

Министерство науки и высшего образования
Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Санкт-Петербургский горный университет»

На правах рукописи

Ревин Илья Евгеньевич



**ГЕОМЕХАНИЧЕСКОЕ ОБОСНОВАНИЕ ПРОЦЕССОВ
ДЕФОРМАЦИИ ТЕХНОГЕННО НАРУШЕННОГО ГОРНОГО
МАССИВА НА ПРИМЕРЕ ХИБИНСКИХ АПАТИТ-НЕФЕЛИНОВЫХ
МЕСТОРОЖДЕНИЙ**

Научная специальность - 25.00.20 Геомеханика, разрушение горных пород,
рудничная аэрогазодинамика и горная теплофизика.

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель
доктор технических наук, профессор
Господариков А.П.

Санкт-Петербург – 2021

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	5
ГЛАВА 1 КОНЦЕПЦИЯ РАЗВИТИЯ ПРОГНОЗА ПРОЦЕССОВ ДЕФОРМАЦИИ ТЕХНОГЕНО-НАРУШЕННОГО МАССИВА НА ПРИМЕРЕ ХИБИНСКИХ АПАТИТ-НЕФЕЛИНОВЫХ РУД.....	10
1.1. Концепция геомеханического мониторинга состояния массива горных пород.....	10
1.2. Система комплексного геомеханического мониторинга и ее применение для условий апатит-нефелиновых месторождений	12
1.3. Методы геомеханического мониторинга техногенно-нарушенного горного массива.....	14
1.3.1. Методы численного и математического моделирования	15
1.3.2. Методы физического моделирования	16
1.3.3. Предлагаемый подход.....	16
1.4. Анализ существующих систем сейсмического мониторинга Кировского рудника КФ АО «Апатит».....	17
1.5. Выводы по главе 1.....	20
ГЛАВА 2 ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ СЕЙСМИЧЕСКОГО МОНИТОРИНГА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ.	22
2.1 Разведывательный анализ данных сейсмического мониторинга.....	22
2.2 Оценка плотности распределения сейсмических событий во времени	29
2.3. Оценка плотности распределения сейсмических событий в пространстве деформационного мониторинга.....	32
2.4. Анализ временного ряда сейсмического мониторинга с помощью метода SSA с целью обнаружения «предвестников» потенциально опасных геодинамических событий	34
2.4.1. Методы обработки временных рядов	35
2.4.2. SVD разложение	38
2.4.3. Матрица U	38

2.4.4. Матрица V	39
2.4.5. Матрица Σ	40
2.4.6. Ранг траекторной матрицы.....	41
2.4.7. Разделение и группировка компонентов временных рядов	42
2.5. Кластерный анализ сейсмических событий в признаковом пространстве деформационного мониторинга.....	53
2.6. Выводы по главе 2.....	61
ГЛАВА 3. МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ДЕФОРМАЦИИ НА ОСНОВЕ ДАННЫХ СЕЙСМИЧЕСКОГО МОНИТОРИНГА. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СИСТЕМЫ ПРОГНОЗА ПРОЦЕССОВ ДЕФОРМАЦИИ ТЕХНОГЕНО-НАРУШЕННОГО МАССИВА.....	64
3.1. Концепция моделирования процессов деформации на основании данных сейсмического мониторинга	64
3.1.1. Многослойный персептрон (MLP)	65
3.1.2. Генетическое программирование (GP) и эволюционная регрессия	66
3.1.3 системы, основанные на нечетких правилах (FRBS).	67
3.1.4. Теория хаоса и нелинейная динамика.....	67
3.1.5. Модели основанные на данных (DDM)	68
3.2. Композитная модель прогноза процессов деформации в техногенно нарушенном массиве на основе данных его сейсмического мониторинга.....	72
3.3. Программное обеспечение системы прогноза опасных геодинамических явлений при ведении работ в техногенно-нарушенном массиве	75
3.4. Выводы по главе 3.....	77
ГЛАВА 4. ГЕОМЕХАНИЧЕСКИЙ ПРОГНОЗ ПРОЦЕССОВ ДЕФОРМАЦИИ ТЕХНОГЕНОНАРУШЕННОГО МАССИВА НА ПРИМЕРЕ ХИБИНСКИХ АПАТИТ-НЕФЕЛИНОВЫХ РУД.....	79
4.1. Прогнозирование параметров сейсмических и деформационных процессов внутри техногенно-нарушенного горного массива.....	79

4.2. Использование композитной модели для анализа данных сейсмического мониторинга за 2020 год.....	81
4.3. Анализ результатов моделирования кластеров сейсмических событий на основе данных сейсмического мониторинга за 2020–2021 года.....	85
4.4. Выводы по результатам моделирования на основе данных сейсмического мониторинга за 2020 год	88
4.5. Анализ данных сейсмического мониторинга за 2021 год.....	89
4.6. Анализ результатов моделирования кластеров сейсмических событий на основе данных сейсмического мониторинга за 2020 год.....	91
4.7. Анализ данных сейсмического мониторинга за период с 10.02.2021 по 22.03.2021	92
4.8. Выводы по главе 4.....	94
ЗАКЛЮЧЕНИЕ	96
СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ	98

ВВЕДЕНИЕ

Актуальность темы исследования

Геомеханический мониторинг напряженного состояния массива горных пород (далее МГП) – это активно развивающийся раздел геомеханики, в котором на данный момент невозможно выделить единые методологию и подход к решению задач, сбору и анализу данных при разработке систем мониторинга.

В соответствии с правилами промышленной безопасности геомониторинг должен осуществляться в непрерывном режиме, поэтому разработка математического метода анализа больших объемов данных, полученных в результате работы автоматической системы сейсмического мониторинга, является весьма актуальной научно-технической задачей, что и предопределило направленность исследований диссертационной работы.

В настоящее время заметное внимание уделяется разработке способов анализа данных мониторинга с помощью методов машинного обучения.

Степень разработанности темы исследования

Проблемами анализа данных геомеханического мониторинга занимались многие ученые, среди которых Козырев А.А., Корчак П.А., Жукова С.А., Гладырь А.В. и др.

В задачу геомеханического мониторинга может входить поиск взаимосвязей различного рода между широким перечнем различных природных и технических факторов, поскольку состояние техногенно-нарушенного массива горных пород определяется этими факторами. Можно выделить две основных группы таких факторов:

1. Природные факторы: свойства пород, слагающих массив, структурные неоднородности массива и естественное природное поле напряжений.

2. Технические факторы: методы ведения горных работ, порядок строительства объектов, применяемая система разработки месторождений полезных ископаемых, характеристики горных выработок и др.

Так же, согласно исследованиям, в результате отработки горизонтов происходит подработка пород висячего бока, которая вызывает образование зон растяжения в одной части массива и увеличение напряжений сжатия в другой части. Такие разнонаправленные воздействия на массив вызывают его растрескивание и разупрочнение.

Объект исследования – техногенно нарушенный массив горных пород на примере Хибинских апатит-нефелиновых руд.

Предмет исследования – система геомеханического мониторинга техногенно нарушенного массива горных пород.

Цель диссертационной работы – создание методики обработки данных сейсмического мониторинга и прогноза опасных геодинамических явлений с помощью современных методов машинного обучения и алгоритмов оптимизации математических моделей природных сред.

Идея работы. На основе анализа данных сейсмического мониторинга методами математического моделирования и подбором архитектуры композитной модели прогноза геодинамических явлений. Реализовать комплексный подход для обеспечения непрерывного геомеханического мониторинга.

Основные задачи исследований:

- Анализ данных сейсмического мониторинга с помощью методов математического моделирования.
- Подбор архитектуры композитной модели машинного обучения для прогноза геодинамических явлений.

- Реализация выбранной архитектуры в виде программного обеспечения, и его апробация на реальных данных сейсмического мониторинга.

Научная новизна работы:

- Выявлены новые закономерности распределения сейсмических событий во времени и пространстве МГП.
- Получен алгоритм прогноза опасных геодинамических явлений внутри массива горных пород, базирующийся на результатах моделирования разработанной композитной модели анализа данных сейсмического мониторинга.
- Выявлена аналитическая зависимость между минимальным значением тренда сейсмической активности и вероятностью наступления опасного геодинамического явления.

Теоретическая и практическая значимость исследования:

- Разработана модель кластеризации пространственно-временных измерений на основе данных сейсмического мониторинга. Алгоритм является композитным и состоит из комбинации различных алгоритмов машинного обучения. Разработанный алгоритм позволяет моделировать распределения сейсмических событий в массиве горных пород с течением времени.
- Разработаны методические рекомендации по прогнозу опасных геодинамических событий и дискретизации пространства сейсмических событий в массиве горных пород.
- Результаты диссертационной работы реализованы в виде эффективного программного вычислительного комплекса и могут быть использованы в производственном процессе.

Методология и методы исследований

Проведение исследований осуществлялось в соответствии с системным подходом, математическим и имитационным моделированием процессов в программной среде Python. Построение математической модели

кластеризации сейсмических событий в массиве горных пород и ее связи с данными станций деформационного мониторинга базируются на основных положениях теории алгоритмов эволюционной оптимизации, теории временных рядов, теории математической статистики.

На защиту выносятся следующие положения:

- Информационную модель мониторинга массива горных пород Кировского рудника КФ АО «Апатит» на основе данных сейсмического мониторинга с потенциальной возможностью учета данных деформационного мониторинга.
- Разработанную композитную модель машинного обучения, обосновывающую закономерности изменения наблюдаемых величин сейсмического мониторинга и проявления опасных геодинамических явлений рассматриваемой части массива горных пород.
- Прогноз опасных геодинамических явлений в техногенно нарушенном массиве должен включать в себя результаты математического моделирования разработанной композитной модели, обеспечивающую безопасность ведения горных работ.

Степень достоверности и апробация результатов

Обусловлены использованием современных методов математического и имитационного моделирования и удовлетворительной сходимостью результатов имитационного моделирования с экспертными оценками. Проведено широкое сравнение полученных временных зависимостей с натурными замерами датчиков сейсмического мониторинга. Основные положения и результаты работы докладывались на следующих семинарах и конференциях:

- Молодежная конференция по математическому моделированию и информационным технологиям SMIT (27.04.19-30.04.19).
- Летняя школа на базе China university of Mining and Technology (14.06.19-30.06.19).
- XII Российско-Германский сырьевой форум на базе Санкт-

Петербургского горного университета (27.11.19-30.11.19).

Публикации по работе

Результаты диссертационной работы в достаточной степени освещены в 2 печатных работах в изданиях из перечня рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук. Получено 1 свидетельство о регистрации государственной программы для ЭВМ.

Объем и структура диссертации

Диссертационная работа состоит из оглавления, введения, четырех глав с выводами по каждой из них, заключения, библиографического списка, включающего 95 наименований и изложена на 108 страницах машинописного текста и содержит 41 рисунок и 13 таблиц.

Благодарности

Автор выражает признательность научному руководителю д.т.н. А.П. Господарикову; к.т.н. К.В. Морозову за поддержку и советы на всех этапах выполнения работы и предоставленные данные сейсмического мониторинга; д.т.н. А.Н. Шабарову за ценные советы при выборе методологии исследования; также выражаю глубокую признательность коллективу научного центра геомеханики и проблем горного производства.

ГЛАВА 1 КОНЦЕПЦИЯ РАЗВИТИЯ ПРОГНОЗА ПРОЦЕССОВ ДЕФОРМАЦИИ ТЕХНОГЕНО-НАРУШЕННОГО МАССИВА НА ПРИМЕРЕ ХИБИНСКИХ АПАТИТ-НЕФЕЛИНОВЫХ РУД

1.1 Концепция геомеханического мониторинга состояния массива горных пород

Задачу геомеханического мониторинга можно описать как нахождение функциональных, аналитических или иных зависимостей между набором различных факторов, поскольку состояние техногенно нарушенного МГП определяется этими факторами.

Можно выделить две основных группы таких факторов:

- Природные факторы: свойства горных пород, слагающих массив, структурные неоднородности массива и естественное природное поле напряжений;
- Технические факторы: методы ведения горных работ, порядок строительства объектов, применяемая система разработки месторождений полезных ископаемых, характеристики горных выработок и др.

Конечная цель системы геомеханического мониторинга - оценка и контролю состояния МГП, создание единой геомеханической модели, на основе которой можно решать любые геомеханические задачи. При этом, как отмечает ряд исследователей: «Необходимыми частными задачами в ходе достижения конечной цели являются исследование упомянутых природных и технических факторов с построением соответствующих частных моделей, создание систем контроля необходимых параметров исследуемых систем и разработка мероприятий, обеспечивающих безопасную эксплуатацию рассматриваемых природотехнических систем. В ходе ведения горных работ все природные факторы претерпевают изменения, но в различной степени. В процессе отработки массива горных пород наиболее явно проявляются изменения состояния структурных неоднородностей – раскрываются имеющиеся естественные структурные неоднородности; происходят подвижки по разрывным нарушениям (разломам); образуются

новые, техногенные нарушения (трещины), которые сопровождаются изменением естественного напряженного состояния различных блоков массивов пород» [8].

Также весьма актуальной задачей является прогноз динамических проявлений горного давления для месторождений, разрабатываемых подземным способом (Хибинские апатит-нефелиновые рудники в том числе). Поскольку одним из компонентов геомеханического мониторинга является сейсмический мониторинг следует помнить, что: «...все сейсмические явления, в том числе горные удары и техногенные землетрясения, с физической точки зрения являются мгновенными (хрупкими) разрушениями некоторых объемов пород, либо подвижками блоков различных структурных неоднородностей относительно друг друга с выделением (иногда очень интенсивным) накопленной ранее энергии. Отсюда следует вывод, что для прогноза таких событий необходимо получать информацию об уровне энергии, накопленной массивом горных пород...» [8].

Ведение горных работ является причиной техногенной нарушенности МГП, что в свою очередь приводит к образованию новых систем трещин и пустот. Эти системы приводят к перераспределению естественных напряжений и разрушению пород, заполняющих тектонические нарушения. Данный процесс является динамическим и эволюционирует во времени. Это приводит постепенному разрушению всего или наблюдаемой части МГП, что ведет к росту вероятности проявления опасных геодинамических событий.

Накопленный к настоящему времени опыт организации и проведения наблюдений по контролю геомеханического состояния массивов пород в условиях различных месторождений позволяет сформулировать общие принципы организации типовой единой комплексной системы геомеханического мониторинга и обозначить основные направления ее дальнейшего совершенствования.

1.2 Система комплексного геомеханического мониторинга и ее применение для условий апатит-нефелиновых месторождений

В настоящее время при проектировании систем комплексного геомеханического мониторинга для различного рода массивов горных пород и апатит-нефелиновых месторождений в частности, довольно трудно выделить единую методологию разработки.

С одной стороны, это особенностями метода прогнозирования опасных геодинамических явлений, которые основаны на детерминированном эмпирическом подходе. Такой подход является не может быть адаптирован к изменчивым и неопределенным условиям сложной динамической системы, которой является исследуемый массив горных пород.

С другой стороны, это связано со сложностью воспроизводимостью реальных геомеханических процессов. Причины следующие:

- Грубая аппроксимация линейной зависимостью связи между напряжениями и деформациями;
- Разупрочнения горных пород при их запредельной деформации;
- Свойства неоднородных горных пород достаточно трудно учитывать в ходе математического моделирования;

Таким образом, на текущий момент, не удалось выработать четких рекомендаций по организации станций геомеханического мониторинга. Существующие методы, как правило, применяют экспертный подход, на основе эмпирических данных о поведении исследуемого массива.

Такой подход к организации системы геомеханического мониторинга является неоптимальным. Развитие во времени взаимодействия системы «крепь — массив» не может быть представлена серией независимых расчетов, необходимо динамическое отображение состояния системы.

Таким образом, выбор программного обеспечения и методики анализа данных сейсмического мониторинга является весьма актуальной научно-технической задачей. «Опыт выполнения геомеханического мониторинга на различных горных предприятиях Кольского полуострова убеждает в том, что

общая концепция организации геомеханического мониторинга должна предусматривать выполнение следующих обязательных этапов» [8].

Первый этап мониторинга включает в себя результаты моделирования МГП, полученные на этапе геологических исследований. Также на первом этапе формируется единое информационное пространство (база данных), в котором агрегируется информация об каждом из этапов работы системы геомеханического мониторинга. Первый этап состоит из следующих шагов:

- Идентификация блочной структуры массива, районирование зон проявлений опасных геодинамических явлений;
- Выбор модели, которая описывает напряженно-деформированное состояние (далее НДС) МГП с учетом геологического строения;
- Оценка НДС массива горных пород и его изменения во времени;
- Определение опасных участков МГП, который потенциально могут являться зонами опасных геодинамических явлений;

Второй этап мониторинга начинается на стадии проектирования подземного сооружения. Второй этап состоит из следующих шагов:

- Уточнение структуры, расположения, мощности и особенности строения блоков массива горных пород. Подробное рассмотрение активных зон проявлений опасных геодинамических явлений;
- Создание и итерационное обновление модели напряженно — деформированного состояния МГП, с учетом актуальной информации;
- Математическое моделирование и прогноз движения существующих и формирования новых блоков. Отображение существующих деформаций в массиве горных пород под действием природных и искусственных факторов.

Третий этап мониторинга проводят при строительстве подземного сооружения. Третий этап состоит из следующих шагов:

- Определение горного давления в массиве методами разгрузки и геофизическими методами;

- Измерение величины модуля деформации и коэффициента Пуассона;
- Контроль сейсмической активности МГП.

Четвертый этап мониторинга связан с работами по вводу в эксплуатацию и работой системы в реальном времени. На четвертом этапе точных шагов не существует, поскольку этот этап заключается в постоянном мониторинге состояния МГП.

Таким образом, алгоритм мониторинга можно представить следующим образом:

- Создается начальная модель МГП;
- При появлении новых данных проводится новое численное моделирование. Эти данные передаются в единую базу данных и по каждой контрольной точке можно узнать, что происходило в данном месте в массиве горных пород с момента первого моделирования;
- Затем, по результатам анализа, можно сделать вывод о том, требуется ли дополнительные методы мониторинга в заданной точке пространства. Таким образом, в течение всего срока эксплуатации подземного сооружения в выбранной точке будет накапливаться информация, которая позволит своевременно предотвратить проявления опасных геодинамических явлений.

1.3 Методы геомеханического мониторинга техногенно-нарушенного горного массива

Ввиду сложности исследуемого объекта для учета изменчивости и неопределенности геологических и инженерных условий при прогнозировании проявления горного давления следует учитывать эмпирические и априорные знания о массиве горных пород. Например: «...в условиях Хибинского массива по мере увеличения выработанных пространств в проявлениях сейсмичности все более четко просматривается влияние крупных структурных неоднородностей – тектонических нарушений (разломов) в пределах шахтных полей. Это подтверждает необходимость

организации геомеханического мониторинга по всей иерархии структурных блоков, слагающих рассматриваемый массив горных пород. Этим объясняется наличие акустической эмиссии, предшествующей разрушению испытуемого под прессом образца горных пород, а также «шумы» массива перед горным ударом или форшоковые явления перед землетрясением. Об этом же свидетельствует достаточно четкая последовательность развития динамических проявлений в массивах пород по мере возрастания степени напряженности массива пород» [8].

- Начальные формы – шелушения и стреляния;
- Горные и горнотектонические удары;
- Техногенные землетрясения.

1.2.1 Методы численного и математического моделирования

Алгоритмы математического и численного моделирования являются хорошо изученными методами, которые используются при ведении геомеханического мониторинга. Как следует из работы ряда исследователей, занимавшихся вопросами моделирования опасных геодинамических явлений в техногенно-нарушенном массиве горных пород: «...при решении геомеханических задач, как правило, наиболее широко используется метод конечных элементов (МКЭ), при этом применяемые модели разделяются в зависимости от типа решаемой задачи (линейной или нелинейной, плоской или объемной) и метода получения нелинейных решений. В ряде случаев для получения решений используется итерационный принцип, поскольку итерационные методы обладают надлежащей сходимостью» [7].

Выбор одного или набора геомеханических параметров которые будут рассчитываться в ходе моделирования, как правило: «...зависит от физического закона, который устанавливает взаимосвязь между напряжением и деформацией. Иногда невозможно выявить взаимосвязь между полной деформацией и напряжением, в то время как соотношение для их приращений может быть выведено. Широкое применение компьютерного

моделирования для решения геомеханических задач, и тенденция к автоматизации расчетов требуют устойчивости алгоритмов к различным типам ошибок, возникающих в результате несоблюдения граничных условий, неправильной дискретизации расчетной области, округления вычислений» [9].

1.2.2 Методы физического моделирования

С другой стороны, существуют методы физического моделирования на эквивалентных материалах, основанный на теориях подобия и размерности, который позволяет воспроизводить фактический процесс выемки полезного ископаемого на заданной глубине его залегания. При таком подходе параметры областей разрушения определяются с помощью геофизических методов, в частности методов сейсмического контроля, ультразвукового профилирования, сейсмоакустических методов, сейсмотомографии и др.

Из этих методов в настоящее время в наибольшей степени изучены и потому находят широкое применение сейсмические и сейсмотомографические методы. Исходя из вышесказанного, эффективным является ансамблирование деформационных методов (геодезических) с сейсмическими методами контроля состояния массива. Если деформационные методы позволяют получить информацию о формировании энергонасыщенных зон, связанных с начальными этапами разрушений, то сейсмические методы отражают динамику реализации разрушений в структурных неоднородностях МГП и позволяют создавать системы, обеспечивающие контроль на разных иерархических уровнях.

1.2.3 Предлагаемый подход

В итоговом прогнозе следует учитывать данные, полученные в результате геомеханического мониторинга, включающие в себя результаты математического и физического моделирования контролируемых параметров. Используя данный подход, а также принцип Парето-оптимальности, можно реализовать концепцию адаптивного поведения

системы геомеханического мониторинга. Таким образом, можно выполнить переход от статической постановки задачи к динамической.

При таком подходе проблема неопределенности геомеханических моделей решается сочетанием результатов физического и компьютерного моделирования методами ансамблирования. «Результаты моделирования используются для определения статистических зависимостей, необходимых для прогнозирования проявлений горного давления с допустимой вероятностью и принятия инженерных решений по обеспечению устойчивости выработок. Идея такого подхода заключается не в выборе методов, связанных с конкретными формами проявления горного давления, а в итеративном улучшении прогноза за счет адаптивного управления, т. е. путем адаптации технологии к окружающей среде» [10].

Подводя итог данной главы, функционирование системы геомеханического можно описать следующими шагами:

- Геомеханический мониторинг должен основываться на обработке данных о состоянии МГП (например, напряжения, деформации, сейсмическая активность и т. д.).
- Работа система геомеханического мониторинга по оценке и контролю изменений в состоянии МГП должна осуществляться в бесперебойном режиме;
- Прогноз опасных геодинамических явлений должен базироваться на результат комплексного моделирования, исследуемого МГП, с учетом данных системы геомеханического мониторинга.

1.3 Анализ существующих систем сейсмического мониторинга

Кировского рудника КФ АО «Апатит»

На данный момент достаточно трудно выделить комплексное решение проблемы геомеханического мониторинга объектов горной промышленности, хотя примеры решения локальных задач были реализованы в достаточной мере. Как отмечает ряд исследователей: «..не является исключением и состояние указанной проблемы на рудниках АО

«Апатит», где имеются примеры успешного решения отдельных частных задач и реализации методических разработок, но еще предстоит выполнить большой объем работ для создания единой комплексной системы мониторинга, способной решать различные геомеханические вопросы, в том числе и основной практический вопрос по повышению надежности прогноза динамических проявлений горного давления – горных ударов и техногенных землетрясений...» [8].

Следует отметить, что на данный момент выполнен достаточный объем исследований, связанных с апатит-нефелиновыми месторождениями.

- Вопросы тектоники массива в областях сильного техногенного воздействия рассмотрены в работе [13],
- Физико-механические свойства исследованы в работе [1],
- Проблема, связанная с изучением структурных неоднородностей массивов достаточно полно раскрыта в работе Ф. М. Онохина [12],
- Определение и анализ естественного напряженного состояния работы И. А. Турчанинова, А. А. Козырева [15].

Многочисленные исследования [12-15] показывают, что техногенная деятельность может быть и причиной, и источником сейсмических волн. Техногенную сейсмичность можно рассматривать как: «Микроколебания, сейсмические толчки и землетрясения, возникающие в земной коре при любых антропогенных воздействиях на окружающую среду. При этом источниками энергии колебаний могут быть как непосредственное техногенное воздействие, так и собственные энергетические запасы в самих структурах земной коры, либо оба источника одновременно» [16].

В настоящее время на рудниках АО «Апатит»: «...развернута сейсмическая сеть, обеспечивающая контроль состояния массива пород в зоне проведения горных работ с разрешением в пространстве порядка от нескольких до сотен метров. Совершенствование методов фиксации разрушений в массиве, т. е. сейсмической системы контроля, должно быть направлено на повышение информативности получаемых результатов, на

разработку методики определения соответствия наблюдаемых разрушений – очагов сейсмических явлений конкретным рангам структурных неоднородностей, а также на методы определения пороговых значений энергии разрушения для них» [10].

Рассмотрим результаты геомеханического мониторинга для условий техногенно нарушенного массива апатит-нефелиновых месторождений, которые наиболее полно отображены в статье Корчака П. А. и Жуковой С. А.

- «Контроль за сейсмичностью на рудниках ОАО «Апатит» ведут специалисты центра геомеханического мониторинга. Мониторинг сейсмичности на подземных рудниках осуществляется с помощью автоматизированной системы контроля состояния массива (АСКСМ), способной регистрировать геодинамические явления с энергией $>10^2$ Дж, и погрешностью определения гипоцентра в зоне повышенной точности не ниже 25 м» [5];

- «Технологические процессы при ведении горных работ оказывают существенное влияние на сейсмический режим рудников. К настоящему времени на рудниках ОАО «Апатит» документально зарегистрировано около 40 горных ударов, а в последние пять лет в районе ведения крупномасштабных горных работ фиксируются сейсмические события (с/с) с энергией до 10^{12} Дж. В сложных условиях отработки месторождений для обеспечения безопасности работ, а также выполнения принятых технологических решений с предсказуемой реакцией массива на тип воздействия необходим контроль геодинамических явлений» [5];

- «За время наблюдения за сейсмичностью в массиве были выявлены такие закономерности:

- реакция массива на массовые взрывы;
- сезонный рост сейсмичности в периоды таяния снега;
- рост сейсмичности Центрального рудника в периоды затяжных дождей;

- рост сейсмичности перед обрушением консоли налегающих пород;
- взаимосвязь сейсмичности на разных структурных блоках» [5].

Так как «...на КФ АО «Апатит» в течение долгого периода службой прогноза и предотвращения горных ударов проводился ряд региональных и локальных видов мониторинга, среди которых выделяется сейсмический» [2], в данной работе в качестве метода контроля состояния массива горных пород рассматриваются данные сейсмического мониторинга. Пример данных приведен в таблице 1.4.1

Таблица 1.4.1 - Пример формата таблицы сейсмического мониторинга

Дата регистрации	Класс события	Координата по оси X	Координата по оси Y	Координата по оси Z	Величина энергии сейсмического события
ELocTime	ETypeId	EX	EY	EZ	EEnergy
02.11.2018	13	1206,421	502,3779	7,84479	1529,814
02.11.2018	0	2204,33	464,1606	148,638	388,1517
02.11.2018	13	3052,182	586,7765	73,39877	17624,06

1.4 Выводы по главе 1

По результатам первой главы можно сделать следующие выводы:

- Произведен обзор предметной области. Сформулирована постановка задачи геомеханического мониторинга и проведен анализ процессов деформации техногенно нарушенного массива на примере Хибинских апатит-нефелиновых руд;
- Сделан обзор актуальных исследований по теме диссертации;
- Произведен анализ существующих систем геомониторинга, рассмотрена существующая методология проведения геомониторинга и ее недостатки;
- Приведена терминология, а также изложены известные результаты ряда разделов математического моделирования природных сред;

- Произведен анализ существующих математических и физических методов геомеханического мониторинга и предложен комплексный метод моделирования;
- Произведен анализ существующей системы сейсмического мониторинга Кировского рудника КФ АО «Апатит».

ГЛАВА 2 ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ СЕЙСМИЧЕСКОГО МОНИТОРИНГА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ.

