

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Школа Информационных технологий и робототехники
Направление подготовки 09.03.01 «Информатика и вычислительная техника»
Отделение школы (НОЦ) Отделение информационных технологий

БАКАЛАВРСКАЯ РАБОТА

Тема работы
Разработка программного обеспечения для исследования корреляционных зависимостей метеорологических параметров

УДК 004.415:551.508:519.23

Студент

Группа	ФИО	Подпись	Дата
8В4Б	Маркова Ксения Дмитриевна		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Ботыгин И.А.	к.т.н		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Старикова Е.В.	к.ф.н		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент	Авдеева И.И.			

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.03.01 Информатика и ВТ	Погребной А.В.	к.т.н		

Томск – 2018 г.

**ЗАПЛАНИРОВАННЫЕ РЕЗУЛЬТАТЫ ПО ОСНОВНОЙ
ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ ПОДГОТОВКИ БАКАЛАВРОВ
09.03.01 «ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА», ИК
ТПУ, ПРОФИЛЬ «ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ, КОМПЛЕКСЫ,
СИСТЕМЫ И СЕТИ»**

Код результата тов	Результат обучения (выпускник должен быть готов)
<i>Профессиональные компетенции</i>	
P1	Применять базовые и специальные естественнонаучные и математические знания в области информатики и вычислительной техники, достаточные для комплексной инженерной деятельности.
P2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных задач.
P3	Ставить и решать задачи комплексного анализа, связанные с созданием аппаратно-программных средств информационных и автоматизированных систем, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
P4	Разрабатывать программные и аппаратные средства (системы, устройства, блоки, программы, базы данных и т. п.) в соответствии с техническим заданием и с использованием средств автоматизации проектирования.
P5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области создания аппаратных и программных средств информационных и автоматизированных систем.
P6	Внедрять, эксплуатировать и обслуживать современные программно-аппаратные комплексы, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.
<i>Универсальные компетенции</i>	
P7	Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.
P8	Владеть иностранным языком на уровне, позволяющем работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.
P9	Эффективно работать индивидуально и в качестве члена группы, состоящей из специалистов различных направлений и квалификаций, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.
P10	Демонстрировать знания правовых, социальных, экономических и культурных аспектов комплексной инженерной деятельности.
P11	Демонстрировать способность к самостоятельному обучению в течение всей жизни и непрерывному самосовершенствованию в инженерной профессии.

Министерство образования и науки Российской Федерации
 федеральное государственное автономное образовательное учреждение
 высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
 ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Школа Информационных технологий и робототехники
 Направление подготовки 09.03.01 «Информатика и вычислительная техника»
 Отделение школы (НОЦ) Отделение информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП
 _____ Погребной А.В.
 (Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Бакалаврской работы

(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
8В4Б	Маркова Ксения Дмитриевна

Тема работы:

Разработка программного обеспечения для исследования корреляционных зависимостей метеорологических параметров	
Утверждена приказом директора (дата, номер)	от 20.03.2018 № 1970/с

Срок сдачи студентом выполненной работы:	31.05.18
------------------------------------------	----------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Объектно-ориентированный язык программирования Java; свободная интегрированная среда разработки приложений NetBeans; база данных Всероссийского научно исследовательского института гидрометеорологической информации – Мирового центра данных (ВНИИГМИ–МЦД).
Перечень подлежащих исследованию, проектированию и разработке вопросов	1 Обзор методов корреляционного анализа; 2 Проектирование архитектуры программного обеспечения; 3 Реализация программного обеспечения; 4 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение; 5 Социальная ответственность.
Перечень графического материала	– Функциональная структура программного обеспечения

Консультанты по разделам выпускной квалификационной работы <i>(с указанием разделов)</i>	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Старикова Екатерина Васильевна
Социальная ответственность	Авдеева Ирина Ивановна

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	1.02.18
-------------------------------------------------------------------------------------------------	---------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Ботыгин Игорь Александрович	кандидат технических наук		1.02.18

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8В4Б	Маркова Ксения Дмитриевна		1.02.18

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Школа Информационных технологий и робототехники
Направление подготовки 09.03.01 «Информатика и вычислительная техника»
Уровень образования бакалавриат
Отделение школы (НОЦ) Отделение информационных технологий
Период выполнения (осенний / весенний семестр 2017/2018 учебного года)

Форма представления работы:

Бакалаврская работа

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

**КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы**

Срок сдачи студентом выполненной работы:	31.05.2018
------------------------------------------	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
1.02.18	Аналитический обзор	15
2.03.18	Проектирование программного обеспечения	20
18.04.18	Реализация программного обеспечения	45
16.05.18	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
25.05.18	Социальная ответственность	10

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Ботыгин И.А.	к.т.н.		

СОГЛАСОВАНО:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.03.01 Информатика и ВТ	Погребной А. В.	к.т.н.		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

Группа	ФИО
8В4Б	Маркова Ксения Дмитриевна

Школа	ИШИТР	Отделение	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	09.03.01 Информатика и вычислительная техника

Исходные данные к разделу «Социальная ответственность»	
<p>1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения</p>	<p>Объект исследования – программное обеспечение, реализующее корреляционную обработку метеорологических данных. Рабочая зона – аудитория с естественным и искусственным освещением, оборудованная системой отопления и кондиционирования воздуха. Область применения – любая организация, работающая с метеорологическими данными.</p>
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<p>1. Производственная безопасность 1.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности 1.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности</p>	<p>1.1. Анализ выявленных вредных факторов при разработке и эксплуатации включает: Микроклимат, слабая освещенность, повышенный уровень шума, умственное перенапряжение, монотонность труда. Требования к помещению описаны в СанПиН 2.2.4.3359-16. 1.2. Анализ выявленных опасных факторов при разработке и эксплуатации включает: Удар электрическим током, короткое замыкание, статическое электричество</p>
<p>2. Экологическая безопасность:</p>	<p>Утилизация использованной техники (компьютера и других составляющих аппаратно-программного комплекса). Утилизация канцелярских принадлежностей и бумаги, использованных лампочек.</p>
<p>3. Безопасность в чрезвычайных ситуациях:</p>	<p>Возможная чрезвычайная ситуация для данного помещения – пожар, вследствие короткого замыкания. Необходимо установить общие правила поведения при пожаре, ознакомить с планом эвакуации, иметь в наличии</p>

	исправного огнетушителя, назначить ответственного по пожарной безопасности в данном помещении.
4. Правовые и организационные вопросы обеспечения безопасности:	Основные проводимые правовые и организационные мероприятия по обеспечению безопасности трудящихся в аудиториях. Анализ правильного расположения и организации рабочего места, а так же режима работы. Регулирования организации рабочего места при выполнении работы сидя проводятся согласно ГОСТ 12.2.032-78. Трудовые отношения регулируются согласно ТК РФ ФЗ–197 от 30.12.2001.

Дата выдачи задания для раздела по линейному графику	01.03.18
-------------------------------------------------------------	----------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент	Авдеева Ирина Ивановна			01.03.18

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8В4Б	Маркова Ксения Дмитриевна		01.03.18

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ,
РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
8В4Б	Маркова Ксения Дмитриевна

Школа	ИШИТР: Инженерная школа информационных технологий и робототехники	Отделение	Информационных технологий
Уровень образования	Бакалавриат	Направление/специальность	09.03.01 Информатика и вычислительная техника

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:	
Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Работа с информацией, представленной в российских и иностранных научных публикациях, аналитических материалах, статистических бюллетенях и изданиях, нормативно-правовых документах.
Нормы и нормативы расходования ресурсов.	
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения.	Оценка потенциальных потребителей исследования, SWOT-анализ.
Определение возможных альтернатив проведения научных исследований.	Определение возможных альтернатив с помощью морфологического подхода.
Планирование научно-исследовательских работ.	Планирование структуры работ, определение трудоемкости работы и построение календарного графика.

Дата выдачи задания для раздела по линейному графику	01.03.2018
-------------------------------------------------------------	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент отделения социально – гуманитарных наук	Старикова Екатерина Васильевна	к.ф.н.		01.03.2018

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8В4Б	Маркова Ксения Дмитриевна		01.03.2018

РЕФЕРАТ

Выпускная квалификационная работа 75с., 15таб., 18 рис., 37 источников.

Ключевые слова: корреляционный анализ, метеорологические данные, программирование, статистический анализ, базы данных.

Объект исследования: временные метеорологические поля.

Цель работы – разработка программного комплекса для статической обработки и визуализации произвольных рядов метеорологических данных.

Область применения: аналитика, прогнозирование, расчёты.

В результате исследования был разработан программный комплекс для статистической обработки и визуализации произвольных рядов метеорологических данных. По итогу работы программным комплексом были рассчитаны коэффициенты корреляции и проведен их сравнительный анализ. Так же в ходе исследования была рассчитана описательная статистика, необходимая для расчета коэффициентов корреляции, и были реализованы построение графиков корреляции и экспорт рассчитанных данных.

Оглавление

РЕФЕРАТ	9
ВВЕДЕНИЕ.....	13
1. АНАЛИТИЧЕСКИЙ ОБЗОР СОВРЕМЕННЫХ СРЕДСТВ СТАТИСТИЧЕСКОГО АНАЛИЗА ДЛЯ ПРОГРАММНОГО КОМПЛЕКСА СТАТИСТИЧЕСКОЙ ОБРАБОТКИ И ВИЗУАЛИЗАЦИИ ДАННЫХ	14
1.1. Статистический анализ.....	14
1.1.1. Высокоуровневый язык программирования общего назначения Python	14
1.1.2. Язык программирования для статистической обработки данных R	16
1.1.3. Объектно-ориентированный язык программирования Java...	19
1.1.4. Альтернативные статистические пакеты.....	20
1.2. Функции описательной статистики.....	21
1.3. Корреляция данных.....	23
2. ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ПРОГРАММНОГО КОМПЛЕКСА СТАТИСТИЧЕСКОЙ ОБРАБОТКИ И ВИЗУАЛИЗАЦИИ ДАННЫХ	28
2.1. Функциональная структура программного обеспечения.....	28
2.2. Проектирование базы данных.....	30
3. ТЕХНОЛОГИЧЕСКОЕ ИСПОЛЬЗОВАНИЕ ПРОГРАММНОГО КОМПЛЕКСА СТАТИСТИЧЕСКОЙ ОБРАБОТКИ И ВИЗУАЛИЗАЦИИ МЕТЕОРОЛОГИЧЕСКИХ ДАННЫХ	32
4. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ.....	42
4.1 Введение.....	42

4.2	Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения.....	42
4.2.1.	Потенциальные потребители результатов исследования	42
4.2.2.	Анализ конкурентных технических решений.....	43
4.2.3.	Технология QuaD	44
4.2.4.	SWOT-анализ.....	45
4.3.	Определение возможных альтернатив проведения научных исследований	47
4.4.	Планирование научно-исследовательских работ	48
4.4.1.	Структура работ в рамках научного исследования	48
4.4.2.	Определение трудоемкости выполнения работ	49
4.4.3.	Разработка графика проведения научного исследования.....	49
4.4.4.	Бюджет научно-технического исследования (НТИ)	51
4.5.	Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	54
5.	СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ	57
5.1.	Производственная безопасность	57
5.1.1.	Требования к микроклимату.....	57
5.1.2.	Расчет уровня шума	59
5.1.3.	Освещенность.....	60
5.1.4.	Монотонность труда и умственное перенапряжение.....	62
5.1.5.	Техника электробезопасности	62
5.2.	Экологическая безопасность.....	65
5.3.	Безопасность в чрезвычайных ситуациях.....	66

5.4. Правовые и организационные вопросы обеспечения безопасности	67
ЗАКЛЮЧЕНИЕ	70
СПИСОК ПУБЛИКАЦИЯ СТУДЕНТА	71
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	72

ВВЕДЕНИЕ

С развитием компьютерных технологий в мире появилась необходимость обработки большого потока данных. В математике анализом данных, а так же их интерпретацией занимается математическая статистика. Для решения множества различных задач требуется статистическая обработка, которая имеет множество подход к обработке данных. Одним из таких подходов является корреляция данных. Корреляционный анализ проводится для выявления зависимости одного параметра от другого.

