

Enhanced crustal and intermediate seismicity in the Hindu Kush-Pamir region revealed by attentive deep learning model



Satyam Pratap Singh^a, Vipul Silwal^{b,*}

^a EarthByte Group, School of Geosciences, The University of Sydney, NSW, 2006, Australia

^b Department of Earth Science, Indian Institute of Technology, Roorkee, Uttarakhand, 247667, India

ARTICLE INFO

Keywords:
Deep-learning
Earthquake detection
STEAD
Hindu kush-pamir
Seismicity

ABSTRACT

The Hindu Kush-Pamir region (HKPR) is characterized by complex ongoing deformation, unique slab geometry, and intermediate seismic activity. The availability of extensive seismological data in recent decades has prompted the use of deep learning algorithms to extract valuable insights. In this study, we present a fully automated approach for augmenting earthquake catalogue within the HKPR. Our method leverages an attention mechanism-based deep learning architecture to simultaneously detect events, perform phase picking, and estimate magnitudes. We applied this model to a ten-month dataset (January 2013–October 2013) from 83 stations in the region. Utilizing a robust criterion to evaluate the model's probabilities, we associated phases at different stations and pinpointed earthquake locations in the region. Our results demonstrate a significant enhancement, revealing nearly four and a half times more earthquakes than previously documented in the International Seismological Center (ISC) catalogue. A notable portion of these newly detected events falls within the category of very low-magnitude earthquakes (<3), which were absent in the ISC catalogue. Notably, our spatiotemporal analysis reveals a concentration of crustal seismicity along poorly mapped neotectonic north and northeast-oriented faults in the western Pamir, as well as the Vakhsh Thrust System and the Darvaz Karakul Fault. These findings underscore potential sources of future seismic hazards. Furthermore, our expanded earthquake catalogue facilitates a deeper understanding of the interplay between crustal and intermediate seismic activity in the HKPR, shedding light on the deformation and active faulting resulting from Eurasian-Indian plate interactions.

1. Introduction

Seismological data is growing continuously over time (and space) and outpacing processing algorithms' development (Kong et al., 2018). When processed cautiously, this dataset can immensely improve our understanding of earthquake mechanisms and underlying tectonic processes, particularly in regions with high seismicity, such as HKPR. The HKPR is a 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 that are primarily associated with oceanic plate subduction rather than intracontinental (Kufner et al., 2021; Pavlis and Das, 2000; Pegler and Das, 1998). Recent studies using seismic imaging from a local network reveal the cold Asian mantle lithosphere beneath the Pamir subducts, but only with the bottom crust

connected (Sippl et al., 2013). In the past two decades, there has been an enormous increase in recorded waveform data with many continuous seismic station networks distributed throughout the region, but many of the low magnitude and low signal-to-noise ratio earthquakes are still undetected. Hence, detecting and locating these undetected earthquakes can significantly improve the understanding of the regional tectonics and the peculiar geometry of the HKPR.

Automatic Detection and Phase Picking of earthquakes is one of the most common problems encountered in seismology and has been tackled in various ways for more than five decades (Freiburger, 1963; Joswig, 1990). Traditionally, human analysts manually carried out seismic phase picking, but the exponential increase in data made this option difficult. Many algorithms have been developed to automate detection, and these detection algorithms fall into three main categories:

* Corresponding author.

E-mail addresses: satyampratap.singh@sydney.edu.au (S.P. Singh), vipul.silwal@es.iitr.ac.in (V. Silwal).

A. STA/LTA: Using an energy detector, such as a short-term average/long-term average, commonly known as STA/LTA, is one technique to automate this procedure (Allen, 1978; Freiberger, 1963). They apply to a wide variety of earthquakes as they do not require prior waveform information. However, the significant drawbacks of this algorithm are that they fail to work for low SNR signals, overlapping arrivals, and cultural noise-rich signals and may produce false or no detection in such cases (Yoon et al., 2015).

B. Template matching: Waveform similarity is also effective in detecting and phase-picking problems. Waveform cross-correlation can be used to find the similarity between the continuous seismic records and known waveform templates. This template-matching technique has been compelling and versatile in finding undetected events in many regions around the globe (Peng and Zhao, 2009; Skoumal et al., 2014). This method has been shown to detect low-magnitude earthquakes, low-frequency earthquakes in tremors, foreshocks,

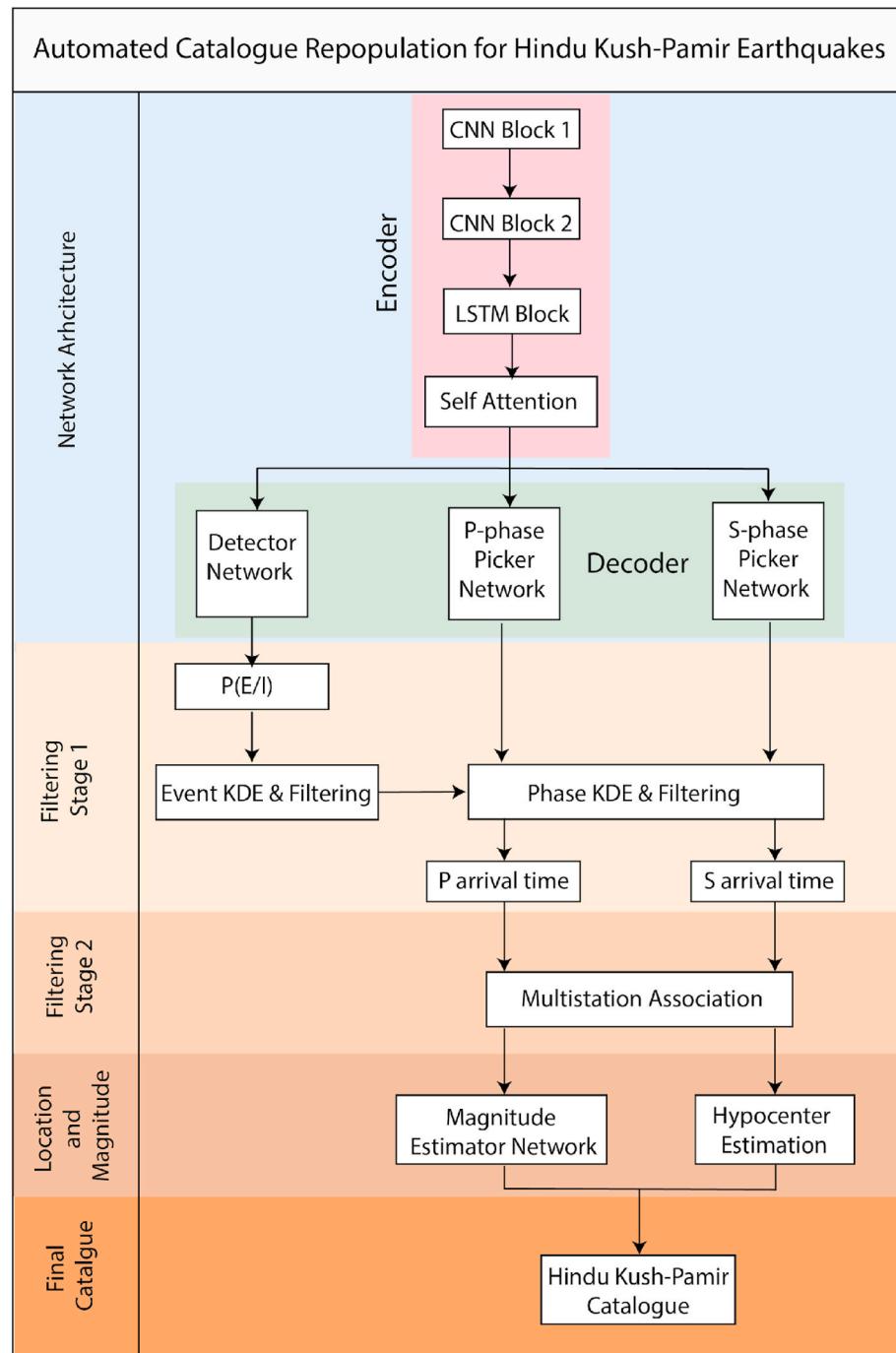


Fig. 1. a) Schematic of the methodology employed for the automated repopulation of earthquakes in the Hindu Kush-Pamir region (HKPR). A transformer-based neural network is utilized for earthquake detection and phase picking, providing us with $P(E/I)$ and $P(E_p/I)$ for raw waveform data at each station. In the initial filtering stage, high-probability detections and picks are retained at each station. In the second stage, a simple phase association method is employed to associate these detections across at least four stations simultaneously, effectively eliminating remaining false positives and true negatives. To ensure the inclusion of only reliable events in our catalogue, we exclusively located only events that are associated events with a minimum of five picks. These events are then used for hypocenter estimation using travel time residual calculations, resulting in low root mean square travel time residuals. Furthermore, our model incorporates a magnitude estimator to estimate the magnitudes of the detected events.

aftershocks, and microseismic monitoring problems. Nevertheless, a major drawback of this technique is that it requires a known waveform for creating a template, hence requiring prior waveform information.

C. Autocorrelation: Dividing the entire time series into overlapping windows and further performing cross-correlation of all possible pairs can be used to solve the requirement of the template (Brown et al., 2008). However, it involves many-to-many waveform correlations, making it computationally intensive and ineffective to work on a sizable amount of data.

Deep learning, a specialized machine learning technique, can be helpful in a wide range of problems like image detection and speech recognition (LeCun et al., 2015). In the same way, deep learning has been used to detect earthquakes and phases in seismic records. Deep learning uses multiple layers in stages to extract higher-level features from the input data (Kong et al., 2018). Implementing diverse types of deep learning models to effectively find events and phase arrival is the new front of the application of machine learning in seismology (Bergen Karianne et al., 2019; Perol et al., 2018). Despite different approaches and models, numerous studies have shown that deep learning-based models are more efficient than traditional approaches for such a large volume of data (Dokht et al., 2019; Mousavi et al., 2019b; Mousavi et al., 2020; Ross et al., 2018). Hence, we apply an attention-based deep learning model, to detect previously unseen P and S-phases generated by earthquakes in the study area (Mousavi et al., 2020) as well as to estimate the magnitude of these events (Fig. 1). Paired with a subsequent location technique, magnitude estimation and GPS data, we ameliorated the understanding of regional tectonics, the current deformation field, and the peculiar slab geometry of the HKPR.

