

INVERSE ANALYSIS WITH VARIATIONAL AUTOENCODERS: A COMPARISON OF SHALLOW AND DEEP NETWORKS

Hao Wu,^{1,*} Daniel O'Malley,¹ John K. Golden,² & Velimir V. Vesselinov¹

¹Computational Earth Science, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

²Information Sciences, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

*Address all correspondence to: Hao Wu, Computational Earth Science, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA; Tel.: +1 505 667 8756, E-mail: wu_hao@lanl.gov

Original Manuscript Submitted: 11/29/2021; Final Draft Received: 1/24/2022

Inverse problems are applied to determine unknown properties by matching observational data with a physical model that often takes many parameters as input. To overcome the underconstrained nature of inverse problems and achieve good performance, an approach is presented involving regularization with a technique known as a variational autoencoder (VAE), which is trained to map a high-dimensional parameter space with a complex structure to a low-dimensional latent space with a simple structure. We apply this approach to unconditioned realizations of the parameters (heterogeneous hydraulic fields) for a hydrogeological inverse problem. Two types of hydraulic conductivity fields are used to evaluate the characterization for the different levels of heterogeneity complexity of the physical inputs. This approach keeps the computational cost of generating the training data low. The reason is unconditioned realizations neither rely on the observational data used to perform the inverse analysis nor require any groundwater flow model forward runs. In addition, this approach applies regularization on a low-dimensional latent space from the VAE and increases optimization efficiency through automatic differentiation. Furthermore, two different neural network (NN) structures are tested for their utility in using the VAE for inverse analysis. The performance of a deep, convolutional neural network strongly depends on the dimensionality of the latent space, which requires tuning. In contrast, a shallow, dense neural network provides consistently accurate characterization without tuning. Our approach evaluates the advantages of the shallow, dense neural network over the deep, convolutional one and enables future application to a wide range of inverse problems.

KEY WORDS: regularization, variational autoencoder, inverse analysis, machine learning, hydrogeology

1. INTRODUCTION

Characterization of reservoir properties, e.g., hydraulic conductivity, is of essential importance in hydrogeology, such as groundwater resources management and subsurface contaminant

remediation. Moreover, spatial heterogeneities of reservoir properties are present from pore to reservoir scale, and understanding the heterogeneity effects on fluid flow is critical (Doughty and Pruess, 2004; Jayne et al., 2019). Inverse analysis aims at determining unknown parameters (heterogeneous) by matching observational data with predictions from a forward model calculation that takes the parameters as input (Linde et al., 2015; Zhou et al., 2014). For example, our study focuses on estimating an aquifer's spatially heterogeneous hydraulic conductivity parameters based on observational hydraulic heads at a sparse collection of points within the flow domain. However, inverse problems are often inherently difficult and computationally expensive (Carrera et al., 2005). In addition, another challenge of inverse problems is the feature of being underconstrained or ill-posed, which results in many viable solutions (Zagst and Pöschik, 2008). There are many feasible methods for inverse problems, such as the traditional geostatistical approach (Kitanidis and Lee, 2014; Lee and Kitanidis, 2014) and more recent machine learning (ML) inverse techniques (Sinha et al., 2020; Zhou et al., 2019). In this study, we are contributing the recently-developed VAE-based inverse methods (Laloy et al., 2017; Lopez-Alvis et al., 2021), which combine some of the strengths of the traditional approaches and the ML approaches.

Different inverse analyses have specific advantages and drawbacks. The stochastic geostatistical inversion approach has been widely applied to estimate unknown parameters (Tartakovsky et al., 2021), and naming one of the recent developed for hydrogeology is the principal component geostatistical approach (PCGA). PCGA processes the inverse analysis through the principal components of the covariance matrix, which overcomes the challenge of high computational cost for large-scale systems (Kitanidis and Lee, 2014; Lee and Kitanidis, 2014) by reducing the dimension parameter space. Even with the low computational cost improvement of PCGA, it can struggle with highly complicated subsurface structures. Recent advances in ML provide an opportunity to conduct inverse analysis in the geological area (Barajas-Solano and Tartakovsky, 2019; Kadeethum et al., 2021) without using a physics model after the initial training. This technique can produce excellent results but comes with a steep up-front cost to generate the training data. Often, the cost of generating the training data is greater than performing the inverse analysis, which can be self-defeating. The underlying limitation of these approaches is that each training data point needs to run through the forward model. Meanwhile, these methods are sometimes seen as being untrustworthy since no physical model calibration is involved. The works of He et al. (2020), Tartakovsky et al. (2020), Geneva and Zabarar (2020), and Zhu et al. (2019) have testified to the physical calibration effect during inverse analysis, which improves the final modeling results. More importantly, the model's success depends on the appropriate selection of ML structures. The "black box" has been used to describe this process based on the uncertainty of the results from different structures (McGovern et al., 2019). To overcome the challenges discussed above, we propose an approach to regularization and optimization by utilizing a variational autoencoder (VAE). This approach combines some of the strengths of traditional inverse methods like the geostatistical approach with more recent ML-based techniques. It has a low up-front cost (similar to methods like PCGA) but can accurately recreate complex subsurface structures (similar to ML methods with a high up-front cost).

A VAE is a machine learning methodology that transforms a high-dimensional object into a set of low-dimensional latent variables with a simple structure and back again. "Encoding" characterizes the mapping process from the high-dimensional object to the latent variables, and oppositely, "decoding" depicts the mapping from latent variables back to the high-dimensional object (Mo et al., 2019). The encoding and decoding processes are ideally the inverses of one another (i.e., the output of the VAE is the same as the input). After VAE training, the latent

variables create an approximately normally distributed latent parameter space that conforms to the existing conceptual model, as exemplified by the training data. A benefit of VAE training is that during optimization, the predicted parameters which need to fit the observational data can be performed in a lower-dimensional space for computational efficiency increase.

Regularization is a technique that enhances the performance of inverse problems by adding a term to the objective function, where a small objective function value is indicative of good performance. Specifically, regularization seeks to impose additional desired features on the solution, such as smoothness, to incorporate prior information (i.e., different from the observations that will be used in the inverse analysis). In the hydrogeologic context that we consider, regularization often smooths out these fields by encouraging the fields' gradient to be small for the heterogeneous fields to be parameterized (Bredies et al., 2010; Rudin et al., 1992). Another commonly used regularization technique is Tikhonov regularization (Franklin, 1974; Tikhonov, 1963).

Our method aims to achieve an easy and fast inverse analysis for complex heterogeneous hydrogeologic reservoir properties and provide highly accurate inverse results. Specifically, our method employs the combined application of VAE, regularization, and automatic differentiation. For VAE training, there are far fewer latent variables than input variables for training and this feature is identified as dimensionality reduction (Laloy et al., 2017). Dimensionality reduction reduces the computational cost when fast gradients are unavailable through automatic differentiation or adjoint methods. Furthermore, the VAE training requires no forward model for dataset generation. An additional benefit of our approach is the application of regularization, which encourages the inverse result to conform to prior knowledge (such as an understanding of a site's geology). Besides these two features, we explored the benefits of automatic differentiation for automatic gradient calculation aiming for fast optimization. In the work of Xu et al. (2021), the application of automatic differentiation through the neural network during inverse analysis shows its huge potential for parametric calibration of geological problems. Automatic differentiation negates the speed benefits of dimensionality reduction, but provides an even greater acceleration in its place. We call our approach RegAE since it does *Regularization with a variational AutoEncoder* for an easy, fast, and accurate inverse problem.

Different neural network (NN) architectures provide different results when used within RegAE's VAE. For example, a deep neural network containing several hidden layers is designed for better feature capture than a shallow network. In addition, the convolutional layer extracts information from fewer parameters through a data filter, forcing the input values to share and the network to learn the relation to the output. In contrast, the dense layer denotes all the inputs through the function (Wu et al., 2021). A convolutional layer is representative for analyzing spatial features from image inputs with high resolution. While considering the specialty of the convolutional layer of reaching a balance between computational efficiency and neglecting "redundant" data (O'Shea and Nash, 2015), it is essential to evaluate the performance of the convolutional layer for various training data. In our study, we compare two different neural network architectures in the inverse analysis: shallow, dense versus deep, convolutional, to test their performance stability and accuracy in the context of hydrologic inverse analysis with a VAE. In addition, we include a geostatistical method compared with the two different neural network architectures to better understand the landscape of methods for hydrogeologic inverse modeling.

In the following content, we describe the detailed method of the RegAE model, the results of RegAE solving a hydrogeologic inverse problem, and the benefits and improvement of this approach in Sections 2, 3, and 4, respectively. Finally, we present our conclusion about the application of the RegAE model in Section 5.

2. METHODS

The RegAE modeling process starts with generating the input data, which are random realizations of the hydraulic conductivity field. These realizations should incorporate prior knowledge of the hydraulic conductivity field. The second step is VAE training, and the low-dimensional latent variables characterize the distribution of high-dimensional input features after training. The final step is performing the inverse analysis combining the VAE and a physical model for gradient-based optimization. The detailed RegAE workflow is illustrated in Fig. 1. In this study, \mathbf{p} and $\hat{\mathbf{p}}$ are used to represent a physical parameter field in vector form; \mathbf{z} shows the latent variables from the VAE; \mathbf{h} and $\hat{\mathbf{h}}$ represent a vector of observations for inverse analysis; $e(\cdot)$ and $d(\cdot)$ denote the encoder and decoder, respectively; $f(\cdot)$ is the forward model. In addition, n_p , n_z , and n_h are the number of components of \mathbf{p} , \mathbf{z} , and \mathbf{h} , respectively. One of the features of this approach here is that the number of latent variables is much less than the number of physical parameters; i.e., $n_z \ll n_p$. We employ the simulation of single-phase flow in heterogeneous porous media as the forward model for inverse analysis. In this study, \mathbf{p} is the hydraulic conductivity of the porous medium, and the observations, \mathbf{h} , are of the fluid pressure at various locations throughout the domain.

