

Министерство науки и высшего образования Российской Федерации
Санкт-Петербургский политехнический университет Петра Великого
Физико-механический институт

Высшая школа теоретической механики и математической физики

Работа допущена к защите

Директор ВШТМиМФ

Д.ф.-м.н., чл.-корр. РАН

_____ А. М. Кривцов

«__» _____ 2023 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

**Разработка математической модели и алгоритма обработки данных в
процессе функционирования автоматизированного робототехнического
комплекса**

по направлению подготовки

01.03.03 «Механика и математическое моделирование»

Направление

01.03.03_03 «Математическое моделирование процессов нефтегазодобычи»

Выполнил

студент гр. 5030103/90301

А. А. Северюхина

Руководитель

доцент ВШИСиСТ, к.т.н.

М. В. Болсуновская

Консультант

по нормоконтролю

Е. А. Хайбулова

Санкт-Петербург

2023

**САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО**
Физико-механический институт
Высшая школа теоретической механики и математической физики

УТВЕРЖДАЮ

Директор ВШТМиМФ

А.М. Кривцов

«__»_____20__г.

ЗАДАНИЕ

**на выполнение выпускной квалификационной
работы**

студентке Северюхиной Анастасии Андреевне, гр. 5030103/90301

1. Тема работы: Разработка математической модели и алгоритма обработки данных в процессе функционирования автоматизированного робототехнического комплекса.
2. Срок сдачи студентом законченной работы: 05.06.2023
3. Исходные данные по работе: актуальные научные публикации по теме работы, экспериментальные данные с датчиков микроклимата.
4. Содержание работы (перечень подлежащих разработке вопросов): анализ подходов к выявлению аномальных значений во временных рядах, прочностной расчет и поиск оптимальной толщины платформы автоматизированного робототехнического комплекса, оценка состояния микроклимата помещения.
5. Перечень графического материала (с указанием обязательных чертежей): геометрическая модель робототехнического комплекса.
6. Консультанты по работе: отсутствуют.
7. Дата выдачи задания 27.02.2023

Руководитель ВКР _____ М. В. Болсуновская, доцент ВШИСиСТ, ИКНиТ

Задание принял к исполнению 27.02.2023

Студент _____ А. А. Северюхина

РЕФЕРАТ

На 43 с., 32 рисунка, 1 таблица, 1 приложение.

КОНЕЧНО-ЭЛЕМЕНТНОЕ МОДЕЛИРОВАНИЕ, ПРОЧНОСТНОЙ РАСЧЕТ, ПОИСК АНОМАЛИЙ В ДАННЫХ

Тема выпускной квалификационной работы: «Разработка математической модели и алгоритма обработки данных в процессе функционирования автоматизированного робототехнического комплекса»

Данная работа посвящена созданию оптимальной модели автоматизированного комплекса с системой датчиков экологического и санитарно-эпидемиологического мониторинга. На устройстве расположены лидар, датчики дистанции, температуры, влажности, барометрического давления, уровня шума, освещенности, запыленности и концентрации опасных газов.

Задачи, решаемые в ходе исследования:

1. Анализ методов расчета прочности конструкций;
2. Прочностной расчет и поиск оптимальной толщины платформы;
3. Анализ подходов к выявлению аномальных значений во временных рядах;
4. Оценка состояния микроклимата помещения.

В результате работы построена модель оптимальной конструкции автоматизированного робототехнического комплекса, проведены эксперименты для исследования корректности работы устройства, выделены участки с отличными от нормы показателями, которые помогут составить более точную карту помещения.

ABSTRACT

On 43 pages, 32 figures, 1 table, 1 application.

FINITE-ELEMENT MODELING, STRENGTH CALCULATION, SEARCH FOR ANOMALIES IN DATA

The topic of final qualifying work: «Development of a mathematical model and an algorithm for data processing in the process of functioning of an automated robotic complex»

This work is devoted to the creation of an optimal model of an automated complex with a system of sensors for environmental and sanitary-epidemiological monitoring. Lidar, sensors of distance, temperature, humidity, barometric pressure, noise level, lightning, dustiness and concentration of dangerous gases are located on the device.

Tasks solved during the study:

1. Analysis the strength calculation methods for constructions;
2. Strength calculation and search for the optimal thickness of the platform;
3. Analysis of approaches to identifying anomalous values in time series;
4. Assessment of the microclimate of the room.

As a result of the work, a model of the optimal structure of an automated robotic complex was built, experiments were conducted to study the correctness of the device, areas with indicators different from the norm were identified, they will help to make a more accurate map of the room.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	5
ГЛАВА 1. ПРОЧНОСТНОЙ РАСЧЕТ КОНСТРУКЦИИ	8
1.1 Конечно-элементное моделирование.....	8
1.2 Проведение прочностного анализа	10
ГЛАВА 2. ПОИСК АНОМАЛИЙ	16
2.1 Методы поиска аномальных значений	16
2.2 Устройство автоматизированного робототехнического комплекса	20
2.3 Визуализация данных эксперимента	24
2.4 Применение методов поиска аномалий в данных	28
ЗАКЛЮЧЕНИЕ	36
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	38
Приложение А. Визуализация результатов экспериментов	41

ВВЕДЕНИЕ

Проблема использования роботов для создания комфортных условий в умных городах становится все более актуальной. Их интеграция в жилые и рабочие пространства позволит повысить эффективность работы городской инфраструктуры.

Основная сложность создания автоматизированных устройств связана с необходимостью выбора безопасной траектории в условиях внешней среды с препятствиями. Информация о текущем состоянии среды, накопленный опыт, а также применение методов анализа данных помогут снизить уровень неопределенности при планировании оптимальной траектории, а также снизить временные затраты на принятие решений [9].

Кроме того, при проектировании устройства необходимо получить оптимальную и простую конструкцию с высоким коэффициентом запаса прочности. Также следует учитывать какая должна быть грузоподъемность механизма, площадь рабочей зоны, с какими ускорениями могут работать отдельные части.

Изготовление некоторых деталей является дорогостоящим, занимает много времени, поэтому предприятию важно знать, как поведет себя изделие под действием различных нагрузок, выдержит ли оно условия эксплуатации. Для этого применяются различные методы анализа прочности конструкций. Прочностные расчеты сейчас применяются во многих отраслях промышленности: атомной, нефтегазовой, автомобильной и др. Они необходимы для установления и исправления слабых мест конструкций.

Программные пакеты помогают проводить анализ напряженно-деформируемого состояния объектов любой сложности при статическом или динамическом действии нагрузок, а также содержат функционал, который поможет рассчитать оптимальные размеры деталей и характеристики их соединений.

Создание автоматизированного комплекса поможет в решении задач охраны труда на предприятиях, где требуется поддержание микроклимата рабочего места в пределах допустимых норм

В соответствии с ГОСТом микроклимат помещения – это состояние внутренней среды, оказывающее воздействие на человека, характеризуемое такими показателями, как температура воздуха и ограждающих конструкций, влажность и подвижность воздуха [21]. Расчетные параметры микроклимата нормируются в зависимости от функционального назначения помещения, интенсивности деятельности и продолжительности пребывания в них людей.

Длительное и систематическое воздействие неблагоприятных параметров окружающей среды на человека может вызвать ухудшение самочувствия и снижение работоспособности.

Очень важной частью для прогнозирования возникновения нештатных ситуаций и анализ данных.

Аналитика данных (Data Science) – одна из передовых и быстро развивающихся отраслей знаний. На сегодняшний день уже разработано большое количество программных платформ, модулей и библиотек, которые реализуют наиболее применяемые алгоритмы анализа данных.

Для работы в области науки о данных очень важно иметь представление об основных понятиях линейной алгебры, математической статистики и теории вероятностей [1].

Основной целью изучения поведения процессов является поиск аномалий. Их обнаружение относится к проблеме поиска экземпляров данных, которые не соответствуют нормальному поведению. Задачи поиска аномалий часто интерпретируются по-разному в зависимости от поставленной цели и характера данных.

Целью данной работы является проведение прочностного анализа модели автоматизированного робототехнического комплекса и создание алгоритма обработки данных экологического мониторинга.

Вследствие чего можно выделить следующие основные задачи, которые решаются в работе:

1) Провести прочностной анализ модели автоматизированного робототехнического устройства и определить оптимальную толщину платформы для создания легкой и прочной конструкции;

2) Изучить методы анализа данных и поиска аномальных значений;

3) Провести эксперимент для проверки корректности работы устройства;

4) Проанализировать полученные данные и составить алгоритм обработки параметров микроклимата, который поможет в создании более точной карты помещения, по которому движется устройство.

ГЛАВА 1. ПРОЧНОСТНОЙ РАСЧЕТ КОНСТРУКЦИИ

1.1 Конечно-элементное моделирование

Для работы с анализом прочности конструкций на практике используют аналитические и численные методы. Как правило, аналитические расчеты позволяют получить решение задач для тел, которые имеют простую геометрию и схему нагружения. Численные решения позволяют получить результаты не зависимо от сложности геометрии и способа приложения нагрузки [7].

Наибольшее распространение при анализе деформаций и напряжений в инженерных конструкциях получил метод конечных элементов. Это способ представления тела с помощью конечного числа элементов малого размера. Метод позволяет упростить модель и дает возможность проводить расчеты с помощью компьютера, получая результаты с небольшим отклонением от реальных данных.

Геометрия модели разбивается на малые элементы, которые могут иметь разные размеры в зависимости от расположения и заданных настроек. Таким образом задача сводится к системе уравнений для связанных точек. Чем на более мелкие элементы разбивается модель, тем более точными получатся результаты расчета.

Конечно-элементное моделирование применяется для:

- расчета конструкций на прочность под действием статических нагрузок;
- модального анализа – расчета собственных частот и анализа вибраций;
- прочностного расчета под действием циклических нагрузок;
- решения задач механики жидкостей;
- теплового анализа [17].

Каждый отдельный элемент модели должен быть достаточно простым, чтобы была возможность легко определить перемещения и напряжения в любой его части по заданным перемещениям узлов. Количество перемещений, однозначно определяющих положение элемента, называют числом степеней свободы. Связь между перемещениями узлов и силами, действующими на них, задается матрицей жесткости. Матрица жесткости системы формируется в единой глобальной системе осей координат.

Кроме расчета воздействия статических нагрузок, важно определить действие динамической нагрузки, изменяющейся во времени. Когда на конструкцию действует переменное возбуждающее воздействие, может происходить резонанс. Если поглощения вибрации практически нет, то можно наблюдать гармонический резонанс, подобный тому, что привел к обрушению Тэкомского моста в 1940 году [15].

Наиболее эффективным способом предотвращения подобной неприятности является создание такой конструкции, собственная частота колебаний которой выше или ниже, чем ее рабочая частота.

1.2 Проведение прочностного анализа

Для данной работы была взята модель автономного робототехнического устройства (рисунок 1.1).

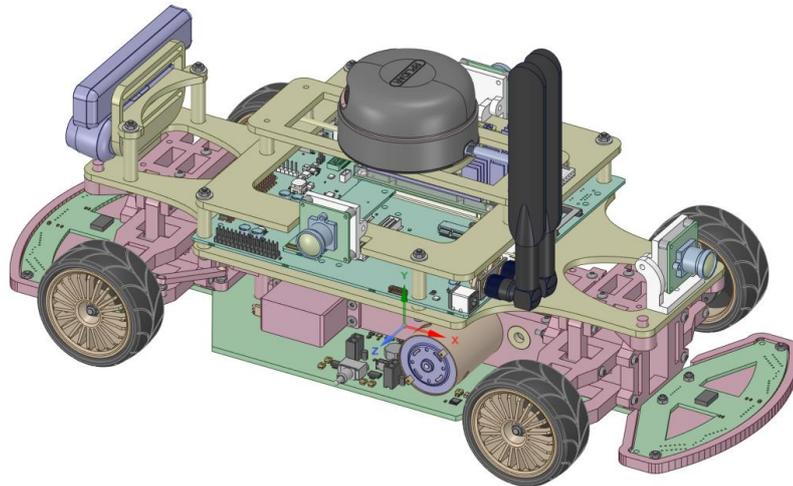


Рисунок 1.1 - Геометрическая модель устройства

Геометрия модели состоит из 1201 детали, поэтому для более простых вычислений модель была упрощена с помощью метода combine, а также с помощью замены взаимодействия деталей эквивалентными нагрузками.

Для соединения деталей болты были представлены соответствующим болтовым соединением. Создана группа контактов типа beam и условие затяжки болта с помощью инструмента Bolt Pretention. Преднатяжение задано силой 100Н [18].

Для того, чтобы задать граничные условия использован инструмент Remote Displacement. В качестве точек приложения заданы центры колес, заблокированы перемещения и разрешены повороты (рисунок 1.2).

