Reconstructing missing seismic traces on BP 2007 and Viking Line 12

Comparing U-Net, SwinV2, and SFM on synthetic and field data

CSE3000 Research Project
June 2026

Author	Jaouad Hidayat
Affiliation	EEMCS, Delft University of Technology, The Netherlands
Repository	https://github.com/jaouadhidayat/seismic-trace-reconstruction

Supervisors

Dr. Jing Sun	EEMCS, Delft University of Technology, The Netherlands
Dr. Eric Verschuur	Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands
Dr. Tiexing Wang	R&D Department, Shearwater GeoServices, United Kingdom
Jiahua Zhao	Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands

Examiner

Dr. Petr Kellnhofer	EEMCS, Delft University of Technology, The Netherlands

Reconstructing missing seismic traces on BP 2007 and Viking Line 12

Comparing U-Net, SwinV2, and SFM on synthetic and field data

Jaouad Hidayat
EEMCS, Delft University of Technology, The Netherlands
Supervisors: Jing Sun, Eric Verschuur, Tiexing Wang, and Jiahua Zhao

Abstract

Marine seismic surveys can contain missing or unusable receiver traces. This paper tests how well those traces can be reconstructed. It compares zero fill, linear interpolation, a U-Net trained from random initialization on BP, SwinV2 with ImageNet weights, and SFM with seismic pretraining. BP 2007 supplies complete synthetic shots, so receiver traces can be removed from the input and scored against the target. Viking Line 12 is a field line; the test removes observed field traces and scores their reconstruction after the models are trained only on BP. All learned models use inputs computed from visible traces, predict a correction to a linear interpolation estimate, copy measured traces back, and are scored only on removed traces. On BP 2007, with 75% of receiver traces removed in groups of eight, the U-Net has the lowest RMSE on removed traces, 1.170 ± 0.467. SwinV2 with LoRA is the best pretrained method on BP, 2.247 ± 0.196. On Viking Line 12, using BP-trained weights without field retraining, the U-Net has the lowest mean RMSE, 16.38 ± 22.88, but repeat variation is large. Frozen SFM is the best pretrained method on Viking by RMSE, 19.98 ± 0.62, and has the highest SSIM.

1 Introduction

When receiver traces are missing from marine seismic data, a gather no longer sits on a regular trace grid. Interpolation restores that grid for inspection and downstream processing. The important rule is that recorded samples are measurements, not predictions: a useful reconstruction estimates what is missing without quietly changing the measured traces.

The two datasets play different roles. BP 2007 provides complete synthetic shots, so traces can be removed from the input and scored against the target [1]. Viking Line 12 is a real marine line [2]. Because it has no target for traces that were not recorded, the Viking test removes observed traces, reconstructs them with models trained on BP, and scores those observed traces.

This paper compares zero fill, linear interpolation, a U-Net trained from random initialization on BP, SwinV2 pretrained on ImageNet, and SFM pretrained on seismic data. SwinV2 and SFM are tested with frozen encoders and with LoRA. Every learned model receives the same visible-trace inputs, uses the same output rule, and is evaluated only where traces were removed.

The results split by setting. On BP 2007, the U-Net has the lowest RMSE by a large margin. On Viking Line 12, the U-Net has the lowest mean RMSE but the widest repeat variation; frozen SFM is the strongest pretrained model and has the highest SSIM. The paper therefore separates a complete synthetic target from a field line test instead of collapsing them into one model ranking.

2 Background

Traditional seismic interpolation methods exploit event continuity while preserving the samples that were actually recorded. F-X interpolation and POCS remain common reference points for interpolation methods that use seismic structure while preserving recorded samples [3, 4]. In this paper, zero fill and linear interpolation set the scale of the task. Linear interpolation is also one of the inputs given to each learned model.

Neural models are not new to seismic reconstruction. CNNs and U-Net style models have been used for seismic interpolation and denoising [5, 6], and the original U-Net remains a useful dense prediction reference because it combines context with high-resolution skip paths [7]. Image restoration transformer designs such as SwinIR and Restormer mark the wider restoration design space [8, 9]. This paper compares U-Net, SwinV2, and SFM on trace reconstruction.

SwinV2 supplies a hierarchical vision backbone [10]; the weights used here are pretrained on ImageNet-1K, a natural-image classification corpus [11]. SFM supplies a transformer pretrained on seismic data and evaluated on geophysical tasks, including interpolation [12]. LoRA adds trainable low-rank matrices to frozen transformer layers [13]. Recent work has also tested foundation model adaptation across geophysical and seismic processing tasks [14, 15].

