

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

---

Инженерная школа информационных технологий и робототехники  
Направление подготовки 09.04.01 Информатика и вычислительная техника  
Отделение информационных технологий

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

Тема работы
<b>Построение регрессионных моделей для прогнозирования метеорологических величин</b>

УДК 004.415:519.233:551.508

Студент

Группа	ФИО	Подпись	Дата
8ВМ6Б	Солтаганов Николай Андреевич		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Ботыгин Игорь Александрович	к.т.н.		

**КОНСУЛЬТАНТЫ:**

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Шаповалова Наталья Владимировна	-		

По разделу «Социальная ответственность»

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

**ДОПУСТИТЬ К ЗАЩИТЕ:**

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.04.01 Информатика и вычислительная техника	Спицын Владимир Григорьевич	д.т.н., профессор		

## ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ООП

Код результатов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
<b>Общепрофессиональные компетенции</b>		
Р1	Воспринимать и самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте.	Требования ФГОС 3+ (ОПК-1; ПК 3-6; ОК-4), критерий 5 АИОР (п. 1.1), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р2	Владеть и применять методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях.	Требования ФГОС 3+ (ОПК-5; ПК-7; ОК-7), критерий 5 АИОР (п. 1.1, 1.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р3	Демонстрировать культуру мышления, способность выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных из разных областей науки и техники, выносить суждения на основании неполных данных, анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями.	Требования ФГОС 3+ (ОПК-6; ПК-1,2; ОК-1,2), критерий 5 АИОР (п. 1.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р4	Анализировать и оценивать уровни своих компетенций в сочетании со способностью и готовностью к саморегулированию дальнейшего образования и профессиональной мобильности. Владеть, по крайней мере, одним из иностранных языков на уровне социального и профессионального общения, применять специальную лексику и профессиональную терминологию языка.	Требования ФГОС 3+ (ОПК-3,4; ПК-11,12; ОК-3), критерий 5 АИОР (п. 1.6, п. 2.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
<b>Профессиональные компетенции</b>		
Р5	Выполнять инновационные инженерные проекты по разработке аппаратных и программных средств автоматизированных систем различного назначения с использованием современных методов проектирования, систем автоматизированного проектирования, передового опыта разработки конкурентноспособных изделий.	Требования ФГОС 3+ (ПК-8–12; ОПК-2, ПК-7,6), критерий 5 АИОР (п. 1.3), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
Р6	Планировать и проводить теоретические и экспериментальные исследования в области проектирования аппаратных и программных средств автоматизированных систем с использованием новейших достижений науки и техники, передового отечественного и зарубежного опыта. Критически оценивать полученные данные и делать выводы.	Требования ФГОС 3+ (ПК-1–7; ОПК-6; ОК-4,9), критерий 5 АИОР (п.1.4), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.

Код результат ов	Результат обучения (выпускник должен быть готов)	Требования ФГОС ВО (ФГОС 3+), критерии АИОР, заинтересованных работодателей и студентов
P8	Использовать на практике умения и навыки в организации исследовательских, проектных работ и профессиональной эксплуатации современного оборудования и приборов, в управлении коллективом.	Требования ФГОС 3+ (ОК-5,8; ОПК-1,6; ПК-6,7,11,12), критерий 5 АИОР (п. 2.1, п. 2.3, п. 1.5), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P9	Осуществлять коммуникации в профессиональной среде и в обществе в целом, активно владеть иностранным языком, разрабатывать документацию, презентовать и защищать результаты инновационной инженерной деятельности, в том числе на иностранном языке.	Требования ФГОС 3+ (ОК-2,9; ОПК-4; ПК-1), критерий 5 АИОР (п. 2.2), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P10	Совершенствовать и развивать свой интеллектуальный и общекультурный уровень. Проявлять инициативу, в том числе в ситуациях риска, брать на себя всю полноту ответственности.	Требования ФГОС 3+ (ОК-1,6; ОПК-2; ПК-1,2), критерий 5 АИОР (п. 2.4, п. 2.5), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.
P11	Демонстрировать способность к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности, способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности, способность к педагогической деятельности.	Требования ФГОС 3+ (ОК-3,4,7; ОПК-3; ПК-7), критерий 5 АИОР (п. 2.6), соответствующий международным стандартам EUR-ACE и FEANI. Запросы студентов, отечественных и зарубежных работодателей.

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

---

Инженерная школа информационных технологий и робототехники  
Направление подготовки 09.04.01 Информатика и вычислительная техника  
Отделение информационных технологий

УТВЕРЖДАЮ:  
Руководитель ООП  
\_\_\_\_\_  
(Подпись)      \_\_\_\_\_ (Дата)      Спицын В.Г.  
(Ф.И.О.)

**ЗАДАНИЕ**  
**на выполнение выпускной квалификационной работы**

В форме:

Магистерской диссертации

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

Студенту:

Группа	ФИО
8ВМ6Б	Солтаганов Николай Андреевич

Тема работы:

<b>Построение регрессионных моделей для прогнозирования метеорологических величин</b>	
Утверждена приказом директора (дата, номер)	07.03.2018, №1548/с

Срок сдачи студентом выполненной работы:	08.06.2018
--	------------

**ТЕХНИЧЕСКОЕ ЗАДАНИЕ:**

<p><b>Исходные данные к работе</b></p> <p><i>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).</i></p>	<p>Язык программирования: Python 3.4 и комплект разработчика: Anaconda Scientific Python Distribution. Разработать метод прогнозирования на основании регрессионной модели и выполнить программную реализацию алгоритмов.</p>
---	---

<p><b>Перечень подлежащих исследованию, проектированию и разработке вопросов</b></p> <p><i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<p>Обзор и классификация моделей и методов прогнозирования временных рядов.</p> <p>Исследование и программные эксперименты по построению прогнозных моделей.</p> <p>Архитектура программного комплекса по построению регрессионных моделей для прогнозирования метеорологических величин.</p> <p>Технологическая схема развертывания и использование программного комплекса по построению регрессионных моделей для прогнозирования метеорологических величин.</p> <p>Обсуждение результатов выполненной работы.</p> <p>Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.</p> <p>Социальная ответственность.</p>
--	---

<p><b>Перечень графического материала</b></p> <p><i>(с точным указанием обязательных чертежей)</i></p>	<ol style="list-style-type: none"> <li>1. Схема взаимодействия модулей программного комплекса.</li> <li>2. Физическая структура файловой системы метеорологической базы данных</li> <li>3. Логическая структура взаимодействия в базе данных.</li> <li>4. Схемы работы алгоритмов.</li> </ol>
--	---

**Консультанты по разделам выпускной квалификационной работы**

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Старший преподаватель Шаповалова Н.В.
Социальная ответственность	Ассистент Авдеева И.И.
Раздел на иностранном языке	Старший преподаватель Куркан Н.В.

<p><b>Названия разделов, которые должны быть написаны на русском и иностранном языках:</b></p>
<p>Обзор и классификация моделей и методов прогнозирования временных рядов</p>

<p><b>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</b></p>	
--	--

**Задание выдал руководитель:**

Должность	ФИО	Ученая степень, звание	Подпись	Дата
-----------	-----	------------------------	---------	------

Доцент	Ботыгин И.А.	к.т.н.		
--------	--------------	--------	--	--

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ВМ6Б	Солтаганов Николай Андреевич		

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

---

Инженерная школа информационных технологий и робототехники  
Направление подготовки 09.04.01 Информатика и вычислительная техника  
Уровень образования Магистратура  
Отделение информационных технологий  
Период выполнения осенний / весенний семестр 2017/2018 учебного года

Форма представления работы:

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

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

Срок сдачи студентом выполненной работы:	14.06.2018 г.
--	---------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
	Основная часть	
	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	
	Социальная ответственность	
	Раздел на иностранном языке	

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Ботыгин Игорь Александрович	к.т.н.		

**СОГЛАСОВАНО:**

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.04.01 Информатика и вычислительная техника	Спицын Владимир Григорьевич	д.т.н., профессор		

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

Студенту:

<b>Группа</b>	<b>ФИО</b>
8ВМ6Б	Солтаганов Николай Андреевич

<b>Школа</b>	<b>ИШИТР</b>	<b>Отделение</b>	<b>ОИТ</b>
<b>Уровень образования</b>	Магистрант	<b>Направление/специальность</b>	09.04.01 «Информатика и вычислительная техника»

<b>Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»</b>	
1. <i>Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	Оклады участников проекта, нормы рабочего времени, ставки налоговых отчислений во внебюджетные фонды, районный коэффициент по г. Томску.
2. <i>Нормы и нормативы расходования ресурсов</i>	
3. <i>Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	
<b>Перечень вопросов, подлежащих исследованию, проектированию и разработке</b>	
1. <i>Оценка коммерческого и инновационного потенциала НТИ</i>	Анализ потенциальных потребителей результатов работы, проработка анализа конкурентных технических разработок
2. <i>Разработка устава научно-технического проекта</i>	Проработка целей и результатов исследования, определение участников разработки
3. <i>Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок</i>	Планирование этапов работ, построение графика проведения работ, расчет бюджета проведения работ
4. <i>Определение ресурсной, финансовой, экономической эффективности</i>	Оценка эффективности проекта
<b>Перечень графического материала</b>	
1. <i>Матрица SWOT.</i> 2. <i>Линейный график работ.</i> 3. <i>График проведения и бюджет НТИ.</i>	

<b>Дата выдачи задания для раздела по линейному графику</b>	
---	--

**Задание выдал консультант:**

<b>Должность</b>	<b>ФИО</b>	<b>Ученая степень, звание</b>	<b>Подпись</b>	<b>Дата</b>
Старший преподаватель ШИП	Шаповалова Наталья Владимировна	-		

**Задание принял к исполнению студент:**

<b>Группа</b>	<b>ФИО</b>	<b>Подпись</b>	<b>Дата</b>
8ВМ6Б	Солтаганов Николай Андреевич		





	<p>специальных мероприятий;</p> <ul style="list-style-type: none"> <li>• снижение вероятности возникновения короткого замыкания с помощью проведения специальных мероприятий;</li> <li>• снижение величин тока статического электричества.</li> </ul>
<b>2. Экологическая безопасность:</b>	<p>2. Полученный в результате выполнения магистерской диссертации, программный продукт не наносит вред окружающей среде. Нанести вред окружающей среде могут средства, которые используются при разработке и эксплуатации продукта, следовательно, для предотвращения такого вреда, рекомендуется организовать утилизацию отходов оргтехники, люминесцентных ламп и макулатуры.</p>
<b>3. Безопасность в чрезвычайных ситуациях:</b>	<p>3. Во время разработки программного продукта и его дальнейшего использования типичной ЧС является:</p> <p>3.1 Пожар.</p> <p>3.2 Действия и меры, направленные на предотвращение возникновения пожаров.</p>
<b>4. Правовые и организационные вопросы обеспечения безопасности:</b>	<p>4. Правовые и организационные вопросы для пользователя программного продукта:</p> <p>4.1 Правовые нормы ТК РФ для рабочей зоны пользователя с ПЭВМ</p> <p>4.2 Организация рабочего места для работы с ПЭВМ согласно СанПиН 2.2.2/2.4.1340-03.</p>

<b>Дата выдачи задания для раздела по линейному графику</b>	
---	--

**Задание выдал консультант:**

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

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
8ВМ6Б	Солтаганов Николай Андреевич		

## **Реферат**

Выпускная квалификационная работа содержит 131 с., 37 рис., 18 табл., 26 формул, 50 источников, 1 прил.

Ключевые слова: метеорологическая база данных, метеорологические данные, метеорологические станции, статистический анализ, алгоритмы статистического анализа, авторегрессионный анализ.

Цель работы – путем построения регрессионных моделей изучение возможности прогнозирования метеорологических величин.

Для достижения этой цели были поставлены и решены следующие задачи.

1. Осуществить обзор моделей и методов прогнозирования временных рядов с выявлением их достоинств и недостатков.
2. Исследовать модели прогнозирования временных рядов.
3. Разработать метод прогнозирования на основании предложенной модели и выполнить программную реализацию алгоритмов.
4. Оценить эффективность моделей прогнозирования при решении ими задачи прогнозирования временных рядов разных масштабов с последующим их сравнением.

Объектом исследования являются метеорологические величины.

Предметом исследования являются модели прогнозирования.

Методы исследования. При решении поставленных задач в работе использованы: регрессионный анализ, методы математического моделирования, анализ временных рядов, методы объектно-ориентированного программирования.

Научная новизна. В диссертации получены следующие основные результаты, которые выносятся на защиту.

1. Прогнозная модель сезонной авторегрессии временных рядов с шестью основными компонентами.
2. Прогнозная аддитивная модель авторегрессии с четырьмя основными компонентами.

Практическая ценность работы. Программная реализация алгоритмов может применяться для прогнозирования временных рядов различных предметных областей, не только для метеорологических величин.

Область применения: формирование отчетов по прогнозу временных рядов.

## **Определения**

**Метеорологические данные** — результаты метеорологических наблюдений на станциях государственной сети наблюдений и автоматических измерений, осуществляемых на железнодорожных станциях, узлах и перегонах.

**Статистический анализ** — является расчетом некоторых статистических показателей, который позволяет описать изучаемое явление, показать его динамику, взаимосвязь с другими явлениями, структуру, закономерности, сделать прогнозы на будущее.

**Модель прогнозирования** — модель объекта прогнозирования, исследование которой позволяет получить информацию о возможных состояниях объекта прогнозирования в будущем и (или) путях и сроках их осуществления.

**Прогноз** — обоснованное суждение о возможном состоянии объекта в будущем или альтернативных путях и сроках достижения этих состояний.

**Метод прогнозирования** — способ исследования объекта прогнозирования, направленный на разработку прогноза. Методы прогнозирования являются основанием для методик прогнозирования.

**Временный ряд** — собранный в разные моменты времени статистический материал о значении каких-либо параметров (в простейшем случае одного) исследуемого процесса.

## **Обозначения и сокращения**

**SARIMA** – сезонная интегрированная модель авторегрессии - скользящего среднего.

**Fbprophet** – библиотека Facebook Prophet.

**ПО** – программное обеспечение.

**БД** – База данных.

**ООП** – Объектно ориентированное программирование.

**ВНИИГМИ-МЦД** – Всероссийский научно исследовательский институт гидрометеорологической информации – мировой центр данных.

