Monitoring CO_2 in Seismic Data Using Neural Network

Elena Gondyul(✉), Vadim Lisitsa, and Dmitry Vishnevsky

Sobolev Institute of Mathematics SB RAS, 4 Acad. Koptug Avenue, 630090 Novosibirsk, Russia
gondyulea@ipgg.sbras.ru,gondyulea@gmail.com

Abstract. Accurate monitoring of CO_2 migration in subsurface reservoirs is critical for understanding the behavior of injected greenhouse gases. This study proposes a neural network-based approach to improve the accuracy of seismograms used in time-lapse seismic monitoring. The method consists of two stages: first, a neural network predicts changes in seismograms corresponding to velocity model variations between consecutive monitoring steps, allowing for the approximation of spatio-temporal dependencies and facilitating wavefield extrapolation. The seismograms at this stage are generated using a coarse computational grid to reduce computational costs.

In the second stage, a neural network is employed to mitigate numerical dispersion in the predicted seismogram differences generated via classical modeling under the assumption of an unchanged velocity model. The trained network is then applied to all seismograms obtained in the first stage. This approach enables a more precise estimation of CO_2 migration patterns, providing valuable insights into subsurface dynamics.

The proposed approach significantly accelerates seismic modeling and its application to monitoring greenhouse gases in reservoir rocks.

Keywords: Deep Learning · Monitoring CO_2 · Seismic modeling

1 Introduction

The capture and flooding of carbon dioxide (CO_2) and other greenhouse gases in geological formations is a crucial strategy to mitigate the effects of human impacts on the environment. To ensure contentment of the injection process, it is essential to select the injection point and rate, assess the long-term storage stability, and prevent leakage. This requires the use of numerical methods for simulation of the long-term evolution of CO_2, as well as seismic modeling for monitoring through changed seismograms. Time-lapse seismic surveys and their simulation provide valuable information about the spatio-temporal evolution of the CO_2 plume. However, numerical dispersion, an artifact arising from the discretization of wave equations in computational simulation often degrades the accuracy of synthetic seismograms used for interpretation and forecasting.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
O. Gervasi et al. (Eds.): ICCSA 2025 Workshops, LNCS 15888, pp. 330–342, 2026.
https://doi.org/10.1007/978-3-031-97596-7_22

Traditional numerical methods, such as finite-difference and finite-element approaches, attempt to mitigate dispersion errors through refined discretization and higher-order schemes. However, these solutions are computationally expensive, especially when modeling large-scale reservoirs with complex geological structures. As an alternative, deep learning-based methods such as neural network NDM-net (Numerical Dispersion Mitigation neural network) have demonstrated the ability to correct numerical dispersion [4,6]. Furthermore, the neural network NDM-net has demonstrated good performance in correcting errors during velocity model disturbances due to changes in its discretization.

In this study, it is suggested to employ a framework utilizing seismic modeling and the neural network NDM-net based on pix2pix to suppress numerical dispersion in seismograms and to predict seismograms at the next stage of monitoring. The aim is to accelerate monitoring CO_2 emission by the classical calculation of seismograms using a coarse computational grid for the total amount of sources in an unmodified velocity model and part of seismograms using a fine computational grid to form a training dataset. The proposed approach offers a computationally efficient alternative to traditional direct modeling methods, which makes real-time seismic monitoring more feasible.

1.1 Related Works

The most effective method for monitoring the condition of a reservoir during the process of fluid substitution, such as CO2 injection, is through the use of seismic monitoring techniques [17,18,24]. At the same time, numerical modeling is essential to understanding and analyzing the manifestations of variations in the petrophysical properties of reservoirs in seismic areas. Furthermore, seismic modeling allows the testing of algorithms for processing seismic data and assessing their suitability for mapping such changes.

Classical methods, such as finite difference [12,14,16,20,23], finite elements [8], discontinuous Galerkin method [2,15], spectral elements [11,22] and others require a significant amount of computational resources and time, particularly when it comes to calculating multiple sets of sources for different velocity models. But it is the finite difference and their variants that are most commonly used in seismic exploration and seismic monitoring due to their high efficiency, coupled with acceptable accuracy. However, it still remains a cost in the task of monitoring greenhouse gases.

Neural networks have recently begun to be applied in the field of hydrocarbon monitoring, for example, through the use of the Fourier neural operator [25] as a proxy for the fluid-flow simulator [26]. They are also used in the detection of hydrocarbon plumes from seismic images [13,21]. In real conditions, it is suggested to use alternative techniques such as gravitational sensing in combination with classical methods like surface seismic monitoring and neural networks [1].

