



Prestack simultaneous inversion of P-wave impedance and gas saturation using multi-task residual networks

Wenjing Sang^{1,2,3,4} · Zhiqiang Ding⁴ · Mingxuan Li⁴ · Xiwu Liu^{1,2,3} · Qian Liu^{1,2,3} · Sanyi Yuan^{1,2,3,4}

Received: 7 February 2023 / Accepted: 13 June 2023 / Published online: 31 July 2023

© The Author(s) under exclusive licence to Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2023

Abstract

It is challenging to realize the gas saturation (GS) estimation via the integration of seismic data and more complementary data (e.g., elastic attributes) in a traditional physical framework. Machine learning, especially multi-task learning (MTL), provides an alternative way for the fuse of multiple information and simultaneous inversion of two or multiple reservoir parameters without model-driven limitations and interactive operators. To improve the estimation accuracy of GS, we propose the prestack simultaneous inversion of P-wave impedance (PI) and GS using the multi-task residual network (MT-ResNet). The designed MT-ResNet consists of two task-related subnets. The first subnet establishes the nonlinear links among low-frequency PI, prestack seismic data, and well-log derived PI. Furthermore, seismic data and the inverted PI via the first subnet are jointly entered into the second subnet and evolved into the well-log interpreted GS. A model based on measured petrophysical parameters associated with the field deep tight dolomite reservoir is used to test the proposed method. Tests on the synthetic data example and the field example demonstrate that the MT-ResNet can simultaneously estimate PI and GS models with the highest reliability, in comparison with single-task residual network (ST-ResNet) and the conventional seismic inversion and rock-physics equations-based method. And the MT-ResNet inverted PI can be utilized as complementary information for improving the prediction accuracy of MT-ResNet inverted GS. Our proposed MT-ResNet has the potential to guide the design of the MTL-based multiple reservoir parameters prediction and practical application.

Keywords Gas saturation prediction · P-wave impedance inversion · Multi-task learning · Machine learning · Prestack simultaneous inversion

Introduction

Reservoir characterization aims to delineate subsurface rock and fluid properties (e.g., lithology, porosity, permeability, and hydrocarbon saturation) based on available geophysical and petrophysical data (Eidsvik et al. 2004; Bosch et al. 2010; Liu et al. 2022). In this paper, we focus on the precise assessment of gas saturation (GS), which is defined as the fraction of the total gas volume over the volume of connected pores in the underground rock (Radke and Gillis 1990). As one of reservoir parameters, GS is highly critical for uncovering the potential favorable gas distribution, assessing gas reserves, designing well locations, and optimizing hydraulic fracture stimulation and completion (Singh et al. 2009; Lucier et al. 2011; Rezaee 2015).

GS determination contains well-log interpretation at the borehole scale and seismic inversion at the spatial scale. It is challenging to directly measure GS via conventional well-logging tools; therefore, the construction

Edited by Dr. Qamar Yasin (ASSOCIATE EDITOR) /
Prof. Gabriela Fernández Viejo (CO-EDITOR-IN-CHIEF).

✉ Sanyi Yuan
yuansy@cup.edu.cn

- ¹ State Key Laboratory of Shale Oil and Gas Enrichment Mechanisms and Effective Development, Beijing 100083, China
- ² Sinopec Key Laboratory of Seismic Elastic Wave Technology, Beijing 100083, China
- ³ Sinopec Petroleum Exploration and Production Research Institute, Beijing 100083, China
- ⁴ College of Geophysics, China University of Petroleum, Beijing, Beijing 102249, China

of GS curves is involved with rock-physics analysis and calibrated with laboratory measurements (Lucier et al. 2011; Ahmad and Haghghi 2013; Qi et al. 2017). This type of methodology firstly constitutes empirical calculation models (e.g., total shale model) between site-specific well logs and water saturation measured from core data (Schlumberger 1989; Yu et al. 2022). And GS is eventually obtained by subtracting water saturation from unity under the hypothesis of two-phase mediums (Qi et al. 2021). For instance, the Archie formula utilized porosity and electrical resistivity logs to evaluate GS of a pure sandstone formation with medium-to-large-size pores (Archie 1942). Qi et al. (2017) investigated compressional (P) to shear (S)-wave attenuation ratio (Q_p^{-1}/Q_s^{-1}) estimated from full-waveform sonic logs as a discriminate tool for shale GS with the criterion of $Q_p^{-1}/Q_s^{-1} < 1$. They showed that the Q_p^{-1}/Q_s^{-1} has a positive correlation with GS and is more sensitive than the P- to S-wave velocity ratio (V_p/V_s) as a quantitative gas indicator in the studied shale gas formation (Qi et al. 2021). Most current GS models depend on calibration to in situ rocks and are time-consuming, expensive, and spatially limited (Morgan et al. 2012). In addition, they have poor GS computation effect for some types of reservoirs characterized by low resistivity or high total organic carbon content (Xu et al. 2017; Qi et al. 2021; Yu et al. 2022).

Well-determined GS logs illustrate the gas distribution around the borehole, geophysicists further integrate seismic, well log, and geology data (e.g., interpreted target horizons) to predict the spatial variability of GS. Since seismic data do not directly reflect the variation of GS in different underground rock or stratum, GS or hydrocarbon-bearing prediction is historically implemented by two successive steps including seismic inversion and petrophysical modeling (Bosch et al. 2007; Aleardi et al. 2018). First, elastic parameters are inverted from seismic-reflection data via poststack impedance inversion, amplitude variation with offset/angle (AVO/AVA) inversion, and other prestack simultaneous inversion (Connolly 1999; Mazzotti and Zamboni 2003; Ma et al. 2023). These sensitive properties include P-wave impedance (PI), P-wave velocity (V_p), S-wave velocity (V_s), density, lambda-rho/mu-rho, and so on (Goodway et al. 1997). Subsequently, these inversion models and petrophysical equations are utilized to implement statistical analysis and characterize the in situ gas-bearing or GS (Figueiredo et al. 2018; Weinzier and Wiese 2021; Liu et al. 2022). For instance, Mazzotti and Zamboni (2003) delineated the interactive relationship between elastic parameters of V_p , V_s , and density retrieved from nonlinear AVO inversion and rock properties of depth, porosity, and saturation, and carried out multiple linear regression to evaluate GS from these

seismic parameters. Hampson et al. (2005) investigated the prestack simultaneous inversion to invert reliable V_p , V_s , and density models, and then calculated the V_p/V_s for specifying the gas-saturated sands of the shallow Cretaceous layers. Nevertheless, this classical stepwise procedure is limited by both seismic inversion errors and rock physics modeling errors, which increase the uncertainty and non-uniqueness of GS inference. In addition, the rock-physics templates usually use no more than three elastic parameters recovered from seismic inversion to predict GS in a low-dimension space. The transformation pattern cannot make full use of more information to express the complicated relationship between GS-sensitive parameters and GS in a high-dimension and nonlinear space.

Machine learning (ML) algorithms, as efficient solution tools, have been recently introduced into geophysical exploration realm again. Their typical application scenarios for seismic inverse problems include seismic noise attenuation (Yu et al. 2019; Sang et al. 2021), high-resolution processing (Gao et al. 2022), seismic inversion (Wu et al. 2021; Yuan, et al. 2022; Bürkle et al. 2023), and reservoir prediction (Sang et al. 2023). Some researchers have combined multi-component or prestack seismic data with different ML algorithms (e.g., self-organizing neural network) for gas-bearing prediction or GS characterization (Zhang et al. 2022a). Gao et al. (2020) utilized convolutional neural networks (CNNs) for gas-bearing estimation of a deep tight dolomite reservoir, and illustrated that transfer learning can improve the generalization ability of CNNs and generates more reliable gas-bearing distribution from AVA gathers. Song et al. (2022) adopted k -nearest neighbor (k NN) for gas-bearing estimation in a tight sand reservoir. Compared with the root-mean-square attribute, the k NN employed both seismic data and GS curves that reflect the gas-bearing corresponding to the near-well seismic traces, and obtained better gas-bearing delineation of the non-well seismic trace according to the average gas-bearing of its first k most similar near-well seismic traces. Zhang et al. (2022b) investigated deep neural network (DNN) and multi-component composite attributes for gas-bearing prediction in the Sichuan Basin, and generated more favorable gas-bearing results than the single PP-wave data. Weinzier and Wiese (2021) developed a three-layer neural network and rock-physics models for porosity and GS curves evaluation using elastic and attenuation attributes. In summary, most current ML-based studies have mainly focused on GS interpretation of specific well locations or qualitatively characterizing the spatial distribution of gas-bearing property. However, there are few researchers have explored quantitative spatial GS prediction using ML algorithms and prestack seismic data.

It seems that adding more complementary information available from the same stratum is a possible approach to improve the estimation accuracy of GS. Previous model-driven and data-driven methodologies merely use seismic data (or seismic attributes) and well-log data for gas-bearing prediction. These methods do not make full use of seismic data and elastic properties to boost the prediction accuracy of GS. Proposed by Caruana (1993 and 1997), multi-task learning (MTL) provides a new approach to flexibly utilize various information and realize multiple associated tasks in parallel, such as simultaneous inversion of elastic properties and joint interpretation of multiple seismic processing tasks (Wu et al. 2019; Zhang et al. 2023). Compared to single-task learning (STL), MTL has advantage over learning efficiency, prediction precision, and sample sharing by leveraging the differences and similarities among these tasks and sharing information and representations. (Caruana 1997; Ruder 2017; Wu et al. 2019). Currently, hard parameter sharing (Caruana 1993; Wu et al. 2019; Li et al. 2021), soft parameter sharing (Duong et al. 2015; Ruder 2017), and hierarchical sharing (Chen et al. 2022; Li et al. 2023) are main parameter sharing approaches used in MTL. For example, Zhang et al. (2023) adopted hard parameter sharing that shares some hidden layers among different tasks and uses several task-specific layers for the simultaneous inversion of V_p , V_s , and density from partially stacked angle stacks. They incorporated prior knowledge of elastic parameters into the MTL framework through a prior-based loss function term, and realized more stable and reliable inversion results compared with traditional learning-based methods. Hard parameter sharing can decrease the risk of overfitting at the expense of artificially adjusting the percentage of shared layers, and it is subject to negative transfer when tasks are interfered or uncorrelated (Li et al. 2023). Soft parameter sharing is an alternative approach to hard parameter sharing when the connection among tasks is weak (Li et al. 2021). When a progressive relationship exists between tasks, hierarchical sharing can learn effective information from one low-level task to improve the accuracy of another more complex task. Li et al. (2023) chose hierarchical sharing approach and investigated the progressive multi-task learning network (PMLN) for low-frequency extension of seismic shot gathers, elastic parameter (V_p) inversion, and image super-resolution in a progressive manner. Their method exhibits higher efficiency and precision compared with traditional full-waveform inversion.

Currently, MTL is mainly employed for intelligent prestack seismic inversion. The MTL framework for prestack simultaneous inversion of GS-related parameter and GS has

not been investigated yet. Generally, gas-bearing strata have lower velocity and lower density compared to surrounding rock due to the existence of pores, resulting in a reduced PI (Weinzier and Wiese 2021). Therefore, PI and GS are two interrelated but heterogeneous properties. We take the simultaneous prediction of PI and GS as an example and investigate the potential of integrating seismic gathers and GS-related parameter to improve the accuracy and stability of GS estimation via MTL.

In this present paper, we based on hierarchical sharing proposed the multi-task residual network (MT-ResNet) to realize the prestack simultaneous inversion of PI and GS. The designed MT-ResNet adopts two subsets and two types of labels (i.e., PI and GS curves) to establish a high-dimensional and nonlinear network template, which considers the task hierarchies and maps seismic data into well-log derived PI and well-log interpreted GS curves of the same formation in a progressive procedure. Specifically, each subset consists of two residual blocks, a convolution layer, and three regression layers. The first subnet learns the physical connection among prestack seismic gathers, low-frequency PI, and PI. Then the inverted PI together with prestack seismic gathers is entered into the second subnet, which further converts them into the desired GS curves. Furthermore, a single-task residual network (ST-ResNet) is employed as a contrast method. The ST-ResNet can be learned to capture the internal relationship between prestack seismic gathers and GS curves. We use a synthetic data example and a field data example to testify the effectiveness of the MT-ResNet-based joint inversion of PI and GS. The test results indicate that the optimized MT-ResNet involving two regressors can be generalized to unseen seismic data to simultaneously predict PI and GS. And the estimation precision of GS via the MT-ResNet is higher than the ST-ResNet.