**MAPE** – средняя абсолютная ошибка в процентах.

**MAE** – средняя абсолютная ошибка.

**DF-тест** – теста Дики-Фуллера.

## Оглавление

Введение.....	18
1 Обзор и классификация моделей и методов прогнозирования временных рядов .....	20
1.1 Регрессионные модели .....	21
1.2 Авторегрессионные модели.....	24
1.3 Модели экспоненциального сглаживания.....	27
1.4 Нейросетевые модели.....	29
1.5 Модели на базе цепей Маркова.....	30
1.6 Модели на базе классификационно-регрессионных деревьев.....	30
1.7 Другие модели и методы прогнозирования .....	31
1.8 Комбинированные модели.....	32
1.9 Сравнение моделей прогнозирования .....	34
1.10 Вывод по разделу.....	35
2 Исследование и программные эксперименты по построению прогнозных моделей .....	37
2.1 Аддитивная регрессионная модель.....	37
2.2 Оценка качества аддитивной регрессионной модели .....	42
2.3 Сезонная интегрированная модель авторегрессии - скользящего среднего (SARIMA).....	43
2.4 Вывод по разделу.....	51
3 Архитектура программных модулей вычислительного комплекса по построению регрессионных моделей для прогнозирования метеорологических величин	52
3.1 Среда разработки .....	52
3.2 Метеорологические данные.....	52

3.3	Взаимодействие модулей программного комплекса .....	55
4	Программная реализация модулей комплекса по построению регрессионных моделей.....	60
4.1	Программная реализация модели Prophet .....	61
4.2	Программная реализация модели SARIMA .....	68
4.3	Вывод по разделу .....	76
5	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	78
5.1	Предпроектный анализ.....	78
5.1.1	Потенциальные потребители результатов исследования .....	78
5.2	SWOT-анализ .....	78
5.3	Инициация проекта.....	79
5.3.1	Организация и планирование работ.....	80
5.3.2	Продолжительность этапов работ.....	82
5.3.3	Расчет нарастания технической готовности работ.....	85
5.4	Расчет сметы затрат на разработку ПО и проведение эксперимента	87
5.4.1	Расчет затрат на материалы .....	87
5.4.2	Расчет основной заработной платы .....	87
5.4.3	Расчет отчислений от заработной платы.....	88
5.4.4	Расчет затрат на электроэнергию .....	89
5.4.5	Расчет прочих расходов .....	90
5.4.6	Расчет общей себестоимости разработки.....	90
5.5	Оценка экономической эффективности проекта.....	90
5.5.1	Оценка научно-технического уровня НИР .....	91
5.6	Вывод по разделу .....	95
6	Социальная ответственность .....	96



6.1	Введение .....	96
6.2	Производственная безопасность .....	96
6.2.1	Вредные производственные факторы.....	98
6.2.2	Опасные производственные факторы.....	104
6.2.3	Минимизация воздействия вредных и опасных факторов .	105
6.3	Экологическая безопасность .....	107
6.4	Безопасность в чрезвычайных ситуациях .....	109
6.4.1	Минимизация возможности возникновения пожара .....	109
6.5	Правовые и организационные вопросы обеспечения безопасности 111	
6.5.1	Правовые нормы трудового законодательства.....	111
6.5.2	Организация рабочего места при работе с ПЭВМ.....	114
6.6	Вывод по разделу .....	114
	Заключение .....	116
	Список источников .....	117
	<b>ПРИЛОЖЕНИЕ А</b> .....	123
	<b>РАЗДЕЛЫ НА ИНОСТРАННОМ ЯЗЫКЕ</b> .....	123
	1 Overview and classification of models and methods of time series forecasting... ..	124
	1.1 Regression models .....	125
	1.2 Autoregressive models.....	128
	1.3 Models of exponential smoothing.....	130

## **Введение**

Актуальность темы. Прогнозирование будущих значений временного ряда на основе его исторических значений является ядром для решения таких задач, как: планирования складского контроля, оптимизации и управления объемами производства, финансового планирования в торговле и экономике [1,2].

В развитых странах наблюдения за климатом и прогноз погоды давно стали категориями экономическими. В мире каждый год умирают около 250 000 человек из-за стихийных бедствий. Также наносится ущерб имуществу во всем мире, который оценивается в пределах 3100-6200 млрд. рублей. Согласно мировой статистике: если доверять гидрометеорологической информации и реагировать на нее адекватно, то можно полностью избежать жертв среди людей и предотвратить 30-40% ущерба имуществу [1,2].

Использование метеорологической информации в сельском хозяйстве, авиации, рыболовстве и судоходстве, энергетике, строительстве, также дает заметный экономический эффект.

С точки зрения науки предсказание погоды – одна из сложнейших задач физики атмосферы.

Для прогнозирования метеорологических явлений и их величин имеются различные методы, такие как, статистические методы, численные, синоптические, но пока в полном объеме ни один из методов не обеспечивает точного прогноза. Именно поэтому исследования в области прогнозирования метеорологических величин являются полезными и важными, а тема – актуальной.

В настоящее время различными организациями по наблюдению за погодой осуществляется накопление исторических значений метеорологических величин в базах данных, что значительно приумножает количество подающейся на входе информации для задач прогнозирования. Вместе с тем, развитие программных и аппаратных средств дает возможность

реализации сложных алгоритмов прогнозирования. Наряду с развитием технологического прогресса задача прогнозирования временных рядов усложняется. Кроме того, ввиду увеличения стихийных бедствий, к точности прогнозирования погоды предъявляют все более строгие требования.

В настоящее время неотъемлемой частью ежедневной работы многих организаций является задача обработки и анализа временных рядов с целью построить прогнозную модель, адекватно описывающую исследуемый процесс. В этом им способствует большое множество моделей прогнозирования. классы Авторегрессионные и нейросетевые модели являются наиболее популярными и широко используемыми [3]. Большое количество свободных параметров, идентификация которых неоднозначна и ресурсоемка, является существенным недостатком авторегрессионного класса [4]. Недоступность промежуточных вычислений, выполняющихся в «черном ящике», является существенным недостатком класса нейросетевых моделей. Поэтому сложно интерпретировать результаты моделирования. Кроме того, еще одним недостатком данного класса моделей является сложность выбора алгоритма обучения нейронной сети [5].

Диссертация посвящена сравнению и исследованию авторегрессионных моделей прогнозирования.

В первом разделе произведен обзор рассматриваемых классов моделей прогнозирования, установлены достоинства и недостатки каждого класса.

Во второй разделе диссертации описаны две модели и проведены программные эксперименты с последующим сравнением результатов экспериментов.

В третьем разделе описывается основная архитектура программных модулей вычислительного комплекса по построению регрессионных моделей для прогнозирования метеорологических величин.

В четвертом разделе диссертации описана программная реализация предложенных моделей для решения задач прогнозирования. Приведены скриншоты и способы работы с данным программным вычислительным комплексом.

## **1 Обзор и классификация моделей и методов прогнозирования временных рядов**

В работе [14] произведен анализ методов, где все методы прогнозирования были поделены на две группы: интуитивные и формализованные.

Интуитивное прогнозирование используется для объектов прогнозирования, которые слишком просты, либо, напротив, очень сложны. Под сложностью понимается наличие большого количества внешних факторов, которые невозможно учесть. Интуитивные методы прогнозирования отражают индивидуальные мнения экспертов относительно перспектив развития процесса и не предполагают разработку моделей прогнозирования. В основе интуитивных методов лежат интуиция и мобилизация профессионального опыта. Такие методы используются для анализа процессов, развитие которых либо полностью, либо частично не поддается математической формализации, то есть для которых трудно разработать адекватную модель. В работе [6] сказано, что к таким методам относятся методы экспертных оценок, исторических аналогий, предвидения по образцу. На сегодняшний день широко распространено применение экспертных систем, в том числе с использованием нечеткой логики [15].

Подробно описаны интуитивные методы прогнозирования в работе [16].

Модели прогнозирования разделяются на статистические модели и структурные модели [9].

В статистических моделях функциональная зависимость задана аналитически между будущими и фактическими значениями временного ряда, а также внешними факторами.

В структурных моделях функциональная зависимость задана структурно между фактическими и будущими значениями временного ряда, а также внешними факторами.

Статистическими моделями являются: модели экспоненциального сглаживания; авторегрессионные модели; регрессионные модели.

Структурными моделями являются: нейросетевые модели; модели на базе цепей Маркова; модели на базе классификационно-регрессионных деревьев.

Стоит упомянуть о том, что для узкоспециализированных задач иногда применяются специальные модели прогнозирования. К примеру, для задачи прогнозирования уровня сахара крови человека используются модели в основе, которых лежат дифференциальные уравнения [8]. Или к примеру, актуальная задача в последние несколько лет для мегаполисов – прогнозирование транспортного потока. В этой задаче используются гидродинамические модели [17]. Также, к примеру, для задачи прогнозирования природных явлений, таких как наводнение, используется, например, модель, в основу которой положены нелинейные клетки (или соты), находящиеся под воздействием внешнего поля, и у которых есть внутреннее состояние, изменяющееся во времени под воздействием этого поля [18]. Аналогичные модели разрабатываются и применяются для специальных процессов и систем.

## **1.1 Регрессионные модели**

Существует множество задач, где требуется исследовать отношения и найти связь между двумя и более переменными. Регрессионный анализ [19] применяется для решения таких задач. На сегодняшний день регрессия активно применяется в разных задачах, включая задачи управления и прогнозирования. Цель регрессионного анализа – определение зависимости между исходной переменной и масса внешних факторов (регрессоров). При этом допустимо определять коэффициенты регрессии по методу наименьших квадратов [19] или методу максимального правдоподобия [20].

Линейная регрессионная модель. Линейная регрессия – самый простой вариант регрессионной модели. Основа модели – гипотеза, что существует дискретный внешний фактор  $X(t)$ , который оказывает влияние на исследуемый

процесс  $Z(t)$ , при этом процесс и внешний фактор имеют между собой линейную связь. Такая модель прогнозирования описывается уравнением

$$t_{\text{ож}} = \frac{3t_{\text{min}}+2t_{\text{max}}}{5}Z(t) = \alpha_0 + \alpha_1 X(t) + \varepsilon_t \quad (1.1)$$

где  $\alpha_0$  и  $\alpha_1$  — коэффициенты регрессии;  $\varepsilon_t$  — ошибка модели. Для получения прогнозных значений  $Z(t)$  в момент времени  $t$  необходимо иметь значение  $X(t)$  в тот же момент времени  $t$ , что на практике является редкостью.

Множественная регрессионная модель. На практике на процесс  $Z(t)$  оказывают влияние целый ряд дискретных внешних факторов  $X_1(t), \dots, X_s(t)$ . В таком случае модель прогнозирования имеет вид

$$t_{\text{ож}} = \frac{3t_{\text{min}}+2t_{\text{max}}}{5}Z(t) = \alpha_0 + \alpha_1 X_1(t) + \alpha_2 X_2(t) + \dots + \alpha_s X_s(t) + \varepsilon_t \quad (1.2)$$

У данной модели есть существенный недостаток. Для вычисления будущего значения процесса  $Z(t)$  необходимо знать будущие значения всех факторов  $X_1(t), \dots, X_s(t)$ , что на практике является редкостью.

Основа нелинейной регрессионной модели — гипотеза о том, что существует известная функция, описывающая зависимость между исходным процессом  $Z(t)$  и внешним фактором  $X(t)$

$$t_{\text{ож}} = \frac{3t_{\text{min}}+2t_{\text{max}}}{5}Z(t) = F(X(t), A) \quad (1.3)$$

В рамках построения модели прогнозирования необходимо определить параметры функции  $A$ . Например, можно предположить, что

$$t_{\text{ож}} = \frac{3t_{\text{min}}+2t_{\text{max}}}{5}Z(t) = \alpha_1 \cos(X(t)) + \alpha_0 \quad (1.4)$$

Чтобы построить модель достаточно определить параметры  $A = [\alpha_1, \alpha_0]$ . Большая редкость встретить на практике процессы, для которых вид функциональной зависимости между процессом  $Z(t)$  и внешним фактором  $X(t)$  заранее известен. Поэтому нелинейные регрессионные модели применяются редко.

Модель группового учета аргументов (МГУА) была разработана Ивахтенко А.Г. [21]. МГУА принимает вид

$$Z(t) = \alpha_0 + \sum_{i=1}^s a_i X_i(t) + \sum_{i=1}^s \sum_{j=1}^s a_{i,j} X_i(t) X_j(t) + \sum_{i=1}^s \sum_{j=1}^s \sum_{k=1}^s a_{i,j,k} X_i(t) X_j(t) X_k(t) + \dots \quad (1.5)$$

Уравнение (1.5) называется опорной функцией. Для построения различных вариантов моделей для некоторых или всех аргументов используют опорную функцию. Например, строятся полиномы с одной переменной, полиномы со всевозможными парами переменных, полиномы со всевозможными тройками переменных и т. д. Методом регрессионного анализа для каждой модели узнаются её линейные коэффициенты  $a_{i,j,k}, \dots$ . Происходит выбор среди всех моделей несколько наилучших, в пределах от 2 до 10. При этом, например, качество моделей определяется среднеквадратичным отклонением или иным критерием. Процесс перебора моделей прекращается, если среди выбранных имеется модель, качество которой достаточно для использования полученных прогнозных значений. В противном случае отобранные модели используются в качестве аргументов  $X_1(t) \dots X_s(t)$  для опорных функций следующего этапа итерации. Иначе говоря, уже найденные модели принимают участие в формировании более сложных.

Важными достоинствами данного класса моделей являются их простота, гибкость, а также единообразие их анализа и проектирования [19]. При использовании линейных регрессионных моделей результат прогнозирования может быть получен значительно быстрее, чем при использовании остальных моделей. А также, достоинством является прозрачность моделирования [5], т. е. доступность для анализа всех промежуточных вычислений. К основному недостатку нелинейных регрессионных моделей нужно отнести их сложность определения вида функциональной зависимости [14], а также сложность и

трудоемкость определение параметров модели. Кроме того, недостатками линейных регрессионных моделей являются низкая адаптивность и отсутствие способности моделирования нелинейных процессов [28].

## 1.2 Авторегрессионные модели

Основа авторегрессионных моделей – гипотеза о том, что значение процесса  $Z(t)$  линейно зависит от некоторого количества предыдущих значений того же процесса  $Z(t - 1), \dots, Z(t - p)$ .

Авторегрессионная модель скользящего среднего. Одной из наиболее используемых в области анализа временных рядов является модель авторегрессии (autoregressive, AR) и модель скользящего среднего (moving average, MA) [1,5].

Согласно работе [1], модель авторегрессии является исключительно полезной для описания некоторых встречающихся на практике временных рядов. В этой модели текущее значение процесса выражается как конечная линейная совокупность предыдущих значений процесса и импульса, который называется «белым шумом»,

$$t_{ож} = \frac{3t_{min}+2t_{max}}{5}Z(t) = C + \varphi_1 Z(t - 1) + \varphi_2 Z(t - 2) + \dots + \varphi_p Z(t - p) + \varepsilon_t \quad (1.6)$$

Формула (1.6) описывает процесс авторегрессии порядка  $p$ , который в литературе часто обозначается  $AR(p)$ , здесь  $C$  — вещественная константа,  $\varphi_1, \dots, \varphi_p$  — коэффициенты,  $\varepsilon_t$  — ошибка модели. Для определения  $\varphi_i$  и  $C$  используют метод наименьших квадратов [19] или метод максимального правдоподобия [20].

Здесь же стоит упомянуть о модели скользящего среднего порядка  $q$ , которая имеет большое значение в описании временных рядов и часто используется совместно с авторегрессией. Она описывается уравнением

$$t_{ож} = \frac{3t_{min}+2t_{max}}{5}Z(t) = \frac{1}{q}(Z(t - 1) + Z(t - 2) + \dots + Z(t - q)) + \varepsilon_t \quad (1.7)$$



$$q) + \varepsilon_t)$$

В литературе процесс (1.7) часто обозначается  $MA(q)$ ; здесь  $q$  — порядок скользящего среднего,  $\varepsilon_t$  — ошибка прогнозирования. По сути дела, модель скользящего среднего является фильтром низких частот. Нужно отметить, что существуют простые, взвешенные, кумулятивные, экспоненциальные модели скользящего среднего.

Согласно работе [1], часто целесообразно объединить в одной модели авторегрессию и скользящее среднее для достижения большей гибкости в подгонке модели. Общая модель обозначается  $ARMA(p,q)$ . Она соединяет в себе авторегрессию фильтрованных значений процесса порядка  $p$  и фильтр в виде скользящего среднего порядка  $q$ . Если в качестве входных данных используются не сами значения временного ряда, а их разность  $d$ -того порядка (на практике  $d$  необходимо определять, однако в большинстве случаев  $d \leq 2$ ), то модель носит название авторегрессии проинтегрированного скользящего среднего. В литературе данная модель известна под названием  $ARIMA(p,d,q)$  (autoregression integrated moving average).

Развитием модели  $ARIMA(p,d,q)$  является модель  $ARIMAX(p,d,q)$ , которая описывается уравнением [1]

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} Z(t) = AR(p) + a_1 X_1(t) + \dots + a_s X_s(t) \quad (1.8)$$

Здесь  $a_1, \dots, a_s$  — коэффициенты внешних факторов  $X_1(t), \dots, X_s(t)$ . В данной модели чаще всего процесс  $Z(t)$  является результатом модели  $MA(q)$ , то есть отфильтрованными значениями исходного процесса. Далее для прогнозирования  $Z(t)$  используется модель авторегрессии, в которой введены дополнительные регрессоры внешних факторов  $X_1(t), \dots, X_s(t)$ .