2 Framework

Changes in seismic data during injection of CO_2 demonstrate how reservoir properties change [10]. In this study, we assume that reservoir properties change relatively slowly over time and do not take into account the physics of multiphase fluid flow, i.e. density ρ and velocities V_p, V_s change on average over the entire layer, as shown in Fig. 1. The velocity model and its properties are denoted as $M^i = \{V_p^i, V_s^i, \rho^i\}$, where $i \in [1, N]$ is the index signifying the stage of the greenhouse injection and N is the final stage of monitoring.

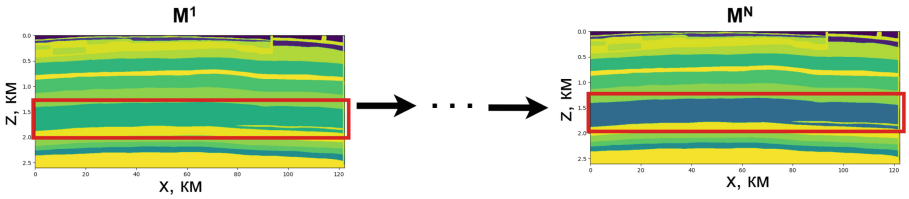

Fig. 1. Changing the properties of one of the layers of the Vanavar velocity model.

The aim of this research is to obtain the evolution of seismograms in the most cost-effective manner possible, therefore, in most cases, it is proposed to use a coarse computational grid to generate a training sample using the finite difference method. As shown in the schematic diagram (Fig. 2), our proposed methodology consists of the two-stage process using a neural network NDM-net (Numerical Dispersion Mitigation neural network) in order to enhance the efficiency of seismic monitoring for CO_2 migration in an oil reservoir.

The first stage involves using a neural network to generate changes in seismograms corresponding to the modified velocity model, that is, the one that was changed at a subsequent monitoring stage and the velocity model at the previous monitoring stage. This allows the network to approximate the spatio-temporal dependence, facilitating the extrapolation of wave field evolution. The seismograms used at this stage are modeled using coarce computational grid to reduce costs.

At the second stage, a neural network is used to reduce the numerical dispersion in the changes in the seismograms generated by the classical method, assuming the use of an unchanged velocity model M^1 in which migration has not yet occurred and velocity model M^2. And then the trained neural network is applied to all generated seismograms in the first stage. This process enables effective estimation of CO_2 migration, providing valuable insights into the behavior of greenhouse gas movement.

The entire proposed algorithm contains the following steps:

1. Classical seismic modeling is carried out for the unmodified velocity model M^1 for the entire set of sources using a coarse computational grid with characteristic step h_1, in order to obtain seismograms $\boldsymbol{u}_{h_1}^1(t, x_o, x_s^k), k \in [1, N_s]$, where N_s is the total number of sources.

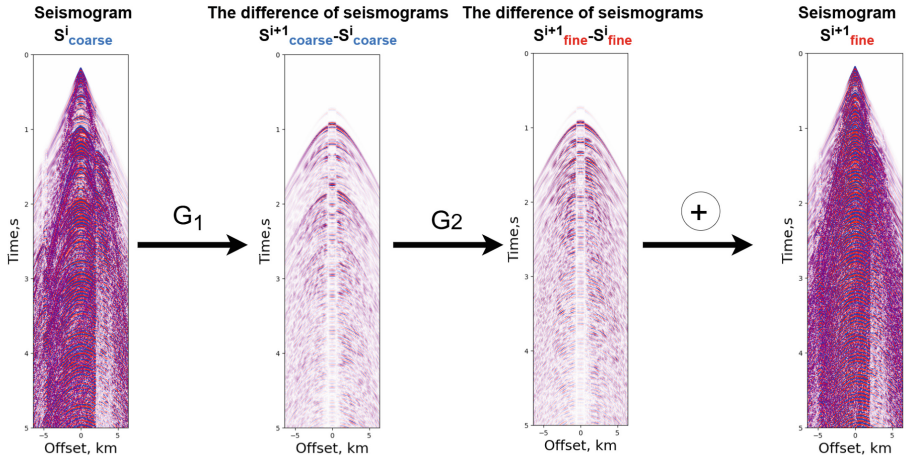

Fig. 2. Framework for monitoring CO_2 during changes in seismograms using neural networks approximating operators G_1 and G_2.

