Bayesian Inversion Of Generative Models For Geologic Storage Of Carbon Dioxide

Gavin H. Graham, Yan Chen
Total
Aberdeen, Scotland, UK.
{gavin.graham, yan.chen}@total.com

Abstract

Carbon capture and storage (CCS) can aid decarbonization of the atmosphere to limit further global temperature increases. A framework utilizing unsupervised learning is used to generate a range of subsurface geologic volumes to investigate potential sites for long-term storage of CO_2. Generative adversarial networks are used to create geologic volumes, with a further neural network used to sample the posterior distribution of a trained Generator conditional to sparsely sampled physical measurements. These generative models are further conditioned to historic dynamic fluid flow data through Bayesian inversion to improve the resolution of the forecast of the storage capacity of injected CO_2.

1 Introduction

Limiting global warming to well below $2°C$, but preferably limiting the temperature increase to $1.5°C$ above pre-industrial levels, was the key outcome from the 2015 United Nations International Climate Change Conference [1]. While this requires a substantive effort in reducing emissions from energy generation and other industrial activities, removing carbon dioxide (CO_2) from the atmosphere is also necessary to ensure that limit is not exceeded [2, 3]. One method to aid decarbonization of the atmosphere is through the use of carbon capture and storage (CCS) [4]. CO_2 storage by injection into deep geologic formations at depleted hydrocarbon fields has been demonstrated as a safe and effective method [5]. To adequately assess the potential site for long-term storage of CO_2, the subsurface geologic volume of interest must be characterized in order to provide numerical models for fluid flow simulation experiments of CO_2 injection.

Detailed characteristics of subsurface geologic properties are only available from sparse physical measurements, such as borehole logs which sample only a small fraction of the subsurface volume. It is therefore necessary to infer geologic properties away from these available data points. Statistical methods for creating geologic patterns and features include geostatistics [6]. Such methods can be used to generate samples of the subsurface that mimic observed geologic features that are far from multivariate Gaussian random fields, but exhibit clear spatial correlations (Fig. 1). In addition to borehole data, historic fluid flow data from the depleted hydrocarbon field can be used to further calibrate the geologic models in order to reduce uncertainty in the subsurface geologic properties [7]. Markov Chain Monte Carlo (MCMC) methods could be used to sample conditional realizations, but are computational costly as a significant number of fluid flow simulations are typically required [8]. Alternatively, ensemble-based data assimilation methods are shown to provide satisfactory approximate sampling with a limited amount of computational cost in applications ranging from weather forecasting to subsurface characterization [15].

However, geologic model parameterization is challenging for inversion techniques which rely on the Gaussian assumption [14]. In this paper we utilize Generative Adversarial Networks (GANs) as a method for dimensionality reduction and re-parameterization of the geologic models. A pre-

trained Generator is then combined with a neural network, the *inference network*, that is trained to sample the posterior distribution of the latent input to the Generator, conditional to the data at known borehole locations. This preserves the parametrization of the Generator and reduces the number of computationally expensive fluid flow simulations required for uncertainty quantification and inversion problems. The Generator and Inference networks are then used to produce a range of plausible geologic models conditional to the borehore data. An ensemble-based data assimilation method is then used to obtain samples of the geologic model that also honor historic fluid flow data. Finally, the calibrated models are used to forecast the subsurface response to CO_2 injection.
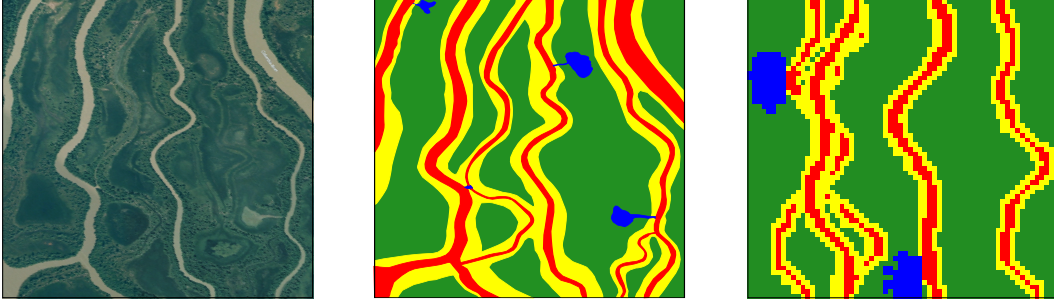

Figure 1: Left: Image of the Columbia river, British Columbia, Canada [9], which can be used as a modern analog to ancient sedimentary systems. Middle: Interpretation of sedimentary features, with river channels in red, levees in yellow and crevasse splays in blue. Right: A stochastic geostatistical simulation of the sedimentary features [6].