The data split matters. Nearby seismic crops from the same shot or gather are correlated, so random crop splitting can make performance look stronger than it is. Spatial dependence between training and test samples is a known source of inflated accuracy in spatial prediction [16]. Here, complete shots or gathers are split before crops are cut, and the main score is computed only on removed traces, not on traces copied from the input.

3 Experiment

Figure 1 shows the reconstruction task before the implementation details. A seismic gather is a two-dimensional array with time on one axis and receiver trace on the other. A mask removes trace columns from the model input. The model receives values derived from visible traces, reconstructs the removed columns, and is scored only on the removed samples. Measured traces are copied back into the output.

3.1 Data and masks

The synthetic benchmark uses all four BP 2007 shot parts [1]. Each shot has 1151 time samples and 800 receiver traces, with an 8 ms time interval and 12.5 m receiver spacing. The scored examples are non-overlapping 256 by 256 center crops cut after the split. Learned models receive a surrounding 512 by 512 context and predict the center crop.

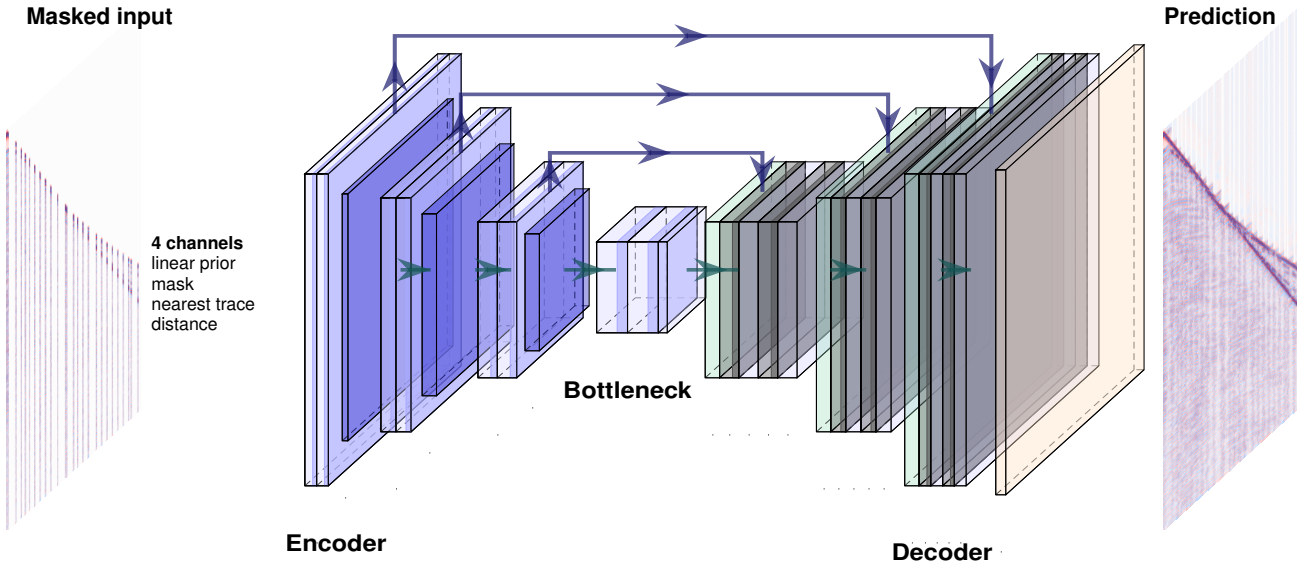

Figure 1: Reconstruction setup. The model sees four channels computed from visible traces: an estimate from linear interpolation, a mask for visible traces, nearest visible trace interpolation, and distance to the nearest visible trace. It predicts a correction for the missing region; measured traces are copied back before scoring. The figure shows the U-Net path, while SwinV2 and SFM use the same inputs and output rule.

BP 2007 is split by independent unit before patch extraction: 70% for training, 15% for validation, and 15% for testing. Normalization is fitted on training units only and then reused. The BP mask removes 75% of receiver traces in groups of eight traces: one group of eight traces is visible, then the next three groups of eight traces are removed. Masks are made on the full shot before crops are cut, so neighboring crops inherit the same receiver pattern.

