# Оценка неоднозначности мультимодальной инверсии фазовой скорости поверхностной волны с использованием машинного обучения

 $A. B. Яблоков^{1,2,3*}, A. C. Сердюков^{1,2,3}$ 

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

<sup>2</sup> Институт горного дела им. Н.А. Чинакала СО РАН, , г. Новосибирск, Российская Федерация <sup>3</sup> Новосибирский государственный университет, г. Новосибирск, Российская Федерация \*e-mail: yablokovav@ipgg.sbras.ru

Аннотация. Работа посвящена количественной оценке неопределённости решения обратной задачи в методе многоканального анализа поверхностных волн — инверсии кривых зависимости фазовой скорости от частоты. Подход к оценке неопределённости основан на применении стратегии семплирования Монте-Карло и использовании многослойной полносвязной искусственной нейронной сети для аппроксимации нелинейной зависимости скорости поперечной волны и мощности слоёв от значений фазовой скорости поверхностной волны Релея. На результат решения обратной задачи проецируются частотно-зависимый шум в данных и ошибки обратного оператора. Сопоставляются результаты унимодальной и многомодальной инверсии на примере обработки синтетических данных. Результаты экспериментов показывают, что использование подходов машинного обучения даёт возможность быстро и точно оценить апостериорную плотность вероятности восстановленных параметров скоростной модели.

**Ключевые слова:** поверхностные волны, инверсия, искусственная нейронная сеть, неоднозначность решения обратной задачи

# Uncertainty quantification of phase velocity surface waves multy-modal inversion using machine learning

A. V. Yablokov<sup>1,2,3,\*</sup>, A. S. Serdyukov<sup>1,2,3</sup>

<sup>1</sup> Trofimuk Institute of Petroleum Geology and Geophysics SB RAS, Novosibirsk, Russian Federation

<sup>2</sup> Chinakal Institute of Mining SB RAS, Novosibirsk, Russian Federation <sup>3</sup>Novosibirsk State University, Novosibirsk, Russian Federation \*e-mail: yablokovav@ipgg.sbras.ru

**Absract.** The paper is devoted to uncertainty quantification of the inverse problem solution of the multichannel analysis of surface waves method – the inversion of the curves of the phase velocity via frequency dependence. The uncertainty estimation approach is based on the Monte Carlo sampling strategy and a multilayer fully connected artificial neural network to approximate nonlinear dependence of shear wave velocity and layers thickness via values of phase velocity surface waves. Frequency-dependent noise in the data and errors of the inverse operator are projected onto the inverse problem solution. The results of unimodal and multimodal inversion are compared on the example of synthetic data processing. The experimental results show that using of machine learning approaches makes it possible to quickly and accurately estimate the posterior probability density of the reconstructed velocity model parameters.

Keywords: surface waves, inversion, artificial neural network, uncertainty of the inversion problem

#### Введение

В работе рассматривается ключевая проблема метода многоканального анализа поверхностных волн (MASW) — одновременное восстановление значений скорости поперечной волны ( $V_S$ ) и мощности (h) путём обращения (инверсии) дисперсионных кривых фазовой скорости ( $V_R$ ) [1]. Преимуществами использования искусственных нейронных сетей (ИНС) для решения обратной задачи метода MASW являются их высокая вычислительная эффективность (при использовании подобранной архитектуры) и отсутствие необходимости включения модельных ограничений в функционал невязки, в отличие от методов глобальной оптимизации [2; 3]. Инверсия дисперсионных кривых поверхностных волн является некорректной и неоднозначной задачей [4]. Для снижения неоднозначности целесообразно включать в решение обратной задачи моды высшего порядка [5; 6]. Основная нерешенная проблема при этом — правильная идентификация различных мод на спектре, осложненным интерференционными помехами [7].

