Pacific Northwest
NATIONAL LABORATORY

PNNL-33394

Deep Learning At Depth

Estimating subsurface parameters from geophysical monitoring data.

September 2022

Kirsten N. Chojnicki
James V. Koch
Timothy C. Johnson

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

**Available to DOE and DOE contractors from
the Office of Scientific and Technical
Information,
P.O. Box 62, Oak Ridge, TN 37831-0062**
www.osti.gov
**ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov**

**Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov**

PNNL-33394

Deep Learning At Depth

Estimating subsurface parameters from geophysical monitoring data.

September 2022

Kirsten N. Chojnicki
James V. Koch
Timothy C. Johnson

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Abstract

Geophysical imaging techniques are a non-invasive way to image the subsurface and understand both subsurface solid (rock/soil) and fluid property distributions and their evolution in time. Inversions of the geophysical data, such as Electrical Resistance Tomography (ERT) data, are solved to estimate the subsurface property distributions, such as conductivity, and many inversion techniques smooth out sharp gradients in rock or fluid property distributions. Sharp gradients in subsurface properties tend to be present in situations with complex subsurface structures, which are common in many subsurface applications. We have successfully demonstrated that it is possible to inform, or constrain, inversions with neural networks trained on synthetic data with complex subsurface structures. Initial results suggest this process may be optimizable to yield property distributions that better represent the true property distributions than the same inversion process without the neural network constraint. Future work would optimize the neural network performance for this application and then apply the synthetic-data trained neural network to real data to understand the utility and performance of this technique for real data sets.

Summary

For users and managers of subsurface resources, critical decision-making hinges on understanding the current state of the subsurface system and how it may evolve in time. For complex systems building that understanding of the system state and how it may evolve is challenging. Geophysical imaging techniques have been developed as lower cost and non-invasive ways to image the subsurface to characterize both the solid (rock/soil) and fluid property distributions and their evolution in time. However, these techniques rely on inversions of the geophysical data which tend to smooth out the sharp changes in property distributions that define these complex systems, which are common in many subsurface applications. In this project we explored whether incorporating deep learning techniques in the geophysical inversion process could be one way to retain sharp subsurface property. We focused on the geophysical technique of Electrical Resistance Tomography (ERT) which is used to estimate 3-D bulk electrical conductivity (BEC) distributions in the subsurface from electrical potential difference measurements. BEC distribution and evolution are governed by parameters that control or describe subsurface dynamics making it useful to understand the transport of material underground. ERT also has a large dynamic measurement range which enables simultaneous measurements of BEC distributions spanning multiple orders of magnitude within a region of interest, making it an important tool for complex systems. For this feasibility study we generated synthetic BEC distributions using Sequential Gaussian Simulation. A suite of 100 synthetic BEC data sets was generated by using the open source geostatistical software package SGeMs (http://sgems.sourceforge.net/). The BEC range in these data spans 5 orders of magnitude from 0.0001 to 1 S/m. These BEC distributions are static to understand the role of deep learning in informing structure alone. We constructed and trained Deep Neural Networks to map between the BEC distributions and/or residuals. We demonstrate that it is possible to inform, or constrain, inversions with neural networks trained on synthetic data with complex subsurface structures. Initial results suggest this process may be optimizable to yield property distributions that better represent the true property distributions than the same inversion process without the neural network constraint. Future work would optimize the neural network performance for this application and then apply the synthetic-data trained neural network to real data to understand the utility and performance of this technique for real data sets.

Acknowledgments

This research was supported by the Energy and Environmental Directorate (EED) Mission Seed, under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Contents

Figures

Tables

Introduction

For users and managers of subsurface resources, critical decision-making hinges on understanding the current state of the subsurface system and how it may evolve in time. As the systems or their processes increase in complexity, such as rapidly changing temporal evolutions (highly dynamic systems) or large changes in material properties over small spatial scales (highly heterogeneous systems), building that understanding of the system state and how it may evolve is challenging. Furthermore, while measurements from techniques such as well logging would support some single-point characterization of these systems, these measurements are often limited because of their expense.

In contrast, geophysical imaging techniques have been developed as lower cost, multipoint and non-invasive ways to image the subsurface to characterize both the solid (rock/soil) and fluid property distributions and their evolution in time. However, these techniques rely on inversions of the geophysical data which tend to smooth out the sharp changes in property distributions that define these complex systems, which are common in many subsurface applications. Thus, improvements in the inversion process are needed to support the characterization, and thereby understanding and prediction, of complex systems in the subsurface and the improvement must allow operation on timescales that support decision making.

Deep learning is a tool that is well suited for this challenge (e.g., Yu and Ma 2021; Yeung et al. 2022). It allows for the combination of disparate data sets which typically characterize subsurface sites and can simultaneously incorporate both the sparse well logging data as well as the more spatially dense and frequent geophysical data by fusing information from each source and then learning the complex non-linear relationships between them. It can also retain information, such as geologic structure, which may be important in complex systems. Furthermore, evaluation of deep learning models is computationally inexpensive, enabling their deployment on decision-making timescales.

