

Depth estimation of weak narrowband sources using a hybrid interference-deep learning method

WANG Tong SU Lin[†] LIU Pengchao REN Qunyan GUO Xinyi HU Tao MA Li

Key Laboratory of Underwater Acoustic Environment, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China

Received December 1, 2025; accepted December 24, 2025

Abstract Interference-structure-based passive source localization provides computational efficiency and physical interpretability but suffers performance degradation for weak acoustic targets, where low signal-to-noise ratios (SNRs) obscure the observable interference patterns necessary for reliable depth estimation. Narrowband time-window limitations and environmental mismatches further introduce blind zones and nonlinear distortions between the source depth and the measured interference structure. To mitigate these effects, this study proposes a hybrid depth estimation method that integrates arrival-angle interference features with a deep learning (DL) framework based on a residual network. Sound intensity in the beam domain is employed to exploit multidimensional and nonlinear relationships in the acoustic field, thereby enhancing robustness under weak target and low SNR conditions. Simulation and sea experiment results demonstrate that the proposed method achieves improved depth estimation accuracy compared with conventional interference-based techniques. The findings indicate that the hybrid interference-DL approach effectively extends the applicability of interference-structure-based localization to weak acoustic targets.

Keywords Source depth estimation, Deep sea, Deep learning, Narrowband source

PACS: 43.30, 43.60

DOI: [10.3724/0217-9776.3025004](https://doi.org/10.3724/0217-9776.3025004)

CSTR: [32050.14.11-2066.3025004](https://cstr.cn/32050.14.11-2066.3025004)

1 Introduction

Passive underwater source localization methods can be broadly classified into three categories: matched field processing (MFP)^[1], multipath arriving structure-based methods^[2], and sound field interference structure-based methods^[3-9]. Traditional MFP approaches rely heavily on accurate environmental information, making them sensitive to parameter mismatches. In contrast, interference-structure-based methods provide rich source-related information with better computational efficiency and robustness, attracting considerable attention. However, for weak acoustic targets under low signal-to-noise ratio (SNR) conditions, the observable interference features often degrade, making accurate localization particularly challenging.

Previous studies^[3-6] have investigated source depth estimation using interference patterns such as Lloyd's mirror, exploiting the oscillatory characteristics of signal spectra. These methods depend on the interference between the direct (D) and the surface-reflected (SR) arrivals. For narrowband sources, the limited time window restricts observable arrival-

angle variations, resulting in blind zones in depth estimation. Moreover, discrepancies between real and assumed environments, particularly in the sound speed profile (SSP), induce nonlinear distortions between source depth and interference structures. Under weak target conditions, interference fringes become blurred or submerged in noise, further degrading performance.

To alleviate these limitations, this paper introduces a deep learning (DL)-assisted depth estimation method based on sound intensity in the beam domain. Although DL cannot fundamentally overcome low-SNR limitations, it can leverage multidimensional acoustic features and capture complex nonlinear relationships beyond the reach of traditional approaches. Benefiting from its success in various ocean acoustic applications such as source localization^[10-14], geoaoustic inversion^[15,16], and target classification^[17-20], DL methods can partially enhance the robustness and stability of interference-based localization in weak target conditions.

Building on these advantages, this study proposes a hybrid depth estimation method that combines the arrival-angle interference structure with a residual network (ResNet)-

This work was supported by the Beijing Nova Program.

[†] Corresponding author: SU Lin, sulin807@mail.ioa.ac.cn

based DL model. Simulation and sea experiments demonstrate that the proposed approach significantly improves weak-target depth estimation, effectively mitigating the performance degradation observed in traditional interference-based methods.

2 Performance analysis in conventional methods

2.1 Matched field processing

The MFP estimates the source position by computing the correlation between the replica field and the received acoustic field. The spatial coherence of the replica field causes constructive fluctuations at specific locations. Due to strong vertical coherence in the near-field region, the MFP applied with a vertical line array (VLA) exhibits a blind zone in close-range localization, leading to significant depth estimation errors at short distances. Furthermore, under the reliable acoustic path (RAP) condition, as the source–receiver distance increases, the influence of seabed reflections on the acoustic field becomes increasingly significant. When the distance exceeds a certain threshold, the effects of bottom reflections can no longer be neglected, contributing to increased depth estimation errors. These inaccuracies in the bottom-reflection zone are primarily due to environmental mismatches, particularly in SSP and bottom parameters.

Using Hamilton's nine typical sediment types^[20] in combination with the Munk sound speed profile, Fig. 1 illustrates the amplitude ratio between the direct and boundary-reflected (BR) arrivals at a frequency of 200 Hz. The observed trend shows an initial increase in this ratio with distance, followed by a decline, highlighting the non-linear impact of seabed reflections on signal amplitude.

2.2 Multi-Fourier transform depth estimation

For narrowband moving sources, the interference between the direct and SR paths gives rise to a periodic structure as a function of the vertical arrival angle within the direct-sound region. This angular oscillation pattern contains essential in-

formation about the source depth. The multi-Fourier transform depth estimation (MSTDE) method^[6,7] exploits this mechanism by mapping the periodic variations of sound intensity into the depth domain through multiple Fourier transformations. The theoretical principles of the interference structure are elaborated in Ref. [4], and the corresponding key equation is expressed as

$$M(z, \omega) = \sum_{l=1}^L P(\omega, \sin\theta_{tr}(t_l)) e^{-i2kz\sin\theta_{tr}(t_l)} \left| \Delta\sin\theta_{tr}(t_l) \right|, \quad (1)$$

where k is wavenumber, $\theta_{tr}(t_l)$ is the grazing angle varying with the discrete snapshot time t_l , P is the beamformed signal power as a function of frequency ω and grazing angle, and M is the two-dimensional output in the depth z –frequency ω domain. Then, the oscillation period for narrowband sources can be expressed as $T_{\sin\theta_0} = 2fz_s / c$, where the source must move for a sufficiently long duration to generate an observable variation in the vertical arrival angle. Because the interference structure relies on the angular evolution of the received signal, the limited effective time window imposes a fundamental constraint on the resolvable oscillation period. Consequently, a critical depth exists beyond which the angular interference pattern becomes indistinguishable, resulting in a blind zone in the depth estimation.