Статистический анализ в целом, и корреляционный анализ в частности проводят многие математические пакеты. Большинство пакетов имеют закрытый код, следовательно, у пользователя нет возможности адаптировать программу под свои нужды. Пакеты с открытым кодом предоставляют такую возможность.

Данная работа посвящена созданию программного комплекса для исследования корреляционной обработке метеорологических данных.

1. АНАЛИТИЧЕСКИЙ ОБЗОР СОВРЕМЕННЫХ СРЕДСТВ СТАТИСТИЧЕСКОГО АНАЛИЗА ДЛЯ ПРОГРАММНОГО КОМПЛЕКСА СТАТИСТИЧЕСКОЙ ОБРАБОТКИ И ВИЗУАЛИЗАЦИИ ДАННЫХ

1.1. Статистический анализ

В современном обществе, при большом потоке различных данных, существует острая необходимость в их анализе и представлении, в удобном для восприятия любым пользователем. Решением данной проблемы занимается статистика, в частности статистический анализ данных.

Статистический анализ занимается сбором, исследованием и представлением больших объемов данных, для обнаружения основных моделей и тенденций. Статистический анализ, от традиционного анализа дисперсии и линейной регрессии до статистической визуализации, имеет большое значение для принятия решений на основе имеющихся данных во всех областях.

Существует множество фреймворков и систем компьютерной алгебры, в которых встроены пакеты для статистической обработки данных. Рассмотрим некоторые из них.

1.1.1. Высокоуровневый язык программирования общего назначения Python

Python является высокоуровневым языком программирования общего назначения [1]. Существует множество встроенных и сторонних библиотек для математического анализа. Рассмотрим несколько наиболее распространенных библиотек.

Statistics является встроенной библиотекой Python [2]. Предоставляет функции для вычисления математической статистики числовых данных. Так же предоставляет функции, вычисляющие среднее или типичное значение из выборки, меру того, насколько образец имеет тенденцию отклоняться от типичных или средних значений.

NumPy – самый фундаментальный пакет для научных вычислений. Он поддерживает многомерные массивы и матрицы, а так же высокоуровневых математических функций, предназначенных для работы над N-массивами и матрицами [3]. Благодаря механизму векторизации, NumPy повышает производительность и, соответственно, ускоряет выполнение операций. Одним из главных достоинств данной библиотеки считается наличие своего типа, для работы с большими многомерными массивами данных – ndarray. Работать с ndarray очень удобно, так как объекты данного класса можно использовать в выражениях как операнды, при этом осуществляются поэлементные операции со всеми элементами массива

SciPy является бесплатной библиотекой Python с открытым исходным кодом, используемой для научных и технических вычислений [4]. Данная библиотека основывается на рассмотренной ранее NumPy. Основной структурой данных, так же как и в NumPy, является многомерный массив.

SciKits является набором отдельных пакетов объединенных в одно пространство имен. Скачать всю библиотеку разом нельзя [5]. В SciPy данные пакеты не попали по трем причинам: несовместимость лицензий, нестабильный статус библиотеки и узкоспециальная направленность. Лишь часть пакетов содержат хорошую документацию и готова к использованию.

Библиотека для статистических исследований StatsModels предоставляет классы и функции для оценки различных статистических моделей, а также для проведения статистических испытаний и исследования статистических данных [6]. Предоставляются функции базовой статистики, остаточной диагностики, отклонений и мер воздействия, пригодности тестов для одномерных распределений, непараметрических тестов, множественных тестов и процедур сравнения, а так же расчета мощности и размера выборки. Так же представлены функции, вычисляющие ковариационные матрицы и стандартные ошибки для оценок параметров, устойчивых к автокорреляции ошибок.

Modular toolkit for Data Processing (MDP) является пакетом для анализа данных [7]. Содержит метод опорных векторов, различные классификаторы,

факторный анализ, метод k-средних, и множество других. Особенность библиотеки заключается в том, что все алгоритмы инкапсулируются в классы (узлы в терминологии MDP) с общим интерфейсом, что позволяет объединять их «пропуская» данные через связку узлов.

Orange – еще один пакет по анализу данных. Содержит алгоритмы классификации, кластеризации и регрессионного анализа из ранее не встречавшихся следует отметить алгоритм построения деревьев решений C45 [8]. Отличительной особенностью является наличие удобного GUI позволяющего пользоваться всеми алгоритмами даже не зная, Питон.

SymPy – пакет, реализующий работу с символьными выражениями, написан без использования сторонних библиотек [9]. Проект молодой и активно развивается, и уже сегодня позволяет решать многие распространенные задачи, включая: дифференцирование, интегрирование, пределы, дифференциальные уравнения и др.

Естественно это не все библиотеки для математической обработки данных на Python, однако даже рассмотренные библиотеки охватывают весомый спектр задач.

1.1.2. Язык программирования для статистической обработки данных R

R представляет интегрированный набор программных средств для обработки и расчета данных, а так же их графического отображения [10]. Он быстро развился и был расширен сторонними пакетами, так как предоставляет исходный код в открытом доступе, что позволяет пользователю самому создавать и публиковать пакеты. Поэтому большинство программ, написанных в R, по сути, являются эфемерными, написанными для одного фрагмента анализа данных.

Распространяются R и созданные пользователями пакеты через Comprehensive R Archive Network (CRAN) [11]. Существует около 25 стандартных пакетов, поставляемых с R, а многие другие доступны через CRAN. В настоящее время в репозитории пакетов CRAN реализовано около

12674 доступных пакетов. Одной из сильных сторон языка является возможность создавать хорошо продуманные графики, включающие математические символы и формулы. R обеспечивает взаимосвязанный набор средств, которые упрощают сбор статистических моделей. Базовый вывод минимален, и нужно уточнять детали, вызывая функции экстрактора. Так же в R рассматриваются линейные и нелинейные модели, анализ дисперсии, сравнения различных моделей, а так же несколько нестандартных моделей регрессии. Как было сказано ранее, для статистической обработки R предоставляет множество пакетов. Рассмотрим некоторые пакеты для работы с корреляцией.

Пакет `Stats4` поддерживает функции с использованием классов `S4` [12]. Пакет `Stats` является обширным пакетом, содержащей функции для статистических расчетов и генерации случайных чисел [13]. Данный пакет содержит 11 функций проводящих различные виды корреляционных анализов. Рассмотрим несколько наиболее популярных.

Функция `cancor()` вычисляет канонические корреляции между двумя матрицами данных [14]. Канонический корреляционный анализ ищет линейные комбинации y переменных, которые хорошо объясняются линейными комбинациями x переменных.

Функции `var()`, `cov()` и `cor()` вычисляют дисперсию x и корреляцию (или ковариацию) x и y , если они являются векторами [15]. Если x и y являются матрицами, тогда вычисляются корреляции (или ковариации) между столбцами столбцов x и столбцов y .

Функция `cov2cor()` эффективно масштабирует матрицу ковариации в соответствующую корреляционную матрицу.

Пакет `Hmisc` [16]. Пакет `Hmisc` содержит множество функций, полезных для анализа данных, вычисления описательной статистики, высокоуровневой графики, служебных операций, функций для вычисления размера и мощности выборки, импорта и аннотирования наборов данных, вменения отсутствующих значений, расширенной таблицы, переменной кластеризации, манипуляции с

символьной строкой, преобразования объектов R в LaTeX и html-код, и переменные перекодировки.

Функция `gcorr()` в данном пакете создает корреляции (ковариации) и уровни значимости для корреляций Пирсона и Спирмана [17]. Однако вход должен быть матрицей, и использовать попарное удаление.

Так же для проведения корреляционного анализа необходим расчет описательной статистики. Рассмотрим некоторые функции расчета описательной статистики на языке R.

Встроенные функции получения описательной статистики. R предоставляет широкий спектр функций для получения сводной статистики. Одним из методов получения описательной статистики является использование функции `sapply()` с указанной статистикой [18]. Возможные функции, используемые в `sapply`, включают в себя среднее значение, максимальное и минимальное значение, дисперсию, медиану, ранг и квантиль. Существует также множество R-функций, предназначенных для одновременного предоставления описательной статистики.

Так же несколько пакетов предоставляют полную описательную статистику: `Hmisc`, `pastecs` и `psych`. Пакет `pastecs` отвечает за регуляризацию, декомпозицию и анализ пространственно-временных рядов [19]. Пакет `psych` был создан как инструментальный общего назначения для экспериментальной психологии [20]. Функции в основном предназначены для многомерного анализа и построения шкалы с использованием факторного анализа, анализа основных компонентов, кластерного анализа и надежности, хотя другие пакеты предоставляют базовую описательную статистику. Функции для анализа данных на нескольких уровнях включают в себя и групповую статистику, включая корреляцию и факторный анализ. Включены функции для моделирования и тестирования отдельных элементов, а так же графическое отображение данных.

1.1.3. Объектно-ориентированный язык программирования Java

Java является типизированным объектно-ориентированным языком программирования [21]. Приложения Java транслируются в специальный байт-код и могут работать на любой компьютерной архитектуре, с помощью виртуальной Java-машины. Рассмотрим некоторые библиотеки, позволяющие производить статистические расчеты на Java.

Colt – это набор библиотек с открытым исходным кодом для высокопроизводительных научно-технических вычислений, написанный на Java [22]. Был разработан с акцентом на физику высоких технологий, но применим ко многим другим проблемам, так же, и к статистическому анализу данных. Имеет 3 основных пакета: Вероятность, Гамма и Описательный. Для анализа данных рассматривается лишь один пакет – Описательный, отвечающий за основную описательную статистику. В данной пакеты находятся методы, выводящие всю стандартную описательную статистику, ковариацию, корреляцию и многое другое.

