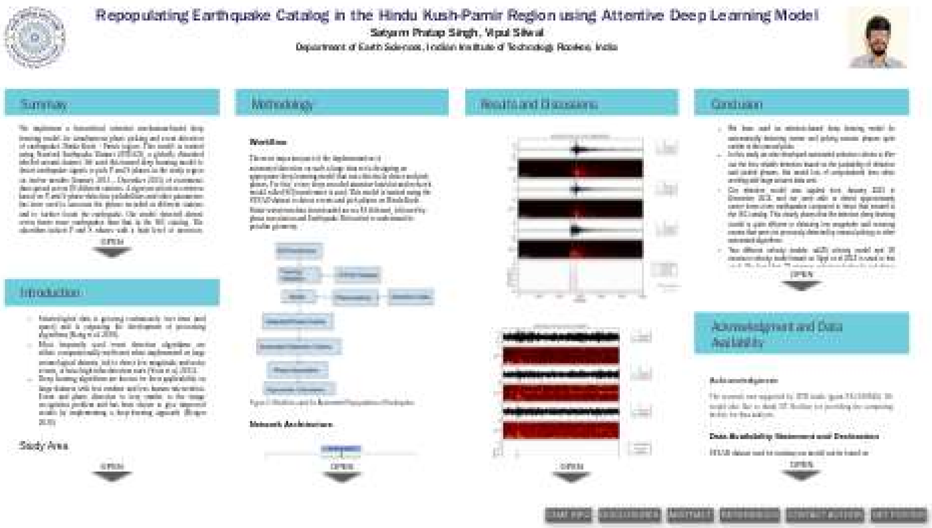
Repopulating Earthquake Catalog in the Hindu Kush-Pamir Region using Attentive Deep Learning Model



Satyam Pratap Singh, Vipul Silwal

Department of Earth Sciences, Indian Institute of Technology Roorkee, India

PRESENTED AT:

SUMMARY

We implement a hierarchical attention mechanism-based deep learning model for simultaneous phase picking and event detection of earthquakes Hindu Kush - Pamir region. This model is trained using Stanford Earthquake Dataset (STEAD), a globally disturbed labeled seismic dataset. We used this trained deep learning model to detect earthquake signals to pick P and S phases in the study region on twelve months (January 2013 – December 2013) of continuous data spread across 83 different stations. A rigorous selection criterion based on P and S phase-detection probabilities and other parameters has been used to associate the phases recorded at different stations and to further locate the earthquake. Our model detected almost seven times more earthquakes than that in the ISC catalog. The algorithm picked P and S phases with a high level of precision, comparable to human analyst picks. Furthermore, pinpointing these events allowed us to define the S-shaped seismic zone in the Pamir-Hindu Kush region and better comprehend the deformation caused by Eurasian- Indian plate motion.

INTRODUCTION

- Seismological data is growing continuously over time (and space) and is outpacing the development of processing algorithms (Kong et al. 2018).
- Most frequently used event detection algorithms are either: computationally inefficient when implemented on large seismological datasets, fail to detect low magnitude and noisy events, or have high false detection rates (Yoon et al, 2015).
- Deep learning algorithms are known for their applicability on large datasets with less runtime and less human intervention. Event and phase detection is very similar to the image recognition problem and has been shown to give improved results by implementing a deep-learning approach (Bergen 2019).