2. Study area

The HKPR is in a wide complex deformation zone in the western Himalayan syntaxis formed by the indentation of the Indian Plate into Eurasian. Three distinct stages have been used to describe the region's tectonic evolution (Pegler and Das, 1998). During the Mesozoic, several plates migrated northward from the Gondwana continent, causing the formation of a magmatic arc and the closure of the Tethys Ocean. Around 50–52 Ma (Guillot et al., 2003; van Hinsbergen et al., 2011), the collision was initiated, and the Indian plate accreted to the southern border of the Eurasian plates via the Indus Suture Zone. The region's high mountain ranges and plateaus resulted from the ongoing convergence at 3–5 cm/yr along the western syntaxis (Ischuk et al., 2013). For the past 25 Ma (Sobel and Dumitru, 1997), the Pamir Ranges have shifted northwards, evident from the terrane accretion history. The Pamir-Plateau crust is currently moving northward, colliding with the Tian Shan, and collapsing and extruding westward into the Tajik depression, creating the Tajik Fold Thrust Belt (Fig. 9). Several strike-slip faults accommodate this northward migration (Metzger et al., 2021; Schurr et al., 2014). The Himalayan Tibet system has accommodated almost similar Cenozoic convergence as the Himalayan Hindu Kush Pamir system. However, the latter has a shorter meridional width due to a significant shortening of the upper plate and much of the crust being subducted in the mantle (Pegler and Das, 1998; Searle et al., 2001). The seismicity associated with the region is quite different from the Himalayan-Tibet collision and other continental collision zones worldwide. The seismicity of the Himalayan-Tibet zone is mainly located in the crust and uppermost mantle (<90 km) and is aseismic below this depth, unlike that in the Hindu Kush, which has intermediate-depth seismicity (70–300 km). The tectonic evolution and geodynamics of the region are complicated and include the subduction of the Indian lithosphere, which produces many large intermediate-depth earthquakes. Many studies have relocated earthquakes to understand the eccentric geometry and intermediate-depth seismicity caused by the complex lithospheric deformation in the

region (Kufner et al., 2021; Poli et al., 2016; Sippl et al., 2013; Zhan and Kanamori, 2016). As outlined by seismic activity, the overturning of the dip along the strike often raises the question of whether it is a single structure (Billington et al., 1977) or formed from two different structures (Pavlis and Das, 2000; Pegler and Das, 1998). The convergence rate at the surface is far less than the slip rate required to produce such frequent earthquakes, suggesting some significant internal deformation of the slab, which may be because of the localised stretching and sinking of the slab (Kufner et al., 2021; Zhan and Kanamori, 2016).

3. Methodology

The most crucial aspect of earthquake catalogue repopulation involves implementing automated detection on a large dataset by designing an appropriate deep-learning model capable of effectively identifying seismic signals amidst various non-seismic signals recorded by seismometers. Detection entails the identification of these seismic signals. Conversely, phase picking involves estimating the arrival timings of different seismic phases, such as the P-wave and S-wave phases, within the seismic signal. Once we have both the detections and picked phases for a single event, recorded simultaneously at several stations, we can utilize them to estimate the hypocenter location and magnitude of that event (Fig. 1).

3.1. Detection and phase picking

All the deep learning detection and picking models are based on high-level representation learning containing certain common features of the earthquake waveform and seismic phase. Realistically, these models should incorporate the fact that not all parts of the seismic signal are useful for the specific detection/picking job. This feature is implemented using an attention-based mechanism in the model (Mousavi et al., 2020). Humans try to focus more on a specific part of the image with high resolution while seeing the surrounding with a lesser resolution to discover an object in an image. Similarly, this neural network 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 (Supplementary Fig. 1). Based on the attention mechanism, a deep learning architecture called the EQtransformer is utilized for simultaneous identification and picking of P and S phases in waveform data acquired at several stations in the HKPR (Mousavi et al., 2020). The structure of this deep learning model is multi-task. It comprises 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 (Supplementary Fig. 1; Mousavi et al., 2020). The encoder takes a time-domain seismic signal and extracts a high-level representation and contextual information. The three decoders are then utilized to map this high-level data into detection probability and phase selection (Fig. 1).

The detection and phase picking are similar problems, but the goal of detection differs from that of picking. The main goal of detection is minimizing false negatives and false positives, while phase picking focuses on the temporal accuracy of the arrival time. The network first learns to distinguish the earthquake signal within the time series. These outputs are then sent to local attention to focus on the smaller section within the earthquake waveform to make the P and S phase picking. So, the hierarchical attention mechanism can be thought of as conditional probability (Mousavi et al., 2020). Moreover, the network's alignment score can be considered the probability distribution of detection and phase picking.

$$P(E | I) = \text{encoder}(I) \quad (1)$$

$$P(E_P | I) = P_{\text{decoder}}(P(E | I)) \quad (2)$$

where $P(E|I)$ is probability of earthquake given the input waveform time-series and $P(E_p|I)$ is probability of p-phase pick given the input waveform.

A profound structure such as that of EQTransformer is useful for HKPR because it is less sensitive to noise and works well for small events with high background noise (Mousavi et al., 2020). In addition, the alignment scores calculated by the model are normalized and can be thought of as the probability distribution of detection and phase picking, which has been used to automate selection and elimination of unreliable detections effectively.

3.2. Automatic selection criteria

The probability parameter can be quite handy in eliminating unreliable detection. Given shear amount of waveform data and corresponding detections, it is quite manually intensive to individually review each detection as deep learning model may often produce erroneous detections. So, we applied stage wise filtering of detected earthquake in order to repopulate catalogue. We implemented the following simple selection criteria. The first step was to select a threshold value for the probability of detection calculated using Equation (1), below which we eliminated all the detection. Choosing a threshold value for detection probability can be particularly challenging since there is a significant tradeoff between false and true positive events. A low threshold value means many false positive events may be included in our detections, and a high value will probably filter out many true positive detections. However, our goals maximize true positive detections. So, we used a Kernel Density Estimation (KDE) plot for all the detections made by the model at all different stations. KDE is a non-parametric statistical technique used for estimating the probability density function (PDF) of a random variable. It is commonly employed to visualize the underlying distribution of a dataset when we have several data points. KDE is particularly useful for visualizing data distributions, identifying modes (peaks) in the data, and providing a continuous estimate of the PDF, especially when the underlying distribution is not known.

$$f(P) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{P - P_i}{h}\right) \quad (3)$$

K is Gaussian Kernel Function; P is the respective probability; n is the number of data points and h controls the bandwidth parameter which is calculated as

$$h = n^{-\frac{1}{5}} \quad (3.1)$$

KDE of detection probability shows the distribution of how sure the deep model feels is in its detection. We used this distribution to select a threshold value for detection probability. Such criteria solve the problem by filtering out many but not all false-positive events from the waveform data. Therefore, we used a second threshold on P and S phase pick probability obtained using Equation (2), which eliminates those detections in which P and S phases are not apparent. This threshold value is selected after manual inspection of events detected by the model across different stations. These two-step selection criteria ensure that most false-positive events are filtered out and significantly save computation costs for next filtering stage.

3.3. Phase association and hypocenter estimation

After applying the automated selection criteria, we may still encounter both erroneous detections. To address this, we implement a second stage of filtering, wherein we link the detections across multiple seismic stations simultaneously. We establish a minimum requirement of 4 stations for an association to be considered valid. In other words, an event will only be recognized if it is detected at a minimum of 4 different stations simultaneously. This additional criterion serves to further refine our results by reducing the incidence of false positives and false

negatives since detections must occur across multiple stations simultaneously.

Once events have been associated, we proceed with hypocenter inversion, utilizing the residual travel time calculation from our picked P and S phase arrival (Klein, 2002). This process is resolved through a series of iterative steps, employing linear techniques to converge towards the minimum Root-Mean-Square (RMS) travel time residual, which stands as the best estimate of the hypocenter (Klein, 2002). We used two different velocity models: the ak135 reference model and a regional 1D minimum velocity model closely related to the Hindu Kush-Pamir, based on Sippl et al. (2013) study for hypocenter location (Fig. 6a). Only close stations were used for locating events; we further eliminated the stations >250 km away from the epicenter after the third iteration in hypocenter inverse if it cannot reach its threshold RMS minima (Klein, 2002). Primarily because the model was trained on the STEAD dataset, which contains only local earthquakes, and secondly because at a farther distance, there may be refracted phase arrival for shallow crustal events, which our model may falsely pick as P Phase. Furthermore, we located only those events that were detected at a minimum of five different P and S picks after applying all these criteria. This is primarily to account for linear dependability of our picks. We used both the P and S phases for locating earthquakes.

3.4. Magnitude estimation

The local magnitude scale, denoted as M_L and introduced by Richter (1935), is defined as follows:

$$M_L = \log(A) - \log(A_0) + S \quad (4)$$

In this equation, A represents the peak amplitude on the horizontal component, while the terms $-\log(A_0)$ and S are corrections empirically determined based on distance and station characteristics. Estimating $-\log(A_0)$ and S for each event can be time-consuming and error-prone, particularly for smaller magnitude earthquakes. To accurately estimate magnitude, it's crucial to model the average attenuation and site functions, and a neural network can learn these relationships from training data. However, the main challenge is providing the network with the necessary information to effectively learn these relationships (Lomax et al., 2019). In Equation (4), the amplitude information plays a pivotal role. Nonetheless, data normalization is typically essential for achieving good performance in a convolutional network. One approach to tackle this challenge is to include the maximum amplitude as an additional numerical input into the network, as suggested by Lomax et al. (2019). However, Mousavi and Beroza (2020) found this approach to be less effective. Consequently, we adopted a network primarily composed of convolutional and Bidirectional LSTM layers (Supplementary Fig. 1). Interestingly, in our network, the convolutional layers do not utilize activation functions but are employed solely for dimensionality reduction and feature extraction. The utilization of LSTM units for magnitude estimation provides a notable advantage due to their ability to effectively handle unnormalized inputs, due to their distinctive gated mechanism involving tanh and sigmoid activation functions (Mousavi and Beroza, 2020).