The group width keeps the scale along the receiver axis close to Viking. Eight traces are 1.0% of an 800-trace BP shot; one trace is 0.83% of a 120-trace Viking gather.

The field test uses patches from Viking Line 12, a public marine field line with 1500 time samples and 120 traces per gather [2]. No Viking training is performed. Each model trained on BP is evaluated with its BP training settings and BP normalization record. For scoring, 75% of observed Viking traces are removed as individual traces and reconstructed.

3.2 Inputs, output rule, and metrics

Every learned model receives four input channels derived only from visible traces: linear interpolation along the receiver axis, a binary channel marking visible traces, nearest visible trace interpolation, and normalized distance to the nearest visible trace. Removed amplitudes are not used as input.

Each learned model predicts a correction to the linear interpolation channel. If x is the target crop, m is one on visible samples and zero on removed samples, and r_θ is the model correction, the reconstruction is

$$x_{\mathrm{rec}} = mx + (1 - m)(x_{\mathrm{linear}} + r_\theta).$$

This copies visible traces exactly and lets the model change only the removed region. The main metric is RMSE on removed samples after the inverse amplitude transform:

$$\mathrm{RMSE}_{\mathrm{removed}} = \sqrt{\frac{1}{N_{\mathrm{removed}}} \sum_{i \in \Omega_{\mathrm{removed}}} (x_i - (x_{\mathrm{rec}})_i)^2}.$$

SSIM is computed on reconstructed crops after the same inverse transform and after measured traces have been copied back [17]. It is a secondary structure metric; the main ranking uses RMSE on removed traces.

3.3 Methods and training

Table 1 lists the methods. The simple baselines have no learned weights. The U-Net is trained from random initialization. SwinV2 uses the SwinV2-Small ImageNet-1K weights [10, 11]. SFM uses the SFM-Base seismic weights [12]. Frozen variants train the input adapter and reconstruction head. LoRA variants also train rank-16 LoRA matrices in attention Q, K, V, and output projections [13].

All learned BP runs use 10 epochs, batch size 32, AdamW [18], automatic mixed precision on CUDA [19], gradient clipping at norm 1 after AMP unscaling, 5% learning rate warmup, and cosine decay to zero. The base learning rate is 10^{-4}. Matrix weights in the reconstruction head use weight decay 10^{-4}; biases and LoRA matrices use no weight decay.

The loss is a Charbonnier amplitude term on removed center samples plus trace-direction derivative terms after measured traces are copied back [20]. The derivative term has weight 1.0 at visible/removed boundaries and 0.25 inside removed groups.

Table 1: Methods in the comparison. The last column reports trainable parameters and total parameters. The reconstruction head is the trainable decoder attached to the backbone.

Method	Trainable weights	Params
Zero fill	none	0 / 0
Linear interpolation	none	0 / 0
U-Net	all weights	4.85M / 4.85M
SwinV2 frozen	adapter and reconstruction head	1.29M / 50.25M
SwinV2 + LoRA	adapter, reconstruction head, LoRA (0.87M)	2.16M / 51.11M
SFM frozen	adapter and reconstruction head	1.29M / 86.70M
SFM + LoRA	adapter, reconstruction head, LoRA (0.88M)	2.18M / 87.58M

Within each repeat, all methods share the same BP sample and mask manifests. BP results use 246 test units, 4 masks, and 3 repeats, giving 2952 matched unit/mask/repeat cases. The Viking test uses 64 patches and 4 masks. RMSE on visible traces is 0.0 because visible traces are copied exactly. The reported run set used 28.1 GPU-hours on NVIDIA L40, A40, and Tesla V100 GPUs on the Delft AI Cluster.

4 Results

The main result is a split: the U-Net trained for this task dominates the BP RMSE test, while Viking changes the order of the pretrained models. On BP 2007, the U-Net has much lower RMSE on removed traces than the pretrained variants. On Viking, the U-Net has the lowest mean RMSE but high repeat variation; frozen SFM is the best pretrained model. Figure 2 places the two rankings next to each other.

4.1 BP 2007: U-Net has the lowest RMSE

On BP 2007, the U-Net has the lowest RMSE on removed traces by a large margin (Table 2). Its mean RMSE is 1.170 ± 0.467, compared with 2.247 ± 0.196 for SwinV2 + LoRA and 2.483 ± 0.169 for frozen SFM. Across matched BP cases, the U-Net beats SwinV2 + LoRA on 100% of cases and SFM + LoRA on 100% of cases.

