

УДК 550.34.06

ПРИМЕНЕНИЕ СИНТЕТИЧЕСКИХ ВОЛНОВЫХ ФОРМ ДЛЯ СОЗДАНИЯ МОДЕЛИ ГЛУБОКОГО МАШИННОГО ОБУЧЕНИЯ С ЦЕЛЬЮ ВЫДЕЛЕНИЯ ПЕРВЫХ ВСТУПЛЕНИЙ НА СЕЙСМИЧЕСКИХ ЗАПИСЯХ

© 2023 г. Н. А. Барышников^{1, *}, И. А. Абзалилов^{1, 2}, С. Б. Турунтаев^{1, 2}¹Институт динамики геосфер имени академика М. А. Садовского РАН, Москва, Россия²Московский физико-технический институт (национальный исследовательский университет), Москва, Россия

*E-mail: nabarysh@gmail.com

Задача данной работы – построение метода анализа записей сигналов группы станций, полученных в ходе микросейсмического мониторинга с использованием нейронной сети с целью выделения моментов первых вступлений. Новизна предлагаемого метода заключается в использовании *Wavelet Scattering* для выделения признаков, что существенно упрощает процесс обучения модели по сравнению с традиционно применяемыми свёрточными нейронными сетями. Используемая в модели архитектура *Transformer* позволила реализовать обмен информацией между произвольным числом принимающих станций и, тем самым, определять времена вступлений на отдельных станциях на основе информации всей группы. Обучение модели проводилось на полностью синтетических данных. Такой подход по сравнению с обучением на реальных данных может позволить получить модель с лучшей обобщающей способностью, так как синтетические данные отражают механику процесса возникновения и распространения упругих колебаний и при этом не искажены локальными условиями и особенностями регистрации сигналов. Исследован ряд стратегий аугментации (модификации данных обучающей выборки) путём наложения шума с целью увеличения устойчивости модели к шуму. Наилучшее соотношение устойчивости к шуму и точности модели было получено при наложении Гауссовского шума со случайным стандартным отклонением. Результаты работы могут быть использованы при построении моделей глубокого обучения, обучаемых как на синтетических, так и на реальных сейсмических записях.

Ключевые слова: микросейсмический мониторинг, машинное обучение, синтетические волновые формы, выделение фаз.

Для цитирования: Барышников Н.А., Абзалилов И.А., Турунтаев С.Б. Применение синтетических волновых форм для создания модели глубокого машинного обучения с целью выделения первых вступлений на сейсмических записях // Динамические процессы в геосферах. 2023. Т. 15. № 3. С. 38–53. http://doi.org/10.26006/29490995_2023_15_3_38

Введение

Микросейсмический мониторинг является одним из методов анализа процессов в геофизических объектах (месторождениях полезных ископаемых, крупномасштабных подземных объектах и т. д.). Он осуществляется с помощью сети станций, фиксирующих слабые сейсмические сигналы. В результате анализа характеристик облака микросейсмических событий могут быть получены важные сведения о процессах, протекающих внутри исследуемого объекта, определены механические и фильтрационные свойства среды, получены данные о её структуре и наличии неоднородностей. Качество результатов, получаемых при микросейсмическом мониторинге, определяется возможностью регистрировать события с как можно меньшей магнитудой. Это может быть достигнуто путём увеличения количества регистрирующих станций: современные сети станций могут состоять из сотен и даже тысячи приборов. Метод анализа данных такой сети должен строиться на основе параллельного анализа всего массива записей, что делает актуальной разработку методов, основанных на использовании нейронных сетей.

Нейронные сети для анализа сейсмических сигналов делятся на основе подхода к выделению признаков. Как правило, используются различные характеристические функции временного представления сигнала [Guo et al., 2021], либо его частотно-временные представления [Zhang et al., 2020]. Для классификации частотно-временных представлений сигналов наиболее успешно используются сверточные нейронные сети [Zhu et al., 2019], а также применяется *transfer learning*, при котором переобучаются существующие модели для анализа изображений [Zhu, Beroza, 2019]. По мере развития архитектур рекуррентных нейронных сетей, они также стали активно применяться для обработки сейсмических сигналов [Zheng et al., 2018]. В настоящее время наиболее эффективно используются сети на основе трансформеров [Stepnov et al., 2021]. Также встречаются и различные сочетания перечисленных методов [Mousavi et al., 2020]. В целом, использование нейронных сетей для задач обработки сейсмических сигналов развивается в общем русле адаптации к данной задаче новых алгоритмов нейронных сетей по мере их появления в виде общедоступных программных продуктов [Anikiev et al., 2023]. Среди основных вызовов на данный момент следует выделить низкую обобщающую способность получаемых моделей: модель, обученная на специфичном наборе данных (данные определённой сети станций, региона наблюдения и т. д.), показывает хорошие результаты при валидации, но оказывается не применима для работы с данными из другого набора. При этом задача получения достаточно представительной обучающей выборки данных для конкретных условий мониторинга может оказаться невыполнимой.

Эта проблема может быть решена при помощи использования синтетических данных при обучении модели. Преимущество такого подхода состоит в том, что для синтетических данных заранее известны все необходимые параметры каждого сейсмического события (время вступления, координаты источника, магнитуда, параметры механизма источника, скоростная модель среды). Это дает возможность создавать и обучать модели на данных, сформированных с учётом особенностей, характерных для заданного района проведения мониторинга, получаемые модели при этом могут обладать большей обобщающей способностью, чем обученные на реальных волновых формах. Кроме того, такой подход, в отличие от использования банков реальных волновых форм, исключает возможность искажения результатов, связанных с применением ручной разметки данных. Главный недостаток использования синтетических данных – сложность адаптации получаемых моделей под реальные данные. Моделируемые сигналы должны быть достаточно реалистичны, для этого необходимо предельно точно моделировать параметры механизма источника и характеристики среды. В данной работе была предпринята попытка построить модель для выделения времён вступления сейсмических фаз на основе нейронной сети, обученной на полностью синтетических волновых формах.

Данные для обучения модели

Синтетические волновые формы, использованные в данной работе, были созданы при помощи Pyrocko – набора библиотек для сейсмологического моделирования с открытым исходным кодом [Heimann et al., 2018]. Распространение сейсмических волн моделировалось для упруго-вязкой слоистой среды. Скоростная модель среды была взята из работы [Málek et al., 2005]. Выбор определялся наличием в свободном доступе предварительно рассчитанного банка функций Грина [https://Greens-Mill.Pyrocko.Org/Vogtland_malek2004_surface-3dc953, n.d.], необходимого для получения волновых форм. Источники сейсмических сигналов моделировались двойной парой сил со случайным равномерным распределением направлений смещения (*strike, deep, rake*) и моментными магнитудами, равномерно распределёнными в заданных границах (0–2). Эпицентры и глубины источников были случайным образом распределены внутри области с максимальным эпицентральной расстоянием в 1.5 км от начала координат и глубиной от 1 до 2 км. Волновые формы (смещение, рис. 1) были смоделированы для пяти станций (четыре симметрично расположены на расстоянии 500 метров от начала координат и одна в центре (рис. 2)) для трёх каналов (два горизонтальных *N*, *E* и вертикальный *Z*) с частотой дискретизации 100 Гц (частота определялась параметрами использованного банка функций Грина),

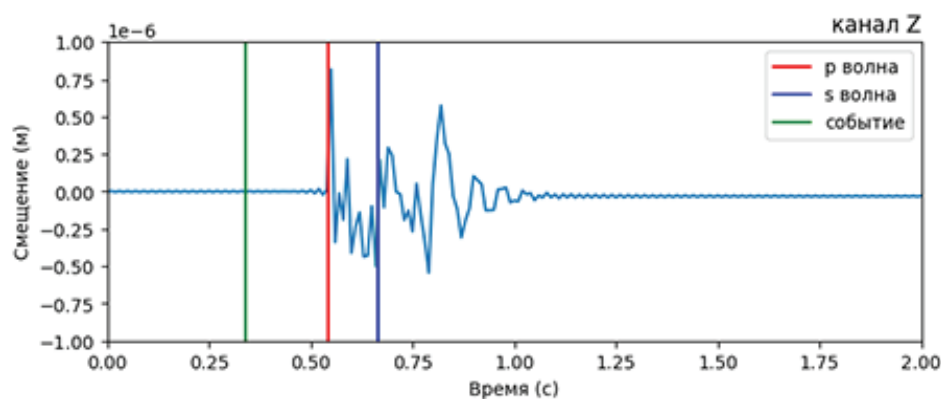


Рис. 1. Пример модельной волновой формы. Вертикальными линиями отмечены момент события и моменты прихода p - и s -волн

