Machine learning in microseismic monitoring

Anikiev, Denis; Birnie, Claire; Waheed, Umair bin; Alkhalifah, Tariq; Gu, Chen; Verschuur, Dirk J.; Eisner, Leo

DOI
10.1016/j.earscirev.2023.104371

Publication date
2023

Document Version
Final published version

Published in
Earth-Science Reviews

Earth-Science Reviews 239 (2023) 104371

Contents lists available at ScienceDirect

Earth-Science Reviews

journal homepage: www.elsevier.com/locate/earscirev

Machine learning in microseismic monitoring

Denis Anikiev [a,*], Claire Birnie [b], Umair bin Waheed [c], Tariq Alkhalifah [b], Chen Gu [d], Dirk J. Verschuur [e], Leo Eisner [f]

[a] Section 4.5: Basin Modelling, Department 4: Geosystems, Helmholtz Centre Potsdam – GFZ German Research Centre for Geosciences, Telegrafenberg, Potsdam, 14473 Brandenburg, Germany
[b] King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia
[c] Department of Geosciences, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
[d] Tsinghua University, Haidian District, Beijing 100084, China
[e] Delft University of Technology, Postbus 5, Delft 2600, The Netherlands
[f] Seismik s.r.o., Kubišova 1265/8, Prague 182 00, Czech Republic

ARTICLE INFO

Keywords:
Microseismic monitoring
Machine learning
Neural networks
Induced seismicity
Passive seismic
Earthquake early warning

ABSTRACT

The confluence of our ability to handle big data, significant increases in instrumentation density and quality, and rapid advances in machine learning (ML) algorithms have placed Earth Sciences at the threshold of dramatic progress. ML techniques have been attracting increased attention within the seismic community, and, in particular, in microseismic monitoring where they are now being considered a game-changer due to their real-time processing potential. In our review of the recent developments in microseismic monitoring and characterisation, we find a strong trend in utilising ML methods for enhancing the passive seismic data quality, detecting microseismic events, and locating their hypocenters. Moreover, they are being adopted for advanced event characterisation of induced seismicity, such as source mechanism determination, cluster analysis and forecasting, as well as seismic velocity inversion. These advancements, based on ML, include by-products often ignored in classical methods, like uncertainty analysis and data statistics. In our assessment of future trends in ML utilisation, we also see a strong push toward its application on distributed acoustic sensing (DAS) data and real-time monitoring to handle the large amount of data acquired in these cases.

1. Introduction

As we listen to the Earth's heartbeats and analyse the sources for these beats, we gain more information on the Earth's workings and its content. On the global scale, studying earthquakes has been instrumental in understanding the Earth, its dynamic and makeup properties. On the regional scale, listening to induced earthquakes, mostly known as microseismic events (although some events are not "quite" micro), allows us to monitor subsurface projects like oil and gas production; hydraulic fracturing for unconventional resources, such as geothermal energy; or the reaction of the Earth's crust to impoundment and storage of water in dams.

In the last decade, machine learning (ML) – a field of knowledge originating from artificial intelligence (AI) and computer science – has become an important interdisciplinary numerical tool that has advanced science in general and geoscience in particular (Lary, 2010; Karpatne et al., 2019; Dramsch, 2020). Recent advances in medical imaging, akin to seismic imaging, have demonstrated state-of-the-art performance of ML on tasks that were once considered almost unique to humans (Barragán-Montero et al., 2021; Greiner, 2022). ML techniques are becoming highly widespread in the geophysical community (e.g., Kong et al., 2019) and attract increasing attention in microseismic imaging, with applications ranging from detection of weak signal signatures and patterns to extraction of features that help to improve our physical understanding of the related phenomena.

Several review articles summarise ML applications in various fields of geoscience and reveal the growing importance of AI methods; for instance, in solid earth geoscience (Bergen et al., 2019), in fault rupture studies (Ren et al., 2020), in earthquake prediction (Mignan and Broccardo, 2020), in seismology (Kong et al., 2019; Mousavi and Beroza, 2022), as well as in geoscience (Dramsch, 2020) and geophysics (Yu and Ma, 2021) in general.

The goal of this review is similar to the other reviews, but here we aim to acquire a deeper understanding of the use of ML in applications

* Corresponding author.
E-mail address: denis.anikiev@gfz-potsdam.de (D. Anikiev).

https://doi.org/10.1016/j.earscirev.2023.104371
Received 18 August 2022; Received in revised form 6 February 2023; Accepted 27 February 2023
Available online 5 March 2023

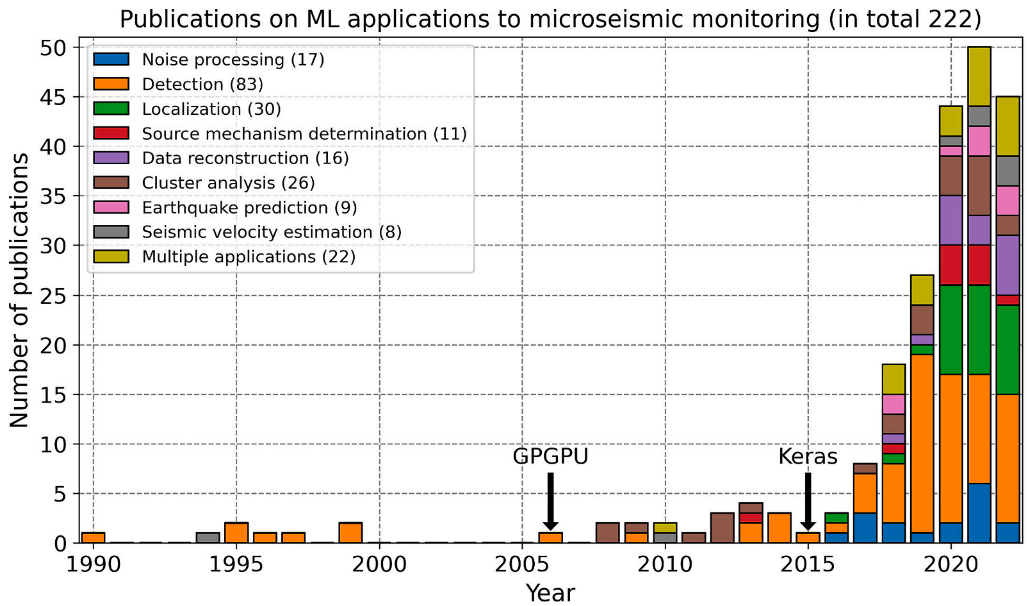

Fig. 1. Peer reviewed articles, conference abstracts, and theses on ML applications published in 1990–2022 that are related to microseismic monitoring, in the form of colour-coded histograms. Bars are categorised by the field of application. Legend entries show the appropriate fields of application (or combination) and the total number of publications across the years.

developed for microseismic monitoring and analysis. The ML applications in microseismic monitoring are generally similar to earthquake monitoring, but there are several important aspects in which they differ. One of the most important factors here is that microseismic monitoring mainly deals with weak seismic signals (e.g., Duncan and Eisner, 2010; Foulger et al., 2018). This often results in the absence of visible signals on individual receivers and requires array processing, which is not usually applied in classical earthquake seismology. In addition, induced seismicity/microseismicity depends on activities in the region of interest. Often, microseismic monitoring starts with the onset of these activities, and thus, past seismic events that could be valuable, for example, for training an ML algorithm, may not exist in the study area.

Microseismic monitoring, in this review, is defined as the analysis and acquisition using specially designed monitoring arrays to record seismic waveforms of weak microseismic events. These events are typically induced by local subsurface processes or more specifically, subsurface anthropogenic activities. The usual goal of microseismic monitoring is to detect, locate and characterise these weak microseismic events (i.e., a cluster of them) to provide geometric and more general information about the considered subsurface processes (Warpinski, 2009). Part of the focus on weak events arises from the Gutenberg-Richter Law stating that the number of events grows exponentially with decreasing magnitude. As such, by tackling lower magnitude events, the volume of analysed events significantly increases, and with it does our ability to extract knowledge of subsurface processes.

The majority of ML techniques discussed herein require special neural networks and utilise monitoring arrays, such as borehole arrays or dense surface arrays over the area of underground operations. Microseismic imaging is used to detect as weak signals as possible in real- or near-real-time – a rare challenge in earthquake seismology. Moreover, it is used to provide better insights into underground processes, for example, optimization of the hydraulic fracturing process during injection, as well as for real-time risk evaluation of induced seismicity. ML in microseismic monitoring is considered a game-changer because many applications require real-time processing, e.g., in traffic light systems for induced seismicity (Foulger et al., 2018). Again, this is different from earthquake seismology, where earthquake early warning (EEW) systems face different challenges (generally focused events with high signal-to-noise ratio (SNR)). The speed and consistency of ML algorithms is a big advantage for induced seismicity analysis.

Taking into account all these factors, this paper summarises major ML applications in microseismic monitoring, discusses their advantages and limitations, and explains the importance of this technology for future developments. Last but not least, for new researchers working at the intersection of ML and microseismic monitoring, this article should provide a holistic overview for this emerging domain of research, as well as provide a vision where ML in microseismic monitoring will head in the future.

While preparing this review, we collected a database of 222 publications (peer reviewed articles, conference abstracts and theses) dealing with machine learning applications related to microseismic monitoring. The number of publications by year (Fig. 1) illustrates that the idea of using ML in microseismic monitoring has existed for a long time, but initially a combination of slow algorithms and limited computer capacity did not allow for practical applications of ML in general, especially in microseismic monitoring. The term graphical processing unit (GPU) was in use since the 1980s and got popularized in 1999 (Peddie, 2022), but the first unified graphical and computing GPU architecture programmed in C with CUDA was released only in 2006 (Nickolls and Dally, 2010). This period can be regarded as a start of the era of General-Purpose GPU (GPGPU, marked by the arrow in Fig. 1), when many researchers and developers started to enthusiastically adopt CUDA and GPU computing for a diverse range of applications requiring massive vector operations and mathematically intensive problems in industry, finances and science (Nickolls and Dally, 2010; Peddie, 2022). The rapid development of GPU-dedicated libraries and tools and the increasing availability of GPUs during the last decade preceded exponential growth in the number of research articles on ML in microseismic monitoring. Another important factor supporting the exponential growth of publications since 2015 has been the first public release of the Keras deep learning library (Chollet et al., 2015), also indicated by the arrow in Fig. 1.

Colour bars in Fig. 1 display different fields of microseismic monitoring applications. It is generally consistent with early stages of acceptance of any new technology, where the first applications were focused on the fundamental processes of microseismic monitoring such as detection and localisation; whereas development of more advanced processing methods targeting noise suppression/classification, source

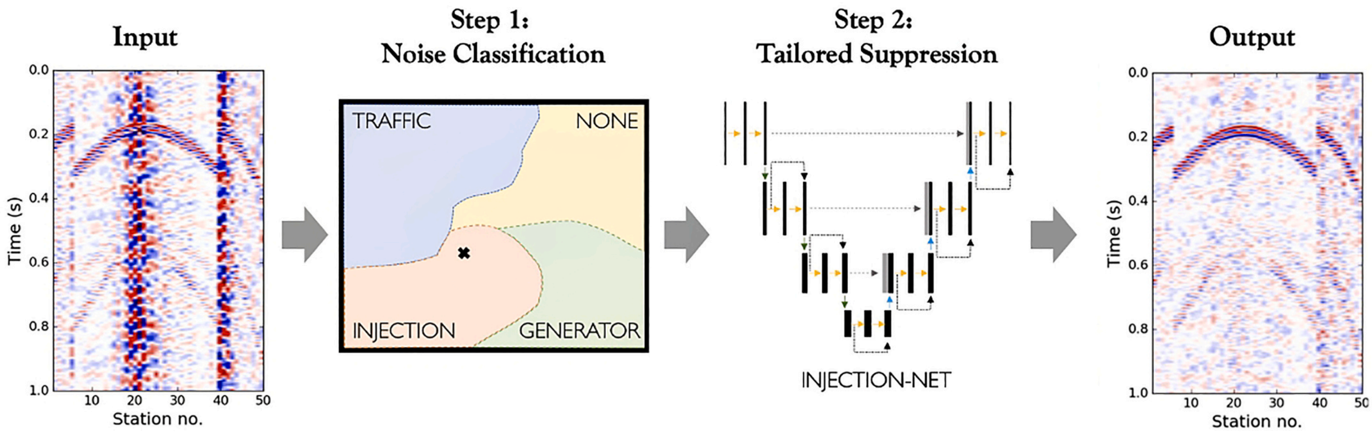

Fig. 2. A conceptual ML workflow for noise classification using clustering prior to tailored noise suppression using a neural network. In this example, the data is contaminated by injection noise, which is identified via a classification scheme and illustrated within the feature space, then subsequently suppressed using a tailor-made neural network termed 'injection-net'. The final product is that injection noise has been suppressed from the input data.

mechanism determination, and cluster analysis started later. We performed subjective evaluation of the research articles to categorise our database. Thus, some publications were targeting multiple applications without a specific focus, sometimes also in the form of a review. We accordingly marked these articles in the histogram in Fig. 1 as "Multiple applications" category.

In this article, we review recent developments in utilising ML in microseismic monitoring and inversion. After this introduction, Section 2 represents an overview of machine-learning applications related to microseismic monitoring. We start the overview with pre-processing applications adopting ML techniques used to enhance the SNR considering the weak signals, namely dealing with noise classification and suppression (Section 2.1), and wavefield reconstruction (Section 2.2). Microseismic monitoring starts with detection of microseismic events, and the role of ML in this task is reviewed in Section 2.3. Section 2.4 investigates the applications for localisation of hypocentres of microseismic events, followed by Section 2.5 which provides an overview of applications for the determination of source mechanisms of microseismic events. Next, Section 2.6 targets post-processing applications for cluster analysis of seismicity. Prediction of earthquakes, another hot topic, is discussed in Section 2.7. The accuracy of locating the microseismic events and characterising them depends heavily on the accuracy of the velocity model. Therefore, we review the importance of having an accurate description of the subsurface in which monitoring is performed, and ML techniques used to determine such a description from passive seismic data in Section 2.8. After this overview that forms the main content of this paper, Section 3 summarises the current best practices and discusses the outstanding challenges and future directions before concluding the review.

The list of acronyms in Appendix A provides a list of common abbreviations used in the text together with their meanings in the form of a glossary.

2. Overview of applications

In this section, we provide an overview of the applications of machine learning applied to microseismic monitoring (detection, localisation, and determination of source mechanism of microseismic events), ranging from data pre-processing (noise classification and suppression, and wavefield reconstruction) to post-processing (e.g., source mechanisms determination and cluster analysis), including seismic velocity inversion based on microseismic data analysis. The review is organised by the different microseismic applications.

2.1. Noise classification and suppression

Noise is broadly defined as anything other than the desired signal. Therefore, in the case of microseismic monitoring noise is often referring to ambient noise caused by the environment or anthropogenic sources. However, depending on our interest it may also refer to other 'seismic' events that are not related to the microseismic objectives. For example, data collected from the Aquistore carbon capture and storage (CCS) permanent monitoring array was contaminated by nearby pot-ash mining activity, yet the seismic events of interest for this monitoring setting would have been induced seismicity with respect to the CO_2 injection (Stork et al., 2018). For microseismic monitoring purposes, noise is particularly troublesome due to the events' characteristically low-magnitudes resulting in low signal-to-noise ratios in seismic recordings. As such, accurate classification and suppression of noise are fundamental for successful microseismic monitoring. Machine learning has been shown to be particularly powerful for pattern recognition tasks, therefore it naturally adapts itself to noise classification and subsequent suppression; which combined could lead to a full, ML-based noise classification and suppression pipeline as illustrated in Fig. 2. Below, we detail a number of proposed ML approaches for the classification and suppression of noise in passive seismic data.

Whilst not utilising the ML component in the suppression of noise, the use of ML for distinguishing between different noise types present within a seismic recording allows for the automatic application of tailored noise suppression algorithms (Birnie, 2018). For microseismic recordings, where the signal is considerably weak, tailored noise suppression allows us to take a conservative approach. This makes most sense for microseismic signals as we want to maintain a fine balance between damaging the intended signal versus leaving some noise. Outside of the field of microseismic monitoring, clustering methods have proved popular for early noise classification. Galvis et al. (2017) examined the myriad of features used in previous studies for detection, characterising and suppression of surface waves and utilised a k-means clustering approach to help identify the best combination of previously determined features. Huot et al. (2017) also use previously established approaches to determine noise features prior to using a clustering procedure to classify noise types within their data. However, unlike the large range of features considered by Galvis et al. (2017), Huot et al. (2017) only use 100 continuous wavelet transform scale factors as their data features, which they then feed into a hierarchical clustering procedure to determine the noise types present within a seismic window. Similar to Galvis et al. (2017), Johnson et al. (2020a) also utilised k-means clustering applied onto waveform features to classify 1 s recordings of seismic data into different noise classes. These studies showed promising results for noise classification through the use of clustering approaches. More recently, with their rise in popularity, neural network applications have been proposed as an alternative means to clustering. For example, applied to a fibre optic recording for interferometry purposes, Huot et al. (2018) train a NN for the detection of traffic signals, that are subsequently suppressed by muting high frequency components after the application of Continuous Wavelet Transforms (CWTs). Similarly, Snover et al. (2021) utilised conventional auto-encoders to identify urban noise signals, such as airport traffic, within 161 h of seismic data recorded on an array of 5200 geophones.

In the context of microseismic monitoring, until now little has been published on the use of ML for pure noise classification, i.e., distinguishing between different noise signals within the data. The majority of microseismic 'noise' classification schemes focus on classification between noise and signal, which will be later discussed in the event detection section. Considering the pure noise classification task, the aforementioned NN procedures are supervised and therefore require noise labels associated with the training data – a limitation the clustering algorithms do not have. However, as discussed by Huot et al. (2018), clustering algorithms typically require seismic attributes, such as coefficients from CWTs, which are computationally expensive to compute, hindering the real-time applicability of clustering approaches. As such, Huot et al. (2018) proposes the use of such clustering procedures to build the labelled training dataset (in a pseudo-unsupervised manner) prior to training a NN model for detection on the raw, noisy seismic signals, resulting in a noise classification scheme that can be applied real-time.

Microseismic data typically undergoes substantially less processing than active seismic data due to a preference to keep noise rather than permit any signal leakage, which would alter the already low magnitude seismic signal. Despite this, a handful of ML approaches have been proposed for noise attenuation in seismic data. Working on a trace-by-trace basis, Zhu et al. (2019b) propose the use of U-Net-style architecture (styled after Ronneberger et al. (2015)), termed DeepDenoiser, that takes the real and imaginary amplitude spectra as input and produces noise and signal masks that are subsequently applied to the noisy field data. Trained on previously recorded high SNR seismic events with additive field noise, the authors advocate the benefits that DeepDenoiser could bring to microseismic monitoring after its application to a number of low SNR earthquake arrivals. If the regional noise for the array to be processed is added into the training dataset, the suppression quality can be significantly improved, even with the same network architecture, by the means of transfer learning. For example, Yang et al. (2022) tuned DeepDenoiser with the rich noise sources recorded by the urban dense array to create a model called UrbanDenoiser. The method achieves a better denoising performance compared to the original one and further reveals the mantle seismicity beneath Los Angeles. Zhang et al. (2020b) also utilise a U-Net-style architecture for the prediction of a noise mask to subsequently be applied to the noisy data. Unlike Zhu et al. (2019b), the training data employed by Zhang et al. (2020b) is semi-synthetic, composed of noise collected in the field and synthetically generated waveform data. Alongside the noise mask prediction, the network is also trained to detect the duration of the seismic arrivals. Alternatively, an advanced unsupervised deep learning approach has been proposed by Saad et al. (2022a). In contrast to the supervised methods, it does not require any labelled data. The method utilises time–frequency representation of seismic records after the short-time Fourier transform (STFT) and a customised loss function to reconstruct the signal binary mask by a DL network. The proposed method shows a robust denoising performance and outperforms the benchmark supervised denoising method of Zhu et al. (2019b).

