3D NDM-Net Training Dataset Construction

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

Sobolev Institute of Mathematics SB RAS, 4 Koptug ave., Novosibirsk 630090, Russia
gondyulea@gmail.com

Abstract. One of the important stages in training neural networks is the construction of a training dataset, especially when it is necessary to minimise its size in order to achieve efficiency. Using the example of a neural network that approximates the fine-grid solution from a coarse-grid solution of a seismic modelling problem, we use a clustering algorithm to form a representative training dataset. In this case, the metric used is the distance between sources. This achieves a solution accuracy that is 10% higher than when using other training dataset construction scenarios.

1 Introduction

Seismic modelling is getting a common tool to study peculiarities of wave propagation in media with complex rheology, including anisotropic media [16,19,30], viscoelastic media [1,36], poroelastic media [2,3,21]. Also, seismic modelling results may be used to design acquisition system in seismic exploration and seismic monitoring, especially in the case of CO_2 sequestration where plum monitoring is crucial for preventing CO_2 leakage [28,32]. However, seismic modelling is computationally intense and time-consuming, especially if applied as a part of seismic processing algorithms [34,35].

A common tool to simulate seismic waves in the Earth's interior is the grid-based numerical methods such as finite differences [16,23,30,33], finite elements [39], spectral elements [31,40], discontinuous Galerkin [5,18,29], pseudo-spectral methods [6,27] and others. However, as it follows from the general theory of numerical methods, to compute an accurate enough solution, one needs to use fine enough grids, because the numerical error is supposed to vanish if the grid step tends to zero. However, the size of the discretized problem increased dramatically if the grid step was reduced. Indeed, if the grid step gets one-half of the original step, the number of grid points increases in 16 times.

The classical way to improve the accuracy of the numerical method is the increase the formal order of approximation (order of the local basis). This approach allows reducing the numerical error even for coarse discretizations [14,17,25]. However, it has two drawbacks. First, increase of the formal order of approximation reduces the Courant number making the stability condition stricter. Second, stencil length (for finite differences) or size of the local mass and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16761, pp. 70–83, 2027.
https://doi.org/10.1007/978-3-032-30527-5_5

stiffness matrices for the discontinuous Galerkin method also increases requiring a higher number of floating point operations per grid node. The other way to improve the accuracy of, at least, finite differences is the use of dispersion suppression schemes [4,11,20]. This approach over-performs the high-order schemes if coarse enough grids are used. However, it has just the same drawbacks as mentioned above; i.e., the high number of flops per grid node and strict stability condition.

The two described approaches allow to reduce the numerical dispersion (numerical error) due to spatial discretization. It is done by the increase of the degrees of freedom when the spatial operator is approximating. However, the approximation of the temporal derivative is typically restricted by the second-order accurate operator. A modern approach to reduce the temporal dispersion in pre- and post-processing of the input signal and output data, respectively. The approach was firstly proposed in [12], and then developed in [22,24]. In combination with the spatial dispersion suppression schemes, this approach may produce accurate enough results on relatively coarse grids with as few as three grid points per minimal grid points per wavelength [24]. However, if the number of flops is considered, the approach is still computationally intense.

A completely different and rapidly developing approach to mitigate numerical error is based on the data post-processing using Deep Learning methods. In particular, the Numerical Dispersion Mitigation Network (NDM-net) was recently suggested in [9]. The idea of the approach is to compute the entire seismic dataset (wavefileds corresponding to all source positions which is about 10^5) using conventional low-order finite difference schemes. After that, a small number of seismograms can be generated using high-order schemes or a fine enough grid to form a training dataset. Then a network is trained and applied to map low-accuracy solution to the accurate one. It was shown in [9] that the application of the NDM-net allows accelerating seismic modelling by a factor of five in 2D in comparison with conventional approaches. However, according to the previous study, the efficiency of the NDM-net strongly depends on the training dataset [8,10].

One of the essential points is to prepare training data and reduce its size without losing the quality of training. Various methods are used to optimize the training set. For example, [37] suggests a method for evaluating the quality of a training sample using the output of the loss function in the learning process, and [38] developed a modified method of RUS (Random Under Sampling) using a random forest classifier. The authors of the article [26] use t-distributed stochastic neighbor embedding (CAE t-SNE) before training a neural network to analyze the similarities between training and target data sets.

