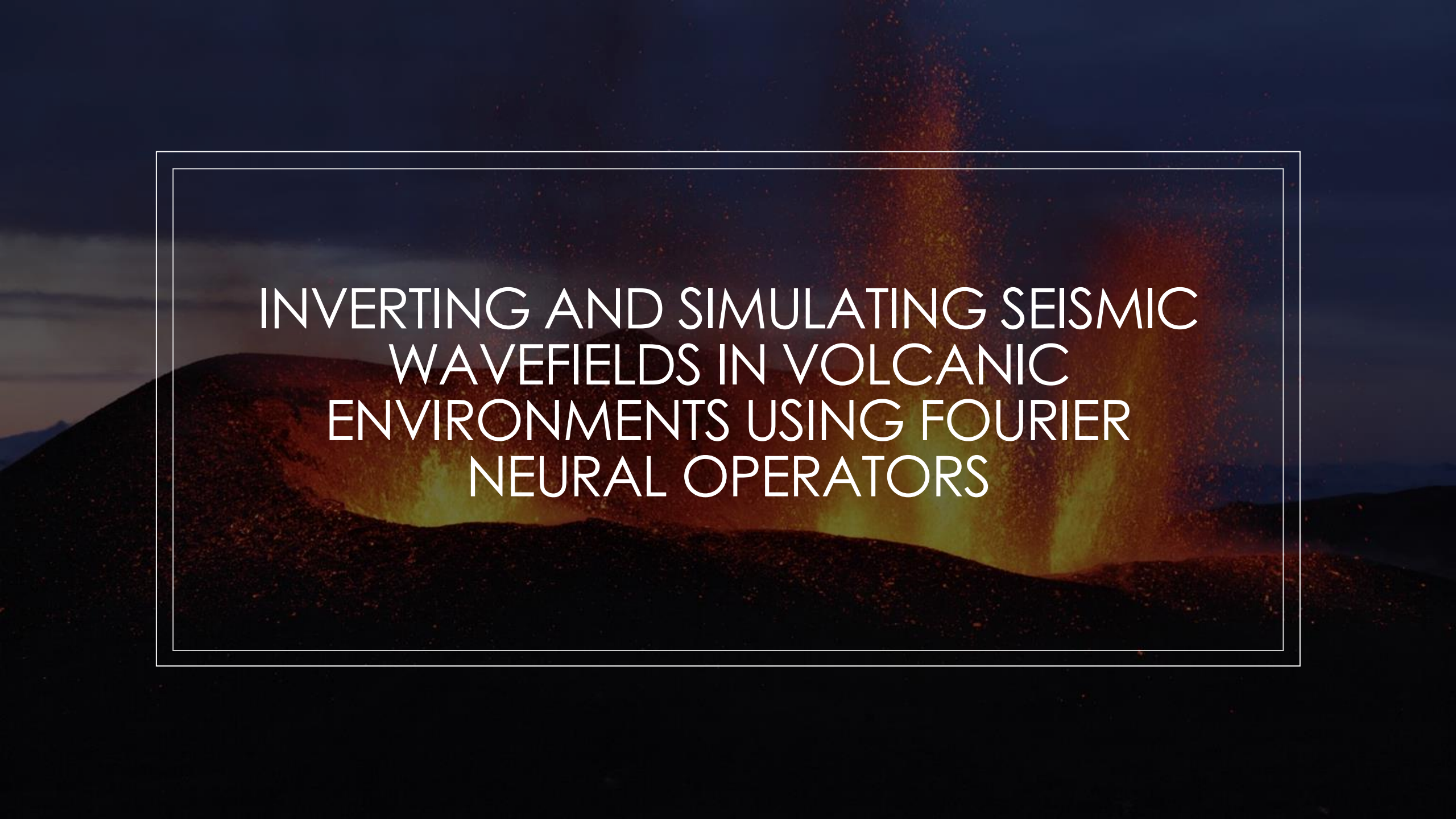
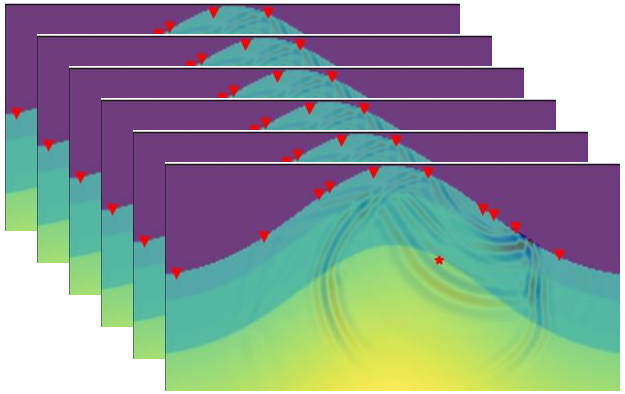
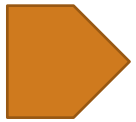
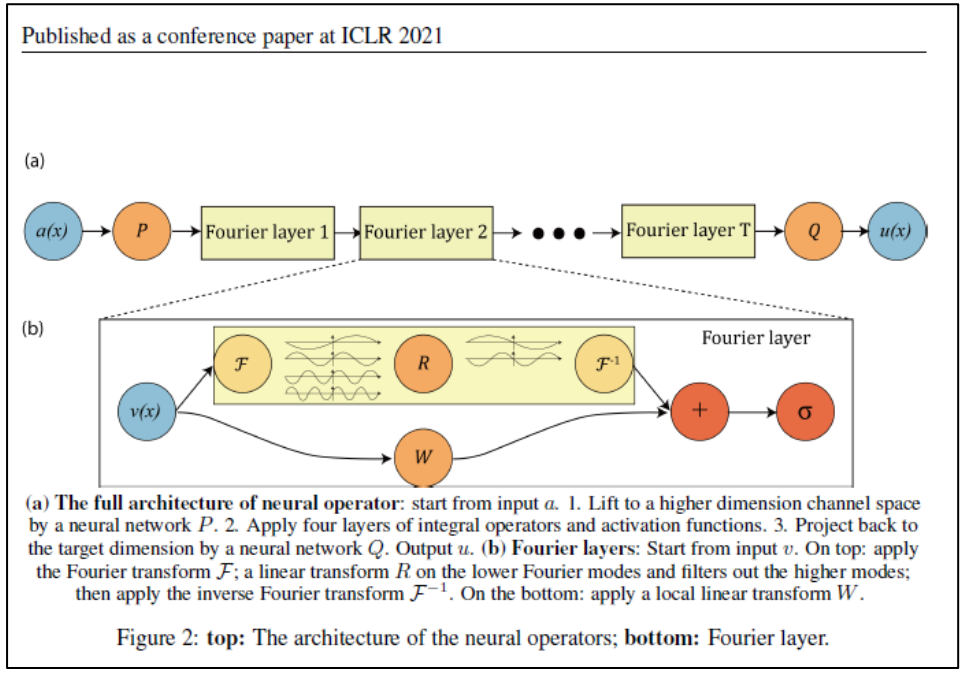


