Applicability of NDM-Net Simulated Data to Prestack Depth Migration

Elena Gondyul[1], Valeria Dubrovina[2], Vadim Lisitsa[1(✉)], and Dmitry Vishnevsky[1]

[1] Sobolev Institute of Mathematics SB RAS, Novosibirsk 630090, Russia
e.gondyul@yandex.ru, Valeria.Dubrovina@ptgeos.com
[2] LLC PetroTrace, Moscow 115114, Russia

Abstract. We present an analysis of the applicability of previously reported approach (NDM-net) to the construction of depth seismic images. The NDM-net allows us to simulate seismic data by combining finite differences with deep learning method. Coarse computational grids are used for the fast generation of inaccurate datasets. After that, the training dataset is simulated for a small number of shot points on fine grids. Then, the NDM-net is trained to map inaccurate data to accurate common shot gathers. Previously, we illustrated the high quality of the generated shot gathers. In this research, we applied the prestack depth migration to estimate the effect of the NDM-net numerical errors on the resulting seismic images.

Keywords: Seismic modeling · Machine learning · Prestack depth migration

1 Introduction

The development of mathematical apparatus and computational technologies has led to the spread of specialized data processing procedures, such as full waveform inversion or FWI [4,31]; depth migration based on the one-way wave equation [25,26]; reverse time migration or RTM [27,30]. However, the computational complexity and resource intensity of seismic modeling, as the main component of these procedures, are significant limiting factors in the development of these methods.

Seismic modeling can also be used to preliminarily assess the applicability of certain data processing procedures in specific complex geological settings [2,11,17]. This requires obtaining high-quality synthetic seismic data to enable testing of processing workflows with confidence that the processing result will not be affected by numerical errors associated with numerical modeling. Calculating high-quality common shot gathers requires the use of modern numerical meth-

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16761, pp. 58–69, 2027.
https://doi.org/10.1007/978-3-032-30527-5_4

ods, such as the finite difference method [14,30], the spectral element method [12,29], the discontinuous Galerkin method [10,18] in combination with sufficiently detailed grids for discretizing the model and wavefield. Restricting the computational grid steps is the primary way to ensure acceptable computational accuracy: the smaller the grid step (the higher the number of degrees of freedom in the discretized problem), the higher the accuracy. Note that the numerical error, in general, decreases proportionally to the square of the grid step, while the problem size increases proportionally to the fourth power of the grid step. Therefore, attempts to suppress the numerical error solely by reducing the grid step quickly lead to excessive demands on computer memory and computation time, without significantly improving the quality of the solution.

Principally different results can be achieved by postprocessing the simulated data, aimed specifically at suppressing the specific numerical error expressed as numerical dispersion. There are approaches based on analytical estimates of dispersion followed by spectrum correction [20,24], and machine learning methods [5–9,33,34]. The quality measure of the resulting solution with such correction is the norm of the difference between common shot gathers, for example, the L_2 norm or NRMS [13]. However, these are integral measures applied to common shot gathers, which do not always allow for monitoring the impact of postprocessing on seismic events. In this work, prestack depth migration is performed on simulated, corrected, and uncorrected data, and the resulting seismic images are analyzed.

2 Seismic Modeling with NDM-Net

The classical formulation of a seismic modeling problem involves solving the elastic wave equation:

$$\begin{gathered} \rho \frac{\partial \boldsymbol{u}}{\partial t} = \nabla \cdot \sigma, \\ \frac{\partial \varepsilon}{\partial t} = \frac{1}{2}\left[\nabla u + \nabla u^T\right], \\ \sigma = C\varepsilon, \end{gathered}$$

where $\boldsymbol{u}$ is the velocity vector, σ and ε are the stress and strain tensors, respectively, ρ is the mass density, and C is the stiffness tensor. In this research we deal with the isotropic models, where tensor C has a simple form with only two independent parameters. The problem is stated in the half-space $x_3 > 0$ with free-surface boundary conditions at $x_3 = 0$. To avoid the reflections from the other boundaries the Perfectly matched layers are applied [1,3,19].

The calculation is performed for a set of sources located on the free surface at specified points $\boldsymbol{x}_j^s$. Wavefields (displacement velocities) are recorded on the free surface at receiver locations $\boldsymbol{x}_k^r$. In this case, the modeling can ensure a fixed geometry of the active array, identical for all source positions. As a result, all common shot gathers will have the same size, determined by the number of receivers in the active array and the number of time samples.

