Laboratório
Nacional de
Computação
Científica

COMOD
Coordenação de Modelagem Computacional

Relatório de Atividades

Servidor: Marcio Borges

Petrópolis-RJ
September 3, 2023

1 Introdution

Natural reservoirs display a high degree of variability in their hydraulic properties in multiple length scales (Dagan, 1989; Gelhar, 1993). Such variability plays a strong impact on fluid flow patterns in subsurface formations (Glimm et al., 1992). Due to operational difficulties and high cost, direct measurements of the properties of interest are scarce and, therefore, a deterministic description of the reservoirs cannot be achieved. Alternatively, a stochastic approach should be adopted to characterize uncertainties in reservoir parameters that, according to Efendiev et al. (2006), are the primary factor contributing to reservoir forecasting uncertainties. In other words, the predictability of computational models is severely limited by the lack of an adequate description of reservoir properties.

Data acquisition along with the use of high-performance computing has encouragedThe growing data acquisition along with the use of high-performance computing has encouraged the use of dynamic data directly in simulations to reduce the associated uncertainties and increase the predictability of the models. Well tests, production history, pressure variation on monitors, and tracer tests, among others, are direct measures of reservoir responses and can provide important information about the flow processes.

Matching field dynamic data with reservoir simulation results is a stochastic inverse problem that can be formalized in terms of Bayesian inference and Markov chain Monte Carlo methods (MCMC). The Bayesian inference is convenient in quantifying the added value of information from several sources, while MCMC methods provide a computational framework for sampling from *a posteriori* distribution (Robert and Casella, 2005; Liu et al., 2008). Within this context MCMC methods have been widely used in the last decade in porous media flow problems (Liu and Oliver, 2003; Efendiev et al., 2005, 2006; Ma et al., 2008; Dostert et al., 2009; Das et al., 2010; Mondal et al., 2010; Ginting et al., 2011, 2012; Iglesias et al., 2013; Emerick and Reynolds, 2013).

1.1 MCMC

In 2000, Computing in Science & Engineering, a joint publication of the IEEE Computer Society and the American Institute of Physics, placed the Metropolis Algorithm on the list of the 10 algorithms with the greatest impact on the development of science and engineering practice in the 20th century (" Top Ten Algorithms of the Century" (Madey et al., 2005)).

1.2 Random permeability fields

Due to incomplete knowledge about the rock properties that show variability at multiple length scales, input parameters such as the permeability field, $\boldsymbol{\kappa}(\mathbf{x},\omega)$, are treated as random space functions with statistics inferred from geostatistical models (here $\mathbf{x} = (x_1, x_2, x_3)^{\mathsf{T}} \in \mathbb{R}^3$ and ω is a random element in the probability space). In line with Dagan (1989) and Gelhar (1993) the permeability field is modeled as a log-normally distributed random space function

$$\boldsymbol{\kappa}(\mathbf{x},\omega) = \beta \exp\Big[\rho \mathsf{Y}(\mathbf{x},\omega)\Big], \tag{1}$$

where $\beta, \rho \in \mathbb{R}^+$ and $\mathsf{Y}(\mathbf{x}, \omega) \sim \mathbb{N}(\boldsymbol{\mu}_\mathsf{Y}, \boldsymbol{\mathcal{C}}_\mathsf{Y})$ is a Gaussian random field characterized by its mean $\boldsymbol{\mu}_\mathsf{Y} = \langle \mathsf{Y}(\mathbf{x}) \rangle$ and two-point covariance function

$$\boldsymbol{\mathcal{C}}_\mathsf{Y}(\mathbf{x}, \mathbf{y}) = \mathsf{Cov}(\mathsf{Y}(\mathbf{x}), \mathsf{Y}(\mathbf{y})) = \mathsf{E}\Big[\big(\mathsf{Y}(\mathbf{x}) - \langle \mathsf{Y}(\mathbf{x}) \rangle\big)\big(\mathsf{Y}(\mathbf{y}) - \langle \mathsf{Y}(\mathbf{y}) \rangle\big)\Big]. \tag{2}$$

Moreover, in this work, Y is a second-order stationary process (Gelhar, 1993), that is:

$$\begin{aligned} \langle \mathsf{Y}(\mathbf{x}) \rangle &= \mu_\mathsf{Y}, \quad \text{(constant)} \\ \boldsymbol{\mathcal{C}}_\mathsf{Y}(\mathbf{x}, \mathbf{y}) &= \boldsymbol{\mathcal{C}}_\mathsf{Y}(\|\mathbf{x} - \mathbf{y}\|) = \boldsymbol{\mathcal{C}}_\mathsf{Y}(d). \end{aligned} \tag{3}$$