Another challenge particularly faced in the attenuation of noise in microseismic data is the unknown nature of the seismic noise field. Noise in seismic data is inherently complex due to the range of natural and anthropogenic sources from which it originates, resulting in an overall noise field that exhibits random and coherent components, as well as stationary and non-stationary elements. This complexity is further amplified in the microseismic context due to the high volume of noise in comparison to the microseismic signals (i.e., due to the characteristically low SNR of microseismic data). Accurately recreating such a complex noise field is non-trivial (Birnie et al., 2016) – this makes the use of synthetic datasets for training much less favourable than in active seismic scenarios. Due to the challenges of selecting an appropriate training dataset Saad et al. (2021) and Liu et al. (2021) have both chosen approaches that do not require noisy-clean pairs of training data. Saad et al. (2021) utilise a Variational AutoEncoder (VAE), where the latent features related to the microseismic signals are extracted by the encoder, and the decoder reconstructs noise-free microseismic data. Illustrated on a field data example from hydraulic fracturing of the Marcellus gas shale in Pennsylvania, the proposed VAE procedure is shown to outperform two conventional denoising techniques - namely, *fx*-deconvolution and damped multichannel singular spectrum analysis. Liu et al. (2021) also directly used the noisy data for training a blind-spot denoising scheme. By 'blinding' the network to a central pixel's value, the network must learn to use neighbouring samples to predict the blinded pixel's value. Under the assumption noise is independent across neighbouring samples, the blind-spot network cannot learn to replicate the noise component of blinded pixel's noisy value and therefore only predicts the signal component. The proposed procedure was benchmarked on a synthetic dataset with random noise and a semi-synthetic dataset with recorded noise, prior to being successfully applied to field data. In contrast to these approaches, Birnie and Alkhalifah (2022) leverage domain adaptation procedures to incorporate features from unlabelled field data into the supervised training of a network using noisy-clean pairs of synthetic data. Through a series of correlations and convolutions, the authors illustrate how a network can be trained on pairs of synthetic data contaminated with coloured, Gaussian noise and later applied to a field dataset exhibiting varying degrees of coherent noise.

Finally, whilst the denoising procedures themselves here are not derived from ML-procedures, a number of authors have considered the inclusion of conventional denoising techniques into their ML training workflows, where the ML component focuses on another step in the microseismic processing pipeline such as event detection. Li et al. (2021) illustrate an 'ML' workflow for seismic denoising and detection where they implement a Graph-based Bilateral Filter (GraphBF) for denoising (after a conventional bandpass filter) prior to computing seismic features to be fed into a random forest to determine if an event is present or not. In contrast, Othman et al. (2021) apply an infinite impulse response (IIR) Wiener filter-based denoising procedure only after an event has been detected by their pre-trained recurrent neural network (RNN). The detection and denoising are then performed as an iterative scheme, to enhance the microseismic signal prior to inputting into subsequent tasks, like moment tensor analysis.

2.2. Wavefield reconstruction and interpolation

The seismic wavefield recorded by geophones during seismic acquisition is regularly sampled in time, but it is often non-uniformly sampled in spatial directions. This leads to irregularly and sparsely populated seismic wavefield representations.

Active seismic acquisition suffers from large gaps between recording profiles, e.g. between consecutive sail lines in 3D marine surveys (e.g., Greiner et al., 2021). In passive seismic acquisition the sparsity and irregularity of geophone distribution can result from environmental, sociological or political reasons, but mostly due to economic limitations as well as road accessibility. Regularly distributed dense surface arrays are significant in microseismic monitoring, especially in post-processing phase, e.g. for constraining non-double-couple source mechanisms of induced events (e.g., Pesicek et al., 2016) or quantifying fracture networks from microseismic event clouds (e.g., McKean et al., 2019). Reconstruction or regularisation of seismic wavefields to a dense and regular grid (which also implies interpolation and extrapolation) is an

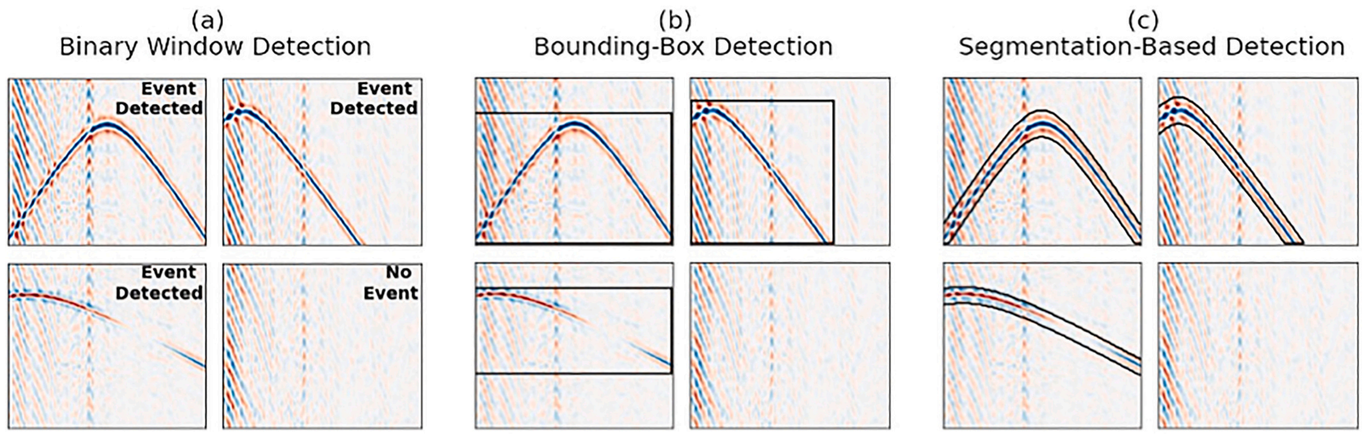

Fig. 3. A schematic illustration of computer vision techniques applied to surface microseismic data for (a) binary event detection, (b) windowed event detection, and (c) segmented detection of event arrivals.

essential step to improve imaging resolution (Yu and Ma, 2021).

Wavefield reconstruction can be regarded as a super-resolution problem (e.g., Li et al., 2022a) and/or as a compressed sensing problem (e.g., Huang et al., 2022), at the same time sharing many similarities with image processing (Greiner, 2022), where AI techniques, and especially deep learning, play a key role (e.g., Yang et al., 2019).

Most of the ML-based applications related to wavefield reconstruction deal with pre-stack seismic data gathered with active acquisition. Among many other publications, a deep neural network (DNN) was proposed for reconstruction of regularly missing seismic data by Wang et al. (2019a). They performed training using a combination of synthetic and field data and achieved better regularisation results for regular seismic data when compared to the classical f-x prediction. An interpolation algorithm originating from image restoration, the projection onto convex sets (POCS) method has been introduced by Abma and Kabir (2006) for field seismic data regularisation. Zhang et al. (2020) considered CNNs in the context of the POCS framework for seismic interpolation and showed that the CNN denoisers pretrained on natural images could efficiently improve the seismic interpolation results. In the same track of using POCS and DL to interpolate and denoise the seismic data, DenseNet and POCS have been proposed to reconstruct the 3D seismic data in an unsupervised scheme by Saad et al. (2022b). In this study, the authors reconstruct the data despite the large number of irregularly missing traces.

Although the aforementioned methods (as well as many others related to seismic processing using ML, as recently discussed by Mousavi and Beroza (2020)) are considered in the active acquisition context, their basic ideas can be applied to passive seismic data interpolation and reconstruction.

In the passive data context, Zhang and van der Baan (2020) proposed to use an unsupervised ML method to recover the microseismic signal from the noisy data with missing traces based on dictionary learning. In their non-parametric model, the dictionary is learned from the observed data without the need to subdivide data into training and evaluation sets.

In the work of Greiner et al. (2021), who investigated unsupervised deep learning based on a CNN for multidimensional wavefield reconstruction of irregularly populated traces, the problem of reconstructing the complete seismic wavefield from a coarsely sampled and incomplete seismic records is formulated as an underdetermined inverse problem. On the other hand, the wavefield reconstruction can be regarded in the context of a partial differential equation (PDE) optimisation as a regularisation term to partially mitigate the nonlinearity of this seismic inverse problem. The so-called efficient wavefield reconstruction inversion (WRI) that aims to mitigate cycle skipping in full-waveform inversion (FWI) via physics-informed neural networks (PINNs) was proposed by Song and Alkhalifah (2020b); Song and Alkhalifah, 2022. The key objective of WRI here is to reconstruct a frequency-domain wavefield that fits the data (specifically scattered seismic data) and the wave equation (Song and Alkhalifah, 2020b), so the workflow can potentially be extended to make use of microseismic data as well.

One of potential trends for the wavefield construction and interpolation of passive seismic data is the adoption of deep generative models. Recently Gan et al. (2022) proposed the EWR-Net – a deep generative model for earthquake waveform regularisation, and showed that the method obtains a higher SNR than a curvelet-based method. They demonstrated that generative network design is feasible for the regularisation of irregular station data, which is often the case in microseismic monitoring.

2.3. *Event detection and picking*

Accurate event detection and arrival picking are longstanding challenges in the field of microseismic monitoring. Typically of low magnitude, microseismic events are often masked by noise, making visual detection impossible. Alongside this, with the ever increasing volume of microseismic data being collected, performing the task manually is unfeasible and, therefore, requires automated algorithms. Among a number of algorithms developed over the years, STA/LTA (Allen, 1978) and template matching (Gibbons and Ringdal, 2006) are the most widely used. While STA/LTA is generalised and computationally efficient, its lack of sensitivity to time-varying noise and low-magnitude events, as well as its strong dependence on the selection of parameters (e.g., triggering and de-triggering thresholds, short and long time intervals), makes it less effective for robust event detection, especially for microseismicity. Although template matching does not suffer from these limitations and can robustly detect smaller events, its major drawbacks are insensitivity to events with waveforms that are dissimilar to the master event as well as a high computational cost, severely limiting its use for real-time microseismic event detection (Yoon et al., 2015). In the instances when noise levels obscure arrivals, stacking has been shown to significantly improve the SNR of the data and aid event detection, as well as subsequent tasks such as migration (Duncan and Eisner, 2010).

With the growing interest of the seismological community in using advances in the field of machine learning (ML), numerous efforts have been made in the recent past to improve event detection and picking capabilities. Below, we outline some of the most promising advances on the topic.

The idea to use machine learning for seismic event detection and picking has been around since the 1990s. A number of methods were proposed based on artificial neural networks using features extracted from the recorded waveforms (Dai and MacBeth, 1995; Wang and Teng, 1995; Musil and Plešinger, 1996; Fedorenko et al., 1999; Tiira, 1999; Gentili and Michelini, 2006; Maity et al., 2014; Akram et al., 2017; Qu et al., 2020).

These algorithms demonstrated improved robustness in detection and arrival picking for low-magnitude events. More recently, a data mining approach, referred to as Fingerprint and Similarity Thresholding (FAST) (Yoon et al., 2015), was developed to reduce the computational complexity of template matching. FAST converts an entire dataset into binary fingerprints, which compactly represent short segments of a continuous waveform. These fingerprints are organised in a special dictionary structure for faster lookup, allowing efficient processing of large datasets. With the fast emergence of deep learning as a disruptive tool to tackle longstanding research problems across science and engineering disciplines, a number of supervised deep learning methods were also recently developed for event detection and arrival picking. Such techniques can be divided based on the volume of information returned. For example, whether the algorithm provides information on the presence of an event within a given window of data or it returns the arrival time of the event at a specific time value. Fig. 3 illustrates this concept for different computer vision methodologies that have been employed for microseismic detection and arrival picking on microseismic data collected on a surface array.

A number of procedures have been proposed for binary window classification. Such approaches involve a time window being passed into a machine learning model and the model returns the prediction of the presence or absence of an event in the window. Wilkins et al. (2020) developed a custom CNN model for detecting microseismic events related to mining. They showed that their trained CNN model was able to detect ten times more events than that found by a human expert. While these methods focused on detecting events by considering waveforms at a single geophone, Shaheen et al. (2021) trained a CNN model to accurately detect low-magnitude events by using the entire shallow borehole network at Groningen, the Netherlands. This allowed them to train the CNN model effectively using moveout patterns of energy travelling across the borehole sensors to discriminate between events originating in the subsurface and local noise arriving from the surface. For earthquake detection, Ross et al. (2018b) developed a generalised phase detection (GPD) method by training a CNN model on hand labelled dataset from the Southern California Seismic Network to classify windows of data as belonging to P-wave, S-wave, or noise. GPD has been shown to robustly detect P- and S-waves for low-magnitude events and is applicable to a range of datasets recorded in different tectonic regimes. This is particularly useful for cases when no seismicity catalogue exists to be used for template matching.

For array-based methods, an extension of the binary window classification is bounding-box detection which detects the time range and observed offset of an events arrival, as illustrated in Fig. 3(b). Prior to the take-off of neural network approaches, Horne et al. (2019) utilised Haar cascades (HC) for detecting microseismic arrivals. HC utilise adjoining rectangles to identify features within an image, and have been heavily used in face-detection tasks. Horne et al. (2019) illustrated that, after minimal signal conditioning, they could accurately detect areas containing microseismic arrivals across 3800 DAS channels for a 4 h recording. The identified events have a close correspondence with known hydraulic fracturing activities. Focusing on the same DAS recording, Stork et al. (2020) advances on the work of Horne through their consideration of neural networks, in particular the common CNN architecture called YOLOv3. Trained on synthetic data which has been contaminated by field noise (i.e., a semisynthetic dataset), the trained YOLO network successfully detects $>$ 80% of manually detected events and the majority of missed events are low SNR. The authors speculate the network's performance could be improved further with the inclusion of more low SNR events during training.

For trace-based detection, both non-neural network and neural network methods have been shown to outperform conventional procedures. Benchmarked against the common STA/LTA trigger, Chen (2020) developed an unsupervised fuzzy clustering procedure that utilised the STA/LTA trigger, alongside power and mean computed over the window, as features in their clustering algorithm. The model is shown to be robust to high noise levels and accurately detects microseismic events in both synthetic and field data. For trace-based neural network methods, Zheng et al. (2017) and Birnie and Hansteen (2022) both utilised the long-short-term-memory (LSTM) recurrent neural network architecture. Zheng et al. (2017) trained their LSTM network using lab-stimulated microseismic events and successfully applied the trained network to field recordings. Birnie and Hansteen (2022) adapted the architecture of Zheng et al. (2017) to include bidirectionality and trained their network on a wide-ranging synthetic dataset contaminated by bandpassed noise. The trained network was successfully applied on two different field datasets: one land and one from ocean bottom nodes. The successful application on the significantly different field datasets highlights the robustness of the network due to the large range of synthetic events it was exposed to during training.

In a similar vein, Mousavi et al. (2019) developed a CNN-RNN earthquake detector (CRED) that combines convolutional layers and bi-directional LSTM units in a residual structure for robust detection of microearthquakes. They showed improved performance of CRED compared to STA/LTA, template matching, and FAST algorithms in lowering the detection threshold while minimising false positive detections. As opposed to considering detection as a time-series task, Birnie et al. (2021) considered it a semantic segmentation task for array recordings, where the seismic recording is considered as an image and the network determines if an event is, or is not, present at each pixel location. Utilising the popular U-Net architecture of Ronneberger et al. (2015) and trained on synthetic data, the authors utilised distributed deep learning to handle the irregular geometry of their ocean bottom node network and illustrated how the resulting model could be applied real-time on a single GPU.

In addition to event detection, a number of deep learning based algorithms were developed for phase picking. Ross et al. (2018a) trained a CNN model on a hand-picked dataset from the Southern California Seismic Network to pick P-wave arrival and first-motion polarity. They showed that the trained model could perform the task more accurately than a professional seismic analyst. Zhu and Beroza (2019) trained a fully convolutional network, referred to as PhaseNet, for arrival time picking of P- and S-waves. They demonstrated state-of-the-art picking capabilities, even in challenging circumstances for a human analyst. Another end-to-end deep learning approach for picking P and S arrivals was proposed by Wang et al. (2019b). Their deep CNN model was trained on manually picked three-component seismograms from the high-sensitivity seismic network deployed on the islands of Japan. In an effort to pick the phase of microseismic events from laboratory-scale hydraulic fracturing experiments, Chai et al. (2020) construct a deep NN initially trained on a large amount of events from global earthquakes, followed by transfer learning towards the scale-model data. They show that the new workflow provides a better seismic catalogue and a larger amount of phase picks compared to human seismic analysts. Johnson et al. (2020b) applied the CNN model of Ross et al. (2018a) on a microseismic dataset related to mining-induced seismicity. They showed remarkable performance through fine-tuning using their dataset-specific training. Moreover, He et al. (2021) showed the application of capsule network for P arrival picking. They showed improved generalisation capabilities and robust performance of capsule neural networks using fewer training data than CNNs. Sequence models, a slightly different technology than CNNs, was also used for phase picking by Kirschner et al. (2019). They trained a Long-Short Term Memory (LSTM) network by formulating the problem as a sequence-sequence classification task. They demonstrated the efficacy of the method on local earthquakes.

Array-based phase picking methods have also been developed recently that harness the spatial coherence of seismic phases among different stations in a seismic array. In this regard, Chen and Li (2022) proposed CubeNet which considers the spatial correlation of individual picks at different stations to improve picking accuracy. They demonstrated that the method is robust against local spurious noises that are incorrectly picked by traced-based pickers. In a similar vein, Feng et al. (2022) developed EdgePhase, a multi-station phase picking model that is obtained by integrating Edge Convolutional module with EQTranformer. Compared to the standard EQTransformer, EdgePhase was shown to increase the F1 score by 5% on the Southern California data set. Moreover, performance tests in regions of different tectonic settings showed its strong generalisation ability in real-world applications.

The aforementioned works cover either event detection or arrival picking tasks separately. However, a number of deep learning approaches have also been developed that integrate the two tasks, such as Zhu and Beroza (2019) that combine phase identification and picking. Moreover, Zhou et al. (2019) developed a hybrid algorithm based on CNN and RNN to detect events and pick phases in two steps. Zhu et al. (2019a) presented a CNN-based Phase-Identification Classifier (CPIC) designed for phase detection and picking, and applied it to aftershock sequences of the 2008 Wenchuan earthquake. An eight-layer CNN model is trained to first detect events, which are then sent to a two-layer bi-directional RNN to pick P and S arrival times. Mousavi et al. (2020) trained a deep neural network with an attention mechanism (EQTransformer) for event detection and arrival picking. The trained model was shown to generalise well and demonstrated performance similar to a human analyst. Similarly, other deep learning approaches were developed for simultaneous event detection and arrival picking tasks (Zhang et al., 2020a; Saad and Chen, 2021). An interesting aspect of the work of Zhang et al. (2020a) is that the CNN does not input the raw data, but its time–frequency transform. In this way, each input seismic time signal becomes a 2D time–frequency panel, which allows better extraction of both temporal and frequency-related features. As a result, this gives an improved network performance on noisy data. Moreover, in a recent publication, Münchmeyer et al. (2022) conducted a large-scale benchmark study by comparing six of these mostly used deep learning models. Using a variety of datasets, they showed that EQTransformer, GPD, and PhaseNet yield similar performance.

While many of the algorithms highlighted above were originally developed in the global earthquake community, they are also applicable in microseismic monitoring. However, one needs to be cognizant of the particular challenges associated with microseismic data. This includes data with typically low SNRs and potentially having multiple events within a given time window, which is more common in microseismic than global earthquake datasets. As indicated by publications referenced above, the microseismic community has made numerous advances in the use of ML on detection and picking problems. Nevertheless, there is still plenty of room to learn from the advances in the global earthquake community for obtaining robust detection and picking performance in microseismic monitoring.