The numerical modeling is performed using the finite difference method. The accuracy of the calculations is determined by the grid step, which is selected based on the acceptable level of numerical dispersion. In this research we use the

second order accurate in time and the fourth order accurate in space staggered grid scheme [14]. This scheme possesses an acceptable level of numerical dispersion when approximately five points per dominant wavelength are used, which can range from 1 to 5 m, depending on the velocity model and the frequency of the probing signal [16,21]. As already noted, halving the grid step reduces the error by approximately a factor of four, since the numerical solution converges to the exact solution in the second order. In fact, the estimate is even coarser for heterogeneous models where the convergence rate may degenerate to the first order [15,22,23,28,32]. At the same time, the problem size increases by a factor of eight, and the computational time increases by 16 in the three-dimensional case (4 and 8 times, respectively, in the two-dimensional setting). Increasing the formal order of the approximation does not yield any significant improvement [23].

To reduce the amount of computation and running time, several studies [5–8] proposed an approach that combines the finite difference method and machine learning methods. This approach, called NDM-net, consists of the following steps:

- Calculation of a complete set of common shot gathers $V_c(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)$ on a coarse grid. In these notations, the subscript c denotes the solution computed on a coarse grid or coarse-grid solution, $\boldsymbol{x}_j^s$ is the coordinate of the $j-$th source, where $j = 1, ..., J$, and $\boldsymbol{x}^o$ is the offset of the receiver with respect to the considered source. This solution can be obtained in a short time; however, it has a high numerical error

$$\epsilon_c = \|V_c(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o) - V(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)\|$$

  where $V(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)$ is the exact solution.
- Construction of a training sample that is the calculation of a set of shot gathers $V_f(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)$ on a fine grid for a small number of sources with $j \subset \{1, ..., J\}$. The calculation time for one common shot gather in this case is significantly higher than when using a coarse grid, and the total time for constructing the training sample is comparable to the time for calculating the full dataset on a coarse grid. However, the numerical error of the reference solution

$$\epsilon_f = \|V_f(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o) - V(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)\|$$

  is significantly lower than ϵ_c. From now on we consider V_f as the exact solution.
- Training the NDM-net - the neural network based on the U-net architecture to map noisy data to fine grid solution, i.e.

$$\mathcal{N} : V_c(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o) \rightarrow V_f(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o).$$

  The aim of mapping is to minimize the error between common shot gathers:

$$\epsilon_{\mathcal{N}} = \|V_f(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o) - V_c(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)\| \rightarrow \min$$

Note that there are two modifications of NDM-net. One is focused on working with source data in the time domain [7], the second is a modified network that works with data in the frequency domain [6]. A description of the advantages of each approach is presented in the above-mentioned works. This stage is also quite resource-intensive and comparable with time needed to the calculation of wavefields on a coarse grid.

– The final stage is the application of the trained neural network to correct the entire dataset. This part of the method is the least computationally intensive.

3 Prestack Depth Migration

In our previous studies, we showed that the use of the NDM-net can reduce the numerical error in synthetic data from 70–100% to 20–25%. However, these estimates were obtained by directly comparing common shot gathers, which still does not allow us to judge the suitability of these data for constructing seismic images. In this study, prestack depth migration was applied to three sets of synthetic data: those calculated on a coarse grid $V_c(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)$, those calculated on a fine grid (for all source positions) $V_f(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)$, and those obtained using NDM-net $\mathcal{N}\left[V_c(t, \boldsymbol{x}_j^s, \boldsymbol{x}^o)\right]$.

The Kirhgoff prestack depth migration was used to construct the seismic images. This algorithm involves summing the amplitude values along a diffraction hyperbola and placing the resulting value at the point corresponding to the hyperbola's apex. The aim of depth migration is to obtain seismic data at a depth scale that correctly accounts for spatial offset and, therefore, accurately determines the coordinates of the reflector, as well as provides the most complete representation of deep tectonics.