Thus, in this project we explored whether incorporating deep learning techniques in the geophysical inversion process could be one way to retain sharp subsurface property gradients while operating at time-scales that could support decision making. We focused on the geophysical technique of Electrical Resistance Tomography (ERT) which is used to estimate 3-D BEC distributions in the subsurface from electrical potential difference measurements (e.g., Johnson et al. 2017). BEC distribution and evolution are governed by parameters that control or describe subsurface dynamics making it useful to understand processes that involve the transport of material underground. ERT also has a large dynamic measurement range which enables simultaneous measurements of BEC distributions spanning multiple orders of magnitude within a region of interest, making it an important tool for complex systems.

Deep learning models traditionally require a large amount of data to perform well (thousands of examples per parameter of interest), and subsurface event data is comparatively rare. Thus, for this feasibility study, we generated synthetic data characterized by sharp changes in property distributions. Traditional inversions of this type of data are known to smooth these gradients and thus it might be straight forward to observe and assess whether the gradients were instead preserved by including deep learning in the inversion process. We also started with a static case to understand the role of deep learning in informing structure before we considered its evolution over time.

Deep Learning Approaches and Results

Background

Inverse problems in Partial Differential Equations (PDEs) involve the reconstruction of some part of a PDE – a Quantity of Interest (QoI) such as a scalar field, closure term, coefficient, etc. – from the system's solutions or observations thereof (Isakov 2006). Recovery of this QoI is deemed an 'inversion' since one begins from a solution to a PDE and then performs the inverse of the forward problem to derive some aspect of the governing PDE. PDE inversions are a primary method by which practitioners can interrogate physical processes without having complete knowledge of said physics or physical properties involved. Inversions are an important tool for disciplines like geophysics which tend to make observations of the earth system for the purpose of deriving dynamics that are otherwise difficult to observe or quantify since they occur underground.

Many inverse problems in the physical sciences are ill-posed, meaning that the solution to an inverse problem is highly sensitive to the final state. Typically, because PDEs are discretized in space (with an associated mesh or grid) and time, there is some degree of truncation and numerical round-off errors introduced to the modeling procedure. For forward modeling, many PDEs of interest are well-posed and the associated numerical algorithms are stable. The ill-posedness of the inverse problem, together with truncation and round-off errors, means that the inverse problem is ill-conditioned and will likely suffer from numerical instability. Thus, many state-of-the-art inversion procedures rely upon *regularizers* to enforce regularity upon the inverse problem – e.g. promoting smooth solutions or through a Tikhonov-type regularization (Golub et al. 1999).

These additional constraints make the inverse problem tractable at the expense of solution resolution. A smoothing constraint, for example, will discourage solutions with sharp discontinuities. In geophysical problems, where identifying such discontinuities is of critical importance, regularization typically inhibits the ability of an inverse problem to resolve these boundaries.

Sought are methods for 'informed' regularization techniques that retain the ability to resolve important geologic features (e.g., discontinuities). Recent advances in the fields of Artificial Intelligence (AI) and data science have provided the opportunity to construct new methods for performing and informing inversions that otherwise are intractable. In seeking to leverage existing validated and verified computational tools (i.e. E4D; Johnson et al. 2017), we consider here only approaches that are deterministic (vs. probabilistic) and unobtrusive (offline vs. online learning) to promote quick prototyping for coupled E4D/ML approaches. Specifically, for this study, we examine utilizing *Principal Component Analysis* and *Deep Neural Networks* in novel ML architectures coupled with the E4D software. First, we look at the applicability of a DNN to inform the depths and conductivities of geologic layers given surface measurements obtained in a pole-pole resistivity survey. Second, we look at dimensionality reduction techniques – such as PCA – as a method to encode spatial structure. We couple PCA with a DNN to guide an inversion to adhere to a reduced-rank approximation of the geology. Note that many other methods fall outside of the scope of this study but still may be valuable to pursue in future work, such as *Physics-Informed Neural Networks* and, more generically, the class of models called *Generative Models*.

Approach #1: Predicting Layer Properties with DNNs

Artificial Deep Neural Networks (DNNs)

Artificial Deep Neural Networks (DNNs) are trainable universal function approximators (Hornik et al. 1989). DNNs are comprised of composed layers of nodes (or artificial neurons), each of which manipulates its input via a parameterized activation function (introduces nonlinearity) and a bias term. These activations are chosen to be easily differentiable such that each parameter of the network has a corresponding gradient function. Thus, in training a neural network to approximate data, residuals can be backpropagated to each of the parameters, which can then be updated according to an optimization algorithm (e.g. gradient descent).