Jama – библиотека функций линейной алгебры [23]. JAMA не содержит ничего, кроме заголовочных файлов с шаблонами, поэтому библиотека не требует компиляции. Поскольку все классы используют шаблоны, одинаково легко использовать матрицы и вектора с элементами типа float, double или описанных пользователем типов. JAMA не является полной средой линейной алгебры. Например, не существует положений для матриц с определенной структурой или для более специализированных разложений, комплексные матрицы не включены. Наконец, JAMA не является классом универсального массива. Вместо этого JAMA фокусируется на принципах математической функциональности, необходимых для выполнения численной линейной алгебры. В результате нет методов для операций с массивами, таких как изменение или применение элементарных функций.

Apache Commons Math – это библиотека автономных математических и статистических компонентов, решающих самые распространенные проблемы [24]. Все алгоритмы использованные в данной библиотеке документированы и

соответствуют общепринятым методам. Если происходит ситуация, когда существуют несколько стандартных алгоритмов решения, то для нескольких реализаций используется шаблон стратегии.

Apache Commons Math содержит множество математических пакетов. Рассмотрим несколько пакетов, отвечающих за расчет описательной статистики и корреляции данных:

Пакет `stat` отвечает за хранилище данных, манипуляции и сводные процедуры [25]. Данный пакет содержит два класса: `Frequency`, отвечающий за частотное распределение, и `StatUtils`, отвечающий за предоставления описательной статистики.

Пакет `stat.descriptive` отвечает за расчет и вывод описательной статистики [26]. Класс `DescriptiveStatistics` поддерживает набор данных одной переменной и вычисляет описательную статистику на основе хранимых данных. Класс `SummaryStatistics` вычисляет сводную статистику для потока значений данных, добавленных с помощью `addValue` метода.

Пакет `stat.correlation` отвечает за вычисления ковариации и корреляции [27]. Содержит пять классов:

1. `Covariance` вычисляет ковариации для пар массивов или столбцов матрицы;
2. `KendallsCorrelation` реализует корреляцию рангов Кендалла;
3. `PearsonsCorrelation` вычисляет коэффициент корреляции Пирсона для пар массивов или столбцов матрицы;
4. `SpearmanCorrelation` реализует Спирменскую ранговую корреляцию;
5. `StorelessCovariance` реализует ковариации, которые не требуют ввода входных данных в память.

1.1.4. Альтернативные статистические пакеты

Помимо библиотек различных платформ, существуют так же пакеты прикладных программ. Рассмотрим некоторые из них.

Matlab – пакет прикладных программ для решения задач технических вычислений [28]. Пакет работает на большинстве современных операционных систем. Statistics Toolbox – пакет расширения Matlab, предлагающий широкий спектр инструментов для статистических вычислений. Основные возможности включают: регрессионный анализ и диагностика с выбором переменной, нелинейное моделирование, моделирование вероятностей и оценка параметров, анализ чувствительности с использованием генератора случайных чисел, управление статистическими процессами и планирование эксперимента. Пакет включает 20 различных распределений вероятностей, включая T, F и Хи-квадрат.

Mathematica – система компьютерной алгебры, широко используемая в научных, инженерных, математических и компьютерных областях [29]. Основные аналитические возможности: решение систем полиномиальных и тригонометрических уравнений и неравенств, а также трансцендентных уравнений, сводящихся к ним, решение рекуррентных уравнений, упрощение выражений, нахождение пределов, интегрирование и дифференцирование функций, нахождение конечных и бесконечных сумм и произведений, решение дифференциальных уравнений и уравнений в частных производных, преобразования Фурье и Лапласа, а также Z-преобразование, преобразование функции в ряд Тейлора, операции с рядами Тейлора: сложение, умножение, композиция, получение обратной функции, вейвлетный анализ. Система также осуществляет численные расчёты.

1.2. Функции описательной статистики

При обработке временных рядов обязателен набор описательной статистики. Рассмотрим распространенные функции описательной статистики [30].

Среднее арифметическое данных

Среднее арифметическое множества чисел – это число, равное сумме всех чисел множества деленное на их количество (1). Данное понятие не относится к

робастной статистике, то есть среднее арифметическое подвержено сильному влиянию больших отклонений.

$$\frac{x_1 + \dots + x_n}{n} \quad (1)$$

Геометрическое среднее данных

Средним геометрическим нескольких положительных вещественных чисел называется такое число, которым можно заменить каждое из этих чисел так, чтобы их произведение не изменилось (2).

$$G(x_1, x_2, \dots, x_n) = \sqrt[n]{x_1 x_2 \dots x_n} = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} \quad (2)$$

Среднее геометрическое двух чисел также называется их средним пропорциональным, поскольку среднее геометрическое g двух чисел a_1 и a_2 обладает следующим свойством: среднее геометрическое относится к первому числу так же, как второе число к среднему геометрическому.

Медиана данных

Медиана в статистике – это число характеризующее выборку. Если все элементы выборки различны, то медиана – это такое число выборки, что ровно половина из элементов выборки больше него, а другая половина меньше него. В более общем случае медиану можно найти, упорядочив элементы выборки по возрастанию или убыванию и взяв средний элемент.

Мода данных

Мода – значение во множестве наблюдений, которое встречается наиболее часто. Для интервального ряда мода определяется по формуле (3):

$$M_o = X_{M_o} + h_{M_o} * \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})}, \quad (3)$$

где X_{M_o} – левая граница модального интервала, h_{M_o} – длина модального интервала, f_{M_o-1} – частота премодального интервала, f_{M_o} – частота модального интервала, f_{M_o+1} – частота послемодального интервала.

Дисперсия данных

Дисперсия в статистике – это способ описания того, как распространяются данные. Общими показателями являются среднеквадратическое отклонение и дисперсия случайной величины.

Среднеквадратическое отклонение в статистике наиболее распространённый показатель рассеивания значений случайной величины относительно её математического ожидания. Рассчитывается по формуле (4):

$$S_0 = \sqrt{\frac{n}{n-1} S^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (4)$$

где S^2 – выборочная дисперсия, x_i – i -й элемент выборки, n – объем выборки, \bar{x} – среднее арифметическое выборки.

Дисперсия случайной величины в статистике является мерой разброса значений случайной величины относительно её математического ожидания. Рассчитывается по формуле (5):

$$D[X] = M[(X - M[X])^2], \quad (5)$$

где X – случайная величина, M – математическое ожидание.

Математическое ожидание находится по формуле (6):

$$M[X] = \sum_{i=1}^n x_i p_i, \quad (6)$$

где p_i – вероятность i -ого элемента выборки.

1.3. Корреляция данных

Большой раздел современной математической статистики занимает корреляционный анализ данных. Данный метод обработки весьма популярен в метеорологии, экономике, социальных науках, контроле качества и др.

Корреляция – статистическая взаимосвязь двух или более случайных величин. При этом изменения значений одной или нескольких из этих величин сопутствуют систематическому изменению значений другой или других величин. Корреляция полезна, поскольку может указывать на прогностические отношения, которые могут быть использованы на практике.

Существует несколько коэффициентов корреляции. Метод вычисления коэффициента корреляции зависит от вида поступивших данных. Для

переменных с интервальной и количественной шкалами необходимо использовать коэффициент корреляции Пирсона. Коэффициент Пирсона является наиболее распространенным коэффициентом корреляции. Данный коэффициент чувствителен только к линейной зависимости между двумя величинами, что делает его менее надежным, по сравнению с другими коэффициентами корреляции. Другие коэффициенты корреляции были разработаны более чувствительными к нелинейным отношениям между величинами, а потому являются более надежными.

В том случае, если, по меньшей мере, одна из двух переменных имеет порядковую шкалу, либо не является нормально распределённой, необходимо использовать ранговую корреляцию Спирмана или Кендалла. Ранговые коэффициенты корреляции, такие как коэффициенты корреляции Спирмана и Кендалла, измеряют степень, в которой по мере увеличения одной переменной другая переменная имеет тенденцию к увеличению, не требуя, чтобы увеличение представлялось линейной зависимостью. Если, по мере увеличения одной переменной, другая уменьшается, коэффициенты ранговой корреляции будут отрицательными. Общепринято рассматривать эти коэффициенты корреляции как альтернативы коэффициенту Пирсона, используемые либо для уменьшения количества вычислений, либо для того, чтобы коэффициент был менее чувствительным к ненормальности в распределениях. Однако этот взгляд не совсем верный, поскольку коэффициенты ранговой корреляции измеряют другой тип отношений, чем коэффициент корреляции Пирсона, и лучше всего воспринимаются как меры другого типа, а не как альтернативная мера корреляции.

В том случае, если, обе переменные являются дихотомическими, то используется четырёхполевая корреляция. Расчёт коэффициента корреляции между двумя недихотомическими переменными имеет смысл только тогда, когда связь между ними линейна (однонаправлена).

В данной работе используются переменные с интервальной и количественной шкалами, поэтому расчеты проводятся для коэффициентов

Пирсона, Спирмана и Кендалла. Рассмотрим расчет рассматриваемых коэффициентов корреляции:

1. Коэффициент корреляции Пирсона рассчитывался по формуле (7):

$$r_{XY} = \frac{\text{cov}_{XY}}{\sigma_X \sigma_Y} = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}}, \quad (7)$$

где $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$, $\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$ – среднее значения выборок.

2. Коэффициент ранговой корреляции Кендалла рассчитывался по формуле (8):

$$\tau = \frac{2S}{n(n-1)}, \quad (8)$$

где n – число наблюдений. Параметр S определялся как $S=P-Q$, где P – суммарное число наблюдений, следующих за текущими наблюдениями с большим значением рангов, Q – суммарное число наблюдений, следующих за текущими наблюдениями с меньшим значением рангов. Если исследуемые данные повторялись, то использовался скорректированный коэффициент корреляции Кендалла (9):

$$\tau = \frac{S}{\sqrt{\left[\frac{n(n-1)}{2} - U_x\right] \left[\frac{n(n-1)}{2} - U_y\right]}}, \quad (9)$$

$$U_x = \frac{\sum t(t-1)}{2},$$

$$U_y = \frac{\sum t(t-1)}{2},$$

где t – число связанных рангов в рядах X и Y соответственно.

3. Коэффициент ранговой корреляции Спирмана рассчитывался по формуле(10):

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \quad (10)$$

где d_i – разности рангов каждой пары сопоставляемых значений.