2.1 Data Generation

A sequence of realizations \mathbf{p}_i for $i = 1, 2, \dots, N$ is generated as input data during the data generation. These realizations used for VAE training should be independent of the observational data, $\hat{\mathbf{h}}$, used for calibration by the inverse analysis, but should include prior knowledge of the

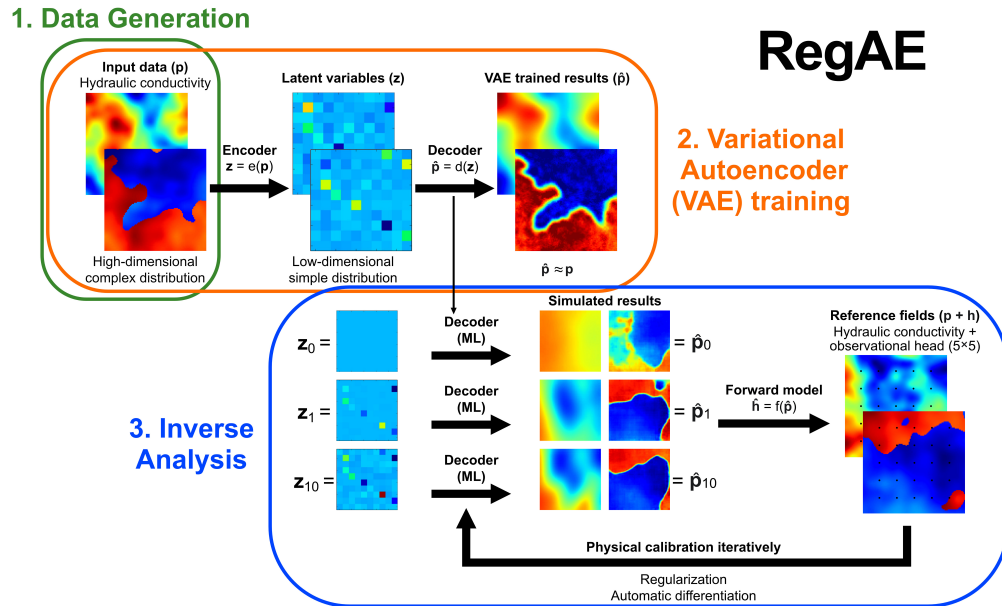


FIG. 1: The RegAE inverse method leverages a VAE for dimensionality reduction and an automated approach to regularization. Calibration is performed on the low-dimensional latent variables (\mathbf{z}) rather than the high-dimensional medium properties, making gradient computations efficient. The simple distribution of the latent variables makes regularization simple.

site. A sufficient number of realizations about hydraulic conductivity distribution must be generated to train the VAE effectively.

This work utilizes a statistical model of \mathbf{p} that results in the hydraulic conductivity field as input data. For the examples in this manuscript, we study a $100 \text{ m} \times 100 \text{ m}$ subsurface aquifer with a unit thickness, where the resolution of pixels for VAE training has been set up for 100×100 . The pixels are used as the input of the neurons for the VAE training. To test the effects of the heterogeneity of hydraulic conductivity distribution on training results and estimate the applicability of different neural network architectures, two different levels of heterogeneity complexity are generated for the parameter field. First, a multivariate Gaussian model has been utilized to create a heterogeneous \mathbf{p} distribution covering the entire study area, which we call a Gaussian field. Increasing the complexity of the heterogeneity, which makes the inverse problem more challenging, the second type of hydraulic conductivity field has two hydrogeologic facies with distinct properties to represent a higher level of heterogeneity or more realistic characterization. In generating the second type of parameter field, three random fields are generated for each realization, all of which have a multivariate Gaussian structure. Two of the fields represent each of the facies and the third field is used to indicate which of the facies is present at a given location. This gives these fields a bimodal hydraulic conductivity distribution.

To generate the multivariate Gaussian fields, we use the `GaussianRandomFields.jl` package for the Julia programming language (Robbe, 2020). The parameters used for the three statistical models are given in Table 1. Each of these models is characterized by a specific mean, variance, and correlation length. Conductivity 1 is implemented in Gaussian field generation, creating one hydrogeological face with continuous heterogeneity, shown in Fig. 2(a). Except for the mean value, the two models (Conductivity 1 and Conductivity 2) have the same variance and correlation length. As a result, these two models have the same characteristic appearance, but one has a higher value and the other a lower value. For the bimodal fields, these two models are applied to generate the situation of two hydrogeological facies and create a comparable higher hydraulic conductivity subsection in the research domain. Meanwhile, inside each subsection of the bimodal type, the statistical models also yield heterogeneous distribution, respectively. As a result, the bimodal type adds complexity to the hydraulic conductivity distribution shown in Fig. 2(d), which supports the study about the effects of dataset complexity on inverse problems. The third model (“Split”) determines which of the two hydrogeologic facies is present at a given location. The split model has the same mean and variance value as model 2 but four times larger correlation length to characterize a more smoothly varying hydraulic conductivity distribution.

For more specific discussion about the bimodal field generation, \mathbf{p} , all the three statistical models are combined. First, a random number, q , is drawn uniformly between $1/4$ and $3/4$. At each point in space, the comparison between q and the “Split” field determines the value of the hydraulic conductivity field. Specifically, if the value of the “Split” field is below the q th quantile, then hydraulic conductivity draws from the “Conductivity 1” model, and “Conductivity

TABLE 1: The statistical parameters used to generate the hydraulic conductivity data

Model name	Mean (m/s)	Variance (m^2/s^2)	Correlation length (m)
Conductivity 1	10^{-5}	1.0	50.0
Conductivity 2	10^{-8}	1.0	50.0
Split	10^{-8}	1.0	200.0

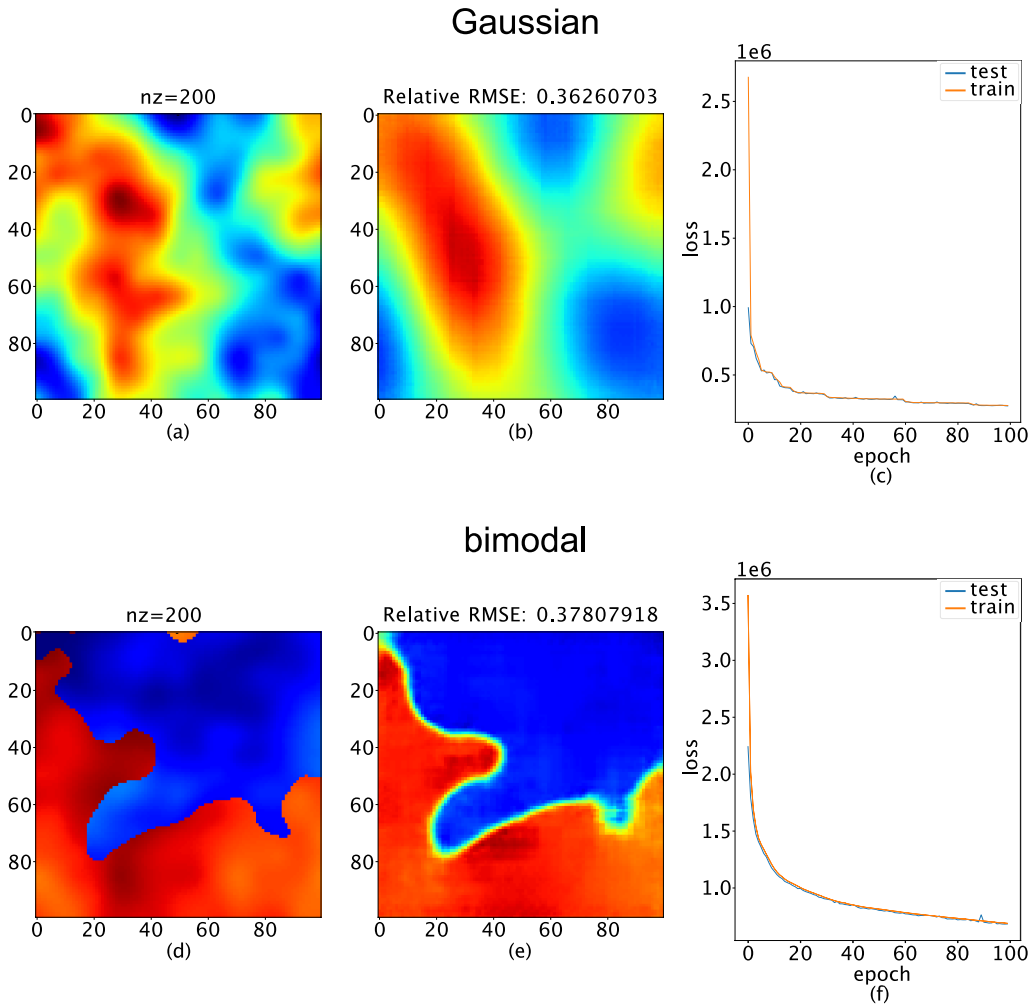


FIG. 2: VAE training results for $n_z = 200$ of Gaussian and bimodal fields. From left to right, the subfigures (a), (d) demonstrate the reference fields; (b), (e) show VAE prediction results; and (e), (f) demonstrate the loss function with epochs for train and test sets. The relative RMSE errors above predicted figures represent the difference between related reference and prediction figures.

2” otherwise. As a result, a fraction q of the domain comes from the “Conductivity 1” field, and a fraction $1 - q$ comes from the “Conductivity 2” field, which generates a bimodal distribution.

The optimal results of the inverse analysis are shown for the two hydraulic conductivity distribution types, Gaussian and bimodal. Afterward, different sizes of datasets are explored with $N = 10^4$, $N = 10^5$, and $N = 10^6$ realizations. We found some improvement increasing from $N = 10^4$ to $N = 10^5$, but rising to $N = 10^6$ did not provide further significant improvements. The number of realizations needed to train the VAE will vary depending on the complexity of both the distribution of \mathbf{p} and the VAE. Typically, the more complex either is, the larger N will need to be. We picked the dataset $N = 10^5$ for the following training and results representation, balancing the complexity of value distribution and computational efficiency. Eighty percent of

the dataset has been applied for VAE training, and the other 20% is utilized as the test set. The VAE training results are presented for both train and test sets.

2.2 Variational Autoencoder

VAEs (Kingma and Welling, 2013) are generative machine learning models with neural network architecture, characterized as unsupervised learning. There is the widespread application of VAEs at image data, such as handwriting and facial characterization (Doersch, 2016), and, in addition, it worked as well in the hydrogeological properties distribution in this study. There are two parts of VAEs: an encoder network $e(\cdot)$ and a decoder network $d(\cdot)$. The encoder maps a high-dimensional space (such as pixels in an image) onto a smaller space, and this feature is described as dimension reduction. In particular, the smaller space contains a certain number of latent variables z , and we usually call it the latent space. The decoder network works in the reverse way as the encoder, which maps the latent variables back to their original higher-dimension space. The encoder and decoder networks are trained in tandem such that $d(e(\cdot)) \approx \text{id}(\cdot)$.

