Combining autoencoder neural network and Bayesian inversion algorithms to estimate heterogeneous fracture permeability in enhanced geothermal reservoirs

Zhenjiao Jiang[1], Siyu Zhang[1], Chris Turnadge[2], and Tianfu Xu[1]

[1]Jilin University
[2]Commonwealth Scientific and Industrial Research Organisation (CSIRO)

November 26, 2022

Abstract

While hydraulic fracturing is widely used to enhance the permeability of deep geothermal, gas and oil reservoirs, it remains challenging to infer the heterogeneous distribution of permeability in the fractured zone. Typically, a limited number of boreholes are available at which reservoir imaging and tracer testing can be conducted. The number of observations is often far fewer than the number of estimable permeabilities, making model inversion ill-posed. To overcome this problem, this study combined the autoencoder neural network (a deep learning approach) with Bayesian inversion algorithm (using Markov Chain Monte Carlo, MCMC sampling) to estimate permeability in the enhanced geothermal reservoir, based on a single-well-injection-withdrawal test (SWIW). The autoencoder neural network was used to reduce parameter dimensionality into low-dimension codes by four orders of magnitude, while MCMC sampling was used to update the low-dimension codes according to the SWIW observations. The spatial distribution of permeability was then reconstructed from these low-dimension codes using the original autoencoder neural network. Application of the approach to a synthetic enhanced geothermal system demonstrated that the methodology achieved rapid stabilization of the Bayesian inversion. When the root mean square error (RMSE) between modelled and observed borehole temperature and flow rate values was less than unity, estimated permeability values were comparable to the synthetic reference case, with a mean square error lower than 0.001 mD. The combination of the deep-learning based dimension reduction technique and Bayesian inversion algorithm allow the estimate of permeability distribution in deep artificial reservoirs based on limited number of boreholes.

Combining autoencoder neural network and Bayesian inversion algorithms to estimate heterogeneous fracture permeability in enhanced geothermal reservoirs

Zhenjiao Jiang [1,2*], Siyu Zhang [1], Chris Turnadge [2], Tianfu Xu [1],

1 Key Laboratory of Groundwater Resources and Environment, Ministry of Education, College of Environment and Resources, Jilin University, Changchun, 130021, China

2 CSIRO Land & Water, Locked Bag 2, Glen Osmond, SA 5064, Australia

Corresponding author: Zhenjiao Jiang (jiangzhenjiao@hotmail.com)

Key Points

- Convolutional neural network was developed for 3D hydraulic parameters dimensionality reduction
- Neural network released the prior Gaussian assumption and shortened the burn-in period of Bayesian inversion
- Joint neural network and Bayesian inversion enabled permeability estimation in deep reservoirs based on single-well test

Abstract

While hydraulic fracturing is widely used to enhance the permeability of deep geothermal, gas and oil reservoirs, it remains challenging to infer the heterogeneous distribution of permeability in the fractured zone. Typically, a limited number of boreholes are available at which reservoir imaging and tracer testing can be conducted. The number of observations is often far fewer than the number of estimable permeabilities, making model inversion ill-posed. To overcome this problem, this study combined the autoencoder neural network (a deep learning approach) with Bayesian inversion algorithm (using Markov Chain Monte Carlo, MCMC sampling) to estimate permeability in the enhanced geothermal reservoir, based on a single-well-injection-withdrawal test (SWIW). The autoencoder neural network was used to reduce parameter dimensionality into low-dimension codes by four orders of magnitude, while MCMC sampling was used to update the low-dimension codes according to the SWIW observations. The spatial distribution of permeability was then reconstructed from these low-dimension codes using the original autoencoder neural network. Application of the approach to a synthetic enhanced geothermal system demonstrated that the methodology achieved rapid stabilization of the Bayesian inversion. When the root mean square error (RMSE) between modelled and observed borehole temperature and flow rate values was less than unity, estimated permeability values were comparable to the synthetic reference case, with a mean square error lower than 0.001 mD. The combination of the deep-learning based dimension reduction technique and Bayesian inversion algorithm allow the estimate of permeability distribution in deep artificial reservoirs based on limited number of boreholes.

Key words

Markov chain Monte Carlo inversion; Deep learning; Tracer test; Enhanced geothermal system; Permeability

Plain Language Summary

Hydraulic fracturing is widely used to enhance permeability in deep oil, shale and geothermal reservoirs. However, the number of boreholes for hydraulic tests is often limited, which make the permeability inversion in deep reservoirs ill-posed. Deep learning approach is employed in this study to reduce the parameters dimensionality, prior to the Bayesian inversion, which can balance the number of known and unknown and shorten the inversion period toward the optimum solutions. The joint application of deep learning and Bayesian inversion is tested to be an effective way to infer heterogeneous distribution of permeability in fractured zone in deep reservoirs based on single-well test.

1 Introduction

Hydraulic fracturing is a widely-used technique for enhancing the permeability of deep gas, oil and geothermal reservoirs [e.g. *Hubbert and Willis*, 1972; *Legarth et al.*, 2005; *Sovacool*, 2014]. Characterization of the permeability of these fractured reservoirs is critical when locating production bores to maximize energy production. Currently, a limited number of methods are available for imaging fractures in deep reservoirs. These include downhole imaging [*Prensky*, 1999; *Vidal et al.*, 2017], which can image near-well fractures. Surface or downhole geophysical prospecting, such as micro-seismic monitoring [*Majer et al.*, 2007; *Maxwell et al.*, 2010], can monitor fracturing processes and estimate the fractured volume. However, neither downhole nor surface geophysical prospecting can determine the spatial distribution of internal fracture structures, e.g. fracture aperture or equilibrium permeability. Characterization of these features is necessary for the prediction of the mass and energy transport in reservoirs.

Tracer test can help inferring the internal fracture properties by monitoring surface tracer responses (e.g. temperature, fluid flux, chemical tracer concentrations) during cycling of featured fluid in a reservoir [*LeBlanc et al.*, 1991; *Sanjuan et al.*, 2006]. It has been used in the inverse estimation of inter-bore connectivity, water-rock interaction area, and permeability and dispersivity values [*Garabedian et al.*, 1991; *Maloszewski et al.*, 1999]. Most applications assume homogeneous reservoir properties when simulating mass and heat transport in the subsurface using analytical solutions [*Tsang*, 1995]. Conversely, one advanced approach to the estimation of heterogeneous reservoir properties is hydraulic tomography [*Lee and Kitanidis*, 2014; *T C J Yeh and Liu*, 2000]. This methodology, however, requires extensive cross-bore tracer tests in order to obtain a sufficient number of observations to undertake well-posed model inversion. Due to the high cost of drilling in deep reservoirs, the number of deep bores (and thus tracer tests) is typically limited, making the inference of heterogeneous reservoir parameters highly ill-posed [*McLaughlin and Townley*, 1996]. Alternatively, Bayesian inversion approaches can be used to estimate the summary statistics (e.g. mean and standard deviation) of reservoir parameters based on tracer observations [*Vogt et al.*, 2012]. However, estimated parameters may be subject to considerable uncertainty, as the ill-posedness of the problem permits non-unique solutions.

For solving the ill-posed problems, it is necessary to reduce parameter dimensionality, and thereby constrain uncertainty in the inversed parameters [*Asher et al.*, 2015; *W W G Yeh*, 1986]. Widely-employed dimensionality reduction methods include principal component analysis [*Ding and He*, 2004; *Kambhatla and Leen*, 1997], discrete wavelet transforms [*Bruce et al.*, 2002], and singular value decomposition [*Doherty and Hunt*, 2010; *Wall et al.*, 2003]. Most of these methods assume that the model parameters are normally distributed. To avoid this limitation, Laloy et al. [2018; 2017] introduced two approaches using deep convolutional neural networks: variational autoencoder (VAE) and generative adversarial network (GAN), for dimensionality reduction of non-Gaussian parameters in channelized aquifer by a factor of 500.

The deep-learning neural network above makes it possible to learn a customized low-dimension latent space on which all prior models lie, that is, the prior model is to be learned from complex geological realizations that cannot be described mathematically in the original parameter space by, for instance, a multivariate Gaussian model. When working with Bayesian framework, deep learning opens a door for faster and robust inversions of complex geological structures, for instance, continuous fracture permeability distribution in the reservoir. In this study, we establish a 3D autoencoder neural network, to express original high-dimension permeabilities into low-dimension codes; the correlations between codes and original permeability space in both horizontal and vertical direction can be addressed. Bayesian inversion is then used to estimate these codes based on single-well injection-withdrawal test, prior to the estimation of high-dimension permeability directly. The method is tested on a synthetic fractured reservoir, which provides a new tool to estimate uneven distribution of fracture permeability in the reservoir based on the cost-flexible single-well test.

2 Methods

The methodology in this study is composed of three key parts: First, a deep learning approach featuring a 3D autoencoder neural network is used for the parameter dimensionality reduction and full-dimension parameter reconstruction; Second, forward fluid and heat transport model is used to relate full-dimension parameters to observations in tracer test (here the borehole temperature and fluid flux in a single-well injection-withdrawal test, SWIW); Third, a Bayesian inversion algorithm is used to update the low-dimension 'codes' in the autoencoder neural network by minimizing the root mean square error between modelled and observed data in tracer test (Fig. 1).