2.1 Разведывательный анализ данных сейсмического мониторинга

Горный удар — это чрезвычайно сложное динамическое явление. Для прогнозирования и предотвращения горных ударов в последние десятилетия были применены различные методы, такие как полевой мониторинг, лабораторные испытания, теоретические модели, эмпирические модели, численные модели, интеллектуальные методы и т. д. Из-за сложности и неопределенности горного удара, механизм его формирования до сих пор четко не изучен. горных выработок и горных работ. На формирование горного удара влияет множество факторов, такие как физико-механические свойства массива горных пород, напряженное состояние, геологическое строение, инженерное положение [85-95].

Несмотря на то, что было разработано большое количество методов для прогнозирования или оценки горного удара, за последние несколько десятилетий не было достигнуто большого прогресса и не было получено общепринятого метода прогнозирования горного взрыва [73-84]. Одним из общепринятых подходов к прогнозу горных ударов является использование единичных индикаторов и мультииндикаторов [17, 33]. В основном они основаны на свойствах горных пород, энергии, глубине выработки и т. д.

Поэтому возникает необходимость в разработке специальных методов обработки информации данных сейсмического мониторинга с целью определения характеристик техногенно-нарушенного массива. Данные сейсмического мониторинга однозначно определяются четырьмя переменными: пространственными координатами $\langle X, Y, Z \rangle$ и временной координатой t . Поэтому в большинстве случаев задача обработки данных мониторинга сводится к анализу пространственно-временных свойств данных, которые можно представить в виде:

- временных рядов (пространственные координаты фиксированы);

- полей (временная координата фиксирована);
- пространственно-временных полей.

Таким образом, при прогнозе проявления горного давления в виде горного удара приходится сталкиваться с ситуациями, когда в определенной точке пространства и времени система мониторинга фиксирует события с высоким значением энергии, которые, однако не сопровождаются опасными геодинамическими проявлениями. В таких ситуациях возникает необходимость проведения интеллектуального анализа данных мониторинга на основе специфических методов математического моделирования. «Классический подход в проведении анализа информации состоит в применении параметрических методов. Однако, использование параметрических методов связано с формулировкой предположений о виде закона распределения наблюдаемых случайных величин. При ограниченном объёме информации нельзя указать каких-либо веских причин, по которым конкретное распределение результатов наблюдений должно входить в то или иное параметрическое семейство» [22]. Следует отметить что: «наиболее общей характеристикой, описывающей поведение одномерной случайной величины, является ее плотность распределения $f(t)$. Зная плотность распределения случайной величины, можно однозначно определить такие характеристики как вероятность наступления определенного сейсмического события, интенсивность таких событий, среднее время между событиями. Среди непараметрических методов построения плотности (функции) распределения широкое распространение находят гистограммные (построение гистограммы или вариационного ряда), проекционные, ядерные и вейвлет оценки. Таким образом, зная плотность или интегрируемую функцию распределения случайной величины, можно перейти к определению характеристик массива горных пород» [22]. В таблице 2.2.1 приведен пример данных сейсмического мониторинга. На рисунке 2.1.1 показана гистограмма распределения сейсмических событий по величине энергии с января 2018 по декабрь 2018 года.

Таблица 2.1.1 - Пример формата таблицы сейсмического мониторинга

X	Y	Z	Время	Энергия, Дж
30269.0	40140.0	464.0	2018.01.05 00:00:01	5900.0
30250.0	40201.0	419.0	2018.01.05 00:00:30	3150.0
30511.0	40294.0	416.0	2018.01.05 00:21:37	12900.0

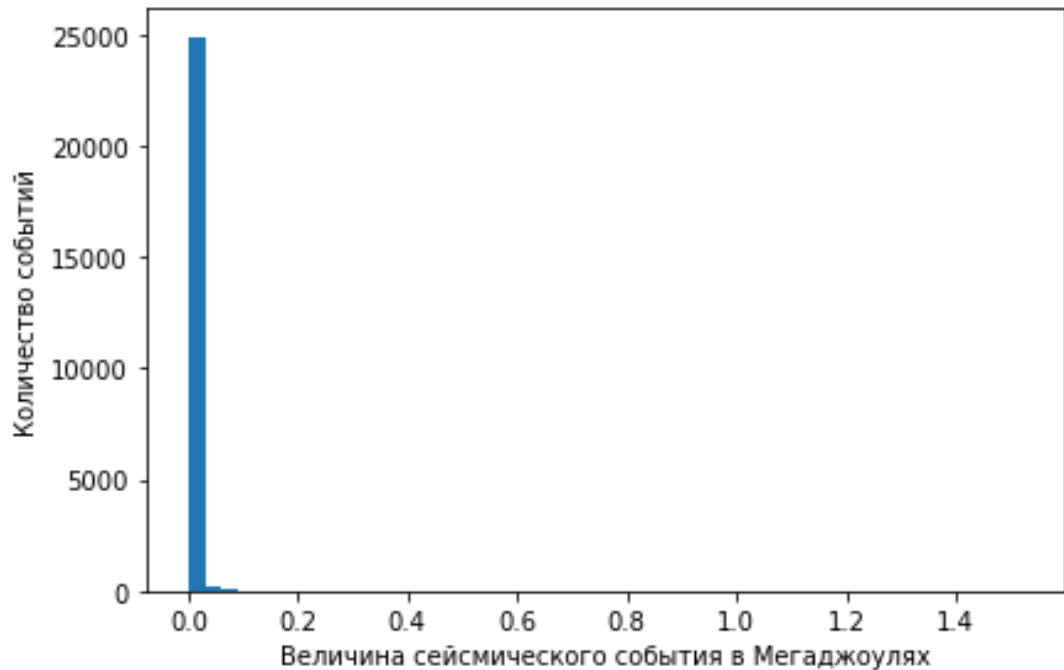


Рисунок 2.1.1 – Распределение сейсмических событий по величине энергии с января 2018 по декабрь 2018 года

Ниже приведены статистические характеристики распределения:

- Число сейсмических событий – 25128 шт.
- Среднее значение – 1480725 Дж.
- Дисперсия – 14824970 Дж.
- Минимальное значение – 0 Дж.
- 25% квартиль (Q_1) - 1000 Дж.
- 50% квартиль (Q_2) – 4695 Дж.
- 75% квартиль (Q_3) – 41000 Дж.
- Максимальное значение – 152000000 Дж.

Как видно из гистограммы, приведенной на рисунке 2.1.1, и значений характеристик, в нашем распределении присутствуют аномально высокие значения, из-за которых распределение имеет эксцесс.

С точки зрения задачи прогноза опасных геодинамических явлений, необходимо знать значения сейсмического мониторинга в диапазоне от 10^2 до 10^5 Дж. Все значения, которые находятся выше этого диапазона, как правило, относятся к событиям, связанным с проведением горных работ (взрывные и очистные работы). Для того, чтобы определить границы статистически значимой выборки (без выбросов), можно использовать значение межквартильного размаха (IQR), которое лежит в основе статистического приема анализа данных под названием «диаграмма размаха». Границами такой диаграммы служат первый и третий квартили (25-й и 75-й квартили соответственно), линия в середине диаграммы — медиана. Наиболее распространённым значением, определяющим значение IQR, является разность первого квартиля и полутора межквартильных расстояний; сумма третьего квартиля и полутора межквартильных расстояний приведена в формуле 2.1.1:

$$X = Q_3 + K(Q_3 - Q_1), \quad (2.1.1)$$

где K , как правило равно 1,5. Таким образом, верхняя граница нашей выборки X равна – 101000 Дж. В качестве нижней границы принято эмпирическое значение, равное 100 Дж.

На рисунке 2.1.2 приведена гистограмма распределения сейсмических событий в заданном диапазоне.

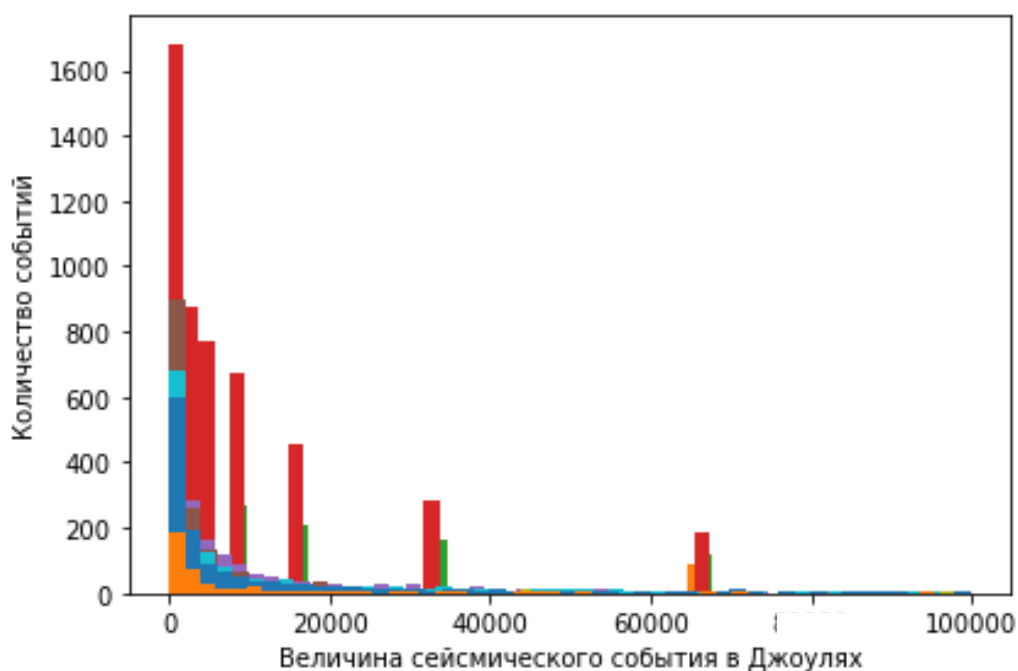


Рисунок 2.1.2 – Распределение сейсмических событий по величине энергии в диапазоне от 100 до 100000 Джоулей с января 2018 по декабрь 2018 года

Как видно из рисунка 2.1.2, форма распределения сейсмических событий сохраняется от месяца к месяцу, однако, следует отметить наличие аномальных значений, которые встречаются в ряде месяцев. В таблице 2.1.2 приведены значения, где N – число сейсмических событий за месяц; P – доля уникальных значений энергетической сигнатуры сейсмического события.

Таблица 2.1.2 - Пример формата таблицы сейсмического мониторинга

Месяц	N	P	$\frac{P}{N}$
Январь	1885	469	0.86
Февраль	1885	457	0.24
Март	2438	452	0.18
Апрель	5084	428	0.084
Май	2137	1355	0.63
Июнь	1775	1139	0.64
Июль	1013	784	0.77
Август	1149	865	0.75
Сентябрь	810	655	0.80

Продолжение таблицы 2.1.2

Октябрь	1679	1185	0.70
Ноябрь	1418	1012	0.71
Декабрь	459	401	0.87

Как следует из таблицы 2.1.2, наиболее аномальными с точки зрения количества уникальных сейсмособытий являются месяцы с февраля по апрель. На гистограмме, изображенной на рисунке 2.1.3, показано распределение сейсмических событий за текущие месяца.

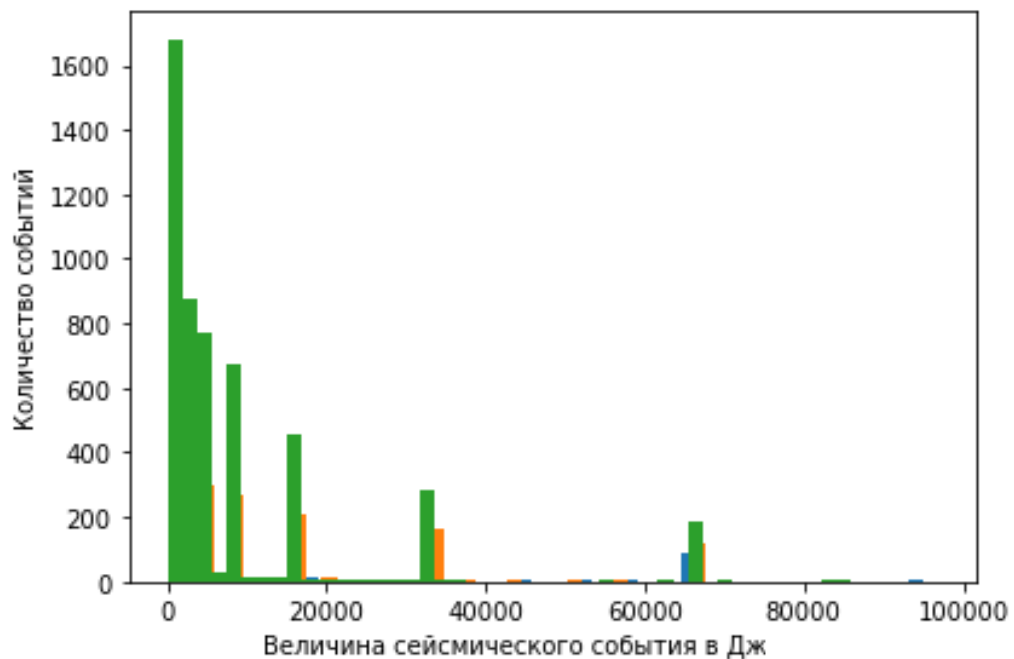


Рисунок 2.1.3 – Распределение сейсмических событий по величине энергии с февраля 2018 по апрель 2018 года

Как видно из графика 2.1.3 присутствует четко выраженный конечный набор значений, которые следует либо более подробно исследовать, либо изъять из исходного набора данных, либо дать им какую-либо интерпретацию. Поскольку в значениях, данных присутствует кратность, а также данные значения не имеют четкой привязки во времени и местоположению, их можно изъять, связав их с особенностями работы существующей системы сейсмического мониторинга. Список значений приведен ниже.

- 2000.0 Дж – 1382 события;

- 4100.0 Дж– 1211 событий;
- 8200.0 Дж – 1083 события;
- 16000.0 Дж – 761 событие;
- 33000.0 Дж – 528 событий;
- 66000.0 Дж– 392 события;

На гистограмме, изображенной на рисунке 2.1.4, показано распределение сейсмических событий по величине энергии с января 2018 по декабрь 2018 года, после фильтрации исходного набора данных.

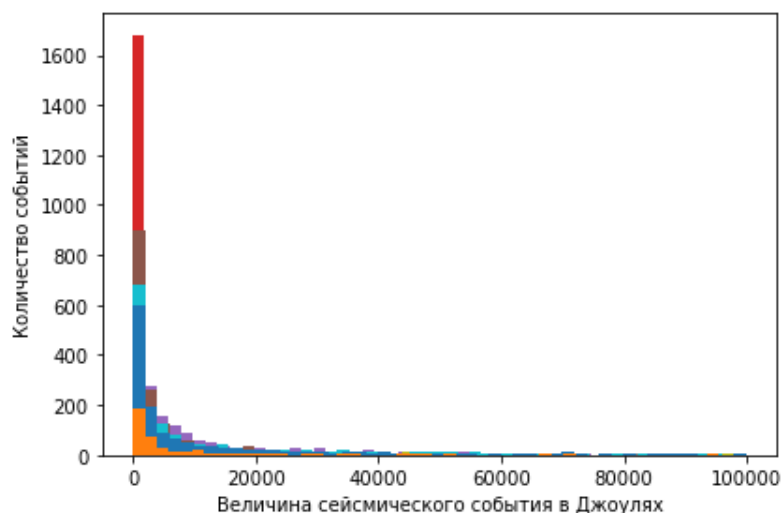


Рисунок 2.1.4 – Распределение сейсмических событий по величине энергии с января 2018 по декабрь 2018 года, после фильтрации исходного набора данных

В таблице 2.1.3 приведены значения наиболее значимых статистических характеристик полученного распределения

Таблица 2.1.3 – Пример формата таблицы сейсмического мониторинга

Название характеристики	Обозначение характеристики	Значение характеристики в Дж
Среднее значение	$\bar{x}$	9620
Среднеквадратичное отклонение (СКО)	$s = \sqrt{D}$	17990
Медиана	Q_{50}	1610

Продолжение таблицы 2.1.3

Усеченное среднее	$\bar{x}$ из которого исключены k % наибольших и k % наименьших значений	4909
Среднее абсолютное отклонение	$MAD = \frac{1}{n} \sum_{i=1}^n (x_i - m(X))$	3032

На данном этапе анализа сделаем следующие предположения:

1. Месяцы с января по декабрь имеют сходный режим протекания сейсмических событий в массиве горных пород.
2. Среди значений сейсмических событий присутствует набор значений, связанный с особенностью работы системы сейсмического мониторинга.

2.2 Оценка плотности распределения сейсмических событий во времени

Ядерная оценка плотности (рисунок 2.2.1) является непараметрическим способом оценки функции плотности вероятности случайной величины. Оценка плотности ядра — это фундаментальная задача сглаживания данных, в которой делаются выводы на основе конечной выборки данных.

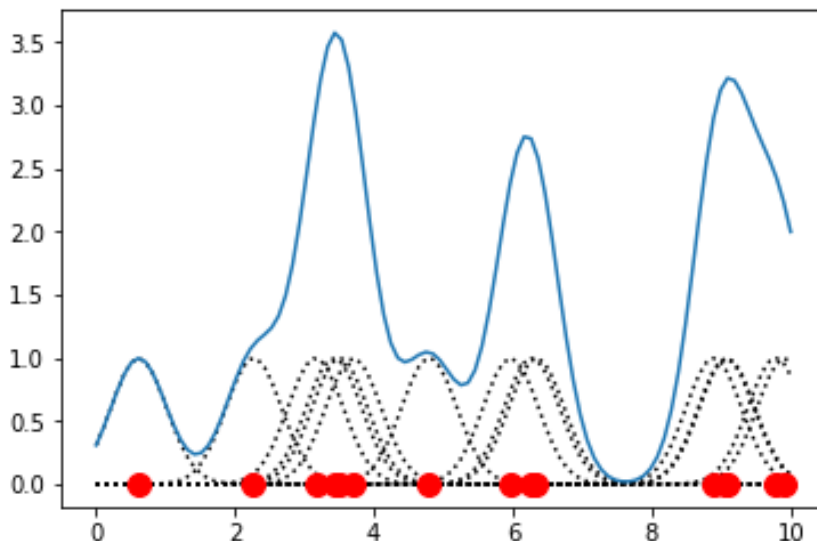


Рисунок 2.2.1 – Пример применения ядерной оценки плотности
распределения с центрами в выборочных значениях

Для получения «гладких» оценок плотности распределения часто применяется метод ядерного сглаживания. Ядерных функций много, но

обычно достаточно гауссовой функции. Суть ядерной оценки плотности распределения состоит в замене прямоугольных непересекающихся столбцов гистограммы суммой кривых (функций) с центрами в выборочных значениях:

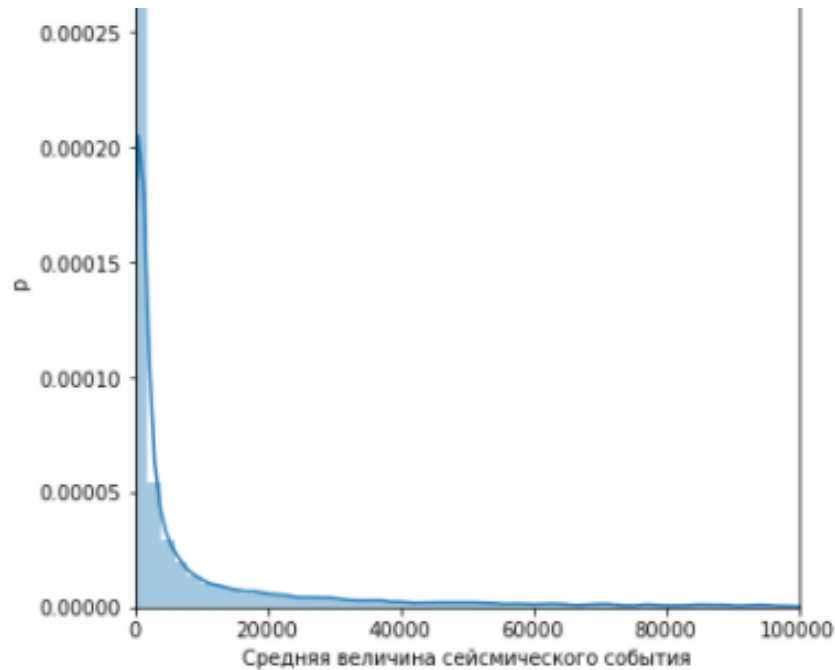


Рисунок 2.2.2 – Сглаженное распределение сейсмических событий по величине энергии с января 2018 по декабрь 2018 года, после фильтрации исходного набора данных

Проведя визуальный анализ гистограммы, было сделано предположение, что имеющийся набор данных может быть описан логнормальным распределением (2.2.1).

$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{\ln(x-\mu)^2}{2\sigma^2}\right), \quad (2.2.1)$$

где μ – параметр, который смещает распределение влево или вправо относительно оси x ; а σ – параметр формы распределения. Для подбора параметров этого распределения воспользуемся методом оптимизации под названием «Метод максимального правдоподобия (Maximum Likelihood)». Данный метод предпочтителен, поскольку результирующая оценка, как известно, обладает хорошими теоретическими свойствами. Опишем механизм применения метода:

1. Допустим, что значения энергетических сигнатур сейсмических событий, представленные в виде $x = x_1, x_2, \dots, x_n$, распределяются в соответствии с неким распределением $\Pr(Y_i = y_i | \Theta)$;

2. Произведение $\prod_{i=1}^n \Pr(y_i | \Theta)$ дает оценку того, насколько «вероятно» наблюдать значения $x = x_1, x_2, x_3, \dots, x_n$ при заданных значениях параметра Θ ;

3. Поиск максимального правдоподобия состоит в выборе соответствующей функции $L = \Pr(Y | \Theta)$, которую необходимо максимизировать для данного набора наблюдений;

4. Эта функция называется функцией правдоподобия, потому что является мерой того, насколько вероятны наблюдения, если модель истинна.

В ходе оптимизации были получены параметры и итоговая формула функции распределения сейсмических событий в пространстве МГП, отраженные в формуле (2.2.2):

$$f(x, \mu, \sigma) = \frac{1}{7.91 * x * \sqrt{2\pi}} \exp\left(-\frac{\ln(x-1.81)}{2 * 7.91^2}\right). \quad (2.2.2)$$

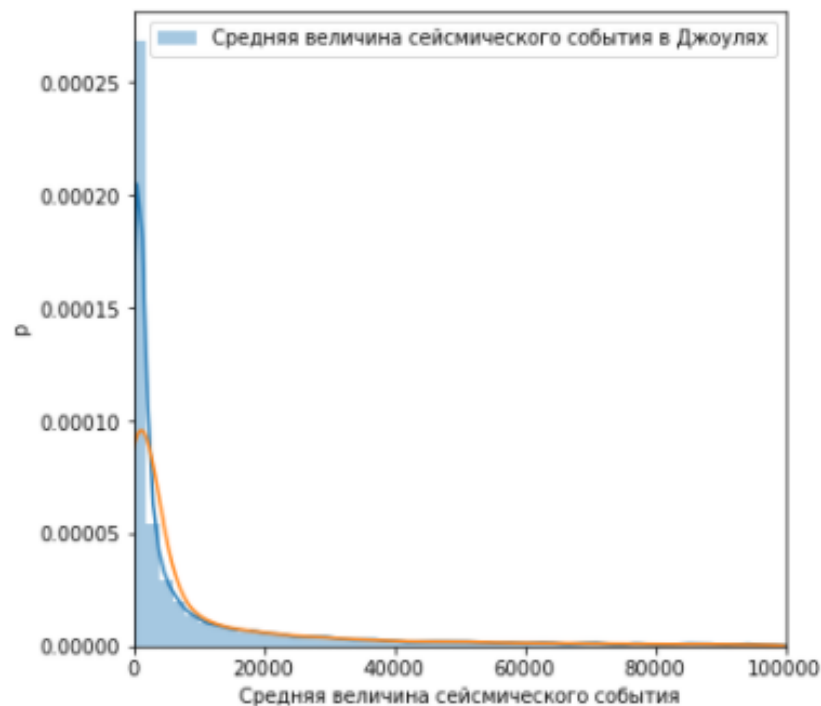


Рисунок 2.2.3 - Оценка распределения величины сейсмических событий в диапазоне свыше Q4 за период февраль-сентябрь 2018 года. Желтая линия — кривая плотности распределения с заданными параметрами

2.3 Оценка плотности распределения сейсмических событий в пространстве деформационного мониторинга

В данном разделе рассмотрен алгоритм создания единого линейного пространства деформационного и сейсмического мониторинга. В качестве такого используем евклидово пространство сейсмических событий, линейная оболочка которого образуется базисом векторов X , Y , Z и обозначим его L_{seism} . Каждый из этих векторов однозначно задает координату сейсмического события в пространстве. Выделим в пространстве L_{seism} подпространство деформационных событий (которое также является евклидовым), линейная оболочка которого образуется базисом векторов X_1 , Y_1 , Z_1 , и обозначим его как L_{deform} . Упрощенная схема выбора подпространства L_{deform} приведена на рисунке 2.3.1. Поскольку известны координаты наших датчиков деформационного мониторинга, в качестве начала координат L_{deform} используем среднее значение по каждой из координат 6 датчиков сейсмического мониторинга. Введем далее понятие евклидова расстояния используя формулу (2.3.1):

$$E = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} \quad (2.3.1)$$

Поскольку на данном этапе не можем делать предположений о структуре, нарушенности и прочих свойствах массива, использование других метрик (например, Манхеттоновское расстояние или расстояние Чебышева) не является целесообразным. Евклидова метрика, по сути, будет являться кратчайшим расстоянием для распространения энергии сейсмической волны от выбранной точки данного датчика. Таким образом, получен новый признак для признакового пространства деформационных событий. Проанализируем распределение значений, которые принимает этот признак, для выявления каких-либо паттернов в кривой плотности распределения.

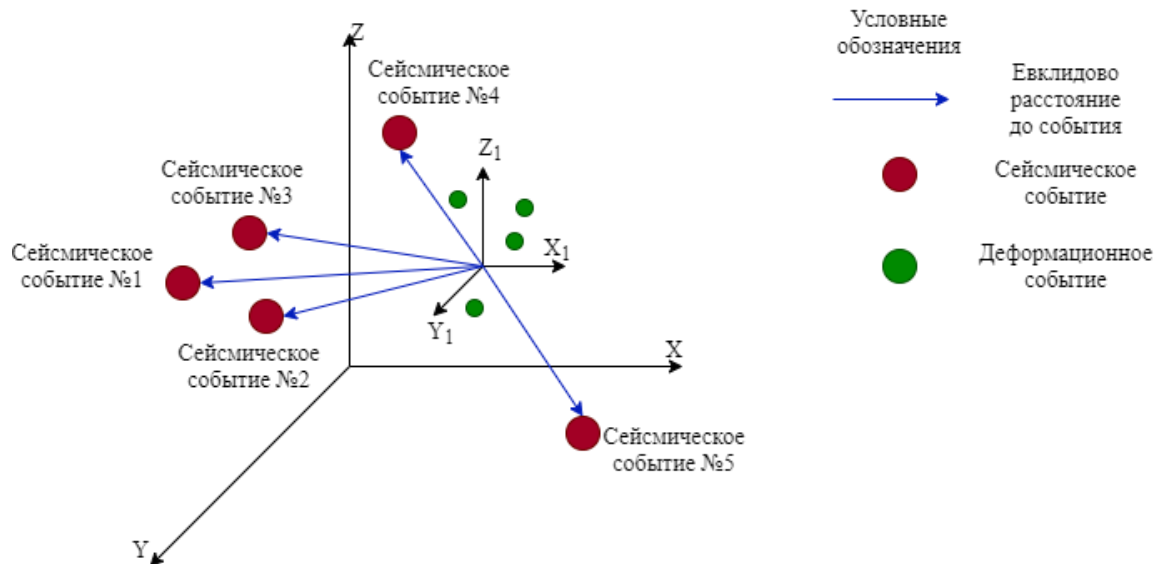


Рисунок 2.3.1 – Схема выбора подпространства деформационных событий

L_{deform} в пространстве сейсмических событий L_{seism}

На рисунках 2.3.2. и 2.3.3 по оси X отложено значение расстояния в метрах от условного «центра масс» деформационного мониторинга до каждого зарегистрированного сейсмического события, а по оси Y-количество событий в откалиброванных условных единицах (0,0005 = 5 событиям, 0,001 = 10 событиям, 0,0015 = 15 событиям).