1.3 Karhunen-Loève expansion

The Gaussian field Y can be represented as a series expansion involving a complete set of deterministic functions with correspondent random coefficients using the Karhunen-Loève (KL) expansion proposed independently by Karhunen (1946) and Loève (1955). It is based on the eigen-decomposition of the covariance function. Depending on how fast the eigenvalues decay one may be able to retain only a small number of terms in a truncated expansion and, consequently, this procedure may reduce the search to a smaller parameter space. In uncertainty quantification methods for porous media flows, the KL expansion has been widely used to reduce the number of parameters used to represent the permeability field (Efendiev et al., 2005, 2006; Das et al., 2010; Mondal et al., 2010; Ginting et al., 2011, 2012). Another advantage of KL expansion lies on the fact that it provides orthogonal deterministic basis functions and uncorrelated random coefficients, allowing for the optimal encapsulation of the information contained in the random process into a set of discrete uncorrelated random variables (Ghanem and Spanos, 1991). This remarkable feature can be used to simplify the Metropolis-Hastings McMC Algorithm in the sense of the search may be performed in the space of discrete uncorrelated random variables ($\boldsymbol{\theta}$), no longer in the space of permeabilities which have a more complex statistical structure.

2 Introdution

Variational auto-encoder (VAE) model is a stochastic inference and learning algorithm based on variational Bayes (VB) inference proposed by Kingma and Welling (2014). This is a generative technique whose central idea is the development of representations in a low-dimensional latent space that can be mapped back into a realistic-looking image.

Higgins et al. (2016) introduced the β-VAE, a modification of the original VAE, that introduces an adjustable hyperparameter β to balance latent channel capacity and independence constraints with reconstruction accuracy. They demonstrate that with tuned values of β ($\beta > 1$) the β-VAE outperforms VAE ($\beta = 1$).

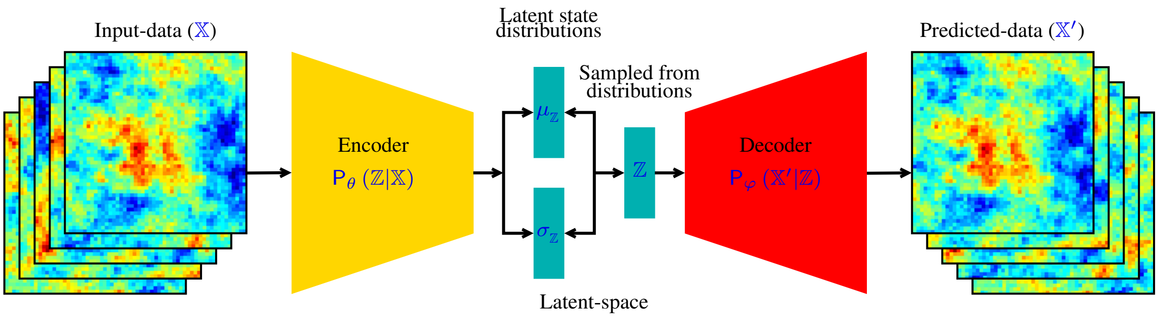
Makhzani et al. (2016); Louizos et al. (2017); Burda et al. (2016); Zheng et al. (2019); Vahdat and Kautz (2020)

3 Zhang et al. (2022)

Variational auto-encoder (VAE) model is a generative network based on variational Bayes (VB) inference proposed by Kingma and Welling (2014).

Zhang et al. (2022) proposed a method to reconstruct porous media based on VAE and Fisher information with good quality and efficiency.

Consider the input data $\mathbb{X} = \left\{\boldsymbol{x}^{(i)}\right\}_{i=1}^{\mathsf{N}}$ consisting of N *i*ndependent and *i*dentically *d*istributed (*i.i.d.*) samples of the continuous (or discrete) variable $\boldsymbol{x}$.



$$\mathbb{Z} = \mu_{\mathbb{Z}} + \sigma_{\mathbb{Z}} \cdot \varepsilon, \quad \text{where} \quad \varepsilon \sim \mathbb{N}(0,1) \tag{4}$$

The Kullback–Leibler divergence (also called relative entropy and I-divergence) is a measure of divergence between two distributions (Kullback and Leibler, 1951; Csiszar, 1975):

$$\mathcal{D}_{\mathsf{KL}}\left(f_{\mathbb{P}}||f_{\mathbb{Q}}\right) = \tag{5}$$

REFERENCES

Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders, 2016.

I. Csiszar. I-Divergence Geometry of Probability Distributions and Minimization Problems. *The Annals of Probability*, 3(1):146 – 158, 1975. doi: 10.1214/aop/1176996454. URL https://doi.org/10.1214/aop/1176996454.

G. Dagan. *Flow and transport in porous formations*. Springer-Verlag, 1989.

N. N. Das, B. P. Mohanty, and Y. Efendiev. Characterization of effective saturated hydraulic conductivity in an agricultural field using Karhunen-Loève expansion with the Markov chain Monte Carlo technique. *Water Resources Research*, 46, W06521 2010. doi: 10.1029/2008WR007100.

P. Dostert, Y. Efendiev, and B. Mohanty. Efficient uncertainty quantification techniques in inverse problems for richards' equation using coarse-scale simulation models. *Advances in Water Resources*, 32(3):329–339, 2009. doi: DOI:10.1016/j.advwatres.2008.11.009.

