

Восточно-Казахстанский технический университет им. Д. Серикбаева

УДК 004:504.3.054

На правах рукописи

**БУГУБАЕВА АЛИНА ЖАНАТБЕКОВНА**

**Методы и алгоритмы интеллектуального анализа данных для системы  
мониторинга атмосферного загрязнения**

6D070300 – Информационные системы (по отраслям)

Диссертация на соискание степени  
доктора философии (PhD)

Отечественный научный  
консультант  
кандидат технических наук,  
ассоциированный профессор  
С.Ж. Рахметуллина

Зарубежный научный  
консультант  
Prof. dr hab., inż.  
W. Wójcik

Республика Казахстан  
Усть-Каменогорск, 2022

## СОДЕРЖАНИЕ

<b>НОРМАТИВНЫЕ ССЫЛКИ</b> .....	4
<b>ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ</b> .....	5
<b>ВВЕДЕНИЕ</b> .....	6
<b>1 АНАЛИЗ СОСТОЯНИЯ ИССЛЕДОВАНИЙ В ПРЕДМЕТНОЙ ОБЛАСТИ</b> .....	11
1.1 Анализ инструментов цифровой трансформации в области вопросов экологии .....	11
1.2 Обзор существующего математического обеспечения систем мониторинга загрязнения атмосферного воздуха .....	16
1.3 Обоснование выбора инструментария интеллектуального анализа данных экологического мониторинга .....	21
Выводы по первому разделу.....	22
<b>2 ИНТЕЛЛЕКТУАЛЬНОЕ ОБЕСПЕЧЕНИЕ ПРОЦЕССОВ ЦИФРОВОГО УПРАВЛЕНИЯ КАЧЕСТВОМ ЖИЗНИ НАСЕЛЕНИЯ РЕГИОНА</b> .....	24
2.1 Объективные предпосылки к системной интеллектуализации формальных процессов принятия решений в мультипараметрической среде .....	24
2.2 Атмосферный массоперенос и разбавление загрязнителей.....	27
2.3 Влияние экологических факторов на здоровье и качество жизни населения .....	30
2.4 Использование искусственного интеллекта в экологии и здравоохранении.....	42
2.5 Математическое обеспечение системы интеллектуальной обработки экологической и медицинско-профилактической информации.....	49
Выводы по второму разделу.....	54
<b>3 ИМИТАЦИОННАЯ МОДЕЛЬ КОЛИЧЕСТВЕННОЙ ОЦЕНКИ КАЧЕСТВА ИНСТРУМЕНТАЛЬНОГО КОНТРОЛЯ В СИСТЕМЕ МОНИТОРИНГА ВОЗДУШНОЙ СРЕДЫ ГОРОДА</b> .....	56
3.1 Количественная оценка качества инструментального контроля в системе мониторинга воздушной среды города.....	56
3.2 Интеллектуальная обработка данных.....	63
3.3 Алгоритм предобработки данных для интеллектуального анализа.....	66
Выводы по третьему разделу.....	69
<b>4 ОПИСАНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ МОНИТОРИНГА АТМОСФЕРНОГО ЗАГРЯЗНЕНИЯ</b> .....	70
4.1 Общее описание и концептуальная основа интеллектуальной информационной системы.....	70
4.2 Архитектура интеллектуальной информационной системы .....	71
4.3 Программная реализация результатов исследования.....	73
Выводы по четвертому разделу.....	80
<b>ЗАКЛЮЧЕНИЕ</b> .....	81
<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b> .....	83

<b>ПРИЛОЖЕНИЕ А</b> – Свидетельство об авторском праве .....	90
<b>ПРИЛОЖЕНИЕ Б</b> – Акт внедрения результатов.....	91

## **НОРМАТИВНЫЕ ССЫЛКИ**

В настоящей диссертации использованы ссылки на следующие стандарты:

Инструкция по оформлению диссертации и автореферата. МОН РК, Внешний аттестационный комитет.

Закон Республики Казахстан. О науке: принят 18 февраля 2011 года, №407-IV.

ГОСТ 7.1-2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления.

Национальный проект «Зеленый Казахстан» 2021-2025 годы.

Стандарт ISO 14064-1. Парниковые газы. Часть 1: Требования и руководство по количественному определению и отчетности о выбросах и удалении парниковых газов на уровне организации.

Стандарт ISO 14064-2. Парниковые газы. Часть 2: Технические требования и руководство для проектировщиков по определению количества, мониторингу и отчетности о сокращении эмиссии парниковых газов и удалении превышенного количества.

Стандарт ISO 9001. Системы менеджмента качества.

Стандарт ISO 31000. Риск-менеджмент.

## ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

РК	– Республика Казахстан
НАО	– Некоммерческое акционерное общество
МОН РК	– Министерство образования и науки Республики Казахстан
ИИ	– Искусственный интеллект
ИИС	– Интеллектуальная информационная система
ППР	– Поддержка принятия решений
IT	– Information technology (информационные технологии)
ПДК	– Предельно допустимая концентрация загрязняющего вещества в атмосфере
ПДВ	– Предельно допустимый выброс загрязняющего вещества в атмосферу
СЭС	– Санитарно-эпидемиологическая служба
ГИС	– Географическая информационная система
ISO	– International Organization for Standardization (Международная организация по стандартизации)
СМК	– Система менеджмента качества
ЭС	– Экспертная система
ЭК	– Экологический контроль
API	– Application-programming interface (программный интерфейс приложения, интерфейс прикладного программирования)
НД	– Набор данных
БД	– База данных
СУБД	– Система управления базами данных
БЗ	– База знаний

## ВВЕДЕНИЕ

**Актуальность работы.** Процветание и конкурентоспособность Республики Казахстан на фоне мировых тенденций, технологических инноваций и научных достижений связано с качественной разработкой, внедрением и эффективной реализацией государственных программ, стратегических планов и проектов, охватывающих все сферы человеческой деятельности. Так, например, национальный проект «Зеленый Казахстан» направлен на создание благоприятной среды проживания для населения и улучшение экологической ситуации, в том числе: улучшение качества атмосферного воздуха, эффективное обращение с отходами производства и потребления, эффективное и бережное использование воды, модернизация экологического сознания населения [1].

В условиях быстрого развития информационных технологий наряду с классическим математическим аппаратом моделирования экологических процессов методы интеллектуального анализа данных позволяют генерировать и накапливать новые знания, путем выявления скрытых закономерностей.

Вопросы разработки моделей и методов для систем мониторинга качества атмосферного воздуха освещены в фундаментальных исследованиях таких зарубежных ученых как Берлянд М.Е., Марчук Г.И., Монин А.С., Яглом А.М., Обухов А.М., Пененко В.В., Алоян А.Е. и другие.

Современные научные исследования, связанные с моделированием и оценкой загрязнения атмосферного воздуха описаны в трудах отечественных ученых, таких как Мутанов Г.М., Темирбеков Н.М., Данаев Н.Т., Султангазин У., Бакирбаев Б., Закарин Э.А., Айдосов А., Мадияров М.Н., Абдолдина Ф.Н., Швец О.Я., Ипалакова М.Т., Султангазин У.М., Жумагулов Б.Т. и многие другие. Анализ показал, что в большинстве исследований используется математический аппарат, основанный на дифференциальных уравнениях и методах математической статистики. Кроме того, экологические процессы являются сложными и нелинейными, что усложняет процессы оценки и прогнозирования экологических показателей. Поэтому необходимы эффективные методы, позволяющие генерировать и накапливать новые знания, путем выявления скрытых закономерностей.

Важнейшими критериями качества мониторинга атмосферного воздуха, являются точность и достоверность и для повышения качества необходимо применение интеллектуальных алгоритмов для снижения рисков связанных с ошибками измерений.

Важным блоком в исследованиях, связанных с вопросами экологии, в том числе загрязнения атмосферного воздуха, является вопрос влияния загрязнения на здоровье человека, но поскольку это процесс сложный многопараметрический, то важно построение интегрированной модели, учитывающей наиболее важные факторы влияния.

Таким образом, для формирования комплекса научно обоснованных практических рекомендаций, которые будут использоваться для принятия оптимальных решений в региональной системе управления экологической

обстановкой, необходимы эффективные методы, алгоритмы и инструменты интеллектуальной поддержки на платформе интенсивно развивающихся средств в этой области.

Исследование направлено на получение наиболее полной картины мониторинга состояния объекта, что требует разработки моделей, методов, а также алгоритмов состояния на основе измерения значимых параметров объекта, включая неполные, несовершенные и небрежные данные, для их использования при поддержке принятия решений.

Эти модели могут быть реализованы с помощью систем искусственного интеллекта, аппарата нечеткой логики, нейронных сетей, генетических алгоритмов и их комбинации. Модель нейронной сети может быть проинструментирована путем измерения данных и позволяет суммировать неполные зашумленные данные. Методы нечеткой логики и нейронные сети способны обрабатывать информацию высокого качества, что является их преимуществом.

**Идея диссертационной работы** заключается в повышении качества процесса управления экологической безопасностью в городской среде.

**Цель работы:** разработка методики интеллектуального анализа данных для системы мониторинга атмосферного загрязнения с целью повышения достоверности экологического контроля.

Для достижения указанной цели, в работе сформулированы следующие **основные задачи:**

1. Разработка математического обеспечения системы мониторинга атмосферного загрязнения на основе построения многофакторной модели.
2. Разработка имитационной модели оценки качества инструментального контроля в системе мониторинга воздушной среды города.
3. Разработка имитационного алгоритма оценки рисков инструментального контроля.
4. Разработка нейросетевой модели анализа данных в многопараметрической среде.
5. Разработка архитектуры интеллектуальной информационной системы оценки негативного влияния промышленных выбросов на здоровье населения.

**Объектом исследования** является процесс мониторинга уровня загрязнения воздуха в городской среде обитания.

**Предметом исследования** являются методы и алгоритмы интеллектуального анализа данных загрязнения воздушной среды города.

**Методы исследования.** Методологической базой исследования является системный подход. В работе для решения поставленных задач используется аппарат теории управления системами, методы статистического анализа, интеллектуальный анализ данных, технологии проектирования информационных систем.

**Научная новизна диссертационного исследования** заключается в том, что впервые предложена методика интеллектуального анализа данных в системе мониторинга атмосферного загрязнения, включающая разработку математического и интеллектуального обеспечения информационной системы,

для оценки и прогноза влияния атмосферного загрязнения на здоровье населения.

**Научные положения, выносимые на защиту:**

1. Методика оценки качества здоровья населения на основе построения интегрированной многофакторной модели.
2. Имитационная модель и алгоритм количественной оценки качества инструментального контроля в системе мониторинга воздушной среды города.
3. Архитектура интеллектуальной информационной системы оценки негативного влияния промышленных выбросов на здоровье населения на основе применения нейросетевой технологии.

**Обоснованность и достоверность** научных положений, выводов и рекомендаций обеспечиваются экологическими стандартами, проверенными средами статистических исследований, техническими характеристиками инструментального обеспечения, типовыми положениями научных исследований.

**Практическая ценность работы** обеспечивается разработанным программным продуктом «Интеллектуальная информационная система оценки негативного воздействия промышленных выбросов на здоровье человека (на примере данных г. Усть-Каменогорск)» с удостоверяющими документами интеллектуальной собственности (Приложение А).

Результаты исследования могут быть использованы в системах мониторинга промышленных городов, подверженных негативному влиянию выбросов промышленных предприятий, для повышения эффективности и достоверности мониторинга качества атмосферного воздуха.

**Реализация результатов работы.** Разработанное в рамках выполнения диссертационной работы программное приложение внедрено в производственный процесс ТОО «ПЕТРОЭКОЦЕНТР-Логистики». Акт внедрения прилагается (Приложение Б).

**Апробация работы.** Материалы диссертационного исследования докладывались на научных семинарах школы, на международных и республиканских научно-практических конференциях, таких как: III Международная научно-практическая конференция «Global science and innovations 2018: Central Asia», Нур-Султан, Казахстан; LXVIII International correspondence scientific and practical conference «European research: innovation in science, education and technology» Лондон, Англия. Также материалы диссертационных исследований докладывались на конференциях, материалы которых рецензированы в базе данных Scopus, таких как: «Communications in Computer and Information Science», Усть-Каменогорск, Казахстан; «Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments», Вильга, Польша; «The International Society for Optical Engineering», Вильга, Польша.

**Публикации.** По теме диссертационной работы опубликовано 11 работ, в том числе 2 статьи в журналах, рецензируемых в базе данных Scopus, таких как: Journal of Theoretical and Applied Information Technology, Пакистан (перцентиль по CiteScore 34%); Przegląd Elektrotechniczny (перцентиль по CiteScore 22%), 3



статьи в издании, рекомендованном Комитетом по обеспечению качества в сфере образования и науки МОН РК, таком как Вестник Восточно-Казахстанского технического университета им. Д. Серикбаева, 6 работ в сборниках материалов международных конференций (4 из которых рецензируются в базе данных Scopus).

**Структура и объем диссертационной работы.** Диссертация состоит из введения, 4 разделов, заключения, списка использованных источников из 106 наименования, изложенных на 89 страницах компьютерного текста, включает 43 рисунка, 5 таблиц и 2 приложения.

**Первый раздел** диссертационной работы содержит аналитическое исследование в форме обзора теоретических и практических подходов к решению задач по изучаемой тематике. Указывается, что одной из перспективных и слабо изученных областей применения искусственного интеллекта является экология, и в настоящее время экология приобретает роль интегратора проблем здравоохранения и хозяйственной деятельности человека.

Проводится анализ инструментов цифровой трансформации в области вопросов экологии и развития аппаратно-программной составляющей систем экологического мониторинга, показывающий, что в настоящее время актуальны методы и средства интеллектуальной обработки данных.

В разделе приводится обзор существующего математического обеспечения систем мониторинга загрязнения атмосферного воздуха, показывающий, что в большинстве исследований используется математический аппарат, основанный на дифференциальных уравнениях и методах математической статистики, а также обосновывается необходимость в эффективных методах, позволяющих генерировать и накапливать новые знания, путем выявления скрытых закономерностей.

Указывается, что в системе экологического мониторинга возможно наличие некачественных, неточных, искаженных данных, поэтому существует необходимость использования специальных интеллектуальных методов и алгоритмов, позволяющих учесть неполноту и зашумленность исходных данных, а также привлечь для анализа качественные экспертные оценки.

**Во втором разделе** проведен анализ объективных предпосылок к системной интеллектуализации формальных процессов принятия решений в многопараметрической среде, выявлено, что разрабатываемая система в обязательном порядке должна содержать блок количественного оценивания рисков в форме мониторинга.

В результате приведенного анализа выявлены главные ограничения и недостатки приведенных методов и предложен подход для реализации процесса интеллектуальной обработки информации. Идея подхода состоит в использовании не статичных коэффициентов корреляции, а коэффициентов и функции автокорреляции с автоматически изменяемой начальной точкой в векторе данных.

В результате статистического эксперимента выявлены следующие значимые факторы, влияющие на продолжительность жизни и качество

здоровья: «загрязнение атмосферы», «состояние здравоохранения», «экономическое благосостояние». Используя методику экспертных оценок и программу автоматизации этой технологии, были найдены весомости каждого из перечисленных факторов здоровья.

Исследован атмосферный массоперенос и разбавление загрязнителей в среде обитания. Разработана методика оценивания и прогнозирования качества здоровья населения, включающая математическое обеспечение системы контроля и принятия решений в процессе экологического мониторинга на основе построения интегрированной многофакторной модели.

**В третьем разделе** разработана имитационная модель и алгоритм количественной оценки качества инструментального контроля в системе мониторинга воздушной среды города. Разработана нейросетевая модель анализа данных для поддержки принятия решений в многопараметрической среде.

Разработана методика интеллектуальной обработки данных измерений загрязнения атмосферного воздуха, включающая имитационную модель оценки рисков ошибок измерений и нейросетевую модель оценки влияния загрязнения воздуха на здоровье населения.

**В четвертом разделе** приводится описание архитектуры интеллектуальной информационной системы оценки негативного воздействия промышленных выбросов на здоровье человека. Архитектура ИИС состоит из трех уровней: уровень загрузки и хранения данных, уровень вычислений и интеллектуальной обработки данных, уровень пользовательского интерфейса ИИС. На каждом уровне архитектуры ИИС происходит формирование интегрированной базы знаний. Подробно описана структура и алгоритм работы программного обеспечения интеллектуальной информационной системы.

**В заключении** диссертационной работы приводится перечень основных результатов и выводов диссертационного исследования, исходя из которых положения, выносимые на защиту, были подтверждены, дается оценка научной новизны и практической значимости исследования.

**Приложения.** Свидетельство о государственной регистрации программ для ЭВМ и Акт внедрения результатов диссертационного исследования.

# **1 АНАЛИЗ СОСТОЯНИЯ ИССЛЕДОВАНИЙ В ПРЕДМЕТНОЙ ОБЛАСТИ**

## **1.1 Анализ инструментов цифровой трансформации в области вопросов экологии**

Бурный прогресс в научно-практической сфере внедрения инновационных подходов в управление бизнес-процессами промышленности, экономики, образования, медицины и экологии в значительной мере определен развитием цифровых технологий. В передовых в этом отношении странах цифровизация декларируется и продвигается специальными программами государственного уровня, в том числе, и в Республике Казахстан [2]. В последнее время цифровизация приобретает новый качественный содержательный образ под названием «цифровая трансформация» [3, 4].

Цифровая трансформация породила глобальную конкуренцию на всех социально-экономических уровнях, преодолевая и разрушая государственные и континентальные границы. В подобной среде важнейшим фактором конкурентного преимущества проявляется необходимость и способность обработки и анализа больших объемов данных [5-7]. Системная динамика изменяющихся потребностей определяет устойчивость и перспективы развития бизнеса.

С точки зрения специалистов предметный признак цифровой трансформации, который ее выделяет из фона равнофункциональных понятий и определений, например, цифровизация – это интеграция платформ и переход от количества к качеству, решение сложных системных проблем во всех без исключения социально-экономических средах.

Отраслевая специфика цифровой трансформации присутствует во всех бизнес-процессах и бизнес-средах экологии, порождая новые, не только информационные, но и правовые связи. Как показали опросы экспертов и оценки НИУ ВШЭ на текущий период потребность и спрос на передовые цифровые технологии составляет десятки миллиардов долларов США с перспективой роста в 15,8 раза к 2030 году.

Отраслевая специфика может приводить к существенному варьированию в подходах и понимании цифровой трансформации. Цифровые активы определяют уровень биржевой стоимости компаний, в том числе, таких цифровых исполинов как Facebook, Google, Microsoft, Apple, Amazon, которые достигли рекордной рыночной капитализации. «Цифровая трансформация» это новое малоизученное динамичное явление, которое только начинает проявлять все свои достоинства в реальной жизни (таблица 1.1).

В некоторых случаях цифровая трансформация рассматривается как программно-управленческий подход на базе цифровых инструментов в некоторой локальной технологической группе, в нескольких отраслях, либо путем технологического форсажа при множестве технологически разнородных решений. Например, в программе развития искусственного интеллекта (ИИ) в Сингапуре предусматривалась разработка данной идеи во многих сферах и отраслях. В США и Европейском Союзе в 2020 году были внедрены различные

программы цифровой трансформации в сфере здравоохранения. Отмечается, что медленно внедряется новая цифровая стратегия в экологии.

Таблица 1.1 – Идентификация цифровой трансформации

Источник	Определения
World Bank Group, 2018	Системная диверсификация бизнес-процессов, в смещении очагов возникновения добавленной стоимости в сферу сквозных цифровых процессов
OECD, 2019	Диверсификация существующих бизнес-процессов на платформе данных и цифровых технологий
ITU, 2018	Внедрение инновационных систем на основе цифровых технологий
UNCTAD, 2019	Радикальная трансформация экономики на базе цифровых инструментов
ITU, 2019	Системный процесс мультимодального изменения составляющих жизненного цикла сервисов государственного и частного сектора
European Commission, 2019	Коренные всеаспектные изменения человеческой жизни на базе цифровых технологий
Примечание – Составлено по источнику [4, с. 82-92]	

В среде «цифровая трансформация» в настоящее время новый виток интенсивного развития получили методы и средства интеллектуальной обработки данных [8]. Главным теоретическим инструментом в обработке данных была прикладная математическая статистика, как средство математического обеспечения автоматизированных информационных систем. Обработывались выборки объемом в десятки и сотни единиц числовой информации. Главным недостатком этих методов является отсутствие возможности генерировать новые знания, хотя бы на уровне гипотез. Под интеллектуальным анализом понимается процесс формирования новых знаний путем выявления скрытых от визуального анализа закономерностей.

Революционные изменения произошли сначала в технологиях полупроводниковой техники с производством средств оперативного и постоянного хранения данных огромных и даже неограниченных размеров и размерностей. Появилась возможность создания, так называемых, баз данных и программных средств СУБД для работы с этими базами. Но результативность работы в части принятия решений не только не возросла, а наоборот даже упала, так как время и стоимости работы с данными резко возросли, а интеллектуальная отдача не появилась. В связи с этими причинами активизировались работы по созданию программных средств с элементами искусственного интеллекта.

Цифровая трансформация включает в себя искусственный интеллект, компоненты робототехники, технологии виртуальной и дополненной реальности, квантовые технологии, технологии больших данных, нейротехнологии и другое. К примеру, цифровая трансформация хозяйствующего субъекта поддерживает производственные, вспомогательные и управленческие бизнес-процессы; в финансах реализует новые способы

взаимодействия между контрагентами; в социальной сфере порождает новые форматы коммуникации.

Одним из широко и часто используемых средств интеллектуальной обработки данных является Data Mining [9-11]. В среде специалистов считается, что Data Mining является наиболее быстро развивающейся технологией обнаружения в «сырых» данных, неочевидных скрытых полезных знаний. Из результатов обзора современных средств интеллектуальной обработки больших данных следуют сделать следующее заключение – Data Mining – это технология обработки «сырых» материалов с целью выявления новых «полезных» актуальных, но скрытых знаний.

При использовании этого интеллектуального продукта возникают достаточно много вопросов у пользователя, например, что такое «полезный». Критерий полезности должен обосновываться, и каким-то образом, конвертироваться к формализованной форме. Некоторые критерии, рекомендуемые к применению, являются известными и типовыми, тогда задача упрощается. В отличие от экспертных методов, Data Mining предлагает методику «объективной» селекции новых знаний. Возникает вопрос, что значит «объективных», чем и как измеряется объективность. В качестве инструментов выбора новых знаний в данной технологии используется методика использования шаблонов «паттернов». В отличие от статистических методов, где связь ищется между заранее подготовленными выборками путем расчета коэффициента корреляции, в Data Mining эти информационные совокупности, определяются и формируются программно, что можно отнести к понятию интеллектуальности [12-15].

В системах искусственного интеллекта в настоящее время довольно широко и успешно внедряются нейросетевые технологии [16-18]. Главной особенностью нейрона является способность к обучению. На вход нейрона поступает совокупность (вектор) данных или объектов. В нейроне формируется сумма «взвешенных» данных. Указанная сумма сравнивается с пороговым значением, если результат сравнения по заданному алгоритму не удовлетворяет условиям задачи, то входной вектор модифицируется. Главный недостаток этой технологии состоит в том, что «вес» входных элементов необходимо оценивать экспертно.

В искусственном интеллекте используется также, так называемый, «метод ближайшего соседа». Данный алгоритм «измеряет» сходство между объектами, по метрике расстояния. Этот прием используется также в экспертных оценках при измерении степени совпадения мнений экспертов по коэффициенту корреляции Спирмена [19].

Известен алгоритм под названием «деревья решений». Эта технология является в настоящее время достаточно востребованной в задачах интеллектуального анализа данных. Основа метода состоит в иерархической многоуровневой детализации. Действия подобных алгоритмов базируется на регрессионном и корреляционном анализе.

Известны «генетические алгоритмы», которые считаются универсальными и используются в оптимизационных задачах, как при

многокритериальном случае, так и при многокритериальном поиске оптимума в среде больших данных.

Одна из самых важных, и даже повседневных задач, как в жизни отдельного субъекта, так и жизненном цикле хозяйствующего субъекта – это предсказание или прогнозирование будущего. Эта функция реализуется двумя методами: методом экспертного анализа, и формальными методами с привлечением компьютерной техники [20]. Для формализации процессов прогноза привлекаются математические инструменты [21].

Многочисленными исследованиями, например, американским метеорологом Эдвардом Лоренцом, доказана крайняя чувствительность формальных методов к начальным данным, что играет огромную роль в системах искусственного интеллекта [22]. Эта чувствительность, как правило, ведет к хаосу, что тем не менее успешно используется в нейронных сетях. Лоренц не мог понять, почему беспрецедентно бурное развитие компьютерных технологий не привело к повышению качества прогнозов, то есть, как было много лет тому назад, так и сегодня, точный прогноз ограничивается 2-3 днями. Эта проблема актуальна в среде активного использования больших данных. Причина этого явления состоит в динамическом хаосе, что крайне важно при работе с большими данными.

Система Лоренца имеет конечный горизонт прогнозирования. Прогнозируемые процессы имеют свойство и тенденцию расходиться с тем большей скоростью, чем дальше от точки начала прогнозирования ставится прогноз. Это определяется, так называемым, показателем Ляпунова [23]. От этого показателя зависит интервал времени прогнозирования, что крайне важно в задачах временной кластеризации данных. Нелинейная динамика лишает многих иллюзий в области искусственного интеллекта, и создает многие ограничения в функциональности данных систем [24, 25].

Однако, нелинейная динамика, наряду с многими проблемами, открывает и большие возможности особенно в нейросетевой технологии. Так, нелинейная динамика позволяет минимизировать количество данных во входном векторе нейрона для качественного мониторинга жизнедеятельности системы. В данном случае появилась возможность перейти от системного анализа к системному синтезу и извлекать из массы переменных наиболее информативные. Таким образом, прогноз из области чистой науки перемещается в область практических технологий.

Как уже отмечалось, одной из перспективных и слабо изученных областей применения искусственного интеллекта является экология. В настоящее время экология приобретает роль интегратора проблем здравоохранения, энергетики, особенно с появлением, так называемой «зеленой энергетики», сельским хозяйством в связи появлением экологически чистых продуктов, автомобильным транспортом, являющимся одним из главных загрязнителей атмосферы [26].

Поистине, неограниченные горизонты открываются для принципов интеллектуального анализа больших данных в технологии, которая стала называться «точное земледелие» [27]. Применение искусственного интеллекта

в этой области порождает огромное количество задач теоретического и практического плана и вопросов, связанных со спецификой применения цифровых технологий, особенно интеллектуального уровня.

Одна из ключевых задач состоит в разработке формальных методов для количественного оценивания рисков принятия решений в условиях неопределенности агентов управления в современной земледелии. В системном плане, разработка только формальных подходов, кроме необоснованных затрат, к другому результату не приведет. Цифровая системность в сложных проектах, к которым следует отнести и точное земледелие, ставит перед необходимостью в комплексе с разработкой и внедрением программного, технического, организационно-методического обеспечения, идти на ряд других непредусмотренных специфических мероприятий, особенно в части применения слабо изученной междисциплинарной области искусственного интеллекта и больших данных.

Интеллектуальное обеспечение на базе идеологии «цифровой трансформации» должно сыграть роль «интегратора оцифрованных человеческих знаний». Выражение «точное земледелие» появилось сравнительно недавно. В данном выражении первое слово-понятие «точное», концентрирует внимание на метрологических аспектах предметной области – «земледелие». Само понятие «точное», как строгая количественная оценка в метрологии отсутствует, так как это качественная оценка такая же как: «хороший», «длинный», «стройный» и так далее. В литературных источниках и интернет-ресурсах в разделе земледелия такое понятие часто носит узкий технологический смысл, но декларируется авторами, как общий метрологический показатель в русле функциональной цели решаемой задачи.

Понятие точности имеет множество определений, особенно, когда речь идет о моделировании. Если речь идет о геометрии технического объекта, то задача оценки точности упрощается, так как многие моделируемые параметры (размеры) измеримы, но если речь идет о естественных или искусственных, например, экологических, антропогенного происхождения, то истина в принципе недоступна исследователю и неизмерима. В подобных случаях исследуют косвенные параметры, которые называются диагностическими индикаторами.

Довольно часто используется оценка «адекватность» модели, например, в регрессионном, факторном, имитационном или нечетком моделировании. В данном случае адекватность рассматривается, как соответствие результата выборочным статистическим данным, по которым строится модель, но не как степень приближения к истине. Поэтому любая модель замещает оригинал лишь в строго ограниченных рамках размерности, времени и имеющихся ресурсов. Ключевое значение при этом имеют цель и точка зрения исследователя. Большие данные – это числовая статистическая модель некоторой среды или пространства. Выборка – это часть больших данных. Если сконцентрировать внимание на некоторые частные показатели, то говорят о генеральной совокупности. Поэтому для одного объекта может быть построено бесчисленное множество моделей.

Важной составляющей цифровой трансформации в области экологического мониторинга является развитие современных автоматизированных систем мониторинга. Аппаратно-программная составляющая этих систем позволяет в режиме реального времени передавать данные о состоянии наблюдаемого объекта, что приводит к накоплению большого массива ретроспективных данных об объекте наблюдения. Так РГП «Казгидромет» имеет 140 экологических постов по 45 населенным пунктам страны, большая часть которых оснащена автоматизированными газоанализаторами по наблюдаемым загрязняющим веществам. Интерактивная карта отражает фактическое состояние атмосферного воздуха: пыль, концентрацию загрязняющих веществ, степень загрязнения воздуха и прогноз загрязнения. Наличие таких автоматизированных систем мониторинга формирует большие массивы данных мониторинга, что создает все предпосылки к использованию методов и средств интеллектуальной обработки данных.

Таким образом, анализ инструментов цифровой трансформации в области вопросов экологии и развитие аппаратно-программной составляющей систем экологического мониторинга показал, что в настоящее время актуальны методы и средства интеллектуальной обработки данных.

Исследование и выбор наиболее эффективной технологии интеллектуальной обработки данных для повышения качества процесса управления экологической безопасностью в городской среде является целью настоящей работы и излагается в последующих материалах.

## **1.2 Обзор существующего математического обеспечения систем мониторинга загрязнении атмосферного воздуха**

Любые экологические системы состоят из множества динамических процессов, взаимосвязанных компонентов биологического, климатического, химического, геологического, физического или социального характера. При проведении анализа состояния экологических систем или проблем, которые связаны с такими системами, возникает сложность обработки данных, полученных из различных источников и имеющих разный формат и размерность. На рисунке 1.1 приведена структурная модель современной экологии, которая отражает многообразие и разноплановость экологической информации.