VAE is one type among all the autoencoders, with a unique feature that the encoder outputs a probability distribution of the latent variables rather than concrete values. The VAE produces a collection of means $\mu(X)$ and standard deviations $\sigma(X)$ for a given input X . The main benefit of this approach is that it creates a locally continuous latent space, which describes all the possible variances of the input values to the decoder. Compared to a traditional autoencoder (AE), VAE yields better training results resulting from fewer input gaps. Another beneficial feature of VAE is creating a globally continuous latent space approximately following a normal distribution with mean 0 and covariance matrix being the identity. As a result, one term, called Kullback–Leibler (KL) divergence (Kullback, 1997), has to be considered regarding the difference between the distribution of the latent variables during training and an $N(0, I)$ distribution. For better characterization of the input values, the KL divergence tends to be minimized during VAE training. More specifically, when calculating the loss function, we include a term for the error in the reconstruction after the encoding and decoding process and a term for the KL divergence. This globally continuous nature of the latent space in a VAE is critical for the inverse analysis for an easy global minimum finding through the entire latent space of the optimization algorithm.

In the context of inverse analysis and regularization, VAEs thus address two significant issues. First, there are far fewer latent variables than input variables, significantly reducing computational overhead for calculating gradients. This nonlinear dimensionality reduction provides the potential for model accuracy improvement resulting from less misleading data noise. Second, the space of latent variables follows a normal distribution, thus dramatically simplifying the regularization process. The normal distribution of latent variables ensures hydraulic conductivity distribution is captured efficiently through VAE training.

We employ the VAE relatively simply and implemented in the Flux (Innes, 2018) machine learning framework. We explored VAEs with $n_z = 25$, $n_z = 50$, $n_z = 100$, and $n_z = 200$. In each case, the number of hidden neurons was 125, 250, 500, and 1000, respectively. We found that the effects of n_z to VAE results are dependable, while the difference is minimal. The specific observation is presented in Section 3.

2.3 Inverse Analysis

Our inverse analysis then combines the VAE and a forward model with an objective function. As shown in Fig. 1, step 3, the decoder's output (simulated hydraulic conductivity) goes through

a forward model for hydraulic head calculation. Then the results are calibrated iteratively using the observational data. We applied this inverse analysis to estimate the hydraulic conductivity of three synthetic aquifers for method evaluation for both the two Gaussian and bimodal types. These six referenced “true” hydraulic conductivity fields [depicted in Figs. 3(a)–3(c) and 4(a)–4(c)] were generated similarly to the training data for Gaussian and bimodal types, respectively. Of course, these reference conductivity fields were not part of the training set used to train the VAE. However, they are generated using the same conceptual models of aquifer heterogeneity. The observations used to inform the inverse analysis consist of the hydraulic head and hydraulic conductivity on a 5×5 regular grid within the domain. The black dots in Figs. 3(a)–3(c) represent the positions of the observation points. The forward model includes a steady-state flow simulation that produces predictions of the hydraulic head given a hydraulic conductivity input. The hydraulic head observations were obtained using the model with the reference field as the input hydraulic conductivity. At the same time, there should be observations from drilling wells or boreholes when solving realistic problems. The hydraulic conductivity observations were obtained directly from the reference field. The forward model predictions for the hydraulic conductivity are obtained directly from the decoder.

The optimization problem of inverse analysis is formulated in terms of the latent variables, \mathbf{z} :

$$f(\mathbf{z}) = [h(d(\mathbf{z})) - \hat{\mathbf{h}}]^T \Sigma_h^{-1} [h(d(\mathbf{z})) - \hat{\mathbf{h}}] + [\mathbf{z} - \bar{\mathbf{z}}]^T \Sigma_z^{-1} [\mathbf{z} - \bar{\mathbf{z}}], \quad (1)$$

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} f(\mathbf{z}), \quad (2)$$

where Σ_h is the covariance matrix of the observations, Σ_z is the covariance matrix for the latent variables, $\bar{\mathbf{z}}$ is the mean of the latent variables, and we call $f(\mathbf{z})$ the objective function. In many inverse analyses, it is assumed that Σ_h is diagonal—effectively asserting that observational noises are uncorrelated. We make this assumption and denote the i th diagonal element of Σ_h as $\sigma_{h,i}^2$. The VAE seeks to make $\bar{\mathbf{z}} = \mathbf{0}$ and $\Sigma_z = I$, but we estimate these from the test data to account for imperfections in the training process. After an estimate of the latent variables, $\hat{\mathbf{z}}$, is obtained, the parameter field can then be estimated with the decoder,

$$\hat{\mathbf{p}} = d(\hat{\mathbf{z}}). \quad (3)$$

We utilize automatic differentiation to compute $\nabla f(\mathbf{z})$. The automatic differentiation enables us to compute the gradient efficiently and is equivalent to using an adjoint method to compute the gradient. The DPFEHM subsurface flow and transport code (O’Malley et al., 2020), which is compatible with the Zygote.jl automatic differentiation library (Innes et al., 2019), simulates the groundwater flow. The application of automatic differentiation makes the computation of the gradients efficient—gradients can be computed at about the cost of a single forward model run. When automatic differentiation is used, the dimensionality reduction of VAE does not result in any acceleration, and the automatic differentiation accelerates the gradient calculations more than the dimensionality reduction. Therefore, automatic differentiation or adjoint methods should be used when available, but they are not required. Using the objective function from Eq. (1) makes the regularization process easier for highly parameterized models (i.e., n_p is very large) because n_z is small. Since the latent variables approximately follow an $N(\bar{\mathbf{z}}, \Sigma_z)$ distribution, the natural regularization is given by the rightmost term in Eq. (1). The difficulties that are typically encountered in the formulation and computation of the regularization component of the objective function were essentially transferred to the data generation and VAE training

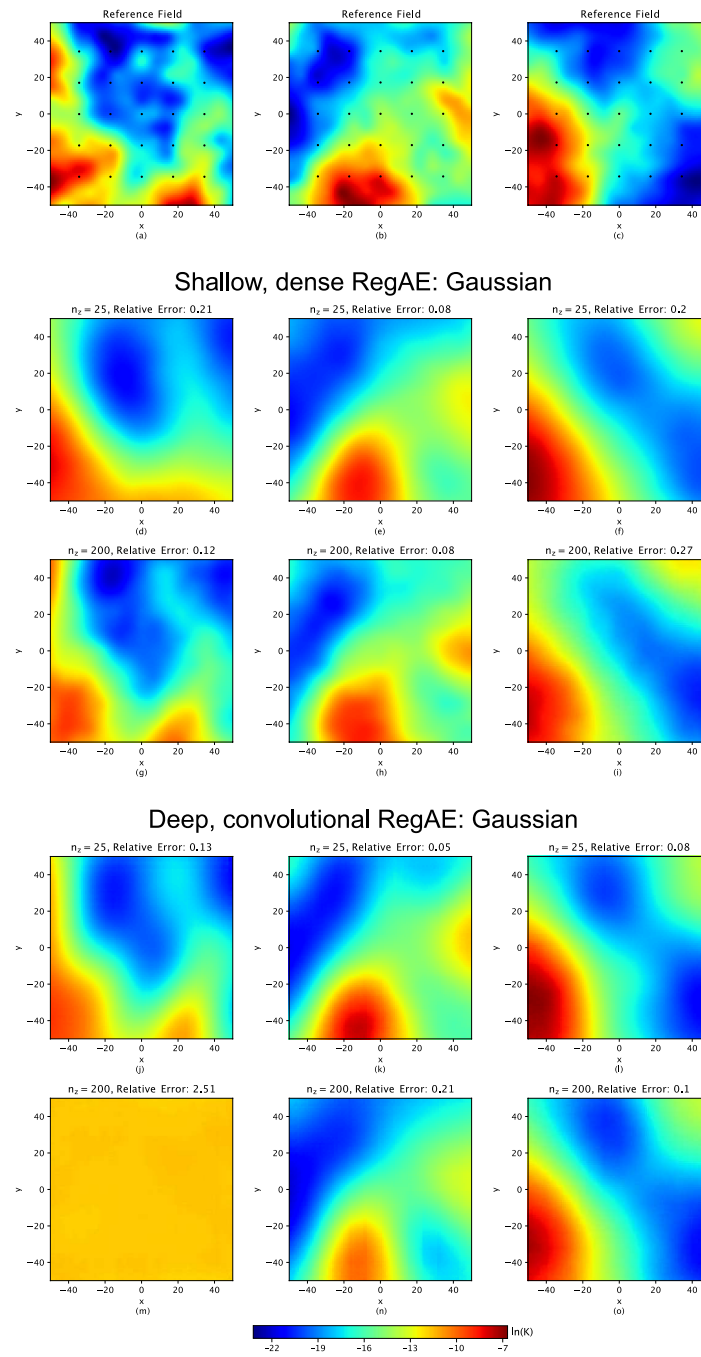


FIG. 3: Three Gaussian reference conductivity fields are shown in subfigures (a)–(c). The black dots in the first row represent the 5×5 regular grids for observation points. The corresponding inverse results of shallow, dense RegAE and deep, convolutional RegAE for $n_z = 25$ and $n_z = 200$ with seed = 1 are in subfigures (d)–(o). The hydraulic conductivity shows on the \ln scale. The relative errors above inverse figures represent the difference between related reference and simulation figures.

components of our approach. Our results show that even a simple VAE (such as the ones using a shallow, dense network) can be used effectively to construct an appropriate regularization term.

To optimize the objective function, we apply an existing method for gradient-based optimization, which is the limited-memory Broyden–Fletcher–Goldfarb–Shanno (Liu and Nocedal, 1989) (L-BFGS) method with a Hager–Zhang line search (Hager and Zhang, 2005). We use the Optim.jl (Mogensen and Riseth, 2018) software package for this. In all cases, to start the inverse analysis, the initial guess of \mathbf{z}_0 is set to be 0.

