Petroleum Science 22 (2025) 4520–4537

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Uncertainty-aware neural networks with manual quality control for hydraulic fracturing downhole microseismic monitoring: From automated phase detection to robust source location

Yi-Lun Zhang [a,b,c], Zhi-Chao Yu [d], Chuan He [c,*]

[a] School of Electrical and Information Engineering, Yunnan Minzu University, Kunming, 650504, Yunnan, China
[b] Yunnan Key Laboratory of Unmanned Autonomous System, Yunnan Minzu University, Kunming, 650504, Yunnan, China
[c] School of Earth and Space Sciences, Peking University, Beijing, 100871, China
[d] Key Laboratory of Deep Petroleum Intelligent Exploration and Development, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, 100029, China

ARTICLE INFO

Article history:
Received 19 April 2025
Received in revised form
29 May 2025
Accepted 22 July 2025
Available online 28 July 2025

Edited by Meng-Jiao Zhou

Keywords:
Microseismic monitoring
Phase detection
Phase arrival picking
Source location
Deep learning
Uncertainty estimation

ABSTRACT

Passive microseismic monitoring (PMM) serves as a fundamental technology for assessing hydraulic fracturing (HF) effectiveness, with a key focus on accurate and efficient phase detection/arrival picking and source location. In PMM data processing, the data-driven paradigm (deep learning based) outperforms the model-driven paradigm in characteristic extraction but lacks quality control and uncertainty quantification. Monte Carlo Dropout, a Bayesian uncertainty quantification technique, performs stochastic neuron deactivation through multiple forward propagation samplings. Therefore, this study proposes a deep learning neural network incorporating uncertainty quantification with manual quality control integration, establishing an optimized workflow spanning automated phase detection to robust source location. The methodology implementation comprises two principal components: (1) The MD-Net employing Monte Carlo Dropout strategy enabling simultaneous phase detection/arrival picking and uncertainty estimation; (2) an integrated hybrid-driven workflow with a traveltime-based inversion method for source location. Validation with field data demonstrates that MD-Net achieves superior performance under low signal-to-noise ratio conditions, maintaining detection accuracy exceeding 99% for both P- and S-waves. The phase arrival picking precision shows significant improvement, with a 40% reduction in standard deviation compared to the baseline model (P-S time difference decreasing from 12.0 ms to 7.1 ms), while providing quantifiable uncertainty metrics for manual calibration. Source location results further reveal that our hybrid-driven workflow produces more physically plausible event distributions, with 100% of microseismic events clustering along the primary fracture expanding direction. This performance surpasses traditional cross-correlation methods and single/multi-trace data-driven methods in spatial rationality. This study establishes an interpretable, high-precision automated framework for HF-PMM applications, demonstrating potential for extension to diverse geological settings and monitoring configurations.

1. Introduction

Recently, significant progress has been made in the exploration and development of unconventional oil and gas resources globally, especially reservoirs like tight sandstones and shales, which have gradually become important exploration areas (McGlade et al., 2013; Zheng et al., 2018). Among the newly developed reservoir modification technologies, hydraulic fracturing (HF) enhances hydrocarbon recovery by injecting high-pressure fluids into underground formations, creating fracture networks that alter local stress fields (Li et al., 2015, 2019; Bao and Eaton, 2016). These stress perturbations induce rock failures along pre-existing or newly formed fractures, generating microseismic events. The

* Corresponding author.
E-mail address: chuanheus@163.com (C. He).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2025.07.017

spatiotemporal distribution of these events directly reflects the propagation and complexity of HF-induced fractures. Passive microseismic monitoring (PMM) thus serves as a critical tool for evaluating HF effectiveness, enabling the characterization of fracture geometry, fluid migration pathways, and stimulated reservoir volume (SRV) (Maxwell and Urbancic, 2001). Downhole PMM captures signals by arranging multiple receivers in one or multiple monitoring wells adjacent to the treatment well, which is considered to be the most reliable monitoring method due to the proximity of the monitoring system to seismic source, low environmental noise, and high signal-to-noise ratio (SNR). Notably, while various unconventional reservoirs exhibit distinct lithological properties, reservoir types share critical commonalities in HF applications: ultralow permeability necessitating artificial fracture networks, stress-dependent fracture propagation behaviors, and microseismic signal attenuation characteristics. These commonalities validate their consolidated analysis in the PMM framework.

The comprehensive processing and interpretation of microseismic data constitutes the core of PMM, and its task lies in the analysis of useful signals/noises and the precise location of seismic sources. Currently, the mainstream microseismic data processing methods can be divided into two types according to whether they rely on the detection of vaild seismic phase: traveltime-based and waveform-based inversion methods (Pesicek et al., 2014; Li et al., 2020). For the traveltime-based inversion method, the kinematic parameters are used as constraints and optimization algorithms (e.g., nonlinear inversion or global search) are applied to realize the source location by constructing the objective function of the residuals between theoretical traveltimes and observed values (Jones et al., 2014). Microseismic signals with significantly high SNR in downhole observation provide a favorable condition for the phase detection and arrival picking, and thus the traveltime-based inversion method is highly applicable in such scenarios. Studies has demonstrated that the accurate phase detection (classification task) and arrival picking (regression task) with consistent standards significantly enhances the reliability of subsequent source location and fracture network inversion. A series of algorithms have been proposed and modified with improvement strategies focusing on: (1) constructing and optimizing objective detection functions in terms of energy distribution, spectral characteristics, polarization attributes, statistical properties and waveform similarity (Akram and Eaton, 2016; Shang et al., 2022); (2) constructing multi-attribute joint detection strategies to enhance characteristic differences; (3) implementing noise suppression preprocessing to improve SNRs (Mousavi and Langston, 2017); (4) using the waveform similarity characteristics among receivers (Tan and He, 2016; Bergen and Beroza, 2019); and (5) correlation detection of strong energy events based on template matching (Meng et al., 2018). These algorithms are all constructed with physics priori and characteristic parameterization, which could generally refer to as the model-driven paradigm. However, the limitations of this paradigm are restricted feature extraction capability, empirical dependence on parameter tuning, and insufficient generalization capability.

Recent advancements in emerging observation technologies have precipitated exponential growth in seismic monitoring datasets. Meanwhile, innovations of GPU-accelerated parallel computing have made quantum improvements in the training efficiency of deep learning (DL) models. Therefore, the data-driven paradigm based on adaptive data mining and modeling have made great progress (Bergen et al., 2019; Yu and Ma, 2021; Kong et al., 2019; Anikiev et al., 2023). In seismic phase detection and arrival picking tasks, the data-driven paradigm achieves optimal performance compared to other geophysical scenarios. The clear pattern recognition characteristics of this task combined with the availability of extensively well-labeled datasets enable direct implementation of mature DL architectures, including Convolutional Neural Networks (CNN), Recursive Neural Networks (RNN), U-Net, and Transformer, etc (Perol et al., 2018; Ross et al., 2018; Zhou et al., 2019; Mousavi et al., 2020; Zhu and Beroza, 2019; Zhang et al., 2021). In the source location task, the data-driven paradigm establishes end-to-end correlation models between seismic waveforms and spatial coordinates through the powerful nonlinear mapping capability of DL, effectively avoiding the complex forward modeling and iterative inversion of traditional methods. The typical applications include waveform-based fast location, source parameter-velocity structure joint inversion, etc. (Zhang et al., 2020; Wamriew et al., 2022). In addition, these methods have shown potential applications in seismic data denoising, focal mechanism resolution, first-motion polarity determination, magnitude prediction, and tectonic interpretation et al. (Kuang et al., 2021; Hu et al., 2021; Chen et al., 2024; Navarro-Rodríguez et al., 2025). Nevertheless, the data-driven paradigm has significant limitations: (1) the performance drops dramatically in data-scarce scenarios; (2) the black-box nature of pure data-driven paradigm leads to a lack of physical guidance and manual quality control; (3) the current research overly pursues the optimization of accuracy metrics while ignoring uncertainty quantification, model interpretability and other key issues, which seriously constrains the application of practical data processing.

Some researchers have begun to focus on the above limitations in geophysics fields (Wu et al., 2023; Zhou et al., 2025; Lan et al., 2023; Sun et al., 2024). Mousavi and Beroza (2020) proposed a single-earthquake location DL method based on Bayesian framework for uncertainty estimation. Sun et al. (2021) developed a hybrid training network by minimizing a data-driven model misfit and a physics-guided data residual simultaneously to determine subsurface elastic property. Wu et al. (2023, 2024) developed a fast Bayesian inversion of airborne electromagnetic data based on the invertible neural network (INN) for uncertainty evaluation and physics-guided neural network (PGNN) to mine physical constraint information for the inversion. However, existing methods mainly rely on data recordings from historical geophysics catalogs for model training, consequently exhibiting inadequate adaptability to the domain-specific requirements of HF monitoring scenarios. Furthermore, these methods neglect the inherent methodological disparities between data-driven and model-driven paradigms in establishing multi-stage processing workflow.

To address the above issues, this study proposes a comprehensive optimization framework integrating data-driven and model-driven paradigms, designed as a full-workflow solution with manual quality control specifically for the downhole PMM scenario in HF. Firstly, one robust DL model for seismic phase detection and arrival picking are adaptively constructed in order to maximize the automatic mapping advantage of the data-driven paradigm. A 2D U-Net processes multi-trace data as 3D volumes, enabling joint learning of “Intra-trace” waveform characteristics and “Inter-trace” phase arrival correlations. We employ the Monte Carlo Dropout strategy (MC-Dropout) to introduce the Bayesian uncertainty quantification (Gal and Ghahramani, 2016). This design achieves enhanced phase detection completeness and arrival picking accuracy, while generating confidence metrics for manual calibration. Secondly, based on the significant advantage of the model-driven paradigm in dealing with geophysical problems with sparse data and clear mechanisms, the subsequent velocity model construction and source location are performed using the traveltime-based inversion method to improve the reliability of the full-workflow.

2. Methodology

Microseismic data processing exhibits inherent nonlinearity characteristics and high uncertainty, convolutional data-driven paradigm based on point estimation is difficult to effectively capture uncertainty information, consequently diminishing interpretability of prediction results. To address above limitations, it is of great value to systematically carry out the uncertainty source analysis of the DL-based method and realize the uncertainty quantification. In this section, we first construct a data-driven framework integrating uncertainty quantification strategy for phase detection and arrival picking, and then form a hybrid-driven workflow for comprehensive microseismic data processing.

2.1. Uncertainty-aware DL framework for phase detection and arrival picking

Microseismic events induced by HF are usually characterized by low energy magnitude and high waveform similarity with ambient noise, consequently enhancing detection sensitivity becomes paramount, particularly for low SNR events. Currently, in the field of artificial intelligence (AI) seismology, the application of DL for seismic phase detection and arrival picking has gained wide academic and industrial recognition compared to other application scenarios. This comparative advantage originates from two key factors: (1) the inherent similarity between seismic phase detection tasks and computer vision-based pattern recognition, enabling straightforward migration of mature algorithms that significantly reduce the technical barriers of model construction; and (2) the availability of substantial labeled samples that facilitate training supervised machine learning models with strong generalization abilities. Meanwhile, regarding downhole microseismic monitoring systems, research focusing on DL-based model constructions and associated scientific issues has emerged as a critical direction.

2.1.1. Monte Carlo Dropout strategy

Uncertainty primarily encompasses two types in DL (Fig. 1): epistemic uncertainty and aleatoric uncertainty (Kiureghian and Ditlevsen, 2009). The former stems from the model's limited knowledge and can be mitigated by enhancing the model architecture or expanding training data, while the latter arises from inherent noises or ambiguity in the data, requiring explicit incorporation of uncertainty quantification during model design. The distinction between these two uncertainties in a regression framework is illustrated in Fig. 1(a). The black dashed curve represents the ground-truth function $y = f(x)$. Blue error bars denote observed data points contaminated with Gaussian noises, characterizing aleatoric uncertainty inherent to measurement processes. The gray shaded area delineates the 2σ confidence interval of a model's predictions, reflecting epistemic uncertainty due to incomplete knowledge. The uncertainty peaks near $x = 5$, simulating a localized knowledge gap (e.g., sparse training data or model inadequacy). Critically, aleatoric uncertainty persists globally as irreducible noises, while epistemic uncertainty diminishes with improved models or additional data. Consequently, our research focuses on quantifying the latter through systematic analysis.

Dropout is a widely adopted regularization technique in DL, first proposed by Srivastava et al. (2014) as a means to mitigate overfitting in neural networks. The core principle involves the stochastic deactivation of neurons during the training phase, where each neuron in a fully connected layer is temporarily dropped with a predefined probability p. This process introduces randomness that prevents the DL model from becoming overly dependent on specific neurons or local features, thereby improving its generalization performance.