Methodology

Numerical model generation

Since few well-logging curves are available, the inverted cross-well results via deep learning-based inversion cannot be evaluated and optimized by an expressible objective function. Consequently, extremely biased sample problems caused by abundant seismic data but rare well-log data have limited the rapid application of ML, especially supervised learning for geophysical/ petrophysical parameter inversion. To relieve this restriction, a relatively comprehensive dolomite reservoir data set is generated via geostatistical simulation. It can meet the demand of sufficient samples and their

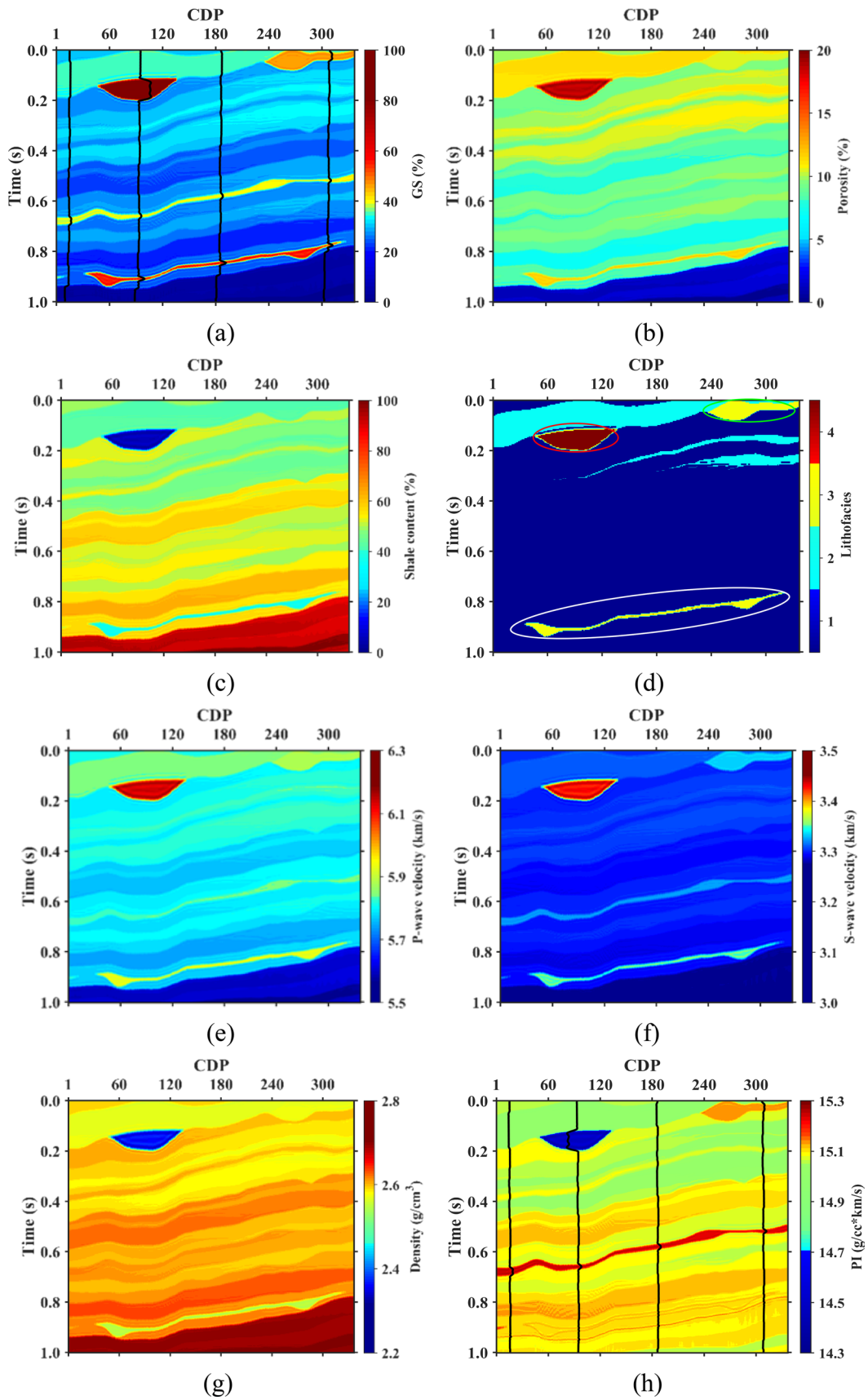


Fig. 1 The synthetic petrophysical, lithological and elastic models for describing the same subsurface tight gas-bearing dolomite reservoir. (a)–(h) represents the gas saturation (GS), porosity, shale content, lithofacies, P-wave velocity, S-wave velocity, density, and P-wave impedance (PI) model, respectively. Numbers 1–4 in (d) refer to mudstone, water-bearing dolomite, gas–water symbiotic dolomite, and gas-bearing dolomite, respectively. Four extracted GS and PI logs (black curves in a and h) are used as the training labels of the multi-task residual network (MT-ResNet)

labels for deep neural network modeling, and meanwhile offer different parameter models for MTL-based reservoir parameter inversion.

Based on available well-log data, geological and petrophysical understanding, and rock-physics or wave theory, the designed numerical models consist of lithofacies, petrophysical parameters, elastic parameters, and prestack seismic data. These four types of geophysical data simulate the same deep tight dolomite gas reservoir located in a basin of western China from different perspectives. Firstly, petrophysical parameters measured from the interested real tight dolomite reservoir are filled into the structure of the Marmousi2 model (Martin et al. 2006) to generate the petrophysical models, such as GS (Fig. 1a), porosity (Fig. 1b), and shale content (Fig. 1c). According to the content of petrophysical properties, lithofacies (Fig. 1d) are assigned into four categories, namely mudstone, water-bearing dolomite, gas–water symbiotic dolomite, and gas-bearing dolomite. And they are denoted as numbers 1–4 shown in Fig. 1d. The bowl-shaped gas-bearing dolomite (red circle in Fig. 1d), gas–water symbiotic dolomite formed by a banded formation (white circle in Fig. 1d), and a spoon-like formation (green circle in Fig. 1d) are dominant gas-bearing areas that GS is over 40%. And their reservoir characteristics are favorable porosity and low shale content. Then the Kuster-Toksöz model (Kuster and Toksöz, 1974) is utilized to convert these petrophysical parameters into elastic parameter models, including V_p (Fig. 1e), V_s (Fig. 1f), and density (Fig. 1g). Figure 1h shows the PI model (i.e., the product of V_p and density). The gas-bearing dolomite has a distinguishable characteristic of low PI value. Next, the angle-dependent reflectivities are derived from the synthetic elastic parameters via the Aki-Richards approximate formula (Aki and Richards 2002); the clean prestack AVA gathers are generated by convolving the reflectivities at different angles with a Ricker wavelet. The main frequency of the Ricker wavelet is 35 Hz. Finally, the noisy prestack AVA gathers (Fig. 4a) were obtained by adding 30% random noise to noise-free prestack AVA gathers. These models can be used for different reservoir-associated parameter prediction tasks with flexible combinations, such as simultaneous prediction of PI and GS.

MT-ResNet-based simultaneous inversion of PI and GS

In the designed dolomite reservoir data set, these parameter models (Fig. 1) reveal the stratum properties of the same underground gas reservoir from different perspectives. For instance, the bowl-shaped gas-bearing dolomite (red circle in Fig. 1d) has the characteristics of high GS, high porosity, low PI, and low shale content. It means that these properties are interrelated and complementary, and one or multiple gas-related parameters can be chosen to help predict the target parameter of GS. In our studied synthetic and field cases, PI and GS have a high correlation. Figure 2 shows the cross-plot between all points in the PI model (Fig. 1h) and all points in the GS model (Fig. 1a). The blue, cyan, yellow, and red points in Fig. 2 mean that these PI or GS values correspond to mudstone, water-bearing dolomite, gas–water symbiotic dolomite, and gas-bearing dolomite, respectively. The relationship between GS and PI can be approximately expressed by a nonlinear fitting equation:

$$GS = 46.4PI^2 + 1256.1PI8361.7, \quad (1)$$

and the correlation coefficient between estimated GS and true GS is 0.6. Therefore, based on the framework of MTL, we take the simultaneous inversion of PI and GS as an example to demonstrate that the alliance of seismic data and sensitive attributes can improve the prediction accuracy of target parameters.

Once PI and GS are determined as two reservoir parameters to be inverted, the configuration of the MT-ResNet will be designed. Figure 3a shows the network architecture of the MT-ResNet. Compared with traditional ML-based method for unique parameter estimation, the MT-ResNet can directly realize data-driven simultaneous inversion of PI and GS from prestack seismic gathers. Wherein, the main task of MT-ResNet is GS prediction, and the auxiliary task of MT-ResNet is PI inversion. In addition, the MT-ResNet has the potential to improve the prediction accuracy of GS by leveraging the GS-related information (i.e., PI) estimated by the auxiliary task. It is worth noting that these two tasks are trained at the same time. The MT-ResNet adopts two subnets to express the intrinsic physical correlation among seismic data and reservoir parameters and establish the complex nonlinear mapping to convert seismic data into PI and GS simultaneously. The first subnet is a data-driven PI inversion solver. The input of the first subnet is the low-frequency PI log and prestack AVA or offset gathers, and its output is the inverted PI curve. Here, the low-frequency PI log controls the low-frequency trend and enhances the stability of data-driven PI inversion (Wu et al. 2021). The role of the first subnet is mainly to provide the second subnet with PI, which is a sensitive seismic attribute of GS here. The second subnet

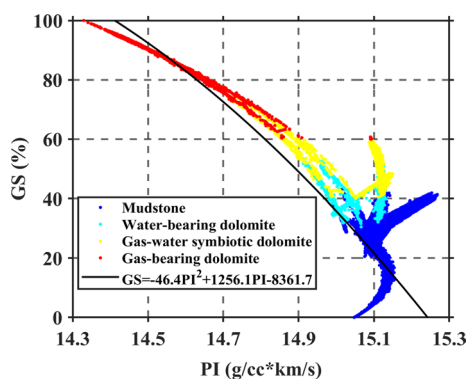


Fig. 2 The cross-plot between all PI values in Fig. 1h and all GS values in Fig. 1a

is designed to integrate the prestack seismic data with the GS-related seismic attribute to form a multi-information fused GS estimator. And its input is not only the raw prestack AVA or offset gather but also the PI curve inverted by the first subnet. And the output of the second subnet is the GS curve. In the framework of MT-ResNet, the inverted PI via the first subnet and the estimated GS via the second subnet can be mutually verified. Their structural similarity and corresponding physical relationship can be used to illustrate the reliability of the reservoir parameters predicted by the MT-ResNet. The hidden layers of the first or second subnet consist of two residual units, three convolution layers, and three fully connected layers (FCL). Each residual unit has two convolution layers and one addition layer. The convolution layer includes the convolution (Conv) operations with the kernel size of 3×3 , the batch normalization (BN) operations, and the activation functions of rectified linear units (ReLU). The Conv operations are used to extract high-dimensional “seismic attributes” related to PI or GS; the BN operations can maintain the stability of convergence and accelerate the training process of MT-ResNet (Ioffe and Szegedy 2015). Here, the ReLU can enhance the nonlinear expression capability of the MT-ResNet. The addition layer aims to add the extracted features of shallow layers to the extracted features of deep layers. The introduction of the residual units can improve the flow of information between different layers and thus avoid performance degradation (He et al. 2016).