2.2.1 Time-window condition

Based on the oscillation period $T_{\sin\theta_0}$ given by $T_{\sin\theta_0} = 2fz_s / c \leq 1 / \Delta\sin\theta_0$, where θ is the grazing angle, f is the source frequency, z_s is the source depth, and c is the reference sound speed, the critical depths that can be accurately estimated at different frequencies of 10, 25, 50, 100, 160, and 200 Hz are shown in Fig. 2. It can be observed that, for the same source depth, decreasing the center frequency of the narrowband source increases the required $\Delta\sin\theta_0$. The corresponding critical depths for these center frequencies are 76, 30, 16, 8, and 4 m, respectively. Depths shallower than the critical depth are identified as the blind zone.

The value of $\Delta\sin\theta_0$ is governed by the initial source range, source speed, and the time window used in the algo-

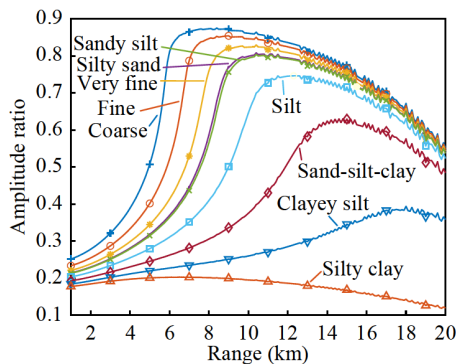


Fig. 1 Amplitude ratio between the direct and the BR arrivals under Hamilton's nine typical sediment layers.

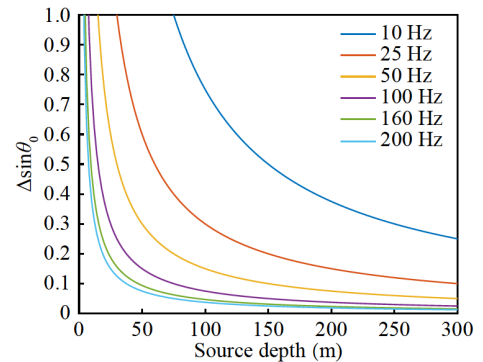


Fig. 2 $\Delta\sin\theta_0$ required for accurate estimation at various source depths and frequencies.

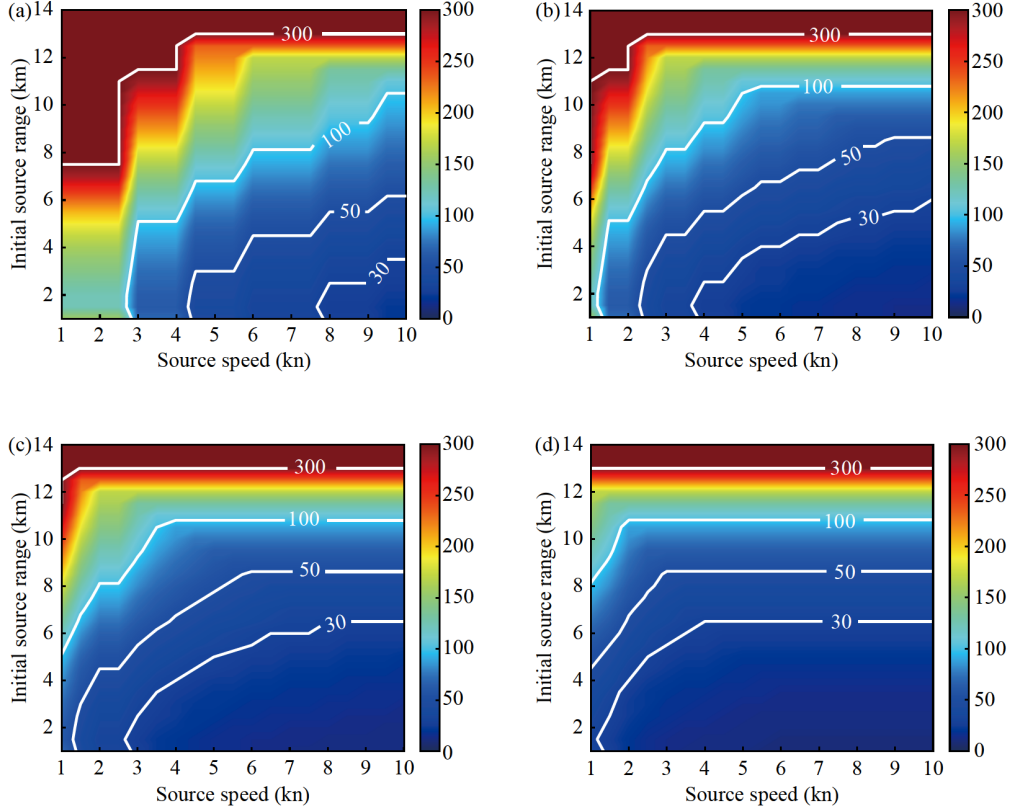


Fig. 3 The critical depths for a 100 Hz narrowband source at different time windows: (a) 10 min; (b) 20 min; (c) 30 min; (d) 40 min.

rithm. When the center frequency of the narrowband source is fixed at 100 Hz, Fig. 3 presents the critical depths obtained for different time windows. Therefore, to balance real-time performance and practical applicability, the time window is set to 30 min.

