Computers & Geosciences 188 (2024) 105622

Contents lists available at ScienceDirect

Computers and Geosciences

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

Deep generative networks for multivariate fullstack seismic data inversion using inverse autoregressive flows

Roberto Miele [a,*], Shiran Levy [b], Niklas Linde [b], Amilcar Soares [a], Leonardo Azevedo [a]

[a] DER/CERENA, Instituto Superior Técnico, 1049-001, Lisbon, Portugal
[b] Institute of Earth Sciences, University of Lausanne, CH-1015, Lausanne, Switzerland

ARTICLE INFO

Keywords:
Stochastic seismic inversion
Multivariate modeling
Inverse autoregressive flows
Uncertainty quantification

ABSTRACT

The simultaneous prediction of the subsurface distribution of facies and acoustic impedance (I_P) from fullstack seismic data requires solving an inverse problem and is fundamental in natural resources exploration, carbon capture and storage, and environmental risk management. In recent years, deep generative models (DGM), such as variational autoencoders (VAE) and generative adversarial networks (GAN), were proposed to reproduce complex facies patterns honoring prior geological information. Variational Bayesian inference using inverse autoregressive flows (IAF) can be performed to infer the solution to a geophysical inverse problem from the encoded latent space of such pre-trained DGM. Successful applications of such approach on crosshole ground-penetrating radar synthetic data inversion demonstrated that the technique's accuracy is comparable to that of Markov chain Monte Carlo (MCMC) inference methods, while significantly reducing the computational cost. Nonetheless, these application examples did not account for the spatial uncertainty affecting the facies-dependent continuous physical property, from which the geophysical data are calculated. This uncertainty can significantly affect the inversion accuracy and its applicability to real data. In this work, specific VAE and GAN architectures are proposed to simultaneously predict facies and co-located I_P, while accounting for their spatial uncertainties. The two types of generative networks are used in Bayesian inversion with IAF for the inversion of seismic data. The results are found to reproduce the statistics of the training images and solve the seismic inversion problem accurately, comparably to MCMC inversion. Furthermore, advantages and limitations of the two DGMs are evaluated by comparing the results obtained.

1. Introduction

Predicting the spatial distribution of subsurface rock types, or facies, and collocated rock's acoustic impedance (I_P) is a main requirement for the evaluation of natural resources potential, or for fluid flow simulations in environmental tasks (e.g., Grana et al., 2021; Linde et al., 2015; Strebelle, 2021). In geophysical inverse modeling, fullstack seismic reflection data is often used to this purpose, with the I_P distribution being dependent on the correlation between facies and seismic domains (Doyen, 2014; Grana et al., 2021; Tarantola, 2005). Seismic inversion problems are ill-posed and account for non-unique solutions, therefore assessing the uncertainty on the predictions is key for well-informed decisions (e.g., Grana et al., 2022). Several probabilistic inversion algorithms are based, e.g., on Bayesian inference (e.g., Buland and Omre, 2003; Connolly and Hughes, 2016; De Figueiredo et al., 2019, 2018; Grana et al., 2017; Grana and Della Rossa, 2010; Sen and Stoffa, 1996) or stochastic optimization based on geostatistical simulations (Azevedo and Soares, 2017; González et al., 2008; Grana et al., 2012; Miele et al., 2022, 2023).

Within this framework, geostatistical modeling allows to reproduce complex subsurface scenarios. Modeling methods based on multiple-points statistics (MPS) (Mariethoz et al., 2010; Strebelle, 2002, 2021) is particularly suitable for facies modeling, as they sample high-order geological patterns from a prior training image (TI). Nonetheless, their use in inverse modeling is computationally expensive due to the large number of subsurface parameters to be optimized (e.g., González et al., 2008).

Deep generative models (DGMs), such as generative adversarial networks (GANs) (Goodfellow et al., 2014) or variational autoencoders (VAEs) (Kingma and Welling, 2014) were proposed as valid alternatives for dimensionality reduction in facies modeling (Laloy et al., 2017, 2018). Analogously to MPS, such DGMs encode complex facies patterns

* Corresponding author.
E-mail address: roberto.miele@tecnico.ulisboa.pt (R. Miele).

https://doi.org/10.1016/j.cageo.2024.105622
Received 8 February 2024; Received in revised form 26 March 2024; Accepted 16 May 2024
Available online 17 May 2024

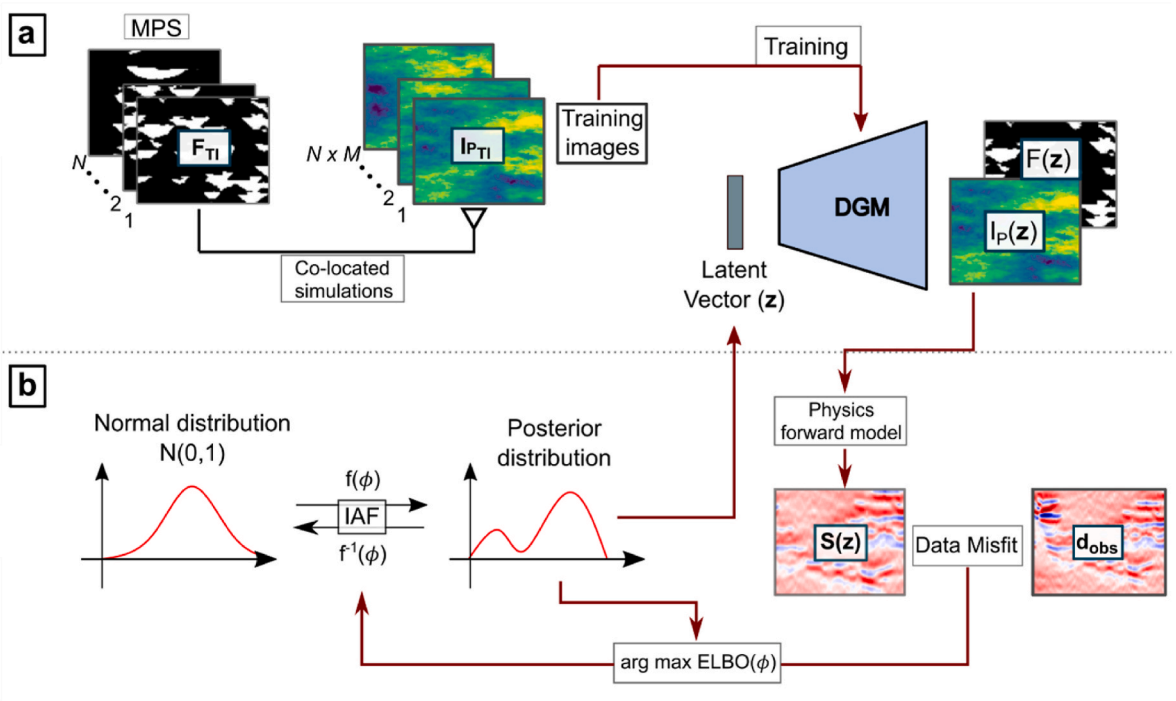

Fig. 1. Schematic representation of the proposed method. (a) The DGM are first trained to reproduce facies and collocated I_P distributions. (b) The seismic inversion solutions are found by means of IAF.

stored in a TI, into a low-dimensional latent space (Azevedo et al., 2020; Dupont et al., 2018; Jordão et al., 2023; Mosser et al., 2018; Pan et al., 2021; Zhang et al., 2021). This represents a main advantage for inverse modeling, as the trained networks reduce the number of unknowns of the inverse problem. The trained DGMs' parametrization of the prior can be used in a subsequent step in a Bayesian inverse modeling framework. The posterior distribution conditioned on the observed data ($\mathbf{d}_{obs}$) can be inferred by either sampling methods, for example, using Markov chain Monte Carlo (MCMC) methods (Laloy et al., 2017, 2019; Levy et al., 2021; Mosser et al., 2020) or by optimization approaches (Chan and Elsheikh, 2019). The latter approach uses variational inference (VI) (Blei et al., 2017) to transform iteratively a surrogate density distribution by maximizing the evidence lower bound (ELBO).

The normalizing flows method (Rezende and Mohamed, 2015) consists in modeling such transformation through a sequence of invertible parametric functions (Hoffman et al., 2019; Kingma et al., 2017; Levy et al., 2023; Papamakarios et al., 2021). Levy et al. (2023) propose the use of inverse autoregressive flows (IAF) networks (Kingma et al., 2017) as a type of normalizing flows, to be used with pre-trained DGMs for geophysical inversion. They refer to this approach as neural transport (NT) (Hoffman et al., 2019). The method was successfully tested for the inference of facies patterns in synthetic crosshole ground-penetrating-radar (GPR) data inversion; constant facies-dependent velocities were used in the physics forward modeling. Compared to MCMC methods, the NT demonstrated comparable accuracy and improved efficiency, although showing expected limitations when handling the non-linear transformations of GANs (Laloy et al., 2019; Lopez-Alvis et al., 2021).

The NT method, as well as similar approaches for facies inverse modeling with DGMs (Laloy et al., 2017, 2019; Lopez-Alvis et al., 2021; Mosser et al., 2020), overlook the spatial uncertainty of facies-dependent continuous properties responsible for the observed geophysical data, i.e., spatial petrophysical and elastic properties uncertainty (e.g., Brunetti and Linde, 2018). In complex scenarios and real case studies, such uncertainty can significantly impact inversion results and should be incorporated into the subsurface modeling (Bosch et al., 2010; Brunetti and Linde, 2018; Doyen, 2007; Friedli and Linde, 2023; Grana et al., 2022).

This work builds upon that of Levy et al. (2023) and proposes two specific DGMs based on VAE and GAN architectures, to model a multivariate prior of facies patterns and collocated I_P spatial distributions, including the elastic property's uncertainty patterns assumed. The trained DGMs are found to be comparable to geostatistical facies simulation based on MPS and co-simulation based on two-point statistics (Nunes et al., 2017). Then, the use of NT with the proposed DGMs is demonstrated in cases of fullstack seismic inversion for the inference of facies patterns in presence of I_P spatial uncertainty. The efficiency and accuracy of the proposed method is further compared to those of MCMC-based inversion.

2. Method

Our methodology builds upon the work of Levy et al. (2023), who used the NT method in combination with pre-trained VAE and GAN generative networks for facies modeling, to infer the distribution of subsurface parameters. The DGMs proposed herein are first trained using a training data set reproducing both facies patterns and the collocated I_P spatial uncertainty (Section 2.1). The considered VAE and GAN networks (Section 2.2) are trained to simultaneously model both domains. The NT method (Section 2.3) is then used on the trained DGMs for inverse modeling using fullstack seismic data (Section 2.4). A schematic representation of the DGM training and the subsequent inversion process is provided in Fig. 1.