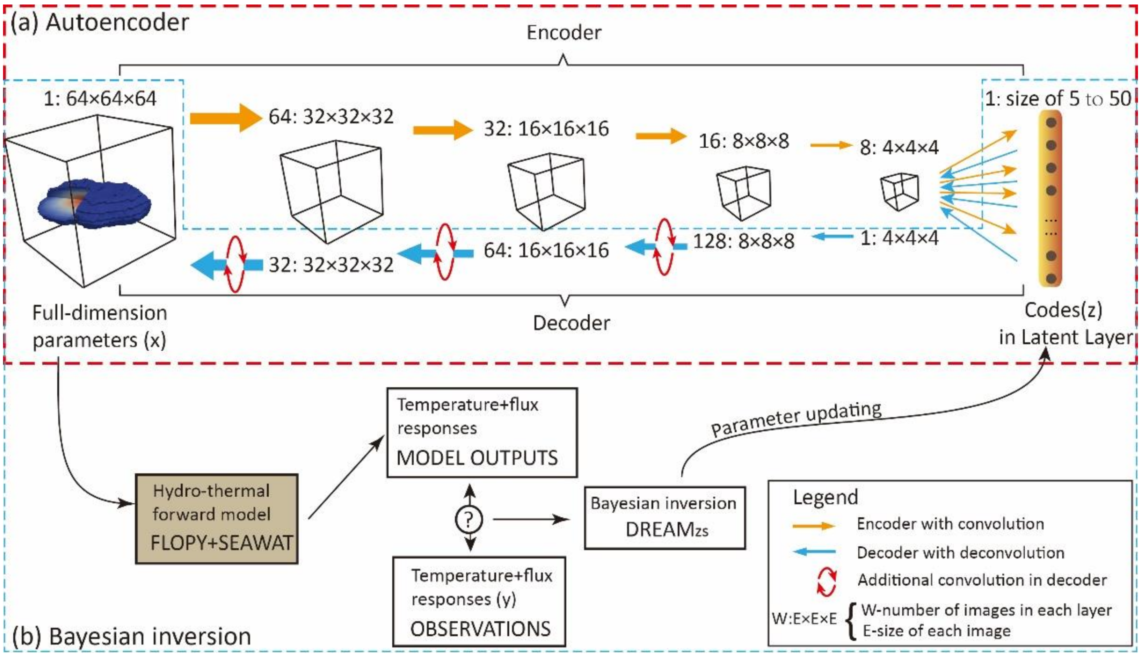

Figure 1. (a) 3D autoencoder neural network composed of (1) "encoder", featuring full-dimension parameters as input and low-dimension 'codes' as output, and (2) "decoder", reconstructing the full-dimension parameters from low-dimension codes; (b) Bayesian inversion model updating low-dimension codes according to the root mean square error between modelled and observed data (borehole temperature and flow fluxes) in tracer testing.

2.1 Autoencoder neural network

The autoencoder is a neural network that can be trained to compress high-dimension input parameters (here, a 3D spatial distribution of fracture permeability) into the low-dimension representative parameters (which are hereafter referred to as 'codes'), and then uncompress the codes into original high-dimension parameters (Fig. 1a). The neural network connecting the input high-dimension parameters to the low-dimension code is called ‘encoder’, while the one converting the code to high-dimension parameters is called ‘decoder’. The central layer containing the code is referred to as ‘latent layer’.

The encoder is generally composed of two to four convolution layers, in which each layer converts the high-dimension input **x** into the low-dimension output **h** via [*Hinton and Salakhutdinov*, 2006]:

$$\mathbf{h}(\mathbf{x}) = f(\mathbf{w} \cdot \mathbf{x} + \boldsymbol{b}) \tag{1}$$

where f is a nonlinear function that is referred to as an “activation function”, **w** is a 3D weights matrix (filter) and $\boldsymbol{b}$ is a bias vector; both follow a uniform distribution ranging from zero to unity. The output images of each convolutional layer provide the input for the next layer, and each layer is allowed to contain multiple images (Fig. 1a). For example, when the input to the encoder is a 3D image, a 3D weight filter of size $N_i \times N_j \times N_k$ is used to calculate the p-th element in the output layer by [*Laloy et al.*, 2017]:

$$h^p_{m,n,o}(\mathbf{x}) = f\left(\sum_{i=1,j=1,k=1}^{N_i,N_j,N_k} w^p_{i,j,k}\, x_{m+i,n+j,o+k} + b^p\right), \tag{2}$$

where m, n, o indicate the voxel position in the 3D image.

The encoder starts from full-dimension parameters and output the low-dimension codes. Conversely, the decoder is composed of transposed convolution layers converting the low-dimension codes to original full-dimension parameters. To enhance the mixing of the data flow via the decoder network, we added additional convolution layers after each transposed convolution layer in the decoder (Fig. 1a).

Weight matrixes and the bias vectors in the encoder are updated to ensure the generated low-dimension codes following the standard normal distribution, by minimizing the Kullback–Leibler divergence (i.e. encoder loss function, $L1$) [*Kullback and Leibler*, 1951]:

$$L1 = \frac{1}{2}\left[\sum_{i=1}^{d}(\mu_i^2 + \sigma_i^2 - \log\sigma_i^2)\right] - \frac{d}{2}, \tag{3}$$

where d is the number of training images used to train the neural network in each training, and μ and σ are the mean and standard deviation of the codes, respectively.

The weights and biases used in the decoder are optimized to minimize the discrepancy between the original full-dimension parameters and those reconstructed by decoder; i.e. the decoder loss function ($L2$). This is defined as:

$$L2 = \sum_{i=1}^{d} \| g(\boldsymbol{z}; \mathbf{w}, \boldsymbol{b}) - \mathbf{x} \|^2, \tag{4}$$

where **x** is the original full-dimension parameters, $\boldsymbol{z}$ is the code, and $g(\cdot)$ represents the calculations in the decoder. Both weight and bias values used in the encoder and decoder are optimized using a stochastic gradient descent algorithm, i.e. the adaptive moment estimation algorithm [*Kingma and Ba*, 2014]. The autoencoder neural network above are implemented using the Tensorflow Python library [*Abadi et al.*, 2016].

2.2 Bayesian inversion

Using a Bayesian inversion approach, a posterior distribution of parameters is proportional to the product of a prior distribution and a likelihood function as [*Vrugt*, 2016]:

$$p(\boldsymbol{z}|\boldsymbol{y}) \propto p(\boldsymbol{z})L(\boldsymbol{z}|\boldsymbol{y}). \tag{5}$$

In this study, $\boldsymbol{y}$ is the vector of tracer test observations (e.g. borehole temperature and fluid flux), $p(\boldsymbol{z})$ is the prior distribution of the code $\boldsymbol{z}$ in the autoencoder (following a standard normal distribution), and $L(\boldsymbol{z}|\boldsymbol{y})$ is the likelihood function, which is assumed to follow a log-normal distribution, i.e.:

$$L(\boldsymbol{z}|\boldsymbol{y}) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - F_i(\mathbf{z})}{\sigma}\right)^2, \tag{6}$$

where n is the number of tracer test observations, σ is the standard deviation of the measurement error of the observations, $F(\mathbf{z})$ are the state variables (e.g. temperature and fluid flux) simulated via forward modelling.

A Markov chain Monte Carlo algorithm implemented by the software $\text{DREAM}_{(zs)}$ rewritten in Python by Laloy et al. [2017] (https://github.com/elaloy/VAE_MCMC) is used to infer low-dimension codes in the autoencoder, which is then converted to the full-dimension parameters via the decoder and is used in the forward model to calculate the tracer test observations. Specifically, a random walk Metropolis-Hastings algorithm is used to draw and accept or reject random samples from prior distributions [*Chib and Greenberg*, 1995; *Hastings*, 1970; *Metropolis et al.*, 1953].

2.3 Forward model

Heat and fluid transport in the subsurface are simulated using SEAWAT [*Langevin et al.*, 2008]. The governing equation for the density-dependent fluid flow is:

$$\rho S\frac{\partial h}{\partial t} + \theta\frac{\partial\rho}{\partial t} = \nabla\left[\frac{\rho g k}{\mu}\rho_0\left(\nabla h + \frac{\rho-\rho_0}{\rho_0}e\right)\right], \tag{7}$$

where ρ is fluid density [kg/m^3], S is specific storage [m^{-1}], h is the hydraulic head [m] of fluid at a reference temperature of 25 ^{o}C, t is time [s], θ is porosity [unitless], k is permeability [m^2], μ is dynamic viscosity [kg/m.s], g is gravitational acceleration [m/s^2], ρ_0 is fluid density [kg/m^3] at a reference temperature of 25 ^{o}C, and e is equal to unity in the vertical direction and zero in other directions [m].

