all the earthquakes located by REAL were used as input and processed in VELEST, this time utilizing the updated velocity model. To further refine the earthquake locations, HypoDD (Waldhauser & Ellsworth, 2000) was employed to find the relative locations. HypoDD utilizes the double-difference concept to relocate a cluster of earthquakes by minimizing the disparity between the observed travel-time differences and the calculated travel-time differences between the event pair and station (Zhang et al., 2022). The updated velocity model obtained from VELEST was also utilized in the HypoDD for improved accuracy and consistency throughout the refinement process.

4. RESULTS AND DISCUSSION

4.1. Hypocenter determination

PhaseNet obtained 623,582 phase picks with a probability of 0.4 or higher using 80 days of data, wherein 442,065 corresponded to P-phases and 181,517 picks corresponded to S-phases. After the final analysis step (HypoDD), 5,017 earthquakes were obtained. This earthquake catalog captures a significantly higher number of seismic events (~69% increase), and the hypocenters are more clustered than the original earthquake count available on the DOST-PHIVOLCS website within that time range (Figure 1).

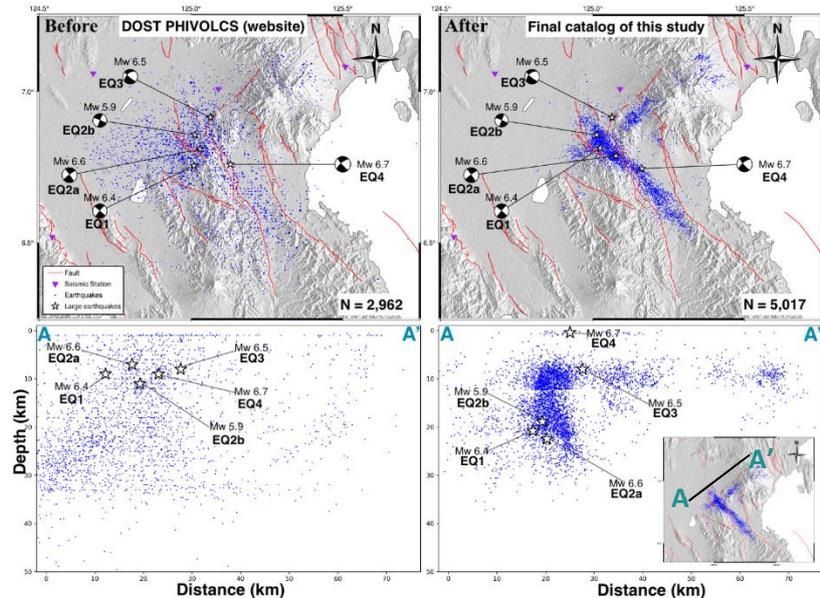


Figure 1. Distribution of hypocenters. Routine catalog by DOST-PHIVOLCS (website) vs. final catalog of this study after applying HypoDD.

4.2. Updated fault structure

Following the EQ1 (as depicted in Figure 2-a), the distribution pattern of aftershocks displays an NW-SE, corresponding to the nodal plane of the left-lateral focal mechanism. This alignment is thought to be the fault plane of the EQ1. Additionally, a secondary trend is discernible on the left side of the primary trend, indicating a NE-SW pattern. These trends suggest the possibility of two intersecting faults: one with an NW-SE trend and another with a NE-SW trend. In Figure 3-a, cross-section, the earthquakes (represented by red circles) after EQ1 display a pattern directed towards the southwest, indicating an inclination of the fault in that direction.

Following the occurrence of the EQ2a and EQ2b earthquakes, a consistent NW-SE becomes evident in the seismicity trend, resembling the pattern observed after the EQ1 (Figure 2-b). However, no visible secondary trend is observed on the left side. Instead, a cluster of earthquakes is apparent on the right side, possibly linked to the EQ3. The cross-sectional analysis (Figure 3-b), represented by cyan color circles, further reveals a similar orientation observed after the EQ1.

