

Nonintrusive reduced basis approximation to the solution of the Helmholtz equation: The magnetotellurics case

Alejandro Quiaro¹, Dawei Liu¹, and Mauricio D. Sacchi¹

ABSTRACT

Electromagnetic wave propagation is commonly modeled using the Helmholtz partial differential equation, which plays a significant role in geophysical studies, such as magnetotellurics forward modeling. Although analytical solutions exist for layered media, most geophysical applications depend on numerical finite-difference, finite-element, or finite-volume solvers. These traditional methods are computationally demanding, particularly for large-scale problems and workflows that require repeated evaluations, such as real-time or probabilistic inversions. Reduced basis (RB) techniques have been developed to accelerate finite-element solvers by reducing the stiffness matrix and nodal forces vector size. However, these methods rely on explicit access to the stiffness matrix, which can limit their applicability. We present a nonintrusive data-driven approach, adapted for the first time to magnetotellurics forward modeling, that eliminates the need for

explicit stiffness matrix availability and is compatible with various numerical solvers. Using a predefined parameter domain, we construct a snapshot matrix from high-fidelity solutions generated for a subset of model parameters. Proper orthogonal decomposition is then applied to extract RB, and a neural network is trained to map the model space to the reduced coefficient space. This enables rapid evaluation of the Helmholtz equation, achieving a speed-up of four orders of magnitude compared with traditional solvers, with average median errors of 9% transverse magnetic (TM) mode and 2% transverse electric (TE) mode. Further accuracy improvements are achieved by incorporating a minimal set of high-fidelity observations and leveraging the fast evaluation to regularize an inverse problem, reducing errors to 2% for the TM mode and 1.5% for the TE mode. These results highlight this approach's potential to dramatically decrease computational costs while maintaining accuracy, making it a flexible and scalable tool for efficient geophysical forward evaluations.

INTRODUCTION

Many problems across various branches of science are addressed by solving partial differential equations (PDEs), which model fields that describe physical phenomena. In geophysical applications, PDEs can be solved analytically for idealized conditions, such as homogeneous media or layered media; however, for real-world cases in which the subsurface is complex, a direct analytical evaluation is not feasible, and one must resort to a numerical solution (LeVeque, 2007; Durran, 2013). Standard techniques for the numerical evaluation of PDEs include finite difference (FD) (Smith, 1985; Vidale, 1988), finite element (FE) (Strang and Fix, 1973; Virieux et al., 2011; Castillo-Reyes et al., 2018; Rochlitz et al., 2019), and finite volume (FV) (Jahandari and Farquharson, 2015), which can be prohibitively expensive for some applications such as

real-time evaluations or uncertainty quantification using Monte Carlo techniques (Dai, 2019; Ramnath, 2023).

Reduced basis (RB) methods address these computational challenges by constructing a low-dimensional approximation of the solution space (Peterson, 1989; Benner et al., 2015). Proper orthogonal decomposition (POD) is a widely used RB technique (Berkooz et al., 1993; Smith et al., 2005) in which high-fidelity solutions (snapshots) are calculated to build RB. A successful application of an RB method in geosciences can be found in Manassero et al. (2020), where an RB solution is found by reducing the dimensionality of the FE linear system of equations via Galerkin projection (Kunisch and Volkwein, 2002; Kahlbacher and Volkwein, 2007; Wang et al., 2013) and then using the fast evaluation of the forward modeling into a Markov chain Monte Carlo scheme for probabilistic inversion. However, traditional RB techniques

Manuscript received by the Editor 15 August 2024; revised manuscript received 2 February 2025; published ahead of production 19 February 2025; published online 28 April 2025.

¹University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: quiarosa@ualberta.ca; 409791715@qq.com (corresponding author); msacchi@ualberta.ca.

© 2025 Society of Exploration Geophysicists. All rights reserved.

are considered “intrusive,” requiring access to the stiffness matrix and the nodal force vector, which limits the applicability for some geophysical applications, especially when using commercial software or when not using an FE numerical solver. As an alternative to intrusive methods, nonintrusive solutions have been proposed to address cases in which access to the stiffness matrix is impossible. In the nonintrusive case, the set of reduced coefficients that maps onto the RB is often found via a regression method such as an artificial neural network (Hesthaven and Ubbiali, 2018; Fresca et al., 2021; Kadeethum et al., 2022) or cubic spline (Li et al., 2021). Successful applications of this method can be found from examples using the Navier-Stokes equation, heat equation, Poisson equation (Hesthaven and Ubbiali, 2018), and geothermal (Degen et al., 2023) and ground motion applications (Rekoske et al., 2023).

In this study, we propose to test a nonintrusive reduced basis (NIRB) technique to accelerate the evaluation of the Helmholtz equation, specifically adapted for the first time to model the response of the 2D magnetotellurics forward operator. Magnetotellurics forward modeling is often deemed computationally expensive to solve due to its size when media is nonhomogeneous (Deng et al., 2023). The proposed workflow first establishes the low rankness of the magnetotellurics data embedded into the snapshot matrix. Based on the observations, RB by POD for the forward response for transverse electric (TE) and transverse magnetic (TM) modes is built. Finally, we exploit the ability of artificial neural networks to model complex nonlinear relationships and train a neural network that maps samples from the model space to RB.

The paper is organized as follows. In the “Theory” section, we describe the theoretical framework for the NIRB technique using artificial neural networks and the PDE that govern the magnetotellurics forward response. Then, in the “Experiments” section, we discuss how to build the snapshot matrix, select the basis, and train the neural network to analyze the results of applying the workflow on a validation data set of entirely random resistivity models. Finally, in the “Discussion” section, we show that it is possible to build a NIRB forward operator without access to the source code of the high-fidelity numerical solver and gain four orders of magnitude in computational efficiency while keeping acceptable levels of accuracy. The results suggest that this method could be integrated into a probabilistic inversion workflow or any other workflow that requires multiple and fast evaluations.

THEORY

Let us consider a model space \mathcal{M} that contains a specific family of models \mathbf{m}_i with size $M \times 1$. This family of models is extracted from a specific prior distribution and possesses any other desired feature, such as smoothness or a specific discretization. We can consider the forward problem:

$$\mathcal{F}[\mathbf{m}_i] = \mathbf{d}_i, \quad (1)$$

where \mathbf{d}_i is the output of the forward operator \mathcal{F} in the data space \mathcal{D} and has dimensions $N \times 1$. In our case, the forward operator \mathcal{F} represents the solution to a PDE via any analytical or numerical method, such as FD, FE, or FV. Suppose a representative number of samples is extracted from the space \mathcal{M} , and every resulting \mathbf{d}_i is sufficiently smooth. We can assume that the data space \mathcal{D} can be considered low rank. This means that a linear approximation space $\mathcal{D}_{\text{approx}}$ could reasonably approximate \mathcal{D} . The existence of a

low-rank linear approximation space translates into the fact that any point in the data space can be reasonably represented by a matrix-vector multiplication of a basis and a vector of reduced coefficients, which significantly reduces the computation complexity and enables faster modeling.

Our goal in this paper is to find a computationally inexpensive approximation of the operator \mathcal{F} that maps \mathcal{M} onto $\mathcal{D}_{\text{approx}}$, which is the low-rank approximation to the original data space \mathcal{D} . More specifically, we look into a forward operator \mathcal{F} that solves the Helmholtz equation for the magnetotellurics TE and TM modes, where $m(x, z)$ represents the subsurface properties (resistivities), which are continuous but are mapped via discretization to a vector of model parameters \mathbf{m}_i . The M elements of vector \mathbf{m}_i represent each discrete element of our model. The data space is represented by modeled observations of impedance, measured at N_r stations at the surface for N_ω number of frequencies. The data are packed in a vector with N number of elements, where $N = N_r \times N_\omega$.

The prior

Let us consider one model belonging to the model space \mathcal{M} with M samples, which are uniform and independently distributed as

$$m_j \sim \mathcal{U}(\log(\rho_{\min}), \log(\rho_{\max})), \quad (2)$$

and that also satisfies

$$\|\nabla \mathbf{m}\|_2^2 < \epsilon, \quad (3)$$

where ∇ is the gradient operator. By satisfying these two conditions, we only admit smooth models (depending on the parameter ϵ) and resistivities within a certain range. Because resistivities are log-normal distributed, the uniform distribution is on the logarithm of the resistivities. The M elements in each model correspond with the number of elements in our computational grid required to calculate the high-fidelity solution (the number of cells in our FD grid or elements in our FE mesh). The smoothness constraint introduces a covariance between samples of the model vector, implying that neighboring cells should exhibit similar resistivity values. One way to impose smoothness and reduce the dimensionality of the model space simultaneously is to model our parameters using a Gaussian process (Ray and Myer, 2019). The Gaussian process allows us to define a nonparametric regression function that maps from a surrogate model with a minimum set of samples to the full numerical mesh while also fulfilling the smoothness constraint.

If we assume that a sufficiently good smooth representation of our model can be obtained from a Gaussian process regression, the models extracted from the prior are

$$\mathbf{m}_{*,i} = [m_*^1, m_*^2, m_*^3, \dots, m_*^{n_l}], \quad (4)$$

with $n_l \ll M$. But the snapshot matrix is built with the models obtained by applying a Gaussian process regression operator (\mathcal{K}), which in our case is an ordinary kriging operator, such that

$$\mathbf{m} = \mathcal{K}(\mathbf{m}_*). \quad (5)$$

The model matrix

To generate the required samples to build the snapshot matrix, it is necessary to explore thoroughly the model space. A simple approach could be random sampling; however, for high-dimensional systems, it may require a large number of samples to explore sufficiently the model space (Renardy et al., 2021), having the risk of leaving areas uncovered. In our case, we ignore the areas of the model that are more significant in building the RB. Therefore, the sampling strategy should guarantee that the model space is evenly sampled while keeping the number of samples small. A solution is found with the Latin hypercube sampling (McKay, 1992), which provides an optimal pseudorandom sampling technique that reduces the number of models needed to estimate the statistical properties of the model space. This method ensures that our minimal set of samples uniformly covers the model space ranges for each parameter, using bins of equal probability while maintaining lower variance compared with entirely random sampling (McKay et al., 2000). Another advantage is that if one sample is dropped from a Latin hypercube, the resulting subset is still a Latin hypercube (Viana, 2016), giving more flexibility to the design of the experiments.

To build our model matrix \mathbf{M} , we first extract a surrogate model Latin hypercube of 10,000 samples from the probability distribution described in equation 2. This surrogate model matrix is fed to the Gaussian process regression operator, which maps the surrogate model onto the computational grid. In Figure 1, we show some models reconstructed from the surrogate models extracted by the Latin hypercube sampling.