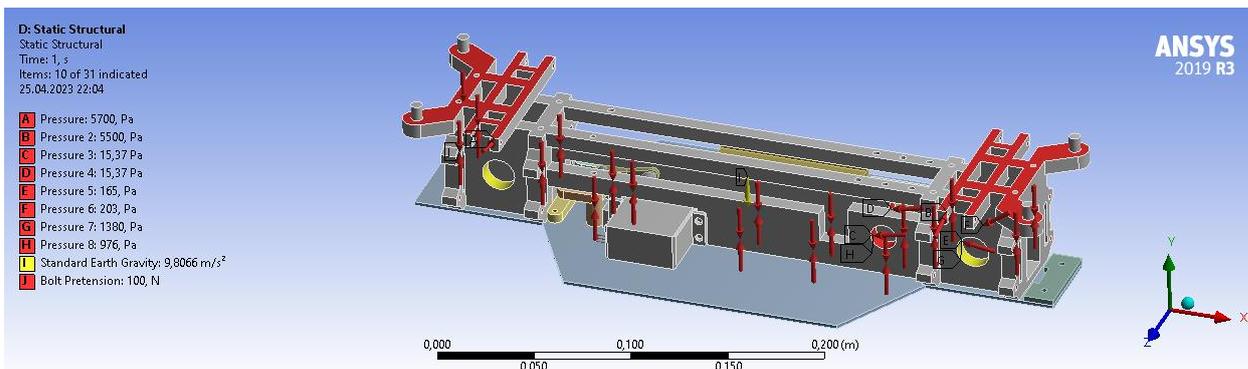


Рисунок 1.2 - Граничные условия и нагрузки на модель

Для иллюстрации полученных результатов были выведены перемещения Total Deformation, напряжения Equivalent Stress и запас прочности Safety Factor.

Пример вывода результатов представлен на рисунке 1.3.

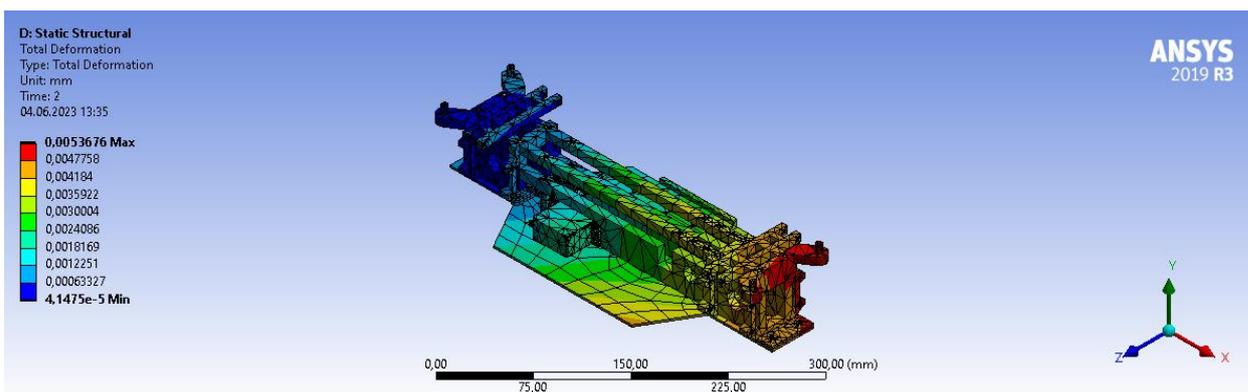


Рисунок 1.3 - Перемещения в модели под действием статической нагрузки при толщине платформы 2.1 мм

Несимметричность полученных результатов связана с несимметричностью нагрузки, действующей на модель.

Для того, чтобы сделать расчеты более точными, необходимо построить более точную расчетную конечно-элементную сетку. Для этого создан новый модуль Static Structural. В геометрии этого модуля оставлена только платформа (рисунок 1.4).

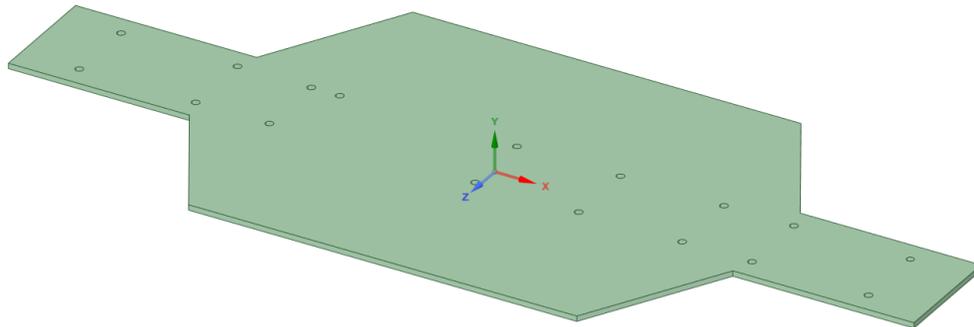


Рисунок 1.4 - Геометрия платформы

Решения из блока Solution были привязаны к блоку расчетов Setup нового модуля. Для того, чтобы задать нагрузки был использован метод Submodeling.

Для правильных расчетов была задана гравитация Земли Standard Earth Gravity и включен режим Large Deflection.

Для иллюстрации полученных результатов были выведены перемещения Total Deformation, напряжения Equivalent Stress и запас прочности Safety Factor.

Пример вывода результатов представлен на рисунке 1.5.

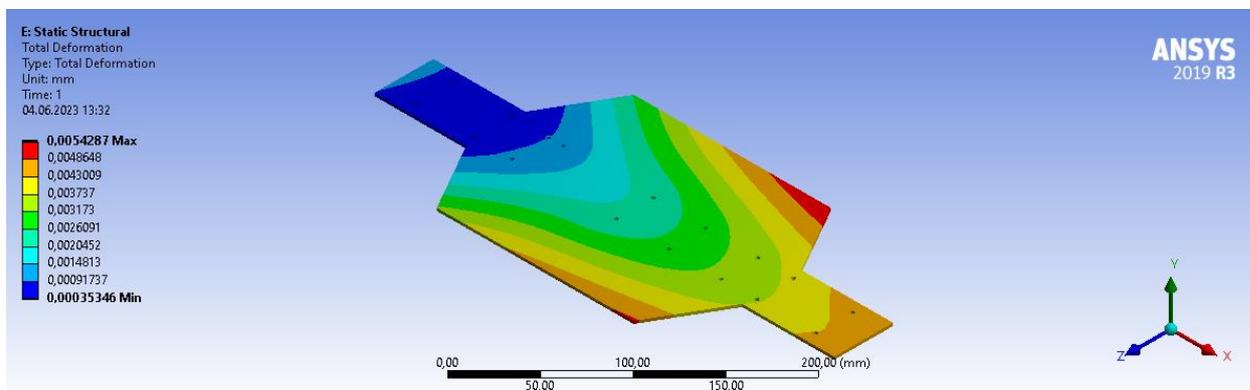


Рисунок 1.5 - Перемещения в платформе толщиной 2.1 мм

Далее было необходимо определить собственные частоты колебания платформы. Для этого создан модуль Modal. Решения из блока Solution были привязаны к блоку расчетов Setup нового модуля.

Результаты расчетов 5 первых собственных частот для толщины платформы 2.1 мм приведены на рисунках 1.6 – 1.10 .