On the basis of the network architecture, we further illustrate the workflow of MT-ResNet. First, prestack seismic data are normalized and clipped into seismic patches. The size of each seismic patch is $m \times n$, where m and n represent the number of time samples and the number of angles or offsets, respectively. The low-frequency PI log corresponds to each seismic patch is also extracted, and its size is $m \times 1$. We use normalized seismic patches and normalized low-frequency PI logs as two kinds of source information to form the input of the MT-ResNet. Therefore, the input one of the MT-ResNet has the size of $m \times (n + 1)$.

Then under the supervision of PI logs, the first subnet hierarchically learns latent reservoir characteristic or attribute information from low-frequency PI logs and seismic patches, and abstracts them into feature maps at different scales or levels via convolution layers and residual units. The extracted high-dimensional features are weighted and regressed into the inverted PI log via three fully connected layers in the first subnet. Subsequently, the evaluated PI log with the size of $m \times 1$ and the raw prestack seismic patch with the size of $m \times n$ are flowed into the second subnet. The second subnet can be viewed as the implicit expression of sensitive parameter models or templates, which finally maps seismic gathers and gas-associated attribute (i.e., PI) into the desired GS log in a high-dimensional and nonlinear space. As displayed in Fig. 3a, the output size of each convolution layer or the residual unit is $m \times (n + 1) \times c$, wherein c represents the channel number of feature maps. And the output size of each FCL is $p \times 1$, where p is the node number of the FCL. The objective function of MLT-ResNet is defined to measure the distance between estimated two properties and their ground truths, which is expressed as:

$$L_1 = \frac{1}{N} \sum_{i=1}^N \left[\lambda \left\| \text{PI}_i - \text{Net}_1(d_i; \text{PI}_i^{\text{low}}; \Theta^1) \right\|_2^2 + (1 - \lambda) \left\| \text{GS}_i - \text{Net}_1(d_i; \text{PI}_i^{\text{pre}}; \Theta^1) \right\|_2^2 \right], \quad (2)$$

where N denotes the number of training samples, Θ^1 represents the network parameters of the MT-ResNet, and $\text{Net}_1(\bullet)$ stands for the nonlinear mapping of the MT-ResNet. $d_i, \text{PI}_i^{\text{low}}, \text{PI}_i$, and GS_i stand for the i th prestack seismic patch, i th low-frequency PI log, i th true PI log, and i th true GS log, respectively. The total loss L_1 consists of two terms, and each of them uses mean square error to calculate the estimated errors and jointly monitor the learning process of the MT-ResNet. The first term computes the loss between the estimated impedance PI_i^{pre} and the impedance label PI_i derived from well logs. The second term calculates the loss between the predicted gas saturation $\text{Net}_1(d_i; \text{PI}_i^{\text{pre}}; \Theta^1)$ and the gas saturation label GS_i interpreted from well logs. λ can control the relative weight between the PI inversion task and the GS inference task. And $1 - \lambda$ is usually larger than λ due to that now GS is a more crucial reservoir parameter. Here, the predicted impedance can be written as:

$$\text{PI}_i^{\text{pre}} = \text{Net}_1(d_i; \text{PI}_i^{\text{low}}; \Theta^1). \quad (3)$$

The Adam (Kingma and Ba 2014) algorithm is adopted to iteratively update Θ^1 until the estimate errors converge to small values without a rising trend. Eventually, we apply the optimized MT-ResNet to the test data and predict PI and GS models simultaneously. The above workflow of the MT-ResNet is summarized in Algorithm 1.

Algorithm 1. The workflow of prestack simultaneous inversion of PI and GS using MT-ResNet.**Data preparation process**

Input: \mathbf{D} : raw prestack seismic data, \mathbf{PI} : synthetic PI models or PI curves derived from well log data, \mathbf{PI}^{low} : low-frequency PI model, \mathbf{GS} : synthetic GS models or GS curves interpreted from well log data, N : number of the training set, J : number of the valid set, K : number of the test set.

Output: $\{d_i, \text{PI}_i, \text{PI}_i^{\text{low}}, \text{GS}_i\}_{i=1}^N$: the training set, $\{d_j, \text{PI}_j, \text{PI}_j^{\text{low}}, \text{GS}_j\}_{j=1}^J, J \leq N$: the valid set, $\{d_k, \text{PI}_k^{\text{low}}\}_{k=1}^K$: the test set.

Training process

Input: $\{d_i, \text{PI}_i, \text{PI}_i^{\text{low}}, \text{GS}_i\}_{i=1}^N, \{d_j, \text{PI}_j, \text{PI}_j^{\text{low}}, \text{GS}_j\}_{j=1}^J, \text{Epo}$: Epoch, bs_1 : batch size for network training, Θ^1 : network parameters, η : learning rate.

Initialization: $L_{\text{train}} = 0, L_{\text{valid}} = 0$

for $l_1 = 1$ to Epo **do**

for $n_1 = N/bs_1$ **do**

$\text{PI}_i^{\text{pre}} = \text{Net}_1(d_i; \text{PI}_i^{\text{low}}; \Theta^1)$; // predicted PI via the first subnet of the MT-ResNet,

$\text{GS}_i^{\text{pre}} = \text{Net}_1(d_i; \text{PI}_i^{\text{pre}}; \Theta^1)$; // estimated GS via the second subnet of the MT-ResNet,

$L_{\text{train}} = \frac{1}{N} \sum_{i=1}^N \left[\lambda \|\text{PI}_i - \text{PI}_i^{\text{pre}}\|_2^2 + (1-\lambda) \|\text{GS}_i - \text{GS}_i^{\text{pre}}\|_2^2 \right]$; // training loss of MT-ResNet.

end

$\Theta_{n+1}^1 = \text{Adam}(\Theta_n^1, L_1, \eta)$; // update network parameters using Adam algorithm.

for $j_1 = J/bs_1$ **do**

$L_{\text{valid}} = \frac{1}{J} \sum_{j=1}^J \left[\lambda \|\text{PI}_j - \text{PI}_j^{\text{pre}}\|_2^2 + (1-\lambda) \|\text{GS}_j - \text{GS}_j^{\text{pre}}\|_2^2 \right]$; // valid loss of MT-ResNet.

end

end

Output: $\text{Net}_1^{\text{best}}$: optimal network model, Θ_1^{best} : optimal network parameters.

Testing process

Input: $\{d_k, \text{PI}_k^{\text{low}}\}_{k=1}^K, \text{Net}_1^{\text{best}}, \Theta_1^{\text{best}}, bs_2$: batch size for network test.

for $l_2 = 1$ to K/bs_2 **do**

$\text{PI}_k^{\text{pre}} = \text{Net}_1^{\text{best}}(d_k; \text{PI}_k^{\text{low}}; \Theta_1^{\text{best}})$,

$\text{GS}_k^{\text{pre}} = \text{Net}_1^{\text{best}}(d_k; \text{PI}_k^{\text{pre}}; \Theta_1^{\text{best}})$.

end

Output: Inverted PI and GS models by recombining patches $\{\text{PI}_k^{\text{pre}}\}_{k=1}^K$ and $\{\text{GS}_k^{\text{pre}}\}_{k=1}^K$ into profiles.

As shown in Fig. 3b, the ST-ResNet, as the comparison approach of the MT-ResNet, is constructed to directly transform the input prestack seismic patch with the size of $m \times n$ into the output estimated GS with the size of $m \times 1$. There are two residual units, three convolution layers, and three FCLs in the ST-ResNet. Both the ST-ResNet and the first or second subnet of the MT-ResNet have the same amount of kernels in the convolution layer and the same number of nodes in the FCL. As a result, the training time of the ST-ResNet is around half that of the MT-ResNet. The ST-ResNet, in contrast to the MT-ResNet, is a standard supervised learning system that merely utilizes prestack seismic

data and well-log derived GS curves to implement network training. GS-related information (e.g., sensitive gas-bearing properties) is not integrated into the ST-ResNet for GS evaluation, resulting in GS estimation errors of the ST-ResNet are commonly larger than that of the MT-ResNet. The objective function of the ST-ResNet is expressed as:

$$L_2 = \frac{1}{N} \sum_{i=1}^N \left\| \text{GS}_i - \text{Net}_2(d_i; \Theta^2) \right\|_2^2, \quad (4)$$

where Θ^2 represents the network parameters of the ST-ResNet, $\text{Net}_2(\bullet)$ represents the nonlinear mapping of the ST-ResNet, and $\text{Net}_2(d_i; \Theta^2)$ means the i th estimated GS log

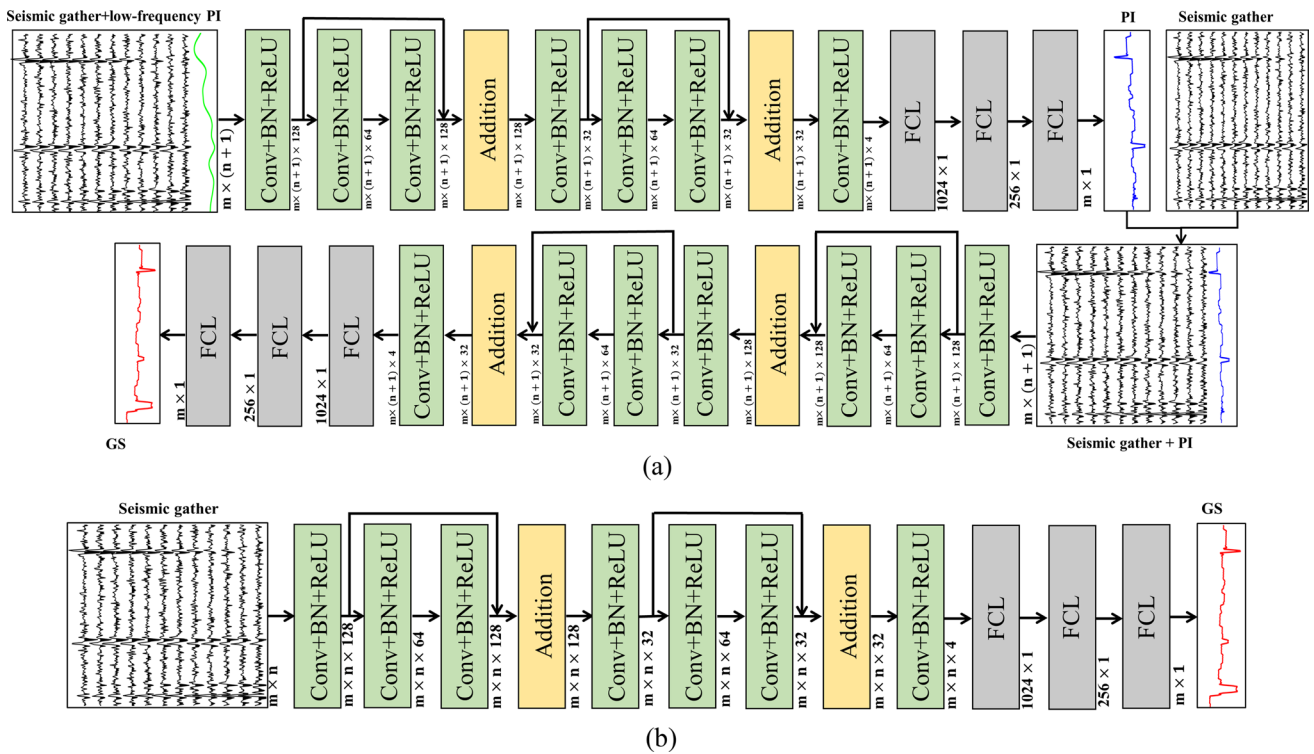


Fig. 3 The network architecture of reservoir parameter prediction. **a** The MT-ResNet for simultaneous inversion of PI and GS, and **b** the ST-ResNet for GS prediction

via the ST-ResNet. The contrast of Eqs. (3) and (4) or Fig. 3a and b illustrates that the ST-ResNet only predicts GS and cannot achieve the simultaneous inversion of PI and GS as the MT-ResNet. As same as the MT-ResNet, the ST-ResNet also employs the Adam algorithm for its optimization.