In this article, we look at a strategy for construction a dataset using a Voronoi diagram and cluster analysis. Thus, the representativeness of the training dataset increases, thereby increasing the quality of the neural network application.

2 Seismic Modelling

One of the important stages is seismic modelling and simulation of synthetic seismograms. In this section, we will present the system of equations and the numerical method used in the study. In this case, we simulate seismograms for all sources only when using a coarse computational grid for fast calculations. To construct the training dataset, we use a fine computational grid and only a small portion of the sources (up to 20% of the total number of sources).

The solvable system is a system of differential equations of elasticity theory in terms of velocity and components of the stress tensor:

$$\begin{gathered}
\rho\frac{\partial u_1}{\partial t} = \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{12}}{\partial x_2} + \frac{\partial \sigma_{13}}{\partial x_3} + f_1(t)\delta(\boldsymbol{x} - \boldsymbol{x}_s), \\
\rho\frac{\partial u_2}{\partial t} = \frac{\partial \sigma_{12}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + \frac{\partial \sigma_{23}}{\partial x_3} + f_2(t)\delta(\boldsymbol{x} - \boldsymbol{x}_s), \\
\rho\frac{\partial u_3}{\partial t} = \frac{\partial \sigma_{13}}{\partial x_1} + \frac{\partial \sigma_{23}}{\partial x_2} + \frac{\partial \sigma_{33}}{\partial x_3} + f_3(t)\delta(\boldsymbol{x} - \boldsymbol{x}_s), \\
\frac{\partial \sigma_{11}}{\partial t} = (\lambda + 2\mu)\frac{\partial u_1}{\partial x_1} + \lambda\frac{\partial u_2}{\partial x_2} + \lambda\frac{\partial u_3}{\partial x_3}, \\
\frac{\partial \sigma_{22}}{\partial t} = \lambda\frac{\partial u_1}{\partial x_1} + (\lambda + 2\mu)\frac{\partial u_2}{\partial x_2} + \lambda\frac{\partial u_3}{\partial x_3}, \\
\frac{\partial \sigma_{33}}{\partial t} = \lambda\frac{\partial u_1}{\partial x_1} + \lambda\frac{\partial u_2}{\partial x_2} + (\lambda + 2\mu)\frac{\partial u_3}{\partial x_3}, \\
\frac{\partial \sigma_{23}}{\partial t} = \mu\left(\frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2}\right), \\
\frac{\partial \sigma_{13}}{\partial t} = \mu\left(\frac{\partial u_3}{\partial x_1} + \frac{\partial u_1}{\partial x_3}\right), \\
\frac{\partial \sigma_{12}}{\partial t} = \mu\left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\right),
\end{gathered} \tag{1}$$

where ρ is the mass density, λ and μ are the Lame parameters, $\boldsymbol{u} = (u_1, u_2, u_3)^T$, is the velocity vector, σ_{ij} are the components of the stress tensor, $\boldsymbol{f}(t)$ is the vector of the external forces, δ is the Dirac delta function, $\boldsymbol{x_s}$ are the coordinates of the source position.

The system is solved in half-space $D = [x_s - L_x, x_s + L_x] \times [0, Z]$, where L_x is the maximal offset. It should be noted that this system is solved for a set of right-hand sides, i.e. source positions. It is precisely because of this that the task is resource-intensive.

We used staggered-grid finite differences [13,15,33,34] to solve system (1), which is described in detail in [9]. We will denote $\boldsymbol{u_h}(t, x_0, x_s^k)$ as a numerical solution of system (1) using a grid step $h = h_x = h_z$.

For our study, we will use synthetic seismograms recorded at the surface at zero depth by receivers. The acquisition system will be described in the section Numerical Experiments. Let us denote the seismogram corresponding to the source position x_s as $S(t, x_s)$.

3 Neural Network

To solve system 1 with the lowest computational cost, we propose using a neural network to suppress numerical dispersion. A slightly modified U-net architecture is used as the neural network architecture. This architecture allows not only to solve the image-to-image problem, but also to train relatively quickly.

The network consists of four consecutive convolution operations and a nonlinear LeakyReLU activation function, and four reverse convolution operations and a nonlinear ReLU activation function (Fig. 1). The kernel of both types of convolutions has a dimension of 2 by 2. Concatenation are also used to preserve spatial information.