Такие задачи экологического мониторинга, как сбор экологической информации, ее последующая обработка и интерпретация, контроль предельно-допустимых параметров, решаются достаточно успешно. Методики сбора и хранения информации рассматриваются в работах таких отечественных и зарубежных авторов как Дж. Парк, С. Маккей, В.В. Смородин, Е.В. Волкова [28, 29]. Например, статистическая обработка экологических данных освещается в классических источниках (Берлянд М.Е., Вуколов Э.А., Цефель П. и других [30-32]).





Рисунок 1.1 – Структурная модель современной экологии

Во многих системах экологического мониторинга анализ экологической информации происходит с применением отдельных не связанных между собой методов или ограничивается сравнением измерений экологических параметров с предельно-допустимыми нормами. Проведение параметрического контроля не подразумевает привлечения специальных математических методов в данной области, к примеру, такие задачи решаются в работах Застрогина Ю.Ф., Сулова А.Г. и других [33-35]. В зарубежной и отечественной литературе большое место также уделяется этой проблематике [36-40]. Например, детальная процедура анализа риска для региона была представлена в работе [41, 42].

При рассмотрении систем для экологического мониторинга, которые основаны на контроле абиотических показателей, выделяют два основных подхода при организации процессов мониторинга по основным загрязнителям окружающей среды [43]:

- системы мониторинга основываются на точном контроле предельно допустимых выбросов (ПДВ) с моделированием распространения этих загрязнений и учетом метеорологических условий;
- система базируется на точном контроле предельно допустимых концентраций (ПДК).

Для расчета параметров загрязнения атмосферы обычно используются дифференциальные уравнения [44, 45] Это связано с тем фактом, что такие системы уравнений позволяют получать формулы, описывающие поведение шлейфа загрязнения, при правильно выбранных граничных условиях. В подходе, который основывается на точном контроле предельно-допустимых

выбросов, должна быть информация об источнике этих выбросов, общем объеме, времени выбросов, полученных ингредиентах концентраций, метеорологических условиях. Если известны первоначальные данные, то имеется информация для принятия решений и управления в данной области. Для проведения прогноза возможных изменений окружающей среды, в основном используются модели, которые получают при помощи решения задач турбулентной диффузии.

Поскольку источники вредных выбросов в Республике Казахстан и других странах СНГ не вынесены за черту города и вокруг них часто располагаются жилые массивы, как в городе Усть-Каменогорск, то при ветре в направлении жилых комплексов с предприятий переносится большое количество шлейфа с вредными веществами и создаются зоны с повышенной концентрацией загрязнителей. Теория распространения примесей при штиле была изложена в работе [46]. Численное решение задач с учетом интегрированного уравнения турбулентной диффузии получено в работах М.Е. Берлянда, и Е.Л. Генриховича [47]. Авторы развивали теорию пограничного слоя в области с криволинейной границей. Работа в данном направлении была продолжена авторами Пененко А.В., Рахметуллиной С.Ж.

Исследования А.А. Айдосова являются интересными для нефтегазоконденсатдобывающих регионов с точки зрения анализа процессов распространения вредных выбросов в пограничном слое атмосферы [48, 49]. Для учета таких специфических условий были разработаны математические модели распространения примесей от источника при переменном профиле скорости ветра [50]. Теоретико-методологической основой разработки и применения рационального метода для оценки геометрии распространения потоков и учета концентрации вредных веществ в окружающей среде могут служить именно рассмотренные в данной работе модели.

Дано теоретическое обоснование построения границ циркуляционных зон, а также механизма загрязнения с использованием молекулярной и турбулентной диффузий и метода Эйлера в работе В.Т. Серкина, который исследовал распространение вредных примесей вблизи границы циркуляционной зоны [51].

Одним из современных способов объединения разнотипных данных и представления целостной картины, полученных результатов, то есть формирования единого информационного пространства мониторинга атмосферного воздуха и почвенных загрязнений, а также визуализации с помощью ГИС-технологий рассмотрены в работах отечественных авторов Мутанова Г.М., Темирбекова Н.М., Мадиярова М.Н., Абдолдиной Ф.Н., Швец О.Я., Ипалаковой М.Т., Султангазина У.М., Жумагулова Б.Т. [52-54]. Жижимов О.Л., Протасов К.Т., Ноженкова Л.Ф. и другие являются выдающимися представителями сибирского отделения Российской Академии наук, которые занимаются исследованиями в данной области [55-57].

Большой популярностью пользуется другое направление расчета загрязнения атмосферы, которое связано со статистическими методами [58]. Например, для расчета и мониторинга загрязнения воздуха в городах в целом и

отдельных промышленных районах, ведутся различные статистические проработки на основе анализа материалов наблюдений, обусловленных действием многих источников выбросов. В большинстве из них допускается, что за рассматриваемый период, к которому относится поступивший материал, а также за срок расчета сама величина выбросов и местоположение источников практически не изменяется. Если рассматривать большое число источников, то допускается, что увеличение объема выбросов от какой-то части из них достигается уменьшением объема выбросов от оставшейся части. В связи с этим, увеличение среднего и общего загрязнения воздуха города связывается с изменением метеоусловий или синоптической ситуации.

Предложенный Ивахненко А.Г. [59] для прогноза загрязнения воздуха метод группового учета аргументов можно считать, как математическое обобщение и графическую регрессию, в котором возможно заменить субъективное проведение кривых на построение по методу наименьших квадратов объективных зависимостей.

Использование обобщенных характеристик концентрации вредных веществ в целом по городу позволяет справиться с трудностями установления корреляционных зависимостей между загрязнениями воздуха и соответствующими предикторами. Такие составленные характеристики, основанные на большом количестве наблюдений в нескольких пунктах за несколько сроков, будут достовернее, чем единичные данные о концентрации выбросов и будут менее подвержены случайным колебаниям. Они показывают вклад загрязнения воздуха, источниками, которые преобладают, фоновую концентрацию в целом по городу, которая создается вследствие объединения многих выбросов. Этот метод Вавилова Н.Г. и другие [58, с. 27] использовали для проведения статистического анализа данных загрязнения воздуха в городе.

Практический интерес для ведущих ученых не только в области геоэкологии, экологии и медицины, но и информационных технологий представляет вопрос о влиянии промышленных выбросов в атмосферу на окружающую среду и, особенно, на здоровье населения. Первые работы по использованию этой методики применялись в США и относятся к 80-м годам XX века. Классическое исследование основано на оценке риска воздействия различных вредных и опасных химических веществ на здоровье населения [60, 61].

С того времени было изучено и разработано большое количество методов и моделей для определения типов риска и причин необходимости такой оценки [62-65]. Примером может служить методология анализа рисков, разработанная в США в Центре безопасности химической промышленности, более подробно описанная в учебнике РГУ нефти и газа им. И.М. Губкина, где в методе оценки рассмотрен расчет концентраций вредных веществ в атмосфере от промышленных предприятий.

Среди отечественных авторов, занимающихся исследованием причинно-следственной связи в системе «среда-здоровье» можно отметить Аманжол И.А., Аманбекову А.У. и других. Авторы комплексно исследовали проблемы общественного здоровья и здравоохранения – заболеваемости населения с

учетом международных рекомендаций, предложена методология оценки риска. Исследования, направленные на выявление и оценку связи загрязнения воздуха и здоровье населения рассмотрены недостаточно полно. Данные научные труды также основаны на медицинских статистических исследованиях [66].

Всемирная Организация Здравоохранения применяет метод расчёта величины нескольких последствий для здоровья, связанных с воздействием наиболее релевантных загрязнителей воздуха в данной популяции, основанный на статистическом логлинейном анализе.

РГП «Казгидромет» использует прогностическую модель качества атмосферного воздуха на базе модели SILAM (System for Integrated modeLling of Atmospheric coMposition), которая представляет собой интерактивную карту по мониторингу качества атмосферного воздуха [67]. На рисунке 1.2 представлена интерактивная карта по мониторингу качества атмосферного воздуха на базе модели SILAM.

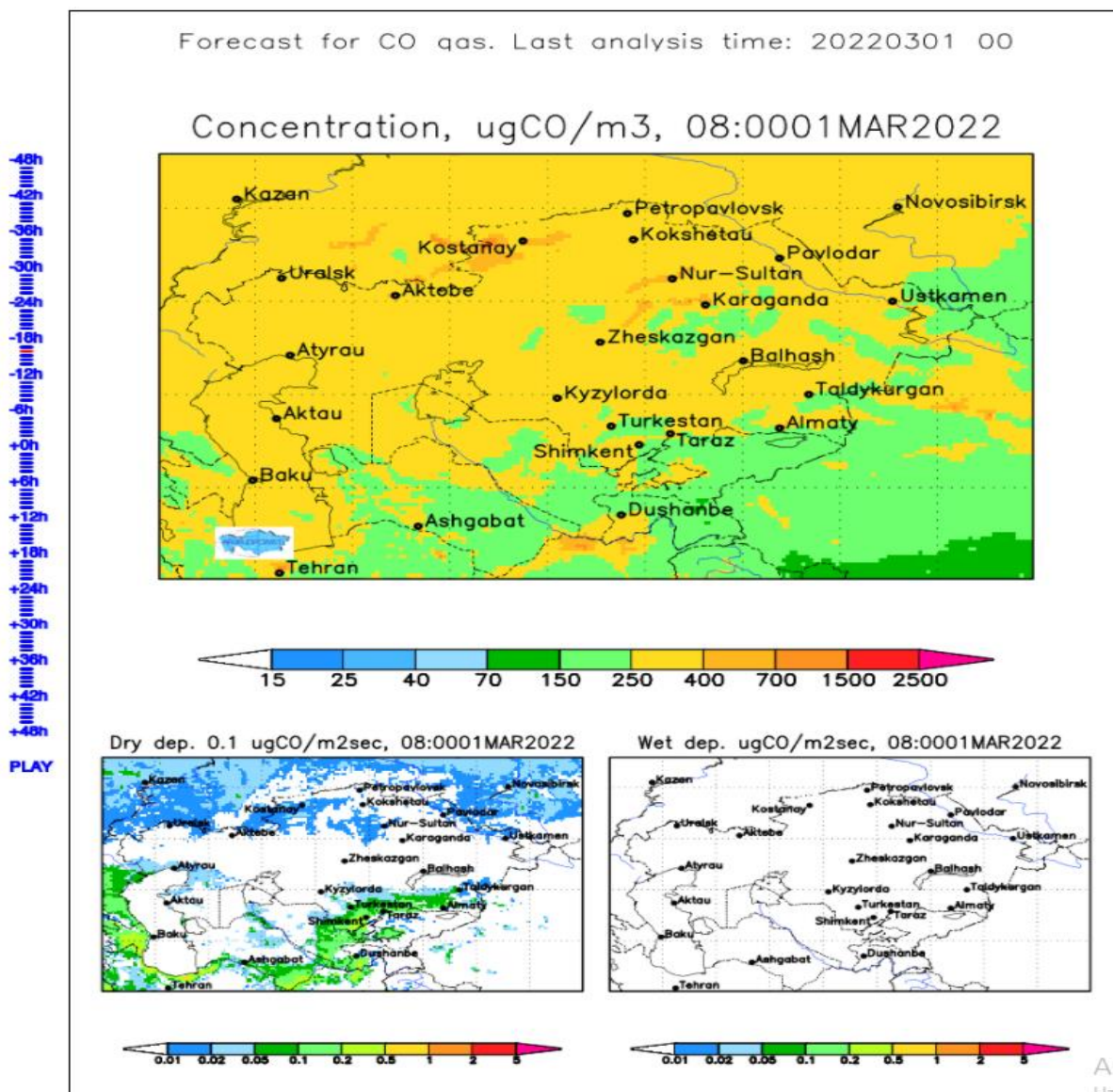


Рисунок 1.2 – Интерактивную карту по мониторингу качества атмосферного воздуха на базе модели SILAM

В качестве методов интеллектуального анализа оценки качества воздуха в системе SILAM используются: алгоритм k-means, агрегация, геопространственные закономерности и репрезентативность эпизодов.

Таким образом, обзор существующего математического обеспечения систем мониторинга загрязнения атмосферного воздуха показывает, что в большинстве исследований используется математический аппарат, основанный на дифференциальных уравнениях и методах математической статистики. Необходимы эффективные методы, позволяющие генерировать и накапливать новые знания, путем выявления скрытых закономерностей.

### **1.3 Обоснование выбора инструментария интеллектуального анализа данных экологического мониторинга**

Для проведения наиболее полного исследования и разработки рабочих алгоритмов, которые будут рассчитывать значения состояния процессов на основе измерения основных параметров исследуемого объекта, а также с учетом неполных, неточных и нечетких данных необходимы новые технические решения. Реализовать такие решения можно с помощью систем искусственного интеллекта: методов на основе нечетких условий, нейросетей, генетических алгоритмов, и их различных комбинаций [68]. Модели нейронных сетей могут быть обучены на данных измерений и дают возможность обобщать не полные и зашумленные данные, модели нечеткого вывода могут обрабатывать данные по лингвистической информации, а также нейросети тоже обладают указанными преимуществами.

В последние годы из-за сложности и нелинейности процессов, протекающих в экологической области, внимание экологов и ученых, занимающихся прогнозированием и обработкой экологических данных, привлекают искусственные нейронные сети [69-71]. Поскольку нейронные сети все чаще используются для моделирования, аппроксимации функций, прогнозирования, классификации и интеллектуального анализа данных, используем этот аппарат для интеллектуального анализа в диссертационной работе.

Из литературного анализа схемотехники нейронов можно составить приближенное кибернетическое, представление в виде структурно-функционального узла, используя понятие «черный ящик», на вход которого поступает некоторый вектор данных  $x_1, x_2, \dots, x_n$ . Каждому элементу вектора присваивается «вес», найденный при помощи метода экспертных оценок. Ядро этого узла выполняет функцию преобразователя информации, чаще роль «свертки» с принятием решения.

Принятие решение состоит в сравнении взвешенной суммы с пороговым значением  $\alpha_0$ . Здесь же происходит логическая обработка, если агрегированный результат не превышает «порог», то следует одно выходное действие – решение, в противном случае следует альтернативное решение. Математически это можно выразить следующим отношением:

$$Y = \begin{cases} 1, \text{ если } \sum_{i=1}^n \alpha_i X_i \leq \alpha_0; \\ 0, \text{ если } \sum_{i=1}^n \alpha_i X_i > \alpha_0. \end{cases} \quad (1.1)$$

Нейроны в реальных случаях объединяются в нейронные сети с целью реализации технологии обучения без учителя. Этот принцип реализует свойство кластеризации, без которого невозможно создание интеллектуальных систем обработки информации.

Большие перспективы в нейросетевой технологии открываются с использованием нечеткой логики, где элемент кластера может принадлежать одновременно и другим кластерам, что крайне усложняет задачу. В нечетком подходе приоритетность отнесения элемента к тому или иному кластеру решает «степень принадлежности», которая является главным свойством в нечетких алгоритмах. Однако возникает проблема точности исходной информации. Повышение качества кластеризации многие авторы связывают с использованием новых идей из области нелинейных систем и теории хаоса [72, 73].

Традиционным методам обработки информации свойственны многие недостатки, в частности, использование понятия и свойства средних величин, что приводит к потере информации. Однако, практика эксплуатации сложных систем показывает, что малозначимый элемент при определенных условиях может сыграть решающую роль [74].

Современные средства моделирования и вычислительные мощности компьютеров позволяют строить работоспособные модели нейронных сетей и прогнозировать возможные результаты. Это подтверждают исследования и примеры реализации L. Li, W. Chu, J., Langford, D. Liben-Nowell, J. Kleinberg и других [75, 76].

Интегрирующей системой, объединяющей социальную систему, систему управления экологической средой, здравоохранение и органы административного управления регионом является, так называемая, интеллектуальная информационная система, в которой интеллектуальная обработка в связи с ее универсальностью и выявлением скрытых закономерностей существенно повысит качество принимаемых решений.

### **Выводы по первому разделу**

Одной из перспективных и слабо изученных областей применения искусственного интеллекта является экология. В настоящее время экология приобретает роль интегратора проблем здравоохранения и хозяйственной деятельности человека.

Анализ инструментов цифровой трансформации в области вопросов экологии и развитие аппаратно-программной составляющей систем экологического мониторинга показал, что в настоящее время актуальны методы и средства интеллектуальной обработки данных.

Обзор существующего математического обеспечения систем мониторинга загрязнения атмосферного воздуха показывает, что в большинстве исследований используется математический аппарат, основанный на дифференциальных уравнениях и методах математической статистики. Необходимы эффективные методы, позволяющие генерировать и накапливать новые знания, путем выявления скрытых закономерностей.

На практике, в системе экологического мониторинга возможно наличие некачественных, неточных, искаженных данных, поэтому существует необходимость использования специальных интеллектуальных методов и алгоритмов, позволяющих учесть неполноту и зашумленность исходных данных, а также привлечь для анализа качественные экспертные оценки.

Современные средства моделирования и вычислительные мощности компьютеров позволяют строить работоспособные модели нейронных сетей и прогнозировать возможные результаты.

## **2 ИНТЕЛЛЕКТУАЛЬНОЕ ОБЕСПЕЧЕНИЕ ПРОЦЕССОВ ЦИФРОВОГО УПРАВЛЕНИЯ КАЧЕСТВОМ ЖИЗНИ НАСЕЛЕНИЯ РЕГИОНА**

### **2.1 Объективные предпосылки к системной интеллектуализации формальных процессов принятия решений в мультипараметрической среде**

В настоящее время существует много методик и работ по анализу влияния вредных веществ на заболеваемость, однако в основу всех работ положено изучение влияния конкретных химических (органических и неорганических) веществ на организм человека, на его отдельные органы, иммунную систему и на динамику изменения структуры клеток в зависимости от дозы тех или иных вредных веществ. При этом, несмотря на серьезные достижения, полученные по расчетам концентраций вредных веществ в городском секторе и области, степень влияния их на появление заболеваний в реальном времени достаточно слабо изучена.

В диссертации проведены исследования распространения шлейфа загрязнения, использующие теоретические и практические данные по состоянию атмосферы, определен и нанесен на карту города сектор негативного воздействия выбросов вредных веществ [77, 78]. Данные динамики загрязнения обрабатываются и автоматически картографируются с целью передачи их в соответствующие организации.

В отличие от традиционных методик расчета влияния выбросов на заболевания в диссертации решалась задача, по данным о пролеченных больных выявлялась корреляция с выбросами, что можно использовать при прогнозе возникновения заболеваний в зависимости от работы промышленных предприятий [79].

Проводимые в настоящее время исследования в области экологии на практике используют в качестве автоматизации процедур принятия решений так называемую «лоскутную автоматизацию», то есть, производятся локальные несложные технические расчеты по автономным небольшим программам, которые обеспечивают и ускоряют некоторую локальную вычислительную работу. В предлагаемой диссертационной работе предприняты попытки разработки информационной системы, основанной на современных интегрированных информационно-технологических подходах с использованием принципов искусственного интеллекта.

Для исследования были выбраны процессы, происходящие в экологической среде и системе здравоохранения на административно-территориальном образовании Восточного Казахстана, где антропогенные проблемы представляют серьезную опасность для населения. Решение обозначенной проблемы, как показали аналитические исследования в первом разделе работы, представляется возможным с глубоким и обширным анализом больших массивов информации. Указанные данные находятся в учреждениях здравоохранения и экологии Восточно-Казахстанской области.



Как показала практика, применение традиционной диалоговой точечной запросной методики и технологии требует огромного количества ручных процедур и машинного времени. Использование средств OLAP для этих целей позволило частично решить некоторые проблемы и повысить производительность труда пользователя системы, путем информационных разрезов исследуемых массивов данных. Но выявление оптимальных траекторий информационных сечений по-прежнему определяется субъектом путем непроизводительного предварительного анализа ситуации. Задача крайне усложняется, если данные являются нечеткими и мониторингового характера, что превращает работу с данными в рутинный долговременный процесс, сопровождаемый частыми ошибками в моменты принятия решений, особенно в оптимизационных задачах.

Все эти факторы предопределили цель и задачи исследований. В задачи входило на первом этапе определиться с подходом в формализации принятия решения и методов интеллектуализации системных управленческих этапов. Одними из многообещающих, и доказавших свою результативность в подобных задачах инструментах являются: нейросетевые технологии, теория нечетких множеств, искусственный интеллект, генетические алгоритмы и принципы синергетики, которые приводят к неожиданным положительным, порой даже трудно объяснимым, результатам.

Нейросетевые алгоритмы и модели как технологии привлекательны тем, что в них заложен процесс обучения свойственный человеку. Обучение построено на фильтрации данных, что крайне важно в условиях зашумления полезной информации, используются лингвистические описания и выводы, характерные для человека, группирование данных по выявленным самим алгоритмом признакам, имеется возможность построения на этих участках прогнозных путей развития процесса [80, 81].

Одной из серьезных ошибок, которая приводила к снижению качества формальных математических моделей и не показывала тех результатов, которые ожидалась на стадии постановки задачи, являлись линейные и детерминированные подходы, приводившие к тривиальным конечным результатам. Вместе с тем, многочисленными исследованиями уже доказано, что в реальности все процессы нелинейны и взаимосвязаны, имея неочевидные связи, и статистически неоднородны, что приводит к неадекватности моделей к «историческим» статистическим данным и предпосылкам [82].

Особенно часто и явно указанные причины, и аргументы отражаются на явлениях и процессах в сфере экологии и медицине, где решающую роль играет трудно прогнозируемый антропогенный фактор. Качество и однородность информации в значительной мере зависит от внешних факторных причин искусственного происхождения, что снижает возможность «обучения» системы. Поэтому в разрабатываемой системе необходимо предусмотреть алгоритм систематической проверки на однородность вновь появляющейся информации. В качестве критериев однородности наиболее приемлемыми является критерий Фишера [83].

В августе 2006 года технический комитет TC 207 «Environmental management» Международной организации по стандартизации (ISO) разработал ряд международных стандартов, которые могут помочь организациям уменьшить их воздействие на окружающую среду, включая серию стандартов ISO 14064 на выбросы парниковых газов.

Стандарт ISO 14064-1 «Парниковые газы. Часть 1: Требования и руководство по количественному определению и отчетности о выбросах и удалении парниковых газов на уровне организации» рассматривает некоторые ключевые вопросы количественной оценки парниковых газов, связь между ними и возобновляемыми источниками энергии включает в себя более стандартизированные рамки отчетности.

Стандарт ISO 14064-2 «Парниковые газы. Часть 2: Технические требования и руководство для проектировщиков по определению количества, мониторингу и отчетности о сокращении эмиссии парниковых газов и удалении превышенного количества» имеет более широкую область распространения, чтобы быть применимым к углеродным кредитам и модифицированным технологиям, а также к внутренним и технологическим проектам по парниковым газам организаций. Данный стандарт устанавливает требования и критерии относительно экологических самодеклараций, описывает общую методологию оценки и верификации экологических самодеклараций, конкретные методы оценки и проверки экологических заявлений относительно экологических аспектов продукции, ее компонентов или упаковки.

Стандарт ISO 9001 устанавливает критерии системы менеджмента качества. Модифицированный вариант ISO 9001:2015 включает некоторые фундаментальные изменения. В новой процессной редакции ISO формула «системы» приобретает качественно новый смысл, и будет звучать как «система – это совокупность взаимосвязанных процессов (не элементов), объединенных единой целью».

Решающим отличием в ISO 9001:2015 по сравнению с предшествующими версиями ISO 9001:2000 и ISO 9001:2008 также следует считать то, что «риски больше не присутствуют в стандарте скрыто (в виде предупреждающих действий) и не локализованы отдельными процессами СМК. Риск теперь пронизывает весь стандарт и встроен в целом в систему, становится ее неотъемлемым свойством, а не представляет собой ее часть». Жизненный цикл любого проекта, опирающегося на «риск-менеджмент», теперь является концептуальным доминантом системной динамики процессного подхода в цифровых технологиях: анализа, оценки и принятия решений в условиях неопределенности на всех проектных этапах.

В 2016 был разработан стандарт ISO 31000 под названием «Риск-менеджмент». В составе Риск-менеджмента появились четыре новых документа. Важным фактом и отличием новой редакции является то, что в модифицированном стандарте используется требование к оценке рисков (risk-based thinking) для всей технологической цепочки управления рисками.

Это дополнение особенно актуализирует проблему контроля рисков, что делает необходимым и обязательным количественное измерение рисков, так

как именно на этом этапе происходит «цифровизация» риска. В системе цифрового управления бизнес-процессами требование количественного измерения текущего значения риска теряет свою значимость без процедуры прогнозирования и описания динамики риска, что может быть реализовано путем привлечения математического аппарата и компьютерного моделирования [84-86].

Положения стандарта ISO 31000 «Менеджмент рисков» могут использоваться по отношению ко всем типам формирования риска, вне зависимости от природы их происхождения. Стандарт ISO 31000 может использоваться на предприятии, как в целом, так и в отдельных его структурных звеньях.

Обработка риска может помочь в прогнозе опасного события и даже полностью предотвратить риск или уменьшить его путем локализации и блокирования опасного источника риска, когда выявлены источники неопределенности агентов управления и статистические свойства параметров исследуемых агентов. Авторы делят условия, в которых принимаются решения, на три типа. Так, Вентцель Е.С. предлагает делить условия принятия решений на условия полной определенности, условия статистической неопределенности и условия полной неопределенности [87].

Таким образом, из приведенного анализа следует очень важный вывод – разрабатываемая система в обязательном порядке должна содержать блок количественного оценивания рисков в форме мониторинга. Для реализации указанной задачи возникает новое функциональное требование в постоянной модификации кластеризованной структуры увеличивающегося массива информации, мониторинговая проверка на однородность данных, модификация статистических характеристик кластерных выборочных образований, количественный расчет рисков, построение динамики рисков и текущие выводы и прогнозы по локальным системным агентам и общая интегрированная оценка динамической ситуации.

## **2.2 Атмосферный массоперенос и разбавление загрязнителей**

Метеорологические условия. Воздействие метеорологии на загрязнение воздуха определяют атмосферный перенос и распространение загрязняющих веществ, и поэтому сильно влияет на их воздействие. Атмосфера обеспечивает диспергирование и разбавление выбрасываемого материала за счет процессов турбулентного перемешивания, которые происходят между поверхностным и верхним слоями воздуха. При определенных условиях вероятность рассеивания ограничена, и может произойти ситуация, в которой загрязняющие вещества будут улавливаться и скапливаться у земли.

В настоящее время универсальной модели нет. Однако в результате тщательных эмпирических исследований за последние несколько десятилетий основные механизмы атмосферной диффузии и воздухопереноса описаны и формализованы.

Механизмы атмосферной диффузии. Различают выброс загрязняющих веществ и их объем транспорта от источников сброса. Такое различие

проистекает из природы турбулентного движения атмосферы, которое можно рассматривать как серию колебаний, которые накладываются на средний поток. Перемешивание облака загрязняющих веществ вызвано смешиванием облака с его локальной атмосферной средой. Процесс смешивания или диффузия в первую очередь связана с атмосферной турбулентностью, которая может быть просто рассмотрена как случайное движение воздушных масс или «водоворотов».

Эффект смешивания молекулярной диффузией относительно незначителен. Вихревые движения происходят на очень широком диапазоне размеров и частот, а также степени растворения облака или шлейфа загрязняющих веществ очень сильно зависит от масштаба турбулентных колебаний в сравнении с размером облака загрязняющих веществ. В основном есть три случая:

1. Когда размер облака превышает масштаб турбулентности, вихревая дисперсия происходит, и облако турбулентно перемешивается с окружающей средой.

2. Когда размер облака меньше масштаба турбулентности, вихревая диффузия ограничена, и облако переносится через среду с относительно небольшим разбавлением.

3. Когда размер облака и масштаб турбулентности одного порядка, вихревые движения просто искажают облако без значительного изменения концентрации.

Атмосферная турбулентность охватывает широкий спектр частот и масштабов, поэтому все описанные выше механизмы диспергирования и переноса происходят одновременно.

Следуя за облаком по мере его развития от небольшого концентрированного объема до большого разбавленного объема, обычно наблюдается продолжающееся усиление механизма разбавления по мере того, как постепенно увеличивается эффективность завихрения, смешанное движение способствует образованию более мелких водоворотов, которые еще разбавляют расширяющееся облако загрязняющих веществ.

Чтобы объяснить и предвидеть неблагоприятные метеорологические условия, необходимо учитывать, как вертикальное перемешивание, так и горизонтальную диффузию, и перенос загрязняющих веществ.

Вертикальное перемешивание загрязнителей. Смешивание и разбавление атмосферных загрязнителей в вертикальном направлении может происходить при очень разных скоростях, в зависимости от эффектов плавучести, которые вызывают вертикальный термический поток. Эти скорости тесно связаны с изменениями вертикальной температуры, или градиентов, которые могут быть классифицированы как стабильные, нейтральные или нестабильные, по порядку возрастающего потенциала диффузии загрязняющих веществ.

Стабильность атмосферы демонстрирует большие суточные колебания, циклические колебания, которые зависят от солнечного излучения и различий в характеристиках теплопередачи в соответствии с особенностями местности. Также может возникать вертикальное перемешивание в нижних пограничных

слоях атмосферы. Наибольший интерес представляет очень устойчивая конфигурация, инверсия, в которой температура фактически увеличивается с высотой. Инверсионные слои могут быть настолько эффективными в уменьшении перемешивания, что, в большинстве случаев, вертикальная диффузия загрязняющих веществ за счет вихря колебания практически полностью подавляются.

Горизонтальный перенос и диффузия загрязнителей. Горизонтальный перенос и распространение загрязняющих веществ происходит в результате горизонтальных потоков воздуха, которые характеризуются следующими закономерностями:

1. Глобальные метеорологические закономерности. На них влияет региональная климатология, определяющая распределение атмосферного давления.

2. Эффект Кориолиса из-за вращения Земли.

3. Поверхностное трение.

4. Механически создаваемая горизонтальная турбулентность.

5. Местные потоки.

Чистый эффект этих факторов заключается в том, что горизонтальные потоки вблизи поверхности Земли никогда не бывают устойчивыми, и их скорость и направление сильно различаются. Загрязнение распространяется с постоянно меняющимся ветром. Высокочастотные колебания рассеивают загрязнитель, и относительно устойчивый «средний» поток отвечает за объем переноса загрязняющих веществ.

Для непрерывного точечного источника концентрация загрязняющего вещества в окружающей среде обратно пропорциональна скорости ветра. Роль горизонтального переносчика особенно высока в периоды низких, сильных условий инверсии, когда вертикальное перемешивание загрязняющих веществ эффективно поглощается на своем уровне. Загрязняющие вещества становятся разбавленными субинверсионным воздухом, но отсутствие перемешивания движения в самом инверсионном слое предотвращает вертикальное разбавление. Если условия останутся неизменными, концентрация загрязняющих веществ будет продолжать расти в этом фиксированном объеме. Введение горизонтального потока обеспечивает разбавление воздуха в районе источника через его ветреную границу и в то же время удаляет равный объем загрязненного воздуха с подветренной стороны. Этот чистый перенос воздуха через зону источника прекращает неограниченное накопление загрязняющего вещества над районом с источником выброса.