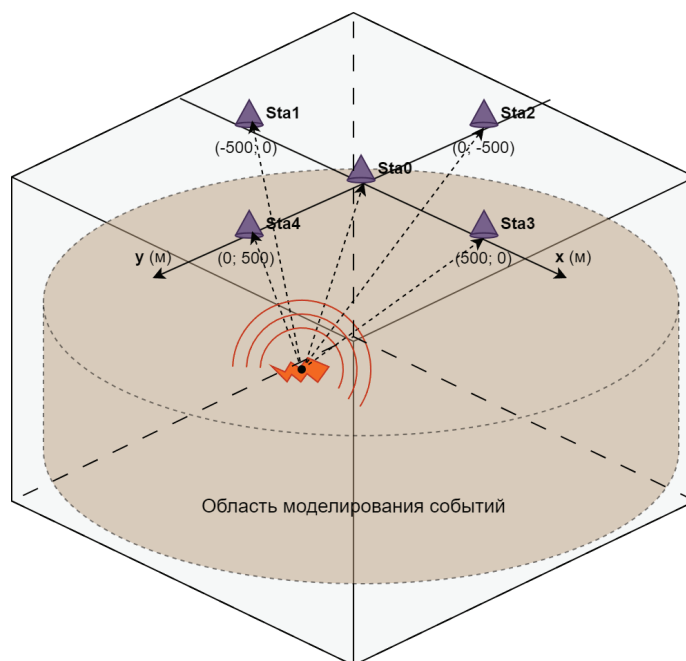


Рис. 2. Схема моделирования волновых форм

длительность каждой записи – 4 секунды. Для каждой волновой формы были получены точные моменты прихода p - и s -волн. В результате моделирования были сформированы обучающая и тестовая выборки из 106 и 103 событий соответственно (15 волновых форм для каждого события).

Полученные в результате моделирования данные проходили несколько стадий первичной обработки (в порядке применения):

- **Нормировка сигнала.** При моделировании волновых форм задавались магнитуды от 0 до 2. Это привело к тому, что амплитуды разных записей различаются на несколько порядков. Нормировка производилась для каждого события, так что максимальная амплитуда сигнала в каждом событии была всегда меньше либо равна 1. Параметр нормировки сохранялся для возможности дальнейшего использования в модели.

- **Преобразование времён моментов прихода фаз (p -, s -волн) в «one-hot» представлении:** вектора, совпадающие по длине с векторами волновых форм, в которых значение 1 соответствует моменту прихода соответствующей фазы, все остальные значения – 0. В случае, когда момент прихода фазы «попадает» между двумя точками волновой формы (что бывает почти всегда), метка 1 присваивалась обоим соседним точкам. По сути такое представление можно рассматривать как классификацию каждого интервала между соседними точками записи по отношению к соответствующему вступлению. Такое представление необходимо для адаптации модели для работы с записями, содержащими несколько

сигналов, наложенных друг на друга. Оно позволяет получать, на выходе из модели, временной ряд с условной вероятностью того, что в данной точке существует приход соответствующей фазы.

- Аугментация сигнала: наложение гауссовского шума малой амплитуды и «спайков» (будет обсуждаться в разделе, посвященном устойчивости модели к шуму).
- Вейвлет фильтрация высокочастотной осцилляции, присутствующей в сигнале и, по-видимому, являющейся артефактом моделирования. Для этого использовалось кратно-масштабное разложение – реконструкция сигнала с применением вейвлета Добеши 4-го порядка [Daubechies, 1992] и занулением детализирующих коэффициентов первого уровня декомпозиции.

Модель глубокого обучения на базе архитектур *deep scattering network* и *transformer encoder*

Схема предложенной модели показана на рисунке 3. Модель была реализована с использованием библиотеки Pytorch. Для формирования пространства признаков используется *Log-wavelet Scattering*

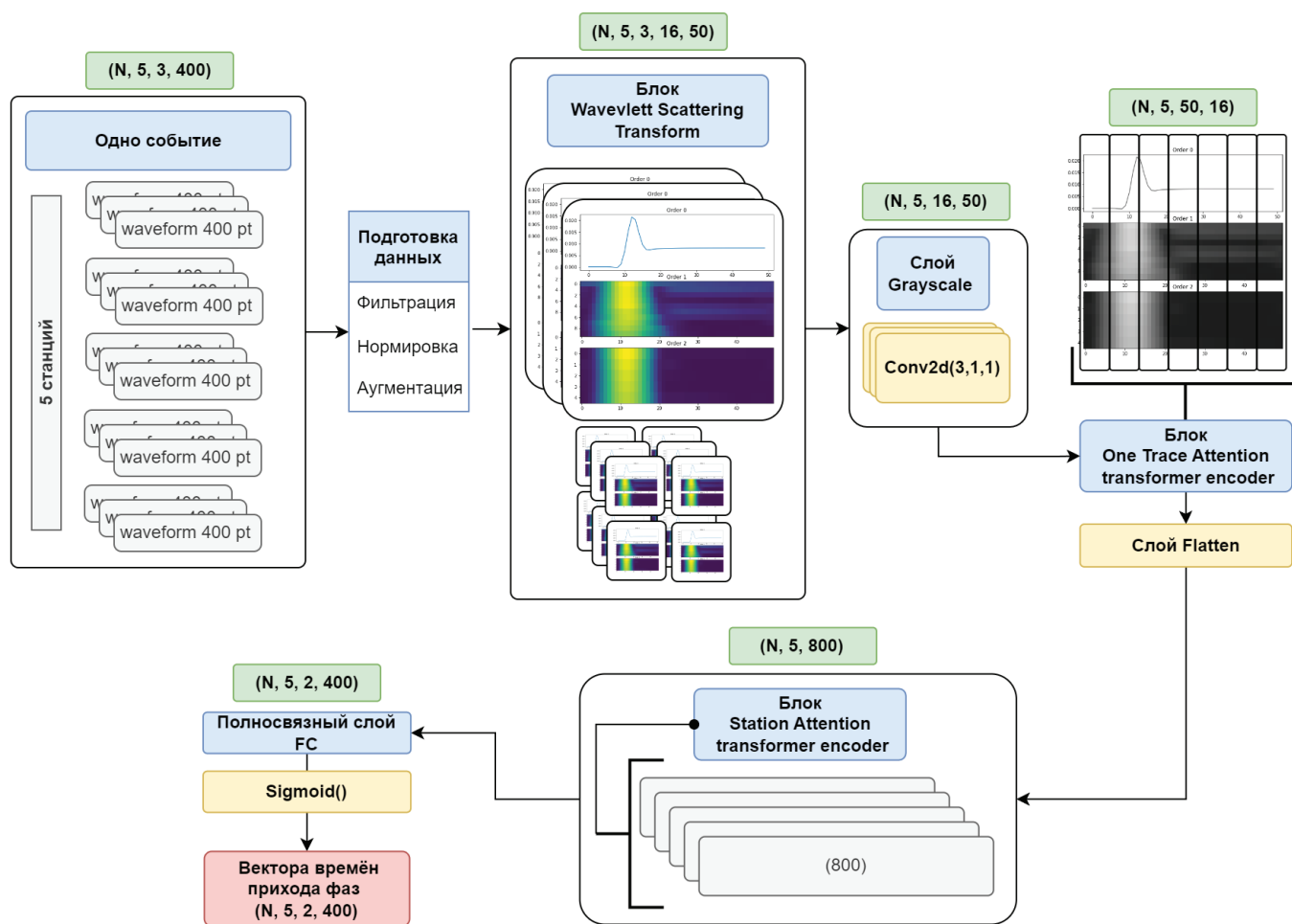


Рис. 3. Схема реализованной модели. В скобках указаны размерности тензоров данных, получаемых на каждом этапе работы модели; N – количество событий в одном пакете

transform исходного сигнала [Andén, Mallat, 2014]. Блок *Wavelet Scattering* (*WS*) формирует частотно-временное представление сигнала на основе вейвлет разложения, отчасти похожее на результат оконного преобразования Фурье. При этом *WS* обладает необходимыми свойствами устойчивости к искажениям исходного сигнала, сдвигам по времени и лучшему сохранению информации при

разложении. Сам алгоритм представляет собой рекуррентную процедуру, в которой к исходному сигналу последовательно применяются операции вейвлет-пакетного разложения, взятия комплексного модуля и усреднения. С точки зрения нейронных сетей данный слой является аналогом сверточных слоев и операций пуллинга. Но, если в сверточном слое фильтры проходят процедуру обучения, то в случае использования *WS* они представлены набором вейвлетов, что снимает неопределённость, связанную с необходимостью обучения фильтров. В качестве реализации *WS* использовалась библиотека Kymatio [Andreux et al., 2020].