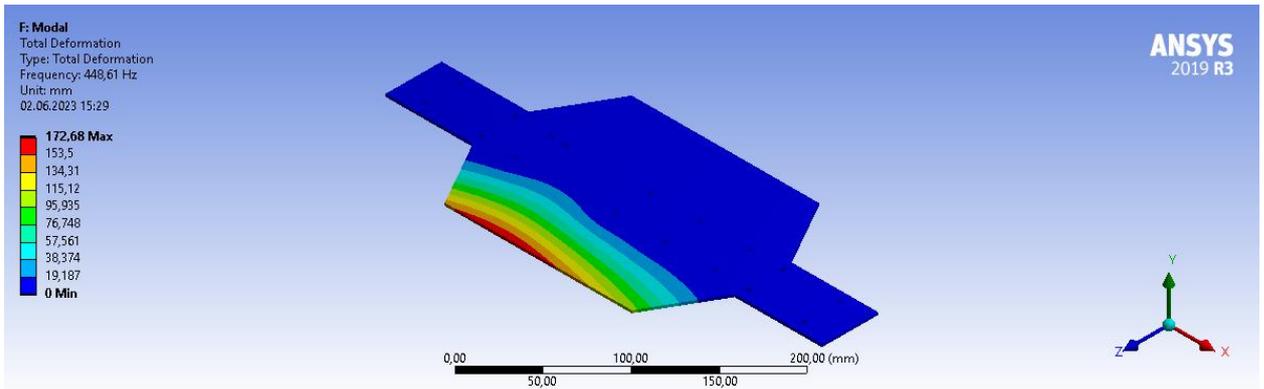


Рисунок 1.6 - Первая собственная частота платформы толщиной 2.1 мм

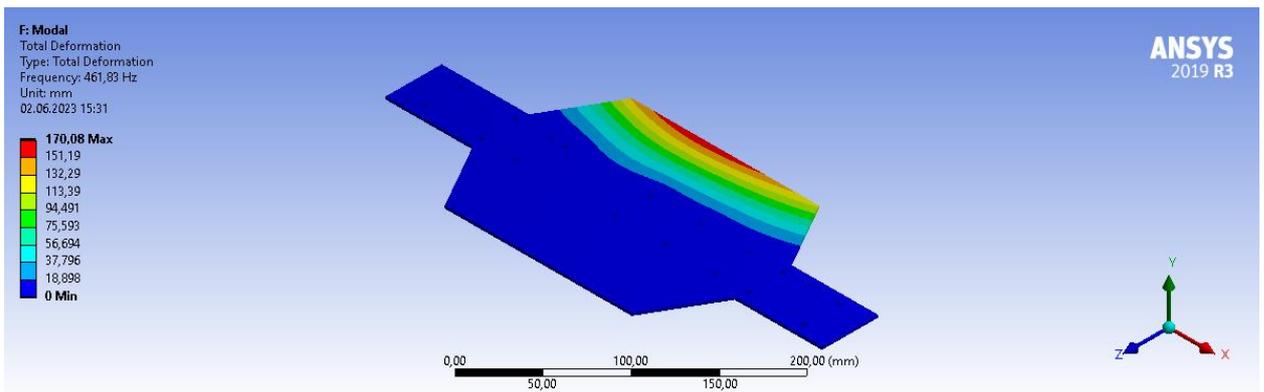


Рисунок 1.7 - Вторая собственная частота платформы толщиной 2.1 мм

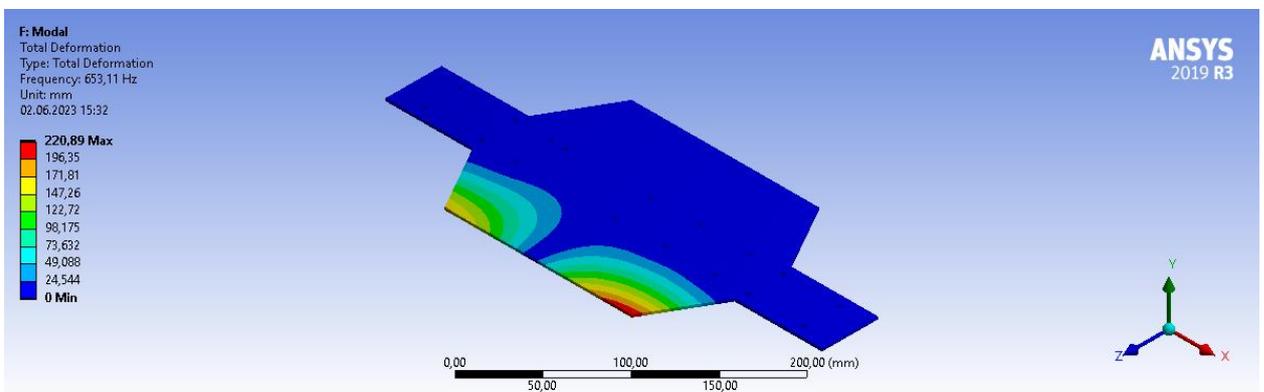


Рисунок 1.8 - Третья собственная частота платформы толщиной 2.1 мм

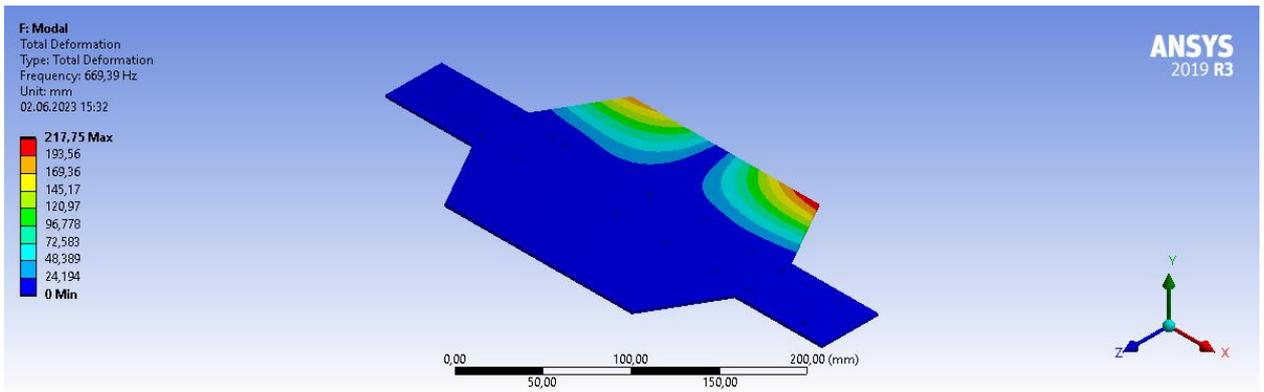


Рисунок 1.9 - Четвертая собственная частота платформы толщиной 2.1 мм

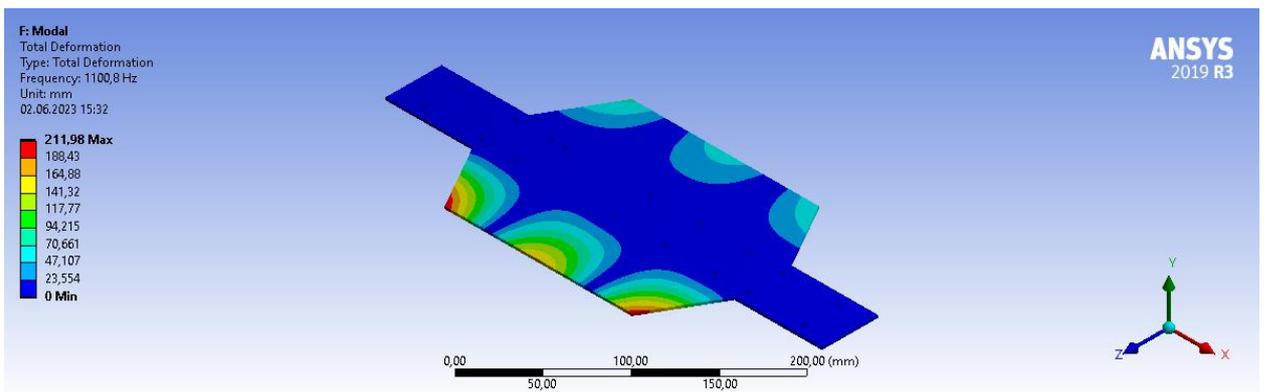


Рисунок 1.10 - Пятая собственная частота платформы толщиной 2.1 мм

Полученные результаты также выводились параметрически (Таблица 1). В качестве входных параметров задана толщина платформы. В выходных параметрах: максимальные перемещения, максимальные напряжения, коэффициент запаса прочности и первые три собственные частоты платформы.

Таблица 1 - Результаты расчетов при различной толщине платформы

Толщина платформы мм	Максимальные перемещения $m \cdot 10^{-6}$	Максимальные напряжения $Pa \cdot 10^7$	Запас прочности	Первая собственная частота, Hz	Вторая собственная частота, Hz	Третья собственная частота, Hz
3	4,8755	3,5429	6,6472	634,15	652,92	923,77
2.9	4,8622	3,7829	6,6087	613,78	631,9	894
2.8	4,8578	3,8244	6,537	593,31	610,82	864,14
2.7	4,8534	3,8651	6,4681	572,8	589,69	834,21
2.6	4,9033	3,9051	6,4019	552,23	568,52	804,2
2.5	4,9839	3,9440	6,3387	531,61	547,28	774,13
2.4	5,0745	3,9819	6,2784	510,94	526	743,98
2.3	5,1772	4,0185	6,2212	490,22	504,66	713,76
2.2	5,2944	4,0538	6,167	469,44	483,27	683,47
2.1	5,4287	4,0877	6,116	448,61	461,83	653,11
2	5,5708	3,7327	5,9736	428,37	440,99	623,56

Для того, чтобы сравнить минимальную собственную частоту платформы был проведен расчет частоты периодической нагрузки вращения шасси.

Диаметр колес составляет 66 мм, соответственно длина окружности шины составляет примерно 207 мм. Максимальная скорость передвижения робототехнического комплекса составляет 15 см/сек, с помощью нее можно найти частоту вращения колес, она равна примерно 434 Гц.

Таким образом, было определено, что оптимальная толщина платформы составляет 2.1 мм. Ее минимальная собственная частота выше максимальной частоты действующей нагрузки. Коэффициент запаса прочности находится в зоне упругих деформаций, значит в детали не возникает остаточных деформаций.

ГЛАВА 2. ПОИСК АНОМАЛИЙ

2.1 Методы поиска аномальных значений

Методы анализа данных позволяют оценить и выявить тенденции временных рядов, оценить влияние различных факторов на изменения, а также проанализировать корреляцию значений и выявить зависимости между ними.

Анализ данных включает три основных этапа (рисунок 2.1) [6]:



Рисунок 2.1. Этапы анализа данных

Во многих областях исследований данные представляют собой последовательность наблюдений с различных датчиков. Поэтому один из наиболее важных методов в аналитике данных – это анализ временных рядов [14].

Временной ряд представляет собой последовательность значений, описывающих протекающий во времени процесс, измеренных в последовательные моменты времени, обычно через равные промежутки.

Выявление аномалий в поведении системы позволяет повысить качество процессов, а также предотвратить нештатные ситуации и аварии на ранних этапах.

Аномалия – это отклонение поведения системы от стандартного (ожидаемого). Они могут возникать в результате действия факторов внешней среды, аварий, технических сбоев, преднамеренных взломов и т.д.

Существуют разные виды аномалий:

- выброс;
- сдвиг;
- изменение характера распределения значений;

- отклонение от «повседневного» (для данных с сезонностью);
- совместные аномалии.

Их можно распределить в три основных типа:

1) Точечные аномалии (выбросы).

Отдельных экземпляр рассматривается как аномальный по отношению к остальным данным. Другими словами, это редкое событие.

Например, можно рассмотреть набор данных по финансовым транзакциям. Если при анализе расходов наблюдается потраченная сумма, которая очень сильно отличается по сравнению с нормальным диапазоном переводов, то такая операция будет точечной аномалией [8].

2) Контекстуальные (условные) аномалии.

Экземпляр является аномальным лишь в определенном контексте. Такие аномалии обычно носят временный характер.

Например, увеличение числа покупателей в магазинах обычно происходит в период отпусков, поэтому внезапный рост в период праздников или распродаж можно читать контекстуальными аномалиями [20].

3) Коллективные аномалии.

Последовательность связанных экземпляров данных является аномальной по отношению к целому набору данных.

Например, большие технологические компании, как правило, развиваются и увеличивают размеры прибыли, но если сразу несколько компаний показывают снижение выручки за один и тот же промежуток времени, то можно определить это как коллективную аномалию.

Процесс выявления аномалий может проводиться для данных различного формата: поток данных (в реальном времени) или архив данных [13].

Существует множество методов, которые помогают определить аномалии. Для того, чтобы выбрать оптимальный, необходимо разобраться в чем заключаются эти алгоритмы.

1) Статистические тесты.

Как правило, применяются для отдельных признаков и отлавливают экстремальные значения. Для этого используют, например, Z-value или Kurtosis measure.

2) Модельные тесты.

Строится модель, которая описывает данные. Точки, которые сильно отклоняются от модели и есть аномалии. Такие методы хороши для определения новизны, но хуже работают при поиске выбросов.

3) Итерационные методы.

Методы состоят из итераций, на каждой из которых удаляется группа «подозрительных» объектов. Например, в n-мерном пространстве можно удалять выпуклую оболочку точек-объектов. Как правило, такие методы достаточно трудоёмки.

4) Метрические методы.

Для данных методов определяется метрика в пространстве объектов, которая помогает найти аномалии.

Хорошей мерой аномальности может служить алгоритм поиска расстояния до k-го соседа. Он отличается простотой реализации и возможностью вводить дополнительные параметры [11].

5) Методы подмены задачи.

Основная идея таких методов заключается в том, чтобы решить новую задачу методами, ориентированными на уже решенные задачи. Например,

можно свести задачу к кластеризации, тогда маленькие кластеры, скорее всего, состоят из аномалий.

6) Методы машинного обучения.

Задачи нахождения аномалий могут быть сведены к задачам машинного обучения.

Например, в системах, где граница между нормальным поведением системы и аномальными значениями является довольно размытой, часто применяется метод опорных векторов. В нем рассматривается удаление экземпляров аномалий от центра масс нормального интервала. Описание данного подхода к решению задач поиска аномальных данных приведено в статье [12].

7) Ансамбли алгоритмов.

Так как задача поиска аномалий - выбор признаков, которые характеризуют те или иные отклонения от нормы, ансамбли строят, пытаясь угадать хорошие пространства. Здесь популярны следующие алгоритмы:

Feature Bagging – для каждого алгоритма берут случайное признаковое подпространство,

Rotated Bagging – в выбранном случайном признаковом подпространстве совершают случайный поворот [16].

В зависимости от применяемого алгоритма обработки данных результатом поиска аномальных значений может быть метка экземпляра как аномального или оценка степени вероятности, что экземпляр является аномальным.

2.2 Устройство автоматизированного робототехнического комплекса

Для того, чтобы получить данные для исследований было проведено несколько экспериментов.

Чтобы проверить правильность сбора данных и выявить аномальные зоны помещения было запущено автономное робототехническое устройство, которое перемещалось по заданной траектории.

В состав системы экологического и санитарно-эпидемиологического мониторинга входят следующие датчики:

- датчик освещенности (TSL2561);
- датчик наличия пылевой взвеси в воздухе (PPD42NS);
- датчик температуры, влажности и барометрического давления (BME280);
- датчик утечки опасных газов (в том числе сжиженных);
- датчик уровня громкости звука;
- лидар, стереокамера, датчики дистанции и одометрии для составления карты помещения.

Расположение экологических датчиков на устройстве можно увидеть на рисунке 2.2.

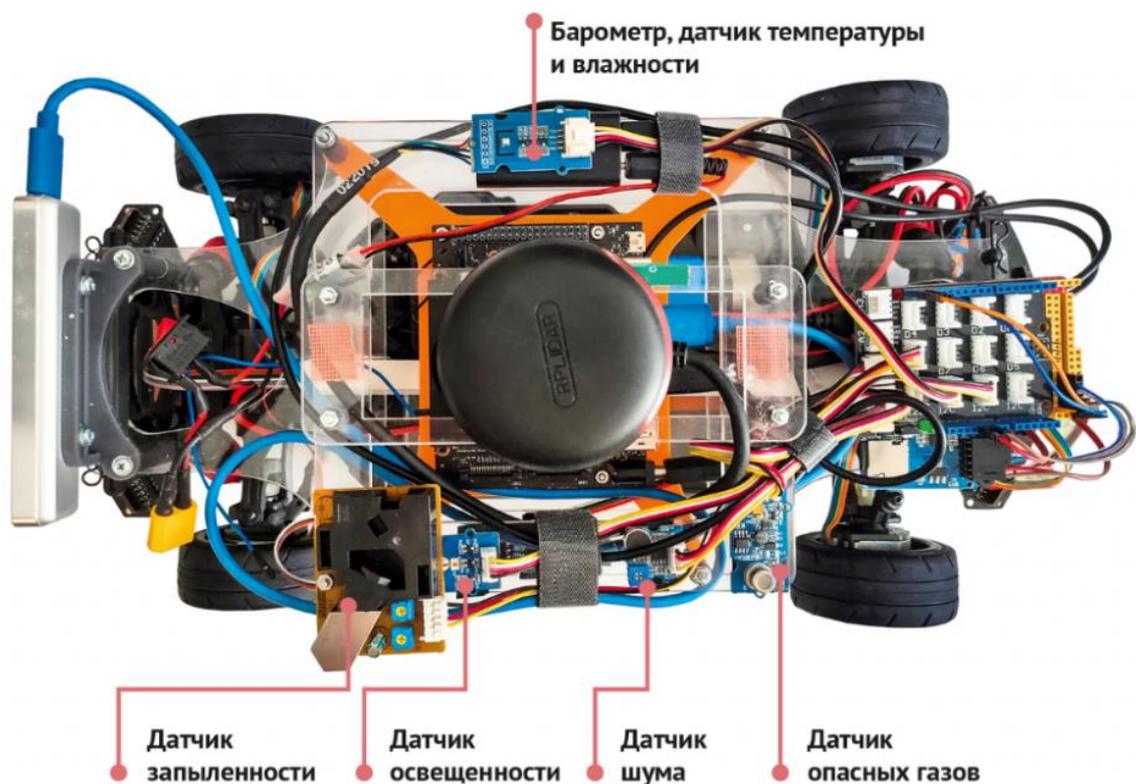


Рисунок 2.2 - Размещение датчиков на устройстве

Датчики работают на частоте до 20 Гц. Их рабочая температура находится в промежутке 0...+45°C, влажность воздуха должна быть менее 95%, за исключением работы датчика температуры, влажности и барометрического давления, которые измеряют данные в диапазоне - 40...+85°C, 0...100% и 300...1100 гПа соответственно.

