Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

На правах рукописи

Евсюткин Иван Викторович

ИНТЕЛЛЕКТУАЛЬНАЯ ИНФОРМАЦИОННАЯ СИСТЕМА ДЛЯ УПРАВЛЕНИЯ ФОНДОМ СКВАЖИН НЕФТЕГАЗОДОБЫВАЮЩЕГО ПРЕДПРИЯТИЯ

05.13.01 – Системный анализ, управление и обработка информации
(промышленность)

Диссертация на соискание ученой степени кандидата
технических наук

Научный руководитель
д.т.н., профессор Марков Н. Г.

Томск – 2021

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 5

1. ПРОБЛЕМА АВТОМАТИЗАЦИИ УПРАВЛЕНИЯ ФОНДОМ СКВАЖИН МЕСТОРОЖДЕНИЙ НЕФТИ И ГАЗА 11

1.1. Задачи управления фондом скважин промысла 11

1.1.1. Методы интенсификации добычи УВС 12

1.1.2. Основные бизнес-процессы управления фондом скважин 14

1.1.3. Особенности реализации бизнес-процессов управления ГТМ 16

1.2. Анализ существующих методов и методик управления фондом скважин . 20

1.2.1. Традиционные методы управления фондом скважин 20

1.2.2. Методы ИАД для управления фондом скважин 26

1.2.3. Информационные системы управления фондом скважин 30

1.3. Концепция интеллектуальных месторождений 33

1.4. Цель и задачи диссертационного исследования 36

1.5. Основные результаты и выводы по разделу 38

2. КОНЦЕПЦИЯ СОЗДАНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ 40

2.1. Основные требования к ИИС и принципы её создания 40

2.2. Концепция сервисно-ориентированной архитектуры программного обеспечения ИИС 45

2.3. Проектирование интерфейса пользователей ИИС 50

2.4. Концептуальная модель базы данных ИИС 55

2.5. Основные результаты и выводы по разделу 57

3. МОДЕЛИ, МЕТОДЫ И АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ИИС 59

3.1. Постановки задач исследований 59

3.1.1. Слабоформализуемые задачи управления фондом скважин 59

3.1.2. Методы и подходы к предварительной подготовке исходных данных 61

3.1.3. Выбор типа ИНС для решения слабоформализуемых задач 63

3.1.4. Подходы и алгоритмы обучения ИНС 70

3.1.5. Выбор библиотеки для формирования и обучения ИНС 76

3.2. Модели, методы и алгоритмы для решения задачи прогноза технологических параметров скважин 78

3.2.1. Алгоритмы и методика подготовки данных для их последующего анализа 79

3.2.2. Эксперименты по прогнозу значений дебитов скважин с помощью ИНС 91

3.2.3. Исследование статистических методов прогноза значений дебитов скважин 99

3.3. Модели, методы и алгоритмы решения задач управления ГТМ 103

3.3.1. Предварительная подготовка данных при решении задач классификации скважин и типов ГТМ 104

3.3.2. Разработка моделей ИНС для решения задач классификации скважин и типов ГТМ 108

3.3.3. Результаты исследования эффективности моделей ИНС 111

3.3.4. Методы и алгоритмы кластерного анализа при решении задачи выбора скважин-кандидатов для ГТМ 118

3.4. Оценки эффективности ГТМ, проводимых на фонде скважин 124

3.5. Алгоритм оптимизации плана-графика работы бригад КРС 127

3.6. Основные результаты и выводы по разделу 130

4. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ИИС 133

4.1. Выбор базовых программных средств ИИС 133

4.2. Выбор системы управления БП 134

4.3. Разработка специализированной сервисной шины 136

4.3.1. Современные ESB и интеграционные платформы 136

4.3.2. Особенности специализированной ESB 141

4.3.3. Исследование эффективности специализированной ESB 143

4.4. Реализация бизнес-процессов управления фондом скважин в среде системы ELMA BPM 145

4.4.1. Веб-сервисы, реализованные в системе ELMA BPM 145

4.4.2. Формирование журнала на остановку скважин средствами системы ELMA BPM 148

4.4.3. Реализация в системе ELMA BPM бизнес- процесса «Управление и проведение ГТМ» 151

4.5. Основные разработанные веб-сервисы ИИС 156
4.5.1. Веб-сервисы для решения сложных вычислительных задач 156
4.5.2. Архитектура программного обеспечения ИИС 158
4.6. Основные результаты и выводы по разделу 161
5. ПРИМЕНЕНИЕ ИИС К РЕШЕНИЮ ЗАДАЧ В НЕФТЕГАЗОВОЙ ОТРАСЛИ 163
5.1. Апробация ИИС в нефтегазодобывающем предприятии АО «Газпром добыча Томск» 163
5.1.1. Организация единого информационного пространства АО «Газпром добыча Томск» 163
5.1.2. Результаты опытной эксплуатации ИИС в АО «Газпром добыча Томск» 166
5.2. Результаты внедрения ИИС в ООО «Центр нефтегазовых технологий».. 168
5.2.1. Результаты применения методов ИАД 168
5.2.2. Организация единого информационного пространства ООО «Центр нефтегазовых технологий» 170
5.3. Методика применения ИИС при анализе данных 173
5.3.1. Общие положения 173
5.3.2. Методика применения ИНС при решении слабоформализуемых задач 175
5.4. Основные результаты и выводы по разделу 178
ЗАКЛЮЧЕНИЕ 180
ЛИТЕРАТУРА 182
ПРИЛОЖЕНИЕ 1 Письмо о передаче результатов работы в АО «Газпром добыча Томск» 200
ПРИЛОЖЕНИЕ 2 Акт о внедрении результатов в ООО «Центр нефтегазовых технологий» 201

ВВЕДЕНИЕ

Актуальность темы. Несмотря на резкое падение цен на нефть, нефтегазовая отрасль по-прежнему является основополагающей при формировании бюджета нашего государства. С другой стороны, газ является стратегически важным ресурсом: за счет газификации сел и деревень России, а также поставок зарубежным потребителям, востребованность данного вида сырья только растет [1]. Фонд скважин нефтегазодобывающего предприятия только на одном месторождении может включать несколько сотен и даже тысяч скважин. Столь крупный объект управления требует серьёзного внимания со стороны субъектов управления – служб предприятия с целью формирования и оказания различных управляющих воздействий на скважины фонда и продуктивные пласты для поддержания стабильной добычи углеводородного сырья (УВС).

Управление фондом скважин промысла требует значительных трудовых и временных затрат квалифицированных специалистов добывающего предприятия, которые на сегодняшний день практически вручную анализируют большие объемы разнородных геологических и технологических данных, получаемых при мониторинге продуктивных пластов месторождения и скважин фонда. Отсутствие соответствующих инструментов ведёт к увеличению вероятности возникновения ошибок вследствие человеческого фактора. Часть решаемых специалистами задач при управлении фондом скважин – это слабоформализуемые задачи. Под ними здесь и далее понимается такой класс задач, для решения которых не существует математической постановки задач и формального алгоритмического решения, а если даже и существует, то поиск точного решения займёт столь большое время, что при имеющихся ресурсах это невозможно будет осуществить [2]. Данные в таких задачах имеют очень большой объём и характеризуются неточностью, неоднозначностью, неполнотой и ненадёжностью. Их решение возможно лишь с помощью интуитивно полученных в ходе экспериментов эвристик. При этом происходит поиск наиболее рационального решения, следуя пути отсева неоптимальных решений.

При решении таких задач специалистам необходимы современные инструменты в виде интеллектуальных информационных систем (ИИС), в которых реализованы интеллектуальные методы поддержки принятия решений. Однако существующие сегодня на добывающих предприятиях производственные информационные системы (ИС) являются в большинстве своём узкоспециализированными [3, 4] и в них отсутствуют интеллектуальные методы поддержки принятия решений специалистами при управлении фондом скважин.

Всё это указывает на актуальность разработки новых моделей, методов и алгоритмов интеллектуального анализа данных (ИАД) и программной реализации их в виде перспективных ИИС для высокоавтоматизированного управления фондом скважин в условиях постоянно обновляющихся на промыслах данных [5]. Особенно такие ИИС будут необходимы как средства поддержки принятия решений специалистами при решении слабоформализуемых задач.

Степень изученности проблемы. Сегодня для анализа данных мониторинга скважин и продуктивных пластов специалисты предприятий используют при реализации процессов управления фондами скважин несложные и далеко не точные методы. Наиболее распространённый из них – метод геологического потенциала скважин. В работах отечественных и зарубежных исследователей Султанова А. С., Перминова В. Е., Келлера Ю. А., Кайлинг К., Валеева С. В. и других с использованием методов кластерного анализа и простых моделей искусственных нейронных сетей (ИНС) при анализе геологических и технологических данных при добыче УВС решён ряд частных задач и получены первые обнадёживающие результаты. Однако комплексные исследования точности основных методов ИАД применительно к задачам управления фондом скважин практически отсутствуют. Более того, в рамках концепции интеллектуального месторождения, развиваемой в России коллективом исследователей под руководством академика РАН Дмитриевского А. Н., сегодня только формируются постановки задач комплексных исследований методов ИАД в нефтегазовой отрасли.

Развитию архитектуры и программного обеспечения (ПО) производственных ИС добывающих предприятий, в которых реализованы частные перспективные методы и алгоритмы ИАД, посвящены работы Аршиновой Н. М., Ахмедова К. С., Кудинова А. В., Маркова Н. Г., Джонса Дж., Формана Р. и других. Анализ показал, что многие из этих ИС являются узкоспециализированными, имеют устаревшую архитектуру и малопригодны для интеграции с унаследованными ИС предприятий.

Цель работы – создание моделей, алгоритмического и программного обеспечения интеллектуальной информационной системы для управления фондами нефтяных и газовых скважин добывающих предприятий нефтегазовой отрасти, основанных на современных моделях и методах интеллектуального анализа данных.

Область исследований в диссертации соответствует специальности 05.13.01 «Системный анализ, управление и обработка информации (промышленность)» по пунктам:

2. Формализация и постановка задач системного анализа, оптимизации, управления, принятия решений и обработки информации.

4. Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации.

9. Разработка проблемно-ориентированных систем управления, принятия решений и оптимизации технических объектов.

10. Методы и алгоритмы интеллектуальной поддержки при принятии управленческих решений в технических системах.

Объект исследования – бизнес-процессы управления фондами нефтяных и газовых скважин.

Предмет исследования – современные модели, методы и алгоритмы ИАД, применяемые при управлении фондом скважин.

Методы исследования. Использованы методы статистического анализа и теории погрешностей, методы имитационного моделирования, глубокие ИНС и методы объектно-ориентированного программирования. Анализ данных при

решении одной из слабоформализуемых задач проведён с помощью статистических методов, реализованных в пакете программ MATLAB. Разработка моделей глубоких ИНС осуществлена с помощью библиотеки MS CNTK.

Научная новизна полученных результатов:

1. Для точного решения слабоформализуемых задач классификации скважин-кандидатов и выбора для них геолого-технических мероприятий (ГТМ) разработан набор моделей глубоких ИНС, обеспечивающих в 2–4 раза более короткие сроки принятия управленческих решений по сравнению с традиционными методами.

2. Для решения слабоформализуемой задачи прогноза значений дебитов скважин разработан набор моделей глубоких ИНС, обеспечивающих в 2–3 раза меньшую погрешность прогноза по сравнению с традиционными методами.

3. Разработаны и экспериментально обоснованы адаптивные алгоритмы для предварительной подготовки технологических и геологических данных по фонду скважин для последующего ИАД, позволяющие корректировать ошибки и неполноту исходных данных с относительной погрешностью не более 4,9 %.

4. Предложена и реализована SOA-модель (англ. Service-Oriented Architecture – SOA) программного обеспечения ИИС, позволяющая создавать единое информационное пространство добывающих предприятий и быстро адаптировать ИИС к изменениям, происходящим на таких предприятиях.

Теоретическая и практическая значимость диссертации. Теоретическая значимость работы заключается в создании моделей глубоких ИНС прямого распространения. Полученные в результате исследований рациональные архитектуры и значения гиперпараметров этих моделей ИНС дают высокую точность и скорость решения трёх важных слабоформализуемых задач при управлении фондом скважин. Разработанные адаптивные алгоритмы предварительной подготовки исходных данных по фонду скважин опираются на принципы «Big Data» и позволяют с высокой точностью корректировать ошибки и пропуски в исходных данных.

Практически значимый результат – программное обеспечение ИИС для управления фондом скважин, успешно апробированное при ИАД по нескольким фондам скважин месторождений и при интеграции с программными средствами унаследованных ИС двух предприятий. Это ПО позволяет специалистам с высокой точностью решать слабоформализуемые задачи и в итоге в сжатые сроки принимать обоснованные управленческие решения. SOA-модель ПО и её высокопроизводительные компоненты позволяют также формировать единое информационное пространство добывающих предприятий отрасли и гибко подстраивать такое пространство и ИИС как его часть под изменения на предприятии технологического и экономического характера.

Практически значимыми для добывающих предприятий нефтегазовой отрасли являются методика применения ИИС при управлении фондами скважин и методика предварительной подготовки данных по дебитам скважин.

Реализация и внедрение результатов работы. Результаты исследований использованы при выполнении проекта по гранту РФФИ № 18-47-700010. ИИС внедрена в ООО «Центр нефтегазовых технологий», о чём получен акт внедрения, и проходит опытную эксплуатацию в АО «Газпром добыча Томск». Отдельные результаты диссертации используются в учебном процессе Отделения информационных технологий Томского политехнического университета.

Достоверность и обоснованность полученных результатов и сделанных выводов обусловлена корректным применением и развитием методов ИАД и количественным сравнением результатов, полученных с помощью предложенных моделей и алгоритмов, с результатами известных методов. Адекватность разработанных моделей ИНС и алгоритмов подтверждается результатами их полномасштабных исследований на реальных данных по фондам скважин трёх месторождений и результатами их апробации на двух предприятиях.

Основные положения, выносимые на защиту:

1. Модели глубоких ИНС с полученными в результате исследований архитектурами и значениями гиперпараметров дают точность при решении слабоформализуемых задач классификации скважин-кандидатов для ГТМ и

выбора типа ГТМ для них, аналогичную точности решения этих задач опытным геологом и позволяют принимать решения по фонду скважин в 2–4 раза быстрее, чем при использовании традиционных методов и способов.

2. Модели глубоких ИНС с полученными в ходе исследований архитектурами и значениями гиперпараметров при решении слабоформализуемой задачи прогноза значений дебитов скважин дают погрешность прогноза в два и более раза меньшую погрешности решения такой задачи известными методами.

3. Адаптивные алгоритмы предварительной подготовки исходных технологических и геологических данных для их последующего интеллектуального анализа позволяют путём адаптации к видам ошибок и пропусков в данных корректировать ошибки и пропуски в исходных данных с относительной погрешностью не более 4,9 %.

4. SOA-модель программного обеспечения ИИС и её высокопроизводительные компоненты позволяют управлять всеми бизнес-процессами фонда скважин и легко формировать единое информационное пространство добывающего предприятия нефтегазовой отрасли, гибко подстраивая его и ИИС под изменения технологического и экономического характера.

Первый раздел диссертации содержит анализ существующего состояния проблемы автоматизации управления фондом скважин месторождений нефти и газа. Во втором разделе изложена концепция создания ИИС. В третьем разделе описаны разработанные модели и алгоритмы ИАД и приведены результаты исследования их эффективности. Четвертый раздел посвящён описанию разработанного программного обеспечения ИИС. В пятом разделе показаны результаты апробации ИИС на предприятиях нефтегазовой отрасли и сформулированы основные положения методики применения ИИС при анализе технологических и геологических данных.

Автор выражает благодарность научному руководителю д.т.н., профессору Маркову Н. Г. за помощь при написании диссертационной работы.

1. ПРОБЛЕМА АВТОМАТИЗАЦИИ УПРАВЛЕНИЯ ФОНДОМ СКВАЖИН МЕСТОРОЖДЕНИЙ НЕФТИ И ГАЗА

Новые интеллектуальные методы анализа данных и эффективные технологии добычи в нефтегазовой отрасли становятся особо актуальными вследствие значительных затрат на освоение новых месторождений, разработки месторождений на последних стадиях эксплуатации и колебаний на рынке спроса на УВС. Поэтому вместо бурения новых скважин сегодня на эксплуатируемых месторождениях всё чаще применяют современные технологии интенсификации добычи УВС на существующем фонде скважин. Управление объектами инфраструктуры промысла вообще, а фондом скважин в частности, осложняется множеством факторов – разрозненностью объектов, суровыми климатическими условиями, вводом в эксплуатацию новых объектов и т. п. В таких условиях необходимы наиболее оптимальные управленческие решения при интенсификации добычи УВС, гарантирующие минимальные риски и ощутимый результат. Ниже анализируется состояние проблемы автоматизации управления фондом скважин промысла по добыче нефти и газа.

1.1. Задачи управления фондом скважин промысла

Основными единицами фонда скважин месторождений нефти и газа являются комплексы «Добывающая скважина – Пласт (Пласты)», причём каждый из комплексов характеризуется двумя основными типами параметров: геологические и технологические параметры [3]. Геологические параметры описывают эксплуатируемый продуктивный пласт (пласты) месторождения, а технологические – зависят от выбранного технологического режима скважины и описывают её работу.

В целом, работы на скважинном фонде месторождения можно разделить на капитальный и текущий ремонты скважин и иной инфраструктуры фонда. Отличительной чертой капитального ремонта скважин является изменение структуры фонда (изменение числа работающих комплексов), в то время как текущий ремонт скважин связан, чаще всего, с небольшими работами на

скважине и (или) на инфраструктуре куста скважин или с переводом скважины с одного способа эксплуатации на другой для обеспечения заданного технологического режима.

Режимы работы скважин с залежами месторождения (продуктивными пластами) даже в своём лучшем случае не обеспечивают полный коэффициент отдачи УВС. Это вызывает необходимость искусственного воздействия со стороны человека на продуктивные пласты или добывающие скважины с целью интенсификации добычи. Работы такого рода, призванные поддержать целевой уровень добычи, названы геолого-техническими мероприятиями (ГТМ). Основное отличие ГТМ от других видов работ на месторождениях состоит в том, что их реализация приводит к приросту добычи УВС. Набор конкретных видов ГТМ для того или иного фонда скважин может существенно отличаться в различных нефтегазодобывающих предприятиях.

Почти все ГТМ относятся к капитальному ремонту скважин. Но как уже говорилось, предприятия самостоятельно определяют перечень работ, попадающих в категорию ГТМ. Несмотря на то, что ГТМ проводятся на всех этапах разработки месторождений, наиболее значимыми эти виды работ становятся на месторождениях, находящихся на заключительной стадии эксплуатации и исчерпавших ресурс своего естественного источника энергии. Они имеют лишь один путь для добычи – использование методов интенсификации добычи, реализуемых через проведение различных ГТМ.

1.1.1. Методы интенсификации добычи УВС

Все методы повышения интенсивности добычи УВС можно разделить на группы, например, для нефтяных скважин это 5 групп методов [3, 4].

1. *Методы поддержания пластового давления.* Данная категория методов рассчитана на тот факт, что при добыче углеводородного сырья в пласте появляются зоны, в которых из-за пустот перемещение нефти становится невозможным. В связи с этим добывающие скважины располагают по месторождению по специальной сетке, где каждая добывающая скважина окружается нагнетательными скважинами, поддерживающими давление внутри

пласта на целевом значении. При этом применяются такие методы, как законтурного заводнения, приконтурного заводнения, внутриконтурного заводнения, закачки газа в газовую шапку нефтяного пласта и т. п.

2. *Методы, повышающие проницаемость пласта* ***в призабойной зоне***. Во-первых, это механистические методы: гидроразрыв пласта (ГРП), тартание, поршневание, гидропескоструйная перфорация, торпедирование, ремонтно-изоляционные работы (РИР). Во-вторых, в эту группу входят химические методы обработки призабойной зоны: соляная кислота, плавиковая кислота, уксусная кислота, серная кислота, угольная кислота, другие реагенты и растворители. В третьих, применяются физические методы: тепловая обработка, вибровоздействие, перевод на вышележащий (или нижележащий) горизонт.

3. *Интенсификация нефтедобычи при воздействии на* ***пласты***: введение в пласт растворителей, растворов полимеров, поверхностно-активных веществ, углекислоту и прочих веществ, которые вытесняют нефть, увеличение температуры пласта. Используются вещества (пропан, бутан), которые растворяясь в нефти, уменьшают её вязкость.

4. *Одновременно раздельная эксплуатация.* Среди методов в этой группе используются зарезка бокового ствола (бурение бокового ствола) и установка пакеров. Пласты в скважине разобщаются так, чтобы каждый из них имел отдельный канал для добычи УВС.

5. *Перевод скважины с одного способа (режима) эксплуатации на другой.* Используются фонтанный, газлифтный и насосный режимы. При газлифтном режиме подъём жидкости из пласта осуществляется путём закачки в колонну сжатого газа. Насос откачивает жидкость из скважины, что уменьшает давление на забое. Применяются различные виды насосов: штанговый глубинный насос (ШГН), электроцентробежный насос (ЭЦН), штанговый винтовой насос (ШВН), электродиафрагменный насос (ЭДН), гидропоршневые и струйные насосы.

При интенсификации добычи газа и газового конденсата из газоконденсатных месторождений также используется большое число методов, в первую очередь, воздействия на призабойные зоны и разрабатываемые пласты.

Степень влияния каждого метода на добычу УВС и условия применения методов изучены в разной степени. Так, тепловые обработки или вытеснение нефти полимерами на практике применяются давно и уже продемонстрировали свою эффективность. Комбинированные методы, закачка в пласт азотных пен, внутрипластовый каталитический крекинг тяжёлой нефти – примеры экспериментальных методов.

Нельзя однозначно сказать, приведёт ли к успеху какой-либо указанный метод интенсификации добычи сырья; во многом каждое месторождение и каждая связка скважина–пласт является уникальным для рассмотрения объектом. Ключевым фактором при выборе конкретного ГТМ является на данный момент интуиция, опыт и знания геолога, отвечающего за определённый участок месторождения.

1.1.2. Основные бизнес-процессы управления фондом скважин

Сегодня управление фондом скважин добывающего предприятия – это управление совокупностью весьма сложных бизнес-процессов (БП). Кратко рассмотрим основные из этих БП.

1. *БП «Сбор предложений по проведению мероприятий, ведущих к остановкам скважин».* Сбор предложений осуществляется от различных служб предприятия: производственно-технологическое управление, служба автоматизации, служба метрологии, служба связи, служба главного геолога и других. Помимо служб предприятия предложения поступают и от сторонних организаций, ведущих те или иные работы на месторождении в интересах предприятия. При этом в заявке указывается объект (скважина), необходимая работа (например, исследования в скважине или ГТМ) и сроки для мероприятия.

2. *БП «Ежемесячное планирование технологических режимов работы каждой скважины фонда».* По сути, реализация этого БП сводится к решению задачи прогноза значений дебита (суточного уровня добычи) каждого из

компонентов УВС (нефти, газа, газового конденсата (жидкая фракция, поэтому далее – жидкости) и воды), то есть прогнозируются четыре вида дебитов. Задача прогноза является слабоформализуемой задачей из-за большого числа факторов, влияющих на режимы работы скважины (измеряются десятки параметров при мониторинге скважины и куста скважин).

3. *БП «Планирование и проведение геофизических (ГИС), гидродинамических (ГДИС) и промыслово-геофизических исследований (ПГИ) в скважинах».* Планирование по каждому типу исследований ведётся по фонду скважин на месяц, квартал и год. Проведения ряда исследований невозможно без остановок скважин, что должно учитываться при планировании и реализации других БП.

4. БП «Управление и проведение ГТМ». Этот сложный БП в последние годы всё чаще выполняется геологами с использованием современных методов и алгоритмов, а также реализующих их программных средств. Часть его основных подпроцессов связана с подбором скважин-кандидатов на выполнение ГТМ и выбором конкретных ГТМ для них. Выбор осуществляется с учётом критериев экономической и технологической эффективности проводимых ГТМ.

5. *БП «Формирование оперативного плана-графика всех мероприятий на фонде скважин на месяц».* В итоговый план-график попадают все возможные мероприятия с определённым сроком выполнения службами промысла и подрядными организациями по каждой скважине: исследования (ГДИС, ГИС, и ПГИ), ГТМ, капитальный ремонт скважин (включая ликвидации аварий) и прочие работы на фонде. На каждое мероприятие назначается подрядная бригада капитального ремонта скважин (КРС) или специализированная организация. План-график постоянно редактируется и оптимизируется по суммарному времени простоя скважин. При возможности мероприятия комплексируются или привязываются друг к другу и выполняются впоследствии одновременно. С помощью плана-графика происходит разрешение конфликтов при наложении работ и нехватке бригад КРС.

6. *БП «Построение ежедневных и ежемесячных отчётов по обслуживанию фонда скважин»*. Данный БП включает в себя получение и рассылку отчётов всем заинтересованным службам и специалистам промысла и аппарата управления предприятия с определённой периодичностью. При формировании отчёта указывается срок, за который агрегируются данные, перечень объектов и их характеристики (параметры продуктивных пластов, скважин и оборудования, проведённые исследования на скважинах и/или КРС). Для отчётов обычно используется табличный формат и/или графики. Отчёты высылаются заинтересованным в них сотрудникам предприятия на электронную почту или доступны в корпоративной информационной системе (если таковая имеется у предприятия). Формирование отчётов должно производиться полностью автоматически, то есть ручной ввод не допустим. При формировании отчётов за конкретный месяц обычно предусматриваются следующие стандартные формы: валовая добыча УВС и закачка воды в пласты, ранжирование скважин, отклонение уровня добычи по причине остановки скважин на КРС, отчёт по добыче по видам УВС, время простоя каждой скважины и т. д.

1.1.3. Особенности реализации бизнес-процессов управления ГТМ

При реализации методов интенсификации добычи УВС особое место отводится ГТМ [5, 6]. Они являются управляющими воздействиями на элементы объекта управления – скважины фонда и продуктивные пласты. Решение о проведении ГТМ принимает субъект управления – специалисты различных служб промысла, в первую очередь, геологической. За окончательный выбор скважин-кандидатов на проведение ГТМ и собственно выбор самих ГТМ для таких скважин отвечает геологическая служба промысла. Обычно список ГТМ формируется каждый год с последующим его ежемесячным уточнением в рамках бюджета предприятия. Принятые специалистами решения о ГТМ на той или иной скважине реализуют бригады КРС или при проведении сложного ГТМ - специалисты подрядной организации. Важным реализуемым в системе управления фондом скважин процессом является оценка эффективности проведённых ГТМ. Такая оценка должна опираться на решение задачи прогноза

дебитов УВС после ГТМ. Оценка эффективности позволяет сформировать обратные связи и, если это требуется, корректировать управляющие воздействия на скважины или пласты [7]. Иногда одни и те же управляющие воздействия повторяются несколько раз. Например, довольно часто выполняется на одном и том же блоке месторождения многократный ГРП – один из перспективных методов ГТМ [8].

С другой стороны, управление ГТМ требует решения ряда слабоформализуемых задач. В первую очередь к ним относятся задача выбора скважин-кандидатов для ГТМ и задача определения вида ГТМ на таких скважинах. Для решения этих задач требуются значительные трудовые и временные затраты квалифицированных специалистов, которые на сегодняшний день на многих предприятиях практически вручную анализируют большие объемы разнородной геологической и технологической информации, что способствует увеличению вероятности возникновения ошибок вследствие человеческого фактора. Это ещё раз подчёркивает необходимость использования в контуре управления методов ИАД как методов поддержки принятия решений субъектом управления в реализации таких слабоформализуемых задач.

БП, которые должны выполнять службы предприятия при управлении ГТМ, показаны на рисунке 1.1. Номера конкретных подпроцессов на рисунке 1.1 указывают на то, что они описаны в нотации IDEF0 [9]. Нетрудно заметить, что этот комплекс подпроцессов весьма сложен и поэтому его невозможно реализовать службам без развитых средств автоматизации, то есть без использования многофункциональной информационной системы (ИС) [6]. Основными из этих подпроцессов являются следующие: сбор и актуализация информации, отбор скважин-кандидатов для проведения мероприятий, выбор ГТМ для выбранных скважин-кандидатов, определение технологической и экономической целесообразности (эффективности) выбранных ГТМ и, наконец, планирование работы бригад КРС для реализации ГТМ.

Рисунок 1.1 – Декомпозиция БП «Управление и проведение ГТМ» на подпроцессы

Рассмотрим подробнее эти подпроцессы и особенности средств для их реализации.

- *Сбор и актуализация данных*. ИС для управления ГТМ должна иметь доступ к исходным данным по фонду скважин и для этого должна интегрироваться с уже существующими на промысле АСУ ТП, получающими значения технологических параметров скважин и геологических параметров продуктивных пластов. В неё также должны заноситься рекомендации служб промысла на проведение ГТМ, ежегодные и ежемесячные планы проведения ГИС, ГДИС и ПГИ на скважинах фонда, которые ведут к остановке работы тех или иных скважин. Необходима также актуализация данных в связи с тем, что некоторые из них являются ошибочными, устаревшими и т. п. Таким образом, для реализации этого подпроцесса необходимо иметь интеграцию ИС для управления ГТМ с различными АСУ ТП промысла, осуществлять сбор данных от таких АСУ и из других источников и актуализировать собранные данные.
- *Отбор скважин-кандидатов для ГТМ.* Среди основных методов при решении задач для выбора скважины-кандидата можно отметить следующие автоматические и автоматизированные методы и алгоритмы: статистический анализ информации об оперативных остановках скважины, метод геологического потенциала и методы анализа основных тенденций изменений параметров скважин и продуктивных пластов. Также в ИС должны быть реализованы алгоритмы учёта скважин резерва, на которых по различным причинам мероприятия ещё не проводились, хотя ранее планировались.
- *Выбор ГТМ*. Для каждой скважины-кандидата, выбранной при выполнении предыдущего подпроцесса, с учётом специфики типа скважины (добывающая нефтяная или газовая, нагнетательная и т. д.) происходит выбор определённых ГТМ с помощью различных методов. Такие методы могут быть

автоматическими или автоматизированными, но важно, чтобы в определённой ситуации ИС позволила специалисту использовать наиболее эффективный метод, тем самым оказав поддержку ему в принятии решения о выборе ГТМ.

- *Оценка технологической и экономической эффективности ГТМ.* В этом подпроцессе производится проверка эффективности решений, принятых при реализации предыдущих подпроцессов, по технологическим и экономическим критериям. Например, в качестве простого экономического критерия используется объем инвестиций, запланированный предприятием на проведение ГТМ на текущий год. Если оценки инвестиций на проведение ГТМ превысят запланированный для них объем, то придётся пересмотреть перечень скважин-кандидатов для ГТМ и ГТМ для них.
- *Планирование работы бригад КРС.* Построение оптимального графика работ по проведению ГТМ на скважинах с точки зрения минимального времени простоя скважин фонда и необходимых временных интервалов для перемещения оборудования в места проведения ГТМ и выполнения работ бригадами КРС является важным подпроцессом управления ГТМ.

Анализ основных БП управления фондом скважин и подпроцессов управления ГТМ позволяет сказать, что практически все БП являются сложными и для их выполнения требуются средства автоматизации [5]. Анализ большого числа опубликованных работ показывает, что уровень автоматизации БП при управлении фондом скважин и, в частности, при управлении ГТМ является на многих добывающих предприятиях весьма низким на текущий момент времени [10]. Зачастую решение тех или иных задач управления фондом зависит от опыта и интуиции специалистов служб нефтегазодобывающего предприятия. Поэтому повышение степени автоматизации БП при управлении фондом скважин сегодня является актуальной проблемой, решение которой обеспечит добывающим предприятиям нефтегазовой отрасли конкурентные преимущества. При этом высокий уровень автоматизации наиболее важен при решении так называемых слабоформализуемых задач. Именно для них сегодня характерна высокая роль специалиста при решении, большая неопределённость влияния принятого

решения о ГТМ на результат работы скважин в виде прироста дебитов УВС вследствие недостатка исходной информации и отсутствие общепринятых методик.

1.2. Анализ существующих методов и методик управления фондом скважин

1.2.1. Традиционные методы управления фондом скважин

Одним из традиционных и часто используемых методов управления фондом скважин является метод сбора и экспертного обобщения предложений на остановки скважин для проведения ГТМ на фонде скважин от специалистов всех заинтересованных служб промысла. К таковым службам относятся геологическая, производственно-технологическая, отделы автоматизации, главного механика и главного энергетика. Обычно все предложения накапливаются в течение месяца в специальном журнале потенциальных остановок скважин. Оперативный сбор предложений от промысловых геологов обычно происходит на основе анализа значений параметров скважин, в первую очередь, дебитов УВС. В итоге опытный геолог (геологи) на основе анализа всех предложений производит выбор скважин-кандидатов на проведение ГТМ экспертным путем, то есть основываясь на своём опыте и интуиции.

ГИС, ГДИС и ПГИ на скважинах, если они ведут к их остановке, также оказывают влияние на итоговый перечень скважин на остановки и ГТМ. По каждому виду ГТМ графики остановок скважин во времени строятся на каждый месяц и на год.

В основе метода геологического потенциала [11, 12] лежат исходные (текущие) геологические и технологические данные о продуктивном пласте и скважине плюс особые критерии (сценарные условия), которые вручную задаются специалистом. Метод требует полноты исходных данных, поэтому отсутствие даже их части является препятствием к использованию метода и вновь требует участия специалиста. Метод позволяет рассчитать прогноз технологического режима добывающей скважины (средний для расчетного месяца уровень добычи нефти в сутки). Все скважины, у которых расчётное значение суточного уровня

добычи (дебита) нефти оказалось меньше утверждённого по плану, становятся кандидатами для ГТМ. Итоговое решение принимается по сценарным условиям, основными из которых для каждой скважины являются PI – индекс доходности, индекс «КН» и дебит нефти $Q_н$, а также был ли уже проведён ГРП на скважине.

История остановок скважин, в том числе оперативных, является основой для использования метода статистического анализа [13, 14] при решении различных задач управления фондом скважин, в том числе и слабоформализуемых задач выбора скважин-кандидатов и вида ГТМ. Отсутствие истории остановок рассматриваемой скважины не является препятствием для использования этого метода, в таком случае анализ производится по скважине (эталону), имеющей наиболее близкие параметры к анализируемой скважине в том же самом месторождении, а также имеющей в течение года остановки или ГТМ. Алгоритм, реализующий этот метод:

Начало

Шаг 1. Выявить на основе истории остановок периодичность проведения конкретного вида ГТМ по скважине или эталону.

Шаг 2. Прибавить к дате последнего проведённого вида ГТМ выявленный на предыдущем шаге период.

Шаг 3. Если расчётная дата попадает в период планирования таких мероприятий, то скважина становится кандидатом на ГТМ и заносится в список остановок скважин.

Конец.

Зачастую при применении ряда методов выбора скважин-кандидатов на проведение ГТМ используют системы критериев с установленными для них весами, определяющие в итоге экономическую и технологическую эффективность выбора скважины-кандидата для ГТМ и ГТМ для них. В качестве критериев могут выступать значения технологических и геологических параметров скважин и пластов, данные 3D-моделирования месторождения, экономические показатели, по которым можно понять, как распределены текущие запасы УВС с целью выделения перспективных для разработки участков. В качестве примера

использования совокупности нескольких критериев можно привести методику ОАО «Газпром» [15], созданную на основе отраслевого стандарта газовой подотрасли.

В принципе, любой метод выбора перспективного участка (блока) или скважин-кандидатов для ГТМ и собственно ГТМ для них обязательно должен включать следующие этапы [16, 17]: оценка основополагающих геологических и технологических факторов; выбор критериев для оценки различного рода рисков; обоснование критериев эффективности ГТМ. Проблема здесь в том, что часть используемых параметров (из проектной документации месторождений, карточек работы скважин, геолого-технической информации) несут большую неопределённость, а определение их достоверности – это отдельная нетривиальная задача.

Сегодня в нефтегазовой отрасли используется значительное число методов и алгоритмов выбора скважин-кандидатов для проведения одного из часто встречающихся видов ГТМ – ремонтно-изоляционных работ (РИР) [18]: трендовый анализ, экспресс-анализ состояния фонда нефтяных скважин, графический метод диагностики водопротоков, расчёт функции целесообразности проведения РИР, функция ожидания «обводнённость - продуктивность», скважинная диагностика водопритоков, подбор скважин-кандидатов по причинам невязок и другие. Каждый из них имеет свои преимущества и недостатки, а также различного рода ограничения: по типу скважин (нефтяные добывающие или нагнетательные и т. п.), учёт взаиморасположения скважин, наличие накопленной статистики, высокие требования к временным и трудовым затратам, наличие программных пакетов, реализующих методы, и другие.

Успешность планирования РИР определяется рядом факторов. Во-первых, необходим правильный выбор скважин для РИР, а также качество данных ГДИС, ГИС и ПГИ. Во-вторых, нужно правильно подобрать технологию РИР и соответствующий изоляционный материал. Также должен быть проведён техническо-экономический анализ для определения целесообразности ГТМ, режим скважины после РИР должен быть рентабельным. Технологическая

успешность подбора РИР для скважин оценивается после проведения самого ГТМ по данным ГИС и ГДИС, составу притока и продуктивности скважины при выходе её на режим. Неудачная доля РИР достаточно высока (до 95 %) [18].

Иногда применяется методика проактивного блочного анализа при разработке месторождений [17], которая позволяет определить скважины, нуждающиеся в ГТМ. Суть методики заключается в том, что весь фонд скважин разбивается на блоки. После сбора информации выявляются и ранжируются проблемные зоны в зависимости от степени несоответствия фактических значений параметров целевым. Затем подбираются ГТМ или иные воздействия с последующим мониторингом их успеха. Под блоком здесь понимается набор скважин, эксплуатирующих один и тот же продуктивный пласт и характеризующихся близкими значениями обводнения и пластового/забойного давления. Граница блока проходит по контуру нагнетательных скважин.

Основная информация о блоках берётся из месячных эксплуатационных рапортов (МЭР). Из МЭР берётся изменение дебита жидкости и нефти, рост обводнённости, время простоя скважин, среднее пластовое давление и пористость, PVT-свойства, проницаемость, начальные и остаточные нефтенасыщенность и водонасыщенность. Далее выполняется проактивный блочный анализ по алгоритму, включающему следующие шаги [17]:

Начало

Шаг 1. Адаптация PVT-параметров (расчёт объёмного коэффициента нефти и воды, сжимаемости нефти и воды),

Шаг 2. Определение величины задержки реакции добычи жидкости на изменение закачки,

Шаг 3. Адаптация модели материального баланса,

Шаг 4. Адаптация модели характеристики вытеснения по фактическим данным (геологические и извлекаемые запасы УВС, конечный коэффициент извлечения нефти – КИН),

Шаг 5. Ретроспективный анализ характеристики вытеснения (рост обводнённости),

Шаг 6. Прогноз эксплуатационных параметров и показателей разработки (дебит жидкости, эффективная вязкость, обводнённость),

Шаг 7. Расчёт целевой закачки воды и компенсации.

Конец.

Подбор конкретных ГТМ (зарезка бокового ствола – ЗБС, ГРП, перевод на выше- или нижележащий горизонт, реперфорация, РИР) может осуществляться на основе использования 3Д гидродинамических моделей [19]. В основе выбора лежит оценка характеристик пласта и величин запасов УВС на участках месторождения, а также различных показателей эффективности ГТМ. Такой подход требует от специалиста серьёзной квалификации и умений работать со специализированным ПО для гидродинамического моделирования. С точки зрения трудоёмкости этот процесс является весьма затратным по времени. Также моделирование связано с накоплением дополнительных ошибок и неточностей, связанных с процессом самого моделирования и имеющихся неточностей в истории разработки. Цифровая модель месторождения позволяет определить перспективность скважины-кандидата на ГТМ, в первую очередь, по величине остаточных извлекаемых запасов УВС. При этом в качестве основополагающих параметров зачастую используют следующие.

- Наличие подвижных извлекаемых запасов нефти. Плотность запасов определяется из 3Д модели.
- Карта трещин нефтенасыщенности. Определяется суммированием по модели.
- Мощность пластов нефтеколлекторов между водо- и нефтенасыщенными пропластками.
- Карты среднего пластового давления, проницаемости и пористости. Рассчитывается как средневзвешенное по поровому объёму по столбцу нефтенасыщенных ячеек.
- Карта расчленённости пласта – количество переходов коллектор–неколлектор вдоль столбца нефтенасыщенных ячеек. Важно при ЗБС.
- Наименьшее расстояние до фронта нагнетаемой воды.

- Расстояние до ближайшей нагнетательной скважины от скважины-кандидата или от точки предполагаемой ЗБС.
- Компенсация по целевому блоку.
- Лучший исторический дебит нефти (или максимальный коэффициент продуктивности).
- Карты прогнозной продуктивности.
- Прогнозный уровень добычи нефти. Рассчитывается по продуктивности и картам пластового и забойного давлений.
- Карта невязок модельной накопленной нефти (отклонения расчёта от факта).
- Критические значения оценочных параметров (максимально и минимально допустимые их значения). Если фактическое значение параметра не попадает в указанный диапазон, то это стоит оценивать как риск.
- Ретроспективный анализ выполненных ГТМ (оценка эффективности проведённых работ в предыдущие месяцы).

На некоторых нефтегазодобывающих предприятиях используются метод и реализующие его алгоритмы, сутью которых является задание специфических границ для технологических параметров скважин и геологических параметров продуктивных пластов [20]. Недостатком этого метода является сложность определения границ интервалов таких параметров для конкретного месторождения. Поэтому качество результатов будет зависеть от интуиции и опыта специалиста, определяющего эти границы или интервалы изменения параметров. Значение того или иного параметра может быть ограничено интервалом, а также только сверху или снизу. Например, для нефтяных скважин могут быть выбраны следующие критерии в виде заданных границ [16]: высокая обводнённость $\overline{O} > 5\,\%$, падение добычи за год $L_{\text{год}} > 3\,\%$, низкий эффект от ГТМ $Q < Q_{\text{цел}}$, низкий темп отбора $\frac{\sum Q_{\text{н}}^{\text{нак за 1 год}}}{tQ_{\text{изв}}} < 0{,}05$ (т/сут.), высокая плотность УВС $p_{\text{з}} > 49$, высокое пластовое давление $\overline{P_{\text{пл}}} \geq 0{,}85P_{\text{пл}}^{\text{цел}}$ (атм.) и т. д.

В п. 1.1.2. показано, что для реализации БП «Ежемесячное планирование технологических режимов работы каждой скважины фонда» важно прогнозировать значения дебитов четырех компонентов УВС каждой добывающей скважины. Однако существующие методы прогноза дебитов в лучшем случае позволяют выполнить экстраполяцию данных путём линейной регрессии [10]. Столь простые методы не способны учесть множество аспектов, заложенных в данных и взаимовлияние их друг на друга в рамках анализируемых временных рядов для дебитов. Соответственно, точность прогноза дебитов не всегда удовлетворяет специалистов предприятий.

Учитывая всё изложенное выше, можно считать, что практически все традиционные методы, алгоритмы и методики не позволяют решать задачи управления фондом скважин с требуемой точностью. Более того, при принятии решения по результатам вычислений специалисты вынуждены опираться на свои опыт и интуицию.

1.2.2. Методы ИАД для управления фондом скважин

В сопредельных с добычей УВС областях, таких как нефтегазовая геология и геофизика уже в течение нескольких лет применяют методы ИАД. Результаты исследований профессора Приезжева И. И. с сотрудниками указывают на перспективность применения искусственных нейронных сетей (ИНС) при решении задач в геофизике [21]. Прорывные результаты при использовании ИНС в этих областях получены также в транснациональной компании Schlumberge, однако особенности применяемых моделей ИНС в научной литературе практически не описаны.

В последние годы для решения отдельных задач при управлении фондом скважин также начали применять методы и алгоритмы ИАД. Проанализируем основные полученные при этом результаты. В РГУ нефти и газа имени И. М. Губкина было проведено исследование [8] применимости ИНС для прогноза параметров работы скважин после проведения ГРП в ООО «Газпромнефть-Ноябрьскнефтегаз». На нескольких нефтяных месторождениях этого предприятия был проведен достаточно большой объём работ по ГРП. В качестве параметров

скважин для прогноза их значений были выбраны следующие: дебит жидкости, дебит нефти и обводнённость. На вход ИНС подавались значения следующих параметров: мощность пласта, пластовое давление, планируемые забойное давление и скин-фактор после ГРП и другие. В обучающую выборку попало 117 ГРП, 10 же мероприятий стали контрольными для проверки качества обучения модели ИНС в виде персептрона. Точность работы модели ИНС оказалась невысокой, поэтому требуется ее последующее улучшение и исследование с учётом различных аспектов процесса эксплуатации скважин после ГРП.

В работе [22] был проведён анализ операций по выравниванию профиля приёмистости (ВПП) для нагнетательных скважин, а также оценено влияние геолого-геофизических характеристик продуктивных пластов и технологических параметров скважин на результат с точки зрения прироста уровня добычи нефти вследствие уменьшения обводнённости УВС. Предложена модель ИНС для предсказания технологической эффективности от проведения ВПП. Обводнённость УВС может являться следствием различных причин. Две из них – заколонные перетоки и нарушение герметичности колонны – решаются за счёт РИР. Другие две – обводнение подошвенной водой и продвижение нагнетаемой воды по пропласткам с высокой проницаемостью – относятся к системе поддержания пластового давления. ВПП по сути является одним из видов ГТМ и достигается за счёт особой технологии воздействия на пласт специальными реагентами, которые уменьшают проницаемость промытых слоёв, что увеличивает рентабельный срок эксплуатируемой скважины, а значит, и обеспечивает дополнительное получение УВС.

В качестве входных параметров для ИНС было выбрано девять характеристик: пористость пласта, проницаемость пласта, дебит жидкости добывающей скважины до ВПП, обводнённость УВС до ВПП, объём закачки реагента, депрессия на добывающей скважине, среднее расстояние между нагнетательными и добывающими скважинами, потенциал динамического уровня и средняя компенсация на участке. Выходными итоговыми характеристиками

стали ожидаемое изменение обводнённости продукции и ожидаемая продолжительность эффекта в днях.

Для обучающей выборки значения входных параметров должны обязательно находиться в том же диапазоне, что и фактические значения при проверке работы модели ИНС. В качестве выходного сигнала ИНС использовался показатель обводнённости УВС на добывающих скважинах, так как успех ВПП означает снижение этого параметра, а, стало быть, происходит увеличение добычи нефти. Оценка производилась по разнице расчётных и фактических значений обводнённости. Для обучения и тестирования ИНС использовалось 106 скважин, причем всё это множество было разбито на две части: 90 % для обучения и 10 % для тестирования. Апробация разработанной модели ИНС проходила на последующих ВПП, выполненных по той же технологии. Точность прогноза при использовании модели ИНС оказалась сопоставима с таковой у метода гидродинамического моделирования, но скорость работы модели гораздо выше. Этими же авторами при решении той же задачи предварительно был использован метод кластеризации для группировки исходных данных на нефтяных месторождениях Поволжского района [23]. Это позволило уменьшить ошибку в предсказаниях характеристик добывающих скважин при проведении ВПП. С предварительной кластеризацией данных прогноз обводнённости оказался почти в 2 раза точнее, чем при использовании метода гидродинамического моделирования. Однако все полученные результаты прогноза далеки от результатов, представляющих практический интерес.

В работе [24] применялся кластерный анализ для разделения фонда скважин на кластеры (группы) в зависимости от приоритетности проведения ГТМ, затем и на подкластеры – по видам ГТМ для каждой скважины-кандидата. Можно сказать, что исследователи смогли извлечь по исходным данным особый кластер скважин-кандидатов на ГТМ. Это позволяет считать методы кластерного анализа перспективными с точки зрения повышения уровня автоматизации БП управления фондом скважин, а стало быть, включение таких методов в ИИС должны быть включены в перечень инструментов для обработки данных.

Основными недостатками применённого подхода являются использование для анализа только геологических данных и применение простейшего алгоритма кластеризации, который на одном и том же наборе исходных данных может показывать различные результаты [25, 26]. Этот алгоритм также чувствителен к шумам и требует априорного задания количества кластеров.

В работах [27–31] показаны результаты исследований при решении задачи выбора скважин-кандидатов для ГТМ с использованием метода нечёткой логики. Если рассматривать процесс выбора скважин-кандидатов с точки зрения классической логики, то скважина оказывается либо кандидатом (истина, единица), либо нет (ложь, ноль), то есть граница между выбираемыми скважинами для ГТМ и не выбираемыми – чёткая. Теория нечётких множеств вводит характеристическую функцию принадлежности, которая принимает значения непрерывно от нуля до единицы, то есть определяет вероятность отнесения скважины к группе (классу) кандидатов, также можно определить порог этой вероятности, когда скважину стоит относить к классу кандидатов на ГТМ. Рассматривать можно не только каждую скважину индивидуально, но и блоки скважин в целом. При этом в качестве геологического критерия авторы предложили использовать набор параметров пласта: мощность пласта, пористость, нефтенасыщенность (ГИС), песчанистость, а также вариативно: проницаемость, расчленённость, вертикальная анизотропия. Важно, чтобы выбираемые параметры коррелировали между собой, иначе брать их не имеет смысла. Для всех параметров определяются функции принадлежности и их веса, которые путём перемножения друг с другом образуют единый геологический критерий. Для оценки технологического критерия анализируются: пластовое давление, работа скважин окружения, анализ закачки, скин-фактор, техническое состояние скважин. Для этих параметров также определяются функции принадлежности и их веса. Недостатком метода нечеткой логики при решении задачи выбора скважин на ГТМ является то, что на практике оценки функций принадлежности и весов выполняются экспертным путем опытными геологами на основе их интуиции и опыта.

Подводя итоги анализа используемых сегодня в России и за рубежом для управления фондом скважин методов ИАД, можно считать, что получены первые обнадеживающие результаты при решении частных задач. Однако точность этих результатов не высока, а решение ряда практически важных задач управления фондом с помощью таких методов даже не рассматривалось.

1.2.3. Информационные системы управления фондом скважин

На нефтегазодобывающем предприятии можно выделить три режима использования ИС, соответствующих трём уровням развития ИС управления фондом скважин, в рамках которых могут решаться описанные задачи:

- «ручной»;
- автоматизированный;
- автоматический.

«Ручной» режим подразумевает выполнение цикла работ по анализу исходных данных и принятия решений специалистом геологической (в некоторых случаях и технологической) службы, основываясь его на индивидуальном опыте. Роль информационных систем в данном случае сводится лишь к сбору информации из баз данных (БД) различных ИС и (или) АСУ ТП предприятия и представлении её в удобной для визуального восприятия форме. Далее специалист ведет анализ данных и принимает решение при управлении фондом скважин в соответствии с его опытом и интуицией. Субъективная интерпретация данных часто приводит к возникновению ошибок из-за человеческого фактора, особенно при решении слабоформализуемых задач, например, выбора скважин-кандидатов на проведение ГТМ.

В ряде информационных систем, используемых в России, реализованы простые методы для «полуручного» выбора скважин-кандидатов для проведения ГТМ. Затем для выбранных скважин автономно в других программах ведется автоматизированный расчет экономической эффективности ГТМ на них и автоматизированное планирование работы бригад КРС для выполнения ГТМ. Например, для нефтяных скважин при расчете эффективности ГТМ большой популярностью пользуется автономная программная реализация метода

потенциала скважины с экономической точки зрения [32]. На основе полученных результатов таких расчетов специалистом проводится окончательный анализ исходных и расчетных данных и принимаются соответствующие решения. При таком автоматизированном режиме решения крупной задачи специалисту при пользовании узкоспециализированными ИС и программами приходится не один раз менять форматы данных с учётом требований этих ИС и принимать окончательные решения по выбору скважин-кандидатов, руководствуясь своими опытом и интуицией.

Автоматический режим использования ИС подразумевает реализацию в виде многофункциональной ИС современных методов и алгоритмов анализа исходных данных, позволяющих при управлении фондом скважин на выходе такой ИС получать готовые результаты, например, получать список скважин-кандидатов для ГТМ и (или) виды ГТМ для конкретных скважин из этого списка. При решении других укрупнённых задач управления фондом скважин в рамках этой ИС также должны использоваться автоматический и (или) автоматизированный режимы работы специалиста с данными. Конечно, такая многофункциональная ИС будет идеальным инструментом для специалистов добывающих предприятий нефтегазовой отрасли.

Среди рассмотренных выше укрупнённых БП управления фондом скважин наиболее трудоемкими и слабоформализуемыми являются процессы выбора скважин-кандидатов на проведение ГТМ и подбора подходящих видов ГТМ для выбранных скважин, а также процесс решения задачи планирования технологических режимов работы скважин. Поэтому на данный момент на большинстве добывающих предприятий России и за рубежом бизнес-процессы, направленные на решение этих задач, выполняются в «ручном» режиме, и лишь на некоторых предприятиях отдельные подпроцессы этих процессов реализуются с помощью узкоспециализированных ИС в автоматизированном режиме. Методы и алгоритмы для автоматического режима решения таких задач только начинают развиваться в силу специфичности исходных данных. Действительно, для каждого типа скважин (нефтяные добывающие и газовые) процесс автоматизации

поиска скважин-кандидатов для ГТМ может существенно отличаться в связи с геологическими особенностями разных видов эксплуатируемых месторождений и из-за различных используемых наборов параметров, описывающих работу скважины или комплекса «скважина – пласт (пласты)».

Некоторые крупные зарубежные компании, такие как Halliburton [33] и Schlumberger [34], известны на мировом рынке ИС для управления разработкой месторождений своими корпоративными продуктами. Эти ИС представляют собой интегрированную среду для работы с различными данными, включают в себя множество различных функций и подсистем, в том числе подсистемы для управления ГТМ, прогноза значений параметров скважин и продуктивных пластов, 3Д-моделирования пластов и другие. Такая полнофункциональная система позволяет специалисту, в первую очередь геологу, вести комплексный анализ данных с высокой степенью автоматизации. Однако методы, реализованные сегодня в таких системах, часто не обладают достаточной точностью, особенно при решении слабоформализуемых задач управления фондом скважин. Также столь многофункциональные корпоративные системы имеют большую совокупную стоимость владения (складывается из стоимости лицензии при покупке и затрат на внедрение и сопровождение), что делает их недоступными для многих российских добывающих предприятий нефтегазовой отрасти.

Однако в последние годы в развитии производственных ИС добывающих предприятий нефтегазовой отрасли наметилась следующая тенденция. Постепенно несложные методы и алгоритмы, используемые на предприятиях при решении частных задач управления фондом скважин, начали дополняться современными статистическими и нейросетевыми методами и алгоритмами обработки и анализа данных, реализуемыми в рамках существующих ИС. С помощью этих методов получены первые обнадеживающие результаты, однако точность последних пока невысока.