Значительная корреляция между двумя случайными величинами всегда является свидетельством существования некоторой статистической связи в данной выборке, но эта связь не обязательно должна наблюдаться для другой выборки и иметь причинно-следственный характер. Часто заманчивая простота корреляционного исследования подталкивает исследователя делать ложные интуитивные выводы о наличии причинно-следственной связи между парами признаков, в то время как коэффициенты корреляции устанавливают лишь статистические взаимосвязи. В то же время, отсутствие корреляции между двумя величинами ещё не значит, что между ними нет никакой связи. Например, зависимость может иметь сложный нелинейный характер, который корреляция не выявляет.

Некоторые виды коэффициентов корреляции могут быть положительными или отрицательными. В первом случае предполагается, что мы можем определить только наличие или отсутствие связи, а во втором – также и её направление. Если предположить, что значения переменных задано отношением строгого порядка, то отрицательный коэффициент корреляции говорит о наличии противоположной связи, т.е. чем выше значение одной переменной, тем ниже значение другой. При этом положительное значения коэффициента корреляции говорит о такой связи, при которой увеличение одной переменной связано с увеличением другой. Если значения коэффициента корреляции до 0.3, то это очень слабая корреляция, до 0.5 – слабая корреляция, до 0.7 – средняя корреляция, до 0.9 – высокая корреляция, больше 0.9 – очень высокая корреляция. Возможна так же ситуация отсутствия взаимосвязей, например для независимых случайных величин.

Корреляционный анализ – метод обработки статистических данных, с помощью которого измеряется теснота связи между двумя или более переменными. Корреляционный анализ тесно связан с регрессионным анализом, с его помощью определяют необходимость включения тех или иных факторов в уравнение множественной регрессии, а также оценивают полученное уравнение регрессии на соответствие выявленным связям

(используя коэффициент детерминации). Применение возможно при наличии достаточного количества наблюдений для изучения. На практике считается, что число наблюдений должно не менее чем в 5-6 раз превышать число факторов (также встречается рекомендация использовать пропорцию, не менее чем в 10 раз превышающую количество факторов). В случае если число наблюдений превышает количество факторов в десятки раз, в действие вступает закон больших чисел, который обеспечивает взаимопогашение случайных колебаний.

Данный метод обработки статистических данных весьма популярен в метеорологии, экономике, астрофизике и социальных науках (в частности в психологии и социологии), хотя сфера применения коэффициентов корреляции обширна: контроль качества промышленной продукции, металловедение, агрохимия, гидробиология, биометрия и прочие. В различных прикладных отраслях приняты разные границы интервалов для оценки тесноты и значимости связи.

Популярность данного метода обусловлена относительной простотой в подсчете коэффициентов корреляции, отсутствием необходимости специальной математической подготовки и простотой интерпретации полученных результатов. Сочетание этих факторов привело к его широкому распространению в сфере анализа статистических данных.

Вывод по разделу. В данном разделе были исследованы фреймворки для разработки программного обеспечения, а так же методы описательной статистики и корреляционного анализа. По итогу исследования был выбран пакет математической статистической обработки Apache Commons Math на языке Java.

4. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

4.1 Введение

Целью данной главы ВКР является анализ перспективности разработки программного обеспечения для исследования корреляционных зависимостей метеорологических данных. Это позволит выделить преимущества и недостатки разработки и использования данной программного обеспечения, определить возможные варианты повышения эффективности и результативности его работы.

Достижение такой цели обеспечено выполнением следующих задач:

- оценка коммерческого потенциала и перспективности проведения научных исследований;
- определение возможных альтернатив проведения научных исследований, отвечающих современным требованиям в области ресурсоэффективности и ресурсосбережения;
- планирование научно-исследовательских работ;
- определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования.

4.2 Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения

4.2.1. Потенциальные потребители результатов исследования

Разрабатываемая информационная система предназначена для внутреннего пользования НИ ТПУ. Пользователями программы выступают студенты или сотрудники ТПУ.

Чтобы перейти к рассмотрению потенциальных потребителей необходимо установить целевой рынок и провести его сегментирование.

Объектом исследования выступает программное обеспечение, позволяющее проводить исследования корреляционных зависимостей метеорологических данных, потому группу потребителей могут составить

научные организации, занимающиеся метеорологическими исследованиями. В Западной Сибири было выделено три организации занятыми в этой сфере деятельности: “Среднесибирское управление по гидрометеорологии и мониторингу окружающей среды” (Красноярск), “Томский центр по гидрометеорологии и мониторингу окружающей среды”, “НИЦ «Планета»” (Новосибирск).

Сегментирование рынка можно произвести посредством построения карты сегментирования. Критериями сегментирования были выбраны вид деятельности и территориальная принадлежность.

		Вид деятельности		
		Прогнозирова- ние	Монитор- инг	Исследова- ние
Территория (область)	Краснояр- ский край	/ / / / /	/ / / / /	/ / / / /
	Томская			
	Новоси- бирская	_ _ _ _ _		_ _ _ _ _

Рисунок 18 – Карта сегментирования рынка организаций метеорологической сферы:

 СУГиМОС
  ТЦГиМОС
  «Планета»

Исходя из предлагаемой карты сегментирования рынка можно сказать о том, что на целевом рынке организаций метеорологической сферы деятельности участками с низкой конкуренцией являются Новосибирская область по виду деятельности «мониторинг» и Томская область по виду деятельности «исследования». В рамках выполнения данного проекта ориентирование было именно на участок низкой конкуренции в Томской области.

4.2.2. Анализ конкурентных технических решений

В качестве конкурентных технических решений были выбраны следующие разработки:

- Разработка программного обеспечения для выполнения корреляционного анализа на платформе Java (данная работа) (1);
- Разработка программного обеспечения для выполнения корреляционного анализа на платформе C# (2);
- Разработка программного обеспечения для выполнения корреляционного анализа на платформе Python (3).

Анализа выполнялся с помощью оценочной карты. Результаты конкурентного анализа приведены в Таблице 1.

Таблица 1 – Оценочная карта

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б ₁	Б ₂	Б ₃	К ₁	К ₂	К ₃
1	2	3	4	5	6	7	8
Технические критерии оценки ресурсоэффективности							
1. Скорость работы	0,3	5	4	5	1,5	1,2	1,5
2. Гибкость платформы	0,2	5	3	5	1	0,6	1
3. Простота эксплуатации	0,1	5	4	4	0,5	0,4	0,4
4. Потребность в ресурсах	0,1	4	4	5	0,4	0,4	0,5
5. Функциональные возможности	0,1	5	4	4	0,5	0,4	0,4
Экономические критерии оценки эффективности							
6. Предполагаемый срок эксплуатации	0,05	5	4	4	0,25	0,2	0,2
7. Обслуживание программного обеспечения	0,05	5	3	3	0,25	0,15	0,15
8. Цена	0,1	5	5	5	0,5	0,5	0,5
Итого:	1	39	31	35	4,9	3,85	4,65

Конкурентоспособность решения 1 к решению 2 равна 1,27.

Конкурентоспособность решения 1 к решению 3 равна 1,05.

На основе вышеприведенных оценочной карты и значений конкурентоспособности можно сделать вывод, что решение 1 имеет конкурентные преимущества и является наиболее практичной для реализации. Именно такое решение и использовалось при выполнении текущего проекта.

4.2.3. Технология QuaD

Данная технология была применена для того, чтобы оценить характеристики качества и перспективности программного обеспечения для исследования корреляционных зависимостей. В основе этой технологии лежит нахождение средневзвешенной величины групп показателей потенциала и качества разработки.

Оценочная карта для сравнения конкурентных технических решений приведена в Таблице 2.

Таблица 2 – Оценочная карта QuaD

Критерии оценки	Вес критерия	Баллы	Максимальный балл	Относительное значение	Средневзвешенное значение
Показатели оценки качества разработки					
1. Скорость работы	0,2	95	100	0,95	0,19
2. Гибкость платформы	0,2	95	100	0,95	0,19
3. Простота эксплуатации	0,1	85	100	0,85	0,085
4. Потребность в ресурсах	0,1	80	100	0,8	0,08
5. Функциональные возможности	0,1	90	100	0,9	0,09
Показатели оценки коммерческого потенциала разработки					
6. Перспективность	0,15	65	100	0,65	0,0975
7. Обслуживание	0,05	85	100	0,85	0,0425
8. Цена	0,1	100	100	0,1	0,01
Итого:					0,785

Оценка качества и перспективности определяется P_{cp} – средневзвешенным значением показателя качества и перспективности научной разработки, который определяется по формуле:

$$P_{cp} = \sum V_i * B_i,$$

где V_i – вес показателя (в долях единицы);

B_i – средневзвешенное значение i -го показателя.

Полученное значение 0,785 позволяет отнести разработку к категории проектов со средней перспективностью (показатель от 0,79 до 0,60).

4.2.4. SWOT-анализ

SWOT-анализ является одним из самых частоиспользуемых методов в менеджменте и маркетинге. Он позволяет комплексно рассмотреть состояние момента на текущий момент и позволяет достичь понимания того, какие действия должны быть предприняты для нейтрализации слабых сторон и угроз, а также для максимизирования возможностей на основе сильных сторон проекта. Результаты SWOT-анализа представлены в Таблице 3.

Таблица 3 – SWOT-анализ

	<p>Сильные стороны проекта: С1. Удобная навигация по NetCDF-файлам. С2. Графическое отображение результатов. С3. Разные варианты базисов вейвлет-анализа. С4. Гибкость платформы проекта.</p>	<p>Слабые стороны проекта: Сл1. Необходимо наличие доступа к сети Internet для получения файлов наборов данных. Сл2. Необходимость ручного задания некоторого числа параметров. Сл3. Большое число этапов перед получением результатов. Сл4. Строго заданная структура файлов наборов данных.</p>
<p>Возможности: В1. Рост спроса на подобные разработки. В2. Рост рыночной стоимости подобных разработок. В3. Малое число конкурентных разработок. В4. Рост спроса на аналитические статьи, которые можно генерировать исходя из результатов приложения.</p>	<p>В1В2С2С3 Разработка и общий доступ к документации к разработке упростит интеграцию новыми потребителями. В3В4С2С3 Проведение активной маркетинговой кампании с целью популяризации программы среди научных сотрудников и преподавателей.</p>	<p>В1В2В4Сл2Сл3 Автоматизация определения параметров анализа, повышение уровня интерактивности процесса.</p>
<p>Угрозы: У1. Сбои в работе сети Internet. У2. Изменение структуры файлов наборов данных. У3. Появление типов анализов данных, имеющих преимущество над используемым. У4. Появление конкурентного продукта.</p>	<p>У1С2С3С4 Введение логов процесса работы ПО. У2С1С3 Обеспечение текстового вывода результатов. У3С3С4 Расширение числа поддерживаемых методов корреляционного анализа. У4С4 Отслеживание рыночной ситуации с целью внедрения актуальных и</p>	<p>У1Сл1Сл2Сл4 Обеспечение поддержки локального выбора файла набора данных. У2Сл2Сл4 Избежание прямолинейности и строгого задания извлекаемой структуры получаемых данных.</p>

	востребованных функций данного рода разработок.	
--	-------------------------------------------------	--

Таким образом, в результате SWOT-анализа были выявлены направления, по которым необходимо развивать проект. Наиболее важными задачами на краткосрочный период являются расширение количества используемых методов анализов данных и оптимизация интерфейса программного обеспечения для упрощения взаимодействия с пользователем.