Данные с датчиков обрабатываются и сохраняются в памяти устройства. По завершении маршрута показания датчиков, связанные с координатами точки, в которой производились замеры, выгружаются в виде CSV-файла и отдельного файла карты помещения (рисунок 2.3).

	Pose X	Pose Y	Pose Z	Lightning, Lx	Air quality PPM	Temperature C	Humidity %	Pressure Pa	Altitude m	Loudness	Dust PPM
0	-0.190054	0.000118	-0.057564	242	8	24.56	16.0	103571.0	-185.31	125	NaN
1	-0.195088	-0.009187	-0.057946	242	8	24.58	16.0	103573.0	-185.47	96	NaN
2	-0.197101	-0.010439	-0.058829	242	8	24.58	16.0	103573.0	-185.47	110	NaN
3	-0.196405	-0.011524	-0.058319	242	8	24.58	16.0	103573.0	-185.47	100	NaN
4	-0.157417	-0.010931	-0.057799	242	8	24.60	16.0	103579.0	-185.96	103	NaN
...
237	0.060567	1.121930	-0.058870	187	7	28.45	14.0	103575.0	-185.63	109	NaN
238	0.012980	1.125270	-0.059238	187	7	28.41	14.0	103574.0	-185.55	110	1776.45
239	0.002212	1.132000	-0.056255	172	7	28.36	14.0	103570.0	-185.23	109	NaN
240	-0.138197	1.280610	-0.055382	126	7	28.34	14.0	103571.0	-185.31	138	NaN
241	-0.578750	1.208200	-0.068505	172	7	28.28	14.0	103574.0	-185.55	113	NaN

Рисунок 2.3 - Представление полученных данных в табличном виде

Карта помещения строится на базе алгоритмов SLAM (Simultaneous Localization and Mapping). Данный метод используется для определения местоположения и ориентации автономных устройств на неизвестной местности, а также, для обновления известных карт окружающего пространства.

На устройстве находятся лидар, датчики дистанции и одометрии, которые постоянно получают информацию о расстоянии до объектов на пути следования робота.

Максимальное расстояние, которое видит лидар - 8 метров. Если дистанция превышает эту отметку, то возвращается infinity, если возвращается число, значит на данном расстоянии находится препятствие, а все что расположено между точкой отсчёта и препятствием будет считаться свободным пространством.

Данные карты хранятся в виде матрицы, где в качестве значений выступают целочисленные эквиваленты понятий «точка занята препятствием», «точка не исследована» и «точка исследована и свободна».

Шасси устройства совершают 2 оборота в секунду. После каждого оборота поступает информация о расстоянии до всех объектов вокруг, то есть двумерное "облако точек". В процессе движения, ориентируясь на показания

датчиков скорости в каждом из колёс и данные, полученные с лидара, информация преобразуется в единую карту помещения.

По ходу движения происходит заполнение матрицы карты необходимыми числами. Специальный пакет `frontier explorer` получает информацию о карте и создает точки следования (фронтиры) с помощью поиска «в ширину». Поиск фронтиров выполняется каждую секунду.

Наиболее выгодное на данный момент направление перемещения определяется в зависимости от того, сколько клеток неисследованной области вошло во фронтир, на каком расстоянии и под каким углом от машины он находится [19].

На рисунке 2.4 показан пример карты помещения, созданной при проведении экспериментов.

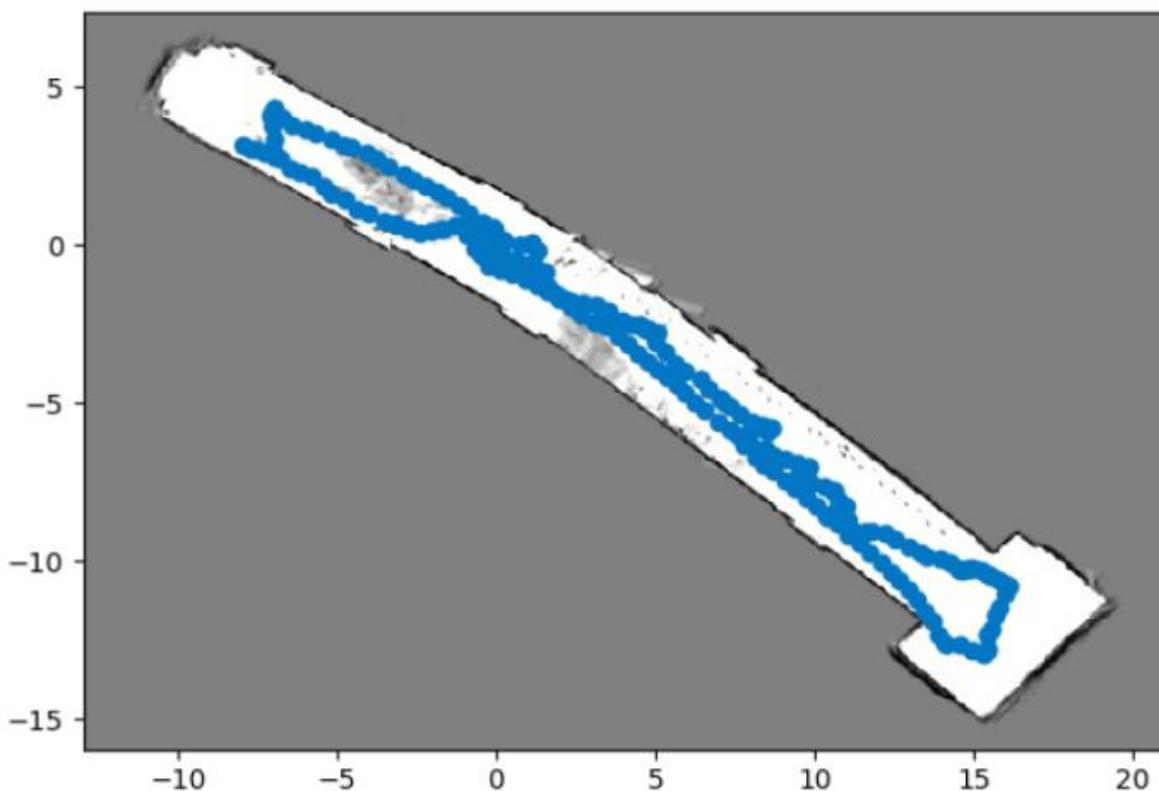


Рисунок 2.4 - Карта помещения

2.3 Визуализация данных эксперимента

В первом эксперименте, для получения данных, искусственного влияния на изменения параметров не проводилось.

Далее, в процессе движения устройства создавались определенные условия, которые могут повлиять на данные:

- Изменение уровня шума.

Для этого в нескольких местах помещения при подъезде устройства была включена громкая музыка.

- Изменение температуры.

Для проверки работы датчика температуры, в определенных момент эксперимента было открыто окно, чтобы температура воздуха изменилась.

- Изменение влажности.

Перед проведением эксперимента воздух в помещении был увлажнен с помощью пульверизатора для воды.

- Изменение освещенности.

В нескольких местах помещения проведено воздействие на параметр освещенности. Для увеличения значений – дополнительная подсветка фонарем, для уменьшения – длинный лист картона над местом, где проезжало устройство.

- Изменение концентрации опасных газов.

В определенный момент в помещение в небольшом количестве была впрыснута газовая смесь, содержащая в составе опасный газ.

Исследуя полученные данные, необходимо будет определить корректность сбора данных, а также выделить участки, где наблюдались отличные от нормы показатели. Такие участки покажут наличие определенных

предметов и содержание веществ в воздухе и помогут составить более точную карту помещения, по которому движется устройство.

Так как датчики имеют определенные диапазоны рабочих параметров и измерений, исследование данных помогает определить корректность сбора информации о микроклимате помещения. Определение тенденций распределения и изменения параметров поможет предотвратить поломку датчиков в случае превышения показателей его рабочей среды.

Визуализация данных позволяет быстрее и эффективнее анализировать полученные результаты.

Графическое представление данных одного из экспериментов представлено на рисунке 2.5:

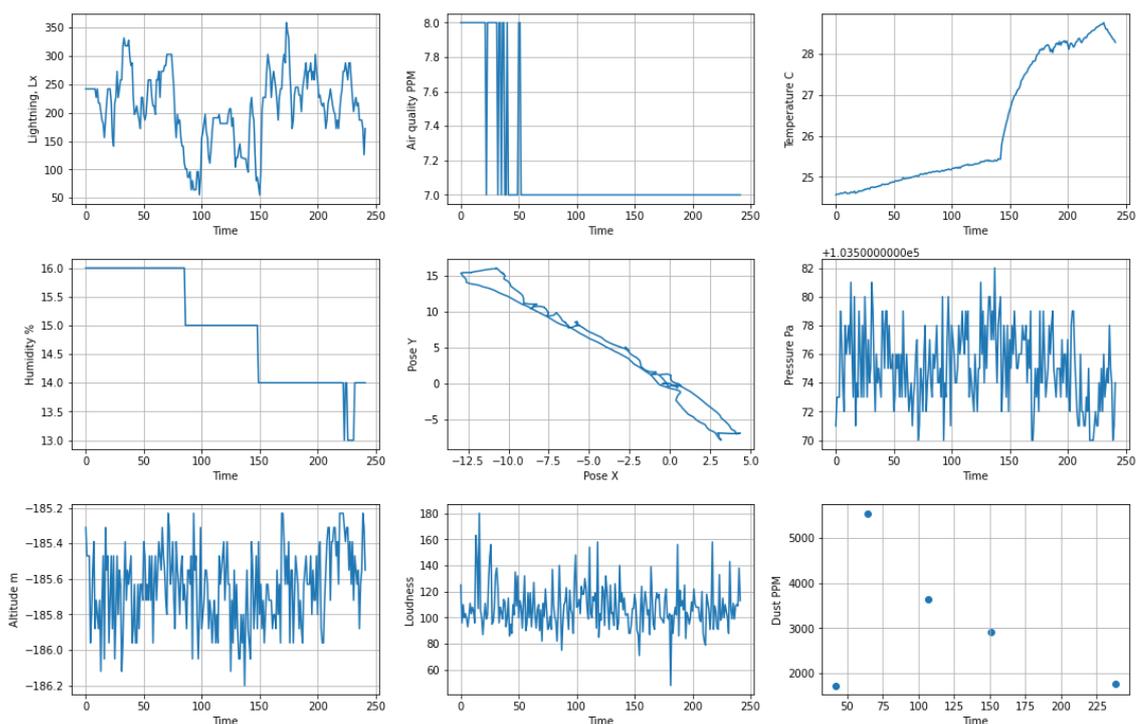


Рисунок 2.5 - Визуализация данных эксперимента

Визуализация данных других экспериментов приведена в приложении 1.

Для того, чтобы посмотреть отличия в данных, полученных для определенного параметра микроклимата можно посмотреть несколько распределений, полученных при различных экспериментах в одном и том же помещении.

На рисунках 2.6–2.8 представлены данные освещенности, полученные при проведении трех различных экспериментов, в одном из которых в помещении были поставлены дополнительные источники света.