В результате применения *WS* каждый вектор волновой формы из пачки записей сигналов преобразуется в двухмерную матрицу признаков. После этого применяется двумерный сверточный слой с единичным размером ядра, линейно преобразующий 3-х канальное изображение каждой записи станции в 1-но канальное (блок «*Grayscale*»). На следующем этапе 2D представление записи каждой станции делится на временные фреймы и через слой кодирования позиции подаётся в первый модуль трансформер-энкодера [Vaswani et al., 2017] – модели, реализующей механизм самовнимания, и обеспечивающей лучшее выделение области, в которой находится сигнал. Далее записи поступают во второй модуль трансформер-энкодера, реализующий механизм взаимного внимания на записях группы станций. Применяемая в данной работе конкретная реализация механизма самовнимания основана на скалярном произведении векторов записей, поэтому данный модуль, по сути, накладывает на вектора данных маску их кросс-корреляций. Это позволяет осуществлять обмен информации между всеми записями всех станций. Данные, полученные на выходе второго модуля трансформер-энкодера, разворачиваются в линейные записи (слой «*Flatten*»), поступают в полносвязный слой и далее в логистическую функцию (сигмоиду), приводящую данные на выходе полносвязного слоя к диапазону (0, 1). Таким образом, на выходе из модели для каждого события получается матрица из пар векторов, отражающих условную вероятность принадлежности каждой точки записи сигнала к моменту первого вступления каждой фазы (рис. 4).

Код модели доступен по адресу: https://github.com/idglgfd/synth_wf_phase_picking

Обучение модели

При обучении модели в качестве функции потерь была использована бинарная кросс-энтропия с увеличенным штрафом за несовпадение с положительными метками:

$$l = -\frac{1}{N} \sum_{i=1}^N [\alpha^{-1} \cdot y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)],$$

где $\hat{y}$ – вектор предсказаний модели, y – целевой вектор, α – примерная доля положительных меток в целевом векторе. Это позволило избежать ошибок обучения, связанных с несбалансированностью целевых векторов (два значения с меткой «1» на 398 нулевых значениях). Для обучения модели был использован оптимизатор Adam. Обучение проводилось с использованием технологии CUDA на видеокарте пользовательского уровня (NVIDIA GeForce 1660 GTX Super). Время обучения 1 эпохи составляло порядка 20 минут. Для удовлетворительного обучения модели на использованной обучающей выборке было достаточно одной-двух эпох, приращение точности на последующих эпохах было незначительным. В процессе обучения был проведён подбор параметров *Wavelet Scattering*, обеспечивающих минимальное значение функции потерь при обучении в течение одной эпохи.

В дальнейшем в модели использовались значения показателя масштаба вейвлета $J = 3$ (максимальный масштаб $2^J = 8$ точек) и количество вейвлетов на октаву в первом слое $Q = 6$. Также было подобрано оптимальное количество слоёв в модулях трансформер-энкодера (2 и 3 соответственно). Для проведения испытаний была использована библиотека для распределённых вычислений Ray. Результаты испытаний представлялись при помощи пакета Wandb.

Результаты обучения

На рисунке 4 показан пример выделения моментов прихода фаз на отдельном канале записи трёх станций. Сплошными линиями отмечены фактические моменты прихода фаз. Сплошным тоном

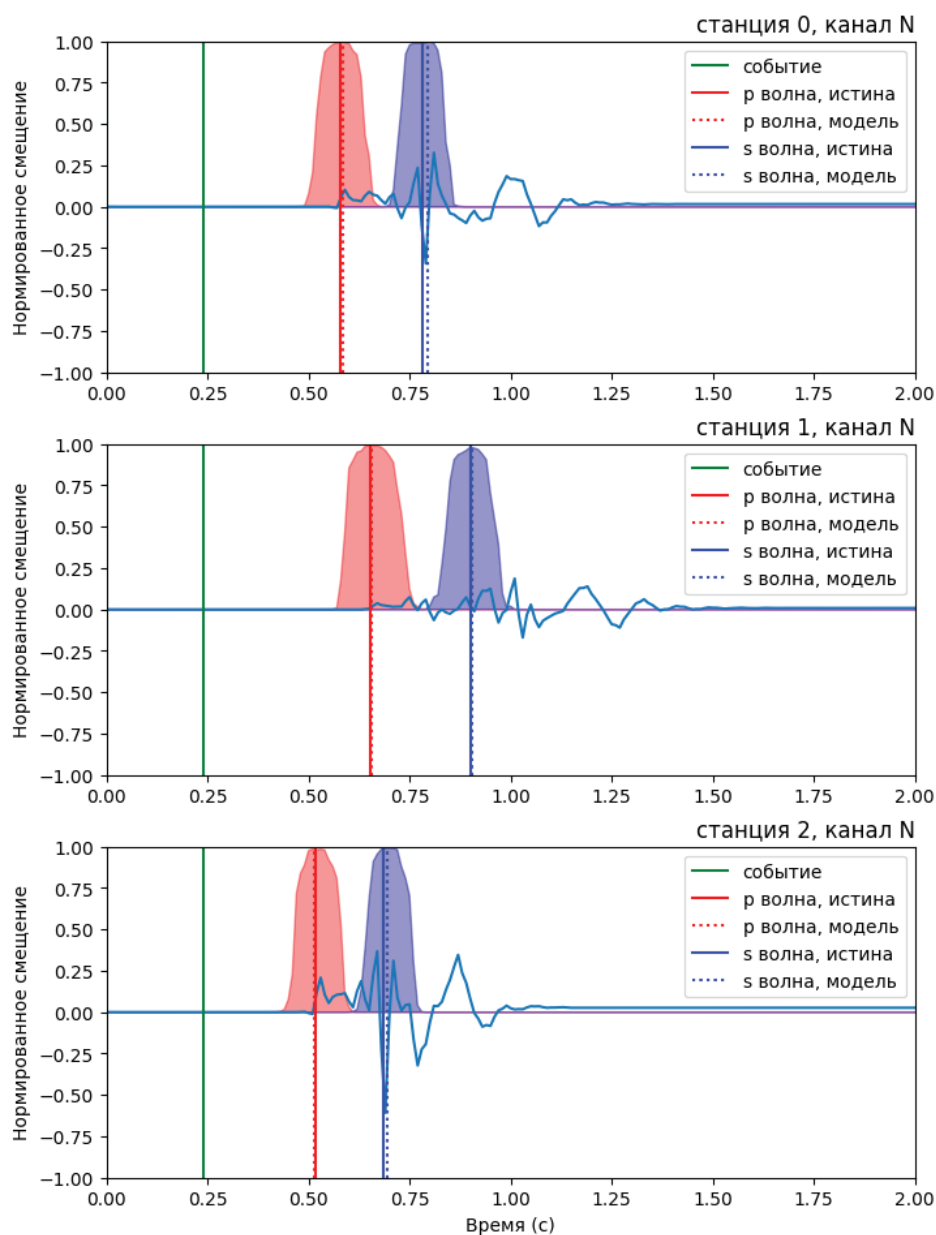


Рис. 4. Результат определения моментов прихода фаз для трёх станций. Сплошными линиями отмечены фактические моменты прихода фаз; сплошным тоном – полученные моделью вектора условных вероятностей для моментов прихода; пунктирными линиями – моменты времени, полученные из вероятностного представления

показаны полученные моделью вектора условной вероятности (степень уверенности модели в том, что момент прихода соответствующей фазы находится в данной точке). Пунктирными линиями показаны моменты времени, полученные из условных вероятностей. Каждая линия соответствует интервалу с максимальной суммой значений вектора вероятностей на его краях:

$$t_{p,s} = \arg \max(\hat{y}_i + \hat{y}_{i+1}).$$

Стандартное отклонение определения моментов прихода на тестовой не зашумлённой выборке составило 0.014 и 0.012 секунды для p - и s -волн соответственно (чуть больше одного отсчёта при частоте дискретизации записи 100 Гц).