Формирование инверсии. Эффект рассеивания усиливается частым возникновением одновременного подавления горизонтального перемешивания. Это означает, что выбросы из дымовых труб или вентиляционных отверстий будут быстро рассеиваться и, если выбросы не обладают достаточным тепловым «давлением», загрязнитель не проникает вертикально в устойчивые пласты атмосферы. Следовательно, есть возможность улавливания загрязнителей и их скопления в приземном слое. Загрязняющие вещества чаще встречаются в составе «коктейлей», чем индивидуально. Например,

Великобританский и Американский смог 1950-х и 1960-х годов был вызван в основном прямым сжиганием угля и привел к образованию твердых частиц и газов, таких как сера диоксид, оксиды азота, сероводород и оксид углерода. Исследования токсических эффектов такого смога, как правило, направлены на борьбу с ним в целом, а не отдельными элементами коктейля. Законодательства о чистом воздухе в обеих странах запретили сжигание угля низкого качества. Одной из наиболее важных особенностей исследований загрязнения и здоровья, по-видимому, является относительным недостатком недавних работ по токсическому действию отдельных загрязнителей.

Большинство исследований воздействия загрязнителей на здоровье, как правило, сосредоточено вокруг загрязнителей, которые регулярно встречаются на промышленных предприятиях.

Исследования, включающие связь загрязнения окружающей среды и здоровье, можно классифицировать по следующим признакам:

- по методам (будь то лабораторные исследования на животных, человеке, клинические исследования или популяционные исследования);
- по заболеваниям (бронхит, рак легких и так далее) в зависимости от рассматриваемого исхода (заболеваемость, смертность, использование медицинских услуг или отсутствие на работе);
- по характеристикам группы изучения (национальность, возраст, пол, раса и так далее);
- по загрязняющим веществам (диоксид серы, углерод монооксид и так далее) или продолжительность их действия (повседневный, острый или хронический).

Из-за законодательства о чистом воздухе в некоторых странах значительные токсические эффекты загрязняющих веществ обычно связаны со случайными или разовыми событиями.

Таким образом, из приведенного анализа следует, что большинство исследований воздействия загрязнителей на здоровье, как правило, сосредоточено вокруг загрязнителей, которые регулярно встречаются на промышленных предприятиях. Исследования, проводимые в диссертационной работе основаны на связи загрязнения окружающей среды и здоровье, путем измерения загрязняющих веществ.

### **2.3 Влияние экологических факторов на здоровье и качество жизни населения**

Загрязнение воздуха представляет собой проблему, вызывающую растущий национальный и международный интерес. Общественная осведомленность и озабоченность продолжает расти, был сделан ряд крайних предложений по регулированию выбросов загрязняющих веществ потому, что многие политики неправильно понимают основные физические и экономические аспекты борьбы с загрязнением воздуха.

Многие загрязнители являются естественными составляющими воздуха. Растения, животные и природная деятельность вызовет некоторое загрязнение даже без человека и его технологий. Животные выдыхают углекислый газ.

Гниющая растительность производит метан и разлагающиеся животные производят аммиак. Вулканическая деятельность выделяет двуокись серы и ветер приводит к образованию взвешенных частиц. В каком-то смысле любое вещество, обычно не присутствующее в окружающей среде или концентрация, превышающая нормальную, может считаться загрязнителем. Однако вещество обычно считается загрязнителем, если оно отрицательно влияет на окружающую среду, изменяя скорость роста видов, вмешивается в пищевую цепочку, токсично или оказывает неблагоприятное воздействие на здоровье, комфорт, удобства или стоимость имущества людей. Загрязняющие вещества попадают в окружающую среду в значительных количествах как промышленные отходы, случайные выбросы, бытовые отходы и сточные воды, или как побочный продукт производственных процессов или другой деятельности человека.

Термин «загрязнение воздуха» ограничен ситуацией, в которой внешняя среда – атмосфера – содержит вещества в концентрациях, вредных для человека и окружающей среды. Загрязнение воздуха обычно представляет собой смесь диоксида серы, твердых частиц, оксида азота, оксида углерода, углеводородов, озона и аммиака. Объединение их в единую категорию «загрязнение воздуха» игнорирует их различные свойства и скрывает различия в степени ущерба, который каждый из них может нанести. Например, одни загрязнители вызывают атмосферную дымку и ухудшают видимость, другие обесцвечивают предметы, а некоторые могут стать причиной болезни. Загрязнение воздуха используется как обобщающий термин для этих соединений, но на протяжении всего исследования будет уделено внимание изучению особенно пагубного воздействия каждого загрязнителя, когда это возможно. В данном исследовании ограничимся связью между загрязнением воздуха и здоровьем населения.

Загрязнение окружающей среды может повлиять на здоровье любым из следующих способов [88]:

- интенсивное воздействие токсичных материалов или агентов, вызывающих немедленную острую реакцию и последствия для здоровья;
- низкий уровень воздействия токсичных материалов или агентов, которые могут вызвать острые или хронические заболевания спустя долгое время после контакта;
- воздействия, которые могут вызвать генетические изменения;
- пониженная сопротивляемость инфекциям;
- вызывать субклиническое раздражение, неприятные ощущения или дискомфорт;
- способствуют обострению имеющегося заболевания;
- создавать условия, несовместимые или унижающие для достижения физического, психического и социального благополучия.

Трудно четко различить последствия загрязнения воздуха от различных источников, и важно учитывать влияние смесей веществ. Эффект от таких смесей может быть просто аддитивным, в большинстве случаев часто обнаруживают, что они синергичны, хотя иногда они могут быть антагонистическими.

Диоксид серы ( $\text{SO}_2$ ) и другие соединения серы раздражает дыхательные пути и увеличивает частоту респираторных заболеваний. Это может вызвать отек легких или голосовой щели и вызвать респираторный паралич. Это может вызвать повышенную кислотность мочи и нарушение обоняния при повышенной утомляемости. Фарингит, бронхит, тонзиллит, простуда и язвы гортани также связаны с  $\text{SO}_2$ . Диоксид серы менее распространен в атмосфере, чем сероводород ( $\text{H}_2\text{S}$ ), который может оказывать раздражающее действие на глаза, парализует обоняние и вызывает мгновенное бессознательное состояние. Это может повлиять на нервную систему, вызывая головную боль, головокружение, возбуждение и потерю зрения.

В источниках изучались острые эффекты вдыхания диоксида серы ( $\text{SO}_2$ ) в группе тщательно подобранных здоровых подростков (3 девочки-подростка и 5 мальчиков-подростков в возрасте 12-14 лет). У здоровых испытуемых наблюдались небольшие статистически значимые изменения, соответствующие с «бронхоконструкцией» после воздействия диоксида серы ( $\text{SO}_2$ ). После воздействия в состоянии покоя до  $1,0 \text{ ppm SO}_2 + 1 \text{ мг/м}^3$  капельного аэрозоля хлорида натрия ( $\text{NaCl}$ ), у здоровых подростков отмечалось небольшое (3%), но значительное снижение объема форсированного выдоха за одну секунду (на английском – FEV1 или на русском – ОФВ1) Существенных изменений не наблюдалось после воздействия только  $\text{NaCl}$  или  $\text{SO}_2$  в состоянии покоя. Воздействие  $\text{SO}_2$  во время умеренных упражнений дало большие изменения. После воздействия либо смеси  $\text{SO}_2 + \text{NaCl}$  или только  $1,0 \text{ ppm SO}_2$  снижение ОФВ1 было немного больше (6%) и продолжительнее по сравнению с наблюдаемыми в состоянии покоя. Маленькое снижение (4%) ОФВ1 наблюдалось после воздействия  $\text{NaCl}$  в течение одного упражнения, но изменение не было статистически значимым. Изменения, наблюдаемые в группе здоровых подростков, были незначительными по сравнению с подростками, страдающими астмой (67%). Таким образом, астматические подростки гораздо более чувствительны к воздействию вдыхаемого  $\text{SO}_2$ , чем здоровые подростки.

Авторы [89] изучали двадцать четыре добровольца-астматика, подвергавшихся воздействию 0, 0,3 и 0,6 ppm диоксида серы ( $\text{SO}_2$ ) в очищенном фоновом воздухе при каждой из трех температур:  $21^\circ\text{C}$ ,  $7^\circ\text{C}$  и  $6^\circ\text{C}$  в камере с контролируемой средой. Относительная влажность была примерно 80%. Экспозиции включали 5-минутные периоды тяжелых упражнений плюс короткие периоды разминки и расслабления. Сопротивление дыхательных путей, объем вдыхаемого газа и симптомы измерялись непосредственно до и после воздействия. Для группы увеличение концентрации  $\text{SO}_2$  и снижение температуры ( $6^\circ\text{C}$ ) были связаны со статистически значимым неблагоприятным влиянием на сопротивление дыхательных путей и респираторными симптомами. Эффекты  $\text{SO}_2$  и температуры обычно оказывались комбинированными аддитивным или менее аддитивным способом, было мало свидетельств синергизма. Образцы ответов людей были разными: несколько предполагали синергизм, но другие предположили смягчающее действие



холода на бронхоконструктивную реакцию к  $SO_2$ . Подобные исследования проводились и в других странах.

Угарный газ вызывает два вида воздействия на здоровье: краткосрочное и долгосрочное. Увеличение концентрации CO повышает уровень карбоксигемоглобина CoHb в крови. Гемоглобин имеет большую чувствительность к CO, примерно в 240 раз больше, чем у кислорода, и он поглощает CO, даже когда его концентрация очень низкая. Первичный результат избытка CoHb в крови – кислородное голодание, оно приводит к снижению внимания, сильной головной боли, слабости, головокружению, тошноте, бронхитическому кашлю с мокротой, рвоте, снижению сенсорной дискриминации, потере остроты зрения, и в итоге к смерти. Длительное воздействие CO может вызвать, бронхит и снижение дыхательной функции.

Концентрация оксида углерода (CO), а также оксидов азота (NO, NO<sub>2</sub>) была выше в домах, использующих газ, чем в домах, использующих электричество для приготовления еды. Объем форсированного выдоха за 1 секунду (FEV1) был использован как показатель вентиляционной функции легких. Использование кухонного вытяжного вентилятора было связано с низкой функцией легких. Большая доля женщин при низком уровне FEV1 для приготовления пищи использовала газ, но эта разница не была статистически существенной. Рентгенологические данные также показали, что частота заболеваний легких была выше в загрязненной зоне.

Используя методологию, которая напрямую рассматривает характер данных временного ряда, в ряде исследований изучались проблемы со здоровьем под влиянием ежедневных уровней загрязнения воздуха с использованием таких событий, как смертность и заболеваемость. Углеводороды оказывают прямое и косвенное воздействие на здоровье. Они очень канцерогенны. При их распаде образуется большое количество CO. Они мешают образованию красных кровяных телец в костном мозге, вызывают лейкоз. Углеводороды вызывают бессознательное состояние, гипоксию и судороги, они раздражают глаза и дыхательные пути, что отражено в работе [89, p. 11176332-11176335]. В данной работе приводятся методы повышения эффективности медико-социальной реабилитации инвалидов с заболеваниями органов дыхания.

Оксиды азота напрямую влияют на здоровье, а также трансформируются в сочетании с другими загрязнителями воздуха и превращаются во вредные вторичные загрязнители. Оксид азота (NO) имеет эффекты, аналогичные эффектам CO, вызывая кислородное голодание. NO<sub>2</sub> раздражает респираторные ткани и вызывает боли, отек легких и смерть. Это также снижает сопротивляемость организма к инфекциям.

При низких концентрациях в воздухе пары аммиака (NH<sub>3</sub>) раздражают глаза, нос и горло и воспаляют верхние дыхательные пути. При высоких концентрациях вызывает судорожный кашель, сильное раздражение глаз, носа и горла, ожог дыхательных путей и чувство удушья. При высоких концентрациях может быть смертельным за считанные минуты. Считается, что газообразный аммиак в выдыхаемом воздухе нейтрализует кислотные

загрязнения воздуха и тем самым потенциально смягчает легочные эффекты загрязнения. Для исследования этого загрязнителя на здоровье был проведен специальный эксперимент. Четырнадцать субъектов (от 21 до 45 лет) выполнили один или несколько из следующих гигиенических маневров: (а) кислотное полоскание полости рта ( $\text{pH} < 2,5$ ); (б) чистка зубов с последующим полосканием ротовой полости кислотой; (с) чистка зубов с последующим ополаскиванием дистиллированной водой; и (д) ополаскивание дистиллированной водой.

Первоначальные уровни аммиака варьировались от 120 до 1280 частей на миллиард (ppm). Кислое полоскание привело к немедленному сокращению выдыхаемого аммиака на 90% у всех субъектов, и возврат к 50% от исходного уровня произошел в течение 1 часа. Результат от чистки зубов или только дистиллированная вода не имели значения. Был сделан вывод, что кислотное полоскание полости рта – эффективный метод снижения содержания аммиака в дыхательных путях, но его следует повторять для поддержания постоянного состояния с низким содержанием аммиака в выдыхаемом воздухе во время исследований воздействия кислотных аэрозолей.

Для характеристики состояния здоровья населения автором работы предлагается использовать понятия «качество жизни» и «благополучие населения», под которыми можно понимать совокупность социальных, экономических и медицинских показателей. В последние двадцать лет по наблюдениям экспертов городской СЭС и комитета общественной экспертизы наблюдается заметное ухудшение здоровья населения в Восточно-Казахстанской области. По их данным особенно тяжелое положение складывается в районе железнодорожной станции «Защита», а в микрорайоне комбината шелковых тканей содержание бора у детей превышает норму примерно в 40 раз.

Проведенная Комитетом общественной экспертизы оценка воздействия загрязнения окружающего воздуха на уровень заболеваемости людей показала, что загрязнение атмосферного воздуха повлияло на общую заболеваемость в целом, и ее вклад составляет примерно 37% у детей и около 10% заболеваний у взрослого населения. Сравнение результатов исследования, начиная с 2015 года, по анализам крови детей на тяжелые металлы показало значительное повышение содержания металлов в крови детей, в первую очередь – свинца.

Следовательно, ухудшение состояния окружающей среды в ВКО и связанное с ней снижение показателей здоровья жителей, проживающих в этом районе, требуют проведения комплексных исследований в этом регионе. Рост онкологических и других заболеваний – неизбежная расплата за бесхозяйственное отношение к природе и объясняется экологическими законами, согласно которым, антропогенные трансформации атмосферы неизбежно ведут к негативным последствиям и резкому снижению качества здоровья, что требует внедрения новых методологий в управление экологической системой города, а также повышению требований к качеству контроля.

Это взаимодействие нелинейно и, как отмечает лауреат Нобелевской премии И. Пригожин, чаще всего непредсказуемо, что также было отмечено в свое время и Лоренцом. Даже незначительное изменение одного из показателей экологической системы может привести к катастрофическим последствиям. Эта ситуация наблюдается в Восточно-Казахстанской области.

Был проделан анализ состояния рождаемости и смертности в регионе Восточно-Казахстанской области (таблица 2.1). С 2006 года рождаемость в Восточно-Казахстанской области снижалась по 2016 год включительно с 32261 человек до 16306 человек.

Таблица 2.1 – Родившиеся, умершие и естественный прирост населения ВКО (человек)

Год	Родившиеся	Умершие	Естественный прирост убыль(-)	Коэффициент рождаемости на 1000 человек	Коэффициент смертности на 1000 человек
2006	32261	16492	15769	18,1	9,3
2007	30224	16832	13392	16,9	9,4
2008	27595	19739	7856	15,5	11,1
2009	26629	20618	6011	15,2	11,8
2010	23132	21574	1558	13,5	12,6
2011	20993	21295	-302	12,5	12,7
2012	18687	20695	-2008	11,3	12,5
2013	17808	19639	-1831	11,0	12,1
2014	16870	18530	-1660	11,1	12,2
2015	17025	18822	-1797	11,3	12,5
2016	16306	18059	-1753	10,9	12,1
2017	17013	18130	-1117	11,5	12,3
2018	18288	18644	-356	12,5	12,8
2019	19397	18398	999	13,4	12,7
2020	19638	19055	583	13,67	13,26

На основании таблицы 2.1 построена следующая диаграмма динамики рождаемости и смертности по ВКО за 2005-2020 гг. (рисунок 2.1).

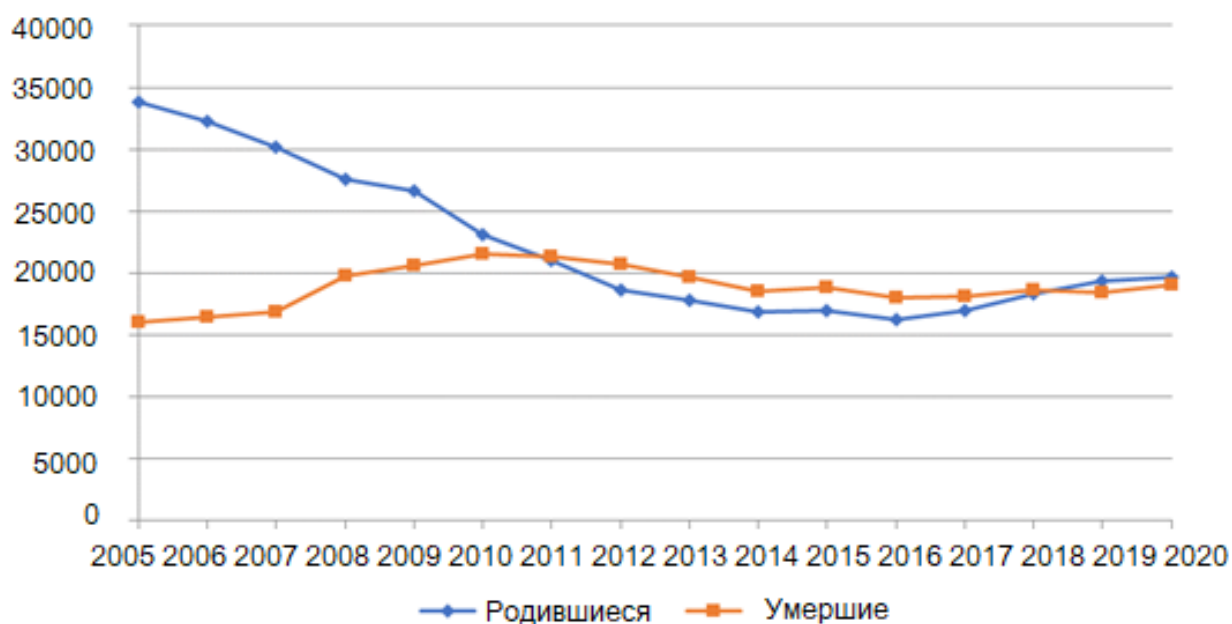


Рисунок 2.1 – Диаграмма динамики рождаемости и смертности по ВКО

Начиная с 2015 года, естественная убыль населения снижалась и к 2018 году составила 356 человек в год. Причем, в городской местности данный показатель, начиная с 2015 года, снижался менее интенсивно, а в сельской местности начиная с 2018 года увеличивался и имел положительное значение.

Таким образом, в 2019 году естественный прирост населения Восточно-Казахстанской области составлял 999 человек ежегодно, а в 2020 году составил соответственно 583 человека (таблица 2.2).

Таблица 2.2 – Показатели естественного прироста (убыли) населения ВКО (человек)

Год	Естественный прирост (убыль)	Город	Село
2015	1797	2876	1079
2016	1753	2739	986
2017	1117	2064	947
2018	356	1663	1307
2019	999	729	1728
2020	583	789	1372

Последствия в форме динамики естественного прироста (убыли) населения породились социально-экономическими факторами, которые возникли еще в середине 90-х годов (рисунок 2.2).

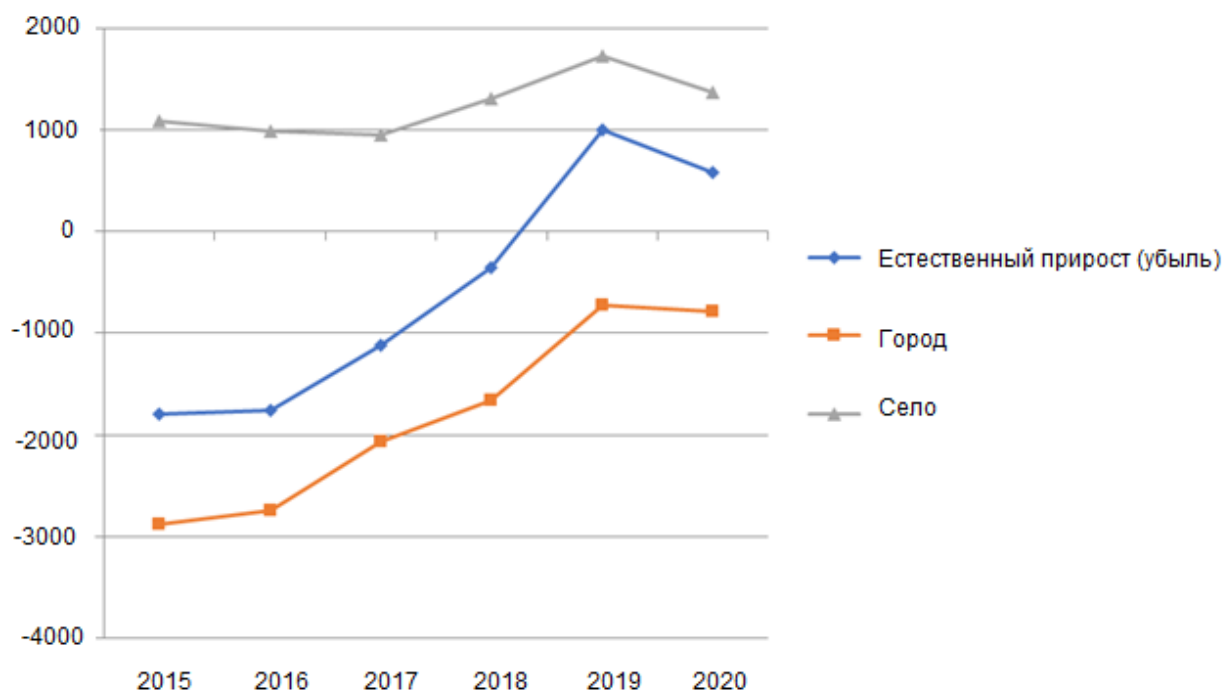


Рисунок 2.2 – Динамика естественного прироста (убыли) населения по ВКО

Коэффициент рождаемости в ВКО возрастает в среднем на 9,9 % по отношению к 2016 году (рисунок 2.3).

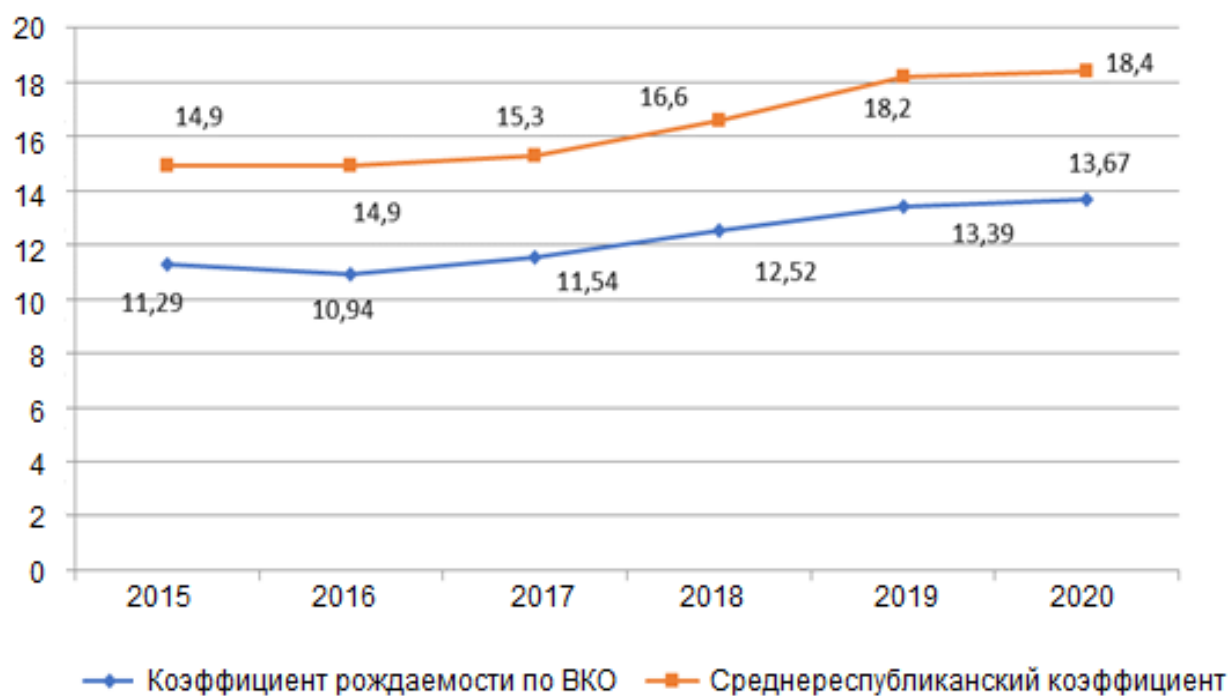


Рисунок 2.3 – Диаграмма коэффициента рождаемости по ВКО и РК

Темпы роста рождаемости в ВКО приведены в таблице 2.3.

Таблица 2.3 – Темп роста рождаемости ВКО

Год	2015	2016	2017	2018	2019	2020
Коэффициент рождаемости	11,29	10,94	11,54	12,52	13,39	13,67
Темп роста, в % к 2010 году		-3,1	2,2	10,8	18,6	21,08

Несмотря на то что, коэффициент рождаемости в ВКО возрастает, его значение ниже среднереспубликанского уровня (рисунок 2.4).

Убыль населения в Восточно-Казахстанской области в 2015 году составляла 1190 человек, на протяжении четырех лет отрицательное значение естественного прироста населения сокращалось и в 2019 положительное сальдо составило 690 человек.

Тем не менее, естественный прирост населения Восточно-Казахстанской области значительно меньше среднереспубликанского значения этого показателя (рисунок 2.5).

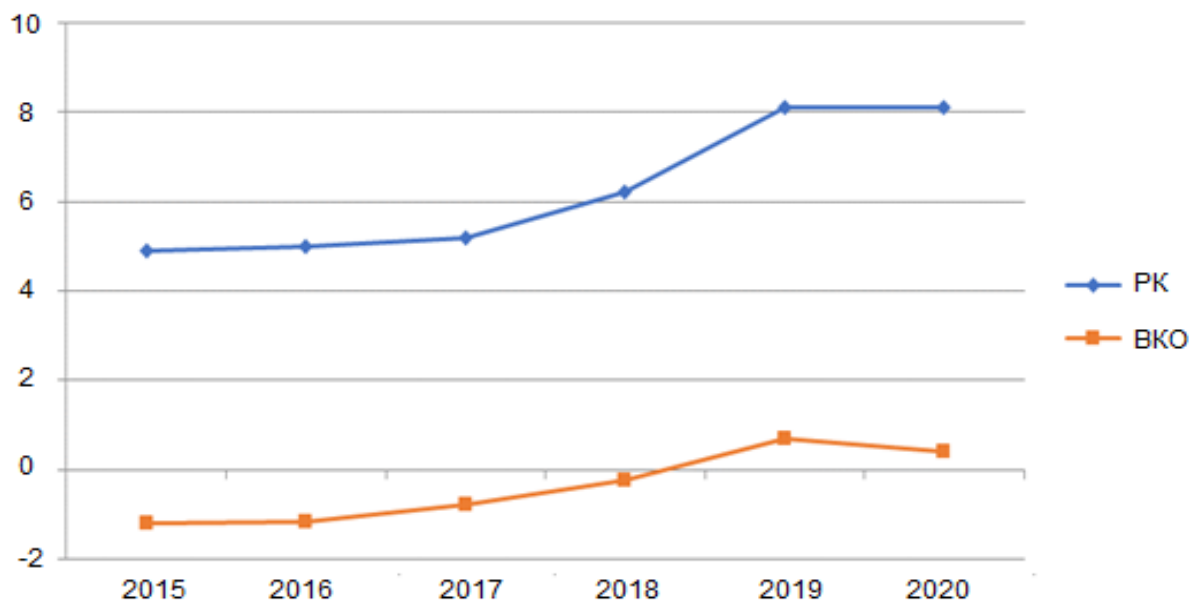


Рисунок 2.4 – Диаграмма динамики естественного прироста (убыли) населения по ВКО и РК (тыс. человек)

У городского населения уровень продолжительности жизни при рождении значительно ниже, чем у сельского населения.

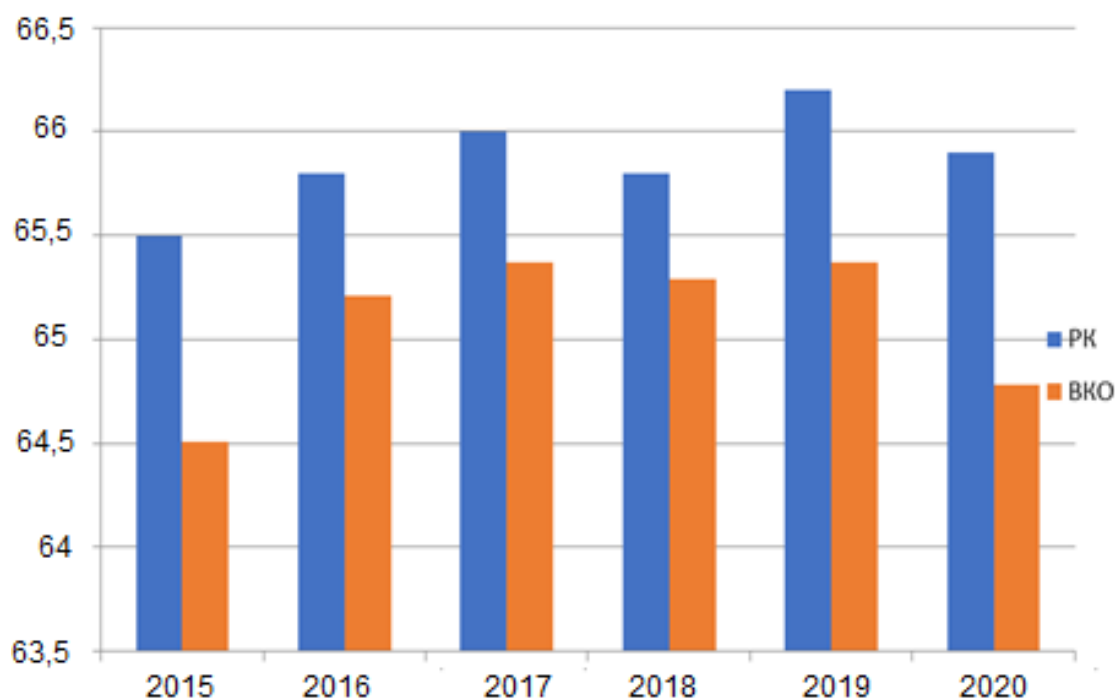


Рисунок 2.5 – Уровень продолжительности жизни при рождении

У сельского населения VKO данный показатель на протяжении пяти лет возрастал и в 2019 году составил 66,87 лет, напротив, у городского населения под воздействием экологической обстановки, множественны заболевания (сердечно-сосудистой системы, органов дыхания, злокачественных образований и так далее) ожидаемая продолжительность жизни при рождении снижается и в 2020 году этот показатель достиг своей критической отметки и составил 64 года (рисунок 2.6).