Авторегрессионная модель с условной гетероскедастичностью (autoregressive conditional heteroskedasticity, GARCH) была разработана в 1986 году Тимом Петером Борреслевым и является моделью остатков для модели  $AR(p)$  [22]. На первом этапе для исходного временного ряда определяется

модель AR(p) (1.10). Далее предполагается, что ошибка модели (1.10)  $\varepsilon_t$  имеет две составляющие

$$\varepsilon_t = \sigma_t \cdot \vartheta_t \quad (1.9)$$

где  $\sigma_t$  — зависимое от времени стандартное отклонение;  $\vartheta_t$  — случайная величина, имеющая нормальное распределение, среднее значение, равное 0, и стандартное отклонение, равное 1. При этом зависимое от времени стандартное отклонение описывается уравнением

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} \sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \dots + \beta_q \varepsilon_{t-q}^2 + \gamma_1 \sigma_{t-1} + \dots + \gamma_p \sigma_{t-p} \quad (1.10)$$

Здесь  $\beta_0, \dots, \beta_q$  и  $\gamma_0, \dots, \gamma_p$  — коэффициенты. Уравнение (1.10) называется моделью GARCH(p,q) и имеет два параметра: p характеризует порядок авторегрессии квадратов остатков; q — количество предшествующих оценок остатков. Очень часто данная модель используется в финансовом секторе, где с помощью нее моделируется волатильность. Существует ряд модификаций модели под названиями NGARCH, IGARCH, EGARCH, GARCH-M и другие [22].

В статье [3], в частности, утверждается: «Без сомнений модели ARIMA и GARCH являются самыми популярными для прогнозирования временных рядов. На сегодняшний день главную конкуренцию данным моделям составляют модели на основе ANN».

Авторегрессионная модель с распределенным лагом (autoregressive distributed lag models, ARDLM) достаточно коротко описана в литературе. Чаще всего она упоминается в книгах по эконометрике [23]. Часто при моделировании процессов на изучаемую переменную влияют не только текущие значения процесса, но и его лаги, то есть значения временного ряда, предшествующие изучаемому моменту времени. Такая модель описывается уравнением

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} Z(t) = \varphi_0 + \varphi_1 Z(t-l-1) + \dots + \varphi_p Z(t-l-p) \quad (1.11)$$

$$l - p) + \varepsilon_t$$

Здесь  $\varphi_0, \dots, \varphi_p$  — коэффициенты,  $l$  — величина лага. Модель (1.11) называется ARDLM(p,l) и чаще всего применяется для моделирования экономических процессов [23].

К достоинствам данных моделей относят простоту и прозрачность моделирования, единообразие анализа и проектирования [1]. Данный класс моделей является одним из наиболее популярных на нынешний день [3], поэтому в открытом доступе легко найти примеры решения задач прогнозирования временных рядов различных предметных областей с применением авторегрессионных моделей. Недостатками данного класса моделей являются: большое число параметров модели, идентификация которых неоднозначна и ресурсоемка [4]; низкая адаптивность моделей, а также линейность и, как следствие, отсутствие способности моделирования нелинейных процессов, часто встречающихся на практике [26].

### 1.3 Модели экспоненциального сглаживания

Модели экспоненциального сглаживания (exponential smoothing, ES) используются для моделирования экономических и финансовых процессов [24]. В основа экспоненциального сглаживания лежит идея постоянного пересмотра прогнозных значений по мере поступления фактических. По мере старения наблюдений модель ES присваивает им экспоненциально убывающие веса. А это значит, что последние доступные наблюдения имеют большее влияние на прогнозное значение, чем старшие наблюдения.

Функция модели ES принимает вид

$$\begin{aligned} Z(t) &= S(t) + \varepsilon_t \\ S(t) &= \alpha * Z(t - 1) + (1 - \alpha) * S(t - 1) \end{aligned} \tag{1.12}$$

где  $\alpha$  — коэффициент сглаживания,  $0 < \alpha < 1$  ; начальные условия определяются как  $S(1) = Z(0)$ . В данной модели каждое последующее сглаженное

значение  $S(t)$  является взвешенным средним между предыдущим значением временного ряда  $Z(t)$  и предыдущего сглаженного значения  $S(t - 1)$ .

Для моделирования процессов, имеющих тренд, используется модель Хольта или двойное экспоненциальное сглаживание. В таком случае в модели необходимо рассматривать две составляющие: тренд и уровень [24]. Уровень и тренд сглаживаются отдельно

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} Z(t) = S(t) + \varepsilon_t$$

$$S(t) = \alpha * Z(t - 1) + (1 - \alpha) * (S(t - 1) - \beta(t - 1)) \quad (1.13)$$

$$S(t) = \gamma * (S(t - 1) - S(t - 2)) - (1 + \gamma) * \beta(t - 1)$$

Здесь  $\alpha$  — коэффициент сглаживания уровня, как и в модели (1.16),  $\gamma$  — коэффициент сглаживания тренда.

Для моделирования процессов, имеющих тренд и сезонную составляющую, используется модель Хольта-Винтерса или тройное экспоненциальное сглаживание.

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} Z(t) = (R(t) + G(t)) * S(t) \quad (1.14)$$

Здесь  $R(t)$  — сглаженный уровень без учета сезонной составляющей

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} R(t) = \frac{\alpha * Z(t-1)}{S(t-L)} + (1 + \alpha) * (R(t - 1) + G(t - 1)) \quad (1.15)$$

$G(t)$  — сглаженный тренд

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} G(t) = \beta * (S(t - 1) - S(t - 2)) + (1 - \beta) * G(t - 1) \quad (1.16)$$

а  $S(t)$  — сезонная составляющая

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5}$$

$$S(t) = \frac{\gamma * Z(t - 1)}{S(t - L)} + (1 - \gamma) * S(t - L) \quad (1.17)$$

Величина  $L$  – длина сезона исследуемого процесса. Для долгосрочного прогнозирования данный класс моделей используется чаще других [24].

К достоинствам данного класса моделей относят единообразие и простоту их анализа и проектирования. Недостатком данного класса моделей является отсутствие гибкости [35].

#### 1.4 Нейросетевые модели

Среди структурных моделей самая популярная на сегодняшний день – модель на основе искусственных нейронных сетей (artificial neural network, ANN) [5]. Нейронные сети состоят из нейронов.

Для описания модели нейрона необходима пара уравнений

$$U(t) = \sum_{i=1}^m \omega_i * Z(t - i) + b \quad (1.18)$$
$$Z(t) = \varphi(U(t))$$

где  $\omega_1, \dots, \omega_m$  — синаптические веса нейрона;  $Z(t - i), \dots, Z(t - m)$  — входные сигналы;  $(U(t))$  — функция активации;  $b$  — порог. Функция активации бывает трех основных типов (сигмоидальная функция, функция единичного скачка, кусочно-линейная функция) [25].

В зависимости от способа связи нейронов (архитектуру такой сети определяет способ связи нейронов), сети бывают трех типов: рекуррентные сети, однослойные или многослойные сети прямого распространения [25].

Основным достоинством нейросетевых моделей является нелинейность, т. е. они могут устанавливать нелинейные зависимости между будущими и фактическими значениями процессов. Другими важными достоинствами являются: адаптивность, масштабируемость (параллельная структура ANN ускоряет вычисления) и единообразие их анализа и проектирования [25].

Недостатки ANN – сложность выбора архитектуры; отсутствие прозрачности моделирования; высокие требования к непротиворечивости обучающей выборки; сложность выбора алгоритма обучения и ресурсоемкость процесса их обучения [5].

## 1.5 Модели на базе цепей Маркова

В основе модели прогнозирования на базе цепей Маркова (Markov chain model) лежит гипотеза, что будущее состояние процесса зависит только от его текущего состояния и не зависит от предыдущих [29]. В связи с этим процессы, моделируемые цепями Маркова, должны относиться к процессам с короткой памятью.

При построении цепи Маркова определяется масса состояний и вероятности переходов. Если текущее состояние процесса  $S_i$ , то в качестве будущего состояния процесса выбирается такое состояние  $S_j$ , вероятность перехода в которое (значение  $\lambda_{ij}$ ) максимальна.

Это значит, что структура цепи Маркова и вероятности перехода состояний определяют зависимость между будущим значением процесса и его текущим значением.

Достоинства моделей – единообразие и простота анализа и проектирования. Недостаток данных моделей – неосуществимость моделирования процессов с длинной памятью [29].

## 1.6 Модели на базе классификационно-регрессионных деревьев

Еще одной популярной структурной моделью прогнозирования временных рядов являются классификационно-регрессионные деревья (classification and regression trees, CART) [30]. Она изначально была создана для моделирования процессов, на которые оказывают влияние, как непрерывные внешние факторы, так и категориальные. Если внешние факторы, влияющие на процесс  $Z(t)$ , непрерывны, то используются регрессионные деревья; если факторы категориальные, то — классификационные деревья. Существуют также смешанные CART, их используют, если при моделировании необходимо учитывать факторы обоих типов.

CART моделирует зависимость будущей величины процесса  $Z(t)$  при помощи структуры дерева, а также пороговых констант и подмножеств.

Достоинствами данного класса моделей являются: быстрота и однозначность процесса обучения дерева (в отличие от ANN) [9]; масштабируемость, за счет которой возможна быстрая обработка сверхбольших объемов данных, а также возможность использовать категориальные внешние факторы. Недостатки данных моделей – неоднозначность алгоритма построения структуры дерева; отсутствие единообразия их анализа и проектирования; сложность вопроса о том, когда стоит прекратить дальнейшие ветвления [30].

### **1.7 Другие модели и методы прогнозирования**

Помимо классов моделей прогнозирования, рассмотренных выше, существуют другие модели и методы прогнозирования, но менее распространенные. Главный недостаток моделей и методов – недостаточная методологическая база, т. е. отсутствие полной документации и недостаточно подробное описание подходов определения параметров моделей и их возможностей. Также, в открытом доступе присутствует лишь небольшое количество статей, посвященных работе с данными методами.

Метод опорных векторов (support vector machine, SVM) часто применяется для прогнозирования движения цен на электроэнергию [5] и рынков [31]. Основа метода – классифицирование, которое происходит за счет перевода исходных временных рядов в пространство более высокой размерности и поиска разделяющей гиперплоскости с максимальным зазором в этом пространстве. SVM применяет алгоритм, который использует гипотезу о том, что чем больше расстояние или разница между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора [32]. В результате чего на этапе обучения классификатора выявляются независимые переменные (внешние факторы), будущие значения которых определяют в какой из определенных ранее подклассов попадет прогноз  $Z(t)$ .

Генетический алгоритм (genetic algorithm, GA) был разработан для решения задач оптимизации, а также поисковых задач. Однако некоторые

модификации GA позволяют решать задачи прогнозирования. В статье [33] указано, что алгоритм прогнозирования на основе GA позволяет принимать в расчет более 15 внешних факторов, используя базовый GA. Принцип работы – исходные значения внешних факторов  $X_1(t), \dots, X_5(t)$  и процесса  $Z(t)$  раскладывают в наборы (состоящие из 0 и 1), которые называют генотипами. Далее применяют ряд преобразований: мутирование и скрещивание для формирования преобразованных наборов, которые называются фенотипами. С использованием функции приспособленности исследуются полученные и исходные наборы. Скрещивание и мутирование будет происходить пока результат не будет удовлетворительным.

Модель на основе передаточных функций (transfer function, TF) применяется для прогнозирования процесса  $Z(t)$  с учетом внешнего фактора  $X(t)$  [34]. Уравнение, отражающее зависимость будущего значения принимает вид

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} Z(t) = v(B)X(t) + \eta(t) \quad (1.19)$$

где  $B$  — оператор сдвига  $BZ(t) = Z(t - 1), \dots, B^k Z(t) = Z(t - k)$ . Временной ряд  $\eta(t)$  характеризует внешнее возмущение. При этом функция  $v(B)$  принимает вид

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} Z(t) = \alpha_0 + \alpha_1 X(t) + \varepsilon_t \quad (1.20)$$

$$t_{ож} = \frac{3t_{min} + 2t_{max}}{5} v(B) = v_0 + v_1 B + v_2 B^k + \dots \quad (1.21)$$

Коэффициенты функции (1.21)  $v_i$  описывают динамические отношения между процессами  $Z(t)$  и  $X(t)$ .

## 1.8 Комбинированные модели

Сотворение комбинированных моделей и методов в области создания моделей прогнозирования – одной из популярных современных тенденций. Подобный подход дает возможность компенсировать недостатки одних



моделей при помощи других и направлен на повышение точности прогнозирования, как одного из главных критериев эффективности модели. Статья [37] является одной из первых работ в этой области. В ней говорится о подходе, в котором временного ряд прогнозируется в два этапа. Первый этап основывается на моделях распознавания образов, где выделяются гомогенные группы временного ряда. После чего для каждой гомогенной группы строится отдельная модель прогнозирования. В статье утверждается, что при комбинированном подходе удастся увеличить точность прогнозирования временных рядов.

При помощи вейвлет преобразования доступные значения временного ряда разделяются на несколько последовательностей, для каждой из которых строится отдельная модель ARIMA. Подобный подход использовался для прогнозирования цен на электроэнергию в Евросоюзе [13]

Различные типы комбинаций (нечеткая логика + ANN, нечеткая логика + GA + ANN, SARIMA + ANN, регрессия + ANN, нечеткая логика + регрессия) рассматривается в обзоре моделей прогнозирования энергопотребления [35].

Модели на основе ANN в большинстве комбинаций используются для решения задачи кластеризации, а далее для каждого кластера строится отдельная модель прогнозирования на основе SARIMA, GA, нечеткой логики и других. В работе [35] автором утверждается, что применение комбинированных моделей, является наиболее перспективным направлением развития моделей прогнозирования. Модели выполняют предварительную кластеризации и последующее прогнозирование внутри определенного кластера.

Авторы работы [38] заявляют, что для повышения точности прогнозирования необходимо применять кластеризацию. Для кластеризации предлагается метод К-средних и метод нечетких С-средних. Получение полезной информации из временного ряда для последующего прогнозирования, является целью обоих методов.

При корректном подходе применение комбинированных моделей дает повышение точности прогнозирования. Ресурсоемкость и сложность их разработки – это главный недостаток комбинированных моделей.

Некоторые исследователи пошли по другому пути и разработали модели авторегрессии. Модели основываются на предположении – временной ряд состоит из patterns (последовательность повторяющихся кластеров). Но исследователи не стали создавать комбинированные модели, так как определили кластеры, тем самым выполнили прогноз на основании одной модели.

## 1.9 Сравнение моделей прогнозирования

Для моделей и методов достоинства и недостатки систематизированы в таблицу 1.1.

Таблица 1.1 - Сводная таблица достоинств и недостатков

Модель и метод	Достоинства	Недостатки
Модели и методы авторегрессии	масса примеров применения; прозрачность моделирования, простота; единообразие проектирования и анализа	низкая адаптивность; ресурсоемкость идентификации моделей и трудоемкость; невозможность моделирования нелинейностей
Модели и методы регрессии	прозрачность моделирования, гибкость, простота; единообразие проектирования и анализа	невозможность для нелинейной регрессии моделирования нелинейных процессов; сложность определения зависимости функциональной; трудоемкий процесс

Модель и метод	Достоинства	Недостатки
		нахождения коэффициентов зависимости
Модели и методы экспоненциального сглаживания	единообразия проектирования и анализа; простота моделирования	узкая применимость моделей; недостаточная гибкость
Нейросетевые модели и методы	единообразия проектирования и анализа; высокая адаптивность, масштабируемость; нелинейность моделей; масса примеров применения	сложность выбора алгоритма обучения; сложность выбора архитектуры; отсутствие прозрачности; ресурсоемкость процесса обучения; строгие требования к обучающей выборке
Модели и методы на базе классификационно регрессионных деревьев	простота и быстрота процесса обучения; масштабируемость; возможность учитывать категориальные переменные	сложность вопроса останова; неоднозначность алгоритма построения дерева
Модели и методы на базе цепей Маркова	простота моделирования; единообразия анализа и проектирования;	узкая применимость моделей; неосуществимость моделирования процессов с длинной памятью;

### 1.10 Вывод по разделу

Был выполнен обзор моделей и методов прогнозирования временных рядов. Приведены достоинства и недостатки каждого из рассматриваемого

метода прогнозирования. Отмечено что на сегодняшний день наиболее распространенными моделями прогнозирования являются нейросетевые модели (ANN), а также авторегрессионные модели (ARIMA).

