



doi 10.26089/NumMet.v25r213

УДК 550.34.01

Оптимизация тренировочного набора данных для подавляющей численную дисперсию нейронной сети NDM-net

Е. А. Гондюл

Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН,
Новосибирск, Российская Федерация

ORCID: 0009-0009-0619-1198, e-mail: gondyulea@ipgg.sbras.ru

В. В. Лисица

Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН,
Новосибирск, Российская Федерация

ORCID: 0000-0003-3544-4878, e-mail: lisitsavv@ipgg.sbras.ru

К. Г. Гадьльшин

Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН,
Новосибирск, Российская Федерация

ORCID: 0000-0001-7247-6911, e-mail: gadylshinkg@ipgg.sbras.ru

Д. М. Вишневский

Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН,
Новосибирск, Российская Федерация

ORCID: 0000-0002-1439-4552, e-mail: vishnevskydm@ipgg.sbras.ru

Аннотация: Предлагается оригинальный способ построения обучающего набора данных для нейронной сети NDM-net (Numerical Dispersion Mitigation neural network), подавляющей численную дисперсию при моделировании сейсмических волновых полей. NDM-net обучается отображать вычисленное на грубой сетке решение системы уравнений динамической теории упругости в рассчитанное на мелкой сетке. Данные сейсмограмм для обучения NDM-net предварительно рассчитываются на мелкой сетке, что является трудоемкой процедурой. Для снижения вычислительных затрат алгоритма время обучения необходимо сокращать без потери точности. В качестве эффективной метрики для генерации обучающего набора данных рассматривается линейная комбинация трех метрик: расстояния между источниками, меры сходства сейсмограмм и меры сходства скоростных моделей. Коэффициенты линейной комбинации определяются с помощью глобального анализа чувствительности.

Ключевые слова: численная дисперсия, сейсмическое моделирование, глубокое обучение.

Благодарности: Гондюл Е.А. провела статистический анализ результатов, Гадьльшин К.Г. разработал архитектуру NDM-net, Лисица В.В. поставил задачу и проанализировал результаты, Вишневский Д.М. выполнил конечно-разностное моделирование волновых полей при поддержке РФФ, грант № 22–11–00004.

Для цитирования: Гондюл Е.А., Лисица В.В., Гадьльшин К.Г., Вишневский Д.М. Оптимизация тренировочного набора данных для подавляющей численную дисперсию нейронной сети NDM-net // Вычислительные методы и программирование. 2024. 25, № 2. 155–174. doi 10.26089/NumMet.v25r213.



Optimization of the training dataset for NDM-net (Numerical Dispersion Mitigation neural network)

Elena A. Gondyul

Trofimuk Institute of Petroleum Geology and Geophysics SB RAS,
Novosibirsk, Russia

ORCID: 0009-0009-0619-1198, e-mail: gondyulea@ipgg.sbras.ru

Vadim V. Lisitsa

Trofimuk Institute of Petroleum Geology and Geophysics SB RAS,
Novosibirsk, Russia

ORCID: 0000-0003-3544-4878, e-mail: lisitsavv@ipgg.sbras.ru

Kirill G. Gadylshin

Trofimuk Institute of Petroleum Geology and Geophysics SB RAS,
Novosibirsk, Russia

ORCID: 0000-0001-7247-6911, e-mail: gadylshinkg@ipgg.sbras.ru

Dmitry M. Vishnevsky

Trofimuk Institute of Petroleum Geology and Geophysics SB RAS,
Novosibirsk, Russia

ORCID: 0000-0002-1439-4552, e-mail: vishnevskydm@ipgg.sbras.ru

Abstract: In this paper we present a new approach to building a training dataset for the NDM-net (Numerical Dispersion Mitigation neural network) neural network, which suppresses numerical dispersion in modeling seismic wave fields. The NDM-net is trained to display the solution of the system of equations of the dynamic theory of elasticity, calculated on a coarse grid, into a solution modeled on a fine grid. However, in order to train an NDM-net, it is necessary to pre-calculate seismograms on a fine grid, which is a time-consuming procedure. To reduce the computational costs of the algorithm, an original approach is proposed that allows to reduce the learning time without loss of accuracy. It is proposed to consider a linear combination of three different parameters: the distance between sources, the similarity of seismograms and the similarity of velocity models as an effective metric for generating a training dataset. The weights of the linear combination are determined using a global sensitivity analysis.

Keywords: numerical dispersion, seismic modelling, deep learning.

Acknowledgements: E. Gondyul performed the statistical analysis of the results and K. Gadylshin designed the NDM-net architecture, V. Lisitsa stated the problem and analyzed the results, D. Vishnevsky performed seismic modelling under the support of RSCF, grant No. 22–11–00004.

For citation: E. A. Gondyul, V. V. Lisitsa, K. G. Gadylshin, and D. M. Vishnevsky, “Optimization of the training dataset for NDM-net (Numerical Dispersion Mitigation neural network),” Numerical Methods and Programming. **25** (2), 155–174 (2024). doi 10.26089/NumMet.v25r213.

1. Введение. Сейсмическое моделирование применяется для изучения распространения волн в слож-
ноустроенных средах: анизотропных [1–3], вязкоупругих [4, 5] и пороупругих [6–8], для проектирования
системы наблюдения и сейсмического мониторинга. В частности, мониторинг необходим для предотвра-
щения утечки CO₂ из коллектора [9, 10]. В то же время моделирование — самая затратная процедура из
этапов алгоритма сейсмической обработки [11, 12].



Для расчета волнового поля в земной коре используются сеточные методы: конечно-разностные [1, 2, 13–15], конечных элементов [16], спектрального элемента [17, 18], разрывный метод Галеркина [19–21], псевдоспектральные методы [22, 23] и др. Как следует из теории численных методов, для вычисления приближенного решения, погрешность которого мала, необходимо использовать мелкую сетку, так как численная погрешность уменьшается вместе с шагом сетки [24]. Тогда размер определенной на сетке задачи резко возрастает: с уменьшением шага в два раза количество узлов увеличивается в 8 раз для двумерного случая и в 16 раз для трехмерного.

Точность численного метода растет с увеличением порядка аппроксимации, таким образом численная ошибка подавляется при грубой дискретизации [25–27]. Однако для высоких порядков аппроксимации число Куранта невелико, т.е. условие устойчивости более строгое, чем для схем низкого порядка. А также увеличивается длина шаблона для конечных разностей или размер локальных матриц массы и жесткости для разрывного метода Галеркина, из-за чего операций с плавающей запятой на узел сетки становится больше.

Для повышения точности конечно-разностного метода также используются схемы подавления дисперсии [28–30]. Для грубых расчетных сеток такие методы больше подходят в сравнении со схемами высокого порядка. Их недостаток — большое количество операций на один узел расчетной сетки и строгое условие устойчивости.

За счет увеличения степеней свободы при аппроксимации пространственного оператора уменьшается пространственная численная дисперсия. Вместе с тем производная по времени аппроксимируется оператором второго порядка. Для подавления временной дисперсии применяется предобработка входного или постобработка выходного сигнала [31–33]. В сочетании со схемами подавления пространственной дисперсии такие подходы обеспечивают достаточную точность решения на относительно грубой сетке, содержащей всего три точки на длину волны. Но из-за большого количества операций вычислительные затраты и в этом случае высоки.

В настоящее время распространяется постобработка сейсмических данных методом глубокого обучения, например алгоритм на основе нейронной сети NDM-net, подавляющей численную дисперсию в сейсмических полях [34, 35]. С использованием конечно-разностной схемы низкого порядка рассчитывается весь набор сейсмических данных, т.е. волновые поля для каждого из источников, число которых достигает порядка 10^5 . Данные для небольшого количества сейсмограмм рассчитываются или с использованием схемы высокого порядка, или на мелкой расчетной сетке, для формирования обучающего набора нейронной сети. Обученная нейронная сеть применяется для сопоставления приближенного решения с численной дисперсией с более точным численным решением. Применение NDM-net позволяет ускорить сейсмическое моделирование в пять раз по сравнению с традиционными подходами в двумерном случае. При этом эффективность NDM-net сильно зависит от качества обучающего набора данных [36–38].

Подготовка обучающих данных и уменьшение их размера без снижения репрезентативности критически важна. Для оптимизации набора обучающих данных используются различные способы, такие как оценка качества обучающей выборки с использованием выходных данных функции потерь в процессе обучения [39] или модифицированный метод RUS (Random Under Sampling) с применением классификатора случайного леса [40]. Существует метод уменьшения размера обучающей выборки, основанный на естественном предположении, что относительная энтропия между целевым набором и набором обучающей выборки сводится к минимуму [41]. Для оценки сходства между обучающим и целевым наборами данных используется t-распределенное стохастическое вложение соседей (CAE t-SNE) [42].

Здесь рассматриваются три варианта построения обучающего набора данных для NDM-net: с учетом расстояния между позициями источников [34], меры сходства сейсмограмм [37] и меры сходства используемых скоростных моделей [36]. Значимость трех метрик оценивается при помощи глобального анализа чувствительности [43]. Обучающий набор данных оптимизируется для достижения наивысшей точности NDM-net при фиксированном числе сейсмограмм в обучающем наборе данных.

Основные принципы сейсмического моделирования и NDM-net излагаются в разделе 2. Различные показатели для построения обучающего набора данных обсуждаются в разделе 3. В разделе 4 приводится статистический анализ наборов данных, построенных с использованием различных метрик, и предлагается оптимальная метрика. Иллюстрирующие эффективность предлагаемого подхода численные эксперименты представлены в разделе 5.

2. Предварительные замечания. Рассмотрим систему уравнений динамической теории упругости в терминах скорости смещения и тензора напряжения в декартовой системе координат:

$$\begin{aligned}
 \rho \frac{\partial u_x}{\partial t} &= \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z}, \\
 \rho \frac{\partial u_z}{\partial t} &= \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z}, \\
 \frac{\partial \sigma_{xx}}{\partial t} &= (\lambda + 2\mu) \frac{\partial u_x}{\partial x} + \lambda \frac{\partial u_z}{\partial z} + \hat{f}_{xx}, \\
 \frac{\partial \sigma_{zz}}{\partial t} &= \lambda \frac{\partial u_x}{\partial x} + (\lambda + 2\mu) \frac{\partial u_z}{\partial z} + \hat{f}_{zz}, \\
 \frac{\partial \sigma_{xz}}{\partial t} &= \mu \frac{\partial u_x}{\partial z} + \mu \frac{\partial u_z}{\partial x} + \hat{f}_{xz},
 \end{aligned} \tag{1}$$

где $\mathbf{u} = (u_x, u_z)^T$ — вектор скорости смещения, $\sigma_{xx}, \sigma_{zz}, \sigma_{xz}$ — компоненты тензора напряжения, $\rho = \rho(x, z)$ — плотность, $\lambda = \lambda(x, z)$, $\mu = \mu(x, z)$ — коэффициенты Ламе, $\hat{f}_{xx} = f_{xx}(t)\delta(x - x_s)\delta(z - z_s)$, $\hat{f}_{zz} = f_{zz}(t)\delta(x - x_s)\delta(z - z_s)$, $\hat{f}_{xz} = f_{xz}(t)\delta(x - x_s)\delta(z - z_s)$ — компоненты тензора сейсмических моментов. Рассматривается источник в виде дельта-функции Дирака, где (x_s, z_s) — координаты источника. Вейвлет $f_{xx}(t), f_{zz}(t), f_{xz}(t)$ ограничивается полосой пропускания.

Система (1) определяется либо во всей плоскости \mathbb{R}^2 , либо в полуплоскости $\mathbb{R}_+^2 = \{z \geq 0\}$. Расчетная область задается как $D = [x_s - L_x, x_s + L_x] \times [0, Z]$, где L_x — максимальное смещение и Z — глубина модели. Таким образом, расчетная область зависит от положения источника. Чтобы модель не включала искусственные отражения, используется идеально согласованный слой PML (Perfectly Matched Layer) [44–46].

