




Machine Learning-Based Numerical Dispersion Mitigation in Seismic Modelling

Kirill Gadylshin¹, Vadim Lisitsa¹ , Kseniia Gadylshina²,
Dmitry Vishnevsky², and Mikhail Novikov²

¹ Sobolev Institute of Mathematics SB RAS, 4 Koptug ave.,
630090 Novosibirsk, Russia
lisitsavv@ipgg.sbras.ru

² Institute of Petroleum Geology and Geophysics SB RAS,
3 Koptug ave., 630090 Novosibirsk, Russia

Abstract. We present an original approach to improving seismic modelling performance by applying deep learning techniques to mitigate numerical error. In seismic modelling, a series of several thousand simulations are required to generate a typical seismic dataset. These simulations are performed for different source positions (equidistantly distributed) at the free surface. Thus, the output wavefields that corresponded to the nearby sources are relatively similar, sharing common peculiarities. Our approach suggests simulating wavefields using finite differences with coarse enough discretization to reduce the computational complexity of seismic modelling. After that, solutions for 1 to 10 percents of source positions are simulated using fine discretizations to obtain the training dataset, which is used to train the deep neural network to remove numerical error (numerical dispersion) from the coarse-grid simulated wavefields. Later the network is applied to the entire dataset. Our experiments illustrate that the suggested algorithm in the 2D case significantly (up to ten times) speeds up seismic modelling.

Keywords: Deep learning · Seismic modelling · Numerical dispersion

1 Introduction

Seismic modelling becomes a common tool to investigate peculiarities of wave propagation in realistic complex models of the Earth's interior [2, 12, 24] verification of the seismic processing and inversion algorithms, and as a part of the inversion methods, [21]. However, simulation of seismic wave propagation in complex media is one of the most computationally demanding problems requiring intense use of high-performance computing. In particular, if a typical seismic

K. Gadylshin and V. Lisitsa—are grateful to Mathematical Center in Akademgorodok, the agreement with Ministry of Science and High Education of the Russian Federation number 075-15-2019-1613 for the financial support. MN is supported by the Agency of the Precedent of Russian Federation, grant no. MK-3947.2021.1.5.

acquisition system is considered, one has to simulate wavefields corresponding to hundreds of thousands of source positions (right-hand sides). Each simulation of a single shot gather is performed in a domain of about 10^3 km, which corresponds to 100^3 wavelength. Thus, up to $8 \cdot 10^9$ grid points are needed to obtain accurate enough numerical results. Reduction of the problem size by increasing the grid step leads to numerical error growth, which may completely destroy the solution. There are several ways to reduce the numerical dispersion, including use of high order finite-difference schemes [11], dispersion-suppression schemes [14], high-order finite element and discontinuous Galerkin methods [1, 8, 13]. However, the increase of the approach accuracy imminently leads to high computational intensity, including increased flops and RAM access operations.

The other option to reduce the numerical dispersion in the simulated wavefields is a post-processing [9, 23]. However, the standard waveform correction procedures used in seismic processing are not efficient for numerical dispersion mitigation. The error associated with the numerical dispersion depends on the wave propagation path, velocity model etc. Thus, it can not be compensated by a single-phase shift. In this paper, we suggest an approach to post-processing based on using the deep learning technique.

Deep learning finds wide application in various fields of science. Providing a large representative training dataset, deep neural networks (DNNs) can approximate complex non-linear operators within the supervised learning workflow. These DNNs can learn about highly non-linear physics and usually provide much faster computational time than traditional simulation [6, 17].

To develop an efficient algorithm for numerical dispersion mitigation, we use the following peculiarity of seismic modelling. The entire seismic dataset includes wavefields corresponding to different source positions. These positions are relatively close to each other (10 to 100 m apart). Thus, the velocity models and the simulated wavefields are similar if the source is situated nearby. It allows using a small number of sources to simulate accurate solution to be used as a training dataset. At the same time, we can simulate the entire dataset using a coarse enough grid, train the deep neural network, and then post-process the data.

The remainder of the paper has the following structure. In Sect. 2 we remind the basic concepts of seismic modelling, including the main estimates of the numerical dispersion, depending on the grid size. The description of the numerical dispersion mitigation network (NDM-net) is provided in Sect. 3. Numerical experiments illustrating the applicability of the NDM-net to the synthetic seismic data enhancement are presented in Sect. 4.

2 Seismic Modelling

Seismic wave propagation in 2D isotropic elastic media is governed by the elastic wave equation:

$$\begin{aligned} \rho \frac{\partial u_1}{\partial t} &= \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{13}}{\partial x_3}, & \frac{\partial \sigma_{11}}{\partial t} &= (\lambda + 2\mu) \frac{\partial u_1}{\partial x_1} + \lambda \frac{\partial u_3}{\partial x_3} + f_{11}(t) \delta(\mathbf{x} - \mathbf{x}_s), \\ \rho \frac{\partial u_3}{\partial t} &= \frac{\partial \sigma_{13}}{\partial x_1} + \frac{\partial \sigma_{33}}{\partial x_3}, & \frac{\partial \sigma_{33}}{\partial t} &= \lambda \frac{\partial u_1}{\partial x_1} + (\lambda + 2\mu) \frac{\partial u_3}{\partial x_3} + f_{33}(t) \delta(\mathbf{x} - \mathbf{x}_s), \\ & & \frac{\partial \sigma_{13}}{\partial t} &= \mu \frac{\partial u_1}{\partial x_3} + \mu \frac{\partial u_3}{\partial x_1} + f_{13}(t) \delta(\mathbf{x} - \mathbf{x}_s), \end{aligned} \quad (1)$$