The baselines show why scoring only the removed traces matters. Under the BP grouped 75% removal pattern, linear interpolation has higher RMSE than zero fill even though it has higher SSIM. The smoother linear estimate can look structurally plausible while carrying larger amplitude errors. The learned models use linear interpolation as an input to correct, not as the final reconstruction.

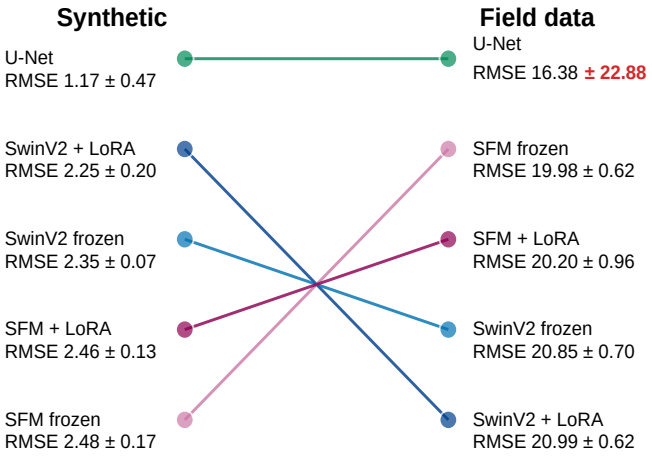

Figure 2: Rank shift between the BP synthetic test and the Viking field line test. Vertical position shows RMSE order within each dataset; the two sides do not share a numeric axis. BP strongly favors the U-Net trained for this task. On Viking, the U-Net has the lowest mean but large repeat variation, while frozen SFM is the strongest pretrained model.

4.2 Pairing shows the size of each difference

The BP comparison is paired by repeat, unit, crop, and mask identity. Figure 3 and Table 3 use the same sign convention: positive values mean the first method has lower RMSE.

The large effect is the U-Net gap. LoRA changes SwinV2 more than it changes SFM on BP: SwinV2 + LoRA wins over frozen SwinV2 in 82.6% of matched cases, while SFM + LoRA is almost split against frozen SFM. Those adaptation differences are small next to the U-Net advantage.

Table 2: Main scores. RMSE is measured only where traces were removed from the input; lower is better. Values are mean ± 95% half-width across the three repeats. SSIM is a secondary structure metric.

Setting	Method	RMSE	SSIM	Main point
BP 2007	U-Net	**1.170 ± 0.467**	0.6909	Lowest BP RMSE.
	SwinV2 + LoRA	2.247 ± 0.196	0.7065	Best pretrained method on BP.
	SwinV2 frozen	2.350 ± 0.071	0.7043	Close to LoRA on BP.
	SFM + LoRA	2.455 ± 0.133	0.6828	Slightly lower RMSE than frozen SFM on BP.
	SFM frozen	2.483 ± 0.169	0.6798	Behind the SwinV2 variants on BP.
	Zero fill	2.980 ± 0.009	0.5891	Lower RMSE than linear under this mask.
	Linear interpolation	3.905 ± 0.006	0.6534	Higher SSIM, larger amplitude error.
Viking Line 12	U-Net	**16.38 ± 22.88**	0.7116	Lowest mean; very large repeat variation.
	SFM frozen	19.98 ± 0.62	0.7345	Best pretrained method on Viking.
	SFM + LoRA	20.20 ± 0.96	0.7100	Below frozen SFM on RMSE and SSIM.
	SwinV2 frozen	20.85 ± 0.70	0.7228	Better than SwinV2 + LoRA on Viking RMSE.
	SwinV2 + LoRA	20.99 ± 0.62	0.7083	LoRA does not improve the Viking mean.