DNNs are attractive for their flexibility and expressivity: according to the universal approximation theory of neural networks, architectures of sufficient width and depth can capture functional forms to arbitrary precision, enabling their use in a wide variety of applications with minimal user supervision. For these reasons, this approach was the first tested for this study.

DNN-based Inversion of Synthetic Data

Recognizing that many geologic structures of interest are fundamentally layered, we define a fictitious set of layered geologies to prototype a machine learning method to predict layer depths and layer conductivities. In this framework, we seek to train a deep neural network to act as a map between observations (pole-pole ERT survey measurements) to feature space, defined by the number of layers and their corresponding conductivities.

Figure 1 shows a description of the problem and the associated training and deployment diagrams. Assumed is geology containing three layers suspended in a pseudo-infinite domain. Each n-th layer is associated with a layer transition depth, d_n, and a conductivity, σ_n, except for the last layer (assumed infinite depth). Stacking these geologic features for each of the layers results in a vector of length $2n$-1. We define this to be the output space for our DNN. The input is the vector of concatenated measurements from a real or synthetic pole-pole ERT survey. For this example, 120 measurements were recorded from an array of sensors equally spaced on the surface of a (synthetic) test site. The synthetic dataset contains 300 unique geologic instances, each with three layers.

The DNN to be trained contains three hidden layers of 10 nodes with Exponential Linear Unit (ELU) activation and a sigmoidal output layer:

$$NN(\boldsymbol{x};\theta): \mathbb{R}^{120} \rightarrow \mathbb{R}^{5}$$

where $\boldsymbol{x}$ is the vector of measurements and θ are the trainable parameters of the neural network. An implicit constraint is therefore imposed on the network in that the number of geologic features (5) and measurements (120) are fixed. Each variation in geologic features or measurements would require instantiation of a new network.

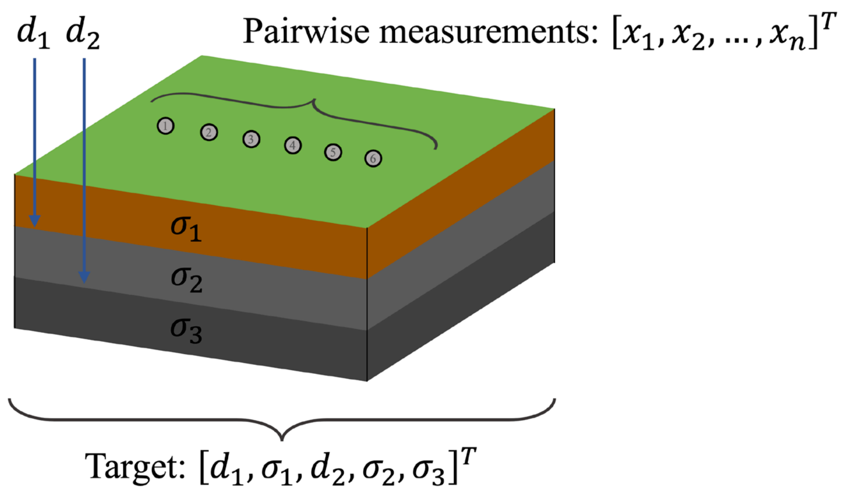

Figure 1: Notional schematic for mapping measurement space to feature space for an assumed-layered geologic structure. Each layer is associated with a depth and a conductivity. A DNN mapping between measurement space and the target feature vector is sought. For this study, the number of measurements and layers was fixed.

Each of the 300 unique geologic instances were subjected to a synthetic pole-pole ERT survey using the E4D software package, generating the 120 measurements per geologic instance. This constitutes the entirety of the dataset for this task. The dataset was split 70/30 into training and testing subsets. Because the proposed network is of a simple feed-forward architecture, training the network can proceed with standard off-the-shelf optimizers (such as Adam). The complete set of training hyperparameters is listed in Table 1. The loss history through the training iterations is shown in Figure 2.

Table 1. Hyperparameters for DNN in layered geology example.

Parameter	Value
N Train	210
N Test	90
Network architecture	3 hidden layers w/ 10 nodes, ELU activation, sigmoid output layer
Loss function	MSE(target,NN(measurements))
Optimizer and LR	Adam, 0.0001
Batch size	10
Epochs	1000
Data normalizing/preprocessing	[0,1] rescaling per feature

Performance of the trained network is evaluated by gathering the NN-predicted features for the measurements of the geologic instances in the test set. The average approximation error for each the five features aggregated over the test set is shown in Table 2.

Table 2. Aggregated performance metrics for layered geology example.

Feature	d_1	σ_1	d_2	σ_2	σ_3
Mean approximation error	93.5%	459%	49.1%	1010%	923%

A comparison of the solutions from the inversions with and without the neural network are shown in Figure 3. Note that in this learning task, obtaining reasonable conductivity values is a far more difficult task than recovering layer depths.

