Data-driven two-dimensional near-surface seismic imaging

Joseph P. Vantassel[1#], and *Sanish* Bhochhibhoya[1]

[1]*Virginia Polytechnic Institute and State University, Department of Civil and Environmental Engineering, Blacksburg, Virginia, USA*
[#]*Corresponding author: jpvantassel@vt.edu*

ABSTRACT

Non-invasive site characterization techniques have the potential to rapidly evaluate large subsurface volumes to guide subsequent invasive geotechnical site investigation. Among these methods, seismic full waveform inversion (FWI) stands out for its potential to recover detailed two-dimensional (2D) images of the subsurface. However, FWI's need for substantial computational resources and sensitivity to the initial starting model has limited its utilization as a general-purpose geotechnical site characterization tool. Addressing this, prior studies have shown data-driven methods can predict 2D subsurface structures composed of soil over rock. In the present study, we aim to generalize these findings to all near-surface conditions. We propose a novel model generation framework that utilizes techniques from geostatistics to generate complex 2D subsurface models. The generated models include dipping soil and rock layers, soil lenses, boulders, and underground utilities; none of which have been considered previously. We use our model generation framework to simulate 100,000 2D subsurface models. We simulate field data acquisition along these 100,000 synthetic models, by numerically solving the elastic wave equation using an impulse source at the model's center surrounded by 24 receivers (12 on either side). The data-driven predictive model, trained on 90% of the simulated data, achieved a mean absolute percent error on the testing set of 19%. Furthermore, these predictions are made within fractions of a second circumventing the computational and starting-model-related challenges associated with traditional 2D FWI. These results demonstrate that data-driven methods can predict complex images of the subsurface to enable rapid subsurface imaging for geotechnical applications.

Keywords: full waveform inversion; geostatistics; subsurface imaging, deep learning

1. Introduction

Traditional geotechnical site characterization relies on interpolating between limited one-dimensional (1D) measurements of subsurface stratigraphy (e.g., via borings). The sparsity of traditional geotechnical measurements presents challenges in geological settings with rapid spatial variation (e.g., alluvial deposits) and/or subsurface anomalies (e.g., dissolution cavities in karst). A promising alternative is to employ non-invasive imaging methods to analyze the subsurface in 2D or 3D, to inform subsequent 1D geotechnical site investigations. Among various non-invasive imaging techniques, seismic full waveform inversion (FWI) is a compelling alternative, as its utilization of both phase and amplitude information from the seismic wavefield enables it to yield high-resolution 2D and 3D subsurface models.

FWI involves two steps: acquisition and inversion. Acquisition involves the controlled application of repeated impacts from a seismic source (e.g., sledgehammer) in the vicinity of strategically positioned sensors (e.g., geophones) to record the ensuing seismic wave propagation. Inversion has traditionally involved an iterative process of refining an assumed subsurface model or set of subsurface models such that the resulting simulated waveforms match those observed during acquisition (Tarantola 1984). FWI's ability to utilize both phase and amplitude information from the observed seismic waveforms, enables it to have higher resolution

and better applicability to heterogenous models than other imaging techniques (Pan, Gao, and Bohlen 2019).

Despite its potential for geotechnical site characterization (Kallivokas et al. 2013; Fathi et al. 2016; Mirzanejad and Tran 2019), traditional FWI faces barriers to its widespread adoption, in particular, the substantial computational resources required and its sensitivity to the initial starting model. Traditional FWI involves a multi-step non-linear data-fitting procedure. First, an initial subsurface model (i.e., a starting model) is assumed. Second, synthetic seismic waves are propagated through the assumed model by numerically solving the appropriate wave equation using computational methods such as finite difference (Levander 1988; Virieux 1986) or spectral element (Tromp, Komatitsch, and Liu 2008; Peter et al. 2011; Vai et al. 1999). Third, the assumed model is refined using the selected numerical optimization algorithm and insight from the mismatch between the simulated waveforms and those acquired in the field. Fourth, the process is repeated using the modified model as the new starting model until the simulated and measured seismic waveforms are in acceptable agreement. The FWI process is guided by the selected optimization algorithm and, as such algorithms are problem-specific, they can have a notable impact on the resulting solution (Nocedal and Wright 2006). Hence, the selected optimization algorithm plays a pivotal role in determining the quality of the ultimate solution (i.e., the subsurface image).

We divide numerical optimization methods into three categories: global, local, and data-driven. First, algorithms using global optimization explore a broad range of potential solutions by iteratively sampling the solution space. Global optimization algorithms are generally considered the most robust but are also the most computationally intensive. As a result, global optimization algorithms are relatively uncommon in the FWI literature, only being used in a few publications (e.g., Datta and Sen 2016; Mojica and Kukreja 2019). Second, local optimization, while less computationally demanding than global optimization, relies on accurate initial starting models that contain the site's general structure. If the starting model does not accurately reflect the site's general subsurface structure, local optimization methods can produce inaccurate subsurface images. In addition, in some cases, local optimization algorithms can become trapped in a local minimum or on a saddle point that prevents them from converging to the optimal solution (Abbas et al. 2023; Feng, Lin, and Wohlberg 2022; Monteiller et al. 2015). Nevertheless, due to their computational efficiency local optimization methods have been the algorithm of choice for the vast majority of FWI literature. Third, data-driven optimization is a recently developed category of algorithm where large quantities of data are used with artificial intelligence (AI) to directly model the inversion process. In the context of FWI, data-driven optimization involves simulating large quantities of subsurface models and the associated experimental data to train an AI predictive model to recover the subsurface model from the experimental data. Once trained, a data-driven model produces predictions almost immediately at negligible computation cost, even compared to local optimization. However, there are two major challenges: (1) the need for large datasets, which are not yet readily available, and (2) the need for complex AI model architectures, which require an understanding of deep learning.