3.5. Dataset

To train our networks, we used the Stanford Earthquake Dataset (STEAD). STEAD is a global dataset of earthquakes and non-earthquake signals that has been labelled at a large scale (Mousavi et al., 2019a). The earthquake data consists of approximately 1,050,000 three-component seismograms, each lasting 1 min. These seismograms are linked to around 450,000 earthquakes that occurred between January 1984 and August 2018. The earthquakes were captured by a network of 2613 seismometers situated worldwide, all located within 350 km of the seismic events. In addition to the seismic recordings, the earthquake data includes a wide array of information about each event.

This includes details such as the origin time of the earthquake, the arrival times of the P and S phases, magnitude, magnitude type, its epicentral location, depth, focal mechanism. It also provides estimates of errors and measures the signal-to-noise ratio for each seismogram component (Mousavi et al., 2019a). Furthermore, it notes the end of the signal's dominant energy (coda-end), and the epicentral distance. The majority of the seismograms in the dataset were collected within a proximity of 110 km from the earthquake source. Most of these earthquakes are relatively shallow, with depths less than 50 km. While the dataset encompasses earthquakes of various sizes, it predominantly includes smaller earthquakes with magnitudes less than 2.5. Magnitude information is reported in 23 different scales, with local (M_L) and duration (M_d) magnitudes being the most prevalent (Mousavi et al., 2019a). For the purpose of training magnitude estimation models, the dataset primarily utilizes data that includes information on the local M_L magnitude.

In order to repopulate HKPR catalogue, we collected ten months of continuous seismic data, spanning from January 1, 2013, to November 5, 2013, from multiple sources, including IRIS, GEOFON data centers, and publicly accessible networks: CB, II, KC, KN, TJ, KR, KZ, 5C, 6C, and CK. This dataset comprises a total of 83 distinct stations distributed across the entire region (Supplementary Table 1 and Fig. 2). The majority of these networks remained operational throughout the entire data collection period, with the exception of network 6C, whose stations were predominantly established in May 2013. The recorded data during this timeframe was sampled at rates of both 40 and 100 Hz. To prepare the data for subsequent detection and phase picking tasks, we conducted preliminary processing akin to the procedures applied to the training dataset. This initial processing involved the conversion of raw waveform data from each station into 1-min-long, three-component seismograms. Subsequently, we fed this waveform data into our deep learning architecture, which provided us with probabilities of event detection and arrival times. We then perform phase association and hypocenter estimation. Upon successfully locating the events, we further processed the waveform data associated with these detected events by passing it through our magnitude estimator network for magnitude determination.

4. Result and discussions

The training on the STEAD dataset took 68 h to generate the model parameters on a 12-core 2.20 GHz with 32 GB of memory, which once trained can also be used for the study of other regions around the world as well. Fig. 2 shows the distribution of earthquakes detected by each station. Applying this model on continuous recorded data, on average, we detected 5066 earthquakes at each station in 58 min for ten months of data which indicates the promptness of the deep learning model.

4.1. Comparison with conventional detection algorithms

Fig. 3 presents a comparison between detections made by various conventional methods and their comparison with our deep learning model. Upon analyzing the raw waveform data (Fig. 3a), one can notice three potential events occurring at GARM station on the dates and times of 2013-02-11 17:31:06, 2013-02-11 17:39:30, and 2013-02-11

17:44:24. Within the Fig. 3b, a similarity assessment is illustrated for two distinct templates employed in template matching algorithms. These templates are derived from past earthquake data recorded at the GARM station. The first template yields similarity values at three different times, registering values of 0.12, 0.2, and 0.16 when a threshold value of 0.12 is chosen. Conversely, the second template exhibits notably lower similarity values, rendering it insufficient for detection at the specified threshold value. This underscores a significant limitation of the template matching technique, which relies on prior waveform information and necessitates the setting of a threshold value to trigger detections. Fig. 3d showcases the STA/LTA ratio values and the corresponding detections. The primary challenges associated with STA/LTA entail determining appropriate window sizes for STA and LTA to ensure accurate value computation and deciding on an optimal trigger threshold for detection. Utilizing a larger threshold may eliminate true positive events, while a very low threshold may result in numerous false detections (Figs. 3d and 4a).

Fig. 4a provides an example of a false event detected by STA/LTA when a low threshold value is employed. Conversely, Fig. 3f demonstrates that the deep learning model identifies several additional events, alongside the prominent earthquake. Upon closer examination, a smaller amplitude event becomes apparent during the same period (Fig. 4b). Furthermore, unlike STA/LTA, which detected events without pinpointing pick times, our deep learning model not only detects events but also determines the arrival times of P and S phases, along with the associated probability values, showcasing the model's confidence in this task (Fig. 4b).

4.2. Comparison with ISC catalogue

The ISC catalogue shows only 1877 events recorded during this period. The contributing source of the ISC catalogue generally detects events based on the manual picking or STA/LTA approach and fails to detect low magnitude, low SNR events. The application of a deep learning-based model solves these limitations quite efficiently and can detect many small events that were previously unnoticed. Our automated selection criterion filters out earthquakes with low detection, and phase pick probabilities. Fig. 5c and d shows the KDE of the detection and P and S phase respectively. The model is more robust towards the detection problem than picking. We have several detections with high detection probability values in the range of 0.9–1.0 (Fig. 5a, b, c). Based on this KDE distribution, we set a threshold value of 0.8 for detection probability. Several S phase picks have a higher probability than the P phase because the model is often less sure of P phase arrival in noisy traces (Fig. 5d). Therefore, P and S phases threshold selection requires manual inspection across different stations. Our P and S thresholds for this study ranged from 0.5 to 0.6. Based on this selection criterion (Table 1), we eliminate false events that would increase our time complexity during phase association. We detected 4,20,405 events at all stations combined, but our selection criteria filtered out low probability detection. Finally, we performed phase association on 2,73,723 filtered detections. A total of 21,229 associated events were detected, more than eleven times that in the ISC catalogue, with only 1877 events. The phase association allows to select only events that are simultaneously detected at atleast 4 station thereby filtering remaining false positives and false negatives. Out of this, 8467 events were successfully relocated using hypocenter inversion, 4.5 times that in the ISC catalogue (Figs. 6 and 7).

Figs. 6 and 7 compares the ISC catalogue and one generated by the deep learning model. We can see many crustal events in our catalogue which does not exist in the ISC catalogue. Of all the events, 999 events were common in ISC and our generated catalog.

Magnitude comparison shows that most of the earthquakes detected by our model has very low magnitude (<3) while that of ISC has magnitude >2.5 (Fig. 7d). We could not retain all the events because our model only detects the local events that lie <250 km from the station (Mousavi et al., 2019a, 2020), have low magnitude (Mousavi and

Table 1
Parameters used in our selection criteria.

Selection Parameter	Criteria
Detection probability threshold	0.8
P-probability	0.5 - 0.6 (obtained manually using KDE)
S-probability	0.5 - 0.6 (obtained manually using KDE)
Minimum stations for association	4
Minimum phases for hypocenter estimation	5

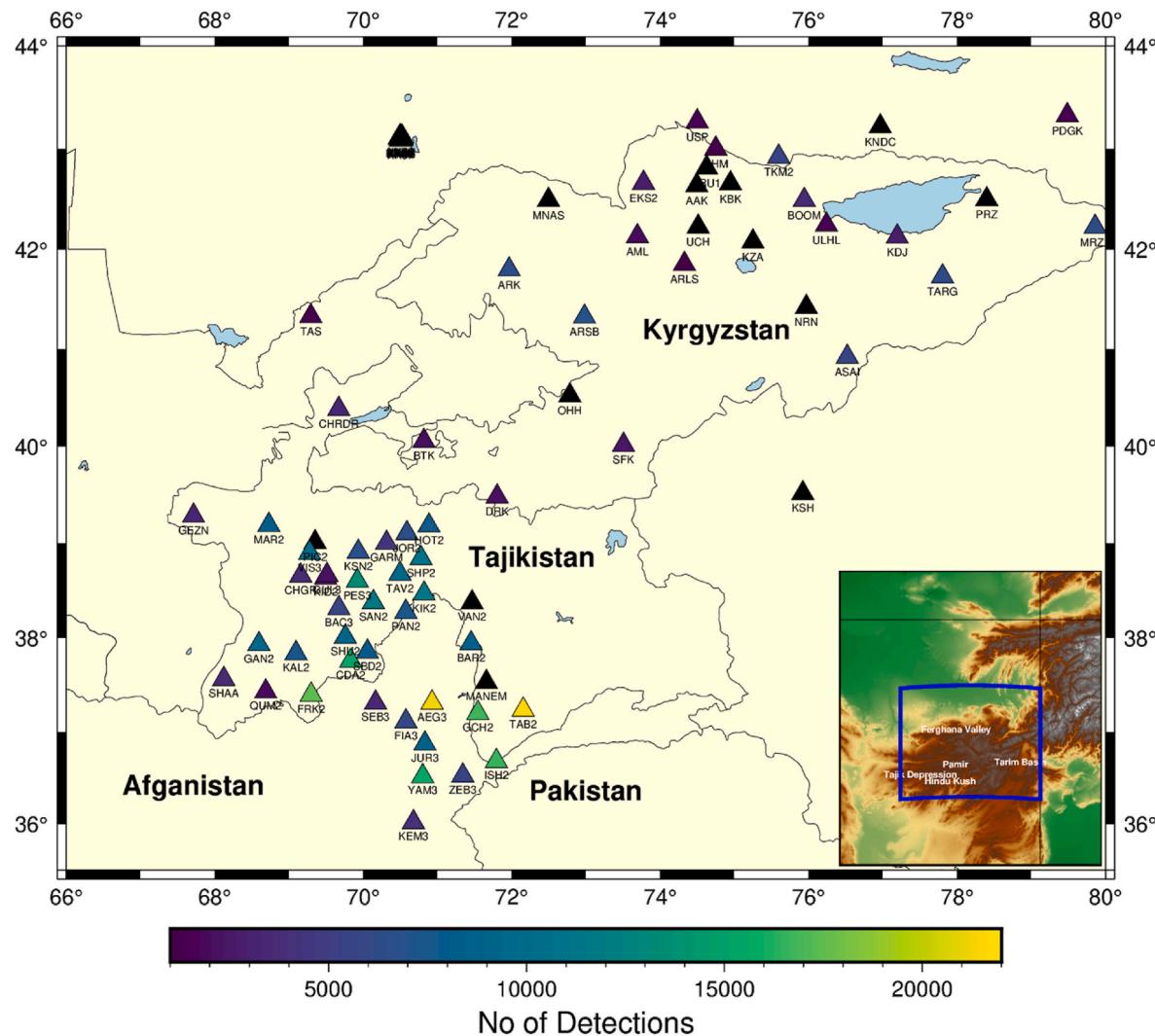


Fig. 2. Distribution of regional seismic stations used in this study, along with the number of detections made by each station based on our model from January 1st, 2013, to November 5th, 2013. Only stations with unrestricted data access were utilized and are depicted in the Fig. The waveform data for these stations ([Table S1](#)) was obtained from the IRIS and GEOFON data centers.