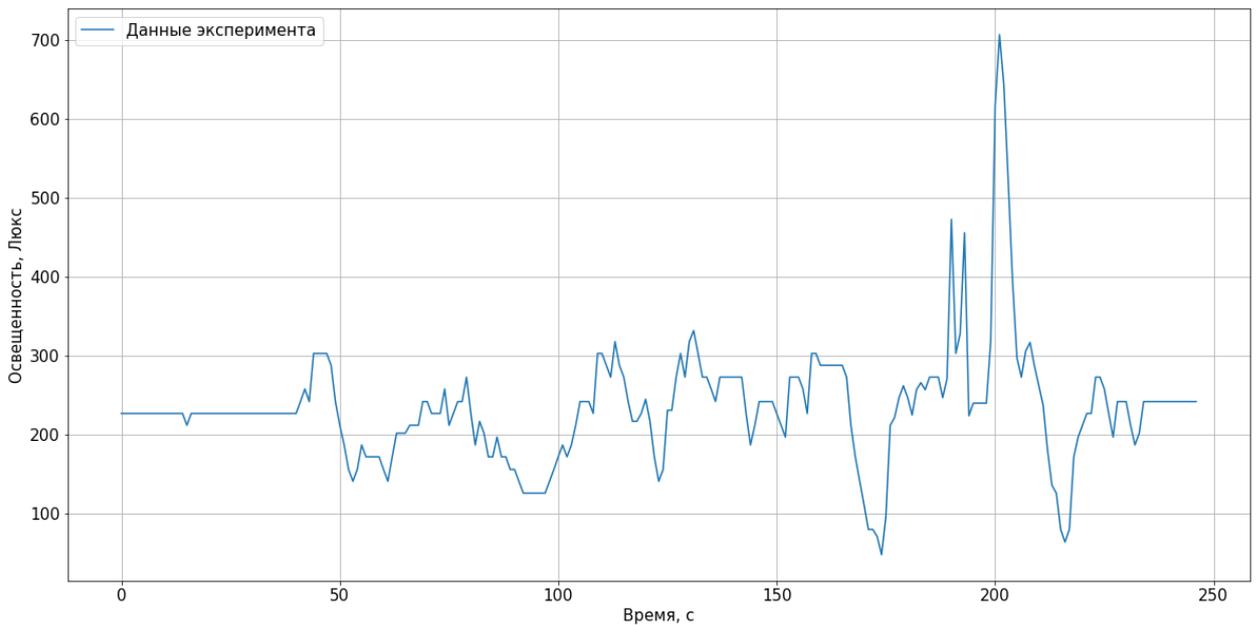


Рисунок 2.6 - Визуализация данных освещенности в эксперименте без влияния на данные

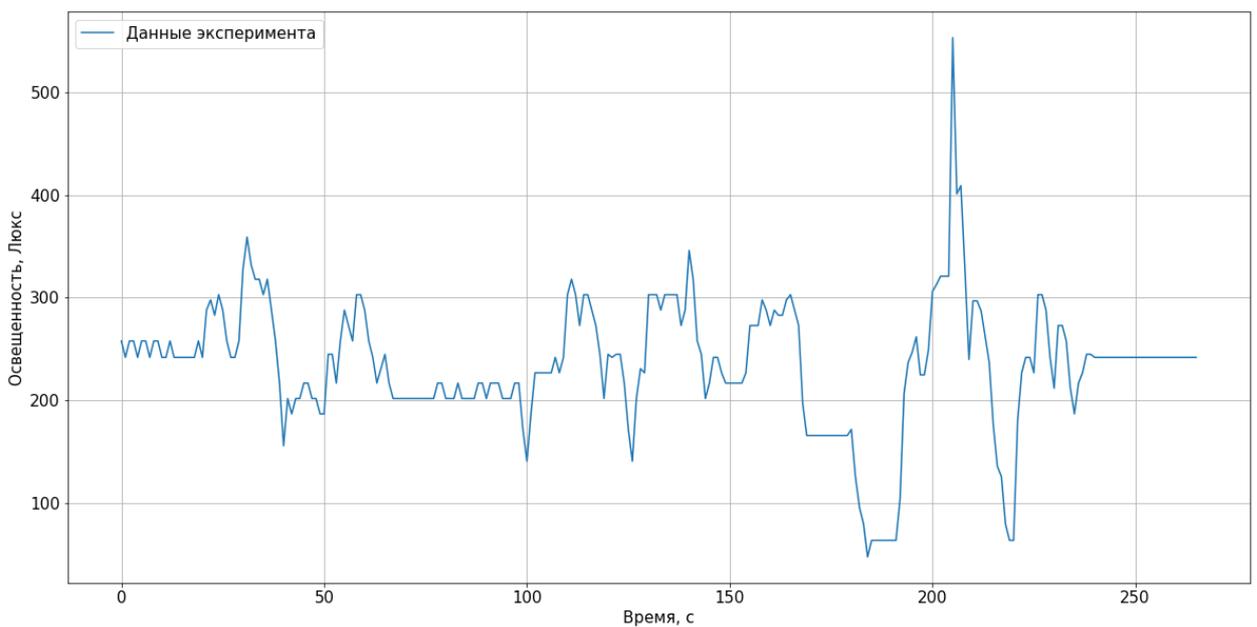


Рисунок 2.7 - Визуализация данных освещенности в эксперименте с влиянием на шум и влажность

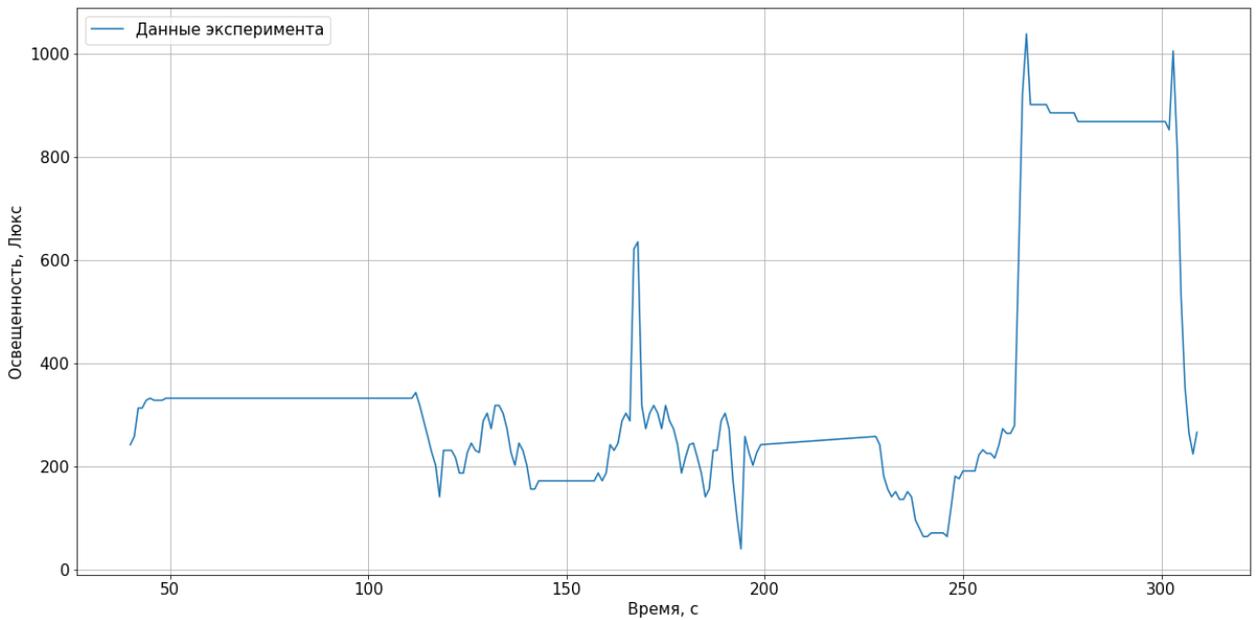


Рисунок 2.8 - Визуализация данных освещенности в эксперименте с влиянием на изменение освещенности

Данные графики показывают чувствительность датчика освещенности на действие внешних факторов. Полученные данные помогут определить местоположение стационарных источников освещения, появление нового источника света и исключить выбросы.

2.4 Применение методов поиска аномалий в данных

К анализу данных относятся разные вычислительные методы обработки данных, в том числе и математико-статистические. Задача исследований при этом заключается в объяснении распределений, порожденных некоторым механизмом, прогнозировании и принятии решений [3].

Различные методы теории вероятности и математической статистики играют большое значение для обработки данных. Эти методы применяются для анализа связи между факторами, измерения ошибки выборки, оценки корректности полученных результатов.

Рассмотрим основные методы статистической обработки данных.

- Среднее арифметическое значение.

Значение, которое определяется как равное сумме числовых значений каждого наблюдения, деленное на общее число наблюдений, представлено в формуле (1) [2].

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n) = \frac{\sum x_i}{n} \quad (1)$$

- Медиана.

Это значение признака, которое делит статистическую совокупность на две равные части: половина единиц совокупности имеет значения признака не меньше медианы, другая половина - значения признака не больше медианы [4].

На рисунке 2.9 приведены результаты вычисления среднего значения, усеченного среднего значения для 95 перцентиля и медианы для анализа данных изменения уровня освещенности в помещении с течением времени.

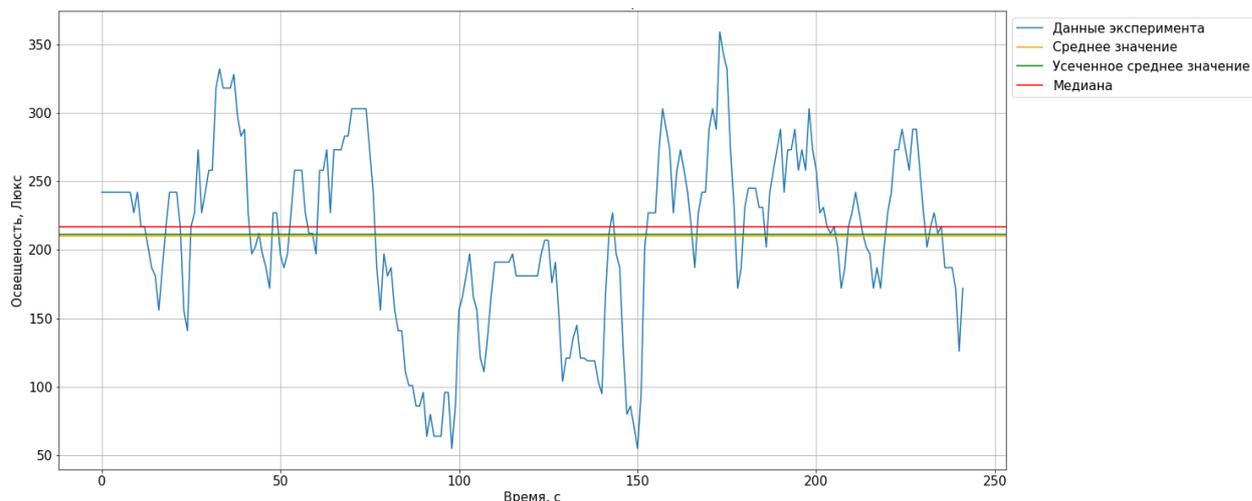


Рисунок 2.9 - График изменения освещенности со временем, величины среднего и усеченного среднего значения и медианы

Используя знания о среднем значении показателей микроклимата в эксперименте, можно определить аномальные зоны. На рисунке 2.10 представлен анализ данных, полученных с датчика освещенности. Зеленым цветом отмечены области, где наблюдалось увеличение уровня освещенности на целом промежутке. Красным цветом отмечено краткосрочное увеличение уровня освещения – такое отклонение является выбросом в данных.

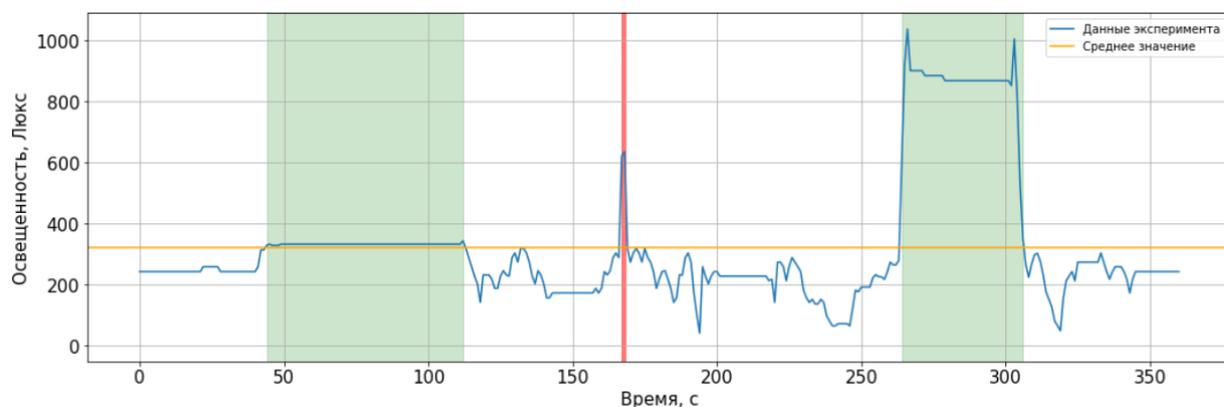


Рисунок 2.10 - Изменение освещенности со временем

На рисунке 2.11 можно увидеть другой тип визуализации данных, на котором видны области увеличения освещенности. В ходе эксперимента в данных частях помещения действительно находились объекты, которые могут повлиять на распределение данных – окно и фонарь.