Study Area

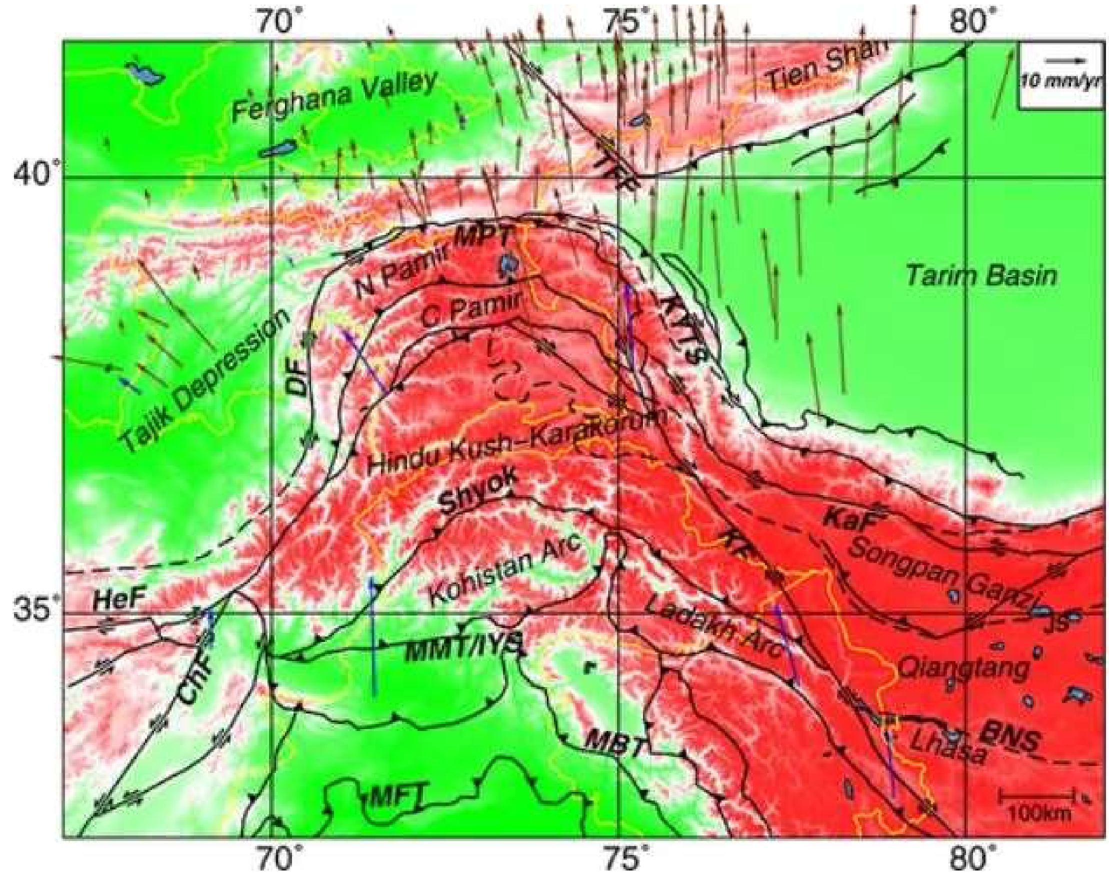


Fig 1: Map of the study region, showing GPS velocities and tectonic features

.

- Hindu Kush-Pamir is part of the ongoing convergence between the Indian and Eurasian plates and is one of the most seismically active regions in the world.
- The seismicity pattern of the region is quite different from other parts of the world where a continent-continent collision occurs, as it has frequent intermediate-to-deep earthquakes (Pegler and Das, 1998; Sippl et al., 2013).
- In the past two decades, there has been an enormous increase in recorded waveform data with numerous continuous seismic station networks being distributed throughout the region but many of the low magnitude and signal-to-noise ratio earthquakes remain undetected.
- Hence automatic detecting and locating earthquakes can significantly improve the understanding of the regional tectonics therefore the peculiar geometry of the Hindu Kush Pamir.

METHODOLOGY

Workflow

The most important part of the implementation of automated detection on such a large data set is designing an appropriate deep learning model that can effectively detect and pick phases. For this, a very deep encoded attention-based neural network model called EQ transformer is used. This model is trained using the STEAD dataset to detect events and pick phases on Hindu Kush Pamir waveform data downloaded across 83 different, followed by phase association and Earthquake Relocation to understand its peculiar geometry.

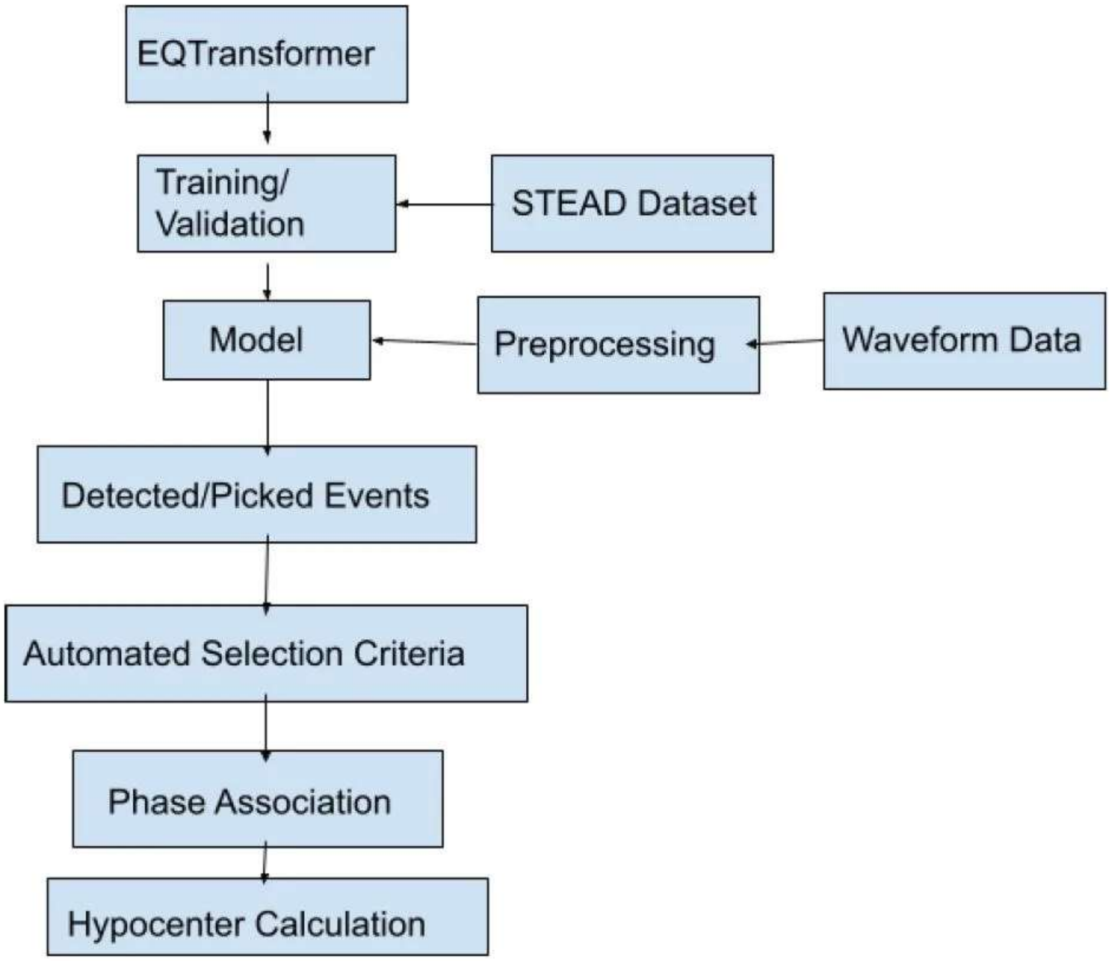


Figure 2: Workflow used for Automated Repopulation of Earthquake.