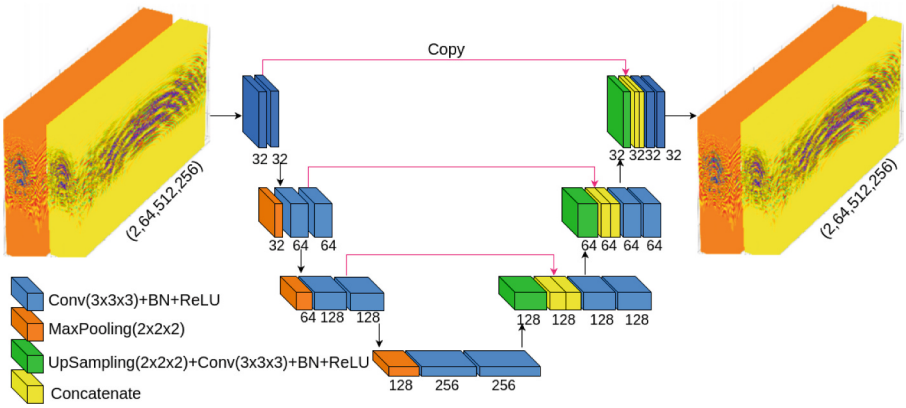

Fig. 1. 3D NDM-net architecture. Input and output data are two channels for the real and imaginary parts of the FFT. Red arrows represent copying skip connections, while black arrows represent operations specified by the colored boxes, such as Conv (convolution), BN (batch-normalization), ReLU (rectified linear unit) and MaxPooling (max pooling). (Color figure online)

As a result of using a neural network, seismograms are obtained, designated as $\mathcal{N}[S_c(t, x_s^j)]$, where $\mathcal{N} : S_c(t, x_s^j) \rightarrow S_f(t, x_s^j)$ is the map operator, $S_c(t, x_s^j)$ is the coarse-grid seismogram for $j - th$ source and $S_f(t, x_s^j)$ is the fine-grid seismogram.

4 Training Dataset Construction Algorithms

This section discusses three methods for construction a training dataset for a neural network that suppresses numerical dispersion in 3D synthetic seismograms.

4.1 Equidistantly Distributed Dataset

One obvious method is to use source positions located at equal distances from each other. This allows us to construct a dataset that is appropriately represented and covers the entire velocity model. Indeed, in 2D seismic modeling one may simply determine the distance between source positions as

$$d_d^{ij} = |\boldsymbol{x}_s^i - \boldsymbol{x}_s^j|,$$

where we assume all sources are placed at the same depth (typically $z = 0$).

In 3D surveys, acquisition typically employs a regular 2D grid of sources and receivers. Sources are positioned on a uniform rectangular lattice with inline spacing d_x and crossline spacing d_y, so $\boldsymbol{x}_s^{i,j} = (i \cdot d_x, j \cdot d_y, 0)$. The Euclidean distance between any two sources becomes

$$d_d^{i_1 j_1, i_2 j_2} = \sqrt{((i_1 - i_2)d_x)^2 + ((j_1 - j_2)d_y)^2},$$

which reduces to the 1D case when $d_x = d_y = d$ and $j_1 = j_2$. The resulting dataset D_d maintains regularity, facilitating efficient interpolation or decimation in processing workflows like NDM-net.

4.2 Randomly Distributed Dataset

The second method considered is to randomly select the source positions. This is the most popular construction method used in deep learning due to its simplicity of implementation. In addition, the construction of such training dataset does not depend on the properties of the data, which affects the representativeness. A uniform distribution is used to create a set of indexes. The seismograms included in this training dataset are denoted as $S(t, x_s^j) \in D_r$.

4.3 Dataset Based on Voronoi Diagram

The last method is to build using the Voronoi diagram and the clustering method. Since we have the distances between all the sources, we can construct a Voronoi diagram for the combined set of positions in terms of the k-mean. The first step of the algorithm is to randomly select the cluster centroids according to a uniform distribution.. Each segment of the Voronoi diagram corresponding to the selected centroids becomes a separate cluster. Next, the position of the centroids is updated in accordance with the minimization of the functional:

$$\sum_{i=1}^{K} \sum_{x \in S_i} (x - \mu_i)^2 \to min,$$

where K is the number of clusters, S_i is the obtained clusters in $i - th$ step, x is the vector of source positions and μ_i is the center of mass of all vectors from cluster S_i.

Finally, the algorithm stops when the relative decrease in the functional becomes less than the set value. To create the training dataset, the final centroids of the clustering algorithm are taken. The dataset constructed using the Voronoi diagram is denoted as D_{cl}.

