ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

# Interpretable intelligent gas-bearing reservoir prediction using time–frequency analysis and manifold-regularized semi-supervised GAN

**Shuying Ma[1], Junxing Cao[2,3]*, Rong Wang[1]*, Xudong Jiang[3], Jun Wang[2,3], Lingsen Zhao[3], Hong Li[3], and Xin Tang[3]**

[1]School of Network and Communication Engineering, Chengdu Technological University, Chengdu, Sichuan, China

[2]State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu, Sichuan, China

[3]College of Geophysics, Chengdu University of Technology, Chengdu, Sichuan, China

## Abstract

As hydrocarbon exploration advances toward deep and complex reservoirs, the identification accuracy of traditional time–frequency (TF) analysis is constrained by strongly heterogeneous geological conditions. Concurrently, while deep learning has shown great potential, mainstream supervised models commonly face the dual challenges of scarce labeled samples and the lack of interpretability in their black-box decision-making processes. To address these challenges, this study proposes an innovative, intelligent prediction framework integrating high-precision TF analysis, manifold-regularized semi-supervised generative adversarial networks, and SHapley Additive exPlanations (SHAP) for interpretability. First, the Fourier-based synchrosqueezing transform was employed to extract two-dimensional TF features with superior energy concentration, effectively overcoming the resolution limits imposed by the Heisenberg uncertainty principle. Subsequently, the manifold-regularized semi-supervised generative adversarial network was developed. By incorporating manifold regularization constraints, the discriminator captures the intrinsic topological structure of large-scale unlabeled samples, effectively leveraging data geometry to significantly enhance generalization capability under sparse-label conditions. Finally, the SHAP method was utilized to conduct a post hoc interpretation. Experimental results on the Marmousi II model demonstrate a remarkable testing accuracy of 98.4%. In a real-world application to deep marine reservoirs in the Sichuan Basin, the framework achieved an 85.0% testing accuracy using only 5% labeled samples. Compared to baseline models, the semi-supervised strategy and manifold regularization contributed accuracy gains of 18.8% and 5.0%, respectively. SHAP analysis further confirms the model's adaptive capability to extract geophysical features, enabling it to deconstruct the tuning-effect patterns in synthetic data and the low-frequency enhancement/high-frequency attenuation patterns in real data, respectively. This validation of geophysical consistency provides a theoretical foundation for the application of artificial intelligence in complex hydrocarbon exploration.

***Keywords:*** Gas-bearing prediction; Semi-supervised learning; Synchrosqueezing transform; Generative adversarial network; SHapley Additive exPlanations

***Corresponding authors:**
Junxing Cao
(caojx@cdut.edu.cn)
Rong Wang
(wrong1@cdtu.edu.cn)

**Citation:** Ma S, Cao J, Wang R, *et al*. Interpretable intelligent gas-bearing reservoir prediction using time–frequency analysis and manifold-regularized semi-supervised GAN. *J Seismic Explor*. 2026;35(2):025450102. doi: 10.36922/JSE025450102

**Received:** November 5, 2025

**Revised:** January 9, 2026

**Accepted:** February 10, 2026

**Published online:** April 24, 2026

**Copyright:** © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note:** AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

As global hydrocarbon exploration shifts toward deep, ultra-deep, and complex hidden reservoirs, seismic signal identification and fluid prediction face unprecedented challenges. When seismic waves propagate through fluid-saturated porous media, they undergo significant energy absorption, scattering, and attenuation due to the combined effects of fluid viscosity and reservoir heterogeneity. These phenomena typically manifest as localized energy anomalies and dispersion signatures in the time–frequency (TF) domain, forming the theoretical foundation for geophysical gas-bearing identification.

Traditional identification methods primarily rely on amplitude anomalies in the time domain (e.g., bright spots[1], dim spots[2], and flat spots[3]) or characteristic changes in the frequency domain (e.g., low-frequency shadows[4], fluid mobility indicators[5], and high-frequency attenuation gradients[6]). While these qualitative methods have achieved some success in simple geological scenarios such as shallow and thick reservoirs, they struggle to support high-reliability quantitative evaluation for deep and thin interbedded targets due to low signal-to-noise ratios (SNR), tuning effects, and strong multi-solution ambiguity.

In recent years, the data-driven paradigm centered on deep learning has demonstrated significant application potential in the intelligent detection of gas-bearing reservoirs, owing to its exceptional capability to uncover non-linear patterns. Numerous studies have attempted to incorporate architectures such as long short-term memory[7,8], deep neural networks[9,10], and convolutional neural networks[11], combined with multiple seismic attributes[12,13], achieving positive progress in specific geological scenarios.

Nevertheless, existing deep learning approaches face substantial challenges in practical production applications: on the one hand, certain models exhibit excessive reliance on long sequences or pre-stack data, resulting in substantial computational overhead; on the other, the manual pre-selection of seismic attributes is prone to introducing subjective bias. A more fundamental constraint lies in the fact that existing models predominantly follow a supervised learning paradigm, whose performance is highly dependent on large-scale, high-confidence labeled samples. In actual oil and gas exploration, constrained by high drilling and coring costs, confirmed gas-bearing well-point labels are extremely scarce and sparsely distributed. This severely limits the generalization capability and engineering applicability of supervised learning models.

To overcome the label scarcity bottleneck, semi-supervised learning (SSL)[14] has emerged as a research focus in geophysical intelligent recognition. This approach leverages both limited labeled data and vast unlabeled datasets for training. Among these, semi-supervised models based on generative adversarial networks (GANs)[15,16] demonstrate unique advantages in feature representation learning and data distribution modeling, showing initial promise in seismic phase identification[17], reservoir prediction[18], and lithology prediction.[19]

Nevertheless, achieving efficient and robust detection of gas-bearing reservoirs requires overcoming three major bottlenecks:

(i) Effective characterization of sensitive features: Extracting TF representations from non-stationary raw signals that are most responsive to gas content and most information-rich forms the foundation for determining model learning efficiency.
(ii) Robustness and generalizability of model training: When confronted with local perturbations and strong background noise, enhancing the model's smoothness and stability within the manifold space is paramount.
(iii) Interpretability of decision mechanisms. Oil and gas exploration constitutes a high-risk decision-making domain. If the decision-making mechanisms of a model (i.e., its black-box characteristics) cannot be understood and trusted, the practical value of its predictive results will be significantly diminished.

By addressing the aforementioned challenges, this study proposes a novel intelligent detection framework for gas-bearing reservoirs that integrates high-precision Fourier-based synchrosqueezing transform (FSST)[20], a manifold-regularized semi-supervised GAN (MR-SSGAN), and SHapley Additive exPlanations (SHAP).[21] The core innovations include:

(i) High-precision TF feature extraction strategy: The FSST is introduced, which uses an instantaneous frequency rearrangement method to squeeze distributed energy back onto actual frequency ridges, significantly reducing the TF blurring inherent in conventional analysis.[22,23] This produces two-dimensional (2D) highly concentrated TF features for deep learning.
(ii) Construction of the MR-SSGAN model: The MR-SSGAN network architecture is developed. This model not only uses GANs to create high-quality pseudo-samples for training set augmentation, but it also incorporates a novel manifold regularization[24] term based on the Laplace norm into the discriminator loss function.[25,26] This constraint compels the model's decision boundary to traverse low-density regions within the data distribution while maintaining

smoothness on the data manifold. Consequently, it efficiently extracts topological structural information from vast unlabeled seismic datasets under conditions of minimal labeled samples, significantly enhancing the model's robustness in complex geological environments.

(iii) Verification of physical interpretability: The SHAP method from game theory is employed to generate post hoc explanations for the trained MR-SSGAN model. By quantifying the marginal contribution of frequency- and time-domain features to the final prediction, the geophysical logic underlying the model's decisions is elucidated, validating the consistency of the model's predictions with fundamental rock physics principles.

## 2. Methods

The overall workflow proposed in this paper comprises three complementary modules, as illustrated in Figure 1: (i) data preprocessing: synchrosqueezing transform (SST)-based TF representation; (ii) construction and training of the MR-SSGAN; and (iii) decision attribution analysis by SHAP.

### 2.1. Data preprocessing: Synchrosqueezing transform-based time–frequency representation

Seismic signals are non-stationary signals whose frequency components undergo significant temporal variation. The traditional short-time Fourier transform (STFT) maps one-dimensional (1D) time signals $f_t \in L^2(R)$ into a 2D TF domain through a sliding window function:

$$STFT_f(f,t) = \int f(t) w(t-\tau) e^{-i2\pi ft} dt \tag{1}$$

where $t$ is time, $f$ is frequency, and $w(t)$ is the window function. The energy distribution in the STFT is significantly influenced by the window function, exhibiting a blurring effect. According to Heisenberg's uncertainty principle, optimal time resolution and frequency resolution cannot be achieved simultaneously, thereby limiting its ability to provide a detailed characterization of complex reservoir signals.

Although the continuous wavelet transform (CWT) incorporates multiscale analysis, it remains constrained in temporal resolution at low frequencies and frequency resolution at high frequencies, while also being sensitive to the selection of wavelet basis functions.

To address the energy blurring effect, Daubechies *et al.*[20] proposed the SST. As a post-processing reassignment technique based on signal phase derivatives, SST squeezes energy dispersed across the TF plane back onto the true instantaneous frequency ridge. For a signal $f_t \in L^2(R)$, its

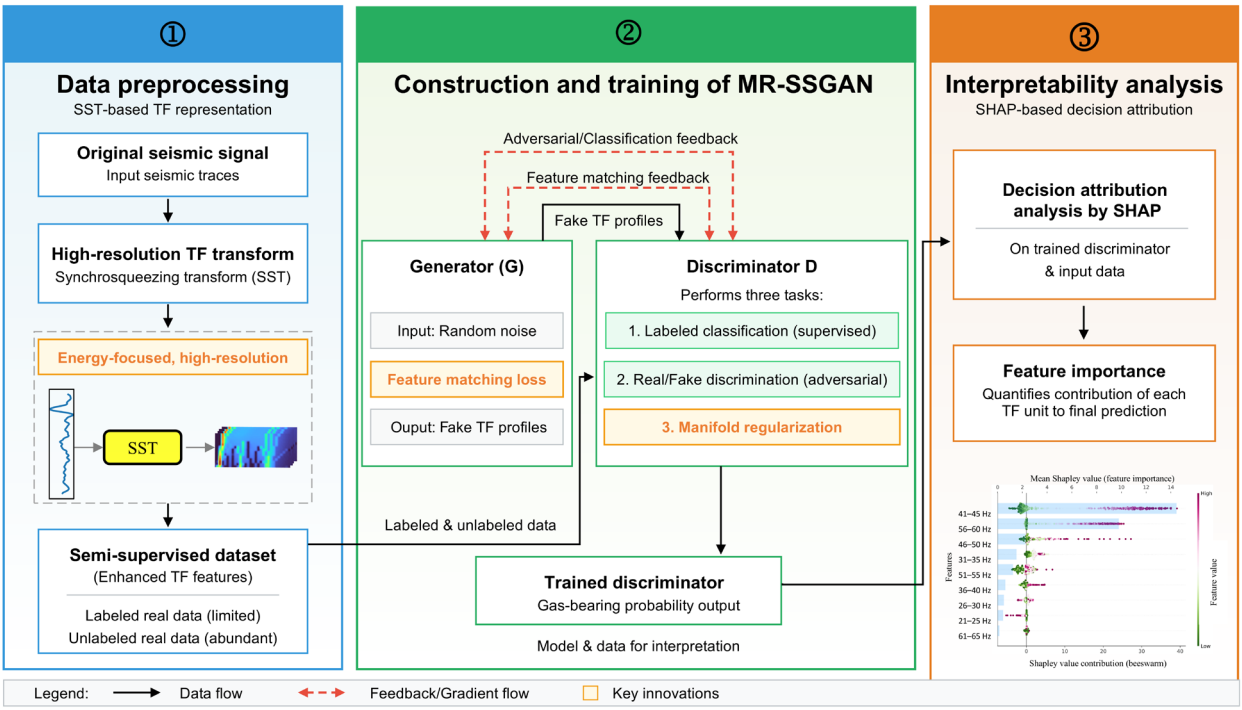


**Figure 1.** Workflow of the intelligent gas-bearing reservoir prediction framework
Abbreviations: MR-SSGAN: Manifold-regularized semi-supervised generative adversarial network; SHAP: SHapley Additive exPlanations; SST: Synchrosqueezing transform; TF: Time–frequency.

instantaneous frequency can be estimated using the partial derivatives of the STFT coefficients:

$$w_f(f,t) = -i\frac{\partial_t \hat{STFT}_f(f,t)}{\hat{STFT}_f(f,t)} \tag{2}$$

where $\hat{STFT}_f(f,t)$ is the complex-valued STFT coefficient. Using this instantaneous frequency estimate, the energy at $(t,f)$ is redistributed to $w_f(f,t)$. For the STFT-based FSST[27], the rearrangement operation is defined as:

$$FSST_f(w_l,t) = \int_{\{f:w_f(f,t)\in F_l\}} \left|STFT_f(f,t)\right| f^{-3/2} df \tag{3}$$

where $F_l$ is the frequency interval and $w_l$ is the center frequency. This operation localizes the originally dispersed energy around the true instantaneous frequency, thereby producing a high-resolution TF spectrum while preserving signal reconstructibility.

