ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ

ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

###### Факультет Высшая школа бизнеса

Сторожок Мария Константиновна

**ОЦЕНКА ПОТРЕБНОСТЕЙ МАЛОГО БИЗНЕСА СЕЙСМООПАСНЫХ РАЙОНОВ В РЕСУРСАХ**

Выпускная квалификационная работа - БАКАЛАВРСКАЯ РАБОТА

по направлению подготовки *38.03.05*

образовательная программа «Бизнес-информатика»

|  |  |
| --- | --- |
|  | Научный руководитель:доктор технических наук,заслуженный профессорАлескеров Фуад Тагиевич |

Москва, 2023

**ОГЛАВЛЕНИЕ**

[ВВЕДЕНИЕ 3](#_Toc135937644)

[ОСНОВНАЯ ЧАСТЬ 6](#_Toc135937645)

[1 Что мы знаем о землетрясениях? 6](#_Toc135937646)

[1.1 Что такое землетрясение и почему оно происходит? 6](#_Toc135937647)

[1.2 Как землетрясения влияют на человечество и инфраструктуру? Исторические примеры 15](#_Toc135937648)

[1.3 Какие данные сейчас собираются для прогнозирования землетрясений? 30](#_Toc135937649)

[2 Как сегодня прогнозируются землетрясения 33](#_Toc135937650)

[2.1 Обзор литературы по теме 33](#_Toc135937651)

[2.2 Проблемы более сложных моделей 45](#_Toc135937652)

[3 Прогнозирование последствий землетрясений в условиях глубокой неопределенности 48](#_Toc135937653)

[4 Построение модели для прогнозирования наступления землетрясения 50](#_Toc135937654)

[4.1 Данные 50](#_Toc135937655)

[4.2 Обучение и тестирование модели 60](#_Toc135937656)

[4.3 Метрики качества 61](#_Toc135937657)

[4.4 Результаты и обсуждение 62](#_Toc135937658)

[4.5 Рекомендации по улучшению результатов 63](#_Toc135937659)

[4.6 Ограничения настоящего исследования 64](#_Toc135937660)

[ЗАКЛЮЧЕНИЕ 65](#_Toc135937661)

[СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ 67](#_Toc135937662)

[ПРИЛОЖЕНИЕ А 71](#_Toc135937663)

[ПРИЛОЖЕНИЕ Б 73](#_Toc135937664)

[ПРИЛОЖЕНИЕ В 74](#_Toc135937665)

[ПРИЛОЖЕНИЕ Г 75](#_Toc135937666)

[ПРИЛОЖЕНИЕ Д 76](#_Toc135937667)

[ПРИЛОЖЕНИЕ Е 77](#_Toc135937668)

[ПРИЛОЖЕНИЕ Ж 78](#_Toc135937669)

[ПРИЛОЖЕНИЕ И 79](#_Toc135937670)

# **ВВЕДЕНИЕ**

Землетрясения представляют собой одну из основных опасностей во многих регионах мира, в том числе в Калифорнии. Более того, на ликвидацию последствий землетрясений и вызванных ими наводнений, пожаров, оползней и лавин расходуется значительная часть национального бюджета. Это служит причиной появления обширных научных исследований, посвященных проблеме предсказуемости параметров землетрясений, таких как дата, время и место возникновения, а также магнитуда, и изучению использования различных методов машинного обучения, нейронных сетей и других математических подходов для ее решения. Тем не менее, несмотря на высокую степень научной проработанности темы, на сегодняшний день не существует технологии, обеспечивающей высокую точность сейсмологических прогнозов, что, в свою очередь, не позволяет предсказывать потребность в ресурсах в сейсмоопасных областях с желаемым уровнем надежности. Таким образом, тема, выбранная в рамках настоящего исследования, является актуальной, – предложение эффективного метода прогнозирования землетрясений, который не обсуждался ранее, позволит получать более аккуратные оценки последствий землетрясений и необходимых для их ликвидации ресурсов, что лежит в основе решения проблемы сильных землетрясений.

Успешное прогнозирование землетрясений и последующая точная предварительная оценка их последствий важны как для бизнесов, чьи офисы и склады расположены в сейсмоопасных районах, так и для государственных организаций и жителей, проживающих в зоне сейсмической опасности. Землетрясения становятся причиной многочисленных человеческих жертв, разрушают как жилые постройки, так и заводы и производственные и складские помещения, нарушают подачу электроэнергии на производствах и не только. Это влечет огромные потери для бизнеса и ведет к сокращению рабочих мест. Помимо прочего зачастую также наблюдается уменьшение предложения на рынке труда, так как определенная часть потенциальной рабочей силы пострадала от землетрясения. Таким образом, создание новой модели предсказания землетрясений позволит решить проблему, являющуюся общей как для компаний, так и для обычных людей, – предоставит возможность заранее рассчитывать последствия землетрясения и нивелировать ущерб.

Цель работы заключается в повышении качества получаемых оценок потребностей в ресурсах, необходимых для устранения последствий сейсмических событий, путем расширения исследований применимости различных методологий и типов исходных данных для прогнозирования сильных землетрясений.

Для достижения поставленной цели в рамках настоящей работы формулируются следующие задачи: исследование методов и типов данных, используемых в существующих подходах к предсказанию землетрясений и оценке на основе данных прогнозов потребности в ресурсах, разработка нового метода, базирующегося на применении случайного леса к двум наиболее распространенным типам данных – геофизическим и сейсмологическим, создание набора данных для работы с моделью, а также изучение влияния комбинирования двух типов данных на качество модели.

Таким образом, объектом исследования является более точное предсказание землетрясений, что позволит повысить надежность прогнозов потребностей в ресурсах. Предметом исследования является применимость метода случайного дерева и двух разных типов данных для построения более точных прогнозов сильных землетрясений.

Методы исследования – сбор и обработка временных рядов значений предикторов, вычисленных на основе геофизических и сейсмологических данных, обучение и тестирование метода случайного леса на отдельных частях предсказательных переменных.

Основным источником информации, использованным при проведении исследования, является ГИС-платформа, хранящая временные ряды значений рассматриваемых признаков.

Хотя метод случайного леса уже не раз рассматривался в рамках решения обозначенной задачи, ранее он не применялся к сейсмологическим и геодинамическим данным одновременно – в этом заключается научная новизна текущей работы. Практическая же значимость настоящего исследования состоит в готовности к немедленному использованию новой концепции в жизни, так как рассматриваемые в ней данные доступны для большого числа сейсмоопасных регионов в достаточном количестве. Это позволит незамедлительно получать прогнозы наступления скорого землетрясения, необходимые для более точного определения потребностей в ресурсах.

Работа имеет определенную структуру. Раздел 1 рассказывает о причинах возникновения землетрясений, мерах оценки их мощности, их влиянии на человечество и инфраструктуру, принципе работы систем раннего предупреждения, современных сейсмоустойчивых постройках и некоторых других аспектах, связанных с появлением землетрясений. Раздел 2 посвящен обзору ранее опубликованных работ, рассматривающих разнообразные методики прогнозирования возникновения сейсмических толчков. Раздел 3 предназначен для ознакомления с подходом к прогнозированию последствий землетрясений в условиях глубокой неопределенности. Раздел 4 описывает процесс сбора и обработки данных и построение модели, а также представляет результаты применения обученной модели к разным выборкам предсказательных переменных. Наконец, в заключении дается краткое изложение и критический анализ результатов.

# **ОСНОВНАЯ ЧАСТЬ**

1. **Что мы знаем о землетрясениях?**
   1. **Что такое землетрясение и почему оно происходит?**

Чтобы лучше понять, как, где и почему зарождаются землетрясения, взглянем на строение Земли, обратившись для этого к учебнику [1], глава 2. Наша планета состоит из внутреннего и внешнего ядер, мантии и коры. Рисунок 1 демонстрирует данную структуру. При этом мантию разделяют на верхнюю и нижнюю. Кора и верхняя часть мантии образуют холодный твердый слой, называемый литосферой. Литосфера разбита на тектонические плиты – сегодня это восемь крупных, десятки средних и множество мелких плит.



Рисунок 1 – Строение Земли

Как и почему происходят землетрясения? - BBC Russian [Электронный ресурс]. – 2015. – URL: https://www.youtube.com/watch?v=5PQR9HOOD7o (дата обращения 11.07.2022)

Под литосферой находится слой горячей жидкой расплавленной породы, которая называется астеносферой, что иллюстрирует рисунок 2, – это самый пластичный слой верхней мантии Земли. Астеносфера жидкая благодаря так называемой конвекции мантии – это очень медленное ползучее движение твердой мантии Земли, вызванное конвекционными потоками, переносящими тепло из недр на поверхность планеты. Мантийная конвекция, показанная на рисунке 3, непрерывна, из-за чего тектонические плиты скользят по слою мантии, и поверхность Земли постоянно движется, очень медленно, со скоростью примерно нескольких сантиметров в год. Поскольку все плиты плотно прилегают друг к другу, движение любой из них действует на окружающие плиты, заставляя и их постепенно перемещаться. Из-за вызванного конвекцией мантии движения плит в разных направлениях и происходят землетрясения. Места контакта крупных плит образуют так называемые сейсмоактивные пояса, в пределах которых происходит подавляющая часть сильных землетрясений.

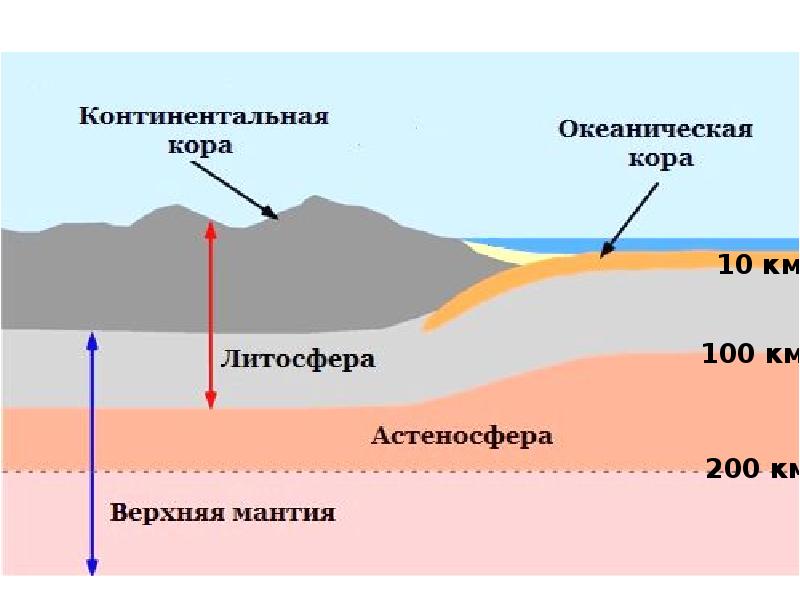


Рисунок 2 – Строение литосферы Земли

Внутреннее ядро. Внешнее ядро. Нижняя мантия. Верхняя мантия. Земная кора [Электронный ресурс]. – URL: https://myslide.ru/presentation/vnutrennee-yadroxAvneshnee-yadroxAnizhnyayaxAmantiyaxAverxnyayaxAmantiyaxAzemnaya-kora (дата обращения 11.07.2022)

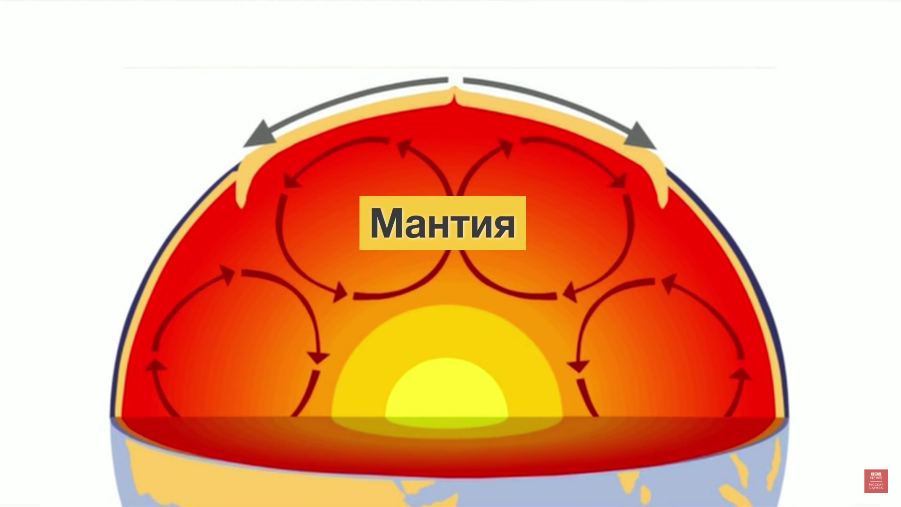


Рисунок 3 – Мантийная конвекция

Как и почему происходят землетрясения? - BBC Russian [Электронный ресурс]. – 2015. – URL: https://www.youtube.com/watch?v=5PQR9HOOD7o (дата обращения 11.07.2022)

Для лучшего понимания сейсмологического процесса важно отметить, что плиты могут взаимодействовать по-разному: они могут сталкиваться, расходиться или двигаться бок о бок [1], глава 3.

Плиты могут двигаться навстречу друг другу. Причем это могут быть две континентальные (материковые) плиты, плиты континентальной и океанической коры или две океанические плиты. В первом случае, поскольку обе плиты примерно одинаковой плотности, ни одна из них «не хочет уступать». Поэтому они сгибаются, деформируются и выдавливаются вверх, образуя горы. Такое явление получило название коллизии. Данный тип столкновения литосферных плит изображен на рисунке 4а. Так например, десятки миллионов лет назад столкновение Индо-Австралийской и Евразийской плит образовало Гималаи. Однако, эти плиты до сих пор продолжают сталкиваться, поэтому Гималаи становятся выше на несколько миллиметров каждый год. Второй вариант схождения плит состоит в столкновении океанической и материковой плит. Более тяжелая океаническая плита погружается под материковую, формируя при этом глубоководные желоба – длинные, узкие и очень глубокие впадины, – достигает мантии и переплавляется в магму, а более легкая материковая плита выдавливается вверх, образуя горные хребты. Так плита Наска уходит под Южноамериканскую плиту с образованием Анд. Помимо этого, движение плит друг по другу нередко вызывает землетрясения, чему свидетельствуют частые землетрясения в Чили. Кроме того, при опускании одной плиты под другую образуется магма, что приводит к возникновению вулканов в местах, где какая-либо плита уходит под другую. Поддвиг океанической плиты под континентальную представлен на рисунке 4б. В случае столкновения двух плит с океанической земной корой одна из плит заползает под другую, из-за чего формируются глубоководные желоба, а также образуются подводные вулканы, которые, с накоплением застывшей лавы, затем нередко превращаются в группы островов. Примером здесь может служить Марианская впадина, или Марианский желоб, являющаяся самой глубокой частью океана – ее глубина равна 11 км, что больше высоты Эвереста, которая составляем чуть меньше 9 км над уровнем моря, – которая образована схождением Тихоокеанской и Филиппинской плит. А неподалеку – Марианские острова. Процесс поддвига океанической плиты под континентальную или другую океаническую называется субдукцией, а границы, вдоль которых происходит столкновение плит, - конвергентными границами. Интересно, что на субдукционные границы приходится около 80% протяженности всех конвергентных границ.

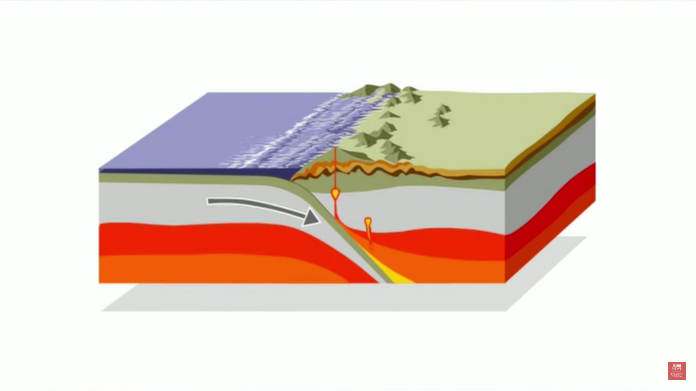
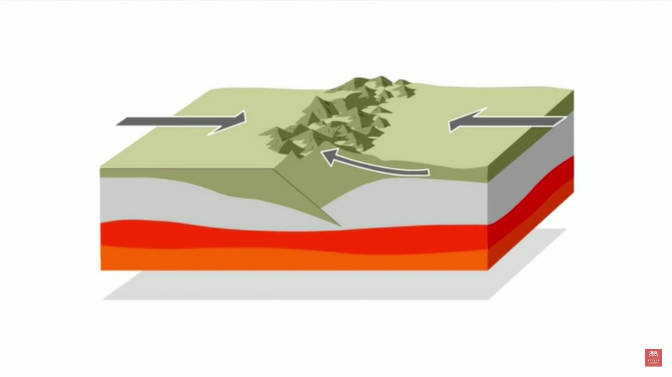


Рисунок 4 – Столкновение литосферных плит

Как и почему происходят землетрясения? - BBC Russian [Электронный ресурс]. – 2015. – URL: https://www.youtube.com/watch?v=5PQR9HOOD7o (дата обращения 11.07.2022)

Плиты могут отдаляться друг от друга. Это позволяет магме подниматься к поверхности. Магма застывает, образуя новую земную кору на краях плит. В результате под водой образуются горные хребты, т.е. трещины прямо посередине дна океана. Примером является Срединно-Атлантический хребет, пролегающий по Атлантическому океану. На суше же образуются рифты – крупные разломы в земной коре. Например, Великая рифтовая долина в Африке. Иллюстрация такого типа взаимодействия тектонических плит приведена на рисунке 5. Границы, вдоль которых происходит раздвижение плит, называются дивергентными границами, а само расхождение литосферных плит называется спредингом.