2. A set of indexes of the source positions $D_t \subset \{1, N_s\}$ is generated in order to create a training set.
3. Classical seismic modeling is carried out for the modified models $M^i, i \in [2, N]$ using a coarse computational grid with characteristic step h_1, in order to obtain $\boldsymbol{u}^i_{h_1}(t, x_o, x^k_s), k \in D_t$.
4. The difference in seismograms between current monitoring step and the next step $d\boldsymbol{u}^i_{h_1}(t, x_o, x^k_s) = \boldsymbol{u}^{i+1}_{h_1}(t, x_o, x^k_s) - \boldsymbol{u}^i_{h_1}(t, x_o, x^k_s), k \in D_t$ are calculated to create training set.
5. After that, numerical simulation of wave fields is carried out to generate seismograms $\boldsymbol{u}^1_{h_2}(t, x_o, x^k_s)$ and $\boldsymbol{u}^2_{h_2}(t, x_o, x^k_s), k \in D_t$ using a fine computational grid with characteristic step $h_2 < h_1$ for a limited number of sources using velocity models M^1 and M^2.
6. The difference of the refined seismograms for the first and second step of greenhouse gas injection $d\boldsymbol{u}^i_{h_2}(t, x_o, x^k_s) = \boldsymbol{u}^{i+1}_{h_2}(t, x_o, x^k_s) - \boldsymbol{u}^i_{h_2}(t, x_o, x^k_s), k \in D_t$ are calculated.
7. Training of the NDM-net for operator approximation:

$$G_1 : \boldsymbol{u}^i_{h_1}(t, x_o, x^k_s) \rightarrow d\boldsymbol{u}^i_{h_1}(t, x_o, x^k_s).$$

8. Training NDM-net to approximate the following operator:

$$G_2 : \hat{d}\boldsymbol{u}^i_{h_1}(t, x_o, x^k_s) \rightarrow d\boldsymbol{u}^i_{h_2}(t, x_o, x^k_s),$$

where $\hat{d}\boldsymbol{u}^i_{h_1}(t, x_o, x^k_s) = G_1(\boldsymbol{u}^i_{h_1}(t, x_o, x^k_s))$.
9. Applying sequentially to all seismograms corresponding to velocity model M^{i+1} to obtain seismograms:

$$\hat{\boldsymbol{u}}^{i+1}_{h_2}(t, x_o, x^k_s) = \hat{\boldsymbol{u}}^i_{h_2}(t, x_o, x^k_s) + \hat{d}\boldsymbol{u}^i_{h_2}(t, x_o, x^k_s).$$

2.1 Neural Network

To correct the numerical solution and reduce the numerical dispersion of the recorded signal as well as predicted seismograms for the next step of the monitoring CO_2 the Numerical Dispersion Mitigation neural network (NDM-net) is used [5]. Initially, a neural network was used for training to convert coarse grid data into a more accurate one. In particular, this neural network reduced the numerical dispersion in the seismograms. Later, the same neural network demonstrated its effectiveness in reducing errors caused by rough velocity model sampling, i.e. when it is perturbed [7]. Thus, under the condition of a small change in the properties of the medium and the corresponding velocity model, the NDM-net neural network can transfer seismograms from the current monitoring step to the next.

We used pix2pix architecture [9] which consists of two sub-neural networks: generator and discriminator. Generator is a slightly modified U-net [19], which contains an encoder-decoder with interconnecting skip connections. The encoder converts the input data into a representation with a smaller dimension in hidden layers using convolution operations. The decoder is structurally identical to the encoder and recovers data back using reverse convolution operators. Downsampling and upsampling layers include the convolutional layer with kernel size $(4 \times 4 \times 4)$, activation function $ReLU(\cdot) = \max(0, \cdot)$ for decoder and $LeakyReLU(\cdot) = \max(0, \cdot) + c\min(0, \cdot)$ with negative slope coefficient $c = 0.2$ for encoder. Discriminator if a simple classifier based in convolution operations.

A mini-batch stochastic gradient descent (SGD) is used to speed up the algorithm. To estimate the number of epochs, the method of early stopping is used, checking the loss function on the validation dataset, which has a size of 10% of the training dataset.

A neural network for the approximation of the operator G_1 was trained by minimizing the loss function in order to select parameters Θ_1:

$$\begin{aligned}\mathcal{L}_1(\Theta_1) = arg\min_{G_1} max_{D_1} E_{k\in D_t}[\log D_1(G_1(\boldsymbol{u}_{h_2}(t, x_o, x_s^k), \Theta_1)) \\ + log(1 - D_1(d\boldsymbol{u}_{h_1}^i(t, x_o, x_s^k))) \\ + \lambda(\left\| d\boldsymbol{u}_{h_1}^i(t, x_o, x_s^k) - G_1(\boldsymbol{u}_{h_2}(t, x_o, x_s^k), \Theta_1)\right\|_1)],\end{aligned} \tag{1}$$

where $\Theta_1 = \{\mathbf{W_1}, \mathbf{b_1}\}$ is the network parameters for G_1, which includes the weight matrix $\mathbf{W_1}$ and bias $\mathbf{b_1}$, $\mathbb{E}_{k\in D_t}$ is a mean over the training dataset, $||\cdot||_1$ is the L_1- norm.