### 2.2. Construction and training of the manifold-regularized semi-supervised generative adversarial networks

Standard GANs comprise a generator *G* and a discriminator *D*, which undergo adversarial training through a minimax game. In a semi-supervised setting (SSGAN), the discriminator is transformed into a *K* + 1-class classifier (*K* real classes plus one generated class). For gas-bearing prediction, $K = 2$ (gas-bearing vs. gas-free). *D* must not only determine whether the input data are real or generated, but also predict the physical category (gas-bearing or gas-free) to which genuine samples belong.

Although SSGAN uses unlabeled data, it does not explicitly exploit the topological geometric structure of the data. In contrast, the core innovation of MR-SSGAN is that it fully exploits GANs' potential for modeling natural data manifolds by introducing manifold regularization terms during discriminator *D* training and imposing feature-matching strategies on the generator *D*, resulting in deeper exploitation of the geometric–topological structure of the data.

The loss function $L_D$ for the MR-SSGAN discriminator comprises supervised loss $L_L$, unsupervised loss for genuine unlabeled data $L_U$, unsupervised loss $L_F$ for generated pseudo-data , and the manifold-regularization term $R_{manif}$, as detailed below:

$$L_D = L_L + \omega(L_U + L_F) + \lambda R_{manif} \tag{4}$$

where $\omega = 0.5$ is the weight balancing the contributions of labeled and unlabeled samples, and $\lambda = 10^{-3}$ is the manifold-regularization weight.

For real labeled samples $(x,y)$, the task of the discriminator is to minimize the cross-entropy between the ground-truth labels and the model's predicted distribution. Its loss function $L_L$ is the standard cross-entropy loss:

$$L_L = -E_{x,y\sim p_{data}(x,y)}\left[\log p_D(y \mid x, y < K+1)\right] \tag{5}$$

where $p_D(y \mid x, y < K+1)$ is the probability of the correct label *L* given input *x*.

For real unlabeled data and generated pseudo-data, the corresponding loss terms are unsupervised. $p_D(y = K+1 \mid x)$ provides the probability that input *x* belongs to class *K* + 1 (i.e., pseudo-data). If the input is pseudo-data, the discriminator should identify it as pseudo-data; in this case, the task of the discriminator is to maximize $p_D(y = K+1 \mid x)$. If the input is real unlabeled data, the discriminator should identify it as real data (i.e., the label is not *K* + 1), and its task becomes to minimize $p_D(y = K+1 \mid x)$. Therefore, the corresponding two unsupervised loss functions are defined as follows:

$$L_U = -E_{x\sim p_{data}(x)}\left[\log\left[1 - p_D(y = K+1 \mid x)\right]\right] \tag{6}$$

$$L_F = -E_{x\sim G}\left[\log\left[p_D(y = K+1 \mid x)\right]\right] \tag{7}$$

In the experiment, to simplify the discriminator output while maintaining consistency with the actual number of classes, the property of the Softmax function where subtracting a constant vector from the logits does not change the output probabilities was utilized. The final dimension $l_{K+1}$ of the *K* + 1-dimensional logit vector produced by the discriminator was set to zero, without affecting the final probability outputs.

The manifold hypothesis posits that high-dimensional data are actually distributed on a low-dimensional manifold structure. To explicitly utilize this structure within SSGAN, MR-SSGAN introduces a manifold regularization term $R_{manif}$. According to manifold learning theory, the smoothness of the discriminator can be measured by the integral of the squared gradient over the data manifold (i.e., the Laplacian norm $\|D\|_L^2$). For manifolds parameterized by the generator, samples $z^{(i)}$ may be drawn from the latent representation space *z*. The discriminator's gradient on the manifold can then be estimated via Monte Carlo integration:

$$\|D\|_L^2 \approx \frac{1}{n}\sum_{i=1}^{n}\left\|J_z D\left(G\left(z^{(i)}\right)\right)\right\|_F^2 \tag{8}$$

In **Equation 8**, *n* is the number of samples, which corresponds to the batch size in stochastic gradient descent. *J* represents the Jacobian matrix, used to estimate the

gradient of the discriminator output on the data manifold. $\|\cdot\|_F$ denotes the Frobenius norm. A smaller value for this term indicates smoother discriminator outputs on the manifold.

In practical training, considering the substantial computational overhead associated with directly computing the Jacobian matrix, this study employs a stochastic finite-difference method to approximate this gradient term. If a small perturbation $\varepsilon\bar{\delta}$ is applied to the generator's input $z$, the generated sample should move a small distance along the manifold, at which point the discriminator $D$'s predicted output should remain stable. This yields the final manifold regularization term:

$$R_{manif} = E_{z\sim p_z(z)} d\left(G(z)\right) - D\left(G\left(z+\varepsilon\bar{\delta}\right)\right)_2^2 \tag{9}$$

where $\bar{\delta}$ is the perturbation term, whereas $\bar{\delta} = \frac{\delta}{\|\delta\|}$, $\delta \sim N(0, I)$, and $\varepsilon = 10^{-5}$ represents the perturbation norm.

To address the issues of non-convergence and instability encountered by generators in conventional adversarial training, this study introduces a feature-matching strategy[28] to construct the generator's loss function. The core of this strategy lies in shifting the generator's optimization objective away from solely maximizing the discriminator's misclassification probability. Instead, it seeks to achieve alignment between the distributions of generated and real data within the discriminator's feature space.

Specifically, the expectation of the activation features $h(G[z])$ of the pseudo-samples $G(z)$ generated by the generator at a certain intermediate layer of the discriminator $D$ is designed to match the expectation of the activation features $h(x)$ of the real data $x$ at the same layer. The loss function is defined as follows:

$$L_G = \left\| E_{x\sim p_{data}(x)} h(x) - E_{z\sim p_z(z)} h\left(G(z)\right) \right\|_1 \tag{10}$$

where $h(x)$ represents the activation output of an intermediate layer of the discriminator. In this study, to capture deep and stable semantic representations, the output from the penultimate layer of the discriminator was selected as the feature-matching representation.

The specific network topology and parameter configuration of MR-SSGAN are illustrated in Figure 2. The generator's architecture adopts the classic up-sampling paradigm proposed by Ioffe and Szegedy[29], which has

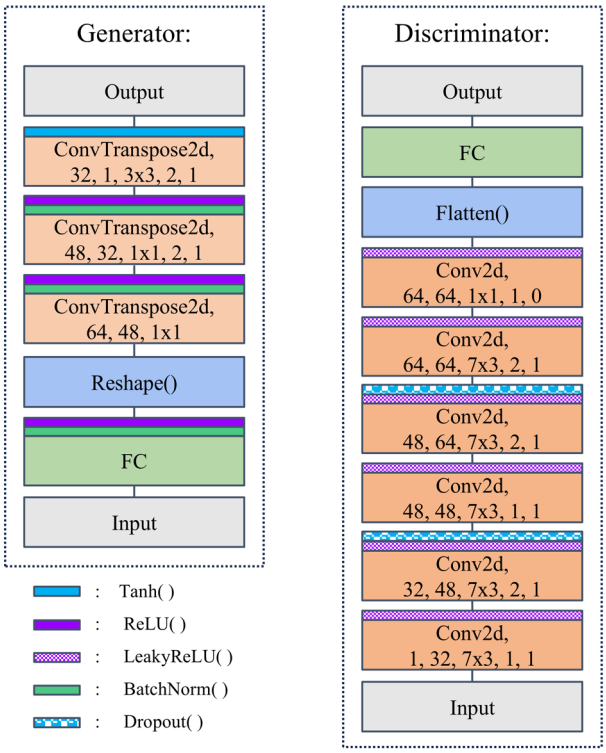


**Figure 2.** Generator and discriminator network architecture of the manifold-regularized semi-supervised generative adversarial network
Abbreviations: BatchNorm: Batch normalization; Conv2d: Convolutional 2D layer; FC: Fully connected layer; ReLU: Rectified linear unit; Tanh: Hyperbolic tangent function.

been adaptively modified to accommodate the input characteristics of 2D TF profiles. It consists of a three-layer transposed convolutional network incorporating batch normalization. In contrast, the discriminator is designed with reference to the robust architecture introduced by Salimans *et al.*[28] for improving GAN training. It is structured as a six-layer 2D convolutional neural network featuring a dropout randomization mechanism, aimed at enhancing the robustness of feature extraction and mitigating overfitting.

By alternately and iteratively optimizing the discriminator loss and the generator loss, the MR-SSGAN framework can efficiently mine potential discriminative information from large-scale unlabeled data while leveraging a small number of critical labeled samples for precise supervision. After training convergence, the discriminator learns the essential characteristic responses of gas-bearing reservoirs in the TF domain, enabling its direct application to gas-bearing prediction of practical 2D seismic profiles.

### 2.3. Decision attribution analysis by SHapley Additive exPlanations

Although the MR-SSGAN model, with its deep non-linear architecture, can capture complex TF features and their interactive effects within seismic signals, its intricate internal parameter mapping results in a decision-making process that exhibits typical black-box properties. To enhance the credibility of model predictions and elucidate its decision-making mechanisms, this study introduces the game-theoretic SHAP attribution analysis.

The core concept of this method is to treat the model's prediction for a specific sample as the result of the combined effects of all input features. By calculating Shapley values, it equitably quantifies the marginal contribution of each feature to the prediction outcome. The computational process follows the classical Shapley formula:

$$\phi_j = \sum_{S \subseteq F \setminus (j)} \left[ \frac{|S|!(M-|S|-1)!}{M!} \times \left[ f\left(S \cup \{j\}\right) - f\left(S\right) \right] \right] \tag{11}$$

where $M$ is the total number of features, $F$ is the set of all input features, and $f$ is the discriminator model being explained. $S$ denotes any possible subset derived from the feature set $F$ excluding feature $j$ (i.e., $S \subseteq F \setminus \{j\}$). The difference $f\left(S \cup \{i\}\right) - f\left(S\right)$ represents the marginal contribution of adding feature $j$ to the existing subset $S$. By calculating the weighted average of this marginal contribution over all possible subsets $S$, the resulting $\phi_j$ is the Shapley value for feature $j$, which comprehensively measures its contribution to the model's output.