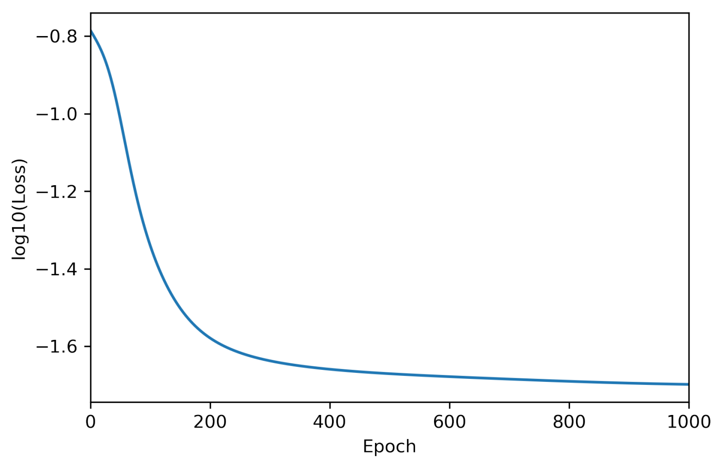

Figure 2: Loss history for training the DNN.

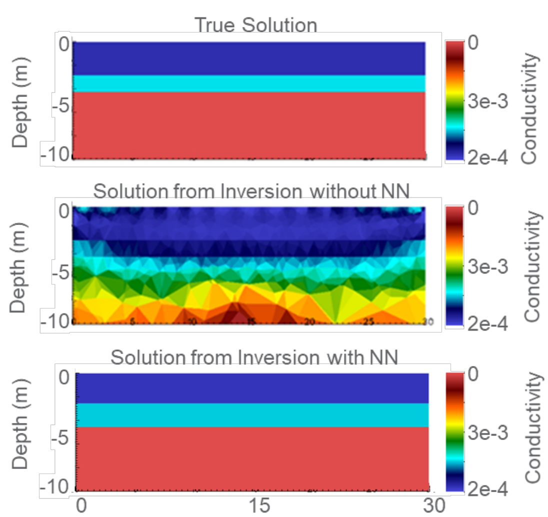

Figure 3: Example DNN-informed inversion. The DNN ingests measurement values and is trained to output layer depths and conductivities. Using those depths, standard inversion tools (such as E4D) can be used to resolve the conductivity values.

Approach #2: Predicting Layer Properties with PCA

Principal Components Analysis (PCA)

The DNN of Approach #1 maps measurements to geologic features explicitly, i.e. a practitioner chooses the measurement and feature vectors to construct and train the network. While this can be demonstrated for relatively simple geologies (such as prototypically layered geology), this becomes intractable with more realistic geology. Here, we shift to using *Principal Components Analysis* (PCA) as a tool for encoding structure for downstream deep learning tasks. Specifically, we focus on *dimensionality reduction* via PCA – that is, projecting high-dimensional data onto a lower-dimensional manifold defined by the first several or most significant orthogonal *principal components*.

PCA is a linear transformation of data into a coordinate system where the variance of the data is expressed by projections of data onto the coordinate axes. Principal components – orthonormal basis vectors - are the unit-normalized eigenvectors associated with the eigenvalues of the covariance matrix of the mean-centered data. *Rank reduction* or *dimensionality reduction* is the procedure of retaining a subset of principal components for the reconstruction of data. Typically, those that are retained describe the majority of variance observed in the dataset or for a particular data snapshot. For a description of the linear algebra formalisms for PCA and rank truncation, we refer to Brunton et al. 2022. Note that for this task, we use an SVD-based formulation as opposed to explicitly calculating eigenvalues and eigenvectors of the covariance matrix.

We seek to encode the dominant geologic features in a synthetically-generated dataset via a rank reduction procedure performed through PCA. For this feasibility study we generated synthetic BEC distributions using Sequential Gaussian Simulation. A suite of 100 synthetic BEC data sets was generated using the open source geostatistical software package SGeMs (http://sgems.sourceforge.net/). The BEC range in these data spans 5 orders of magnitude from 0.0001 to 1 S/m. An example distribution is shown in Figure 4. The distributions also contain a range of spatial complexities over which to test the performance of the neural network informed inversion process. Also, these BEC distributions are static to understand the role of deep learning in informing structure before considering its evolution.

Figure 5 displays the first singular values associated with the principal components for the training data. For rank reduction, a certain energy threshold is usually defined (e.g., “capturing” 99% of the energy of the dataset, where energy, or a surrogate, is defined as the cumulative sum of the singular values ratioed to their total sum, as seen in Figure 6) for a minimum rank to describe the major structural features of the data. Rules of thumb exist (e.g., capturing 99% of energy), but ultimately these are heuristics that are practitioner-tuned for specific purposes. In the present exploratory study, we select a low rank, 15, to (i) reduce the number of trainable parameters in our downstream tasks, and (ii) reduce the influence of excessively small spatial features.