A neural network for the approximation of the operator G_2 was alse trained by minimizing the loss function in order to select parameters Θ_2:

$$\begin{aligned}\mathcal{L}_2(\Theta_2) = arg\min_{G_2} max_{D_2} E_{k\in D_t}[\log D_2(G_2(\hat{d}\boldsymbol{u}_{h_1}^i(t, x_o, x_s^k), \Theta_2)) \\ + log(1 - D_1(d\boldsymbol{u}_{h_2}^i(t, x_o, x_s^k))) \\ + \lambda(\left\| d\boldsymbol{u}_{h_2}^i(t, x_o, x_s^k) - G_1(\hat{d}\boldsymbol{u}_{h_1}^i(t, x_o, x_s^k), \Theta_2)\right\|_1)],\end{aligned} \tag{2}$$

where $\Theta_2 = \{\mathbf{W_2}, \mathbf{b_2}\}$ is the network parameters for G_2, which includes the weight matrix $\mathbf{W_2}$ and bias $\mathbf{b_2}$.

3 Numerical Experiments

We used Vanavar velocity model, which corresponds to an Eastern Siberian region to calculate seismograms. It is a multi-layered structure with prominent vertical intrusions. The model spans 220 km in length and 2.6 km in width. At the same time, we only used the first half of the original model prior to the intrusion as shown in Fig. 1. Therefore, the size of the modified model is 110 km by 2.6 km. The acquisition system comprises 1222 sources spaced 100 m apart. The wavefield is recorded by 512 receivers for each shot, with a maximum source-receiver separation of 6.4 km. The receivers are spaced 25 m apart. The source is a Ricker wavelet with a central frequency of 30 Hz. For seismic monitoring the properties of the medium, such as density ρ and velocities V_p and V_s, change gradually. Specifically, they change by 3% at each step in the range $[1, N]$, where $N = 4$.

To create the training set, seismic simulation is performed, which involves solving a system of equations from the dynamic theory of elasticity using the finite difference method on staggered grids. This has been discussed previously in the publications [3,5,6]. We calculated two sets of seismograms using a coarse grid with a characteristic pitch of 5 m and a fine grid with a characteristic pitch of 2.5 m. The recording time of the seismogram was 4 s with a time discreteness of 1 ms.

The sample of seismograms used for training is created so that the size of the sample for approximating the operator G_1 is 30% of the total number of seismograms at each step in the seismic monitoring process, i.e., sources multiplied by the number of monitoring steps, and the size of the sample used for approximating operator G_2 is 30% of the number of source locations. After that, the data is normalized for stable operation of the neural network, i.e. the values of the seismograms are converted to the range $[-1, 1]$. The training parameters are presented in Table 1.

Table 1. Parameters of the neural networks

Number of epochs	100	Batch size	5
Learning rate	1e-4	Momentum parameters	$\beta_1 = 0.5, \beta_2 = 0.999$
Input data dimension	512×4096	Coefficient λ	10000
Output data dimension	512×4096		

3.1 Results and Discussion

One of the key stages of the proposed algorithm is the training of a neural network to approximate operator G_1 and discriminator D_1. Learning and validation

curves (Fig. 3A) indicate that the neural network has learned well. Similar learning and validation patterns are observed when training a network that approximates operator G_2 and discriminator D_2 (Fig. 3B). At the same time, training time on an NVIDIA GeForce RTX 3090 with 8 GB memory and Intel Core CPU with 256 GB RAM card was approximately 300 and 100 min for the first and second scenarios, respectively.

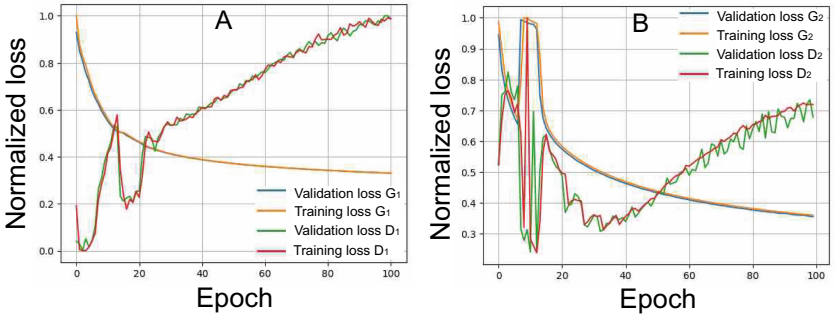

Fig. 3. The normalized training and validation losses for neural networks G_1, D_1 (A) and neural networks G_2, D_2 (B).