Mathematically, for a neural network layer l (totally L) with 2D weight matrix $\mathbf{W}^{(l)} \in \mathbf{C}^{m\times n}$, input $\mathbf{H}^{(l-1)}$ and bias vector $\mathbf{b}^{(l)}$, the conventional output $\mathbf{H}^{(l)}$ can be denoted as:

$$\mathbf{H}^{(l)} = \sigma\left(\mathbf{W}^{(l)}\mathbf{H}^{(l-1)} + \mathbf{b}^{(l)}\right) \tag{1}$$

where σ is the activation function. Then, the above form of dropout is modified by element-wise multiplication with a Bernoulli-distributed mask $\mathbf{\Phi}^{(l)}$:

$$\mathbf{H_D}^{(l)} = \sigma\left(\mathbf{W_D}^{(l)}\mathbf{H_D}^{(l-1)} + \mathbf{b}^{(l)}\right)$$

$$\mathbf{W_D}^{(l)} = \mathbf{W}^{(l)} \odot \mathbf{\Phi}^{(l)} = \begin{bmatrix} w_{11}\phi_{11} & w_{12}\phi_{12} & \cdots & w_{1n}\phi_{1n} \\ w_{21}\phi_{21} & w_{22}\phi_{22} & \cdots & w_{2n}\phi_{2n} \\ \cdots & \cdots & & \cdots \\ w_{m1}\phi_{m1} & w_{m2}\phi_{m2} & \cdots & w_{mn}\phi_{mn} \end{bmatrix}$$

$$\mathbf{\Phi}^{(l)} \sim \text{Bernoulli}(1-p) \tag{2}$$

Here, $\odot$ denotes the Hadamard product (element-wise multiplication), where each entry of the resulting matrix is the product of the corresponding entries in $\mathbf{W}^{(l)}$ and $\mathbf{\Phi}^{(l)}$. During training, the mask is regenerated for each mini-batch, creating an ensemble effect where the network effectively learns multiple subnetworks simultaneously. The theoretical foundation of dropout can be interpreted through two complementary perspectives:

(a) Co-adaptation prevention: By disrupting fixed interaction patterns between neurons, dropout forces individual units to develop robust features independently of specific collaborators.
(b) Implicit model averaging: The training process approximates the optimization of an ensemble of multiple subnetworks, effectively performing approximate Bayesian marginalization over network structures.

Monte Carlo Dropout is an extension of the traditional dropout technique, proposed by Gal and Ghahramani (2016). Its core idea lies in extending the stochastic deactivation mechanism from the training phase to the prediction phase, enabling approximate Bayesian inference through multiple forward propagation samplings to quantify prediction uncertainty. Formally, this corresponds to learning a variational distribution that minimizes the KL divergence from the true parameter posterior. Fig. 1(b) illustrates the difference between dropout and MC-Dropout strategies during the training and prediction phases. Unlike dropout, which is only active during training, MC-Dropout retains stochastic neuron deactivation during prediction, forming a Monte Carlo estimation of the posterior distribution.

Let the trained neural network parameters (weight matrix $\overline{\mathbf{W}}$ and bias $\overline{\mathbf{b}}$) operate on the input x. In traditional inference, the output $\mathbf{H}^{(L)}$ is computed via deterministic forward propagation. In contrast, MC-Dropout generates independent Bernoulli masks during each forward pass and repeats sampling T times. The predictive distribution is denoted as:

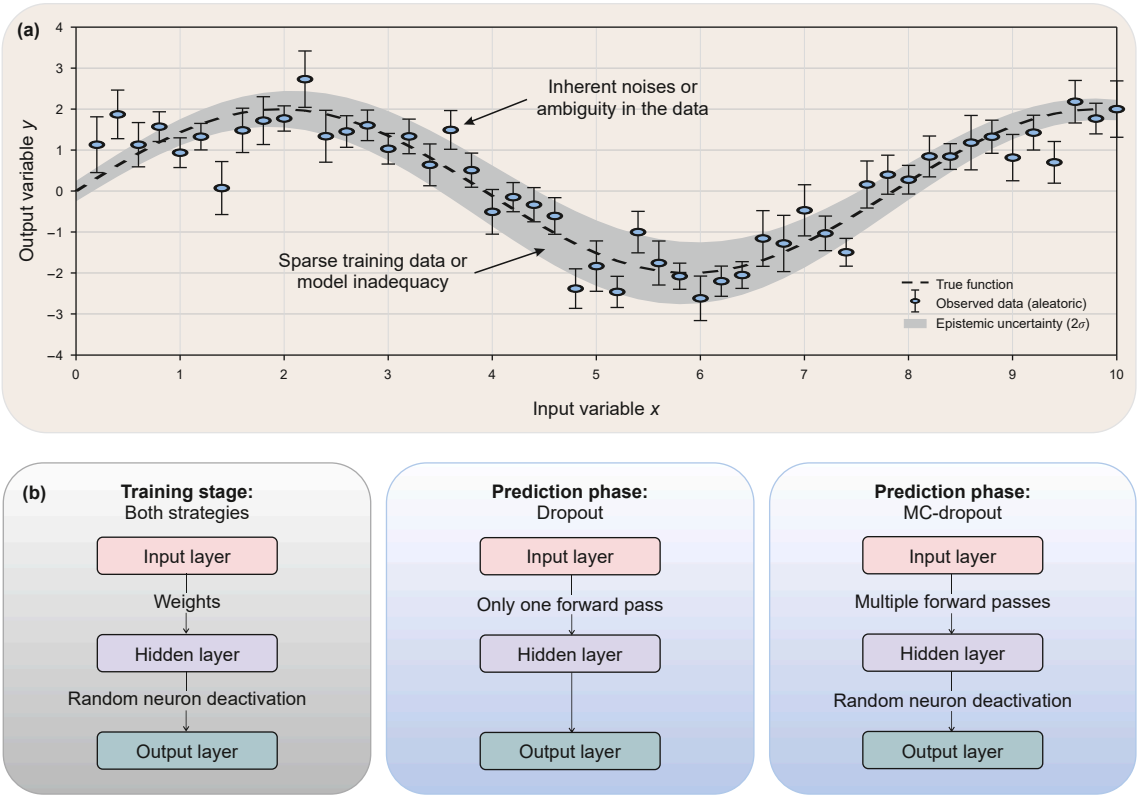


Fig. 1. Uncertainty characterization in DL: **(a)** aleatoric vs epistemic uncertainties and **(b)** dropout/MC-Dropout strategies during the training and prediction phases.

$$\begin{aligned} \mathbf{H}_{\mathrm{MD}}{}^{(L)}(t) &= \sigma(\mathbf{H}_{\mathrm{MD}}|x, \overline{\mathbf{W}}(t), \overline{\mathbf{b}}(t)) \\ \overline{\mathbf{W}}(t) &= \overline{\mathbf{W}} \odot \boldsymbol{\varepsilon}(t) \end{aligned} \tag{3}$$

where $\overline{\mathbf{W}}(t)$ and $\overline{\mathbf{b}}(t)$ represent the stochastic weights and bias for the t-th sampling, respectively. The mean (Me) and standard deviation (Sd) of the output $\mathbf{H}_{\mathrm{MD}}$ can be calculated as:

$$\begin{aligned} \mathrm{Me}(\mathbf{H}_{\mathrm{MD}}) &\approx \frac{1}{T}\sum_{t=1}^{T} \mathbf{H}_{\mathrm{MD}}{}^{(L)}(t), \\ \mathrm{Sd}(\mathbf{H}_{\mathrm{MD}}) &\approx \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{H}_{\mathrm{MD}}{}^{(L)}(t) - \mathrm{Me}(\mathbf{H}_{\mathrm{MD}})\right)^2} \end{aligned} \tag{4}$$

2.1.2. MD-Net

The U-Net model, a convolutional neural network architecture, was initially proposed by Ronneberger et al. (2015) for biomedical image segmentation tasks, demonstrating remarkable sensitivity in detecting edge features. Analogous to image segmentation, the phase detection and arrival picking of P-waves and S-waves can be conceptualized as a multi-class classification optimization problem. Within this framework, each sampling point in seismic waveforms should be categorized into one of three classes: P-wave, S-wave, or noise. Our study aims to achieve accurate classification of all sampling points while detecting edge characteristics at the boundaries between noise and valid seismic phases effectively, thereby enabling precise determination of phase arrival times. Based on the U-Net, we propose a new framework named MD-Net (Monte Carlo Dropout-Net) which constructs a feature-enhanced model using multi-trace microseismic recordings. By inputting target multi-trace waveform data into this established model, the workflow of phase detection and arrival picking can be finished automatically.

The network architecture of MD-Net is illustrated in Fig. 2(a). The model takes multi-trace segmented waveforms extracted from continuous microseismic recordings as input units, where three channels correspond to two horizontal components (E–W and N–S) and one vertical component (Z) of waveforms. The output layer generates probability distribution maps with identical dimensions to the input layer, comprising three independent channels representing probability distributions for P-wave, S-wave, and noise, respectively. Structurally, both input and output data are organized as 3D tensors of size $N \times M \times K$. The first dimension (N) denotes the "Intra-Trace" dimension, specifying the number of sampling points per trace, which should theoretically ensure complete inclusion of double-phase seismic events (e.g., full waveforms of both P- and S-phases). The second dimension (M) represents the "Inter-Trace" dimension, indicating the number of traces for joint processing, with its value determined by the actual spatial arrangement of receivers in the observation system. The third dimension (K) corresponds to the "Channel" dimension, representing three-component seismic recordings (two horizontal components and one vertical component), hence maintaining a fixed value of 3.

The MD-Net architecture is composed of symmetrical encoding and decoding paths, implementing hierarchical feature extraction and reconstruction through multiple downsampling and upsampling stages. The encoding path consists of five cascaded downsampling stages, each containing a convolutional layer and a pooling layer: (a) The 2D convolutional layer employs multiple convolutional kernels to extract spatiotemporal features. Convolution operations along the first dimension (temporal axis) and second dimension (spatial axis) jointly model "Intra-Trace" waveform characteristics and "Inter-Trace" phase arrival

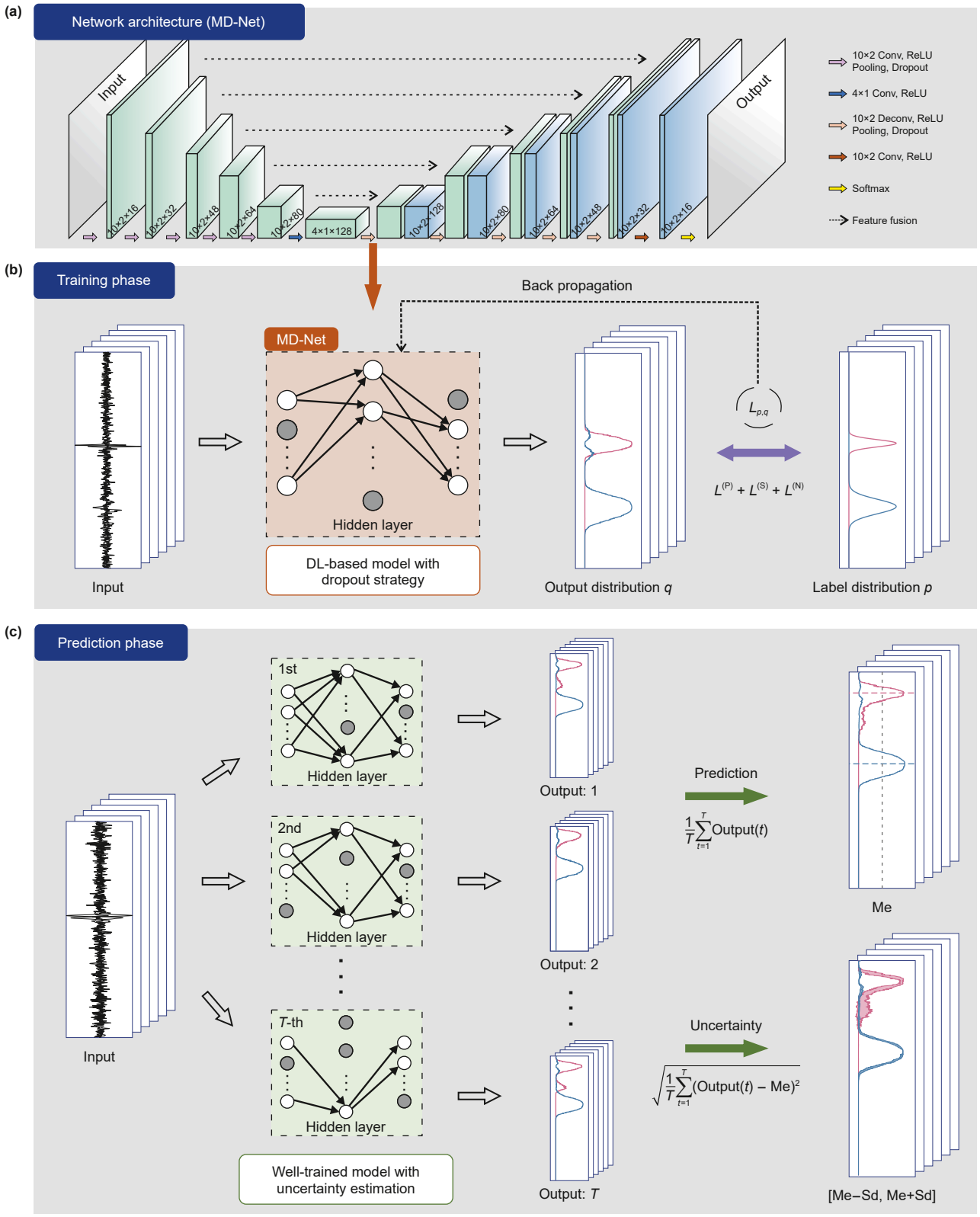


Fig. 2. MD-Net framework for uncertainty-aware phase detection and arrival picking: **(a)** network architecture; **(b)** training phase integrating Monte Carlo Dropout sampling and multi-loss optimization; and **(c)** prediction phase with uncertainty quantification (mean value [Me] and confidence interval [Me−Sd, Me + Sd] derived from *T* stochastic outputs).

correlations. (b) The max-pooling layer performs spatial downsampling on feature maps, preserving salient features while strategically reducing spatial resolution to control parameter growth. All convolutional kernels are configured with a size of (10, 2). The pooling factor is set to (5, 1) for the first two pooling layers and (2, 1) for the subsequent three layers. Through the decoding path, the single-channel sampling resolution of input data is progressively reduced to 1/200 of the initial value (N/200). To compensate for information loss, the network implements a channel expansion strategy synchronously to increase feature channels from 3 to 128. Each upsampling stage in the decoding path contains a transposed convolutional layer and a feature fusion layer: (a) Spatial dimensions are restored by the transposed convolution operation through deconvolutional kernels, with magnification factors strictly equivalent to the corresponding pooling factors in the encoding path. (b) The skip connection operation for feature fusion establishes cross-layer feature reuse by concatenating the same-scale feature maps from the downsampling and upsampling stages, ultimately constructing an end-to-end encoder-decoder architecture analogous to the U-Net framework. The deconvolutional kernels maintains parameter symmetry with convolutional kernels (kernel size = (10, 2)) to progressively recover the feature map to the original spatial resolution ($N \times M \times K$). All convolutional and deconvolutional convolutional layers employ ReLU nonlinear activation functions (Glorot et al., 2011). Finally, a Softmax layer converts the output into a probability distribution within the interval [0, 1]. The specific functional form adopted in this study is:

$$q_i(\text{label}) = \text{Softmax}(Z) = \frac{e^{Z_i(\text{label})}}{\sum_{j=1}^{3} e^{Z_j}} \tag{5}$$

where Z_i denotes the input value at the i-th sampling point of the Softmax layer. The labels 1–3 correspond to the P-wave, S-wave, and noise channels respectively, with q representing the output probability distribution among these three channels.