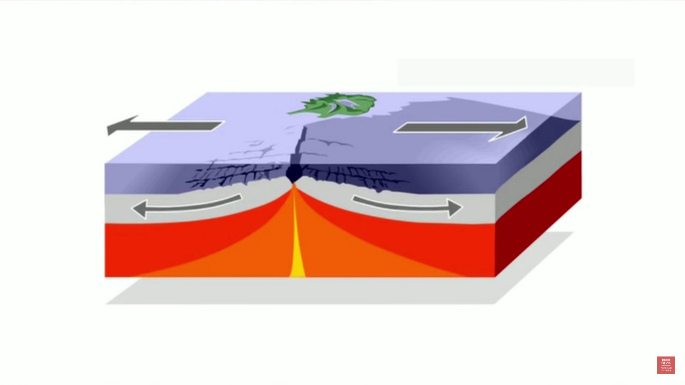


Рисунок 5 – Расхождение литосферных плит

Как и почему происходят землетрясения? - BBC Russian [Электронный ресурс]. – 2015. – URL: https://www.youtube.com/watch?v=5PQR9HOOD7o (дата обращения 11.07.2022)

Наконец, плиты могут скользить мимо друг друга, в противоположных направлениях с разной скоростью. Это явление получило название трансформного разлома, а границы, вдоль которых происходят сдвиговые смещения плит, - трансформных границ. Такой способ сдвига плит представлен на рисунке 6. Примером здесь является знаменитый разлом Сан-Андреас, расположенный на месте встречи Североамериканской и Тихоокеанской литосферных плит. Из-за трения плит друг об друга создается напряжение, которое постепенно нарастает. Это напряжение также называют энергией литосферы или геотермальной энергией. Она представляет собой естественное тепло нашей планеты и находит широкое применение для обогрева жилых помещений и теплиц, в лечебных целях. Напряжение растет до тех пор, пока не превысит сопротивление контактирующих пород. В результате происходит разрыв пород и их смещение. В этот момент внутриземное напряжение скачкообразно сбрасывается и энергия высвобождается в разные стороны от места разрыва в виде сейсмических волн, что ощущается на поверхности как землетрясение.

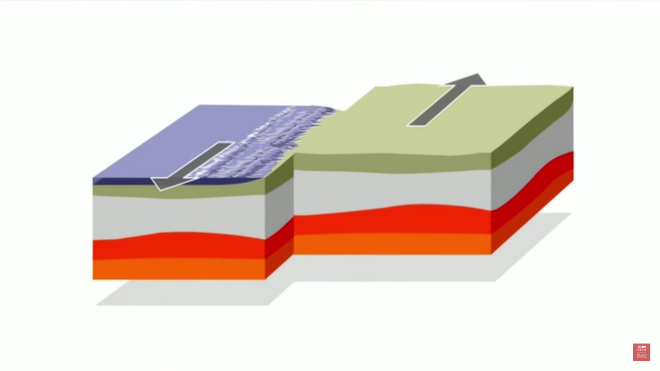


Рисунок 6 – Движение литосферных плит бок о бок

Как и почему происходят землетрясения? - BBC Russian [Электронный ресурс]. – 2015. – URL: https://www.youtube.com/watch?v=5PQR9HOOD7o (дата обращения 11.07.2022)

Сейсмический разрыв земной коры всегда зарождается на глубине, чаще всего 30-40 км и очень редко на меньшей глубине. Точка под поверхностью Земли, в которой начинается землетрясение, называется его гипоцентром, или очагом. Это пространство внутри Земли, где произошло разламывание горных пород.

При этом энергия землетрясения выделяется в двух видах: посредством объемных и поверхностных волн, – согласно [2].

Объемные волны бывают продольными и поперечными. Первичные Р-волны, также продольные волны или волны сжатия, ощущаются как внезапный толчок. Вторичные S-волны, по-другому поперечные волны или волны сдвига, доходят на несколько секунд позже и ощущаются как более продолжительные поперечные колебания.

Поверхностные волны расходятся от эпицентра – точки на поверхности Земли, находящейся ровно над гипоцентром, – и доходят последними, после P- и S-волн. Можно выделить два типа поверхностных волн, вызывающих наибольшие разрушения, – волны Релея и волны Лява. Волны Релея создают вращающиеся движения частиц вверх вниз. Волны Лява, названные в честь математика Августа Лява, заставляют поверхность Земли колебаться из стороны в сторону.

Таким образом, два типа волн, объемные и поверхностные, движутся в разных направлениях. Это можно увидеть на рисунке 7.

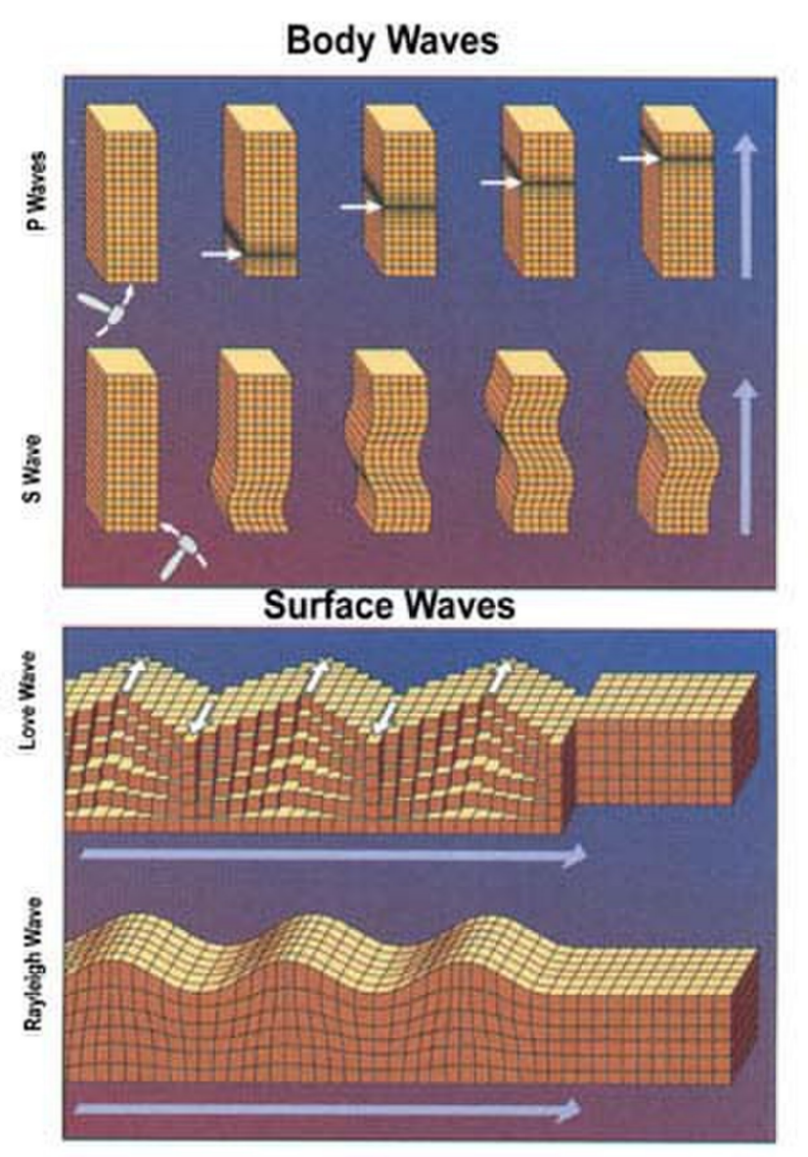


Рисунок 7 – Виды сейсмических волн

Abednego M. Seismic earthquake tomography inversion using artificial neural network in Fribourg, Switzerland. – 2009.

Чтобы оценить мощность землетрясения, согласно [2], вычисляется его магнитуда. Магнитуда – это величина, характеризующая энергию, выделившуюся при разломе горных пород в очаге землетрясения в виде сейсмических волн.

Магнитуда землетрясения измеряется с использованием шкалы Рихтера, которая проходит от 1 до 9,5, где 9,5 соответствует наимощнейшему землетрясению. Каждое увеличение целого числа на шкале означает в 10 раз большее изменение движения грунта, что видно на рисунке 8. Существуют и менее популярные шкалы для измерения магнитуды землетрясения, которые были предложены позже и используются гораздо реже.

Стоит отметить, что в зарегистрированной истории мир лишь однажды испытывал событие магнитудой 9,5 по шкале Рихтера, о чем написано в подразделе 1.2, но сильные землетрясения магнитудой 8 и выше происходят на Земле по крайней мере один раз в год.

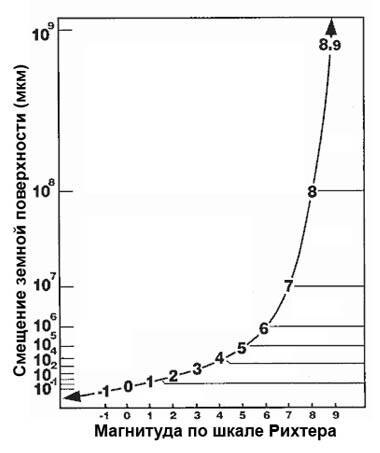


Рисунок 8 – Шкала Рихтера для оценки магнитуды землетрясения

Трагедия в Японии: Из-под земли, из-под воды [Электронный ресурс]. – 2011. – URL: https://www.techinsider.ru/science/11434-tragediya-v-yaponii-iz-pod-zemli-iz-pod-vody/ (дата обращения: 10.05.2023)

Сила землетрясения, интенсивность вызванных им сотрясений на поверхности, как рассказывается в [2], оценивается в баллах по шкале Меркалли. Она связана со внешними проявлениями подземных толчков — воздействием на здания, людей и так далее.

Шкала Меркалли имеет 12 делений, по нарастающей силе проявления землетрясений, где

* 1 балл (незаметное) – землетрясение, при котором только приборы улавливают колебания почвы,
* 12 баллов (катастрофа) – происходит сильное изменение рельефа местности, образуются многочисленные трещины, огромные обвалы и оползни, изменяются русла рек, образуются водопады и озера.

Сила землетрясения зависит от величины магнитуды и расстояния до определенной точки поверхности Земли от очага землетрясения. При одинаковой магнитуде землетрясения, т.е. при одинаковой энергии, высвободившейся при разломе горных пород, сила землетрясения может быть разной в зависимости от глубины очага землетрясения. Чем глубже находится гипоцентр землетрясения, тем меньше сила землетрясения, вызванного подземными толчками, при заданной магнитуде.

Необходимо подчеркнуть, что интенсивность землетрясения зависит не только от его магнитуды и глубины эпицентра. Тип и состояние почвы, геологические особенности региона также влияют на интенсивность.

Обобщая вышесказанное, когда внутри земных недр возникает подземный толчок, выделяется энергия, которая измеряется в магнитудах, возникают сейсмические волны, которые распространяются во все стороны и, достигнув поверхности Земли, вызывают колебания земной поверхности, сила которых измеряется в баллах и определяет последствия, к которым привели эти колебания.

* 1. **Как землетрясения влияют на человечество и инфраструктуру? Исторические примеры**

Большая часть земного шара сейсмически безопасна, только пограничные области между литосферными плитами являются сейсмически опасными. Местности, наиболее подверженные землетрясениям, называются сейсмическими областями или сейсмическими поясами.

Согласно [2], землетрясения скапливаются преимущественно в двух поясах:

* Средиземноморско-Трансазиатском, охватывающим страны Европы (Португалия, Италия, Греция), Ближнего и Среднего Востока (Турция, Иран), Центральной Азии (Северная Индия, Индонезия),
* Тихоокеанском, охватывающем Японию, Китай, Дальний Восток, Камчатку, Сахалин, Курильские острова.

Их можно увидеть на рисунке 9.



Рисунок 9 – Сейсмические пояса Земли

Строение земной коры. Землетрясение [Электронный ресурс]. – 2015. – URL: https://shareslide.ru/geografiya/prezentatsiya-k-uroku-stroenie-zemnoy-koryzemletryasenie (дата обращения 11.07.2022)

В сейсмоопасных районах, таких как, например, Узбекистан, ежедневно происходят небольшие землетрясения, которые не ощущаются людьми, а фиксируются только с помощью специальной аппаратуры. Гораздо реже происходят сильные толчки, которые могут вызвать на поверхности сотрясения с интенсивностью 5−6 баллов и выше. В Узбекистане, например, такие сильные землетрясения происходят в среднем один раз в год. Именно такие землетрясения представляют опасность для людей и инфраструктуры. Цель этой работы – предложить новый метод их предсказания.

Но сначала рассмотрим несколько примеров землетрясений, имевших место в разных точках земного шара и вошедших в историю как самые масштабные, разрушительные и смертоносные.

Самым мощным землетрясением в истории наблюдений землетрясений на планете стало Великое Чилийское землетрясение, или  Вальдивское землетрясение, 1960 года. Его магнитуда составила 9,5, а глубина очага равнялась 33 километрам. Сильные толчки потрясли территорию в 200 000 квадратных километров. Для сравнения, площадь Чили составляет 756 102 квадратных километра. В результате главного удара, эпицентр которого пришелся на окрестности города Вальдивия, и следующей за ним серии более слабых толчков было уничтожено около 20% промышленного потенциала Чили. Непрекращающиеся землетрясения вместе с вызванными ими вулканами и цунами опустошили более 100 000 квадратных километров сельской местности в [Андах](https://ru.wikipedia.org/wiki/%D0%90%D0%BD%D0%B4%D1%8B). Некоторые города были почти полностью разрушены. Последовавшее за землетрясением цунами, волны которого достигали высоты 10 метров, полностью уничтожило некоторые прибрежные города и деревни, смыло дороги. В отдельных городах цунами отбросило часть домов вглубь материка на три километра. Отхлынув, волна двинулась в противоположном направлении и достигла берегов Японии, где разрушила 6 000 и затопила еще 40 000 зданий, а после, пройдя еще 10 000 километров, достигла гавани Хило на Гавайях. После волны цунами дошли до Калифорнии, удалённой от Чили на 9 000 километров. Остатки цунами наблюдали также у берегов Филиппин, Китая, восточной части Новой Зеландии, юго-востока Австралии и Алеутских островов. Землетрясение вызвало также многочисленные оползни, в основном в ледниковых долинах южных Анд.

Число жертв в этом землетрясении составило примерно 6 000 человек. Землетрясение таких масштабов привело бы к еще более многочисленным человеческим жертвам, если бы основной удар не пришелся на малозаселенные местности. Ущерб, нанесенный землетрясением и вызванными им стихийными бедствиями, в ценах 1960 года составил около полумиллиарда долларов. Чтобы проиллюстрировать последствия землетрясения подобной силы, ниже приведен рисунок 10.



Рисунок 10 – Улица в городе Вальдивии после землетрясения 22 мая 1960 года

Великое Чилийское землетрясение [Электронный ресурс]. – 2023. – URL: https://ru.wikipedia.org/wiki/Великое\_Чилийское\_землетрясение (дата обращения 10.05.2023)

Самое же смертоносное землетрясение в истории человечества произошло в провинции Шэньси 23 января 1556 года. Оно получило название «Великое китайское землетрясение». В нем погибло более 830 000 человек. В эпицентре землетрясения открылись 20-метровые провалы и трещины, которые привели к разрушению всех построек. Разрушения затронули даже территории, находившиеся от эпицентра землетрясения в 500 километрах. Кадры схожего по мощности землетрясения в Китае можно увидеть на рисунке 11. Магнитуда этого землетрясения оценивается примерно в 8 по шкале Рихтера, сила же равна 9 по шкале Меркалли, глубина гипоцентра оценивается в 32 километра.

По количеству погибших — это третья катастрофа за всю письменную историю человечества, а из землетрясений — сильнейшее в истории по количеству жертв и одно из сильнейших по магнитуде.

Столь огромные человеческие жертвы обусловлены несколькими причинами. Во-первых, территория Китая исторически всегда была плотно заселена. Во-вторых, основная часть населения провинции жила в непрочных домах, возводившихся на рыхлых, сильно обводненных грунтах, и пещерах, вырытых прямо в склонах холмов. После первых же толчков здания начали погружаться в разжиженный сейсмическими колебаниями грунт, а пещеры мгновенно обрушились либо были затоплены водяными потоками. Сотни тысяч человек оказались погребенными под толстым слоем земли и развалинами своих домов. В-третьих, землетрясение произошло в 5 часов утра, когда большинство людей еще находились в своих нестойких к подземным ударам жилищах. Неспособность предсказать приближающееся землетрясение привела к тому, что почти никто не смог заблаговременно покинуть помещение.



Рисунок 11 – Землетрясение в Китае 1976 года

Землетрясение в Китае 1976 года [Электронный ресурс]. – 2023. – URL: https://bangkokbook.ru/foto/zemletryasenie-v-kitae-1976-god.html (дата обращения 10.05.2023)

Самым смертоносным землетрясением в современной истории является подводное землетрясение, произошедшее в Индийском океане 26 декабря 2004 года в 07:58 по местному времени и вызвавшее цунами, ставшее самым смертоносным в истории человечества. От его волн, высота которых превышала 15 метров и в некоторых случаях достигала 20 метров, погибло, по разным подсчетам, от 225 000 до 300 000 человек в четырнадцати странах, включая Индонезию, Шри-Ланку, Индию, Таиланд, Сомали, Мьянму, Мальдивы, Малайзию, Танзанию, Сейшельские острова, Бангладеш, Южную Африку, Йемен и Кению. Помимо огромного количества погибших, цунами привело к бесчисленным разрушениям, его последствия ощущались даже в Порт-Элизабет, ЮАР, в 6900 километрах от эпицентра. Масштаб воздействия вызванного подземным толчком цунами показан на рисунке 12. Магнитуда землетрясения составила, по разным оценкам, от 9,1 до 9,3, благодаря чему землетрясение входит в тройку самых сильных землетрясений за всю историю наблюдения.