Fluid density and viscosity are both functions of temperature; i.e.:

$$\rho = \rho + \rho_0\beta(T - T_0), \tag{8}$$

$$\mu = a \cdot 10^{\left(\frac{b}{T+c}\right)}, \tag{9}$$

where β is thermal expansion coefficient [-3.75×10^{-4} 1/^{o}C], T_0 is the reference temperature, and a, b and c are unitless empirical coefficients, equal to 239.4×10^{-7}, 248.37, and 133.15, respectively [*Voss*, 1984]

Heat transport is simulated by solving the equation:

$$\left(\theta\rho c_f + (1-\theta)\rho_s c_s\right)\frac{\partial T}{\partial t} = \nabla\left[\left(\lambda + \alpha v\rho c_f\right)\nabla T\right] - \rho c_f\nabla(vT), \tag{10}$$

where c_f and c_s are the specific heat capacity[J/kg.^{o}C] of fluids and solids, respectively, ρ_s is rock density [kg/m^3], T is temperature [^{o}C], α is thermal dispersivity [m], v is the fluid velocity [m/s], and λ is the bulk thermal conductivity [W/m.^{o}C].

2.4 Performance metrics of autoencoder and Bayesian inversion

The performance of the autoencoder in reconstructing full-dimension parameters from low-dimension codes is assessed using three metrics: peak signal to noise ratio, which is a voxel-wise independent criterion defining the accuracy of parameter estimation on each voxel; structure similarity index metric, which expresses the similarity between spatial correlations of calculated and real parameters; and the coefficient of variation, which defines the sensitivity of output parameters to the input variables.

The peak signal to noise ratio (PSNR) is given as [*Wang and Bovik*, 2002]:

$$PSNR = -10\log_{10}\left[\frac{1}{N}\sum_{i=1}^{N}(\tilde{x}_i - x_i)^2\right], \tag{11}$$

where x represents the real full-dimension parameters, $\tilde{x}$ are the full-dimension parameters estimated using decoder, N is the total number of parameters in space, i is the index of the voxel. The term enclosed in square brackets is the mean square error and N is the number of full-dimension parameters. A high PSNR value represents the high-quality of full-dimension parameters being reconstructed.

The structure similarity index metric (SSIM) is given as [*Wang et al.*, 2004]:

$$SSIM = \frac{2\mu_x\mu_{\tilde{x}}+\varepsilon}{\mu_x^2+\mu_{\tilde{x}}^2+\varepsilon}\cdot\frac{2cov(x,\ \tilde{x})+\varepsilon}{\sigma_x^2+\sigma_{\tilde{x}}^2+\varepsilon}, \tag{12}$$

where μ_x and σ_x^2 are the mean and variance of real full-dimension parameters, $\mu_{\tilde{x}}$ and $\sigma_{\tilde{x}}^2$ are the mean and variance of full-dimension parameters estimated by decoder, cov is the covariance of the original or reconstructed parameters, and ε is a small number (10^{-6}) to avoid zero in denominator. SSIM ranges from zero to unity, with higher values indicating a better reconstruction.

The coefficient of variation (CV) is given as:

$$CV = \frac{\sigma_{\tilde{x}}}{\mu_{\tilde{x}}}, \tag{13}$$

where $\sigma_{\tilde{x}}$ and $\mu_{\tilde{x}}$ are the standard deviation and mean, respectively, of full-dimension parameters estimated by decoder at each spatial position under varying sets of low-dimension codes. Larger CV (often >0.5) indicates the greater sensitivity of full-dimension parameters to low-dimension codes.

Moreover, the MCMC inversion is monitored by the misfit between modelled and observed tracer test observations (here borehole temperature and fluid flux), expressed by the root mean square errors metric:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\|y_i - F_i(\mathbf{z})\|^2}. \tag{14}$$

where y and $F(\mathbf{z})$ are the observed and modelled tracer test observations.

3 Application

Although several enhanced geothermal systems were constructed in the world, including such as France Soultz-Sous-Forets, Switzerland Basel, US Desert Peak, Australia Habanero, the heterogeneous distribution of fractured permeability in these geothermal reservoirs is still not fully determined. Micro-seismic monitoring during hydraulic fracturing indicates that fracture patterns in existing enhanced geothermal reservoirs typically feature an ellipsoid-like shape [e.g. *Calò and Dorbath*, 2013; *Cuenot et al.*, 2008; *Maurer et al.*, 2015], centred on the fracturing well, with major axes oriented in the direction of maximum stress, and minor axes oriented in the direction of minimum stress. These ellipsoid fracture zones can be skewed where mechanical properties of the rock are heterogeneous. Following these basic patterns, 3D training images of fracture probability (which represents normalized equilibrium porosity and can be converted to equilibrium permeability) are generated. The generated fracture probability ranges from unity near the fracturing bore to zero at a given distance. Further details are given in the Appendix. Consequently, 6000 images (forming training dataset) of fracture probability are used to train weights and biases in the autoencoder neural network. Another 100 images (forming validation dataset) independent to the training images are used to assess the efficacy of the autoencoder in parameter dimensionality reduction and reconstruction.

For verification, the methodology in Section 2 is implemented to a synthetic case, with an ellipsoid-like fractured zone and a single vertical bore in geothermal reservoir. First, a single-well injection-withdrawal (SWIW) test is simulated in order to generate synthetic observations; Second, the autoencoder neural network is trained to assure the efficacy in conversion between low-dimension codes and full-dimension parameters; Third, low-dimension codes are estimated using a Bayesian inversion algorithm according to the SWIW test observations, which are then converted to the permeability by autoencoder to compare with the synthetic permeability distribution.

3.1 Conceptual model

The synthetic case study features an ellipsoid-like fractured zone created by hydraulic fracturing via single vertical well in a geothermal reservoir (Fig. 2a). Equilibrium permeability of the fractured zone decreases outward from the fracturing well, ranging from 0.01 to 10 mD (i.e. 10^{-17} to 10^{-14} m^2) in the fracture zone, and up to 0.01 mD outside the fracture zone (detail in Appendix). The other parameters affecting the heat and fluid transport are listed in Table 1 and fixed in this study [*Langevin*, 2009].

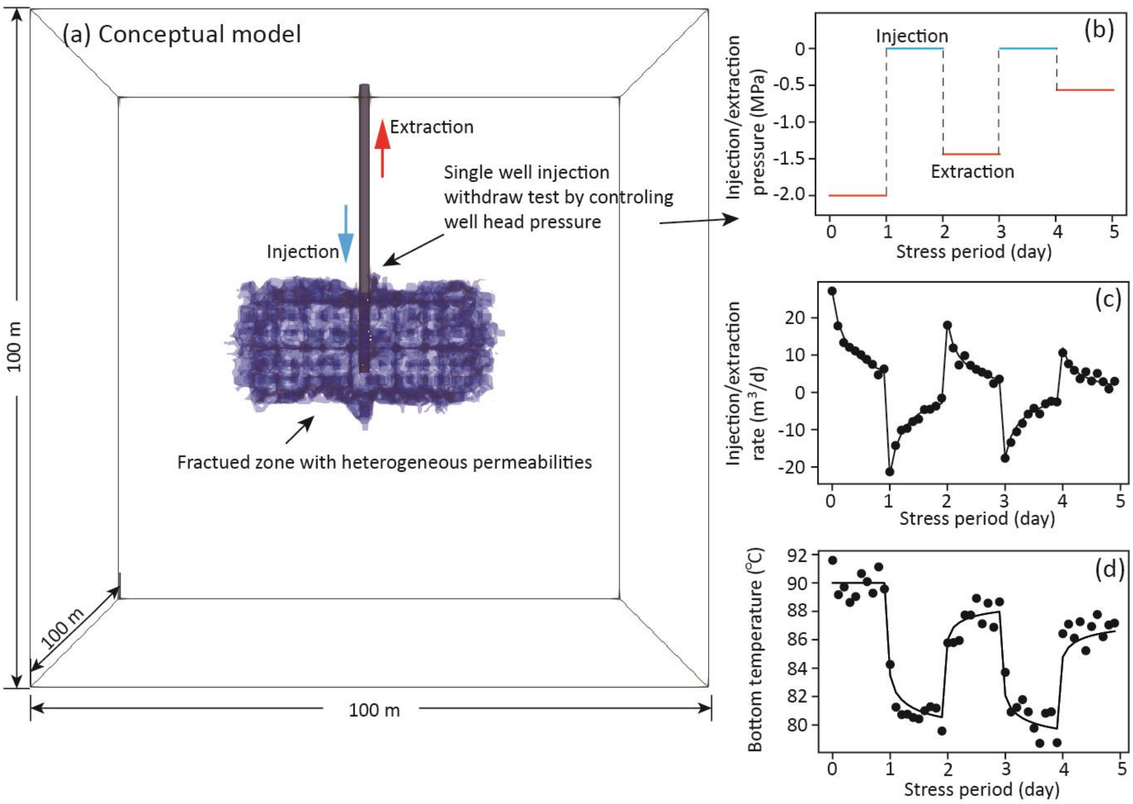

Figure 2. (a) Conceptual model of a single-well injection-withdrawal test in a fractured zone, under constant-pressure operations on the well head in five stress periods (b), leading to transient extraction/injection rates (c) and bottom borehole temperatures (d), artificially adding white noise with standard deviation of 1.0 to mimic the measurement errors.

