

*Е. А. Гондюл<sup>1\*</sup>, В. В. Лисица<sup>1</sup>*

## **Обучающая выборка в задаче подавления численной дисперсии в сейсмических данных при использовании нейронных сетей**

<sup>1</sup> Институт нефтегазовой геологии и геофизики СО РАН, г. Новосибирск, Российская Федерация

**Аннотация.** В данной работе используются методы машинного обучения для подавления численной дисперсии, которая возникает при моделировании сейсмических полей методом конечных разностей с использованием крупной сетки. С целью уменьшения времени расчёта предлагается аппроксимировать нелинейный оператор перехода из данных с численной ошибкой в данные с меньшей ошибкой с помощью нейронных сетей. Приводятся три способа формирования обучающей выборки и результаты обучения.

**Ключевые слова:** численная дисперсия, нейронные сети, сейсмика

*E. A. Gondyul<sup>1\*</sup>, V. V. Lisitsa<sup>1</sup>*

## **The training set in the dispersion mitigation problem in seismic data using neural networks**

<sup>1</sup> Institute of Petroleum Geology and Geophysics SB RAS, Novosibirsk, Russia

**Abstract.** In this paper, machine learning methods are used to suppress the numerical dispersion that occurs when modeling seismic fields by the finite difference method using a coarse grid. In order to reduce the calculation time, it is proposed to approximate the non-linear transition operator from data with a numerical error to data with a smaller error using neural networks. Also, the question of how to form a training sample based on the properties of the seismic data themselves, the velocity model, and more are being investigated. Three ways of creating a training sample and learning outcomes are given.

**Keywords:** numerical dispersion, neural network, seismology

### ***Введение***

Численное моделирование сейсмических полей сопровождается численной ошибкой, называемой численной дисперсией, в случае, когда используется грубая расчётная сетка. Численная дисперсия возникает из-за использования производных нечётного порядка в аппроксимации пространственных производных. Для подавления этой численной дисперсии было предложено несколько алгоритмов предварительной и последующей обработки [1]. Однако эти алгоритмы требовательны к вычислительным ресурсам и могут обрабатывать только дисперсию, вызванную дискретизацией временных производных. Также можно использовать конечно-разностные схемы высокого порядка, разрывный метод Галёркина [2, 3] и методы спектральных элементов [4]. Но такие методы не обязательно приводят к снижению вычислительной сложности. По мере увеличения порядка аппроксимации количество операций и используемая память также уве-

личиваются. В нашем исследовании предлагается аппроксимировать нелинейный оператор перехода от данных с большой числовой ошибкой к данным с меньшей с помощью нейронных сетей. Этот подход позволяет в значительной степени подавить дисперсию и артефакты с меньшим потреблением памяти и времени, чем другие методы постобработки.

Способность нейронной сети обучаться и выдавать точные результаты на тестовых данных в значительной степени зависит от данных, на которых она обучается. Одним из важных аспектов построения нейронных сетей является создания репрезентативного обучающего набора данных. Хорошо спроектированный набор данных может обеспечить высокую степень обобщения. Недостаточно репрезентативный набор данных приводит к неправильному принятию решений и снижению эффективности алгоритмов обучения. Целью данной работы является проектирование различных видов оптимальных выборок для нейронной сети NDM-net.

### *Нейронная сеть*

За основу архитектуры NDM-net взята одна из простых нейронных сетей – U-net. Такая полносвязная нейронная сеть хорошо себя показала не только в сегментации изображений [5], но и генерации данных, аппроксимации нелинейных операторов. U-net архитектура состоит из кодера, уменьшающего пространственное разрешение входных изображений с помощью ряда свёрточных слоёв для получения высокоуровневых признаков, и декодера, восстанавливающего исходное пространственное разрешение. Глубина нейронной сети задавалась исходя из размерности входных данных и их количества, а также целеполагания задачи. Так NDM-net состоит из 16 свёрточных слоёв: 8 слоёв повышающей дискретизации (кодер) и 8 слоёв понижающей (декодер). Слои состоят из оператора свёртки (convolution operator), пакетной нормализации (batch normalization). Слой свёртки производит сумму поэлементного произведения фрагмента входных данных и ядра свёртки. Весовые коэффициенты ядра свёртки определяются в ходе обучения нейронной сети. Используется LeakyReLU (Leaky Rectified Linear Unit) в качестве функции активации для первых восьми свёрточных слоёв с отрицательным коэффициентом наклона 0,2 и ReLu для декодера.