5 Numerical Experiments

This section discusses the results of using three methods to construct training datasets to calibrate the NDM-net to suppress numerical dispersion in seismograms.

5.1 Training and Test Data

For experiments, we use the acoustic 3D SEG/EAGE Overthrust velocity model, which is shown in Fig. 2. The geological structure is a complex features with a fault. The size of the model is 20 km with a resolution of 25 m. The acquisition system uses an area system with lines of sources that are located 200 m apart, and the distance between the sources is 25 m. Perpendicular to the source lines are receiver lines, which are also located 200 m apart with a distance of 25 m between the receivers. For each source, a set of receivers is fixed with a maximum offsets of 6.4 km. Seismograms are recorded up to 3.2 s with a sampling rate of 2 ms.

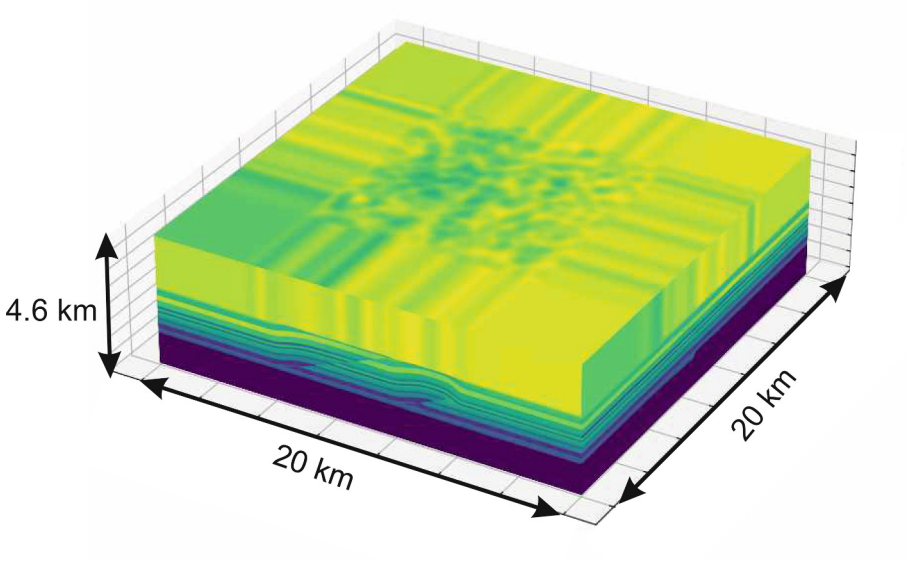

Fig. 2. The 3D SEG/EAGE Overthrust velocity model.

When calculating two sets of seismic data (a fine-resolution solution and a coarse-resolution solution), the Ricker pulse with centered frequency of 25 Hz is employed. To obtain a fine-resolution solution, a computational grid with a step of 2.5 m is used, for a coarse-resolution solution 5 m. The size of the training dataset is 2% of the total number of sources in the acquisition system.

For seismic modeling, as well as neural network training, a local server with 8 Nvidia A100 GPUs with 80 Gb of RAM was used. The duration of computation for each seismogram was 2 min when using a coarse computational grid, and 16 min when using a fine computational grid. The time required to train NDM-net was 20 h for a subset of 4% of the total number of sources. Therefore, the speed-up factor is between 5 and 6.

5.2 Training Scheme

The first step is data preprocessing, namely the mapping of seismograms from the time domain to the frequency domain using the Fast Fourier Transform (FFT). Due to the fact that seismograms are recorded with a fairly dense time grid, the high-frequency part of the signal is noise. Therefore, we are cutting off a significant unnecessary part of the signal. This procedure allows to significantly reduce the size of the data, and hence the RAM used during training. In addition, as shown in the [7], the use of FFT makes it possible to improve the quality of the neural network while suppressing numerical dispersion. Thus, one seismogram as a results of this procedure has a size of $32 \times 512 \times 256 \times 2$, where 32×512 is the number of traces of area receivers, 256 is the number of frequency samples, 2 is the real and imaginary part of the signal, instead of $32 \times 512 \times 1600$, where 1600 is the number of time samples.