To estimate the spatial distribution of reservoir permeability, a single-well injection-withdrawal (SWIW) test can be conducted by manipulating the well-head pressure. An initial hydraulic head in the reservoir equal to 3000 m (representing a fracture zone at depth of 3000 m) is specified. A wellhead pressure of -2 MPa (corresponding to a hydraulic head of 2800 m) is assigned on the first day, followed by zero MPa (3000 m), -1.5 MPa (2850 m), zero MPa

(3000 m) and -0.5 MPa (2950 m) on subsequent days (Fig. 2b). A model domain of 100 × 100 × 100 m^3 is discretised into 64 × 64 × 64 pixels (i.e. a total of 262,144 pixels). Zero-flux Neumann boundary conditions for both fluid and heat are specified on all model boundaries. Modelling testing demonstrated that these boundary conditions did not affect borehole temperature and fluid fluxes during five-day periodic injection and extraction scheme. A spatially uniform initial temperature of 90 ^oC is specified across the model domain at a scale of 100 m without considering the vertical temperature changes under natural geothermal gradient.

Table 1. Hydraulic and thermal parameters used in the fluid and heat transport modelling [*Langevin*, 2009]

Parameters	Values	Parameters	Values
Permeability [mD]	<10	Heat conductivity of water [W/m.^{o}C]	0.58
Effective porosity [unitless]	<0.1	Thermal dispersivity [m]	50
Heat capacity of granite [J/kg.^{o}C]	790	Reference temperature [^{o}C]	25
Heat capacity of water [J/kg.^{o}C]	4.2	Water density at reference temperature [kg/m^3]	1000
Heat conductivity of granite [W/m.^{o}C]	3.59	Granite density at reference temperature [kg/m^3]	2750

Simulated borehole temperatures and fluid fluxes are presented in Fig. 2c and 2d. Temperature and fluid flux observations are recorded at a frequency of 10 cycles per day over the total simulated period of five days, resulting in 50 observations of either type. These observations represent a tracer test observation dataset. Gaussian white noises with standard deviations of 1.0 ^oC or 1.0 m^3/d are added to the observations of temperature and fluid flux, respectively, to represent measurement errors. Since the value of outflow flux and temperature are on the same order of magnitude, they are mixed in RMSE calculation in Eq. (14) to monitoring inversion processes.

In summary, the model domain features over 200,000 parameters, to be estimated based on 100 observations in tracer test. Autoencoder is used to reduce the parameter dimensionality.

3.2 Autoencoder optimization

A classical autoencoder neural network [*Hinton and Salakhutdinov*, 2006] consisting of four layers in both the encoder and decoder is used in this study. To enhance the mixing of data flow (data transferring among multiple layers in the neural network) in the decoder, we add one convolution layers after each transposed convolution layers in the decoder. The number and size of the images contained in each layer is illustrated in Fig. 1a. Since the number of weights is sufficient to recreate the original parameters after parameter dimension reduction and reconstruction, we do not investigate the influence of network structure on the accuracy of the result. Details of network structure selection can be found in e.g. Liu et al. [2017] and Goodfellow et al. [2016]. Alternatively, this study focuses on minimizing the number of codes in the latent layer that is able to fully represent the full-dimension parameters. This can help release the burden of Bayesian inversion and constrain the inversion uncertainties.

As shown in Fig. 3, encoder loss function (Eq. 3) finalizes at a value < 0.01 after 6000 iterations, while the value of decoder loss function (Eq. 4) finalizes at < 1000/262144 (0.004) after 10,000 iterations. The former indicates that the resulting codes from the encoder follow the standard normal distribution, while the latter indicates the calculated 3D fracture probability via the decoder aligning with original full-dimension parameters. PSNR and SSIM metric values exceed 20 and 0.9, respectively, after 10,000 iterations (Fig. 3c and 3d). Moreover, the PSNR and SSIM calculated on the validation dataset fit with those resulting from the training dataset, meaning that (1) the trained autoencoder is applicable in the dataset independent to the training dataset without overfitting problem, and (2) the trained autoencoder is capable to reconstruct 3D images outside the training image pool. This is important, as it indicates that the autoencoder neural network, learned from 6000 training images, is applicable to fracture patterns beyond the training dataset, since the validation images are generated independent to the training images, by varying number and ranges of free variables (Appendix).

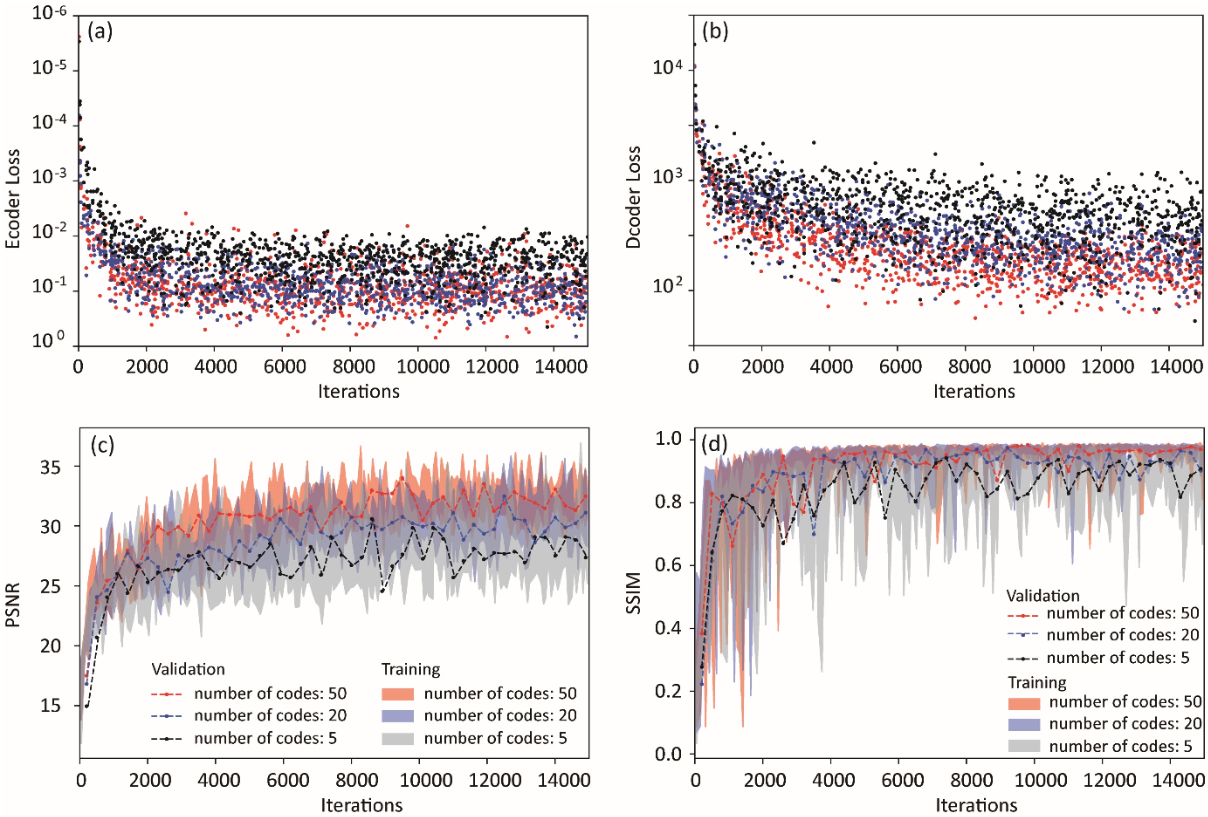

Figure 3. Monitoring loss functions of encoder (Eq. 3) (a) and decoder (Eq. 4) (b); and (c) PSNR and (d) SSIM between the original and reconstructed fracture probability, with the number of codes in the latent layer of 5, 20, 50, respectively.

Both finalized PNSR and SSIM values increase with the number of codes, and tend to be less variable, suggesting that the performance of the autoencoder in reconstructing the full-dimension parameters can be enhanced by increasing the number of codes (Fig. 3c and 3d). However, the lager number of codes can increase the computational burden and uncertainty of Bayesian inversion. Thus, an optimized number of codes is desirable. To this end, the finalized PNSR and SSIM for the validation images are then calculated under varying numbers of codes by five parallel computations. In each computation, the PNSR and SSIM increase with the number of codes, but are highly variable (Fig. 4). This is due to the applied method of stochastic optimization for autoencoder weights, which do not consistently identify the global minimum encoder and decoder losses (Eqs. 3 and 4, respectively). Based on five parallel computations, it is found that the PNSR and SSIM can reach 30 and 0.95, respectively, at minimum number of codes of 20, and the reconstructed 3D fracture probability compares well with the original values, with a mean square error < 0.001. Further increasing the number of codes does not improve the accuracy of parameter reconstruction significantly.