Y. Efendiev, A. Datta-gupta, V. Ginting, X. Ma, and B. Mallick. An efficient two-stage Markov chain Monte Carlo method for dynamic data integration. *Water Resources Research*, 41: 12423, 2005.

Y. Efendiev, T. Hou, and W. Luo. Preconditioning Markov Chain Monte Carlo Simulations Using Coarse-Scale Models. *SIAM J. Sci. Comput.*, 28:776–803, 2006.

A. A. Emerick and A. C. Reynolds. Investigation of the sampling performance of ensemble-based methods with a simple reservoir model. *Computational Geosciences*, 17(2):325–350, 2013. ISSN 1420-0597. doi: 10.1007/s10596-012-9333-z. URL http://dx.doi.org/10.1007/s10596-012-9333-z.

L. W. Gelhar. *Stochastic subsurface hydrology*. Englewood Cliffs. Prentice-Hall, 1993.

R. Ghanem and P.D. Spanos. *Stochastic Finite Element: A Spectral Approach*. Springer, New York, 1991.

V. Ginting, F. Pereira, M. Presho, and S. Wo. Application of the two-stage Markov chain Monte Carlo method for characterization of fractured reservoirs using a surrogate flow model. *Computational Geosciences*, 15:691–707, 2011.

V. Ginting, F. Pereira, and A. Rahunanthan. Multiple Markov chains Monte Carlo approach for flow forecasting in porous media. *Procedia Computer Science*, 5:707–716, 2012.

J. Glimm, W. B. Lindquist, F. Pereira, and R. Peierls. The multi-fractal hypothesis and anomalous diffusion. *Mat. Aplic. Comp.*, 11(2):189–207, 1992.

REFERENCES

Irina Higgins, Loïc Matthey, Arka Pal, Christopher P. Burgess, Xavier Glorot, Matthew M. Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. In *International Conference on Learning Representations*, 2016.

M. A. Iglesias, K. J. H. Law, and A. M. Stuart. Evaluation of gaussian approximations for data assimilation in reservoir models. *Computational Geosciences*, pages 1–35, 2013. doi: 10.1007/s10596-013-9359-x.

K. Karhunen. Zur spektraltheorie stochastischer prozesse. *Ann. Acad. Sci. Fennicae*, 1946.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *International conference on learning representations*, pages 14–27, 2014.

S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79 – 86, 1951. doi: 10.1214/aoms/1177729694. URL `https://doi.org/10.1214/aoms/1177729694`.

Gaisheng Liu, Yan Chen, and Dongxiao Zhang. Investigation of flow and transport processes at the made site using ensemble kalman filter. *Advances in Water Resources*, 31(7):975–986, 2008. doi: 10.1016/j.advwatres.2008.03.006.

N. Liu and D. S. Oliver. Evaluation of monte carlo methods for assessing uncertainty. *SPE Journal*, 8:188–195, 2003. doi: 10.2118/84936-PA.

M. M. Loève. *Probability Theory*. Princeton, N.J., 1955.

Christos Louizos, Kevin Swersky, Yujia Li, Max Welling, and Richard Zemel. The variational fair autoencoder, 2017.

X. Ma, M. Al-Harbi, A. Datta-Gupta, and Y. Efendiev. An efficient two-stage sampling method for uncertainty quantification in hystory matching geological models. *SPE Journal*, 2008.

G. Madey, X. Xiang, S. E. Cabaniss, and Y. Huang. Agent-based scientific simulation. *Computing in Science & Engineering*, 2(01):22–29, jan 2005. ISSN 1558-366X. doi: 10.1109/MCSE.2005.7.

Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, and Ian Goodfellow. Adversarial autoencoders. In *International Conference on Learning Representations*, 2016. URL `http://arxiv.org/abs/1511.05644`.

A. Mondal, Y. Efendiev, B. Mallick, and A. Datta-Gupta. Bayesian uncertainty quantification for flows in heterogeneous porous media using reversible jump Markov chain Monte Carlo methods. *Advances in Water Resources*, 33(3):241 – 256, 2010.

C. P. Robert and G. Casella. *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

REFERENCES

Arash Vahdat and Jan Kautz. Nvae: a deep hierarchical variational autoencoder. In *NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, number 1650, page 19667–19679, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Ting Zhang, Tu Hongyan, Xia Pengfei, and Du Yi. Reconstruction of porous media using an information variational auto-encoder. *Transport in Porous Media*, 143(2):271–295, 2022. doi: 10.1007/s11242-022-01769-5. URL `https://doi.org/10.1007/s11242-022-01769-5`.

Huangjie Zheng, Jiangchao Yao, Ya Zhang, Ivor Tsang, and Jia Wang. Understanding vaes in fisher-shannon plane. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33: 5917–5924, 07 2019. doi: 10.1609/aaai.v33i01.33015917.

INDEX