Only the decoder step of VAE is included in the inverse analysis, and it also involves the ML technique. We test two different neural network architectures in the VAE during inverse analysis: the first is a shallow, dense network, and the other is a deep, convolutional network. There is only one connection layer for the shallow, dense network, while for the deep, convolutional network, we implement three convolutional layers. We name them shallow; dense RegAE, and deep, convolutional RegAE, respectively. Beyond the discussion of different architectures in the inverse analysis, we compared other inverse analyses as well. Traditional geostatistical inverse problems require efficient computational methods to calculate regularization components in the objective function for large-scale problems. Further, the creativity in formulating the conceptual model in the geostatistical approach is usually limited to formulating a variogram. As a result, ML inverse models (Kadeethum et al., 2021) often outperform traditional inverse approaches at results accuracy. Our work brings the advantages of these ML inverse approaches into a more traditional physical calibration involved inverse analysis framework. When comparing our model to the pure ML inverse analysis frameworks, it does not require a forward model run to train the ML model, only realizations of the model parameters.

3. RESULTS

This study provides an inverse analysis model for hydraulic conductivity characterization with regularization and variational autoencoder. Meanwhile, two types of neural network architectures built in the inverse step have been evaluated based on the modeling results for Gaussian and bimodal hydraulic conductivity fields. In addition, a traditional geostatistical inverse analysis method, PCGA, has been compared with our RegAE model as a reference to illustrate the benefits and advantages of using a VAE. Figure 2 presents the results of VAE training for Gaussian and bimodal fields. Figures 3 and 4 are the results of two neural network architectures (shallow, dense and deep, convolutional) of the inverse analysis step for Gaussian and bimodal fields. The convergence results of the four models mentioned above are shown in Fig. 5. We conducted 300 inverse analyses for each field type (Gaussian or bimodal) and each VAE architecture (shallow, dense or deep, convolution) to test the performance of the two network structures for Gaussian and bimodal fields, and the statistical results of cumulative distribution function (CDF) for relative error are shown in Figs. 6–8. For PCGA, the modeling results of Gaussian and bimodal fields are demonstrated in Figs. 9 and 10, respectively.

3.1 VAE Results

VAE training (step 2 in Fig. 1) provides an application of low-dimensional latent space that represents the features of a high-dimensional parameter field. Even though we test two VAEs with different neural network architectures, the VAE training is identical for different runs, except the seed value for the random number generator (the training of a VAE involves generating random numbers). Three seed values have been applied to test the stability of the VAE training (of the

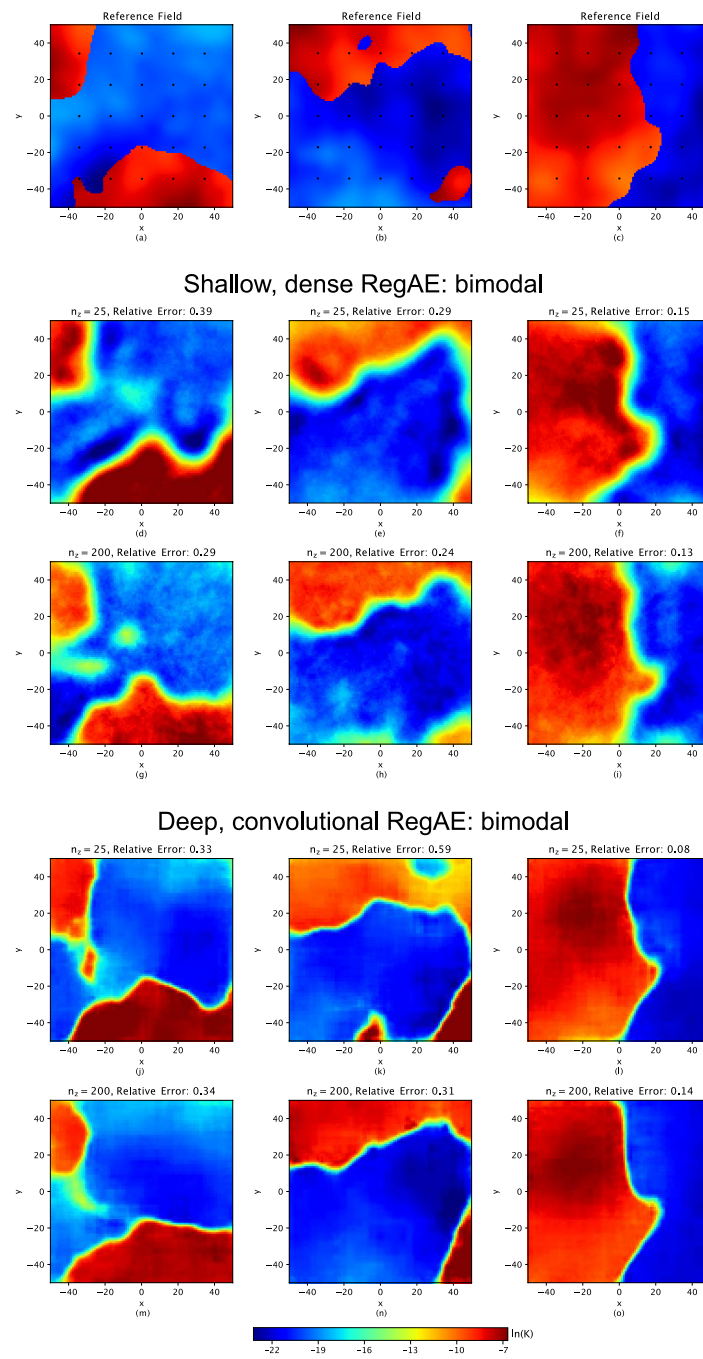


FIG. 4: Three bimodal reference conductivity fields are shown in subfigures (a)–(c). The black dots in the first row represent the 5×5 regular grids for observation points. The corresponding inverse results of shallow, dense RegAE and deep, convolutional RegAE for $n_z = 25$ and $n_z = 200$ with seed = 1 are in subfigures (d)–(o). The hydraulic conductivity shows on the \ln scale. The relative errors above inverse figures represent the difference between related reference and simulation figures.

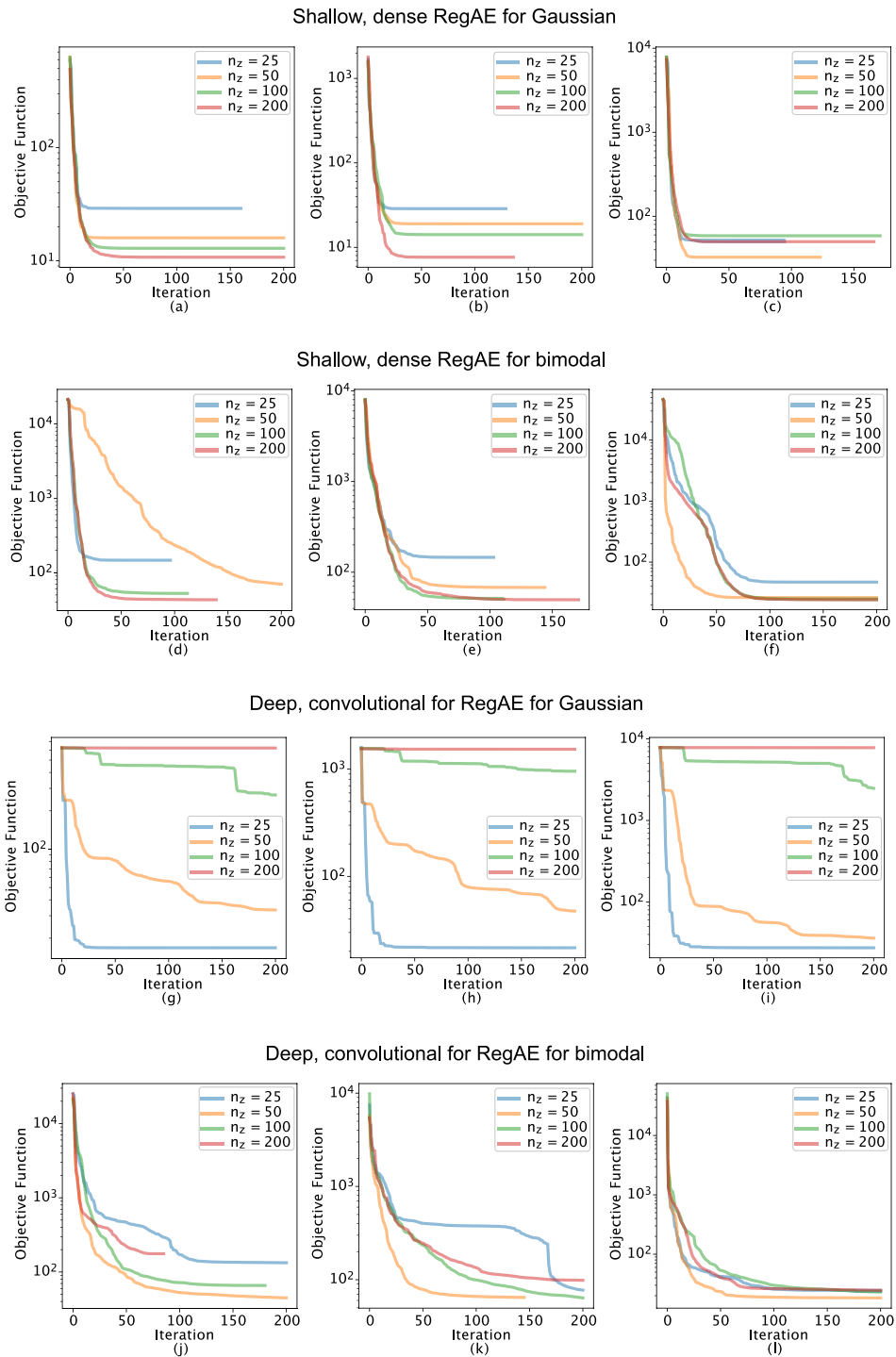


FIG. 5: The convergence of the inverse analysis is shown for different values of n_z for shallow, dense RegAE and deep, convolutional RegAE of Gaussian and bimodal fields