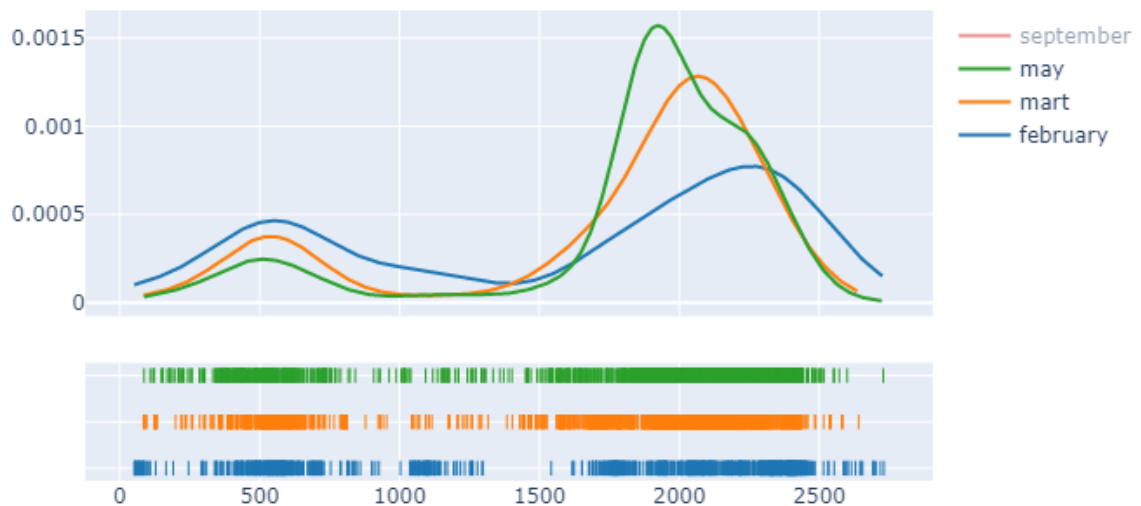


Рисунок 2.3.2 - Оценка распределения значения евклидовой метрики за период февраль-май 2018 года

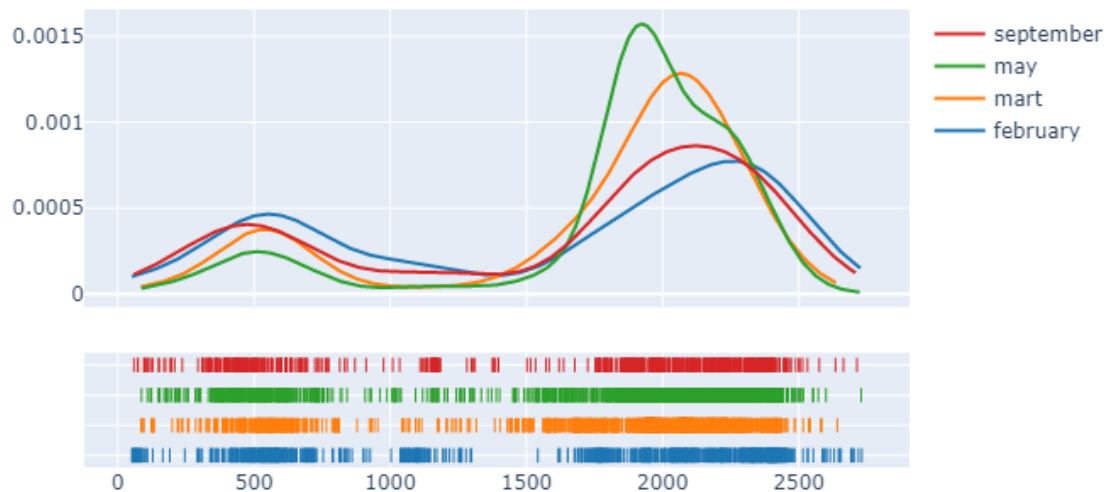


Рисунок 2.3.3 - Оценка распределения значения евклидовой метрики за период февраль-сентябрь 2018 года

Проанализируем кривые оценки распределения значения евклидовой метрики за период февраль-сентябрь 2018 года:

- Все распределения являются бимодальными и имеют две «четко выраженных» моды;
- В период с февраля по май 2018 года присутствуют рост и смещение диапазона от 1500 до 2500 м в сторону значения 1600–1700 метра;
- Распределения в диапазоне от 300 до 700 метров сохраняют свой тип и частоту наблюдений на протяжении всего периода мониторинга;
- В период с мая по сентябрь 2018 года присутствует существенное снижение (примерно в 2 раза) частоты значений в диапазоне от 1600 до 2500 м.

2.4 Анализ временного ряда сейсмического мониторинга с помощью метода SSA с целью обнаружения «предвестников» потенциально опасных геодинамических событий

В настоящее время для изучения свойств сложных систем, в том числе и при экспериментальных исследованиях, широко используется подход, основанный на анализе сигналов, произведенных системой. Это очень актуально в тех случаях, когда математически описать изучаемый процесс практически невозможно, но в нашем распоряжении имеется некоторая характерная наблюдаемая величина. Поэтому анализ систем, особенно при

экспериментальных исследованиях, часто реализуется посредством обработки регистрируемых сигналов. Например, в сейсмологии — запись колебаний земной коры, в метеорологии — данные метеонаблюдений и т. п. Обычно такой сигнал называется откликом, а метод исследования — реконструкцией динамических систем. Этот раздел теории динамических систем называется анализом временных рядов.

Однако следует отметить, что «временной ряд существенно отличается от простой выборки данных, так как при таком анализе учитывается взаимосвязь измерений со временем, а не только статистическое разнообразие и статистические характеристики выборки. Анализ временных рядов — это совокупность математико-статистических методов анализа, предназначенных для определения структуры ряда динамики и его прогнозирования в будущее» [2].

2.4.1. Методы обработки временных рядов

Несмотря на то, что нелинейные системы могут значительно отличаться в конкретных проявлениях и деталях, существуют глубокие аналогии в их организации и функционировании. Это предопределило интерес к методам, которые развиваются в рамках теории динамических систем, как универсальному инструменту исследования объектов самой различной природы. Наиболее явно такой объединяющий подход проявляется при анализе временных рядов. В настоящее время существует два качественно различных подхода к исследованию временных рядов:

- Статистические (вероятностные модели, авторегрессионные модели);
- Динамические (теорема Такенса о размерности вложения, теории информации, топологии, дифференциальной динамики, динамических систем).

При статистической обработке временного ряда используются авторегрессионные модели со скользящим средним (ARIMA). При

использовании таких моделей необходимо привести исходный ряд к стационарному виду, что можно осуществить с помощью операций вида:

- Удаление тренда;
- Применение статистических оценок;
- Фильтрация шума.

В случаях, требующих быстрого обновления прогноза на основе вновь поступивших данных, используются адаптивные методы прогноза. К ним относится, например, метод экспоненциального сглаживания (метод Брауна). Следуя ему, каждому значению ряда в процессе идентификации модели присваивается весовой коэффициент, экспоненциально убывающий во времени.

Динамические методы обработки временных рядов. Статистические методы не позволяют отличить конечномерный процесс от бесконечномерного, но это вполне можно сделать динамическими методами. Для многих систем конечномерное описание вполне допустимо. Тогда применение динамических моделей сводится к нахождению размерности вложения — минимального числа динамических переменных, однозначно описывающих поведение исследуемой системы.

В данной работе для анализа временных рядов сейсмического мониторинга применяется динамический метод обработки временных рядов Singular Spectrum Analysis.

Базовый вариант метода состоит в следующем: «преобразовании одномерного ряда в многомерный с помощью однопараметрической сдвиговой процедуры; исследовании полученной многомерной траектории с помощью анализа главных компонент (сингулярного разложения); восстановлении (аппроксимации) ряда по выбранным главным компонентам» [29].

Таким образом: «...результатом применения метода является разложение временного ряда на простые компоненты: медленные тренды, сезонные и другие периодические или колебательные составляющие, а также

шумовые компоненты. Полученное разложение может служить основой прогнозирования как самого ряда, так и его отдельных составляющих. SSA метод допускает естественное обобщение на многомерные временные ряды» [30].

В данной работе рассмотрен вариант алгоритма, предназначенный для анализа одномерного временного ряда. Для анализа временного ряда выбирается параметр L .

Отметим что: «этот параметр может выбираться достаточно произвольно. При достаточно большой длине временного ряда и достаточно большом L результаты не зависят от длины окна» [31].

Первым этапом метода является сопоставление временного ряда с последовательностью многомерных векторов-задержек. Выбирается такое целое число L (длина окна) что, $2 \leq L \leq N/2$, где N -длина временного ряда. Таким образом, получаем следующий набор векторов (2.4.1):

$$\begin{aligned} X_0 &= (x_i, \dots, x_{L-1})^T \\ X_1 &= (x_{i+1}, \dots, x_L)^T \\ X_2 &= (x_{i+2}, \dots, x_{L+2})^T \\ X_3 &= (x_{i+3}, \dots, x_{L+3})^T \end{aligned} \quad (2.4.1)$$

Эти вектора образуют траекторную матрицу X (2.4.2) исходного временного ряда:

$$X = \begin{bmatrix} f_0 & \cdots & f_{N-L} \\ \vdots & \ddots & \vdots \\ f_{L-1} & \cdots & f_{N-1} \end{bmatrix} \quad (2.4.2)$$

Из приведенной выше матрицы видно, что элементы антидиагоналей (то есть диагонали, идущие снизу слева направо) равны. Этот тип матрицы известен как матрица Ганкеля. Положим $K = N-L+1$ и введем следующие обозначения:

- Столбцы матрицы X - вектора с L -задержкой;
- Строки матрицы X - вектора с K -задержкой

2.4.2. SVD разложение

Следующим этапом алгоритма является разложение матрицы X с помощью сингулярного разложения (SVD), приведенного в формуле (2.4.3):

$$X = U\Sigma V^T, \quad (2.4.3)$$

где

- U - унитарная матрица размера $L \times L$, содержащая ортонормированный базис левых сингулярных векторов X ;
- Σ - прямоугольная диагональная матрица размера $L \times K$, содержащая L сингулярных значений траекторной матрицы X , упорядоченных от наибольшего к наименьшему;
- V^T - унитарная матрица размера $K \times K$, содержащая ортонормированный базис правых сингулярных векторов X

SVD разложение может быть переписано в следующей форме (2.4.4):

$$X = \sum_{i=0}^{d-1} \sigma_i U_i V_i^T \quad (2.4.4)$$

где σ_i - i -е сингулярное значение, U_i и V_i - векторы, представляющие i -ые столбцы матриц U и V соответственно, $d \leq L$ - ранг траекторной матрицы, а $X_i = \sigma_i U_i V_i^T$ - i -ая элементарная матрица X . Набор $\{U_i, \sigma_i, V_i\}$ будет обозначаться i -ой собственной тройкой. Для более детального анализа элементов разложения матрицы X , рассмотрим по очереди рассмотрим матрицы U , V , Σ .

2.4.3. Матрица U

U - матрица размера $L \times L$, столбцы которой ортонормированы (2.4.5), то есть:

$$U_i \cdot U_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (2.4.5)$$

Это означает, что $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{1}$, что делает U унитарной матрицей. Чтобы прояснить роль, которую U играет в разложении матрицы X (2.4.6), обозначим $Z_i = \sigma_i V$ - вектор столбцов, такой что

$$\mathbf{X} = \sum_{i=0}^{d-1} U_i Z_i^T, \quad (2.4.6)$$

и каждый вектор-столбец с L -задержкой (обозначим его X_j), задается формулой (2.4.7)

$$X_j = \sum_{i=0}^{d-1} z_{j,i} U_i, \quad (2.4.7)$$

Выражение для X_j предполагает, что $U = \{U_0, \dots, U_{d-1}\}$ является базисом для пространства, которое образуют столбцы траекторной матрицы, а $z_{i,j}$ - это i -ый коэффициент вектора-задержек X_j , представленного в базисе U .

Другими словами, столбцы матрицы U образуют ортонормированный базис, который описывает временные подмножества $\{f_i, \dots, f_{i+L-1}\}_{i=0}^{N-L}$ в пространстве, которое образованно столбцами траекторной матрицы.

2.4.4. Матрица V

Матрица V , представляет собой матрицу размера $K \times K$ с ортонормированными столбцами (как и матрица U), что делает ее унитарной.

Чтобы проинтерпретировать столбцы матрицы V полученные в ходе SVD разложения траекторной матрицы, отметим, что для любых матриц A B $(AB)^T = B^T A^T$. Таким образом:

$$X^T = V \Sigma^T U^T \quad (2.4.8)$$

в котором имеется набор $Y_i = \sigma_i U_i$, такой что:

$$X_j^{(T)} = \sum_{i=0}^{d-1} y_{j,i} V_i, \quad (2.4.9)$$

где X_j^T , - j-ый столбец X^T , а $y_{j,i}$ - j-ая составляющая вектора V_i . Используя формулы (2.4.8) и (2.4.9) можно сделать предположение, что $V = \{V_0, \dots, V_{d-1}\}$ является базисом пространства, которое образуют столбцы матрицы X^T , а $y_{j,i}$ - i-ый коэффициент запаздывающего вектора X_j^T , в базисе V . Таким образом, столбцы матрицы V образуют ортонормированный базис, который описывает временные подмножества $\{f_i, \dots, f_{i+N-L}\}_{i=0}^{L-1}$ в пространстве, которое образовано строками траекторной матрицы.

2.4.5. Матрица Σ

Матрица Σ представляет собой прямоугольную диагональную матрицу размера $L \times K$, содержащую сингулярные значения матрицы X . Сингулярные значения упорядочены от наибольшего к наименьшему. Эти значения могут быть интерпретированы как коэффициенты масштабирования, которые определяют относительную важность собственной тройки $\{U_i, \sigma_i, V_i\}$ в траекторной матрице $X = \sum_{i=0}^{d-1} \sigma_i U_i V_i^T$.

Обозначим норму Фробениуса матрицы X следующим образом (2.4.10):

$$\mathbf{X}_F = \sqrt{\sum_{j=0}^{L-1} \sum_{k=0}^{K-1} x_{j,k}^2}, \quad (2.4.10)$$

где $x_{j,k}$ обозначает элемент в j-ой строке и k-ом столбце матрицы X .

Рассмотрим элементарные матрицы $\mathbf{X}_i = \sigma_i U_i V_i^T$. Из-за нормализации U_i и V_i их произведение $\|U_i V_i^T\|_F = \|U_i\|_F \|V_i\|_F$ равно 1. Таким образом, $\|\mathbf{X}_i\|_F = \sigma_i$.

Также оказывается, что

$$\mathbf{X}^2 = \sum_{i=0}^{d-1} \sigma_i^2 \quad (2.4.11)$$

Таким образом, квадрат нормы Фробениуса траекторной матрицы равен сумме квадратов сингулярных значений (2.4.11). Это говорит о том, что допустимо принять соотношение $\sigma_i^2 / \|\mathbf{X}\|_F^2$ в качестве меры вклада,

который вносит элементарная матрица X_i . Домножим исходную матрицу X на транспонированную ей.

$$XX^T = U\Sigma V^T X^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T \quad (2.4.12)$$

Домножим левую часть на матрицу U (2.4.13) .

$$(XX^T)U = U\Sigma^2 \quad (2.4.13)$$

Принимая во внимание, что Σ^2 является диагональной матрицей с элементами σ_i^2 , столбцы U являются собственными векторами матрицы XX^T (2.4.12) с собственными значениями $\{\sigma_0^2, \dots, \sigma_{L-1}^2\}$, можно сделать вывод что умножение матрицы X на X^T , доказывает, что столбцы матрицы V являются собственными векторами матрицы $X^T X$, также с собственными значениями $\{\sigma_0^2, \dots, \sigma_{L-1}^2\}$.

2.4.6. Ранг траекторной матрицы

Столбцы траекторной матрицы представляют собой последовательность векторов с L -задержками, которые образуют фазовое пространство временного ряда. Размерность этого пространства как минимум равна L однако, если столбцы в X имеют линейную зависимость, то размерность этого пространства не будет превышать L . Это проявляется в виде одного или нескольких нулевых сингулярных значений в Σ . Ранг X — это максимальное значение i , такое, что $\sigma_i > 0$. Иными словами, ранг матрицы X , $d = \text{rank}\{X\}$, можно рассматривать как значение размерности фазового пространства исходного временного ряда.

Важно отметить, что каждая элементарная матрица X_i имеет ранг равный 1, и что данная матрица может быть записана в общем виде следующим образом.

$$X^{(r)} = \sum_{i=0}^r X_i, \quad (2.4.14)$$

для $r < d$ - наилучшее приближение ранга r траекторной матрицы X , такое что, $\|X - X^{(r)}\|_F$ сведено к минимуму. Таким образом, просуммировав

первые r элементарных матриц, получается оптимальное приближение матрицы X с меньшей размерностью (2.4.14).

2.4.7. Разделение и группировка компонентов временных рядов

После получения набора собственных троек $\{U_i, \sigma_i, V_i\}$ временного ряда, актуальным становится вопрос об их группировке между собой. Наиболее простой и очевидный вариант – группировка с помощью визуального осмотра, критерием сходства здесь является внешний вид компоненты. Этот метод хорошо подходит для коротких ($N < 10^4$) и простых временных рядов, однако для более длинных и сложных временных рядов требуется метод, который количественно определяет, можно ли считать реконструированный компонент F_i отличным от другого компонента F_j . Для двух восстановленных компонент F_i и F_j исходного временного ряда с длиной N и длиной окна L , определим взвешенное скалярное произведение (F_i, F_j) как:

$$(\tilde{F}_i, \tilde{F}_j)_w = \sum_{k=0}^{N-1} w_k \tilde{f}_{i,k} \tilde{f}_{j,k}, \quad (2.4.15)$$

где w_k определяется следующим образом (2.4.16)

$$w_k = \begin{cases} k+1 & 0 \leq k \leq L-1 \\ L & L \leq k \leq K-1 \\ N-k & K \leq k \leq N-1 \end{cases} \quad (2.4.16)$$

Учитывая, что $K=N-L+1$, вес w_k показывает количество раз, когда $f_{i,k}$ и $f_{j,k}$ появляются в матрицах X_i и X_j из которых были получены временные ряды F_i и F_j .

Итак, если выражение 2.4.15 равно нулю, то F_i и F_j являются ортогональными друг другу и компоненты временного ряда разделимы между собой. Определим взвешенную корреляционную матрицу W_{corr} размера $d \times d$, которая описывает корреляцию F_i и F_j . Элементы W_{corr} задаются

$$W_{i,j} = \frac{(\tilde{F}_i, \tilde{F}_j)_w}{\|\tilde{F}_i\|_w \|\tilde{F}_j\|_w} \quad (2.4.17)$$

Интерпретация $W_{i,j}$ (2.4.17) довольно проста. Если компоненты близки друг к другу (но не идентичны), то $(\tilde{F}_i, \tilde{F}_j)_w \rightarrow \|\tilde{F}_i\|_w \|\tilde{F}_j\|_w$, и следовательно $W_{i,j} \rightarrow 1$. Умеренные значения $W_{i,j}$ между 0 и 1 (допустим $W_{i,j} \geq 0,3$), указывают на компоненты, которые, возможно, потребуется сгруппировать вместе.

Таким образом весь алгоритм можно описать следующими 3 этапами.

Вложение. Строится траекторная матрица размерности $L \times K$ ряда X следующим образом:

$$X = [x_i : \dots : x_N] = \begin{pmatrix} x_1 & \cdots & x_n \\ \vdots & \ddots & \vdots \\ x_L & \cdots & x_N \end{pmatrix}, \quad (2.4.18)$$

где $X = (x_i, \dots, x_{i+L-1})$ — векторы вложения длины L ($1 \leq i \leq N$).

Матрица X (2.4.18) является ганкелевой, то есть имеет одинаковые элементы на антидиагоналях ($i+j=const$).

Сингулярное разложение (SVD). Далее производится сингулярное разложение (SVD) траекторной матрицы X . Положим $S = XX^T$ и введем обозначения:

- собственные числа $\lambda_1, \dots, \lambda_L$ матрицы S , взятые в невозрастающем порядке
- система собственных векторов $U_1, \dots, U_L$ матрицы S , соответствующие собственным числам.

Данные векторы называются левыми и правыми, соответственно, сингулярными векторами матрицы и по аналогии с анализом главных векторов, называются векторами главных компонент, числа — сингулярные числа (они составляют сингулярный спектр, что и дало название Singular Spectrum Analysis методу).

Диагональное усреднение. Наиболее точно процедура диагонального усреднения и последующей группировки собственных троек, описана в следующих работах [29-31]. Рассмотрим этот этап более подробно: «Каждая матрица сгруппированного разложения ганкелизуется (усредняется по антидиагоналям), а затем полученная ганкелева матрица трансформируется в новый временной ряд длины n на основе взаимно однозначного соответствия между ганкелевыми матрицами и временными рядами. Диагональное усреднение, применённое к каждой результирующей матрице, создает восстановленные временные ряды. Таким образом, исходный ряд раскладывается в сумму восстановленных рядов. Данное разложение является главным результатом алгоритма SSA для анализа временного ряда. Это разложение имеет смысл если каждая из его компонент может быть интерпретируема как либо тренд, либо колебания (периодики), либо шум. Предположим, что исходный ряд является суммой нескольких рядов. Теоретические результаты позволяют определить по виду собственных чисел, собственных и факторных векторов, что это за слагаемые и каким элементарным матрицам они соответствуют. Суммируя элементарные матрицы внутри каждого набора и переходя от результирующих матриц к ряду, мы получаем разбиение исходного ряда на аддитивные составляющие, например, на сумму тренда, цикла и шума или сумму низкочастотной и высокочастотной составляющих. Возможность разбить совокупность элементарных матриц на группы, соответствующие интерпретируемым аддитивным составляющим тесно связана с понятием разделимости ряда».

Таким образом, целью данного метода является разложение исходного ряда в сумму аддитивных компонент. При этом метод не требует стационарности ряда, знания тренда, а также знания о наличии в ряде сезонных и циклических составляющих. В итоге данный метод может применяться для обнаружения тренда, циклики, сглаживания ряда, построение полного разложенного ряда в сумму тренда, циклики и шума.

Для начала выпишем координаты и даты всех документально зафиксированных опасных геодинамических проявлений в таблице 3.2.1.1.

Таблица 3.2.1.1. Зафиксированные опасные геодинамические проявления за весь период сейсмического мониторинга

Координата сейсмического события по оси X	Координата сейсмического события по оси Y	Координата сейсмического события по оси Z	Дата сейсмического события
39491,00	31290,00	250,00	13.05.2009
38750,00	31287,00	170,00	21.10.2010
39001,00	31887,00	235,00	27.01.2016

Для первых двух явлений использован временной ряд длиной 365 дней, ширину окна равную 30 дням (параметр L), а в качестве единиц отсчета среднее значение всех сейсмических событий за день.

Так как была выбрана длина окна, равная 30 дням, поэтому получаем 30 элементарных матриц, которые описываются собственными значениями матриц, оценивая которые можно предположить их вклад в исходный ряд.

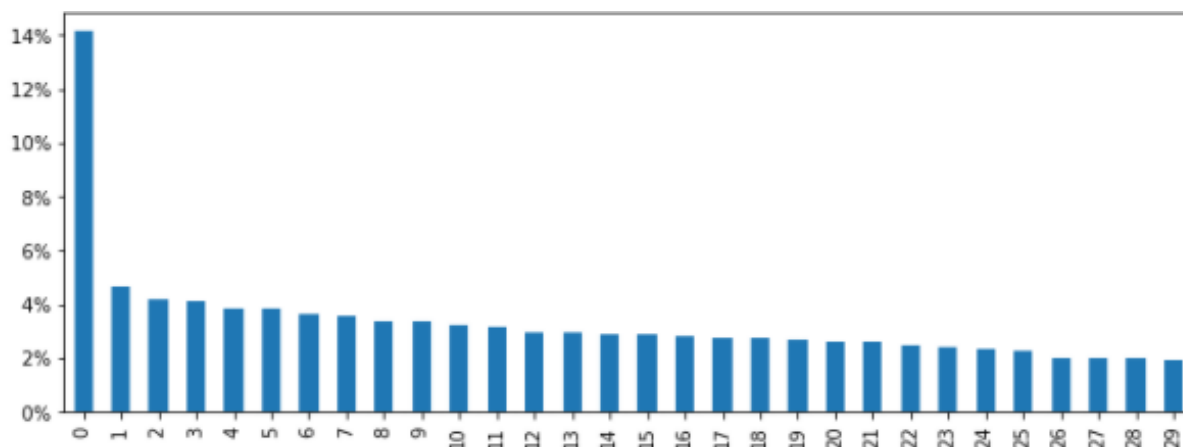


Рисунок 2.4.2 – Вклад каждой элементарной матрицы в исходный временной ряд в процентах

Как видно из графика 2.4.2. около 14% от всего ряда дает 1 элементарная матрица, а остальные 29 матриц дают вклад от 4 до 1%. Задача исследования найти в полученных 30 элементарных матрицах группы,

которые отвечают за цикличность процесса, за шумовую составляющую и за тренд.

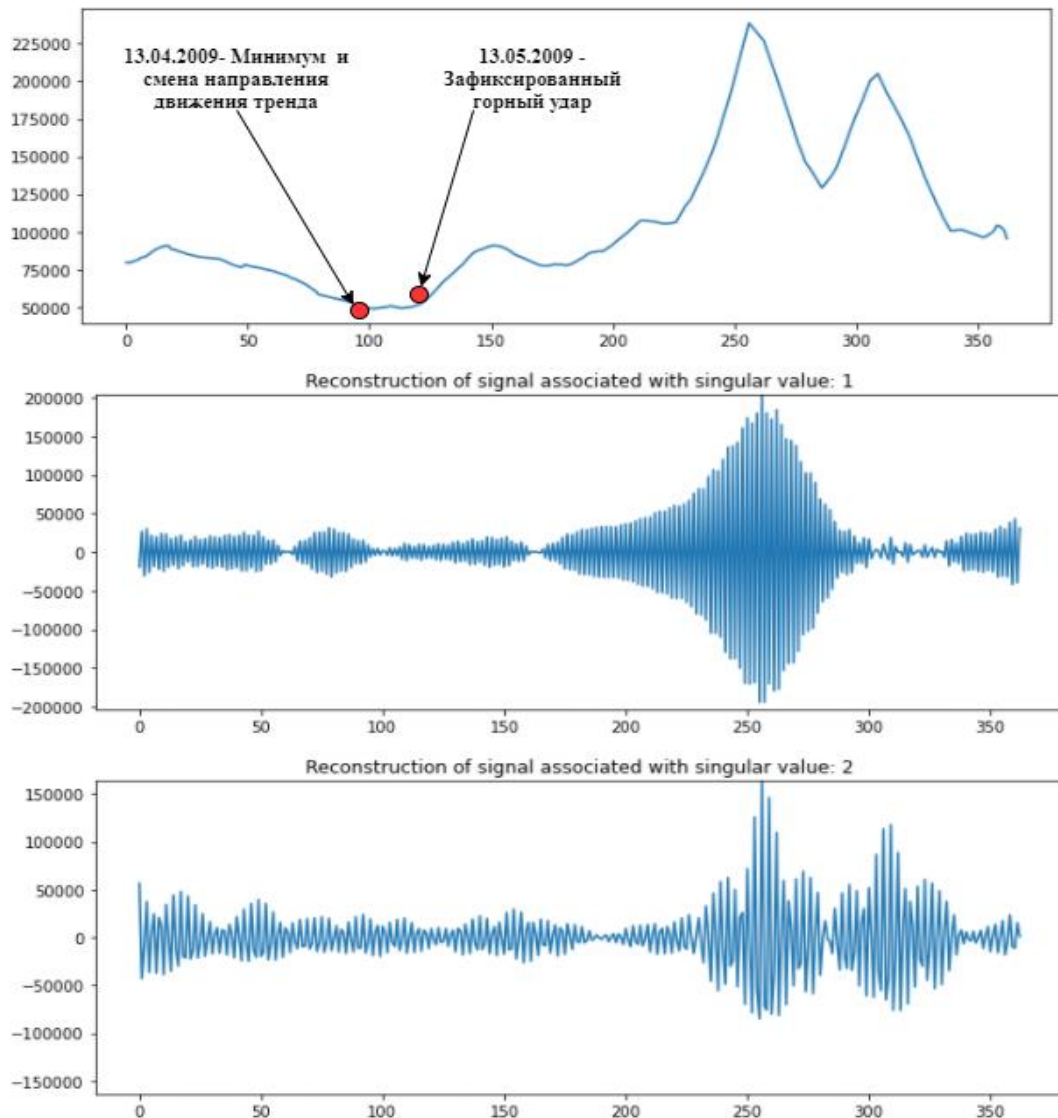


Рисунок 2.4.3 – Первая, вторая и третья элементарные матрицы, отвечающие соответственно за тренд и шум временного ряда сейсмического мониторинга за 2009 год

Первая элементарная матрица отвечает за тренд среднего значения сейсмической активности. Как видно из графика 2.4.3 на 100 единице отсчета (13 апреля 2009 года) достигается минимальное значение тренда сейсмической активности за весь период наблюдений. Данное событие не сопровождается какими-либо аномалиями в других временных рядах, поэтому скорее всего для его идентификации требуются дополнительные сведения о состоянии МГП в данный период. Ключевой компонентой в

исходному ряду является тренд (12% всей информации). Анализ тренда не позволяет явным образом установить связь между опасным геодинамическим явлением и сейсмической активностью. После наступления данного события не наблюдается ни изменений в росте и направления тренда, ни аномальных отклонений в компонентах, отвечающих за цикличность и шум в исходном ряду. Таким образом опасному геодинамическому явлению может предшествовать так называемый период «затишья» (плавное снижение тренда) и изменения в направлении тренда процесса.