Network Architecture

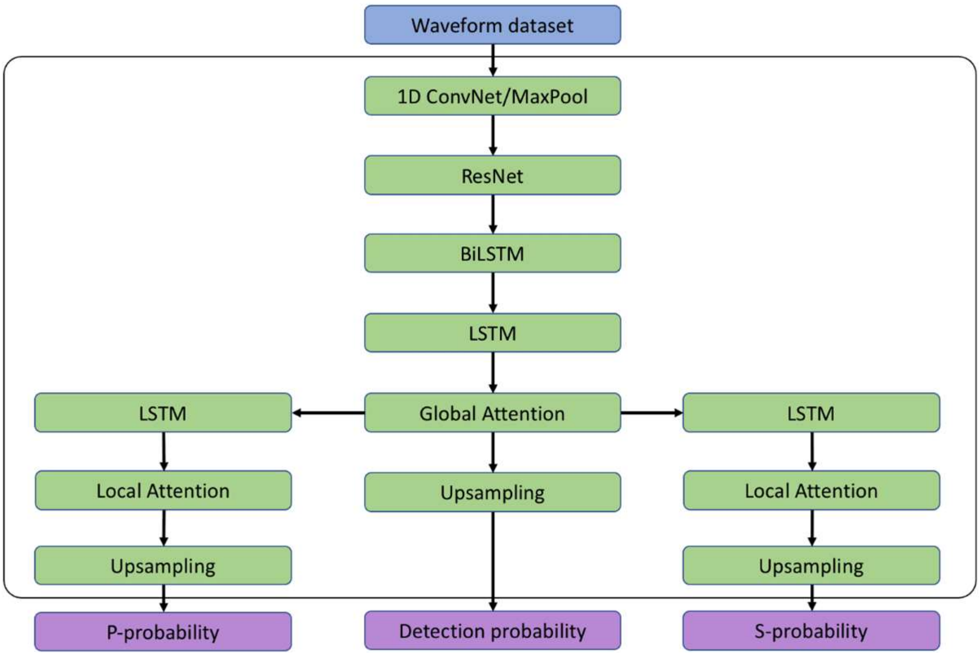


Figure 3: Summary of Network architecture

- All the deep learning detection/picking models are based on high-level representation learning containing certain common features of the earthquake waveform and seismic phase.
- This feature is implemented using an attention-based mechanism in the model (Mousavi et al., 2020). The model contains two levels of attention mechanisms: one at the global level to recognize events from the dataset and one at the local level to detect the P and S Phases.
- The structure of this deep learning model is multi-task. It is made up of one relatively deeper encoder and three different decoders that use 1D convolutions, residual connections, unidirectional and bidirectional long-short-term memory (LSTM), Network in Network, feed-forward layer, and self-attentive model.
- The encoder takes a time-domain seismic signal and extracts a high-level representation as well as contextual information. The three decoders are then utilized to map this high-level data into detection probability and phase selection.

Dataset

To train our network, we used the Stanford Earthquake Dataset (STEAD). STEAD is a global dataset of earthquakes and non-earthquake signals that has been labeled at a large scale (Mousavi et al., 2019).

Twelve months of continuous data from 1st January 2013 to 31st December 2013 were downloaded from IRIS and GEOFON data centers. This consists of 83 different stations spread across the whole region. Using our trained model detection has been at all these different stations.

Automated Selection Criteria

- A simple criterion is used to select a valid detection based on the probability of the detection and phase picks. Kernel Density Estimation is plotted for all the detections made by our neural network.
- It shows how sure our model is in picking and detection. This distribution is then used to select a threshold value for detection probability.
- Such criteria solve the problem to filter out many false-positive events (detections with low probability) from the waveform data and hence save the computation time for phase association.

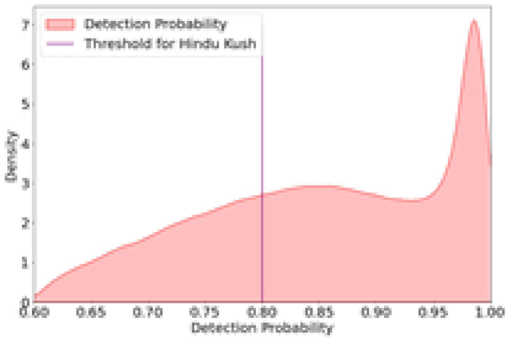
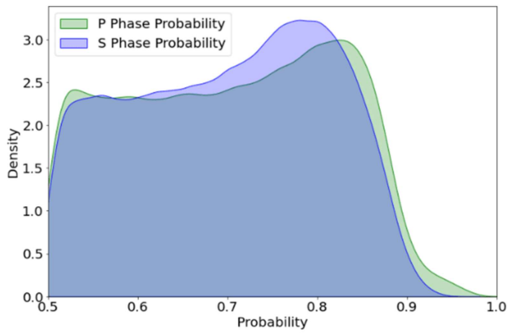