2.1. Multivariate training data set

The training dataset used represents a multivariate prior distribution of facies (categorical) and collocated I_P (continuous) patterns, honoring assumed prior data statistics obtained, e.g., from direct well-logs observations and prior knowledge on the subsurface geology. The expected facies distributions are represented through a set of N unconditional MPS realizations, honoring the spatial statistics of a geological

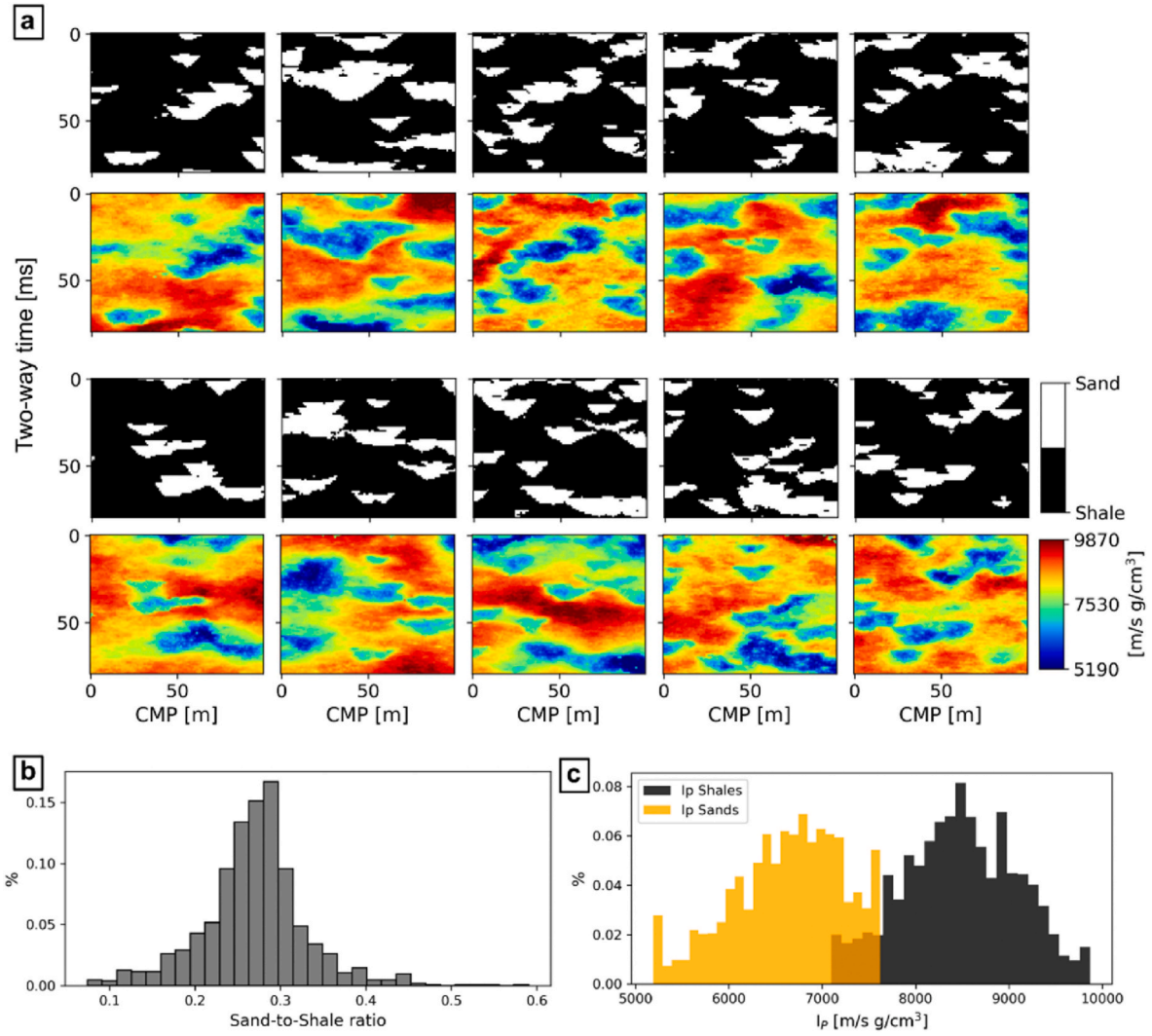

Fig. 2. Representation of training data set; (a) samples of pairs of facies and collocated I_P; (b) sand-to-shale ratio distribution from the facies samples; (c) facies-dependent distribution of I_P assumed and honored in each I_P realization.

conceptual model used as TI (Mariethoz et al., 2010). The I_P data distribution is represented for each of the N facies realizations, using geostatistical direct sequential co-simulation with multi-local distribution functions (co-DSS) (Nunes et al., 2017). At each location of the simulation grid, the I_P realizations honor the data histogram and variogram models defined for each conditioning facies class. To represent the spatial uncertainty in this conditioned domain, a set of M equiprobable I_P realizations are generated for each of the N facies. Therefore, the resulting training data set is a nested set of facies and corresponding I_P distributions (Fig. 1a). During the training of the DGMs, for each epoch, the N facies patterns are paired with a single corresponding I_P realization, randomly sampled from the M realizations. The resulting distribution is used as training data set for that epoch.

2.2. Deep generative adversarial networks

2.2.1. VAE

The VAE architecture introduced for facies and I_P modeling is derived from the framework presented by Lopez-Alvis et al. (2021). Here, the encoder is composed of a series of convolutional layers followed by a fully connected layer, while the discriminator shares a specular architecture, using a sequence of transposed convolutions. The proposed architecture incorporates two main modifications. First, the number of convolutional layers within both the encoder and decoder is increased from four to six. Second, two parallel branches in the last two layers of the decoder are used, for the independent generation of facies and I_P data. The encoder's input is an image with two channels, representing facies and collocated I_P, respectively, which features are mapped into a 1D latent space of 60 variables. A schematic representation of the VAE architecture is given in the Supplementary material (Fig. A1).

The optimization of the network's parameters follows the common practice (e.g., Kingma and Welling 2014; Lopez-Alvis et al., 2021) of maximizing the ELBO

$$L(\theta,\Theta)=\mathbb{E}_{q_\theta(\mathbf{z}|\mathbf{X})}[\log(p_\Theta(\mathbf{X}|\mathbf{z}))]-D_{KL}(q_\theta(\mathbf{z}|\mathbf{X})\|p(\mathbf{z})), \tag{1}$$

where $q_\theta(\mathbf{z}|\mathbf{X})$ and $p_\Theta(\mathbf{X}|\mathbf{z})$ represent the parametrical transformations of the encoder and the decoder, respectively, and KL is the Kullback-Leibler divergence (Kullback and Leibler, 1951) between the encoded latent vector and a target standard normal distribution. Following Lopez-Alvis et al. (2021), the first term of Eq. (1) is approximated by the mean squared error estimator (MSE) between the generated images and the input data. In the multivariate case, Eq. (1) is rewritten as

$$L(\theta,\Theta)=-\frac{1}{N}\sum_{i=1}^{N}\frac{\|g_\Theta(\mathbf{z})_1-\mathbf{x}_1\|}{N}\alpha_1-\frac{1}{N}\sum_{i=1}^{N}\frac{\|g_\Theta(\mathbf{z})_2-\mathbf{x}_2\|}{N}\alpha_2 - \beta\, D_{KL}(q_\theta(\mathbf{z}|\mathbf{X})\|p(\mathbf{z})), \tag{2}$$

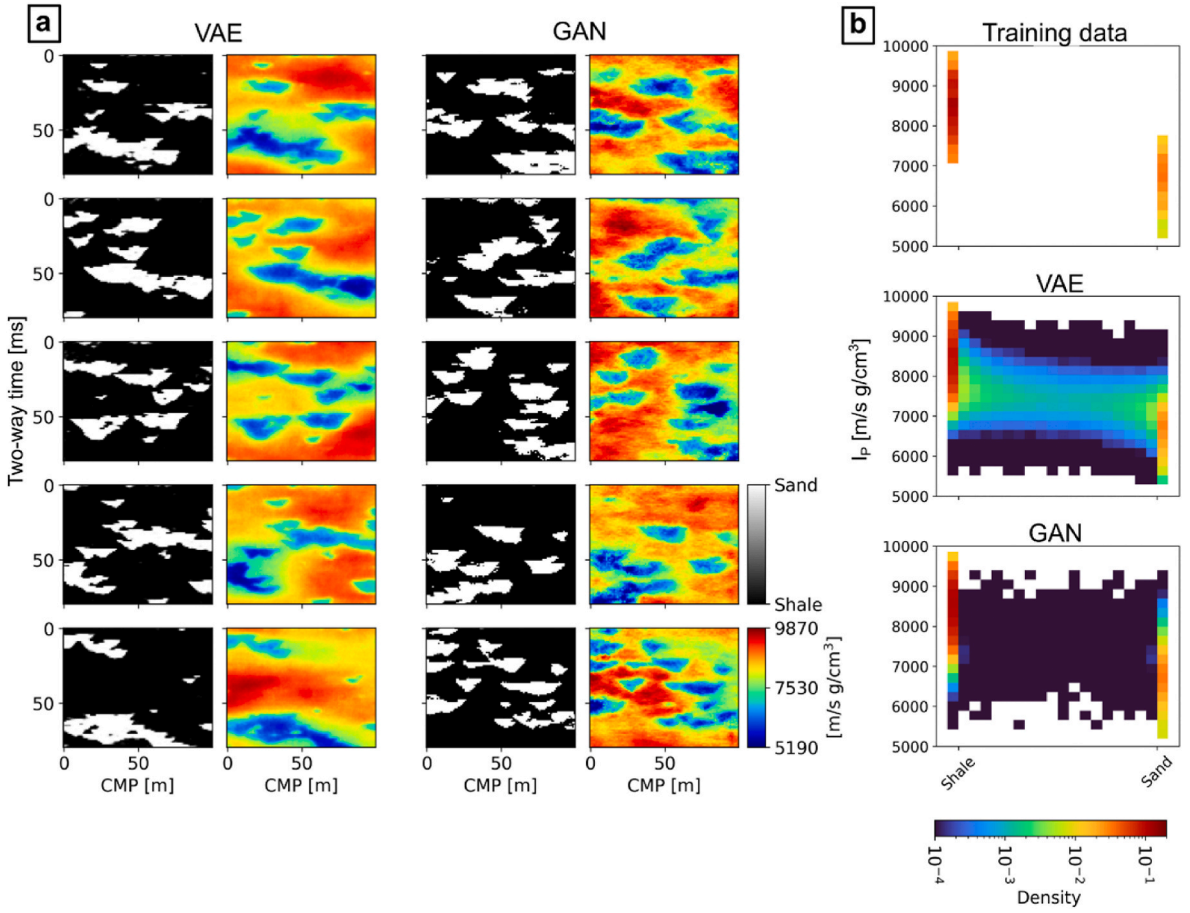

Fig. 3. (a) Five realizations of facies and collocated I_P generated by the trained VAE and GAN; (b) comparison between the training (i.e., target) multivariate distribution and 100 random realizations of the VAE and GAN.

where $g_\Theta(\mathbf{z})$ is the VAE's decoder, $\mathbf{x} \in \mathbf{X}$, the labels 1 and 2 represent the facies and the I_P data domains, respectively; α_1 and α_2 are rescaling factors used to keep the distances (for facies and I_P, respectively) within the same orders of magnitude; and β is the regularization weight used to better condition the encoder mapping into the desired target distribution $p(\mathbf{z})$. Following Lopez-Alvis et al. (2021), the MSE loss is used for the facies domain instead of the more common binary cross-entropy (see Lopez-Alvis et al. (2021) for further details).

2.2.2. GAN

The GAN used in this work is composed of an unconditional generator (G) with an architecture that is analogous to the VAE decoder described above, accounting for five transposed convolutional layers followed by two separate branches of two transposed convolutional layers each. Here, a latent vector ($\mathbf{z}$) of 60 variables is mapped into high-dimensional facies and I_P $G(\boldsymbol{z})$ images. Differently from conventional GANs, the discriminator uses an architecture with three branches, designed to encode the features of facies and I_P domains both separately and jointly. Here, two branches encode the input into two scores, one for the facies domain (D_F), one for the I_P domain (D_{I_P}). At each layer, they also pass the encoded information to the middle convolutional branch through skip connections, where the joint features from the two domains are encoded together to evaluate a joint distribution score (D_J). While the first two *external* branches can be seen as conventional GAN's discriminators, the central one allows to link two domains, reinforcing the simultaneous learning of marginal and joint features in complex multivariate scenarios. A schematic representation of the GAN architecture is given in the Supplementary material (Fig. A2).