Устойчивость модели к шуму в данных

Обобщающая способность модели в значительной степени определяется её устойчивостью по отношению к шуму в обрабатываемых данных. Анализ работы представленной модели показал, что даже небольшой шум (амплитуда порядка 0.1% от максимальной амплитуды сигнала), наложенный на подаваемые в модель записи сигналов, приводит к критическому ухудшению качества определения момента прихода волн (рис. 5). Для увеличения толерантности модели к шуму применяют различные стратегии аугментации данных – внесение дополнительных шумов или искажений в обучающую выборку [He et al., 2021]. Само по себе использование синтетических данных для обучения модели можно рассматривать как способ аугментации данных реальных наблюдений.

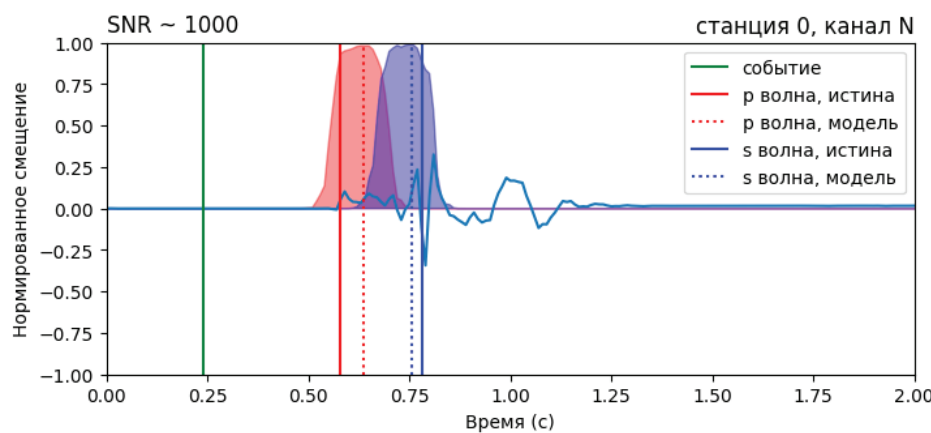


Рис. 5. Результат работы модели, обученной на неаугментированной обучающей выборке при анализе данных с незначительным ($SNR \sim 1000$) шумом

В ходе данной работы было рассмотрено несколько способов внесения шума в обучающие данные с целью повышения устойчивости модели к шуму в тестовых данных:

- гауссовский шум со стандартным отклонением σ_n (имитация «шума прибора»);
- гауссовский коррелированный шум со стандартным отклонением σ_{n_corr} (одинаковый шум на всех каналах одной станции, но разный на разных станциях, имитация «проехавшей рядом со станцией машины»);
- «спайки» – определённая доля точек вектора исходных данных заменяется на случайную величину со стандартным отклонением σ_{spike} .

Кроме того, были рассмотрены два способа задания стандартного отклонения шума: постоянное значение для всей выборки либо случайное значение для каждого события (величины стандартного отклонения обоих видов шумов для каждого события выбирались случайным образом в границах от 0 до заданного максимального значения).

Для выбора оптимального способа внесения шума в данные обучающей выборки была проведена серия испытаний. В каждом испытании модель в течение 2-х эпох обучалась на данных, зашумлённых определёнными сочетаниями перечисленных способов – всего порядка 50 испытаний (сводная таблица параметров представлена в Приложении 1). Затем полученная модель испытывалась на данных тестовой выборки, на которую также был наложен шум (случайный и коррелированный) с различной амплитудой. Для каждого испытания были построены зависимости величины стандартных отклонений предсказанных времен прихода фаз от стандартного отклонения σ_{test} шума, наложенного на данные тестовой выборки (рис. 6).

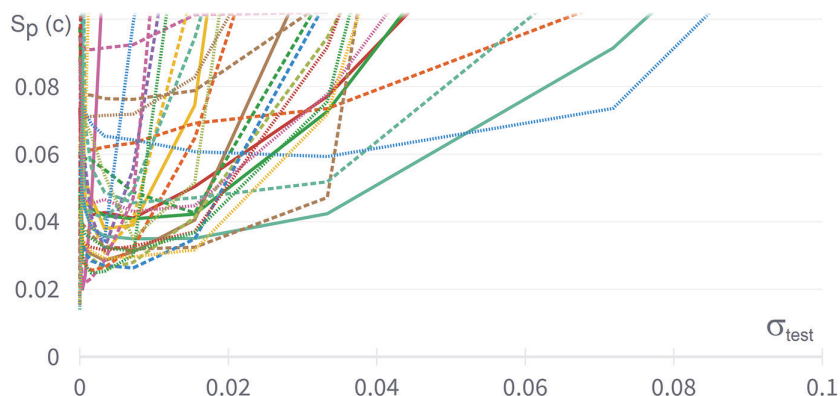


Рис. 6. Стандартная ошибка s_p выделения времени прихода p -волны для различных стратегий аугментации. Каждая линия на графике соответствует модели, обученной на данных с определённым сочетанием шумов

Влияние различных видов шума в данных обучающей выборки на устойчивость модели

Наложение всех видов шума на данные обучающей выборки в той или иной мере расширяет область устойчивости модели к шуму в данных тестовой выборки. В таблице 1 сведены качественные результаты испытаний модели, обученных на данных с различным видом шума, далее каждый пункт рассмотрен подробнее.

Таблица 1.

Качественный эффект от различных способов наложения шума на устойчивость модели

	В отсутствие обычного Гауссовского шума:	В дополнение к обычному Гауссовскому шуму:
Гауссовский шум	Существенно увеличивает точность модели	—
Гауссовский коррелированный шум	Увеличивает точность модели	Незначительно ухудшает точность модели
«Спайки»	Увеличивает точность модели	Сказывается незначительно

На рисунке 7 показаны усреднённые для всех испытаний кривые зависимости стандартной ошибки s_p выделения времен прихода p -волны от величины стандартного отклонения шума σ_{test} , наложенного на данные тестовой выборки, для трёх величин стандартного отклонения σ_n гауссовского некоррелированного шума, наложенного на данные при обучении модели (для s -волны здесь и далее картина схожая). С увеличением амплитуды шума обучающей выборки кривые в среднем становятся ниже, то есть увеличивается точность модели при работе с зашумлёнными тестовыми данными. Однако максимум точности (минимум кривой) при этом растёт, что хорошо заметно при сравнении кривых с σ_n , равными 0.1 и 0.01. На графике (рис. 7) этому значению соответствуют, в основном, испытания с переменной $0 < \sigma_n \leq 0.1$.

Существенного положительного эффекта от присутствия коррелированного шума в дополнение к обычному в данных обучающей выборки обнаружено не было. Общая тенденция от присутствия коррелированного шума такая же, как и в случае простого шума: сам по себе он существенно увеличивает область устойчивости модели (рис. 8), однако в сочетании с обычным шумом даёт эффект скорее обратный от желаемого (рис. 9).