1.3. Концепция интеллектуальных месторождений

Сегодня уже общеизвестно, что интеллектуальные технологии как часть перспективных информационных технологий – главный фактор обеспечения эффективных управленческих решений при добыче УВС [35]. В результате использования таких технологий сегодня за рубежом у ряда исследовательских центров начинают появляться крупные ИС для автоматического сбора, обработки и анализа данных с целью прогноза добычи УВС, решения задач поддержки принятия решений при управлении фондом скважин и т. п. [35]. Наиболее важная особенность таких систем заключается в возможности работы с большими объемами данных в рамках концепции «Big Data» [36]. Сегодня ведущие нефтегазовые компании мира всё чаще используют методы ИАД и уже начали разрабатывать свои методы и подходы к комплексному интеллектуальному управлению эксплуатируемыми месторождениями. При этом они также разрабатывают соответствующие аппаратные и программные средства [37]. Причём каждая из таких компаний реализует свою уникальную концепцию интеллектуального месторождения. Например, компания BP (Англия) – «Месторождение будущего», Shell (Дания) – «Умное месторождение», Chevron (США) – «Интеллектуальное месторождение», Halliburton (США) – «Управление в режиме реального времени», крупнейшая нефтесервисная транснациональная компания Schlumberger – «Умные скважины» и т. д. [38]. Все указанные концепции можно объединить под одним общим названием – концепция интеллектуальных месторождений.

В этой концепции большую роль играют информационные технологии, позволяющие производить мониторинг объектов месторождения (фонд скважин, оборудование по подготовке добытого УВС и т. д.) в режиме реального времени [39]. При этом все используемые информационные технологии, включая интеллектуальные технологии, и реализующие их ИС должны тесно интегрироваться при выполнении всех производственных БП предприятия для повышения эффективности управления его производством. Также важна ориентированность в концепции на сокращение издержек и быстрое принятие

решений в режиме реального (квазиреального) времени [7]. Интегрированная модель месторождения включает актуальные математические модели продуктивных пластов, скважин фонда и его наземной инфраструктуры. Так как целью разработки интеллектуального месторождения является эффективность всего актива предприятия в целом, то при автоматизации контроля и мониторинга эксплуатируемых месторождений эти модели должны быть интегрированы друг с другом, а также все их параметры должны быть актуальны и доступны в режиме реального времени.

Можно считать, что эффективность интеллектуальных технологий, встроенных в информационные технологии предприятия и являющихся их частью, обуславливается тесной их связью с БП предприятия, оборудованием и информационно-аналитическими системами за счёт интеграции их по функциям и данным. Режим реального времени должен осуществляться для предотвращения осложнений при эксплуатации месторождения, так как использование информации о событии после её сбора из-за отложенного анализа данных – это уже риск.

Отметим, что в России фундаментальные основы концепции интеллектуальных месторождений развиваются в Институте проблем нефти и газа РАН под руководством академика РАН Дмитриевского А. Н. [40]. В НТЦ ПАО «Газпромнефть» начаты исследовательские работы по использованию современных ИНС при решении слабоформализуемых задач в управлении эксплуатируемыми месторождениями [41]. К сожалению, эти работы, посвящённые теоретическим аспектам построения систем управления эксплуатируемыми месторождениями и основам применения методов ИАД, пока ещё не вышли за рамки теоретических исследований. Однако такая концепция в части других направлений исследований уже начинает частично реализовываться на практике и доказывать свою эффективность на месторождениях ряда крупных российских предприятий: Ачимовское, Северо-Комсомольское, Уренгойское нефтяные месторождения (ПАО НК «Роснефть»), Южно-Тамбейское газоконденсатное месторождение (ПАО «НОВАТЭК») [37] и т. д.

Кратко остановимся на некоторых особенностях реализации концепции интеллектуальных месторождений. Во-первых, при реализации концепции необходимо с учётом принципов «Big Data» разработать алгоритмы сбора, хранения, структурирования и, главное, предварительной обработки (корректировка ошибочных данных, замена пропущенных значений и т. п.) технологических и геологических данных, собираемых на эксплуатируемом месторождении. Во-вторых, важно разработать интегрированную модель месторождения, а также создать и правильно использовать инструментальные средства моделирования всех её элементов. В-третьих, в рамках концепции необходимо создать современные методы и алгоритмы ИАД для поддержки принятия решений специалистами предприятия при эксплуатации месторождения, в том числе при управлении фондом скважин. В-четвертых, создание и использование всех этих моделей, методов и алгоритмов должно идти в рамках разрабатываемой на предприятии многофункциональной корпоративной ИС. При реализации такой ИС также важно разработать не только подсистему предварительной обработки собираемой геолого-технологической информации, но и для сбора этих данных разработать средства интеграции этой системы со всеми унаследованными ИС предприятия: АСУ ТП на промыслах, ИС химико-аналитических лабораторий и т. д.

Результаты реализации концепции интеллектуальных месторождений на конкретном добывающем предприятии обязательно должны включать в себя наряду с созданной многофункциональной ИС рекомендации по оптимизации управления месторождением. Примером таких рекомендаций могут служить рекомендации по оптимизации технологических режимов работы скважин фонда при использовании разработанных средств этой ИС и имеющихся технико-экономических ограничений и необходимости возврата инвестиций за определённый временной интервал. Среди таких рекомендаций могут быть также рекомендации по решению слабоформализуемых задач управления ГТМ для каждой скважины, рекомендации по бурению новых скважин, оптимизации простоев скважин и т. п.

Принятие оптимальных управленческих решений при эксплуатации и развитии инфраструктуры месторождений должно опираться на новые интеллектуальные методы поддержки принятия решений. Такие методы сегодня находятся в процессе разработки в нескольких исследовательских центрах за рубежом [35].

К сожалению, сегодня в России и за рубежом на добывающих предприятиях нефтегазовой отрасли применяются ИС, зачастую предназначенные для решения узкоспециализированных задач. Используя их, специалисты практически в ручном режиме [10] решают слабоформализуемые задачи управления фондом скважин. То есть степень автоматизации является достаточно низкой. Чаще всего эти системы производят доставку необходимой информации специалистам из других ИС предприятия и предоставляют её в нужной визуальной форме (в виде таблиц, графиков). После этого специалист анализирует данные и принимает окончательное решение, руководствуясь своим опытом, а иногда и интуицией. Поэтому разработка новых интеллектуальных моделей, методов и алгоритмов интеллектуального анализа геологических и технологических данных и создание интеллектуальных информационных систем (ИИС) для управления фондом скважин, включая управление ГТМ, по-прежнему остаются актуальными.

1.4. Цель и задачи диссертационного исследования

Опираясь на результаты проведённого выше анализа методов, алгоритмов и информационных систем, используемых сегодня для управления фондом скважин нефтегазодобывающих предприятий, можно сделать вывод об актуальности развития новых и адаптации существующих моделей, методов и алгоритмов ИАД и создания ИИС для управления фондом скважин. Более того, в этих многофункциональных ИИС необходимо предусмотреть возможность гибкой интеграции с унаследованными ИС и программными системами добывающих предприятий и возможность лёгкой модернизации таких ИИС.

В настоящее время всё более актуальной становится концепция интеллектуальных месторождений, которая включает в себя разработку целого

комплекса эффективных технологий добычи УВС, интеллектуальных методов обработки и анализа больших объёмов геолого-технологических данных, а также интеграцию всех программных и информационных систем нефтегазодобывающего предприятия по функциям и данным. Это ещё раз подчеркивает актуальность создания многофункциональных ИИС для управления фондом скважин на основе ИАД и указывает на то, что тема диссертационной работы совпадает с основным вектором реализации концепции интеллектуальных месторождений.

Цель работы – создание моделей, алгоритмического и программного обеспечения интеллектуальной информационной системы для управления фондами нефтяных и газовых скважин добывающих предприятий нефтегазовой отрасти, основанных на современных моделях и методах интеллектуального анализа данных.

Для достижения поставленной цели необходимо решить следующие **задачи:**

1. Проанализировать современное состояние проблемы автоматизации процессов управления фондами нефтяных и газовых скважин добывающих предприятий.
2. Разработать концепцию ИИС для управления фондами скважин как системы с реализованными в ней моделями, методами и алгоритмами ИАД, способной легко интегрироваться с унаследованными ИС и адаптироваться к изменяющимся условиям конкретных добывающих предприятий отрасли.
3. Разработать и исследовать новые модели и алгоритмы решения ряда задач, в том числе слабоформализуемых, при управлении фондами скважин.
4. Создать программное обеспечение (ПО) ИИС с учётом разработанных её концепции, моделей и алгоритмов ИАД.
5. Провести апробацию ИИС на предприятиях нефтегазовой отрасли, в том числе путём интеграции её с существующими ИС и АСУ ТП, и создать методику применения ИИС при управлении фондами скважин.

1.5. Основные результаты и выводы по разделу

Управление фондом скважин – это управление совокупностью весьма сложных БП. При реализации ряда БП специалистам предприятия также необходимо решать слабоформализуемые задачи, а для этого следует использовать современные методы анализа данных. Анализ методов и средств автоматизации при управлении фондом скважин позволил получить следующие результаты и сделать выводы.

1. Сегодня на большинстве нефтегазодобывающих предприятий России и на ряде добывающих предприятий за рубежом специалисты продолжают использовать традиционные простые методы, алгоритмы и методики при решении многих задач управления фондом скважин. Практически все эти методы и алгоритмы не позволяют решать такие задачи с требуемой точностью, а при принятии окончательных решений по результатам проведённых с их помощью расчетов специалисты вынуждены опираться на свои опыт и интуицию. Всё это указывает на актуальность разработки современных более точных методов, алгоритмов и программных средств анализа геологических и технологических данных и поддержки принятия решений специалистами при управлении фондом скважин.

2. Методы и алгоритмы ИАД являются альтернативными традиционным простым методам и алгоритмам и должны позволять автоматически получать решения при управлении фондом скважин, в частности при управлении ГТМ. Сегодня получены первые обнадёживающие результаты при применении таких методов и алгоритмов при решении ряда частных задач при управлении фондом скважин, однако их точность пока не высока.

3. Анализ используемых на добывающих предприятиях ИС показывает, что в большинстве из них реализованы несложные, чаще всего традиционные методы и алгоритмы управления фондом скважин. При этом многие ИС являются узкоспециализированными. На некоторых предприятиях специалисты продолжают анализировать практически в «ручном» режиме огромные массивы геологических и технологических данных по фонду скважин, извлекаемых из

различных АСУ ТП и других ИС предприятий. Это требует значительных временных затрат труда специалистов и может привести к ошибкам из-за человеческого фактора. По результатам анализа можно сделать вывод, что такие ИС обеспечивают лишь частичную автоматизацию БП управления фондом скважин.

4. Всё изложенное в п. 2 и п. 3 указывает на актуальность создания многофункциональных ИИС для управления фондом скважин и проведения исследования эффективности положенных в их основу современных методов и алгоритмов ИАД. Более того, при решении трудоёмких и наукоёмких слабоформализуемых задач при управлении фондом скважин такие ИИС должны выполнять функции систем поддержки принятия решений. В этих ИИС необходимо также предусмотреть возможность гибкой интеграции с унаследованными ИС и АСУ ТП и иными программными системами добывающих предприятий и возможность лёгкой модернизации ИИС для обеспечения высокого уровня автоматизации БП управления фондом скважин.

5. В настоящее время всё более актуальной становится концепция интеллектуальных месторождений, которая включает разработку целого комплекса эффективных технологий добычи УВС, интеллектуальных методов обработки и анализа больших объёмов геологических и технологических данных, а также интеграцию всех программных и информационных систем нефтегазодобывающего предприятия. Опыт реализации отдельных положений такой концепции на нескольких месторождениях ведущими в мире компаниями также подчеркивает актуальность создания многофункциональных ИИС для управления месторождениями, в том числе их фондами скважин.

6. Сформулированы цель и задачи диссертационного исследования, направленные на создание именно такой ИИС. Следует отметить, что развиваемое в диссертационной работе научное направление совпадает с основным вектором реализации концепции интеллектуальных месторождений, продвигаемой сегодня мировым научным сообществом и ведущими добывающими компаниями нефтегазовой отрасли.

2. КОНЦЕПЦИЯ СОЗДАНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

Одной из задач, решаемых для достижения цели диссертационной работы, является разработка концепции ИИС управления фондом скважин нефтегазодобывающего предприятия. Под концепцией здесь и далее понимается совокупность требований, способов (подходов) и принципов создания ИИС. Ниже рассматриваются основные сформированные требования к ИИС и принципы её создания. Предлагается использовать принципы сервисно-ориентированной архитектуры при разработке ПО системы. На основе этих принципов разрабатывается модель её ПО, а также проектируются интерфейсы пользователей и концептуальная модель базы данных ИИС.

2.1. Основные требования к ИИС и принципы её создания

До этапа проектирования любой ИС необходимо сформировать и структурировать требования к ней в виде технического задания (ТЗ). Оно позволяет всем заинтересованным лицам совместно сформировать желаемый образ ИС и отразить свои взгляды в ТЗ. При этом важно подойти к составлению этого документа комплексно. Все требования должны быть чётко сформулированы и классифицированы по определённым признакам, также должны быть указаны взаимосвязи между требованиями. Не допускается использовать противоречивые и нереалистичные требования. Каждое требование должно соответствовать только одной неделимой единице и иметь уникальный идентификатор для простоты адресации.

Одной из известных методологий разработки ТЗ в инженерии требований к системам является методология А. В. Симкина [42]. В соответствии с ней были сформированы группы требований к ИИС для управления фондом скважин. При этом в качестве основы при формировании перечня требований к функциональности ИИС была взята функциональность следующих отечественных ИС для добывающих предприятий нефтегазовой отрасли: отраслевой стандарт ИС, принятый в ПАО «Газпром» [15], Корпоративная геоинформационная

система управления (КГСУ) производством ОАО «Томскгазпром» [32], а также производственная информационно-аналитическая система ПАО «Сургутнефтегаз» [43].

Основные группы требований к ИИС, включаемые в ТЗ в нашем случае, разработаны в соответствии с методологией Симкина А. В. и приведены в таблице 2.1. Они содержат не только функциональные группы (F, FA), но и группы нефункциональных требований (I, DR, DF, A и другие), варианты использования ИИС (V) и т. д. Все приведённые группы включают более двухсот требований. Дадим краткое обоснование необходимости и специфики формирования основных групп требований к ИИС.

Таблица 2.1 – Группы требований к ИИС

Символ	**Группа требований**
F	Общие функциональные требования
V	Варианты использования ИИС
FA	Требования к методам и алгоритмам обработки и анализа данных
I	Требования к интерфейсу пользователя
DR	Требования к составу данных
DF	Требования к представлению данных
D	Требования к описанию данных
R	Требования к отчетам
C	Требования к управлению справочниками и классификаторами
A	Требования к администрированию, управлению доступом и безопасностью системы
AR	Требования к правам доступа
TS	Требования к техническому обеспечению
IS	Требования к информационной безопасности системы

Общие функциональные требования (группа F), во-первых, содержат требования к способам взаимодействия пользователя с ИИС: с помощью интернет-браузеров, при чем должны использоваться версии браузеров с минимально достаточным набором функций. Это позволит разработать в ИИС

интерфейсы пользователя для решения каждой задачи управления фондом скважин. Доступ к ИИС осуществляется по URL, сервер располагается в корпоративной компьютерной сети нефтегазодобывающего предприятия.

Во-вторых, ИИС должна иметь обязательную процедуру авторизации и аутентификации пользователей (требования группы F в части этой процедуры взаимосвязаны с требованиями групп A, AR и IS). Корректность введённых данных должна проверяться на стороне клиента за счёт регулярных выражений и на СУБД сервера. При этом на стороне клиента должны сохраняться в файле cookie все пользовательские настройки, касающиеся оформления и используемого языка.

В ИИС должны быть реализованы наборы функций (группы требований F и FA), позволяющие решать все поставленные задачи управления фондом скважин, в том числе с использованием наиболее перспективных моделей, методов и алгоритмов ИАД. При сборе и актуализации исходных геологических и технологических данных важно иметь программные средства для предварительной обработки больших массивов данных на основе принципов «Big Data». Требуются средства для заполнения справочников и классификаторов (группа C), ведения форм для ввода предложений от служб промысла по остановкам скважин, в том числе для ГТМ, ведения журналов остановок скважин и расписания работы бригад КРС и т. п. При этом необходимы средства, реализующие функции дополнения и корректировки введённой или полученной информации.

Основным общесистемным требованием из группы F к ИИС является требование интеграции ИИС с унаследованными ИС и АСУ ТП предприятия, которые для неё являются источниками данных. Для его реализации в первую очередь следует разработать средства интеграции ИИС с различными АСУ ТП или с производственными ИС класса MES и через них получить данные из этих АСУ ТП. Интеграция ИИС с такими системами необходима для доставки в неё технологических и части геологических данных мониторинга скважин и продуктивных слоев. Требуется также интеграция ИИС с ИС, где хранятся

данные о проведённых и планируемых ГИС, ГДИС и ПГИ, паспортные данные об объектах эксплуатации и утверждённые планы-графики работ бригад КРС.

Разработан первый принцип создания ИИС – принцип минимизации её функциональности. В соответствии с ним в ИИС не следует реализовывать традиционные (известные) методы и алгоритмы управления фондом скважин, однако должна быть осуществлена её интеграция с ИС предприятия, в которых они реализованы. Среди таких методов и алгоритмов проанализированные в разделе 1 и наиболее часто используемые на практике: расчёт геологического потенциала, анализ трендов значений параметров скважин и продуктивных пластов и метод задания допустимых границ интервалов для значений каждого параметра, а также методы статистического анализа оперативных остановок скважин. Интеграция систем необходима специалистам, в первую очередь, геологической службы, для сравнения результатов, полученных с помощью таких методов, с результатами, полученными в ИИС методами ИАД (методами кластерного анализа, ИНС и т. д.). Для удобства работы пользователей с такими системами через ИИС необходимо реализовать принцип одного окна: вызов традиционного метода (алгоритма), реализованного в другой ИС, и сравнение результатов, полученных этим методом и методом ИАД должны осуществляться в ИИС. Этот принцип необходимо учесть при разработке интерфейса пользователя. Также при его реализации интегрировать системы необходимо не только по данным, но и по функциям.

Для оценки технологической и экономической эффективности планируемых ГТМ на выбранных скважинах-кандидатах должны быть реализованы функции получения профилей планируемой добычи, расчета показателей инвестиционного проекта, оценки возможных отклонений от планируемых технологических режимов скважин, а также функция оценки равномерности разработки месторождения. ИИС должна позволять формировать оперативный план-график работы бригад КРС с возможностью его оптимизации.

В требованиях к вариантам использования ИИС (группа V) описываются роли каждого пользователя и действия, выполняемые им в ходе работы в виде

сценариев (последовательности шагов) [44]. При этом шаги сценария могут содержать ссылки на требования из других групп FA, DF, DR, I, если это необходимо для уточнения. Среди ролей пользователей ИИС можно выделить 3 типа ролей.

- Администратор. Несёт ответственность за мониторинг состояния ИИС и управление пользователями, а также бесперебойное функционирование системы.
- Специалист, отвечающий за ввод данных в БД системы через заполнение ряда форм и справочников и за задание значений некоторых параметров методов и алгоритмов ИАД. Решает также с помощью ИИС основные задачи по управлению фондом скважин.
- Конечный пользователь, участвующий в наполнении и актуализации БД системы данными из других ИС, интегрированных с ИИС. Ведет также согласование и утверждение полученных, в том числе на промежуточном этапе, результатов.

Требования к интерфейсу пользователя (группа I) необходимы для лучшего понимания программистами-разработчиками ИИС структур данных (требования групп DR, DF, D) при работе пользователей в соответствии со сценариями. Любой пользователь должен иметь дружественный интерфейс для доступа к информации с помощью браузера (управление вкладками, окнами, инструменты навигации), переходя по гиперссылкам и кнопкам пользовательского интерфейса. При этом ИИС должна следить за корректностью введённых пользователем данных и выдавать соответствующее оповещение при некорректном вводе. Табличное представление данных должно иметь функции сортировки и фильтрации. Любой выполняемый процесс должен сопровождаться отображением прогресса его выполнения с оценкой оставшегося времени ожидания.

Основным комплексным (системным) требованием из группы F к ИИС является её гибкость во многих аспектах: по отношению к перечню сформулированных требований к ней, к моделям месторождения и бизнес-процессам, выполняемым на промысле, к используемым в ней форматам данных и

программным платформам, а также к перечню реализуемых методов и алгоритмов, в том числе ИАД. ИИС должна легко интегрироваться с унаследованными ИС (MES, системы 3Д-моделирования продуктивных пластов, построения геологических разрезов и т. д.) и с АСУ ТП предприятия, из которых в неё поступают данные. Такая гибкая ИИС должна быть многокомпонентной системой, при этом должна состоять из отдельных слабосвязанных модулей, которые предназначены для решения прикладных задач управления фондом скважин. Такие модули должны иметь возможность легко заменяться, модифицироваться и дополняться так часто, как этого требует стратегия развития предприятия, учитывающая необходимость внедрения инновационных технологий и ситуацию на рынке УВС.

В настоящее время на многих добывающих предприятиях нефтегазовой отрасли используется большое число разнообразных и разнородных производственных ИС и АСУ ТП. В некоторых из этих ИС реализованы также отдельные традиционные методы и алгоритмы решения задач управления фондом скважин. Выше выдвигалось требование интегрировать создаваемую ИИС с теми из систем, которые являются источниками данных для нее. Для эффективного управления производством предприятия, в том числе фондом скважин, можно пойти дальше и реализовать предлагаемый принцип комплексной интеграции на основе программного обеспечения ИИС всех унаследованных производственных ИС и АСУ ТП предприятия. В итоге можно получить единое информационное пространство предприятия [45], в котором центральная связующая роль отводится ИИС. Реализация такого пространства даёт специалистам предприятия ряд преимуществ, например, отсутствие необходимости двойного ввода данных, упрощение технологий поддержания целостности баз данных и т. п.

2.2. Концепция сервисно-ориентированной архитектуры программного обеспечения ИИС

Разнообразие ИС, применяемых на предприятиях нефтегазовой отрасли, – это результат влияния многих факторов при их разработке и внедрении:

применение разных СУБД, использование различных языков программирования при реализации ПО этих систем, различные разработанные архитектуры систем и т. п. При этом все данные предприятия хранятся в виде множества файлов и (или) в БД этих ИС. Чтобы поддерживать целостность такой гетерогенной системы специалистам предприятия необходимо прикладывать значительные усилия. Поэтому возникает задача интеграции всех ИС предприятия по функциям и данным и, в итоге, реализуется концепция единого информационного пространства предприятия. Эта задача решается в основном тремя способами [46, 47]: интеграция между отдельными ИС (точка-к-точке); использование центрального элемента с логикой взаимодействия систем и преобразования форматов данных; стандартизированная и унифицированная организация взаимодействия систем предприятия, реализованных в виде сервисов (веб-сервисов).

Последний способ реализуется на основе принципов концепции сервисно-ориентированной архитектуры ПО [48, 49]. Он позволяет удовлетворить перечисленным выше основным требованиям, предъявляемым к ИИС добывающего предприятия: гибкость ИИС во многих аспектах, возможность легко интегрироваться с унаследованными ИС и АСУ ТП предприятия и возможность стать основой единого информационного пространства предприятия. Среди принципов концепции SOA, наиболее важных в наших исследованиях, можно отметить: модульный принцип построения ПО ИС в виде сервисов (веб-сервисов), способность архитектуры ПО ИС подстраиваться под изменения рыночного и технологического характера и минимум зависимости от форматов данных, протоколов передачи информации и платформы. Для нас SOA – это способ создания различных систем, в основе которого лежит применение легко заменяемых, распределённых, слабо связанных модулей, называемых сервисами [50].

Кратко рассмотрим компоненты SOA, важные при разработке архитектуры программного обеспечения ИИС управления фондом скважин. Первая компонента SOA – сервис. Под ним понимается самостоятельный программный

модуль, выполняющий определённую задачу (алгоритм, бизнес-процесс, подпроцесс) и предоставляющий результат своей работы другому сервису или информационной системе, которые этот сервис вызывают. Сервисы снабжены стандартизированными сетевыми интерфейсами и взаимодействуют только по стандартизированным протоколам [50, 51]. Сервисы могут использоваться как сами по себе, так и в комбинации с другими сервисами с последовательным или одновременным запуском, также они могут быть комплексными, то есть состоять из множества других сервисов.

Другой компонентой SOA является реестр сервисов. По сути – это каталог, он содержит описания сервисов (физическое месторасположение сервисов, версии, а также длительность их действия и документацию по сервисам).

Следующий компонент SOA – это сервисная шина предприятия (англ. Enterprise Service Bus – ESB) [50]. Шина представляет собой центральное связующее программное обеспечение, позволяющее организовать унифицированный и централизованный событийно-ориентированный обмен сообщениями между всеми сервисами. Шина обеспечивает при интеграции безопасность, маршрутизацию сообщений, преобразование их форматов, логирование событий, поддержку транзакций и множества протоколов сетевого взаимодействия.

Ещё одна обязательная компонента SOA – это система управления бизнес-процессами (англ. BPMS – Business Process Management System), она представляет собой ПО с возможностями для запуска, координирования, управления, администрирования и отладки бизнес-процессов. При этом каждый БП рассматривается как некоторая последовательность логически связанных действий, которые преобразуют входные данные в результат (обычно выходные данные). Именно БП определяют через BPMS последовательность запуска сервисов с целью решения определённой бизнес-задачи. BPMS призвана помочь реализовать эту задачу с помощью понятных даже не разработчику ИИС нотаций, самой популярной из которых является BPMN (англ. Business Process Model and Notation – BPMN) [45, 52]. Язык BPEL (англ. Business Process Execution Language

– BPEL) [53] используется для определения координации запуска бизнес-процессов, реализованных в виде веб-сервисов, но зачастую сама BPMS имеет возможность конвертировать структуры на этом языке по диаграммам BPMN.

Самыми часто используемыми технологиями в концепции SOA являются [54]: язык WSDL (англ. Web Services Description Language – WSDL), который обеспечивает исчерпывающее описание функциональных возможностей сервиса и способ доступа; протокол SOAP (англ. Simple Object Access Protocol – SOAP), содержащий XML-формат сообщений как отправляемых, так и принимаемых сообщений. В отличие от протокола SOAP, REST является архитектурным стилем, при реализации которого обязательным является использование протокола HTTP и какого-либо обменного формата (JSON, XML). Считается, что Restful-сервисы реализовать проще [55], так как они не требуют обязательного описания службы на языке WSDL, а также имеется возможность работать с менее тяжеловесными форматами. Это предъявляет менее серьёзные требования к пропускной способности каналов связи и используемым ресурсам, в отличие от использования SOAP. Однако REST не покреплён стандартом и основывается на соглашениях, что делает его использование менее безопасным.

Таким образом, краткий анализ составляющих SOA и их основных возможностей позволяет считать, что разрабатываемая на принципах и компонентах SOA архитектура программного обеспечения ИИС позволит удовлетворить многим комплексным требованиям, предъявляемым к проектируемой ИИС.

С учётом всего вышеизложенного на основе принципов SOA была разработана архитектура программного обеспечения ИИС. Укрупнённая схема архитектуры ПО этой ИИС, входящей в состав единого информационного пространства добывающего предприятия, изображена на рисунке 2.1.

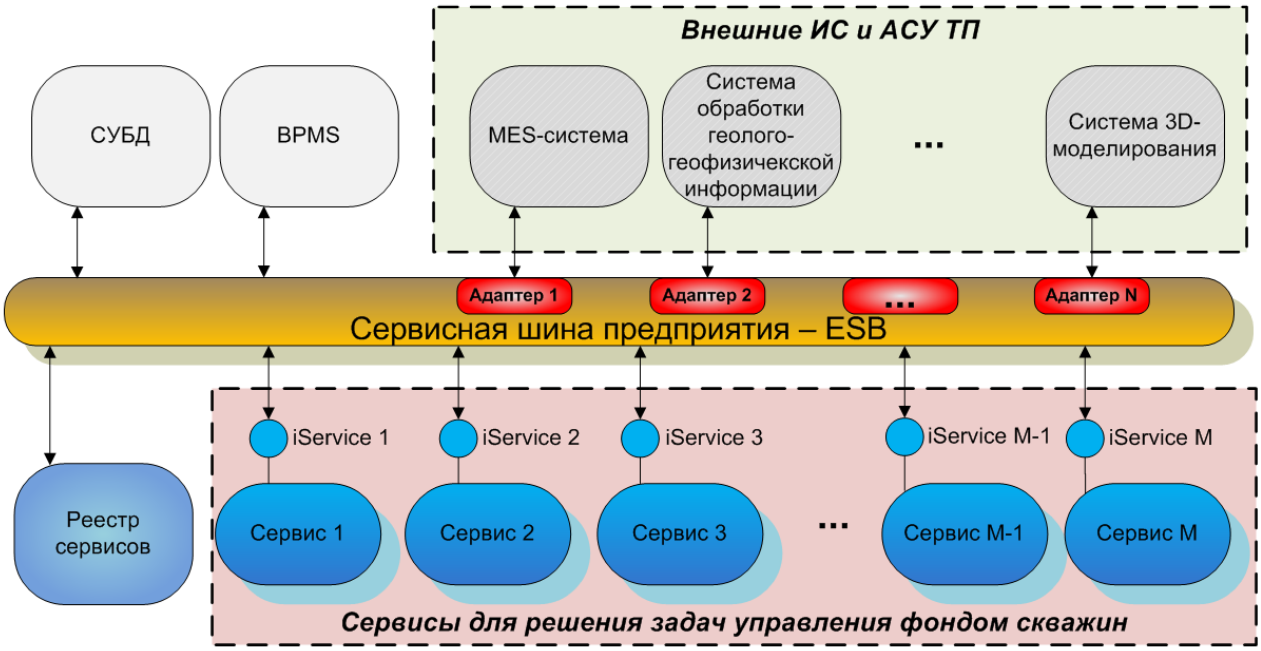

Рисунок 2.1 – Укрупнённая схема программного обеспечения единого информационного пространства добывающего предприятия нефтегазовой отрасли

Рассмотрим особенности архитектуры ПО такой ИИС. Основная компонента этой архитектуры – сервисная шина предприятия ESB. Другими основными компонентами архитектуры ПО являются BPMS для управления сервисами и СУБД для управления данными. За счёт сервисной шины предприятия происходит обмен сообщениями между сервисами ИИС и внешними по отношению к создаваемой ИИС унаследованными ИС и АСУ ТП нефтегазодобывающего предприятия. Как пример внешних ИС на рисунке 2.1 приведены три системы из N систем (они наряду с СУБД и BPMS расположены выше сервисной шины). Первая из этих систем – MES (англ. Manufacturing Execution System – MES) [56] – предназначена для оперативного управления производственными процессами нефтегазодобывающего предприятия. Вместо столь крупной корпоративной системы, как MES, предприятие может иметь ряд более простых систем, например, систему диспетчерского управления, ИС для химического анализа УВС и т. д. Из таких систем и (или) АСУ ТП сервисы ИИС должны получать данные об остановках скважин, времени их простоя и т. д. Другими примерами внешних ИС являются система обработки (интерпретации) геолого-геофизической информации и система 3D-моделирования продуктивных пластов. Из них сервисы ИИС получают, например, значения геологических

параметров продуктивных пластов. Если БД других внешних ИС хранят данные, нужные для выполнения сервисов ИИС, то существует возможность подключения этих ИС, обеспечиваемая шиной предприятия за счёт стандартизированных интерфейсов. Многоточие показывает, что может быть подключён ещё ряд ИС.

Внешние по отношению к ИИС информационные системы и АСУ ТП не всегда имеют стандартизированные интерфейсы, так как они могли создаваться не в рамках концепции SOA. Поэтому для них должны быть реализованы специальные модули в ESB, называемые адаптерами. На рисунке 2.1показаны три адаптера, а три точки означают, что могут быть разработаны и использоваться ещё адаптеры для других ИС и АСУ ТП предприятия. На рисунке 2.1 также наряду с реестром сервисов показаны М сервисов для решения задач управления фондом скважин. Каждый сервис снабжён стандартизированным интерфейсом, обозначенным на рисунке 2.1 как «iService 1», «iService 2» и т. д.

Разработанная на основе концепции SOA архитектура программного обеспечения ИИС названа SOA-моделью ПО создаваемой системы. Нетрудно видеть, что в приведённой на рисунке 2.1 укрупнённой схеме единого информационного пространства добывающего предприятия SOA-модель играет основную связующую роль при интеграции её с ПО всех унаследованных ИС и АСУ ТП предприятия.

2.3. Проектирование интерфейса пользователей ИИС

В соответствиями с требованиями к ИИС на этапе её проектирования были разработаны эскизы пользовательского интерфейса системы. На рисунке 2.2 показан эскиз пользовательского интерфейса главной страницы одной из подсистем ИИС, на котором можно выделить следующие области.

1. Панель основных БП управления фондом скважин (на рисунке 2.2 в качестве примера приведены основные БП управления ГТМ на фонде скважин);

2. Панель вспомогательных бизнес-процессов. К ним относится также процесс «Чат», позволяющий общаться пользователям в чате ИИС. Здесь же размещаются кнопки запроса к другим ИС предприятия: «Система 3Д-

моделирования», «Создание отчётности» и т. д. При выполнении запроса ИИС подключается к той или иной ИС. Пользователь, выполняющий роль Администратора, входит в подсистему администрирования через нажатие кнопки «Администрирование»;

3. Рабочая область, где отображается содержимое (в виде результатов выполнения БП, промежуточных данных при проведении вычислений, сообщений и т. п.) при работе с любыми основными и вспомогательными БП и справочной информацией;

4. Справочная панель. Позволяет получить справку об ИИС и контактах её разработчиков.

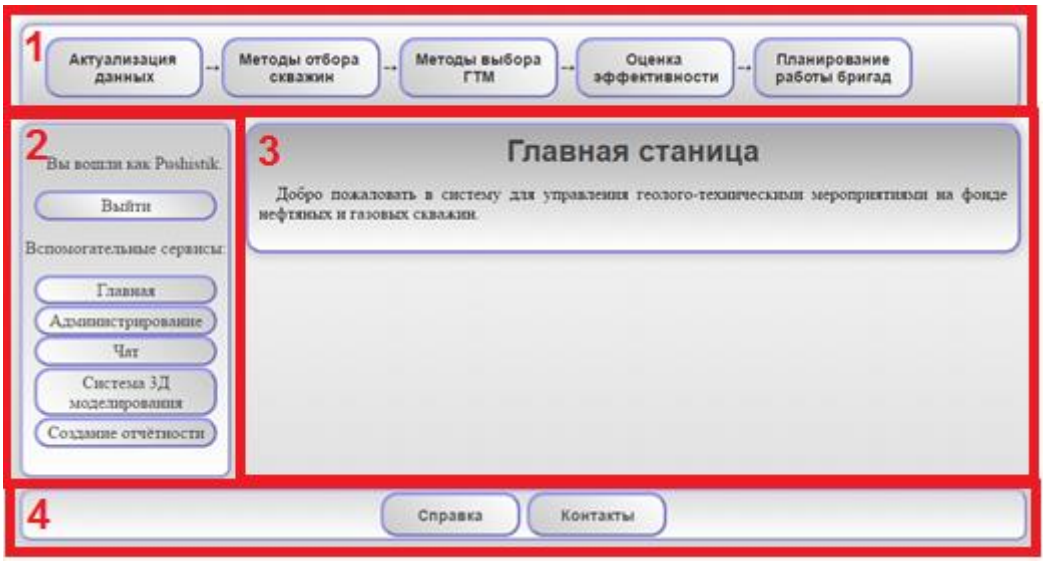

Рисунок 2.2 – Интерфейс главной страницы подсистемы «Управление ГТМ»

При работе с любыми объектами ИИС, справочниками и классификаторами должны поддерживаться все операции CRUD с данными: чтение, редактирование, добавление и удаление. На рисунке 2.3 в качестве примера показан пользовательский интерфейс, где каждый атрибут соответствующей таблицы отображается в корректной для него форме (текст, числовой, дата-время, список и т. д.).

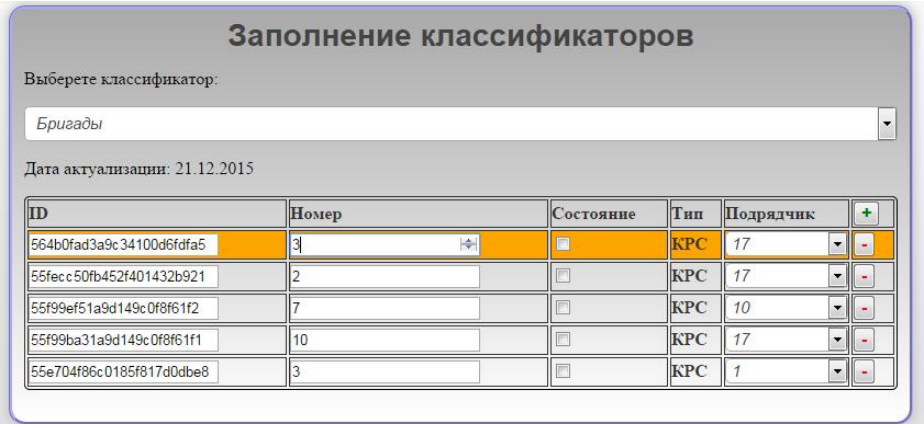

Рисунок 2.3 – Интерфейс при заполнении классификаторов и корректировке данных

При решении задачи отбора скважин-кандидатов для ГТМ важно дать пользователю возможность выбрать интересующее его месторождение и скважины, а также отметить конкретный выполняемый автоматизированно или автоматически метод или алгоритм, который будет использован при выборе скважин после нажатия на кнопку «Начать» (рисунок 2.4). Если будет осуществлён выбор метода или алгоритма, реализованного в унаследованной ИС предприятия, то сначала происходит вызов этой системы.

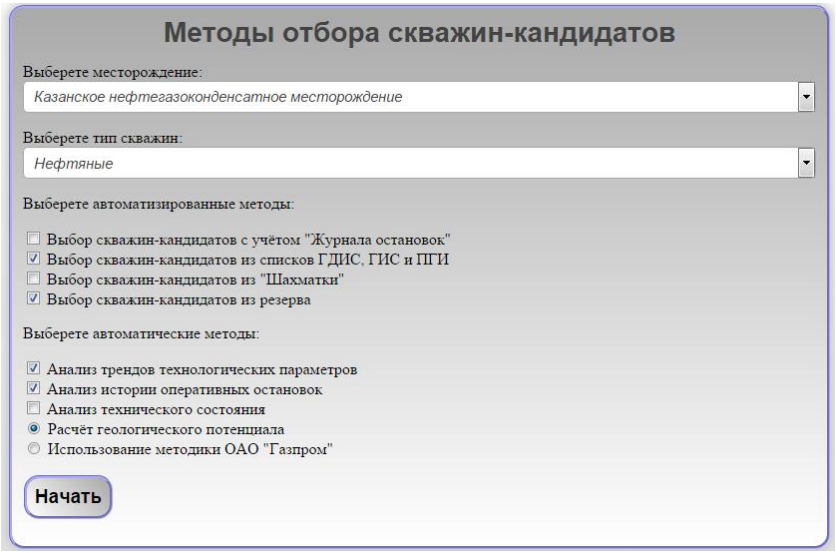

Рисунок 2.4 – Выбор методов (алгоритмов) для отбора скважин-кандидатов для ГТМ

В полученном списке скважин (рисунок 2.5) специалист может увидеть результаты работы выбранного метода или алгоритма или результаты применения ряда методов (алгоритмов) выбора скважин-кандидатов и результаты выбора ГТМ для них. Он может произвести требуемые CRUD-операции, а также сохранить полученные результаты, загрузить уже существующий список скважин-

кандидатов для ГТМ или отправить такой список на согласование всем заинтересованным лицам. Стоит отметить имеющуюся возможность оставить комментарий к каждой скважине-кандидату, выбрать планируемый временной интервал для проведения ГТМ и скомпоновать несколько подобных списков.

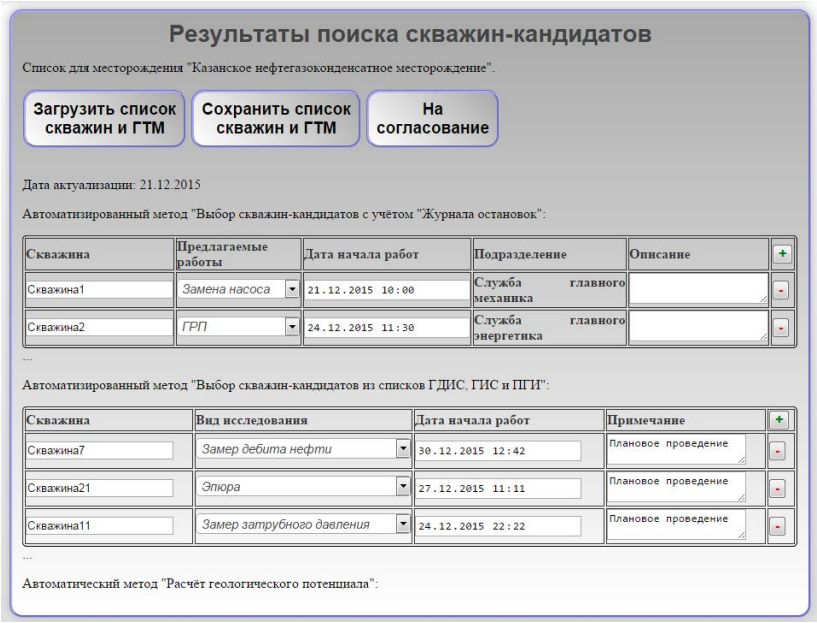

Рисунок 2.5 – Представление результатов выбора скважин-кандидатов на ГТМ различными методами и выбора конкретных ГТМ для них

После отправки на согласование, список становится доступным для просмотра всем заинтересованным службам добывающего предприятия. При согласовании можно отредактировать записи в списке, добавить новые, удалить старые, оставить комментарий в зависимости от предпочтений той или иной службы (рисунок 2.6).

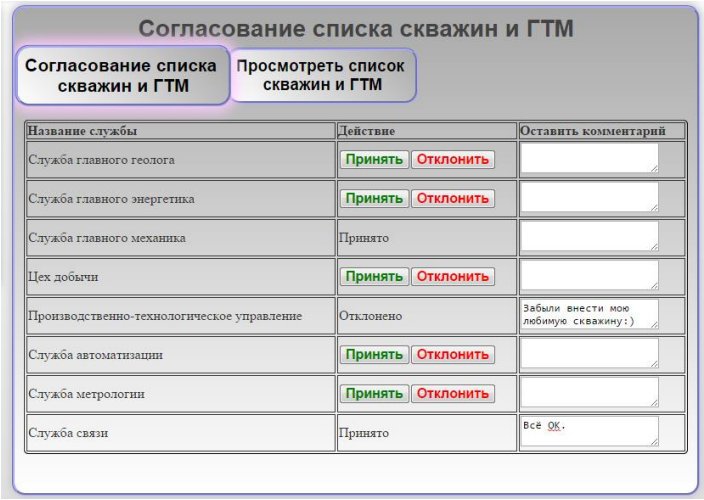

Рисунок 2.6 – Интерфейс пользователя при согласовании списка скважин и ГТМ

При оценке технологической и экономической эффективности каждой скважины из списка скважин-кандидатов для ГТМ и ГТМ для них ИИС проводит расчеты и демонстрирует пользователю различные метрики по коэффициентам инвестиционного проекта и (или) показателям технологического режима по каждой скважине (рисунок 2.7). Можно пересмотреть список всех скважин-кандидатов и при необходимости внести коррективы и уточнения.

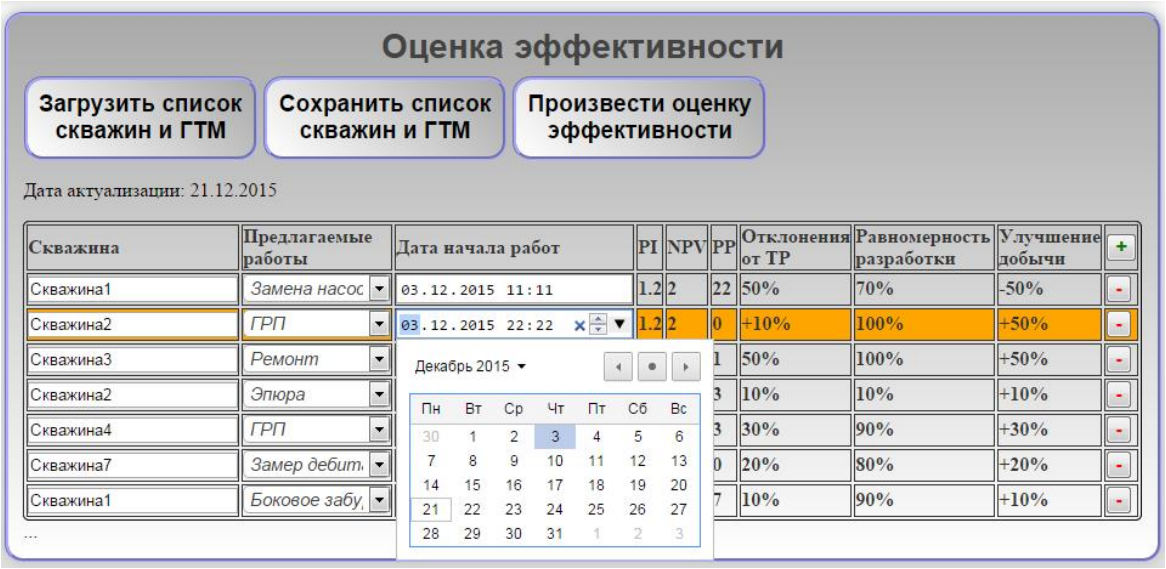

Рисунок 2.7 – Результаты расчета эффективности ГТМ и уточнения списка скважин и ГТМ

Составление плана-графика работы бригад КРС для проведения ГТМ происходит автоматически после загрузки конкретного списка скважин-кандидатов для ГТМ и ГТМ для них (рисунок 2.8). При этом должны учитываться планы-графики работ по исследованию (ГИС, ГДИС и ПГИ) тех же скважин, связанные с их остановками. Такой план-график может быть отредактирован, сохранён, а также загружен вариант ранее сохранённого плана-графика. В ИИС должны быть реализованы различные методы и алгоритмы для оптимизации такого графика. При возникновении конфликтов при распределении работ по бригадам КРС и бригадам, выполняющим работы по исследованию скважин, система должна выделять цветом такие работы. Например, на рисунке 2.8 графики конфликтующих работ на двух разных месторождениях, распределяемые на бригаду КРС № 1, показаны красным цветом. Пользователь должен иметь возможность разрешать конфликты вручную, а также иметь возможность применять специальные средства оптимизации плана-графика.

Оптимизированный и проверенный на наличие конфликтов план-график отсылается на утверждение.

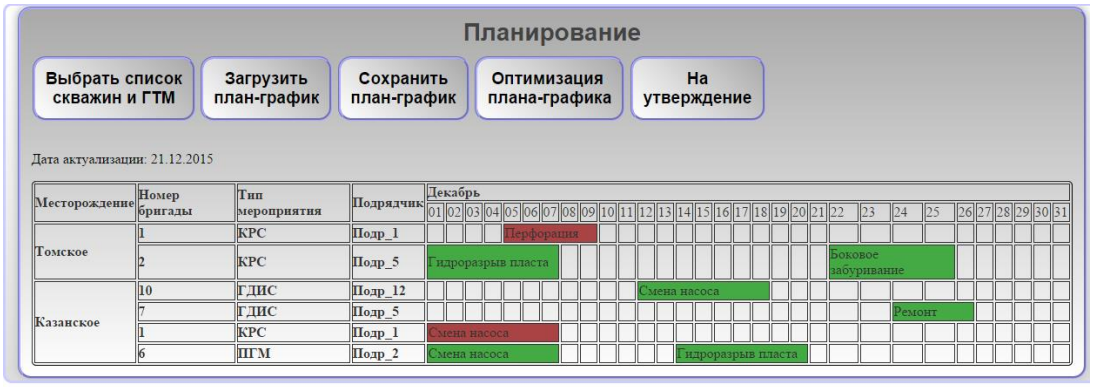

Рисунок 2.8 – Интерфейс при формировании оперативного плана-графика работ

2.4. Концептуальная модель базы данных ИИС

Для формирования концептуальной модели и в последующем структуры БД ИИС изначально необходимо определить перечень сущностей (объектов), которые следует описать и использовать в модели. Среди наиболее значимых из них можно отметить: Месторождение, Куст скважин, Скважина, ГТМ, Работа на скважине, Остановка скважины, Исследование, МЭР, Параметр скважины, Параметр пласта, Плановое значение параметра и другие сущности.

Атрибутами сущности Скважина являются следующие данные о каждой скважине: текущий забой, диаметр штуцера, категория, диаметр насосно-компрессорных труб (НКТ), состояние, радиус, время работы в сутки, давление буферное, давление забойное, давление затрубное, интервал перфорации, радиус контура питания, дебиты по нефти, газу, газовому конденсату и воде, межремонтный промежуток и прочие атрибуты. Для описания Кустов скважин предусмотрены атрибуты: извлекаемые и плановые объемы (количества) нефти, газа, газового конденсата и воды, количество скважин в кусте по плану. У Параметров пластов должны быть атрибуты – характеристики вязкости, скин-фактор, пластовое давление, температура, обводнённость, плотность, глубина, мощность, проницаемость, объёмный коэффициент и наличие пропластков. Аналогичным образом описаны атрибуты других сущностей и выявлены и учтены связи между сущностями.

Также в модель БД включаются различные справочники и справочные значения ряда параметров (атрибутов) сущностей, данные оборудования, операции (работы) на фонде, сценарные условия и нормативы. Кроме того, были введены дополнительные сущности: Отчёты, Расписания бригад и т. д. Каждый атрибут в сущности имеет определённый домен и тип: целочисленный, числовой, дата-время, файловый, текстовый, логический, список и т. п.

ER-модель (концептуальна модель) БД была создана в дизайнере схем Toad Data Modeler в нотации Crow’s Foot [57]. В соответствии с правилами формирования таких схем внешние ключи в сущностях не указываются, связи один-к-одному и многие-ко-многим преобразуются соответственно путём объединения сущностей в одну, либо путем формирования промежуточной сущности. Фрагмент разработанной ER-модели БД приведён на рисунке 2.9.

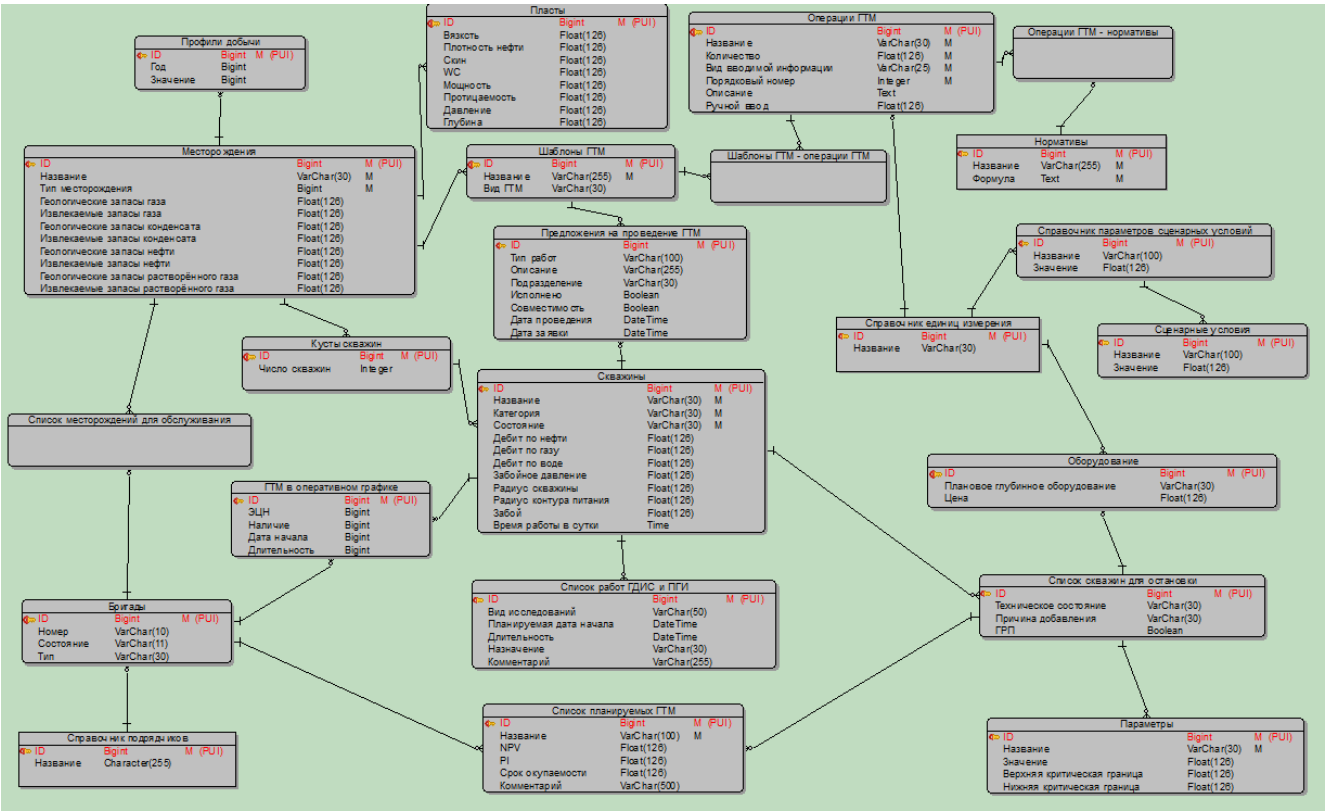

Рисунок 2.9 – Фрагмент концептуальной модели данных БД системы

Дизайнер схем Toad Data Modeler даёт возможность разработчику по созданной ER-модели сформировать код и впоследствии развернуть физическую модель данных под выбранную СУБД. Это будет сделано на одном из следующих этапов создания ИИС.

2.5. Основные результаты и выводы по разделу

Одной из важных задач, решаемых для достижения цели диссертационной работы, является разработка концепции ИИС управления фондом скважин добывающего предприятия нефтегазовой отрасли. В разделе рассматриваются основные аспекты предлагаемой концепции такой ИИС. При этом получен ряд новых результатов и сделаны соответствующие выводы.

