

К. Г. Гадьльшин^{1}, В. В. Лисица¹, К. А. Гадьльшина¹, М. Д. Вишневецкий¹*

Подавление численной дисперсии в частотной области на основе глубокого обучения

¹ Институт нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН,
г. Новосибирск, Российская Федерация
* e-mail: GadylshinKG@ipgg.sbras.ru

Аннотация. Предлагается метод подавления численной дисперсии в частотной области, осуществляемый искусственной нейронной сетью NDM-net. Обработка в частотной области имеет несколько преимуществ в сравнении с обработкой во временной области, в частности, позволяет извлекать больше информации из больших объёмов данных, эффективнее бороться с их зашумленностью и корректировать амплитудные и фазовые искажения. Для демонстрации подхода используется реалистичная синтетическая модель ВР, при проведении численных экспериментов численная дисперсия снижается в три раза, а производительность вычислений увеличена на порядок в сравнении с традиционными подходами. Обработка в частотной области осуществляется для всей сейсмограммы одновременно, чего не позволяла реализация метода уменьшения численной дисперсии с применением NDM-net во временной области. В целом, применение NDM-net в частотной области повышает качество и эффективность обработки сейсмических данных.

Ключевые слова: NDM-net, частотная область, глубокое обучение, сейсмическое моделирование, численная дисперсия, преобразование Фурье

K. G. Gadylshin^{1}, V. V. Lisitsa¹, K. A. Gadylshina¹, M. D. Vishnevsky¹*

Numerical dispersion mitigation network in time-frequency domain

¹ Institute of Petroleum Geology and Geophysics SB RAS, Novosibirsk, Russian Federation
* e-mail: GadylshinKG@ipgg.sbras.ru

Abstract. We propose a method for suppressing numerical dispersion in the frequency domain, implemented by an artificial neural network NDM-net. Frequency-domain processing has several advantages over time-domain processing, such as extracting more information from large amounts of data, dealing with data noise more effectively, and correcting amplitude and phase distortions. To demonstrate the approach, a realistic synthetic BP model is used; in numerical experiments, the numerical variance is reduced by a factor of three, and the computational performance is increased by an order of magnitude in comparison with traditional approaches. Processing in the frequency domain is carried out for the entire seismogram simultaneously, which was not allowed by the implementation of the numerical dispersion reduction method using NDM-net in the time domain. In general, the use of NDM-net in the frequency domain improves the quality and efficiency of seismic data processing.

Keywords: NDM-net, frequency domain, deep learning, seismic modeling, numerical dispersion, Fourier transform

Введение

Одним из способов снизить вычислительные ресурсы и уменьшить время на моделирование сейсмограмм является уменьшение размерности задачи за счёт

увеличения шага пространственной численной сетки, что, в свою очередь, влечёт увеличение ошибки, проявляющейся в виде численной дисперсии. Таким образом, вычисленные на грубых сетках сейсмограммы требуют постобработки, которая производится, в частности, при помощи искусственной нейронной сети, как это предлагается в статьях [1, 2]. В предыдущей работе [3] описывается глубокая свёрточная нейронная сеть NDM-net, предназначенная для подавления численной дисперсии в предварительно смоделированных волновых полях, записанных на свободной поверхности.

С целью моделирования реалистичных сейсмических данных, волновые поля должны рассчитываться для большого числа источников, обычно их количество насчитывает порядка 10^4 . При плотном расположении источников сейсмограммы для двух соседних источников практически неотличимы друг от друга [4]. Эта особенность модели позволяет для обучения искусственной нейронной сети строить репрезентативный набор сейсмограмм, соответствующих небольшому числу источников. Моделирование для этих источников произведено с использованием очень мелкой сетки, что позволяет получать близкое к точному решение. В предыдущей работе [3] этот подход применяется к двумерным моделям, и показано, что численная ошибка в данных, вычисленных на грубой сетке, существенно уменьшается при постобработке искусственной нейронной сетью, для обучения которой использовались сейсмограммы от 10% источников. Так же демонстрируется, насколько сократилось время расчета полного набора сейсмограмм.

Анализ различных частотных компонент сейсмического сигнала позволяет идентифицировать особенности недр, которые трудно обнаружить во временной области [5]. Существует множество подходов в частотной области, которые улучшают качество данных за счет подавления шума и коррекции амплитудных и фазовых искажений, вызванных системой сбора сейсмических данных [6, 7]. Еще одним преимуществом является то, что обработка в частотной области обеспечивает эффективный способ извлечения информации из сейсмических данных при работе с большими наборами данных.