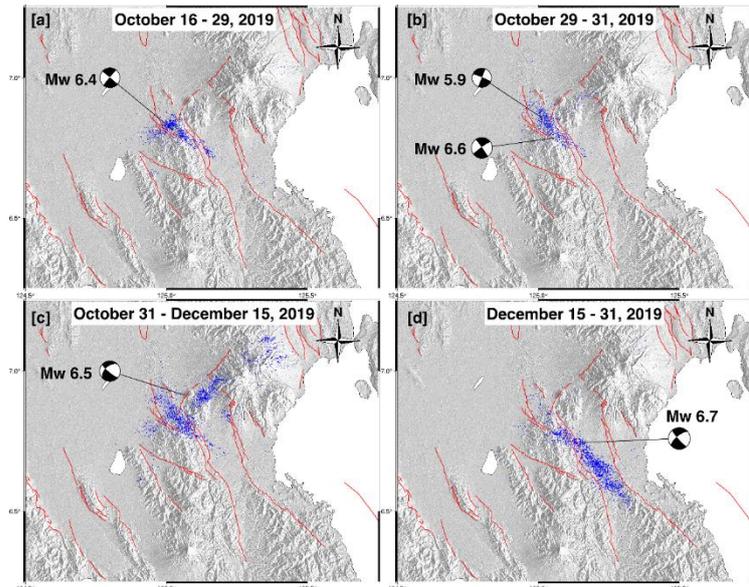


Figure 2. Spatial and temporal changes following the occurrence of a significant earthquake. The beach ball represents the location of a large earthquake.

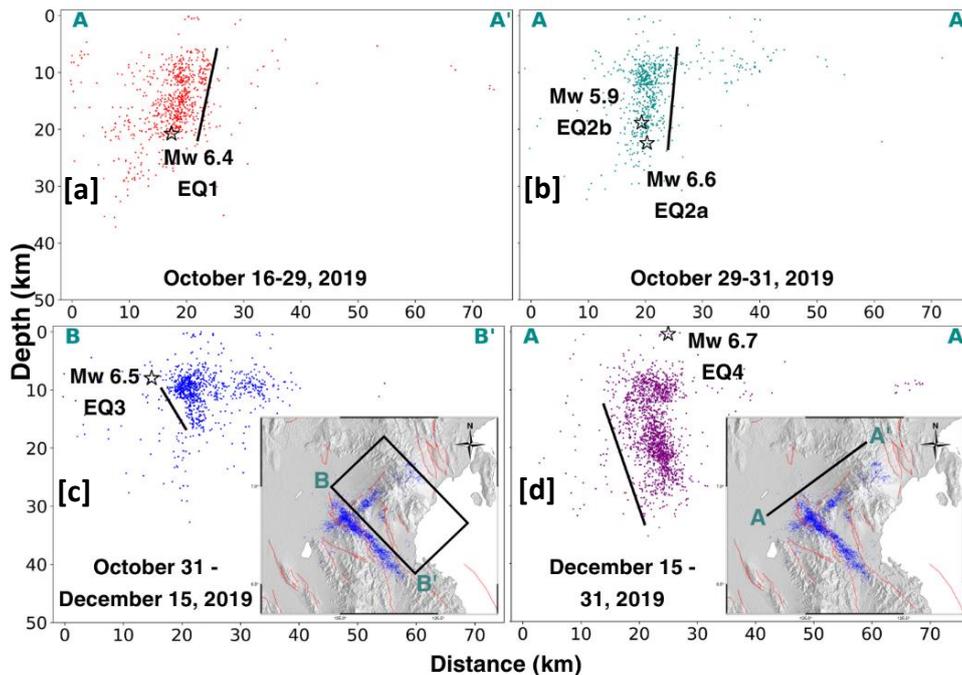


Figure 3. Spatial and temporal changes in a cross-section view following the occurrence of EQ 1 (red circles), EQ 2 a-b (cyan circles), EQ3 (blue circles), and EQ4 (purple circles).

The EQ3 exhibits an interesting characteristic where foreshocks are already evident after the EQ2a and EQ2b events. The hypocenters of these earthquakes coincide with those after EQ3 (Figure 2-c), suggesting a strong association between these seismic events. A distinct NE-SW trend emerges in Figure 2-c, potentially indicating the fault responsible for generating the EQ3.

In the case of EQ4, a consistent NW-SE trend of earthquakes can be observed (Figure 2-d), similar to the main trend observed after EQ1, EQ2a, and EQ2b. Although it has a similar trend in other earthquakes, aftershocks of the EQ4 extended towards the southeastern region and a completely different trend of the aftershock's alignment compared to the EQ3. Moreover, the cross-sectional analysis in Figure 3-d reveals a distinct pattern in the earthquakes after the EQ4, indicating a trend towards the east or northeast, different from the other EQ1 and EQ2a-b. This observation suggests that the fault responsible for these earthquakes may be inclined in the northeast direction.

The seismicity extracted in this study exhibits a distinct distribution pattern characterized by alignments along cross-faults, known as conjugate faults. These phenomena have been documented in other regions, such as southern California. Notable instances include the 1987 Superstition Hills earthquake sequences (Hudnut et al., 1989) and the 2019 M7.1 Ridgecrest earthquake (Milliner et al., 2022), all displaying strike-slip faulting in their focal mechanism.