Решение системы (1) соответствует записи расположенных согласно системе наблюдения серий приемников (рис. 1). Обычно для наземной системы горизонтальные координаты источников и приемников определяются как $z_s = 0$ и $z_r = 0$, т.е. координаты приемников — $(x_r^j(x_s), 0)$ и источников — $(x_s, 0)$. После замены переменных $x_o^j = x_r^j - x_s$ координаты для всех положений источников изменяются с фиксиро-

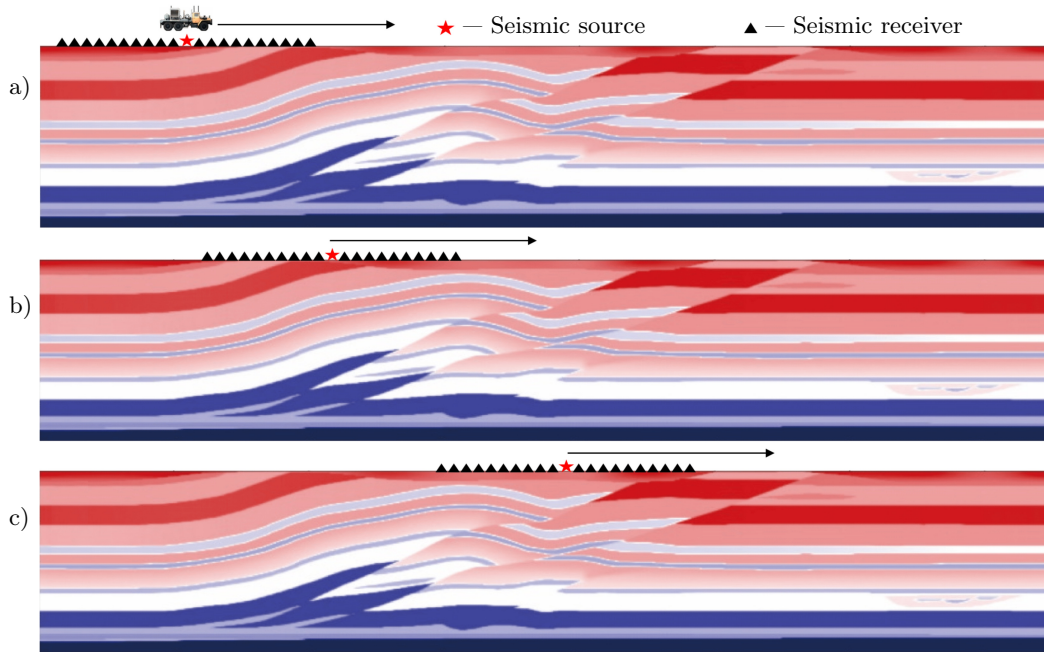


Рис. 1. Схема системы наблюдения. Красной звездой маркируется положение сейсмического источника, черными треугольниками — положения приемников. С увеличением индекса сейсмического источника его горизонтальная координата увеличивается и система сбора данных смещается вправо

Fig. 1. Surveillance system diagram. The position of the seismic source is marked with a red star, and the positions of the receivers are marked with black triangles. As the seismic source index increases, its horizontal coordinate increases and the data acquisition system shifts to the right



ваным интервалом $x_o \in [-L_o, L_o]$, где L_o — максимальное смещение. При этом $L_o \leq L_x$, т.е. локальная часть модели, где система (1) решается для одной правой части, шире максимального смещения.

Полученное решение $\mathbf{u}(t, x_o, x_s)$ называется сейсмограммой, для каждого положения источника каждая составляющая вектора скорости смещения — это функция от двух переменных, времени и смещения.

2.1. Конечно-разностная схема. Для численного решения системы (1) используется конечно-разностная схема на сдвинутых сетках [12, 13, 25, 47]. Для построения схемы вводится сдвинутая сетка: $(u_x)_{i+1/2,j}^n$ определяется в точках $x_{i+1/2} = h_x(i + 1/2)$, $z_j = h_z j$, $t^n = \tau n$; $(u_z)_{i,j+1/2}^n$ в точках $x_i = h_x i$, $z_{j+1/2} = h_z(j + 1/2)$, $t^n = \tau n$; $(\sigma_{xx})_{i,j}^{n+1/2}$ и $(\sigma_{zz})_{i,j}^{n+1/2}$ в точках $x_i = h_x i$, $z_j = h_z j$, $t^{n+1/2} = \tau(n + 1/2)$; $(\sigma_{xz})_{i+1/2,j+1/2}^{+1/2}$ в точках $x_{i+1/2} = h_x(i + 1/2)$, $z_{j+1/2} = h_z(j + 1/2)$, $t^{n+1/2} = \tau(n + 1/2)$; τ — шаг по времени, h_x и h_z — шаги по пространству, индексы i, j, n — целые числа.

Для схемы второго порядка по времени и четвертого порядка по пространству производные аппроксимируются следующими конечно-разностными операторами:

$$D_t[f]_{I,J}^N = \frac{f_{I,J}^{N+1/2} - f_{I,J}^{N-1/2}}{\tau} \approx \frac{\partial f}{\partial t},$$

$$D_x[f]_{I,J}^N = \frac{9}{8} \frac{f_{I+1/2,J}^N - f_{I-1/2,J}^N}{h_x} - \frac{1}{24} \frac{f_{I+3/2,J}^N - f_{I-3/2,J}^N}{h_x} \approx \frac{\partial f}{\partial x},$$

$$D_z[f]_{I,J}^N = \frac{9}{8} \frac{f_{I,J+1/2}^N - f_{I,J-1/2}^N}{h_z} - \frac{1}{24} \frac{f_{I,J+3/2}^N - f_{I,J-3/2}^N}{h_z} \approx \frac{\partial f}{\partial z}.$$

С учетом выражений для всех операторов аппроксимирующая уравнение (1) конечно-разностная схема записывается следующим образом:

$$\rho D_t[u_x]_{i+1/2,j}^{n+1/2} = D_x[\sigma_{xx}]_{i+1/2,j}^{n+1/2} + D_z[\sigma_{xz}]_{i+1/2,j}^{n+1/2},$$

$$\rho D_t[u_z]_{i,j+1/2}^{n+1/2} = D_x[\sigma_{xz}]_{i,j+1/2}^{n+1/2} + D_z[\sigma_{zz}]_{i,j+1/2}^{n+1/2},$$

$$D_t[\sigma_{xx}]_{i,j}^n = (\lambda + 2\mu) D_x[u_x]_{i,j}^n + \lambda D_z[u_z]_{i,j}^n + (\hat{f}_{xx})_{i,j}^n,$$

$$D_t[\sigma_{zz}]_{i,j}^n = \lambda D_x[u_x]_{i,j}^n + (\lambda + 2\mu) D_z[u_z]_{i,j}^n + (\hat{f}_{zz})_{i,j}^n,$$

$$D_t[\sigma_{xz}]_{i+1/2,j+1/2}^n = \mu D_x[u_z]_{i+1/2,j+1/2}^n + \mu D_z[u_x]_{i+1/2,j+1/2}^n + (\hat{f}_{xz})_{i+1/2,j+1/2}^n.$$

Коэффициенты $\lambda(x, z)$, $\mu(x, z)$ и $\rho(x, z)$ изменяются по методу баланса [14, 48, 49–51].

С помощью дисперсионного анализа оценивается расхождение в фазовой скорости решения плоской волны [2, 27, 52]. В частности, для рассматриваемой схемы четвертого порядка по пространству и второго порядка по времени фазовая скорость численного решения равна

$$c^{fd} = \pm \frac{N_w c}{\alpha \pi} \arcsin \left(\alpha \sqrt{\hat{k}_1^2 + \hat{k}_2^2} \right),$$

$$\hat{k}_1 = \frac{9}{8} \sin \left(\frac{\pi \cos(\beta)}{N_w} \right) - \frac{1}{24} \sin \left(\frac{3\pi \cos(\beta)}{N_w} \right), \quad \hat{k}_2 = \frac{9}{8} \sin \left(\frac{\pi \sin(\beta)}{N_w} \right) - \frac{1}{24} \sin \left(\frac{3\pi \sin(\beta)}{N_w} \right),$$

где N_w — число точек на длину волны, α — число Куранта, β — угол, определяющий направление распространения, и c — фазовая скорость сейсмической волны. Фазовая скорость c^{fd} зависит от частоты сигнала, поэтому форма сигнала изменяется во времени, и расхождение распространения волны увеличивается, что трудно подавить математическими методами.

2.2. Нейронная сеть NDM-net. Чтобы скорректировать численное решение и подавить численную дисперсию сигнала, используется искусственная нейронная сеть NDM-net [34]. Подход основан на особенностях сейсмического моделирования, при выполнении которого решения вычисляются для нескольких правых частей, т.е. для нескольких положений источников. Весь набор данных быстро рассчитывается на грубой расчетной сетке, чтобы получить неточное решение. В то же время небольшое количество задач решается с использованием достаточно мелкой сетки, чтобы получить более точное решение. После этого

нейронная сеть обучается сопоставлять данные с грубой сетки с данными с мелкой сетки и в итоге применяется для уменьшения численной дисперсии во всем наборе сейсмических данных. Формально алгоритм выглядит следующим образом.

1. Для всех положений источников x_s^k , $k = 1, \dots, N_s$, где N_s — количество источников, рассчитать сейсмограммы $\mathbf{u}_{h_2}(t, x_o, x_s^k)$, используя грубую сетку с характерным шагом сетки h_2 .
2. Сформировать набор положений источников, соответствующих сейсмограммам, которые будут входить в тренировочную выборку, т.е. сформировать набор индексов $D_t \subset \{1, \dots, N_s\}$.
3. Рассчитать сейсмограммы $\mathbf{u}_{h_1}(t, x_o, x_s^k)$, $k \in D_t$, с использованием мелкой расчетной сетки с шагом дискретизации $h_1 < h_2$.
4. Обучить нейронную сеть $G : \mathbf{u}_{h_2}(t, x_o, x_s^k) \rightarrow \mathbf{u}_G(t, x_o, x_s^k)$ так, чтобы для всех $k \in D_t$ выполнялось следующее условие:

$$\|G[\mathbf{u}_{h_2}(t, x_o, x_s^k)] - \mathbf{u}_{h_1}(t, x_o, x_s^k)\| \ll \|\mathbf{u}_{h_2}(t, x_o, x_s^k) - \mathbf{u}_{h_1}(t, x_o, x_s^k)\|,$$

где G — оператор перехода из данных с грубой сетки в данные с мелкой сетки.

5. Применить обученную нейронную сеть на весь набор сейсмограмм:

$$\mathbf{u}_G(t, x_o, x_s^k) = G[\mathbf{u}_{h_2}(t, x_o, x_s^k)], \quad k = 1, \dots, N_s.$$

NDM-net — это модификация архитектуры U-net [53], которая содержит кодер-декодер с пропускными соединениями. В блоке кодера с использованием свертки входные данные конвертируются в карту признаков. Декодер структурно похож на кодер: с использованием обратной свертки карта признаков преобразуется в изображение. Повышающие и понижающие дискретизацию слои включают слои свертки с ядром размером $(4 \times 4 \times 4)$, функцию активации $ReLU(\cdot) = \max(0, \cdot)$ для кодера и $LeakyReLU(\cdot) = \max(0, \cdot) + c \min(0, \cdot)$ с отрицательным коэффициентом наклона $c = -0.2$ для декодера. Для ускорения алгоритма применяется мини-пакетный стохастический градиентный спуск (SGD). Для оценки количества оптимальных эпох используется метод ранней остановки: функция потерь проверяется на валидационной выборке, размер которой составляет 10% от обучающей. Параметры обучения приведены в табл. 1.

В процессе обучения параметры нейронной сети θ оптимизируются минимизацией функции потерь:

$$L(\Theta) = \mathbb{E}_{k \in D_t} \left[\left\| \mathbf{u}_{h_1}(t, x_o, x_s^k) - G[\mathbf{u}_{h_2}(t, x_o, x_s^k), \Theta] \right\|_1 \right],$$

где $\Theta = \{W, \mathbf{b}\}$ — параметры сети, включающие матрицу весов W и вектор смещений \mathbf{b} , $\mathbb{E}_{k \in D_t}$ — среднее по тренировочной выборке, $\|\cdot\|_1$ — норма L_1 .