Prestack depth migration (PSDM) processes seismic data before stacking, working directly with common-shot or common-offset gathers to position reflections accurately. Unlike time migration, which assumes a flat earth and constant velocity, PSDM uses a depth-velocity model to trace ray paths from sources to receivers through laterally varying velocities. This corrects for complex structures like faults. In Kirchhoff PSDM, traveltime tables are first computed for every source-receiver pair to every image point in the subsurface grid, often via wavefront reconstruction or Eikonal solvers. For each trace, the algorithm then applies a weighted summation of amplitudes along the diffraction traveltime surface, imaging energy at the apex where specular reflection occurs.

The depth migration was performed with the following parameters: the depth-velocity model is the original BP velocity model. We used the wavefront reconstruction to calculate traveltimes. The aperture diameter was 6 km, the output offset ranged between 0–3200 m with a step of 25 m, the maximum frequency was 90 Hz, and the maximum depth was 12 km.

4 Numerical Experiment

The described seismic modeling and seismic imaging procedure was performed for the BP velocity model, the most complex and representative among the two-dimensional models available. The model is presented in Fig. 1. The model size is 67.4 km in the horizontal direction and 11.9 km in the vertical direction. The acquisition system includes 2696 sources with a step of 25 m between them. The active array is central, with 512 receivers with a step of 25 m and a maximum offset of 6.4 km. A Ricker wavelet with a central frequency of 30 Hz was used as the source. A coarse grid with a step of 6 m was used, while the step of the fine grid was equal to 3 m. The sample interval of the common shot gathers was 3 ms and the record length was 6 s. Ten percent of the total number of sources were used to construct the training dataset. As a result, the initial error between the data from the coarse and fine grids was 92%; after data correction using NDM-net, the error decreased to 31%, which confirms the results presented in [5].

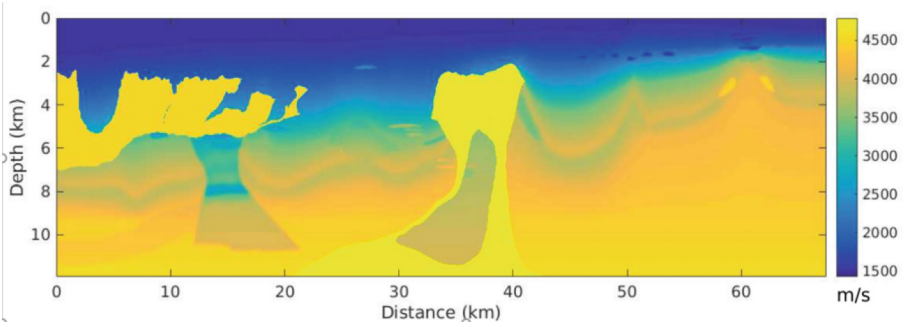

Fig. 1. BP velocity model (P-wave velocity).

The shot gather corresponding to the source located at $x_1 = 26.500$ m are shown in Fig. 2. We also provide the difference between the fine grid solution and the other two gathers to illustrate the numerical error before and after the application of the NDM-net (Fig. 3). The difference of the two shoth gathers consistently show an increase in numerical error with increasing recording time. This is because the main error is the numerical dispersion, which leads to deformation and decay of the wavepacket. Furthermore, a phase delay is evident, which increases with increasing time traveling. NDM-net effectively suppresses signal phase changes, but does not accurately correct its amplitude.

The depth seismic sections obtained for the three datasets are shown in Figs. 4, 5 and 6. The section constructed using the coarse grid data (Fig. 5)