Figure 4: KDE for Detection and P, S phase pick probabilities of Hindu Kush- Pamir



Earthquake Location

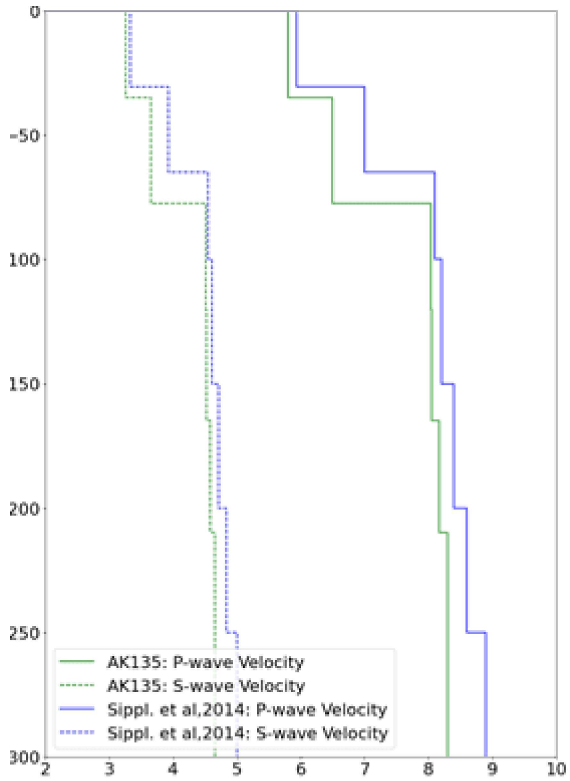
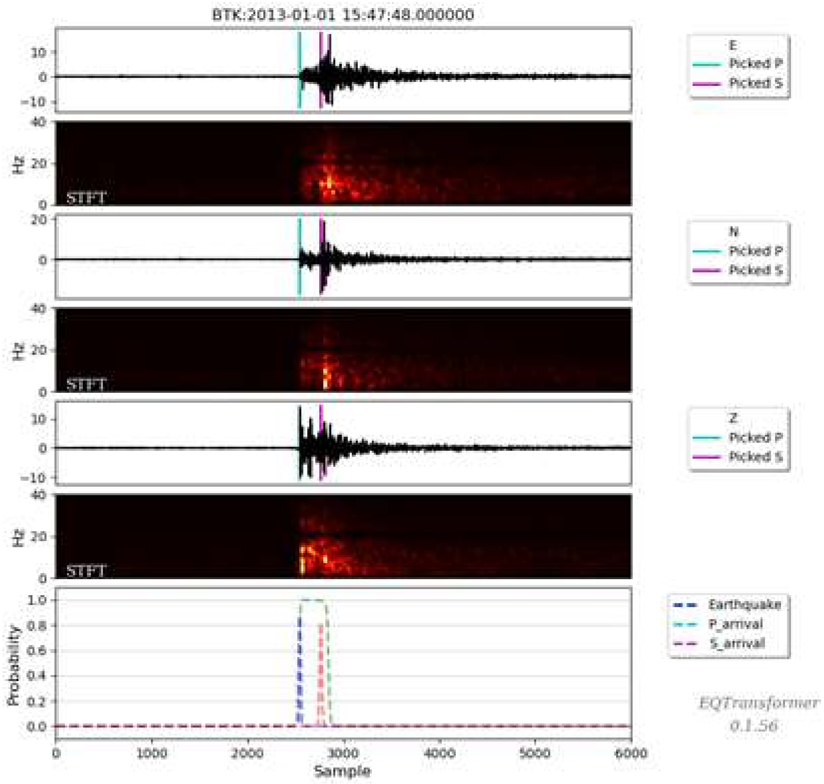


Figure 5: Velocity model used for locating the Events

The solution to the earthquake location problem is not linear but converges to the minimum root mean square value of offset time through a series of linear step length iterations, which is the best estimate of the seismic source location.

The arrival time and location is calculated by using two different velocity model: ak135 reference model; and a regional 1D minimum velocity model which is closely related to the Hindu Kush-Pamir is based on Sippl et al, 2013 study.