3. Способы формирования обучающей выборки. От качества обучающего набора зависит эффективность работы нейронной сети. Репрезентативные выборки из большого набора данных размером N_s строятся с учетом параметров скоростной модели и сейсмических данных. Выделяются три типа параметров задачи сейсмического моделирования, влияющих на репрезентативность выборки. Прежде всего необходимо учитывать геометрию системы сбора данных, а именно координаты источников. Во-вторых, так как для анализа доступен весь набор вычисленных на грубой сетке сейсмограмм, возможно их прямое сравнение при помощи метрики на пространстве сейсмических данных. В-третьих, как упоминалось выше, каждый расчет выполняется для части скоростной модели, располагающейся под источником, т.е. скоростные модели для разных сейсмограмм различаются, а значит, мера схожести скоростных моделей — это один из характерных параметров. Далее приводится формальное описание рассматриваемых параметров.

Таблица 1. Гиперпараметры и параметры обучения

Table 1. Hyperparameters and training parameters

Parameter	Value
Learning rate	from 0.01 to 0.0001 during epochs
Momentum parameters	$\beta_1 = 0.5, \beta_2 = 0.999$
Batch size	10
The number of epochs	500
Optimization algorithm	Adaptive Momentum (Adam)
Optimality criterion	Mean Absolute Error (MAE)



3.1. Расстояние между источниками. Первый и очевидный способ сформировать обучающий набор — использовать геометрию системы наблюдения. Расстояние между положениями источников определяется как

$$d_d^{ij} = d_d(x_s^i, x_s^j) = |x_s^i - x_s^j|.$$

Пусть все источники располагаются на одинаковой глубине. Так как используется стандартная система наблюдения, то

$$d_d^{ij} = |i - j|d,$$

где d — расстояние между двумя соседними источниками.

3.2. Расстояние между сейсмограммами. Как сказано выше, сейсмограмма для фиксированного положения x_s^i источника — это функция $\mathbf{u}(t, x_o, x_s^i)$ от двух переменных t и x_o . Тогда вторая метрика вводится на основе L_2 -нормы следующим образом:

$$d_s^{ij} = d_s(\mathbf{u}(t, x_o, x_s^i), \mathbf{u}(t, x_o, x_s^j)) = 2 \frac{\|\mathbf{u}(t, x_o, x_s^i) - \mathbf{u}(t, x_o, x_s^j)\|_2}{\|\mathbf{u}(t, x_o, x_s^i)\|_2 + \|\mathbf{u}(t, x_o, x_s^j)\|_2},$$

где

$$\|\mathbf{u}(t, x_o, x_s^i)\|_2^2 = \sum_{n=1}^{N_t} \sum_{m=1}^{N_o} [u_x^2(t_n, x_o^m, x_s^i) + u_z^2(t_n, x_o^m, x_s^i)],$$

здесь N_t — количество отсчетов по времени и N_o — количество трасс в сейсмограмме, $\|\cdot\|_2$ — норма L_2 .

3.3. Расстояние на пространстве моделей. Третья метрика задает расстояние между скоростными моделями. При сейсмическом моделировании для фиксированного положения источника волновое поле рассчитывается в подобласти, для которой горизонтальная координата центра совпадает с горизонтальной координатой источника. Для каждого расчета рассматривается модель $M(x, z, x_s^i) \subset M(x, z)$, где $M(x, z)$ — вся скоростная модель.

В соответствии с этим расстояние между моделями вводится следующим образом:

$$d_m^{ij} = d_m(M(x, z, x_s^i), M(x, z, x_s^j)) = 2 \frac{\|M(x, z, x_s^i) - M(x, z, x_s^j)\|_2}{\|M(x, z, x_s^i)\|_2 + \|M(x, z, x_s^j)\|_2},$$

где

$$\|M(x, z, x_s^i)\|_2^2 = \int_0^Z \int_{x_s^i - L_x}^{x_s^i + L_x} [v_p^2(x, z) + v_s^2(x, z)] dx dz.$$

Скорость продольной волны v_p и скорость поперечной волны v_s — параметры скоростной модели, которые зависят от коэффициентов Ламе как $v_p = \sqrt{\frac{\lambda + 2\mu}{\rho}}$ и $v_s = \sqrt{\frac{\mu}{\rho}}$.

3.4. Хаусдорфова метрика. Введены метрики на трех пространствах: геометрическое расстояние между источниками, L_2 -расстояние между сейсмограммами и L_2 -расстояние между скоростными моделями. Расстояние между объектом и набором объектов для любой из введенных метрик определяется следующим образом:

$$D_l(q^j, B) = \min_{q^i \in B} d_l(q^j, q^i),$$

где q^j и q^i — либо положения источника, либо сейсмограммы, либо скоростные модели и B — набор рассматриваемых объектов, $d_l, l \in \{d, s, m\}$ — метрика в соответствующем пространстве.

Хаусдорфово расстояние между двумя множествами:

$$\delta_l(A, B) = \max \left\{ \max_{q^j \in A} D_l(q^j, B), \max_{q^i \in B} D_l(q^i, A) \right\},$$

где A и B — множества в рассматриваемом пространстве. Так как обучающий набор данных — это подмножество всего набора сейсмограмм, то расстояние Хаусдорфа упрощается следующим образом:

$$\delta_l(B_0 \subset B, B) = \max_{q^i \in B} D_l(q^i, B_0) = \max_{q^i \in B} \min_{q^j \in B_0} d_l(q^i, q^j).$$

Этой формулой задается расстояние между всем набором данных и обучающим набором.

4. Статистический анализ. Для оценки зависимости качества обучающей выборки от каждой из трех введенных метрик проводится статистический анализ, а именно глобальный анализ чувствительности [54, 55]. Для статистического анализа требуются параметры скоростных моделей и соответствующих систем наблюдений.

4.1. Скоростные модели.

§ 4.1.1. Модель ВР. Для валидации алгоритмов обработки сейсмических данных часто используется скоростная модель ВР (рис. 2). Модель включает контрастное соляное тело в левой и центральной части (выделено пунктиром), изменение скорости в правой части происходит плавно. Скорость варьируется от 1400 до 4800 м/с. Размер модели — 67 км в горизонтальном направлении и 14 км в вертикальном. Система наблюдения состоит из 2696 источников, располагающихся на расстоянии 25 м друг от друга. Для каждого источника записывается волновое поле 2401 приемником с выбросом в 15 км. Расстояние между приемниками — 12.5 м. Источник — вейвлет Рикера с центральной частотой 30 Гц.

Рассчитываются два набора сейсмограмм: с использованием мелкой расчетной сетки с пространственным шагом $h = h_x = h_z = 3$ м и грубой сетки с шагом $h = h_x = h_z = 6$ м. Сейсмограммы обрезаются до размера 512×512 , что соответствует выбросу $L_o = 3200$ м и временному интервалу в 4.608 с (диапазон от 2.370 до 6.978 с), шаг по времени — 9 мс. Временной интервал выбирается таким образом, чтобы не учитывать прямую волну, которая не дает информации о недрах Земли, но имеет амплитуду, превышающую амплитуду отраженных волн.

§ 4.1.2. Модель Ванавары. Скоростная модель Ванавары характерна для Восточной Сибири, имеет слоистую структуру с вертикальным включением по центру (рис. 3). Скорость продольной волны принимает значения от 3000 до 6000 м/с. Размер модели — 220 на 2.6 км. Система наблюдения состоит из 1901

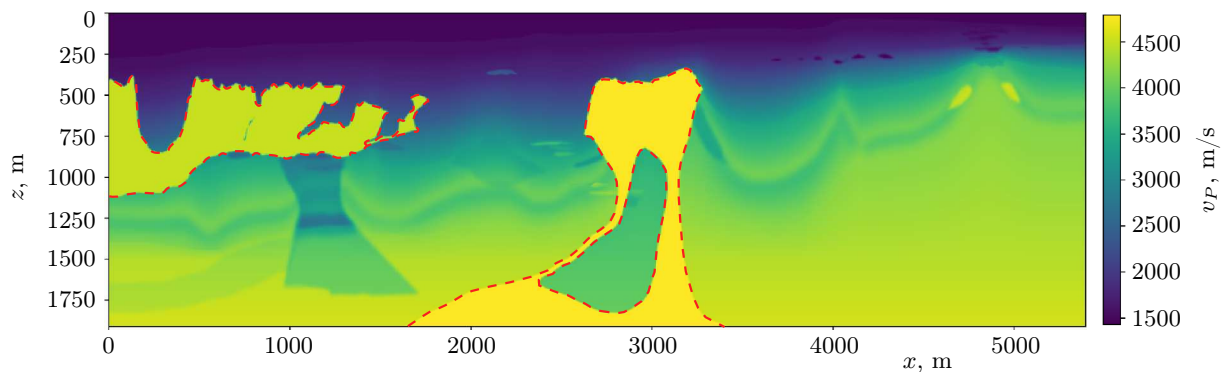


Рис. 2. Скоростная модель ВР, значения скорости продольной волны

Fig. 2. BP velocity model, longitudinal wave velocity values

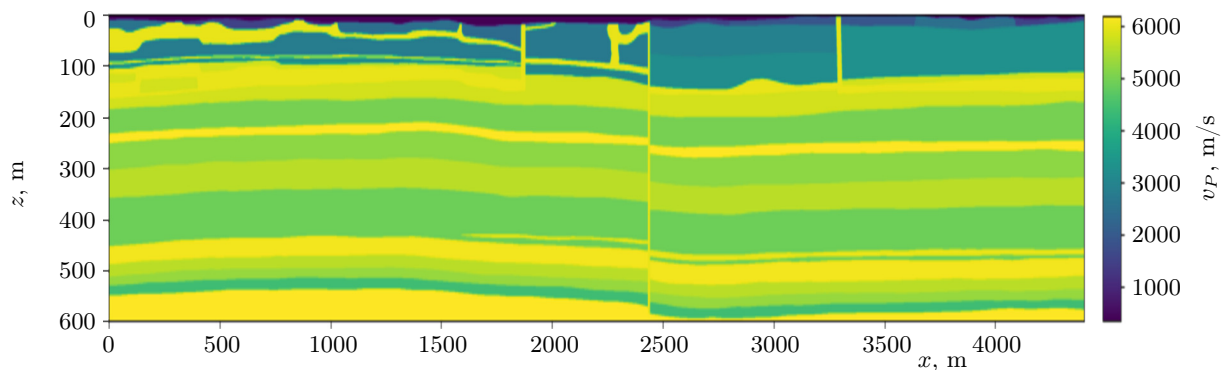


Рис. 3. Скоростная модель Ванавары, значения скорости продольной волны

Fig. 3. Vanavara velocity model, longitudinal wave velocity values



источника, располагающихся на расстоянии 100 м друг от друга. Волновое поле записывалось 512 приемниками для каждого источника с максимальным выбросом 6.4 км. Расстояние между приемниками — 25 м. Источник — вейвлет Рикера с центральной частотой 30 Гц.

Так же как и для модели ВР, моделируются два набора сейсмограмм с использованием мелкой сетки с шагом $h = 2.5$ м и грубой сетки с шагом $h = 5$ м. Сейсмограммы выбираются во временном диапазоне от 1.300 до 1.812 с, шаг по времени — 1 мс, размер сейсмограмм — 512×512 .

4.2. Формирование случайной обучающей выборки. Для статистического анализа выбирают наборы с фиксированным числом сейсмограмм, которые не коррелируют ни по одной из рассматриваемых метрик. Пусть N_s — количество сейсмограмм во всем наборе данных и N — количество сейсмограмм в обучающем наборе. Для генерации некоррелированной последовательности N_s случайных величин, принимающих значения 1 и 0 с вероятностями $p = N/N_s$ и $1 - p$ соответственно, используется биномиальное распределение. После этого формируется набор индексов D_t : в него входят те индексы из последовательности $\{1, \dots, N_s\}$, которым в построенной последовательности случайных величин соответствуют единицы. Затем NDM-net обучается, и рассчитывается средняя ошибка между двумя наборами:

$$\tilde{\varepsilon} = \frac{1}{N_s} \sum_{i=1}^{N_s} d_s(G[\mathbf{u}_{h_2}(t, x_o, x_s^i)], \mathbf{u}_{h_1}(t, x_o, x_s^i)). \quad (2)$$

Для каждой случайным образом сгенерированной выборки оцениваются четыре параметра: общее количество $N = \dim\{D_t\}$ сейсмограмм в обучающей выборке и три расстояния Хаусдорфа между полным набором данных и обучающим:

$$\delta_d = \max_{i \in D} \min_{j \in D_t} d_d(x_s^i, x_s^j),$$

$$\delta_s = \max_{i \in D} \min_{j \in D_t} d_s(\mathbf{u}(t, x_o, x_s^i), \mathbf{u}(t, x_o, x_s^j)),$$

$$\delta_m = \max_{i \in D} \min_{j \in D_t} d_m(M(x, z, x_s^i), M(x, z, x_s^j)),$$

где D — полный набор индексов данных, D_t — набор индексов, формирующий обучающий набор данных.