В процессе обучения определяются весовые коэффициента ядра свёртки. Преимуществом данной нейронной сети является использование skip connections для решения проблемы деградации точности, именно они помогают обновлять веса в случае, когда градиент приближается к нулю. Для избежания переобучения применяется метод ранней остановки.

В качестве входных данных используется набор сейсмограмм, рассчитанных на грубой сетке, и обозначается здесь как  $X = \{x\}_{i=1}^N$  где  $N$  – количество сейсмограмм во всём наборе данных. Выходные данные нейронной сети обозначаются как  $Y = \{y\}_{i=1}^N$  и представляют собой изображения, имитирующие сейсмограммы, рассчитанные на мелкой сетке.

В ходе обучения NDM-net минимизируется следующая функция потерь:

$$L = \sum_i \|G(x_i) - y_i\|_1,$$

где  $G: x_i \rightarrow y_i$  – оператор перехода, который аппроксимируется нейронной сетью.

Сама нейронная сеть обновляет свои параметры по цепному правилу:

$$\theta_l = \theta_l - \gamma \left( \frac{\partial l}{\partial \theta_l} \right),$$

где  $\theta_l$  – параметры нейронной сети на слое  $l$ ;  $\gamma$  – коэффициент скорости обучения, значение которого меняется в ходе обучения.

### **Описание выходных данных**

Рассмотрим упругую ВР модель (рис. 1), которая имеет контрастное соляное тело и плавно меняющую свои скоростные значения правую часть. Размер модели – 67 км на 12 км. Набор сейсмических данных состоит из 2696 источников с расстоянием между ними в 25 м. Волновое поле регистрируется 2401 приёмниками для каждого источника с выносом 15 км. Расстояние между приёмниками – 12,5 м. Источник описывается вейвлетом Рикера с центральной частотой 30 Гц.

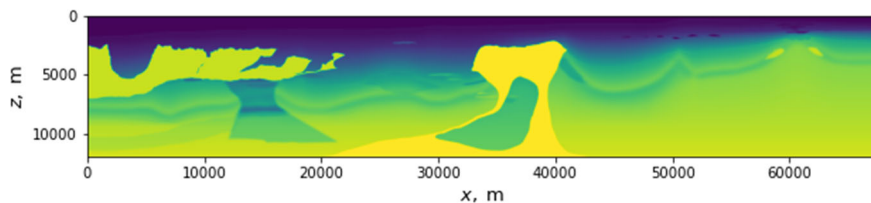


Рис. 1. Упругая ВР модель, используемая для расчёта набора сейсмограмм

В качестве входных данных для нейронной сети использовались два набора сейсмограмм, рассчитанных на грубой сетке с шагом  $h_x = 6$  м и на мелкой сетке с шагом  $h_x = 3$  м. Синтетические наборы были вычислены с помощью схем на сдвинутых сетках четвертого порядка [6]. Сейсмические данные были обрезаны таким образом, чтобы прямая волна не входила в сейсмические изображения и имели размер  $512 \times 512$ .

### **Равноудалённая выборка**

Самый простой способ построить обучающую выборку – это брать равноудалённые друг от друга источники:

$$J_k^{eq} = 1, k, 2k, \dots, q(k)k,$$

где  $k \in \mathbb{N}, k > 1$  и  $q(k) = \frac{N}{k}$ ,  $N$  – количество сейсмограмм в наборе данных.

