Submission code: 8NZ468WAN7

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Workflow for timelapse seismic data interpolation with machine learning

Edwin Fagua Duarte (UFRN), Daniel Pinheiro (Universidade Federal do Rio Grande do Norte), João Medeiros Araújo (Universidade Federal do Rio Grande do Norte), Danyelle Da Silva (UFRN), Gilberto Corso (Universidade Federal do Rio Grande do Norte), Tiago Barros (Universidade Federal do Rio Grande do Norte)

Workflow for timelapse seismic data interpolation with machine learning

Abstract Summary

We developed a seismic data interpolation workflow, which includes data preprocessing, network training, and measuring the quality of the interpolated data. The goal is to train the network in such a way that it can interpolate Monitor datasets, which could be acquired with a sparse geometry. We used a realistic synthetic 2D Ocean-Bottom Nodes (OBN) dataset which simulates non-repeatability effects. We found out that the network was capable of interpolating 50% and 75% of the sources, maintaining the timelapse information.

Introduction

Seismic data interpolation using machine learning techniques is a widely researched topic. Methodologies using many architectures and frameworks have been proposed, such as U-Net (Chai et al., 2021), GAN (Collazos et al., 2024), and diffusion models (Wang et al., 2024), for example.

Most works focus on the comparison between the interpolated and the original seismic data, aiming to minimize their difference by optimizing metrics like RMS and SSIM, and do not deal with timelapse applications, where the interpolation errors become noises that may be larger than the timelapse signals of interest.

We developed an interpolation workflow of timelapse data which includes data preprocessing, network training, and measuring the quality of the interpolated data.

Methodology

We interpolate the seismic data with a U-Net with input and output sizes of 256×256. We normalize the seismic data to the interval $[-1, 1]$ and split the seismograms into windows using the strategy presented by Zhu et al. (2023). The training process is done by decimating and interpolating the decimated traces, using a dense Baseline dataset. The goal is to train the network in such a way that it can interpolate source traces from Monitor datasets, which could be acquired with larger shot spacing, accelerating the acquisition. We also use the hyperparameter search strategy proposed by Pinheiro et al. (2024) to tune the U-Net hyperparameters.

We use a realistic synthetic 2D OBN dataset which simulates a Baseline and a Monitor acquisition with four non-repeatability effects presented by Duarte et al. (2025). This dataset is based on a Brazilian pre-salt model that was obtained by simulating 26 receiver gathers, spaced by about 375 m, located on the water bottom, and 353 sources spaced by about 50 m, using a Ricker wavelet with a peak frequency of 14 Hz. In Figure 1, we present the Baseline and Monitor P-wave velocity models, as well as the difference between them. Note that the velocities range between around 1,500 and 5,500 m/s, while the maximum difference between the velocity models is around 90 m/s.

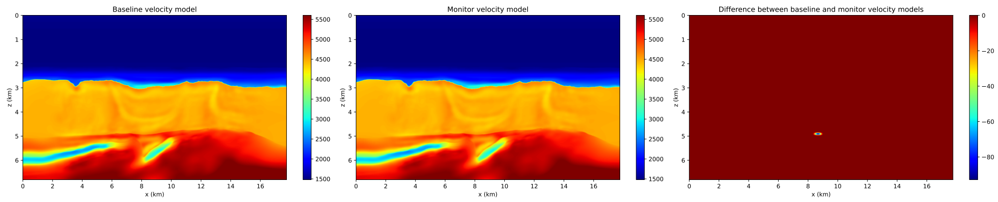

Figure 1: Baseline velocity model, Monitor velocity model, and the difference between the Baseline and Monitor models.

In order to measure the quality of interpolation, we perform the migration of the seismic data with the Reverse Time Migration with Inverse Scattering Image Condition (Kiyashchenko et al., 2007). We then visually compare the migrated images and compute the following metrics:

$$NRMS = 200 \times RMS(Monitor - Baseline)/(RMS(Baseline) + RMS(Monitor)), \quad (1)$$

$$NdRMS = 2 \times (RMS(Monitor) - RMS(Baseline))/(RMS(Baseline) + RMS(Monitor)), \quad (2)$$

$$NRMS\_Diff = 2 \times RMS(Monitor - Baseline)/(RMS(Baseline) + RMS(Monitor)). \quad (3)$$

NRMS is calculated over a section of the migrated images, while NdRMS and NRMS_Diff are calculated for each trace under a given region.

Experiments

We trained the U-Net using the Baseline dataset by consistently decimating and interpolating source traces. After that, we used the trained network to infer the decimated traces of the Monitor dataset. We selected three different regular decimation ratios: 50%, 75%, and 87.5%.

In Figure 2, we present the original seismograms, decimated seismograms, interpolated seismograms, and the difference between the original and the interpolated seismograms. We call attention to the fact that the original and interpolated seismograms are very similar, but, as the decimation ratio increases, the metrics MAE and RMS increase and the metric SSIM decreases. It is difficult to assert whether the interpolation has good quality or not, based only on the values of these metrics. It is even harder to estimate the quality of the interpolation visually, as it is necessary to apply a color mapping gain so that we can observe the interpolation errors more clearly.