4.3. Определение возможных альтернатив проведения научных исследований

В качестве морфологических характеристик объекта исследования можно выделить тип данных, язык программирования, тип анализа данных и тип приложения. Морфологическая матрица с рассмотрением альтернативных решений приведена в Таблице 4.

Таблица 4 – Морфологическая матрица альтернативных решений

Альтернативы	1	2	3
А. Тип данных	NetCDF	Grib	Grib2
Б. Язык программирования	Python	C#	Java
В. Тип анализа данных	Корреляционный	Статистический	На основе искусственных нейронных сетей
Г. Тип приложения	Веб-приложение	Десктопное приложение	Мобильное приложение

Данную работу можно описать как выбор наиболее желательных функционально конкретных решений вида А1Б3В1Г2:

- тип данных: NetCDF – является стандартным типом представления климатический данных реанализа мировыми центрами данных. Имеет четкую структуру и описание содержащихся данных;
- язык программирования: Java – универсален, кроссплатформенен, имеет крупную базу готовых решений, библиотек и инструкций;
- тип анализа данных: корреляционный – перспективное направление, позволяющее провести анализ как стационарных, так и

нестационарных данных, выявить и отобразить закономерности, аномалии и прочую информацию об исходном ряде данных;

- тип приложение: десктопное приложение – необходимо обеспечить доступ и навигацию по файлам. Так же, необходимо предусмотреть сложный интерфейс с большим набором функций. Удобнее всего реализовать такое решение в виде стандартного десктопного приложения.

4.4. Планирование научно-исследовательских работ

4.4.1. Структура работ в рамках научного исследования

Перечень этапов и работ в рамках научного исследования представлен в Таблице 5.

Таблица 5 – Перечень этапов, работ, распределение исполнителей

Основной этап	№	Содержание работ	Исполнитель
Разработка ТЗ	1	Разработка требований	Научный руководитель (100%)
	2	Анализ требований	Научный руководитель (25%), инженер-программист (75%)
	3	Анализ предметной области	Научный руководитель (50%), инженер-программист (50%)
	4	Выбор инструментов и методологий	Научный руководитель (50%), инженер-программист (50%)
Проектирование ПО	5	Создание архитектуры ПО	Инженер-программист (100%)
	7	Проектирование компонент и классов ПО	Инженер-программист (100%)
Разработка ПО	8	Создание логики доступа к NetCDF-файлам	Инженер-программист (100%)
	9	Создание пользовательского интерфейса	Инженер-программист (100%)
	10	Создание логики корреляционного анализа	Инженер-программист (100%)
Тестирование и отладка	11	Тестирование ПО с различными данными	Инженер-программист (100%)
Обобщение и оценка	12	Оценка полученных результатов работы	Научный руководитель (75%), инженер-программист (25%)

результатов			
Оформление отчета по НИР	13	Оформление расчетно-пояснительной записки	Инженер-программист (100%)
Подведение итогов	14	Оценка выполненной работы	Научный руководитель (50%), инженер-программист (50%)

4.4.2. Определение трудоемкости выполнения работ

Необходимо произвести оценку трудоемкости выполнения вышеперечисленных работ. Для этого необходимо рассмотреть минимальное и максимальное время выполнения каждой из этих работ. Расчет ожидаемого значения трудоемкости будет выполняться по следующей формуле:

$$t_{ож,i} = \frac{(3t_{min i} + 2t_{max i})}{5},$$

где $t_{min i}$ – минимальное время выполнение i -ой работы чел.-дн.;

$t_{max i}$ – максимальное время выполнения i -ой работы чел.-дн.

Параллельность работ не учитывается, потому как инженер-программист является единственным исполнителем задач.

4.4.3. Разработка графика проведения научного исследования

Для подсчета продолжительности выполнения i -ой работы в календарных днях применяется формула:

$$T_{ki} = T_{pi} * k_{кал},$$

где T_{pi} – продолжительность выполнения i -ой работы в рабочих днях;

$k_{кал}$ – коэффициент календарности.

Коэффициент календарности считается по формуле:

$$k_{кал} = \frac{T_{кал}}{T_{кал} - T_{вых} - T_{пр}},$$

где $T_{кал}$ – количество календарных дней в году;

$T_{вых}$ – количество выходных дней в году;

$T_{пр}$ – количество праздничных дней в году.

$$k_{кал 2018} = \frac{365}{365 - 118} = 1,48$$

Временные показатели проведения научного исследования приведены в Таблице 6.

Таблица 6 – Временные показатели научного исследования

Этап	Исполнители	Продолжительность работ, дни			Длительность работ, чел/дн.			
		t_{min}	t_{max}	$t_{ож}$	T_{pi}		T_{ki}	
					НР	ИП	НР	ИП
Разработка требований	НР	1	2	1,4	1,4	-	2	-
Анализ требований	НР, ИП	1	2	1,4	0,35	1,05	1	2
Анализ предметной области	НР, ИП	3	5	3,8	1,9	1,9	3	3
Выбор инструментов и методологий	НР, ИП	5	10	7	3,5	3,5	6	6
Создание архитектуры ПО	ИП	2	10	5,2	-	5,2	-	8
Проектирование компонент и классов ПО	ИП	4	8	5,6	-	5,6	-	9
Создание логики доступа к NetCDF-файлам	ИП	10	15	12	-	12	-	18
Создание пользовательского интерфейса	ИП	15	20	17	-	17	-	26
Создание логики корреляционного анализа	ИП	10	15	12	-	12	-	18
Тестирование ПО с различными данными	ИП	5	10	7	-	7	-	11
Оценка полученных результатов работы	НР, ИП	2	2	2	1,5	0,5	3	1
Оформление расчетно-пояснительной записки	ИП	15	20	17	-	17	-	26
Оценка выполненной работы	НР, ИП	5	10	7	3,5	3,5	6	6
Итого:				98,4	12,15	86,25	21	134

НР – Научный руководитель; ИП – Инженер-программист.

На основе приведенной таблицы строится календарный план-график. В таком плане-графике предусмотрены разбиения по декадам (10 дней) и месяцам.

Таблица 7 – Календарный план-график работ

№	Вид работ	Исполнители	T_{ki} , кал.-дн.		Продолжительность выполнения работ														
			НР	ИП	январь			февраль			март			апрель			май		
					2	3	1	2	3	1	2	3	1	2	3	1	2	3	
1	Разработка требований	НР	2	-	█														
2	Анализ требований	НР, ИП	1	2	█														

3	Анализ предметной области	НР, ИП	3	3	█														
4	Выбор инструментов и методологий	НР, ИП	6	6	█														
5	Создание архитектуры ПО	ИП	-	8		█													
7	Проектирование компонент и классов ПО	ИП	-	9			█												
8	Создание логики доступа к NetCDF-файлам	ИП	-	18				█											
9	Создание пользовательского интерфейса	ИП	-	26					█										
10	Создание логики корреляционного анализа	ИП	-	18						█									
11	Тестирование ПО с различными данными	ИП	-	11							█								
12	Оценка полученных результатов работы	НР, ИП	3	1															█
13	Оформление расчетно-пояснительной записки	ИП	-	26															█
14	Оценка выполненной работы	НР, ИП	6	6															█

4.4.4. Бюджет научно-технического исследования (НТИ)

Расчет материальных затрат НТИ

Материальные затраты определяются материалами и ресурсами, используемыми в течении процесса разработки текущего научно-исследовательского проекта. К таким, в данном случае, можно отнести только электроэнергию, обеспечивающую работу ЭВМ инженера-программиста. Канцелярские принадлежности предусмотрены расходами научной организации.

В час компьютер потребляет около 0,3 кВт. Возьмем стоимость 1 кВт*ч электричества для юридических лиц равной 5,8 рублей. В день компьютер используется в среднем 6 часов. Тогда, стоимость материальных затрат:

$$Z_{\text{мат}} = 0,3 \text{ кВт} * 5,8 \frac{\text{руб}}{\text{кВт} * \text{ч}} * 6 \text{ ч} * 134 \text{ д} = 1398,96 \text{ руб.}$$

Расчет на специальное оборудование для научных - экспериментальных работ

Затраты на специальное оборудование будет состоять из монитора, системного блока, периферийного оборудования. В совокупности затраты выйдут в размере 11500 рублей.

Основная заработная плата исполнителей темы

Расчет бюджета НТИ состоит из расчета материальных затрат, затрат на зарплаты руководителю и инженера. Материальные затраты составляют только расходные материалы и амортизация оборудования.

Оклад руководителя от ТПУ (доцента, к.т.н) составляет 26300 рубля (без учета районного коэффициента).

Оклад младшего научного сотрудника составляет 17000 руб. (без учета районного коэффициента).

Таблица 8 – Распределение рабочего времени

Показатели рабочего времени	Дни
Календарное число	365
Количество нерабочих дней (праздники/выходные)	118
Действительный годовой фонд рабочего времени	247

С учетом районного коэффициента, равного 30% от оклада, получается месячная заработная плата:

$$z_{\text{м}}^{\text{рук}} = 26300 * 1,3 = 34190 \text{ руб.}$$

$$z_{\text{м}}^{\text{разр}} = 17000 * 1,3 = 22100 \text{ руб.}$$

Зная месячную заработную плату каждого участника проекта, можно рассчитать соответствующую среднедневную заработную плату. Количество месяцев работы без отпуска принимается равным 11,2 (считается отпуск длиной 24 рабочих дня при 6-дневной рабочей неделе):

$$z_{\text{дн}}^{\text{рук}} = \frac{34190 * 11,2}{247} = 1550,31 \text{ руб.},$$

$$z_{\text{дн}}^{\text{разр}} = \frac{22100 * 11,2}{247} = 1002,1 \text{ руб.}$$

Тогда, основная зарплата за период НТИ:

$$Z_{\text{осн}}^{\text{рук}} = 1550,31 * 21 = 32556,51 \text{ руб.},$$

$$Z_{\text{осн}}^{\text{разр}} = 1002,1 * 134 = 134292,12 \text{ руб.}$$

За статью расходов «Затраты по основной зарплате»:

$$Z_{\text{осн}} = 32556,51 + 134292,12 = 166848,63 \text{ руб.}$$

Дополнительная заработная плата исполнителей темы

С учётом основной заработной платы, можно посчитать дополнительную заработную плату в размере 12 % от основной:

$$Z_{\text{доп}}^{\text{рук}} = k_{\text{доп}} * Z_{\text{осн}} = 0,12 * 32556,51 = 3906,78 \text{ руб.}$$

$$Z_{\text{доп}}^{\text{разр}} = k_{\text{доп}} * Z_{\text{осн}} = 0,12 * 134292,12 = 16115,05 \text{ руб.}$$

За статью расходов «Затраты по дополнительной зарплате»:

$$Z_{\text{доп}} = 3906,78 + 16115,05 = 20021,83 \text{ руб.}$$

Отчисления во внебюджетные фонды (страховые отчисления)

Величина отчислений во внебюджетные фонды определяется как:

$$Z_{\text{внеб1}} = k_{\text{внеб}} (Z_{\text{осн}} + Z_{\text{доп}}) = 0,302 * (32556,51 + 3906,78) = 11011,91 \text{ руб.}$$

$$Z_{\text{внеб1}} = k_{\text{внеб}} (Z_{\text{осн}} + Z_{\text{доп}}) = 0,302 * (134292,12 + 16115,05) = 45422,97 \text{ руб.}$$

За статью расходов «Отчисления во внебюджетные фонды»:

$$Z_{\text{внеб}} = 11011,91 + 45422,97 = 56434,88 \text{ руб.}$$

Расчет затрат на научные и производственные командировки

Научных и производственных командировок в данном объекте исследования не производилось.