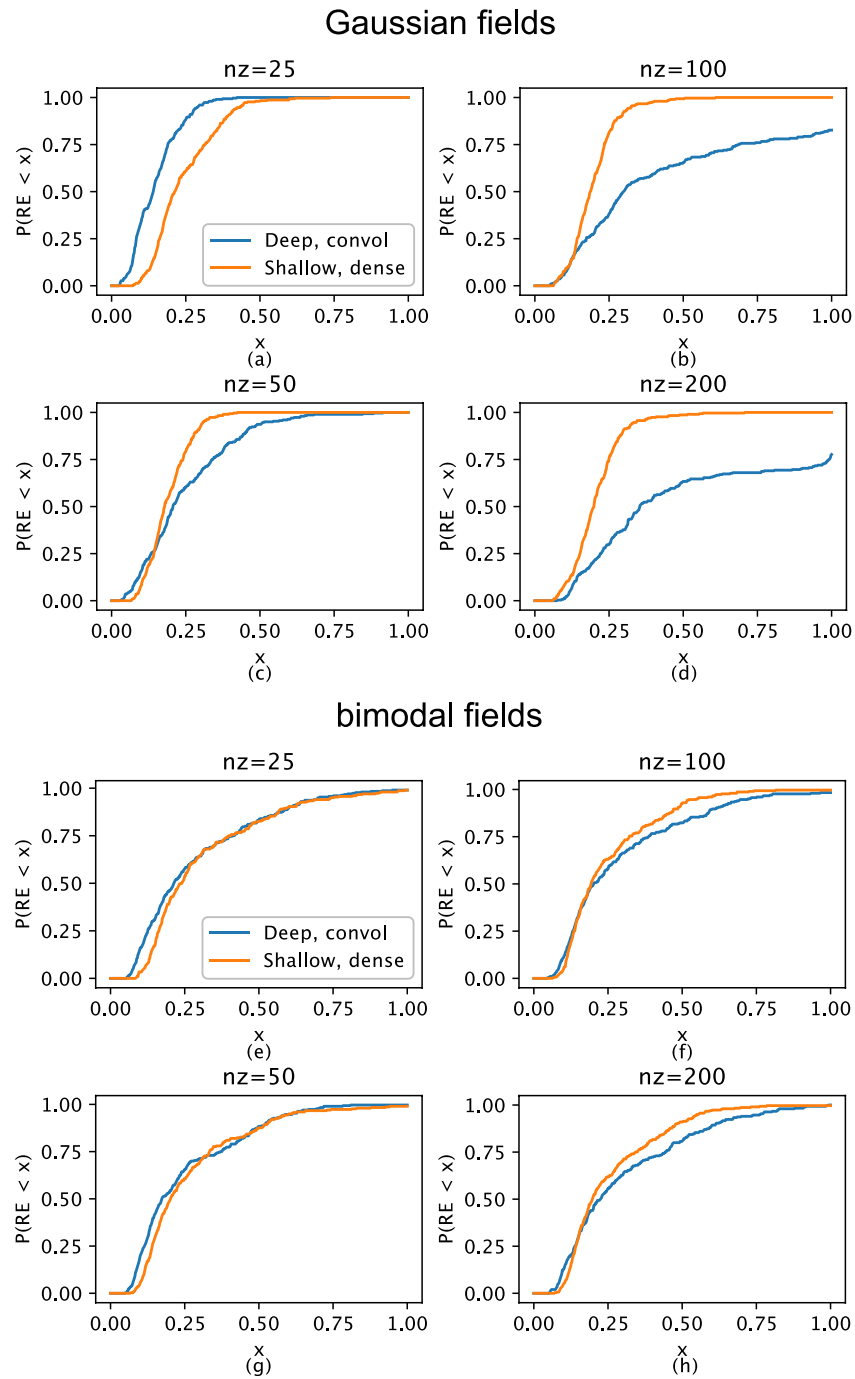


FIG. 6: Cumulative distribution function (CDF) of relative error for inverse analyses of all 300 reference fields with seed = 1, 2, and 3. The results are for two deep, convolutional and shallow, dense machine learning structures during the inverse analysis. The comparisons of the two structures are presented for four n_z options separately. The results are presented for Gaussian and bimodal fields.

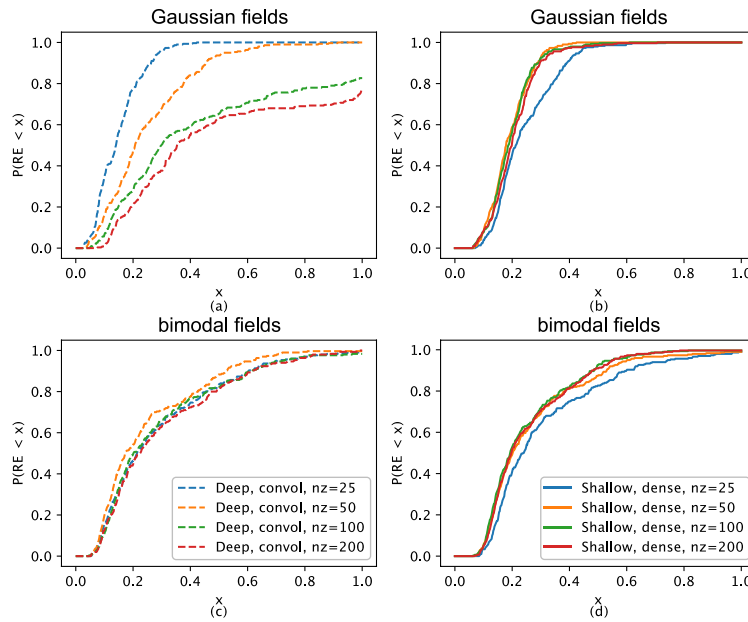


FIG. 7: Cumulative distribution function (CDF) of relative error for inverse analyses of all 300 reference fields with seed = 1, 2, and 3. The results are for two deep, convolutional and shallow, dense machine learning structures during the inverse analysis. The comparisons are presented for four n_z options. The results are presented for Gaussian and bimodal fields.

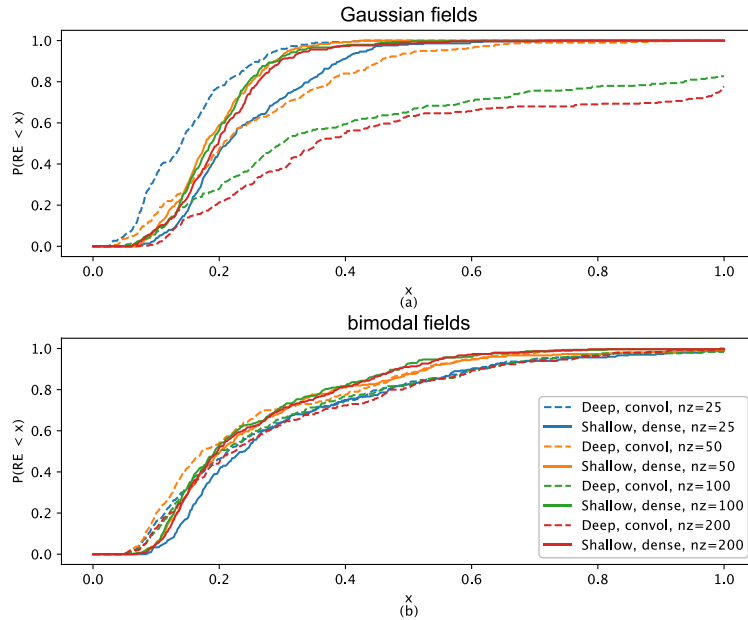


FIG. 8: Cumulative distribution function (CDF) of relative error for inverse analyses of all 300 reference fields with seed = 1, 2, and 3. The results are for two deep, convolutional and shallow, dense machine learning structures during the inverse analysis. The comparisons are presented for two structures with four n_z options. The results are presented for Gaussian and bimodal fields.

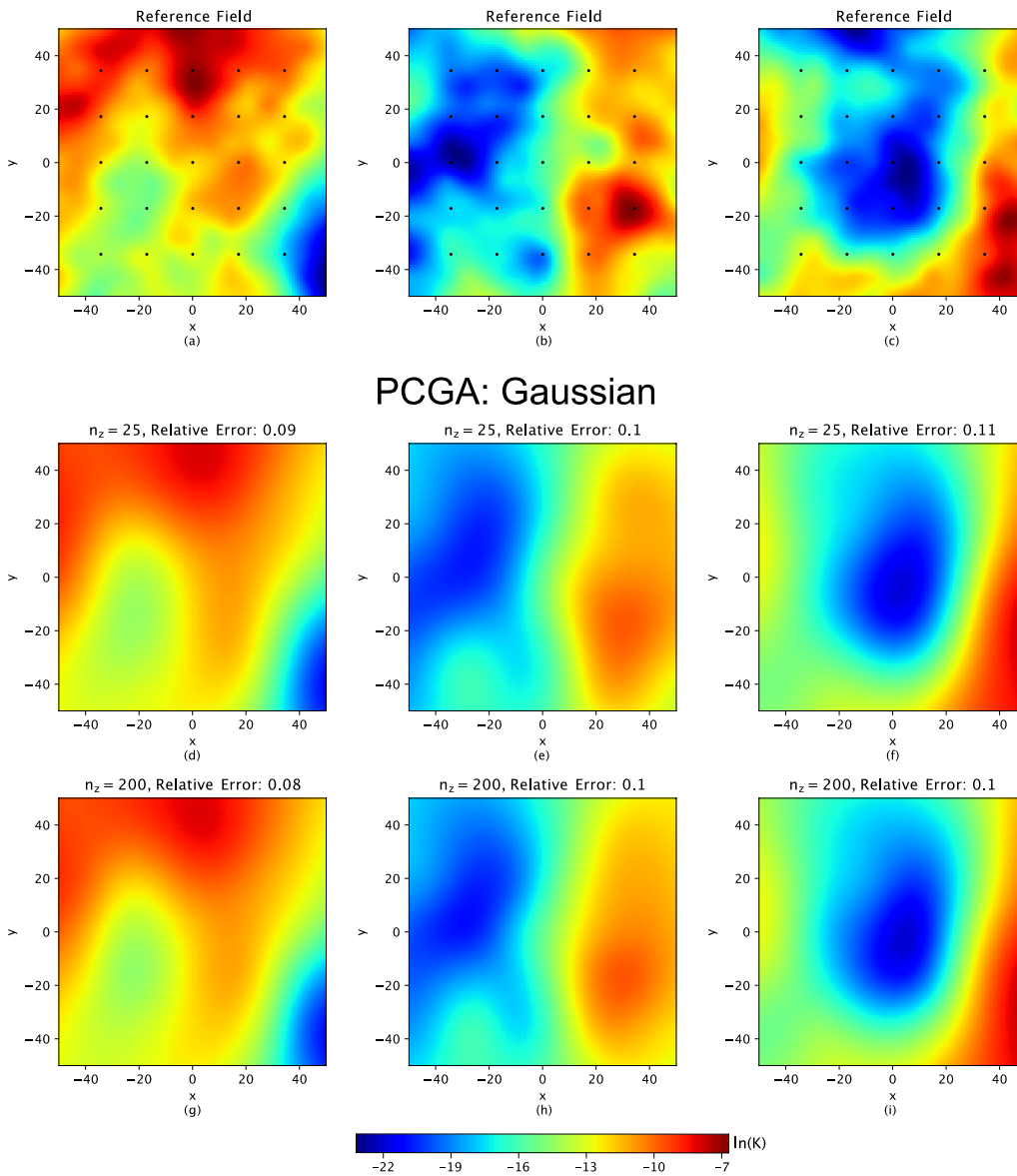


FIG. 9: Three Gaussian reference conductivity fields are shown in subfigures (a)–(c). The corresponding inverse results of PCGA for $n_z = 25$ and $n_z = 200$ are in subfigures (d)–(i). The hydraulic conductivity is shown on the ln scale.