Consistent with the output probability distribution q, the sample labels also adopt probability distributions for the P-wave, S-wave, and noise channels. For samples including P- or S-wave arrivals requiring simultaneous phase detection and arrival picking, the three-channel probability distributions $p(\text{P})$, $p(\text{S})$, and $p(\text{N})$ are defined using manual phase arrival pickings as reference standards. Taking a typical single-trace double-phase sample label (Fig. 2(b)) as an example, we account for potential inaccuracies in manual phase arrival pickings (represented by red vertical lines in the figure: solid for P-wave arrival, dashed for S-wave arrival). The P-wave and S-wave labels are designed as Gaussian probability distributions centered at their respective arrivals ($\mu(\text{P})$ for P-wave, $\mu(\text{S})$ for S-wave), with standard deviations $\sigma(\text{P})$ and $\sigma(\text{S})$. The formulas for $p(\text{P})$, $p(\text{S})$, and $p(\text{N})$ at the i-th sampling point are given by:

$$\begin{cases} p_i(\text{P}) = \exp\left(-\dfrac{(i-\mu(\text{P}))^2}{2\sigma(\text{P})^2}\right) \\ p_i(\text{S}) = \exp\left(-\dfrac{(i-\mu(\text{S}))^2}{2\sigma(\text{S})^2}\right) \\ p_i(\text{N}) = 1 - p_i'(\text{P}) - p_i'(\text{S}) \end{cases} \tag{6}$$

For noise samples, they just need to be detected correctly. Thus, we set $p(\text{P}) = 0$, $p(\text{S}) = 0$, and $p(\text{N}) = 1$.

To integrate the Monte Carlo Dropout strategy into the network for uncertainty estimation, we introduce five dropout layers after each pooling operation in all downsampling stages and after skip connection operations in all upsampling stages. During both training and prediction phases of the MD-Net model, the dropout layers must remain activated (Gal and Ghahramani, 2016). The dropout rate directly influences the sensitivity of uncertainty estimation. If set too high, the model may become overly conservative, potentially obscuring genuine seismic phase characteristics (e.g., weak P-wave signals). Conversely, if set too low, the model may fail to adequately capture uncertainty caused by noise interference, thereby reducing robustness for low SNR events. Generally, low-SNR samples require higher dropout rates. In practice, differentiated dropout rates should be implemented based on model complexity and dataset qualities with varying SNR levels.

During training of the MD-Net model, a loss function is defined to quantify the discrepancy between model outputs and sample labels, with the optimization objective of minimizing this function. For the multi-class classification task in this study (identifying P-wave, S-wave, and non-phase noise in seismic waveforms), the standard categorical cross-entropy loss combined with a Softmax activation function can measure the discrepancy between the predicted probability distribution and true labels, expressed as:

$$L_{p,q} = -\sum_{\text{label}} \sum_{i} p_i(\text{label}) \cdot \log q_i(\text{label}) \tag{7}$$

However, this task suffers from a severe class imbalance issue: non-phase noise segments dominate the waveform data (e.g., in double-phase events shown in Fig. 2(b), P- and S-wave phases constitute ~10%, while noise accounts for ~90%). To address this, we propose a weighted cross-entropy loss function that dynamically assigns distinct weighting factors to phase and non-phase segments, thereby increasing the penalty for errors in phase segments to enhance the loss function's sensitivity to phase-related discrepancies (Fig. 2(b)). The loss function L is formulated as:

$$L_{p,q} = L_{p,q}^{(\text{P})} + L_{p,q}^{(\text{S})} + L_{p,q}^{(\text{N})} \tag{8}$$

where the sub-terms are defined as:

$$\begin{cases} L_{p,q}^{(\text{P})} = -\sum_i [\lambda_0 \cdot p_i(\text{P}) \cdot \log q_i(\text{P}) + \lambda_1 \cdot (1 - p_i(\text{P})) \cdot \log(1 - q_i(\text{P}))] \\ L_{p,q}^{(\text{S})} = -\sum_i [\lambda_0 \cdot p_i(\text{S}) \cdot \log q_i(\text{S}) + \lambda_1 \cdot (1 - p_i(\text{S})) \cdot \log(1 - q_i(\text{S}))] \\ L_{p,q}^{(\text{N})} = -\sum_i [\lambda_1 \cdot p_i(\text{N}) \cdot \log q_i(\text{N}) + \lambda_0 \cdot (1 - p_i(\text{N})) \cdot \log(1 - q_i(\text{N}))] \end{cases} \tag{9}$$

Here, λ_0 and λ_1 are dynamic weighting factors with $\lambda_0 > \lambda_1$, ensuring that prediction errors for phase segments (P- and S-waves) contribute more significantly to the total loss.

For a microseismic input sample, performing T forward propagation processes during prediction phase yields T output volumes with dimensions $N \times M \times 3$ (Fig. 2(c)). The final probability distribution is derived by computing Me and Sd across these outputs. Key points of MD-Net include:

(a) Phase detection: For one microseismic channel of a sample, we record the maximum value of the mean distribution and its corresponding sampling point position. For a sample, we compare the maximum values of the P-wave and S-wave mean probability distributions against preset thresholds to classify events as double-phase, single-phase, or noise. The setting of thresholds for P-wave and S-wave detection is critical. Excessively high thresholds may result in the misclassification of low SNR seismic phases as noise, while overly low thresholds could erroneously detect noise events

as valid seismic signals. The specific values for thresholds must be determined through case-by-case analysis, tailored to the specific characteristics of datasets and scenarios.

(b) Phase arrival picking: For one double-phase (including P-wave and S-wave) or single-phase event (including only P-wave or S-wave), the position corresponding to the maximum value in the mean probability distribution is defined as the predicted phase arrival time. The theoretical foundation lies in the fact that the statistical average of multiple sampling results can effectively suppress random noise interference, causing the position of the probability peak to converge towards the true arrival time.

(c) Uncertainty awareness: The interval $U = [\mathrm{Me}-\mathrm{Sd}, \mathrm{Me} + \mathrm{Sd}]$ provides a quantitative estimate of prediction uncertainty for each valid seismic phase. The range of U reflects the magnitude of uncertainty for the prediction results. By overlaying U onto microseismic waveforms, we can distinguish among "high-uncertainty phase arrivals" (likely caused by noise interference) and "low-uncertainty phase arrivals" (enough reliable seismic phases). By utilizing the acquired uncertainty information, we can improve the accuracy of phase arrival picking through manual calibration of model prediction results. Furthermore, applying the calibrated results to subsequent inversion processes contributes to enhancing the reliability of source location.

2.2. Workflow optimization for source location

The traditional model-driven paradigm suffer from inherent limitations due to their heavy dependence on a priori knowledge and oversimplification of complex geological conditions. In contrast, the data-driven paradigm significantly improve performance in complex scenarios by learning implicit patterns from large-scale datasets. Nevertheless, traditional paradigm can still be useful in interpretable and small-sample scenarios in combination with data-driven paradigm. The comparison of different paradigms is summarized in Table 1.

In PMM, rapid phase detection, arrival picking, and event location are needed and its core objective is to establish an automated processing workflow that is efficient, accurate, and has minimal human experience intervention. The MD-Net proposed in Section 2.1 can be automatically built to establish a mapping relationship between microseismic waveforms and detection/arrival picking results, realizing real-time processing of continuous recordings. On this basis, this study innovatively proposes a source location workflow that integrates data-driven and model-driven paradigms, taking into account the temporal characteristics of data acquisition. As shown in Fig. 3, for the phase detection and arrival picking tasks, the pure data-driven paradigm is adopted; for the subsequent tasks (velocity modeling, polarization analysis and source localization), the model-driven paradigm based on traveltime-based inversion method is adopted because the existing data-driven paradigm for these tasks are not mature enough. The specific steps are as follows:

(1) Initial dataset construction and benchmark model training. Firstly, upon acquiring the data of the first HF stage, the suspected microseismic events are quickly identified by loading the seismic phase detection and arrival picking pre-trained models or automatic model-driven methods (e.g., multi-trace cross-correlation method). Notably, pre-trained models from other regions or other HF projects in the same region can be transferred. If pre-trained models are not available, automatic model-driven methods can be utilized alone. Secondly, the suspected microseismic events are identified and classified manually, and valid seismic phases are picked, the results of which are used to create sample labels. The above process utilizes cross-validation of data-driven and model-driven paradigms to improve the labeling quality of the dataset. Finally, the training of the MD-Net benchmark model is completed using the framework detailed in Section 2.1.

(2) Progressive dataset augmentation and transfer learning. The benchmark model is applied to the second fracture section, and the continuous recordings are scanned to obtain detection and arrival picking results as well as associated uncertainty information of suspected microseismic events. The events with small uncertainty are directly cataloged, while the samples with large uncertainty undergo manual calibration. After removing misidentified samples and correcting the phase arrival picking results manually, the training dataset are augmented adaptably. Then, based on the idea of transfer learning, the model parameters are iteratively updated using the fine-tuning strategy to obtain the modified MD-Net, which realizes the progressive updating across fracturing stages (Hu et al., 2021; Zhang et al., 2022). The above process is looped until the end of the HF project. During this process, the size of the dataset is continuously expanded, and the detection and arrival picking abilities are correspondingly improving.

(3) Source inversion and location. Using the P/S-wave traveltime information obtained in the above steps as constraints and combining with the velocity modeling and azimuth information, the geophysical inversion algorithm is used to locate the source of all microseismic events. The velocity modeling can be implemented by establishing the initial velocity structures derived from sonic logging data and calibrated using the perforation events during HF. Source azimuths are resolved via the vector endpoint analysis of the polarization characteristics of P-waves. Hypocentral coordinates are determined through the traveltime grid search method with coordinate transformation. Specifically, theoretical traveltimes from each receiver to the hypothetical source are calculated and compared with the observed traveltimes to define the residual objective

Table 1
The comparison of different processing paradigms.