Так, например, образуются обучающие выборки  $D_{10}^{eq}, D_5^{eq}, D_{2,6}^{eq}$ , соответствующие 10, 5 и 2,6% от всего набора данных соответственно. Были рассчитаны средние расстояния  $\bar{R}(Y^{ML}, Y)$  между сейсмограммами, сгенерированными нейронной сетью, и сейсмограммами, смоделированными на мелкой сетке, с использованием  $L_2$  метрики (табл. 1.). Расстояние между исходными наборами  $R(X, Y) = 0.469$ .

Таблица 1

Результаты обучения NDM-net на равноудалённой выборке

Набор данных	Количество источников	$\bar{R}(Y^{ML}, Y)$
$D_{11}^{eq}$	300	0.122
$D_{10}^{eq}$	270	0.140
$D_{8,3}^{eq}$	225	0.138
$D_{6,3}^{eq}$	169	0.190
$D_5^{eq}$	135	0.156
$D_{2,9}^{eq}$	78	0.194
$D_2^{eq}$	54	0.264

### Формирование выборки на основе Хаусдорфова расстояния между сейсмограммами

Предполагается, что сейсмограммы, соответствующие ближайшим друг к другу источникам, схожи [7]. Чтобы проиллюстрировать это, рассчитывается матрица расстояний для всего набора данных, смоделированных на грубой сетке (рис. 2). Расстояние между источниками выражается следующим образом:

$$nMSE(u(s_j), u(s_{j+1})) = \frac{0.5 * \|u(s_j) - u(s_{j+1})\|}{\|u(s_j)\|_2 + \|u(s_{j+1})\|_2},$$

где  $u(s_j)$  – решение системы уравнений динамической теории упругости для источника  $s_j$ .

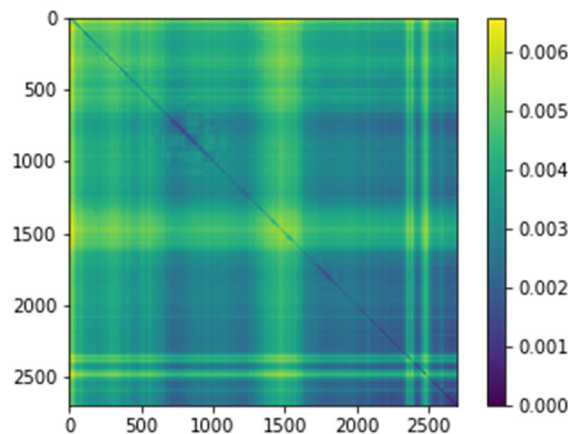


Рис. 2. Матрица расстояний для сейсмограмм

Значения матрицы равны нулю в случае, когда сейсмограммы эквивалентны друг другу. Вблизи главной диагонали расстояния между источниками растут достаточно резко. Однако, зависимость расстояния между сейсмограммой и всеми остальными источниками (рис. 3) показывает, что  $nMSE$  растёт почти линейно вблизи источника и выходит на предельное значение, после которого уже практически не зависит от расстояния между позициями источника.