The size of the training datasets, regardless of the method of its construction, is 2% of the total number of sources in the acquisition system. In the process of training the NDM-net, the mean-square error between the fine-resolution solution and the generated solution is minimized. Adam (Adaptive Moment Estimation) is used with increments of 10^{-4} to minimize the loss function. Training and validation losses are depicted in Fig. 3. Here, the early stopping and dropout techniques are used to prevent overfitting.

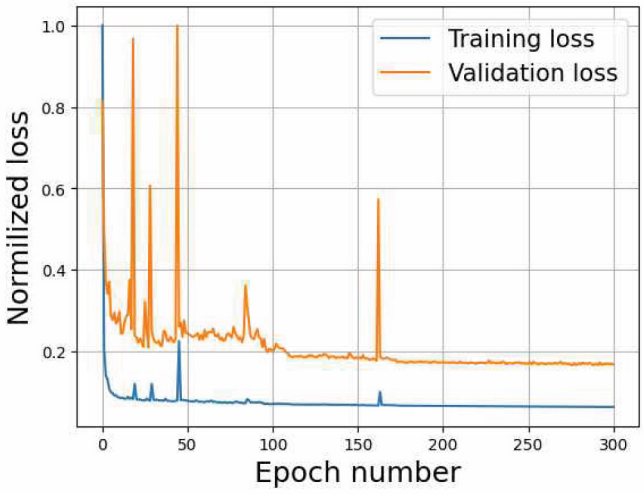

Fig. 3. Training and validation loss of NDM-net.

5.3 Results

This section presents the results of numerical experiments using three training datasets: equidistributed (D_d with source spacing $d = 25$ m), random (D_r), and Voronoi diagram-based (D_{cl} with $K = 200$ clusters initialized from D_d centroids). Figure 4 shows the seismograms obtained after applying three neural networks trained on three different training datasets. Thus, the seismogram on Fig. 4C was obtained using clustering in the formation of training data and has the lowest numerical dispersion compared to other seismograms (Fig. 4A, 4B, 4E). Figure 4 also shows the seismograms as the results of numerical modeling on a fine and coarse grid, as well as their difference. It is noticeable that numerical dispersion is more successfully suppressed when using a training dataset based on the Voronoi diagram.

A relative error is used to measure the accuracy of the results obtained from applying a neural network to seismograms:

$$\varepsilon^j = \varepsilon(S_f^j(t, x_s), \mathcal{N}[S_c^j(t, x_s)]) = \frac{\|S_f(t, x_s^j) - \mathcal{N}[S_c(t, x_s^j)]\|_2}{\|S_f(t, x_s^j)\|_2} \times 100\%,$$

$$\|S(t, x_s^j)\|_2^2 = \int_0^T \int_{-X}^{X} \int_{-Y}^{Y} u^2(x_s^j, t, x, y) dy dx dt,$$

where x and y are offsets along corresponding directions, X and Y are maximal offsets, and t is time.

The relative error for each source in the three source lines is shown in Fig. 5 when using different training datasets. It is noticeable that when using random sampling and equally distributed sampling, the error of using a neural network is almost the same. However, when using a Voronoi-based sample, we get the lowest relative error value for all sources.

We also calculate relative error maps for all the source positions in the acquisition system (Fig. 6). The NDM-net was observed to successfully reduce numerical dispersion by approximately 2-fold in all cases. However, it is recommended to use clustering when creating a training dataset based on information regarding the distance between sources, as this approach yielded the best results. The maps also indicate that numerical error fluctuates significantly in areas with complex structures, while it remains relatively stable in areas with homogeneous structures. Additionally, increasing the representativeness of the training dataset by using of a Voronoi diagram reduced this variation by 10%.