Paradigm	Scenario	Characteristic	Methods/Citations
Model-driven	Data scarcity with clear mechanisms	Relying on physical theories without requiring large datasets.	Multi-trace cross-correlation (Tan and He, 2016); template matching (Meng et al., 2018)
Data-driven	Data-rich with complex mechanisms	Leveraging the nonlinear fitting capability of data.	CNN (Perol et al., 2018); Transformer (Mousavi et al., 2020)
Hybrid-driven	Balancing efficiency and interpretability	Combining physical constraints and manual quality control with data-driven advantages.	Proposed workflow: MD-Net + traveltime-based inversion

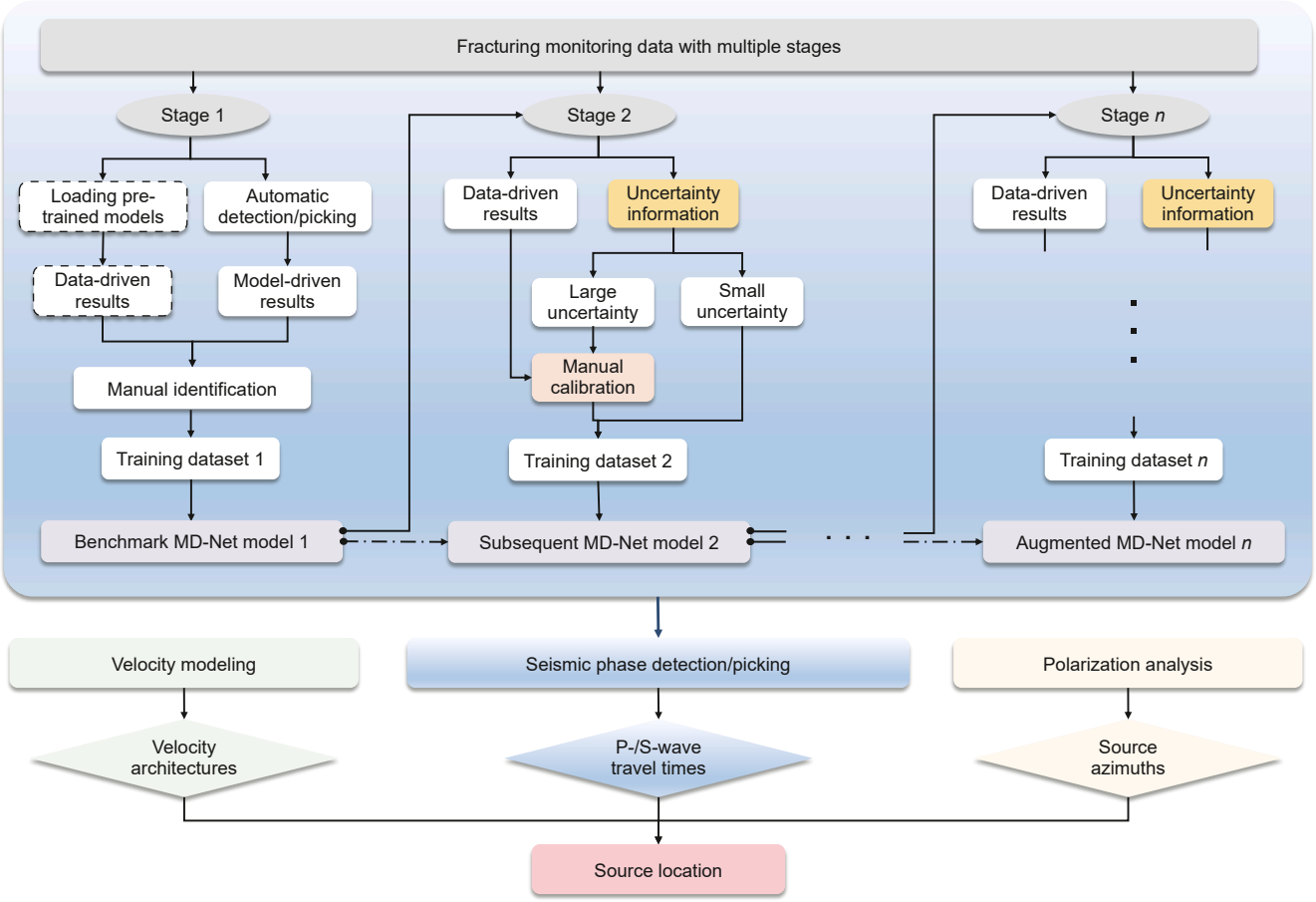


Fig. 3. Workflow optimization for source location of multi-stage fracturing monitoring using the hybrid driven paradigm.

function. The optimal source location is subsequently determined through iterative optimization search within the 3D space. For specifics on source azimuth determination, objective function formulation, and the search algorithm utilized, see Tan et al. (2017). The advantages of this hybrid-driven workflow are: (1) Utilizing the law of data acquisition in real-time data processing to construct an adaptive and automated process. The model of phase detection and arrival picking are constructed progressively during HF, and the model parameters are modified and the generalization ability is continuously enhanced. (2) The synergy between manual quality control and automated methods improves the accuracy while significantly reducing the uncertainty. (3) The innovative combination of model-driven and data-driven paradigms realizes the complementary advantages and forms a domain-adaptable hybrid-driven optimization framework.

In the workflow, manual quality control is systematically designed to minimize subjective judgment through standardized procedures. For samples predicted by MD-Net as high-uncertainty (confidence intervals exceeding predefined thresholds), independent manual annotations are performed by multiple geophysical analysts. These analysts adhere to unified guidelines for phase detection and arrival picking, ensuring consistency in labeling criteria. Their individual annotations are subsequently aggregated through consensus-based reconciliation. This multi-operator cross-validation approach effectively mitigates individual biases, reduces arbitrariness in ambiguous cases, and enhances the reproducibility of processing results. The finalized annotations are then fed back to update model, establishing a closed-loop workflow.

3. Field applications

3.1. Dataset and preprocessing

The dataset is acquired from a PMM program during one horizontal well HF project in a tight oil and gas field in eastern China, containing continuous monitoring recordings of 10 fracturing stages over approximately 16 h. The observation system consists of 15 three-component receivers deployed in an adjacent vertical monitoring well with 10 m spacing, operating at the sampling rate of 2000 Hz. Fig. 4(a) shows the observation system where target fracturing zones are located at about 2700 m depth, while the receiver array covers the interval of 2443–2673 m, with the location of ball-drop events indicated by red asterisks.

The preprocessing for raw microseismic recordings comprises three steps: (1) Spectral filtering: Application of a 30–350 Hz band-pass filter to suppress low-frequency ambient noise and high-frequency instrumental artifacts. (2) Receiver orientation: Alignment of horizontal components using perforation events (Fig. 4 (b)), resulting in enhanced waveform coherence and sharper phase arrivals. (3) Data standardization: Segmentation using a 1200-sample window (0.6 s) based on the observed P-S arrival time differences (<0.25 s), followed by demeaning and amplitude normalization via standard deviation.

The MD-Net architecture features an encoder-decoder structure with the input dimension of 1200 × 15 × 3. Firstly, the number of sampling points N within a single channel is reduced from 1200 to 6, and the number of channels K is increased from 3 to 128 by 5 downsampling stages; and then the original dimension is recovered gradually by 5 upsampling stages. The weighting factors $\lambda_0 = 6.7$ and $\lambda_1 = 1.2$ are derived from the inverse proportion of seismic phase segments (P-/S-waves, ~15%) to noise segments (~85%) in one waveform sample, prioritizing penalties for seismic phase segments.

Four sample categories are defined: background noise, noise event, single-phase event, and double-phase event. Following monitoring practice, only single- and double-phase events are retained as valid samples, with others classified as noises in the dataset construction. Manual arrival picking uncertainties were set at $\sigma_P = 10$ ms (20 samples) and $\sigma_S = 20$ ms (40 samples) respectively, reflecting the superior P-wave detectability due to higher SNR level (Fig. 2(b)).

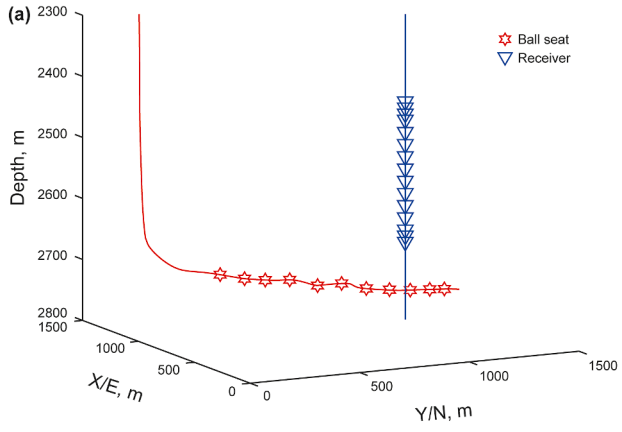
The majority of the segmented recordings consist of noises, with valid seismic phase samples representing only a small fraction. If all samples from the continuous recordings are used as inputs for model training, a typical sample imbalance problem in DL may occur, resulting in poor performance for the valid seismic phase samples (López et al., 2013). To address this issue, a sampling strategy was employed to select a subset of noise samples alongside all available seismic phase samples for the training dataset. Ultimately, it was determined that the training dataset originated from the first seven fracturing stages, comprising 32 single-phase samples (480 traces), 436 double-phase samples (6540 traces), and 815 noise samples (12,225 traces). The testing dataset are from the last three fracturing stages, containing a total of 29,307 samples. As the evaluation criteria, the monitoring recordings from these four stages are manually detected and picked, yielding 8 single-phase samples (120 traces) and 85 double-phase samples (1275 traces).

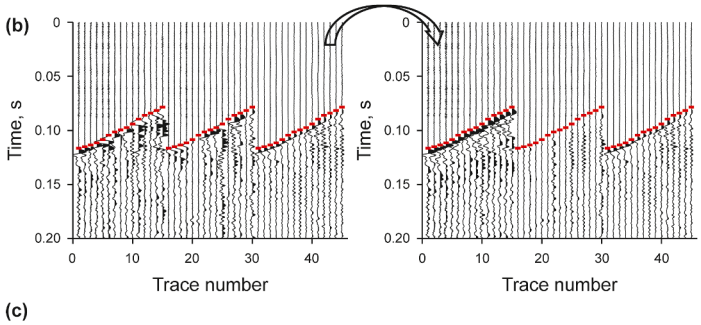
The model implementation utilizes the Keras API within the TensorFlow framework. During the training phase, a mini-batch gradient descent method is employed with a batch size of 32, allowing samples to be fed into the model in batches for loss function calculation and network parameter updates. The Adam optimizer, with an initial learning rate of 0.001, is used to update the model parameters through adaptive moment estimation, which enhances the convergence speed of the model (Kingma and Ba, 2015). The entire training dataset undergoes multiple iterations, and the training is concluded when the loss function no longer decreased significantly. Additionally, 10% of the samples are split as a validation dataset to fine-tune the model hyperparameters. The adjusted and optimized MD-Net serves as the final training model for subsequent testing.

3.2. Phase detection/arrival picking and uncertainty estimation

In the MD-Net architecture, four distinct dropout rates (0, 0.1, 0.2, 0.3) are implemented based on empirical optimization from (Lan et al. (2023)) and (Mousavi and Beroza (2020)), who demonstrated their efficacy in balancing uncertainty quantification and computational efficiency for CNN-based geophysical inversion tasks. These values mitigate over-regularization while preserving sensitivity in low-SNR conditions. Fig. 5 demonstrates accuracy curves and loss function convergence curves of the four models during training. All models exhibit stable decreasing trends in loss functions with increasing iterations, ultimately converging to minimal values. During testing, the number of forward propagation times T is theoretically proportional to prediction accuracy. In this study, T is empirically set to 75 through systematic optimization. For phase arrival picking, the final arrival time prediction is derived from the ensemble mean of T stochastic forward propagation, yielding three experimental groups: T-MD1 (MC-Dropout, dropout rate = 0.1), T-MD2 (0.2), T-MD3 (0.3). To validate the effectiveness of the MD-Net proposed in this paper, we establish three control groups T-Dr1 (dropout, dropout rate = 0.1), T-Dr2 (0.2), and T-Dr3 (0.3), all of which keep the dropout active only during training. Another baseline group T-Base (dropout rate = 0) is set up with the dropout not enabled in both training and testing phases. Architectural differences among experimental groups are shown in Table 2. For phase detection, threshold setting for the output probability distribution of P- and S-wave channels is critical. Overly conservative thresholds risk misclassifying low SNR events as noises, while lenient thresholds may induce false positives. Through comprehensive sensitivity analysis, optimal thresholds for both P- and S-wave channels in our case are set at 0.5.

Three representative double-phase waveforms from the testing dataset (Figs. 6, 8 and 10) are analyzed to evaluate phase detection and arrival picking performance of all experimental groups. Corresponding uncertainty estimations (Sd of output probability distributions) of T-MD series are presented in Figs. 7, 9 and 11. The first high SNR example (Figs. 6 and 7) shows all models successfully detecting both phases above thresholds. However, T-Base exhibits insignificant amplitude variations (<0.5) near the true P-





(c)

Field data	For training							For testing		
Fracturing stage	1	2	3	4	5	6	7	8	9	10
Valid event number	105	116	78	31	34	36	66	60	18	7

Fig. 4. Field dataset from a PMM program with multiple fracturing stages: **(a)** observation system; **(b)** receiver orientation (horizontal component 1: channels 1–15, horizontal component 2: channels 16–30, and vertical component: channels 31–45; red markers: P-wave arrivals); and **(c)** valid double-phase events for model training and testing.

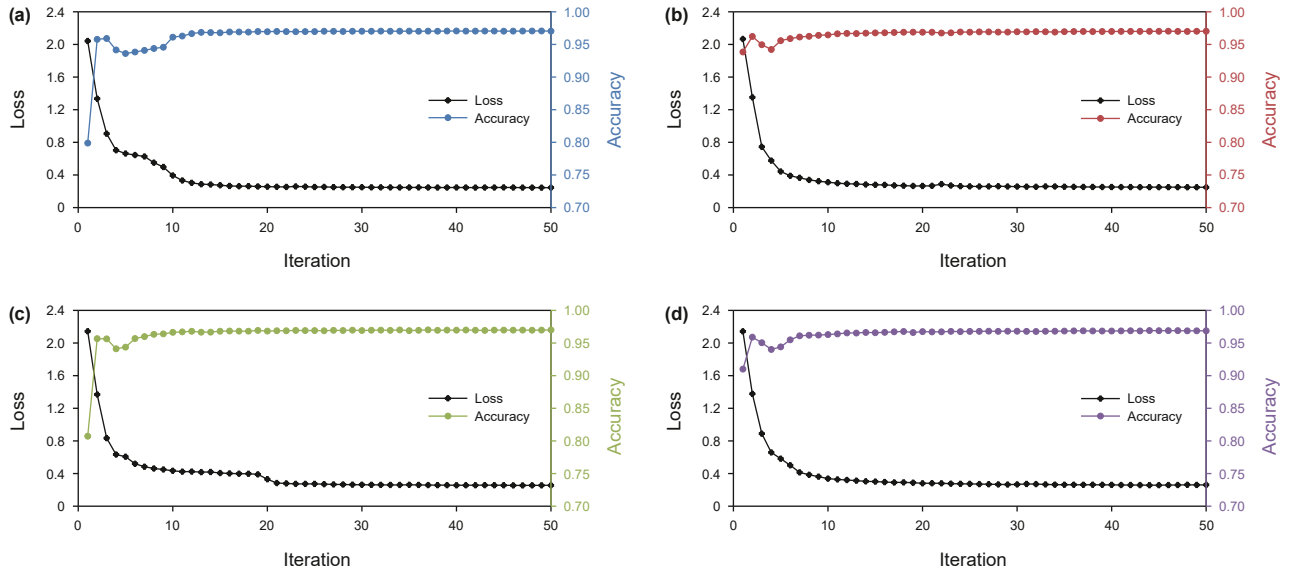


Fig. 5. Loss and accuracy curves during training: **(a)** dropout rate = 0, **(b)** dropout rate = 0.1, **(c)** dropout rate = 0.2, and **(d)** dropout rate = 0.3.