2 Methodology

2.1 Generative Adversarial Networks

Generative Adversarial Networks (GANs) [10] consist of two competing functions (typically neural networks), a Generator $G(z)$ and a Discriminator $D(y)$, with an objective function to find the Nash equilibrium between the two networks. The Generator $G(z)$ maps an input noise vector z from a simple Gaussian distribution to a synthetic sample y. The goal of $G(z)$ is to generate samples with a distribution p_g that is close to the data generating distribution p_{data}. The Discriminator, $D(y)$, is a classifier that takes a sample as input and tries to determine if this sample is real or synthetic. The output of $D(y)$ is a scalar representing the probability of y coming from the data. This interplay between G and D is formulated as a two-player minimax game:

$$\min_G \max_D \mathbb{E}_{x \sim p_{data}}[\log D(x)] + \mathbb{E}_{\tilde{x} \sim p_g}[\log(1 - D(\tilde{x}))] \tag{1}$$

where p_{data} is the data distribution and p_g is the model distribution implicitly defined by $\tilde{x} = G(z), z \sim p(z)$. The training process consists of simultaneous application of Stochastic Gradient Descent on D and G. Training alternates between k steps of optimizing D and one step of optimizing G. The process of training stops when D is unable to distinguish p_g and p_{data} i.e. $D(y) = 1/2$ or when $p_g = p_{data}$. Due to the opposing nature of the objective function, training GANs is inherently unstable and finding stable training methods remains an open research problem. In particular, if the discriminator D is optimally trained, it becomes saturated and provides no useful information for improvement of G. On the other hand, over-training G might result in mode collapse of the generated distribution toward a single sample that D always accepts. Wasserstein GAN (WGAN) [11] minimizes the Wassterstein distance rather than the Jensen-Shannon divergence in [10] and provides a meaningful loss metric during training that correlates with the quality of generated samples. This is important for assessing training progress and providing a criteria for convergence and was used in this study. For more detailed discussion on WGAN, the reader is referred to [11].

2.2 Inference network

Generated images must be calibrated to borehole data as geologic properties are known at these locations *a priori* to image generation. Given a pre-trained generator G, the goal is to find z such that

$G(z)$ honours the borehole data. An *inference network*, I, introduced by [12] was used as a basis to retain the parameterization of G. Let d_{obs} denote the borehole data and $G(z)_{\mathrm{obs}}$ the generated image values at the borehole locations given $G(z)$. The loss can then be defined as:

$$\mathcal{L}(z) = \|G(z)_{\mathrm{obs}} - d_{\mathrm{obs}}\|^2 + \lambda\|z\|^2 \tag{2}$$

where $\lambda = \sigma^2$ and σ is the measurement standard deviation. The first term in Eq. 2 represents the error that is the difference between the observed borehole data and the generated images. The second term of Eq. 2 ensures the distribution of z remains close to the prior. For more information the reader is referred to [12]. This method ensures that conditional samples of G are generated without the need to repeatedly use a local optimizer and different initial guesses for z, or sampling the full posterior using Markov Chain Monte Carlo methods.

2.3 Iterative ensemble smoother

A particular type of the ensemble-based data assimilation methods, Ensemble Smoother with Multiple Data Assimilation (see procedure of ESMDA in Table 1), is used to calibrate the input of the inference network to honor dynamic fluid flow data observed at borehole locations. The vector of model parameters $\mathbf{m}$ in Table 1 is the input of the inference network. It was shown that for linear Gaussian cases, α at all iterations need to satisfy the following condition, $\sum_{\ell=1}^{N_{\mathrm{a}}} 1/\alpha_\ell = 1$, in order to obtain the correct posterior mean and covariance at the final iteration in the limit of an infinite ensemble size [13]. An easy choice of α is $\alpha_\ell = N_{\mathrm{a}}$, often referred to as ESMDA with equal weights.

---

1. **Initialization:** generate initial ensemble, $\left\{\mathbf{m}_j^0\right\}_{j=1}^{N_{\mathrm{e}}}$, by sampling the prior distribution of model parameters. Choose number of data assimilations, N_{a} and coefficients α_ℓ .