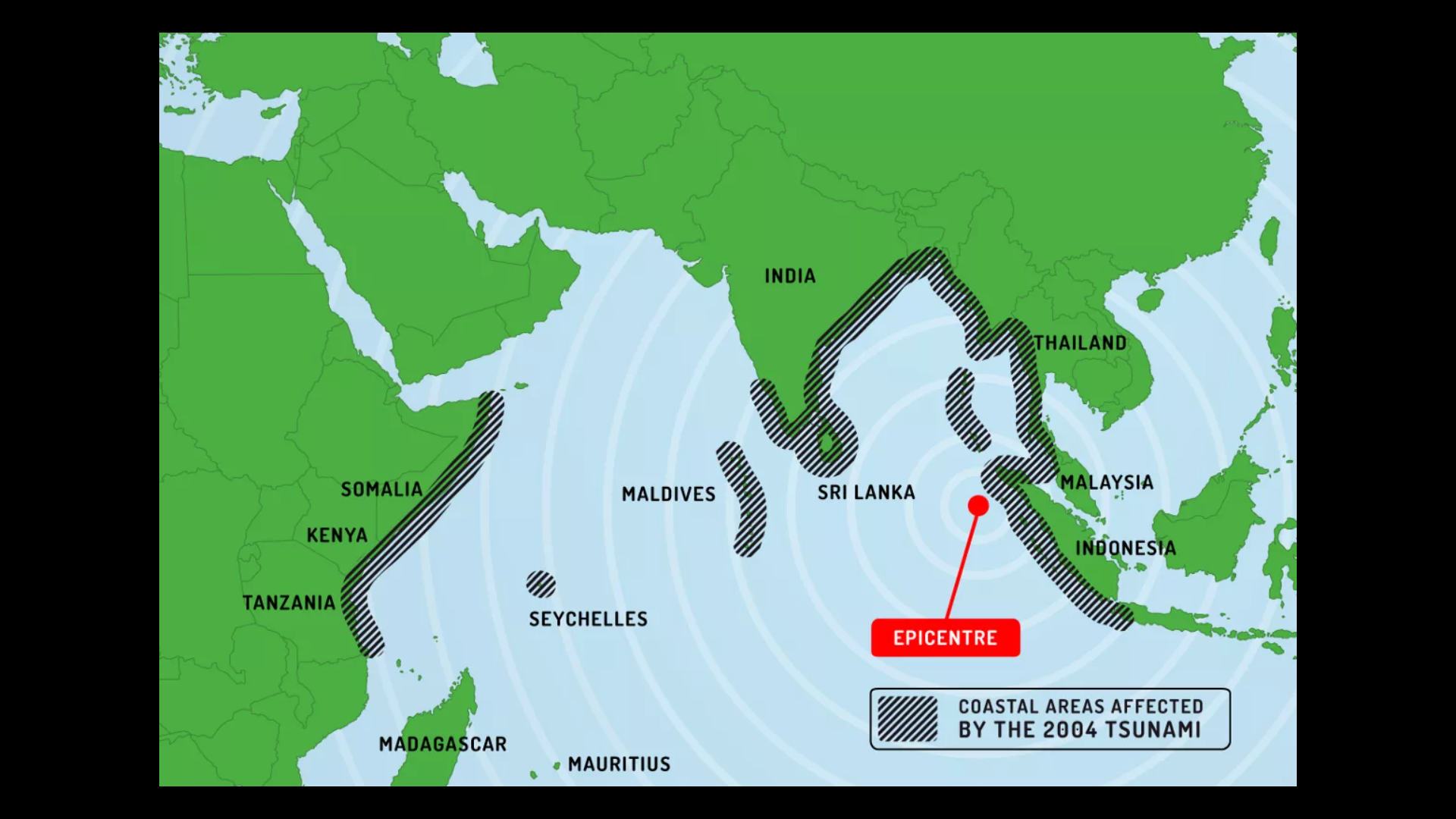


Рисунок 12 – Карта последствий цунами 2004 года

Цунами 2004 года в Малайзии: психологическое воздействие [Электронный ресурс]. – 2017. – URL: https://medium.com/caium/tsunami-de-2004-en-malaysia-un-impact-psychologique-6e81aae0746e (дата обращения 10.05.2023)

Гипоцентр основного землетрясения находился на глубине 30 км от уровня моря в западном конце Тихоокеанского кольца огня (или Тихоокеанского сейсмического пояса, о котором упоминалось ранее) – пояса, в котором происходит свыше 80% всех землетрясений в мире. Кольцо огня продемонстрировано на рисунке 13.

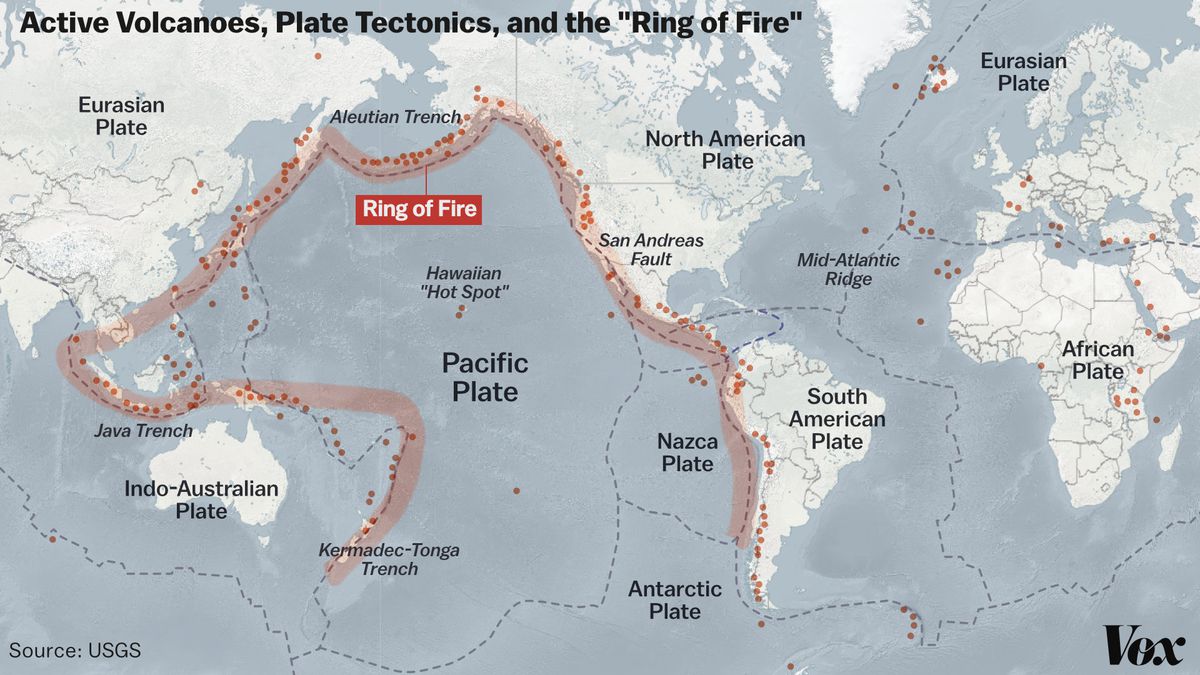


Рисунок 13 – Тихоокеанское кольцо огня

The Geo-Science Behind The Taal Volcano Eruption [Электронный ресурс]. – 2020. – URL: https://nestheprint.com/general/2020/02/12/the-geo-science-behind-the-taal-volcano-eruption/ (дата обращения 10.05.2023)

Землетрясение последовало в результате сдвига около 1200 километров, по некоторым оценкам, 1600 километров, породы на расстояние в 15 метров ввиду поддвига Индийской плиты под Бирманскую плиту.

Еще одним примером разрушительного землетрясения может стать Таншаньское землетрясение, произошедшее в китайском городе Таншане, провинция Хэбэй, 28 июля 1976 года. От первого же удара было разрушено 90% всех городских построек.  Вследствие землетрясения около 5,3 миллионов домов оказались разрушенными или поврежденными настолько, что в них было невозможно жить. Город практически сравнялся с землей. Некоторые районы покрылись множеством огромных трещин. Одна из таких трещин поглотила здание больницы и переполненный пассажирами поезд. Землетрясение магнитудой 8,2 по шкале Рихтера с очагом на глубине 22 километров ощущалось на удалении до 800 километров от эпицентра. По официальным данным властей КНР, землетрясение унесло 242 419 жизней, по неофициальным – до 655 000. Огромное число человеческих жертв объясняется тем, что главный удар произошёл ночью, в 03:42 по местному времени, когда почти все жители спали, и тем, что жилые постройки имели низкое качество и были перенаселены. Именно  отсутствие необходимых мер сейсмозащиты при строительстве стало основной причиной колоссальных разрушений. Землетрясение в Таншане является вторым в истории по количеству жертв после землетрясения в Шэньси в 1556 году.

Помимо цунами, землетрясения зачастую становятся первопричинами многих других катаклизмов, например, пожаров или оползней, что отмечается в главе 26 справочника [3].

Ярким примером здесь может стать землетрясение в Ганьсу, произошедшее 16 декабря 1920 года в китайском уезде Хайюань. Магнитуда первого удара составила 7,8 по шкале Рихтера, его очаг располагался на глубине 17 километров. Интенсивность последовавших за ним толчков в эпицентре достигла максимума в XII баллов по шкале Меркалли. Сейсмические колебания ощущались на территории площадью в четыре миллиона квадратных километров. Наиболее значительные разрушения охватили площадь в 67 500 квадратных километров. Общее число погибших составило 270 000 человек, 20 000  из которых умерли от холода, лишившись своих жилищ, еще 73 000 похоронил под собою вызванный землетрясением оползень.

Другим наглядным примером является Великое землетрясение Канто. Землетрясение магнитудой 8,3 произошло 1 сентября 1923 года у берегов Японии (его очаг располагался в море на глубине 23 километров), охватив площадь около 56 000 квадратных километров, и практически полностью разрушило Токио, Иокогаму, Йокосуку и ещё 8 менее крупных городов, расположенные в регионе Канто, став самым разрушительным за всю историю Японии. Последовавшее за подземными толчками цунами, волны которого достигали высоты в 12 метров, полностью опустошило прибрежные поселения. Начавшиеся повсюду пожары, разносимые сильным ветром, нанесли еще больший ущерб. В Токио только пожаром было уничтожено свыше 300 000 зданий (из миллиона), в Иокогаме подземными толчками было разрушено 11 000 зданий и ещё 59 000 сгорело. Из 675 мостов 360 было уничтожено огнем.

Официальное число погибших — 142 800. Часть из них – жители, пытавшиеся спастись от подземных толчков и пожаров на открытых пространствах, таких как площади или парки. На одной из площадей Токио погибло около 40 000 человек — они задохнулись, когда загорелись окружающие площадь дома. Общее число пострадавших составило около 4 миллионов.

Материальный ущерб, понесенный Японией от землетрясения Канто, оценивается в 4,5 миллиарда долларов, что составляло на тот момент два годовых бюджета страны и в пять раз превышало расходы Японии в Русско-японской войне. Ситуацию ухудшило то, что выдаваемые пострадавшим для восстановления домов деньги и векселя, общая сумма которых превысила 43 миллиона, послужили причиной обесценивания японской валюты и инфляции, которые привели к несостоятельности населения и невозможности вернуть государству долги, что вскоре повергло страну в финансовый кризис. Попытки же правительства прибегнуть к помощи иностранного капитала вызвали резкий рост государственного долга.

Подводя итог, можно сказать, что основными причинами столь пагубного влияния, оказываемого землетрясениями на людей и инфраструктуру, являются неспособность предсказать приближающееся сейсмическое событие, а также отсутствие требований к сейсмоустойчивости зданий при строительстве и учета строительных свойств грунтов. Землетрясения были и остаются одной из основных угроз во многих странах мира. И даже если мы знаем, что в каком-то месте, разломе земной коры на стыке тектонических плит, рано или поздно должно произойти землетрясение, пока что мы не можем предсказать, когда именно это случится. Так произошло в Турции, расположенной на стыке сразу трех тектонических плит: Анатолийской, Аравийской и Африканской. Хоть ученые и ожидали в месте разлома землетрясение, основываясь на методе, описанном далее в настоящем подразделе, они не смогли определить время его наступления. Это привело к трагедии, случившейся 6 февраля 2023 года. Землетрясение было признано самым мощным в Турции после землетрясения 1939 года в Эрзинджане. Более 57 000 погибли и более 121 000 ранены в Турции и Сирии. Экономический ущерб оценивается в более чем 104 миллиарда долларов. Но таких потерь можно было бы избежать, если бы человечество умело предсказывать землетрясения. Это в очередной раз подтверждает актуальность выбранной темы работы.

Ведущие мировые специалисты утверждают, что современная наука не умеет предсказывать землетрясения. Множество методов было предложено: по выходу газа радона [4, 5], электромагнитным сигналам [6-8], изменениям в ионосфере [9-11], сейсмическому затишью [12], уровню и газохимическому составу подземных вод [13, 14], последовательностям подземных толчков и даже нестандартному поведению животных, но ни один не заработал, и регулярных, надежных и точных предсказаний до сих пор не делает никто.

Землетрясение может произойти где угодно, но в некоторых местах его появление вероятнее, чем в других. В регионах, где за всю историю наблюдений не было значимых землетрясений, специалисты не ожидают событий высокой магнитуды. Для регионов, где землетрясения бывали раньше – тут важно отметить, что подробные данные научное сообщество начало собирать лишь с появлением первых современных сейсмографов в 19 веке, – ученые создают карты сейсмической опасности. Пример такой карты представлен на рисунке 14. Однако, полагаться на подобные карты в своих решениях не стоит. Например, жители Москвы, находящейся в белой зоне на карте ниже, в 1977 году ощутили колебания довольно большой магнитуды, являвшиеся отголосками Карпатского землетрясения, до сих пор одного из самых разрушительных в истории Румынии. До Москвы от эпицентра землетрясения по прямой около 2 000 километров. На некоторых здания столицы тогда появились трещины, и мы можем только гадать, сколько зданий выстояло бы, окажись колебания немного сильнее. Тем не менее, именно в местах встреч тектонических плит происходит большинство землетрясений, поэтому все внимание сосредоточено в первую очередь на них.

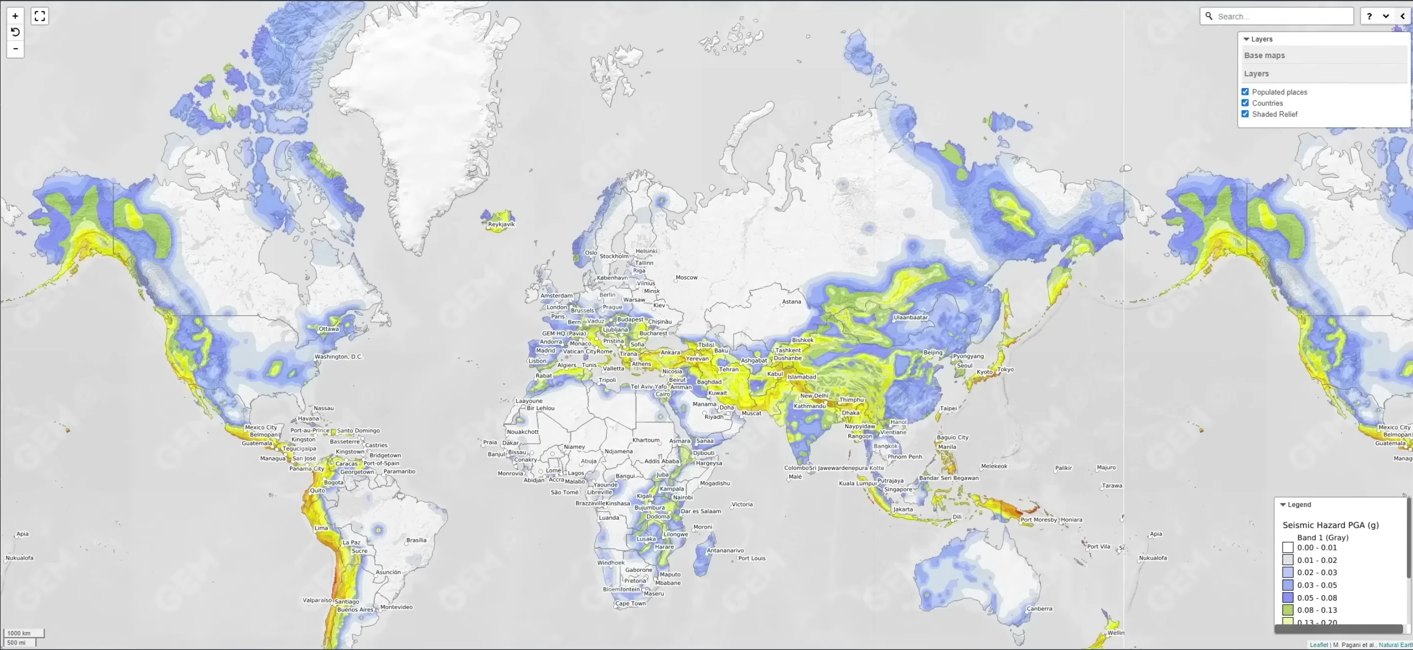


Рисунок 14 – Карта сейсмической опасности

OpenQuake Map Viewer [Электронный ресурс]. – 2023. – URL: https://maps.openquake.org/map/global-seismic-hazard-map/#2/48.3/39.9 (дата обращения 10.05.2023)

Разломы же разделены на сегменты, в определенный момент времени в разных сегментах накоплено разное напряжение вдоль одного разлома. Когда один сегмент смещается и сотрясается, высвобождая огромную энергию в виде землетрясения, это изменяет напряжение на соседних участках того же разлома и других разломах поблизости. Это увеличивает напряжение в одних местах, приближая их к разрушению, но ослабляет напряжение в других, делая их более безопасными на данный момент. Изменения напряжения вдоль разлома, вызываемые новыми землетрясениями, отображаются в модели продолжающегося движения плит, состоящей в составлении карты напряжений вдоль разлома, – этот метод известен как Кулоновский перенос напряжения (Coulomb stress transfer), он рассматривался во многих работах, нацеленных объяснить появление землетрясений и предсказать их [15, 16]. И именно с помощью этого метода уже давно были предсказаны место и даже магнитуда случившегося в Турции в 2023 году землетрясения [17], однако этот метод не предусматривает расчета времени наступления толчка.

Итак, когда разлом приходит в движение, напряжение непременно меняется в других местах вдоль разлома. Это приводит к тому, что каждое, даже самое маленькое, землетрясение меняет сразу всю ситуацию во всех масштабах: на локальном, региональном, глобальном уровнях. И это только часть картины, ведь на каждый разлом влияет сразу большое количество факторов одновременно, от состава пород на разных глубинах на разных участках разлома, а одних только минералов 5 000 видов, и все они по-разному себя ведут при разных нагрузках, до лунной гравитации. А это значит, что взаимодействие больших и малых землетрясений на одном и между разными разломами может быть настолько сложным, что даже если мы будем знать все о конкретном разломе или даже целом регионе, сейсмические события на стыке плит будут так хаотичны, что мы не сможем предугадать их наперед. Это ограничение называют эффектом бабочки, когда самое малое событие, как взмах крыла бабочки, может иметь самые большие последствия.