To visually demonstrate the decision-making basis of the MR-SSGAN discriminator for individual gas-bearing prediction, SHAP approximates the original model's decision behavior by constructing a linear local surrogate model $g(x)$. Its additive decomposition expression is as follows:

$$g\left(x\right) = \phi_0 + \sum_{j=1}^{p} \phi_j \tag{12}$$

where $x$ is a simplified binary input vector, $P$ represents the total dimensionality of the input TF features, $\phi_0$ denotes the baseline prediction mean for all samples, and $\phi_j$ represents the attribution contribution value corresponding to feature $j$. Through this linear additive decomposition, SHAP can break down complex non-linear decisions into the sum of contributions from each TF unit, thereby clearly identifying the key TF patterns that drive the model's judgment of gas-bearing or gas-free.

## 3. Results

### 3.1. Synthetic data results using the Marmousi II model

This study first validates the algorithm's performance based on the Marmousi II model.[30] Building upon the classical Marmousi model, this iteration incorporates reservoir simulation capabilities, rendering it suitable for assessing the performance of gas-bearing identification algorithms. The core research target comprises a representative gas-bearing sandstone lens located within the common depth point (CDP) 420–550 interval and spanning 1.3–1.4 s, as illustrated in Figure 3A.

To simulate complex field acquisition conditions and assess model robustness, random Gaussian white noise was introduced into the original seismic record, creating a noisy dataset with an SNR of 0 dB. During model training, both the original noise-free data and the 0 dB SNR noisy data were utilized simultaneously to enhance MR-SSGAN's robustness and adaptability to noise.

#### *3.1.1. Time–frequency concentration comparison*

For the CDP 475 single-channel seismic signal, this study compared the feature characterization accuracy of four TF transformation methods (Figure 4). As the performance of CWT is highly dependent on the mother wavelet, the generalized Morse wavelet was selected. Its flexibly adjustable TF window shape enables effective localized characterization. As shown in Figure 4A, although the CWT TF spectrum possesses extremely high temporal resolution, its energy exhibits significant dispersion along the frequency axis.

For the STFT, to overcome the severe spectral leakage

inherent in traditional window functions (e.g., the Hanning window) during detailed analysis, this study employed the discrete phase shell sequence as the window function. This sequence exhibits extremely low sidelobes and rapid sidelobe attenuation characteristics.

Additionally, the findings demonstrated that the frequency resolution of the discrete phase shell sequence-based STFT significantly outperforms that of the CWT (Figure 4B). However, constrained by the Heisenberg uncertainty principle, its temporal resolution remains limited. To further enhance the energy concentration and resolution of the TF representation, the SST was introduced to rearrange the energy in the aforementioned results. Figure 4C, D confirm that both wavelet SST (WSST) and FSST substantially enhance TF concentration: WSST achieves frequency ridge convergence while preserving high temporal resolution, whereas FSST significantly strengthens energy concentration in the temporal dimension while maintaining high frequency resolution.

Further examination of the characterization performance of the aforementioned methods across the profile data was conducted using the 45 Hz frequency-

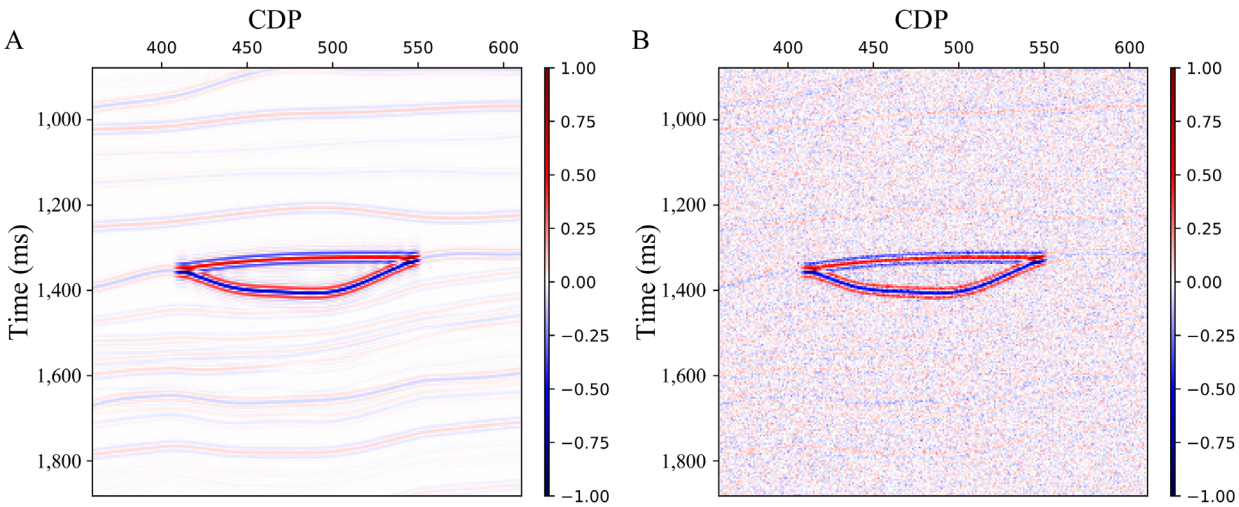


**Figure 3.** Partial data from the Marmousi II model. (A) Clean profile. (B) Noisy profile.
Abbreviations: CDP: Common depth point; SNR: Signal-to-noise ratio.

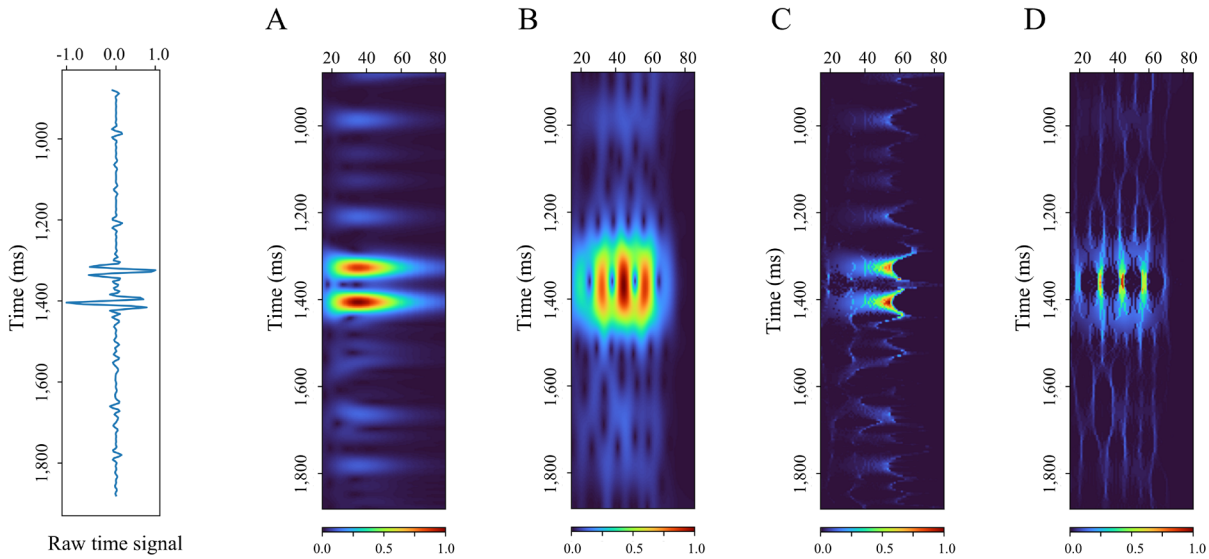


**Figure 4.** The original time-domain signal of the common depth point-475 and its corresponding time–frequency profiles: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

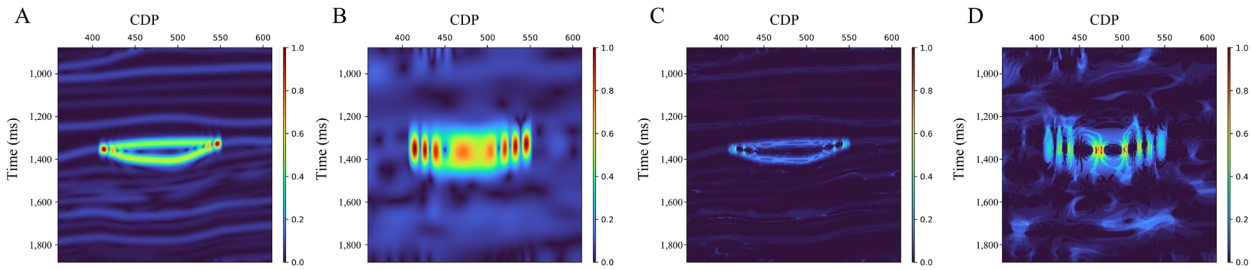


**Figure 5.** 45 Hz frequency slices: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform
Abbreviation: CDP: Common depth point.

separated profile (Figure 5). Results indicate that, while the CWT (Figure 5A) accurately delineates the reservoir top and bottom interfaces, its ability to reveal hydrocarbon enrichment within the reservoir is limited. The STFT (Figure 5B) identifies energy anomaly zones extending far beyond actual reservoir boundaries. Moreover, due to temporal domain blurring effects, the predicted results exhibit pronounced lateral discontinuities and spatial dispersion.

Following synchrosqueezing processing, energy focusing was substantially enhanced, yet identification limitations persisted: WSST (Figure 5C) tended to amplify reservoir boundaries, while FSST (Figure 5D) only displayed high-energy responses within the central reservoir zone. This series of qualitative comparisons demonstrates that, although advanced TF analysis techniques can significantly enhance the physical concentration of seismic characteristics, manual interpretation relying solely on traditional spectral decomposition attributes still suffers from interpretational ambiguity and uncertainty, making it difficult to achieve precise delineation of gas boundaries and intensity.

### *3.1.2. Ablation and robustness analysis*

Traditional gas-bearing prediction based on spectral decomposition relies heavily on manual qualitative interpretation of TF slices. This process is not only highly subjective and inefficient but also struggles to capture subtle, non-linear, sensitive features within seismic signals. In contrast, deep learning frameworks can automatically extract higher-order representations from complex seismic data through multi-layer non-linear mappings. Addressing the challenges of sparse labeling and vast unlabeled datasets in practical geophysical exploration, this study developed an MR-SSGAN model. The model's core advantage lies in leveraging the adversarial interaction between the discriminator and the generator to simultaneously extract prior distribution features from unlabeled data, while strengthening the model's generalization boundaries in sparse labeling scenarios through manifold constraints.

To achieve synchronous representation of TF features, a 2D sliding window sampling strategy $\left(f_{lgh}, t_{lgh}\right)$ was designed to construct 2D TF slice samples. Along the time dimension ($t_{lgh}$), adjacent sampling points before and after the target point are captured, balancing localized temporal information with the integrity of contextual feature learning and preventing degradation in localization accuracy. In the frequency dimension ($f_{lgh}$), the window range was determined based on the instantaneous spectral energy distribution. To capture core signal features while filtering out interference, the frequency interval corresponding to the cumulative energy between the 5th and 95th percentiles at each instantaneous frequency was statistically determined. Through this approach, each sample point was mapped onto a 2D slice sample carrying adjacent spatiotemporal information. Within the experimental configuration of the Marmousi II model (sampling rate of 250 Hz), the extracted TF slice dimensions were set to 45 × 5. The frequency window spans 21–65 Hz ($f_{lgh} = 45$), while the time window covers 20 ms ($t_{lgh} = 5$). This parameter selection is based on statistical analysis of the reservoir response energy accumulation distribution, ensuring coverage of over 90% of the target layer's reflected energy while physically suppressing interference from low-frequency background noise and spurious high-frequency components.

The semi-supervised dataset comprises labeled training data, unlabeled training data, a validation set, and a test set. Based on prior geological information such as logging constraints, gas-bearing and gas-free slices were extracted from well-side traces to form the labeled training set $D_L = \left\{\left(x_i; y_i\right)\right\}_{i=1}^{N_L}$ (sample size $L_{synt}$), while samples distant from the well locations were selected to constitute the unlabeled training set $D_U = \left\{\left(x_j\right)\right\}_{j=1}^{N_U}$ (sample size $U$). The

labels (gas-bearing or gas-free) for 2D TF slices in the labeled samples are determined by the actual gas content at their center points.

