

СТАТИСТИЧНА КОРЕКЦІЯ ПРИПОВЕРХНЕВИХ СПОТВОРЕНЬ ДЛЯ ПІДВИЩЕННЯ НАДІЙНОСТІ ІНТЕРПРЕТАЦІЇ СЕЙСМІЧНИХ ДАННИХ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

О. В. Піскунов

аспірант

О. К. Тяпкін

докт. геол. наук

Національний технічний університет «Дніпровська політехніка»,
49000, Дніпро, пр. Д.Яворницького, 19

Надійність інструментів машинного навчання для сейсмічної інтерпретації обмежується якістю вхідних даних, яка часто погіршується через спотворення у верхній частині геологічного розрізу (ВЧР). У роботі представлено статистичний метод для характеристики та корекції цих спотворень шляхом декомпозиції сейсмічних амплітуд. Точність цього методу підтверджено на синтетичній моделі, що доводить його здатність кількісно оцінювати властивості приповерхневих шарів. Далі, ми використовуємо контекстно-залежну мережу з довгою короткочасною пам'яттю (LSTM) як приклад, демонструючи, що такі інструменти можуть проводити інтерпретацію на «чистих», геологічно складних даних. Зроблено висновок, що гібридний робочий процес, де надійна, фізична підготовка даних передусе аналізу за допомогою ШІ, є необхідною умовою для успішного застосування машинного навчання в реальних сейсмічних задачах.

Ключові слова: статистична обробка, глибинне навчання, геологічна інтерпретація сейсмічних даних, нейронні мережі, обробка сейсмічних даних.

STATISTICAL CORRECTION OF NEAR-SURFACE DISTORTIONS FOR INCREASING THE RELIABILITY OF SEISMIC DATA INTERPRETATION USING MACHINE LEARNING

O. V. Piskunov

PhD student

O. K. Tiapkin

Doctor of Geological Sciences, Professor

Dnipro University of Technology, 49005, Dnipro, D. Yavornytskoho ave., 19

The reliability of ML tools for seismic interpretation is fundamentally limited by input data quality, which is often degraded by distortions from the near-surface (UPGS). This paper presents a statistical method to characterize and correct for these distortions by decomposing seismic amplitudes. The accuracy of this technique is validated on a synthetic model, confirming its ability to quantify near-surface properties. We then use a context-aware Long Short-Term Memory (LSTM) network as a case study, demonstrating that such tools can achieve accurate interpretations on clean, geologically complex data. We conclude that a hybrid workflow, where robust, physics data conditioning precedes AI-based analysis, is a prerequisite for enabling the successful application of machine learning in real-world seismic exploration

Keywords: statistical processing, deep learning, geological interpretation of seismic data, neural networks, seismic data processing

Глобальна необхідність у забезпеченні надійними енергетичними ресурсами продовжує бути стимулом для розвідки на все більших глибинах і геологічно складніших умовах. Зі зростанням обсягу та складності сейсмічних даних виникає потреба в передових автоматизованих системах, які можуть допомогти геофізикам в ефективній та точній інтерпретації глибоких горизонтів. Методи штучного інтелекту (ШІ), такі як нейронні мережі з довгою короткочасною пам'яттю (LSTM) та згорткові нейронні мережі (CNN), демонструють потенціал для структурної інтерпретації сейсмічних даних [1], але їх практичне застосування стикається з серйозними перешкодами.

Однією з проблем як для традиційної, так і для машинної інтерпретації є значне спотворення сигналу, викликане приповерхневим шаром – верхньою частиною геологічного розрізу (ВЧР) [2]. Цей шар діє як «лінза», вносячи затримки часу приходу (статика) і викликають значне згасання амплітуд та спотворення форми хвилі, що є критичним для подальшої обробки. Ці ефекти погіршують якість сейсмічного зображення, порушуючи безперервність відбиваючих горизонтів, на яку спираються більшість автоматизованих систем. Представлено дослідження вдосконаленої архітектури LSTM моделі, що враховує просторовий контекст, для інтерпретації горизонтів. Проаналізовано її ефективність на синтетичних даних, та виявлено обмеження та артефакти, що виникають під час застосування. Такі артефакти можуть бути аналогічні ефектам некоригованих спотворень у вхідних даних, які в реальних умовах часто спричинені ВЧР. Для вирішення цих потенційних проблем ми демонструємо спосіб характеристики приповерхневих умов шляхом статистичної декомпозиції сейсмічних амплітуд. Припущено, що інтеграція такого кроку підготовки даних є важливим фактором для підвищення надійності будь-якого робочого процесу на основі ШІ.