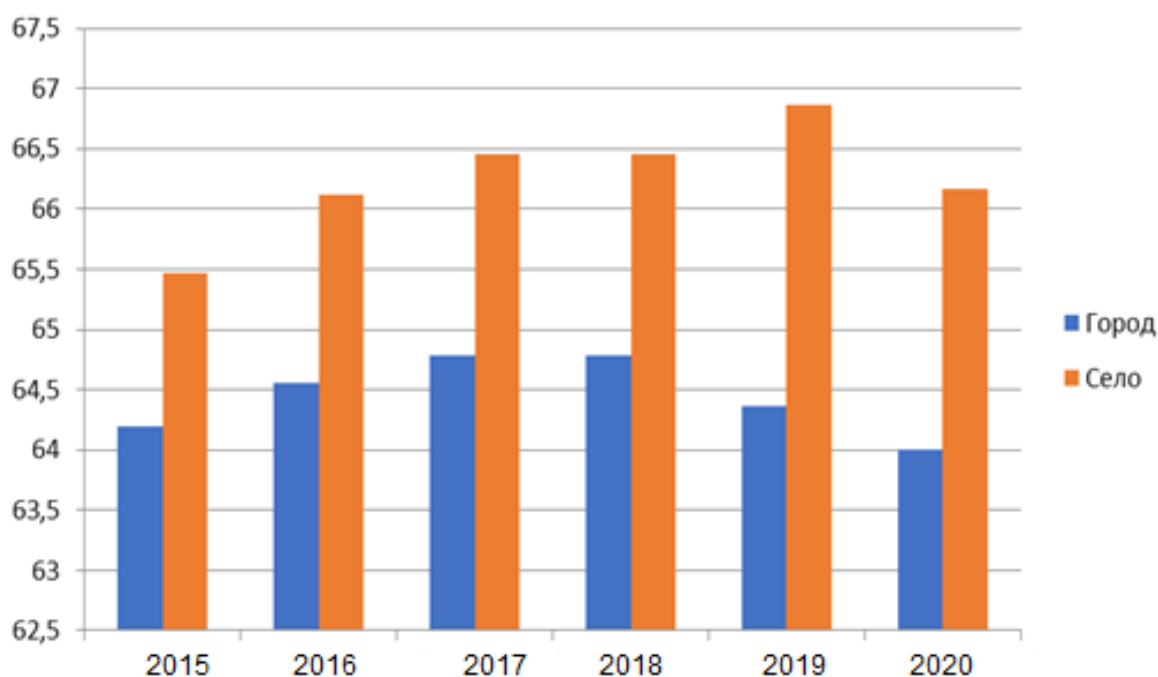


Рисунок 2.6 – Диаграмма ожидаемой продолжительности жизни при рождении по VKO

На рисунке 2.7 изображена диаграмма продолжительности жизни в гендерном аспекте.

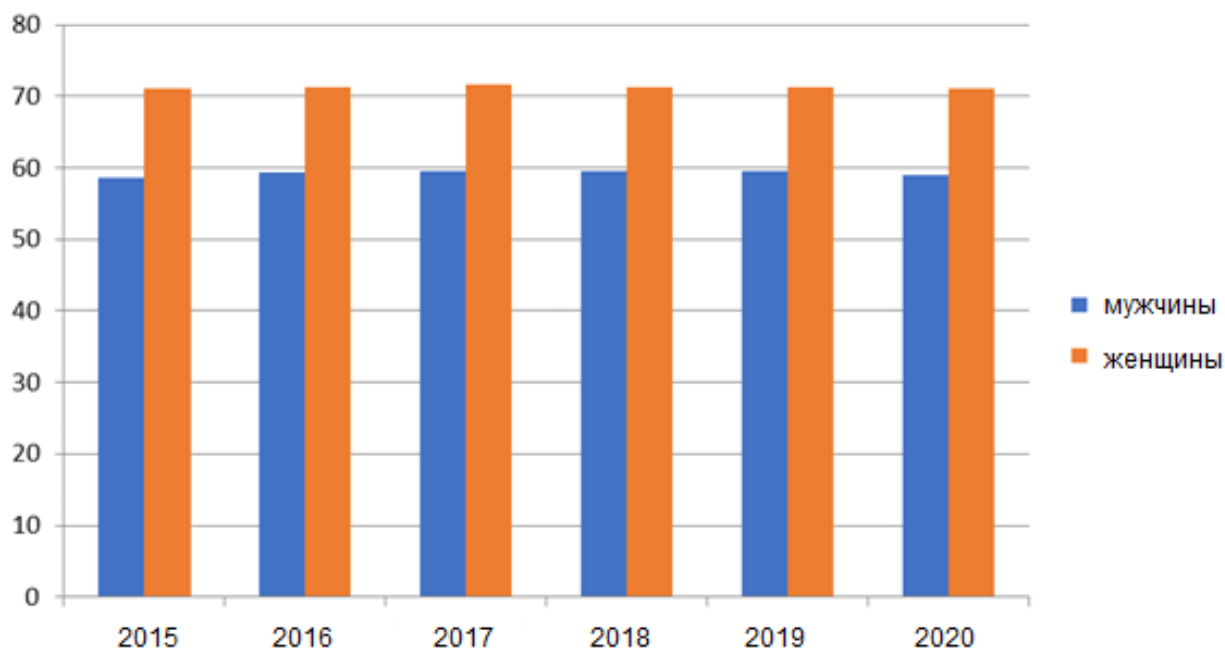


Рисунок 2.7 – Прогнозируемая гендерная информация

На основе диаграммы продолжительности жизни в гендерном аспекте, то есть по полу (рисунок 2.7), можно сделать вывод, что женщины Восточно-Казахстанской области в среднем имеют ожидаемую продолжительность жизни 71,34 лет, ниже этого уровня ожидаемая продолжительность жизни при рождении только в Акмолинской области – 69,92 лет и в Северо-Казахстанской области – 70,94 лет.

Наивысший показатель по Республике Казахстан среди женщин достигнут в г. Нур-Султан – 75,82 лет. У мужчин ВКО в среднем данный показатель составляет 59,25 лет, что отличается от женщин на 12,09 лет. Ниже чем в ВКО ожидаемая продолжительность жизни при рождении в Карагандинской области – 57,65 лет и Акмолинской области – 57,84. Наивысший показатель по Республике Казахстан среди мужчин достигнут в г. Нур-Султан – 73,80 лет. Ожидаемая продолжительность жизни при рождении у мужчин и у женщин ВКО, значительно ниже среднереспубликанского значения.

Особую роль среди негативных характеристик психофизиологического потенциала играет младенческая смертность (до 1 года) (рисунок 2.8).



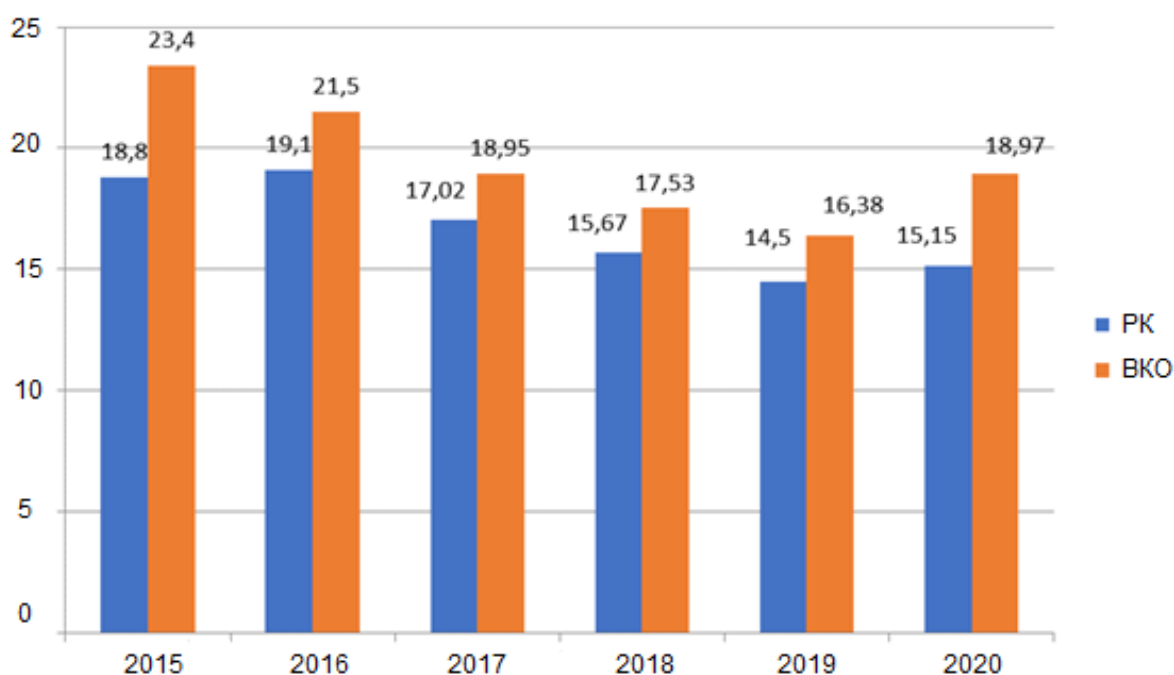


Рисунок 2.8 – Младенческая смертность на 100 родившихся

В 2015 году этот показатель увеличился в десяти регионах Республики Казахстан, лидером явилась Кызылординская область – коэффициент смертности на 1000 родившихся составил 23,89, второе место по республике занимает Восточно-Казахстанская область – 18,97. Атырауская область имеет наименьший коэффициент младенческой смертности 11,49, который постоянно снижается. Остаются высокими показатели детской смертности (до 5 лет).

Так в 2020 году в Кызылординской области – коэффициент смертности на 1000 родившихся составил 29,07, Мангистауской области – 22,63, Восточно-Казахстанской области – 22,20. Ниже среднереспубликанского значения (18,83), коэффициенты детской смертности в Акмолинской – 18,06, Актюбинской – 17,80, Алматинской – 15,97, Атырауской – 15,18, Западно-Казахстанской – 18,41, Карагандинской – 13,91, Павлодарской – 14,53, Северо-Казахстанской – 18,29 областях и в городах Астана и Алматы – 15,48 и 15,54 соответственно.

Таким образом, из приведенного анализа рождаемости и смертности в Восточно-Казахстанской области выявлено, что с 2015 года снижение рождаемости приобретает системную тенденцию.

Ухудшение состояния окружающей среды в ВКО и связанное с ней снижение показателей здоровья жителей, проживающих в этом районе, требуют проведения комплексных исследований в этом регионе.

Термин «загрязнение воздуха» ограничен ситуацией, в которой внешняя среда – атмосфера – содержит вещества в концентрациях, вредных для человека и окружающей среды. Антропогенные трансформации атмосферы неизбежно ведут к негативным последствиям и резкому снижению показателей здоровья населения, что требует внедрения новых методологий в управление

экологической системой города, а также повышению требований к качеству контроля.

Для характеристики состояния здоровья населения автором работы предлагается использовать понятия «качество здоровья», «благополучие населения», «уровень загрязнения атмосферы» под которыми можно понимать совокупность социальных, экономических и медицинских показателей.

#### **2.4 Использование искусственного интеллекта в экологии и здравоохранении**

Для проведения исследования были использованы статистические данные по результатам лечебно-диагностической деятельности областной больницы Восточно-Казахстанского областного медицинского объединения по демографическим и статистическим данным Департамента здравоохранения. Предварительно были изучены публикации, где приводились результаты использования искусственного интеллекта в экологии и здравоохранении.

Связь между загрязнением окружающей среды и заболеваниями населения изучалась многими авторами, и как было выяснено, степень влияния уровня загрязнения, структуры вредных составляющих для организма человека во внешней среде, по-разному проявляются на качестве жизни среднестатистического человека.

Так в работе [90] автором приводятся доказательства и эффективность аппарата регрессионного анализа для контроля  $\text{NO}_2$  и классического «байесовского подхода к моделированию пространственно-временной изменчивости загрязнителей воздуха».

В работе [91] приводятся результаты испытания ультразвукового микроконтроллерного устройства для измерения расстояния между мусоросборником и контейнером для твердых бытовых отходов. Разработка данного устройства входило в план инициативной программы технического обеспечения системы «Экология».

В работе [92] были обоснованы диапазоны уровня достоверности решений для нейронной обработки данных: от 0,00 до 0,40 СЛАБОЕ, от 0,41 до 0,60 УМЕРЕННОЕ, от 0,61 до 0,80 ОБЩИЕ требования, от 0,81 до 1,00 СТРОГОЕ.

В работе [93] приводятся данные по исследованию алгоритма «дерева решений», с целью выявления последствий от выбросов автомобилей на новом топливе. Были обнаружены закономерности, которые невозможно выявить классическими методами при большом объеме данных за несколько лет.

В исследовании [94] анализируются результаты применения искусственного интеллекта на базах данных города Куритиба, Бразилия. Были получены неожиданные результаты и обнаружены назревающие социальные катастрофы.

В работе [95] с помощью AirData и аппарата временных рядов на основе шейплетов, были изучены региональные профили воздействия на окружающую среду. Бинарный классификатор, предсказал вспышку заболеваний, которые было трудно обнаружить известными методами.

В работе [79, р. 3743-3747] исследовались системы на нейронных сетях и возможности искусственного интеллекта для прогнозных целей. Выявлена связь антропогенных и естественных факторов с качеством здоровья в условиях неочевидности внутрисистемных закономерностей.

В результате приведенного анализа были выявлены главные ограничения и недостатки приведенных методов и предложен новый подход для реализации процесса интеллектуальной обработки информации.

Идея подхода состоит в использовании не статичных коэффициентов корреляции, а коэффициентов и функции автокорреляции с автоматически изменяемой начальной точкой в векторе данных.

Предварительно была исследована картина заболеваний в местных урбанизированных условиях города. На рисунке 2.9 представлена структурная диаграмма заболеваний.

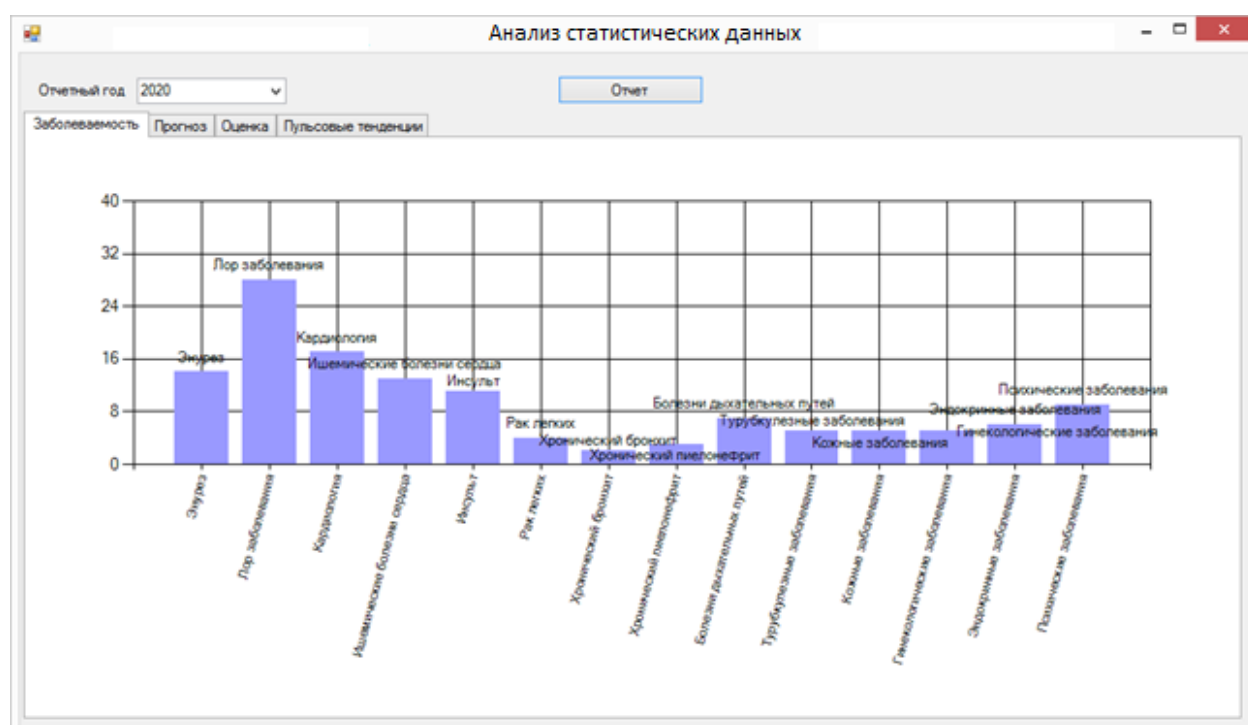


Рисунок 2.9 – Структурная статистика заболеваемости по г. Усть-Каменогорск

Из рисунка 2.9 и дополнительных исследований выявлено, что наиболее частые заболевания это сердечно-сосудистые, предвестниками которых являются высокое артериальное давление и пульсовые характеристики. Однако, в силу индивидуальности людей показатели нормы могут значительно отличаться. В настоящее время на рынке медицинских приборов предлагается большой спектр различных тонометров.

Как показали исследования достоверности результатов измерения давления различными средствами, только 60-70% дают относительно «удовлетворительные» результаты. Тогда возникает вопрос – каков риск при измерении артериального давления, если его оценить по методике,

предложенной в данной работе. Для реализации этой задачи были проведены экспериментально-статистические исследования и компьютерный эксперимент для количественной оценки вероятных ошибок  $P_{\text{ло}}$  (ложный отказ) и  $P_{\text{но}}$  (необнаруженный отказ).

Статистические исследования показали следующие результаты, которые в виде гистограммы представлены на рисунке 2.10.

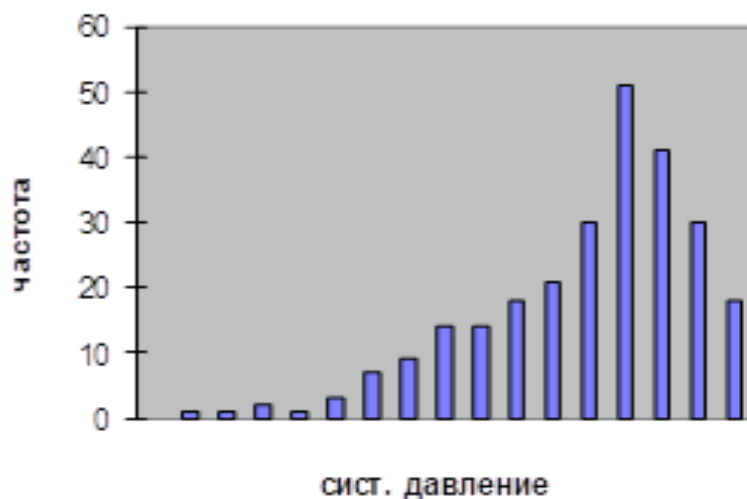


Рисунок 2.10 – Гистограмма распределения систолического давления у группы населения возрастом до 40 лет

Обработка приведенных экспериментальных результатов показала, что среднее и среднее квадратическое отклонение для рисунка 2.10 равны:  $S_{\text{ср}} = 26.3$  мм рт.ст;  $\sigma = 0.6$  мм рт.ст. Аппроксимация эмпирического распределения теоретическим законом показала, что гипотеза о нормальном законе распределения по критерию Пирсона с доверительной вероятностью 0,95 не отвергается.

Так как выборка по возрастной группе формировалась не одновременно, а в течении нескольких опытов и состояла из совокупностей, полученных в разное время, то возникла необходимость оценки однородности данных совокупностей при композиции в единый массив данных. Для статистической проверки гипотезы о качественной неоднородности, приведенной выборки, были использованы два метода: метод, основанный на использовании критерия Фишера и метод качественного различия средних арифметических, использующий t-статистику. Расчеты показали, что гипотезы об однородности для выборки не отвергаются с вероятностью 0,95.

Прежде чем рекомендовать к практическому использованию индикаторов, необходимо построить его модель. Моделирование индикаторов качества определяется системой мониторинга качества клинической деятельности и поддерживаемой версией истории болезни.

В основе всех известных методик разработки индикаторов качества лежат статистические наблюдения и их соответствующая обработка. Но главным и принципиальным недостатком подобной технологии является тот факт, что

элементы вектора данных (выборки) одного и того же индикатора качества за определенный интервал времени изменяется и не отвечают требованиям однородности.

С позиций статистической надежности результатов интеллектуального контроля данных на примере задач управления процессами, демонстрируют методы использования корреляций в динамических рядах временных данных, таких как, автокорреляционная и взаимокорреляционная функции. Методология данного подхода опирается на предсказуемость данных с определенным лагом изменений элементов вектора информации.

Использование автокорреляционной функции для временных рядов облегчается тем, что данные «привязаны» к определенному моменту времени и временные корреляционные расчетные результаты имеют гладкий скользящий вид. В случае качественного изменения внутри или внесистемного характера на функции отразится этот факт заметным количественным откликом, что будет индикатором изменения качества следующей последовательности данных. Для интеллектуальной системы этот «скачок» на поле данных является диагностическим индикатором явления.

Формула расчета автокорреляционной связи осуществляется по формуле (2.1):

$$r_{xy} = \frac{\sum_{t=2}^n (y_t - \bar{y}) \cdot (y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^n (y_t - \bar{y}_1)^2 \cdot \sum_{t=2}^n (y_{t-1} - \bar{y}_2)^2}}, \quad (2.1)$$

где средние для переменных  $y$  и  $x$  производятся по известным выражениям, и как правило, в автоматическом компьютерном режиме.

Для решаемой в работе задачи нормативной кривой является «типовая» коррелограмма, которая предположительно должна быть индикатором исследуемого процесса, в данном случае, интегральным показателем качества здоровья или качество атмосферы. Типовая коррелограмма формируется в результате, так называемого, «обучения» интеллектуальной системы по накопленным данным. Наличие режима «обучения» является одним из общепринятых признаков искусственного интеллекта. В процессе обучения обязательным требованием является проверка вновь появившихся данных на однородность по  $F$  и  $t$  критериям. Данные проверки на однородность предлагается представлять в виде временного ряда и программно исследовать на автокорреляционную связь, что будет также являться классификационным индикатором.

Надежные индикаторы процессов являются базой для оценки и прогнозирования степени риска, а также создания системы клинического и экологического мониторинга, включающего процедуры: измерения, сравнения результата измерения с нормативом и принятия решений в реальном масштабе времени.

На рисунке 2.11 представлена диаграмма, которая дает возможность визуально проследить тенденции от дня поступления к каждому дню лечения в

стационаре как всех больных, так и конкретного больного, и больных, прикрепленных к конкретному врачу.

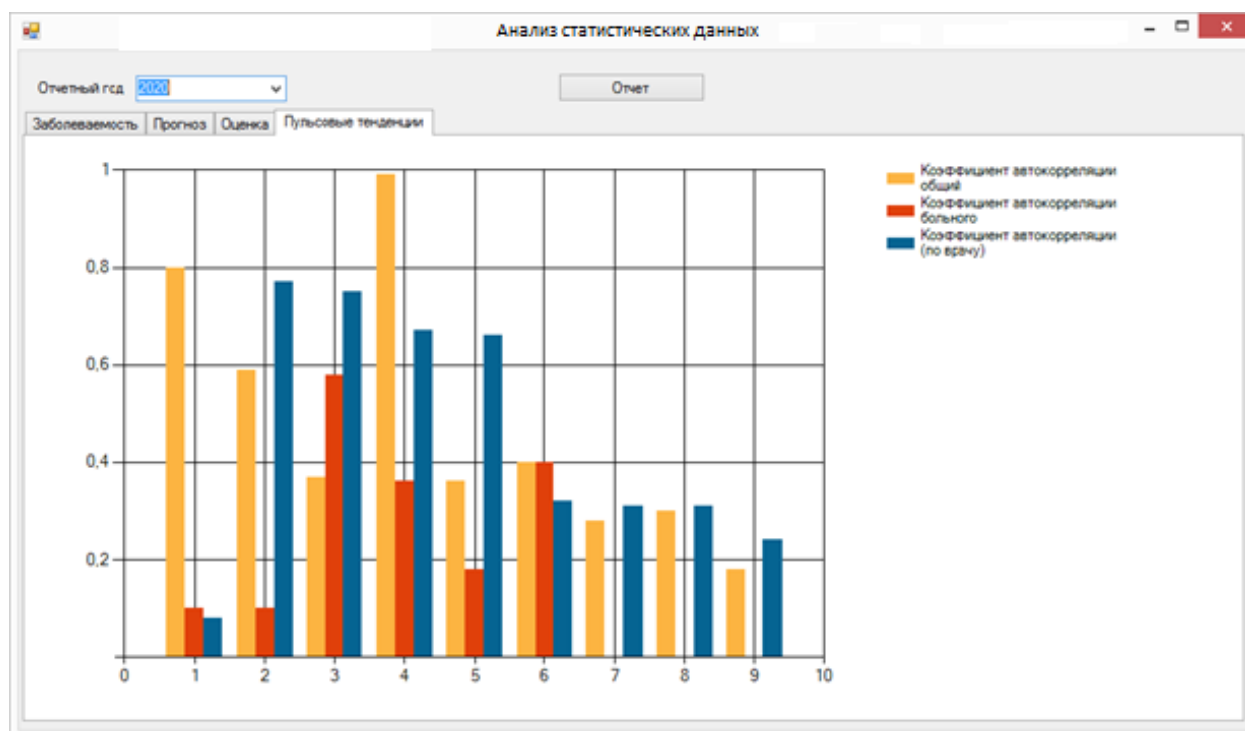


Рисунок 2.11 – Сравнение коэффициентов автокорреляции

Входящий вектор данных – личная карта врача, удостоверение личности пациента (пенсионное удостоверение, свидетельство о рождении) и направление на лечение (если пациент направлен другим медицинским учреждением) преобразующиеся в процессе поликлинических исследований в историю болезни, которая, далее проходя через специалистов пополняется температурным листом, и на основе этих данных преобразуется в выходящие документы: карта выбывшего из стационара и анализ лечебно-диагностической деятельности проведенный руководителем предприятия. Предусмотрено ведение справочной информации, которую можно изменять и пополнять.

Для исследования эффективности предложенной методики, были проведены экспериментально-статистические наблюдения в реальных условиях. По результатам наблюдений рассчитаны и построены корреляционные эмпирические кривые, которые демонстрируются на рисунках 2.12, 2.13.

### Систолическо давление

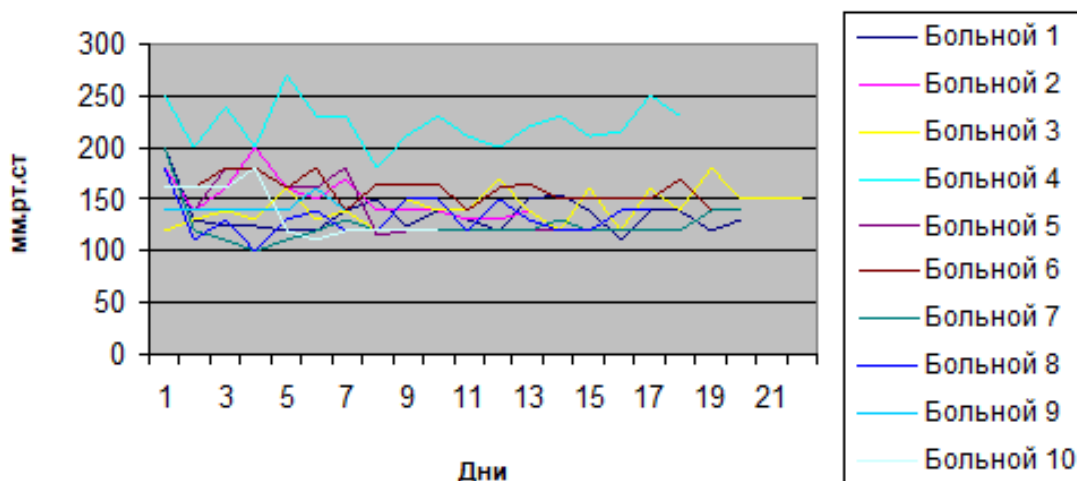


Рисунок 2.12 – Случайные реализации систолического давления

Была сформирована статистическая выборка из пациентов областной больницы в составе объемом 10 человек. В качестве контролируемого индикатора здоровья было исследовано артериальное давление. Контрольная группа содержала пациентов разного смешанного возраста от 25 до 55 лет.

### Диастолическое давление

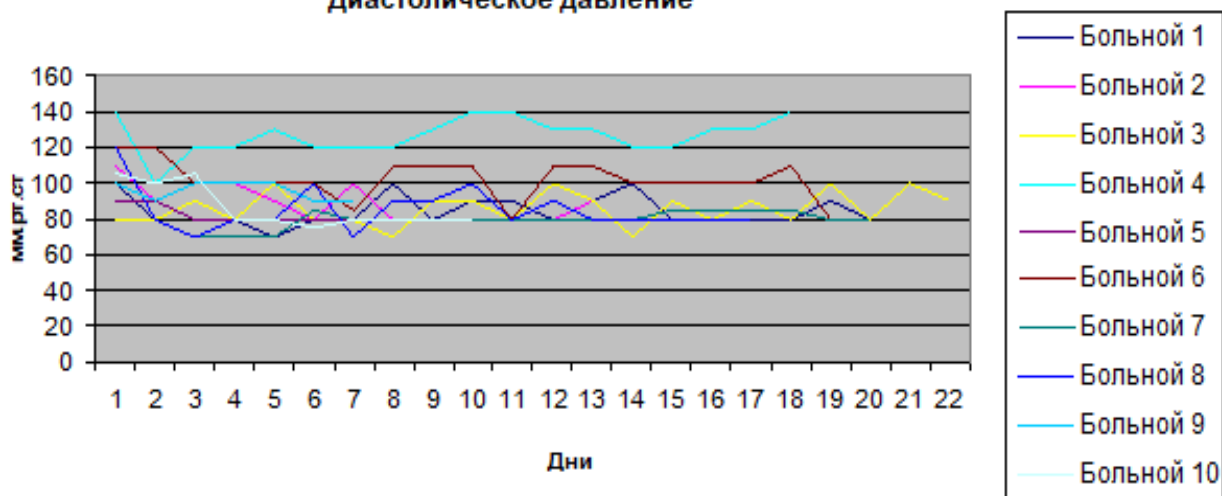


Рисунок 2.13 – Случайные реализации диастолического давления

Анализ семейства корреляционных кривых показывает на проблематичность персонального автоматизированного контроля в связи с отсутствием объективных нормативов. Вместе с тем, результаты индицируют возможность получить эти нормативы путем изменения корреляционного лага, и по признакам «микробифуркаций» при определенном лаге фиксировать конец однородности группы и начала следующей группы. Эту процедуру представляется возможным реализовать только программным интеллектуальным способом.