Чтобы предсказать землетрясение, нужно назвать заранее точные дату и время, место и магнитуду. Можно подумать, что чтобы минимизировать разрушения и жертвы, нам достаточно знать, где случится землетрясение. Тогда мы сможем ввести подходящие строительные нормы и подготовиться. Это спасет множество жизней и инфраструктуру. В реальности это не так. Строить надежно и безупречно – вовсе не страховка от крушений. В этом можно убедиться на примере Японии. В 2011 году по стране ударило сильнейшее землетрясение магнитудой 9. Власти официально насчитали 127 000 полностью разрушенных строений, и все они возводились по стандартам, считавшимся самыми передовыми и безопасными в то время. В сейсмологии существует явление, получившее неформальное название «убийственный пульс землетрясения» [18]. Оно заключается в том, что если подземные толчки следуют друг за другом каждые несколько секунд, то землетрясение даже не самой высокой магнитуды становится намного более разрушительным. Многие ученые утверждают, что от тотального коллапса зданий во время землетрясения 2011 года спасла именно невысокая частота колебаний, а вовсе не хорошая по тем стандартам сейсмоустойчивость построек. Таким образом, говоря о сейсмической безопасности зданий сейчас, эксперты имеют в виду не то, что строения удастся сохранить, а то, что более крепкие здания помогут выиграть время для спасения людей.

Сейчас же сейсмологи по всему миру учатся все точнее засекать начало больших землетрясений, то есть создают системы раннего предупреждения. Как сообщается в источнике [3], для передачи энергии землетрясения через горные породы и почву требуется больше времени, чем для того, чтобы электронное предупреждение достигло места назначения. Имея возможность засечь сейсмические волны начинающегося землетрясения на самой ранней стадии, жители ближайших городов приобретут время, пусть это будет всего несколько секунд или минут, чтобы остановить скоростные поезда и другие транспортные средства, открыть двери лифтов, отключить подачу газа или другие чувствительные системы, а также начать эвакуацию. Пожарные и бригады скорой помощи могут заранее выехать за пределы депо, которые могут обрушиться, чтобы быть готовыми оказать помощь.

Задача же инженеров заключается в том, чтобы снизить колебания зданий. На сегодняшний день разработана такая архитектура, как дом с изолированным фундаментом. Толчки магнитудой 9 ощущаются как легкая дрожь от проходящего мимо поезда в небоскребах, у которых изолирован фундамент. Из-за сложности и дороговизны такого метода только в 9 000 из миллиона зданий в Японии использована такая технология. В зданиях с изолированным основанием практически всю вызванную землетрясением тряску принимает на себя подложка – слой из стальных элементов или резины, – расположенная под фундаментом здания. С той же целью – снизить колебания строения – используют инерционные демпферы – многотонные шары, подвешенные как маятники или емкости с водой в высотном здании. Иллюстрация приведена на рисунке 15. В обоих случаях принцип один и тот же – подложка или демпфер раскачивается не синхронно со всем остальным зданием, а в противоположном направлении, за счет чего гасит колебания, как противовес [19].

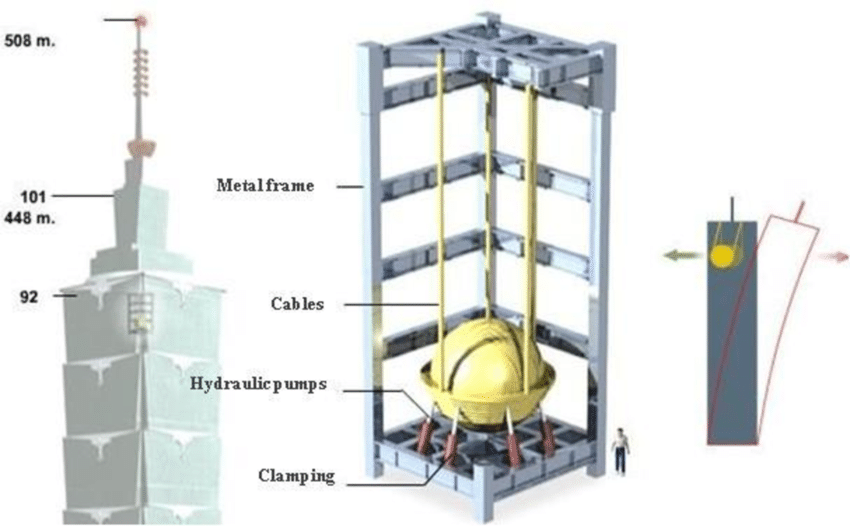


Рисунок 15 – Принцип работы инерционного демпфера

Кто изобрел массовые демпферы? [Электронный ресурс]. – 2023. – URL: https://isaacantisismica.com/ru/кто-изобрел-массовые-демпферы/ (дата обращения 10.05.2023)

Таким образом, в данном подразделе было рассмотрено, к каким последствия может привести землетрясение и как если не избежать их, то хотя бы уменьшить.

### **1.3 Какие данные сейчас собираются для прогнозирования землетрясений?**

Сейсмическую активность фиксируют сейсмические станции. Они регистрируют колебания земной поверхности, вызываемые землетрясениями. Современные сейсмические станции состоят из трехкомпонентного сейсмоприемника (одного вертикального и двух горизонтальных), по-другому сейсмографа, который улавливает движение земли вверх-вниз, а также в северо-южном и западно-восточном направлениях, то есть вдоль соответствующих линий сечения земной поверхности, называемых долготой (вспомним нулевой меридиан) и широтой (вспомним экватор), соответственно. Сейсмографы оснащены электромагнитными датчиками, которые преобразуют колебания грунта в электрические изменения, которые обрабатываются и регистрируются аналоговыми или цифровыми схемами приборов.

Когда происходит землетрясение, фронт упругой волны расширяется от гипоцентра, устремляясь во все стороны от очага, и проходит каждую сейсмическую станцию в определенное время, которое мы регистрируем. Как рассказывается в [1], волны разного типа движутся с разными скоростями, и из наблюдаемых данных можно узнать время прихода P- и S-волн. Разница во времени прибытия P- и S-волн пропорциональна расстоянию от гипоцентра до станции. Имея правильные показания времени прихода продольных и поперечных волн по крайней мере для четырех сейсмостанций, не расположенных на одной прямой, мы можем локализовать источник землетрясения, рассчитав координаты местоположения события и время его начала. При этом характеристики зарегистрированных сейсмических волн – время их появления, амплитуда, период колебаний, длительность затухания и другие параметры – позволяют определять не только положение эпицентра землетрясения, но и магнитуду и интенсивность события. Так, с помощью данных, полученных некоторым числом сейсмологических станций, составляются каталоги землетрясений, в которых хранится информация о уже произошедших сейсмических событиях: дата и время возникновения землетрясений, местоположение гипоцентра, а именно, широта, долгота и глубина очага, магнитуда – основные параметры, присутствующие в таких архивах. Таким образом, для сейсмологов и других ученых и специалистов становятся доступными не только сырые данные в виде сигналов со станций, но и каталоги исторических землетрясений. На их основе исследователи вычисляют самые разные параметры, которые после используют в своих прогностических моделях, рассмотренных более подробно в разделе 2.

Стоит отметить, что именно сейсмометры, являющиеся, в большинстве случаев, частью систем раннего оповещения о землетрясениях, о которых говорилось ранее, вместе с  акселерометрами – устройствами, фиксирующими уже не колебания, а ускорение земли, обнаруживают землетрясение вблизи территорий, на которые сразу отправляется оповещение.

Другим типом данных, к которому прибегают специалисты по всему миру с целью выявления надвигающегося землетрясения, являются данные о движении грунта, регистрируемые с помощью спутниковых систем.

Спутниковая система навигации – это система, предназначенная для определения местоположения (географических координат) объектов, находящихся на Земле. Она также позволяет получить скорость и направление движения объекта, получающего сигналы со спутников станции – приемника сигнала. Принцип работы данных систем состоит в том, что спутник посылает сигнал, в котором, помимо других данных, транслируются время отправки сигнала и координаты спутника на момент отправки сигнала. Приемник принимает этот сигнал, считает разницу между временем отправки и временем получения сигнала и затем, предполагая известной скорость распространения радиоволн, исходящих от спутников, вычисляет расстояние, на котором он находится от спутника. Каждый спутник также передаёт в своём сигнале альманах – таблицу положений всех спутников. Таким образом, получив сигналы от не менее чем 4 спутников и зная расстояния до каждого из них, на основе альманаха, навигационный приемник сможет вычислить свое местонахождение. Все остальные параметры движения приемника (скорость, направление, пройденное расстояние) вычисляются на основе измерения времени, которое он затратил на перемещение между двумя или более точками с определёнными координатами. Так, с помощью систем, включающих сразу большое количество наземных приемников, ведется наблюдение за движением плит, что отмечается в справочнике [3], а также определяется скорость этого движения. Геодезический контроль состоит в постоянном наблюдении ежедневных координат, получаемых в тектонически активных регионах.

Данным, описанным в данном подразделе, посвящается наибольшее внимание в публикуемых попытках научиться предсказывать землетрясения. Тем не менее, в некоторых научных работах находят применение данным, получаемым из ионосферного слоя Земли, электромагнитным сигналам и прочим источникам наблюдений, способных содержать информацию о предвестниках сейсмического события. Как было отмечено ранее, предсказывать землетрясения с высокой точностью не смогли ни с одним из предложенных методов.

## **2 Как сегодня прогнозируются землетрясения**

### **2.1 Обзор литературы по теме**

В последние десятилетия появляется все больше и больше исследований, нацеленных на изучение применения самых разных методов для решения различных сейсмологических проблем. В одних работах рассматривается применение методов машинного обучения при прогнозировании сильных землетрясений, в других – более сложных архитектур, таких как нейронные сети, нередко предпринимаются попытки использования математических моделей для решения данной задачи.

В одних случаях для обучения и последующего тестирования моделей используются сигналы, полученные с сейсмических станций, а также всевозможные рассчитанные по ним статистики. В других же случаях на основе баз данных, хранящих сведения о произошедших в прошлом землетрясениях, или данных, описывающих движения земной поверхности, регистрируемые спутниковыми системами, вычисляются сейсмические и/или геофизические показатели, способные указать на возможность скорого наступления землетрясения в исследуемом регионе.

В следующих работах, которые мы рассмотрим, на вход изучаемым моделям подаются необработанные исходные сигналы, полученные с региональных сейсмических станций, а также некоторые статистики временных рядов, которые и представляют собой дискретизированные сигналы, или вычисленные с помощью этих рядов параметры.

В [20] описывается метод, основанный на сверточной нейронной сети, для прогнозирования 5 измерений интенсивности сотрясения грунта при землетрясении. Входные данные представляют собой необработанные трехкомпонентные сигналы, записанные 39 сейсмическими станциями в одно и то же время – и на выходе модель предсказывает значения 5 измерений интенсивности сотрясения грунта для всех станций. Используемые сигналы начинаются в момент возникновения землетрясения и продолжаются в течение следующих 10 секунд. Концепция заключается в использовании сигналов, содержащих всего несколько секунд после начала события, P-волны которого способно зарегистрировать только небольшое количество ближайших станций, таким образом, можно предсказать измерения интенсивности сотрясения земли для более удаленных станций.

Следует отметить, что для обучения и последующего тестирования модели используются необработанные данные, чтобы повторить ситуацию потоковой передачи данных в реальном времени. По той же причине как в обучающую, так и в тестовую выборки добавляются дополнительные шумовые примеры.

Было обнаружено, что предложенная модель дает полезные оценки движения грунта в течение 15–20 секунд после начала землетрясения, что подтверждает полученное во время тестирования модели значение среднеквадратичной ошибки (Mean Squared Error, MSE), равное 0,176.

Из-за существующего дисбаланса магнитуд, присутствующих в обучающем наборе данных, модель занижает значения прогнозируемых величин для событий больших магнитуд. Использование аугментации данных или трансферного обучения может стать возможным решением проблемы.

В [21] представлен регрессор, который оценивает магнитуду землетрясения по необработанным сигналам одной станции. Входными данными для модели являются трехканальные сейсмограммы сигналов длительностью 30 секунд. Набор данных включает как сигналы, на которых записаны землетрясения, так и сигналы сейсмического шума. Нейронная сеть состоит из двух сверточных слоев, роль которых заключается в уменьшении размерности и извлечении признаков – так как такие архитектуры способны улавливать самые разные паттерны, хранящиеся в исходных данных, – двунаправленного LSTM-слоя, где уже происходит обучение, и полносвязного слоя, состоящего из одного нейрона, служащего для получения выхода модели – предсказанной магнитуды.

Как пишут сами авторы, имея информацию об амплитуде сигнала и расстоянии между регистрирующей станцией и местом события, которое можно узнать из сейсмограммы землетрясения и зафиксированного в ней времени прихода P- и S-волн, как описывалось в подразделе 1.3, нейронная сеть может изучить модель среднего затухания силы волны по мере удаления от места события и, тем самым, определить магнитуду землетрясения.

Модель с такой архитектурой способна предсказывать магнитуду землетрясения со средней ошибкой, близкой к нулю, и стандартным отклонением, близким к 0,2. Коэффициент детерминации (R2, R-квадрат), рассчитанный по предсказаниям обученной модели, равен 0,9.

Производительность можно улучшить, отчистив данные от шума на этапе предварительной обработки или построив более глубокую сеть, однако для этого потребуется больше обучающих данных для предотвращения переобучения. Чтобы адаптировать представленную методологию к предсказанию больших магнитуд, необходимо использовать обучающий набор данных, содержащий большее число больших событий. Использовать аугментацию данных – еще одно потенциальное решение.

Несмотря на то что подходы, предложенные в работах [20] и [21], предоставляют довольно точные предсказания, с их помощью можно прогнозировать землетрясения непосредственно за очень короткий период времени до их наступления. Таким образом, описанные методологии могут применяться в системах раннего предупреждения о землетрясениях, однако они не подойдут для построения долгосрочных сейсмических прогнозов, например, на день, неделю или даже месяц вперед.

Целью [22] является поиск модели, с помощью которой можно предсказывать время, оставшееся до следующего лабораторного землетрясения. В рамках исследования используются разные форматы наборов данных для обучения и тестирования изучаемых моделей. Обучающие данные представляют собой единую сейсмограмму, содержащую многочисленные лабораторные толчки. Тестовый же датасет состоит из набора небольших сегментов той же сейсмограммы.

Авторы описывают наиболее удачные из изученных в рамках проведенного исследования методы, а также некоторые способы преобразования исходных сигналов, благодаря которым удается повысить значения метрик.

Лучшим из изученных методов является градиентный бустинг Light Gradient Boosted Machine (LightGBM) в сочетании с нейронной сетью. К входным сейсмическим данным перед расчетом признаков и обучением модели добавляется синтетический шум. Это помогает сделать модель более устойчивой к шуму и более способной к обобщению. Сгенерированные на основе зашумленных сигналов признаки включают в себя количество пиков в сигнале, а также 20-й перцентиль стандартного отклонения в скользящем окне размера 50. Использование такого предиктора, как «время, прошедшее с момента последнего землетрясения», также улучшает прогностическую способность модели. Средняя абсолютная ошибка (Mean absolute error, MAE) модели равняется 2,2650.

Еще одним методом, показавшим хороший результат – MAE = 2,2960, – является уже другая реализация градиентного бустинга – CatBoost. Главным признаком, используемым при обучении модели, является стандартное отклонение амплитуды сигнала после удаления из него пиков большой амплитуды.

Схожие значения метрики показывают подходы, основанные на применении комбинации алгоритма LightGBM и рекуррентной нейронной сети (Recurrent neural network, RNN), MAE = 2,2968, или алгоритма LightGBM, в который включен подбор гиперпараметров, MAE = 2,2975. При этом для обучения последнего в качестве прогностических параметров используются не только скользящее стандартное отклонение, но и скользящие моменты 3-6 порядков, автокорреляция.

При работе со всеми описанными выше методами исследователи рассматривают только те признаки, которые имеют одинаковое распределение как в обучающей, так и в тестовой выборках, что можно проверить с помощью теста Колмогорова-Смирнова. Такой подход к отбору признаков, во-первых, обеспечивает невозможность использования нестационарных признаков в процессе обучения, а во-вторых, улучшает обобщающую способность модели.

Последняя из оказавшихся успешными методик заключается в построении простой нейронной сети с прямой связью, MAE = 2,2980, чьи результаты значительно улучшаются за счет применения масштабирования времени, оставшегося до следующего лабораторного толчка. Оно заключается в изменении целевой переменной и предсказании не абсолютного времени до следующего землетрясения, а оставшейся доли сейсмического цикла – отрезка времени между двумя сильными землетрясениями в одном месте. Цикл землетрясений относится к явлению, при котором землетрясения неоднократно происходят на одном и том же разломе в результате непрерывного накопления напряжения и его периодического высвобождения. Он может длиться 10, 100 и более лет. Примером здесь может служить участок Паркфилда разлома Сан-Андреас, где землетрясения магнитудой в районе 6 с аналогичным расположением регистрируются приборами каждые 30-40 лет.

Такой метод поспособствовать обучению модели, возможно, является наиболее интересным, с физической точки зрения, из всех описанных в работе.

Главным недостатком полученных моделей является то, что на реальных данных они могут работать не так хорошо, как на искусственных. Еще одним местом для доработки может являться то, что предложенные архитектуры способны предсказывать время до следующего толчка, однако они не подразумевают предсказания магнитуды предстоящего сейсмического события. Это делает применение любой из них крайне затруднительным. Если в регионе в скором времени ожидается землетрясение магнитудой 2-3, скорее всего, не будет предпринято никаких мер, так как такое землетрясение либо будет неощутимо для людей вовсе, либо проявится в раскачивании некоторых подвешенных предметов, таких как картины, и будет заметно для некоторых людей, находящихся преимущественно на верхних этажах высотных зданий. Повреждений зданий подобное землетрясение не вызовет, так как события такой магнитуды обычно вызывают колебания интенсивностью 2-4 балла по шкале Меркалли. Если же в регионе предсказано наступление землетрясения магнитудой 7-8, будет начата немедленная эвакуация людей, ведь за событием такой мощности последуют повреждения и разрушения некоторых зданий, появление трещин в земной поверхности, а также значительных оползней и обвалов, повреждения дорог, разрывы трубопроводов, камнепад. Также возможны искривление железнодорожных и трамвайных рельсов и разрушение мостов и платин, некоторые водоемы могут выйти из берегов. Сила такого толчка будет оценена в 9-11 баллов по шкале Меркалли, а само землетрясение приведет к огромным человеческим жертвам и экономическому ущербу. Однако, сильные землетрясения, имеющие столь ужасающие последствия, случаются относительно редко, гораздо чаще происходят толчки малой магнитуды, не представляющие угрозы для людей и инфраструктуры. Этот феномен объясняется законом Гутенберга-Рихтера, который упоминается чуть позже. Таким образом, знание предсказанного времени до следующего события, не подкрепленное сведениями о его магнитуде, не принесет большой пользы.

