Received: 1 June 2023 | Accepted: 12 July 2023

DOI: 10.1111/bre.12811

Basin Research | IAS EAGE | WILEY

RESEARCH ARTICLE

Introducing conceptual geological information into Bayesian tomographic imaging

Hugo Bloem[1] | **Andrew Curtis**[1] | **Daniel Tetzlaff**[2]

[1]School of GeoSciences, University of Edinburgh, Edinburgh, UK

[2]Westchase Software Corporation, Houston, Texas, USA

Correspondence
Hugo Bloem, School of GeoSciences, University of Edinburgh, Edinburgh, UK.
Email: hugo.bloem@ed.ac.uk

Abstract

Geological process models typically simulate a range of dynamic processes to evolve a base topography into a final two-dimensional cross section or three-dimensional geological scenario. In principle, process parameters may be updated to better align with observed geophysical or geological data. However, it is hard to find any process model realisations that fit all observations if data sets are complex and sparse in space or time because the simulations typically depend highly non-linearly on base topography and dynamic parameters. As an alternative, geophysical probabilistic tomographic methods may be used to estimate the family of models of a target subsurface structure that are consistent both with information obtained from previous experiments and with new data (the Bayesian posterior probability distribution). However, this family seldom embodies geologically reasonable images. Here we show that the posterior distribution of tomographic images obtained from travel time data can be fully geological by injecting geological prior information into Bayesian inference and that we can do this near-instantaneously by using trained mixture density networks (MDNs). We invoke two geological concepts as prior information about the possible depositional environment of an imaged target structure: a braided river system and a set of marine parasequences. Each concept is parameterised by the latent parameters of a generative adversarial network. Data from a target structure can then be used to infer the family of compatible latent parameter values using either geological concept using MDNs. Our near-instantaneous MDN solutions closely resemble those found using relatively expensive Monte Carlo methods. We show that while the use of incorrect geological conceptual models provides significantly less accurate results, a classifier neural network can infer which geological conceptual model is most consistent with the data. It is thus demonstrated that even apparently barely related geophysical data may contain information about abstract geological concepts, and that geological conceptual models are key to creating reasonable images from geophysical data.

Westchase Software Corporation: https://wsoftc.com

KEYWORDS
forward stratigraphic modelling, geological process modelling, imaging, inversion

1 | INTRODUCTION

Geological process models simulate dynamic processes to evolve an initial topography into a geologically plausible two-dimensional cross section or three-dimensional geological model (Burgess et al., 2001; Hill et al., 2009; Paola, 2000; Tetzlaff & Harbaugh, 1989). The input parameters to the geological process model may be updated to align the model output with geophysical data such as recorded seismic travel times or waveforms. However, a number of problems may occur: some geological process models are chaotic in their behaviour (Burgess & Emery, 2004; Tetzlaff & Harbaugh, 1989) such that a small change in the input could lead to a significantly different output. In addition, the model output is never uniquely constrained by data so that an infinite family of dynamic models is consistent with observations. Finding the family of realisations that are consistent with observed data may be infeasible due to the computational expense involved. And finally, the true structure of the Earth always deviates from the model output, even when compared to the level of detail of the model. As an alternative, we may look to probabilistic inversion methods to identify the family of models that fit geophysical data within their uncertainties (Tarantola, 2005). Unfortunately, as applied to date, geophysical inversion methods do not impose geological realism as a criterion for the solution. As a result, the inferred models are usually geologically implausible. In this paper, we aim to combine the geological prior information embodied within process models with probabilistic inversion methods of geophysics to obtain a set of geological models that both fit the recorded data and are geologically reasonable.

Seismic travel time tomography is commonly applied to image the Earth's subsurface (Aki & Lee, 1976; Dziewonski & Woodhouse, 1987; Lee et al., 1995; Tsekhmistrenko et al., 2021; Zhang & Curtis, 2020). By measuring the time taken for waves to travel between pairs of points on the Earth's surface, tomographic methods estimate maps of subsurface velocities in up to three dimensions. The subsurface is usually described by a finite-dimensional parameter matrix m which often consists of seismic velocities at each of a set of subsurface locations. We study the situation where those parameters are to be inferred from a vector of recorded data d which describe the travel times of the seismic energy between a set of sources and receivers.

Highlights

- Combined geological process models and probabilistic geophysical tomography.
- Solutions obtained are probability distributions over geologically plausible models.
- Near-instantaneous inversion of data for model parameters using trained neural networks.
- Possible to constrain most appropriate geological conceptual model from geophysical data.
- First application of a fast stratigraphic forward model: SedSimple.

Estimating subsurface velocities from travel times is a non-linear inverse problem (Aki & Lee, 1976). The unknown inverse function is potentially complicated and ill-posed and always has a non-unique solution which means that infinitely many subsurface parameter matrices fit the data to within measurement uncertainty. It is therefore impossible to infer which particular parameter matrix produced the recorded data; the most that one can achieve is to constrain the family of parameter matrices that are consistent with measured data as tightly as possible.

In this context, Bayesian inversion provides a general method to define the statistical distribution of parameter matrices that fit the data and assigns the probability density of each parameter matrix given the data $\rho(m|\, d)$, known as the posterior probability distribution function (pdf) which here is referred to simply as the *posterior*. Bayes theorem allows us to calculate the posterior as follows:

$$\rho(m|\, d) = \frac{\rho(d|\, m)\rho(m)}{\rho(d)} \tag{1}$$

where $\rho(d|\, m)$ is called the *likelihood* which describes the probability of observing data d if parameter matrix m is true, $\rho(m)$ is the prior probability distribution of m (here referred to as the *prior*), and $\rho(d)$ describes the marginal probability of the data post-experiment—also called the *evidence*. If parameters m are real-valued, then each of these distributions is a probability density, and Tarantola (2005) gives a clear exposition of how they should be defined.

Our aim is to improve knowledge about the parameter values. Following Bayes theorem, this can be achieved by increasing the amount of relevant information in the observed data set which is represented by the likelihood,

or increasing information in the prior pdf. The likelihood and prior have equal mathematical weight in Equation 1, and while most work focuses on adding information through better data or improved data processing and thus targets the likelihood, in this study we aim to inject more geological information about the parameters through the prior probability distribution. This paper focuses on the development and demonstration of a methodology that introduces the information and solves the resulting inverse problem efficiently.

Conceptual models are the hierarchically highest level of information in most geological studies. In the current context, they describe our understanding of which geological processes have influenced the current subsurface structure and composition. We consider two geological conceptual models in this study: sedimentary structures created either by terrestrial river channel systems or alternatively by marine parasequences. We represent each conceptual model by large sets of different geometries of rock types that might be generated by processes invoked in that model. These geometries are in turn represented by a group of neural networks which are trained to regress through each set, to allow other representative geometries to be generated efficiently.

The generation of river channel geometries and subsequent training of neural networks was already performed by Laloy et al. (2018): the resulting networks are available online and produce rudimentary maps that depict possible geometries of river channels in a background medium, parameterised by these two binary facies. We additionally introduce prior information about marine parasequence structures created by a geological process forward model called SedSimple (Tetzlaff, 2022). SedSimple simulates sedimentary deposition, erosion and transport over geological timescales given a base topography and relative sea level curve, to create a three-dimensional geological conceptual model simulation of the subsurface. Compared to other, often commercially available process models, such as *SLB's GPM* (Courtade et al., 2021; Otoo & Hodgetts, 2021), *DionisosFlow*™ (Al-Wazzan et al., 2021; Borgomano et al., 2020; Hamon et al., 2021), *SedsimX* (Snieder et al., 2021), or *CarboCAT* (Masiero et al., 2021), SedSimple requires less computational resource making it possible to run a large number of simulations, but at the cost of reduced complexity in modelled processes and hence in the produced simulations. We train neural networks to represent the information in a large set of geometries obtained from SedSimple simulations, to produce networks for shallow marine environments that mirror the fluvial networks of Laloy et al. (2018).

Parameterising the geological prior information using neural networks is important because geological process forward models cannot be used in inversion schemes directly. Their relationships between simulated geological geometries and controlling dynamic parameters and base (initial) topographies are typically both chaotically complex and very high dimensional. This makes it very expensive, if not practically impossible, to find a set of dynamic and topographic controls that produce geometries which fit observed data. Re-parameterising the geological simulations into a more convenient (neural network based) form allows us to find geometries similar in character to those produced by the GPM, but which also fit the geophysical data.

More generally, geological data are commonly available as examples (called statistical *samples*) of a conceptual model. These might be facies maps from geological cross sections, from field outcrops or from geological process model simulations, and each of these might require a specific set of physical processes (a conceptual model) to be invoked to explain their geological origin. While the true parameter matrix or image in our tomographic volume may be explained using the same conceptual model, it will never exactly match one of those observed or simulated samples. Therefore, the parameterisation method must be able to generate other geological cross sections or three-dimensional facies maps that conform to the same concept, in other words, which are similar but not identical to the given set of samples. In addition, the method must facilitate our goal to explore the space of possible subsurface geometries to find those that are consistent with the observed geophysical data.

Exploring high-dimensional parameter spaces is extraordinarily computationally demanding, a phenomenon referred to as the curse of dimensionality (Curtis & Lomax, 2001). In order to make this feasible, we must represent the geological information using fewer representative parameters, usually called *latent* parameters. In principle, we expect that this is possible because different points on geological facies maps are not spatially independent (indeed they are strongly correlated in space, as observed in all geological outcrops—Arnold et al. (2019)) and so may be supported by a lower-dimensional manifold of latent parameters (Arjovsky et al., 2017). Two common mathematical constructs that can be used to parameterise outcrops into latent variables are Variational Auto Encoders (VAEs) and generative adversarial networks (GANs)—both being types of Neural Networks. In this research, we choose GANs for their demonstrated generational quality over VAEs and their near-instantaneous generation of samples which enables more rapid inversions in our applications (Goodfellow, 2016).

We invert seismic arrival times for the latent parameter posterior distribution using two methods: Markov

chain Monte Carlo (McMC) and mixture density networks (MDN). The former method is computationally expensive but tends towards the correct solution in the limit of infinite sampling (Haberland et al., 2023; Mosegaard & Tarantola, 1995; Zhang & Curtis, 2020). The latter method is another neural network method which has been used successfully to obtain marginal posterior distributions in a travel time tomographic setting (Earp et al., 2020; Earp & Curtis, 2020; Meier et al., 2007a, 2007b; Meier et al., 2009; Taroudakis & Smaragdakis, 2016). Post-training, the MDN produces an estimate of the posterior distribution in near-real time, and when combined with the GAN described above, we produce a method that rapidly inverts new data to estimate solutions of the Bayesian non-linear tomography problems that include geological prior information.

The importance of selecting *appropriate* prior information is highlighted by Kass and Wasserman (1996). Previous methods for including geological information in prior distributions include using a multi-point statistical method to simulate geology (González et al., 2008; Lochbühler et al., 2015), Hidden Markov Models (Feng et al., 2018; Moja et al., 2019; Nawaz & Curtis, 2016) and more recently using Neural Networks (Laloy et al., 2018; Mosser et al., 2020; Song et al., 2021a, 2021b). In this paper, we also analyse cases where inappropriate geological prior information is imposed on the problem and demonstrate that in principle such cases can be detected and corrected.

In subsequent sections, we introduce our methodology in six sub-sections: first travel time tomography, then generative adversarial networks, followed by Markov chain Monte Carlo, and mixture density networks, then posterior classification probabilities, and lastly geological information. Thereafter, we describe our specific worked example, followed by the results, a discussion and a summary of our conclusions.

2 | METHODOLOGY

2.1 | Travel time tomography

The time that energy takes to travel between two points in a medium contains information about the part of the medium through which it propagated. In seismic or acoustic tomography, the travel time stores information about variations in wave slowness (the reciprocal of wave speed) averaged over the propagation path. If multiple energy source and receiver locations are used, each travel time corresponds to a different path. In seismic tomography, we use different travel times to estimate the spatial distribution of slowness or velocity in the Earth's subsurface (Aki & Lee, 1976).

The likelihood $\rho(d|\,m)$ in Equation 1 compares the travel times that would occur through a proposed parameter matrix to the observed travel times. We therefore need to compute synthetic travel times from any proposed parameter matrix, which we achieve by solving the Eikonal equation

$$(\nabla t)^2 = s^2 \tag{2}$$

with $s(x)$ the medium slowness and $t(x)$ the arrival time from a fixed source location to any other location x. Equation 2 can be solved efficiently using a finite difference approximation (Podvin & Lecomte, 1991; Rawlinson & Sambridge, 2004), where a finer discretization of the simulation provides more accurate results.