RESULTS AND DISCUSSIONS



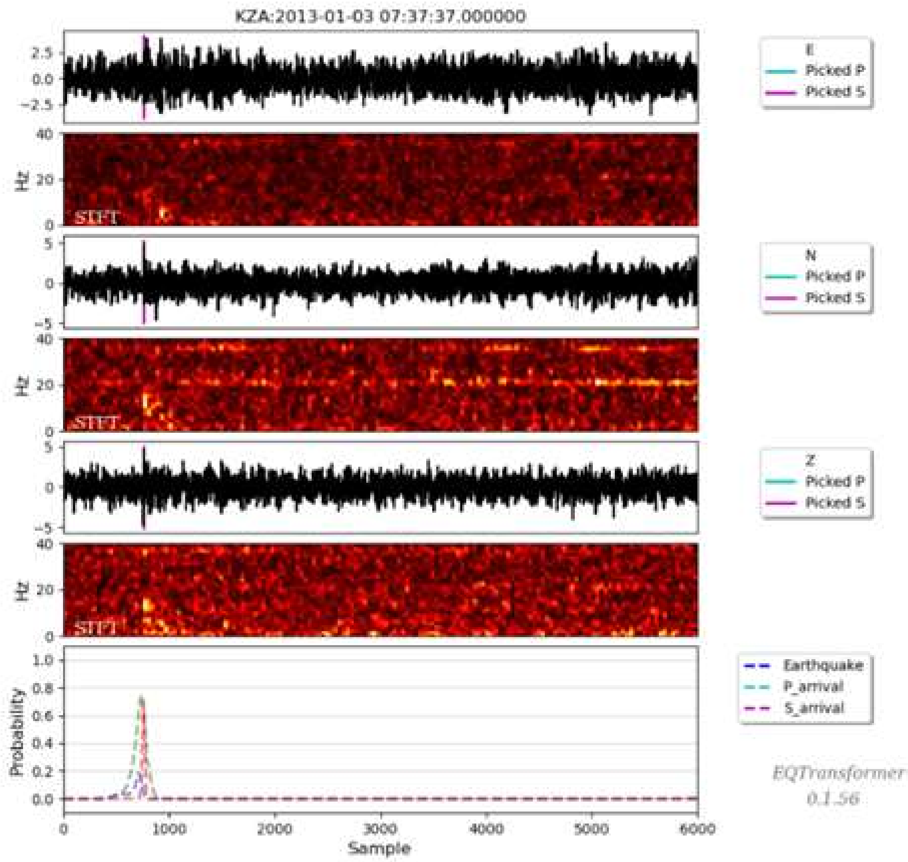


Figure 6: Detection made by network: A Event with high detection probability and A Event with Low detection probability

- Most of the event detection have high probability values between the range of 0.9 to 1.0 (See Figure 4)