Прогнозирование времени возникновения, магнитуды или интенсивности землетрясений было не единственным предметом научных исследований в последние годы. В [23] исследуется эффективность линейной регрессии (Linear least square regression, LSR) и четырех нелинейных методов машинного обучения: случайного леса (Random forest, RF), градиентного бустинга (Gradient Boosting, GB), метода опорных векторов (Support Vector Machines, SVM) и метода k ближайших соседей (K-Nearest Neighbors, kNN), – при прогнозировании структурного дрейфа по первым секундам P-волн, записанным системами раннего предупреждения землетрясений. Коэффициент дрейфа рассчитывается как разница между пиковым смещением в верхней части здания и пиковым смещением в нижней части здания, деленная на расстояние между ними. Модели обучаются и тестируются на конкретных параметрах P-волн, которые обычно рассчитываются в системах предупреждения о землетрясениях на основе необработанных сигналов, – это пиковое смещение, интеграл квадрата скорости и интеграл квадрата смещения. В рамках данной работы каждый из параметров рассчитывается на основе сигнала длительности 1, 2 и 3 секунды, таким образом, получается 9 предикторов. Модели обучаются с использованием различных подмножеств исходного набора данных о строениях в Японии, начиная с данных, относящихся к определенному виду зданий, характеризующихся одинаковыми условиями местности и особенностями строения (здания, построенные с использованием железобетона в районе Шиодоме в Токио, Япония), затем включая больше зданий, относящихся к тем же типологии и региону (все железобетонные постройки, расположенные по всей Японии и, следовательно, характеризующиеся разными условиями местности), а после включая конструкции из того же региона, но других строительных типологий (весь здания в Японии). Наконец, модели, обученные на наборе данных, собранных в Японии, применяются к записям о зданиях в США. По мере того, как сложность данных увеличивается из-за различных условий местности и различных типов конструкции, наблюдается общее ухудшение метрик R2 и RMSE, где RMSE – корень из среднеквадратичной ошибки, или Root Mean Squared Error. Тем не менее, следует отметить, что во всех случаях, во-первых, прогнозы моделей улучшаются с увеличением длины временного окна, и во-вторых, метод SVM показывает лучшие результаты по сравнению с другими моделями. Коэффициент детерминации (R2), рассчитанный по прогнозам SVM, обученного на трех параметрах, полученных при помощи 3-секундных сигналов, равняется 0,64, 0,58 и 0,47 для трех подмножеств японских зданий, соответственно.

Кроме того, еще один феномен, который можно обнаружить при исследовании применимости первых трех подвыборок – это недооценка дрейфа при наибольших магнитудах сейсмических событий, которая, вероятно, связана как с отсутствием достаточного количества данных в диапазоне больших магнитуд, так и с максимальной длиной окна, равной 3 секундам, что является слишком коротким интервалом для записи Р-волн и оценки их характеристик.

Применение моделей, обученных на наборе данных Японии, к набору данных США приводит к общему смещению прогнозов в сторону занижения оценок дрейфа, что может быть связано с тектоническими и геологическими различиями между странами, а также с разными строительными нормами и правилами проектирования в Японии и Соединенных Штатах. Однако, применяя некоторые поправки к оценкам моделей, смещение можно устранить, и в этом случае линейная регрессия, использующая всего один из девяти признаков, работает лучше, чем нелинейные методы машинного обучения. R2 равняется в этом случае 0,30.

Данное исследование показывает, что надежные предсказания могут быть получены путем группировки данных, относящихся к схожим зданиям, и увеличения временного окна регистрируемых сигналов.

В статье, которую мы обсудим далее, идея состоит в том, чтобы использовать при работе с моделями не сигналы сейсмографов или полученные с их помощью статистические параметры, а рассчитанные уже на базе каталога всех прошлых землетрясений, произошедших в регионе, смысловые признаки, описывающие сейсмический потенциал территории, для которой мы хотим получить предсказание, на ближайшее время.

Авторы [24] исследуют три различные нейронные сети, а именно: нейронную сеть обратного распространения Левенберга-Марквардта (Levenberg-Marquardt backpropagation, LMBP), рекуррентную нейронную сеть (RNN) и нейронную сеть радиальной базисной функции (Radial basis function, RBF), – для прогнозирования магнитуды крупнейшего сейсмического события в следующем месяце на основе восьми индикаторов сейсмичности в Южной Калифорнии и заливе Сан-Франциско.

Входные параметры рассчитываются для определенного региона с использованием информации о заданном количестве последних землетрясений пороговой магнитуды и выше, произошедших до начала месяца, для которого осуществляется прогноз. К вычисляемым сейсмическим индикаторам относятся время, прошедшее за последними событиями пороговой магнитуды и выше, средняя магнитуда по шкале Рихтера последних событий, величина квадратного корня из суммарной сейсмической энергии, выделившейся за время, прошедшее за последними событиями, наклон регрессионной кривой зависимости логарифма частоты землетрясений от магнитуды (параметр основан на законе Гутенберга-Рихтера), а также суммирование среднеквадратичного отклонения от линии регрессии, разница между наибольшей наблюдаемой магнитудой и наибольшей ожидаемой магнитудой согласно соотношению Гутенберга-Рихтера, среднее время, наблюдаемое между характерными событиями – сильными землетрясениями, имеющими близкие значения магнитуд, например, от 7 до 7,5, – среди последних толчков, и коэффициент вариации среднего времени между характерными событиями.

На основе восьми описанных показателей каждая из моделей выдает один из двух возможных прогнозов – возникновение (1) или отсутствие (0) землетрясения заранее определенной пороговой магнитуды или выше в течение следующего месяца. Работа модели повторяется, и на каждой новой итерации пороговая магнитуда увеличивается на 0,5, до тех пор пока выход модели не станет равным 0. Наибольшей магнитудой землетрясения в течение следующего месяца является конкретная пороговая магнитуда, при увеличении которой выход модели меняется на 0.

Как показывает исследование, RNN обеспечивает наилучшую точность прогнозирования, что проявляется в значениях вероятности обнаружения (Probability of detection, POD), равной 1,00, коэффициента ложных тревог (False alarm rate, FAR), равного 0,00, и оценки R (R score, R), равной 1,00, для событий магнитудой 7,0 против 0,00, 0,00 и 0,00, полученных моделью LMBP, и 0,50, 0,50 и 0,00, полученных моделью RBF, соответственно. Важным замечанием здесь является то, что тестовый набор данных, используемый для оценки моделей, содержит довольно мало событий больших магнитуд.

Важно отметить, что для всех моделей значимость каждого входного параметра для точности предсказаний исследуется путем удаления этого параметра из входного вектора, что иногда приводит к сильному ухудшению метрик. Это позволяет сделать вывод о том, что определенные индикаторы действительно несут в себе смысл и описывают паттерны, лежащие в основе сейсмической активности Южной Калифорнии и залива Сан-Франциско.

Наконец, результаты могут быть улучшены путем вычисления индикаторов сейсмичности для временных периодов короче одного месяца и для регионов меньшей площади.

Отличный от описанных выше подход используется в [25], где землетрясения предсказываются при помощи анализа спутниковых данных. В рамках исследования рассматривается метод машинного обучения, который называется методом минимальной площади тревоги. Область, для которой требуется сделать прогноз, в нашем случае, это сразу два региона, Япония и Калифорния, делится на отдельные участки размером 0,1◦ × 0,075◦ (в западно-восточном и северо-южной направлениях, соответственно), для каждого из которых затем вычисляются пространственно-временные признаки на основе временных рядов смещений земной поверхности, полученных по данным с определенного количества ближайших к участку приемных станций.

При помощи исходных данных, представляющих ежедневные записи о координатах станций, вычисляются скорости движения, или деформации, земной поверхности в двух горизонтальных направлениях. Уже на их основе для каждого участка рассчитываются три инварианта скоростей деформации: дивергенция (расхождение) скоростей деформации – параметр описывает сужение или расширение небольшой горизонтальной области, – ротация скоростей деформации – признак определяет направление и интенсивность закручивания участка вокруг вертикальной оси, – а также сдвиг скоростей деформации. Следующие три признака представляют изменения инвариантов скоростей деформации во времени. Затем авторы рассматривают пространственные корреляции изменений инвариантов скоростей деформации в скользящем окне размером 75 × 75 км2, что эквивалентно 70–80 участкам. На заключительном этапе расчета признаков от каждой пары пространственных корреляций берется минимум, который служит последним предиктором.

В конечном итоге участки с аномально высокими значениями компонентов вектора признаков, согласно методу минимальной площади тревоги, помечаются на карте региона как тревожные зоны с высокой вероятностью стихийного бедствия.

Помимо геофизических данных, в работе также рассматриваются сейсмологические предикторы, сконструированные на основе информации о эпицентрах и магнитудах прошлых землетрясений в регионе. Использование в процессе прогнозирования как признаков каждого типа по отдельности, так и их объединения, позволяет увидеть, что геофизические параметры особенно хорошо отражают сейсмическую ситуацию в регионе и повышают точность оценок опасности его отдельных участков. Если для Японии лучшие прогнозы осуществляются при совместном использовании двух типов данных, то для Калифорнии наиболее высокая вероятность предсказания землетрясений достигается при применении математического метода лишь только к геофизическим данным. Добавление же к ним сейсмических параметров не оказывает положительного влияния на значение метрики.

Вероятность предсказания землетрясений в Японии равна 0,57, 0,71 и 0,71 при использовании отдельно геофизических, отдельно сейсмологических и обоих типов предсказательных переменных одновременно, соответственно. В случае Калифорнии значения метрики равняются 0,83, 0,75 и 0,92, соответственно.

Таким образом, полученные значения метрики подтверждают эффективность использования геофизических данных при прогнозировании землетрясений.

Следующая работа [26] также использует данные спутниковой системы, но рассматривает уже совсем другой подход, впервые предложенный в [27].

Если в предыдущей работе ученые пытаются определить деформации, связанные с землетрясениями, исследуя необычные вариации скорости движение земной поверхности, здесь рассматриваются ориентации смещений плит.

Для извлечения из спутниковых данных сигналов, вызванных землетрясениями, используется преобразование Гильберта-Хуанга (Hilbert‐Huang transform, HHT) для фильтрации данных от долгосрочных движений – некоторые смещения вызываются надвигающимися землетрясениями, но не стоит забывать, что литосферные плиты сами по себе находятся в постоянном движении относительно друг друга, и эти подвижки не должны использоваться при построении прогнозов, – краткосрочного шума и сезонных (годовых и полугодовых) вариаций. Данное преобразование состоит из нескольких этапов. Во-первых, исходные данные о смещениях обрабатываются с помощью эмпирической модовой декомпозиции (Empirical mode decomposition, EMD). В результате мы получаем несколько собственных модовых функций, а также остаток, сложив которые мы получим исходный сигнал. Во-вторых, мы переводим полученные собственные модовые функции в частотную область с помощью преобразования Гилберта, где рассчитываем мгновенную частоту и ее амплитуду в каждой точке функции. Таким образом, каждая собственная модовая функция представляется некоторой совокупностью частот и соответствующих амплитуд. Далее выбирается частотный диапазон для отсеивания нежелательных влияний (т.е. шума, долговременных движений и т.д.). После применения фильтрации смещения, соответствующие оставшимся частотам, накладываются друг на друга для создания нового временного ряда смещений.

Наконец, отфильтрованные смещения в направлениях север-юг и запад-восток, рассчитанные для каждой станции, используются для оценки горизонтальных смещений поверхности, ассоциирующихся с землетрясениями, по их направлениям.

Наблюдение показало, что ориентации горизонтальных азимутов должны быть обычно случайными до тех пор, пока в коре не накапливается значительное напряжение, связанное с землетрясением. За несколько дней – недель до возникновения землетрясения ориентации азимутов выравниваются в сходном направлении вокруг эпицентров землетрясений. Затем ориентации вновь становятся неупорядоченными, после чего наступает землетрясение. После азимуты выравниваются в противоположном направлении, а позже вновь становятся неупорядоченными.

Таким образом, мы приходим к выводу, что существует связь между возникновением землетрясений и выравниванием азимутом. При помощи данного метода удается предсказать 26 из 32 событий тестовой выборки, причем 6 не предсказанных толчков появляются в воде, что, вероятно, влияет на успешность работы алгоритма. Области с наиболее параллельными ориентациями азимутов появляются вокруг эпицентров за несколько дней – недель до возникновения 26 землетрясений. Причем время, прошедшее с момента выравнивания азимутов до возникновения землетрясения, примерно пропорционально магнитуде события.

Данный подход так же используется в ряде других более поздних работ [28-30], где так же показывает свою эффективность.

Тем не менее, декларируемая во многих работах успешность более сложных моделей, предложенных за последние несколько лет, не является подтверждением того, что они действительно обладают лучшей предсказательной способностью по сравнению с более простыми подходами [31].

### **2.2 Проблемы более сложных моделей**

На основе анализа 77 разных подходов, предложенных за период с 1994 по 2019 гг., к решению задачи предсказания землетрясений можно заключить, что у применения более сложных моделей есть несколько проблем [31].

Первая проблема связана с доступностью лишь небольших наборов обучающих данных по сравнению с размером предлагаемых моделей, что в свою очередь приводит к переобучению. Кроме того, редкость крупных событий в сочетании с недостаточной выборкой может привести к любому возможному результату. А значит нельзя быть уверенными в сообщаемых улучшенных значениях метрик. Решение данной проблемы заключается в сборе и последующем использовании большего количества данных.

Отсутствие надлежащих базовых моделей (baseline) для сравнения с новыми подходами к прогнозированию землетрясений является второй проблемой современных исследований. Наиболее распространенной референсной моделью является нулевая гипотеза Пуассона, несмотря на то что, как известно, сейсмичность накапливается как в пространстве, так и во времени, а значит, такая модель не подходит для описания сейсмических процессов и не может быть применена. Также нередко предлагаемые методы сравниваются с базовым классификаторами, такими как метод опорных векторов, ансамбль решающих деревьев (случайный лес и бустинг), наивный байесовский классификатор и др., и в большинстве случаев базовые модели не так хорошо настраивают, в отличие от новых методов, тщательно подгоняемых под данные. Как было замечено в [31], при правильной настройке очень часто базовые классификаторы обеспечивают лучшие результаты, чем более сложные архитектуры. Это, вероятно, способствует тому, что в последнее время ансамбли деревьев решений вызывают все больший интерес специалистов, стремящихся научиться предсказывать землетрясения. В [31] предлагается решение проблемы использования неподходящих базовых моделей – использование в качестве базовой модель, основанную на законе Гутенберга-Рихтера.

Еще одним недостатков сложных моделей является их не интерпретируемость по сравнению с другими классификаторами. И если мы можем самостоятельно извлечь геофизические параметры из исходных данных, основываясь на теоретических знаниях в области сейсмологии и геодезии, чтобы подать их на вход нейронной сети, контролировать процесс обучения мы уже никак не сможем – заглянуть в «черный ящик» мы на данный момент не в состоянии.

В совокупности все приведенные исследования показывают, что большинство усилий на данный момент направлено на применение методов машинного и глубокого обучения для прогнозирования магнитуды события во временном или пространственно-временном окне, реже времени возникновения главного толчка. Чаще всего используется один тип данных.

В настоящей работе представлен и протестирован подход, основанный на применении случайного леса. Прогнозирование магнитуды следующего землетрясения осуществляется при помощи использования сразу сейсмологических и геофизических данных региона.

## **3 Прогнозирование последствий землетрясений в условиях глубокой неопределенности**

Получения прогноза магнитуды приближающегося землетрясения недостаточно. Для оценки потребностей в ресурсах, необходимых для устранения следов предстоящего сейсмического события и восстановления инфраструктуры, необходимо просчитать, какие последствия событие спрогнозированной силы может иметь в конкретном регионе.

Существуют разные инструменты для быстрого прогнозирования последствий стихийных бедствий. Мы рассмотрим описанный в [32] подход, основанный на сценарном анализе, который разработан для оценки инфраструктурного ущерба, числа погибших и раненных при землетрясении, а также потребности во временных убежищах на уровне района и его отдельных частей при различной силе и времени наступления землетрясения.

В качестве объектов для оценки угрозы в предложенной модели рассматриваются так называемые кластеры. Каждый кластер представляет собой некоторое количество «зданий одного типа, используемых для одной и той же цели» (например, все жилые каменные дома, имеющие от одного до трех этажей и расположенные в конкретной части района) или «одно здание, которое одновременно используется многими людьми в данный момент времени» (например, крупные торговые центры или школы).

В модели используются три основных параметра, а именно типология строительства, количество этажей и год постройки, для разработки сценариев и расчета последствий землетрясения ожидаемой силы. Различные сценарии могут иметь место в зависимости от интенсивности возможного землетрясения.

В работе модели можно выделить три основных этапа. Во-первых, для кластера с определенной типологией застройки, в зависимости от интенсивности землетрясения, оценивается степень повреждения зданий. Во-вторых, используя оценки повреждения зданий, полученные для кластеров с известным распределением людей в них в течение дневного и ночного периодов, можно установить число погибших и раненых. В-третьих, основываясь на полученных цифрах, можно определить количество людей в каждой отдельной части района, которым потребуется временное убежище.

Таким образом, использование предложенного метода позволит получить ответы на наиболее важные вопросы по управлению землетрясениями, в том числе оценить потребности в ресурсах.

## **4 Построение модели для прогнозирования наступления землетрясения**

Своевременное прогнозирование землетрясений является важным предметом изучения в мировой научной литературе. Многочисленные исследования по построению сейсмологических прогнозов изучают роль предсказательных переменных в получении надежных результатов. Несмотря на то что можно заметить большое разнообразие в различных типах предикторов, используемых в исследованиях последних лет, вопрос об их сочетании исследователями мало обсуждается. В данном исследовании используются как геофизические, так и сейсмологические параметры, предложенные ранее в [25, 33]. Оба исследования показали достаточно хорошие результаты для региона Калифорнии, и улучшение их результатов за счет использования метода, ранее не применявшегося ни к спутниковым данным в отдельности, ни к их объединению с сейсмическими данными, представляет основной интерес для данной работы.