INVERTING AND SIMULATING SEISMIC WAVEFIELDS IN VOLCANIC ENVIRONMENTS USING FOURIER NEURAL OPERATORS

Models In

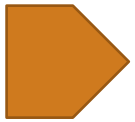




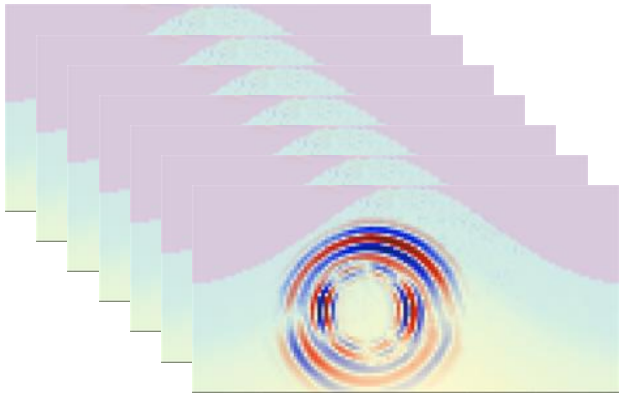


(a) **The full architecture of neural operator**: start from input a. 1. Lift to a higher dimension channel space by a neural network P. 2. Apply four layers of integral operators and activation functions. 3. Project back to the target dimension by a neural network Q. Output u. **(b) Fourier layers**: Start from input v. On top: apply the Fourier transform $\mathcal{F}$; a linear transform R on the lower Fourier modes and filters out the higher modes; then apply the inverse Fourier transform $\mathcal{F}^{-1}$. On the bottom: apply a local linear transform W.

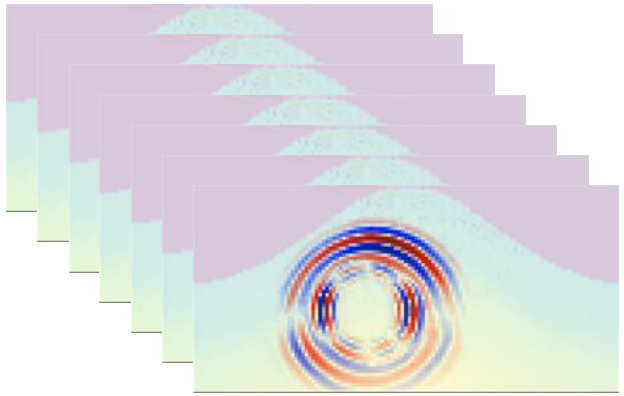
Figure 2: **top:** The architecture of the neural operators; **bottom:** Fourier layer.