## 5 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

### 5.1 Предпроектный анализ

#### 5.1.1 Потенциальные потребители результатов исследования

Для анализа потребителей результатов необходимо рассмотреть целевой рынок и провести его сегментирование.

Результатом выполнения магистерской диссертации является программный продукт, служащий инструментом для научно-исследовательских вычислений с последующим анализом полученных данных. Данный программный продукт применяется в области метеорологии и климатологии.

Целевым рынком, на котором будет реализовываться в будущем разработка, являются коммерческие организации, научно-исследовательские институты, гидрометцентры предоставляющие услуги по прогнозированию погоды. Потенциальными потребителями являются аэропорты, городские, региональные и федеральные службы.

### 5.2 SWOT-анализ

SWOT – Strengths (сильные стороны), Weaknesses (слабые стороны), Opportunities (возможности) и Threats (угрозы) – представляет собой комплексный анализ научно-исследовательского проекта.

Эта часть раздела посвящена описанию сильных и слабых сторон проекта, в выявлении возможностей и угроз для реализации проекта, которые проявились или могут появиться в его внешней среде. Слабые и сильные стороны, возможности и угрозы представлены в табличном виде (таблица 5.1).

Таблица 5.1 – Матрица SWOT-анализ

	Сильные стороны научно-исследовательского проекта: С1. Повышение производительности и эффективности работы	Слабые стороны научно-исследовательского проекта: Сл1. Необходимость дополнительного обучения кадров для работы с программным продуктом. Сл2. Большой набор
--	---	---

	пользователей. С2. Оптимизация различных задач. С3. Заявленная экономичность и энергоэффективность технологии.	требований к продукту со стороны заказчика. Сл3. Отсутствие необходимого оборудования для проведения испытания опытного образца.
Возможности: В1. Появление новых сфер применения программного продукта. В2. Интерес ведущих научно-исследовательских институтов. В3. Повышение стоимости конкурентных разработок.	Новые сферы применения увеличат финансирование разработки. Интерес предприятий поможет понять конечные требования к продукту.	Новые сферы применения технологии помогут найти область с легко выполнимыми требованиями.
Угрозы: У1. Отсутствие спроса на разрабатываемую систему. У2. Развитие альтернативных систем и технологий.	Повышенная производительность работы системы, за счет оптимизации решения различных задач и новых технологий могут стать важным аргументом для перехода к новой системе.	Главной угрозой является отсутствие спроса, что решается поиском новых сфер применения и демонстрацией достоинств программного продукта.

Исходя из результатов SWOT-анализа, можно сделать вывод о том, что при разработке системы большое внимание должно уделяться задаче оптимизации и повышения производительности работы, поскольку благодаря таким сильным сторонам системы обеспечивается защита от угроз, а также открываются возможности по расширению такой системы.

### 5.3 Инициация проекта

Группа процессов инициации состоит из процессов, которые выполняются для определения нового проекта или новой фазы существующего. В рамках процессов инициации определяются изначальные цели и содержание и фиксируются изначальные финансовые ресурсы. Определяются внутренние и внешние заинтересованные стороны проекта, которые будут взаимодействовать

и влиять на общий результат научного проекта. Данная информация закрепляется в уставе проекта.

Ниже, в таблицах 5.2-5.3 представлены необходимые данные, которые входят в устав проекта.

Таблица 5.2 – Заинтересованные стороны проекта

Заинтересованные стороны проекта	Ожидания заинтересованных сторон
ФГБУН «Томский научный центр Сибирского отделения Российской академии наук»	Разработка программного продукта для прогнозирования метеорологических величин.

Таблица 5.3 – Цели и результат проекта

Цели проекта:	Разработка программного продукта для построения регрессионных моделей для прогнозирования метеорологических величин.
Ожидаемые результаты проекта:	<ul style="list-style-type: none"> <li>• Улучшение эффективности системы прогнозирования метеорологических величин.</li> <li>• Оптимизация работы пользователей с программным продуктом.</li> <li>• Повышение производительности и эффективности работы.</li> </ul>
Критерии приемки результата проекта:	Разработка окончательного функционала для нормальной работы пользователей.
Требования к результату проекта:	Удобный интерфейс и функционал для пользователей продуктом. Автоматизация всего процесса организации.

### 5.3.1 Организация и планирование работ

При организации процесса реализации конкретного проекта необходимо оптимально планировать занятость каждого из его участников и сроки проведения отдельных работ.

На данном этапе составляется полный перечень проводимых работ, и определяются их исполнители и оптимальная продолжительность. Результатом планирования работ является сетевой, либо реализации проекта. Наиболее удобным, простым и наглядным способом для этих целей является

использование линейного графика. Для его построения составим перечень работ и соответствие работ своим исполнителям, продолжительность выполнения этих работ и сведем их в таблицу 5.4.



Таблица 5.4 - Перечень работ и продолжительность их выполнения

Этапы работы	Исполнители	Загрузка исполнителей
Постановка целей и задач, получение исходных данных	НР, Ст	НР – 100% Ст – 30%
Составление и утверждение ТЗ	НР, Ст	НР – 100% Ст – 10%
Подбор и изучение материалов по тематике	НР, Ст	НР – 30% Ст – 100%
Разработка календарного плана	НР, Ст	НР – 30% Ст – 100%
Обсуждение литературы	НР, Ст	НР – 20% Ст – 100%
Выбор алгоритмов для решения поставленной задачи	НР, Ст	НР – 20% Ст – 100%
Планирование архитектуры программного обеспечения	НР, Ст	НР – 10% Ст – 100%
Разработка программного обеспечения	Ст	Ст – 100%
Тестирование и отладка программного обеспечения	Ст	Ст – 100%
Оформление расчетно-пояснительной записки	Ст	Ст – 100%
Подведение итогов	НР, Ст	НР – 70% Ст – 100%

НР – научный руководитель.

Ст – студент

### 5.3.2 Продолжительность этапов работ

Расчет продолжительности этапов работ осуществляется двумя методами:

- технико-экономическим;
- опытно-статистическим.

В данном случае используется опытно-статистический метод, который реализуется двумя способами:

- аналоговый;
- экспертный (вероятностный).

Для определения ожидаемого значения продолжительности работ  $t_{ож}$  применяется вероятностный метод – метод двух оценок  $t_{min}$  и  $t_{max}$ .

$$t_{\text{ож}} = \frac{3t_{\text{min}} + 2t_{\text{max}}}{5}, \quad (5.1)$$

где  $t_{\text{min}}$  – минимальная трудоемкость работ, чел/дн.;

$t_{\text{max}}$  – максимальная трудоемкость работ, чел/дн.

Для выполнения перечисленных в таблице 1 работ требуются специалисты:

- инженер – в его роли действует исполнитель НИР (ВКР);
- научный руководитель.

Для построения линейного графика необходимо рассчитать длительность этапов в рабочих днях, а затем перевести в календарные дни.

Расчет продолжительности выполнения каждого этапа ведется по формуле:

$$T_{\text{РД}} = \frac{t_{\text{ож}}}{K_{\text{ВН}}} \cdot K_{\text{Д}}, \quad (5.2)$$

где  $t_{\text{ож}}$  – трудоемкость работы, чел/дн.;

$K_{\text{ВН}}$  – коэффициент выполнения работ ( $K_{\text{ВН}} = 1$ );

$K_{\text{Д}}$  – коэффициент, учитывающий дополнительное время на компенсации и согласование работ ( $K_{\text{Д}} = 1.2$ ).

Расчет продолжительности этапа в календарных днях ведется по формуле:

$$T_{\text{КД}} = T_{\text{РД}} \cdot T_{\text{К}}, \quad (5.3)$$

где  $T_{\text{РД}}$  – продолжительность выполнения этапа в рабочих днях;

$T_{\text{КД}}$  – продолжительность выполнения этапа в календарных днях;

$T_{\text{К}}$  – коэффициент календарности.

Коэффициент календарности рассчитывается по формуле:

$$T_{\text{К}} = \frac{T_{\text{КАЛ}}}{T_{\text{КАЛ}} - T_{\text{ВД}} - T_{\text{ПД}}}, \quad (5.4)$$

где  $T_{\text{КАЛ}}$  – календарные дни ( $T_{\text{КАЛ}} = 365$ );

$T_{\text{ВД}}$  – выходные дни ( $T_{\text{ВД}} = 52$ );

$T_{\text{ПД}}$  – праздничные дни ( $T_{\text{ПД}} = 10$ ).

$$T_{\text{К}} = \frac{365}{365 - 52 - 10} = 1,205, \quad (5.5)$$

Таблица 5.5 – Трудозатраты на выполнение проекта

Этап	Исполнители	Продолжительность работ, дни			Длительность работ, чел/дн.			
					$T_{РД}$		$T_{КД}$	
		$t_{min}$	$t_{max}$	$t_{ож}$	НР	Ст	НР	Ст
Постановка целей и задач, получение данных	НР, Ст	2	4	2,8	2,35	1,34	3	2
Составление и утверждение ТЗ	НР, Ст	4	6	4,8	5,76	0,58	7	1
Подбор и изучение материалов по тематике	НР, Ст	12	16	13,6	4,90	16,32	6	20
Разработка календарного плана	НР, Ст	1	3	1,8	0,65	1,51	1	2
Обсуждение литературы	НР, Ст	2	5	3,2	0,77	3,84	1	5
Выбор алгоритмов для решения поставленных задач	НР, Ст	15	20	17	4,08	20,4	5	25
Планирование архитектуры программного обеспечения	НР, Ст	10	20	14	1,68	16,8	2	20
Разработка программного обеспечения	Ст	40	60	48	0	57,6	0	69
Тестирование и отладка программного обеспечения	Ст	10	20	14	0	16,8	0	20
Оформление расчетно-пояснительной записки	Ст	5	7	5,8	0	6,96	0	8
<b>Итого:</b>				<b>127,8</b>	<b>22,5</b>	<b>145,5</b>	<b>25</b>	<b>172</b>

### 5.3.3 Расчет нарастания технической готовности работ

Величина нарастания технической готовности работы показывает процент выполнения работы на каждом этапе. Расчетная формула:

$$H_i = \frac{t_{Hi}}{t_0} \cdot 100\%, \quad (5.6)$$

где  $t_{Hi}$  – нарастающая трудоемкость с момента начала работы  $i$ -го этапа;  
 $t_0$  – общая трудоемкость.

Общая трудоемкость вычисляется по формуле:

$$t_0 = \sum_{i=1}^n t_{OЖi}, \quad (5.7)$$

где  $t_{OЖi}$  – ожидаемая продолжительность  $i$ -го этапа.

Удельный вес каждого этапа  $Y_i$  определяется по формуле:

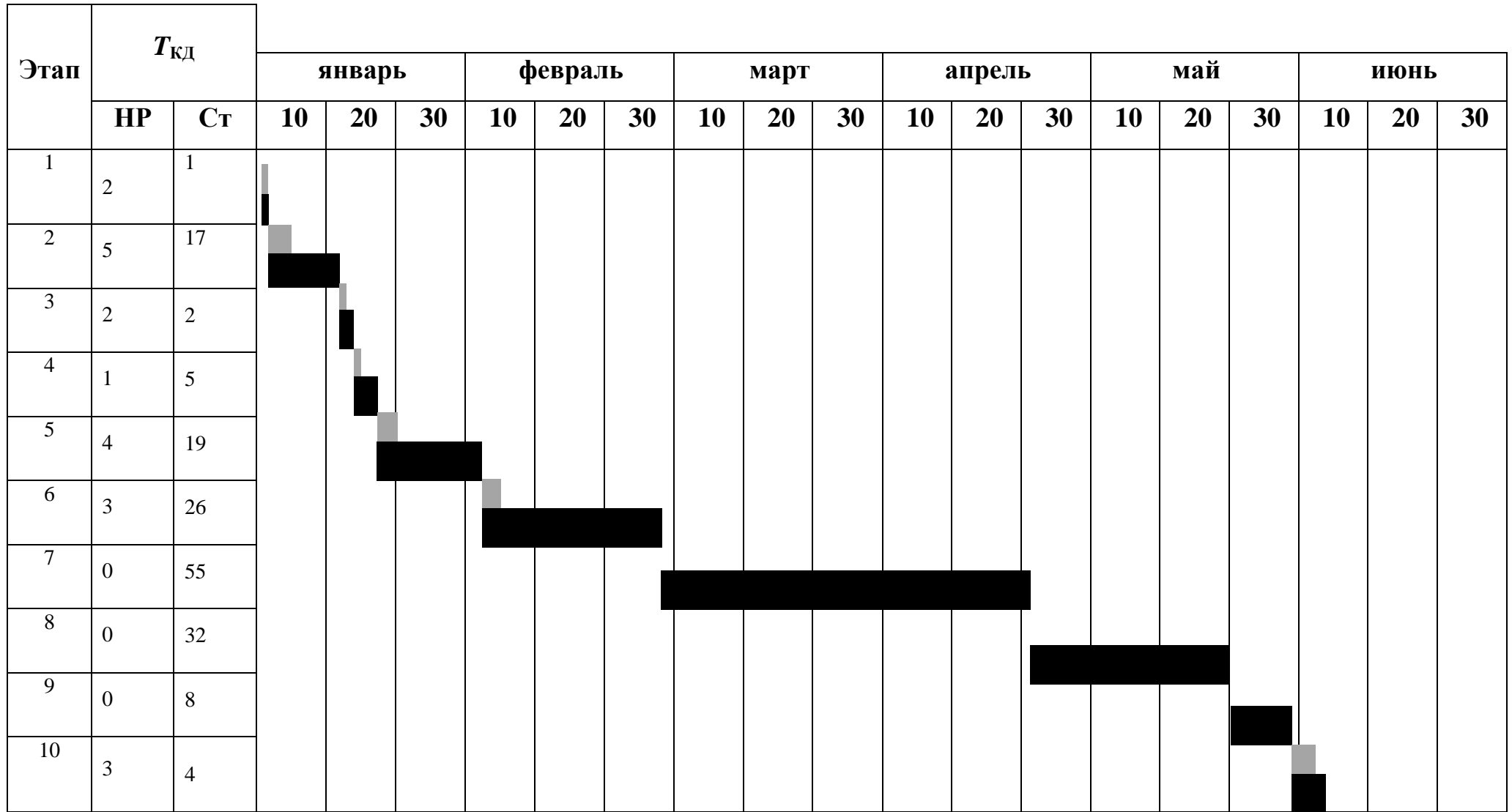
$$Y_i = \frac{t_{OЖi}}{t_0} \cdot 100\% \quad (5.8)$$

Результаты вычислений  $H_i$  и  $Y_i$  отражены в таблице 5.6.

Таблица 5.6 – Нарастание технической готовности и удельный вес этапов

Этап	$H_i, \%$	$Y_i, \%$
Постановка целей и задач, получение исходных данных	2,19	2,19
Составление и утверждение ТЗ	5,95	3,76
Подбор и изучение материалов по тематике	16,59	10,64
Разработка календарного плана	18,00	1,41
Обсуждение литературы	20,50	2,50
Выбор алгоритмов для решения поставленных задач	33,80	13,30
Планирование архитектуры программного обеспечения	44,76	10,95
Разработка программного обеспечения	82,32	37,56
Тестирование и отладка программного обеспечения	93,27	10,95
Оформление расчетно-пояснительной записки	97,81	4,54
Подведение итогов	100	2,19

Таблица 5.7 – Линейный график работ



■ НР    ■ Ст

## **5.4 Расчет сметы затрат на разработку ПО и проведение эксперимента**

В состав затрат на разработку ПО и проведение эксперимента включается стоимость всех расходов, необходимых для реализации комплекса работ, составляющих содержание данной разработки. Расчет сметной стоимости на выполнение данной разработки производится по следующим статьям затрат:

- материалы и покупные изделия;
- основная заработная плата;
- отчисления в социальные фонды;
- расходы на электроэнергию;
- амортизационные отчисления;
- работы, выполняемые сторонними организациями;
- прочие расходы.