В настоящем разделе предлагается применить метод случайного леса, актуальность и возрастающая популярность которого были отмечены ранее, к некоторым геодинамическим и сейсмическим показателям, оказавшимся лучшими предикторами для Калифорнии в [25, 33]. Мы рассмотрим, какие именно признаки наилучшим образом описывают сейсмическую перспективу Калифорнии, поговорим, как использование двух типов данных в отдельности и вместе влияет на результаты работы модели, а также затронем способы ее улучшения.

Повышение точности прогнозов мощности будущего землетрясения, которое мы хотим достичь, позволит лучше предсказывать последствия землетрясения, а следовательно, и потребность в ресурсах.

### **4.1 Данные**

В рамках исследования анализируемая территория представляется в виде сетки, для каждого поля которой вычисляются пространственно-временные признаки, аномальные значения которых могут быть предвестниками сильных землетрясений. Размер каждого такого участка равен 0,1◦ × 0,075◦ , где первая компонента соответствует шагу по долготе, а вторая – по широте. Область, выбранная для проведения настоящего исследования, расположена в границах 119,5–117,4° западной долготы (з.д.) и 38–39° северной широты (с.ш.) в Калифорнии. В работе рассматривается период с 7 августа 2009 по 26 января 2023 года.

Временные ряды значений признаков, рассчитанных для отдельных участков в рамках рассматриваемой территории, а также каталог землетрясений, произошедших за указанный период времени, взяты с ГИС-платформы GeoTime, предложенной в [33]. Ниже представлена краткая характеристика используемых параметров, а также принцип их расчета.

4.1.1 Изучим отдельно геодинамические и сейсмологические признаки, используемые в работе, а также первичные данные, которые берутся за их основу, и операции, которые к ним применяются.

4.1.1.1 Рассмотрим геодинамические параметры, показавшие себя особенно хорошо при прогнозировании землетрясений в Калифорнии ранее [25, 33].

Исходные данные представляют собой временные ряды суточных координат x(t) и y(t) приемных станций GPS, расположенных в анализируемой области, в западно-восточном и северо-южном направлениях за выбранный временной период. По двум координатам наземной станции, зарегистрированным с временным интервалом Т0 дней, определяются суточные горизонтальные смещения земной поверхности gx(t) и gy(t) по следующей формуле:

где gx(t) – суточная скорость смещения земной поверхности в направлении З-В в месте (точке) расположения приемной станции,

x(t) – координата долготы приемной наземной станции GPS в момент времени t,

x(t – Т0) – координата долготы приемной наземной станции GPS в момент времени t – Т0.

Скорость деформации земной поверхности в направлении С-Ю gy(t) рассчитывается аналогично.

Во временных рядах координат присутствуют пропуски, из-за чего пропущенные значения появляются так же в значениях скоростей смещения. Чтобы ограничить их количество, значения координат линейно интерполируются в промежутках, меньших или равных T0. При промежутках более Т0 дней расчет скоростей заканчивается на последнем значении координаты станции перед началом разрыва и продолжается, начиная с первого значения координаты станции после разрыва.

Для расчета суточных скоростей используется интервал T0 = 30 дней.

Далее на основе полученных расчетов вычисляются, так же ежедневно, скорости смещения Vx и Vy в направлениях З-В и С-Ю для каждого участка сетки размером 0,1◦ × 0,075◦, при этом рассматриваются только те временные ряды, в которых отсутствуют пропуски:

где Vxn(t) – значение скорости деформации в направлении З-В в участке сетки n в момент времени t,

K – максимальное число ближайших к участку n станций в окружности радиуса Rmax, значения которых использовались для вычисления скорости деформации,

gxk(t) – скорость деформации в направлении З-В для станции k, k = 1, …, K, в момент времени t,

rk ≤ Rmax – расстояние от k-й станции до участка сетки n,

p – степень, определяющая зависимость веса станции от ее расстояния до участка сетки.

Для расчета используемых в настоящем исследовании признаков используются следующие значения параметров: K = 5, Rmax = 50 км, p = 1. Если rk = 0, то Vxn(t) = gxk(t). Соответствующие расчеты скорости деформации в направлении С-Ю проводятся аналогичным образом.

Первые три геофизических параметра, рассмотренные в [25, 33], являются инвариантами скоростей деформации.

Первый признак, F1, представляет дивергенцию (расходимость) скоростей деформации и вычисляется по следующей формуле:

Данный параметр характеризует отношение скоростей изменения значений скоростей деформации в направлении З-В и С-Ю в участке сетки n. Максимальное и минимальное значения параметра соответствуют местам, где происходит относительное сужение или расширение размера небольшой горизонтальной области.

Второй параметр, F2, описывает ротор скоростей деформации:

Значения признака определяют направление и интенсивность закручивания участка сетки вокруг вертикальной оси.

Еще один признак, F3, определяет сдвиг скоростей деформации и вычисляется следующим образом:

Следующие три признака показывают изменения инвариантов скоростей деформации F1, F2 и F3 во времени. Каждый из них равен отношению разности средних значений определенного инварианта в двух последовательных временных интервалах к стандартному отклонению этой разности и вычисляется раз в 30 дней.

F4 – признак временных вариаций дивергенции скоростей деформации:

где рассчитывается по значениям параметра F1 на интервале (t − T2, t), – по значениям параметра F1 на интервале (t − T2 − T1, t − T2), T1 = T2 = 360 дней.

Соответствующие признаки для F2 и F3 вычисляются по следующим формулам:

Именно параметры F4, F5, F6 оказались ключевыми в успешности предсказаний землетрясений ранее [25, 33], и именно они используются в настоящей работе для исследования применимости иного подхода к решению данной задачи.

На рисунке 16 приведены временные ряды значений параметров F4, F5, F6, рассчитанных для участка в границах 117,9–117,8° з.д. и 38,103–38,178° с.ш. в Калифорнии. Красной пунктирной линией отмечены соответствующие значения признаков, рассчитанные 11 мая 2020 года и используемые для составления сейсмологического прогноза на следующие 30 дней. Спустя 4 дня, 15 мая 2020 года, в Калифорнии произошло землетрясение. Координаты его эпицентра – 117,849° з.д. и 38,169° с.ш.

Изображение выглядит как текст, диаграмма, линия, График

Автоматически созданное описание

Рисунок 16 – Временные ряды значений признаков F4, F5, F6, рассчитанных для области с границами 117,9–117,8° з.д. и 38,103–38,178° с.ш. в Калифорнии

4.1.1.2 С целью сравнения с прогнозами, полученными посредством использования признаков, рассчитанных на основе данных спутниковой системы навигации, так же исследуется применение предложенных в [25] сейсмических параметров, вычисленных по каталогу исторических сейсмических событий, состоящему из событий с магнитудой m ≥ 2,0 и глубиной гипоцентра H ≤ 160 км, произошедших с 1 января 1995 года по 26 января 2023 года.

Первый сейсмологический признак, S1, – плотность эпицентров землетрясений (оценивается с использованием трех измерений, где третье измерение – время). Признак вычисляется методом локальной ядерной регрессии. Функция ядра для q-го землетрясения имеет вид:

где rq < R ϵ – расстояние между эпицентром q-ого землетрясения и участком сетки,

tq < Tϵ – временной интервал между эпицентром q-ого землетрясения и участком сетки.

Для расчета признака параметры берутся равными следующим значениям: ϵ = 2, R = 50 км, T = 100 дней.

Следующий признак, который используется в настоящей работе, имеет обозначение S8 и равняется квантилю плотности эпицентров землетрясений, который соответствует значению плотности эпицентров землетрясений на момент проведения расчетов.

На основе последних двух признаков вычисляется еще один параметр, равный отношению значения плотности эпицентров землетрясений к значению соответствующего ему квантиля плотности эпицентров землетрясений, рассчитанных для того же участка сетки:

Наконец, последний сейсмологический параметр, который мы рассмотрим, – S11 – основывается на t-статистике Стьюдента средней магнитуды землетрясений и определяется для каждого участка сетки как отношение разности средних значений текущего интервала T2 и фонового интервала T1 к стандартному отклонению этой разности. Используются временные промежутки T1 = 2310 и T2 = 120 дней.

Для предсказания землетрясений в Калифорнии мы будем рассматривать параметры S9 и S11, так как они оказались лучшими из предложенных в [25] сейсмологических предикторов для данного региона.

4.1.2 Таким образом, в настоящем исследовании используются два набора данных, полученных после обработки исходных данных с платформы GeoTime.

Первый датасет содержит значения признаков, рассчитанные для каждого участка рассматриваемой территории в конкретный момент времени.

Вначале территория, ограниченная координатами 119,5–117,4° з.д. и 37,952–39,082° с.ш., делится на отдельные участки размером 0,1◦ × 0,075◦. Общее число шагов составляет 21 по долготе и 15 по широте. Таким образом, получается 315 полей.

Далее для каждого из полей осуществляется расчет признаков F4, F5, F6, S9, S11 каждые 30 дней, начиная с 7 августа 2009 года и заканчивая 26 января 2023 года. Итого, мы имеем 165 временных шагов.

Собрав все значения, посчитанные за указанный промежуток времени для 5 параметров, мы получим двумерный массив размером 51975 × 10. В таблице 1 приведены первые 5 строк из собранного набора данных:

Таблица 1 – Строки из набора данных, содержащего значения признаков

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Год | Месяц | День | Долгота | Широта | F4 | F5 | F6 | S9 | S11 |
| 2009 | 8 | 7 | -119,5 | 39,007 | 0,503 | -0,219 | 1,245 | 0,446 | 2,142 |
| 2009 | 8 | 7 | -119,4 | 39,007 | 2,261 | 3,551 | 8,822 | 0,523 | 2,231 |
| 2009 | 8 | 7 | -119,3 | 39,007 | -8,884 | 8,894 | -8,575 | 0,524 | 2,312 |
| 2009 | 8 | 7 | -119,2 | 39,007 | -20,893 | -0,0514 | -15,676 | 0,654 | 2,377 |
| 2009 | 8 | 7 | -118,5 | 39,007 | 14,297 | -10,377 | -15,117 | 0,249 | 2,5127 |

Указанные в таблице 1 значения долготы и широты соответствуют левому нижнему углу конкретного участка. Таким образом, первая строка описывает поле с границами 119,5–119,4° з.д. и 39,007–39,082° с.ш.

После удаления записей (строк), содержащих пропуски (пустые значения признаков), в датасете останется 43413 строк.

Скрипт кода, с помощью которого были обработаны исходные данные и получен датасет со значениями признаков, приведен в приложении А.

Второй набор данных представляет каталог исторических землетрясений, имевших место в рассматриваемой области в охваченный период времени. Из начального архива сейсмических ударов были удалены записи о событиях, произошедших в тот же год и месяц рядом с более сильными ударами – их разделяли не более 0,1◦ по долготе и не более 0,075◦ по широте, – которые, вероятно, являлись главными толчками, в то время как события, информация о которых была удалена, являлись их автершоками.

На исследуемой территории за период с 7 августа 2009 года по 26 января 2023 года произошло 8 землетрясений. Информация о них приведена в таблице 2:

Таблица 2 – Землетрясения, произошедшие на территории с координатами 119,5-117,4° з.д. и 38-39° с.ш. с 7 августа 2009 года по 26 января 2023 года

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Год | Месяц | День | Широта | Долгота | Глубина | Магнитуда |
| 2013 | 2 | 13 | 38,0222 | -118,055 | 7,5 | 5,1 |
| 2016 | 12 | 28 | 38,3904 | -118,897 | 12,2 | 5,6 |
| 2020 | 4 | 11 | 38,0525 | -118,733 | 8,45 | 5,24 |
| 2020 | 5 | 15 | 38,1689 | -117,85 | 2,7 | 6,5 |
| 2020 | 6 | 30 | 38,1539 | -117,958 | 8,3 | 5 |
| 2020 | 11 | 13 | 38,1693 | -117,853 | 4,8 | 5,3 |
| 2020 | 12 | 2 | 38,1637 | -118,084 | 5 | 5,1 |
| 2021 | 7 | 9 | 38,5075 | -119,5 | 7,45 | 6 |

Скрипт соответствующего кода расположен в приложении Б.

Наконец, два набора данных сопоставляются, и каждой строке в первом датасете ставится в соответствие 1, если по расположенным в ней значениям параметров необходимо предсказать землетрясение в следующие 30 дней, иначе – 0. Ввиду малого числа объектов во втором датасете предсказание более сложных целевых переменных, таких как, например, магнитуда, затруднительно. Тем не менее, большая часть землетрясений в Калифорнии имеет магнитуду 5-6,5, более сильные толчки происходят крайне редко, а значит, получив предсказание 1, можно с большой долей вероятности ожидать событие магнитудой не более 6-6,5. Таким образом, такого формата предсказаний в данном регионе может быть вполне достаточно.

Скрипт кода, с помощью которого осуществляется соединение двух наборов данных, можно увидеть в приложении В.

После все значения признаков нормализуются по методу минимакса, целью применения которой является преобразование исходного набора в диапазон от 0,0 до 1,0. Для такого типа нормализации используется следующая формула:

где Xnorm – нормализованное значение признака,

X – исходное значение признака,

Xmin – минимальное значение временного ряда значений признака (представленного в соответствующем столбце двумерного массива данных),

Xmax – максимальное значение временного ряда значений признака.

Часть кода, посвященная процессу нормализации данных, приведена в приложении Г.

Последним шагом подготовки датасета является уменьшение дисбаланса присутствующих в нем классов. Для этого генерируется некоторое число новых объектов необходимого класса, значения признаков для которых случайным образом выбираются из значений реальных объектов того же класса из набора данных. В рамках данного исследования было сгенерировано 20 новых элементов.

Соответствующий скрипт кода размещен в приложении Д .

### **4.2 Обучение и тестирование модели**

Метод случайного леса был выбран из-за его относительной простоты по сравнению с нейронными сетями и высокой точности прогнозов, получаемых при его использовании.

Из-за наблюдаемого дисбаланса классов – даже после искусственного увеличения числа объектов класса 1, описанного ранее, дисбаланс классов остается существенным – обучение на всем наборе данных не является правильным. По этой причине каждое дерево обучается на случайном подмножестве объектов, в котором присутствует одинаковое количество элементов обоих классов. Тестирование проводится так же с использование сбалансированной выборки, но уже меньшего размера.

Обучающие и тестовые выборки составляются таким образом, чтобы 70% всех объектов класса 1 попадали в набор для обучения, а остальные 30% – шли на тестирование модели. Объектов же класса 0 в обоих случаях берется столько же, сколько элементов класса 1 попало в выборку. Элементы класса 1, используемые в каждой из выборок, фиксируются в начале работы и не меняются. В рамках обучения объекты класса 0 выбираются случайным образом для каждого дерева модели. Перед началом тестирования конкретное число случайных представителей класса 0, равное количеству заранее определенных объектов класса 1 в тестовой выборке, фиксируется и далее используется на протяжении тестирования всех деревьев случайного леса.

В рамках предлагаемого подхода множество альтернатив, оцениваемых по части параметров – только геодинамическим, только сейсмологическим, по обоим типам параметров, – подается на вход дереву решений. На выходе ожидается получить предсказание классов соответствующих участков: 1, если в них ожидается сильное землетрясение в следующие 30 дней, иначе 0. Акцент здесь делается на построении определенного количества таких деревьев, объединение выходов которых будет итоговым предсказанием – прогнозом для участка является класс, который был предсказан большинством решающих деревьев. Всего в работе рассматривается 100 деревьев.

Части кода, в которых происходят обучение и тестирование модели, находятся в приложениях Е и Ж, соответственно.

### **4.3 Метрики качества**

Для оценивания деревьев решений, обучаемых на случайно выбранных для них выборках объектов, используются упомянутые ранее метрики: POD, FAR, R. Чтобы понять, как именно и почему они рассчитываются, рассмотрим сначала представленную таблицей 3 матрицу путаницы (Confusion matrix), из которой можно узнать, сколько объектов классов 0 и 1 модель предсказала верно, а сколько отнесла к неправильным классам – для участков, в которых не наблюдалось землетрясения, предсказала скорое сейсмическое событие, а для участков, в которых должен был наступить удар, спрогнозировала его отсутствие:

Таблица 3 – Матрица путаницы

|  |  |  |
| --- | --- | --- |
|  | Observed | Not Observed |
| Predicted | True Positive (TP) | False Positive (FP) |
| Not Predicted | False Negative (FN) | True Negative (TN) |

Зная элементы матрицы путаницы, рассмотрим формулы, по которым вычисляются метрики:

В то время как POD указывает на долю объектов класса 1, которые были верно предсказаны, а FAR вычисляет частоту ложных прогнозов землетрясений, R оценивает общее умение модели прогнозировать землетрясения, балансируя между значениями двух первых показателей качества работы модели. Эта оценка является более предпочтительной, по сравнению с POD и FAR, поскольку она включает в себя равное представление как правильных, так и неправильных прогнозов.

Чтобы получить итоговое значение определенной метрики для случайного леса, состоящего из нескольких решающих деревьев, значения нужной метрики для всех входящих в него деревьев усредняются.

### **4.4 Результаты и обсуждение**

С целью изучения применимости данного метода и эффективности его использованию вместе с разными наборами признаков обучение модели проводится на трех частях параметров – рассчитанных на основе спутниковых данных, рассчитанных на основе информации о прошлых землетрясениях в рассматриваемой области, обоих типах параметров в одно время.

При тестировании модели, обученной на наборе данных, содержащем исключительно геодинамические признаки, мы получаем следующие значения метрик: POD = 0,751, FAR = 0,184, R = 0,567. С учетом малого размера обучающей выборки, генерируемой для обучения каждого дерева, в среднем решающие деревья, входящие в модель, научились хорошо справляться с задачей, предсказывая 75% землетрясений. Стоит отметить, что 18% объектов класса 0 получают неверные предсказания, что является серьезным ограничением в использовании полученной модели.

Модель, обученная только на сейсмологических признаках, получает POD = 0,462, FAR = 0,04, R = 0,422, что говорит о меньшей точности модели при прогнозировании объектов класса 1, и в то же время, значительно меньшей доле ложных тревог. Тем не менее, общая оценка навыка прогнозирования данной модели ниже.