Within the Marmousi II synthetic model, the elastic parameters of gas-bearing sandstone lenses are explicitly defined. Gas-bearing samples are defined as sampling points with a $V_p / V_s$ ratio below 1.5 that are located within known sandstone lens boundaries; gas-free samples are defined as sampling points with a $V_p / V_s$ ratio exceeding 1.5 and located outside known sandstone lens boundaries. To ensure model generalization capability and category balance, the ratio of gas-bearing to gas-free samples, as well as raw noise-free to noisy (SNR = 0 dB) samples, was maintained at 1:1. In the experiments, the baseline values for $L_{synt}$, $U_{real}$, and $T_{synt}$ were set to 40, 2,400, and 400, respectively.

To systematically evaluate the impact of labeled sample quantity on model performance, four traces with CDPs of 450, 475, 500, and 510 were selected as training wells. Samples corresponding to quantities of $1\times L_{synt}$, $2\times L_{synt}$, $3\times L_{synt}$, and $4\times L_{synt}$ were sequentially extracted from these wells to form the labeled training set (denoted as $1\times L_{synt}$ to $4\times L_{synt}$), while the unlabeled training and test sets retained their original sample sizes. The labeled and unlabeled training sets collectively constituted the model's training data. Additionally, approximately 10% of the total training samples were selected as the validation set during training, with the discriminator loss monitored in real time using an early stopping strategy.

The model was trained using a learning rate of $5 \times 10^{-4}$ with a batch size of 10, combined with an early stopping strategy (patience = 5) to ensure efficient convergence of the discriminator. To mitigate random fluctuations, all experiments were repeated using three random seeds, and the results were averaged (see Table 1).

Evaluation results demonstrate that, compared to purely supervised GANs, SSGAN effectively enhances the extraction of gas-bearing features by incorporating unlabeled samples into the learning process. Particularly when labeled samples are extremely scarce ($1\times L_{synt}$, i.e., only 40 labeled samples), SSGAN achieves accuracy and F1 scores that are 10.0% and 0.071 higher than those of SGAN, respectively. The performance of MR-SSGAN, which incorporates manifold-regularization constraints, was further improved, achieving additional gains of 4.7% in accuracy and 0.040 in F1 score over SSGAN.

Ablation experiments confirmed that manifold regularization significantly enhances the training stability and generalization of the discriminator under minimal sample sizes by reinforcing the model's predictive smoothness within data manifold neighborhoods. When the labeled data volume increased to $4\times L_{synt}$, MR-SSGAN achieved its optimal performance. Noise resistance analysis conducted under this configuration (Figure 6) demonstrates that the framework maintained high alignment with the ground truth model even under strong background noise at SNR = −5 dB.

To investigate the impact of feature representation quality on classification performance, this study systematically compared the performance differences of four TF transforms—CWT, STFT, WSST, and FSST—as model inputs under fixed data volume ($4\times L_{synt}$-labeled data) (Table 2). The findings revealed that, in the vast majority of cases, the MR-SSGAN model based on FSST input delivers optimal performance, with its predicted profiles (Figure 7) exhibiting the best alignment with the true gas reservoir morphology. This advantage stems primarily from FSST's superior frequency resolution and its more focused energy distribution within the TF domain. These characteristics facilitate the network's learning of clearer feature representations, thereby generating higher-resolution predictive outputs.

In contrast, while STFT also possesses good frequency resolution and its network can partially compensate for temporal resolution limitations through short-time window feature learning, its overall performance was inferior to FSST. CWT and WSST, however, fail to adequately reveal internal details due to energy concentration near reservoir edges in their TF representations (Figure 5), resulting in trained models struggling to accurately identify central reservoir zones. The ablation experiment demonstrates the critical importance of the deep coupling between high-resolution TF characterization and manifold-regularized SSL for enhancing prediction accuracy in gas-bearing reservoirs.

#### *3.1.3. SHapley Additive exPlanations analysis on synthetic data*

To gain deeper insight into how the trained MR-SSGAN model (based on $4\times L_{synt}$-labeled FSST data) makes gas content decisions from 2D TF inputs, this study employed SHAP analysis. SHAP values quantify the contribution of each specific frequency component and time sampling point within the TF data to the prediction outcome for an individual sample. Figure 8 illustrates the feature dependency patterns across both frequency and time dimensions when the model detects gas-bearing states on the Marmousi II model synthetic data, presented through feature importance bar charts and scatter plots.

Analysis revealed that the model assigns the highest

**Table 1. Performance comparison of different models using Fourier-based synchrosqueezing transform data on the Marmousi II model**

| Network architecture | Number of labeled samples | Performance | | Gain vs. SGAN | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| SGAN | $1 \times L_{synt}$ | 82.2 | 0.858 | - | - |
| | $2 \times L_{synt}$ | 86.3 | 0.892 | - | - |
| | $3 \times L_{synt}$ | 91.6 | 0.928 | - | - |
| | $4 \times L_{synt}$ | 95.0 | 0.955 | - | - |
| SSGAN | $1 \times L_{synt}$ | 92.2 | 0.929 | 10.0 | 0.071 |
| | $2 \times L_{synt}$ | 94.1 | 0.947 | 7.8 | 0.055 |
| | $3 \times L_{synt}$ | 96.3 | 0.963 | 4.7 | 0.035 |
| | $4 \times L_{synt}$ | 97.2 | 0.972 | 2.2 | 0.017 |
| MR-SSGAN | $1 \times L_{synt}$ | 96.9 | 0.969 | 14.7 | 0.111 |
| | $2 \times L_{synt}$ | 95.9 | 0.962 | 9.7 | 0.069 |
| | $3 \times L_{synt}$ | 97.5 | 0.976 | 5.9 | 0.048 |
| | $4 \times L_{synt}$ | 98.4 | 0.985 | 3.4 | 0.030 |

Abbreviations: MR: Manifold-regularized; SGAN: Supervised generative adversarial network; SSGAN: Semi-supervised generative adversarial network.

**Table 2. Performance comparison of different time–frequency data using MR-SSGAN on the Marmousi II model**

| Input | Number of labeled samples | Performance | | Loss vs. FSST | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| CWT | $1 \times L_{synt}$ | 95.9 | 0.960 | 0.9 | 0.009 |
| | $2 \times L_{synt}$ | 96.9 | 0.971 | −0.9 | −0.009 |
| | $3 \times L_{synt}$ | 95.9 | 0.962 | 1.6 | 0.014 |
| | $4 \times L_{synt}$ | 96.6 | 0.968 | 1.9 | 0.017 |
| STFT | $1 \times L_{synt}$ | 73.1 | 0.803 | 23.8 | 0.165 |
| | $2 \times L_{synt}$ | 96.3 | 0.964 | −0.3 | −0.003 |
| | $3 \times L_{synt}$ | 97.2 | 0.973 | 0.3 | 0.003 |
| | $4 \times L_{synt}$ | 97.2 | 0.973 | 1.3 | 0.012 |
| WSST | $1 \times L_{synt}$ | 80.6 | 0.817 | 16.3 | 0.151 |
| | $2 \times L_{synt}$ | 83.1 | 0.831 | 12.8 | 0.130 |
| | $3 \times L_{synt}$ | 84.1 | 0.841 | 13.4 | 0.135 |
| | $4 \times L_{synt}$ | 85.3 | 0.855 | 13.1 | 0.129 |
| FSST | $1 \times L_{synt}$ | 96.9 | 0.969 | - | - |
| | $2 \times L_{synt}$ | 95.9 | 0.962 | - | - |
| | $3 \times L_{synt}$ | 97.5 | 0.976 | - | - |
| | $4 \times L_{synt}$ | 98.4 | 0.985 | - | - |

Abbreviations: CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

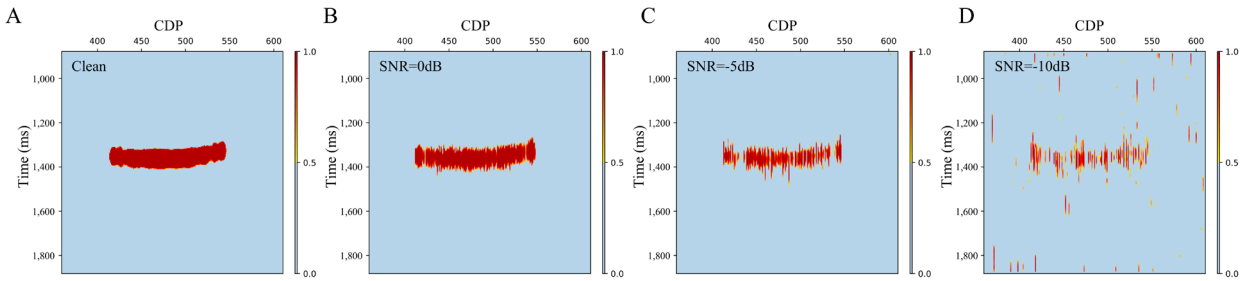


**Figure 6.** Prediction probability of gas-bearing reservoirs by the Fourier-based synchrosqueezing transform manifold-regularized semi-supervised generative adversarial network under varying SNR conditions: (A) clean, (B) 0 dB, (C) −5 dB, and (D) −10 dB
Abbreviations: CDP: Common depth point; SNR: Signal-to-noise ratio.

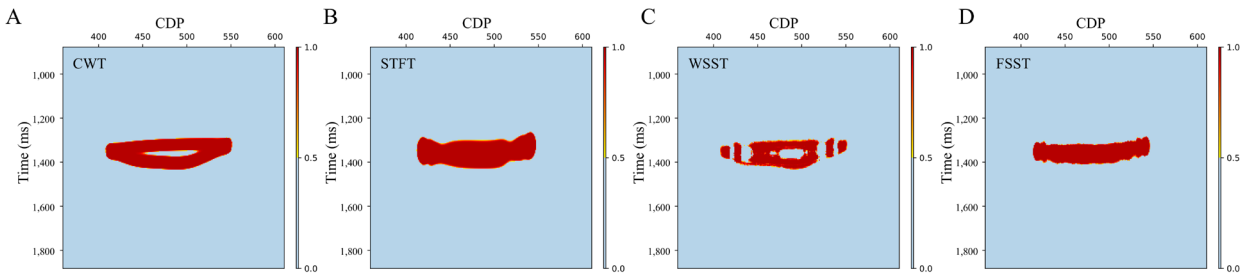


**Figure 7.** 45 Hz frequency slices of synthetic data: (A) CWT, (B) STFT, (C) WSST, and (D) FSST
Abbreviations: CDP: Common depth point; CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

decision weights to the mid-to-high frequency bands (e.g., 41–45 Hz, 46–50 Hz, 51–55 Hz, and 56–60 Hz). When energy in these bands is elevated, SHAP values predominantly cluster on the positive half-axis, strongly favoring gas-bearing predictions; conversely, they tend toward the negative half-axis, indicating gas-free. Notably, the Marmousi II model employs the Ormsby source wavelet (5–80 Hz), with 41–60 Hz falling precisely within its primary frequency response range. The pattern captured by the model, where enhanced energy indicates gas presence, likely stems from relative energy amplification in specific frequency bands within gas-bearing sandstone lenses. This amplification arises from thin-layer tuning effects or constructive wavefield interference. These results demonstrate that MR-SSGAN is not constrained by a single linear attenuation assumption but possesses the capability to uncover complex, non-intuitive seismic physical characteristics.

In the temporal dimension, analysis focused on the current time $T_0$ and five surrounding sampling points ($T_0 - 2$, $T_0 - 1$, $T_0$, $T_0 + 1$, and $T_0 + 2$). SHAP results revealed that the model focuses most on the energy characteristics at the three time points: $T_0$, $T_0 + 1$, and $T_0 + 2$. When $T_0$ and subsequent sampling points exhibit high energy values while earlier points (e.g., $T_0 - 2$) display relatively low energy, the model tends to classify the event as gas-bearing. This observed rearward shift in the energy centroid aligns with the instantaneous phase delay and energy envelope tailing characteristics induced by gas-bearing reservoirs, reflecting the physical process of kinetic energy dissipation as seismic waves propagate through fluid-saturated media.