Анализ данных сейсмического мониторинга за 2010 год

Как видно из графика 2.4.4 около 30% от всего ряда дает 1 элементарная матрица, отвечающая за тренд, остальные матрицы дают от 5 до 1% всей информации.

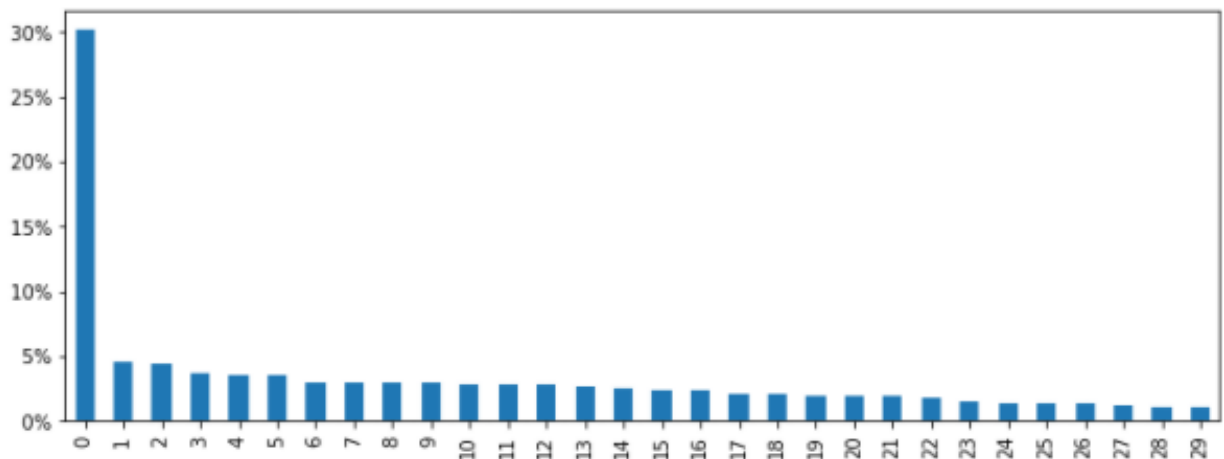


Рисунок 2.4.4 – Вклад каждой элементарной матрицы в исходный временной ряд в процентах

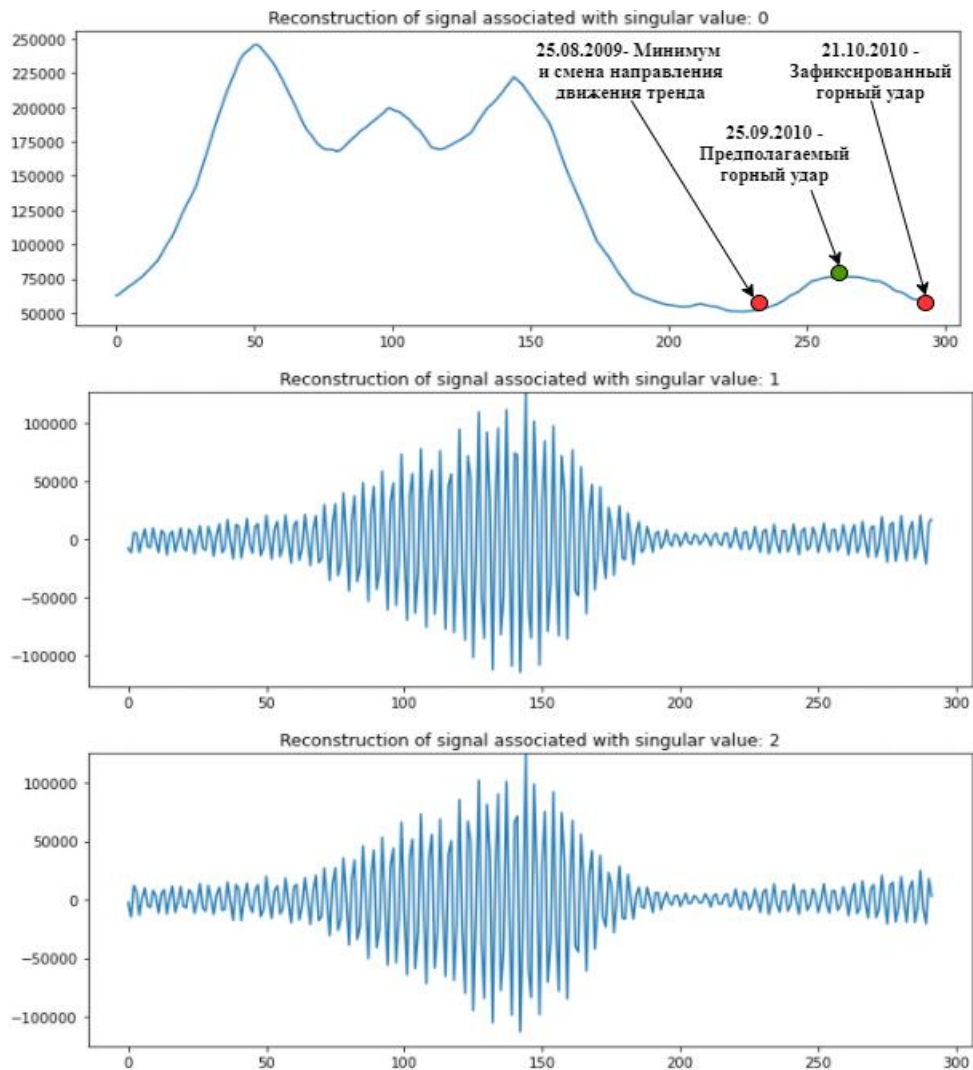


Рисунок 2.4.5 – Первая, вторая и третья элементарные матрицы, отвечающие соответственно за тренд и шум временного ряда сейсмического мониторинга за 2010 год

Как видно из графика 2.4.5 на 240 единице отсчета (25 августа 2010 года) достигается минимальное значение тренда сейсмической активности за весь период наблюдений. Далее, следуя предположению о разнице в 30 дней между проявлением опасных геодинамических событий и сменой тренда сейсмической активности, было предсказана дата возможного горного удара 25 сентября 2010 года. Однако, в данную дату не было зафиксировано никаких горных ударов или иных опасных событий. Тем не менее следует заметить, что после этой даты начинается очередная смена тренда сейсмической активности. Ключевой компонентой в исходному ряду по-прежнему является тренд (30% всей информации). Анализ тренда позволяет

явным образом установить связь между опасным геодинамическим явлением и сейсмической активностью. После наступления данного события наблюдается изменение в направлении тренда (нисходящий тренд), а также когерентность изменений в других компонентах, отвечающих за цикличность процесса и шум. Таким образом, можно сделать предположение, что точкой отсчета нового цикла до наступления возможного горного удара является 25.09.2010.

Анализ данных сейсмического мониторинга за 2015 и 2016 год

Отметим, что событие, зафиксированное 27.01.2016, следует рассматривать в рамках временного ряда 2015 и 2016 года, т. к. нас интересуют предшествующие данному явлению «предвестники».

Как видно из графика 2.4.6 около 25% информации дает первая элементарная матрица, а следующие 29 матриц дают вклад от 3 до 1%. Используя для восстановления исходного ряда первые 30 матриц, можно сохранить до ~83% процентов полезной информации от всего ряда.

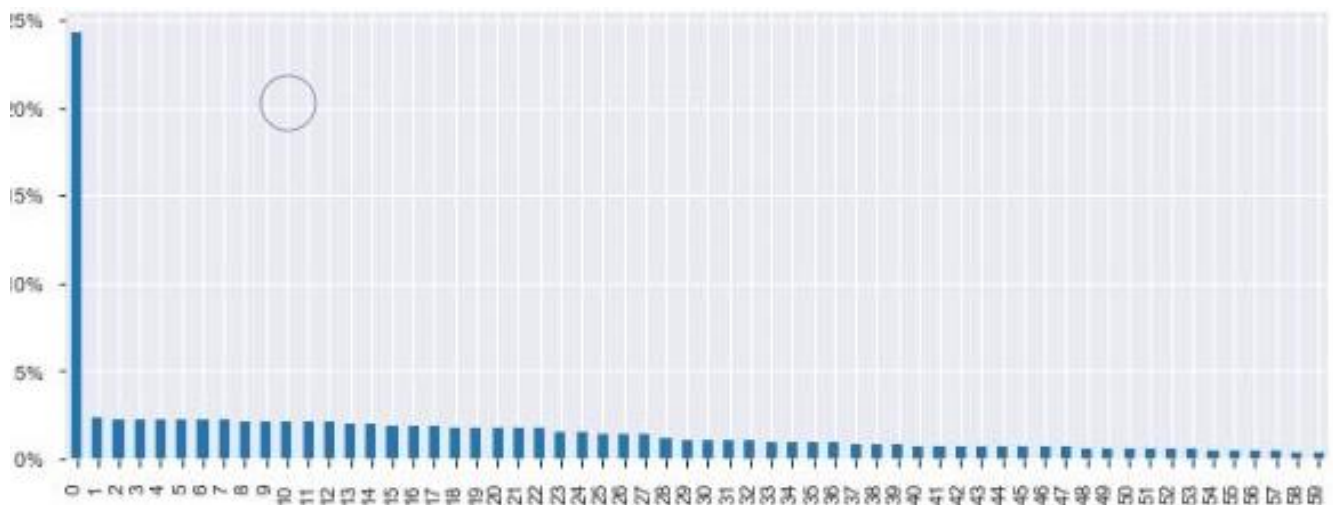


Рисунок 2.4.6 – Вклад каждой элементарной матрицы в исходный временной ряд в процентах

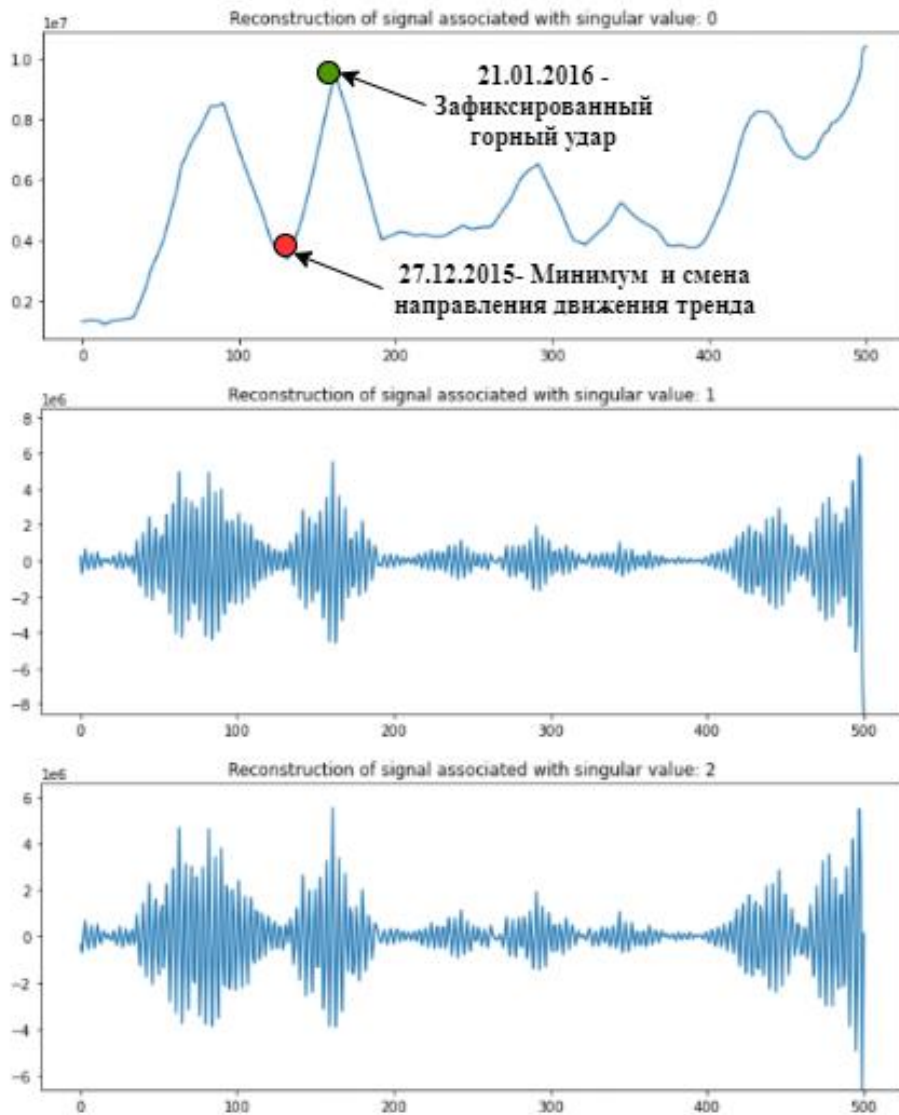


Рисунок 2.4.7 – Первая, вторая и третья элементарные матрицы, отвечающие соответственно за тренд и шум временного ряда сейсмического мониторинга за 2015–2016 год

График 2.4.7 представляет собой сумму первого и второго случаев (2009 и 2010 года). С одной стороны, есть повтор ситуации 2009 года (падение до минимального значения тренда сейсмической активности, и точка смены направления роста тренда) на точке отсчета 220. С другой стороны, в период с 220 по 310 (с 10 августа по 10 ноября 2015 года) точку отсчета, где наблюдается резкий рост тренда, и точка конца роста в 3 раза больше точки начала роста (эвристика «трех сигм»). Ключевой компонентой в исходному ряду является тренд (25% всей информации). Анализ тренда позволяет явным образом установить связь между опасным геодинамическим

явлением и сейсмической активностью. После наступления данного события наблюдается изменение в направлении тренда (нисходящий тренд снижение уровня значения, предшествующего релизу напряжений в МГП в виде горного удара от 27.01.2016).

Анализ данных сейсмического мониторинга за 2015 и 2018 год

Целью дальнейшего исследования является поиск предполагаемых «предвестников» опасных геодинамических событий для отслеживания их связи с данными датчиков деформационного мониторинга. Как видно из графика 2.4.8, около 45% информации дает первая элементарная матрица, а следующие 29 матриц дают от 3 до 1%. Используя для восстановления исходного ряда первые 30 матриц, можно сохранить до приблизительно 97% процентов полезной информации от всего ряда.

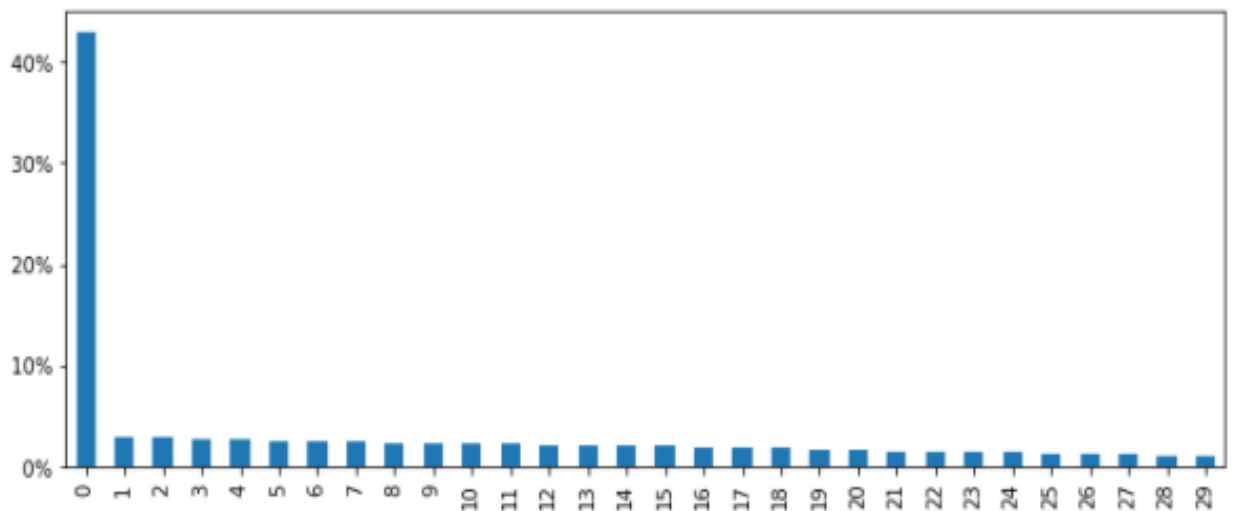


Рисунок 2.4.8 – Вклад каждой элементарной матрицы в исходный временной ряд в процентах

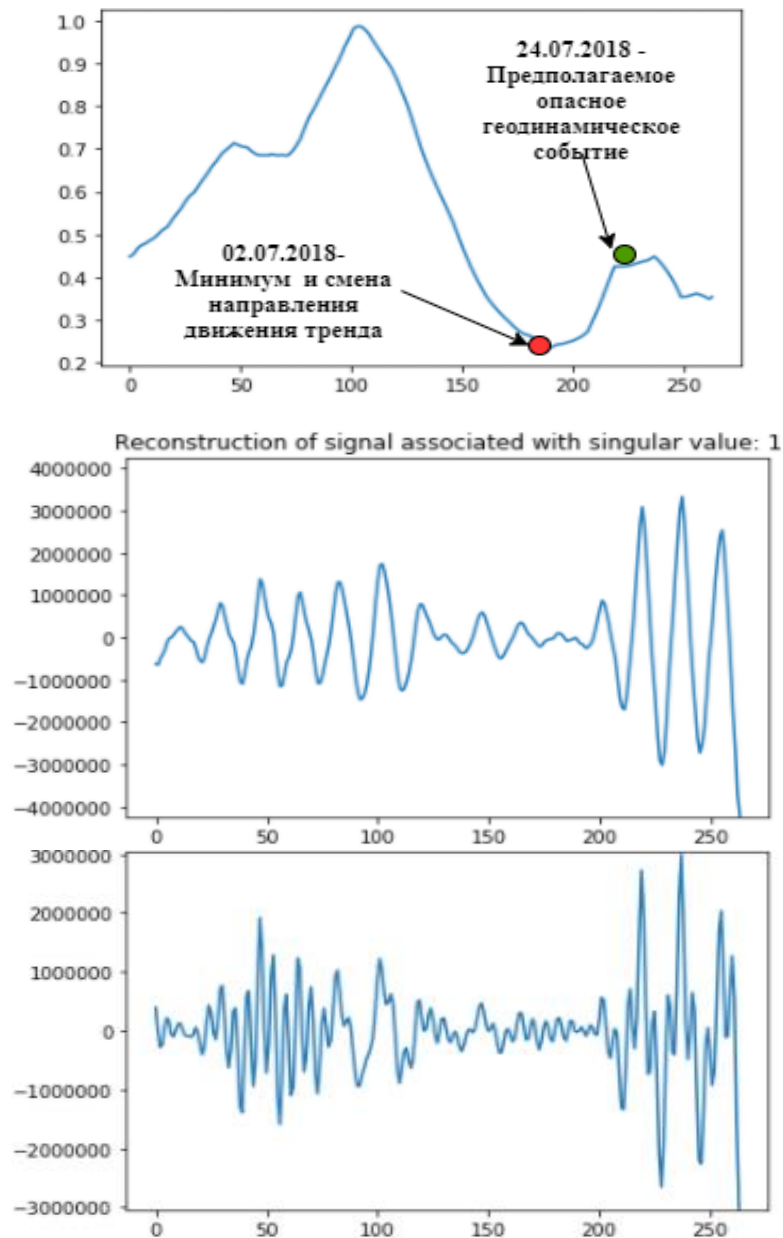


Рисунок 2.4.9 – Первая, вторая и третья элементарные матрицы, отвечающие соответственно за тренд и шум временного ряда сейсмического мониторинга за 2018 год

График тренда (первая элементарная матрица) представляет собой отмасштабированную версию паттерна графика тренда 2015–2016 года. С одной стороны, есть падение значения сейсмической активности до минимального значения тренда сейсмической активности и точка смены направления роста тренда на точке отсчета 180 (02.07.2018). С другой стороны, в период с 185 по 208 точку отсчета (с 2 июля по 24 июля 2018

года), наблюдается резкий рост тренда, и точка конца роста в 3 раза больше точки начала роста (эвристика «трех сигм»).

Ключевой компонентой в нашем исходном ряду является тренд (45% всей информации). Анализ тренда позволяет явным образом установить связь между опасным геодинамическим явлением и сейсмической активностью. После наступления данного события наблюдается изменение в направлении тренда (нисходящий тренд снижение уровня почти в 2 раза), что связано с релизом напряжений в МГП в виде горного удара от 24.07.2018.

2.5 Кластерный анализ сейсмических событий в признаковом пространстве деформационного мониторинга

Задачей кластерного анализа является применение методов кластеризации к данным сейсмического мониторинга. Применение таких методов поможет детектировать эпицентры сейсмических событий и отслеживать во времени этапы формирования потенциально опасных зон проявления горного давления в массиве горных пород. Большая часть подходов к оценке и прогнозу проявлений опасного горного давления выделяет несколько стадий разрушения горной породы, причем эти стадии могут переходить друг в друга.

Так, например: «...изучение сейсмического режима Кукисвумчоррского месторождения за длительный период наблюдений (2008–2016 гг.) показало, что кластеры сейсмических событий приурочены к местам активного ведения горных работ, к разрывным нарушениям, к формирующимся разрывным нарушениям в консоли пород висячего бока, а также могут образовываться под влиянием других факторов как природных, так и техногенных.» [6].

Можно выделить два основных подхода при выборе алгоритмов кластеризации. Либо: «...на представлениях о физических процессах, происходящих в процессе нагружения и последующей деформации горного массива» [5], использовать детерминированные модели кластеризации (иерархического или графовый модели). Такие модели разбивают исходное

множество сейсмических событий на несколько непересекающихся множеств. Такой подход относится к алгоритмам «четкой» кластеризации. Его использование позволяет довольно хорошо идентифицировать структуру МГП с помощью кластеризации сейсмических событий.

Однако такой подход не учитывает случайные процессы и их стохастический характер [21]. Поэтому для идентификации местоположения и момента перехода потенциально опасной зоны проявления горного давления от одной стадии разрушения к другой целесообразно использовать вероятностные методы кластеризации. Такие методы могут относить выбранное сейсмическое событие к одному и более кластерам с разными оценками вероятности принадлежности.

В настоящей работе обосновано было принято решение применить алгоритм HDBSCAN (Hierarchical Density Based Spatial Clustering of Applications with Noise). Как следует из названия, данный алгоритм является иерархическим пространственным алгоритмом кластеризации данных с шумом на основе использования плотности распределения [22-27].

Необходима очень «недорогая», с точки зрения вычислительного процесса, оценка плотности, и самая простая — это расстояние до k -го ближайшего соседа в n -мерном пространстве. Формируя, таким образом, матрицу расстояний для наших данных, достаточно просто получить ее, если выбранная метрика поддерживается алгоритмом и размерность признакового пространства низкая, при использовании *kernel density trees*. Это алгоритм классификации, относящий исследуемый объект к определенному классу (кластеру) и использующий алгоритм решающих деревьев, где предикатами являются значения плотности распределения. Расстояние, определенное для параметра k и точки x , в работе обозначим как $core_k(x)$. Далее необходимо найти способ разделения точек с низкой плотностью (соответственно, с высоким расстоянием между ядрами). Самый простой способ — это ввести метрику нового расстояния между точками, называемую расстоянием

взаимной достижимости, то есть расстояние взаимной достижимости определяется следующим образом (2.5.1):

$$\text{dmreach} - k(a, b) = \max\{\text{corek}(a), \text{corek}(b), d(a, b)\} \quad (2.5.1)$$

где $d(a, b)$ — исходное метрическое расстояние между точками a и b . В соответствии с этой метрикой плотные точки (с низким расстоянием между ядрами) остаются на одном и том же расстоянии друг от друга, а более отдаленные точки отодвигаются так, чтобы быть на расстоянии, равным *core distance* от любой другой точки.

На примере, приведенном на рисунке 2.5.1, используется $k=5$. Для данной точки можно изобразить круг, радиусом которого будет являться величина *core distance*.

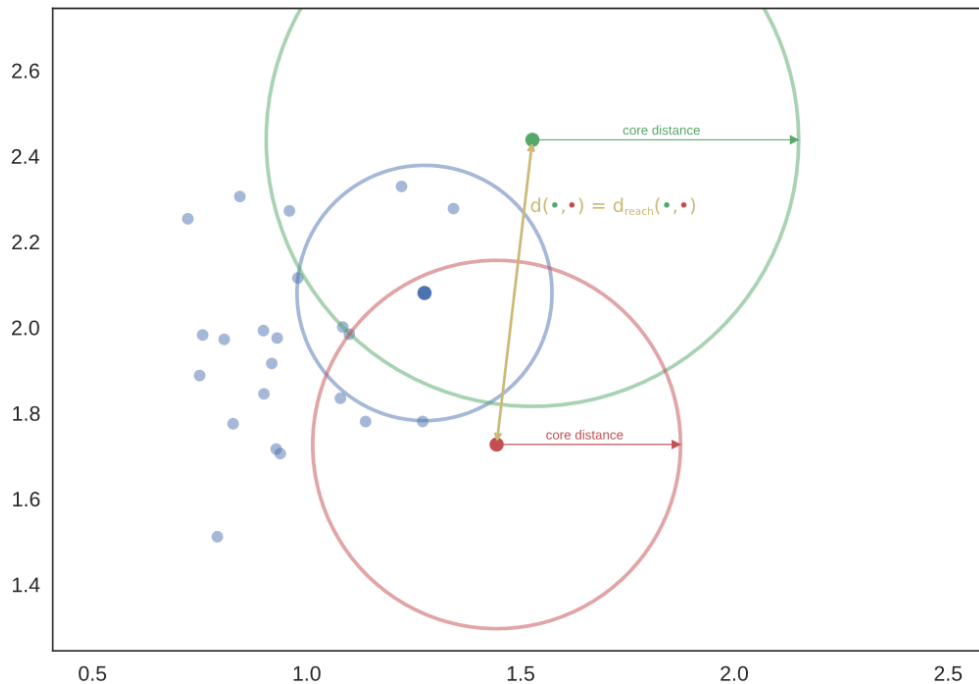


Рисунок 2.5.1 - Пример работы алгоритма HDBSCAN для $k=5$

Расстояние взаимной достижимости как преобразование хорошо работает, позволяя кластеризации с одной связью более точно аппроксимировать иерархию наборов уровней любого истинного распределения плотности, из которого были выбраны точки.

Таким образом, есть новая метрика взаимной достижимости для данных. Очевидно, что области «плотного» распределения относительно, и разные кластеры могут иметь разную плотность. Концептуально — это

приводит к следующему: рассмотрим данные как взвешенный граф с точками данных в виде вершин и ребром между любыми двумя точками с весом, равным расстоянию взаимной достижимости этих точек.

Теперь рассмотрим пороговое значение, начинающееся высоко, и неуклонно понижающееся. Отбрасывая любые ребра с весом выше этого порога, по мере удаления ребер разъединяем граф на связанные компоненты. В конечном итоге получим иерархию связанных компонентов (от полностью связанных до полностью изолированных) на различных пороговых уровнях.