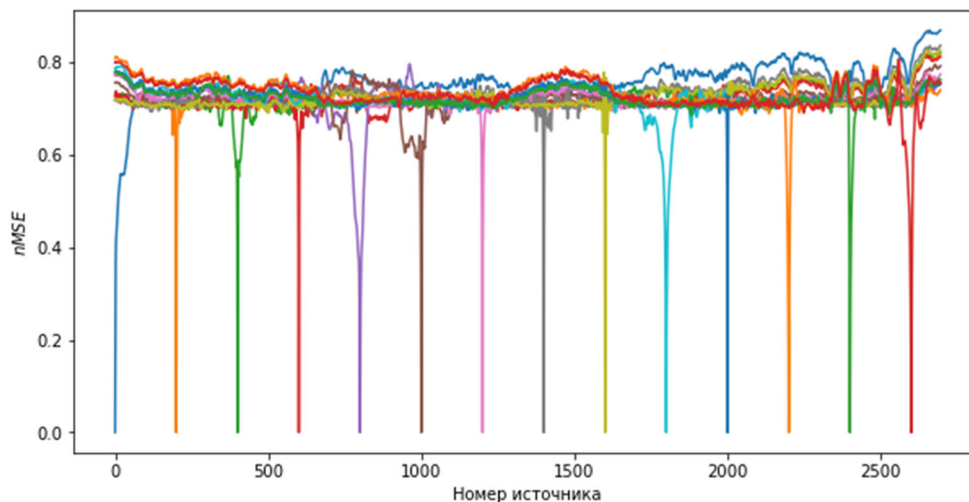


Рис. 3. Расстояния для 14 источников

Предлагается построить оптимизированную обучающую выборку путем решения min-max задачи между обучающей выборкой и всем набором данных:

$$\max_{k \in 1, \dots, N} \min_{j \in J_t} nMSE(u(s_k), u(s_j)) \leq Q,$$

где  $J_t$  – набор индексов источников обучающего набора,  $Q$  – требуемый уровень ошибки.

Результаты усреднённой ошибки между  $Y^{ML}$  и  $Y$  приведены в табл. 2.

Таблица 2

Результаты обучения NDM-net на оптимальной выборке  $D^S$

Набор данных	Количество источников	$\bar{R}(Y^{ML}, Y)$
$D_{6,8}^S$	185	0.143
$D_{5,5}^S$	148	0.170
$D_{3,9}^S$	105	0.177
$D_{2,4}^S$	68	0.206
$D_2^S$	55	0.291

По сравнению с результатами, полученными с использованием равноудаленных источников, оптимизированный набор данных обеспечивает в среднем

меньшую погрешность для заданного количества обучающих данных. Однако, следует заметить, что качество обучения зависит от распределения источников. Такой подход позволяет сократить время обучения нейронной сети за счет уменьшения количества сейсмограмм в обучающей выборке.

### **Формирование выборки на основе свойств скоростной модели**

Предлагается также построить обучающий набор на основе свойств скоростной модели с использованием Хаусдорфовой метрики. Предполагается, что части модели, которые соответствуют соседним источникам, подобны. Введем множество  $M = M_1, \dots, M_{N_s}$  с моделями, соответствующими положениям источников. Для такого набора моделей была построена матрица расстояний с использованием nMSE (рис. 4).

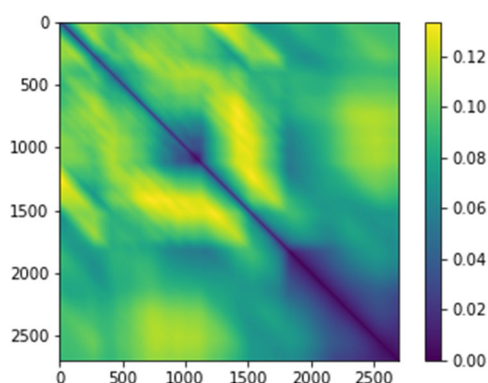


Рис. 4. Матрица расстояний для скоростной модели

Такая матрица показывает, как меняется модель в зависимости от расстояния между источниками. Вблизи главной диагонали расстояние между  $M_i$  и  $M_j$  невелико, так как источники являются соседними. В области плавного изменения скоростной модели расстояние между частями модели растет медленно. Приведем несколько столбцов матрицы расстояний (рис. 5), чтобы показать, как меняется nMSE с расстоянием между источниками.