The comprehensive SHAP analysis revealed that the MR-SSGAN establishes a logically rigorous decision criterion on the synthetic data. In the frequency domain, the model relies critically on energy resonance within specific sensitive frequency bands, while in the time domain, it identifies evolving patterns of energy distribution asymmetry. This process not only statistically validates the model's high accuracy but also geophysically confirms its decision logic aligns closely with tuning effects and wavefield attenuation patterns.

### 3.2. Results of real seismic data

To validate the practical applicability of the proposed framework, this study employed actual data from deep marine carbonate reservoirs of the Leikoupu Formation in the Triassic strata of the western Sichuan Basin, China. The target interval (T1–T2), buried at depths exceeding 5,000 m, exhibits strong heterogeneity, complex pore structures, and extremely thin interbeds due to multiple

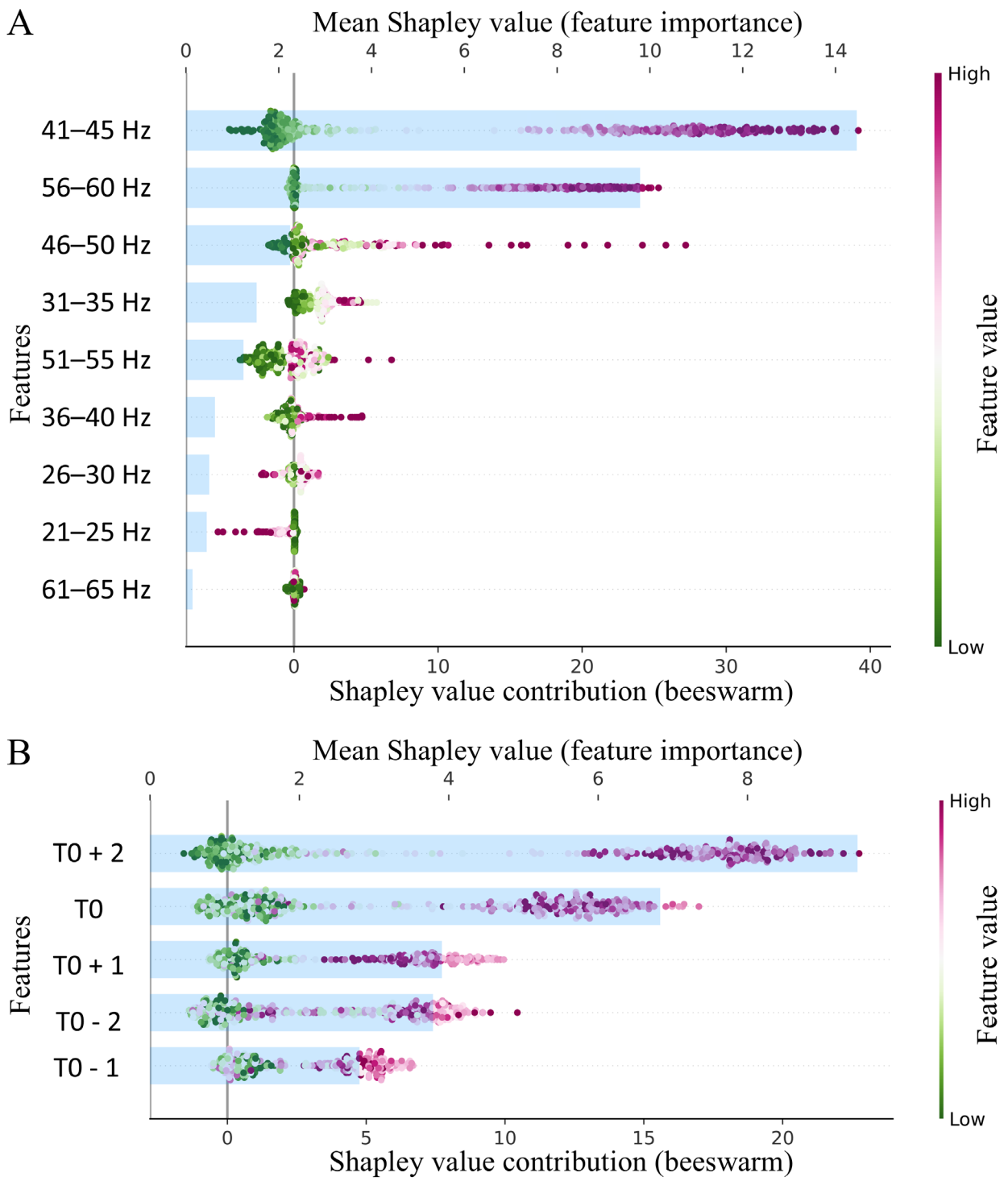


**Figure 8.** SHapley Additive exPlanations (SHAP) analysis of the manifold-regularized generative adversarial network on synthetic data. Each point represents a single feature from one sample, with purple and green corresponding to high and low feature values, respectively. The *x*-axis position denotes the SHAP value for that feature (positive values indicate promotion of gas-bearing predictions, whereas negative values indicate promotion of gas-free predictions). The bar lengths on the left intuitively reflect the average absolute SHAP value for each feature, indicating its importance. On the frequency dimension, the frequency axis of the TF plot is divided into bands at 5 Hz intervals.

episodes of tectonic deformation. This has resulted in severe absorption attenuation of seismic signals and an extremely low SNR environment. Furthermore, complex lateral facies transitions and uneven gas–water distribution pose significant challenges to gas content prediction due to non-uniqueness.

A representative 2D seismic profile (Figure 9) was selected, traversing three key calibration wells: Well A (gas-poor), Well B (gas-rich), and Well C (gas-bearing). From top to bottom, four main reflection horizons were identified within the profile: T1 (weathered crust at the top of the Leikoupo Formation), T2 (top of the fourth member of the Leikoupo Formation), T3 (top of the third member of the Leikoupo Formation), and T4 (bottom of the third member of the Leikoupo Formation). Intelligent detection focused specifically on gas content within the T1–T2 interval, aiming to validate the model's capability for detailed characterization of actual gas presence within complex geological settings.

#### 3.2.1. Feature response in real data

This study conducted TF characterization analysis on seismic records from the bypass Well B. The results (Figure 10)demonstrated high consistency with the synthetic data experiments: STFT remained constrained by the Heisenberg uncertainty principle, leading to TF energy dispersion, whereas FSST and WSST significantly enhanced feature concentration and resolution through their phase rearrangement mechanisms. Further comparison of the 15 Hz frequency slices (Figure 11) revealed that the energy clusters identified by traditional STFT and CWT in the T1–T2 interval were primarily distributed between Wells A and B, failing to effectively characterize the known gas-bearing zone near Well C.

Following application of SST, the WSST frequency slice (Figure 11C) revealed scarcely any pronounced zones of high energy concentration within the target stratum. While the FSST frequency slice (Figure 11D) demonstrates improved local energy contrast, it still lacked physical responses highly consistent with drilling conditions at the critical Well B and Well C intervals. This phenomenon clearly demonstrates that, for deep, strongly heterogeneous thin reservoirs in western Sichuan, conventional spectral decomposition properties remain constrained by seismic multi-solutions. Consequently, even with advanced synchrosqueezing operators, precise localization of gas-bearing sweet spots remains challenging.

#### 3.2.2. Accuracy and section evaluation

Given the complexity of seismic data from the field area, this study employed a 2D sliding TF window technique consistent with the synthetic experiments to construct samples, systematically evaluating the predictive performance of MR-SSGAN under sparse labeling conditions.

During feature extraction, FSST was applied to perform TF transformations on each raw seismic trace. Given the low dominant frequency (approximately 20 Hz) characteristic of deep reflection signals, the frequency truncation window was set to 1–45 Hz to ensure coverage of the core energy range spanning 5th–95th percentiles of most samples. The time window comprised the target sampling point flanked by two preceding and two following points. Given the 500 Hz sampling rate of the field data, the resulting feature slice dimensions remained 45 × 5, corresponding to a time-domain length of 10 ms.

To scientifically evaluate the model's generalization capability in strongly heterogeneous environments, a neighborhood $N_{W_i}(p)$ of $p$ units around well $W_i$ was defined in the coordinate system composed of xline and

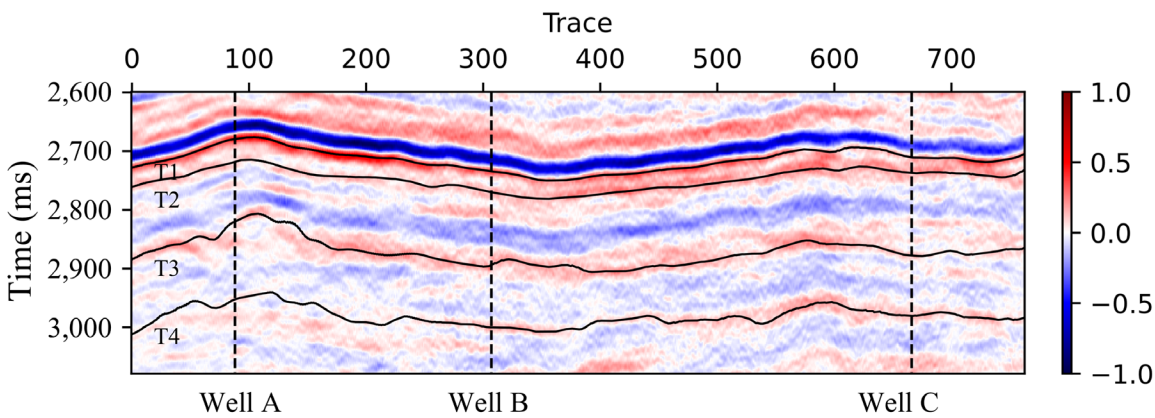


**Figure 9.** A post-stack seismic profile from the field area

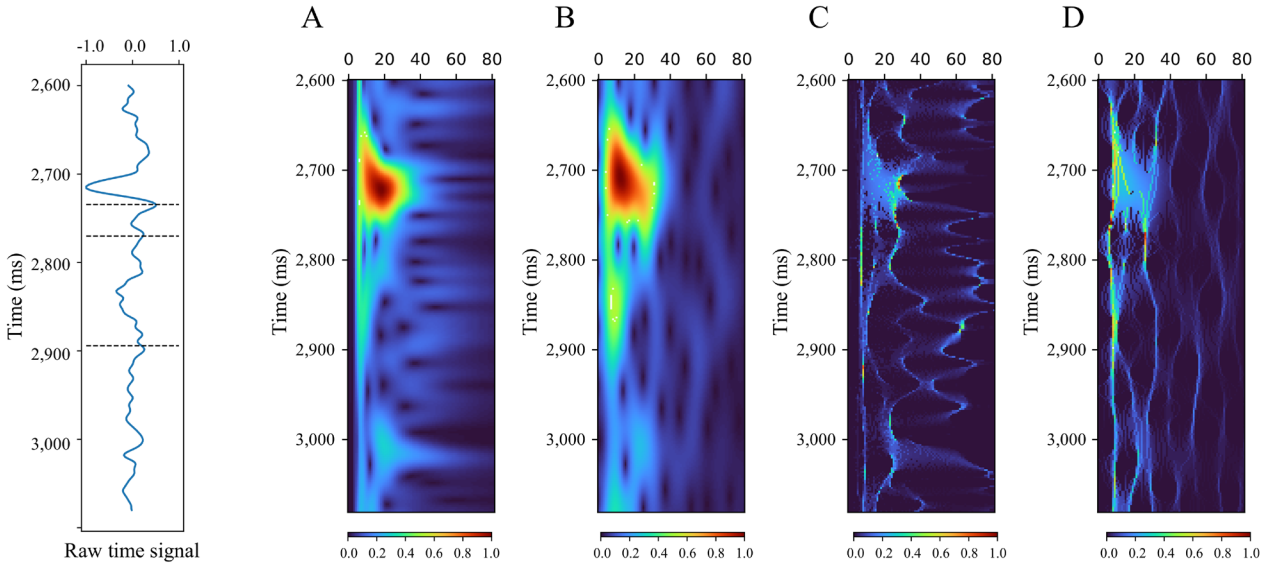


**Figure 10.** The original time-domain signal of the well-side trace at Well B and its corresponding time–frequency analysis results: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