300 inverse analyses, 100 are performed with each seed). The VAE training lasts for 100 epochs, and we use the VAE that performs the best on the test set to avoid overfitting. In Fig. 2, the left column subfigures (a), (d) are the input hydraulic conductivity fields. The middle column subfigures (d), (e) are the predicted outputs of VAE training for Gaussian and bimodal fields. The loss function of the train and test sets are shown in the subfigures (c), (f) in the right column. The relative error above the predicted figures used to measure the difference between input and

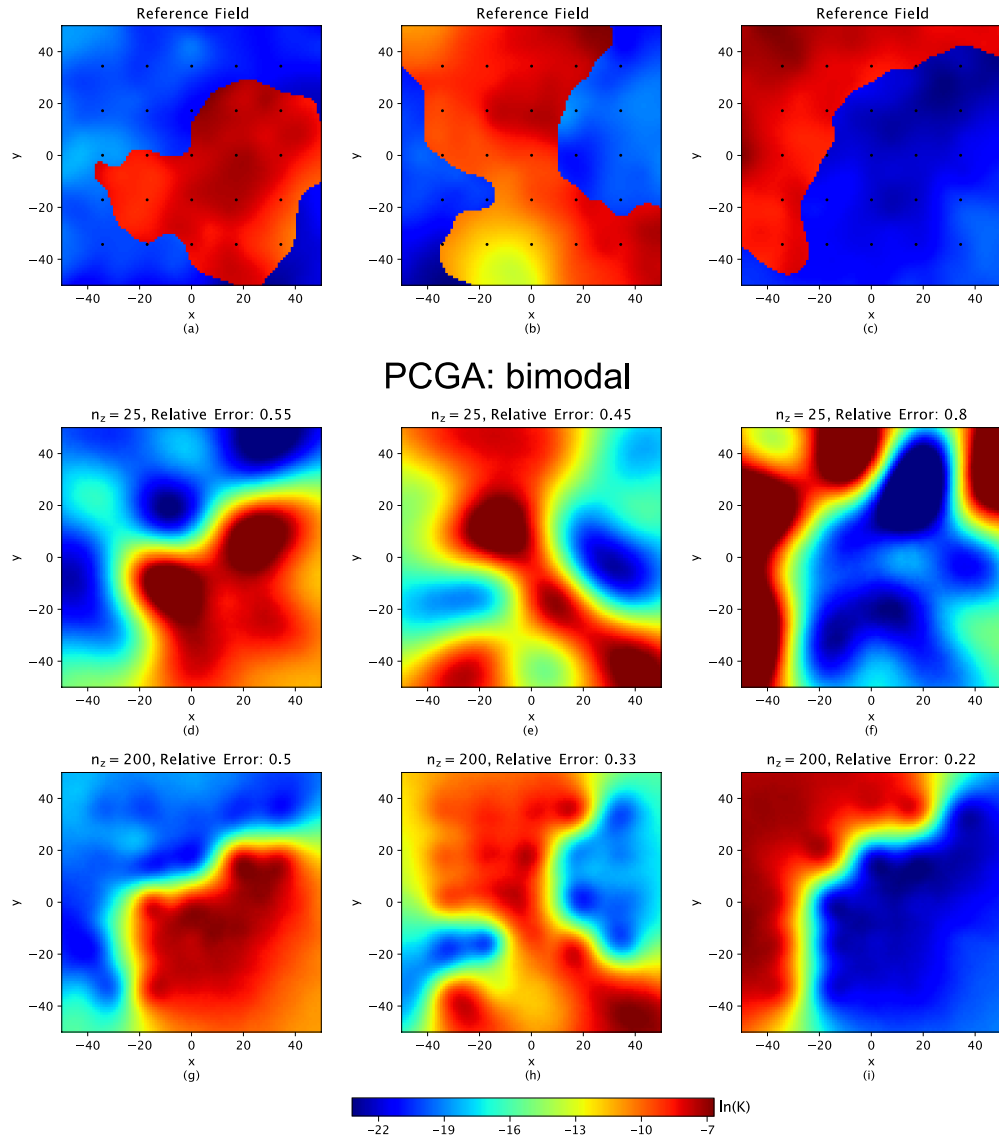


FIG. 10: Three bimodal reference conductivity fields are shown in subfigures (a)–(c). The corresponding inverse results of PCGA for $n_z = 25$ and $n_z = 200$ are in subfigures (d)–(i). The hydraulic conductivity is shown on the $\ln(K)$ scale.

output data is the root relative squared error, defined as

$$\sqrt{\frac{\sum(\mathbf{p}_t - \mathbf{p}_v)^2}{\sum(\mathbf{p}_v)^2}}, \quad (4)$$

where \mathbf{p}_t is the “true” field, and \mathbf{p}_v is the predicted field.

As shown in Fig. 2, the similarity between the predicted fields and the reference fields and the minor relative errors leads us to believe that the VAE training successfully captures the

distribution of hydraulic conductivity for both Gaussian and bimodal fields. Regarding the two fields, the VAE training of Gaussian fields has minor relative errors than bimodal fields, which can be identified as the easier parameter distribution (Gaussian) leading to better training results. In addition, the comparable small decrease at loss function implies the hardness of VAE training for bimodal fields compared to Gaussian fields even though the loss function was identified as a continuous decline through 100 epochs.

3.2 RegAE Results

The inverse analysis results of the RegAE model are shown in Figs. 3 and 4 for Gaussian and bimodal fields, respectively. We also present the results of the two neural network architectures: shallow, dense and deep, convolutional. The first rows of Figs. 3 and 4 are the three “true” reference hydraulic conductivity fields for Gaussian and bimodal fields. The black dots illustrate the 25 observational points for physical calibration. To test the stability of the VAE training, we launched different runs with three values of seed. In Figs. 3 and 4, the results come from runs with seed = 1, 2, and 3, and we only present the runs when $n_z = 25$ and 200 (the smallest and largest values of n_z). The results for all seed options and n_z values are demonstrated in supplementary figures from S1 to S12. On top of each simulated result is the relative error, which measures how close the inverse result is to the reference field and is defined as

$$\frac{\|\mathbf{p}_r - d(\hat{\mathbf{z}})\|^2}{\|\mathbf{p}_r - \bar{\mathbf{p}}_r\|^2}, \quad (5)$$

where \mathbf{p}_r is the reference field, and $\bar{\mathbf{p}}_r$ is the mean of the reference field.

The inverse analysis results of the shallow, dense RegAE model for Gaussian and bimodal fields are shown in Figs. 3 and 4, respectively. The broad similarity between the reference fields and simulated results for Gaussian fields implies that the shallow, dense RegAE captures the salient aspects with a maximal relative error of 0.27 in these cases. Even if the relative error is comparably higher than those from Gaussian fields, the phenomena lie in the higher complexity of bimodal fields. The averaged error is 0.25, which is less than the maximum error of Gaussian fields and good performance among inverse problems. Similarly, the shallow, dense RegAE model characterizes the features of bimodal fields as well. Generally, having n_z larger than 25 shows improved results with lower relative error, concluding the negative effect of making n_z too small. However, increasing beyond $n_z = 50$ does not significantly improve the results. Among different runs in supplementary figures (S1, S2, S5, S6, S9, S10), shallow, dense RegAE provides consistent modeling behavior and proves its stability at good results quality. In conclusion, shallow, dense RegAE turns out to be a good application for different types of heterogeneous hydraulic conductivity fields based on having consistently good results for both the Gaussian and bimodal fields.

Compared to shallow, dense RegAE, deep, convolutional RegAE tells a surprisingly different story based on the large variability of final results. Specifically, even if deep, convolutional RegAE achieves good performance for $n_z = 25$ with an averaged error of 0.09, one result of $n_z = 200$ indicates a failure with the highest relative error reaching over 200% for Gaussian fields. This illustrates that increasing n_z can result in information loss, which leads to the needed n_z selection. Meanwhile, deep, convolutional RegAE also shows its insatiability considering various modeling results with a wide relative error range from 0.08 to 0.59 for bimodal fields. The similar behavior of different results for deep, convolutional RegAE is shown in the supplementary figures (S3, S4, S7, S8, S11, S12). A previous work of Lopez-Alvis et al. (2021), which

conducted a gradient-based inverse problem with VAE, observed similar results as ours where a complex VAE model may not offer good inverse modeling performance. The reason can be understood as the possible conflict between the accuracy of results and the feasibility of inverse models. At the same time, when applying an inverse model, the tradeoff between generative accuracy and a well-behaved generator should be considered. Deep, convolutional RegAE has the capability for the high quality of inverse analysis that has some desirable features (e.g., sharper edges between the two facies in the bimodal images). However, the performance is unstable and it is hard to understand when it will perform poorly, making it potentially untrustworthy.

Figure 5 depicts the convergence results with $n_z = 25$, $n_z = 50$, $n_z = 100$, and $n_z = 200$ of shallow, dense RegAE and deep, convolutional RegAE for Gaussian and bimodal fields. This figure shows that convergence of Gaussian fields is generally obtained after 20 iterations, except for the low-quality results (discussed above). In contrast, note the convergence of bimodal fields obtained after 50 iterations. The convergence plots also illustrate the trend of improvement in the objective function with increasing n_z values. Sometimes, for $n_z = 200$, the objective function value increases again. While the objective functions are not strictly comparable, this indicates that there is little or no benefit in improving the latent space beyond a certain point, as is found in other dimensionality reduction approaches (Kitanidis and Lee, 2014; Lee and Kitanidis, 2014; Lin et al., 2017).