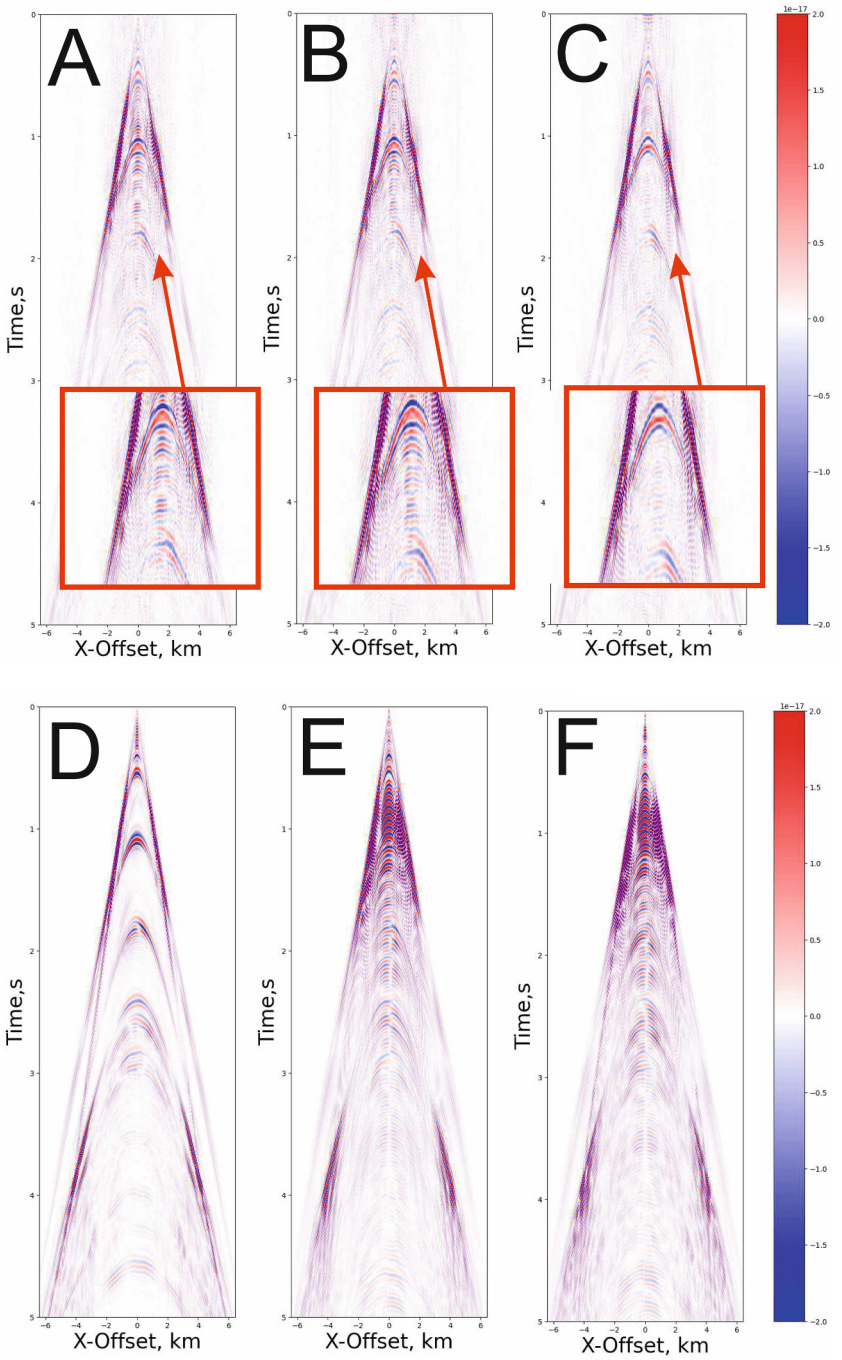

Fig. 4. Seismograms obtained using a neural network trained on a random dataset D_r (A), an equidistributed dataset D_d (B), a dataset using a Voronoi diagram D_{cl} (C) (the main differences are highlighted in red boxes), and obtained using seismic modeling on a fine grid (D), on a coarse grid (E), the difference between a fine-grid solution and a coarse solution (F). All seismograms were obtained for the same source with a position of $X = 2600$ m, $Y = 400$ m for Y-offset $= 0$ m. (Color figure online)