In response, our study seeks to address these two challenges. First, we develop a novel model-generation framework to generate 100,000 realistic multi-layered 2D subsurface models. We then simulate the corresponding experimental seismic data using the open-source spectral element software SPECFEM2D. Second, we utilize the shear wave velocity (V_s) image from the simulated models and the corresponding waveform to develop a deep learning (DL) architecture for performing FWI. We note that both research objectives of this study builds substantially on previous work by Vantassel, Kumar, and Cox (2022). In regards to the first objective, this study develops an entirely new model generation framework that develops subsurface models that more completely define the range of complexity expected in geotechnical practice. In regard to the second objective, this study develops a new DL model architecture that better extracts seismic wavefield information. To demonstrate these improvements, the new model architecture is compared to the architecture proposed previously by Vantassel, Kumar, and Cox (2022) using the newly developed dataset.

2. Generating Realistic Subsurface Models

In geotechnical engineering, the subsurface is predominantly modeled as homogenous layers with horizontal boundaries (e.g., Jaksa et al. 2005; Naghibi, Fenton, and Griffiths 2014; Bhochhibhoya et al. 2023), or relatively simple two-layered profiles with somewhat intricate boundaries (e.g., Crisp, Jaksa, and Kuo 2017; Vantassel, Kumar, and Cox 2022; Abbas et al. 2023; Yust et al. 2023). However, actual soil profiles are multi-layered and have complex layer boundaries. Therefore, we propose a novel model generation framework for generating such realistic soil profiles to address limitations in existing methods. The framework, which utilizes techniques from geostatistics, is able to generate diverse models with multiple layers, arbitrary boundaries, soil lenses, and underground utilities.

The model generation framework operates in a multi-step process. First, the number of layers (n) is selected following a user-defined probability mass function (PMF). For this study, we allow up to eight layers, but bias the PMF such that the majority of models have fewer than four layers. Second, the framework defines the rough (i.e., unsmoothed) boundaries between each layer. To do this, the framework defines the starting point of each of the $n-1$ layer boundaries on the left-most edge of the 2D models, as shown by the two filled circles in Figure 1a. These starting points are then used to anchor the generation of the rough boundary. The generation of each of the $n-1$ rough boundaries involves repeatedly selecting from a uniform distribution that defines the vertical increment by which each subsequent point is shifted from the previous. We use a constant horizontal spacing between points (1.5 m in this study), see Figure 1a. This rough boundary is then smoothed to more accurately reflect standard geological processes (e.g., erosion). To smooth the rough boundary in this study, we use the Savitzky and Golay (1964) filter with a window comprising 20 data points and a polynomial order of 3.
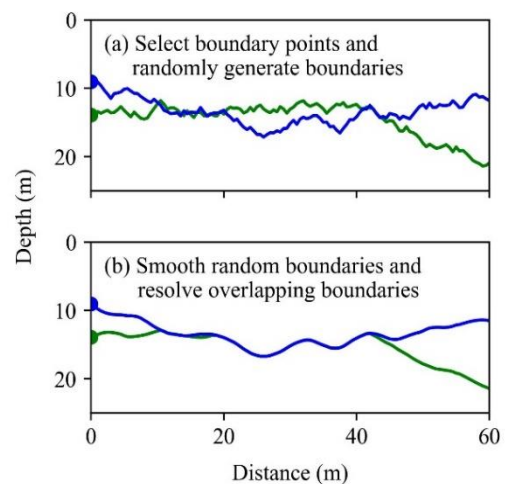
Figure 1. Simulation of layer boundaries for a three-layered model using the proposed model generation framework. (a) two points (solid-colored circles) are randomly selected on the left boundary to anchor the associated random rough layer boundary (solid-colored lines). (b) depicts the smoothed layer boundaries after applying the filter by Savitzky and Golay to the rough boundary in (a) and resolving boundary overlaps by preferentially selecting the upper-most material.

Finally, overlapping boundaries are resolved by preferentially selecting the upper material. This preferential selection of the upper-most material is consistent with geological processes associated with erosion and subsequent deposition (Crisp, Jaksa, and Kuo 2017). The layer boundaries after smoothing and layer boundary resolution are shown in Figure 1b. Notably, the model generation parameters aforementioned were selected through an iterative, albeit subjective, process to achieve the desired tradeoff between subsurface model complexity and geologic realism.