2.4. Localisation

Following the detection of (induced) events, it is important to find out where the detected seismic event is originating from (Warpinski, 2009).

Traditionally, localisation methods rely on retrieving the (relative) arrival times from the measured seismic records, and calculate from these arrival times the source locations, e.g. by forward calculating the ray travel times from a grid of potential source locations and matching these with the observed times (see, e.g., Lomax et al., 2009). For an overview of traditional travel time inversion methods for source localisation, see Thurber and Engdahl (2000).

As an alternative to these travel-time based methods, full waveform approaches have been utilised to locate earthquakes, where the waveforms are directly used in a stacking or mapping process in order to retrieve the source location. These can be divided into method using parts of the wavefield or the full wavefield. Part of the wavefield is usually used in diffraction stacking, often done in a grid-search method, where the measurements are focused for all possible source locations (e.g., Kao and Shan, 2004; Duncan and Eisner, 2010; Chambers et al., 2010; Anikiev et al., 2014). This process sums all recorded data along the traveltimes related to a diffraction response, where the diffractor is the seismic source. As an alternative source localisation process, Gharti et al. (2010) describe an efficient global optimisation method based on stacking the envelopes of the signals.

The full wavefield is used in methods based on reverse modelling, or back-propagation, within a gridded velocity model, provided that the measurements are acquired at a sufficiently dense spatial sampling (Gajewski and Tessmer, 2005). If the used velocity model is accurate enough, the back-propagated wavefield will focus at the hypocentre. As a hybrid solution, Willacy et al. (2019) propose a grid-search method for both location and moment-tensor type based on comparing elastic finite difference-forward modelled responses with the full-waveform microseismic events, assuming a known subsurface velocity model.

A common advantage of the time-based localisation methods lies in lower requirements on the seismic velocity model, as long as dealing with traveltimes leads to averaging over the velocities. Wavefield-based localisation algorithms do not require picking of wave arrivals but require more detailed knowledge of the velocities to model the waveforms, which may lead to larger errors where such an accurate velocity model is not available.

To avoid dependency on requiring a prior velocity model, Eisner et al. (2008) and Grigoli et al. (2016) use prior detected master events to relatively locate similar events in a data-driven focusing process. This approach benefits from naturally correcting local heterogeneities (like statics) as they are contained in the master event, as well as some source polarisation effects.

In recent years many of these deterministic steps have been replaced or augmented by machine learning. Such machine learning solutions are already widely applied in seismology for the earthquake localisation problem (e.g., Kong et al., 2019). The global seismology methods generally use the P- and S-phase of the recorded signal and sometimes are based on data from one seismic station. As an example, Mousavi and Beroza (2020) use a single channel measurement to directly find the earthquake's epicentre and depth. ML applications in the global seismology field are reviewed by Ren et al. (2020).

The field of microseismic monitoring differs from the global seismology in magnitudes, thereby, the signal-to-noise ratio of the events is less favourable. In addition, large earthquakes radiate strong signal in lower frequencies where the earth is more laterally homogeneous and velocity models are better constrained, while microseismic events are more sensitive to local heterogeneities in the overburden. This results in less information to be extracted from a microseismic event recorded by a single station measurement. Therefore, microseismic event localisation methods are mostly based on dense arrays of receivers, which provides more information about the location of the source and also its characteristics (see next subsection).

Secondly, microseismic arrays often use only single vertical component geophones (especially at the surface) and with such arrays – and the epicentres being mostly located within the array extent – it is harder to identify S-wave arrivals. Therefore, most of the microseismic localisation methods (using surface arrays) are based on P-wave arrivals. However, this is expected to change, as recent reports indicate the use of multi-component data for the detection and localisation of small earthquakes, for example in analysing weak aftershocks down to magnitude -2 (Li et al., 2022b).

The ML-based methods for this process, similarly to classical localisation techniques, can be divided into two categories: traveltime-based and waveform-based methods. The former method uses wave arrival picking, whereas the latter takes full waveforms and directly maps them to subsurface locations.

The arrival time based locations use several ML-based methods to extract the arrival times from the raw data. Huang et al. (2018) proposed a CNN to take the microseismic events during mining activities and

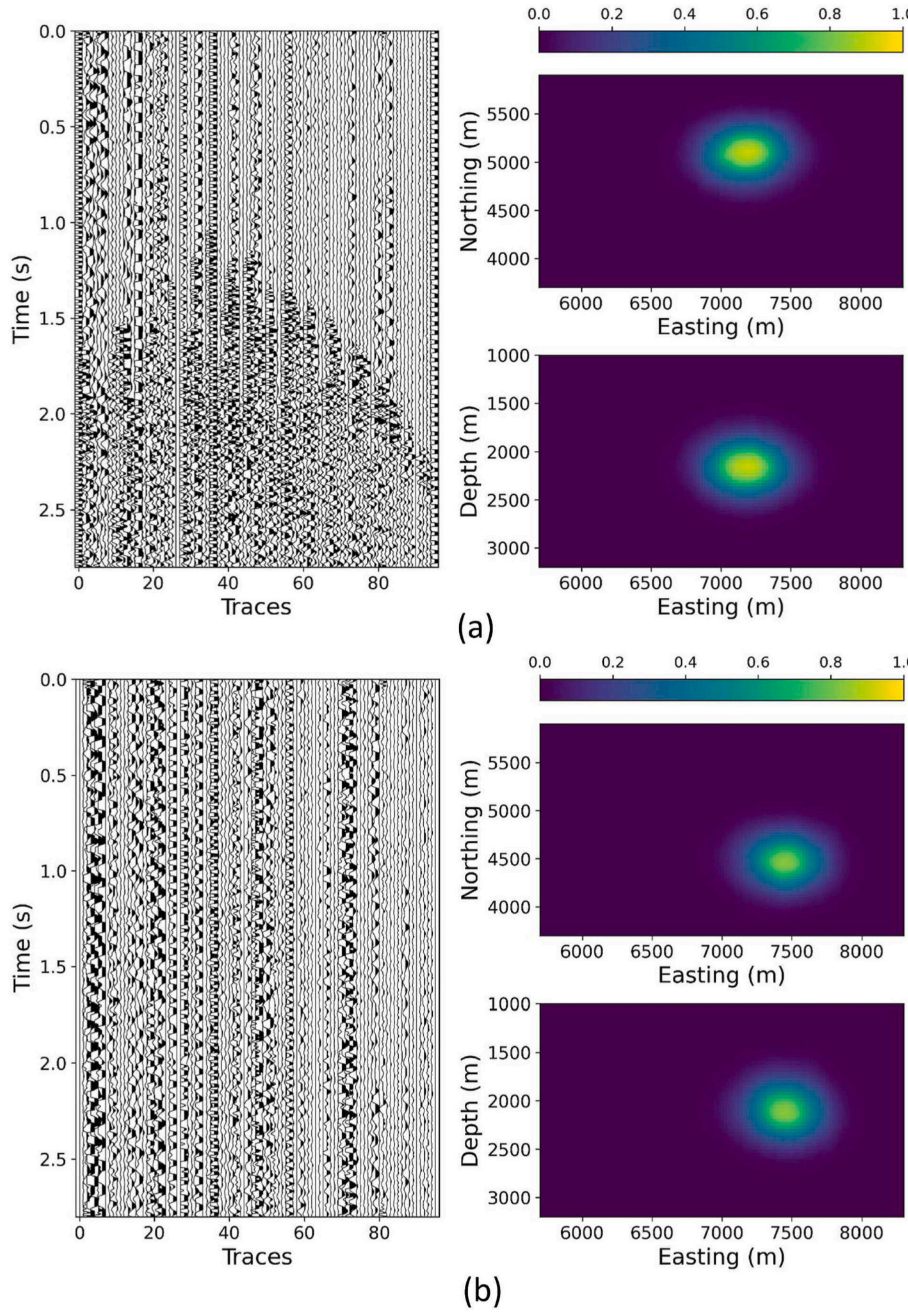

Fig. 4. An example of the probability density functions in 3D directly mapped via a NN from the microseismic field data events, for (a) a good quality event and (b) an event with poorer SNR. The 3D distribution is shown via two cross-sections, while the microseismic response is shown with amplitude normalisation on the left.

convert them in Time Delay of Arrival (TDOA), after which they can be mapped to source locations via a traditional calculation scheme, assuming a known speed of sound model of the subsurface. One interesting aspect is that they do not use the actual recorded data as input for their CNN, but the power spectrum and phase spectrum of the cross-wavelet transform calculated from the recordings.

Traveltime-based localisation methods use determined arrivals for microseismic location determination. Hao et al. (2020) proposed a simple feedforward artificial NN for hypocentre estimation using existing wave arrival picks, which was later extended (Anikiev et al., 2021) and benchmarked (Anikiev et al., 2022) against a classical traveltime-based method using a real microseismic monitoring dataset from a hydraulic fracturing site in Oklahoma, USA. The training is done on synthetics and does not require any historical seismicity. Along a similar strategy, Chen et al. (2022) pick the arrival times of the microseismic events in the traditional manner and then use these arrival times as the input to a ML algorithm – based on Random Forest (RF) – to find the source location. The training is done by ray-tracing arrival times of all possible source locations computed in a 3D velocity model.

Another application of machine learning is by using so-called physics-informed neural networks (PINNs), where the loss function in the NN is augmented by a physics law. This can help to make forward modelling of wavefields (Moseley et al., 2020) or travel time functions (Waheed et al., 2021b) more robust. Grubas et al. (2021) suggested utilising an eikonal equation based neural network to find the location that minimises the misfit between the picked traveltimes and the predicted ones. Yildirim et al. (2022) predict the location by finding the minimum of the NN trained traveltime function that fits the data and the eikonal equation where no optimisation is required. Izzatullah et al. (2022) supports this novel utilisation of PINNs with uncertainty quantification using a predictive Bayesian implementation using the Laplace approximation. The PINN utilisation in microseismic localisation is new, but shows considerable promise, especially in flexibility as these NN functions are continuous (i.e., no grid points are involved) and thus, any recordings along any type of surface, regular or irregular, can be utilised for microseismic event localisation.

ML approaches have also been proposed for full waveform methods using stacking. One such approach is to use ML for denoising the data, in order to get cleaner focusing results using traditional methods. Saad and Chen (2020) used an NN to select the desired waveforms from the raw microseismic data while masking all the noise, and use these pre-processed signals in a back-propagation and imaging algorithm, with a modified imaging condition that multiplies the contribution from various detectors, rather than summing them (Nakata and Beroza, 2016).

As the diffraction stacking – in combination with some grid search – is the most time-consuming task, this process is well-suited for a ML implementation. Thus, ML is used to directly map the raw – or pre-processed – seismic data into an estimate of the source location. Along this line of thought, Gu et al. (2019) construct a Bayesian deep convolutional neural network to output the source location and its corresponding uncertainty directly from the full waveforms recorded with a seismic array. Kriegerowski et al. (2018) demonstrate the possibility to find the source location from the raw data – without additional pre-processing – when using three-component seismic measurements. Their full waveform windows contain both P-wave and S-wave contributions. The method is trained on records from known locations in the same area. Although this work is in the context of global seismology, it may also have applicability to multi-component microseismic recordings.

Wang and Alkhalifah (2021) cross-correlated a reference trace with the rest of the data within its segment to reduce the size of the input data, which is input to a convolutional neural network to predict the location of the microseismic event and some of its characteristics. They later demonstrated that with training on synthetic data with field data noise added to them, they can locate events from field data in high accuracy (Wang et al., 2022). Finally, Alkhalifah et al. (2022) utilise a domain adaptation method to help improve the location accuracy on field data. Along these lines, Wamriew et al. (2022) propose a CNN-based method to jointly invert for source locations of microseismic events and the vertically varying velocity model using real-time downhole seismic waveform recordings. Wamriew et al. (2021) apply such direct waveform methods, using a deep CNN, on distributed acoustic sensing data and show the effectiveness in handling these large data volumes.

All these methods can provide near real-time locations when the network is properly trained using an accurate velocity model of the region, when velocity update is not part of the process (Wang et al., 2022). Another advantage of these ML approaches is that the user-dependent pre-processing is largely avoided, making the method more automatised.

It is natural to combine detection and localisation in one unique process. Zhang et al. (2018) propose two networks: one CNN for classifying if a certain time window contains an event and a second CNN that takes any detected event and locates it. They are able with – using small training data sets and networks with few layers – to get localisation results that match with results from traditional grid-search methods.

ML can also provide uncertainty of the located event. For this, Zhang et al. (2020) design a CNN that outputs a probability density function of the source location as a 3D volume. For the training data, suitable 3D Gaussian functions are used to represent the expected accuracy of the training data. Zhang et al. (2022b) modify their earlier work to avoid focusing problems related to polarity reversals caused by source mechanisms or other poor imaging conditions. The work of Zhang et al. (2020) is followed up by Vinard et al. (2020) and Vinard et al. (2022), who show that such a network can be applied to field data, by training it with synthetics contaminated with realistic noise. Fig. 4 shows examples of such a NN for field data from a hydraulic fracture monitoring project in Texas, where the SNR of the data is represented in the maximum amplitude of the obtained 3D function. Vinard et al. (2021) extend this work to localise lower magnitude events via transfer learning using detected weaker events on a network pre-trained with large-amplitude events. Fig. 4b illustrates an example of an event without high SNR arrivals on the individual receivers. Such waveforms illustrate the challenge of the microseismic monitoring where, unlike in earthquake seismology, we can't assume high SNR arrivals that can be easily detected. The ML algorithms should and need to overcome such challenge for all aspects of processing (detection, localisation, source mechanism determination, etc.).

Note that this transfer-learning approach shows similarities with the technique used by Münchmeyer et al. (2020) and Münchmeyer et al. (2021) for global seismology earthquake localisation. There, a Transformer Neural Network (TNN) is pre-trained with field data from a wide range of earthquake locations, after which transfer learning is applied to predict source locations from observed data in the target area.

An extensive comparative study on various neural networks for microseismic event localisation, including a posterior probability distribution, can be found in Mancini et al. (2021). This demonstrates that ML approaches can provide considerable speed in evaluating the uncertainty, which can also deviate from strict Gaussian processes.

In the same vein, Perol et al. (2018) use the three geophone components at a single measurement position to first detect an event and, next, estimate a location by assigning a label, indicting a limited number of geographic regions. Because the data from only one multi-component geophone is used, limited source location information can be retrieved and, therefore, this approach is better referred to as 'catalogisation'. However, the output provides a probabilistic distribution of these regions.

It is not clear what kind of ML technique is optimal for source localisation. Usually, three main methods are employed: traditional Random Forest (RF), Support Vector Machine (SVM), or a neural network which can handle more complex, nonlinear relationships between input and output. A comparison was carried out by Yang et al. (2021a), who describe a situation where microseismic events are expected either from a shallow or a deep region. They use detected signals from a single seismic sensor as input to a variety of classifiers to detect either shallow or deep source locations. Their conclusion is that the deep learning CNN outperforms the more traditional ML approaches like SVM or RF.

Finally, the question is whether ML methods outperform the traditional techniques. To investigate this, Zhang et al. (2022a) consider natural earthquakes at the San Andreas fault and develop a workflow, which they call LOC-FLOW, around a deep learning module to extract arrival times of detected earthquakes by an existing PhaseNet (Zhu and Beroza, 2019), and augment this by a set of steps to both detect and locate earthquakes. With LOC-FLOW they detect and locate almost four times more earthquakes than with traditional methods.

Although, there may still be reasons why a deterministic inversion would be preferred, ML could provide additional input that is not easily retrieved otherwise. As an example, Käufl et al. (2016) do not use a NN for the direct mapping of earthquake data into the required source parameters, but use the NN to build stochastic information on the desired parameters, which then can be used in a deterministic Bayesian inversion.

Besides only the source location, ML could give more information. As examples, Song and Alkhalifah (2020a) used a support vector machine (SVM) to develop a classification algorithm of microseismic events in time-reversed source images and van den Ende and Ampuero (2020) proposed a graph neural network (GNN) approach to seismic source characterisation (location and magnitude estimation), based on multi-station waveform recordings. In the next subsection, focus will be put on the extraction of source mechanism information.

2.5. Source mechanism determination

Induced microseismicity can have both pure shear and non shear mechanisms as the human activity result in volumetric changes. Initial observation of induced seismicity in unconventional and conventional oil/gas fields was explained by double-couple (DC) source mechanisms (Rutledge et al., 2004; Li et al., 2011a; Li et al., 2011b; Li, 2013). However, other studies found evidence of both pure shear as well as non-shear source mechanisms in induced seismicity. Therefore, inversion of the full moment tensor for induced microseismic events is preferable.

The two most common methods of inversion, of either pure shear or full moment tensor source mechanisms, can be carried out either by inversion of amplitudes of seismic waves (e.g., Šílený et al., 1992) or full waveforms (Sipkin, 1982).

Both methods have their advantages and drawbacks. For example, amplitude inversion requires accurate picking of the arriving waves but allows simpler velocity models while full waveform modelling provides robust inversion of recorded waveforms but is limited in fitting low amplitude arrivals (typically P-waves, S-wave arrivals are used if 3-C data are available) and requires a very accurate velocity model suitable for modelling of multiply reflected arrivals.

In the following paragraphs we will review ML approaches to source mechanism inversion using mainly equivalents of the full waveform methods as this is the current state of art in the ML applications to source mechanism inversion, but we wanted to provide a broader perspective to the reader.

Käufl et al. (2013) implemented a Bayesian approach using Mixture Density Networks (MDNs, introduced by Bishop (1994), which is a class of neural networks outputting parameters of Gaussian mixture models) for early warning recognition of source mechanisms of large earthquake centroids. The methodology uses displacement data from GPS local stations to constrain centroid location and event magnitude assuming deviatoric moment tensor (DMT) for source mechanisms with posterior probability densities derived using the MDN trained on synthetics in a form of parameters of a Gaussian mixture model (GMM). The results in this study revealed robust solutions that are relatively insensitive to variations in the 1-D crustal earth model. A promising application from earthquake seismology which uses a CNN-based approach for sparse surface network of receivers was proposed by Kuang et al. (2021), who constrained the source mechanisms to pure shear and used full waveforms recorded on 16 three-component stations. To lower the sensitivity to attenuation and possible cycle skipping, the seismograms were filtered between 0.05 and 0.1 Hz and the application is aiming at an automated determination of large earthquakes with Mw>5; the velocity model is 1D calibrated with previous studies without attenuation (not mentioned). This implementation illustrates well the need for low frequency response to simplify wave propagation. However, such an approach is difficult to adopt for microseismicity as the low frequency signal is hidden below the noise for most induced seismic events.

For microseismicity, Ovcharenko et al. (2018) implement single well inversion in homogeneous isotropic medium using a trained ANN (multilayer perceptron). The inversion is based on least-squares method (LSM) for the P and S peak amplitudes on all components, modelling does not include attenuation. The authors test sensitivity to the number of hidden layers in the ANN concluding the three layers are needed to achieve satisfactory inversion measured by misfit of the moment tensor components. The conclusions have been somewhat surprising, given the fact that the full moment tensor is very poorly constrained (essentially only thanks to near field zone) from a single vertical monitoring borehole in homogeneous medium. Alternatively, Carrizo Mascarell (2020) changed the methodology to use full waveforms simulated in a 1D layered structure (using discrete wave number (Bouchon and Aki, 1977)) and simulated response on a horizontal and vertical array of receivers to better constrain the inverted mechanisms. The misfit function was the least squares difference between the data and synthetics. They obtained very good results on a synthetic test for low noise environment, reliably inverting full moment tensors. Revelo Obando (2021) used the same methodology and applied it to full moment tensor inversion from a horizontal array of receivers with very encouraging results for synthetic datasets. It is theoretically possible to obtain full moment tensor from horizontal array of receivers in 1D medium with full waveform modelling, however such an inversion requires very accurate and precise velocity and attenuation models. It remains to be seen if such results can be also obtained for real data where models are not known, but the ANN algorithm works well.