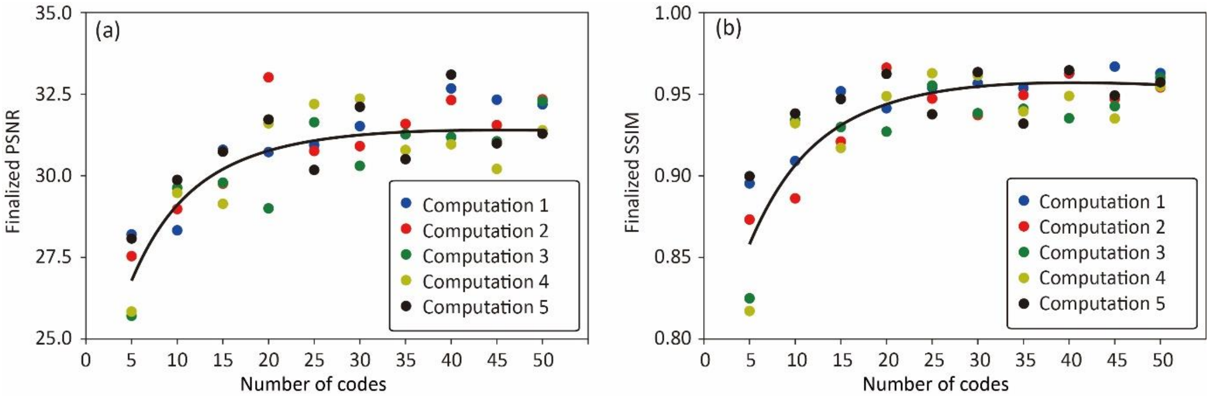

Figure 4. The average PNSR (a) and SSIM (b) for 100 validation images under the number of codes in the latent layer varying from 5 to 50, based on five parallel computations.

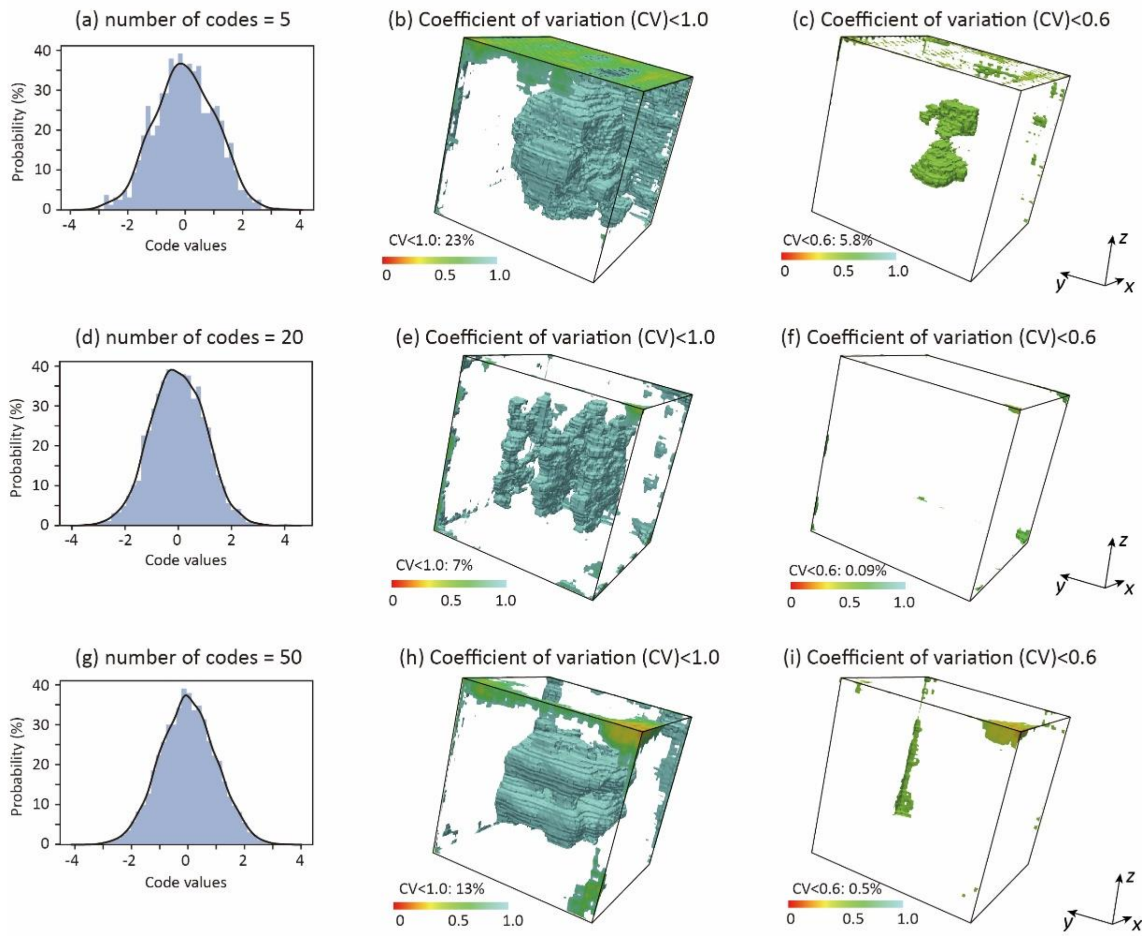

Figure 5. Coefficient of variation values of the reconstructed full-dimension parameters (fracture probability) by the decoder with the number of codes in the latent layer of (a)-(c) 5, (d)-(f) 20, and (g)-(i) 50, respectively.

Furthermore, to ensure that the full-dimension parameters are sensitive to codes in the latent layer (rather than controlled merely by the weights in the autoencoder neural network), a sensitivity analysis is conducted with 5, 20, 50 codes in the latent layer. For each number of codes, 1000 sets of codes are generated from a standard normal distribution (Fig. 5a, 5d and 5g) which result in 1000 sets of full-dimension parameters via the decoder. Coefficient of variation (CV) of full-dimension parameters are subsequently calculated at each voxel (Eq. 13). When five codes are used in the latent layer, CV values are less than 1.0 for 20 % of voxels (Fig. 6b) and are less than 0.5 for 5% of voxels (Fig. 5c). When 20 codes are used, only 7 % of voxels feature CV values less than 1.0 (Fig. 5e) and CV are greater than 0.5 at all voxels (Fig. 5f). This indicates that the sensitivity of output full-dimension parameters to the input codes in the decoder can be enhanced by increasing the number of codes in the latent layer from 5 to 20. Further increasing the number of codes from 20 to 50 do not enlarge the CV values (Fig. 5h and 5i).

3.3 Bayesian inversion

The decoder is then combined with the Bayesian inversion, with 20 codes in the latent layer to update according to the discrepancies between calculated borehole temperatures and fluid flux with the observations (Fig. 2c and 2d). Markov Chain Monte Carlo (MCMC) inversion approaches with five chains are employed. During the inversion processes, the objective function expressed by RMSE (Eq. 14) between observed and calculated tracer test observations decreases rapidly along each chain during the inversion (Fig. 6a). After a burn-in period of 800 iterations, objective function values decrease to about 1.0, which is equilibrium to the measured errors of the tracer observations. Spatial distributions of permeability generated by the decoder are examined after every 100 iterations of the Bayesian inversion algorithm (Fig. 6). Prior to inversion, high fracture permeabilities (with values >0.01 mD) are sparsely distributed throughout the model domain (Fig. 6b). After 200 iterations of the inversion algorithm, high permeability values are clustered at the centre of the domain (Fig. 6c). After 800 iterations, the extent of the high-permeability zone is further constrained at the centre of the model domain (Fig. 6d).

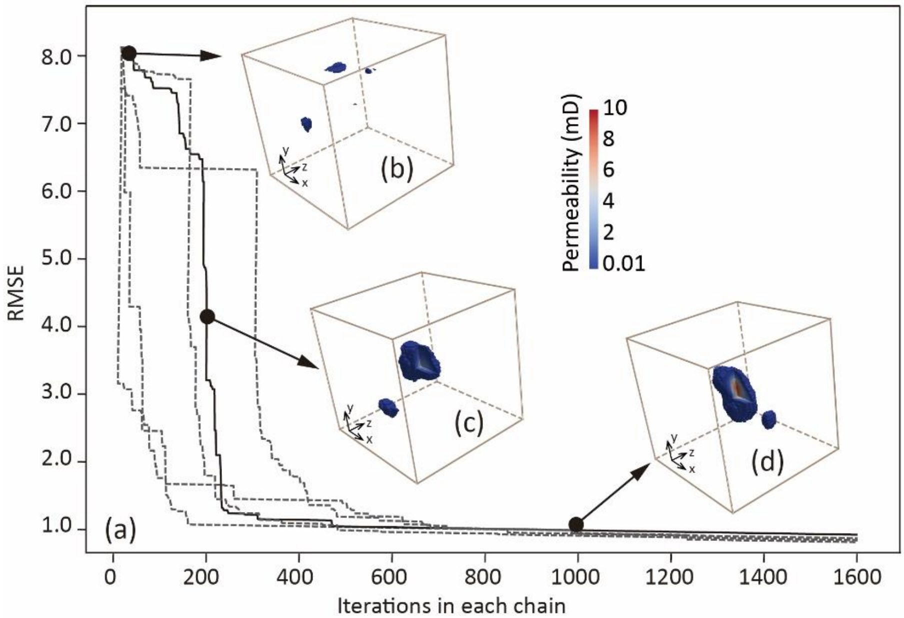

Figure 6. (a) Trace plot of RMSE values along five Markov chains over 1600 iterations and spatial distributions of fractured zone permeability values higher than 0.01 mD after 1 (b), 200 (c), and 2000 (d) iterations.