The training objective of the GAN is defined as follows. First, the discriminator is trained to maximize its three output scores when the input data are the training images samples (i.e., the real samples) and to minimize them when the input are the generated models (i.e., fake samples). Moreover, the learning process is further reinforced by evaluating cases where fake and real samples from the two domains are combined (i.e., evaluate the score when the input is a combination of fake facies samples with real I_P samples and vice-versa). In such cases, the D is trained to maximize only the score corresponding to the real input and minimize the other two. The corresponding loss function for the discriminator training (L_D) is defined as

$$\begin{aligned} L_D = &\left(-\log[D_F(\mathbf{x_1})] - \log[D_J(\mathbf{x_1}, \mathbf{x_2})] - \log[D_S(\mathbf{x_2})]\right) + \left(-\log\left[1 \right.\right. \\ &\left. - \left(D_F\left(G(\mathbf{z})_1\right)\right)\right] - \log\left[1 - \left(D_J\left(G(\mathbf{z})_1, G(\mathbf{z})_2\right)\right)\right] - \log\left[1 \right. \\ &\left.\left. - \left(D_F\left(G(\mathbf{z})_2\right)\right)\right]\right) + \left(-\log\left[1 - \left(D_F\left(G(\mathbf{z})_1\right)\right)\right] - \log\left[1 \right.\right. \\ &\left. - \left(D_J\left(G(\mathbf{z})_1, \mathbf{x_2}\right)\right)\right] - \log[D_S(\mathbf{x_2})]\right) + \left(-\log[D_F(\mathbf{x_1})] - \log\left[1 \right.\right. \\ &\left.\left. - \left(D_J\left(\mathbf{x_1}, G(\mathbf{z})_2\right)\right)\right] - \log\left[1 - \left(G(\mathbf{z})_2\right)\right]\right), \end{aligned} \tag{3}$$

where $\mathbf{x} \in \mathbf{X}$, the labels 1 and 2 represent the facies and the I_P data domains, respectively. Given the training conditions of D, the score D_J can be maximized only if both input images are realistic and their joint distribution of features is the same as the one represented in the training data set. Therefore, the training of G aims at the maximization of the D_J score only.

2.2.3. Code implementation and training

The networks were implemented in Python 3.8.13 using PyTorch 1.13 with CUDA 11.7. Both the VAE and the GAN were trained using a training data set of 1024 facies model realizations, each coupled with 16

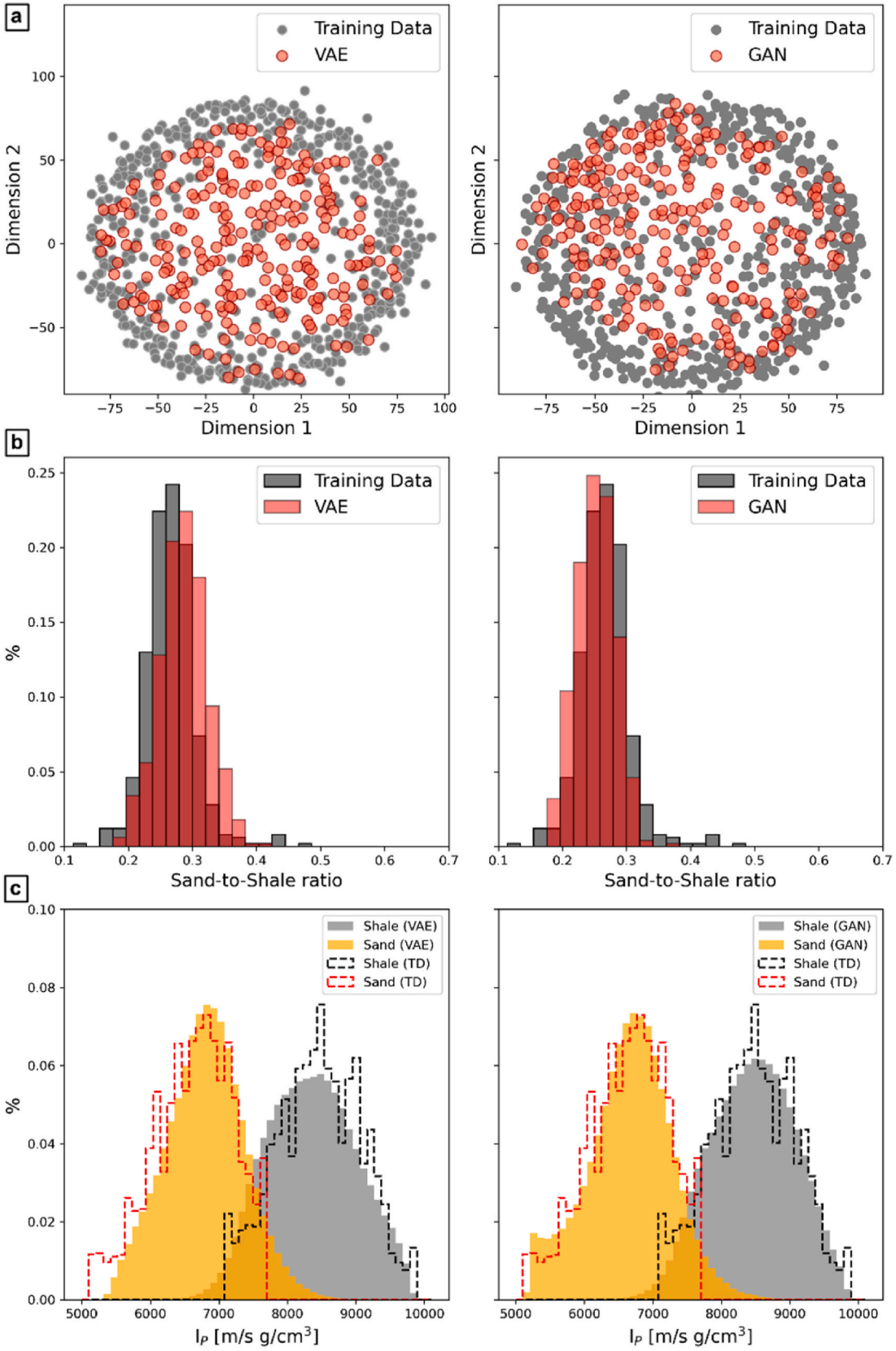

Fig. 4. Comparison between 250 models retrieved from the training data set and DGM realizations: (a) MDS representation of the pointwise Euclidean distances between facies models; (b) comparison between sand-to-shale ratio from the training data set and predicted with the DGM; (c) comparison between facies-dependent I_P distribution from the training data set and predicted with the DGM.

Table 1
KL divergences between the I_P density distributions in the training data set (TD) and reproduced by the DGM (Fig. 4c); and average SSIM values of generated I_P and facies images compared to the geostatistical references (see section 3.1.2).

	VAE	GAN
$I_{P\,shale}$ $KL(DGM\|\|TD)$	0.038	0.024
$I_{P\,sand}$ $KL(DGM\|\|TD)$	0.116	0.015
$SSIM_F$	0.87±0.03	0.73±0.03
$SSIM_{IP}$	0.68±0.03	0.54±0.03

collocated geostatistical realizations of I_P. The networks were trained for 1000 epochs, using ADAM optimizer (Kingma and Ba, 2017), considering an initial learning rate of 0.001 with a step decay of 50% every 100 epochs. The VAE is trained using $\beta = 0.03$, $\alpha_1 = 1$, $\alpha_2 = 5$.

The training process ran on a computer with Windows 11, with a CPU Intel® Core™ i7-8750H and a GPU NVIDIA® GeForce™ GTX 1060. Using PyTorch with CUDA parallelization, the training of the VAE took 40 s per epoch, while the training of the GAN took 55 s.

2.3. Neural transport with inverse autoregressive flows

This section summarizes the neural transport (NT) method as proposed by Levy et al. (2023) in the context of DGM. We refer to the original work for a detailed description of the methodology.

Neural transport is a type of VI method using normalizing flows (Rezende and Mohamed, 2015). It considers a set of continuous, invertible and volume conserving transformations applied to transform a known probability density function, or PDF, (here chosen as the prior distribution) into a more complex one (i.e., the targeted posterior). This is possible by repeatedly transforming a random variable $\mathbf{z}$ from the base distribution $q(\mathbf{z})$ into a random variable $\mathbf{m}$, from $q(\mathbf{m})$, by applying the rule for change of variables

$$q(\mathbf{m}) = \mathrm{q}(\mathbf{z}) \left| \det \frac{d\, f(\mathbf{z})}{d\, \mathbf{z}} \right|^{-1}, \tag{4}$$

where the second term represents the relative change in volume after the transformation f. In NT, f is represented by a sequence of transformations, known as inverse autoregressive flows (IAF) (Kingma et al., 2017), represented by scale and shift functions having trainable parameters φ. In IAF each transformation is conditional to the previous instances. Given f, a sequence of K transformations using the logarithmic form of Eq. (4) is defined as

$$\log q(\mathbf{m}) = \log \mathrm{q}\left(\mathbf{z}^{(0)}\right) - \sum_{k=1}^{K} \log \left| \det \frac{d\, f_k\left(\mathbf{z}^{(k-1)}\right)}{d\, \mathbf{z}^{(k-1)}} \right|, \tag{5}$$

where $\mathbf{z}^{(0)}$ is the random variable sampled from the base distribution.

In NT, the parameters φ of the IAF are iteratively updated by means of gradient-based optimization using variational Bayesian inference, where the posterior distribution is approximated by minimizing the ELBO (Blei et al., 2017), defined as

$$\mathrm{L}(\varphi) = \mathbb{E}_{\mathbf{m}\sim \mathrm{q}} \left[\log \frac{p(\mathbf{m}, \mathbf{d})}{q_\varphi(\mathbf{m})} \right], \tag{6}$$

where $p(\mathbf{m}, \mathbf{d})$ is the joint distribution between the model parameters $\mathbf{m}$ (i.e., the target random variable of the IAF) and the observed data $\mathbf{d}$. Maximizing the ELBO (Eq. (6)) minimizes the KL divergence between the surrogate distribution $q_\varphi(\mathbf{m})$ and the targeted, intractable posterior. The joint distribution $p(\mathbf{m}, \mathbf{d})$ is evaluated using the high-dimensional domain of the DGM, as described in the following section.

2.4. Fullstack seismic data inversion with neural transport

After training the two proposed DGMs to reproduce the prior, the seismic data inversion is carried out by optimizing the IAF parameters. First, N_s random draws (or *particles*) are sampled from the base distribution and transformed into a first approximation of the surrogate distribution $q_\varphi(\mathbf{m})$. The transformed distribution is then mapped into N_s realizations of facies and collocated I_P using the trained DGM ($G(\mathbf{m})$). In the forward modeling used for fullstack seismic data inversion, we assume that $\mathbf{d}_{\mathrm{obs}}$ is the result of the normal incidence of a plane wave (or source wavelet) with a series of interfaces determined by I_P contrasts of impedance in the inversion grid (Sen, 2006). Hence, the I_P are transformed into the corresponding seismic reflection data ($g(G(\mathbf{m}))$) by calculating the contrasts of impedance between one location (i) and the subsequent one (i + 1),

$$\mathbf{r}(i) = \frac{\mathrm{I_{P(i+1)}} - \mathrm{I_{P(i)}}}{\mathrm{I_{P(i+1)}} + \mathrm{I_{P(i)}}}, \tag{7}$$

and then convolving the resulting $\mathbf{r}(i)$ with a known source seismic wavelet (Sen, 2006).