Beroza, 2020). So, firstly many events are eliminated in the region where station density is sparse (Fig. 2). Because our seismic network was devoid of stations in the Chinese Pamir for most of the recording period, our catalogue for this region may be insufficient, implying that we may have underestimated the actual seismicity rate there (Fig. 2). In Fig. 7a, we can see that events around the area 72–73° E and 39–40° N are present in the ISC catalogue but not in our catalogue (Fig. 7b) because very few stations are available around the region. Moreover, the accuracy and reliability of any deep learning also depend on the training dataset and since our model was trained on the STEAD dataset, which contains small magnitude events (Mousavi et al., 2019a); therefore, our deep learning model is limitation of detecting only small magnitude earthquake (Mousavi et al., 2020). Fig. 7c shows the month-wise seismicity located by our model, which significantly outnumbers ISC. We further explored spatiotemporal seismicity in HKPR (see section 4.3.3, Fig. 8). We were able to locate many previously unnoticed events. It demonstrates that deep learning is quite powerful and outweighs the traditional approaches.

4.3. Seismicity

We implemented two models for locating our hypocenters: a generalised ak135 global 1D velocity model and a 1D minimum velocity

model based on Sippl et al., 2013 (Fig. 6a). Both models show that there is a significant improvement in crustal seismicity (Figs. 6 and 7). Earthquakes typically occur in the upper 15 km of the crust, which corresponds to the normal seismogenic layer of the Pamir's brittle crust. The various active faults generate this seismicity in the region. The distribution and number of earthquakes in the 1D minimum velocity model is much better than in the ak135 because an HVZ mostly dominates the Hindu Kush Pamir beneath its upper mantle that extends up to 600 km in the central Hindu Kush (Kufner et al., 2021). We tried to map and understand the associated seismicity by dividing the events into two parts: shallow events and intermediate events (Figs. 8 and 9).

4.3.1. Shallow seismicity

We mapped the shallow seismicity generated by our catalogue to the fault system in the area. The observed seismicity pattern supports the region's large-scale deformation and fault kinematics. There is distributed seismicity associated with high seismic deformation in the south-western Pamir and west of the Sarez-Karakul Fault system (Fig. 9b). Although this region is active, knowledge about deformation is scarce; only a few active features have been identified so far because the region has widespread glaciation and the highest topography in the Pamir, making it inaccessible (Schurr et al., 2014). The bulk of the Cenozoic east-west structures is not consistent with the north or north-east

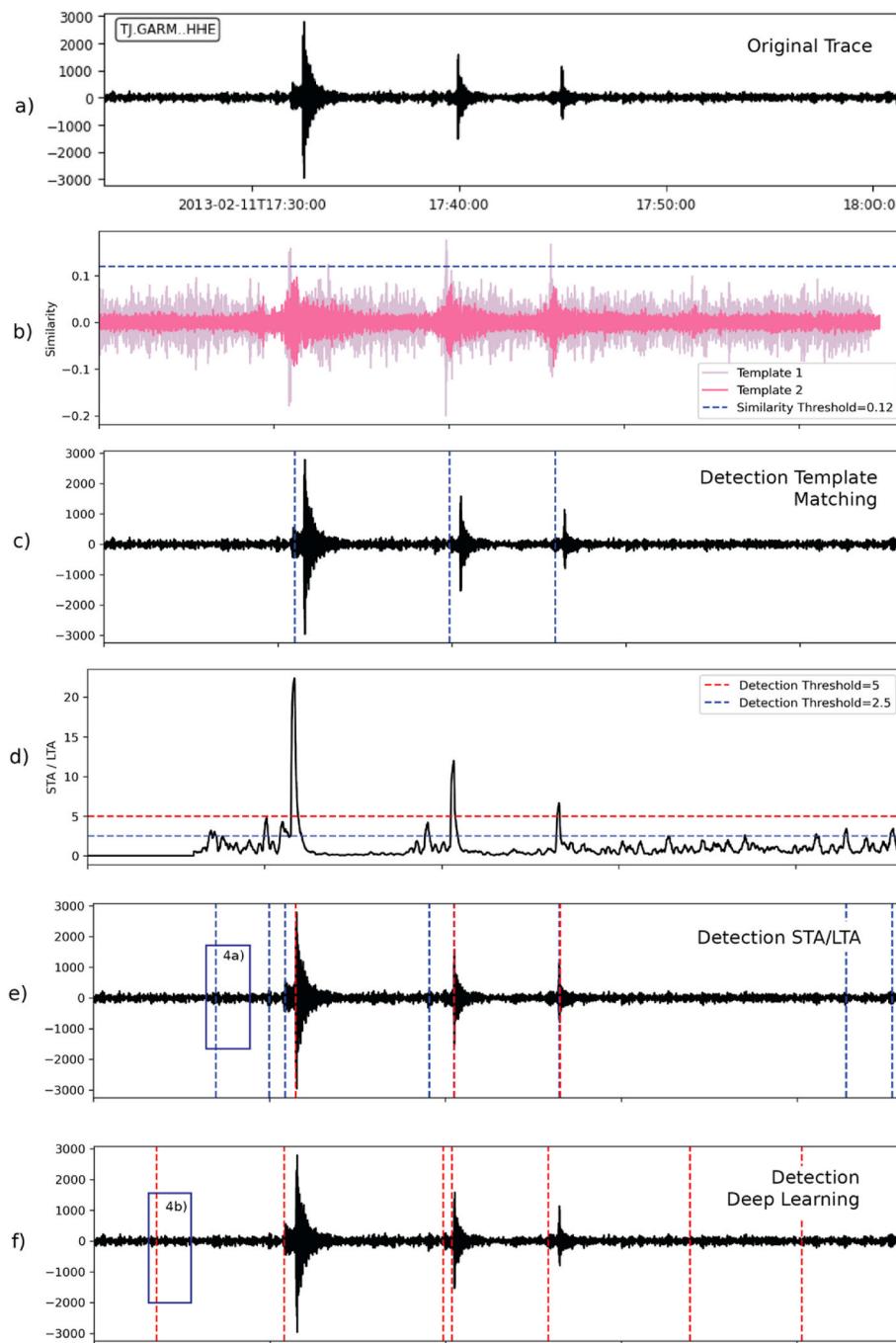


Fig. 3. A comparative analysis of various earthquake detection methods and their performance when applied to raw waveform data recorded at the GARM station. (a) Raw waveform data reveals potential earthquake events occurring on 2013-02-11 at 17:31:06, 17:39:30, and 17:44:24. (b) Template matching methods employ two distinct templates derived from past earthquake data, with the first template detecting events at 0.12, 0.2, and 0.16 similarity values when a threshold of 0.12 is chosen. (c) The corresponding detection by template matching based on threshold. Template 2 shows no detection (d) STA/LTA ratio values and their corresponding trigger point. It is quite a challenge of selecting appropriate window sizes and threshold values. (e) The corresponding detection for STA/LTA. A smaller threshold (blue) results in larger detections compared to higher threshold (red). A detailed inspection of one such low threshold event is shown in Fig. 4(a) and is perhaps a false positive. (f) The deep learning model identifies additional events, including smaller amplitude events (see 4(b)), alongside the three primary earthquake, demonstrating its enhanced detection capabilities.

oriented active faults, indicating that the deformation is new. Ruleman et al. (2007) have identified these northwest-trending strike-slip faults as structures with a mild expression in the landscape or with limited or poor representation in quaternary deposits, which may be a source of seismic hazards using satellite photography. The deformation pattern along these poorly mapped sinistral strike-slip faults along the northeast strike or conjugate plane and the normal fault along the north-south trend indicate north-south shortening and east-west extension. This

can be explained by the western extrusion of Pamir rock in the Tajik depression caused by the ongoing collapse of the western margin of the Pamir Plateau (Metzger et al., 2021; Schurr et al., 2014). In the Tajik depression, GPS velocity vectors shift anticlockwise from NNW in the eastern Pamir to WNW (Fig. 9b) (Ischuk et al., 2013; Zubovich et al., 2010). This rotation is a combination of two main motions that are perpendicular to one another, namely the main northward motion caused by the tectonic forces of the northern push and the main