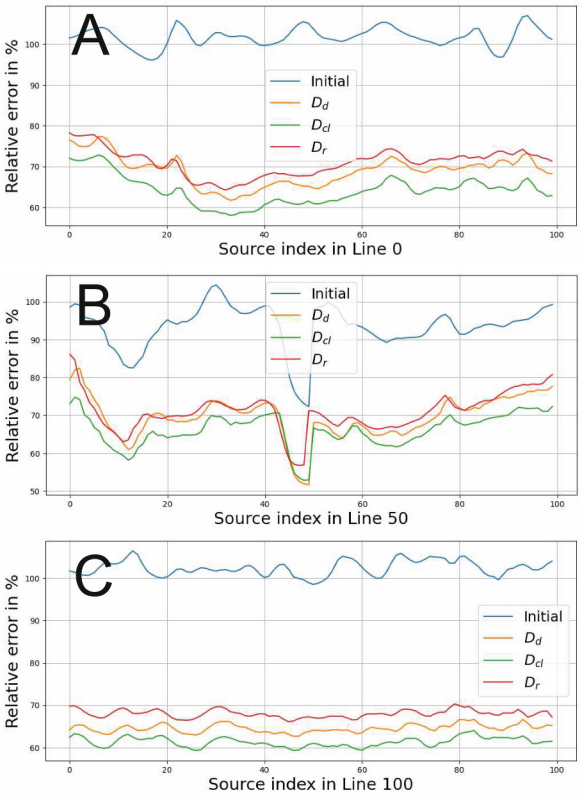

Fig. 5. The relative error between seismogram calculated on a coarse grid and fine grid solution (blue plot), and the relative errors between NDM-net post-processed seismogram and fine grid solution for different training datasets (D_d, D_{cl}, D_r) and lines of acquisition system with position Y-offset = 0 m (A), Y-offset = 10 km (B), Y-offset = 20 km (C). (Color figure online)

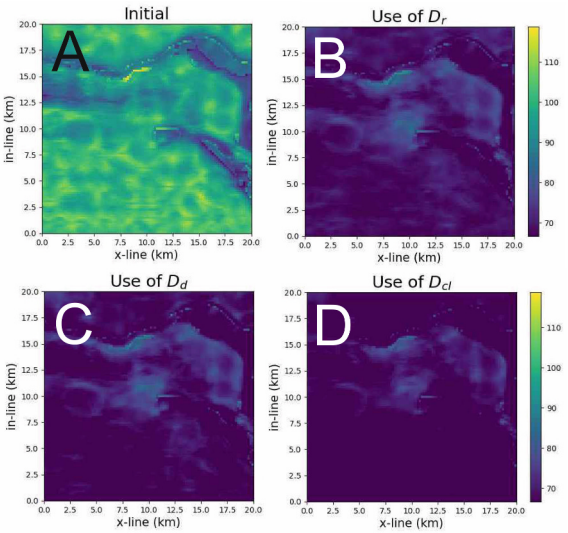

Fig. 6. Maps for the relative error (in %) between solutions on a coarse and a fine computational grid (A), between the fine grid solutions and 3D NDM-net postprocessing results for a training datasets D_r (B), D_d (C), D_{cl} (D).

6 Conclusion

We presented a detailed study of the effect of the training dataset construction on the accuracy of the NDM-net. The NDM-net was designed to reduce the numerical error in seismic modeling, where the elastic wave equation is solved for the multiple right-hand sides. Thus, the solution for all source positions can be computed using a low-order finite-difference scheme on a coarse grid. The training dataset is generated using a fine enough mesh for a small number of source positions. After that, the NDM-net is trained to map the coarse grid solution to the fine grid solution. Generation of the training dataset is the most time-consuming part of the algorithm.

It has been demonstrated that the use of a more representative dataset obtained from the Voronoi diagram and clustering techniques reduces the relative error of applying the NDM-net by 10% compared to a conventional, equally distributed training sample. Additionally, numerical dispersion has been successfully reduced by a factor of two in areas with complex structures, which was not possible using other training sets.

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. Carcione, J.M., Cavallini, F.: A rheological model for anelastic anisotropic media with applications to seismic wave propagation. Geophys. J. Int. **119**, 338–348 (1994)
2. Carcione, J.M., Morency, C., Santos, J.E.: Computational poroelasticity – a review. Geophysics **75**(5), 75A229-75A243 (2010)
3. Caspari, E., et al.: Attenuation mechanisms in fractured fluid-saturated porous rocks: a numerical modelling study. Geophys. Prospect. **67**(4), 935–955 (2019)
4. Chen, J.B.: A 27-point scheme for a 3D frequency-domain scalar wave equation based on an average-derivative method. Geophys. Prospect. **62**(2), 258–277 (2014)
5. 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
6. Fornberg, B.: The pseudospectral method: accurate representation of interfaces in elastic wave calculations. Geophysics **53**(5), 625–637 (1988)
7. 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
8. 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
9. Gadylshin, K., Vishnevsky, D., Gadylshina, K., Lisitsa, V.: Numerical dispersion mitigation neural network for seismic modeling. Geophysics **87**(3), T237–T249 (2022)
10. 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 (2023)
11. He, Z., Zhang, J., Yao, Z.: Determining the optimal coefficients of the explicit finite-difference scheme using the remez exchange algorithm. Geophysics **84**(3), S137–S147 (2019). https://doi.org/10.1190/geo2018-0446.1
12. Koene, E.F.M., Robertsson, J.O.A., Broggini, F., Andersson, F.: Eliminating time dispersion from seismic wave modeling. Geophys. J. Int. **213**(1), 169–180 (2017)
13. Kostin, V., Lisitsa, V., Reshetova, G., Tcheverda, V.: Local time-space mesh refinement for simulation of elastic wave propagation in multi-scale media. J. Comput. Phys. **281**, 669–689 (2015)
14. Levander, A.R.: Fourth-order finite-difference P-SV seismograms. Geophysics **53**(11), 1425–1436 (1988)
15. Levander, A.: Fourth-order finite-difference P-S. Geophysics **53**(11), 1425–1436 (1988). https://doi.org/10.1190/1.1442422
16. 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
17. 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