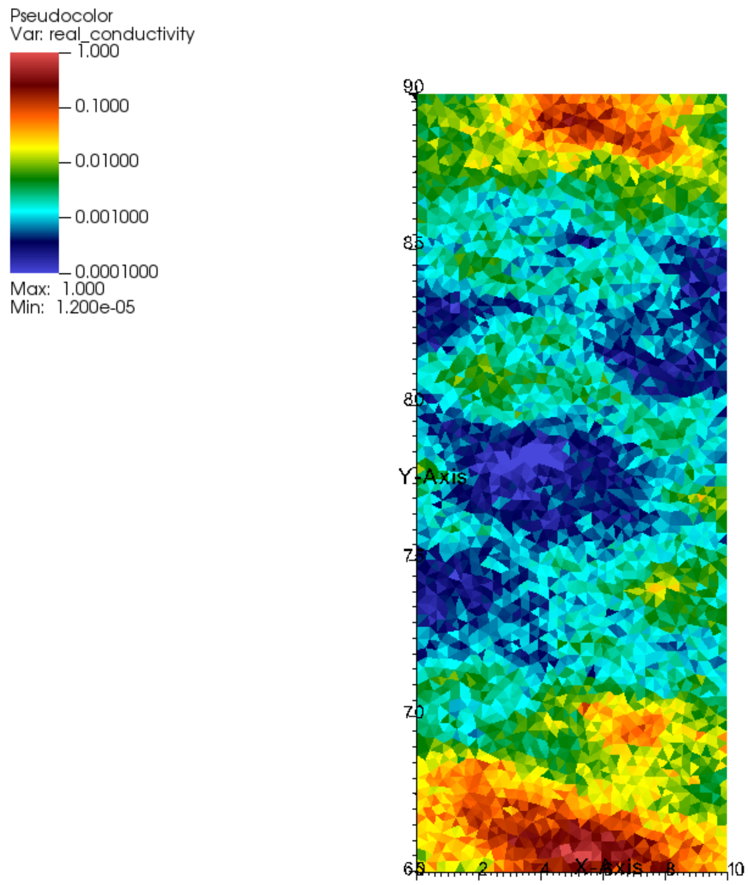

Figure 4. Example of a BEC distribution generated using Sequential Gaussian Simulation.

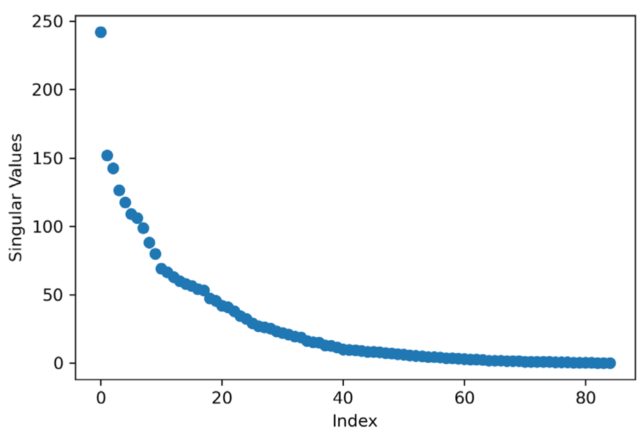

Figure 5: Singular value decay for the synthetic BEC dataset.

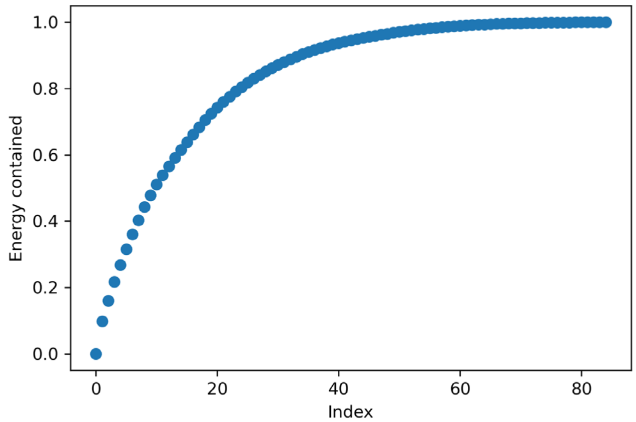

Figure 6: Cumulative energy distribution through singular value index.

PCA-based Inversion of Synthetic Data

The DNN of Approach #1 was constructed to map measurement space for feature space. In this section, an analogous DNN is constructed and trained to map from a measurement space to feature space, where feature space is defined by the reduced-rank PCA coefficients. This allows for the PCA to handle the spatial structure encoding without input from a practitioner, except for selecting a rank cutoff.