2.2.2 Sound speed profile

The mismatch between the iso-velocity SSP and the real SSP leads to variations in the mean grazing angle and the phase difference between the direct and the SR arrivals. These disparities magnify with an increase in source range,

leading to a progressively rising relative error in depth estimation when employing the MSTDE method. As shown in Fig. 4, when the grazing angle error $\Delta\theta$ is substantial, the relative error (ζ_r) exhibits a growing trend as the range extends, which can be expressed as Eq. (2). As depicted in Fig. 5, the relative error (ζ_r) exhibits a non-linear trend concerning the source range.

$$\zeta_r = \frac{dz_s}{z_s} = \frac{\frac{ccos\theta}{2f\sin^2\theta}d\theta}{\frac{c}{2f\sin\theta}} = \cot\theta d\theta. \quad (2)$$

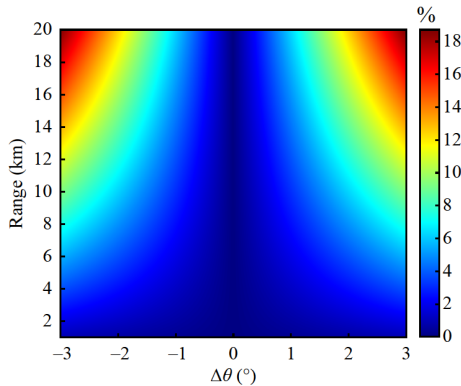


Fig. 4 The relative error as a function of the range and $\Delta\theta$ when z_s is 100 m.

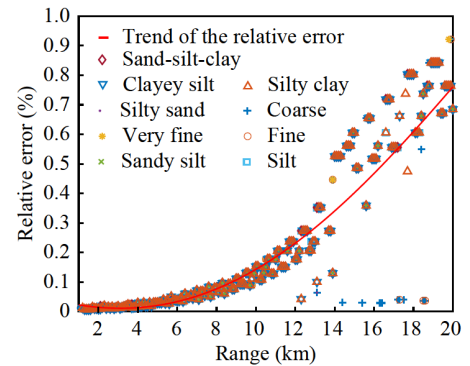


Fig. 5 The relative error when z_s is 100 m under Munk SSP using MSTDE.

3 The hybrid source localization method based on deep learning

Conventional localization methods are limited in detection range by vertical correlation fluctuations and bandwidth or time-window constraints. Their accuracy is also influenced by seabed reflections, variations in the SSP, and grazing angle errors. To address these issues, a hybrid source localization method based on an optimized residual network (ORN) is proposed. Using sound intensity in the time-beam domain as input, the method accounts for seabed and SSP effects and mitigates the performance degradation of conventional techniques, particularly for long-range and weak-target scenarios.

3.1 Input and label

The hybrid method adopts the beam-time domain sound intensity $\mathbf{I}_{\text{cbf}}$ as a model input. Figure 6 shows the simulated sound intensity field of an 89 Hz narrowband source at a depth of 50 m. A conventional beamformer (CBF) is applied using a full-aperture array with a steering resolution of 1° , covering the negative vertical arrival angle region. To ensure consistency in amplitude scale across time blocks, amplitude normalization is applied to the beamformed signals. Before inputting the matrix, normalization and a threshold of -30 dB are set to focus on the interference structure. The input sound intensity matrix $\mathbf{I}_{\text{cbf}}$ size is $\theta \times T$, where θ is the number of beamforming angles, and T is the number of time bins. As discussed in Sec. 2, T is set to 30 min. Instead of using a single scalar value, the ground truth source depth is encoded into a Gaussian-distributed vector $\mathbf{L}_{\text{Gauss}}$, centered at the true depth, to serve as the training label:

$$\mathbf{L}_{\text{Gauss}} = \left[\exp \frac{(\Delta d - d_{\text{truth}})^2}{2\sigma^2}, \exp \frac{(2\Delta d - d_{\text{truth}})^2}{2\sigma^2}, \dots, \exp \frac{(N\Delta d - d_{\text{truth}})^2}{2\sigma^2} \right]. \quad (3)$$

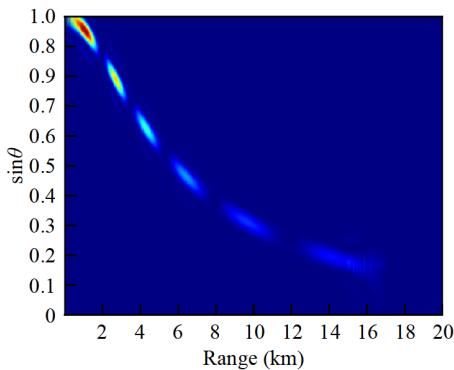


Fig. 6 The simulated sound intensity in the beam-range domain at 50 m, 89 Hz.

As proposed in Ref. [6], the Gaussian-shaped label provides a probabilistic representation that enhances robustness and stability in training by accommodating minor depth estimation errors. In Eq. (3), the discretized depth vector corresponds to a detection range of 2–200 m with a step size of $\Delta d = 2$ (i.e., $N = 100$). Here, d_{truth} represents the true depth, and σ is the standard deviation of the normal distribution, which represents the tolerance of the model to depth estimation errors. A higher value of σ indicates greater model robustness, yet it may lead to reduced accuracy, which will be fine-tuned during model construction and debugging based on performance.