Examples of results obtained using two operators, i.e. generated seismograms $\hat{\boldsymbol{u}}^2_{h_2}(t, x_o, x^k_s) = \hat{\boldsymbol{u}}^1_{h_2}(t, x_o, x^k_s) + G_2(G_1(\boldsymbol{u}^1_{h_1}(t, x_o, x^k_s))$ corresponding to the second velocity model M^2 (Figures 4A, 4B, 4C), demonstrate that they visually resemble the numerical solution to the problem of seismic wave propagation on a fine computational grid $\boldsymbol{u}^2_{h_2}(t, x_o, x^k_s)$ (Figures 4D, 4E, 4F).

In order to evaluate the quality of neural networks, we used the error in terms of the L_2-norm:

$$\varepsilon(\hat{\boldsymbol{u}}^{i+1}_{h_2}(t, x_o, x^k_s), \boldsymbol{u}^{i+1}_{h_2}(t, x_o, x^k_s)) = \frac{||\hat{\boldsymbol{u}}^{i+1}_{h_2}(t, x_o, x^k_s) - \boldsymbol{u}^{i+1}_{h_2}(t, x_o, x^k_s)||_2}{||\boldsymbol{u}^{i+1}_{h_2}(t, x_o, x^k_s)||_2} \times 100\%,$$

where $\hat{\boldsymbol{u}}^{i+1}_{h_2}(t, x_o, x^k_s) = \hat{\boldsymbol{u}}^{i}_{h_2}(t, x_o, x^k_s) + G_2(G_1(\boldsymbol{u}^{i}_{h_1}(t, x_o, x^k_s))$. Using this formula, we can also calculate the initial discrepancy between the solutions on the fine and coarse computational grids in a similar manner.

The overall trend in the error ε and initial error ε_{init} for each source (Fig. 5) indicates that numerical dispersion is reduced average by a factor of 4 (from 114.6% to 27.6% in terms of L_2-norm with standard deviation 5.6% and 10.6%, respectively). However, some seismograms have recovered relatively poorly after using neural networks. This may be because the training set is equally distributed and does not include a wide variety of seismograms. This problem was discussed in [6].

For the velocity model M^4 after sequential application of neural networks, we also constructed a relative error similar to the velocity model M^2 (Fig. 6.