The parameters of the IAF are evaluated using the ELBO loss in Eq. (6), with the joint distribution $p(\mathbf{m}, \mathbf{d})$ being calculated using the transformed latent and the observed seismic data $\mathbf{d}_{\mathrm{obs}}$ (i.e., $\mathbf{d} = \mathbf{d}_{\mathrm{obs}}$). The logarithmic form of the joint distribution is expressed as $\log p(\mathbf{m}, \mathbf{d}) = \log p(\mathbf{d}|\mathbf{m}) + \log p(\mathbf{m})$. By assuming the measurement data errors to be uncorrelated and follow a normal distribution with zero mean, the log-likelihood function $\log p(\mathbf{d}|\mathbf{m})$ can be expressed as (Tarantola, 2005)

$$\log p(\mathbf{d}|\mathbf{m}) = -\frac{N_\mathrm{d}}{2} \log(2\pi) - \frac{1}{2} \log\left(\prod_{i=1}^{N_\mathrm{d}} \sigma_i^2 \right) - \frac{1}{2} \sum_{i}^{N_\mathrm{d}} \left(\frac{g_i(G(\mathbf{m})) - \mathbf{d}_\mathrm{i}}{\sigma_i} \right)^2, \tag{8}$$

where N_d is the number of samples in $\mathbf{d}$ and $\sigma_i{}^2$ is the assumed variance of the data error for the i-th datum. To represent errors in the data domain due to measurement errors and approximations to the physical system under investigation during data processing, the standard deviation σ_i is defined as the sum of relative errors with standard deviation r and a given constant σ_C: $\sigma_i = r\, \mathbf{d}_\mathrm{i} + \sigma_C$. Since both DGMs use a standard-normal prior with mean $\mu_m = 0$ and standard deviation $\sigma_m = 1$, $\log p(\mathbf{m}, \mathbf{d})$ can be expressed as

$$\log p(\mathbf{m}, \mathbf{d}) = -\frac{1}{2} \left(N_d \log(2\pi) + \log\left(\prod_{i=1}^{N} \sigma_i^2 \right) + \sum_{i}^{N} \left(\frac{g_i(G(\mathbf{m})) - \mathbf{d}_\mathrm{i}}{\sigma_i} \right)^2 + N_m \log(2\pi) + 2\, N_m \log(\sigma_m) + \sigma_m^{-2} \sum_{i=1}^{N_z} m_i^2 \right), \tag{9}$$

where N_z is the number of $\mathbf{m}$ parameters, equal to 60 for the proposed DGM. The loss function used to optimize the parameters of the IAF are expressed using Eq. (6), as

$$\mathrm{L}(\varphi) = \mathbb{E}_{\mathbf{m}\sim \mathrm{q}} \left[\log \frac{p(\mathbf{m}, \mathbf{d})}{\log \mathrm{q}(\mathbf{z}^{(0)}) - \sum_{k=1}^{K} \log \left| \det \frac{d\, f_k\left(\mathbf{z}^{(k-1)}\right)}{d\, \mathbf{z}^{(k-1)}} \right|} \right] \tag{10}$$

In this work, the architecture of the IAF is the same as the one used by Levy et al. (2023). The IAF parameters are optimized using the ADAM optimizer using $\beta_{IAF_1} = 0.9$ and $\beta_{IAF_2} = 0.999$ (Kingma and Ba, 2017).

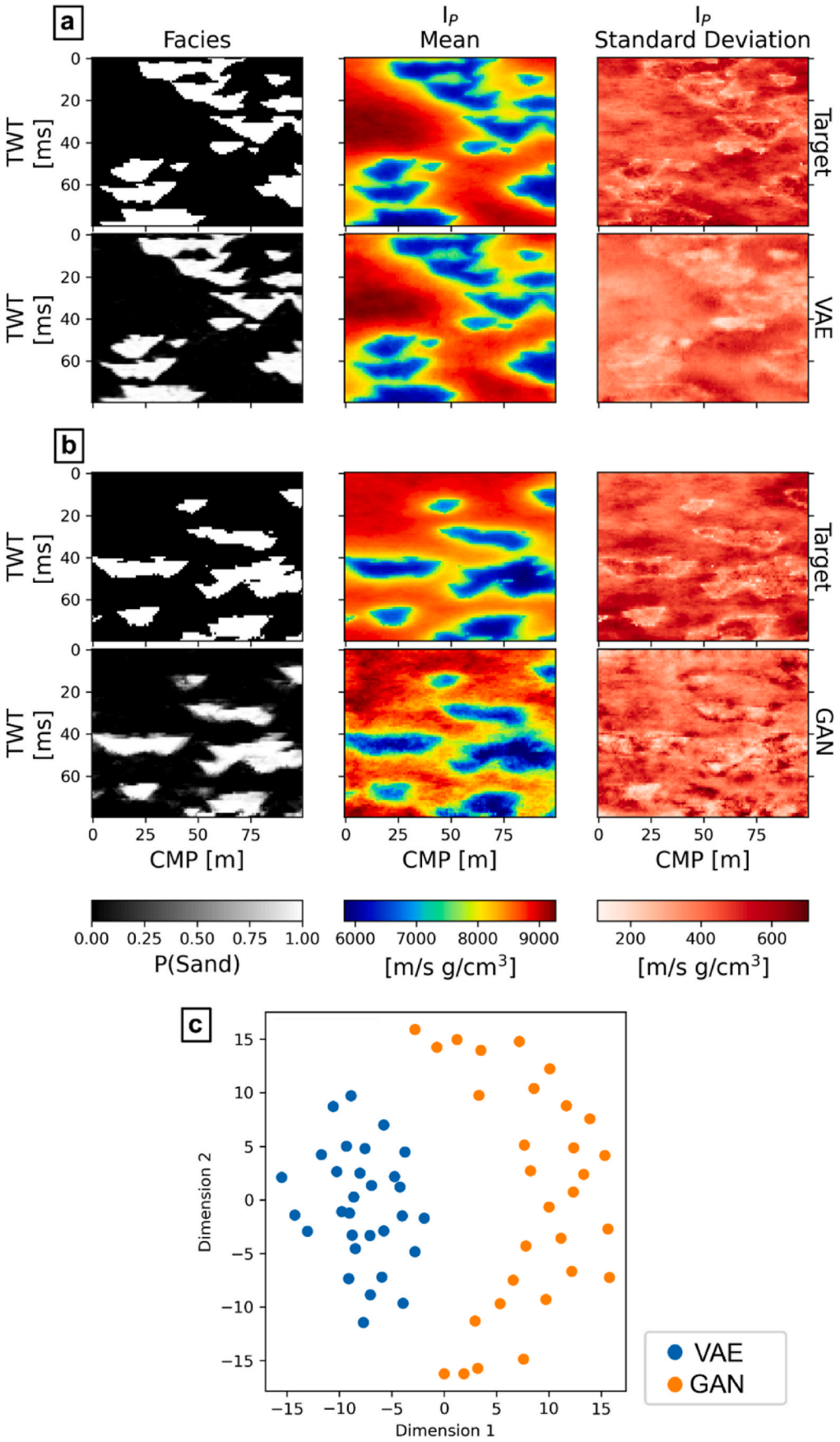

Fig. 5. Comparison between the I_P variance and spatial uncertainty in the prior distribution and that reproduced by the (a) VAE and (b) GAN; (c) comparison of the sampled **z** vectors for each I_P realization, projected in MDS space.

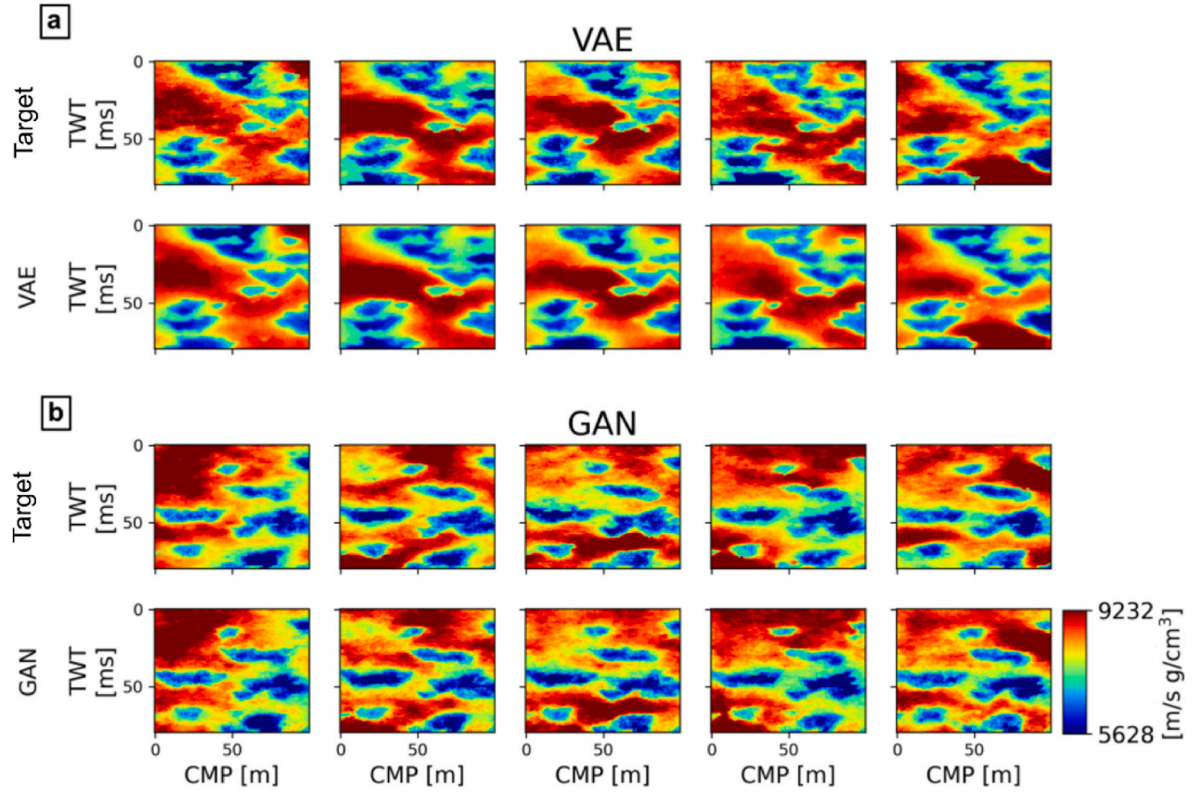

Fig. 6. Comparison between geostatistical co-simulations of I_P and corresponding reproduction of the I_P distributions using the pre-trained (a) VAE and (b) GAN; the realizations are obtained with the NT method.

Further information on the inversion performance is provided in Section 3.2. The fullstack seismic inversion and IAF parameters update algorithm is summarized in Fig. 1b.

2.5. Performance evaluation metrics

2.5.1. Deep generative models

The similarity and variability of facies patterns generated by the DGMs are assessed by computing the pairwise Euclidean distances between all the facies samples in the training data set and a set of DGM realizations. The data points are projected into a 2-D metric space using multi-dimensional scaling (MDS) (Cox and Cox, 2000); data points clustering in the same region identify similar manifolds of facies patterns. The quality of collocated I_P realizations is assessed by measuring the KL divergence between simulated and training I_P distributions.

The ability of the DGMs to reproduce multiple I_P distributions conditioned on the same facies pattern, hence simulate the geostatistical co-simulations, is further assessed. Given a set of I_P co-DSS realizations conditioned on a given facies pattern (*targets*), the latent space of the trained DGMs is explored to find all the possible I_P realizations fitting the targets. The NT method is employed for this purpose, where the fit can be determined by evaluating the I_P data misfit within the objective function (Eq. (10)). The resulting facies and I_P simulations are compared to the targets by measuring the root mean squared error (RMSE) distance and as the structural similarity index (SSIM) (Wang et al., 2004). The SSIM parameters are set to $M = 7$, $C_1 = 0.01$, and $C_2 = 0.03$, values which are commonly adopted, and were used in Levy et al. (2023) and Wang et al. (2004). The search for a best-fitting solution is done iteratively using the NT method considering the geostatistical realizations of I_P as the observed data $\mathbf{d}_{obs}$ (see Eq. (9)). Many particles (30) and a relatively low learning rate (0.005) are adopted to minimize the risk of convergence towards local minima (see e.g., Levy et al., 2023).

2.5.2. Seismic data inversion

The RMSE between the predicted and observed data $\mathbf{d}_{obs}$ ($RMSE_d$), models ($RMSE_F$ and $RMSE_{I_P}$ for facies and I_P, respectively), and latent vector in the DGM latent space ($RMSE_z$) are evaluated to assess the quality of the predicted posterior distribution. The convergence of the inversion is defined when: i) the average $RMSE_d$ from all the NT particles is equal to the average in the last 10% iterations of the algorithm; and ii) the average data misfit weighted by the standard deviation of the data noise, $WRMSE = \sqrt{\frac{1}{N_d}\sum\left[\frac{d_i - g_i(G(\mathbf{z}))}{\sigma_i}\right]^2}$, is lower than 1.1 (Levy et al., 2023). The SSIM distance of the average predicted facies and I_P is further assessed.