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)
19. 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)
20. Liu, Y., Sen, M.K.: Acoustic VTI modeling with a time-space domain dispersion-relation-based finite-difference scheme. Geophysics **75**(3), A11–A17 (2010)
21. Masson, Y.J., Pride, S.R.: Finite-difference modeling of biot's poroelastic equations across all frequencies. Geophysics **75**(2), N33–N41 (2010)
22. Mittet, R.: Second-order time integration of the wave equation with dispersion correction procedures. Geophysics **84**(4), T221–T235 (2019)
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. O'Reilly, O., Lundquist, T., Dunham, E.M., Nordstrom, J.: Energy stable and high-order-accurate finite difference methods on staggered grids. J. Comput. Phys. **346**, 572–589 (2017). https://doi.org/10.1016/j.jcp.2017.06.030
26. Park, J., Choi, J., Jee Seol, S., Byun, J., Kim, Y.: A method for adequate selection of training data sets to reconstruct seismic data using a convolutional u-net. Geophysics **86**, V375–V388 (2021)
27. 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)
28. Prasad, M., Glubokovskikh, S., Daley, T., Oduwole, S., Harbert, W.: Co2 messes with rock physics. Lead. Edge **40**(6), 424–432 (2021)
29. Puente, J., Kaser, M., Dumbser, M., Igel, H.: An arbitrary high-order discontinuous galerkin method for elastic waves on unstructured meshes iv. anisotropy. Geophys. J. Int. **169**(3), 1210–1228 (2007). https://doi.org/10.1111/j.1365-246X.2007.03381.x
30. 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)
31. Tromp, J., Komatitsch, D., Liu, Q.: Spectral-element and adjoint methods in seismology. Commun. Comput. Phys. **3**(1), 1–32 (2008)
32. Verliac, M., Calvez, J.L.: Microseismic monitoring for reliable co2 injection and storage – geophysical modeling challenges and opportunities. Lead. Edge **40**(6), 418–423 (2021)
33. Virieux, J.: P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method. Geophysics **51**(4), 889–901 (1986)
34. 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
35. Virieux, J., et al.: Seismic wave modeling for seismic imaging. Lead. Edge **28**(5), 538–544 (2009)
36. Vishnevsky, D.M., Solovyev, S.A., Lisitsa, V.V.: Numerical simulation of wave propagation in 3d elastic media with viscoelastic formations. Lobachevskii J. Math. **41**(8), 1603–1614 (2020)

37. Wang, R., Meng, Y., Zhang, W., Li, Z., Hu, F., Meng, L.: Remote sensing semantic segregation for water information extraction: optimization of samples via training error performance. IEEE Access (2019). https://doi.org/10.1109/ACCESS.2019.2894099
38. Yan, S., Zhang, L., Liu, D.: An empirical study on optimization of training dataset in harmfulness prediction of code clone using ensemble feature selection model. In: 018 5th International Conference on Information and Communication Technologies for Disaster Management, pp. 1–4 (2018). https://doi.org/10.1109/ICT-DM.2018.8636375
39. Zhang, J., Verschuur, D.J.: Elastic wave propagation in heterogeneous anisotropic media using the lumped finite-element method. Geophysics **67**(2), 625–638 (2002)
40. Zhang, Y., Fisser, J., Gerritsma, M.: A hybrid mimetic spectral element method for three-dimensional linear elasticity problems. J. Comput. Phys. **433**, 110179 (2021). https://doi.org/10.1016/j.jcp.2021.110179