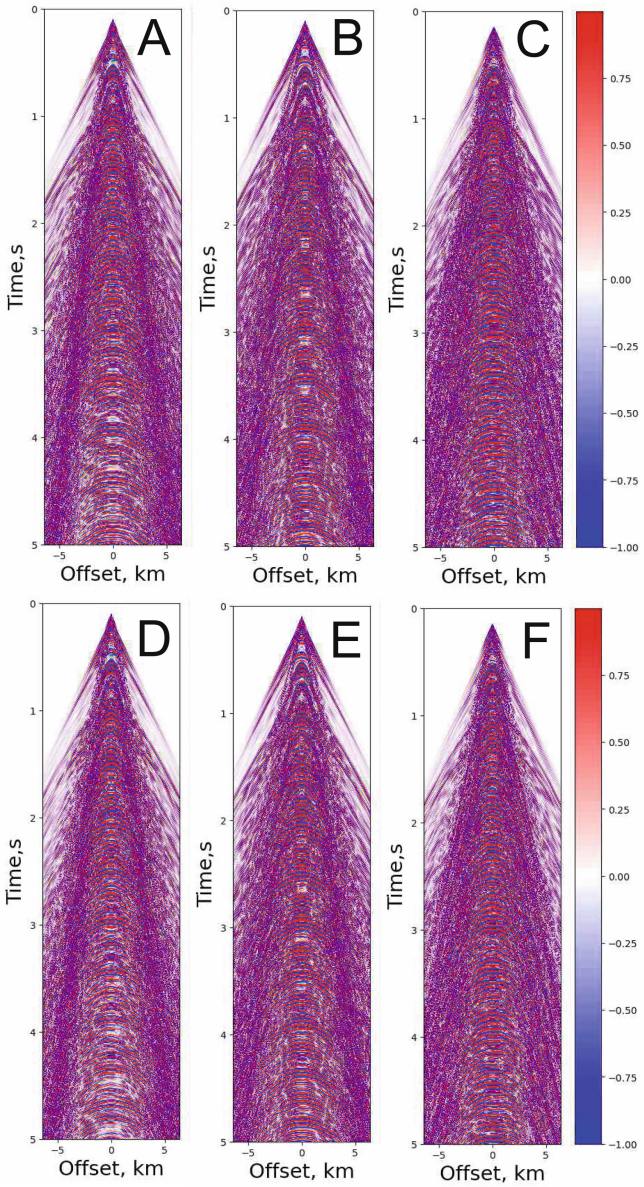

Fig. 4. Generated seismograms $\hat{\boldsymbol{u}}^2_{h_2}(t, x_o, x_s^k)$ (A,B,C) for the random source positions and corresponding seismograms calculated on a fine grid $\boldsymbol{u}^2_{h_2}(t, x_o, x_s^k)$ (D,E,F).

However, in this case, there is a significant drop in the quality of the results of using neural networks. This is due to the fact that when neural networks are consistently applied to the seismograms of each velocity model, an error accumu-

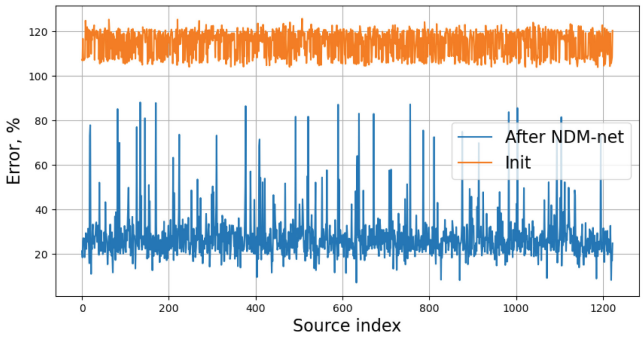

Fig. 5. The relative error between seismogram calculated on a fine grid and coarse grid solution (orange line) and the relative error between neural network post-processed seismogram and fine solution (blue line) for each source position and velocity model M^2. (Color figure online)

lates. At the same time, the numerical dispersion, in this case, is suppressed on average in L_2-norm from 134.8% to 27.4% with a standard deviation of 39.6% and 13.4%, respectively.

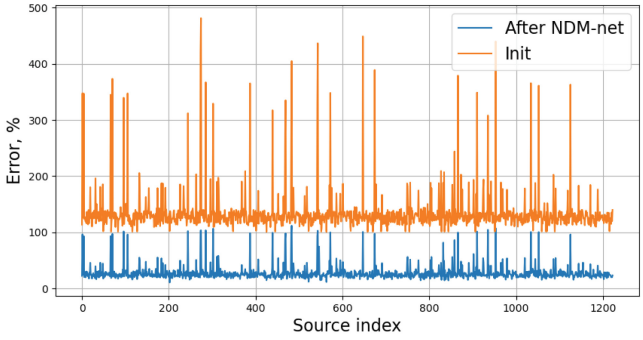

Fig. 6. The relative error between seismogram calculated on a fine grid and coarse grid solution (orange line) and the relative error between neural network post-processed seismogram and fine solution (blue line) for each source position and velocity model M^4. (Color figure online)

Random generated seismograms after sequential application of neural networks for the velocity model M^4 are shown in Fig. 7. Subtle noise appears on the generated seismograms, due to the reusable sequential use of neural networks.

Therefore, the proposed approach allows for a significant reduction in the computational time required for solving the direct seismic problem. Specifically, it will take approximately 29352 min to compute all seismograms for both the