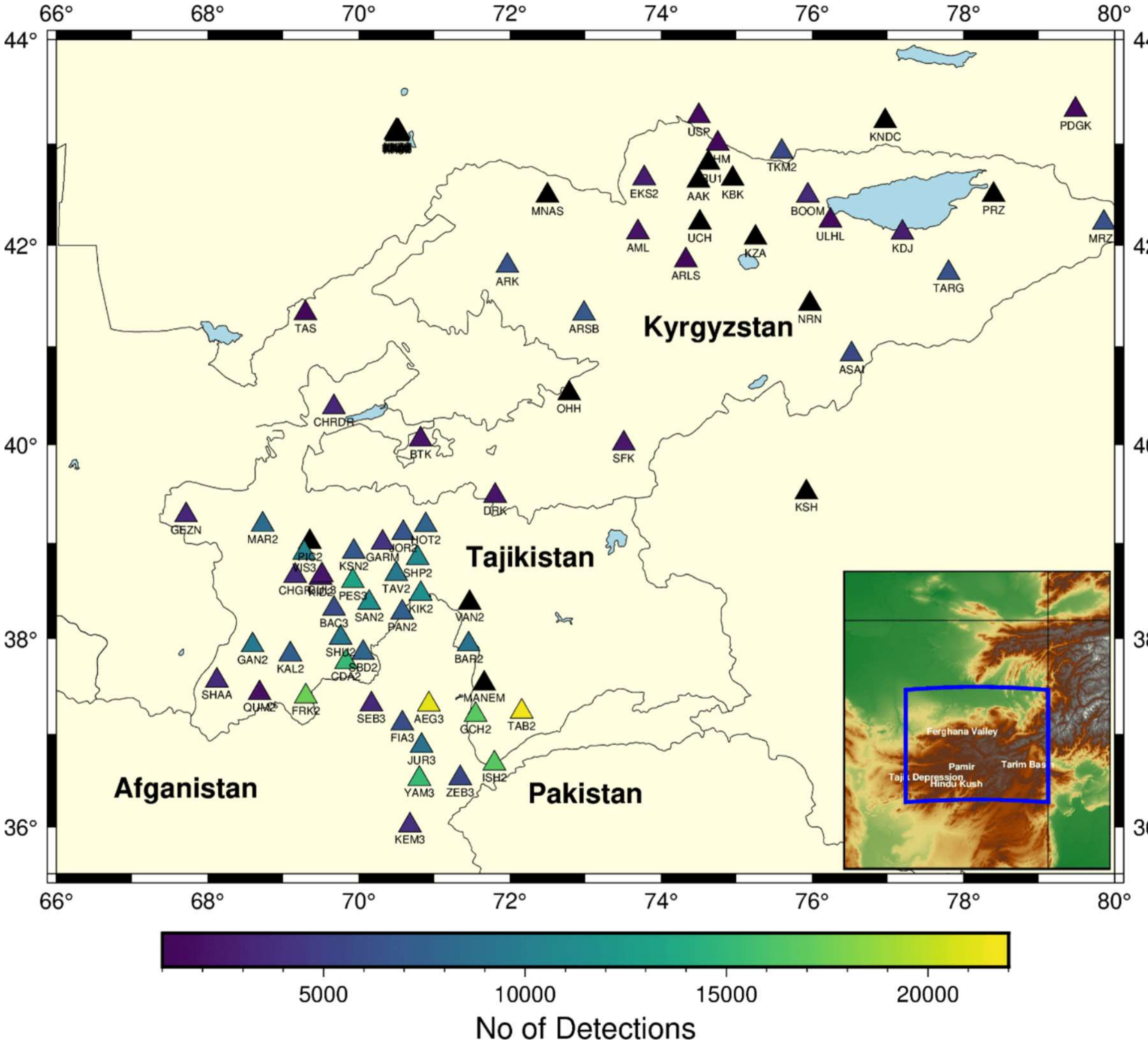


Figure 7: Distribution of Detection made at different Stations

- On average 5066 earthquakes were detected at each station in just 58 minutes on a machine with an octa-core 3.5 GHz Intel processor with 32GB memory.
- .A total of 21,229 associated events were detected which is more than seven times that exist in the ISC catalog which has only 2,959 events. Out of this in total 12,025 events were successfully relocated using hypoinverse.

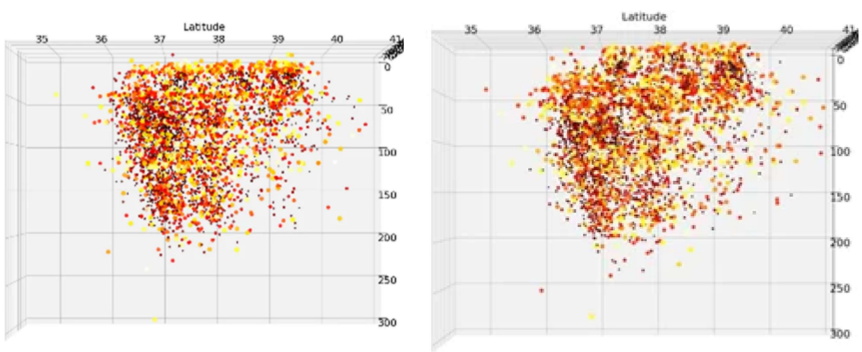


Figure 8: Depth Distribution of Located Earthquakes using ak135 and Sippl. 1D minimum Velocity Model.

The distribution of earthquakes in the 1D minimum velocity model is much better compared to the ak135 because Hindu Kush Pamir is mostly dominated by an HVZ beneath its mantle that extends up to 600 km in the central Hindu Kush (Kufner et al., 2021). 1D minimum velocity model has more detections with more minor depth calculation error 60% of the detection has an error less than 35 km. The ak135 model has a higher error in depth calculation, with most of the error being greater than 20 km. The error in latitude and longitude calculation is quite similar for both models.