Следующая гипотеза, которая исследовалась в эксперименте, ставила задачу выявить диагностические признаки групповой значимости у таких индикаторов, как вариация параметра. В качестве вариации было выбрано среднееквадратическое отклонение.

На рисунках 2.14 и 2.15 иллюстрируются функции отклонений давления и пульсовых характеристик пациента.

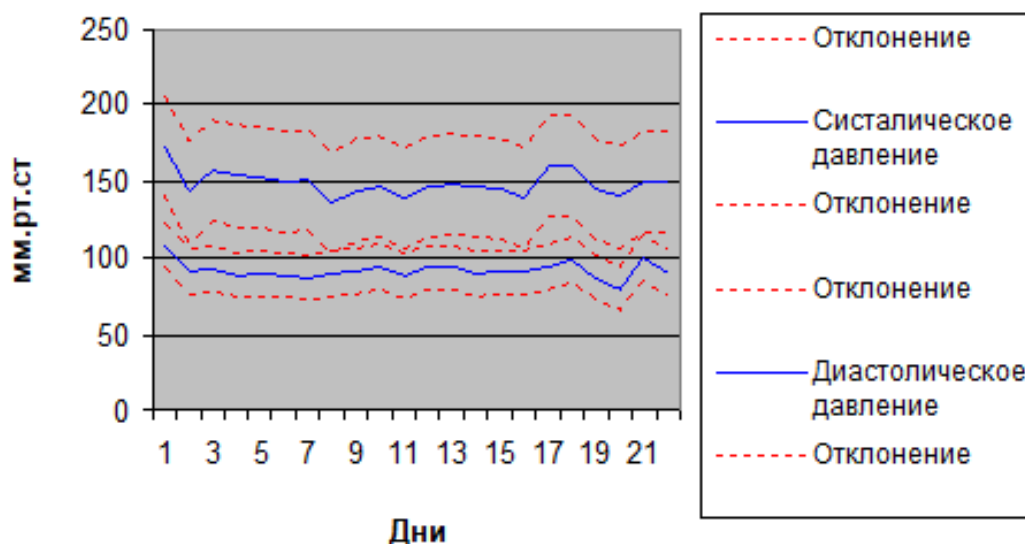


Рисунок 2.14 – Зависимости средних и вариаций нижнего и верхнего значений артериального давления

Для реализации нейронной технологии контроля качества экологического и лечебно-диагностического процесса предлагается использовать групповые средние и вариационные показатели в качестве нормативных функций автокорреляции, которые могут иметь аналитическую форму.

На графиках рисунка 2.15 иллюстрируется пульсовая кривая среднего и среднееквадратического отклонения.

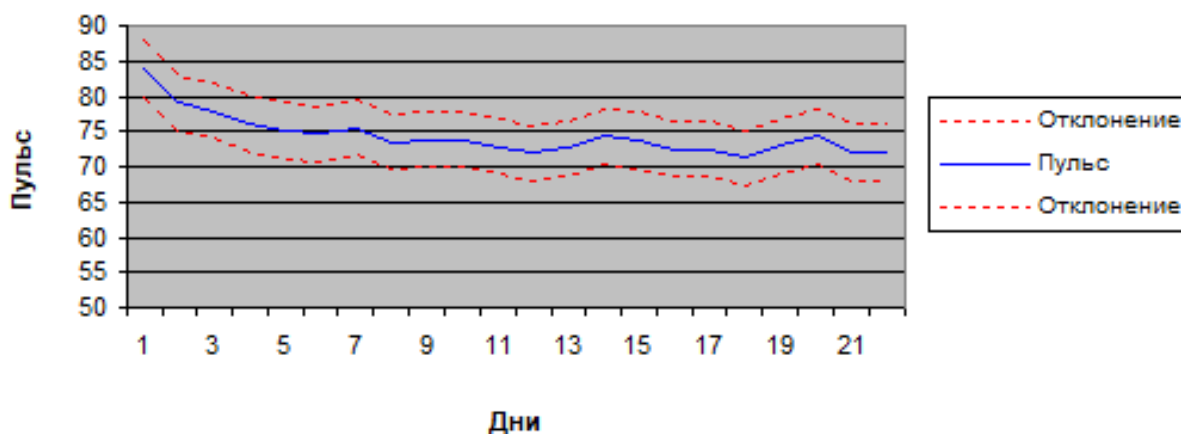


Рисунок 2.15 – Пульсовая кривая среднего и среднееквадратического отклонения



На рисунке 2.16 приводится автокорреляционная динамика пульса. Возникает вопрос, что дает эта временная статистика.

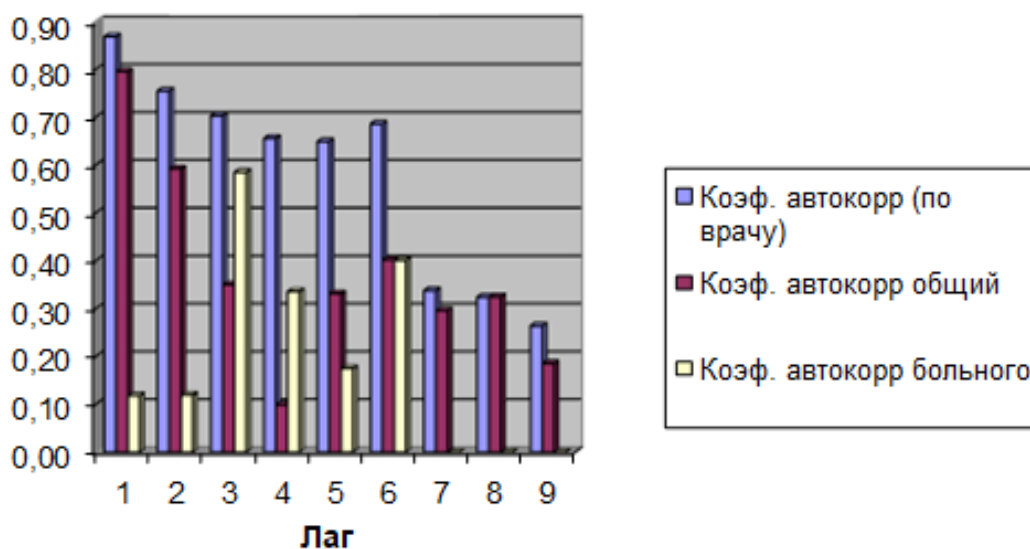


Рисунок 2.16 – Автокорреляционные функции пульса

Алгоритм интеллектуальной обработки данных, реализованный на предложенной математической основе и выявления в корреляционной непрерывной кривой качественных динамик (скачков), позволяет ретроспективно «заглянуть» в информационное прошлое и сопоставить с внешними организационными, экологическими данными.

## 2.5 Математическое обеспечение системы интеллектуальной обработки экологической и медицинско-профилактической информации

Опыт развитых стран показал, что экономическая поддержка, какой бы она ни была значительной, отдельных направлений здравоохранения вне системы не приводит к какому-либо ожидаемому результату. Важность, актуальность и безальтернативность внедрения системы качества на базе международных стандартов ISO независимо от отраслевой принадлежности и форм собственности была отмечена стать основным концептуальным подходом для тех, кто стремится интегрировать ограниченные ресурсы общества для кардинального решения проблем здоровья населения региона.

Современная отечественная и международная политика в области общественного здравоохранения требуют количественной оценки здоровья людей в масштабах регионов и стран. Анализ последних публикаций показал, что эта задача еще более актуальна применительно к так называемому «качеству здоровья» как средний показатель для группы населения, так и отдельного жителя.

В процессе выполнения диссертационной работы был предложен метод формального анализа и количественного оценивания социально-экономических критериев проживания населения.

Графическая модель (рисунок 2.17) демонстрирует картину снижения качества здоровья. Количественно критерий снижения качества здоровья измеряется в диапазоне 0-1. Единица в данном подходе обозначает полное использование жизненных ресурсов оцениваемого субъекта. Потеря ресурсов качества жизни и здоровья на рисунке 2.17 изображена прямой линией, которая имеет математическую форму. Реально эта математическая функция определяется регрессионными методами. Регрессионная модель может иметь нелинейную форму.

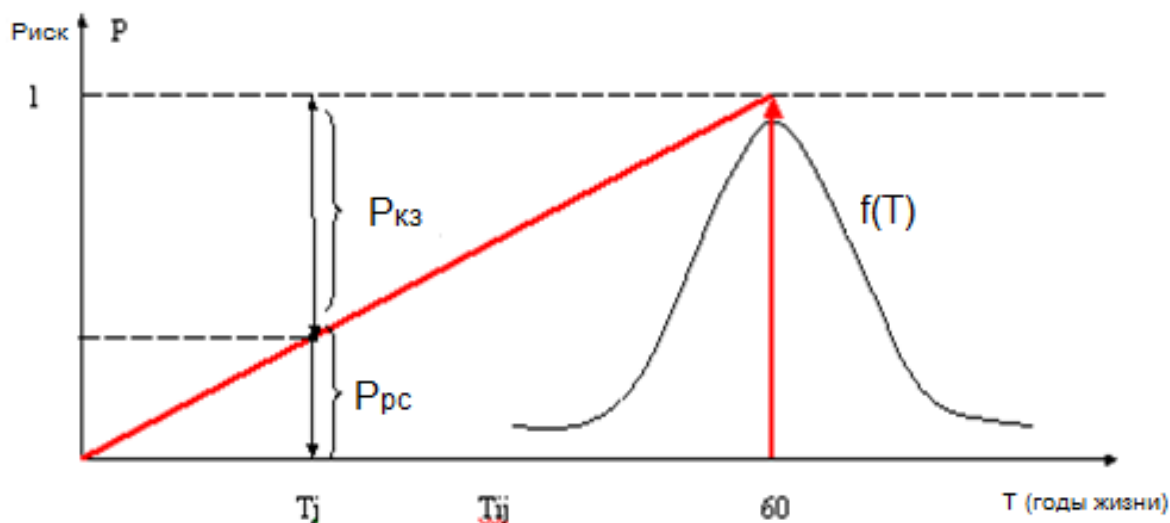


Рисунок 2.17 – Модель прогнозирования вероятности снижения качества жизни

На рисунке 2.17 вероятность  $P_{кз}$  определяет остаточный вероятный ресурс здоровья,  $P_{рс}$  – вероятность использованного жизненного ресурса (ресурса здоровья),  $f(T)$  – функция плотности распределения величина продолжительности жизни. Предполагается, что закон распределения  $f(T)$  при уровне значимости 0,05 не противоречит гипотезе о нормальном распределении.

Генерируя время продолжительности жизни отдельного жителя региона ( $T_i$ ), и проводя через эту точку прямую линию, получается индивидуальная функция качества здоровья по которой можно вычислит вероятности  $P_{кз}$  и  $P_{рс}$ . Эта ситуация демонстрируется рисунком 2.18. Данные вероятности имеют смысл жизненных рисков.

В результате статистического эксперимента, проведенного в Областном медицинском объединении г. Усть-Каменорске исследовались следующие значимые факторы, влияющие на продолжительность жизни и качество здоровья: загрязнение атмосферы промышленными объектами и автотранспортом, состояние здравоохранения, наследственность, экономическое благосостояние и образ жизни.

Используя методику экспертных оценок и программу автоматизации этой технологии, были найдены весомости каждого из перечисленных факторов

здоровья. Тогда, аналитическое выражение для прогнозирования рисков (качества здоровья) можно выразить следующим образом:

$$P_{KЗ} = 1 - \sum_{i=1}^n \alpha_i P_{iPC}, \quad (2.2)$$

где  $\alpha_i$  – «вес» фактора;

$P_{iPC}$  – весовая вероятность влияния  $i$ -го фактора.

Представленная формула показывает способ вероятностной оценки деградации жизненного ресурса в функции времени жизни.

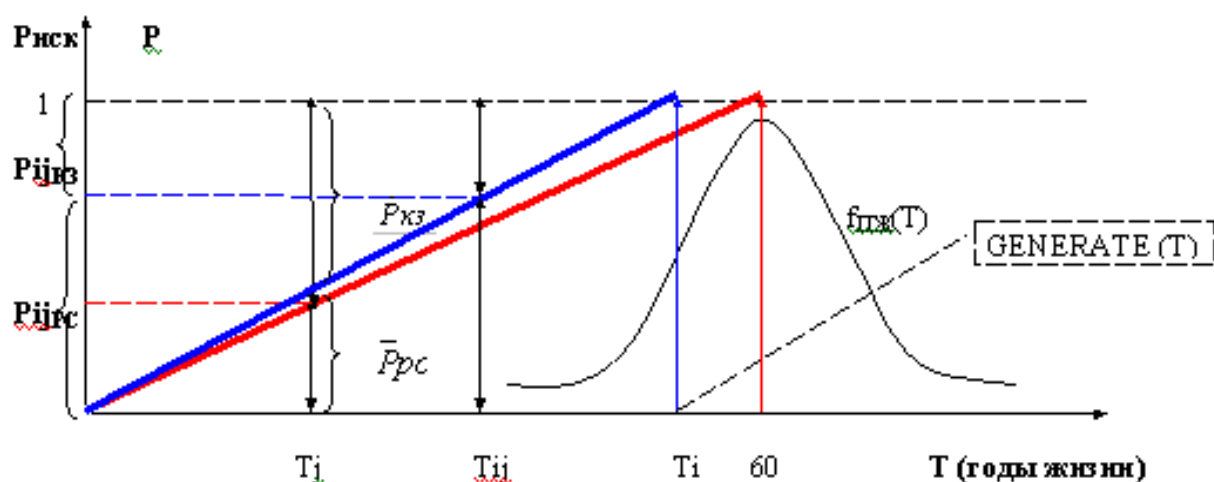


Рисунок 2.18 – Модель зависимости рисков снижения качества здоровья

В качестве математического аппарата для графических построений работы использовался регрессионный анализ в программной среде STATISTICA [96]. Эмпирические модели были построены по экспериментально-статистическим данным из документов отчетности Восточно-Казахстанском областном медицинском объединении (ВК ОМО).

На рисунках (рисунки 2.19, 2.20, 2.21, 2.22,) в графической форме демонстрируются результаты экспериментально-статистических исследований и компьютерного эксперимента. С помощью разработанной программы определяется математическое выражение для функции влияния качество здравоохранения на риск снижения качества здоровья. Данные для построения эмпирической функции, были выявлены в отчетно-статистической документации ВК ОМО.

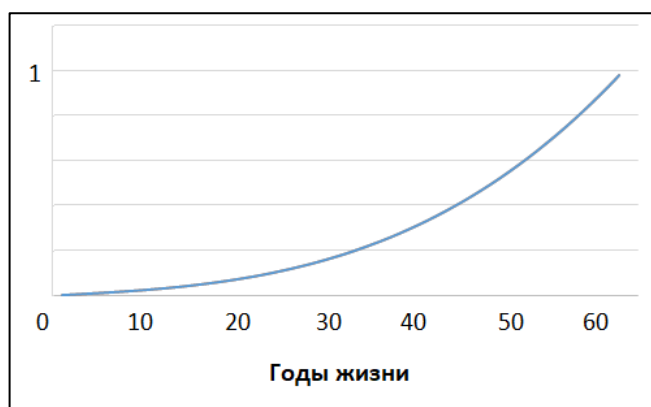


Рисунок 2.19 – Линия тренда кривой риска потери здоровья в функции качества системы здравоохранения

Получена следующая регрессионная модель:

$$Y = -0,0175 x^3 + 0,0198 x^2 + 10,445 x - 117,58 \quad (2.3)$$

Расчетная величина критерия Фишера равна:

$$F = 149,71 \gg F_{\text{таб}} . \quad (2.4)$$

В данном случае эмпирическое значение F-критерия существенно выше порогового табличного, из чего можно сделать заключение, что гипотеза о полиномиальной зависимости влияния здравоохранения на риск снижения качества здоровья от возраста при уровне значимости 0,05 не отвергается.

Вторым исследуемым системным фактором влияния на качество здоровья анализировался экологический. На рисунке 2.20 иллюстрируется эмпирическая зависимость продолжительности жизни в функции качества экологической обстановки на оцениваемой местности.

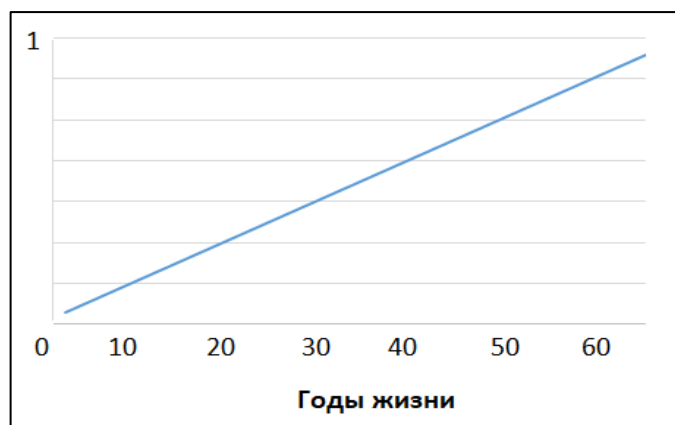


Рисунок 2.20 – Эмпирическая зависимость продолжительности жизни в функции качества экологии

Регрессионная модель изучаемого явления имеет следующую аналитическую форму:

$$Y = 5x + 15. \quad (2.5)$$

Для данной модели критерий Фишера равен:

$$F = 684,45 \gg F_{\text{таб}}. \quad (2.6)$$

Следовательно, гипотеза о линейной связи исследуемых показателей не отвергается с уровнем значимости 0,05.

Рассмотрим графики влияния экономического благосостояния на риск снижения качества здоровья. Для определения влияния экономического благосостояния используем полиномиальный тренд второй степени (рисунок 2.21).

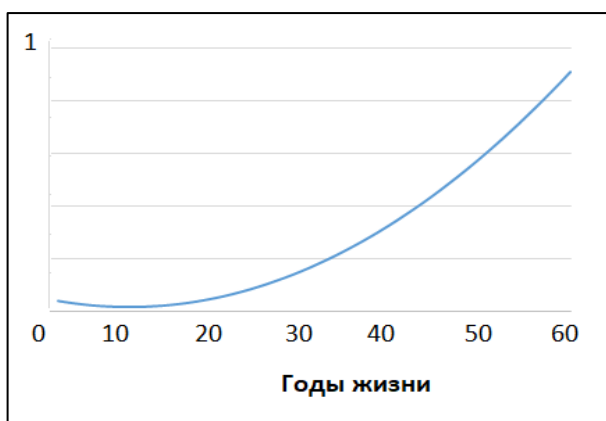


Рисунок 2.21 – Влияние экономического фактора риска на качество здоровья

Получена следующая регрессионная модель:

$$Y = 0,8682 x^2 - 15,928 x + 110,15. \quad (2.7)$$

Используя полученное выражение, определяем критерий Фишера:

$$F = 434,19 \gg F_{\text{таб}}. \quad (2.8)$$

Следовательно, гипотеза о нелинейной зависимости фактора риска в продолжительности жизни от экономического благосостояния не отвергается при уровне значимости 0,05.

Для построения многофакторной модели интегрированной оценки влияния факторов антропогенной социально-экономической урбанизированной среды обитания населения на качество здоровья, были выделены наиболее значимые показатели влияния. Для оценки весомости системных показателей

был проведен экспертный анализ, результаты которого приведены в таблице 2.4.

Таблица 2.4 – Экспертный факторный анализ качества здоровья

Факторы качества здоровья	Коэффициент весомости факторов $\alpha$
Качество здравоохранения	0,12
Уровень загрязнения атмосферы	0,67
Благосостояние	0,21

Для проведения экспертизы экспертам были представлены статистические данные органов экологии города за пятилетний период. Таким образом, с учетом полученных эмпирических данных и двухфакторных связей была получена интегрированная многофакторная модель риска, для субъекта, обладающим высоким экономическим благосостоянием:

$$P_{K3} = 1 - 0,01(\alpha_1(-0,017x^3 + 0,0198x^2 + 10,445x - 117,58) + \alpha_2(5x + 15) + \alpha_3(0,8682x^2 - 15,928x + 110,15) + \sum \alpha_i P_i), \quad (2.9)$$

где  $x$  – годы жизни;

$P_i$  – неучтенные риски.

Анализ результатов компьютерного моделирования связей рисков для здоровья и продолжительности жизни имеет сложный нелинейный многопараметрический характер. Приведенное математическое и программное обеспечение является базовыми агентами интеллектуальной информационной системы оценки негативного влияния промышленных выбросов на здоровье населения.

В процессе выполнения диссертационной работы был предложен метод формального анализа и количественного оценивания социально-экономических критериев проживания населения. В результате статистического эксперимента выявлены следующие значимые факторы, влияющие на продолжительность жизни и качество здоровья: «состояние здравоохранения», «загрязнение атмосферы», «экономическое благосостояние».

Используя методику экспертных оценок и программу автоматизации этой технологии, были найдены весомости каждого из перечисленных факторов здоровья. Таким образом, с учетом полученных эмпирических данных и многофакторных связей была получена интегрированная многофакторная модель риска снижения качества здоровья.

### **Выводы по второму разделу**

В ходе анализа объективных предпосылок к системной интеллектуализации формальных процессов принятия решений в мультипараметрической среде выявлено, что разрабатываемая система в

обязательном порядке должна содержать блок количественного оценивания рисков в форме мониторинга.

В результате приведенного анализа были выявлены главные ограничения и недостатки приведенных методов и предложен подход для реализации процесса интеллектуальной обработки информации.

Идея подхода состоит в использовании не статичных коэффициентов корреляции, а коэффициентов и функции автокорреляции с автоматически изменяемой начальной точкой в векторе.

В результате статистического эксперимента выявлены значимые факторы, влияющие на продолжительность жизни и качество здоровья. Используя методику экспертных оценок и программу автоматизации этой технологии, были найдены весомости каждого из перечисленных факторов здоровья.

Разработана методика оценивания и прогнозирования качества здоровья населения, включающая математическое обеспечение системы контроля и принятия решений в процессе экологического мониторинга на основе построения интегрированной многофакторной модели.

### 3 ИМИТАЦИОННАЯ МОДЕЛЬ КОЛИЧЕСТВЕННОЙ ОЦЕНКИ КАЧЕСТВА ИНСТРУМЕНТАЛЬНОГО КОНТРОЛЯ В СИСТЕМЕ МОНИТОРИНГА ВОЗДУШНОЙ СРЕДЫ ГОРОДА

#### 3.1 Количественная оценка качества инструментального контроля в системе мониторинга воздушной среды города

Контроль является неотъемлемым и ключевым процессом любой системы управления [97]. Как отмечено во многих работах [98, 99], «Экологический контроль является неотъемлемым звеном в системе инженерно-экологического обеспечения». Основная задача экологического контроля (ЭК) состоит в формировании информационной базы состояния и изменения окружающей среды. Важнейшими критериями качества контроля, как указывает Мазур И.И. и соавторы, являются точность и достоверность [100].

Вопрос о «мерах точности» имеет равнозначное отношение и к результатам измерений, и к практической эффективности реализуемых решений. Интегральным критерием, отмечается авторами, должна служить «информационная емкость», из которой следуют требования к точности и достоверности при заданных нормативах.

Качество контроля является интегрированным критерием, который является агрегированным показателем, зависящим от композиции статистических характеристик агентов контроля, а именно: неопределенности, включая закон распределения, контролируемого параметра, неопределенности, включая закон распределения, результата измерения, формы нормативных ограничений.

Системный результат данной композиции – это достоверность контроля и ошибки контроля. Ошибки контроля бывают двух типов: вероятность ложного отказа и вероятность необнаруженного отказа. Эти вероятные ошибки являются рисками производителя и потребителя.

Графически в форме 3D эти ошибки демонстрируются следующим образом (рисунок 3.1).

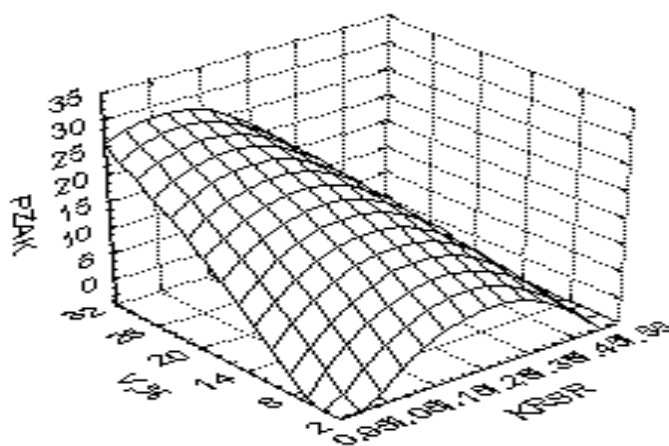


Рисунок 3.1 – Вероятность риска заказчика

Примечание – Составлено по источнику [100, с. 149-150]



Вероятность риска изготовителя представлена на рисунке 3.2.

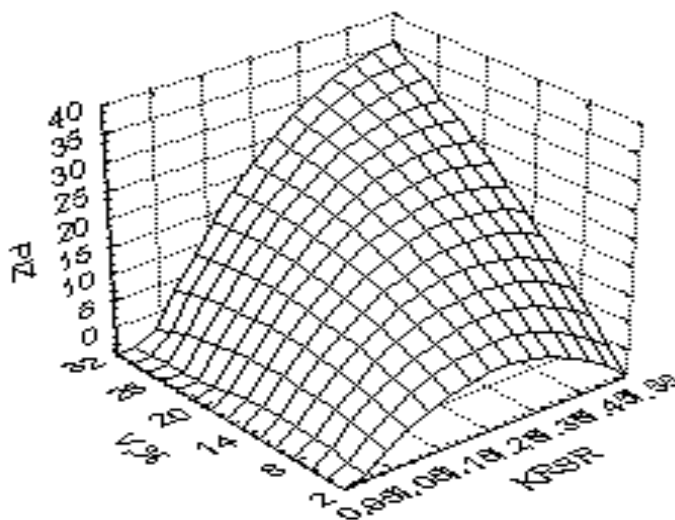


Рисунок 3.2 – Вероятность риска изготовителя

Примечание – Составлено по источнику [100, с. 151-152]

Так как на практике и системах интеллектуально обработки информации данные имеют более нечеткую, чем статистическую форму, а также более употребительны и применимы лингвистические понятия и оценки, был разработан нечеткий алгоритм для указанных целей, в форме графической схемы, представленный на рисунке 3.3.

На рисунке 3.3 представлена нечетко-имитационная схема моделирования. Контролируемый параметр в общем случае будем обозначать символом  $S_i$ . Оценка вероятных ошибок, возникающих в процессе контроля может быть найдена только путем компьютерного моделирования.

Как следует из рисунка 3.3, после ввода исходной нечеткой информации, к которой относятся значения трапеций функций принадлежности  $A_1, A_2, C_1, B_2, D_1, A_3, C_2, B_3, D_2$ , начинается имитационный процесс генерирования случайных чисел – параметров и некоторых действий – процедур. На рисунке 3.3 оператор (GENERATE) разыгрывает первое случайное значение  $S_i$  контролируемого параметра. В данном алгоритме из трех возможных законов распределения: нормального, равномерного и Вейбулла – предполагается, параметр  $S$  аппроксимируется нормальным распределением.

Следующей процедурой генерируются параметры измерительного прибора, а именно,  $A$  и  $B$ . Затем, имитируется процесс измерения. В алгоритме предусмотрены два варианта точности измерения. Первый вариант, когда погрешность прибора подчиняется нормальному закону, а второй вариант равномерного распределения погрешности.

Затем следует логика – анализируется в какой зоне находится измеряемый параметр, а именно, в зоне допуска, или за допустимыми значениями. От результата анализа начинается «разветвление IF», в траектории алгоритма и в программных счетчиках аккумулируются вероятные эквиваленты ошибок контроля.