After definition of the model's layer boundaries, each layer's three elastic material properties, shear wave velocity (V_s), compression wave velocity (V_p), and mass density (ρ) were defined. Each subsurface model was represented as a collection of three, 2D images at a 0.25 m pixel, one image of V_s, one image of V_p, and one image of ρ. We will present the generation of each in turn.

Each V_S image was generated by simulating a correlated random field. The mean shear wave velocity ($V_{s,mean}$) of the correlated random field was selected between 50 and 1500 m/s using a user-defined probability density function (PDF). The values of V_S represent a range of geomaterials spanning from soft organic soils to hard rock. The coefficient of variance of V_S ($V_{S,COV}$) for each random field was randomly defined between 5% and 20% and conditioned on the selected value of $V_{s,mean}$, see Table 1. Lateral correlations were developed using an exponential correlation function with vertical and horizontal correlation lengths of 0.5 and 100 m, respectively.

The V_p and ρ images were derived from the V_s image. The V_p image was computed using:

$$V_p = V_s \times \sqrt{1 + \frac{2}{(1-2v)}}, \tag{1}$$

from elasticity theory where, v is Poisson's ratio and all other terms are defined previously. The range of v was defined between 0.1 to 0.495 and conditioned on the value of $V_{s,mean}$ of the layer, see Table 1. Finally, the ρ image was defined in the range of 2000 kg/m³ to 2100 kg/m³, and again conditioned on the selected value of $V_{s,mean}$, see Table 1. Note that for simplicity ρ of each layer, unlike V_s and V_p, was assumed constant for each layer. The use of constant ρ in each layer is similar to Vantassel, Kumar, and Cox (2022) and is appropriate because elastic wave propagation is only minimally sensitive to ρ. Note that the range of values defined in Table 1 represents the range from which a single value was selected following a uniform distribution.

Following V_S, V_P, and ρ image generation, we introduce additional complexity by including underground utilities. We assume, for simplicity, that all utilities are circular with uniform elastic material properties. As utilities are most likely to be near the ground surface, we bias their location to the upper 3 m, with an increased probability that they occur in the upper 1 m. For each model, the number of utilities was randomly selected between 0 and 8, following a user-defined PMF. For this study, we bias our PMF such that the majority of profiles have fewer than 2 utilities. The ranges of different parameters of the simulated utilities are summarized in Table 2. As before, a uniform distribution is used to randomly select a specific value from each parameter range.

Using the model generation framework, 100,000 soil models were generated for AI predictive model training and evaluation (discussed next). Each model was 60 m in length and 25 m in depth and discretized at a 0.25 m pixel. To focus on those portions of the model directly beneath the sensing array the models were truncated to 50 m by 20 m for AI prediction. Figure 2 presents the V_s image from eight models generated using the proposed framework, with the thin red lines defining the subset of the model utilized for training the AI predictive model (discussed next).

Note to show the complexity of the dataset more completely, the eight models presented in Figure 2a through 2h correspond to randomly generated profiles with 1 to 8 layers, respectively. As illustrated in Figure 2, the proposed model generation framework can generate a wide variety of subsurface models with highly variable numbers of layers and realistic boundaries.

3. Simulation of Seismic Data Acquisition

Synthetic seismic wavefield data was generated for each generated subsurface model using the open-source, spectral-element-based software SPECFEM2D (Tromp, Komatitsch, and Liu 2008). More specifically, we solved

Table 1. Range of values for the parameters of each layer used in this study. Note that all parameters are conditioned on the mean shear wave velocity ($V_{s,mean}$) of the layer. For each layer a specific value is randomly selected from within each specified range following a uniform distribution.

Parameter	Range
$V_{S,COV}$	0.05 – 0.15 if $V_{s,mean} > 1200$ m/s 0.05 – 0.20 if $V_{s,mean} > 460$ m/s 0.10 – 0.20 otherwise.
v	0.10 – 0.20 if $V_{s,mean} > 1200$ m/s 0.15 – 0.30 if $V_{s,mean} > 760$ m/s 0.25 – 0.35 if $V_{s,mean} > 460$ m/s 0.30 – 0.495 otherwise.
ρ	2100 kg/m³ if $V_{s,mean} > 760$ m/s 2050 kg/m³ if $V_{s,mean} > 360$ m/s 2000 kg/m³ otherwise.

Table 2. Range of values for the parameters describing the utilities placed into the generated models. For each utility, a specific value is randomly selected from within each specified range following a uniform distribution.