Присутствие «спайков» в обучающей выборке в дополнение к обычному шуму (рис. 10) существенно не сказывается на устойчивости модели (как и в случае коррелированного шума). В условиях

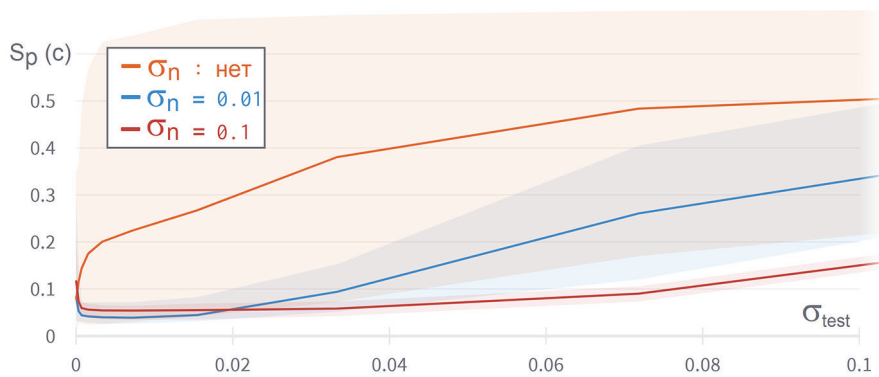


Рис. 8. Ошибка выделения времени прихода p -волны в зависимости от величины стандартного отклонения шума σ_{test} , наложенного на данные тестовой выборки для двух величин стандартного отклонения σ_{corr} гауссовского коррелированного шума, наложенного на данные при обучении модели. Кривые усреднены по всем испытаниям с данным значением σ_{corr} , в которых отсутствовал некоррелированный шум ($\sigma_n = 0$)

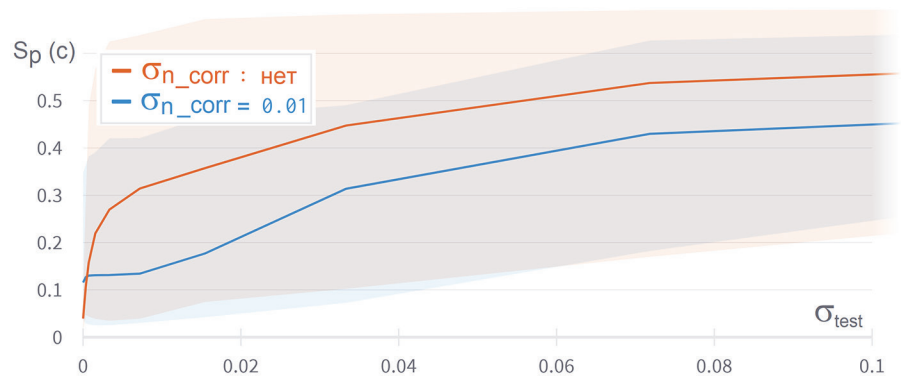


Рис. 9. Ошибка выделения времени прихода p -волны в зависимости от величины стандартного отклонения шума σ_{test} , наложенного на данные тестовой выборки для двух величин стандартного отклонения σ_{corr} гауссовского коррелированного шума, наложенного на данные при обучении модели. Кривые усреднены по всем испытаниям с данным значением σ_{corr}

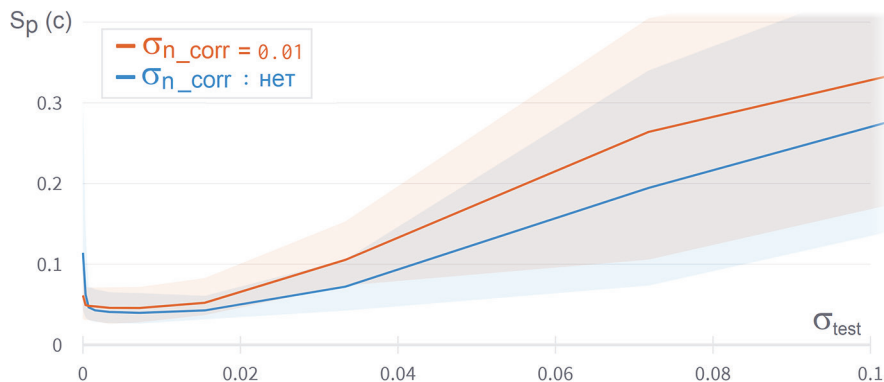
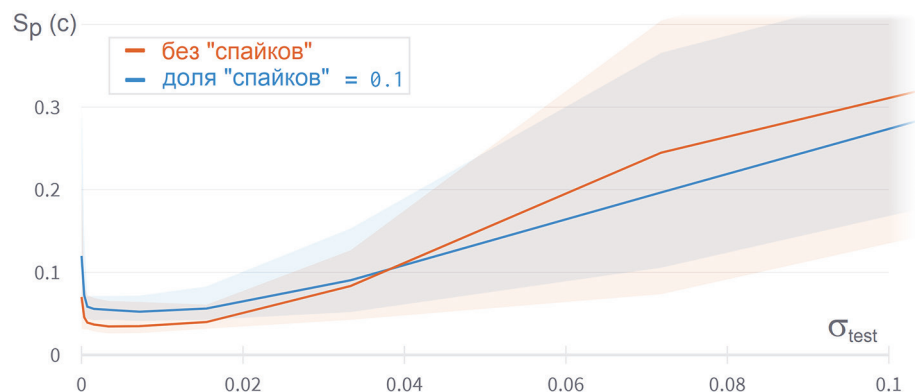


Рис. 10. Ошибка выделения времени прихода p -волны в зависимости от величины стандартного отклонения шума σ_{test} , наложенного на данные тестовой выборки для моделей, обученных на данных со «спайками» и без них в дополнение к обычному шуму. Кривые усреднены по испытаниям с присутствием либо отсутствием «спайков» в обучающей выборке



отсутствия обычного шума присутствие «спайков» увеличивает устойчивость модели (рис. 11). Сами по себе «спайки» при отсутствии прочих шумов, при одинаковой амплитуде, работают существенно хуже обычного шума.

Варьирование амплитуды накладываемых шумов практически никак не сказывалось на устойчивости

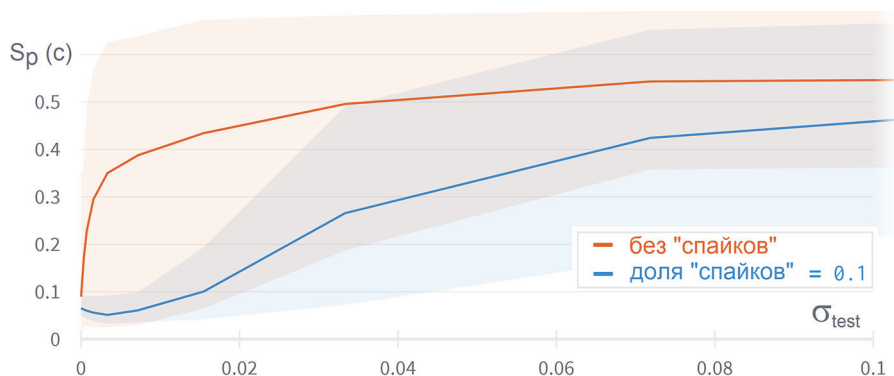
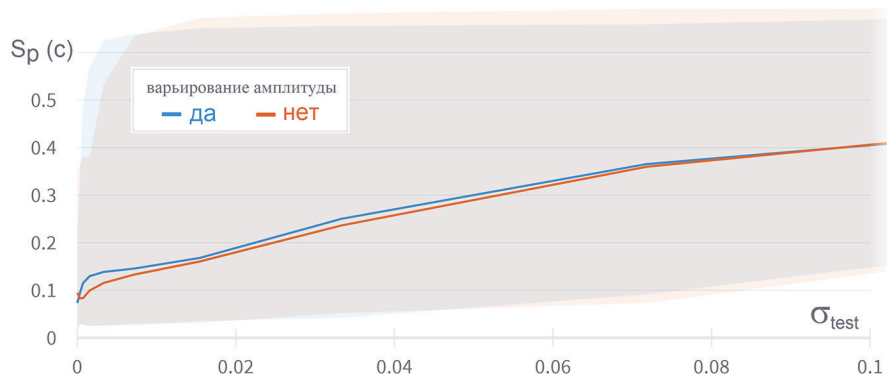


Рис. 11. Ошибка выделения времени прихода p -волны в зависимости от величины стандартного отклонения шума σ_{test} , наложенного на данные тестовой выборки для моделей, обученных на данных со «спайками» и без них в отсутствие обычного шума. Кривые усреднены по испытаниям с присутствием либо отсутствием «спайков» в обучающей выборке.