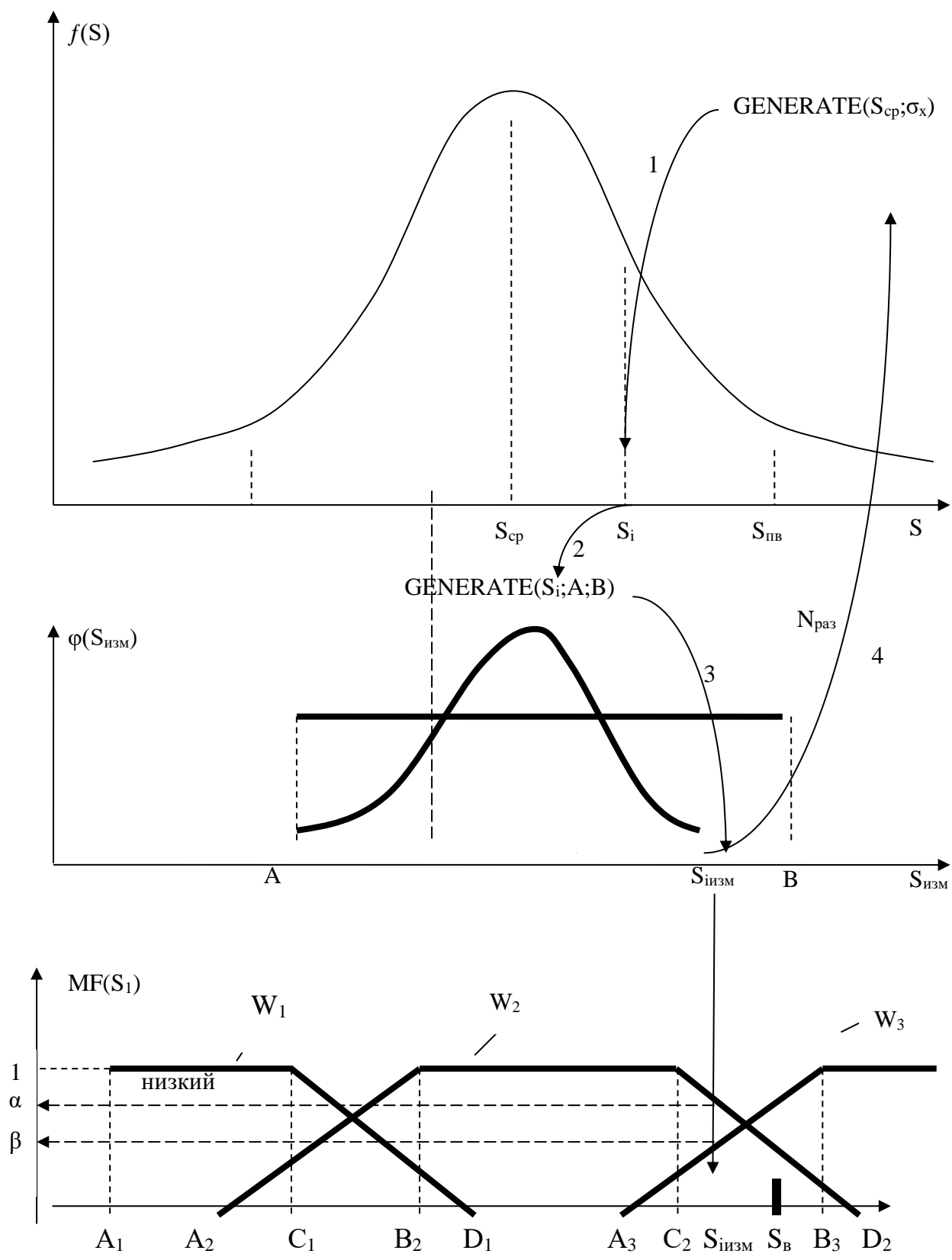


Рисунок 3.3 – Графическая 2D модель имитационной траектории интеллектуального контроля

Вероятности  $P_{но}$  и  $P_{ло}$  вычисляются по формулам (3.1), (3.2):

$$P_{\text{ло}} = \frac{N_{\text{ло}}}{N}, \quad (3.1)$$

$$P_{\text{но}} = \frac{N_{\text{но}}}{N}, \quad (3.2)$$

где  $N_{\text{ло}}$  и  $N_{\text{но}}$  – содержимое счетчиков ложного и необнаруженного отказов.

Функция  $MF(S_i)$  является лингвистическая переменная  $i$ -го индикатора загрязнения воздуха. Данная переменная представлена тремя функциями потерь: «низкий», «средний» и «высокий». Потери связаны с загрязнением среды и независимо от уровня они всегда существуют. Условность (неопределенность) нормативов в нечеткой методологии показана интервалами  $A_2-D_1$  и  $A_3-D_2$ .

В районе указанных интервалов функции принадлежности имеют плавный линейный переход из одной зоны допуска в другую, таким образом создается интервал неопределенности норматива. Результат измерения некоторого индикатора –  $S_{\text{изм}}$ , если он окажется в зоне допуска, будет принадлежать с «вероятностью»  $b$  среднему значению загрязнения и с «вероятностью»  $a$  высокому интервалу. Это приведет к потерям, которые суммарно для  $S_{\text{изм}}$  могут быть определены из выражения:

$$C_i = \alpha \cdot W_2 + \beta \cdot W_3. \quad (3.3)$$

Значения потерь  $W_i$  для каждого из интервалов индикатора  $S_i$  могут быть определены теоретическими расчетами, экспериментально или экспертным методом, к чему чаще всего и прибегают [101].

На рисунке 3.4 алгоритм, реализующий имитационную модель, приведенную на рисунке 3.3.

В первом блоке осуществляется ввод статистических характеристик всех агентов системы интеллектуальной обработки информации процесса контроля и принятия решений:

$\sigma_s$  – среднее квадратическое отклонение статистической функции распределения контролируемого параметра  $S$ ;

$\sigma_\phi$  – среднее квадратическое отклонение статистической функции распределения погрешности измерения;

$S_{\text{ср}}$  – среднее арифметическое статистической функции распределения контролируемого параметра  $S$ ;

$N$  – количество испытаний модели (количество имитаций). Все статистические параметры находятся эмпирически. Оптимальное количество испытаний  $N$  находится итерационно, потому что чем больше  $N$ , тем большее время моделирования.

Во втором блоке открывается цикл моделирования по переменной  $i$ .

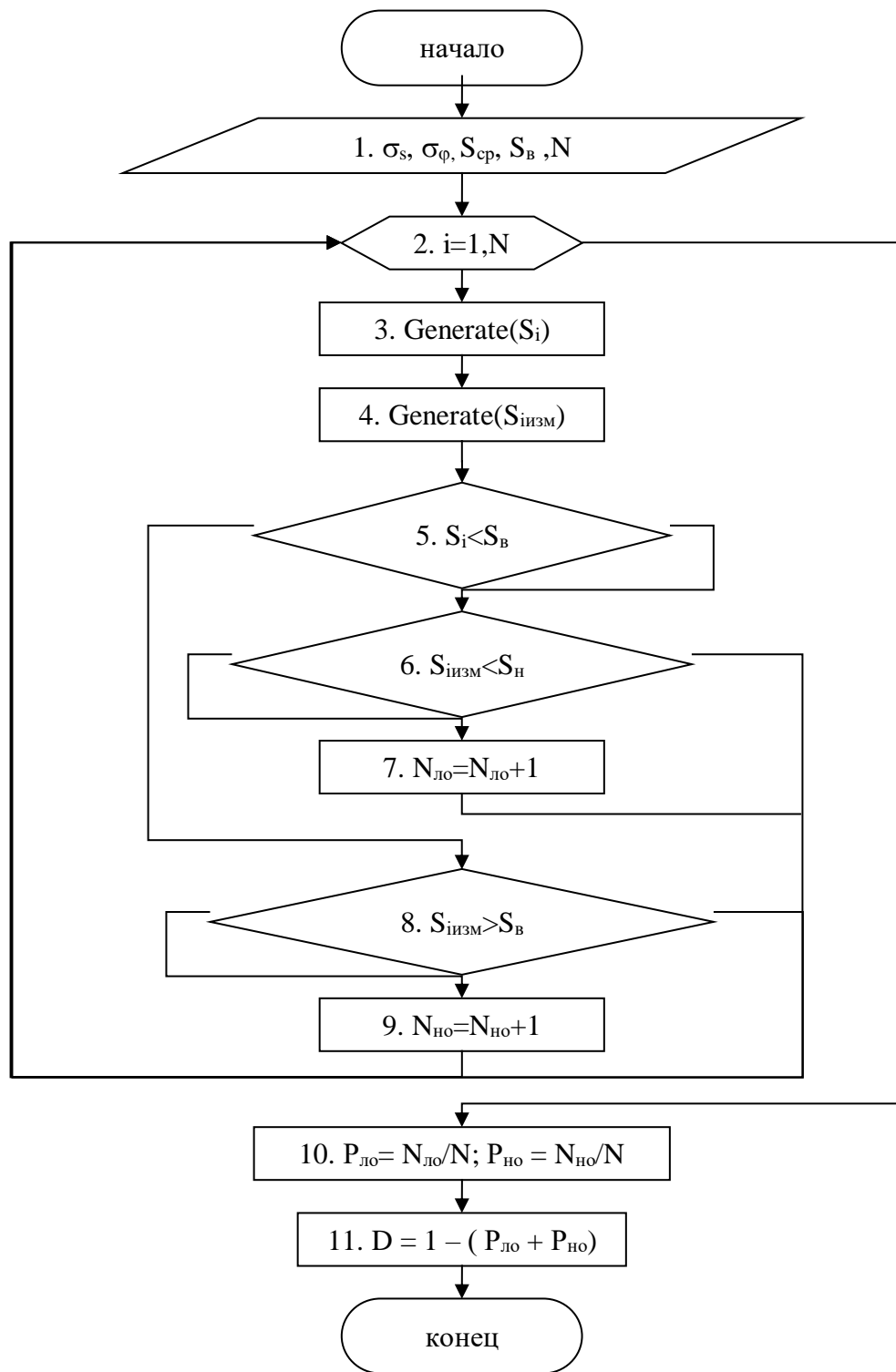


Рисунок 3.4 – Имитационный алгоритм формирования рисков контроля

В третьем блоке «разыгрывается» очередное значение параметра  $S_i$ . При использовании генераторов случайных чисел надо обращать внимание на качество генераторов. В основе алгоритма генерирования находится математический метод, который в начале формирования случайного числа используется тоже некоторое число. Каждый генератор генерирует псевдослучайные числа с циклом повторения чисел.

В четвертом блоке разыгрывается событие «измерения». Эта процедура осуществляется также путем генерирования. И в алгоритме происходит обращение к подпрограмме с предварительной передачей в эту программу параметров закона распределения погрешности измерения.

В пятом и шестом блоках осуществляет логическая операция анализа по следующей логике:

– если в результате генерирования значение  $S_i$  оказалось в зоне допуска (больше  $S_n$ , но меньше  $S_b$ ), а результат измерения прибором случайно – ошибочно регистрируется за зоной допуска, то происходит ошибка контроля – ложный отказ;

– если в результате генерирования значение  $S_i$  оказалось вне зоны допуска (больше  $S_b$  или меньше  $S_n$ ), а результат измерения прибором случайно – ошибочно регистрируется результат в зоне допуска, то происходит ошибка контроля – необнаруженный отказ;

В процессе контроля с большей вероятностью происходят правильные результаты, но в данной работе внимание концентрируется на рискованных ситуациях.

В девятом и десятом блоках подводятся результирующие оценки с вычислением соответствующих вероятностей по формулам (3.1) и (3.2), где:

$N_{ло}$  – содержимое счетчика ложного отказа;

$N_{но}$  – содержимое счетчика необнаруженного брака;

$N$  – общее количество имитационных повторений.

В одиннадцатом блоке находится интегральный показатель качества контроля надежность  $D$  по формуле (3.4):

$$D = 1 - (P_{ло} + P_{но}) \quad (3.4)$$

Один из известных алгоритмов, который использовался для обучения функционирования совместной работы с алгоритмом (рисунок 3.4) состоял в следующем:

– осуществляется экспертиза всех параметров модели, приведенной на рисунке 3.3, и некоторых элементов входного вектора и получают результат по формуле (3.5):

$$y_i = \omega_i \cdot MF(x_i), \quad (3.5)$$

где  $\omega_i$  – вес  $i$ -го объекта;  $MF(x_i)$  – лингвистический моделируемый результат;

– пороговая функция либо вычисляется, или задается из вне и обрабатывается, как это указывается в алгоритме 3.4. Осуществление контрольно-измерительных функций в системе экологического контроля реализуется с использованием различных инструментальных средств, рассмотрение которых в настоящем исследовании не предусматривается.

Как следует из приведенного обзора и соответствующего анализа, в доступных литературных источниках, довольно широко представлены исследования рисков во многих сферах человеческой деятельности. В

интересующей нас области – экологии и здравоохранении, рассматривается только качественная сторона рисков ситуаций. Количественная оценка рисков если и предпринимается, то только на уровне общих вероятностных подходов. Вместе с тем, в условиях рыночной экономики и внедрения системы качества проблема оценки и прогнозирования рисков стала крайне актуальной, требующей скорейшего решения.

Системные предпосылки диагностирования качества принятия решений в экологии здравоохранении состоят в следующем: лечебно-диагностическому процессу вне зависимости от сложности клинического случая, профессиональной компетентности персонала и используемой технологии сопутствуют ошибки в принятии решений; принятие решений, как уже было указано в предыдущих разделах, является неотъемлемой процедурой человеческой деятельности. Ошибочные решения являются конечным результатом в рискованной ситуации.

В медицинской практике, как правило, рассматриваются только риски пациента. Но в математической статистике известны две категории рисков, тесно связанных между собой, – ошибки (риски) 1-го и 2-го рода. Так как риск является следствием процесса принятия решений в условиях неопределенности и его природа вероятностная, то должна присутствовать и вторая составляющая риска. Этим рискам дается определение «ложного» и «необнаруженного» брака и соответственно вероятности указанных видов брака обозначены как  $P_{лб}$  и  $P_{нб}$ .

В медицине все несколько сложнее, например, если под ложным браком понимать случай, когда здорового человека диагностируют как больного, то страдает пациент, а также, если больного диагностируют как здорового, то и в этом случае опять страдает пациент.

В случае управленческих ошибок хозяйственного или финансового характера последствия будут иные, поэтому имеет смысл сохранить терминологию и обозначения в разрабатываемых аналитических выражениях, а потери от неправильных решений оценивать в каждом случае индивидуально с учетом специфики. Качество процесса принятия решений следует рассматривать в контексте деловых процессов лечебно-диагностического учреждения как системный процесс, который можно исследовать как в целом (интегрировано), так и при необходимости разлагать его на простые составляющие, то есть, подвергать его декомпозиции до нужного уровня. При этом возникают как минимум две задачи: во-первых, клинике необходимо оценить текущий уровень качества и рисков; во-вторых, прогнозировать данные показатели на определенный период или в определенных сочетаниях внешних и внутренних условий. Принятие решения, если не рассматривать последствия, является заключительной процедурой на каждом шаге или этапе технологии управления и в значительной степени определяет уровень риска.

Контроль является неотъемлемым и ключевым процессом любой системы управления. Важнейшими критериями качества контроля являются точность и достоверность. Вопрос о «мерах точности» имеет равнозначное отношение и к результатам измерений, и к практической эффективности реализуемых решений.

Так как на практике и системах интеллектуальной обработки экологической информации данные имеют более нечеткую, чем статистическую форму, а также более употребительны и применимы лингвистические понятия и оценки, была разработана имитационная модель и нечеткий алгоритм для указанных целей, позволяющий оценивать риски ошибок измерений атмосферного загрязнения.

### 3.2 Интеллектуальная обработка данных

При разработке информационной системы было проанализировано большое количество литературы, а также существующие стандарты по информационным технологиям. Как установлено выше, многие экологические системы типа «экология – здоровье» включают процессы, которые еще малоизвестны, и для которых нет формальных моделей до настоящего времени. Поскольку последствия изменений экологической системы или работа в неблагоприятных условиях может быть серьезной проблемой для здоровья населения, есть большая потребность в новых знаниях.

Доступны большие объемы информации, но при этом требуется логический и качественный анализ больших массивов данных, генерируемых антропогенной средой. Подходы, выходящие за рамки простого применения обычных классических техник, необходимы для системного решения проблемы экологической безопасности, и как следствие, здоровья людей.

При современных объемах данных и динамики их качественного изменения, без современных методов интеллектуального анализа данных, обеспечивающих высокую достоверность решить указанную проблему невозможно. Эффективные инструменты для извлечения полезной информации из больших баз данных и способных идентифицировать и фиксировать ключевые факторы, управляющие этими сложными системами используют искусственный интеллект.

Максимальной эффективности в решении рассматриваемой проблемы, как показывают исследования, достигаются при помощи нейросетевых технологий и математического обеспечения на базе теории графов. Один из вариантов подобного решения представлен на рисунке 3.5 [79, р. 3739].

В данной модели функциональные агенты имеют следующие обозначения:

- (X(H)) – накопитель данных;
- (A) – агент анализа данных;
- (ПР) – агент принятия решений;
- (O) – генератор отчетов;
- (K) – агент контроля.

Соединения обозначают информационные связи (дуги). Условно агенты (A), (ПР), (K) реализуют основные функции управления и определяют качество принятия решений и последующих рисков.

В данной системе дуги имеют следующие обозначения: (X(H), O), (X(H), A)), (X(H), K)), (O, ПР), (O, A), (O, K)), (A, K), (A, ПР)), (K, A) и (K, ПР)).

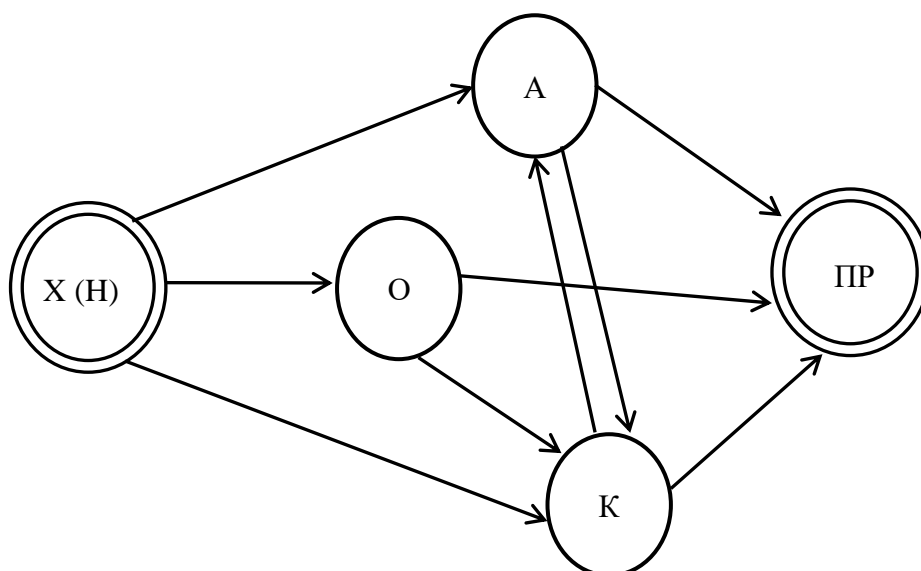


Рисунок 3.5 – Модель анализа и принятия решений

Процесс классификации является необходимой функцией нейронного решения. Данный процесс содержит процедуры дифференциации множества объектов на классы. Для этой процедуры нужны критерии. Объектом классификации в нейронной модели являются данные.

Формальная постановка задачи классификации выглядит как:

$$X_{pj} = \{p_1, p_2, \dots, p_n, \dots, p_m, y\}. \quad (3.6)$$

Показатель  $y$  определяется множеством значений заранее определенных:

$$S = \{s_1, s_2, \dots, s_r, \dots, s_k\}. \quad (3.7)$$

Эта задача может быть реализована известными математическими выражениями в виде:

$$y_j = \omega_0 + \omega_1 p_1 + \omega_2 p_2 + \dots + \omega_m p_m. \quad (3.8)$$

Весомости  $\omega_i$  оцениваются экспертно:

$$a_i = \frac{\sum_{j=1}^m \frac{B_{ij}}{B_{cj}}}{m}, \quad (3.9)$$

где  $j$  – номер эксперта;

$m$  – количество экспертов в группе;

$B_{ij}$  – балл, присвоенный  $i$ -му параметру  $j$ -м экспертом;

$B_{ci}$  – сумма баллов, присвоенных  $j$ -м экспертом всем параметрам объекта.

Сначала каждый эксперт находит соотношение между параметрами попарно. Если весомость данного параметра, по мнению эксперта, выше



другого, с которым сравнивается данный параметр, ему присваивается два балла. Если весомость параметров одинакова, данному параметру присваивается один балл. И если весомость данного параметра ниже другого, то первому параметру баллов не дается.

Наиболее часто уровень компетенции эксперта определяется самими экспертами в бальной форме:

$$a_i = \frac{\sum_{i=1}^m K_i(x_j) i}{\sum_{i=1}^m K_i}, \quad (3.10)$$

где  $(x_j)_i$  – относительная оценка  $j$ -го фактора  $i$ -м экспертом.

Проверка согласованности мнений экспертов проводится до начала процесса определения важности рассматриваемых параметров. Цель проверки – выяснить, являются ли экспертные оценки достоверными, либо их нельзя считать таковыми в силу большого расхождения мнений экспертов. Для проверки согласованности можно использовать коэффициент ранговой корреляции Спирмена, или коэффициент конкордации Кендела.

Второй важной функцией является кластеризация, которая позволяет определить логически и качественно однородные данные [102]. При построении нового множества кластеров из вектора исследуемой информации необходимо сформировать множество критериев, по которым будет осуществляться фильтрация данных и проверяться на близость к тем или иным подмножествам. Используется довольно часто «метод ближайшего соседа», метод опорных векторов [103, 104]. В работе автор предлагает использовать для временной селекции подгрупп информации, например, из вектора экологических данных, и из вектора данных по заболеваниям, коэффициент взаимной корреляции между указанными двумя векторами. На временной функции взаимокорреляции будут наблюдаться «микробифуркации», которые будут определять качественные изменения в потоке информации.

Загрязнение окружающей среды может повлиять на здоровье любым из следующих способов:

- интенсивное воздействие токсичных материалов или агентов, вызывающих немедленную острую реакцию и последствия для здоровья;
- низкий уровень воздействия токсичных материалов или агентов, которые могут вызвать острые или хронические заболевания спустя долгое время после контакта;
- воздействия, которые могут вызвать генетические изменения;
- пониженная сопротивляемость инфекциям;
- вызывать субклиническое раздражение, неприятные ощущения или дискомфорт;
- способствуют обострению имеющегося заболевания;
- создавать условия, несовместимые или унизительные для достижения физического, психического и социального благополучия.

Трудно четко различить последствия загрязнения воздуха от различных источников, и важно учитывать влияние смесей веществ. Эффект от таких смесей может быть просто аддитивным, в большинстве случаев часто обнаруживают, что они синергичны, хотя иногда они могут быть антагонистическими. Для характеристики состояния здоровья населения использованы следующие критерии – понятия: «качество здравоохранения», «благополучие населения», «качество воздуха».

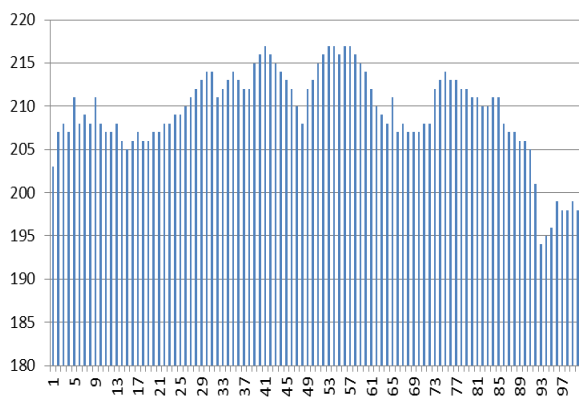
Во многих системах экологического мониторинга отсутствует интеллектуальный анализ данных, а компьютерная обработка носит выборочный, локальный характер. Осуществление текущего параметрического контроля не предусматривает привлечение специальных математических методов, актуальной задачей является объединение разнотипных данных с целью изучения целостной картины данных обработки и формирования единого информационного пространства при мониторинге атмосферного воздуха.

### **3.3 Алгоритм предобработки данных для интеллектуального анализа**

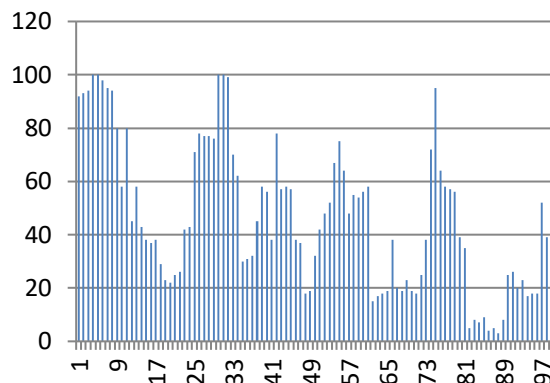
Методика выбора, обоснования, и особенно оптимизация нормативов, как таковой общепринятой, а в данной предметной области в частности, всегда представляла крайне сложную задачу, не поддающуюся формализации. Существует несколько методов решения этой проблемы, которые в каждой задаче требуют нетиповых подходов. В нейронных технологиях часто используется экспертный подход. Одним из самых эффективных способов считается статистический, который в некоторой мере присутствует и в экспертных оценках [105]. Причем, как правило, по истечению определенного времени указанные нормировки пересматриваются в сторону ужесточения. Процесс нормирования в нейронной сети необходим, для реализации функции сравнения результата «суммирования» с порогом и выработки управляющей реакции устройства [106].

Для подбора наиболее оптимального способа нормировки проведем экспериментальную нормировку усредненного взвешенного вредных выбросов в атмосферу и соответствующее им число прироста заболеваний (рисунок 3.6), которые готовим для подачи на входы нейронной сети. Эти данные являются разнородными, так как усредненная величина выбросов изменяется в диапазоне от 196 до 216, а значения прироста индикатора заболеваемости от 0 до 100 пунктов.

На рисунке 3.6 демонстрируется средневзвешенный метод оптимального выбора на примере оценки влияния загрязнения атмосферы на прирост заболеваний с использованием нейронной сети.



а

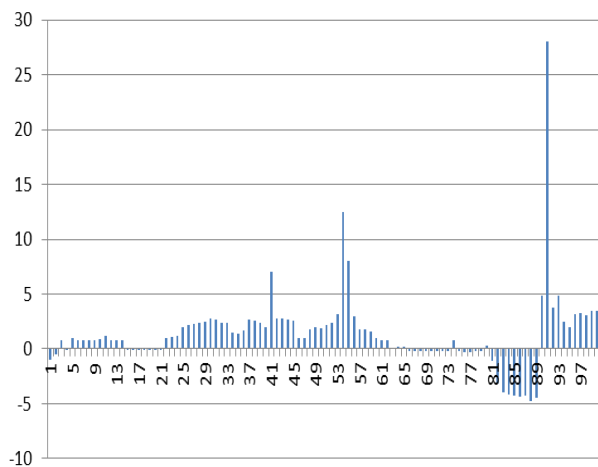


б

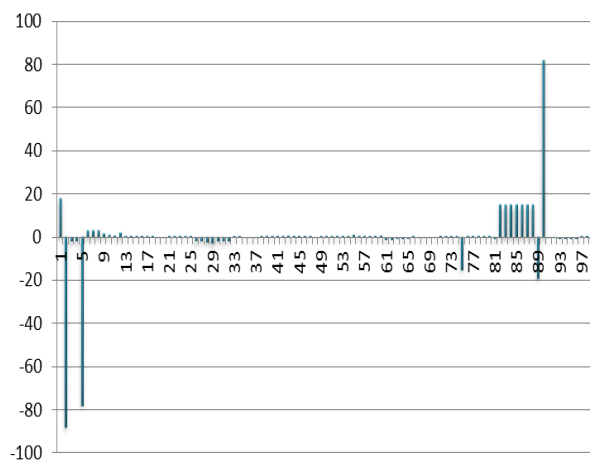
а – ежедневные значения усредненной величины выбросов; б – индикатор прироста заболеваний

Рисунок 3.6 – Ежедневные значения усредненной величины выбросов (слева) и индикатор прироста заболеваний (справа) с 1.02.2015 по 11.05.2015

В данном случае использовалась некоторая функция активации нейрона, с целью выбора параметрических коэффициентов и анализа реакции системы (рисунок 3.7).



а



б

а – ежедневные значения усредненной величины выбросов; б – индикатор прироста заболеваний

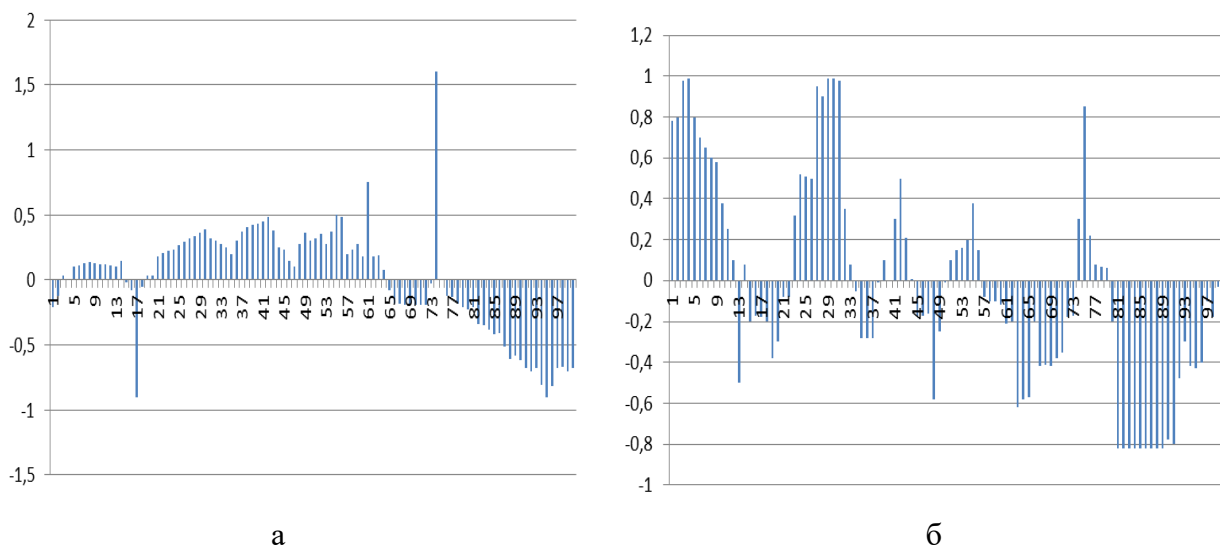
Рисунок 3.7 – Нелинейное нормирование усредненной величины

Для устранения ненормируемых возмущений используется специальный алгоритм:

- осуществит селекцию по МАХ среди данных входного вектора;
- оценить значение параметра  $Y$  по формуле (3.11):

$$\alpha = \frac{y_{max} \cdot 4}{\pi}. \quad (3.11)$$

Входной вектор ранжируется по убыванию нормированных и центрированных данных в интервале  $[-1;1]$  (рисунок 3.8).



а – ежедневные значения усредненной величины выбросов; б – индикатор прироста заболеваний

Рисунок 3.8 – Нелинейное нормирование значения усредненной величины выбросов

Нормированные и центрированные данные прироста заболеваний становятся нормативной базой в системе нейросетевого анализа (рисунок 3.9).