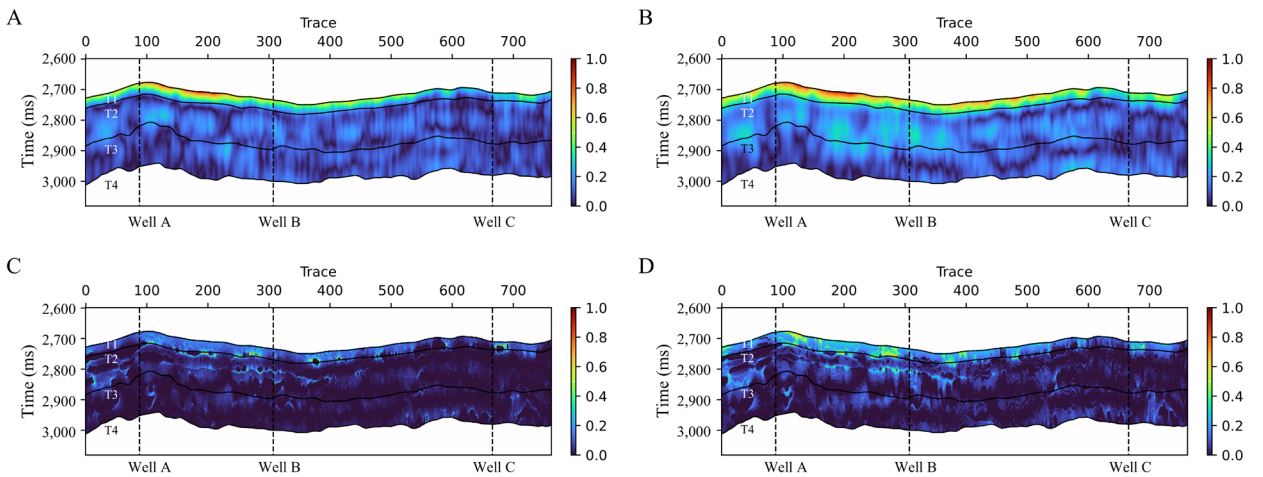


**Figure 11.** 15 Hz frequency slices of real data: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

inline:

$$\left\{(xl,il)\in R^2 \middle| p\le\sqrt{(xl-xl_w)^2-(il-il_w)^2}<p+1\right\} \tag{13}$$

where the well location $(xl_{Wi},il_{Wi})$ is represented by $N_{Wi}(0)$. Besides the well location, the neighborhood one unit away from the well is represented by $N_{Wi}(1)$, which includes eight seismic traces, namely $(xline\pm1,inline)$, $(xline\pm1,inline-1)$, $(xline\pm1,inline+1)$, and $(xline,inline\pm1)$. Wells B and C were used as training wells, with labeled samples selected from the well-side traces. Unlabeled training samples, validation samples, and test samples were derived from seismic traces within a neighborhood of three units outside the training well locations. The specific sample quantity allocation was as follows: labeled training samples $L_{real}=40$, unlabeled training samples $U_{real}=800$, test samples $T_{real}=40$, and the number of validation set samples was approximately 10% of the total training data. The learning rate, batch size, and early stopping strategy (patience = 20) for model training were generally consistent with the synthetic data experiments.

Table 3 presents a comparative analysis of different

model architectures on real seismic datasets (averaged across three randomized experiments). Under conditions of extremely sparse labeling, the semi-supervised strategy demonstrates significant advantages: SSGAN achieves an 18.8% improvement in accuracy (with an F1 score increase of 0.200) over the purely supervised GAN, robustly demonstrating the auxiliary role of unlabeled data. Building upon this, the introduction of manifold regularization in MR-SSGAN yields an additional 5.0% gain in accuracy. This phenomenon aligns closely with the trend observed in the synthetic experiments, further validating that manifold constraints effectively suppress prediction fluctuations caused by noise in the field area by enhancing model smoothness on the data manifold.

To investigate the coupled effects of feature quality and semi-supervised architecture, the classification performance was compared using four TF representations as input (see Table 4). Quantitative results confirm that the MR-SSGAN model trained with FSST significantly outperforms other approaches in both accuracy and F1 score. Prediction accuracy based on STFT, CWT, and WSST exhibits a substantial decline, primarily due to spectral leakage or spatiotemporal blurring during the processing of deep, strongly inhomogeneous seismic signals.

To systematically evaluate the impact of different TF inputs on model predictive performance, Figure 12 presents the predicted probability profiles of MR-SSGAN trained using four TF representations.

First, predictions based on CWT (Figure 12A) and STFT (Figure 12B) exhibit significant deviations from known geological interpretations. Although CWT possesses multiscale analysis capabilities, its energy distribution becomes excessively dispersed when applied to real data with low SNR. This resulted in missed detections within the critical gas-bearing zones of Wells B and C, while simultaneously generating false continuous high-probability bands in the gas-depleted zone of Well A. The STFT, constrained by the inherent resolution limitation due to fixed window lengths, produces blurred boundaries for predicted gas-bearing zones and generates large-scale artefactual anomalies within non-target intervals (T2–T3 sections). These observations indicate that traditional TF transforms were unable to effectively mitigate TF blurring caused by the Heisenberg uncertainty principle, rendering them incapable of providing robust discriminative features for models under complex background noise.

Second, although synchrosqueezing technology aims to enhance energy concentration, the WSST-based model (Figure 12C) exhibits suboptimal performance in practical applications. The predicted profile exhibits severe fragmentation, poor lateral continuity, and significant misclassification in the middle–lower section. This is because, when processing complex wavefields, the energy reassignment operator of WSST is highly sensitive to strong noise interference, thereby compromising the physical consistency of the input features. This further indicates that, within the strongly heterogeneous conditions of the field area, an algorithm's robustness to noise is equally critical to feature aggregation.

**Table 3. Performance comparison of different models based on Fourier-based synchrosqueezing transform data (real data)**

| Network architecture | Number of labeled samples | Performance | | Gain vs. SGAN | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| SGAN | $1 \times L_{real}$ | 61.3 | 0.613 | - | - |
| SSGAN | $1 \times L_{real}$ | 80.0 | 0.813 | 18.8 | 0.200 |
| MR-SSGAN | $1 \times L_{real}$ | 85.0 | 0.885 | 23.8 | 0.272 |

Abbreviations: MR: Manifold-regularized; SGAN: Supervised generative adversarial network; SSGAN: Semi-supervised generative adversarial network.

**Table 4. Performance comparison of different time–frequency data based on MR-SSGAN (real data)**

| Input | Number of labeled samples | Performance | | Loss vs. FSST | |
|---|---|---|---|---|---|
| | | Accuracy (%) | F1-score | Accuracy (%) | F1-score |
| CWT | $1 \times L_{real}$ | 75.0 | 0.767 | 10.0 | 0.118 |
| STFT | $1 \times L_{real}$ | 77.9 | 0.784 | 7.1 | 0.100 |
| WSST | $1 \times L_{real}$ | 70.6 | 0.707 | 14.4 | 0.177 |
| FSST | $1 \times L_{real}$ | 85.0 | 0.885 | - | - |

Abbreviations: CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; SGAN: Supervised generative adversarial network; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

Third, the FSST-based MR-SSGAN architecture (Figure 12D), in contrast, exhibits superior predictive performance, with results highly consistent with drilling calibration. This model not only accurately captured the gas anomalies in the target intervals of Wells B and C, presenting clear and laterally coherent high-probability responses, but also precisely depicted the gas-depleted conditions near Well A. The predicted reservoir morphology is well-defined, with thickness characterization better aligned with the geological development features of thin interbedded formations, demonstrating exceptionally high vertical and spatial resolution.

In summary, the experimental results further validate the synergistic effect between FSST and MR-SSGAN. FSST provides a highly focused and more noise-resistant feature benchmark through energy reassignment, while MR-SSGAN establishes a stable smoothing constraint in the feature space via manifold regularization. The deep integration of this high-quality input and robust learning paradigm enables the architecture to effectively overcome seismic ambiguity. Even under the constraint of extremely sparse labels, it achieves high-reliability detection of gas-bearing sweet spots in deep thin reservoirs in western Sichuan, providing effective technical support for the intelligent evaluation of complex hydrocarbon reservoirs.

### *3.2.3. SHapley Additive exPlanations analysis on real data*

SHapley Additive exPlanations analysis applied to field seismic data (Figure 13) further elucidates the decision logic of MR-SSGAN within complex geological environments, revealing attribution patterns markedly distinct from those observed in synthetic data.

Along the frequency dimension (Figure 13A), SHAP attribution results maintain high consistency with classical gas-bearing response theory. The feature importance bar chart indicates that the model primarily relies on information within the 6–30 Hz frequency band for decision-making, with the 6–10 Hz low-frequency range contributing most significantly. The scatter plot further reveals that energy enhancement (high feature values) in this low-frequency band is positively correlated with gas-bearing discrimination, accurately capturing the typical low-frequency shadow feature. In contrast, within the 11–30 Hz mid-to-high-frequency range, energy attenuation (low feature values) contributes positive Shapley values to gas-bearing prediction. This combined signature of low-frequency resonance and high-frequency absorption physically corroborates the expected dynamic evolution of seismic waves as they propagate through gas-bearing porous media. Compared to the mid-to-high-frequency

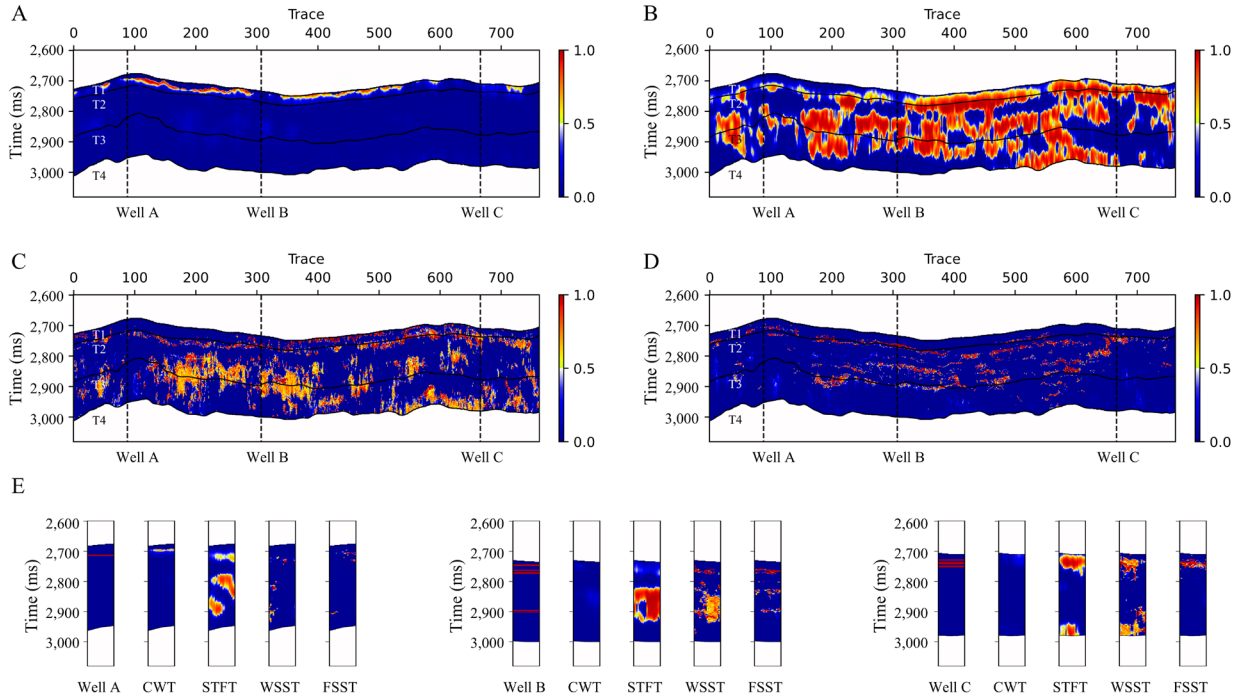


**Figure 12.** Gas-bearing reservoir prediction probability of the manifold-regularized generative adversarial network based on different time–frequency data (input is real seismic data): (A) CWT, (B) STFT, (Cc) WSST, and (D) FSST. (E) Comparison of labels versus predictions across time–frequency inputs. Abbreviations: CWT: Continuous wavelet transform; FSST: Fourier-based synchrosqueezing transform; SGAN: Supervised generative adversarial network; STFT: Short-time Fourier transform; WSST: Wavelet synchrosqueezing transform.

enhancement pattern dominated by thin-layer tuning effects in synthetic data, the analysis results from real data align more closely with classical wavefield attenuation theory, reflecting the intelligent discriminative adaptation of the deep learning framework when addressing idealized models versus complex real-world environments.