### **5.4.1 Расчет затрат на материалы**

К данной статье расходов относится стоимость материалов, покупных изделий, полуфабрикатов и других материальных ценностей, расходуемых непосредственно в процессе выполнения работ. Так как основным рабочим инструментом при разработке данного проекта выступал персональный компьютер, затраты на материалы и оборудование составляют 0 рублей в связи с отсутствием необходимости в расходных материалах.

### **5.4.2 Расчет основной заработной платы**

Данная статья расходов включает заработную плату научного руководителя и студента, а также премии, входящие в фонд заработной платы. Расчет основной заработной платы выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя.

Среднедневная заработная плата рассчитывается по формуле:

$$\text{Дневная з/плата} = \frac{\text{Месячный оклад}}{25,17 \text{ дней}} \quad (5.9)$$

Расчеты затрат на основную заработную плату приведены в таблице 6. При расчете учитывалось, что в году 302 рабочих дня и, следовательно, в месяце 25,17 рабочих дня. Затраты времени на выполнение работы по каждому исполнителю брались из таблицы 2. Также был принят во внимание коэффициент, учитывающий коэффициент по премиям  $K_{\text{доп.ЗП}} = 1,188$ ,  $K_{\text{ПР}} = 1,1$  и районный коэффициент  $K_{\text{РК}} = 1,3$  ( $K = 1 * K_{\text{ПР}} * K_{\text{РК}} * K_{\text{РК}} = 1 * 1,1 * 1,3 * 1,188 = 1,699$ ).

Расчет основной заработной платы исполнителей системы выбирается на основе системы оплаты труда в ТПУ (для руководителя). Для исполнителя (студент) предусматривается расчет оплаты труда исходя из системы оплаты труда предприятия.

Таблица 5.8 – Затраты на основную заработную плату

Исполнитель	Оклад, руб./мес.	Среднедневная ставка, руб./день	Затраты времени, дни	Коэффициент	Фонд з/платы, руб.
НР	33664	1337,46	22,54	1,699	51218,64
Ст	9489	376,99	145,51	1,62	88616,02
<b>Итого:</b>					139834,66

#### 5.4.3 Расчет отчислений от заработной платы

Затраты по этой статье составляют отчисления в социальные фонды.

Отчисления по заработной плате определяются по следующей формуле:

$$C_{\text{соц}} = K_{\text{соц}} \cdot C_{\text{осн}} \quad C_{\text{соц}} = K_{\text{соц}} * C_{\text{зп}}, \quad (5.10)$$

где  $K_{\text{соц}}$  – коэффициент, учитывающий размер отчислений из заработной платы. Данный коэффициент составляет 30% от затрат на заработную плату и включает в себя:

- отчисления в пенсионный фонд;

- на социальное страхование;
- на медицинское страхование.

Итак, отчисления из заработной платы составили:

$$C_{\text{соц}} = 0,3 * 139834,66 = \mathbf{41950,4 \text{ руб.}}$$

#### 5.4.4 Расчет затрат на электроэнергию

Данный вид расходов включает в себя затраты на электроэнергию при работе оборудования. Затраты на электроэнергию при работе оборудования для технологических целей рассчитываются по формуле:

$$Э_{\text{ОБ}} = P_{\text{ОБ}} \cdot Ц_{\text{Э}} \cdot t_{\text{ОБ}}, \quad (5.11)$$

где  $Э_{\text{ОБ}}$  – затраты на электроэнергию, потребляемую оборудованием, руб.;

$P_{\text{ОБ}}$  – мощность, потребляемая оборудованием, кВт;

$Ц_{\text{Э}}$  – тарифная цена за 1 кВт·час,  $Ц_{\text{Э}} = 5,782 \text{ руб/кВт·час}$ ;

$t_{\text{ОБ}}$  – время работы оборудования, час.

Время работы оборудования вычисляется на основе данных для  $T_{\text{РД}}$  таблицы 4 для инженера из расчета, что продолжительность рабочего дня равна 8 часов.

Мощность, потребляемая оборудованием, определяется по формуле:

$$P_{\text{ОБ}} = P_{\text{УСТ.ОБ}} \cdot K_{\text{С}}, \quad (5.12)$$

где  $P_{\text{УСТ.ОБ}}$  – установленная мощность оборудования, кВт;

$K_{\text{С}}$  – коэффициент загрузки, зависящий от средней степени использования номинальной мощности.

Для технологического оборудования малой мощности  $K_{\text{С}} = 1$ .

Затраты на электроэнергию для технологических целей приведены в таблице 6.

Таблица 5.9 – Затраты на электроэнергию для технологических целей

Наименование оборудования	Время работы оборудования $t_{\text{ОБ}}$ , час	Потребляемая мощность $P_{\text{ОБ}}$ , кВт	Затраты $Э_{\text{ОБ}}$ , руб.
Персональный компьютер	989,5	0,3	1716,4
<b>Итого:</b>			<b>1716,4</b>



#### 5.4.5 Расчет прочих расходов

В статье «Прочие расходы» отражены расходы на разработку проекта, которые не учтены в предыдущих статьях.

Прочие расходы составляют 10 % от единовременных затрат на выполнение технического продукта и проводятся по формуле:

$$C_{\text{проч}} = (C_{\text{зп}} + C_{\text{соц}} + C_{\text{эл.об.}} + C_{\text{ам}}) * 0,1 \quad (5.14)$$

$$C_{\text{проч}} = (139834,2 + 41950,4 + 1716,4) * 0,1 = 18350,1 \text{ руб.}$$

#### 5.4.6 Расчет общей себестоимости разработки

Проведя расчет сметы затрат на разработку, можно определить общую стоимость разработки проекта «Разработка и исследование алгоритмов распараллеливания многомерного статистического анализа на примере метеорологических полей».

Таблица 5.10 – Смета затрат на разработку проекта

Статья затрат	Условное обозначение	Сумма, руб.
1 Материалы и покупные изделия	$C_{\text{МАТ}}$	0
2 Основная заработная плата	$C_{\text{ЗП}}$	139834,66
3 Отчисления в социальные фонды	$C_{\text{СОЦ}}$	41950,4
4 Расходы на электроэнергию	Э	1716,4
5 Прочие расходы	$C_{\text{ПРОЧ}}$	18350,1
<b>Итого:</b>		<b>201850,5</b>

Таким образом, расходы на разработку составили  $C = 201\ 850,5$  руб.

#### 5.5 Оценка экономической эффективности проекта

Выполнение научно-исследовательских работ оценивается уровнями достижения экономического, научного, научно-технического и социального эффектов.

Научный эффект характеризует получение новых научных знаний и отображает прирост информации, предназначенной для внутри научного потребления. Научно-технический эффект характеризует возможность использования результатов в других проектах и обеспечивает получение информации, необходимой для создания новой техники. Экономический эффект характеризуется в частности выраженной в стоимостных показателях экономией живого общественного труда, а также в возможности применения полученных знаний для создания новых разработок. Социальный эффект проявляется в улучшении условий труда.

Экономическая эффективность проекта обусловлена острой необходимостью и потребностью глобальных вычислений метеорологических данных за максимально короткие сроки. Достижение необходимой производительности путем использования алгоритмов для анализа больших данных является важнейшей темой в условиях технического прогресса на сегодняшний день.

Конечная система позволяет производить расчеты значительно быстрее обычных способов.

Использование подобной системы в коммерческой и социальной сфере позволяет решать такие задачи, как оперативная обработка данных, анализ данных своевременно, и своевременно давать оценку происходящим процессам в природе.

Подводя итог вышесказанному, экономический эффект от реализации проекта может быть выражен в снижении затрат на покупку подобных систем от сторонних разработчиков, а также повышение эффективности обработки данных за счет применения разработанного решения на текущем оборудовании.

### **5.5.1 Оценка научно-технического уровня НИР**

Научно-технический уровень характеризует, в какой мере выполнены работы и обеспечивается научно-технический прогресс в данной области.

Для оценки научной ценности, технической значимости и эффективности, планируемых и выполняемых НИР, используется метод бальных оценок. Бальная оценка заключается в том, что каждому фактору по принятой шкале присваивается определенное количество баллов. Обобщенную оценку проводят по сумме баллов по всем показателям или рассчитывают по формуле. На этой основе делается вывод о целесообразности НИР.

Сущность метода заключается в том, что на основе оценок признаков работы определяется коэффициент ее научно-технического уровня по формуле:

$$K_{НТУ} = \sum_{i=1}^3 R_i \cdot n_i, \quad (5.15)$$

где  $K_{НТУ}$  – коэффициент научно-технического уровня;

$R_i$  – весовой коэффициент  $i$ -го признака научно-технического эффекта;

$n_i$  – количественная оценка  $i$ -го признака научно-технического эффекта, в баллах.

Таблица 5.11 – Весовые коэффициенты признаков НТУ

Признак НТУ	Примерное значение весового коэф-та $n_i$	
1 Уровень новизны	Систематизируются и обобщаются сведения, определяются пути дальнейших исследований	0,4
2 Теоретический уровень	Разработка способа (алгоритм, программа мероприятий, устройство, вещество и т.п.)	0,1
3 Возможность реализации	Время реализации в течение первых лет	0,5

Таблица 5.12 – Баллы для оценки уровня новизны

Уровень новизны	Характеристика уровня новизны	Баллы
Принципиально новая	Новое направление в науке и технике, новые факты и закономерности, новая теория, вещество, способ	8 – 10

Новая	По-новому объясняются те же факты, закономерности, новые понятия дополняют ранее полученные результаты	5 – 7
Относительно новая	Систематизируются, обобщаются имеющиеся сведения, новые связи между известными факторами	2 – 4
Не обладает новизной	Результат, который ранее был известен	0

Таблица 5.13 – Баллы значимости теоретических уровней

Теоретический уровень полученных результатов	Баллы
1 Установка закона, разработка новой теории	10
2 Глубокая разработка проблемы, многоспектральный анализ, взаимодействия между факторами с наличием объяснений	8
3 Разработка способа (алгоритм, программа и т. д.)	6
4 Элементарный анализ связей между фактами (наличие гипотезы, объяснения версии, практических рекомендаций)	2
5 Описание отдельных элементарных факторов, изложение наблюдений, опыта, результатов измерений	0,5

Таблица 5.14 – Возможность реализации научных, теоретических результатов по времени и масштабам

Время реализации	Баллы
В течение первых лет	10
От 5 до 10 лет	4
Свыше 10 лет	2

Результаты оценок признаков научно-технического уровня приведены в таблице 5.15.

Таблица 5.15 – Количественная оценка признаков НИОКР

<b>Признак научно-технического эффекта НИР</b>	<b>Характеристика признака НИОКР</b>	<b><math>R_i</math></b>
Уровень новизны	Систематизируются, обобщаются имеющиеся сведения, новые связи между известными факторами	00,4
Теоретический уровень	Разработка способа (алгоритм, программа мероприятий, устройство, вещество и т.п.)	00,1
Возможность реализации	Время реализации 5-10 лет	00,5

Обоснование оценки признаков НИОКР приводится в таблице 5.16.

Таблица 5.16 – Оценки научно-технического уровня НИР

<b>Фактор НТУ</b>	<b>Значимость</b>	<b>Уровень фактора</b>	<b>Выбранный балл</b>	<b>Обоснование выбранного балла</b>
Уровень новизны	0,4	Относительно новая	4	Увеличит быстродействие обработки метеоданных
Теоретический уровень	0,1	Разработка способа	6	Разработана и реализован алгоритм, решающий поставленные в рамках задания задачи
Возможность реализации	0,5	В течение первых лет	7	Полученный продукт находится на стадии тестирования и внедрения

Исходя из оценки признаков НИОКР, показатель научно-технического уровня для данного проекта составил:  $K_{НТУ} = 0,5 \cdot 4 + 0,3 \cdot 6 + 0,2 \cdot 12 = 6,2$ . Таким образом, исходя из данных в таблице 5.17, проект «Построение регрессионных моделей для прогнозирования метеорологических величин» имеет средний уровень научно-технического эффекта.

Таблица 5.17 – Оценка уровня научно-технического эффекта

Уровень НТЭ	Показатель НТЭ
Низкий	1-4
Средний	4-7
Высокий	8-10

## 5.6 Вывод по разделу

Проведенный SWOT-анализ показ на что в первую очередь стоит обратить внимание при разработке программного продукта для построения регрессионных моделей для прогнозирования метеорологических величин

На разработку данного проекта потребовалось 167 дн., для подробного описания трудозатрат был представлен календарный план проекта. Итоговый бюджет проекта, исходя из расчетов всех затрат, составил около 201,85 тыс. руб.

В заключении стоит отметить, что в результате внедрения данной системы, организация улучшит процесс обработки информации и эффективность математических вычислений, что является основной целью организации, путем использования разработанных алгоритмов и автоматизации основных работ. Данные улучшения позволят снизить затраты на некоторые виды работ пользователей, что позволит направить сохраненное время для решения других задач.

## **6 Социальная ответственность**

### **6.1 Введение**

Разрабатываемый в рамках магистерской диссертации проект является программным вычислительным комплексом, который предназначен для построения регрессионных моделей для прогнозирования метеорологических величин, а как следствие увеличение предсказуемости этих параметров во времени. Работа над данным проектом проходит с использованием персонального компьютера. Взаимодействие с данной программой осуществляется с помощью персонального компьютера. Данный раздел посвящен анализу различных вредных и опасных факторов производственной среды для пользователей персональных компьютеров при работе с разработанным программным продуктом.

### **6.2 Производственная безопасность**

Для того чтобы обеспечить необходимый уровень производственной безопасности, нужно провести анализ возможных воздействий на человека вредных и опасных производственных факторов, которые могут возникнуть при разработке программного продукта, а также при его эксплуатации.

Производственный фактор считается вредным, если воздействие этого фактора на работника может привести к его заболеванию. Производственный фактор считается опасным, если его воздействие на работника может привести к его травме [47].

Производственные факторы классифицируются по группам:

- физические;
- химические;
- биологические;
- психофизические;

Так как данная работа не связана с химическими и биологическими факторами, то их мы рассматривать не будем. Следовательно рассмотрим физические и психофизические факторы связанные как с разработчиком

продукта, так и с конечными пользователями, и их рабочими зонами.  
Факторы представлены в таблице 6.1.



Таблица 6.1 – Вредные и опасные производственные факторы при выполнении работ за ПЭВМ [42].

Источник фактора, наименование видов работ	Факторы (по ГОСТ 12.0.003-2015)		Нормативные документы
	Вредные	Опасные	
1. Персональный компьютер, работа за персональным компьютером.	1. Недостаточная освещенность рабочей зоны. 2. Микроклимат. 3. Шум. 4. Умственное напряжение.	1. Опасность поражения электрическим током 2. Короткое замыкание. 3. Статическое электричество	1. СанПиН 2.2.4.548-96. 2. СанПиН 2.2.2/2.4.1340-03. 3. СП 52.13330.2011. 4. ГОСТ Р 12.1.019-2009 ССБТ. 5. СНиП 21-01-97.

## 6.2.1 Вредные производственные факторы

### 6.2.1.1 Недостаточная освещенность рабочей зоны

Большую часть информации человек получает посредством зрения. Вследствие этого важную роль имеет используемое освещение. При недостаточной освещенности рабочего места, вынуждает работника работать с контрастным фоном, а это вызывает переутомление, снижает производительность и концентрацию. Аналогичное происходит при избыточном освещении. Требования к освещению регламентируются СП 52.13330.2011[43].

Работа за персональным компьютером относится к зрительным работам согласно СП 52.13330.2011[43] высокой точности и соответствует третьему разряду зрительных работ. В связи с этим недостаточная или избыточная освещенность снижает производительность труда, повышает утомляемость, количество допускаемых ошибок и т.д. Требования к освещению помещений представлены в таблице 6.2.

Таблица 6.2 – Требования к освещению помещений[44].