3.2 Optimized residual network

ResNets are widely used in various fields, particularly in image classification tasks due to their excellent performance. They address the issue of gradient vanishing or exploding during the training of deep networks by introducing skip connections between layers, which reduces the difficulty of training deep networks. To address the source depth estimation problem, we adopt ResNet50 as the backbone model with moderate depth and strong performance among convolutional neural networks and make alterations based on it.

In order to tailor ResNet50 for our specific task, certain modifications are implemented. Firstly, we adapted the convolution block conv1 to the input data size and removed the bottom classifier, retaining only the feature extraction part of the model. Secondly, we added a global average pooling layer and a flatten layer to transform the extracted features into a one-dimensional vector. To prevent overfitting, we introduced a dropout layer with a rate of 0.5, randomly selecting neurons. We then added three fully connected layers containing 1024, 512, and 256 neurons, respectively, with the rectified linear unit (ReLU) activation function applied to the first two layers. Finally, the output was connected to a softmax layer to output the probability values of each sound source depth. The improved network structure is depicted in Fig. 7. The ORN can perform deep prediction tasks, and the related training parameter settings and training process will be described in the next subsection.

3.3 Model training

The primary factors that influence the interference

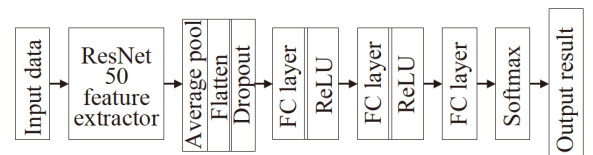


Fig. 7 The optimized ResNet50 network.

structures should be considered since they are handled as an input to our network. The SNR is taken into account while creating the datasets. The SNR of the received signal is defined as $R_{\text{SN}} = 10\log_{10}(I_{\text{po}}/I_{\text{no}})$, where I_{po} and I_{no} are the sound intensity and noise intensity at the reference element.

The simulation settings are carefully matched to the corresponding experimental environments to ensure consistency and improve generalizability. For each experiment, the primary simulation scenario is defined using the measured SSP collected via conductivity-temperature-depth system (CTD) during the sea trials. To reflect environmental variability, nine sub-scenarios are created by combining the measured SSP with nine different sediment types, as summarized in Table 1. Each sub-scenario generates 19100 samples at 0 dB SNR, resulting in a total of 171900 training samples (100 distances $\times$ 191 depths $\times$ 9 sediment types), forming Training Set 1. In contrast, Training Set 2 uses the same SSP but only one randomly selected sediment type, producing 19100 samples in total. For validation, the model is tested on synthetic data under variable SNR conditions, ranging from -20 dB to 10 dB (in 2 dB steps). For each SNR level, 400 samples are generated with random sediment types and parameter perturbations, which are not seen during training. It is important to note that the real experimental datasets are not included in any training set. They are exclusively used as independent test cases to evaluate generalization performance and avoid data leakage. These measures ensure that the training process reflects real conditions while maintaining robust model evaluation.

The model was trained with a batch size of 128 and optimized using the Adam optimizer with an initial learning rate of 0.001. The learning rate was reduced by a factor of 0.1 every 50 epochs, while the momentum and weight decay were set to 0.9 and 0.00001, respectively. Training continued for a maximum of 500 iterations. The model was trained on a Linux system using the PyTorch^[21] framework and four NVIDIA Tesla V100 GPUs, requiring approximately 6 h for 500 epochs.

$$A = \frac{\sum_{i=1}^N \eta(i)}{N} \times 100\%,$$

$$\begin{cases} \eta(i) = 1, & |\hat{y}_i - y_i|/y_i \leq 0.2, \\ \eta(i) = 0, & \text{else.} \end{cases} \quad (4)$$

To quantify the performance, we defined root mean square error (RMSE) and accuracy as evaluation metrics on the test set. Assuming $\hat{y}_i$ and y_i are the estimated depth and true depth, N is the number of statistical samples in the dataset, the accuracy is assessed as a measure of result reliability, considering results with a relative error below 20% as reliable. The method's accuracy is based on the percentage of such reliable outcomes within the test set.

4 Performance evaluation

4.1 Experimental dataset

Passive underwater source localization experiments were conducted in a typical deep sea area in 2021 and 2025. The 2021 experiment employed a 32-element VLA with 10-m spacing deployed at a depth of 3760 m, during which narrowband signals at [63, 79, 105, 126] Hz were continuously transmitted from a towed source to emulate a high-SNR target with a source level of 115 dB. The 2025 experiment served as a follow-up campaign, utilizing a 64-element VLA with the same 10-m spacing at a depth of 4300 m, where narrowband signals at [63, 97, 157, 199, 260, 370] Hz were transmitted to emulate a low-SNR target with a source level of 105 dB.

Overviews of the environmental conditions, source-receiver geometry, and received acoustic characteristics for both experiments are shown in Figs. 8 and 9. Because the performance of the hybrid method depends on the clarity of the interference structure, a quantitative metric is required to assess fringe visibility. To this end, fringe contrast (C) is employed, defined in Eq. (5) as the normalized difference between the mean intensity around the dominant interference trajectory and that in background regions. A larger C indicates a more distinct and well-defined interference track, whereas a smaller value reflects degradation due to noise.

$$C = \frac{\mu_{\text{signal}} - \mu_{\text{background}}}{\mu_{\text{signal}} + \mu_{\text{background}}}. \quad (5)$$

In our experiments, the high-SNR dataset yields a fringe contrast of 0.843, substantially higher than the value of 0.611 obtained from the low-SNR dataset, confirming that the interference structure is considerably clearer and more coherent under high-SNR conditions.