Magnetotellurics forward operator and the snapshot matrix

The magnetotellurics forward operator is based on solving the Helmholtz equation, which is a linear PDE with the following basic form:

$$\nabla^2 f - k^2 f = 0. \quad (6)$$

Two propagation modes are modeled in the magnetotellurics method (Berdichevsky et al., 1998). In the TE mode, the electric field is polarized along the strike of the structures and can be described using the following Helmholtz equation:

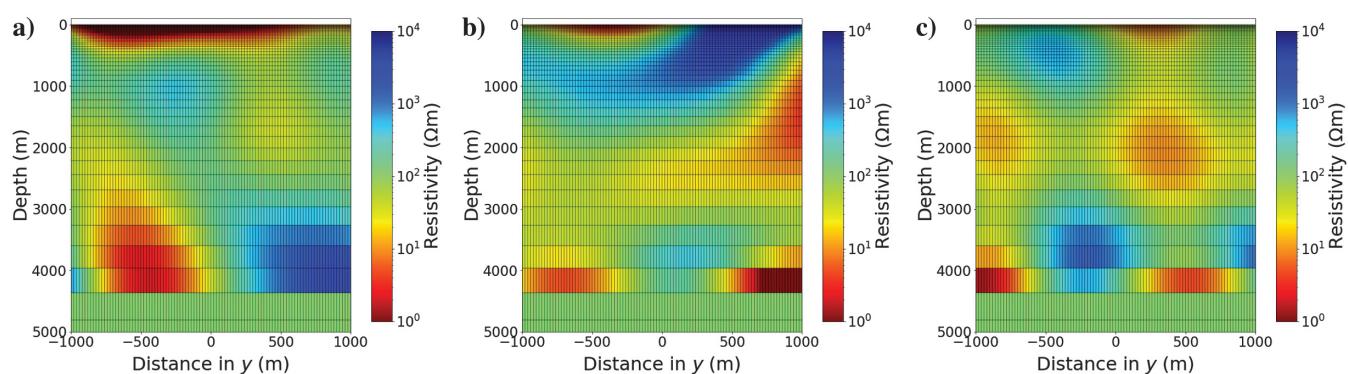


Figure 1. Three models extracted from the 10,000 samples Latin hypercube used for training. (a) Model 1, (b) model 5000, and (c) model 10,000. Notice the wide range of behaviors simulated for the training data set.

$$\nabla^2 E(y, z) - i\omega\mu\sigma E(x) = 0. \quad (7)$$

In the TE mode, there is no formation of surface electric charges because the electric current does not cross boundaries with different resistivities. However, if we consider the TM mode, the magnetic field is polarized along the strike, and the numerical solution is more challenging because electric current densities should be conserved at the interfaces (Unsworth et al., 1993). This mode can be modeled using the following PDE:

$$\nabla^2 H(y, z) - i\omega\mu\sigma H(x) = 0. \quad (8)$$

The impedance at the surface for each mode is then the ratio between orthogonal orientations of the electric and magnetic field at the surface (Z). The data are given by the apparent resistivity

$$\rho_a(\omega) = \frac{|Z|^2}{\omega\mu} \quad (9)$$

and phase

$$\phi(\omega) = \arctan\left(\frac{\text{Re}(Z)}{\text{Im}(Z)}\right), \quad (10)$$

with ω being the angular frequency and μ the magnetic permeability of free space. Although other geophysical phenomena can be described using the Helmholtz equation, such as acoustic wave propagation in the frequency domain (Song and Williamson, 1995; Plessix, 2007; Signorini et al., 2017), we focus only on magnetotellurics applications in this paper.

We build the snapshot matrix and use the 2D MT forward operator of the package SimPEG (Cockett et al., 2015; Heagy et al., 2017). We consider an FE mesh of 164×80 cells ($M = 13,120$) represented by surrogate models of 50 samples. We consider a minimum frequency of 0.01 Hz and a maximum frequency of 100 Hz, generating five samples per decade. The primary cell size in x is 20 m, and the starting layer size in z is 5 m, with an expansion rate with a depth of 10%. This cell size guarantees stability in the numerical solution because the prior is defined between 1 and 3000 $\text{ohm} \cdot \text{m}$. For the highest frequency and smallest resistivity case, we have a skin depth

of approximately 50 m, which makes our smallest cell one-tenth of our skin depth. A padding area in x is also included, with an expansion rate of 20% for 12 cells at each lateral edge of the model. Our 2D example contains 100 stations between -1000 and 1000 m in x . After forward modeling, we obtain apparent resistivity and phase for the TE and the TM modes. Figures 2, 3, and 4 show the numerical solution for the models shown in Figure 1. As expected, the responses are smooth for all the propagation modes and all the frequencies in apparent resistivity and phase, and the TM mode displays less frequency-dependent behavior than the TE mode.

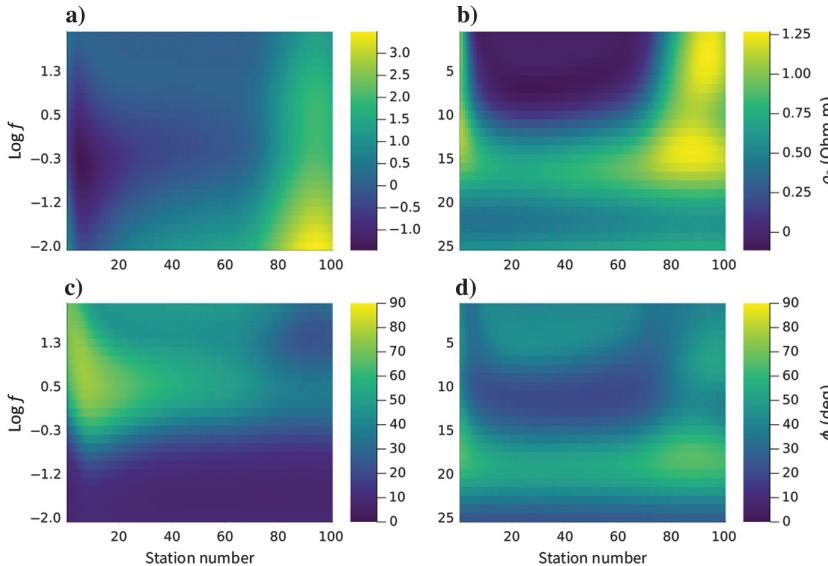


Figure 2. Forward response corresponding to training model 1 (Figure 1a). (a) TM apparent resistivity, (b) TE apparent resistivity, (c) TM phase, and (d) TE phase. Here, f is frequency in Hz, ρ_a is apparent resistivity, and ϕ is phase.

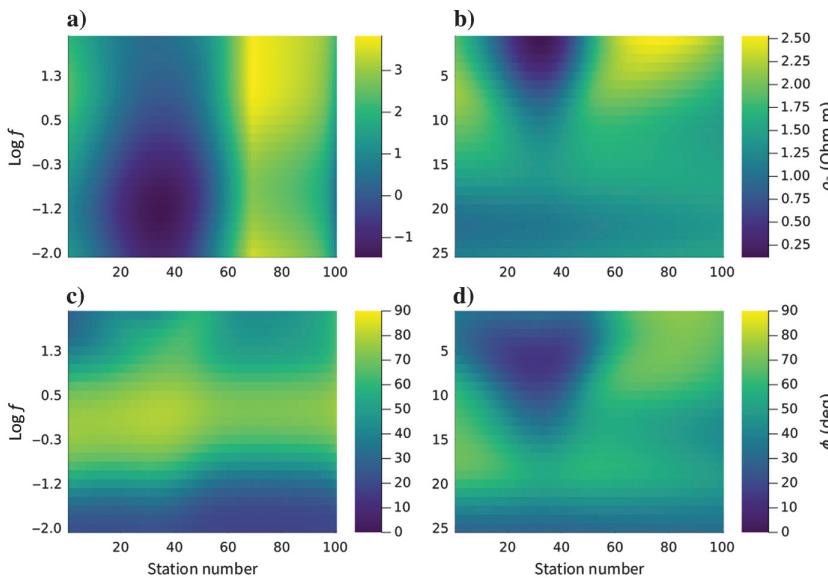


Figure 3. Forward response corresponding to training model 5000 (Figure 1b). (a) TM apparent resistivity, (b) TE apparent resistivity, (c) TM phase, and (d) TE phase. Here, f is frequency in Hz, ρ_a is apparent resistivity, and ϕ is phase.

Proper orthogonal decomposition

In this section, we explore the rank of the MT data matrix. The purpose of POD is to decompose a physical field into a set of deterministic basis functions, which provide a compact approximation of a high-dimensional process (Berkooz et al., 1993). Given a set of models \mathbf{m}_i with $i = 1:k$, we can use the forward modeling algorithm to compute $\mathbf{d}_i = \mathcal{F}[\mathbf{m}_i]$. The “snapshot matrix” (\mathbf{D}) is a matrix whose columns are given by every \mathbf{d}_i such that $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K]$ with size $N \times K$. If \mathbf{D} is sufficiently large to provide a good approximation of the data space and is also

low rank, it should be possible to identify a set of parameter-independent basis functions using low-rank decomposition techniques. The singular value decomposition (SVD) decomposes the matrix \mathbf{D} as

$$\mathbf{D}_p = \mathbf{U}_p \Sigma_p \mathbf{V}_p^T, \quad (11)$$

where p is the rank of the matrix. Moreover, if we choose to truncate the SVD, the Eckart-Young theorem (Eckart and Young, 1936) shows that a low-rank approximation with only the p first basis functions is an optimal approximation in the least-squares sense. Notice that equation 11 can also be written as

$$\mathbf{D} = \mathbf{U}_p \mathbf{A}, \quad (12)$$

and therefore any column of \mathbf{D} can be written as

$$\mathbf{d}_j = \mathbf{U}_p \mathbf{a}_j, \quad j = 1:k. \quad (13)$$

This proves that the matrix \mathbf{U}_p constitutes a basis to represent the data with acceptable fitting errors. Moreover, because the vector of coefficients \mathbf{a}_j has length $p \ll N$, we have effectively found a compressed representation of the data matrix. A diagram with the steps to build the basis is shown in Figure 5. One advantage of knowing the basis is that if we are solving the FE equilibrium equation, we can project the stiffness matrix and the nodal force vector onto the basis, reducing the size of the equilibrium equation and therefore making it more computationally affordable to solve the PDE (Manassero et al., 2020).

However, this approach is only possible if there is a linear approximation space $\mathcal{D}_{\text{approx}}$. Therefore, the first question that needs to be answered is if the snapshot matrix can be considered low rank. The rank of the matrix is the number of linearly independent basis functions that allow the reproduction of all the columns of the matrix by a linear combination of the basis. By inspecting the singular values of the matrix, we can define the most energetic basis functions and select the number of basis functions required to concentrate a certain percentage of the variance of the original matrix.