The 3D permeability in the iterations of 1400 to 1600 are selected to calculate the mean and CV at each voxel. As illustrated in Fig. 7b, the inversion based on multiple-stage SWIW tests can recreate the pattern of permeability, which decreases outwards from the centre of model domain and is comparable to the synthetic realisation in Fig. 7a. PNSR and SSIM metric values between inversed permeabilities and synthetic values are greater than 30 and 0.70, respectively.

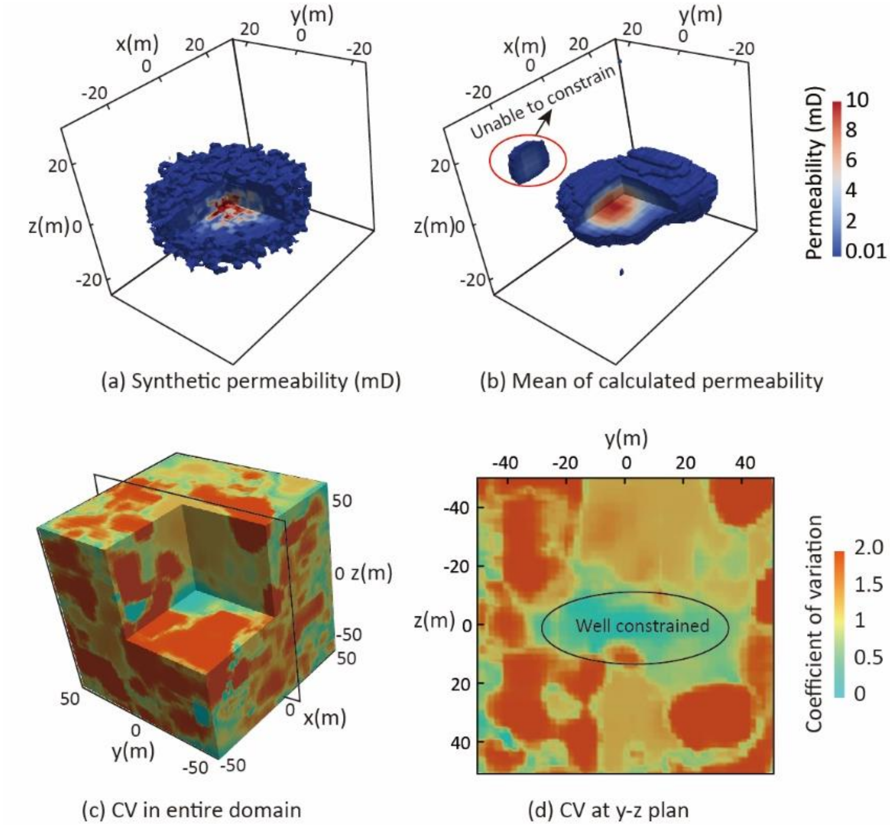

Figure 7. (a) Synthetic permeability in the illustrated case with value > 0.01 mD, which is comparable to (b) mean permeability generated by combining Bayesian inversion and decoder in the last 200 iterations; and (c) 3D and (d) 2D cross-sectional coefficient of variation (CV) of the generated permeability which is low in central permeable zones and high in the low-permeability area.

4 Discussion

4.1 Improvement of autoencoder structure

To improve the efficacy of the autoencoder, we added the convolution layers in the decoder after each transposed convolution layer (Fig. 1a). A small filter size of 2 ×2 ×2 was used. In contrast to a decoder without additional convolution layers, PSNR and SSIM metric values increased by more than 10 and 0.1, respectively (Fig. 8). This suggests that the performance of the autoencoder is improved significantly, because the additional convolution layers can enhance the mixing of the data flow via the network. Even when a small filter was used, the relationship between codes and full-dimension parameters can be improved by adding additional convolution layers.

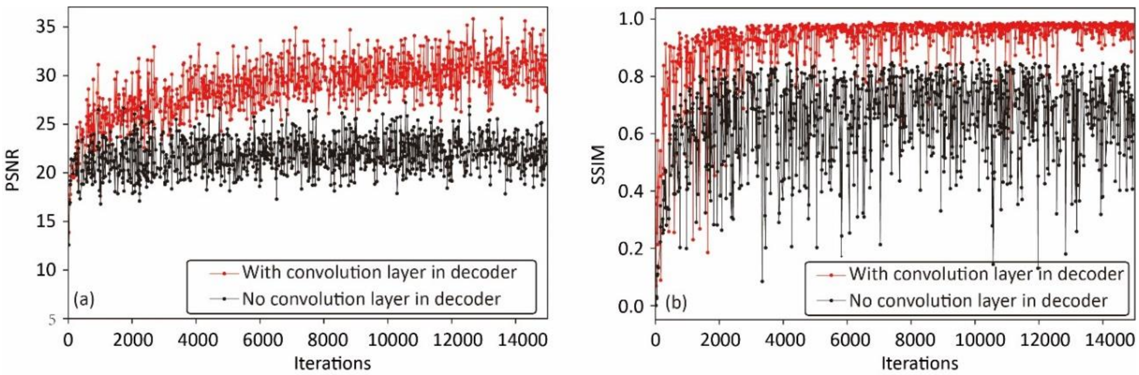

Figure 8. The performance of autoencoder network expressed by (a) PSNR and (b) SSIM by adding convolution layers in the decoder and without convolution layers in the decoder.

4.2 Calculation efficiency

Training of the autoencoder network using 15,000 iterations on a high-performance computer (Tesla P-100-SXM2-M-16GB) required approximately one hour of computation time. Once trained, the conversion of codes to full-dimension parameters via the decoder on a desktop computer (Intel Core i7-7600 CPU-2.80GHz-M-16GB) required less than five seconds of computation time. The estimates of low-dimension codes via MCMC were also fast, requiring less than two seconds on the same desktop computer. The most time-consuming aspect of the methodology presented was the forward modelling of fluid and heat transport. A single forward model run required more than 200 seconds. Bayesian inversion with 1600 iterations required 95.5 hours on a desktop machine. However, as demonstrated in Fig. 6, the burn-in period (i.e. approximately 800 iterations) was relatively short. Less iteration was required when estimating merely 20 codes in the Bayesian model rather than high-dimension parameters directly. Therefore, the considerable reduction of parameter dimensionality (i.e. from 262,144 voxels to 20 codes) achieved using the autoencoder network enabled efficient Bayesian inversion.

4.3 Advantages and Limitations

By combining deep learning methods of parameter dimensionality reduction with Bayesian inversion, it becomes possible to infer the heterogeneous distribution of fracture permeability in deep reservoirs based on single-well injection-withdrawal tests. It was observed in Fig. 7c and 7d that the CV of permeability values in the central zone of the model domain was <0.5, suggesting that permeability in this zone was well constrained by SWIW observations. However, as the influence range of SWIW testing was limited, model inversion based on SWIW observations cannot robustly constrain permeability values beyond the fracture zone. CV values (Fig. 7d) increased with distance from the central zone, indicating that permeability values fluctuated considerably during Bayesian inversion. Because permeability values distant to the fractured zone had limited influence on SWIW responses, it was difficult for the inversion model to update these parameters according to the discrepancies between tracer test observations and calculations. As a consequence, in Fig. 7b, there appeared an isolated high-permeable area which was disconnected to the major permeable zone. This can be overcome by adding additional well(s) and conducting cross-well tests or multiple single-well tests to assure that the tracer test observations can reflect the properties in the entire domain, which however depends on the deep borehole availability.

This study focused on the methodology development and tested it with a local-scale model. The size of model domain was given as 100 m×100 m×100 m with the fractured zone occupying the central part. A successful enhanced geothermal system may involve a fractured zone larger than the scale demonstrated in this work. In that case the model domain can be discretised into more voxels to assure the convergence of the forward model. Correspondingly, the training images with large number of voxels should be generated. Also, the hydraulic fracturing is often operated via horizontal wells, which creates parallel-plate-shape fracture zones. To train the neural network, the training images pool should be updated following the prior knowledges on the engineering realizations. However, if the structure of autoencoder network and number of weights are same as those in this work, the time taken to train the network will not increase with the size of training images.

The generation of training images is critical for training the autoencoder. An ideal method for the training images generation is to simulate the hydraulic fracturing processes by coupled fluid flow, heat transport and mechanical modelling, under extensively varying boundary and initial conditions. However, the coupled modelling is time consuming and the limited number of results from constrained boundary and initial conditions may not yield representative training images. We here proposed an empirical approach to generate a large training image pool, based on prior knowledges of artificial fracture patterns in hot dry rocks. The validation of joint autoencoder-Bayesian approach based on images independent to the training images suggested that the relationship between low-dimension codes and full-dimension parameters established on training images is reproducible.