Here, we use the root mean square error (RMSE) to measure the distance between the true model X (i.e., PI or GS) and the predicted model via the ST-ResNet or MT-ResNet:

$$RMSE = \frac{\sqrt{\|X - \bar{X}\|_2^2}}{\sqrt{J \times K}}, \tag{5}$$

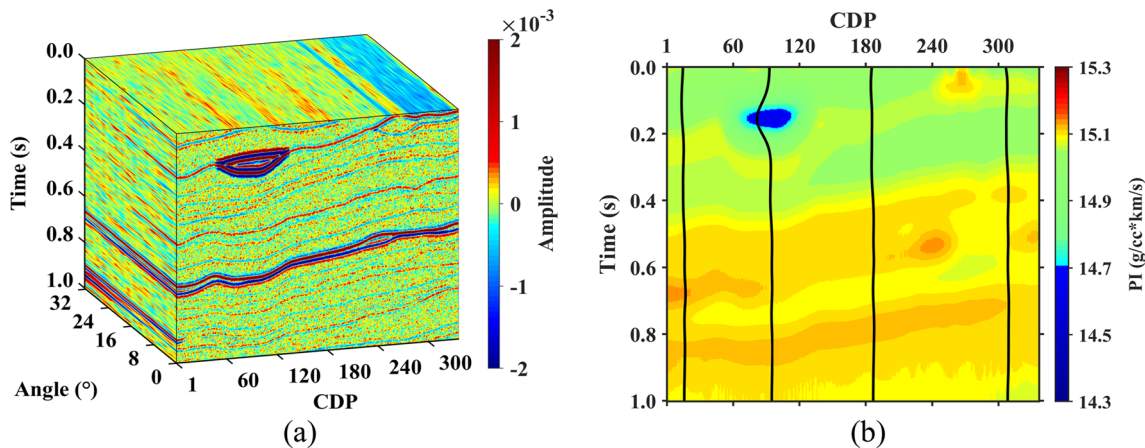


Fig. 4 The synthetic seismic data and the established initial model. **a** Noisy prestack seismic data generated by adding 30% (i.e., the ratio of noise energy to signal energy) random noise to the noise-free seismic data, and **b** the initial PI model with the frequency band

of 0–5 Hz. The initial PI logs (black curves in b) and corresponding near-well AVA gathers are utilized as the input training samples of MT-ResNet

where J and K denote the number of rows and columns in \mathbf{X} or $\bar{\mathbf{X}}$, respectively. In addition, we utilize structural similarity index measurement (SSIM) to qualitatively evaluate the similarity between \mathbf{X} and $\hat{\mathbf{X}}$:

$$\text{SSIM} = \frac{(2\mu_{\mathbf{X}}\mu_{\bar{\mathbf{X}}} + c_1)(2\text{Cov}_{\mathbf{X}\bar{\mathbf{X}}} + c_2)}{(\mu_{\mathbf{X}}^2 + \mu_{\bar{\mathbf{X}}}^2 + c_1)(\sigma_{\mathbf{X}}^2 + \sigma_{\bar{\mathbf{X}}}^2 + c_2)}, \quad (6)$$

where $\mu_{\mathbf{X}}$ and $\mu_{\bar{\mathbf{X}}}$ are the mean values of \mathbf{X} and $\bar{\mathbf{X}}$, respectively. $\sigma_{\mathbf{X}}$ and $\sigma_{\bar{\mathbf{X}}}$ are the standard deviations of \mathbf{X} and $\bar{\mathbf{X}}$, respectively. $\text{Cov}_{\mathbf{X}\bar{\mathbf{X}}}$ denotes the covariance between \mathbf{X} and $\bar{\mathbf{X}}$. c_1 and c_2 are constants. Generally, c_1 is equal to 0.01, and c_2 is equal to 0.03 (Wang et al. 2004).

Examples

Synthetic data example

In this subsection, we will first go through the process of preparing training, validation, and test datasets. The network training for the ST-ResNet and MT-ResNet is then implemented. Finally, we use the synthetic data test to illustrate the effectiveness and superiority of the MT-ResNet over the ST-ResNet for GS estimation.

In addition to the PI model (Fig. 1h), the GS model (Fig. 1a), and the prestack seismic data (Fig. 4a), we further apply a low pass filter of 0–5 Hz to Fig. 1h and acquire the low-frequency PI model (Fig. 4b) for the synthetic data test. The size of PI, low-frequency PI, or GS model is 1000×337 , and the size of synthetic AVA gathers is $1000 \times 24 \times 337$. Wherein, 1000, 24, and 337 refer to the time sampling number, the number of angles, and the number of common depth points, respectively. It is apparent that the continuity of seismic events in Fig. 4a has been disrupted, and the S/N ratio

(Sang et al. 2021) is 5.2 dB. Additionally, the noisy prestack AVA gathers have 24 incident angles that are evenly spaced and vary from 0° to 32° . The time sample interval of these earth models or seismic data is 1 ms.

Before generating the training set, we investigate the physical link among seismic data, PI, and GS, providing a foundation for evaluating the prediction results of the two networks. Figure 5 describes the near-well AVA gathers, PI logs (blue curves), and GS logs (red curves) at the well positions of CDP 92 and CDP 306. As illustrated in Fig. 5, the seismic and well-log responses of the gas-bearing dolomite are markedly different from other lithofacies. It has the characteristics of low PI, high GS, high amplitude anomaly, and amplitude decreases with incidence angles. Gas–water symbiotic dolomite shows normal amplitude in small and medium angles, and amplitude increases with the angle in the range of large angles. Water-bearing dolomite and mudstone exhibit weak amplitude in all angles. Different lithofacies have different AVA features, which provide a physical basis for PI and GS estimation. PI and GS show the best correlation in gas-bearing dolomite compared with other lithologies. In comparison with the ST-ResNet, the main advantage of MT-ResNet may be that it can ameliorate the prediction performance of GS around gas-bearing dolomite.

For methodological comparison, we begin to extract seismic gathers and well-log curves and construct the training, valid, and test sets of these two networks. Before preparing these sets, all parameter models and prestack seismic data are individually normalized to $[0, 1]$ by means of min–max normalization (Sang et al. 2021). Four pseudo wells are randomly selected to prepare the training sets of these two networks, and their concrete locations are CDP 13, CDP 92, CDP 184, and CDP 306. Taking the data sets construction of the MT-ResNet as an example, normalized PI and GS curves of four selected wells (black lines in Fig. 1h and a)

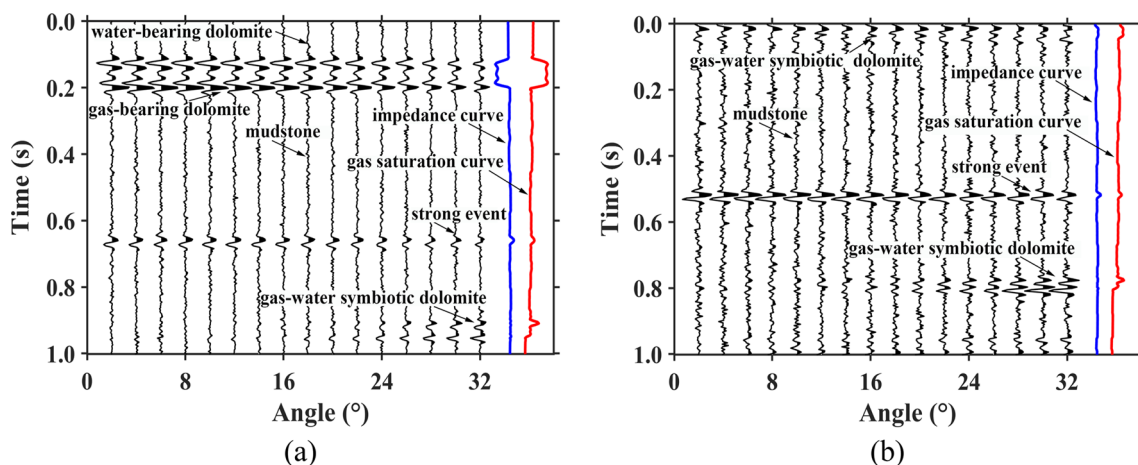


Fig. 5 Near-well prestack seismic gathers at the training well locations of **a** CDP 92 and **b** CDP 306. The blue and red lines in **a** or **b** are PI and GS logs corresponding to the seismic gathers

are firstly extracted for generating training labels. Four normalized low-frequency PI logs (black lines in Fig. 4b) and corresponding normalized near-well AVA gathers in Fig. 4b are extracted for producing the training samples. Next, these extracted low-frequency PI, PI, and GS curves are clipped into patches by means of the local temporal window with the stride of 1. Each low-frequency PI, PI, or GS patch has the size of 65×1 . The corresponding AVA gathers are also chopped into patches via a similar process. The size of each seismic patch is 65×24 . Subsequently, we use these seismic patches and low-frequency PI patches as the training samples of the MT-ResNet, PI patches as the training labels of the first subnet of the MT-ResNet, and GS patches as the training labels of the second subnet of the MT-ResNet. Specifically, the training data set has 3744 training samples, 3744 PI labels, and 3744 GS labels. The size of each training sample is 65×25 , and the size of each PI or GS label is 65×1 . In addition, all seismic patches and GS labels in the training set of the MT-ResNet are also utilized as the training set of the ST-ResNet. Finally, the valid and test data sets of these two networks are prepared in a similar way to the training data set of them. After the preparation of these sets, the training samples and labels are used to train the MT-ResNet with a batch size of 256 and a learning rate of 0.001. Under the joint implementation of training data, optimizer, objective function, and backpropagation algorithm, the network completes its preliminary learning and training by setting the maximum number of epochs in advance. The loss curves are plotted to observe whether the training loss and the validation loss converge to a minimum value at the same time. And if not, we further adjust the epochs or the regularization parameter in Eq. 2 until two types of losses simultaneously satisfy the convergence conditions. By repeated experiments, the loss weight λ in the objective function of the MT-ResNet is set to 0.4. The optimal MT-ResNet model for prestack simultaneous inversion of PI and GS can be obtained when the maximum number of iterations is determined to be 400. The ST-ResNet is also trained with seismic patches and GS labels in terms of other conditions unchanged.

We apply the well-trained ST-ResNet and MT-ResNet to the test data set and obtain the PI or GS results with the size of $315,432 \times 65$, where 315,432 represents the number of test samples. The predicted PI or GS model is then obtained by averaging the test results over the overlapped parts. Figure 6a shows the estimated GS model via the ST-ResNet, and Fig. 6b shows the residuals between Fig. 6a and the true GS model (Fig. 1a). The RMSE and SSIM between Figs. 6a and 1a are 0.07 and 0.59, respectively. Figure 6a and b illustrate that the estimation accuracy of the ST-ResNet is highly influenced by the quality of prestack seismic gathers. The negative interference of random noise results in that the estimated GS model via the ST-ResNet shows terrible

lateral continuity (Fig. 6a) and distinct deviation (Fig. 6b). Therefore, when the seismic data is heavily contaminated by noise, neural networks rely on more gas-sensitive attributes to improve the accuracy and stability of GS prediction. It is worth noting that the inverted GS model clearly delineates the spatial morphology of gas-bearing dolomite. The potential reasons are that the seismic gathers around the gas-bearing dolomite have higher local S/N ratio compared with the surrounding rock, and the GS labels in the training data set are comprehensive and are already enough to represent the gas-bearing features of gas-bearing dolomite. Figure 6c and e displays the inverted PI model and the inverted GS model via the MT-ResNet, respectively. The differences between Fig. 6c and the reference PI model (Fig. 1h) are shown in Fig. 6a, and the differences between Fig. 6e and the reference GS model (Fig. 1a) are shown in Fig. 6f. And the RMSE and SSIM between Figs. 6e and 1a are 0.06 and 0.81, respectively. It can be illustrated from Fig. 6c and d that the MT-ResNet can precisely estimate the PI model using seismic amplitude and low-frequency PI information. The RMSE and SSIM between Figs. 6c and 1h are 0.02 and 0.97, respectively. On the one hand, the inverted PI model (Fig. 6c) can reduce the influence of seismic noise for the MT-ResNet based GS prediction. On the other hand, it can provide accuracy gas-sensitive properties for boosting the estimation precision of GS. Therefore, the inverted GS model (Fig. 6e) via the MT-ResNet is significantly superior to the inverted GS model (Fig. 6a) via the ST-ResNet. And the former shows more favorable consecutiveness and higher precision.