We can start exploring the rank of the magnetotellurics data matrix by inspecting three cases: forward modeling on random 1D models, random smooth 1D models, and forward modeling on random smooth 2D models. The snapshot matrix for the 1D case can be computed efficiently using the Wait et al. (1964) recursion. Let us initially consider data matrices with 1 million 1D models with 90 50 m cells. Figure 6 shows the model matrices for the entirely random and smooth random cases and their corresponding basis. Considering that both snapshot matrices represent entirely different families of data (Figure 6a and 6b), a closer inspection of the resulting basis (Figure 6c and 6d) surprisingly shows that the principal modes are virtually the same. In addition, the corresponding singular values, as shown in Figure 7, demonstrate that 99.9% of the total variance of the data matrix is explained by the 10 most

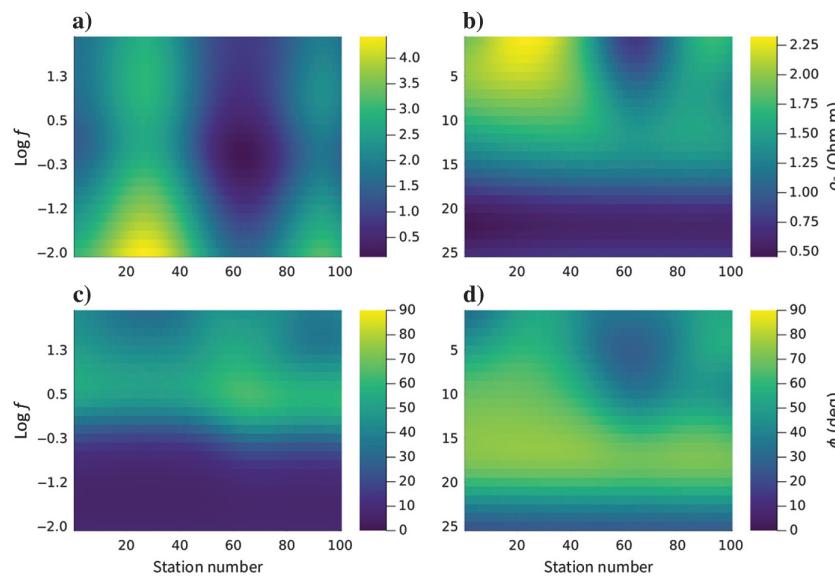


Figure 4. Forward response corresponding to training model 10,000 (Figure 1c). (a) TM apparent resistivity, (b) TE apparent resistivity, (c) TM phase, and (d) TE phase. Here, f is frequency in Hz, ρ_a is apparent resistivity, and ϕ is phase.

energetic modes of the snapshot matrix. We can conclude that both snapshot matrices can be reproduced on the same model-independent basis, and the model dependency is entirely contained in the reduced coefficients. These observations show that the 1D magnetotellurics data matrix can be considered low rank, at least for the 2 million models (1 million random and 1 million smooth random).

However, we find an even more noteworthy result when we inspect the 2D case. Figure 8 shows the singular values of decomposing the snapshot matrix of the 10,000 Latin hypercube samples extracted from our model prior. Notice how it is clear that the data space of the TE and TM modes are low rank; moreover, more than 99% of the variance of the data space is explained by the first 100 left eigenvectors (basis). Some of the most energetic bases for the 2D case are in Figures 9, 10, and 11 for apparent resistivity and phase in both propagation modes. A close inspection of mode 1 in Figure 9 shows that the basis captures the main difference in the physics of both modes. Notice that for apparent resistivity mode 1, the response can be mainly considered frequency independent. This may be due to the basis capturing the galvanic effect generated by the electric currents crossing boundaries between regions. Because galvanic effects are frequency independent, they can be observed at low frequencies. This type of physics is not present in the TE mode. In addition, a response splitting can be interpreted for low frequencies between both modes for basis 1, a behavior expected in the quarter-space experiment. Lower energy modes are more challenging to interpret. However, it is clear that those modes contain more significant wave numbers and explain local variations of the responses.

Projection operator

Following the workflow proposed by Hesthaven and Ubbiali (2018), we train a neural network to act as the projection operator (which they call the NN-POD approach). This method

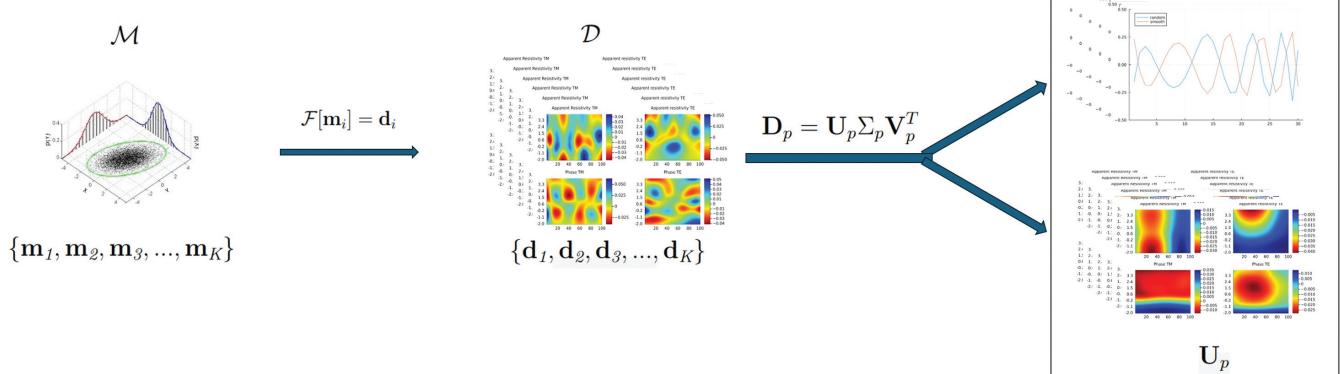


Figure 5. Diagram showing the process of building the basis. A model is extracted from a prior, and the forward-modeling operator is applied to build the snapshot matrix. The basis is built from the left eigenvector matrix obtained using the SVD.

uses a supervised learning paradigm between the subset of models $\{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \dots, \mathbf{m}_K\} \subset \mathcal{M}$ and a training set of reduced coefficients $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_K\}$. Due to the orthogonality of matrix \mathbf{U}_p , it is sufficient to project the snapshot matrix onto the basis to obtain the corresponding reduced coefficients for the training data. By projecting equation 12 onto the basis, we obtain

$$\mathbf{A} = \mathbf{U}_p^T \mathbf{D}. \quad (14)$$

In the NIRB approach, we must infer the coefficients \mathbf{a} from a resistivity model \mathbf{m} . In other words, we need to find the model parameters θ (weights and biases) for the network \mathcal{N} that minimize the cost function:

$$\hat{\theta} = \operatorname{argmin}_{\theta} \sum_j \|\mathbf{U}_p \mathbf{a}_j - \mathbf{U}_p \mathcal{N}[\theta, \mathbf{m}_j]\|_2^2. \quad (15)$$

Notice that instead of minimizing the difference between the training and predicted reduced coefficients as proposed by Hesthaven and Ubbiali (2018), our cost function is defined on the data space (after projection onto the basis). This is because small error values in low-energy modes can introduce small wavenumber anomalies in the forward modeling, even if the error is small in the least-squares sense. If the cost function is defined in the data space, we ensure that the network will find the coefficients that minimize the difference directly in the data space. After the model is properly trained, we can say that the solution to our forward problem becomes

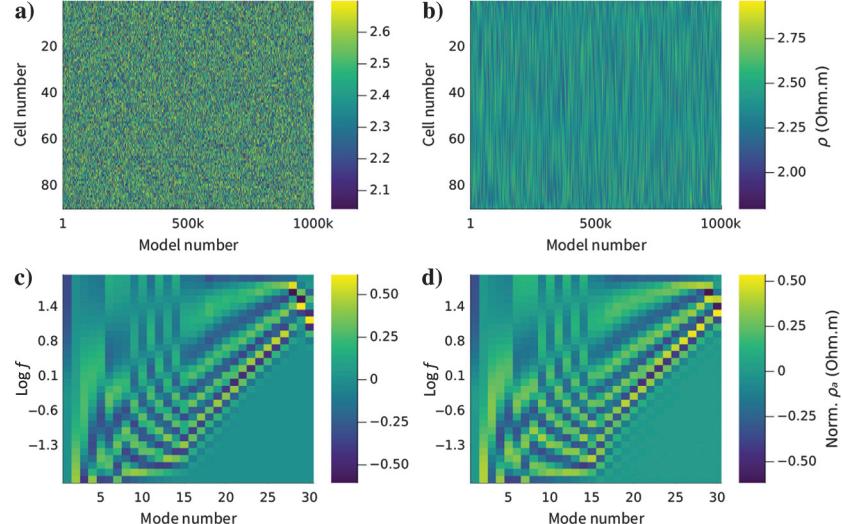
$$\hat{\mathbf{d}} = \mathbf{U}_p \mathcal{N}[\hat{\theta}, \mathbf{m}], \quad (16)$$

which should be more efficient to evaluate than $\mathcal{F}[\mathbf{m}] = \mathbf{d}$. A diagram showing the main steps to train and apply the approximated operator is shown in Figure 12.

Artificial neural networks

Deep learning is renowned for its ability to model complex nonlinear relationships (Hansen and Finlay, 2022; Liu et al., 2022; Wu et al., 2022). Given the highly nonlinear relationship between \mathbf{a} and

Figure 6. (a) A subset of random models, (b) subset of smooth random models, (c) basis functions of the snapshot matrix corresponding to (a), and (d) basis functions of the snapshot matrix corresponding to (b). Notice that (a) and (b) can be described by almost the same basis despite clear differences in model characteristics.



\mathbf{m} , we opt for a deep-learning approach as our model \mathcal{N} . We choose a shallow multilayer perceptron (MLP) for this task, prioritizing its simplicity and effectiveness in capturing the essential features necessary for mapping the entire model space to the reduced model space without overfitting. This selection balances model complexity and computational efficiency, providing a robust yet economical solution for approximating the forward operator of the Helmholtz equation in 2D magnetotelluric surveys.

The MLP architecture consists of an input layer, one hidden layer with 2048 neurons, one hidden layer with 512 neurons, and an output layer. The input layer aligns with the dimensionality of the numerical grid. The hidden layers use Gaussian error linear units (GELUs) to introduce nonlinearity, enhancing the network's ability to model the desired nonlinear relationship. The output layer corresponds to the dimensionality of the reduced coefficients vector (number of basis functions). We normalize the generated input-label pairs \mathbf{m} and \mathbf{a} to have a mean of zero and a standard deviation of one. The Adam optimizer is then used over multiple epochs to iteratively update the network's weights and biases, minimizing the

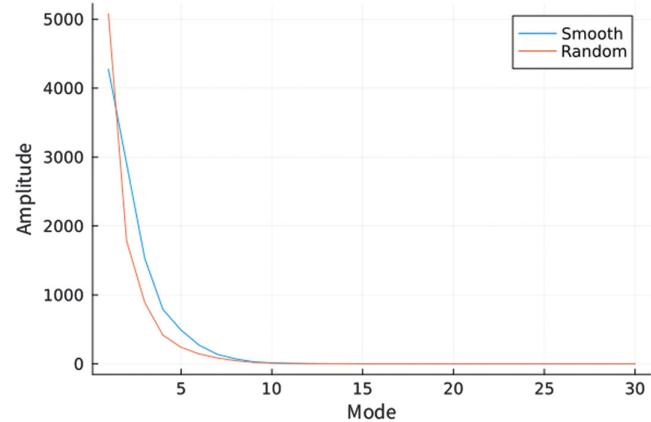


Figure 7. Singular values for the snapshot matrices shown in Figure 6a and 6b. Notice how, for both cases, the 10 most energetic modes explain almost the entire variance of the data space.

loss function defined in equation 15 and obtaining an optimal neural network $\mathcal{N}[\hat{\theta}, \mathbf{m}]$.