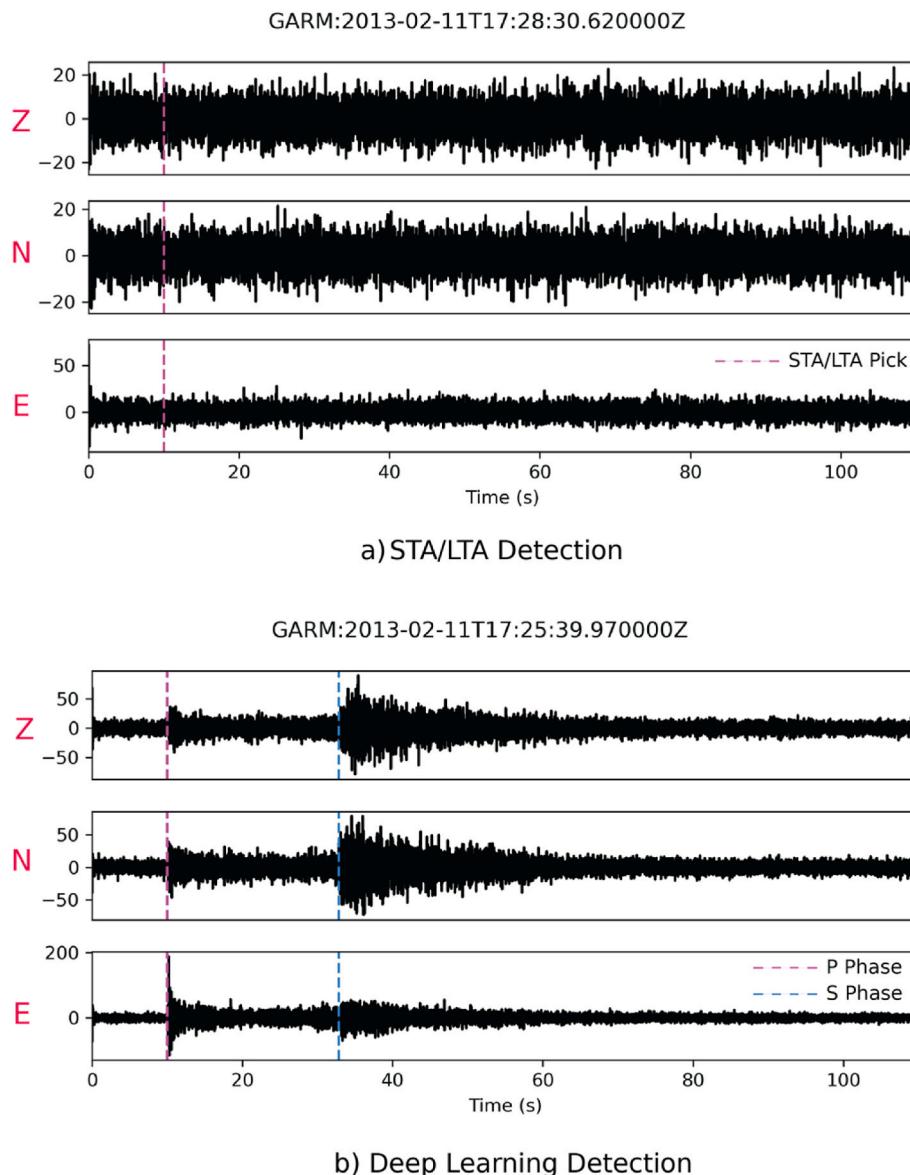


Fig. 4. Detailed analysis of additional events obtained using STA/LTA technique and deep-learning model in Fig. 3. (a) Example of a false detection when a low STA/LTA threshold = 2.5 is applied. (b) The deep learning model not only detects events but also accurately determines P and S phase arrival times, accompanied by associated probability values, exemplifying its confidence in event detection.

westward component caused by the gravitational forces from the elevation contrast between the 4 km high plateau and the 0.5 km high Tajik depression, causing an east-west extension (Schurr et al., 2014).

In the north, with 13–19 mm/yr of shortening and 7 mm/yr of dextral shear, the Pamir Frontal Thrust has the highest differential velocity and hence the most seismicity (Ischuk et al., 2013; Metzger et al., 2021). The Vakhsh thrust system, which is an active thrust front in the Pamir thrust system, continues further into the dextral-reverse lateral ramp of the Illiac fault zone and borders the thrust-fold belt of the Tajik depression (Kufner et al., 2018; Metzger et al., 2021; Schurr et al., 2014). Vakhsh Thrust exhibits a 15 + 4/-2 mm/yr compression with dextral shear of up to 16 mm/yr (Metzger et al., 2021). Offset streams suggest that recent fault zone activity along the sinistral transpressive Darvaz fault zone extends into the Pamir thrust system's rear (Metzger et al., 2021). Crustal seismicity can be traced along the Vakhsh Thrust System and the Darvaz Karakul Fault. Rates decrease from north to south across the Darvaz fault zone, from 15 to 7–9 mm/yr sinistral shear, and from 10 to 4 mm/yr extension can be correlated with decreasing seismicity (Metzger et al., 2021). A cluster of earthquakes exists beneath the

dextral transpressive Peter I. Range caused by the internal deformation and westward extrusion of the decollement between the Vakhsh and Darvaz Karakul Faults (Hamburger et al., 1992; Metzger et al., 2021). To our surprise, the westward motion also prevails on the northern edge of the Pamir, as indicated by the strike-slip fault and crustal seismicity in the Trans-Alai Range, although the GPS data predicts significant crustal shortening (Ischuk et al., 2013; Zubovich et al., 2010). A cluster of events can be found in the Vanch Range (SEDA site). These shallow events traces along the eastern edge of the Badakhshan Faults. This cluster was the epicenter of the Vanch Mw 5.3 earthquake in January 2010, causing much damage to Vanch City (Schurr et al., 2014). Several earthquakes can be seen along the Ferghana Faults and the Tian Shan faults in the north. A cluster of shallow events lies around 71.5° E, 43.5° N. This is because of the dense coverage of the KZ network near the Fergana Fault in the region.

4.3.2. Intermediate seismicity

The intermediate earthquakes located in the region form an S-shaped band. An obvious gap in seismicity that coincides with a change in the

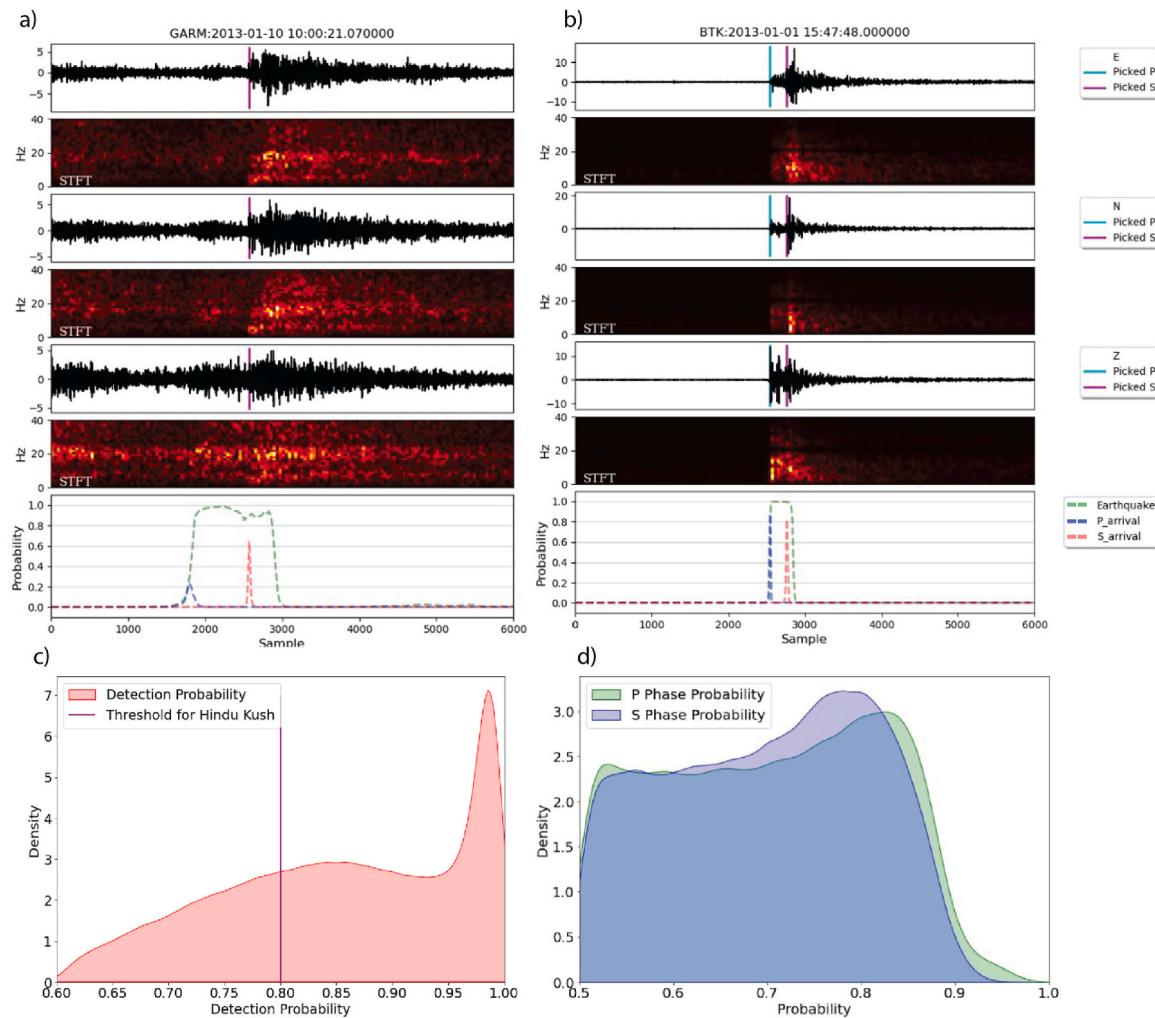


Fig. 5. a) Detection and phases marked by the deep learning model and corresponding short-term Fourier transform shows an event with low P phase pick probability; Although, the overall probability of detection is high such events are eliminated using the selection criteria. b) A detection with a significantly high probability of detection and P and S picks. c) The KDE distribution of the probability of detection and threshold detection probability we used in these study d) P and S phases of all the 4,20,405 events detected by the model.

dip direction of the Benioff zone allows for the first-order split of intermediate seismicity into the Pamir and Hindu Kush seismic zones (Fig. 9a). This gap narrows as it approaches shallow depths, where both zones adjoin just beyond the assumed Moho position. The eastern part of the Pamir appears to be less seismically active than the western part. The deepest earthquake in the eastern part is about 165 km, while in the western part, it is 235 km. Intermediate earthquakes in the Pamir Zone are restricted to a curvilinear structure with a steady dip in changing from south to east, and for earthquakes greater than 150 km, the dip steepens to become sub-vertical. The Pamir is a single curved plane with increasing curvature and vertical extension to the southwest. The Hindu Kush seismic zone is aligned east-west with a dip north direction. The earthquakes in the western section of this zone are restricted to a tight cluster. At its eastern end, the structure seems to be a planar with seismicity from crustal to about 278 km further it turns northward to join the main trend of the western Pamir zone. Only a few shallow earthquakes lie above the western and central Hindu Kush intermediate depth zones, which shows the lateral variability in crustal deformation style and intensity may be associated with slab stretching in the mantle lithosphere, which causes this crust-mantle decoupling (Kufner et al., 2021). Various imaging results have shown that the gradual eastward thinning, deepening, and steepening of the subducted slab indicates a slab breakoff process, with the breakoff most advanced beneath the central Hindu Kush (Kufner et al., 2021; Sippl et al., 2013). This

geometry best fits the scenario of a north-dipping Indian lithospheric slab beneath the Hindu Kush, with its focused seismicity and unambiguous dip direction. Earthquakes (deeper than 150 km) lying in the central and eastern Hindu Kush may be due to this slab stretching going on beneath (Zhan and 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 earthquakes. The hypothesized start of slab breakoff at mantle depths coincides spatially with the transition from assumed upper plate coupling to decoupling (sparsity of crustal earthquakes).