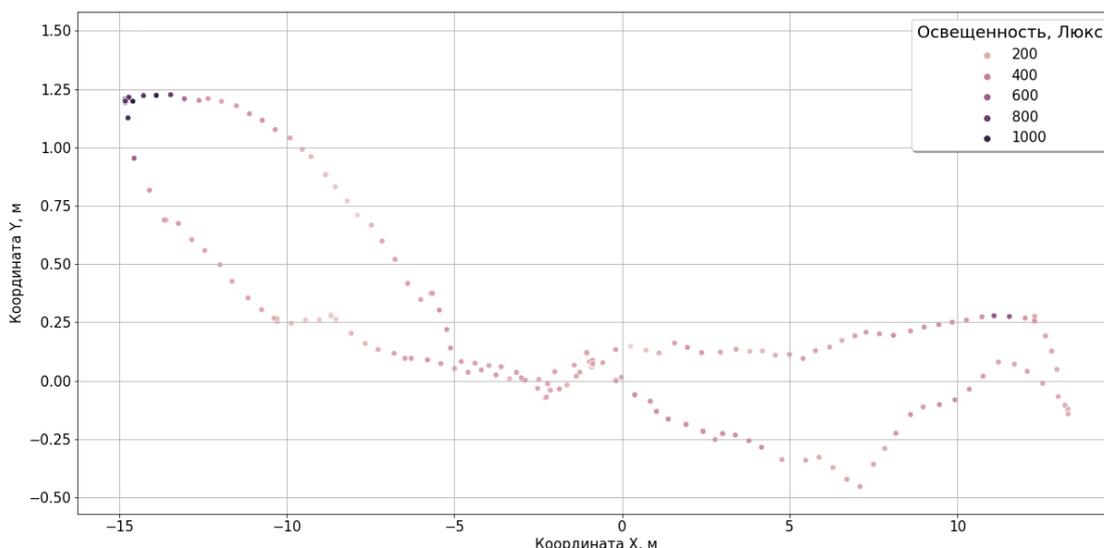


Рисунок 2.11 - Изменение освещенности в пространстве

Представленные графики изменения освещенности показывают, как различные типы визуализации данных, а также применение методов математической статистики позволяют увидеть аномалии в данных и отличить действие внешних факторов на изменение параметров от выбросов.

- Стандартное отклонение (или среднее квадратическое отклонение).

Является мерой изменчивости (вариации) признака. Оно показывает, на какую величину в среднем отклоняются данные от среднего значения признака. Математическое описание стандартного отклонения представлено в формуле (2).

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}} \quad (2)$$

Для устранения аномальных наблюдений используется правило «трех сигм»: наблюдение признается аномальным и отбрасывается, если его отклонение от выборочной средней более чем в 3 раза превышает среднеквадратическое отклонение выборки, т. е. удовлетворяют неравенству (3):

$$|\bar{x} - x_i| > 3\sigma \quad (3)$$

На рисунке 2.12 можно увидеть, как применяется правило трех сигм для поиска выбросов.

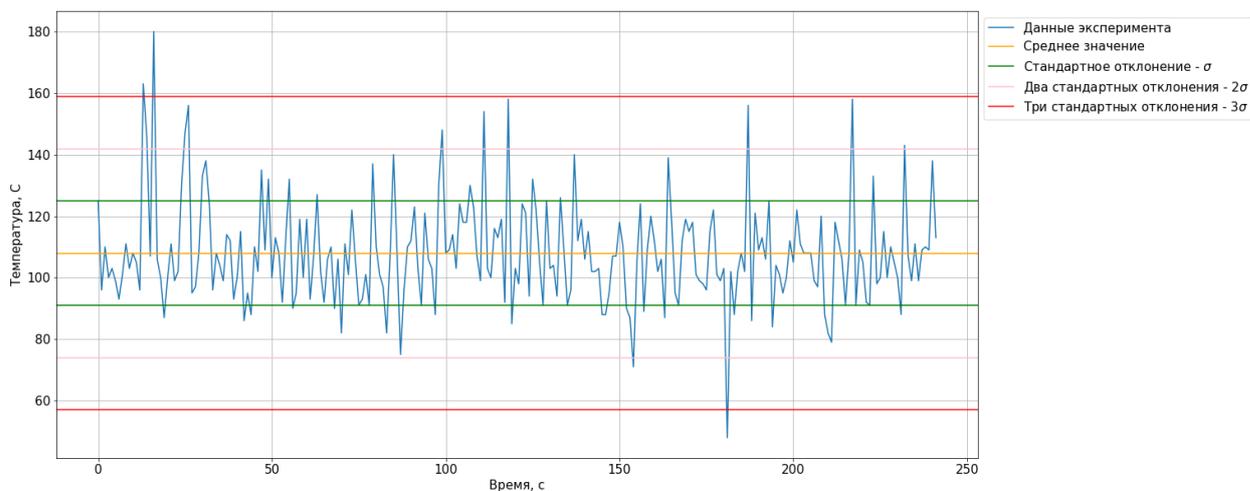


Рисунок 2.12 - График изменения уровня шума со временем, визуализация метода трех сигм

На графике видны три значения уровня шума, значение отклонений которых более, чем в три раза превышает среднеквадратичное отклонение. Такие значения можно считать выбросами и не рассматривать их влияние при анализе распределения значений и поиске тенденций изменений в данных.

- Стандартная ошибка среднего.

Это величина, на которую отличается среднее значение выборки от среднего значения генеральной совокупности при условии, что распределение близко к нормальному. Рассматривая эту величину вводят доверительный интервал. Это интервал, который покрывает неизвестный параметр с заданной надёжностью. Например, 95%-й доверительный интервал для среднего - интервал, в который с вероятностью 0,95 попадает среднее значение признака генеральной совокупности [5].

Рисунок 2.13 иллюстрирует различие экспериментальных данных и данных, которые находятся в пределах доверительного интервала 95%.

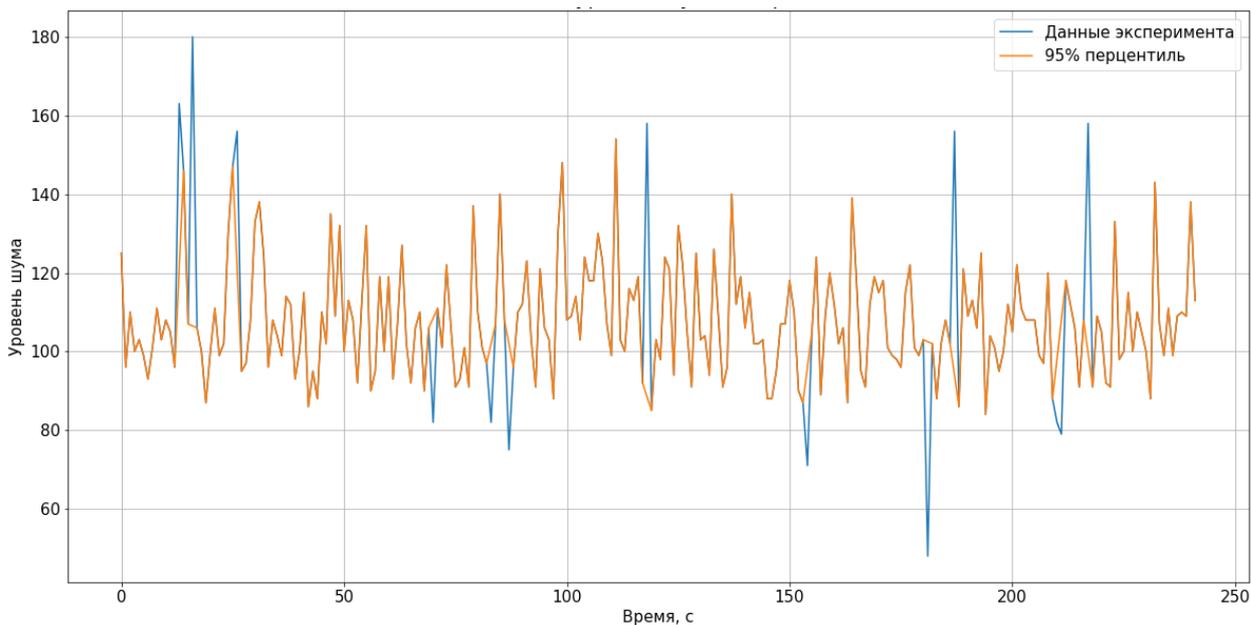


Рисунок 2.13 - Распределение данных эксперимента и данных, входящих в доверительный интервал 95%

Данный рисунок показывает, что использование доверительного интервала при проведении исследований помогает избавиться от выбросов и снизить возможность ошибки при оценке состояния среды.

Анализ распределения данных во многом связан с действием неоднородных по силе и направлению факторов на изучаемый процесс. Рассматривая динамические ряды, пытаются разделить факторы на случайные, постоянно действующие и факторы, формирующие новую тенденцию.

Простейший способ обработки данных – метод укрупнения интервалов. Основная идея – перейти от исходных интервалов и периодов времени к более продолжительным.

Рисунки 2.14 и 2.15 иллюстрируют применение метода укрупнения интервалов.

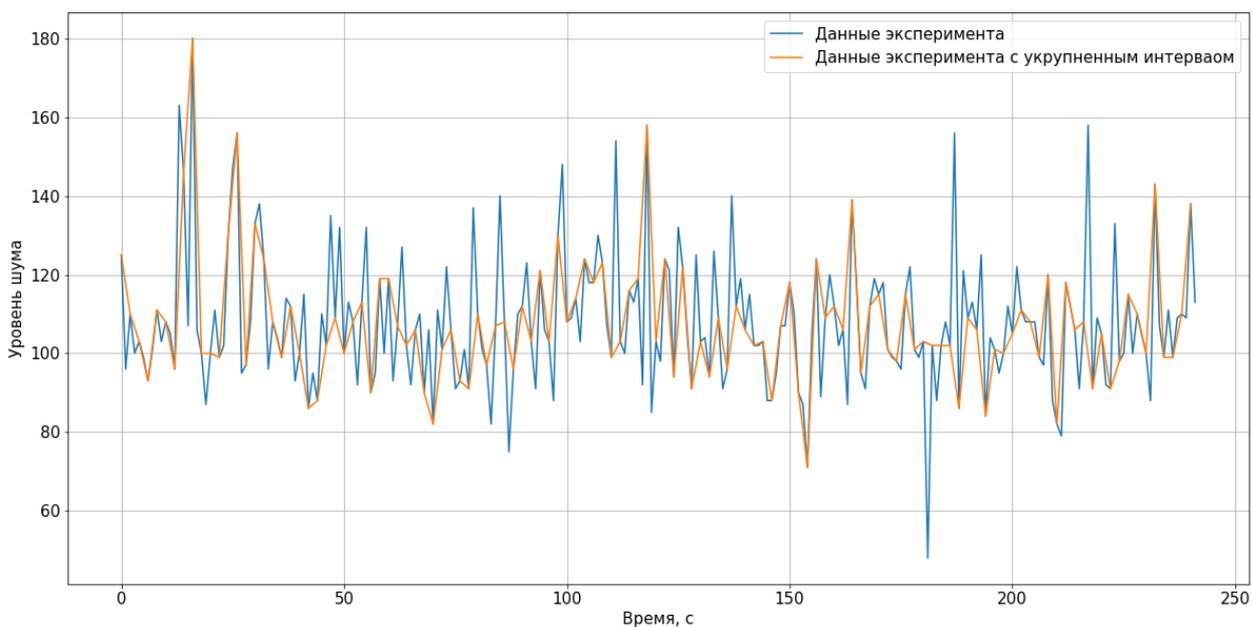


Рисунок 2.14 - Изменение уровня шума. Визуализация данных, метод укрупнения данных с шагом 3



Рисунок 2.15 - Изменение уровня шума. Визуализация данных, метод укрупнения данных с шагом 5

Из визуализации метода укрупнения интервалов можно увидеть, что в сглаженном ряду закономерности изменения данных более наглядные.

Другой способ – метод скользящей средней. Суть метода заключается в замене средними значениями, вычисленными по определенным правилам. Например, вычисление среднего значения на интервале.

Метод скользящей средней не позволяет получить численные оценки для выражения основной тенденции в ряду, а дает лишь наглядное графическое представление [10].

На рисунке 2.16 представлена визуализация применения метода скользящей средней по средним значениям.