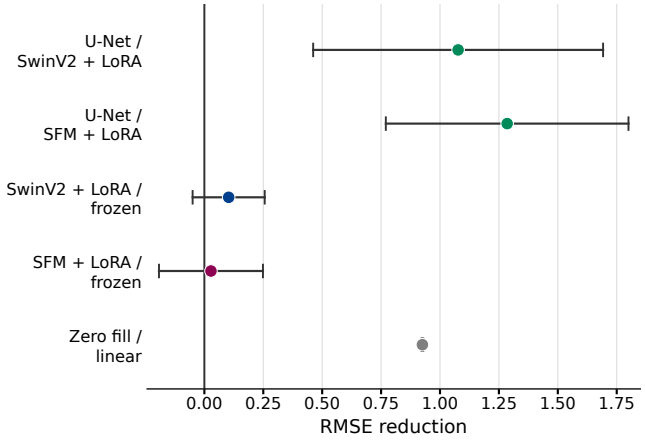

Figure 3: Paired RMSE reductions on BP 2007. Positive values favor the first method. Table 3 gives the exact mean reductions, 95% half-widths, win rates, and denominator.

Table 3: Paired BP differences over 2952 matched unit/mask/repeat cases. Win rate is the percentage of cases where the first method has lower RMSE.

Comparison	RMSE reduction	Win rate
U-Net over SwinV2 + LoRA	1.077 ± 0.615	100%
U-Net over SFM + LoRA	1.285 ± 0.335	100%
SwinV2 + LoRA over frozen SwinV2	0.103 ± 0.153	82.6%
SFM + LoRA over frozen SFM	0.028 ± 0.221	50.5%
Zero fill over linear interpolation	0.925 ± 0.004	100%

4.3 Viking changes the pretrained ranking

On Viking Line 12, the U-Net has the lowest mean RMSE, 16.38, but its 95% half-width is very large (± 22.88). Frozen SFM is therefore the steadier pretrained result by RMSE, 19.98 ± 0.62, and has the highest SSIM. The Viking test puts frozen SFM above the other pretrained methods.

The Viking test uses one field line, no Viking training, BP normalization reused, and individual Viking traces removed instead of grouped BP trace blocks. Within that condition, the models trained on BP still reconstruct observed field traces well enough.

4.4 Selected examples

Figure 4 shows one BP case and one Viking case. The BP row makes the grouped trace removal visible. The Viking row shows the same models trained on BP on a field gather.

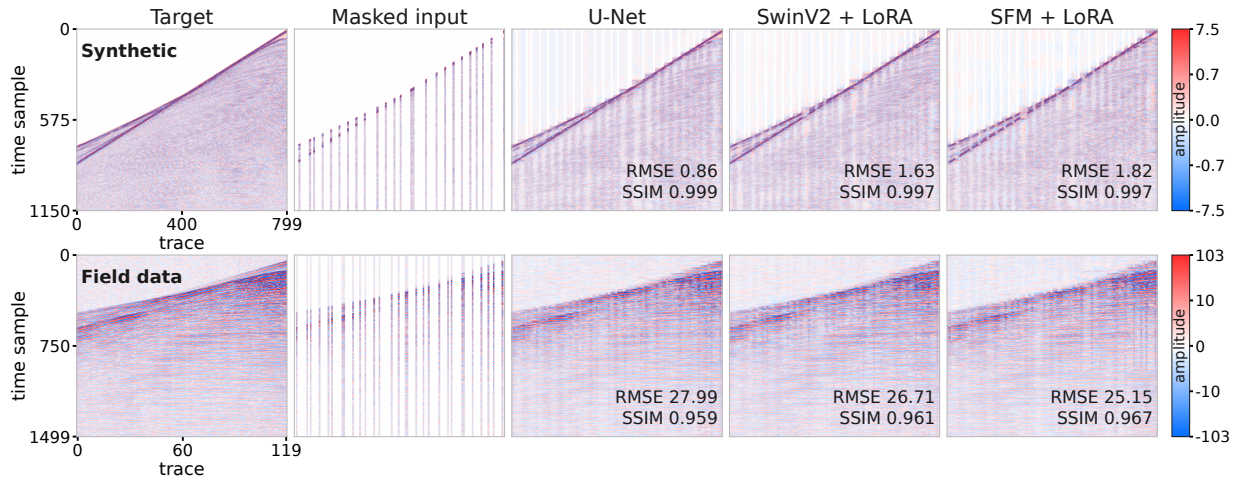

Figure 4: Selected BP 2007 and Viking Line 12 reconstructions. The panels show the reference, masked input, and model predictions for one BP case and one Viking case. Numbers inside panels describe only these examples; aggregate results are in Table 2.

5 Discussion

BP rewards the U-Net trained for this task. The pretrained backbones carry many more total parameters, but this BP test is a local correction problem. The U-Net has dense convolutional skip paths and all of its weights are trained on the BP reconstruction loss, which fits the result.