Таким образом, NDM-net рассматривается как оператор, отображающий набор данных, зависящий от четырех параметров $N, \delta_d, \delta_s, \delta_m$ выборки D_t , на набор данных, характеризуемый одним параметром — средней ошибкой. Ниже приводится анализ чувствительности этого оператора.

4.3. Глобальный анализ чувствительности. Применив глобальный анализ чувствительности, оценим влияние каждого входного параметра относительно ошибки (2) нейронной сети. Глобальный анализ чувствительности широко используется для изучения нейронных сетей [56]. В результате анализа определяются параметры, которыми можно пренебречь и тем самым уменьшить размерность задачи, а также строятся качественные оценки поведения модели [57, 58]. Глобальный анализ чувствительности основывается на регрессионном анализе или дисперсионном анализе [58]. Проведем анализ чувствительности на основе метода Монте-Карло с использованием индексов Соболя [43, 54].

Пусть выходные данные обозначаются как $\tilde{\varepsilon} = f(\mathbf{y})$, где $\mathbf{y} = (N, \delta_d, \delta_s, \delta_m)^T$. Введем индексы Соболя первого порядка

$$S_i = \frac{\text{Var}(\mathbb{E}_{\mathbf{y} \sim i}[\tilde{\varepsilon}|y_i])}{\text{Var}(\tilde{\varepsilon})},$$

где \mathbb{E} — математическое ожидание, Var — дисперсия, y_i — i -й параметр, $\mathbf{y} \sim i$ — набор всех параметров, кроме y_i . Индексы Соболя первого порядка характеризуют изменчивость рассматриваемой функции по одному входному параметру y_i , т.е. оценивают значимость каждого параметра.

Применение глобального анализа чувствительности ко всем четырем входным параметрам дает ожидаемый результат: наиболее значимо общее количество сейсмограмм в наборе обучающих данных. Индексы Соболя S_N^{BP} и $S_N^{Vनावар}$ для моделей ВР и Ванавары принимали значения близкие к 0.63. Таким образом, N подавляет вклад остальных параметров. По тренду ошибки $\tilde{\varepsilon}$, рассчитанной для всех рассмотренных реализаций, заметно, что даже при фиксированном N изменение ошибки велико (рис. 4). Следовательно, разумно сформулировать задачу построения оптимального обучающего набора данных для фиксированного N .

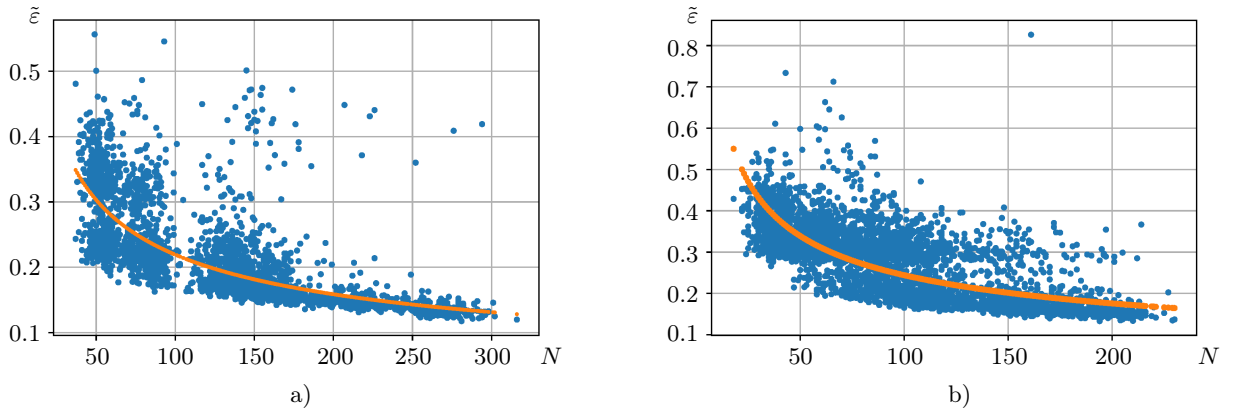


Рис. 4. Зависимости ошибок от размера обучающей выборки: а) для модели ВР; б) для модели Ванавары
 Fig. 4. Trend of error dependence on the size of the training set: а) for BP model, б) for Vanavar model

4.4. Метод главных компонент для фиксированного N . Для оптимизации обучающего набора данных фиксированного размера применяется глобальный анализ чувствительности для каждого N и определяется связь параметров $\delta_d, \delta_s, \delta_m$ при построении обучающего набора данных.

Для глобального анализа чувствительности необходимо, чтобы параметры модели не коррелировали. Если данные не коррелируют, то сумма индексов Соболя равна единице, в противном случае она превышает единицу [59]. Для анализа корреляции между параметрами применяется метод главных компонент.

Математически анализ главных компонент включает спектральное разложение ковариационной матрицы, которое совпадает с сингулярным разложением матрицы исходных данных. Собственные векторы ковариационной матрицы статистически независимы, поэтому ковариационная матрица в новых переменных — диагональная. В трехмерном пространстве параметров $\delta_d, \delta_s, \delta_m$ отношения между исходными параметрами и собственными векторами могут быть представлены в виде корреляционной матрицы. Для обеих рассмотренных моделей генерируется серия наборов данных для различных значений N . После этого для каждого фиксированного N строятся ковариационные матрицы Σ^N с элементами, которые определяются следующим образом:

$$\Sigma_{ij}^N = \text{Cov}(\delta_i^N, \delta_j^N) = \mathbb{E}[(\delta_i^N - \mathbb{E}[\delta_i^N])(\delta_j^N - \mathbb{E}[\delta_j^N])],$$

где \mathbb{E} — среднее значение, δ_i и δ_j — любые из трех расстояний δ_d, δ_s или δ_m , а N — количество сейсмограмм в рассматриваемых наборах данных.

Далее вычисляется разложение по собственным значениям ковариационных матриц и строится график собственных векторов в координатном пространстве $\delta_d, \delta_s, \delta_m$. Чтобы найти общую тенденцию в направлениях собственных векторов, графики для разных N совмещаются (рис. 5).

Параметр δ_m независим для обеих моделей, тогда как два других параметра связаны, поскольку направления двух собственных векторов различаются в пределах плоскости (δ_d, δ_s) . Тем не менее для модели Ванавары собственные векторы образуют кластеры вблизи направлений координат, а значит сумма индексов Соболя превышает единицу. Однако цель исследования — не построение полного набора индексов Соболя, а оценка взаимосвязей между первыми индексами, и, следовательно, данный подход применим непосредственно к рассматриваемой системе.

4.5. Глобальный анализ чувствительности для фиксированного N . Далее будем предполагать, что ошибка может быть представлена в виде

$$\tilde{\varepsilon}_{N_c}(\delta_1, \delta_2, \delta_3) = \tilde{\varepsilon}_0 + \alpha_1 \delta_1 + \alpha_2 \delta_2 + \alpha_3 \delta_3 + o(\delta_1, \delta_2, \delta_3),$$

где $\tilde{\varepsilon}_0$ — среднее значение ошибки, $\tilde{\varepsilon}_{N_c}$ — ошибка для фиксированного размера обучающей выборки $N_c = \text{const}$, $(\delta_1, \delta_2, \delta_3) = (\delta_d, \delta_s, \delta_m)$. Минимизация ошибки соответствует минимизации линейной комбинации

$$l(\delta_1, \delta_2, \delta_3) = q(\alpha_1 \delta_1 + \alpha_2 \delta_2 + \alpha_3 \delta_3) \rightarrow \min,$$

где q — любое положительное вещественное число.



Коэффициенты Соболя для линейной функции:

$$S_{\delta_i} = \frac{\text{Var}(\mathbb{E}_{\delta \sim i}[\tilde{\varepsilon}|\delta_i])}{\text{Var}(\tilde{\varepsilon})} = \frac{(\alpha_i L_{\delta_i})^2}{\sum_j (\alpha_j L_{\delta_j})^2},$$

где L_{δ_i} — разброс переменной δ_i . Видно, что безразмерные переменные изменяются в промежутке $[0, 1]$. Так, индексы Соболя для линейной функции считаются с помощью первых производных.

Коэффициенты α_1, α_2 и α_3 рассчитываются с применением регрессионного анализа, затем введенная метрика нормируется и осредняется по формуле

$$\tilde{\alpha}_i = \frac{1}{N_r} \sum_{k=0}^{N_r} \frac{\sqrt{S_i^k \sum_j (\alpha_j^k L_j^k)^2}}{L_i^k},$$

где N_r — количество случайных обучающих выборок.

Глобальный анализа чувствительности выполняется для обеих скоростных моделей при размере выборки 5–10% от общего количества сейсмограмм (табл. 2). Оценки коэффициентов линейной части функции ошибки приведены в табл. 2.

Таблица 2. Оценки коэффициентов линейной части ошибки

Table 2. Estimates of the coefficients of linear part of the error

Model	S_{δ_d}	S_{δ_s}	S_{δ_m}	$\tilde{\alpha}_d$	$\tilde{\alpha}_s$	$\tilde{\alpha}_m$
BP	0.499	0.0442	0.4565	0.815	0.089	0.095
Vanavar	0.489	0.139	0.3714	0.812	0.112	0.075

5. Численные эксперименты. Выполняются две серии численных экспериментов, по одной для каждой скоростной модели. Для обеих моделей рассматривается линейная комбинация метрик со средними коэффициентами (табл. 2), т.е. исследуется задача минимизации:

$$0.813\delta_d + 0.1\delta_s + 0.085\delta_m \rightarrow \min.$$

В соответствии с линейной комбинацией рассчитываются матрицы расстояний, затем для формирования обучающей выборки с фиксированным N и минимальным расстоянием применяется иерархическая кластеризация [60]. Далее на полученных наборах данных обучается NDM-net. Процедура повторяется для различных N . Так формируется оптимизированная обучающая выборка с набором индексов $D^{H_{new}}$. Также формируются три выборки с наборами индексов $D^{H_d}, D^{H_s}, D^{H_m}$: на основе расстояния между положениями источников d_d , между сейсмограммами d_s , между скоростными моделями d_m .

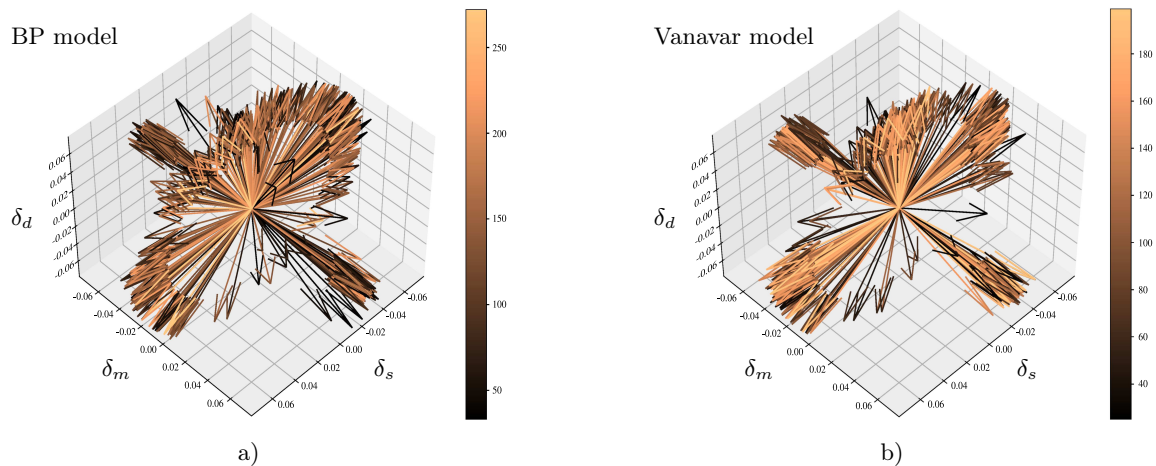


Рис. 5. Собственные векторы ковариационной матрицы Σ : а) для модели ВР; б) для модели Ванавары