4.3.3. Spatiotemporal variation in seismicity

Fig. 8 illustrates the spatiotemporal variations in shallow and intermediate seismic activity within the HKPR. In the months of January and February, a notable concentration of crustal seismic activity is observed, primarily clustered around the Vakhsh Thrust System, the Darvaz Karakul Fault, and the neotectonic faults oriented to the north and northeast in the western Pamir region. During the period spanning March to May, seismic activity in the Vakhsh Thrust System and Darvaz Karakul Fault region subsides, while a distinct increase in crustal seismicity becomes evident along the north and northeast-oriented faults in the western Pamir. From June onwards, seismic activity experiences a significant resurgence along the Vakhsh Thrust System and the Darvaz

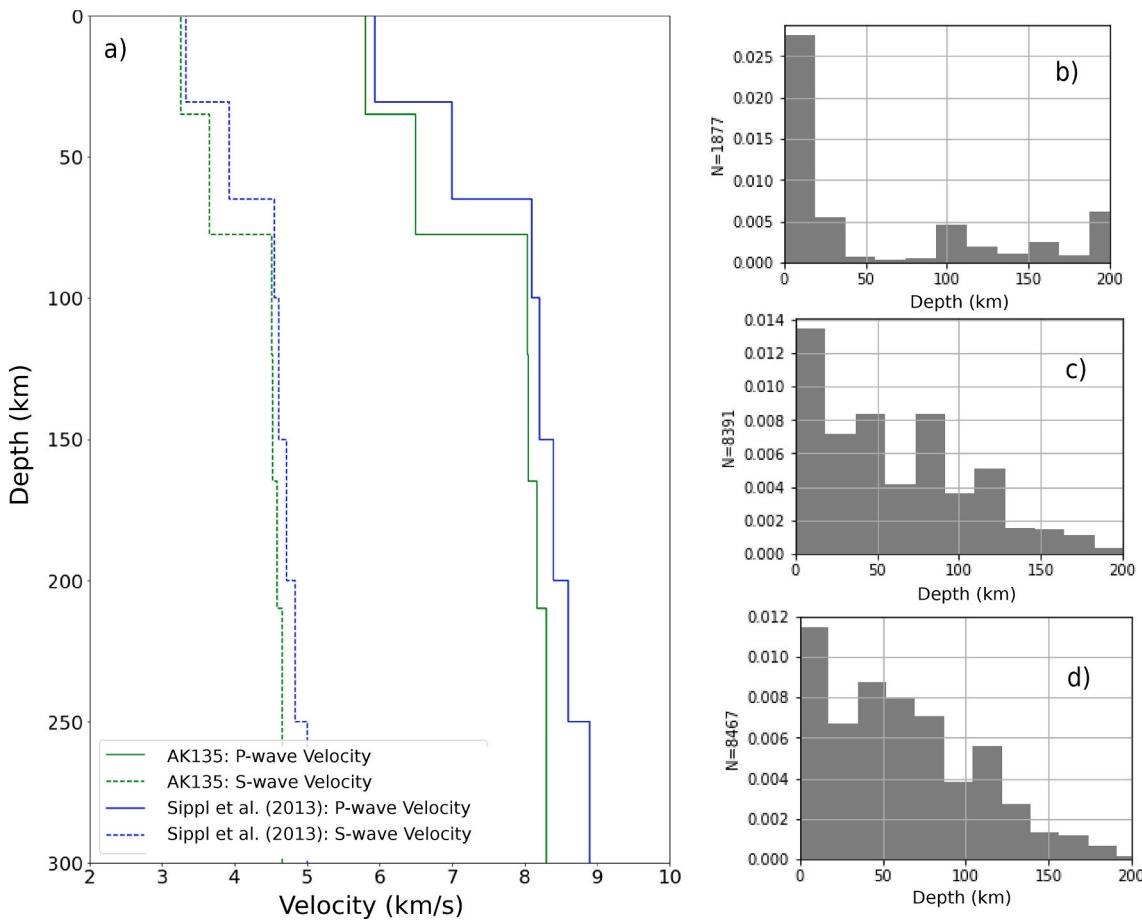


Fig. 6. a) The velocity models used for this study b) Depth Distribution of earthquakes in ISC catalogue c) and d) are depth distribution of earthquakes located by our model using the ak135 model and Sippl. et al. (2013) velocity model respectively.

Karakul Fault, with peak seismicity occurring in September and October. Concurrently, there is a notable uptick in shallow seismic activity in the western Pamir zone. Our model also successfully locates deeper earthquakes, demonstrating a significant correlation with the ISC catalogue. The month of May witnesses an increased occurrence of intermediate-depth seismic events, corroborated by both our model and the ISC catalog. Interestingly, our model identifies several low-magnitude intermediate earthquakes during January and February. It's worth emphasizing that throughout each month, there is a consistent presence of significant seismic activity in the intermediate zones of the Hindu Kush and Pamir regions. This suggests ongoing and continuous deformation in these areas.

Although the exact peculiar geometry and deformation of the HKPR might require further study using subsequent relocation and imaging techniques on much larger datasets. Nevertheless, such catalogue generation method using more advanced deep learning techniques can act as a first order processing to extract the maximum possible information from existing data in a faster and more efficient way with less manual intervention. This can in turn provide a fruitful geological understanding of the region when combined with existing geophysical techniques.

5. Conclusion

We have used an attention-based deep learning model for automatically detecting events, picking seismic phases and estimating magnitude. 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 sizeable seismic data sets. Our deep learning based model was

applied from January 2013 to October 2013, and we were able locate approximately 4.5 times more earthquakes than those in the ISC catalogue within a very short time. Magnitude estimation 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. Our model enhanced the region's seismicity and could map several earthquakes along the neotectonic likely active faults of western Pamir, which further supports the western extrusion of Pamir rocks due to the collision of the Pamir Plateau in the Tajik Depression. This earthquake may be a source of future seismic hazard as it is poorly mapped and studied due to its high altitude and widespread glaciation. The intermediate seismicity of the region was studied using these located events, showing two distinct zones with different dip directions. The deep intermediate earthquake (>150 km) forms a single sub-vertical planar structure along the Pamir. The western Hindu Kush has a tight cluster, while the eastern part shows a planar structure due to slab stretching. Further, the absence of crustal seismicity above the western and central Hindu Kush supports the idea of crust-mantle decoupling caused by slab breakoff beneath the Hindu Kush.

Data availability statement and declaration

STEAD dataset used for training our model is available here: <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 was carried out using the Hypoinverse-2000 (Klein,

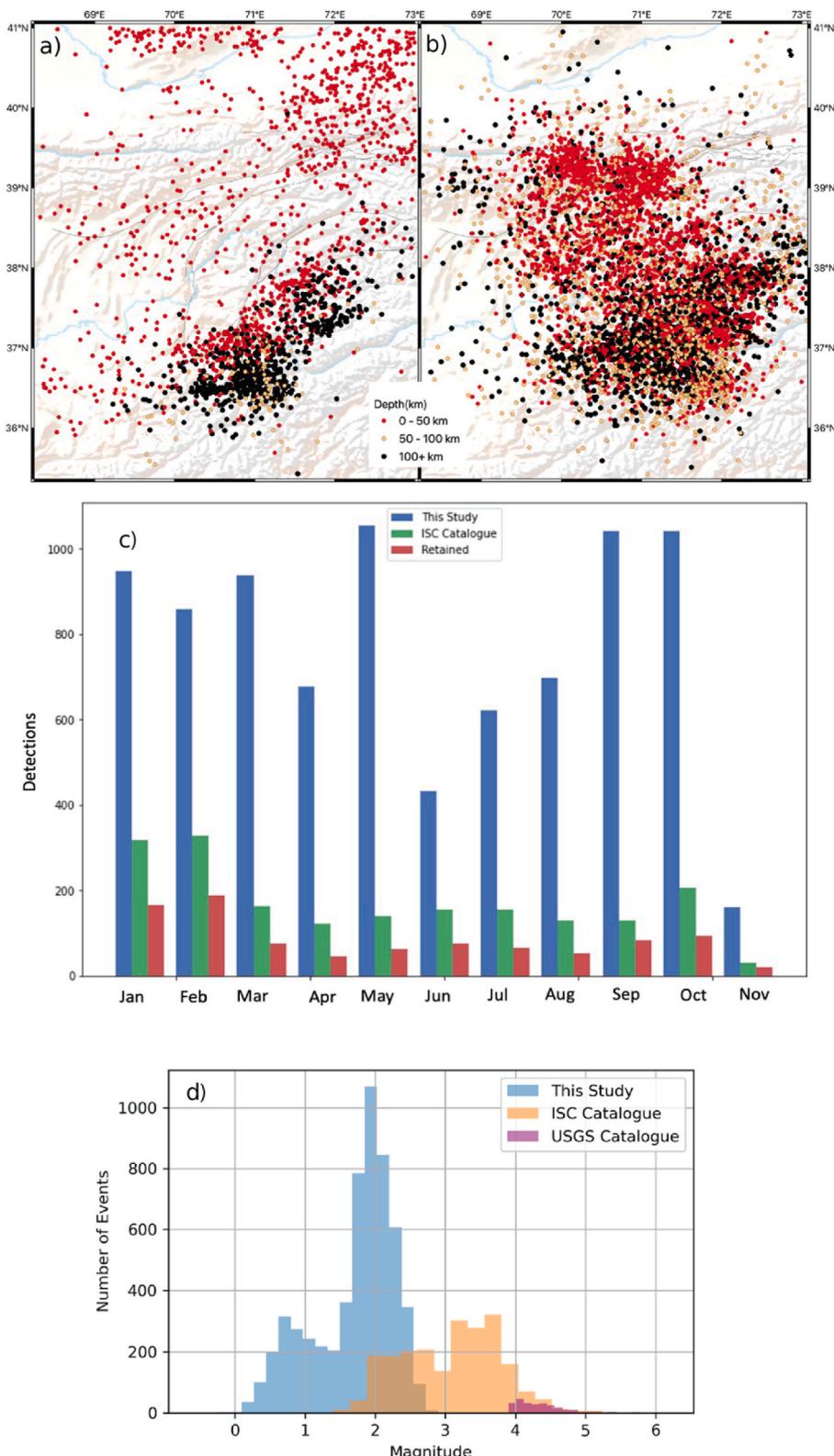


Fig. 7. a) Seismicity Distribution presents the distribution of seismic events as documented in the ISC (International Seismological Centre) catalogue. b) Seismicity Distribution predicted by our model; we showcase the distribution of seismic events as predicted by our computational model. c) Monthly Earthquake Counts Fig. illustrates the monthly count of earthquakes detected by our model, categorized into those also present in the ISC catalogue and those that are common to both our model and the ISC catalogue. d) Earthquake Magnitudes provides a comparison of earthquake magnitudes as determined by our model versus those recorded in the ISC catalogue and the USGS (United States Geological Survey) catalogue. It's noteworthy that our model predominantly detects smaller-magnitude earthquakes, primarily due to the influence of our training data (STEAD), which primarily comprises smaller magnitude earthquake instances.