Wamriew et al. (2020) designed a joint localisation and moment tensor inversion scheme, where the locations of microseismic events are estimated with a multi-layer 2D CNN, whereas their source mechanisms are inverted using a multi-head 1D CNN. Both networks are trained on synthetic data with low SNRs, mimicking the real data. The results show that the proposed joint approach is capable of finding meaningful relations between the input data and source characteristics – locations and source mechanisms – and gives an accuracy comparable with classical methods.

Choi et al. (2022) proposed a CNN-based source mechanism inversion method with domain adaption. The synthetic peak amplitudes and first arrivals of P- and S-waves are used as training data sets. The CNN model is pre-trained with a homogeneous velocity model. The pre-trained CNN model is then fine-tuned with the domain adaptation technique.

The currently prevailing methodology for using neural networks for source mechanism inversion is using synthetic full waveform seismograms with modelled source mechanisms as labels for training and validation and then applying such trained neural network to invert real observed detected events. This is reliable and successful if the modelled training dataset uses realistic velocity models. Such models are generally available for large earthquakes in low frequencies because low frequency signals are less sensitive to local heterogeneity. But this is not generally available for microseismic data as the low frequency part of their signals have generally very low SNR and the velocity models allowing to model high frequency signal are generally not good enough and furthermore must be calibrated locally. Earthquake seismologists generally overcome velocity model limitations by using attributes of the waveforms such as P- or S-wave phase polarity and amplitudes or their ratio. We believe neural networks for source mechanism inversion of microseismic events have potential to identify such features in waveforms, but algorithms may have to include uncertainty of the velocity model or some kind of specific misfit functions dependent on the attributes.

2.6. Cluster analysis

Technical and methodological advances in microseismic monitoring result in larger amounts of detected and located microseismic events and, therefore, call upon using statistical analysis of the distribution of seismic events (clusters or clouds) in space and time. Spatiotemporal cluster analysis of microseismic event clouds is a powerful tool that helps to get information on the dimensions, orientation, and complexity of the fracture system, as well as to predict the fracture growth (e.g., Andrade and van der Baan, 2021b). On the other hand, accumulated seismicity datasets can be used to forecast seismic events and predict earthquake trends (e.g. Geller, 1997), which is particularly important in seismic hazard assessment, early-warning, and seismic traffic-light systems. Such sophisticated and challenging tasks can benefit a lot from ML and data science techniques.

As far as seismic sensors can record all types of signals propagated, classification of this signal is an important first step to focus the analysis only on a certain type of seismic data, according to the application. There are several ML-based approaches to discriminate between earthquakes and microseismic activities. For instance, Duan et al. (2021) use four different techniques – random forest (RF), support vector machine (SVM), deep convolutional neural network (DCNN), and residual neural

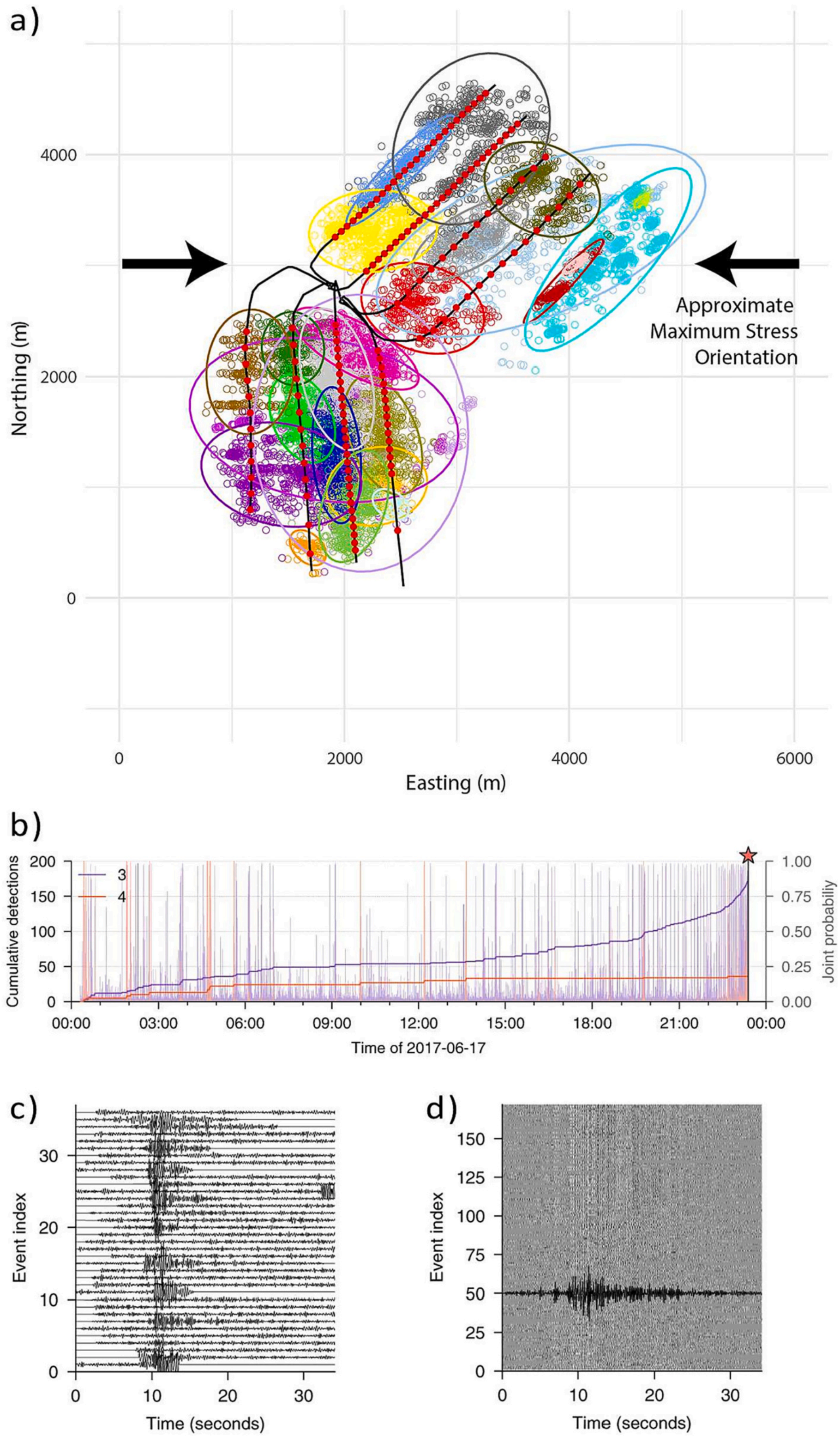

Fig. 5. Seismic cluster analysis in space (a) and time (b,c,d): a) Spatial cluster analysis results from the application of the GMM algorithm to the microseismic dataset recorded during hydraulic fracturing in Western Canada (plan view). The clusters are differentiated by colour. The black arrows show the orientation of the maximum principal stress (from McKean et al., 2019); b-d) Temporal cluster analysis of a landslide in Greenland: within-cluster cumulative number of event detections (b) in two clusters (purple and red). The relative probability for each time window to belong to each cluster is represented with lighter bars. The waveforms extracted within the clusters are extracted and aligned with respect to a reference waveform within the cluster, for cluster 4 (purple, c) and cluster 3 (red, d). The stack of the waveforms (panel d) is shown in black solid line (from Seydoux et al., 2020).

network (ResNN) – to classify the signals in a longwall coal mine in Australia. The results showed that microseismic events could be classified with an accuracy of over 90%. For training, a primary database was established with eight thousand signals in total, with two thousand for each of the four event classes manually labelled as a ground truth. An overview of the performances of ten frequently-used ML models for microseismic/blasting events recognition in underground excavations was prepared by Pu et al. (2020).

Seismic cluster analysis is used to determine groups of seismic events (clusters) in such a way that certain event properties assigned to the same cluster are similar to each other. Clusters are, therefore, identified either as groups of events that minimise internal and maximise external distances, or as dense event regions separated by sparse regions (Piegari et al., 2022). Distance-based algorithms minimise the distances between hypocentres of located earthquakes, using, for instance, k-means (e.g., Ouillon et al., 2008), fuzzy clustering (e.g., Ansari et al., 2009), Gaussian Mixture models (GMMs) (e.g., Ouillon and Sornette, 2011) as well as hierarchical clustering (e.g., Trugman and Shearer, 2017; Kamer et al., 2020). However, the analysis of spatial features of seismicity by density rather than distance seems to be more advantageous (e.g., Piegari et al., 2022; Fan and Xu, 2019). Density-based algorithms can identify clusters of an arbitrary shape and are more efficient on large datasets. Moreover, they allow to take into account multiple event characteristics, such as moment tensors or even waveforms and their spectra.

The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, introduced by Ester et al. (1996), is one of the most popular density-based clustering approaches used in seismology. Konstantaras et al. (2012) presented a graphical tool based on various clustering algorithms, including the DBSCAN algorithm, Fuzzy C-Means, Quantum Clustering, and a self-developed dynamic spatial clustering algorithm, and applied them to analyse seismicity in the region of Aegean Sea. Cesca et al. (2013) adopted the DBSCAN algorithm to classify focal mechanisms from large moment tensor catalogues taking into account variable uncertainties of different moment tensor components. They developed a set of software tools (Cesca, 2020; Petersen et al., 2020) based on DBSCAN implementing multi-dimensional clustering that accounts for various properties such as origin time, focal mechanism (moment tensor), and waveform similarity. Schoenball and Ellsworth (2017) also used the DBSCAN to identify clusters in the catalogue based on the spatial proximity of earthquakes in Oklahoma and Southern Kansas and performed a detailed analysis of their spatiotemporal evolution. More recently, Piegari et al. (2022) investigated the DBSCAN algorithm and its extension OPTICS (Ordering Points To Identify Cluster Structure) (Ankerst et al., 1999) for analysis of the spatial distribution of seismicity and applied them to seismic catalogues of earthquake sequences in Italy and Japan. They showed how cluster solutions help to identify 3D features of tectonic structures that were activated in a seismic sequence and generalise the analyses for arbitrary seismic sequences.

While supervised ML algorithms rely on the quality of the predefined labels to determine specific data classes known a priori, unsupervised machine learning strategies explore seismic data without using any explicit assumptions and, therefore, are more suitable for seismic cluster analysis. However, in both cases, the keystone to success lies in the data representation, namely one needs to define an appropriate set of features for solving the task of interest (e.g., Seydoux et al., 2020).

Unsupervised learning has been widely applied in volcanic monitoring systems (e.g., Esposito et al., 2008; Hammer et al., 2012; Soubestre et al., 2018; Giudicepietro et al., 2021), as well as in geothermal induced seismicity analysis (e.g., Holtzman et al., 2018). Among unsupervised neural networks, the Self-Organising Map (SOM) is suitable for the discrimination of seismic signals generated by different sources in a composite seismic wavefield (Giudicepietro et al., 2021). Sick et al. (2015) used SOM combined with Principal Component Analysis (PCA) to estimate the applicability of a single-station seismic clustering.

Beyreuther et al. (2012) applied a Hidden Markov Model (HMM) adopted from speech recognition to detection (triggering) of seismic events, and showed general applicability of so-called state clustering for earthquake classification of volcano-induced seismicity.

McKean et al. (2019) introduced a probabilistic clustering method based on a GMM algorithm with physical constraints. They applied the method to the microseismic dataset recorded during the hydraulic fracturing in western Canada in order distinguish natural and anthropogenic processes and identify fracture networks in the subsurface. An example of clustering analysis resulting in the identification of 25 clusters is shown in Fig. 5a.

Seydoux et al. (2020) proposed a strategy for clustering and detection of seismic events utilising unsupervised machine learning that combines a deep scattering network (based on a deep CNN) and the GMM. They applied the method to the continuous seismograms collected during the massive landslide in Greenland and showed that the method is capable of unsupervised detection and recovery of the repeating precursory seismicity recorded before the main landslide rupture. Examples of waveforms extracted during the clustering analysis are shown in Fig. 5c and d, whereas the temporal evolution of the two of the clusters (clusters 3 and 4 from Seydoux et al. (2020)) together with the within-cluster cumulative detections are presented in Fig. 5b. The study of Seydoux et al. (2020) nicely shows that the unsupervised learning can also be used for informative forecasting of seismic activity, which is the topic of the following subsection.

2.7. *Earthquake prediction*

Earthquake prediction, or, more generally, probabilistic forecasting of the seismicity trends, is another big research field. This field has had a long and controversial history, sometimes sordid, but remains a very important and hot topic with ML giving geoscientists a fresh hope to find approximate solutions to problems that were considered unsolvable (Geller, 1997). We believe that the vast amount of recent papers on ML applications relevant to that topic deserve a separate review study. Here we provide just a brief overview of the current state-of-the-art of earthquake prediction in the context of applicability to microseismic monitoring.

In global seismology, there are two general approaches to predicting earthquakes, one is based on precursors, and the other takes into account seismicity trends (Bhandarkar et al., 2019). While earthquake precursor studies focus on various phenomena, e.g., radon gas emissions, unusual animal behaviour or electromagnetic anomalies (Geller, 1997), which might indicate an impending earthquake, trend-based methods tend to identify specific patterns of seismicity that precede an earthquake and thus are more suitable for ML applications and microseismic monitoring. For instance, Bhandarkar et al. (2019) proposed a Long Short-Term Memory (LSTM) recurrent neural network (RNN) to predict an attempt to forecast earthquakes, and trends using a dataset containing a series of past earthquakes, and benchmarked it against a feedforward neural network (FNN) approach. The LSTM model predicts earthquake time, location and magnitude; it is also able to capture certain data trends and clearly outperforms the FNN method (Bhandarkar et al., 2019).

Fluid injections into underground formations and associated induced seismicity jeopardize the sustainable use of the subsurface. Forecasting of potential induced seismicity, as well as understanding the fault behaviour is the key to successful management and mitigation of injection-induced seismic risks (Ji et al., 2022). In this context, one of important research directions is the analysis of acoustic emission data based on laboratory experiments on fractured rocks. Given the fact the loading condition is controllable and the recorded seismicity sequence is long and complete, the 'labquake' data is a perfect target for ML-based analysis (e.g., Laurenti et al., 2022). Recent works demonstrate that ML can predict the timing and magnitude of laboratory earthquakes using statistics of acoustic emissions (Rouet-Leduc et al., 2017; Hulbert et al., 2018; Bolton et al., 2020, among others). At the same time applying ML to the prediction problem raises several challenges related to proper validation performance on rare events, generalisation potential of the studied ML models, as well as handling the outputs of black-box ML methods (Johnson et al., 2021). Recent advances in laboratory earthquake studies have important implications for ML-based prediction of microseismic activity and precursors to failure, which have a fundamental impact on the ability to improve earthquake early warning systems and possibly earthquake forecasting (Bolton et al., 2020).

As for microseismic studies outside of the lab, Andrade and van der Baan (2021a) analysed the spatiotemporal distribution of microseismicity induced by hydraulic fracturing and studied methods to forecast the microseismic cloud size in real time. They benchmarked the proposed CNN-based approach with the analytical diffusion model method of Shapiro et al. (1997). Results show that the CNN-based approach

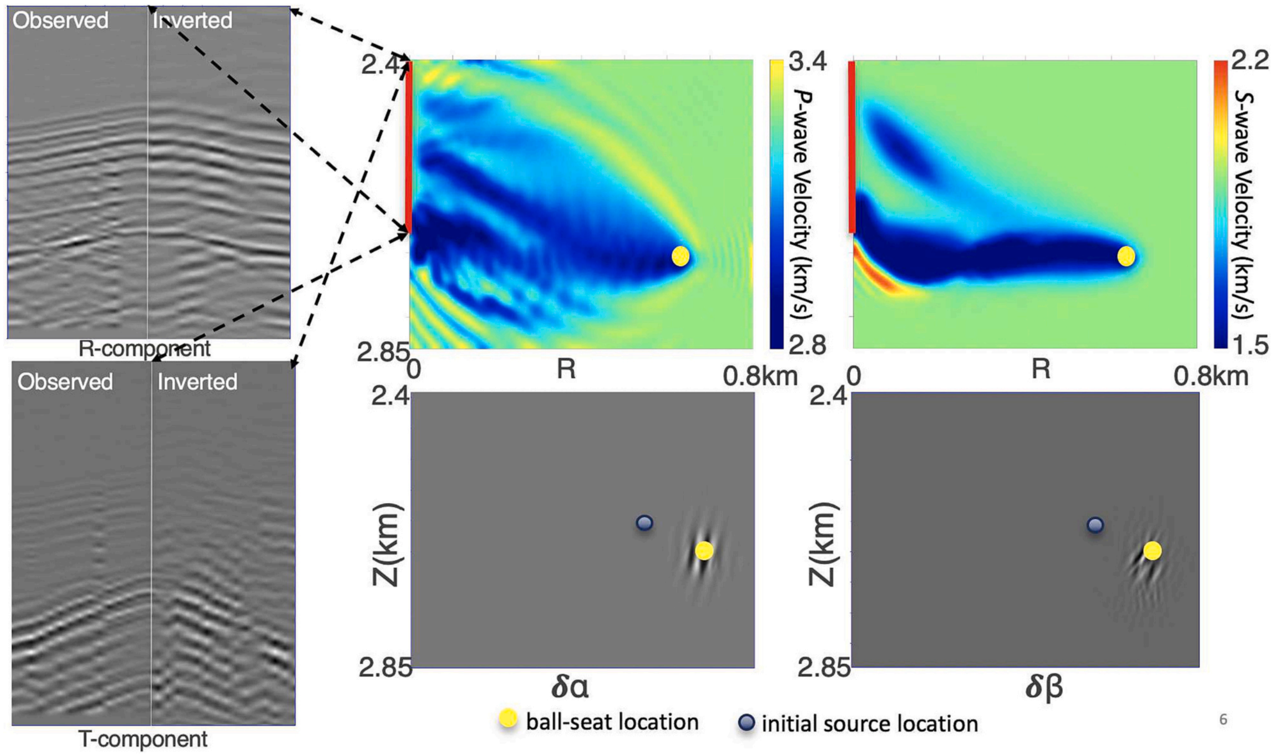

Fig. 6. An example of velocity and source image locations extracted from multi-component seismic data recorded in a well at the locations given by red on the upper left side of the velocity inversion plots. The initial estimated location of the microseismic event based on the constant background velocity model is shown as a grey dot, while the actual location given by the ball-seat event is shown as a yellow dot. The inverted model with updates in the illuminated region by this source provide a source image for *P*- and *S*-waves at the ball seat location. On the left, the measured data and the simulated ones from the inverted source and velocity model for the radial (top) and transverse (bottom) components.

outperforms the quality of predictions of the physics-based models but also reveal a reduced prediction capability. More recently, Andrade and van der Baan (2021b) extended the study and applied a physics-based approach relying on diffusivity estimates derived from the microseismic observations and a convolutional neural network (CNN) trained with the engineering curves to forecast the microseismic cloud size in real-time.

Earthquake prediction is one of the hottest topic in global seismology which explains the rich variety of different ML applications (e.g., Asim et al., 2018; Bhatia et al., 2018; Fabregas et al., 2020; Berhich et al., 2021; Doğan and Demir, 2022), even such an exotic studies as adaptation of biological immune models (Zhou et al., 2022). Mignan and Broccardo (2020) found two emerging trends while analysing ML applications to earthquake prediction in the period from 1994 to 2019: an increasing interest in this domain over time and a complexification of the NN models toward deep learning (DL). However, they have also concluded that potential of DL in significant improvement of earthquake forecasting remains unproven. As it was nicely summarised by Kong et al. (2019), although ML methods provide seismology with new tools, combining them with classical techniques might lead to radically new discoveries. The same is true for microseismic monitoring, where one of the important aspects is whether the underlying velocity model is optimal. This is a fundamental question that influences the estimates of seismic event characteristics as well as subsequent seismicity analysis. These questions are discussed in the following subsection.