Решение обратной задачи метода MASW в слоистой постановке носит вероятностный характер, что требует оценки области эквивалентности (неоднозначности) восстановленных параметров скоростной модели или т.н. апостериорной плотности вероятности (PPD). Для решения такой задачи успешно применяется метод Монте-Карло с цепями Маркова [8]. Однако, такой подход требует больших вычислительных ресурсов из-за большого (около 10 000) количества отсчетов (решений обратной задачи), необходимых для достижения стабильного PPD. В работе [2] предложена стратегия семплирования Монте-Карло с применением остаточной нейронной сети (ResNet) для обращения двумерных амплитудных спектров в пространстве фазовая скорость – частота и восстановления скоростной модели через оценку PPD. В настоящей работе используется стратегия семплирования Монте-Карло и алгоритм инверсии на основе ИНС [3], развитый до многомодального (одновременная инверсия нескольких мод) подхода. Обученная ИНС позволяет спроецировать шум дисперсионных данных и ошибки обратного оператора на пространство параметров скоростной модели. Таким образом, адаптируется стратегия семплирования Монте-Карло для оценки неопределенности многомодальной инверсии с использованием ИНС. В работе представлены результаты тестирования алгоритма оценки неоднозначности решения обратной задачи и сопоставления унимодального и многомодального подходов на примере обращения синтетических дисперсионных кривых.

### Методы

Двумерный амплитудный спектр в пространстве фазовая скорость — частота (V-f спектр) рассчитывается путём SFK-преобразования сейсмограмм [9]. Значения фазовых скоростей ( $V_{R,obs}^{i,j}$ , где i=[1,2,...,n] — номера мод, j=[1,2,...,m] — номера частот) извлекаются вручную пикированием максимумов амплитуд визуально различаемых мод. Для дальнейшего выполнения многомодальной инверсии необходима одинаковая длина векторов фазовых скоростей различных

мод. Для этого недостающие значения высших мод на нижних частотах заполняются нулями, что не несет физического смысла, а является техническим требованием. Матрица извлечённых мод поверхностной волны преобразуется в вектор последовательно записанных мод в порядке возрастания номера моды:  $V_{R,obs} = [V_{R,obs}^{i,1}, V_{R,obs}^{i,2}, ..., V_{R,obs}^{i,n}]$ .

Алгоритм решения обратной задачи подробно изложен в статье [3]. Модификация алгоритма заключается в расчёте обучающей выборки для заранее заданного ограниченного числа мод и изменении формата входных данных для подбора и применения весов ИНС. Входной слой ИНС имеет формат извлечённых дисперсионных кривых:  $V_{R,train} = [V_{R,train}^{i,1}, V_{R,train}^{i,2}, \dots, V_{R,train}^{i,n}]$  и длину  $p = n \cdot m$ . Выходной слой – конкатенация векторов  $[V_S, h]$ , соответствующих рассчитанным фазовым скоростям, имеет длину 2l-1, где l- число слоёв.

Для корректной оценки неоднозначности решения обратной задачи шум в данных и ошибки обратного оператора проецируются на пространство решения [10]. Предлагается ввод в значения фазовой скорости частотно-зависимого шума из нормального распределения  $\mathcal{N}(0, \sigma_{\text{шума}}(f))$ , размер которого  $p \cdot u$  (для статистически значимой выборки u=10~000 примеров). Границы  $\sigma_{\text{шума}}(f)$  – пределы ошибок построения кривой фазовой скорости, к которым спектральные амплитуды затухают вдвое относительно максимума на фиксированной частоте. Для определения ошибок обратного оператора применяются веса обученной ИНС (W) для обращения всего обучающего набора фазовых скоростей (D) и расчета ошибки: E=M-W(D), где M — набор обучающих параметров скоростной модели, соответствующий D. Ошибки обратного оператора предполагаются нормально распределёнными —  $\mathcal{N}(\mu_{\text{ошибки}}, \sigma_{\text{ошибки}})$ ,  $\mu_{\text{ошибки}}$  и  $\sigma_{\text{ошибки}}$  вычисляются из E.

## Результаты экспериментов

Для тестирования алгоритма и сопоставления результатов унимодальной и многомодальной инверсии рассматривается горизонтально-слоистая скоростная модель с тремя слоями:  $V_S = [200, 300, 500], \ V_P = [400, 600, 1000], \ \rho = [1800, 1930, 2110], \ h = [3, 7, \infty].$  Для такой модели методом матричных пропагаторов рассчитываются 4 моды фазовых скоростей. Сейсмограмма, объединяющая волновые формы рассчитанных мод, и её V-f спектр изображены на рис.1. Сейсмограмма также содержит случайный шум Гаусса с SNR = -2 дБ. Красными доверительными интервалами изображены пределы ошибки пикирования для определения  $\sigma_{\text{шума}}(f)$  (длина интервала  $= 2 \cdot 3\sigma_{\text{шума}}(f)$ ).