Table 2
Architectural differences among experimental groups.

Group	Training phase	Prediction phase	Forward propagation samplings	Uncertainty quantification
T-Base	No dropout	No dropout	Deterministic (single time)	Not applicable
T-Dr series	Dropout (rate: 0.1/0.2/0.3)	No dropout	Deterministic (single time)	Not applicable
T-MD series	Dropout (rate: 0.1/0.2/0.3)	Dropout (rate: 0.1/0.2/0.3)	Stochastic (multiple times)	Enabled via ensemble mean

wave arrival (Fig. 6(b)). For phase arrival picking, T-Base mistakenly picks the P-wave to a wrong position, but the picking results of the rest of groups do not notably deviate from the true arrival. Further analysis shows that there is a typical noise interference near the position of P-wave arrival picked by T-Base, and the jumping phenomenon can also be observed there in Fig. 6(c), (d), (f), and (h). The uncertainty estimation (Fig. 7) reveals elevated uncertainties near phase arrivals, particularly for the P-wave, mostly attributable to the typical noise interference mentioned above. This demonstrates MC-Dropout's double advantages: mitigating large picking errors through stochastic regularization while providing quantitative reliability metrics for prediction results, which is crucial for the subsequent manual calibration.

The second sample presents a low SNR trace (Figs. 8 and 9), where the P-wave arrival is entirely obscured by background noise (Fig. 8(a)). This challenging case demonstrates phase detection capability even for manual analysis, requiring integration with the similarity of adjacent trace waveforms or template matching techniques for reliable detection. As shown in Fig. 8(b)–(h), all groups exhibit significant misclassification manifested through phase confusion between P- and S-waves, irregular output probability distribution morphologies, and substantial deviation of arrival time picks. The T-MD series display elevated prediction uncertainties throughout the seismic phases (Fig. 9(b)–(d)). Detailed analysis of the true P-wave arrival position reveals distinct behavioral patterns: The T-MD1 shows no significant jumping in either mean probability distribution (Fig. 8(d)) or uncertainty estimation (Fig. 9(b)), indicating background noise. In contrast, T-MD2 and T-MD3 demonstrate measurable probability jumping (Fig. 8(f) and (h)) accompanied by elevated uncertainties (Fig. 9(c) and (d)), indicating uncertain phase. This case highlights the necessity for synthesizing uncertainty metrics from multiple MD-Net configurations with different dropout rates and the waveform characteristics among traces for manual analysis.

The third sample presents another low SNR trace (Figs. 10 and 11). The P-wave phase could be correctly detected by all experimental groups, while the T-Base, T-Dr1, and T-MD1 do not have a significant jump near the S-wave arrival, erroneously classifying this sample as a single-phase event. Groups with higher dropout rates (T-Dr2/3, T-MD2/3) maintain better detection performance near the S-wave arrival, consistent with the generalization enhancement effect of stronger stochastic regularization. In picking accuracy, T-Dr2 and T-MD2 achieve precise S-wave picking, whereas T-Dr3 and T-MD3 exhibit insufficient probability jumping at true S-wave arrival with notable picking deviations. This sensitivity to dropout rate hyperparameter underscores the importance of maintaining appropriate regularization intensity during model configuration. Uncertainty visualization (Fig. 11(b)–(d)) confirms the essential role of manual calibration for such ambiguous low SNR cases, with prediction reliability metrics providing crucial decision-support information for practical scenarios.

From results, all experimental groups achieved zero false positives in detecting noise samples as valid microseismic events (including single-phase and double-phase events) across the entire testing dataset, demonstrating MD-Net's robustness against false alarms under diverse noise conditions. Systematic evaluation of microseismic samples across Figs. 6–11 reveals two critical findings: Firstly, the MC-Dropout technique enables effective uncertainty quantification without compromising phase detection or arrival picking accuracy. Secondly, the relationship between dropout rate and model generalization follows a nonlinear pattern, necessitating systematic optimization through cross-validation procedures.

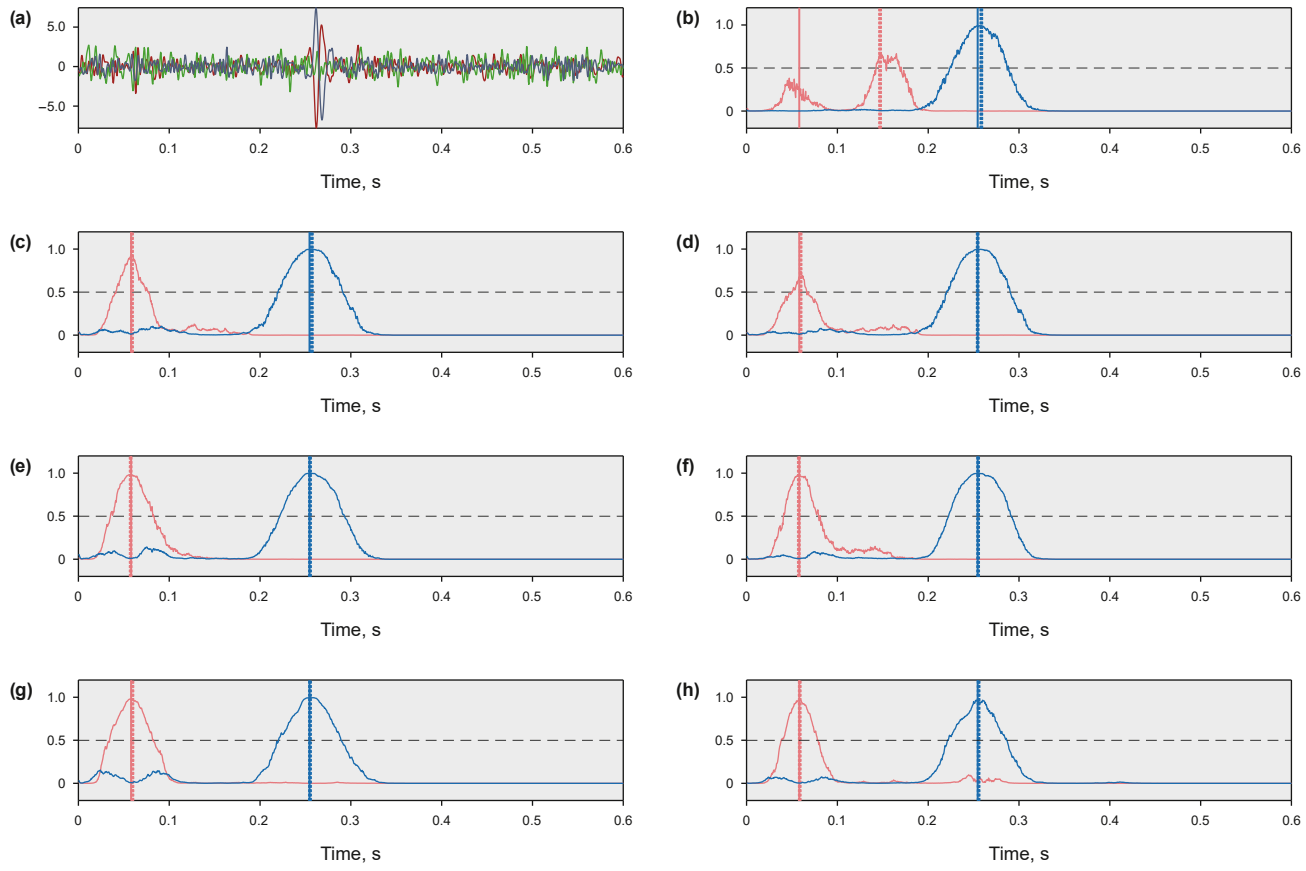


Fig. 6. Phase detection and arrival picking results of different groups (double-phase example 1): **(a)** three component single-trace seismogram, **(b)** T-Base, **(c)** T-Dr1, **(d)** T-MD1, **(e)** T-Dr2, **(f)** T-MD2, **(g)** T-Dr3, and **(h)** T-MD3. In **(b)**, **(c)**, **(e)**, and **(g)**, the red and blue curves represent the output probability distributions of P- and S-waves, respectively; while in **(d)**, **(f)**, and **(h)**, the curves represent the Me values of the output probability distributions over *T* iterations. All subfigures show manual picks (solid vertical lines) and model-predicted picks (dashed vertical lines), with probability thresholds marked by horizontal black dashed lines.

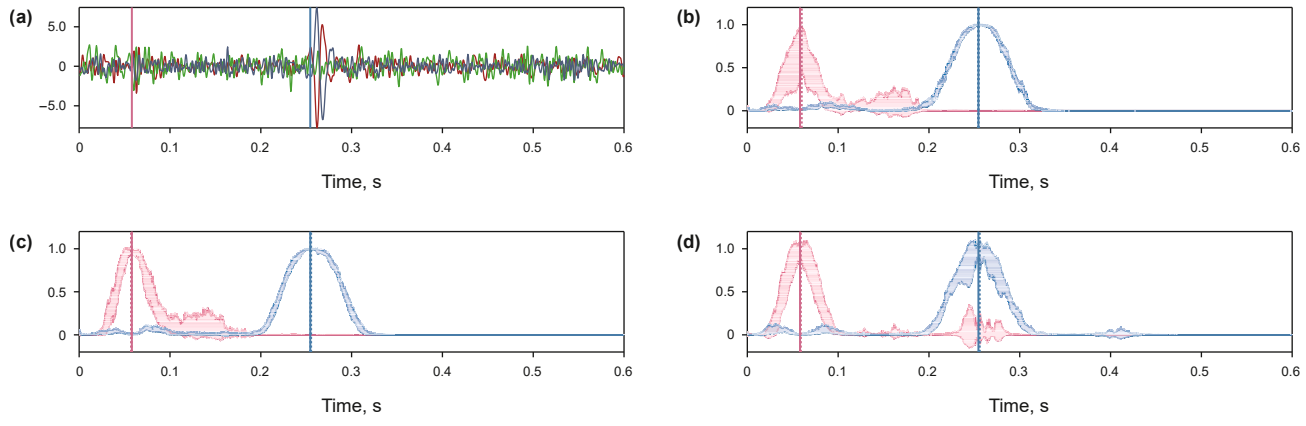


Fig. 7. Uncertainty quantification of T-MD series (double-phase example 1): **(a)** three component single-trace seismogram, **(b)** T-MD1, **(c)** T-MD2, and **(d)** T-MD3. In **(b)**, **(c)**, and **(d)**, intervals [Me–Sd, Me + Sd] between red/blue curves represent uncertainty bounds of P-/S-waves.

The detection accuracy metric is defined as the ratio of correctly detected traces to total double-phase traces of the testing dataset. The detection performance of double-phase samples is summarized in Table 3. All experimental groups exhibited detection omissions, with T-Base demonstrating the poorest performance at only 1244 correctly detected traces. The T-Dr and T-MD series show progressive accuracy improvements. With a detection accuracy of 99.00% as the threshold, the high-accuracy group (>99.00%) consists of four configurations with dropout rates of 0.2 and 0.3 (T-Dr2/3, T-MD2/3), while the low-accuracy group (<99.00%) consists of three configurations with dropout rates of 0 and 0.1 (T-Base, T-Dr1, and T-MD1). This statistical analysis demonstrates that the dropout rate serves as a critical controlling factor for model detection performance.