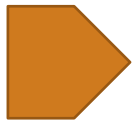


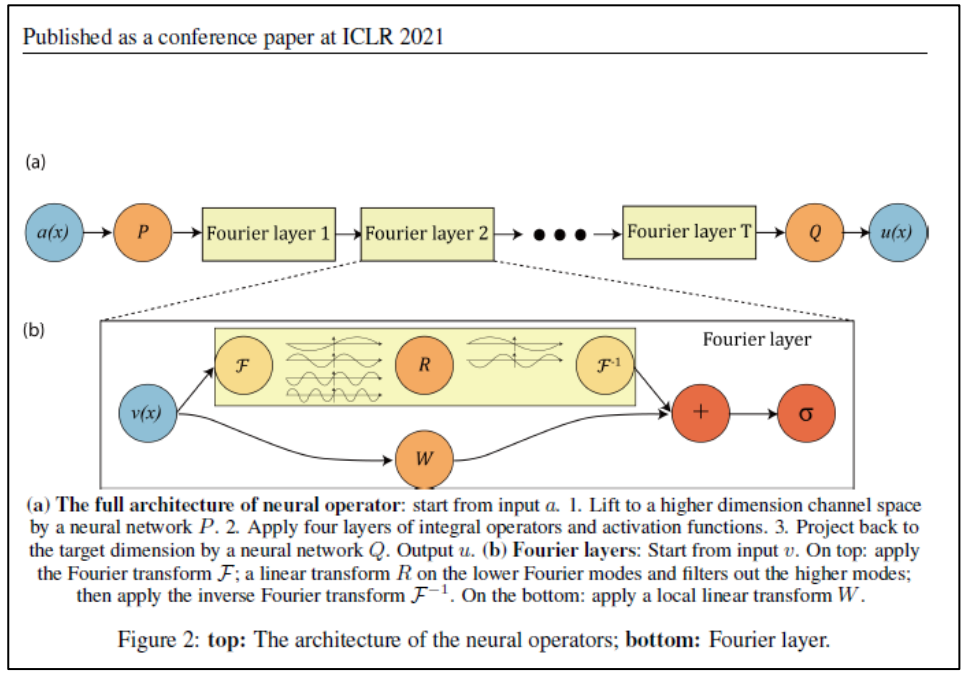
Wavefields Out



Wavefield In

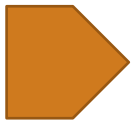




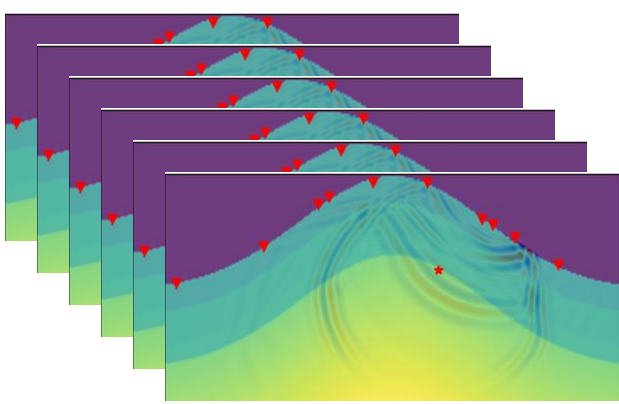


(a) **The full architecture of neural operator**: start from input a. 1. Lift to a higher dimension channel space by a neural network P. 2. Apply four layers of integral operators and activation functions. 3. Project back to the target dimension by a neural network Q. Output u. **(b) Fourier layers**: Start from input v. On top: apply the Fourier transform $\mathcal{F}$; a linear transform R on the lower Fourier modes and filters out the higher modes; then apply the inverse Fourier transform $\mathcal{F}^{-1}$. On the bottom: apply a local linear transform W.