Теория и метод

В наших предыдущих работах по применению NDM-net во временной области накладывались жесткие условия на длину сейсмограммы и не всегда было возможным обрабатывать сейсмограмму целиком вследствие того, что оперативная память графических ускорителей ограничена, и мы вынуждены были разбивать сейсмограммы вдоль временной оси таким образом, чтобы количество отсчетов в одном фрагменте данных было равно 512. Обычно при моделировании используется вейвлет Рикера с доминирующей частотой порядка 25-30 Гц, так как сейсмические данные имеют ограниченный частотный диапазон. Это приводит к тому, что в частотной области основная энергия сигнала сконцентрирована в полосе вокруг пиковой частоты зондирующего сигнала.

В данной работе мы будем использовать NDM-net, описанный в работе [3]. Эта нейронная сеть является модификацией сети U-net, полностью свёрточной

нейронной сети. Основным отличием будет являться вход и выход нейронной сети. Теперь данные – это преобразование Фурье по времени от сейсмограммы во временной области, где в окрестности пиковой частоты выбирается 512 значащих частот. Таким образом тензор имеет размерность $512 \times 512 \times 2$: 512 частот, 512 приемников и 2 компоненты – вещественная и мнимая.

Численный эксперимент

Численный эксперимент был проведен для реалистичной синтетической модели ВР (рис. 1). Левая часть модели построена с использованием геологического разреза, проходящего через западную часть Мексиканского залива. В центральной части модели изображены упрощенные представления геологических структур в восточной/центральной части Мексиканского залива и на шельфе Анголы. Правая часть модели представляет собой комбинацию структур, встречающихся в таких регионах, как Каспийское море, Северное море или Тринидад. Эта модель часто используется в качестве эталона для глубинной миграции до суммирования, таких как миграция в обратном времени.

Модель простирается на 67.4 км по горизонтали и на 11.9 км по вертикали. Геометрия системы наблюдения состоит из 2696 источников, расположенных на равном расстоянии друг от друга по поверхности воды с шагом 25 м. Для каждого источника мы генерируем синтетическую сейсмограмму, записанную 512 приемниками, максимальное удаление источник-приемник составляет 6.4 км. Форма импульса зондирующего сигнала – вейвлет Рикера с пиковой частотой 30 Гц. Для моделирования мы использовали конечно-разностный алгоритм 4-го порядка по пространству с поглощающими граничными условиями на всех сторонах модели. Мы рассчитывали решения на мелкой и грубой численных сетках с шагом 3 м и 6 м соответственно.

Используя характеристики скоростной модели ВР (см. черную кривую на рис. 1), мы проводим дополнительное преобразование во временной области, чтобы выровнять времена первого отражения от дна для всех положений источников. Поскольку в модели скорость в воде остается постоянной, любые изменения скорости начинаются ниже дна воды. Это позволяет нам вычислить двойное время по вертикали от источника до соответствующего положения на дне воды и соответствующим образом ввести статические поправки в данные.

Обучение NDM-net проводилось на сервере, оснащенный двумя графическими ускорителями NVIDIA RTX3090. В качестве функции потерь была выбрана среднеквадратическая ошибка. Количество эпох в процессе обучения было равно 50, что заняло примерно 45 минут. Впоследствии мы использовали веса NDM-net, полученные после 47-й эпохи, для проведения численного подавления дисперсии по всему набору данных, который был рассчитан на грубой сетке. Результаты прогнозирования для случайной сейсмограммы из тестового набора данных можно просмотреть на рис. 2. Примечательно, что после постобработки NDM-net наблюдается значительное уменьшение численной дисперсии.

В этой работе мы рассчитали относительную ошибку между решением на мелкой сетке и решением на грубой сетке (или его NDM-net аппроксимацию) в процентах, используя следующую формулу:

$$error_{relative}(A_i, B_i) = \frac{norm(A_i - B_i)}{norm(B_i)} \times 100\%, \quad (1)$$

здесь A_i представляет решение на грубой сетке для источника с индексом i , B_i представляет соответствующее решение на мелкой сетке, а $norm$ представляет евклидову норму. Евклидова норма является широко используемой метрикой для измерения ошибки между двумя векторами. В данном случае мы использовали его для расчета относительной ошибки между решениями. Результирующий график ошибок, который показывает уровень численной дисперсии, вызванной грубой конечно-разностной аппроксимацией в зависимости от положения источника, показан на рис. 3 красным графиком. График указывает на значительный уровень численной дисперсии в решении на грубой сетке, о чем свидетельствует высокая относительная ошибка.