The MLP training uses the 100 most energetic basis functions over 500 epochs. The network performance is validated posttraining using a separate validation set that was not involved in the training process. Error metrics with the Huber norm are applied to the training cost function to promote robust handling of outliers.

The NIRB data as regularization

We can balance computational gains and desired accuracy by using the NIRB solution as a regularization instead of relying on it as the final result of the modeling. Let us consider that we have access to a minimal set of high-fidelity observations and know the RB; we can propose a least-squares solution to the coefficients that satisfies these observations using the NIRB results as regularization. A similar technique has recently been applied to reconstruct flow and heat fields by Zhao et al. (2024) in the context of gappy POD.

Consider the following inverse problem:

$$\hat{\mathbf{a}} = \operatorname{argmin}_{\mathbf{a}} \{ \| \mathbf{T} \mathbf{U} \mathbf{a} - \mathbf{d} \|_2^2 + \lambda \| \mathbf{a} - \mathcal{N}[\hat{\theta}, \mathbf{m}] \|_2^2 + \mu \| \mathbf{D} \mathbf{U} \mathbf{a} \|_2^2 \}, \quad (17)$$

where \mathbf{T} is a sampling operator that specifies the available high-fidelity frequencies and \mathbf{D} is a first derivative operator to enforce smoothness on the data space. The solution to this inverse problem is

$$\hat{\mathbf{a}} = (\mathbf{U}^T \mathbf{T}^T \mathbf{T} \mathbf{U} + \lambda \mathbf{I} + \mu \mathbf{U}^T \mathbf{D}^T \mathbf{D} \mathbf{U})^{-1} (\mathbf{U}^T \mathbf{T}^T \mathbf{d} + \lambda \mathcal{N}[\hat{\theta}, \mathbf{m}]). \quad (18)$$

Notice that as long as the basis and the sampling operator are not modified, the Hessian is the same for repeated evaluations, and only the right side of the expression changes (the high-fidelity observations and the output of the NIRB operator). This allows further computational improvement by performing Cholesky decomposition (Benoit, 1924). Consider the following:

$$\mathbf{A} = \mathbf{U}^T \mathbf{T}^T \mathbf{T} \mathbf{U} + \lambda \mathbf{I} + \mu \mathbf{U}^T \mathbf{D}^T \mathbf{D} \mathbf{U}, \quad (19)$$

$$\mathbf{z} = \mathbf{U}^T \mathbf{T}^T \mathbf{d} + \lambda \mathcal{N}[\hat{\theta}, \mathbf{m}]. \quad (20)$$

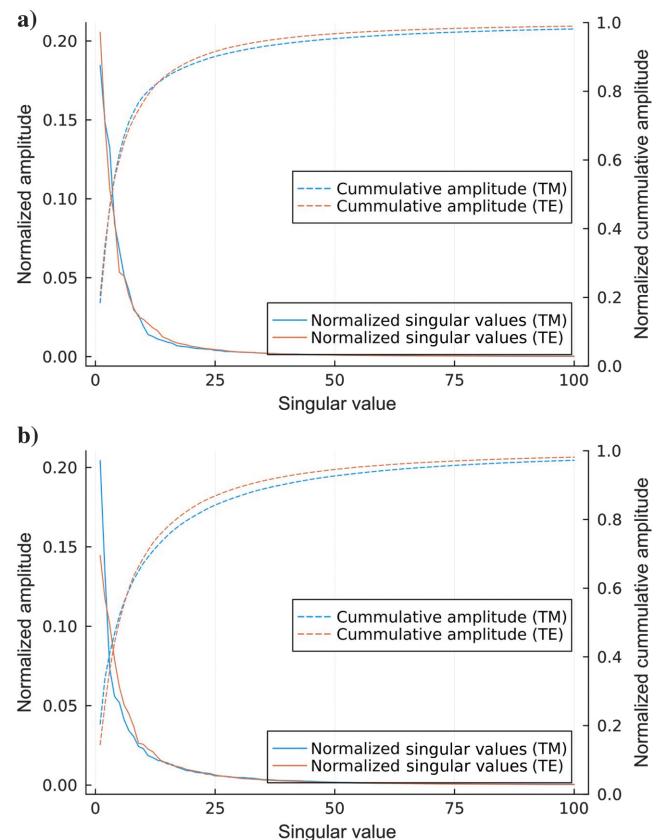


Figure 8. Normalized singular values and cumulative amplitude for (a) the apparent resistivity and (b) phase for snapshot matrix of the 2D magnetotellurics case. Notice how the singular values of the first 100 most energetic modes explain almost the entire data variance.

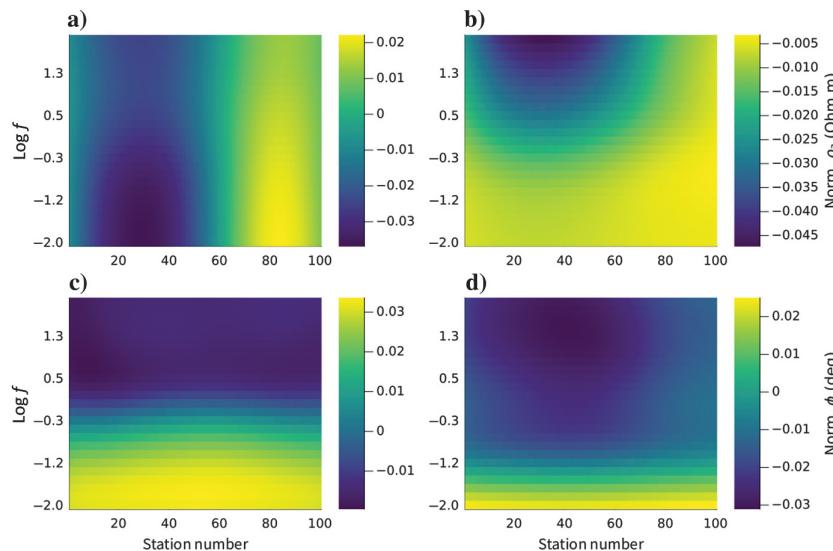


Figure 9. The first mode of the 2D RB. (a) Apparent resistivity TM, (b) apparent resistivity TE, (c) phase TM, and (d) phase TE.

This means that we are interested in solving the linear system of equations $\mathbf{A}\hat{\mathbf{a}} = \mathbf{z}$. Cholesky decomposition allows us to decompose a symmetric positive definite matrix into the multiplications of a lower triangular matrix and its transpose ($\mathbf{A} = \mathbf{LL}^T$). To find \mathbf{A}^{-1} , we can introduce an auxiliary vector \mathbf{y} and find \mathbf{L}^{-1} by solving the following linear system of equations by forward substitution:

$$\mathbf{Ly} = \mathbf{z}. \quad (21)$$

Then, we find $(\mathbf{L}^T)^{-1}$ solving the following linear system of equations by backward substitution:

$$\mathbf{L}^T \hat{\mathbf{a}} = \mathbf{y}. \quad (22)$$

The computational cost of inverting a matrix using Gaussian elimination is $O(n^3)$, whereas the computational cost of performing Cholesky decomposition is $(1/3)O(n^3)$, and then calculating \mathbf{A}^{-1} becomes $O(n^2)$. Notice that the decomposition is only done once;

Figure 10. The 10th mode of the 2D RB. (a) Apparent resistivity TM, (b) apparent resistivity TE, (c) phase TM, and (d) phase TE.

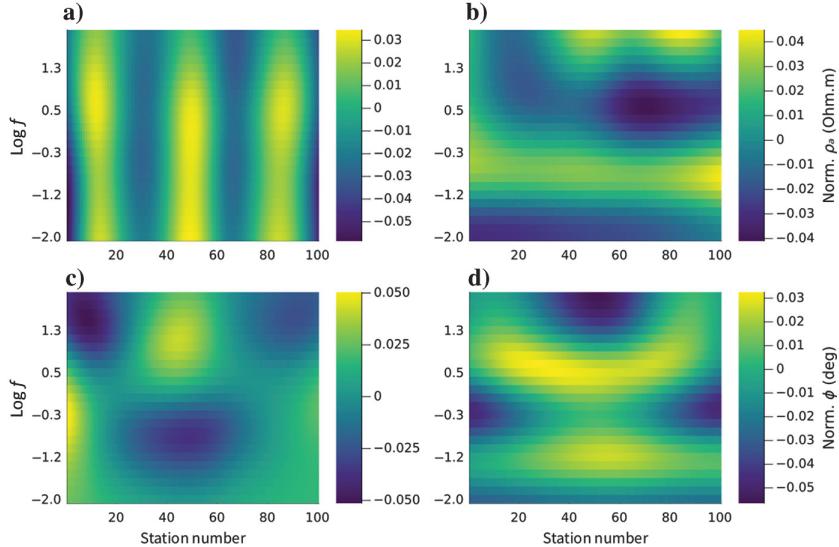
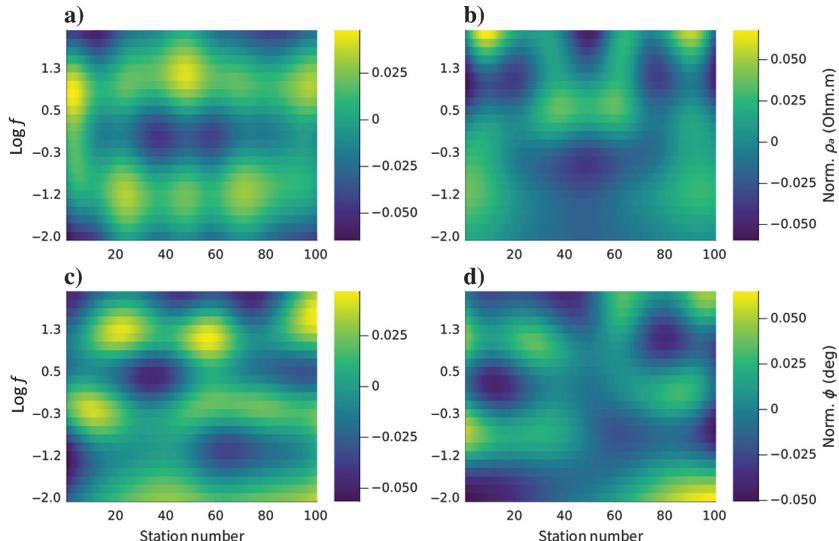


Figure 11. The 20th mode of the 2D RB. (a) Apparent resistivity TM, (b) apparent resistivity TE, (c) phase TM, and (d) phase TE.



therefore, the final computational cost for repeated evaluations is $O(n^2)$ for each iteration.