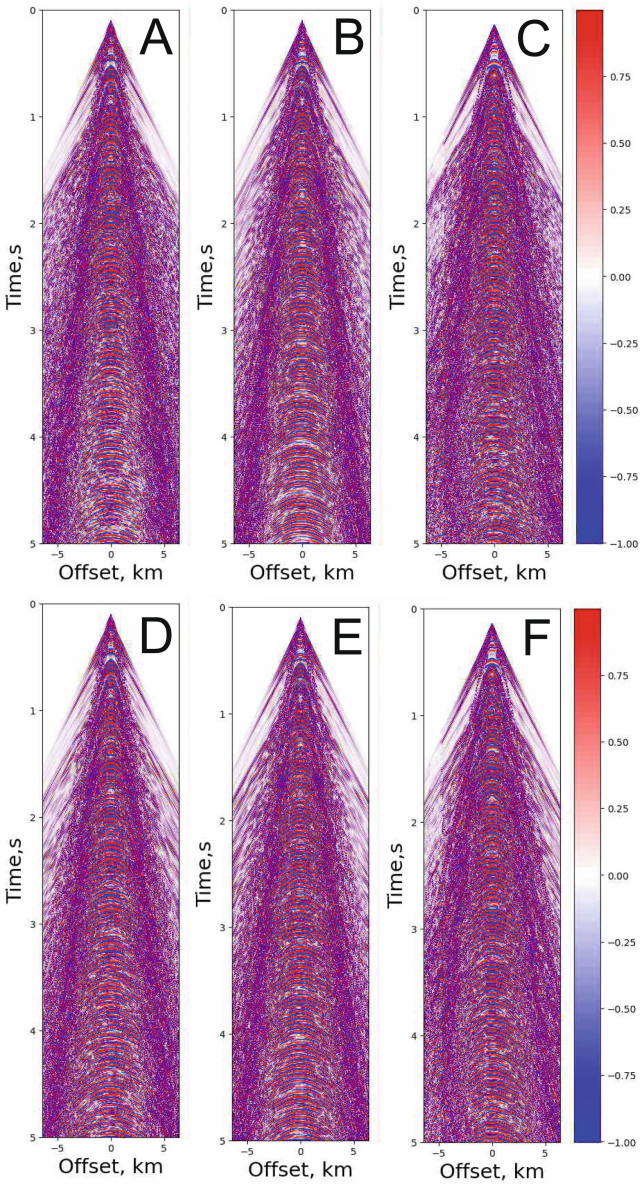

Fig. 7. Generated seismograms $\hat{\boldsymbol{u}}^4_{h_2}(t, x_o, x_s^k)$ (A,B,C) for the random source positions and corresponding seismograms calculated on a fine grid $\boldsymbol{u}^4_{h_2}(t, x_o, x_s^k)$ (D,E,F).

modified and unmodified velocity models. In contrast, the training and implementation of a neural network would take approximately 4926 min, resulting in a 6x acceleration of the overall process.

4 Conclusion

This study presents a neural network-based approach to improve the accuracy of seismograms when monitoring a CO_2 reservoir. The proposed technique includes two-stage neural network training, which makes it possible to predict changes in the seismogram due to changes in the velocity model and refine numerical solutions obtained on a coarse computational grid. The results of numerical experiments confirm that the neural network copes well with the test data, while the training time at the first stage and the second, along with the formation of a training sample, is 6 times less than in the case of seismic modeling of a complete set of seismograms for modified velocity models using a fine computational grid. Quantification using the normal error indicator L_2 demonstrates a significant decrease in numerical variance, while average errors decrease fourfold. These results indicate that the proposed method effectively improves the accuracy of frame-by-frame seismic monitoring while maintaining computational efficiency.

Future work may focus on extending this approach to three-dimensional seismic data, incorporating multiphase flow physics to improve generalization, and exploring alternative neural network architectures to further reduce numerical distortion.

Acknowledgments. This study was funded by Russian Science Foundation grant no. 22-11-00004П.

References

1. Celaya, A., Araya-Polo, M.: Joint inversion of time-lapse surface gravity and seismic data for monitoring of 3d co_2 plumes via deep learning (2023). https://arxiv.org/abs/2310.04430
2. Dumbser, M., Kaser, M.: An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - ii. the three-dimensional isotropic case. Geophys. J. Int. **167**(1), 319–336 (2006). https://doi.org/10.1111/j.1365-246X.2006.03120.x
3. Gadylshin, K., Lisitsa, V., Gadylshina, K., Vishnevsky, D.: Optimization of the training dataset for numerical dispersion mitigation neural network. In: Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A.C., Garau, C. (eds.) Computational Science and Its Applications – ICCSA 2022 Workshops, pp. 295–309. Springer International Publishing (2022)
4. Gadylshin, K., Lisitsa, V., Gadylshina, K., Vishnevsky, D.: Frequency domain numerical dispersion mitigation network. In: Gervasi, O., et al. (eds.) Computational Science and Its Applications – ICCSA 2023 Workshops, pp. 31–44. Springer Nature Switzerland, Cham (2023)
5. Gadylshin, K., Vishnevsky, D., Gadylshina, K., Lisitsa, V.: Numerical dispersion mitigation neural network for seismic modeling. Geophysics **87**(3), T237–T249 (2022)
6. Gondyul, E., Lisitsa, V., Gadylshin, K., Vishnevsky, D.: Numerical dispersion mitigation neural network with the model-based training dataset optimization. In: Gervasi, O., et al., (eds.) Computational Science and Its Applications – ICCSA 2023 Workshops, pp. 19–30. Springer Nature Switzerland (2023)