Следовательно, надо найти минимальный набор ребер таким образом, чтобы удаление любого ребра из набора приводило к разъединению компонентов. В этом случае необходимо, чтобы этот набор был таким, чтобы не было ребра с низким значением веса, который мог бы соединять компоненты. Теория графов предоставляет такую возможность: минимальное связующее дерево графа. На рисунке 2.5.2 показано минимальное связующее дерево для расстояния взаимной достижимости.

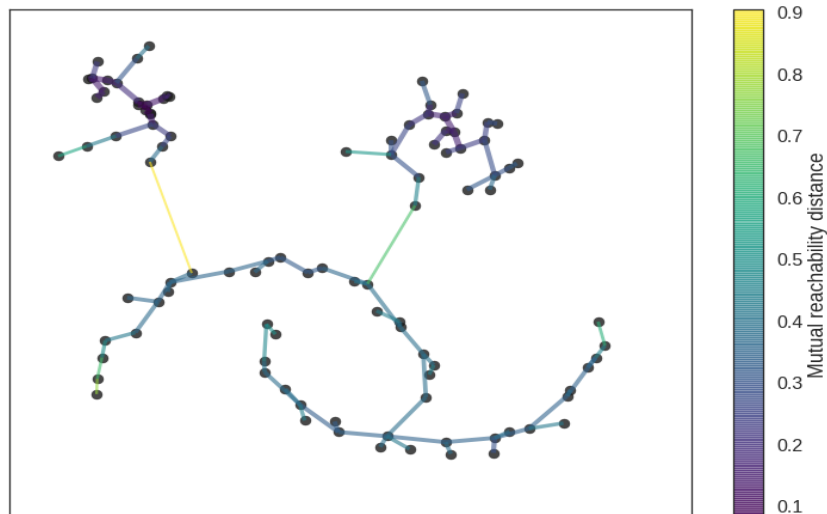


Рисунок 2.5.2 - Минимальное связующее дерево для расстояния взаимной достижимости

Объявляя все конечные узлы выбранными кластерами, сделаем далее обход дерева с помощью обратного топологического порядка сортировки. Если сумма взаимных расстояний достижимости дочерних кластеров больше, чем взаимное расстояние достижимости кластера, то взаимное расстояние достижимости кластера считается равным сумме дочерних взаимных

расстояний достижимости. В обратном случае кластер объявляется конечным и снимается выделение со всех его потомков. Как только достигается корневой узел, возвращается текущий набор выбранных кластеров.

Итак, для того чтобы определить форму и количество кластеров необходимо знать следующие две величины:

- *min cluster size*. Это минимальный размер нашего кластера. Данная величина определяет размер кластера, используя величину *core distance* [40]. Физическая интерпретация данного параметра — радиус распространения сейсмической волны от центра очага до его границы;

- *min points*. Это минимальное число точек в нашем кластере. Данная величина описывает сколько событий в нашем признаковом пространстве достаточно для того, чтобы отнести их к одному кластеру. Физическая интерпретация данного параметра — разделение «шумовых» точек от границ очага сейсмического события. Количество точек определяется либо эвристикой ($n=2*dim$, где *dim* количество признаков), либо с помощью настройки гиперпараметров алгоритма [41].

Далее отслеживается появление новых точек по месяцам с целью определения: есть ли временная компонента в появлении кластеров «очагов». В качестве осей координат *X* и *Y* принимаются координаты сейсмических событий по *X* и *Y*.

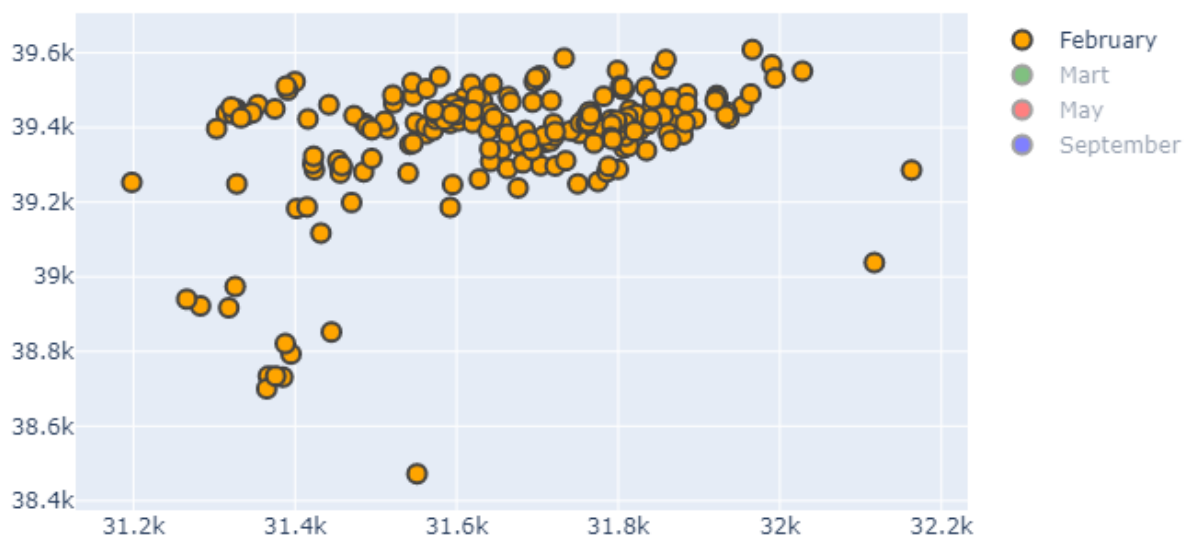


Рисунок 2.5.3 - Распределение сейсмических событий за февраль 2018

Как видно из рисунка 2.5.3, определить наличие каких-либо кластеров в массиве горных пород практически невозможно.

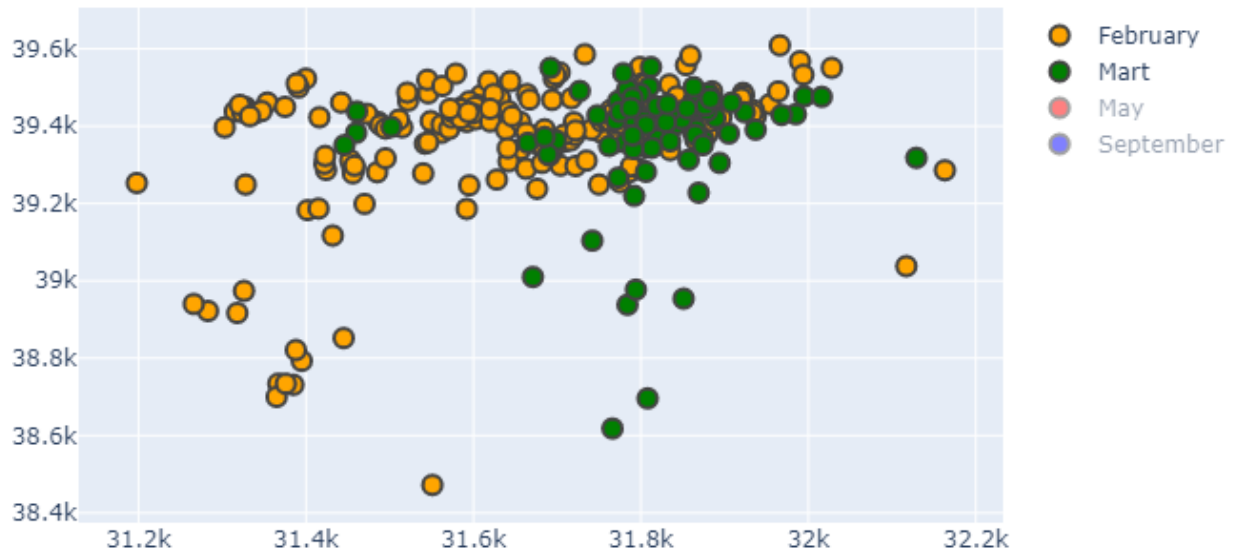


Рисунок 2.5.4 - Кластеризация сейсмических событий за февраль март 2018 года

На рисунке 2.5.4 показано, что наши сейсмические события начинают группироваться в один кластер, что, как правило связано с ведением горных работ. Однако говорить о наличии каких-либо дополнительных кластеров пока преждевременно.

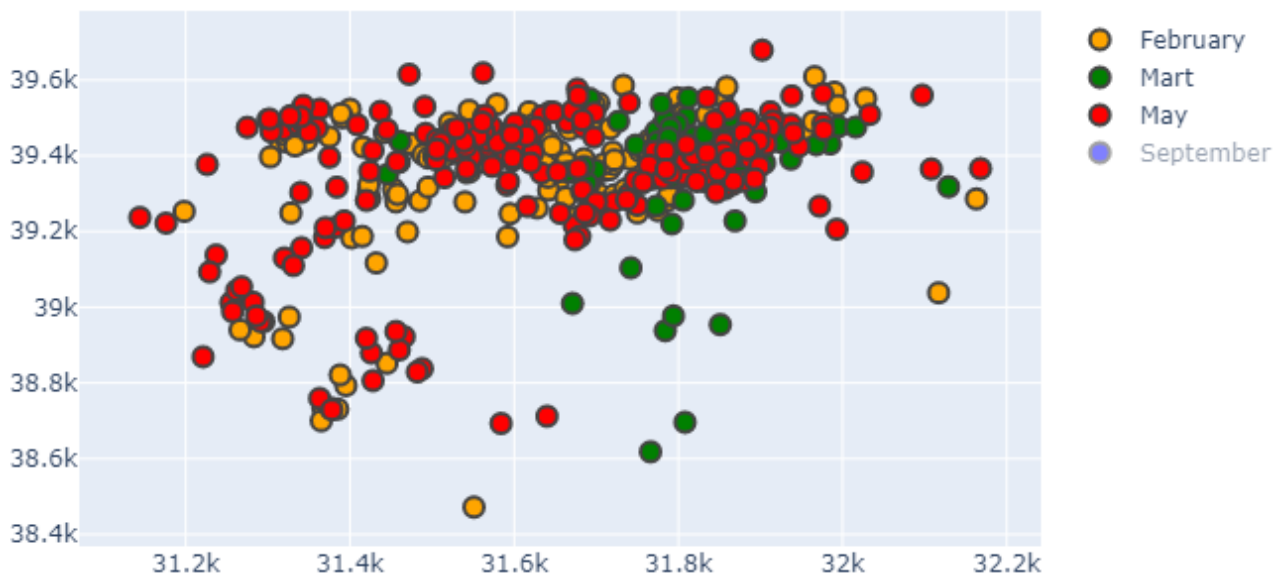


Рисунок 2.5.5 - Кластеризация сейсмических событий за февраль май 2018 года

В результате визуального анализа данных, изображенных на рисунке 2.5.5, было сделано предположение о начале формирования новых кластеров сейсмических событий, которые могут являться индикаторами начала формирования процессов деформации в массиве горных пород. Данные кластеры выделены фиолетовой кривой на рисунке 2.5.6.

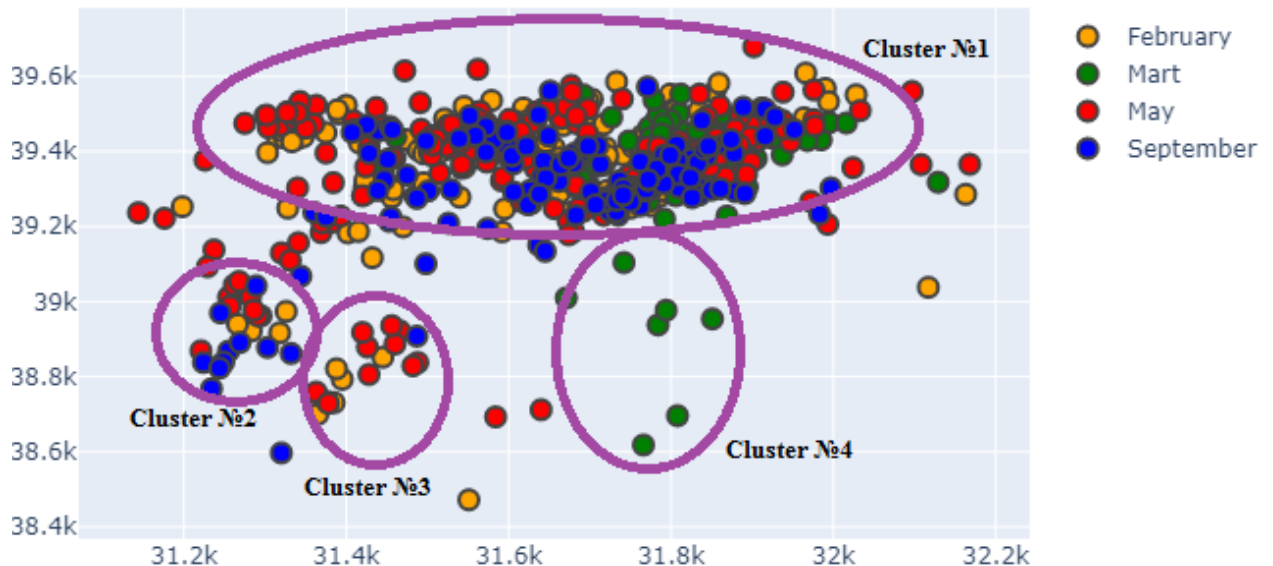


Рисунок 2.5.6 - Метки возможных кластеров сейсмических событий за февраль-сентябрь 2018 года

Отметим, что, зоны кластеров были сформированы еще в феврале и в дальнейшем либо «насыщались» за счет новых сейсмических событий, либо расширяли свои границы. Так же следует отметить, что визуально разделить кластеры №1 и №5 не представляется возможным из-за большой плотности данных в этой области. Однако, на рисунке 2.5.7 можно отметить «направление» распространения сейсмических событий. Получив такие предварительные оценки, было принято решение применить алгоритм для этих данных, учитывая отсутствие информации о количестве кластеров или количестве точек внутри кластера. Результаты моделирования приведены в таблице 2.5.2.

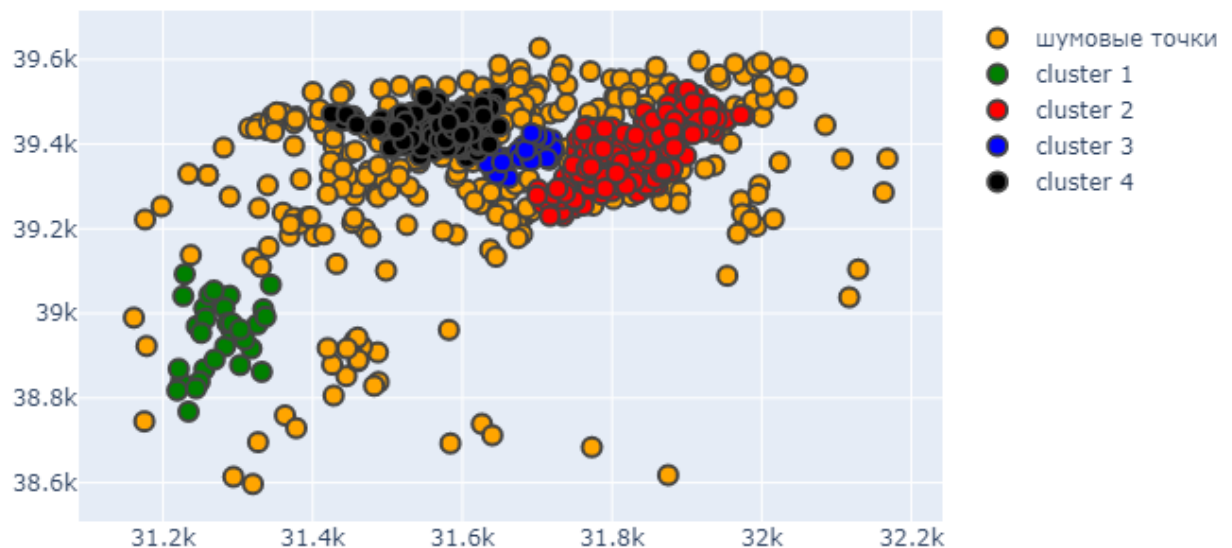


Рисунок 2.5.7 - Кластеризация сейсмических событий за февраль-сентябрь 2018 года

Использование кластерного анализа позволяет провести детальное исследование структуры множества сейсмических событий.

Таблица 2.5.2 - Характеристика кластеров сейсмических событий с февраля по сентябрь 2018 года

Кластер	Min cluster size	Min points	Количество точек внутри кластера
№	Eps, м	п, шт	п, шт
Шумовые точки	-	-	252
Cluster 1	27	5	32
Cluster 2	27	5	266
Cluster 3	27	5	32
Cluster 4	27	5	283

Таким образом, на основании проведенного анализа можно сделать следующие выводы:

- Визуально выделенные кластеры №3,4,6 были отнесены алгоритмом к «шумовым» точкам. Причинами, не позволяющими объявить

данные области кластерами, являются недостаточное число сейсмических событий и их разрозненность в пространстве.

- Алгоритм обнаружил, выделенный визуально кластер №2 и определил его границы, радиус и количество точек внутри (Cluster №1 в таблице 2.5.2).
- Алгоритм выделил кластер №1 и идентифицировал его как 3 независимых кластера. Данный факт может служить основой для теоретических предположений о структуре горного массива в данном регионе.

2.6 Выводы по главе 2

По результатам исследований, произведенных в данной главе, можно сделать следующие выводы и рекомендации:

1. Произведен ретроспективный анализ сейсмических событий и получены частотные распределения сейсмических событий по месяцам и произведена фильтрация событий исходного набора данных.
2. Проанализированы данные сейсмического мониторинга предшествующих всем документально зафиксированным опасным геодинамическим явлениям (2009, 2010, 2015, 2016, 2018 года). Была выявлена единая для всех рядов структура, состоящая из аддитивных компонентов (тренда, цикличности и шума). Это значит, что все ряды порождены единой динамической системой и их параметры могут быть применены друг относительно друга..
3. Предложена модель анализа временного ряда сейсмического мониторинга на основе метода SSA (singular spectrum analysis). Цель данной модели – поиск во временном ряду участков являющих данной комбинацией факторов с целью обнаружения потенциально опасных геодинамических явлений.
4. Сформированы кластеры сейсмических событий, основываясь на данных за весь период наблюдений. Для поиска и идентификации потенциальных «кластеров очагов» сейсмических событий был использован

алгоритм кластеризации HDBSCAN на основе данных сейсмического мониторинга за 2018 год. Получены оценки минимально допустимого размера кластера. Форма, положение в пространстве и размер кластеров были интерпретированы с точки зрения физических процессов в массиве горных пород.

5. Сформировано предположение о статистически значимом диапазоне значений сейсмических событий. Диапазон от 10^2 до 10^5 Дж, где значение 10^2 – выбрано эмпирически, а значение 10^5 – является статистически надежной границей.

6. Схожесть типов распределений, а также наличие взаимосвязей изменения частоты значения от времени наблюдения позволяют сформировать параметр для рассматриваемой модели расстояния в метрах от условного «центра масс» деформационного мониторинга до каждого зарегистрированного сейсмического события (евклидово расстояние, далее по тексту Euc_{dist}). Наличие бимодальности в графике распределения сейсмических событий в пространстве деформационного мониторинга позволяет выделить два интервала значений предложенной метрики:

- $300 < Euc_{dist} < 700$ м;
- $1500 < Euc_{dist} < 2500$ м

7. Сформировано предположение о связи между минимальным значением тренда сейсмической активности и наступлением опасного геодинамического события. Резкая смена тренда, без наличия периода стационарности тренда, может служить триггером для проявления опасных геодинамических явлений.

8. Сформировано предположение о связи между разницей значений трендов сейсмической активности в точке начала роста и в точке наступления опасного события. Данное предположение согласуется с эвристикой «трех сигм» и его можно использовать в качестве «предвестника» опасных геодинамических событий. Поскольку наступлению опасного периода предшествует период «затишья» длиной от 15 до 30 дней, данная

методика позволит своевременно обнаруживать и предпринимать меры по предотвращению проявления опасных геодинамических событий.

9. Использование кластерного анализа позволяет провести детальное исследование структуры множества сейсмических событий и выделить потенциально опасные зоны геодинамических событий.

ГЛАВА 3 МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ДЕФОРМАЦИИ НА ОСНОВЕ ДАННЫХ СЕЙСМИЧЕСКОГО МОНИТОРИНГА. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СИСТЕМЫ ПРОГНОЗА ПРОЦЕССОВ ДЕФОРМАЦИИ ТЕХНОГЕНО-НАРУШЕННОГО МАССИВА

3.1 Концепция моделирования процессов деформации на основании данных сейсмического мониторинга

В точки зрения моделирования динамических систем геомеханические модели можно отнести к классу сосредоточенных систем (число степеней свободы системы конечно). Моделирование и изучение таких систем довольно хорошо изучено и может быть представлено следующим образом:

- Определение характеристик системы (энтропия, показатель устойчивости Ляпунова);
- Выбор уравнений и законов взаимодействия системы;
- Процедура прогноза и оценка результатов моделирования;
- Калибровка гиперпараметров системы.

Однако на практике в реальном мире преобладают так называемые распределенные системы.

Рассмотрим данный случай в контексте геомеханической модели. Фазовое пространство такой модели будет бесконечномерным (требуется определить начальные условия в каждой точке пространства). Каждая точка в МГП отвечает определенному распределению величин, характеризующих систему: полю напряжений, деформаций, энергии и т. п.

Как правило для изучения таких систем используют мощные численные методы. Но зачастую они оказываются некорректными, и проверить численно полученные результаты невозможно. Альтернативным подходом является приведение распределенных систем к конечномерным и их изучение без составления конечной модели. Большая часть природных систем диссипативна: с течением времени все фазовые траектории стягиваются к некоторому подмножеству исходного бесконечномерного

пространства. Как правило, это подмножество конечномерно (описывается конечным набором переменных). Наименьшее число независимых переменных, однозначно определяющих установившееся движение исходной диссипативной распределенной системы, называют размерность вложения. Очевидно, подмножество, к которому стягиваются траектории системы — это аттрактор. Размерность вложения аттрактора — это минимальная размерность фазового пространства, в которое без самопересечений может быть помещено гладкое многообразие, целиком содержащее этот аттрактор. Таким образом, можно смоделировать динамику исходной системы, изучая это подмножество.

Для определения размерности вложения необходимо использовать порождаемые системой временные ряды, которые фиксируют изменения одной переменной в какой-либо точке.

Большинство инженерных задач формулируются как задача предсказания (прогноза) некоторой вещественной переменной. Это проблема регрессии, которую можно описать как поиск функции, которая наилучшим образом аппроксимирует некоторые данные. Это побуждает к использованию соответствующих методов, уже доступных, таких как линейная регрессия, полиномиальные функции, сплайны или ортогональные полиномиальные функции. Большинство моделей, основанных на данных, используют комбинации многих простых функций. По сути, обучение направлено на оптимизацию количества этих функций и значений их параметров (с учетом класса функций). Рассмотрим некоторые из наиболее эффективных решений.

3.1.1. Многослойный персептрон (MLP)

Многослойный персептрон (MLP) является типичным примером искусственной нейронной сети (ANN). Он состоит из нескольких слоев взаимосвязанных узлов (нейронов), каждый из которых получает несколько наборов входных данных, вычисляет их взвешенную сумму и затем передает результат в нелинейную функцию “классификации”. Таким образом, входные

данные для модели MLP подвергаются многопараметрическому нелинейному преобразованию, так что полученная модель способна аппроксимировать сложные нелинейные взаимосвязи. Обучение MLP, по сути, решает проблему минимизации ошибки модели (как правило, среднеквадратичной ошибки) путем определения оптимального набора весов. Поскольку принцип обратного распространения для обучения MLPs был найден и усовершенствован в 1970–80-х годах, этот тип ANN стал самым популярным инструментом машинного обучения. Различные типы ANN широко используются для прогнозирования и классификации. Обратное распространение ошибки — это принцип, который позволил использовать градиентные методы для обучения MLP, и это позволяет использовать различные алгоритмы оптимизации – от упрощенных версий градиентного спуска, до более эффективных методов, таких как сопряженный градиент или метод Бroyдон–Флетчер–Гольдфарб–Шанно (BFGS)

3.1.2. Генетическое программирование (GP) и эволюционная регрессия

GP — это метод символической регрессии, в котором конкретная структура модели не выбирается априори, а является результатом процесса поиска. Различные элементарные математические функции, константы и арифметические операции объединены в одну функцию, и алгоритм пытается построить модель, рекомбинирующую эти строительные блоки в одной формуле. Структура функции представлена в виде дерева, и поскольку результирующая функция сильно нелинейная и часто недифференцируема, оптимизируется методом рандомизированного поиска –обычно GA. Одно из критических замечаний в адрес GP связано с тем, что формулы, созданные на основе комбинации множества элементарных функций, часто чрезвычайно сложны и не несут в себе физического смысла. Для решения этой проблемы была предложена дополненная версия GP – GP с учетом размеров. Она ограничивает поиск и гарантирует, что выходные данные имеют ожидаемое физическое измерение, позволяя использовать только формулы с переменными определенных размеров (представляющими собой комбинацию

длины, времени, массы и.т.д.). Это приводит к формуле с размерной семантикой и увеличивает вероятность того, что у них есть некоторые физический смысл. В эволюционной регрессии элементарные функции выбираются из ограниченного набора, а структура общей функции фиксируется. Как правило, используется полиномиальное уравнение регрессии, и коэффициенты определяются генетическим алгоритмом. Этот метод преодолевает некоторые недостатки GP, такие как вычислительные требования, количество настраиваемых параметров и сложность результирующих символьных моделей.

3.1.3 Системы, основанные на нечетких правилах (FRBS).

Нечеткая логика была введена Лотфи Заде (1965), и с тех пор она нашла множество успешных применений, главным образом, в теории управления. Системы, основанные на нечетких правилах, могут быть построены путем опроса экспертов-людей или путем обработки исторических данных и формирования модели, управляемой данными. Эти наборы локальных моделей перекрываются по всему пространству параметров, используя своего рода интерполяцию на более низком уровне для представления паттернов в сложных нелинейных отношениях.

3.1.4. Теория хаоса и нелинейная динамика

Данная теория применяется для прогнозирования временных рядов, когда временной ряд содержит достаточно информации о поведении системы. Пусть задан временной ряд $X = (x_1, \dots, x_n)$ (например, последовательность сейсмических событий). Состояние системы в момент времени t может быть представлено вектором y_t в m -мерном пространстве состояний $X = (x_t, \dots, x_{t-(m-1)z})$ — где z — время задержки. Тогда весь временной ряд может быть представлен последовательностью таких векторов y_t . Если исходный временной ряд проявляет так называемые хаотические свойства (проявляющиеся в виде эквивалентной траекторией в фазовом пространстве по квазипериодической схеме), то методы теории хаоса могут

быть использованы для предсказания будущих значений y , a , следовательно, и x . Для этого в фазовом пространстве должна быть построена так называемая локальная модель, предсказывающая будущее значение y ; это модель обучения на основе экземпляра или регрессионная модель (линейная или нелинейная), построенная на основе точек, представляющих “движения в фазовом пространстве” соседей текущего y . Прогностическая способность теории хаоса основана на идее что система ведет себя в будущем так же, как и в (далеком) прошлом. В практических приложениях время задержки z и размерность m должны быть выбраны соответствующим образом, чтобы точно уловить динамическую структуру временного ряда. Методы, основанные на хаосе, не имеют универсальной применимости: они могут быть успешно применены только тогда, когда временные ряды (или их комбинация) имеют некоторые свойства, например, являются периодическими или действительно проявляют свойства хаотического поведения (или близки к нему), и когда временные ряды имеют достаточную (значительную) длину.