Viking changes the ranking of the pretrained models. Frozen SFM is behind the U-Net on BP 2007, but it is the strongest pretrained model on Viking Line 12 and has the highest Viking SSIM. The two tests should therefore be read separately: BP scores reconstruction against a complete synthetic target, while Viking removes observed traces from one real line and scores those traces.

The BP and Viking masks are matched in scale along the trace axis. Eight removed BP traces cover 1.0% of an 800-trace shot. One removed Viking trace covers 0.83% of a 120-trace gather. The grouping therefore gives the synthetic shots a scale of removed receiver traces close to the field gathers, rather than making BP artificially fine-grained.

LoRA depends on the backbone and dataset here. It helps SwinV2 on BP mean RMSE and wins most matched SwinV2 cases there. It does not clearly improve SFM on BP, and on Viking both LoRA variants are below their frozen counterparts by mean RMSE and SSIM. The practical choice is to use LoRA per backbone and dataset, not as a default.

Zero fill and linear interpolation are simple references. They show the scale of the task and supply the starting estimates used by the learned models. Operational seismic methods such as POCS and F-X interpolation remain outside these runs, so the claim is about the evaluated methods, starting estimates, masks, and models.

The limits are narrow. Viking is one field line. BP removes traces in groups of eight, while Viking removes individual observed traces. The Viking result therefore combines a change of data source with a change of mask pattern. The three repeats measure variation across repeats for these runs.

For practical use, reconstructed traces must remain labeled as reconstructed. They can support inspection and processing experiments, but they are not measurements. Before a reconstruction affects imaging, reservoir interpretation, acquisition infill, or safety-critical work, it needs domain quality control and validation against the downstream task.

6 Conclusion

Under the BP test with grouped trace removal used here, the U-Net trained for this task is the clear RMSE winner on removed traces. On the Viking field line test, the U-Net has the lowest mean RMSE but large repeat variation, while frozen SFM is the strongest pretrained model and has the highest SSIM.

The useful separation is between a complete synthetic target and a field line test where observed traces are removed from the input and scored. BP 2007 gives exact scoring against a complete target. Viking Line 12 tests the same BP-trained models on a real line without field retraining. Together, the results show that synthetic reconstruction accuracy and the Viking test can point to different model choices.

The narrow lesson is to choose the U-Net trained for this task for the BP mask used here, treat frozen SFM as the strongest pretrained Viking option in these runs, and not treat LoRA as a default improvement.

Acknowledgements

I thank Dr. Jing Sun, Dr. Eric Verschuur, Dr. Tiexing Wang, and Jiahua Zhao for supervision, feedback, and domain guidance during the project.

The experiments ran on the Delft AI Cluster (DAIC) [21].

The BP 2007 benchmark was created by Hemang Shah and provided courtesy of BP Exploration Operation Company Limited. Mobil AVO Viking Graben Line 12 is used through the public SEG/Open Source Geoscience release.

Responsible research and use of generative AI

The main integrity risk is an inflated reconstruction score. The experiment addresses that risk by splitting complete units before patch extraction, fitting normalization on training units only, copying visible traces into the output, and reporting the main error only on traces removed from the input. Sample and mask manifests are saved so paired comparisons can be audited.

The code and experiment materials are available in the project repository. The repository contains the implementation, configuration files, split records, masks, normalization records, seeds, analysis code, result tables, and figure scripts used for this paper. Raw seismic files are not redistributed where source terms do not permit it; the paper cites the public data sources instead.

A separate use risk is mistaking a reconstructed trace for a measurement. In real processing or interpretation, reconstructed traces should remain labeled, reviewed by domain experts, and tested against downstream objectives before they influence decisions about imaging, reservoir interpretation, acquisition infill, or safety-critical work. The method supports reconstruction experiments; it does not replace acquisition metadata or geophysical judgment.

Generative AI tools were used for coding and editing. All results and claims were checked against experiment outputs, code, or cited sources. The author assumes full responsibility for the content of this document.