Рис. 12. Влияние варьирования амплитуды шумов на стандартную ошибку выделения времени прихода p -волны. Кривые усреднены по испытаниям, в которых варьировалась либо не варьировалась амплитуда накладываемых шумов



модели (рис. 12). Предполагалось, что варьирование поможет уменьшить эффект снижения точности модели в области шумов тестовых данных с низкой амплитудой, однако при обучении модели в течение 2-х эпох этого добиться не удалось.

Результаты испытаний параметров аугментации

В результате анализа проведённых испытаний в качестве наиболее перспективного способа аугментации для дальнейшей работы выбран Гауссовский некоррелированный шум со случайным стандартным отклонением, варьируемым в пределах $0 < \sigma_n \leq 0.1$. Случайная величина дисперсии σ_n в среднем позволяла получить приемлемую погрешность на широком интервале амплитуд шумов тестовой выборки. Меньшее отклонение σ_n не позволило получить приемлемую погрешность модели в области сильных шумов ($\sigma_{test} > 0.01$). При фиксированном отклонении шума $\sigma_n = 0.1$ большая часть попыток обучения модели заканчивалась неудачно (модель не обучалась).

На рисунке 13 показаны результаты работы модели, обученной с указанными параметрами аугментации при различных соотношениях сигнал/шум анализируемого сигнала. Следует отметить некоторое увеличение ошибки выделения при нулевом уровне шума, что было характерно для большинства проведённых испытаний. Это может быть связано с переобучением модели на используемый вид шума. В диапазоне соотношений сигнал/шум порядка 10 стандартная ошибка выделения времён вступлений p - и s -волн составила порядка 0.03 секунды, что может считаться хорошим результатом при учёте того, что в работе использовались записи сигналов с частотой дискретизации 100 Гц.

Сопоставление модели с методом выделения первого вступления на основе критерия Акаике

Для оценки качества полученной модели было проведено сравнение ошибок выделения времени прихода p -волны с широко применяемым методом на основе информационного критерия Акаике [Maeda, 1985]. В качестве его реализации была использована соответствующая функция пакета для

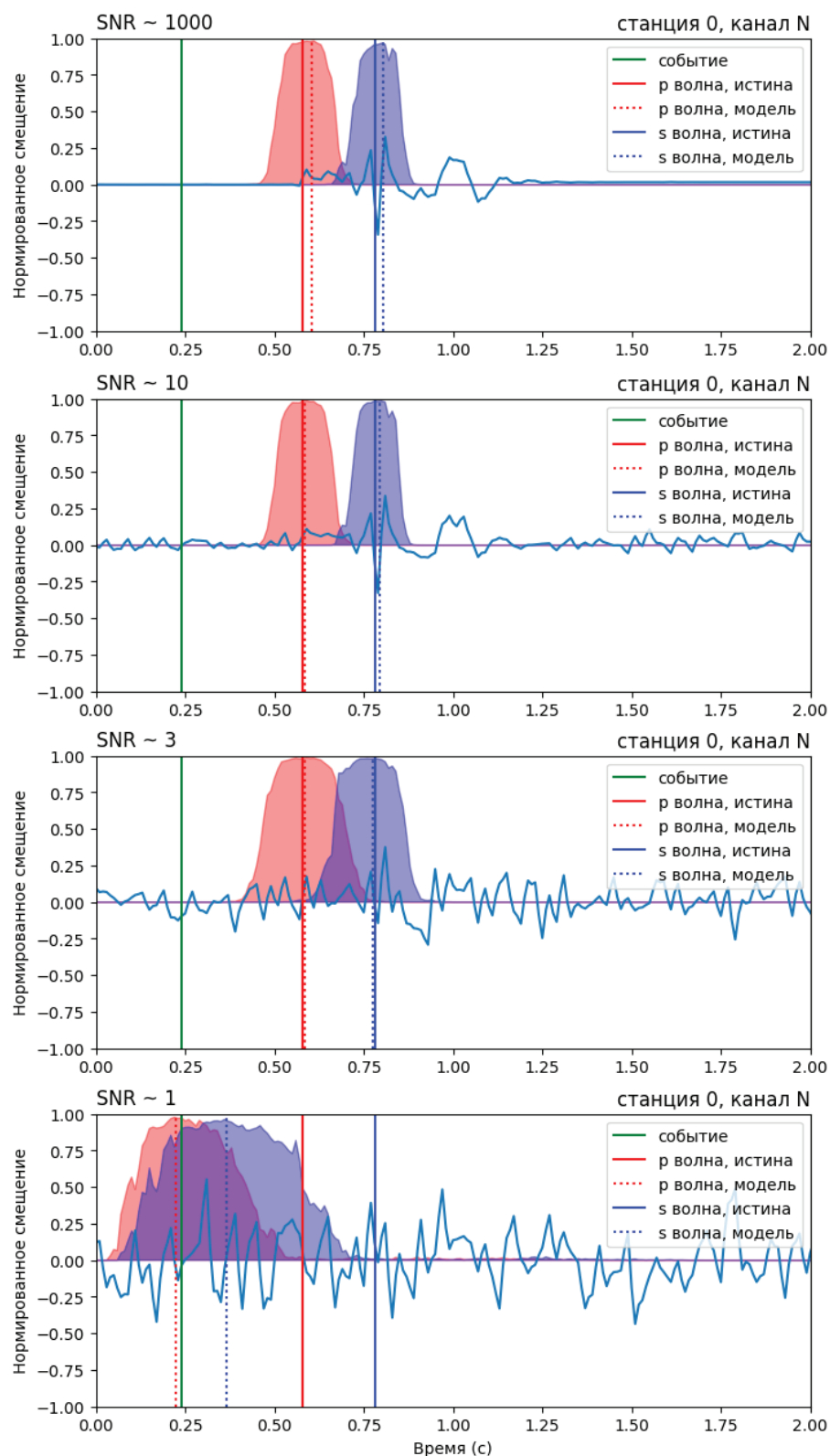


Рис. 13. Результат работы модели, обученной на данных, аугментированных Гауссовским шумом со случайным стандартным отклонением $0 < \sigma_n \leq 0.1$, при различных соотношениях сигнал/шум анализируемого сигнала

обработки сейсмических данных ObsPy [Beyreuther et al., 2010]. В качестве первого вступления в каждой записи сигнала бралась точка с минимальным значением критерия AIC в диапазоне от начала записи до точки с максимальной амплитудой сигнала.

Для исследования ошибок выделения первых вступлений была реализована следующая процедура. На волновые формы тестовой выборки накладывался шум (Гауссовский обычный и коррелированный) со случайной амплитудой в диапазоне от 0 до некоторой величины. Для полученных данных вычислялась величина фактического соотношения сигнал/шум для записи каждой станции $SNR = \max(|x|)/3\sigma_{test}$ (из трёх каналов выбирался канал с максимальной амплитудой сигнала). Полученные зашумлённые данные были использованы для выделения времён прихода p -волны.

На рисунке 14 представлены графики зависимости ошибок выделения времён $\Delta t = t_p - \hat{t}_p$ от фактического соотношения сигнал/шум данных тестовой выборки для обоих методов (модель и AIC). Пунктирными линиями на графиках обозначены медианные значения ошибок $M_e = \text{median}(\Delta t_i)$,