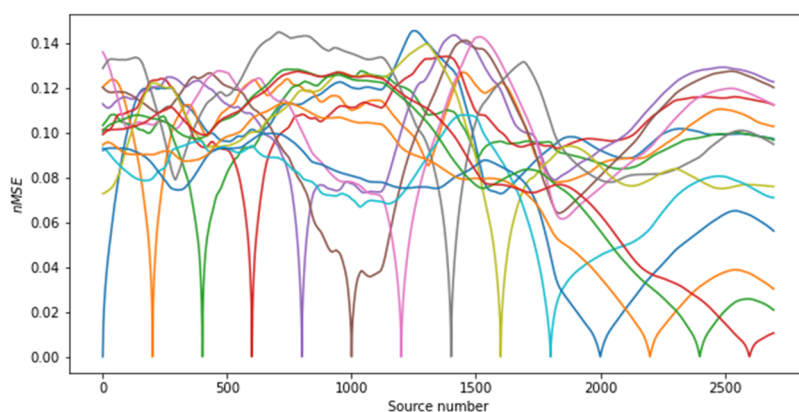


Рис. 5. Расстояние nMSE для 14 скоростных моделей

Также, как и в предыдущем способе построения обучающей выборки, будем ограничивать Хаусдорфово расстояние между обучающей выборкой и всем набором данных:

$$\max_{k \in 1, \dots, N} \min_{j \in J_t} nMSE(M_k, M_j) \leq Q.$$

Результаты обучения на обучающих выборках, основанных на свойствах скоростной модели, приведены в табл.3.

*Таблица 3*

Результаты обучения NDM-net на оптимальной выборке  $D^M$

Набор данных	Количество источников	$\bar{R}(Y^{ML}, Y)$
$D_9^M$	245	0.147
$D_8^M$	216	0.147
$D_{6.4}^M$	173	0.145
$D_{3.8}^M$	104	0.190
$D_{2.4}^M$	55	0.240

Из таблицы видно, что данный способ формирования обучающей выборки не даёт значимого преимущества в результатах по сравнению с другими.

### ***Заключение***

Результаты обучения показали, что нейронные сети могут быть использованы для подавления численной дисперсии в сейсмических данных. Была использована NDM-net, которая преобразует сейсмические данные, смоделированные на грубой сетке, в данные, рассчитанные на мелкой. В этом исследовании рассматриваются три подхода к построению обучающего набора данных: метод, основанный на равноудаленном расположении источников, и методы, которые создают оптимизированные выборки на основе свойств скоростной модели, а также сейсмических данных. Последние способы построения обучающего набора данных заключается в ограничении Хаусдорфова расстояния. Показано, что два последних способа позволяют уменьшить размер обучающей выборки, однако, они сильно зависят от распределения источников.

### ***Благодарности***

В.В. Лисица разработал алгоритм построения оптимального набора обучающих данных при поддержке гранта РФФ №22-11-00004., Е.А. Гондюл выполнила численные эксперименты по обучению NDM-net сети в рамках выполнения работ по гос. заданию FWZZ-2022-0022.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Mittet R. Second-order time integration of the wave equation with dispersion correction procedures // *Geophysics*. – 019. – № 84(4). – P. T221—T235.
2. Lisitsa V. Dispersion analysis of discontinuous Galerkin method on triangular mesh for elastic wave equation // *Applied Mathematical Modelling*. – 2016. – № 40(7-8). – P. 5077—5095.
3. Kaser M., Dumbser M., Puente J., Igel H. An arbitrary high-order Discontinuous Galerkin method for elastic waves on unstructured meshes III. Viscoelastic attenuation // *Geophysical Journal International*. – 2007. – № 168(1). – P. 224—242.
4. Pleshkevich A., Vishnevsky D., Lisitsa V., Levchenko V. Parallel Algorithm for One-Way Wave Equation Based Migration for Seismic Imaging // *Supercomputing*. Springer International Publishing. – 2019. – P. 125—135.
5. Ronneberger O, Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. LNCS. Springer. – 2015. – № 9351. – P. 234—241.
6. Levander A. R. Fourth-order finite-difference P-SV seismograms // *Geophysics*. – 1988. – № 53(11). – P. 1425 —1436.
7. Gadylshin K., Lisitsa V., Gadylshina K., Vishnevsky D. Optimization of the Training Dataset for Numerical Dispersion Mitigation Neural Network // *Computational Science and Its Applications – ICCSA 2022 Workshops*. Springer International Publishing. – 2022. – P. 295—309.

© *Е. А. Гондюл, В. В. Лисица, 2023*