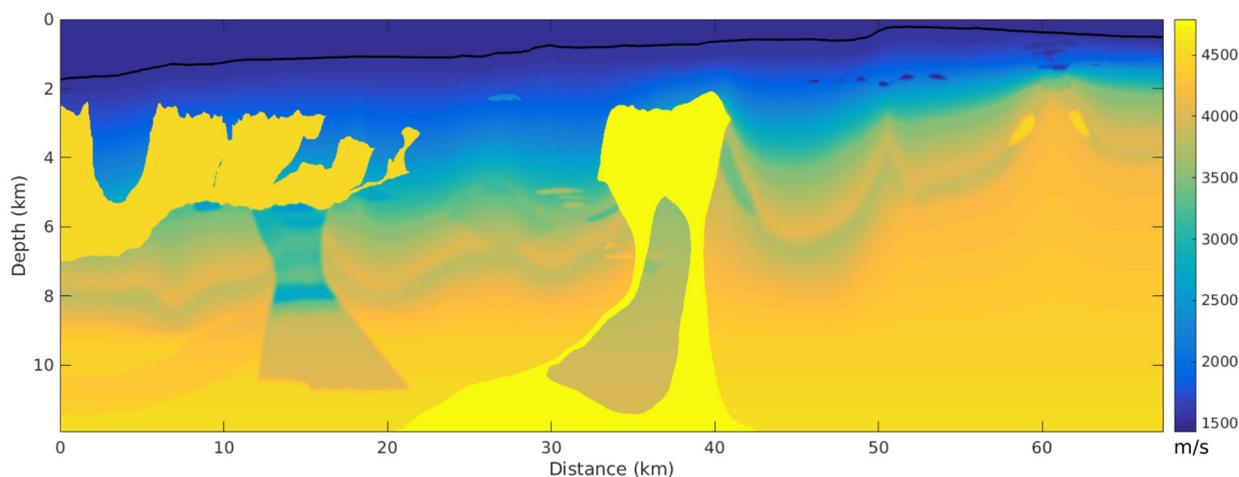


Рис. 1. Скоростная модель ВР. Черной линией в верхней части рисунка обозначен первый отражающий горизонт – морское дно

После выполнения пост-обработки NDM-net и уменьшения численной дисперсии в данных мы смогли построить ту же кривую относительной ошибки между данными, предсказанными NDM-net, и решением на мелкой сетке. Прогнозируемая ошибка сети NDM представлена синей кривой на рис. 3 со средним значением 31.49% и стандартным отклонением 7.13%. Таким образом, мы делаем вывод, что уровень численной дисперсии уменьшился в три раза после NDM-net аппроксимации.

Чтобы лучше сравнить решения на грубой сетке и аппроксимацию NDM-net, мы выбрали два случайных индекса источника из набора тестовых данных, которые сеть NDM-net не использовала во время обучения. На рис. 4 показано преобразование сейсмограмм из частотной области обратно во временную область для этого источника. Следуя тождеству Парсеваля, утверждающему, что евклидова норма сейсмограммы после преобразования Фурье равна ее норме во

временной области, мы наблюдаем аналогичный уровень численной дисперсии во временной области.