The comparison between the travel times of the proposed parameter matrix and the observed travel times allows a gradient direction in parameter space to be computed that should infinitesimally improve the data fit. In linearised travel time tomography, we iteratively update the slowness or velocity parameters by a small perturbation in that direction until a suitable data fit is achieved. Unfortunately, using that approach it is never clear whether an approximately correct parameter matrix has then been found, due to the extensive and complex minima in the data misfit function. Therefore, in MDN tomography, we train a neural network to estimate directly the distribution of all slowness or velocity parameters that fit the data within their uncertainties, as explained below.

2.2 | Generative adversarial networks

We store prior information about the geological concepts inside a generative adversarial network (GAN). GANs were introduced by Goodfellow et al. (2014) to generate high-dimensional samples efficiently from a relatively low-dimensional space of so-called *latent* parameters. A GAN consists of two separate NNs: a generator and a discriminator as shown in Figure 1. We train the generator to generate high-dimensional outputs that approximate simulations from a training distribution (represented by a set of samples from that distribution called the *training set*). We train the discriminator to discriminate between simulations coming from the training distribution and outputs created by the generator. The training is adversarial in the sense that the discriminator is trained to minimise a loss function while the generator is trained to maximise the same loss, such that the distribution of outputs of the generator approaches the training distribution, at which point the discriminator can no longer discriminate between the two distributions. After training, the discriminator is discarded and the generator is used as an efficient

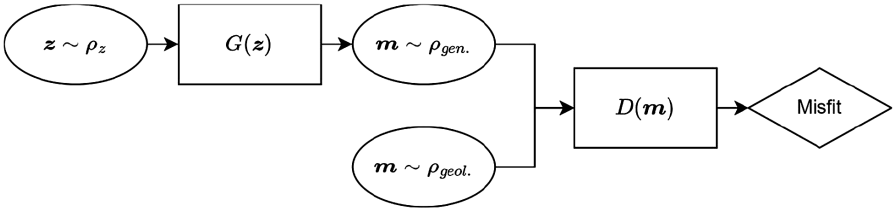


FIGURE 1 Schematic representation of a generative adversarial network. The discriminator D and generator G are two independent Neural Networks. G is trained to map a latent sample $z \sim \rho_z$ to a generated simulation $m \sim \rho_{\text{gen.}}$. D is trained to discriminate between simulations coming from the generator $m \sim \rho_{\text{gen.}}$ and geological simulations coming from the training set $m \sim \rho_{\text{geol.}}$. Symbol ~ means 'distributed according to'.

mapping from the lower-dimensional latent space to a higher-dimensional simulation space. In this study, the high-dimensional simulations are geological parameter matrices m, and we use two specific variants in the type of GAN to represent the generator: the Spatial-GAN (Jetchev et al., 2016) and the Wasserstein-GAN (Arjovsky et al., 2017).

The input to the generator G is a latent vector $z \sim \rho_z$, where ρ_z is a distribution that can be chosen such that samples can be drawn easily (where ~ means 'distributed according to'); typically, a Uniform or Gaussian distribution is chosen. The number of dimensions in ρ_z has to be sufficiently large to describe the low-dimensional manifold of the training distribution. Unfortunately, the dimensionality of this manifold is unknown so a suitable number of dimensions in z is found by trial and error. G can be thought of as a mapping from the latent space to the high-dimensional generated parameter matrix space, denoted as $G(z;\theta_g)$, where θ_g represents the network parameters; this mapping is referred to simply as $G(z)$ for brevity from hereon.

The input to the discriminator is a sample $z \sim \rho_z$ randomly selected from either the generator $G(z)$ or the geological training distribution $m \sim \rho_{\text{geol.}}$. The discriminator outputs a scalar representing the probability that m is a sample from the geological distribution (the training set). We train D to maximise the output probability of $D(m \sim \rho_{\text{geol.}})$ and minimise the output probability of $D(G(z \sim \rho_z))$. Conversely, G is trained with the opposite goal to maximise $\log[1 - D(G(z))]$. The loss for the GAN as a whole can be described by the value function $V(G,D)$:

$$\min_G \max_D V(D,G) = \mathrm{E}_{m\sim p_{\text{geol.}}(m)}\log[D(m)] + \mathrm{E}_{z\sim p_z(z)}\log[1-D(G(z))] \quad (3)$$

We do not directly optimise for $V(G,D)$ but rather update the two networks separately by alternately minimising the following loss functions for D and G respectively:

$$L_D = -\mathrm{E}_{m\sim\rho_{\text{geol}(m)}}\log[D(m)] - \mathrm{E}_{z\sim\rho_z(z)}\log[1-D(G(z))] \quad (4)$$

$$L_G = \mathrm{E}_{z\sim\rho_z(z)}\log[1-D(G(z))] \quad (5)$$

The adversarial objectives of the two networks imply that we seek an equilibrium between the two. Unfortunately, training to find equilibria is difficult (Salimans et al., 2016). There are optimizers that find equilibria but none are available for the non-convex cost functions and the continuous and high-dimensional parameter spaces that occur when training GANs. Instead, a gradient descent algorithm is used that finds a low value of a cost function, and alternating between updating G and D a pseudo-equilibrium is found. Since the optimisation algorithm is sub-optimal, the GAN may fail to converge during training, and it is common for multiple GANs to be trained in order to find one that performs well.

Convergence during training may be promoted by using a different loss function for D that measures the distance between two distributions: ρ_{geol} and $\rho_{\text{gen.}}$. Arjovsky et al. (2017) show that the Jensen-Shannon distance used in Equation 4 may not provide a gradient towards the solution in all scenarios, as it is not always differentiable. Updating Equation 4 to use the so-called Wasserstein distance measure can mitigate this problem. L_D then becomes

$$L_D = -\mathrm{E}_{m\sim\rho_{\text{geol.}}}D(m) + \mathrm{E}_{z\sim\rho_z}D(G(z)) \quad (6)$$

The Wasserstein distance is shown to be continuous and differentiable almost everywhere (Arjovsky et al., 2017). The improvements to GAN training include improved training stability and reduced mode collapse (the latter is the term used to describe situations where multiple high-dimensional features are mapped to the same latent parameter by the GAN, thus restricting its generation capability).

Jetchev et al. (2016) updated the GAN architecture to make it better suited to synthesising textures or maps; the updated architecture is called a Spatial-GAN or SGAN. Texture synthesis is the generation of samples of a given texture, which is defined as repeating patterns with some degree of stochasticity (Georgiadis et al., 2013). Geological

parameter matrices can be similar to textures as they often contain approximately repeating patterns and have some stochasticity (e.g. repeating sedimentary layers with varying thicknesses, or meandering river channels consisting of similar facies). The input to an SGAN is a latent tensor (matrix) rather than simply a latent vector as used in a standard GAN. Furthermore, all the layers in the SGAN are convolutional layers (Jetchev et al., 2016). This enables us to scale the input tensor to obtain a different-sized output, that is in our case, a larger latent tensor results in a larger geological parameter matrix. What is more, individual elements of the latent tensor describe only a single, localised patch of the output parameter matrix. Thus, we can update a single patch of an output parameter matrix while keeping the rest of the matrix constant. The discriminator is also updated to output a loss for each entry of the input tensor.

In this paper, we invert arrival times directly for the latent parameters used in the GAN. The latent parameter distribution used when training the GAN is therefore the prior distribution. Unlike in traditional inversions where the prior distribution is unknown or assumed to be known, in this case, we choose the distribution of the latent parameters. After training, the latent distribution represents the complex distribution over the training images, therefore the latent distribution represents $\rho(m)$ in Equation 1. Note that this approach does introduce a systematic error if the GAN produces incorrect simulations. We therefore encourage the use of large training sets and an appropriate number of latent parameters to encapsulate all variability within the geological prior distribution.

2.3 | Markov chain Monte Carlo

We aim to make a reasonably accurate estimate of certain statistics of the posterior distribution $\rho(m|\,d)$ in Equation 1. We could sample the prior distribution $\rho(m)$ and compute the likelihood for those samples to obtain an estimate of the posterior distribution. However, the maxima in $\rho(m)$ do not necessarily align with the maxima in $\rho(m|\,d)$ which would make sampling of the posterior inefficient. What is more, the likelihood $\rho(d|\,m)$ may introduce non-linear relationships that further impede finding a representative sampling of $\rho(m|\,d)$. We would therefore like to sample $\rho(m|\,d)$ directly, which is approximately possible using Markov chain Monte Carlo (McMC) sampling (Mosegaard & Tarantola, 1995).

McMC sampling creates a chain of samples, where each sample is found using a two-step process: first, we sample a parameter according to a proposal distribution $q(m'|\,m)$ which describes the probability that we should consider a move to m' given that we are at m. Second, the proposed sample is accepted or rejected depending on the probability of acceptance

$$P_{\text{accept}} = \left\{ 1, \frac{\rho(m')\rho(d|\,m')q(m|\,m')}{\rho(m)\rho(d|\,m)q(m'|\,m)} \right\} \tag{7}$$

where we assume an explicit form for the likelihood ratio (central terms in the fraction) of $\exp -\frac{\Delta S}{\sigma^2}$, where $\Delta S = S(m') - S(m)$ which is the difference in misfits S to travel time data between the proposed and current samples m' and m, and σ^2 is the variance or noise on the observed travel time data. Thus, for similar prior probabilities and if q is symmetric in m' and m, if the misfit of the proposed parameter m' is lower than the current parameter m, then m' is always accepted as a new sample of the posterior. In the reverse case, m' is accepted as a posterior sample with a probability based on the difference between the misfits of the proposed and current model. If m' is rejected, the current model is repeated (duplicated) in the chain.

From an initial parameter sample, consecutive samples are found iteratively using Equation 7. Metropolis et al. (1953) and Mosegaard and Tarantola (1995) show that after an infinite number of samples, the set of parameter samples is distributed according to the posterior distribution. However, an infinite number of samples is computationally infeasible, as is a number of samples that is sufficiently large to approximate this case in practical, non-linear tomographic problems. Therefore, multiple chains are computed in parallel, each with different initial parameters, such that we obtain a greater diversity of samples more rapidly. What is more, having multiple chains allows for more resilience if a chain gets stuck in a maximum.

Although the acceptance probability in Equation 7 ensures that we end up with a set of samples that estimate the posterior distribution if sampled infinitely many times, we would like a finite set of samples that are representative of the posterior distribution. What is more, we would like to obtain this finite set efficiently. Therefore, we must design our proposal distribution q such that we minimise the number of rejected samples while still spanning the parameter space. We define our proposal distribution as randomly selected perturbation to parameters of the previous parameter matrix in the chain. We can vary both the magnitude of the perturbation and the number of parameters updated to optimise the efficiency of posterior distribution sampling. Even after such provisions, samples are only ever approximately distributed according to the posterior. We therefore often only analyse statistics of the sample set (usually moments such as means, variances, etc.). These are assumed to be more robust estimators of properties of the posterior distribution than are individual samples.

2.4 | Mixture density networks

Mixture density networks (MDN) are a type of neural network that can be trained to infer the posterior distribution of parameters directly from measured data (Bishop & Nasrabadi, 2006). The posterior $\rho(m|\,d)$ is approximated by a weighted sum of multiple Gaussians

$$\rho(m|\,d) = \sum_{k=1}^{K} \left(m|\,\mu_k(d), \sigma_k^2(d)\pi_k(d)\right)N \tag{8}$$

where K is the number of Guassians, π_k the kth Gaussian's mixing coefficient or weight, N the Gaussian or Normal distribution, μ_k the Gaussian mean and σ_k^2 the Gaussian standard deviation. The vector parameters π, μ and σ are inferred from the data using a neural network which may have any of a range of architectures and complexities.

MDN training is performed with N parameter–data pairs $\{(m_n, d_n): n = 1, \ldots, N\}$, which are generated by selecting m_n according to the prior pdf and calculating the corresponding measured data d_n using a synthetic forward model. The neural network weights θ_{MDN} are optimised by minimising a cost function E which for independent training data is taken to be

$$E(\theta_{\mathrm{MDN}}) = -\sum_{n=1}^{N} \ln\left[\sum_{k=1}^{K} \pi_k(d_n, \theta_{\mathrm{MDN}}) N\left(m|\mu_k(d_n, \theta_{\mathrm{MDN}}), \sigma_k^2(d_n, \theta_{\mathrm{MDN}})\right)\right] \tag{9}$$

with N the number of pairs in the training set (Bishop & Nasrabadi, 2006). To optimise the network, we calculate derivatives of the cost function with respect to each network weight, which are obtained using a backpropagation procedure (Bishop & Nasrabadi, 2006). Due to the sum over all data points in Equation 9 we backpropagate the derivatives for each data point and then sum the resulting N derivatives to find the derivative of $E(\theta_{\mathrm{MDN}})$. If samples from the prior distribution are used as training data for the network then the prior becomes implicit in the training procedure. After training, the resultant network then directly estimates the complete posterior probability distribution given any data set as input.