The deep learning pipeline is depicted in Figure 7. E4D is used as a forward simulator during training; i.e. to solve the appropriate physics and produce an output consistent with an observation and measurement strategy – in this case, this is a pole-pole ERT survey done synthetically with the results of the physics simulation. To begin, two forward simulations are performed with E4D: (i) one on a geologic instance reconstructed from random PCA coefficients, and (ii) one on a geologic instance from the data set. Each forward run results in a vector of measurements, which can be compared. The residual error between these two forward runs is the input to the DNN. The DNN maps this residual vector to an *update* of the PCA coefficients (i.e. the DNN output is added to the random PCA coefficients). The reconstruction residual between the updated PCA reconstruction and the ground-truth geology is computed, and this residual is backpropagated through the DNN such that the DNN’s weights can be updated accordingly.

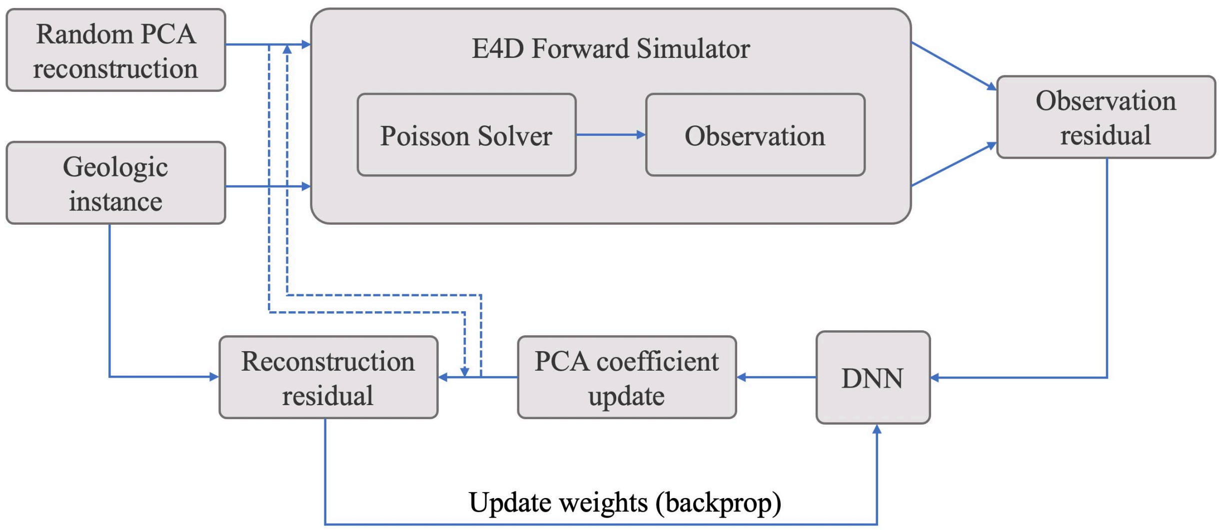

Figure 7: Training pipeline for PCA-based inversion. A standalone physics simulator (E4D) is used to generate synthetic measurement data for both ground-truth geologic instances and random PCA reconstructions. The measurement residual is fed into a DNN, which generates a PCA coefficient update. The reconstruction residual is used to train the weights and biases of the DNN.

Because the E4D simulator is not differentiable, the DNN is trained to map measurement residuals (between that of a random PCA reconstruction and a ground-truth geologic instance) to an update to the random PCA coefficient. Residuals cannot be backpropagated through E4D. Table 3 lists all hyperparameters associated with this approach and Figure 8 shows the loss history for training the DNN with these hyperparameters.

Table 3. Hyperparameters for DNN in PCA-based inversion example.

Parameter	Value
N Train	85
N Test	9
N Measurements	496
Rank	15
Network Architecture	Two hidden layers of size 200 and 75, respectively. ELU() activation and linear output.
Loss function	MSE(PCA reconstruction, G.T. geology)
Data preprocessing/norm	Log10() and [0,1] rescaling
Optimizer and LR	Adam, 0.01
Batch size	Full batch
Epochs	200

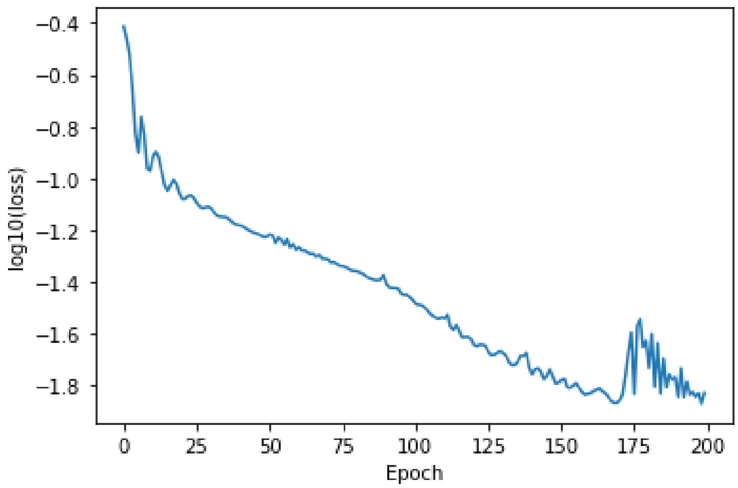

Figure 8: Loss history during training for the PCA-based inversion.