In Fig. 7, we ulteriorly compare the differences between the ST-ResNet and the MT-ResNet in terms of the frequency distribution of the estimated GS models. Figure 7a, b and c show the histograms of the true GS model (Fig. 1a), the ST-ResNet retrieved GS model (Fig. 6a), and MT-ResNet retrieved GS model (Fig. 6e), respectively. It can be seen from Fig. 6a that the true GS model obeys the non-Gaussian distribution. Both ST-ResNet and MT-ResNet can approximately fit this non-Gaussian distribution due to the non-linear fitting ability of neural networks. On the whole, the GS distribution estimated by the MT-ResNet (Fig. 7c) is closer to the actual GS distribution (Fig. 7a) than that estimated by the ST-ResNet (Fig. 7b). We divide the range of GS into four segments according to the content of GS in Fig. 7a, and they are denoted by purple, cyan, green, and red circles in Fig. 7, respectively. The ranges of GS represented by purple, cyan, green, and red circles are 0–20%, 20–40%, 40–60%, and 80–100%, respectively. In addition, different ranges of GS correspond to different lithofacies. Purple, cyan, green, and red circles correspond to mudstone, mudstone or water-bearing dolomite, gas–water symbiotic dolomite or water-bearing dolomite, and gas-bearing dolomite, separately. As shown in Fig. 6e and the cyan circles in

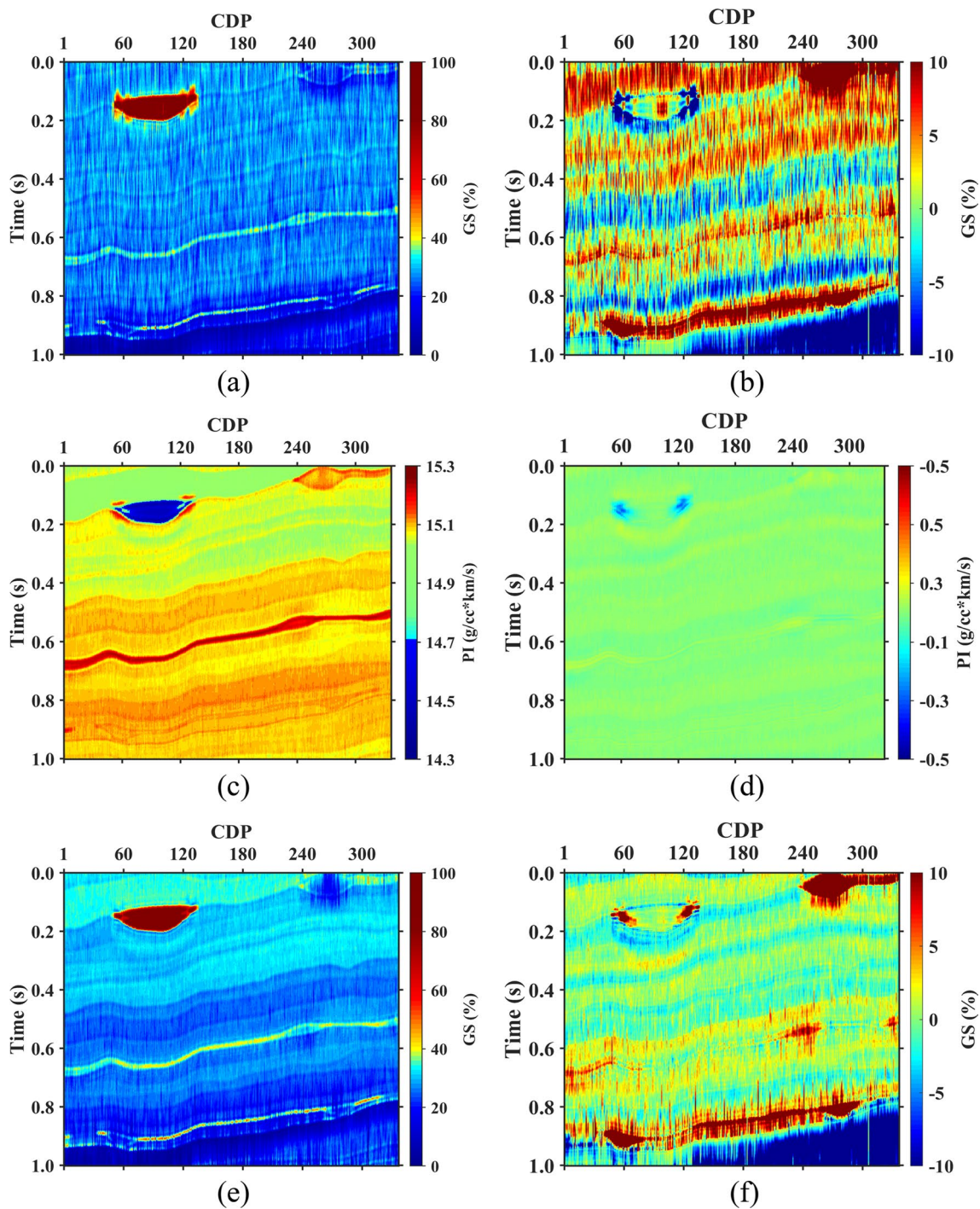


Fig. 6 Comparisons between the ST-ResNet and MT-ResNet estimated parameter models. **a** The inverted GS model via the ST-ResNet, **b** the residuals between true GS model (Fig. 1a) and (a), **c**

the inverted PI model via the MT-ResNet, **d** the differences between true PI model (Fig. 1e) and (c), **e** the inverted GS model via the MT-ResNet, and **f** the residuals between Fig. 1a and e

Fig. 7, the MT-ResNet has better prediction performance of GS than the ST-ResNet in some mudstone and water-bearing dolomite. The underlying reason may be these two lithofacies occupy a larger proportion in both training and test sets. And the MT-ResNet mainly improves the estimation

accuracy of GS in mudstone and water-bearing dolomite by the predicted PI with high precision. The GS prediction accuracy of two networks in purple and green circles is relatively low, which may be caused by the that the AVA characteristics of gas–water symbiotic dolomite (or water-bearing

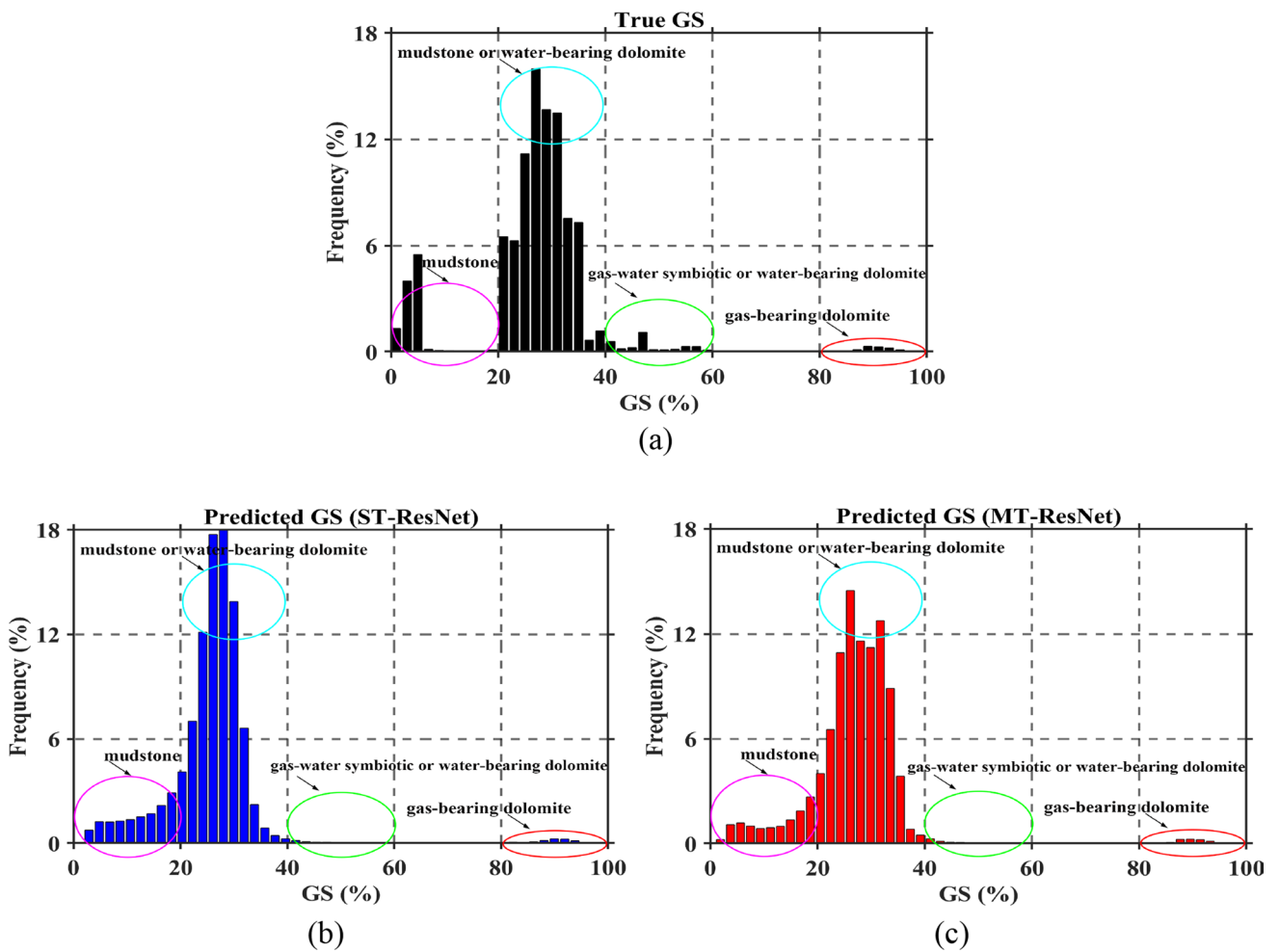


Fig. 7 The histograms of **a** the true GS model of Fig. 1a, **b** the predicted GS model (Fig. 6a) via the ST-ResNet, and **c** the predicted GS model (Fig. 6e) via the MT-ResNet. The range of GS within the pink,

cyan, green, and red circles in **a–c** corresponds to mudstone, mudstone or water-bearing dolomite, gas–water symbiotic or water-bearing dolomite, and gas-bearing dolomite, respectively

Fig. 8 Comparisons between ST-ResNet and MT-ResNet predicted GS curves at the blind well locations of **a** CDP 80, **b** CDP 160, and **c** CDP240. The black, blue, and red lines in **(a–c)** represent the true, ST-ResNet inverted, and MT-ResNet inverted GS curves, separately