7. Gondyul, E., Lisitsa, V., Gadylshin, K., Vishnevsky, D.: Numerical dispersion mitigation neural network with velocity model correction. Comput. Geosci. **196**, 105806 (2025). https://doi.org/10.1016/j.cageo.2024.105806, https://www.sciencedirect.com/science/article/pii/S0098300424002899
8. Idesman, A., Pham, D.: Finite element modeling of linear elastodynamics problems with explicit time-integration methods and linear elements with the reduced dispersion error. Comput. Methods Appl. Mech. Eng. **271**, 86–108 (2014)
9. Isola, P., Zhu, J., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. CoRR **abs/1611.07004** (2016), http://arxiv.org/abs/1611.07004
10. Ivandic, M., Bergmann, P., Kummerow, J., Huang, F., Juhlin, C., Lueth, S.: Monitoring co2 saturation using time-lapse amplitude versus offset analysis of 3d seismic data from the ketzin co2 storage pilot site, Germany. Geophys. Prosp. **66**(8), 1568–1585 (2018). https://doi.org/10.1111/1365-2478.12666, https://www.earthdoc.org/content/journals/10.1111/1365-2478.12666
11. Komatitsch, D., Vilotte, J.P.: The spectral element method; an efficient tool to simulate the seismic response of 2D and 3D geological structures. Bull. Seismol. Society Am. **88**(2), 368–392 (1998)
12. Levander, A.R.: Fourth-order finite-difference P-SV seismograms. Geophysics **53**(11), 1425–1436 (1988)
13. Li, B., Li, Y.E.: Neural network-based co2 interpretation from 4d Sleipner seismic images. J. Geophys. Res.: Solid Earth **126**(12), e2021JB022524 (2021)
14. Lisitsa, V., Vishnevskiy, D.: Lebedev scheme for the numerical simulation of wave propagation in 3d anisotropic elasticity. Geophys. Prospect. **58**(4), 619–635 (2010). https://doi.org/10.1111/j.1365-2478.2009.00862.x
15. Lisitsa, V., Tcheverda, V., Botter, C.: Combination of the discontinuous Galerkin method with finite differences for simulation of seismic wave propagation. J. Comput. Phys. **311**, 142–157 (2016)
16. Lisitsa, V., Tcheverda, V., Vishnevsky, D.: Numerical simulation of seismic waves in models with anisotropic formations: coupling Virieux and Lebedev finite-difference schemes. Comput. Geosci. **16**(4), 1135–1152 (2012)
17. Paffenholz, J.: Introduction to this special section: The role of advanced modeling in enhanced carbon storage. Lead. Edge **40**(6), 408–412 (2021)
18. Pevzner, R., et al.: Feasibility of time-lapse seismic methodology for monitoring the injection of small quantities of co2 into a saline formation, co2crc otway project. Energy Proc. **37**, 4336–4343 (2013). https://doi.org/10.1016/j.egypro.2013.06.336
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). LNCS, vol. 9351, pp. 234–241. Springer (2015). http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a
20. Saenger, E.H., Gold, N., Shapiro, S.A.: Modeling the propagation of the elastic waves using a modified finite-difference grid. Wave Motion **31**, 77–92 (2000)
21. Sheng, H., Wu, X., Sun, X., Wu, L.: Deep learning for characterizing co2 migration in time-lapse seismic images. Fuel **336**, 126806 (2023)
22. Tromp, J., Komatitsch, D., Liu, Q.: Spectral-element and adjoint methods in seismology. Commun. Comput. Phys. **3**(1), 1–32 (2008)
23. Virieux, J.: P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. Geophysics **51**(4), 889–901 (1986)
24. Wang, Z., Harbert, W.P., Dilmore, R.M., Huang, L.: Modeling of time-lapse seismic monitoring using co2 leakage simulations for a model co2 storage site with realistic geology: Application in assessment of early leak-detection capabilities. Int. J.

Greenhouse Gas Control **76**, 39–52 (2018). https://doi.org/10.1016/j.ijggc.2018.06.011
25. Wen, G., Li, Z., Azizzadenesheli, K., Anandkumar, A., Benson, S.M.: U-FNO–an enhanced Fourier neural operator-based deep-learning model for multiphase flow. Adv. Water Resour. **163**, 104180 (2022)
26. Yin, Z., Siahkoohi, A., Louboutin, M., Herrmann, F.J.: Learned coupled inversion for carbon sequestration monitoring and forecasting with fourier neural operators. In: Second International Meeting for Applied Geoscience; Energy. p. 467–472. Society of Exploration Geophysicists and American Association of Petroleum Geologists (Aug 2022). https://doi.org/10.1190/image2022-3722848.1, http://dx.doi.org/10.1190/image2022-3722848.1