Контрагентные расходы

Данная статья расходов описывает затраты, связанные с привлечением сторонних организаций для выполнения работ связанных с текущим объектом исследования. Привлечения сторонних организаций не производилось, потому контрагентные расходы равны нулю.

Накладные расходы

Накладные расходы учитывают прочие затраты организации. Коэффициент учета накладных расходов берется равным 0,16. Тогда, расчет накладных расходов:

$$Z_{\text{накл}} = (\text{сумма статей } 1 - 7) * 0,16 = 40992,68 \text{ руб.}$$

Формирование бюджета затрат научно-исследовательского проекта

Таблица 9 – Бюджет затрат по каждому исполнению НТИ

Наименование статьи	Сумма руб.		
	Исп. 1	Исп. 2	Исп. 3
1. Материальные затраты НТИ	1398,96	1357,20	1461,60
2. Затраты на спец. оборудование	11500		
3. Затраты по основной з/п	166848,63	170581,06	171300,20
4. Затраты по доп. з/п	20021,83	20469,72	20556,02
5. Отчисления во внебюджетные фонды	56434,88	57697,36	57940,58
6. Затраты на научные и производственные командировки	0		
7. Контрагентские расходы	0		
8. Накладные расходы	40992,68	41856,85	42041,34
9. Бюджет затрат НТИ	297196,98	303462,19	304799,74

4.5. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Определение эффективности происходит посредством нахождения двух средневзвешенных величин: финансовой эффективности и ресурсоэффективности.

Интегральный финансовый показатель разработки рассчитывается как:

$$I_{\text{финр}}^{\text{исп.}i} = \frac{\Phi_{p,i}}{\Phi_{\text{max}}},$$

Где $\Phi_{p,i}$ – стоимость i -го варианта исполнения;

Φ_{max} – максимальная стоимость исполнения научно-исследовательского проекта.

Тогда, согласно Таблице 10:

$$I_{\text{финр}}^{\text{исп1}} = \frac{297196,98}{304799,74} = 0,975,$$

$$I_{\text{финр}}^{\text{исп2}} = \frac{303462,19}{304799,74} = 0,996,$$

$$I_{\text{финр}}^{\text{исп3}} = 1$$

Интегральный показатель ресурсоэффективности можно определить следующим образом:

$$I_{p,i} = \sum a_i b_i$$

где $I_{p,i}$ – интегральный показатель ресурсоэффективности для i -го варианта разработки;

a_i – весовой коэффициент i -го варианта разработки;

b_i – балльная оценка i -го варианта исполнения разработки, устанавливаемая экспертным путем по выбранной шкале оценивания.

Расчет интегральных показателей ресурсоэффективности приведен в Таблице 10.

Таблица 30 – Сравнительная оценка характеристик вариантов исполнения проекта.

Критерии оценки	Вес критерия	Баллы		
		Исп. 1	Исп. 2	Исп. 2
1. Скорость работы	0,3	5	4	5
2. Гибкость платформы	0,25	5	3	5
3. Простота эксплуатации	0,15	5	4	4
4. Потребность в ресурсах	0,1	4	4	5
5. Функциональные возможности	0,2	5	4	4
Итого:	1	4,9	3,75	4,65

$$I_{p-\text{исп1}} = 0,3 * 5 + 0,25 * 5 + 0,15 * 5 + 0,1 * 4 + 0,2 * 5 = 4,9;$$

$$I_{p-\text{исп2}} = 0,3 * 4 + 0,25 * 3 + 0,15 * 4 + 0,1 * 4 + 0,2 * 4 = 3,75;$$

$$I_{p-\text{исп3}} = 0,3 * 5 + 0,25 * 5 + 0,15 * 4 + 0,1 * 5 + 0,2 * 4 = 4,65.$$

Интегральный показатель эффективности вариантов исполнения разработки рассчитывается на основании интегрального показателя ресурсоэффективности и интегрального финансового показателя согласно формуле:

$$I_{\text{исп.i}} = \frac{I_{p-\text{исп}i}}{I_{\text{финр}}^{\text{исп.i}}}$$

Сравнительная эффективность проекта считается согласно формуле:

$$\mathcal{E}_{\text{ср}} = \frac{I_{\text{исп.1}}}{I_{\text{исп.2}}}$$

Расчет сравнительной эффективности разработок приведет в Таблице 11.

Таблица 11. Сравнительная эффективность разработки

Показатели	Исп.1	Исп.2	Исп.3
Интегральный финансовый показатель разработки	0,975	0,996	1
Интегральный показатель ресурсоэффективности разработки	4,9	3,75	4,65
Интегральный показатель эффективности	5,026	3,765	4,65
Сравнительная эффективность вариантов исполнения	1,335	1	1,235

Вывод по разделу. Исходя из проведенного анализа сравнительной эффективности вариантов исполнения можно сделать вывод, что Исполнение 1 является более эффективным и предпочтительным вариантом из всех предложенных. Так же, согласно подсчету бюджета затрат по каждому из вариантов следует подчеркнуть, что Исполнение 1 является так же самым выгодным. Таким образом, в настоящей работе реализована самая подходящая вариация проекта.

5. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

5.1. Производственная безопасность

В стандарте ГОСТ 12.0.003–74 «Опасные и вредные производственные факторы» рассматриваются вредные и опасные факторы, подразделяющиеся по природе действия на следующие группы [32]:

- Физические;
- Химические;
- Биологические;
- Психофизиологические.

Так как разработка приложения ведется в кабинете при использовании персонального компьютера (далее ПК), выделяют следующие вредные и опасные факторы (Таблица 12).

Таблица 12– Опасные и вредные факторы при выполнении работ

Факторы (по ГОСТ 12.0.003-74)		Нормативные документы
Вредные	Опасные	
1. Отклонение показателей микроклимата;	1. Удары электрическим током;	1. ГОСТ 12.0.003–74; 2. СанПиН 2.2.4–548–96 [33]; 3. СанПиН 2.2.2/2.4.1340–03 [34]; 4. СанПиН 2.2.1/2.1.1.1278–03 [35]; 5. ГОСТ Р 50739-95 [36].
2. Недостаточная освещенность рабочей зоны;	2. Короткое замыкание;	
3. Повышенный уровень шума и вибрации от вентиляторов охлаждения компьютера;	3. Статическое электричество.	
4. Умственное перенапряжение и монотонность труда.		

5.1.1. Требования к микроклимату

Показатели микроклимата должны обеспечивать сохранение теплового баланса человека с окружающей средой и поддержание оптимального или

допустимого теплового состояния организма. Показателями, характеризующими микроклимат в помещении, являются:

- Температура воздуха;
- Температура поверхностей;
- Относительная влажность воздуха;
- Скорость движения воздуха;

Оптимальные показатели температур воздуха, поверхностей, относительной влажности и скорости воздуха должны соответствовать значениям, указанным в таблице 13 и распространяться на всю рабочую зону. Температура воздуха в рабочей зоне в течение смены не должна выходить за пределы оптимальных величин, указанных в таблице 13 для отдельных категорий работ. Выполняемая работа относится к категории Ia – производимая сидя и сопровождающаяся незначительным физическим напряжением. Перепады температуры воздуха по высоте и по горизонтали, а также изменения температуры воздуха в течение смены, при обеспечении оптимальных величин микроклимата на рабочих местах, не должны превышать 2 °С и выходить за пределы величин, указанных в таблице 13.

Таблица 13 – Параметры микроклимата для помещений, где установлен компьютер согласно СанПиН 2.2.4–548–96

Время года	Параметр микроклимата			
	Температура воздуха, °С	Температура поверхностей	Влажность, %	Скорость движения воздуха, м/с
Холодный	22-24	21-25	40-60	0,1
Теплый	23-25	22-26	40-60	0,1

В целях профилактики неблагоприятного воздействия микроклимата должны быть использованы защитные мероприятия, такие как системы местного кондиционирования воздуха, компенсация неблагоприятного воздействия одного параметра микроклимата изменением другого, помещения для отдыха и обогрева, регламентация времени работы, в частности, перерывы в работе, сокращение рабочего дня.

5.1.2. Расчет уровня шума

Согласно СанПиН 2.2.2/2.4.1340–03 в производственных помещениях с использованием ПК уровни шума на рабочих местах не должны превышать предельно допустимых значений. Шум на рабочем месте вызван следующим оборудованием: винчестером в системном блоке, вентиляторами, кулерами охлаждения процессора ПК, монитор, клавиатура. Уровень шума на рабочем месте не должен превышать 50 дБ. В таблице 14 приведены уровни шума из различных источников.

Таблица 14 – Уровень звукового давления различных источников

Источник шума	Уровень шума, дБ
Жесткий диск	40
Вентилятор	45
Монитор	17
Клавиатура	10
Принтер	45
Сканер	42

Уровень шума, возникающий от нескольких некогерентных источников, работающих одновременно, подсчитывается на основании принципа энергетического суммирования излучений отдельных источников.

$$L_{\Sigma} = 10 \lg \sum_{i=1}^{i=n} 10^{0.1L_i} \quad (1)$$

где L_i – уровень звукового давления, i – го источника шума, n = количество источников шума.

Подставив в формулу значения уровня звукового давления для каждого оборудования, получим:

$$L_{\Sigma} = 10 * \log (10^4 + 10^{4.5} + 10^{1.7} + 10^1) = 46,2 \text{ дБ}$$

Полученное значение не превышает допустимую норму, поэтому использование специальных средств защиты не требуется. В случае превышения допустимой нормы для снижения уровня шума стены и потолок помещений, где установлены компьютеры, могут быть оснащены звукопоглощающими материалами.