Parameter	Range
Radius	0.25 - 1.00 m
V_s	25 – 70 m/s
V_p	340 – 1500 m/s
ρ	800 - 2000 kg/m³
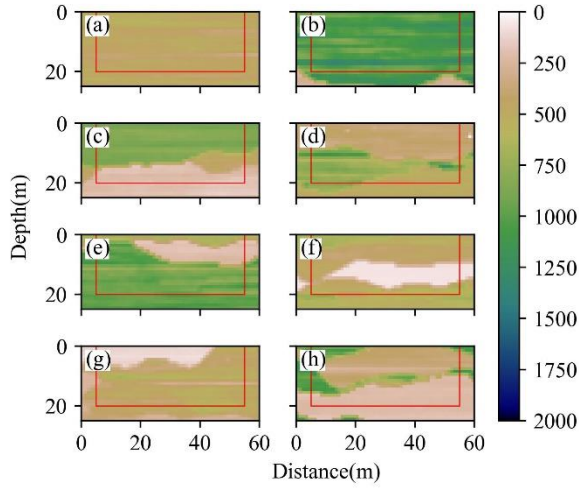
Figure 3. Randomly selected V_s images from the 100,000 generated for training the data-driven model. Each panel shows an image generated with a different number of layers, from (a) one to (h) eight. The thin red line shown in all panels indicates the portion of the V_s image directly beneath the surface array used and used to train the data-driven predictive model.

the elastic wave equation assuming 2D, plan-strain, and P-SV conditions. The input source was a 30 Hz Ricker wavelet that contains significant seismic energy between 10 and 100 Hz (i.e., a range representative of impact-type sources commonly used in near-surface seismic data acquisition). The source was positioned at the model's center (i.e., at 30 m) at the ground surface and oriented vertically. The top of the domain was modeled as a free surface (i.e., zero-stress boundary), whereas the sides and bottom were modeled as infinite using absorbing boundary conditions (ABCs). As mentioned previously, the domain was discretized into 0.25-m pixels to enable the modeling of appropriately short wavelengths in the domain. A suitably short time step (1E-5 seconds) was selected to ensure numerical stability. The output data, in the form of the vertical component of particle velocity, were recorded at 24 receivers' locations (12 on each side of the source) at a time step of 1 ms (i.e., at 1000 Hz) for a total duration of 1 second. Receivers were spaced at 2-m intervals and spanned between 7 and 53 m. Note that a time step shorter than the minimum required by the Nyquist theorem was used intentionally, to reduce signal variance and enable a more reliable capture of the "true" (i.e., noise-free) signal.
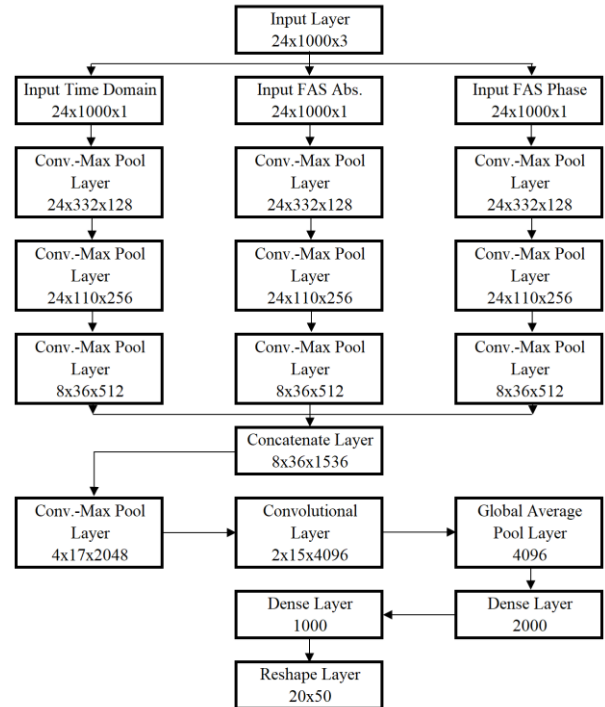
4. Developing the Model Architecture

DL model architectures are composed of a series of layers that each perform a relatively simple and differentiable numerical operation. These layers can be of many types, for example: convolution layers, dense layers, recurrent layers, dropout layers, pooling layers, and many others. For the sake of brevity, we will not detail the more than 40 model architectures (i.e., unique combinations of layers) considered throughout the course of this study. Instead, we present the selected architecture that was selected based primarily on its performance on the validation set (i.e., a subset of the training set not used directly for model training), but also considering

trainable parameter count, training time, and model elegance.

The selected DL model architecture, depicted in Figure 3, integrates both time-domain and frequency-domain representations of the experimental data. Here it is worth noting that early work by, Vantassel, Kumar, and Cox (2022) used solely a time-domain representation of the data, whereas later work by Abbas et al. (2023) used solely a frequency-velocity (i.e., via an F-K transform) representation of the data. By utilizing both time- and frequency-domain representations of the data we will show that we have developed a novel architecture with enhanced model performance and generalization capability.

The selected DL model takes a 3D tensor of size 24x1000x3 where the third (or channels) dimension corresponds to the raw seismic wavefield (time domain), amplitude of the Fourier amplitude spectrum (FAS) (frequency domain), and phase of the FAS (frequency domain) as input. Each channel of the input is of size 24x1000, corresponding to the 24 receivers and 1000 time-/frequency- domain samples. Each channel of the input data undergoes a sequence of three convolutional and max-pooling layers. Note that in Figure 3, a convolution layer followed by a max-pooling layer is denoted as a "Conv.-Max Pool Layer" for brevity. After the initial convolution, the three input channels are channel-wise concatenated. The concatenated data is then transformed by a convolution layer, followed by a max-pooling layer, and then another convolution layer. Followed by global-average pooling and two dense layers of size 2000 and 1000. Finally, the output layer of 1000 nodes is reshaped to 20x50 to correspond to the predicted seismic image of 20 m by 50 m with a 1-m pixel.