A fixed permeability-porosity relationship (Appendix) was assumed in this study, which can be modified according to the specific situation (e.g. tectonic stress field, lithology, temperature and pressure) in the geothermal reservoir. We can also consider the permeability-porosity equation as an unknown factor to update in the inversion model, the accuracy of the results could be further improved. Tackling these problems is beyond the scope of this contribution and requires further investigation.

5 Conclusions

A deep learning approach featuring an autoencoder neural network was combined with a Bayesian inversion algorithm to estimate the high-dimension hydraulic parameters based on a single-well injection-withdrawal test. The following conclusions were drawn:

(1) An autoencoder network composed by four convolution layers in the encoder and four transposed convolution layers and three additional convolution layers in the decoder reduced parameter dimensionality by at least four orders of magnitude and was able to reconstruct the high-dimension parameters from low-dimension codes accurately. Importantly, there was no prior assumption on the probability distribution of the original high-dimension parameters, while the low-dimension codes can be designed to follow normal distribution, which made the estimates of these codes via subsequently Bayesian inversion much easier.

(2) Bayesian inversion based on Markov Chain Monte Carlo sampling was used to estimate low-dimension codes (rather than high-dimension parameters). By minimizing the misfit between modelled and observed tracer test responses (borehole temperature and fluid flux), parameter values were estimated after a short burn-in period, and parameter uncertainty was constrained.

(3) Application of the methodology to a simulated single-well injection-withdrawal test demonstrated that 262,144 high-dimension parameters could be represented using only 20 low-dimension codes. High-dimension parameters were subsequently reconstructed successfully using a decoder. The estimation of 20 codes via Bayesian inversion reproduced a heterogeneous spatial distribution of fracture zone permeability after 2000 iterations. Coefficient of variation values observed in the fractured zone were < 0.5.

The proposed methodology provided an efficient way to characterise complex spatial distributions of fracture zone permeability in deep reservoirs based on a single-well test.

Appendix Training dataset generation

Training images pool in this study is generated based on an empirical approach. Given an arbitrary position *x*, *y*, *z* in space, its spherical coordinate can be written as:

$$r = \sqrt{x^2 + y^2 + z^2}$$
$$\alpha = \mathrm{atan}(\frac{y}{x + \varepsilon}) \quad \text{(A1)}$$

$$\beta = \mathrm{atan}(\frac{z}{\sqrt{x^2 + y^2} + \varepsilon})$$

where r is the radial distance (< 50 m), β is the polar angle measured in zenith direction ($-\pi/2$, $\pi/2$), α is the azimuth angle measured on a reference plane orthogonal to the zenith (0, 2π). The reference position is located at the centre of the model, with r, α and β equal to zero.

Assuming that the fracture zone has an ellipsoid-like shape, the radial distance between model centre and outline of ellipsoid (R) at the direction of β and α can be expressed by:

$$R = \frac{1}{\sqrt{(cos\alpha \cdot cos\beta/a)^2 + (sin\alpha \cdot cos\beta/b)^2 + (sin\beta/c)^2}}, \tag{A2}$$

where R is defined as the range. a, b and c are main axes in x, y, z direction, respectively, which are randomly generated in 1 to 50 m following a uniform distribution.

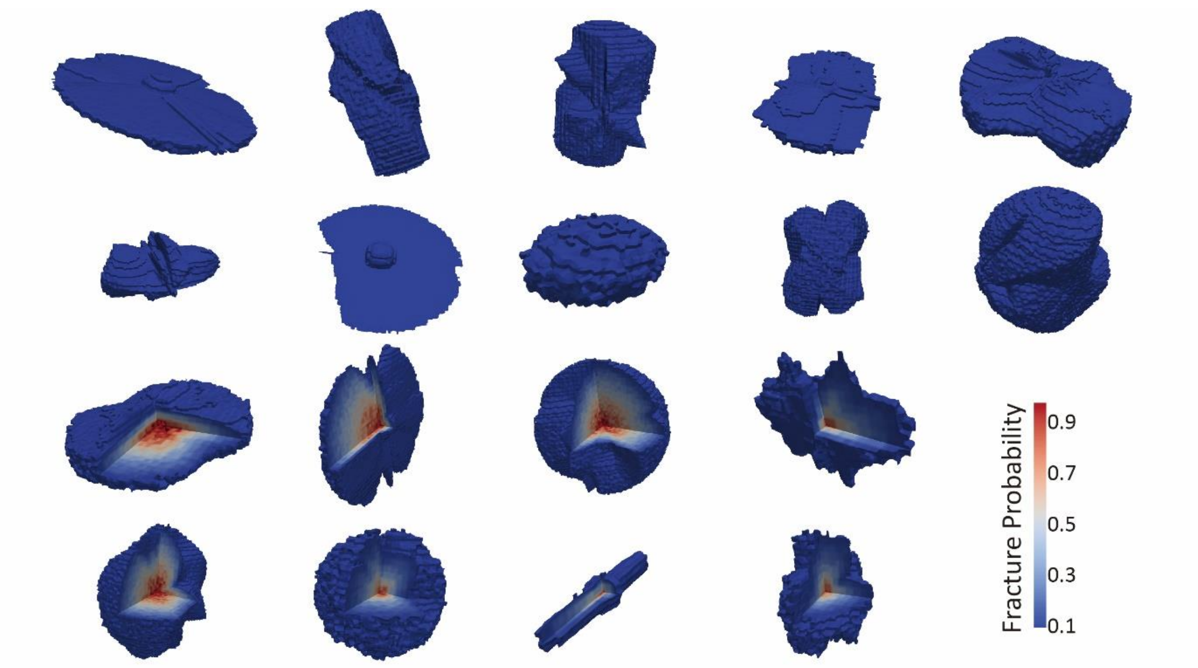

Figure A1. Training images with fracture zone of irregular shape and fracture probability reducing outward.

R at the angle (α and β) of every 45$^{\mathrm{o}}$, 30$^{\mathrm{o}}$ and 20$^{\mathrm{o}}$, respectively, are then interrupted by noise following uniform distribution in 0 to 25m, and R at other angles are obtained by the bicubic interpolation. Thus, fracture zones of irregular shapes can be created to represent the influence of heterogeneous mechanical rock properties, stresses and initial fractures (Fig. A1). In such manner, the freedom of the fracture zone shape is determined by the frequency of R variation. For instance, R varying among α and β by every 20$^{\mathrm{o}}$ corresponds to a freedom of 18 (in α direction)$\times$ 9 (in β direction) +1(interpolation method).

Furthermore, from the centre to the outline of fracture zone, the fracture probability (P) is defined to decrease from 1.0 to zero following linear and exponential randomly:

$$P(r, \alpha, \beta) = \begin{cases} \exp\left(-\frac{5r^2}{R^2}\right), Gaussian \\ \exp\left(-\frac{5r}{R}\right), exponential. \\ 1 - \frac{r}{R}, linear \end{cases} \tag{A3}$$

The resultant *P* values are contaminated by 10% random noise following a uniform distribution. Based on the method above, we created 6000 3D images of fracture probability with values ranging from 0.0 to 1.0 to train the autoencoder neural network.

For the Bayesian inversion, the fracture probability needs to be converted to the porosity and permeability. Since the permeability in the enhanced geothermal reservoir is often lower than 10.0 mD and the effective porosity is lower than 0.1 [*Ghassemi and Kumar*, 2007; *Zimmermann and Reinicke*, 2010], *P* is simply regarded as a normalized porosity, and converted to the porosity by:

$$\theta(r, \alpha, \beta) = P/10, \tag{A4}$$

and the porosity is converted to the permeability by a modified Kozeny–Carman equation [*Latief and Fauzi*, 2012; *Ma*, 2015]:

$$k = 10(\frac{\theta}{0.1})^{11}, \tag{A5}$$

Radom variation of *R* directions, interpolation method and fracture probability formula allow the resulting training images to follow the basic fracture permeability patterns created by artificial hydraulic fracturing via single vertical well. A single training image of fracture probability (or permeability) is of high freedom and do not necessarily follow a standard distributions (e.g. uniform or Gaussian). Thus, these images of fracture permeability cannot be compressed by zonation or pilot points approach, but require an advanced new tool like deep-learning neural network to reduce the dimensionality. Autoencoder neural network allows the automatic conversion between low-dimension codes (following standard normal distribution) and full-dimension parameters (with unknown probability distribution). Then, updating the codes following standard normal distribution made the Bayesian inversion well constrained and stabilized rapidly.

To guarantee the generic feature of training images, the fracture patterns in the synthetic study (Fig. 2) and validation images pool are generated by randomly varying *R* at angles of every 10^{o} and 60^{o} (which is not involved in training images generation) and the probability is calculated by a Gaussian model in Eq. A3, while the other factors controlling the fracture pattern are proceeded following the method above.

Acknowledgement

Funding support for this study was provided by the National Key R&D Program of China (2018YFC0604305-03), and the National Natural Science Foundation of China (No. 41572215 and No. 41502222) and also the China Postdoctoral funding (No. 2016M591483) and the International Postdoctoral Exchange Fellowship Program (2017) from the China Postdoctoral Council in combination with CSIRO funding through the Land and Water Business Unit and the Future Science Platform *Deep Earth Imaging*. We thank Lei Gao and Uli Kelka from CSIRO for the internal review on the manuscript, which helped improving the article quality.