Fig. 5. Eigenvectors of the covariation matrix Σ : a) for BP model; b) for Vanavar model

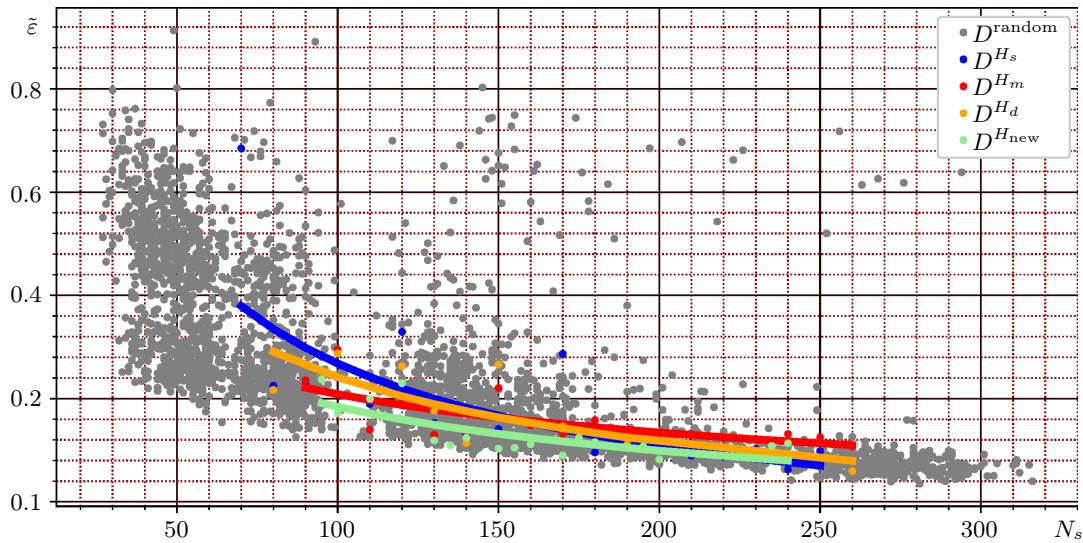


Рис. 6. Средняя ошибка после применения NDM-net ко всему набору данных в зависимости от размера обучающей выборки для различных типов выборки для модели ВР

Fig. 6. The average error of applying NDM-net to all data, depending on the size of training dataset for different types of samples for BP model

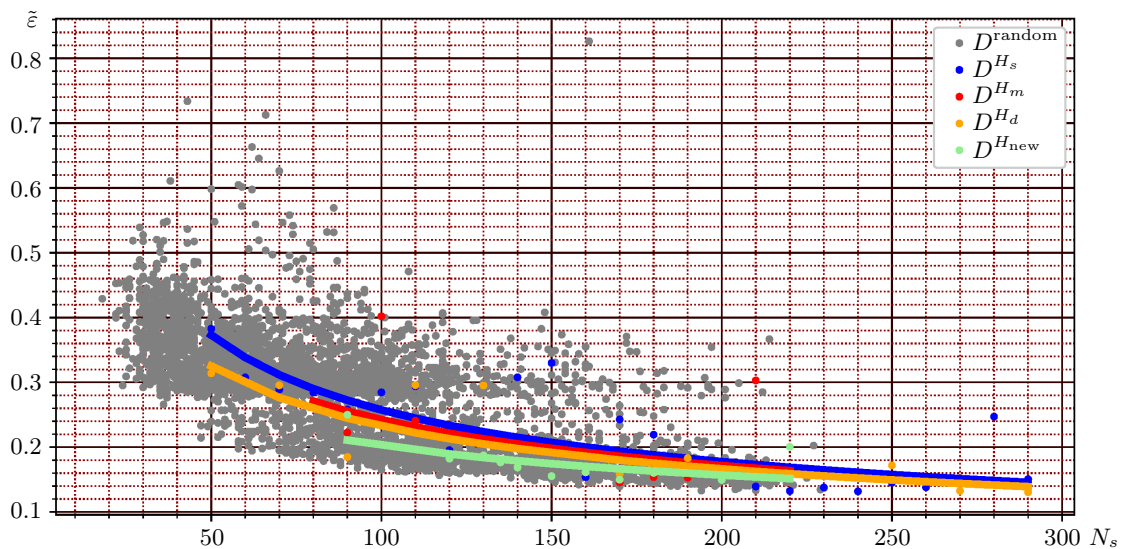


Рис. 7. Средняя ошибка после применения NDM-net ко всему набору данных в зависимости от размера обучающей выборки для различных типов выборки для модели Ванавары

Fig. 7. The average error of applying NDM-net to all data, depending on the size of the training dataset for different types of samples for the Vanavar model

По формуле (2) вычисляются ошибки $\tilde{\epsilon}$ для оптимизированных наборов данных в зависимости от N для моделей ВР (рис. 6) и Ванавары (рис. 7) и ошибки, соответствующие минимизации только одного расстояния, для случайным образом сгенерированных наборов данных, индексы которых обозначаются как D^{random} . Графики демонстрируют, насколько меньшую ошибку обеспечивает оптимизированный набор данных в сравнении с остальными способами построения обучающего набора данных.

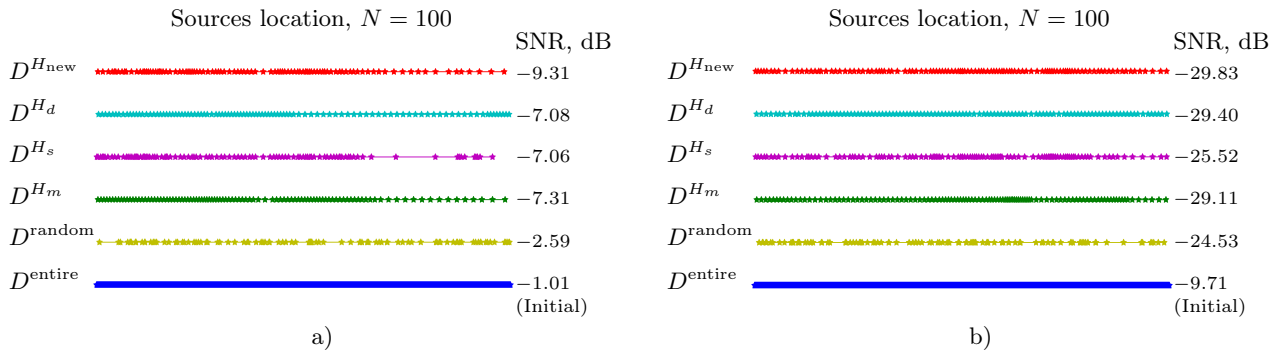


Рис. 8. Распределение источников в пространстве для каждого типа выборок и соответствующее им результирующее среднее отношение сигнал/шум: а) для модели ВР; б) для модели Ванавары

Fig. 8. Distribution of sources in space for each type of samples and the resulting average signal-to-noise ratio corresponding to them: a) for the BP model; b) for the Vanavara model

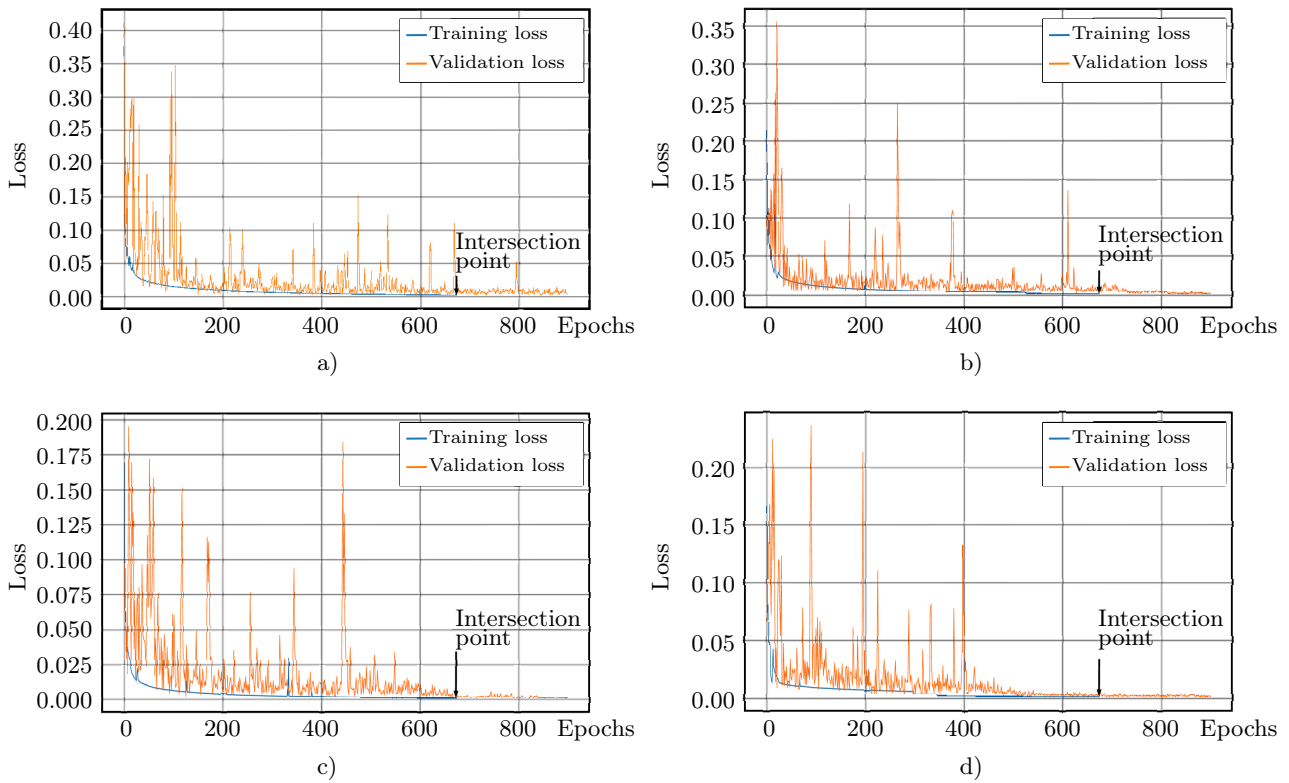


Рис. 9. Кривые обучения NDM-net для модели Ванавары с использованием различных тренировочных наборов данных: а) D^{H_d} ; б) D^{H_s} ; в) D^{H_m} ; г) $D^{H_{new}}$

Fig. 9. Loss curves for NDM-net training using Vanavar model using different datasets: a) D^{H_d} ; б) D^{H_s} ; в) D^{H_m} ; г) $D^{H_{new}}$

Кроме того, для проверки результатов используется отношение сигнал/шум (SNR), которое определяется по формуле

$$SNR = 20 \lg \left(\frac{\|\mathbf{u}_{h_1}(t, x_o, x_s^i) - G[\mathbf{u}_{h_2}(t, x_o, x_s^i)]\|_2}{\|\mathbf{u}_{h_1}(t, x_o, x_s^i)\|_2} \right).$$

Среднее значение SNR между полным сгенерированным нейронной сетью набором данных и набором данных, рассчитанных с использованием мелкой сетки, приводится на рис. 8 для различных типов обучающей выборки совместно с расположением источников. Для модели ВР и модели Ванавары размеры

выборки составляют $N = 100$ и $N = 110$ соответственно. Для модели ВР оптимизация набора данных позволила уменьшить SNR в 9 раз, в то время как использование других обучающих наборов данных уменьшает SNR в 7 раз. Для модели Ванавары применение нового набора данных уменьшает SNR в 3 раза, на других наборах данных — почти в 2.5 раза.

Кривые обучения и проверки NDM-net с использованием различных наборов данных для модели Ванавары показывают, что нейронная сеть эффективно работает без переобучения (рис. 9). Кривые обучения для модели ВР аналогичны.

6. Заключение. Представлен новый способ оптимизации обучающей выборки для NDM-net. Нейронная сеть NDM-net разрабатывается для подавления численной дисперсии в результатах сейсмического моделирования, при котором система уравнений динамической теории упругости решается для нескольких правых частей. Решение для всех положений источников, т.е. для всех правых частей, вычисляется с использованием конечно-разностной схемы низкого порядка аппроксимации на грубой сетке. Обучающий набор данных генерируется с использованием достаточно мелкой сетки для небольшого числа источников. Затем NDM-net обучается отображать решение на грубой сетке в решение на мелкой сетке. Генерация обучающего набора данных — наиболее трудоемкая часть алгоритма применения NDM-net.