Figure 2. Schematic representation of the model architecture for the proposed model. Each box denotes one or multiple layers within the model, with the name and output shape of each layer explicitly listed inside the respective box.
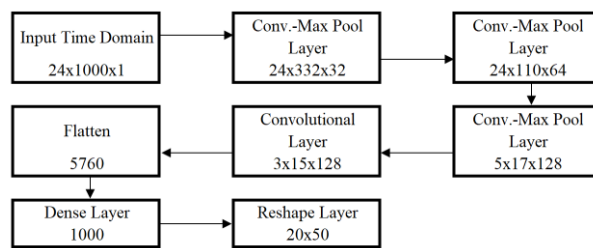
For comparison, the model architecture proposed by Vantassel, Kumar, and Cox (2022) follows a more typical structure, featuring a sequence of five convolution layers interspersed with four max-pooling layers and culminating in a single dense layer. This architecture demonstrated good performance on models composed of soil-over-rock, however, as this study uses complex subsurface models, it would be unfair to directly compare it to the model developed by Vantassel, Kumar, and Cox (2022). Therefore, to facilitate a more- fair performance evaluation, a slightly modified version of the Vantassel, Kumar, and Cox (2022) model was trained using the aforementioned (more complex) dataset. The modified model is summarized in Figure 4. It is noteworthy that all analyses, such as hyperparameter tuning and data augmentation, conducted in the proposed model, were also applied to the updated Vantassel, Kumar, and Cox (2022) model. Such that the updated model developed here is undoubtedly an improvement over the original given the more complex dataset and more advanced preprocessing (discussed next).

5. Model Training

Model training and evaluation were facilitated using the open-source deep learning library Keras, developed by Chollet and others (2018). The candidate model architectures were developed using the 100,000 seismic wavefield-image pairs aforementioned. The 100,000 pairs were split for training (80%), validation (10%), and testing (10%). Data were effectively expanded, using data augmentation, and pre-processed, using band-pass filtering and normalization. Regarding the former, data augmentation is a crucial step that enhances the generalization capacity of the model. Three data augmentation techniques were used in this study: flipping the image-wavefield data left to right (i.e., simulating a change in perspective), zero-padding of the wavefields (i.e., simulating a delay of the seismic records in time), and injecting channel-wise random noise (i.e., simulating channel-specific levels of noise during data acquisition). In addition, each seismic wavefield was bandpass filtered between 10 and 100 Hz and then normalized using the maximum absolute value observed across all receivers. Performing normalization maintains the relative amplitude between receivers, while also enabling the model architecture to learn more efficiently. While not able to be explained in detail here all of the data augmentation and pre-processing techniques aforementioned were shown to improve the model's predictive performance and generalization ability.

Model training also involved tuning each model's hyperparameters such as initial learning rate, learning rate decay function, batch size, number of epochs, and model optimizer. For the sake of brevity, we will not discuss the dozens of combinations of hyperparameters considered during this study, instead, we summarize that the final selected hyperparameters for the proposed model were a learning rate of 0.002 with exponential decay by a factor of 0.1 after 10 epochs, a batch size of 16, 60 training epochs, and the use of the Adam optimizer by Kingma and Ba (2015). For the proposed DL model, mean absolute error (MAE) was used as the loss function

Figure 4. Schematic representation of the model architecture, proposed by Vantassel, Kumar, and Cox (2022) that has been adjusted for the input and output shapes of the dataset used in this study.

due to its decreased sensitivity to outliers and its ability to converge to the median of a dataset. It should be noted that all procedures used in developing the proposed model, including hyperparameter tuning, were also applied to the modified Vantassel, Kumar, and Cox, (2022) model shown in Figure 4.
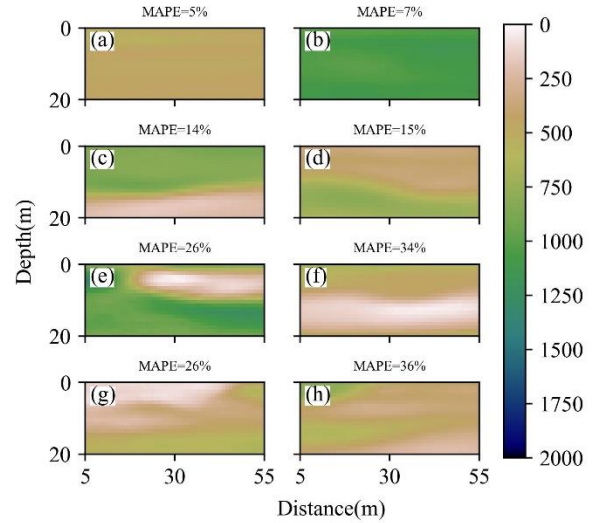
6. Testing the Model