1. В соответствии с методологией разработки требований к системам, развитой Симкиным А. В., сформированы основные группы требований к ИИС. Показано, что кроме общесистемных функций в создаваемой ИИС должны быть реализованы наборы функций, позволяющие решать все поставленные задачи управления фондом скважин, в первую очередь, с использованием наиболее перспективных моделей, методов и алгоритмов ИАД. Такие наборы функций позволят обеспечить высокую степень автоматизации БП при управлении фондом скважин. Показано, что ИИС должна быть не только многофункциональной, но и должна удовлетворять ещё целому ряду требований.

2. Показано, что среди сформированных наборов требований к ИИС для управления фондом скважин наиболее важными требованиями, составляющими основу концепции её создания, являются комплексные требования гибкости ИИС во многих аспектах и возможности легко интегрироваться с унаследованными ИС и АСУ ТП предприятия. ИИС должна иметь механизмы гибкой модернизации при изменениях технологического характера на предприятии и при изменениях на рынке УВС.

3. Разработаны три основополагающих принципа в концепции построения ИИС. Среди них принцип минимизации набора функций ИИС и принцип одного окна в ИИС для работы пользователя через её среду с другими производственными ИС. Третьим является принцип комплексной интеграции на основе программного обеспечения ИИС всех унаследованных производственных ИС и АСУ ТП добывающего предприятия при создании его единого информационного пространства.

4. Рассмотрены основные способы решения задачи создания единого информационного пространства предприятия. Сделан вывод, что наиболее перспективным является способ, опирающийся на принципы концепции сервисно-ориентированной архитектуры (англ. SOA) ПО.

5. Проанализированы компоненты SOA и их основные возможности, наиболее важные при разработке концепции ИИС. Анализ позволил сделать вывод, что разрабатываемая на принципах и компонентах SOA архитектура программного обеспечения ИИС позволит удовлетворить основным системным требованиям, предъявляемым к создаваемой ИИС.

6. На основе принципов SOA разработана архитектура программного обеспечения ИИС, названная SOA-моделью ПО такой системы, и описаны её компоненты. Показано, что SOA-модель, являясь главным связующим звеном, позволит создать единое информационное пространство добывающего предприятия с большим числом унаследованных производственных ИС и различных АСУ ТП.

7. Для эффективного взаимодействия пользователей с ИИС при анализе больших объёмов технологических и геологических данных разработаны эскизы интерфейсов пользователей.

8. С помощью дизайнера схем Toad Data Modeler спроектирована концептуальная модель базы данных ИИС.

3. МОДЕЛИ, МЕТОДЫ И АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ИИС

В этом разделе рассматриваются модели, методы и алгоритмы обработки и анализа значений технологических и геологических параметров, реализуемые в ИИС. Приводятся результаты исследования их эффективности.

3.1. Постановки задач исследований

3.1.1. Слабоформализуемые задачи управления фондом скважин

Рассмотрим постановки ряда задач, при решении которых необходимы исследования:

Задача 1. При управлении фондом скважин одним из важных выполняемых бизнес-процессов является БП «Ежемесячное планирование технологических режимов работы каждой скважины фонда». Реализация этого БП, в основном, сводится к решению слабоформализуемой задачи прогноза значений дебитов УВС каждой добывающей скважины. В принципе, при решении этой задачи можно использовать различные методы прогнозирования временных рядов. Один из перспективных методов – это использование моделей ИНС разных классов. На наш взгляд, перечень моделей ИНС должен выбираться исходя из реальной ситуации на фонде скважин месторождения. Поэтому сформулируем задачу исследования эффективности различных моделей ИНС при прогнозе значений дебитов УВС. То есть необходимо решать слабоформализуемую задачу прогноза значений дебитов скважин (далее для краткости *Задача 1)*. При этом в качестве исходных данных для ИНС будем использовать реальные данные с эксплуатируемых месторождений. Для экспериментов в качестве исходных данных из класса технологических данных выберем исключительно параметры дебитов (ежесуточные объёмы добычи компонентов УВС – нефти, газа, газового конденсата и воды), так как они, по мнению многих исследователей [58], являются наиболее значимыми интегральными показателями добычи УВС [59]. Для них в БД каждого предприятия обычно имеется наибольший объём данных.

Задача 2. Другой слабоформализуемой задачей является задача выбора скважин-кандидатов на ГТМ, сводящаяся к задаче бинарной классификации всех

добывающих скважин фонда. Она решается с помощью методов кластерного анализа и (или) ИНС. В первый класс при исследованиях должны попасть все скважины, которые планируются к проведению определённого типа (вида) ГТМ. Это класс «скважин-кандидатов для ГТМ». Основанный на опыте и интуиции специалистов и проведённый ранее на фонде ручной выбор специалистами-геологами скважин-кандидатов будет считаться эталоном при формировании обучающей выборки, если после проведённого ГТМ произошло существенное увеличение дебита скважины. Во второй класс скважин отнесём все остальные скважины фонда (класс «скважин не для ГТМ»), на которых в момент времени, когда принималось решение, по той или иной причине ГТМ проводить не следовало. Например, это высокодебитные скважины, дающие основной объём добычи УВС на месторождении. Отметим, что решение этой слабоформализуемой задачи важно при выполнении БП «Управление и проведение ГТМ».

Задача 3. Другой задачей, решаемой при управлении ГТМ, является слабоформализуемая задача разделения ГТМ на классы и выбора класса (типа) ГТМ для каждой скважины-кандидата. Среди ГТМ нами на основе анализа мероприятий на нескольких месторождениях было выделено четыре класса (типа) мероприятий, обычно примерно в одинаковых пропорциях, проводимых на скважинах фонда:

1. ремонтно-изоляционные работы (РИР) (замена НКТ, ликвидация межколонных давлений и перетоков);
2. использование электроцентробежных насосов (ЭЦН) (перевод скважин на механизированный способ добычи, спуск ЭЦН, смена ЭЦН);
3. ГРП;
4. прочие мероприятия (ремонт наземного оборудования скважин, смена фонтанной арматуры или пакеров и т. п.).

По сути, эта задача является задачей классификации для четырёх классов.

При решении *Задачи 2* предлагается использовать различные методы кластерного анализа и модели ИНС, а при решении *Задачи 3* – использовать модели ИНС. В качестве входных параметров для таких методов и моделей ИНС

будем использовать различные наборы геологических и технологических параметров. А в качестве эталонных меток при обучении ИНС и проведении кластерного анализа данных будем использовать результаты решения этих задач классификации специалистами предприятий. Для проведения исследований в качестве исходных данных использовались данные из архивов различных эксплуатируемых месторождений.

Итак, необходимо провести исследования возможностей ряда современных методов ИАД при решении указанных выше трёх слабоформализуемых задач.

3.1.2. Методы и подходы к предварительной подготовке исходных данных

Геологические и технологические параметры фонда скважин в целом обладают рядом особенностей. Во-первых, значения некоторых из них могут быть в те или иные моменты времени не измерены из-за остановок скважины и т. п. Во-вторых, появляются ошибочные по разным причинам значения параметров. В итоге, существует проблема актуализации значений параметров добывающих скважин и продуктивных пластов. Рассмотрим её подробнее.

Исходные данные для анализа в ИИС с целью выработки управляющих воздействий на фонд скважин могут поступать (собираться) в неё из БД различных ИС и (или) АСУ ТП предприятия. Зачастую эти БД содержат ошибочную, пропущенную и искажённую информацию о параметрах скважин и продуктивных пластов. Рассмотрим основные подходы к предварительной подготовке (обработке) таких данных. Известно, что каждый параметр скважины или продуктивного пласта месторождения должен находиться в определённом диапазоне значений. Этот диапазон зависит от особенностей месторождения и условий его эксплуатации. Используя такие диапазоны значений параметров, можно определить ошибочные или пропущенные данные, а для их исправления использовать различные зависимости. Например, для некоторых геологических параметров использовать закон Дарси [11] или метод интерполяции для исправляемых (пропущенных) значений по соседним корректным значениям технологического параметра. Ряд параметров, в первую очередь геологических, в силу своей природы не может изменяться резко. Так, например, вязкость,

эффективная мощность продуктивного пласта могут считаться постоянными в течение довольно продолжительного времени (обычно до года). Тогда в результатах ручного ввода значений таких параметров в БД легко распознать ошибку порядка числа и автоматически её исправить. Некоторые из этих подходов и алгоритмов уже использовались нами при подготовке данных для последующего ИАД [60].

При исправлении значений технологических параметров можно использовать рассчитанные технологические режимы (запланированные (прогнозные) на следующий месяц значения уровней добычи компонентов УВС). В идеале измеренные значения уровней добычи УВС должны совпадать с этими прогнозными значениями или незначительно отличаться от них. В действительности же различия могут быть существенны, особенно если в течение месяца произошли остановки скважины или на ней проводились ГТМ. Такие события необходимо отслеживать и учитывать в виде правил корректировки. Различные виды работ (текущие ремонты скважин и ГТМ) по-разному влияют на значения параметров скважины. Всё это должно быть учтено при разработке методов, алгоритмов и правил обнаружения в данных ошибок и их корректировки.

Для того чтобы каждый параметр оказывал влияние на конечный результат при анализе данных в ряде планируемых нами исследований, необходима нормализация исходных данных – преобразование их таким образом, чтобы шкалы разных параметров, характеризующих скважину и продуктивный пласт, оказались в сопоставимых интервалах. В таких исследованиях исходные значения по каждому параметру будут нормироваться так, чтобы они лежали в диапазоне от 0 до 1, причём значение, равное 1, получает максимальное значение конкретного параметра за весь период рассмотрения эксплуатируемого месторождения [61]. Это позволяет изначально уравнять потенциальную значимость каждого параметра перед исследованием, а также сохранить соотношения величин как по всему месторождению, так и между скважинами конкретных кустов.

Вместо нормализации исходных данных может быть применен метод нечеткой логики. При этом диапазон значений, которые принимает параметр, разбивается на лингвистические переменные, описываемые нечеткими множествами. В итоге для каждого из значений определяется степень принадлежности его к той или иной лингвистической переменной [24]. Однако этот метод ввиду его сложности нами не применялся.

Для предварительной обработки (актуализации) исходных данных с эксплуатируемых месторождений можно использовать различные подходы, методы и алгоритмы обработки постоянно растущих объёмов данных по фонду скважин в соответствии с концепцией «Big Data» [36]. Например, одним из источников дополнительной информации относительно геологических и технологических исходных данных могут служить результаты анализа изменения значений параметров скважин и продуктивных пластов во времени. Такие методы и алгоритмы анализа временных рядов частично использовались нами далее при предварительной подготовке (обработке) исходных данных для последующего исследования методов и алгоритмов ИАД. Они, в принципе, могут выступать и в качестве самостоятельного инструмента для анализа трендов, например, технологических параметров скважин и выявления в них критических тенденций.

Подводя итоги, можно сказать, что существует несколько различных методов и подходов к предварительной подготовке (обработке) данных для их последующего анализа. Выбор конечных подходов (методов) и алгоритмов зависит от природы исходных геологических и технологических данных и может быть осуществлен только в результате дополнительных исследований. Таким образом, весь накопленный в доступных нам БД объём геологических и технологических данных предварительно необходимо привести в форму, пригодную для решения слабоформализуемых задач методами ИАД и исследования эффективности этих методов.

3.1.3. Выбор типа ИНС для решения слабоформализуемых задач

Проанализируем ИНС различных типов с точки зрения решения с их помощью слабоформализуемых задач при управлении фондом скважин.

1. *Сети прямого распространения* (англ. Feed forward neural networks – FFNN) и *перцептроны* (англ. Perceptrons – P) [62]. Первая ИНС такого типа была предложена как простая модель хранения и запоминания информации в биологических нервных системах. В последующем она стала основой для многих современных нейросетей. ИНС этого типа весьма просты: они содержат слои сенсорных ассоциативных и реагирующих нейронов, то есть передают информацию от входа к выходам. Есть входной слой, промежуточные (скрытые) и выходной слои. Нейроны одного слоя не имеют связей между собой, при этом каждый нейрон имеет связь с каждым нейроном из соседних слоёв.

Особенность перцептронов как подтипа сетей прямого распространения состоит в том, что каждый P имеет только один скрытый слой. Это сильно ограничивает его функциональные возможности при решении задач прогноза и классификации, так как число признаков, которые может извлечь P является незначительным. Стоит отметить, что даже статистические методы при прогнозировании имеют явные функциональные преимущества по сравнению с перцептронами [63]. Более того, инвариантные представления образов не доступны для разграничения в P, либо в таких случаях они будут требовать нереальных вычислительных мощностей (память и (или) производительность процессора) [62].

FFNN, которые относятся к глубоким ИНС [64], в отличие от P имеют 2 и более скрытых слоёв. При выборе числа скрытых слоёв, адекватному сложности решаемой задачи, теоретически такие ИНС всегда способны установить связь между любыми входными и выходными данными. Большим преимуществом таких сетей является независимость от линейной и (или) нелинейной корреляции параметров на входе [65, 66]. Сложность использования FFNN заключается в том, что отсутствуют критерии выбора входных и выходных параметров таких сетей, также нет чёткого правила, сколько должно быть скрытых слоёв и каково число нейронов в них. В соответствии с результатами работ ряда исследователей [64, 67] FFNN часто используются в других областях знаний для решения задач

классификации и прогноза, что указывает на их потенциальную перспективность при решении слабоформализуемых задач управления фондом скважин.

Другим подтипом сетей прямого распространения являются *сети радиально-базисных функций* (англ. Radial basis function – RBF) [68]. Отличие их от описанных ранее ИНС заключается в использовании особых функций активации нейронов: если в обычных сетях прямого распространения чаще всего используются пороговые и сигмоидальные функции, то в этих сетях в хотя бы одном слое должны использоваться специальные радильно-базисные функции (функция Гаусса, мультиквадратичная, полигармоничный сплайн и другие). Эта особенность сетей радиально-базисных функций позволяет справиться с ситуациями с линейной неразделимостью классов, а также такие ИНС хорошо себя проявляют при прогнозировании временных рядов. Однако для устойчивой работы таких сетей и для получения высокоточных результатов число нейронов в слое радиально-базисных функций должно быть очень велико (в идеале – неограниченно), причём их центры не должны перекрываться и должны быть заданы заранее так, чтобы они равномерно покрывали всю область определения искомой функции. Так как эта область заранее неизвестна, то предугадать количество нейронов в слое и центры радиально-базисных функций очень сложно [69]. По этим причинам такой подтип нейросетей при решении поставленных задач использовать не рекомендуется.

2. *Нейронная сеть Хопфилда* [70], *Цепи Маркова* (англ. Markov Chain-MC) [71], *Машины Больцмана* (англ. Boltzmann machines – BM) [72], *Ограниченная машина Больцмана* (англ. Restricted Boltzmann machine – RBM) [73]. Данные ИНС являются аналогами ассоциативной памяти человека, они представляют собой полносвязные модели (каждый нейрон соединён с каждым – кроме нейросети RBM). Каждый нейрон является входным, скрытым и выходным одновременно (сеть Хопфилда) или же только часть нейронов функционирует как входная (нейросети BM и RBM). Один раз обученная ИНС будет сходиться при подаче входных данных к одному из известных образов, что можно использовать при решении задач классификации. Однако такие сети не имеют возможности

запоминать большое число различных образов. Также эти ИНС могут сходиться к ложным состояниям, именуемым «химерами» [74] (они состоят из фрагментов разных образов и не гарантируют правильный ответ). Более того, в процессе функционирования сети может произойти её зацикливание из-за использования коррелированных векторов-образцов [75]. Всё это не позволяет использовать такие нейросети при решении поставленных слабоформализуемых задач.

3. *Автоэнкодеры* (англ. Autoencoders – AE) [76], *Разряжённый автоэнкодер* (англ. Sparse autoencoder – SAE) [77], *Вариационный аутоэнкодер* (VAE) [78], *Шумоподавляющие автоэнкодеры* (англ. Denoising – DAE) [79]. Архитектура AE симметрична относительно среднего скрытого слоя, самого узкого по числу входящих в него нейронов, поэтому внешне схема ИНС напоминает песочные часы. Это позволяет извлекать из входных данных наиболее существенную информацию. SAE, напротив, необходим для извлечения деталей, его скрытые слои содержат больше нейронов, чем входные и выходные. Для обучения на выходе используются не просто входные данные, а они же, но плюс штраф за число активированных нейронов. VAE обучается путём приблизительного вероятностного распределения входных образцов. DAE отличается от AE лишь тем, что на вход подаются зашумлённые сигналы, а на выходе ожидаются данные без шума.

Автоэнкодеры и их модификации [70–73, 76–79] могут позволить очищать данные от шума при их предварительной подготовке. Однако обученные модели ИНС крайне неустойчивы к тем данным, которые не встречались в исходной выборке, то есть при наличии проблемы в работе сети, её источник сложно определить для локального дообучения. Из-за этого приходится дополнять обучающую выборку и переобучать нейросеть заново [80].

4. *Свёрточные нейронные сети* (англ. Convolutional neural networks – CNN) [81]. В основном CNN используются для обработки изображений, аудиоинформации, а иногда одномерный вариант CNN используется для анализа других видов данных (временных рядов). Типичное использование CNN – классификация изображений. Особенностью такой ИНС является наличие

свёрточных слоёв. Это слои, которые снижают размерность входящих сигналов, оставляя только существенные признаки (границы, яркость, цветовые характеристики). То есть слой входных сигналов не является полносвязным со следующим скрытым слоем свёртки, преобразованию подлежат только группы пикселов по правилам, определяемым в матрице свёртки. В конце к CNN часто присоединяют FFNN для интерпретации результатов анализа.

В качестве недостатков CNN можно отметить сложность её архитектуры и большое число варьируемых параметров. Поэтому заранее невозможно сказать, какие вычислительные мощности и какие настройки сети требуются при решении определённой задачи. Сложность архитектуры нейросети влечёт за собой низкую скорость обучения.

5. *Рекуррентные нейронные сети* (англ. Recurrent neural network – RNN) [82]. RNN подобны FFNN, но со смещением во времени: нейроны получают сигналы не только от предыдущего слоя, но и от самих себя в результате предыдущего прохода. Поэтому большое значение приобретает порядок подачи векторов обучающей выборки. RNN полезны во всех случаях, когда речь идёт об анализе данных, связанных с течением времени (звук, аудио, текст, временной ряд из значений какого-либо параметра). Проблема в RNN заключается во «взрывном/исчезающем градиенте», как и в глубоких FFNN с большим числом скрытых слоев. Это ведет к тому, что информация в сети со временем теряется (веса становятся слишком большими или слишком малыми).

Более сложным подтипом RNN является нейросеть с долгой краткосрочной памятью (англ. Long short term memory – LSTM) [83]. LSTM решает проблему градиента, используя фильтры (англ. gates) и блоки памяти (англ. memory cells). Задача фильтров – сохранять информацию, управляя её потоком, приостанавливая или возобновляя его. Входной фильтр (входной input gate) определяет количество информации, которое будет храниться в блоке памяти. Выходной фильтр (англ. output gate) контролирует информацию, которую получит следующий слой. Фильтр забывания (англ. forget gate) решает, когда следует справляться со проблемой «взрывного/исчезающего градиента».

Управляемый рекуррентный блок (англ. Gated recurrent units – GRU) [84]. GRU – разновидность LSTM, в которой система фильтров переработана. Вместо входного, выходного фильтров здесь используется фильтр обновления (англ. update gate), который определяет, сколько информации получить от предыдущего слоя и сколько информации сохранить от его последнего состояния. Фильтру забывания соответствует фильтр сброса (англ. reset gate). Выходного фильтра тут нет – на последующий слой отправляется полная информация о состоянии. В отличие от LSTM работает быстрее и проще, однако обеспечивает не столь большую функциональность.

В соответствии с результатами исследований, изложенными в работах [82–84], все эти нейросети перспективны для решения сложных классификационных задач, распознавания образов, задач нечёткого поиска и прогноза. Они могут использоваться при решении наших слабоформализованных задач, причём наиболее перспективны LSTM при решении задачи прогноза технологических параметров. Однако подготовка и обучение RNN очень сложны и вычислительно затратны.

6. *Самоорганизующаяся карта Кохонена* (англ. Selforganising (feature) Kohonen map – SOKM) [85]. Одной из самых важных особенностей карт Кохонена является тот факт, что ИНС обучается без учителя: нейрон, максимально отреагировавший на эталонный входной сигнал, становится центром кластера, плюс к этому веса соседних для него нейронов также корректируются, чтобы больше походить на входные сигналы. Имеется небольшой опыт использования карт Кохонена при кластеризации скважин-кандидатов для ГТМ [23]. Для ИНС такого типа требуется изначальный выбор количества кластеров(в нашем случае классов скважин), а также первый раз требуется обучающая выборка, подаваемая на вход сети в качестве ядер кластеров. При визуальной интерпретации полученных кластеров на выходе карты эксперт может установить количество их потенциальных классов в заранее неизвестных по содержанию данных и повторить процесс кластеризации скважин. К сожалению, полученные результаты имеют невысокую точность.

Так как каждая модель ИНС имеет преимущества и недостатки, то выбор конкретного типа ИНС и тем более модели ИНС в любом случае ограничит возможности метода ИНС для анализа геологических и технологических данных. В принципе, модель ИНС должна быть индивидуально подобрана при управлении фондом скважин определённого нефтяного или газоконденсатного месторождения. Большинство исследователей при выборе модели ИНС руководствуется правилом соответствия сложности задачи и функциональной мощности выбираемого инструмента (модели ИНС) для её решения: не следует использовать сложную модель ИНС при решении простой задачи. Основываясь на всех сделанных при анализе различных типов ИНС промежуточных выводах и на этом правиле, было принято решение использовать FFNN. Такой тип моделей ИНС является компромиссом между невысокой сложностью моделей ИНС, но достаточными их функциональными возможностями при решении поставленных слабоформализуемых задач. По сути, делается допущение, что возможностей таких нейросетей будет достаточно при решении всех трёх задач. Поэтому далее эффективность FFNN должна быть исследована для ответа на вопрос, может ли этот тип ИНС дать точность при решении задач прогноза и классификации, приемлемую с практической точки зрения.

Выбранный тип ИНС имеет большое число варьируемых характеристик архитектуры и гиперпараметров, что позволяет создавать и обучать различные модели ИНС этого типа. Такие модели описываются парой множеств:

$$<Architecture, Hyperparameters>, \qquad (3.1)$$

где Architecture – множество характеристик архитектуры ИНС (число входов и набор подаваемых на вход параметров, число скрытых слоёв, число нейронов в скрытых слоях, функция активации); Hyperparameters – множество гиперпараметров ИНС при обучении (скорость обучения, алгоритм обучения, размер минивыборки, число эпох).

Среди наиболее часто варьируемых характеристик – число скрытых слоёв FFNN, число значений геологических и технологических параметров во входном слое ИНС и функция активации нейронов. Рациональный выбор характеристик и параметров в (3.1) должен обеспечивать формирование модели (моделей) ИНС и её (их) обучение с целью решения с требуемой точностью той или иной

слабоформализуемой задачи. Число сочетаний различных параметров и характеристик модели ИНС может быть весьма велико, поэтому очень важно ограничить задачу поиска только теми из них и, соответственно, теми моделями ИНС, которые с высокой вероятностью дают гарантированный результат при решении слабоформализуемых задач.

3.1.4. Подходы и алгоритмы обучения ИНС

В настоящее время при задании гиперпараметров ИНС существует ряд эвристических рекомендаций, которые позволяют относительно быстро получить и обучить лучшие модели ИНС [86, 87]. Воспользуемся этими рекомендациями при формировании и обучении моделей ИНС выбранного типа FFNN. Среди таких рекомендаций основными являются следующие.

- Инициализацию значений для весов связей рекомендуется производить малыми случайными значениями, а не одинаковыми значениями или нулями, что вызывает асимметрию реагирования нейронов, за счёт чего веса обновляются по-разному в ходе обучения [88]. Также при инициализации необходимо учитывать число входов и выходов, чтобы нормализовать дисперсию для всех нейронов ИНС. Это создаёт во всех связях примерно одинаковое распределение и повышает скорость сходимости. Примером такой инициализации может быть задание значений для весов, рассчитываемых по формулам [89]:

$$\sqrt{\frac{6}{N_{in}+N_{out}}} \text{ или } \sqrt{\frac{2}{N_{in}+N_{out}}},$$

где N_{in} и N_{out} *– числа входных и выходных связей для нейрона.*

- Скорость обучения обычно выбирается равной *0,1*. Но если на валидационном наборе не наблюдается прогресса в обучении ИНС, то скорость уменьшается в *2* или *5* раз. При слишком большой скорости ИНС не обучается в принципе. При большой скорости велика вероятность оказаться в локальном минимуме. При слишком малой скорости потребуется значительное количество эпох и времени на обучение [90]. Поэтому важно применять различные методики, позволяющие изменять значение скорости в процессе обучения сети.

- Следует использовать функцию активации нейронов. Это привносит в модель ИНС компонент нелинейности.
- Число нейронов в скрытых слоях рекомендуется выбирать с использованием различных эмпирических правил [91], например, по формуле:

$$N_h = N_{in} \times 2 + 2, \tag{3.2}$$

где N_h *– число нейронов в скрытых слоях;* N_{in} *– число входных нейронов.*

- Для предотвращения переобучения ИНС используются методы регуляризации. Первый из них – ограничение максимального по модулю значения весовых коэффициентов значением 3 или 4. Метод дропаут (англ. dropout) [92] выделяет подсеть нейронов, в которой во время обучения случайным образом задействуются нейроны (обычно 50 % вероятность).
- Для расчёта ошибки между эталонными и полученными значениями весовых коэффициентов используется функции потерь – особая метрика, которая зависит только от значений на выходе сети и показывает, насколько в целом ИНС хорошо работает [93]. Наиболее известными функциями потерь являются: среднеквадратичная погрешность, кросс-энтропия, экспоненциальная, расстояние Кульбака – Лейблера.
- Если число обучающих примеров неравномерно распределено по классам, то есть некоторые классы имеют много обучающих представлений, а другие – значительно меньше, то необходимо произвести аугментацию недостающих примеров [94]. Если это сделать по каким-либо причинам невозможно, то одним из вариантов является использование взвешенной функции потерь. Также такую проблему можно решить следующим способом: сначала обучить модель на данных, обучающих примеров которых больше, а затем на данных с малым количеством примеров [95].
- Существует довольно большое число алгоритмов обучения. Целью любого из них является поиск таких весовых коэффициентов связей между нейронами, чтобы входные сигналы соответствовали определённым выходным сигналам. То есть алгоритм должен обеспечить нахождение значения определённой метрики (ошибки или погрешности) в глобальном минимуме.

Известны два подхода к обучению ИНС: обучение без учителя и обучение с учителем [96]. Второй подход более перспективный и подразумевает использование заранее сформированного множества обучающих примеров, содержащих соответствие входных и выходных сигналов.

Рассмотрим подробнее наиболее популярный из алгоритмов обучения с учителем – алгоритм *обратного распространения ошибки*, в котором в качестве метрики для минимизации используется отклонение (среднеквадратичная погрешность) [97] реальных значений выходных сигналов ИНС от эталонных:

$$E(w) = \frac{1}{2}\sum_{i,k}\left(f_{i,k} - y_{i,k}^{(T)}\right)^2, \tag{3.3}$$

где $f_{i,k}$ *– значение выходного сигнала на нейроне* k *при подаче обучающего примера* i *из набора обучающих данных;* $y_{i,k}^{(T)}$ *– требуемое значение сигнала на нейроне* k *при подаче обучающего примера* i *из набора данных для обучения.*

При этом для коррекции весов используется метод градиентного спуска, который позволяет изменить вес связи в соответствии с её вкладом в ошибку [98, 99]:

$$\Delta w_{ij}^{(q)} = -\eta\delta_i x_i, \tag{3.4}$$

где $\Delta w_{ij}^{(q)}$ *– величина изменения веса связи, соединяющей нейрон* i *из слоя* q*–1 с нейроном* j *из слоя* q*;* η *– коэффициент скорости обучения в диапазоне [0, 1];* δ_i *– значение ошибки нейрона* j *в слое* q*;* x_i *– значение выходного сигнала* i *для нейрона* j *в слое* q*.*

Для нейрона выходного слоя и нейрона скрытого слоя вычисление значения ошибки происходит следующим образом [97]:

$$\delta_i = \left(f_{i,k}(S)\right)'\left(f_{i,k} - y_{i,k}\right) \tag{3.5}$$

$$\delta_i^{(q)} = \left(f_i^{(q)}(S)\right)'\sum_j w_{ij}\,\delta_j^{(q+1)} \tag{3.6}$$

При этом из выражений (3.5) и (3.6) видно, что алгоритм требует применения дифференцируемых функций активации. Таким образом, для работы алгоритма на вход ИНС предъявляются значения из обучающей выборки, затем вычисляются выходные сигналы ИНС. После этого происходит подсчёт ошибок в выходном слое и скрытых слоях. Веса однократно корректируются. Условием

остановки алгоритма может быть либо достижение определённого значения ошибки, либо прохождение заданного числа эпох. Существует несколько модификаций алгоритма обратного распространения ошибки [100], которые позволяют увеличить скорость сходимости, избежать застревания в локальных минимумах, учесть ландшафт целевой функции и подобрать оптимальную скорость обучения в самом процессе обучения, при этом оставаясь вычислительно не накладными:

1. Первая из них – стохастический градиентный спуск (англ. Stochastic Gradient Descend – SGD) [101] отличается тем, что градиент оптимизационной функции на каждом шаге считается от одного случайно выбранного элемента, а не как сумма всех градиентов от всех элементов выборки.

2. Следующая модификация – адаптивный градиент (англ. AdaGrad). Данный алгоритм позволяет учесть редкие, но при этом информативные, признаки в обучающем наборе. Идея состоит в том, чтобы уменьшить обновление элементов, которые часто обновляются в процессе обучения. Например, для этого каждая связь хранит сумму квадратов его обновлений G_{t+1} [102]:

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{G_{t+1} + e}} g_t \tag{3.7}$$

$$g_t = \nabla_w E(w_t)$$

$$\Delta w = -\eta g_t$$

$$G_{t+1} = G_t + g_t^2,$$

где g_t *– направление уклона градиента;* $E(w_t)$ *– целевая функция потерь;* w *– параметр сети (вес связи);* η *– коэффициент скорости обучения в диапазоне [0, 1];* e *– сглаживающий параметр [10^{-8}; 10^{-6}], чтобы избежать деления на 0;* w_{t+1} *– новое значение веса нейрона.*

Это позволяет не подбирать точно скорость обучения, достаточно изначально задать её большой, она, в свою очередь, в ходе обучения будет постепенно уменьшаться. Однако скорость алгоритма может упасть настолько, что обучение прекратиться.

3. Алгоритм AdaDelta призван избежать недостатка, который имеет AdaGrad. Идея его такая же, но дополнительно введён стабилизирующий член [102]:

$$\Delta w = -\frac{RMS[\Delta w]_{t-1}}{RMS[g]_t} g_t \tag{3.8}$$

$$RMS[g]_t = \sqrt{G[g^2]_t + e}$$

$$w_{t+1} = w_t - \frac{RMS[\Delta w]_{t-1}}{RMS[g]_t} g_t$$

$$G[\Delta w^2]_t = \gamma G[\Delta w^2]_t - (1-\gamma)\Delta w_t^2$$

$$RMS[\Delta w]_{t-1} = \sqrt{G[\Delta w^2]_t + e},$$

где $RMS[g]_t$ *– корень из среднего квадратов градиентов;* $RMS[\Delta w]_{t-1}$ *– стабилизирующий член;* γ *– коэффициент сохранения [0; 1], обычно выбирают 0,9;* w_{t+1} *– новое значение веса нейрона.*

4. Алгоритм Adaptive Moment Estimation (Adam) основан на сочетании идей накопления движения и более слабого обновления весов для типичных признаков [102]:

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{\hat{u}_t + e}} \hat{m}_t \tag{3.9}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}$$

$$\hat{u}_t = \frac{u_t}{1-\beta_2^t}$$

$$u_t = \beta_1 u_{t-1} + (1-\beta_1) g_t^2$$

$$m_t = \beta_2 u_{t-1} + (1-\beta_2) g_t,$$

где u_t *– нецентрированная дисперсия;* β_2 *– коэффициент сохранения, 0,999 (авторская рекомендация);* β_1 *– коэффициент сохранения, 0,9 (авторская рекомендация);* e *– сглаживающий параметр,* 10^{-8} *(авторская рекомендация).*

Во многих случаях обучения ИНС алгоритм Adam показывает лучшие результаты среди других алгоритмов при минимальном подгоне параметров.

Существует большое число функций активации, каждая из них имеет свои достоинства и недостатки. Кратко рассмотрим основные из них [103].

1. Ступенчатая функция активации позволяет выходному значению принять значение «1» (нейрон активирован), если её аргумент превышает определённое значение (например, выше нуля), в ином случае выходное значение становится равным «0» (нейрон не активирован). Модели с такой функцией сложнее обучать:

функция не дифференцируема, что делает невозможным применение алгоритма обратного распространения ошибки при обучении.

2. Линейная функция – функция, пропорциональная входу. Она позволяет получать промежуточные значения, а не только бинарный ответ как в случае ступенчатой функции. Так как производная данной функции постоянна, то спуск по градиенту при обучении будет постоянным, что в большинстве случаев является недостатком. Другой отрицательной чертой этой функции является тот факт, что несколько последовательных линейных слоёв теряют смысл, так как комбинация линейных слоёв остаётся по-прежнему линейной функцией. Это ограничивает сложность модели ИНС и, соответственно, сложность решаемых задач.

3. Сигмоидальная функция (англ. Sigmoid) вычисляется по формуле [103]:

$$Sigmoid(x) = \frac{1}{1 + e^{x}} \tag{3.10}$$

Она нелинейна по своей природе, и комбинация сигмоид также является нелинейной функцией, что позволяется использовать многослойные модели ИНС. Для сигмоиды характерна небинарность и гладкий градиент, что позволяет её использовать в методах градиентного спуска. В отличие от линейной функции, диапазон выходных значений лежит в диапазоне [0, 1], а не [-inf, +inf], что не приводит к ошибкам в случае больших значений активации. Однако данная функция подвержена насыщению (приводит к затуханию градиентов) и не центрирована относительно нуля, что затрудняет обучение и приводит к зигзагообразному обновлению весов.

4. Гиперболический тангенс (англ. hyperbolic tangent – TanH) – это модификация сигмоиды, вычисляется в соответствии с выражением:

$$TanH(x) = 2 \times Sigmoid(2x) - 1 \tag{3.11}$$

Он тоже подвержен насыщению, но центрирован относительно нуля (функция активации не перегрузится от больших значений), поэтому всегда из этих двух вариантов в задачах классификации предпочтительнее использовать его.

5. Выпрямитель (англ. rectified linear unit – ReLU) – линейный порог в нуле, менее ресурсоёмкий, не подвержен насыщению, скорость сходимости примерно в 6 раз выше, чем у предыдущих двух вариантов, так как часть нейронов не активируется, что особо важно при обучении глубоких ИНС. Однако нейроны с этой функцией активации склонны выходить из строя в ходе обучения, это происходит на горизонтальной части ReLU (градиент равен «0»), из-за чего на этой части веса нейрона не корректируются (проблема умирающего ReLU). Для решения этой проблемы используются такие модификации, как LeakyReLU, ParametricReLU, RandomizedReLU. Они отличаются от ReLU тем, что на отрицательной оси функция имеет угловой коэффициент. В целом, ReLU и его модификации работают как хороший аппроксиматор, их также будем использовать в разрабатываемых моделях ИНС.

В итоге следует сказать, что результаты анализа всех указанных рекомендаций при формировании моделей ИНС и приёмы задания значений гиперпараметров при их обучении далее будут использованы при решении слабоформализуемых задач с помощью ИНС типа FFNN.

3.1.5. Выбор библиотеки для формирования и обучения ИНС

Проблема реализации ИНС и их обучения решается использованием готовых библиотек (платформ), которые максимально инкапсулировано и оптимизированно позволяют взаимодействовать напрямую с GPU, что в десятки раз ускоряет процесс обучения ИНС по сравнению с использованием процессора персонального компьютера, так как задействует техническое обеспечение, специально предназначенное для обработки параллельных структур.

Различаются библиотеки, в основном, синтаксическими особенностями, свойственными используемому языку программирования, и наименованиями функций в библиотеке. Стоит сказать, что стандартом написания кода практически для всех библиотек является язык Python, но иногда библиотеки включают в себя API с доступом для других языков: Keras (Python), Google TensorFlow (Python), Caffe (Python, C++), FaceBook PYTORCH (Python),

Accord.NET, Microsoft CNTK (Python, C++, BrainScript, позже .NET), cuDNN (команды GPU), DeepLearning4J (Java) и другие.

Как уже было сказано, многие библиотеки примерно в равной степени удовлетворяют всем критериям, необходимым для глубокого обучения ИНС: *скорость* (производительность при обучении ИНС), *точность* (наличие максимально выверенных моделей глубокого обучения без каких-либо ошибок), *структура API* (наличие доступа к функциональности библиотеки сторонними системами), *масштабируемость* (задействование нескольких параллельных графических процессоров), *расширяемость* (возможности использования пользовательских слоёв в ИНС с уникальными характеристиками) и наличие *встроенных модулей считывания* (оптимизация передачи данных из ОЗУ в GPU и средства чтения данных из различных форматов) [104].

Исследования [104, 105] показывают небольшой выигрыш Microsoft CNTK (англ. Cognitive Toolkit – CNTK) по сравнению с остальными библиотеками почти по всем указанным критериям. Библиотека находится в свободном доступе с открытым исходным кодом [106], достаточно просто устанавливается и используется [107].

Таким образом, с учётом опыта разработки программ в среде Visual Studio 2019 на языке C# 7.0 и платформе .NET 4.7 нами был выбран соответствующий API библиотеки CNTK 2.7.

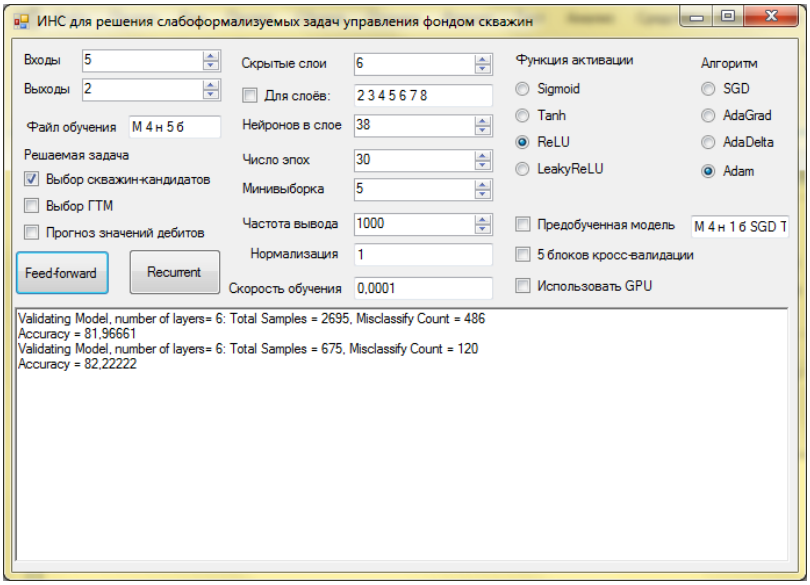

Рисунок 3.1 – Интерфейс пользователя для построения моделей ИНС

На рисунке 3.1 представлен интерфейс пользователя ПО, созданного нами на основе этой библиотеки и позволяющего в интерактивном режиме задавать параметры построения модели ИНС для последующего использования в веб-сервисе.

ПО, созданное на основе этой библиотеки, позволяет выбрать заранее приготовленные файлы для обучения и тестирования ИНС, задать число входов и выходов для нейронной сети, установить число скрытых слоёв и число нейронов в слое. Также можно настроить число эпох, задать объём минивыборки, скорость обучения, оптимизационный алгоритм, функцию активации. При необходимости указывается целая цепочка тестов с различным числом слоёв, функциями активации и алгоритмами обучения.

Данный интерфейс позволяет строить модель ИНС для решения сформулированных выше трёх слабоформализуемых задач и сохранять её в виде файла, доступном для дальнейшего использования при дообучении ИНС, а также валидации и тестировании её на соответствующей выборке.

3.2. Модели, методы и алгоритмы для решения задачи прогноза технологических параметров скважин

Целью планирования работы производства нефтегазодобывающего предприятия является определение таких технологических режимов работы фонда скважин и технологического оборудования, которые позволят достичь заданного объёма выпуска продукции (кондиционных нефти, газа и газового конденсата). При этом необходимо планирование технологических режимов работы каждой добывающей скважины фонда и технологических режимов каждой установки подготовки продукции из добываемого УВС, а также планирование режимов транспортировки этой продукции до магистральных нефте- и газопроводов. При планировании технологических режимов скважин фонда необходимо решать слабоформализуемые задачи прогноза ряда технологических параметров скважин. Наиболее востребованными являются прогнозы значений дебитов (суточных уровней добычи) каждого из четырех компонентов добываемого УВС (дебит

нефти $Q_н$, дебит газа $Q_г$, дебит газового конденсата (жидкости) $Q_ж$ и дебит воды $Q_в$) каждой скважины фонда. Это означает, что необходимо решать *Задачу 1*. Однако существующие методы прогноза дебитов в лучшем случае позволяют выполнить экстраполяцию данных путём линейной регрессии [10]. Разумеется, столь простые методы не способны учесть множество аспектов, заложенных в данных и взаимовлияние их друг на друга в рамках анализируемых временных рядов для дебитов [21]. Соответственно, точность прогноза дебитов не всегда удовлетворяет специалистов нефтегазодобывающих предприятий. Это ещё раз указывает на актуальность разработки новых методов и алгоритмов, позволяющих существенно повысить точность прогноза значений параметров скважин.

Точное предсказание дебитов скважин является необходимой частью решения задачи поддержки принятия решений при управлении фондом скважин любого нефтегазодобывающего предприятия. Важно прогнозировать (планировать) дебиты не просто на конкретный день, а на определённый промежуток времени (например, равный месяцу), чтобы понимать, сколько в среднем УВС будет добываться из скважины в сутки. В этом случае говорят о планировании оперативного технологического режима работы скважины. Знание такого технологического режима скважины позволяет рассчитать экономический эффект от добычи УВС и понять рентабельность эксплуатации конкретной скважины, построить перспективные планы поставок продукции и учесть ограничения, наложенные органами надзора за пользованием недрами.

3.2.1. Алгоритмы и методика подготовки данных для их последующего анализа

Решение *Задачи 1* осуществлялось с использованием архивов данных в виде БД по двум эксплуатируемым месторождениям: нефтегазоконденсатному месторождению в Томской области (далее, для краткости, – Месторождение 1) и газоконденсатному месторождению в Томской области (далее – Месторождение 2). Фонд скважин Месторождения 1 имеет 142 добывающих скважины, на которых было проведено 436 различных ГТМ за 6 лет. Фонд скважин Месторождения 2 имеет 60 добывающих скважин, на которых было проведено

313 различных ГТМ за 5 лет. В БД за несколько лет эксплуатации месторождений накоплен значительный объём данных по дебитам. Для каждого из четырех дебитов всех скважин в анализируемой БД имеются прогнозное и фактическое значения для всего периода мониторинга фонда скважин. На предприятии использовалась методика прогноза (планирования) значений дебитов, основанная на методе экстраполяции по скользящей средней [108]. Этот метод производит точечную оценку, выравнивая временной (динамический) ряд по средним характеристикам ряда в прошлом. Как и любой статистический метод, он даёт высокий по точности результат только при краткосрочном планировании.

Формирование обучающей и тестовой выборок для выбранного типа ИНС предлагается вести с использованием разработанного адаптивного алгоритма предварительной подготовки (обработки) исходных данных. Укрупнённая схема алгоритма приведена на рисунке 3.2.

Рисунок 3.2 – Укрупнённая схема адаптивного алгоритма предварительной подготовки данных при формировании обучающей и тестовой выборок для решения *Задачи 1*

Рассмотрим основные блоки схемы алгоритма более подробно.

- *Ввод данных из БД* (блок 1). В этом блоке с помощью SQL-запросов к БД выбранного месторождения происходит извлечение (сбор) данных. Термин

«сбор данных» чаще употребляется, когда данные по месторождению извлекаются из БД нескольких унаследованных ИС. При этом собираются следующие сведения:

- паспортные данные о кусте скважин и скважине (название или номер, дата ввода в эксплуатацию);
- фактические (измеренные) значения дебитов $Q_н$, $Q_г$, $Q_ж$ и $Q_в$ (название, дата получения, значение);
- плановые (прогнозные) значения дебитов (название, дата планирования, значение), рассчитанные по методике прогнозирования дебитов предприятия;
- данные об остановках скважин, включая остановки на ГТМ и для исследования скважин (номер скважины, время остановки и время запуска).

Собранные (извлечённые из БД) данные вводятся в специально выделенную буферную память.

- *Отбраковка данных и формирование таблиц* (блок 2). В этом блоке ведется выбраковка некоторых собранных данных. Дело в том, что прогноз дебитов, выполненный геологической службой предприятия, судя по имеющимся данным, осуществлялся от минус 8 до 229 дней до измерения фактического значения дебита. То есть в БД имеются явно ошибочные значения параметров, по крайней мере, при планировании значений дебитов, так как нельзя планировать значение дебита уже после получения его фактического значения. Также вызывают сомнения прогнозы на большие временные промежутки (на 2 и более месяцев). Все такие данные подлежат отбраковке и в наших экспериментах не используются.

Часть данных по дебитам также выбраковывалась из-за принятой нами методики проведения исследований моделей ИНС и необходимости сравнения наших результатов исследований с результатами прогноза дебитов, полученными на предприятии. Так по нашей методике учитывалось, что все фактические (измеренные) значения дебитов из БД на дату планирования и ранее подаются на входы ИНС (признак, англ. feature). Фактическое значение для каждого дебита из

БД через месяц после этой даты является эталоном, то есть такие значения служат эталоном выходным значениям ИНС (метка, англ. label) при её обучении и тестировании. Данные по дебитам, которые находятся между датой, в которую происходит планирование, и датой получения фактического значения дебита-эталона, не используются, несмотря на их наличие в БД. Это позволит находиться в тех же самых условиях, что и находилась геологическая служба предприятия, планировавшая технологический режим работы скважины, и позволит корректно осуществить сравнение полученных при исследовании результатов прогноза с имеющимися в БД значениями прогнозных (плановых) дебитов. Более того, при проведении наших экспериментов было решено брать значения дебитов от даты планирования в течение трёх – пяти недель до получения фактического значения. Это практически важно при формировании технологических режимов скважин.

После отбраковки все оставшиеся в буфере данные сводятся в две таблицы базы данных ИИС. Первая из них содержит название (номер) куста, номер скважины, дату ввода её в эксплуатацию, название дебита, дату планирования, плановое (прогнозное) значение дебита, дату получения фактического значения этого дебита и его фактическое значение. Данные сортируются по названию куста, затем номеру скважины, потом по дате получения фактического значения дебита.

Вторая таблица включает данные: название куста, номер скважины, дату ввода её в эксплуатацию, название мероприятия на остановку скважины или остановку по иным причинам, дату начала и дату окончания мероприятия (остановки). Эти данные сортируются по нескольким столбцам: сначала по кусту, затем по скважине, потом по дате начала проведения мероприятия.

- *Анализ данных на корректность* (блок 3). Некоторые полученные данные из БД предприятия могут быть некорректными, поэтому в этом блоке алгоритма ведется их анализ на корректность. Так, встречаются ошибки ручного ввода, которые характеризуются наличием внезапного выброса. Они проявляются в виде неверной записи порядка числа. Возможны выбросы значений (аномально большие отсчёты) дебитов из-за сбоев измерительной аппаратуры. Все выбросы

выявляются на основании сравнения в соответствии с положениями концепции «Big Data» с соседними во временном ряду значениями параметров с учётом того, что они резко отличаются (иногда в десятки раз) от соседних, причём не закономерно. Такого рода значения из последовательности удаляются и доступны в дальнейшем для восстановления интерполяционными методами. В случае остановки скважины при одновременном проведении на ней нескольких видов мероприятий время окончания зачастую бывает только одного из этих мероприятий, в то время как их начало было индивидуальным для каждого мероприятия. Подобные мероприятия следует объединять в одно укрупнённое событие, имеющее единое время начала и окончания. Возможны дубликаты тех или иных событий (получение фактического значения параметра или проведения мероприятия на скважине), их необходимо удалять из общего списка. Дубликаты не всегда легко автоматически распознать в данных. Например, для такого ГТМ, как ГРП, можно встретить следующие варианты: «ГРП», «Гидроразрыв пласта», «гидро-разрыв пласта», «гидро разрыв пл.» и прочее. Поэтому приходилось контролировать результаты работы этого блока после их получения и дополнять в алгоритм новые правила выявления и отсева дубликатов.

Наконец, отсутствующие значения (NULL) не всегда являются признаком некорректности, так как служат впоследствии при выполнении алгоритма признаком для применения методов интерполяции. Однако для части данных значение «NULL» является ошибкой: дата начала и окончания мероприятия, дата прогноза параметра и другие. Для формата данных «дата-время» иногда встречаются значения вида «01.01.1900:00.00.00», то есть значение по умолчанию. Дата-время получения того или иного параметра является ключевым фактором, без которого данные восстановить невозможно, поэтому записи с таким некорректным содержимым удаляются из общего перечня записей.

- *Адаптивное восстановление пропусков в данных* осуществляется с использованием блоков 4, 5, 6. Во время проведения различных мероприятий по остановке скважин данные не измеряются, что логично, ведь добыча не идёт, и значения дебитов отсутствуют. Основой для восстановления пропущенных

данных или отсутствующих данных на месте обнаруженных и удалённых выбросов служат значения дебитов, полученные в периоды работы скважин: в промежутки между различными остановками скважин, в период времени с момента начала эксплуатации скважины до первой её остановки, а также в интервал времени после последней её остановки. Если на скважине ГТМ не проводились, то берётся весь массив значений фактических дебитов для анализа без каких-либо ограничений. Блок 4 определяет характер пропусков в данных, при этом в нём анализируются различные возможные ситуации (единичный пропуск, совокупность пропущенных значений и т. п.). Например, наиболее нестандартная из ситуаций, когда между двумя остановками скважины проведено только одно измерение дебита или оно вообще отсутствует. В таких условиях провести интерполяцию нельзя, и прогноз на этом интервале становится невозможным. Важно определить число пропусков N в совокупности и местоположение его в ряду анализируемых значений дебитов относительно других возможных единичных пропусков и (или) блоков пропусков. В решающих правилах блоков 4 и 5 учитываются все возможные ситуации с пропусками и наличием измеренных значений дебитов. Например, если точность того или иного метода интерполяции лучше, чем у других методов, однако для его применения не достаточно точек в исходных данных, то применяется другой метод интерполяции с характеристиками по точности, ближайшими к приоритетному методу. Это позволяет адаптировать восстановление данных к особенностям анализируемого ряда значений дебитов и, в том числе в зависимости от числа N пропусков в блоке, перенаправить вычислительный процесс на соответствующий субблок блока 6. Учитывая, что восстановление значений дебитов необходимо осуществлять с высокой точностью, в соответствии с рекомендациями концепции «Big Data» в блоке 6 необходимо использовать различные методы интерполяции и, соответственно, различные субблоки. Последние должны обеспечивать заданную точность вычислений при различных размерах N блоков пропусков. Ниже будут исследованы и показаны возможности каждого метода интерполяции в зависимости от числа пропусков в исходных данных.

- *Формирование обучающих примеров (блок 7).* Текущий или капитальный ремонт скважин, остановка их по другим причинам или на проведение ГТМ вносят непредсказуемость в изменение параметров скважин. Это должно быть учтено при составлении обучающих примеров, чтобы в вектор признаков на входах ИНС по каждой скважине не попадали значения дебитов одновременно до и после проведения тех или иных ГТМ или остановок скважин. То есть планирование технологического режима скважины должно осуществляться исключительно с учётом периодов её эксплуатации.

В качестве входных значений для ИНС использовались только фактические значения прогнозируемого дебита. Не ясно, какое оптимальное число значений дебита из анализируемого временного ряда нужно подавать на входы модели ИНС, чтобы получить наиболее точное прогнозируемое значение дебита. Это предстоит определить экспериментально. Более подробно процесс выбора значений дебитов описан ниже в предложенной методике формирования обучающих примеров.

- *Распределение обучающих примеров по выборкам (блок 8).* Все обучающие примеры необходимо разделить на два блока – обучающая выборка и тестовая выборка. В текущей реализации алгоритма предлагается использовать соотношение 4:1, то есть тестовая выборка составляет 20 % от общего числа примеров, а обучающая – 80 %. Таким образом, при использовании метода кросс-валидации по методу К блоков должно быть выбрано К = 5. При формировании блока в тестовую выборку попадает каждый пятый пример их общего объёма выборки, а для построения очередного блока происходит то же самое, но со смещением на единицу. Остальные примеры попадают в обучающую выборку для соответствующего блока.

Были исследованы и затем использованы в адаптивном алгоритме следующие три наиболее часто применяемых на практике метода интерполяции.

1. Линейная интерполяция [109]. Это наиболее простой метод интерполяции, в котором неизвестные значения параметра вычисляются на основе соседних известных значений по следующей формуле:

$$f(x) = f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}(x - x_i), \quad (3.12)$$

где $f(x)$ *– искомое значение параметра;* x_i *– момент времени в левой границе рассматриваемого интервала;* $f(x_i)$ *– значение параметра в* x_i *момент времени;* x_{i+1} *– момент времени в правой границе рассматриваемого интервала;* $f(x_{i+1})$ *– значение параметра в* x_{i+1} *момент времени.*

2. Локальная кусочно-квадратичная интерполяция [110]. Особенностью этого метода является то, что для интерполяции должно использоваться три опорных точки с известными значениями. Например, могут использоваться два значения дебита до пропущенных значений и одно – после них.

$$F(x) = a_0 + a_1 x + a_2 x^2, \quad (3.13)$$

где $x_{i-1} \le x \le x_{i+1}$; $a_2 = \frac{f(x_{i+1}) - f(x_{i-1})}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)} - \frac{f(x_i) - f(x_{i-1})}{(x_i - x_{i-1})(x_{i+1} - x_i)}$; $a_1 = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} - a_2(x_i + x_{i-1})$; $a_0 = f(x_{i-1}) - a_1 x_{i-1} - a_{2x_{i-1}} x_{i-1}^2$

3. Одномерный кубический сплайн [111]. Сплайн производит интерполяцию максимально плавно, в получаемой функции отсутствуют резкие изменения и острые углы. В нашей задаче интерполяции это соответствует процессу постепенного убывания дебита с течением времени.

$$g_i(x) = a_i + b_i(x - x_i) + \frac{c_i}{2}(x - x_i)^2 + \frac{d_i}{2}(x - x_i)^3, \quad (3.14)$$

где $a_i = g_i(x_i)$; $b_i = g_i'(x_i)$; $c_i = g_i''(x_i)$; $d_i = g_i'''(x_i)$; $[x_i; x_{i+1}]$ *– подотрезки кубического сплайна;* $g_i'(x_i) = g_{i+1}'(x_i)$; $g_i''(x_i) = g_{i+1}''(x_i)$ *– граничные условия, периодический сплайн;* $g_i''(x_1) = 0$; $g_i''(x_n) = 0$ *– граничные условия, «естественный сплайн»;* $g_i'''(x_1) = g_i'''(x_n) = 0$ *– граничные условия, «непрерывность второй производной»;*

Следует помнить, что качество получаемой обучающей и тестовой выборок для ИНС будет зависеть от точности исходных данных и корректности выбора метода корректировки этих данных и точности метода восстановления пропусков в них. В связи этим, главным критерием оценки полученных при интерполяции значений дебитов является их *точность*. Точность методов определялась с помощью имитационного моделирования на участках с достоверными данными (без пропусков и ошибок), где искусственно изымалось определённое число значений дебитов, а затем происходило сравнение полученных результатов

интерполяции с изъятыми значениями. При большом объёме данных сложные методы интерполяции могут существенно замедлить процесс обработки временных рядов для дебитов, при этом обеспечив не столь значимый прирост по точности. Поэтому важно обращать внимание и на *быстродействие* при выполнении метода. Быстродействие определяется скоростью выполнения каждого из методов интерполяции при неизменной вычислительной мощности устройства при проведении экспериментов.