2. For $\ell = 1$ to N_{a}:

   (a) **Forecast step:** run the forward model to compute the vector of simulated data

   $$\mathbf{d}_{\mathrm{sim},j}^\ell = \mathbf{g}(\mathbf{m}_j^\ell) \quad \text{for } j = 1, 2, \ldots, N_{\mathrm{e}} \tag{3}$$

   (b) Perturb the vector of historic data using

   $$\mathbf{d}_{\mathrm{hist},j}^\ell = \mathbf{d}_{\mathrm{hist}} + \sqrt{\alpha_\ell}\mathbf{C}_{\mathrm{D}}^{1/2}\mathbf{z}_j^\ell \quad \text{for } j = 1, 2, \ldots, N_{\mathrm{e}}, \text{ where } \mathbf{z}_j^\ell \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \tag{4}$$

   (c) **Analysis step:** update the vector of model parameters using

   $$\mathbf{m}_j^{\ell+1} = \mathbf{m}_j^\ell + \mathbf{C}_{\mathrm{md}}^{\mathrm{e}\,\ell}(\alpha_\ell\mathbf{C}_{\mathrm{D}} + \mathbf{C}_{\mathrm{dd}}^{\mathrm{e}\,\ell})^{-1}\left(\mathbf{d}_{\mathrm{hist},j}^\ell - \mathbf{d}_{\mathrm{sim},j}^\ell\right) \quad \text{for } j = 1, 2, \ldots, N_{\mathrm{e}} \tag{5}$$

   End (for)

Description of symbols	
Scalar	ℓ is the iteration index, N_{a} is the total number of iterations, α is a iteration tuning parameter (a easy choice is $\alpha_\ell = N_{\mathrm{a}}$ for all ℓs), j is the index for realizations, N_{e} is the ensemble size, N_{d} is the number of data and N_{m} is the number of model parameters
Vector	$\mathbf{m}_j$ of length N_{m} is the jth realization of model parameters, $\mathbf{d}_{\mathrm{hist}}$ of length N_{d} is the historic data, $\mathbf{g}(\mathbf{m}_j)$ of length N_{d} is the simulated data from $\mathbf{m}_j$ ($\mathbf{g}(\cdot)$ is the flow equations represented by the reservoir simulator), $\mathbf{z}_j$ of length N_{d} is white noise sampled from $\mathcal{N}(\mathbf{0}, \mathbf{I})$
Matrix	Data noise is assumed to follow multi-variate Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathrm{D}})$, where $\mathbf{C}_{\mathrm{D}}$ is a diagonal matrix of size $N_{\mathrm{d}} \times N_{\mathrm{d}}$ with σ_{d}^2 on its diagonal, $\mathbf{C}_{\mathrm{md}}^{\mathrm{e}}$ is the sample covariance between realizations of $\mathbf{m}_j$ and $\mathbf{g}(\mathbf{m}_j)$, and $\mathbf{C}_{\mathrm{dd}}^{\mathrm{e}}$ of size $N_{\mathrm{d}} \times N_{\mathrm{d}}$ is the sample covariance of realizations of $\mathbf{g}(\mathbf{m}_j)$. The superscript "e" indicates these covariances are computed from the ensemble

Table 1: Algorithm of ESMDA (adapted from [13]).

3   Results

As a first step, the methodology was tested on simplified binary representations of the geologic features in Figure 1. A reference case was generated, from which borehole data were sampled and historic fluid flow data were simulated using a multi-phase fluid flow simulator. The reference geologic features were not part of the training dataset but were generated using the same parameters through a geostatistical simulation method. The images produced by the conditional Generator were geologically consistent in comparison to the training data upon visual inspection and respected the borehole data (Fig. 2). An ensemble with 100 members conditional to the borehole data was generated, followed by 20 iterations of ESMDA. In each ESMDA iteration 100 flow simulations were completed in parallel. The final calibrated models were obtained after 2000 fluid flow simulations. The match to historic fluid flow data was satisfactory (last panel of Fig. 2). Four realizations from the final ensemble of 100 calibrated models are shown in the middle panel of Fig. 2. Due to the constraints imposed by the borehole data and the time series of phase rates and pressure data, the final realizations share many similar features. Although there are still subtle variations in the local positioning of river channel features among the final realizations, it is very likely that the final ensemble of calibrated realizations under-estimates the true uncertainty. The input space of the inference network I and the latent space of the generator z also become correlated as a result of conditioning to both static and dynamic data.
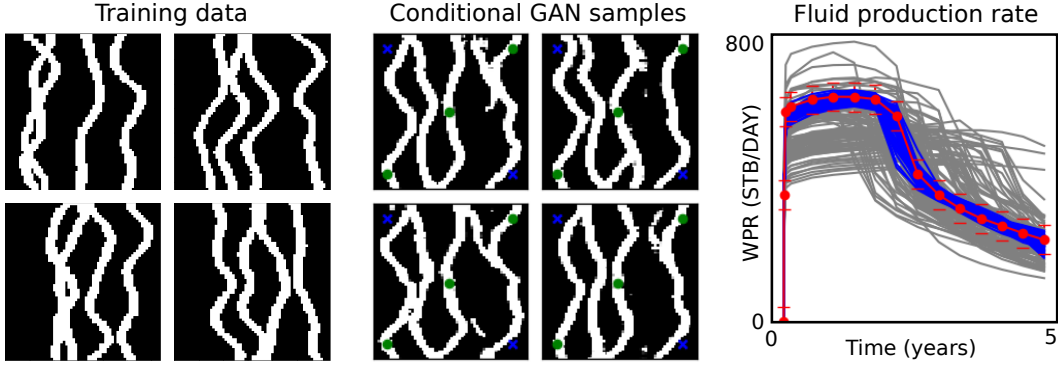