A comprehensive evaluation of the optimized and tuned predictive AI model was performed using the 10,000 seismic wavefield-image pairs reserved for testing. As mentioned above, Figure 2 illustrates 8 V_s images chosen for comparing predictions by the proposed AI model; these images were selected from the testing set and therefore not used during model training. Figure 5 illustrates the prediction of V_s images using the proposed AI model and the associated seismic waveforms. To quantitatively demonstrate the efficacy of the proposed AI model the mean absolute percent error (MAPE) (i.e., the average of the pixel-by-pixel MAPE of the true and predicted V_s image) is shown. The MAPE for the eight example models ranges from 6% to 30%, with MAPE increasing with model complexity. The average MAPE across the entire 10,000 member testing dataset was 19% (i.e., predictions within 19% of the true value). Figure 6 shows the distribution of pixel-by-pixel residuals for the same models. We generally observe that residuals (i.e., the difference between the true and predicted image) increase with depth and distance from the model's center. This is expected from utilizing only a source at the center of the model's center. Future work will investigate the added benefit of using additional seismic sources located elsewhere in the model.

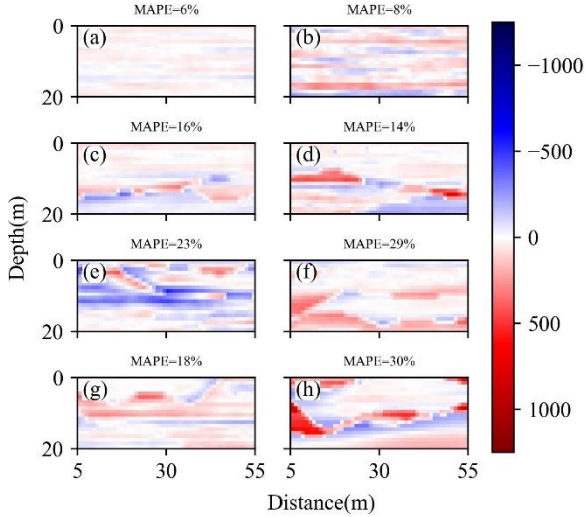## 7. Comparison of the Proposed Model Architecture with that by Vantassel, Kumar, and Cox (2022)

To enable a direct comparison with the architecture proposed by Vantassel, Kumar, and Cox (2022), we now present their predictions using the same eight models. Again, we note that all preprocessing steps, including data normalization, augmentation, and hyper-parameter-tuning, were executed for this model in the same manner as the previous one to ensure an unbiased comparison. Figures 7 and 8 present the prediction and pixel-by-pixel residual plot for the same eight models shown in Figure 2. The comparison of MAPE for predictions by Vantassel, Kumar, and Cox (2022) with those predicted
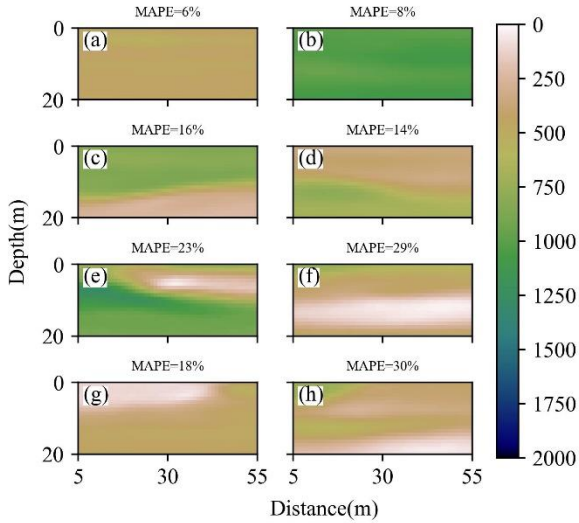
Figure 5. V_s images predicted by the proposed model from the seismic waveforms associated with the subsurface model's whose true V_s image is shown in Figure 2. Above each predicted V_s image is the associated mean absolute percent error (MAPE).
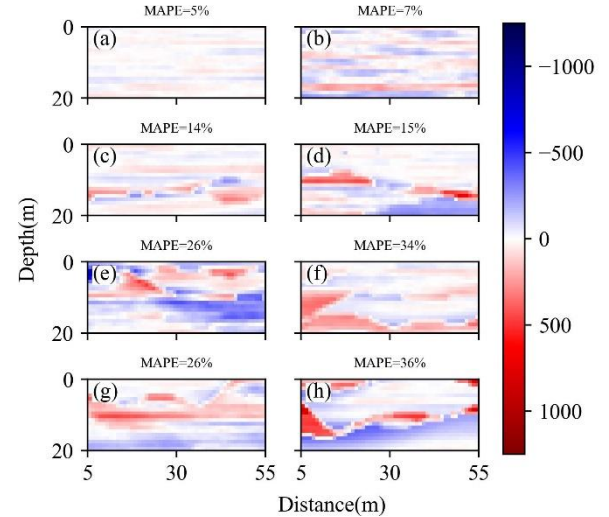

Figure 6. Pixel-by-pixel V_s residuals for the predictions made by the proposed model. The residual images here are the difference between the V_s images in Figures 2 and 5. Above each residual image is the predictions associated mean absolute percent error (MAPE).