The results of the NT-based seismic inversion were also compared to those obtained by exploring the latent space of the DGM through MCMC, using differential evolution adaptive Metropolis algorithm ($DREAM_{(ZS)}$) (Laloy and Vrugt, 2012; Ter Braak and Vrugt, 2008; Vrugt et al., 2009). The algorithm uses several chains in parallel where, at each step, the model proposal distribution, sampled from an archive of past states, is updated based on differential evolution. The acceptance of the proposed distributions is based on the Metropolis acceptance probability. The predictive power of both MCMC and NT methods are compared based on the logarithmic scoring rule (LogS) (Good, 1952). The latter is measured as the negative logarithm of the probability of the true latent ($\mathbf{z}$), calculated using the predicted posterior distribution ($\widehat{p}$), that is, $\log S(\widehat{p}, \mathbf{z}) = -\log\widehat{p}(\mathbf{z})$. Lower values of $\log S(\widehat{p}, \mathbf{z})$ indicate a better approximation of the considered test model.

3. Results

The geological setting assumed for the test represents an environment with lenticular channels deposited in a shale background. Training images and inversion grid share the same size of 100 × 80 cells, which in our study correspond to 100 common mid points (CMP) and 80ms Two-

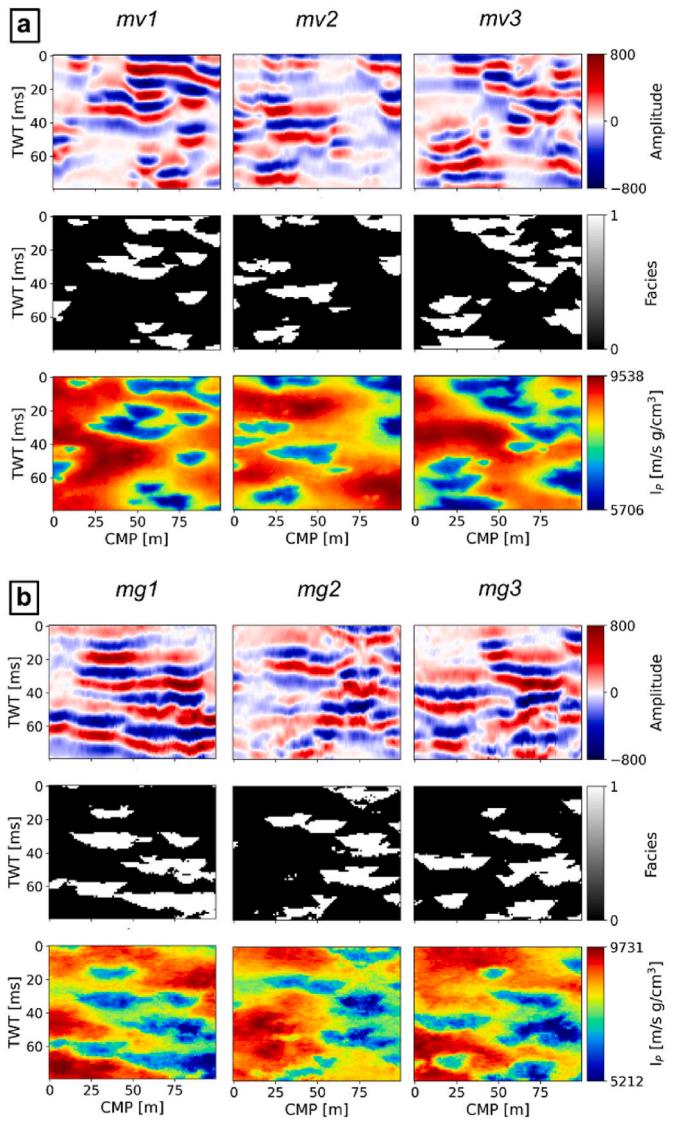

Fig. 7. Test scenarios used for the seismic data inversion with NT with VAE (a) and GAN (b). The facies and I_P images are generated by the DGM using a random, known, latent vector. The corresponding seismic reflection data is obtained using a Ricker wavelet (see Section 3.2 for further details).

way time (TWT). The facies models from the training data set were generated through MPS using a binary geological model as training image (Fig. 2a); depending on the DGM output activation functions, shales and sands were classified respectively as 0 and 1 for the VAE, or as −1 and 1 for the GAN. The average sand-to-shale ratio in each facies model is 0.269 ± 0.06 (Fig. 2b). The collocated I_P (Fig. 2a) was simulated through Co-DSS conditioning to facies-dependent distributions (Fig. 2c), sampled from two normal distributions having mean $\mu_{shale} = 8500 \pm 600 \left[\frac{m}{s}\frac{g}{cm^3}\right]$ and $\mu_{sand} = 6500 \pm 600 \left[\frac{m}{s}\frac{g}{cm^3}\right]$. The I_P spatial uncertainty is modeled per facies, through two spherical variogram models, with no nugget effect, and a range of 80m in the horizontal direction for both sand and shale. Along the vertical direction, the variograms' ranges are 30 ms two-way time (TWT) for the shale and 50 ms TWT for the sand.

3.1. Multivariate modeling with deep generative networks

Fig. 3a shows five realizations generated by the trained VAE and GAN. Geologically realistic facies models can be generated by both DGM, with no evident artifacts (Fig. 3a). The joint distributions of facies and collocated I_P values from 100 training images are compared to those generated by the DGMs (Fig. 3b). Here, the VAE's simulated facies show a distribution of intermediate values simulated between one facies class and the other. Contrarily, this effect is negligible for the GAN (Fig. 3b). To approximate the following calculations for the VAE, facies classes below 0.5 were considered shales, and those above this threshold were considered sands.

In the 2-D MDS space (Fig. 4a), the data points corresponding to 250 training facies patterns and DGM realizations lie in the same regions of the MDS space, identifying a good similarity and comparable variability. The distribution of points for the VAE's realizations, however, are significantly skewed towards the center of the MDS space, suggesting an under-exploration of the facies model parameter space. This difference is expected, as it is influenced by the continuous transition between facies classes in the VAE facies realizations. The sand-to-shale ratio distributions calculated from each facies model generated by the DGM are in close agreement with the training data (Fig. 4b).

The facies-dependent I_P distributions and their spatial correlation patterns are reproduced by both DGM (Figs. 3 and 4c). Nonetheless, the I_P realizations generated by the VAE result smoother (Fig. 3a), lacking the small-scale spatial variability present in GAN's realization and training data. Comparing the distribution of facies-dependent I_P from 250 random realizations (Fig. 4c), the training data ($\mu_{shale} = 8460 \pm 610 \left[\frac{m}{s}\frac{g}{cm^3}\right]$ and $\mu_{sand} = 6630 \pm 580 \left[\frac{m}{s}\frac{g}{cm^3}\right]$) are better reproduced by the GAN ($\mu_{shale} = 8460 \pm 608 \left[\frac{m}{s}\frac{g}{cm^3}\right]$ and $\mu_{sand} = 6650 \pm 630 \left[\frac{m}{s}\frac{g}{cm^3}\right]$) than by the VAE ($\mu_{shale} = 8350 \pm 650 \left[\frac{m}{s}\frac{g}{cm^3}\right]$ and $\mu_{sand} = 6800 \pm 570 \left[\frac{m}{s}\frac{g}{cm^3}\right]$). The KL divergence computed between the distributions in Fig. 4c are summarized in Table 1 and confirm the better performance of the GAN in reproducing the I_P marginal distribution.

The ability of the DGM to reproduce the geostatistical uncertainty represented in the training dataset was assessed using the conditional facies patterns in Fig. 5a and b. 30 co-DSS realizations of I_P are generated for each facies distribution. The pointwise average and standard deviation of these realizations are shown in Fig. 5a and b as *Target* I_P. An example of five realizations is shown in Fig. 6a and b as *Target*. Using the NT to explore the latent space (see Section 2.5.1) the results obtained for the trained VAE show that the network can fit the average I_P realizations with high accuracy, with a $RMSE_{I_P} = 235 \pm 15 \left[\frac{m}{s}\frac{g}{cm^3}\right]$ (Fig. 5a and 6a). Areas of lower standard deviation in the geostatistical simulations, mostly corresponding to the facies boundaries, and those of higher standard deviation are generally approximated well by the VAE, although smoothed. The variance reproduced by this DGM is approximately 72% of the variance of the geostatistical realizations. At the same time, the facies realizations generated simultaneously to the I_P distribution match the conditioning facies with an $RMSE_F = 0.13 \pm 0.01$, reproducing well the local co-dependency of the two variables (Fig. 5a). In the same test, the facies and co-simulated I_P realizations obtained with the GAN reproduce the target with generally poorer accuracy compared to the VAE (Fig. 5b). On average, the resulting I_P realizations match the geostatistical targets with an $RMSE_{I_P} = 343 \pm 19 \left[\frac{m}{s}\frac{g}{cm^3}\right]$. The pointwise I_P mean model of the GAN realizations present small-scale patterns mismatching with the corresponding mean calculated from the geostatistical realizations (Fig. 5b). The I_P realizations reproduce the spatial patterns of the corresponding targets with an $RMSE_{I_P} = 364 \pm 20 \left[\frac{m}{s}\frac{g}{cm^3}\right]$ (Fig. 6b). The variance reproduced by the GAN is approximately 83% of that of the target I_P spatial distributions. The facies patterns, co-generated with I_P, have a relatively larger variability and an $RMSE_F = 0.22 \pm 0.02$, compared to the VAE. The differences in performance observed between VAE and GAN are also reflected in the SSIM values calculated for both properties and shown in Table 1. A qualitative

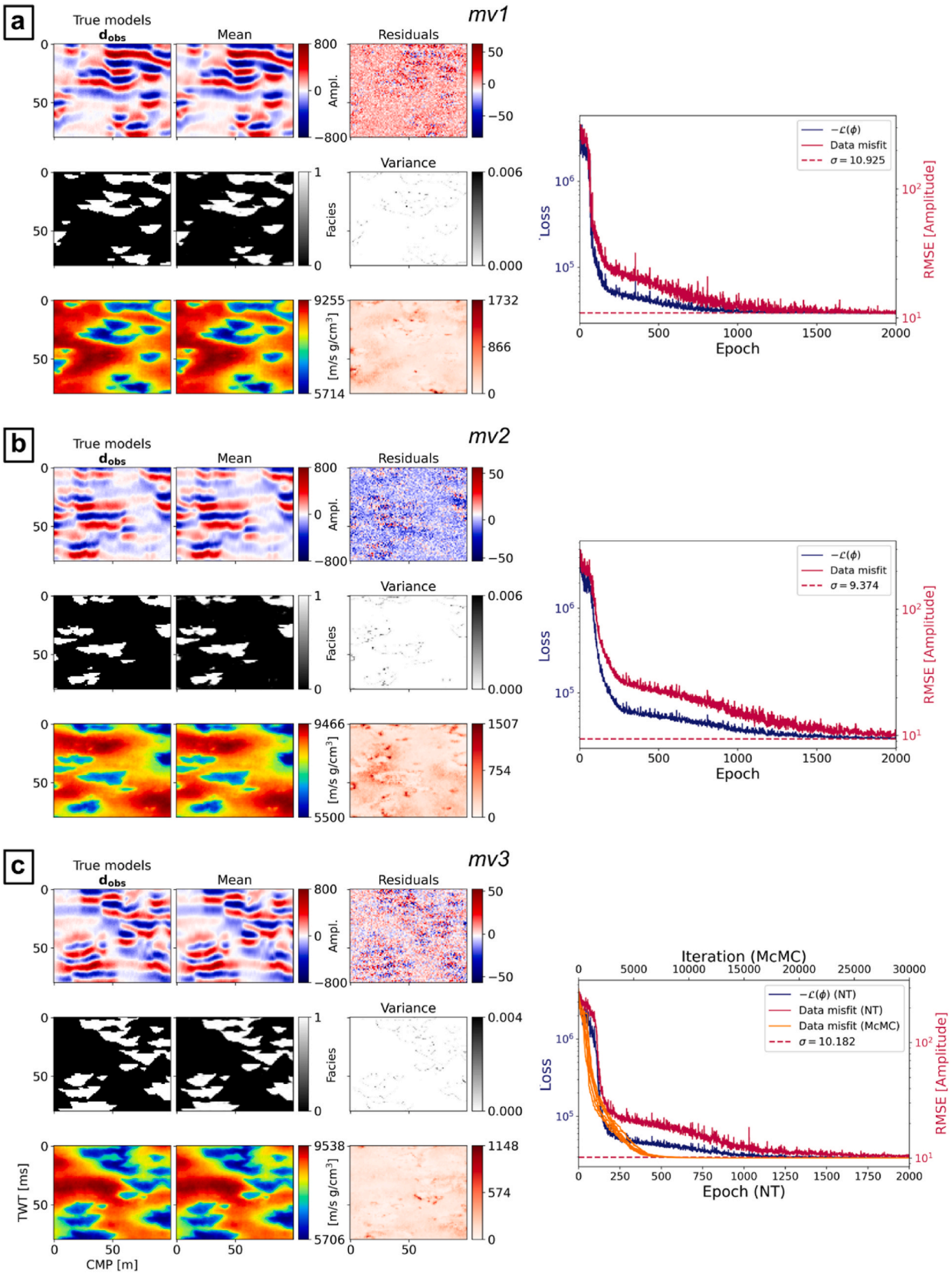

Fig. 8. Results of the seismic data inversion using NT with VAE and progression of data misfit and loss per inversion epoch. The data misfit for seismic data inversion with MCMC is also shown for mv3 (c). The retrieved subsurface spatial models are shown as pointwise mean facies, I_P and seismic data predicted at the last iteration, together with the corresponding seismic residuals and variance.