Підхід представлений у двох частинах. Спочатку описано статистичний метод для характеристики приповерхневих ефектів, що є основним фокусом. Друга частина – представлення вдосконаленої контекстно-залежної LSTM моделі як приклад для демонстрації критичного впливу якості даних на інтерпретацію за допомогою ШІ.

Запропонований статистичний метод для характеристики приповерхневих ефектів базується на фізичному принципі, згідно з яким амплітуда зареєстрованого сейсмічного запису є добутком кількох незалежних та розділюваних факторів, включаючи енергію джерела, умови прийому та збудження, а також приповерхнєве згасання. Для спрощеної двошарової моделі, що представляє ВЧР (шар 1) та нижчу частину розрізу (шар 2), кінцеву зареєстровану амплітуду $A(t)$ можна виразити як

$$A(t) \approx A_0 * \frac{V_0}{V_1^2 * t_1 + V_2^2 * t_2} * e^{-\alpha_1 * V_1 * t_1 - \alpha_2 * V_2 * t_2} \quad (1)$$

де V_i , t_i та α_i представляють інтервальну швидкість, подвійний час пробігу (TWT) та ефективний коефіцієнт згасання для i -го шару відповідно.

Ця фізична модель показує, що властивості ВЧР (V_1 , t_1 , α_1) мають першочерговий та мінливий вплив на кінцеву амплітуду. Щоб виділити ці ефекти, моделюємо амплітуду $A_{ij}(t)$ для джерела i та приймача j як добуток незалежних компонентів:

$$A_{ij}(t) = A_0 * \frac{B_{i+j}}{2} * C_i(t) * D_j(t) * E_{rand} \quad (2)$$

Де змінні представляють базову енергію (A_0), глибинну геологію (B), ефекти джерела (C), ефекти приймача (D) та випадковий шум (E_{rand}). Шляхом побудови та розв'язання системи лінійних рівнянь, отриманої з цієї моделі, можна виділити компоненту приймача D_j , яка служить надійним індикатором амплітудних спотворень, викликаних ВЧР у кожній точці прийому.

Для демонстрації важливості якості даних, ми використовуємо контекстно-залежну LSTM – архітектуру нейронної мережі, розроблену для послідовних даних, таких як сейсмічні траси. Оновленням цієї моделі є система з двома входами, яка обробляє як поточну сейсмічну трасу, так і результат інтерпретації з попередньої траси. «Контекстний» вхід дозволяє моделі вивчати геологічний принцип горизонтальної безперервності, що робить її інструментом для автоматизованої та допоміжної інтерпретації.

Точність статистичного підходу була перевірена на синтетичній моделі з відомими параметрами ВЧР. Рисунок 1 показує крос-плот істинної товщини ВЧР із синтетичної моделі проти розрахованого нами коефіцієнта d_j . Висока кореляція демонструє, що метод може характеризувати приповерхневі умови, надаючи надійний інструмент для підготовки даних. Цей результат підтверджує, що статистичний коефіцієнт є валідним індикатором фізичних властивостей ВЧР.

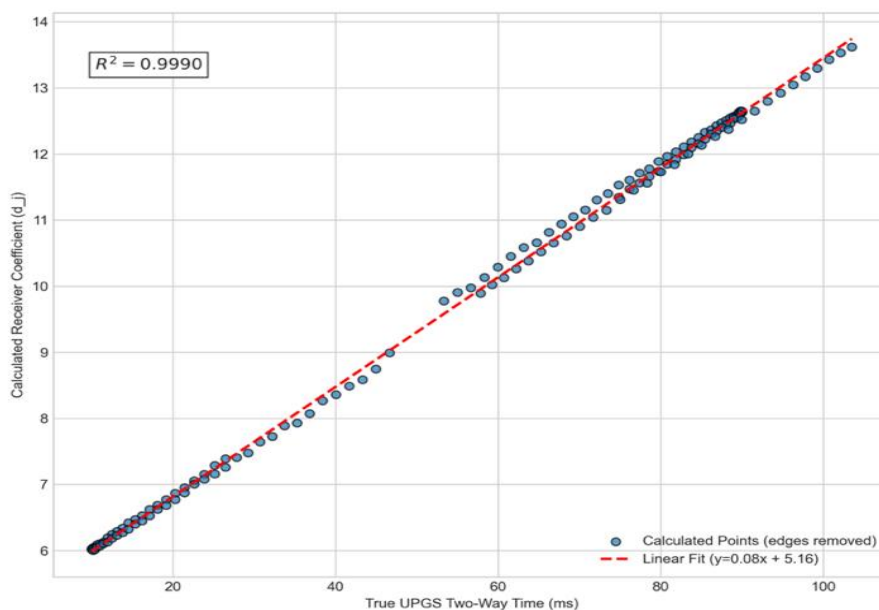


Рис. 1. Крос-плот, що порівнює істинну товщину ВЧР із синтетичної моделі з розрахованим коефіцієнтом приймача d_j

Модель при застосуванні на «чистих» синтетичних сейсмічних даних працює з достатньою точністю, успішно простежуючи горизонти (Рисунок 2). Це представляє сценарій, коли якість вхідних даних є ідеальною. Однак ефективність такої моделі може бути значно знижена за наявності некоригованих приповерхневих спотворень, які вносять шум і порушують безперервність сейсмічних даних.