While a single MDN can predict the posterior distribution directly, we opt to train multiple networks with different layer sizes. As MDNs are prone to mode collapse, different networks may find different modes of the posterior distribution. The ensemble of predictions is combined linearly in a so-called mixture of experts (Dietterich, 2000).

2.5 | Posterior classification probabilities

We aim to assign any input travel time data to one of a set of discrete classes, each defined by the geological concept model under which similar training data were generated. A neural network tasked with the classification of its inputs into a set of discrete classes is often trained to output a score for each possible class. The classification derived by the network is then the class with the greatest score. If the network is trained with particular cost functions, the outputs for each class become estimates of the Bayesian posterior probabilities (Richard & Lippmann, 1991), provided that the network is trained using a one-hot encoding scheme (i.e. the true classification is encoded as 1 for the correct class and 0 for the other classes) and that the network has a sufficient number of trainable weights. Richard and Lippmann (1991) prove this for a squared error and cross-entropy cost function.

Assume that we have a training set containing a parameter matrix m which belongs to one of N classes in $\{C_n: n = 1, \cdots, N\}$. Let $\{y_n: n = 1, \cdots, N\}$ be the network output and $\{c_n: n = 1, \cdots, N\}$ the desired output. Then, we can construct the squared error cost function

$$E(\theta_{\mathrm{class}}) = \mathrm{E}\left\{\sum_{n=1}^{N} \left[y_n(m) - c_n\right]^2\right\} \tag{10}$$

with θ_{class} the network weights, $y_n(m)$ the network output for class n and c_n the desired output for class n (either 1 or 0) for each of the N classes (Richard & Lippmann, 1991). The neural network weights are then optimised using a similar backpropagation algorithm to the MDN (see Section 2.4). For the prior distribution, we assume a Uniform distribution over the different geological concept classes.

2.6 | Geological information

The conceptual geological model usually describes our beliefs about the tectonic setting, depositional environment of sediments, geographical relationships to continents and marine waters and other high-level information. This model ultimately governs lower-level information about the exact geometry, abundance of different sediment types, etc. We use two conceptual models: a braided river system and marine parasequences. The former model has already been described in detail in Laloy et al. (2018), so here we discuss the relatively new marine parasequences conceptual model.

Our choice of conceptual model defines what processes are included in the SedSimple (Tetzlaff, 2022)

geological process modeller (GPM) runs used herein. A GPM simulates geological processes through time computationally to obtain a three-dimensional distribution of geological facies such as that shown in Figure 2 (Tetzlaff & Harbaugh, 1989). Simulations are started from an initial topography, for which in this study we use a sigmoidal topography that represents the transition from the continental shelf to the abyssal plain. Relative sea level through time is defined to be sinusoidal and we include an influx of water and sediment on the shelf representing a river. Lastly, we define two types of clastic sediments by setting their relative transportability; one sediment has double the transportability of the other. Note that in this manuscript, we aim to establish and demonstrate the methodology rather than to apply it in a real setting. We therefore chose these values to obtain a thick sediment with interesting features rather than to emulate a scenario matching a certain geographical location and time interval as would be the case when this method was deployed in practise. All parameter values used herein are defined in Table 1.

After the initial parameters are defined, SedSimple simulates different geological processes at each of a set of small time steps. Such processes are sedimentary dispersion, erosion, transportation and deposition. Dispersion is the process that simulates sediment particles moving downhill from high to low elevations, and similarly from high to low concentrations when suspended in a fluid. The governing equation for sediment dispersion is

$$\frac{\partial z}{\partial t} = D\nabla^2 z + s \tag{11}$$

with elevation z, time t, diffusion coefficient D and the sediment source term s. Sedimentary erosion, transport and deposition are also dependent on the fluid flow. SedSimple simulates fluid flow to determine whether sediment is eroded, transported, and deposited, for which a simplified version of the Navier–Stokes equation is employed. Erosion and deposition are then calculated based on empirical formulae (see Tetzlaff & Harbaugh, 1989 for more information).

Although GPM algorithms are deterministic, they are still chaotically non-linear: a small change in the input variables can cause large changes in the output simulation.

TABLE 1 Parameters used for the geological process model simulation are shown in Figure 2. Transportability in the table is also known as the diffusion coefficient.

Parameter	Value
Geological time	500 kA
Manning coefficient	0.3
Water source	1000 m^3/s
Sediment 1 influx	60 mL/s
Sediment 2 influx	160 mL/s
Transportability sediment 1	100
Transportability sediment 2	200

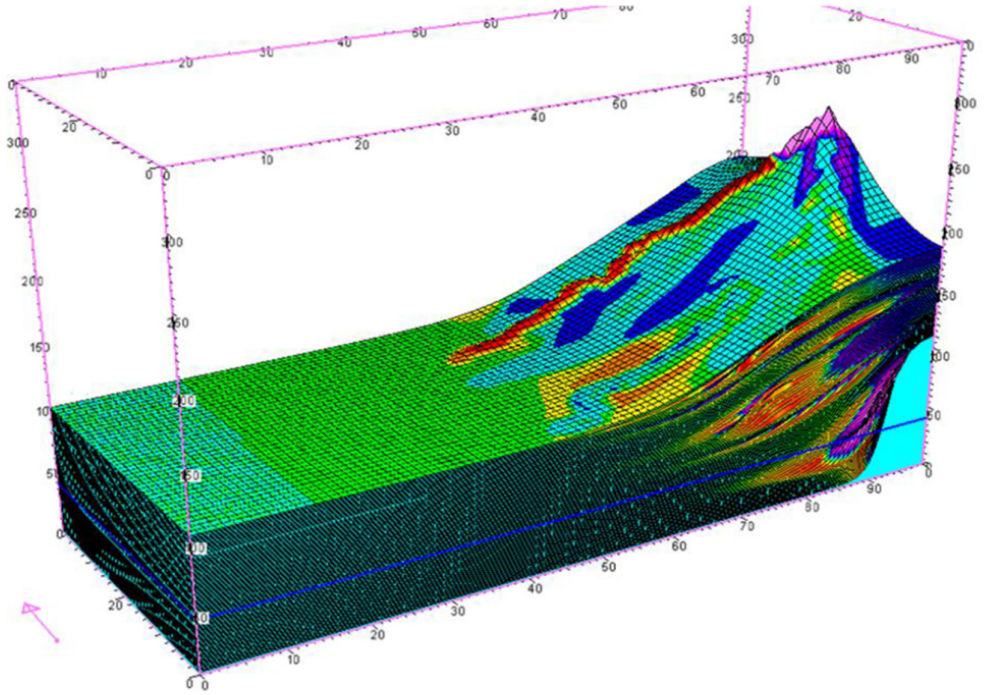


FIGURE 2 3D view of example output from the geological process model SedSimple. Colours in the plot represent the relative concentrations of different facies in the simulation.

For example, if the initial simulation has two parallel river channels, then tiny changes in the input parameters create entirely different braided river systems. It is therefore difficult, if not impossible, to use a GPM directly in an inversion algorithm to fit specific measured data. We therefore train a GAN to emulate the spatial sedimentary patterns produced by the GPM (Mosser et al., 2020). Since the generator in a GAN has an analytic structure, the input latent parameters can be varied so that the distribution of simulations produced emulates the prior distribution represented by any given set of simulations from the GPM. This method allows multiple GPM simulations with different initial parameters to be used to capture the chaotic nature of sedimentary distributions.

3 | PROBLEM DESCRIPTION

In this paper, we focus on how to introduce geological information into tomographic studies, in particular in cases where multiple conceptual geological models may describe the geological scenario under which the imaged structure was created. Given that we can deploy two such models (one parameterised by Laloy et al. (2018) and one using SedSimple), we can investigate the impact on geophysical tomography if different (and potentially misleading) prior information is injected.

The SedSimple simulation in Figure 2 was used to construct training data for the GANs that embody the second conceptual model. In this simulation, there are two different sediments, both siliciclastics but with different grain sizes. The colours in the plot represent the abundance of the two sediments. We extracted 1800 two-dimensional 32-by-32 pixel slices in both the x and y direction from the three-dimensional simulation that created a training set of parameter matrices for the GAN. For details of the braided river system conceptual model and SGAN training, we refer the reader to Laloy et al. (2018).

We trained a GAN to emulate the marine parasequences conceptual model using codes from Kang and Park (2020). We tested different network architectures including BigGAN (Brock et al., 2018), ReACGAN (Kang et al., 2021), ICRGAN (Zhao et al., 2020) and WGAN-GP (Gulrajani et al., 2017). For each network, we minimised the number of latent parameters by visual inspection of the output and trained the same architecture multiple times to reduce the effect of poor (random) network initialisation. The trained networks are visually compared after which we selected a single best network, in our case, a WGAN with eight latent parameters. Training on this network took 7 h and 30 min on a single NVIDIA TITAN X.

The parameter matrices generated by both the braided river and the marine parasequence GAN are in a value range of [−1, 1] which we rescaled to [1, 2] km/s to represent a reasonable range of seismic velocities. Four samples from each GAN are shown in Figure 3, the braided river channel realisations on the left and marine parasequences on the right represent samples from the two prior pdfs used in this study. Although the parameter matrices are high-dimensional, there is a low-dimensional latent representation for each such matrix. The braided river channel prior is encoded by nine latent parameters and the marine parasequences prior is encoded by eight latent parameters. The respective GANs create mappings from the low-dimensional latent parameters to the prior pdfs in high-dimensional geophysical parameter matrices. Each set of latent parameter values selected from the latent distributions produces an approximate sample from the corresponding geological prior pdf.

To represent a geophysical tomographic survey, we defined a square data acquisition geometry with corners at

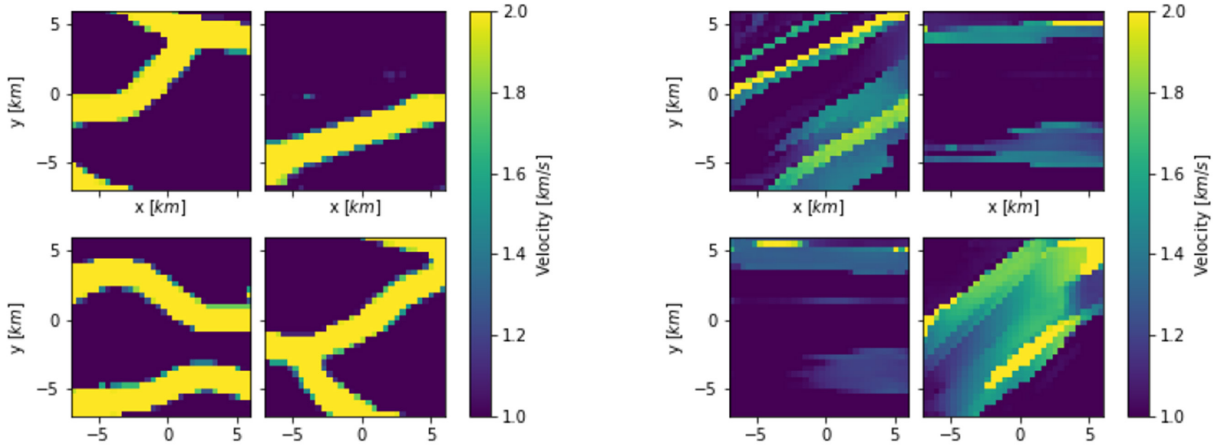


FIGURE 3 Four realisations from the braided river system (left) and marine parasequence (right) GANs. Colours represent seismic velocities.

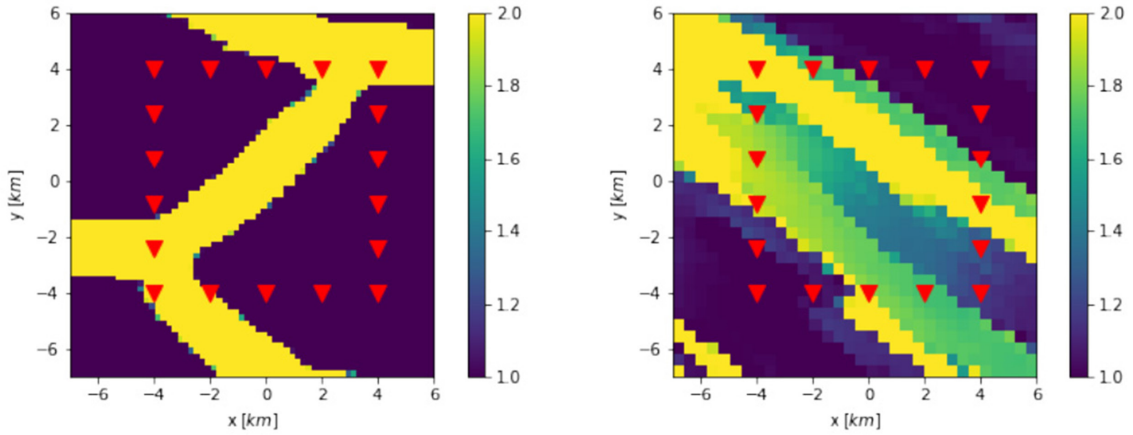


FIGURE 4 The two 'true' parameter matrices for which we simulate measured data, one from each of the conceptual models: a braided river system (left) and marine parasequences (right). Red triangles indicate the data acquisition geometry—the locations of both sources and receivers between which travel time data are simulated.