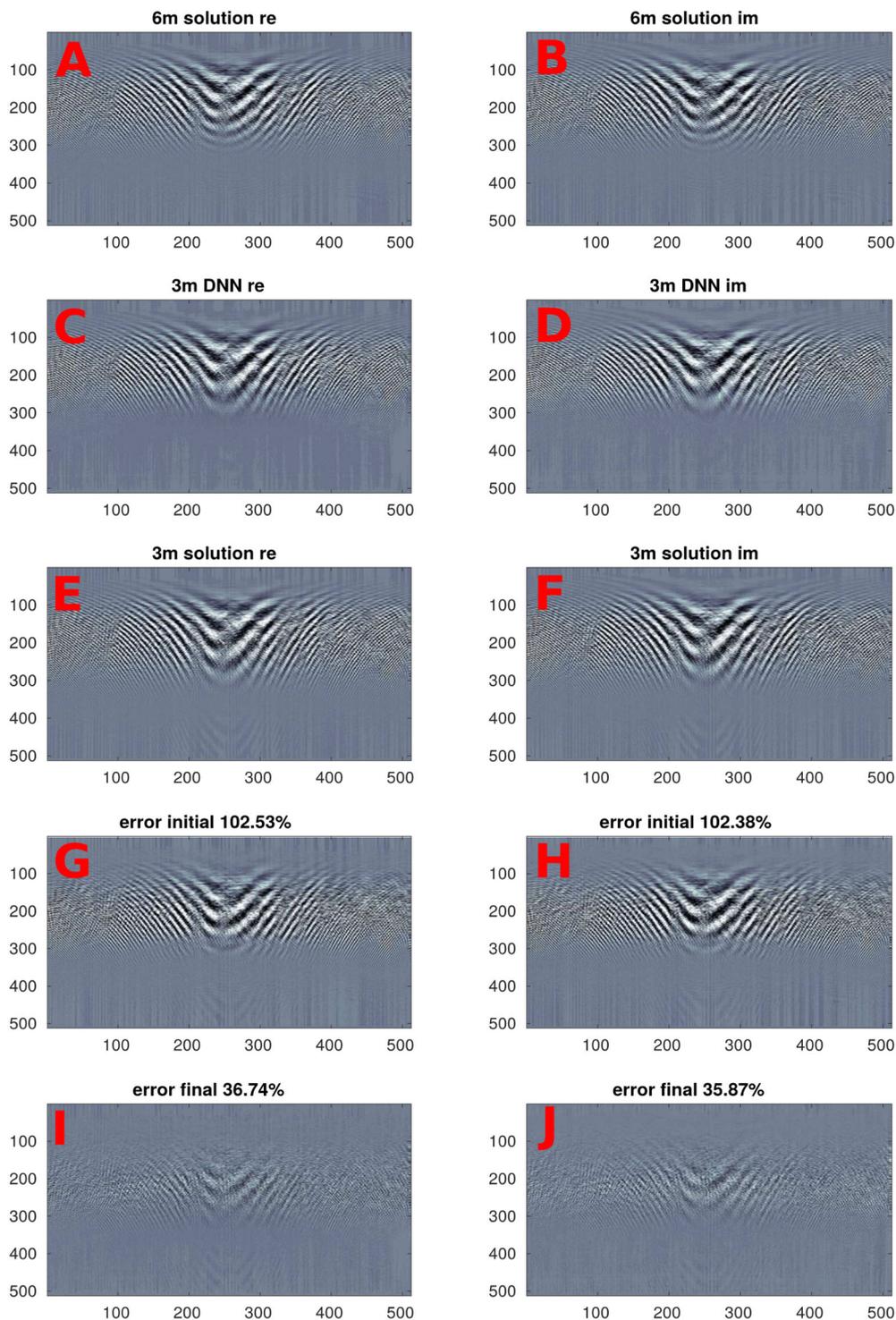


Рис. 2. NDM-net аппроксимация для источника, расположенного в точке $X=3.2\text{км}$. Левая колонка – вещественные части, а правая колонка – мнимые части решения, рассчитанного на грубой сетке (A, B), на мелкой сетке (E, F), NDM-net аппроксимация (C, D), дисперсия до подавления (G, H) и дисперсия после NDM-net аппроксимации