Характеристика зрительной работы	Наименьший или эквивалентный размер объекта различения, мм	Разряд зрительной работы	Подразряд зрительной работы	Контраст объекта с фоном	Характеристика фона	Искусственное освещение		
						Освещённость, лк		
						При системе комбинированного освещения		При системе общего освещения
						всего	В том числе от общего	
Высокой точности	От 0,3 до 0,5	III	г	Средний, большой	Светлый, средний	400	200	200

Рабочее место разработчика было оснащено естественным и искусственным освещением. Для рассеивания избыточного естественного света в помещении предусмотрены жалюзи. Искусственное освещение обеспечивали люминесцентные лампы по всей площади потолка, создавая равномерное освещение рабочих мест. Исходя из нормативных требований к освещению, можно сделать вывод, что помещение, в котором происходила работа разработчика удовлетворяет им.

### 6.2.1.2 Микроклимат рабочего места

К основными показателями микроклимата в производственных помещениях относятся: температура воздуха, относительная влажность воздуха, температура поверхностей, скорость движения воздуха, интенсивность теплового облучения [45]. Оптимальные параметры микроклимата на рабочих местах должны соответствовать величинам, приведенным в таблице 6.3, применительно к выполнению работ различных категорий в холодный и теплый периоды года.

Категории работ разграничиваются исходя из энерготрат организма в ккал/ч (Вт). Работы, производимые сидя и сопровождающиеся незначительным физическим напряжением, исходя из данных СанПиН

2.2.4.548-96 [44], относятся к категории работ Ia. В данную категорию можно отнести работу разработчика системы и ее пользователей.

Таблица 6.3 – Оптимальные показатели микролимата на рабочих местах[44]

Период года	Категория работ по уровню затрат, Вт	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	Ia (до 139)	22-24	21-25	60-40	0,1
Теплый	Ia (до 139)	23-25	22-26	60-40	0,1

В данном случае температура воздуха и температура поверхностей составляют 22<sup>0</sup>С и 21<sup>0</sup>С при относительной влажности 45% в холодный период года; 24<sup>0</sup>С и 23<sup>0</sup>С при относительной влажности воздуха 50% в теплый период года, что соответствует нормам.

В процессе работы с компьютером и разработчик, и пользователи сталкиваются с воздействием на них тепла, которое выделяет вычислительная техника, и как следствие способствует бстрой утомляемости организма. Для обеспечения оптимальной температуры воздуха в помещения где происходит разработка продукта производится проветривание помещения (естественная вентиляция). В теплое время года, в помещении используется вентилятор (механическая вентиляция). В холодное время производится отопление с помощью системы водяного отопления. Все перечисленные меры направлены на установление отимальных параметров микролимата на рабочем месте.

### 6.2.1.3 Шум

Важным фактором, который влияет на самочувствие человека на работе, его трудоспособности является уровень шума в помещении. В помещениях, где разрабатывался программный продукт, и в помещениях, где предполагается использовать данный продукт, к источникам возникновения

шума можно отнести: работу всех офисных электроприборов, наружный шум, например, уличный.

Требования к предельно допустимому уровню шуму на рабочих местах установлены в СН 2.2.4/2.1.8.562-96. В таблице 6.4 представлены допустимые значения уровней звукового давления для рабочих мест в зависимости от трудовой деятельности.

Таблица 6.4 – Предельно допустимые уровни звукового давления, уровни звука и эквивалентные уровни звука для основных наиболее типичных видов трудовой деятельности и рабочих мест [44].

Вид трудовой деятельности, рабочее место	Уровни звукового давления, дБ, в октавных полосах со среднегеометрическими частотами, Гц									Уровни звука и эквивалентные уровни (в дБА)
	31,5	63	125	205	500	1000	2000	4000	8000	
Творческая деятельность, руководящая работа с повышенными требованиями, научная деятельность, конструирование и проектирование, программирование, преподавание и обучение, врачебная деятельность. Рабочие места в помещениях дирекции, проектно-конструкторских бюро, расчетчиков, программистов вычислительных машин, в лабораториях для теоретических работ и обработки данных, приема больных в здравпунктах	86	71	61	54	49	45	42	40	38	50
Работа, выполняемая с часто получаемыми указаниями и акустическими сигналами; работа, требующая постоянного слухового контроля; операторская работа по точному графику с инструкцией; диспетчерская работа. Рабочие места в помещениях диспетчерской службы, кабинетах и помещениях наблюдения и дистанционного управления с речевой связью по телефону; машинописных бюро, на участках точной сборки, на телефонных и телеграфных станциях, в помещениях мастеров, в залах обработки информации на вычислительных машинах	96	83	74	68	63	60	57	55	54	65

Уровень шума помещения, в котором рассматривается рабочее место, не превышает 50 дБа, что соответствует нормативным документам.

#### 6.2.1.4 Умственное перенапряжение

Главный инструмент пользователя или разработчика при работе с программным продуктом – это умственная деятельность. Такого рода деятельность является наиболее сложной формой труда, протекающей с повышенным эмоциональным напряжением. Как правило, повышенное эмоциональное напряжение ведет к нервно-психическим перегрузкам людей.

Поэтому меры, позволяющие снизить воздействие этого вредного производственного фактора, которые регулируются СанПиН 2.2.2/2.4.1340-03[45], являются важными в работе оператора ПЭВМ. Они позволяют увеличить производительность труда и предотвратить появление профессиональных болезней.

Работа разработчика, рассматриваемая в данной работе системы, можно отнести к группам А и Б, в то время, как деятельность пользователя, который будет использовать систему в профессиональной деятельности, относится к группе В. Категории трудовой деятельности различаются по степени тяжести выполняемых работ. Для снижения воздействия рассматриваемого вредного фактора предусмотрены регламентированные перерывы для каждой группы работ – таблица 6.5.

Таблица 6.5 – Суммарное время регламентированных перерывов в зависимости от продолжительности работы, вида категории трудовой деятельности с ПЭВМ[4].

Категория работы с ПЭВМ	Уровень нагрузки за рабочую смену при видах работ с ПЭВМ			Суммарное время регламентированных перерывов, мин.	
	группа А, количество знаков	группа Б, количество знаков	группа В, ч	при 8-часовой смене	при 12-часовой смене
I	до 20 000	до 15 000	до 2	50	80
II	до 40 000	до 30 000	до 4	70	110
III	до 60 000	до 40 000	до 6	90	140

Для разработчика программного продукта были предусмотрены перерывы, составляющие 60 мин., что удовлетворяет показателям таблицы 5.

К мерам предупреждения умственного напряжения, можно отнести производственную гимнастику, изменение рабочей позы в процессе работы, а

также смену деятельности. В последнее время, все больше предприятий организуют активные перерывы: прогулка по улице, катание на велосипедах, посещение бассейна и т.д.

## **6.2.2 Опасные производственные факторы**

### **6.2.2.1 Опасность поражения электрическим током**

Поражение электрическим током является опасным производственным фактором и, поскольку оператор ПЭВМ имеет дело с электрооборудованием, то вопросам электробезопасности на его рабочем месте должно уделяться много внимания. Нормы электробезопасности на рабочем месте регламентируются СанПиН 2.2.2/2.4.1340-03[45], вопросы требований к защите от поражения электрическим током освещены в ГОСТ Р 12.1.019-2009 ССБТ[46]

Помещение, где расположено рабочее место оператора ПЭВМ, относится к помещениям без повышенной опасности поскольку отсутствуют: сырость, токопроводящая пыль, токопроводящие полы, высокая температура, возможность одновременного прикосновения человека к имеющим соединение с землей металлоконструкциям зданий, технологическим аппаратам, механизмам и металлическим корпусам электрооборудования.

Для оператора ПЭВМ при работе с электрическим оборудованием обязательны следующие меры предосторожности:

- Перед началом работы нужно убедиться, что выключатели и розетка закреплены и не имеют оголённых токоведущих частей.
- При обнаружении неисправности оборудования и приборов необходимо, не делая никаких самостоятельных исправлений, сообщить человеку, ответственному за оборудование.

### **6.2.2.2 Короткое замыкание**

Последствия короткого замыкания приводит к весьма серьезным последствиям. Во-первых, достаточно часто оно сопровождается выходом из строя электроустановок и возникновением в них пожаров. Во-вторых, из-за резкого увеличения силы тока в цепи отдельные части кабеля могут быть подвергнуты механическому воздействию, в результате чего появятся механические и термические повреждения. В-третьих, достаточно часто короткое замыкание сопровождается значительным падением напряжения в цепи или на отдельных ее участках. Это, в свою очередь, ведет к ухудшению работы электрооборудования. В-четвертых, это явление оказывает крайне негативное влияние на находящиеся поблизости приборы, провода и другое электрическое оборудование. Защита от короткого замыкания включает в себя целый комплекс мер, исходным пунктом в которых является профилактика повреждений линий электропередач и оборудования. Кроме того, чтобы предотвратить возникновение пожара, используют специальные приборы – плавкиеставки, которые при замыкании сгорают и размыкают электрическую цепь.

### **6.2.2.3 Статическое электричество**

Токи статического электричества, наведенные в процессе работы компьютера на корпусах монитора, системного блока и клавиатуры, могут приводить к разрядам при прикосновении к этим элементам. Такие разряды опасности для человека не представляют, но могут привести к выходу из строя компьютера. Для снижения величин токов статического электричества используются нейтрализаторы, местное и общее увлажнение воздуха, использование покрытия полов с антистатической пропиткой [47].

### **6.2.3 Минимизация воздействия вредных и опасных факторов**

Обеспечить поддержку оптимальных условий освещений позволяют правила, описанные в СанПиН 2.2.2/2.4.1340-03[45]. Рабочее место должно быть оснащено естественным и искусственным освещением и



соответствовать показателям, отображенным в таблице 3. Для рассеивания избыточного естественного помещения рекомендуется использовать жалюзи, рулонные шторы и др. При искусственном освещении рекомендуется использовать люминесцентные лампы и лампы накаливания, а также светодиодные лампы.

Для поддержания оптимальных параметров микроклимата, рабочие места необходимо оснащать в теплый период системами кондиционирования воздуха, вентиляторами (механической вентиляция), а так же проветривать помещение (естественная вентиляция). В зимний период системами отопления. В помещения где проходит работы на персональных компьютерах необходима влажная уборка, так как пыль способствует ухудшению работы оборудования, перегреву и др.

Для снижения уровня шума предусматриваются следующие меры: различные звукоизоляционные конструкции, например, звукоизоляционные перегородки; уплотнение притворов окон, дверей; укрытия для источников шума.

Когда основная работа подразумевает взаимодействие с компьютером долгое время, необходимо организовать перерывы на каждые 10-15 минут раз в 45-60 минут, это позволит снизить утомляемость и перенапряжение.

Для минимизации возможности поражения электрическим током необходимо проведение мероприятий:

1. При производстве монтажных работ необходимо использовать только исправный инструмент, аттестованный службой КИПиА.
2. С целью защиты от поражения электрическим током, возникающим между корпусом приборов и инструментом при пробое сетевого напряжения на корпус, корпуса приборов и инструментов должны быть заземлены
3. При включенном сетевом напряжении работы на задней панели должны быть запрещены

4. Все работы по устранению неисправностей должен производить квалифицированный персонал

5. Необходимо постоянно следить за исправностью электропроводки.

### **6.3 Экологическая безопасность**

Разработанный программный продукт не несет вред окружающей среде, однако, в результате вывода из эксплуатации эта продукция превращается в отходы, которые содержат токсичные вещества, представляющие собой существенную угрозу для окружающей среды, жизни и здоровья людей.

Согласно ГОСТ Р 55102-2012[48] персональный компьютер содержит следующие компоненты, которые должны быть отдельно собраны при выводе его из эксплуатации, представленные в таблице 6.6.

Таблица 6.6 – Перечень основных компонентов некоторых типов электротехнического и электронного оборудования

Компоненты	Тип оборудования	
	Системный блок и клавиатура персонального компьютера	Монитор персонального компьютера
Металлы	✓	
Двигатель	✓	
Пластик	✓	✓
ЭЛТ		✓
ЖК-экран		✓
Электропровода	✓	
Трансформатор	✓	
Печатные платы	✓	✓
ХИТ	✓	✓
Внешние электропровода	✓	✓

Перечисленные компоненты требуют специальной утилизации, поскольку игнорирование таких работ оказывает значительный вред человечеству и окружающему миру в целом. Для обеспечения правильной утилизации необходимо обращаться в специальные организации, имеющие государственные сертификаты, документация и профессиональное оборудование для осуществления услуг утилизации отходов оргтехники.

Для поддержания оптимальных требований к освещению зачастую используются люминесцентные лампы, а такие лампы содержат опасное химическое вещество – ртуть. Поскольку ртуть может служить отравлением

для живых организмов, такие лампы требуют специальной утилизации, а юридические лица должны вести паспорт для данного вида отходов.

Стоит отметить, что работа в помещениях предполагает использования большого количества бумаги и другой макулатуры, которая также подлежит соответствующей утилизации.

#### **6.4 Безопасность в чрезвычайных ситуациях**

Чрезвычайные ситуации (ЧС) – возникшие обстоятельства, оказывающие влияние на всех людей вне зависимости от их рода деятельности. По природе своего воздействия ЧС делят на следующие группы [6]:

1. Техногенные – пожары, взрывы, обрушение зданий, аварии на системах жизнеобеспечения.
2. Природные – землетрясения, наводнения, бури и др.
3. Экологические – кислотные дожди, аномальное природное загрязнение атмосферы, опустынивание земель и др.
4. Биологические – эпидемии, эпизоотии, эпифитотии.
5. Социальные – межнациональные конфликты, терроризм, войны и др.
6. Антропогенные – ЧС в следствии ошибочных действий людей.

Самой распространенной ЧС в помещениях является пожар. По пожарной опасности помещения, где происходит работа на персональных компьютерах, относятся к категории В [47].

##### **6.4.1 Минимизация возможности возникновения пожара**

В первую очередь, для того, чтобы предотвратить возникновение пожара в помещениях с ПЭВМ, должен осуществляться комплекс работ по обеспечению пожарной безопасности. Требования по обеспечению пожарной безопасности устанавливаются согласно НПБ [49].

В помещении с ПЭВМ должен быть определен порядок обесточивания электрооборудования по окончании рабочего дня, порядок

прохождения противопожарного инструктажа, помещения должны оснащаться углекислотными огнетушителями. Дополнительно к огнетушителям на каждые 200 м<sup>2</sup> площади рекомендуется иметь: грубошерстную ткань или войлок размером не менее 1х1 м, асбестовое полотно и пожарный стенд с емкостью для песка не менее 0,1 м<sup>3</sup>.

Система электропитания ЭВМ должна иметь блокировку, обеспечивающую отключение ее в случае остановки системы охлаждения и кондиционирования. Следует предусматривать пожарную автоматику и датчики пожарной сигнализации. Целесообразно разработать план эвакуации имущества и людей на случай пожара. Эффективность эвакуации людей может быть обеспечена достаточным количеством запасных выходов с хорошим освещением и многочисленными сигнальными огнями и указателями на стенах, полах и т.д.

В случае возникновения пожара в помещении или в здании необходимо предпринять меры по эвакуации людей из него, используя для этого план эвакуации. При необходимости устранения очага возгорания в помещениях с ПЭВМ использовать огнетушители. Требуется отметить, что в таких помещениях запрещается использовать химические пенные огнетушители.

## **6.5 Правовые и организационные вопросы обеспечения безопасности**

### **6.5.1 Правовые нормы трудового законодательства**

Сотрудник, постоянно работающий с ПЭВМ, сталкивается с различными родами негативными воздействия на организм. Для минимизирования таких воздействий законодательно закреплено установление перерывов.

Согласно классификации видов трудовой деятельности с персональным компьютером, представленной в ТОИ Р-45-084-01[50], работу разработчика и пользователей с программным продуктом следует отнести к группе В, предполагающая работу в режиме диалога с компьютером. Согласно данной классификации, суммарное время регламентированных перерывов составляет от 30-120 минут в соответствии с категорией работ.