$x = -4, 4$ km and $y = -4, 4$ km and source spacing of 1.4 km as shown by the red triangles in Figure 4. This geometry defines the locations of sources and receivers: each location in turn acts as a seismic source, whereas all other locations act as receivers. This generated 153 unique travel times—the data corresponding to each parameter matrix. We add Gaussian noise with a standard deviation of 1% of the mean arrival times to the synthetic arrival times to simulate measurement uncertainties.

Figure 4 shows what we used as the true parameter matrices composed of a terrestrial river system (left) and marine parasequences (right). The acquisition geometry is indicated by red triangles. In the examples below, we inverted the travel times corresponding to these true parameter matrices for the Bayesian posterior pdf in the low-dimensional latent space in each case. Samples of this pdf could then be translated to the high-dimensional parameter space using each respective GAN. This indirect approach ensures that we need only estimate the posterior pdf over eight or nine latent parameters rather than across the high-dimensional 32-by-32 parameter matrix.

For each data set, we computed two estimates of the posterior pdf: a benchmark solution found using McMC and a large number of samples and a rapid estimate using MDNs. We validated the convergence of the McMC runs by monitoring the posterior marginal pdf estimates for each parameter in latent space. Any chains that were obviously stuck in local minima (span a relatively narrow range of parameter values) were removed and we validated that a reasonable number of samples have been taken by ensuring that the posterior distribution is essentially the same for the complete set as well as half of the set of posterior samples. We computed 40 chains with around 2 million samples each which took approximately 3 days to run. This is a lot of samples for McMC runs with ca. 8 parameters, but the latent parameter space is more information dense than model parameter space and posterior pdfs in the latent space can therefore be strongly multimodal. We therefore run the chains for many samples to avoid remaining trapped in a subset of the modes.

A real basin example is provided in the form of geological models from a mapped outcrop from Last Chance Canyon (New Mexico, USA) interpreted by Sonnenfeld and Cross (1993). Geological models were extracted from the interpretation by digitising their outcrop model, assigning velocities to each of the facies, selecting suitable outcrop partitions, and generating synthetic arrival times for these partitions. We invert the arrival times from these geological models using the MDN to obtain geological posterior estimates. This process is likely to create a geological model that lies significantly outside of our prior information and training sets, since those were generated using a significantly different and simplified set of geological prior information (that encoded in the GPMs, their initial conditions and dynamic parameters). This therefore represents a particularly stringent test of our methods.

For each prior pdf, we trained an MDN to invert the data using around 3 million samples in each training set. During training, we monitored progress by evaluating a validation data set of 25% of the size of the training set. Multiple network architectures were used and each architecture was trained multiple times to eliminate bias due to their random initialisation. Using a test set of size 5% of the training set size we measured the network performance. To better generalise the outputs, we selected five networks and combined them in a linear mixture of experts (Dietterich, 2000). We found that training the MDN for the complete, multi-dimensional posterior resulted in the MDN not finding all the posterior modes present in the 'correct' McMC posterior estimate. We therefore

opted to train multiple networks to infer the marginal pdf of each individual latent parameter, similar to the tomographic work of Earp and Curtis (2020). Thus, for a single prior, we have $N_{\text{latent parameters}} \times M_{\text{mixture of experts}}$ neural networks (e.g. the river channel prior pdf has nine latent parameters giving $9 \times 5 = 45$ networks). Training of a single network took about 90 min using a single NVIDIA Tesla K80. We had access to multiple GPUs so we could train the networks in parallel.

Finally, we trained a classifier NN that estimates the posterior probability that each conceptual model (embodied in one of the two prior pdfs) pertained to a certain set of travel times. For this, we combined the river and marine training sets such that we obtained a total of 6 million data points. We trained multiple networks with different architectures and selected the one that performed best on a validation dataset. We randomly selected a number of samples for each training epoch. For the optimal network, we used 8 batches with 4000 samples each and trained the network for 100 epochs. This took 80 s on a NVIDIA Tesla T4.

4 | RESULTS

Figure 5 shows the estimated geological posterior statistics for a river channel true parameter matrix where the inversion is performed using the correct prior; latent posterior distributions and summary statistical geological models are found in Appendix A. Figure 5 shows from left to right: the true parameter matrix, the posterior mean from the MDN (top) and McMC (bottom), the posterior estimate of the standard deviation for MDN (top) and McMC (bottom) and lastly a histogram of the arrival time misfits for 5000 posterior samples from MDN (top) and McMC (bottom) solutions (the MDN samples are independently selected from the posterior marginal pdf of each latent parameter). The histogram shows $|d_{\text{obs}} - d_{\text{sample}}|$ over different data points and different models. The means for both posterior estimates are close to the true parameter matrix. The standard deviations show high uncertainty loops around the boundaries of features in the true parameter matrix, similar to the results of Earp and Curtis (2020); these are expected, and quantify uncertainty in the location of the edges of those features (Galetti et al., 2015). The posterior estimates thus show that there is a river channel running diagonally over the parameter matrix but that the exact boundary of the channel is uncertain. The channels outside of the survey acquisition area have broader uncertainties since fewer rays travel outside of the acquisition array. The travel time misfits are centred around one travel time measurement standard deviation σ, and almost all are within $\pm 2\sigma$ from the true arrival times. Finally, all statistics shown here are consistent between our benchmark McMC posterior and the rapid MDN method, and since these are completely independent methods this attests to the robustness of both. All of the above indicates that both the McMC and MDN solutions are approximately correct, and conform to intuition about probabilistic solutions derived from analyses in previous studies.

Figure 6 shows similar plots to Figure 5 but for the marine parasequence true parameter matrix and prior information. Again, there is a close match between the posterior mean and the true parameter matrix. The McMC posterior resolves the high-velocity feature at location (1, 1) km slightly better than the MDN posterior mean

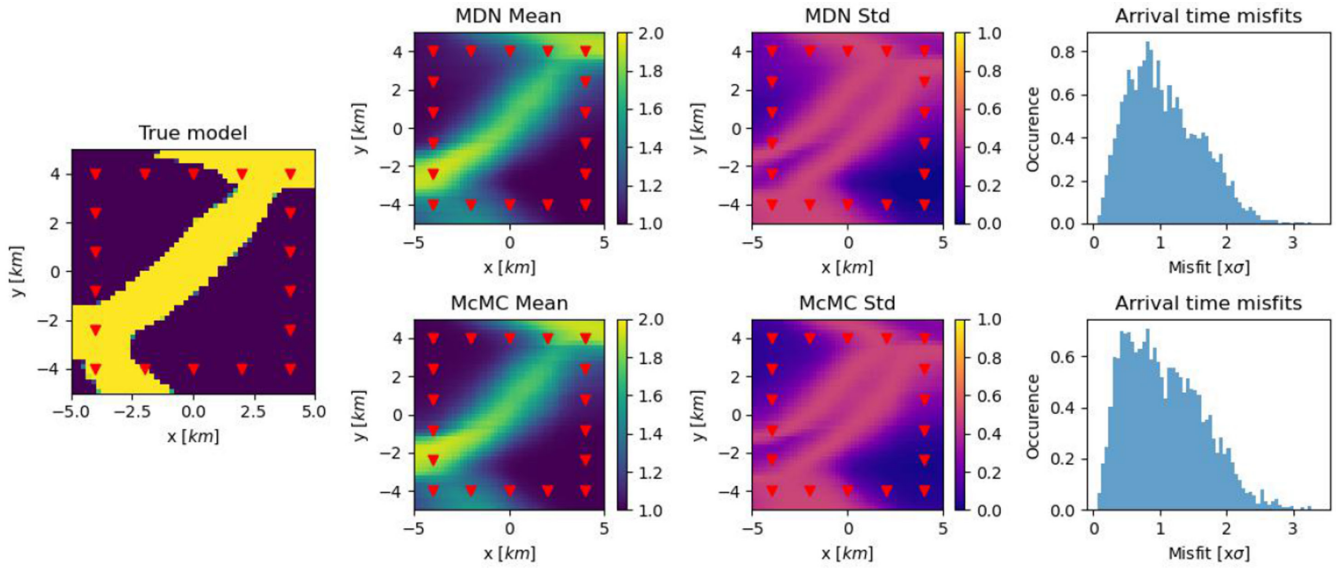


FIGURE 5 The true braided river channels parameter matrix (left) and the summary statistics for the posterior estimates from the MDN (top) and McMC (bottom). The statistics are the posterior mean, posterior standard deviation and the travel time misfits.

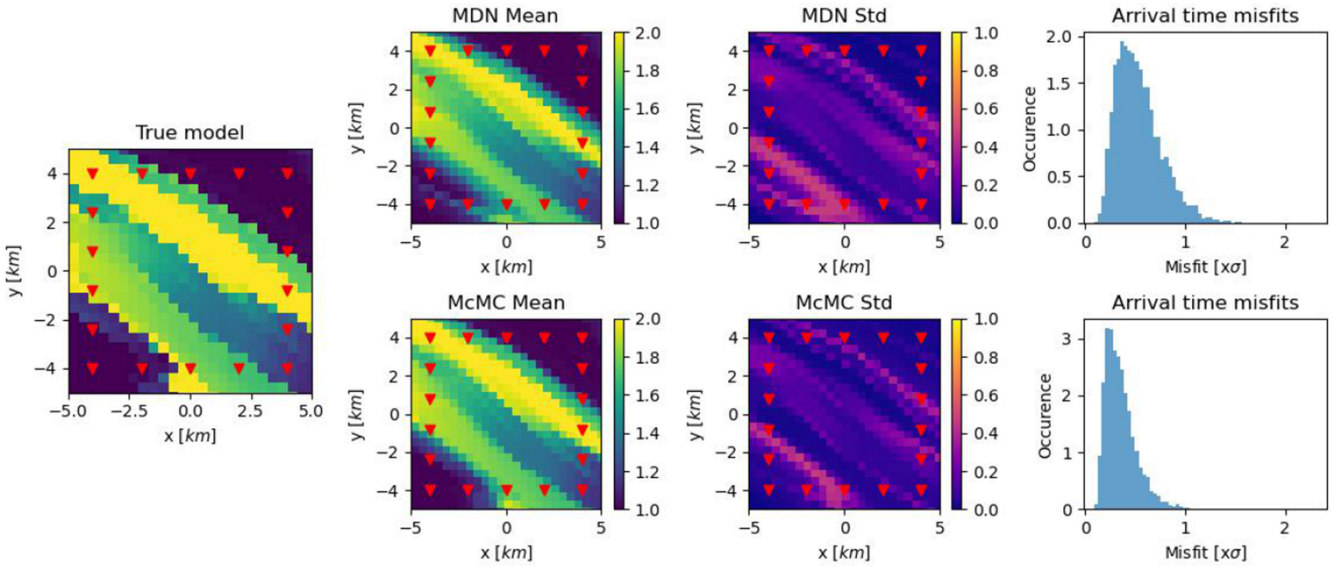


FIGURE 6 The true marine parasequences parameter matrix (left) and the summary statistics for the posterior estimates from the MDN (top) and McMC (bottom). The statistics are the posterior mean, posterior standard deviation and the travel time misfits.