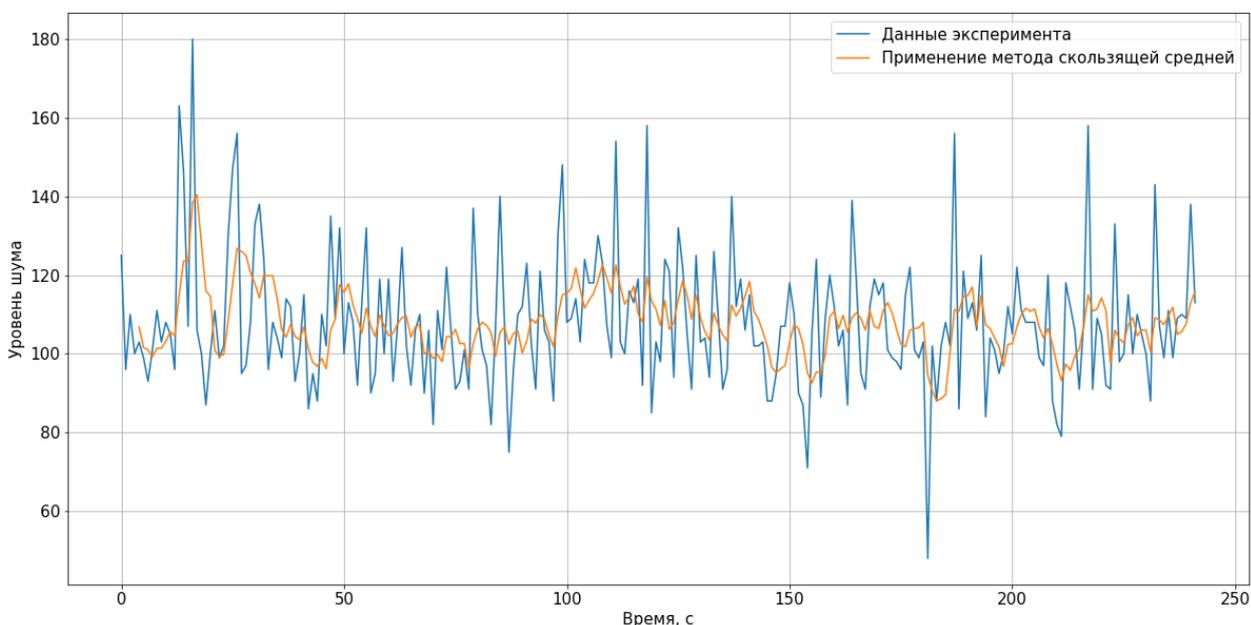


Рисунок 2.16 - Изменение уровня шума. Визуализация данных, метод скользящей средней

Применение метода скользящей средней помогает выявить тенденции изменения данных в динамических рядах.

Исследование различных типов визуализации данных также может помочь выделить области аномальных значений. На рисунках 2.17 и 2.18 приведено распределение данных, полученных с датчика концентрации опасных газов с течением времени и в пространстве. Из этих рисунков видно, что в определенный момент времени в помещение был впрыснут газ и на графиках это отражено резким скачком концентрации газа. По мере удаления от источника концентрация снижается.

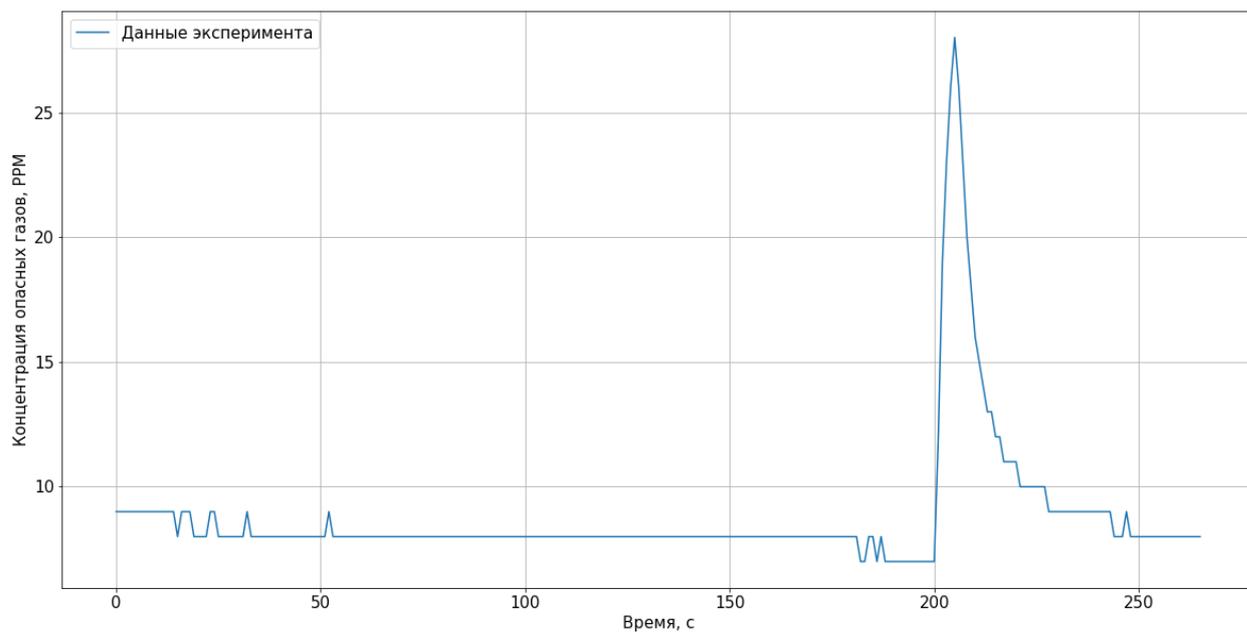


Рисунок 2.17 - Изменение концентрации опасных газов с течением времени

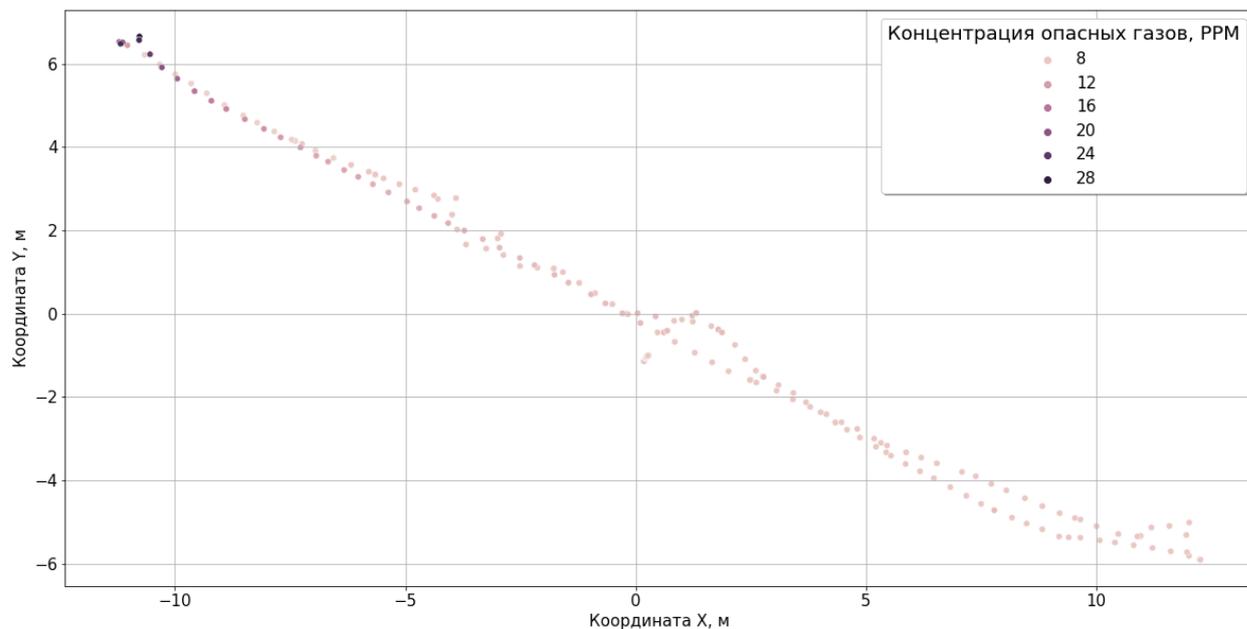


Рисунок 2.18 - Изменение концентрации опасных газов в пространстве

Выявление такого вида аномалий поможет выявлять утечки опасных газов в производственных помещениях, чтобы быстро решить проблему и предотвратить возникновение внештатной ситуации.

ЗАКЛЮЧЕНИЕ

В данной выпускной квалификационной работе был проведен прочностной анализ модели автоматизированного робототехнического комплекса в программном пакете Ansys, а также реализован алгоритм обработки данных экологического мониторинга на языке программирования Python.

Для создания оптимальной модели был проведен расчет прочности под действием статических и динамических нагрузок. Определена минимальная толщина платформы 2.1 мм. Коэффициент запаса прочности данной детали находится в зоне упругих деформаций, а минимальная собственная частота выше максимальной частоты действующей периодической нагрузки вращения колес.

После проведения прочностных расчетов была поставлена серия экспериментов для того, чтобы проверить корректность работы устройства и получить необходимые данные для дальнейшей обработки.

Для анализа полученных данных был составлен алгоритм, в котором применялись статистические оценки параметров и различные способы визуализации. Были выделены участки помещения, где наблюдались отличные от нормы показатели. Данные области совпадают с местоположением объектов, которые производили влияние на экологические параметры в ходе проведения экспериментов.

Полученные алгоритм имеет следующую структуру:

- 1) На вход подаются данные с датчиков в виде CSV-файла.
- 2) Полученные данные визуализируются.

Для того, чтобы увидеть распределение данных проводится визуализация. Она помогает увидеть к какому набору параметров лучше применить те или иные математические методы.

3) Обработка данных каждого датчика.

Для обработки данных применяются различные методы теории вероятности и математической статистики. Определяется доверительный интервал значений, который позволяет делать оценку распределения параметров исключая из рассмотрения выбросы.

4) Визуализация полученных данных.

Для того, чтобы посмотреть полученные результаты и наглядно показать расположение аномальных зон проводится визуализация полученных результатов.

5) Вывод результатов анализа данных.

По результатам применения математических методов и визуализации полученных результатов проводится оценка микроклимата помещения. Эта оценка может использоваться для составления более точной карты помещения, выявления нештатных ситуаций при превышении допустимых параметров и быстрого реагирования на данные проблемы.

Полученные результаты были презентованы на конференции бакалавров Высшей школы теоретической механики и математической физики.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Грас Дж. Data Science. Наука о данных с нуля: пер. с англ. Андрея Логунова. - СПб.: БХВ-Петербург, 2017.– 336 с.
2. Калинин А.Г. Обработка данных методами математической статистики. – М.: Чита : ЗИП СибУПК, 2015. – 106 с.
3. Миркин Б. Г. Введение в анализ данных. — М.: Юрайт. 2020. 175 с.
4. Мойзес Б. Б., Плотникова И. В., Редько Л. А. Статистические методы контроля качества и обработка экспериментальных данных: учебное пособие. – М.: Юрайт, 2019. – 118 с.
5. Никифорова Ю.Ю. Статистические методы в экологии и природопользовании: учеб. пособие / под. общ.ред. И. С. Белюченко. – Краснодар: КубГАУ, 2019. – 88 с.
6. Поручиков М.А. Анализ данных: учеб. пособие. – Самара: Изд-во Самарского университета, 2016. – 88 с.
7. Шимановский А.О., Путято А.В. Применение метода конечных элементов в решении задач прикладной механики: учеб.-метод. пособие для студентов технических специальностей. – Гомель: БелГУТ, 2008. – 61 с.
8. Chandola V., Banerjee A., Kumar V. “Anomaly detection: A survey” // ACM Computing Surveys. – vol. 41(3). – 2009. – pp. 1–58.
9. Бурдаков С.Ф., Харадин Н.А. Планирование и оптимизация программных движений роботов методом конфигурационного пространства // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Сер.: Информатика. Телекоммуникации. Управление. - 2012. - № 3 (150). - С. 17-23
10. Заварзин Д.В. К вопросу поиска аномалий во временных рядах: Инновации в науке: сб. ст. по матер. XXIX междунар. науч.- практ. конф. № 1(26). – Новосибирск: СибАК, 2014

11. Зубков Е. В., Белов В. М., «Методы интеллектуального анализа данных и обнаружение вторжений» // Вестник СибГУТИ. — 2016. — № 1.
12. Суханов А. В. Интеллектуальные методы обнаружения и прогнозирования аномальных событий в темпоральных данных.: Диссертация на соискание ученой степени кандидата технических наук. // Ростов-на-Дону: РГУПС, 2015. – 159 с.
13. Шкодырев В.П. [и др.] Обзор методов обнаружения аномалий в потоках данных. // Proc. of the Second Conference on Software Engineering and Information Management, Санкт-Петербург. — 2017. — Vol. 1864. — 7 с.
14. Deepthi Cheboli. Anomaly Detection of Time Series // Faculty Of The Graduate School Of The University Of Minnesota. // Minneapolis: Engineering University of Minnesota, 2010. – 75 p.
15. Лайрд Джордж Конечно-элементный анализ для всех. Часть 1 [Электронный ресурс]. – URL: <https://sapr.ru/article/21943> (дата обращения: 12.02.2023)
16. Поиск аномалий. Анализ малых данных [Электронный ресурс]. – URL: <https://alexanderdyakonov.wordpress.com/2017/04/19/поиск-аномалий-anomaly-detection/> (дата обращения 05.12.2022)
17. Расчет на прочность методом конечных элементов [Электронный ресурс]. – URL: <https://tech-structure.ru/articles/raschet-na-prochnost-metodom-konechnykh-elementov> (дата обращения 09.02.2023)
18. Ansys Workbench. Статистический конструкционный анализ [Электронный ресурс]. – URL: https://mircompozitov.ucoz.ru/05_ansys_workbench-staticheskij_konstrukcionnyj_an.pdf (дата обращения 07.02.2023)
19. Explore_lite. Package Summary [Электронный ресурс]. – URL: http://wiki.ros.org/explore_lite (дата обращения: 25.04.2023)
20. What Is Anomaly Detection in Machine Learning? Yulia Gavrilova [Электронный ресурс]. – URL: <https://serokell.io/blog/anomaly-detection-in-machine-learning> (дата обращения: 25.11.2022)

21. ГОСТ 30494–2011 Межгосударственный стандарт. Здания жилые и общественные. Параметры микроклимата в помещениях. Введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 12 июля 2012 г. N 191-ст.