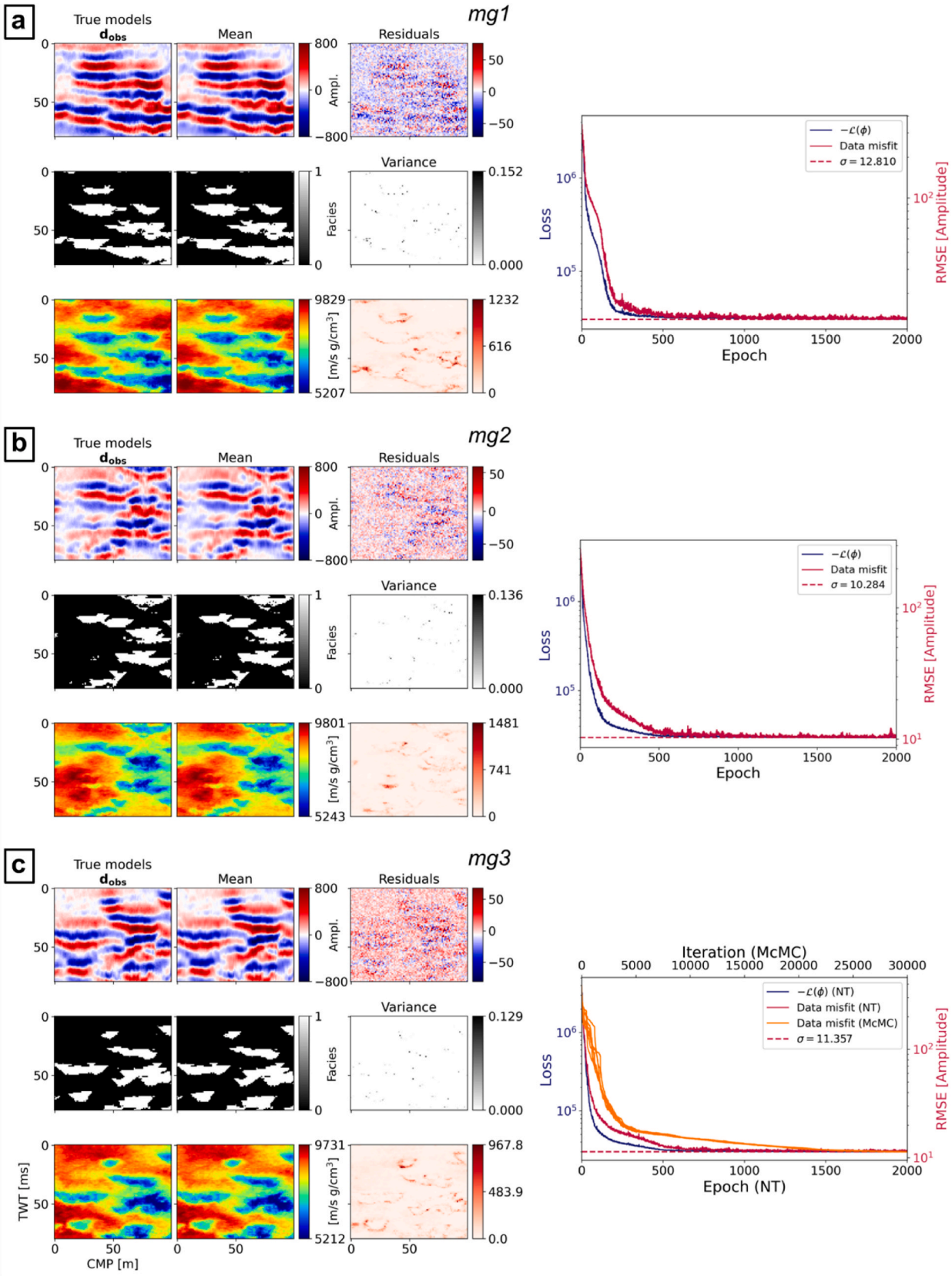

Fig. 9. Results of the seismic data inversion using NT with GAN and progression of data misfit and loss per inversion epoch. The data misfit for seismic data inversion with MCMC is also shown for *mg3* (c). The retrieved subsurface spatial models are shown as pointwise mean facies, I_P and seismic data predicted at the last iteration, together with the corresponding seismic residuals and variance.

Table 2
Summary of the statistical metrics used to evaluate the seismic inversion though NT with both VAE and GAN.

Model	VAE			GAN		
	mv1	*mv2*	*mv3*	*mg1*	*mg2*	*mg3*
RMSEz	0.04	0.04	0.06	2.33	2.34	2.15
$RMSE_d$ *[Ampl.]*	10.94	9.61	10.23	13.43	10.49	11.92
$RMSE_F$	0.05	0.05	0.05	0.02	0.01	0.01
$RMSE_{I_P}$ $\left[\frac{m}{s}\frac{g}{cm^3}\right]$	10.46	13.02	12.91	7.55	7.15	9.73
Convergence step	1646	1748	1645	874	1334	930
$SSIM_F$	0.95	0.92	0.97	0.99	1.00	0.99
$SSIM_{I_P}$	1.00	1.00	1.00	1.00	1.00	1.00

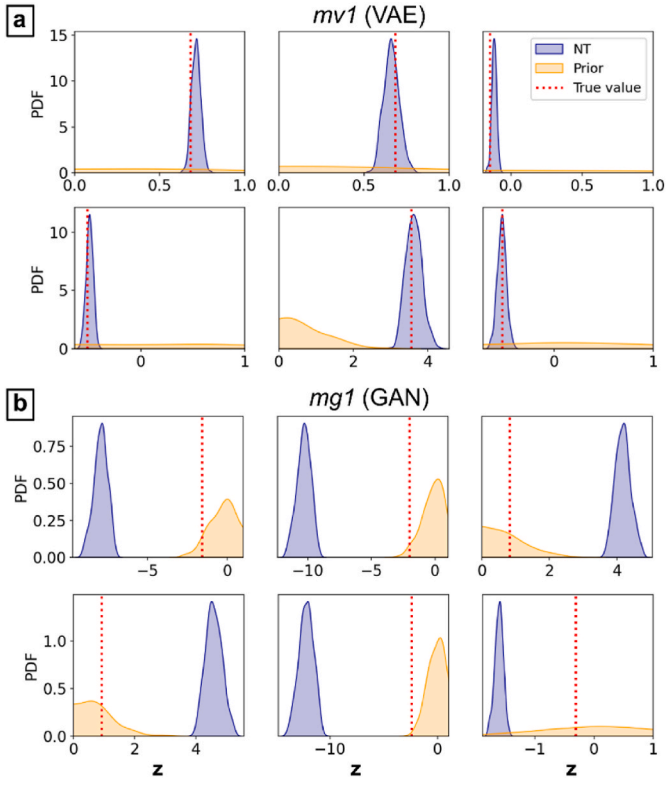

Fig. 10. Comparison of the posterior approximation through NT for the *mv1* (a) and *mg1* (b) test cases. The figure shows a subset of 6 of the 60 samples of the latent vectors.

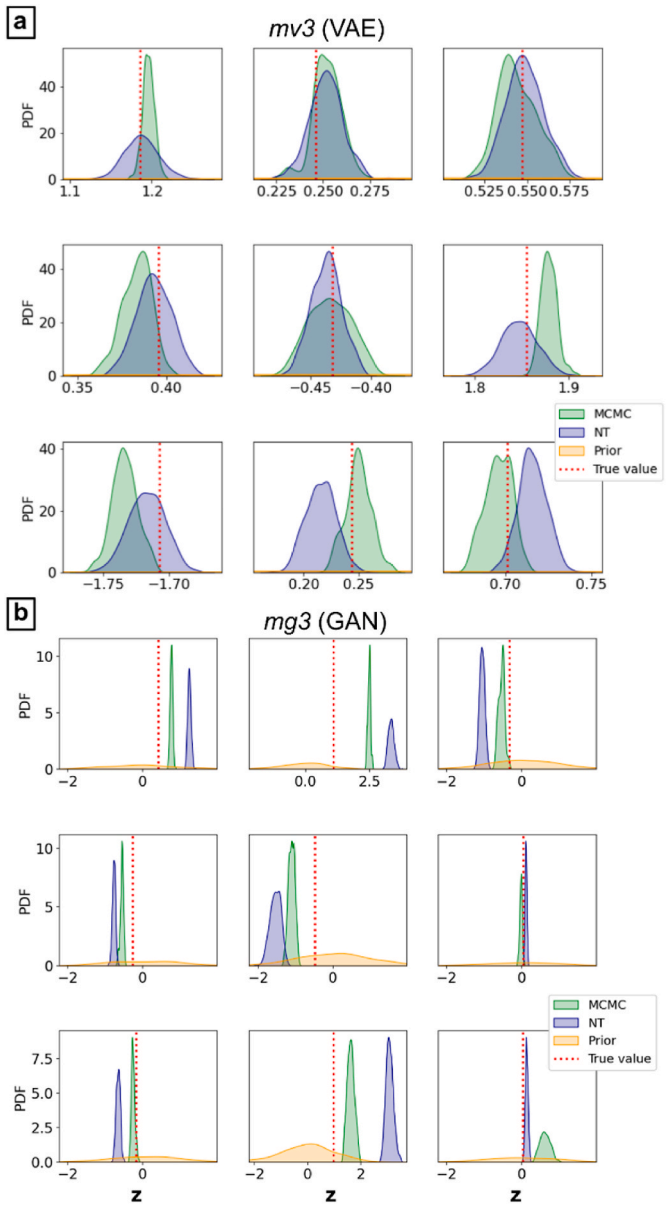

Fig. 11. Comparison of the posterior distributions in the latent space as predicted by the NT and MCMC, for both *mv3* (a) and *mg3* (b) test cases. The figure shows a subset of 6 of the 60 samples of the latent vectors.

representation of the areas of the latent space where the NT converged for each I_P realization target is given in Fig. 5c, in MDS space, for both VAE and GAN.