В таблице 3.1 приведены результаты исследований возможностей выбранных выше методов интерполяции, полученные с использованием данных по дебитам для фонда скважин Месторождения 1. Точность вычислений и скорость интерполяции, полученные при использовании метода линейной интерполяции (базовый метод) для восстановления одного пропуска (N=1) в данных, приняты за единицу (1,000). Отметим, что относительная погрешность [112] при вычислениях с помощью базового метода не превышает 4,9 %. Полученные значения точности вычислений с помощью других двух методов представляют собой доли от результатов базового метода при N =1.

Таблица 3.1 – Зависимость точности и скорости интерполяции при различном числе пропусков

Число пропусков N	**Линейная интерполяция**	**Кусочно-квадратичная интерполяция**	**Одномерный кубический сплайн**
	Точность / Скорость интерполяции		
1	**1,000 / 1,000**	1,002 / 0,995	1,002 / 0,830
2	0,996 / 0,999	**1,005 / 0,994**	0,991 / 0,821
3	0,972 / 0,999	**1,011 / 0,994**	0,988 / 0,820
4	0,987 / 0,998	**1,030 / 0,993**	0,992 / 0,812
5	0,969 / 0,998	**1,027 / 0,993**	1,012 / 0,806
6	0,961 / 0,998	1,021 / 0,993	**1,031 / 0,793**
7	0,955 / 0,997	1,025 / 0,991	**1,045 / 0,791**
8	0,941 / 0,996	1,031 / 0,985	**1,042 / 0,785**
9	0,946 / 0,996	1,020 / 0,985	**1,048 / 0,741**
10	0,953 / 0,995	0,989 / 0,981	**1,013 / 0,740**
20	0,914 / 0,988	0,978 / 0,939	**1,005 / 0,718**

Исходя из этих результатов, рекомендовано применять методы интерполяции в адаптивном алгоритме в зависимости от размера (числа пропусков N) блока пропусков следующим образом:

1. $\boldsymbol{N} = \mathbf{1}$ (одно недостающее значение дебита). Это наиболее простой случай, для которого *линейная интерполяция* показала лучшие значения как по точности, так и по быстродействию.

2. $\mathbf{2} \leq \boldsymbol{N} \leq \mathbf{5}$ (средний по размеру блок пропусков). При таком числе пропусков в блоке *локальная кусочно-квадратичная интерполяция* показала лучшую точность с небольшим отставанием по скорости обработки данных от линейной интерполяции.

3. В случаях большого числа пропусков в блоке, когда $\mathbf{N} > \mathbf{5}$, был выбран *одномерный кубический сплайн.* Он требует высоких вычислительных мощностей, но при этом даёт существенный прирост точности по сравнению с другими методами интерполяции для большого числа пропусков в блоке.

Именно поэтому блок 6 в укрупнённой схеме алгоритма на рисунке 3.2 реализован в виде трёх субблоков.

Рассмотрим в дополнение к описанию блока 7 адаптивного алгоритма (рисунок 3.2) предложенную нами методику формирования обучающих примеров. Суть её состоит в том, чтобы увеличивать при прогнозе число входных значений дебитов для ИНС за счёт расширения числа известных из БД фактических значений дебитов от 6 (6 дней назад от даты планирования) с постепенным добавлением таких значений за предыдущие дни до 15. Схематично процесс выбора данных изображён на рисунке 3.3, где показано движение скважины во времени. Видно, что скважина была два раза остановлена на разные промежутки времени. Жирными штрихами отмечены взятые в эксперимент 13 фактических значений дебита на скважине, получаемые ежедневно. Пусть дата прогноза значения дебита будет отмечена знаком «!» (метка для ИНС), и на эту дату в БД должно быть фактическое значение дебита. Дата, с которой планируется значение дебита, показана знаком «?». С даты планирования и по дату прогноза должен пройти один месяц (период планирования). Все фактические значения дебита (от 6 до 15 значений) до даты планирования становятся признаками, подаваемыми на входы ИНС (отмечены жирными штрихами).

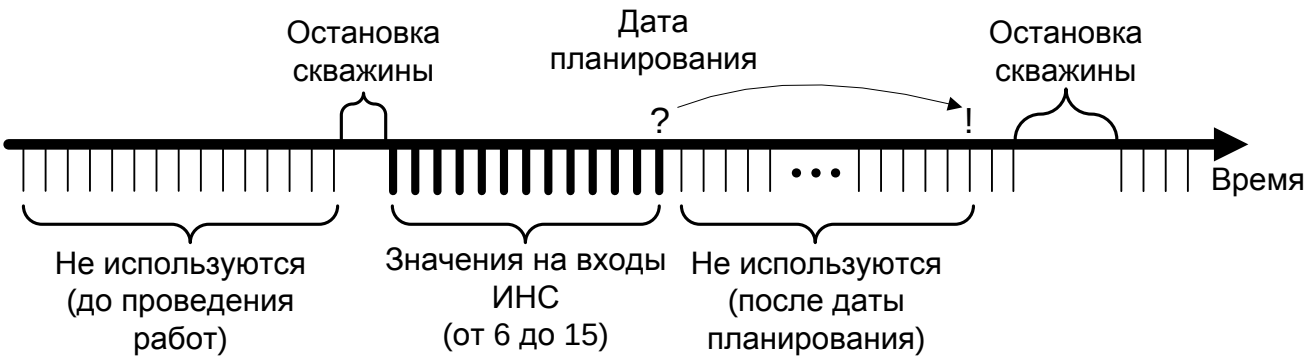

Рисунок 3.3 – Схема выбора значений измеряемых дебитов

Дате планирования может предшествовать остановка скважины для проведения каких-либо работ (например, ГТМ). Поэтому число признаков для ИНС, большее некоторого имеющегося числа, взять не получится. Например, из рисунка 3.3 следует, что выбрать 14 признаков не удастся. Это значит, что в общем случае обучающие выборки будут не сопоставимы по объёму, так как для малого числа признаков на входах ИНС (например, 6) будет всегда больше обучающих примеров, чем для большого числа признаков (например, 15).

Для реализации адаптивного алгоритма, начиная с извлечения данных о дебитах и остановках скважин из архивных БД и заканчивая распределением обучающих примеров по выборкам было создано специальное ПО. Это ПО позволяет провести всю необходимую предварительную обработку данных, выбрать вид дебита (дебит нефти, газа, жидкости (газового конденсата) или воды), число признаков на входе ИНС (от 6 до 15), указать число блоков для кросс-валидации данных методом K блоков [113], а также создать файл для последующего обучения или тестирования ИНС. Пользователь видит при загрузке данных информацию о характеристиках выборки в текстовом поле. При формировании файла указывается, сколько было задействовано скважин при построении выборки и общее число обучающих и тестовых примеров.

Для Месторождения 1 исходное число записей в БД предприятия составляло 17117, после выбраковки ошибочных данных и последующей корректировки данных записей осталось 13032. Аналогично и для Месторождения 2 исходное число записей 9433, а после отбраковки и

корректировки данных стало 5323 записи. Видно, что такие действия уменьшили объём данных, однако это позволило нам получить более точную оценку результатов при решении *Задачи 1* на месяц по существующей на предприятии методике.

Для оценки точности моделей ИНС и статистических методов при решении *Задачи 1* использовалась взвешенная средняя абсолютная погрешность (англ. – WAPE) [114]:

$$WAPE = \frac{\sum_{i=1}^{N} |\Phi_i - \Pi_i|}{\sum_{i=1}^{N} \Phi_i}, \tag{3.15}$$

где Φ_i *– фактическое (измеренное) значение дебита на дату прогноза (из БД предприятия);* Π_i *– прогнозное значение дебита, полученное для i-го обучающего либо тестового примера при использовании ИНС или других методов; N – число обучающих/тестовых примеров.*

Отметим, что формула (3.15) использовалась также при оценке погрешности прогноза по методике предприятия, но в качестве значений Π_i выбирались плановые значения дебитов из БД предприятия.

Число анализируемых записей и рассчитанные нами значения взвешенных средних абсолютных ошибок при прогнозе каждого вида дебита по каждому фонду месторождений по методике предприятия указаны в таблице 3.2. Стоит отметить, что для Месторождения 2 дебиты нефти, газа и жидкости отдельно не измеряются, поэтому в его БД есть только значения суммарного дебита УВС (включают дебиты $Q_н$, $Q_г$, $Q_ж$) и дебита воды.

Таблица 3.2 – Число анализируемых записей и оценки погрешности прогноза дебитов по методике предприятия

Дебит	**Обозначение**	**Единица измерения**	**Число записей для Месторождения 1**	**Число записей для Месторождения 2**	**WAPE для Месторождения 1, %**	**WAPE для Месторождения 2, %**
дебит нефти	$Q_н$	т/сут	3557	2958	27,347	8,098
дебит газа	$Q_г$	тыс. м^3/сут	3458		35,133	
дебит жидкости	$Q_ж$	м^3/сут	3459		26,734	
дебит воды	$Q_в$	т/сут	2558	2365	67,044	17,140

Для проведения экспериментов с помощью созданного ПО было построено 40 пар файлов (для обучения и тестирования ИНС) – по 10 пар на каждый вид дебита для Месторождения 1. Для Месторождения 2 построено 20 пар файлов.

Для Месторождения 1 объёмы пар «признаки – метка» для разного числа входов ИНС и для всех дебитов указаны в таблице 3.3.

Таблица 3.3 – Объёмы выборок для различных дебитов и разного числа входов ИНС, полученные для Месторождения 1

Дебит	**Число входов ИНС**									
	6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$Q_н$	1248	1248	1246	1246	1244	1244	1243	1243	1243	1235
$Q_г$	1199	1199	1197	1197	1195	1195	1194	1194	1194	1188
$Q_ж$	1199	1199	1197	1197	1195	1195	1194	1194	1194	1188
$Q_в$	880	877	875	875	873	873	872	872	872	868

3.2.2. Эксперименты по прогнозу значений дебитов скважин с помощью ИНС

Программная реализация моделей ИНС для проведения экспериментов выполнена на языке C# с использованием библиотеки Microsoft CNTK [106], выбранной из ряда наиболее известных библиотек и описанной в п. 3.1.5. С её помощью были автоматизированы операции по изменению архитектуры и гиперпараметров моделей ИНС, обучены и тестированы ИНС на вышеуказанных выборках. Обучение моделей ИНС проводилось с учётом рекомендаций, изложенных в п. 3.1.4.

Укажем пределы изменения характеристик и параметров моделей ИНС прямого распространения. Число скрытых слоёв моделей ИНС выбиралось равным 2, 3, 4, 5, 6, 7, 8, 10, 15. Исследовалась также эффективность применения различных функций активации (TanH (формула 3.11) и LeakyReLU), позволяющих получать связи между нейронами с отрицательными воздействиями. Использовался алгоритм оптимизации при обучении ИНС Adam, он основан на выражении (3.9). Каждый обучающий пример уникален и должен осуществлять воздействия на ИНС индивидуально, но скорость обучения тем больше, чем меньше объём мини-выборки, поэтому был выбран компромиссный вариант, равный 5. Число эпох выбиралось динамически в соответствии с достигаемой точностью обучения, если она несколько эпох существенно не менялась. Число нейронов в каждом скрытом слое выбиралось в соответствии с эвристическим правилом (формула 3.2).

На входы модели ИНС подаются фактические значения дебита за определённое число дней (от 6 до 15). Далее в модели идёт ряд скрытых слоёв – от 2 до 15. Отметим, что в ходе экспериментов было выявлено, что число слоёв, большее 8, не даёт прироста точности прогноза. Поэтому далее результаты показаны только для моделей ИНС с числом скрытых слоев от 2 до 8. В выходном слое ИНС находится один нейрон (прогнозируемое значение дебита). На рисунке 3.4 в качестве примера приведена модель ИНС с 3 скрытыми слоями, 12 фактическими значениями дебита на входах ИНС и 26 нейронами в скрытых слоях, используется функция активации LeakyReLU [115].

В экспериментах необходимо определить рациональные значения характеристик и параметров моделей ИНС.

Рисунок 3.4 – Пример архитектуры модели ИНС

В таблицах 3.4–3.7 приведены результаты экспериментов с различными моделями ИНС при решении на тестовых и обучающих выборках *Задачи 1* для нефти $Q_н$, газа $Q_г$, жидкости (газового конденсата) $Q_ж$ и воды $Q_в$ для Месторождения 1. Переменными являются как функции активации и число скрытых слоёв ИНС, так и число входов ИНС. Значение погрешности прогноза WAPE вычисляется как среднее по всем погрешностям, полученным по схеме кросс-валидации по 5 блокам, и по всем погрешностям для добывающих скважин в выборках. Жирным отмечается наименьшая погрешность, результаты округлены до тысячных [116].

Таблица 3.4 – Оценки погрешности прогноза $Q_н$ с помощью ИНС для Месторождения 1

Число входов ИНС	**Функция активации**	**Выборка**	**WAPE, %** **Для указанного числа скрытых слоёв**						
			2	3	4	5	6	7	8
6	TanH	обучающая	13,619	13,260	12,467	12,747	12,469	12,365	13,080
		тестовая	16,723	16,859	17,795	15,179	16,713	17,912	17,511
	LeakyReLU	обучающая	16,748	14,311	14,155	13,554	13,800	13,670	13,530
		тестовая	20,560	17,997	18,405	16,662	18,166	17,464	19,119
7	TanH	обучающая	14,853	13,459	12,796	13,084	12,556	12,250	12,790
		тестовая	18,217	17,742	17,375	17,636	15,437	17,127	18,592
	LeakyReLU	обучающая	14,845	13,721	14,321	13,222	13,700	12,877	13,009
		тестовая	18,565	17,961	17,582	17,519	17,141	15,794	17,220
8	TanH	обучающая	13,660	12,861	12,836	12,023	17,155	13,648	12,554
		тестовая	16,299	17,403	17,160	17,166	20,241	19,101	17,068
	LeakyReLU	обучающая	14,881	14,244	14,179	13,595	13,185	13,045	12,617
		тестовая	17,011	17,515	17,640	16,855	16,186	17,389	17,446
9	TanH	обучающая	13,640	12,970	13,019	10,540	12,479	11,893	16,571
		тестовая	16,569	17,399	17,375	18,858	17,697	15,827	17,707
	LeakyReLU	обучающая	14,206	14,810	13,619	12,940	13,360	12,927	12,116
		тестовая	17,058	18,310	16,738	19,012	15,875	15,484	16,886
10	TanH	обучающая	14,316	14,860	12,931	13,481	14,431	14,565	12,632
		тестовая	14,127	14,674	12,623	13,296	13,985	14,523	13,298
	LeakyReLU	обучающая	15,004	15,010	13,922	13,983	13,329	12,674	13,450
		тестовая	13,889	13,849	**12,477**	13,337	13,456	13,351	13,768
11	TanH	обучающая	14,839	13,356	14,942	14,753	13,769	13,575	13,674
		тестовая	13,250	12,808	14,595	14,189	13,658	16,746	15,482
	LeakyReLU	обучающая	15,354	14,644	13,798	12,708	13,548	13,311	17,283
		тестовая	13,733	13,477	13,694	12,698	13,236	13,453	14,005
12	TanH	обучающая	13,323	13,228	13,025	13,918	13,185	14,853	13,063
		тестовая	15,585	16,023	15,165	17,042	16,985	16,431	17,145
	LeakyReLU	обучающая	15,344	14,755	13,471	13,067	13,394	11,781	12,100
		тестовая	15,875	15,676	15,729	15,634	16,474	15,332	15,910
13	TanH	обучающая	13,788	13,240	13,161	13,337	15,007	13,536	17,180
		тестовая	16,888	15,196	15,394	16,141	16,017	16,447	16,334
	LeakyReLU	обучающая	15,565	14,465	13,234	12,457	12,390	11,946	11,816
		тестовая	16,559	15,415	14,617	15,648	14,348	15,257	14,429
14	TanH	обучающая	13,552	13,268	13,587	13,540	18,320	15,855	14,185
		тестовая	15,654	15,196	15,979	15,985	16,287	16,117	20,807
	LeakyReLU	обучающая	15,896	14,388	12,587	11,956	11,572	11,651	11,284
		тестовая	17,513	16,005	15,011	15,628	14,859	15,035	14,713
15	TanH	обучающая	13,309	13,396	12,798	13,924	12,911	13,315	13,571
		тестовая	15,376	15,240	14,979	17,218	15,068	15,139	17,018
	LeakyReLU	обучающая	14,814	13,733	13,061	12,682	11,741	11,548	11,398
		тестовая	15,388	14,747	14,443	14,772	14,066	13,883	13,459

Таблица 3.5 – Оценки погрешности прогноза $Q_г$ с помощью ИНС для Месторождения 1

Число входов ИНС	**Функция активации**	**Выборка**	**WAPE, %** **Для указанного числа скрытых слоёв**						
			2	3	4	5	6	7	8
6	TanH	обучающая	21,740	23,069	21,803	21,970	23,998	22,995	22,884
		тестовая	20,244	21,491	19,794	20,403	24,466	19,538	20,604
	LeakyReLU	обучающая	24,644	22,535	24,707	23,411	24,967	19,044	22,056
		тестовая	23,099	22,349	24,855	21,629	24,488	22,137	24,406
7	TanH	обучающая	22,739	22,520	35,902	22,509	27,447	21,881	25,297
		тестовая	24,083	20,923	29,393	23,589	27,273	21,632	23,500
	LeakyReLU	обучающая	25,268	20,611	22,089	20,429	18,825	21,191	19,044
		тестовая	23,756	22,245	24,302	23,188	20,693	23,196	19,431
8	TanH	обучающая	22,093	21,714	27,792	21,116	24,625	23,329	23,931
		тестовая	25,598	23,853	33,667	25,795	31,379	27,284	30,741
	LeakyReLU	обучающая	27,908	29,307	27,410	27,863	27,599	30,711	21,948
		тестовая	31,123	33,448	31,632	32,575	32,025	34,413	26,976
9	TanH	обучающая	23,396	22,468	23,567	29,935	25,605	28,160	25,765
		тестовая	30,102	27,184	24,859	34,116	33,061	34,004	28,312
	LeakyReLU	обучающая	25,813	31,587	25,631	27,741	26,753	23,196	20,103
		тестовая	30,366	34,651	29,281	31,970	30,856	27,250	24,454
10	TanH	обучающая	23,311	24,016	21,647	24,102	24,844	29,029	29,029
		тестовая	23,303	23,077	21,239	21,365	24,737	30,247	30,247
	LeakyReLU	обучающая	22,464	21,428	19,319	19,512	18,900	17,968	19,238
		тестовая	21,952	21,606	20,641	21,009	21,139	22,342	21,257
11	TanH	обучающая	21,719	21,628	22,533	20,339	26,478	24,407	25,850
		тестовая	26,487	24,310	23,931	23,543	28,753	27,469	26,281
	LeakyReLU	обучающая	19,852	18,656	19,083	16,979	14,945	17,164	16,619
		тестовая	22,553	22,153	22,372	21,671	**17,129**	22,728	22,481
12	TanH	обучающая	26,036	33,570	29,809	26,133	33,407	22,097	36,058
		тестовая	24,198	32,144	25,394	26,181	28,647	21,933	37,391
	LeakyReLU	обучающая	21,302	20,819	21,064	19,100	18,142	19,341	18,161
		тестовая	22,085	22,193	23,069	21,339	19,538	22,215	20,934
13	TanH	обучающая	22,527	26,218	32,913	46,369	25,988	27,603	30,091
		тестовая	23,411	24,265	29,229	38,263	24,369	25,182	29,022
	LeakyReLU	обучающая	21,053	20,875	19,924	20,392	19,980	18,116	18,116
		тестовая	22,011	22,524	21,558	21,881	21,941	21,647	21,770
14	TanH	обучающая	26,731	30,425	38,074	27,432	24,748	28,869	31,171
		тестовая	25,364	28,383	35,980	26,964	24,811	27,414	27,507
	LeakyReLU	обучающая	20,611	21,087	20,671	19,976	19,256	18,903	19,939
		тестовая	22,019	22,602	22,126	21,306	21,239	21,651	22,308
15	TanH	обучающая	25,880	22,149	29,200	22,750	27,302	25,999	28,357
		тестовая	26,281	24,280	26,786	23,975	26,797	28,873	25,747
	LeakyReLU	обучающая	21,618	20,552	18,562	18,272	18,595	18,202	17,526
		тестовая	24,087	23,084	21,870	22,104	22,806	21,933	22,271

Таблица 3.6 – Оценки погрешности прогноза $Q_ж$ с помощью ИНС для Месторождения 1

Число входов ИНС	**Функция активации**	**Выборка**	**WAPE, %** **Для указанного числа скрытых слоёв**						
			2	3	4	5	6	7	8
6	TanH	обучающая	16,501	15,478	21,056	18,322	14,486	16,216	17,921
		тестовая	16,722	14,052	20,989	17,246	13,597	14,281	17,597
	LeakyReLU	обучающая	16,145	13,611	14,498	13,801	14,447	13,388	13,923
		тестовая	15,956	14,897	15,087	14,447	15,887	14,188	15,634
7	TanH	обучающая	15,488	27,513	16,301	15,869	17,115	47,482	16,746
		тестовая	15,116	27,515	15,721	15,790	16,491	46,277	15,255
	LeakyReLU	обучающая	17,008	14,077	13,896	13,619	12,764	13,795	13,013
		тестовая	15,970	14,666	15,543	16,205	13,589	15,361	14,480
8	TanH	обучающая	46,749	14,745	14,968	45,856	15,168	13,840	15,950
		тестовая	50,587	17,220	19,037	50,496	18,275	18,320	17,566
	LeakyReLU	обучающая	15,177	13,601	13,354	13,485	12,866	12,547	15,657
		тестовая	18,123	16,716	16,977	17,394	15,709	14,532	18,988
9	TanH	обучающая	12,635	12,390	14,054	14,848	13,704	14,921	16,013
		тестовая	16,347	17,222	18,057	16,610	16,183	15,869	16,019
	LeakyReLU	обучающая	15,258	14,002	13,309	12,750	12,479	12,624	13,267
		тестовая	17,789	16,841	16,293	16,100	15,114	16,392	16,211
10	TanH	обучающая	14,433	14,956	18,115	15,284	19,740	16,272	15,942
		тестовая	15,381	14,162	16,928	14,674	17,704	15,488	15,286
	LeakyReLU	обучающая	13,628	12,611	11,936	11,784	11,632	11,048	11,129
		тестовая	14,749	14,117	13,813	13,916	14,036	**13,584**	13,776
11	TanH	обучающая	14,581	15,284	14,348	15,667	15,195	13,935	17,295
		тестовая	15,004	16,513	14,123	17,404	14,682	15,385	17,419
	LeakyReLU	обучающая	14,443	13,135	12,740	11,926	12,693	11,849	11,754
		тестовая	15,019	14,002	13,811	13,473	14,154	13,364	13,163
12	TanH	обучающая	14,990	13,698	14,938	16,278	15,758	18,484	15,691
		тестовая	16,240	15,146	15,752	16,019	16,694	17,498	15,596
	LeakyReLU	обучающая	14,358	13,872	13,056	12,238	13,196	12,543	12,460
		тестовая	15,122	15,174	15,258	14,723	15,065	15,148	14,536
13	TanH	обучающая	46,089	15,644	21,249	18,032	16,165	46,188	54,966
		тестовая	45,983	16,013	19,934	17,407	17,182	45,082	52,779
	LeakyReLU	обучающая	13,805	12,977	12,519	12,179	12,560	11,826	11,749
		тестовая	15,160	14,731	14,429	14,972	15,229	14,767	14,530
14	TanH	обучающая	45,880	20,627	17,625	14,593	17,457	20,175	17,127
		тестовая	45,606	20,631	17,652	15,296	17,279	18,622	17,078
	LeakyReLU	обучающая	14,287	13,141	12,631	12,392	12,120	12,339	12,710
		тестовая	15,466	14,692	14,794	14,613	14,571	14,923	14,699
15	TanH	обучающая	15,881	18,286	13,662	46,255	16,858	15,373	17,532
		тестовая	17,078	19,825	15,904	45,568	19,669	16,594	17,793
	LeakyReLU	обучающая	12,973	12,547	11,968	11,470	11,833	10,956	10,733
		тестовая	16,159	16,185	15,713	15,095	15,989	14,824	14,994

Таблица 3.7 – Оценки погрешности прогноза $Q_в$ с помощью ИНС для Месторождения 1

Число входов ИНС	**Функция активации**	**Выборка**	**WAPE, %** **Для указанного числа скрытых слоёв**						
			2	3	4	5	6	7	8
6	TanH	обучающая	38,506	35,293	33,383	28,912	29,693	30,127	26,568
		тестовая	55,566	51,268	46,623	44,410	51,789	41,067	41,371
	LeakyReLU	обучающая	41,848	32,992	31,733	28,955	29,606	26,220	27,523
		тестовая	60,689	50,748	47,622	78,183	44,019	40,459	47,188
7	TanH	обучающая	43,758	36,292	32,732	34,555	31,213	33,036	32,341
		тестовая	43,151	36,378	36,378	33,036	31,820	33,644	32,906
	LeakyReLU	обучающая	50,878	41,067	36,509	32,862	27,306	31,516	31,560
		тестовая	43,541	42,847	40,893	34,251	29,042	30,735	35,206
8	TanH	обучающая	36,205	34,859	31,126	26,611	32,602	28,217	23,312
		тестовая	44,583	45,625	40,546	39,330	46,320	40,372	40,372
	LeakyReLU	обучающая	40,937	35,858	32,906	30,605	30,518	31,039	26,828
		тестовая	49,836	44,236	44,713	43,151	36,899	51,486	39,547
9	TanH	обучающая	36,422	34,382	31,473	28,391	29,780	24,223	33,036
		тестовая	43,541	45,972	46,623	40,112	43,672	44,106	48,881
	LeakyReLU	обучающая	43,281	41,241	34,208	30,388	30,865	25,135	24,657
		тестовая	54,047	64,596	51,789	64,075	45,886	37,637	41,284
10	TanH	обучающая	38,537	32,175	32,070	31,329	28,773	28,009	31,029
		тестовая	42,718	41,338	47,539	49,412	42,769	36,373	45,891
	LeakyReLU	обучающая	41,037	32,977	27,588	29,446	19,727	24,225	21,300
		тестовая	32,977	47,574	37,360	41,141	**29,513**	33,326	37,456
11	TanH	обучающая	37,551	32,385	36,899	27,566	28,868	27,132	27,479
		тестовая	45,538	40,763	48,143	40,676	43,932	42,890	44,713
	LeakyReLU	обучающая	39,808	38,158	28,000	23,659	27,175	24,397	25,526
		тестовая	41,067	40,633	47,752	28,651	38,289	32,732	33,296
12	TanH	обучающая	41,501	31,560	31,213	29,910	30,388	23,833	30,301
		тестовая	50,227	43,194	44,887	50,530	62,425	43,672	46,320
	LeakyReLU	обучающая	43,411	34,772	28,000	26,698	28,391	21,749	26,177
		тестовая	52,180	50,530	36,986	40,025	50,791	38,202	40,372
13	TanH	обучающая	40,372	34,295	29,910	33,340	29,606	29,693	39,374
		тестовая	44,236	41,675	44,323	42,369	47,231	57,476	47,144
	LeakyReLU	обучающая	46,841	31,733	28,391	26,481	30,301	28,087	27,740
		тестовая	50,270	44,800	40,546	44,366	47,839	43,975	62,946
14	TanH	обучающая	36,205	29,129	27,262	30,301	30,388	36,552	30,084
		тестовая	43,498	46,971	39,678	46,971	47,275	57,476	41,458
	LeakyReLU	обучающая	40,329	30,561	27,392	27,653	28,521	29,693	29,259
		тестовая	46,016	38,896	40,242	42,673	51,138	40,155	44,279
15	TanH	обучающая	34,034	27,349	32,645	29,910	23,138	44,279	29,042
		тестовая	36,986	32,385	36,075	33,600	31,343	41,371	33,166
	LeakyReLU	обучающая	37,941	28,999	25,916	24,050	22,921	22,704	23,268
		тестовая	43,020	33,513	33,036	26,785	26,481	30,258	29,780

Лучшие результаты из всех четырёх таблиц в виде значений взвешенной средней абсолютной погрешности WAPE при прогнозе дебитов, полученные при исследованиях различных моделей ИНС на тестовых выборках, занесены в таблицу 3.8. Сравнительный анализ лучших результатов из таблицы 3.8 с результатами для взвешенной средней абсолютной погрешности WAPE при прогнозе дебитов из таблицы 3.2. показал, что для каждого из дебитов значения погрешности прогноза с помощью моделей ИНС ниже, чем полученные методом экстраполяции по скользящей средней (методика предприятия). Такое уменьшение погрешности прогноза дебитов с помощью ИНС указано для тестовых выборок в сравнении с результатами из таблицы 3.2 в последнем столбце таблицы 3.8. Видно, что лучший результат прогноза каждого из дебитов, полученный с использованием моделей ИНС с рациональными параметрами, имеет погрешность в два и более раза меньшую, чем у результата прогноза каждого из дебитов по методике, принятой на предприятии [115, 116].

Таблица 3.8 – Лучшие результаты решения *Задачи 1* с помощью ИНС для Месторождения 1

Дебит	**Взвешенная средняя абсолютная погрешность WAPE, %**	**Рациональные параметры моделей ИНС**	**Уменьшение погрешности, раз**
$Q_н$	12,477	10 входов ИНС; 4 скрытых слоя; функция активации LeakyReLU	2,192
$Q_г$	17,129	11 входов ИНС; 6 скрытых слоёв; функция активации LeakyReLU	2,051
$Q_ж$	13,584	10 входов ИНС; 7 скрытых слоёв; функция активации LeakyReLU	1,968
$Q_в$	29,513	10 входов ИНС; 6 скрытых слоёв; функция активации LeakyReLU	2,272

Как и для фонда скважин Месторождения 1, были проведены эксперименты для фонда Месторождения 2 с различными моделями ИНС при решении на обучающих и тестовых выборках *Задачи 1* для дебита УВС и дебита воды $Q_в$ на данных, сгруппированных в виде 20 пар файлов. Использовалось от 3060 до 3803 обучающих примеров в зависимости от числа входных параметров ИНС и вида дебита. Результаты исследования для дебита УВС в качестве примера представлены в таблице 3.9, а лучшие результаты для дебита УВС и дебита воды – в таблице 3.10.

Таблица 3.9 – Оценки погрешности прогноза дебита УВС с помощью ИНС для Месторождения 2

Число входов ИНС	**Функция активации**	**Выборка**	**WAPE, %** **Для указанного числа скрытых слоёв**						
			2	3	4	5	6	7	8
6	TanH	обучающая	5,622	7,448	9,567	11,509	6,119	10,787	9,872
		тестовая	5,695	7,239	9,629	12,044	6,016	10,848	9,512
	LeakyReLU	обучающая	3,361	3,446	3,159	3,100	3,032	2,951	3,293
		тестовая	3,498	3,595	3,394	3,405	3,320	3,225	3,526
7	TanH	обучающая	5,978	7,109	7,361	8,208	8,327	7,273	10,077
		тестовая	5,592	6,855	6,734	7,941	7,997	6,562	9,171
	LeakyReLU	обучающая	3,403	3,497	3,341	3,462	3,134	3,513	3,858
		тестовая	3,383	3,418	3,301	3,346	3,168	3,451	3,803
8	TanH	обучающая	15,691	10,027	10,959	14,335	12,108	12,051	35,883
		тестовая	16,003	10,477	11,320	14,808	11,981	12,440	35,823
	LeakyReLU	обучающая	4,198	4,277	4,027	4,330	3,588	3,408	3,514
		тестовая	4,348	4,637	4,448	4,823	4,227	4,068	4,214
9	TanH	обучающая	9,271	5,852	8,890	10,393	8,949	13,233	39,432
		тестовая	9,036	6,185	9,333	10,508	9,518	13,337	39,324
	LeakyReLU	обучающая	3,657	3,935	3,645	3,739	3,627	3,213	3,732
		тестовая	4,052	4,350	3,985	4,014	4,010	3,642	4,215
10	TanH	обучающая	27,720	22,608	13,602	15,265	14,242	40,524	16,087
		тестовая	27,810	22,533	13,662	15,385	14,649	40,154	15,948
	LeakyReLU	обучающая	3,513	3,452	3,595	3,207	3,115	2,979	2,948
		тестовая	3,912	4,006	3,995	3,666	3,520	3,479	3,510
11	TanH	обучающая	9,560	40,193	14,672	22,841	11,742	40,014	39,305
		тестовая	10,165	40,610	15,458	23,482	11,899	40,447	39,823
	LeakyReLU	обучающая	3,927	3,392	3,354	3,147	**2,960**	3,098	3,208
		тестовая	3,982	3,404	3,492	3,197	**3,125**	3,236	3,392
12	TanH	обучающая	10,466	20,121	18,143	39,919	41,081	40,282	41,484
		тестовая	10,656	20,281	19,495	40,828	41,389	40,894	41,694
	LeakyReLU	обучающая	3,790	3,290	4,761	3,946	2,848	5,850	2,846
		тестовая	4,135	3,696	5,136	4,340	3,240	6,270	3,198
13	TanH	обучающая	10,276	40,237	10,806	12,027	12,330	39,238	41,772
		тестовая	10,148	41,158	40,445	12,381	12,667	39,940	41,761
	LeakyReLU	обучающая	4,062	4,319	3,019	3,021	3,184	4,070	3,222
		тестовая	4,309	4,596	3,292	3,375	3,560	4,277	3,571
14	TanH	обучающая	39,849	18,494	39,664	8,106	41,552	41,576	41,407
		тестовая	40,706	18,281	40,371	8,478	41,715	41,715	41,256
	LeakyReLU	обучающая	3,967	3,208	3,414	2,994	2,922	2,881	2,869
		тестовая	4,236	3,518	3,778	3,334	3,423	3,209	3,353
15	TanH	обучающая	26,403	40,031	17,965	41,567	40,791	41,786	41,799
		тестовая	26,727	40,779	17,982	41,454	41,276	41,602	41,677
	LeakyReLU	обучающая	3,949	3,135	3,073	2,956	2,954	2,943	2,867
		тестовая	4,169	3,326	3,391	3,506	3,512	3,565	3,416

Из сравнения результатов в таблицах 3.2 и 3.9 и результатов по дебиту воды $Q_в$ следует, что использование моделей ИНС с рациональными параметрами для фонда Месторождения 2 позволило, как и в случае фонда Месторождения 1, уменьшить погрешность решения *Задачи 1* в два и более раза по сравнению с результатами прогноза каждого из дебитов по методике, принятой на предприятии (таблица 3.10).

Таблица 3.10 – Лучшие результаты решения *Задачи 1* с помощью ИНС для Месторождения 2

Дебит	**Взвешенная средняя абсолютная погрешность WAPE, %**	**Рациональные параметры моделей ИНС**	**Уменьшение погрешности, раз**
$Q_н + Q_г + Q_ж$	8,098	11 входов ИНС; 6 скрытых слоёв; функция активации LeakyReLU	2,591
$Q_в$	17,140	10 входов ИНС; 6 скрытых слоёв; функция активации LeakyReLU	3,680

3.2.3. Исследование статистических методов прогноза значений дебитов скважин

Чтобы понять, насколько высокую точность решения *Задачи 1* дают предложенные модели ИНС и используемый на предприятии метод экстраполяции по скользящей средней (методика предприятия), задачу прогноза необходимо решить с помощью других наиболее известных статистических методов. Среди них нами выбраны аппарат регрессионных моделей и статистические методы экстраполяции.

Задачей регрессионных моделей является описание поведения того или иного процесса с помощью определённой функции. В нашем случае необходимо описать зависимость компонента дебита скважины ($Q_н$, $Q_г$, $Q_ж$, или $Q_в$) от времени (t, день). Наиболее простой является линейная модель *Q(t)* для компонента дебита скважины

$$Q(t) = a + bt, \tag{3.16}$$

где a, b *– искомые коэффициенты модели.*

Такая модель регрессии позволит по полученным коэффициентам в формуле (3.16) посчитать интересующее значение дебита через месяц и тем самым спланировать технологический режим работы скважины. Это значение можно будет сравнить с фактическим (измеренным) значением дебита через месяц от даты планирования и, посчитав погрешность WAPE, понять насколько адекватно модель описывает процесс прогноза в сравнении с методикой предприятия и моделями ИНС.

Схожим образом работают методы экстраполяции, позволяя использовать тенденции прошлого для предсказания будущих значений дебитов. Предполагается, что тренд при этом не меняется. Также следует понимать, что чем дальше прогнозируемое значение во времени, тем больше вероятность потенциальной смены тренда и, соответственно, увеличения погрешности прогноза.

При проведении исследований были сделаны следующие допущения. Если модель выдавала значение ниже нуля, то прогнозное значение приравнивалось к нулю. Если корреляция между временем и дебитом отсутствовала (дебит держится на одном и том же уровне), то в качестве прогнозного значения выбиралось значению этой константы.

Для поиска коэффициентов a и b модели линейной регрессии (3.16) использовался наиболее простой критерий наименьших квадратов [117]. К недостаткам модели можно отнести жесткую фиксацию тренда моделью, небольшой период упреждения, сложность подбора уравнения регрессии [118]. Для её построения использовался весь набор данных по дебитам скважин Месторождения 1, что и при обучении моделей ИНС, представляющий собой объединённую обучающую и тестовую выборки.

В экспериментах варьировалось число известных имеющихся значений дебита (от 6 до 15 значений), полученных до даты, в которую происходило планирование. Результаты исследований представлены в таблице 3.11. В ней и в последующих таблицах жирными цифрами показаны наименьшие значения погрешности прогноза для каждого из дебитов.

Таблица 3.11 – Оценки погрешности решения *Задачи 1* с помощью модели линейной регрессии

Число значений дебита	**$Q_н$**	**$Q_г$**	**$Q_ж$**	**$Q_в$**
	WAPE, %			
6	50,858	80,690	67,923	143,173
7	47,089	73,395	60,296	129,097
8	44,546	68,376	55,168	119,423
9	**42,904**	**64,844**	51,583	112,625
10	44,409	70,939	52,462	106,792
11	44,978	73,616	52,337	102,960
12	45,187	74,807	51,902	100,054
13	45,228	74,962	51,316	97,787
14	45,207	74,636	50,697	95,951
15	45,018	74,044	**49,917**	**94,301**

В таблицах 3.12–3.15 представлены результаты исследований, полученные для тех же наборов данных с помощью различных методов экстраполяции (интерполяции), которые имеются в известной системе MatLAB [119].

Таблица 3.12 – Оценки погрешности решения *Задачи 1* при использовании сохраняющего форму кусочно-кубического полинома Эрмита (pchip)

Число значений дебита	**$Q_н$**	**$Q_г$**	**$Q_ж$**	**$Q_в$**
	WAPE, %			
6	405,701	447,293	377,672	154,313
7	63,175	65,810	64,866	79,660
8	55,534	60,850	63,269	113,076
9	**34,021**	**47,623**	**34,690**	77,289
10	122,273	117,757	158,451	184,924
11	34,981	48,733	35,480	**75,378**
12	91,908	172,428	108,454	76,141
13	36,338	51,433	36,609	76,846
14	36,821	51,939	37,010	77,595
15	210,996	573,986	266,022	78,078

Таблица 3.13 – Оценки погрешности решения *Задачи 1* при использовании модифицированного Акима метода Эрмита (makima)

Число значений дебита	**$Q_н$**	**$Q_г$**	**$Q_ж$**	**$Q_в$**
	WAPE, %			
6	715,036	799,864	600,651	186,271
7	51,120	58,363	49,512	82,394
8	34,151	**47,568**	35,074	77,395
9	**34,021**	47,623	**34,690**	77,289
10	104,744	103,862	160,150	1050,4
11	34,981	48,733	35,480	**75,378**
12	119,424	229,606	173,465	76,270
13	36,338	51,433	36,609	76,846
14	36,821	51,939	37,010	77,595
15	77,348	97,251	85,339	96,060

Таблица 3.14 – Оценки погрешности решения *Задачи 1* при использовании сплайн-интерполяции (spline)

Число значений дебита	**$Q_н$**	**$Q_г$**	**$Q_ж$**	**$Q_в$**
	WAPE, %			
6	2327,5	1814,8	2324,7	3079,9
7	1789,3	2685,5	1787,4	3892,3
8	485,3	285,7	368,2	475,1
9	195,9	287,4	208,8	354,6
10	8542,5	2844,9	1132,2	874,6
11	**83,0**	94,948	118,1	809,6
12	1318,7	4406,8	1739,5	947,983
13	102,2	**201,2**	**179,9**	129,1
14	1082,6	3447,0	1389,0	140,2
15	238,7	658,2	308,0	**81,1**

Таблица 3.15 – Оценки погрешности решения *Задачи 1* при использовании интерполяции методом ближайшего соседа (nearest)

Число значений дебита	**$Q_н$**	**$Q_г$**	**$Q_ж$**	**$Q_в$**
	WAPE, %			
6	**22,194**	**31,356**	**23,714**	**53,931**
7	22,396	31,635	23,815	54,132
8	22,606	31,996	23,963	54,422
9	22,827	32,407	24,150	54,730
10	24,350	37,068	26,054	54,573
11	24,566	37,426	26,240	54,822
12	24,741	37,726	26,385	55,138
13	24,932	37,954	26,542	55,472
14	25,134	38,197	26,712	55,840
15	25,391	38,526	26,842	56,126

Все наилучшие результаты (наименьшие значения погрешности прогноза значений дебитов) исследованных методов и модели линейной регрессии сведены в таблицу 3.16. Сюда же внесены минимальные значения этой погрешности, полученные для методики предприятия (таблица 3.2) и для моделей ИНС в случае Месторождения 1 (таблица 3.8).

Таблица 3.16 – Сводные результаты по оценке погрешности при решении *Задачи 1*, полученные с помощью методики предприятия, статистических методов и моделей ИНС для Месторождения 1

Дебит	**Методика предприятия**	**Линейная регрессия**	**Полином Эрмита**	**Полином Акима**	**Сплайн экстраполяция**	**Метод ближайшего соседа**	**Модели ИНС**
	WAPE, %						
$Q_н$	27,347	42,904	34,021	34,021	83,0	22,194	12,477
$Q_г$	35,133	64,844	47,623	47,568	201,2	31,356	17,129
$Q_ж$	26,734	49,917	34,690	34,690	179,9	23,714	13,584
$Q_в$	67,044	94,301	75,378	75,378	81,1	53,931	29,513

Из таблицы 3.16 видно, что используемый на предприятии метод экстраполяции по скользящей средней даёт меньшую погрешность при прогнозе каждого из дебитов по сравнению с другими исследуемыми статистическими методами, но уступает методу ближайшего соседа. При этом полученные путём исследований модели ИНС с рациональными параметрами дают результаты прогноза дебитов с гораздо меньшей погрешностью, чем статистические методы и применяемая на предприятии методика. Это ещё раз подчеркивает перспективность моделей ИНС по сравнению с другими методами решения слабоформализуемой *Задачи 1*.

В заключение отметим, что полученные модели ИНС с рациональными параметрами могут использоваться при прогнозе значений других технологических параметров скважин при соответствующей предварительной подготовке значений этих параметров.

3.3. Модели, методы и алгоритмы решения задач управления ГТМ

Так как проведение большинства ГТМ финансово затратно, то качество управления ГТМ должно быть очень высоким. Использование сложных аналитических или вероятностных методов, в частности, использование геолого-гидродинамических моделей залежей, требует серьёзных затрат вычислительных ресурсов, однако не всегда даёт точные результаты, позволяющие принять адекватные управленческие решения о необходимости проведения тех или иных ГТМ.

В качестве альтернативы простым и далеко не точным методам управления ГТМ, проанализированным в п. 1.2.1, а также появляющимся сложным аналитическим методам, сегодня начинают применять методы ИАД. Такие методы имеют меньшую трудоёмкость, в отличие от существующих аналитических методов и подходов, но оставляют открытым вопрос об их точности при решении задач управления ГТМ.

Основная идея использования методов ИАД во всех областях знаний состоит в том, чтобы программно-технический комплекс без помощи человека

позволил решать тот или иной полезный, но слабоформализуемый класс задач. В нашем случае – выбирать скважины-кандидаты для ГТМ и конкретные ГТМ для них (см. п. 3.1.1). Проблема здесь состоит не в том, чтобы полностью заменить существующие методы и средства работы специалиста промысла интеллектуальным инструментом, а оказать ему информационно-аналитическую поддержку в принятии сложных решений при управлении ГТМ.

3.3.1. Предварительная подготовка данных при решении задач классификации скважин и типов ГТМ

Предварительная подготовка данных для последующего решения задач управления ГТМ осуществлялась в соответствии с методами и подходами, изложенными в п. 3.1.2.

Обучающая и тестовая выборки для моделей ИНС формировались на основе БД технологических и геологических параметров фонда скважин Месторождения 1 и отдельно на основе БД, содержащей меньшую номенклатуру таких параметров фонда Месторождения 2.

В качестве наборов данных для обучения и тестирования различных моделей ИНС на Месторождении 1 были выбраны значения указанных ниже параметров для каждого комплекса «скважина-пласт». Использовались следующие геологические и технологические параметры: дебит нефти $Q_н$, т; дебит газа $Q_г$, тыс. м^3; дебит жидкости $Q_ж$, м^3; дебит воды $Q_в$, т; давление забойное $P_{заб}$, атм; давление буферное $P_{буф}$, атм; давление затрубное $P_{зат}$, атм; температура на устье скважины $T_у$, °С; давление пластовое $P_{пл}$, атм; пластовая температура $T_{пл}$, °С; обводнённость W, %; эффективная мощность пласта H, м; газовый фактор B; диаметр штуцера D, мм. К сожалению, не для всех этих параметров на фонде скважин Месторождения 2 имеется достаточный объём данных для обучения ИНС, поэтому для его комплексов «скважина-пласт» в исследованиях использовался урезанный набор из таких параметров. Скважины, признанные специалистами геологических служб предприятий кандидатами на ГТМ, и проведённые ГТМ считались эталонными и имели соответствующие метки.

Формирование обучающей и тестовой выборок для исследования моделей ИНС проводилось по разработанному адаптивному алгоритму предварительной подготовки (обработки) данных, подобному адаптивному алгоритму (рисунок 3.2), описанному в п. 3.2.1. Часть блоков этого алгоритма подобна блокам адаптивного алгоритма из п. 3.2.1 и позволяет отбраковывать ошибочные данные, анализировать корректность данных, определять характер пропусков в данных и адаптивно восстанавливать их с требуемой точностью. Однако в этом алгоритме присутствуют новые блоки, например, блок нормализации данных, часть правил формирования обучающих примеров и часть решающих правил при обнаружении ошибок и восстановлении пропусков в данных отличается от правил первого алгоритма. Далее будем рассматривать только особенности этого алгоритма. Отметим, что при предварительной подготовке значений дебитов скважин будет использоваться первый адаптивный алгоритм за исключением его последних двух блоков. Затем формирование наборов параметров, включая дебиты, для обучения и тестирования моделей ИНС проводится по второму алгоритму.

По данным ряда геологических служб добывающих предприятий известно, что планирование ГТМ происходит не ранее, чем за 4 месяца до его проведения. Более того, нет той чёткой границы, когда происходит переход скважины из одного класса в другой класс (из класса рентабельных в класс нерентабельных и требующих ГТМ). Чаще всего это происходит в процессе её эксплуатации от одного проведённого ГТМ до другого путём постепенного изменения значений её параметров и геологических параметров продуктивного пласта до критических. Поэтому средний интервал времени между моментами времени проведения соседних ГТМ на скважине разумно считать «сомнительным для принятия решения о необходимости ГТМ» и исключить его из рассмотрения. В общем случае такой интервал не симметричен относительно середины интервала времени эксплуатации скважины между выполненными на ней ГТМ. Экспериментально, по данным из архивов данных предприятий было установлено, что «сомнительный» интервал времени может составлять до 30 % от интервалов времени между ГТМ:

$$\frac{T_{\text{ГТМ}_{n-1}} + T_{\text{ГТМ}_n}}{2} - \left(T_{\text{ГТМ}_n} - T_{\text{ГТМ}_{n-1}}\right) \times 0{,}15 < t_i < \frac{T_{\text{ГТМ}_{n-1}} + T_{\text{ГТМ}_n}}{2} + \left(T_{\text{ГТМ}_n} - T_{\text{ГТМ}_{n-1}}\right) \times 0{,}15, \quad (3.17)$$

где T_i *– время получения параметра в векторе значений (исключается из рассмотрения, при соблюдении условий);* $T_{ГТМn-1}$ *– время проведения предыдущего ГТМ или запуска скважины в эксплуатацию;* $T_{ГТМn}$ *– время проведения очередного ГТМ.*

Это было учтено при формировании обучающей и тестовой выборок для моделей ИНС и при оценке их достаточного для исследований объёма.

Фонд скважин месторождения часто имеет относительно небольшое количество добывающих скважин, особенно если это месторождение эксплуатируется не так давно. Ещё меньше в таком фонде тех скважин, которые подвергались ГТМ в выбранный период времени. Учитывая это, предложено использовать аугментацию данных – методику создания дополнительных обучающих данных из имеющихся данных. Применялся один из основных методов аугментации – добавление к ряду значений вышеперечисленных параметров исходной обучающей выборки случайного шума [120]. При пятипроцентном шуме значения параметров в векторе обучающего примера рассчитывались следующим образом:

$$p_i = p_i + rnd(-0{,}05; 0{,}05) \times p_i, \quad (3.18)$$

где p_i *– значение геологического или технологического параметра p в векторе значений параметров для конкретного обучающего примера; rnd() – функция получения случайного числа в выбранном диапазоне, в нашем случае диапазон* $\pm$*5 %. от исходного значения параметра p.*

Значения различных параметров скважин и продуктивных пластов измеряются в разных диапазонах, и для того чтобы каждый параметр оказывал влияние на конечный результат в экспериментах, осуществлялась нормализация данных [61]. Значения по каждому параметру были нормированы так, чтобы они находились в диапазоне [0, 1]:

$$p_i = \frac{p_i}{p_{max}}, \quad (3.19)$$

где p_i *– это i-е значение параметра p в векторе значений для конкретного обучающего примера;* p_{max} *– максимальное значение параметра p анализируемого фонда скважин, вошедшее в обучающую выборку.*

При исследовании важно понять, какое сочетание геологических и (или) технологических параметров содержит наиболее существенные признаки,

которые может извлечь модель ИНС для решения поставленных задач классификации. Формировались следующие обучающие и тестовые наборы параметров, подаваемых на вход ИНС.

- Набор № 1 – все измеряемые параметры скважин и продуктивных пластов: $Q_н$; $Q_г$; $Q_ж$; $Q_в$; $P_{заб}$; $P_{буф}$; $P_{зат}$; $T_у$; D; $P_{пл}$; $T_{пл}$; W; H; B.
- Набор № 2 – все технологические параметры скважин: $Q_н$; $Q_г$; $Q_ж$; $Q_в$; $P_{заб}$; $P_{буф}$; $P_{зат}$; $T_у$; B.
- Набор № 3 – параметры дебитов скважин: $Q_н$; $Q_г$; $Q_ж$; $Q_в$.
- Набор № 4 – технологические параметры скважин (без дебитов): $P_{заб}$; $P_{буф}$; $P_{зат}$; $T_у$; B.
- Набор № 5 – параметры дебитов скважин и параметры продуктивных пластов: $Q_н$; $Q_г$; $Q_ж$; $Q_в$; $P_{пл}$; $T_{пл}$.

Значения параметров в общем случае измеряются в разное время для каждой конкретной скважины. Чтобы получить вектор всех значений на определённый момент времени, необходимо было интерполировать значения некоторых параметров по соседним имеющимся значениям. Для этого использовались методы линейной интерполяции, кусочно-квадратичной интерполяции и одномерный кубический сплайн (формулы (3.12), (3.13) и (3.14)). Более того, было известно, что в БД Месторождения 1 часть данных временных рядов для дебитов скважин прореживалась для сокращения занимаемого объёма памяти на сервере. Для восстановления таких данных использовался метод линейной интерполяции и вычисления их производились по формуле (3.12).

Разработанный адаптивный алгоритм был программно реализован. При этом для извлечения информации из БД предприятий была создана хранимая процедура на языке SQL в СУБД Microsoft SQL Server. Блоки алгоритма по предварительной подготовке данных и формированию обучающих примеров реализованы на платформе .NET, в интегрированной среде разработки Microsoft Visual Studio 2019, на языке C#. Результаты применения созданных программных средств и хранимой процедуры представляются в новой БД в виде совокупности строк, очищенных и скорректированных, упорядоченных по кустам скважин и

отдельным скважинам, по времени измерения каждого параметра из набора и дате проведения каждого ГТМ.

Число обучающих и тестовых примеров, включая данные, полученные за счёт аугментации, составляет 2016 пар «признаки – метки» для каждого указанного выше обучающего и тестового набора. Объём обучающей выборки составляет 80 % от общего используемого объёма данных набора, а объем тестовой выборки – 20 % этого объёма данных.