References

[1] SEG Wiki, "2007 BP anisotropic velocity benchmark," Open seismic benchmark dataset, 2007, dataset created by Hemang Shah and provided courtesy of BP Exploration Operation Company Limited. [Online]. Available: https://wiki.seg.org/wiki/2007_BP_Anisotropic_Velocity_Benchmark

[2] Open Source Geoscience, "Mobil AVO Viking Graben Line 12," SEG File Publications No. 4, data released for the 1994 SEG workshop, 1994, R. G. Keys and D. J. Foster, eds. [Online]. Available: https://s3.amazonaws.com/open.source.geoscience/open_data/Mobil_Avo_Viking_Graben_Line_12/mobil_avo.html

[3] S. Spitz, "Seismic trace interpolation in the F-X domain," *Geophysics*, vol. 56, no. 6, pp. 785–794, Jun. 1991. [Online]. Available: https://doi.org/10.1190/1.1443096

[4] R. Abma and N. Kabir, "3D interpolation of irregular data with a POCS algorithm," *Geophysics*, vol. 71, no. 6, pp. E91–E97, Nov. 2006. [Online]. Available: https://doi.org/10.1190/1.2356088

[5] S. Mandelli, V. Lipari, P. Bestagini, and S. Tubaro, "Interpolation and denoising of seismic data using convolutional neural networks," arXiv:1901.07927, 2019. [Online]. Available: https://arxiv.org/abs/1901.07927

[6] W. Fang, L. Fu, M. Zhang, and Z. Li, "Seismic data interpolation based on U-Net with texture loss," *Geophysics*, vol. 86, no. 1, pp. V41–V54, 2021. [Online]. Available: https://doi.org/10.1190/geo2019-0615.1

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241. [Online]. Available: https://doi.org/10.1007/978-3-319-24574-4_28

[8] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "SwinIR: Image restoration using Swin Transformer," in *Proc. IEEE/CVF Int. Conf. Computer Vision Workshops (ICCVW)*, 2021, pp. 1833–1844. [Online]. Available: https://doi.org/10.1109/ICCVW54120.2021.00210

[9] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 5718–5729. [Online]. Available: https://doi.org/10.1109/CVPR52688.2022.00564

[10] Z. Liu, H. Hu, Y. Lin, Z. Yao, Z. Xie, Y. Wei, J. Ning, Y. Cao, Z. Zhang, L. Dong, F. Wei, and B. Guo, "Swin transformer V2: Scaling up capacity and resolution," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 12 009–12 019. [Online]. Available: https://doi.org/10.1109/CVPR52688.2022.01170

[11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 2009, pp. 248–255.

[12] H. Sheng, X. Wu, X. Si, J. Li, S. Zhang, and X. Duan, "Seismic foundation model (SFM): a next generation deep-learning model in geophysics," *Geophysics*, vol. 90, no. 2, pp. IM59–IM79, 2025. [Online]. Available: https://doi.org/10.1190/geo2024-0262.1

[13] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "LoRA: Low-rank adaptation of large language models," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2022. [Online]. Available: https://openreview.net/forum?id=nZeVKeeFYf9

[14] Z. Guo, X. Wu, L. Liang, H. Sheng, N. Chen, and Z. Bi, "Cross-domain foundation model adaptation: Pioneering computer vision models for geophysical data analysis," arXiv:2408.12396, 2024. [Online]. Available: https://arxiv.org/abs/2408.12396

[15] F. Fuchs, M. R. Fernandez, N. Ettrich, and J. Keuper, "Foundation models for seismic data processing: An extensive review," arXiv:2503.24166, 2025. [Online]. Available: https://arxiv.org/abs/2503.24166

[16] N. Karasiak, J.-F. Dejoux, C. Monteil, and D. Sheeren, "Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing," *Machine Learning*, vol. 111, no. 7, pp. 2715–2740, 2022. [Online]. Available: https://doi.org/10.1007/s10994-021-05972-1

[17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004. [Online]. Available: https://doi.org/10.1109/TIP.2003.819861

[18] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2019. [Online]. Available: https://openreview.net/forum?id=Bkg6RiCqY7

[19] P. Micikevicius, S. Narang, J. Alben, G. Diamos, E. Elsen, D. Garcia, B. Ginsburg, M. Houston, O. Kuchaiev, G. Venkatesh, and H. Wu, "Mixed precision training," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2018. [Online]. Available: https://openreview.net/forum?id=r1gs9JgRZ

[20] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 298–311, Feb. 1997. [Online]. Available: https://doi.org/10.1109/83.551699

[21] Delft AI Cluster (DAIC), "The Delft AI Cluster (DAIC), RRID:SCR_025091," 2024. [Online]. Available: https://daic.tudelft.nl/