This strategy makes this inverse problem computationally inexpensive because it involves solving a system of size $k \times k$, where k is the number of basis functions (a small number in this study). Solving this inverse problem translates into finding the coefficients that fit a few high-fidelity observations while remaining close to the neural network's output, with an extra regularization term that keeps the data space low rank by enforcing smoothness (the primary assumption of this paper). This approach allows for significant computational savings, reducing the need to generate high-fidelity observations for all frequencies, requiring forward modeling only for a select few. We will refer to this approach as hybrid-NIRB.

EXPERIMENTS

The offline phase

The computational burden of the offline phase is mainly related to three stages: the snapshot matrix building, the basis building

via SVD, and the neural network training. From these three stages, the snapshot matrix is by far the most expensive. This study generated two snapshot matrices with 10,000 forward evaluations (for TM and TE modes). Each forward model evaluation of the FE solver takes 35 s on an Intel Core i5-14400 workstation for a total running time of approximately 97 h. The final size for each snapshot matrix is 5000×10000 (containing apparent resistivity and phase). The basis generation involves calculating four bases, one for apparent resistivity and one for phase, for each propagation mode, with a total running time of 14.4 s. We train the neural networks across 500 epochs; the computational times are 3.1 h for each network and 12.4 h for both propagation modes. We can also make a naïve estimation of the cost of building the snapshot matrix for the 3D case. Solving a PDE with FE involves two principal stages: assembling the stiffness matrix and solving the linear system of equations. Assembling the stiffness matrix involves a small fixed number of operations with cost $O(1)$, which makes the total cost $O(N)$. The number of elements for the hypothetical 3D case is 2,151,680 ($164 \times 164 \times 80$), whereas, for the 2D case, it is 13,120 elements (164×80), making the stiffness matrix assembly roughly 162 times more expensive. With iterative solvers, each iteration has an approximate computational cost of $O(N)$ because the dominant computation is a matrix-vector multiplication; therefore, for 100 internal iterations, the computational cost of finding the solution is approximately 330 times more in the 3D case. Notice that this estimation is based on a naive implementation, and real 3D implementation would require advanced computational tools, such as graphics processing units (GPUs) and parallelization.

Architecture selection

The architecture selection process involves evaluating various hyperparameters to optimize the model's generalization performance, quantified through mean error, maximum error, training loss, and validation loss. Through systematic hyperparameter optimization (detailed in Appendix A), we identify 20% dropout rate and GELU activation paired with Huber norm as the optimal configuration for outlier robustness. The final architecture comprises an input layer, a hidden layer with 2048 neurons, a subsequent hidden layer with 512 neurons, and an output layer. This optimized architecture demonstrates stable convergence across TE and TM modes, as evidenced by the training curves in Figures 13 and 14. Notably, the validation loss is typically larger than the training loss, indicating some underfitting, particularly for the TM mode. This underfitting likely results from the straightforward architecture of the proposed neural network, suggesting untapped potential for advanced architectures. Although more complex models such as generative adversarial network or diffusion models may offer superior performance, this study aims to establish a simple baseline prioritizing computational transparency over complexity. Readers are encouraged to consult Appendix A for a complete parameter selection methodology and comparative tables.

Forward modeling

This section compares the results from the high-fidelity solution (FE) and those of the NIRB and hybrid-NIRB operators. Two hundred entirely random models are generated and modeled using an FE high-fidelity operator. For this validation data set, the models are altogether

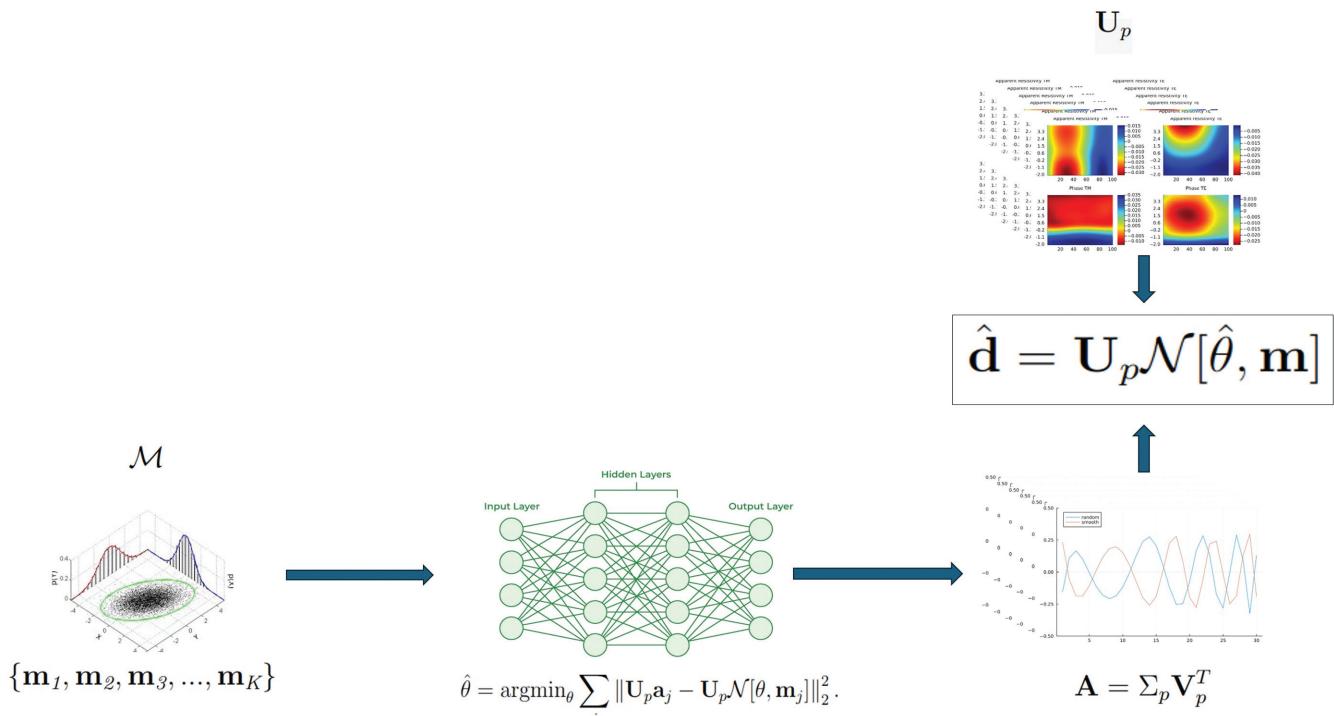


Figure 12. Diagram showing the nonintrusive RB forward modeling. The model is the input of the trained neural network, which predicts the reduced coefficients. Finally, the data are obtained by multiplying the reduced coefficients and the basis.

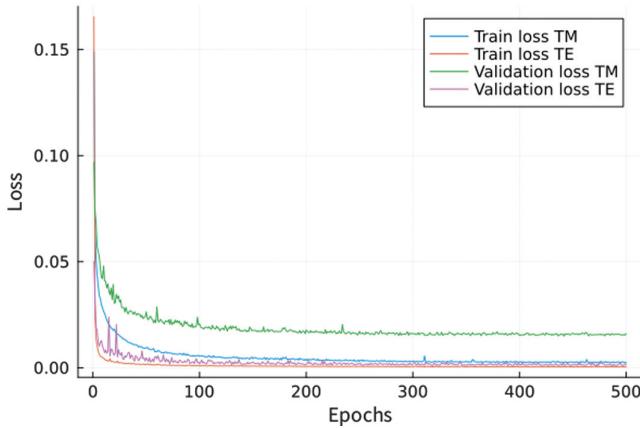


Figure 13. Training and validation loss for the final proposed architecture for apparent resistivity.

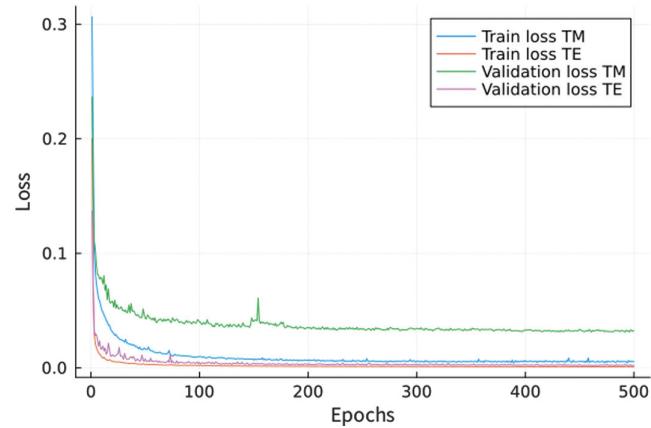


Figure 14. Training and validation loss for the final proposed architecture for phase.

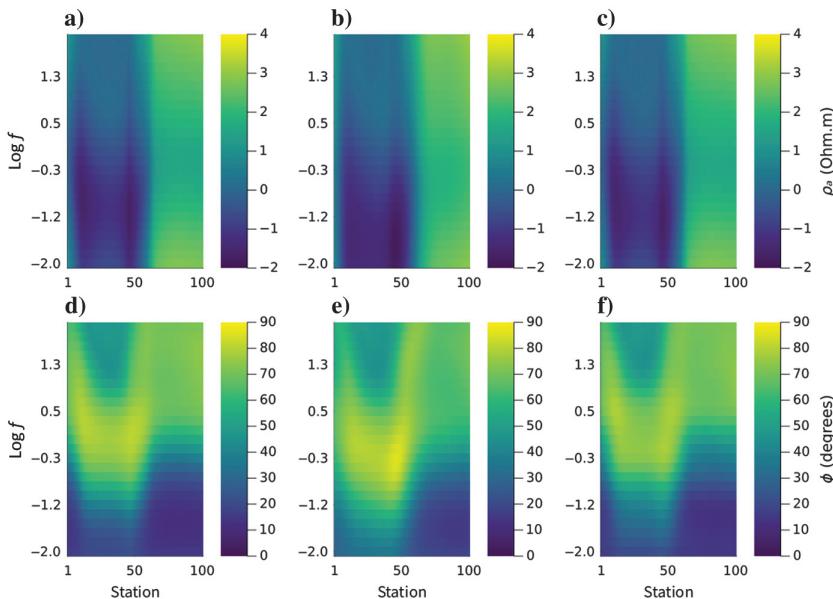


Figure 15. The TM forward response for the validation model 1. (a) High-fidelity apparent resistivity, (b) NIRB apparent resistivity (error 18.8%), (c) hybrid-NIRB apparent resistivity (error 2.1%), (d) high-fidelity phase, (e) NIRB phase (error 16.7%), and (f) hybrid-NIRB phase (error 3.8%).