Quantitatively, the inverse results of the RegAE model for bimodal fields capture its large-scale trend as it illustrates where the “Conductivity 1” (red) and “Conductivity 2” (blue) parts are present. It also captures some of the variability within each of these two fields at the medium scale. As a result, the RegAE model shows its ability to characterize the complex spatial heterogeneity of hydraulic conductivity distribution. The edges of the inverse results of shallow, dense RegAE are blurred compared to the sharp edges of deep, convolutional RegAE. However, even with the advantages of deep, convolutional RegAE at edge capture, the behavior of the inconsistent results cannot be ignored when evaluating its performance compared to shallow, dense RegAE.

3.3 Statistical RegAE Results

Aiming to test the variance of results performance, we conduct the inverse analyses on 100 reference fields of the two neural network structures for Gaussian and bimodal fields. In addition, we generate the runs for three seeds that result in different random numbers being used in the VAE training process. Afterward, the 300 results of each model are collected to illustrate the statistical evaluation and characterized as CDF about the relative error in Figs. 6–8. On the CDF line, these figures show the cumulative probability for the relative error [Eq. (5)] as smaller than the corresponding number on the X axis of the 300 results for each model. Figure 6 presents the aggregate results of each model for the four different n_z options. For Gaussian fields, except for $n_z = 25$, shallow, dense RegAE depicts the hydraulic conductivity distribution better than deep, convolutional RegAE, especially for $n_z = 100$ and 200, which shows consistency with the results in Fig. 3. However, for bimodal fields, the two shallow, dense and deep, convolutional RegAE illustrate a slight difference between them. The deep, convolutional RegAE performs better for the smaller relative error part (blue line above the yellow line for small relative error), while the shallow, dense RegAE wins for the larger relative error part (yellow line above the blue line for large relative error).

To compare the different models, the results of different n_z options are shown in Fig. 7. For Gaussian fields, the deep, convolutional RegAE provides a large variability compared to the

shallow, dense RegAE where only the case $n_z = 25$ shows the lack of good results with the relative error from 0.2 to 0.5. In contrast, the variances for both deep, convolutional and shallow, dense are limited for bimodal fields. Among all the models for Gaussian fields in Fig. 8(a), deep, convolutional RegAE has the best performance. At the same time, with $n_z = 100$ and 200, it also presented the worst results and related the situation in Fig. 3 for relative error over 200%. Oppositely, the bimodal fields do not show a significant variance in Fig. 8(b). Deep, convolutional RegAE with $n_z = 50$ and shallow, dense RegAE with $n_z = 100$ are comparable good results. In conclusion, small n_z of deep, convolutional RegAE can give good results, while the results are not trustworthy if an appropriate value of n_z is not known. However, shallow, dense RegAE does not provide the best results, but its performance is stable with only a modest dependence on n_z . At the same time, the performance of shallow, dense RegAE is reliable, while the performance of deep, convolutional RegAE requires hyperparameter optimization to choose n_z .

The field's complexity should be considered during the hyperparameter optimization for optimal n_z . The relationship that optimal n_z may depend on the complexity of the field was disclosed in the studies by Tartakovsky et al. (2021) and He et al. (2020). Deciding the best choice of n_z is not always intuitive (not necessarily the larger, the better). It, to some extent, depends on the field's complexity, where sometimes smaller n_z performs better than larger n_z . As shown in Fig. 8(a), the modeling results of Gaussian fields reach the best performance as n_z is a small number (25), while for large n_z values (100 and 200), the results are not acceptable. In contrast, large n_z values (50 and 100) are the best options for bimodal fields. Considering bimodal fields have a higher level of complexity, it leads us to pick large n_z values when modeling complicated geological fields and vice versa.

3.4 Traditional Geostatistical Results

PCGA is a recently developed traditional geostatistical inverse problem for hydrogeology with improved computational cost saving by applying the principle components of covariance. Besides the RegAE model, we also evaluated the inverse analysis performance of PCGA for both Gaussian and bimodal fields, and the results are shown in Figs. 9 and 10, respectively. Similarly, only the smallest and largest n_z values are represented, while the completed results are shown in supplementary figures (S13, S14). PCGA shows its strength in the characterization of the Gaussian field based on low relative error. However, PCGA struggles with the characterization of the bimodal field, where the relative error is comparably higher no matter the selection of n_z values. In the work of Tartakovsky et al. (2021), they tested the traditional geostatistical method on a simple bimodal field (Fig. 5), and the good results approve its application. However, the bimodal fields in this study are more complex than theirs because, in their problem, the permeability is constant within each zone. Hence, the traditional geostatistical inverse problems, such as PCGA, prove their capability at Gaussian fields or the relatively simple bimodal fields, which is ideal for the regularization based on the covariance matrix (since the mean and covariance fully characterize the distribution). However, at the same time, the results are more concerning when dealing with the more complex and heterogeneous situations (complicated bimodal fields). One of the advantages that PCGA has compared to the VAE is the ability to quantify the information lost. However, this is more challenging for the VAE because of the added complexity of the neural network. As a result, the VAE aiming to reach good results needs more careful parameter tuning than PCGA.

4. DISCUSSION

The RegAE approach provides a means to perform inverse analysis effectively. Our goal here is to demonstrate its application for different study fields and achieve good modeling results by studying the performance of two different neural network architectures—a deep, convolutional network compared with a shallow, dense network. We now focus our discussion on results and benefits which offer the potential for future application. Three inverse problems are compared in terms of accuracy and stability for two different parameter fields. Generally, the results indicate that it is easier to conduct inverse analysis on the Gaussian field than the bimodal field, which has more complex heterogeneity. Specifically, PCGA shows consistent modeling results with good quality in Gaussian fields. In contrast, the PCGA struggles with the bimodal fields because the statistics of these fields cannot be accurately characterized through a two-point covariance model. In addition, we test two different neural network architectures in the inverse analysis step of the RegAE model. For the shallow, dense RegAE, the inverse results accurately recreate the “true” conductivity fields with relatively low relative error. At the same time, it shows consistent results among a large ensemble of inverse problems. As a result, shallow, dense RegAE illustrates an insight for characterizing heterogeneous parameter space with accurate and stable results. However, for another architecture, the deep, convolutional RegAE yields various performances with good results when n_z is chosen appropriately. In some scenarios, the relative error of inverse analysis exceeds 100% demonstrating the potential for bad outcomes if n_z is not well-chosen. Therefore, applying the deep, convolutional RegAE needs careful selection of the number of latent variables or hyperparameter tuning.

Our analysis was performed on a machine with an Intel(R) Core(TM) i9-9960X CPU @ 3.10 GHz with 32 threads and an NVIDIA RTX 2080 Ti GPU. The data generated in the case of $N = 10^5$ realizations (used for the VAE training) required ~ 10 min of wall time using 32 threads. We used a GPU to train the VAE, and this process took 25 min for each of the cases using 100 training epochs. The time to perform the inverse analysis varied somewhat depending on the reference field. Averaging the time required to perform the inverse analysis on each of the three reference fields shown in Figs. 3 and 4, it needed 3–12 min for cases with different n_z selections. Note that with more complex physical models (e.g., transient flow, multiphase flow, reactive transport, etc.), the cost of performing the inverse analysis will increase significantly because the cost of running the forward model will dramatically increase. However, the computational cost of the data generation and training of the VAE will remain about the same. Therefore, the use of the VAE adds a modest computational cost to the inverse analysis—the cost of the physical model runs (and computing derivatives of the physical model) dominates, and the cost of the VAE becomes negligible.

RegAE provides an application of inverse problems with regularization and VAE for fast and easy inverse modeling. Among all the features of RegAE, the VAE reduces the dimensionality and does not require any forward model simulations during data generation, and the application of regularization supports easy optimization; these features all illustrate their vast potential at the computational cost saving. Afterward, the computational cost of the inverse analysis is dominated by gradient calculation, which will dominate the computational cost of the inverse analysis when using expensive physical models. We note that while the dimensionality reduction here makes it possible to perform inverse analysis on highly parameterized fields with low computational cost, our model also tests the possibility of utilizing automatic differentiation to compute the gradients efficiently. However, the performance benefits of dimensionality reduction go away when applying automatic differentiation, which provides even more significant

performance benefits. It is worth noting that automatic differentiation only shows its advantage at computational acceleration, and its improvement of the quality of the inverse results is minimal in these problems.

RegAE provides a framework for efficiently characterizing heterogeneous hydraulic conductivity fields and demonstrates accurate and trustworthy prediction results when a shallow, dense NN is used within the VAE. In addition, we found that a shallow and dense network outperforms a deep and convolutional network considering its easy application without hyperparameter tuning and consistently good results. This is surprising since deep, convolutional NNs would generally be expected to give better performance than shallow, dense NNs. Our study illustrates that a deep and complex network may not be applicable for accurate and consistent inverse modeling results. In addition, the dense neural network provides another advantage at dealing with unstructured grids, where the convolutional one shows difficulty resulting in its grids-based information extraction process. However, while the dense neural network has these specialties, there are also potential downsides. Notably, He et al. (2020) show that the size of a fully connected feed-forward deep neural network exponentially increases with the decreasing correlation length, so this could act as a limit on how small the correlation length can be or how smooth the property distribution is in fields. The application of this modeling approach also provides new avenues of support for groundwater source management and underground contamination treatment in complex heterogeneous reservoirs. Furthermore, the workflow developed here for inverse problems is broadly applicable to numerous geological conditions, e.g., fracture-dominated networks, reservoirs with extremely low permeability intrusive body, and karst systems.