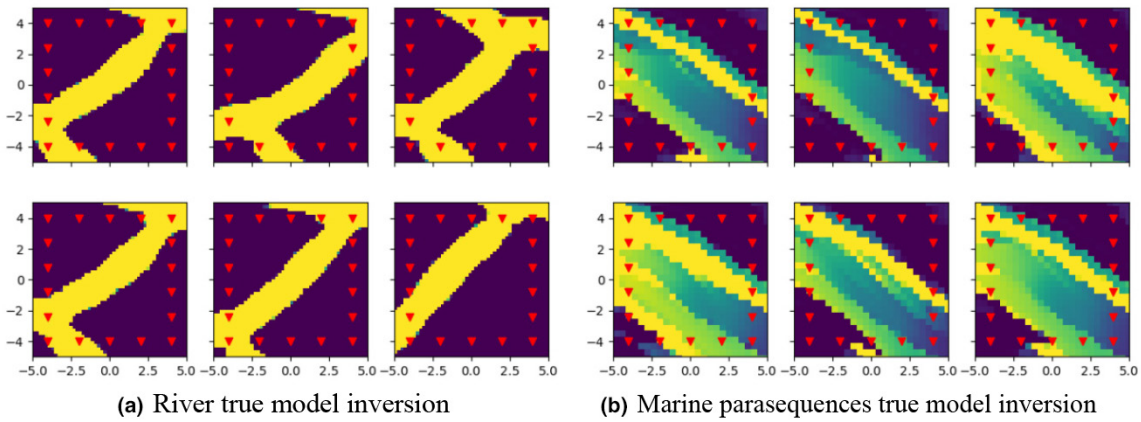


FIGURE 7 Samples from the MDN posterior marginal pdf estimates. (a) shows samples from the inversion in Figure 5 and (b) shows samples from the inversion in Figure 6.

velocities. The MDN posterior standard deviations are wider for the transition from low to high velocity at (−3, −2.5) km compared to the McMC standard deviations. The McMC posterior estimate is thus more narrowly concentrated around the true parameter matrix. This is confirmed by the arrival time misfit histograms: the McMC posterior data misfits show a narrower peak at a lower misfit value compared to the broader peak of the MDN posterior data misfit. This illustrates that the MDN posterior marginal pdf estimates do not capture all of the information that is contained in the McMC posterior. This is likely to be because we infer only single-parameter MDN marginal distributions which therefore do not contain information about correlations between latent parameters; this contrasts with, the Monte Carlo samples which are taken in the full latent space and so do contain correlation information. While in principle, it is possible to train MDNs to represent the fully correlated posterior pdf, we found such networks extremely difficult to train reliably. Therefore, this slight reduction in posterior information is the price paid for obtaining stable solutions for any travel time data set in ca. 1 s from an MDN rather than from days of computation when using McMC.

The statistics in Figures 5 and 6 only show summary statistics of the posterior pdf. What is more, the posterior mean is not in itself a geological parameter matrix selected from the posterior distribution (the values shown are an integral over all parameter matrix samples). We therefore show six random samples from the MDN posterior estimate in Figure 7a,b for the river and marine inversion respectively. The samples in each set are slightly different but all do resemble their respective true parameter matrix. What is more, all samples are geological: the samples from the river inversion all show reasonable representations of

river channels and all samples from the marine parasequence inversion show reasonable geological marine parasequences. Geological realisations are highly unlikely to occur in any inversion conducted using a non-geological prior (e.g. Bodin & Sambridge, 2009; Earp & Curtis, 2020; Galetti et al., 2015).

So far we have applied the correct set of prior information for each of our target true parameter matrices. However, it has been shown that prior choices between conceptual geological models (Bond et al., 2015) and their parameters (Curtis & Wood, 2004) are subject to natural human biases (Bond et al., 2012; Curtis, 2012; Polson & Curtis, 2010). It is therefore of interest to assess the effects of using incorrect prior information: inverting a river channel true model using a marine prior pdf and vice versa. Figures 8 and 9 show the inversions using the incorrect prior pdf (top) versus the correct one (bottom), and Figure 10a,b shows corresponding example posterior samples for the river and marine true parameter matrices, respectively. The posterior statistics show that the inversions with the correct prior perform better than those with the incorrect prior, as expected. However, the incorrect prior

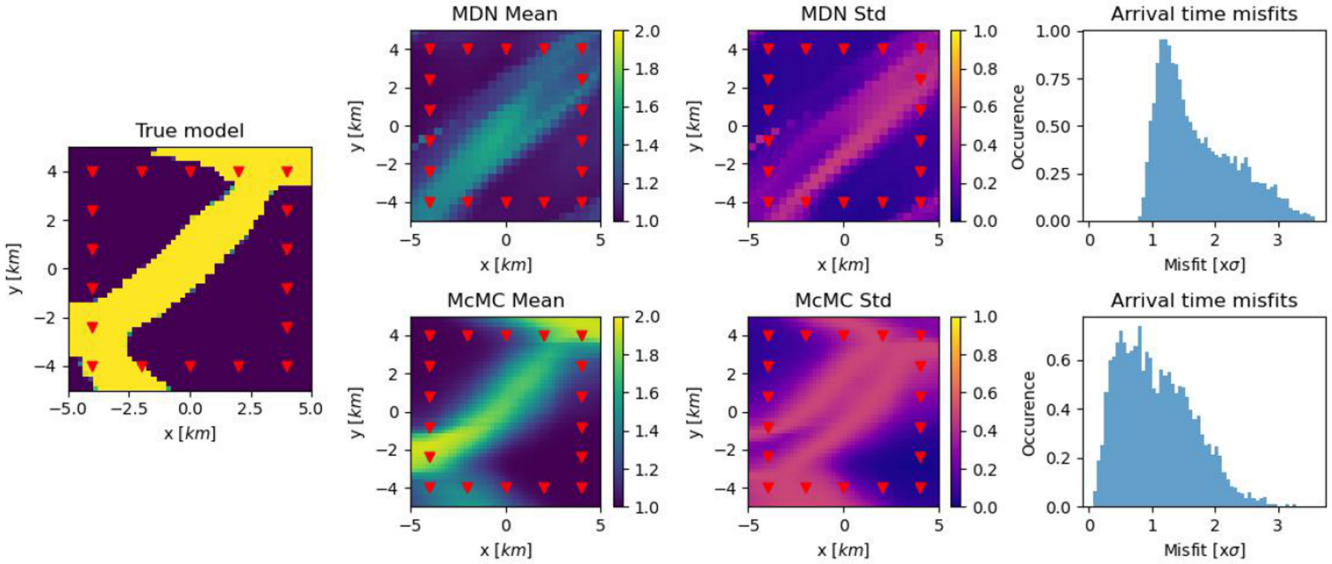


FIGURE 8 The true parameter matrix taken from the river prior (left) and the summary statistics for the posterior estimates using the unsuitable marine prior pdf (top) and suitable river prior pdf (bottom). The unsuitable posterior pdf estimates are computed using the MDN.

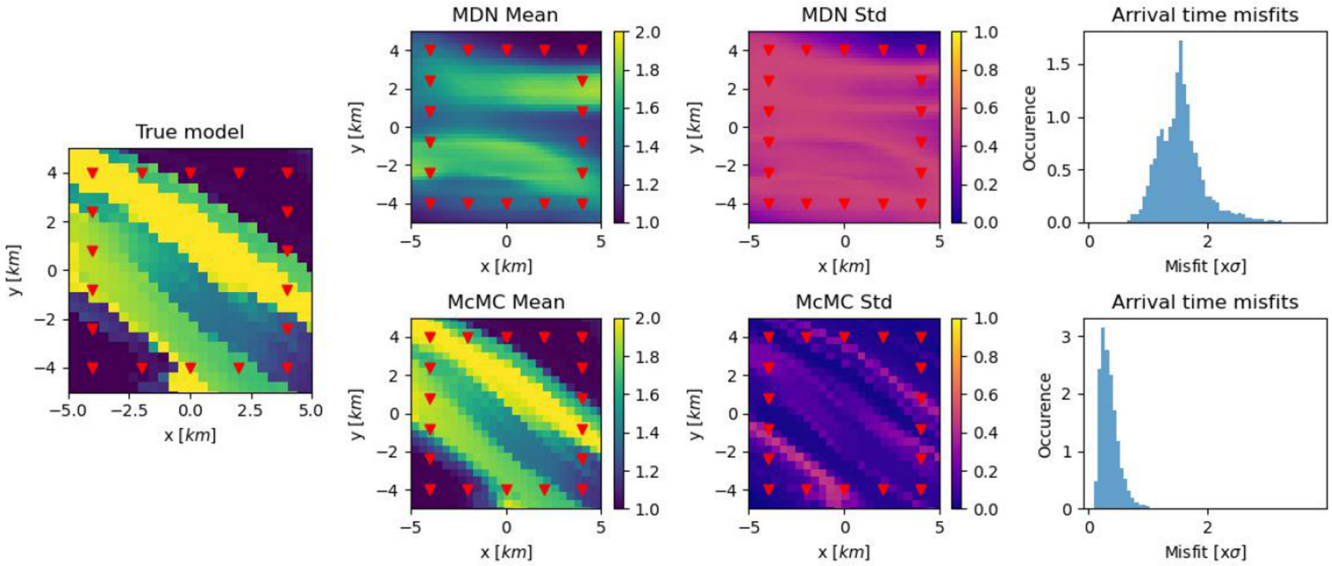


FIGURE 9 The true parameter matrix is taken from the marine parasequences prior (left) and the summary statistics for the posterior estimates using the unsuitable river prior pdf (top) and suitable marine parasequences prior pdf (bottom). The unsuitable posterior pdf estimates are computed using the MDN.

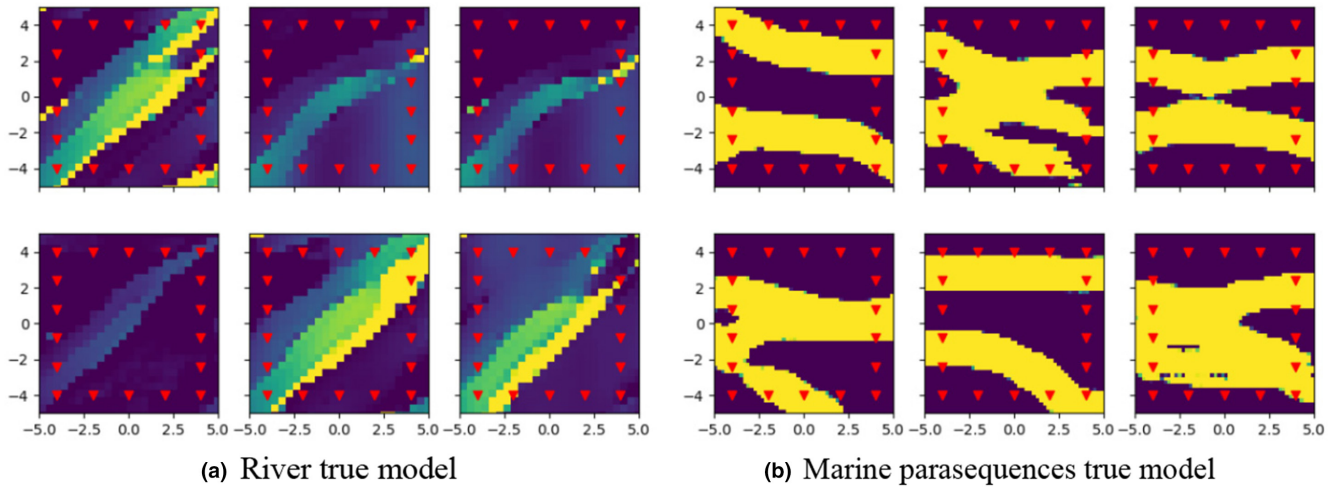


FIGURE 10 Samples from inversions using unsuitable prior information. (a) shows the samples from the inversion in Figure 8 and (b) shows the samples from the inversion in Figure 9.

inversions do still retrieve approximately recognisable estimates of the overall structure. The inversion of the river channel data produces a diagonal channel-like feature in all samples in Figure 10a, but the channels sub-parallel to the x-axis at the top and bottom of the true model are not inferred. The posterior samples from the inversion of the marine data show that the posterior does somewhat capture the high-velocity feature in the true model, but that there is a lot of variation. However, in both cases when inappropriate prior information is used, the expected loop-like uncertainty structures created by high uncertainties occurring on the boundaries of the main velocity anomalies do not appear.

As a test in which the true model is even further outside of the prior information space, we invert three models that are sub-areas of the cross-sectional interpretation of Last Chance Canyon (NM, USA) from Sonnenfeld and Cross (1993). Inversion results using the marine prior are shown in Figure 11. From left to right, we show the true model, maximum a posteriori (MAP) model, the mean, the standard deviation and the arrival time misfits. The latter three are calculated over geological posterior distribution samples. The top row shows a fair agreement between the true model and the posterior models. The middle ca. 1.75 km/s zone is recovered well, with uncertainty as to where the bottom boundary of that zone lies. The second row shows results for a more complicated true model. The two low-velocity features are represented by a single low-velocity zone in posterior estimates. Furthermore, the high-velocity zone at the top in the posterior results is an overestimation of what is seen in the true model. Lastly, the bottom row true model has a large low-velocity zone diagonally over the model which is well represented in the posterior estimates. The boundaries of the low-velocity zone have higher uncertainty with a narrower band of uncertainty for the bottom boundary versus the top boundary. Nevertheless, the overall structure of each model is recovered reasonably for this real setting, despite prior information injected from the combination of conceptual and process models certainly being oversimplified.