In order to estimate the quality of interpolation, we performed migrations with the Baseline dataset and the Monitor dataset. In Figure 3, we present the comparisons between migrated images using the Monitor and Baseline datasets. We present results using the complete Monitor dataset, and the Monitor with 50%, 75%, and 87.5% interpolated traces. Looking at the difference between migrations with the complete Monitor dataset, we note that only the small region around $(x, z) = (8.7, 4.9)$ is due to the timelapse change in the reservoir, and everything else is due to the noise generated by the non-repeatability effects. We observe that the results with the decimation ratios of 50% and 75% are visually similar to the ones with the complete Monitor, and they also have similar values of NdRMS and NRMS_Diff around the reservoir, which indicates a satisfying interpolation for these cases. However, we observe that the timelapse signal around $(x, z) = (8.7, 4.9)$, due to the change in the reservoir, is significantly lost when we interpolate 87.5% of the traces. The NdRMS and NRMS_Diff values around $x = 8.7$ are also greatly reduced with this level of interpolation.

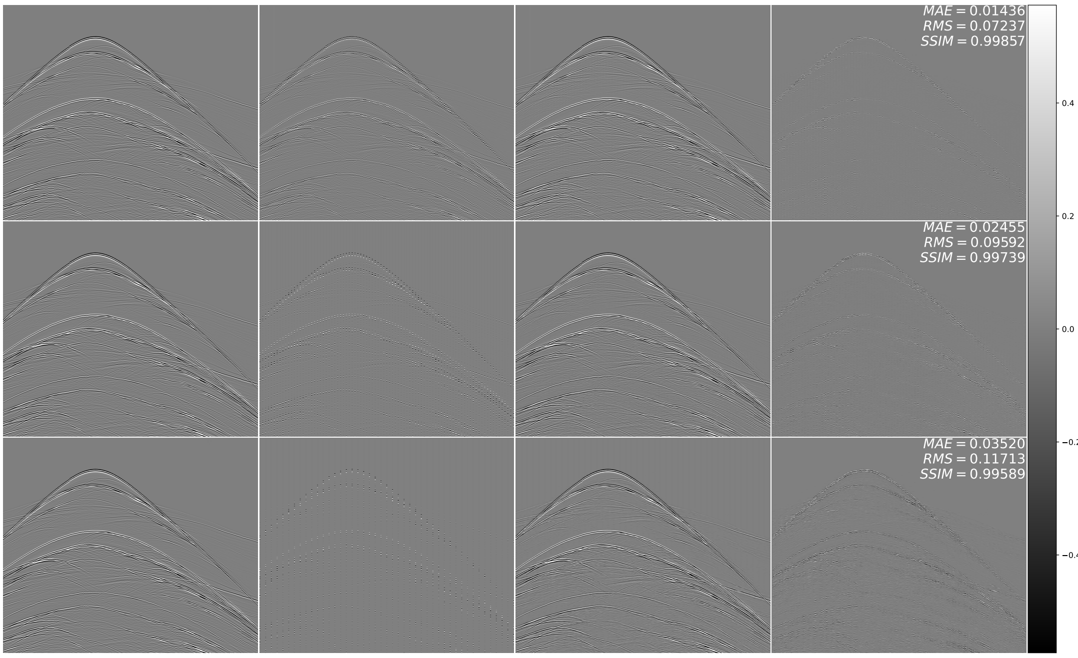

Figure 2: Seismic data interpolation. The four columns of figures refer to the original, decimated, and interpolated seismograms, and the difference between the original and the interpolated seismograms, respectively. The three rows refer to the decimation ratios of 50%, 75%, and 87.5%, indicating 100 m, 200 m, and 400 m between traces, respectively. The values of the metrics MAE, RMS and SSIM are calculated between the original and interpolated seismograms. A color mapping gain of 25x was applied for better visualization.

Conclusions

We developed an interpolation workflow for data preprocessing, network training, and measuring the quality of the interpolated data. The network was trained with a Baseline dataset, and it was used to infer traces from a Monitor dataset. We measured the quality of the interpolation by performing migrations with the interpolated data. We found out that the network was capable of interpolating 50% and 75% of the traces, maintaining the timelapse information.

Acknowledgments

We gratefully acknowledge support from Shell Brasil Petróleo Ltda. through the “New computationally scalable methodologies for target-oriented 4D seismic in pre-salt reservoirs” project (ANP 22681-1) and the strategic importance of the support given by the Brazilian Petroleum Agency (ANP) through the R&D levy regulation. We acknowledge NPAD/UFRN for the computational support.