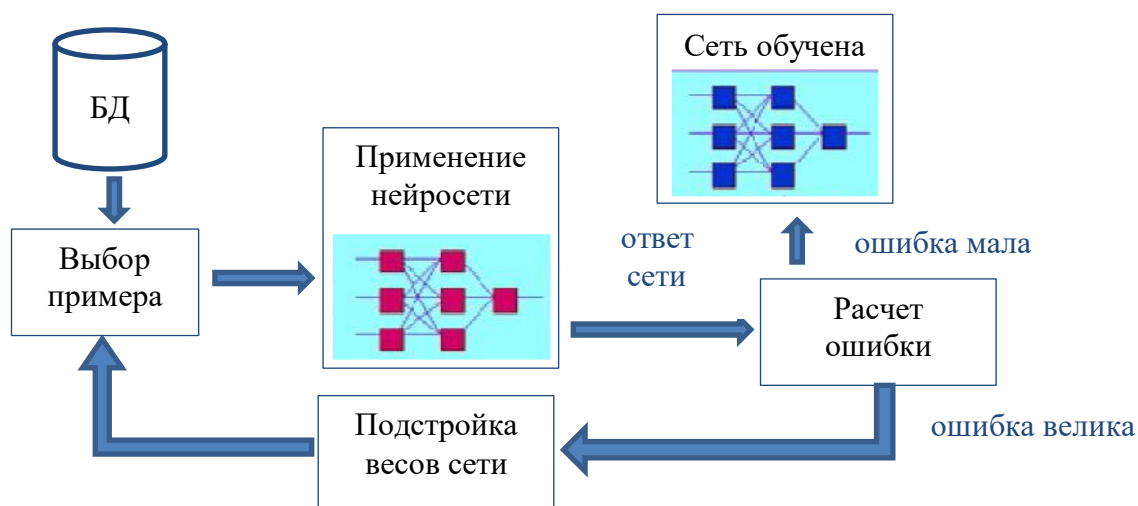


Рисунок 3.9 – Концептуальная модель процесса обучения нейросети

Таким образом, разработана модель нейронной сети, позволяющая оценить и спрогнозировать степень влияния загрязнения воздуха на заболеваемость.

### **Выводы по третьему разделу**

Разработана имитационная модель и алгоритм количественной оценки качества инструментального контроля в системе мониторинга воздушной среды города.

Разработана нейросетевая модель анализа данных для процесса принятия решений в многопараметрической среде.

Разработана методика интеллектуальной обработки данных измерений загрязнения атмосферного воздуха, включающая имитационную модель оценки рисков ошибок измерений и нейросетевую модель оценки влияния загрязнения воздуха на здоровье населения.

## 4 ОПИСАНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ МОНИТОРИНГА АТМОСФЕРНОГО ЗАГРЯЗНЕНИЯ

### 4.1 Общее описание и концептуальная основа интеллектуальной информационной системы

Современные научные направления, связанные с исследованием атмосферного загрязнения промышленного города, в настоящее время сложно представить без использования надёжных, точных и быстрых методов интеллектуального анализа данных и прогнозирования. Использование технологий интеллектуального анализа данных дают возможность для извлечения новых знаний из накапливаемых наборов данных. Такая подход способствует быстрому принятию решений и оценки текущей экологической ситуации.

Аналитические системы на основе технологий интеллектуального анализа данных успешно применяются в различных ГИС. Выполнение аналитической работы без использования методов и алгоритмов интеллектуального анализа данных, только человеком может оказаться трудоемким и длительным. Тем более, скрытые закономерности в данных не всегда могут быть обнаружены человеком визуально.

На рисунке 4.1 изображена концепция разработки интеллектуальной информационной системы.

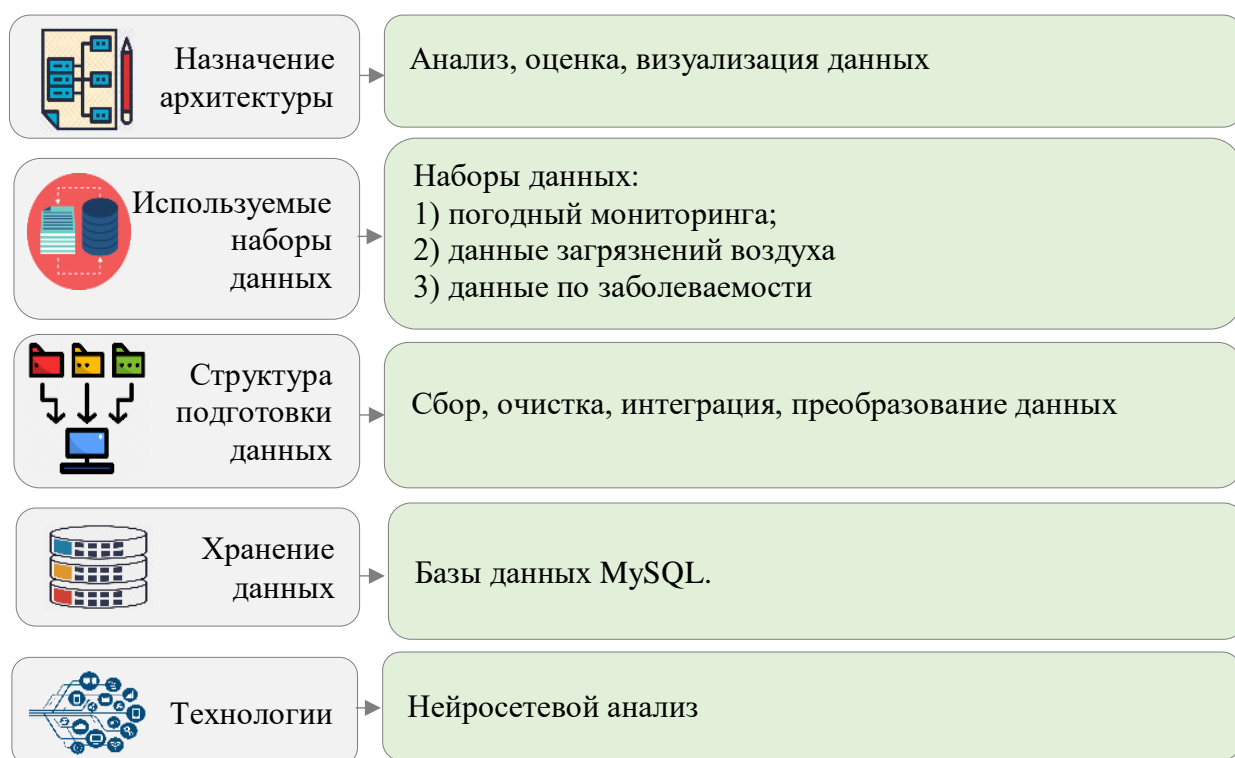


Рисунок 4.1 – Концепция разработки ИИС

В рамках диссертационной работы была разработана архитектура интеллектуальной информационной системы оценки негативного влияния атмосферного загрязнения на здоровье населения. В отличие от уже

существующих информационных систем, данная архитектура является экономической, оперативной и открытой.

Концепция данной интеллектуальной информационно системы поддерживает следующие возможности:

- сбор и хранение необработанных данных;
- хранение многих типов данных в одном хранилище данных;
- выполнение преобразований данных;
- выполнение аналитических операций на основе модели оценки риска снижения качества здоровья;
- интеллектуальная обработка данных на основе нейросетевой модели оценки влияния промышленных выбросов на здоровье.

Спроектированная архитектура предназначена для создания и реализации интеллектуальной информационной системы оценки влияния промышленных выбросов на здоровье населения.

Данная архитектура включает в себя методы и алгоритмы интеллектуального анализа данных мониторинга атмосферного воздуха в виде многофакторной и имитационной модели.

#### **4.2 Архитектура интеллектуальной информационной системы**

На рисунке 4.2 представлена архитектура интеллектуальной информационной системы поддержки принятия решений.

На первом уровне архитектуры ИИС осуществляется загрузка и хранение данных с различных источников: данные погодного мониторинга (Казгидромет), данные загрязнения воздуха (департамент экологии), данные по заболеваемости (департамент здравоохранения).

По средствам технологии ETL разнотипные разноформатные данные подготавливаются для загрузки в хранилище данных MySQL. По средствам SQL-запросов происходит выборка данных, необходимых для интеллектуальной обработки.

На втором уровне архитектуры ИИС происходят вычисления, и интеллектуальная обработка данных, здесь осуществляется предобработка данных для последующего интеллектуального анализа, затем строится модель, модель проходит кросс-проверку, и оценивается.

На третьем уровне архитектуры ИИС представлен пользовательский интерфейс разработанного программного приложения. Пользователю доступны пять вкладок: экологические посты, экологический мониторинг, погодный мониторинг, данные по заболеваемости, нейросеть и прогнозирование.

Посредством работы с программным приложением EnvironmentalDataAnalysis пользователь может осуществлять мониторинг и анализ полученной экологической информации и подготавливать соответствующие отчеты.

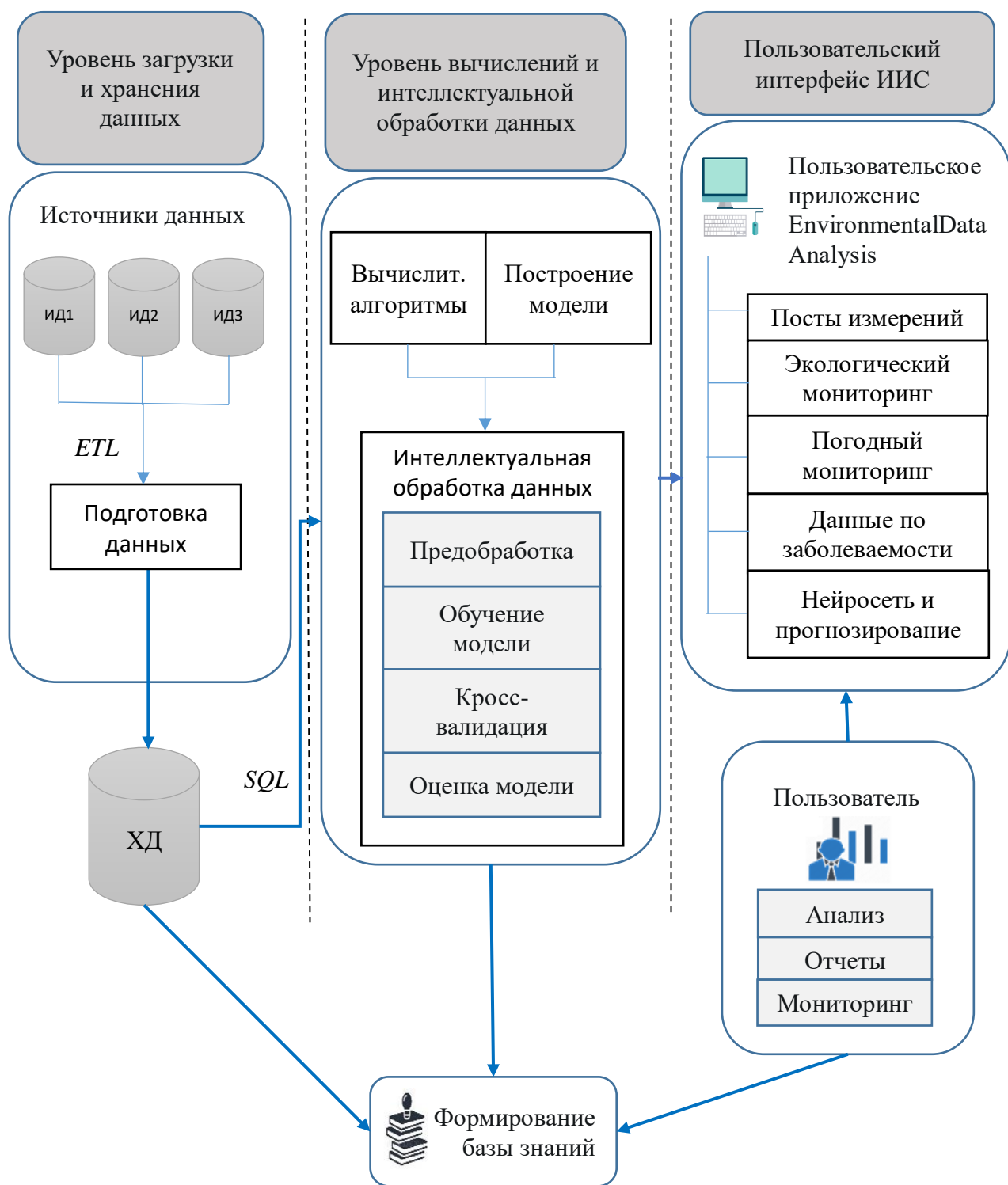


Рисунок 4.2 – Архитектура интеллектуальной ИС

На каждом уровне архитектуры ИИС происходит формирование базы знаний интеллектуальной информационной системы, включающей проблемно-функциональный процессор и интегрированную базу знаний.

Схема формирования базы знаний интеллектуальной ИС представлена на рисунке 4.3.



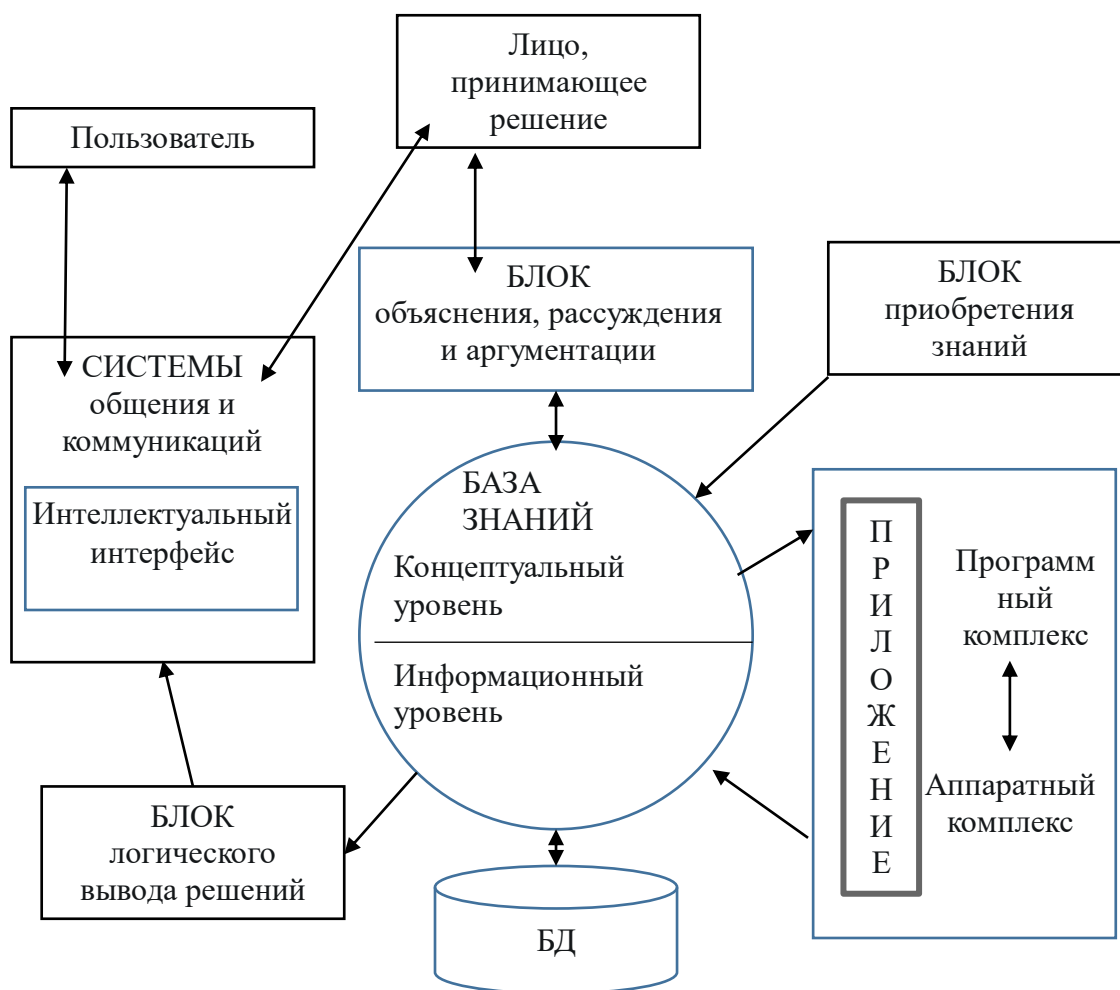


Рисунок 4.3 – Схема формирования базы знаний

Концептуальный уровень базы знаний представлен методами интеллектуальной обработки данных, в том числе работа нейронной сети, используемые модели и алгоритмы.

### 4.3 Программная реализация результатов исследования

Основные технические характеристики программного приложения следующие: системные требования – Microsoft Windows 10 любой версии, процессор x64, x86, разрешение экрана не менее 1536 на 864 пикселей при масштабе 100%.

Запуск программного продукта осуществляется нажатием на иконку EnvironmentalDataAnalysis.

Доступен выбор одной из пяти основных вкладок программного продукта:

1. Экологические посты.
2. Экологический мониторинг.
3. Погодный мониторинг.
4. Данные по заболеваемости.
5. Нейросеть и прогнозирование.

На вкладке «Экологические посты» имеется возможность просмотреть данные с постов экологического мониторинга города Усть-Каменогорск. На постах ведется мониторинг следующих вредных компонентов:

- углекислый газ CO<sub>2</sub>;
- монооксид углерода CO;
- углеводороды C<sub>x</sub>H<sub>y</sub>;
- хлороводород HCl;
- формальдегид HCOH;
- фтороводород HF;
- диоксид азота NO<sub>2</sub>;
- диоксид серы SO<sub>2</sub>.

Имеется возможность использования данных с девяти экологических постов: Промзона; станция Защита; УК Вокзал; КШТ; понтонный мост; поселок Мирный; объездная; Согра; Абая.

В окне отображается информация осуществления/неосуществления мониторинга за вредными компонентами (рисунок 4.4).

EnvironmentalDataAnalysis

Экологические посты   Экологический мониторинг   Погодный мониторинг  

Пост экологического мониторинга	Контроль Cl2	Контроль CO	Контроль CxHy	Контроль HCl	Контроль HCOH	Контроль HF	Контроль NO2	Контроль SO2
Промзона	Нет	Да	Да	Да	Да	Нет	Да	Да
Защита	Нет	Да	Да	Да	Да	Нет	Да	Да
Вокзал	Нет	Да	Да	Да	Да	Нет	Да	Да
КШТ	Нет	Да	Да	Да	Да	Нет	Да	Да
Понтонный	Нет	Да	Да	Да	Да	Нет	Да	Да
Мирный	Нет	Да	Да	Да	Да	Нет	Да	Да
Объездная	Нет	Да	Да	Да	Да	Нет	Да	Да
Согра	Да	Да	Да	Нет	Да	Нет	Да	Да
Абая	Нет	Да	Да	Нет	Да	Да	Да	Да

Рисунок 4.4 – Экологические посты

Вкладке «Экологический мониторинг» позволяет просмотреть данные экологического мониторинга и строить графики загрязнения воздуха.

В левой части окна отображается дата и время замера, пост экологического мониторинга, а также содержание вредных компонентов в (рисунок 4.5).

Загрузить данные

Дата и время замера	Пост экологического мониторинга	Cl2	CO	CxHy	HCl	HCOH	HF	NO2	SO2
19.03.2021 9:48:00	Промзона	0,000	1,520	13,000	0,037	0,000	0,000	0,069	0,013
19.03.2021 9:48:00	Защита	0,000	3,164	53,000	0,000	0,000	0,000	0,015	0,109
19.03.2021 9:48:00	Вокзал	0,000	1,133	0,000	0,000	0,000	0,000	0,054	0,095
19.03.2021 9:48:00	КШТ	0,000	4,141	33,250	0,000	0,000	0,000	0,029	0,000
19.03.2021 9:48:00	Понтонный	0,000	0,963	66,500	0,000	0,000	0,000	0,022	0,125
19.03.2021 9:48:00	Мирный	0,000	2,891	0,000	0,000	0,000	0,000	0,091	0,130
19.03.2021 9:48:00	Объездная	0,000	0,830	33,250	0,000	0,000	0,000	0,059	0,001
19.03.2021 9:48:00	Согра	0,009	1,141	1,055	0,000	0,000	0,000	0,049	0,000
19.03.2021 9:48:00	Абая	0,000	0,834	20,875	0,000	0,000	0,006	0,000	0,000

Рисунок 4.5 – Левая часть окна с данными экологического мониторинга

В правой части окна имеется возможность выбора, по каким вредным веществам и с каких постов экологического мониторинга строить график с помощью установки флагов. При нажатии на кнопку «Построить график», осуществляется визуализация данных (рисунок 4.6).

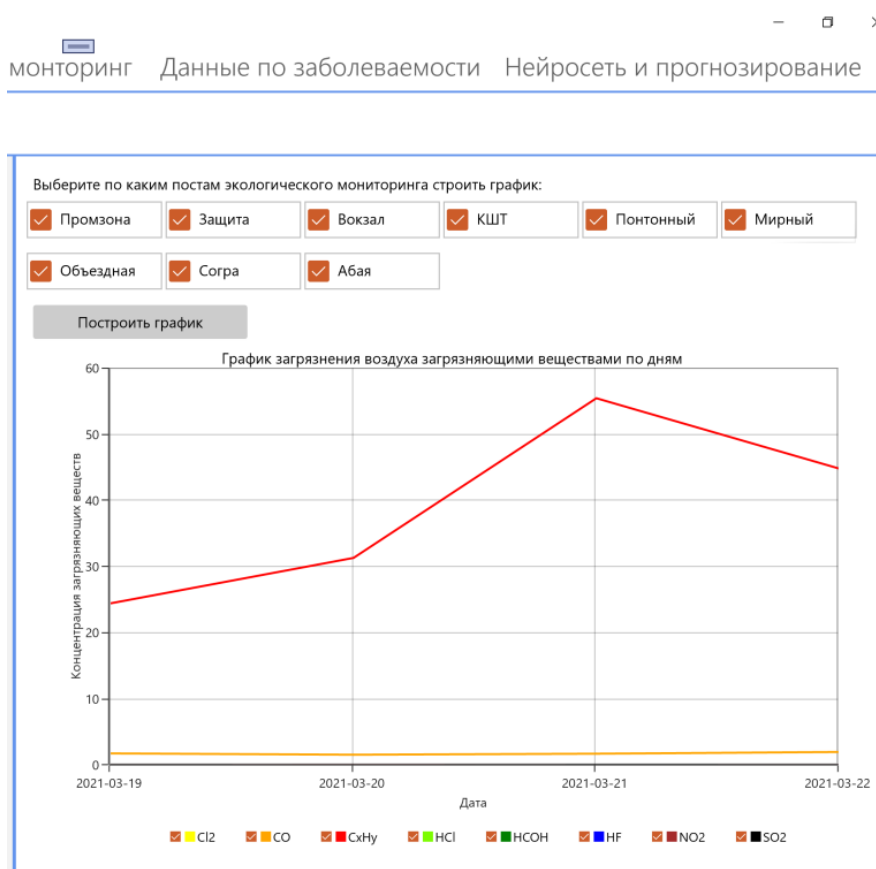


Рисунок 4.6 – Правая часть окна с данными экологического мониторинга

На вкладке «Погодный мониторинг» отображается информация метеорологической службы Казгидромет, включающая дату, минимальную и максимальную температуру воздуха, атмосферное давление, направление и скорость ветра.

В левой части окна отображаются данные погодного мониторинга, а в правой части окна – график колебаний температуры по дням (рисунки 4.7, 4.8).

EnvironmentalDataAnalysis

Экологические посты Экологический мониторинг Погода

Загрузить данные

Дата	Минимальная температура, °С	Максимальная температура, °С	Атмосферное давление, мм рт. ст.	Направление ветра	Скорость ветра, м/с
01.03.2021	-6	1	739	ЮВ	6
02.03.2021	-2	1	736	В	6
03.03.2021	-15	2	742	ЮЗ	2
04.03.2021	-20	-4	743	В	8
05.03.2021	-7	2	745	З	6
06.03.2021	-9	2	741	ЮВ	2
07.03.2021	-9	2	740		0
08.03.2021	-11	7	738	Ю	6
09.03.2021	-8	2	743		0

Рисунок 4.7 – Левая часть окна с графиком колебаний температуры по дням

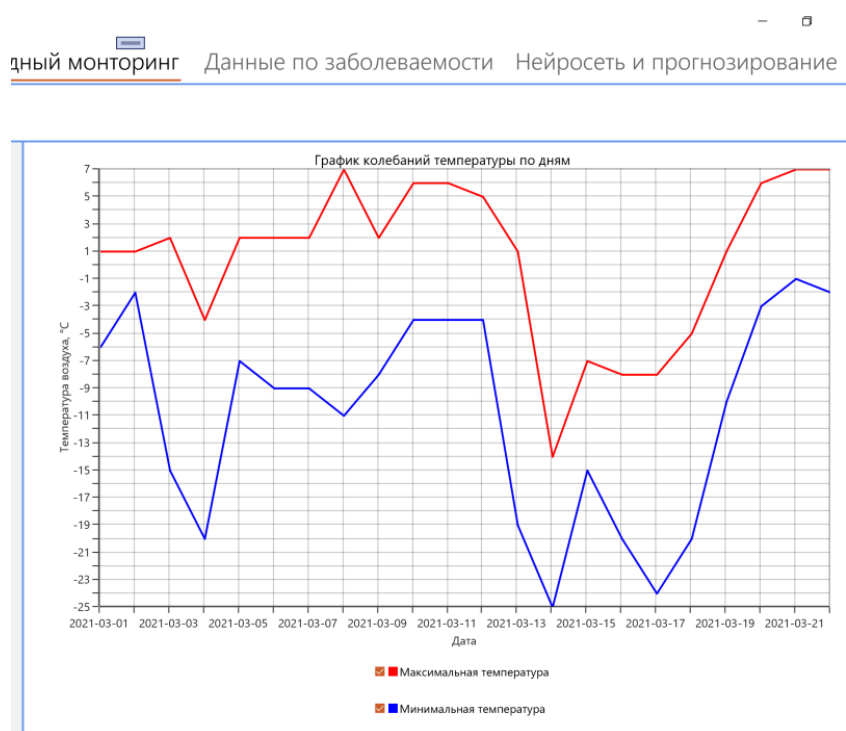


Рисунок 4.8 – Окно с графиком колебаний температуры по дням

На вкладке «Данные по заболеваемости» представлена информация о дате, количестве заболевших в разрезе взрослых и детей и общее количество заболевших. Для примера взяты данные по заболеванию острыми респираторными вирусными инфекциями с симптоматикой заболевания дыхательных путей.

В правой части окна отображаются графики заболевших по дням в зависимости от установки соответствующих флагов: взрослые, дети и общее количество человек (рисунки 4.9, 4.10).

Загрузить данные			
Дата	Количество заболевших ОРВИ взрослых	Количество заболевших ОРВИ детей	Общее количество заболевших ОРВИ
01.03.2021	11	17	28
02.03.2021	12	15	27
03.03.2021	13	20	33
04.03.2021	15	23	38
05.03.2021	18	25	43
06.03.2021	20	28	48
07.03.2021	21	30	51
08.03.2021	20	27	47
09.03.2021	18	24	42
10.03.2021	19	25	44
11.03.2021	18	26	44
12.03.2021	16	23	39
13.03.2021	13	20	33
14.03.2021	12	18	30
15.03.2021	10	17	27

Рисунок 4.9 – Окно с графиками заболевших по дням

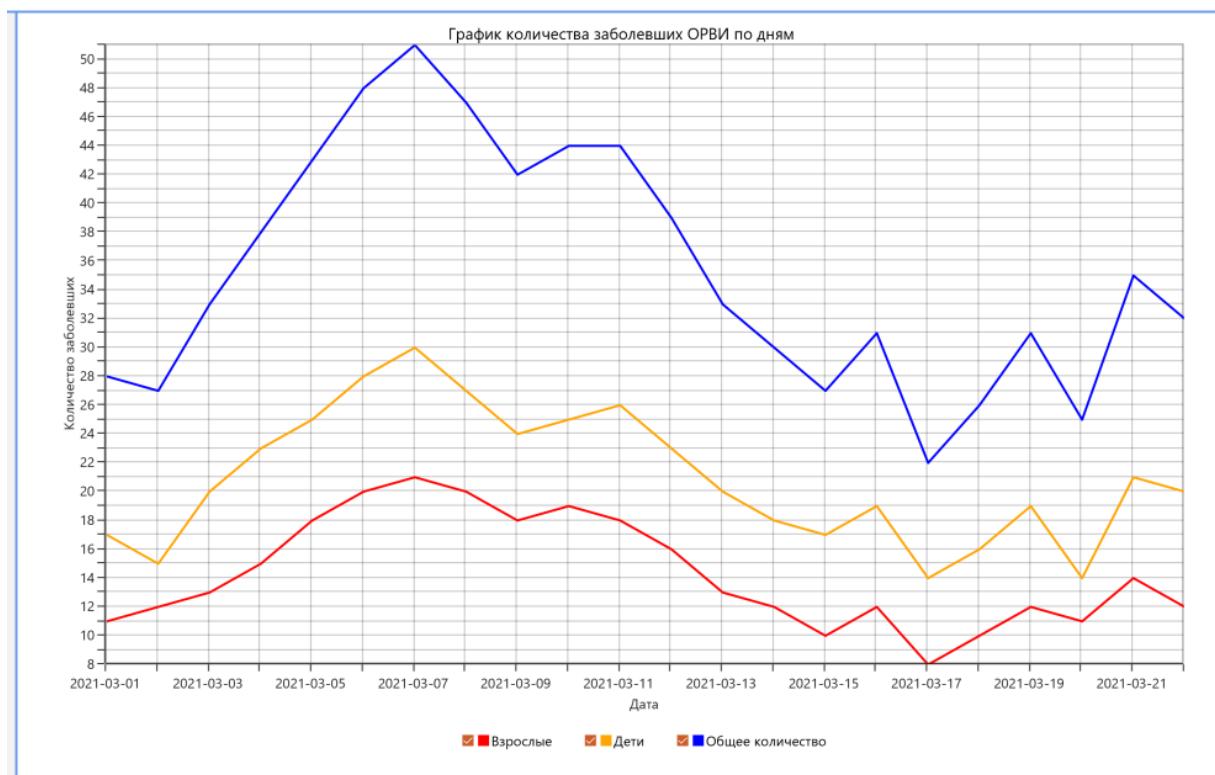


Рисунок 4.10 – Окно с графиками заболевших по дням

На вкладке «Нейросеть и прогнозирование» реализованы методы прогнозирования, построения и обучения нейронной сети. Обучение модели проводится в несколько этапов (рисунок 4.11). На первом этапе на входы модели подаются усредненные данные по загрязняющим веществам по постам экологического мониторинга, выбранным для построения графика. Выходами модели могут быть по выбору количество взрослых заболевших, детей или общее количество заболевших. При нажатии на кнопку «Подготовить набор данных для обучения» в окне «Лог событий» появляется оповещение: «Начата подготовка набора данных для обучения».

При обработке данных используются алгоритмы, описанные в предыдущих главах. По окончании подготовки набора данных в окне лога событий также появляется соответствующее оповещение: «Закончена подготовка набора данных для обучения» и количество загруженных записей. Для второго этапа обучения модели требуется нажать кнопку «Провести обучение модели», при этом будут рассчитаны метрики модели для оценки близости и адекватности моделирования.