The data misfit in Figures 8 and 9 highlight the difference between using correct and incorrect conceptual models. The top rows show posterior statistics obtained using unsuitable prior information and the bottom rows show comparable results when using appropriate prior information. Differences between the misfit distributions (rightmost panels) when changing the prior pdf for the same data set indicate that there may be information in the data that could discriminate which class of prior information is more likely to be appropriate. Assuming the prior probability over the geological concept classes is Uniform, we sample the prior of each geological concept and generate arrival times. The obtained set of arrival times is the training data for the classifier neural network. We train the classifier neural network that provides the posterior class probability of each conceptual model C given the observed travel times $\rho(C|\, d)$. Given the posterior class probability and the posterior model parameter probability (from previous results), we can also compute the joint posterior probability of the conceptual class and the corresponding model parameters $\rho(C, m \mid d)$. The results are shown in Figures 12 and 13 for the river and marine true parameter matrices respectively. The top panels show the following geological parameter matrices from left to right: the true parameter matrix, the posterior mean using marine prior information, the posterior mean using river prior information, the posterior mean given (marginalised over) the class probabilities from the classifier neural network and lastly, a posterior mean if the class probabilities were Uniform (i.e. the posterior class probability is equal

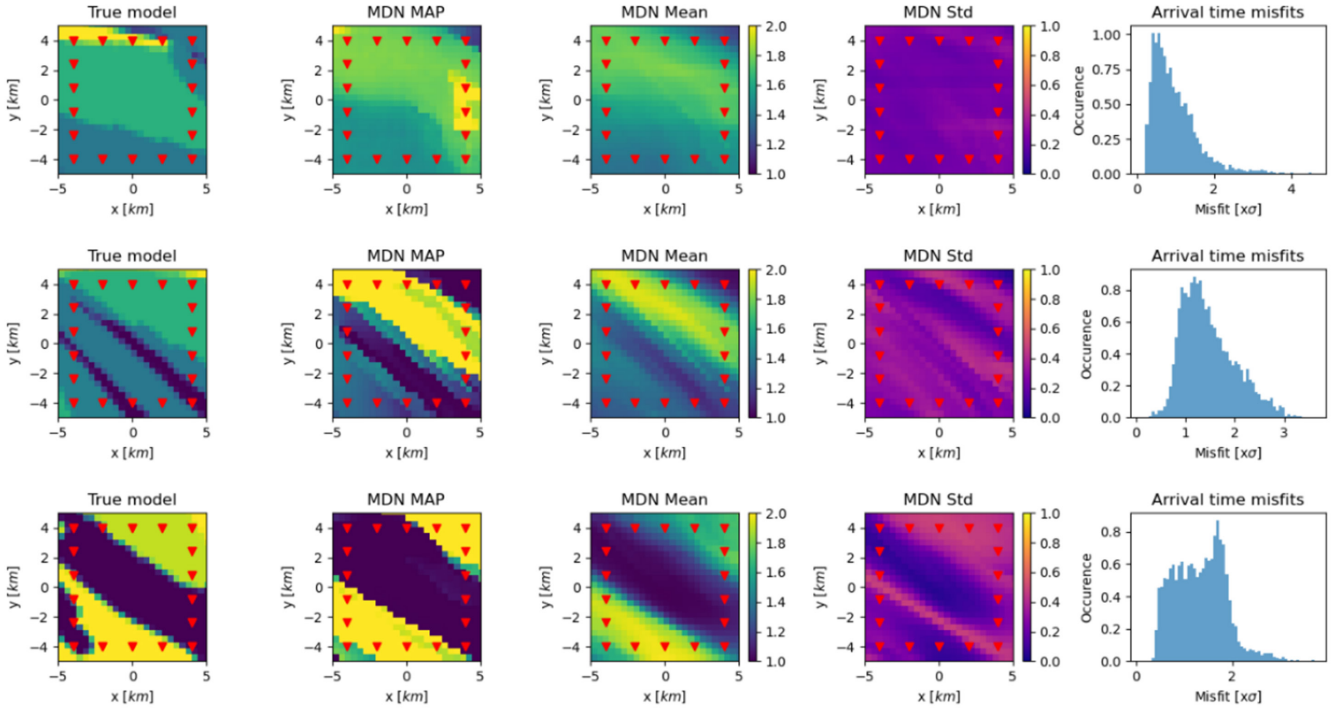


FIGURE 11 Posterior distributions for three true models coming from real basin data. The true models are constructed from the Last Chance Canyon (NM, USA) outcrop interpreted by Sonnenfeld and Cross (1993). From left to right, the panels show the true model, Maximum A Posteriori model, posterior mean, posterior standard deviation and arrival time misfits.

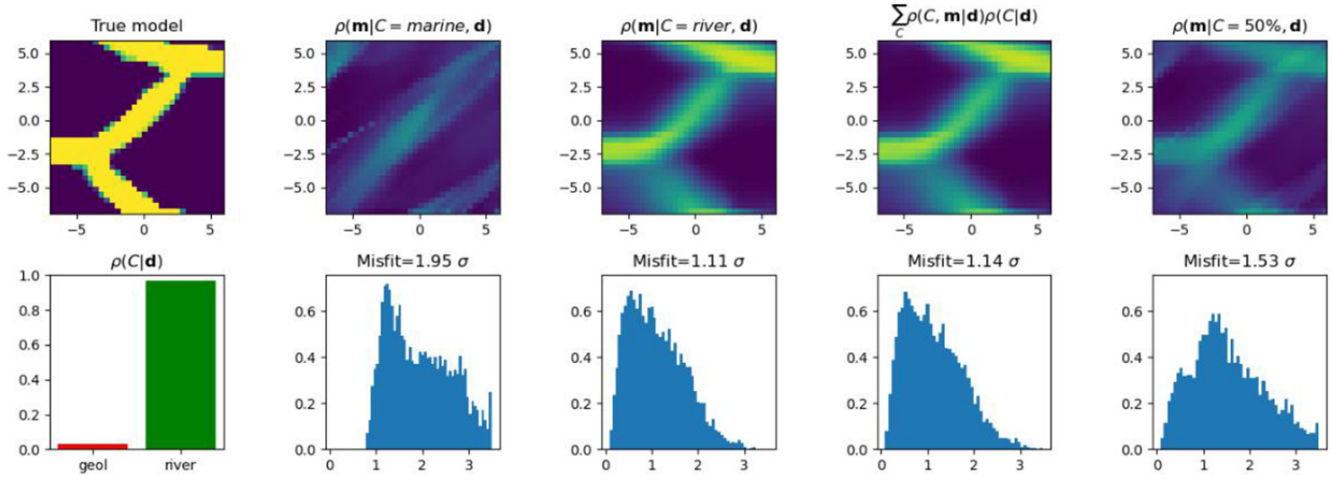


FIGURE 12 Posterior statistics for different prior pdfs. Left column shows the true river parameter matrix (top) and the posterior probability of the conceptual model class given the observed travel times (bottom). Successive columns show the posterior mean (top) and travel time misfit (bottom) for the following scenarios: marine parasequence conceptual model, river channel conceptual model, mixed posterior given the posterior class probability and the posterior mean if the posterior probability for each conceptual model class was Uniform.

to the prior class probability). The bottom row shows the posterior class probabilities on the left and then histograms of the data misfits corresponding to each posterior pdf in the top row. First, the posterior class probability provides clear information about which conceptual model pertains to which data set. As above, the posterior misfits are lower when the correct conceptual model is used for the inversion. If we combine the class probabilities and the parameter probabilities into the joint posterior probability $\rho(C, m \mid d)$, the misfits are higher than for the correct prior but significantly lower than those obtained using either the wrong class or Uniform class probabilities. For the data corresponding to the Last Chance Canyon models (Figure 11) the network correctly predicts the marine prior as the more appropriate prior for the travel time data. The likelihood of the data coming from marine environments

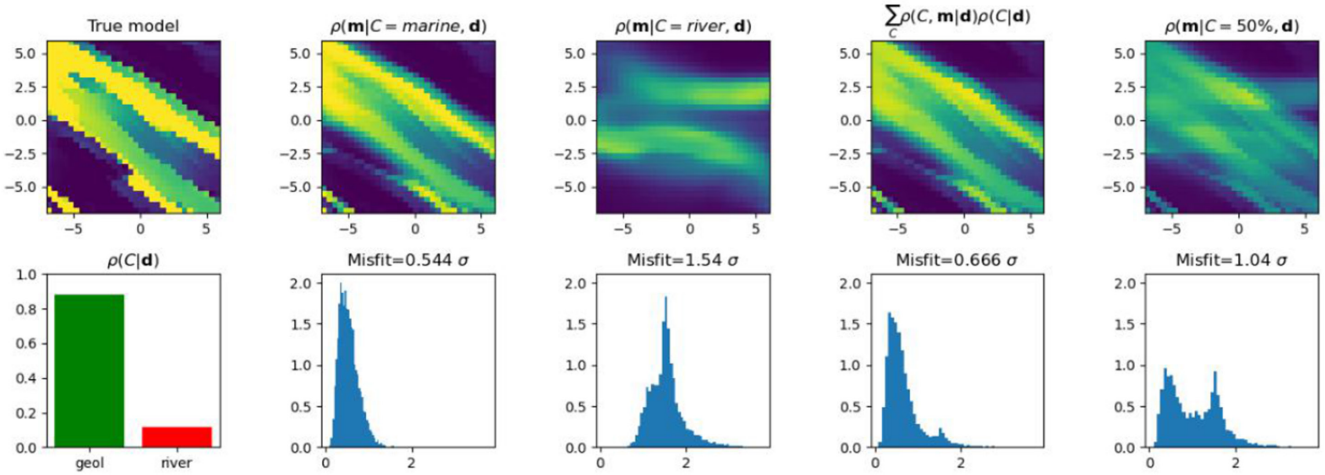


FIGURE 13 Posterior statistics for different prior pdfs. Left column shows the true marine parasequences parameter matrix (top) and the posterior probability of the conceptual model class given the observed travel times (bottom). Successive columns show the posterior mean (top) and travel time misfit (bottom) for the following scenarios: marine parasequence conceptual model, river channel conceptual model, mixed posterior given the posterior class probability, and the posterior mean if the posterior probability for each conceptual model class was Uniform.

are 95%, 61% and 79% for the top, middle and bottom true model in Figure 11 respectively.

5 | DISCUSSION

Our aim in this paper is to develop and demonstrate a methodology to inject high-level (conceptual) and lower-level (parameter) geological prior information into a Bayesian inversion scheme and to investigate their effects. Figures 5 and 6 show that we can successfully inject geological prior information into Bayesian inversion using the GAN. For both the river and the marine prior pdfs we found posterior pdfs that closely match the true parameter matrix. In these cases, the rapid MDN inversions obtain similar posterior estimates to those from McMC in a fraction of the time, post-training. Furthermore, we have shown that the GAN prior parameterisation is reasonably agnostic to the Bayesian inversion scheme used (e.g. McMC or MDN) to estimate parameter pdfs. Thus, if faster or more accurate inversion methods become available then this solution to injecting geological prior information can still be used.

However, the true parameter matrices that we have used in the above examples result in so-called inverse crimes. An inverse crime is a situation where a parameter matrix that has been simplified to a certain degree, is used to generate data, which are then inverted to try to recover a parameter matrix that contains the same simplifications (Kaipio & Somersalo, 2007). Inverse crimes show the best case scenario for any inversion method because the method is not tested against the complexity of true data uncertainties, such tests might overestimate the performance of an algorithm. We therefore also show inversions of the true parameter matrix using the incorrect prior conceptual model. In that scenario, there is no inverse crime as the simplifications made in the true parameter matrix and in the prior information are different. Although these inversions performed worse than their inverse crime equivalents, we argue that the results are potentially still recognisable, particularly when using Monte Carlo methods.

The three inversions on the real basin data do not commit inverse crimes. These inversions show that we are also able to invert for real structures. In these examples, we have not used real measured travel time data because whenever such data are available it is seldom the case that the true model is known so that would not constitute a test of the inversion method. Our examples therefore provide a reasonable test of our inversion scheme in a real-world setting. What is more, our classifier NN correctly predicts the shallow marine setting is more likely to be correct from the arrival time data.

Another benefit of using geological prior pdfs is that the resulting posterior pdf consists only of geologically reasonable parameter matrices. Each sample of the posterior pdf is a geological parameter matrix in itself. Posterior samples from most tomographic methods can contain high-probability parameter matrices that are not geological (Earp & Curtis, 2020); when interpreting such samples, one may opt to alter the distribution to make it more geological which introduces unknown uncertainties. Another option is to interrogate the whole set of parameter samples to answer specific geological questions (Arnold & Curtis, 2018; Zhang &

Curtis, 2022; Zhao et al., 2022). Our method solves this problem directly, as each parameter matrix in the posterior is automatically geological, at least to the extent that the GAN has been trained to represent the prior pdf. High-probability samples can therefore be interpreted directly, and estimates of their relative probabilities are known.