3.1.5. Модели основанные на данных (DDM)

Геомеханические модели можно охарактеризовать как физические, математические и эмпирические. Последний класс моделей, в отличие от первых двух, включает математические уравнения, которые выводятся не из физических процессов, наблюдаемых, например, при лабораторных исследованиях, а из анализа данных временных рядов. Последние разработки в области искусственного интеллекта, в частности в области машинного обучения, значительно расширили возможности эмпирических моделей. Область, которая охватывает эти новые подходы, называется моделированием на основе данных (DDM). Как следует из названия, DDM основан на анализе данных о системе, в частности на нахождении связей между переменными состояния системы (входными, внутренними и выходными переменными) без явного знания физического поведения этих

систем. Эти методы представляют собой значительный прогресс в области традиционного эмпирического

моделирования и включают вклад из следующих пересекающихся областей:

- искусственный интеллект (AI), который является всеобъемлющим исследованием того, как человеческий интеллект может быть включен в компьютеры;
- вычислительный интеллект (CI), который включает в себя нейронные сети, нечеткие системы и эволюционные вычисления, а также другие области искусственного интеллекта и машинного обучения;
- мягкие вычисления (SC), которые близки к CI, но с особым акцентом на системы, основанные на нечетких правилах, индуцированных из данных;
- машинное обучение (ML), которое когда-то было подотраслью AI, концентрирующейся на теоретических основах, используемых CI и SC;
- интеллектуальный анализ данных (DM) и обнаружение знаний в базах данных (KDD).

Поскольку природные процессы сложны, иногда невозможно построить единую глобальную модель, адекватно отражающую поведение системы. Вместо этого обучающие данные могут быть разделены на несколько подмножеств, и на каждом подмножестве могут быть построены отдельные специализированные модели. Эти модели называются локальными или экспертными моделями, и этот тип модульной модели иногда называют композитной моделью (CM).

Моделирование на основе данных и вычислительный интеллект в целом доказали свою применимость к различным проблемам: моделирование, краткосрочное прогнозирование, классификация данных. Модели, основанные на данных, были бы полезны при решении практической проблемы или моделировании конкретной системы или процесса, если:

- имеется значительный объем данных, описывающих эту систему;
- В течение периода, охватываемого моделью, в моделируемой системе не произошло существенных изменений.

При построении СМ необходимо принять два ключевых решения. Во-первых, как разделить данные, а во-вторых, как объединить отдельные модели для получения конечного результата. Существует несколько общепринятых подходов.

- Группа статистически управляемых подходов с “мягкими” разделениями входного пространства
- Построение ансамбля моделей и объединение результатов модели по некоторой схеме усреднения; этот подход широко используется в метеорологии.

- Группа методов не объединяет результаты различных моделей,

Такие методы используют “жесткие” разбиения одного пространства на области. Каждая отдельная локальная модель обучается индивидуально на подмножествах экземпляров, содержащихся в этих регионах, и, наконец, учитывается результат только одной специализированной модели.

Такие модели особенно эффективны, если трудно построить имитационные модели, основанные на знаниях (например, из-за отсутствия понимания базовых процессов), или доступные модели недостаточно адекватны. Конечно, всегда полезно иметь альтернативы моделированию и проверять результаты моделирования физически основанных моделей с помощью моделей, основанных на данных, или наоборот. Однако, такие модели, как правило, на самом деле не отражают физику моделируемого процесса; это всего лишь устройства, используемые для фиксации взаимосвязей между соответствующими входными и выходными переменными.

Однако такие устройства могут быть более точными, чем модели процессов, поскольку они основаны на объективной информации (т. е. данных), и последние часто могут страдать от неполноты представления

моделируемого процесса. Современная тенденция заключается в объединении моделей, основанных на данных, т. е. объединении моделей различных типов, которые следуют различным парадигмам моделирования (таким образом, образуя гибридные модели), включая оптимальное сочетание с моделями, основанными на физических данных.

Достоверный анализ и моделирование процессов деформации техногенно-нарушенного массива базируется на применении математических моделей. Однако, получение качественных и достоверных результатов при реализации таких моделей без непосредственного участия эксперта предметной области исследований всегда затруднено, так как последний выбирает соответствующие стратегии моделирования и предварительной обработки данных. Одной из таких стратегий будут являться системы управления технологическими объектами и процессами на базе математических моделей и цифровых моделей, управляемыми данными [43-44]. Поэтому, для исключения влияния субъективного подхода, был выбран метод создания математической модели, управляемой данными сейсмического мониторинга, т. е. DDM подход. Искомую математическую модель можно разработать, используя как одиночную модель машинного обучения, так и гибридный (композиционный) подход. В настоящее время идентификация DDM моделей со сложной гетерогенной структурой остается нерешенной проблемой. Она тесно связана с задачей автоматического машинного обучения (также называемой AutoML), но в то же время, в первую очередь, ориентирована на идентификацию оптимальных моделей, а не на оптимизацию параметров. Структура композитной модели может быть представлена в виде ориентированного ациклического графа (DAG), а наиболее подходящий вариант структуры может быть найден с использованием оптимизационных подходов. Элементами этого графа являются модели машинного обучения.

3.2 Композитная модель прогноза процессов деформации в техногенно-нарушенном массиве на основе данных его сейсмического мониторинга

Анализ данных сейсмического мониторинга и кластеризация сейсмических событий требует применения современных методов математического моделирования. В процессе изучения сейсмического режима Кукисвумчоррского месторождения за длительный период было установлено что: «кластеры сейсмических событий приурочены к местам активного ведения горных работ, к дизъюнктивным нарушениям в консоли пород висячего бока, а также могут образовываться под влиянием других факторов, как природных, так и техногенных» [41, 42].

На основе проведенных исследований разработан алгоритм системы анализа и прогноза опасных геодинамических явлений техногенно-нарушенного горного массива (рисунок 3.2.1.), заключающийся в одновременном применении различных моделей машинного обучения, наиболее подходящих для определенного типа данных, в единой композитной модели [32-41].

Такой подход позволил получить набор сейсмических кластеров, которые могут быть потенциальными триггерами опасных проявлений горного давления. Задачей модели является поиск и детектирование кластеров-"очагов разрушения горной породы". Для оценки качества моделирования в работе предлагается ввести два критерия оценки качества итоговой модели, полученных в процессе многокритериальной оптимизации [42-50]. Для решения данной задачи в данной работе было принято использовать алгоритм эволюционной многокритериальной оптимизации (ЕМО) [51-59]. Алгоритмы эволюционной оптимизации (ЕО) используют подход, основанный на принципе, в котором на каждом шаге алгоритма учувствует более одного решения и в результате создается новая совокупность решений на каждой итерации.

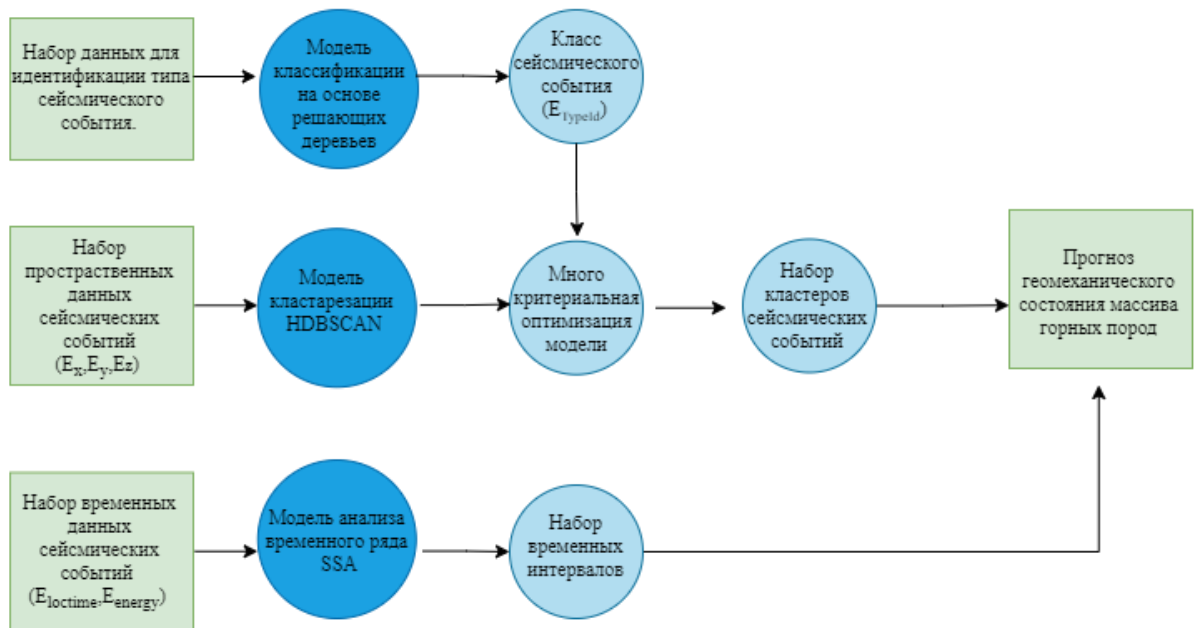


Рисунок 3.2.1 - Композитная модель прогноза опасных геодинамических явлений

Данный подход довольно популярен по ряду причин:

- ЕМО не требуют никакой производной информации;
- ЕМО относительно просты в реализации;
- ЕМО являются гибкими широко распространенными в прикладных задачах алгоритмами.

Следует отметить, что для решения задач однокритериальной оптимизации, особенно при поиске единственного оптимального решения, использование совокупности решений может показаться избыточным, однако при решении задач многокритериальной оптимизации процедура ЕО является идеальным выбором.

Решение, которое является оптимальным по отношению к одному критерию, требует компромисса с другими критериями [60-71]. Такой подход запрещает выбирать решение, оптимальное только для одного критерия.

Таким образом работу алгоритма многокритериальной оптимизации можно представить следующим образом:

- Поиск набор решений, которые лежат на оптимальном по Парето фронте

- Поиск набора решений, которые достаточно разнообразны, чтобы представлять весь диапазон оптимального по Парето фронта.

Поскольку оптимальным является набор решений довольно нетривиальной является задача выбора наилучшего решения. Довольно часто она решается с помощью эвристического подхода основанного на опыте исследователя. В данной работе были введены два критерия качества кластеризации, которые оптимизируются в ходе работы многокритериального алгоритма.

1. Критерий ***Silhouette***. Силуэтом выборки называется средняя величина силуэта объектов данной выборки. Таким образом, силуэт показывает, насколько среднее расстояние до объектов своего кластера отличается от среднего расстояния до объектов других кластеров. Данная величина лежит в диапазоне от -1 до 1. Значения, близкие к -1, соответствуют плохим (разрозненным) кластерам, значения, близкие к нулю, говорят о том, что кластеры пересекаются и накладываются друг на друга, значения, близкие к 1, соответствуют "плотным", четко выделенным кластерам, то есть, чем больше силуэт, тем более четко выделены кластеры, и они представляют собой компактные, плотно сгруппированные облака точек.

2. Критерий ***гомогенности кластера***. Данный критерий определяет насколько однородными являются полученные кластеры с точки зрения распределения в них классов сейсмических событий. Данная величина лежит в диапазоне от 0 до 1. Значения, близкие к 1, соответствуют однородным кластерам, которые можно интерпретировать с точки зрения механики разрушения горных пород, значения близкие к нулю, говорят о том, что кластеры неоднородны и не могут быть интерпретированы с точки зрения геомеханики.

3.3 Программное обеспечение системы прогноза опасных геодинамических явлений при ведении работ в техногенно нарушенном массиве

Программное обеспечение системы прогноза опасных геодинамических явлений, при ведении работ в техногенно-нарушенном массиве разработано и реализовано в среде разработки *Python* для автоматического определения основных параметров сейсмических и деформационных процессов внутри массива горных пород, адаптированный к конфигурации, существующей сети станций деформационного и сейсмического мониторинга и особенностям сейсмичности региона.

На первом этапе работы программы происходит дискретизации данных сейсмического мониторинга в пространстве и времени. Эта позволяет группировать сейсмические события в связанные кластера («очаги») в пространстве и через регулярные промежутки времени. Как видно из рисунка 3.3.1, общей проблемой для данных, поступающих в этот модуль, является то, что точки не пересекаются друг с другом.

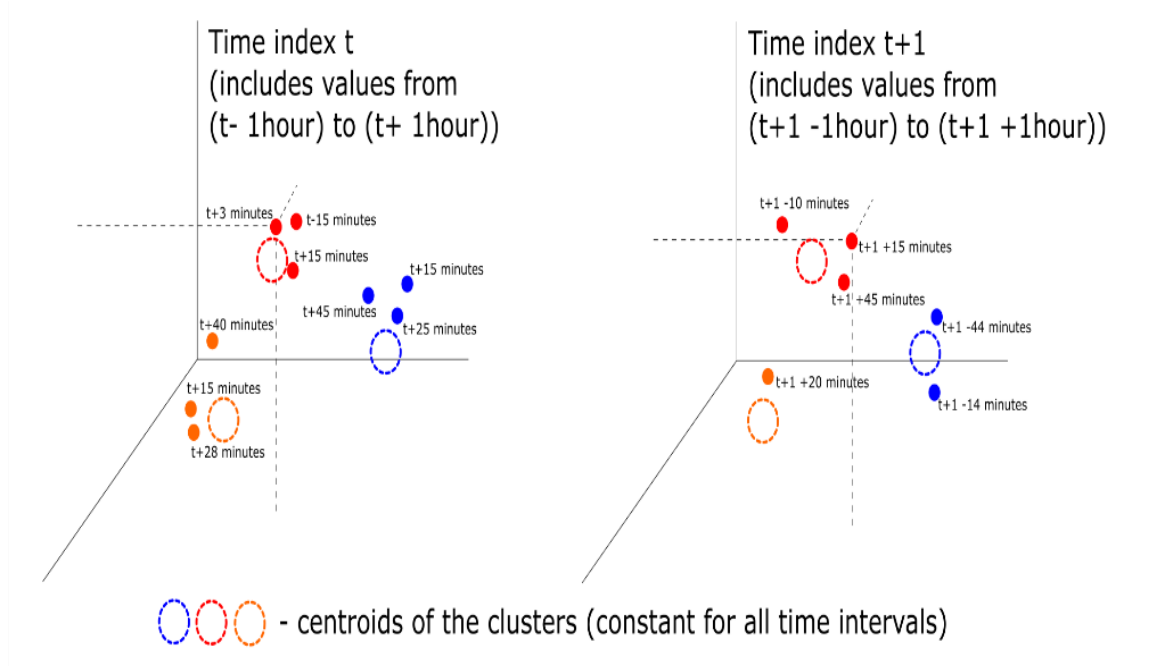


Рисунок 3.3.1 – Первый этап работы программы, на котором выбирается временной шаг дискретизации пространства сейсмических событий

На их основе невозможно построить временные ряды даже с нерегулярным шагом. Если все объекты в таблице имеют разные временные привязки и интервалы временных меток не являются постоянными, то нам нужно поместить значения в стационарную временную сетку. Для этого, при помощи алгоритма DBSCAN, объединяются близко расположенные (по трем координатам X , Y , Z) сейсмические события в один кластер. События, расположенные в одном кластере, могут быть интерпретированы как один объект, для которого значения энергии фиксировались в разное время немного в разных местах, и тем не менее, показатели в окрестности характеризуют именно один этот объект. Раз объект один, то и координаты для него должны быть определены однозначно и это координаты центроида нашего кластера.

На втором этапе у нас имеются сейсмические события с назначенными метками пространственных кластеров и новыми регулярными временными индексами. Работу программы на этом этапе можно представить в виде следующей цепочки действий:

- Происходит оценка значения некоторых переменных (непрерывных или категориальных) в полученных пространственных кластерах. В данном случае прогнозируются переменные «Энергия» (вещественный признак) и «Тип сейсмического события» (категориальный признак)). Для этого используется два алгоритма К-ближайших соседей: один алгоритм по соседям вычисляется значение «Энергии» для кластера, другой по мажоритарному классу типа события определит «Тип сейсмического события» для кластера;
- С помощью композитной модели кластеризации вычисляются координаты центроида кластера- X , Y , Z . Для каждого временного индекса на регулярной сетке определяется набор точек (строк из таблицы). Эти точки относятся к одному и тому же периоду времени, но могут принадлежать к разным пространственным кластерам. Таким образом, на данном этапе мы точно знаем координаты центроида кластера и значение индекса времени для

каждого сейсмического события;

- Рассчитывается разность для каждого сейсмического события (X , Y , Z и времени) в виде: "расстояние по координате $x = x$ центроида кластера - координата X этого события". Время для этой операции преобразуется в абсолютные значения, например, в минутах в пределах заданного временного блока;

- На основе полученных «расстояний» прогнозируется значение целевой переменной в центре кластера для новой временной метки с использованием алгоритма К-ближайшего соседа, который передает значения расстояний для всех четырех координат в качестве признаков.

В результате этих действий получен массив точек, расположенных на регулярной сетке, как в пространстве, так и во времени. Упрощенная демонстрация алгоритма представлена на рисунке 3.3.2. Итогом работы программы является создание набора данных, в котором каждая координата центроида кластера связана с определенным значением целевой переменной в определенный временной индекс.

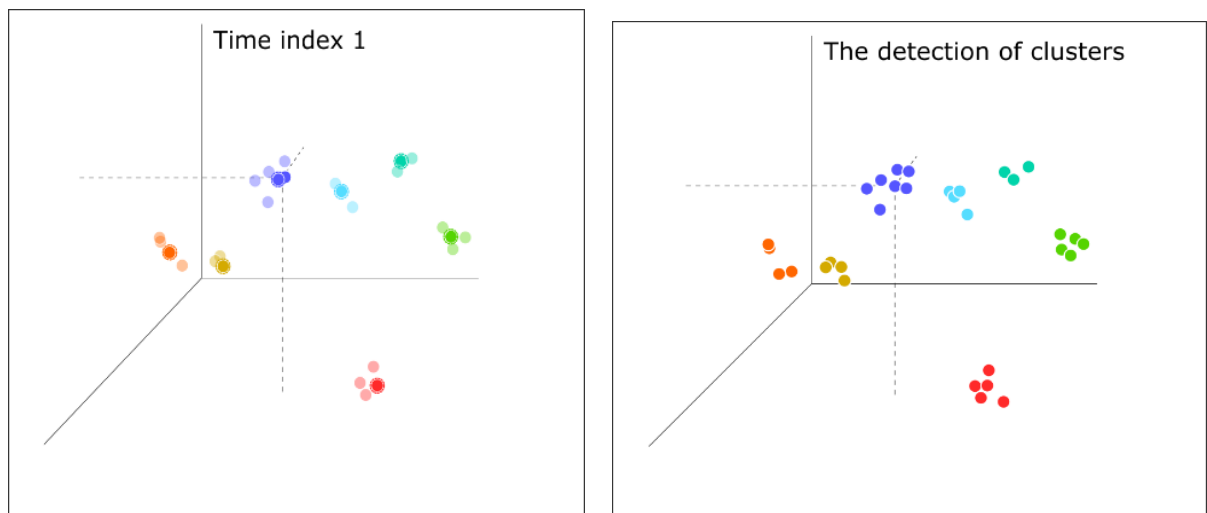


Рисунок 3.3.2. – Визуализация 1 (слева) и 2 (справа) этапов работы программы

3.4 Выводы по главе 3

По результатам исследований, произведенных в данной главе, можно сделать следующие выводы и рекомендации:

- Произведен анализ существующих математических методов моделирования динамических систем. Рассмотрены различные реализации прикладных алгоритмов из теории машинного обучения.

- Реализовано программное обеспечение системы прогноза опасных геодинамических явлений в среде разработки Python для автоматического определения основных параметров сейсмических и деформационных процессов внутри массива горных пород. Данная реализация адаптирована к конфигурации, существующей сети станций деформационного и сейсмического мониторинга АО «Апатит» и запатентована под следующим названием: «Программа «SeDef» для обработки данных сейсмического мониторинга и поиска кластеров-очагов сейсмических событий с использованием методов иерархической кластеризации» [4].

- Реализована композитная модель машинного обучения. Цель данной модели – ансамблирование результатов работы атомарных моделей машинного обучения. Итоговый прогноз является комбинацией факторов цель которого - обнаружение потенциально опасных геодинамических явлений.

- Внедрена реализация алгоритма многокритериальной оптимизации для подбора гиперпараметров композитной модели. Разработан собственный критерий качества («критерий гомогенности кластера»), позволяющий учесть физические особенности процессов деформации в техногенно нарушенном массиве.

ГЛАВА 4 ГЕОМЕХАНИЧЕСКИЙ ПРОГНОЗ ПРОЦЕССОВ ДЕФОРМАЦИИ ТЕХНОГЕНОНАРУШЕННОГО МАССИВА НА ПРИМЕРЕ ХИБИНСКИХ АПАТИТ-НЕФЕЛИНОВЫХ РУД

4.1 Прогнозирование параметров сейсмических и деформационных процессов внутри техногенно-нарушенного горного массива

Длина анализируемого временного ряда составляет 365 дней, ширина окна равна 30 дням (параметр L), а в качестве наблюдаемой величины используется среднее значение всех сейсмических событий за один день. В качестве оси X выбрана временная шкала сейсмического мониторинга, в качестве оси Y шкала величины среднего значения сейсмических событий в Дж.

Поскольку выбрана ширина окна равная 30 дням, получено 30 элементарных матриц, которые описываются собственными значениями матриц, оценивая которые можно предположить их вклад в исходный ряд. Параметр предполагаемой периодичности процесса выбираем равным 15 дням. Значение этого параметра и его интерпретация крайне важны при анализе, т. к. он равен половине длине окна, что означает наличие двух подпроцессов в рамках 1 месяца. И если в первом подпроцессе (первые 15 дней) не было опасных проявлений ГД, то вероятность их появления во втором подпроцессе (вторые 15 дней) растет экспоненциально с каждым днем. На рисунке 4.1.1 и на рисунке 4.1.2 показаны восстановленные по первой элементарной матрице временные ряды для 2009, 2010 и 2016 годов, соответственно.

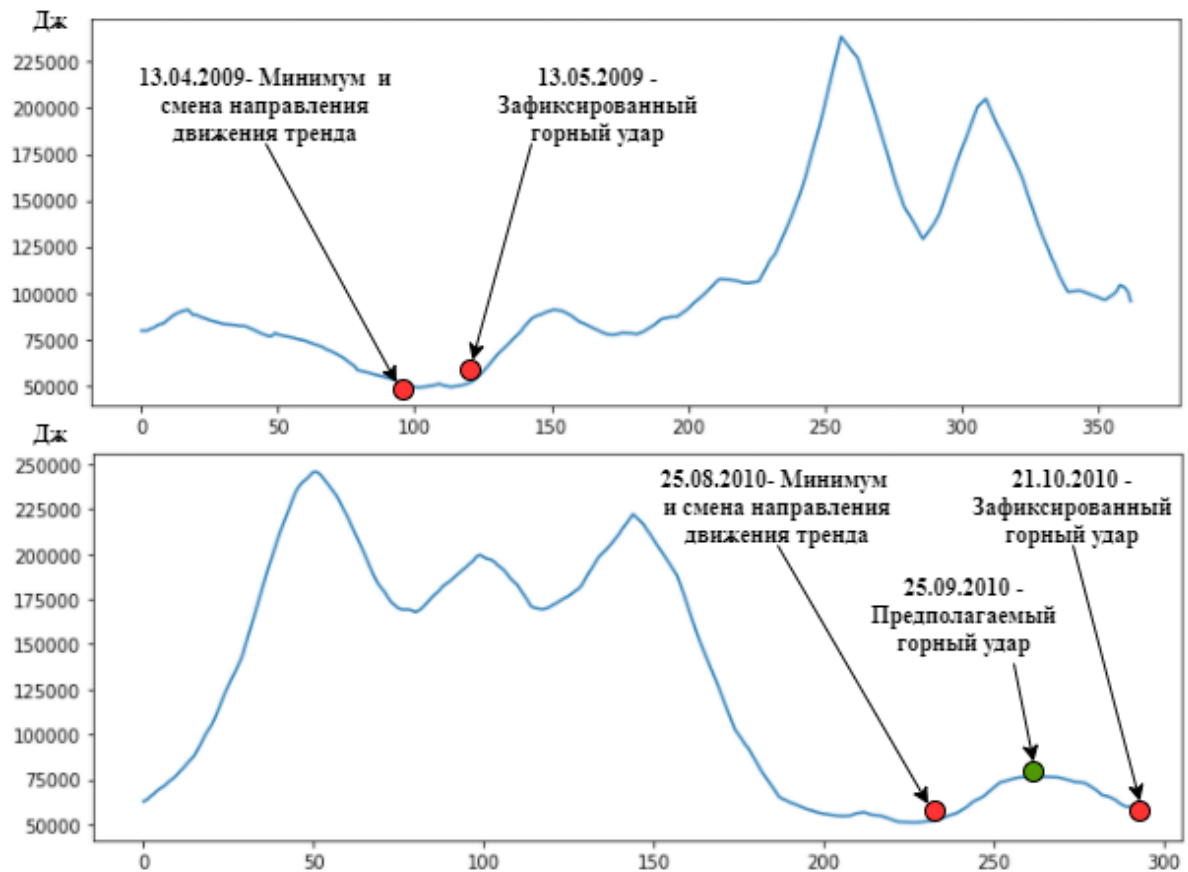


Рисунок 4.1.1 - Тренд временного ряда сейсмического мониторинга за 2009 и 2010 год соответственно

Первая элементарная матрица отвечает за тренд сейсмической активности. Как видно из рисунка 4.1.1, 13 апреля 2009 года и 25 августа 2010 года достигается минимальное значение тренда сейсмической активности в выбранные периоды наблюдений.



Рисунок 4.1.2 - Тренд временного ряда сейсмического мониторинга за 2016 год

Рисунок 4.1.2. представляет из себя сумму первого и второго случаев (2009 и 2010 год). С одной стороны, имеется повтор ситуации 2009 года (падение до минимального значения тренда сейсмической активности, и точка смены направления роста тренда) 27 декабря 2015 года. Таким образом, анализ тренда сейсмической активности, полученного с помощью метода SSA, позволяет явным образом установить связь между опасным проявлениями в виде горного удара от 21.01.2016, и сейсмической активностью, выраженной нисходящим трендом.

Сформируем предположение о том, что опасному проявлению ГД может предшествовать плавное снижение тренда до точки глобального минимума тренда с последующим изменением в направлении тренда. В соответствии с предположением о цикличности процесса, время до наступления опасного проявления ГД события равно 15 и 30 дням, соответственно.

4.2 Использование композитной модели для анализа данных сейсмического мониторинга за 2020 год

В таблице 4.2.1 приведены даты документально зафиксированных горных ударов за период 2009–2018 гг, пространственные координаты были отмасштабированы по шкале от 0 до 1.