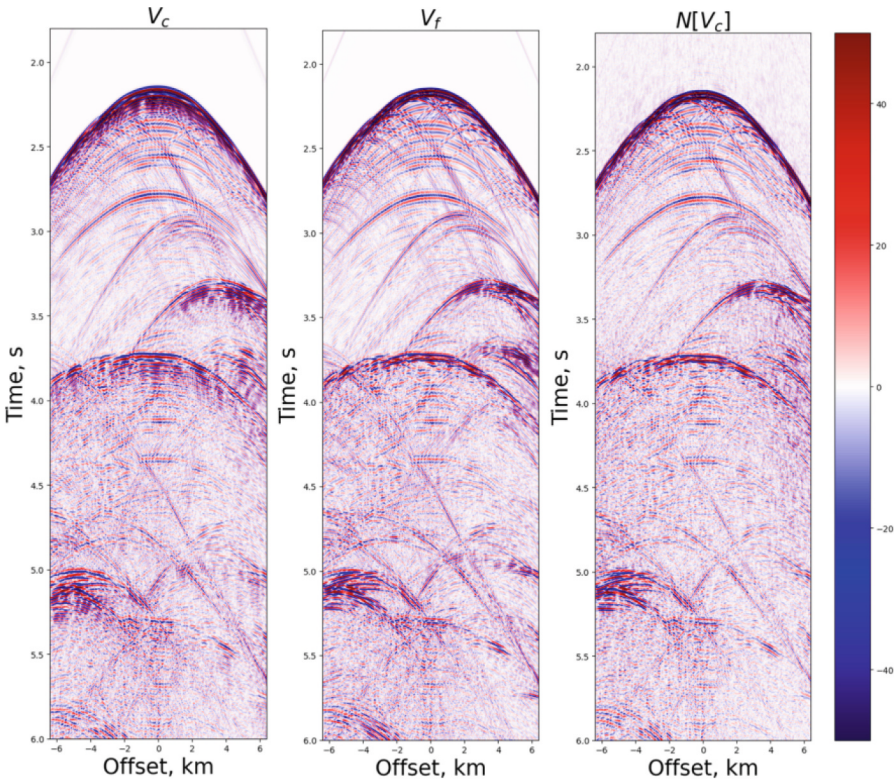

Fig. 2. Sample shot gather corresponding to the source position $x_1 = 26.500$. The coarse grid solution is at the left, the fine grid solution is at the center, and the solution corrected with the NDM-net is at the right.

reveals pronounced oscillations below the main seismic events. This effect is caused by numerical dispersion, which deforms the wavepacket. Some seismic events, clearly visible in the reference section (Fig. 4), are indistinguishable in Fig. 5 due to signal interference. This effect is observed in areas of thin layering (on the right side of the section and in the central part between the intrusive bodies). This effect is pronounced in the section constructed using the NDM-net-corrected data (Fig. 6).

For a more detailed analysis of the effect of the NDM-net correction, we considered the differences between the constructed sections as presented in Figs. 7 and 8. Note that all provided images are of the same gain to simplify visualization and comparison. The use of a coarse grid results in an extremely high error in the migrated data, which is evident in almost all seismic events, starting from the earliest ones. The use of NDM-net allows us to significantly suppress this error.

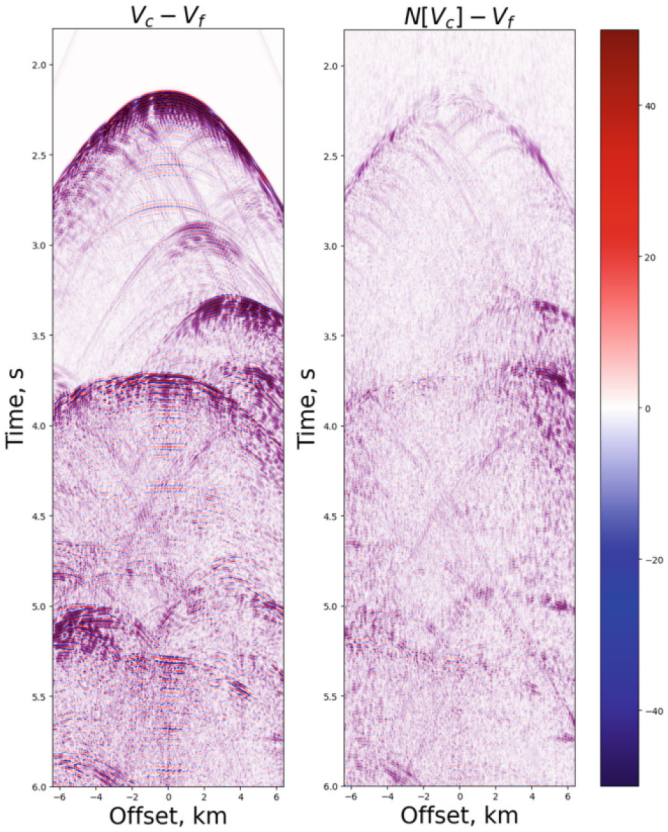

Fig. 3. The difference between the fine grid solution and the coarse grid solution is at the left. The difference between the fine grid solution and the NDM-net corrected solution is at the right.