In the temporal dimension (Figure 13B), the model exhibits sampling point dependency patterns distinct from synthetic data. Analysis indicates that the present time $T_0$ carries the highest importance weight, followed by $T_0 - 2$ and $T_0 + 2$. The model identifies a specific temporal evolution pattern: samples exhibiting localized energy enhancement at $T_0 - 2$ followed by energy attenuation at $T_0$ and $T_0 + 2$ are more readily identified as gas-bearing. This complex response pattern likely originates from the intricate wavefield interference and phase delay phenomena within deep, thin interbedded reservoirs. Compared to the simple energy center shift observed in synthetic data, the SHAP attribution patterns in real data reveal how actual stratigraphic absorption characteristics modulate the waveform envelope morphology.

The comprehensive SHAP attribution analysis demonstrates that the MR-SSGAN model not only achieves statistically high-precision predictions but also captures seismically plausible phenomena. The model adaptively adjusts its decision logic based on data context—transitioning from tuning-effect–driven behavior in synthetic data to absorption-attenuation–driven behavior in real data. This cross-dataset consistency in decision-making, coupled with transparent physical logic, effectively rules out the possibility of model classification arising from overfitting noise. Consequently, it significantly enhances the reliability and guidance value of artificial intelligence technology in high-risk decision-making within oil and gas exploration.

## 4. Discussion

### 4.1 Selection and parameter impact of synchrosqueezing transform

During the feature characterization phase, experimental results (Tables 2 and 4) show that FSST produces downstream models with more discriminative input features than WSST. Although the wavelet-based WSST has advantages in multiscale analysis, its energy rearrangement operator is sensitive to random noise when processing seismic signals with a low SNR. This results in fractured TF ridges and energy artifacts. In contrast, FSST integrates the DPSS window function, which has greater sidelobe suppression capabilities, resulting in a relatively robust analysis framework. This design successfully mitigates random disturbances during energy compression, guaranteeing that extracted gas-bearing anomaly responses are statistically stable. While this study provides initial validation of the chosen window parameters, future work should involve comprehensive sensitivity analyses to develop tailored optimization strategies for SST, targeting specific applications such as ultra-deep or tight gas reservoirs. This will further advance the capabilities of high-resolution TF characterization.

### 4.2. Critical role of time–frequency coupled features

Seismic signals are inherently non-stationary, complex wave processes. Traditional 1D attributes (e.g., instantaneous frequency or amplitude) represent time and frequency information separately, resulting in the loss of rich context-coupled characteristics. The comparative experiment (Figure 14) clearly illustrates this limitation: model predictions based solely on 1D instantaneous frequency exhibit pronounced fragmentation, lack geological lateral continuity, and struggle to characterize the fine boundaries of thin reservoirs. In contrast, 2D TF slices simultaneously capture complete spectral evolution patterns at specific moments and within their adjacent temporal windows. SHAP analysis (Figure 13B) further confirms, from an attribution perspective, that the model's gas-bearing classification heavily depends on the energy distribution ratio within the narrow time window from $T_0 - 2$ to $T_0 + 2$. This cross-dimensional spatiotemporal coupling characteristic cannot be expressed by 1D attributes and constitutes the key physical prerequisite for achieving high-precision identification of complex thin reservoirs.

### 4.3. Comparison of learning paradigms

When addressing the challenge of sparse labeling, the proposed MR-SSGAN demonstrates greater endogeneity and robustness compared to data augmentation, which depends on manually generated samples[31,32] or transfer learning[11,33,34], which relies on external source domain knowledge. Although the data augmentation strategy (Figure 15A) enhances coverage, its predicted profiles are often accompanied by false positive indications of questionable geological plausibility and are highly dependent on the quality of generated pseudo-samples. Transfer learning (Figure 15B) focuses on sweet-spot regions, yet suffers from weak continuity representation in inter-well areas due to domain differences between source and target domains.

In contrast, MR-SSGAN achieves end-to-end knowledge extraction by directly mining the distributional manifold of vast unlabeled data within the target field area. Its predictions strike a favorable balance between smoothness in macro-trends and focus on local details (Figure 12D).

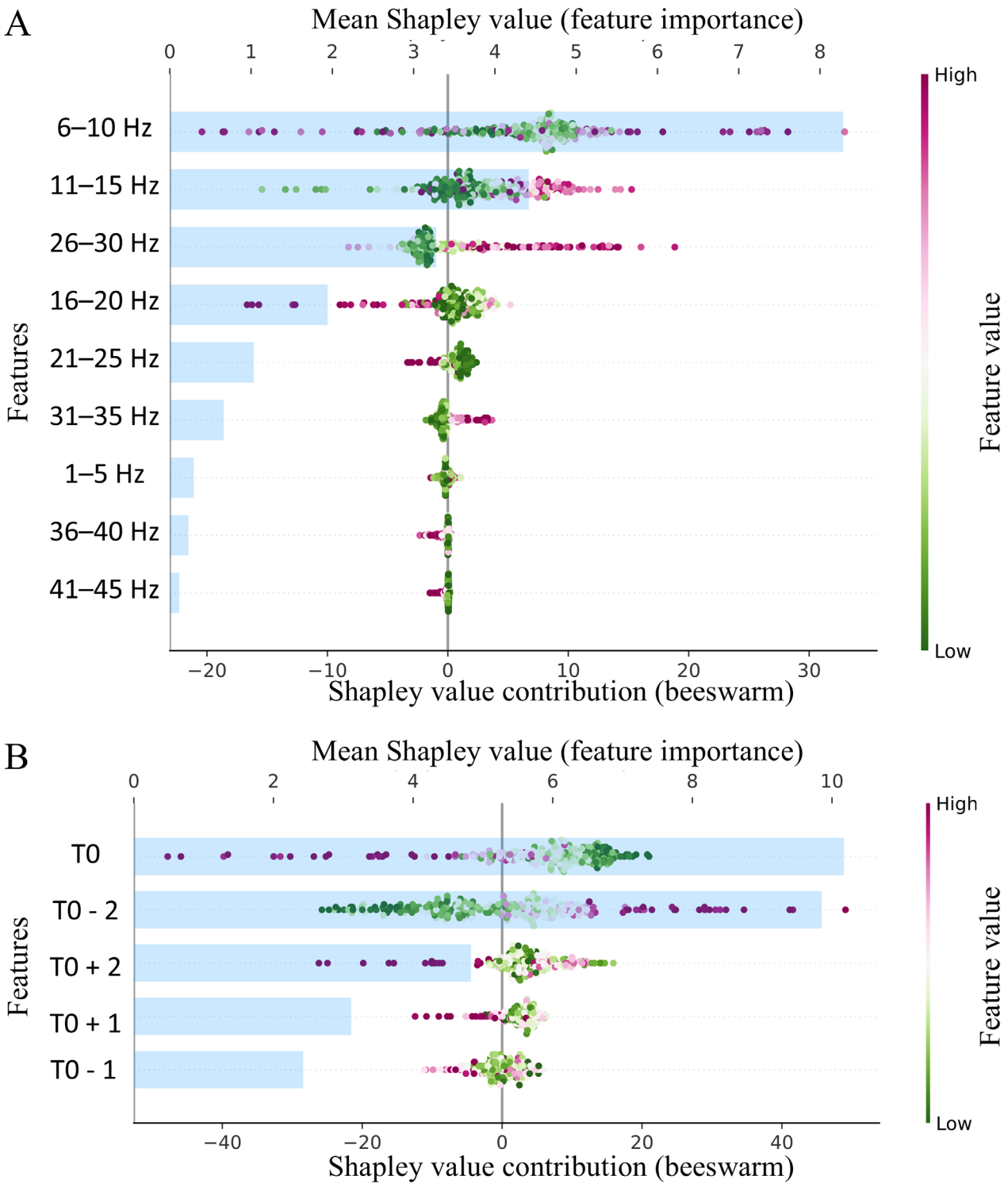


**Figure 13.** SHapley Additive exPlanations analysis of the manifold-regularized generative adversarial network on real seismic data

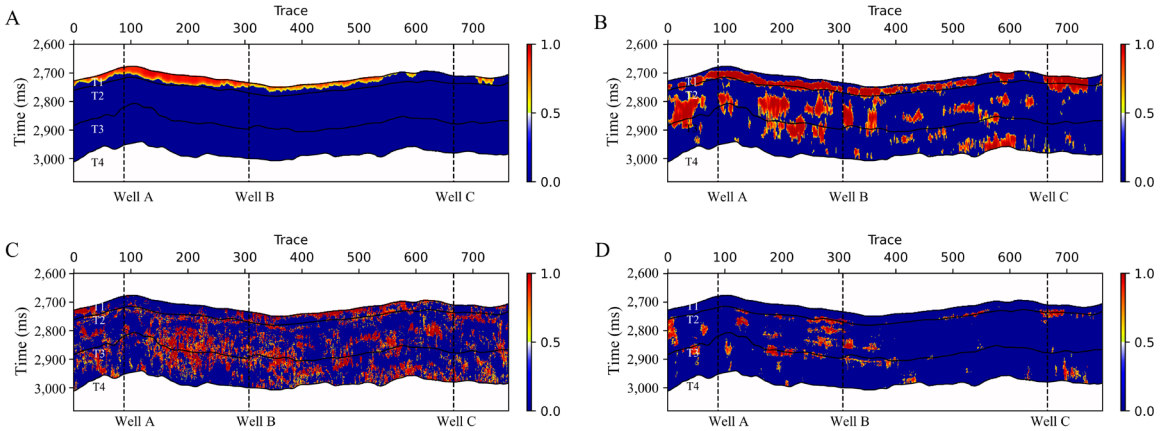


**Figure 14.** Gas-bearing reservoir prediction probability of the manifold-regularized generative adversarial network based on different one-dimensional instantaneous frequency inputs from real seismic data: (A) continuous wavelet transform, (B) short-time Fourier transform, (C) wavelet synchrosqueezing transform, and (D) Fourier-based synchrosqueezing transform

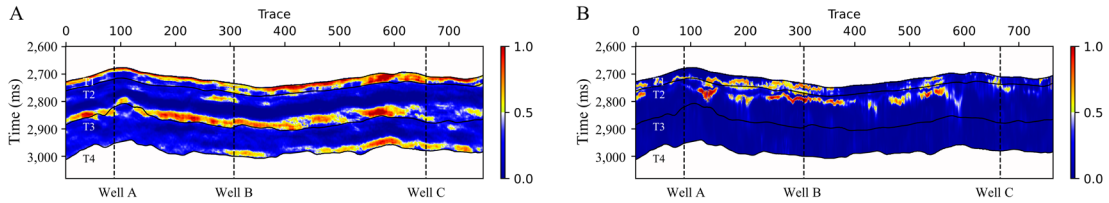


**Figure 15.** Gas-bearing reservoir prediction probability using (A) a dual-branch spatio-temporal fusion model with data augmentation and (B) an efficient transformer model with transfer learning

This semi-supervised paradigm, which directly acquires prior information from the target domain's topological geometric features, not only reduces reliance on external training datasets but also enhances the algorithm's generalizability in highly heterogeneous environments. Future exploration of embedding small quantities of high-quality augmented data into the semi-supervised training process holds promise for further improving the intelligent evaluation efficiency of complex reservoirs.