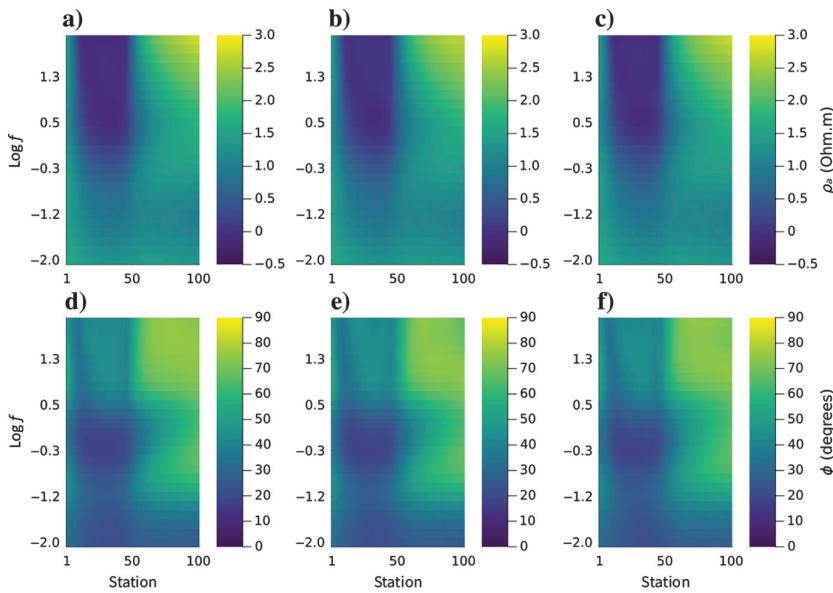


Figure 16. The TE forward response for the validation model 1. (a) High-fidelity apparent resistivity, (b) NIRB apparent resistivity (error, 3.1%), (c) hybrid-NIRB apparent resistivity (error, 1.9%), (d) high-fidelity phase, (e) NIRB phase (error, 2.4%), and (f) hybrid-NIRB phase (error, 1.9%).

random to avoid bias related to sampling the training data set used to build the basis (Latin hypercube sampling). Each NIRB forward calculation is evaluated in approximately 4 ms, as opposed to the 35 s on average for the FE numerical solution, reducing computational costs by three orders of magnitude. The median errors over the validation data set are 10% for apparent resistivity, 9% for phase on the TM mode, 2% for apparent resistivity, and 2% for phase on the TE mode. The error values are relatively close to the error thresholds accepted by Manassero et al. (2020), which is 7%, showing that it is possible to get close to those error levels with a nonintrusive workflow. One advantage of the method proposed by Manassero et al. (2020) is a deterministic way to assess the accuracy of the RB approximation without requiring a high-fidelity solution, which gives a strategy to decide when to enrich the basis. This deterministic assessment needs to be developed in our method and is the objective of current research.

Results also show that, in general, the evaluation of the TE mode yields lower values of error compared with the TM mode. The per-

formance gap between TE and TM modes reflects fundamental differences in their governing physics. TE mode's smoother parameter space enables robust NIRB approximations. In contrast, the TM numerical solution is more computationally complicated and more difficult to capture in the reduced coefficient space, necessitating hybrid correction. In addition, phase prediction usually yields smaller errors than apparent resistivity for TE and TM cases. When the hybrid-NIRB case is considered, further improvement is obtained. In some cases, error percentages are decreased by one order of magnitude, and the improvement is particularly noticeable for the TM mode.

For the hybrid-NIRB experiments, the observed data are modeled frequencies using the high-fidelity operator. In the examples shown, we initially run the forward operator for only five frequencies of the 35 originally modeled. By solving this small inverse problem, we obtain the set of coefficients that solve for all the frequencies, minimizing the misfit with the observations at the high-fidelity

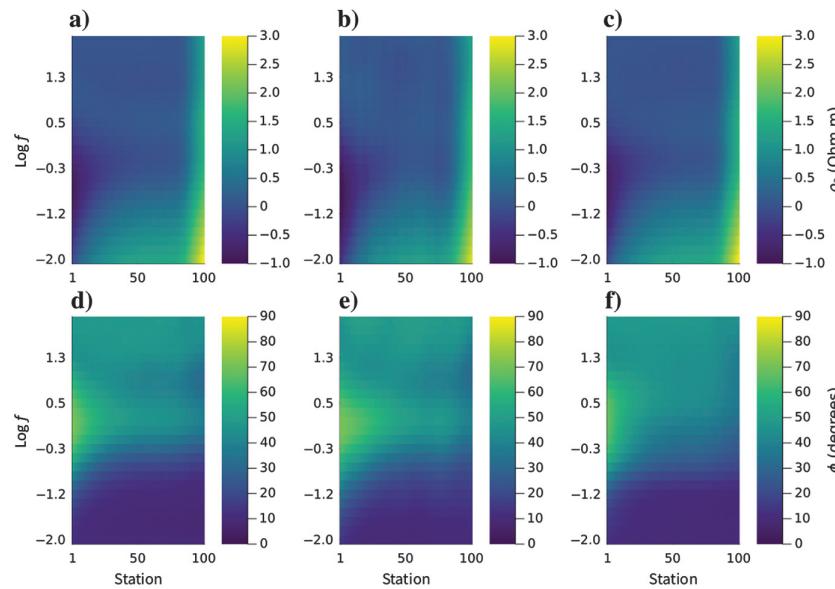


Figure 17. The TM forward response for the validation model 100. (a) High-fidelity apparent resistivity, (b) NIRB apparent resistivity (error 14.5%), (c) hybrid-NIRB apparent resistivity (error 3%), (d) high-fidelity phase, (e) NIRB phase (error 11.2%), and (f) hybrid-NIRB phase (error 6.3%).

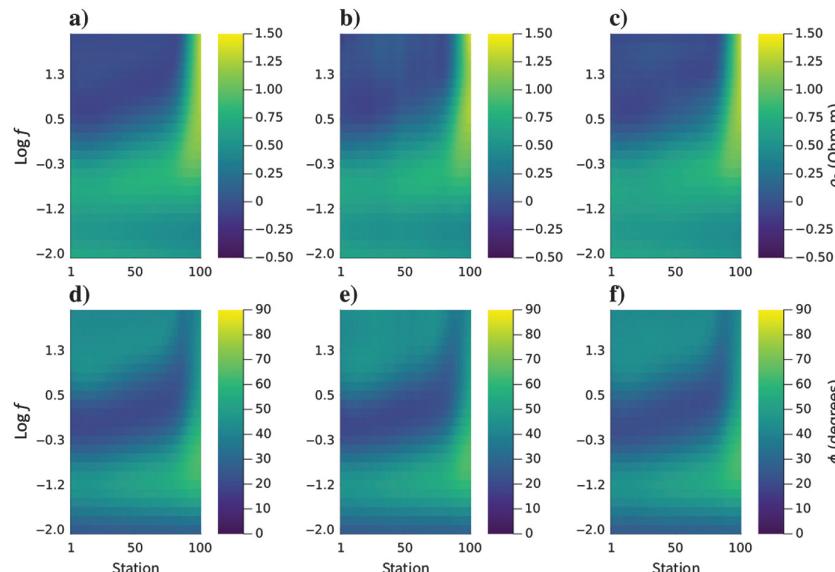


Figure 18. The TE forward response for the validation model 100. (a) High-fidelity apparent resistivity, (b) NIRB apparent resistivity (error 7.3%), (c) hybrid-NIRB apparent resistivity (error 5.1%), (d) high-fidelity phase, (e) NIRB phase (error 2%), and (f) hybrid-NIRB phase (error 1.9%).

frequencies. For the high-fidelity solution of five frequencies, we require only 14% of the original computational requirements, and the error reduces to an average of 1% on apparent resistivity and 3% for phase on the TM mode and 1.4% for apparent resistivity and 1.6% for phase on the TE mode. The hybrid-NIRB method is good at handling outliers, making a maximum error of 9% drop to 5% on the TE mode and a maximum error of 51% drop to 7.9% for the TM mode. In Figures 15, 16, 17, 18, 19, and 20, we show the comparisons for the high-fidelity solution, the NIRB solution, and the hybrid-NIRB solution for validation models 1, 100, and 200. Error distribution analysis in Figures 21 and 22 confirms these trends: whereas NIRB alone suffices for TE mode approximations, hybrid-NIRB becomes essential for TM mode accuracy. Detailed comparisons of forward responses across validation models (1, 100, 200) at stations 25 and 75 are provided in supplemental file S1. No clear frequency-dependent error patterns emerge, suggesting

model complexity, rather than measurement characteristics, drives approximation challenges. Extreme conductivity contrasts or specific anomaly geometries appear to disproportionately affect TM mode responses, though systematic error sources require further investigation.

Although our nonintrusive workflow matches the accuracy thresholds of intrusive methods, three key directions emerge for future research: (1) development of TM-optimized snapshot matrices using adaptive sampling strategies, (2) implementation of ensemble methods to address initialization sensitivity, and (3) online updating frameworks combining basis and operator refinement during field deployments. Computational enhancements through GPU acceleration and parallelization could further bridge the performance gap between modes. More details to bridge the performance gap between TE and TM modes while maintaining the method's nonintrusive advantages can be found in supplemental file S1.

Figure 19. The TM forward response for the validation model 200. (a) High-fidelity apparent resistivity, (b) NIRB apparent resistivity (error 8.4%), (c) hybrid-NIRB apparent resistivity (error 1.5%), (d) high-fidelity phase, (e) NIRB phase (error 6.3%), and (f) hybrid-NIRB phase (error 4%).

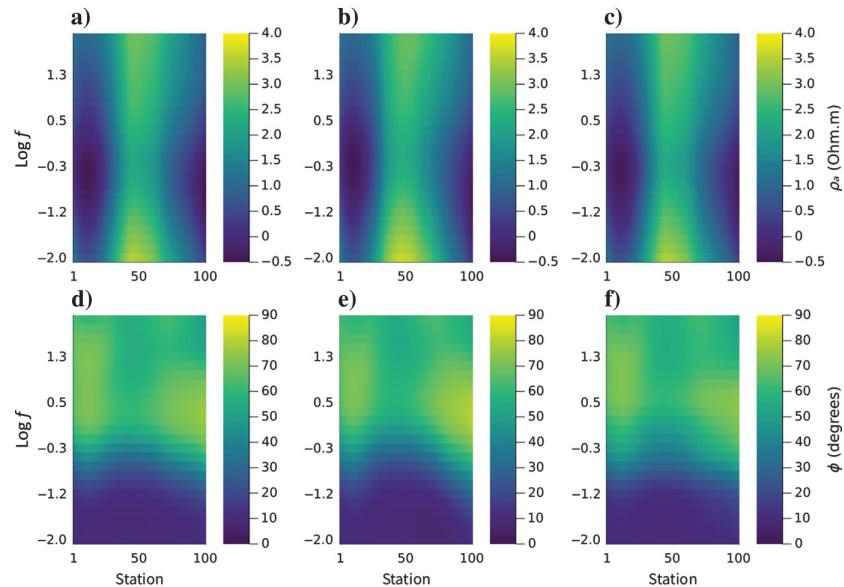
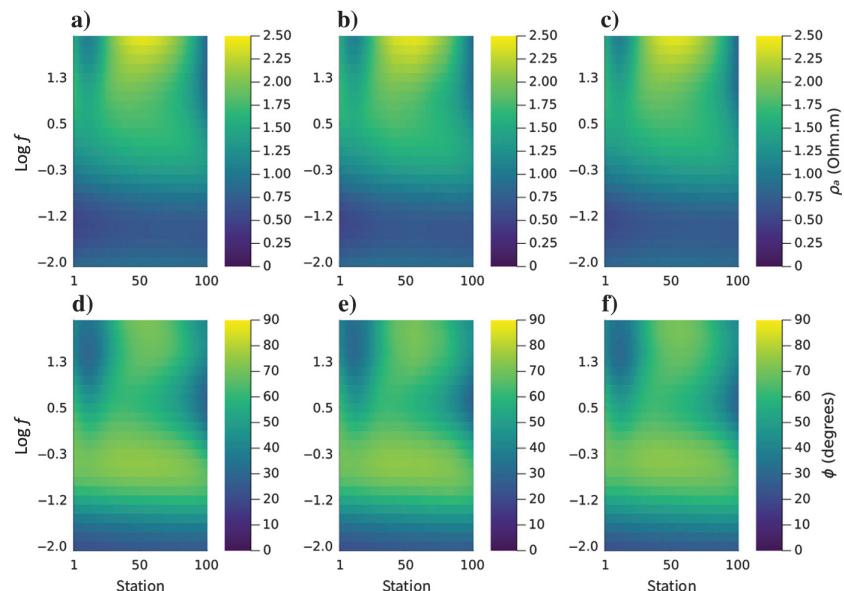