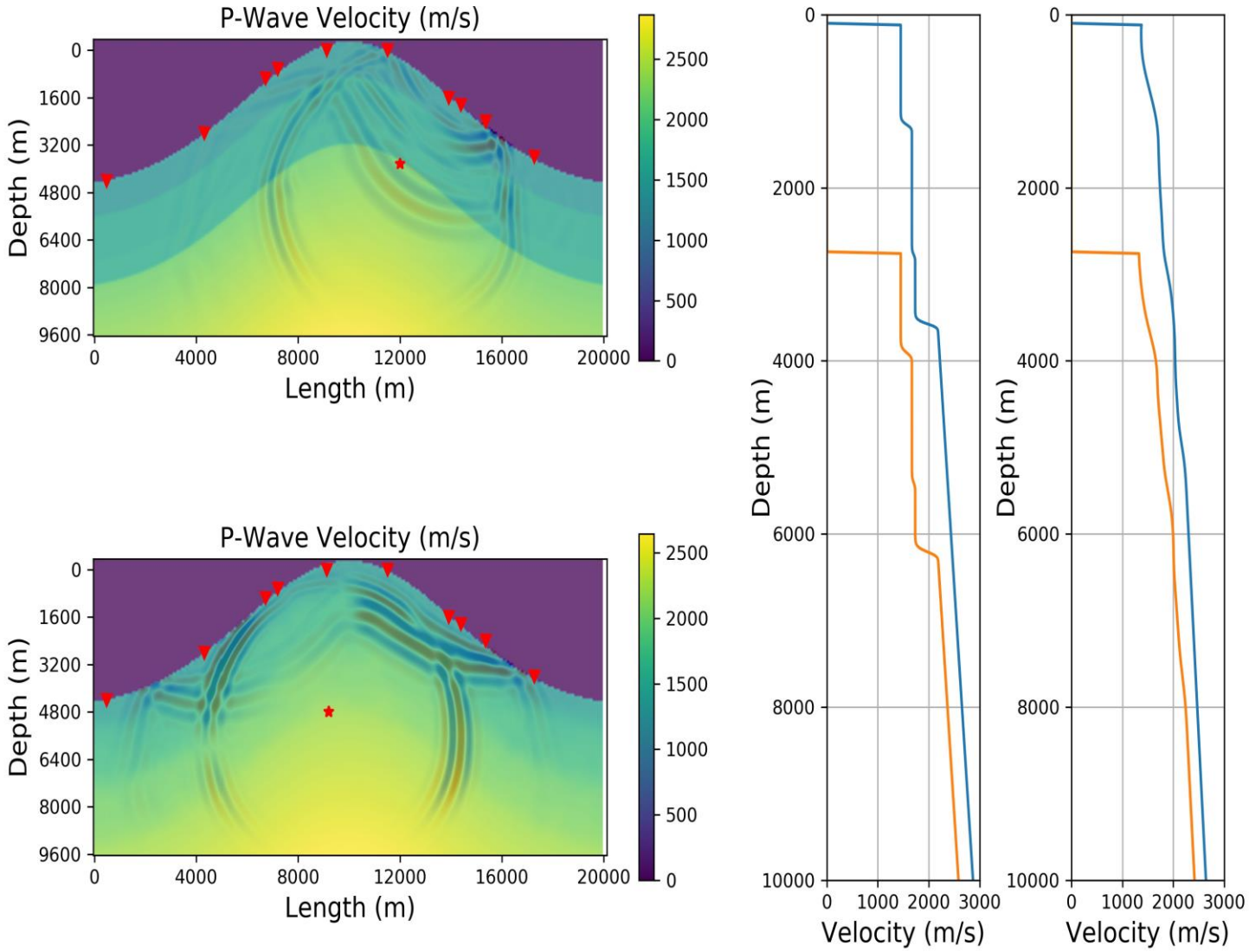
Figure 2: **top:** The architecture of the neural operators; **bottom:** Fourier layer.