If more prior pdfs become available, for example conceptual models for salt diapirs, carbonates or specific geographic regions, it may become infeasible to invert the data for each of these priors due to the computational expense that comes with inverting the data. As a solution, we introduced a classifier neural network that infers posterior probabilities of which prior pdf is most consistent with the arrival time data. This information may be especially useful in the case when there is a wish to obtain the most accurate possible McMC posterior estimates, but where the posterior parameter estimation is costly to compute, as the arrival times only have to be inverted for the parameters associated with the prior information class found by the classifier NN. This neural network must be retrained for each additional prior pdf but its training is cheap relative to the cost of training the MDNs or computing an McMC posterior estimate. Furthermore, it can be used to combine different prior pdfs to obtain a posterior parameter estimate given different classes of prior information, which could be useful in a scenario where the true parameter matrix is in fact best represented by a mixture between two canonical conceptual models.

Finally, this work highlights the improvements to imaging using MDNs offered by combining them with geological prior information. Earp and Curtis (2020) used MDNs to invert directly for parameters that represent pixels in the tomographic image and obtained marginal posteriors for each parameter individually. By introducing the geological prior pdf, we also obtain uncorrelated marginal pdfs, but on latent parameters of the GANs; varying those latent parameters within their respective posterior marginals translates through the GAN to correlated estimates of the posterior variation of the image parameters. This is demonstrated by the fact that the posterior samples look geological and are thus highly correlated in space. This occurs because approximately correct intra-parameter correlations are stored in the GAN architecture. Ideally, the fully correlated posterior pdf over latent parameters would be estimated directly, and this may be possible in the future using different neural network architectures within the MDN, or different training methods.

6 | CONCLUSION

In this study, we inject geological prior information into a travel time tomography Bayesian inversion scheme to improve the posterior knowledge about parameters that describe a tomographic image. We evaluate two different geological conceptual models: a braided river system and the formation of marine parasequences. Each is parameterised inside a generative adversarial network (GAN) for rapid generation of prior samples. To create a computationally efficient inversion method, we use a mixture density network (MDN) to perform the inversions and use Markov chain Monte Carlo inversion to validate the results. We successfully inject geological prior information using the GAN, and the rapid MDN posterior estimates closely approximate the benchmark McMC posterior estimates.

Furthermore, we are able to analyse the effects of using inappropriate prior information for a given set of travel times (travel times from a braided river system inverted using prior information from marine parasequences and vice versa). We find that the posterior estimates with inappropriate prior information are worse compared to appropriate prior information, as expected. However, the posterior estimates also contain information about the underlying true parameter matrix, so we train a neural network to find the posterior class probability that describes which prior conceptual model information is most consistent with a given set of travel time data. We thus demonstrate that we can rapidly invert tomographic travel times with rich geological prior information and that we are able to discriminate between a set of geological conceptual models to find which is most appropriate for the area under consideration.

ACKNOWLEDGEMENTS

We thank the sponsors of the Edinburgh Imaging Project (blogs.ed.ac.uk/imaging): TotalEnergies and BP for enabling this study. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk/). Furthermore, we would like to thank Niklas Linde for directing us to the river network GAN published by Laloy et al. (2018). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

FUNDING INFORMATION

The Edinburgh Imaging Project is funded by TotalEnergies and BP.

CONFLICT OF INTEREST STATEMENT

The authors declare that there was no conflict of interest while creating this paper.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/bre.12811.

DATA AVAILABILITY STATEMENT

The data used in this paper are created using the SedSimple geological process model and are available at https://wsoftc.com/index.php/software. GAN codes and the trained models will be made available at https://github.com/hugobloem/FSM-prior-information.

ORCID

Hugo Bloem https://orcid.org/0000-0002-4205-0597

REFERENCES

Aki, K., & Lee, W. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first p arrival times from local earthquakes: 1. A homogeneous initial model. *Journal of Geophysical Research*, *81*(23), 4381–4399.

Al-Wazzan, H. A., Hawie, N., Al-Haggan, H., Al-Mershed, M. K., Al-Sahlan, G., & Al-Wadi, M. (2021). 3D forward stratigraphic modelling of the lower jurassic carbonate systems of Kuwait. *Marine and Petroleum Geology*, *123*, 104699.

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks. In *International Conference on Machine Learning* (pp. 214–223). PMLR.

Arnold, D., Demyanov, V., Rojas, T., & Christie, M. (2019). Uncertainty quantification in reservoir prediction: Part 1—Model realism in history matching using geological prior definitions. *Mathematical Geosciences*, *51*(2), 209–240.

Arnold, R., & Curtis, A. (2018). Interrogation theory. *Geophysical Journal International*, *214*(3), 1830–1846.

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. *4*). Springer.

Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, *178*(3), 1411–1436.

Bond, C., Lunn, R., Shipton, Z., & Lunn, A. (2012). What makes an expert effective at interpreting seismic images? *Geology*, *40*(1), 75–78.

Bond, C. E., Johnson, G., & Ellis, J. (2015). Structural model creation: The impact of data type and creative space on geological reasoning and interpretation. *Geological Society, London, Special Publications*, *421*(1), 83–97.

Borgomano, J., Lanteaume, C., Léonide, P., Fournier, F., Montaggioni, L. F., & Masse, J.-P. (2020). Quantitative carbonate sequence stratigraphy: Insights from stratigraphic forward models. *AAPG Bulletin*, *104*(5), 1115–1142.

Brock, A., Donahue, J., & Simonyan, K. (2018). Large scale GAN training for high fidelity natural image synthesis. *arXiv preprint* arXiv:1809.11096.

Burgess, P., Wright, V., & Emery, D. (2001). Numerical forward modelling of peritidal carbonate parasequence development: Implications for outcrop interpretation. *Basin Research*, *13*(1), 1–16.

Burgess, P. M., & Emery, D. J. (2004). Sensitive dependence, divergence and unpredictable behaviour in a stratigraphic forward model of a carbonate system. *Geological Society, London, Special Publications*, *239*(1), 77–94.

Courtade, S., Lejri, M., Dutra, H. P. L., & Tetzlaff, D. (2021). Reproducing sedimentary processes and architecture in turbidite deposits with forward stratigraphic modeling. In *In 82nd EAGE Annual Conference & Exhibition* (Vol. *2021*, pp. 1–5). EAGE Publications BV.

Curtis, A. (2012). The science of subjectivity. *Geology*, *40*(1), 95–96.

Curtis, A., & Lomax, A. (2001). Prior information, sampling distributions, and the curse of dimensionality. *Geophysics*, *66*(2), 372–378.

Curtis, A., & Wood, R. (2004). Optimal elicitation of probabilistic information from experts. *Geological Society, London, Special Publications*, *239*(1), 127–145.

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Multiple Classifier Systems: First International Workshop, MCS 2000 Cagliari, Italy, June 21–23, 2000 Proceedings 1* (pp. 1–15). Springer.

Dziewonski, A. M., & Woodhouse, J. H. (1987). Global images of the earth's interior. *Science*, *236*(4797), 37–48.

Earp, S., & Curtis, A. (2020). Probabilistic neural network-based 2D travel-time tomography. *Neural Computing and Applications*, *32*(22), 17077–17095.

Earp, S., Curtis, A., Zhang, X., & Hansteen, F. (2020). Probabilistic neural network tomography across Grane field (North Sea) from surface wave dispersion data. *Geophysical Journal International*, *223*(3), 1741–1757.

Feng, R., Luthi, S. M., Gisolf, D., & Angerer, E. (2018). Reservoir lithology classification based on seismic inversion results by hidden Markov models: Applying prior geological information. *Marine and Petroleum Geology*, *93*, 218–229.

Galetti, E., Curtis, A., Meles, G. A., & Baptie, B. (2015). Uncertainty loops in travel-time tomography from nonlinear wave physics. *Physical Review Letters*, *114*(14), 148501.

Georgiadis, G., Chiuso, A., & Soatto, S. (2013). Texture compression. In *2013 Data Compression Conference* (pp. 221–230). IEEE.

González, E. F., Mukerji, T., & Mavko, G. (2008). Seismic inversion combining rock physics and multiple-point geostatistics. *Geophysics*, *73*(1), R11–R21.

Goodfellow, I. (2016). NIPS 2016 tutorial: Generative adversarial networks. *arXiv preprint* arXiv:1701.00160.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, *27*. https://doi.org/10.48550/arXiv.1406.2661

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of wasserstein GANs. *arXiv Preprint* arXiv:1704.00028.

Haberland, C., Ryberg, T., Riedel, M., & Bauer, K. (2023). Bayesian seismic travel-time cross-hole tomography in vertically transversely isotropic media. *Journal of Applied Geophysics*, *209*, 104917.

Hamon, Y., Bachaud, P., Granjeon, D., Bemer, E., & Carvalho, A. M. A. (2021). Integration of diagenesis in basin-scale, stratigraphic forward models using reactive transport modeling: Input and scaling issues. *Marine and Petroleum Geology*, *124*, 104832.

Hill, J., Tetzlaff, D., Curtis, A., & Wood, R. (2009). Modeling shallow marine carbonate depositional systems. *Computers & Geosciences*, *35*(9), 1862–1874.

Jetchev, N., Bergmann, U., & Vollgraf, R. (2016). Texture synthesis with spatial generative adversarial networks. *arXiv preprint* arXiv:1611.08207.

Kaipio, J., & Somersalo, E. (2007). Statistical inverse problems: Discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, *198*(2), 493–504.

Kang, M., & Park, J. (2020). ContraGAN: Contrastive learning for conditional image generation. In *Conference on Neural Information Processing Systems (NeurIPS)*. Curran Associates, Inc.

Kang, M., Shim, W., Cho, M., & Park, J. (2021). Rebooting ACGAN: Auxiliary classifier GANs with stable training. In *Conference on Neural Information Processing Systems (NeurIPS)*. Curran Associates, Inc.

Kass, R. E., & Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, *91*(435), 1343–1370.

Laloy, E., Hérault, R., Jacques, D., & Linde, N. (2018). Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research*, *54*(1), 381–406.

Lee, D. S., Stevenson, V. M., Johnston, P. F., & Mullen, C. (1995). Time-lapse crosswell seismic tomography to characterize flow structure in the reservoir during the thermal stimulation. *Geophysics*, *60*(3), 660–666.

Lochbühler, T., Vrugt, J. A., Sadegh, M., & Linde, N. (2015). Summary statistics from training images as prior information in probabilistic inversion. *Geophysical Journal International*, *201*(1), 157–171.

Masiero, I., Burgess, P., Hollis, C., Manifold, L., Gawthorpe, R., Lecomte, I., Marshall, J., & Rotevatn, A. (2021). Syn-rift carbonate platforms in space and time: Testing and refining conceptual models using stratigraphic and seismic numerical forward modelling. *Geological Society, London, Special Publications*, *509*(1), 179–203.

Meier, U., Curtis, A., & Trampert, J. (2007a). A global crustal model constrained by nonlinearised inversion of fundamental mode surface waves. *Geophysical Research Letters*, *34*, L16304.

Meier, U., Curtis, A., & Trampert, J. (2007b). Global crustal thickness from neural network inversion of surface wave data. *Geophysical Journal International*, *169*(2), 706–722.

Meier, U., Trampert, J., & Curtis, A. (2009). Global variations of temperature and water content in the mantle transition zone from higher mode surface waves. *Earth and Planetary Science Letters*, *282*(1–4), 91–101.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092.

Moja, S. S., Asfaw, Z. G., & Omre, H. (2019). Bayesian inversion in hidden Markov models with varying marginal proportions. *Mathematical Geosciences*, *51*(4), 463–484.

Mosegaard, K., & Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research: Solid Earth*, *100*(B7), 12431–12447.

Mosser, L., Dubrule, O., & Blunt, M. J. (2020). Stochastic seismic waveform inversion using generative adversarial networks as a geological prior. *Mathematical Geosciences*, *52*(1), 53–79.

Nawaz, M. A., & Curtis, A. (2016). Bayesian inversion of seismic attributes for geological facies using a hidden Markov model. *Geophysical Journal International*, *208*(2), 1184–1200.

Otoo, D., & Hodgetts, D. (2021). Porosity and permeability prediction through forward stratigraphic simulations using GPM™ and Petrel™: Application in shallow marine depositional settings. *Geoscientific Model Development*, *14*(4), 2075–2095.

Paola, C. (2000). Quantitative models of sedimentary basin filling. *Sedimentology*, *47*, 121–178.

Podvin, P., & Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity models: A massively parallel approach and its associated tools. *Geophysical Journal International*, *105*(1), 271–284.

Polson, D., & Curtis, A. (2010). Dynamics of uncertainty in geological interpretation. *Journal of the Geological Society*, *167*(1), 5–10.

Rawlinson, N., & Sambridge, M. (2004). Multiple reflection and transmission phases in complex layered media using a multistage fast marching method. *Geophysics*, *69*(5), 1338–1350.

Richard, M. D., & Lippmann, R. P. (1991). Neural network classifiers estimate bayesian a posteriori probabilities. *Neural Computation*, *3*(4), 461–483.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training GANs. *Advances in Neural Information Processing Systems*, *29*, 2234–2242.