Значения метрик, полученных наилучшей из моделей, обученных сразу на геофизических и сейсмологических признаках, равняются POD = 0,762, FAR = 0,135, R = 0,627. Точность предсказания класса 1 возросла, по сравнению с первой моделью, в то время как процент ложных тревог понизился, что говорит о более высокой надежности предсказаний, получаемой данной моделью.

Таким образом, в рамках данного исследования наилучшие результаты показывает метод, основанный на применении случайного леса к двум типам предикторов одновременно. Заметим, что при обучении решающих деревьев, входящих в состав двух последних моделей, наиболее значимым признаком оказывается S11. Для самой первой модели признаком, имеющим наибольшую важность, является F6.

Скрипт кода, в рамках которого происходит тестирование модели на разных наборах предикторов, расположен в приложении И.

Важно отметить, что несмотря на не самые хорошие результаты, продемонстрированные моделями, применение предложенного в подразделе 4.2 подхода имеет одно веское преимущество. В то время как зона тревоги, получаемая в [25, 33], охватывает сразу множество соседних участков, данный метод строит прогноз отдельно для каждого из них, а значит, полученные с его помощью предсказания позволят принимать срочные меры на относительно небольшой территории, что можно будет сделать быстро, не потратив при этом большое количество ресурсов.

### **4.5 Рекомендации по улучшению результатов**

Существуют способы, с помощью которых возможно повышение прогностической способности описанной модели.

Во-первых, использование набора данных большего размера, то есть анализирование не отдельной области в рамках региона Калифорнии, а всей его территории или даже ряда регионов, позволит лучше обучить модели. Следовательно, прогнозы модели, обученной на достаточном объеме данных, будут более статистически значимыми и надежными.

Во-вторых, ввиду наличия сильного дисбаланса в подобных данных, связанного с малочисленностью сильных сейсмических событий, рекомендуется также использование разнообразных способов синтеза и аугментации данных, что поспособствует повышению обобщающей способности модели.

Использование обеих методик позволит перейти от решения задачи двухклассовой классификации (наличие или отсутствие сильного землетрясения в ближайшее время) к предсказанию более сложных целевых переменных. Примером может стать прогнозирование уровня опасности, равного 0, если землетрясение магнитудой 5,0 и выше не ожидается, 1 – вероятно появление землетрясения магнитудой 5,0-6,0, 2 – вероятно появление землетрясения магнитудой 6,0-7,0, 3 – вероятно появление землетрясения магнитудой 7,0-8,0, и 4 – вероятно появление землетрясения магнитудой 8,0 и более. Прогнозирование значения подобного параметра позволит более точно определить сценарий происходящего, а значит, более точно распланировать действия и рассчитать потребности в ресурсах.

Наконец, использование улучшенного случайного леса, основанного на пороговой процедуре и принципе суперпозиции, который доказал свою полезность в прошлом при прогнозировании торнадо [34], может стать главным способом улучшения результатов.

### **4.6 Ограничения настоящего исследования**

В данном работе мы отказались от использования метода, предложенного в [34], ввиду ограничений в виде мощности компьютера. По этой же причине было принято решение в пользу использования ограниченного набора данных – рассматриваемая территория ограничена координатами, указанными ранее.

# **ЗАКЛЮЧЕНИЕ**

Основная цель настоящего исследования состояла в том, чтобы предложить подход, позволяющий спрогнозировать потребность в ресурсах при наступлении землетрясения. В рамках достижения поставленной цели был рассмотрен подход, основанный на применении случайного леса и определении эффекта использования двух совершенно разных типов данных, в которых могут быть зарегистрированы аномалии, указывающие на скорое появление землетрясения. Получение прогнозов сейсмических явлений с помощью предложенного метода позволит провести более точные оценки инфраструктурного ущерба, числа погибших и раненных при землетрясении, а также потребности во временных убежищах, воспользовавшись методикой, описанной в разделе 3.

Исходя из того, что подавляющая часть землетрясений в Калифорнии имеет магнитуду 5,0-6,0 и только в редких случаях достигает значений в районе 7,0, применение предложенного в настоящем исследовании метода, базирующегося на применении случайного леса при решении задачи двухклассовой классификации – предсказании наступления или отсутствие сильного землетрясения в ближайший период времени, – и последующее использование полученного прогноза при предсказании последствий сейсмического события и оценке потребности в ресурсах, необходимых для ликвидации воздействий спрогнозированного сейсмического удара и восстановления инфраструктурного капитала, для региона Калифорнии является хорошим решением и позволяет получить надежные расчеты.

Наконец, необходимо учитывать ряд важных ограничений. Во-первых, полученные результаты справедливы только для отдельного района Калифорнии и не предполагают успеха применения модели к данным, собранным на других территориях, характеризующихся иными геодезическими и тектоническими закономерностями. Во-вторых, при обучении и тестировании моделей использовались наборы данных небольшого размера, а значит, полученные результаты требуют дополнительной проверки.

В заключение необходимо упомянуть, что последующая доработка модели позволит в дальнейшем исследовать ее применимость для прогнозирования иных стихийных бедствий, таких как наводнения, ураганы, лесные пожары, оползни и прочее.

# **СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ**

1. Короновский Н.В. Общая геология – 3-е издание. – М.: КДУ, 2012. – 578 с.

2. Уолтхем Т. Катастрофы: неистовая Земля. – М.: Недра, 1982. – 223 с.

3. Lomnitz C., Wisner B. The Routledge handbook of hazards and disaster risk reduction. – Routledge, 2012. – 12 с.

4. Negarestani A., Setayeshi S., Ghannadi-Maragheh M., Akashe B. Layered neural networks based analysis of radon concentration and environmental parameters in earthquake prediction // Journal of environmental radioactivity. – 2002. – Vol. 62, № 3. – P. 225–233.

5. Külahcı F., İnceöz M., Doğru M., Aksoy E., Baykara O. Artificial neural network model for earthquake prediction with radon monitoring // Applied Radiation and Isotopes. – 2009. – Vol. 67, № 1. – P. 212–219.

6. Varotsos P., Lazaridou M. Latest aspects of earthquake prediction in Greece based on seismic electric signals // Tectonophysics. – 1991. – Vol. 188, № 3. – P. 321–347.

7. Hayakawa M., Hattori K., Ohta K. Monitoring of ULF (ultra-low-frequency) geomagnetic variations associated with earthquakes // Sensors. – 2007. – Vol. 7, № 7. – P. 1108–1122.

8. Hayakawa M., Hobara Y. Current status of seismo-electromagnetics for short-term earthquake prediction // Geomatics, Natural Hazards and Risk. – 2010. – Vol. 1, № 2. – P. 115–155.

9. Guo J., Li W., Yu H., Liu Z., Zhao C., Kong Q. Impending ionospheric anomaly preceding the Iquique Mw8. 2 earthquake in Chile on 2014 April 1 // Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society. – 2015. – Vol. 203, № 3. – P. 1461–1470.

10. Masci F., Thomas J.N., Villani F., Secan J.A., Rivera N. On the onset of ionospheric precursors 40 min before strong earthquakes // Journal of Geophysical Research: Space Physics. – 2015. – Vol. 120, № 2. – P. 1383–1393.

11. Carter B.A., Kellerman A.C., Kane T.A., Dyson P.L., Norman R., Zhang K. Ionospheric precursors to large earthquakes: A case study of the 2011 Japanese Tohoku Earthquake // Journal of Atmospheric and Solar-Terrestrial Physics. – 2013. – Vol. 102. – P. 290–297.

12. Wyss M., Bodin P., Habermann R.E. Seismic quiescence at Parkfield: an independent indication of an imminent earthquake // Nature. – 1990. – Vol. 345, № 6274. – P. 426–428.

13. Itaba S., Koizumi N. Earthquake-related changes in groundwater levels at the Dogo hot spring, Japan // Terrestrial Fluids, Earthquakes and Volcanoes: The Hiroshi Wakita Volume II. – 2008. – P. 2397–2410.

14. Lee H.A., Hamm S.Y., Woo N.C. The abnormal groundwater changes as potential precursors of 2016 M L 5.8 Gyeongju earthquake in Korea // Economic and Environmental Geology. – 2018. – Vol. 51, № 4. – P. 393–400.

15. Stein R.S., Barka A.A., Dieterich J.H.. Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering // Geophysical Journal International. – 1997. – Vol. 128, № 3. – P. 594–604.

16. Stein R.S. The role of stress transfer in earthquake occurrence // Nature. – 1999. – Vol. 402, № 6762. – P. 605–609.

17. Nalbant S.S., McCloskey J., Steacy S., Barka A.A. Stress accumulation and increased seismic risk in eastern Turkey // Earth and Planetary Science Letters. – 2002. – Vol. 195, № 3-4. – P. 291–298.

18. Ichiguchi T., Matsumura S. Building Damage Depending on Earthquake Vibration Period and New Technology Issues // Science & Technology Trends. – 2013. – Vol. 46, № 5. – P. 64–80.

19. Ribakov Y. Reduction of structural response to near fault earthquakes by seismic isolation columns and variable friction dampers // Earthquake Engineering and Engineering Vibration. – 2010. – Vol. 9, № 1. – P. 113–122.

20. Jozinović D., Lomax A., Štajduhar I., Michelini A. Rapid prediction of earthquake ground shaking intensity using raw waveform data and a convolutional neural network // Geophysical Journal International. – 2020. – Vol. 222, № 2. – P. 1379–1389.

21. Mousavi S.M., Beroza G.C. A machine‐learning approach for earthquake magnitude estimation // Geophysical Research Letters. – 2020. – Vol. 47, № 1.

22. Johnson P.A., Rouet-Leduc B., Pyrak-Nolte L.J., Beroza G.C., Marone C.J., Hulbert C., Howard A., Singer P., Gordeev D., Karaflos D., Levinson C.J. Laboratory earthquake forecasting: A machine learning competition // Proceedings of the national academy of sciences. – 2021. – Vol. 118, № 5.

23. Iaccarino A.G., Gueguen P., Picozzi M., Ghimire S. Earthquake early warning system for structural drift prediction using machine learning and linear regressors // Frontiers in Earth Science. – 2021. – Vol. 9, № 666444.

24. Panakkat A., Adeli H. Neural network models for earthquake magnitude prediction using multiple seismicity indicators // International journal of neural systems. – 2007. – Vol. 17, № 1. – P. 13–33.

25. Gitis V., Derendyaev A., Petrov K. Analyzing the performance of GPS data for earthquake prediction // Remote Sensing. – 2021. – Vol. 13, № 9.

26. Chen C.H., Yeh T.K., Liu J.Y., Wang C.H., Wen S., Yen H.Y., Chang S.H. Surface deformation and seismic rebound: implications and applications // Surveys in geophysics. – 2011. – Vol. 32. – P. 291–313.

27. Huang N.E., Shen Z., Long S.R., Wu M.C., Shih H.H., Zheng Q., Yen N.-C., Tung C.C., Liu H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis // Proceeding Royal Society London. – 1998. – Vol. 454. – P. 903–995.

28. Chen C. H., Wen S., Yeh T. K., Wang C. H., Yen H. Y., Liu J. Y., Hobara Y., Han P. Observation of surface displacements from GPS analyses before and after the Jiashian earthquake (M= 6.4) in Taiwan // Journal of Asian Earth Sciences. – 2013. – Vol. 62. – P. 662–671.

29. Chen C. H., Wen S., Liu J. Y., Hattori K., Han P., Hobara Y., Wang C. H., Yeh T. K., Yen H. Y. Surface displacements in Japan before the 11 March 2011 M9. 0 Tohoku-Oki earthquake // Journal of Asian Earth Sciences. – 2014. – Vol. 80. – P. 165–171.

30. Chen C. H., Yeh T. K., Wen S., Meng G., Han P., Tang C. C., Liu J. Y., Wang C. H. Unique pre-earthquake deformation patterns in the spatial domains from GPS in Taiwan // Remote Sensing. – 2020. – Vol. 12, № 3. – P. 366–386.

31. Mignan A., Broccardo M. Neural network applications in earthquake prediction (1994–2019): Meta‐analytic and statistical insights on their limitations // Seismological Research Letters. – 2020. – Vol. 91, № 4. – 2330–2342.

32. Aleskerov F., Say A. I., Toker A., Akin H. L., Altay G. A cluster‐based decision support system for estimating earthquake damage and casualties // Disasters. – 2005. – Vol. 29, № 3. – P. 255–276.

33. Gitis V., Derendyaev A. A Technology for Seismogenic Process Monitoring and Systematic Earthquake Forecasting // Remote Sensing. – 2023. – Vol. 15, № 8. – P. 2171–2195.

34. Aleskerov F., Baiborodov N., Demin S., Shvydun S., Trafalis T., Richman M., Yakuba V. Constructing an efficient machine learning model for tornado prediction // International Journal of Information Technology & Decision Making. – 2020. – Vol. 19, № 5. – P. 1177–1187.

# **ПРИЛОЖЕНИЕ** **А**

**Подготовка набора данных со значениями признаков**

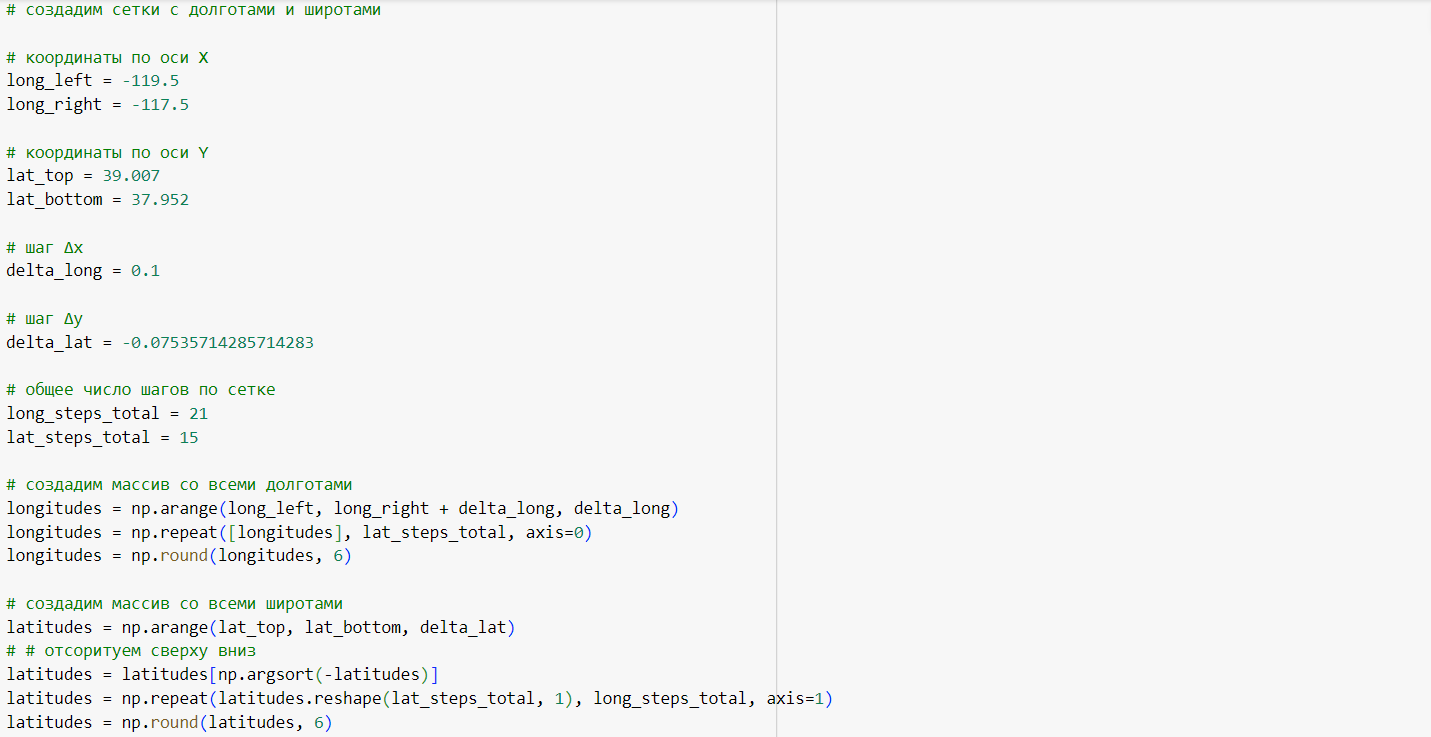


Рисунок А.1 – Создание сетки географических координат

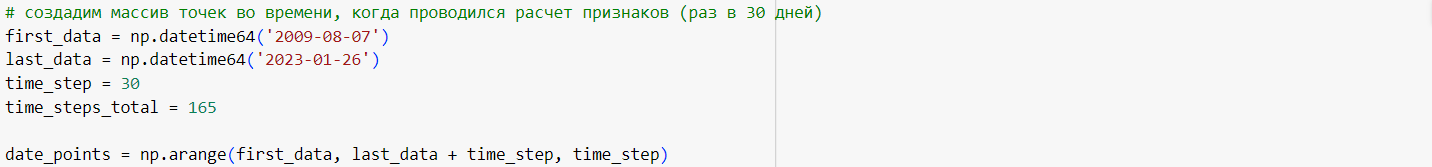


Рисунок А.2 – Создание массива точек во времени



Рисунок А.3 – Загрузка отдельных файлов со значениями признаков и объединение в общий массив