The events at shallow depths contain mainly no numerical error after applying NDM-net. The magnitude of the residual error, on the one hand, increases with depth, and, on the other hand, it manifests itself at sharply contrasting boundaries (salt body boundaries). The first type of error is associated with the increase in error in the original data with depth due to the longer wave propagation path. The second type of error is related to the fact that NDM-net primarily corrects the most significant component of the error (the phase shift) while the amplitude error remains virtually uncorrected. However, NDM-net correction made a deep seismic event in the central part of the model out of phase. This

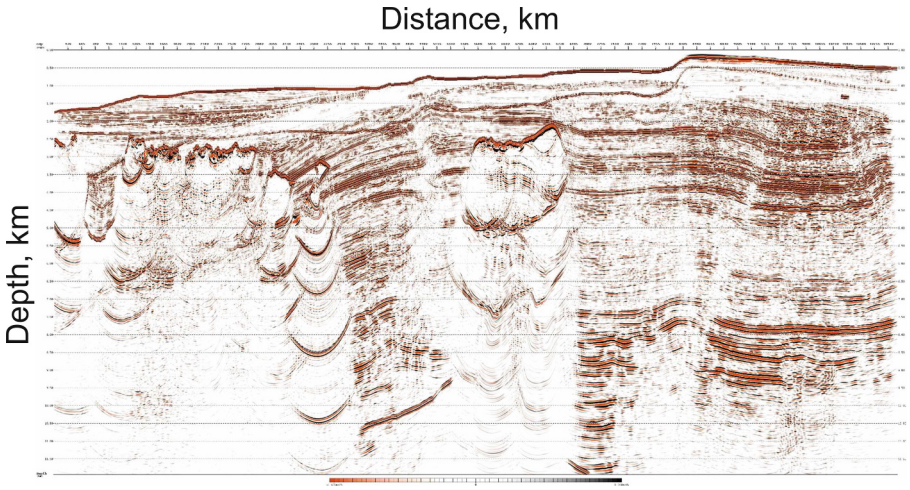

Fig. 4. The seismic depth section constructed from the fine grid data.

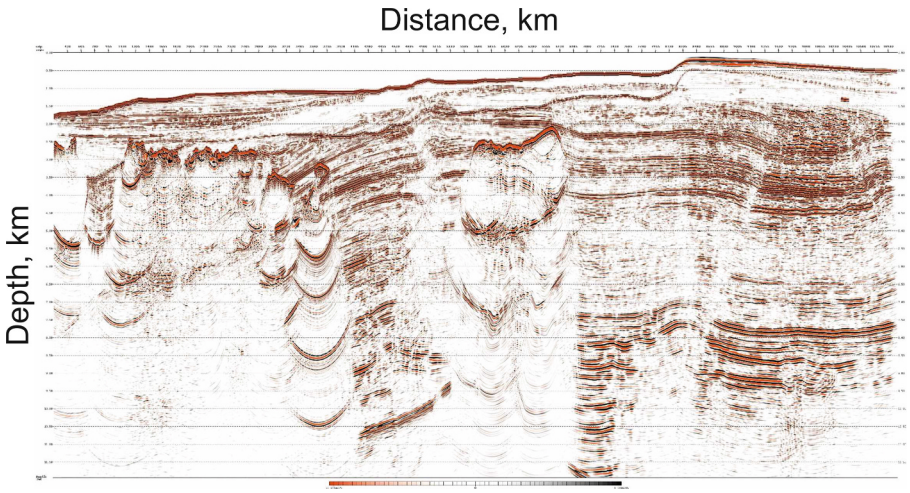

Fig. 5. The seismic depth section constructed from the coarse grid data.

is undoubtedly an error introduced by the neural network that requires further refinement.