where ρ is the mass density, λ and μ are the Lamé parameters, $\mathbf{u} = (u_1, u_3)^T$ is the particle velocity vector, σ is the stress tensor, $f_{ij}(t)$ are the components of the source wavelet function, $\delta(\mathbf{x})$ is the Kronecker delta-function, \mathbf{x} is the vector of spatial coordinates, and \mathbf{x}_s is the source coordinate. The seismic modelling is stated in half-space $x_3 > 0$ and within bounded time interval $t \in [0, T]$.

A common way to approximate the elastic wave equation is the use of staggered grid finite differences [11, 20], where the different components of the wavefield are defined at different spatial and temporal points with the use of symmetric stencils to approximate the derivatives:

$$\begin{aligned}
 \rho D_t[u_1]_{i+1/2,j}^{n-1/2} &= D_1[\sigma_{11}]_{i+1/2,j}^{n-1/2} + D_3[\sigma_{13}]_{i+1/2,j}^{n-1/2}, \\
 \rho D_t[u_3]_{i,j+1/2}^{n-1/2} &= D_1[\sigma_{13}]_{i,j+1/2}^{n-1/2} + D_3[\sigma_{33}]_{i,j+1/2}^{n-1/2}, \\
 D_t[\sigma_{11}]_{i,j}^n &= (\lambda + 2\mu)D_1[u_1]_{i,j}^n + \lambda D_3[u_3]_{i,j}^n + f_{11}(t^n)[\delta(\mathbf{x} - \mathbf{x}_s)]_{i,j}, \\
 D_t[\sigma_{33}]_{i,j}^n &= \lambda D_1[u_1]_{i,j}^n + (\lambda + 2\mu)D_3[u_3]_{i,j}^n + f_{33}(t^n)[\delta(\mathbf{x} - \mathbf{x}_s)]_{i,j}, \\
 D_t[\sigma_{13}]_{i+1/2,j+1/2}^n &= \mu D_1[u_3]_{i+1/2,j+1/2}^n + \mu D_3[u_1]_{i+1/2,j+1/2}^n + \\
 &\quad + f_{13}(t^n)[\delta(\mathbf{x} - \mathbf{x}_s)]_{i+1/2,j+1/2},
 \end{aligned} \tag{2}$$

where finite-difference operators are

$$\begin{aligned}
 D_t[g]_{I,J}^N &= \frac{g_{I,J}^{N+1/2} - g_{I,J}^{N-1/2}}{\tau}, \\
 D_1[g]_{I,J}^N &= \frac{1}{h_1} \sum_{m=0}^M \alpha_m \left(g_{I+m+1/2,J}^N - g_{I-m-1/2,J}^N \right), \\
 D_2[g]_{I,J}^N &= \frac{1}{h_2} \sum_{m=0}^M \alpha_m \left(g_{I,J+m+1/2}^N - g_{I,J-m-1/2}^N \right),
 \end{aligned} \tag{3}$$

where indices I, J, N can be either integer or half-integer, and g is a smooth enough scalar function. The operator D_t approximates the temporal derivatives with the second order to the time step τ . The operators D_1 and D_2 approximate the spatial derivatives. However, depending on the choice of α_m one may construct a high-order approximation up to $2m + 2$ [11], or use these degrees of freedom to suppress numerical dispersion [14]. Note, that we are not discussing the approximation of the right-hand sides, as it is presented in [7], and the model parameters treatment because it is studied in [16, 22].

The use of symmetric stencils to approximate derivatives ensures an even order of approximation with zero coefficients of odd degrees in the differential approximation of the finite difference scheme (2). Thus the numerical error appears in the solution as a numerical dispersion without dissipation, see [19], and [3] for the details. It means that the emitted impulse will deteriorate, propagating through the media. An example of the impulse deformation due to the dispersion is presented in Fig. 1. We plot the true pulse and that travelled 30 wavelengths simulated by the second-order scheme with a spatial discretization of 10, 20, and 40 points per wavelength and Courant number equal to 0.8. Note that the maximum impulse shifts backwards in time, leading to an overestimation of the reflecting intervals depth in seismic processing and interpretation. Refining the mesh, one gets the convergence of the numerical solution to the true one, however, refining a spatial step by the factor of two leads to the increase of the problem size by the factor of 8 in 3D and 4 in 2D. Moreover, the number

of flops increases by 16 in 3D and 8 in 2D because of the temporal step refinement. On the other hand, the solution obtained on the grid with 20 points per wavelength is accurate enough, and simple processing may turn it into the true one.