Figure 7. V_s images predicted by the revised Vantassel, Kumar, and Cox (2022) model using the seismic waveforms from subsurface model's whose true V_s image is shown in Figure 2. Above each predicted V_s image is the associated mean absolute percent error (MAPE).

Figure 8. Pixel-by-pixel V_s residuals for the predictions made by the Vantassel, Kumar, and Cox (2022) model. The residual images here are the difference between the V_s images in Figures 2 and 7. Above each residual image is the predictions associated mean absolute percent error (MAPE).

by the proposed model, as depicted in Figures 5 and 6, shows that the newly proposed model performs better than the Vantassel, Kumar, and Cox (2022) model. To demonstrate this more quantitatively, Figure 9 depicts the comparison of the average MAPE of the proposed model with that of Vantassel, Kumar, and Cox (2022) across the entire testing dataset. The proposed model is shown to outperform the Vantassel, Kumar, and Cox (2022) architecture by reducing MAPE by 3%, a substantive reduction.

8. Conclusion

Herein we present ongoing advancements in data-driven near-surface 2D seismic imaging. In particular, we present a novel model development framework for producing complex subsurface models and a new data-driven predictive model architecture for subsurface

imaging. The former has enabled the development of 100,000 subsurface models with a level of complexity not previously attainable. The generated models include dipping soil and rock layers, soil lenses, boulders, and underground utilities; none of which have been considered previously. We simulate field data acquisition along these 100,000 synthetic models, by numerically solving the elastic wave equation using an impulse source at the model's center surrounded by 24 receivers (12 on either side) using the open-source, spectral-element-based software SPECFEM2D. A novel DL model architecture was developed through a rigorous evaluation of more than 40 model architectures and extensive hyperparameter tuning. The proposed DL model demonstrated strong performance on the testing dataset of 10,000 wavefield-image pairs with a MAPE of 19%. Moreover, the proposed model outperformed the previous, best-performing model architecture, developed
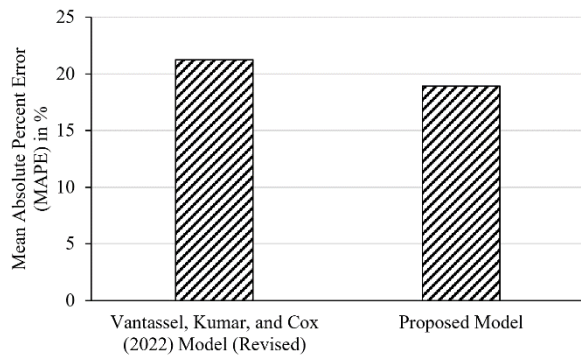
Figure 9. Quantitative comparison of the test accuracy between the revised Vantassel, Kumar, and Cox (2022) model and the model proposed in this study. The model proposed in this study performs better on the testing set with a reduction in the mean absolute percent error (MAPE) of 3% (i.e., from 22% to 19%).

by Vantassel, Kumar, and Cox (2022). Future work will focus on applying this model to real field data and incorporating multiple source positions to increase model resolution. These preliminary results demonstrate the proposed model's robustness in predicting 2D near-surface seismic images and highlight its potential for rapid and reliable non-invasive subsurface imaging for geotechnical engineering applications.

Acknowledgments

References

Abbas, A., J. P. Vantassel, B. R. Cox, K. Kumar, and J. Crocker. 2023. "A Frequency-Velocity CNN for Developing near-Surface 2D vs Images from Linear-Array, Active-Source Wavefield Measurements." *Computers and Geotechnics* 156 (April): 105305. https://doi.org/10.1016/j.compgeo.2023.105305.

Bhochhibhoya, S., R. Maharjan, R. Adhikari, S. Gurung, P. Adhikari, S. Shrestha, and M. S. K. C. 2023. "Suitability Assessment of Stone Columns to Improve Soil Foundation in Kathmandu Valley, Nepal." *Geotechnical and Geological Engineering* 41 (1): 57–73. https://doi.org/10.1007/s10706-022-02262-y.

Chollet, F., and others. 2018. "Keras: The Python Deep Learning Library." *Astrophysics Source Code Library, Record Ascl:1806.022.*

Crisp, M. P., M. B. Jaksa, and Y. L. Kuo. 2017. "The Influence of Site Investigation Scope on Pile Design in Multi-Layered, 2D Variable Ground." In *Geo-Risk 2017*, 390–99. Reston, VA: American Society of Civil Engineers. https://doi.org/10.1061/9780784480717.037.

Datta, D., and M. K. Sen. 2016. "Estimating a Starting Model for Full-Waveform Inversion Using a Global Optimization Method." *GEOPHYSICS* 81 (4): R211–23. https://doi.org/10.1190/geo2015-0339.1.

Fathi, A., B. Poursartip, K. H. Stokoe II, and L. F. Kallivokas. 2016. "Three-Dimensional P- and S-Wave Velocity Profiling of Geotechnical Sites Using Full-Waveform Inversion Driven by Field Data." *Soil Dynamics and Earthquake Engineering* 87 (August): 63–81. https://doi.org/10.1016/j.soildyn.2016.04.010.