The primary aim of the developed algorithm is to speed up seismic modelling. Therefore, we present a performance analysis of various scenarios. For the velocity model considered in this study, the computation time was 1 min per seis-

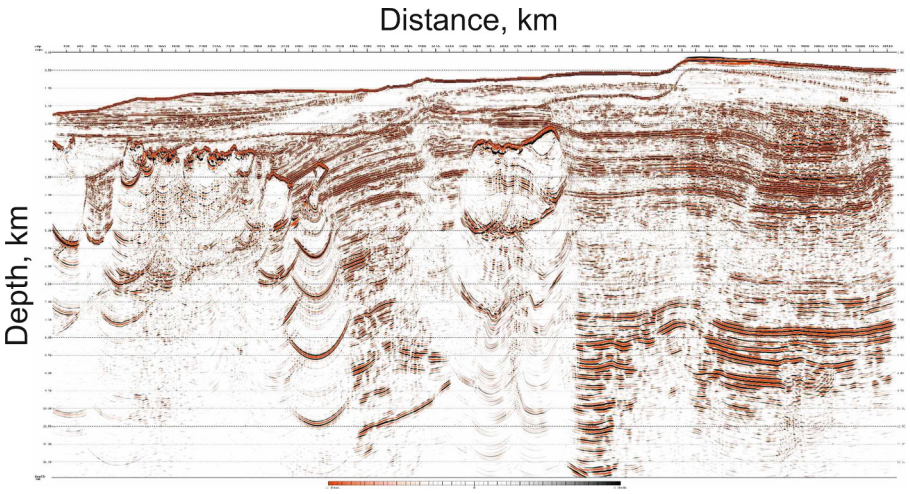

Fig. 6. The seismic depth section constructed from the NDM-net corrected data.

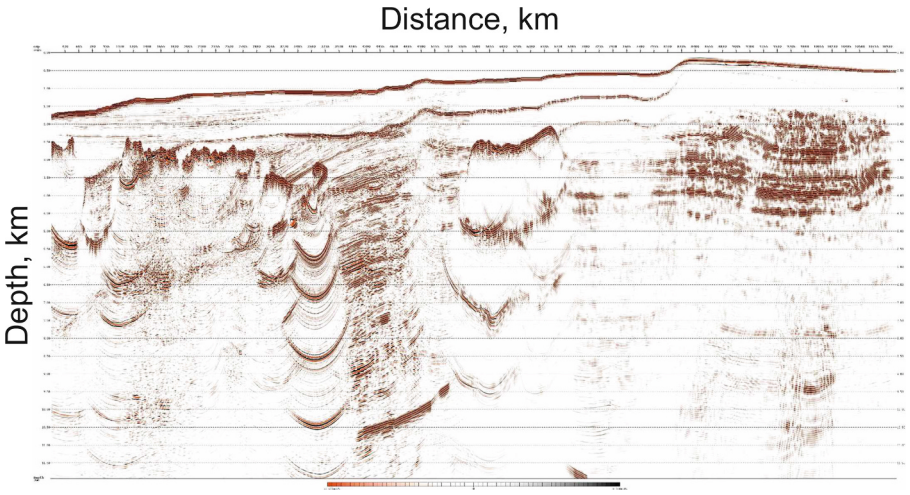

Fig. 7. The difference between seismic depth section constructed from the coarse grid data and fine grid data.

mogram (single souce) when using a coarse computational grid and 6 min when using a fine computational grid. The calculations and subsequent training of the neural network were carried out on a local server with 8 Nvidia A100 GPUs with 80 Gb of RAM. The training time for NDM-net is 60 min for a sample of 10% of the total number of sources, Thus, the speed-up factor is 3-4.

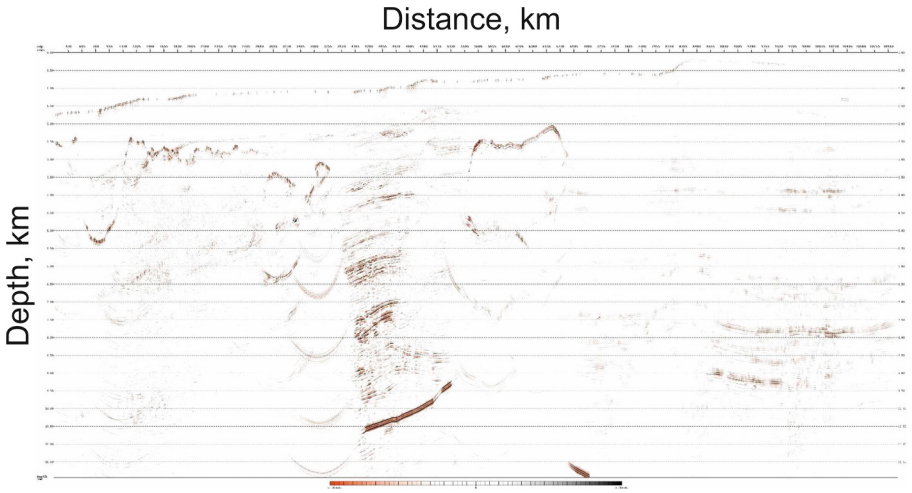

Fig. 8. The difference between seismic depth section constructed from the NDM-net corrected data and fine grid data.