2.8. The role of velocity

An essential ingredient to properly locate and characterise microseismic events, whether we use conventional or ML methods, is knowledge of the elastic or even simply the acoustic properties of the subsurface in the region of interest (Duncan and Eisner, 2010). Without an accurate wave propagation description of the subsurface, the localisation and characterisation of microseismic events using passive seismic data are prone to errors (e.g., Gajek and Malinowski, 2021). Many studies have been devoted to quantifying the quality of the microseismic event locations and the uncertainty involved due to potential errors in the velocity model (e.g., Kocon and van der Baan, 2012). For example, Poliannikov et al. (2014) and Gesret et al. (2014), among others, utilised a Bayesian formulation to quantify such uncertainty when the sources of information are the traveltime picks. With traveltime picks, Zhang et al. (2017) also used Bayesian inference to estimate both the velocity model and the microseismic location. The common theme behind all the above studies is that without an accurate description of the subsurface, locating microseismic events is potentially biased.

The sources for the Earth property information (a velocity, including an anisotropic and an attenuation, model of the subsurface) necessary for the microseismic localisation and characterisation objectives have varied over the years. The common theme is to extract such information from any source possible (Collins et al., 2014). The sources of information include prior active seismic experiments supported, possibly, by measurements at wells, as well as, a priori information from our geological knowledge and from other geophysical methods. Depending on the available sources of velocity information, we can build detailed and complex velocity models for microseismic localisation and characterisation tasks (Grechka and Duchkov, 2011; Das et al., 2021). Recently, we have even utilised the microseismic events themselves to estimate the velocity in a multi-parameter (source locations and velocity) inversion, handled either sequentially (e.g., Li et al., 2013) or simultaneously (Grechka and Heigl, 2017; Wang and Alkhalifah, 2018). In fact, for microseismic localisation objectives using P-waves or P-wave traveltimes, the acoustic assumption for describing wave propagation often suffices. Along these lines, Choi et al. (2018) and Barthwal and van

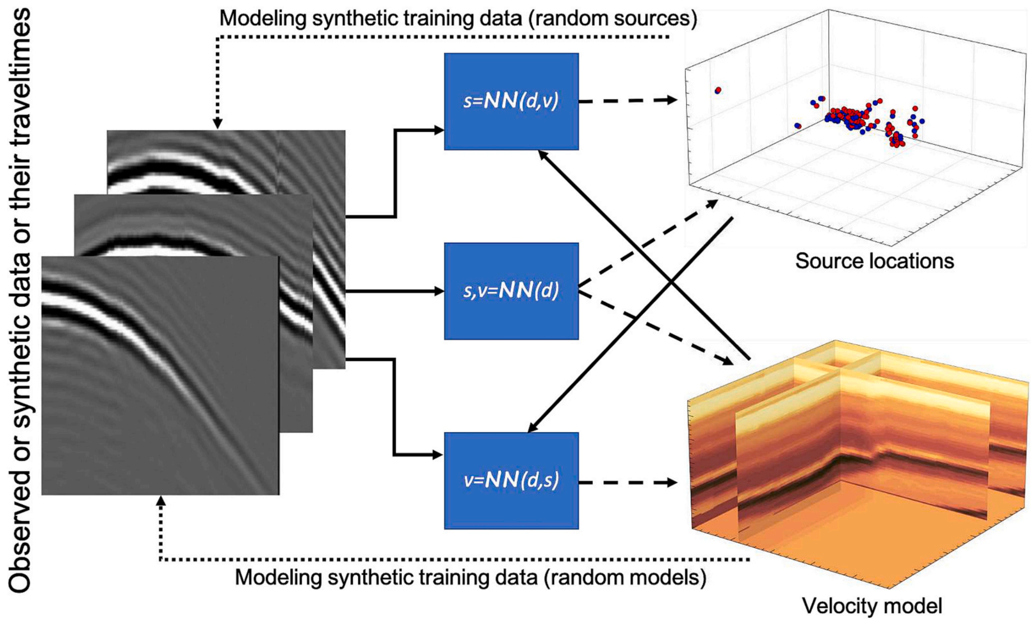

Fig. 7. A diagram outlining the general options we have in utilising ML in velocity inversion from microseismic events. Here, *NN* depicts a neural network model with inputs: either data, *d*, source locations, *s*, or velocity, *v*, or a mix of them. The dashed arrows point to the outputs of the trained model.

der Baan (2019) inverted for the sources and velocity using picked traveltimes in a tomographic fashion. Traditionally, calibration shots are used to calibrate an a priori built velocity model. For example, Bardainne and Gaucher (2010) utilised calibration shots to constrain the tomographic inversion and tested several methods of calibration. Tan et al. (2018) used a neighbourhood algorithm (Sambridge, 1999) along with setting a station as a reference, and they quantified the accuracy of their approach on real data using ball-seat events where the source is known. Such ball-seat events were used to test a waveform based inversion that utilises a source function independent objective function to avoid the nonlinearity induced by the unknown source (Wang et al., 2020). Along the line of waveforms, Song and Alkhalifah (2019) used a wavefield inversion approach to simultaneously determine the source location and invert for the velocity model. Such inversions, time- or waveform-based, are common in the global scale as other sources of Earth elastic property information are usually not available, and most of the information comes from the earthquakes themselves (Montelli et al., 2004).

Illustration of the importance of the velocity model constrain is shown in Fig. 6, extracted from the work of Wang et al. (2020), which shows multi-parameter inversion of data acquired along a well. This work demonstrated that important velocity information is embedded in the passive seismic recording. They utilised a source independent objective function to reduce the nonlinearity, and thus, were able to invert for the *P*- and *S*-wave velocities, as well as the *P*- and *S*-wave source images for a seat-ball event along a horizontal well, where the source location is known and used here to evaluate the accuracy of the inversion. Compared to the background initial constant velocity location, predicting the event much closer to the well, the inverted low velocity information in the illuminated areas managed to place the event in its true location and that produced multi-component data that matched the observed ones. This example demonstrates the importance of using microseismic data for velocity estimation, and with machine learning we will have an opportunity to utilise a swarm of events efficiently for such a task.

Machine learning methods for determining the location and attributes of the microseismic events are not immune from a requirement of an accurate velocity model in the area of investigation. Whether the velocity models are needed to generate the synthetic examples for training the ML model, or are embedded in the ML implementation, the subsurface information accuracy is crucial to the performance of the ML. One may argue that we can train a neural network to locate and characterise microseismic events using previously determined locations, like in mature fields (i.e. Groningen Gas field), with a history of passive seismic studies (Daniel et al., 2016; Shaheen et al., 2021). However, we have to keep in mind that these training samples at some point in the past required a velocity model to determine the microseismic event locations.

Using machine learning to determine the velocity model goes back to the early work of Röth and Tarantola (1994), who have achieved compelling results at the time. Moya and Irikura (2010) applied velocity inversion using neural networks granted the earthquake locations were already determined. This has been a general theme in building the velocity model in the global scale, as we rely on seismic, often traveltime, tomography (Montelli et al., 2004). In spite of the aforementioned work, ML methods for velocity inversion based on microseismic events have been only mildly explored. Part of the reason is the challenge of the inherently large model space for such a problem, which includes the source location, the velocity model, and in the ML case, the neural network parameters. Among the recent attempts to utilise ML for passive seismic data model building, there is one performed by Wamriew et al. (2021). They applied their approach on distributed acoustic sensing (DAS) data, and considered a simple layered model as they predict the location of the event and the velocity values at that location. They trained their model on synthetic data mixed with real seismic ambient noise collected from field data, and applied the approach on real DAS data collected in the same area. The approach showed relatively low errors. Delplancke et al. (2020) utilised mining events in a passive seismic tomography approach. They specifically use stochastic gradient decent methods promoted in machine learning to implement a Bayesian inference of the velocity model.

Utilising physics-informed neural networks (PINNs) for this objective will be the next frontier in machine learning. PINNs are neural network functions of space, and potentially time, representing the solution of a partial differential equation, like the eikonal or wave equations, by using these equations as loss functions to train the neural network. PINNs offer high degrees of flexibility and have already made their imprint in locating microseismic events (Grubas et al., 2021; Yildirim et al., 2022). Their utilisation in inverting for velocity models have been demonstrated for active seismic events (Waheed et al., 2021a; Song and

Alkhalifah, 2022). It is a matter of time that the two tasks converged to provide a mechanism to invert for the velocity model using passive seismic data of microseismic events. In addition to PINNs, some of its variants, like DeepOnet and neural operators (Rosofsky and Huerta, 2022; Liu and Cai, 2021; Yang et al., 2021b) are potential tools in the machine learning sphere that could help us invert for the velocity model using neural networks. Unlike, PINNs, these methods train neural network operators rather than functions, and thus, are potentially applicable to wider problems.

Despite the limited work on utilising machine learning to estimate velocity models from microseismic events, we can not dispute the fact that the microseismic data have plenty of velocity information in them. Since the wavepath between these almost point-source microseismic events and the recording surface is generally direct, such velocity information may include high resolution components similar to those advertised for diffractions (Bauer et al., 2017). The utilisation of ML in velocity inversion from microseismic events will generally take one of the paths outlined in Fig. 7. We can either invert for the velocity considering the source locations are known, which is often, but not always practised in Earthquake seismology, or invert for the sources considering the velocity is known, which is common in micro-seismic monitoring as we saw in the section dedicated to detection, or invert for them simultaneously or in sequence as two sub-problems. In using physics (simulation) to train the NN model for velocity inversion, we have the complexity of generating synthetic data for many random models (we have done that with active seismic data). As opposed to random source locations used to train the network to locate microseismic sources, velocity models are often high dimensional functions, which will pose a challenge to this problem. However, we have learned fast that complexity can be addressed by the constant increase in available computational resources and by our ability to develop smarter algorithms, like those that may utilise a reduction in the dimensional representation of the velocity model (the latent space) (Razak et al., 2022).

3. Discussion and concluding remarks

The reviewed publications show certain trends in the way they solve the challenges of microseismic monitoring using machine learning. For example, often the lack of pre-existing seismicity data recorded on the specifically designed network is overcome by training on synthetic seismograms that replicate as much as possible the real data, including the addition of realistic noise. This approach provides us with ample labels for training, and these labels can be considered as accurate (within the accuracy of the physics used to generate the synthetic data). Such labels (e.g., event locations) can be confined to an expected area of human activity, e.g., expected locations at the reservoir or monitored overburden. This provides a significant advantage over training on real data, where training events can be limited to only some part of the model space. However, a great drawback of using synthetic data for training is the need for modelling of realistic waveforms to mimic the real data. There are reasons why the synthetic waveforms hardly mimic real waveforms ranging from the source representation to media assumptions. The generalisation of synthetic data training to field data falls under a category of machine learning referred to as domain adaptation. Examples of that in our field, which includes discussions on such synthetic data limitations, are provided in Alam et al. (2018), Alkhalifah et al. (2022) and Birnie and Alkhalifah (2022). However, for the majority of the current applications in the microseismic monitoring this limitation is neglected and we hope future work will address this issue. Another issue related to this is overfitting on limited data if it is available. Whilst still an area of active research, various methodologies have already been proposed to tackle overfitting, for instance, early stopping (e.g., Chollet et al., 2015) tracks the validation loss and aims to stop the training before overfitting occurs. Future methods will probably allow for combinations of field and synthetic data while avoiding overfitting on limited pre-exiting data. Another interesting direction is to augment the limited seismic data by learning its distribution through generative adversarial networks (GANs). The trained GAN model can then be used to generate realistic seismic data for the augmentation task (Wang et al., 2021; Novoselov et al., 2021). This approach can be particularly useful for overcoming the limitation in the availability of microseismic data for supervised learning tasks.

Another common feature in the currently applied methodologies is training and processing full waveforms of the detected events (Münchmeyer et al., 2021; Wang and Alkhalifah, 2021) rather than using only some human-defined features of the seismograms, like arrival times, amplitudes or phases of the microseismic events. ML algorithms can be 'smart' enough to define their own features that are informative for the desired inversion result, where these features are hidden in the layers of the NN. Our current understanding is that the features of seismograms not only allow humans to simplify the inversion, but also simplify requirements on accuracy and precision of the input parameters such as velocity model. Maybe NN will enable us to discover new features. However, this brings a concern we face in all ML based applications, including microseismic monitoring and characterisation tasks, which is our limited understanding on what our networks are actually learning or doing, i.e., the "black box" phenomenon (Guidotti et al., 2019). We see a trend in general in ML towards understanding more about our networks inner workings, or, in other words, towards explainable AI (Linardatos et al., 2020). This includes understanding the lower dimensional representations and features responsible for the predictions. These features in microseismic data might be alternative representations of wavefields, like the geometrical shapes of arrivals or their phases, or their arrival times, and so on. Such knowledge of how neural networks operate can increase our confidence in them and help us better understand the uncertainty involved – an issue that we found lacking in the papers we have reviewed. We anticipate that more attention will be paid to understanding our networks as this trend is gaining steam in the machine learning world, in general.

In addition to the above, we foresee that future research in ML applications will result in better estimations of the uncertainties of the inverted results. This will hopefully include also uncertainties resulting from the neural network model and/or the Earth model, as well as, microseismic event parameters. Considering that the use of full waveforms is now the preferred choice, we envision at least three ways where neural networks may contribute to this theme: 1) using neural networks to evaluate uncertainty resulting from the model; 2) using seismicity to invert more complex models that would allow fitting of the synthetics to the measured waveforms; and 3) using neural networks to identify the parts of seismograms that are most relevant to the model space we seek. It is even possible that future neural networks may identify new characteristic features of observed seismograms that could provide better inversion of source parameters than those used today.

Another big area in which future neural network applications will evolve is in the use of new instrumentation. For example, the DAS acquisition provides not only information similar to geophones, but also measurements of the strain rate. Furthermore, the measurement with DAS is not a point measurement but nearly continuous in space and time. This promises large amounts of data in which ML algorithms would be able to handle and adapt to better. As illustrated by Birnie et al. (2021) for microseismic event detection, such large datasets can be easily utilised for training (without the need for data cropping) through the implementation of distributed deep learning - which is readily available in both the PyTorch and TensorFlow deep learning frameworks. From denoising, localisation and characterisation of events to velocity model building and even mapping CO_2 plumes and fracture evolution, ML has the capability of learning these tasks regardless of the size of the inference data.

As discussed in Section 2.8, the estimation or updating of propagation velocity models with ML in the microseismic context is a largely unexplored territory, where much development can be expected,

Table A.1
Acronyms and abbreviations used in text with brief explanation of their meaning.

Acronym	Description	Meaning
AE	Autoencoder	A class of ANN used in unsupervised ML with unlabeled input data; AEs are mainly used for data compression
AI	Artificial intelligence	Ability of a computer program or a machine to think and learn
ANN	Artificial neural network	A computing system inspired by biological NNs that constitute animal brains
BNN	Bayesian neural network	A special class of ANN where the weights are considered to have a probability distribution under a Bayesian inference concept, which helps to avoid overfitting and allows to account for uncertainty of the model parameters
CAE	Convolutional autoencoder	An autoencoder with shared weights
CCS	Carbon capture and storage	Capturing carbon dioxide before it enters the atmosphere, transporting it, and storing it
CLVD	Compensated linear vector dipole	A component of the seismic moment tensor (SMT) that describes a source mechanism representing a shear slip along two or more fault planes
CNN	Convolutional neural network	A deep NN with shared weights most commonly applied to analyze visual imagery
CRED	CNN-RNN earthquake detector	An earthquake detecting method based on deep neural networks that uses a combination of convolutional layers and bi-directional LSTM units in a residual structure.
CUDA	Compute unified device architecture	A parallel computing platform and application programming interface that allows software to use certain types of GPUs.
CWT	Continuous wavelet transform	A mathematical function used to represent a waveform using a set of analytical functions – wavelets, by comparing the waveform with their shifted and scaled versions
DAS	Distributed acoustic sensing	A technology that enables continuous, real-time acoustic measurements along the entire length of a fiber optic cable
DBSCAN	Density-based spatial clustering of applications with noise	A popular density-based clustering method
DC	Double-couple	A component of the seismic moment tensor that corresponds to a pure shear source mechanism
DMT	Deviatoric moment tensor	A type of the seismic moment tensor with trace equal to zero, meaning that there are no volumetric changes introduced by the corresponding source mechanism
DNN	Deep neural network	An ANN with many hidden layers between the input and output layers
DCNN	Deep convolutional neural network	A CNN with many hidden layers between the input and output layers
DT	Decision tree	A concept that uses a tree-like model of decisions and their possible consequences; commonly used in statistics, data mining, and supervised ML for both classification and regression tasks
EEW	Earthquake early warning	A system designed to detect a seismic event, determine its parameters and issue an alert to sites/areas where necessary actions should be taken before destructive seismic energy arrivals
FAST	Fingerprint and similarity thresholding	A seismic event detection method based on a data mining approach that involves extraction of discriminative features of seismic waveforms into compact "fingerprints" that are stored in a database
FIR filter	Finite impulse response filter	A type of filter whose impulse response is of finite duration
FNN	Feedforward neural network	The simplest type of ANN where the information moves in only one direction (forward) from the input nodes through the hidden nodes (if any) to the output nodes
FWI	Full waveform inversion	A method to estimate subsurface parameters from observed seismic waveforms using inversion theory
GAN	Generative adversarial network	A class of ML frameworks where two neural networks compete with each other in order to generate new synthetic instances of data that can pass for input real data.
GMM	Gaussian mixture model	A probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters; typically used for data clustering tasks
GNN	Graph neural network	A class of deep NNs designed to find relations in data described by graphs
GPD	Generalized phase detection	A CNN-based method for recognition of signal phases in a seismic waveform time series
GPU	Graphics processing unit	A specialized processor originally designed to accelerate graphics rendering; GPUs are very useful for ML because they can process many portions of data simultaneously
GPGPU	General-purpose graphics processing unit	A GPU used to perform computations in applications requiring massive vector operations and mathematical intensive problems (e.g., in science, finances, industry) traditionally handled by the central processing unit (CPU)
GraphBF	Graph-based bilateral filter	An image filter that smooths images while preserving edges by taking the weighted average of the nearby pixels; represented using graphs
HC	Haar cascades	ML object detection algorithms utilizing Haar features to determine the likelihood of a certain point being part of an object; Haar features are sequence of rescaled square shape functions proposed by Alfred Haar in 1909.
HMM	Hidden Markov model	A statistical Markov model (MM) in which the system being modeled is assumed to be a Markov process (MP) with unobserved (hidden) states
IIR filter	Infinite impulse response filter	A type of filter whose impulse response continues indefinitely (unlike the FIR filter)
K-means	K-means clustering	A distance-based clustering algorithm, where k is the number of clusters
LFS	Latent feature space	An abstract multi-dimensional space containing feature values for internal digital representation of observed data
LSM	Least squares method	A statistical method to find the best fit for a set of data values by minimizing the sum of the squared residuals – differences between an observed value and the fitted value provided by a model
LSTM	Long short-term memory	A type of RNN capable of learning order dependence in sequence prediction problems
MDN	Mixture density network	A class of NN obtained by combining a conventional ANN with a mixture density model – a model of probability

(continued on next page)

Table A.1 (*continued*)

Acronym	Description	Meaning
		distributions built up with a weighted sum of more simple distributions
MC	Markov chain	A stochastic model used to model pseudo-randomly changing systems, i.e. systems where it is assumed that future states do not depend on past states
ML	Machine learning	Machines or computer programs are learned to perform tasks that require natural intelligence displayed by animals including humans
MM	Markov model	A stochastic model used to model pseudo-randomly changing systems, i.e. systems where it is assumed that future states do not depend on past states
MP	Markov process	Same as Markov chain (MC)
NN	Neural network	A network or circuit of neurons, either organic or artificial in nature
PCA	Principal component analysis	A fast and flexible unsupervised method for data dimensionality reduction
PDE	Partial differential equation	An equation which imposes relations between multiple independent variables, an unknown function on those variables, and its partial derivatives
PINN	Physical informed neural network	A type of NN where a physical equation is used as a constraint
POCS	Projections onto convex sets	An algorithm used to find a point in the intersection of two closed convex sets
ResNN	Residual neural network	A DCNN model with residual blocks which aims to enhance the generalisation capability of NN
RF	Random forest	A commonly-used ML algorithm combining multiple decision trees (DT); designed for classification and regression tasks
RNN	Recurrent neural network	A class of ANN where connections between nodes form a directed or undirected graph along a temporal sequence
SMT	Seismic moment tensor	A mathematical representation of the moments generated by a seismic event, as symmetric tensor composed of 6 independent elements; SMT can be decomposed (non-uniquely) into three components: volumetric, DC and CLVD
SNR	Signal-to-noise ratio	A measure that compares the level of a desired signal to the level of background noise
SOM	Self-organizing map	A class of ANN that is trained using unsupervised learning to produce a low-dimensional discretized representation of the input, so-called Kohonen map; used for reduction of dimensionality
STA/LTA	Short-time-average through long-time-average	An algorithm designed for triggered seismic data acquisition; in other words, an algorithm for detection of seismic events based on analysis of amplitudes in short and long time windows
STFT	Short-time frequency transform	A sequence of Fourier transforms of a windowed signal changing over time; the procedure provides time-localized frequency information addressing signal non-stationarity.
SVM	Support vector machines	A set of supervised learning methods used for classification, regression and outliers detection
TNN	Transformer neural network	A class of NN that are able to efficiently track relationships in sequential data, e.g. text, signals, time series; it is mostly used in natural language processing and computer vision
U-Net	U-shaped NN	A CNN with a U-shaped architecture consisting of contracting and expansive paths; has been originally developed for image segmentation
VAE	Variational autoencoder	A type of autoencoder that addresses the issue of non-regularized latent feature space (LFS); VAEs are mainly used for data generation
WRI	Wavefield reconstruction inversion	A method that allows to mitigate non-linearity of the FWI by reconstructing a frequency-domain seismic wavefield using PDE

especially in light of the aforementioned new acquisition methods, providing more data with unique velocity information.