Строится обучающий набор данных, минимизирующий ошибку выхода нейронной сети для фиксированного числа правых частей. Рассматриваются три метрики расхождения между данными: физическое расстояние между позициями источников, прямое расхождение между сейсмограммами для различных источников и разница в скоростных моделях, используемых для моделирования сейсмических волновых полей. Для каждой метрики вводится хаусдорфово расстояние между обучающим набором данных и всем набором сейсмограмм. В итоге предлагается линейная комбинация рассматриваемых метрик с коэффициентами, полученными на основе глобального анализа чувствительности ошибки NDM-net к каждой из этих метрик.

Результаты численных экспериментов показали, что обученная на наборе данных, построенном на основе линейной комбинации метрик, NDM-net подавляет численную дисперсию в несколько раз эффективнее, чем при использовании других обучающих наборов данных для двух скоростных моделей. Оценки для коэффициентов линейной комбинации метрик, полученные глобальным анализом чувствительности для двух рассматриваемых моделей, близки. Тот факт, что эти скоростные модели характеризуются принципиально различным геологическим строением, позволяет предположить универсальность полученных результатов и возможность их обобщения и применения для других сейсмогеологических моделей.

Список литературы

1. Saenger E.H., Gold N., Shapiro S.A. Modeling the propagation of elastic waves using a modified finite-difference grid // *Wave Motion*. 2000. **31**, N 1. 77–92. doi 10.1016/S0165-2125(99)00023-2.
2. Lisitsa V., Vishnevskiy D. Lebedev scheme for the numerical simulation of wave propagation in 3D anisotropic elasticity // *Geophysical Prospecting*. 2010. **58**, N 4. 619–635. doi 10.1111/j.1365-2478.2009.00862.x.
3. Lisitsa V., Tcheverda V., Vishnevskiy D. Numerical simulation of seismic waves in models with anisotropic formations: coupling Virieux and Lebedev finite-difference schemes // *Computational Geosciences*. 2012. **16**, N 4. 1135–1152. doi 10.1007/s10596-012-9308-0.
4. Carcione J.M., Cavallini F. A rheological model for anelastic anisotropic media with applications to seismic wave propagation // *Geophys. J. Int.* 1994. **119**, N 1. 338–348. doi 10.1111/j.1365-246X.1994.tb00931.x.
5. Вишневецкий Д.М., Луцица В.В., Решетова Г.В. Численное моделирование распространения сейсмических волн в средах с вязкоупругими включениями // *Вычислительные методы и программирование*. 2013. **14**, № 1. 155–165. <https://num-meth.ru/index.php/journal/article/view/593>. Cited March 30, 2024.
6. Carcione J.M., Morency C., Santos J.E. Computational poroelasticity — a review // *Geophysics*. 2010. **75**, N 5. 75A229–75A243. doi 10.1190/1.3474602.
7. Masson Y.J., Pride S.R. Finite-difference modeling of Biot’s poroelastic equations across all frequencies // *Geophysics*. 2010. **75**, N 2. N33–N41. doi 10.1190/1.3332589.
8. Caspari E., Novikov M., Lisitsa V., et al. Attenuation mechanisms in fractured fluid-saturated porous rocks: a numerical modelling study // *Geophysical Prospecting*. 2019. **67**, N 4. 935–955. doi 10.1111/1365-2478.12667.
9. Verliac M., Calvez J.L. Microseismic monitoring for reliable CO₂ injection and storage — geophysical modeling challenges and opportunities // *The Leading Edge*. 2021. **40**, N 6. 418–423. doi 10.1190/tle40060418.1.



10. Prasad M., Glubokovskikh S., Daley T., et al. CO₂ messes with rock physics // The Leading Edge. 2021. 40, N 6. 424–432. doi 10.1190/tle40060424.1.
11. Virieux J., Operto S., Ben-Hadj-Ali H., et al. Seismic wave modeling for seismic imaging // The Leading Edge. 2009. 28, N 5. 538–544. doi 10.1190/1.3124928.
12. Virieux J., Calandra H., Plessix R.É. A review of the spectral, pseudo-spectral, finite-difference and finite-element modelling techniques for geophysical imaging // Geophysical Prospecting. 2011. 59, N 5. 794–813. doi 10.1111/j.1365-2478.2011.00967.x.
13. Virieux J. P–SV wave propagation in heterogeneous media: velocity–stress finite-difference method // Geophysics. 1986. 51, N 4. 889–901. doi 10.1190/1.1442147.
14. Moczo P., Kristek J., Gális M. The finite-difference modelling of earthquake motions: waves and ruptures. Cambridge: Cambridge University Press, 2014. doi 10.1017/CB09781139236911.
15. Садовский В.М., Садовская О.В. Вычислительный алгоритм для расчета вязкоупругих волн в среде Кельвина–Фойхта // Вычислительные методы и программирование. 2014. 15, № 1. 98–108. <https://en.num-meth.ru/index.php/journal/article/view/753>. Cited March 29, 2024.
16. Zhang J., Verschuur D.J. Elastic wave propagation in heterogeneous anisotropic media using the lumped finite-element method // Geophysics. 2002. 67, N 2. 625–638. doi 10.1190/1.1468624.
17. Tromp J., Komatitsch D., Liu Q. Spectral-element and adjoint methods in seismology // Commun. Comput. Phys. 2008. 3, N 1. 1–32. <https://authors.library.caltech.edu/records/stavm-ptm61>. Cited March 29, 2024.
18. Zhang Y., Fisser J., Gerritsma M. A hybrid mimetic spectral element method for three-dimensional linear elasticity problems // J. Comput. Phys. 2021. 433, Article Identifier 110179. doi 10.1016/j.jcp.2021.110179.
19. Dumbser M., Käser M. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - II. The three-dimensional isotropic case // Geophys. J. Int. 2006. 167, N 1. 319–336. doi 10.1111/j.1365-246X.2006.03120.x.
20. Puente J., Käser M., Dumbser M., Igel H. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes – IV. Anisotropy // Geophys. J. Int. 2007. 169, N 3. 1210–1228. doi 10.1111/j.1365-246X.2007.03381.x.
21. Lisitsa V., Tcheverda V., Botter C. Combination of the discontinuous Galerkin method with finite differences for simulation of seismic wave propagation // J. Comput. Phys. 2016. 311. 142–157. doi 10.1016/j.jcp.2016.02.005.
22. Fornberg B. The pseudospectral method: accurate representation of interfaces in elastic wave calculations // Geophysics. 1988. 53, N 5. 625–637. doi 10.1190/1.1442497.
23. Pleshkevich A., Vishnevsky D., Lisitsa V., Levchenko V. Parallel algorithm for one-way wave equation based migration for seismic imaging // Communications in Computer and Information Science. Vol. 965. Cham: Springer, 2019. 125–135. doi 10.1007/978-3-030-05807-4_11.
24. Луцуца В.В. Дисперсионный анализ разрывного метода Галеркина в применении к уравнениям динамической теории упругости // Вычислительные методы и программирование. 2015. 16, № 3. 387–406. doi 10.26089/NumMet.v16r338.
25. Levander A.R. Fourth-order finite-difference P–SV seismograms // Geophysics. 1988. 53, N 11. 1425–1436. doi 10.1190/1.1442422.
26. O’Reilly O., Lundquist T., Dunham E.M., Nordström J. Energy stable and high-order-accurate finite difference methods on staggered grids // J. Comput. Phys. 2017. 346, 572–589. doi 10.1016/j.jcp.2017.06.030.
27. Lisitsa V. Dispersion analysis of discontinuous Galerkin method on triangular mesh for elastic wave equation // Applied Mathematical Modelling. 2016. 40, N 7–8. 5077–5095. doi 10.1016/j.apm.2015.12.039.
28. Liu Y., Sen M.K. Acoustic VTI modeling with a time–space domain dispersion-relation-based finite-difference scheme // Geophysics. 2010. 75, N 3. A11–A17. doi 10.1190/1.3374477.
29. Chen J.-B. A 27-point scheme for a 3D frequency-domain scalar wave equation based on an average-derivative method // Geophysical Prospecting. 2014. 62, N 2. 258–277. doi 10.1111/1365-2478.12090.
30. He Z., Zhang J., Yao Z. Determining the optimal coefficients of the explicit finite-difference scheme using the Remez exchange algorithm // Geophysics. 2019. 84, N 3. S137–S147. doi 10.1190/geo2018-0446.1.
31. Koene E.F.M., Robertsson J.O.A., Brogгинi F., Andersson F. Eliminating time dispersion from seismic wave modeling // Geophys. J. Int. 2018. 213, N 1. 169–180. doi 10.1093/gji/ggx563.
32. Mittet R. Second-order time integration of the wave equation with dispersion correction procedures // Geophysics. 2019. 84, N 4. T221–T235. doi 10.1190/geo2018-0770.1.
33. Mulder W.A. Temporal dispersion correction for wave-propagation modelling with a series approach // Geophysical Prospecting. 2024. 72, N 2. 301–314 doi 10.1111/1365-2478.13411.

34. *Gadylshin K., Vishnevsky D., Gadylshina K., Lisitsa V.* Numerical dispersion mitigation neural network for seismic modeling // *Geophysics*. 2022. **87**, N 3. T237–T249. doi [10.1190/geo2021-0242](https://doi.org/10.1190/geo2021-0242). 1.
35. *Гадьльшина К.А., Лисица В.В., Вишневецкий Д.М., Гадьльшин К.Г.* Искусственная нейронная сеть, уменьшающая численную дисперсию, для постобработки результатов сейсмического моделирования // *Геофизические технологии*. 2022. № 1. 99–109. doi [10.18303/2619-1563-2022-1-99](https://doi.org/10.18303/2619-1563-2022-1-99).
36. *Gondyul E., Lisitsa V., Gadylshin K., Vishnevsky D.* Numerical dispersion mitigation neural network with the model-based training dataset optimization // *Lecture Notes in Computer Science*. Vol. 14106. Cham: Springer, 2023. 19–30. doi [10.1007/978-3-031-37111-0_2](https://doi.org/10.1007/978-3-031-37111-0_2).
37. *Гадьльшина К.А., Вишневецкий Д.М., Гадьльшин К.Г., Лисица В.В.* Построение тренировочной обучающей выборки на основе хаусдорфовой метрики в пространстве сейсмограмм для подавляющей численную дисперсию нейронной сети // *Вычислительные методы и программирование*. 2023. **24**, № 2. 195–212. doi [10.26089/NumMet.v24r215](https://doi.org/10.26089/NumMet.v24r215).
38. *Gadylshin K., Lisitsa V., Vishnevsky D., Gadylshina K.* Hausdorff-distance-based training dataset construction for numerical dispersion mitigation neural network // *Computers and Geosciences*. 2023. **180**, Article Number 105438. doi [10.1016/j.cageo.2023.105438](https://doi.org/10.1016/j.cageo.2023.105438).
39. *Wang R., Meng Y., Zhang W., et al.* Remote sensing semantic segregation for water information extraction: optimization of samples via training error performance // *IEEE Access*. 2019. **7**, 13383–13395. doi [10.1109/ACCESS.2019.2894099](https://doi.org/10.1109/ACCESS.2019.2894099).
40. *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 // *Proc. 5th Int. Conf. on Information and Communication Technologies for Disaster Management, Sendai, Japan, December 04–07, 2018*. Published by IEEE Press in <https://ieeexplore.ieee.org/xpl/conhome/8634895/proceeding>. Cited March 29, 2024. doi [10.1109/ICT-DM.2018.8636375](https://doi.org/10.1109/ICT-DM.2018.8636375).
41. *Everaert D., Potts C.* GIO: gradient information optimization for training dataset selection // doi [10.48550/arXiv.2306.11670](https://doi.org/10.48550/arXiv.2306.11670).
42. *Park J., Choi J., Seol S.J., et al.* A method for adequate selection of training data sets to reconstruct seismic data using a convolutional U-Net // *Geophysics*. 2021. **86**, N 5. V375–V388. doi [10.1190/geo2019-0708](https://doi.org/10.1190/geo2019-0708). 1.
43. *Sobol' I. M., Tarantola S., Gatelli D., et al.* Estimating the approximation error when fixing unessential factors in global sensitivity analysis // *Reliab. Eng. Syst. Saf.* 2007. **92**, N 7. 957–960. doi [10.1016/j.ress.2006.07.001](https://doi.org/10.1016/j.ress.2006.07.001).
44. *Berenger J.-P.* A Huygens subgridding for the FDTD method // *IEEE Trans. Antennas Propag.* 2006. **54**, N 12. 3797–3804. doi [10.1109/TAP.2006.886519](https://doi.org/10.1109/TAP.2006.886519).
45. *Collino F., Tsogka C.* Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media // *Geophysics*. 2001. **66**, N 1. 294–307. doi [10.1190/1.1444908](https://doi.org/10.1190/1.1444908).
46. *Lisitsa V.* Optimal discretization of PML for elasticity problems // *Electron. Trans. Numer. Anal.* 2008. **30**. 258–277. <http://eudml.org/doc/130538>. Cited March 29, 2024.
47. *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.* 2015. **281**, 669–689. doi [10.1016/j.jcp.2014.10.047](https://doi.org/10.1016/j.jcp.2014.10.047).
48. *Самарский А.А.* Теория разностных схем. М.: Наука, 1983.
49. *Moczko P., Kristek J., Varyčuk V., et al.* 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.* 2002. **92**, N 8. 3042–3066. doi [10.1785/0120010167](https://doi.org/10.1785/0120010167).
50. *Vishnevsky D., Lisitsa V., Tcheverda V., Reshetova G.* Numerical study of the interface errors of finite-difference simulations of seismic waves // *Geophysics*. 2014. **79**, N 4. T219–T232. doi [10.1190/geo2013-0299](https://doi.org/10.1190/geo2013-0299). 1.
51. *Lisitsa V., Podgornova O., Tcheverda V.* On the interface error analysis for finite difference wave simulation // *Computational Geosciences*. 2010. **14**, N 4. 769–778. doi [10.1007/s10596-010-9187-1](https://doi.org/10.1007/s10596-010-9187-1).
52. *Ainsworth M.* Discrete dispersion relation for *hp*-version finite element approximation at high wave number // *SIAM J. Numer. Anal.* 2004. **42**, N 2. 553–575. doi [10.1137/S0036142903423460](https://doi.org/10.1137/S0036142903423460).
53. *Ronneberger O., Fischer P., Brox T.* U-Net: convolutional networks for biomedical image segmentation // *Lecture Notes in Computer Science*. Vol. 9351. Cham: Springer, 2015. 234–241. doi [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
54. *Sobol' I.M.* Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates // *Math. Comput. Simul.* 2001. **55**, N 1–3. 271–280. doi [10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6).
55. *Saltelli A., Ratto M., Andres T., et al.* Sensitivity Analysis: from Theory to Practice // *Global Sensitivity Analysis. The Primer*. Chichester: John Wiley, 2008. 237–275. doi [10.1002/9780470725184.ch6](https://doi.org/10.1002/9780470725184.ch6).
56. *Yeung D., Cloete I., Shi D., Ng W.W.Y.* Sensitivity analysis for neural networks. Berlin: Springer, 2010. doi [10.1007/978-3-642-02532-7](https://doi.org/10.1007/978-3-642-02532-7).