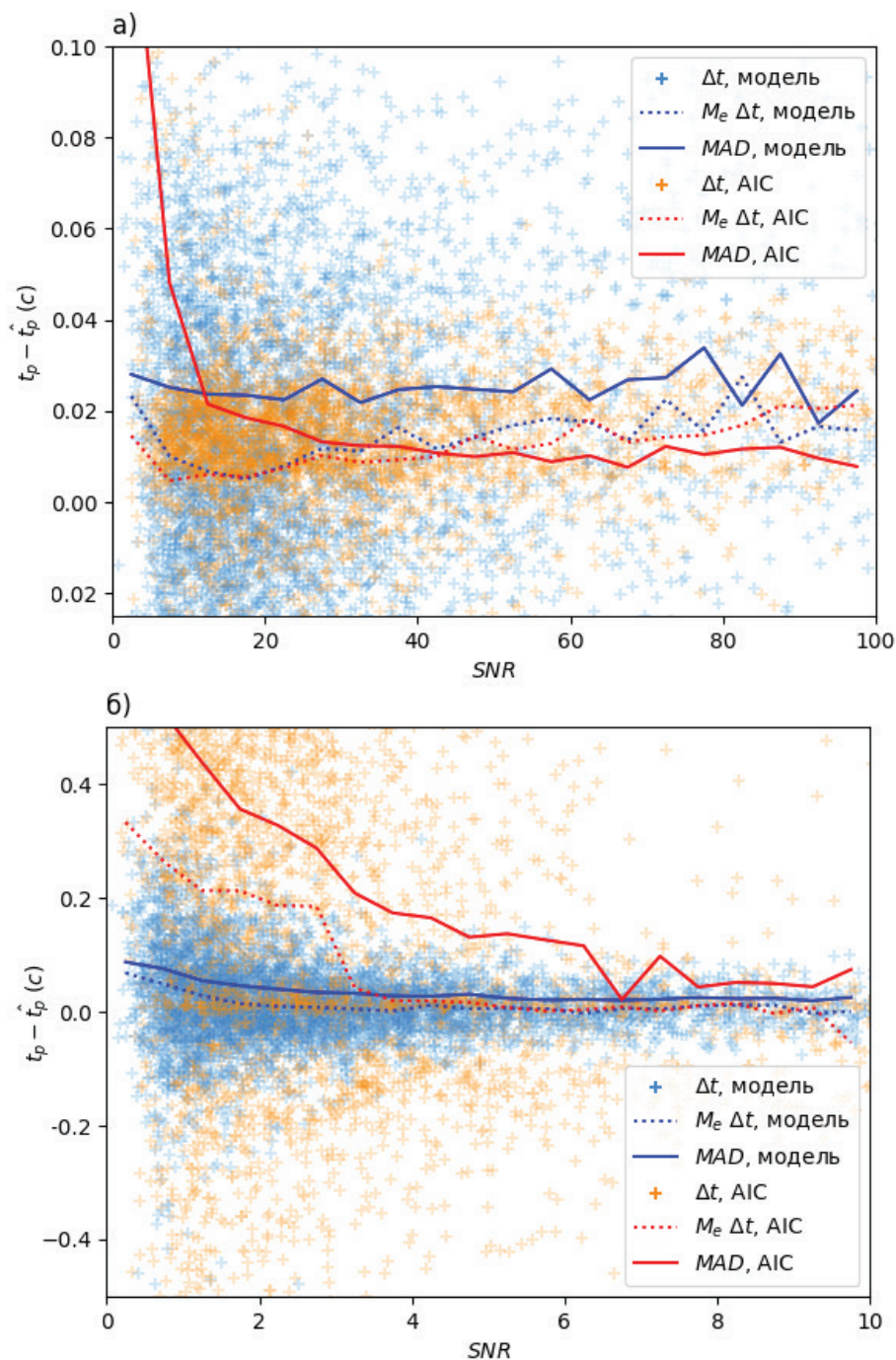


Рис. 14. Зависимости невязки выделения времён Δt от фактического соотношения сигнал/шум для рассматриваемых методов (представленная модель и AIC): а) для высоких соотношений сигнал/шум ($SNR < 100$); б) для низких соотношений сигнал/шум ($SNR < 10$). M_e – медианные значения невязки; MAD – медианное абсолютное отклонение

сплошными линиями – величины медианного абсолютного отклонения ошибок $MAD = \text{median}(\Delta t_i - M_e)$. Мы использовали данные величины в качестве робастных оценок средней ошибки и её стандартного отклонения (стандартное отклонение нормально распределённой величины может быть оценено как $\sigma \approx MAD/0.6745$). Прежде всего такая необходимость была обусловлена большим количеством ложных «срабатываний» метода на основе критерия *AIC*. На первом графике (рис. 14а) показаны ошибки выделения времён прихода на данных с относительно большим соотношением сигнал/шум (следует помнить, что величина *SNR* вычислялась для пиковой амплитуды сигнала, амплитуда самой *p*-волны была, как правило, существенно меньше). При соотношении сигнал/шум выше 20 оценка *MAD* стандартного отклонения ошибки выделения вступлений с помощью модели примерно вдвое превышает результаты работы метода *AIC*. С увеличением амплитуды шума *MAD* выделения по методу *AIC* увеличиваются, и уже при *SNR* порядка 5 становятся в несколько раз больше оценки стандартного отклонения ошибок выделения с помощью модели (рис. 14б). Отчасти, столь низкое качество работы стандартного метода на сильно зашумлённых данных может объясняться недостаточно высокой частотой дискретизации использованных данных. Мы планируем исправить данный недостаток в ходе будущей работы. Тем не менее, мы считаем, что сопоставление можно считать корректным, так как оба испытываемых метода находятся в одинаковых условиях.

Результаты и выводы

Предложена модель глубокого обучения для выделения вступлений отдельных фаз на записях волновых форм группы приёмных станций при проведении микросейсмического мониторинга. Для обучения модели использованы полностью синтетические данные. Новизна предлагаемого метода заключается в использовании в модели *Wavelet Scattering* для выделения признаков и механизма самовнимания типа *Transformer encoder*, что позволило реализовать обмен информацией между произвольным числом принимающих станций и, тем самым, предсказывать времена вступлений на отдельных станциях на основе информации всей группы.

Исследован ряд стратегий аугментации обучающей выборки путём наложения шума с целью увеличения устойчивости модели к шуму в данных. Наилучшее соотношение устойчивости к шуму и точности модели было получено при наложении Гауссовского шума со случайным стандартным отклонением, варьируемым в пределах $0 < \sigma_n \leq 0.1$.

Стандартная ошибка выделения времён вступлений *p*- и *s*-волн полученной модели составила порядка 0.03 секунды при соотношении сигнал/шум порядка 10, что может считаться хорошим результатом с учётом того, что в работе использовались записи сигналов с частотой дискретизации 100 Гц. Повышение точности модели может быть достигнуто за счёт использования других видов шума для аугментации обучающей выборки, в том числе и реальных записей сейсмического шума.

Финансирование

Работа выполнена в рамках государственного задания Министерства науки и высшего образования РФ (тема № 122032900167-1).

Список литературы

- Andén J., Mallat S. Deep scattering spectrum // IEEE Transactions on Signal Processing. 2014. Vol. 62 (16). P. 4114–4128. <https://doi.org/10.1109/TSP.2014.2326991>
- Andreux M., Angles T., Exarchakis G., Leonarduzzi R., Rochette G., Thiry L., Zarka J., Mallat S., Andén J., Belilovsky E., Bruna J., Lostanlen V., Chaudhary M., Hirn M.J., Oyallon E., Zhang S., Cella C., Eickenberg M. Kymatio: Scattering transforms in python // Journal of Machine Learning Research. 2020. Vol. 21. P. 1–6. <https://doi.org/10.48550/arXiv.1812.11214>
- Anikiev D., Birnie C., Waheed U., Alkhalifah T., Gu C., Verschuur D.J., Eisner L. Machine learning in microseismic monitoring // Earth-Science Reviews. 2023. Vol. 239. P. 104371. <https://doi.org/10.1016/j.earscirev.2023.104371>