Fig. 12 displays phase arrival picking error distributions relative to manual benchmarks, with quantitative metrics detailed in

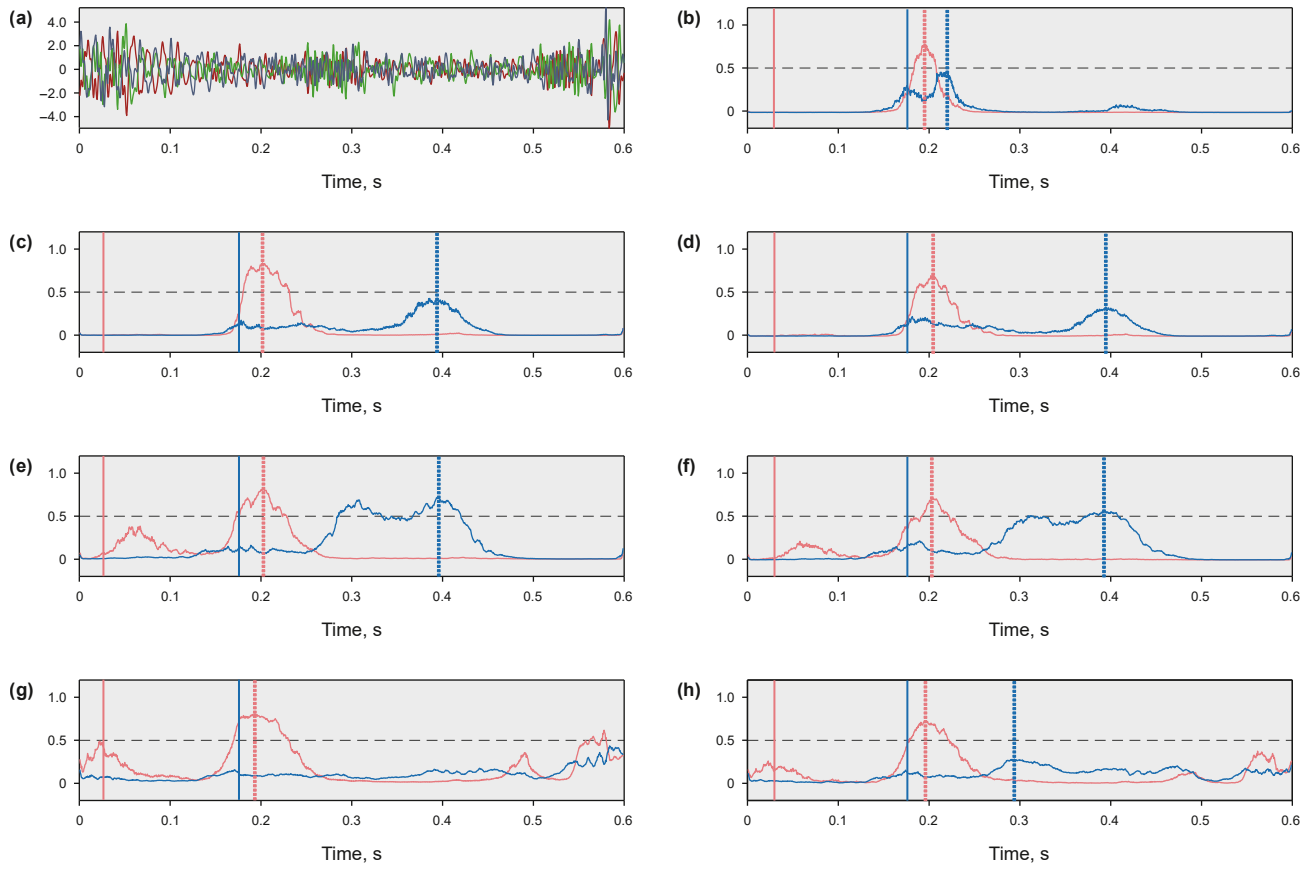


Fig. 8. Phase detection and arrival picking results of different groups (double-phase example 2): **(a)** three component single-trace seismogram, **(b)** T-Base, **(c)** T-Dr1, **(d)** T-MD1, **(e)** T-Dr2, **(f)** T-MD2, **(g)** T-Dr3, and **(h)** T-MD3.

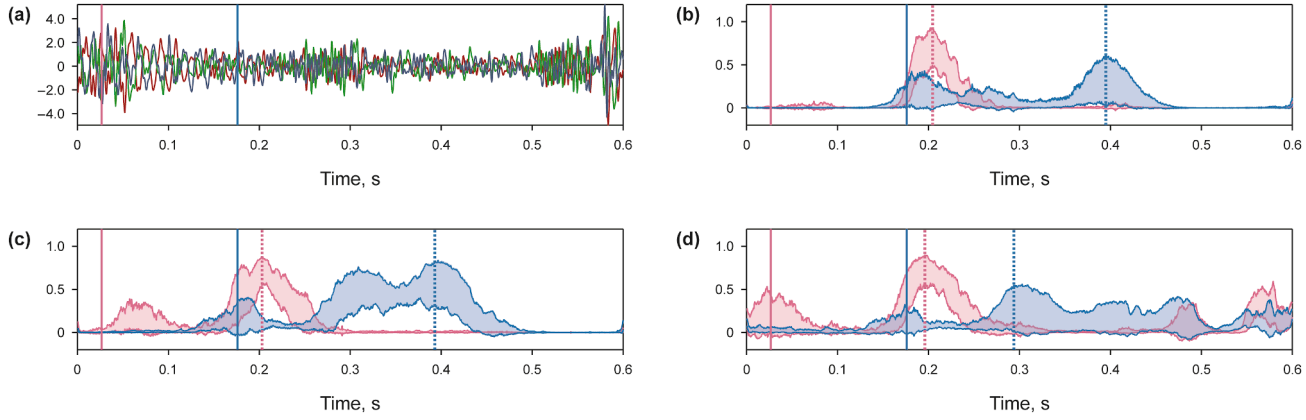


Fig. 9. Uncertainty quantification of T-MD series (double-phase example 2): **(a)** three component single-trace seismogram, **(b)** T-MD1, **(c)** T-MD2, and **(d)** T-MD3.

Table 4. All groups achieve the mean of picking errors (ME) within 0–1.7 ms for P-wave, S-wave, and P-S time difference, satisfying microseismic monitoring precision requirements. Standard deviation of picking errors (SD) analysis demonstrates significant error reduction compared to the baseline group: P-wave decreases from 14.3 ms (T-Base) to 9.1 ms (T-Dr2), S-wave from 14.3 ms to 11.7 ms (T-MD2, T-Dr3), and P-S time difference from 12.0 ms to 7.1 ms (T-MD3). Among uncertainty-aware T-MD series, T-MD1 showed degraded S-wave performance compared to the baseline group, while T-MD2/3 exhibited significant improvements. Notably, T-MD3 achieves about 40% reduction in the SD of P-S time difference versus T-Base, attaining accuracy comparable to the multi-trace cross-correlation method (Tan and He, 2016). These results confirm that the proposed MD-Net framework successfully balances double objectives: maintaining phase detection and arrival picking accuracy meeting monitoring requirements while extracting crucial uncertainty metrics. The achieved precision enhancement, particularly in critical P-S time difference measurements, ensures reliable data for subsequent processing workflow.

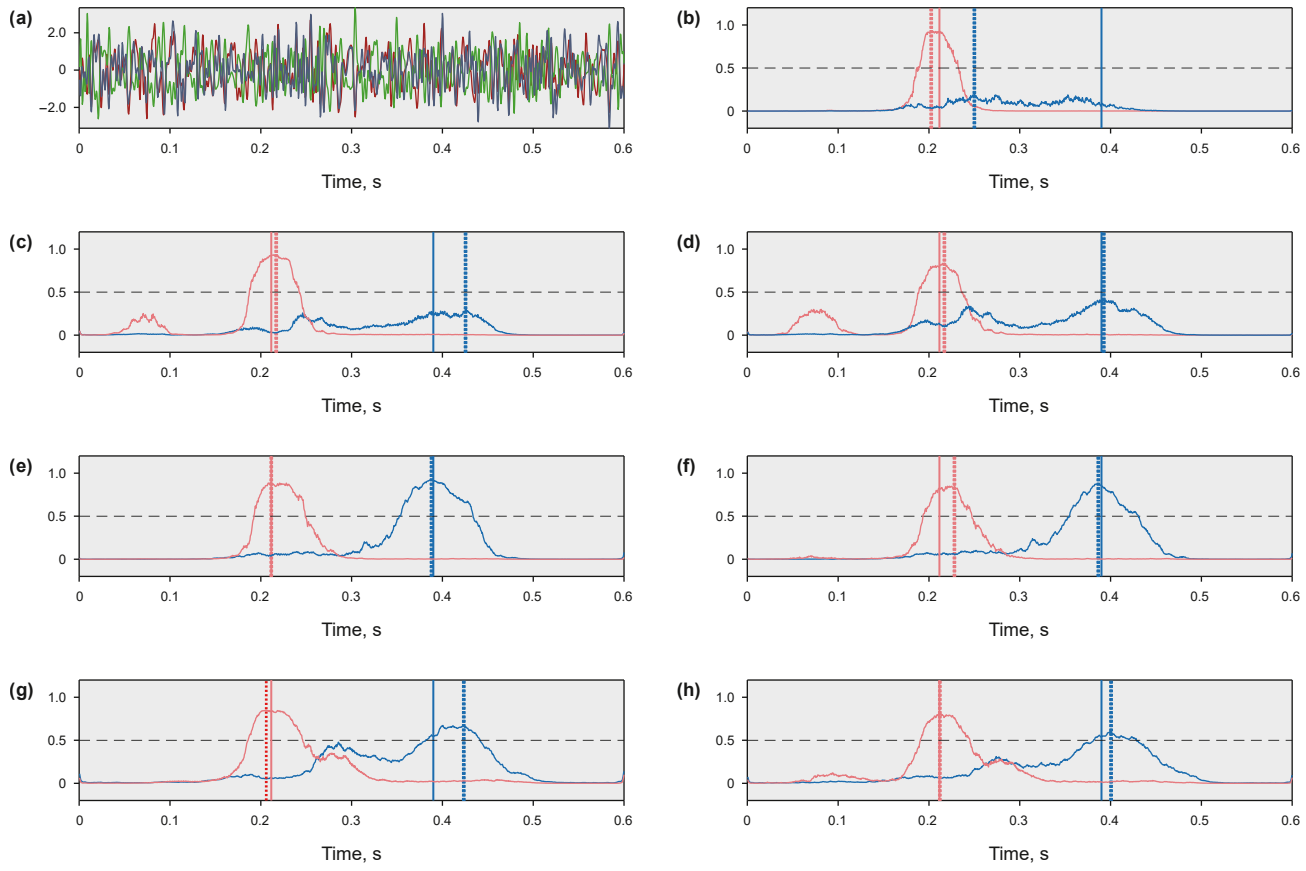


Fig. 10. Phase detection and arrival picking results of different groups (double-phase example 3): **(a)** three component single-trace seismogram, **(b)** T-Base, **(c)** T-Dr1, **(d)** T-MD1, **(e)** T-Dr2, **(f)** T-MD2, **(g)** T-Dr3, and **(h)** T-MD3.

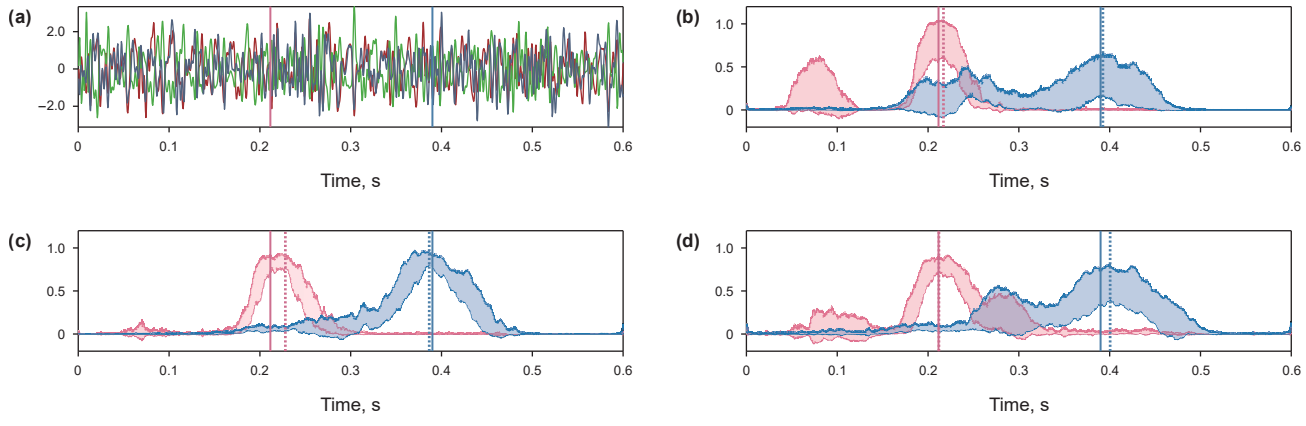


Fig. 11. Uncertainty quantification of T-MD series (double-phase example 3): **(a)** three component single-trace seismogram, **(b)** T-MD1, **(c)** T-MD2, and **(d)** T-MD3.

Table 3
Statistics on the detection results of double-phase samples.