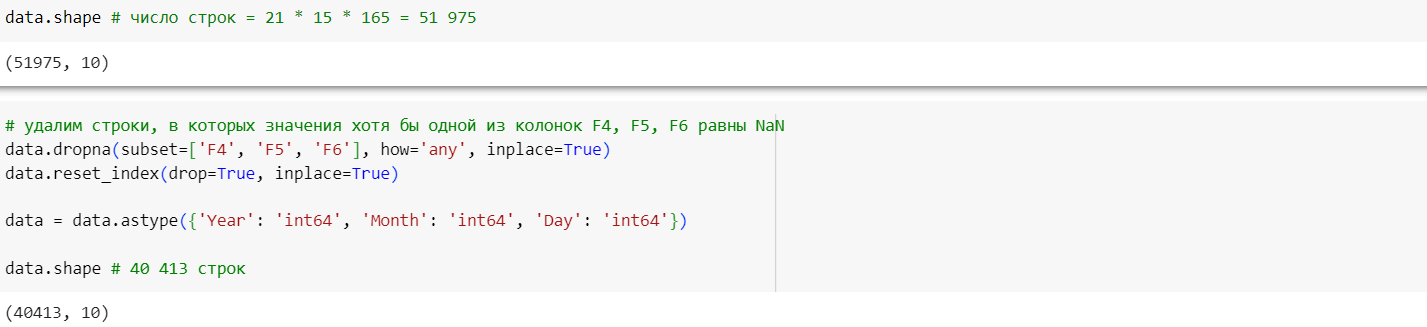


Рисунок А.4 – Удаление из массива строк, содержащих пропуски

# **ПРИЛОЖЕНИЕ Б**

**Подготовка набора данных со целевыми землетрясениями**

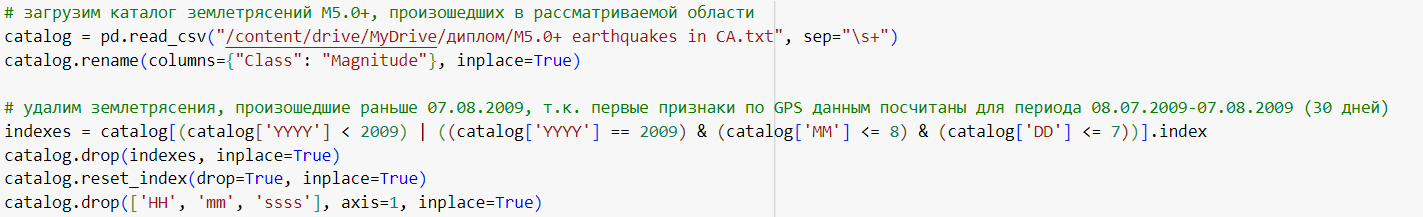


Рисунок Б.1 – Загрузка каталога всех землетрясений магнитудой 5,0 и больше, произошедших в Калифорнии

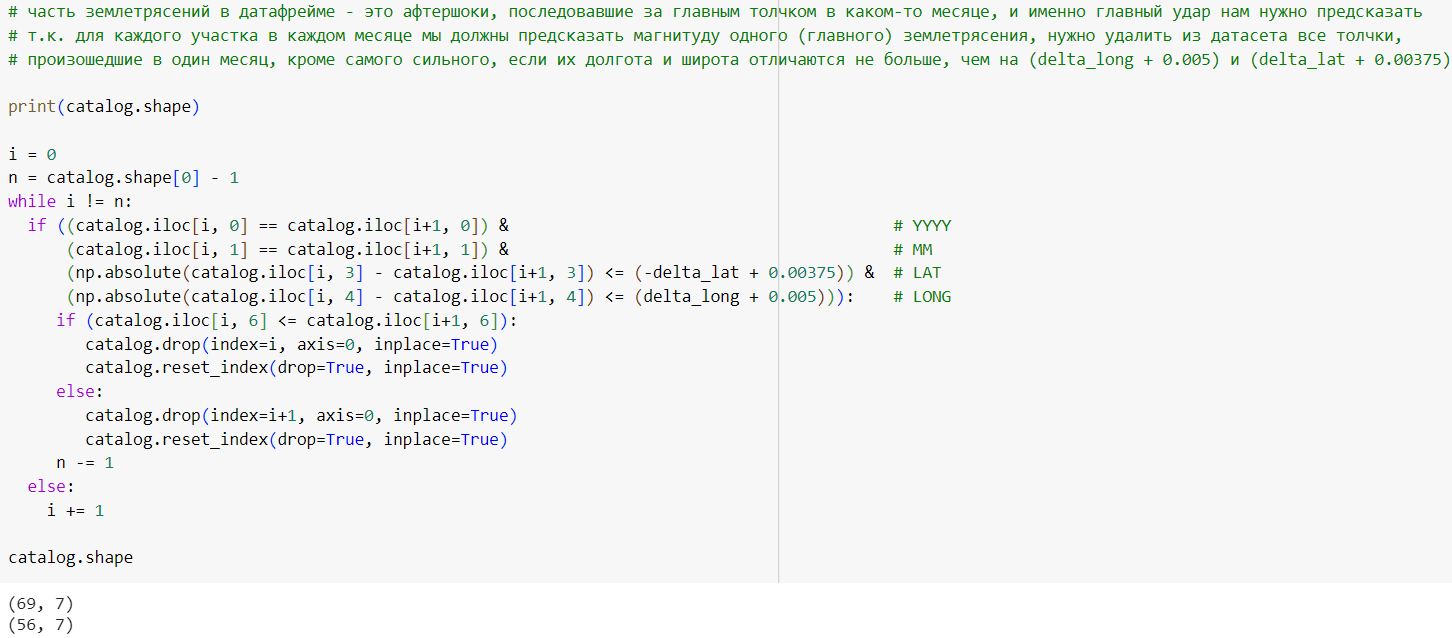


Рисунок Б.2 – Удаление из каталога афтершоков

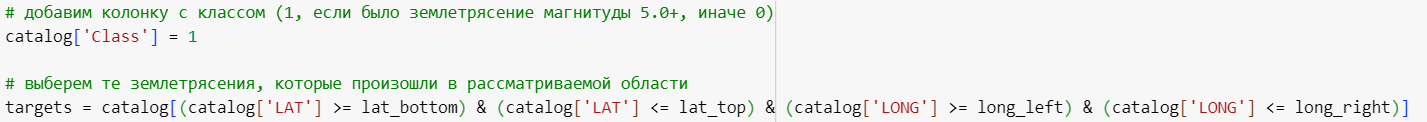


Рисунок Б.3 – Добавление колонки со значением метки класса

# **ПРИЛОЖЕНИЕ В**

**Объединение наборов данных в общий датасет**



Рисунок В.1 – Объединение массивов данных

### **ПРИЛОЖЕНИЕ Г**

**Нормализация данных**

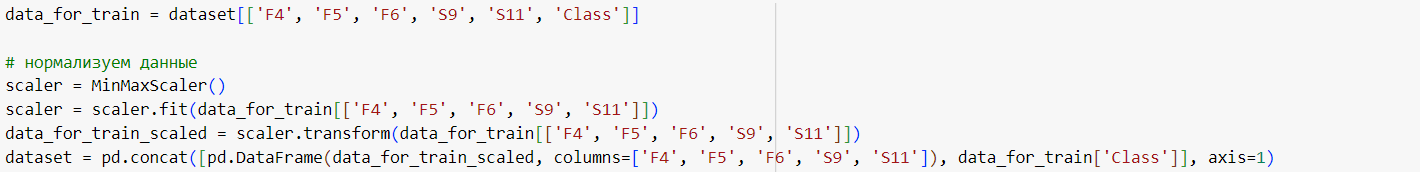


Рисунок Г.1 – Нормализация значений признаков

# **ПРИЛОЖЕНИЕ Д**

**Уменьшение дисбаланса классов в наборе данных**

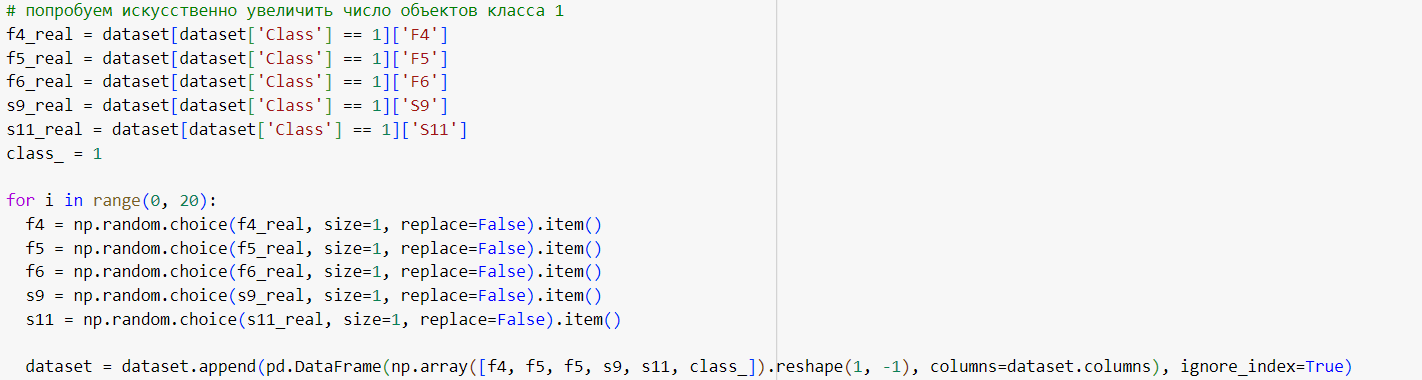


Рисунок Д.1 – Генерация искусственных объектов класса 1

# **ПРИЛОЖЕНИЕ Е**

**Обучение модели**

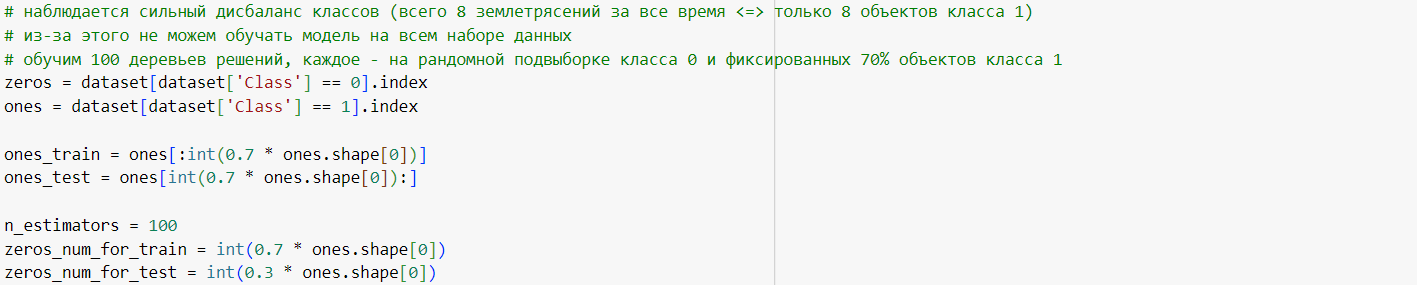


Рисунок Е.1 – Определение размеров обучающей и тестовой выборок



Рисунок Е.2 – Обучение модели

# **ПРИЛОЖЕНИЕ Ж**

**Тестирование модели**

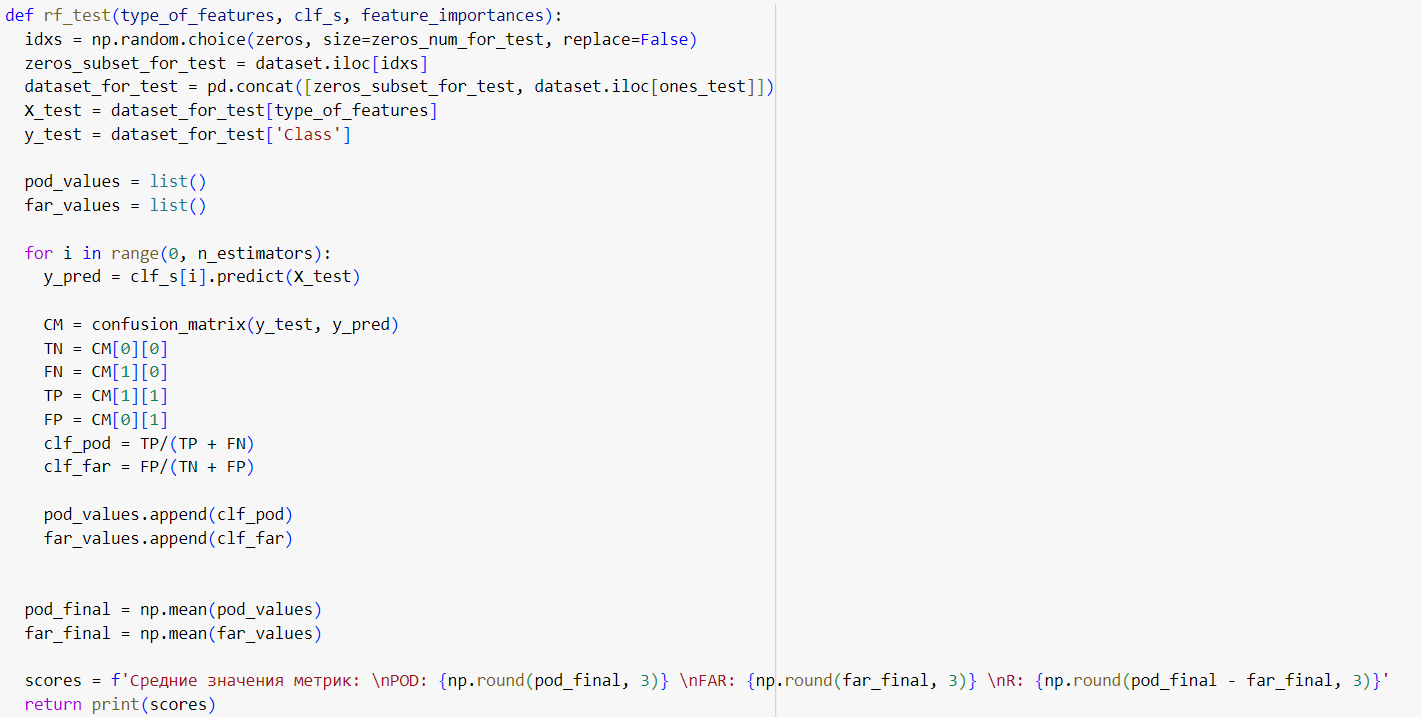


Рисунок Ж.1 – Тестирование модели

# **ПРИЛОЖЕНИЕ И**

**Применение модели**

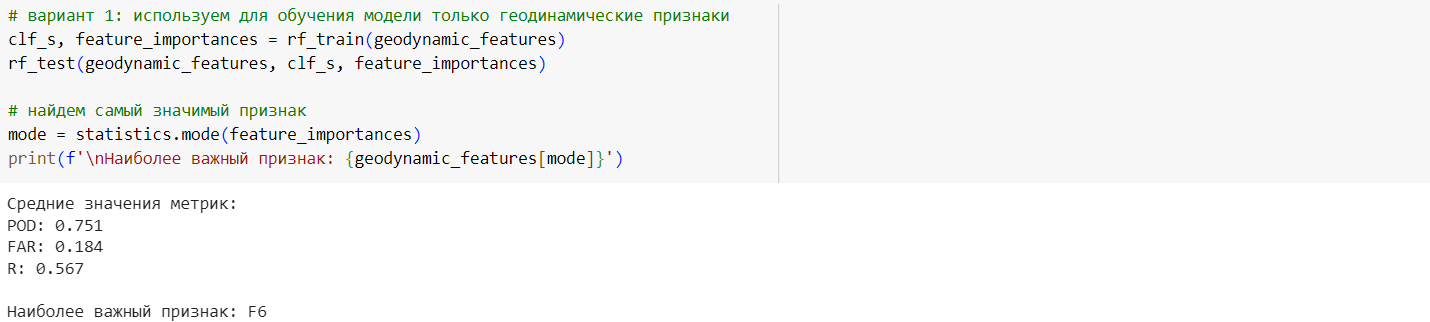
****

Рисунок И.1 – Тестирование модели, обученной на геодинамических параметрах

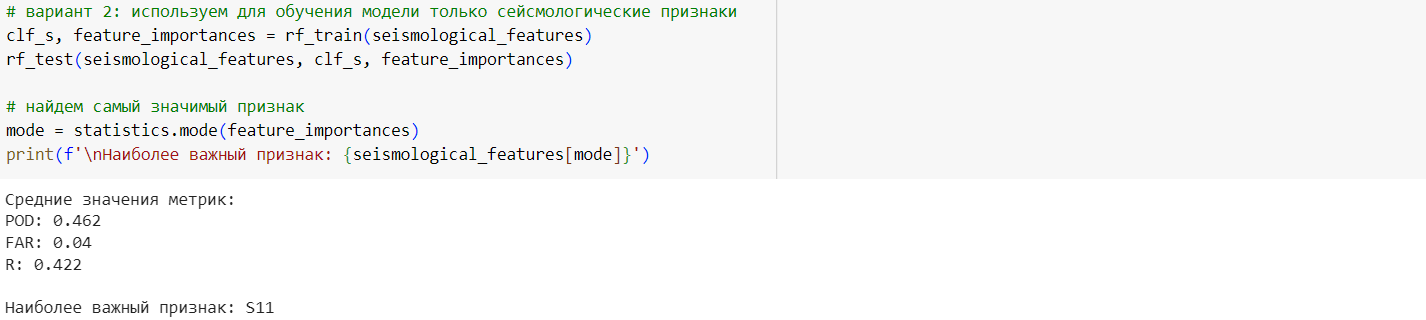


Рисунок И.2 – Тестирование модели, обученной на сейсмологических параметрах

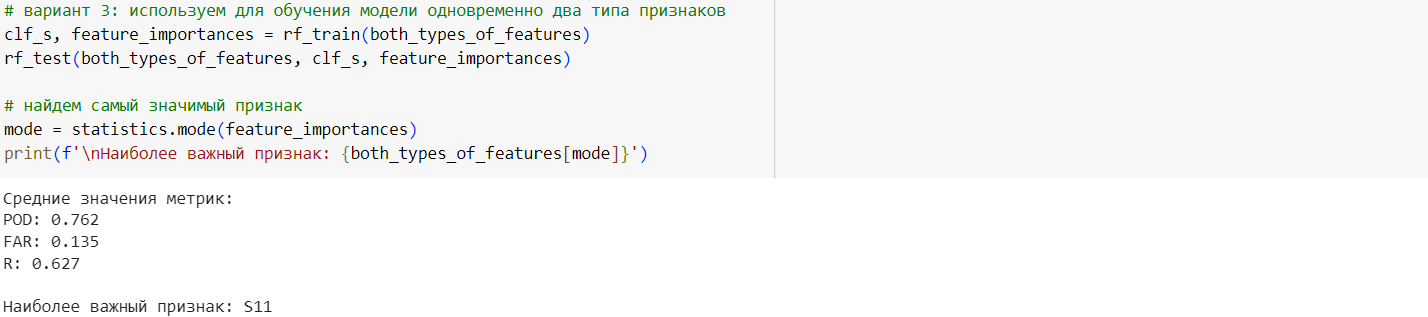


Рисунок И.3 – Тестирование модели, обученной на геодинамических и сейсмологических параметрах