3.2. Seismic data inversion

The NT-based seismic inversion was evaluated for both VAE and GAN using different test scenarios (Fig. 7a and b, respectively). The test spatial distributions of facies and I_P were generated randomly by sampling the latent space of the trained DGM, so their representation in the latent space of the DGMs is known. The I_P distributions were used to compute the reflectivity coefficients (Eq. (7)); then, a Ricker wavelet with a central frequency of 25 Hz, length of 120 ms and sampling rate of 2 ms was used to calculate $\mathbf{d}_{obs}$. The seismic data were contaminated with random noise normally distributied as, $\mathcal{N}(0, \sigma)$, with $\sigma = 0.03\,\mathbf{d}_{obs} + 5$ [Amplitude] (see Section 2.4.2). The test models are named using the label '*mv*', for the VAE, and '*mg*' for those of the GAN.

To compare the performances of the NT inversions with the two DGM, the inversion process ran for 2000 epochs in both cases. The calibration of the hyperparameters involved testing different learning rates and number of particles. For both the VAE and GAN cases, a learning rate of 0.01 provided the fastest inversion at the highest accuracy. For the NT with VAE, one particle per epoch was sufficient to obtain accurate predictions. Comparable performances for the NT with

Table 3
Comparison of the PDFs in the latent space **z** for test models *mv3* and *mg3*: logS of the posterior distributions predicted by NT and McMC, and the Prior distribution; and the KL divergence between the predicted distributions and the prior.

	mv3			mg3		
	NT	MCMC	Prior	NT	MCMC	Prior
logS	−2.62	−2.45	1.38	427.39	129.39	1.38
	NT||Prior		**MCMC||Prior**	**NT||Prior**		**MCMC||Prior**
$D_{KL}(Q\|\|P)$	3.66		4.42	9.25		4.05

GAN were found when using five particles per epoch. The MCMC method was applied for the seismic inversions in scenarios *mv3* and *mg3* to enable comparison of the results with a reference solution. Here, eight parallel chains were used per test model and the number of samples per chain was set to 30000. The NT and MCMC inversion algorithms for seismic inversion are implemented in Python 3.18. The inversions ran on a machine with Windows 11, with a CPU Intel® Core™ i7-8750H, on the sole CPU. Approximately, each epoch (or inversion iteration) of the NT took 0.05 s for the IAF in combination with the VAE (using one particle), and 0.4 s for the IAF with GAN (using five particles). The inversion with MCMC took approximately 0.17 s per iteration when combined with VAE and 0.19 s per iteration when combined with GAN (using eight chains in both cases). The actual inversion time depends on each specific case study.

Figs. 8 and 9 summarize the results of the NT inversion for VAE and GAN, respectively, while Table 2 gives the performance metrics (Section 2.5) applied to the last iteration of the NT inversion. The NT inversion with VAE converged in all the three considered scenarios (Fig. 8). The seismic data misfit between the average seismic data and $\mathbf{d}_{obs}$ ($RMSE_d$), is in statistical agreement with the assumed σ_d (Table 2) and the resulting 2-D residual images are mostly populated by the relative data noise (Fig. 8). In all three scenarios, the facies and I_P models agree well with the true test scenarios, showing low values of misfit to the target models and posterior's variance. The least good performance is found for *mv2*. In this case, the NT inversion converged after 1748 iterations, against the 1646 and 1645 for *mv1* and *mv3*, respectively; the predicted average models of facies and I_P also show relatively higher $RMSE_d$ and $RMSE_{I_P}$ values compared to the other two test models. Overall, the facies and I_P patterns are well reproduced in all scenarios with values of SSIM above 0.92 (Table 2). The NT inversion results obtained with the GAN are of comparable quality to those of the VAE in terms of high-dimensional model reconstruction. As for the previous case, the $RMSE_d$ agrees with the assumed data errors at the last iteration (Fig. 8 and Table 2). In all the cases, the predicted facies and I_P match the spatial distributions of the target (Fig. 8) with a higher accuracy than for the VAE cases (Table 2). This is further confirmed by the SSIM values for facies and I_P, that are between 0.99 and 1.

In agreement with Levy et al. (2023), the VAE outperforms the GAN in approximating the posterior distribution in the latent space: the average $RMSE_Z$ between the true latent vector and the predicted distributions is two orders of magnitude lower for the VAE case (Table 2). Fig. 10 compares the distributions of the posterior as predicted by the NT with VAE and GAN, for *mv1* and *mg1*, in a subset of 6 variables of the latent space. Here, the distribution predicted by the NT with VAE always includes the true values of the target latent vector, contrarily to the GAN case, where the predicted distributions converge towards different values.

In the MCMC inversion, the time required to reach the burn-in for *mv3* (Fig. 8c) (495 iterations) is considerably lower than for *mg3* (Fig. 9c) (1490 iterations). After convergence, the $RMSE_d$ of the predicted data is comparable to that of the NT methods (Fig. 8c and 9c) and equals the assumed data uncertainty σ_d. The facies and I_P predicted by the MCMC with VAE have $RMSE_F = 0.05298 \pm 4e-5$ and $RMSE_{I_P} = 7.55 \pm 0.35 \left[\frac{m}{s}\frac{g}{cm^3}\right]$; and those predicted by the MCMC with GAN have an average $RMSE_F = 0.0534 \pm 6 \times 10^{-4}$ and $RMSE_{I_P} = 23.05 \pm 0.21 \left[\frac{m}{s}\frac{g}{cm^3}\right]$. The latent vector is predicted accurately for *mv3* by both NT and MCMC, while both methods fail at converging towards the correct latent vector for *mg3* (Fig. 11). As shown in Table 3, the smaller logS value of the posterior predicted by NT for *mv3*, compared to that of MCMC, indicates larger accuracy. On the other hand, the larger logS values for *mg3* are a consequence of the large discrepancy between the predicted and true distributions. While both methods converge to a specific solution (see KL divergence in Table 3), the relative difference between NT and MCMC can be explained by the smaller variance of the posterior inferred by the MCMC method and its distance to the true value of the known latent vector.

4. Discussion

The evaluations of the proposed VAE and GAN show that these DGM can be considered as alternative approaches to multivariate geostatistical modeling of facies and collocated I_P. The two DGMs can be further used as priors in variational Bayesian inversion of seismic reflection data, through the NT method. As the encoded parameters have a well-defined prior probability density function, it is possible to calculate the gradient of a misfit function with respect to each parameter, making it possible to generate conditional simulations efficiently.

Prior modeling and, consequently, inverse modeling performance with these networks are characterized by specific advantages and drawbacks, mostly due to the types of architectures and training (i.e., variational and adversarial for VAE and GAN, respectively). While both networks encode the given prior distribution with a satisfactory degree of approximation, the GAN generates sharper and more accurate spatial distributions of facies and I_P, compared to the VAE (Figs. 3 and 4). The lower spatial resolution of the VAE realizations and the modeling of intermediate facies values are both characteristics typical of this network. Similar results were found, for example, in Bao et al. (2022) and Levy et al. (2023).

The generative models considered are deterministic, that is, to each point in the latent space corresponds a unique realization of facies and I_P. The use of two independent convolutional branches at the end of the generative networks favors the modeling of several manifolds of I_P spatial patterns, while modeling the same facies pattern (with good approximation), within the assumed uncertainty. The VAE outperformed the GAN when testing their ability to reproduce I_P facies-dependent spatial uncertainty (Section 3.1.2), generating I_P realizations that better matched the targets, while honoring the facies patterns (Figs. 5 and 6). Nonetheless, the better performances of the VAE are, in fact, partially due to the characteristic smoothing of the original spatial patterns, averaging out the small-scale spatial variability of the features. Consequently, the differences between one I_P realizations predicted for each target are relatively smaller; this is also reflected in corresponding smaller regions sampled in the latent space (Fig. 5c). In this test, the GAN was able to match the targets with sufficiently good approximation. Nonetheless, the overestimation of small-scale variability visible in the realizations' mean (Fig. 5b) suggests an underestimation of the spatial variance of the modeled I_P. According to the assumed physics governing the propagation of the source seismic wavelet, the recorded seismic data are entirely dependent on the spatial distribution of I_P. Therefore, the results of this test allow to assess the reliability of facies predictions using the NT with the proposed DGM, when considering facies-dependent I_P data uncertainty. A possible improvement to the modeling ability of the proposed networks would be to use probabilistic generative networks (e.g., Feng et al., 2022), to quantify the uncertainty in the parameters modeling. Nonetheless, the adoption of these networks in inverse problems may require complex inference framework, due to the significantly larger number of unknown parameters.

The NT inversion results demonstrate that it is possible to combine IAF with the proposed DGM and solve a seismic inverse problem efficiently. The smaller number of iterations required for the convergence of the inversion results, compared to the VAE case, can be imputable to the use of a larger number of particles. While both networks provide comparable results in the approximation of the high-dimensional spatial patterns of facies and I_P (Figs. 8 and 9, and Table 2), the two networks differ in their ability to approximate the posterior distribution in the latent space. In fact, the posterior distributions predicted by the GAN converge towards different values than the true one (Fig. 10). The test conducted with MCMC confirms these results (Fig. 11), suggesting that the issue is related to the generative model rather than to the inference

method used. These results are consistent with those obtained by Levy et al. (2023) on GPR data inversion and agree with the findings of Laloy et al. (2019) and Lopez-Alvis et al. (2021). This is also explains the need of a larger number of particles, as demonstrated by Levy et al. (2023). The authors conclude, through different analyses, that the adversarial training favors highly non-linear transformations, resulting in a model parameter space characterized by a rough surface with several local minima. A further indication of this observation can be interpreted by the pseudo-inversion analysis provided in Section 3.1.2: compared to the VAE, similar I_P realizations are sampled from a wider range of points in the latent space (Fig. 5c). Overall, the NT demonstrated to be faster than the MCMC for the inversion of the tested models, requiring less iterations. For the VAE case, the NT is significantly less computational demanding (the inversion is possible with just one particle) and provides more accurate predictions.

The accurate training of the DGMs proposed is a fundamental aspect, as their modeling ability directly influences the accuracy of subsurface parameters predictions. Further research on this topic should aim at studying the applicability of the proposed framework in real case scenarios, for the prediction of both I_P and facies spatial patterns.

5. Conclusions

The aim of the proposed research is to improve the prediction of subsurface facies distributions through the neural transport (NT) method by accounting for experimental uncertainty in facies-dependent properties on forward modeling. Two deep generative models based on variational autoencoder (VAE) and generative adversarial network (GAN) architectures are proposed to simultaneously learn the spatial arrangement of facies and the corresponding collocated acoustic impedance (I_P). The training is based on a dataset of geostatistical realizations of both facies and collocated I_P. Both networks reproduce the prior data joint distribution and spatial patterns of facies and I_P spatial uncertainty, mapping a latent vector into multivariate realizations. Both VAE and GAN have characteristic limitations and advantages in agreement with previous studies when used for subsurface modeling. The VAE can encode similar manifolds of facies and collocated I_P spatial distributions into close regions of the latent space, although with limited capabilities in modeling the small-scale spatial features in the training images. The GAN can better reproduce the desired spatial statistics of the training images, but shows poorer encoding capabilities, scattering similar I_P manifolds into larger regions of its latent space. The proposed applications of NT with the two generative models demonstrate that this method can indeed be used for multivariate parameter inversion of seismic reflection data. The posterior distribution, parametrically defined, and provides a quantification of the actual uncertainty on the subsurface spatial models. As in the baseline studies, the NT combined with the GAN required a relatively larger number of parallel predictions to update its parameters: the non-linear mapping of the GAN affects the accurate prediction of the posterior in its latent space. However, the inference of the subsurface facies and I_P distribution is accurate using NT with both generative models. The results of the NT are comparable to those obtained through Markov chain Monte Carlo sampling, with the advantage of needing a significantly lower number of iterations.

Computed code availability

The python codes for the GAN, VAE training and IAF inversion algorithms are available together with training and test models, at https://github.com/romiele/DGM-for-multivariate-modeling-and-inversion.

Competing interest declaration

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of generative AI in scientific writing

The authors declare that no AI or AI-assisted technology was used in the writing process of this work.