One aspect not yet highlighted in this review paper is the use of multi-physics information in order to get more accurate results. As en example, it is usually assumed there is a strong relation between induced seismic magnitude and fluid injection (McGarr, 2014). However, recent investigations exposed more complex relationships with various factors, like fluid-pressure rates, temperature variations and tectonic conditions (Cacace et al., 2021). We foresee that ML can assist in combining all available information in order to improve our microseismic analysis. An example can be found in Wozniakowska and Eaton (2020), who use tectonic, geological and geomechanical information in a logistic regression for predicting hydraulic-fracturing induced seismicity. Thus, it seems such an ML approach is fruitful and requires more investigation.

Finally, the biggest role we anticipate ML will play is in real-time monitoring. Processing in real-time helps in decision making during underground observation while post-processing of microseismicity allows deeper understanding of induced seismicity. The speed in which ML algorithms compute outputs in the inference stage make them prime candidates for real-time applications. The consistency and robustness of these outputs will depend on the NN model training and the adaptation. ML methods requiring training dataset from the whole dataset are generally difficult to apply in the real-time processing and are more suitable for post-processing where ML methods provide a high level of consistency without human bias. ML methods using synthetic datasets for training are generally more suitable for real-time processing. Real-time applications in ML are finding a home in self driving cars and medical monitoring, among many other fields. In the seismic arena, we anticipate that real-time microseismic monitoring will be firmly realised with machine learning.

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.

Data availability

No data was used for the research described in the article.

Acknowledgments

We thank the College of Petroleum Engineering and Geosciences (CPG) at KFUPM, KAUST, TU Delft, Tsinghua University, and the Basin Modelling Section at GFZ Potsdam for support. We are grateful to Nicolas Vinard for generating Fig. 4 and to the authors of publications whose figures we have re-used in our publication. We appreciate the valuable comments of the two anonymous reviewers who helped us to substantially improve the manuscript. The research resulted from collaboration between Seismik s.r.o. and CPG of KFUPM with contract

number CPG-21-0157.

Appendix A. List of acronyms

References

Abma, R., Kabir, N., 2006. 3d interpolation of irregular data with a POCS algorithm. GEOPHYSICS 71, E91–E97. https://doi.org/10.1190/1.2356088.

Akram, J., Ovcharenko, O., Peter, D., 2017. A robust neural network-based approach for microseismic event detection. In: SEG Technical Program Expanded Abstracts 2017Society of Exploration Geophysicists, pp. 2929–2933. https://doi.org/10.1190/segam2017-17761195.1.

Alam, F., Joty, S., Imran, M., 2018. Domain adaptation with adversarial training and graph embeddings. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, Melbourne, Australia, pp. 1077–1087. https://doi.org/10.18653/v1/P18-1099. URL:https://aclanthology.org/P18-1099.

Alkhalifah, T., Wang, H., Ovcharenko, O., 2022. MLReal: Bridging the gap between training on synthetic data and real data applications in machine learning. Artif. Intell. Geosci. 3, 101–114. https://doi.org/10.1016/j.aiig.2022.09.002.

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.

Andrade, J.N., van der Baan, M., 2021a. Real-time analysis and forecasting of the microseismic cloud size: Physics-based models versus machine learning. GEOPHYSICS 86, KS175–KS186. https://doi.org/10.1190/geo2021-0094.1.

Andrade, J.N., van der Baan, M., 2021b. Real-time prediction of the microseismic cloud size: A comparison between a physics-based model and a machine learning approach. In: First International Meeting for Applied Geoscience & Energy Expanded Abstracts. Society of Exploration Geophysicists, pp. 2026–2030. https://doi.org/10.1190/segam2021-3583194.1.

Anikiev, D., Valenta, J., Staněk, F., Eisner, L., 2014. Joint location and source mechanism inversion of microseismic events: benchmarking on seismicity induced by hydraulic fracturing. Geophys. J. Int. 198, 249–258. https://doi.org/10.1093/gji/ggu126.

Anikiev, D., Waheed, U.b., Staněk, F., Alexandrov, D., Eisner, L., 2021. Microseismic event location using artificial neural networks. In: First International Meeting for Applied Geoscience & Energy Expanded Abstracts. Society of Exploration Geophysicists, pp. 1661–1665. https://doi.org/10.1190/segam2021-3582729.1.

Anikiev, D., Waheed, U.b., Staněk, F., Alexandrov, D., Hao, Q., Iqbal, N., Eisner, L., 2022. Traveltime-based microseismic event location using artificial neural network. Front. Earth Sci. 10 https://doi.org/10.3389/feart.2022.1046258.

Ankerst, M., Breunig, M.M., Kriegel, H.P., Sander, J., 1999. OPTICS: Ordering points to identify the clustering structure. ACM SIGMOD Record 28, 49–60. https://doi.org/10.1145/304181.304187.

Ansari, A., Noorzad, A., Zafarani, H., 2009. Clustering analysis of the seismic catalog of Iran. Comput. Geosci. 35, 475–486. https://doi.org/10.1016/j.cageo.2008.01.010.

Asim, K.M., Idris, A., Iqbal, T., Martínez-Álvarez, F., 2018. Earthquake prediction model using support vector regressor and hybrid neural networks. PLOS ONE 13, e0199004. https://doi.org/10.1371/journal.pone.0199004.

Bardainne, T., Gaucher, E., 2010. Constrained tomography of realistic velocity models in microseismic monitoring using calibration shots. Geophys. Prospect. 58, 739–753. https://doi.org/10.1111/j.1365-2478.2010.00912.x.

Barragán-Montero, A., Javaid, U., Valdés, G., Nguyen, D., Desbordes, P., Macq, B., Willems, S., Vandewinckele, L., Holmström, M., Löfman, F., Michiels, S., Souris, K., Sterpin, E., Lee, J.A., 2021. Artificial intelligence and machine learning for medical imaging: A technology review. Phys. Med. 83, 242–256. https://doi.org/10.1016/j.ejmp.2021.04.016.

Barthwal, H., van der Baan, M., 2019. Passive seismic tomography using recorded microseismicity: Application to mining-induced seismicity. GEOPHYSICS 84, B41–B57. https://doi.org/10.1190/geo2018-0076.1.

Bauer, A., Schwarz, B., Gajewski, D., 2017. Utilizing diffractions in wavefront tomography. GEOPHYSICS 82, R65–R73. https://doi.org/10.1190/geo2016-0396.1.

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 363. https://doi.org/10.1126/science.aau0323.

Berhich, A., Belouadha, F.Z., Kabbaj, M.I., 2021. LSTM-based earthquake prediction: enhanced time feature and data representation. Int. J. High Perform. Syst. Archit. 10, 1. https://doi.org/10.1504/ijhpsa.2021.115499.

Beyreuther, M., Hammer, C., Wassermann, J., Ohrnberger, M., Megies, T., 2012. Constructing a hidden markov model based earthquake detector: application to induced seismicity. Geophys. J. Int. 189, 602–610. https://doi.org/10.1111/j.1365-246x.2012.05361.x.

Bhandarkar, T., K, V., Satish, N., Sridhar, S., Sivakumar, R., Ghosh, S., 2019. Earthquake trend prediction using long short-term memory RNN. Int. J. Electr. Comput. Eng. (IJECE) 9, 1304. https://doi.org/10.11591/ijece.v9i2.pp1304-1312.

Bhatia, A., Pasari, S., Mehta, A., 2018. EARTHQUAKE FORECASTING USING ARTIFICIAL NEURAL NETWORKS. Int. Archiv. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-5, 823–827. https://doi.org/10.5194/isprs-archives-xlii-5-823-2018.

Birnie, C., Alkhalifah, T., 2022. Leveraging domain adaptation for efficient seismic denoising. In: Energy in Data Conference, Austin, Texas, 20–23 February 2022. Energy in Data, pp. 11–15. https://doi.org/10.7462/eid2022-04.1.

Birnie, C., Chambers, K., Angus, D., Stork, A.L., 2016. Analysis and models of pre-injection surface seismic array noise recorded at the Aquistore carbon storage site. Geophys. J. Int. 206, 1246–1260. https://doi.org/10.1093/gji/ggw203.

Birnie, C., Hansteen, F., 2022. Bidirectional recurrent neural networks for seismic event detection. GEOPHYSICS 87, KS97–KS111. https://doi.org/10.1190/geo2020-0806.1.

Birnie, C., Jarraya, H., Hansteen, F., 2021. An introduction to distributed training of deep neural networks for segmentation tasks with large seismic data sets. GEOPHYSICS 86, KS151–KS160. https://doi.org/10.1190/geo2021-0130.1.

Birnie, C.E., 2018. Statistical methods for ambient noise characterisation, modelling and suppression: theory and applications for surface microseismic monitoring. Ph.D. thesis. University of Leeds.

Bishop, C.M., 1994. Mixture density networks. Neural Computing Research Group Report: NCRG/94/004.

Bolton, D.C., Shreedharan, S., Rivière, J., Marone, C., 2020. Acoustic energy release during the laboratory seismic cycle: Insights on laboratory earthquake precursors and prediction. J. Geophys. Res.: Solid Earth 125. https://doi.org/10.1029/2019jb018975.

Bouchon, M., Aki, K., 1977. Discrete wave-number representation of seismic-source wave fields. Bull. Seismol. Soc. Am. 67, 259–277. https://doi.org/10.1785/bssa0670020259.

Cacace, M., Hofmann, H., Shapiro, S.A., 2021. Projecting seismicity induced by complex alterations of underground stresses with applications to geothermal systems. Sci. Rep. 11 https://doi.org/10.1038/s41598-021-02857-0.

Carrizo Mascarell, M., 2020. Extracting geomechanical information from borehole microseismic data with machine learning. Master of science thesis. Delft University of Technology.

Cesca, S., 2020. Seiscloud, a tool for density-based seismicity clustering and visualization. J. Seismolog. 24, 443–457. https://doi.org/10.1007/s10950-020-09921-8.

Cesca, S., Şen, A.T., Dahm, T., 2013. Seismicity monitoring by cluster analysis of moment tensors. Geophys. J. Int. 196, 1813–1826. https://doi.org/10.1093/gji/ggt492.

Chai, C., Maceira, M., Santos-Villalobos, H.J., Venkatakrishnan, S.V., Schoenball, M., Zhu, W., Beroza, G.C., Thurber, C., 2020. Using a deep neural network and transfer learning to bridge scales for seismic phase picking. Geophys. Res. Lett. 47 https://doi.org/10.1029/2020gl088651.

Chambers, K., Kendall, J.M., Barkved, O., 2010. Investigation of induced microseismicity at Valhall using the Life of Field Seismic array. Lead. Edge 29, 290–295. https://doi.org/10.1190/1.3353725.

Chen, G., Li, J., 2022. CubeNet: Array-based seismic phase picking with deep learning. Seismol. Res. Lett. 93, 2554–2569. https://doi.org/10.1785/0220220147.

Chen, Y., 2020. Automatic microseismic event picking via unsupervised machine learning. Geophys. J. Int. 222, 1750–1764. https://doi.org/10.1093/gji/ggaa186.

Chen, Y., Saad, O.M., Savvaidis, A., Chen, Y., Fomel, S., 2022. 3D Microseismic Monitoring Using Machine Learning. J. Geophys. Res.: Solid Earth 127. https://doi.org/10.1029/2021jb023842.

Choi, J., Byun, J., Seol, S.J., Lee, S.K., 2022. Convolutional neural network-based moment tensor inversion using domain adaptation for microseismicity monitoring. Explor. Geophys. 1–11 https://doi.org/10.1080/08123985.2022.2086798.

Choi, W., Kim, W., Pyun, S., 2018. A simple inversion algorithm to estimate a linearly increasing velocity model for microseismic monitoring. Explor. Geophys. 49, 647–654. https://doi.org/10.1071/EG17104.

Chollet, F., et al., 2015. Keras: Deep learning library for Theano and TensorFlow. URL: https://keras.io.

Collins, D., Toya, Y., Pinnock, I., Shumila, V., Hosseini, Z., 2014. 3d velocity model with complex geology and voids for microseismic location and mechanism. In: Hudyma, M., Potvin, Y. (Eds.), Deep Mining 2014: Proceedings of the Seventh International Conference on Deep and High Stress Mining. Australian Centre for Geomechanics, Perth, pp. 681–688. https://doi.org/10.36487/ACG_rep/1410_48_Collins.

Dai, H., MacBeth, C., 1995. Automatic picking of seismic arrivals in local earthquake data using an artificial neural network. Geophys. J. Int. 120, 758–774. https://doi.org/10.1111/j.1365-246x.1995.tb01851.x.

Daniel, G., Fortier, E., Romijn, R., Oates, S., 2016. Location results from borehole microseismic monitoring in the Groningen gas reservoir, Netherlands. In: Sixth EAGE Workshop on Passive Seismic. European Association of Geoscientists & Engineers, pp. 1–6. https://doi.org/10.3997/2214-4609.201600016.

Das, S., Hobson, M.P., Feroz, F., Chen, X., Phadke, S., Goudswaard, J., Hohl, D., 2021. Microseismic event detection in large heterogeneous velocity models using bayesian multimodal nested sampling. Data-Centric Eng. 2, e1 https://doi.org/10.1017/dce.2021.1.

Delplancke, C., Fontbona, J., Prado, J., 2020. A scalable online algorithm for passive seismic tomography in underground mines. GEOPHYSICS 85, WA201–WA211. https://doi.org/10.1190/geo2019-0440.1.

Doğan, A., Demir, E., 2022. Structural recurrent neural network models for earthquake prediction. Neural Comput. Appl. https://doi.org/10.1007/s00521-022-07030-w.

Dramsch, J.S., 2020. 70 years of machine learning in geoscience in review. Adv. Geophys. 61, 1–55. https://doi.org/10.1016/bs.agph.2020.08.002.

Duan, Y., Shen, Y., Canbulat, I., Luo, X., Si, G., 2021. Classification of clustered microseismic events in a coal mine using machine learning. J. Rock Mech. Geotech. Eng. 13, 1256–1273. https://doi.org/10.1016/j.jrmge.2021.09.002.

Duncan, P.M., Eisner, L., 2010. Reservoir characterization using surface microseismic monitoring. GEOPHYSICS 75, 75A139–75A146. https://doi.org/10.1190/1.3467760.