Feng, S., Y. Lin, and B. Wohlberg. 2022. "Multiscale Data-Driven Seismic Full-Waveform Inversion With Field Data Study." *IEEE Transactions on Geoscience and Remote Sensing* 60: 1–14. https://doi.org/10.1109/TGRS.2021.3114101.

Jaksa, M. B., J. S. Goldsworthy, G. A. Fenton, W. S. Kaggwa, D. V. Griffiths, Y. L. Kuo, and H. G. Poulos. 2005. "Towards Reliable and Effective Site Investigations." *Géotechnique* 55 (2): 109–21. https://doi.org/10.1680/geot.55.2.109.59531.

Kallivokas, L. F., A. Fathi, S. Kucukcoban, K. H. Stokoe, J. Bielak, and O. Ghattas. 2013. "Site Characterization Using Full Waveform Inversion." *Soil Dynamics and Earthquake Engineering* 47 (April): 62–82. https://doi.org/10.1016/j.soildyn.2012.12.012.

Kingma, D. P., and J. Ba. 2015. "Adam: A Method for Stochastic Optimization." In *3rd International Conference for Learning Representations*. San Diego. https://doi.org/https://doi.org/10.48550/arXiv.1412.6980.

Levander, A. R. 1988. "Fourth-order Finite-difference *P-SV* Seismograms." *GEOPHYSICS* 53 (11): 1425–36. https://doi.org/10.1190/1.1442422.

Mirzanejad, M., and K. T. Tran. 2019. "3D Viscoelastic Full Waveform Inversion of Seismic Waves for Geotechnical Site Investigation." *Soil Dynamics and Earthquake Engineering* 122 (July): 67–78. https://doi.org/10.1016/j.soildyn.2019.04.005.

Mojica, O. F., and N. Kukreja. 2019. "Towards Automatically Building Starting Models for Full-Waveform Inversion Using Global Optimization Methods: A PSO Approach via DEAP + Devito," May.

Monteiller, V., S. Chevrot, D. Komatitsch, and Y. Wang. 2015. "Three-Dimensional Full Waveform Inversion of Short-Period Teleseismic Wavefields Based upon the SEM–DSM Hybrid Method." *Geophysical Journal International* 202 (2): 811–27. https://doi.org/10.1093/gji/ggv189.

Naghibi, F., G. A. Fenton, and D. V. Griffiths. 2014. "Serviceability Limit State Design of Deep Foundations." *Géotechnique* 64 (10): 787–99. https://doi.org/10.1680/geot.14.P.40.

Nocedal, J., and S. J. Wright. 2006. *Numerical Optimization*. New York, NY: Springer New York.

Pan, Y., L. Gao, and T. Bohlen. 2019. "High-Resolution Characterization of Near-Surface Structures by Surface-Wave Inversions: From Dispersion Curve to Full Waveform." *Surveys in Geophysics* 40 (2): 167–95. https://doi.org/10.1007/s10712-019-09508-0.

Peter, D., D. Komatitsch, Y. Luo, R. Martin, N. Le Goff, E. Casarotti, P. Le Loher, et al. 2011. "Forward and Adjoint Simulations of Seismic Wave Propagation on Fully Unstructured Hexahedral Meshes." *Geophysical Journal International* 186 (2): 721–39. https://doi.org/10.1111/j.1365-246X.2011.05044.x.

Savitzky, Abraham., and M. J. E. Golay. 1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." *Analytical Chemistry* 36 (8): 1627–39. https://doi.org/10.1021/ac60214a047.

Tarantola, A. 1984. "Inversion of Seismic Reflection Data in the Acoustic Approximation." *GEOPHYSICS* 49 (8): 1259–66. https://doi.org/10.1190/1.1441754.

Tromp, J., D. Komatitsch, and Q. Liu. 2008. "Spectral-Element and Adjoint Methods in Seismology." *Communications in Computational Physics* 3 (1): 1–32.

Vai, R., J. M. Castillo-Covarrubias, FranciscoJ. Sánchez-Sesma, D. Komatitsch, and J.-P. Vilotte. 1999. "Elastic Wave Propagation in an Irregularly Layered Medium." *Soil Dynamics and Earthquake Engineering* 18 (1): 11–18. https://doi.org/10.1016/S0267-7261(98)00027-X.

Vantassel, J. P., K. Kumar, and B. R. Cox. 2022. "Using Convolutional Neural Networks to Develop Starting Models for Near-Surface 2-D Full Waveform Inversion." *Geophysical Journal International* 231 (1): 72–90. https://doi.org/10.1093/gji/ggac179.

Virieux, J. 1986. "*P-SV* Wave Propagation in Heterogeneous Media: Velocity-stress Finite-difference Method." *GEOPHYSICS* 51 (4): 889–901. https://doi.org/10.1190/1.1442147.

Yust, M. B. S., B. R. Cox, J. P. Vantassel, P. G. Hubbard, C. Boehm, and L. Krischer. 2023. "Near-Surface 2D Imaging via FWI of DAS Data: An Examination on the Impacts of FWI Starting Model." *Geosciences* 13 (3): 63. https://doi.org/10.3390/geosciences13030063.