Таблица 6.7 – Регламентированное время перерывов при работе с ПЭВМ

Категории работы	Уровень нагрузки за рабочую смену при видах работ с компьютером	Суммарное время регламентированных перерывов, мин.	
	Группа В, час.	при 8-часовой смене	при 12-часовой смене
I	До 2,0	30	70
II	До 4,0	50	90
III	До 6,0	70	120

Работу разработчика и пользователей системы можно отнести к III категории работ, тогда, при 8-часовой рабочей смене, регламентированные перерывы следует устанавливать через 1,5 - 2,0 часа от начала рабочей смены и через 1,5 - 2,0 часа после обеденного перерыва продолжительностью 20 минут каждый или продолжительностью 15 минут через каждый час работы.

При 12-часовой рабочей смене регламентированные перерывы должны устанавливаться в первые 8 часов работы аналогично перерывам при 8-часовой рабочей смене, а в течение последних 4 часов работы, независимо от категории и вида работ, каждый час продолжительностью 15 минут.

Продолжительность рабочего дня не должна быть меньше указанного времени в договоре, но не больше 40 часов в неделю. Для работников до 16 лет – не более 24 часов в неделю, от 16 до 18 лет и инвалидов I и II группы – не более 35 часов.

При работе в ночное время продолжительность рабочей смены сокращается на один час. К работе в ночную смену не допускаются беременные женщины; работники, не достигшие возраста 18 лет; женщины, имеющие детей в возрасте до трех лет, инвалиды, работники, имеющие детей-инвалидов, а также работники, осуществляющие уход за больными членами их семей в соответствии с медицинским заключением, матери и отцы-одиночки детей до пяти лет.

Организация обязана предоставлять ежегодный отпуск продолжительностью 28 календарных дней. Дополнительные отпуска предоставляются работникам, занятым на работах с вредными или опасными условиями труда, работникам имеющими особый характер работы, работникам с ненормированным рабочим днем и работающим в условиях Крайнего Севера и приравненных к нему местностях.



## **6.5.2 Организация рабочего места при работе с ПЭВМ**

Правильная организация рабочего места пользователя ПЭВМ позволяет повысить эффективность работ и облегчить работу пользователей, сделать ее максимально удобной и безопасной. Сам разрабатываемый продукт не несет никакого влияния на организацию рабочей зоны.

Требования к помещениям для работы с ПЭВМ устанавливаются согласно СанПиН 2.2.2./2.4.1340-03. К таким требованиям относятся:

- При размещении рабочих мест с ПЭВМ расстояние между рабочими столами с видеомониторов, должно быть не менее 2,0 м, а расстояние между боковыми поверхностями видеомониторов - не менее 1,2 м.
- Экран видеомонитора должен находиться от глаз пользователя на расстоянии 600-700 мм, но не ближе 500 мм.
- Конструкция рабочего стула (кресла) должна обеспечивать поддержание рациональной рабочей позы при работе на ПЭВМ, позволять изменять позу с целью снижения статического напряжения мышц шейно-плечевой области и спины для предупреждения развития утомления.

## **6.6 Вывод по разделу**

В данном разделе рассматривались такие части как:

- Производственная безопасность.
- Экологическая безопасность.
- Безопасность в чрезвычайных ситуациях.
- Правовые и организационные вопросы обеспечения безопасности.

В производственной безопасности был проведён анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения, в результате было установлено, что все пункты, которые включены в данную часть соответствуют нормативным документам. Приведены меры по снижению вреда, оказывающие негативное воздействие на человека в процессе работы.

В экологической безопасности были рассмотрены вопросы утилизации отходов оргтехники, макулатуры. В ходе работы было установлено, что из самых распространённых источников ртутного загрязнения является вышедшие из эксплуатации люминесцентные лампы, которые также подлежат соответствующей утилизации.

При рассмотрении безопасности в чрезвычайных ситуациях было выявлено следующее, что наиболее типичной ЧС для помещения, в котором производится выполнение работы, является пожар.

Правовые и организационные вопросы обеспечения безопасности включают в себя требования к организации рабочих мест пользователей, которые соответствуют ГОСТу.

## **Заключение**

Был произведен анализ открытых источников метеорологических данных, в частности использовались данные Всероссийским научно-исследовательским институтом гидрометеорологической информации. Данные собраны за последние пятьдесят лет. Данные собраны по основным метеостанциям стран СНГ. Данные взяты за различные интервалы времени (сроки, года, месяцы, дни и часы). Структурирована и разработана метеорологическая база данных.

Был осуществлен обзор моделей и методов прогнозирования временных рядов, были выявлены достоинства и недостатки каждого класса моделей.

Произведен анализ базовых алгоритмов статистического анализа данных. Разработаны улучшенные алгоритмы для построения прогнозных моделей.

Разработан программный вычислительный комплекс с возможностью построения регрессионных моделей для прогнозирования метеорологических величин.

Построенные прогнозные модели не показали себя довольно эффективными. Возможно, нужно увеличить выборку временного ряда или прибегнуть к другим моделям прогнозирования, т.к. модуль SARIMA пакета statsmodels и библиотека Facebook Prophet достаточно новые. В целом для исследованные подходы вполне пригоден для научной работы.

## Список источников

1. Бокс Дж., Дженкинс Г.М. Анализ временных рядов, прогноз и управление. М.: Мир, 1974. 406 с.
2. Егошин А.В. Анализ и прогнозирование сложных стохастических сигналов на основе методов ведения границ реализаций динамических систем: Автореферат диссертации ... канд. техн. наук. Санкт-Петербург, 2009. 19 с.
3. Gheyas I.A., Smith L.S. A Neural Network Approach to Time Series Forecasting // Proceedings of the World Congress on Engineering, London, 2009, Vol 2. P. 1292 – 1296. URL: [www.iaeng.org/publication/WCE2009/WCE2009\\_pp1292-1296.pdf](http://www.iaeng.org/publication/WCE2009/WCE2009_pp1292-1296.pdf) (дата обращения: 28.05.2018).
4. Morariu N., Iancu E., Vlad S. A neural network model for time series forecasting // Romanian Journal of Economic Forecasting. 2009, No. 4. P. 213 – 223.
5. Mazengia D.H. Forecasting Spot Electricity Market Prices Using Time Series Models: Thesis for the degree of Master of Science in Electric Power Engineering. Gothenburg, Chalmers University of Technology, 2008. 89 p.
6. Нормативные системы в прогнозировании развития предпринимательского сектора экономики / Л.И. Муратова [и др.] // Управление экономическими системами. 2009, №20. URL: <http://uecs.mcnp.ru/modules.php?name=News&file=print&sid=145> (дата обращения: 28.05.2018).
7. Parzen E. Long memory of statistical time series modeling // NBER-NSF Time Series Conference, USA, Davis, 2004. 10 p. URL: <http://www.stat.tamu.edu/~eparzen/Long%20Memory%20of%20Statistical%20Time%20Series%20Modeling.pdf> (дата обращения: 28.05.2018).
8. Методы прогнозирования оптимальных доз инсулина для больных сахарным диабетом I типа. Обзор / С.А. Чернецов [и др.] // Наука и

образование. 2009, №9. URL: <http://technomag.edu.ru/doc/119663.html> (дата обращения: 28.05.2018).

9. Jingfei Yang M. Sc. Power System Short-term Load Forecasting: Thesis for Ph.d degree. Germany, Darmstadt, Elektrotechnik und Informationstechnik der Technischen Universität, 2006. 139 p.

10. Extrapolation // The free encyclopedia «Wikipedia». URL: <http://en.wikipedia.org/wiki/Extrapolation> (дата обращения: 28.05.2018).

11. Norizan M., Maizah Hura A., Zuhaimy I. Short Term Load Forecasting Using Double Seasonal ARIMA Model // Regional Conference on Statistical Sciences, Malaysia, Kelantan, 2010. P. 57 – 73.

12. Collantes-Duarte J., Rivas-Echeverriat F. Time Series Forecasting using ARIMA, Neural Networks and Neo Fuzzy Neurons // WSEAS International Conference on Neural Networks and Applications, Switzerland, 2002. 6 p. URL: [www.wseas.us/e-library/conferences/switzerland2002/papers/464.pdf](http://www.wseas.us/e-library/conferences/switzerland2002/papers/464.pdf) (дата обращения: 28.05.2018).

13. Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models / A.J. Conejo [at al.] // IEEE transaction on power systems. 2005, Vol. 20, No. 2. P. 1035 – 1042.

14. Тихонов Э.Е. Прогнозирование в условиях рынка. Невинномысск, 2006. 221 с.

15. Леоненков А. Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб: БХВ-Петербург, 2005. 736 с.

16. Armstrong J.S. Forecasting for Marketing // Quantitative Methods in Marketing. London: International Thompson Business Press, 1999. P. 92 – 119.

17. Семенов В.В. Математическое моделирование динамики транспортных потоков мегаполиса. М.: ИПМ им. М.В.Келдыша РАН, 2004. 44 с.

18. Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and

forecasting / J.B. Rundle [at al.] // Colloquium of the National Academy of Sciences, Irvine, USA, 2002. P. 2514 – 2521.

19. Draper N., Smith H. Applied regression analysis. New York: Wiley, In press, 1981. 693 p.

20. Maximum likelihood // The free encyclopedia «Wikipedia». URL: [http://en.wikipedia.org/wiki/Maximum\\_likelihood](http://en.wikipedia.org/wiki/Maximum_likelihood) (дата обращения: 28.05.2018).

21. Ивахненко А.Г. Обзор задач, решаемых по алгоритмам Метода Группового Учета Аргументов (МГУА) // Group Method of Data Handling. URL: <http://www.gmdh.net/articles/rus/obzorzad.pdf> (дата обращения: 28.05.2018).

22. Autoregressive conditional heteroskedasticity // The free encyclopedia «Wikipedia». URL: [http://en.wikipedia.org/wiki/Autoregressive\\_conditional\\_heteroskedasticity](http://en.wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity) (дата обращения: 28.05.2018).

23. Эконометрия: Учебное пособие / В.И. Сулов [и др.] Новосибирск: Издательство СО РАН, 2005. 744 с.

24. Prajakta S.K. Time series Forecasting using Holt-Winters Exponential Smoothing // Kanwal Rekhi School of Information Technology Journal. 2004. 13 p. URL: [http://www.it.iitb.ac.in/~praj/acads/seminar/04329008\\_ExponentialSmoothing.pdf](http://www.it.iitb.ac.in/~praj/acads/seminar/04329008_ExponentialSmoothing.pdf) (дата обращения: 28.05.2018).

25. Хайкин С. Нейронные сети: полный курс. М.: ООО «И. Д. Вильямс», 2006. 1104 с.

26. Pradhan R.P., Kumar R. Forecasting Exchange Rate in India: An Application of Artificial Neural Network Model // Journal of Mathematics Research. 2010, Vol. 2, No. 4. P. 111 – 117.

27. Yildiz B., Yalama A., Coskun M. Forecasting the Istanbul Stock Exchange National 100 Index Using an Artificial Neural Network // An International Journal of Science, Engineering and Technology. 2008, Vol. 46. P.36 – 39.

28. An Artificial Neural Network Approach for Day-Ahead Electricity Prices Forecasting / J. Catalao [at al.] // 6th WSEAS international conference on

Neural networks, USA, Stevens Point, 2005. P. 80 – 83. 29. Kumar M. Short-term load forecasting using artificial neural network techniques: Thesis for Master of Science degree in Electrical Engineering. India, Rourkela, National Institute of Technology, 2009. 48 p.

29. Zhu J., Hong J., Hughes J.G. Using Markov Chains for Link Prediction in Adaptive Web Sites // 1st International Conference on Computing in an Imperfect World, UK, London, 2002. P. 60 – 73.

30. Hannes Y.Y., Webb P. Classification and regression trees: A User Manual for Identifying Indicators of Vulnerability to Famine and Chronic Food Insecurity // International Food Policy Research Institute. 1999. 59 p. URL: [http://www.fao.org/sd/erp/toolkit/BOOKS/classification\\_and\\_regression\\_trees\\_intro.pdf](http://www.fao.org/sd/erp/toolkit/BOOKS/classification_and_regression_trees_intro.pdf) (дата обращения: 28.05.2018).

31. Huang W., Nakamoria Y., Wang S. Forecasting stock market movement direction support vector machine // Elsevier: computers and operation research. 2005, Vol. 32. P. 2513 – 2522.

32. Support vector machine // The free encyclopedia «Wikipedia». URL: [http://en.wikipedia.org/wiki/Support\\_vector\\_machine](http://en.wikipedia.org/wiki/Support_vector_machine) (дата обращения: 28.05.2018).

33. Mahfoud S., Mani G. Financial Forecasting Using Genetic Algorithms // Applied Artificial Intelligence. 1996, Vol. 10, No.6. P. 543 – 565.

34. Nogales F.J., Conejo A.J. Electricity price forecasting through transferfunction models // Journal of the Operational Research Society. 2006, Vol. 57, No. 4. P. 350 – 356.

35. Alfares H.K., Nazeeruddin M. Electric load forecasting: literature survey and classification of methods // International Journal of Systems Science. 2002, Vol 33. P. 23 – 34.

36. Hinman J., Hickey E. Modeling and forecasting short term electricity load using regression analysis // Journal of Institute for Regulatory Policy Studies. 2009. 51 p. URL: <http://www.irps.ilstu.edu/research>

/documents/LoadForecastingHinman-HickeyFall2009.pdf (дата обращения: 28.05.2018).

37. Fogler H.R. A pattern recognition model for forecasting // Management science. 1974, No.8. P. 1178 – 1189.

38. Discovering Patterns in Electricity Price Using Clustering Techniques / F. Martínez Álvarez [at al.] // ICREPQ International Conference on Renewable Energies and Power Quality, Spain, Sevilla, 2007. 8 p. URL: <http://www.icrepq.com/icrepq07/245-martinez.pdf> (дата обращения: 28.05.2018).

39. Prophet // Forecasting at scale. URL: <https://facebook.github.io/prophet/> (дата обращения: 28.05.2018).

40. Statsmodels Statistics in Python // Statsmodels Documentation. URL: <https://www.statsmodels.org/stable/index.html> (дата обращения: 28.05.2018).

41. scikit-learn Machine Learning in Python // scikit-learn. URL: <http://scikit-learn.org/stable/index.html> (дата обращения: 28.05.2018).

42. ГОСТ 12.0.003-2015. Система стандартов безопасности труда (ССБТ). Опасные и вредные производственные факторы. Классификация // Электронный фонд правовой и нормативно-технической документации. 2017. URL: <http://docs.cntd.ru/document/1200136071> (дата обращения: 10.05.2018 г.)

43. СП 52.13330.2011. Естественное и искусственное освещение // Электронный фонд правовой и нормативно-технической документации. 2011. URL: <http://docs.cntd.ru/document/1200084092> (дата обращения: 10.05.2018 г.)

44. СанПиН 2.2.4.548-96. Гигиенические требования к микроклимату производственных помещений // Электронный фонд правовой и нормативно-технической документации. 1996. URL: <http://docs.cntd.ru/document/901704046> (дата обращения: 10.05.2018 г.)

45. СанПиН 2.2.2/2.4.1340-03. Гигиенические требования к персональным электронно-вычислительным машинам и организации работы // Электронный фонд правовой и нормативно-технической документации. 2003. URL: <http://docs.cntd.ru/document/901865498> (дата обращения: 12.05.2018 г.)



46. ГОСТ Р 12.1.019-2009. Система стандартов безопасности труда (ССБТ). Электробезопасность. Общие требования и номенклатура видов защиты // Электронный фонд правовой и нормативно-технической документации. 2009. URL: <http://docs.cntd.ru/document/1200080203> (дата обращения: 12.05.2018 г.)

47. Назаренко О. Б. Безопасность жизнедеятельности: учебное пособие / О. Б. Назаренко, Ю. А. Амелькович; Томский политехнический университет. – 3-е изд., перераб. и доп. – Томск: Изд-во Томского политехнического университета, 2013. – 178 с.