Eisner, L., Abbott, D., Barker, W.B., Lakings, J., Thornton, M.P., 2008. Noise suppression for detection and location of microseismic events using a matched filter. SEG Techn. Program Expand. Abstr. 1431–1435 https://doi.org/10.1190/1.3059184.
van den Ende, M.P.A., Ampuero, J.P., 2020. Automated seismic source characterization using deep graph neural networks. Geophys. Res. Lett. 47 https://doi.org/10.1029/2020gl088690.
Esposito, A.M., Giudicepietro, F., D'Auria, L., Scarpetta, S., Martini, M.G., Coltelli, M., Marinaro, M., 2008. Unsupervised neural analysis of very-long-period events at Stromboli volcano using the self-organizing maps. Bull. Seismol. Soc. Am. 98, 2449–2459. https://doi.org/10.1785/0120070110.
Ester, M., Kriegel, H.P., Sander, J., Xu, X., et al., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Knowledge Discovery and Data Mining, pp. 226–231. doi:10.5555/3001460.3001507.
Fabregas, A.C., Arellano, P.B.V., Pinili, A.N.D., 2020. Long-short term memory (LSTM) networks with time series and spatio-temporal approaches applied in forecasting earthquakes in the philippines. In: Proceedings of the 4th International Conference on Natural Language Processing and Information Retrieval. ACM, pp. 188–193. https://doi.org/10.1145/3443279.3443288.
Fan, Z., Xu, X., 2019. Application and visualization of typical clustering algorithms in seismic data analysis. Proc. Comput. Sci. 151, 171–178. https://doi.org/10.1016/j.procs.2019.04.026.
Fedorenko, Y.V., Husebye, E.S., Ruud, B.O., 1999. Explosion site recognition; neural net discriminator using single three-component stations. Phys. Earth Planet. Inter. 113, 131–142. https://doi.org/10.1016/s0031-9201(99)00023-0.
Feng, T., Mohanna, S., Meng, L., 2022. EdgePhase: A deep learning model for multi-station seismic phase picking. Geochem. Geophys. Geosyst. 23 https://doi.org/10.1029/2022gc010453.
Foulger, G.R., Wilson, M.P., Gluyas, J.G., Julian, B.R., Davies, R.J., 2018. Global review of human-induced earthquakes. Earth Sci. Rev. 178, 438–514. https://doi.org/10.1016/j.earscirev.2017.07.008.
Gajek, W., Malinowski, M., 2021. Errors in microseismic events locations introduced by neglecting anisotropy during velocity model calibration in downhole monitoring. J. Appl. Geophys. 184, 104222 https://doi.org/10.1016/j.jappgeo.2020.104222.
Gajewski, D., Tessmer, E., 2005. Reverse modelling for seismic event characterization. Geophys. J. Int. 163, 276–284. https://doi.org/10.1111/j.1365-246x.2005.02732.x.
Galvis, I.S., Villa, Y., Duarte, C., Sierra, D., Agudelo, W., 2017. Seismic attribute selection and clustering to detect and classify surface waves in multicomponent seismic data by using k-means algorithm. Lead. Edge 36, 239–248. https://doi.org/10.1190/tle36030239.1.
Gan, H., Pan, X., Tang, K., Hu, N., Zhang, W., 2022. EWR-net: Earthquake waveform regularization network for irregular station data based on deep generative model and ResNet. J. Geophys. Res.: Solid Earth 127. https://doi.org/10.1029/2022jb024122.
Geller, R.J., 1997. Earthquake prediction: a critical review. Geophys. J. Int. 131, 425–450. https://doi.org/10.1111/j.1365-246x.1997.tb06588.x.
Gentili, S., Michelini, A., 2006. Automatic picking of P and S phases using a neural tree. J. Seismolog. 10, 39–63. https://doi.org/10.1007/s10950-006-2296-6.
Gesret, A., Desassis, N., Noble, M., Romary, T., Maisons, C., 2014. Propagation of the velocity model uncertainties to the seismic event location. Geophys. J. Int. 200, 52–66. https://doi.org/10.1093/gji/ggu374.
Gharti, H.N., Oye, V., Roth, M., Kühn, D., 2010. Automated microearthquake location using envelope stacking and robust global optimization. GEOPHYSICS 75, MA27–MA46. https://doi.org/10.1190/1.3432784.
Gibbons, S.J., Ringdal, F., 2006. The detection of low magnitude seismic events using array-based waveform correlation. Geophys. J. Int. 165, 149–166. https://doi.org/10.1111/j.1365-246x.2006.02865.x.
Giudicepietro, F., Esposito, A.M., Spina, L., Cannata, A., Morgavi, D., Layer, L., Macedonio, G., 2021. Clustering of experimental seismo-acoustic events using self-organizing map (SOM). Front. Earth Sci. 8 https://doi.org/10.3389/feart.2020.581742.
Grechka, V., Duchkov, A.A., 2011. Narrow-angle representations of the phase and group velocities and their applications in anisotropic velocity-model building for microseismic monitoring. GEOPHYSICS 76, WC127–WC142. https://doi.org/10.1190/geo2010-0408.1.
Grechka, V., Heigl, W.M., 2017. Microseismic Monitoring. Society of Exploration Geophysicists, Geophysical references series.
Greiner, T.A.L., 2022. Deep learning in seismic wavefield reconstruction and resolution enhancement. Ph.D. thesis. University of Oslo.
Greiner, T.A.L., Lie, J.E., Kolbjørnsen, O., Evensen, A.K., Nilsen, E.H., Zhao, H., Demyanov, V., Gelius, L.J., 2021. Unsupervised deep learning with higher-order total-variation regularization for multidimensional seismic data reconstruction. GEOPHYSICS 87, V59–V73. https://doi.org/10.1190/geo2021-0099.1.
Grigoli, F., Cesca, S., Krieger, L., Kriegerowski, M., Gammaldi, S., Horalek, J., Priolo, E., Dahm, T., 2016. Automated microseismic event location using master-event waveform stacking. Sci. Rep. 6, 25744. https://doi.org/10.1038/srep25744.
Grubas, S., Yaskevich, S., Duchkov, A., 2021. Localization of microseismic events using the physics-informed neural-network for traveltime computation. In: 82nd EAGE Annual Conference & Exhibition. European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.202113191.
Gu, C., Marzouk, Y.M., Toksoz, M.N., 2019. Bayesian deep learning and uncertainty quantification applied to induced seismicity locations in the Groningen gas field in the Netherlands: What do we need for safe AI?. In: SEG Technical Program Expanded Abstracts 2019Society of Exploration Geophysicists, pp. 2548–2553. https://doi.org/10.1190/segam2019-3216455.1.
Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D., 2019. A survey of methods for explaining black box models. ACM Comput. Surv. 51, 1–42. https://doi.org/10.1145/3236009.
Hammer, C., Ohrnberger, M., Fäh, D., 2012. Classifying seismic waveforms from scratch: a case study in the alpine environment. Geophys. J. Int. 192, 425–439. https://doi.org/10.1093/gji/ggs036.
Hao, Q., Waheed, U., Babatunde, M., Eisner, L., 2020. Microseismic hypocenter location using an artificial neural network. In: 82nd EAGE Annual Conference & Exhibition. European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.202010583.
He, Z., Peng, P., Wang, L., Jiang, Y., 2021. PickCapsNet: Capsule network for automatic p-wave arrival picking. IEEE Geosci. Remote Sens. Lett. 18, 617–621. https://doi.org/10.1109/lgrs.2020.2983196.
Holtzman, B.K., Paté, A., Paisley, J., Waldhauser, F., Repetto, D., 2018. Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field. Sci. Adv. 4 https://doi.org/10.1126/sciadv.aao2929.
Horne, S., Baird, A., Stork, A., Naldrett, G., 2019. Machine learning for DAS microseismic event detection. In: 81st EAGE Conference and Exhibition 2019 Workshop Programme. European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.201901972.
Huang, L., Li, J., Hao, H., Li, X., 2018. Micro-seismic event detection and location in underground mines by using convolutional neural networks (CNN) and deep learning. Tunn. Undergr. Space Technol. 81, 265–276. https://doi.org/10.1016/j.tust.2018.07.006.
Huang, X., Zhang, J., Du, X., Zhu, Z., Li, W., Zhang, Y., Ding, J., Liu, Z., Zhang, J., 2022. Compressive sensing aided seismic geometry design for offshore acquisition. Acta Geophys. https://doi.org/10.1007/s11600-022-00748-0.
Hulbert, C., Rouet-Leduc, B., Johnson, P.A., Ren, C.X., Rivière, J., Bolton, D.C., Marone, C., 2018. Similarity of fast and slow earthquakes illuminated by machine learning. Nat. Geosci. 12, 69–74. https://doi.org/10.1038/s41561-018-0272-8.
Huot, F., Ma, Y., Cieplicki, R., Martin, E., Biondi, B., 2017. Automatic noise exploration in urban areas. In: SEG Technical Program Expanded Abstracts 2017. Society of Exploration Geophysicists, pp. 5027–5032. https://doi.org/10.1190/segam2017-17774369.1.
Huot, F., Martin, E.R., Biondi, B., 2018. Automated ambient-noise processing applied to fiber-optic seismic acquisition (DAS). In: SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists, pp. 4688–4692. https://doi.org/10.1190/segam2018-2997880.1.
Izzatullah, M., Waheed, U.b., Yildirim, I.E., Alkhalifah, T., 2022. Predictive uncertainty quantification for bayesian physics-informed neural network (PINN) in hypocentre estimation problem. In: 83rd EAGE 2022 Annual Conference & Exhibition. European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.202210063.
Ji, Y., Hofmann, H., Duan, K., Zang, A., 2022. Laboratory experiments on fault behavior towards better understanding of injection-induced seismicity in geoenergy systems. Earth Sci. Rev. 226, 103916 https://doi.org/10.1016/j.earscirev.2021.103916.
Johnson, C.W., Ben-Zion, Y., Meng, H., Vernon, F., 2020a. Identifying different classes of seismic noise signals using unsupervised learning. Geophys. Res. Lett. 47, e2020GL088353 https://doi.org/10.1029/2020gl088353.
Johnson, P.A., Rouet-Leduc, B., Pyrak-Nolte, L.J., Beroza, G.C., Marone, C.J., Hulbert, C., Howard, A., Singer, P., Gordeev, D., Karaflos, D., Levinson, C.J., Pfeiffer, P., Puk, K. M., Reade, W., 2021. Laboratory earthquake forecasting: A machine learning competition. Proc. Nat. Acad. Sci. 118 https://doi.org/10.1073/pnas.2011362118.
Johnson, S.W., Chambers, D.J.A., Boltz, M.S., Koper, K.D., 2020b. Application of a convolutional neural network for seismic phase picking of mining-induced seismicity. Geophys. J. Int. 224, 230–240. https://doi.org/10.1093/gji/ggaa449.
Kamer, Y., Ouillon, G., Sornette, D., 2020. Fault network reconstruction using agglomerative clustering: applications to southern californian seismicity. Nat. Hazards Earth Syst. Sci. 20, 3611–3625. https://doi.org/10.5194/nhess-20-3611-2020.
Kao, H., Shan, S.J., 2004. The Source-Scanning Algorithm: mapping the distribution of seismic sources in time and space. Geophys. J. Int. 157, 589–594. https://doi.org/10.1111/j.1365-246x.2004.02276.x.
Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A., Kumar, V., 2019. Machine learning for the geosciences: Challenges and opportunities. IEEE Trans. Knowl. Data Eng. 31, 1544–1554. https://doi.org/10.1109/tkde.2018.2861006.
Kirschner, D., Howes, N., Daly, C., Mukherjee, J., Li, J., 2019. Detecting p- and s-wave arrivals with a recurrent neural network. In: SEG Technical Program Expanded Abstracts 2019Society of Exploration Geophysicists, pp. 2659–2662. https://doi.org/10.1190/segam2019-3215081.1.
Kocon, K., van der Baan, M., 2012. Quality assessment of microseismic event locations and traveltime picks using a multiplet analysis. Lead. Edge 31, 1330–1337. https://doi.org/10.1190/tle31111330.1.
Kong, Q., Trugman, D.T., Ross, Z.E., Bianco, M.J., Meade, B.J., Gerstoft, P., 2019. Machine learning in seismology: Turning data into insights. Seismol. Res. Lett. 90, 3–14. https://doi.org/10.1785/0220180259.
Konstantaras, A.J., Katsifarakis, E., Maravelakis, E., Skounakis, E., Kokkinos, E., Karapidakis, E., 2012. Intelligent spatial-clustering of seismicity in the vicinity of the hellenic seismic arc. Earth Sci. Res. 1, 1. https://doi.org/10.5539/esr.v1n2p1.
Kriegerowski, M., Petersen, G.M., Vasyura-Bathke, H., Ohrnberger, M., 2018. A Deep Convolutional Neural Network for Localization of Clustered Earthquakes Based on Multistation Full Waveforms. Seismol. Res. Lett. 90, 510–516. https://doi.org/10.1785/0220180320.
Kuang, W., Yuan, C., Zhang, J., 2021. Real-time determination of earthquake focal mechanism via deep learning. Nat. Commun. 12 https://doi.org/10.1038/s41467-021-21670-x.

Käufl, P., Valentine, A.P., O'Toole, T.B., Trampert, J., 2013. A framework for fast probabilistic centroid-moment-tensor determination—inversion of regional static displacement measurements. Geophys. J. Int. 196, 1676–1693. https://doi.org/10.1093/gji/ggt473.
Käufl, P., Valentine, A.P., de Wit, R.W., Trampert, J., 2016. Solving probabilistic inverse problems rapidly with prior samples. Geophys. J. Int. 205, 1710–1728. https://doi.org/10.1093/gji/ggw108.
Lary, D.J., 2010. Artificial intelligence in geoscience and remote sensing. In: Imperatore, P., Riccio, D. (Eds.), Geoscience and Remote Sensing New Achievements. InTech, Rijeka. https://doi.org/10.5772/9104. Chapter 7.
Laurenti, L., Tinti, E., Galasso, F., Franco, L., Marone, C., 2022. Deep learning for laboratory earthquake prediction and autoregressive forecasting of fault zone stress. Earth Planet. Sci. Lett. 598, 117825 https://doi.org/10.1016/j.epsl.2022.117825.
Li, J., 2013. Study of Induced Seismicity for Reservoir Characterization. Ph.D. thesis. Massachusetts Institute of Technology.
Li, J., Sadi Kuleli, H., Zhang, H., Nafi Toksöz, M., 2011a. Focal mechanism determination of induced microearthquakes in an oil field using full waveforms from shallow and deep seismic networks. GEOPHYSICS 76, WC87–WC101. https://doi.org/10.1190/geo2011-0030.1.
Li, J., Stankovic, L., Pytharouli, S., Stankovic, V., 2021. Automated platform for microseismic signal analysis: Denoising, detection, and classification in slope stability studies. IEEE Trans. Geosci. Remote Sens. 59, 7996–8006. https://doi.org/10.1109/tgrs.2020.3032664.
Li, J., Wu, X., Hu, Z., 2022a. Deep learning for simultaneous seismic image super-resolution and denoising. IEEE Trans. Geosci. Remote Sens. 60, 1–11. https://doi.org/10.1109/tgrs.2021.3057857.
Li, J., Yao, H., Wang, B., Yang, Y., Hu, X., Zhang, L., Ye, B., Yang, J., Li, X., Liu, F., Chen, G., Guo, C., Yang, W., 2022b. A real-time AI-assisted seismic monitoring system based on new nodal stations with 4g telemetry and its application in the yangbi MS 6.4 aftershock monitoring in southwest china. Earthq. Res. Adv. 2, 100033 https://doi.org/10.1016/j.eqrea.2021.100033.
Li, J., Zhang, H., Kuleli, H.S., Toksöz, M.N., 2011b. Focal mechanism determination using high-frequency waveform matching and its application to small magnitude induced earthquakes. Geophys. J. Int. 184, 1261–1274. https://doi.org/10.1111/j.1365-246x.2010.04903.x.
Li, J., Zhang, H., Rodi, W.L., Toksöz, M.N., 2013. Joint microseismic location and anisotropic tomography using differential arrival times and differential backazimuths. Geophys. J. Int. 195, 1917–1931. https://doi.org/10.1093/gji/ggt358.
Linardatos, P., Papastefanopoulos, V., Kotsiantis, S., 2020. Explainable AI: A review of machine learning interpretability methods. Entropy 23, 18. https://doi.org/10.3390/e23010018.
Liu, L., Cai, W., 2021. Multiscale deeponet for nonlinear operators in oscillatory function spaces for building seismic wave responses. ArXiv abs/2111.04860. doi:10.48550/ARXIV.2111.04860.
Liu, S., Birnie, C., Alkhalifah, T., 2021. Self-supervised learning for random noise removal in seismic data. In: 2nd Edition Artificially Intelligent Earth Exploration: Between Expectations and the Current Situation 2021. Society of Exploration Geophysicists, pp. 1–4.
Lomax, A., Michelini, A., Curtis, A., 2009. Earthquake Location, Direct, Global-Search Methods. Springer New York, New York, NY. https://doi.org/10.1007/978-0-387-30440-3_150, pp. 2449–2473.
Maity, D., Aminzadeh, F., Karrenbach, M., 2014. Novel hybrid artificial neural network based autopicking workflow for passive seismic data. Geophys. Prospect. 62, 834–847. https://doi.org/10.1111/1365-2478.12125.
McGarr, A., 2014. Maximum magnitude earthquakes induced by fluid injection. J. Geophys. Res.: Solid Earth 119, 1008–1019. https://doi.org/10.1002/2013jb010597.
McKean, S.H., Priest, J.A., Dettmer, J., Eaton, D.W., 2019. Quantifying fracture networks inferred from microseismic point clouds by a gaussian mixture model with physical constraints. Geophys. Res. Lett. 46, 11008–11017. https://doi.org/10.1029/2019gl083406.
Mignan, A., Broccardo, M., 2020. Neural network applications in earthquake prediction (1994–2019): Meta-analytic and statistical insights on their limitations. Seismol. Res. Lett. 91, 2330–2342. https://doi.org/10.1785/0220200021.
Montelli, R., Nolet, G., Masters, G., Dahlen, F.A., Hung, S.H., 2004. Global P and PP traveltime tomography: rays versus waves. Geophys. J. Int. 158, 637–654. https://doi.org/10.1111/j.1365-246X.2004.02346.x.
Moseley, B., Markham, A., Nissen-Meyer, T., 2020. Solving the wave equation with physics-informed deep learning. arXiv doi:10.48550/arXiv:2006.11894.
Mousavi, S.M., Beroza, G.C., 2020. A machine-learning approach for earthquake magnitude estimation. Geophys. Res. Lett. 47 https://doi.org/10.1029/2019gl085976.
Mousavi, S.M., Beroza, G.C., 2022. Deep-learning seismology. Science 377. https://doi.org/10.1126/science.abm4470.
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, 1–12. https://doi.org/10.1038/s41467-020-17591-w.
Mousavi, S.M., Zhu, W., Sheng, Y., Beroza, G.C., 2019. CRED: A deep residual network of convolutional and recurrent units for earthquake signal detection. Sci. Rep. 9 https://doi.org/10.1038/s41598-019-45748-1.
Moya, A., Irikura, K., 2010. Inversion of a velocity model using artificial neural networks. Comput. Geosci. 36, 1474–1483. https://doi.org/10.1016/j.cageo.2009.08.010.
Münchmeyer, J., Woollam, J., Rietbrock, A., Tilmann, F., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., et al., 2022. Which picker fits my data? A quantitative evaluation of deep learning based seismic pickers. J. Geophys. Res.: Solid Earth, e2021JB023499.
Musil, M., Plešinger, A., 1996. Discrimination between local microearthquakes and quarry blasts by multi-layer perceptrons and Kohonen maps. Bull. Seismol. Soc. Am. 86, 1077–1090. https://doi.org/10.1785/BSSA0860041077.
Münchmeyer, J., Bindi, D., Leser, U., Tilmann, F., 2020. The transformer earthquake alerting model: a new versatile approach to earthquake early warning. Geophys. J. Int. 225, 646–656. https://doi.org/10.1093/gji/ggaa609.
Münchmeyer, J., Bindi, D., Leser, U., Tilmann, F., 2021. Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network. Geophys. J. Int. 226, 1086–1104. https://doi.org/10.1093/gji/ggab139.
Nakata, N., Beroza, G.C., 2016. Reverse time migration for microseismic sources using the geometric mean as an imaging condition. GEOPHYSICS 81, KS51–KS60. https://doi.org/10.1190/geo2015-0278.1.
Nickolls, J., Dally, W.J., 2010. The GPU computing era. IEEE Micro 30, 56–69. https://doi.org/10.1109/mm.2010.41.
Novoselov, A., Sinkovics, K., Bokelmann, G., 2021. This earthquake doesn't exist. In: NeurIPS 2021 AI for Science Workshop. URL:https://openreview.net/forum?id=YJnpUHXNtQf.
Othman, A., Iqbal, N., Hanafy, S.M., Waheed, U.b., 2021. Automated event detection and denoising method for passive seismic data using residual deep convolutional neural networks. IEEE Trans. Geosci. Remote Sens. 60, 1–11. https://doi.org/10.1109/tgrs.2021.3054071.
Ouillon, G., Ducorbier, C., Sornette, D., 2008. Automatic reconstruction of fault networks from seismicity catalogs: Three-dimensional optimal anisotropic dynamic clustering. J. Geophys. Res. 113 https://doi.org/10.1029/2007jb005032.
Ouillon, G., Sornette, D., 2011. Segmentation of fault networks determined from spatial clustering of earthquakes. J. Geophys. Res. 116 https://doi.org/10.1029/2010jb007752.
Ovcharenko, O., Akram, J., Peter, D., 2018. Feasibility of moment tensor inversion from a single borehole data using artificial neural networks. Search Discov.
Peddie, J., 2022. The History of the GPU - Eras and Environment. Springer International Publishing. https://doi.org/10.1007/978-3-031-13581-1.
Perol, T., Gharbi, M., Denolle, M., 2018. Convolutional neural network for earthquake detection and location. Sci. Adv. 4 https://doi.org/10.1126/sciadv.1700578.
Pesicek, J.D., Cieślik, K., Lambert, M.A., Carrillo, P., Birkelo, B., 2016. Dense surface seismic data confirm non-double-couple source mechanisms induced by hydraulic fracturing. GEOPHYSICS 81, KS207–KS217. https://doi.org/10.1190/geo2016-0192.1.
Petersen, G.M., Niemz, P., Cesca, S., Mouslopoulou, V., Bocchini, G.M., 2020. Clusty, the waveform-based network similarity clustering toolbox: concept and application to image complex faulting offshore Zakynthos (Greece). Geophys. J. Int. 224, 2044–2059. https://doi.org/10.1093/gji/ggaa568.
Piegari, E., Herrmann, M., Marzocchi, W., 2022. 3d spatial cluster analysis of seismic sequences through density-based algorithms. Geophys. J. Int. https://doi.org/10.1093/gji/ggac160.
Poliannikov, O.V., Prange, M., Malcolm, A.E., Djikpesse, H., 2014. Joint location of microseismic events in the presence of velocity uncertainty. GEOPHYSICS 79, KS51–KS60. https://doi.org/10.1190/geo2013-0390.1.
Pu, Y., Apel, D.B., Hall, R., 2020. Using machine learning approach for microseismic events recognition in underground excavations: Comparison of ten frequently-used models. Eng. Geol. 268, 105519 https://doi.org/10.1016/j.enggeo.2020.105519.
Qu, S., Guan, Z., Verschuur, E., Chen, Y., 2020. Automatic high-resolution microseismic event detection via supervised machine learning. Geophys. J. Int. 222, 1881–1895. https://doi.org/10.1093/gji/ggaa193.
Razak, S.M., Jiang, A., Jafarpour, B., 2022. Latent-space inversion (lsi): a deep learning framework for inverse mapping of subsurface flow data. Comput. Geosci. 26, 71–99. https://doi.org/10.1007/s10596-021-10104-8.
Ren, C.X., Hulbert, C., Johnson, P.A., Rouet-Leduc, B., 2020. Machine learning and fault rupture: A review. Adv. Geophys. 61, 57–107. https://doi.org/10.1016/bs.agph.2020.08.003.
Revelo Obando, E.D., 2021. Moment tensor estimation through inversion of borehole microseismic data with machine learning. Master of science thesis. Delft University of Technology.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer, pp. 234–241.
Rosofsky, S., Huerta, E.A., 2022. Applications of physics informed neural operators. ArXiv doi:10.48550/ARXIV.2203.12634.
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/10.1785/0120180080.
Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C.J., Johnson, P.A., 2017. Machine learning predicts laboratory earthquakes. Geophys. Res. Lett. 44, 9276–9282. https://doi.org/10.1002/2017gl074677.
Rutledge, J.T., Phillips, W.S., Mayerhofer, M.J., 2004. Faulting induced by forced fluid injection and fluid flow forced by faulting: An interpretation of hydraulic-fracture microseismicity, Carthage Cotton Valley gas field, Texas. Bull. Seismol. Soc. Am. 94, 1817–1830. https://doi.org/10.1785/012003257.
Röth, G., Tarantola, A., 1994. Neural networks and inversion of seismic data. J. Geophys. Res. 99, 6753–6768. https://doi.org/10.1029/93jb01563.