## 5. Conclusion

This study proposes an intelligent detection framework that incorporates the STFT-based FSST, MR-SSGAN, and SHAP analysis to address core challenges in gas content identification within complex reservoirs, namely scarce labeled samples, significant environmental noise interference, and opaque model decision-making. This paradigm shows strong prediction advantages in both synthetic models and real seismic data. Under extreme label scarcity (5% labeled samples), the SSL technique and manifold-regularization constraint improved accuracy by 18.8% and 5.0%, respectively (with simultaneous significant gains in F1 scores). Ablation experiments show that 2D TF coupled information better describes transient dynamic anomaly responses in thin reservoirs than 1D transient features. Furthermore, unlike transfer learning, which relies on external previous information or data augmentation using artificial rules, the proposed semi-supervised architecture recovers latent distributional patterns directly from real seismic data. This demonstrates higher generalization capacity when dealing with low SNR settings, such as deep marine reservoirs in the Sichuan Basin. Further SHAP analysis demonstrates that the MR-SSGAN model adapts its decision logic based on data context, shifting from tuning-effect–driven in synthetic data to absorption-attenuation–driven in real data. This capability not only verifies the model's ability to capture true physical principles but also aligns its decision-making process with conventional rock physics theories. This considerably improves the dependability of data-driven strategies in high-risk decision-making for oil and gas exploration.

Looking ahead, research may deepen in two directions: first, introducing automatic optimization algorithms—

such as particle swarm optimization or Bayesian optimization—for adaptive tuning of key parameters to further enhance the framework's generalizability; second, integrating advanced architectures like Transformers with multi-source physical information to uncover deeper sensitivity patterns, ultimately constructing a more precise and reliable intelligent evaluation system for hydrocarbon exploration.

## Acknowledgments

None.

## Funding

This work was supported by the National Natural Science Foundation of China (Grant 42330813 and Grant 42030812), and the Natural Science Foundation of Sichuan Province (Grant 2026NSFSC1138).

## Conflict of interest

The authors declare they have no competing interests.

## Author contributions

*Conceptualization:* Shuying Ma, Junxing Cao, Rong Wang
*Formal analysis:* Jun Wang, Lingsen Zhao, Hong Li, Xin Tang
*Investigation:* Junxing Cao, Xudong Jiang
*Methodology:* Shuying Ma
*Writing–original draft:* Shuying Ma
*Writing–review & editing:* Shuying Ma, Junxing Cao, Rong Wang

## Availability of data

Data supporting the findings of this study are available from the corresponding authors upon reasonable request.

## References

1. Hammond AL. Bright Spot: Better Seismological Indicators of Gas and Oil. *Science*. 1974;185(4150):515-517.

   doi: 10.1126/science.185.4150.515

2. Brown AR. Dim spots: Opportunity for future hydrocarbon discoveries? *Lead Edge*. 2012;31(6):682-683.

   doi: 10.1190/tle31060682.1

3. Backus MM, Chen RL. Flat spot exploration. *Geophys Prospect*. 1975;23(3):533-577.

   doi: 10.1111/j.1365-2478.1975.tb01547.x

4. Korneev VA, Goloshubin GM, Daley TM, Silin DB. Seismic low-frequency effects in monitoring fluid-saturated reservoirs. *Geophysics*. 2004;69(2):522-532.

   doi: 10.1190/1.1707072

5. Silin DB, Korneev VA, Goloshubin GM, Patzek TW. *A Hydrologic View on Biot's Theory of Poroelasticity*. Office of Scientific and Technical Information (OSTI); 2004.

   doi: 10.2172/822181

6. Mitchell JT, Derzhi N, Lichman E, Lanning EN. Energy absorption analysis: A case study. In: *SEG Technical Program Expanded Abstracts 1996*. 66th SEG Annual International Meeting; January 1, 1996; Denver, Colorado, USA. ASME; 1996:1785-1788.

   doi: 10.1190/1.1826480

7. Santos LF, Silva RMGE, Gattass M, Silva AC. Direct hydrocarbon indicators based on long short-term memory neural network. In: *SEG Technical Program Expanded Abstracts 2019*. 89th SEG Annual International Meeting; August 1, 2019; San Antonio, Texas, USA. ASME; 2019:2373-2377.

   doi: 10.1190/segam2019-3215628.1

8. Júnior DAD, Batista da Cruz L, Otávio Bandeira Diniz J, *et al*. Detection of potential gas accumulations in 2D seismic images using spatio-temporal, PSO, and convolutional LSTM approaches. *Expert Syst Appl*. 2023;215:119337.

   doi: 10.1016/j.eswa.2022.119337

9. Ma S, Cao J, Liu Z, Jiang X, Su Z, Xue Y. Gas-bearing prediction of deep reservoir based on DNN embeddings. *Front Earth Sci*. 2023;11.

   doi: 10.3389/feart.2023.1117797

10. Jiang X, Cao J, Zu S, Xu H, Wang J. Detection of hidden reservoirs under strong shielding based on bi-dimensional empirical mode decomposition and the Teager–Kaiser operator. *Geophys Prospect*. 2021;69(5):1086-1101.

    doi: 10.1111/1365-2478.13073

11. Gao J, Song Z, Gui J, Yuan S. Gas-Bearing Prediction Using Transfer Learning and CNNs: An Application to a Deep Tight Dolomite Reservoir. *IEEE Geosci Remote Sensing Lett*. 2022;19:1-5.

    doi: 10.1109/lgrs.2020.3035568

12. Yang J, Lin N, Zhang K, Zhang D, Wang D, Zhang J. Variable Selection and Oil-Gas-Bearing Evaluation Using Multicomponent Seismic Data and AMPSO-ML Approaches. *SPE J*. 2023;28(03):1183-1201.

    doi: 10.2118/214298-pa

13. Yang JQ, Lin NT, Zhang K, Cui Y, Fu C, Zhang D. Deep learning CNN-APSO-LSSVM hybrid fusion model for feature optimization and gas-bearing prediction. *Pet Sci*. 2024;21(4):2329-2344.

    doi: 10.1016/j.petsci.2024.02.012

14. Yang X, Song Z, King I, Xu Z. A Survey on Deep Semi-Supervised Learning. *IEEE Trans Knowl Data Eng*. 2023;35(9):8934-8954.

doi: 10.1109/tkde.2022.3220219

15. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, *et al*. Generative adversarial networks. *Commun ACM*. 2020;63(11):139-144.

doi: 10.1145/3422622

16. Odena A. Semi-supervised learning with generative adversarial networks. *arXiv*. Preprint online 2016.

doi: 10.48550/arXiv.1606.01583

17. Liu M, Jervis M, Li W, Nivlet P. Seismic facies classification using supervised convolutional neural networks and semisupervised generative adversarial networks. *Geophysics*. 2020;85(4):O47-O58.

doi: 10.1190/geo2019-0627.1

18. Song L, Yin X, Zhang R, Li J, Zhang J, Li J. Reservoir Discrimination Based on Physic-Informed Semi-Supervised Learning. *IEEE Trans Geosci Remote Sens*. 2024;62:1-14.

doi: 10.1109/tgrs.2024.3409578

19. Song L, Yin X, Yin L. Reservoir Lithology Identification Based on Improved Adversarial Learning. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

doi: 10.1109/lgrs.2023.3281545

20. Daubechies I, Lu J, Wu HT. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl Comput Harmon Anal*. 2011;30(2):243-261.

doi: 10.1016/j.acha.2010.08.002

21. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Guyon I, Von Luxburg U, *et al*., editors. *Advances in Neural Information Processing Systems 30*. 31st International Conference on Neural Information Processing Systems (NIPS 2017); December 4–9, 2017; Long Beach, CA, USA. Curran Associates Inc; 2017:4766-4777.

22. Huang Z lai, Zhang J, Zhao T hu, Sun Y. Synchrosqueezing S-Transform and Its Application in Seismic Spectral Decomposition. *IEEE Trans Geosci Remote Sens*. 2016;54(2):817-825.

doi: 10.1109/tgrs.2015.2466660

23. Mao X. A Concentrated Time-Frequency Method for Reservoir Detection Using Adaptive Synchrosqueezing Transform. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/lgrs.2022.3160930

24. Belkin M, Niyogi P, Sindhwani V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. J Mach Learn Res. 2006;7(85):2399–2434.

25. Ni Y, Koniusz P, Hartley R, Nock R. Manifold Learning Benefits GANs. In: *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). June 19–24, 2022; New Orleans, Louisiana. IEEE; 2022:11255-11264.

doi: 10.1109/cvpr52688.2022.01098

26. Lecouat B, Foo CS, Zenati H, Chandrasekhar VR. Semi-Supervised Learning with GANs: Revisiting Manifold Regularization. *arXiv*. Preprint posted online May 23, 2018. Last accessed November 4, 2024.

doi: 10.48550/arXiv.1805.08957

27. Oberlin T, Meignen S, Perrier V. The fourier-based synchrosqueezing transform. In: *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*). May 4–9, 2014; Florence, Italy. IEEE; 2014:315-319.

doi: 10.1109/icassp.2014.6853609

28. Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X. Improved Techniques for Training GANs. *arXiv*. Preprint posted online June 10, 2016. Last accessed November 1, 2024.

doi: 10.48550/arXiv.1606.03498

29. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: *Proceedings of the 32nd International Conference on Machine Learning*. International Conference on Machine Learning; July 7-9 2015; Lille, France. PMLR; 2015;37:448-456.

30. Martin GS, Marfurt KJ, Larsen S. Marmousi-2: An updated model for the investigation of AVO in structurally complex areas. In: *SEG Technical Program Expanded Abstracts 2002*. 72th SEG Annual International Meeting; January 1, 2002; Salt Lake City, Utah, USA. ASME; 2002:1979-1982.

doi: 10.1190/1.1817083

31. Ma S, Cao J. A Two-Branch Neural Network for Gas-Bearing Prediction Using Latent Space Adaptation for Data Augmentation—An Application for Deep Carbonate Reservoirs. *IEEE Geosci Remote Sens Lett*. 2024;21:1-5.

doi: 10.1109/lgrs.2024.3436836

32. Zhang K, Lin N, Yang J, Zhang D, Cui Y, Jin Z. An Intelligent Approach for Gas Reservoir Identification and Structural Evaluation by ANN and Viterbi Algorithm—A Case Study From the Xujiahe Formation, Western Sichuan Depression, China. *IEEE Trans Geosci Remote Sens*. 2023;61:1-12.

doi: 10.1109/tgrs.2023.3247183

33. Song Z, Li S, He S, Yuan S, Wang S. Gas-Bearing Prediction of Tight Sandstone Reservoir Using Semi-Supervised Learning and Transfer Learning. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/lgrs.2022.3177314

34. Ma S, Cao J. An Efficient Transformer Model Enhanced by S-Transform and Transfer Learning for Predicting Gas Distribution in Deeply Buried Reservoirs. *IEEE Trans Geosci Remote Sens*. 2025;63:1-14.

doi: 10.1109/tgrs.2025.3558731