5 Conclusions

This paper presented an analysis of the results of numerical modeling of seismic wavefields using a method based on a combined neural network approach and the finite difference method. Using the BP model as an example, numerical experiments were conducted using fine and coarse grids to compute the seismorgams, and NDM-net was also applied to correct noisy data. Prestack depth migration with a known velocity model was performed for all three datasets and depth sections were constructed. The images constructed from coarse grid data are shown to contain a significant error in seismic events. Data correction using NDM-net can significantly improve data quality. However, for events at high depths, the effectiveness of NDM-net may decrease. One possible way to improve the effectiveness of NDM-net in this case could be to modify the training set by increasing the number of sampling sources in the vicinity of target objects (increasing the illumination of objects within the training set).

Acknowledgments. Seismic modelling and NDM-net correction was done by the IM SB RAS group under the support of the Russian Science Foundation grant no. 22-11-00004-*П*.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Berenger, J.P.: A perfectly matched layer for the absorption of electromagnetic waves. J. Comput. Phys. **114**, 185–200 (1994). https://doi.org/10.1006/jcph.1994.1159
2. Cheverda, V., et al.: Three-dimensional model of oil and gas reservoirs based on gaussian beam processing of scattered seismic waves. Russ. Geol. Geophys. **63**(1), 109–123 (2022). https://doi.org/10.2113/rgg20204231
3. Collino, F., Tsogka, C.: Application of the perfectly matched layer absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media. Geophysics **66**, 294–307 (2001)
4. Egorov, A., et al.: Elastic full-waveform inversion of vertical seismic profile data acquired with distributed acoustic sensors. Geophysics **83**(3), R273–R281 (2018)
5. Gadylshin, K., Gondyul, E., Lisitsa, V., Gadylshina, K., Vishnevsky, D.: Mitigation of numerical dispersion in seismic data in spectral domain with neural networks. Soil Dyn. Earthq. Eng. **187**, 109028 (2024). https://doi.org/10.1016/j.soildyn.2024.109028
6. Gadylshin, K., Lisitsa, V., Vishnevsky, D., Gadylshina, K.: Hausdorff-distance-based training dataset construction for numerical dispersion mitigation neural network. Comput. Geosci. **180**, 105438 (2023). https://doi.org/10.1016/j.cageo.2023.105438
7. Gadylshin, K., Vishnevsky, D., Gadylshina, K., Lisitsa, V.: Numerical dispersion mitigation neural network for seismic modeling. Geophysics **87**(3), T237–T249 (2022). https://doi.org/10.1190/geo2021-0242.1
8. 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
9. Han, Y., Wu, B., Yao, G., Ma, X., Wu, D.: Eliminate time dispersion of seismic wavefield simulation with semi-supervised deep learning. Energies **15**(20), 7701 (2022)
10. Kaser, M., Dumbser, M.: An arbitrary high-order discontinuous galerkin method for elastic waves on unstructured meshes - i. the two-dimensional isotropic case with external source terms. Geophys. J. Int. **166**(2), 855–877 (2006)
11. Kolyukhin, D.R., et al.: Seismic imaging and statistical analysis of fault facies models. Interpretation **5**(4), SP71–SP82 (2017)
12. 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. Soc. Am. **88**(2), 368–392 (1998). https://doi.org/10.1785/BSSA0880020368
13. Kragh, E., Christie, P.: Seismic repeatability, normalized RMS, and predictability. Lead. Edge **21**(7), 640–647 (2002)
14. Levander, A.R.: Fourth-order finite-difference P-SV seismograms. Geophysics **53**(11), 1425–1436 (1988). https://doi.org/10.1190/1.1442422
15. Lisitsa, V., Podgornova, O., Tcheverda, V.: On the interface error analysis for finite difference wave simulation. Comput. Geosci. **14**(4), 769–778 (2010). https://doi.org/10.1007/s10596-010-9187-1
16. Lisitsa, V.: Dispersion analysis of discontinuous galerkin method on triangular mesh for elastic wave equation. Appl. Math. Model. **40**, 5077–5095 (2016). https://doi.org/10.1016/j.apm.2015.12.039
17. Lisitsa, V., Kolyukhin, D., Tcheverda, V.: Statistical analysis of free-surface variability's impact on seismic wavefield. Soil Dyn. Earthq. Eng. **116**, 86–95 (2019)