Saad, O.M., Bai, M., Chen, Y., 2021. Uncovering the microseismic signals from noisy data for high-fidelity 3D source-location imaging using deep learning. GEOPHYSICS 86, KS161–KS173. https://doi.org/10.1190/geo2021-0021.1.

Saad, O.M., Chen, Y., 2020. Automatic waveform-based source-location imaging using deep learning extracted microseismic signals. GEOPHYSICS 85, KS171–KS183. https://doi.org/10.1190/geo2020-0288.1.

Saad, O.M., Chen, Y., 2021. Earthquake detection and p-wave arrival time picking using capsule neural network. IEEE Trans. Geosci. Remote Sens. 59, 6234–6243. https://doi.org/10.1109/tgrs.2020.3019520.

Saad, O.M., Chen, Y., Savvaidis, A., Chen, W., Zhang, F., Chen, Y., 2022a. Unsupervised deep learning for single-channel earthquake data denoising and its applications in event detection and fully automatic location. IEEE Trans. Geosci. Remote Sens. 60, 1–10. https://doi.org/10.1109/tgrs.2022.3209932.

Saad, O.M., Fomel, S., Abma, R., Chen, Y., 2022b. Unsupervised deep learning for 3d interpolation of highly incomplete data. GEOPHYSICS 1–99. https://doi.org/10.1190/geo2022-0232.1.

Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm—i. searching a parameter space. Geophys. J. Int. 138, 479–494. https://doi.org/10.1046/j.1365-246x.1999.00876.x.

Schoenball, M., Ellsworth, W.L., 2017. A systematic assessment of the spatiotemporal evolution of fault activation through induced seismicity in Oklahoma and Southern Kansas. J. Geophys. Res.: Solid Earth 122, 10189–10206. https://doi.org/10.1002/2017jb014850.

Seydoux, L., Balestriero, R., Poli, P., de Hoop, M., Campillo, M., Baraniuk, R., 2020. Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. Nat. Commun. 11 https://doi.org/10.1038/s41467-020-17841-x.

Shaheen, A., bin Waheed, U., Fehler, M., Sokol, L., Hanafy, S., 2021. GroningenNet: Deep learning for low-magnitude earthquake detection on a multi-level sensor network. Sensors (Basel, Switzerland) 21, 8080. https://doi.org/10.3390/s21238080.

Shapiro, S.A., Huenges, E., Borm, G., 1997. Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site. Geophys. J. Int. 131, F15–F18. https://doi.org/10.1111/j.1365-246x.1997.tb01215.x.

Sick, B., Guggenmos, M., Joswig, M., 2015. Chances and limits of single-station seismic event clustering by unsupervised pattern recognition. Geophys. J. Int. 201, 1801–1813. https://doi.org/10.1093/gji/ggv126.

Šílený, J., Panza, G.F., Campus, P., 1992. Waveform inversion for point source moment tensor retrieval with variable hypocentral depth and structural model. Geophys. J. Int. 109, 259–274. https://doi.org/10.1111/j.1365-246x.1992.tb00097.x.

Sipkin, S.A., 1982. Estimation of earthquake source parameters by the inversion of waveform data: synthetic waveforms. Phys. Earth Planet. Inter. 30, 242–259. https://doi.org/10.1016/0031-9201(82)90111-x.

Snover, D., Johnson, C.W., Bianco, M.J., Gerstoft, P., 2021. Deep clustering to identify sources of urban seismic noise in long beach, california. Seismol. Res. Lett. 92, 1011–1022. https://doi.org/10.1785/0220200164.

Song, C., Alkhalifah, T., 2019. Microseismic event estimation based on an efficient wavefield inversion. IEEE J. Select. Top. Appl. Earth Obs. Remote Sens. 12, 4664–4671. https://doi.org/10.1109/jstars.2019.2946903.

Song, C., Alkhalifah, T., 2020a. Identifying microseismic events in time-reversed source images using support vector machine. In: EAGE 2020 Annual Conference & Exhibition Online. European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.202010241.

Song, C., Alkhalifah, T., 2020b. Wavefield reconstruction inversion via machine learned functions. In: SEG Technical Program Expanded Abstracts 2020Society of Exploration Geophysicists, pp. 1710–1714. https://doi.org/10.1190/segam2020-3427351.1.

Song, C., Alkhalifah, T.A., 2022. Wavefield reconstruction inversion via physics-informed neural networks. IEEE Trans. Geosci. Remote Sens. 60, 1–12. https://doi.org/10.1109/tgrs.2021.3123122.

Soubestre, J., Shapiro, N.M., Seydoux, L., de Rosny, J., Droznin, D.V., Droznina, S.Y., Senyukov, S.L., Gordeev, E.I., 2018. Network-based detection and classification of seismovolcanic tremors: Example from the klyuchevskoy volcanic group in kamchatka. J. Geophys. Res.: Solid Earth 123, 564–582. https://doi.org/10.1002/2017jb014726.

Spurio Mancini, A., Piras, D., Ferreira, A.M.G., Hobson, M.P., Joachimi, B., 2021. Accelerating bayesian microseismic event location with deep learning. Solid Earth 12, 1683–1705. https://doi.org/10.5194/se-12-1683-2021.

Stork, A.L., Baird, A.F., Horne, S.A., Naldrett, G., Lapins, S., Kendall, J.M., Wookey, J., Verdon, J.P., Clarke, A., Williams, A., 2020. Application of machine learning to microseismic event detection in distributed acoustic sensing data. GEOPHYSICS 85, KS149–KS160. https://doi.org/10.1190/geo2019-0774.1.

Stork, A.L., Nixon, C.G., Hawkes, C.D., Birnie, C., White, D.J., Schmitt, D.R., Roberts, B., 2018. Is CO2 injection at Aquistore aseismic? A combined seismological and geomechanical study of early injection operations. Int. J. Greenhouse Gas Control 75, 107–124. https://doi.org/10.1016/j.ijggc.2018.05.016.

Tan, Y., He, C., Mao, Z., 2018. Microseismic velocity model inversion and source location: The use of neighborhood algorithm and master station method. GEOPHYSICS 83, KS49–KS63. https://doi.org/10.1190/geo2017-0308.1.

Thurber, C.H., Engdahl, E.R., 2000. Advances in global seismic event location. In: Thurber, C.H., Rabinowitz, N. (Eds.), Advances in Seismic Event Location. Springer, Netherlands, Dordrecht, pp. 3–22. https://doi.org/10.1007/978-94-015-9536-0_1.

Tiira, T., 1999. Detecting teleseismic events using artificial neural networks. Comput. Geosci. 25, 929–938. https://doi.org/10.1016/s0098-3004(99)00056-4.

Trugman, D.T., Shearer, P.M., 2017. GrowClust: A Hierarchical Clustering Algorithm for Relative Earthquake Relocation, with Application to the Spanish Springs and Sheldon, Nevada, Earthquake Sequences. Seismol. Res. Lett. 88, 379–391. https://doi.org/10.1785/0220160188.

Vinard, N., Drijkoningen, G., Verschuur, E., 2020. Real-data earthquake localization using convolutional neural networks trained with synthetic data. In: SEG Technical Program Expanded Abstracts 2020Society of Exploration Geophysicists, pp. 1576–1580. https://doi.org/10.1190/segam2020-3428107.1.

Vinard, N., Drijkoningen, G., Verschuur, E., 2021. Fine tuning a deep neural network to localize low magnitude earthquakes. In: 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET). IEEE, pp. 1–6. https://doi.org/10.1109/icecet52533.2021.9698462.

Vinard, N.A., Drijkoningen, G.G., Verschuur, D.J., 2022. Localizing microseismic events on field data using a U-Net-based convolutional neural network trained on synthetic data. GEOPHYSICS 87, KS33–KS43. https://doi.org/10.1190/geo2020-0868.1.

Waheed, U.b., Alkhalifah, T., Haghighat, E., Song, C., Virieux, J., 2021a. PINNtomo: Seismic tomography using physics-informed neural networks. arXiv preprint arXiv:2104.01588.

Waheed, U.b., Haghighat, E., Alkhalifah, T., Song, C., Hao, Q., 2021b. PINNeik: Eikonal solution using physics-informed neural networks. Comput. Geosci. 155, 104833 https://doi.org/10.1016/j.cageo.2021.104833.

Wamriew, D., Charara, M., Pissarenko, D., 2022. Joint event location and velocity model update in real-time for downhole microseismic monitoring: A deep learning approach. Comput. Geosci. 158, 104965 https://doi.org/10.1016/j.cageo.2021.104965.

Wamriew, D., Pevzner, R., Maltsev, E., Pissarenko, D., 2021. Deep neural networks for detection and location of microseismic events and velocity model inversion from microseismic data acquired by distributed acoustic sensing array. Sensors 21, 6627. https://doi.org/10.3390/s21196627.

Wamriew, D.S., Charara, M., Maltsev, E., 2020. Deep neural network for real-time location and moment tensor inversion of borehole microseismic events induced by hydraulic fracturing. In: SPE Russian Petroleum Technology Conference. SPE, pp. 1–15. https://doi.org/10.2118/201925-ms.

Wang, B., Zhang, N., Lu, W., Wang, J., 2019a. Deep-learning-based seismic data interpolation: A preliminary result. GEOPHYSICS 84, V11–V20. https://doi.org/10.1190/geo2017-0495.1.

Wang, H., Alkhalifah, T., 2018. Microseismic imaging using a source function independent full waveform inversion method. Geophys. J. Int. 214, 46–57. https://doi.org/10.1093/gji/ggy121.

Wang, H., Alkhalifah, T., 2021. Direct microseismic event location and characterization from passive seismic data using convolutional neural networks. GEOPHYSICS 86, KS109–KS121. https://doi.org/10.1190/geo2020-0636.1.

Wang, H., Alkhalifah, T., Waheed, U.b., Birnie, C., 2022. Data-driven microseismic event localization: An application to the oklahoma arkoma basin hydraulic fracturing data. IEEE Trans. Geosci. Remote Sens. 60, 1–12. https://doi.org/10.1109/TGRS.2021.3120546.

Wang, H., Guo, Q., Alkhalifah, T., Wu, Z., 2020. Regularized elastic passive equivalent source inversion with full-waveform inversion: Application to a field monitoring microseismic data set. GEOPHYSICS 85, KS207–KS219. https://doi.org/10.1190/geo2019-0738.1.

Wang, J., Teng, T.L., 1995. Artificial neural network-based seismic detector. Bull. Seismol. Soc. Am. 85, 308–319. https://doi.org/10.1785/bssa0850010308.

Wang, J., Xiao, Z., Liu, C., Zhao, D., Yao, Z., 2019b. Deep learning for picking seismic arrival times. J. Geophys. Res.: Solid Earth 124, 6612–6624. https://doi.org/10.1029/2019jb017536.

Wang, T., Trugman, D., Lin, Y., 2021. SeismoGen: Seismic waveform synthesis using GAN with application to seismic data augmentation. J. Geophys. Res.: Solid Earth 126, e2020JB020077. https://doi.org/10.1029/2020jb020077.

Warpinski, N., 2009. Microseismic monitoring: Inside and out. J. Petrol. Technol. 61, 80–85. https://doi.org/10.2118/118537-jpt.

Wilkins, A.H., Strange, A., Duan, Y., Luo, X., 2020. Identifying microseismic events in a mining scenario using a convolutional neural network. Comput. Geosci. 137, 104418 https://doi.org/10.1016/j.cageo.2020.104418.

Willacy, C., van Dedem, E., Minisini, S., Li, J., Blokland, J.W., Das, I., Droujinine, A., 2019. Full-waveform event location and moment tensor inversion for induced seismicity. GEOPHYSICS 84, KS39–KS57. https://doi.org/10.1190/geo2018-0212.1.

Wozniakowska, P., Eaton, D.W., 2020. Machine learning-based analysis of geological susceptibility to induced seismicity in the Montney Formation, Canada. Geophys. Res. Lett. 47 https://doi.org/10.1029/2020gl089651.

Yang, D.H., Zhou, X., Wang, X.Y., Huang, J.P., 2021a. Mirco-earthquake source depth detection using machine learning techniques. Inf. Sci. 544, 325–342. https://doi.org/10.1016/j.ins.2020.07.045.

Yang, L., Liu, X., Zhu, W., Zhao, L., Beroza, G.C., 2022. Toward improved urban earthquake monitoring through deep-learning-based noise suppression. Sci. Adv. 8 https://doi.org/10.1126/sciadv.abl3564.

Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J.H., Liao, Q., 2019. Deep learning for single image super-resolution: A brief review. IEEE Trans. Multimedia 21, 3106–3121. https://doi.org/10.1109/tmm.2019.2919431.

Yang, Y., Gao, A.F., Castellanos, J.C., Ross, Z.E., Azizzadenesheli, K., Clayton, R.W., 2021b. Seismic wave propagation and inversion with neural operators. ArXiv abs/2108.05421. doi:10.48550/ARXIV.2108.05421.

Yildirim, I., Waheed, U.b., Izzatullah, M., Alkhalifah, T., 2022. PINNhypo: Hypocenter localization using physics informed neural networks. In: 83rd EAGE 2022 Annual Conference & Exhibition. European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.202210773.

Yoon, C.E., O'Reilly, O., Bergen, K.J., Beroza, G.C., 2015. Earthquake detection through computationally efficient similarity search. Sci. Adv. 1, e1501057 https://doi.org/10.1126/sciadv.1501057.

Yu, S., Ma, J., 2021. Deep learning for geophysics: Current and future trends. Rev. Geophys. 59 https://doi.org/10.1029/2021rg000742.
Zhang, C., van der Baan, M., 2020. Microseismic denoising and reconstruction by unsupervised machine learning. IEEE Geosci. Remote Sens. Lett. 17, 1114–1118. https://doi.org/10.1109/lgrs.2019.2943851.
Zhang, G., Lin, C., Chen, Y., 2020a. Convolutional neural networks for microseismic waveform classification and arrival picking. GEOPHYSICS 85, WA227–WA240. https://doi.org/10.1190/geo2019-0267.1.
Zhang, H., Ma, C., Pazzi, V., Zou, Y., Casagli, N., 2020b. Microseismic signal denoising and separation based on fully convolutional encoder–decoder network. Appl. Sci. 10, 6621. https://doi.org/10.3390/app10186621.
Zhang, H., Yang, X., Ma, J., 2020. Can learning from natural image denoising be used for seismic data interpolation? GEOPHYSICS 85, WA115–WA136. https://doi.org/10.1190/geo2019-0243.1.
Zhang, M., Liu, M., Feng, T., Wang, R., Zhu, W., 2022a. LOC-FLOW: An end-to-end machine learning-based high-precision earthquake location workflow. Seismol. Res. Lett. https://doi.org/10.1785/0220220019.
Zhang, Q., Zhang, W., Wu, X., Zhang, J., Kuang, W., Si, X., 2022b. Deep learning for efficient microseismic location using source migration-based imaging. J. Geophys. Res.: Solid Earth 127. https://doi.org/10.1029/2021jb022649.
Zhang, X., Yuan, C., Zhang, J., Liu, S., Chen, Z., Li, W., 2018. Automatic microseismic detection and location via the deep-convolutional neural network. In: SEG Technical Program Expanded Abstracts 2018Society of Exploration Geophysicists, pp. 3057–3061. https://doi.org/10.1190/segam2018-2997548.1.
Zhang, X., Zhang, J., Yuan, C., Liu, S., Chen, Z., Li, W., 2020. Locating induced earthquakes with a network of seismic stations in oklahoma via a deep learning method. Sci. Rep. 10, 1941. https://doi.org/10.1038/s41598-020-58908-5.
Zhang, Z., Rector, J.W., Nava, M.J., 2017. Simultaneous inversion of multiple microseismic data for event locations and velocity model with bayesian inference. GEOPHYSICS 82, KS27–KS39. https://doi.org/10.1190/geo2016-0158.1.
Zheng, J., Lu, J., Peng, S., Jiang, T., 2017. An automatic microseismic or acoustic emission arrival identification scheme with deep recurrent neural networks. Geophys. J. Int. 212, 1389–1397. https://doi.org/10.1093/gji/ggx487.
Zhou, W., Liang, Y., Wang, X., Ming, Z., Xiao, Z., Fan, X., 2022. Introducing macrophages to artificial immune systems for earthquake prediction. Appl. Soft Comput. 122, 108822 https://doi.org/10.1016/j.asoc.2022.108822.
Zhou, Y., Yue, H., Kong, Q., Zhou, S., 2019. Hybrid event detection and phase-picking algorithm using convolutional and recurrent neural networks. Seismol. Res. Lett. 90, 1079–1087. https://doi.org/10.1785/0220180319.
Zhu, L., Peng, Z., McClellan, J., Li, C., Yao, D., Li, Z., Fang, L., 2019a. Deep learning for seismic phase detection and picking in the aftershock zone of 2008 Mw7.9 Wenchuan Earthquake. Phys. Earth Planet. Inter. 293, 106261 https://doi.org/10.1016/j.pepi.2019.05.004.
Zhu, W., Beroza, G.C., 2019. PhaseNet: A deep-neural-network-based seismic arrival time picking method. Geophys. J. Int. 216, 261–273. https://doi.org/10.1093/gji/ggy423.
Zhu, W., Mousavi, S.M., Beroza, G.C., 2019b. Seismic signal denoising and decomposition using deep neural networks. IEEE Trans. Geosci. Remote Sens. 57, 9476–9488. https://doi.org/10.1109/tgrs.2019.2926772.