References

Chai, X., G. Tang, S. Wang, K. Lin, and R. Peng, 2021, Deep learning for irregularly and regularly missing 3-D data reconstruction: IEEE Transactions on Geoscience and Remote Sensing, **59**,

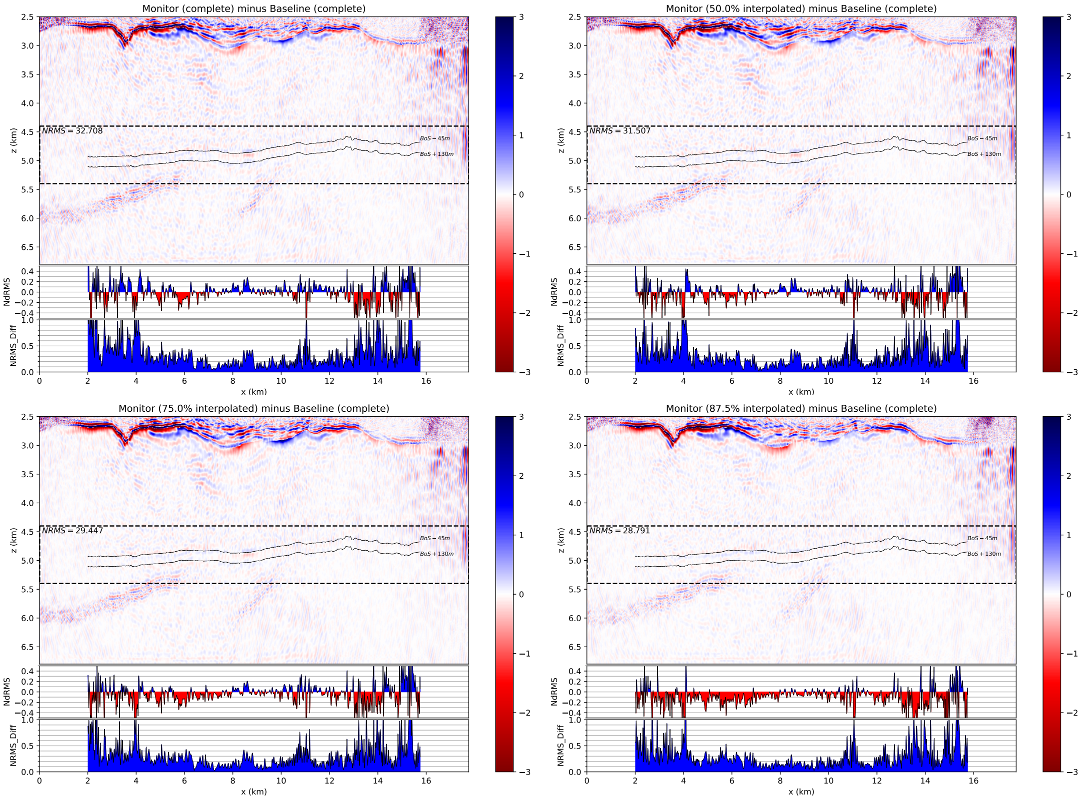

Figure 3: Comparisons between the migrated images using the Monitor and Baseline datasets. The NRMS values were calculated using the window delimited by the dashed lines. The NdRMS and NRMS_Diff values were calculated for each trace within the horizons $BoS - 45m$ and $BoS + 130m$.

6244–6265.

Collazos, J. A., K. D. J. Rincon, D. N. Pinheiro, M. G. Gebre, C. A. N. da Costa, G. Corso, T. Barros, J. M. de Araújo, and Y. Wang, 2024, Seismic trace interpolation based on the principle of reciprocity using a conditional generative adversarial network (cGAN): Journal of Geophysics and Engineering, **21**, 1775–1790.

Duarte, E. F., D. da Silva, J. Schwarzrock, R. C. F. Araujo, G. Corso, J. M. de Araujo, J. L. Lopez, and D. Kiyashchenko, 2025, Realistic 4D OBN seismic data simulation for brazilian pre-salt settings: Presented at the International Meeting for Applied Geoscience & Energy 2025. (In press).

Kiyashchenko, D., R. Plessix, B. Kashtan, and V. Troyan, 2007, A modified imaging principle for true-amplitude wave-equation migration: Geophysical Journal International, **168**, 1093–1104.

Pinheiro, D. N., J. C. Gonzalez, G. Corso, M. G. Gebre, C. A. da Costa, S. X. de Souza, and T. Barros, 2024, Hyperparameter determination for GAN-based seismic interpolator with variable neighborhood search: Computers & Geosciences, **192**, 105689.

Wang, S., F. Deng, P. Jiang, Z. Gong, X. Wei, and Y. Wang, 2024, Seisfusion: Constrained diffusion model with input guidance for 3-D seismic data interpolation and reconstruction: IEEE Transactions on Geoscience and Remote Sensing, **62**, 1–15.

Zhu, Y., J. Cao, H. Yin, J. Zhao, and K. Gao, 2023, Seismic data reconstruction based on Attention U-net and transfer learning: Journal of Applied Geophysics, **219**, 105241.