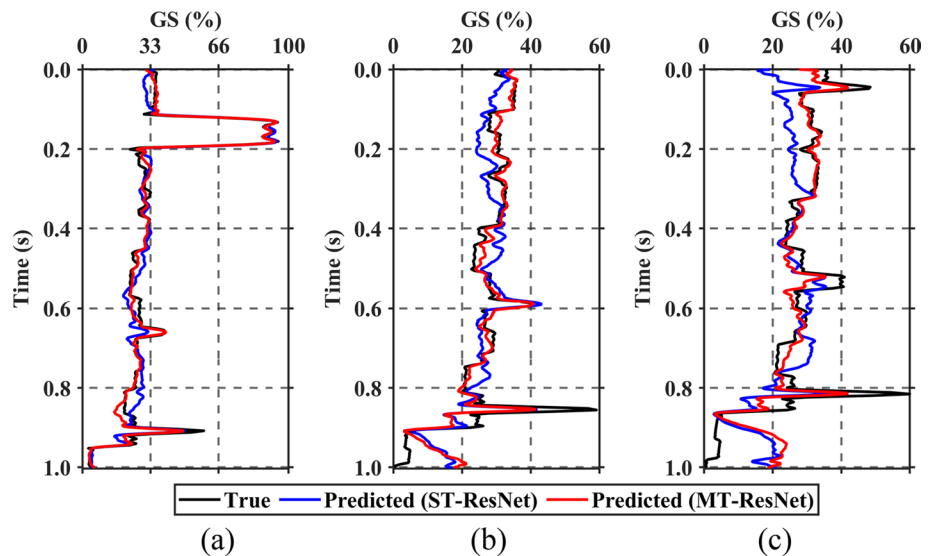


Table 1 The RMSE and SSIM between true parameters and inverted parameters using the ST-ResNet or the MT-ResNet for synthetic prestack seismic data with different S/N ratios

S/N (dB)	GS (ST-ResNet)	GS (MT-ResNet)	PI (MT-ResNet)
$+\infty$ (clean)	0.067/0.914	0.060/0.923	0.0173/0.992
10	0.068/0.753	0.060/0.882	0.0173/0.983
7.0	0.070/0.667	0.061/0.847	0.0174/0.975
5.2	0.072/0.594	0.061/0.808	0.0177/0.967
4.0	0.075/0.537	0.064/0.775	0.0179/0.960
3.0	0.077/0.489	0.064/0.746	0.0182/0.954

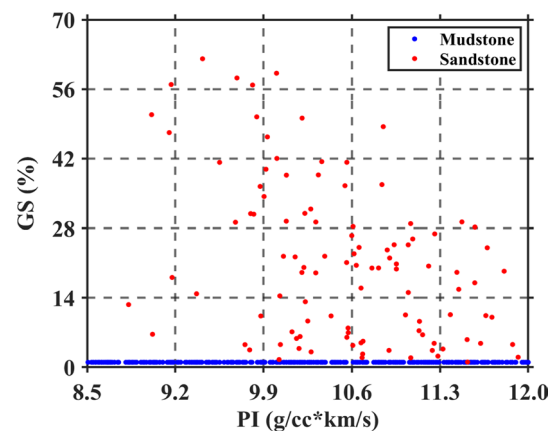
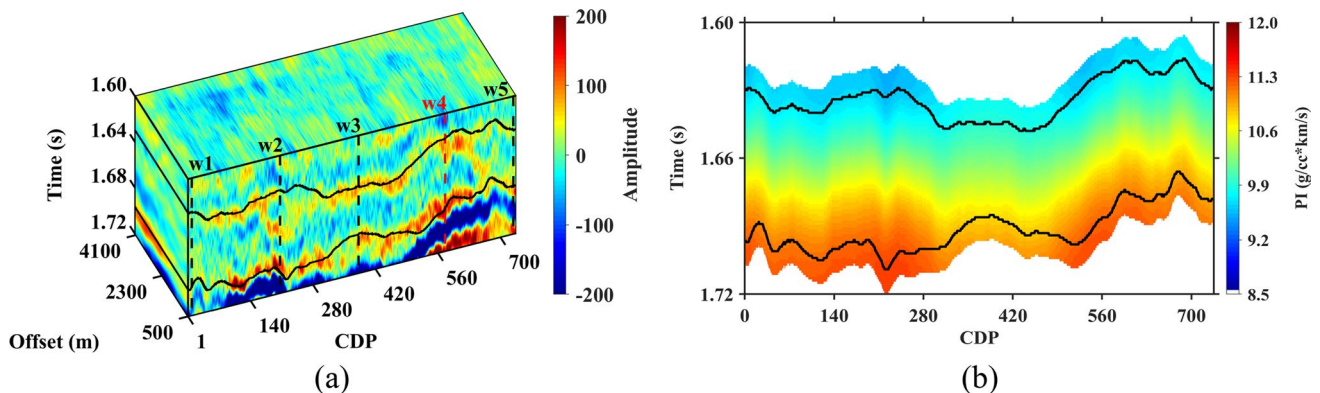
dolomite) are similar to the mudstone in small and middle angles, as illustrated in Fig. 5. This similarity may cause the two networks to misestimate the GS of gas–water symbiotic dolomite (or water-bearing dolomite) as the GS range of mudstone, as displayed in purple and green circles of Fig. 7.

To further compare single-trace predictions via two networks, Fig. 8 shows the comparisons among true GS curves (black lines), ST-ResNet inverted GS curves (blue lines), and MT-ResNet inverted GS curves (red lines) at three blind well locations of CDP 80, CDP 160, and CDP 240. Compared with ST-ResNet estimated GS curves, the overall trend and local variation of MT-ResNet estimated GS curves are more consistent with true GS curves. The above-mentioned ST-ResNet and MT-ResNet are trained by synthetic seismic data with the S/N ratio of 5.3 dB. Finally, we add 0–50% Gaussian white noise to the clean seismic data and generate noisy seismic data with different S/N ratios. Noise-free and noisy seismic data are tested to evaluate the noise-resistance and generalization ability of the two networks. Table 1 summarizes the RMSE and SSIM between inverted results via the two methods and true models for clean or noisy seismic data. As can be seen from Table 1, the RMSE and SSIM of estimated PI or GS models via the two methods decrease with the decline of S/N ratios. However, the ST-ResNet is less

stable and less accurate than the MT-ResNet with the gradually increase of noise levels. Their difference is prominent in the scenes of low S/N ratios. Figures 6, 7 and 8 and Table 1 show that the MT-ResNet is better than the ST-ResNet and can obtain more precise and stable inversion results in both noise-free and low S/N cases.

Real data example

In this subsection, we adopt a field data example from Northern China to further verify the effectiveness of prestack simultaneous inversion of PI and GS using the MT-ResNet. The working area is a tight sandstone gas-bearing reservoir, characterized by low porosity and low permeability. The gas reservoir is mainly composed of sandstone and mudstone. Figure 9a shows the prestack seismic data that passes through five wells (named w1–w5 from left to right). The size of prestack seismic data is $121 \times 16 \times 735$. Wherein, 121, 16, and 735 refer to the time sampling number, the number of offsets, and the number of common depth points,

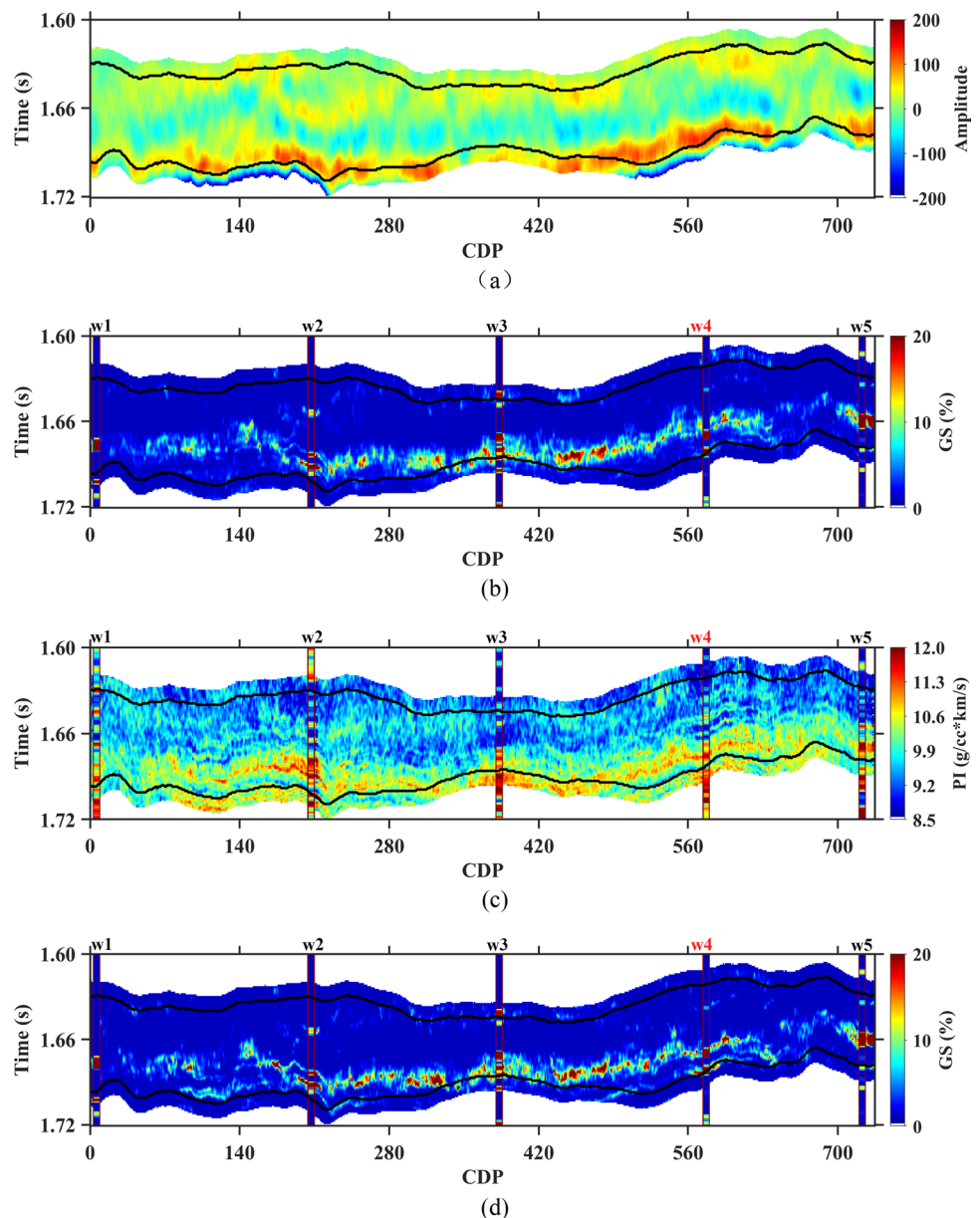
**Fig. 10** The cross-plot between PI values derived from five wells and GS values interpreted from five wells**Fig. 9** The field seismic data and the established initial model. **a** The field prestack seismic data overlaid with five wells (w1–w5) and two interpreted horizons (i.e., two lateral black curves), and **b** the low-frequency PI model

respectively. The time range and the offset range of prestack seismic data are 1.60–1.72 s and 500–4100 m, and the time sampling interval is 1 ms. The strata between two target horizons (black lateral curves in Fig. 9a) stand for the interested reservoir units with a burial depth of about 3 km. The thickness of the reservoir units is approximately 110–150 m. Geological characteristics of the target reservoir interval are severe interbedding of sand and mud, and developed single thin sand bodies are about 10 m. It can be seen from Fig. 9a that the seismic data is low resolution and only develops one trough in the reservoir units. After depth-time conversion, we up-sample the PI curves derived from the measured well logs and the GS curves interpreted from the measured well logs to 1 ms. Figure 9b shows the low-frequency PI model

interpolated by seismic horizons and five PI curves, and its frequency band is 0–8 Hz.