5. CONCLUSION

We have presented the application of an approach called RegAE for performing inverse analysis, which provides an efficient method of regularizing inverse problems where the parameter fields are high-dimensional and have different heterogeneous structures. RegAE is designed for inverse problems and leverages a VAE to learn how to regularize these inverse problems based on the unconditioned realizations of the high-dimensional parameter fields. This study aims to determine the application of the RegAE method by testing different neural network structures and hyperparameter tuning to achieve good modeling results. We demonstrated the approach for a hydrogeologic inverse problem with two types of heterogeneity of hydraulic conductivity, and we call them Gaussian (simple heterogeneity) and bimodal (more complex heterogeneity) fields. More importantly, two neural network structures (shallow, dense and deep, convolutional) were evaluated for their performance during inverse analysis. The shallow, dense RegAE provides accurate and consistent performance, while deep, convolutional RegAE only shows promising results when the number of latent variables is tuned correctly. This tendency for the shallow, dense neural network to outperform the deep, convolutional network is surprising since more complex neural networks generally outperform simpler neural networks. Additional comparison with a traditional geostatistical method illustrates that the VAE-based approach performs similarly for the Gaussian fields and provides better performance for the bimodal fields. This is not terribly surprising since the VAE can represent more complex statistical structures than the covariance matrix that is the heart of the geostatistical approach. Our VAE-based approach is computationally efficient and obtains an excellent solution to the inverse problem by easing the regularization process, reducing the dimensionality of the parameter space, and applying automatic differentiation.

ACKNOWLEDGMENT

Hao Wu and Daniel O’Malley acknowledge support from Los Alamos National Laboratory’s Laboratory Directed Research and Development Early Career Award (20200575ECR).

DATA AVAILABILITY

A computer program automatically generated all the data used in this manuscript. The code for generating the data, training the VAE, and performing the inverse analysis is available at <https://github.com/madsjulia/RegAE.jl>.

REFERENCES

- Barajas-Solano, D.A. and Tartakovsky, A.M., Approximate Bayesian Model Inversion for PDEs with Heterogeneous and State-Dependent Coefficients, *J. Comput. Phys.*, vol. **395**, pp. 247–262, 2019.
- Bredies, K., Kunisch, K., and Pock, T., Total Generalized Variation, *SIAM J. Imaging Sci.*, vol. **3**, no. 3, pp. 492–526, 2010.
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L.J., Inverse Problem in Hydrogeology, *Hydrogeol. J.*, vol. **13**, no. 1, pp. 206–222, 2005.
- Doersch, C., Tutorial on Variational Autoencoders, 2016. arXiv: 1606.05908
- Doughty, C. and Pruess, K., Modeling Supercritical Carbon Dioxide Injection in Heterogeneous Porous Media, *Vadose Zone J.*, vol. **3**, no. 3, pp. 837–847, 2004.
- Franklin, J.N., On Tikhonov’s Method for Ill-Posed Problems, *Math. Comput.*, vol. **28**, no. 128, pp. 889–907, 1974.
- Geneva, N. and Zabarar, N., Modeling the Dynamics of PDE Systems with Physics-Constrained Deep Auto-Regressive Networks, *J. Comput. Phys.*, vol. **403**, p. 109056, 2020.
- Hager, W.W. and Zhang, H., A New Conjugate Gradient Method with Guaranteed Descent and an Efficient Line Search, *SIAM J. Optimiz.*, vol. **16**, no. 1, pp. 170–192, 2005.
- He, Q., Barajas-Solano, D., Tartakovsky, G., and Tartakovsky, A.M., Physics-Informed Neural Networks for Multiphysics Data Assimilation with Application to Subsurface Transport, *Adv. Water Res.*, vol. **141**, p. 103610, 2020.
- Innes, M., Flux: Elegant Machine Learning with Julia, *J. Open Source Software*, vol. **3**, no. 25, p. 602, 2018.
- Innes, M., Edelman, A., Fischer, K., Rackauckas, C., Saba, E., Shah, V.B., and Tebbutt, W., A Differentiable Programming System to Bridge Machine Learning and Scientific Computing, 2019. arXiv: 1907.07587
- Jayne, R.S., Wu, H., and Pollyea, R.M., Geologic CO₂ Sequestration and Permeability Uncertainty in a Highly Heterogeneous Reservoir, *Int. J. Greenhouse Gas Control*, vol. **83**, pp. 128–139, 2019.
- Kadeethum, T., O’Malley, D., Fuhg, J.N., Choi, Y., Lee, J., Viswanathan, H.S., and Bouklas, N., A Framework for Data-Driven Solution and Parameter Estimation of PDEs Using Conditional Generative Adversarial Networks, 2021. arXiv: 2105.13136
- Kingma, D.P. and Welling, M., Auto-Encoding Variational Bayes, 2013. arXiv: 1312.6114
- Kitanidis, P.K. and Lee, J., Principal Component Geostatistical Approach for Large-Dimensional Inverse Problems, *Water Res. Res.*, vol. **50**, no. 7, pp. 5428–5443, 2014.
- Kullback, S., *Information Theory and Statistics*, North Chelmsford, MA: Courier Corporation, 1997.

- Laloy, E., Héroult, R., Lee, J., Jacques, D., and Linde, N., Inversion Using a New Low-Dimensional Representation of Complex Binary Geological Media Based on a Deep Neural Network, *Adv. Water Res.*, vol. **110**, pp. 387–405, 2017.
- Lee, J. and Kitanidis, P.K., Large-Scale Hydraulic Tomography and Joint Inversion of Head and Tracer Data Using the Principal Component Geostatistical Approach (PCGA), *Water Resour. Res.*, vol. **50**, no. 7, pp. 5410–5427, 2014.
- Lin, Y., Le, E.B., O'Malley, D., Vesselinov, V.V., and Bui-Thanh, T., Large-Scale Inverse Model Analyses Employing Fast Randomized Data Reduction, *Water Resour. Res.*, vol. **53**, no. 8, pp. 6784–6801, 2017.
- Linde, N., Renard, P., Mukerji, T., and Caers, J., Geological Realism in Hydrogeological and Geophysical Inverse Modeling: A Review, *Adv. Water Res.*, vol. **86**, pp. 86–101, 2015.
- Liu, D.C. and Nocedal, J., On the Limited Memory BFGs Method for Large Scale Optimization, *Math. Progr.*, vol. **45**, nos. 1-3, pp. 503–528, 1989.
- Lopez-Alvis, J., Laloy, E., Nguyen, F., and Hermans, T., Deep Generative Models in Inversion: The Impact of the Generator's Nonlinearity and Development of a New Approach Based on a Variational Autoencoder, *Comput. Geosci.*, vol. **152**, p. 104762, 2021.
- McGovern, A., Lagerquist, R., Gagne, D.J., Jergensen, G.E., Elmore, K.L., Homeyer, C.R., and Smith, T., Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning, *Bull. Am. Meteorol. Soc.*, vol. **100**, no. 11, pp. 2175–2199, 2019.
- Mo, S., Zhu, Y., Zabarav, N., Shi, X., and Wu, J., Deep Convolutional Encoder-Decoder Networks for Uncertainty Quantification of Dynamic Multiphase Flow in Heterogeneous Media, *Water Resour. Res.*, vol. **55**, no. 1, pp. 703–728, 2019.
- Mogensen, P.K. and Riseth, A.N., Optim: A Mathematical Optimization Package for Julia, *J. Open Source Soft.*, vol. **3**, no. 24, p. 615, 2018.
- O'Malley, D., Vesselinov, V.V., and Harp, D., DPFEHM.jl, accessed from <https://github.com/OrchardLANL/DPFEHM.jl>, 2020.
- O'Shea, K. and Nash, R., An Introduction to Convolutional Neural Networks, 2015. arXiv: 1511.08458
- Robbe, P., Gaussianrandomfields.jl, accessed from <https://github.com/PieterjanRobbe/GaussianRandomFields.jl>, 2020.
- Rudin, L.I., Osher, S., and Fatemi, E., Nonlinear Total Variation Based Noise Removal Algorithms, *Phys. D: Nonlinear Phenom.*, vol. **60**, nos. 1-4, pp. 259–268, 1992.
- Sinha, S., de Lima, R.P., Lin, Y., Sun, A.Y., Symons, N., Pawar, R., and Guthrie, G., Normal or Abnormal? Machine Learning for the Leakage Detection in Carbon Sequestration Projects Using Pressure Field Data, *Int. J. Greenhouse Gas Control*, vol. **103**, p. 103189, 2020.
- Tartakovsky, A.M., Barajas-Solano, D.A., and He, Q., Physics-Informed Machine Learning with Conditional Karhunen-Loève Expansions, *J. Comput. Phys.*, vol. **426**, p. 109904, 2021.
- Tartakovsky, A.M., Marrero, C.O., Perdikaris, P., Tartakovsky, G.D., and Barajas-Solano, D., Physics-Informed Deep Neural Networks for Learning Parameters and Constitutive Relationships in Subsurface Flow Problems, *Water Resour. Res.*, vol. **56**, no. 5, p. e2019WR026731, 2020.
- Tikhonov, A.N., Solution of Incorrectly Formulated Problems and the Regularization Method, *Soviet Math.*, vol. **4**, pp. 1035–1038, 1963.
- Wu, H., Lubbers, N., Viswanathan, H.S., and Pollyea, R.M., A Multi-Dimensional Parametric Study of Variability in Multi-Phase Flow Dynamics during Geologic CO₂ Sequestration Accelerated with Machine Learning, *Appl. Energy*, vol. **287**, p. 116580, 2021.
- Xu, K., Tartakovsky, A.M., Burghardt, J., and Darve, E., Learning Viscoelasticity Models from Indirect Data Using Deep Neural Networks, *Comput. Methods Appl. Mech. Eng.*, vol. **387**, p. 114124, 2021.
- Zagst, R. and Pöschik, M., Inverse Portfolio Optimisation under Constraints, *J. Asset Manage.*, vol. **9**, no. 3,

- pp. 239–253, 2008.
- Zhou, H., Gómez-Hernández, J.J., and Li, L., Inverse Methods in Hydrogeology: Evolution and Recent Trends, *Adv. Water Res.*, vol. **63**, pp. 22–37, 2014.
- Zhou, Z., Lin, Y., Zhang, Z., Wu, Y., Wang, Z., Dilmore, R., and Guthrie, G., A Data-Driven CO₂ Leakage Detection Using Seismic Data and Spatial–Temporal Densely Connected Convolutional Neural Networks, *Int. J. Greenhouse Gas Control*, vol. **90**, p. 102790, 2019.
- Zhu, Y., Zabaras, N., Koutsourelakis, P.S., and Perdikaris, P., Physics-Constrained Deep Learning for High-Dimensional Surrogate Modeling and Uncertainty Quantification without Labeled Data, *J. Comput. Phys.*, vol. **394**, pp. 56–81, 2019.