В дальнейшем эксперимент развивается параллельно для двух случаев: 1 — использование только фундаментальной моды и 2 — использование всех 4х мод. Архитектура ИНС и количество примеров в обучающей выборке одинаковы в обоих случаях. После извлечения дисперсионных кривых и подготовки данных в формате, описанном выше, рассчитывается тренировочный набор данных: 100 000 равномерно выбранных скоростных моделей в пределах оценённых диапазонов и соответствующих им кривых фазовых скоростей. Кривые обучения

ИНС, т.е. значения функционала невязки на каждой эпохе, для случаев использования различного количества мод поверхностной волны представлены на рис. 2. Значение функционала невязки (средняя абсолютная ошибка, МАЕ) после 300 эпох обучения при использовании только фундаментальной моды равна 0.034, а при использовании 4х мод 0.014. Меньшие значения невязок говорят о более глубоком обучении ИНС при использовании дополнительных (высшие моды) данных.

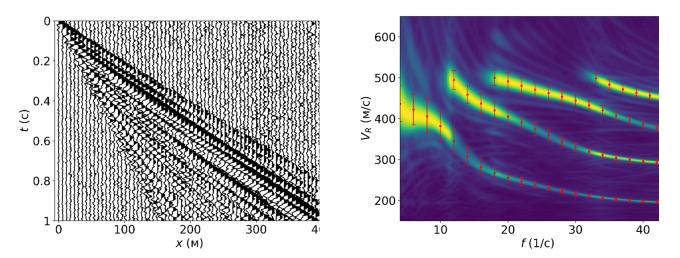


Рис. 1. Сейсмограмма, содержащая 4 моды фазовых скоростей поверхностной волны, (слева) и её двумерный амплитудный V-f спектр (справа).

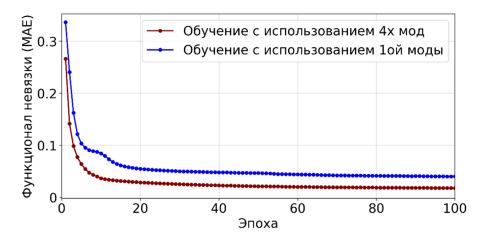


Рис. 2. Кривая обучения ИНС при различных тренировочных данных: для одной (фундаментальной) моды и 4х мод.

После подбора весов ИНС и применения алгоритма семплирования Монте-Карло [10] для инверсии дисперсионных кривых определяются распределения восстановленных значений скорости S-волны  $V_S$  и мощности h (показаны цветом на рис. 3, яркий цвет — большее число попаданий в точку на пространстве глубина-скорость). Математическое ожидание полученного распределения — искомая скоростная модель S-волны (показана серой кривой на рис. 3).

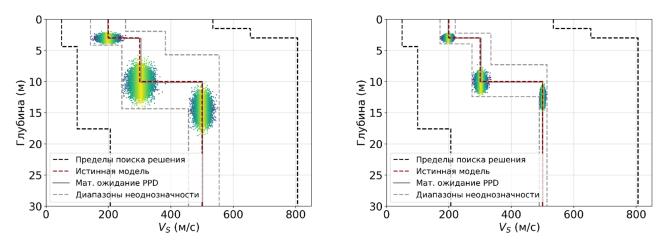


Рис. 3. Результаты количественного определения неопределенности: распределения ошибок S-волн для унимодальной (слева) и многомодальной (справа) инверсии. Более яркий цвет показывает большее число попаданий в точку на пространстве глубина-скорость

Апостериорная плотность вероятности (PPD) вычисляется ядерной оценкой плотности (KDE) для распределений  $V_S$  и h. Вычисленные значения PPD для унимодальной и многомодальной инверсии показаны на рис. 4. Из сопоставления результатов следует вывод об уменьшении области неоднозначности и увеличении достоверности определения всех параметров при переходе от унимодальной к многомодальной инверсии.