Table 1 Parameters used for the training set and validation set

Parameter	Training Set 1	Training Set 2	Validation Set
Source depth (m)	2:2:200	2:2:200	10:10:200
Source distance (km)	1:0.1:20	1:0.1:20	1:1:20
R_{SN} (dB)	0	0	$-20:2:10$
Sediment	Hamilton nine types of sediments	One random type of sediment	1%:1%:1% (Random)
Frequency (Hz)	[63, 97, 157, 199, 260, 370]	[63, 97, 157, 199, 260, 370]	[63, 97, 157, 199, 260, 370]
SSP	One SSP measured by CTD	One SSP measured by CTD	One SSP measured by CTD

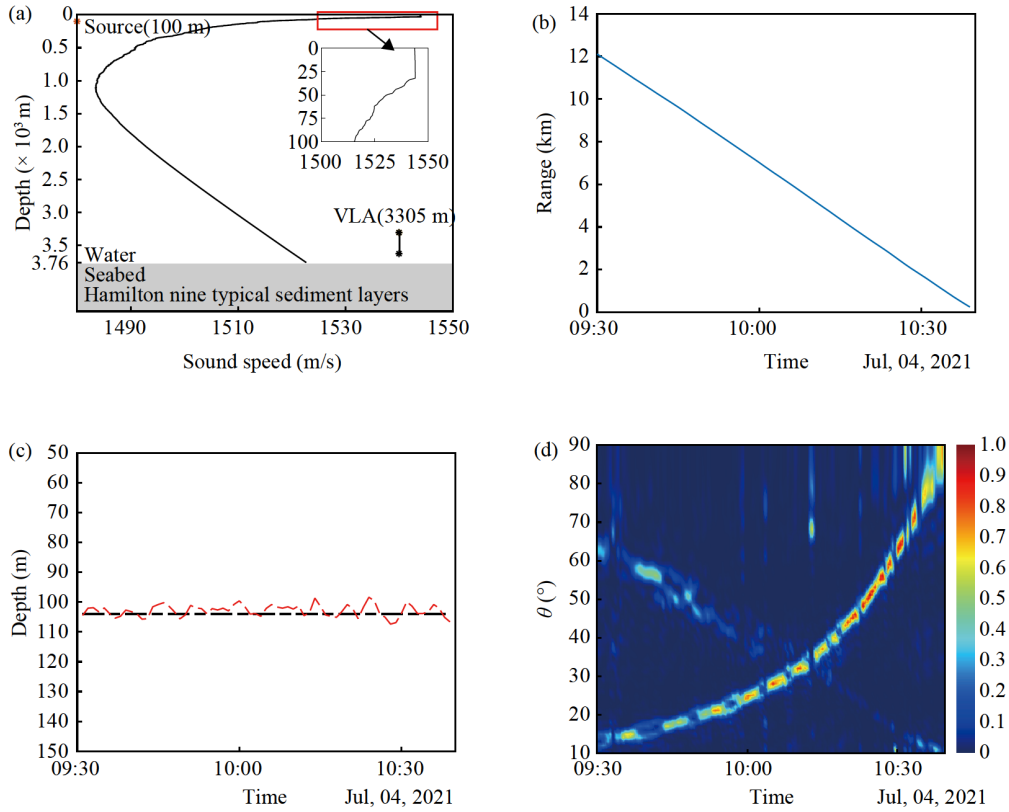


Fig. 8 Overview of deep-sea experiment in 2021: (a) Environment parameters of experiment area; (b) Source range; (c) Source depth; (d) Sound intensity in the beam-time domain of 126 Hz.

4.2 Comparison between different training sets

To evaluate the impact of different preprocessing strategies, we compare depth estimation performance using models trained with the same architecture but different input types. The model based on CBF-preprocessed data is referred to as ORN-CBF. When trained on Training Set 1 (which includes nine sediment types), it is labeled ORN-CBF-1; when trained on Training Set 2 (with one randomly selected sediment type), it is labeled ORN-CBF-2. Similarly, the model using standardized frequency-domain data without CBF preprocessing is denoted as ORN-Field, with ORN-Field-1 and ORN-Field-2 corresponding to the two training sets.

Figure 10 compares the performance of these four models on the validation set. Overall, ORN-CBF-1 achieves significantly higher accuracy and lower RMSE than ORN-CBF-2, indicating that incorporating diverse sediment types during training improves generalization. In contrast, both ORN-Field models perform worse, largely due to the lack of spatially coherent interference features and increased sensitivity to noise in the frequency domain.

As discussed in Sec. 2, the interference pattern in the negative vertical arrival angle domain is influenced by both source location and bottom sediment properties. Therefore, including multiple sediment types in model

training enhances robustness under uncertain environmental conditions.

4.3 Comparison with different CNN models

To evaluate the effectiveness of the proposed model in source depth estimation under uncertain environments, we compare it with two widely used convolutional neural network (CNN) architectures: VGG16 and GoogLeNet, both trained using Training Set 1.

Figure 11 shows that ORN-CBF-1 consistently achieves lower RMSE and higher accuracy across different SNR levels, outperforming both baseline models. While VGG16 and GoogLeNet have demonstrated strong performance in various tasks, ResNet50's residual connections enable deeper network training without compromising stability, making it particularly effective for extracting complex features in low-SNR acoustic environments. Its architecture has inspired many variants and remains a preferred backbone in deep learning applications.