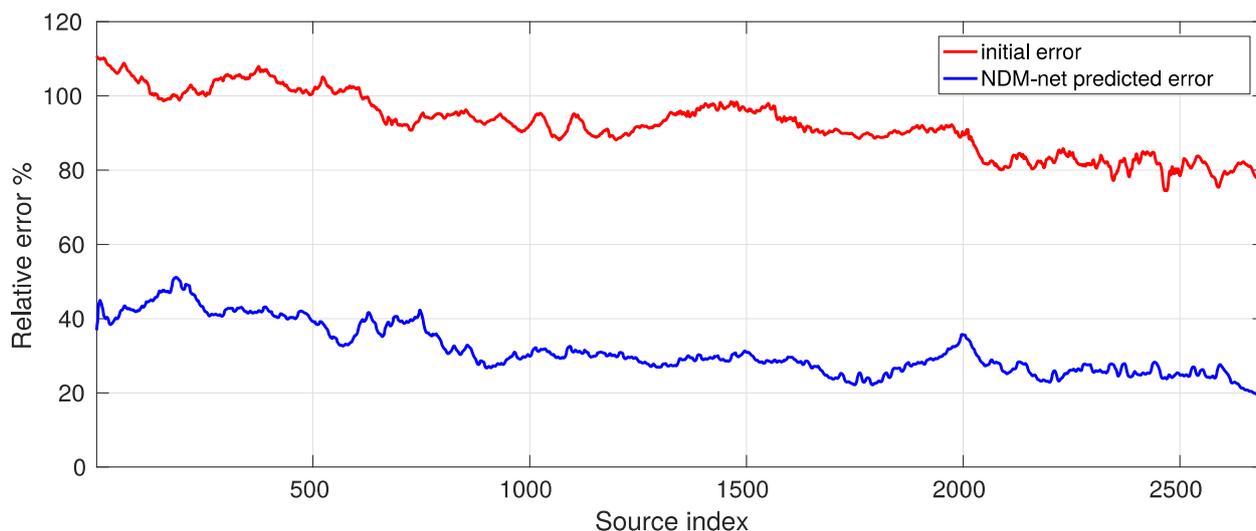


Рис. 3. Относительная ошибка между сейсмограммами, рассчитанными на мелкой и грубой сетках (красный график), и относительная ошибка между сейсмограммами, рассчитанными на мелкой сетке и пост-обработанными NDM-net сейсмограммами (синий график)

Мы также провели сравнение сейсмических трасс во временной и частотной областях для большей наглядности. Эти сравнения показывают, что аппроксимация NDM-net и решение на мелкой сетке имеют хорошую корреляцию как в частотной, так и во временной областях (см. рис. 4F). Красный график представляет решение на сетке 6 м, черный график представляет решение на сетке 3 м, а синий график представляет аппроксимацию NDM-net. Мы наблюдаем хорошее соответствие между сейсмической трассой, рассчитанной на мелкой сетке, и аппроксимацию NDM-net, что говорит об эффективности NDM-net пост-обработки сейсмических данных в контексте снижения уровня численной дисперсии.

Заключение

Представлен подход к подавлению численной дисперсии при сейсмическом моделировании на основе NDM-net в частотной области. В отличие от обработки во временной области обработка в частотной области предлагает несколько преимуществ, в частности позволяет обрабатывать большие сейсмограммы целиком, что не всегда возможно во временной области. На примере реалистичной синтетической модели ВР было достигнуто снижение численной дисперсии в среднем в три раза и увеличение производительности вычислений на порядок.