Figure 20. The TE forward response for the validation model 200. (a) High-fidelity apparent resistivity, (b) NIRB apparent resistivity (error 1.5%), (c) hybrid-NIRB apparent resistivity (error 1%), (d) high-fidelity phase, (e) NIRB phase (error 1.5%), and (f) hybrid-NIRB phase (error 1.7%).



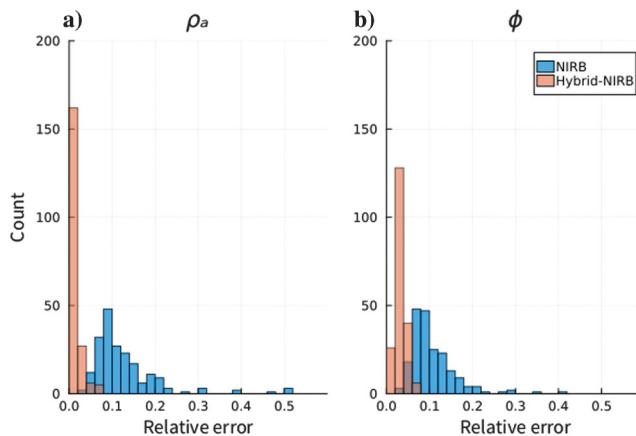


Figure 21. Comparison of error for the NIRB and hybrid-NIRB approaches on the testing data sets. (a) Histogram of median error on apparent resistivity, TM mode and (b) histogram of median error on phase, TM mode.

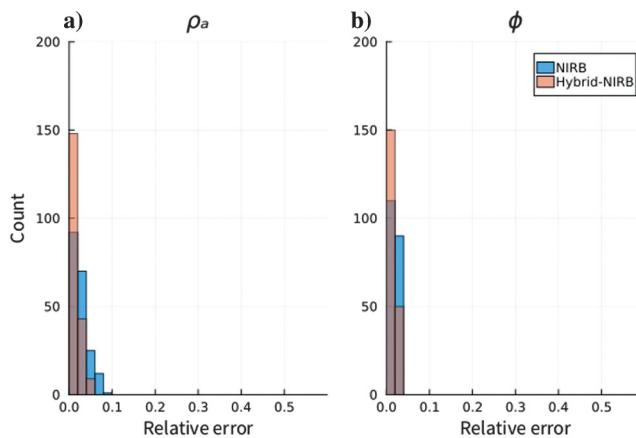


Figure 22. Comparison of error for the NIRB and hybrid-NIRB approaches on the testing data sets. (a) Histogram of median error on apparent resistivity, TE mode and (b) histogram of median error on phase, TE mode.

CONCLUSION

We have presented the first approximation to the solution of the Helmholtz equation using an NIRB solution for magnetotellurics forward modeling. Unlike conventional RB workflows, explicit access to the stiffness matrix is unnecessary. Moreover, the proposed workflow is compatible with any numerical solver. An offline stage of four days was used to create a meaningful basis. The analysis of the singular values of the snapshot matrices in the 1D and 2D cases shows that the data matrix of the magnetotellurics method can be effectively represented using a low-rank approximation. This basis could be used to train the projection operators. In the online stage, a single evaluation of the forward model takes 4 ms versus 35 s of running the high-fidelity solution, evaluating the model almost four orders of magnitude faster. The median errors for the NIRB solution are 10% for the apparent resistivity, 9% for the phase in the TM mode, 2% in apparent resistivity, and 2% for the phase in the TE mode. Better levels of accuracy are possible by sacrificing some computational time by using the hybrid-NIRB approach because high-fidelity forward modeling is

only required for a subset of frequencies. The advantage of hybrid-NIRB is that it does not rely entirely on the neural projection operator, which makes the result more accurate, and this accuracy depends on the number of high-fidelity observations considered in the inverse problem. Error values of 1% are reached for the apparent resistivity and 3% for phase on the TM mode, but error values were 1.4% for apparent resistivity and 1.6% for phase on the TE mode. The gains in computational time can facilitate extensive uncertainty quantification studies, Monte Carlo simulations, or Bayesian inversion approaches, which are generally prohibitively expensive. The method can be extended beyond magnetotellurics forward modeling. It could have applications in other geophysical problems with similar PDEs, such as seismic wave propagation, groundwater flow, or controlled-source electromagnetics, that could benefit from faster evaluations.

ACKNOWLEDGMENTS

We gratefully acknowledge the sponsors of the Signal Analysis and Imaging Group at the University of Alberta for their generous support. We also sincerely thank the Exploration Team at YPFB-Chaco for all the fruitful discussions.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be accessed via the following GitHub repository: <https://github.com/quiaro21/NIRB>. This repository is actively maintained and subject to updates as the study progresses. The snapshot matrix was built using SimPEG (<https://github.com/simpeg>). For further details or clarifications, please contact the corresponding author.

APPENDIX A

ARCHITECTURE SELECTION METHODOLOGY

The architecture selection is based on calculating each model's mean error, maximum error, training loss, and validation loss. Fifteen different MLP architectures were tested across 50 epochs, changing parameters such as network depth, network width, batch normalization, optimizer, dropout rate, and activation function. However, we find that the most influential parameters that help improve the model's generalization are the dropout rate, data misfit norm, and the activation function. In the first test, we select the baseline architecture with a GELU activation function, Huber loss, while testing various dropout rates.

The results in Table A-1 show that a dropout rate of 20% provides the best baseline architecture results. It achieves a balanced trade-off

Table A-1. Comparison of different dropout rates on approximation errors.

Dropout rate (%)	Training loss	Validation loss	Mean error (%)	Maximum error (%)
15	0.0077	0.027	13	100.2
20	0.0086	0.026	13.5	79
25	0.0089	0.025	13.4	87
30	0.0093	0.026	13.4	87

Bold values are the best results.

Table A-2. Comparison of activation functions on approximation errors.

Activation function	Training loss	Validation loss	Mean error (%)	Maximum error (%)
ReLU	0.0159	0.025	13.3	93
GELU	0.0086	0.026	13.5	79
Tanh	0.016	0.035	16	85

Bold values are the best results.

Table A-3. Comparison of data misfit norms on approximation errors.

Data misfit norm	Training loss	Validation loss	Mean error (%)	Maximum error (%)
\mathcal{L}_1	0.085	0.026	13.5	79
\mathcal{L}_2	0.017	0.05	13	75
Huber	0.0086	0.026	13.5	79

Bold values are the best results.

between regularization and network capacity, providing the most robust result. In contrast, other dropout rates lead to higher maximum errors and less robust performance against outliers.

Moreover, we test different activation functions, such as the rectified linear unit (ReLU), GELU, and the hyperbolic tangent. As shown in Table A-2, the GELU activation function drastically reduces the overfitting, making the training and validation losses closer and further decreasing the error of the outliers. This results in a marked reduction in maximum errors compared with ReLU and Tanh, highlighting GELU's capacity to maintain smooth gradient flow and enhance overall model generalization.

A final test was carried out to change the norm of the data misfit (Table A-3). We test the model's performance using the \mathcal{L}_2 norm, the \mathcal{L}_1 norm, and the Huber norm to make the weight inversion more robust. Although the training and validation loss cannot be directly compared between different norms (because they are calculated differently), we see that \mathcal{L}_1 norm and Huber norm are producing comparable results in terms of the median and the maximum error. However, the \mathcal{L}_2 norm produces slightly better handling of outliers, reducing the maximum error an extra 4%. However, when trained for more epochs, using the Huber norm produced a better handling of outliers, giving the smallest maximum error. The Huber norm reduces overfitting by approximating an \mathcal{L}_2 norm for small misfit and an \mathcal{L}_1 for large values of misfit, providing a robust handling of outliers. Therefore, the final network is trained for a total of 500 epochs using a Huber norm, which also demonstrates that longer training improves the generalization ability of the network.

REFERENCES

- Benner, P., S. Gugercin, and K. Willcox, 2015, A survey of projection-based model reduction methods for parametric dynamical systems: SIAM Review, **57**, 483–531, doi: [10.1137/130932715](https://doi.org/10.1137/130932715).
- Benoit, C., 1924, Note sur une méthode de résolution des équations normales provenant de l'application de la méthode des moindres carrés à un système d'équations linéaires en nombre inférieur à celui des incon-

nues (Procédé du Commandant Cholesky): Bulletin Géodésique, **2**, 67–77, doi: [10.1007/BF03031308](https://doi.org/10.1007/BF03031308).

Berdichevsky, M., V. Dmitriev, and E. Pozdnjakova, 1998, On two-dimensional interpretation of magnetotelluric soundings: Geophysical Journal International, **133**, 585–606, doi: [10.1046/j.1365-246X.1998.01333.x](https://doi.org/10.1046/j.1365-246X.1998.01333.x).

Berkooz, G., P. Holmes, and J. L. Lumley, 1993, The proper orthogonal decomposition in the analysis of turbulent flows: Annual Review of Fluid Mechanics, **25**, 539–575, doi: [10.1146/annurev.fl.25.010193.002543](https://doi.org/10.1146/annurev.fl.25.010193.002543).

Castillo-Reyes, O., J. de la Puente, and J. M. Cela, 2018, PETGEM: A parallel code for 3D CSEM forward modeling using edge finite elements: Computers & Geosciences, **119**, 123–136, doi: [10.1016/j.cageo.2018.07.005](https://doi.org/10.1016/j.cageo.2018.07.005).