4.4 Comparison with the conventional methods

Table 2 summarizes the prediction performance of the

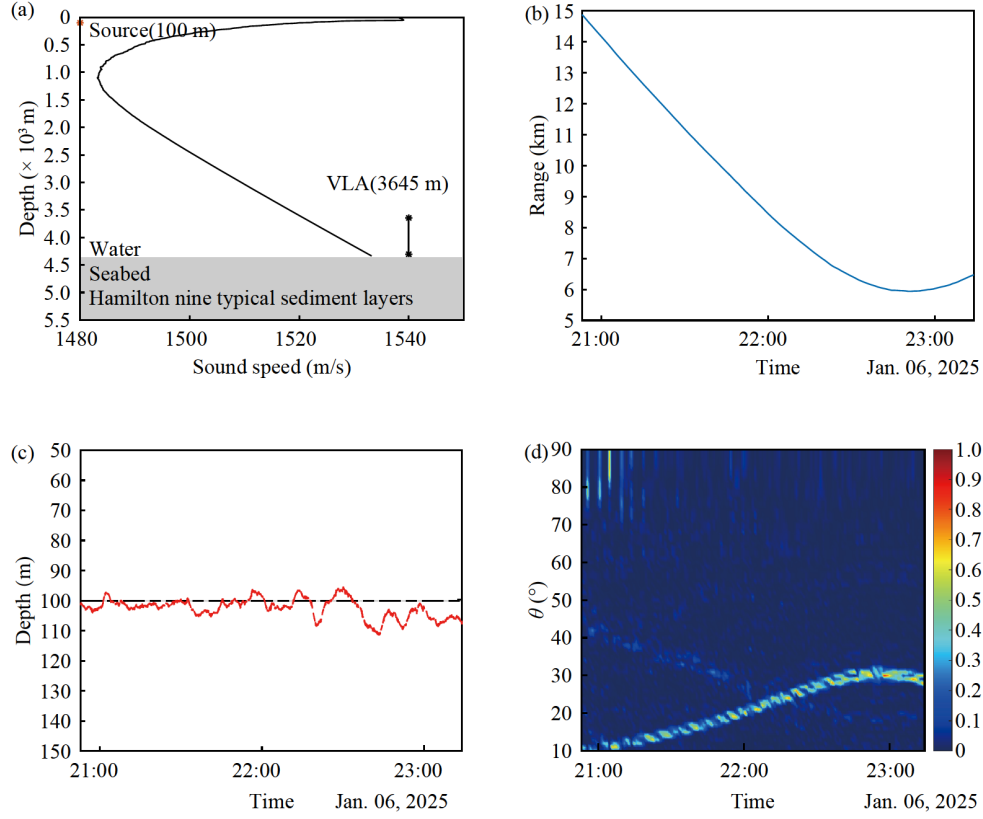


Fig. 9 Overview of deep-sea experiment in 2025: (a) Environment parameters of experiment area; (b) Source range; (c) Source depth; (d) Sound intensity in the beam-time domain of 157 Hz.

Table 2 Prediction performance of different models

Method	Experimental dataset-2021 (High SNR)		Experimental dataset-2025 (Low SNR)	
	RMSE	Accuracy	RMSE	Accuracy
ORN-CBF-1	12.8	98.6%	15.6	83.7%
MSTDE-C	22.5	69.2%	45.9	57.1%

two evaluated methods on both the high-SNR (2021) and low-SNR (2025) experimental datasets. The MSTDE-C method, although incorporating nonlinear error compensation within the MSTDE framework to alleviate the mismatch

caused by the assumed iso-velocity profile, still exhibits limited performance, particularly under weak-target, low-SNR conditions. In contrast, the proposed ORN-CBF-1 model achieves consistently lower RMSE and higher accuracy across both experiments, with its advantage becoming especially pronounced in the low-SNR case.

This performance difference is further illustrated in Figs. 12 and 13. In the high-SNR 2021 experiment (Fig. 12), both MSTDE-C and ORN-CBF-1 produce reasonable depth estimates, although ORN-CBF-1 shows smaller fluctuations and consistently lower relative errors. The contrast between

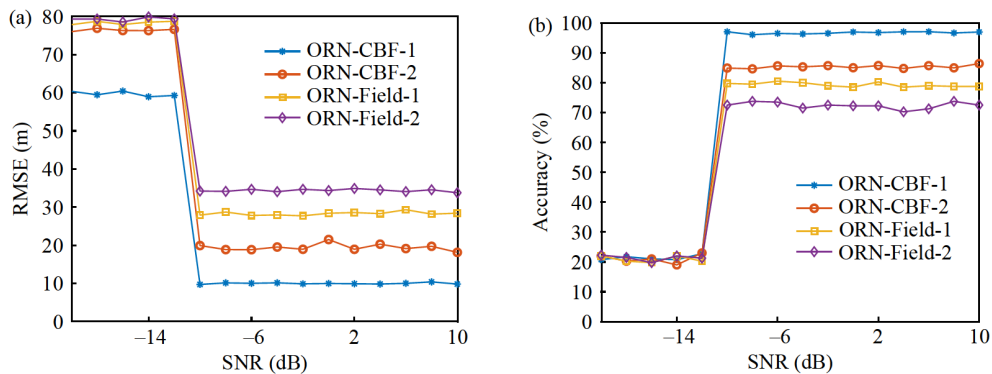


Fig. 10 Performance comparison between ORN-CBF and ORN-Field models trained on different datasets: (a) RMSE curves; (b) Accuracy curves of source depth estimation for the validation set.