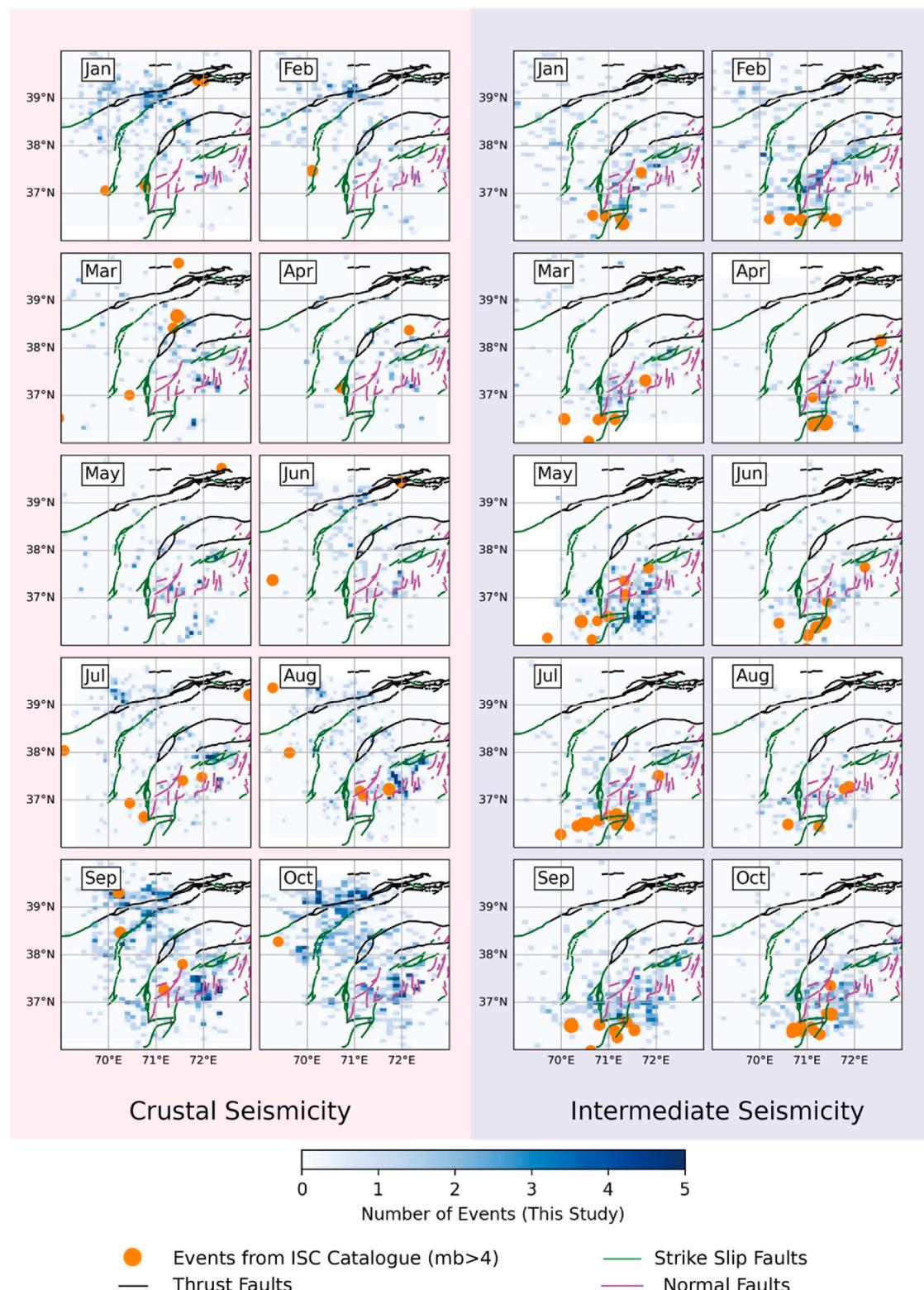


Fig. 8. Spatiotemporal Seismic Activity Trends portrays the changing seismic activity patterns across the Hindu Kush-Pamir region over several months. In January and February, we observe concentrated crustal seismic events around the Vakhsh Thrust System, Darvaz Karakul Fault, and north-northeast-oriented faults in the western Pamir region. From March to May, seismicity decreases in the Vakhsh Thrust System and Darvaz Karakul Fault area, while crustal seismicity rises in western Pamir's north-northeast-oriented faults. Starting in June, seismic activity surges again in the Vakhsh Thrust System and Darvaz Karakul Fault, peaking in September and October. Shallow seismic activity increases in the western Pamir. Our model successfully detects deep earthquakes, closely aligning with the ISC catalog. May shows more intermediate-depth seismic events, with our model identifying lower-magnitude ones in January and February. Importantly, each month exhibits ongoing seismic activity in the intermediate zones of the Hindu Kush and Pamir regions, indicating continuous deformation processes.

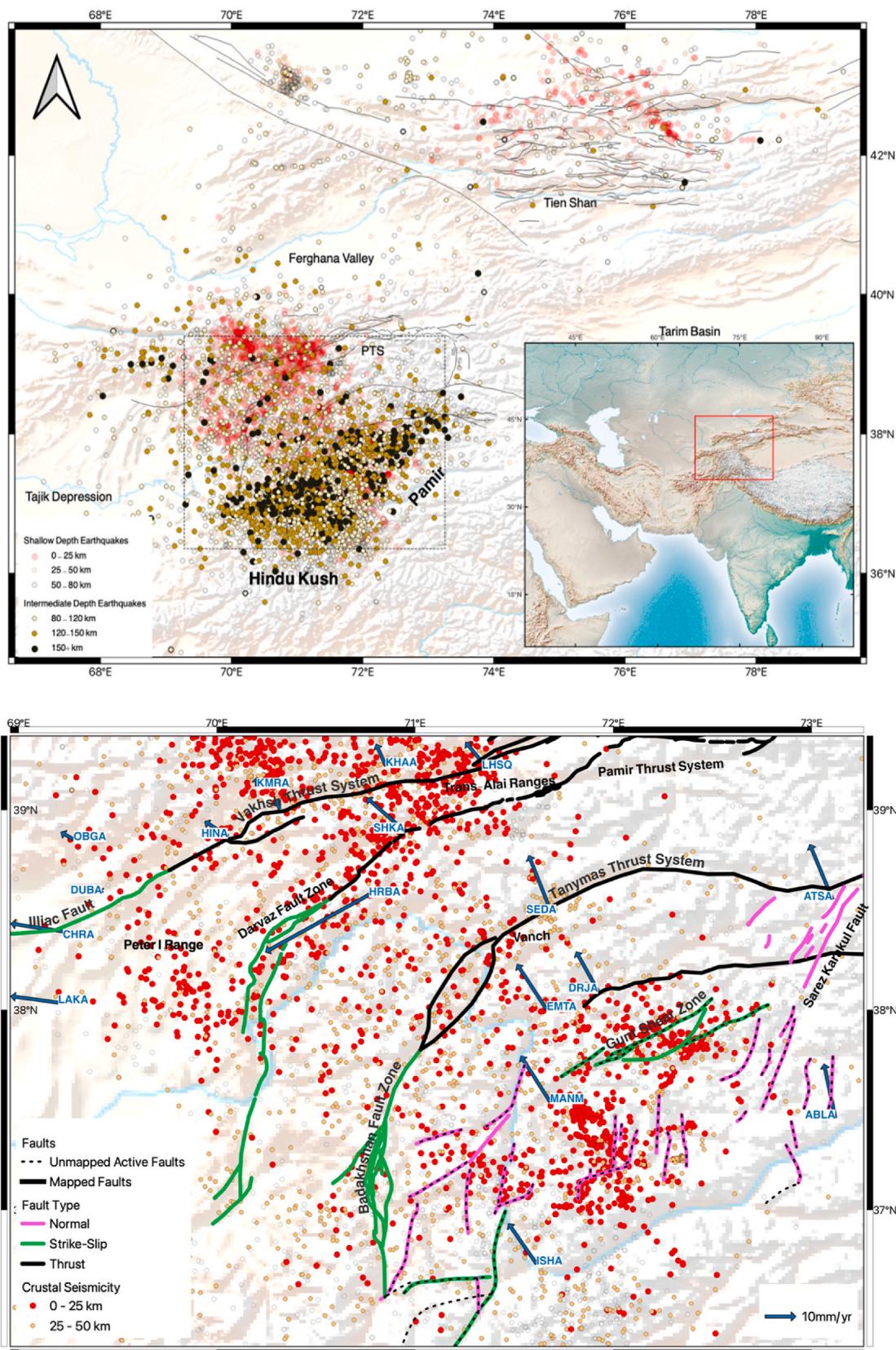


Fig. 9. a) Intermediate seismicity generated in this study. We can see two intermediate seismic zones: the Hindu Kush Zone and the Pamir Zone, and their orientation and different dip direction are profound at a depth greater than 120 km. Also, note that the crustal seismicity in the central part of Hindu Kush is significantly sparse. b) The crustal seismicity in the region shows likely active north-south normal fault and northeast trending strike-slip. There is a change in faulting type and the orientation of GPS velocity showing the two dominant motions causing such a deformation.

2002). Seismicity data for comparison with our results were obtained from the ISC catalogue.

CRediT authorship contribution statement

Satyam Pratap Singh: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Vipul Silwal:** Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by Science and Engineering Research Board (SERB), Department of Science and Technology, under project grant no: SRG/2021/002244. We would also like to thank IIT Roorkee for providing research facility and funds under project grant no: FIG/100840.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jaiig.2023.10.002>.

