

Response to Reviewer 2

Manuscript geosciences-4554261, “Time-Lapse Seismic Inversion Using a 3D Neural Operator”
Xiao Ma and Tariq Alkhalifah

We thank the reviewer for the careful reading and constructive comments. Changes in the manuscript are highlighted (additions in blue, deletions in red). Section, equation, figure and table numbers below refer to the revised manuscript.

Comment 1

The forward mapping implemented in this study is based on post-stack reflectivity and wavelet convolution. It would be helpful to state the neural operators more explicitly in the Introduction and Discussion so that readers clearly understand that the current neural operator approximates a post-stack convolutional forward operator rather than a full wave-equation propagator.

Response. We agree and have made this explicit. The abstract no longer motivates the method with the cost of wave-equation simulations. The last paragraph of the Introduction now states that the training data are generated with the normal-incidence convolutional model, so the network approximates a post-stack convolutional forward operator and not a wave-equation propagator. A new Discussion section (Section 4.1, “What the neural operator approximates”) describes what this operator does and does not model (multiples, transmission losses, offset-dependent amplitudes, diffractions), and what role it plays in the inversion.

Changes. Abstract; Introduction, last two paragraphs; Section 4.1.

Comment 2

Please provide more details on the two-channel input strategy. During inversion the first channel is iteratively updated and the second channel remains the initial model? What’s the useful for the unchanged second channel?

Response. Thank you for pointing out this ambiguity; the original text was wrong on this point. During training, the first channel holds a Gaussian-smoothed copy of the true model, $S_\sigma \mathbf{I}$, and the second channel holds the true model $\mathbf{I}$. During inversion, both channels start from the initial model; the *first* channel is kept fixed and the *second* channel (the one that held the true model during training) is updated. The revised Equations (8)–(10) show both arguments of the operator explicitly.

The fixed channel plays the role it had during training. It tells the operator which smooth background the model in the second channel is a perturbation of, and because the network has only seen inputs in which the second channel is a detailed version of the first, it anchors the update to the background trend. To test this, we added an ablation on the Kimberlina data (Section 3.3.1, Table 2). It compares the proposed scheme with two alternatives: both channels updated together, and a first channel recomputed at every iteration as a smoothed copy of the updated model. The results support the fixed channel:

- When both channels are updated together, the inversion fails. The data misfit falls only to 22% of its initial value (2% with the fixed channel), and the model ends up less accurate than the initial model.
- Recomputing the first channel as a smoothed copy of the update converges almost as well, but gives a slightly less accurate model (relative error 3.10% versus 3.01%).

The first channel therefore has to carry a smooth background that does not follow every detail of the update, and fixing it at the initial model gives the most accurate result.

Changes. Section 2.4 and Equations (8)–(10); ablation in Section 3.3.1 and Table 2.

Comment 3

The construction of the initial impedance models should be described more explicitly. The training-data generation procedure should provide the ranges of impedance, layer thickness, fault displacement, fold geometry, and noise levels.

Response. Section 2.3 now reports the parameters of the training models: model size, layer thicknesses, number of faults, maximum dip and noise levels. The initial models are now described in detail:

- *Sleipner*: the well log of well 15/9-13 is smoothed and extended laterally, so the initial model contains no lateral (time-slice) information. The same model is used for all four vintages.
- *Kimberlina*: the true impedance is smoothed with a 3D Gaussian filter ($\sigma = 15$ samples, i.e., 150 m), which removes the layering and keeps only the large-scale trend.

The inversion settings (100 Adam iterations, step size, amplitude scalar and TV weight) are listed for both data sets and are the same for all vintages of a data set.

Changes. Sections 2.3 and 3.3; Figure 6.

Comment 4

The ground truth is available for the Kimberlina synthetic experiment, the authors may consider reporting one or two quantitative metrics, such as relative error, RMSE, or correlation coefficient, for the recovered impedance or time-lapse impedance differences?

Response. We added Table 1. It lists the relative error, RMSE and correlation coefficient of the recovered impedance for Years 0, 20 and 200, over the whole volume and over the reservoir interval. The initial model is included as a reference row, so the table shows what the inversion adds. The main results are:

- For all three vintages, the inversion reduces the error of the initial model. Over the whole volume, the relative error falls from 3.1–3.2% to 2.8–2.9% and the correlation rises from 0.978–0.979 to 0.982–0.983. In the reservoir interval, the RMSE falls by 8–10%.
- The correlation between the model update and the true residual is 0.53–0.63, so the update points in the right direction.
- The improvement for Years 20 and 200, which were not used for fine-tuning, is as large as for Year 0.

Changes. Section 3.3, Table 1.

Comment 5

It would be useful to show inversion-loss curves for representative Kimberlina and Sleipner examples.

Response. We added two figures showing the normalized data misfit versus iteration for every patch (median and 10–90% range): Figure 8 for the Kimberlina Year-20 and Year-200 inversions, and Figure 13 for the Sleipner 1994, 2001 and 2006 inversions. On Kimberlina, the misfit falls to about 2% of its initial value; on Sleipner, it falls to 22–27% after the 100 iterations used for all field results.

Changes. New Figures 8 (Section 3.3) and 13 (Section 3.4).

Comment 6

Some figure-label descriptions should be made consistent throughout the manuscript.

Response. Thank you. We regenerated or relabelled all multi-panel figures so that the panel letters and titles use the same font and printed size, with the letter at the left and the title centred over the panel, and we used one vocabulary and one label format throughout:

- panel titles now read “(x) Initial model”, “(x) True AI, Year n ”, “(x) Inverted AI, 2001”, “(x) Observed data, 2001”, “(x) Predicted data” and “(x) Difference”;
- “initial model” is used throughout, instead of alternating with “background model”;
- colorbars read “AI (m/s·g/cm³)”, “ Δ AI (m/s·g/cm³)” or “Amplitude”;
- axes read “Time (ms)” or “Depth (km)”, “Crossline” and “Inline”.

Response. The network is now called U-NO (U-shaped neural operator) throughout, including in the schematic of Figure 3, consistent with the cited architecture (Rahman et al., 2022); the submitted version used “UFNO”, which is the name of a different architecture.

Comment 7

Page 7, Equations (8)–(10): There seems to be a notational inconsistency between the equations and the actual network input.

Response. Thank you. Equations (8)–(10) now write the operator with both inputs, $\mathcal{G}_\theta(\mathbf{I}_0, \mathbf{I})$. The training objective uses $\mathbf{I}_0 = S_\sigma \mathbf{I}_i$, and the inversion objective fixes $\mathbf{I}_0 = \mathbf{I}_{\text{init}}$. The inversion objective now also shows the amplitude scalar and the total-variation term used in the implementation, which the original Equation (10) omitted.

Comment 8

Page 14, Figure 9 caption: “inlince” is a typographical error?

Response. Corrected (this figure is now Figure 10). We also proofread the manuscript for similar errors.

Comment 9

Page 16, Lines 367–374: The expressions “bypasses this trade-off entirely” and “parameter-free behavior” are somewhat too strong.

Response. We agree. The paragraph now says that the neural operator is “less exposed to” the trade-off. It states that the inversion does have settings (amplitude scalar, TV weight, step size, number of iterations and smoothing of the initial model), that these were chosen once per data set and kept fixed for all vintages, and that different values were used for Kimberlina and Sleipner. “Parameter-free” has been replaced by “fixed settings” throughout. We also toned down other statements in the same spirit (abstract, Figure 6 and Figure 10 discussions, and the Figure 14 caption, in the submitted numbering).