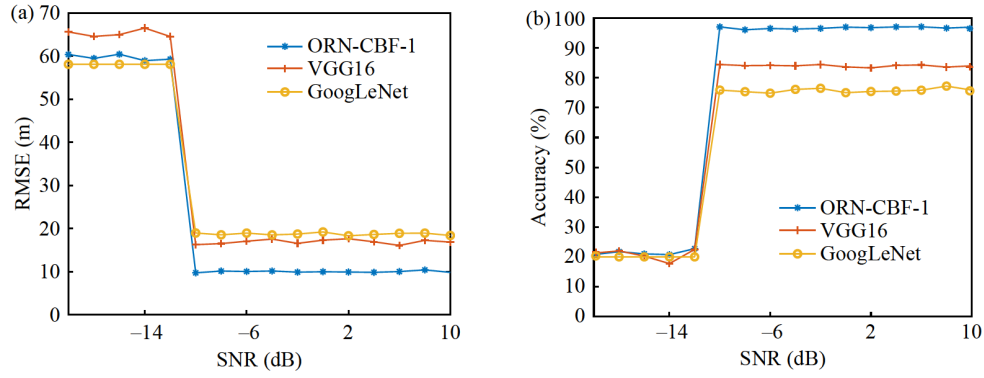


Fig. 11 Performance comparison between ORN-CBF-1, VGG16, and GoogLeNet: (a) RMSE curves; (b) Accuracy curves of source depth estimation for validation set.

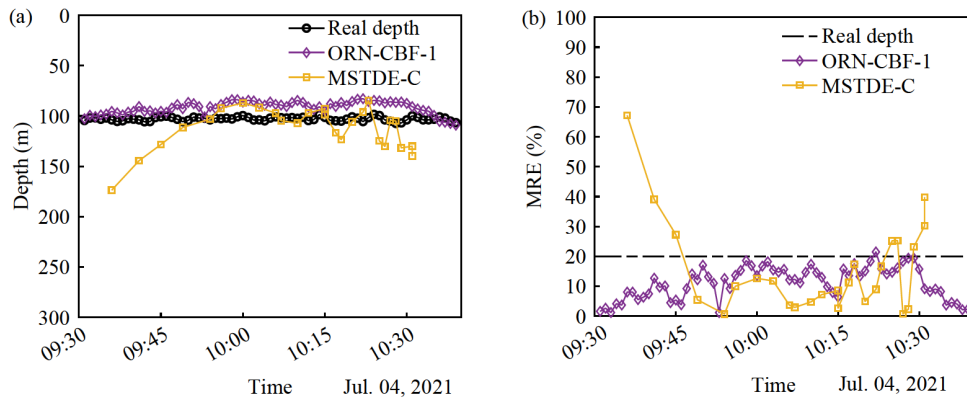


Fig. 12 Depth estimation results from the 2021 high-SNR experiment: (a) Predicted depths; (b) Relative errors.

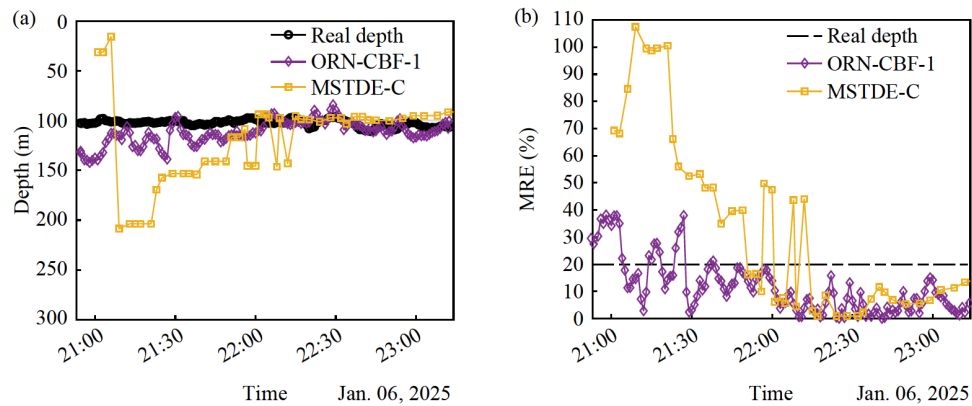


Fig. 13 Depth estimation results from the 2025 low-SNR experiment: (a) Predicted depths; (b) Relative errors.

the two methods becomes markedly clearer in the low-SNR 2025 experiment. As shown in Fig. 13(a), MSTDE-C exhibits large deviations and frequent outliers, particularly at long ranges where the weak-target interference structure is difficult to extract. In contrast, ORN-CBF-1 provides stable estimates that closely track the true source depth throughout the experiment. The corresponding mean relative error (MRE) curves in Fig. 13(b) reinforce this observation: MSTDE-C

reaches errors above 80% in the far-range region and remains unstable until the source approaches within approximately 10 km, whereas ORN-CBF-1 generally maintains an MRE below 20%. These results demonstrate that the proposed deep-learning-assisted method not only improves performance under favorable high-SNR conditions but, more importantly, effectively mitigates the severe degradation experienced by conventional interference-based approaches under weak-tar-

get, low-SNR scenarios.