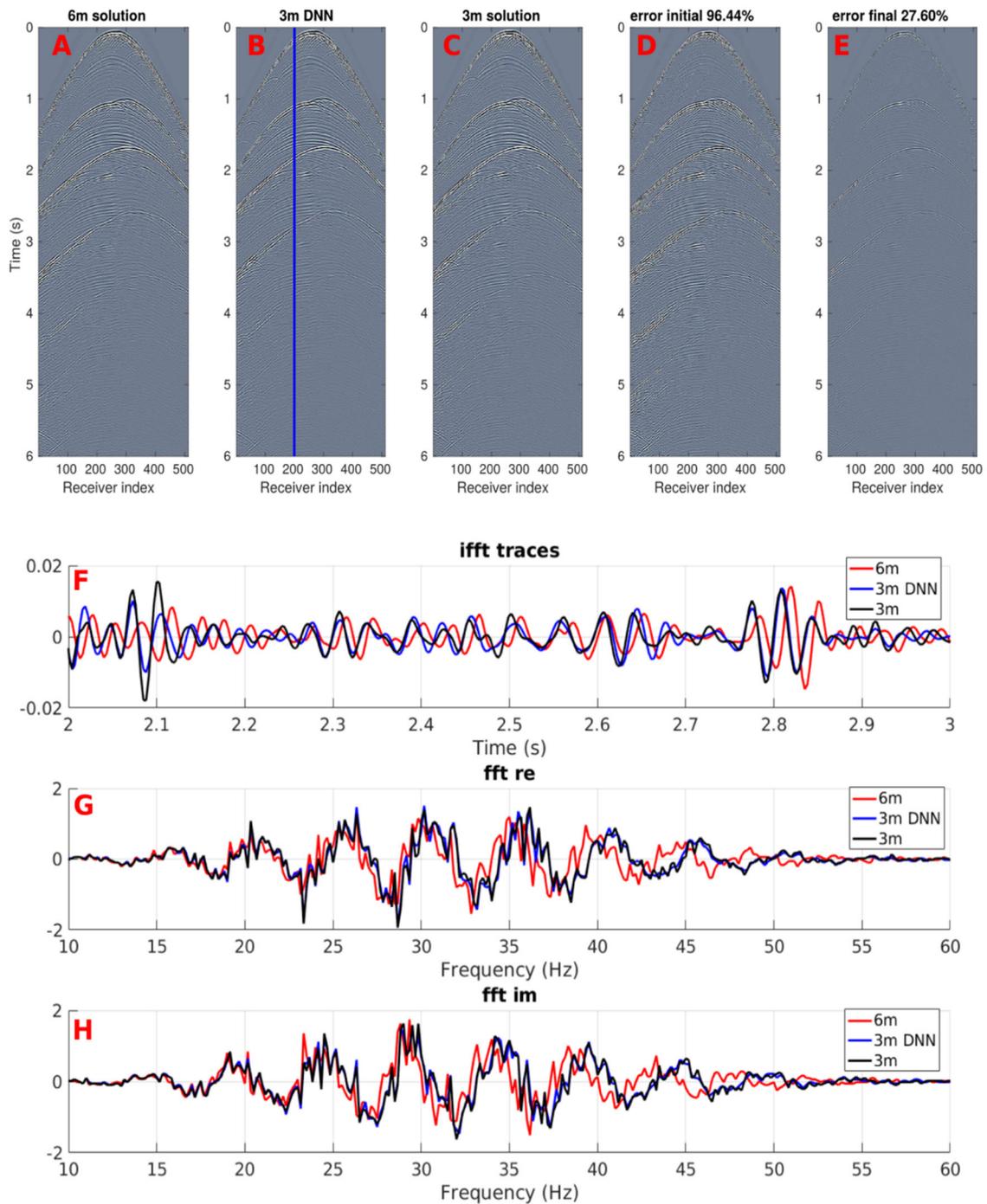


Рис. 4. Результаты работы сети NDM для источника, расположенного в точке $X=3,2$ км, преобразованные обратно из частотной области во временную. Сейсмограмма, смоделированная на грубой (A) и мелкой сетке (C), начальная численная дисперсия (D) (разница между (C) и (A)), сейсмограмма после постобработки NDM-net (B), и подавленная численная дисперсия (E) (разница между (C) и (B)). Сравнение сейсмических трасс во временной области (F) (индекс трассы 200, смотри вертикальную синюю линию на рис. (B)); сравнение трасс в частотной области: вещественная (G) и мнимые (H) части, соответственно. Красная линия – решение на грубой (6м) сетке, черная линия – на мелкой сетке (3м), и синяя линия – NDM-net аппроксимация

Благодарности

К. Гадыльшин разработал архитектуру сети, а Д. Вишнеvский выполнил конечно-разностное моделирование волновых полей в рамках работ по гос. заданию № FWZZ-2022-0022. В. Лисица разработал алгоритм подавления численной дисперсии в частотной области, а К. Гадыльшина провела численные эксперименты при поддержке РФФ (грант № 22-11-00004).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Siahkoohi A., Louboutin M., Herrmann F.J. The importance of transfer learning in seismic modeling and imaging // *Geophysics*. – 2009. – Vol. 84. – P. A47–A52.
2. Kaur H., Fomel S., Pham N. Overcoming numerical dispersion of finite-difference wave extrapolation using deep learning // *SEG Technical Program Expanded Abstracts*. – 2019. P. 2318–2322.
3. Gadylyshin K., Vishnevsky D., Gadylyshina K., Lisitsa V. Numerical dispersion mitigation neural network for seismic modeling // *Geophysics*. – 2022. Vol. 87(3). – P. T237–T249.
4. Gadylyshin K., Lisitsa V., Gadylyshina K., Vishnevsky D. Optimization of the training dataset for numerical dispersion mitigation neural network // *Computational Science and Its Applications – ICCSA 2022 Workshops*. – 2022. P. 295–309.
5. Yilmaz O., Doherty S. *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data* // Crisp Fifty-Minute Books Society of Exploration Geophysicists. – 2001. – Vol. 10. – No. 10.
6. Elboth T., Geoteam F., Qaisrani H.H., Hertweck T. De-noising seismic data in the time-frequency domain // *SEG Expanded abstracts*. – 2008. – P. 2622–2626.
7. Zhao Y., Liu Y., Li X., Jiang N.: Time–frequency domain snr estimation and its application in seismic data processing // *Journal of Applied Geophysics*. – 2014. – Vol. 107. – P. 25–35.

© К. Г. Гадыльшин, В. В. Лисица, К. А. Гадыльшина, Д. М. Вишнеvский, 2023