18. 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). https://doi.org/10.1016/j.jcp.2016.02.005
19. Lisitsa, V.V.: Optimal grids for numerical solution of wave equation with variable coefficients. Sib. Zh. Vychisl. Mat. **8**(3), 219–229 (2005). in Russian
20. Mittet, R.: Second-order time integration of the wave equation with dispersion correction procedures. Geophysics **84**(4), T221–T235 (2019)
21. Moczo, P., Kristek, J., Halada, L.: 3D fourth-order staggered-grid finite-difference schemes: stability and grid dispersion. Bull. Seismol. Soc. Am. **90**(3), 587–603 (2000)
22. Moczo, P., Kristek, J., Vavrycuk, V., Archuleta, R.J., Halada, L.: 3D heterogeneous staggered-grid finite-difference modeling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities. Bull. Seismol. Soc. Am. **92**(8), 3042–3066 (2002). https://doi.org/10.1785/0120010167
23. Moczo, P., Kristek, J., Galis, M.: The Finite-Difference Modelling of Earthquake Motion: Waves and Ruptures. Cambridge University Press (2014)
24. Mulder, W.A.: Temporal dispersion correction for wave-propagation modelling with a series approach. Geophys. Prospect. **72**(2), 301–314 (2024). https://doi.org/10.1111/1365-2478.13411
25. Pleshkevich, A., Lisitsa, V., Vishnevsky, D., Levchenko, V.: A multi-level parallel algorithm for seismic imaging based on one-way wave equation migration. Comput. Math. Appl. **97**, 344–354 (2021)
26. Pleshkevich, A., Vishnevskiy, D., Lisitsa, V.: Sixth-order accurate pseudo-spectral method for solving one-way wave equation. Appl. Math. Comput. **359**, 34–51 (2019)
27. Shragge, J., Girard, A.J., Oren, C., Krahenbuhl, R.: Elastic time-reverse imaging of backscattered surface-wave data. Geophysics **89**(4), KS119–KS128 (2024). https://doi.org/10.1190/geo2023-0569.1
28. Symes, W., Vdovina, T.: Interface error analysis for numerical wave propagation. Comput. Geosci. **13**(3), 363–371 (2009)
29. Tromp, J., Komatitsch, D., Liu, Q.: Spectral-element and adjoint methods in seismology. Commun. Comput. Phys. **3**(1), 1–32 (2008)
30. Virieux, J., Calandra, H., Plessix, R.E.: A review of the spectral, pseudo-spectral, finite-difference and finite-element modelling techniques for geophysical imaging. Geophys. Prospect. **59**(5), 794–813 (2011). https://doi.org/10.1111/j.1365-2478.2011.00967.x
31. Virieux, J., Operto, S.: An overview of full-waveform inversion in exploration geophysics. Geophysics **74**(6), WCC1–WCC26 (2009)
32. Vishnevsky, D., Lisitsa, V., Tcheverda, V., Reshetova, G.: Numerical study of the interface errors of finite-difference simulations of seismic waves. Geophysics **79**(4), T219–T232 (2014). https://doi.org/10.1190/geo2013-0299.1
33. Xiang, Y., Wang, Z., Song, Z., Huang, R., Song, G., Min, F.: Seismictransformer: an attention-based deep learning method for the simulation of seismic wavefields. Comput. Geosci. **190**, 105629 (2024). https://doi.org/10.1016/j.cageo.2024.105629
34. Yu, S., Ma, J.: Deep learning for geophysics: current and future trends. Rev. Geophys. **59**(3), e2021RG000742 (2021)