References

- Allen, R.V., 1978. Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.* 68, 1521–1532. <https://doi.org/10.1785/BSSA0680051521>.
- Bergen Karianne, J., Johnson Paul, A., de Hoop Maarten, V., Beroza Gregory, C., 2019. Machine learning for data-driven discovery in solid Earth geoscience. *Science* 363, <https://doi.org/10.1126/science.aau0323> eaau0323.
- Billington, S., Isaacs, B.L., Barazangi, M., 1977. Spatial distribution and focal mechanisms of mantle earthquakes in the Hindu Kush–Pamir region: a contorted Benioff zone. *Geology* 5, 699–704. [https://doi.org/10.1130/0091-7613\(1977\)5<699:SDAFMO>2.0.CO;2](https://doi.org/10.1130/0091-7613(1977)5<699:SDAFMO>2.0.CO;2).
- Brown, J.R., Beroza, G.C., Shelly, D.R., 2008. An autocorrelation method to detect low frequency earthquakes within tremor. *Geophys. Res. Lett.* 35 <https://doi.org/10.1029/2008GL034560>.
- Dokht, R.M.H., Kao, H., Visser, R., Smith, B., 2019. Seismic event and phase detection using time-frequency representation and convolutional neural networks. *Seismol. Res. Lett.* 90, 481–490. <https://doi.org/10.1785/0220180308>.
- FREIBERGER, W.F., 1963. AN approximate method in signal detection. *Q. Appl. Math.* 20, 373–378.
- Guillot, S., Garzanti, E., Baratoux, D., Marquer, D., Mahéo, G., de Sigoyer, J., 2003. Reconstructing the total shortening history of the NW Himalaya. *G-cubed* 4. <https://doi.org/10.1029/2002GC000484>.
- HAMBURGER, M.W., SAREWITZ, D.R., PAVLIS, T.L., POPANDOPULO, G.A., 1992. Structural and seismic evidence for intracontinental subduction in the peter the first range, central asia. *GSA Bulletin* 104, 397–408. [https://doi.org/10.1130/0016-7606\(1992\)104<0397:SASEFI>2.3.CO;2](https://doi.org/10.1130/0016-7606(1992)104<0397:SASEFI>2.3.CO;2).
- Ischuk, A., Bendick, R., Rybin, A., Molnar, P., Khan, S.F., Kuzikov, S., Mohadjer, S., Saydullaev, U., Ilyasova, Z., Schelochkov, G., Zubovich, A.V., 2013. Kinematics of the Pamir and Hindu Kush regions from GPS geodesy. *J. Geophys. Res. Solid Earth* 118, 2408–2416. <https://doi.org/10.1002/jgrb.50185>.
- Joswig, M., 1990. Pattern recognition for earthquake detection. *Bull. Seismol. Soc. Am.* 80, 170–186. <https://doi.org/10.1785/BSSA0800010170>.
- Klein, F.W., 2002. User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes (USGS Numbered Series No. 2002-171), User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes, Open-File Report. U.S. Geological Survey. <https://doi.org/10.3133/ofr02171>.
- Kong, Q., Trugman, D.T., Ross, Z.E., Bianco, M.J., Meade, B.J., Gerstoft, P., 2018. Machine learning in seismology: turning data into insights. *Seismol. Res. Lett.* 90, 3–14. <https://doi.org/10.1785/0220180259>.
- 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. *Nat. Commun.* 12, 1685. <https://doi.org/10.1038/s41467-021-21760-w>.
- Kufner, S.-K., Schurr, B., Ratschbacher, L., Murodkulov, S., Abdulhameed, S., Ischuk, A., Metzger, S., Kakar, N., 2018. Seismotectonics of the Tajik basin and surrounding mountain ranges. *Tectonics* 37, 2404–2424. <https://doi.org/10.1029/2017TC004812>.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>.
- Lomax, A., Michelini, A., Jozinović, D., 2019. An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network. *Seismol. Res. Lett.* 90, 517–529. <https://doi.org/10.1785/0220180311>.
- Metzger, S., Gagala, Ł., Ratschbacher, L., Lazecký, M., Maghsoudi, Y., Schurr, B., 2021. Tajik depression and greater Pamir neotectonics from InSAR rate maps. *J. Geophys. Res. Solid Earth* 126, e2021JB022775. <https://doi.org/10.1029/2021JB022775>.
- Mousavi, S.M., Beroza, G.C., 2020. A machine-learning approach for earthquake magnitude estimation. *Geophys. Res. Lett.* 47, e2019GL085976 <https://doi.org/10.1029/2019GL085976>.
- Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y., Beroza, G.C., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* 11, 3952. <https://doi.org/10.1038/s41467-020-17591-w>.
- Mousavi, S.M., Sheng, Y., Zhu, W., Beroza, G.C., 2019a. STanford EAarthquake dataset (STEAD): a global data set of seismic signals for ai. *IEEE Access* 7, 179464–179476. <https://doi.org/10.1109/ACCESS.2019.2947848>.
- Mousavi, S.M., Zhu, W., Sheng, Y., Beroza, G.C., 2019b. CRED: a deep residual network of convolutional and recurrent units for earthquake signal detection. *Sci. Rep.* 9, 10267 <https://doi.org/10.1038/s41598-019-45748-1>.
- Pavlis, G.L., Das, S., 2000. The Pamir-Hindu Kush seismic zone as a strain marker for flow in the upper mantle. *Tectonics* 19, 103–115. <https://doi.org/10.1029/1999TC900062>.
- Pegler, G., Das, S., 1998. An enhanced image of the Pamir–Hindu Kush seismic zone from relocated earthquake hypocentres. *Geophys. J. Int.* 134, 573–595. <https://doi.org/10.1046/j.1365-246x.1998.00582.x>.
- Peng, Z., Zhao, P., 2009. Migration of early aftershocks following the 2004 Parkfield earthquake. *Nat. Geosci.* 2, 877–881. <https://doi.org/10.1038/ngeo697>.
- Perol, T., Gharbi, M., Denolle, M., 2018. Convolutional neural network for earthquake detection and location. *Sci. Adv.* 4, e1700578 <https://doi.org/10.1126/sciadv.1700578>.
- Poli, P., Prieto, G., Rivera, E., Ruiz, S., 2016. Earthquakes initiation and thermal shear instability in the Hindu Kush intermediate depth nest. *Geophys. Res. Lett.* 43, 1537–1542. <https://doi.org/10.1002/2015GL067529>.
- Richter, C., 1935. An Instrumental Earthquake Magnitude Scale. *Bulletin of the Seismological Society of America*.
- Ross, Z.E., Meier, M.-A., Hauksson, E., 2018a. P wave arrival picking and first-motion polarity determination with deep learning. *J. Geophys. Res. Solid Earth* 123, 5120–5129. <https://doi.org/10.1029/2017JB015251>.
- Ross, Z.E., Meier, M.-A., Hauksson, E., Heaton, T.H., 2018b. Generalized seismic phase detection with deep learning. *Bull. Seismol. Soc. Am.* 108, 2894–2901. <https://doi.org/10.1785/012018080080>.
- Ruleman, C.A., Crone, A.J., Machette, M.N., Haller, K.M., Rukstales, K.S., 2007. Map and Database of Probable and Possible Quaternary Faults in Afghanistan. U.S. Geological Survey Open-File Report, 1103.
- Schurr, B., Ratschbacher, L., Sippl, C., Gloaguen, R., Yuan, X., Mechic, J., 2014. Seismotectonics of the Pamir. *Tectonics* 33, 1501–1518. <https://doi.org/10.1002/2014TC003576>.
- Searle, M., Hacker, B.R., Bilham, R., 2001. The Hindu Kush seismic zone as a paradigm for the creation of ultrahigh-pressure diamond- and coesite-bearing continental rocks. *J. Geol.* 109, 143–153. <https://doi.org/10.1086/319244>.
- Sippl, C., Schurr, B., Yuan, X., Mechic, J., Schneider, F.M., Gadoev, M., Orunbaev, S., Oimahmadov, I., Haberland, C., Abdybakaev, U., Minaev, V., Negmatullaev, S., Radjabov, N., 2013. Geometry of the Pamir–Hindu Kush intermediate-depth earthquake zone from local seismic data. *J. Geophys. Res. Solid Earth* 118, 1438–1457. <https://doi.org/10.1029/jgrb.50128>.
- Skoumal, R.J., Brudzinski, M.R., Currie, B.S., Levy, J., 2014. Optimizing multi-station earthquake template matching through re-examination of the Youngstown, Ohio, sequence. *Earth Planet Sci. Lett.* 405, 274–280. <https://doi.org/10.1016/j.epsl.2014.08.033>.
- Sobel, E.R., Dumitru, T.A., 1997. Thrusting and exhumation around the margins of the western Tarim basin during the India-Asia collision. *J. Geophys. Res. Solid Earth* 102, 5043–5063. <https://doi.org/10.1029/96JB03267>.
- van Hinsbergen, D.J.J., Kapp, P., Dupont-Nivet, G., Lippert, P.C., DeCelles, P.G., Torsvik, T.H., 2011. Restoration of cenozoic deformation in asia and the size of greater India. *Tectonics* 30. <https://doi.org/10.1029/2011TC002908>.
- Yoon, Clara E., O'Reilly, Bergen Karianne, J., Beroza Gregory, C., 2015. Earthquake detection through computationally efficient similarity search. *Sci. Adv.* 1, e1501057 <https://doi.org/10.1126/sciadv.1501057>.
- Zhan, Z., Kanamori, H., 2016. Recurring large deep earthquakes in Hindu Kush driven by a sinking slab. *Geophys. Res. Lett.* 43, 7433–7441. <https://doi.org/10.1002/2016GL069603>.
- Zubovich, A.V., Wang, X., Scherba, Y.G., Schelochkov, G.G., Reilinger, R., Reigber, C., Mosienko, O.I., Molnar, P., Michajlow, W., Makarov, V.I., Li, J., Kuzikov, S.I., Herring, T.A., Hamburger, M.W., Hager, B.H., Dang, Y., Bragin, V.D., Beisenbaev, R. T., 2010. GPS velocity field for the Tien Shan and surrounding regions. *Tectonics* 29. <https://doi.org/10.1029/2010TC002772>.