Figure 2: Left: Stochastic geostatistical simulations of simplified river channels used as training data for a GAN. Middle: Random samples from a trained Generator conditional on borehole data and historic fluid production data. Green circles represent boreholes that intersect river channels and blue crosses represent boreholes that intersect the background. Note that each conditional sample respects the known data points. Right: Fluid flow simulation data. Simulated data from initial set of generated models prior to inversion are shown in grey, and simulated data after Bayesian inversion shown in blue. Historic fluid flow data are shown as red dots with their associated uncertainty shown as error bars.

4   Conclusions

Subsurface geologic volumes can be calibrated to known data in order to reduce uncertainty in numerical models for computational fluid flow experiments of CO_2 injection. Conditional sampling of a Generator network through an inference network produced geologically consistency features while honoring known static borehole data. The ensemble-based data assimilation method was effective in calibrating the input of the inference network to provide further conditioning to dynamic fluid flow observations. The proposed methodology can be extended to more complex and realistic geologic models and can provide a basis for assessing the capacity of CO_2 injection and storage in subsurface geologic formations.

References

[1] Adoption of the Paris Agreement Report No. FCCC/CP/2015/L.9/Rev.1. United Nation Framework Convetion on Climate Change, 2015.

[2] D. P. Schrag. Preparing to Capture Carbon. *Science*, 315:5813 812–813, 2007. doi: 10.1126/science.1137632.

[3] Anna B. Harper, Tom, Powell, et al. Land-use emissions play a critical role in land-based mitigation for Paris climate targets. *Nature Communications*, 9:1, 2018. doi: 10.1038/s41467-018-05340-z

[4] D. C. Thomas and S. M. Benson. Carbon dioxide capture for storage in deep geological formations-results from the CO2 capture project, 2005. Elsevier Science.

[5] C. R. Jenkins, P. J. Cook, et al. Safe storage and effective monitoring of CO2 in depleted gas fields, 2012. *Proceedings of the National Academy of Sciences*, 109:2 E35–E41. doi: 10.1073/pnas.1107255108.

[6] C. V Deutsch, A. G. Journel, et al. Geostatistical software library and users guide, 1992. New York.

[7] M. Alfi and S. A. Hosseini. Integration of reservoir simulation, history matching, and 4D seismic for CO_2-EOR and storage at Cranfield, Mississippi, USA, 2016. *Fuel*, 175, 116–128 doi: 10.1016/j.fuel.2016.02.032

[8] E. laloy, R. Hérault et al. Training-image based geostatistical inversion using a spatial generative adversarial network, 2018. *Water Resources Research*, 54:1, 381–406.

[9] Google Earth. 2013.

[10] I. Goodfellow, J. Pouget-Abadie et al. Generative Adversarial Nets, 2014 *Advances in Neural Information Processing Systems*, 27 2672–2680.

[11] M. Arjovsky, S. Chintala and L. Bottou. Wasserstein GAN, 2017. *arXiv preprint arXiv:1701.07875*

[12] S. Chan and A. H. Elsheikh Parametric generation of conditional geological realizations using generative neural networks. *arXiv preprint arXiv:1807.05207*

[13] A. A. Emerick and A. C. Reynolds. Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55:3–15, 2013.

[14] S. W. Canchumuni, A. A. Emerick, and M. A. C. Pacheco. History matching geological facies models based on ensemble smoother and deep generative models. *Journal of Petroleum Science and Engineering*, 177:941 – 958, 2019.

[15] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer Verlag, second edition, 2009.