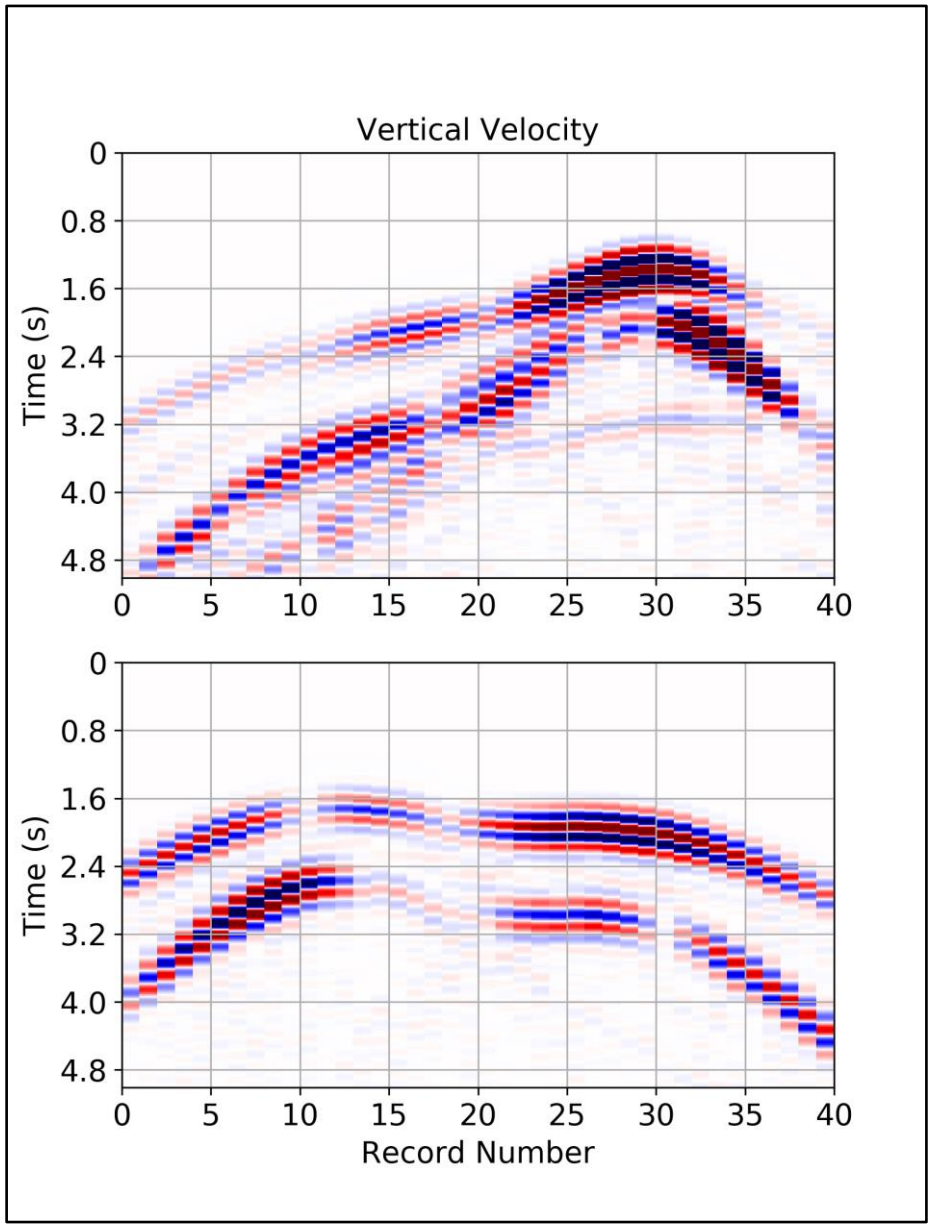
Model Out



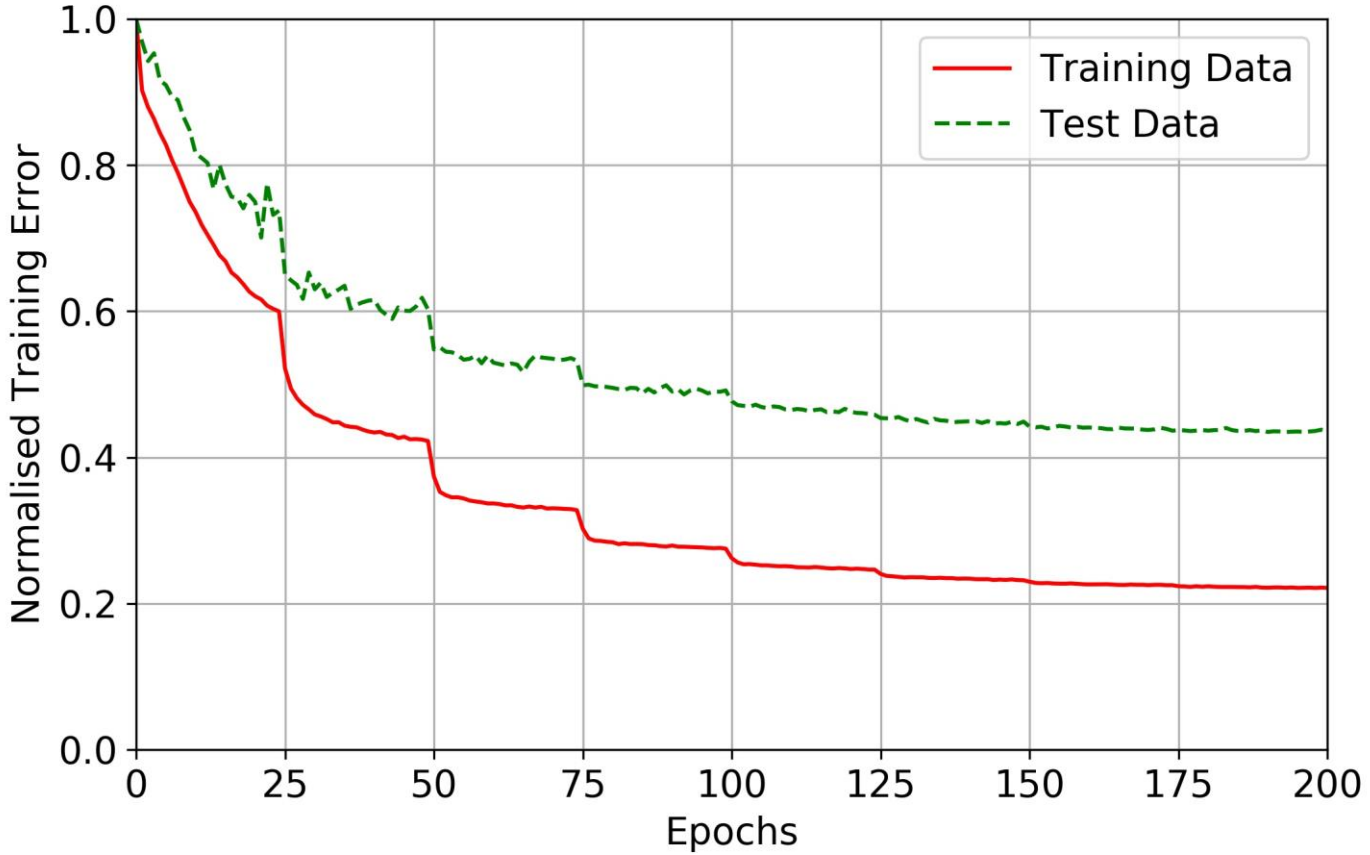
Training Datasets



Two different examples of elastic models used in this study are shown and are taken from two datasets of 20,000+ different models. The upper model is generated with a random velocity gradient and four distinct velocity layers with random thicknesses and velocities. The left-hand profile shows the velocity from two different locations in this model. The bottom velocity model represents a smooth tomography-like model with random smooth perturbations embedded into a random velocity gradient. The right-hand figure shows the velocity profiles from two different locations in this model. A sample numerically simulated seismic wavefield is shown overlain on the models which is generated from a random location (shown by the *) with a random moment tensor.