Таблица 4.2.1 - Зафиксированные горные удары (по данным карточек горных ударов)

Координата сейсмического события по оси X	Координата сейсмического события по оси Y	Координата сейсмического события по оси Z	Время наступления зафиксированного геодинамического события.
0,456	0,413	0,325	13.05.2009
0,381	0,464	0,228	21.10.2010
0.603	0.441	0.293	27.01.2016

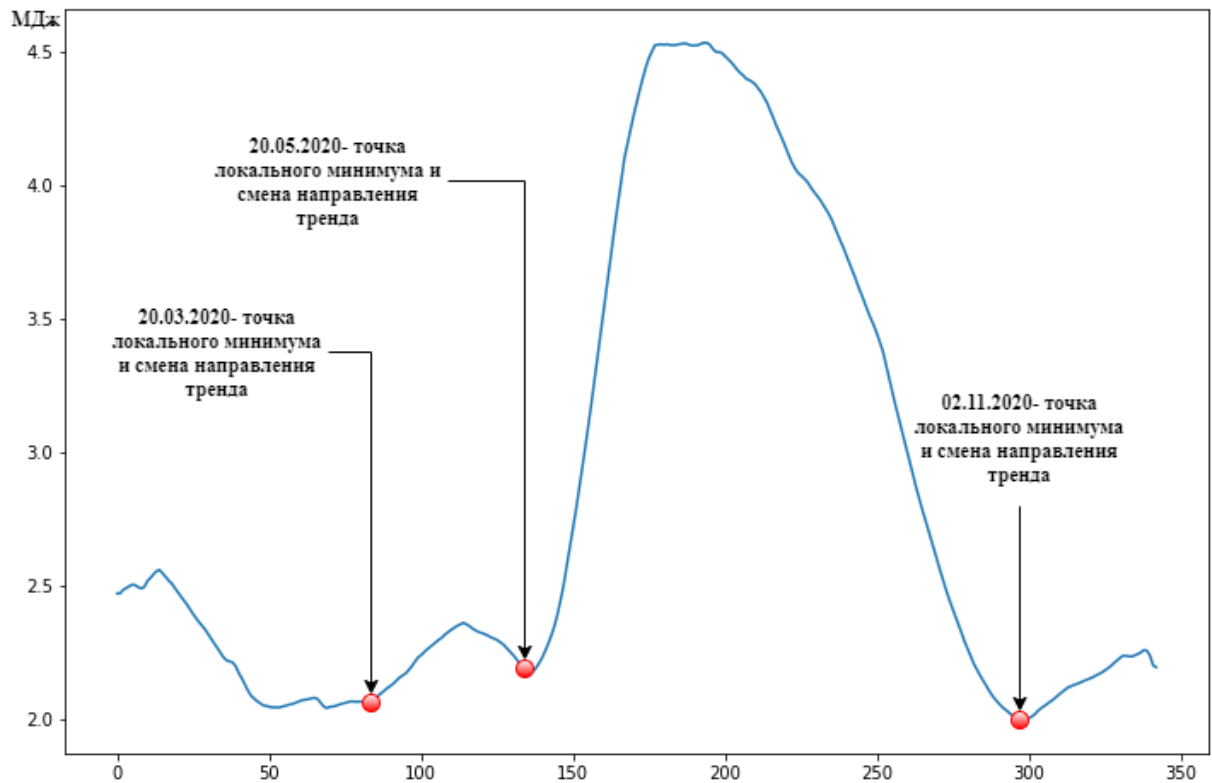


Рисунок 4.2.1 - Тренд временного ряда сейсмического мониторинга за 2020
год

Как видно из рисунка 4.2.1, можно выделить 3 ключевых точки в тренде:

- 20 марта 2020 года достигается минимальное значение тренда сейсмической активности на текущий момент наблюдений. Однако, вероятность наступления горного удара – умеренная, т. к. росту тренда предшествует период стационарности (отсутствие роста среднего значения тренда), равный длине окна (30 дней). Данный период можно интерпретировать как стабилизацию процессов деформации в МГП, что позволяет предотвратить возможные опасные проявления ГД;
- 20 мая 2020 года достигается локальный минимум тренда сейсмической активности и начинается резкий рост тренда в противоположном направлении. Вероятность наступления горного удара в следующие 30 дней достаточно велика, т. к. процесс достиг своего локального минимума, и начался рост сейсмической активности;

- 2 ноября 2020 года достигается глобальный минимум значения тренда сейсмической активности за весь период наблюдений, а затем начинается умеренный рост тренда сейсмической активности. Вероятность наступления горного удара в следующие 30 дней крайне велика, т. к. процесс достиг своего минимального исторического значения, и начался рост тренда сейсмической активности.

Таблица 4.2.2 - Сводная таблица прогнозов времени наступления потенциальных горных ударов

Время появления точки локального минимум и смены тренда	Время потенциального горного удара. Степень, риска-умеренная.	Время потенциального горного удара. Степень риска-крайне высокая.	Время наступления зафиксированного геодинамического события.
20.03.2020	06.04.2020	21.04.2020	08.04.2020
20.05.2020	06.06.2020	20.06.2020	06.07.2020
02.11.2020	17.11.2020	03.12.2020	04.12.2020

В двух из трех зафиксированных горных ударах удалось спрогнозировать приблизительное (с разницей в 1–2 дня) время наступления горного удара. Однако в случае, произошедшем 06.07.2020 ошибка прогноза составила 15 дней, что требует от авторов уточнения прогноза. Далее рассмотрим детально случай и восстановим исходный временной ряд, с использованием данных с 01.01.2020 по 05.07.2020 (за один день до фактического горного удара).



Рисунок 4.2.2.1 - Тренд временного ряда сейсмического мониторинга за период 01.01.2020-05.07.2020 год



Рисунок 4.2.2.2 - Циклическая компонента временного ряда сейсмического мониторинга за период 01.01.2020-05.07.2020 год

Как видно из рисунка 4.2.2.1 в точке 30.06.2020 достигается локальный максимум, и начинается смена направления движения тренда. На рисунке 4.2.2.2 видно, что в период с 30.06.2020 по 05.07.2020 зафиксировано максимальное значение амплитуды циклической составляющей. Совокупность этих факторов позволяет спрогнозировать наступление потенциального горного удара 06.07.2020. Таким образом, можно сделать

Вывод о том, что анализ тренда целесообразно дополнять анализом циклической составляющей.

4.3 Анализ результатов моделирования кластеров сейсмических событий на основе данных сейсмического мониторинга за 2020–2021 года

Как отмечается в ряде работ: «Существенная часть подходов к прогнозным оценкам проявлений опасного горного давления основывается на следующей концепции: по мере разрушения горной породы происходит формирование нескольких стадий разрушения, с постепенным переходом от одной стадии к другой» [48, 49]. Размеры и диапазон трещин, которые образуются в результате деформации и разрушения МГП, могут варьироваться от миллиметра до десятых долей метра. В реальности, наиболее часто, опасные проявления ГД выражаются в виде горных ударов и других проявлений техногенной сейсмичности.

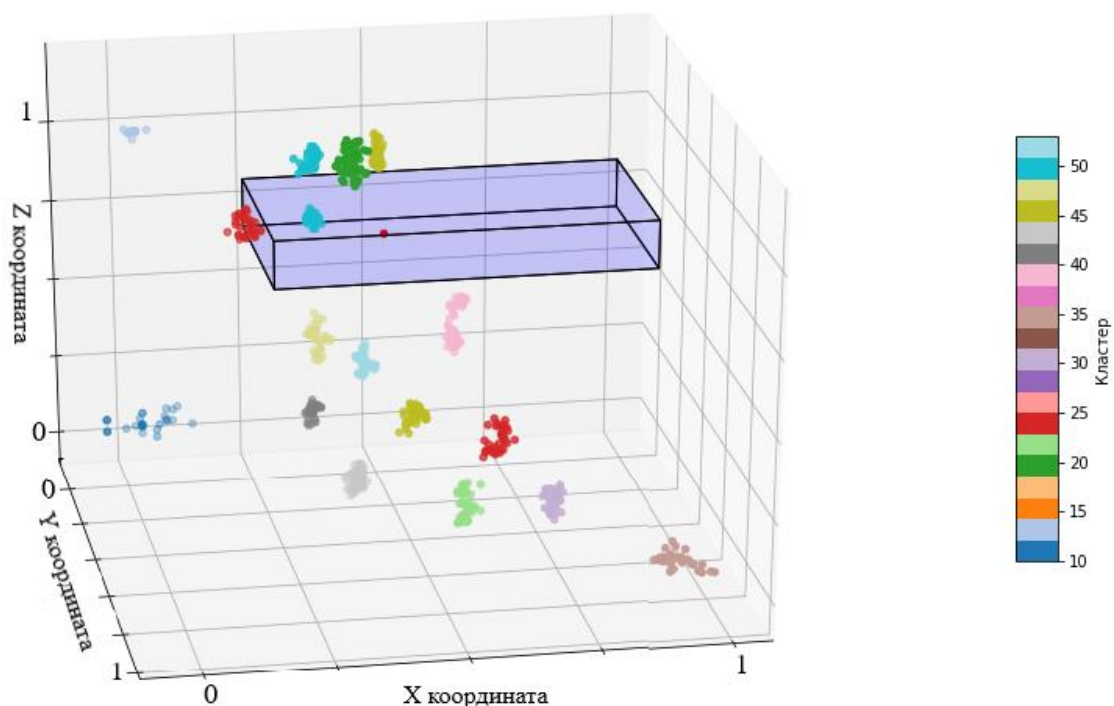


Рисунок 4.3.1 - Выделение всех «микрочластеров» (не более 100 событий в год) сейсмических событий. Фиолетовый параллелепипед-опасная зона (зона горных работ +100 м по осям X и Y и 50 метров запаса по оси Z)

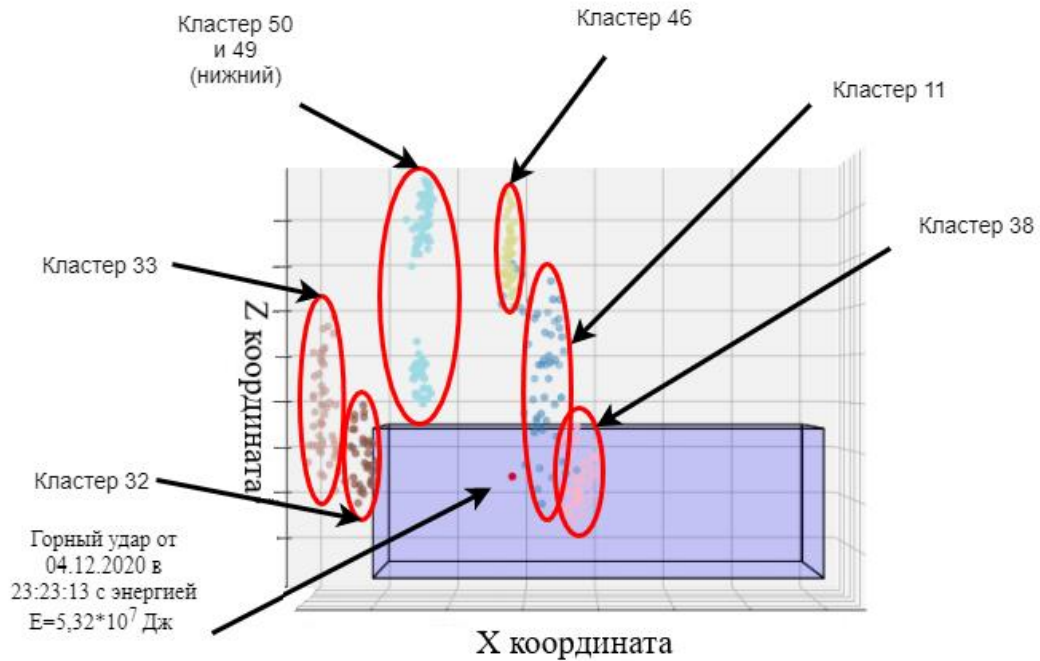


Рисунок 4.3.2 - Кластеры-"очаги разрушения горной породы"

С.А.Игнатъев в своей статье [48] утверждает, что: «Кластерный анализ опирается на два основных предположения:

- Выявленные признаки объекта должны допускать разбиение некоторой совокупности объектов на кластеры;
- Правильность выбора должного масштаба или необходимых единиц величин, отражающих признаки объекта (в некоторых случаях требуется применение стандартизованных величин)».

Группа авторов в своих работах предлагает «использование методов кластерного анализа для анализа данных сейсмического мониторинга показывает, что перед началом кластеризации следует выбрать подмножества данных, находящихся в областях с высокой плотностью точек. Для этого предлагается оценить плотность распределения в каждой из точек с использованием непараметрической оценки плотности. Исходя из описанной выше модели, решение данной задачи позволит осуществить дифференциацию значений, относящихся к регионам с низкой или высокой плотностью распределения. При этом значения, отнесенные к регионам с низкой плотностью, однозначно относятся к «шумовым» точкам» [49]. Так

же одним из этапов кластерного анализа является использование статистических критериев полученных, в результате анализа данных сейсмического мониторинга Кукисвумчоррского месторождения за длительный период наблюдений (2000–2020 гг.), приведенных в таблице 4.3.1.

Таблица 4.3.1 - Статистические значимые критерии распределения сейсмических событий в ходе мониторинга Кукисвумчоррского месторождения за период 2000–2020 г.

Название критерия	Значение в Дж
Среднее значение	9620+/241.76
Медиана	1610
Среднеквадратичное отклонение	18000

Таким образом, все сейсмические события, превышающие выборочное среднее (для простоты расчетов было выбрано значение 10^4 Дж), рассматриваются как сейсмические события способные привести к опасным проявлениям ГД. Используя в качестве критериев качества, предложенные авторами метрики, в ходе многокритериальной оптимизации с помощью разработанного алгоритма были получены кластеры «неклассических» геометрических форм (эллипсоиды), которые условно можно поделить на 3 группы по расположению относительно произошедшего горного удара. Кластеры 32-33 находятся севернее точки удара, кластеры 50,49,46 находятся «сверху» над точкой удара. Кластеры 11 и 38 находятся на юго-востоке от точки удара. Следует отметить, что все выделенные кластеры являются «устойчивыми» во времени, т. е. включают в себя сейсмические события в течении всего периода мониторинга в 2020 году. Таким образом, в таблице 4.3.2 приведены кластеры, которые с высокой долей вероятности могут являться триггерами горных ударов, и мониторинг сейсмических событий в этих кластерах в дальнейшем является целесообразным.

Таблица 4.3.2 - Характеристики сейсмических событий выбранных кластеров, превышающих 10^4 Дж

Время наступления сейсмического события	Величина энергии сейсмического события	Номер кластера	Время предполагаемого горного удара
2020-03-11 22:13:27	976000	50	6 апреля 2020
2020-04-12 17:53:49	153000	50	21 апреля 2020
2020-05-21 02:55:17	406000	50	6 июня 2020
2020-06-07 21:27:14	94600	46	6 июня 2020
2020-06-15 06:33:30	18400	50	21 июня 2020
2020-11-08 08:31:49	22300	32	17 ноября 2020
2020-11-21 01:24:33	15000	32	3 декабря 2020
2020-11-21 02:23:31	27400	33	3 декабря 2020
2020-11-28 13:48:17	24000	32	3 декабря 2020

4.4 Выводы по результатам моделирования на основе данных сейсмического мониторинга за 2020 год

Анализируя кластеризацию сейсмических событий на основе данных сейсмического мониторинга за 2020 год, можно сделать следующие выводы:

1. Выявлена связь между опасным проявлением ГД и плавным снижением тренда сейсмической активности до точки глобального минимума с последующим изменением в направлении тренда. В таблице 4.3.2 приведены события-предвестники фактических горных ударов. Для каждого из горных ударов были найдены свои события предвестники. Их количество, разница во времени и значения энергий напрямую связаны с вероятностью наступления горного удара.

2. Предложена схема композитной модели анализа данных сейсмического мониторинга. На основании данных сейсмического мониторинга за 2020 год, были получены экспериментальные данные, подтверждающие выявленную связь. Полученные результаты являются

развитием идеи об «использовании методов машинного обучения при работе с математическими моделями временных рядов, которые базируются на данных сейсмического мониторинга» [31].

3. Анализ полученных данных, их распределения во времени и пространстве, выявил даты двух потенциально возможных опасных проявлений ГД 06.04.2020 и 03.12.20, соответственно. При этом фактические даты горных ударов от 08.04.2020 и 04.12.2020 отличаются от предполагаемых на 1–2 дня. В случае с фактическим горным ударом от 06.07.2020 был проведен дополнительный анализ временного ряда сейсмического мониторинга и внесены поправки в прогноз опасных проявлений ГД. Таким образом можно утверждать, что композитная модель, основанная на анализе временного ряда сейсмического мониторинга и прогнозе распределения кластеров- «очагов» сейсмических событий в МГП, является эффективным средством контроля опасных проявлений ГД.

4.5 Анализ данных сейсмического мониторинга за 2021 год

Используем временной ряд длиной 343 дней, ширину окна равную 7 дням (параметр L), а в качестве единиц отсчета-среднее значение всех сейсмических событий за день. В качестве оси X используется временная шкала сейсмического мониторинга, в качестве оси Y шкала величины среднего значения сейсмических событий в МДж.

Поскольку выбрано окно равное 7 дням, получено 7 элементарных матриц, которые описываются собственными значениями матриц, оценивая которые можно предположить их вклад в исходный ряд. Первая элементарная матрица отвечает за тренд среднего значения сейсмической активности. Гипотеза, лежащая в основе методики анализа временного ряда сейсмического мониторинга, говорит о том, что опасному геодинамическому явлению может предшествовать так называемый период «затишья» (плавное снижение тренда) и изменения в направлении тренда. Анализ тренда позволяет явным образом установить связь между потенциальными горными ударами и сейсмической активностью

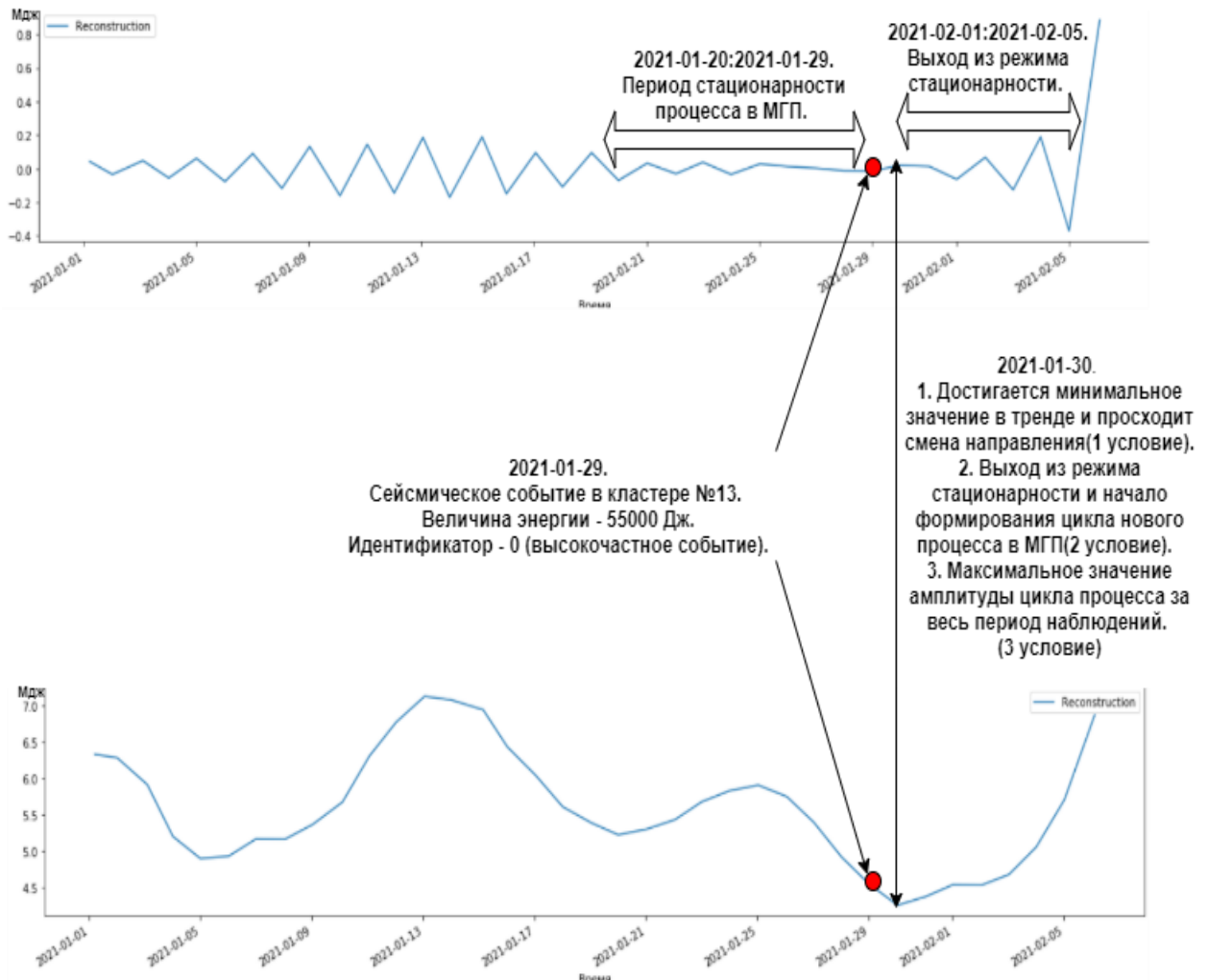


Рисунок 4.3.3 – Периодическая компонента(сверху) и компонента, отвечающая за тренд (рисунок внизу) временного ряда сейсмического мониторинга за январь-февраль 2021 года

Как видно из рисунка 4.3.3, можно выделить 3 таких периода:

- В период с 21 января 2021 года по 29 января 2021 года в МГП идет переход к процессу стационарности. Это обусловлено снижением уровня тренда и стационарным режимом в периодической компоненте;
- 30 января 2021 года достигается глобальный минимум значения тренда сейсмической активности за весь период наблюдений и затем начинается умеренный рост тренда сейсмической активности;
- Вероятность наступления горного удара в следующий 21 день крайне велика, т. к. процесс достиг своего минимального исторического значения и начался рост тренда сейсмической активности.

4.6 Анализ результатов моделирования кластеров сейсмических событий на основе данных сейсмического мониторинга за 2020 год

В результате применения композитной модели были получены следующие значения гиперпараметров. Минимальный диаметр кластера – 24 метра, минимальное число сейсмических событий внутри одного кластера – 12. На рисунке 4.3.4. показаны полученные кластеры, каждый из которых включает в себя не менее 30 сейсмических событий. В результате визуального анализа кластеров, изображенных на рисунке 4.3.4, было сделано предположение о 2 потенциально опасных кластерах сейсмических событий, которые могут являться индикаторами начала формирования процессов деформации в массиве горных пород.

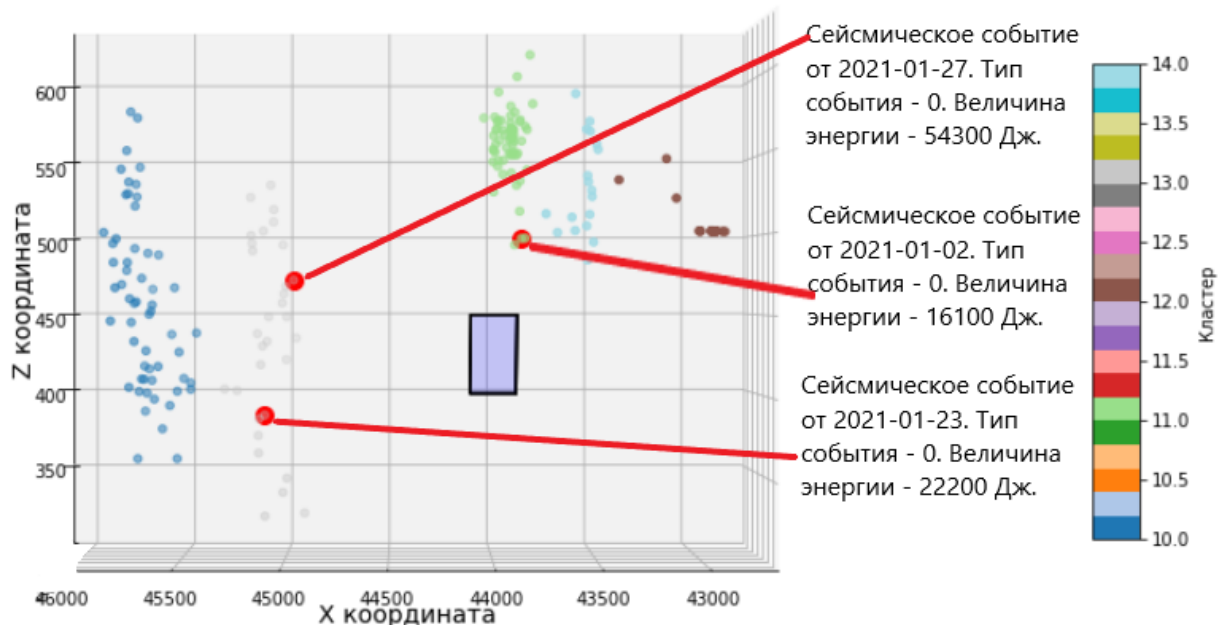


Рисунок 4.3.4 – Выделение всех «микрочластеров» (не более 100 событий в год) сейсмических событий. Фиолетовый параллелепипед- Блок 5/8

Так как сейсмические события произошедшие 23 и 27 января располагаются в значительном расстоянии от зоны ведения горных работ в блоке 5/8, кластером за которым стоит установить наблюдение, является кластер № 13, в котором 2 января 2021 года произошло сейсмическое событие со значением энергии равным 16100 Дж. Все сейсмические события, превышающие выборочное среднее, рассматриваются как сейсмические события способные привести к опасным геодинамическими явлениям. В

таблице 4.3.3. приведены кластеры, которые с высокой долей вероятности могут являться триггерами горных ударов и мониторинг сейсмических событий в этих кластерах в дальнейшем является приоритетным.

Таблица 4.3.3. – Сводная таблица характеристик с сейсмических кластеров

Номер кластера	Количество сейсмических событий за весь период	Количество сейсмических событий за весь период, превышающих 10^4 Дж	Сумма энергии всех сейсмических событий в МДж
11	12	1	0.019
13	29	2	0.09
14	14	-	0.011
10	2	-	0.002

В таблице 4.3.4. приведены сейсмические события, которые могут быть интерпретированы как «предвестники» горных ударов.

Таблица 4.3.4. – Характеристики сейсмических событий выбранных кластеров, превышающих 10^4 Дж

Идентификатор	Номер кластера	Время	Энергия
0	11	2021-01-02 00:36:00	16400
0	13	2021-01-23 12:25:00	22200
0	13	2021-01-27 13:34:00	54300

4.7 Анализ данных сейсмического мониторинга за период с 10.02.2021 по 22.03.2021

Исходя из предположения, сформулированного в предыдущем разделе, вероятность наступления опасного геодинамического события растет и достигает максимального значения 20.02.21. Однако при анализе актуальных данных сейсмического мониторинга в этот период не было зафиксировано сейсмических событий, которые могут являться опасными с точки зрения

геодинамики. Более этого, если проанализировать тренд (рисунок 4.3.5.1) и циклическую компоненты (рисунок 4.3.5.2) временного ряда сейсмического мониторинга можно, сделать вывод, что после точки глобального минимума (29.01.21) тренд плавно достиг локального пика и далее начиная с 25.02.21 стабилизировался вокруг среднесуточного значения 6 Мдж.

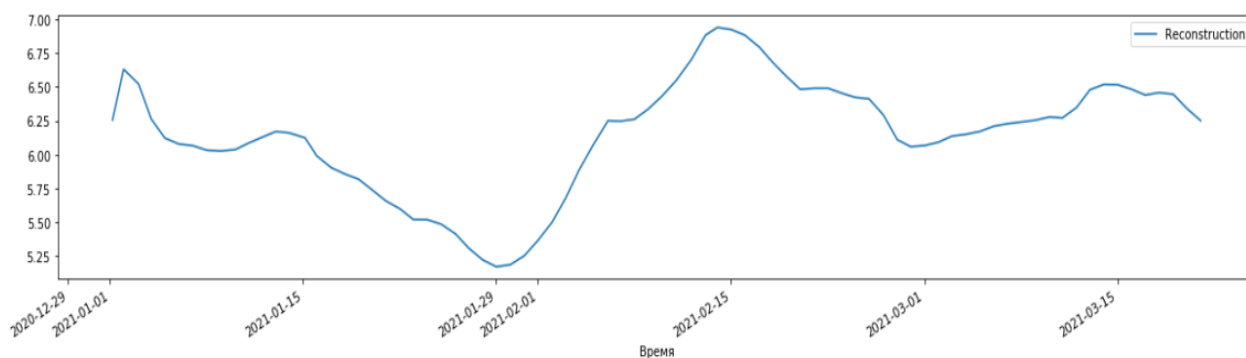


Рисунок 4.3.5.1 – Компонента, отвечающая за тренд временного ряда сейсмического мониторинга за январь-март 2021 года

Визуальный анализ циклической компоненты так же указывает на отсутствие потенциально опасных геодинамических событий.

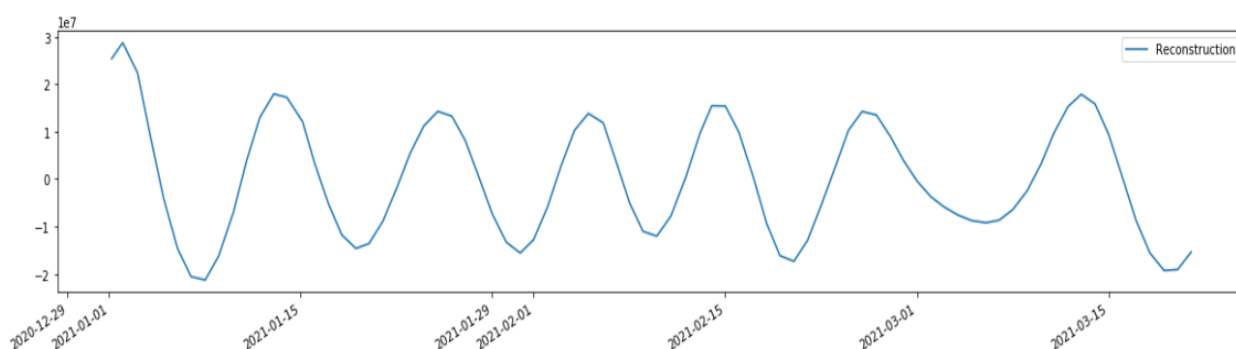


Рисунок 4.3.5.2 – Периодическая компонента временного ряда сейсмического мониторинга за январь-март 2021 года

При этом стоит отметить, что в кластерах сейсмических событий, которые были выявлены в период с января по февраль, появились сейсмические события, которые могут быть интерпретированы как «предвестники» опасных геодинамических событий. Они отображены в таблице 4.3.5. Наиболее интересным для наблюдения являются события от 2021-02-25 мощностью 0,583 МДж и от 2021-03-11 мощностью 1,240 МДж.