5. CONCLUSIONS

This study used a machine learning-based phase picker, called PhaseNet, to analyze the October Cotabato and December 2019 Davao del Sur earthquakes. The analysis spanned 80 days, from October 16 to December 31, 2019. The phase picks obtained from PhaseNet were associated and initially located using REAL. VELEST was employed to refine the hypocenter locations. Finally, HypoDD was used for the relocation.

Using the abovementioned methods, we created an earthquake catalog that contains more earthquakes. This catalog encompasses over 69% of the earthquakes available on the DOST-PHIVOLCS website and reveals the spatial and temporal changes that occur after each significant earthquake, showing the pattern of the cluster of earthquakes associated with the fault structures that may have contributed to these seismic events.

While the obtained results demonstrated significant improvements in seismicity analysis and hypocenter quality, further trials are needed to determine the optimal parameters for each method. Conducting additional experiments with different parameter settings can enhance the reliability and accuracy of the seismic analyses. Incorporating and comparing a local velocity model with various velocity models can also yield valuable insights. These comparisons will help us better understand local seismic characteristics and help refine seismic imaging and locating processes.

In conclusion, this study demonstrated a successful application of ML techniques for phase picking and other relevant approaches. These methods proved to be particularly valuable for analyzing the aftershocks, series, and sequences of earthquakes, offering insights into fault structures and their spatiotemporal evolution.

ACKNOWLEDGEMENTS

This research was conducted during the individual study period of the training course "Seismology, Earthquake Engineering, and Tsunami Disaster Mitigation" by the International Institute of Seismology and Earthquake Engineering – Building Research Institute (IISEE-BRI), Japan International Cooperation Agency (JICA), and National Graduate Institute for Policy Studies (GRIPS). I want to express my sincere gratitude to my supervisor Dr. Saeko Kita for her unwavering guidance and support throughout this research period. I would also like to thank the Department of Science and Technology – Philippine Institute of Volcanology and Seismology (DOST-PHIVOLCS) for giving me this great opportunity.

REFERENCES

- Crotwell, H. P., Owens, T. J., & Ritsema, J. (1999). The TauP Toolkit: Flexible seismic travel-time and ray-path utilities. *Seismological Research Letters*, 70, 154-160.
- Hudnut, K. W., Seeber, L., & Pacheco, J. (1989). Cross - fault triggering in the November 1987 Superstition Hills earthquake sequence, southern California. *Geophysical Research Letters*, 16(2), 199-202.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Kissling, E., Ellsworth, W., Eberhart - Phillips, D., & Kradolfer, U. (1994). Initial reference models in local earthquake tomography. *Journal of Geophysical Research: Solid Earth*, 99(B10), 19635-19646.
- Laske, G., Masters, G., Ma, Z., & Pasyanos, M. (2013). Update on CRUST1. 0—A 1-degree global model of Earth's crust. *Geophysical research abstracts*, vol. 15, no. 15, p. 2658.
- Milliner, C., Aati, S., & Avouac, J.-P. (2022). Fault friction derived from fault bend influence on coseismic slip during the 2019 Ridgecrest Mw 7.1 mainshock. *Journal of Geophysical Research: Solid Earth*, 127(11), e2022JB024519.
- Park, Y., Mousavi, S. M., Zhu, W., Ellsworth, W. L., & Beroza, G. C. (2020). Machine - learning - based analysis of the Guy - Greenbrier, Arkansas earthquakes: A tale of two sequences. *Geophysical Research Letters*, 47(6), e2020GL087032.
- Perez, J., Abigania, M. I., Abitang, M., Amilbahar, R., Bariso, E., Daquipa, H., . . . Veracruz, N. (2019). The October 2019 Cotabato earthquake sequence: Parameters and impacts Conference: 2019 Geological Society of the Philippines Geological Convention (2019 GEOCON),
- Perez, J., Amilbahar, R., Baltazar, H., Bon, D., Buhay, D. J., Dizon, M., . . . Ysulan, A. (2020). Fault Geometry and Geologic Impacts of the 15 December 2019 Magnitude 6.9 Davao del Sur Earthquake Virtual Geological Convention GEOCON 2020 (75 Years of Geosciences in the Philippines: Opportunities, Advances, and Frontiers),
- Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bulletin of the seismological society of America*, 90(6), 1353-1368.
- Zhang, M., Ellsworth, W. L., & Beroza, G. C. (2019). Rapid earthquake association and location. *Seismological Research Letters*, 90(6), 2276-2284.
- 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. *Seismological Society of America*, 93(5), 2426-2438.
- Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1), 261-273.