5.1.3. Освещенность

Производственное освещение – это система устройств и мер, обеспечивающих благоприятную работу зрения человека в процессе труда. Правильно спроектированное и выполненное производственное освещение улучшает условия зрительной работы, снижает утомляемость, способствует повышению производительности труда, благотворно влияет на производственную среду, оказывая положительное психологическое воздействие на работника, повышает безопасность труда и снижает травматизм.

Недостаточность освещения приводит к напряжению зрения, ослабляет внимание, приводит к наступлению преждевременной утомленности. Чрезмерно яркое освещение вызывает ослепление, раздражение и резь в глазах. Неправильное направление света на рабочем месте может создавать резкие тени, блики, дезориентировать работающего.

К системам производственного освещения предъявляются следующие требования:

- соответствие уровня освещённости рабочих мест характеру выполняемой зрительной работы;
- достаточно равномерное распределение яркости на рабочих поверхностях и в окружающем пространстве;
- отсутствие резких теней, прямой и отраженной повышенной яркости светящихся поверхностей, вызывающей ослепление;
- постоянство освещённости во времени;
- оптимальная направленность излучаемого осветительными приборами светового потока;
- долговечность, экономичность, пожаро- и электробезопасность, эстетичность, удобство и простота эксплуатации.

Согласно СанПиН 2.2.2/2.4.1340-03, рабочие столы следует размещать таким образом, чтобы мониторы были ориентированы боковой стороной к световым проемам, чтобы естественный свет падал преимущественно слева,

искусственное освещение в помещениях должно осуществляться системой общего равномерного освещения.

Естественное освещение осуществляется через два оконных проема размером 2 на 2,5 метра по наружной стене. Нормируемые показатели естественного, искусственного и совмещенного освещения в соответствии с СанПиНом 2.2.1/2.1.1.1278–03 указаны в таблице 15.

Таблица 15– Нормируемые показатели естественного, искусственного и совмещенного освещения в соответствии с СанПиНом 2.2.1/2.1.1.1278–03

Помещени я	Рабочая поверхност ь и плоскость нормирова ния КЕО и освещенно сти и высота плоскости над полом, м	Естественное освещение КЕО ен, %		Совмещенное освещение КЕО ен, %		Искусственное освещение			
		при верхне м или комбин ирован ном освеще нии	при боков ом освещ ении	при верхне м или комбин ирован ном освеще нии	при боков ом освещ ении	освещенность , лк		показа тель диско м форта, М, не более	коэффи циент пульса ции освещё нности, Кп, %, не более
						при комб иниро ванно м освещ ении	при обще м освещ ении		
1	2	3	4	5	6	7	8	9	10
Каби неты	Г- 0,8	3,2	1,0	1,8	06	400	300	40	15

Чтобы поддерживать освещение в помещении по всем соответствующим нормам, необходимо хотя бы два раза в год проводить чистку стекол и светильников, а также по мере необходимости заменять перегоревшие лампы. В утреннее и вечернее время вводится общее искусственное освещение. Основными источниками искусственного освещения являются люминесцентные лампы белого и дневного света ЛБ-20 и ЛД-20. Следует ограничивать отраженную блескость на рабочих поверхностях (экран, стол, клавиатура) за счет правильного выбора и расположения светильников, яркость бликов на экране не должна превышать 40 кд/м². Рассматриваемое помещение соответствует указанным нормированным показателям, как в дневное время суток, так и в вечернее.

5.1.4. Монотонность труда и умственное перенапряжение

Монотонный режим работы связан с однообразным повторением рабочих операций. Опасность монотонности заключается в снижении внимания к процессу производства, быстрой утомляемости и снижению интереса к трудовому процессу, что может повлиять на безопасность труда в целом. Для борьбы с монотонностью используются следующие меры:

- расширение круга обязанностей, усложнение работы или объединение ее в комплексы;
- организация 5-ти минутных перерывов;
- увеличение числа поставленных целей, за счет разделения одной общей на несколько промежуточных.

Умственное перенапряжение возникает в результате длительной умственной работе и проявляется в снижении работоспособности. Факторами вызывающими умственное перенапряжение являются:

- Длительное и повышенное внимание и концентрация;
- Высокие требования к самореализации и самоподготовке;
- Ощущение недостатка времени
- Неправильное питание работника (приводит к недостатку энергии для работы мозга).

Для борьбы с умственным перенапряжением используется выполнение различных физических упражнений, а так же организация 5-ти минутных перерывов и перерыва на обед.

5.1.5. Техника электробезопасности

В соответствии с СанПиН 2.2.2/2.4.1340–03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы» помещения, где размещаются рабочие места с ПК, должны быть оборудованы защитным заземлением в соответствии с техническими требованиями по эксплуатации электроустановок и вычислительной техники. Рабочие места с ПК не следует размещать вблизи силовых кабелей и вводов, высоковольтных

трансформаторов, технологического оборудования, создающего помехи в работе ПК.

Помещение, в котором проводилась работа, относится к помещениям без повышенной опасности, так как в нем отсутствуют условия, которые могут создать повышенную или особую опасность.

К организационным мерам электробезопасности относятся:

1. Инструктаж

Целью инструктажа является сообщение работникам знаний, необходимых для правильного и безопасного выполнения ими своих профессиональных обязанностей, а также формирования у работников убеждения в объективной и абсолютной необходимости выполнения правил и норм безопасной жизнедеятельности в производственной среде.

Существуют следующие виды инструктажа:

- вводный инструктаж
- первичный инструктаж
- периодический (повторный)

2. Правильная организация рабочего места

Организация рабочего места заключается в выполнении ряда мероприятий, которые обеспечивают рациональный и безопасный трудовой процесс, и эффективное использование орудий и предметов труда, что повышает производительность и способствует снижению утомляемости работающих. Так, например, правильно выбранная рабочая поза (с возможностью её перемены) исключает или сводит к минимуму вредное влияние выполняемой работы на организм человека.

3. Режим труда и отдыха

Оптимальный режим труда и отдыха – это такое чередование периодов работы с периодами отдыха, при котором достигается наибольшая эффективность деятельности человека и хорошее состояние его здоровья. Он оказывает благотворное влияние на функциональное состояние человека.

Электрические установки, источником работы которых является переменный ток напряжением 220В и частота 50 Гц, к которым относится большинство оборудования ПК, представляют для человека большую потенциальную опасность, так как в процессе эксплуатации (проведение регламентных работ) человек может коснуться частей оборудования, находящихся под напряжением. Специфическая опасность электроустановок состоит в том, что токоведущие проводники, оказавшиеся под напряжением в результате повреждения изоляции, не подают каких-либо сигналов, которые бы предупреждали об опасности. Для защиты от поражения электрическим током все токоведущие части должны быть защищены от случайных прикосновений и заземлены. Питание устройства должно осуществляться от силового щита через автоматический предохранитель, срабатывающий при коротком замыкании нагрузки.

В соответствии с правилами электробезопасности в помещении осуществляется постоянный контроль состояния электропроводки, предохранительных щитов, шнуров, с помощью которых включаются в электросеть ПК, осветительные приборы, другие электроприборы. Также в помещении отсутствуют токопроводящая пыль, электрически активная среда, возможность одновременного прикосновения к металлическим частям прибора и заземляющему устройству, высокая температура и сырость.

Основным опасным фактором является опасность поражения электрическим током. Исходя из анализа состояния помещения, данное помещение по степени опасности поражения электрическим током можно отнести к классу помещений без повышенной опасности. В помещении подавляющая часть электрической проводки является скрытой. Поражение электрическим током возможно только при возникновении оголенных участков на кабеле, а также нарушении изоляции распределительных устройств, однако в помещении кабель имеет двойную изоляцию, поэтому опасность поражения значительно снижается. Не исключается также опасность поражения и от токоведущих частей компьютера в случае их пробоя и нарушении изоляции.

Для устранения опасности поражения электрическим током регулярно проводится осмотр кабелей, проводов, электрических розеток и токоведущих частей компьютера. А также, перед началом работы за компьютером каждый работник проходит инструктаж по технике безопасности.

Возникающие при прикосновении к любому из элементов ПК разрядные токи статического электричества могут привести к выходу ПК из строя. Для снижения величины возникающих зарядов статического электричества в помещении покрытие полов выполнено из однослойного поливинилхлоридного антистатического линолеума. К мерам защиты от статического электричества также можно отнести общее и местное увлажнение воздуха.

Компьютер также является и источником статического электричества. Местами скопления статических зарядов, как правило, служит поверхность экрана монитора. Для уменьшения статического электричества на поверхности монитора следует раз в 6 часов протирать экран влажной материей.

5.2. Экологическая безопасность

В работе по реализации проекта не оказывается значительного влияния на окружающую среду, так как в процессе работы не использовались вредные химические соединения. Самой серьезной проблемой, с которой столкнулись при выполнении работы, является потребление электроэнергии. С увеличением количества компьютерных систем, внедряемых в производственную сферу, увеличится и объем потребляемой ими электроэнергии. Рост энергопотребления приводит к таким экологическим нарушениям, как изменение климата – накопление углекислого газа в атмосфере Земли (парниковый эффект).

Из этого можно сделать простой вывод, что необходимо стремиться к снижению энергопотребления, то есть разрабатывать и внедрять системы с малым энергопотреблением. В современных компьютерах, повсеместно используются режимы с пониженным потреблением электроэнергии при длительном простое.

Техника, вышедшая из строя, утилизируется согласно ГОСТ Р 50739–95[5]. Люминесцентные лампы, вышедшие из строя, сдаются в специализированный пункт приема. Для уменьшения отходов, связанных с расходными материалами (бумага, ручки, картриджи и т.д.), можно использовать повторно переработанную бумагу или использовать двухстороннюю печать.

5.3.Безопасность в чрезвычайных ситуациях

Чрезвычайная ситуация - это обстановка на определенной территории, сложившаяся в результате аварии, опасного природного явления, катастрофы, стихийного или иного бедствия, которые могут повлечь или повлекли за собой человеческие жертвы, ущерб здоровью людей или окружающей природной среде, значительные материальные потери и нарушение условий жизнедеятельности людей.

Мероприятия по обеспечению пожарной безопасности

Пожарная безопасность регламентируется федеральным законом «Технический регламент о требованиях пожарной безопасности». Опасными факторам пожара для людей являются открытый огонь, искры, повышенная температура воздуха и предметов, токсичные продукты горения, дым, пониженная концентрация кислорода, обрушение и повреждение зданий, сооружений, установок, а также взрыв.

Организационными мероприятиями по обеспечению пожарной безопасности являются: обучение рабочих и служащих правилам пожарной безопасности; разработка и реализация норм и правил пожарной безопасности, инструкций о порядке работы в помещениях; изготовление и применение средств наглядной агитации по обеспечению пожарной безопасности.