Figure 10 shows the cross-plot between all PI values and all GS values of five well logs at the range of 1.60–1.72 s. Red and blue points in Fig. 10 correspond to sandstone and mudstone, respectively. GS and PI show certain negative correlation in the sandstone part, but show poor correlation in the mudstone part due to GS is interpreted as minimal constant in the mudstone. It can be seen from Fig. 10 that PI and GS have correlation on the whole. Therefore, we implement MT-ResNet-based prestack simultaneous inversion and use inverted PI to improve the prediction effect of GS. We adopt the well of w4 as the test blind well and use other four wells as the training wells. Four interpreted GS curves and their corresponding borehole-side AVO gathers are used to

Fig. 11 Comparisons between the ST-ResNet and MT-ResNet estimated parameter models for the field data. **a** The poststack seismic profile, **b** the inverted GS profile via the ST-ResNet, **c** the inverted PI profile via the MT-ResNet, and **d** the inverted GS profile via the MT-ResNet



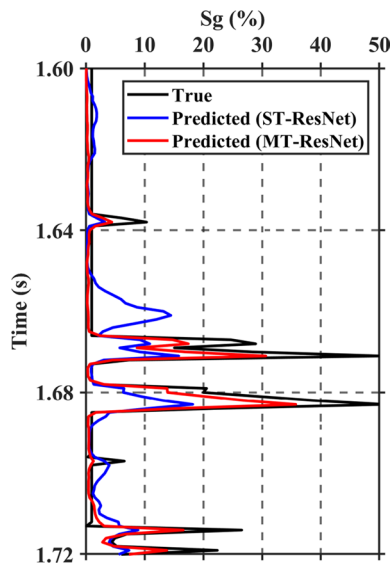


Fig. 12 Comparisons between ST-ResNet and MT-ResNet predicted GS curves at the blind well locations of w4. The black, blue, and red lines represent the true, ST-ResNet inverted, and MT-ResNet inverted GS curves, separately

construct the training set of the field data for the ST-ResNet. Four derived PI curves, low-frequency PI curves, and the training set of the ST-ResNet are utilized to establish the training set of the MT-ResNet. The preparation process of training and test sets and the network architecture of the ST-ResNet or MT-ResNet for the field data is the same as that adopted in the synthetic data. Once the two networks are optimized via 300 epochs, we apply them to the whole test data set.

Figure 11a–d shows the poststack seismic profile, ST-ResNet inverted GS profile, MT-ResNet inverted PI profile, and MT-ResNet inverted GS profile around the reservoir units. The vertical pillars in Fig. 11b–d refer to true PI or GS curves at five well locations. Compared with the seismic data in Fig. 11a, estimated GS results (Fig. 11b and d) via two methods show higher vertical and lateral resolution. Both of them indicate that the bottom of the reservoir section is the domain gas-bearing area, and the top of the reservoir section lacks gas. Figure 11b depicts that the ST-ResNet based GS results identify merely one gas-bearing strata. By comparison, the MT-ResNet-based GS results (Fig. 11d) can successfully identify two gas-bearing strata near the wells of w2 and w4. In addition, compared with the ST-ResNet, the MT-ResNet has a higher coincidence in the blind well position of w4. The contrast of converted PI and GS results (Fig. 11c and d) illustrates that the high gas-bearing region corresponds to the relatively low PI, which is in line with the petrophysical relationship between GS and PI around the wells (Fig. 10). Figure 12 shows the estimated GS results via two methods at the blind test well of w4. In comparison with

the ST-ResNet inverted GS curve (the blue line), the MT-ResNet inverted GS curve (the red line) is more comparable to the true GS curve (the black line). The RMSE between the true GS curve and the ST-ResNet estimated GS curve is 0.07, while the RMSE between the true GS curve and the MT-ResNet estimated GS curve is 0.04. The field data test manifests that the MT-ResNet is more suitable for the actual application and it can utilize the inverted PI profile to further enhance the prediction accuracy and resolution of the target GS results. In addition, the retrieved PI and GS results are mutually verified and have the potential to decrease the risk of drilling decisions.

Discussions

In addition to the learning-based ST-ResNet, we compare our proposed MT-ResNet with the traditional non-learning-based method for GS determination. Conventional methods usually adopt prestack seismic inversion and GS calculation models for calculating the spatial distribution of GS. We first prepare 0–5 Hz initial models of elastic parameters and utilize noisy prestack seismic gathers of Fig. 4a to generate partially stacked angle stacks. Then we implement AVO inversion based on Tikhonov (TK) regularization (She et al. 2018), which primarily adopts the seismic data misfit and regularization term represented by specific prior assumptions to provide estimations of elastic parameters. By setting the regularization parameter as 0.1, the TK-constrained AVO inversion approach predicts V_p , V_s , and density. Figure 13a shows the inverted PI model based on estimated V_p and density models. The RMSE and SSIM between Fig. 13a and the true PI model (Fig. 1h) are 0.03 and 0.84, respectively. Figure 13b shows the difference between Figs. 13a and 1h. By comparing MT-ResNet and conventional AVO inversion estimated PI models (Figs. 6c and 13a), MT-ResNet shows less prediction deviation and is more stable to noise interference. Finally, we employ the nonlinear fitting formula of Eq. 1 to transform the PI model of Fig. 13a into the calculated GS model, as shown in Fig. 13c. The RMSE and SSIM between Fig. 13c and the true GS model (Fig. 1a) are 0.10 and 0.62, respectively. The inaccuracy estimation of PI and approximate rock-physics relation leads to that the estimated GS results of Fig. 13c are worse than the MT-ResNet retrieved GS results (Fig. 6e).

Figures 6, 7, 8, 9, 10, 11, 12 and 13 and Table 1 demonstrate that the proposed MT-ResNet method is superior to both ST-ResNet-based method and conventional non-learning method for GS estimation. Compared with these two methods, the MT-ResNet can not only realize the simultaneous inversion of PI and GS, but also improve the estimation accuracy and stability of GS by integrating seismic data and the GS-related elastic attribute. Apart from these advantages, some potential issues

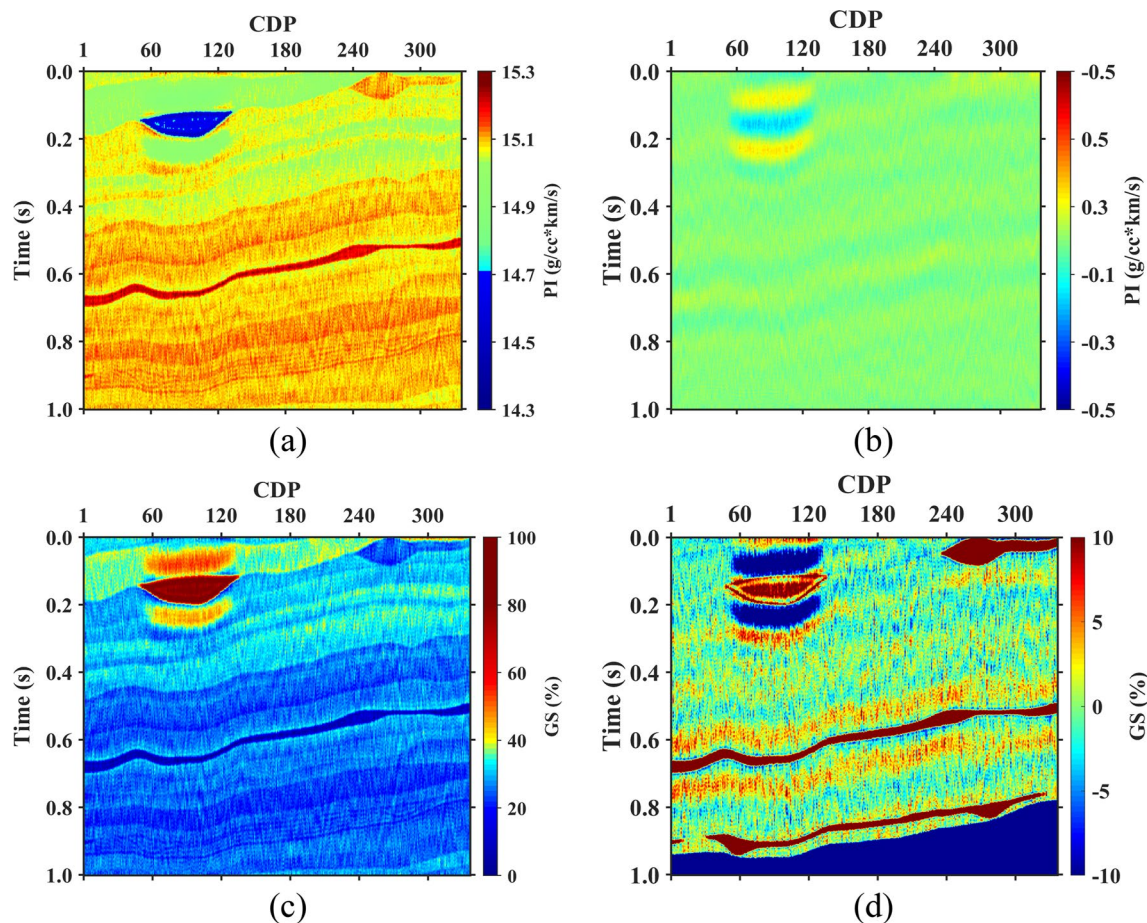


Fig. 13 Estimated parameter models using conventional AVO inversion and the rock-physics model. **a** The inverted PI model based on TK-regularized AVO inversion, **b** the differences between the true

PI model (Fig. 1h) and **(a), c** the GS model calculated from **(a)** via the rock-physics empirical formula of Eq. 1, and **d** the differences between the true GS model (Fig. 1a) and **(c)**

should be noticed and investigated in the future. First, the network parameters of the MT-ResNet are nearly as twice as the ST-ResNet, resulting in the former achieving the simultaneous prediction of two parameters (i.e., PI and GS) at the expense of twice as long as the time of the latter. In the future, we need to optimize the network structure and ensure both efficiency and precision of MTL. Second, the cross-plot of Figs. 2 and 10 shows that PI and GS have a complex relationship. Equation 1 only expresses the major relationship between PI and GS in Fig. 2. Therefore, we merely use PI as a sensitive attribute of a “soft” physical constraint to improve the accuracy of MT-ResNet estimated GS. If the relationship between PI and GS is relatively simple and straightforward (e.g., Fig. 2), an explicit formula like Eq. 1 can be derived and adopted as a physical constraint in the MT-ResNet. The initial GS results are calculated from the predicted PI of the first subnet of the MT-ResNet via this physical constraint. The initial GS results and seismic data can be further utilized to predict the final GS estimation via the second subnet of the MT-ResNet. If the relationship between PI and GS is complicated (e.g., Fig. 10),

more sophisticated rock-physics models and more gas-bearing sensitive properties can be used as the physical constraints to assist the estimation of GS. In the future, we will further improve our method and investigate the feasibility and effectiveness of above strategies for these potentially intractable situations. Lastly, as shown in Fig. 3a, we adopt a hierarchical learning framework to reduce the interference of different tasks as much as possible. Compared with hard parameter sharing and soft parameter sharing, the learned features of shallow layers (i.e., first subnet of MT-ResNet) will not directly influence the features of deep layers (i.e., second subnet of MT-ResNet) for the GS estimation task. In our hierarchical learning framework, the risk of task interference mainly comes from the inverted PI by the first task. The inaccurate PI may introduce negative information into the second task of GS inference and reduce the prediction performance of GS. Estimating a more accurate PI model can reduce the task distraction in our method. In addition, hard parameter sharing with appropriate shared layers and task-specific layers or soft parameter sharing with proper regularization between subnets also can be

studied to reduce the interference risk of multiple reservoir parameter tasks.

Conclusions

We propose a data-driven prestack simultaneous inversion framework (MT-ResNet) that integrates multi-source information including seismic data, well data, and the sensitive attribute for PI and GS estimation. Test on the synthetic data example indicates that MT-ResNet can obtain more accurate and stable PI and GS models compared with ST-ResNet and the traditional AVO inversion and rock-physics formulas-based method. PI and its low-frequency components diminish the negative influence of seismic random noise and improve the accuracy of MT-ResNet estimated GS. The field data example of a tight sandstone reservoir demonstrates that MT-ResNet can mine the reservoir information embedded in seismic waveform and generate reasonable PI and GS models, which conform to well-logging curves and mutually describe the potential gas-bearing regions of subsurface rocks. The future work is to attempt MT-ResNet to predict other multiple reservoir parameters and clarify the mechanism of ML-based multiple information integration for seismic inversion and reservoir characterization.