Additionally, both the 2021 and 2025 sea trials feature a moving source with decreasing source–receiver distance over time (Figs. 8(b) and 9(b)). This variation allows us to investigate error behavior across different ranges. In the 2025 experiment, the MSTDE-C method exhibits large estimation errors at long distances (>12 km), which progressively diminish as the source approaches the array (<10 km). This correlates with the fringe visibility deterioration shown in Fig. 9(d) and quantified by the fringe contrast metric ($C = 0.611$). The ORN-CBF-1 method, by contrast, maintains lower errors throughout, demonstrating its robustness under poor interference conditions and long-range propagation scenarios.

5 Conclusion

This study presents a hybrid source localization method that integrates arrival-angle interference features with an optimized residual network to improve depth estimation performance for weak narrowband targets in deep-sea environments. By leveraging ray-theoretical insights and deep learning-based feature extraction, the proposed method effectively captures corrupted interference structures in the beam domain and establishes a more reliable mapping between these structures and the source depth.

Comprehensive evaluations were conducted using both simulated and experimental datasets. Simulations incorporating a nine-layer sediment model demonstrated that the ORN framework maintains strong robustness under environmental uncertainties, outperforming traditional CNN architectures and validating the advantages of residual learning. Experimental results further confirmed that the ORN-based estimator significantly enhances depth-prediction stability and accuracy compared with conventional interference-based approaches, particularly under low-SNR, long-range conditions where weak-target interference patterns are difficult to resolve.

Overall, the proposed hybrid method mitigates key limitations of traditional techniques and provides a practical and effective solution for passive source localization in weak-target deep-sea scenarios.

References

- [1] Tappert F D, Nghiem-Phu L. Localization of sources of acoustic transients using a broadband matched-field processing technique. *J. Acoust. Soc. Am.*, 1995, **97**(5): 3368

- [2] Duan R, Yang K D, Ma Y, et al. Moving source localization with a single hydrophone using multipath time delays in the deep ocean. *J. Acoust. Soc. Am.*, 2014, **136**(2): EL159–EL165
- [3] McCargar R K, Zurk L M. Depth-based suppression of moving interference with vertical line arrays in the deep ocean. *J. Acoust. Soc. Am.*, 2012, **132**(3): 2081
- [4] McCargar R, Zurk L M. Depth-based signal separation with vertical line arrays in the deep ocean. *J. Acoust. Soc. Am.*, 2013, **133**(4): EL320–EL326
- [5] Kniffin G P, Boyle J K, Zurk L M, et al. Performance metrics for depth-based signal separation using deep vertical line arrays. *J. Acoust. Soc. Am.*, 2016, **139**(1): 418–427
- [6] Wang W B, Wang Z, Su L, et al. Source depth estimation using spectral transformations and convolutional neural network in a deep-sea environment. *J. Acoust. Soc. Am.*, 2020, **148**(6): 3633–3645
- [7] Li H, Wang T, Su L, et al. Passive depth estimation for a narrowband source using a single vector sensor in deep water. *JASA Express Lett.*, 2023, **3**(6): 066002
- [8] Yang K D, Xu L Y, Yang Q L, et al. Striation-based source depth estimation with a vertical line array in the deep ocean. *J. Acoust. Soc. Am.*, 2018, **143**(1): EL8–EL13
- [9] Zhou L S, Zheng G Y, Yang T C. Target depth estimation by frequency interference matching for a deep vertical array. *Appl. Acoust.*, 2022, **190**: 108671
- [10] Niu H Q, Reeves E, Gerstoft P. Source localization in an ocean waveguide using supervised machine learning. *J. Acoust. Soc. Am.*, 2017, **142**(3): 1176–1188
- [11] Niu H Q, Ozanich E, Gerstoft P. Ship localization in Santa Barbara Channel using machine learning classifiers. *J. Acoust. Soc. Am.*, 2017, **142**(5): EL455–EL460
- [12] Huang Z, Xu J, Gong Z, et al. Source localization using deep neural networks in a shallow water environment. *J. Acoust. Soc. Am.*, 2018, **143**(5): 2922–2932
- [13] Niu H Q, Gong Z, Ozanich E, et al. Deep-learning source localization using multi-frequency magnitude-only data. *J. Acoust. Soc. Am.*, 2019, **146**(1): 211–222
- [14] Wang W B, Ni H Y, Su L, et al. Deep transfer learning for source ranging: deep-sea experiment results. *J. Acoust. Soc. Am.*, 2019, **146**(4): EL317–EL322
- [15] Benson J, Chapman N R, Antoniou A. Geoacoustic model inversion using artificial neural networks. *Inverse Probl.*, 2000, **16**(6): 1627–1639
- [16] Caiti A, Jesus S M. Acoustic estimation of seafloor parameters: a radial basis functions approach. *J. Acoust. Soc. Am.*, 1996, **100**(3): 1473–1481
- [17] Zhang X H, Lu Z B, Kang C Y. Underwater acoustic targets classification using support vector machine. International Conference on Neural Networks and Signal Processing, IEEE, Nanjing, China, 2003: 918–921
- [18] Fischell E M, Schmidt H. Classification of underwater targets from autonomous underwater vehicle sampled bistatic acoustic scattered fields. *J. Acoust. Soc. Am.*, 2015, **138**(6): 3773–3784
- [19] Niu H Q, Li Z L, Wang H B, et al. Overview of machine learning methods in underwater source localization. *J. Signal Process.*, 2019, **35**(4): 621–635
- [20] Hamilton E L. Geoacoustic modeling of the sea floor. *J. Acoust. Soc. Am.*, 1980, **68**(5): 1313–1340
- [21] Chen K M, Cofer E M, Zhou J, et al. A PyTorch-based deep learning library for sequence data. *Nat. Methods*, 2019, **16**(4): 315–318