48. ГОСТ Р 55102-2012. Ресурсосбережение. Обращение с отходами. Руководство по безопасному сбору, хранению, транспортированию и разборке отработавшего электротехнического и электронного оборудования, за исключением ртути содержащих устройств и приборов // Электронный фонд правовой и нормативно-технической документации. 2012. URL: <http://docs.cntd.ru/document/1200104723> (дата обращения: 15.05.2018 г.)

49. СНиП 21-01-97. Пожарная безопасность зданий и сооружений // Информационно-правовое обеспечение. 1997 г. URL: <http://base.garant.ru/2305928/> (дата обращения: 15.05.2018 г.)

50. ТОИ Р-45-084-01. Типовая инструкция по охране труда при работе на персональном компьютере // Справочная правовая система КонсультантПлюс. URL: [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_79762/8b2ed343ae4bcf0636cсе936afa1156fb10b78ae/](http://www.consultant.ru/document/cons_doc_LAW_79762/8b2ed343ae4bcf0636cсе936afa1156fb10b78ae/) (дата обращения: 21.05.2018 г.)

# ПРИЛОЖЕНИЕ А

## РАЗДЕЛЫ НА ИНОСТРАННОМ ЯЗЫКЕ

### Раздел 1

#### Обзор и классификация моделей и методов прогнозирования временных рядов

Студент:

Группа	ФИО	Подпись	Дата
8ВМ6Б	Солтаганов Николай Андреевич		

Консультант отделения ИТ

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Дорофеев Вадим Анатольевич	-		

Консультант – лингвист отделения ИЯ:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Куркан Наталия Владимировна	-		

## **1 Overview and classification of models and methods of time series forecasting**

Paper [14] contains the analysis of methods. All methods of forecasting were divided into two groups: intuitive and formal.

Intuitive forecasting is used for forecasting objects that are too simple, or, on the contrary, very complex. Complexity shall be understood to mean availability of a large number of external factors that cannot be taken into account. Intuitive forecasting methods reflect individual opinions of experts with respect to process development prospects and do not imply development of forecasting models. Intuition and mobilization of professional expertise are the basis for intuitive methods. Such methods are used to analyze processes, which development cannot be mathematically formalized in full or in part, i.e. for which it is difficult to develop an adequate model. It is said in paper [6] that such methods include expert evaluation methods, methods of historical analogies, and methods of foresight by model. To date use of expert systems, including those ones with fuzzy logic [15], is widespread.

Intuitive methods of forecasting are described in detail in paper [16].

In paper [9] forecasting models are divided into statistical and structural models.

In statistical models the functional dependence is set analytically between future and actual values of the time series, as well as external factors.

Statistical models include the following groups:

- Regression models.
- Autoregressive models.
- Models of exponential smoothing.

In structural models the functional dependence is set structurally between future and actual values of the time series, as well as external factors. Structural models include the following groups:

- Neural network models.

- Models based on Markov chains.
- Models based on classification-regression trees.

It is worth mentioning that special forecasting models are sometimes used for highly specialized tasks. For example, models based on differential equations [8] are used for the problem of forecasting the sugar level in human blood. Or, for example, forecasting of the traffic flow - an actual problem faced by megapolises over the last few years. Hydrodynamic models are used [17] for this problem. Also, for example, a model based on nonlinear cells influenced by an external field, which has an internal state that varies with time under the influence of this field [18], is used for the problem of natural phenomena forecasting such as flood, for example. Similar models are developed and applied for special processes and systems.

### 1.1 Regression models

There are multiple problems that require studying the relationship and finding a connection between two or more variables. Regression analysis [19] is used to solve such problems. Today, regression is actively used for various problems, including management and forecasting. The purpose of regression analysis is to determine the dependence between the source variable and multiple external factors (regressors). In these conditions it is allowed to determine regression coefficients by the least square method [19] or the maximum likelihood method [20].

Linear regression model. Linear regression is the simplest version of the regression model. The basis for the model is an assumption that there is a discrete external factor  $X(t)$  that influences the process under study  $Z(t)$ , while the process and the external factor have a linear connection between them. Such a forecasting model is described by the following equation

$$t_{ож} = \frac{3t_{\min} + 2t_{\max}}{5} Z(t) = \alpha_0 + \alpha_1 X(t) + \varepsilon_t \quad (1.1)$$

Where  $\alpha_0$  and  $\alpha_1$  are regression coefficients;  $\varepsilon_t$  is the model error. To obtain predicted values  $Z(t)$  at time point  $t$ , it is required to have the value of  $X(t)$  at the same time point  $t$ , which is rarely met in practice.

Multiple regression model. In practice process  $Z(t)$  is influenced by a number of discrete external factors  $X_1(t), \dots, X_s(t)$ . In this case the forecasting model is given by the following formula

$$t_{ож} = \frac{3t_{\min} + 2t_{\max}}{5} Z(t) = \alpha_0 + \alpha_1 X_1(t) + \alpha_2 X_2(t) + \dots + \alpha_s X_s(t) + \varepsilon_t \quad (1.2)$$

This model has a significant drawback. To calculate the future value of process  $Z(t)$ , it is required to know future values of all factors  $X_1(t), \dots, X_s(t)$ , which is rarely met in practice.

The basis for the nonlinear regression model is an assumption that there is a known function describing the relationship between the source process  $Z(t)$  and the external factor  $X(t)$

$$t_{ож} = \frac{3t_{\min} + 2t_{\max}}{5} Z(t) = F(X(t), A) \quad (1.3)$$

It is required to determine the parameters of function  $A$  within the framework of forecasting model development. For example, we can assume that

$$t_{ож} = \frac{3t_{\min} + 2t_{\max}}{5} Z(t) = \alpha_1 \cos(X(t)) + \alpha_0 \quad (1.4)$$

To develop a model, it is enough to determine parameters  $A = [\alpha_1, \alpha_0]$ . It is almost impossible to find in practice such processes, for which the form of the functional dependence between process  $Z(t)$  and external factor  $X(t)$  is known in advance. Therefore, nonlinear regression models are rarely used.

The model of group accounting of arguments (MGAA) was developed by A.G. Ivakhtenko. [21]. The model is given by the following formula

$$Z(t) = \alpha_0 + \sum_{i=1}^s a_i X_i(t) + \sum_{i=1}^s \sum_{j=1}^s a_{i,j} X_i(t) X_j(t) + \quad (1.5)$$

$$\sum_{i=1}^s \sum_{j=1}^s \sum_{k=1}^s a_{i,j,k} X_i(t)X_j(t)X_k(t)+\dots$$

Equation (1.5) is called a support function. To develop various options of models, the support function is used for some or all of the arguments. For example, polynomials with one variable, polynomials with all possible pairs of variables, polynomials with all possible triples of variables, etc. are constructed. The method of regression analysis is used for each model to determine its linear coefficients  $a_{i,j,k},\dots$ . A few best models ranging from 2 to 10 are chosen among all models. And the quality of models, for example, is determined by mean-square deviation or another criterion. The process of models searching is stopped, if there is a model among the selected ones, which quality is sufficient to be used for the obtained predicted values. Otherwise, the selected models are used as arguments  $X_1(t) \dots X_s(t)$  for support functions of the next iteration step. In other words, the models that have already been found take part in the formation of more complex ones.

Important advantages of this class of models are their simplicity, flexibility, and also uniformity of their analysis and design [19]. If linear regression models are used, it is possible to get the result of forecasting much faster as compared to the use of other models. In addition, another advantage is transparency of modeling [5], i.e. availability of all intermediate calculations for the analysis. The main disadvantage of nonlinear regression models is complexity in determining the type of functional dependence [14], as well as complexity and labor intensity of determining model parameters. In addition, disadvantages of linear regression models are low adaptability and the lack of the ability to model nonlinear processes [28].

## 1.2 Autoregressive models

The basis for autoregressive models is an assumption that the value of the process  $Z(t)$  linearly depends on a number of previous values of the same process  $Z(t - 1), \dots, Z(t - p)$ .

Autoregressive model of the moving average. The autoregressive (AR) model and the moving average (MA) model [1,5] are mostly used for the analysis of a time series.

According to paper [1], the autoregressive model is extremely useful for describing some time series encountered in practice. In this model the current process value is expressed as a finite linear aggregate of previous values of the process and the pulse, which is called "white noise",

$$t_{ож} = \frac{3t_{min}+2t_{max}}{5}Z(t) = C + \varphi_1Z(t - 1) + \varphi_2Z(t - 2) + \dots + \varphi_pZ(t - p) + \varepsilon_t \quad (1.6)$$

Formula (1.6) describes the process of autoregression of order  $p$ , which is often expressed as AR( $p$ ) in the literature, here  $C$  is a real constant,  $\varphi_1, \dots, \varphi_p$  - coefficients,  $\varepsilon_t$  - the model error. The least square method [19] or the maximum likelihood method [20] are used for the determination of  $\varphi_i$  and  $C$ .

The model of moving average of order  $q$  is worth mentioning, which is of great importance for the description of time series and is often used in conjunction with autoregression. It is described by the following equation

$$t_{ож} = \frac{3t_{min}+2t_{max}}{5}Z(t) = \frac{1}{q}(Z(t - 1) + Z(t - 2) + \dots + Z(t - q) + \varepsilon_t) \quad (1.7)$$

In the literature process (1.7) is often expressed as MA( $q$ ); here  $q$  is the order of moving average,  $\varepsilon_t$  - the forecasting error. In fact, the moving average model is a low-frequency filter. It should be noted that there are simple, weighted, cumulative, exponential moving average models.

According to paper [1], it is often advisable to combine autoregression and the moving average in a single model to achieve greater flexibility in model

adjustment. The general model is expressed as ARMA(p,q). It combines autoregression of filtered values of the process of order  $p$  and the filter in the form of the moving average of order  $q$ . If difference of  $d$ -th order (in practice,  $d$  is to be determined, but in most cases  $d \leq 2$ ) is used as input data instead of time series values, then the model is called autoregression of the integrated moving average. In the literature this model is known as ARIMA(p,d,q) (autoregression integrated moving average).

The development of ARIMA(p,d,q) model is ARIMAX(p,d,q) model, which is described by the following equation [1]

$$t_{ож} = \frac{3t_{\min}+2t_{\max}}{5}Z(t) = AR(p) + a_1X_1(t)+\dots+a_sX_s(t) \quad (1.8)$$

Here  $a_1, \dots, a_s$  - coefficients of external factors  $X_1(t), \dots, X_s(t)$ . In this model process  $Z(t)$  is often the result of MA(q) model, i.e. filtered values of the source process. Then the autoregressive model, into which additional regressors of external factors  $X_1(t), \dots, X_s(t)$  are included, is used to forecast  $Z(t)$ .

The autoregressive model with conditional heteroskedasticity (GARCH) was developed in 1986 by Tim Peter Borreslev and is a model of residuals for AR(p) model [22]. At the first stage AR(p) model (1.10) is determined for the initial time series. Further it is assumed that the model error (1.10)  $\varepsilon_t$  consists of two components

$$\varepsilon_t = \sigma_t \cdot \vartheta_t \quad (1.9)$$

Where  $\sigma_t$  - time-dependent standard deviation;  $\vartheta_t$ - a random variable with normal distribution, an average value equal to 0, and a standard deviation equal to 1. In this case the time-dependent standard deviation is described by the following equation

$$t_{ож} = \frac{3t_{\min}+2t_{\max}}{5}\sigma_t^2 = \beta_0 + \beta_1\varepsilon_{t-1}^2 + \dots + \beta_q\varepsilon_{t-q}^2 + \gamma_1\sigma_{t-1} + \dots + \gamma_p\sigma_{t-p} \quad (1.10)$$



Where  $\beta_0, \dots, \beta_q$  and  $\gamma_0, \dots, \gamma_p$  are coefficients. Equation (1.10) is called the GARCH(p,q) model and it consists of two parameters:  $p$  characterizes the order of autoregression of residual squares;  $q$  is the number of previous estimates of residuals. Very often this model is used by the financial sector to model volatility. There are a number of model modifications called NGARCH, IGARCH, EGARCH, GARCH-M, etc. [22].

The autoregressive distributed lag model (ARDLM) is described rather briefly in the literature. Most often it is mentioned in books on econometrics [23]. Often during processes modeling not only current values of process, but also its lags, i.e. values of the time series preceding the target point of time, influence the variable under study. This model is described by the following equation

$$t_{ож} = \frac{3t_{\min} + 2t_{\max}}{5} Z(t) = \varphi_0 + \varphi_1 Z(t - l - 1) + \dots + \varphi_p Z(t - l - p) + \varepsilon_t \quad (1.11)$$

Where  $\varphi_0, \dots, \varphi_p$  - coefficients,  $l$  - the lag value. Model (1.11) is called ARDLM(p,l) and it is most often used to model economic processes [23].

Simplicity and transparency of modeling, uniformity of analysis and design [1] are advantages of these models. This class of models is one of the most popular today [3], therefore it is easy to find examples of solving problems of forecasting of time series of different subject areas with the use of autoregressive models. Disadvantages of this class of models are as follows: a large number of model parameters, which identification is ambiguous and resource-intensive [4]; low adaptability of models, as well as linearity and, as a consequence, the lack of the ability to model nonlinear processes often encountered in practice [26].

### 1.3 Models of exponential smoothing

Exponential smoothing (ES) models are used to model economic and financial processes [24]. The basis for exponential smoothing is the idea of continuous revision of predicted values as far as actual ones are obtained. As far as observations become outdated, the ES model assigns exponentially decreasing

weights to them. And this means that the latest available observations have a greater impact on the predicted value than earlier ones.

The function of the ES model is as follows

$$\begin{aligned} Z(t) &= S(t) + \varepsilon_t \\ S(t) &= \alpha * Z(t - 1) + (1 - \alpha) * S(t - 1) \end{aligned} \quad (1.12)$$

Where  $\alpha$  - smoothing factor,  $0 < \alpha < 1$ ; the initial conditions are determined as  $S(1) = Z(0)$ . In this model each subsequent smoothed value  $S(t)$  is the weighted average between the previous value of the time series  $Z(t)$  and the previous smoothed value  $S(t - 1)$ .

Holt's model or double exponential smoothing is used to model processes that have trend. In this case the following two components should be considered: trend and level [24]. Level and trend are smoothed separately

$$\begin{aligned} t_{ож} &= \frac{3t_{\min} + 2t_{\max}}{5} Z(t) = S(t) + \varepsilon_t \\ S(t) &= \alpha * Z(t - 1) + (1 - \alpha) * (S(t - 1) - \beta(t - 1)) \\ S(t) &= \gamma * (S(t - 1) - S(t - 2)) - (1 + \gamma) * \beta(t - 1) \end{aligned} \quad (1.13)$$

Where  $\alpha$  - the coefficient of level smoothing as in model (1.16),  $\gamma$  - the coefficient of trend smoothing.

The Holt-Winters model or triple exponential smoothing is used to model processes that have trend and seasonality.

$$t_{ож} = \frac{3t_{\min} + 2t_{\max}}{5} Z(t) = (R(t) + G(t)) * S(t) \quad (1.14)$$

Where  $R(t)$  - smoothed level without seasonality

$$\begin{aligned} t_{ож} &= \frac{3t_{\min} + 2t_{\max}}{5} R(t) = \frac{\alpha * Z(t-1)}{S(t-L)} + (1 + \alpha) * (R(t - 1) + \\ &G(t - 1)) \end{aligned} \quad (1.15)$$

$G(t)$  - smoothed trend

$$\begin{aligned} t_{ож} &= \frac{3t_{\min} + 2t_{\max}}{5} G(t) = \beta * (S(t - 1) - S(t - 2)) + \\ &(1 - \beta) * G(t - 1) \end{aligned} \quad (1.16)$$

a  $S(t)$  - seasonal component

$$t_{ож} = \frac{3t_{\min} + 2t_{\max}}{5}$$
$$S(t) = \frac{\gamma * Z(t - 1)}{S(t - L)} + (1 - \gamma) * S(t - L) \quad (1.17)$$

Value  $L$  is the length of the season of the process under study. This class of models is used more often than others [24] for long-term forecasting.

Uniformity and simplicity of analysis and design are advantages of this class of models. The disadvantage of this class of models is the lack of flexibility [36].