Таблица 4.3.5. – Характеристики сейсмических событий выбранных кластеров, превышающих 10^4 Дж

Ид	Номер кластера	Время	Энергия
0	26	2021-02-17 11:57:07	12600
0	14	2021-02-25 20:00:53	583000
1	24	2021-03-11 06:23:36	1240000

Так же, интересной с точки зрения геометрической интерпретации является траектория распространения событий предвестников в МГП.

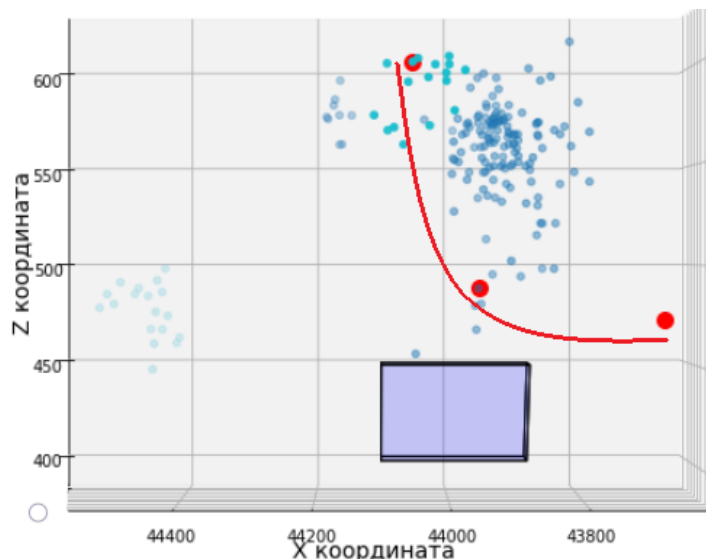


Рисунок 4.3.6 – Наблюдаемые кластеры сейсмических событий и события «предвестники» опасных геодинамических событий (красные точки).

«Предвестники» расположены в порядке убывания по дате сверху вниз. Красная линия-текущая граница распространения событий «предвестников».

4.8 Выводы по главе 4

- Предложены даты возможных горных ударов: 6 февраля 2021 – невысокая вероятность горного удара; 13 февраля 2021 – умеренная вероятность горного удара; 20 февраля 2021 – крайне высокая вероятность горного удара.

- Определено количество кластеров сейсмических событий потенциально опасных для ведения горных работ (2 кластера). Определены их пространственные координаты и размеры.
- В таблице 4 приведены события-«предвестники» предполагаемых опасных геодинамических событий.
- Анализ количества таких событий и их распределения во времени дает возможность предположить о крайне высокой вероятности наступления опасного геодинамического события 20.02.21.
- На данном этапе нет причин ожидать проявления опасных геодинамических событий. Анализ временного ряда и анализа кластеров сейсмических событий подтверждают данное утверждение.
- Предложено ввести границу распространения событий «предвестников», которую следует отслеживать в ходе дальнейшего мониторинга.

ЗАКЛЮЧЕНИЕ

Диссертация представляет собой законченную научно-квалификационную работу, в которой предлагается новое решение актуальной научной задачи - прогноза опасных геодинамических явлений на Кировском руднике КФ АО «Апатит». Выполненные исследования позволяют сделать следующие выводы и рекомендации:

1. Сформировано предположение о связи между минимальным значением тренда сейсмической активности и наступлением опасного геодинамического события. Резкая смена тренда, без наличия периода стационарности тренда, может служить триггером для проявления опасных геодинамических явлений.

2. Сформировано предположение о связи между разницей значений трендов сейсмической активности в точке начала роста и в точке наступления опасного события. Данное предположение согласуется с эвристикой «трех сигм» и его можно использовать в качестве «предвестника» опасных геодинамических событий. Поскольку наступлению опасного периода предшествует период «затишья» длиной от 15 до 30 дней, разработанная методика позволит предпринять меры по предотвращению проявления опасных геодинамических событий.

3. Предложена методика прогноза опасных геодинамических явлений на основе данных сейсмического мониторинга, с помощью методов машинного обучения.

4. На основании проведенных исследований разработан вычислительный программный комплекс алгоритм прогноза опасных геодинамических явлений.

5. На основании данных, полученных в ходе сейсмического мониторинга за 2018 год подтверждена достоверность результатов применения композитной модели прогноза опасных геодинамических явлений. Полученные результаты являются развитием идеи, предложенной в данной работе, об использовании методов машинного обучения при работе с

математическими моделями временных рядов, которые базируются на данных сейсмического мониторинга.

6. Предложена схема композитной модели анализа данных сейсмического мониторинга. На основании данных сейсмического мониторинга за 2020 год, были получены экспериментальные данные. Анализ полученных данных, их распределения во времени и пространстве, выявил даты двух потенциально возможных опасных проявлений ГД 06.04.2020 и 03.12.20 соответственно. При этом, фактические даты горных ударов от 08.04.2020 и 04.12.2020 отличаются от предполагаемых на 1-2 дня.

7. В случае с фактическим горным ударом от 06.07.2020 был проведен дополнительный анализ временного ряда сейсмического мониторинга и внесены поправки в прогноз опасных проявлений ГД. Таким образом можно утверждать, что композитная модель, основанная на анализе временного ряда сейсмического мониторинга и прогнозе распределения кластеров- «очагов» сейсмических событий в МГП, является эффективным средством контроля опасных проявлений ГД.

Перспективы развития темы диссертации преимущественно связаны с использованием современных методов математического и имитационного моделирования с использованием алгоритмов машинного обучения. Поскольку кластеризация сейсмических событий в массиве горных пород и их связь с данными станций деформационного мониторинга базируются на основных положениях теории алгоритмов эволюционной оптимизации, теории временных рядов, теории математической статистик, весьма актуальной и перспективной задачей является построение математических моделей с учетом обратной связи датчиков сейсмического мониторинга.

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

1. Горбунов Г.И. Минеральные месторождения Кольского полуострова / Г.И. Горбунов, И.В. Бельков, С.И. Макиевский – Л.: Наука, 1981. – 272 с.
2. Господариков А.П. О методе обработки данных сейсмического и деформационного мониторинга при ведении подземных горных работ на примере Кикусвумчоррского месторождения АО "Апатит" / А.П. Господариков, К.В. Морозов, **И.Е. Ревин** // ГИАБ. – 2019. – №8 – С. 159–168.
3. **Ревин И.Е.** Композитная модель анализа данных сейсмического мониторинга при ведении горных работ на примере Кикусвумчоррского месторождения АО "Апатит" / **И.Е. Ревин**, А.П. Господариков, Морозов К.В. - DOI: 10.31897/PMI.2021.6.628 // Записки горного института. – 2021. – № 252.
4. Господариков А.П. Программа «SeDef» для обработки данных сейсмического мониторинга и поиска кластеров-очагов сейсмических событий с использованием методов иерархической кластеризации / А.П. Господариков, **И.Е. Ревин** // Свидетельство о регистрации государственной программы для ЭВМ №2021616652, заявл. 16.04.2021, опубл. 23.04.2021.
5. Гладырь А. В. Разработка метода выделения опасных участков в массиве горных пород по данным сейсмоакустических наблюдений / А.В. Гладырь, Г.А. Курсакин, М.И. Рассказов, А.В. Константинов - DOI: 10.25018/0236-1493-2019-08-0-21-32 // Горный информационно-аналитический бюллетень. – 2019. – № 8. – С. 21–32.
6. Журавлева О.Г. Кластеризация сейсмических событий в условиях удароопасных месторождений Хибинского массива / О.Г. Журавлева // Проблемы недропользования. – 2017. – №1 (12).
7. Игнатьев С.А. Современные математические методы прогноза условий поддержания и крепления горных выработок / С.А. Игнатьев, А.Е.

Судариков, А.Ж. Имашев - DOI: 10.31897/PMI.2019.4.371 // Записки Горного института. – 2019. – Т. 238. – С. 371-375.

8. Козырев А.А. Изменения деформаций и наклонов геоструктурного блока в процессе подготовки и реализации техногенного землетрясения / А.А. Козырев, М.М. Каган, К.Н. Константинов, Д.В. Жиров // Записки горного института. – 2012. – т. 199. – С. 230–235.

9. Козырев А.А. Концепция единой системы комплексного геомеханического мониторинга при ведении горных работ в скальных массивах горных пород / А.А. Козырев, Э.В. Каспарьян, Ю.В. Федотова // ГИАБ. – 2016. – №4. – С. 83–91.

10. Кашников Ю.В. Деформационные предвестники техногенных землетрясений при разработке месторождений углеводородов / Ю.А. Кашников, С. Г. Ашихмин, В. Г. Букин, С.В. Гришко, И.В. Гетманов, С.Л. Одинцов, А.В. Горбатиков // ФТПРПИ. – 2011. – № 4. – С. 40–49.

11. Корчак П.А. Методика получения исходных данных для обеспечения сейсмического мониторинга на подземных рудниках ОАО «Апатит» / П.А. Корчак, С.А. Жукова // ГИАБ. – 2014. – №10.

12. Онохин Ф.М. Особенности структуры Хибинского массива и апатит-нефелиновых месторождений / Ф.М. Онохин – Л.: Наука. – 1975. – 106 с.

13. Опарин В.Н. Деструкция земной коры и процессы самоорганизации в областях сильного техногенного воздействия / Отв. ред. Н.Н. Мельников. – Новосибирск: Изд-во СО РАН, 2012. – 632 с

14. Савченко С.Н. Особенности напряженного состояния блочных массивов пород с учетом размеров неоднородностей / С.Н. Савченко // Геодинамика и напряженное состояние недр Земли. – Новосибирск, 2011. – С. 245–250.

15. Турчанинов И.А. Тектонические напряжения и устойчивость горных выработок / И.А. Турчанинов, Г.А. Марков, В.И. Иванов, А.А. Козырев – Л.: Наука. – 1978. – 256 с.

16. Цирель С.В. Закономерности развития техногенной сейсмической активности при ведении горных работ / С.В. Цирель // Записки Горного института. – 20102. – т. 188. – С. 58–62.
17. Bardet, J. P. Finite element analysis of rockburst as surface instability / J. P. Bardet // Computers and Geotechnics. –1989. – vol. 8. – № 3. – pp. 177–193.
18. Chen B.R. “Rock burst intensity classification based on the radiated energy with damage intensity at Jinping II hydropower station, China,” / B.-R. Chen, X.-T. Feng, Q.-P. Li, R.-Z. Luo// Rock Mechanics and Rock Engineering. – 2013. – vol. 48. – pp. 289–303.
19. Calinski, T. A dendrite method for cluster analysis / T. Calinski, J. Harabasz - DOI: 10.1080/03610927408827101 // Communications in Statistics. – 1974. – vol. 3. – pp. 1—27.
20. Черпурко, В. Исследование метода ядерной оценки плотности распределения / В.Черпурко, А.Антонюк, Н.Г.Зюльяева // Надежность. – 2007. – 1(20) . – 4-12.
21. Chen, K. Short-Term Load Forecasting With Deep Residual Networks / K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, J. He - DOI:10.1109/TSG.2018.2844307 // IEEE Transactions on Smart Grid. – 2019. – Vol. 10 –№ 4 – pp. 3943–3952.
22. Campello, R. Density-based clustering / R.Campello, P.Kröger, J.Sander, A.Zimek - DOI: 10. 10.1002/widm.1343 // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.
23. Dowding, C. H. Potential for rock bursting and slabbing in deep caverns / C. H. Dowding. //Engineering Geology. – 1986. – vol. 22. – pp. 265–279.
24. Dickey, D.A. Likelihood ratio statistics for autoregressive time series with a unit root / D.A. Dickey, W.A. Fuller // Econometrica: journal of the Econometric Society. – 1981. – pp. 1057–1072.
25. Elistratov, V.V. Principles of an integrated approach to determining the efficiency of stand-alone wind/diesel power systems / V.V. Elistratov, I.G.

Kudryasheva // Power Technology and Engineering. – 2016. – Vol. 49 – № 6 – pp. 464–467.

26. Ertugrul, Ö.F. Forecasting electricity load by a novel recurrent extreme learning machines approach / Ö.F. Ertugrul // International Journal of Electrical Power & Energy Systems. – 2016. – Vol. 78 – pp. 429–435.

27. Ferdoush, Z. A short-term hybrid forecasting model for time series electrical-load data using random forest and bidirectional long short-term memory / Z. Ferdoush, B.N. Mahmud, A. Chakrabarty, J. Uddin - DOI:10.11591/ijece.v11i1.pp763-771 // International Journal of Electrical and Computer Engineering (IJECE). – 2021. – Vol. 11 – № 1 – p. 763.

28. Golyandina, N. On the choice of parameters in singular spectrum analysis and related subspace-based methods / N. Golyandina // Statistics and Its Interface. – 2010. – Vol. 3. –pp. 259-279.

29. Golyandina, N. Basic singular spectrum analysis and forecasting with R / N. Golyandina, A. Korobeynikov // Computational Statistics & Data Analysis. – 2014. – Vol. 71 – pp. 934–954.

30. Gorban, A. Principal Manifolds for Data Visualisation and Dimension Reduction / A. Gorban, B. Kegl, D. Wunsch, A. Zinovyev. – 2007.

31. Gupta, P.C. A stochastic approach to peak power-demand forecasting in electric utility systems / P.C. Gupta // IEEE Transactions on Power Apparatus and Systems. – 1971. – № 2. – pp. 824–832.

32. .Han, P. Drought forecasting based on the remote sensing data using ARIMA models/ P. Han, P.X. Wang, S.Y. Zhang // Mathematical and computer modelling. – 2010. – Vol. 51 – № 11–12 – pp. 1398–1403.

33. Hastie, T. Principal Curves and Surfaces, Ph.D Dissertation / T. Hastie- DOI:10.21236/ADA148833// – 1984.

34. Hurst, H.E. Long-term storage capacity of reservoirs / H.E. Hurst // Transactions of the American society of civil engineers. – 1951. – Vol. 116 – № 1 – pp. 770–79940. Park, D.C. Electric load forecasting using an artificial neural

network / D.C. Park, M.A. El-Sharkawi, R.J. Marks, L.E. Atlas, M.J. Damborg // IEEE transactions on Power Systems. – 1991. – Vol. 6 – № 2 – pp. 442–449.

35. Kalyan, D. A fast and elitist multiobjective genetic algorithm: NSGA-II/ D.Kalyan, P.Amrit, A.Sameer, T.Meyarivan - DOI: 10.1109/4235.996017 // Evolutionary Computation, IEEE Transactions. – 2002. – Vol. 6. – pp. 182 - 197.

36. Kovalchuk, S. V. Towards management of complex modeling through a hybrid evolutionary identification / S. V. Kovalchuk //Proceedings of the Genetic and Evolutionary Computation Conference Companion. – ACM, 2018. – pp. 255-256.

37. Laurenceau, J. Building efficient response surfaces of aerodynamic functions with kriging and cokriging / J. Laurenceau, P. Sagaut //AIAA journal. – 2008. – T. 46. – № 2. – pp. 498-507.

38. Liu, Y., Gupta H. V. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework / Y. Liu, Gupta H. V //Water Resources Research. – 2007. – T. 43. 78. Lopatoukhin L. I., Yaitskaya N. A. Wave climate of the Caspian sea. The input wind data for hydrodynamical modeling and some results of calculations //Okeanologiya. – 2019. – T. 59. – № 1. – pp. 12-21.

39. Lopez-Ibanez, M. Automatically improving the anytime behaviour of optimisation algorithms / M. Lopez-Ibanez, T. Stützle //European Journal of Operational Research. – 2014. – T. 235. – № 3. – pp. 569-582.

40. Lv X. C. Calibration of SWAN Model for Wave Simulation in Bohai Sea / X. C. Lv //Applied Mechanics and Materials. – 2013. – T. 423. – pp. 1344-1350.

41. Madec, G. NEMO ocean engine / G.Madec // Scientific Notes of Climate Modelling Center. – 2015/ - T. 27. 82. Madsen H. Automatic calibration of a conceptual rainfall–runoff model using multiple objectives //Journal of hydrology. – 2000. – T. 235. – № 3-4. – pp. 276-288.

42. Maier, H. R. Introductory overview: Optimization using evolutionary algorithms and other metaheuristics / H. R. Maier //Environmental modelling & software. – 2018.

43. McPhail, C. Robustness metrics: How are they calculated, when should they be used and why do they give different results? / C. McPhail // *Earth's Future*. – 2018. – T. 6. – № 2. – pp. 169-191.
44. Miller, P.A. Optimization of a sea ice model using basinwide observations of Arctic sea ice thickness, extent, and velocity / P.A. Miller // *Journal of Climate*. – 2006. – T. 19. – № 7. – pp. 1089-1108.
45. Moor, L. P. Proposal of a new autocorrelation function in low wind speed conditions / L. P. Moor // *Physica A: Statistical Mechanics and its Applications*. – 2015. – T. 438. – pp. 286-292.
46. Muller, W. Numerical simulation of rock bursts / W. Muller // *Mining Science and Technology*. – 1991. – vol. 12. – № 1. – pp. 27–42.
47. Murphy, J. M. Quantification of modelling uncertainties in a large ensemble of climate change simulations / J. M. Murphy // *Nature*. – 2004. – T. 430. – № 7001. – C. 768.
48. Neshat, M. A Hybrid Evolutionary Algorithm Framework for Optimising Power Take Off and Placements of Wave Energy Converters / M. Neshat // *arXiv preprint arXiv:1904.07043*. – 2019.
49. Nguyen, A. T. Arctic ice-ocean simulation with optimized model parameters: Approach and assessment / A.T. Nguyen, D. Menemenlis, R. Kwok // *Journal of Geophysical Research: Oceans*. – 2011. – T. 116. – № C4.
50. Nguyen, D. C. H. Framework for computationally efficient optimal crop and water allocation using ant colony optimization / D.C.H. Nguyen // *Environmental Modelling & Software*. – 2016. – T. 76. – pp. 37-53.
51. Nikitin, N.O. Deadline-driven approach for multi-fidelity surrogate-assisted environmental model calibration: SWAN wind wave model case study / N.O. Nikitin // *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. – 2019. – ACM. – pp. 1583–1591.
52. Nikitin, N. O. Evolutionary ensemble approach for behavioral credit scoring / N.O. Nikitin // *International Conference on Computational Science*. – Springer. – 2018. – pp. 825-831.

53. Nikitin, N. O. et al. Statistics-based models of flood-causing cyclones for the Baltic Sea region / N. O. Nikitin // *Procedia Computer Science*. – 2016. – T. 101. – pp. 272-281.
54. . Ortlepp, W. D. Rockburst mechanisms in tunnels and shafts / W. D. Ortlepp, T. R. Stacey // *Tunnelling and Underground Space Technology*. –1994. – vol. 9. – № 1. – pp. 59–65.
55. Osidele, O. O. A random search methodology for examining parametric uncertainty in water quality models / O. O. Osidele, W. Zeng, M. Beck // *Water science and technology*. – 2006. – T. 53. – № 1. – pp. 33-40.
56. Papalexopoulos, A.D. A regression-based approach to short-term system load forecasting / A.D. Papalexopoulos, T.C. Hesterberg // *IEEE Transactions on Power Systems*. – 1990. – Vol. 5 – № 4 – pp. 1535–1547.
57. Park, J.H. Composite modeling for adaptive short-term load forecasting / J.H. Park, Y.M. Park, K.Y. Lee // *IEEE Transactions on Power Systems*. – 1991. – Vol. 6 – № 2 – pp. 450–457.
58. Paenke, I., Branke, J., Jin, Y.: Efficient search for robust solutions by means of evolutionary algorithms and fitness approximation. *IEEE Transactions on Evolutionary Computation* 10(4), 405–420 (2006).
59. Panthadeep, B. A survey of density based clustering algorithms/ B. Panthadeep, M. Pinaki - DOI: 10.1007/s11704-019-9059-3 // *Frontiers of Computer Science*. – 15. 151308.
60. Pearson, K. On lines and planes of closest fit to systems of points in space / K. Pearson // *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. – 1901. – Vol. 2 – № 11 – pp. 559–572
61. Pintér J. Globally optimized calibration of environmental models / J.Pintér // *Annals of Operations Research*. – 1990. – T. 25. – № 1. – pp. 211-221.
62. Ray J. et al. Bayesian calibration of the Community Land Model using surrogates // *SIAM/ASA Journal on Uncertainty Quantification*. – 2015. – T. 3. – № 1. – pp. 199-233.

63. Ricker, R. A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data / R. Rickler // *Cryosphere*. – 2017. – T. 11. – № 4. – pp. 1607-1623.
64. Sander, J. Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications / J. Sander, M. Ester, H. P. Kriegel, X. Xu - DOI: 10.1023/A: 1009745219419 // *Data Mining and Knowledge Discovery*. Berlin: Springer-Verlag. – 1998. – (2). – pp. 169—194.
65. Savard, C. A Suggested Improvement for Small Autonomous Energy System Reliability by Reducing Heat and Excess Charges / C. Savard, E. Iakovleva - DOI:10.3390/batteries5010029 // *Batteries*. – 2019. – Vol. 5 – № 1 –29. p.
66. Su G.S. Rockburst prediction method based on case reasoning pattern recognition” / G. S. Su, X. F. Zhang, and L. B. Yan // *Journal of Mining and Safety Engineering*. –2008. – vol. 25. – No. 1. – pp. 63–67.
67. Shan S. Survey of modeling and optimization strategies to solve highdimensional design problems with computationally-expensive black-box functions / S. Shan, G. G. Wang // *Structural and Multidisciplinary Optimization*. – 2010. – T. 41. – № 2. – pp. 219-241.
68. Simpson, T. Design and analysis of computer experiments in multidisciplinary design optimization: A review of how far we have come-or not / T. Simpson // *ISSMO multidisciplinary analysis and optimization conference*. – 2008. – pp. 5802.
69. Solonen, A. Efficient MCMC for climate model parameter estimation: Parallel adaptive chains and early rejection / A. Solonen // *Bayesian Analysis*. – 2012. – T. 7. – № 3. – pp. 715-736.
70. Sumata, H. A comparison between gradient descent and stochastic approaches for parameter optimization of a sea ice model / H. Sumata // *Ocean Science*. – 2013. – T. 9. – № 4. – pp. 609-630.
71. Sutulo, S. Mathematical models for simulation of manoeuvring performance of ships / S. Sutulo, S. Guedes // *Marine Technology and Engineering*,

G. Soares, C. Garbatov, Y. Fonseca, and AP Teixeira, eds., Taylor & Francis Group, London. – 2011. – pp. 661-698.

72. S. K. Sharan, “A finite element perturbation method for the prediction of rockburst,” *Comput. Struct*, vol. 85, no. 17-18, pp. 1304–1309, 2007.

73. Taylor, J.W. Short-term load forecasting with exponentially weighted methods / J.W. Taylor // *IEEE Transactions on Power Systems*. – 2011. – Vol. 27 – № 1 – pp. 458–464.

74. Tolman, H.L. OpenDA-NEMO framework for ocean data assimilation / H.L. Tolman // *Ocean Dynamics*. – 2016. – T. 66. – № 5. – pp. 691-702.

75. Voevodin, V. V. Supercomputer lomonosov-2: large scale, deep monitoring and fine analytics for the user community / V. V. Voevodin // *Supercomputing Frontiers and Innovations*. – 2019. – T. 6. – № 2. – pp. 4-11.

76. Vychuzhanin, P. Robust Ensemble-Based Evolutionary Calibration of the Numerical Wind Wave Model / P. Vychuzhanin, N. O. Nikitin, A. V. Kalyuzhnaya // *International Conference on Computational Science*. – 2019. – Springer. – pp. 614-627

77. Wang, S. Y. Analytical and numerical study on the pillar rockbursts mechanism / S.Y. Wang, K. C. Lam, S. K. Au, C. A. Tang, W. C. Zhu, and T. H. Yang, // *Rock Mechanics and Rock Engineering*. – 2006. – vol. 39. – № 5. – pp. 445–467.

78. Wheelwright, S. Forecasting: methods and applications / S. Wheelwright, S. Makridakis, R.J. Hyndman. – 1998. – 420.p.

79. Xie, Y. A hybrid short-term load forecasting model and its application in ground source heat pump with cooling storage system / Y. Xie, P. Hu, N. Zhu, F. Lei, L. Xing, L. Xu, Q. Sun // *Renewable Energy*. – 2020. – Vol. 161 – pp. 1244–1259.

80. Wang, C. An evaluation of adaptive surrogate modeling based optimization with two 188 benchmark problems / C. Wang // *Environmental Modelling & Software*. – 2014. – T. 60. – pp. 167-179.

81. Wei, G. Multiobjective adaptive surrogate modeling-based optimization for parameter estimation of large, complex geophysical model / G.Wei// Water Resources Resear. – 2016. – 167.p.
82. Yang, B. Some issues in uncertainty quantification and parameter tuning: a case study of convective parameterization scheme in the WRF regional climate model / B. Yang //Atmospheric Chemistry & Physics. – 2012. – T. 12. – № 5.
83. Williams, J.J. Guidance on setup, calibration, and validation of hydrodynamic, wave, and sediment models for shelf seas and estuaries / J.J. Williams, L.S. Esteves //Advances in Civil Engineering. –2017. – 533.p.
84. Yang, B. Uncertainty quantification and parameter tuning in the CAM5 Zhang-McFarlane convection scheme / B. Yang // Journal of Geophysical Research: Atmospheres. – 2013. – T. 118. – № 2. – pp. 395-415.
85. Yang, L. Automatic calibration of numerical models using fast optimisation by fitness approximation In Neural Networks / L. Yang //IJCNN 2007. –2007. – pp. 1073–1078.
86. Yapo, P. O. Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data / P. O. Yapo, H. V. Gupta//Journal of Hydrologyc1996. – T. 181. – № 1-4. – pp. 23-48.
87. Yiming, Z. Multifidelity Surrogate Based on Single Linear Regression / Z. Yiming, K. Nam, P. Chanyoung // AIAA Journal 56. –2018. –pp. 4944–4952.
88. Zambrano-Bigiarini, M. A model-independent Particle Swarm Optimisation software for model calibration / M. Zambrano-Bigiarini, R. Rojas //Environmental Modelling & Software. – 2013. – T. 43. – pp. 5- 25.
89. Zang, C. A review of robust optimal design and its application in dynamics / C. Zang, M. Friswell, J. Mottershead// Computers & structures. – 2005. – 83. – pp. 315–326.
90. Zitzler, E. SPEA2: Improving the strength Pareto evolutionary algorithm / E. Zitzler, M. Laumanns, L. Thiele //TIK-report. – 2001. – T. 103.

91. Zhao, Z. Directional sharp-point failure mechanism of rocks surrounding underground circular cavities subjected to large-scale failure / X. Guo, Z. Zhao, X. Gao, Z. Ma, and N. Ma // *Mathematical Problem in Engineering*. – 2019. – vol. 19. – № 4. – pp. 19.
92. Zhang, S. J. Mechanism, Warning and Dynamical Control of Rockburst Evolution Process / X. T. Feng, B. R. Chen, C. Q., Zhang, S. J., Li, S. Y. Wu // *Science Press Beijing*. –2013.
93. Zhou, J. Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines / J. Zhou, X. Li, X. Shi // *Safety Science*. –2012. – vol. 50. – № 4. – pp. 629–644.
94. Zhou, J. Classification of rockburst in underground projects: comparison of ten supervised learning methods / J. Zhou, X. B. Li, H. S. Mitri // *Journal of Computing in Civil Engineering*. –2016. – vol. 30. – № 5.
95. Zhu, W. C. Numerical simulation on rockburst of underground opening triggered by dynamic disturbance/ W. C. Zhu, Z. H. Li, L. Zhu, and C. A. Tang // *Tunnelling and Underground Space Technology*. –2010. – vol. 25. – №. 5. – pp. 587–599.