3.3.2. Разработка моделей ИНС для решения задач классификации скважин и типов ГТМ

В п. 3.1.3 проанализированы различные типы ИНС и в качестве наиболее перспективного типа для решения поставленных в п. 3.1.1 слабоформализуемых задач классификации скважин и типов ГТМ выбраны глубокие ИНС прямого распространения FFNN. В рамках этого типа ИНС следует разработать модели ИНС, позволяющие наиболее точно решать эти задачи. По сути, необходимо провести исследование предлагаемых моделей ИНС прямого распространения с целью поиска их рациональной архитектуры и гиперпараметров при решении поставленных задач классификации. В п. 3.1.5 осуществлен выбор библиотеки CNTK для формирования и глубокого обучения ИНС. Разработан интерфейс пользователя ПО, созданного нами на основе этой библиотеки. Эти программные средства позволяют в интерактивном режиме задавать параметры при построении, обучении и тестировании моделей ИНС.

Число скрытых слоёв в моделях ИНС при исследовании выбиралось равным 2, 3, 4, …, 9, 10, 15, 20. Формировались модели ИНС с различным числом входов. В исследованиях осуществлялась поочередная подача на вход ИНС каждого из пяти сформированных наборов геологических и технологических параметров, поэтому число входов модели ИНС определялось числом параметров в наборе. Для всех таких моделей ИНС исследовался также эффект от применения различных функций активации (Sigmoid, TanH, ReLU) и алгоритмов оптимизации при обучении ИНС (SGD, Adam, AdaGrad, AdaDelta), кратко описанных в п. 3.1.4 и дающих лучшую сходимость результатов при обучении [121]. Изменялась также

скорость обучения: 0,2; 0,1; 0,05; 0,01; 0,005; 0,001. Также использовался слой Dropout с различными вероятностями [122].

Схема предложенной модели ИНС, приведённой в качестве примера, представлена на рисунке 3.5 [123]. Эта модель позволяет решать *Задачу 2*. Она включает в себя входной слой, куда подаются значения параметров из того или иного обучающего или тестового набора из пяти перечисленных для исследуемой скважины, совокупность скрытых слоёв (показаны 3 слоя с функцией активации TanH) и выходной слой Softmax. Выходной слой содержит два нейрона, позволяющие сделать вывод, является ли скважина кандидатом на ГТМ. Для предотвращения переобучения перед выходным слоем в ИНС использовался слой Dropout.

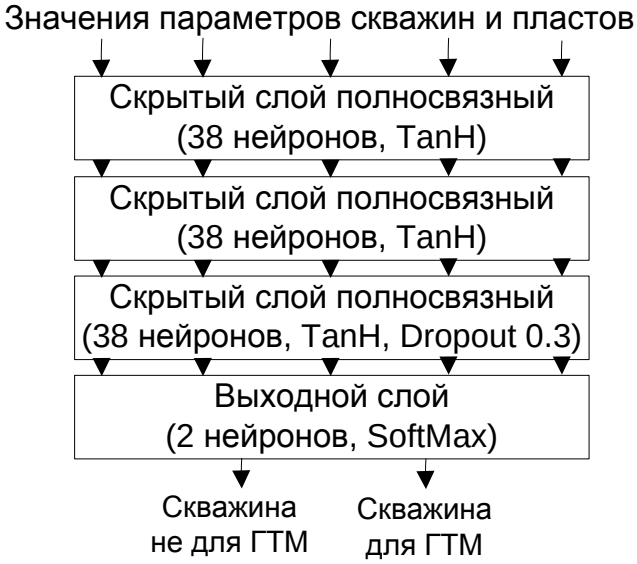

Рисунок 3.5 – Пример модели ИНС для решения *Задачи 2*

На рисунке 3.6 приведен пример модели ИНС для решения *Задачи 3*. Архитектура этой ИНС отличается от архитектуры модели ИНС на рисунке 3.5 тем, что в выходном слое располагается 4 нейрона, соответствующих конкретным видам работ: РИР, ЭЦН, ГРП, прочие ГТМ. При этом на вход ИНС подаются параметры только тех скважин, которые были выбраны в качестве кандидатов при решении задачи бинарной классификации [123].

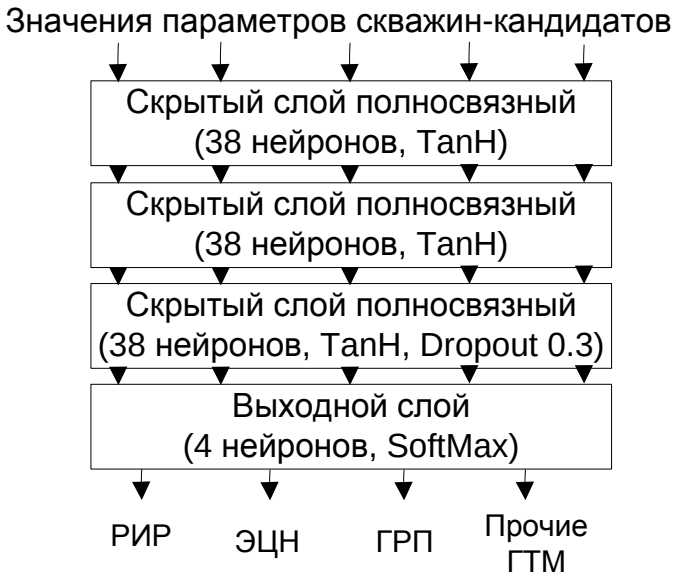

Рисунок 3.6 – Пример модели ИНС для решения *Задачи 3*

Точность классификации с помощью каждой из моделей ИНС оценивалась для обучающей выборки и для тестовой выборки отдельно путем вычисления специальной метрики. Наиболее простой является метрика Accuracy. Под ней понимается отношение скважин, правильно отнесённых к заранее известному классу, к общему числу анализируемых скважин в обучающей и тестовой выборках (в процентах) [96]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\ \%, \tag{3.20}$$

где TP – истинно-положительное решение; TN – истинно-отрицательное решение; FP – ложно-положительное решение; FN – ложно-отрицательное решение.

Несмотря на интуитивную понятность метрики Accuracy, она требует сбалансированной выборки для каждого из классов при бинарной классификации: оба класса должны быть представлены выборками приблизительно равного объёма. В реальности же так не всегда получается, например, на месторождении ГТМ на скважине проводятся гораздо реже, чем скважина работает без ГТМ. С этим можно справиться как с помощью применения к менее мощной выборке какого-либо метода аугментации, так и с помощью использования другой метрики. Предпочтителен второй вариант. Среди часто применяемых при решении задач классификации метрика Precision – доля скважин в реальности принадлежащих данному классу («кандидат для ГТМ» или «скважина, на которой

ГТМ проводить не следует») относительно всех скважин, которые система отнесла к этому классу [96]:

$$Precision = \frac{TP}{TP + FP} \times 100\ \%, \tag{3.21}$$

Далее при оценке точности бинарной классификации использовалась метрика Precision. Она же использовалась при оценке точности классификации в задаче выбора типа ГТМ.

Для оценки соответствия используемым для обучения данным каждой из исследуемых моделей ИНС с её выбранными гиперпараметрами применялась технология кросс-валидации [113]. Использовался метод *К*-блоков, причем *К* = 5. Это означает, что все вычисленные в ходе исследований итоговые значения точности классификации ТА скважин или типов ГТМ являются результатом усреднения по пяти значениям точности, полученным в ходе обучения и тестирования конкретной модели ИНС:

$$TA = \frac{\sum_{i=1}^{5} Precision_i}{5}, \tag{3.22}$$

где $Precision_i$ *– оценка точности классификации в случае i-ой позиции тестового блока данных.*

3.3.3. Результаты исследования эффективности моделей ИНС

Приведем результаты исследований эффективности каждой модели ИНС при решении поставленных задач классификации [124]. Отметим, что результаты получены при использованы как обучающей, так и тестовой выборок. В таблицах 3.17–3.21 показаны результаты экспериментов для Месторождения 1 с различными моделями ИНС при решении *Задачи 2*. Результаты приведены в виде усреднённой метрики Precision, рассчитанной по выражению (3.22), для всех пяти наборов данных по этому месторождению. В экспериментах использованы четыре варианта алгоритмов оптимизации при обучении ИНС, и для каждого из них – три варианта наиболее подходящих функций активации.

Таблица 3.17 – Результаты классификации скважин с использованием моделей ИНС в случае набора № 1

Алгоритм обучения	**Функция активации**	**Выборка**	**Точность классификации скважин для различного числа скрытых слоёв ИНС, %**										
			2	3	4	5	6	7	8	9	10	15	20
SGD	Sigmoid	Обучающая	88,239	71,838	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732
		Тестовая	88,022	70,879	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736
	TanH	Обучающая	93,648	94,021	94,207	98,686	95,695	98,693	**100**	98,882	95,510	94,388	84,741
		Тестовая	92,528	92,528	93,297	96,374	93,352	96,978	**99,231**	96,154	92,528	92,583	84,890
	ReLU	Обучающая	93,648	97,924	**99,060**	98,683	**100**	**100**	**100**	**100**	99,815	96,425	95,737
		Тестовая	93,242	97,088	**98,572**	96,374	**98,517**	**99,231**	**99,286**	**99,231**	97,857	94,945	94,011
AdaGrad	Sigmoid	Обучающая	71,083	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732
		Тестовая	70,879	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736
	TanH	Обучающая	94,022	94,584	94,958	94,958	98,316	**100**	**100**	**100**	**100**	94,022	94,584
		Тестовая	92,528	94,011	94,011	94,011	97,088	**98,517**	**98,517**	**99,231**	**99,231**	97,143	94,121
	ReLU	Обучающая	45,524	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732
		Тестовая	46,052	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736
AdaDelta	Sigmoid	Обучающая	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732	53,732
		Тестовая	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736	53,736
	TanH	Обучающая	53,732	69,395	74,256	72,568	71,827	71,642	72,390	70,339	71,831	71,460	86,373
		Тестовая	53,736	68,022	75,275	72,418	71,648	71,648	72,418	70,879	71,648	71,703	85,000
	ReLU	Обучающая	50,559	60,255	69,951	76,314	69,961	72,005	71,820	76,104	69,402	75,748	76,310
		Тестовая	50,899	59,725	70,220	76,154	70,165	71,703	70,934	76,209	68,681	76,154	76,868
Adam	Sigmoid	Обучающая	94,025	94,954	94,588	94,217	78,941	68,152	69,008	80,402	77,980	53,732	53,732
		Тестовая	92,528	91,703	88,626	90,934	79,011	66,923	66,538	76,868	75,440	53,736	53,736
	TanH	Обучающая	94,954	94,954	97,942	98,868	96,810	100	**100**	**100**	**100**	60,713	53,732
		Тестовая	95,495	95,495	97,802	98,571	97,802	97,802	**100**	**98,572**	**99,286**	60,165	53,736
	ReLU	Обучающая	100	100	**100**	**100**	**100**	100	100	100	100	100	100
		Тестовая	97,802	97,747	**98,517**	**97,088**	**98,517**	97,802	97,033	97,747	97,033	97,802	97,747

Таблица 3.18 – Результаты классификации скважин с использованием ИНС в случае набора № 2

Алгоритм обучения	**Функция активации**	**Выборка**	**Точность классификации скважин для различного числа скрытых слоёв ИНС, %**										
			2	3	4	5	6	7	8	9	10	15	20
SGD	Sigmoid	Обучающая	65,327	52,874	50	50	50	50	50	50	50	50	50
		Тестовая	64,205	52,092	50	50	50	50	50	50	50	50	50
	TanH	Обучающая	93,416	98,250	99,228	97,115	**99,332**	98,870	97,220	99,590	96,497	99,640	91,298
		Тестовая	92,594	96,722	97,734	95,910	**98,359**	97,730	95,268	97,326	95,702	97,334	89,337
	ReLU	Обучающая	94,853	98,148	98,765	98,970	99,127	97,323	99,333	99,024	98,095	98,971	94,539
		Тестовая	93,019	96,305	97,134	97,334	97,947	96,709	97,534	97,938	97,746	95,893	92,827
AdaGrad	Sigmoid	Обучающая	53,232	52,423	50	50	50	50	50	50	50	50	50
		Тестовая	52,708	52,245	50	50	50	50	50	50	50	50	50
	TanH	Обучающая	86,005	93,467	97,839	99,074	99,537	99,280	98,662	99,332	99,536	99,614	99,076
		Тестовая	85,604	93,414	96,913	97,742	96,702	96,913	96,918	96,922	97,334	97,118	96,084
	ReLU	Обучающая	49,692	47,737	56,995	51,594	50,566	46,912	50,051	50,671	46,135	52,881	50,000
		Тестовая	49,800	47,938	57,198	51,641	50,629	47,326	50,208	51,433	46,327	53,291	50,000
AdaDelta	Sigmoid	Обучающая	50,103	50	50	50	50	50	50	50	50	50	50
		Тестовая	50,000	50	50	50	50	50	50	50	50	50	50
	TanH	Обучающая	56,168	63,271	52,470	51,851	48,463	58,180	57,716	59,722	55,194	64,403	63,323
		Тестовая	55,761	62,751	52,462	52,270	48,963	58,019	57,194	59,881	54,736	64,392	63,376
	ReLU	Обучающая	52,830	62,447	61,058	61,266	66,716	60,495	60,805	61,575	58,537	70,113	75,670
		Тестовая	52,474	61,505	60,697	61,101	66,471	60,068	60,276	61,922	58,027	69,116	74,881
Adam	Sigmoid	Обучающая	84,158	89,760	96,039	82,665	81,662	73,243	71,504	55,132	51,751	50,825	50
		Тестовая	83,529	89,107	93,831	79,821	80,672	70,391	69,575	54,362	51,637	50,816	50
	TanH	Обучающая	99,794	**99,640**	98,508	99,383	**99,846**	99,537	99,279	98,256	96,393	51,388	51,540
		Тестовая	97,126	**98,563**	96,922	97,942	**98,767**	97,530	97,334	96,909	95,710	51,442	51,247
	ReLU	Обучающая	**99,486**	97,889	99,332	99,281	99,486	99,332	**99,538**	**99,589**	89,639	59,333	59,227
		Тестовая	**98,146**	96,298	97,326	97,126	97,326	97,942	**98,159**	**98,567**	88,163	59,583	59,184

Таблица 3.19 – Результаты классификации скважин с использованием ИНС в случае набора № 3

Алгоритм обучения	**Функция активации**	**Выборка**	**Точность классификации скважин для различного числа скрытых слоёв ИНС, %**										
			2	3	4	5	6	7	8	9	10	15	20
SGD	Sigmoid	Обучающая	53,321	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
		Тестовая	53,260	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
	TanH	Обучающая	61,685	65,548	64,790	68,598	67,482	73,310	72,857	69,326	71,344	68,659	59,932
		Тестовая	60,628	64,854	63,776	67,632	66,056	71,007	69,916	67,981	70,397	66,911	58,938
	ReLU	Обучающая	68,386	70,802	74,124	78,048	77,538	76,358	77,052	78,473	77,929	77,112	63,465
		Тестовая	68,958	70,892	71,013	74,277	73,905	71,859	72,705	72,823	73,061	72,954	61,231
AdaGrad	Sigmoid	Обучающая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
		Тестовая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
	TanH	Обучающая	59,058	63,949	66,303	66,664	69,533	76,087	77,355	74,368	76,567	73,461	76,663
		Тестовая	58,810	64,124	66,425	67,276	69,077	74,267	72,225	70,891	71,869	72,216	71,856
	ReLU	Обучающая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
		Тестовая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
AdaDelta	Sigmoid	Обучающая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
		Тестовая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
	TanH	Обучающая	53,261	53,261	53,261	53,563	53,804	53,503	53,321	54,318	53,895	54,256	63,798
		Тестовая	53,261	53,261	53,261	53,383	53,624	53,381	52,902	54,230	54,228	54,715	63,888
	ReLU	Обучающая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
		Тестовая	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261	53,261
Adam	Sigmoid	Обучающая	67,241	66,907	67,844	66,848	66,667	69,082	53,261	53,261	53,261	53,261	53,261
		Тестовая	67,386	66,306	68,596	67,146	66,905	67,634	53,261	53,261	53,261	53,261	53,261
	TanH	Обучающая	69,867	85,810	89,403	91,122	91,576	86,805	86,326	82,789	82,391	53,261	53,261
		Тестовая	68,710	78,262	79,236	79,114	79,959	77,665	76,451	74,400	75,380	53,261	53,261
	ReLU	Обучающая	78,835	83,817	90,731	91,666	93,569	91,880	89,764	88,738	88,135	56,097	53,261
		Тестовая	76,444	76,933	78,626	78,627	78,381	78,136	78,027	75,364	78,623	55,806	53,261

Таблица 3.20 – Результаты классификации скважин с использованием ИНС в случае набора № 4

Алгоритм обучения	**Функция активации**	**Выборка**	**Точность классификации скважин для различного числа скрытых слоёв ИНС, %**										
			2	3	4	5	6	7	8	9	10	15	20
SGD	Sigmoid	Обучающая	77,890	77,091	73,217	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
		Тестовая	77,740	76,877	73,186	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
	TanH	Обучающая	83,057	92,068	95,204	94,866	94,251	95,232	93,944	92,157	**95,417**	86,010	74,823
		Тестовая	83,027	92,123	94,706	94,458	94,091	95,085	93,596	91,641	**94,836**	86,710	74,383
	ReLU	Обучающая	91,267	94,772	95,172	95,511	**96,095**	93,882	91,793	93,083	95,726	84,593	88,993
		Тестовая	91,018	94,713	94,957	95,080	**95,572**	92,984	91,744	93,349	95,448	84,384	89,169
AdaGrad	Sigmoid	Обучающая	77,183	75,585	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
		Тестовая	76,753	75,400	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
	TanH	Обучающая	80,503	88,130	93,083	95,448	95,603	94,435	94,035	94,834	95,204	90,067	80,907
		Тестовая	80,446	87,452	93,722	94,467	95,203	94,217	93,846	94,832	94,954	89,424	81,046
	ReLU	Обучающая	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
		Тестовая	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
AdaDelta	Sigmoid	Обучающая	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
		Тестовая	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426	58,426
	TanH	Обучающая	58,426	62,730	65,467	70,234	74,692	70,972	71,833	70,880	71,494	73,217	77,183
		Тестовая	58,426	62,731	65,316	69,990	74,666	70,973	71,834	70,851	71,712	73,309	77,369
	ReLU	Обучающая	58,426	58,426	58,426	58,426	58,426	58,426	65,437	65,529	66,082	67,589	68,822
		Тестовая	58,426	58,426	58,426	58,426	58,426	58,426	65,439	65,438	66,053	67,653	68,604
Adam	Sigmoid	Обучающая	82,318	86,654	87,239	86,099	88,283	76,789	58,426	76,415	58,426	58,426	58,426
		Тестовая	82,411	86,467	87,083	85,983	87,824	76,118	58,426	76,015	58,426	58,426	58,426
	TanH	Обучающая	95,264	**97,478**	96,432	96,586	95,541	96,586	88,801	90,558	72,199	58,426	58,426
		Тестовая	94,957	**97,418**	95,330	96,313	95,328	95,454	88,214	90,658	72,336	58,426	58,426
	ReLU	Обучающая	93,788	96,340	**97,201**	97,140	96,185	96,986	94,709	94,433	96,247	78,191	68,203
		Тестовая	93,603	96,312	**97,173**	96,803	95,456	96,064	94,591	94,101	95,946	78,134	67,904

3.21 – Результаты классификации скважин с использованием ИНС в случае набора № 5

Алгоритм обучения	Функция активации	Выборка	Точность классификации скважин для различного числа скрытых слоёв ИНС, %										
			2	3	4	5	6	7	8	9	10	15	20
SGD	Sigmoid	Обучающая	68,273	56,705	56,666	56,725	56,705	56,705	56,725	56,725	56,705	56,705	56,705
		Тестовая	67,171	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,779	56,725	56,725
	TanH	Обучающая	87,577	91,520	95,469	**97,074**	97,363	97,951	90,215	90,140	89,330	87,588	73,390
		Тестовая	85,338	91,812	93,543	**95,274**	94,686	94,132	89,375	87,777	86,468	87,039	73,119
	ReLU	Обучающая	90,641	96,050	93,860	96,042	90,747	95,744	96,338	92,545	**97,955**	92,692	83,218
		Тестовая	90,064	92,989	91,762	92,972	89,005	94,115	94,115	91,174	**94,686**	91,174	81,186
AdaGrad	Sigmoid	Обучающая	61,263	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
		Тестовая	62,005	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
	TanH	Обучающая	79,975	90,931	95,173	95,611	95,758	97,224	94,280	93,104	92,983	97,657	88,259
		Тестовая	78,968	88,332	93,595	94,166	94,721	94,115	91,258	88,972	90,619	95,309	85,526
	ReLU	Обучающая	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
		Тестовая	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
AdaDelta	Sigmoid	Обучающая	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
		Тестовая	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
	TanH	Обучающая	57,750	56,725	56,725	56,725	60,236	56,729	61,261	59,358	64,767	62,577	66,082
		Тестовая	57,868	56,725	56,725	56,936	60,274	56,154	60,862	59,113	64,341	62,023	66,091
	ReLU	Обучающая	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
		Тестовая	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725	56,725
Adam	Sigmoid	Обучающая	92,246	93,123	83,744	86,525	81,278	74,686	70,010	78,510	69,718	56,725	56,725
		Тестовая	92,399	93,579	79,609	82,016	77,220	74,345	69,657	77,203	69,674	56,725	56,725
	TanH	Обучающая	96,049	99,267	98,680	**99,708**	99,854	99,416	95,468	87,211	**99,416**	74,815	57,313
		Тестовая	93,578	97,040	95,308	**97,057**	95,344	95,898	91,831	84,487	**97,058**	73,843	56,725
	ReLU	Обучающая	97,222	98,976	98,833	**99,562**	98,385	99,268	98,827	99,271	99,562	89,919	95,166
		Тестовая	94,149	96,452	95,864	**97,646**	97,500	96,452	96,470	96,470	95,864	87,107	91,276

На рисунке 3.7 приведены взятые из этих таблиц все лучшие результаты по точности классификации скважин для всех пяти наборов параметров в случае тестовых выборок. Они получены с помощью оптимизационных алгоритмов обучения Adam или SGD и функций активации TanH или ReLU.

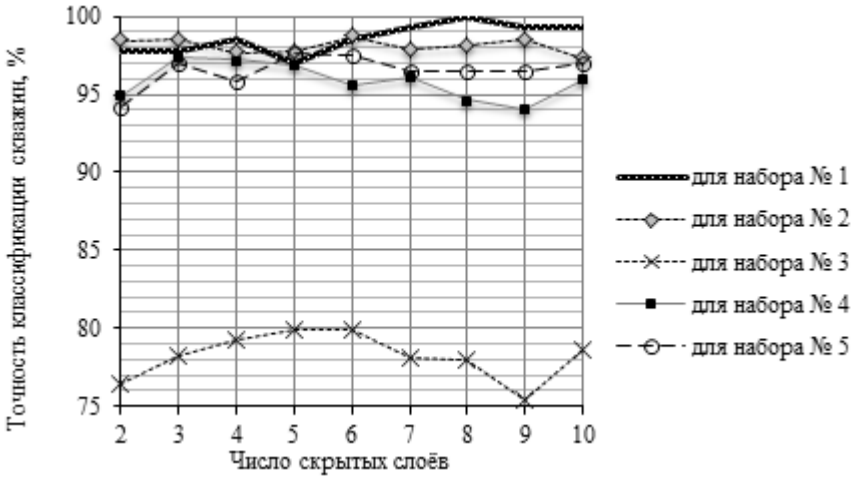

Рисунок 3.7 – Лучшие результаты по точности классификации скважин на тестовых выборках

Проанализируем полученные результаты. Видно, что основным фактором, влияющим на точность классификации, является различие моделей ИНС, в

первую очередь, по числу скрытых слоёв [125]. Функция активации Sigmoid для всех алгоритмов обучения и алгоритм оптимизации при обучении AdaDelta для всех функций активации не позволили получить высокую точность классификации скважин вследствие плохой обучаемости таких моделей ИНС.

Весьма низкие результаты наблюдаются при использовании в ИНС 2 или 3 скрытых слоёв для всех алгоритмов оптимизации при обучении, кроме алгоритма Adam. По-видимому, это связано с тем, что число извлекаемых в этих случаях признаков скважины для отнесения её к тому или иному классу не соответствует сложности задачи. Для числа скрытых слоёв более 10 для этих же алгоритмов оптимизации и функций активации результаты также не высоки. Из рисунка 3.7 следует, что наилучшая точность для разного числа скрытых слоёв ИНС достигнута для набора № 1, а наихудшая – для набора № 3.

В таблице 3.22 показаны результаты экспериментов с различными моделями ИНС при решении *Задачи 2* в случае Месторождения 2. В БД этого месторождения имеются значения параметров, соответствующие набору № 4. С помощью программных средств, реализующих второй адаптивный алгоритм предварительной подготовки исходных данных, из этой БД был извлечен и подготовлен 3371 обучающий пример параметров набора № 4. Кроме того, 20 % этих примеров использовались в качестве тестовой выборки для исследования моделей ИНС. Применение кросс-валидации данных позволило вычислять усреднённую метрику Precision согласно выражению (3.22).

Таблица 3.22 – Результаты классификации скважин для Месторождения 2 (набор № 4)

Алгоритм обучения	**Функция активации**	**Выборка**	**Точность классификации скважин для различного числа скрытых слоёв ИНС, %**										
			2	3	4	5	6	7	8	9	10	15	20
SGD	Sigmoid	Обучающая	78,590	80,722	81,150	81,521	53,210	53,210	53,210	53,210	53,210	53,210	53,210
		Тестовая	79,111	80,277	81,630	81,333	53,185	53,185	53,185	53,185	53,185	53,185	53,185
	TanH	Обучающая	95,621	**98,887**	98,367	99,852	99,147	99,072	98,998	95,399	95,436	89,314	98,924
		Тестовая	94,519	**98,222**	97,926	98,815	97,778	91,926	98,512	95,259	95,111	89,630	97,926
	ReLU	Обучающая	93,989	98,664	99,072	**99,221**	99,221	95,176	98,369	99,369	98,219	98,367	95,139
		Тестовая	93,333	97,926	98,815	**99,259**	98,667	94,815	98,667	98,667	97,630	97,333	95,111
AdaGrad	Sigmoid	Обучающая	77,180	77,403	77,996	75,028	53,210	53,210	53,210	53,210	53,210	53,210	53,210
		Тестовая	77,037	77,630	78,370	74,963	53,185	53,185	53,185	53,185	53,185	53,185	53,185
	TanH	Обучающая	93,061	99,295	**99,703**	99,481	99,629	99,555	97,922	**99,481**	99,369	99,852	96,623
		Тестовая	92,296	98,519	**99,259**	98,815	99,111	98,963	96,889	**99,259**	98,815	98,963	96,148
	ReLU	Обучающая	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210
		Тестовая	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185

Продолжение таблицы 3.22

Алгоритм обучения	**Функция активации**	**Выборка**	**Точность классификации скважин для различного числа скрытых слоёв ИНС, %**										
			2	3	4	5	6	7	8	9	10	15	20
AdaDelta	Sigmoid	Обучающая	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210	53,210
		Тестовая	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185	53,185
	TanH	Обучающая	75,102	73,692	74,286	73,803	74,768	74,805	75,176	75,436	76,030	75,584	79,109
		Тестовая	74,963	73,481	73,630	73,185	74,667	74,222	74,667	74,815	75,407	75,259	79,407
	ReLU	Обучающая	73,692	76,809	76,660	77,403	77,477	77,662	78,553	80,148	80,297	81,076	82,004
		Тестовая	73,481	74,593	76,296	77,630	77,778	77,630	78,519	80,000	80,296	80,889	81,630
Adam	Sigmoid	Обучающая	89,907	92,801	96,994	94,731	92,319	94,063	85,974	87,384	83,265	53,210	53,210
		Тестовая	89,778	91,704	95,407	94,222	92,000	93,481	85,481	87,852	83,259	53,185	53,185
	TanH	Обучающая	99,147	98,924	**99,666**	99,406	99,295	95,250	96,735	93,061	85,566	46,790	46,790
		Тестовая	98,370	98,370	**98,666**	98,666	98,666	95,111	96,296	93,185	85,333	46,815	46,815
	ReLU	Обучающая	97,885	91,169	99,332	99,406	99,518	**99,518**	98,479	99,109	97,440	54,286	53,210
		Тестовая	97,333	90,074	98,222	98,519	98,519	**98,963**	97,778	98,074	97,037	54,222	53,185

Из таблицы 3.22 следует, что для набора параметров № 4 получены высокие результаты для алгоритмов обучения SGD, AdaGrad и Adam, а также для функций активации TanH и ReLU. Из сравнения результатов из таблицы 3.20 с этими результатами следует, что модели ИНС, которые показали невысокие результаты для набора параметров № 4 Месторождения 1, часто дают невысокие результаты и для Месторождения 2, однако есть и исключения.

Также был проведен ряд дополнительных исследований. Во-первых, обученные модели ИНС, полученные для фонда скважин Месторождении 1, были исследованы на тестовой выборке Месторождения 2. При этом наряду с обученными моделями ИНС для фонда Месторождения 1, которые дали самую высокую точность на тестовой выборке этого месторождения, были исследованы также модели ИНС, которые дали на его тестовых данных невысокую точность. Наоборот, обученные модели ИНС для фонда Месторождения 2 были исследованы на тестовой выборке Месторождения 1. Во-вторых, была апробирована идея о том, что слияние обучающих выборок с фондов скважин двух месторождений может дать более высокие результаты по точности классификации скважин. Несмотря на то, что для фонда Месторождения 2 имеется бóльшая обучающая выборка, чем для фонда Месторождения 1, эта идея была апробирована и для Месторождения 2 в надежде, что даже незначительное увеличение размера обучающей выборки может улучшить результат. Результаты

этих исследований представлены в таблице 3.23. Столбцы 1, 2 и 3 (слева направо) этой таблицы характеризуют параметры моделей ИНС для Месторождения 1.

Таблица 3.23 – Результаты классификации скважин на Месторождениях 1 и 2 с использованием данных и моделей ИНС другого месторождения

Алгоритм обучения	**Функция активации**	**Число скрытых слоёв**	**Точность классификации скважин, %**					
			Месторождение 1			**Месторождение 2**		
			Обучение на своих данных	Обучение на объединённой выборке	Использованы обученные модели Мест. 2	Обучение на своих данных	Обучение на объединённой выборке	Использованы обученные модели Мест. 1
SGD	Sigmoid	*3*	76,877	79,558	61,086	80,277	81,630	61,187
	TanH	*10*	94,836	**99,334**	54,102	95,111	**98,963**	58,293
	ReLU	*6*	95,572	**99,533**	72,931	98,667	**99,259**	62,857
AdaGrad	Sigmoid	*2*	76,753	75,691	61,197	77,037	77,481	59,740
	TanH	*6*	95,203	**99,933**	52,239	99,111	98,963	63,636
	ReLU	*4*	58,426	60,773	60,754	53,185	53,185	53,210
AdaDelta	Sigmoid	*4*	58,426	60,773	60,754	53,185	53,185	53,210
	TanH	*8*	71,834	71,242	42,749	74,667	74,667	67,199
	ReLU	*10*	66,053	77,901	56,785	80,296	80,296	53,210
Adam	Sigmoid	*6*	87,824	95,301	61,552	92,000	96,593	55,547
	TanH	*3*	97,418	**99,800**	58,714	98,370	**98,667**	62,894
	ReLU	*4*	97,173	**99,448**	63,282	98,222	**98,963**	62,894

Из таблицы 3.23 видим, что модели ИНС, обученные на объединённой выборке, дают более высокий результат по точности классификации скважин, чем в случае их обучения на данных с одного месторождения. Однако если на обучающей выборке своего месторождения уже были получены высокие результаты по точности классификации, то выигрыша от объединения выборок может не быть. Модели ИНС, обученные на данных с одного месторождения, при тестировании их на данных другого месторождения дают более низкую точность классификации скважин, чем при тестировании их на данных с собственного месторождения. По-видимому, признаки, которыми обладают скважины-кандидаты для ГТМ фонда скважин конкретного месторождения, являются довольно индивидуальными. Это не позволяет применять обученные на одном месторождении модели ИНС при решении задач классификации скважин на фондах других месторождений без дообучения на данных этих фондов скважин.

Все результаты по точности классификации скважин, приемлемые для практического использования, сравнивались по критерию Краскела – Уоллиса:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1), R_i = \sum_{j=1}^{n_i} R_i^j, \quad (3.23)$$

где H – статистика критерия Краскела – Уоллиса (гипотеза о наличии сдвига в параметрах положения двух сравниваемых выборок); N – объём выборки; n_i *– элемент i-ой выборки;* R_i *– ранг i-ой выборки;* R_i^j *– ранг j-ого элемента в i-ой выборки в полученном вариационном ряду.*

Показано, что отличия в результатах при выборочных флуктуациях не являются существенными при уровне значимости, равном 0,05.

В экспериментах по определению точности *Задачи 3* на Месторождении 1 и Месторождении 2 использовались архитектуры и гиперпараметры ИНС, а также обучающие и тестовые наборы данных, давшие наилучшие результаты по точности классификации скважин (см. рисунок 3.7 и таблицы 3.22 и 3.23). В итоге для Месторождения 1 с помощью модели ИНС с 9 скрытыми слоями, для алгоритма обучения Adam и функции активации TanH в случае набора параметров № 2 получена наилучшая точность классификации типов ГТМ при тестировании – 85,598 %. Для Месторождения 2 при использовании в экспериментах лучших моделей ИНС для набора параметров № 4 наилучшая точность составила 85,946 %. Она получена для модели ИНС с тремя скрытыми слоями и функции активации TanH, алгоритм обучения SGD. Небольшой объём выборки ГТМ по каждому из месторождений, особенно для каждого конкретного типа ГТМ, не позволил получить более высокие результаты.

Знание типа ГТМ позволяет специалисту сузить область рассмотрения потенциально возможных ГТМ для каждой скважины-кандидата и тем самым окажет ему поддержку в принятии решения по выбору ГТМ.

3.3.4. Методы и алгоритмы кластерного анализа при решении задачи выбора скважин-кандидатов для ГТМ

В БД некоторых нефтегазодобывающих предприятий хранятся данные о проведённых работах, включая ГТМ, на фонде скважин и весьма небольшие объемы данных по результатам мониторинга отдельных технологических и геологических параметров скважин и продуктивных пластов. В этих случаях практически невозможно сформировать представительную обучающую выборку для моделей ИНС. Однако методы кластерного анализа позволяют вести анализ

данных в подобных ситуациях и получать определенные результаты, облегчающие принятие решений специалистами при управлении фондом скважин. Предположение, лежащее в основе наших исследований возможностей методов кластерного анализа, заключается в том, что скважины-кандидаты для ГТМ обладают особыми качествами, которые выражаются в схожем сочетании значений характеризующих их параметров. Поэтому при применении некоторого метода деления множества (или подмножества в случае куста скважин) скважин фонда на определённое число кластеров появляется возможность выделить и кластер скважин-кандидатов для проведения ГТМ.

Исследовалась эффективность при решении *Задачи 2* следующих методов кластеризации.

- *Иерархический метод* (реализующий его алгоритм Single-link) [126, 127]

$$R_{min}(U,V) = \min \rho(u,v), \tag{3.24}$$

где $u \in U$, $v \in V$; *R(U,V) – расстояние между кластерами U и V по одиночной связи;* $\rho(u,v)$ *– функция расстояния между объектами.*

При этом существует расстояние R(W,S) от образовавшегося кластера W до любого другого S:

$$R(W,S) = 0{,}5 \times R(U,S) + 0{,}5 \times R(V,S) + 0{,}5 \times R(U,V) - 0{,}5 \times |R(U,S) - R(V,S)|,$$

где $W = U \cup V$.

- *Плотностной метод (*алгоритм DB-SCAN) [128, 129]

$$U(x,\varepsilon) = \{y \in V : \rho(x,y) \leq \varepsilon\}, \tag{3.25}$$

где x – рассматриваемый объект; ε – окрестность объекта x; y – плотнодостижимый объект из объекта x; ρ(x,y) – симметричная функция расстояния.

- *Итеративный метод* (алгоритм K-means) [26]

$$V = \min\left(\sum_{i=1}^{k} \sum_{x \in S_i} (x-\mu)^2\right), \tag{3.26}$$

где V – искомое значение, k – число кластеров, S_i *– полученные кластеры,* μ_i *– центр масс всех векторов из кластера* S_i,.

Здесь в качестве метрики используется Эвклидово расстояние:

$$\rho(x,y) = \sqrt{\sum_{p=1}^{n} (x_p - y_p)^2}, \tag{3.27}$$

где $x, y \in R^n$

- *Статистический метод* (алгоритм EM) [130]:

$$p(X) = \prod_{i=1}^{N}\left(\sum_{j=1}^{K} \omega_j \varphi\left(x_i, \theta_j\right)\right); \quad (3.28)$$

$$[\omega, \theta] = \max\left[\sum_{i=1}^{N} \log\left(\sum_{j=1}^{K} \omega_j \varphi\left(x_i, \theta_j\right)\right)\right], \quad (3.29)$$

где $p(X)$ *– вероятностная модель множества точек* X*;* ω_j *– вероятность того, что случайная точка относится к кластеру* j*;* $\varphi(x_i, \theta_j)$ *– вероятность встретить точку* x *в кластере* j*;* $[\omega, \theta]$ *– максимизируемая метрика.*

Эти методы и реализующие их алгоритмы (3.24) – (3.29) исследовались на технологических и геологических данных с учётом проведенных ГТМ из БД по 10 кустам фонда скважин нефтяного Месторождения 3 (116 скважин). В качестве исходных данных для кластерного анализа использовались значения технологических и одного геологического параметров каждого комплекса «скважина-пласт». При этом используются по отдельности дебит нефти $Q_н$, дебит газа $Q_г$, дебит жидкости (газового конденсата) $Q_ж$, сочетание дебитов нефти $Q_н$, газа $Q_г$ и жидкости $Q_ж$, сочетание пластового давления $P_{пл}$ с дебитами жидкости и газа и т. п.

В качестве множества скважин для последующего кластерного анализа было решено использовать скважины, принадлежащие одному кусту. При этом параметры для анализа брались за месяц до проведения какого-либо ГТМ. Следует отметить, что основанием для остановки скважин на проведение того или иного ГТМ являлось исключительно решение геологов-экспертов предприятия. Судя по информации из БД предприятия, скважина либо была кандидатом для проведения ГТМ, либо нет, то есть с точки зрения кластерного анализа эксперты выделяли две группы скважин (бинарная классификация) без промежуточных вариантов. Выбор скважин экспертами считался нами образцом (эталоном), с которым впоследствии происходило сравнение полученных результатов анализа. Кроме такого варианта задания числа кластеров при делении скважин на кусте нами предложено провести эксперименты в случае разбиения множества скважин на три группы. Аргумент при введении третьей группы следующий. Третий кластер отвечает за тот самый промежуточный вариант скважин, на которых по

каким-либо причинам ГТМ проведены не были (скважины резерва): возможно, не хватило бюджета предприятия или отсутствовали свободные бригады КРС. Скважины для анализируемого месяца последнего кластера могут быть в последующих месяцах кандидатами при формировании нового списка скважин-кандидатов для ГТМ.

Перед проведением кластерного анализа данные из БД Месторождения 3 для каждой скважины собирались в виде специального файла. При этом для исключения ошибочных, пропущенных и искажённых значений данных, извлечённые из БД данные проходили предварительную подготовку с использованием подходов и адаптивных алгоритмов из п. 3.1.2, 3.2.1 и 3.3.1.

В экспериментах при оценке результатов работы алгоритмов вычислялась метрика Precision (формула 3.21). В случае деления скважин на три кластера «средний кластер» вариативно может относиться как к скважине для ГТМ, так к скважинам без ГТМ.

В таблице 3.24 в качестве примера приведены результаты кластеризации для куста № 1. Этот куст содержит 7 скважин, в том числе 4 эталонных скважины, на которых через месяц после даты анализа данных экспертами на скважинах куста были проведены ГТМ. Для каждого алгоритма есть области в виде двух или трёх столбцов, что соответствует числу кластеров, на которые разбивалось исходное множество скважин куста. Для каждого параметра выделено две строки: в первой находятся скважины, на которых были проведены ГТМ (эталон), а во второй – скважины без ГТМ.

Таблица 3.24 – Результаты кластеризации для скважин куста № 1

Алгоритмы	**K-means**					**DBSCAN**					**EM**					**Single-link**				
Кластеры	**2**		**3**			**2**		**3**			**2**		**3**			**2**		**3**		
$Q_ж$	**4**	0	**2**	**2**	0	**4**	0	**4**	0	0	**4**	0	**4**	0	0	**4**	0	**4**	0	0
	2	1	0	2	1	2	1	1	1	1	2	1	1	1	1	2	1	1	1	1
$Q_г$	**1**	**3**	0	**2**	**2**	**4**	0	**3**	0	**1**	**4**	0	**3**	0	**1**	**4**	0	**3**	0	**1**
	2	1	2	0	1	1	2	1	2	0	1	2	1	2	0	1	2	1	2	0
$Q_н$	**3**	**1**	**1**	**2**	**1**	**4**	0	**3**	**1**	0	0	**4**	**3**	0	**1**	**4**	0	**3**	**1**	0
	1	2	0	1	2	2	1	2	0	1	2	1	1	1	1	2	1	2	0	1
$Q_ж + Q_г$	**2**	**2**	**2**	**2**	0	**4**	0	**4**	0	0	**4**	0	**4**	0	0	**4**	0	**4**	0	0
	2	1	0	2	1	2	1	0	2	1	1	2	0	2	1	2	1	0	2	1
$Q_ж + Q_г + P_{пл}$	**3**	**1**	**3**	**1**	0	**4**	1	**4**	1	0	**4**	0	**4**	0	0	**4**	2	**4**	0	0
	1	2	0	2	1	0	2	0	1	1	1	2	0	2	1	0	1	0	2	1

Число эталонных скважин показано жирным шрифтом. В строке «Кластеры» показано число (2 или 3) кластеров на выходе программ, реализующих алгоритмы кластеризации. Содержимое кластера для каждого из параметров или их комбинации, полученное в результате выполнения того или иного алгоритма, представлено в виде столбца. Аналогичные таблицы с результатами кластеризации получены для других девяти кустов скважин. Отметим, что у ряда кустов скважин результаты кластеризации при использовании комбинаций тех же параметров, что в случае куста 1 несколько хуже результатов, приведённых в таблице 3.24, но близки к ним при других комбинациях параметров. Эти же результаты, но с оценками в виде метрики Precision (формула 3.21) в процентах представлены в таблице 3.25. В каждой ячейке таблицы верхнее число – это значение метрики Precision для класса скважин-кандидатов (важно для специалиста), а нижнее – для скважин, на которых ГТМ проводить не следует.

Таблица 3.25 – Результаты кластеризации для скважин куста № 1

Алгоритмы	**K-means**		**DBSCAN**		**EM**		**Single-link**	
Кластеры	**2**	**3**	**2**	**3**	**2**	**3**	**2**	**3**
$Q_ж$	67 100	80 100	67 100	80 100	67 100	80 100	66 100	80 100
$Q_г$	33 25	50 33	80 100	75 67	80 100	75 67	80 100	67 67
$Q_н$	75 67	75 75	67 100	60 50	0 20	75 67	67 100	67 50
$Q_ж + Q_г$	50 33	100 100	67 100	100 100	80 100	100 100	67 100	100 100
$Q_ж + Q_г + P_{пл}$	75 67	100 100	100 100	100 100	80 100	100 100	100 100	100 100

Основываясь на анализе результатов по всем десяти кустам, можно сделать следующие выводы [131, 132].

1. Разбиение скважин куста на три группы в большом числе случаев даёт кластеры, содержащие разные группы объектов. То есть скважины-кандидаты для ГТМ и скважины, на которых ГТМ проводить не следует, не смешиваются между собой, за счёт чего точность в отдельных случаях равна 100 %. При разбиении скважин на 2 кластера подобная ситуация имеет место только на кусте № 1 при использовании алгоритмов DBSCAN и Single-link и в случае набора $Q_ж + Q_г + P_{пл}$.

2. Комбинации двух и трёх параметров скважин и продуктивных пластов (дебиты газа и жидкости, пластовое давление с дебитами жидкости и газа и т. п.) позволяют получить более высокие результаты по точности классификации, чем это происходит при кластеризации по одному какому-либо параметру.

3. Статистический метод кластерного анализа EM не для всех кустов скважин демонстрирует высокие показатели точности, в отличие от более стабильной работы плотностного и иерархического методов. Для первого и третьего кустов этот метод даёт высокие результаты, а для остальных кустов точность его значительно ниже. Итеративный метод с использованием алгоритма K-means в сравнении с другими методами показывает худшие результаты.

Таким образом, при решении слабоформализуемой *Задачи 2* наибольшую точность классификации показали алгоритмы DBSCAN и Single-link, причём лучшие результаты для них получены для комбинаций параметров. Ввод дополнительного кластера для «средних» скважин позволяет получить разделимые кластеры и в ряде случаев достичь точность классификации, равную 100 %.Однако результаты кластерного анализа при вводе дополнительного кластера не позволяют понять, к какому из первых двух кластеров принадлежит та или иная скважина третьего кластера, то есть полученные кластеры ещё должен интерпретировать сам специалист-геолог, что вновь может привести к ошибкам из-за субъективного фактора.

В отличие от результатов, полученных с помощью моделей ИНС, методы кластерного анализа дают менее стабильные по точности результаты и требуют от специалиста более высоких трудозатрат при интерпретации результатов кластеризации. Последнее означает, что модели ИНС обеспечивают более высокую степень автоматизации при принятии решений специалистом. Однако положительным аспектом использования методов кластерного анализа является тот факт, что они не требуют формирования обучающей выборки и поэтому могут применяться на месторождениях с небольшим объёмом накопленных данных.

3.4. Оценки эффективности ГТМ, проводимых на фонде скважин

Планируемое проведение ГТМ на скважине рассматривается как инвестиционный проект, то есть после проведения ГТМ скважина должна обеспечивать прибыль предприятию. Как уже было сказано в разделе 1.2.1, для оценки экономической эффективности часто применяют индекс доходности PI и индекс КН. При этом учитывается дебит нефти $Q_н$, получаемый из скважины после ГТМ. Их вычисление производится по следующим формулам [11]:

$$KH = \sum_{i=1}^{n}(k_i h_i) \tag{3.30}$$

$$Q_н = Q_п \times [1 - wc] \times \rho;\ Q_п = \alpha \frac{KH(P_{пл} - P_з)}{B_0 \mu_0 \left(\ln\left(\frac{R_e}{R_w}\right) - 0{,}75 + S\right)}, \tag{3.31}$$

где k_i *– проницаемость пропластка;* h_i *– мощность пропластка;* α *– пересчётный коэффициент (зависит от используемых единиц);* $P_{пл}$ *– пластовое давление;* $P_з$ *– забойное давление;* R_e *– радиус дренирования;* R_w *– радиус скважины;* B_0 *– объёмный коэффициент нефти;* μ_0 *– вязкость пласта; S – скин-фактор; n – число пропластков;* Q_n *– приток к скважине; wc – обводнённость пласта; ρ – плотность нефти.*

$$PI = \frac{\sum_{t=1}^{T} \frac{CF(t)}{(1+r)^t}}{З_{ГТМ}(0) + КВ(0)}, \tag{3.32}$$

где CF(t) – денежный поток за расчётный шаг; r – эффективная ставка (годовая ставка, проценты); $З_{ГТМ}(0)$ *– первоначальные затраты на ГТМ за период t; КВ(0) – первоначальные капитальные вложения за период t; t – порядковый номер шага расчёта; T – продолжительность расчётного периода.*

Денежный поток за расчетный шаг вычисляется по формуле:

$$CF(t) = CF_{in}(t) - CF_{out}(t),$$

где CF_{in} *– приток денег за расчётный период;* CF_{of} *– отток за расчётный период.*

Причем имеем:

$$CF_{in} = \sum\left(П_{товар}(t) + Л(t) + A(t)\right),$$

где $П_{товар}(t)$ *– чистая прибыль от реализации товарной продукции после ГТМ; Л(t) – доход от ликвидации оборудования за вычетом расходов по его ликвидации; A(t) – амортизационные отчисления на оборудование, учитывается, если при выполнении ГТМ происходит установка нового оборудования.*

В свою очередь, $П_{товар}(t)$ вычисляется:

$$П_{товар}(t) = \sum_{t=1}^{n_{продукции}}\left(\bar{V}_{товар(t)}(t) \times У_{тп(i)}\right),$$

где $n_{продукции}$ *– количество видов производимой продукции;* $\bar{V}_{\text{товар}(t)}(t)$ *– ожидаемый объём добычи товарной продукции* i*-ого вида в расчётном периоде с учётом вероятности остановки скважины после ГТМ;* $У_{тп(i)}$ *– показатель удельной чистой прибыли от реализации единицы продукции* i*-ого вида.*

Ожидаемый объём добычи вычисляется по формуле:

$$\bar{V}_{\text{товар}(t)}(t) = V_{\text{товар}(t)}(t)(1-p)^i,$$

где $V_{\text{товар}(t)}(t)$ *– объём добычи товарной продукции* i*-го вида в расчётном периоде;* p *– вероятность остановки скважины после ГТМ.*

При этом для p используется выражение :

$$p = \frac{\sum_{t=1}^{N} p(t)}{N},$$

где $p(t)$ *– вероятность остановки скважины в определённом периоде после ГТМ;* N *– количество лет, за который производится оценка экономической эффективности.*

В свою очередь, p(t) вычисляется следующим образом:

$$p(t) = \frac{N_{\text{ост}}(t)}{N_{\text{ГТМ}}(t)},$$

где $N_{ост}$ *– количество остановленных скважин после ГТМ;* $N_{ГТМ}$ *– общее количество скважин, на которых было произведено ГТМ.*

Амортизационные отчисления учитываются при закупке нового оборудования:

$$A(t) = \text{БС} \times a/100,$$

где БС – базовая стоимость; a *– норма амортизации:* $a = \frac{1}{n} \times 100\ \%$*, где* n *– количество периодов.*

Отток за расчётный период:

$$CF_{out} = \sum_{t=1}^{T} (\text{З}_{\text{ГТМ}}(t) + H_{\text{ИМ}}(t) + \text{КВ}(t)),$$

где $З_{ГТМ}(t)$ *– планируемые затраты на ГТМ скважины;* $Н_{им}$ *– налог на имещество; КВ(t) – затраты на новое оборудование.*

При этом $\text{Н}_{\text{им}}$ вычисляется по формуле:

$$H_{\text{ИМ}} = \text{ОФ}_{\text{ср}} \frac{\text{СтавкаН}_{\text{им}}}{100\ \%},$$

где $ОФ_{ср}$ *– средняя стоимость установленного оборудования за отчётный период;* $СтавкаН_{им}$ *– ставка налога на имущество, %.*

В свою очередь, средняя стоимость установленного оборудования за отчётный период вычисляется так:

$$ОФ_{ср} = \frac{ОФ_{нач} + \sum_{i=2}^{n} ОФ_i + ОФ_{след}}{n+1},$$

где $ОФ_{нач}$ *– остаточная стоимость имущества на начало налогового периода;* $ОФ_i$ *– остаточная стоимость имущества на начало каждого месяца;* $ОФ_{след}$ *– остаточная стоимость имущества на начало следующего месяца;* $n_м$ *– количество месяцев в налоговом периоде;*

Оценка технологической эффективности проведения ГТМ на скважине предполагает использование различных критериев:

- Количественный показатель извлекаемости запасов. Определяется как потенциальное количество УВС (газа, конденсата, нефти), извлекаемого скважиной из дренируемой зоны до наступления одного из граничных условий эксплуатации (достижение заданного коэффициента извлечения, обводнение продуктивного горизонта, консервация/ликвидация газового промысла по технико-экономическим показателям). В частности, используется нормированный показатель эффективной толщины пласта:

$$K_{пр} = \frac{(kh)_{скв}}{(kh)_{мест}}, \quad (3.33)$$

где k *– проницаемость пласта;* h *– эффективная мощность пласта;* $(kh)_{скв}$ *– среднее значение по скважине;* $(kh)_{мест}$ *– среднее значение по месторождению.*

- Показатель продуктивности скважины. Для нефтяных скважин и газовых скважин показатель продуктивности определяется как отношение дебита скважины на депрессию. Для газовых скважин показатель продуктивности оценивается по величине линейного коэффициента фильтрационного сопротивления:

$$(P_{пл}^2 - P_{заб}^2) = aQ_г + bQ_г^2, \quad (3.34)$$

где Рпл – пластовое давление; Рзаб – давление забойное; $Q_г$ *– дебит газа скважины;* a *– линейный коэффициент фильтрационного сопротивления;* b *– квадратичный коэффициент фильтрационного сопротивления.*

- Показатель обводнённости. Определяется как доля обводнённых дренируемых интервалов перфорации к необводнённым; показатель

рассчитывается отдельно по каждому объекту разработки и однотипным конструкциям скважин (длина скважины, диаметр НКТ, форма ствола скважины, способ вскрытия продуктивного пласта).

- Показатель отклонения значений параметров (дебитов) от технологического режима. Рассчитывается как процент отклонения от максимально допустимого значения по технологическому режиму (в долях единицы):

$$К_{min} = \frac{Q_{min} - Q_{скв}}{Q_{min}}, \tag{3.35}$$

где Q_{min} *– минимально допустимый дебит;* $Q_{скв}$ *– фактический дебит скважины.*

- Равномерность освоения месторождения. Депрессия в зоне дренирования скважины от среднего пластового давления по объекту разработки. Рассчитывается следующим образом:

$$N_{P_{пл.скв.}} = \left|\frac{P_{пл.ср.} - P_{пл.скв.}}{P_{пл.ср.}}\right|, \tag{3.36}$$

где $N_{P_{пл.скв.}}$ *– нормированный показатель;* $P_{пл.скв.}$*– среднее пластовое давление по объекту разработки;* $P_{пл.ср.}$ *– пластовое давление в зоне дренирования скважины (куста).*

Формулы (3.30) – (3.36) и приведённые выше выражения для вычисления ряда параметров в этих формулах стали основой для разработанных алгоритмов. Программная реализация этих алгоритмов в виде веб-сервисов описана далее в разделе 4. В свою очередь, веб-сервисы позволят специалистам реализовать БП оценки экономической и технологической эффективности ГТМ.