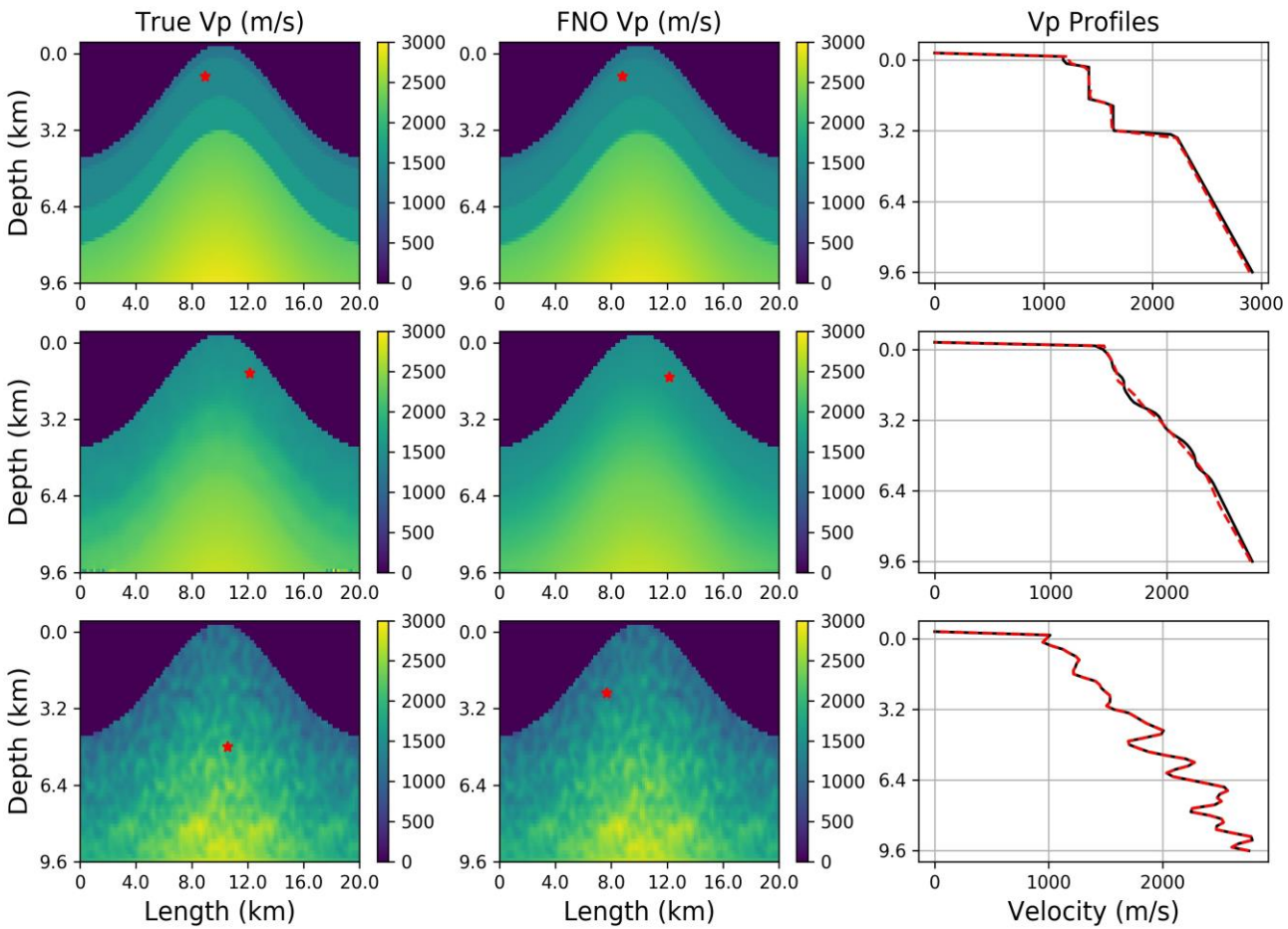


Virtual seismic sensors are distributed across the surface of all models and the simulated seismograms along with the total wavefield is outputted as part of the datasets. The image shows the vertical seismic traces from the models shown in the previous slide for the layered (top) and tomography models (bottom). Having 40 sensors distributed uniformly across the surface is not normally the case so in the study, a random sample of 10 sensors (shown as triangles in the previous image) is chosen as the input data for the networks used in the inversion.



Training error for one of the network learning runs. Here the training data consists of 19,500 input simulations and 500 simulations used for the test data.