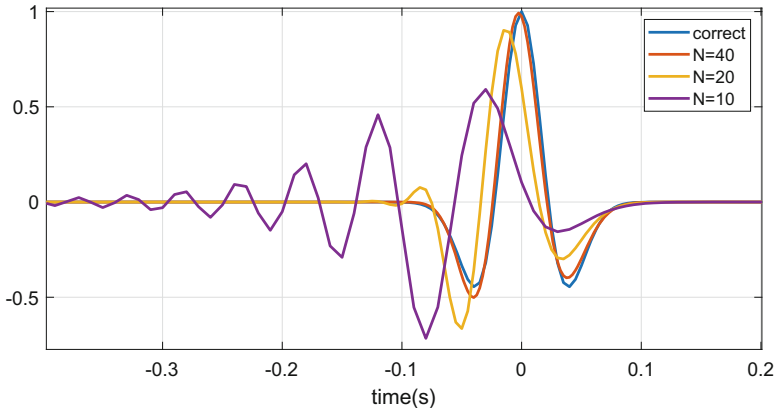


Fig. 1. An example of the pulse deformation due to numerical dispersion.

3 Numerical Dispersion Mitigation Network (NDM-net)

Convolutional Neural Networks (CNN) are usually applied to analyze visual imagery. A particular case of CNN is a U-Net [18], which was originally introduced for biomedical image segmentation. At this moment, the U-Net and its modifications have broad applications in seismic inversion, pre-stack seismic data processing and interpretation. This work suggests using the Numerical Dispersion Mitigation deep neural network (NDM-net) to learn the mapping between the synthetic seismic data modelled on a coarse grid and data modelled on a fine grid. In other words, we plan to eliminate the numerical dispersion using the Deep Learning approach.

The architecture of the network is similar to the one used by [5]. The differences are using a conventional convolutional layer instead of partial convolutions and the different input/output dimensions, see Fig. 2. These DNN contains 16 convolutional layers, eight upsampling layers, and eight concatenation layers (skip connections). The input and output tensors dimensions are $1250 \times 512 \times 2$. An activation function for the first eight convolutional layers (encoding, or feature extracting, part of the DNN) is ReLU, while the last eight convolutional layers (decoding part) have LeakyReLU activation with a negative slope coefficient equals to 0.2. We implemented NDM-net in TensorFlow. The DNN weights were randomly initialized, and Adam stochastic optimization algorithm was exploited during the training process.

In the current implementation, we consider the input/output to be regularly sampled pre-stack seismic data. For training, we used each 10-th common shot gather computed on a fine grid and its corrupted version modelled on a coarse grid. Each common shot is converted to a tensor with a dimension of $1250 \times 512 \times 2$. Here 1250 is the number of time samples in data (4 ms time discretization and 5s record time), 512 is the number of 2C receivers, and 2 is the number of recorded components (vertical and horizontal velocity components). Next, we split this dataset into training and validation datasets. Each common shot is normalized by scaling it to unit variance before being processed by the NDM-net.

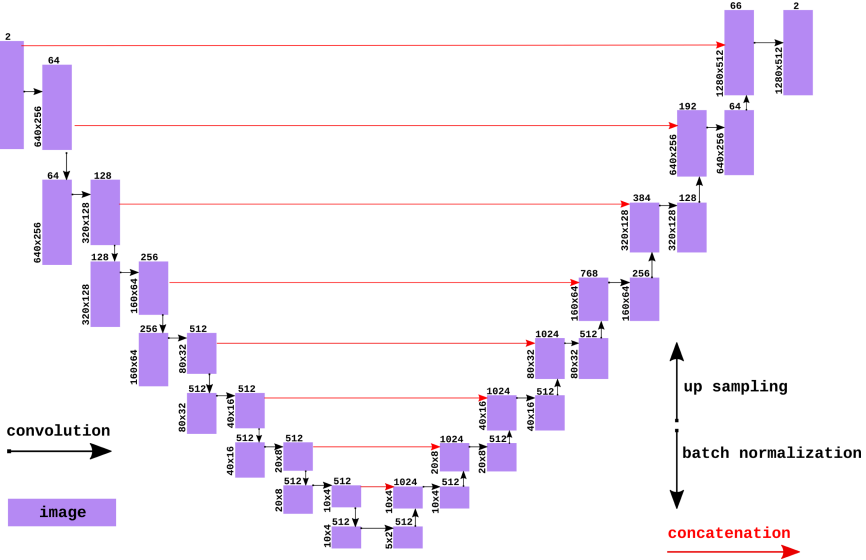


Fig. 2. The architecture of NDM-net. The Black right arrow indicates convolution operation, while the red right arrow indicates concatenation. Up and down arrows indicate upsampling and batch normalization operations correspondingly.

4 Numerical Experiments

We applied our approach to mitigate the numerical dispersion in two datasets. Both simulations were done in 2D to illustrate the applicability of the NDM-net to improve seismic modelling accuracy and efficiency.

4.1 Marmousi2 Model

First we considered the elastic Marmousi2 model [15], as presented in Fig. 3. The size of the model was 17km in the horizontal and 3.6km in the vertical direction. Marmousi2 is the offshore model with water at the top. To make the considerations consistent with land data acquisition, we substitute water with solid used for the ocean bottom in the model. We performed simulations of seismic waves propagation using meshes with steps equal to 1.25 m, 2.5 m, and 5 m, assuming the solution obtained on the 1.25 m grid is the exact one. Such small grid steps were chosen due to the thick low-velocity layer, that was introduced instead of water at the top of the model. Note that the original model was provided on a grid with step size 1.25 m. However, to exclude the effect of model changes when the simulation mesh is coarsening, we map the mode to the mesh with step 5 m. After that 5-meters model was used for all numerical simulations.