3.5. Алгоритм оптимизации плана-графика работы бригад КРС

Поскольку задача составления оптимального плана-графика работы бригад КРС на месторождении является многокритериальной, поэтому решить её точными математическими методами, например методами линейного программирования, довольно сложно. Поэтому для оптимизации плана-графика работы бригад КРС предлагается использовать генетический алгоритм, позволяющий искать глобальные экстремумы для сложных функций многих переменных [133]. Пусть расписание работ на месяц – это отдельная особь в популяции. Её генотип представлен хромосомами – множеством видов работ,

выбранных, оценённых, согласованных и утверждённых на предыдущих этапах управления ГТМ. То есть изначальное число хромосом упорядочено и задаётся числом работ, которые будут проводиться на эксплуатируемом месторождении в том или ином месяце. Каждая хромосома соответствует виду работ на конкретной скважине, выполняемые с определённой длительностью (справочная величина). Хромосома определяется следующими компонентами, поиск оптимальных значений которых происходит с помощью генетического алгоритма: 1. Дата начала работы. Кодируется вещественным числом в формате «дата-время» (без месяца и года), инициализируется изначально случайными значениями в пределах планируемых дат; 2. Бригада, выполняющая работу. Задаётся натуральным числом в диапазоне [0, N] в зависимости от того, какое количество бригад имеет предприятие. Каждой бригаде ставится в соответствие конкретное число из этого диапазона.

Было решено использовать популяцию из 110 особей. В ходе селекции используется принцип элитизма, то есть при очередной оценке приспособленности популяции всегда 10 наилучших особей остаются. Оценка популяции производится путём сортировки особей по их приспособленности. За приспособленность особи отвечает множество различных критериев, которые вычисляются после её формирования. Особь получает штрафные баллы за следующие недостатки в расписании, которые ущемляют интересы бригад КРС, правила проведения работ, а также увеличивают время простоя скважин:

- *Временной промежуток планирования.* Планирование происходит на месяц вперёд. За каждый выход каких-либо работ за пределы месяца начисляется по **10** штрафных баллов.
- *Наличие совместимых видов работ.* В плане-графике необходим учёт возможности проведения нескольких видов работ одновременно, если они совместимы; если на скважине запланированы совместимые работы, но они не были поставлены в одно время, то начисляется **10** штрафных баллов. В плане-графике не должны проводиться одновременно несовместимые виды работ на

одной скважине; если обнаружены подобного рода работы, то за каждый такой случай начисляется **50** штрафных баллов.

- *Соответствие расписанию бригад.* В плане-графике не допускается задействование бригад в различных местах одновременно. У каждой бригады есть расписание, в рамках которого она работает; бригада специализируется на определённом перечне видов работ (ГТМ, ремонт, исследования и прочее). Если работа назначается бригаде, не специализирующейся на этом виде работ, либо в нерабочее время, то начисляется **30** штрафных баллов.
- *Время между работами для конкретной бригады КРС.* Данное время вычисляется на основании расстояния, которое необходимо преодолеть бригаде с возможной погрешностью в 10 %. За наличие таких недочётов особь получает по **20** штрафных баллов за каждый недочёт.
- *Дефектные гены.* В ходе выполнения алгоритма могут быть получены значения генов, для которых не существует соответствующего фенотипа в реальности. За это особь получает **100** штрафных баллов.

Все критерии объединяются в единую функцию приспособленности, которая позволяет регулировать процесс оптимизации, сравнивать полученные особи (расписания) между собой и, в конечном счёте, выбрать наиболее оптимальное по минимальному числу штрафных баллов (в идеале их число должно равняться нулю).

Селекция реализована турнирным отбором. Изначально в массив особей для отбора попадает всё предыдущее поколение, за исключением элитных особей, а затем в нём начинается бинарный турнир, оставляющий половину от всех особей, кроме элитных, (50 штук). Соревнуются особи по своей мере приспособленности, побеждает тот, у кого число штрафных баллов меньше. Особи из популяции выбираются случайно, главное, чтобы они были не одинаковыми.

Скрещивание производится между гомологичными хромосомами. Для целочисленных генов используется оператор одноточечного кроссинговера. При вещественном кодировании – простой арифметический кроссинговер [134]:

$$y_i^1 = \alpha x_i^1 + (1-\alpha)x_i^2$$
$$y_i^2 = \alpha x_i^2 + (1-\alpha)x_i^1, \quad (3.37)$$

где y_i^1 *– значение i-ой компоненты первого потомка;* y_i^2 *– значение i-ой компоненты второго потомка;* x_i^1 *– значение i-ой компоненты мамы;* x_i^2 *– значение i-ой компоненты папы;* α *– случайная константа в диапазоне [0; 1].*

Битовая мутация целочисленных компонент происходит с вероятностью 0,01 %. Для компонент с вещественным типом кодирования добавляется шум в диапазоне 1 % от их значения.

Цикл «оценка – селекция – скрещивание – мутация» выполняется либо до тех пор, пока не будет найдена особь, имеющая 0 штрафных баллов, либо пройдёт заданное число эпох, тогда выбирается лучшее решение.

Предлагаемый алгоритм оптимизации плана-графика работы бригад КРС проверялся на реальных исходных данных о работе бригад КРС для одного из эксплуатируемых месторождений. Он позволил сформировать ежемесячные планы-графики работы шесть раз (в итоге на полгода работы бригад КРС на месторождении), получившие одобрение экспертов предприятия.

3.6. Основные результаты и выводы по разделу

В разделе исследованы модели, методы и алгоритмы обработки и анализа значений технологических и геологических параметров фонда скважин. При этом получен ряд результатов и сделаны следующие выводы.

1. Сформулированы задачи исследования возможностей ряда современных методов ИАД при решении трёх слабоформализуемых задач, возникающих при управлении фондом скважин.

2. Проведен анализ основных типов моделей ИНС применительно к решению слабоформализуемых задач. Показано, что предпочтение следует отдать глубоким ИНС прямого распространения.

3. С учётом специфики исходных технологических и геологических данных по фонду скважин проанализирован ряд методов и подходов к предварительной подготовке (обработке) таких данных для последующего применения к ним методов ИАД. На основе этого анализа и принципов концепции «Big Data»

разработаны оригинальный адаптивный алгоритм и методика предварительной подготовки значений дебитов скважин. Этот алгоритм отличается от известных методов и алгоритмов наличием решающих правил в блоках адаптации к видам пропусков и ошибок в значениях дебитов и позволяет скорректировать ошибки и неполноту исходных данных при формировании обучающей и тестовой выборок для решения слабоформализованной задачи прогнозирования дебитов скважин. Проведённые исследования точности и скорости вычислений трёх методов интерполяции данных позволили выявить пределы их применимости в составе адаптивного алгоритма.

4. Для решения слабоформализуемой задачи прогноза значений дебитов нефти, газа, жидкости (газового конденсата) и дебита воды были разработаны различные модели глубоких ИНС прямого распространения. Проведены масштабные исследования таких моделей ИНС, модели линейной регрессии и четырех статистических методов прогноза дебитов. По результатам исследований сделан вывод о том, что оценка погрешности прогноза значений всех дебитов скважин с помощью моделей ИНС с рациональными архитектурой и гиперпараметрами в два и более раза ниже по сравнению с погрешностью прогноза, даваемой традиционными методами, включая метод экстраполяции по скользящей средней, принятый рядом добывающих предприятий в качестве основного.

5. На основе принципов концепции «Big Data» разработан оригинальный адаптивный алгоритм предварительной подготовки исходных технологических и геологических данных по фонду скважин для решения с помощью методов ИАД двух слабоформализуемых задач классификации скважин-кандидатов для ГТМ и при выборе типа ГТМ для них. Этот алгоритм отличается от вышеописанного адаптивного алгоритма наличием дополнительных блоков и использованием ряда других решающих правил.

6. Для решения двух слабоформализуемых задач классификации при выборе скважин-кандидатов для ГТМ и выборе типа ГТМ для таких скважин разработаны различные модели глубоких ИНС прямого распространения. Масштабные

исследования на подготовленных с помощью адаптивного алгоритма исходных данных по фондам скважин двух месторождений показали, что наилучшие результаты по точности (99 % и выше) выбора скважин-кандидатов для ГТМ достигаются при подаче на вход моделей ИНС с рациональными архитектурами и гиперпараметрами значений полного набора всех измеряемых технологических и геологических параметров комплекса «скважина – пласт» или значений набора всех технологических параметров скважины. Наилучшая точность классификации типов ГТМ для скважин-кандидатов составила по каждому из фондов месторождений около 86 % на тестовых выборках.

7. Результаты исследований основных методов кластерного анализа при решении слабоформализуемой задачи классификации скважин-кандидатов для ГТМ позволили установить, что наиболее высокие значения точности классификации дают иерархический и плотностной методы. При этом для анализа следует брать сочетания значений геологических и технологических параметров. Предложенная идея введения дополнительного кластера «средних» скважин позволила получить разделимые три кластера скважин и в ряде случаев достичь точности классификации 100 %. Однако методы кластерного анализа дают менее стабильные по точности результаты классификации, чем модели ИНС. Более того, при введении дополнительного кластера «средних» скважин от специалиста требуются дополнительные трудозатраты при интерпретации результатов кластеризации.

8. Приведены выражения, наиболее часто используемые при оценке экономической и технологической эффективности ГТМ, проводимых на фонде скважин. На их основе разработаны соответствующие алгоритмы.

9. Разработан алгоритм оптимизации плана-графика работы бригад КРС, основанный на генетическом алгоритме. С его помощью сформирован ряд ежемесячных планов-графиков работы бригад КРС, получивших одобрение экспертов предприятия.

4. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ИИС

На предыдущих этапах разработки ИИС была создана её концепция и разработаны и исследованы модели и алгоритмы предварительной обработки и интеллектуального анализа данных. На этапе программной реализации ИИС учитываются все эти результаты.

4.1. Выбор базовых программных средств ИИС

Под базовым ПО будем понимать средства для разработки оригинального ПО интеллектуальной системы и продукты сторонних производителей для управления созданным ПО и базой данных ИИС.

Главным критерием при выборе среды разработки и языков программирования является полнота средств для реализации выбранной сервисно-ориентированной архитектуры ПО и функциональных возможностей ИИС. При этом среда должна позволять быстро изменять ПО вслед за изменениями БП управления фондом скважин и обеспечивать слабую связность компонентов.

Немаловажным критерием при выборе базовых программных средств ИИС и языков программирования является наличие соответствующей квалификации по ним у разработчиков системы.

При выборе языка программирования существует большое число альтернативных вариантов: C++, C#, Java, Python и другие. Альтернативные варианты имеются и при выборе среды разработки. Анализ языков программирования и сред разработки показал, что наиболее соответствующими критериям в данном проекте являются перспективные среда разработки Microsoft Visual Studio 2015–2019 [135] и языки и технология программирования [136]:

- C# 7.0 для реализации логики программ;
- MS CNTK 2.7 для обучения ИНС;
- C# ASP.NET Core, это кроссплатформенная технология от компании Microsoft, предназначенную для создания различного рода веб-приложений вплоть до крупных веб-порталов и веб-сервисов [137].

Для обращения к данным, хранимым в БД, и манипуляции с ними следует использовать современную СУБД. Многие нефтегазодобывающие предприятия имеют архив данных в виде реляционной БД, для работы с которой (выполнение запросов) применяется язык SQL и его диалекты. Для хранения и администрирования реляционных БД существует довольно большое число СУБД: MS SQL Server, Oracle, MySQL, PostgreSQL, SQLite и другие [138, 139].

MS SQL Server – одна из наиболее популярных СУБД в мире, подходящая для реализации проектов любого масштаба [140]. MS SQL Server работает очень быстро, предоставляет шифрование данных, удобна в работе и администрировании. Эта СУБД была выбрана в качестве базового ПО для работы с данными в ИИС. Используются версии СУБД 2012–2019 годов выпуска.

С точки зрения реализации SOA-модели ПО интеллектуальной системы нет ограничений по базовым программным средствам и языкам программирования, а важным аспектом является только соблюдение стандартов сетевых интерфейсов и протокола обмена данными, чтобы максимально уменьшить число разрабатываемых адаптеров для унаследованных ИС предприятия. Поэтому выбор базовых средств при реализации веб-сервисов SOA-модели ПО может быть сделан с учётом особенностей конкретного реализуемого метода или алгоритма, а также квалификации разработчика. Например, реализация веб-сервиса может проходить на платформе Node.JS и на языке программирования JavaScript.

Выбор ESB и BPMS как обязательных компонентов SOA-модели ПО интеллектуальной системы рассмотрим отдельно.

4.2. Выбор системы управления БП

Для выбора BPMS в составе ИИС был проведён анализ наиболее известных и распространённых BPMS [141]. В их число авторы этой работы включили системы, удовлетворяющие следующим критериям.

1. Полнота реализации той или иной нотации для моделирования БП.
2. Возможность компиляции спроектированных диаграмм БП.

3. Число протоколов, форматов данных, возможности взаимодействия с другими системами, наличие готовых адаптеров и возможности их создания.

4. Стоимость системы.

5. Кроссплатформенность.

На основе нашего анализа систем из [141] было решено оставить из них для дальнейшего рассмотрения те системы, которые способны работать на любых платформах и не имеют стоимости, то есть удовлетворяют критериям 4 и 5. Дополнительно к известным на рынке системам из [141] были рассмотрены такие системы и технологии, как Windows Workflow Foundation, Bonita BPM, Bizagi BPM Suit, а также российская разработка ELMA BPM Community [52].Они также удовлетворяют критериям 4 и 5. Оценки всех отобранных нами систем по критериям 1–3 проводились по документации или путем опытной эксплуатации некоторых из них. Результаты отбора приведены в таблице 4.1 по каждому критерию в десятибалльной шкале [142].

Таблица 4.1 – Результаты оценки BPMS по выбранным критериям

Название BPMS	**Критерий**		
	Полнота реализации BPMN языка	**Возможность компиляции БП**	**Наличие встроенных адаптеров**
Activiti Modeler	*10*	*10*	*5*
Bonita BPM	**10**	**10**	**7**
BeePMN	10	0	0
BPMN Web Modeler	10	1	0
bpmn.io	10	0	1
Camunda Modeler	**10**	**10**	**8**
Eclipse BPMN2 Modeler	10	0	1
HEFLO	10	0	0
GenMyModel	8	0	0
IBM BlueWorks Live	10	2	2
WWF	*5*	*10*	*6*
Bizagi BPM Suit	10	2	2
jBPM	**10**	**10**	**8**
jBPMN	*10*	*8*	*5*
LucidChart	7	2	0
MagicDraw	0	0	3
RunaWFE	**10**	**10**	**7**
simpl4	10	7	5
Stages	9	2	2
SYDLE SEED Community	*10*	*8*	*5*
yEd Live	10	0	1
ELMA BPM Community	**10**	**10**	**8**

Жирным отмечены BPMS, которые удовлетворяют всем критериям: Bonita BPM, Camunda Modeler, jBPM, RunaWFE и ELMA BPM. Выбор между ними зависит от владения разработчиком ИИС тем или иным языком

программирования (Java, C#, JavaScript, MVEL и другими), используемым в системе. Подчёркнутым курсивом обозначены системы, в принципе отвечающие всем критериям, но с определёнными условиями (ограничение на максимальное число работающих пользователей, по числу имеющихся встроенных адаптеров и т. д). Остальные BPMS либо не имеют требуемой от них функциональности, либо предоставляют её только на ограниченный период.

Анализ показал, что из пяти BPMS, удовлетворяющих всем критериям, предпочтение в контексте особенностей разработки ИИС следует отдать системе ELMA BPM. Поддерживаемый этой системой язык C# и нотация BPMN позволят справиться с реализацией всей необходимой функциональности ИИС и идеально подходят под квалификацию разработчиков этой системы [142].

4.3. Разработка специализированной сервисной шины

Другим ключевым компонентом в SOA-модели ПО является сервисная шина предприятия. Сегодня наиболее популярными ESB считаются продукты: Oracle Service Bus компании Oracle, WebSphere Message Broker компании IBM, ActiveMatrix Service Bus компании Tibco, WebMethods ESB компании Software AG, Sonic ESB компании Progress Software и JBoss ESB компании Red Hat [143–148]. Проанализируем их более детально.

4.3.1. Современные ESB и интеграционные платформы

В настоящее время в некоторых промышленных предприятиях и в ряде государственных организаций ESB используется как единая точка для связи между собой большого числа ИС и приложений различного класса. Целью при этом является обеспечение гибкости масштабирования и простоты переноса в единое информационное пространство предприятия созданных ранее (унаследованных) ИС [149]. Таким образом, ESB играет ключевую роль в интеграции различных ИС и в создании единого информационного пространства, в принципе, любого предприятия.

Большинство ESB строятся на схожих по функциональности компонентах, и, следовательно, обладают близкими функциональными возможностями. Часть

этих ESB разрабатывается крупными компаниями, другая же часть – сообществом энтузиастов по принципу открытости исходного кода. Любая ESB включает готовые модули для задач интеграции (адаптеры) к популярным ИС классов ERP (англ. Enterprise Resource Planning – ERP) и MES с известными протоколами и форматами данных. Если же необходимого адаптера в ESB нет, то обычно в ней есть встроенный предметно-ориентированный язык программирования (а иногда и несколько языков), позволяющий создавать новые адаптеры с учётом желаемых аспектов взаимодействия со сторонней уникальной ИС. ESB обязательно содержит систему очередей сообщений с поддержкой транзакций и различными моделями обмена сообщениями. Причём поддержка обмена сообщениями во многих ESB имеет две альтернативы спецификаций – REST и SOAP [55].

С помощью ESB должно связываться множество ИС и сервисов предприятия, поэтому актуальной задачей при эксплуатации систем на основе принципов SOA является создание отказоустойчивой конфигурации, подразумевающей установку, обновление и удаление части сервисов без остановки всей корпоративной системы в целом и зависимых модулей.

В процессе эксплуатации ESB важным является показатель производительности, так как шина не должна являться узким местом в SOA ПО предприятия. Поэтому было проведено исследование [150] наиболее известных на тот момент времени ESB (Oracle SOA Suite, Oracle Service Bus, IBM Integration Bus) с точки зрения пропускной способности, среднего времени получения и сохранения данных, среднего времени работы асинхронных сервисов, загрузки процессора в зависимости от количества потоков данных. Вывод, сделанный по результатам анализа, говорит о том, что все эти нагрузочные показатели удовлетворяют даже самым высоким требованиям, а стало быть, для этих продуктов можно сосредотачиваться на других критериях.

Продукты корпоративных гигантов в области разработки ESB, таких как IBM и Oracle, несмотря на популярность и качество, также имеют ряд недостатков [151]: устаревшие графические интерфейсы пользователя начала 2000-х годов, большие объёмы исторического кода, каждая версия системы имеет разные

показатели эффективности (и не всегда последняя – самая лучшая). Также такие шины очень сложны в изучении и администрировании (даже настройка простых функций требует большого количества не интуитивно понятных действий); поддерживают огромное число различных стандартов, что часто излишне, однако, может быть и плюсом в конкретных условиях больших предприятий-заказчиков.

При выборе конкретной сервисной шины или интеграционной платформы предприятия часто используют следующие критерии для их оценки [152].

1. *Коммерческие и технологические*. Это взаимосвязанные критерии, так как определённая цена ESB подразумевает некоторое обоснование в виде объёма её функциональных возможностей. Среди технологических критериев можно выделить полноту поддержки функций маршрутизации сообщений по различным протоколам, способность преобразования различных форматов данных, число встроенных адаптеров, удобные графические средства администрирования и конфигурирования ESB, производительность, поддержка различных протоколов по передаче данных и другие. Обычно эти критерии в окончательной оценке при выборе ESB составляют до 45 %.

2. *Отраслевые.* Сфера деятельности и отрасль предприятия определяет дополнительные критерии, необходимость поддержки определённых отраслевых стандартов и спецификаций. Для интеграционных проектов важна масштабируемость, наличие компетентных специалистов для разработки и сопровождения ESB, наличие документации и успешных примеров внедрения рассматриваемых технологий.

3. *Проектные.* На выбор интеграционной платформы влияют факторы, связанные с особенностями конкретного предприятия, например, совместимость с программными продуктами, уже существующими у неё и у предприятий-партнёров, простота установки и первоначального конфигурирования. Серьёзным отличием систем класса ESB является их цена. Производители включают в стоимость дополнительные адаптеры, поддержку особых протоколов и форматов данных, упрощённые инструменты для администрирования и мониторинга, графические консоли для подключения плагинов, возможность распределения

нагрузки на ESB посредством использования кластера и прочие вещи, нужные не всем. Поэтому если не хочется переплачивать за всё это, то встроенный язык программирования платформы позволит восполнить недостающую желаемую функциональность за счёт работы программистов предприятия [153]. Согласно результатам сравнительного анализа [151] созданные небольшими предприятиями ESB по функциональности часто не хуже программных продуктов крупных корпораций, а стало быть, весьма конкурентоспособны, особенно в секторах определённой направленности, за счёт ценообразования, учёта особенностей отрасли и гибкого поведения на рынке.

На рынке ESB есть бесплатные результаты open-source проектов. Среди таких шин наиболее известны Mule ESB, JBoss, Apache ServiceMix и ряд других [154]. Сравнительный анализ по выделенным критериям показан в таблице 4.2, «+» в столбце означает удовлетворение критерию, «++» - удовлетворение в большей степени (чем остальные рассматриваемые аналоги), «+/-» же значит частичное удовлетворение критерию. Часть результатов анализа взята из [155].

Таблица 4.2 - Характеристики open-source ESB

Критерии \ ESB	**MuleESB**	**ServiceMix**	**OpenESB**	**Petals ESB**
Поддержка функциональности ESB	+	+	+	+
Распространённость на рынке	++	+	+/-	+/-
Сообщество активных разработчиков и поддержка	++	+	+/-	+
Гибкая и легко расширяемая логика построения продуктов	++	+	+/-	+
Поддержка широкого спектра транспортных протоколов и настроек соединения	+	+	+/-	+
Хорошо написанная документация	+	+/-	+/-	+/-
Интеграция с другими open-source проектами	++	++	+/-	+
Поддержка разработки решений через IDE	+	+	++	+

Видно, что по многим критериям бесплатные ESB могут соперничать с коммерческими продуктами.

По сути, каждая из упомянутых выше шин или платформ может покрыть весь необходимый предприятию набор требований к ESB. Однако ни одна платформа не содержит в себе готовых решений для ESB, все они включают в себя только библиотеки для работы с сервисами, протоколы обмена и

преобразования данных, соединения с БД предприятия и другие полезные инструменты, которыми необходимо грамотно воспользоваться для построения на базе платформы требуемой предприятию ESB.

Разработка корпоративных ИС для управления производством, в том числе фондом скважин добывающего предприятия нефтегазовой отрасли и создание на основе концепции SOA единого информационного пространства предприятия имеет ряд особенностей, характерных для этой отрасли.

1. ESB должна пропускать через себя огромные объёмы технологических и геолого-геофизических данных о скважинах и продуктивных пластах месторождения, а также данные служб промыслов на проведение ГТМ на фонде скважин, ежегодные и ежемесячные планы и результаты ГДИС, ГИС и ПГИ в скважинах, технологические данные с АСУ ТП установок подготовки УВС и т. п. Частота поступления различных типов данных варьируется от 15 секунд до месяца. Это означает, что шина должна быть высокопроизводительной.

2. Усложняет разработку единого информационного пространства предприятия необходимость интеграции корпоративной ИС с большим числом унаследованных ИС (число таких ИС в каждом предприятии может достигать десятков). Среди таких ИС крупная система 3D-моделирования резервуаров месторождений, различные ИС обработки и интерпретации геолого-геофизической информации, ERP-система и т. п. [10].

3. ИС предприятия не обязательно должны иметь стандартизированные интерфейсы, поскольку такие системы могли создаваться не в рамках концепции SOA. Поэтому для них должны быть разработаны адаптеры в ESB для преобразования специфических форматов данных и протоколов их передачи.

Перспективность использования ESB как важного компонента концепции SOA в нефтегазовой отрасли подтверждается успешным примером построения корпоративной шины и единого информационного пространства ПАО «Газпром нефть». В данном проекте стояла задача организации взаимодействия основной ИС предприятия с внешними ИС партнёров и реализации унифицированного

подхода к интеграции между всеми ИС предприятия [156]. Имеются и другие положительные примеры использования ESB в этой отрасли [157, 158].

Анализ достоинств и недостатков существующих на рынке ESB и интеграционных платформ, а также перечисленных выше особенностей использования концепции SOA на предприятиях нефтегазовой отрасли указывает на необходимость создания специализированной ESB, учитывающей специфику информационных процессов предприятий этой отрасли. Кроме того, оценка трудозатрат квалифицированной команды разработчиков показала, что выгоднее разрабатывать собственную специализированную ESB, а не покупать полнофункциональную ESB.

4.3.2. Особенности специализированной ESB

В качестве инструментария для разработки специализированной ESB была выбрана платформа .NET [159]. Она имеет мощную библиотеку классов преобразования форматов данных, протоколов взаимодействия между системами, а также несколько технологий для реализации ESB на языке C#. Для разработки специализированной ESB была выбрана технология ASP.NET MVC (англ. Model-View-Controller – Модель-Представление-Контроллер) проекта, она предоставляет каркас для разработки систем с веб-взаимодействием и уже имеет некоторые необходимые функциональные возможности, свойственные ESB, такие как маршрутизация сообщений и преобразование форматов данных между сервисами. Платформа .NET может быть использована через выбранную в п. 4.1 среду разработки MS Visual Studio. Гибкость MVC позволяет изменять структуру ESB в соответствии с существующими требованиями предприятия-заказчика.

Созданная специализированная ESB предоставляет стандартные возможности интеграции программных продуктов (сервисов) и ИС и разработана для ИИС таким образом, чтобы отсутствовала зависимость от конкретных базовых платформ и программных решений. При необходимости можно добавлять шине новую функциональность, поэтому изначально нет необходимости предприятию платить за функциональность ESB, которая не будет востребована.

Чтобы осуществлять связь между СУБД и ESB, необходим посредник (адаптер). Таким посредником в данном случае является технология ADO.NET (основные классы – *Connection, Command, DataReader, DataSet, DataAdapter*). Разные СУБД могут иметь разные типы данных, диалекты языка SQL, строки подключения, поддерживать объектную или реляционную модель данных и т. д. ADO.NET обеспечивает ESB унифицированным интерфейсом для работы с самыми различными СУБД, среди них: MS SQL Server, OLE DB (Access, DB2, MySQL), Oracle, Realm DB, Mongo DB, PostgreSQL, FireBird и другие.

Для ESB важно иметь возможность взаимодействовать с ИС по протоколам передачи данных различного уровня и назначения. Сегодня в ней реализованы сетевые протоколы Ggp, Icmp, Icmp6, Idp, Igmp, IP, IPv4, IPv6, Ipx, ND, Pup, Raw, Spx, SpxII (класс *System.NET.Sockets.Socket*), потоковые (*TcpClient, TcpListener*) и дейтаграммные сокеты (*UdpClient*), Http (*WebClient – WebRequest, WebResponse, HttpWebRequest, HttpWebResponse, HttpListener*), почтовые протоколы (*SmtpClient,* POP3, IMAP), передача файлов (*FtpWebRequest, FtpWebResponse*). Это позволяет поддерживать интеграцию сервисов различного вида (SOAP и RESTful веб-сервисы, proxy).

Для работы со сложными объектами, передаваемыми по сети в обменных форматах, необходимо производить их сериализацию – преобразование в поток байтов. Можно использовать следующие популярные форматы: бинарный, soap, xml, json, xls, xlsx, doc, docx, cvs, txt. Для части форматов предусмотрен свой класс сериализации: *BinaryFormatter, SoapFormatter, XmlSerializer, DataContractJsonSerializer*. Ряд ИС предприятия в рамках интеграционного проекта содержит большие объёмы подобных данных в разных единицах измерения и иногда с разными наименованиями. Поэтому часть сущностей предметной области была преобразована в структуры классов моделей (англ. Model). Для объектов модели была указана дополнительная информация в виде атрибутов, а благодаря механизму привязки моделей можно не задумываться о сериализации данных из различных форматов в соответствующий класс.

Специализированная ESB поддерживает асинхронные методы. Такие методы позволяют оптимизировать производительность системы в целом, они предназначены, в первую очередь, для запросов, которые могут занять продолжительное время (чтение больших объёмов данных из БД, обращение к внешнему сетевому ресурсу, вычислительно сложные задачи).

На запрашиваемый маршрут в ESB можно накладывать различные ограничения, например, за счёт использования синтаксиса регулярных выражений. Также ограничения могут касаться метода HTTP-запроса и конкретных адресов. При этом контроллеру (англ. Controller) доступны все аспекты веб-взаимодействия (англ. *HttpContext*): идентификационные данные пользователя; данные запроса (переданные параметры, куки, данные пользователя); управление ответом (кодировка, заголовки, тело ответа); данные сессии пользователя и данные маршрута.

Безопасность доступа обеспечивается за счёт разграничения пользователей по ролям с необходимыми правами. Безопасность аутентификаций и работы пользователей в сети обеспечивает технология ASP.NET Identity.

При включении в единое информационное пространство предприятия дополнительных ИС, не имеющих стандартизованных интерфейсов, возможна дополнительная разработка новых адаптеров с использованием созданного шаблона.

4.3.3. Исследование эффективности специализированной ESB

По аналогии с экспериментами по определению производительности при нагрузочном тестировании трёх известных ESB, описанных в [150], был проведён ряд собственных экспериментов с созданной специализированной ESB. Для имитационного моделирования использовался сервер со схожими техническими характеристиками, аналогичные ИС, включающие также сервер БД и сервер приложений. Исследования проводились по описанному в [150] алгоритму для определения эффективности ESB. Повышение нагрузки обеспечивалось путём увеличения числа потоков, эмулирующих одновременную работу определённого числа сторонних пользователей, подключённых к шине (например, ИС

предприятия и/или сервисы). Получены зависимости среднего времени работы сервиса для сохранения клиентских данных от количества потоков (рисунок 4.1) и загрузки процессора от количества потоков (рисунок 4.2). На рисунке 4.1 и рисунке 4.2 для сравнения приведены аналогичные результаты из [150] для случаев использования широко известных ESB – Oracle SOA Suite, Oracle Service Bus (OSB) и IBM Integration Bus.

Из рисунка 4.1 и рисунка 4.2 видно, что показатели производительности при нагрузочном тестировании для разработанной специализированной ESB лучше, чем у продуктов компании Oracle (шины Oracle SOA Suite и OSB), однако хуже, чем у шины компании IBM. Это говорит о том, что созданная ESB вполне конкурентоспособна на рынке по критерию производительности [160].

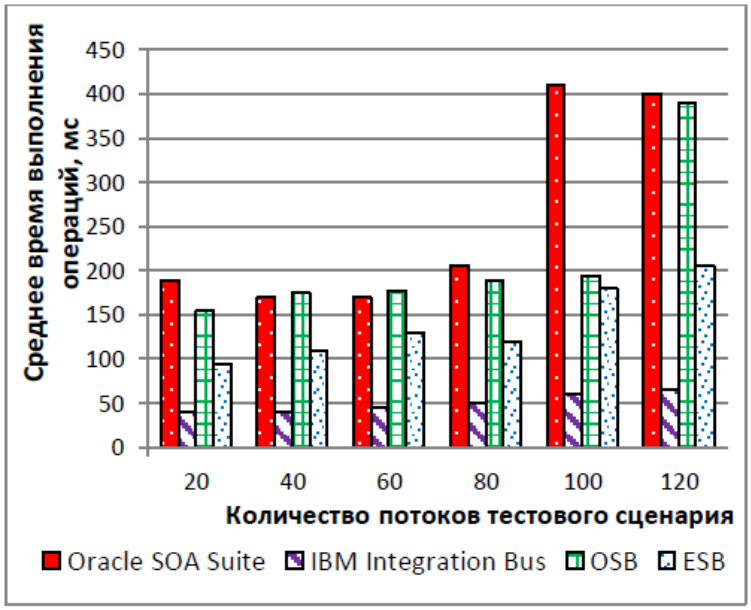

Рисунок 4.1 – Зависимость среднего времени работы сервиса, отвечающего за сохранение клиентских данных, от количества потоков

Специализированная ESB является свободно распространяемой и имеет невысокую стоимость, поэтому по критерию цены она превосходит все известные коммерческие продукты. Отметим, что в отличие от шин указанных компаний специализированная ESB имеет большое число адаптеров, предназначенных для интеграции через неё различных ИС управления производством, применяемых на предприятиях нефтегазовой отрасли. Это делает её привлекательной для предприятий отрасли, создающих у себя единое информационное пространство [161].

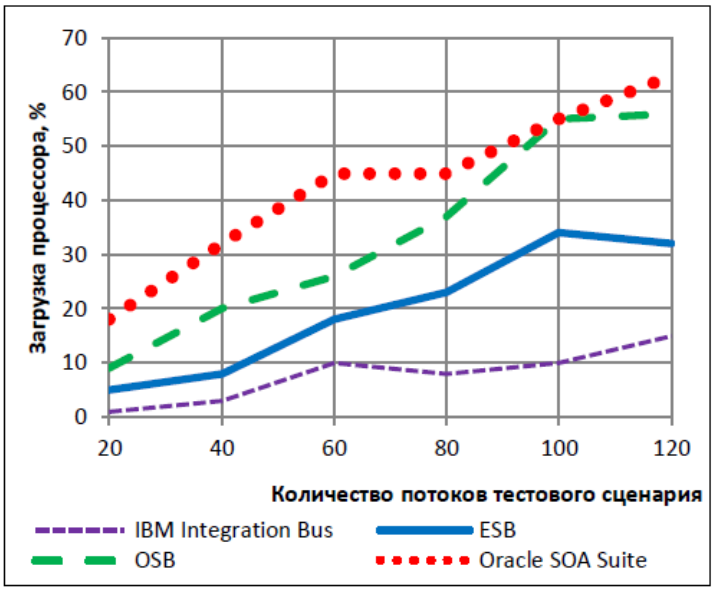

Рисунок 4.2 – Зависимость загрузки процессора от количества потоков

4.4. Реализация бизнес-процессов управления фондом скважин в среде системы ELMA BPM

4.4.1. Веб-сервисы, реализованные в системе ELMA BPM

Для работы специалиста с ИИС используются пользовательские интерфейсы системы ELMA BPM, предоставляющие доступ к функциональности реализуемых веб-сервисов, получивших название «дополнительные сервисы». Веб-сервисы созданы путём адаптации (доработки) ряда шаблонов системы, разработки BPMN-диаграмм и написания дополнительных модулей на языке C#, что позволило считать ELMA BPM модифицированной под особенности БП управления фондом скважин добывающих предприятий нефтегазовой отрасли. При адаптации разработчики ИИС опирались на спроектированные эскизы интерфейсов пользователя ИИС (п. 2.3). Эти интерфейсы дают пользователям доступ ко всем выполняемым в ИИС бизнес-процессам управления фондом скважин в форме запуска и выполнения задач (рисунок 4.3). При адаптации использовался следующий подход. Каждая пользовательская задача на диаграмме в нотации BPMN имеет множество настроек, в частности, задание контекстных данных (объекты и файлы). В соответствии с этими настройками система ELMA BPM автоматически формирует пользовательские интерфейсы, ставя в соответствие определённому типу данных подходящую форму в интерфейсе.

Стрелки, идущие из блока, становятся кнопками, по которым происходит соответствующий переход к следующей задаче. Если шаблон, предлагаемый системой ELMA BPM, не устраивает по каким-либо причинам разработчика ИИС, то можно реализовать свой пользовательский интерфейс с помощью разметки ASP.NET MVC Razor. В рамках создаваемой ИИС сгенерированные пользовательские интерфейсы соответствуют сложности решаемых задач и удовлетворяют принятым эргономическим критериям.

При создании веб-сервисов учитывалось, что с ИИС будут работать пользователи, выполняющие, как показано в п. 2.1., три типа ролей: Администратор, Специалист и Конечный пользователь.

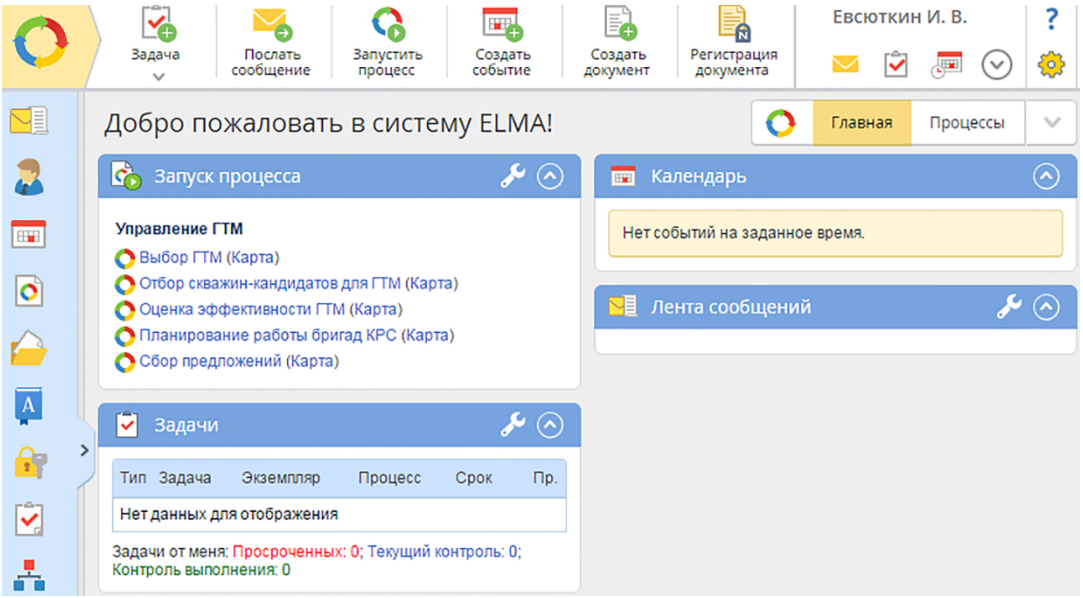

Рисунок 4.3 – Адаптированный под ИИС интерфейс пользователя главной страницы системы ELMA BPM

Поэтому в рамках системы ELMA BPM была создана организационная иерархическая структура служб предприятия, включающая всех возможных пользователей ИИС, привлекаемых для управления фондом скважин. Каждый пользователь или роль обладает определённым набором прав и ответственностью за выполнение конкретных задач, доступа к данным и действий с ними.

Частично данные хранятся в БД, управляемой реляционной СУБД FireBird среды ELMA BPM. Сначала эта БД была создана во встроенном дизайнере среды ELMA BPM в виде логической схемы БД (рисунок 4.4), а после создания физической модели этой БД пользователи ИИС с соответствующей ролью

получили возможность наполнять БД в веб-интерфейсе пользователя. Пример такого наполнения БД приведен на рисунке 4.5. Отметим, что эта БД содержит данные, необходимые исключительно для локального функционирования самой BPM-системы, – это данные об организационной структуре предприятия, элементах BPMN-диаграмм, о правах ролей и их пользовательских настройках в среде ELMA BPM, но эти данные могут использоваться также извне другими системами.

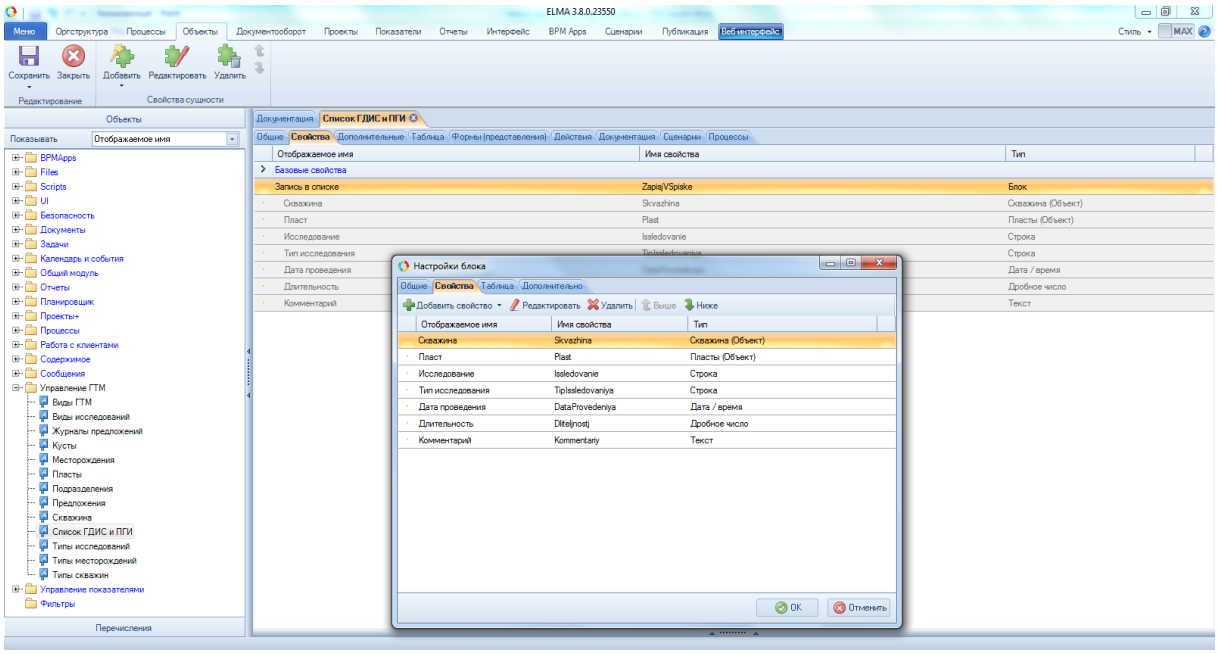

Рисунок 4.4 – Создание логической модели данных в среде ELMA BPM

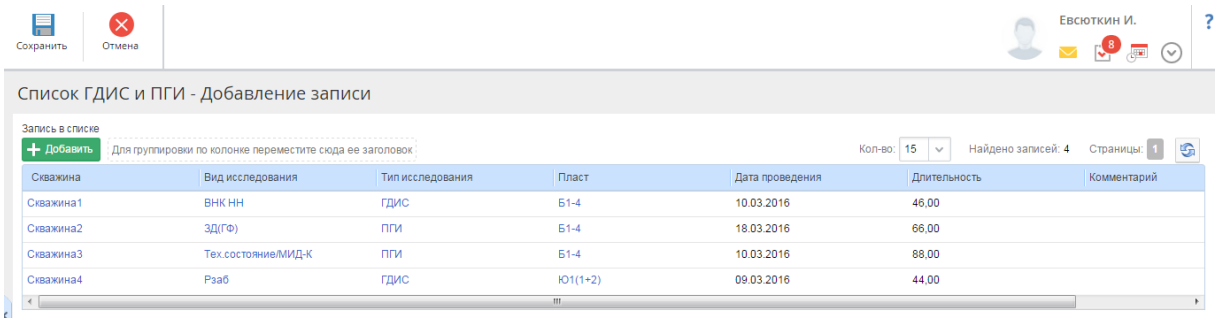

Рисунок 4.5 – Интерфейс пользователя при заполнении справочника исследований в скважинах

Кроме локальной БД в ИИС имеется основная БД, в которой хранятся все паспортные данные скважин и продуктивных пластов, значения технологических и геологических параметров, остановки скважин, данные оборудования, справочные значения, расписание работы бригад КРС, справочники и классификаторы предметной области и т. д. Физическая модель данных основной

БД была получена из концептуальной модели, описанной в п. 2.4, с помощью инструментов Toad Data Modeler для формирования скриптов на языке SQL. Доступ к её данным со стороны веб-сервисов, в том числе системы ELMA BPM осуществляется через СУБД MS SQL Server 2019 (возможны версии 2012, 2014, 2016 и 2017 гг.). Эта СУБД может выполнять запросы и к локальной БД.

Интерфейс пользователя при добавлении нового объекта для исследования скважин показан на рисунке 4.6. Важной особенностью этого интерфейса является то, что все поля преобразуются в форму, соответствующую его типу данных.

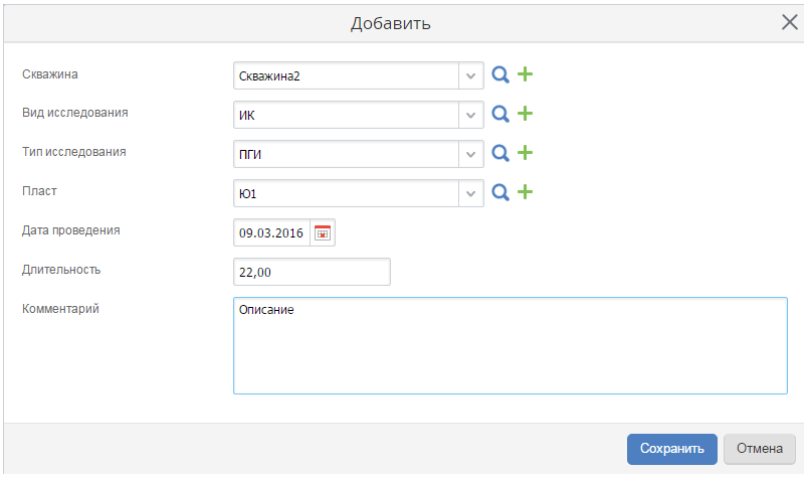

Рисунок 4.6 – Добавление записи в план ГДИС и ПГИ

4.4.2. Формирование журнала на остановку скважин средствами системы ELMA BPM

При реализации *БП «Сбор предложений по проведению мероприятий, ведущих к остановкам скважин»* задействованы не только специалисты-геологи, но и специалисты всех производственных служб, имеющие отношение к фонду скважин. Все они к назначенному сроку формируют предложения на остановку скважин, в том числе для проведения ГТМ. Диаграмма этого БП в нотации BPMN, созданная средствами системы ELMA BPM, показана на рисунке 4.7 [162]. Она размещается на стороне сервера и описывает именно этот отдельный БП. Пользователь с правом запуска такого БП инициирует выполнение цепочки подпроцессов, заложенных в диаграмме. БП выполняется в соответствии с логикой цепочки подпроцессов и силами соответствующих им исполнителей

(специалист-геолог, специалист-технолог, система и т. д.). При этом различные операции BPMS выполняет в соответствии с их типом: если это пользовательская задача, то определённому пользователю становится доступен пользовательский интерфейс задачи, если это запрограммированный сценарий, то система производит необходимые вычисления, обращается к внешнему относительно её веб-сервису через шину ESB или заносит данные в локальную БД.

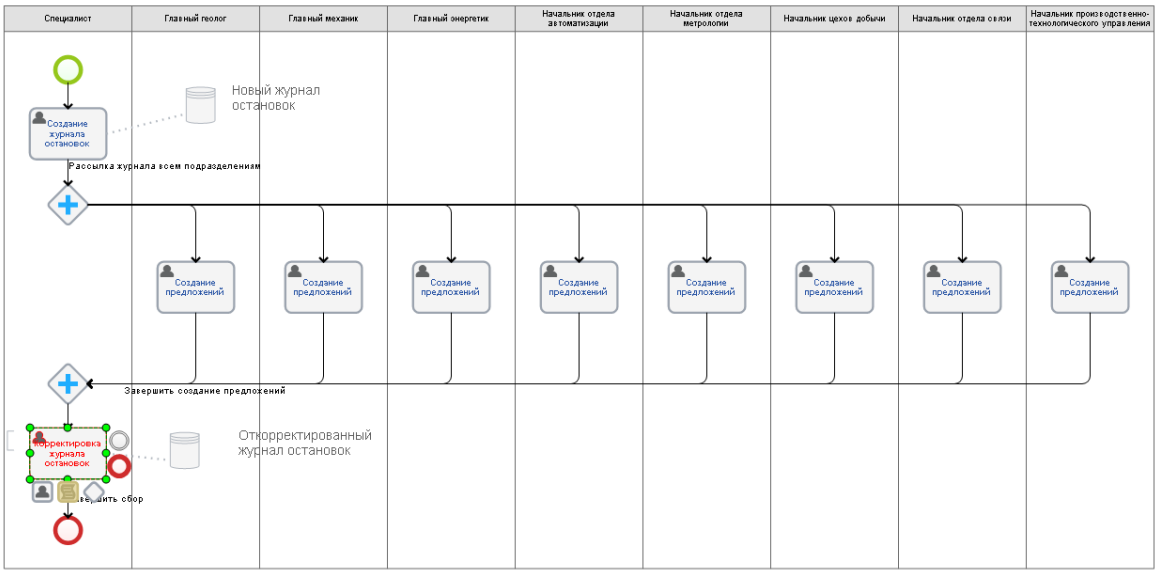

Рисунок 4.7 – Диаграмма БП формирования журнала предложений по остановке скважин в нотации BPMN

В соответствии с этой диаграммой специалист-геолог сначала создаёт новый ежемесячный журнал для сбора предложений, после чего этот журнал становится доступным для редактирования всем заинтересованным службам (рисунок 4.8).

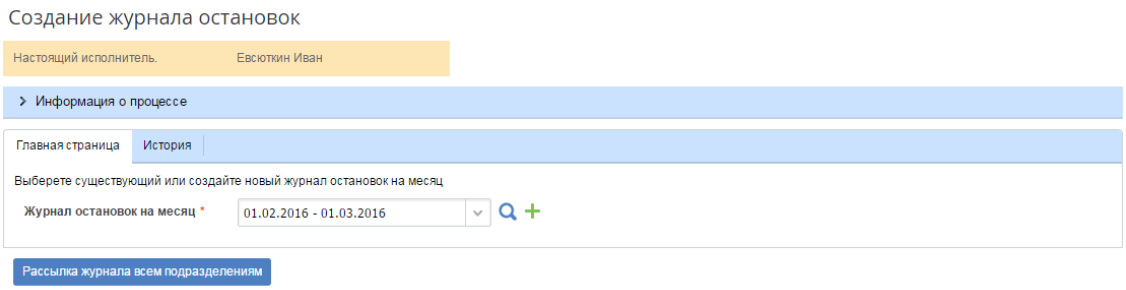

Рисунок 4.8 – Интерфейс пользователя при создании ежемесячного журнала предложений

Получив новый журнал остановок скважин, службы имеют возможность добавлять записи, удалять и редактировать их (рисунок 4.9). Выбор той или иной скважины для остановки на ГТМ, а также ГТМ для неё может производиться

специалистами-геологами на основе их опыта и интуиции. Однако при решении этих двух слабоформализуемых задач предпочтение следует отдать методам ИАД. Это означает, что предварительно в ИИС выбирается веб-сервис, реализующий тот или иной метод ИАД, проводятся расчеты и по их результатам специалистами принимаются решения, которые и заносятся в журнал. Также в список могут заноситься скважины из предыдущих ежемесячных планов, скважины, предварительно выбранные с помощью традиционных методов. Последние запускаются из ИИС, но вычисления по ним ведутся в унаследованных ИС, интегрированных с ИИС.

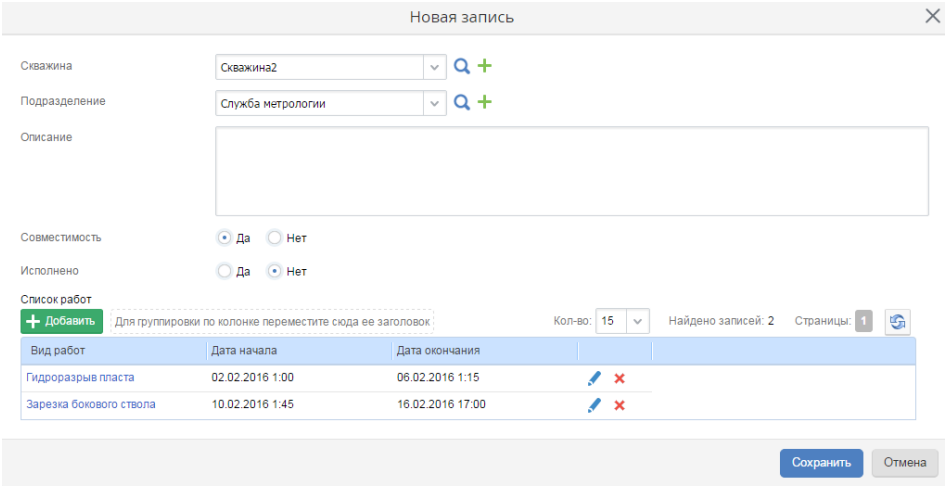

Рисунок 4.9 – Пользовательский интерфейс при создании предложения на остановку скважины службой

После завершения приёма предложений от всех служб журнал возвращается к специалисту-геологу, который принимает окончательное решение и может отредактировать записи в журнале желаемым образом (рисунок 4.10).

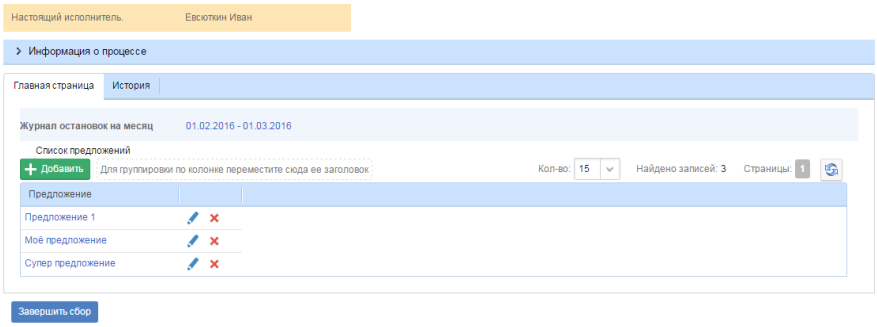

Рисунок 4.10 – Интерфейс пользователя при редактировании окончательной версии журнала предложений

4.4.3. Реализация в системе ELMA BPM бизнес- процесса «Управление и проведение ГТМ»

Рассмотрим реализацию в среде ELMA BPM наиболее сложного, как показано в п. 1.1.3, БП «Управление и проведение ГТМ». Именно этот БП имеет сложную логику, поскольку реализуется как совокупность веб-сервисов на основе методов и алгоритмов решения слабоформализуемых задач при управлении фондом скважин, в том числе на основе методов ИАД (раздел 3). Описание реализации такого БП позволит продемонстрировать все возможности системы ELMA BPM по управлению БП.