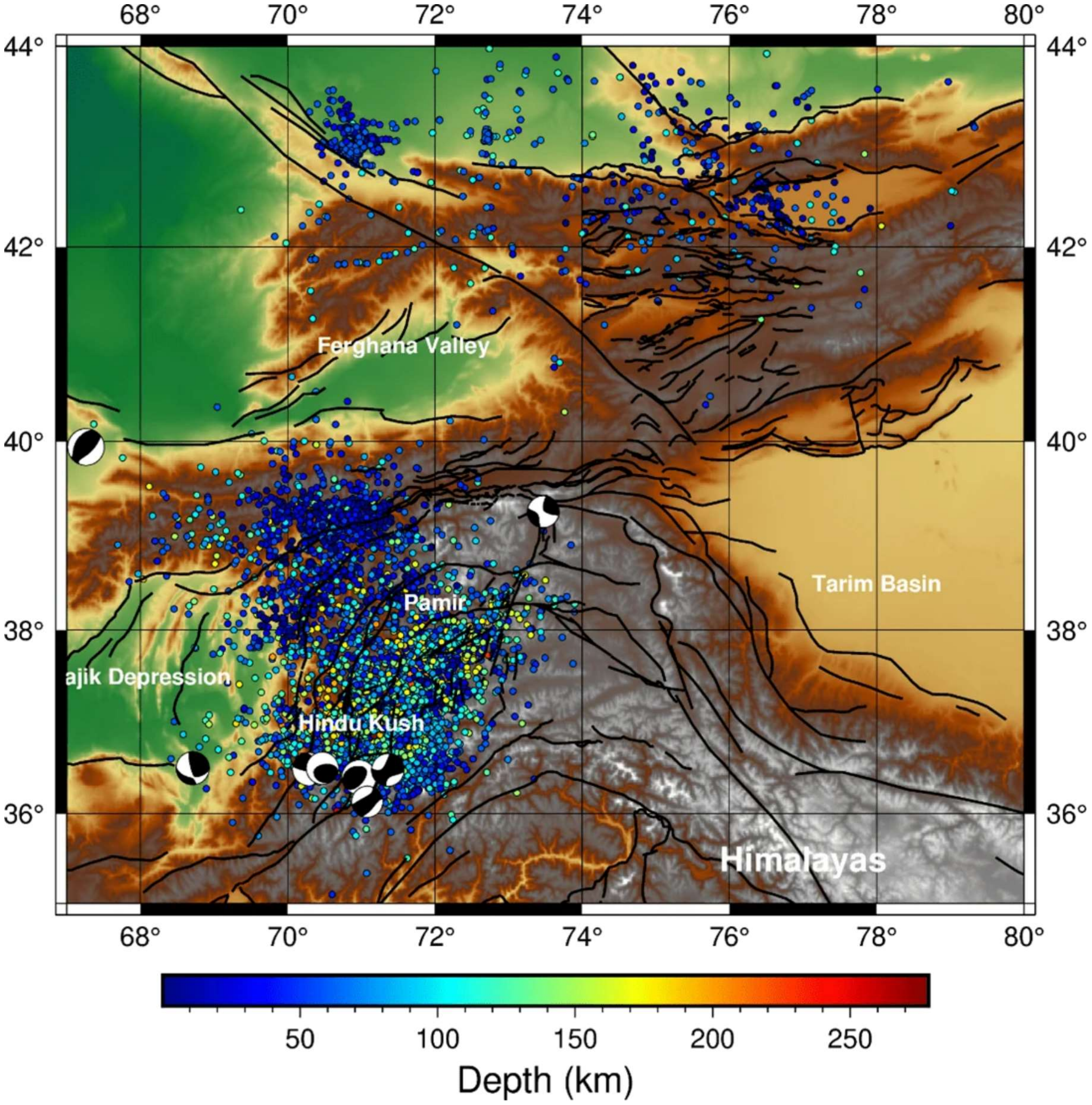


Figure 9: The distribution of the earthquake located using the Sippl et al. (2013) 1D minimum velocity model.

- More earthquakes are concentrated in the uppermost crust while the middle to lower crust appears to be less seismically active.
- The intermediate earthquakes located in the region form an S-shaped band.
- The Eastern part of the Pamir appears to be less seismically active than the Western part.
- Many shallow depth crustal earthquakes are also located throughout the region but mostly cluster in the northwest of this curved arc along the Gissal-Kokshal and Dawaz Karakul Faults.
- Only a few such earthquakes lie above the central Hindu Kush and part of the western Pamir intermediate depth zone. This shows the lateral variability in crustal deformation style and intensity may be associated with slab stretching in the mantle lithosphere.

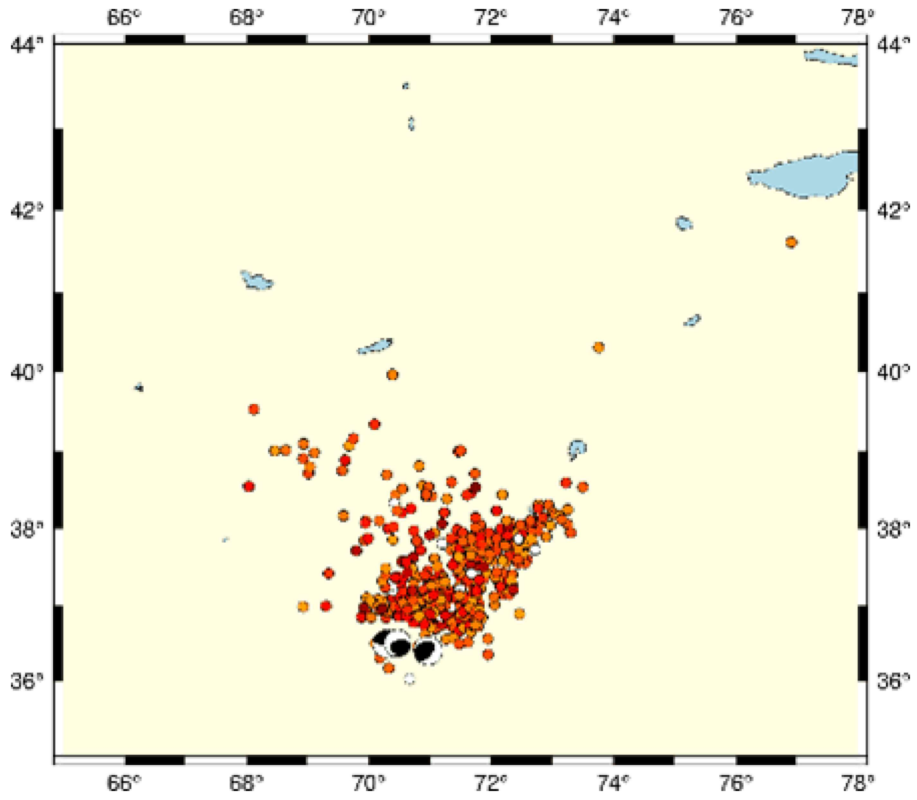


Figure 10: The intermediate-depth earthquake located at a depth greater than 150km.

- Earthquakes (deeper than 150 km) lying in the central Hindu Kush zone and western Pamir form a NE-SW oriented sub-vertical planar structure thinning and deepening in the east direction which might be due to slab stretching going on beneath (Zhan & Kanamori, 2016; Kufner et al., 2021).
- As the deeper part of subducts is faster than the shallow, there may be a possibility of stretching which may result in thinning producing a higher strain rate and hence such intermediate-depth earthquake.