Source location, moment and velocity model inversion results using the FNO



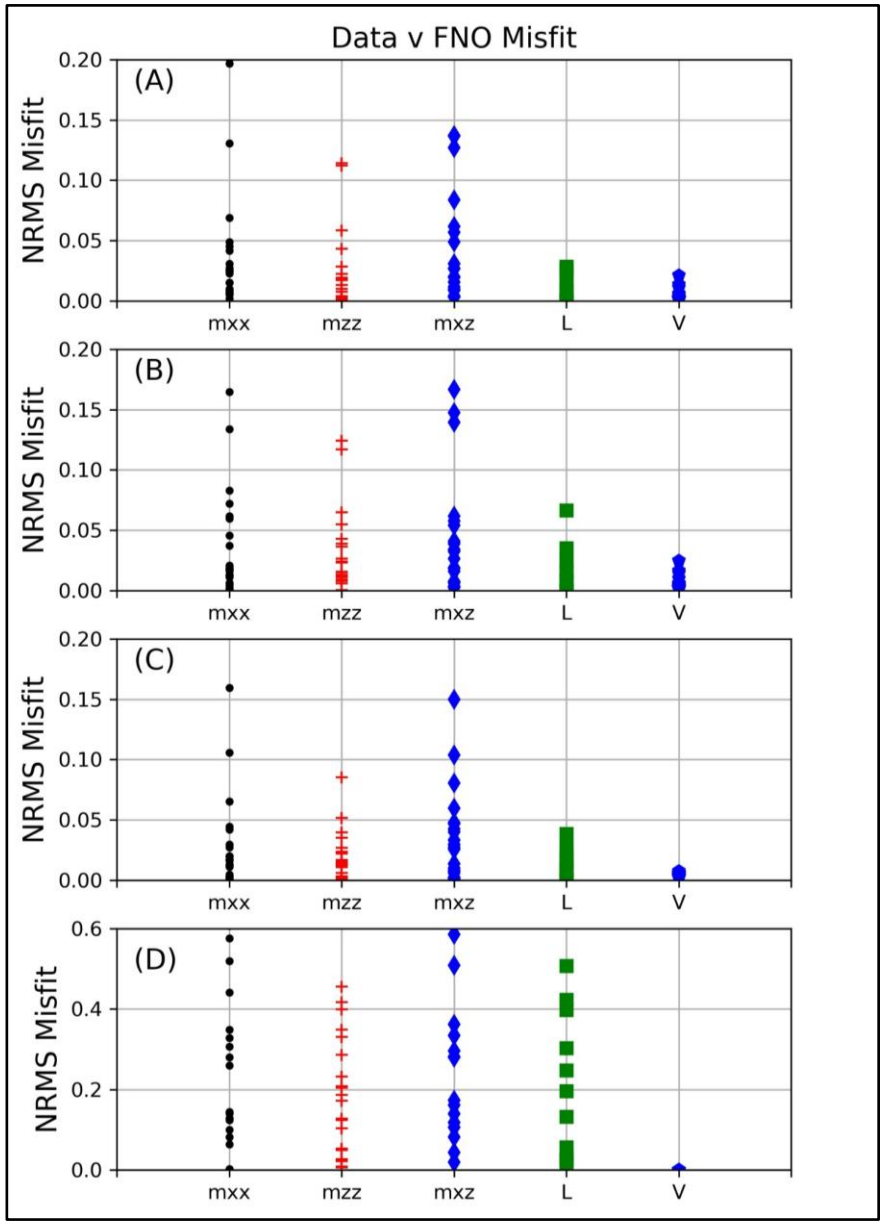
The seismic traces from the 10 random sensors shown in the first figure are used to train the neural networks. In all cases, signal to noise of 15% was added to the input seismic traces.

The left-hand panels show the true models with the seismic source location highlighted by the asterisks. These models were not used in the training dataset. The output results from the networks from the unseen sensor input is shown in the middle panels. The right-hand panels show a profile through the models comparing the true model with the predicted FNO result.

The bottom model is from a different model distribution, but the results are still generated using the network trained with the layered dataset.

The resultant inversion for the velocity is excellent in all cases as is the source location when the network is trained with the same model distribution.