Основной причиной возникновения пожара в помещениях с электронной техникой является неисправность проводки. Вероятность возгорания самих электронных устройств чрезвычайно мала.

В качестве оперативных средств тушения пожара применяются порошковые огнетушители ОПУ–5. Сеть электропитания оборудуется входным

рубильником, позволяющим в оперативном порядке отключить электропитание во всем здании. Для обеспечения эвакуации людей в случае пожара помещения должны иметь не менее двух выходов шириной не менее одного метра и высотой не менее двух метров.

Наиболее частыми причинами пожаров являются нарушения правил пожарной безопасности и технологических процессов, неправильная эксплуатация электросети и оборудования, грозовые разряды.

Одна из главных причин травм, связанных с действием электрического тока, слабые знания правил электробезопасности. Нарушение правил электробезопасности при использовании электроустановок и непосредственное соприкосновение с токоведущими частями электроустановок, находящихся под напряжением, создает опасность поражения электрическим током.

5.4. Правовые и организационные вопросы обеспечения безопасности

Под безопасностью понимаются защитные мероприятия и средства, обеспечивающие снижение опасности до минимальной степени риска, когда негативные факторы не превышают допустимой величины.

Трудовой кодекс регулирует вопросы, относящиеся к охране и организации труда. СанПиН 2.2.2/2.4.1340–03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы» регламентирует аспекты работы на ПК. Так же ГОСТ 12.2.032–78[37] ССБТ «Рабочее место при выполнении работ сидя» регламентирует общие требования к рабочему месту при выполнении работ сидя.

Эргономика рабочего места

Данные правила определяют санитарно – эпидемиологические требования к: проектированию, изготовлению и эксплуатации ПК, используемых на производстве; организации рабочих мест с ПК, производственным оборудованием. Согласно СанПиН 2.2.2/2.4.1340–03 Общие требования к организации рабочих мест пользователей ПК:

- При размещении рабочих мест с ПК расстояние между рабочими столами должно быть не менее 2,0 м, а расстояние между боковыми поверхностями видеомониторов не менее 1,2 м.
- Рабочие места с ПК в помещениях с источниками вредных производственных факторов должны размещаться с организованным воздухообменом.
- Рабочее место сотрудника, требующее значительного умственного напряжения или высокой концентрации внимания, рекомендуется изолировать руг от друга перегородками высотой 1,5–2,0 м.
- Конструкция рабочего кресла должна обеспечивать поддержание рациональной рабочей позы при работе на ПК позволять изменять позу с целью снижения напряжения мышц шейно – плечевой области и спины для предупреждения развития утомления.
- Площадь на одно рабочее место с компьютером для взрослых пользователей должна составлять не менее 6 м², а объем не менее 24 м³.
- Помещения с компьютерами должны оборудоваться системами отопления, кондиционирования воздуха или эффективной приточно-вытяжной вентиляцией.
- Для внутренней отделки интерьера помещений с компьютерами должны использоваться диффузно-отражающие материалы с коэффициентом отражения для потолка – 0,7–0,8; для стен – 0,5–0,6; для пола – 0,3–0,5.
- Поверхность пола в помещениях эксплуатации компьютеров должна быть ровной, без выбоин, нескользкой, удобной для очистки и влажной уборки, обладать антистатическими свойствами.
- В помещении должны находиться аптечка первой медицинской помощи, а также углекислотный огнетушитель для тушения пожара.

Вывод по разделу. При выполнении раздела социальная ответственность были установлены и исследованы на соответствие нормам вредные и опасные факторы. По результатам исследований были рассмотрены меры по противодействию данным факторам. Так же была рассмотрена экологическая

безопасность, где было установлено, как следует утилизировать отходы при выполнении работы. В разделе ЧС были рассмотрены возможные виды ЧС и меры их профилактики. И в последнем разделе были рассмотрены правовые и организационные моменты социальной ответственности. Рабочее место соответствует нормативным требованиям.

ЗАКЛЮЧЕНИЕ

В результате работы был создан программный комплекс, рассчитывающий коэффициенты корреляции и статистические математические функции произвольных рядов метеорологических данных. Данное программное обеспечение ориентировано на работу с метеоданными со станций многих городов. Был сделан обзор фреймворков для разработки программного обеспечения. Так же были исследованы методы описательной статистики и корреляционного анализа. Данные для базы данных были взяты из ВНИИГМИ-МЦД и загружаются в таблицу созданной базы данных. Программный комплекс был реализован с помощью пакета математической статистической обработки Apache Commons Math на языке Java. По результату реализации работы программного обеспечения были рассчитаны коэффициенты корреляции и построен график зависимости выпавших осадков от температуры окружающей среды. По итогам проведенного исследования, в котором сравниваются коэффициенты корреляции, видна хорошая согласованность расчетных значений коэффициентов корреляций для заданного параметра и временного периода.

СПИСОК ПУБЛИКАЦИЯ СТУДЕНТА

- 1) Маркова К.Д. Исследования корреляционных зависимостей метеорологических параметров // Материалы международной научно-практической конференции «Современное общество, образование и наука», г. Тамбов, 30 апреля 2018 г. – Вестник научных конференций. 2018 №4 – С. 68-70.
- 2) Маркова К.Д. Исследования корреляционных зависимостей метеорологических параметров // Материалы международной научно-технической конференции «Системы контроля окружающей среды – 2018», г. Севастополь, 5–9 ноября 2018 г.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Python // Python Software Foundation. URL: <https://www.python.org/> (дата обращения: 15.05.2018).
2. Mathematical statistics functions // Python. URL: <https://docs.python.org/3/library/statistics.html> (дата обращения: 15.05.2018)
3. NumPy // NumPy . URL: <http://www.numpy.org/> (дата обращения: 15.05.2018)
4. Scientific Computing Tools for Python // SciPy. URL: <https://www.scipy.org/> (дата обращения: 15.05.2018)
5. SciKits // SciPy. URL: <https://www.scipy.org/scikits.html> (дата обращения: 15.05.2018)
6. Statistics stats // Statistic in Python. URL: <https://www.statsmodels.org/stable/stats.html> (дата обращения: 15.05.2018)
7. Modular toolkit for Data Processing // Modular toolkit for Data Processing. URL: <http://mdp-toolkit.sourceforge.net/> (дата обращения: 15.05.2018)
8. Data Mining Fruitful and Fun // Orange. URL: <https://orange.biolab.si/> (дата обращения: 15.05.2018)
9. Why SymPy // SymPy. URL: <http://www.sympy.org/ru/> (дата обращения: 15.05.2018)
10. Getting Started // The R Project for Statistical Computing. URL: <https://www.r-project.org/> (дата обращения: 15.05.2018)
11. The Comprehensive R Archive Network // R Project for Statistical Computing. URL: <https://cran.r-project.org/> (дата обращения: 15.05.2018)
12. Statistical Functions using S4 Classes // R Project for Statistical Computing. <https://stat.ethz.ch/R-manual/R-devel/library/stats4/html/00Index.html> (дата обращения: 15.05.2018)
13. Stats-package // The R Stats Package URL: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html> (дата обращения: 15.05.2018)

14. Canonical Correlations // R: Canonical Correlations. URL: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/cancor.html> (дата обращения: 15.05.2018)
15. Correlation, Variance and Covariance // R: Correlation, Variance and Covariance. URL: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/cor.html> (дата обращения: 15.05.2018)
16. Hmisc: Harrell Miscellaneous // CRAN Package Hmisc. URL: <https://cran.r-project.org/web/packages/Hmisc/index.html> (дата обращения: 15.05.2018)
17. rcorr // RDocumentation. URL: <https://www.rdocumentation.org/packages/Hmisc/versions/4.1-1/topics/rcorr> (дата обращения: 15.05.2018)
18. sapply// RDocumentation. URL: <https://www.rdocumentation.org/packages/memisc/versions/0.99.14.9/topics/Sapply> (дата обращения: 15.05.2018)
19. pastecs: Package for Analysis of Space-Time Ecological Series // CRAN Package pastecs. URL: <https://cran.r-project.org/web/packages/pastecs/index.html> (дата обращения: 15.05.2018)
20. psych: Procedures for Psychological, Psychometric, and Personality Research // CRAN Package psych. URL: <https://cran.r-project.org/web/packages/psych/index.html> (дата обращения: 15.05.2018)
21. Что такое Java // Java. URL: <https://www.java.com/ru/> (дата обращения: 09.05.2018)
22. Colt // Colt URL: <http://dst.lbl.gov/ACSSoftware/colt/> (дата обращения: 09.05.2018)
23. Jama // Jama Network. URL: <https://jamanetwork.com/journals/jama> (дата обращения: 09.05.2018)
24. Commons Math: The Apache Commons Mathematics Library // Apache Commons. URL: <http://commons.apache.org/proper/commons-math/> (дата обращения: 09.05.2018)

25. Package org.apache.commons.math3.stat // API Apache Commons Math. URL: <http://commons.apache.org/proper/commons-math/javadocs/api-3.3/org/apache/commons/math3/stat/package-summary.html> (дата обращения: 10.05.2018)
26. Package org.apache.commons.math3.stat.descriptive // API Apache Commons Math. URL: <http://commons.apache.org/proper/commons-math/javadocs/api-3.3/org/apache/commons/math3/stat/descriptive/package-summary.html> (дата обращения: 10.05.2018)
27. Package org.apache.commons.math3.stat.correlation // API Apache Commons Math. URL: <http://commons.apache.org/proper/commons-math/javadocs/api-3.3/org/apache/commons/math3/stat/correlation/package-summary.html> (дата обращения: 10.05.2018)
28. Matlab // Matlab. URL: <https://matlab.ru/products/matlab> (дата обращения: 18.05.2018)
29. WOLFRAM COMPUTATION MEETS KNOWLEDGE // WOLFRAM MATHEMATICA. URL: <https://www.wolfram.com/mathematica/> (дата обращения: 18.05.2018)
30. Елисеева И. И., Юзбашев М. М. Общая теория статистики. М.: Финансы и Статистика, 2002. – 480 с.
31. Специализированные массивы // Всероссийский научно-исследовательский институт гидрометеорологической информации - Мировой центр данных. URL: <http://meteo.ru/data> (дата обращения: 15.05.2018)
32. ГОСТ 12.0.003-74. ССБТ. «Опасные и вредные производственные факторы. Классификация».
33. СанПиН 2.2.4.548–96. «Гигиенические требования к микроклимату производственных помещений».
34. СанПиН 2.2.2/2.4.1340–03. «Санитарно-эпидемиологические правила и нормативы «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы».

35. СанПиН 2.2.1/2.1.1.1278–03. «Гигиенические требования к естественному, искусственному и совмещённому освещению жилых и общественных зданий».

36. ГОСТ Р 50739–95. «Средства вычислительной техники. Защита от несанкционированного доступа к информации».

37. ГОСТ 12.2.032-78 ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования. – М. : Изд-во стандартов.