- Beyreuther M., Barsch R., Krischer L., Megies T., Behr Y., Wassermann J. ObsPy: A Python Toolbox for Seismology // Seismological Research Letters. 2010. Vol. 81 (3). P. 530–533. <https://doi.org/10.1785/gssrl.81.3.530>
- Daubechies I. Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics. 1992. <https://doi.org/10.1137/1.9781611970104>
- Guo C., Zhu T., Gao Y., Wu S., Sun J. AEnet: Automatic Picking of P-Wave First Arrivals Using Deep Learning // IEEE Transactions on Geoscience and Remote Sensing. 2021. Vol. 59 (6). P. 5293–5303. <https://doi.org/10.1109/TGRS.2020.3010541>
- He C., Liu J., Zhu Y., Du W. Data Augmentation for Deep Neural Networks Model in EEG Classification Task: A Review // Frontiers in Human Neuroscience. 2021. Vol. 15. <https://doi.org/10.3389/fnhum.2021.765525>
- Heimann S., Kriegerowski M., Isken M., Cesca S., Nooshiri N. Pyrocko: A Versatile Software Framework for Seismology. 2018. Vol. 20. P. 17314. https://greens-mill.pyrocko.org/vogtland_malek2004_surface-3dc953. (n.d.).
- Maeda N. A Method for Reading and Checking Phase Time in Auto-Processing System of Seismic Wave Data // Zisin (Journal of the Seismological Society of Japan. 2nd Ser.). 1985. Vol. 38 (3). P. 365–379. https://doi.org/10.4294/zisin1948.38.3_365
- Málek J., Horálek J., Janský J. One-Dimensional qP-Wave Velocity Model of the Upper Crust for the West Bohemia/Vogtland Earthquake Swarm Region // Studia Geophysica et Geodaetica. 2005. Vol. 49 (4). P. 501–524. <https://doi.org/10.1007/S11200-005-0024-2>
- Mousavi S.M., Ellsworth W.L., Zhu W., Chuang L.Y., Beroza G.C. Earthquake transformer – an attentive deep-learning model for simultaneous earthquake detection and phase picking // Nature Communications. 2020. Vol. 11 (1). <https://doi.org/10.1038/s41467-020-17591-w>
- Stepnov A., Chernykh V., Konovalov A. The seismo-performer: A novel machine learning approach for general and efficient seismic phase recognition from local earthquakes in real time // Sensors. 2021. Vol. 21 (18). <https://doi.org/10.3390/s21186290>
- Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser Ł. Polosukhin I. Attention is all you need. Advances in Neural Information Processing Systems. 2017. Decem(Nips). P. 5999–6009. <https://doi.org/10.48550/arXiv.1706.03762>
- Zhang G., Lin C., Chen Y. Convolutional neural networks for microseismic waveform classification and arrival picking // Geophysics. 2020. Vol. 85 (4). P. WA227–WA240. <https://doi.org/10.1190/geo2019-0267.1>
- Zheng J., Lu J., Peng S., Jiang T. An automatic microseismic or acoustic emission arrival identification scheme with deep recurrent neural networks // Geophysical Journal International. 2018. Vol. 212 (2). P. 1389–1397. <https://doi.org/10.1093/gji/ggx487>
- Zhu L., Peng Z., McClellan J., Li C., Yao D., Li Z., Fang L. Deep learning for seismic phase detection and picking in the aftershock zone of 2008 Mw7.9 Wenchuan Earthquake // Physics of the Earth and Planetary Interiors. 2019. Vol. 293. <https://doi.org/10.1016/j.pepi.2019.05.004>
- Zhu W., Beroza G.C. PhaseNet: A deep-neural-network-based seismic arrival-time picking method // Geophysical Journal International. 2019. Vol. 216 (1). P. 261–273. <https://doi.org/10.1093/gji/ggy423>

ПРИЛОЖЕНИЕ 1.

Таблица параметров испытаний

ID испытания	σ_n	σ_{n_corr}	Доля «спайков»	σ_{spike}	Случайная амплитуда шумов	Успешное обучение модели	Стандартная ошибка выделения p-волны при $\sigma_{test} \approx 0.015$, с	Стандартная ошибка выделения p-волны при $\sigma_{test} \approx 0.07$, с
c29e8_23	0.1	0.01	0.1	0.1	да	да	0.142	0.210
c29e8_22	0.1	—	0.1	0.1	да	нет	—	—
c29e8_21	0.01	0.01	0.1	0.1	да	да	0.326	0.407
c29e8_20	0.01	—	0.1	0.1	да	да	0.050	0.164
c29e8_19	—	0.01	0.1	0.1	да	да	0.042	0.182
c29e8_18	—	—	0.1	0.1	—	да	0.074	0.652

Продолжение. ПРИЛОЖЕНИЕ 1.

c29e8_17	0.1	0.01	–	–	да	нет	–	–
c29e8_16	0.1	–	–	–	да	да	0.035	0.091
c29e8_15	0.01	0.01	–	–	да	да	0.041	0.219
c29e8_14	0.01	–	–	–	да	да	0.511	0.637
c29e8_13	–	0.01	–	–	да	да	0.466	0.493
c29e8_12	–	–	–	–	–	да	0.651	0.643
c29e8_11	0.1	0.01	0.1	0.01	да	да	0.069	0.106
c29e8_10	0.1	–	0.1	0.01	да	нет	–	–
c29e8_9	0.01	0.01	0.1	0.01	да	да	0.043	0.366
c29e8_8	0.01	–	0.1	0.01	да	да	0.111	0.576
c29e8_7	–	0.01	0.1	0.01	да	да	0.168	0.627
c29e8_6	–	–	0.1	0.01	да	да	0.094	0.409
c29e8_5	0.1	0.01	–	–	да	да	0.032	0.565
c29e8_4	0.1	–	–	–	да	да	0.244	0.399
c29e8_3	0.01	0.01	–	–	да	да	0.041	0.256
c29e8_2	0.01	–	–	–	да	да	0.035	0.257
c29e8_1	–	0.01	–	–	да	да	0.065	0.357
c29e8_0	–	–	–	–	–	да	0.648	0.659
598cf_23	0.1	0.01	0.1	0.1	нет	нет	–	–
598cf_22	0.1	–	0.1	0.1	нет	нет	–	–
598cf_21	0.01	0.01	0.1	0.1	нет	нет	–	–
598cf_20	0.01	–	0.1	0.1	нет	да	0.047	0.120
598cf_19	–	0.01	0.1	0.1	нет	да	0.079	0.260
598cf_18	–	–	0.1	0.1	–	да	0.101	0.170
598cf_17	0.1	0.01	–	–	нет	нет	–	–
598cf_16	0.1	–	–	–	нет	да	0.061	0.074
598cf_15	0.01	0.01	–	–	нет	нет	–	–
598cf_14	0.01	–	–	–	нет	да	0.037	0.318
598cf_13	–	0.01	–	–	нет	да	0.162	0.596
598cf_12	–	–	–	–	–	да	0.425	0.509
598cf_11	0.1	0.01	0.1	0.01	нет	да	0.456	0.456
598cf_10	0.1	–	0.1	0.01	нет	да	0.293	0.293
598cf_9	0.01	0.01	0.1	0.01	нет	да	0.083	0.232
598cf_8	0.01	–	0.1	0.01	нет	да	0.045	0.192
598cf_7	–	0.01	0.1	0.01	нет	да	0.052	0.527
598cf_6	–	–	0.1	0.01	–	да	0.194	0.565
598cf_5	0.1	0.01	–	–	нет	нет	–	–
598cf_4	0.1	–	–	–	нет	нет	–	–
598cf_3	0.01	0.01	–	–	нет	да	0.037	0.405
598cf_2	0.01	–	–	–	нет	да	0.032	0.340
598cf_1	–	0.01	–	–	нет	да	0.385	0.395
598cf_0	–	–	–	–	–	да	0.673	0.692

APPLICATION OF SYNTHETIC WAVEFORMS FOR CREATING A DEEP LEARNING MODEL FOR IDENTIFYING FIRST ARRIVALS IN SEISMIC RECORDS

© 2023 N. A. Baryshnikov^{1,*}, I. A. Abzalilov^{1,2}, S. B. Turuntaev^{1,2}

¹*Sadovsky Institute of Geospheres Dynamics of Russian Academy of Sciences, Moscow, Russia*

²*Moscow Institute of Physics and Technology (National Research University), Moscow, Russia*

**E-mail: nabarysh@gmail.com*

The goal of this study was to develop a method for analyzing signal records from a group of stations acquired during microseismic monitoring using a neural network to identify the moments of first arrivals. The novelty of the proposed method lies in the use of Wavelet Scattering for feature extraction, which significantly simplifies the model training process compared to traditionally used convolutional neural networks. The Transformer architecture used in the model allows for information exchange between an arbitrary number of recording stations, thereby determining arrival times at individual stations based on information from the entire group. The model was trained on fully synthetic data. This approach, compared to training on real data, may lead to a model with better generalization capabilities, as synthetic data reflect the mechanics of the generation and propagation of elastic waves without being influenced by local conditions and signal registration features. Several data augmentation strategies were explored, including adding noise to increase the model's robustness to noise. The best balance between noise robustness and model accuracy was achieved by adding Gaussian noise with random standard deviation. The results of this study can be used in the development of deep learning models trained on both synthetic and real seismic records.

Keywords: microseismic monitoring, machine learning, synthetic waveforms, phase picking.