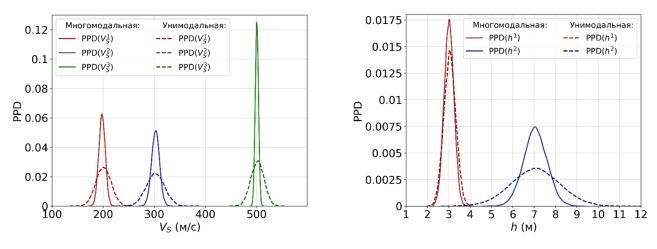


Рис. 4. PPD для скорости S-волны (слева) и мощности (справа).

Результаты определения диапазонов неоднозначности восстановленных параметров скоростной модели обобщены в табл. 1 как отклонения значений  $V_S$  и h от их математического ожидания PPD в процентах в каждом слое. В среднем наблюдается увеличение достоверности определения  $V_S$  в 2.2 раза, а h в 1.5 раза.

Номер слоя	Унимодальная инверсия		Мультимодальная инверсия	
	$V_S$	h	$V_S$	h
1	± 20%	± 40%	± 10%	± 30%
2	± 19%	± 44%	± 11%	± 27%
3	± 9%		± 3%	

#### Заключение

В работе предложен подход к количественной оценке неопределенности восстановленных значений параметров скоростной модели S-волны путём многомодальной инверсии дисперсионных кривых поверхностных волн с помощью ИНС и стратегии семплирования Монте-Карло. Результаты синтетических экспериментов показывают, что использование ИНС дает возможность точно оценить PPD параметров модели без значительных вычислительных затрат, а математическое ожидание PPD дает точное решение обратной задачи. Преимущество многомодальной инверсии над одномодальной подтверждается сопоставлением расчетных PPD. Алгоритм, реализованный в этой работе, может быть применен к любым параметрическим или непараметрическим распределениям ошибок. Вычислительная стоимость применения обученной ИНС пренебрежимо мала и растёт незначительно с увеличением числа дисперсионных кривых.

# Благодарности

Исследование выполнено при финансовой поддержке РНФ в рамках научного проекта № 20-77-10023.

### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- 1. Miller R. D. et al. Multichannel analysis of surface waves to map bedrock //The Leading Edge. -1999. Vol. 18. No. 12. P. 1392-1396.
- 2. Aleardi M., Stucchi E. A hybrid residual neural network–Monte Carlo approach to invert surface wave dispersion data //Near Surface Geophysics. 2021. Vol. 19. №. 4. P. 397-414.
- 3. Yablokov A. V. et al. An artificial neural network approach for the inversion of surface wave dispersion curves //Geophysical Prospecting. 2021. Vol. 69. №. 7. P. 1405-1432.
  - 4. Foti S. et al. Surface wave methods for near-surface site characterization. CRC press, 2014.
- 5. Xia J. et al. Inversion of high frequency surface waыves with fundamental and higher modes //Journal of Applied Geophysics.  $-2003. \text{Vol.} 52. \text{N}\underline{\circ}. 1. \text{P.} 45-57.$
- 6. Luo Y. et al. Joint inversion of high-frequency surface waves with fundamental and higher modes //Journal of Applied Geophysics. 2007. Vol. 62. №. 4. P. 375-384.
- 7. Zhang S. X., Chan L. S. Possible effects of misidentified mode number on Rayleigh wave inversion //Journal of Applied Geophysics. 2003. Vol. 53. №. 1. P. 17-29.
- 8. Aleardi M., Salusti A., Pierini S. Transdimensional and Hamiltonian Monte Carlo inversions of Rayleigh-wave dispersion curves: a comparison on synthetic datasets //Near Surface Geophysics.  $-2020. \text{Vol. } 18. \cancel{N}_{2}. 5. \text{P. } 515-543.$

- 9. Serdyukov A. S. et al. Slant f-k transform of multichannel seismic surface wave data //Geophysics. -2019. Vol. 84. N0. 1. P. A19-A24.
- 10. Яблоков А. В., Сердюков А. С. Исследование неоднозначности обращения дисперсионных кривых поверхностных волн при помощи искусственных нейронных сетей //Интерэкспо Гео-Сибирь.  $-2021.-T.\ 2.-N$  $2.-C.\ 82-89.$

© А. В. Яблоков, А. С. Сердюков, 2022