Data Availability Statement

The python codes for 3D autoencoder neural network were developed by Tensorflow package, which is now provided as supportive materials for review. It is noted that the submitted package also include MCMC python codes developed by Laloy et al., to assure that these codes can run well together. After the final revision on the article and the codes, we would like to submit the final codes in the format as requested by the journal.

References

Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard (2016), Tensorflow: a system for large-scale machine learning, paper presented at OSDI.

Asher, M. J., B. F. Croke, A. J. Jakeman, and L. J. Peeters (2015), A review of surrogate models and their application to groundwater modeling, *Water Resources Research*, *51*(8), 5957-5973.

Bruce, L. M., C. H. Koger, and J. Li (2002), Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction, *IEEE Transactions on geoscience and remote sensing*, *40*(10), 2331-2338.

Calò, M., and C. Dorbath (2013), Different behaviours of the seismic velocity field at Soultz-sous-Forêts revealed by 4-D seismic tomography: case study of GPK3 and GPK2 injection tests, *Geophysical Journal International*, *194*(2), 1119-1137.

Chib, S., and E. Greenberg (1995), Understanding the metropolis-hastings algorithm, *The american statistician*, *49*(4), 327-335.

Cuenot, N., C. Dorbath, and L. Dorbath (2008), Analysis of the microseismicity induced by fluid injections at the EGS site of Soultz-sous-Forêts (Alsace, France): implications for the characterization of the geothermal reservoir properties, *Pure and Applied Geophysics*, *165*(5), 797-828.

Ding, C., and X. He (2004), K-means clustering via principal component analysis, paper presented at Proceedings of the twenty-first international conference on Machine learning, ACM.

Doherty, J. E., and R. J. Hunt (2010), *Approaches to highly parameterized inversion: a guide to using PEST for groundwater-model calibration*, US Department of the Interior, US Geological Survey.

Garabedian, S. P., D. R. LeBlanc, L. W. Gelhar, and M. A. Celia (1991), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 2. Analysis of spatial moments for a nonreactive tracer, *Water Resources Research*, *27*(5), 911-924.

Ghassemi, A., and G. S. Kumar (2007), Changes in fracture aperture and fluid pressure due to thermal stress and silica dissolution/precipitation induced by heat extraction from subsurface rocks, *Geothermics*, *36*(2), 115-140.

Goodfellow, I., Y. Bengio, and A. Courville (2016), *Deep learning*, MIT press Cambridge.

Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications.

Hinton, G. E., and R. R. J. s. Salakhutdinov (2006), Reducing the dimensionality of data with neural networks, *313*(5786), 504-507.

Hubbert, M. K., and D. G. Willis (1972), Mechanics of hydraulic fracturing.

Kambhatla, N., and T. K. Leen (1997), Dimension reduction by local principal component analysis, *Neural computation*, *9*(7), 1493-1516.

Kingma, D. P., and J. Ba (2014), Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*.

Kullback, S., and R. A. Leibler (1951), On information and sufficiency, *The annals of mathematical statistics*, *22*(1), 79-86.

Laloy, E., R. Hérault, D. Jacques, and N. Linde (2018), Training-Image Based Geostatistical Inversion Using a Spatial Generative Adversarial Neural Network, *Water Resources Research*, *54*(1), 381-406.

Laloy, E., R. Hérault, J. Lee, D. Jacques, and N. Linde (2017), Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network, *Advances in Water Resources*, *110*, 387-405.

Langevin, C. D. (2009), SEAWAT: A computer program for simulation of variable-density groundwater flow and multi-species solute and heat transport*Rep. 2327-6932*, US Geological Survey.

Langevin, C. D., D. T. Thorne Jr, A. M. Dausman, M. C. Sukop, and W. Guo (2008), SEAWAT version 4: a computer program for simulation of multi-species solute and heat transport*Rep. 2328-7055*, Geological Survey (US).

Latief, F. D. E., and U. Fauzi (2012), Kozeny–Carman and empirical formula for the permeability of computer rock models, *International Journal of Rock Mechanics and Mining Sciences*, *50*, 117-123.

LeBlanc, D. R., S. P. Garabedian, K. M. Hess, L. W. Gelhar, R. D. Quadri, K. G. Stollenwerk, and W. W. Wood (1991), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 1. Experimental design and observed tracer movement, *Water Resources Research*, *27*(5), 895-910.

Lee, J., and P. K. Kitanidis (2014), Large-scale hydraulic tomography and joint inversion of head and tracer data using the Principal Component Geostatistical Approach (PCGA), *Water Resources Research*, *50*(7), 5410-5427.

Legarth, B., E. Huenges, and G. Zimmermann (2005), Hydraulic fracturing in a sedimentary geothermal reservoir: Results and implications, *International Journal of Rock Mechanics and Mining Sciences*, *42*(7-8), 1028-1041.

Liu, W., Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi (2017), A survey of deep neural network architectures and their applications, *Neurocomputing*, *234*, 11-26.

Ma, J. (2015), Review of permeability evolution model for fractured porous media, *Journal of Rock Mechanics and Geotechnical Engineering*, *7*(3), 351-357.

Majer, E. L., R. Baria, M. Stark, S. Oates, J. Bommer, B. Smith, and H. Asanuma (2007), Induced seismicity associated with enhanced geothermal systems, *Geothermics*, *36*(3), 185-222.

Maloszewski, P., A. Herrmann, and A. Zuber (1999), Interpretation of tracer tests performed in fractured rock of the Lange Bramke basin, Germany, *Hydrogeology Journal*, *7*(2), 209-218.

Maurer, V., N. Cuenot, E. Gaucher, M. Grunberg, J. Vergne, H. Wodling, M. Lehujeur, and J. Schmittbuhl (2015), Seismic monitoring of the Rittershoffen EGS project (Alsace, France), paper presented at Proceedings World Geothermal Congress.

Maxwell, S. C., J. Rutledge, R. Jones, and M. Fehler (2010), Petroleum reservoir characterization using downhole microseismic monitoring, *Geophysics*, *75*(5), 75A129-175A137.

McLaughlin, D., and L. R. Townley (1996), A reassessment of the groundwater inverse problem, *Water Resources Research*, *32*(5), 1131-1161.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *The journal of chemical physics*, *21*(6), 1087-1092.

Prensky, S. E. (1999), Advances in borehole imaging technology and applications, *Geological Society, London, Special Publications*, *159*(1), 1-43.

Sanjuan, B., J.-L. Pinault, P. Rose, A. Gérard, M. Brach, G. Braibant, C. Crouzet, J.-C. Foucher, A. Gautier, and S. Touzelet (2006), Tracer testing of the geothermal heat exchanger at Soultz-sous-Forêts (France) between 2000 and 2005, *Geothermics*, *35*(5-6), 622-653.

Sovacool, B. K. (2014), Cornucopia or curse? Reviewing the costs and benefits of shale gas hydraulic fracturing (fracking), *Renewable and Sustainable Energy Reviews*, *37*, 249-264.

Tsang, Y. (1995), Study of alternative tracer tests in characterizing transport in fractured rocks, *Geophysical Research Letters*, *22*(11), 1421-1424.

Vidal, J., A. Genter, and F. Chopin (2017), Permeable fracture zones in the hard rocks of the geothermal reservoir at Rittershoffen, France, *Journal of Geophysical Research: Solid Earth*, *122*(7), 4864-4887.

Vogt, C., C. Kosack, and G. Marquart (2012), Stochastic inversion of the tracer experiment of the enhanced geothermal system demonstration reservoir in Soultz-sous-Forêts—Revealing pathways and estimating permeability distribution, *Geothermics*, *42*, 1-12.

Voss, C. I. (1984), *A finite-element simulation model for saturated-unsaturated, fluid-density-dependent ground-water flow with energy transport or chemically-reactive single-species solute transport*, US Geological Survey Reston, VA.

Vrugt, J. A. (2016), Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation, *Environmental Modelling & Software*, *75*, 273-316.

Wall, M. E., A. Rechtsteiner, and L. M. Rocha (2003), Singular value decomposition and principal component analysis, in *A practical approach to microarray data analysis*, edited, pp. 91-109, Springer.

Wang, Z., and A. C. Bovik (2002), A universal image quality index, *IEEE signal processing letters*, *9*(3), 81-84.

Wang, Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli (2004), Image quality assessment: from error visibility to structural similarity, *IEEE transactions on image processing*, *13*(4), 600-612.

Yeh, T. C. J., and S. Liu (2000), Hydraulic tomography: Development of a new aquifer test method, *Water Resources Research*, *36*(8), 2095-2105.

Yeh, W. W. G. (1986), Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resources Research*, *22*(2), 95-108.

Zimmermann, G., and A. Reinicke (2010), Hydraulic stimulation of a deep sandstone reservoir to develop an Enhanced Geothermal System: Laboratory and field experiments, *Geothermics*, *39*(1), 70-77.