Group	T-Base	T-Dr1	T-MD1	T-Dr2	T-MD2	T-Dr3	T-MD3
Number of correctly detected traces	1244	1259	1253	1275	1268	1270	1267
Percentage, %	97.57	98.75	98.27	100.00	99.45	99.61	99.37

3.3. Location optimization evaluation

The location optimization evaluation focuses on fracturing stages 8–10 in Fig. 4(c), containing 85 double-phase microseismic event samples. These events undergo location processing through the workflow shown in Fig. 3, combining automated data-driven phase detection and arrival picking based MD-Net and manual quality control, denoted as D-MD mode. Given the limited inter-

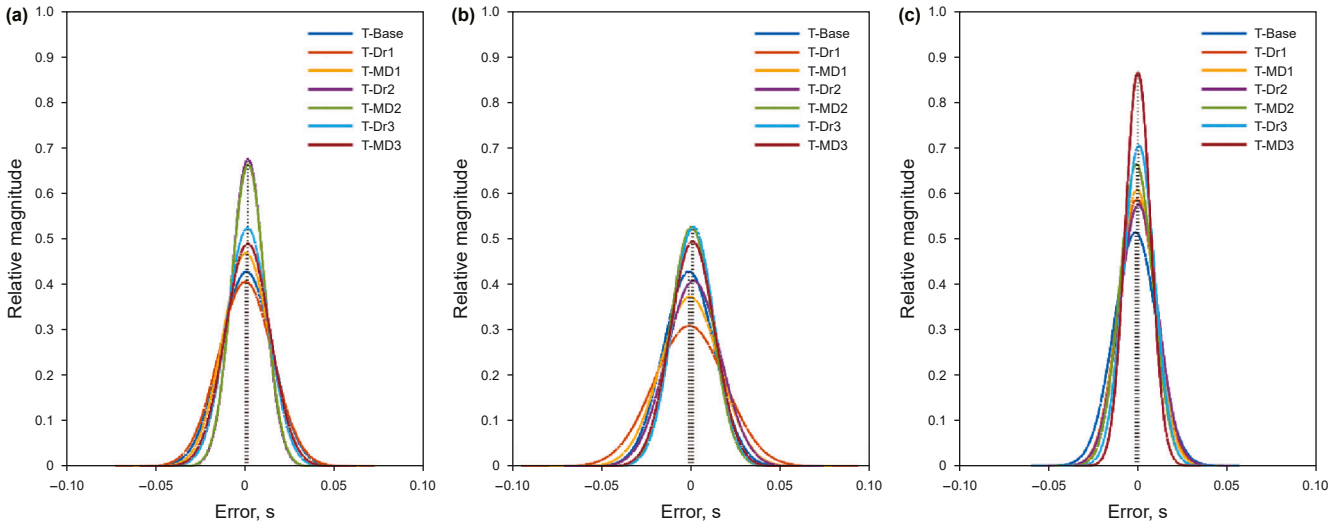


Fig. 12. Gaussian distributions relative to manual benchmarks of phase arrival picking errors of double-phase samples: **(a)** P-wave, **(b)** S-wave, and **(c)** P-S time difference.

well distances and small surface elevation variations, we implement a horizontal laminar velocity model derived from fracturing well sonic logging data, calibrated using perforation events from stages 8–10 (Fig. 13). Field implementation addresses uncertainty variability across dropout rates through parallel model operations: Three MD-Net models with dropout rates 0.1, 0.2, and 0.3 generate simultaneous predictions. Automated quality control triggers manual calibration when the uncertainty estimation of any model's probability distribution exceeds 0.2 (empirically determined threshold).

To verify the workflow performance of our proposed D-MD mode, three control groups are set up to carry out identical source location work without manual quality control after obtaining picking results from one model-driven (cross-correlation method) (Tan and He, 2016), two data-driven method (ST-Net (Wu et al., 2019) and MT-Net (Ma et al., 2020; Zhang et al., 2021, 2022), denoted as M-CC mode, D-ST mode, D-MT mode, respectively. Among them, the M-CC mode, which uses waveform similarity characteristics by calculating correlation functions among multi-trace recordings, is currently recognized as one of the best strategies to improve the picking accuracy. The D-ST mode is based on the 1D U-Net structure with single-trace waveforms as input, and extracting the waveform characteristics using multilevel feature fusion of the U-Net. The D-MT mode is still based on the 2D U-Net structure, and uses multi-trace waveforms as a whole to adaptively learn in "intra-trace" and "inter-trace" dimensions meanwhile. The rest of settings for the D-ST and D-MT modes are identical to those of the D-MD except that the MC-Dropout strategy is not used. All modes share identical velocity models and azimuthal constraints, with source locations determined through grid search inversion.

Fig. 14(a) presents the source location results obtained through the conventional M-CC mode. Spatial analysis reveals distinct clustering patterns: Microseismic events of stage 8 predominantly concentrate near wellbores, events of stage 9 show partial location adjacent to monitoring wells, while events of stage 10 demonstrate sparse distribution with limited spatial coherence. The D-ST mode (Fig. 14(b)) outputs exhibit significant location anomalies, with parts of events erroneously localized beyond target fracturing

Table 4
Statistics of phase arrival picking errors of double phase samples (mean: ME; standard deviation: SD).

Group	P-wave		S-wave		P-S time difference	
	ME, ms	SD, ms	ME, ms	SD, ms	ME, ms	SD, ms
T-Base	0.5	14.3	1.1	14.3	1.7	12.0
T-Dr1	0.1	15.1	0.8	19.9	0.5	10.4
T-MD1	0.2	13.1	0.5	16.5	0.7	10.1
T-Dr2	1.4	9.1	1.4	15.0	0	10.7
T-MD2	1.4	9.3	0.3	11.7	1.1	9.3
T-Dr3	1.1	11.7	1.4	11.7	0.3	8.7
T-MD3	1.2	12.5	1.0	12.4	0.3	7.1

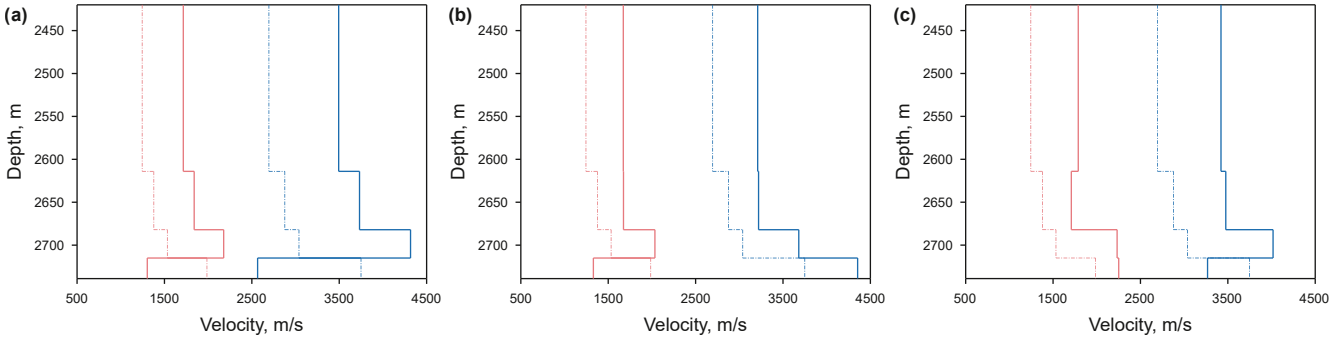


Fig. 13. Velocity model calibration using perforation events: **(a)** fracturing stage 8; **(b)** fracturing stage 9; and **(c)** fracturing stage 10. Blue and red lines denote velocity models of P- and S-wave respectively, with dashed lines representing initial models and solid lines indicating calibrated models.

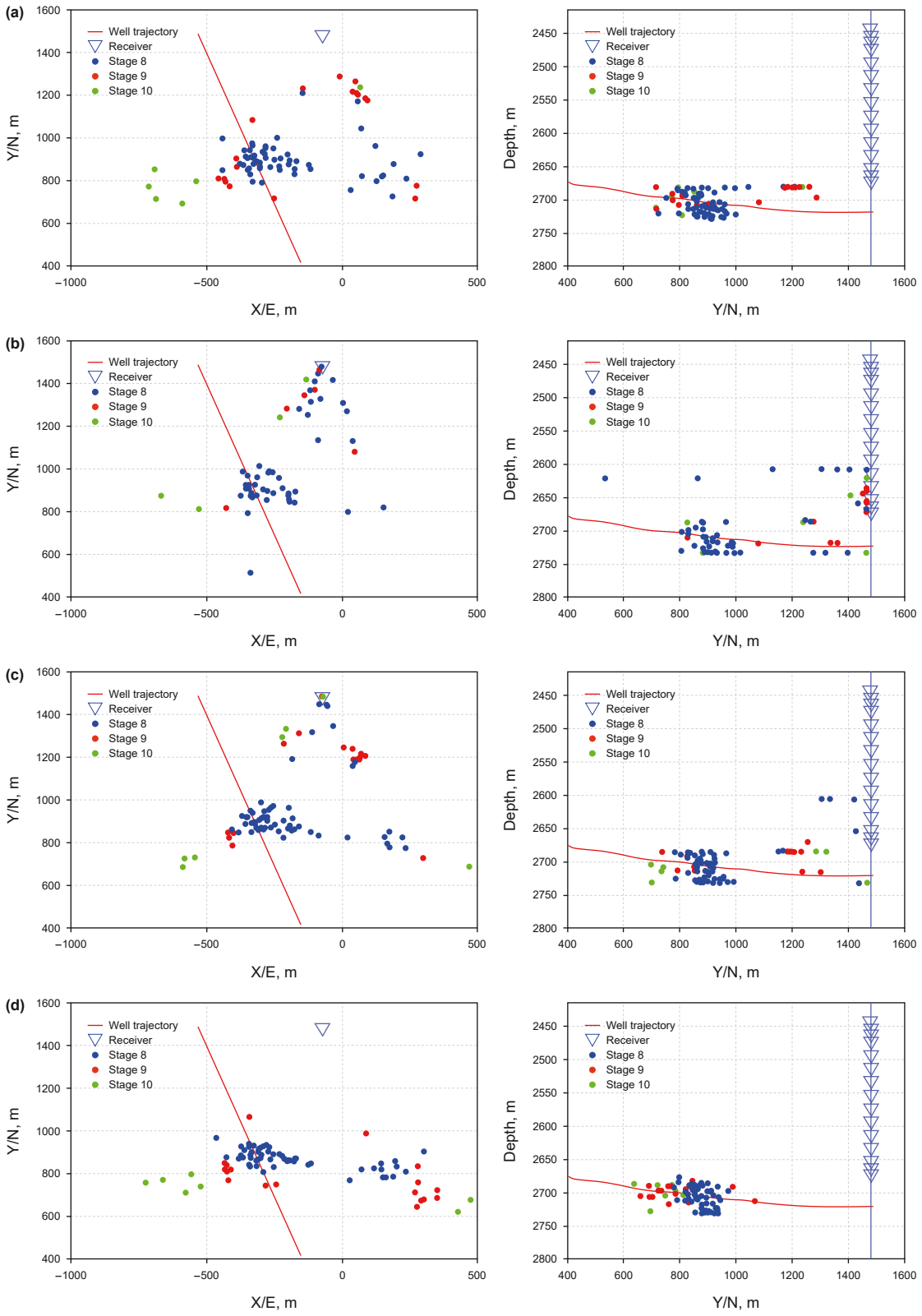


Fig. 14. Spatial distribution of microseismic sources using different modes: **(a)** M-CC, **(b)** D-ST, **(c)** D-MT, **(d)** D-MD. Hypocenters from fracturing stages 8, 9, and 10 are denoted by blue, red, and green scatter points respectively.

zones (exceeding grid search boundaries), likely attributed to errors in P-S time difference measurements of some traces of those events. Although the D-MT mode (Fig. 14(c)) demonstrates improved clustering along fracture expanding direction, minority of events remain mislocated near monitoring wellbore, resulting in loss of microseismic data for fracture interpretation. The

proposed D-MD mode (Fig. 14(d)) achieves optimal performance with all of events concentrated along the primary fracture expanding direction near horizontal wellbore. Compared to other modes, D-MD mode avoids location errors while maintaining geomechanically consistent spatial distributions. This precision enhancement confirms the effectiveness of our proposed workflow in meeting practical requirements for real-time fracturing interpretation of HF scenarios.

4. Discussion

The integration of Monte Carlo Dropout into the MD-Net architecture introduces a paradigm shift in uncertainty quantification for seismic phase detection and arrival picking. By approximating Bayesian inference through stochastic forward sampling, the model not only suppresses noise-induced picking errors but also generates confidence intervals that guide manual calibration. The D-MD series maintains computational efficiency comparable to other experimental groups during training, with only a marginal increase in prediction cost due to multiple sampling. For field microseismic monitoring, this added latency (~0.1 s/event) is negligible compared to operational timescales and does not hinder practical deployment. The framework thus balances uncertainty quantification with minimal computational overhead in field applications.

Furthermore, the hybrid workflow underscores the synergy between data-driven automation and physics-based constraints. Unlike pure data-driven approaches, which may produce geologically implausible outliers, the incorporation of traveltime-based inversion ensures adherence to wave propagation principles, bridging the gap between black-box predictions and interpretable geophysical models. The current study is designed for extensibility to diverse lithologies and noise profiles. Zhang et al. (2022) demonstrated that transfer learning enables effective domain adaptation across distinct fracturing projects and geological regions. By integrating similar TL strategies, MD-Net can be rapidly fine-tuned for new scenarios using minimal target data, significantly reducing deployment barriers.

Despite its strengths, the framework exhibits three primary limitations:

(1) The dropout rate significantly influences uncertainty estimation robustness. Excessively high rates may oversmooth phase characteristic, while low rates inadequately capture epistemic uncertainty. Future implementations could integrate Bayesian optimization to dynamically adjust dropout rates based on real-time SNR assessments.
(2) The current workflow assumes laminar velocity models calibrated via perforation events. In complex reservoirs, this simplification may introduce location biases. Coupling MD-Net integrating distributed acoustic sensing (DAS) data could address such limitations (Tegtow et al., 2025).
(3) The microseismic events processed in field case are events with clear phase arrivals, and the picking error is small. Therefore, the location results using traditional grid search algorithm can basically meet the processing requirements, and other better inversion algorithms should be considered under the condition of low SNR in other monitoring.

Beyond HF monitoring, the hybrid-driven paradigm holds promise for diverse geophysical applications, and further validation under varying conditions remains necessary for comprehensive evaluation of its generalizability. The uncertainty-aware architecture could be adapted for earthquake early warning systems, where false alarms carry significant societal consequences. Similarly, the workflow's emphasis on manual-calibration compatibility provides a blueprint for integrating AI into legacy processing chains without disrupting established practices. As the geosciences increasingly embrace data-driven methodologies, this study exemplifies how domain knowledge and machine learning can coexist to achieve both innovation and reliability.