Приложение А. Визуализация результатов экспериментов

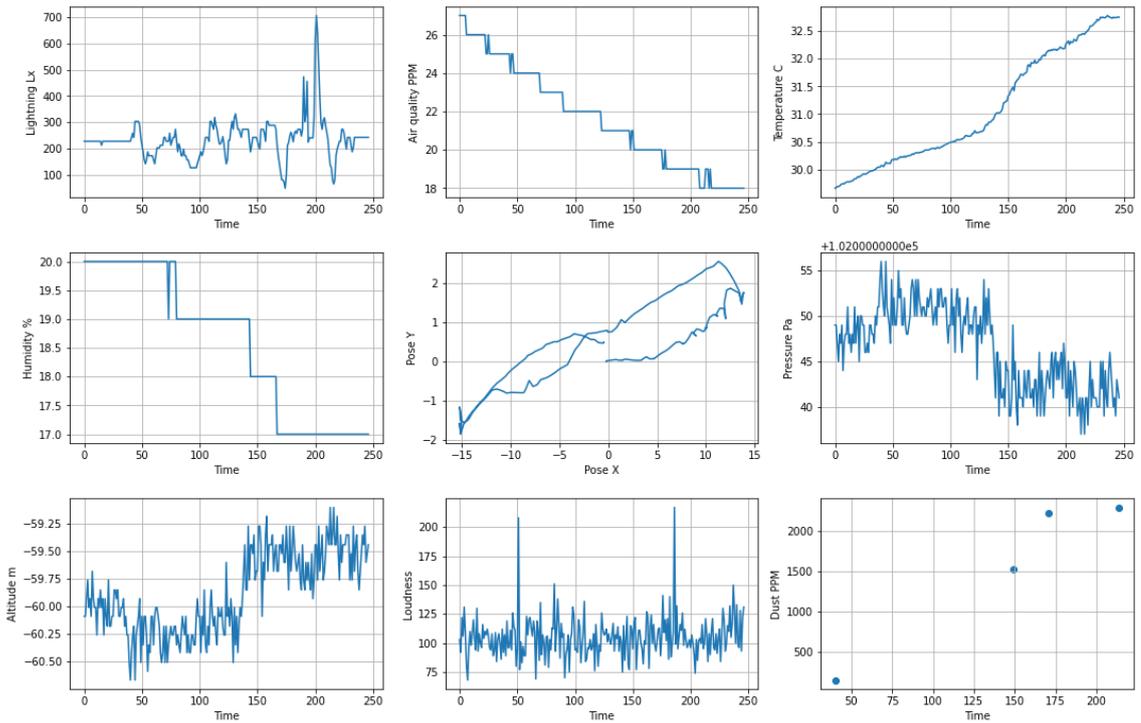


Рисунок А1 - Визуализация данных эксперимента без создания возмущений

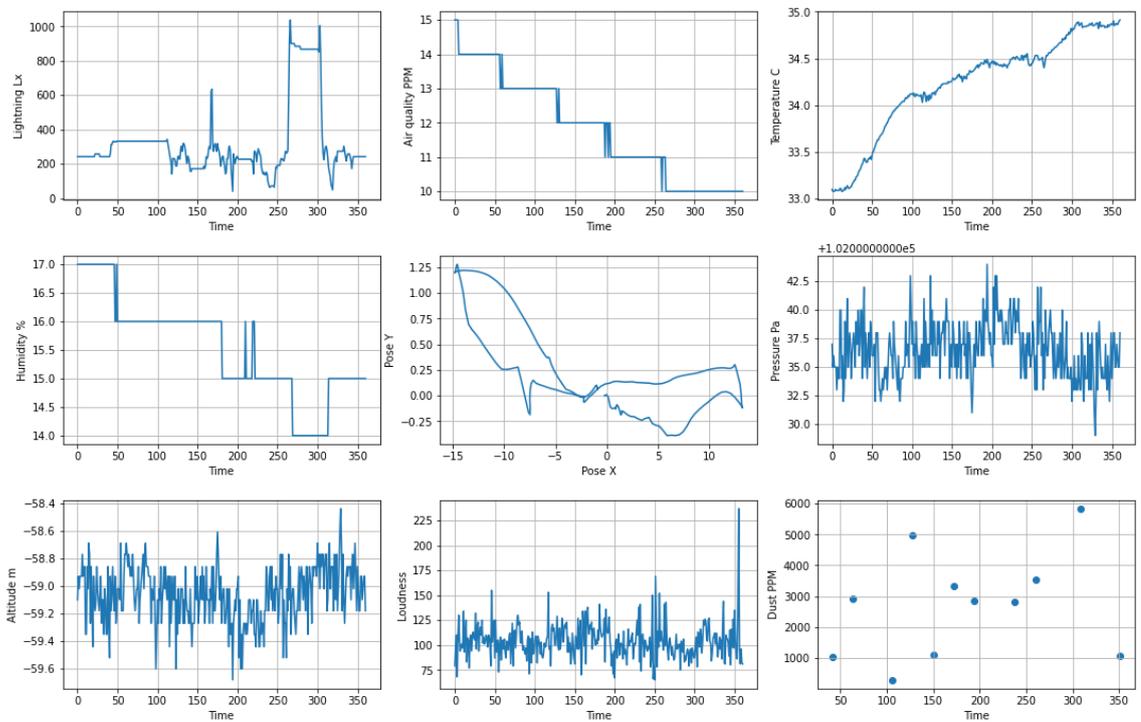


Рисунок А2 - Визуализация данных эксперимента с влиянием на изменение уровня шума и освещенности

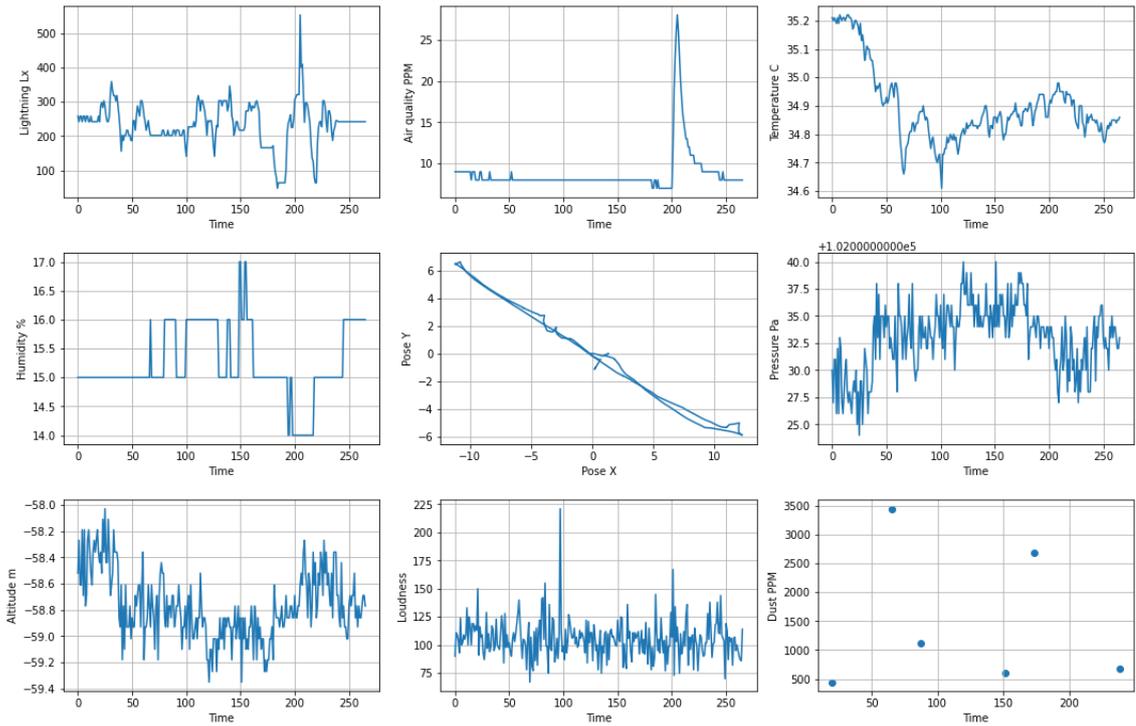


Рисунок А3 - Визуализация данных эксперимента с влиянием на изменение уровня влажности и концентрации газов

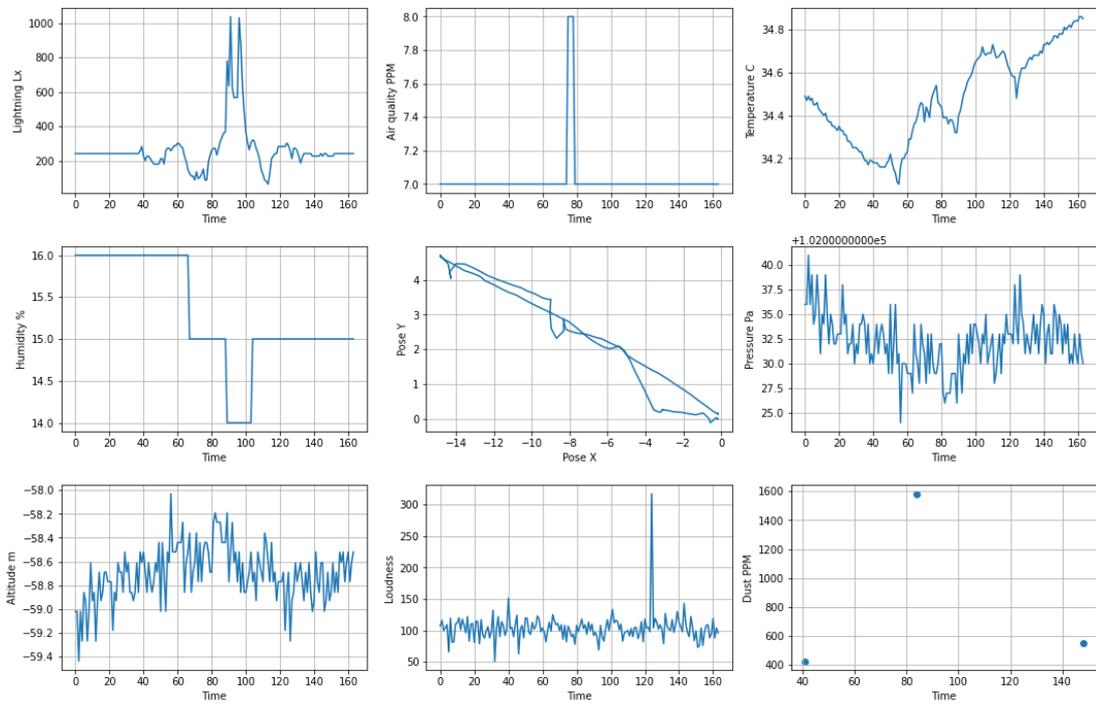


Рисунок А4 - Визуализация данных эксперимента с влияние на изменение температуры