CRediT authorship contribution statement

Roberto Miele: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Conceptualization. **Shiran Levy:** Writing – review & editing, Methodology. **Niklas Linde:** Writing – review & editing, Methodology. **Amilcar Soares:** Writing – review & editing. **Leonardo Azevedo:** Writing – review & editing, Supervision, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge the support of the CERENA (strategic project FCT-UIDP/04028/2020) and that of the FCT (PhD project 2020.05571.BD).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cageo.2024.105622.

References

Azevedo, L., Paneiro, G., Santos, A., Soares, A., 2020. Generative adversarial network as a stochastic subsurface model reconstruction. Comput. Geosci. 24, 1673–1692. https://doi.org/10.1007/s10596-020-09978-x.

Azevedo, L., Soares, A., 2017. Geostatistical methods for reservoir geophysics. Advances in Oil and Gas Exploration & Production. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-53201-1.

Bao, J., Li, L., Davis, A., 2022. Variational autoencoder or generative adversarial networks? A comparison of two deep learning methods for flow and transport data assimilation. Math. Geosci. 54, 1017–1042. https://doi.org/10.1007/s11004-022-10003-3.

Blei, D.M., Kucukelbir, A., McAuliffe, J.D., 2017. Variational inference: a review for statisticians. J. Am. Stat. Assoc. 112, 859–877. https://doi.org/10.1080/01621459.2017.1285773.

Bosch, M., Mukerji, T., Gonzalez, E.F., 2010. Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: a review. Geophysics 75, 75A165–75A176. https://doi.org/10.1190/1.3478209.

Brunetti, C., Linde, N., 2018. Impact of petrophysical uncertainty on Bayesian hydrogeophysical inversion and model selection. Adv. Water Resour. 111, 346–359. https://doi.org/10.1016/j.advwatres.2017.11.028.

Buland, A., Omre, H., 2003. Bayesian linearized AVO inversion. Geophysics 68, 185–198. https://doi.org/10.1190/1.1543206.

Chan, S., Elsheikh, A.H., 2019. Parametric generation of conditional geological realizations using generative neural networks. Comput. Geosci. 23, 925–952. https://doi.org/10.1007/s10596-019-09850-7.

Connolly, P.A., Hughes, M.J., 2016. Stochastic inversion by matching to large numbers of pseudo-wells. Geophysics 81, M7–M22. https://doi.org/10.1190/geo2015-0348.1.

Cox, T., Cox, M., 2000. Multidimensional Scaling, 0 ed. Chapman and Hall/CRC. https://doi.org/10.1201/9780367801700.

De Figueiredo, L.P., Grana, D., Bordignon, F.L., Santos, M., Roisenberg, M., Rodrigues, B. B., 2018. Joint Bayesian inversion based on rock-physics prior modeling for the estimation of spatially correlated reservoir properties. Geophysics 83, M49–M61. https://doi.org/10.1190/geo2017-0463.1.

De Figueiredo, L.P., Grana, D., Roisenberg, M., Rodrigues, B.B., 2019. Gaussian mixture Markov chain Monte Carlo method for linear seismic inversion. Geophysics 84, R463–R476. https://doi.org/10.1190/geo2018-0529.1.

Doyen, P., 2014. Seismic Reservoir Characterization: an Earth Modelling Perspective. EAGE Publications bv.
Doyen, P., 2007. Seismic Reservoir Characterization: an Earth Modelling Perspective, Education Tour Series. EAGE. DB Houten.
Dupont, E., Zhang, T., Tilke, P., Liang, L., Bailey, W., 2018. Generating Realistic Geology Conditioned on Physical Measurements with Generative Adversarial Networks.
Feng, R., Grana, D., Mukerji, T., Mosegaard, K., 2022. Application of bayesian generative adversarial networks to geological facies modeling. Math. Geosci. 54, 831–855. https://doi.org/10.1007/s11004-022-09994-w.
Friedli, L., Linde, N., 2023. Solving geophysical inversion problems with intractable likelihoods: linearized Gaussian approximations versus the correlated pseudo-marginal method. Math. Geosci. https://doi.org/10.1007/s11004-023-10064-y.
González, E.F., Mukerji, T., Mavko, G., 2008. Seismic inversion combining rock physics and multiple-point geostatistics. Geophysics 73, R11–R21. https://doi.org/10.1190/1.2803748.
Good, I.J., 1952. Rational decisions. J. Roy. Stat. Soc. B 14, 107–114. https://doi.org/10.1111/j.2517-6161.1952.tb00104.x.
Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Networks. arXiv preprint arXiv:1406.2661.
Grana, D., Azevedo, L., De Figueiredo, L., Connolly, P., Mukerji, T., 2022. Probabilistic inversion of seismic data for reservoir petrophysical characterization: review and examples. Geophysics 87, M199–M216. https://doi.org/10.1190/geo2021-0776.1.
Grana, D., Della Rossa, E., 2010. Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. Geophysics 75, O21–O37. https://doi.org/10.1190/1.3386676.
Grana, D., Fjeldstad, T., Omre, H., 2017. Bayesian Gaussian mixture linear inversion for geophysical inverse problems. Math. Geosci. 49, 493–515. https://doi.org/10.1007/s11004-016-9671-9.
Grana, D., Mukerji, T., Doyen, P., 2021. Seismic Reservoir Modeling: Theory, Examples, and Algorithms. Wiley Blackwell, Hoboken, NJ Chichester.
Grana, D., Mukerji, T., Dvorkin, J., Mavko, G., 2012. Stochastic inversion of facies from seismic data based on sequential simulations and probability perturbation method. Geophysics 77, M53–M72. https://doi.org/10.1190/geo2011-0417.1.
Hoffman, M., Sountsov, P., Dillon, J.V., Langmore, I., Tran, D., Vasudevan, S., 2019. NeuTra-lizing Bad Geometry in Hamiltonian Monte Carlo Using Neural Transport.
Jordão, H., Sousa, A.J., Soares, A., 2023. Using bayesian neural networks for uncertainty assessment of ore type boundaries in complex geological models. Nat. Resour. Res. 32, 2495–2514. https://doi.org/10.1007/s11053-023-10265-6.
Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.
Kingma, D.P., Salimans, T., Jozefowicz, R., Chen, X., Sutskever, I., Welling, M., 2017. Improving Variational Inference with Inverse Autoregressive Flow. Adv. Neural Inform. Process. Syst. 29, 4743–4751.
Kingma, D.P., Welling, M., 2014. Auto-Encoding Variational Bayes. arXiv preprint arXiv:1312.6114.
Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Ann. Math. Stat. 22, 79–86. https://doi.org/10.1214/aoms/1177729694.
Laloy, E., Hérault, R., Jacques, D., Linde, N., 2018. Training-image based geostatistical inversion using a spatial generative adversarial neural network. Water Resour. Res. 54, 381–406. https://doi.org/10.1002/2017WR022148.
Laloy, E., Hérault, R., Lee, J., Jacques, D., Linde, N., 2017. Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network. Adv. Water Resour. 110, 387–405. https://doi.org/10.1016/j.advwatres.2017.09.029.
Laloy, E., Linde, N., Ruffino, C., Hérault, R., Gasso, G., Jacques, D., 2019. Gradient-based deterministic inversion of geophysical data with generative adversarial networks: is it feasible? Comput. Geosci. 133, 104333 https://doi.org/10.1016/j.cageo.2019.104333.
Laloy, E., Vrugt, J.A., 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM $_{(ZS)}$ and high-performance computing: efficient MCMC for high-dimensional problems. Water Resour. Res. 48 https://doi.org/10.1029/2011WR010608.
Levy, S., Hunziker, J., Laloy, E., Irving, J., Linde, N., 2021. Using deep generative neural networks to account for model errors in Markov chain Monte Carlo inversion. Geophys. J. Int. 228, 1098–1118. https://doi.org/10.1093/gji/ggab391.
Levy, S., Laloy, E., Linde, N., 2023. Variational Bayesian inference with complex geostatistical priors using inverse autoregressive flows. Comput. Geosci. 171, 105263 https://doi.org/10.1016/j.cageo.2022.105263.
Linde, N., Renard, P., Mukerji, T., Caers, J., 2015. Geological realism in hydrogeological and geophysical inverse modeling: a review. Adv. Water Resour. 86, 86–101. https://doi.org/10.1016/j.advwatres.2015.09.019.
Lopez-Alvis, J., Laloy, E., Nguyen, F., Hermans, T., 2021. Deep generative models in inversion: the impact of the generator's nonlinearity and development of a new approach based on a variational autoencoder. Comput. Geosci. 152, 104762 https://doi.org/10.1016/j.cageo.2021.104762.
Mariethoz, G., Renard, P., Straubhaar, J., 2010. The Direct Sampling method to perform multiple-point geostatistical simulations: performing multiple-points simulations. Water Resour. Res. 46 https://doi.org/10.1029/2008WR007621.
Miele, R., Barreto, B.V., Yamada, P., Varella, L.E.S., Pimentel, A.L., Costa, J.F., Azevedo, L., 2022. Geostatistical seismic rock physics AVA inversion with data-driven elastic properties update. IEEE Trans. Geosci. Rem. Sens. 60, 1–15. https://doi.org/10.1109/TGRS.2021.3135718.
Miele, R., Grana, D., Seabra Varella, L.E., Viola Barreto, B., Azevedo, L., 2023. Iterative geostatistical seismic inversion with rock physics constraints for permeability prediction. Geophysics 1–51. https://doi.org/10.1190/geo2022-0352.1.
Mosser, L., Dubrule, O., Blunt, M.J., 2020. Stochastic seismic waveform inversion using generative adversarial networks as a geological prior. Math. Geosci. 52, 53–79. https://doi.org/10.1007/s11004-019-09832-6.
Mosser, L., Dubrule, O., Blunt, M.J., 2018. Conditioning of three-dimensional generative adversarial networks for pore and reservoir-scale models. arXiv:1802.05622 [physics, stat].
Nunes, R., Soares, A., Azevedo, L., Pereira, P., 2017. Geostatistical seismic inversion with direct sequential simulation and Co-simulation with multi-local distribution functions. Math. Geosci. 49, 583–601. https://doi.org/10.1007/s11004-016-9651-0.
Pan, W., Torres-Verdín, C., Pyrcz, M.J., 2021. Stochastic Pix2pix: a new machine learning method for geophysical and well conditioning of rule-based channel reservoir models. Nat. Resour. Res. 30, 1319–1345. https://doi.org/10.1007/s11053-020-09778-1.
Papamakarios, G., Nalisnick, E., Rezende, D.J., Mohamed, S., Lakshminarayanan, B., 2021. Normalizing Flows for Probabilistic Modeling and Inference.
Rezende, D.J., Mohamed, S., 2015. Variational Inference with Normalizing Flows. In International Conference on Machine Learning. PMLR, pp. 1530–1538.
Sen, M.K., 2006. Seismic Inversion. Society of Petroleum Engineers, Richardson, TX.
Sen, M.K., Stoffa, P.L., 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion1. Geophys. Prospect. 44, 313–350. https://doi.org/10.1111/j.1365-2478.1996.tb00152.x.
Strebelle, S., 2021. Multiple-point statistics simulation models: pretty pictures or decision-making tools? Math. Geosci. 53, 267–278. https://doi.org/10.1007/s11004-020-09908-8.
Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. Math. Geol. 21.
Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics, Philadelphia, PA.
Ter Braak, C.J.F., Vrugt, J.A., 2008. Differential Evolution Markov chain with snooker updater and fewer chains. Stat. Comput. 18, 435–446. https://doi.org/10.1007/s11222-008-9104-9.
Vrugt, J.A., Ter Braak, C.J.F., Gupta, H.V., Robinson, B.A., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch. Environ. Res. Risk Assess. 23, 1011–1026. https://doi.org/10.1007/s00477-008-0274-y.
Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13, 600–612. https://doi.org/10.1109/TIP.2003.819861.
Zhang, C., Song, X., Azevedo, L., 2021. U-net generative adversarial network for subsurface facies modeling. Comput. Geosci. 25, 553–573. https://doi.org/10.1007/s10596-020-10027-w.