Cockett, R., S. Kang, L. J. Heagy, A. Pidlisecky, and D. W. Oldenburg, 2015, SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications: Computers & Geosciences, **85**, 142–154, doi: [10.1016/j.cageo.2015.09.015](https://doi.org/10.1016/j.cageo.2015.09.015).

Dai, H., 2019, A review on the exact Monte Carlo simulation, in Bayesian inference on complicated data: IntechOpen, doi: [10.5772/intechopen.88619](https://doi.org/10.5772/intechopen.88619).

Degen, D., D. Caviedes Voullième, S. Buitre, H.-J. Hendricks Franssen, H. Vereecken, A. González-Nicolás, and F. Wellmann, 2023, Perspectives of physics-based machine learning strategies for geoscientific applications governed by partial differential equations: Geoscientific Model Development, **16**, 7375–7409, doi: [10.5194/gmd-16-7375-2023](https://doi.org/10.5194/gmd-16-7375-2023).

Deng, F., J. Hu, X. Wang, S. Yu, B. Zhang, S. Li, and X. Li, 2023, Magnetotelluric deep learning forward modeling and its application in inversion: Remote Sensing, **15**, 3667, doi: [10.3390/rs15143667](https://doi.org/10.3390/rs15143667).

Durran, D. R., 2013, Numerical methods for wave equations in geophysical fluid dynamics: Springer Science and Business Media.

Eckart, C., and G. Young, 1936, The approximation of one matrix by another of lower rank: Psychometrika, **1**, 211–218, doi: [10.1007/BF02288367](https://doi.org/10.1007/BF02288367).

Fresca, S., L. Dede', and A. Manzoni, 2021, A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs: Journal of Scientific Computing, **87**, 1–36, doi: [10.1007/s10915-021-01462-7](https://doi.org/10.1007/s10915-021-01462-7).

Hansen, T. M., and C. C. Finlay, 2022, Use of machine learning to estimate statistics of the posterior distribution in probabilistic inverse problems — An application to airborne EM data: Journal of Geophysical Research: Solid Earth, **127**, e2022JB024703, doi: [10.1029/2022JB024703](https://doi.org/10.1029/2022JB024703).

Heagy, L. J., R. Cockett, S. Kang, G. K. Rosenkjaer, and D. W. Oldenburg, 2017, A framework for simulation and inversion in electromagnetics: Computers & Geosciences, **107**, 1–19, doi: [10.1016/j.cageo.2017.06.018](https://doi.org/10.1016/j.cageo.2017.06.018).

Hesthaven, J. S., and S. Ubbiali, 2018, Non-intrusive reduced order modeling of nonlinear problems using neural networks: Journal of Computational Physics, **363**, 55–78, doi: [10.1016/j.jcp.2018.02.037](https://doi.org/10.1016/j.jcp.2018.02.037).

Jahandari, H., and C. Farquharson, 2015, Finite-volume modelling of geophysical electromagnetic data on unstructured grids using potentials: Geophysical Journal International, **202**, 1859–1876, doi: [10.1093/gji/ggv257](https://doi.org/10.1093/gji/ggv257).

Kadeethum, T., F. Ballarin, Y. Choi, D. O'Malley, H. Yoon, and N. Bouklas, 2022, Non-intrusive reduced order modeling of natural convection in porous media using convolutional autoencoders: Comparison with linear subspace techniques: Advances in Water Resources, **160**, 104098, doi: [10.1016/j.advwatres.2021.104098](https://doi.org/10.1016/j.advwatres.2021.104098).

Kahlbacher, M., and S. Volkwein, 2007, Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems: Discussiones Mathematicae, Differential Inclusions, Control and Optimization, **27**, 95–117, doi: [10.7151/dmdico.1078](https://doi.org/10.7151/dmdico.1078).

Kunisch, K., and S. Volkwein, 2002, Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics: SIAM Journal on Numerical Analysis, **40**, 492–515, doi: [10.1137/S0036142900382612](https://doi.org/10.1137/S0036142900382612).

LeVeque, R. J., 2007, Finite difference methods for ordinary and partial differential equations: Steady-state and time-dependent problems: SIAM.

Li, K., T.-Z. Huang, L. Li, and S. Lanteri, 2021, Non-intrusive reduced-order modeling of parameterized electromagnetic scattering problems using cubic spline interpolation: Journal of Scientific Computing, **87**, 52, doi: [10.1007/s10915-021-01467-2](https://doi.org/10.1007/s10915-021-01467-2).

Liu, D., X. Wang, X. Yang, H. Mao, M. D. Sacchi, and W. Chen, 2022, Accelerating seismic scattered noise attenuation in offset-vector tile domain: Application of deep learning: Geophysics, **87**, no. 5, V505–V519, doi: [10.1190/geo2021-0654.1](https://doi.org/10.1190/geo2021-0654.1).

Manassero, M. C., J. C. Afonso, F. Zyberman, S. Zlotnik, and I. Fomin, 2020, A reduced order approach for probabilistic inversions of 3-D magnetotelluric data I: General formulation: Geophysical Journal International, **223**, 1837–1863, doi: [10.1093/gji/ggaa415](https://doi.org/10.1093/gji/ggaa415).

McKay, M. D., 1992, Latin hypercube sampling as a tool in uncertainty analysis of computer models: Proceedings of the 24th Conference on Winter Simulation, 557–564.

McKay, M. D., R. J. Beckman, and W. J. Conover, 2000, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code: Technometrics, **42**, 55–61, doi: [10.1080/00401706.2000.10485979](https://doi.org/10.1080/00401706.2000.10485979).

Peterson, J. S., 1989, The reduced basis method for incompressible viscous flow calculations: SIAM Journal on Scientific and Statistical Computing, **10**, 777–786, doi: [10.1137/0910047](https://doi.org/10.1137/0910047).

- Plessix, R.-E., 2007, A Helmholtz iterative solver for 3D seismic-imaging problems: *Geophysics*, **72**, no. 5, SM185–SM194, doi: [10.1190/1.2738849](https://doi.org/10.1190/1.2738849).
- Ramnath, V., 2023, Monte Carlo simulations in uncertainty evaluation for partial differential equations, in *Handbook of metrology and applications*: Springer, 1–23, doi: [10.1007/978-981-19-1550-5_124-1](https://doi.org/10.1007/978-981-19-1550-5_124-1).
- Ray, A., and D. Myer, 2019, Bayesian geophysical inversion with trans-dimensional Gaussian process machine learning: *Geophysical Journal International*, **217**, 1706–1726, doi: [10.1093/gji/ggz111](https://doi.org/10.1093/gji/ggz111).
- Rekoske, J. M., A.-A. Gabriel, and D. A. May, 2023, Instantaneous physics-based ground motion maps using reduced-order modeling: *Journal of Geophysical Research: Solid Earth*, **128**, e2023JB026975, doi: [10.1029/2023JB026975](https://doi.org/10.1029/2023JB026975).
- Renardy, M., L. R. Joslyn, J. A. Millar, and D. E. Kirschner, 2021, To Sobol or not to Sobol? The effects of sampling schemes in systems biology applications: *Mathematical biosciences*, **337**, 108593, doi: [10.1016/j.mbs.2021.108593](https://doi.org/10.1016/j.mbs.2021.108593).
- Rochlitz, R., N. Skibbe, and T. Günther, 2019, custEM: Customizable finite-element simulation of complex controlled-source electromagnetic data: *Geophysics*, **84**, no. 2, F17–F33, doi: [10.1190/geo2018-0208.1](https://doi.org/10.1190/geo2018-0208.1).
- Signorini, M., S. Zlotnik, and P. Diez, 2017, Proper generalized decomposition solution of the parameterized Helmholtz problem: Application to inverse geophysical problems: *International Journal for Numerical Methods in Engineering*, **109**, 1085–1102, doi: [10.1002/nme.5313](https://doi.org/10.1002/nme.5313).
- Smith, G. D., 1985, Numerical solution of partial differential equations: Finite difference methods: Oxford University Press.
- Smith, T. R., J. Moehlis, and P. Holmes, 2005, Low-dimensional modelling of turbulence using the proper orthogonal decomposition: A tutorial: *Nonlinear Dynamics*, **41**, 275–307, doi: [10.1007/s11071-005-2823-y](https://doi.org/10.1007/s11071-005-2823-y).
- Song, Z.-M., and P. R. Williamson, 1995, Frequency-domain acoustic-wave modeling and inversion of crosshole data: Part I — 2.5-D modeling method: *Geophysics*, **60**, 784–795, doi: [10.1190/1.1443817](https://doi.org/10.1190/1.1443817).
- Strang, G., and G. Fix, 1973, An analysis of the finite element method: Prentice-Hall.
- Unsworth, M. J., B. J. Travis, and A. D. Chave, 1993, Electromagnetic induction by a finite electric dipole source over a 2-D earth: *Geophysics*, **58**, 198–214, doi: [10.1190/1.1443406](https://doi.org/10.1190/1.1443406).
- Viana, F. A., 2016, A tutorial on Latin hypercube design of experiments: Quality and Reliability Engineering International, **32**, 1975–1985, doi: [10.1002/qre.1924](https://doi.org/10.1002/qre.1924).
- Vidale, J., 1988, Finite-difference calculation of travel times: *Bulletin of the Seismological Society of America*, **78**, 2062–2076, doi: [10.1785/BSSA0780062062](https://doi.org/10.1785/BSSA0780062062).
- Virieux, J., H. Calandra, and R.-É. Plessix, 2011, A review of the spectral, pseudospectral, finite-difference and finite-element modelling techniques for geophysical imaging: *Geophysical Prospecting*, **59**, 794–813, doi: [10.1111/j.1365-2478.2011.00967.x](https://doi.org/10.1111/j.1365-2478.2011.00967.x).
- Wait, J. R., A. Cullen, V. Fock, and H. Hagger, 1964, Electromagnetic waves in stratified media: American Institute of Physics.
- Wang, X., Y. Wang, Z. Cao, W. Zou, L. Wang, G. Yu, B. Yu, and J. Zhang, 2013, Comparison study on linear interpolation and cubic B-spline interpolation proper orthogonal decomposition methods: *Advances in Mechanical Engineering*, **5**, 561875, doi: [10.1155/2013/561875](https://doi.org/10.1155/2013/561875).
- Wu, S., Q. Huang, and L. Zhao, 2022, Instantaneous inversion of airborne electromagnetic data based on deep learning: *Geophysical Research Letters*, **49**, e2021GL097165, doi: [10.1029/2021GL097165](https://doi.org/10.1029/2021GL097165).
- Zhao, X., X. Chen, Z. Gong, W. Yao, and Y. Zhang, 2024, A hybrid method based on proper orthogonal decomposition and deep neural networks for flow and heat field reconstruction: *Expert Systems with Applications*, **247**, 123137, doi: [10.1016/j.eswa.2024.123137](https://doi.org/10.1016/j.eswa.2024.123137).

Biographies and photographs of the authors are not available.