57. Reuter U., Liebscher M. Global sensitivity analysis in view of nonlinear structural behavior // https://www.researchgate.net/profile/Martin_Liebscher/publication/242318342_Global_sensitivity_analysis_in_view_of_nonlinear_structural_behavior/links/58e213c5aca272059ab08f35/Global-sensitivity-analysis-in-view-of-nonlinear-structural-behavior.pdf. Cited March 29, 2024.
58. Morio J. Global and local sensitivity analysis methods for a physical system // Eur. J. Phys. 2011. **32**, N 6. 1577–1583. doi 10.1088/0143-0807/32/6/011.
59. Saltelli A., Tarantola S. On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal // J. Am. Stat. Assoc. 2002. **97**, N 459. 702–709. doi 10.1198/016214502388618447.
60. Merwe D.W., Engelbrecht A.P. Data clustering using particle swarm optimization // Proc. 2003 Congress on Evolutionary Computation, Canberra, Australia, December 08–12, 2003. Published by IEEE Press in Congress on Evolutionary Computation, Vol. 1, pp. 215–220 (2003): <https://ieeexplore.ieee.org/xpl/conhome/9096/proceeding?isnumber=28874&sortType=vol-only-seq&pageNumber=2>. Cited March 29, 2024. doi 10.1109/CEC.2003.1299577.

Поступила в редакцию
 31 января 2024 г.

Принята к публикации
 16 марта 2024 г.

Информация об авторах

Елена Александровна Гондюл — инженер; Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН, пр-кт Академика Коптюга, 3, 630090, Новосибирск, Российская Федерация.

Вадим Викторович Лисица — д.ф.-м.н., заведующий лабораторией; Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН, пр-кт Академика Коптюга, 3, 630090, Новосибирск, Российская Федерация.

Кирилл Геннадьевич Гадильшин — к.ф.-м.н., ст. науч. сотр.; Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН, пр-кт Академика Коптюга, 3, 630090, Новосибирск, Российская Федерация.

Дмитрий Михайлович Вишневецкий — к.ф.-м.н., науч. сотр.; Институт нефтегазовой геологии и геофизики имени А. А. Трофимука СО РАН, пр-кт Академика Коптюга, 3, 630090, Новосибирск, Российская Федерация.

References

1. E. H. Saenger, N. Gold, and S. A. Shapiro, “Modeling the Propagation of Elastic Waves Using a Modified Finite-Difference Grid,” *Wave Motion* **31** (1), 77–92 (2000) doi 10.1016/S0165-2125(99)00023-2.
2. V. Lisitsa and D. Vishnevskiy, “Lebedev Scheme for the Numerical Simulation of Wave Propagation in 3D Anisotropic Elasticity,” *Geophys. Prospect.* **58** (4), 619–635 (2010). doi 10.1111/j.1365-2478.2009.00862.x.
3. V. Lisitsa, V. Tcheverda, and D. Vishnevsky, “Numerical Simulation of Seismic Waves in Models with Anisotropic Formations: Coupling Virieux and Lebedev Finite-Difference Schemes,” *Comput. Geosci.* **16** (4), 1135–1152 (2012). doi 10.1007/s10596-012-9308-0.
4. J. M. Carcione and F. Cavallini, “A Rheological Model for Anelastic Anisotropic Media with Applications to Seismic Wave Propagation,” *Geophys. J. Int.* **119** (1), 338–348 (1994). doi 10.1111/j.1365-246X.1994.tb00931.x.
5. D. M. Vishnevsky, V. V. Lisitsa, and G. V. Reshetova, “Numerical Simulation of Seismic Wave Propagation in Media with Viscoelastic Inclusions,” *Numerical Methods and Programming (Vychislitel’nye Metody i Programmirovaniye)* **14** (1), 155–165 (2013). <https://num-meth.ru/index.php/journal/article/view/593>. Cited March 29, 2024.
6. J. M. Carcione, C. Morency, and J. E. Santos, “Computational Poroelasticity — A Review,” *Geophysics.* **75** (5), 75A229–75A243 (2010). doi 10.1190/1.3474602.
7. Y. J. Masson and S. R. Pride, “Finite-Difference Modeling of Biot’s Poroelastic Equations across all Frequencies,” *Geophysics* **75** (2), N33–N41 (2010). doi 10.1190/1.3332589.
8. E. Caspari, M. Novikov, V. Lisitsa, et al., “Attenuation Mechanisms in Fractured Fluid-Saturated Porous Rocks: A Numerical Modelling Study,” *Geophys. Prospect.* **67** (4), 935–955 (2019). doi 10.1111/1365-2478.12667.
9. M. Verliac and J. L. Calvez, “Microseismic Monitoring for Reliable CO₂ Injection and Storage — Geophysical Modeling Challenges and Opportunities,” *The Lead. Edge* **40** (6), 418–423 (2021). doi 10.1190/tle40060418.1.

10. M. Prasad, S. Glubokovskikh, T. Daley, et al., “CO₂ Messes with Rock Physics,” *The Lead. Edge* **40** (6), 424–432 (2021). doi [10.1190/tle40060424.1](https://doi.org/10.1190/tle40060424.1).
11. J. Virieux, S. Operto, H. Ben-Hadj-Ali, et al., “Seismic Wave Modeling for Seismic Imaging,” *The Lead. Edge* **28** (5), 538–544 (2009). doi [10.1190/1.3124928](https://doi.org/10.1190/1.3124928).
12. J. Virieux, H. Calandra, and R. É. Plessix, “A Review of the Spectral, Pseudo-Spectral, Finite-Difference and Finite-Element Modelling Techniques for Geophysical Imaging,” *Geophys. Prospect.* **59** (5), 794–813 (2011). doi [10.1111/j.1365-2478.2011.00967.x](https://doi.org/10.1111/j.1365-2478.2011.00967.x).
13. J. Virieux, “*P*–*SV* Wave Propagation in Heterogeneous Media: Velocity–Stress Finite-Difference Method,” *Geophysics* **51** (4), 889–901 (1986). doi [10.1190/1.1442147](https://doi.org/10.1190/1.1442147).
14. P. Moczo, J. Kristek, and M. Gális, *The Finite-Difference Modelling of Earthquake Motions: Waves and Ruptures* (Cambridge Univ. Press, Cambridge, 2014). doi [10.1017/CB09781139236911](https://doi.org/10.1017/CB09781139236911).
15. V. M. Sadovskii and O. V. Sadovskaya, “A Numerical Algorithm for the Analysis of Viscoelastic Waves in the Kelvin–Voigt Medium,” *Numerical Methods and Programming (Vychislitel’nye Metody i Programirovanie)* **15** (1), 98–108 (2014). <https://en.num-meth.ru/index.php/journal/article/view/753>. Cited March 29, 2024.
16. J. Zhang and D. J. Verschuur, “Elastic Wave Propagation in Heterogeneous Anisotropic Media Using the Lumped Finite-Element Method,” *Geophysics* **67** (2), 625–638 (2002). doi [10.1190/1.1468624](https://doi.org/10.1190/1.1468624).
17. J. Tromp, D. Komatitsch, and Q. Liu, “Spectral-Element and Adjoint Methods in Seismology,” *Commun. Comput. Phys.* **3** (1), 1–32 (2008). <https://authors.library.caltech.edu/records/stavm-ptm61>. Cited March 29, 2024.
18. Y. Zhang, J. Fisser, and M. Gerritsma, “A Hybrid Mimetic Spectral Element Method for Three-Dimensional Linear Elasticity Problems,” *J. Comput. Phys.* **433**, Article Identifier 110179 (2021). doi [10.1016/j.jcp.2021.110179](https://doi.org/10.1016/j.jcp.2021.110179).
19. M. Dumbser and M. Käser, “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). doi [10.1111/j.1365-246X.2006.03120.x](https://doi.org/10.1111/j.1365-246X.2006.03120.x).
20. J. Puente, M. Käser, M. Dumbser, and H. Igel, “An Arbitrary High-Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes — IV. Anisotropy,” *Geophys. J. Int.* **169** (3), 1210–1228 (2007). doi [10.1111/j.1365-246X.2007.03381.x](https://doi.org/10.1111/j.1365-246X.2007.03381.x).
21. V. Lisitsa, V. Tcheverda, and C. Botter, “Combination of the Discontinuous Galerkin Method with Finite Differences for Simulation of Seismic Wave Propagation,” *J. Comput. Phys.* **311**, 142–157 (2016). doi [10.1016/j.jcp.2016.02.005](https://doi.org/10.1016/j.jcp.2016.02.005).
22. B. Fornberg, “The Pseudospectral Method: Accurate Representation of Interfaces in Elastic Wave Calculations,” *Geophysics* **53** (5), 625–637 (1988). doi [10.1190/1.1442497](https://doi.org/10.1190/1.1442497).
23. A. Pleshkevich, D. Vishnevsky, V. Lisitsa, and V. Levchenko, “Parallel Algorithm for One-Way Wave Equation Based Migration for Seismic Imaging,” in *Communications in Computer and Information Science* (Springer, Cham, 2019), Vol. 965, pp. 125–135 (2019). doi [10.1007/978-3-030-05807-4_11](https://doi.org/10.1007/978-3-030-05807-4_11).
24. V. V. Lisitsa, “Dispersion Analysis of the Discontinuous Galerkin Method as Applied to the Equations of the Dynamic Elasticity Theory,” *Numerical Methods and Programming (Vychislitel’nye Metody i Programirovanie)* **16** (3), 387–406 (2015). doi [10.26089/NumMet.v16r338](https://doi.org/10.26089/NumMet.v16r338).
25. A. R. Levander, “Fourth-Order Finite-Difference *P*–*SV* Seismograms,” *Geophysics* **53** (11), 1425–1436 (1988). doi [10.1190/1.1442422](https://doi.org/10.1190/1.1442422).
26. O. O’Reilly, T. Lundquist, E. M. Dunham, and J. Nordström, “Energy Stable and High-Order-Accurate Finite Difference Methods on Staggered Grids,” *J. Comput. Phys.* **346**, 572–589 (2017). doi [10.1016/j.jcp.2017.06.030](https://doi.org/10.1016/j.jcp.2017.06.030).
27. V. Lisitsa, “Dispersion Analysis of Discontinuous Galerkin Method on Triangular Mesh for Elastic Wave Equation,” *Appl. Math. Model.* **40** (7–8), 5077–5095 (2016). doi [10.1016/j.apm.2015.12.039](https://doi.org/10.1016/j.apm.2015.12.039).
28. Y. Liu and M. K. Sen, “Acoustic VTI Modeling with a Time-Space Domain Dispersion-Relation-Based Finite-Difference Scheme,” *Geophysics* **75** (3), A11–A17 (2010). doi [10.1190/1.3374477](https://doi.org/10.1190/1.3374477).
29. J.-B. Chen, “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). doi [10.1111/1365-2478.12090](https://doi.org/10.1111/1365-2478.12090).
30. Z. He, J. Zhang, and Z. Yao, “Determining the Optimal Coefficients of the Explicit Finite-Difference Scheme Using the Remez Exchange Algorithm,” *Geophysics* **84** (3), S137–S147 (2019). doi [10.1190/geo2018-0446.1](https://doi.org/10.1190/geo2018-0446.1).
31. E. F. M. Koene, J. O. A. Robertsson, F. Brogini, and F. Andersson, “Eliminating Time Dispersion from Seismic Wave Modeling,” *Geophys. J. Int.* **213** (1), 169–180 (2018). doi [10.1093/gji/ggx563](https://doi.org/10.1093/gji/ggx563).
32. R. Mittet, “Second-Order Time Integration of the Wave Equation with Dispersion Correction Procedures,” *Geophysics* **84** (4), T221–T235 (2019). doi [10.1190/geo2018-0770.1](https://doi.org/10.1190/geo2018-0770.1).