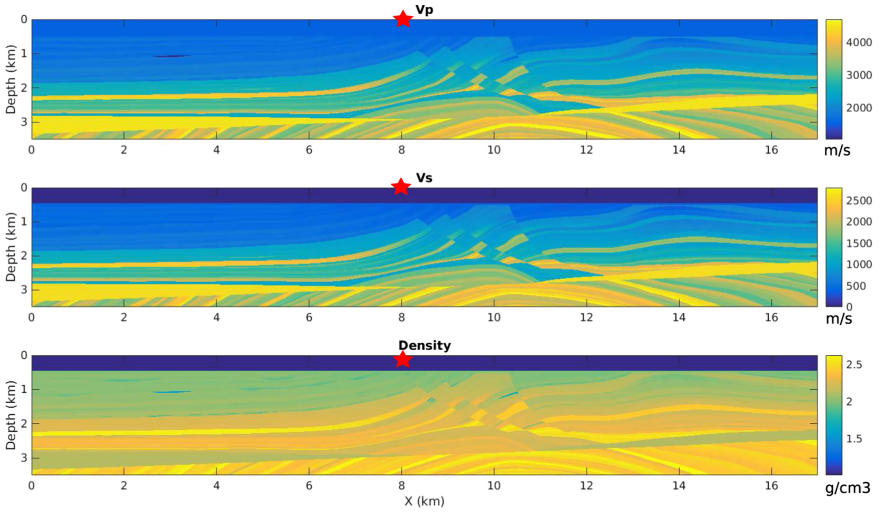


Fig. 3. Marmousi2 elastic velocity model used for synthetic data generation. The marker represents the source position at $x = 8$ km.

The acquisition included 171 sources with the distance between the sources 100 m. We recorded wavefield by 512 2C receivers for each shot with maximal source-receiver offsets equal to 6.4 km. The distance between the receivers was 25 m. Simulations were performed using the fourth-order staggered grid scheme [11]. On average, the simulation time was 5 s per shot if a 5 m grid was used; 40 s per shot for 2.5 m grid; and 4 min for 1.25 m grid, using Nvidia V100 GPU. The example of modelled seismogram on the grid 1.25 m ($X = 9$ km) is presented in Fig. 4.

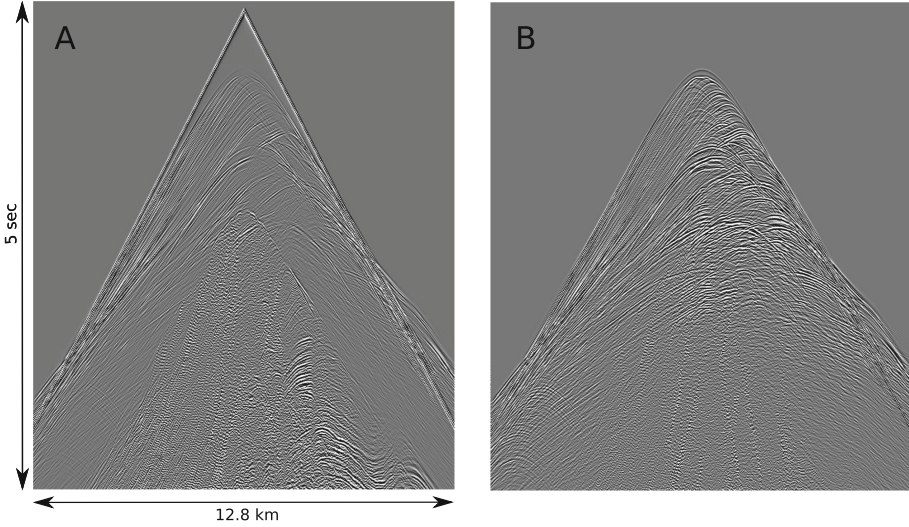


Fig. 4. Synthetic seismograms for shot positioned at $x = 9$ km: horizontal (a) and vertical (b) components calculated on a numerical grid with the spatial step 1.25 m.

We performed two numerical experiments, and for each experiment, we trained NDM-net. One was designed to map the data simulated using a 2.5 m grid to the exact solution (data acquired on the grid with steps 1.25 m). The other NDM-net was trained to map 5 m-data to the 1.25 m-data. The training was performed on the Nvidia V100 GPU. As a regularization, we used an early stopping technique and interrupted the training when the error on the validation dataset started to grow. In both cases (2.5 m to 1.25 m and 5 m to 1.25 m), the training process took about 30min. The prediction time is about 0.7 sec for one full common shot gather, while one forward modelling using FD technique on a GPU took about 40s on 2.5 m grid and about 5 s on 5 m grid, but 5 min for the finest grid of 1.25 m.