Acknowledgements This work was financially supported by the National Natural Science Foundation of China (U19B6003-004), the National Key R&D Program of China (2018YFA0702505), the National Natural Science Foundation of China (41974140), and R&D Department of China National Petroleum Corporation (2022DQ0604-01).

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Ahmad M, Haghghi M (2013) Water saturation evaluation of Murteer and Roseneath shale gas reservoirs, Cooper Basin, Australia using wire-line logs, focused ion beam milling and scanning electron microscopy. SPE Unconv Resour Conf Exhib. <https://doi.org/10.2118/167080-MS>
- Aki K, Richards PG (2002) Quantitative seismology. University Science Books
- Alardi M, Ciabbarri F, Gukov T (2018) A two-step inversion approach for seismic-reservoir characterization and a comparison with a single-loop Markov-chain Monte Carlo algorithm. *Geophysics* 83(3):R227–R244. <https://doi.org/10.1190/geo2017-0387.1>
- Archie GE (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans AIME* 146:54–62. <https://doi.org/10.2118/942054-G>
- Bosch M, Cara L, Rodrigues J, Navarro A, Díaz M (2007) A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes. *Geophysics* 72(6):O29–O39. <https://doi.org/10.1190/1.2783766>
- Bosch M, Mukerji T, Gonzalez EF (2010) Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: a review. *Geophysics* 75(5):75A165–75A176. <https://doi.org/10.1190/1.3478209>
- Bürkle PY, Azevedo L, Vellasco L (2023) Deep physics-aware stochastic seismic inversion. *Geophysics* 88(1):R11–R24. <https://doi.org/10.1190/geo2021-0551.1>
- Caruana R (1997) Multitask Learning. *Mach Learn* 28:41–75. <https://doi.org/10.1023/A:1007379606734>
- Caruana R (1993) Multitask learning: a knowledge-based source of inductive bias. In: International conference on machine learning, 41–48. <https://doi.org/10.1016/B978-1-55860-307-3.50012-5>
- Chen Y, Yu J, Zhao YT, Chen JY, Du XS (2022) Task’s choice: pruning based feature sharing (PBFS) for multi-task learning. *Entropy* 24(3):432. <https://doi.org/10.3390/e24030432>
- Connolly P (1999) Elastic impedance. *Lead Edge* 18(4):438–452. <https://doi.org/10.1190/1.1438307>
- Duong L, Cohn T, Bird S, Cook P (2015) Low resource dependency parsing: Cross-lingual parameter sharing in a neural network parser. In: Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing, pp 845–850. <https://doi.org/10.3115/v1/P15-2139>
- Eidsvik J, Avseth P, Omre H, Mukerji T, Mavko G (2004) Stochastic reservoir characterization using prestack seismic data. *Geophysics* 69(4):978–993. <https://doi.org/10.1190/1.1778241>
- Figueiredo LPD, Grana D, Bordignon FL, Santos M, Roisenberg M, B. Rodrigues B, (2018) Joint Bayesian inversion based on rock-physics prior modeling for the estimation of spatially correlated reservoir properties. *Geophysics* 83(5):M49–M61. <https://doi.org/10.1190/geo2017-0463.1>
- Gao JH, Song ZH, Gui JY, Yuan SY (2020) Gas-bearing prediction using transfer learning and CNNs: an application to a deep tight dolomite reservoir. *IEEE Geosci Remote Sens Lett* 19:1–5. <https://doi.org/10.1109/LGRS.2020.3035568>
- Gao Y, Zhang JL, Li H, Li GF (2022) Incorporating structural constraint into the machine learning high-resolution seismic reconstruction. *IEEE Trans Geosci Remote Sens* 60:1–12. <https://doi.org/10.1109/TGRS.2022.3157064>
- Goodway B, Chen TW, Downton J (1997) Improved AVO fluid detection and lithology discrimination using Lamé petrophysical parameters; “ $\lambda\rho$ ”, “ $\mu\rho$ ”, & “ λ/μ fluid stack”, from P and S inversions. In: SEG technical program expanded abstracts, pp 183–186. <https://doi.org/10.1190/1.1885795>
- Hampson DP, Russell BH, Bankhead B (2005) Simultaneous inversion of pre-stack seismic data. In: SEG Technical Program Expanded Abstracts, pp 1633–1637. <https://doi.org/10.1190/1.2148008>
- He KM, Zhang XY, Ren SQ, Sun J (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition, pp 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Ioffe S, Szegedy C (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. International conference on machine learning, pp 448–456. <https://arxiv.org/abs/1502.03167v3>
- Kingma DP, Ba JL (2014) Adam: a method for stochastic optimization. 3rd International conference for learning representation. <https://doi.org/10.48550/arXiv.1412.6980>
- Kuster GT, Toksöz MN (1974) Velocity and attenuation of seismic waves in two-phase media: Part I. Theoretical formulations. *Geophysics* 39(5):587–616. <https://doi.org/10.1190/1.1440450>
- Li YS, Song JY, Lu WK, Monkam P, Ao YL (2021) Multitask learning for super-resolution of seismic velocity model. *IEEE Trans*

- Geosci Remote Sens 59(9):8022–8033. <https://doi.org/10.1109/TGRS.2020.3034502>
- Li D, Peng SP, Guo YL, Lu YX, Cui XQ, Du WF (2023) Progressive multitask learning for high-resolution prediction of reservoir elastic parameters. *Geophysics* 88(2):M71–M86. <https://doi.org/10.1190/geo2022-0275.1>
- Liu XB, Shen GQ, Chen JY (2022) Amplitude variation with incidence angle inversion of fluid parameters in porous media using an accurate estimation of the Jacobian matrix. *Geophysics* 87(6):N95–N113. <https://doi.org/10.1190/geo2021-0612.1>
- Lucier AM, Hofmann R, Bryndzia LT (2011) Evaluation of variable gas saturation on acoustic log data from the Haynesville shale gas play, NW Louisiana, USA. *Lead Edge* 30(3):300–311. <https://doi.org/10.1190/1.3567261>
- Ma M, Zhang R (2023) Block sparse Bayesian learning-based prestack seismic inversion with the correlation of velocities and density. *Acta Geophys* 71:261–274. <https://doi.org/10.1007/s11600-022-00914-4>
- Martin GS, Wiley R, Marfurt KJ (2006) Marmousi2: An elastic upgrade for Marmousi. *Lead Edge* 25(2):156–166. <https://doi.org/10.1190/1.2172306>
- Mazzotti A, Zamboni E (2003) Petrophysical inversion of AVA data. *Geophys Prospect* 51:517–530. <https://doi.org/10.1046/j.1365-2478.2003.00389.x>
- Morgan EC, Vanneste M, Lecomte I, Baise LG, Longva O, McAdoo B (2012) Estimation of free gas saturation from seismic reflection surveys by the genetic algorithm inversion of a P-wave attenuation model. *Geophysics* 77(4):R175–R187. <https://doi.org/10.1190/geo2011-0291.1>
- Qi Q, Müller TM, Pervukhina M (2017) Sonic Q_p/Q_s ratio as diagnostic tool for shale gas saturation. *Geophysics* 82(3):MR97–MR103. <https://doi.org/10.1190/geo2016-0499.1>
- Qi QM, Fu LY, Deng JX, Cao JX (2021) Attenuation methods for quantifying gas saturation in organic-rich shale and tight gas formations. *Geophysics* 86(2):D65–D75. <https://doi.org/10.1190/geo2020-0291.1>
- Radke CJ, Gillis JV (1990) A dual gas tracer technique for determining trapped gas saturation during steady foam flow in porous media. *SPE Annu Tech Conf Exhib*. <https://doi.org/10.2118/20519-MS>
- Rezaee R (2015) *Fundamentals of gas shale reservoirs*. John Wiley & Sons, Hoboken
- Ruder S (2017) An overview of multi-task learning in deep neural networks. <https://doi.org/10.48550/arXiv.1706.05098>
- Sang WJ, Yuan SY, Yong XS, Jiao XQ, Wang SX (2021) DCNNs-based denoising with a novel data generation for multidimensional geological structures learning. *IEEE Geosci Remote Sens Lett* 10(4):1851–1865. <https://doi.org/10.1109/LGRS.2020.3007819>
- Sang WJ, Yuan SY, Han HW, Liu HJ, Yu Y (2023) Porosity prediction using semi-supervised learning with biased well log data for improving estimation accuracy and reducing prediction uncertainty. *Geophys J Int* 232(2):940–957. <https://doi.org/10.1093/gji/ggac371>
- Schlumberger (1989) *Schlumberger log interpretation principles/applications*. Schlumberger Education Services.
- She B, Wang YJ, Liang JD, Liu ZN, Song CY, Hu GM (2018) Adata-driven amplitude variation with offset inversion method via learned dictionaries and sparse representation. *Geophysics* 83(6):R725–R748. <https://doi.org/10.1190/geo2017-0615.1>
- Singh V, Hegazy M, Fontanelli L (2009) Assessment of reservoir uncertainties for development evaluation and risk analysis. *Lead Edge* 28(3):272–282. <https://doi.org/10.1190/1.3104070>
- Song ZH, Yuan SY, Li ZM, Wang SX (2022) k NN-based gas-bearing prediction using local waveform similarity gas-indication attribute—an application to a tight sandstone reservoir. *Interpretation* 10(1):SA25–SA33. <https://doi.org/10.1190/INT-2021-0045.1>
- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612. <https://doi.org/10.1109/TIP.2003.819861>
- Weinzierl W, Wiese B (2021) Deep learning a poroelastic rock-physics model for pressure and saturation discrimination. *Geophysics* 86(1):MR53–MR66. <https://doi.org/10.1190/geo2020-0049.1>
- Wu XM, Liang LM, Shi YZ, Geng ZC, Fomel S (2019) Multitask learning for local seismic image processing: Fault detection, structure-oriented smoothing with edge-preserving, and seismic normal estimation by using a single convolutional neural network. *Geophys J Int* 219:2097–2109. <https://doi.org/10.1093/gji/ggz418>
- Wu XM, Yan SS, Bi ZF, Zhang SB, Si HJ (2021) Deep learning for multidimensional seismic impedance inversion. *Geophysics* 86(5):R735–R745. <https://doi.org/10.1190/geo2020-0564.1>
- Xu JL, Xu L, Qin YX (2017) Two effective methods for calculating water saturations in shale-gas reservoirs. *Geophysics* 82(3):D187–D197. <https://doi.org/10.1190/geo2016-0462.1>
- Yu SW, Ma JW, Wang WL (2019) Deep learning for denoising. *Geophysics* 84(6):V333–V350. <https://doi.org/10.1190/geo2018-0668.1>
- Yu C, Zhang CM, Shi WR, Li ML, Yan W (2022) Research and comparison of evaluation methods for shale reservoir saturation in Longtan Formation of marine-continent transition facies. *Geophysics* 87(6):M293–M305. <https://doi.org/10.1190/geo2021-0740.1>
- Yuan SY, Jiao XQ, Luo YN, Sang WJ, Wang SX (2022) Double-scale supervised inversion with a data-driven forward model for low-frequency impedance recover. *Geophysics* 87(2):R165–R181. <https://doi.org/10.1190/geo2020-0421.1>
- Zhang K, Lin NT, Tian GP, Yang JQ, Wang DY, Jin ZW (2022a) Unsupervised-learning based self-organizing neural network using multi-component seismic data: application to Xujiahe tight-sand gas reservoir in China. *J Petrol Sci Eng* 209:109964. <https://doi.org/10.1016/j.petrol.2021.109964>
- Zhang K, Lin NT, Yang JQ, Jin ZW, Li GH, Ding RW (2022b) Predicting gas-bearing distribution using DNN based on multi-component seismic data: quality evaluation using structural and fracture factors. *Pet Sci* 19(4):1566–1581. <https://doi.org/10.1016/j.petsci.2022.02.008>
- Zhang J, Zhao XY, Chen YK, Sun H (2023) Domain knowledge-guided data-driven prestack seismic inversion using deep learning. *Geophysics* 88(2):M31–M47. <https://doi.org/10.1190/geo2021-0560.1>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.