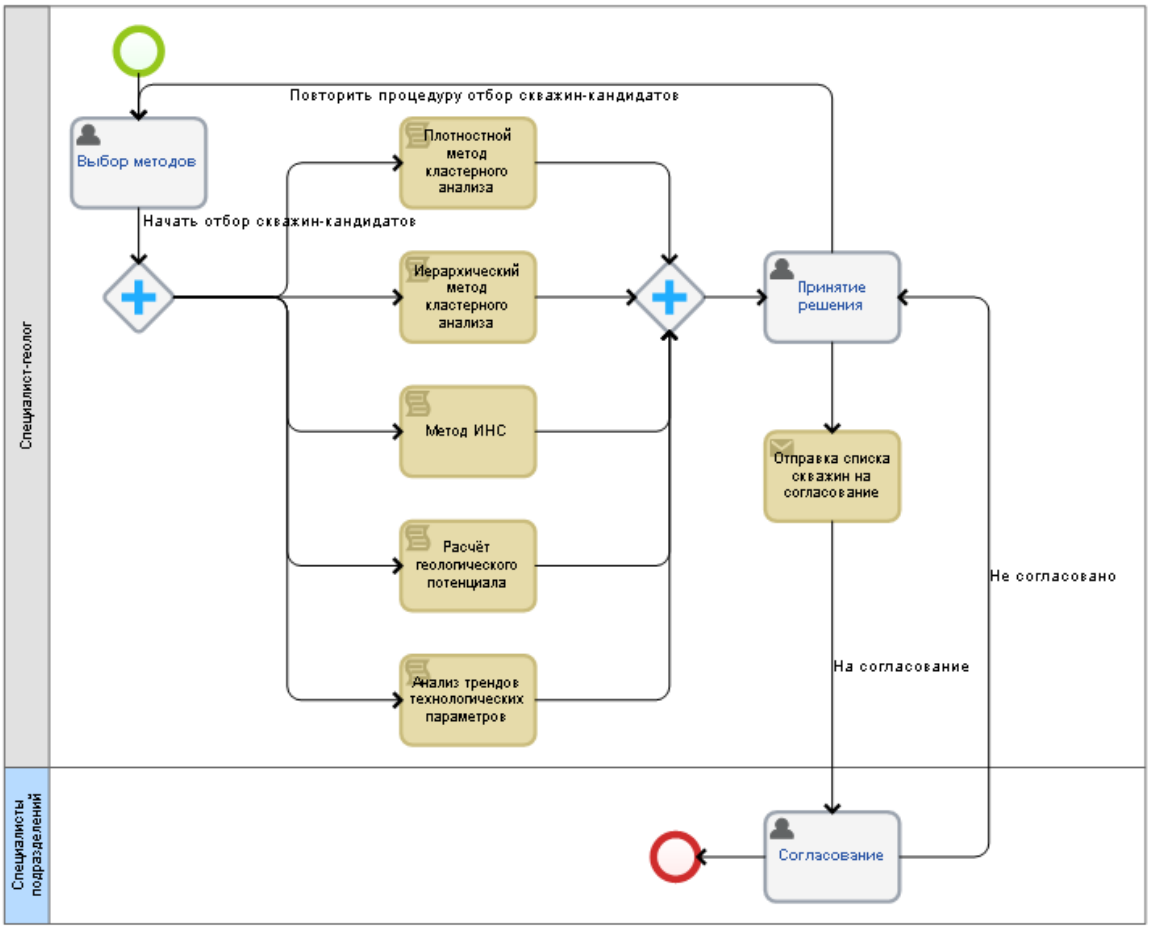

Рисунок 4.11 – Диаграмма подпроцесса отбора скважин-кандидатов для ГТМ в нотации BPMN

BPMN-диаграмма подпроцесса по выбору скважин-кандидатов для ГТМ показана на рисунке 4.11. Аналогичным образом выглядит диаграмма подпроцесса выбора ГТМ с той разницей, что набор методов для этого подпроцесса другой и доступны только те скважины, которые были выбраны как кандидаты при выполнении первого подпроцесса. Каждый реализуемый метод, в

том числе методы ИАД, для удобства оформляется в виде отдельного сценария. На диаграмме показано 5 методов, которые при необходимости можно редактировать: добавлять новые методы, удалять устаревшие или модифицировать существующие. Отметим, что на этом этапе специалист даже не знает, где будет реализовываться метод: в ИИС или в унаследованной ИС, интегрированной с ИИС.

Специалист-геолог при запуске этого подпроцесса производит выбор из перечня требующийся метод (или методы при поочередном их запуске) отбора скважин-кандидатов для ГТМ (задача специалиста-геолога «Выбор методов»), а затем запускает метод, а в итоге процесс собственно выбора скважин-кандидатов. Интерфейс пользователя для этой задачи показан на рисунке 4.12. Каждый метод реализуется либо во внутренней среде ELMA BPM на языке C#, либо в виде отдельного простого сервиса со стандартизированными сетевыми интерфейсами. Традиционные методы анализа для решения этой слабоформализуемой задачи обычно реализованы в унаследованных ИС, интегрированных с ИИС.

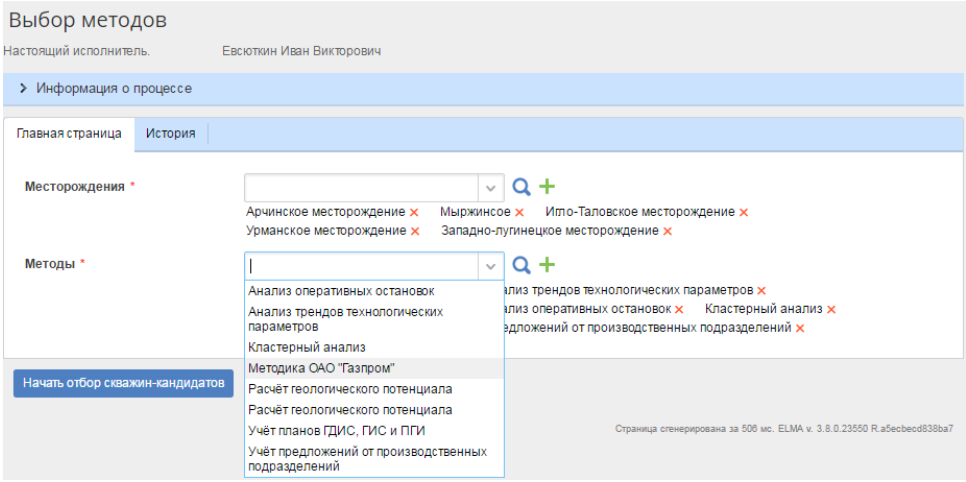

Рисунок 4.12 – Интерфейс пользователя для задания месторождения и методов анализа данных для подпроцесса выбора скважин-кандидатов для ГТМ

После запуска подпроцесса отбора скважин-кандидатов для ГТМ происходит вызов соответствующих веб-сервисов, реализующих конкретный метод или алгоритм. При этом пользователь может видеть на экране состояние прогресса процесса вычислений, так как он может занять длительное время. По завершении специалист может ознакомиться с результатами работы метода (методов): сгруппированные по кустам скважины-кандидаты для ГТМ, их

состояние и причины, почему в итоговый перечень попали те или иные скважины (рисунок 4.13).

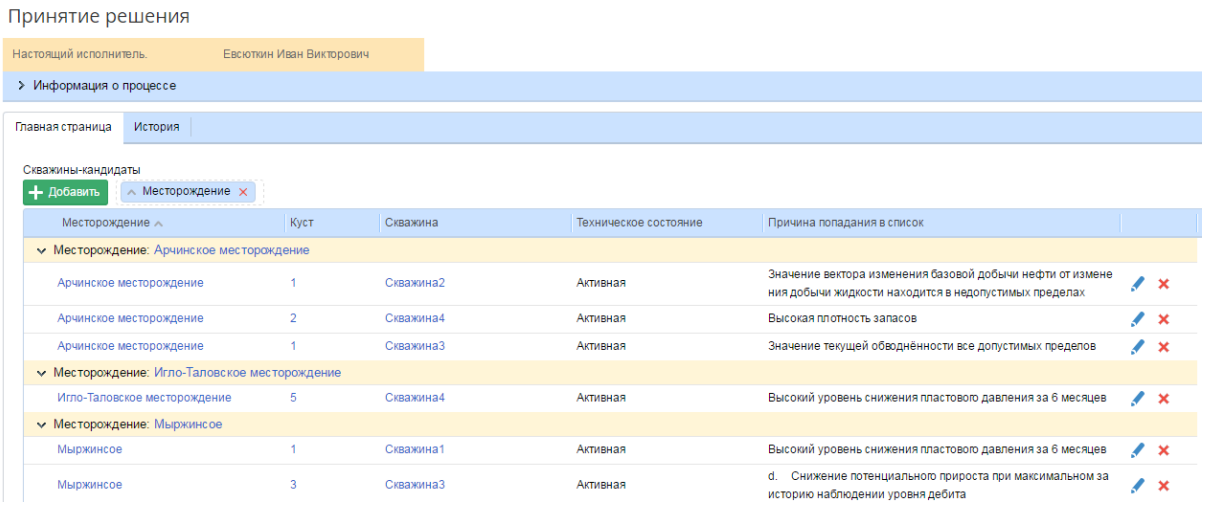

Рисунок 4.13 – Список скважин-кандидатов для ГТМ после применения метода кластерного анализа

Специалист-геолог просматривает полученный перечень скважин-кандидатов для ГТМ, вносит в него свои коррективы и принимает решение, отправить его на дальнейшее согласование всем заинтересованным службам и на утверждение главному геологу (рисунок 4.14), либо он может вернуться на предыдущий шаг для выбора другого метода и проведения новых вычислений.

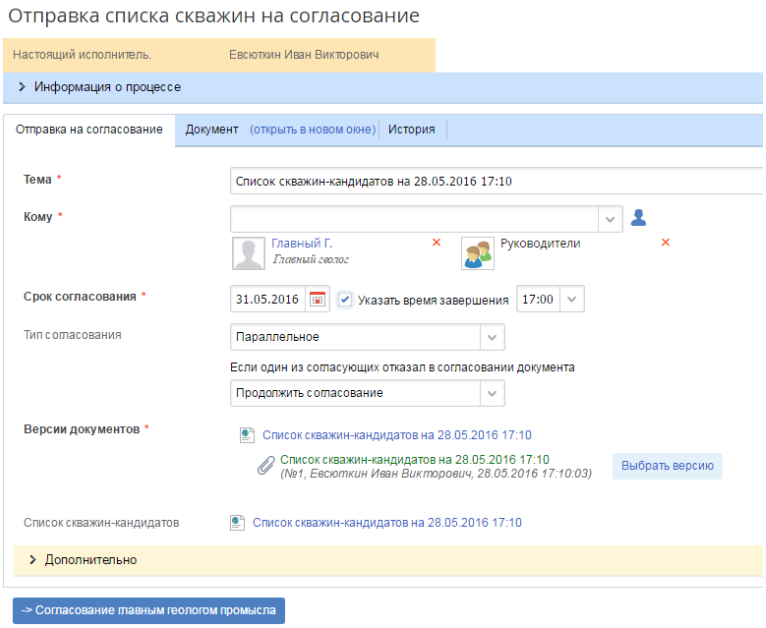

Рисунок 4.14 – Отправка списка скважин-кандидатов для ГТМ на согласование заинтересованным службам и главному геологу

Заинтересованные службы через их ответственных специалистов получают перечень скважин-кандидатов для ГТМ, просматривают его и принимают

решение об отказе или согласовании (рисунок 4.15). Аналогично поступает и главный геолог, который может утвердить или нет представленный список скважин-кандидатов. Своё решение специалисты служб и главный геолог могут дополнить комментарием, в котором будет указана причина отказа. Отказ будет означать, что специалисту-геологу необходимо повторить шаги подпроцесса и учесть пожелания конкретного лица.

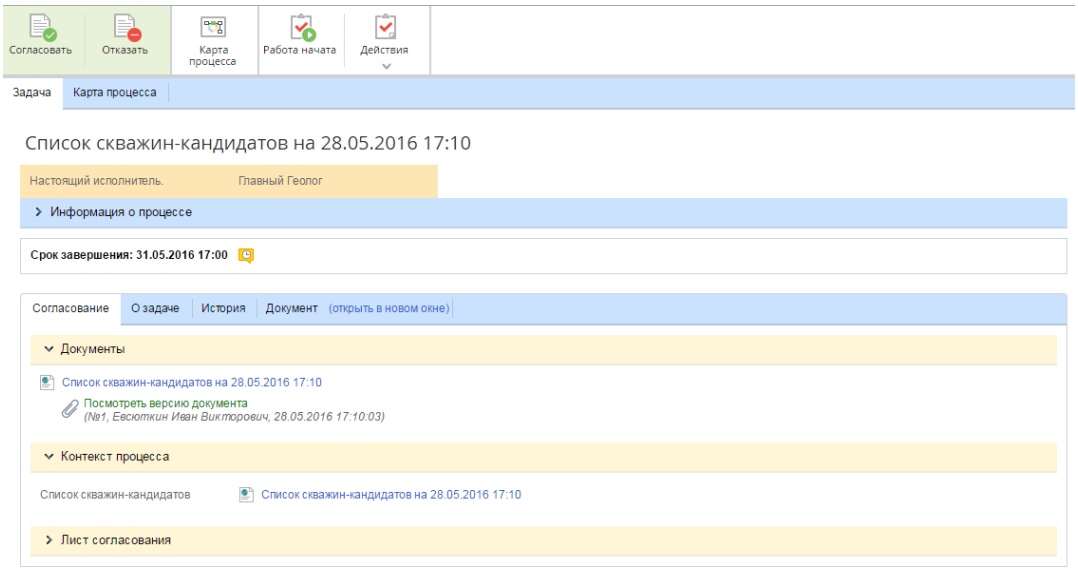

Рисунок 4.15 – Согласование списка скважин-кандидатов для ГТМ службами и главным геологом

Для каждой скважины из согласованного списка, остановка которой предполагает выполнение на ней ГТМ, запускается и выполняется подпроцесс оценки экономической и технологической эффективности ГТМ с помощью соответствующих веб-сервисов. Оценки экономической эффективности ГТМ проводятся по алгоритмам из п. 3.4. При этом специалист может получить результат в виде значения одного из индексов доходности PI или суммарной мощности пласта КН. В качестве технологических показателей эффективности ГТМ специалист может выбрать из меню и запустить расчет следующих из них: количественный показатель извлекаемости запасов, показатель продуктивности скважины, показатель обводнённости скважины, показатель отклонения значений дебитов от технологического режима и равномерность освоения месторождения.

Подпроцесс составления плана-графика работы бригад КРС происходит с использованием системы ELMA BPM в специальном календаре (рисунок 4.16). Календарь позволяет планировать мероприятия (работы) на фонде скважин,

включая ГТМ, на неделю, месяц или на год для каждой бригады КРС отдельно. Список работ формируется автоматически.

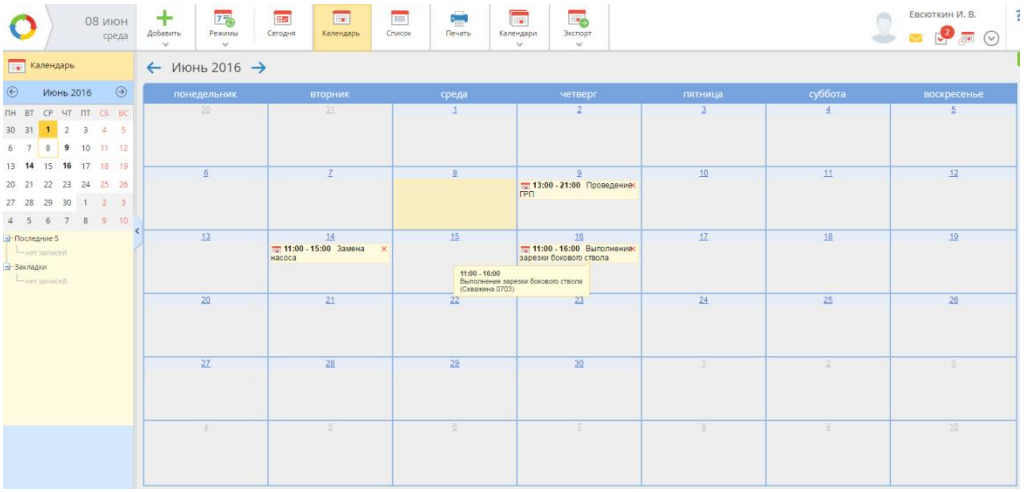

Рисунок 4.16 – Пользовательский интерфейс при планировании работы бригад КРС

Добавление работ вручную для бригады КРС возможно только из согласованного списка скважин-кандидатов для ГТМ и ГТМ для них. Пользовательский интерфейс при добавлении работы показан на рисунке 4.17. При этом задаётся не только вид работы, но и временной промежуток её выполнения и другие желаемые параметры.

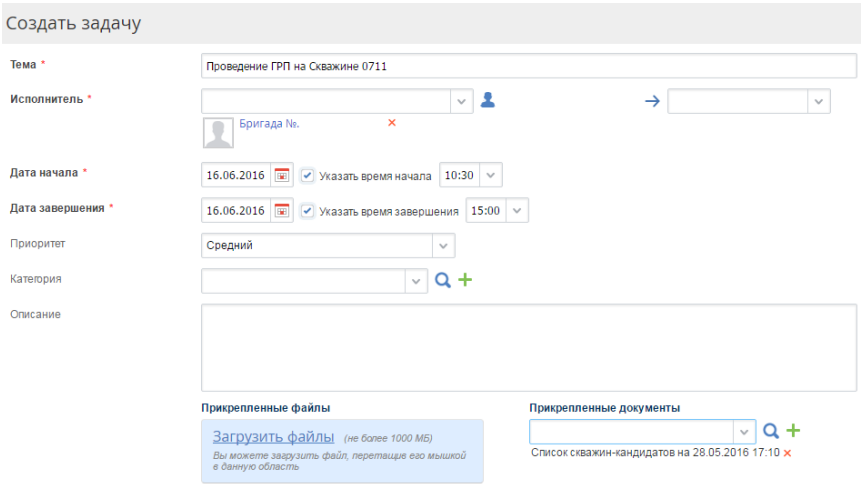

Рисунок 4.17 – Интерфейс пользователя для добавления конкретного вида работ

Построенный план-график работ на месяц может быть оптимизирован как автоматически с использованием разработанного веб-сервиса, реализующего описанный в п. 3.5. алгоритм оптимизации, так и вручную специалистом. Критерием оптимизации считается отсутствие конфликтов между различными

работами на скважине и задействованными бригадами, а также минимальное время простоя скважин, определённых для проведения мероприятий на них.

4.5. Основные разработанные веб-сервисы ИИС

4.5.1. Веб-сервисы для решения сложных вычислительных задач

Помимо дополнительных сервисов, разработанных в среде модифицированной системы ELMA BPM, были созданы другие веб-сервисы, идентифицируемые строкой URL, по адресу которой находится WSDL-файл. Обращение к этим веб-сервисам происходит через элемент BPMN-диаграмм «сценарий» системы ELMA BPM. Если строка URL введена корректно, то в этом элементе формируется класс для вызова всех запрограммированных функций (методов, или алгоритмов), которые реализует веб-сервис. Такие внешние по отношении к системе ELMA BPM веб-сервисы реализуют методы и алгоритмы из раздела 3, подразумевающие сложные и поэтому времяёмкие вычисления. При этом появляется возможность выполнять эти сервисы на ЭВМ высокой мощности и разгрузить компьютеры пользователей ИИС, задействовав для обмена данными специализированную шину ESB. Отметим, что программной реализации подлежали модели, методы и алгоритмы, исследованные в разделе 3 и имеющие значения параметров, позволяющих наиболее точно решать слабоформализуемые задачи управления фондом скважин. Более того, созданное для исследований ПО служило основой при разработке сервисов.

Внешние веб-сервисы были реализованы с использованием описанных в п. 4.1. базовых программных средств (среда разработки Microsoft Visual Studio 2019, платформа .NET, язык C#), а также с использованием протокола SOAP 1.1 и языка описания сервисов WSDL 1.1. Для программной реализации веб-сервисов была выбрана технология WCF (англ. Windows Communication Foundation – WCF) с поддержкой технологий передачи данных AJAX, REST и RSS. Технология WCF создана для разработки приложений, ориентированных на службы. Помимо этого часть внешних веб-сервисов была реализована в среде разработки WebStorm на

платформе Node.JS на языке JavaScript с использованием протоколов http и webSocket.

Разработано шесть веб-сервисов, четыре из которых являются композитными и включают от двух до трёх простых сервисов. Например, композитный веб-сервис «Сбор и актуализация данных» состоит из двух простых: «Сервис предварительной подготовки данных для прогноза дебитов» и «Сервис предварительной подготовки данных для выбора скважин-кандидатов и ГТМ». Первый из них реализует адаптивный алгоритм предварительной подготовки данных при формировании обучающей и тестовой выборок для решения задачи прогноза дебитов скважин (п. 3.2.1.), а второй сервис реализует адаптивный алгоритм предварительной подготовки данных при формировании обучающей и тестовой выборок для решения задачи классификации скважин для ГТМ (п. 3.3.1.). В качестве примера кратко остановимся на особенностях первого из этих простых веб-сервисов. Он включает в себя следующие модули [163]:

- взаимодействия с СУБД MS SQL Server для сбора данных из архивных БД для последующего ИАД и занесения в БД ИИС скорректированных значений параметров;
- визуализации данных по каждому параметру скважины и продуктивного пласта при отображении его в виде временного ряда с указанием событий на графике (остановок скважин, ГТМ и т. п.); визуализация результатов работы по устранению ошибок по двум графикам временных рядов (какие данные были и какие стали после исправления);
- набор модулей, реализующих различные методы, алгоритмы и решающие правила для обнаружения ошибок в данных, восстановления недостающих отсчетов с учётом особенностей обрабатываемого параметра и используемого метода интерполяции и т. п.
- модуль формирования обучающих и тестовых выборок для ИНС.

Отметим, что этот веб-сервис имеет возможность гибкой настройки реализованных алгоритмов и решающих правил. Например, ведется

корректировка границ диапазона изменений параметра, видов интерполяции данных, исправления порядка чисел при ошибках ручного ввода и т. д.

Веб-сервисы, в которых реализована работа с моделями глубоких ИНС, используют нейросетевую библиотеку MS CNTK 2.7 (см. п. 3.1.5.). Кратко рассмотрим механизм взаимодействия веб-сервисов, предназначенных для решения слабоформализуемых задач, с системой ELMA BPM. Со стороны этой системы, являющейся сервером, посылается стандартизированный запрос через ESB, а клиентское приложение (веб-сервис) отвечает на него в зависимости от параметров запроса. Например, для сервиса выбора скважин-кандидатов для ГТМ в теле запроса пересылается: перечень скважин для анализа (для каждой из них указываются «Месторождение», «Куст» и «Номер скважины»), название метода/алгоритма или обученной модели ИНС, с помощью которого будет происходить анализ данных (в том числе обращение к методам в унаследованных ИС, к которым есть доступ из интерфейса системы ELMA BPM (рисунки 4.11, 4.12)), параметры метода/алгоритма или модели ИНС и информацию, о том, где взять из БД значения геологических и технологических параметров на определённый момент времени. Веб-сервис обрабатывает запрос и анализирует полученные значения параметров, периодически отправляя сведения о текущем прогрессе выполнения анализа. По итогам работы он отправляет системе ответ, содержащий информацию по каждой скважине – кандидат (значение 1) или не кандидат (значение 0).

На веб-сервисы для предварительной подготовки данных было получено свидетельство о государственной регистрации программы для ЭВМ «Обработка геолого-технологических данных при формировании обучающей и тестовой выборок для искусственных нейронных сетей» [164].

4.5.2. Архитектура программного обеспечения ИИС

Учитывая все особенности программной реализации ИИС и созданные веб-сервисы, можно провести детализацию предложенной в п. 2.1. концептуальной схемы архитектуры ПО интеллектуальной системы в виде SOA-модели. Как итог разработана детализированная SOA-модель ПО (архитектура ПО на принципах

SOA), приведённая на рисунке 4.18. Штриховкой выделены основные реализованные в ИИС компоненты ПО.

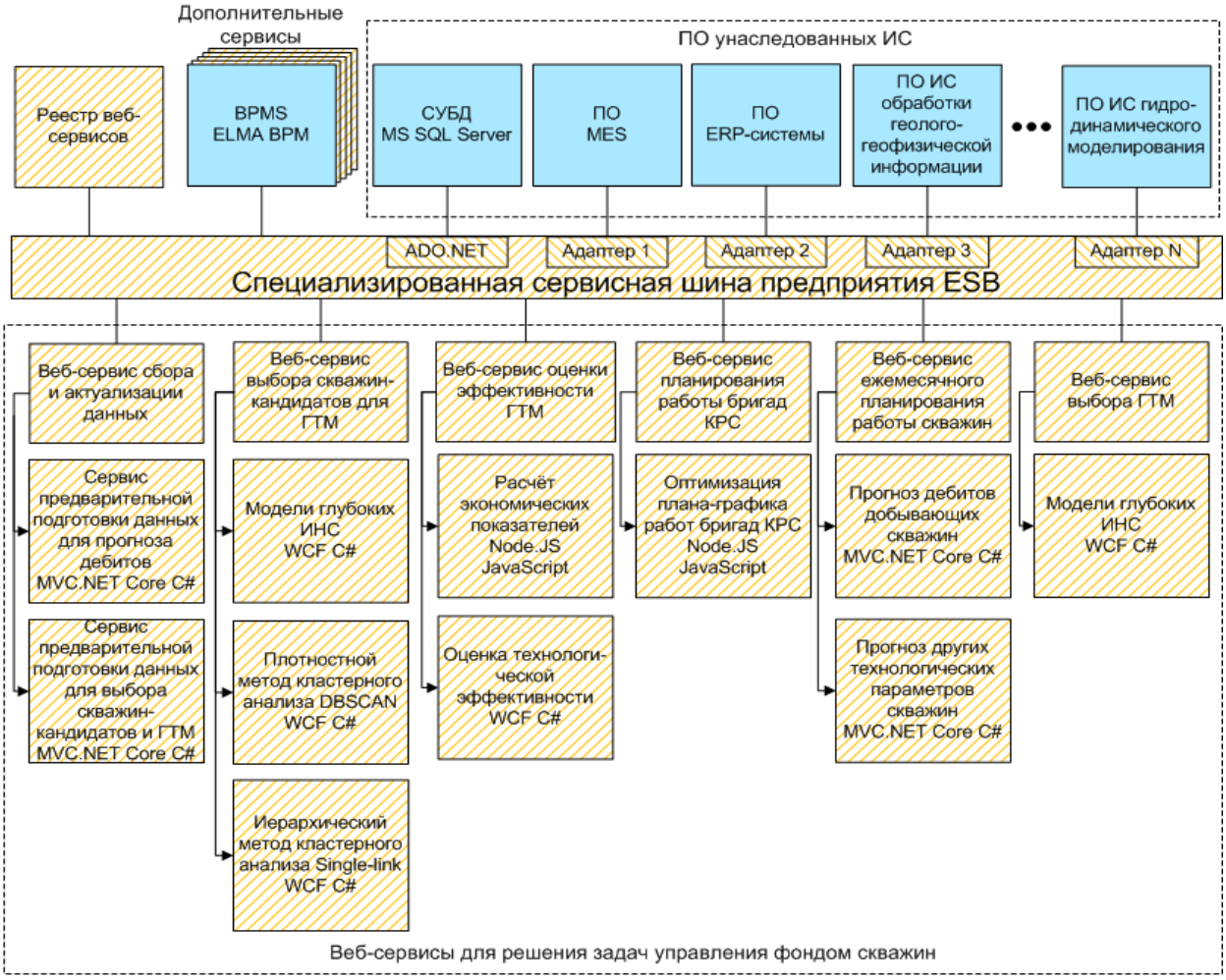

Рисунок 4.18 – SOA-модель программного обеспечения ИИС

Основными составляющими SOA-модели программного обеспечения ИИС являются модифицированная система ELMA BPM, реестр сервисов и созданная специализированная сервисная шина предприятия ESB. Через последнюю ведётся взаимодействие веб-сервисов ИИС между собой и с унаследованными ИС предприятия. Также шина ESB, используя технологию ADO.NET, взаимодействует с СУБД MS SQL Server и позволяет осуществлять с помощью созданных хранимых процедур и представлений на языке SQL сбор из БД данных о скважинах и продуктивных пластах, необходимых для выполнения веб-сервисов.

Веб-сервисы взаимодействия пользователя с ИИС и для реализации ряда БП созданы в системе ELMA BPM с использованием её встроенной среды разработки

на языке C# по технологии ASP.NET MVC. Более подробно основные из них описаны в п. 4.4.1. и п. 4.4.2. На рисунке 4.18 все эти веб-сервисы помечены как «Дополнительные сервисы» и изображены рядом с системой ELMA BPM. Среди них основные:

- веб-сервис для реализации БП «Формирование оперативного плана-графика всех мероприятий на фонде скважин на месяц»;
- веб-сервис для реализации БП «Планирование и проведение ГИС, ГДИС и ПГИ в скважинах»;
- веб-сервис для реализации БП «Сбор предложений по проведению мероприятий, ведущих к остановкам скважин»;
- веб-сервис ведения справочников и классификаторов;
- веб-сервисы запуска и просмотра результатов работы внешних веб-сервисов (п. 4.5.1) и запущенных ими из ИИС унаследованных ИС.

В целом, такие веб-сервисы могут использоваться и другими системами по известному API через систему ELMA BPM.

На рисунке 4.18 ниже сервисной шины ESB показаны шесть внешних по отношению к системе ELMA BPM веб-сервисов (п. 4.5.1.).Четыре из них имеют композитную природу, поскольку включают от двух до трёх простых сервисов. Композитный сервис «Сбор и актуализация данных» используется при реализации БП «Управление и проведение ГТМ» и БП «Ежемесячное планирование технологических режимов работы скважин». В поле каждого простого веб-сервиса указана среда (платформа) разработки и язык программирования, на котором он реализован. Особенности программной реализации веб-сервисов и их вызова на выполнение из среды модифицированной системы ELMA BPM более подробно описаны в п. 4.5.1. и в п. 4.4.3.

В качестве примера на рисунке 4.18 приведено ПО унаследованных ИС предприятия: ИС обработки геолого-геофизической информации, MES, ИС гидродинамического моделирования и ERP-система. Они подключены к шине ESB через соответствующие адаптеры. Три точки на рисунке 4.18 показывают, что таких ИС у предприятия может быть N. Например, к шине ESB могут

подключаться другие унаследованные производственные ИС и АСУ ТП, если они содержат данные, используемые ИИС, или в них реализованы традиционные методы и алгоритмы, применяемые при управлении фондом скважин на предприятии. Такие ИС не имеют стандартизированных интерфейсов, поэтому для них должны быть созданы адаптеры по разработанному шаблону.

С учётом возможности применения при разработке программного обеспечения ИИС различных технологий, языков, платформ и сред разработки становится всё более очевидным преимущество SOA-модели программного обеспечения ИИС (см. раздел 2.), в первую очередь, в части гибкости при взаимодействии программ и легкой интеграции ИИС с разными унаследованными ИС предприятия.

4.6. Основные результаты и выводы по разделу

С учётом концепции ИИС (раздел 2.) и разработанных и исследованных алгоритмов ИАД (раздел 3.) были реализованы все компоненты ПО разрабатываемой ИИС. При этом получены следующие результаты и сделаны основные выводы.

1. На основе проведённого анализа выбраны базовые программные средства для разработки ПО в виде платформы .NET, среды разработки Microsoft Visual Studio 2015–2019 и языка программирования C# 7.0. Предложено использовать также технологии ASP.NET MVC (Core) и WPF, библиотеку для глубокого обучения MS CNTK 2.7 и реляционную СУБД SQL Server.

2. Сформулирован перечень критериев и на его основе выбрана система ELMA BPM для управления БП. Поддерживаемые этой системой язык C# и нотация BPMN позволили модифицировать её путем создания ряда веб-сервисов (названы дополнительными сервисами). Они позволяют пользователям ИИС через средства системы ELMA BPM управлять БП.

3. Проведен анализ используемых сегодня сервисных шин предприятия, который показал, что в случае создания ИИС актуальной является задача разработки недорогой и легко конфигурируемой специализированной шины.

Такая специализированная шина, учитывая специфику нефтегазодобывающих предприятий, была создана с использованием выбранных базовых программных средств. Результаты исследования её эффективности указывают на высокую производительность шины и позволяют сделать вывод о возможности её применения на добывающих предприятиях нефтегазовой отрасли.

4. Создан ряд внешних относительно системы ELMA BPM веб-сервисов, позволяющих решать сложные вычислительные задачи при управлении фондом скважин. При их разработке использовались платформы .NET (среда разработки Microsoft Visual Studio, язык C#) и Node.JS (среда разработки WebStorm, язык JavaScript). Для связи с веб-сервисами, реализованными на платформе .NET, применялись стандартные интерфейсы и протоколы SOA-архитектуры: SOAP 1.1 и WSDL 1.1. Для связи с веб-сервисами, разработанными на платформе Node.JS, использовались технологии AJAX и COMET и протоколы http и webSocket.

Четыре из разработанных веб-сервисов имеют композитную природу, поскольку включают от двух до трёх простых сервисов. Композитный сервис «Сбор и актуализация данных» используется при выполнении двух БП для управления фондом скважин.

При программной реализации моделей глубоких ИНС в соответствующих веб-сервисах использована нейросетевая библиотека CNTK 2.7. ПО, созданное на её основе, позволяет задавать архитектуру и гиперпараметры ИНС и выбрать заранее подготовленные с помощью других веб-сервисов файлы с данными для их обучения и тестирования при решении трёх слабоформализуемых задач.

5. Учитывая все особенности программной реализации ИИС и созданные компоненты предложенной в п. 2.2 концептуальной схемы архитектуры ПО интеллектуальной системы, можно считать, что полностью реализована детализированная SOA-модель ПО этой системы. При этом полнофункциональное ПО удовлетворяет требованиям, предъявляемым к ИИС, и позволяет решать основные задачи управления фондом скважин.

5. ПРИМЕНЕНИЕ ИИС К РЕШЕНИЮ ЗАДАЧ В НЕФТЕГАЗОВОЙ ОТРАСЛИ

Проведена апробация ИИС на предприятиях АО «Газпром добыча Томск» и ООО «Центр нефтегазовых технологий». Ниже показаны её результаты.

5.1. Апробация ИИС в нефтегазодобывающем предприятии АО «Газпром добыча Томск»

5.1.1. Организация единого информационного пространства АО «Газпром добыча Томск»

АО «Газпром добыча Томск» – дочернее предприятие АО «Востокгазпром», входящего в состав ПАО «Газпром». Основной деятельностью предприятия является геологическое изучение, разведка, добыча и подготовка УВС. Для обеспечения непрерывности такого бизнеса предприятие занимается освоением и обустройством новых месторождений УВС. Для реализации своей продукции в виде газа, газового конденсата и нефти АО «Газпром добыча Томск» эксплуатирует большую сеть нефтепроводов и газопроводов для перекачки этой продукции с промыслов в магистральные нефте- и газопроводы. Сегодня предприятие обладает лицензиями на право пользования недрами семи лицензионных участков и эксплуатирует девять месторождений УВС [165].

На основе концепции создания ИИС (раздел 2) и разработанных и описанных в разделе 4 компонентов SOA-модели ПО этой интеллектуальной системы создано единое информационное пространство АО «Газпром добыча Томск». На рисунке 5.1 приведена архитектура ПО этого информационного пространства. Оно включает как унаследованные ИС предприятия, СУБД MS SQL Server 2012, так и компоненты ИИС. К таким компонентам относятся BPMS ELMA BPM с разработанными в её среде дополнительными сервисами, специализированная сервисная шина предприятия ESB и ряд внешних по отношению к системе ELMA BPM веб-сервисов, в том числе композитных.

Веб-сервисы, реализующие бизнес-процессы управления фондом скважин, были созданы на различных платформах и языках программирования и описаны в

разделе 4. В АО «Газпром добыча Томск» были востребованы не все разработанные веб-сервисы. Среди внешних востребованных и изображённых на рисунке 5.1 сервисов: композитный веб-сервис сбора и актуализации данных (разработан на платформе .NET, технология MVC.NET Core, язык программирования C#); веб-сервис прогноза дебитов добывающих скважин (разработан на платформе .NET, технология MVC.NET Core, язык программирования C#, реализует модели глубоких ИНС); три веб-сервиса композитного сервиса по выбору скважин-кандидатов для проведения ГТМ, которые реализованы на платформе .NET по технологии WCF на языке C# (они реализуют модели глубоких ИНС, плотностной DBSCAN и иерархический Single-link алгоритмы кластерного анализа); веб-сервис выбора ГТМ (реализует модели глубоких ИНС, используется технология WCF, язык программирования C#) и веб-сервис оптимизации плана-графика работ бригад КРС (платформа Node.JS, язык программирования JavaScript). При реализации всех моделей глубоких ИНС использовалась библиотека CNTK 2.7.

Рисунок 5.1. – Архитектура ПО единого информационного пространства АО «Газпром добыча Томск»

Программное обеспечение унаследованных ИС предприятия показано на рисунке 5.1 выше специализированной сервисной шины. Число таких ИС в общем случае на предприятии N. Каждая ИС требует адаптер для взаимодействия с ИИС, причём адаптеры были либо изначально включены в некоторые из этих ИС в виде

API, либо были дополнительно разработаны совместно с сотрудниками АО «Газпром добыча Томск» в рамках специализированной шины ESB при создании единого информационного пространства предприятия. Кратко остановимся на основных функциях приведённых на рисунке 5.1 унаследованных ИС.

Корпоративная геоинформационная система управления (КГСУ) «Магистраль – Восток» является основной производственной ИС предприятия и по наличию многих возможностей может быть отнесена к классу MES [56]. Она обеспечивает автоматизацию основных функций управления производством: распределённый мониторинг технологических параметров основного оборудования (включая фонды скважин) промыслов предприятия и мониторинг геологических параметров продуктивных пластов месторождений, сбор данных и составление МЭРов о добыче УВС на фонде скважин, построение различного рода производственных отчетов и прочее. Более того, КГСУ «Магистраль – Восток» интегрирована со всеми АСУ ТП промыслов предприятия и многими ИС для проведения ГДИС, ПГИ, химико-аналитических исследований и т. д., что позволяет собирать данные из таких систем в её БД.

ERP-система Microsoft Dynamics AX 2012 [166] является системой планирования и управления различными ресурсами средних и крупных организаций, повышает эффективность работы и управления изменениями. В АО «Газпром добыча Томск» эта система представляет собой решение по автоматизации БП управления финансами, материальными и человеческими ресурсами, сбытом продукции и т. д.

Программный комплекс (ПК) БАСПРО Оптима (Россия) предназначен для мониторинга разработки месторождений УВС и геологического моделирования [167]. В АО «Газпром добыча Томск» этот ПК используется для решения ряда узкоспециальных задач: проведения детального геолого-промыслового анализа и интерпретации данных ГДИС и ГИС и оценки извлекаемых запасов месторождения по характеристикам вытеснения по нефтяным скважинам. Важной, решаемой с его помощью задачей, ведущей к необходимости интеграции

ПК с ИИС, является задача оценки технологического и финансового эффектов от ГТМ на скважинах и продуктивных пластах.

Программное обеспечение ECLIPSE транснациональной корпорации Шлюмберже – это семейство программных 3D-симуляторов для моделирования многофазных потоков флюидов для всех типов коллекторов месторождений УВС [168].

Программное обеспечение TEMPEST норвежской компании Roxar – интегрированный программный комплекс, позволяющий специалистам создавать детальные гидродинамические модели эксплуатируемых месторождений [169]. Специалист, работая с ИС TEMPEST, может решать практически весь комплекс основных и вспомогательных задач гидродинамического моделирования месторождения, оставаясь при этом в рамках единой модели данных.

По мнению специалистов АО «Газпром добыча Томск» после создания единого информационного пространства предприятия полностью исключено дублирование ввода одних и тех же данных в разные ИС, снизились затраты времени на обнаружение и корректировку ошибок в данных и появилась возможность запуска из ИИС некоторых модулей в составе наследованных ИС.

5.1.2. Результаты опытной эксплуатации ИИС в АО «Газпром добыча Томск»

За более чем 15 лет эксплуатации фондов скважин месторождений АО «Газпром добыча Томск» в БД КГСУ накоплен огромный объём данных, который сегодня составляет несколько терабайт. Структура этой БД содержит около 800 таблиц. БД КГСУ включает все паспортные данные объектов промыслов, справочники и классификаторы, метаданные, запланированные и выполненные работы, включая ГТМ, на фондах скважин, а также архив технологических и геологических параметров всех комплексов «скважина-пласт» каждого из эксплуатируемых месторождений предприятия.

Также в БД имеются сведения о планируемых и проведённых исследованиях в скважинах и продуктивных пластах и всех остановках скважин.

Весь необходимый набор данных для ИИС при её опытной эксплуатации был извлечён из БД КГСУ с помощью веб-сервиса сбора и актуализации данных.

ИИС не только берёт данные из БД, но и с помощью других компонентов этого сервиса ведет предварительную подготовку собранных данных и затем с помощью других сервисов позволяет специалистам предприятия выполнять функции, которых нет в КГСУ. В первую очередь, к ним относятся функции поддержки принятия решений специалистами при решении трёх слабоформализуемых задач методами ИАД, подробно описанных в разделе 3. Если в КГСУ или в других унаследованных ИС имеются программные модули с функциональными возможностями, подобными веб-сервисам ИИС, то внедрение таких сервисов для исключения дублирования функциональности не проводилось. Это видно и из сравнения рисунка 4.18 с рисунком 5.1. Однако специалисты предприятия из ИИС могут вызывать традиционные методы управления фондом скважин, реализованные в КГСУ или в других унаследованных ИС. Так, например, в КГСУ реализована часть традиционных методов по выбору скважин-кандидатов для ГТМ: анализ трендов технологических параметров, статистический анализ частоты остановок скважин, метод геологического потенциала и т. п.

Отметим, что интеллектуальные методы выбора скважин-кандидатов для ГТМ и ГТМ на них, реализованные в ИИС, прекрасно дополнили функциональность КГСУ. Так, методы ИАД, в первую очередь, глубокие ИНС и частично методы кластерного анализа, оказали значительную поддержку в принятии управленческих решений, повысив на промыслах предприятия точность выбора скважин-кандидатов для ГТМ до 97–99 %, а точность выбора типа ГТМ для них до 83 %. Аналогичная точность достигается на предприятии при принятии решений опытным промысловым геологом. Однако решения по фонду скважин при использовании ИИС принимаются в 2–4 раза быстрее.

Суммарное время простоя скважин фонда одного из месторождений предприятия после оптимизации плана-графика работ КРС с помощью соответствующего веб-сервиса оптимизации сократилось в среднем на 12 %, что обеспечило более оптимальный режим работы фонда добывающих скважин этого месторождения.

Стоит отметить, что во время опытной эксплуатации ИИС осуществлялись комплексные испытания ряда её оригинальных компонентов, в первую очередь, модифицированной системы ELMA BPM и специализированной сервисной шины ESB. В случае системы ELMA BPM анализировалось удобство пользовательского интерфейса. В оценке интерфейса участвовали десятки специалистов различных служб предприятия, средний балл – 4,7 из 5, что является весьма высокой оценкой его эргономичности. Испытания специализированной сервисной шины ESB проводились с учётом различного числа и объемов информационных потоков в едином информационном пространстве АО «Газпром добыча Томск». Результаты показали высокие значения характеристик шины ESB, в первую очередь, её производительности.

В настоящее время ИИС находится на стадии опытной эксплуатации в АО «Газпром добыча Томск». Закончена организация единого информационного пространства предприятия. Письмо о передаче результатов диссертационного исследования на предприятие представлено в Приложении 1.

5.2. Результаты внедрения ИИС в ООО «Центр нефтегазовых технологий»

5.2.1. Результаты применения методов ИАД

ООО «Центр нефтегазовых технологий», далее ООО «ЦНГТ», – это инновационное предприятие, занимающееся созданием и внедрением методов интенсификации добычи УВС, разработкой наукоемкого программного обеспечения и методик его использования на добывающих предприятиях нефтегазовой отрасли. Кроме этого предприятие оказывает услуги добывающим предприятиям по обработке и анализу геолого-геофизических и технологических данных с эксплуатируемых месторождений, поэтому имеет геофизические, геологические и технологические данные с их месторождений.

При внедрении ИИС в ООО «ЦНГТ» специалистам предприятия важно было понять, насколько эффективны веб-сервисы, реализующие методы ИАД, при решении слабоформализуемых задач при управлении фондом скважин, на данных, имеющихся у этого предприятия. Ведь далеко не факт, что модели ИНС и

методы кластерного анализа, успешно справляющиеся с решением этих задач на наборах данных одного предприятия, дадут такие же высокие результаты на наборах данных по эксплуатируемым месторождениям другого предприятия.

Первая задача, которая решалась на этапе внедрения ИИС, – задача обнаружения и корректировки ошибок в значениях технологических и геологических параметров, хранимых в БД ООО «ЦНГТ». Как уже было сказано ранее, в БД многих предприятий нефтегазовой отрасли имеется довольно большое количество ошибок, привнесённых при измерениях параметров с помощью различных датчиков (устройств) и при ручном вводе данных специалистами. ООО «ЦНГТ» в данном случае не являлось исключением. Для обнаружения ошибок и восстановления пропусков в геологических и технологических данных с нескольких нефтяных месторождений в Ханты-Мансийском автономном округе использовался композитный веб-сервис сбора и актуализации данных, реализующий адаптивные алгоритмы предварительной подготовки данных для их последующего анализа с помощью методов ИАД. По мнению специалистов предприятия эта задача была успешно решена, причем в довольно сжатые сроки.

Вторым веб-сервисом для исследования стал веб-сервис планирования технологических режимов работы скважин путем ежемесячного прогноза значений дебитов добывающих скважин. Для оценки точности прогноза дебитов нефти, газа, газового конденсата (жидкости) и воды на одном из нефтяных месторождений Ханты-Мансийском автономном округе использовались модели глубоких ИНС из раздела 3 с рациональными архитектурой и гиперпараметрами. Такие модели ИНС обучались на предварительно подготовленных данных с этого месторождения. Результаты тестирования обученных моделей ИНС показали уменьшение погрешности прогноза значений дебитов WAPE в 1,96–2,87 раз по сравнению с методами, используемыми на предприятии. Эти результаты коррелируют с результатами прогноза, полученными на данных по месторождениям Томской области при разработке и исследовании моделей ИНС в разделе 3. Всё это позволяет сделать вывод об успешности решения такой слабоформализуемой задачи с помощью предложенных моделей ИНС прямого

распространения вне зависимости от того, в какой нефтегазоносной провинции находится месторождение: в Томской области или в Ханты-Мансийском автономном округе.

Задачи бинарной классификации скважин-кандидатов для ГТМ и классификации при выборе ГТМ для таких скважин решались для нефтяного месторождения из Ханты-Мансийского автономного округа с помощью моделей глубоких ИНС с использованием соответствующих веб-сервисов. Для формирования набора № 1 в БД ООО «ЦНГТ» этого месторождения оказалось недостаточно данных, поэтому готовился набор № 2 (все измеряемые технологические параметры скважин фонда). На этом наборе данных обучались модели ИНС с рациональными параметрами, показавшие наилучшие результаты для набора № 2 Месторождения 1 (см. п. 3.3.3.). Дополнительные эксперименты по обучению и тестированию моделей ИНС на данных нового месторождения с целью поиска более эффективных значений их параметров не проводились. Точность выбора скважин-кандидатов для ГТМ на фонде нового месторождения оказалась не более 98 %, а реализация идеи объединения обучающих выборок для Месторождения 1 и нового месторождения позволила повысить её до 99 %. Точность выбора типа ГТМ для скважин, выбранных кандидатами для ГТМ, составила 82 %. Последний результат по выбору типа ГТМ несколько хуже результатов, полученных при разработке этих моделей ИНС в случае данных по Месторождению 1 (см. п. 3.3.3). Близкие по точности результаты при решении каждой из этих слабоформализуемых задач смогли получить только опытные специалисты ООО «ЦНГТ», но времени у них на это ушло в несколько раз больше, чем в случае использования моделей ИНС. Отметим, что все эти описанные результаты внедрения ИИС, демонстрирующие её эффективность, получены совместно с сотрудниками ООО «ЦНГТ».

5.2.2. Организация единого информационного пространства ООО «Центр нефтегазовых технологий»

Основные положения концепции создания ИИС (раздел 2) и разработанные и описанные в разделе 4 компоненты SOA-модели ПО этой системы позволили

создать единое информационное пространство ООО «ЦНГТ». Внедрённые компоненты программного обеспечения ИИС, составляющие основу этого пространства, программное обеспечение унаследованных ИС, а также веб-сервисы, разработанные на предприятии, показаны на рисунке 5.2.

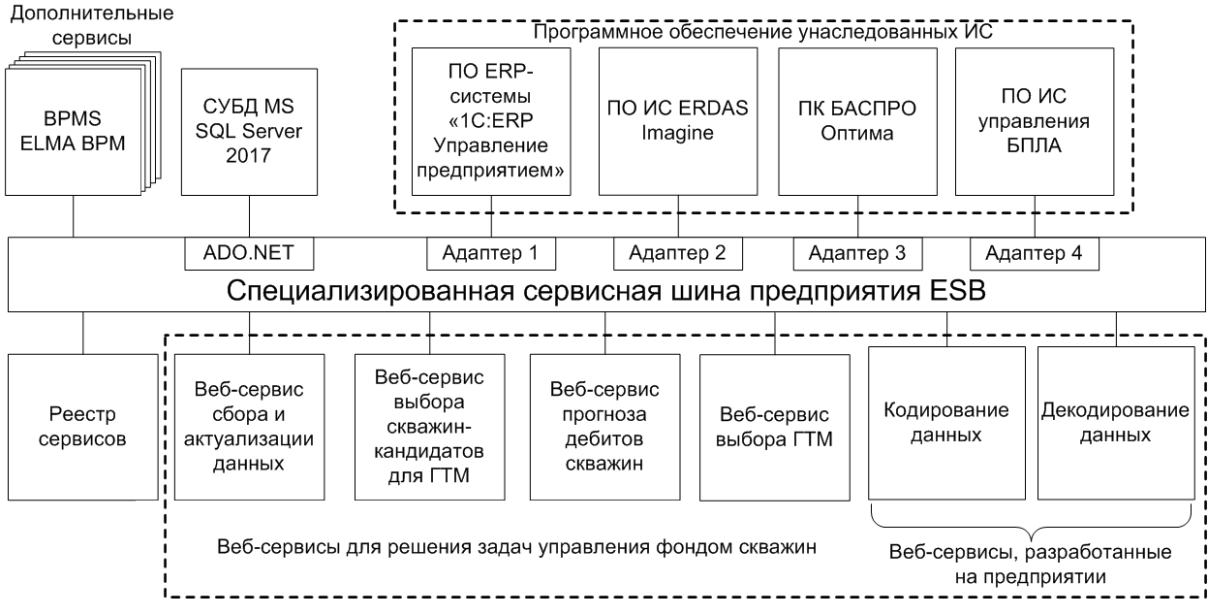

Рисунок 5.2. – Архитектура ПО единого информационного пространства ООО «ЦНГТ»

В состав единого информационного пространства входят следующие внедрённые в ООО «ЦНГТ» компоненты ИИС: специализированная сервисная шина предприятия ESB; БД предприятия под управлением СУБД MS SQL Server 2017; модифицированная система ELMA BPM с реализованными её среде веб-сервисами (на рисунке 5.2 показаны как дополнительные сервисы); композитный веб-сервис сбора и актуализации геологических и технологических данных; композитный веб-сервис для решения задачи бинарной классификации при выборе скважин-кандидатов для ГТМ (реализует модели глубоких ИНС прямого распространения и два метода кластерного анализа), веб-сервис для выбора ГТМ для скважин-кандидатов, а также веб-сервис для прогноза дебитов скважин.

Унаследованные ИС ООО «ЦНГТ» стандартизированных сетевых интерфейсов не имели, поэтому потребовалась реализация адаптеров на стороне ESB (четыре из них показаны на рисунке 5.2). В ООО «ЦНГТ» разработаны также с учётом требований со стороны специализированной шины ESB собственные веб-сервисы. На рисунке 5.2 приведены в качестве примера два веб-сервиса для

помехоустойчивого кодирования/декодирования данных. Кратко рассмотрим примеры унаследованных ИС предприятия, изображённые на рисунке 5.2.

ERP-система «1С:ERP Управление предприятием» [170] создана на основе платформы компании «1С» как комплексная информационная система для управления финансовыми ресурсами и производственной деятельностью предприятия. В этой системе используются комплексный подход к управлению бизнесом, международные методики и многолетняя отечественная практика.

ИС ERDAS Imagine [171] является мощной программной системой, предназначенной для обработки и интерпретации космических данных и аэрофотоснимков. Применяется при обработке таких данных в самых различных отраслях человеческой деятельности, в том числе в нефтегазовой отрасли.

Программный комплекс (ПК) БАСПРО Оптима (Россия) предназначен для мониторинга разработки месторождений УВС и геологического моделирования [167]. В п. 5.1.1 он уже кратко описан. В ООО «ЦНГТ» ПК используется для проведения детального геолого-промыслового анализа и интерпретации данных ГДИС, ПГИ и ГИС, а также для оценка извлекаемых запасов месторождений.

Программное обеспечение ИС для управления беспилотными летательными аппаратами (БПЛА), используемыми для мониторинга инфраструктуры эксплуатируемых месторождений, является собственной разработкой ООО «ЦНГТ».

Специализированная сервисная шина EBS и другие компоненты ИИС были апробированы в составе созданного единого информационного пространства предприятия в течение длительного периода опытной эксплуатации. По результатам эксплуатации отмечена высокая производительность шины EBS. Специалисты отдела ИТ предприятия указывают на простоту интеграции с её помощью разрабатываемых сервисов и унаследованных ИС.

В настоящее время закончено внедрение ИИС в ООО «ЦНГТ», включая создание единого информационного пространства предприятия. Получен акт о внедрении результатов диссертационного исследования (Приложение 2).

5.3. Методика применения ИИС при анализе данных

5.3.1. Общие положения

ИИС позволяет автоматизировать каждый БП управления фондом скважин (п. 1.1.2.), обеспечивая удобную работу с ней специалистам различных служб предприятия через разработанные пользовательские интерфейсы модифицированной системы ELMA BPM. Нотация BPMN построения БП в среде этой системы позволяет специалисту даже без должной квалификации в области систем управления БП перестраивать их под нужды своего предприятия, добавляя, редактируя или удаляя соответствующие элементы диаграммы. Для эффективного использования ИИС при выполнении БП управления фондом скважин необходимо руководствоваться следующими рекомендациями.

БП «Сбор предложений по проведению мероприятий, ведущих к остановкам скважин» описывается в среде системы ELMA BPM в виде отдельного процесса в нотации BPMN. Сбор предложений осуществляется от служб предприятия в интерфейсе веб-сервиса, входящего в набор дополнительных сервисов системы ELMA BPM. Список заинтересованных подразделений и лиц, которые подают свои предложения и согласуют итоговый перечень скважин на остановки, настраивается в соответствии с требованиями предприятия.

При реализации БП «Ежемесячное планирование технологических режимов работы каждой скважины фонда» в ИИС используется композитный веб-сервис, реализующий модели глубоких ИНС прямого распространения и выполняющийся в зависимости от решаемой задачи в виде одного из сервисов, входящих в его состав. При решении задачи ежемесячного планирования дебитов нефти, газа, жидкости (газового конденсата) и воды каждой добывающей скважины пользователь ИИС должен применять сервис «Прогноз дебитов добывающих скважин». При решении задачи ежемесячного планирования значений другого параметра из перечня технологических параметров скважины он должен воспользоваться сервисом «Прогноз других технологических параметров скважин». После вызова нужного сервиса пользователю важно правильно задать

параметры модели ИНС, рекомендации по этому поводу приведены ниже. Все прогнозируемые значения дебитов или других технологических параметров вычисляются на стороне сервиса, попадают в БД ИИС и становятся доступны в пользовательском интерфейсе ИИС для просмотра.