To estimate the quality of DNN prediction, we use the normalized RMS (NRMS) as a measure of datasets similarity. NRMS is a strict sample-by-sample metric used for evaluating repeatability between two datasets in 4D seismic [10]. An acceptable level of NRMS in 4D seismic is about 20–40%. The verification of DNN predictions was performed on a testing dataset that differs from training and validation, i.e. were invisible by DNN during the training process. The NRMS plot calculated trace by trace using a sliding window of 200 ms is presented in Fig. 5. On average, the NRMS between 1.25 m data and 2.5 m-data was 30%. Application of the NDM-net reduced the NRMS down to 14%. The average NRMS between 5-m data and 1.25 m-data was about 59%, and the DNN

managed to construct a prediction with the NRMS of 33%. So one may conclude that in both cases, NDM-net were able to reduce NRMS up to the acceptable level. To illustrate the effect of the NDM-net data enhancement, we provide the plots (see Figs. 6, 7) of a single seismic trace computed using different grids and then improved by the NDM-net.

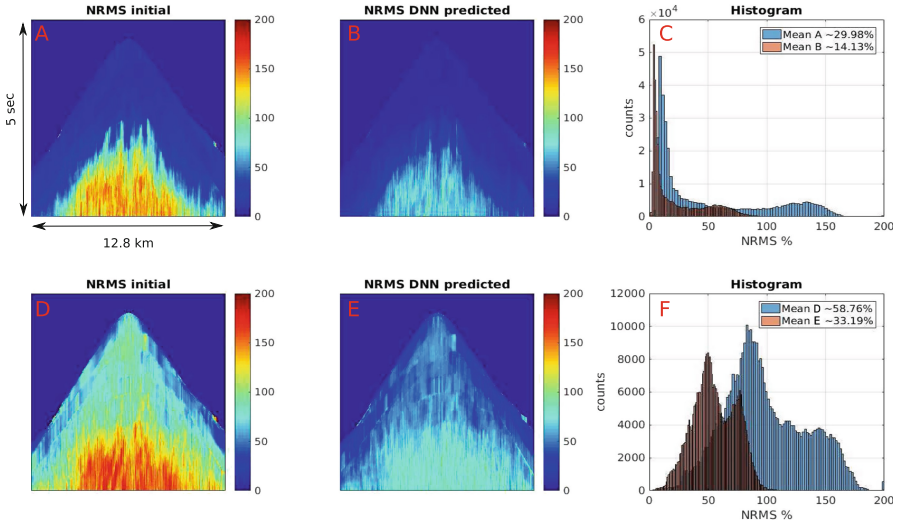


Fig. 5. NRMS plot calculated between seismograms computed on a numerical grid with the spatial steps 1.25 m and 2.5 m (a), 1.25 m and DNN predicted data using 2.5 m data as input (b), 1.25 m and 5 m (d), 1.25 m and DNN predicted data using 5 m data as input, and the corresponding histograms (c,f).

4.2 Model with Vertical Intrusion

The second set of experiments was done for a model with vertical high-contrast intrusions causing lateral heterogeneity as presented in Fig. 8. The size of the entire model was 220 km by 2.6 km. The acquisition included 1901 sources with the distance between the sources 100 m. We recorded wavefield by 512 receivers for each shot with maximal source-receiver offsets equal to 6.4 km. The distance between the receivers was 25 m. In this research, we simulated the wavefield without the surface waves by using a perfectly matched layer for $x < 0$ [4]. The source wavelet was the Ricker pulse with a central frequency 30 Hz.

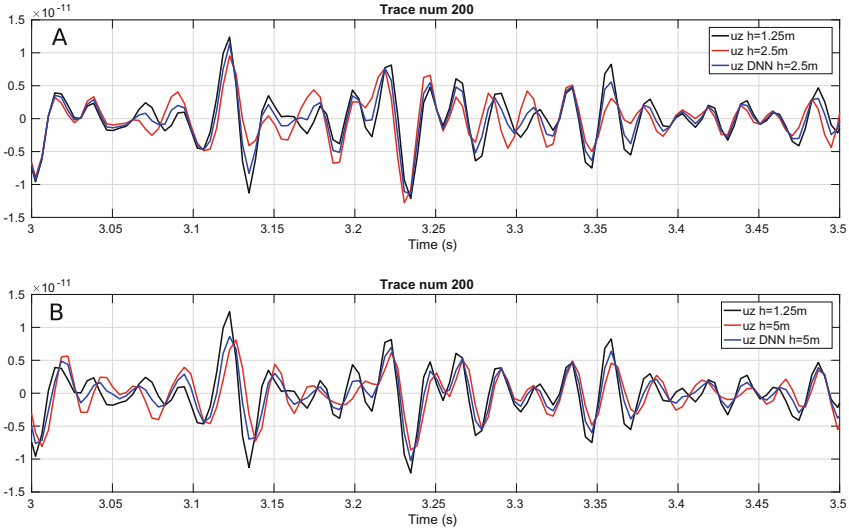


Fig. 6. Seismic traces at different positions and its DNN-predictions for the case 2.5 m-data (a) and 5 m-data (b). Black plot – vertical component on the fine grid, red plot – input data for DNN prediction and blue plot – DNN-predicted data.