33. W. A. Mulder, “Temporal Dispersion Correction for Wave-Propagation Modelling with a Series Approach,” *Geophys. Prospect.* **72** (2), 301–314 (2024). doi [10.1111/1365-2478.13411](https://doi.org/10.1111/1365-2478.13411).
34. K. Gadylyshin, D. Vishnevsky, K. Gadylyshina, and V. Lisitsa, “Numerical Dispersion Mitigation Neural Network for Seismic Modeling,” *Geophysics* **87** (3), T237–T249 (2022). doi [10.1190/geo2021-0242.1](https://doi.org/10.1190/geo2021-0242.1).
35. K. Gadylyshina, V. Lisitsa, D. Vishnevsky, and K. Gadylyshin, “An Artificial Neural Network that Reduces Numerical Variance for Postprocessing the Results of Seismic Modeling,” *Geophys. Technol.*, No. 1, 99–109 (2022). doi [10.18303/2619-1563-2022-1-99](https://doi.org/10.18303/2619-1563-2022-1-99).
36. E. Gondyul, V. Lisitsa, K. Gadylyshin, and D. Vishnevsky, “Numerical Dispersion Mitigation Neural Network with the Model-Based Training Dataset Optimization,” in *Lecture Notes in Computer Science* (Springer, Cham, 2023), Vol. 14106, pp. 19–30. doi [10.1007/978-3-031-37111-0_2](https://doi.org/10.1007/978-3-031-37111-0_2).
37. K. A. Gadylyshina, D. M. Vishnevsky, K. G. Gadylyshin, and V. V. Lisitsa, “Training Data Set Construction Based on the Hausdorff Metric for Numerical Dispersion Mitigation Neural Network in Seismic Modelling,” *Numerical Methods and Programming (Vychislitel’nye Metody i Programirovanie)*. **24** (2), 195–212 (2023). doi [10.26089/NumMet.v24r215](https://doi.org/10.26089/NumMet.v24r215).
38. K. Gadylyshin, V. Lisitsa, D. Vishnevsky, and K. Gadylyshina, “Hausdorff-Distance-Based Training Dataset Construction for Numerical Dispersion Mitigation Neural Network,” *Comput. and Geosci.* **180**, Article Number 105438 (2023). doi [10.1016/j.cageo.2023.105438](https://doi.org/10.1016/j.cageo.2023.105438).
39. R. Wang, Y. Meng, W. Zhang, et al., “Remote Sensing Semantic Segregation for Water Information Extraction: Optimization of Samples via Training Error Performance,” *IEEE Access* **7**, 13383–13395 (2019). doi [10.1109/ACCESS.2019.2894099](https://doi.org/10.1109/ACCESS.2019.2894099).
40. S. Yan, L. Zhang, and D. Liu, “An Empirical Study on Optimization of Training Dataset in Harmfulness Prediction of Code Clone Using Ensemble Feature Selection Model,” in *Proc. Int. Conf. on Information and Communication Technologies for Disaster Management, Sendai, Japan, December 04–07, 2018*. Published by IEEE Press in <https://ieeexplore.ieee.org/xpl/conhome/8634895/proceeding>. Cited March 29, 2024. doi [10.1109/ICT-DM.2018.8636375](https://doi.org/10.1109/ICT-DM.2018.8636375).
41. D. Everaert and C. Potts, “GIO: Gradient Information Optimization for Training Dataset Selection,” doi [10.48550/arXiv.2306.11670](https://doi.org/10.48550/arXiv.2306.11670).
42. J. Park, J. Choi, S. J. Seol, et al., “A Method for Adequate Selection of Training Data Sets to Reconstruct Seismic Data Using a Convolutional U-Net,” *Geophysics* **86** (5), V375–V388 (2021). doi [10.1190/geo2019-0708.1](https://doi.org/10.1190/geo2019-0708.1).
43. I. M. Sobol’, S. Tarantola, D. Gatelli, et al., “Estimating the Approximation Error when Fixing Unessential Factors in Global Sensitivity Analysis,” *Reliab. Eng. Syst. Saf.* **92** (7), 957–960 (2007). doi [10.1016/j.ress.2006.07.001](https://doi.org/10.1016/j.ress.2006.07.001).
44. J.-P. Berenger, “A Huygens Subgridding for the FDTD Method,” *IEEE Trans. Antennas Propag.* **54** (12), 3797–3804 (2006). doi [10.1109/TAP.2006.886519](https://doi.org/10.1109/TAP.2006.886519).
45. F. Collino and C. Tsogka, “Application of the Perfectly Matched Absorbing Layer Model to the Linear Elastodynamic Problem in Anisotropic Heterogeneous Media,” *Geophysics* **66** (1), 294–307 (2001). doi [10.1190/1.1444908](https://doi.org/10.1190/1.1444908).
46. V. Lisitsa, “Optimal Discretization of PML for Elasticity Problems,” *Electron. Trans. Numer. Anal.* **30**, 258–277 (2008). <http://eudml.org/doc/130538>. Cited March 29, 2024.
47. V. Kostin, V. Lisitsa, G. Reshetova, and V. Tcheverda, “Local Time–Space Mesh Refinement for Simulation of Elastic Wave Propagation in Multi-Scale Media,” *J. Comput. Phys.* **281**, 669–689 (2015). doi [10.1016/j.jcp.2014.10.047](https://doi.org/10.1016/j.jcp.2014.10.047).
48. A. A. Samarskii, *Theory of Difference Schemes* (Nauka, Moscow, 1983; Marcel Dekker, New York, 2001).
49. P. Moczo, J. Kristek, V. Varyčuk, et al., “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). doi [10.1785/0120010167](https://doi.org/10.1785/0120010167).
50. D. Vishnevsky, V. Lisitsa, V. Tcheverda, and G. Reshetova, “Numerical Study of the Interface Errors of Finite-Difference Simulations of Seismic Waves,” *Geophysics* **79** (4), T219–T232 (2014). doi [10.1190/geo2013-0299.1](https://doi.org/10.1190/geo2013-0299.1).
51. V. Lisitsa, O. Podgornova, and V. Tcheverda, “On the Interface Error Analysis for Finite Difference Wave Simulation,” *Comput. Geosci.* **14** (4), 769–778 (2010). doi [10.1007/s10596-010-9187-1](https://doi.org/10.1007/s10596-010-9187-1).
52. M. Ainsworth, “Discrete Dispersion Relation for *hp*-Version Finite Element Approximation at High Wave Number,” *SIAM J. Numer. Anal.* **42** (2), 553–575 (2004). doi [10.1137/S0036142903423460](https://doi.org/10.1137/S0036142903423460).
53. O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Lecture Notes in Computer Science* (Springer, Cham, 2015), Vol. 9351, pp. 234–241. doi [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
54. I. M. Sobol’, “Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates,” *Math. Comput. Simul.* **55** (1–3), 271–280 (2001). doi [10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6).

55. A. Saltelli, M. Ratto, T. Andres, et al., “Sensitivity Analysis: from Theory to Practice,” in *Global Sensitivity Analysis. The Primer* (John Wiley, Chichester, 2008), pp. 237–275. doi 10.1002/9780470725184.ch6.
56. D. Yeung, I. Cloete, D. Shi, and W. W. Y. Ng, *Sensitivity Analysis for Neural Networks* (Springer, Berlin, 2010). doi 10.1007/978-3-642-02532-7.
57. U. Reuter and M. Liebscher, “Global Sensitivity Analysis in View of Nonlinear Structural Behavior,” https://www.researchgate.net/profile/Martin_Liebscher/publication/242318342_Global_sensitivity_analysis_in_view_of_nonlinear_structural_behavior/links/58e213c5aca272059ab08f35/Global-sensitivity-analysis-in-view-of-nonlinear-structural-behavior.pdf. Cited March 29, 2024.
58. J. Morio, “Global and Local Sensitivity Analysis Methods for a Physical System,” *Eur. J. Phys.* **32** (6), 1577–1583 (2011). doi 10.1088/0143-0807/32/6/011.
59. A. Saltelli and S. Tarantola, “On the Relative Importance of Input Factors in Mathematical Models: Safety Assessment for Nuclear Waste Disposal,” *J. Am. Stat. Assoc.* **97** (459), 702–709 (2002). doi 10.1198/016214502388618447.
60. D. W. van der Merwe and A. P. Engelbrecht, “Data Clustering Using Particle Swarm Optimization,” in *Proc. 2003 Congress on Evolutionary Computation, Canberra, Australia, December 08–12, 2003*. Published by IEEE Press in Congress on Evolutionary Computation, Vol. 1, pp. 215–220 (2003): <https://ieeexplore.ieee.org/xpl/conhome/9096/proceeding?isnumber=28874&sortType=vol-only-seq&pageNumber=2>. Cited March 29, 2024. doi 10.1109/CEC.2003.1299577.

Received
January 31, 2024

Accepted for publication
March 16, 2024

Information about the authors

Elena A. Gondyul – Engineer; Trofimuk Institute of Petroleum Geology and Geophysics SB RAS, Ac. Koptyug prospekt, 3, 630090, Novosibirsk, Russia.

Vadim V. Lisitsa – Dr. Sci., Head of Laboratory; Trofimuk Institute of Petroleum Geology and Geophysics SB RAS, Ac. Koptyug prospekt, 3, 630090, Novosibirsk, Russia.

Kirill G. Gadylshin – Ph.D., Senior Research Scientist; Trofimuk Institute of Petroleum Geology and Geophysics SB RAS, Ac. Koptyug prospekt, 3, 630090, Novosibirsk, Russia.

Dmitry G. Vishnevsky – Ph.D., Research Scientist; Trofimuk Institute of Petroleum Geology and Geophysics SB RAS, Ac. Koptyug prospekt, 3, 630090, Novosibirsk, Russia.