Обучение модели	Прогнозирование						
<p><b>Входы модели</b></p> <p>На входы подаются усредненные данные по загрязняющим веществам по постам экологического мониторинга, выбранным для построения графика</p> <p><b>Выход модели</b></p> <p><input type="radio"/> Количество заболевших ОРВИ взрослых за день</p> <p><input type="radio"/> Количество заболевших ОРВИ детей за день</p> <p><input checked="" type="radio"/> Общее количество заболевших ОРВИ за день</p> <p>Подготовить набор данных для обучения</p> <p>Провести обучение модели</p> <p><b>Метрики модели</b></p> <table> <tr> <td>Average RMS</td> <td>4,065</td> </tr> <tr> <td>Average Loss Function</td> <td>6,232</td> </tr> <tr> <td>Average R-squared</td> <td>0,801</td> </tr> </table>	Average RMS	4,065	Average Loss Function	6,232	Average R-squared	0,801	<p><b>Содержание загрязняющих веществ в воздухе</b></p> <p>Хлор Cl2 <input type="text" value="0"/></p> <p>Оксид углерода CO <input type="text" value="1"/></p> <p>Суммарный углеводород CxHy <input type="text" value="30"/></p> <p>Хлористый водород HCl <input type="text" value="0"/></p> <p>Формальдегид HCHO <input type="text" value="0"/></p> <p>Фтористый водород HF <input type="text" value="0"/></p> <p>Диоксид азота NO2 <input type="text" value="0,1"/></p> <p>Диоксид серы SO2 <input type="text" value="0,1"/></p> <p><b>Погодные условия</b></p> <p>Минимальная температура воздуха, °C <input type="text" value="-2"/></p> <p>Максимальная температура воздуха, °C <input type="text" value="5"/></p> <p>Направление ветра <input type="text"/></p> <p>Скорость ветра, м/с <input type="text" value="0"/></p> <p>Выполнить прогноз</p> <p><b>Результат прогнозирования</b></p> <p>Количество заболевших ОРВИ за день <b>35</b></p>
Average RMS	4,065						
Average Loss Function	6,232						
Average R-squared	0,801						

Рисунок 4.11 – Окно прогнозирования

Работоспособность ИИС протестирована на исследованных наборах данных в соответствии с разработанным способом, что показывает точные результаты анализа наборов данных, согласующиеся с экспертными оценками.

Получено свидетельство о государственной регистрации на объект авторского права (программа для ЭВМ) Интеллектуальная информационная система оценки негативного влияния промышленных выбросов на здоровье населения (на примере данных города Усть-Каменогорска) №16777 от 20.04.2021.

### **Выводы по четвертому разделу**

Представлена архитектура интеллектуальной информационной системы поддержки принятия решений в системе экологического мониторинга качества атмосферного воздуха. Архитектура ИИС имеет три уровня:

- 1) уровень загрузки и хранения данных;
- 2) уровень вычислений и интеллектуальной обработки данных;
- 3) уровень пользовательского интерфейса.

На каждом из трех уровней архитектуры осуществляется формирование базы знаний интеллектуальной информационной системы. Приведена схема формирования базы знаний.

Создан программный продукт интегрирующий основные результаты, исследования. При обработке данных с целью моделирования используются алгоритмы, рассчитывающие параметры и режимы принятия решений по результатам моделирования.



## ЗАКЛЮЧЕНИЕ

В ходе выполнения диссертационной работы и проведенного исследования, выполнен ряд задач, и получены следующие результаты.

Проведен анализ инструментов цифровой трансформации в области вопросов экологии и систем экологического мониторинга, показывающий, что в настоящее время актуальны методы и средства интеллектуальной обработки данных.

Проведен обзор существующего математического обеспечения систем мониторинга загрязнения атмосферного воздуха, показывающий, что в большинстве исследований используется математический аппарат, основанный на дифференциальных уравнениях и методах математической статистики, а также обосновывается необходимость в эффективных методах, позволяющих генерировать и накапливать новые знания, путем выявления скрытых закономерностей.

В системе экологического мониторинга возможно наличие некачественных, неточных, искаженных данных, обосновывается необходимость использования специальных интеллектуальных методов и алгоритмов, позволяющих учесть неполноту и зашумленность исходных данных, а также привлечь для анализа качественные экспертные оценки.

Проведен анализ объективных предпосылок к системной интеллектуализации формальных процессов принятия решений в мультипараметрической среде выявлено, что разрабатываемая система в обязательном порядке должна содержать блок количественного оценивания рисков в форме мониторинга.

В рамках диссертационной работы реализованы основные задачи:

– разработано математическое обеспечение системы контроля и принятия решений в процессе экологического мониторинга на основе построения многофакторной модели. Идея подхода состоит в использовании не статичных коэффициентов корреляции, а коэффициентов и функции автокорреляции с автоматически изменяемой начальной точкой в векторе данных и увеличение смещений по лагу до определенного момента;

– разработана методика оценивания и прогнозирования качества здоровья населения, включающая математическое обеспечение системы контроля и принятия решений в процессе экологического мониторинга на основе построения интегрированной многофакторной модели;

– разработана методика интеллектуальной обработки данных измерений загрязнения атмосферного воздуха, включающая имитационную модель оценки рисков ошибок измерений и нейросетевую модель оценки влияния загрязнения воздуха на здоровье населения;

– разработана архитектура интеллектуальной информационной системы оценки негативного влияния промышленных выборов на здоровье населения.

Создан программный продукт, агрегирующий основные результаты исследований. Получено свидетельство о государственной регистрации на объект авторского права (программа для ЭВМ) «Интеллектуальная

информационная система оценки негативного влияния промышленных выбросов на здоровье населения (на примере данных города Усть-Каменогорска)» №16777 от 20.04.2021 (Приложение А). Работоспособность разработанной ИИС проверена на исследованных наборах данных загрязнений атмосферного и данных.

Работоспособность ИИС протестирована на исследованных наборах данных в соответствии с разработанным способом, что показывает точные результаты анализа наборов данных, согласующиеся с экспертными оценками.

Разработанное в рамках выполнения диссертационной работы программное приложение внедрено в производственный процесс ТОО «ПЕТРОЭКОЦЕНТР-Логистики». Акт внедрения прилагается (Приложение Б).

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Национальный проект «Зеленый Казахстан» // <https://primeminister.kz/ru/nationalprojects/nacionalnyy-proekt-zelyonyy-kazahstan-159217>. 01.03.2022.
- 2 Государственная Программа «Цифровой Казахстан»: утв. Постановлением Правительства Республики Казахстан от 12 декабря 2017 года, №827 // <https://adilet.zan.kz/rus/docs/P1700000827>. 01.03.2022.
- 3 Абдрахманова Г.И., Быховский К.Б., Веселитская Н.Н. и др. Цифровая трансформация отраслей: стартовые условия и приоритеты // Матер. 22-й междунар. науч. конф. по проблемам развития экономики и общества. – М., 2021. – С. 11-16.
- 4 Алтухов А.И., Дудин М.Н., Анищенко А.Н. Цифровая трансформация как технологический прорыв и переход на новый уровень развития России // Продовольственная политика и безопасность. – 2020. – Т. 7, №2. – С. 81-96.
- 5 Fayyad U., Piatetsky-Shapiro G., Smyth P. From Data Mining to Knowledge Discovery in Databases // AI Magazine. – 1996. – Vol. 17, №3. – P. 37-54.
- 6 Кречетов Н. Продукты для интеллектуального анализа данных // Рынок программных средств. – 1997. – №14-15. – С. 32-39.
- 7 Киселев М., Соломатин Е. Средства добычи знаний в бизнесе // Открытые системы. – 1997. – №4. – С. 41-44.
- 8 Абдрахманова Г.И., Быховский К.Б., Веселитская Н.Н. и др. Цифровая трансформация отраслей: стартовые условия и приоритеты // Матер. 22-й междунар. науч. конф. по проблемам развития экономики и общества. – М., 2021. – С. 19-27.
- 9 Дюк В.А., Самойленко А.П. Data mining: учеб. курс. – СПб.: Питер, 2001. – 368 с.
- 10 Горбань А.Н. Обучение нейронных сетей. – М.: СССР-США СП "ПараГраф", 1990. – 160 с.
- 11 Minsky M., Papert S. Perceptrons: An Introduction to Computational Geometry. – Cambridge MA: The MIT Press, 1972. – 308 p.
- 12 Классификация и кластер / пер. с англ.; под ред. Д.В. Райзин, Ю.И. Журавлева. – М.: Мир, 1980. – 389 p.
- 13 Ormerod C., Bordes N.S., Pailthorpe V.A. Characterising coupled map lattices // Proceed. of HPC-Asia 2001. – Queensland, 2001. – P. 10-16.
- 14 Иванова А.С., Кузнецов С.П. Волны кластеризации в цепочке систем, каждая из которых содержит набор элементов с внутренней глобальной связью // Известия вузов ПНД. – 2003. – Т. 11, №4-5. – С. 80-88.
- 15 Sarangi S.K., Jaglan V., Dash Y. A Review of Clustering and Classification Techniques in Data Mining // International Journal of Engineering, Business and Enterprise Applications. – 2013. – Vol. 4, Issue 2. – P. 140-145.
- 16 Eberhart R., Simpson P., Dobbins R. (1996). Computational Intelligence PC Tools. – Boston: AP Professional, 1996. – 464 p.
- 17 Хант Э. Искусственный интеллект. – М.: Мир, 1978. – 558 с.
- 18 Galushkin I.A. Neural Networks Theory. – Berlin: Springer-Verlag, 2007. – 396 p.

- 19 Корнев В.А., Макенов А.А. Современные методы моделирования процессов принятия решений в системах управления: монография. – Усть-Каменогорск: Изд-во ВКГУ им. С. Аманжолова, 2008. – 148 с.
- 20 Айдосов А. Прогнозирование распространения сероводорода при аварийных выбросах в атмосферу // Гидрометеорология и экология. – 2002. – №4. – С. 37-46.
- 21 Балова Т.Г., Ипалакова М.Т. Статистические модели и методы прогнозирования уровня загрязнения воздуха // Вестник Восточно-Казахстанского государственного технического университета им. Д. Серикбаева. – 2003. – №4. – С. 20-26.
- 22 Lorenz E.N. Deterministic nonperiodic flow // Journal of the Atmospheric Science. – 1963. – Vol. 20. – P. 130-141.
- 23 Чернавский Д.С., Пирогов Г.Г. и др. Динамика экономической структуры общества // Известия вузов. Прикладная нелинейная динамика. – 1996. – Т. 4, №3. – С. 67-75.
- 24 Малинецкий Г.Г., Потапов А.Б. Современные проблемы нелинейной динамики. – М.: УРСС, 2000. – 336 с.
- 25 Малинецкий Г.Г. Хаос. Структуры. Вычислительный эксперимент. Введение в нелинейную динамику. – М.: УРСС, 2000. – 256 с.
- 26 Ландсберг Г.Е. Климат города. – Л.: Гидрометеиздат, 1983. – 248 с.
- 27 Якушев В.П. На пути к точному земледелию. – СПб., 2002. – 458 с.
- 28 Парк Дж., Маккей С. Сбор данных в системах контроля и управления. – М.: Изд-во: Группа ИДТ, 2006. – 505 с.
- 29 Смородин В.В., Волкова Е.В., Алиев А.А. От хранения данных к управлению информацией. – СПб.: Питер, 2010. – 528 с.
- 30 Берлянд М.Е., Генихович Е.Л. Атмосферная диффузия и структура воздушного потока над неоднородной подстилающей поверхностью // В кн.: Метеорологические аспекты загрязнения атмосферы. – Л.: Гидрометеиздат, 1979. – С. 49-69.
- 31 Вуколов Э.А. Основы статистического анализа. – М.: Форум, 2008. – 464 с.
- 32 Бююль А., Цефель П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей / пер. с нем. – М.: ДиаСофт, 2001. – 608 с.
- 33 Застрогин Ю.Ф. Контроль параметров движения с использованием лазеров: методы и средства. – М.: Машиностроение, 1981. – 176 с.
- 34 Суслов А.Г., Корсакова И.М. Назначение, обозначение и контроль параметров шероховатости поверхностей деталей машин. – М.: МГИУ, 2010. – 112 с.
- 35 Понтрягин Л.С. Математическая теория оптимальных процессов. – Изд. 3-е. – М.: Наука, 1976. – 392 с.
- 36 Davis A. Environment Management // In book: Mastering Salesforce DevOps. – Berkeley, CA: Apress, 2019. – P. 141-201.

37 Green D.G., Klomp N.I., Rimmington G. et al. Digital Ecology: New Technologies Are Revolutionizing Ecology // In book: Complexity in Landscape Ecology. – Cham: Springer, 2020. – Vol. 22. – P. 197-224.

38 Киянский В. Новые системы управления охраной окружающей среды // Промышленность Казахстана. – 2000. – №2. – С. 78-80.

39 ГОСТ 17.4.1.02-83 Охрана природы. Почвы. Классификация химических веществ для контроля загрязнения. – Введ. 01-01-1985. – М., 2008. – 4 с.

40 Kulzhanov M., Rechel V. Kazakhstan: Health system review // Health Systems in Transition. – 2007. – Vol. 9, Issue 7. – P. 1-158.

41 Авалиани С.Л., Андрианова М.М., Печенникова Е.В. и др. Окружающая среда: оценка риска для здоровья (мировой опыт). – М., 1996. – 158 с.

42 Владимиров В.А., Воробьев Ю.Л., Малинецкий Г.Г. и др. Управление риском: риск, устойчивое развитие, синергетика. – М.: Наука, 2000. – 432 с.

43 Белов И.В., Беспалов М.С., Ключкова Л.В. и др. Сравнительный анализ некоторых математических моделей для процессов распространения загрязнений в атмосфере // Математическое моделирование. – 1999. – №7. – С. 45-49.

44 Пененко А.В., Рахметуллина С.Ж. Алгоритмы локализации источников загрязнения атмосферного воздуха на основе данных автоматизированной системы экологического мониторинга // Сиб. электрон. матем. известия. – 2013. – Т. 10. – С. 35-54.

45 Park J., Mackay S. Practical Data Communications for Instrumentation and Control. – Oxford: Elsevier science & technology, 2003. – 400 p.

46 Никитин Д.П., Новиков Ю.В. Окружающая среда и человек: учеб. пос. – М.: Высш. школа, 1980. – 424 с.

47 Берлянд М.Е., Генихович Е.Л., Грачева И.Г. и др. Учет влияния застройки при расчетах загрязнения воздуха // Тр. ГГО. – 1987. Вып. 511. – С. 24-38.

48 Айдосов А.А., Заурбеков Н.С. Математическое моделирование процесса распространения вредных примесей в атмосфере при переменном профиле скорости // Новое в охране труда и окружающей среды: матер. 4-й междунар. науч.-техн. конф. – Алматы, 2000. – С. 232-275.

49 Айдосов А., Турганбаева К.С. Моделирование распространения вредных выбросов в пограничном слое атмосферы нефтегазконденсатдобывающих регионов // Гидрометеорология и экология. – 2002. – №2. – С. 21-33.

50 Жумагулов Б.Т., Абдибеков У.С., Ахметов О.А. Турбулентный перенос примеси в искривленных каналах // Вестник Национальной инженерной академии РК. – 2006. – №1. – С. 24-27.

51 Серкин В.Т. Распространение вредной примеси вблизи границы циркуляционной зоны // Инженерные аспекты экологических проблем Казахстана: межвуз. сб. науч. тр. – Алматы: КазГАСА, 2001. – С. 205-211.

52 Мутанов Г.М., Швец О.Я. Математическое обеспечение трансферта информационных технологий для решения задач экологического мониторинга // Роль вузов в формировании инновационной экономики: матер. 3-й междунар. науч.-практ. конф. – Усть-Каменогорск, 2009. – С. 113-117.

53 Темирбеков Н.М., Мадияров М.Н., Абдолдина Ф.Н. и др. Численное моделирование атмосферных процессов на ограниченной территории и их адаптация для моделирования микроклимата г. Усть-Каменогорска // Вычислительные технологии. – 2006. – Т. 11, ч. 2. – С. 41-45.

54 Sultangazin U.M. Mathematical problems connected with construction of algorithms for atmosphere correction in remote sensing // Journal of Inverse and Ill-posed Problems. – 2001. – Vol. 9, Issue 6. – P. 655-668.

55 Жижимов О.Л., Федотов А.М., Шокин Ю.И. Технологическая платформа массовой интеграции гетерогенных данных // Вестник Новосибирского государственного университета. – 2013. – Т. 11, №1. – С. 24-41.

56 Протасов К.Т., Белан Б.Д., Идрисов З.Ф. Адаптивная идентификация оптико-метеорологических характеристик атмосферы // Оптика атмосферы и океана. – 1988. – Т. 1, №10. – С. 95-98.

57 Ничепорчук В.В., Ноженков А.И., Ноженкова Л.Ф. Система распределенного сбора и анализа данных мониторинга чрезвычайных ситуаций // Информатизация и связь. – 2012. – №5. – С. 45-50.

58 Вавилова Н.Г., Генихович Е.Л., Сонькин Л.Р. Статистический анализ данных о загрязнении воздуха в городах с помощью естественных функций // Тр. ГГО. – 1969. – Вып. 238. – С. 27-32.

59 Ивахненко А.Г., Лапа В.Г. Предсказание случайных процессов. – Киев: Наукова думка, 1971. – 416 с.

60 Box G.E.P., Jenkins G.M. Time Series Analysis, Forecasting and Control. – San Francis-co: Holden-Day, 1976. – 575 p.

61 Авдеенко Н.В., Ефимова А.А., Балоболкин И.И. Влияние загрязнения воздушной среды на распространенность и течение аллергических болезней у детей // Педиатрия. – 1999. – №5. – С. 15-24.

62 Ваганов П.А., Ман-Сунг И. Экологические риски: учеб. пос. – Изд. 2-е. – СПб.: Изд-во Санкт-Петербургского университета, 2001. – 152 с.

63 Annesi-Maesano I., Agabiti N., Pistelli R. et al. Subpopulations at increased risk of adverse health outcomes from air pollution // European Respiratory Journal. – 2003. – Vol. 21. – P. 57-63.

64 Куриленко Е.А. Методика оценки негативного влияния промышленных выбросов на состояние здоровья населения (на примере г. Усть-Каменогорска). – Алматы: КазНТУ им. Сатпаева, 2008. – 139 с.

65 Bakhtiari P.H., Nikoo M.R., Golkar F. et al. Design of a high-coverage ground-based CO<sub>2</sub> monitoring layout using a novel information theory-based optimization model // Environ Monit Assess. – 2021. – Vol. 193. – P. 150.

66 Аманжол И.А., Аманбекова А.У., Баттакова Ш.Б. и др. Современные проблемы экологически зависимых заболеваний населения урбанизированных территорий: монография. – Караганда, 2012. – 200 с.

67 РГП «Казгидромет» может прогнозировать загрязнение атмосферного воздуха // <https://www.kazhydromet.kz/ru/post/378>. 01.03.2022.

68 Mossalam A., Arafa M. Using artificial neural networks (ANN) in projects monitoring dashboards' formulation // *HBRC Journal*. – 2018. – Vol. 14, Issue 3. – P. 385-392.

69 Jorgensen S.E., Bendoricchio G. *Fundamentals of Ecological Modelling*. – Ed. 3rd. – Amsterdam: Elsevier, 2001. – 544 p.

70 El-Qawasma F., Elfouly T., Ahmed M. Minimising number of sensors in wireless sensor networks for structure health monitoring systems // *Wireless Sensor Systems IET*. – 2019. – Vol. 9. – P. 94-101.

71 Yu T., Lin C. et al. Wireless sensor networks for indoor air quality monitoring // *Medical Engineering & Physics*. – 2013. – Vol. 35. – P. 231-235.

72 Малинецкий Г.Г., Подлазов А.В. Парадигма самоорганизованной критичности. Иерархия моделей и пределы предсказуемости // *Известия ВУЗов. Прикладная нелинейная динамика*. – 1997. – Т. 5, №5. – С. 89-106.

73 Waldrop M.M. *Complexity: The emerging science at the edge of order and chaos*. – NY., 1993. – 379 p.

74 Додонов А.Г., Ландэ Д.В. *Живучесть информационных систем*. – Киев: Наукова думка, 2011. – 256 с.

75 Li L., Chu W., Langford J. et al. A contextual-bandit approach to personalized news article recommendation // *Nineteenth International Conference on World Wide Web (WWW 2010)*. – Raleigh, 2010. – P. 1-10.

76 Liben-Nowell D., Kleinberg J. The link-prediction problem for social networks // *Journal of the American Society for Information Science and Technology*. – 2007. – Vol. 58, Issue 7. – P. 1019-1031.

77 Рахметуллина С.Ж., Бакланова О.Е., Бугубаева А.Ж. Влияние атмосферных условий на поведение концентрации вредных веществ в атмосфере региона // *Вестник Восточно-Казахстанского технического университета им. Д. Серикбаева*. – 2019. – №4(86). – С. 41-47.

78 Bugubayeva A., Rakhmetullina S. Changes in the concentration of pollutants in the atmosphere of the region under the influence of atmospheric conditions // *Procced. 68th internat. corresp. scient.-pract. conf. «European research: innovation in science, education and technology»*. – London, 2020. – P. 20-25.

79 Shvets O., Bugubayeva A., Rakhmetullina S. et al. Development of intelligent information system of assessing the negative impact of industrial emissions on the public health // *Journal of Theoretical and Applied Information Technology*. – 2021. – Vol. 99, Issue 15. – P. 3736-3752.

80 Рахметуллина С.Ж., Бугубаева А.Ж. Применение технологии Data Mining с целью прогнозирования загрязнения атмосферного воздуха // *Вестник Восточно-Казахстанского технического университета им. Д. Серикбаева*. – 2018. – №4(82). – С. 177-183.

81 Bugubayeva A., Rakhmetullina S. Application of data mining technology for forecasting air pollution // *Global science and innovations 2018: Central Asia: mater. of the 3rd. internat. scient.-pract. conf.* – Астана, 2018. – P. 84-89.

82 Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных условий. – М.: Наука, 1976. – 254 с.

83 Терехов Л.Л. Экономико-математические методы. – М. Статистика 1988. – 340 с.

84 Baklanov A., Dmitrieva T., Bugubayeva A. et al. Environmental threat calculation dealing with the risk of industrial atmospheric emission // Communications in Computer and Information Science – 2019. – Vol. 998. – P. 23-33.

85 Меньшиков В.В., Швыряев А.А., Захарова Т.В. Анализ риска при систематическом загрязнении атмосферного воздуха опасными химическими веществами: учеб. пос. – М.: Изд-во Химич. фак. Моск. ун-та, 2003. – 120 с.

86 Рахметуллина С.Ж., Бакланова О.Е., Бугубаева А.Ж. и др. Расчет экологической опасности, связанной с риском выбросов крупных промышленных предприятий города Усть-Каменогорск // Вестник Восточно-Казахстанского технического университета им. Д. Серикбаева. – 2020. – №4(90). – С. 144-148.

87 Вентцель Е.С. Исследование операций: задачи, принципы, методологии. – М.: Наука, 1988. – 208 с.

88 Игнатъева Л.П., Чирцова М.В., Потапова М.О. Факторы, влияющие на здоровье: учеб. пос. – Иркутск: ИМГУ. – 2014. – 33 с.

89 Bezmertnyi Y., Shevchuk V., Pavlov S. et al. Prognosis of efficacy of medical and social rehabilitation in disabled individuals with respiratory diseases // Proceed. Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments. – Wilga, 2019. – P. 11176331-11176336.

90 Vitolo C., Scutari M., Ghalaieny M. et al. Modeling air pollution, climate, and health data using Bayesian Networks: A case study of the English regions // Earth and Space Science. – 2018. – Vol. 5. – P. 76-88.

91 Bereziuk O., Lemeshev M., Bogachuk V. et al. Ultrasonic microcontroller device for distance measuring between dustcart and container of municipal solid wastes // Przegląd Elektrotechniczny. – 2019. – Vol. 95, Issue 4. – P. 146-150.

92 Dicken R., Rubby S., Naz S. et al. Analysis and classification of respiratory health risks with respect to air pollution levels // Proceed. 16th internat. conf. on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). – Takamatsu, 2015. – P. 1-6.

93 Cui X., E. S., Niu D. et al. Forecasting of Carbon Emission in China Based on Gradient Boosting Decision Tree Optimized by Modified Whale Optimization Algorithm // Sustainability. – 2021. – Vol. 13. – P. 12302-1-12302-18.

94 Souza F.T., Rabelo W.S. A Data Mining Approach to Study the Air Pollution Induced by Urban Phenomena and the Association With Respiratory Diseases // Proceed. 11th internat. conf. on Natural Computation (ICNC). – Curitiba, 2015. – P. 1045-1050.

95 Yoon H.-J., Xu S., Tourassi G. Predicting Lung Cancer Incidence from Air Pollution Exposures Using Shapelet-based Time Series Analysis // Proceed. internat. conf. «Biomed Health Inform» (IEEE EMBS). – Las Vegas, 2016. – P. 565-568.



- 96 Боровиков В.П., Ивченко Г.И. Прогнозирование в системе STATISTICA в среде WINDOWS. – М.: Финансы и статистика, 2000. – 384 с.
- 97 Корнев В., Сагидуллина Г., Кайсарулы Т. Модели прогнозирования достоверности инструментальной клинической диагностики. – Усть-Каменогорск: Изд-во ВКГТУ, 2020. – 160 с.
- 98 Zadeh L.A. Fuzzy Sets // Information and Control. – 1965. – Vol. 8, Issue 3. – P. 338-353.
- 99 Zadeh L.A. Outline of a New Approach to the Analysis of Complex Systems and Decision Processes // Proceed. IEEE Transactions on Systems, Man, and Cybernetics (SMC-3(1)). – Washington, 1973. – P. 28-44.
- 100 Мазур И.И., Молдаванов О.И., Шишов В.Н. Инженерная экология. Общий курс: в 2 т. – М.: Высш. шк. 1996. – Т. 1. – 637 с.
- 101 Орлов А.И. Статистика объектов нечисловой природы и экспертные оценки // Экспертные оценки: вопросы кибернетики: сб. – М., 1979. – Вып. 58. – С. 17-33.
- 102 Задачи Data Mining. Классификация и кластеризация // <http://www.intuit.ru/department/database/datamining/5/>. 04.11.2020.
- 103 Метод ближайшего соседа // <http://www.aiportal.ru/articles/autoclassification/single-link.html>. 04.11.2020.
- 104 Метод опорных векторов // <http://ru.wikipedia.org>. 04.11.2020.
- 105 Олейви Х.З. Метод экспертных оценок как метод оценки хозяйственного риска на предприятии // Актуальные проблемы гуманитарных и естественных наук. – 2014. – №5. – С. 221-224.
- 106 Strikova T.O., Lytyuga O.P., Skorupski K., Bugubayeva A. Stochastic deterministic methods for processing signals and images in optical electronic systems // Proceed. Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments. – Wilga, 2019. – P. 11176081-11176089.



# ПРИЛОЖЕНИЕ А

## Свидетельство об авторском праве

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ

РЕСПУБЛИКА КАЗАХСТАН



**СВИДЕТЕЛЬСТВО**  
**О ВНЕСЕНИИ СВЕДЕНИЙ В ГОСУДАРСТВЕННЫЙ РЕЕСТР**  
**ПРАВ НА ОБЪЕКТЫ, ОХРАНЯЕМЫЕ АВТОРСКИМ ПРАВОМ**  
№ 16777 от «20» апреля 2021 года

Фамилия, имя, отчество, (если оно указано в документе, удостоверяющем личность) автора (ов):  
**ШВЕЦ ОЛЬГА ЯКОВЛЕВНА, Шокарев Алексей Владимирович, Бугубаева Алина Жанатбековна**

Вид объекта авторского права: **программа для ЭВМ**

Название объекта: **Интеллектуальная информационная система оценки негативного влияния промышленных выбросов на здоровье населения (на примере данных города Усть-Каменогорска)**

Дата создания объекта: **14.04.2021**





Құжат тұңғыш рет қолданылған <http://www.kazpatent.kz/> сайтының "Авторлық құқық" бөлімінде тексеруге болады. <https://copyright.kazpatent.kz>

Подлинность документа возможно проверить на сайте [kazpatent.kz](http://www.kazpatent.kz) в разделе «Авторское право» <https://copyright.kazpatent.kz>

Подписано ЭЦП

Оспанов Е.К.



# ПРИЛОЖЕНИЕ Б

## Акт внедрения

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ

Жауапкершілігі шектеулі  
серіктестігі

**ПЕТРОЭКОЦЕНТР-  
Логистика**

150000, Петропавл қаласы, Горький к-сі, 166  
тел./факс 8 (7152) 50-25-25, 50-30-30, 52-75-52  
e-mail: dkedich@yandex.ru  
www.pec.kz



РЕСПУБЛИКА КАЗАХСТАН

Товарищество с ограниченной  
ответственностью

**ПЕТРОЭКОЦЕНТР-  
Логистика**

150000, г. Петропавловск, ул. Горького, 166  
тел./факс 8 (7152) 50-25-25, 50-30-30, 52-75-52  
e-mail: dkedich@yandex.ru  
www.pec.kz

Исх. № 64 от "24" 12 2021г.

### Акт

внедрения результатов диссертационной работы  
Бугубаевой Алины Жанатбековны  
на тему «Методы и алгоритмы интеллектуального анализа данных для  
системы мониторинга атмосферного загрязнения»,  
представленную на соискание степени доктора философии (PhD) по  
специальности 6D070300 – Информационные системы (по отраслям)

Разработанное Бугубаевой А.Ж., Швец О.Я., Шокаревым А.В. программное приложение «Интеллектуальная информационная система оценки негативного влияния промышленных выбросов на здоровье населения», созданное в рамках выполнения диссертационной работы по теме: «Методы и алгоритмы интеллектуального анализа данных для системы мониторинга атмосферного загрязнения» по специальности 6D070300 – Информационные системы (по отраслям) внедрено в производственный процесс ТОО «ПЕТРОЭКОЦЕНТР-Логистика».

Эффективность внедрения: разработанный алгоритм интеллектуального анализа данных позволяет находить закономерности, которые могут быть с успехом использованы для целей экологического мониторинга. Его применение позволяет оценить влияние загрязнения атмосферного воздуха на здоровье населения. Программное приложение позволяет рассчитать метрики модели для оценки близости и адекватности моделирования. Испытания показали высокую точность прогноза заболеваемости – 98%. Также имеются следующие преимущества: применяемые на сегодняшний день в действующей системе экологического мониторинга методы математической обработки данных обладают высокой погрешностью из-за невозможности учета в используемых моделях всех значимых факторов, определяющих состояние объектов мониторинга. В то время как применяемые в разработанном программном приложении методы интеллектуального анализа данных позволяют устранить недостатки применения классических математических методов при обработке мониторинговой информации.

Директор ТОО «ПЕТРОЭКОЦЕНТР-Логистика»



 Кедич Д.В.

0001167