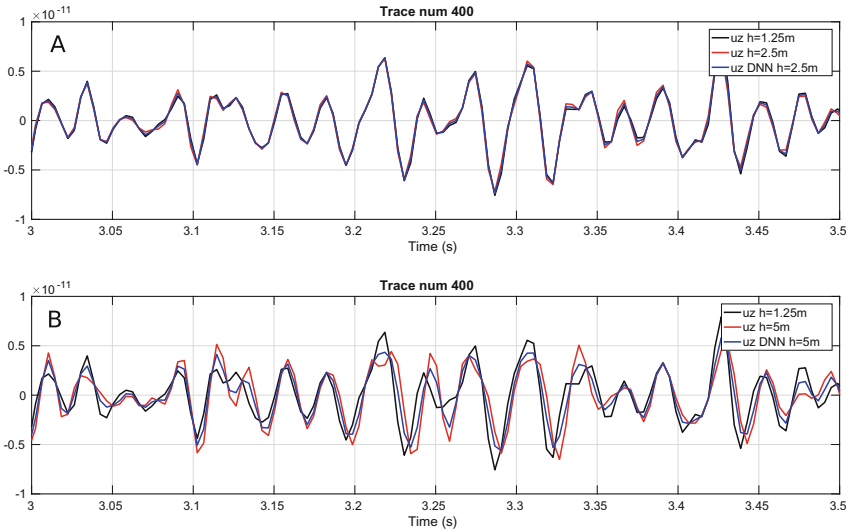


Fig. 7. Seismic traces at different positions and its DNN-predictions for the case 2.5 m-data (a) and 5 m-data (b). Black plot – vertical component on the fine grid, red plot – input data for DNN prediction and blue plot – DNN-predicted data.

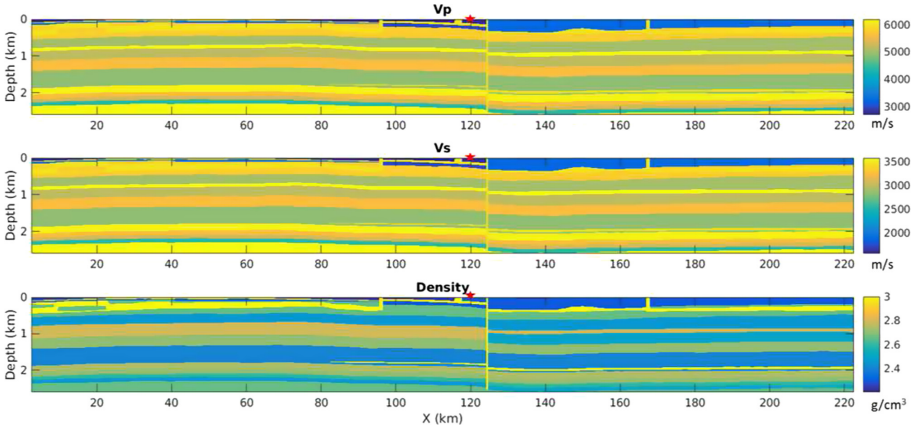


Fig. 8. Elastic velocity model used for synthetic data generation. The marker represents the source position at $x = 120$ km.

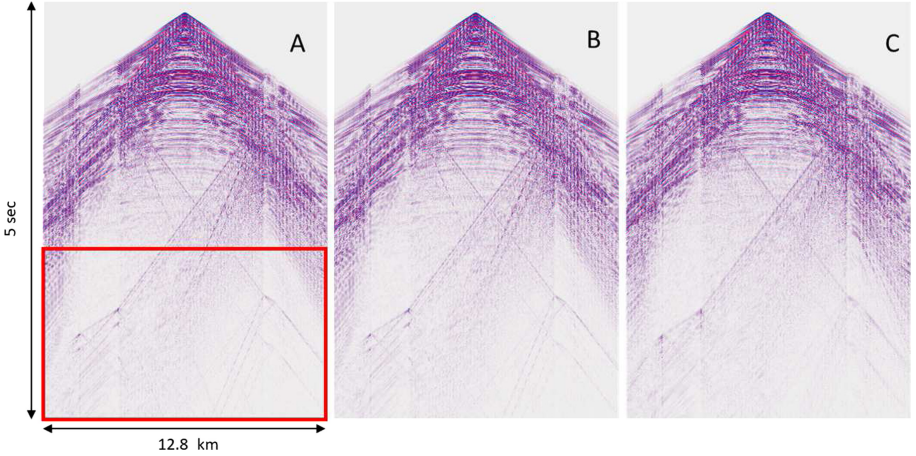


Fig. 9. Synthetic seismograms for shot positioned at $x = 120$ km: vertical component calculated on a numerical grid with the spatial steps 2.5 m (a), 5 m (b) and 10 m (c). (Color figure online)

Originally, the model was provided on a grid with the steps 50 m in horizontal and 5 m in a vertical direction. We computed three datasets using the fourth-order staggered grid scheme [11]. We considered the solution acquired at the grid with steps of 2.5 m as the accurate one, whereas two others generated using grids with 5 m and 10 m spatial steps are polluted. We provide examples of the seismograms in Fig. 9.