5. Conclusion

This study proposes a hybrid-driven framework integrating uncertainty-aware neural networks with manual quality control for full workflow processing of downhole PMM data in HF. The main conclusions are summarized as follows.

(1) Automated phase detection/arrival picking and uncertainty quantification: By incorporating Monte Carlo Dropout strategy into a modified 2D U-Net architecture, the proposed MD-Net achieves multi-task collaborative optimization of phase detection, arrival picking, and uncertainty estimation. Field data validation demonstrates the model's superior performance under low SNR conditions, with detection accuracy exceeding 99% for both P- and S-waves. Compared with the baseline model, the standard deviation of phase arrival picking errors is reduced by 40% (P-S arrival time difference decreasing from 12.0 ms to 7.1 ms), while providing quantifiable uncertainty metrics to support manual calibration decisions. This breakthrough overcomes the bottleneck of traditional data-driven paradigm in uncertainty characterization and manual quality control, establishing a new technical paradigm for microseismic signal processing in complex noise environments.
(2) Workflow optimization for source location: A hybrid workflow balancing automation efficiency and physical interpretability is developed through deep integration of data-driven phase detection/arrival picking and model-driven traveltime-based inversion methods. Field data validation reveals that 100% of microseismic events exhibit spatial clustering along the primary fracture expanding direction, showing significantly better spatial rationality compared with traditional cross-correlation and single/multi-trace data-driven methods. This collaboration mechanism not only avoids the black-box limitations of purely data-driven paradigm in velocity modeling and nonlinear inversion, but also compensates for the characteristic generalization deficiencies of traditional model-driven paradigm, providing a high-precision scheme for fracturing interpretation.
(3) Practical applicability and adaptive optimization: Through progressive dataset enhancement and transfer learning strategies, the model achieves parameter self-adaptation across multiple fracturing stages of one HF project, significantly reducing manual intervention requirements.

This study establishes an interpretable, high-precision framework for HF-PMM, demonstrating significant potential for industrial applications in fracture network characterization and hydraulic fracturing optimization.

CRediT authorship contribution statement

Yi-Lun Zhang: Writing – original draft, Software, Methodology, Visualization, Resources, Investigation, Validation, Project administration, Conceptualization. **Zhi-Chao Yu:** Supervision, Formal analysis, Writing – review & editing, Data curation. **Chuan He:** Writing – review & editing, Formal analysis, Supervision, Data curation, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Chuan He reports financial support was provided by DeepEarth Probe and Mineral Resources Exploration - National Science and Technology Major Project (Grant No. 2024ZD1002503). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the DeepEarth Probe and Mineral Resources Exploration - National Science and Technology Major Project (Grant No. 2024ZD1002503).

References

Akram, J., Eaton, D.W., 2016. A review and appraisal of arrival-time picking methods for downhole microseismic data. Geophysics 81 (2), KS71–KS91. https://doi.org/10.1190/geo2014-0500.1.

Anikiev, D., Birnie, C., Waheed, U., et al., 2023. Machine learning in microseismic monitoring. Earth Sci. Rev. 239. https://doi.org/10.1016/j.earscirev.2023.104371.

Bao, X., Eaton, D.W., 2016. Fault activation by hydraulic fracturing in Western Canada. Science 354. https://doi.org/10.1126/science.aag2583.

Bergen, K.J., Beroza, G.C., 2019. Earthquake fingerprints: extracting waveform features for similarity-based earthquake detection. Pure Appl. Geophys. 176, 1037–1059. https://doi.org/10.1007/s00024-018-1995-6.

Bergen, K.J., Johnson, P.A., de Hoop, M.V., et al., 2019. Machine learning for data-driven discovery in solid Earth geoscience. Science 363. https://doi.org/10.1126/science.aau0323.

Chen, Y., Saad, M., Savvaidis, O., Alexandros, et al., 2024. Deep learning for P-wave first-motion polarity determination and its application in focal mechanism inversion. IEEE Trans. Geosci. Rem. Sens. 62. https://doi.org/10.1109/TGRS.2024.3407060.

Gal, Y., Ghahramani, Z., 2016. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In: International Conference on Machine Learning. PMLR, pp. 1050–1059.

Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. JMLR Workshop and Conference Proceedings, pp. 315–323.

Hu, W., Jin, Y., Wu, X., et al., 2021. Progressive transfer learning for low frequency data prediction in full waveform inversion. Geophysics 86 (4), R369–R382. https://doi.org/10.1190/geo2020-0598.1.

Jones, G.A., Kendall, J.-M., Bastow, I.D., et al., 2014. Locating microseismic events using borehole data. Geophys. Prospect. 62 (1), 34–49. https://doi.org/10.1111/1365-2478.12076.

Kingma, D.P., Ba, J.L., 2015. Adam: a method for stochastic optimization. IEEE, San Diego, USA, pp. 1–15.

Kiureghian, A.D., Ditlevsen, O., 2009. Aleatory or epistemic? Does it matter? Struct. Saf. 31 (2), 105–112. https://doi.org/10.1016/j.strusafe.2008.06.020.

Kong, Q., Trugman, D.T., Ross, Z.E., et al., 2019. Machine learning in seismology: turning data into insights. Seismol Res. Lett. 90 (1), 3–14. https://doi.org/10.1785/0220180259.

Kuang, W., Yuan, C., Zhang, J., 2021. Real-time determination of earthquake focal mechanism via deep learning. Nat. Commun. 12 (1). https://doi.org/10.1038/s41467-021-21670-x.

Lan, X., Zou, C., Peng, C., et al., 2023. Uncertainty quantification in intelligent-based electrical resistivity tomography image reconstruction with Monte Carlo dropout strategy. IEEE Trans. Geosci. Rem. Sens. 61, 1–16. https://doi.org/10.1109/TGRS.2023.3262835.

Li, L., Tan, J., Schwarz, B., et al., 2020. Recent advances and challenges of waveform-based seismic location methods at multiple scales. Rev. Geophys. 58 (1). https://doi.org/10.1029/2019RG000667.

Li, L., Tan, J., Wood, D.A., et al., 2019. A review of the current status of induced seismicity monitoring for hydraulic fracturing in unconventional tight oil and gas reservoirs. Fuel 242, 195–210. https://doi.org/10.1016/j.fuel.2019.01.026.

Li, Q., Xing, H., Liu, J., et al., 2015. A review on hydraulic fracturing of unconventional reservoir. Petroleum 1 (1), 8–15. https://doi.org/10.1016/j.petlm.2015.03.008.

López, V., Fernández, A., García, S., et al., 2013. An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics. Inf. Sci. 250, 113–141. https://doi.org/10.1016/j.ins.2013.07.007.

Ma, Y., Cao, S., Rector, J.W., et al., 2020. Automated arrival-time picking using a pixel-level network. Geophysics 85 (5), V415–V423. https://doi.org/10.1190/geo2019-0792.1.

Maxwell, S.C., Urbancic, T.I., 2001. The role of passive microseismic monitoring in the instrumented oil field. Lead. Edge 20 (6), 636–639. https://doi.org/10.1190/1.1439012.

McGlade, C., Speirs, J., Sorrell, S., 2013. Unconventional gas–a review of regional and global resource estimates. Energy 55, 571–584. https://doi.org/10.1016/j.energy.2013.01.048.

Meng, X., Chen, H., Niu, F., et al., 2018. Microseismic monitoring of stimulating shale gas reservoir in SW China: 1. an improved matching and locating technique for downhole monitoring. J. Geophys. Res. Solid Earth 123 (2), 1643–1658. https://doi.org/10.1002/2017JB014488.

Mousavi, S.M., Beroza, G.C., 2020. Bayesian-deep-learning estimation of earthquake location from single-station observations. IEEE Trans. Geosci. Rem. Sens. 58 (11), 8211–8224. https://doi.org/10.1109/TGRS.2020.2988770.

Mousavi, S.M., Ellsworth, W.L., Zhu, W., et al., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nat. Commun. 11 (1). https://doi.org/10.1038/s41467-020-17591-w.

Mousavi, S.M., Langston, C.A., 2017. Automatic noise-removal/signal-removal based on general cross-validation thresholding in synchrosqueezed domain and its application on earthquake data. Geophysics 82 (4), V211–V227. https://doi.org/10.1190/geo2016-0433.1.

Navarro-Rodríguez, A., Castro-Artola, O.A., García-Guerrero, E.E., et al., 2025. Recent advances in early earthquake magnitude estimation by using machine learning algorithms: a systematic review. Applied Sciences 15 (7). https://doi.org/10.3390/app15073492.

Perol, T., Gharbi, M., Denolle, M., 2018. Convolutional neural network for earthquake detection and location. Sci. Adv. 4 (2). https://doi.org/10.1126/sciadv.1700578.

Pesicek, J.D., Child, D., Artman, B., et al., 2014. Picking versus stacking in a modern microearthquake location: comparison of results from a surface passive seismic monitoring array in Oklahoma. Geophysics 79 (6), KS61–KS68. https://doi.org/10.1190/geo2013-0404.1.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and computer-assisted intervention–MICCAI 2015: 18Th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18. Springer international publishing, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

Ross, Z.E., Meier, M.A., Hauksson, E., et al., 2018. Generalized seismic phase detection with deep learning. Bull. Seismol. Soc. Am. 108 (5A), 2894–2901. https://doi.org/10.1785/0120180080.

Shang, G.J., Huang, W.L., Yuan, L.K., et al., 2022. Automatic microseismic events detection using morphological multiscale top-hat transformation. Pet. Sci. 19 (5), 2027–2045. https://doi.org/10.1016/j.petsci.2022.08.005.

Srivastava, N., Hinton, G., Krizhevsky, A., et al., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15 (1), 1929–1958.

Sun, J., Innanen, K.A., Huang, C., 2021. Physics-guided deep learning for seismic inversion with hybrid training and uncertainty analysis. Geophysics 86 (3), R303–R317. https://doi.org/10.1190/geo2020-0312.1.

Sun, Q.H., Zong, Z.Y., Li, X., 2024. Probabilistic seismic inversion based on physics-guided deep mixture density network. Pet. Sci. 21 (3), 1611–1631. https://doi.org/10.1016/j.petsci.2023.12.015.

Tan, Y., He, C., 2016. Improved methods for detection and arrival picking of microseismic events with low signal-to-noise ratios. Geophysics 81 (2), KS93–KS111. https://doi.org/10.1190/geo2015-0213.1.

Tan, Y.Y., Li, L.L., Zhang, X., et al., 2017. An improved method for microseismic source location based on grid search. Chin. J. Geophys. 60 (1), 293–304. https://doi.org/10.6038/cjg20170124 (in Chinese).

Tegtow, W., Boitz, N., Shapiro, S.A., 2025. Integrating microseismicity and low frequency DAS to characterize hydraulic fracturing. Geophysics 90 (3), 1–37. https://doi.org/10.1190/geo2024-0265.1.

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. https://doi.org/10.1016/j.cageo.2021.104965.

Wu, H., Zhang, B., Li, F., et al., 2019. Semiautomatic first-arrival picking of microseismic events by using the pixel-wise convolutional image segmentation method. Geophysics 84 (4), V143–V155. https://doi.org/10.1190/geo2018-0389.1.

Wu, S., Huang, Q., Zhao, L., 2023. Fast Bayesian inversion of airborne electromagnetic data based on the invertible neural network. IEEE Trans. Geosci. Rem. Sens. 61, 1–11. https://doi.org/10.1109/TGRS.2023.3264777.

Wu, S., Huang, Q., Zhao, L., 2024. Physics-guided deep learning-based inversion for airborne electromagnetic data. Geophys. J. Int. 238 (3), 1774–1789. https://doi.org/10.1093/gji/ggae244.

Yu, S., Ma, J., 2021. Deep learning for geophysics: current and future trends. Rev. Geophys. 59 (3). https://doi.org/10.1029/2021RG000742.

Zhang, X., Zhang, J., Yuan, C., et al., 2020. Locating induced earthquakes with a network of seismic stations in Oklahoma via a deep learning method. Sci. Rep. 10 (1). https://doi.org/10.1038/s41598-020-58908-5.

Zhang, Y., Leng, J., Dong, Y., et al., 2022. Phase arrival picking for bridging mulita-source downhole microseismic data using deep transfer learning. J. Geophys. Eng. 19 (2), 178–191. https://doi.org/10.1093/jge/gxac009.

Zhang, Y.L., Yu, Z.C., Hu, T.Y., et al., 2021. Multi-trace joint downhole microseismic phase detection and arrival picking method based on U-Net. Chin. J. Geophys. 64 (6), 2073–2085. https://doi.org/10.6038/cjg2021O0379 (in Chinese).

Zheng, M., Li, J., Wu, X., et al., 2018. China's conventional and unconventional natural gas resources: potential and exploration targets. Journal of Natural Gas Geoscience 3 (6), 295–309. https://doi.org/10.1016/j.jnggs.2018.11.007.
Zhou, Y., Ding, H., Ghost, A., et al., 2025. AI-PAL: self-supervised AI phase picking via rule-based algorithm for generalized earthquake detection. J. Geophys. Res. Solid Earth 130 (4). https://doi.org/10.1029/2025JB031294.
Zhou, Y., Yue, H., Kong, Q., et al., 2019. Hybrid event detection and phase-picking algorithm using convolutional and recurrent neural networks. Seismol Res. Lett. 90 (3), 1079–1087. https://doi.org/10.1785/0220180319.
Zhu, W., Beroza, G.C., 2019. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. Geophys. J. Int. 216 (1), 261–273. https://doi.org/10.1093/gji/ggy423.