Для организации БП «Планирование и проведение исследований ГДИС, ГИС, ПГИ в скважинах» в ИИС используется разработанный в среде системы ELMA BPM веб-сервис, входящий в число её дополнительных сервисов. Он позволяет составлять планы по каждому типу исследований по фонду скважин на год, квартал и месяц, при этом доступны функции удаления, добавления, редактирования, фильтрации и сортировки записей.

БП «Управление и проведение ГТМ» является наиболее сложным. При его реализации используется несколько веб-сервисов и ряд дополнительных сервисов системы ELMA BPM. Сначала из композитного веб-сервиса сбора и актуализации данных выбирается и используется веб-сервис предварительной подготовки данных для выбора скважин-кандидатов и ГТМ. После выявления ошибок в данных и их корректировки пользователь вызывает композитный веб-сервис и ему предлагается из его состава выбрать веб-сервис, реализующий наиболее подходящий метод ИАД решения слабоформализуемой задачи выбора скважин-кандидатов для ГТМ. Пользователь может в меню выбрать традиционный метод решения этой задачи, тогда из ИИС идет обращение к унаследованной ИС, в которой реализован этот метод. После этого с помощью веб-сервисов решаются слабоформализуемая задача выбора типа ГТМ и задачи оценки эффективности ГТМ. Веб-сервисы после проведения расчетов возвращают результаты в БД ИИС. Дополнительные сервисы и пользовательские интерфейсы системы ELMA BPM позволяют пользователю через СУБД обращаться к БД, выполнять подпроцессы обработки результатов выполнения веб-сервисов и представлять их в требуемом виде. Особенности работы с веб-сервисами при решении слабоформализуемых задач изложены ниже.

БП «Формирование оперативного плана-графика всех мероприятий на фонде скважин на месяц» объединяет все возможные мероприятия с

определённым сроком выполнения от всех подразделений и организаций по каждой скважине: исследования (ГДИС, ГИС, и ПГИ), капитальный ремонт скважин (ГТМ, ликвидация аварий) и прочие работы. В системе ELMA BPM на каждое мероприятие назначается подрядная бригада. План-график доступен для оптимизации по суммарному времени простоя скважин с помощью соответствующего веб-сервиса. Реализована возможность совмещать мероприятия и привязывать их друг к другу. В модуле есть функциональность для разрешения конфликтов при наложении работ и нехватке бригад.

БП «Построение ежедневных и ежемесячных отчётов по обслуживанию фонда скважин» частично реализуется с использованием средств системы ELMA BPM. По запросам, формируемым в системе ELMA BPM, ряд отчетов создаётся в унаследованных ИС предприятия и передается в эту систему. Её функции позволяют делать рассылку отчётов всем заинтересованным специалистам. Для отчётов используется табличный формат Excel.

5.3.2. Методика применения ИНС при решении слабоформализуемых задач

При решении слабоформализуемой задачи ежемесячного планирования технологических режимов работы скважин фонда(ежемесячного прогнозирования значений технологических параметров, включая дебиты скважин**)**, следует использовать модели глубоких ИНС прямого распространения, предложенные и исследованные в п. 3.2.2. При этом на данных конкретного месторождения необходимо обучать модели ИНС со следующими рациональными характеристиками и параметрами: 10 или 11 значений параметра (измеренных отсчетов параметра во времени) во входном слое ИНС; 4 или 6 или 7 скрытых слоёв ИНС и функция активации LeakyReLU. Обучающие и тестовые значения прогнозируемого параметра каждой добывающей скважины фонда должны перед подачей на вход такой модели ИНС предварительно подготавливаться в зависимости от параметра либо по адаптивному алгоритму, описанному в п. 3.2.1. (для дебитов), либо по адаптивному алгоритму из п. 3.3.1. (для других технологических параметров). При подготовке обучающей и тестовой выборок по

каждому из адаптивных алгоритмов используется свой веб-сервис, входящий в композитный веб-сервис сбора и актуализации данных (рисунок 4.18).

Если эти модели ИНС обучены на данных по фонду скважин одного месторождения, то они обязательно должны быть дообучены на данных по другому фонду, параметры скважин которого подлежат прогнозированию. Без дообучения точность прогноза будет невелика. Возможно проведение собственных исследований и выбора рациональных значений параметров моделей глубоких ИНС прямого распространения. Исследования должны проводиться в соответствии с методикой из п. 3.2.2.

На практике для столь серьёзного по финансовым затратам мероприятия, как остановка скважины и выполнение на ней ГТМ, важно иметь как можно большую точность классификации скважин и ГТМ хотя бы по видам. Поэтому средняя точность классификации при решении слабоформализуемой задачи выбора скважин-кандидатов на ГТМ на тестовой выборке должна быть не менее 97 %, а точность классификации при решении слабоформализуемой задачи выбора типов ГТМ для таких скважин на тестовой выборке – не хуже 82 %.Такие пороговые (приемлемые для практического использования) значения точности классификации получены на основе экспертного опроса специалистов двух предприятий, где осуществлялась апробация ИИС. С учётом этого, кратко рассмотрим основные положения разработанной методики применения ИИС при решении этих двух слабоформализуемых задач.

1. В первую очередь необходимо сформировать наборы № 1 и № 2 технологических и геологических параметров фонда скважин (п. 3.3.1.) для обучения и тестирования моделей ИНС, так как для этих наборов результаты исследования многих моделей ИНС, приведённые в п. 3.3.3., сравнимы или выше указанных пороговых значений при решении каждой из двух слабоформализуемых задач. При классификации скважин-кандидатов на ГТМ рекомендуется использовать модели ИНС с функцией активации TanH или ReLU, обученные с помощью оптимизационных алгоритмов Adam или SGD. Число скрытых слоев для моделей ИНС выбирается из таблицы 3.17. Для решения

задачи классификации типов ГТМ с точностью выше пороговой следует использовать модель ИНС с 9 скрытыми слоями и функцией активации TanH, а обучать её необходимо по алгоритму Adam на наборе № 2.

2. Если в наборы № 1 и № 2 не входят (не измерялись) те или иные параметры фонда скважин, то следует формировать наборы параметров № 4 и № 5 для обучения и тестирования моделей ИНС. Для входного набора параметров № 4 модели ИНС с числом скрытых слоев от 2 до 6 и функцией активации TanH или модели ИНС с числом скрытых слоев 2, 4, 5, 6, 7, 8, 9 и 10 и функцией активации ReLU при обучении их с помощью алгоритма Adam могут дать точность классификации скважин-кандидатов для ГТМ выше пороговой. Для определенного числа скрытых слоев (см. таблицу 3.22) и тех же функций активации модели ИНС могут показать точность классификации скважин не хуже пороговой и при обучении их по алгоритмам SGD и AdaGrad. Для набора № 5 модели ИНС с числом скрытых слоев 3, 5, 6, 10 и включающие функции активации TanH или ReLU при их обучении по алгоритму Adam также могут дать точность классификации скважин выше пороговой. Задачу классификации типов ГТМ с точностью выше пороговой следует решать на наборе № 4 с помощью модели ИНС с 3 скрытыми слоями и функцией активации TanH, обученной по алгоритму Adam.

3. Формировать и использовать набор параметров № 3 для решения слабоформализуемых задач не следует, поскольку результаты исследований моделей ИНС (п. 3.3.3.) в случае этого набора всегда ниже пороговых значений.

4. При небольшой обучающей выборке точность решения задачи классификации скважин-кандидатов можно повысить более чем на 1 % дообучением модели ИНС на основе выборки с другого фонда скважин.

5. При отсутствии высокопроизводительных вычислительных средств можно использовать простые модели ИНС (малое число скрытых слоев и параметров на их входе). Так модель ИНС с 2 или 3 скрытыми слоями и функцией активации ReLU или TanH, обученная по алгоритму Adam на наборе параметров

№ 2, даст точность классификации при выборе скважин-кандидатов для ГТМ несколько выше пороговой.

6. Если результаты предлагаемых моделей ИНС не устраивают специалиста по точности, то должны быть построены и обучены дополнительные модели ИНС прямого распространения. При этом необходимо осуществлять построение обучающей выборки и впоследствии обучать модели ИНС на данных собственного предприятия с учётом всех принципов предварительной обработки данных, описанных в разделе 3 в п. 3.2 и 3.3.

7. Если для фонда или куста скважин имеются малый объем технологических и геологических данных и по нему невозможно сформировать обучающую выборку для моделей ИНС, то при решении задачи классификации скважин-кандидатов для ГТМ следует использовать иерархический или плотностной методы кластерного анализа. Для достижения порогового значения точности классификации следует брать сочетания значений геологических и технологических параметров. Можно вводить дополнительный кластер «средних» скважин, что позволяет получить разделимые три кластера скважин и в ряде случаев достичь точности классификации 100 %. Однако от специалиста требуются дополнительные трудозатраты при интерпретации результатов кластеризации. Следует помнить, что методы кластерного анализа дают менее стабильные по точности результаты классификации, чем модели ИНС.

5.4. Основные результаты и выводы по разделу

При апробации ИИС на двух предприятиях нефтегазовой отрасли получены следующие основные результаты и выводы.

1. Проведённая апробация ИИС на данных АО «Газпром добыча Томск» и ООО «ЦНГТ» показала, что применение моделей ИНС и, частично методов кластерного анализа, оказало значительную поддержку в принятии управленческих решений по фондам скважин, повысив точность выбора скважин-кандидатов для ГТМ до 97–99 %, а точность выбора типа ГТМ для них – до 82–83 %. Аналогичная точность достигается на предприятиях при решении этих

задач только опытными геологами. Однако решения по фонду скважин при использовании ИИС принимаются в 2–4 раза быстрее, чем при использовании геологами традиционных методов. Отметим, что при апробации точность классификации скважин при использовании ряда моделей ИНС с рациональными параметрами превосходила пороговое значение 97 %, а точность классификации типов ГТМ превосходила пороговое значение 82 %.

Результаты прогноза значений дебитов скважин фонда одного из нефтяных месторождений Ханты-Мансийского автономного округа с использованием моделей ИНС показали уменьшение погрешности прогноза WAPE в 1,96–2,87 раз по сравнению с методами, используемыми в ООО «ЦНГТ». Это подтверждает результаты исследований, изложенные в п. 3.2.2.

2. Разработанные концепция ИИС и ряд компонентов SOA-модели её программного обеспечения позволили создать единое информационное пространство АО «Газпром добыча Томск» и ООО «ЦНГТ». Во время опытной эксплуатации ИИС в АО «Газпром добыча Томск» осуществлялись комплексные испытания модифицированной системы ELMA BPM и специализированной сервисной шины ESB. Результаты показали удобство пользовательского интерфейса системы ELMA BPM и высокое значение производительности шины ESB.

3. По итогам внедрения результатов диссертационной работы в ООО «ЦНГТ» получен акт внедрения, а от АО «Газпром добыча Томск» – письмо о прохождении опытной эксплуатации ИИС на этом предприятии.

4. Разработана методика применения ИИС при реализации БП управления фондом скважин. Сформулированы основные положения методики при решении трёх слабоформализуемых задач.

5. Апробация ИИС ещё раз подтвердила, что полученные в работе теоретические и практические результаты являются важным вкладом в реализацию концепции интеллектуального месторождения УВС в части управления его фондом скважин.

ЗАКЛЮЧЕНИЕ

Известно, что интеллектуальные технологии – главный фактор обеспечения оптимальных управленческих решений при эксплуатации фондов скважин. В работе для управления такими фондами создана ИИС как перспективный инструмент поддержки принятия обоснованных решений специалистами. При этом получены следующие наиболее важные научные и практические результаты.

1. Анализ уровня автоматизации управления фондом скважин показал, что сегодня на большинстве нефтегазодобывающих предприятий специалисты используют традиционные простые методы и алгоритмы, либо применяют несложные методы ИАД при решении многих задач управления фондом скважин. Практически все они не позволяют решать такие задачи с требуемой точностью. Сделан вывод об актуальности создания современных многофункциональных ИИС для управления фондами скважин. Сформулированы цель и решаемые задачи по разработке такой ИИС на основе современных методов ИАД.

2. Разработана концепция ИИС, при этом показано, что наиболее эффективной архитектурой ПО такой ИИС для удовлетворения комплексных требований к системе является архитектура, построенная на принципах SOA. Разработана укрупнённая SOA-модель программного обеспечения ИИС.

3. Разработаны оригинальные адаптивные алгоритмы предварительной подготовки технологических и геологических данных для последующего применения методов ИАД. Показано, что они позволяют корректировать ошибки и пропуски в исходных данных с относительной погрешностью не более 4,9 %.

4. Предложены модели глубоких ИНС прямого распространения. Проведены исследования на реальных данных их эффективности при решении трёх слабоформализуемых задач для принятия решений в процессе управления фондами скважин. Исследована эффективность четырёх методов кластерного анализа при решении задачи классификации скважин-кандидатов для ГТМ. Погрешность прогноза значений дебитов скважин с помощью разработанных моделей ИНС была снижена по сравнению с погрешностью известных методов в

2–3 раза, точность выбора скважин-кандидатов достигла для ряда моделей ИНС более 99 %, а точность выбора типа ГТМ составила около 86 %. Отметим, что такая точность классификации скважин значительно превосходит пороговое значение 97 %, а точность классификации типов ГТМ – пороговое значение 82 %. Показано, что методы кластерного анализа при выборе скважин-кандидатов для ГТМ не всегда обеспечивают требуемую пороговую точность классификации.

5. С учётом разработанной концепции ИИС и результатов исследований методов ИАД реализованы все компоненты SOA-модели её ПО. Разработана специализированная сервисная шина предприятия, показана её высокая производительность. Обоснован выбор системы управления бизнес-процессами ELMA BPM, которая была затем модифицирована с учетом специфики процессов управления фондом скважин добывающего предприятия нефтегазовой отрасли. Разработаны ряд веб-сервисов для решения слабоформализуемых задач и дополнительные сервисы в среде ELMA BPM при управлении фондами скважин.

6. Проведена апробация ИИС на двух предприятиях нефтегазовой отрасли. Её результаты подтвердили эффективность разработанных моделей, алгоритмов и программного обеспечения ИИС. Более того, продемонстрированы возможность быстрой адаптации системы ELMA BPM к изменяющимся условиям предприятий и высокая производительность специализированной шины ESB. Это позволяет считать, что SOA-модель ПО и её компоненты могут служить основой при создании единого информационного пространства предприятий этой отрасли.

7. С учетом результатов исследований и апробации ИИС разработана методика её применения при управлении фондами скважин.

В итоге можно сделать вывод о том, что полученные в работе теоретические и практические результаты являются важным вкладом в реализацию концепции интеллектуального месторождения УВС в части управления его фондом скважин. Более того, результаты исследований и апробации ИИС позволяют считать, что в перспективе концепция ИИС и оригинальные компоненты SOA-модели её ПО могут быть использованы при создании единого информационного пространства предприятий в нефтеперерабатывающей и нефтехимической отраслях.

ЛИТЕРАТУРА

1. BP Statistical Review of World Energy June 2015 / Official site of BP (British Petroleum) company [Электронный ресурс]. URL: http://www.bp.com/content/dam/bp/pdf/energy-economics/statistical-review-2015/bp-statistical-review-of-world-energy-2015-full-report.pdf (дата обращения 18.12.2020).

2. Основные положения системного подхода к инженерно-технической защите информации [Электронный ресурс]. URL: http://www.bp.com/content/dam/bp/pdf/energy-economics/statistical-review-2015/bp-statistical-review-of-world-energy-2015-full-report.pdf (дата обращения 18.12.2020).

3. Крец В. Г., Лене Г. В. Основы нефтегазодобычи: Учебное пособие / Под ред. канд. геол.-минер. наук Волощука Г. М.. – Томск: изд-во Том. ун-та, 2000. – 220 с.

4. Коршак А. А., Шаммазов А. М. Основы нефтегазового дела: Учебник для вузов. - 3-е изд., испр. и доп. – Уфа.: ООО «ДизайнПолиграфСервис», 2005. – 528 с.

5. Evsyutkin I. V., Markov N. G. The information system of the geological and technical arrangements management on a well stock of an oil-and-gas production enterprise // Advances in Computer Science Research. – 2016. – 2352-53. – P. 379–383.

6. Евсюткин И. В., Марков Н. Г. Информационная система для управления геолого-техническими мероприятиями // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов II Международной конференции (г. Томск, 19–22 мая 2015 г.). – Томск: ТПУ, 2015. – ч. II. – С. 253–254.

7. Березина А. А., Череповицын А. Е. Экономическая концепция нефтегазового «интеллектуального» месторождения // Нефтяное хозяйство. – 2014. – № 14. – С. 14–15.

8. Редикульцев С. А., Липлянин А. В., Палий А. О. Использование метода нейронных сетей для прогноза параметров работы скважин после проведения ГРП // Бурение, разработка, добыча. – 2010. – Т. 1, № 5. – С. 33–37.

9. Нотация IDEF0 и пример её использования [Электронный ресурс]. URL: https://habr.com/ru/company/trinion/blog/322832/ (дата обращения 18.12.2020).

10. Марков Н. Г. Информационно-управляющие системы для газодобывающего производства. - Томск: изд-во Томского политехнического университета, 2016. - 261 с.

11. Гасанов И. Р. Обобщённая формула Дюпюи [Электронный ресурс]. URL: https://moluch.ru/archive/149/42066/ (дата обращения 18.12.2020).

12. Развитие цифрового месторождения за счет использования блока интеллектуализации нижнего уровня скважин [Электронный ресурс]. URL: https://magazine.neftegaz.ru/articles/nefteservis/456282-razvitie-tsifrovogo-mestorozhdeniya-za-schet-ispolzovaniya-bloka-intellektualizatsii-nizhnego-urovnya/ (дата обращения 18.12.2020).

13. Дементьев Л. Ф., Жданов М. А., Кирсанов А. Н. Применение математической статистики в нефтегазопромысловой геологии. – М.: изд-во Недра, 1977. – 255 с.

14. Добрецов Н. Л., Зуенко В. В., Шемякин М. Л. Статистические методы в геологии. – Новосибирск: изд-во Наука, 1974. – 144 с.

15. Ахмедов К. С., Аршинова Н. М., Семеняк А. А. Информационная система планирования и оценки эффективности ГТМ на фонде скважин ОАО «Газпром» // Газовая промышленность. – 2012. – № 7. – С. 51–55.

16. Тимонов А. В. Системный подход к выбору геолого-технических мероприятий для регулирования разработки нефтяных месторождений: дис. канд. техн. наук. – Уфа., 2010. – 151 с.

17. Ситников А. Н., Пустовских А. А., Ахметов А. В и др. Проактивный блочный анализ разработки месторождений // SPE International. Society of Petroleum Engineers. – 2015. – SPE-176572-RU.

18. Габдулов Р. Р., Никишков В. И., Сливка П. И. Обобщение выбора потенциальных скважин-кандидатов и технологий для проведения ремонтно-изоляционных работ // Научно-технический вестник ОАО «Роснефть». - 2009. - № 4. – С. 22–26.

19. Сыртланов В., Межнова Н., Ковалева Е. и др. О методах автоматизации оптимизации заводнения и подбора скважин-кандидатов для ГТМ // SPE International, Society of Petroleum Engineers. – 2015. – SPE-176730-RU.

20. Султанов А. С., Латифуллин Ф. М., Насыбуллин А. В. Автоматизированный подбор скважин-кандидатов для гидравлического разрыва пластов на АРМ геолога «Лазурит» // Нефтяное хозяйство. – 2010. – № 7. – С. 48–51.

21. Приезжев И. И., Егоров С. В. Гибридное обучение нейронных сетей с целью прогноза параметров нефтегазовой продуктивности горных пород // Сейсмические технологии. – 2017. – С. 205–208.

22. Келлер Ю. А. Разработка искусственных нейронных сетей для предсказания технологической эффективности от выравнивания профиля приёмистости // Известия Томского политехнического университета. Информационные технологии. – 2014. – № 5 (325). – С. 60–65.

23. Келлер Ю. А. Применение кластеризации данных на основе самоорганизующихся карт Кохонена при подборе скважин-кандидатов для методов увеличения нефтеотдачи // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2014. – № 3 (28). – С. 32–37.

24. Перминов Д. Е., Валеев С. В. Кластерный анализ с использованием элементов нечеткой логики с целью автоматического поиска скважин-кандидатов для проведения геолого-технических мероприятий // Научно-технический вестник ОАО «НК «Роснефть». – 2013. – № 1 (30). – С. 31–35.

25. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: классификация и снижение размерности. – М.: Финансы и статистика, 1989. – 607 с.

26. Ту Дж., Гонсалес Р. Принципы распознавания образов. – М.: Мир, 1978. – 412 с.

27. Кашапов А. Применение теории нечётких множеств для поиска скважин-кандидатов на проведение геолого-технических мероприятий // SPE International. Society of Petroleum Engineers. – 2015. – SPE-176744-RU.

28. Галиуллин М. М., Зимин П. В., Васильев В. В. Методика выбора скважин-кандидатов для интенсификации добычи с использованием математического аппарата нечёткой логики // Нефтяное хозяйство. – 2011. – № 6. – С. 120–123.

29. Галиуллин М. М., Шабаров А. Б. Применение теории нечетких множеств для подбора скважин с целью геолого-технологических мероприятий на нефтяных месторождениях // Вестник Тюменского государственного университета. Физико-математические науки. Информатика. – 2011. – № 7. – С. 30–37.

30. Odedele T. O., Ibrahim H. D. Oil Well Performance Diagnosis System Using Fuzzy Logic Inference Models // Proceedings of The World Congress on Engineering. – 2014. – P. 80–85.

31. Okwu M. O., Nwachukwu N. A. A review of fuzzy logic applications in petroleum exploration, production and distribution operations // Journal of Petroleum Exploration and Production Technology. – 2018. – Vol. 9. – P. 1555–1568.

32. Кудинов А. В., Марков Н. Г. / Проблемы автоматизации производства газодобывающих компаний: монография. – Томск: изд-во Томского политехнического университета, 2012. – 247 с.

33. Software Integrated Solutions / Сайт компании Schlumberger. Программные продукты для геологоразведки и добычи [Электронный ресурс]. URL: https://www.sis.slb.ru (дата обращения 18.12.2020).

34. Halliburton LandMark / A Halliburton business line is the leading technology solution provider of data and analytics [Электронный ресурс]. URL: https://www.landmark.solutions (дата обращения 18.12.2020).

35. Барбер Э., Шиппен М. Е., Баруа С. и др. Оптимизация добычи: от продуктивного пласта до пункта подготовки нефти и газа // Нефтегазовое обозрение. – 2008. –№ 4 (19). – С. 22–37.

36. Концепция Big Data [Электронный ресурс]. URL: https://rb.ru/howto/chto-takoe-big-data/ (дата обращения 18.12.2020).

37. Маргелов Д. В. Месторождение на ладони – инновационный взгляд на перспективу интеллектуальных месторождений // Инженерная практика. – 2010. – № 9. – С. 43–46.

38. Черкасов М. «Умные» технологии в нефтегазовой отрасли // Control Engineering. Россия. – 2015. - № 3 (57). – С. 20-22.

39. Инженерное программное обеспечение. Материалы презентации компании Petroleum Experts [Электронный ресурс]. URL: http://itps.com/uploads/files/Petex/20IPM/20Brochure/20RUS.pdf (дата обращения 18.12.2020).

40. Дмитриевский А. Н. Цифровизация и интеллектуализация нефтегазовых месторождений // Автоматизация и информационные технологии в нефтегазовой области. – 2016. – № 2 (24). – С. 13–19.

41. Научно-технический центр ПАО «Газпромнефть» [Электронный ресурс]. URL: https://ntc.gazprom-neft.ru/about/company/ (дата обращения 18.12.2020).

42. Симкин А. В. Подход к комплексному применению методологий систематизации требований // Прикладная информатика. –2013. - № 3. – С. 60–75.

43. Официальный сайт ПАО «Сургутнефтегаз» [Электронный ресурс]. URL: https://www.surgutneftegas.ru/ (дата обращения 18.12.2020).

44. Диаграмма вариантов использования (use case diagram) [Электронный ресурс]. URL: http://khpi-iip.mipk.kharkiv.edu/library/case/leon/gl4/gl4.html (дата обращения 18.12.2020).

45. Джонс Дж., Форман Р., Форрестер Д., Дюшарм Д. Модель зрелости ИТ. Сервис-ориентированная архитектура в подразделении разведки и добычи BP // Oil & Gas Journal. – 2017. – № 38 (105). – С. 36–43.

46. Решетников И. С., Козлецов А. П. MES – теория и практика – М.: Российская рабочая группа MESA International, 2010. – 98 с.

47. Евсюткин И. В., Марков Н. Г. Архитектура информационной системы для управления геолого-техническими мероприятиями // Молодёжь и современные информационные технологии: сборник трудов XIII Международной научно-практической конференции студентов, аспирантов, молодых учёных (г. Томск, 9–13 ноября 2015 г.). – Томск: изд-во ТПУ, 2015. – С. 143–144.

48. Козлецов А. П., Решетников И. С. Современные способы организации обмена данными с системами управления // Информационные технологии в проектировании и производстве. – 2010. – № 2. – С. 17–23.

49. Козлецов А. П., Решетников И. С. Стандарты и технологии интеграции производственных информационных систем // Технологии управления. – 2010. – № 2. – С. 24–30.

50. Juric M. SOA approach to integrational. – Birmingham: Packt Publishing Ltd, 2007. – 366 p.

51. Медокс Я., Боркус В., Жуков К. и др. Работать с сервисами сложнее, чем с традиционными приложениями // Intelligent Enterprise. – 2007. – № 16. – С. 18–30.

52. BPMN – модель бизнес-процессов и нотация [Электронный ресурс]. URL: www.elma-bpm.ru (дата обращения 18.12.2020).

53. A Hands-on Introduction to BPEL [Электронный ресурс]. URL: https://www.oracle.com/technical-resources/articles/matjaz-bpel.html (дата обращения 18.12.2020).

54. WP 27. Сервис-ориентированная архитектура в системах управлении производством. Официальные материалы ассоциации MESA International // MES – теория и практика. – 2010. - №2. - С. 5–59.

55. Hayes B. First links in the Markov chain // American Scientist. –2013. – № 2 (101). – P. 252–257.

56. MES системы – современный стандарт управления производством [Электронный ресурс]. URL: http://asapcg.com/press-center/articles/mes-sistemy/ (дата обращения 18.12.2020).

57. Crow’s Foot Notation [Электронный ресурс]. URL: http://www2.cs.uregina.ca/~bernatja/crowsfoot.html/ (дата обращения 18.12.2020).

58. Определение дебита нефтяной скважины: формула и методы расчета [Электронный ресурс]. URL: http://snkoil.com/press-tsentr/polezno-pochitat/debit-neftyanoy-skvazhiny/ (дата обращения 18.12.2020).

59. Методические рекомендации по подготовке технических проектов разработки месторождений углеводородного сырья [Электронный ресурс]. URL: http://www.gkz-rf.ru/sites/default/files/docs/metodicheskie_rekomendacii_po_pravilam_prektirovaniya_uvs.pdf (дата обращения 18.12.2020).

60. Материалы открытого курса OpenDataScience и Mail.Ru Group по машинному обучению [Электронный ресурс]. URL: https://habr.com/company/ods/blog/344044/ (дата обращения 18.12.2020).

61. Han J., Kamber M., Pei J. Data Mining: Concepts and Techniques, 3rd Edition. – Elsevier Science, 2011. – 744 p.

62. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain // Psychological review. – 1958. – № 6 (65). – P. 386–392.

63. Минский М., Пейперт С. Персептроны. – М.: изд-во «Мир», 1971. – 263 с.

64. Искусственные нейронные сети: архитектуры и обучение [Электронный ресурс]. URL: https://postnauka.ru/longreads/155977 (дата обращения 18.12.2020).

65. Bengio Yoshua Learning Deep Architectures for AI // Foundations and Trends in Machine Learning. – 2009. – № 2 (1). – P. 1–127

66. Schmidhuber J. Deep Learning in Neural Networks: An Overview // Neural Networks. – 2015. – № 61. – P. 85–117.

67. Глубокие нейронные сети и почему они требуют так много памяти [Электронный ресурс]. URL: https://habr.com/ru/post/402641/ (дата обращения 18.12.2020).

68. Broomhead D. S., Broomhead D. L. Radial basis functions, multi-variable functional interpolation and adaptive networks // Royal signals and radar establishment malvern. – 1988. – No. RSRE-MEMO-4148.

69. Sahin Ferat. A Radial Basis Function Approach to a Color Image Classification Problem in a Real Time Industrial Application (dis. M. Sc.). – 1997. – Virginia Tech. – 26 p. [Электронный ресурс]. URL: http://web.archive.org/web/20151026005931/http://scholar.lib.vt.edu/theses/available/etd-6197-223641/unrestricted/Ch3.pdf (дата обращения 18.12.2020).

70. Hopfield John J. Neural networks and physical systems with emergent collective computational abilities // Proceedings of the national academy of sciences. – 1982. – № 8 (79). – P. 2554–2558.

71. Hayes B. First links in the Markov chain // American Scientist. – 2013. – № 2 (101). – P. 252–258.

72. Hinton G. E., Sejnowski T. J. Learning and relearning in Boltzmann machines // Parallel distributed processing: Explorations in the microstructure of cognition. – 1986. – Vol. 1. – P. 282–317.

73. Smolensky Paul. Information processing in dynamical systems: Foundations of harmony theory // Colorado university at boulder dept of computer science. – 1986. – No. CU-CS-321-86. – P. 194–281.

74. Kryzhanovsky B. V., Litinskii L. B., Mikaelian A. L. Vector-neuron models of associative memory // Proc. of Int. Joint Conference on Neural Networks IJCNN-04, Budapest-2004. – 2004. – P. 909–1004.

75. Достоинства, недостатки и модификации Сети Хопфилда. [Электронный ресурс]. URL: https://lektsii.com/1-82490.html (дата обращения 18.12.2020).

76. Bourlard Hervé, Yves Kamp. Auto-association by multilayer perceptrons and singular value decomposition. // Biological cybernetics. – 1988. – № 4-5 (59). – P. 291–294.

77. Marc’Aurelio Ranzato, Christopher Poultney, Sumit Chopra, Yann LeCun. Efficient learning of sparse representations with an energy-based model. // Advances in

Neural Information Processing Systems 19 – Proceedings of the 2006 Conference. – 2007. – 1049-5258. – P. 1137–1144.

78. Kingma Diederik P., Max Welling. Auto-encoding variational bayes // arXiv preprint arXiv. – 2013. – 1312.6114. – P. 1–14.

79. Vincent Pascal, et al. Extracting and composing robust features with denoising autoencoders // Machine learning. Proceedings of the 25th international conference, Helsinki, Finland, June 5–9, 2008. – 2008 – P. 1–16.

80. Автоэнкодеры и сильный искусственный интеллект [Электронный ресурс]. URL: https://habr.com/ru/post/417405/ (дата обращения 18.12.2020).

81. LeCun Yann, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE. – 1998. – № 11 (86). – P. 2278–2324.

82. Elman Jeffrey L. Finding structure in time. // Cognitive science. – 1990. – № 2 (14). – P. 179–211.

83. Hochreiter Sepp, Jürgen Schmidhuber. Long short-term memory // Neural computation. – 1997. – № 8 (9). – P. 1735–1780.

84. Chung Junyoung, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling. // arXiv preprint arXiv. – 2014. – 1412.3555. – P. 1–9.

85. Kohonen Teuvo. Self-organized formation of topologically correct feature maps. // Biological cybernetics. – 1982. – № 1 (43). – P. 59–69.

86. Нейросети и глубокое обучение: выбор гиперпараметров нейросети [Электронный ресурс]. URL: https://habr.com/ru/post/460711/ (дата обращения 18.12.2020).

87. Сравнение алгоритмов и выбор гиперпараметров [Электронный ресурс]. URL: https://ru.coursera.org/lecture/supervised-learning/sravnieniie-alghoritmov-i-vybor-ghipierparamietrov-aF79U (дата обращения 18.12.2020).

88. Инициализация ИНС случайными весами [Электронный ресурс]. URL: https://www.machinelearningmastery.ru/why-initialize-a-neural-network-with-random-weights/ (дата обращения 18.12.2020).

89. Parameters and constants [Электронный ресурс]. URL: https://docs.microsoft.com/en-us/cognitive-toolkit/parameters-and-constants (дата обращения 18.12.2020).

90. Коэффициент скорости обучения (Learning rate) [Электронный ресурс]. URL: https://wiki.loginom.ru/articles/learning-rate.html (дата обращения 18.12.2020).

91. Количество слоев и узлов в нейронной сети [Электронный ресурс]. URL: https://www.machinelearningmastery.ru/how-to-configure-the-number-of-layers-and-nodes-in-a-neural-network/ (дата обращения 18.12.2020).

92. Dropout – метод решения проблемы переобучения в нейронных сетях [Электронный ресурс]. URL: https://habr.com/ru/company/wunderfund/blog/330814/ (дата обращения 18.12.2020).

93. Нейронная сеть: алгоритмы, обучение, функции активации и потери [Электронный ресурс]. URL: https://neurohive.io/ru/osnovy-data-science/osnovy-nejronnyh-setej-algoritmy-obuchenie-funkcii-aktivacii-i-poteri/ (дата обращения 18.12.2020).

94. Введение в машинное обучение и искусственные нейронные сети [Электронный ресурс]. URL: https://foobar167.github.io/page/vvedeniye-v-mashinnoye-obucheniye-i-iskusstvennyye-neyronnyye-seti.html#ref2.2.2 (дата обращения 18.12.2020).

95. Рецепт обучения нейронных сетей [Электронный ресурс]. URL: https://id-lab.ru/posts/developers/reczept-obucheniya-nejronnyh-setej/ (дата обращения 18.12.2020).

96. Болотова Л. С. Системы поддержки принятия решений: учебник и практикум для академического бакалавриата. – Ч. 1. – М.: изд-во Юрайт, 2018.– 257 с.

97. Нейронные сети: обучение с учителем [Электронный ресурс]. URL: http://masters.donntu.org/2012/fknt/kovalyov_a/library/article8.htm (дата обращения 18.12.2020).

98. Каллан Р. Основные концепции нейронных сетей: пер. с англ. – М.: изд-во «Вильямс», 2001. – 288 с.

99. Осовский С. Нейронные сети для обработки информации. – М.: Финансы и статистика, 2002. – 344 с.

100. Методы оптимизации нейронных сетей [Электронный ресурс]. URL: https://habr.com/ru/post/318970/ (дата обращения 18.12.2020).

101. Стохастический градиентный спуск [Электронный ресурс]. URL: https://ru.coursera.org/lecture/supervised-learning/stokhastichieskii-ghradiientnyi-spusk-xRY50 (дата обращения 18.12.2020).

102. Stochastic Optimization Techniques [Электронный ресурс]. URL: http://colinraffel.com/wiki/stochastic_optimization_techniques (дата обращения 18.12.2020).

103. Функции активации нейросети: сигмоида, линейная, ступенчатая, ReLU, TanH [Электронный ресурс]. URL: https://neurohive.io/ru/osnovy-data-science/activation-functions/ (дата обращения 18.12.2020).

104. Нейросетевая библиотека CNTK [Электронный ресурс]. URL: https://habr.com/ru/ company/microsoft/blog/336552/ (дата обращения 18.12.2020).

105. Shaohuai Shi, Qiang Wang, Pengfei Xu, Xiaowen Chu Benchmarking State-of-the-Art Deep Learning Software Tools // arXiv:1608.07249v7 [cs.DC] 17 Feb 2017. [Электронный ресурс]. URL: https://arxiv.org/pdf/1608.07249.pdf (дата обращения 18.12.2020).

106. Microsoft/CNTK [Электронный ресурс]. URL: https://github.com/ Microsoft/CNTK (дата обращения 18.12.2020).

107. The Microsoft Cognitive Toolkit [Электронный ресурс]. URL: https://docs.microsoft.com/en-us/cognitive-toolkit/ (дата обращения 18.12.2020).

108. Формула скользящего среднего [Электронный ресурс]. URL: https://ru.photo-555.com/9085155-moving-average-formula (дата обращения 18.12.2020).

109. Линейная интерполяция [Электронный ресурс]. URL: https://www.matematicus.ru/vysshaya-matematika/issledovanie-funktsii/linejnaya-interpolyatsiya (дата обращения 18.12.2020).

110. Интерполяция [Электронный ресурс]. URL: http://aco.ifmo.ru/el_books/numerical_methods/lectures/glava3.html (дата обращения 18.12.2020).

111. Сплайн-интерполяция [Электронный ресурс]. URL: http://www.astro.tsu.ru/OsChMet/7_7.html (дата обращения 18.12.2020).

112. Методические указания к решению задач по интерполяции функций [Электронный ресурс]. URL: http://www.lib.unn.ru/students/src/alkinterpol.pdf (дата обращения 18.12.2020).

113. Кросс-валидация [Электронный ресурс]. URL: https://long-short.pro/post/kross-validatsiya-cross-validation-304/ (дата обращения 18.12.2020).

114. Ошибка прогнозирования: виды, формулы, примеры [Электронный ресурс]. URL: https://shtem.ru/ошибка-прогнозирования-формула/ (дата обращения 18.12.2020).

115. Евсюткин И. В., Марков Н. Г. Глубокие искусственные нейронные сети для прогноза значений дебитов добывающих скважин // Известия ТПУ. Инжиниринг георесурсов. – 2020. –№ 11 (331). – С. 88–95.

116. Евсюткин И. В., Марков Н. Г. Прогноз значений дебитов скважин с использованием искусственных нейронных сетей // Сборник трудов VIII Международной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем» (г. Томск, 27–31 мая 2020 г.). – Томск: ТГУ, 2020. – С. 34–40.

117. Основы линейной регрессии [Электронный ресурс]. URL: http://statistica.ru/theory/osnovy-lineynoy-regressii/ (дата обращения 18.12.2020).

118. Методы экстраполяции [Электронный ресурс]. URL: http://www.uamconsult.com/book_547.html (дата обращения 18.12.2020).

119. Линейная регрессия [Электронный ресурс]. URL: https://docs.exponenta.ru/matlab/ref/interp2.html (дата обращения 18.12.2020).

120. Khandakar M. Times-series data augmentation and deep learning for construction equipment activity recognition / M. Rashid Khandakar, Joseph Louis // Advanced Engineering Informatics. – 2019. – Vol. 42. – Article 100935.

121. Srivastava N., Hinton G., Krizhevsky A., et al. Dropout: A Simple Way to Prevent Neural Networks from overfitting // Journal of Machine Learning Research. – 2014. – Vol. 15 – P. 1929–1958.

122. Darío Baptista Performance comparison of ANN training algorithms for classification / Darío Baptista, Sandy Rodrigues, F. Morgado-Dias // Conference: Intelligent Signal Processing (WISP), 2013 IEEE 8th International Symposium. – 2013. – P. 115–120.

123. Евсюткин И. В., Марков Н. Г. Искусственные нейронные сети для решения задач управления геолого-техническими мероприятиями // Сборник трудов XVI Международной научно-практической конференции студентов, аспирантов, молодых учёных «Молодёжь и современные информационные технологии» (г. Томск, 17–20 февраля 2020 г.). – Томск: ТПУ, 2020. – С. 81–82.

124. Evsyutkin I. V., Markov N. G. The intellectual analysis of geological and technological data during the management of an oil field’s well-stock // Journal of Physics: Conference Series – 2020. – Vol. 1661. – [012033, 7 p.]

125. Евсюткин И. В., Марков Н. Г. Управление геолого-техническими мероприятиями на месторождениях нефти и газа с использованием искусственных нейронных сетей // Доклады ТУСУР / Управление, вычислительная техника и информатика. – 2020. – № 1 (23). – С. 62–69.

126. Guha S., Rastogi R., Shim K. CURE: An Efficient Clustering Algorithm For Large Databases / Proceedings of the ACM-SIGMOD, International Conference on Management of Data, June 1–4 1998, Seattle, Washington. // ACM Press. – 1998. – P. 73–84.

127. Zhang T., Ramakrishnan R., Linvy M. BIRCH: An Efficient Data Clustering Method for Large Databases. Proceedings of the ACM-SIGMOD, International Conference on Management of Data, June 4–6, 1996, Montreal, Quebec. // ACM Press. – 1996. – P. 103–114.

128. Ester M., Kriegel H.-P., Sander J., Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, August 2–4, 1996, Portland, Oregon. // The AAAI Press. – 1996. – P. 226–231.

129. Kailing K., Kriegel H.-P., Kröger P. Density-Connected Subspace Clustering for High-Dimensional Data. Proceedings of 4th SIAM Intern. Conf. on Data Mining, April 22–24, 2004, Lake Buena Vista, Florida. // Society for Industrial and Applied Mathematics. – 2004. – P. 246–257.

130. Нейросети и глубокое обучение. [Электронный ресурс]. URL: https://habr.com/ru/post/456738/ (дата обращения 18.12.2020).

131. Evsyutkin I. V. Intellectual information system for management of geological and technical arrangements during oil field exploitation / Markov N. G., Vasilyeva E. E., Evsyutkin I. V. // Journal of Physics: Conference Series. – 2017. – Vol. 803. – [012093, 6 p.].

132. Васильева Е. Е, Евсюткин И. В. Интеллектуальная информационная система управления геолого-техническими мероприятиями на фонде скважин // Тезисы докладов XVI Всеросс. конфер. молодых ученых по математическому моделированию и информационным технологиям (г. Красноярск, 28–30 октября 2015 г.). – Новосибирск: ИВТ СО РАН, 2015. – С. 63–64.

133. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И. Д. Рудинского. – М.: Горячая линия – Телеком, 2006. – 452 с.

134. Кононюк А. Е., Дискретно-непрерывная математика, Книга 10, Алгоритмы, Часть 3, Генетические алгоритмы [Электронный ресурс]. URL: http://ecat.diit.edu.ua/ft/Algorithms3.pdf (дата обращения 18.12.2020).

135. Visual Studio [Электронный ресурс]. URL: https://visualstudio.microsoft.com/ru/ (дата обращения 18.12.2020).

136. C#/.NET [Электронный ресурс]. URL: https://metanit.com/ (дата обращения 18.12.2020).

137. ASP.NET Core [Электронный ресурс]. URL: https://metanit.com/sharp/aspnet5/1.1.php (дата обращения 18.12.2020).

138. C#/.NET [Электронный ресурс]. URL: https://hsbi.hse.ru/articles/obzor-populyarnykh-subd/ (дата обращения 18.12.2020).

139. Обзор популярных СУБД [Электронный ресурс]. URL: https://proglib.io/p/databases-2019 (дата обращения 18.12.2020).

140. SQL Server и T-SQL [Электронный ресурс]. URL: https://metanit.com/sql/sqlserver/1.1.php (дата обращения 18.12.2020).

141. BPMN Tool Matrix. Электронная таблица BPMS [Электронный ресурс]. URL: https://bpmnmatrix.github.io/ (дата обращения 18.12.2020).

142. Евсюткин И. В., Марков Н. Г. Выбор системы управления бизнес-процессами для нефтегазодобывающего предприятия // Сборник трудов XV Международной научно-практической конференции «Молодежь и современные информационные технологии» (Томск, 04–07 декабря 2017 г.). – Томск: Д-Принт. 2018.– С. 237–238.

143. Oracle Service Bus. Официальный сайт производителя Oracle [Электронный ресурс]. URL: http://www.oracle.com/technetwork/middleware/service-bus/overview/index.html (дата обращения 18.12.2020).

144. WebSphere Message Broker. Официальный сайт производителя IBM [Электронный ресурс]. URL: https://www.ibm.com/support/knowledgecenter/en/SSKM8N/mapfiles/product_welcome.html (дата обращения 18.12.2020).

145. ActiveMatrix Service Bus. Официальный сайт производителя Tibco [Электронный ресурс]. URL: https://docs.tibco.com/pub/activematrix_service_bus/3.3.0_september_2013/doc/pdf/tib_amx_concepts/tib_amx_concepts.pdf (дата обращения 18.12.2020).

146. WebMethods ESB. Официальный сайт производителя Software AG [Электронный ресурс]. URL: https://www.softwareag.com/resources/Enterprise-service-bus (дата обращения 18.12.2020).

147. Sonic ESB. Официальный сайт производителя Progress Software [Электронный ресурс]. URL: http://www.progress-tech.ru/products/sonic/esb (дата обращения 18.12.2020).

148. JBoss ESB. Официальный сайт производителя Red Hat [Электронный ресурс]. URL: http://jbossesb.jboss.org (дата обращения 18.12.2020).

149. Выбор интеграционной платформы: технологии и критерии [Электронный ресурс]. URL: https://www.epam-group.ru/ideas/white-papers/integration-platform-choice (дата обращения 18.12.2020).

150. Выбор сервисной шины предприятия [Электронный ресурс]. URL: https://www.softlab.ru/blog/issledovaniya/5494/ (дата обращения 18.12.2020).

151. Сравнение интеграционной шины Mediator ESB и существующих на рынке решений [Электронный ресурс]. URL: http://www.dasystems.ru/esbComparison.html (дата обращения 18.12.2020).

152. Корпоративная сервисная шина – «бюджетный» подход к решению задач интеграции [Электронный ресурс]. URL: http://citforum.ru/internet/webservice/esb/ (дата обращения 18.12.2020).

153. Обзор ESB-систем ServiceMix и Fuse [Электронный ресурс]. URL: https://habr.com/post/311540/ (дата обращения 18.12.2020).

154. Top Open Source ESB Projects (IBM) [Электронный ресурс]. URL: https://dzone.com/articles/top-open-source-esbs (дата обращения 18.12.2020).

155. Выбор корпоративной сервисной шины с открытым исходным кодом в составе ИС на основе сервис-ориентированной архитектуры [Электронный ресурс]. URL: http://www.ict.edu.ru/vconf/files/11590.pdf (дата обращения 18.12.2020).

156. Опыт внедрения ESB в ПАО «Газпром нефть» [Электронный ресурс]. URL: https://infostart.ru/public/925150 (дата обращения 18.12.2020).

157. Интеграционная платформа «Галактика ESB» [Электронный ресурс]. URL: https://servernews.ru/987101 (дата обращения 18.12.2020).

158. Интеграция автоматизированных систем при помощи корпоративной сервисной шины (ESB) [Электронный ресурс]. URL: https://trinidata.ru/integration.htm (дата обращения 18.12.2020).

159. Общие сведения о платформе .NET [Электронный ресурс]. URL: https://docs.microsoft.com/ru-ru/dotnet/framework/get-started/overview (дата обращения 18.12.2020).

160. Евсюткин И. В., Марков Н. Г. Специализированная сервисная шина для создания единого информационного пространства компаний нефтегазовой отрасли // Программные продукты и системы. – 2019. – Т. 32, № 2. – С. 326–336.

161. Евсюткин И. В., Марков Н. Г. Специализированная сервисная шина предприятия // Сб. трудов V Международной конференции «Информационные технологии в науке, управлении, социальной сфере и медицине» (г. Томск, 17–21 мая 2018 г.). – ч. I. – Томск: ТПУ, 2018. – С. 232–237.

162. Воротов В. Е., Карождей Д. В., Евсюткин И. В. Программные средства для формирования журналов работ на фонде скважин // Технологии Microsoft в теории и практике программирования: сборник трудов VIII Всероссийской научно-практической конференции студентов, аспирантов и молодых учёных (г. Томск, 22–23 марта 2016 г.). – Томск: ТПУ, 2016. – С. 238–240.

163. Евсюткин И. В., Марков Н. Г. Веб-сервис обнаружения и корректировки ошибок в данных на промыслах нефтегазодобывающих компаний // Сборник трудов XVI Международной научно-практической конференции «Молодёжь и современные информационные технологии» (г. Томск, 3–7 декабря 2018 г.). – Томск: ТПУ, 2019. – С. 231–232.

164. Свидетельство о государственной регистрации программы для ЭВМ № 2020663208 (RU); заявка № 2020662243 от 13.10.2020, дата рег. 23.10.2020; Бюл. № 11 от 23.10.2020 // Евсюткин И. В., Марков Н. Г. Обработка геолого-технологических данных при формировании обучающей и тестовой выборок для искусственных нейронных сетей.

165. Официальный сайт «Востокгазпром» [Электронный ресурс]. URL: https://vostokgazprom.gazprom.ru/about/ (дата обращения 18.12.2020).

166. Возможности Microsoft Dynamics AX 2012 [Электронный ресурс]. URL: https://docs.microsoft.com/ru-ru/dynamicsax-2012/appuser-itpro/introduction-to-microsoft-dynamics-ax-2012. (дата обращения 18.12.2020).

167. Программный комплекс БАСПРО ОПТИМА [Электронный ресурс]. URL: https://platforms.su/platform/5973 (дата обращения 18.12.2020).

168. ECLIPSE [Электронный ресурс]. URL: https://sis.slb.ru/products/eclipse/ (дата обращения 18.12.2020).

169. Tempest [Электронный ресурс]. URL: https://www.petroleumengineers.ru/app/8800 (дата обращения 18.12.2020).

170. 1С:Предприятие 8 Система программ [Электронный ресурс]. URL: https://v8.1c.ru/ (дата обращения 18.12.2020).

171. ERDAS Imagine [Электронный ресурс]. URL: https://www.hexagongeospatial.com/products/power-portfolio/erdas-imagine (дата обращения 18.12.2020).

ПРИЛОЖЕНИЕ 1 ПИСЬМО О ПЕРЕДАЧЕ РЕЗУЛЬТАТОВ РАБОТЫ В АО «ГАЗПРОМ ДОБЫЧА ТОМСК»

ВОСТОКГАЗПРОМ

Акционерное общество
«Востокгазпром»
(АО «Востокгазпром»)

ул. Большая Подгорная, д. 73, г. Томск,
Томская область, Российская Федерация, 634009
тел.: +7 (3822) 40-63-03, факс: +7 (3822) 40-69-44, 61-21-93
e-mail: sandervgp@vostokgazprom.ru, www.vostokgazprom.gazprom.ru
ОКПО 49382579, ОГРН 1027000850111, ИНН 7017005296, КПП 701701001

02.02.2021 № 13/138

на № _____ от _____

В диссертационный
совет ДС. ТПУ.15

Настоящим письмом подтверждаем, что программное обеспечение интеллектуальной информационной системы (ИИС) для управления фондом скважин, разработанное ассистентом ТПУ Евсюткиным И.В. при выполнении им кандидатской диссертации, передано в АО «Газпром добыча Томск» — дочернее добывающее предприятие АО «Востокгазпром» в следующем составе:

- специализированная сервисная шина предприятия ESB;
- веб-сервисы сбора и актуализации технологических и геологических данных по фонду скважин, реализующие алгоритмы предварительной обработки этих данных;
- веб-сервисы для выполнения ряда бизнес-процессов управления фондом скважин, реализованные в среде системы ELMA-BPM;
- веб-сервисы, реализующие модели ИНС и методы кластерного анализа, для решения трёх слабоформализуемых задач управления фондом скважин.

Также переданы разработанные структуры базы данных ИИС под управлением СУБД MS SQL Server 2012 и Описание применения ИИС.

В период опытной эксплуатации на данных одного из месторождений АО «Газпром добыча Томск» будет оцениваться эффективность моделей, методов и алгоритмов ИИС при управлении фондом скважин этого месторождения и возможность использования её в качестве ядра единого информационного пространства предприятия. Окончательное решение о внедрении ИИС будет принято по результатам опытной эксплуатации.

Начальник управления ИТ
АО «Востокгазпром», к.т.н. А.В. Сарайкин

Подпись Сарайкина А.В. удостоверяю,
и.о. начальника отдела по работе с персоналом
Управления по кадровой политике Л.П. Прокопьева

АКЦИОНЕРНОЕ ОБЩЕСТВО «ВОСТОКГАЗПРОМ» УПРАВЛЕНИЕ ПО КАДРОВОЙ ПОЛИТИКЕ ТОМСК

ПРИЛОЖЕНИЕ 2 АКТ О ВНЕДРЕНИИ РЕЗУЛЬТАТОВ В ООО «ЦЕНТР НЕФТЕГАЗОВЫХ ТЕХНОЛОГИЙ»

Общество с ограниченной ответственностью
«Центр нефтегазовых технологий»
634029, г. Томск, ул. Никитина 37Б, оф. 3, тел. 8-913-821-43-88
ИНН 7017292107 / КПП 701701001 ОГРН 11117017016016
р/с 40702810164000051430 в ОТДЕЛЕНИИ № 8616 СБЕРБАНКА РОССИИ г. Томск
БИК 046902606 к/с 30101810800000000606
stroyservis_07@mail.ru

УТВЕРЖДАЮ
Генеральный директор
ООО « ЦНГТ »
С.В. Федяева
«02» декабря 2020 г.

АКТ ВНЕДРЕНИЯ
результатов кандидатской диссертации Евсюткина И.В.

Настоящим актом подтверждаем, что в ООО «Центр нефтегазовых технологий» внедрены следующие результаты, полученные Евсюткиным Иваном Викторовичем при выполнении кандидатской диссертации по теме «Интеллектуальная информационная система для управления фондом скважин нефтегазодобывающего предприятия».

1. Концепция создания интеллектуальной информационной системы (ИИС) для управления фондом скважин на основе принципов SOA.
2. SOA-модель программного обеспечения ИИС.
3. Адаптивные алгоритмы предварительной подготовки геологических и технологических данных для анализа этих данных интеллектуальными методами.
4. Специализированная сервисная шина предприятия ESB.
5. Структура базы данных ИИС под управлением СУБД MS SQL Server 2017.
6. Веб-сервисы для реализации бизнес-процессов управления фондом скважин в среде модифицированной системы управления бизнес-процессами ELMA BPM.
7. Веб-сервисы сбора и актуализации геологических и технологических данных, реализующие адаптивные алгоритмы предварительной подготовки таких данных.
8. Модели глубоких искусственных нейронных сетей, методы кластерного анализа и реализующие их веб-сервисы для интеллектуального анализа данных при выборе скважин-кандидатов для геолого-технических мероприятий (ГТМ) и выборе ГТМ для таких скважин, а также для прогнозирования значений дебетов скважин.
9. Методика использования моделей глубоких искусственных нейронных сетей, реализованных в ИИС, для интеллектуального анализа данных.

Эти результаты были использованы при выполнении НИОКР для ряда предприятий-заказчиков и при создании единого информационного пространства ООО «ЦНГТ».

Начальник лаборатории ИТ — К.А. Мартынов

Ведущий инженер лаборатории ИТ — Т.И. Зайцева