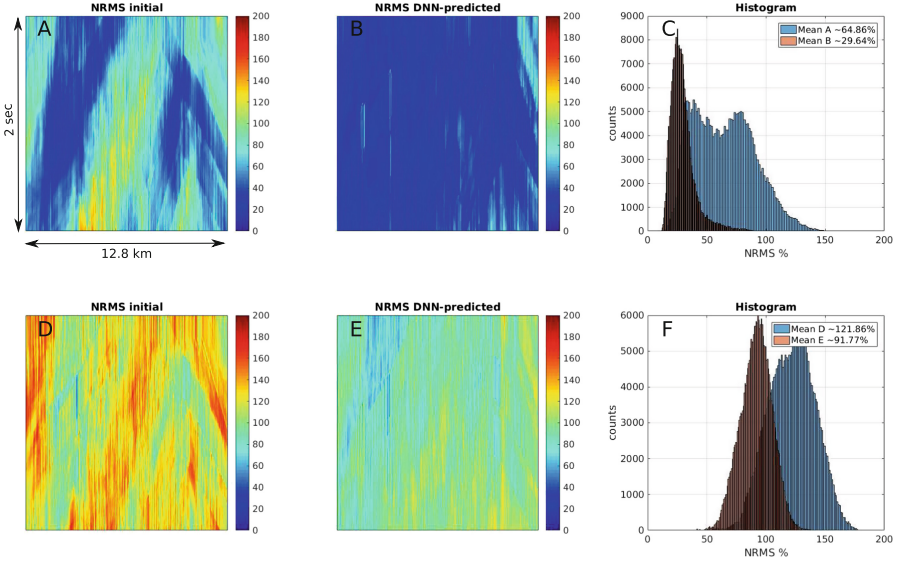


Fig. 10. NRMS plot calculated between seismograms computed on a numerical grid with the spatial steps 2.5 m and 5 m (a), 2.5 m and DNN predicted data using 5 m data as input (b), 2.5 m and 10 m (d), 2.5 m and DNN predicted data using 10 m data as input, and the corresponding histograms (c,f). NRMS were calculated in the area designated by a red rectangle on Fig.2 (vertical component at the time from 3 s to 5 s including all receiver positions)

As in the previous example, we trained two NDM-nets for two synthetic datasets. One was designed to map the data simulated using a 5 m grid to the exact solution (data acquired on the grid with steps 2.5 m). The other NDM-net was trained to map 10 m-data to the 2.5 m-data. In both cases (5 m to 2.5 m and 10 m to 2.5 m), the training process took about 40 min. The prediction time is about 0.7 sec for one full common shot gather, while one forward modelling using FD technique on a GPU took about 40 s on 2.5 m grid and about 5 s on 5 m grid. Since the main error accumulates in the late arrivals, we calculate NRMS for the time range from 3 s to 5 s (red rectangle on the Fig. 9). The corresponding NRMS plot is presented in Fig. 10. On average, the NRMS between 2.5 m data and 5 m-data was 65%. Application of the NDM-net reduced the NRMS down to 30%. The average NRMS between 10-m data and 2.5 m-data was about 120%, which means that the 10-m data are extremely far from the true solution. As a result, the DNN managed to reduce NRMS up to the 90% level. The effect of the NDM-net data enhancement is illustrated in the Figs. 11, 12.

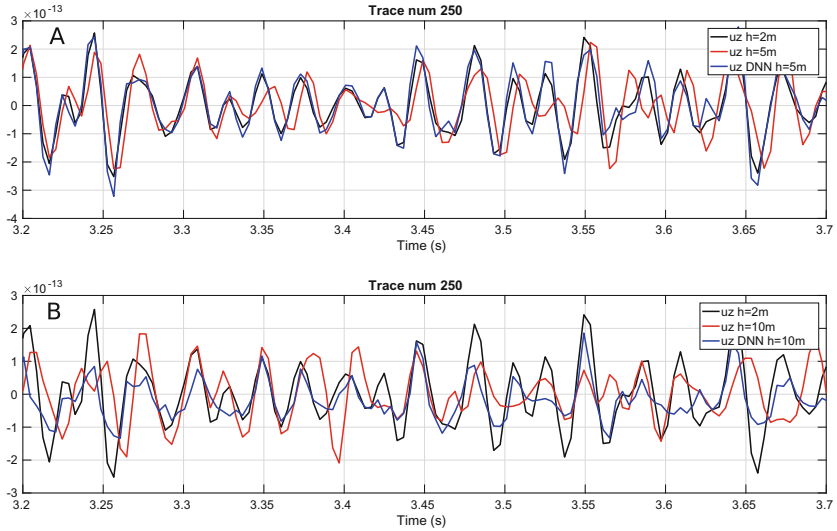


Fig. 11. Seismic traces at different positions and its DNN-predictions for the case 5 m-data (a) and 10 m-data (b). Black plot – vertical component on the fine grid, red plot – input data for DNN prediction and blue plot – DNN-predicted data.