Source location, moment and velocity model inversion normalised RMS misfit



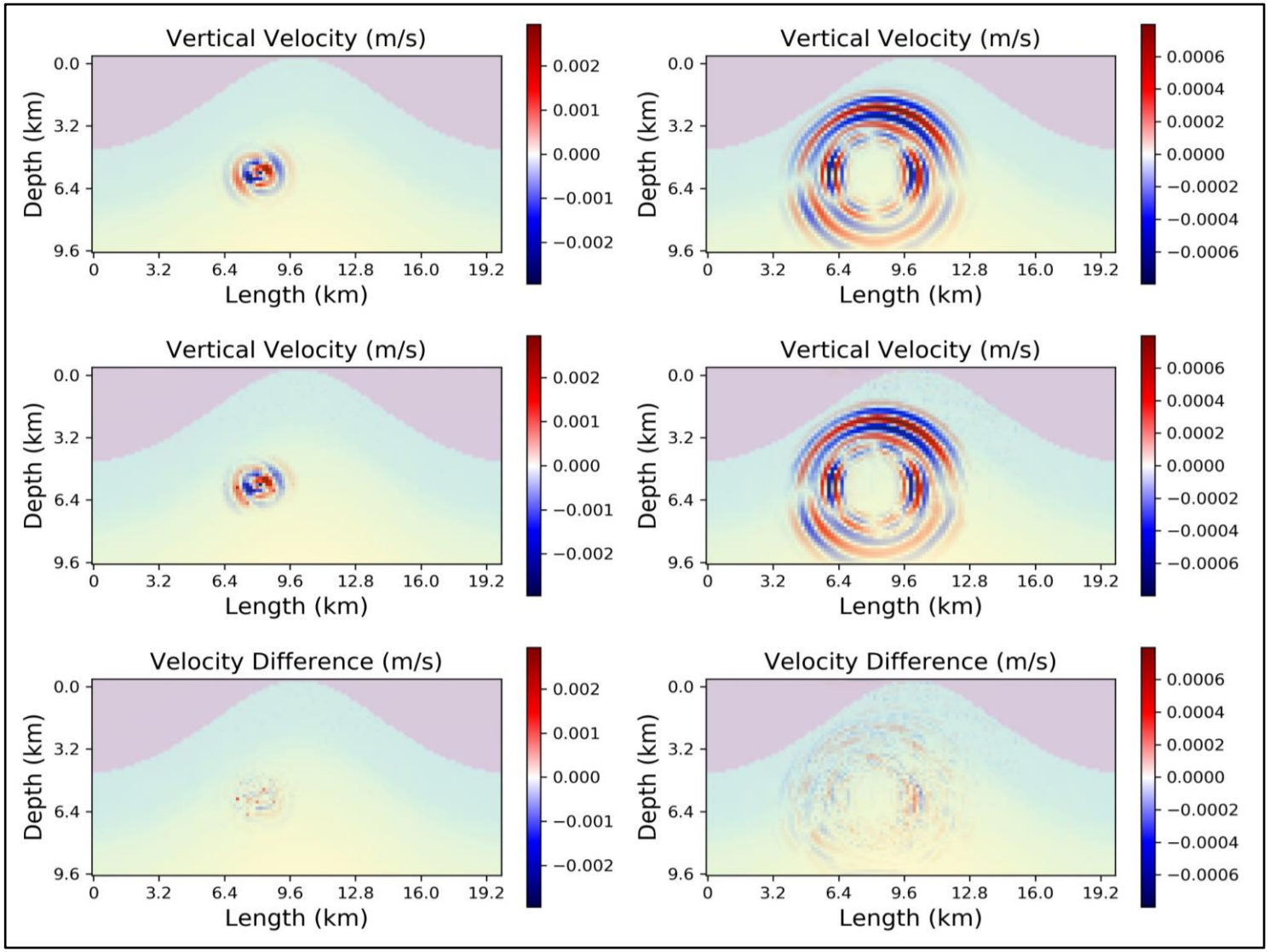
(A) The normalised RMS misfit is shown for 20 different inversions using unseen samples from the layered model. The signal to noise during the training is zero in this case. The velocity model (V) is well resolved by the inversion as is the source location (L). However, the error in the source moment is large across all three tensor components (mxx, mzz, mxz).

(B) The normalised RMS misfit is shown for 20 different inversions using unseen samples from the layered model. The signal to noise during the training is 15% in this case. The velocity model is well resolved by the inversion as is the source location. However, the error in the source moment is large across all three tensor components.

(C) The normalised RMS misfit is shown for 20 different inversions using unseen samples from the tomography model. The signal to noise during the training is zero in this case. The velocity model is well resolved by the inversion as is the source location. However, the error in the source moment is large across all three tensor components

(D) The normalised RMS misfit is shown for 20 different inversions using unseen samples from the unknown model distribution. The unknown model inputs are inverted with the network trained using the layered model with 15% noise. The velocity model is still well resolved by the inversion but the misfit in the source location is large, as is the moment tensor.

Forward modelling seismic wavefields



The inversion results can be interpreted as the network learning the inverse Green's function for the elastic wave operator. The network can also be trained to generate the Green's function for a particular model by using the models as the input and the wavefield (or seismic traces) as the output.

The top panel shows the results from a numerical simulation directly solving the equations with a sample from the unseen layered model data. The left-hand side is for a snapshot after 1 second and the right-hand panel is after 3 seconds. The middle panel shows the results from the FNO network using the layered dataset to train.

The bottom panels show the difference. The overall match is very good though the network can struggle to capture the smaller amplitude scattering. This failure to accurately predict the scattering is greater when using very heterogeneous training models.

In reality, the full wavefield is unknown and is only sampled by surface seismic sensors. Instead of using the full wavefield to train as in the previous case, the 40 seismic traces distributed across the surface are used to train the network.

The top panel (A) shows the results from a numerical simulation directly solving the equations with a sample model from the unseen layered data.

Panel (B) shows the results from the FNO network using the same layered model as in panel (A) to predict the seismic traces.

Panel (C) show the difference between the simulation traces and the predicted network outputs. The overall match is very good as in the full wavefield case and again, the network can struggle to capture the smaller amplitude scattering. Some edge artefacts from the training set-up can also be observed.

The bottom panel (D) shows three simulated traces compared with the FNO predictions. In this plot, the smaller scale amplitude mis-match can be easily observed.

The accuracy of the FNO in predicting the exact result may be considered large relative to output from numerical simulations solving the PDEs. However, the approximation of the Green's function encoded in the network carries enough accurate information to allow for the inversion of an accurate velocity model.

Forward modelling seismic traces