In deployment (Figure 9), the DNN ingests a measurement residual (between E4D in inversion mode a ground-truth measurement) and outputs an updated PCA reconstruction. If running coupled with E4D, the updated PCA reconstruction is fed back into E4D repetitively until convergence is reached. Alternatively, after a single pass through E4D, the PCA reconstruction can be *digitized*, i.e. the values binned into a discrete number of intervals and used as a prior for the spatial structure of the geology during an inversion call via E4D.

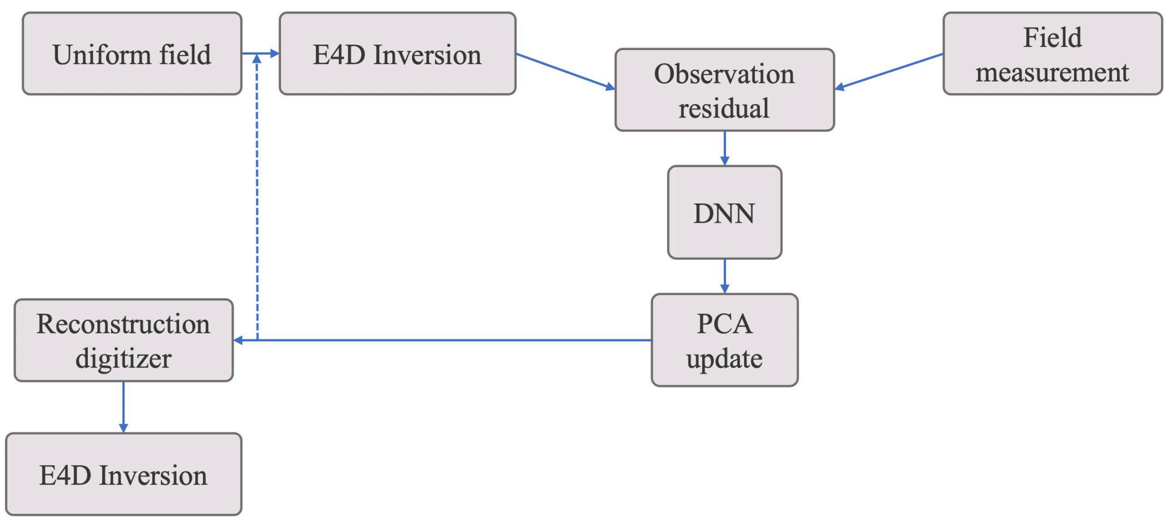

Figure 9: Schematic for deployment of the trained DNN for PCA-based inversions. The DNN provides an update to the PCA coefficients, which can be used iteratively with an inversion code (E4D) or used to produce a spatial encoding to be used as a constraint in an offline inversion.

Example NN-informed inversions using this latter technique are shown in Figure 10 with 4 (user-specified) digitized conductivity bins. The left column represents true BEC fields, whereas the second column represents the encoding of spatial structure by the trained NN. The resulting NN-informed inversion is shown in the third column and can be compared with the fourth column – the simplest 'Occam's Inversion' performed on the same data.