CONCLUSION

- We have used an attention-based deep learning model for automatically detecting events and picking seismic phases quite similar to the manual picks.
- In this study we also developed automated selection criteria to filter out the less reliable detection based on the probability of detection and picked phases, this saved lots of computational time when working with large seismic data sets.
- Our attention model was applied from January 2013 to December 2013, and we were able to detect approximately seven times more earthquakes compared to those that existed in the ISC catalog. This clearly shows that the attention deep learning model is quite efficient in detecting low magnitude and recurring events that were not previously detected by manual picking or other automated algorithms.
- Two different velocity models: ak135 velocity model and 1D minimum velocity model based on Sippl et al 2013 is used in this study. We found that 1D minimum velocity is better for calculating the earthquake location in the Hindu Kush-Pamir region as it is closely related to the velocity model of the region.
- Seismicity of the region was studied using these located events show the deep intermediate earthquake (>150 km) forms a sub-vertical planar structure which may be due to slab stretching.

ACKNOWLEDGMENT AND DATA AVAILABILITY

Acknowledgment

The research was supported by IITR funds (grant FIG/100840). We would also like to thank IIT Roorkee for providing the computing facility for data analysis.

Data Availability Statement and Declaration

STEAD dataset used for training our model can be found on https://github.com/smousavi05/STEAD. The network architecture used was based on EQTransformer (Mousavi et al., 2020). Hindu Kush Pamir waveform data for our study period was downloaded from IRIS and GEOFON datacenters (https://ds.iris.edu/ds/nodes/dmc/). The relocation of earthquakes is carried out using the Hypoinverse-2000 (Klein, 2002). Seismicity data for comparison with our results were obtained from ISC catalog.

DISCLOSURES

Declaration of Competing Interests

The authors declare no competing interests.

ABSTRACT

Seismology data is overgrowing and is outpacing the development of processing algorithms. This tremendous increase in high-quality data can help better understand the earthquake processes related to the geology of active seismic regions such as the Hindu Kush. Most traditional detection algorithms are computationally inefficient compared to the amount of seismological data available and fail to detect low magnitude and noisy events. Deep learning algorithms are known for their applicability on large datasets with less runtime. Event detection and phase detection can be considered a supervised deep-learning problem quite similar to image recognition. In this study, we have implemented a hierarchical attention mechanism-based deep learning model for simultaneously phase picking and earthquake detection. This model is trained using Stanford Earthquake Dataset (STEAD), a globally disturbed labeled seismic dataset. We used this trained deep learning model to detect earthquake signals and pick P and S phases in the Hindu Kush - Pamir region for twelve months of continuous data spread across 83 different stations. A rigorous selection criterion based on detection, P and S phase probabilities, and other parameters has been used to associate the phases from different stations and to locate the earthquake. Our model detected almost seven times more earthquakes than previously existed in the catalog. The algorithm picked P and S phases with a high level of precision, comparable to human analyst picks. Furthermore, pinpointing these events allowed us to define the S-shaped seismic zone in the Pamir-Hindu Kush region and better comprehend the deformation caused by Eurasian- Indian plate motion.

REFERENCES

1. Bergen, K. J., Johnson, P. A., de Hoop, M. V., Beroza, G. C., (2019) Machine learning for data-driven discovery in solid Earth geoscience. Science, Mar 22;363(6433), doi: 10.1126/science.aau0323.
2. Freiberger, W. F. (1963). An approximate method in signal detection, Quaterly Appl. Math., 20, 373-378.
3. Joswig, M., (1990), Pattern recognition for Earthquake Detection, Bulletin of the Seismological Society of America, 80(1), 170-186.
4. Klein, F. W. (2002). User's Guide to HYPOINVERSE-2000, a Fortran Program to Solve for Earthquake Locations and Magnitudes pp. ftp://ehzftp.wr.usgs.gov/klein/hyp1.3/docs/hyp1.3.pdf*U.S. Geol. Surv. Open File Report 02-17*
5. Kufner, S. K., Kakar, N., Bezada, M., Bloch, W., Metzger, S., Yuan, X., Mechie, J., Ratschbacher, L., Murodkulov, S., Deng, Z., & Schurr, B. (2021). The Hindu Kush slab break-off as revealed by deep structure and crustal deformation. *Nature Communications*, *12*(1), 1–11. https://doi.org/10.1038/s41467-021-21760-w (https://doi.org/10.1038/s41467-021-21760-w)
6. Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019a). STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. 179464–179476
7. Mousavi, S.M., Ellsworth, W.L., Zhu, W. *et al.* Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat Commun* **11,** 3952 (2020). https://doi.org/10.1038/s41467-020-17591-w
8. Pegler, G., & Das, S. (1998). An enhanced image of the Pamir-Hindu Kush seismic zone from relocated earthquake hypocentres. *Geophysical Journal International*, *134*(2), 573–595. https://doi.org/10.1046/j.1365-246X.1998.00582.x (https://doi.org/10.1046/j.1365-246X.1998.00582.x)
9. Sippl, C., Schurr, B., Yuan, X., Mechie, J., Schneider, F. M., Gadoev, M., Orunbaev, S., Oimahmadov, I., Haberland, C., Abdybachaev, U., Minaev, V., Negmatullaev, S., & Radjabov, N. (2013). Geometry of the Pamir-Hindu Kush intermediate-depth earthquake zone from local seismic data. *Journal of Geophysical Research: Solid Earth*, *118*(4), 1438–1457.
10. Yoon, C. E., O'Reilly, O., Bergen, K. J., & Beroza, G. C., (2015) Earthquake detection through computationally efficient similarity search. Sci Adv., Dec 4; 1(11), doi: 10.1126/sciadv.1501057