Snieder, S., Griffiths, C. M., Owen, A., Hartley, A. J., & Howell, J. A. (2021). Stratigraphic forward modelling of distributive fluvial systems based on the Huesca system, Ebro basin, northern Spain. *Basin Research*, *33*(6), 3137–3158.

Song, S., Mukerji, T., & Hou, J. (2021a). Bridging the gap between geophysics and geology with generative adversarial networks. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–11.

Song, S., Mukerji, T., & Hou, J. (2021b). GANSim: Conditional facies simulation using an improved progressive growing of generative adversarial networks (GANs). *Mathematical Geosciences*, *53*(7), 1413–1444.

Sonnenfeld, M. D., & Cross, T. A. (1993). Volumetric partitioning and facies differentiation within the permian Upper San Andres Formation of Last Chance Canyon, Guadalupe Mountains, New Mexico. In *Carbonate sequence stratigraphy: Recent developments and applications*. American Association of Petroleum Geologists.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation* (Vol. 89). SIAM.

Taroudakis, M., & Smaragdakis, C. (2016). *Ocean acoustic tomography using a three-phased probabilistic model-based inversion scheme*. Proceedings of the 22nd International Congress on Acoustics, by MCI Group.

Tetzlaff, D. (2023). Stratigraphic forward modeling software package for research and education. *arXiv preprint* arXiv:2302.05272. https://wsoftc.com

Tetzlaff, D. M., & Harbaugh, J. W. (1989). *Simulating clastic sedimentation*. Van Nostrand Reinhold Co. Inc.

Tsekhmistrenko, M., Sigloch, K., Hosseini, K., & Barruol, G. (2021). A tree of Indo-African mantle plumes imaged by seismic tomography. *Nature Geoscience*, *14*(8), 612–619.

Zhang, X., & Curtis, A. (2020). Seismic tomography using variational inference methods. *Journal of Geophysical Research: Solid Earth*, *125*(4), e2019JB018589.

Zhang, X., & Curtis, A. (2022). Interrogating probabilistic inversion results for subsurface structural information. *Geophysical Journal International*, *229*(2), 750–757.

Zhao, X., Curtis, A., & Zhang, X. (2022). Interrogating subsurface structures using probabilistic tomography: An example

assessing the volume of Irish sea basins. *Journal of Geophysical Research: Solid Earth*, *127*(4), e2022JB024098.

Zhao, Z., Singh, S., Lee, H., Zhang, Z., Odena, A., & Zhang, H. (2020). Improved consistency regularization for GANs. *arXiv preprint* arXiv:2002.04724.

How to cite this article: Bloem, H., Curtis, A., & Tetzlaff, D. (2023). Introducing conceptual geological information into Bayesian tomographic imaging. *Basin Research*, *00*, 1–22. https://doi.org/10.1111/bre.12811

APPENDIX A: LATENT POSTERIOR DISTRIBUTIONS

The results shown in Figures 5 and 6 are statistical summaries of the underlying geological posterior distribution. The geological posterior distribution is sampled in latent space, whereafter the geological images for each set of latent variable values are calculated using the associated GAN and then means and standard deviations across the geological image are calculated on a per-pixel basis. These statistics illustrate the information contained in the geological posterior distributions. However, the statistical summaries are not in themselves geological models. In this appendix, we therefore show some posterior samples that are geological models and the latent parameter posterior distributions that describe the geological posterior distributions. While with the geological posteriors, the inversion performance can be judged by visually comparing the true model to images from the posterior pdf (either geological models or summary statistics), we have no intuition about the performance of inferred latent variable posterior pdf's. Indeed, the latent parameters in themselves do not represent any meaningful properties relating to geology. Only after they have been input to the GAN can their performance be judged.

Figure A1a,b shows the latent posterior distributions corresponding to the geological posterior distributions in Figures 5 and 6, respectively. Note that the river channels GAN uses 9 latent parameters versus 8 parameters for the marine parasequences GAN. The orange line indicates the benchmark McMC latent posterior, the blue line is the MDN latent posterior and the vertical black lines are the true values for each latent parameter. For all parameters, the true value is captured by the high-probability region of the posterior distributions. This indicates that both inversion methods are correctly finding the latent values.

The MDN posterior estimates are fairly close to those found by the benchmark McMC method. In most cases, the majority of high-probability density is located in the same intervals. However, the MDN does have some bias compared to McMC: note the skewed peaks in parameters

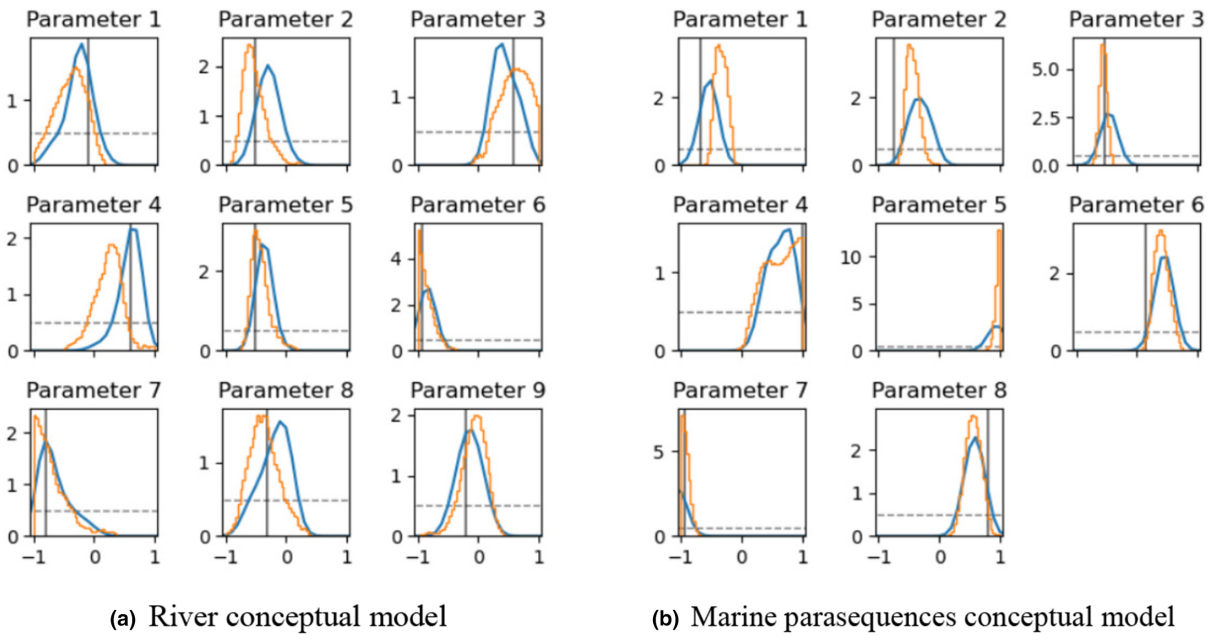
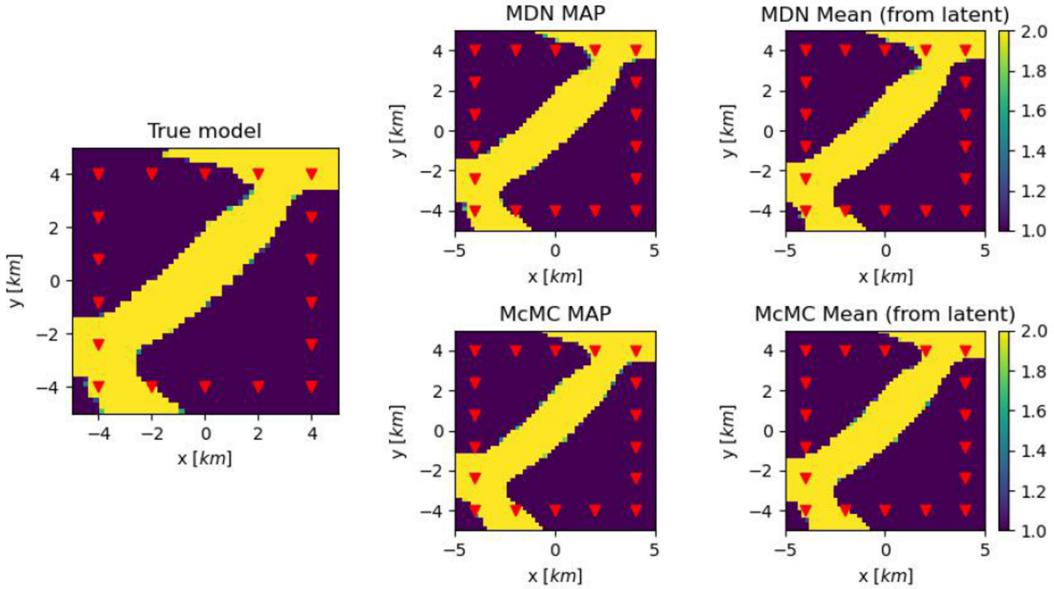


FIGURE A1 Latent posterior distributions for the rivers (a) and marine parasequences (b) conceptual models corresponding to the geological posterior results are shown in Figures 5 and 6, respectively. Orange lines are the benchmark McMC posterior estimates, blue lines indicate the MDN posterior estimates and the black vertical lines show the true value for each latent parameter. The dashed line shows the prior distribution.

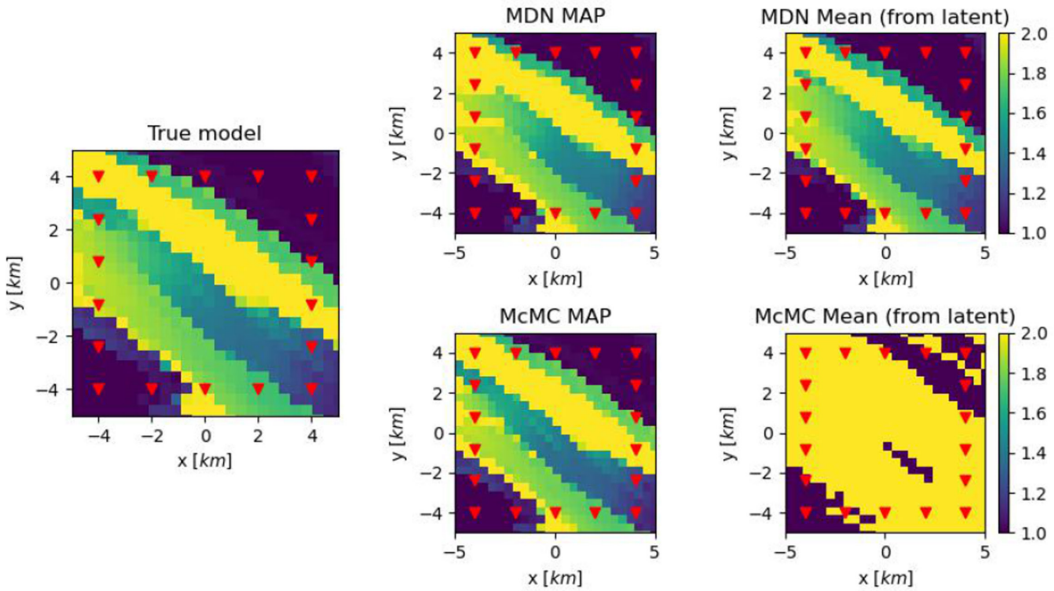
2, 3, 4 and 8 for the river conceptual model and parameter 1 for the marine parasequences. The MDN also fails to capture the multi-modal distribution observed in marine parasequences parameter 4: the MDN distribution is a single-modal with the same width as the multi-modal McMC distribution.

From the posterior distributions in Figure A1a,b, we can calculate the maximum a posteriori (MAP) and mean model in latent space, and by using the GAN, we can translate those parameter values to associated geological models. These are shown in Figure A2a,b. Note that these are intrinsically different from the results in Figures 5 and 6, which show the per-pixel averages of many geological models in the posterior distribution. Hence, those resulting images are not geological but rather summarise the geological distribution.

In Figure A2a,b, the MAPs and mean models are very close to the true models for both conceptual models. What



(a) River conceptual model



(b) Marine parasequences conceptual model

FIGURE A2 For each conceptual model, braided river systems (a) and marine parasequences (b), we show geological models created from latent posterior statistics. From left to right, the following models are shown: true model, maximum a posteriori (MAP) model and mean model (calculated in latent space, then mapped into geological space). The top row in each of the panels (a, b) shows those models from the MDN posterior estimate, and the bottom row shows the models from the McMC posterior estimate.

is more, the McMC and MDN posterior estimates produce similar MAP models and mean models in the river conceptual model. The outlier is the McMC mean model for the marine parasequences model. This model shares the approximate structure with the true model (the diagonal pattern), but the velocities are dissimilar. Rather than identifying a fault in the inversion, we attribute this notable result to the fact that the mean of any distribution need not coincide with a high-probability model. This property is often stated as a caveat to using mean values from McMC results as models, without providing an explicit example, but here is one.