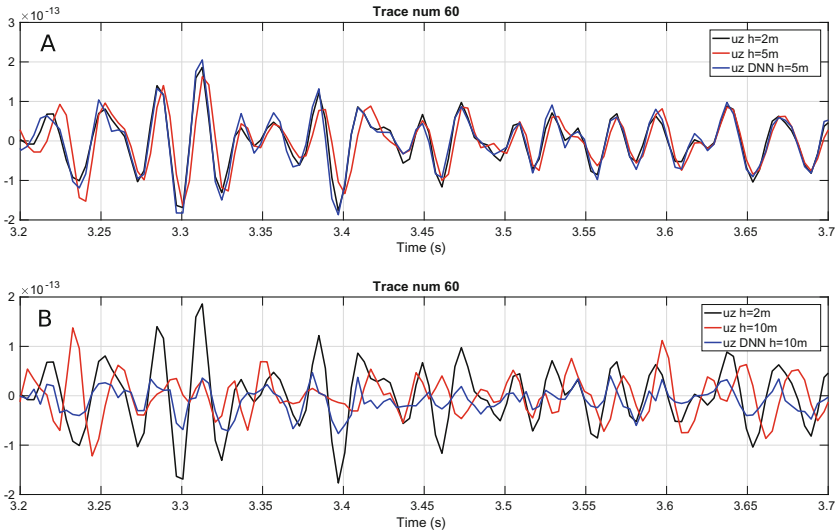


Fig. 12. Seismic traces at different positions and its DNN-predictions for the case 5 m-data (a) and 10 m-data (b). Black plot – vertical component on the fine grid, red plot – input data for DNN prediction and blue plot – DNN-predicted data.

5 Conclusions

We present an original approach to numerical simulation of seismic wavefields. The method combines conventional seismic modelling based on the finite differences with the consequent correction of the data by the DNN-based algorithm called the NDM-net. First, we generate a training dataset simulating wavefields corresponding to at most 10% of the positions of the sources using fine enough spatial discretization (up to 20 points per minimal wavelength - ppw). Second, the full dataset is generated using a coarse mesh with no more than 3–5 ppw. Note that in the 2D case, simulation of the solution using 5 ppw is 64 times faster than that with 20 ppw. Third, the NDM-net is trained to reduce the numerical error in the coarse-grid solution. Then the NDM-net is applied to correct the entire dataset. The presented results demonstrate the ability of the NDM-net to make a high-quality seismic data prediction using the synthetics generated on a coarse grid. In particular, the application of the NDM-net reduced the computational time to simulate the full dataset of 171 common shot gathers for the Marmousi2 model from 684 min to 112 min.

References

1. Baldassari, C., Barucq, H., Calandra, H., Diaz, J.: Numerical performances of a hybrid local-time stepping strategy applied to the reverse time migration. *Geophys. Prospect.* **59**(5), 907–919 (2011). <https://doi.org/10.1111/j.1365-2478.2011.00975.x>
2. Chen, G., Song, L., Liu, L.: 3D numerical simulation of elastic wave propagation in discrete fracture network rocks. *Pure Appl. Geophys.* **176**(12), 5377–5390 (2019)
3. Cohen, G. (ed.): *Methodes numeriques d'ordre eleve pour les ondes en regime transitoire*. INRIA (1994). in French
4. 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)
5. Gadylshin, K., Silvestrov, I., Bakulin, A.: Inpainting of local wavefront attributes using artificial intelligence for enhancement of massive 3-D prestack seismic data. *Geophys. J. Int.* **223**, 1888–1898 (2020)
6. Guo, X., Li, W., Iorio, F.: Convolutional neural networks for steady flow approximation. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, San Francisco, CA, USA, pp. 481–490 (2016). <https://doi.org/10.1145/2939672.2939738>
7. Hicks, G.: Arbitrary source and receiver positioning in finite-difference schemes using kaiser windowed sinc functions. *Geophysics* **67**(1), 156–165 (2002)
8. 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)
9. Koene, E., Robertsson, J.: Removing numerical dispersion artifacts from reverse time migration and full-waveform inversion, pp. 4143–4147 (2017)
10. Kragh, E., Christie, P.: Seismic repeatability, normalized rms, and predictability. *Lead. Edge* **21**(7), 640–647 (2002)
11. Levander, A.R.: Fourth-order finite-difference p-sv seismograms. *Geophysics* **53**(11), 1425–1436 (1988)

12. 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)
13. 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)
14. Liu, Y.: Optimal staggered-grid finite-difference schemes based on least-squares for wave equation modelling. *Geophys. J. Int.* **197**(2), 1033–1047 (2014)
15. Martin, G.S., Wiley, R., Marfurt, K.J.: Marmousi2: an elastic upgrade for marmousi. *Lead. Edge* **25**(2), 156–166 (2006)
16. 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)
17. Moseley, B., Nissen-Meyer, T., Markham, A.: Deep learning for fast simulation of seismic waves in complex media. *Solid Earth* **11**, 1527–1549 (2020)
18. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.I.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
19. Shokin, Y., Yanenko, N.: *Method of Differential Approximation. Application to Gas Dynamics*. Nauka, Novosibirsk (1985). in Russian
20. Virieux, J.: P-sv wave propagation in heterogeneous media: velocity-stress finite-difference method. *Geophysics* **51**(4), 889–901 (1986)
21. 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>
22. 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)
23. Xu, Z., et al.: Time-dispersion filter for finite-difference modeling and reverse time migration, pp. 4448–4452 (2017)
24. Zhu, J., Ren, M., Liao, Z.: Wave propagation and diffraction through non-persistent rock joints: an analytical and numerical study. *Int. J. Rock Mech. Mining Sci.* **132**, 104362 (2020)