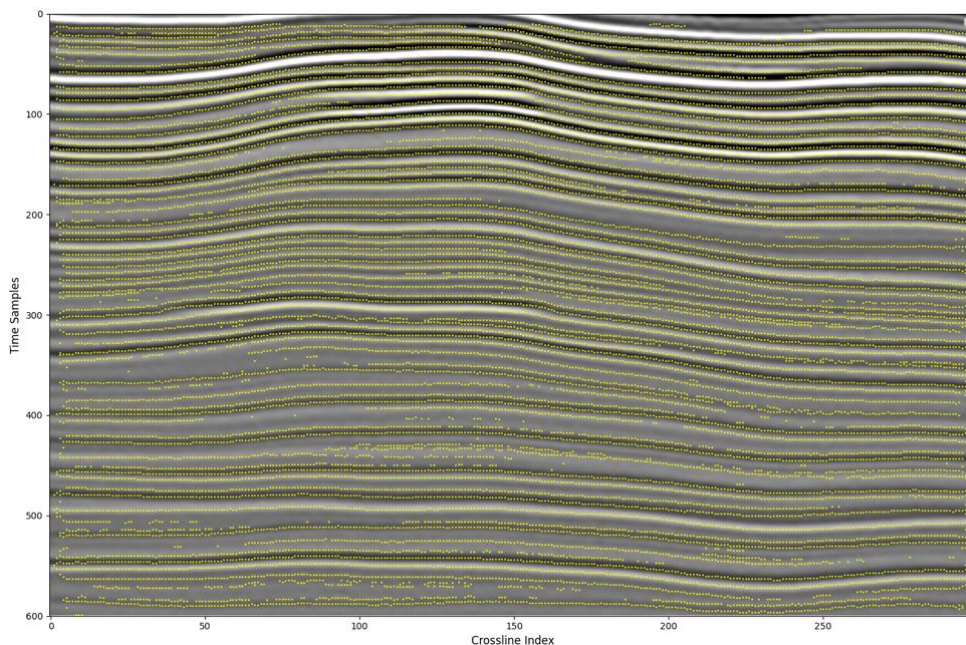


Рис. 2. Результати роботи моделі LSTM на чистих синтетичних сейсмічних даних

Це підкреслює зв'язок між двома частинами досліджуваного підходу. Висока ефективність статистичного методу показує, що ті самі спотворення, які можуть призвести до збою інструменту машинного навчання, можуть бути кількісно оцінені. Застосування такої корекції як кроку попередньої обробки є потенційно критичним фактором для забезпечення надійного застосування інструментів, подібних до цієї LSTM, до реальних сейсмічних даних. Такий гібридний робочий процес є важливим для подолання розриву між продуктивністю моделі на ідеальних синтетичних даних та її застосуванням до реальних завдань розвідки родовищ вуглеводнів.

Отримані нами результати підкреслюють залежність інструментів інтерпретації на основі ШІ від якості вхідних сейсмічних даних. На прикладі контекстно-залежної LSTM було продемонстровано, що хоча такі моделі добре працюють на «чистих» синтетичних даних, їх послідовний, контекстно-залежний характер робить їх вразливими до спотворень, що часто зустрічаються в реальних наборах сейсморозвідувальних даних.

Представлений статистичний метод для характеристики та кількісної оцінки одного зі значних джерел цих спотворень – верхньої частини розрізу (ВЧР). Для успішного та надійного впровадження інструментів на основі ШІ в

сейсмічній розвідці їх не слід розглядати як самостійне рішення. Натомість потрібен гібридний робочий процес, де застосовується попередня обробка, як представлена статистична корекція для забезпечення можливої якості вхідних даних.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Піскунов О.В., Тяпкін О.К. Можливість використання методів глибинного навчання у інтерпретації елементів геологічного розрізу на сейсмічних даних. *Сучасні проблеми гірничої геології та геоекології: Матеріали IV Міжнародної наукової конференції*. Київ. ДУ НЦ ГГГІ НАН України, 2024. С.90-94.
2. Tiapkin O.K., Onyshchenko S.A., Piskunov O.V. Engineering-ecological research of the upper part of geological section using seismic data. *Geofizychnyi Zhurnal*. 2025. 47(2). P. 150-155.