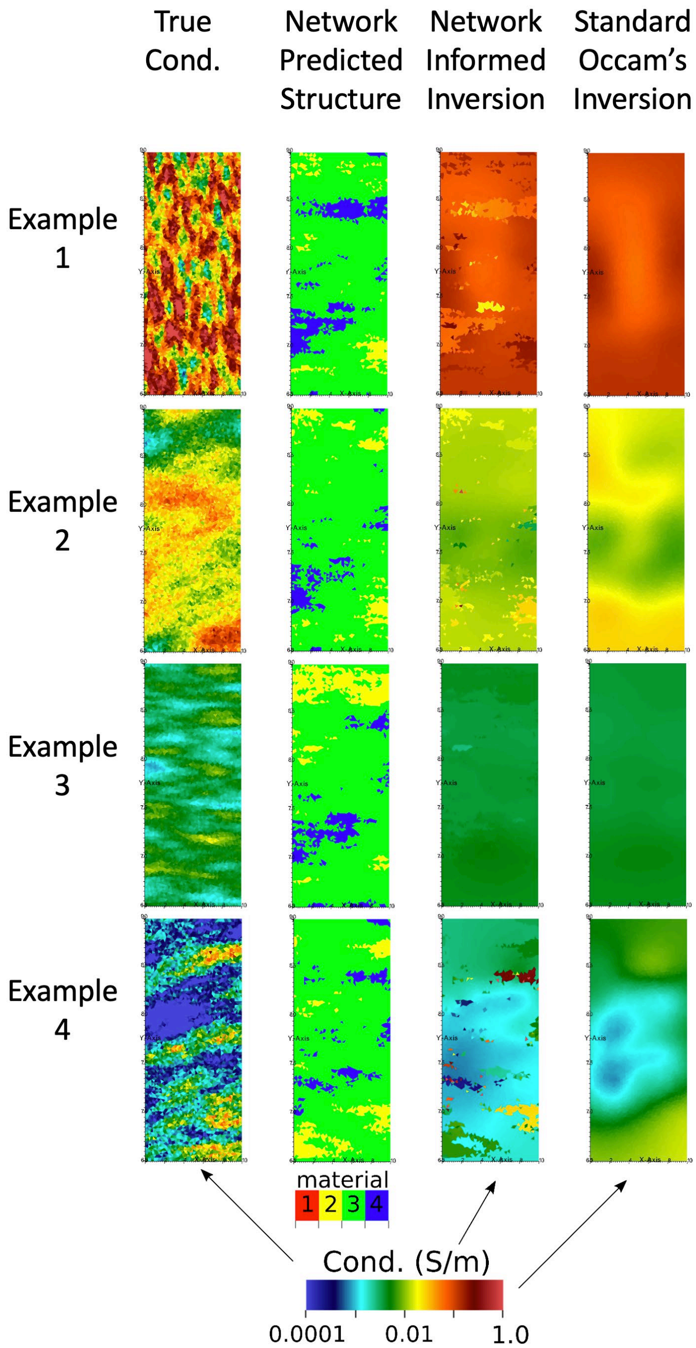

Figure 10: Example DNN-informed inversions with PCA-based spatial encoding. The first column corresponds to the true geologic instance (i.e., the synthetic data). The second column is the spatial encoding via the trained DNN. The third column is the DNN-informed inversion and can be compared with the results in the fourth column, the standard Occam's inversion.

Discussion, Conclusions, and Future Work

Presented were two proof-of-concept ML approaches for encoding spatial structure in geophysical inversions. The first, detailed in Approach #1, focused exclusively on characterizing layers by their depth and conductivity. These geologic features were predicted via a trained DNN whose input was the vector of measurements from a pole-pole ERT survey. In our second ML approach, detailed in Approach #2, we generalized our first approach to adapt to arbitrary geology through spatial encoding via data compression through Principal Components Analysis of a geologic dataset.

Through these proof-of-concept exercises, we have successfully demonstrated that it is possible to inform, or constrain, inversions with neural networks trained on synthetic data with complex subsurface structures. There are two main takeaways from these preliminary results: (i) this particular strategy of spatial encoding is successful in that the inversion results closely adhere to the structural prior, and (ii) this particular strategy for determining spatial structure is insufficient to recover the true geology.

Our initial results suggest this process may be optimizable to yield property distributions that better represent the true synthetic property distributions than the same inversion process without the neural network constraint. While we were successful in encoding spatial structure, success was varied in the accuracy of the said structure. First, we stress that the self-contained studies of Approaches #1 and #2 were each performed for a single set of hyperparameters - the presented results should not be taken out of context from these settings. To better determine the performance of these methods, a comprehensive hyperparameter search should be completed (e.g. network architecture, learning rate, PCA reduced rank, etc.). Future work should consider these strategies for optimizing this process – e.g., a formal hyperparameter tuning procedure - to improve upon the results of this study.

For future work, needed are different approaches that better leverage both existing datasets, such as the synthetic data generated for this study, and for datasets with variable measurement and feature spaces. Such approaches are likely to include *probabalistic* methods as opposed to *deterministic* methods. These approaches can allow for both modeling complex data distributions and efficient distribution sampling. When coupled with inversion routines like E4D, these approaches can both minimize data misfit and maximize likelihood of a particular geology given the data.

References

Brunton, Steven L., and J. Nathan Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

Golub, Gene H., Per Christian Hansen, and Dianne P. O'Leary. "Tikhonov regularization and total least squares." *SIAM journal on matrix analysis and applications* 21, no. 1 (1999): 185-194.

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators." *Neural networks* 2, no. 5 (1989): 359-366.

Isakov, Victor. 2006. “Inverse Problems for Partial Differential Equations,” *Applied Mathematical Sciences*, Vol. 127, https://doi.org/10.1007/0-387-32183-7

Johnson, Timothy C., Glenn E. Hammond, and Xingyuan Chen. 2017. “PFLOTRAN-E4D: A parallel open source PFLOTRAN module for simulating time-lapse electrical resistivity data” *Computers & Geosciences*, Vol. 99:72–80, https://doi.org/10.1016/j.cageo.2016.09.006.

Yeung, Yu-Hong, David A. Barajas-Solano, and Alexandre M. Tartakovsky. 2022. “Physics-Informed Machine Learning Method for Large-Scale Data Assimilation Problems” *Water Resources Research*, Vol. 58 (5), https://doi.org/10.1029/2021WR031023.

Yu, Siwei and Jianwei Ma. 2021. “Deep Learning for Geophysics: Current and Future Trends” *Reviews of Geophysics*, Vol. 59 (3), https://doi.org/10.1029/2021RG000742.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov