

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ  
ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Уфимский государственный нефтяной технический университет»

На правах рукописи



**БАХИТОВ РИНАТ РАДИКОВИЧ**

**ПРОГНОЗ ОСНОВНЫХ ПАРАМЕТРОВ ЭКСПЛУАТАЦИИ  
СКВАЖИН НЕФТЯНОГО ПЛАСТА МЕТОДАМИ  
СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ И МАШИННОГО  
ОБУЧЕНИЯ**

Специальность 2.8.4. – Разработка и эксплуатация нефтяных и газовых  
месторождений

Диссертация  
на соискание ученой степени кандидата технических наук

Научный руководитель:  
доктор технических наук,  
профессор  
Хасанов Марс Магнавиевич

Уфа 2024

## Оглавление

ВВЕДЕНИЕ .....	4
1 АНАЛИЗ ИССЛЕДОВАНИЙ ПО МОДЕЛИРОВАНИЮ ГЕОЛОГИЧЕСКОГО СТРОЕНИЯ И ПРОЦЕССОВ РАЗРАБОТКИ МЕСТОРОЖДЕНИЯ.....	11
1.1 Классические методы и феноменологические подходы.....	11
1.2 Емкостно-резистивные модели - CRM.....	16
1.3 Статистические методы .....	26
1.4 Методы машинного обучения .....	30
1.5 Методы геостатистики .....	33
Выводы по первой главе.....	39
2 РАЗРАБОТКА МЕТОДИКИ ОЦЕНКИ ВЗАИМОВЛИЯНИЯ СКВАЖИН И ОПРЕДЕЛЕНИЯ СВЯЗАННОСТИ ПЛАСТОВЫХ СИСТЕМ.....	43
2.1 Характеристики объектов исследований.....	43
2.2 Предобработка входной информации на основе геофизических критериев .....	48
2.3 Методика синхронного анализа временных рядов.....	52
2.4 Разбиение месторождений на кусты .....	62
2.5 Инструмент дискретного скрытого марковского моделирования (ДСММ) для оценки точек переключения типов процессов .....	64
2.6 Анализ вариограмм и диаграмм рассеяния пространственных индексов.....	67
Выводы по второй главе.....	75
3 МОДЕЛИ ОЦЕНКИ ВЗАИМОВЛИЯНИЯ СКВАЖИН И СВЯЗАННОСТИ ПЛАСТОВЫХ СИСТЕМ .....	77
3.1 Алгоритм построения и тестирования модели байесовской векторной авторегрессии BVAR .....	77
3.2 Тестирование данных на наличие пространственных корреляционных зависимостей.....	85
3.3 Методика построения динамических моделей пространственной статистики с учетом отсроченного эффекта как по времени, так и по пространству ...	100
3.4 Вероятностное прогнозирование прерывистой динамики добычи жидкости обусловленными нормализующими потоками .....	107
Выводы по третьей главе .....	110
4.1 Методика формирования прогнозов добычи на скважинах месторождения .....	112
4.2 Подбор спецификации прогнозных моделей добычи на скважинах месторождения.....	114
4.3 Программная реализации методики получения прогнозов добычи на скважинах месторождения с учётом их взаимного влияния.....	127
4.4 Оценка экономической эффективности моделей прогнозирования с учетом взаимной связности .....	131
Выводы по 4 главе .....	133
ОСНОВНЫЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ .....	135
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	139
Приложение А.....	152

Приложение Б.....	154
Приложение С.....	156
Приложение D.....	158
Приложение Е.....	162
Приложение F.....	164

## ВВЕДЕНИЕ

### **Актуальность темы исследования**

Одной из важных задач при управлении разработкой нефтяных месторождений скважин является достоверный прогноз периода эффективной выработки запасов и сохранение рентабельной добычи нефти.

Прогнозирование дебита жидкости и нефти скважин на различных месторождениях с использованием традиционных методов, в том числе геолого-гидродинамического моделирования, характеристик вытеснения, не всегда обеспечивает качественный «сценарий». Основная проблема прогнозирования добычи углеводородов при сложном геологическом строении залежей заключается в уникальности каждого продуктивного пласта и, как следствие, затрудненность прямого переноса методик оценки свойств с одного эксплуатационного объекта на другой. В связи с этим создание методики прогноза, позволяющей получать достоверные прогнозные оценки независимо от горно-геологических условий залежей нефтяных и нефтегазовых месторождений, является актуальной задачей.

Важной проблемой современных исследований является получение достоверной информации о геологической связности эксплуатационного объекта, преимущественных направлениях фильтрационных потоков и анизотропии геологических свойств, продуктивности пласта для планирования стратегии разработки месторождения, в том числе с бурением уплотняющих скважин. Используемые в настоящее время модели и инструменты прогноза гидродинамической связанности пластовых систем на основе оценки взаимовлияния скважин друг на друга (метод материального баланса, полуаналитические емкостно-резистивные методы и др.) не в полной мере обеспечивают достоверную информацию. Кроме того, данные модели сложно настраиваемые, требуют больших ресурсов и имеют ряд ограничений. В современных условиях развития информационных технологий актуально использование моделей и алгоритмов для анализа дискретных временных рядов, а

также глобальных моделей из связанных временных рядов, применяемых в машинном обучении.

### **Степень разработанности темы исследования**

Значительный вклад в развитие разработки методов и подходов оценки взаимовлияния скважин в рамках одного объекта разработки и моделирования связанности пластовой системы внесли труды зарубежных и отечественных ученых. В частности, резистентно-емкостные модели (CRM) и их модификации разрабатывали Albertoni A., Weber D., Yousef A.A., Lake L., Sayarpour M., Хатмуллин И.Ф., Поспелова Т.А., Данько М.Ю. Модели на основе анализа временных рядов использовали Wiliantoro O., Apergis N., Bakari H.R., Wang O., модели на основе геостатистики разрабатывали Wigwe M.E., Cressie N., модели на основе инструментов машинного обучения и нейронных сетей использовали Gentil P.H., Akande K.O., Bansal Y., Demiryurek U., Nwachukwu A. Разработки этих ученых явились «фундаментом» для исследований, результаты которых представлены в настоящей диссертации. Область применений статистического моделирования машинного обучения постоянно расширяется в связи с повсеместной информатизацией производственных процессов. Однако практика их непосредственного применения свидетельствует о необходимости учета достоверной информации и важности предварительной подготовки данных, особенно динамических данных в виде дискретных временных рядов, а также возможностей по интерпретируемости результатов моделирования.

**Цель работы** – разработка методик оценки и прогноза основных параметров эксплуатации скважин нефтяного пласта с учетом их взаимовлияния с использованием статистического моделирования и машинного обучения.

### **Основные задачи исследований**

1. Выполнить литературный обзор используемых в настоящее время математических моделей и статистических алгоритмов для оценки взаимовлияния скважин и построения моделей связанности пластов, в том числе анализ применения методов аналитического и полуаналитического моделирования, анализа временных рядов, машинного обучения, нейросетевых моделей, моделей

пространственной статистики, учитывающих временное и пространственное запаздывание; выделить ограничения существующих подходов.

2. Разработать методику комплексного предварительного анализа входной информации геолого-промысловых данных работы скважин, позволяющую избежать ошибок моделирования.

3. Выбрать оптимальный метод формирования кустовых площадок на месторождении, учитывающий гидродинамическую (фильтрационную) взаимосвязь эксплуатационных объектов и взаимовлияния скважин.

4. Разработать достоверные модели прогноза дебита скважины, учитывающие взаимовлияние скважин по времени и по пространству, а также связанность пластовых систем изучаемой залежи нефти.

5. Разработать программное обеспечение и провести повариантные эксперименты численной реализации прогнозной модели с учётом взаимовлияния скважин месторождения.

### **Научная новизна**

1. Предложена матрица применимости алгоритмов и методов для разработки методики прогнозирования гидродинамической связанности пластов на основе оценки взаимовлияния скважин.

2. Разработана методика комплексного предварительного анализа входной информации по промысловым данным и параметрам работы скважин с использованием дискретных скрытых марковских процессов, позволяющая выделить участки однородности исторических рядов добычи нефти и существенно повысить достоверность их прогноза .

3. Получена модель байесовской векторной авторегрессии (BVAR), позволяющая комплексно описать динамические процессы взаимовлияния скважин в виде системы одновременных авторегрессионных уравнений с учетом эндогенных и экзогенных геологических свойств и параметров пластов, а также запаздывания динамических показателей скважин при их фильтрационной взаимосвязи.

4. Впервые разработаны панельные модели пространственного лага, позволяющие оценить взаимную связь добычи скважин с учётом влияния внутрипластового давления, забойного давления, пространственного и временного лагов запаздывания, а также индивидуального эффекта, присущего конкретной скважине и временному периоду.

#### **Теоретическая и практическая значимость работы**

Полученная модель байесовской векторной авторегрессии позволяет обосновывать спецификации статистических моделей, что обеспечивает качество мониторинга и прогноза дебита нефти.

Предложен подход к прогнозированию дебита скважины на основе комбинации глубоких нейронных сетей и классической модели авторегрессии, позволяющий достичь высокой точности прогноза в среднесрочной перспективе.

Предложенная комплексная методика обработки входной информации в виде исторических временных рядов по промысловым характеристикам работы скважин, расположенных на одном из месторождений Восточной Сибири, продемонстрировала высокую точность и устойчивость получаемых в дальнейшем прогнозов, позволяющая сделать вывод о возможности её использования в практике разработки месторождения.

Разработанный комплекс моделей прогнозирования и их реализация в виде программного обеспечения позволяет построить в реальном времени точные (с ошибкой менее 10%) прогнозы дебита добывающих скважин с горизонтом упреждения в 3 месяца с учётом их взаимного влияния. Это позволяет формировать эффективную стратегию управления разработкой месторождения и планировать, в том числе проведение геолого-технических мероприятий по интенсификации добычи нефти, консервации или ликвидации скважин.

Разработан программный код для построения численной реализации прогнозной модели BVAR в двух спецификациях: без дополнительных переменных – только на временных рядах дебита жидкости скважин; с дополнительными переменными – дебит скважин с учетом давлений. Программный модуль прошел проверку на период прогноза 30, 60 и 90 дней на

основе геолого-геофизической и промысловой информации месторождений и внедрен в промышленную эксплуатацию.

Разработанный программный модуль был интегрирован в информационную систему ООО «Газпромнефть НТЦ» (г. Санкт-Петербург) для анализа и прогнозирования добычи нефти.

### **Методология и методы исследования**

Настоящая работа выполнена на основе систематизации существующих исследований, методик и наработок из открытых информационных литературных источников в области применения математических моделей и статистических алгоритмов прогнозирования и оценки взаимовлияния скважин и построения моделей связанности пластовых систем, а именно применения инструментов анализа временных рядов, теории коинтеграции, панельного регрессионного анализа, моделей дискретных скрытых марковских процессов, метода кригинга и анализа вариограмм, нейросетевых моделей, моделей пространственной статистики, учитывающих временное и пространственное запаздывание.

Для численного эксперимента использовался комплекс различных исторических промысловых данных по скважинам месторождения Восточной Сибири. В прогнозных моделях взаимовлияния скважин и оценки связанности пластовых систем поведение каждой добывающей скважины описывалась в виде функции истории действия окружающих нагнетательных и добывающих скважин месторождения.

### **Положения, выносимые на защиту**

1. Методика комплексного предварительного анализа входной информации работы скважин позволяет выделить участки однородности в динамике исторических рядов добычи за счет применения алгоритма дискретных скрытых марковских процессов, а также оценить оптимальный лаг запаздывания промысловых характеристик за счет применения инструментов анализа временных рядов и теории их коинтеграции.

2. Статистические модели байесовской векторной авторегрессии (BVAR) и вероятностного прогнозирования DeepAR позволяют количественно



оценить связанность пластовых систем и построить прогноз взаимовлияния скважин исследуемых залежей.

3. Пространственные панельные модели позволяют количественно оценить взаимное влияние добычи нефти из скважин с учётом внешнего влияния внутрипластового давления, забойного давления, пространственного лага и временного лага запаздывания и выявить индивидуальные эффекты, присущие конкретной скважине и временному периоду ее эксплуатации.

### **Соответствие диссертации паспорту научной специальности**

Указанная область исследований соответствует паспорту специальности 2.8.4. «Разработка и эксплуатация нефтяных и газовых месторождений», а именно п. 9: «Научные основы создания цифровых двойников технологических процессов, используемых в компьютерных технологиях интегрированного проектирования и системного мульти-дисциплинарного мониторинга эволюции природно-техногенных систем, создаваемых для эффективного извлечения из недр или хранения в недрах жидких и газообразных углеводородов и водорода путем управления ими с использованием методов и средств информационных технологий, включая методы оптимизации и геолого-гидродинамическое моделирования».

### **Степень достоверности и апробация результатов**

Достоверность результатов обосновывается итерационным подходом при тестировании и адаптации моделей прогноза добычи жидкости и нефти из скважин и достижением приемлемой погрешности. Сравнение прогнозных данных с множеством фактических показателей работы скважин месторождения Восточной Сибири подтверждает высокую сходимость. Алгоритмы, методики и рекомендации, полученные в рамках представленной работы, прошли промышленную апробацию и получили положительный технологический и экономический эффект. Методика синхронизированного анализа временных рядов дебита нефти и жидкости прошла верификацию.

Основные положения и результаты диссертационной работы докладывались на Международной научно-практической конференции «Цифровые технологии и

бизнесе» (г. Уфа, 2023 г.), научно-практической конференции «Актуальные проблемы нефтегазовой отрасли» (г. Москва, 2022 г.), Международной научной конференции «Моделирование в инженерном деле 2021» (МИЕ2021) (г. Москва, 2021 г.), Международной конференции общества SPE «Arctic and Extreme Environments Technical Conference and Exhibition» (Москва, 2013 г.), а также реализованы при проектировании разработки Приобьского, Салымского и Куюмбинского нефтяных месторождений.

По результатам диссертационного исследования была зарегистрирована программа для ЭВМ: № 2023683879 - «Программа среднесрочного прогнозирования показателей продуктивности скважин с учетом влияния ее окружения», дата регистрации 10.11.2023 (совместно с Поповым Д.В., вклад автора диссертации - постановка задачи).

#### **Публикация результатов**

Основные результаты диссертационной работы отражены в 11 научных трудах, 5 из них опубликованы в ведущих рецензируемых научных журналах, рекомендованных ВАК при Министерстве науки и высшего образования РФ, и в изданиях, входящих в международную реферативную базу данных Scopus.

#### **Структура и объем диссертации**

Диссертационная работа состоит из введения, четырёх глав, основных выводов и рекомендаций, библиографического списка использованной литературы, включающего 117 наименования. Работа изложена на 183 страницах машинописного текста, содержит 43 рисунка, 20 таблиц, 6 приложений.

# 1 АНАЛИЗ ИССЛЕДОВАНИЙ ПО МОДЕЛИРОВАНИЮ ГЕОЛОГИЧЕСКОГО СТРОЕНИЯ И ПРОЦЕССОВ РАЗРАБОТКИ МЕСТОРОЖДЕНИЯ

Под моделированием пласта понимается моделирование гидродинамики потоков, получаемого как результат описания поведения эксплуатационных скважин, представленных данными дебитов и приемистостей производственной практики. Для разработки методики прогноза продуктивности пластовых систем необходимо определить модели и инструменты разработки такого прогноза, а также факторы, определяющие каузальное предсказание. По результатам проведённого предварительного анализа можно выделить несколько крупных направлений исследований: классический подход, полуаналитические ёмкостно-резистивные методы, статистические методы, методы машинного обучения, методы геостатистики [43]. Для понимания преимуществ и ограничений использования данных методов проведём их сравнительный анализ.

## 1.1 Классические методы и феноменологические подходы

Самыми распространёнными моделями, описывающими непосредственные физические законы течения жидкости и газа в пластовых системах с учётом их связности, являются системы дифференциальных уравнений в частных производных, называемые гидродинамическими моделями (ГДМ).

Математическая постановка задачи состоит в определении описывающих взаимовлияния скважин неизвестных характеристик, которые входят в систему дифференциальных уравнений, краевых и начальных условий. Существующие методы решения обратных задач позволяют определить робастность ГДМ пластовой системы, провести корректную параметризацию. Однако численная параметризация модели требует использования специальных алгоритмов с подбором оптимальных оценок регуляризирующего коэффициента. При этом особую роль здесь играет наличие дополнительно вводимой информации,

полученной в результате промышленного эксперимента. При этом важен учет фактора погрешностей в экспериментальных данных [16].

Несмотря на то, что «математический аппарат решения уравнений подземной гидромеханики разрабатывается и совершенствуется десятки лет, в конкретных практических ситуациях существуют значительные трудности» [30]: проблемы выбора модели пласта по конкретным кривым восстановления давления; «незначительные вариации в исходных данных, которые всегда присутствуют в замерах, вызывающие значительные изменения в конечных результатах» [38]. Ввиду больших трудозатрат на актуализацию значительных временных и технических ресурсов для реализации, ГДМ нельзя отнести к инструментам оперативного анализа и регулирования процессов разработки месторождений [101]. Данный метод изначально ориентирован на идентификацию механизмов пластовой системы и оптимальной степени её разнообразия [14].

Следующий подход – упрощённый метод материального баланса (МБ).

Авторы исследований [20, 51] отмечают преимущества метода материального баланса (МБ) в качестве мощного инструмент для оперативного исследования и понимания динамики пласта в целом. Модель МБ представляет простейшую форму ГДМ, которая основана на концепции сохранения массы: накопленная добыча нефти равна сумме изменений объёмов воды, нефти и пр.

Метод МБ, в противоположность разностным схемам гидродинамики, подразумевает «интегрирование», когда пласт рассматривается как нуль-мерный «черный ящик». Здесь используется принцип целостной оценки объекта вместо разбивки его на части и анализа каждый из них по отдельности в ГДМ. Эмпирическую базу исследования составляют только данные обобщённых характеристик по добыче, давлению и PVT-свойств. Необходимым условием применения методов МБ в прогнозировании является то, что на разных участках пластовой системы давление должно быть примерно одинаковым. При обратной ситуации польза от уравнения МБ снижается.

В многочисленных публикациях авторы пытаются усложнить вычисления, разделяя пласт на два или три блока и применяя материальный баланс к каждому из них. Такой подход опять же требует учёта перетока флюидов между блоками, а это порождает те же проблемы, что и в численном моделировании. В своих исследованиях авторы используют разные подходы для преодоления указанной проблемы [1, 40, 19].

Использование инструментов материального баланса имеют дополнительные ограничения по сравнению с гидродинамическим моделированием. Параметризация ГДМ позволяет настраивать вариацию параметров по сетке поиска таким образом, чтобы получить оптимальное соответствие модели свойствам пласта. Методом материального баланса такую «настройку» выполнить нельзя, но их использование позволяет предсказывать поведение резервуара. Если падение давления является практически одинаковым на различных участках пласта, то целесообразность применения прогнозных методов имеется, если же пласт неоднородный, то польза от уравнения материального баланса заметно снижается.

Численное моделирование и метод материального баланса не являются конкурентными методиками, а напротив, дополняющими друг друга: ММБ нужен для ретроспективного анализа, а ГДМ для прогнозного анализа.

Основной недостаток метода материального баланса заключается в том, он применим только для однородных систем с высокой проницаемостью. В случае использования матрицы взаимных продуктивностей для неоднородных пластов возникает задача с  $N^2$  неизвестными, которая неразрешима без дополнительной априорной информации.

Методы материального баланса схожи со статистическими методами в принципиальном подходе, который заключается в целостном описании объекта по входным, выходным и экзогенным характеристикам по типу «чёрного ящика». При этом в МБ закладывается физическая закономерность явления, которая интерпретирует преобразования входов в выходы для последующей оценки неизвестных параметров. Задача статистических методов заключается в

выявлении закона преобразования по совокупности объясняющих, выходных и экзогенных параметров объекта.

В научной литературе наряду с классическими методами активно развиваются феноменологические подходы для системного описания взаимовлияния добывающих и нагнетательных скважин пласта. Развивается понятийный аппарат, авторами используются термины – коэффициенты взаимовлияния и связности [31] в линейных уравнениях, связывающих входные и выходные параметры пласта. Для матричной формы введён термин «матрица взаимной продуктивности» (МВП) [41].

Авторы исследования [95] для вертикальных скважин в прямоугольном замкнутом пласте на псевдоустановившемся режиме доказали справедливость матричного уравнения в описании производительности многоскважинной системы:

$$\begin{pmatrix} q_1 \\ q_2 \\ \dots \\ q_N \end{pmatrix} = \begin{pmatrix} J_{11} & J_{12} & J_{1N} \\ J_{21} & J_{22} & J_{2N} \\ \dots & \dots & \dots \\ J_{N1} & J_{N2} & J_{NN} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \dots \\ d_N \end{pmatrix}, \quad (1.1)$$

где  $q_i$  – дебит  $i$ -й скважины;

$J_{ij}$  – элемент матрицы;

$d_i = \bar{p} - p_{wf\_i}$  – депрессия в  $i$ -й скважине;

$p_{wf\_i}$  – забойное давление в  $i$ -й скважине;

$\bar{p}$  – среднее пластовое давление;  $N$  – число скважин.

В матрице (1.1) элементы по диагонали  $J_{ii}$  являются коэффициентом продуктивности  $i$ -й скважины, остальные элементы  $J_{ij}$  отражают степень влияния на дебит  $i$ -й скважины изменения депрессии  $j$ -й скважины. Значение  $\bar{p}$  определяется из уравнения материального баланса и является константой. Недостатком данного подхода является то, что он применим для оценки взаимной связности в системах из множества скважин только в однородных пластах.

В работе Е.В. Юдина [50] предложен инженерный инструмент для анализа производительности многоскважинной системы на основе алгоритма идентификации коэффициентов МВП для общего случая неоднородности пласта.

Однофазная фильтрация в пласте, толщина которого значительно меньше его латеральных значений и изменяется относительно медленно, может быть описана следующим уравнением:

$$(\bar{k}h\nabla p) = \frac{h}{\chi} \frac{\partial p}{\partial t}, \quad (1.2)$$

где  $\bar{k}(x, y) = \frac{1}{h} \int_0^h k(x, y, z) dz$  – поле проницаемости, полученное усреднением проницаемости  $k(x, y, z)$  по вертикальной координате;

$h$  – толщина пласта;

$t$  – время;

$\chi$  – эффективный коэффициент пьезопроницаемости;  $\frac{1}{\chi} = \frac{h(C_f \bar{\varphi} + C_\varphi)}{\lambda_0}$ ;

$\bar{\varphi} = \frac{1}{h} \int_0^h (\varphi_0 dz)$  – усреднённая по толщине пористость пласта  $\varphi_0$ ;  $\lambda_0$  – удельная гидропроводность пласта;  $C$  – сжимаемость [38].

На внешних границах пласта заданы соответствующие условия поддержания постоянного давления или отсутствие перетока. Пласт эксплуатируется  $N$  скважинами с заданными произвольными забойными давлениями. Уравнение (1.2) в векторной форме можно переписать в виде:

$$\bar{q} = \hat{J} \cdot \bar{d}, \quad (1.3)$$

где  $\bar{q} = (q_1, q_2, \dots, q_N)^T$  – вектор дебитов;

$\bar{d} = (d_1, d_2, \dots, d_N)^T$  – вектор депрессий.

Основная сложность задачи заключается в том, что в общем случае вычислить коэффициенты  $J_{ij}$  аналитически не представляется возможным. Е.В. Юдин и соавторы [50] предложили использовать метод граничных элементов (МГЭ) решения уравнения (3) для оценки коэффициентов МВП  $J_{ij}$ . На основании полученного решения авторы находят депрессию  $d_i$  в скважинах, как алгебраическую сумму дебитов  $q_j$  и дебитов найденных мнимых источников  $\tilde{q}_j$ :

$$d_i = \sum_{j=1}^N a_{ij} q_j + \sum_{j=1}^{N_\Omega} b_{ij} \tilde{q}_j, \quad (1.4)$$

где  $N_{\Omega}$  – число ребер многоугольника, которым аппроксимирована внешняя граница пласта;

$a_{ij}, b_{ij}$  – функции взаимовлияния скважин и мнимых источников в бесконечном пространстве.

В матричном виде соотношение (1.4) имеет вид  $\vec{d} = \hat{A} \cdot \vec{q} + B \cdot \vec{q}$ . Тогда вычисляемый прямым счетом элемент матрицы  $\hat{A}^{-1}$  показывает влияние на дебит  $j$ -й скважины изменения депрессии в  $i$ -й скважине. Использование этой информации позволяет качественно оценить распределение фильтрационных потоков на рассматриваемом участке и определить степень интерференции скважин, используя единое среднее пластовое давление [50]. Однако возможности предложенного подхода ограничены самой постановкой задачи – однофазная фильтрация в пласте, толщина которого значительно меньше его латеральных значений, изменяется относительно медленно.

## 1.2 Емкостно-резистивные модели – CRM

Название емкостно-резистивных моделей (capacitance-resistance models, CRM) исходит из аналогии с резистивно-емкостными электросхемами, где напряжение батареи эквивалентно сигналу нагнетания. CRM модель на основе уравнений материального баланса и Дюпюи позволяет рассчитывать гидродинамическую связь между скважинами. Неизвестные параметры системы (коэффициенты взаимовлияния скважин, коэффициент продуктивности, временной параметр) вычисляются с помощью адаптации рассчитанного дебита скважин к истории разработки. В работе [36] показано преимущество применения для прогнозирования на практике CRM по сравнению с моделями гидродинамики в условиях отсутствия полной и надежной информации о строении пластов.

Авторы исследования [52] впервые предложили модель межскважинной связи в качестве попытки упростить моделирование месторождения до системы вводов (нагнетательных скважин) и выходов (добывающих скважин). В последующих вариантах емкостно-резистивных моделей (CRM) авторы



расширили число параметров модели и уточнили их физический смысл [113, 114]. Было предпринято несколько попыток использовать CRM для максимизации прогнозируемой добычи нефти [88, 103]. Для нахождения необходимых параметров модели требуется только исторические темпы закачки и общие темпы добыч. CRM не требует априорной оценки физических свойств пластов. Однако сопоставление расчётов по модели и исторических данных предоставляет ценную информацию о геологической связности объекта, преимущественных направлениях фильтрационных потоков, что может повысить эффективность управленческих решений в процессе заводнения.

Модель межскважинной связности характеризует зависящее от времени влияние нагнетательных скважин на добывающие с использованием коэффициентов связности и постоянных времени для каждой пары вход-выход аналогично линейным динамическим моделям, используемым в управлении химическими процессами. Контрольный объем определяется при этом как дренажная зона вокруг каждой добывающей скважины. Общий дебит добычи (нефти и воды) данной добывающей скважины определяется следующим уравнением непрерывности:

$$q_j(t) = \sum_{i=1}^{n_i} f_{ij} I_i(t) - \tau_j \frac{dq_j}{dt} \quad (1.5)$$

где  $q_j(t)$  – общая производительность (нефть плюс вода) добывающей скважины  $j$  в момент времени  $t$  [объем/время];

$I_i(t)$  – скорость закачки в скважину  $i$  в момент времени  $t$  [объем/время];

$f_{ij}$  – коэффициент связности между «нагнетателем»  $i$  и «производителем»  $j$  [безразмерная];

$\tau_j$  – постоянная времени, связанная с «производителем»  $j$  [время];

$n_i$  – общее количество «нагнетателей».

Физические коэффициенты связности представляют собой долю воды, закачиваемой в «нагнетатель»  $i$ , которая способствует получению как нефти, так и воды в «производителе»  $j$ .

Интегрируя уравнение (1.5) в дискретные периоды времени  $\Delta t$ , в течение которого скорость закачки всех «нагнетателей» принимается постоянной, и предполагая известной начальную скорость добычи для всех «производителей», получаем следующее аналитическое решение для скорости добычи в период времени  $k$ :

$$q_{jk}(t) = q_{j(k-1)} e^{-\frac{\Delta t}{\tau_j}} + \left(1 - e^{-\frac{\Delta t}{\tau_j}}\right) \sum_{i=1}^{n_i} f_{ij} I_{ik}. \quad (1.6)$$

Текущий общий уровень добычи «производителя»  $j$  является средневзвешенным показателем предыдущего уровня добычи и текущей скорости закачки воды у этого «производителя» от всех «нагнетателей». Здесь веса – это экспоненциальные члены, оба положительные и в сумме дают единицу.

Поскольку отсутствуют измерения внутрислоевых параметров, емкостно-резистивная модель использует тот факт, что имеются детальные данные о состоянии пластовой жидкости на скважинах. Количество параметров резко сокращается по сравнению с традиционным симулятором месторождения, хотя бы с десятков тысяч ячеек до сотен «добывающих» и «нагнетающих» скважин. Доля воды, закачиваемой в данную скважину, которая будет поступать в данную добывающую скважину, моделируется как коэффициент связности между двумя скважинами. Точно так же постоянная времени для добывающей скважины зависит от сжимаемости жидкости и соответствует времени, которое требуется для изменения скорости закачки, чтобы повлиять на нее. Модель может быть подогнана под выбранный временной интервал по фактическим данным или с использованием модельных прогнозов из традиционного имитатора месторождения.

Параметры модели (коэффициенты связности и постоянные времени) оцениваются с помощью нелинейной регрессии со следующей целевой функцией:

$$\min z = \sum_{k=1}^{n_t} \sum_{j=1}^{n_p} (q_{jk}^{obs} - q_{jk})^2, \quad (1.7)$$

где  $q_{jk}^{obs}$  – общая наблюдаемая скорость добычи «производителя»  $j$  на временном шаге  $k$ ;

$q_{jk}$  – рассчитанная суммарная скорость добычи «производителя»  $j$  на временном шаге  $k$ , определяемая уравнением (1.6) модели связности;

$n_t$  – это общее количество исторических периодов времени;

$n_p$  – это общее количество добывающих скважин.

Целевая функция, содержащаяся в уравнении (1.6), ограничивается следующими дополнительными условиями:

$$\sum_{i=1}^{n_i} f_{ij} \leq 1, \text{ для всех } j; \quad (1.8)$$

$$f_{ij}, \tau_j \geq 0, \text{ для всех } i, j. \quad (1.9)$$

Ограничение (1.8) представляет собой неравенство материального баланса, учитывающее потерю воды, закачиваемой в пределах контрольного объема, когда сумма прироста меньше, чем один. Поскольку коэффициенты связности в (1.8) суммируются по индексу «нагнетателей»  $i$ , применение этого ограничения требует оценки параметров модели для всех «производителей» одновременно. Невозможно минимизировать члены в (1.7), соответствующие каждому «производителю»  $j$  отдельно, когда присутствует (1.8), поскольку результирующая прибыль может не удовлетворять (1.8). Однако, если условием (1.8) пренебречь, то параметры каждого «производителя» могут быть найдены независимо, и эта процедура используется для определения начальных значений, которые обеспечивают хорошую начальную точку для решения общесистемной модели.

Продолжением модели CRM является вторая модель CRMIP, позволяющая разделить общий дебит добычи на дебит нефти и дебит добычи воды. Модель эмпирической степенной зависимости движения отдельных фаз, описанная [61], применима для уже освоенных месторождений. Для данной эксплуатационной скважины скорость добычи нефти ( $qo_{jk}$ ) определяется как

$$qo_{jk} = \frac{1}{1 + a_j (CWI_{jk})^{b_j}} q_{jk}, \quad \forall j, k, \quad (1.10)$$

где  $CWI_{jk}$  – вся вода, закачиваемая из всех нагнетательных скважин в пласте за все периоды времени, предшествующие периоду времени  $k$ , и достигнутый объем добычи нефти скважины  $j$  [объем], определяемый как:

$$CWI_{jk} = \sum_{k=1}^{n_t} \sum_{i=1}^{n_i} f_{ij} I_{ik} \quad \forall j, k; \quad (1.11)$$

$a_j$  и  $b_j$  параметры степенной модели движения отдельных фаз для каждой добывающей скважины  $j$ .

Модель в уравнении (1.9) может быть преобразована в линейную форму, определяемую как:

$$\ln \left( \frac{q_{jk}}{qo_{jk}} - 1 \right) = \ln \left( \frac{qw_{jk}}{qo_{jk}} \right) = \ln(a_j) + b_j \ln(CWI_{jk}), \quad (1.12)$$

где  $qw_{jk}$  – скорость добычи воды в добывающей скважине  $j$  в период  $k$ .

Особенностью модели учета факторов в степенном виде является возможность перевода ее в линейный вид за счет логарифмирования. Это значительно облегчает процедуру оценки параметров модели, которая включает только по два параметра каждой добывающей скважины. Конечно, существуют и другие модели движения отдельных фаз, и применимость этой конкретной модели к разным месторождениям может быть различной. Однако отсутствуют какие-либо ограничения, связывающие параметры движения отдельных фаз добывающих скважин, поэтому каждый из них может быть подогнан отдельно к историческим значениям натурального логарифма соотношения вода / нефть,

$$\ln \left( \frac{qw_{jk}}{qo_{jk}} \right).$$

Поскольку формула (1.10) справедлива только для более поздних временных данных (эмпирическое правило состоит в том, что отношение скорости производства воды к скорости производства нефти должна быть больше 0,5), то модель CRMIP обычно подходит для меньшего диапазона временного интервала, чем CRM.

Целевой функцией этой задачи для добывающей скважины  $j$  является

$$\min z = \sum_{k=k_{oil}}^{n_t} \left( WOR_{jk} - \left( \ln(a_j) + b_j \ln(CWI_{jk}) \right) \right)^2, \quad (1.13)$$

где  $WOR_{jk}$  – натуральный логарифм наблюдаемого отношения вода/нефть для добывающей скважины  $j$  в период  $k$ ;

$k_{oil}$  – первый шаг по времени, который рассматривается для соответствия модельной скорости добычи нефти;

$\ln(a_j)$  и  $b_j$  – переменные решения.

Единственное ограничение, накладываемое при минимизации (1.13), это

$$b_j \geq 0, \quad (1.14)$$

что обеспечивает увеличение соотношения вода/нефть по мере увеличения суммарно закачиваемой воды.

Таким образом, минимизация (1.13) с учетом (1.14) сводится к минимизации квадратичной функции двух переменных с учетом простых границ переменных. Данная задача легко может быть решена практически любым алгоритмом градиентного спуска.

Получаемые авторами многочисленных работ, например [110], результаты показали, что CRM в различных модификациях способна достаточно хорошо фиксировать общие тенденции добычи пласта. Описанные выше методы значительно уменьшают размерность задачи в обоих случаях. Алгоритмы выбросов удаляют небольшой процент точек данных из рассмотрения в соответствующих целевых функциях и приводят к небольшому увеличению значения коэффициента детерминации  $R^2$  как общего показателя оценки качества моделей в обоих случаях, но большому изменению значений соответствующих параметров.

Модель движения отдельных фаз также способна отражать общие тенденции данных, хотя ее пригодность не всегда одинакова для разных добывающих скважин. Поскольку модель движения отдельных фаз является одновременно простой и эмпирической, в некоторых случаях она может не быть адекватным описанием месторождения.

В дальнейшем множество исследований было посвящено снятию или ослаблению указанных ограничений. В [114] авторы показали возможность

изучения свойств пласта по данным закачки и добычи с помощью ёмкостной модели. Данная модель учитывает влияние сжимаемости, пористости и индекса производительности в нелинейной многомерной регрессии, где для описания временной задержки поступления сигнала закачки к добывающим скважинам используется постоянная времени. Рассматриваются два подхода: сбалансированная (темпы закачки и добычи по сумме закачивающих и производящих скважин равны) ёмкостная модель (VCM) и несбалансированная (имеется разница между темпами закачки/добычи) ёмкостная модель (UCM). Корректный подход будет зависеть от типа заводнённости и имеющихся данных. Оба подхода основаны на материальном балансе, использующем общие объёмы производства (вода+нефть+газ) и закачки в единицах объёма на единицу времени для каждой скважины. При этом в данных о характеристиках местоположения нет необходимости.

В работе [113] рассмотрено применение ёмкостных моделей к геологически сложным месторождениям. Гетерогенность месторождения определяется с помощью кривых Лоренца в координатах «ёмкость потока» и «толщина». Чем выше гетерогенность, тем сильнее отклонение кривой Лоренца от прямой. Коэффициент Лоренца непосредственно рассчитывается по площадям под кривой/прямой. Графики ёмкости потоков, полученные из ёмкостных моделей, отличаются от стандартной кривой Лоренца. Последние основываются на данных проницаемости и, возможно, пористости, измеренных с собранных образцов породы, причем пространственная связь между данными при этом игнорируется. Ёмкостные модели основаны на расчёте  $\lambda$  и  $\tau$  по данным добычи / закачки. Таким образом, графики ёмкостных моделей динамичны и вероятнее отмечают сильные течения и геологические особенности месторождений.

В [85] предложена модель OmegaCRM, как и другие CRM-модели, основанная на уравнении материального баланса и уравнении Дюпюи, но с учетом возможных остановок работы скважин. При этом во время остановок скважины модель рассматривает усеченное уравнение материального баланса без отбора, что позволяет восстанавливать среднее давление в объеме во время

простоев. Данная модель вводит поправку на коэффициент продуктивности после включения скважины на основе модели Чекалюка. На примере конкретного месторождения показаны преимущества использования модели OmegaCRM.

В [104] авторами представлены аналитические решения для фундаментального дифференциального уравнения емкостной модели, основанной на суперпозиции во времени, для трех различных вариантов целевой переменной:

- 1) объем добычи для всего месторождения;
- 2) объем добычи каждой скважины;
- 3) объем добычи между каждой парой «нагнетающая-добывающая скважины».

Три варианта целевой переменной выбраны для того, чтобы показать различные степени сложности, возникающие при моделировании. При получении аналитических решений отдельно рассмотрены две группы допущений: 1) допущение линейного изменения забойного давления при дискретном изменении коэффициента закачки и 2) допущение линейного изменения как коэффициента закачки, так и забойного давления в течение соответствующего интервала времени. Представленные авторами аналитические решения значительно расширяют возможности применения CRM для реальных данных с учетом их дискретного характера. При этом решения каждой аналитической модели дают оценки общего коэффициента добычи для всей жидкости (флюида). Поэтому требуется дополнительное соотношение, позволяющее разделить всю жидкость на нефть и воду. Для расчета добычи нефти авторы использовали эмпирическую модель дробного потока нефти. Эта модель учитывает степенной характер соотношения между мгновенным отношением воды к нефти  $F_{wo}$  и накопленным объемом закачанной воды  $i$ -ой скважины  $W_i$ . Оценка параметров моделей на основе ретроспективных данных осуществляется за счет минимизации средней абсолютной ошибки по всей истории добычи жидкости (или только нефти) в месторождении. В сбалансированной системе заводнения авторы рекомендуют выполнить оптимизацию параметров добычи одновременно для всех добывающих скважин. Таким образом, вместо минимизации ошибки оценки

коэффициента добычи для одной добывающей скважины следует минимизировать ошибку добыче по всем скважинам на месторождении.

В целом исследование [104] представляет новые доработки в моделировании CRM, аналитические решения на основе трех различных контрольных объемов, что облегчает применение модели в реальных условиях. Например, может быть получено общее представление о производительности заводнения во всем месторождении, а также лучшее понимание процессов на уровне связности скважин, что, в свою очередь, может применяться для исследования воздействия на месторождении одной или нескольких нагнетающих скважин на одну или несколько добывающих.

Для апробации модели авторы использовали два кейса смоделированных данных для имитации реальных условий при разных CRM. Были сгенерированы данные с широким диапазоном возмущений коэффициентов закачек и забойного давления, причем дополнительно были сгенерированы данные, включающие случайную ошибку. Параметры моделей оцениваются достаточно хорошо (как с ошибкой, так и без нее), тем самым демонстрируя способность CRM учитывать ошибки при измерении коэффициентов.

Таким образом, результаты моделирования показали, что откалиброванные CRM способны генерировать решения, сравнимые с решениями, полученными для трёхмерных потоковых моделей (flow-simulation models). Ёмкостно-резистивные модели способны быстро выполнять подгонку по историческим данным и прогнозировать последующую производительность, что приводит к перераспределению нагнетаемой воды и остановке малоэффективных нагнетающих скважин. Идентификация модели устойчива к ошибкам в измерении коэффициентов закачек.

В работе [82] исследуются возможности емкостной модели для оценки связности между добывающими и нагнетательными скважинами, причем особое внимание уделено недостаткам, присущим моделям данного вида, которые авторы пытаются преодолеть с помощью следующих модификаций модели:



– рассмотрением забойного давления как варьируемого неизвестного параметра (сегментированная ёмкостная модель);

– обобщением модели на случай изменения количества скважин в месторождении (компенсированная ёмкостная модель).

Главное достоинство модификаций заключается в том, что они могут быть использованы одновременно для построения единой модели. Модель была апробирована автором на двух синтетических кейсах и одном реальном месторождении (20 скважин: 13 добывающих, 7 нагнетающих), где полученные модификации помогли повысить коэффициент детерминации  $R^2$  на 15% по сравнению с обычной ёмкостной моделью.

В работе [47] коллектив отечественных авторов также рассматривает применимость моделей класса CRM и их модификации, адаптированных для применения в условиях реальных месторождений. В работе авторы особое внимание уделяют возможности учёта остановок скважин при оперативной оценке их гидродинамической связи. Преимущество использованного авторами подхода однозначно можно определить адекватностью модели реальному процессу добычи нефти.

Таким образом, CRM-модели являются хорошей альтернативой гидродинамическим моделям, которые позволяют учитывать относительно небольшой объем необходимой входной информации. Для идентификации параметров модели требуются только исторические данные, которые накапливаются компаниями. При этом количество параметров в сравнении с традиционными симуляторами месторождения значительно сокращается.

CRM не требуют никаких априорных оценок физических свойств пластов и модель быть подогнана под любой выбранный временной интервал по фактическим данным, либо по модельным синтетическим прогнозам из традиционного симулятора месторождения. Во всех указанных случаях результаты моделирования показывают, что откалиброванные CRM способны относительно быстро генерировать решения, сравнимые по качеству с решениями, полученными для трёхмерных потоковых гидродинамических

моделей (flow-simulation models). После подгонки параметров модель CRM может быть использована для прогнозирования добычи как для каждой отдельной добывающей скважины, так и месторождения в целом. Модель эффективна и для оптимизации системы управления нагнетательными скважинами посредством изменения скоростей закачки, перераспределения нагнетаемой воды и остановки малоэффективных нагнетательных скважин.

Однако у применения моделей CRM на практике есть ряд ограничений и недостатков. Необходимо отметить, что при выводе базовых моделей CRM используется ряд допущений, среди которых наиболее существенными и ограничительными являются следующие:

- постоянное число добывающих и нагнетательных скважин на всем рассматриваемом периоде времени (не учитываются простои скважин, смена режима работы (нагнетательная становится добывающей), введению в эксплуатацию новых скважин в многоскважинном пространстве, что практически нереально в практических условиях);

- неизменность пластовых и скважинных условий;
- постоянный коэффициент продуктивности скважин;
- необходимость рассмотрения продолжительного периода наблюдений;
- слабая корреляции приемистостей нагнетательных скважин.

В реальных условиях исходные данные могут быть неточными, зашумлёнными, ошибочными, что может приводить к значительным отклонениям полученного по модели значения дебита жидкости от фактического [34]. Кроме этого, задача оптимизации в рамках CRM будет иметь слишком большую размерность при моделировании месторождения с достаточно большим числом скважин.

### 1.3 Статистические методы

Выбор моделей при статистическом подходе во многом опирается на структуру данных. С учётом того, что мониторинг пластовых систем

подразумевает использование данных, представленных временными рядами, как и в CRM-моделях, то необходимо провести исследование применимости инструмента анализа временных рядов (TSA) в нефтегазовой области.

В работе [112] проведены исследования, связанные с применением параметрических и непараметрических моделей для прогнозирования добычи нефти и природного газа в анализе спада производительности скважин. Исследования проводились для 11 скважин, в качестве метрики использовался показатель AARD (absolute average relative deviation). Анализ показал, что непараметрическая модель ARIMA позволяет получить более точные прогнозы в случае наличия достаточной статистики по производительности скважин. В противном случае рекомендуется использовать параметрические модели (экспоненциального спада Арпса и Дуонга).

В работе [109] разработана NMGM-ARIMA модель (нелинейная модель Грея, объединённая с ARIMA) для прогнозирования добычи сланцевой нефти в США. В качестве метрик использовались показатели MAE (средняя абсолютная ошибка), MAPE (средняя абсолютная процентная ошибка), RMSE (квадратный корень из среднеквадратичной ошибки). Предложенная модель показала большую точность по сравнению с другими моделями (модель Грея, модель Грея, объединённая с ARIMA, нелинейная модель Грея, модель ARIMA).

В работе [58] для прогнозирования ежегодной добычи и потребления газа нефтяной компании из Нигерии предложена модель ARIMA. Результаты исследований показали, что лучшая модель для добычи газа – ARIMA (1, 1, 1), для потребления газа – ARIMA (0, 1, 1). Соответствующие цифры в скобках определяют вид модели ARIMA( $p, d, q$ ), где  $p$  – порядок авторегрессионной части модели, равный лагу запаздывания, учитывающий линейную комбинацию от предыдущих значений ряда;  $q$  – порядок части модели, относящейся к процессу скользящего среднего, равный лагу запаздывания в линейной комбинации следующих друг за другом белых шумов;  $d$  – порядок интеграции, равный порядку взятия разниц процесса до получения его стационарности в широком смысле.

Однако в последнее время особое значение приобретают исследования, посвящённые не только проблеме прогнозированию продуктивности скважин, но и исследования, посвящённые прогнозированию с учётом взаимного влияния скважин друг на друга. В работе [55] предложены методы анализа временных рядов, такие как проверка гипотезы наличия единичного корня в процессах, проверка гипотезы о наличии коинтеграции между двумя временными рядами, векторные модели коррекции ошибок для исследования динамической взаимосвязи между добычей нефти, количеством буровых установок и ценами на сырую нефть для шести регионов нефтедобычи США.

В работе [72] для прогнозирования добычи нефти предложен метод SAM-oil, который объединяет алгоритм имитации отжига с эволюционными вычислениями, статистический анализ и моделирование остатков. В качестве метрики использована SMAPE (симметричная средняя абсолютная процентная ошибка). Предложенная модель показала большую точность по сравнению с другими моделями (модель анализа кривой спада, модель ARIMA, нейронные сети, модели экспоненциального сглаживания) как на всем наборе данных, так и для некоторых отдельных кластеров.

В работе [116] для анализа и прогнозирования добычи нефти предлагается использовать инструмент анализа временных рядов – модель ARIMA. Предварительно построены две модели ARIMA (1, 0, 0) и (2, 0, 0), выбор между которыми осуществлялся по информационным критериям Акайке и Шварца. В результате для прогноза, согласно критериям, выбрана первая модель. В качестве метрик использованы показатели MAE (средняя абсолютная ошибка), MAPE (средняя абсолютная процентная ошибка), RMSE (квадратный корень из среднеквадратичной ошибки) и др.

В работе [71] проведены исследования, связанные с применением различных моделей прогнозирования для различных видов ископаемого топлива в Турции: модель кубической регрессии используется для прогнозирования каменного и бурого угля, модель ARIMA – для асфальтита и природного газа, модель SARIMA – для нефти.

В работе индонезийских авторов Сухартоно и др. [106] для прогнозирования добычи нефти на нескольких буровых скважинах Восточной Явы использовались модели векторной авторегрессии (VAR), позволяющей учесть взаимное влияние дебита добычи скважин друг на друга. Однако количество скважин, на которых обучалась модель, крайне мало (всего 3 скважины), а ежемесячная динамика составляла всего 60 наблюдений, соответственно, сделать выводы о возможности универсального применения данной модели не представляется возможным. Однако именно данная методика имеет наибольшие перспективы при получении прогноза связанности пластовых систем на основе оценки взаимовлияния скважин друг на друга при условии корректного объёма данных для обучения соответствующих моделей.

В работе [89] представлена улучшенная авторегрессионная модель связанности скважин с учётом влияния экзогенных переменных (M-ARX), показывающая связь между добывающими и нагнетательными скважинами. Модель была апробирована на 9 скважинах (5 – нагнетательные, 4 – добывающие). Превосходство модели M-ARX над ёмкостной моделью подтверждается численным моделированием при различных сценариях. Указанная модель может прогнозировать производительность месторождения с гораздо меньшим количеством параметров, когда вводятся некоторые искусственные ограничения, например остановка работы добывающей скважины. Наконец, модель может быть легко расширена до более высокого порядка для обработки более сложных и разнородных сценариев.

В работе [52] Альбертони и Лейк одними из первых использовали метод линейной многомерной регрессии с диффузионными фильтрами для прогнозирования общего дебита скважины на основе коэффициентов закачки, используя данные по закачке и добыче.

Автор работы [115] пытается устранить некоторые ограничения моделей, предложенных в [52], в частности устранить возможность наличия отрицательных коэффициентов, возникающих в результате наличия мультиколлинеарности и не имеющих физического смысла. Автор предлагает в качестве решения

использовать гребневую регрессию, что позволяет улучшить представление о связности скважин. Для демонстрации результатов была построена модель на трех скважинах: 1 добывающая и 2 нагнетательных скважины. Также модель применялась к следующим реальным данным: The Chihuido de la Sierra Negra Field: 25 нагнетательных и 16 добывающих скважин; Magnus field: 13 нагнетательных и 14 добывающих; The South Wason Clear Fork Field: 18 нагнетательных и 13 добывающих; The North Buck Draw field: 8 нагнетательных и 12 добывающих. Во всех случаях метод, предложенный в [115], обеспечивал лучшее понимание связности скважин и неоднородностей месторождений, чем предыдущие методы.

#### 1.4 Методы машинного обучения

В последнее время стало популярным использование методов интеллектуального анализа данных, в частности инструментов машинного обучения для моделирования связности пластов и продуктивности скважин. В работе [105] проведены исследования, связанные с прогнозированием производительности скважин, добывающих нефть. В качестве метрики оценки качества прогноза использовался показатель среднеквадратической ошибки. Входными параметрами являются геологические, геофизические параметры и технические характеристики. Для предварительного анализа данных используются: функциональный анализ данных для сглаживания временных рядов и метод главных компонент для снижения размерности. Для прогнозирования производительности скважин применялся алгоритм случайного леса.

Следует отметить, что для решения задачи прогноза связанности пластовых систем на основе оценки взаимовлияния скважин друг на друга часто применяются искусственные нейронные сети (artificial neural network, ANN). Основная проблема такого подхода заключается в переобучении временных рядов. Например, в работе [61] была применена методология, основанная на ANN,

к месторождению, расположенному в Западном Техасе. Литология коллектора представляла собой сочетание детритовых карбонатов, алевролитов и черных сланцев. Исследуемый пласт является довольно прерывистым с несвязанными нефтяными карманами. Авторами разработана авторская архитектура сети. В этой сети было использовано в общей сложности 137 скважин (87 старых скважин + 50 новых скважин). Из них 111 скважин были использованы для обучения ANN, а 13 скважин – для тестирования и валидации ANN. Авторы исследования утверждают, что такие свойства, как сложная геология, неподключенные углеводородные карманы и доступность данных, делают его хорошим кандидатом для применения ANN.

В работе [51] представлено сравнительное исследование и анализ производительности обобщения и прогнозирующей способности методов машинного обучения (ML) при характеристике и моделировании пласта с использованием двух различных наборов данных по нефтяным скважинам. Проницаемость карбонатного пласта была предсказана в этой исследовательской работе с использованием методов искусственного интеллекта. Производительность обобщения и прогнозирующая способность методов применительно к прогнозированию проницаемости сравнивались для отражения их конкретных сильных и слабых сторон. Показано, что ANN (искусственная нейронная сеть) превосходит SVR (регрессию опорных векторов) в случае большого набора данных, в то время как SVR работает лучше в сценарии с небольшим набором данных. Кроме того, для обоих алгоритмов реализована технология выбора признаков, которая значительно повышает их производительность, что приводит к уменьшению количества функций, сокращению вычислительного времени и ресурсов. Можно сделать вывод, что ANN больше подходит для задач с большим, а SVM – для задач с относительно малым набором данных. При этом подход выбора признаков следует проводить на основе корреляции в целях повышения производительности обоих методов машинного обучения.

В [92] авторы построили ANN для оценки межскважинных связей на базе синтетических и реальных данных. Для ускорения подбора гиперпараметров модели использовался генетический алгоритм. Результатом данной модели стали коэффициенты силы межскважинных связей, значения которых коррелировали с геологическими данными и результатами трассировочных исследований месторождения JD (Китай).

Работа [56] сравнивает эффективность ANN в сравнении с CRM-моделью. Оба подхода позволили быстро построить карту связности между скважинами примерно одинакового качества. Однако, как отмечено автором, ANN выигрывает в гибкости к требованиям к исходным данным.

Авторы в [67] произвели анализ чувствительности ANN для определения силы связи между парами закачивающих-производящих скважин. В числе прочего данный подход позволил выявить наиболее и наименее эффективные закачивающие скважины. Помимо ANN и SVM, для решения данных или похожих задач могут быть использованы другие подходы машинного обучения. Так, в [100] авторы использовали градиентный бустинг для выбора местоположения скважин с учётом межскважинных связей. Данный подход с дополнениями был повторён в [99].

В работе [94] представлено использование анализа по дереву классификаций и регрессии (CART) для оценки межскважинной связности на примере месторождения Дацин. Классификация производилась по 10 характеристикам на 20000 наблюдениях. В результате точность классификации связности достигла 85,2%, а продолжительность расчётов – около 50 секунд. В [91] также было использовано дерево классификаций и регрессии для анализа связанности. Классификация производилась по 8 характеристикам, на 2000 тестовых наблюдениях и достигла точности от 80 до 90%.

Для этапа выбора характеристик скважин, на основе которых будет производиться исследование их связанности, можно использовать классификатор случайного леса, что было продемонстрировано в работе [87].



Одним из подходов машинного обучения является иерархическая кластеризация, которая может быть использована и на задачах геостатистики, как показано в [74]. Однако непосредственно к задаче оценки межскважинной связности данный подход не применялся.

## 1.5 Методы геостатистики

Одним из подходов, позволяющих решить задачи оценки и прогнозирования связности пластов и продуктивности скважин, являются методы геостатистики. Суть подхода в том, чтобы, опираясь на имеющиеся данные, находить закономерности пространственных изменений показателей и использовать эти знания для получения наилучших оценок по всему месторождению. Геостатистика основана на простом принципе о том, что геологические данные пространственно-коррелированы [65]. Пространственная корреляция измеряется и используется для определения весов и оценки значений интересующих показателей в точках, где бурение не проводилось [83].

В нефтяной промышленности используются такие методы геостатистической и стохастической оценки, как кригинг, кокригинг, условное моделирование, фрактальные методы и метод вероятностей перехода (Марковские процессы). Любой из этих методов может быть использован в качестве модели для описания пространственной структуры показателя. Одним из преимуществ использования методов крикинга и кокринга является возможность восполнения пропусков в данных или расчёта оценок свойств пласта, которые нужны в последующем для построения корректных статистических моделей. Например, в [62]. В работе [98] используется уникальный подход кригинга для оценки распределения пористости и проведения условного гауссова моделирования для генерации распределения пористости в месторождении Мансури.

Кригинг, или наилучшее линейное несмещённое предсказание (BLUP) стало очень популярным в науках о земле и об окружающей среде. Сам термин

«кригинг» был введен в честь Д. Г. Криге, южноафриканского горного инженера. С его внутренней количественной оценки пространственной изменчивости через ковариационную функцию (или вариограмма) методология кригинга может производить карты оптимальных прогнозов и связанные стандартные ошибки прогнозирования из неполных и зашумлённых пространственных данных. Иногда получить пространственные данные дорого (например, бурение скважин для оценки добычи нефти), в этом случае размер выборки  $n$  обычно мал, и может быть выполнен кригинг прямолинейно. В последнее время, благодаря повсеместному распространению платформ дистанционного зондирования на спутниках, парадигмы баз данных перешли от небольших к массовым (порядка гигабайт в день). Решение уравнений кригинга непосредственно включает инверсию (обращение)  $n \times n$  – ковариационной матрицы  $\Sigma$ , где для  $n$  данных может потребоваться  $O(n^3)$  вычислений для получения  $\Sigma^{-1}$ .

В целом кригинг – это метод для вычисления оценки регионализованной переменной в точке, в области или в объёме. Метод взвешенного скользящего среднего, основанный на двух критериях минимизации дисперсии и несмещённого условия, представлен в [67]. Преимущество кригинга состоит в вычислениях доверительного интервала для каждой точки. В классической статистике это невозможно. Кригинг даёт наилучшие несмещённые эффективные оценки [63]. Эта оценка определяется следующим образом:

$$Z_k^* = \sum_{i=1}^n \lambda_i Z(x_i) \quad (1.15)$$

где  $Z(x_i)$  – выборочное значение (в данном случае – пористость) и  $\lambda_i$  – весовой коэффициент для точки  $i$ ;

$Z_k^*$  – оценка кригинга.

Для расчета весовых коэффициентов  $\lambda_i$  используется ранее описанный критерий.

Геостатистическая оценка состоит из двух этапов: сначала вычисление и моделирование полувариограммы, а затем – оценка интересующего параметра. Полувариограмма определяется как половина дисперсии приращений. Однако,

когда присутствуют сильно коррелированные структуры, такие как скрытые каналы или трещины, невозможно непосредственно представлять их с помощью одной вариограммы, и, если они были выявлены, то предпочтительно использовать метод геологически информированного зонирования. Кроме того, расширение кригинга, т.н. кокригинг, может быть использован для оценки проницаемости на основании результатов испытаний скважины и измерений дополнительных показателей (например, удельной ёмкости, удельного электрического сопротивления и так далее). Необходимо отметить, однако, что кригинг не всегда работает одинаково хорошо [95].

Авторы [64] для применения кригинга к данным из большой пространственной области используют так называемый кригинг с фиксированным рангом с целью снижения вычислительных затрат на кригинг до  $O(n)$ . Кригинг с фиксированным рангом является кригингом в некотором классе нестационарных ковариационных функций. Он основан на вычислительных упрощениях, когда  $n$  очень велико, для получения наилучшего пространственного линейный несмещенный предиктор и его среднеквадратичной ошибки прогноза для скрытого (невидимого) пространственного процесса. Метод, основанный на минимизации взвешенной нормы Фробениуса, даёт наилучшие оценки параметров ковариационной функции, которые затем подставляются в уравнения кригинга с фиксированным рангом.

Для значительных объёмов данных (порядка десятков или сотен тысяч) прямой кригинг может расходиться, и тогда обычно используются специальные локальные кригинг-районы за счет разложения эквивалентного представления в терминах ортогональных базисов и усечённых базисов либо выполнение ковариационного сужения, используя приближенные итерационные методы, такие как сопряжённый градиент, реализующий разреженные аппроксимации с индуцирующими переменными или заменяющими местоположения данных меньшим набором. В работе [80] авторы воспользовались данной идеей при подгонке класса пространственных моделей, которые они назвали геоаддитивными моделями. Другой подход заключается в выборе классов

ковариационных функций, для которых кригинг может быть произведён с достаточной точностью за разумное вычислительное время даже в случае наличия большой размерности данных [79]. В этих статьях многоуровневый пространственный (и пространственно-временной) процесс был построен таким образом, что (простой) кригинг мог быть вычислен итеративно и чрезвычайно быстро. В пространственном случае было достигнуто ускорение порядка  $10^8$  раз относительно непосредственного решения уравнений кригинга. Результаты авторов основаны на использовании класса нестационарных ковариационных функций, которые возникают из пространственной модели случайных эффектов. Байесовский анализ позволяет осуществлять оптимальное пространственное предсказание в нелинейной геостатистике, модели такого рода, были рассмотрены в [68]. Хотя подобные модели могут быть вычислительно затратными для больших пространственных наборов данных, благодаря методам Монте-Карло по схеме цепей Маркова, которые используются в анализе, все еще есть возможность достичь вычислительных результатов ускорения. В качестве доказательства этого, в [76] авторы сравнили стандартную геостатистическую ковариационную модель с моделью, в которой обратная ковариационная матрица была смоделирована непосредственно, что привело к увеличению вычислительной эффективности более чем в 5 раз.

К геостатистическому подходу также следует отнести моделирование Монте-Карло водоносного горизонта, где параметры представляют собой условные реализации неизвестного пространственного распределения параметров. Метод позволяет оценить степень неопределенности системы в предположении, что условные реализации отображают полный спектр неопределённости водоносного горизонта, и что ковариация способна описать всю соответствующую неоднородность. В этих условиях моделирование потока по методу Монте-Карло дает корректную оценку результирующей неопределённости в отношении потока (напор, расход, концентрация и т. д.) [117].

Отдельной задачей является определение гипотез и факторов различной природы, влияющих на продуктивность месторождения и связанность скважин.

Так, в работе [77] выделяют следующие факторы, влияющие на продуктивность добычи нефти: степень извлечения нефти, композитная обводненность, средний перепад давления производства и процент извлечения нефти.

В работе [107] выделяют две группы факторов, влияющие на продуктивность многоразрывных горизонтальных скважин:

1) резервная пропускная способность (толщина продуктивной части пласта, пористость, проницаемость, нефтенасыщенность, гамма-каротаж, длина песчаника в горизонтальном сечении);

2) эффективность завершения (количество стадий, длина трещины, средний объем жидкости для гидроразрыва пласта, количество расклинивающего агента в среднем, эффективность обратного потока).

В работе [78] выделяют факторы, влияющие на продуктивность добычи нефти: общее количество скважин, ежемесячная закачка воды, полное содержание влаги, дебит нефти, процент извлечения, время эксплуатации.

Относительно новым анализом, применяемым в геостатистике, является пространственное регрессионное моделирование, то есть построение моделей с пространственным лагом, которая учитывает наличие связи в пространственном распределении объектов. В данных моделях вводят понятие пространственного лага между двумя объектами, который определяется либо как бинарный признак соседства двух объектов, либо как расстояние между объектами. Когда речь идет о взаимовлиянии скважин в рамках одного месторождения, ключевым, безусловно, становится их месторасположение. В отдельных случаях пространственная компонента в моделях может быть основной при несколько менее влияющей временной составляющей. Такая ситуация в работе [57] объясняется благоприятными пластовыми условиями. Пространственная модель может содержать полезную информацию о ненаблюдаемых факторах, влияющих на интересующую переменную (например, на продуктивность скважины). Если модель неправильно определена и не способна принять во внимание эту пространственно структурированную информацию, это может привести к пространственно автокоррелированным остаткам, что указывает на

несостоятельность модели. Например, авторы [96] отметили, что если выбранный подход к моделированию не включает пространственную зависимость с высоким разрешением, может быть введена значительная погрешность, что приведет к переоценке относительного влияния технологии на повышение производительности скважины.

Учесть пространственные зависимости можно несколькими вариантами. Так, в работе [111] построены географически взвешенная регрессия (GWR) и универсальный кригинг. Montgomery и O'Sullivan (2017) [96] построили модели, контролирующие пространственную неоднородность с помощью фиксированных эффектов на уровне территориальных участков анализа поверхностных трендов (STA), а также модели с учетом пространственной автокорреляции, такие как регрессионный кригинг и модель пространственной ошибки (SEM). В работе [111] были сравнены кригинг с фиксированным рангом (FRK) и пространственно-временная обобщенная аддитивная модель (GAM). В их работе отмечено, что успешна именно пространственно-временная оценка добычи, независимо от конкретного изучаемого бассейна. Важность одновременного учета и пространства и времени определяется тем, что объем добычи нефти изменяется, так дебит скважины обычно достигает пика в первые несколько месяцев, а затем начинает снижаться.

## Выводы по первой главе

Результат проведённого анализа литературных источников позволяет сделать вывод о том, что ёмкостно-резистивное моделирование, применение методов геостатистики (кригинга), статистическое пространственное моделирование, модели на основе временных рядов способны без обращения к классическим физическим моделям получить информацию для прогнозирования пространственно-временного взаимовлияния скважин месторождения. Перспективны модели на основе временных рядов (в частности модели связности временных рядов VAR, BVAR, VECM, ARIMAX), их панельные аналоги, а также регрессионные модели пространственного лага (SAR, SAC, SEM, SDM). Для получения оценочных значений и восполнения данных о свойствах пласта в моделях пространственной регрессии целесообразно применять методы кригинга и кокригинга. Важен комбинированный подход, в качестве сравнения прогнозных свойств построенных статистических моделей целесообразно использование моделей искусственных нейронных сетей (ANN), хотя они существенно проигрывают статистическим моделям с точки зрения объяснимости процессов и последствий переобучения. При построении модели на всем множестве связанных временных рядов чрезвычайно важно сопровождение исследования синхронным анализом для выявления характерных закономерностей как каждой скважины, так существующего временного запаздывания в обусловленном взаимном влиянии скважин.

На основании проведённого литературного обзора была составлена матрица применимости методов и подходов для формирования методики прогноза связанности пластовых систем на основе оценки взаимовлияния скважин друг на друга (Таблица 1.1).

Таблица 1.1 - Матрица применимости методов/алгоритмов по определению связанности пластовых систем

Метод/ алгоритм	Достоинства	Недостатки/Ограничения	Необходимая входная информация	Оценка применимости
<b>Метод материального баланса</b>	1.Простота и эффективность применения для однопластовых систем. 2. Классический подход, хорошо интерпретируемый с физической точки зрения	1.Значительная погрешность в коллекторах с неоднородной структурой или с неустановившимся/постоянно меняющимися режимами. 2.Межскважинное пространственное взаимодействие учитывается в виде матрицы взаимных продуктивностей, которую сложно определить эмпирически 3.В ММБ отсутствует переменная времени, то есть модель позволяет спрогнозировать то, что будет, но не определяет дату когда это произойдет	Толщина пласта, пьезопроницаемость, усредненная по толщине пористость, удельная гидропроводность	Метод не может быть применен в силу существенных ограничений по оценке матрицы взаимных продуктивностей
<b>Модели машинного обучения</b>	Доказано, что искусственные нейронные сети в отдельных случаях превосходят по эффективности другие методы по оценке взаимовлияния скважин Меньше требований к наборам исходных данных (можно работать с практически любыми данными)	Требуется достаточно большой массив данных для обучения и тестирования моделей. Многие модели, основанные на применении инструментов ML, склонны, либо к недообучению, либо к переобучению. Построение качественных и точных моделей требует больших компьютерных ресурсов Не дают информации о реальных причинах, которые приводят к тому или иному результату (отсутствие физических объяснений реальных процессов)	Временные ряды показателей эксплуатации скважин, достаточной длины для построения моделей динамики Геологические и гидродинамические параметры	Могут быть применены в исследовании (в том числе для получения синтетических образцов скважин с заданными показателями)



продолжение таблицы 1.1

Метод/ алгоритм	Достоинства	Недостатки/Ограничения	Необходимая входная информация	Оценка применимости
<b>Емкостно-резистивные модели</b>	<p>1.Позволяют получить достаточно корректную информацию о влиянии закачки на продуктивность (всего месторождения, отдельных скважин);</p> <p>2.Не требуют значительных временных и вычислительных ресурсов, основаны на методах численной линейной оптимизации;</p> <p>3.Относительно небольшой объем необходимой входной информации.</p> <p>4.Уменьшение количества параметров в сравнении с традиционными потоковыми симуляторами месторождения;</p> <p>5.Нет необходимости в априорной оценке физических свойств пластов;</p> <p>6.Модели могут быть подогнаны под любой выбранный временной интервал</p>	<p>1.Достаточно много предпосылок, при которых модель «работает» (линейный закон изменения продуктивности скважины от изменений скоростей закачки, забойное давление постоянное, число скважин и характер их работы не меняется, отсутствуют остановки и повторные запуски) и как следствие значительные требования к выбору временного интервала для подгонки и тестирования; подходящий по условиям временной интервал данных для моделирования зачастую достаточно короткий</p> <p>2.Возможны некорректные оценки коэффициентов связности при наличии мультиколлинеарности скоростей закачек;</p> <p>3.Зашумленность/неточность исходных данных ведет к искажению значений параметров модели; для устранения зашумленности, а также с целью уменьшения размерности задачи на крупных месторождениях часто прибегают к таким средствам, как введение дополнительных переменных, агрегирование и трансформация данных, - например, агрегирование данных по скважинам, которым соответствует относительно небольшая доля закачки/добычи в одну;</p> <p>4.Получаемые в результате моделирования коэффициенты связности фактически не отражают взаимовлияния закачивающих скважин, а только связи между закачивающими и добывающими скважинами;</p>	<p>Для рассматриваемых скважин исторические ретроспективные данные по скоростям закачки и коэффициентам добычи (жидкость, нефть, вода); забойное давление</p>	<p>Применение метода возможно с целью сравнения с другими методами оценки взаимного влияния скважин Ограничения применимости метода достаточно сильные для возможности использования их в данной задаче в качестве основного инструмента (основные сложности: выбор временного интервала, изменение режимов работы скважин, наличие горизонтальных скважин в месторождении)</p>

продолжение таблицы 1.1

Метод/ алгоритм	Достоинства	Недостатки /Ограничения	Необходимая входная информация	Оценка применимости
<b>Статистические модели (модели геостатистики и модели анализа временных рядов)</b>	<p>1. Статистические модели временных рядов в отдельных случаях позволяют получить более точные прогнозы с учетом взаимного влияния скважин в случае наличия достаточной статистики динамического характера по производительности скважин;</p> <p>2. Модель содержит относительно небольшое число параметров в случаях ввода некоторых искусственных ограничений, например, остановка работы добывающей скважины;</p> <p>3. Модель может быть легко расширена до более высокого порядка для обработки более сложных и разнородных сценариев</p> <p>4. Модели пространственной регрессии позволяют количественно оценить взаимосвязанность скважин с учетом расстояния между ними.</p> <p>5. Распространённые среды статистического моделирования (открытые ресурсы) позволяют легко проводить анализ</p>	<p>1. Зависимость от структуры данных;</p> <p>2. Оцененные коэффициенты моделей не всегда поддаются физической интерпретации</p> <p>3. Требуется достаточно большой объем данных для возможности «обучения» статистических моделей их валидации</p> <p>4. Отсутствует качественный подход по учету ГТМ, учету переводов скважин (из добывающих в нагнетательную), требование по кластеризации скважин на сектора, отсутствие физического смысла связности – все влияющие факторы заносятся в одну переменную связанности</p>	<p>Временные ряды показателей эксплуатации скважин, достаточной длины для построения моделей динамики</p>	<p>Могут быть применены в исследовании с использованием современных сред статистического моделирования (R Studio)</p>

## 2 РАЗРАБОТКА МЕТОДИКИ ОЦЕНКИ ВЗАИМОВЛИЯНИЯ СКВАЖИН И ОПРЕДЕЛЕНИЯ СВЯЗАННОСТИ ПЛАСТОВЫХ СИСТЕМ

Ввод в эксплуатацию трудноизвлекаемых запасов, проблема повышения эффективности активов требуют дальнейшего развития современных цифровых технологий анализа данных [6]. Под этим термином предполагается извлечение в результате численной обработки из множества показателей ценной информации о характере пласта, геологической связности объекта, преимущественных направлениях фильтрационных потоков и анизотропии геологических свойств его продуктивности. Необходима системная и комплексная методика анализа откликов на входные сигналы по всему резервуару как средство проверки гипотез и решения задач исследователя [9].

### 2.1 Характеристики объектов исследований

Объектами исследования являлись несколько месторождений. Первое из них (№1) состоит из 115 синтетических скважин, из которых добывающих – 84. На Рисунке 2.1 изображён график плотности распределения начала работы скважин месторождения №1, из которого видно, что в 1989 г. и 1991 г. было активное освоение месторождения с максимальным запуском скважин. С 1993 г. прекратился запуск новых скважин, добыча осуществлялась из уже запущенных. С 1994 г. наблюдалось некоторое оживление по вводу новых скважин. Второе исследуемое месторождение представлено ежемесячными данными, которые отражают динамику пласта с середины 1980-х по 1999-й год.



Рисунок 2.1 – Плотность распределения начала работы скважин месторождения №1

На Рисунке 2.2 представлены графики по уровням добычи нефти и жидкости на месторождении № 1 за весь период наблюдения. Если до 1990 г. объемы добываемых нефти и жидкости сравнимы, то с 1991 г. объем выкачиваемой жидкости резко возрос при плавном снижении объема добываемой нефти и истощении месторождения к 2000 г.

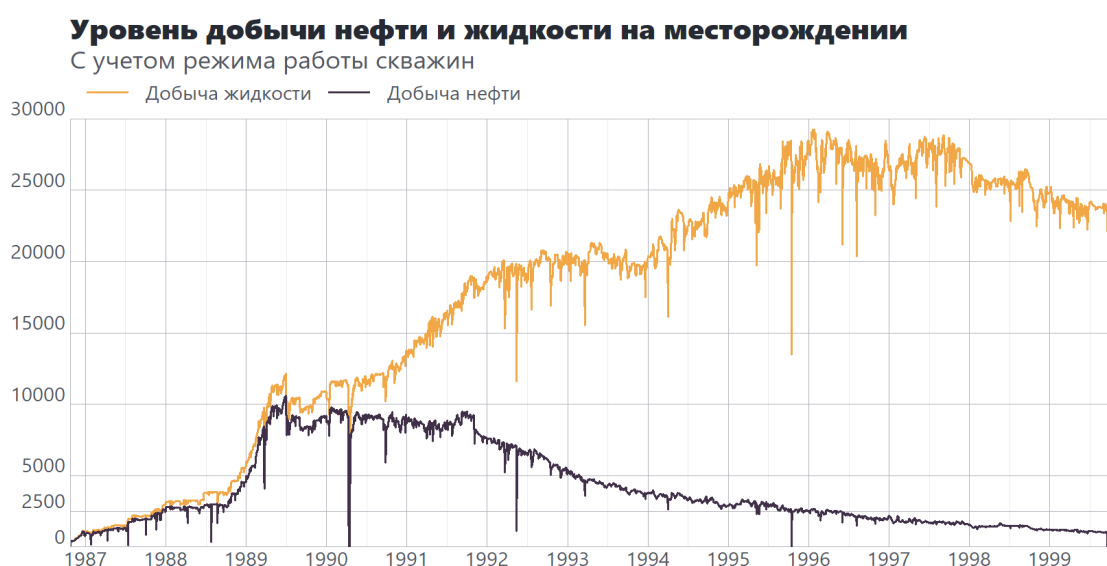


Рисунок 2.2 – Уровень объема добычи нефти и жидкости на месторождении №1 за весь период наблюдения

Исходные данные по месторождению №2 содержат информацию по 312 скважинам, из которых 263 (84%) – вертикальные, 12 (4%) – разведочные, 6 (2%) – боковой ствол (БС), 28 (9%) – горизонтальный ствол (ГС), 2 (0.6%) – ответвление (ПО), 1 (0.4%) – разведочный боковой ствол (РБС). Всего в данных 23371 наблюдение. Из них 21651 приходится на статус “Работает”. У 6 скважин за весь период нет состояния “Работает”. Пропусков в переменных практически нет. В исходных данных представлена разбивка месторождения на 16 кустов. Есть информация с уточнением проводимых мероприятий. На Рисунке 2.3 изображена динамика изменения объемов добычи нефти и закачки жидкости за весь период представленных наблюдений по месторождению №2.

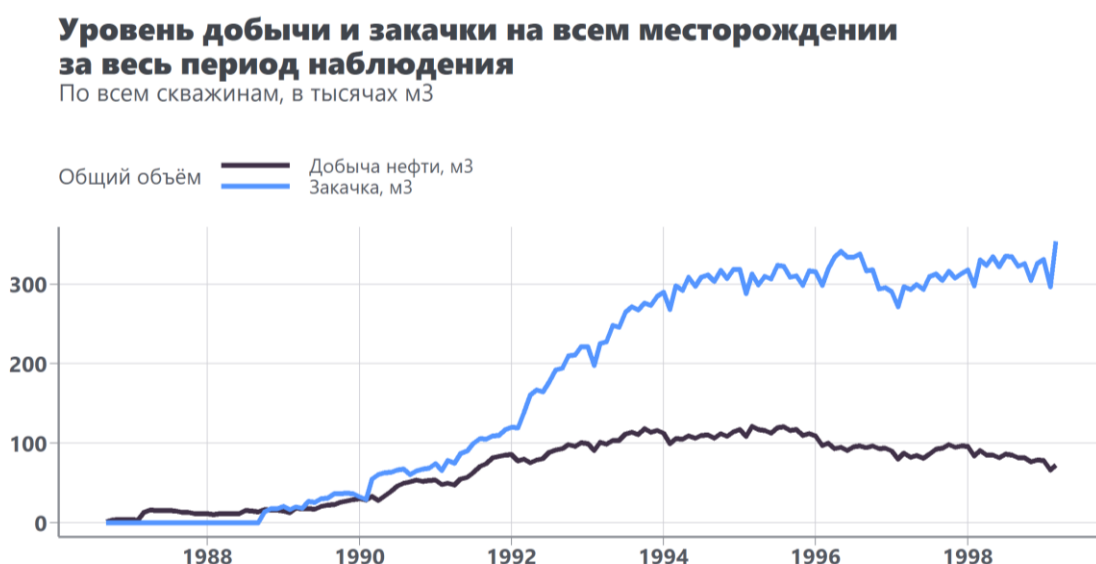


Рисунок 2.3 – Уровень объема добычи нефти и закачки на месторождении №2 за весь период наблюдения

На Рисунке 2.4 представлены распределение дат начала и окончания эксплуатации всех скважин месторождения №2. У большинства скважин последний наблюдаемый месяц – март 1999 года.

## Дата начала работы скважины

График плотности распределения

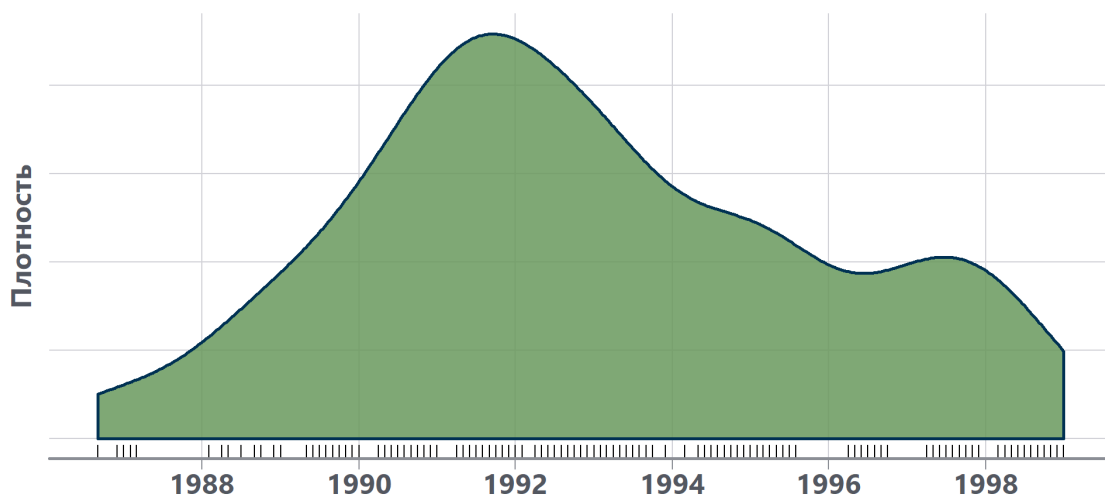


Рисунок 2.4 – Распределение дат начала работы скважин месторождения №2 по времени эксплуатации

В диссертации использовалась база из синтетических данных по месторождениям №1 и №2, которая включает: номер скважины, дата наблюдения, ежедневная добыча нефти (в  $\text{м}^3$ ), Formation Pressure (внутрипластовое давление, МПа), Bottom Hole Pressure (забойное давление, МПа). Информация по нагнетательным скважинам включает: номер скважины, дата наблюдения, объем закачиваемой ежедневно жидкости ( $\text{м}^3$ ). Для каждой скважины также имеется три координаты ее местоположения (абсцисса, ордината и аппликата) [7]. Входные параметры структурно с группировкой переменных по факторам представлены в Таблице 2.1.

На рисунках 1 и 2 Приложения А представлены графики объемов добычи нефти и жидкости, которые демонстрируют производительность скважин.

Таблица 2.1 – Структура входной информации месторождений №1 и №2

Факторы	Переменные
Разное	- Номер скважины - Дата наблюдения - Скин - фактор (показатель мероприятия)
Данные о добыче	- Нефть, тонн - Вода, тонн - Жидкость, тонн - Газ попутный, м <sup>3</sup> - Закачка, м <sup>3</sup> - Обводненность, % - Нефть, м <sup>3</sup> - Вода, м <sup>3</sup> - Газ растворенный, м <sup>3</sup>
Статические показатели	- Песчанность пласта - Пористость, % - Мощность коллектора по стволу, м - Эффективная мощность коллектора, м
Координаты	- Координата забоя X - Координата пластопересечения X - Координата пластопересечения Y - Кровля коллектора по траектории (Z), м
Статусы	- Характер работы - Состояние - Время работы, ч
Давления	- Пластовое давление, атм - Забойное давление, атм

Для анализа бесперебойности работы скважин для каждого типа скважин в зависимости от объема нефти были построены боксплоты, показывающие время эксплуатации в месяцах, позволяющие выдвинуть гипотезы о законе распределения признаков.

## Распределение времени эксплуатации скважин

С учетом характера работы скважин

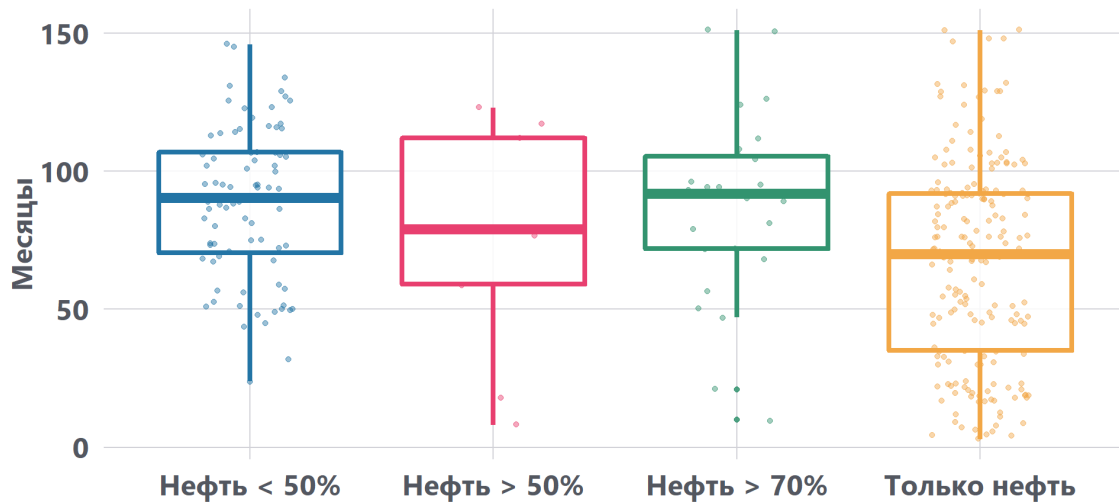


Рисунок 2.5 – Распределение времени эксплуатации скважины

На Рисунке 2.5 можно увидеть, что большинство скважин отработали практически весь период эксплуатации, в данных имеется незначительное смещение [8].

Несмотря на то, что второе месторождение содержало больше скважин, основой для обучения моделей оценки взаимовлияния скважин стало месторождение №1.

### 2.2 Предобработка входной информации на основе геофизических критериев

Одной из проблем при получении первичной информации с месторождения является ее возможная некорректность. Поэтому входные данные по всем месторождениям были проверены на корректность. Для этого использовался геофизический критерий на основании уравнения Дарси [21]:

$$q_{liq} = PI(P_{пл} - P_{заб}^{доб}), q_{inq} = PI(P_{заб}^{наг} - P_{пл}), \quad (2.1)$$

где  $q_{liq}$  – добыча жидкости в единицу времени;



$q_{inq}$  – закачка жидкости в единицу времени;

$P_{пл}$  – пластовое давление;

$P_{заб}^{доб}$  – забойное давление добывающей скважины,  $P_{заб}^{доб} \ll P_{пл}$ ;

$P_{заб}^{нагн}$  – забойное давление нагнетательной,  $P_{заб}^{наг} \gg P_{пл}$ ;

$PI$  – постоянная величина.

Объем жидкости в пласте:

$$Q_{жид} = Q_0 - Q_{изл} + Q_{зак}. \quad (2.2)$$

Для  $wc$  – обводненность:

$$wc = \frac{q_{вод}}{q_{liq}} \quad (2.3)$$

$$q_{liq} = q_{вод} + q_{oil}, \quad (2.4)$$

где  $q_{вод}$  – объем извлеченной воды;

$q_{oil}$  – объем извлеченной нефти;

$$q_{oil} = (1 - wc)q_{liq} = (1 - wc) * PI * (P_{пл} - P_{заб}^{доб}) = (1 - wc) * PI * \Delta P^{доб}.$$

$$PI = \frac{q_{liq}}{(P_{пл} - P_{заб}^{доб})} = \frac{q_{oil}}{(1 - wc) * (P_{пл} - P_{заб}^{доб})} \quad (2.5)$$

Допустимой считается отклонение параметра  $PI$  от среднего значения на 10-15%. Для очистки входных данных для добывающих скважин использовалась формула:

$$PI = \frac{q_{liq}}{(P_{пл} - P_{заб}^{доб})}. \quad (2.6)$$

Расчёт коэффициента  $PI$  на исходных данных показал, что значение критерия во временных рядах принимает отрицательные значения по двум причинам: отсутствовали данные по пластовому и/или забойному давлению; скачок забойного давления.

Помимо проверки коэффициента  $PI$  на неотрицательность, необходимо провести проверку на отсутствие необъяснимых скачков параметра, например, по причине остановок, изменения давления и иных причин. В случае если

происходят необъяснимые скачки, то корректировать информацию следует за счет заполнения значением предыдущего дня. Стоит отметить, что в спорных ситуациях, когда неясно, является ли данный скачок аномалией в данных или же результатом неких мероприятий, то целесообразно данные оставлять без изменений.

Корректировка исходной информации по забойному давлению проведена аналогично. График изменения коэффициента  $PI$  на примере скважины 10 месторождения №1 до и после процедуры очистки (Рисунки 2.6 и 2.7).



Рисунок 2.6 – Коэффициент  $PI$  для скважины до процедуры очистки

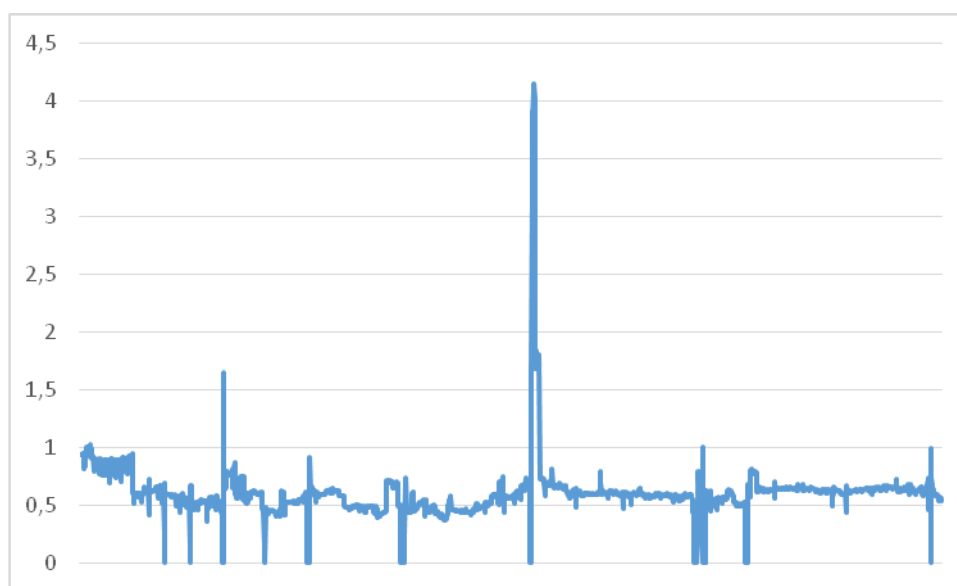


Рисунок 2.7 – коэффициент  $PI$  для скважины 10 после процедуры очистки

Ряд скважин продемонстрировали аномальное количество отрицательных значений коэффициента  $PI$  либо на некотором начальном отрезке, либо за все наблюдение, поэтому было принято решение отбросить такие ряды полностью.

В результате проделанной очистки данных и после удаления аномальных участков удалось сократить вариацию временных рядов коэффициента  $PI$  по скважинам (Таблица 2.2). Стоит отметить, что после процесса очистки осталось некоторое количество рядов со значительной вариацией коэффициента  $PI$ , что, возможно, связано с большим количеством мероприятий на скважинах и определенным количеством ошибок в исходных данных, которые остались после очистки.

Таблица 2.2 – распределение скважин по размеру вариации коэффициента  $PI$

Уровень вариации, %	До очистки данных	После очистки данных
0-20	2	4
20-50	21	46
50-100	35	34
100+	40	13

Сравнение Рисунков 2.8 и 2.9 показывает, что в результате корректировки данных качество входной информации удалось значительно улучшить. Итак, менее половины исходных данных показывает отклонение коэффициента  $PI$  на менее чем 50% (до очистки менее 25%), более 30% скважин имеют погрешность от 50 до 100% (против 35%), более 10% имеют отклонение более 100% (против 40% до корректировок).

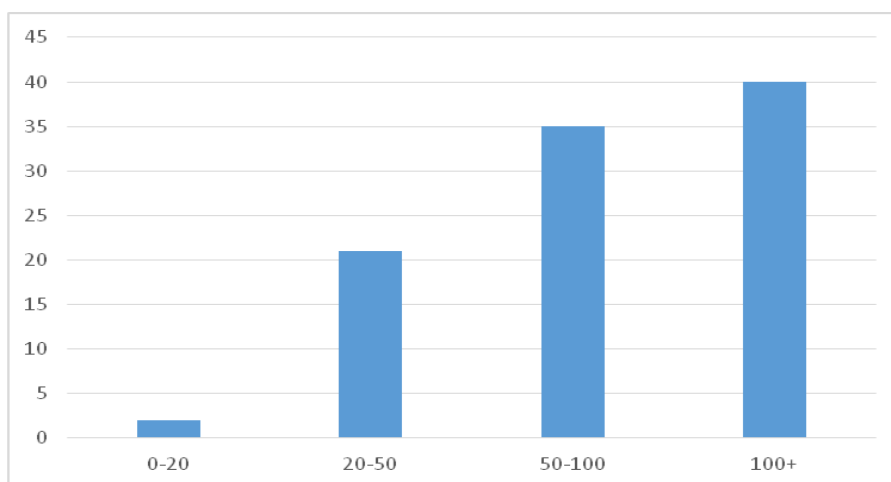


Рисунок 2.8 – Распределение скважин по размеру вариации коэффициента  $PI$  до очистки данных

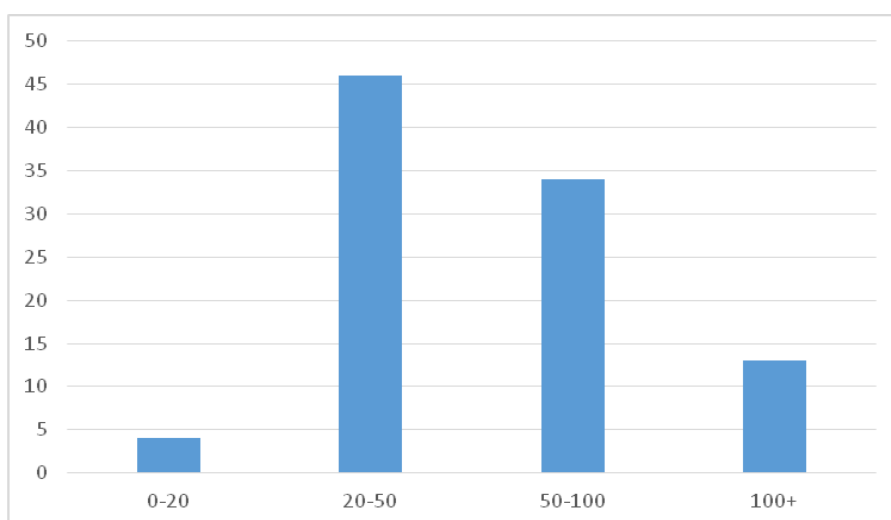


Рисунок 2.9 – Распределение скважин по размеру вариации коэффициента  $PI$  после очистки данных

В результате преобразования входной информации по месторождениям, представленной в виде ретроспективы данных, удалось получить относительно качественную информационную базу достаточного объема.

### 2.3 Методика синхронного анализа временных рядов

Для оценки взаимовлияния скважин необходимо проводить синхронный анализ временных рядов динамических характеристик работы всех скважин

месторождений. Для этого используется следующая последовательность методик анализа:

Первый этап – выявление типа процесса динамических рядов с использованием процедур статистического тестирования [25]. С помощью расширенного теста Дики-Фуллера тестируются следующие уравнения спецификации:

Процесс первого порядка интегрирования с детерминированным линейным трендом:

$$\Delta^2 z_t = \rho_0 + \rho_1 \Delta z_{t-1} + \sum_{i=2}^p p_i \Delta^2 z_{t-i+1} + \eta t + \varepsilon_t \quad (2.7)$$

Процесс нулевого порядка интегрирования с детерминированным линейным трендом:

$$\Delta z_t = \rho_0 + \rho_1 z_{t-1} + \sum_{i=2}^p p_i \Delta z_{t-i+1} + \eta t + \varepsilon_t \quad (2.8)$$

Процесс второго порядка интегрирования:

$$\Delta^3 z_t = \rho_0 + \rho_1 \Delta^2 z_{t-1} + \sum_{i=2}^p p_i \Delta^3 z_{t-i+1} + \varepsilon_t \quad (2.9)$$

Процесс первого порядка интегрирования:

$$\Delta^2 z_t = \rho_0 + \rho_1 \Delta z_{t-1} + \sum_{i=2}^p p_i \Delta^2 z_{t-i+1} + \varepsilon_t \quad (2.10)$$

Процесс нулевого порядка интегрирования:

$$\Delta z_t = \rho_0 + \rho_1 z_{t-1} + \sum_{i=2}^p p_i \Delta z_{t-i+1} + \varepsilon_t \quad (2.11)$$

Во всех уравнениях (2.7) – (2.11) тестируемой спецификации:  $z_t$  – тестируемый случайный дискретный процесс, являющийся временным рядом;  $t$  – временной период;  $i$  – лаг запаздывания;  $\rho_1$  – коэффициент авторегрессии первого порядка;  $p$  – максимальный лаг запаздывания в авторегрессии;  $\varepsilon_t$  – случайная составляющая, являющаяся белым шумом;  $\Delta$  – знак взятия разницы;  $\rho_0$  – константа. Во всех этих спецификациях (2.7) – (2.11) нулевой гипотезой является

предположение о том, что  $\rho_1=0$ , принятие решений приводится на основе проверки сложной гипотезы, то есть пересечения пяти соответствующих гипотез. Расчетное значение получаемых  $t$ -статистик сравнивается с критическим значением Мак-Кинона, что дает основу для итогового принятия решения о типе случайного процесса.

Второй этап – уточнение типа процесса, лежащего в основе временного ряда с учетом возможных структурных изменений ряда (остановки работы скважины, проведения мероприятий по интенсификации добычи нефти и др.) с использованием теста Перрона [25].

Тестировали процессы следующей спецификации согласно тесту (здесь  $H_0$  – нулевая гипотеза, а  $H_1$  – альтернативная):

Процесс первого порядка интегрирования:

$$H_0: X_t = \mu + X_{t-1} + \gamma \cdot D(t_B)_t + \varepsilon_t. \quad (2.12)$$

Против процесса с детерминированным трендом

$$H_1: X_t = a_2 + b \cdot t + (a_1 - a_2) \cdot D_t + \varepsilon_t. \quad (2.13)$$

Процесс первого порядка интегрирования

$$H_0: X_t = \mu_1 + X_{t-1} + (\mu_2 - \mu_1) \cdot D_t + \varepsilon_t. \quad (2.14)$$

Против процесса с детерминированным трендом

$$H_1: X_t = a_1 + b_1 \cdot t + (b_2 - b_1) \cdot D^*_t + \varepsilon_t. \quad (2.15)$$

Процесс первого порядка интегрирования

$$H_0: X_t = \mu_1 + X_{t-1} + \gamma \cdot D(t_B)_t + (\mu_2 - \mu_1) \cdot D_t + \varepsilon_t. \quad (2.16)$$

Против процесса с детерминированным трендом

$$H_1: X_t = a_1 + b_1 \cdot t + (b_2 - b_1) \cdot D^*_t + (a_2 - a_1) \cdot D_t + \varepsilon_t \quad (2.17)$$

Здесь:  $D^*_t = \begin{cases} t - t_B, & t > t_B, \\ 0, & t \leq t_B. \end{cases}$ ,  $D_t = \begin{cases} 1, & t > t_B, \\ 0, & t \leq t_B. \end{cases}$

$D(t_B)_t = \begin{cases} 1, & t_B + 1, \\ 0, & \text{в остальных случаях.} \end{cases}$

$t_B$  – точка излома и/или скачка тенденции.

Для определения  $t_B$  – точки излома и/или скачка тенденции используется тест Квандта-Эндрюса, в котором нулевой гипотезой делается предположение об

отсутствии точки излома в каждой точке временного ряда.

Все временные ряды ежедневной динамики добычи нефти, дебита жидкости, забойного давления, пластового давления и ряда разницы давлений для всех добывающих и нагнетательных скважин месторождений №1 и №2 протестированы с помощью теста Дики-Фуллера и уточнены с помощью тестов Квандта-Эндрюса и Перрона. Пример результатов тестирования временных рядов характеристик работы 9-ой скважины сведены в Таблице 2.3. Здесь также рассматривались и логарифмы дебита скважины.

Таблица 2.3 – Тесты на уточнение типов процессов характеристик добычи на скважине месторождения 1

Номер скважины = переменная	Расширенные тесты Дики-Фуллера					Уточнение типа процесса			Вывод о типе процесса
	Процесс первого порядка интегрирования с детерминированным линейным трендом	Процесс нулевого порядка интегрирования с детерминированным линейным трендом	Процесс второго порядка интегрирования	Процесс первого порядка интегрирования	Процесс нулевого порядка интегрирования	Тест Перрона Исходного ряда	Тест Перрона с трендом	Тест Перрона Первых разностей	
1	2	3	4	5	6	7	8	9	10
9 добыча	$t_{МК}=-25,458$ ; $p=0,0000$ тренд не значим	$t_{МК}=-8,198$ ; $p=0,0000$ тренд значим	$t_{МК}=-22,795$ ; $p=0,0000$	$t_{МК}=-25,461$ ; $p=0,0000$	$t_{МК}=-6,291$ ; $p=0,0000$	$t_{МК}=-9,448$ ; $p<0,01$ точка излома 12/02/1987 излом не значим, дамми-переменная не значима	$t_{МК}=-10,847$ ; $p<0,01$ точка излома 12/02/1987 тренд значим, излом не значим, дамми-переменная не значима	$t_{МК}=-76,257$ ; $p<0,01$ точка излома 10/31/1990 излом не значим, дамми-переменная значима	TS без структурных скачков
9 LN(добычи)	$t_{МК}=-34,896$ ; $p=0,0000$ тренд не значим	$t_{МК}=-5,006$ ; $p=0,0002$ тренд значим	$t_{МК}=-22,571$ ; $p=0,0000$	$t_{МК}=-34,911$ ; $p=0,0000$	$t_{МК}=-3,717$ ; $p=0,0039$	$t_{МК}=-25,663$ ; $p<0,01$ точка излома 7/17/1995 излом значим, дамми-переменная значима	$t_{МК}=-29,567$ ; $p<0,01$ точка излома 12/02/1987 тренд значим, излом значим, дамми-переменная не значима	$t_{МК}=-93,578$ ; $p<0,01$ точка излома 3/09/1993 излом не значим, дамми-переменная значима	TS со структурными скачками



продолжение Таблицы 2.3

1	2	3	4	5	6	7	8	9	10
9 жидкость	$t_{MK}=-16,774$ ; $p=0,0000$ тренд не значим	$t_{MK}=-2,651$ ; $p=0,2576$	$t_{MK}=-23,539$ ; $p=0,0000$	$t_{MK}=-16,775$ ; $p=0,0000$	$t_{MK}=-1,980$ ; $p=0,2960$	$t_{MK}=-6,147$ ; $p<0,01$ точка излома 11/19/1987 излом не значим, дамми-переменная не значима	$t_{MK}=-8,537$ ; $p<0,01$ точка излома 12/02/1987 тренд значим, излом не значим, дамми- переменная не значима	$t_{MK}=-68,129$ ; $p<0,01$ точка излома 11/19/1987 излом не значим, дамми- переменная не значима	DS I(1) без структурн ых скачков
9 Заб. давление	$t_{MK}=-31,951$ ; $p=0,0000$ тренд не значим	$t_{MK}=-5,879$ ; $p=0,0000$ тренд не значим	$t_{MK}=-21,734$ ; $p=0,0000$	$t_{MK}=-31,954$ ; $p=0,0000$	$t_{MK}=-5,880$ ; $p=0,0000$	$t_{MK}=-6,955$ ; $p<0,01$ точка излома 12/01/1987 излом не значим, дамми-переменная не значима	$t_{MK}=-8,306$ ; $p<0,01$ точка излома 9/11/1995 тренд значим, излом значим, дамми- переменная значима	$t_{MK}=-69,308$ ; $p<0,01$ точка излома 11/16/1987 излом незначим, дамми- переменная не значима	DS I(0) со структурн ыми скачками
9 Пласт. Давл	$t_{MK}=-66,254$ ; $p=0,0000$ тренд не значим	$t_{MK}=-2,399$ ; $p=0,3797$	$t_{MK}=-24,922$ ; $p=0,0000$	$t_{MK}=-66,251$ ; $p=0,0001$	$t_{MK}=-2,472$ ; $p=0,1225$	$t_{MK}=-6,955$ ; $p=0,4069$ точка излома 4/30/1999	$t_{MK}=-4,093$ ; $p=0,3164$ точка излома 12/31/1992	$t_{MK}=-75,558$ ; $p<0,01$ точка излома 3/01/1994 излом не значим, дамми- переменная значима	DS I(1) со структурн ыми скачками
9 Разница давлений	$t_{MK}=-43,097$ ; $p=0,0000$ тренд не значим	$t_{MK}=-5,909$ ; $p=0,0000$ тренд не значим	$t_{MK}=-26,443$ ; $p=0,0000$	$t_{MK}=-43,101$ ; $p=0,0000$	$t_{MK}=-5,914$ ; $p=0,0000$	$t_{MK}=-7,267$ ; $p<0,01$ точка излома 9/11/1995 излом значим, дамми-переменная не значима	$t_{MK}=-10,423$ ; $p<0,01$ точка излома 8/12/1995 тренд значим, излом значим, дамми- переменная незначима	$t_{MK}=-69,068$ ; $p<0,01$ точка излома 11/16/1987 излом незначим, дамми- переменная не значима	DS I(0) со структурн ыми скачками

Распределение по типам процессов по всем исследуемым временным рядам приведено в Таблице 2.4.

Таблица 2.4 – Количественное распределение динамических признаков работы скважин месторождения №1 по типам процессов

	TS + DS(I(1))	TS+DS(I(0))	DS(I(2))	DS(I(1))	DS(I(0))
Добыча нефти	1	8	3	16	56
Дебит жидкости	0	1	8	9	66
Забойное давление	0	1	12	5	66
Пластовое давление	0		23	51	10
Разница давлений	0	1	15	9	59

Следует отметить, что подавляющее большинство динамических показателей, относящихся к эксплуатации скважин, относятся к процессам DS(I(0)), то есть по сути они являются стационарными случайными процессами в широком смысле даже при наличии структурных скачков. Выявлен незначительный ряд скважин с типом процесса TS (наличие тренда). Полученный результат говорит о том, что для рядов DS(I(0)) наилучшим прогнозом будет прогноз, основанный на продлении динамики обычной константой, то есть метод «протяжки последнего значения ряда динамики». Дальнейшие численные эксперименты это подтвердили.

Третий этап – проверка коинтеграции. Проверка условия коинтеграции необходима для построения динамических моделей взаимовлияния скважин, а также возможности выявления корректной спецификации моделей [32]. Для проверки нулевой гипотезы об отсутствии коинтеграции (тест Ингла-Гренджера) оценивается коинтеграционное соотношение, как линейная регрессия  $y_t = \hat{\alpha} + \hat{\beta} \cdot x_t + \varepsilon_t$ , где  $x_t$  и  $y_t$  временные ряды. Затем остатки  $\varepsilon_t$  тестируются на стационарность с помощью теста Дики-Фуллера. Необходимым условием для использования теста на коинтеграцию является то, что ряды должны относиться к

одному и тому же типу случайных дискретных процессов. При выполнении обоих условий временные ряды  $x_t$  и  $y_t$  являются коинтегрированными. В тесте для переменных вероятность ошибки отклонения нулевой гипотезы ( $p$ ) сразу рассчитывается на основе сравнения расчетных значений  $\tau$ -статистик и  $z$ -статистик со значения уточненными Мак-Киноном [29].

На четвертом этапе – построение динамических моделей. Здесь осуществляется проверка теста на причинность, то есть определения, какая из переменных является эндогенной и экзогенной. Если  $X_t$  - причина по Гренджеру для  $Z_t$ , то это означает, что между этими процессами есть причинно-следственная связь. Для этого используют тест на причинность Гренджера с двумя нулевыми гипотезами:  $X_t$  – не является причиной по Гренджеру для  $Z_t$ , и  $Z_t$ , - не является причиной по Гренджеру для  $X_t$ .

Для тестирования причинности по Гренджеру строят регрессию  $Z_t$  на его собственные предыдущие значения и на предыдущие значения процесса  $X_t$ :

$$Z_t = \alpha_0 + \sum_{i=1}^k \alpha_i Z_{t-i} + \sum_{i=1}^k \beta_i X_{t-i} + \varepsilon_t. \quad (2.18)$$

Затем проверяем гипотезы относительно коэффициентов уравнения (2.18):

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0;$$

$$H_1: \beta_1^2 + \beta_2^2 + \dots + \beta_k^2 > 0.$$

Поверяют гипотезу на основе расчёта  $F$ -статистик, которую сравнивают с критическими значениями Фишера. Если нулевая гипотеза отвергается, то  $X_t$  является причиной для  $Z_t$ .

В качестве целевых переменных в тесте используются динамические переменные добычи нефти и дебита жидкости, в качестве предикторов (факторов или независимых переменных) забойное давление, пластовое давление и разница давлений. Особенностью теста является то, что можно проверить тест с уравнением, в которое добавлена авторегрессия до задаваемого лага запаздывания включительно. То есть тест на коинтеграцию может стать основой для выбора оптимальной спецификации модели описания взаимного влияния скважин так, чтобы учесть отсроченное влияние таких характеристик как добыча на соседних

скважинах и изменение давлений на дебит жидкости (или нефти) добывающей скважины. Результаты тестирования на коинтеграцию согласно критерию Ингла-Гренджера по скважинам 9, 10, 11 представлены в Таблице 2.5. В таблицу сведены значения  $F$ -статистик проверки нулевых гипотез о влиянии первой на вторую переменную ( $F_1$ ) и наоборот ( $F_2$ ) и соответствующие им  $p$ -уровни. Также сделаны выводы о проверке соответствующих гипотез, указан лаг запаздывания.

Результаты проведения соответствующих тестов для целевых переменных ежедневной добычи нефти и ежедневного дебита жидкости и предикторов забойного давления, пластового давления и разницы давлений также приведены в Таблице 1, приложения 1. Здесь в тесте  $F_1$  – статистика теста, где  $H_1$  объясняющая переменная (фактор, предиктор) является причиной для целевой,  $F_2$  – статистика теста, где  $H_1$  целевая переменная является причиной для объясняющей (фактора). Также проведены тесты на причинность между дебитом нефти текущей и дебитом нефти соседних по кусту скважин с учётом лага запаздывания. Рассматривались тесты на причинность между дебитом жидкости текущей и дебитом жидкости соседних по кусту скважин с учётом лага запаздывания.

Таблица 2.5 – Тесты на причинность по Гренджеру между показателями работы скважин соседних скважин

Номер скважины	Целевая переменная дебит нефти	Целевая переменная дебит жидкости
9	$F_1=6.303; p=0.0003,$ $F_2=0.928; p=0.4263$ На добычу нефти скважины 9 влияет добыча скважины 69 только с лагом =3	$F_1=41,386; p<0,0001,$ $F_2=9,745; p<0,0001$ На добычу нефти скважины 9 влияет добыча скважины 69 только с лагом =3
10	$F_1=2,34; p=0,0397,$ $F_2=3,03; p=0,0099$ Есть зависимость дебитов между 70 и 11 скважинами, Лаг = 5. $F_1 =3,29; p=0,0199,$ $F_2 =1,00; p=0,3917$ Есть зависимость дебита скважины 11 от дебита скважины 71, Лаг = 3. $F_1 =1,92; p=0,0386,$ $F_2 =5,11; p=0,0000$ Есть зависимость дебитов между 70 и 71 скважинами, Лаг = 10.	$F_1 =1,43; p=0,2333,$ $F_2 =3,31; p=0,0194$ Есть зависимость дебита скважины 70 от дебита скважины 11, Лаг = 3. $F_1 =1,52; p=0,2197,$ $F_2 =3,76; p=0,0234$ Есть зависимость дебита скважины 71 от дебита скважины 11, Лаг = 2. $F_1 =9,07; p=0,0000,$ $F_2 =2,99; p=0,0299$ Есть зависимость дебитов между 70 и 71 скважинами, Лаг = 3.
11	$F_1=9,81; p=0,0000, F_2 =39,6; p=0,0000$ Есть зависимость дебитов между 78 и 74 скважинами, Лаг = 2. $F_1=27,09; p<0,0001, F_2=27,05; p<0,0001$ Есть зависимость дебитов между 78 и 21 скважинами, Лаг = 2. $F_1=29,21; p<0,0001, F_2=18,09; p<0,0001$ Есть зависимость дебитов между 78 и 79 скважинами, Лаг = 2. $F_1=42,11; p<0,0001, F_2=25,37; p<0,0001$ Есть зависимость дебитов между 78 и 84 скважинами, Лаг = 2. $F_1=7,81; p=0,0004, F_2=8,87; p=0,0001$ Есть зависимость дебитов между 74 и 21 скважинами, Лаг = 2. $F_1 =5,68; p=0,0034, F_2 =11,38; p<0,0001$ Есть зависимость дебитов между 74 и 79 скважинами, Лаг = 2. $F_1 =12,06; p<0,0001, F_2 =31,61; p=0,0000$ Есть зависимость дебитов между 74 и 84 скважинами, Лаг = 2. $F_1 =15,43; p<0,0001, F_2 =18,08; p<0,0001$ Есть зависимость дебитов между 21 и 79 скважинами, Лаг = 2. $F_1 =17,2; p<0,0001, F_2 =42,1; p<0,0001$ Есть зависимость дебитов между 21 и 84 скважинами, Лаг = 2. $F_1 =26,83; p<0,0001, F_2 =15,26; p<0,0001$ Есть зависимость дебитов между 79 и 84 скважинами, Лаг = 2.	$F_1 =3,74; p=0,0239, F_2 =1,66; p=0,1901$ Есть зависимость дебита скважины 78 от дебита скважины 74, Лаг = 2. $F_1 =6,83; p=0,0011, F_2 =9,02; p=0,0001$ Есть зависимость дебитов между 78 и 21 скважинами, Лаг = 2. $F_1 =4,32; p=0,0134, F_2 =6,62; p=0,0014$ Есть зависимость дебитов между 78 и 79 скважинами, Лаг = 2. $F_1=12,53; p<0,0001, F_2 =11,44; p<0,0001$ Есть зависимость дебитов между 78 и 84 скважинами, Лаг = 2. Между дебитами скважин 74 и 21 нет значимой зависимости при лагах 2-10 $F_1=11,23; p<0,0001, F_2 =10,77; p<0,0001$ Есть зависимость дебитов между 74 и 79 скважинами, Лаг = 2. $F_1 =0,66; p=0,5190, F_2 =12,23; p<0,0001$ Есть зависимость дебита скважины 84 от дебита скважины 74, Лаг = 2. $F_1 =9,27; p=0,0001, F_2 =5,18; p=0,0057$ Есть зависимость дебитов между 21 и 79 скважинами, Лаг = 2. $F_1 =7,31; p=0,0007, F_2 =19,43; p<0,0001$ Есть зависимость дебитов между 21 и 84 скважинами, Лаг = 2. $F_1 =5,72; p=0,0033, F_2 =8,15; p=0,0003$ Есть зависимость дебитов между 79 и 84 скважинами, Лаг = 2.

Следует добавить, что проведённый анализ позволяет определить спецификации моделей, то есть при построении моделей проводится

преобразование временных рядов в соответствии с результатами коинтеграционного анализа. В случае если ряды относились к одному типу процесса и коинтегрированы, то модель строилась на исходных рядах. В случае если процессы в модели (зависимая и независимая переменные) относятся к разным типам процесса, то их необходимо преобразовать до одного типа процесса. В случае если процессы отличаются порядком интегрирования, то один из типов процессов необходимо дифференцировать (взять разность). В случае наличия детерминированного тренда он непосредственно учитывался в моделях в виде добавочного линейного члена по времени ( $\beta^*t$ ).

#### 2.4 Разбиение месторождений на кусты

Характер рассматриваемого месторождения №2 с ежедневными данными таков, что сетка расположения скважин выглядит как равномерно-переменная. Выдвигались различные варианты формирования скважин из взаимовлияющих скважин друг на друга. На Рисунке 2.10 представлен пример кустов, сформированных из добывающих скважин. Стандартно кусты формируются не по учёту взаимовлияния скважин, а по принципу проведения буровых работ [84]. В каждый куст вошли совокупность скважин, устья которых группируются на близком расстоянии друг от друга на общей ограниченной площадке, с учётом результатов синхронного анализа временных рядов. В итоге было сформировано 24 куста.



Рисунок 2.10 – Разделение на кусты месторождения с ежедневными данными

В ходе численных экспериментов выдвигался ряд гипотез по разделению на кусты. Рассматривалась также гипотеза о целесообразности построения кустов для каждой скважины с учётом влияющих скважин. На Рисунке 2.11 представлен такой пример, кусты для 4 добывающих скважин: 88, 11, 58, 110 (выделены синим цветом). Новые кусты (выделенные красным цветом) для скважин (выделенных синим цветом) вместо первоначальных кустов (выделенных черным цветом). Дальнейшие расчёты на основе предложенных моделей и их вариаций показали, что первоначальное разделение является наиболее эффективным для задач исследования, поскольку даёт более точные прогнозы. Добавление следующего круга не улучшает прогнозные качества модели.

## Карта месторождения

С учетом режима работы скважин

● Оба режима    ● Только добыча    ● Только закачка

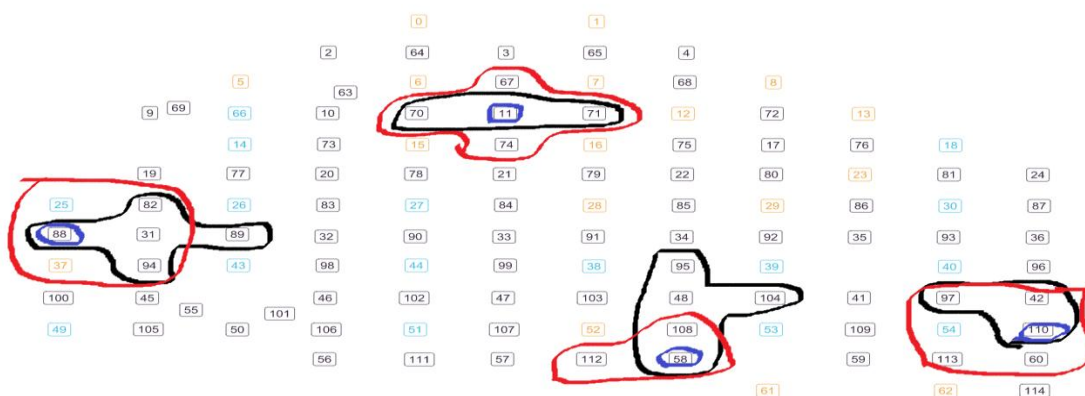


Рисунок 2.11 – Разделение месторождения по принципу образования индивидуального куста

Для формирования принципов вхождения скважин в кусты определялся радиус от целевой скважины: 500 м, 750 м и 1000 м.

## 2.5 Инструмент дискретного скрытого марковского моделирования (ДСММ) для оценки точек переключения типов процессов

Одной из гипотез для построения модели, выдвинутых экспертами была разбивка временных рядов объёмов выкачки нефти на отрезки со своими закономерностями. У временных рядов происходит переход между режимами выхода из структурного скачка (экспоненциальное падение,  $Exp$ ) и случайного блуждания,  $Rnd$  ( $X_t = X_{t-1} + \varepsilon_t$ , где  $\varepsilon_t$  – белый гауссов шум). Для выявления переходов использован механизм дискретных скрытых марковских моделей (ДСММ), относящийся к классу моделей машинного обучения без учителя [28].

Для двух режимов граф переходов между состояниями представлен на Рисунке 2.12.



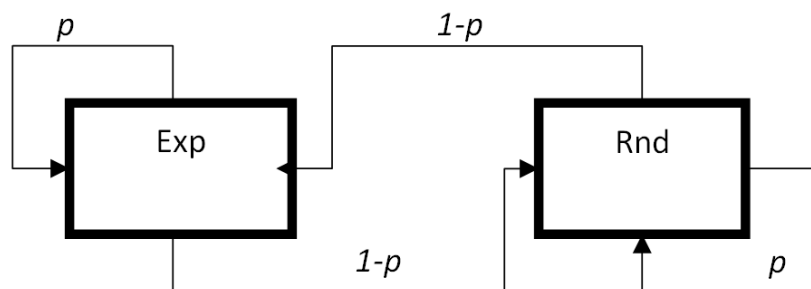


Рисунок 2.12 – Граф переходов между двумя состояниями

Соответственно, матрица переходов (трансмиссий) выглядит как:

	Exp	Rnd
Exp	$P$	$1-p$
Rnd	$1-p$	$P$

На каждый временной период считалась разность между скользящим средним в данном и предшествующим периодами и рассматривался знак данной разности (т.е. варианты  $>0$ ,  $<0$  и  $=0$ ). За исход брали скользящие средние длины  $k$ . Предварительные оценки дали следующую матрицу:

	$<0$	$=0$	$>0$
Exp	0,8	0,01	0,19
Rnd	0,499	0,002	0,499

Для определения последовательности режимов при априорно заданных вероятностях использовались алгоритм Витебри и soft decoding (библиотека `dermix S4` в R). Также были рассмотрены варианты с обучением, когда эти вероятности неизвестны. В результате численных экспериментов был использован исходный алгоритм Витебри. Для месторождения с ежемесячными данными, подобраны  $k = 5$  и  $p = 0,9999$ , месторождения с данными по дням  $k=101$  и  $p = 0,99$ . Результат серий экспериментов показал, что обучающие алгоритмы сильно «режут на кусочки» временной ряд, причем обучение Баума-Уэлча немного лучше обучения Витебри. Это продемонстрировано на Рисунках 2.13 – 2.15 для добычи на скважине 2.

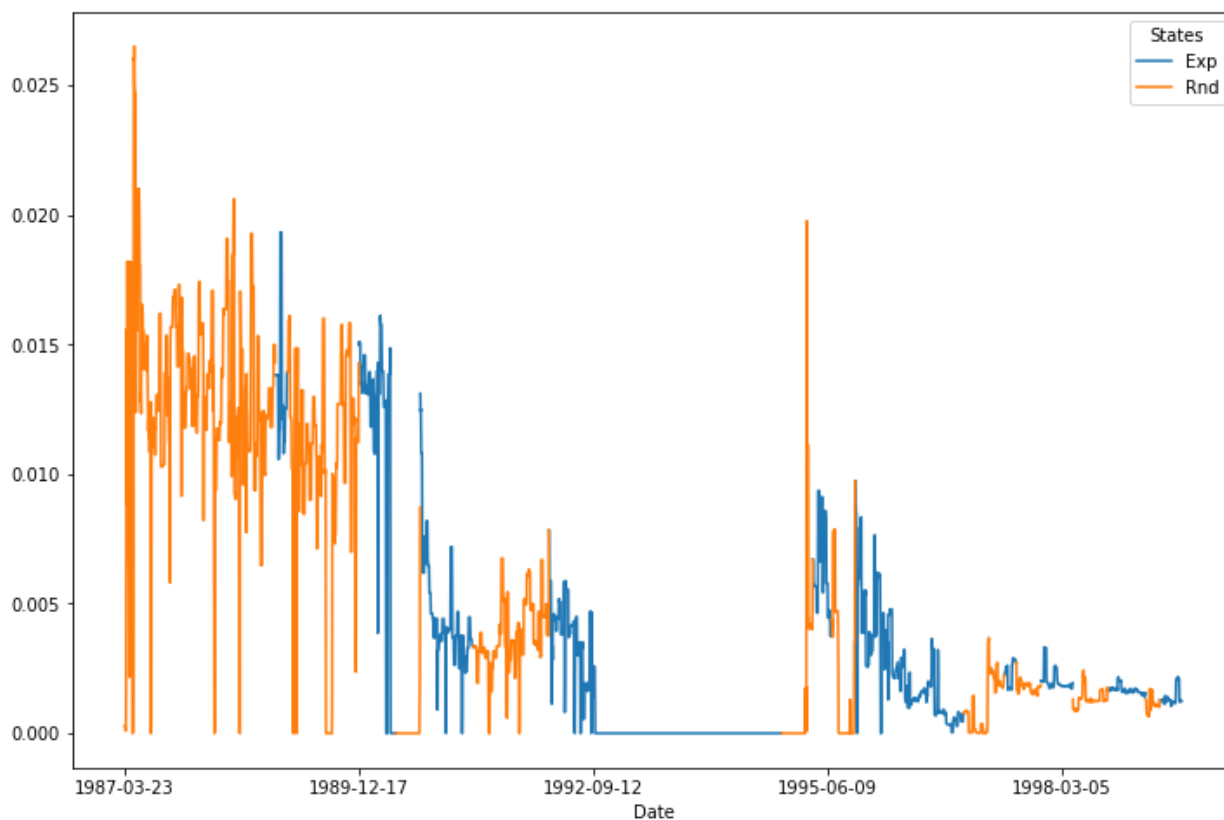


Рисунок 2.13 – Пример применения ДСММ с алгоритмом Soft Decoding (с порогом отсечения – 0.5) для ряда добычи нефти на скважине №2

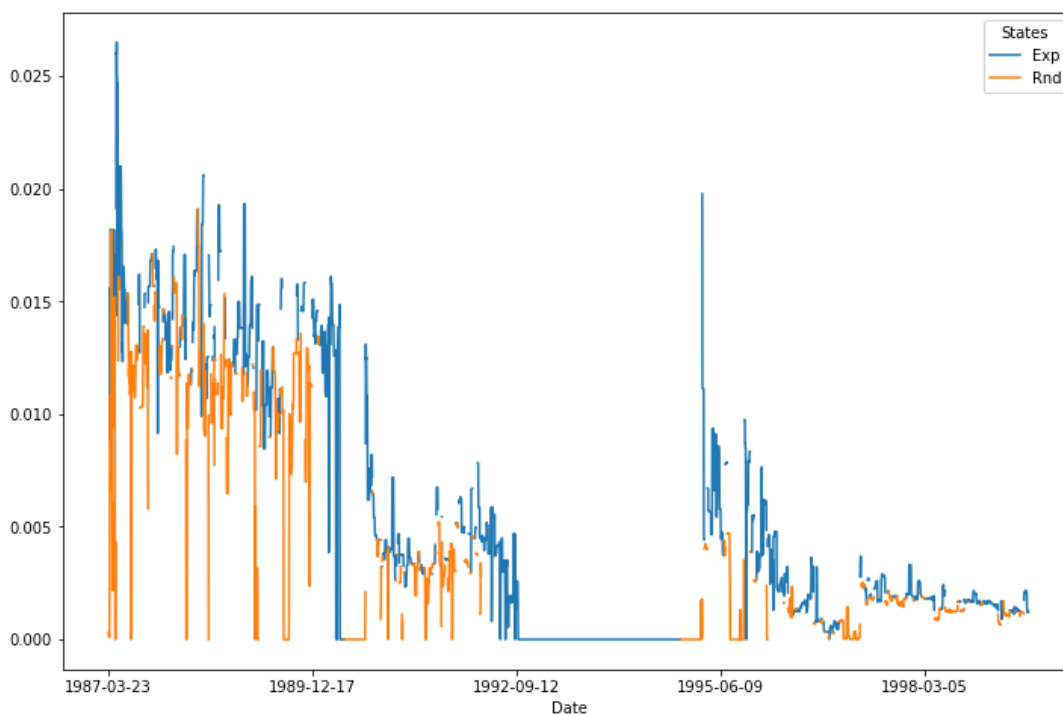


Рисунок 2.14 – Пример применения ДСММ с алгоритмом Витебри (количество итераций – 10) для ряда добычи нефти на скважине №2

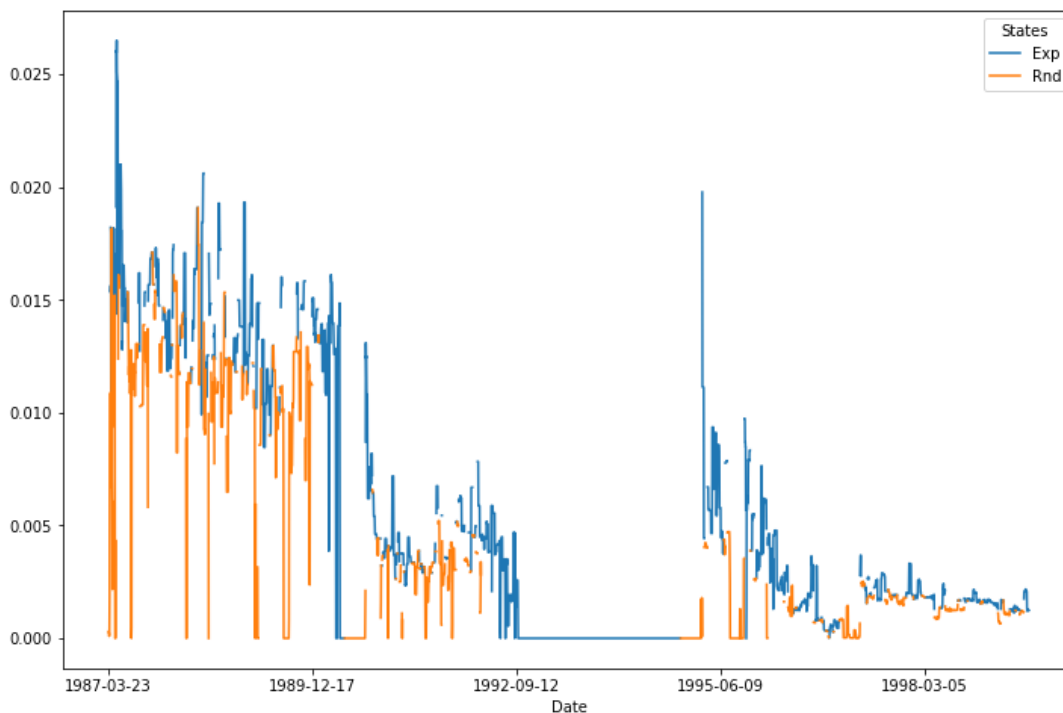


Рисунок 2.15 – Пример применения ДСММ с алгоритмом Баума-Уэлча (количество итераций – 10, с порогом отсечения – 0.5) для ряда добычи нефти на скважине №2

## 2.6 Анализ вариограмм и диаграмм рассеяния пространственных индексов

Для каждой из скважин месторождения (без учета нагнетательных скважин) были также на предварительном этапе проверены гипотезы о пространственных связях на основе анализа вариограмм.

Данный анализ выполнялся для возможности прогнозирования интересующих показателей в тех точках, по которым данных нет, в том числе для восполнения пропусков [5].

Для проверки справедливости гипотез о мультинормальности данных о песчаности и стационарности в пространстве процесса, из которого они получены, используются  $h$ -графики (для которых значение переменной в

пространственной точке  $x$  изображается напротив значения переменной в точке  $x+h$ ) [35].

На Рисунках В.1-В.4 Приложения Б представлены  $h$ -графики для направлений 0, 45, 90 и 135 градусов и длиной векторов  $h$  в интервале от 0 до 5000. Их анализ показал, что графики имеют форму, достаточно близкую к эллипсовидной, что является признаком того, что «исходные данные не противоречат гипотезе о мультинормальности, поскольку для мультинормальной случайной функции  $h$ -графики должны выглядеть как облака эллиптической формы вокруг диагональной прямой» [86]. В целом имеющихся данных достаточно для обобщения и построения по данным эмпирической вариограммы, поверхность которой представлена на Рисунке 2.16.

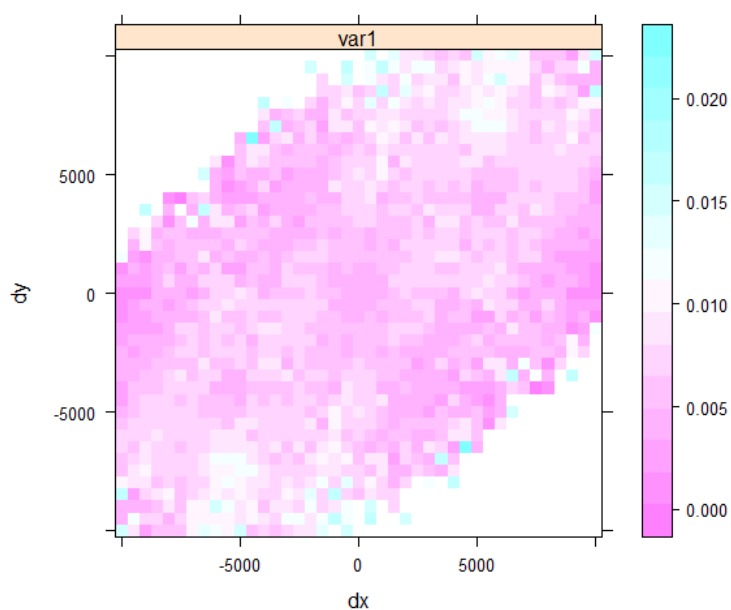


Рисунок 2.16 – Поверхность вариограммы для песчанности

Для анализа пространственных корреляций построены вариограммы для направлений от 0 до 360 градусов с шагом 30 градусов (Рисунки 2.17-2.20).

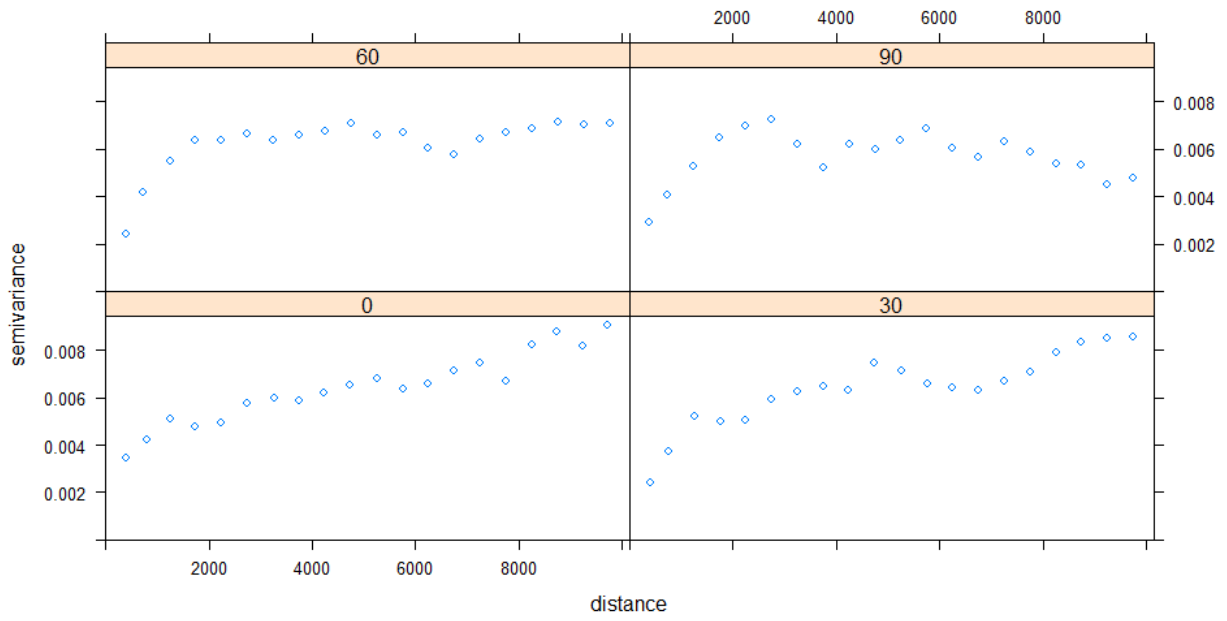


Рисунок 2.17 – от 0 до 90 градусов

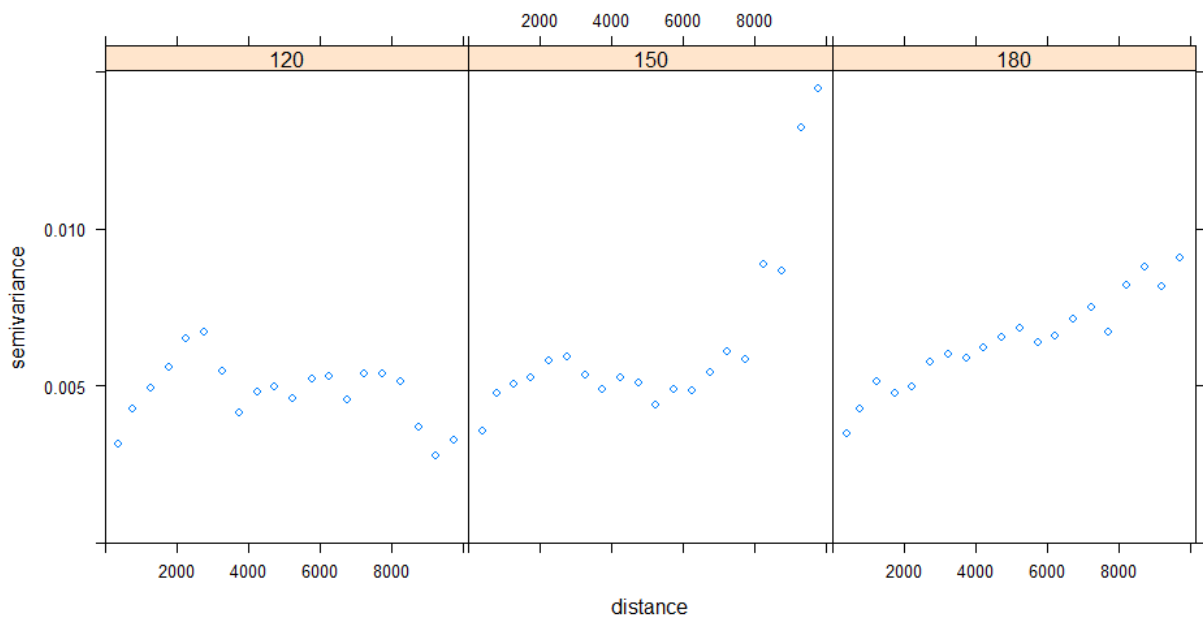


Рисунок 2.18 – от 90 до 180 градусов

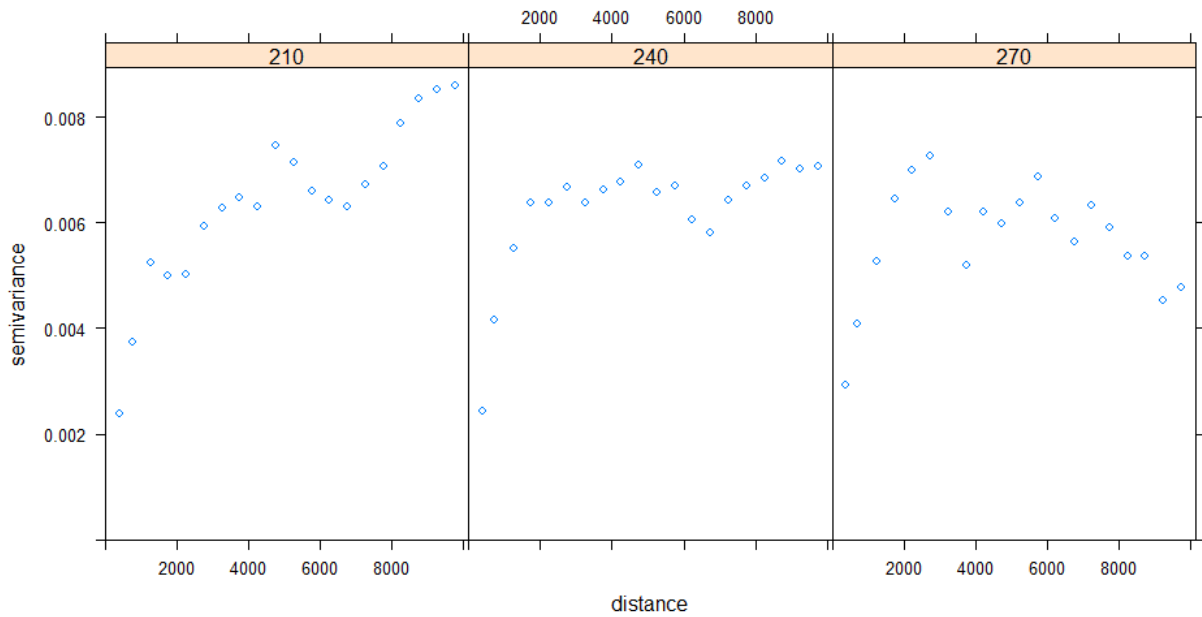


Рисунок 2.19 – от 210 до 270 градусов

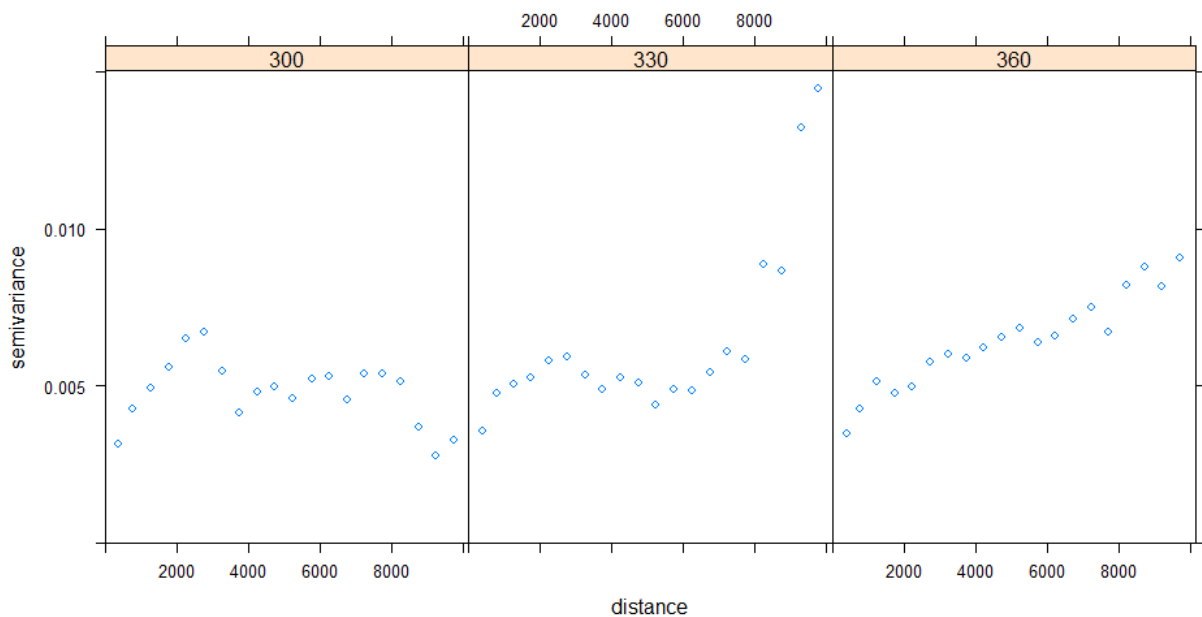


Рисунок 2.20 – от 300 до 360 (0) градусов

Анализ построенных графиков показал, что возможно присутствие «эффекта самородков» на вариограммах по всем направлениям [37] в размере около 0,002. Вариограмма в направлениях 0 и 180 градусов, вероятно, может моделироваться линейной функцией, 30, 60, 210 и 240 – сферической или гауссовой моделью. В целом относительная непрерывность графика при малых значениях расстояния свидетельствует об относительной непрерывности

песчанности. По направлениям 120, 270, 300 и 330 градусов поведение вариограммы описывается белым шумом (отсутствие корреляции между двумя точками независимо от расстояния между ними).

Средствами R (пакет gstat) была проведена оценка эмпирической вариограммы для данных по 4 направлениям [22]: по направлению максимальной анизотропии 30 градусов, перпендикулярному направлению 120 градусов и промежуточным 75 градусов и 165 градусов, с допуском 30 градусов, шагом лагов, равным 500, до максимального расстояния 10000.

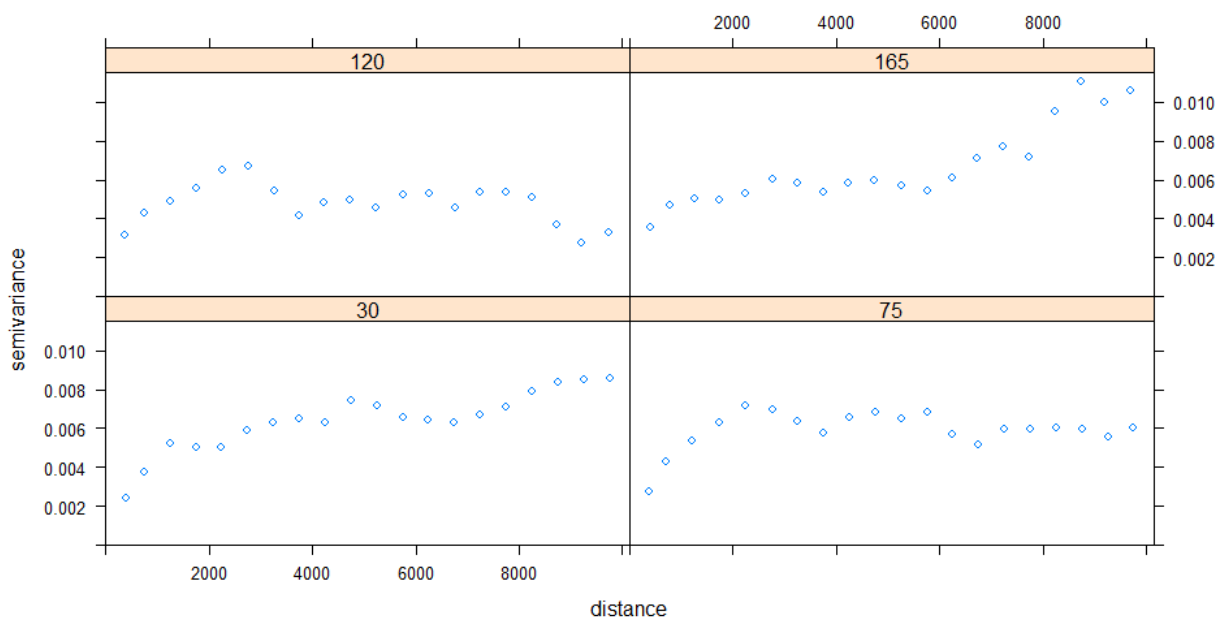


Рисунок 2.21 – Вариограммы по направлениям 30, 75, 120 и 165 градусов

Из Рисунке 2.21 видно, что поведение вариограммы действительно различно при различных направлениях. Оценённая дисперсия песчанности по выборке составляет 0,0625, и во всех случаях вариограммы не сильно выходят за это значение на малых значениях расстояний между точками. Возможно, в вариограммах помимо эффекта самородков (0,002) присутствует вложенная структура с рангом около 3000 и порогом около 0,004.

Для моделирования вариограммы используем модель с вложенной структурой [23]. Первая структура – модель сферическая, порог 0,004, ранг 3000, эффект самородков 0,002.

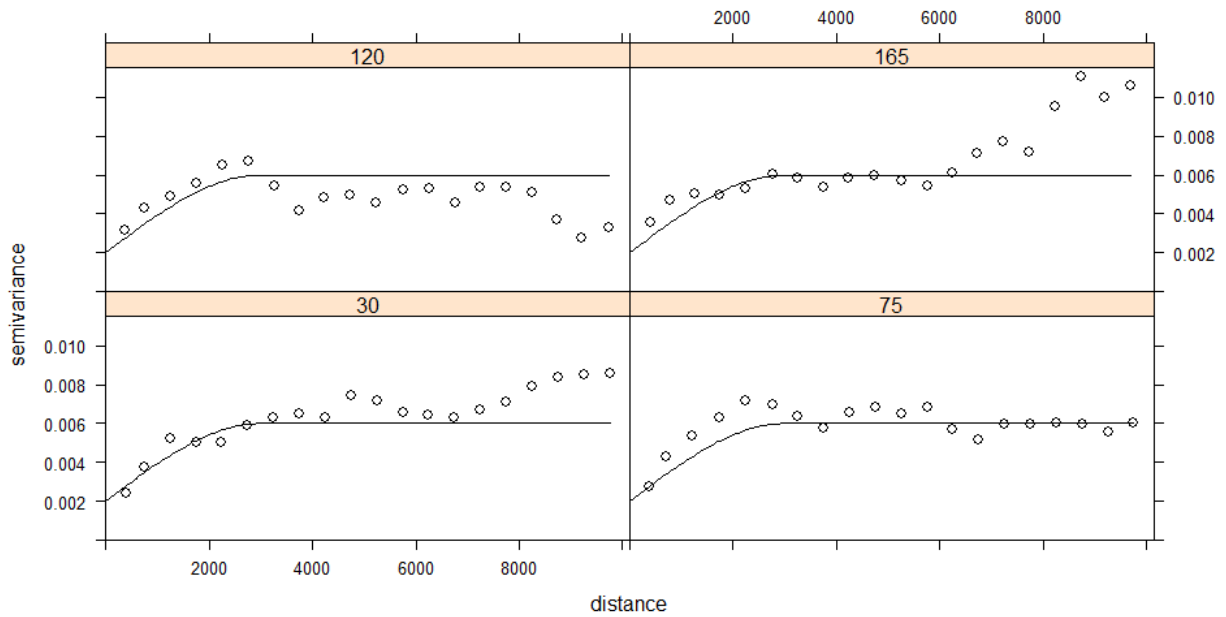


Рисунок 2.22 – Эмпирическая и модельная (сферическая) вариограммы по направлениям 30, 75, 120 и 165 градусов

Поскольку на эмпирической вариограмме возможно наличие иерархической структуры, добавим в построенную модель вторую структуру, для которой базисная модель – линейная модель с параметрами: порог 0,002, ранг 9000 (Рисунок 2.23).

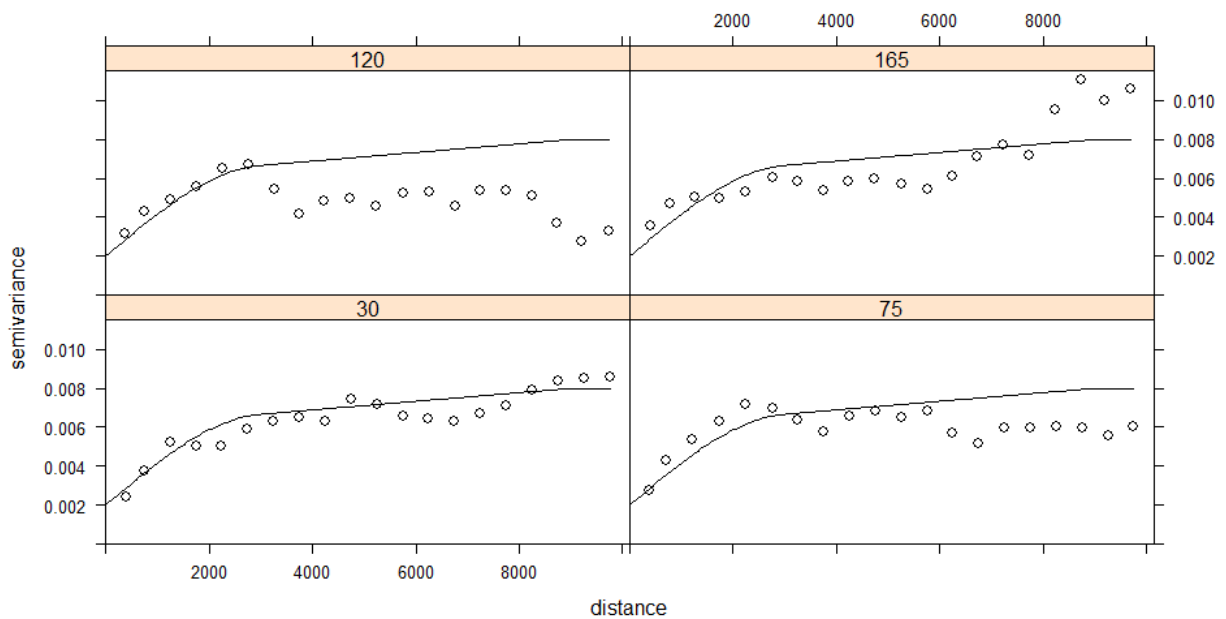


Рисунок 2.23 – Модель для вариограммы (со вложенной структурой)



Данная иерархическая модель (сферическая + линейная) была выбрана в качестве основы. Для уточнения ее параметров была проведена процедура подгонки с помощью функции `fit.variogram` пакета `gstat`. В результате подгонки параметров изменился эффект самородков (с 0,002 до 0,0019225), порог сферической модели уменьшился с 0,004 до 0,003502665, а ранг с 3000 до 1717.254; порог линейной модели вырос с 0,002 до 0,003225594, а ранг с 9000 до 21173,099.

Для перестроенной модели рисунки наложенных на эмпирические вариограммы аппроксимаций представлены на Рисунке 2.24.

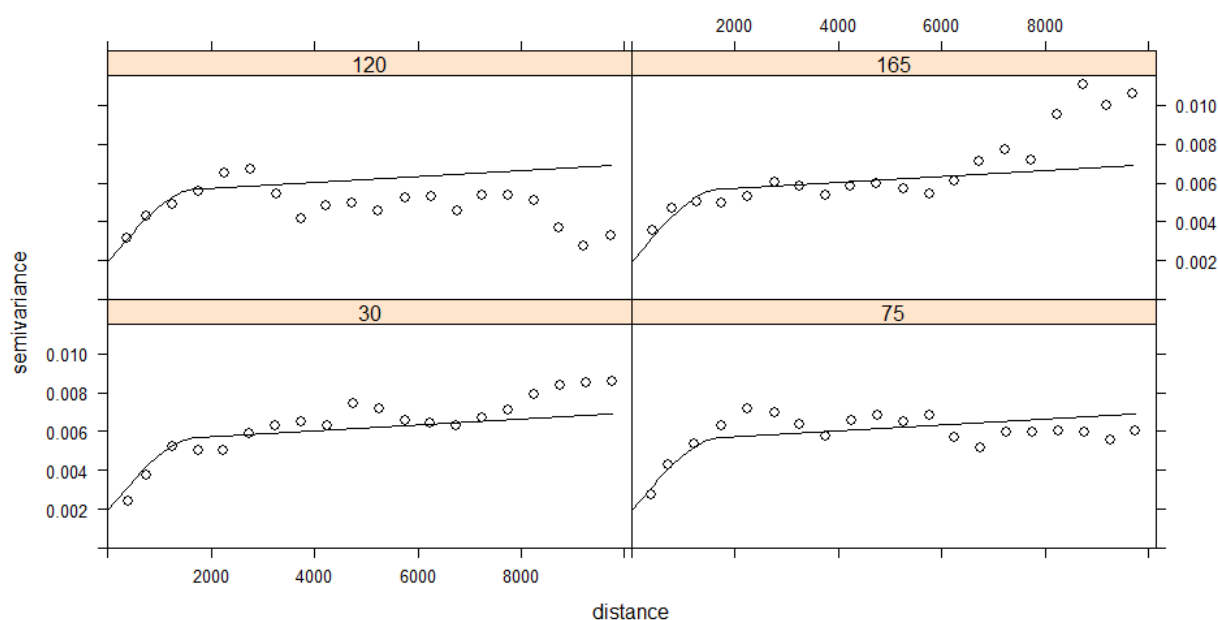


Рисунок 2.24 – Модель для вариограммы с иерархической структурой (сферическая + линейная) и эффектом самородков

На Рисунках 2.25 а) и б) приведены точки, для которых имеются данные о песчанистости, а также точки, для которых методом кригинга такие данные будут получены.

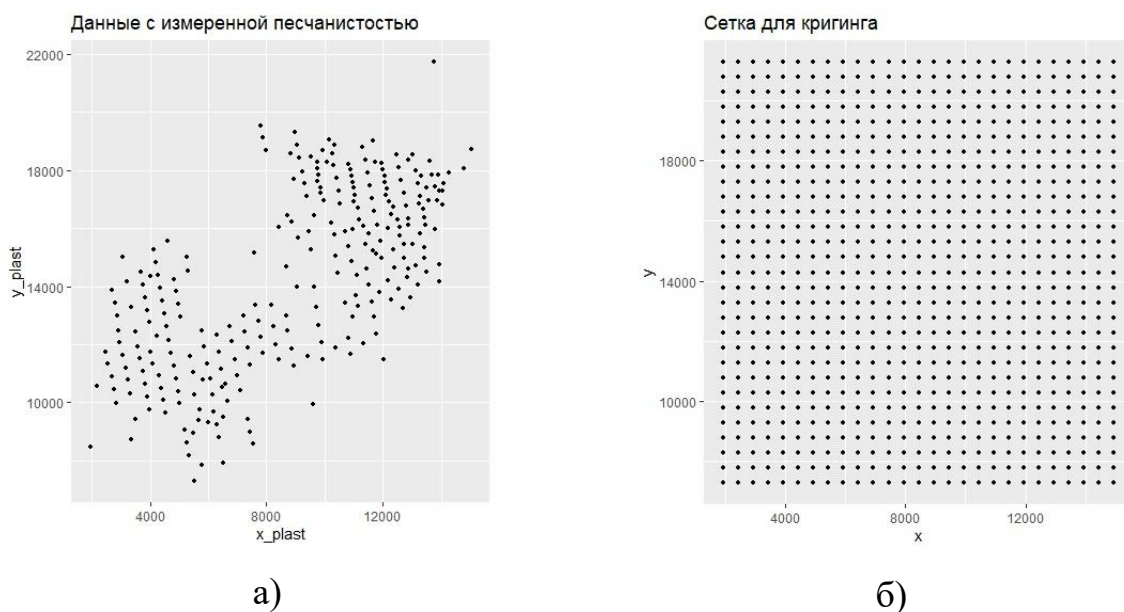


Рисунок 2.25 – Координаты точек, для которых имеются данные о песчаности (а), для которых данные будут получены методом кригинга (б)

Результат кригинга для заданной сетки по построенной модели представлен на Рисунке 2.26. Значения песчаности для каждой точки области восполнены с указанием координат точек и дисперсии оценки.

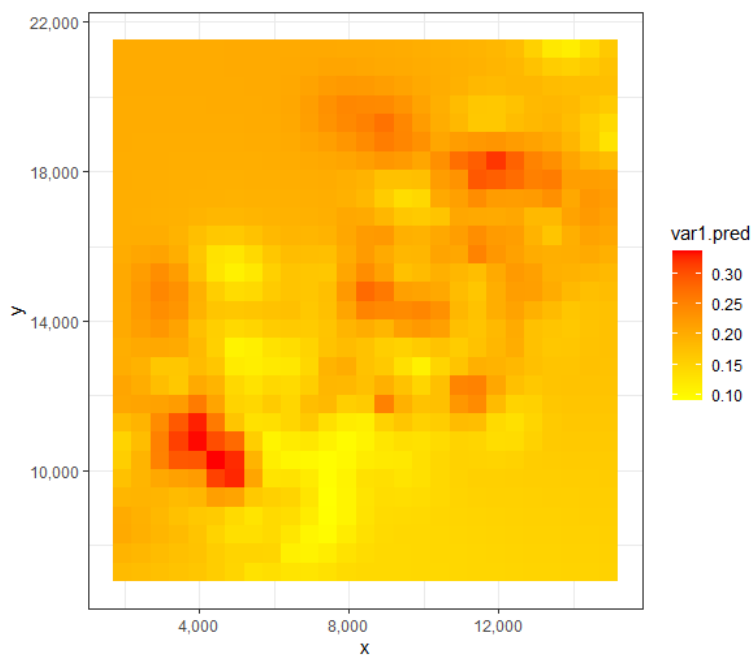


Рисунок 2.26 – Оцененная песчаность по поверхности области, заданной сеткой

## Выводы по второй главе

1. Все динамические данные по добыче на скважинах месторождения были проверены на корректность с использованием геофизического критерия на основании уравнения Дарси. Такой подход позволил определить некорректно записанные данные и внести изменения в них на основе уточнения информации из других баз данных, содержащих промысловые характеристики скважины. Таким образом, удалось избежать ошибок моделирования, получаемых при использовании некорректной входной информации.

2. Исследование временных рядов по промысловым характеристикам работы скважин с использованием статистического подхода позволило определиться с типами процессов, в том числе определить порядок их интеграции. Уточнение процессов с использованием тестов на структурные изменения во временных рядах, позволило учесть влияния мероприятий по повышению интенсивности по добыче нефти с «истории» промысловых показателей «работы» скважины. В результате такая детекция процессов добычи нефти позволяет избежать в дальнейшем ошибок спецификации моделей и получения ложных зависимостей.

3. Проведение тестов на коинтеграцию и причинность по Гренджеру между динамическими показателями по добыче на скважинах месторождения позволили определить длину лага запаздывания между связными показателями определиться с предварительной спецификацией моделей.

4. Различные методы формирования кустов, в том числе с учетом расстояния от нагнетательной скважины, позволили сформировать сценарии рассмотрения кустообразования для обучения в дальнейшем модели связности пластовых систем на основе оценки взаимовлияния скважин друг на друга.

5. Основанный на применении метода ДСММ подход позволил выделить однородные участки в рядах данных и, как следствие, строить более качественные модели для них, отбрасывая участки ряда, относящиеся принципиально к другому типу случайных процессов. Такой подход

предварительного преобразования временных рядов позволил повысить прогнозные свойства моделей, для которых использовался интервал, последний перед прогнозным с однородными данными.

6. Анализ вариограмм позволил выдвинуть гипотезу о наличии пространственных зависимостей, которые необходимо учитывать в модели связности пластовых систем. Проведённый кригинг позволил восполнить недостающую информацию по характеристикам скважины с привязкой к пространству.

### 3 МОДЕЛИ ОЦЕНКИ ВЗАИМОВЛИЯНИЯ СКВАЖИН И СВЯЗАННОСТИ ПЛАСТОВЫХ СИСТЕМ

#### 3.1 Алгоритм построения и тестирования модели байесовской векторной авторегрессии

Модель байесовской векторной авторегрессии (BVAR) имеет наибольшие перспективы для получения прогноза связанности пластовых систем на основе оценки взаимовлияния скважин друг на друга при корректном объёме данных для обучения соответствующих моделей [3]. Модель BVAR представляет систему одновременных авторегрессионных уравнений [80], которые описывают динамику каждой скважины куста с учётом распределения взаимного влияния по лагам:

$$Y_t = \Xi + \Phi_1 \cdot Y_{t-1} + \Phi_2 \cdot Y_{t-2} + \dots + \Phi_p \cdot Y_{t-p} + B_0 \cdot X_t + B_1 \cdot X_{t-1} + B_2 \cdot X_{t-2} + \dots + B_q \cdot X_{t-q} + \varepsilon_t, \quad (3.1)$$

где  $Y_t$  – вектора эндогенных переменных, дебитов жидкости/нефти в момент времени  $t$  скважин куста;

$X_t$  – вектора экзогенных переменных в момент времени  $t$  скважин куста,  $Y_{t-2}, \dots, Y_{t-p}, X_{t-1}, X_{t-2}, \dots, X_{t-q}$  – вектора лаговых переменных соответственно эндогенных и экзогенных;

$\Phi_1, \Phi_2, \dots, \Phi_p, B_0, B_1, \dots, B_q, \Xi$  – вектора коэффициентов, оцениваемые с помощью байесовского подхода.

Принцип моделирования взаимодействия скважин через систему уравнений, каждая из которых описывает поведение каждой скважины с учётом поведения соседних скважин, а также экзогенных переменных, учет лагов запаздывания в их взаимовлиянии позволяет отразить системную динамику куста. Особенность используемого подхода заключается в необходимости задания матрицы переходов для получения надёжных оценок.

Методика построения моделей и проверка её адекватности реальному моделируемому процессу проходит в соответствии с этапами:

1. Коэффициенты модели оцениваются иерархическим байесовским подходом, проверяется статистическая значимость коэффициентов авторегрессии (при переменных  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ ) зависимой (целевой) переменной  $Y_t$  и коэффициентов при лаговых экзогенных переменных (соответствующих признакам влияния  $X_{t-1}, X_{t-2}, \dots, X_{t-q}$ ). При проверке принимается либо отклоняется нулевая гипотеза о равенстве соответствующего коэффициента нулю при накладываемых уровнях значимости. В ходе численных экспериментов исходят из условия, что влияние на целевую переменную по каждой скважине оказывается значимой, если соответствующая рассчитанная статистика Стьюдента превышала по модулю критическое значение распределения на заданном уровне значимости ( $p=0,1; p=0,05$ ).

2. Оценка качества модели проводится на основе коэффициента детерминации  $R^2$ , позволяющего оценить «объясняющую силу» модели в относительных величинах. Коэффициент рассчитывается как разница между 1 и отношением остаточной дисперсии к общей дисперсии показателя. Чем ближе соответствующий коэффициент к единице, тем точнее отобранные признаки (авторегрессия зависимой переменной (дебит жидкости или объем добычи нефти на скважине)) и лаговые переменные независимых признаков (внутрипластовое давление, забойное давление, разница давлений, суммарный объем закачки в нагнетательные скважины) объясняют вариацию целевой переменной. Как показатель качества также рассматривается значение информационного критерия Акайке  $AIC = \ln(\sigma^2) + \frac{2k}{T}$ , где  $T$  – длина временного ряда,  $k$  – число степеней свободы модели (равно числу факторов в модели +1),  $\sigma^2$  – остаточная или объяснённая моделью дисперсия. Критерий используется не в оценочных целях, а для проведения селекции спецификаций модели: выбирается такая спецификация, для которой значение критерия минимально. Также рассматриваются критерии RMSE и MAPE, рассчитанные по всей ретроспективе ряда ( $y_t$  – фактическое значение в  $t$ -ый период времени,  $\hat{y}_t$  – прогнозное значение в  $t$ -ый период времени):

$$MAPE_{series} = \frac{1}{T} \cdot \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\%;$$

$$RMSE_{series} = \sqrt{\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{T}}$$

3. Для оценки состоятельности, эффективности и несмещённости, полученных на основе байесовских оценок коэффициентов анализируются автокорреляционные и кросскорреляционные функции остатков модели по каждому ряду как самой целевой переменной одного уравнения системы модели BVAR, так и их совместные взаимные корреляции [4]. В случае надёжности коэффициентов модели коррелограммы кросскорреляционной и автокорреляционной функций будут соответствовать коррелограммам процесса белого шума (математическое ожидание не отличается от нуля, дисперсия постоянна, автокорреляция между любыми уровнями такого процесса отсутствует). Для оценки устойчивости модели и пригодности ее для прогнозирования оцениваются корни соответствующего процессу характеристического уравнения. В случае сходимости процесса обратные корни по модулю не должны превышать единицы (комплекснозначные обратные корни должны лежать строго внутри единичного круга) [49].

4. Коэффициенты модели оценивались иерархическим байесовским подходом, проверялась статистическая значимость коэффициентов авторегрессии (при переменных  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ ) зависимой (целевой) переменной  $Y_t$  и коэффициентов при лаговых экзогенных переменных (признаков влияния  $X_{t-1}, X_{t-2}, \dots, X_{t-q}$ ). При проверке принималась либо отклонялась нулевая гипотеза о равенстве соответствующего коэффициента нулю при различных уровнях значимости. Считали, что влияние на целевую переменную по каждой скважине оказывается, если соответствующая рассчитанная статистика Стьюдента превышала по модулю критическое значение распределения на заданном уровне значимости ( $p=0,1$ ;  $p=0,05$ ).

Построение и тестирование модели BVAR месторождения с ежедневными данными: зависимая целевая переменная – объем жидкости н скважин;

независимые переменные – авторегрессия до четвертого лага включительно (4 дня) дебита жидкости целевой переменной (лаговые переменные); лаговые переменные до 4-ого лага включительно дебита жидкости на соседних скважинах по кусту; экзогенная переменная – разница давлений внутрислоевого и забойного для скважины, рассматриваемой в качестве целевой переменной, разница давлений внутрислоевого и забойного для скважин – соседей по кусту суммарный дебит закачки по всем нагнетательным скважинам в кусте. Для куста №2 (скважины 19 и 77) был в модель добавлен линейный тренд, так как объем дебита жидкости для этих скважин представлял собой тип процесса  $TS(I(0))$ . Для куста №23 (скважины 42, 97 и 110) динамика дебита жидкости по скважинам была дважды продифференцирована, так как относилась к типу процесса  $DS(I(2))$ , тогда как дебит жидкости по 97 скважине относился к процессу  $DS(I(0))$ .

Результаты оценок модели BVAR куста 12 из трех скважин представлены ниже в Таблице 3.1, метрики качества представлены в Таблице 3.2. Проверка надёжности оценок модели: коррелограммы остатков модели по каждой скважине и кросскоррелограммы взаимной корреляции остатков по каждой из скважин куста (Рисунок 3.1) соответствуют белому шуму; обратные корни по модулю меньше 1 (Рисунок 3.2). Вывод – модель BVAR по кусту №12 даёт надёжные и устойчивые результаты [44].

Графики фактических и прогнозных значений по модели для оценки дебита жидкости на 33, 90 и 91 скважинах соответственно на графиках 3.3, 3.4, 3.5.



Таблица 3.1 – Модель BVAR, куст -№ 12. Скважины 33, 90 и 91 месторождения №1

	Д_Ж_33 <sub>t-1</sub>	Д_Ж_33 <sub>t-2</sub>	Д_Ж_33 <sub>t-3</sub>	Д_Ж_33 <sub>t-4</sub>	Д_Ж_90 <sub>t-1</sub>	Д_Ж_90 <sub>t-2</sub>	Д_Ж_90 <sub>t-3</sub>	Д_Ж_90 <sub>t-4</sub>	
Дебит жидкости 33 скв	0,78948*	0,09816*	0,00157*	-0,01563*	-0,01588*	0,02111*	-0,00013**	-0,00048**	
Дебит жидкости 90 скв	-0,0752*	0,103022,	0,064048,	0,046945,	0,797748*	0,097185*	0,027587*	-0,00805*	
Дебит жидкости 91 скв	-0,02025*	0,012768*	0,005015*	0,002275**	-0,00207**	-0,00613**	0,002592**	0,001693**	
	Д_Ж_91 <sub>t-1</sub>	Д_Ж_91 <sub>t-2</sub>	Д_Ж_91 <sub>t-3</sub>	Д_Ж_91 <sub>t-4</sub>	Р_Дав_33	Р_Дав_90	Р_Дав_91	Нагнет	Const
Дебит жидкости 33 скв	-0,07825,	0,050851,	0,013595*	0,006666*	0,151223*	0,000343*	0,024566*	-0,00099**	-3,39
Дебит Жидкости 90 скв	0,041098,	0,060092,	0,008836,	-0,02207,	-0,00698,	0,159041*	-0,16131,	0,002204**	0,508799
Дебит жидкости 91 скв	0,859015*	0,025503*	0,033833*	0,041393*	0,002821*	0,00973**	-0,00745*	0,001749***	-4,55139

\*\*\* \*\* \* , , , . – p-уровень <0,001, <0,01, <0,05, <0,1 соответственно

Таблица 3.2 – Качество модели Модель BVAR, куст -№ 12. Скважины 33, 90 и 91 месторождения №1: Метрики качества

	$R^2$	F-статистика	MAPE (по всей истории ряда)
Дебит жидкости 33 скв	0,90	476,62***	4,37
Дебит жидкости 90 скв	0,93	760,74***	4,78
Дебит жидкости 91 скв	0,96	1225,07***	6,45

\*\*\* – p-уровень <0,001.

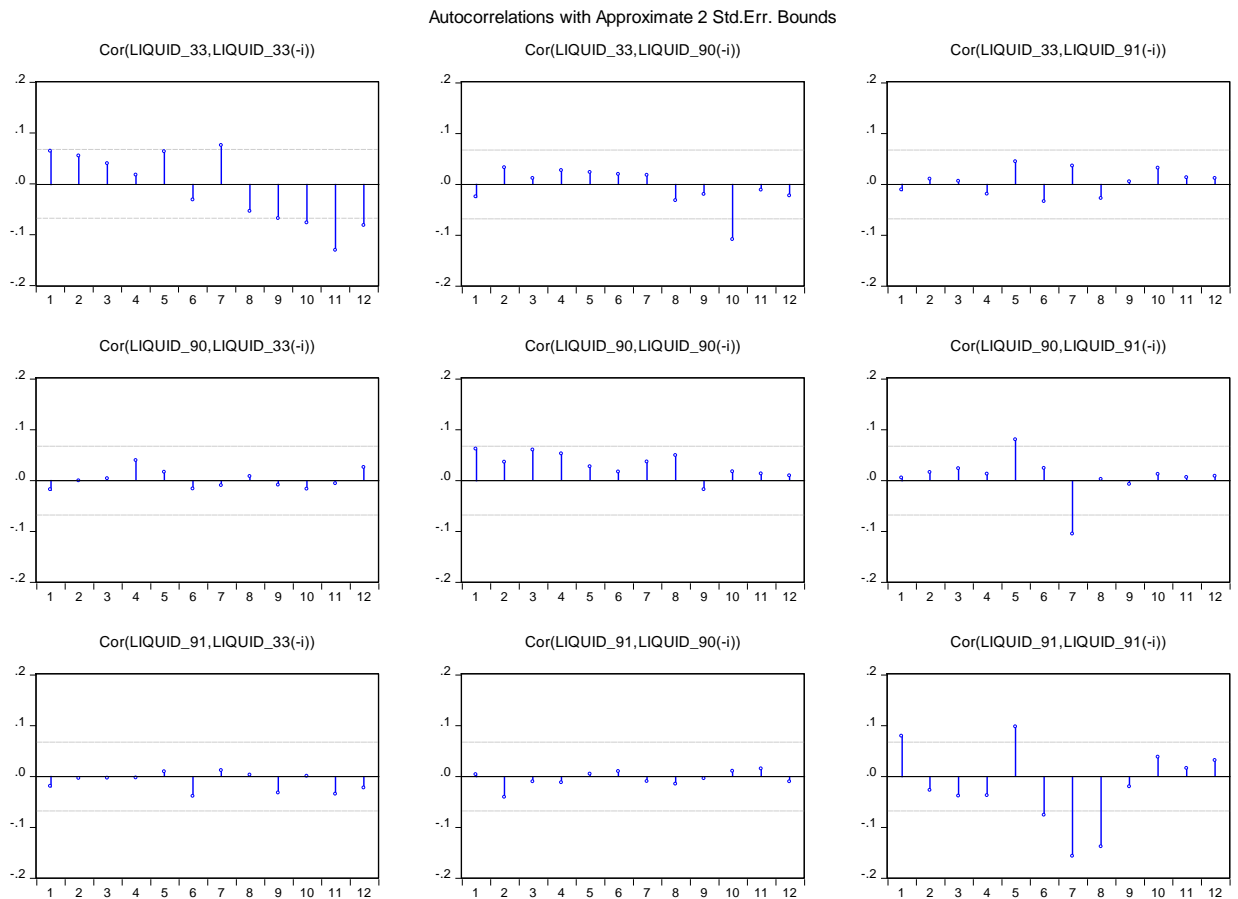


Рисунок 3.1 – Коррелограммы остатков модели и кросскоррелограммы взаимной корреляции остатков по каждой из скважин куста

Корни характеристического уравнения

Корни	Модуль
0,958745	0,958745
0,946333	0,946333
0,860091	0,860091
0,094728 - 0,356307i	0,368684
0,094728 + 0,356307i	0,368684
-0,307404	0,307404
-0,141582 - 0,239838i	0,278509
-0,141582 + 0,239838i	0,278509
0,270207	0,270207
-0,174295 - 0,162694i	0,238429
-0,174295 + 0,162694i	0,238429
0,160576	0,160576

Inverse Roots of AR Characteristic Polynomial

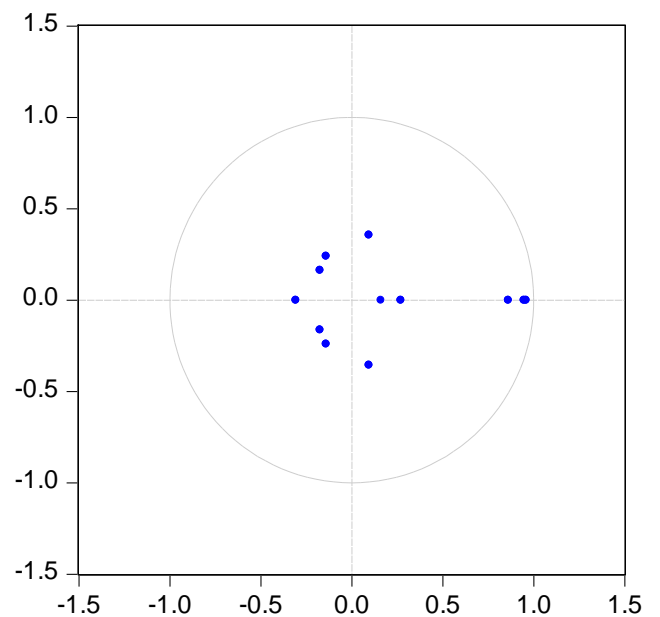


Рисунок 3.2 – Корни характеристического уравнения модели (таблица и график)

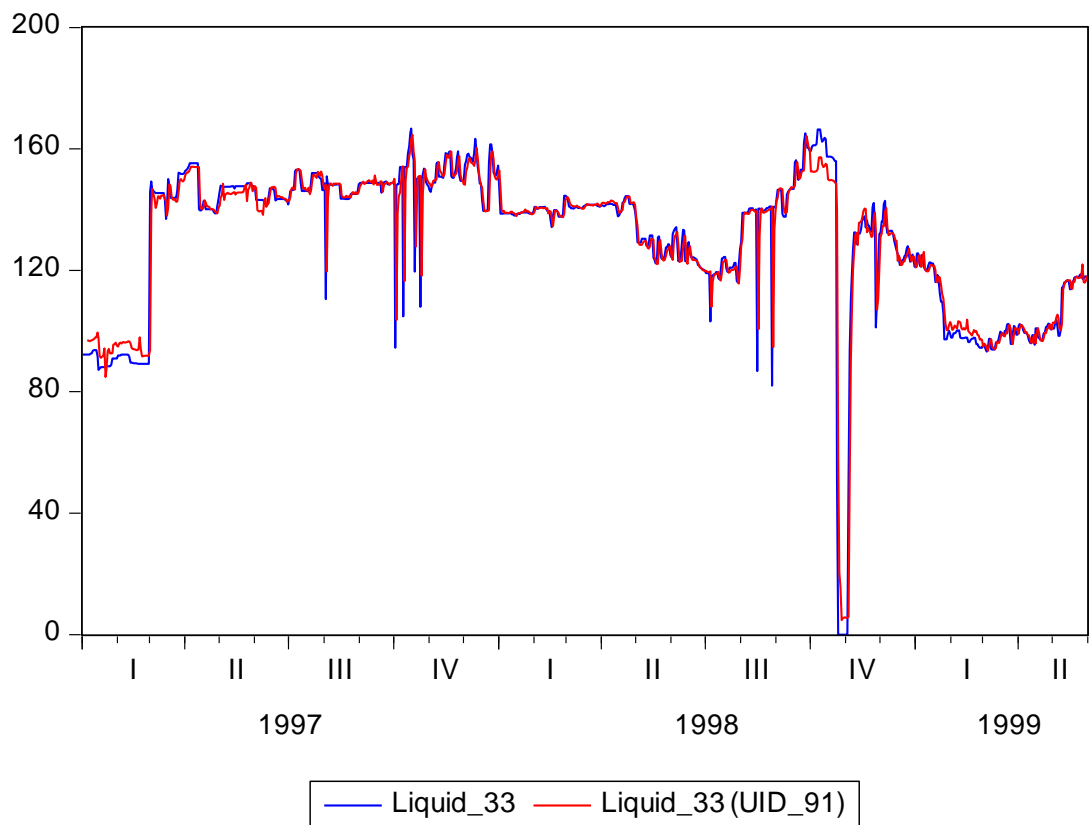


Рисунок 3.3 – График фактических и прогнозных значений по модели для оценки дебита жидкости на 33 скважине

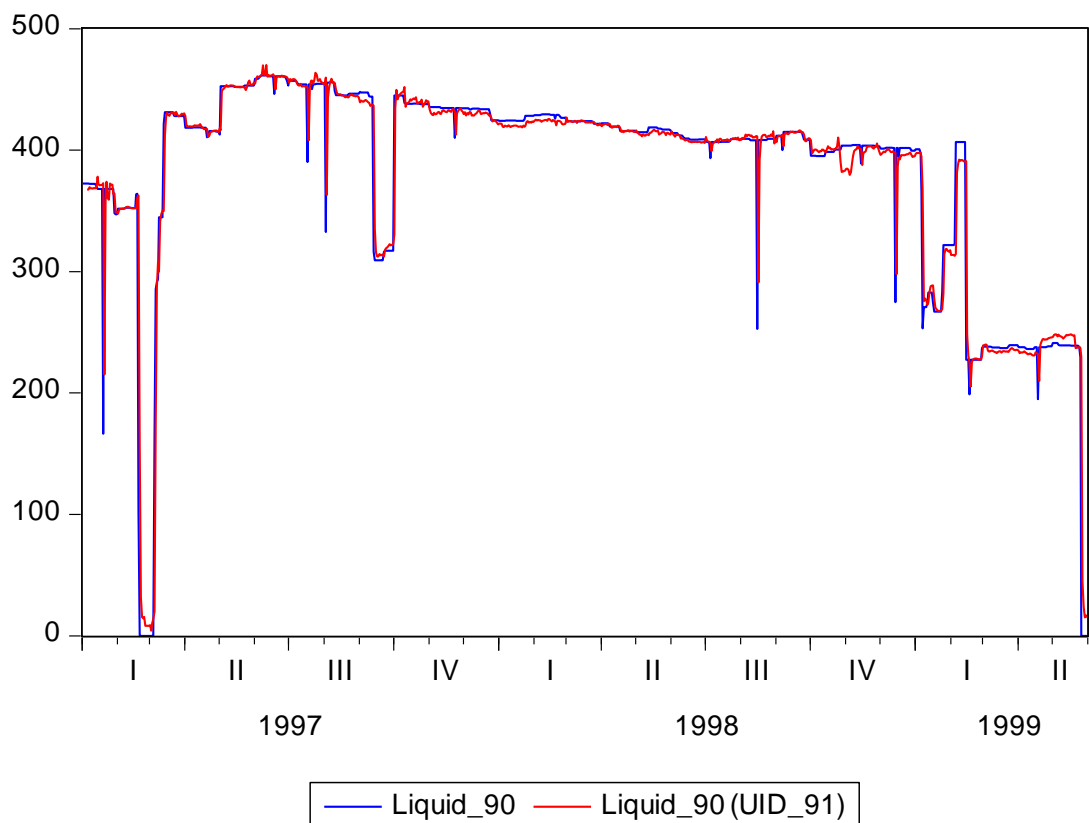


Рисунок 3.4 – График фактических и прогнозных значений по модели для оценки дебита жидкости на 90 скважине



Рисунок 3.5 – График фактических и прогнозных значений по модели для оценки дебита жидкости на 91 скважине

Все построенные модели по 24-м кустам оказались надежными и устойчивыми. Коэффициенты моделей состоятельны и эффективны. Во всех моделях наблюдается статистическая значимость зависимости объема дебета жидкости добывающих скважин от объема жидкости соседних по кусту добывающих скважин с лагом до 4 дней включительно. Также статистическая значимость ( $p < 0,1$ ) имеется для влияния разницы забойного и внутрислоевого давлений на объем добываемой жидкости на скважине в текущий момент времени. Разница давлений на соседних по кусту скважинах иногда оказывает, а иногда нет статистически значимое влияние на дебит жидкости рассматриваемой скважины. Суммарный дебит закачки по всем нагнетательным скважинам в кустах (за исключением кустов 1 (скважины 9 и 69) и 6 (скважины 20 и 73)) оказывает статистически значимое влияние на объем дебита жидкости добывающих скважин в рассматриваемом месторождении №1.

### 3.2 Тестирование данных на наличие пространственных корреляционных зависимостей

Основой методики построения динамических пространственных моделей, позволяющих оценить связанность пласта, является пространственная панельная модель, позволяющая учесть как влияние отсроченных эффектов за счёт введения в модель лагов запаздывания, так и пространственные эффекты влияния в добычи нефти за счёт введения пространственного лага. Под панельными данными здесь понимается совокупность временных рядов, собранных на основе истории «работы» каждой скважины. Панельные данные содержат информацию об одних и тех же объектах, число которых равно  $n$ , наблюдавшихся на протяжении  $T$  периодов времени [18]. Обозначаются панельные данные обычно буквами латинского алфавита с двумя индексами  $i$  (индекс объектов, скважин) и  $t$  (индекс моментов времени). Например,  $y_{it}$ , где  $i=1, \dots, n$ ;  $t=1, \dots, T$ . Преимуществом использования структуры панельных данных является возможность учесть ненаблюдаемые эффекты (закономерности в распределении показателей), присущие исключительно индивидуально для каждого объекта наблюдения (в рассматриваемом случае – скважин) за счёт определения так называемых фиксированных эффектов. Под пространственными моделями понимаются модели, учитывающие влияние на текущее значение показателя для рассматриваемого объекта значений соседних к нему объектов. Под пространственным лагом соответственно понимают величину «соседства» к объекту – это может быть обратное расстояние, а может быть простая смежность (есть непосредственное соседство объектов, или его нет). Учитывается пространственный лаг за счёт весовой пространственной матрицы  $w_{ij}$ , элементы которой обратные расстояния между скважинами или простое соседство в определённом радиусе. Под пространственно-панельной моделью понимается модель панельной структуры, учитывающая пространственные связи за счёт введения весовой пространственной матрицы между объектами.

«С целью определения возможности применения инструментов панельного анализа к показателям добычи нефти были проведены тесты на наличие единичных корней в процессах, описываемых панельными данными» [33]. Данные тестировались для месторождения, состоящего из скважин. На единичные корни проверялась панель по данным дебита жидкости в ежедневной динамике в период с января 1997 по октябрь 1999 по 82 добывающим скважинам из 115 скважин на месторождении. В Таблице 3.3 приведены результаты проверки панельной переменной с использованием различных тестов спецификации (без включения дополнительной константы и тренда, с включением их) [39, 90].

Затем, согласно последовательности, предложенной [26], «были проведены тесты на наличие ненаблюдаемых эффектов в остатках модели (тест Вулдриджа), наличия последовательной корреляции и кросс-секционную зависимость остатков панельной модели с фиксированными эффектами (тест Песарана), наличия авторегрессии первого порядка (AR(1)) в случайных эффектах, наличия авторегрессии первого порядка или процесса скользящего среднего первого порядка в случайных эффектах (тест множителей Лагранжа Балтаги-Ли), тест на локальную устойчивость панельной модели с альтернативной гипотезой о наличии случайных эффектов как авторегрессии первого порядка от ошибок модели (тест Бера-Соса-Эскудеро-Юна), тест Вулдриджа на наличие сериальной корреляции на первых разностях панельных данных (модификация теста Броша-Погана для панельных моделей)». В качестве модели рассматривалось влияние внутрипластового давления на дебит жидкости скважин. В качестве гипотез рассматривали влияние внутрипластового давления на дебит жидкости, также рассматривали влияние забойного давления (но статистическая значимость такого влияния была менее выраженной). Результаты тестов с указанием альтернативных гипотез представлены в Таблице 3.4.

Таблица 3.3 – Результаты тестов на наличии единичных корней в показателе дебит жидкости

Наименование теста на наличие единичных корней	Модель без индивидуальной константы и тренда	Индивидуальная константа	Индивидуальная константа и линейный тренд по времени
Тест Левин-Ли-Чу	Подбор величины лага запаздывания на основе значения теста Шварца		
	$z = -6,711; p < 0,0001$	$z = -21,384; p < 0,0001$	$z = -37,699; p < 0,0001$
	Подбор величины лага запаздывания на основе значения теста Акайке		
	$z = -6,716; p < 0,0001$	$z = -21,138; p < 0,0001$	$z = -37,043; p < 0,0001$
	Подбор величины лага запаздывания на основе значения теста Холла		
	$z = -6,708; p < 0,0001$	$z = -21,102; p < 0,0001$	$z = -37,201; p < 0,0001$
Тест Има-Песарана-Шина	-	$Wtbar = -42,113; p < 0,0001$	$Wtbar = -49,129; p < 0,0001$
Маддала-Бу	Подбор величины лага запаздывания на основе значения теста Шварца		
	$\chi^2 = 233,71; df = 164; p = 0,0003$	$\chi^2 = 2412,3; df = 164; p < 0,0001$	$\chi^2 = 2600,1; df = 164; p < 0,0001$
	Подбор величины лага запаздывания на основе значения теста Айкайке		
	$\chi^2 = 234,05; df = 164; p < 0,0001$	$\chi^2 = 2384,9; df = 164; p < 0,0001$	$\chi^2 = 2567,4; df = 164; p < 0,0001$
	Подбор величины лага запаздывания на основе значения теста Холла		
	$\chi^2 = 197,01; df = 164; p = 0,04$	$\chi^2 = 2382,2; df = 164; p < 0,0001$	$\chi^2 = 2572,3; df = 164; p < 0,0001$
Модифицированный тест Чои	Подбор величины лага запаздывания на основе значения теста Шварца		
	$Pm = 3,8489; p < 0,0001$	$Pm = 124,14; p < 0,0001$	$Pm = 134,51; p < 0,0001$
	Подбор величины лага запаздывания на основе значения теста Айкайке		
	$Pm = 3,868; p < 0,0001$	$Pm = 122,63; p < 0,0001$	$Pm = 132,7; p < 0,0001$
Подбор величины лага запаздывания на основе значения теста Холла			
	$Pm = 3,8381; p < 0,0001$	$Pm = 122,48; p < 0,0001$	$Pm = 132,98; p < 0,0001$
Тест Чои на обратные нормальные единичные корни	Подбор величины лага запаздывания на основе значения теста Шварца		
	$z = -5,6057; p < 0,0001$	$z = -5,6057; p < 0,0001$	$z = -5,6057; p < 0,0001$
	Подбор величины лага запаздывания на основе значения теста Айкайке		
	$z = -5,6177; p < 0,0001$	$z = -5,6177; p < 0,0001$	$z = -5,6177; p < 0,0001$
	Подбор величины лага запаздывания на основе значения теста Холла		
	$z = -5,5991; p < 0,0001$	$z = -5,5991; p < 0,0001$	$z = -5,5991; p < 0,0001$
Тест Хадри	-	$z = 1600,1; p < 0,0001$	$z = 980,5; p < 0,0001$

Таблица 3.4 – Результаты проведённых предварительных тестов для уточнения спецификации панельных моделей регрессии влияния внутрипластового давления на дебит жидкости

Наименование теста	Значение	Альтернативная гипотеза
Критерий Вулдриджа для ненаблюдаемых индивидуальных эффектов. $H_0$ : нет ненаблюдаемых эффектов в остатках.	$z = 5,8262$ ; $p < 0,0001$	Присутствует ненаблюдаемый индивидуальный эффект в остатках
Локально устойчивые тесты Бера-Соса-Эскудеро-Юна	$\chi^2 = 2035,1$ , $df = 1$ , $p < 0,0001$	Наличие случайных эффектов и авторегрессии первого порядка (AR(1)) в остатках
Односторонний тест Бера-Соса-Эскудеро-Юна	$z = 5600,5$ , $p < 0,0001$	Наличие авторегрессии первого порядка (AR(1)) в остатках и случайных эффектов
Совместный тест Балгати и Ли на случайные эффекты и последовательную корреляцию при нормальности и гомоскедастичности идиосинкразических ошибок	$\chi^2 = 31449847$ , $df = 2$ , $p < 0,0001$	Наличие последовательной корреляции и/ или случайных эффектов
Балгаги и Ли двусторонний тест ЛМ	$\chi^2 = 77251$ , $df = 1$ , $p < 0,0001$	AR(1)/MA(1) ошибки в модели со случайными эффектами
Балгаги и Ли односторонний тест ЛМ	$z = 277,94$ , $p < 0,0001$	AR(1)/MA(1) ошибки в модели со случайными эффектами
Тест Вулдриджа на последовательную корреляцию в моделях с фиксированными эффектами	$F = 13180$ , $df_1 = 1$ , $df_2 = 85524$ , $p < 0,0001$	Последовательная корреляция
Тест Вулдриджа на последовательную корреляцию	$F = 7,5206$ , $df_1=1$ , $df_2=85442$ , $p < 0,0001$	Последовательная корреляция в дифференцированных ошибках
Тест Вулдриджа на последовательную корреляцию	$F = 434,95$ , $df_1 = 1$ , $df_2 = 85442$ , $p < 0,0001$	Последовательная корреляция в исходных ошибках
Тест Песарана на кросс-секционную зависимость	$z = 77,375$ , $p < 0,0001$	Кросс-секционная зависимость

Для оценки наличия связи между панельными переменными необходимо провести процедуру для определения панельной коинтеграции. Панельная



коинтеграция позволяет сделать вывод о том, что построение модели между признаками возможно на исходных уровнях данных, не прибегая к процедурам отклонения от тренда либо взятия разностей различного порядка (зависит от типа случайного панельного процесса) [70].

Для проверки «панельной коинтеграции» был проведен тест Педрони для тестирования коинтеграционного соотношения между панельными переменными». «В качестве нулевой гипотезы применяется предположение об отсутствии коинтеграции между панельными переменными, означающее наличие единичных корней в остатках после построения коинтеграционного соотношения между ними» [18]. Результаты теста Педрони представлены в Таблице 3.5. Здесь ADF-статистика – это статистика расширенного теста Дики-Фуллера.

Пары переменных, характеризующих добычу нефти на скважинах, практически все коинтегрированы согласно проведенному тесту Педрони, то есть между данными панельными переменными имеет смысл построения уравнений связной регрессии, не получив при этом ложного результата.

Особенностью именно панельного моделирования является выбор лучшей спецификации модели с учётом введения/не введения в уравнение индивидуальных панельных эффектов. Для этого проводят следующие тесты [59]:

- теста Вальда (модель с фиксированными эффектами сравнивается со сквозной регрессией);

- теста Хаусмана (модель с фиксированными эффектами по периодам и кросс-секциям сравнивается с моделью со случайными эффектами по периодам и кросс-секциям);

- теста Бреуша-Пагана (модель со случайными эффектами по периодам и кросс-секциям сравнивается со сквозной регрессией).

Таблица 3.5 – Результаты теста Педрони на коинтеграцию между панельными переменными

Панельные переменными	Спецификация теста	ADF-статистика ( <i>p</i> -уровень)	Вывод о коинтеграции
1	2	3	4
Нефть – Пластовое давление	Без индивидуальных констант и индивидуального тренда	-11,39 (0,00)	+
	Включение индивидуальных констант	-1,34 (0,09)	–
	Включение индивидуальных констант и индивидуального тренда	-14,87 (0,00)	+
Нефть – Дебит газа	Без индивидуальных констант и индивидуального тренда	-9,90 (0,00)	+
	Включение индивидуальных констант	-9,15 (0,00)	+
	Включение индивидуальных констант и индивидуального тренда	-4,69 (0,00)	+
Нефть – Забойное давление	Без индивидуальных констант и индивидуального тренда	-7,16 (0,00)	+
	Включение индивидуальных констант	-7,33 (0,00)	+
	Включение индивидуальных констант и индивидуального тренда	-13,02 (0,00)	+
Нефть - Закачка	Без индивидуальных констант и индивидуального тренда	-7,86 (0,00)	+
	Включение индивидуальных констант	-4,53 (0,00)	+
	Включение индивидуальных констант и индивидуального тренда	-4,19 (0,00)	+
Нефть - Жидкость	Без индивидуальных констант и индивидуального тренда	-5,79 (0,00)	+
	Включение индивидуальных констант	-6,82 (0,00)	+
	Включение индивидуальных констант и индивидуального тренда	-12,69 (0,00)	+

Результаты тестирования спецификации сведены в Таблицу 3.6.

Результаты проверки сложных гипотез на спецификацию панельных тестов показали необходимость построения двунаправленной модели с фиксированными эффектами по периодам и по объектам (скважинам).

Так как основной целью анализа является получение панельной пространственной модели взаимного влияния дебита жидкости на скважинах были проведены тесты на наличие пространственной связности в панельных моделях. В качестве матрицы пространственной связности скважин использовалась простая матрица соседства.

Матрица соседства – это квадратная матрица, размер которой равен числу рассматриваемых скважин, а элементами являются 0, если скважины не являются соседями (не имеют общей границы), и 1 иначе.

Таблица 3.6 – Результаты тестов на панельную коинтеграцию

Тест	Значение расчетной статистики теста ( <i>p</i> -уровень) для выбранной спецификации			Вывод о спецификации
	для эффектов по периодам	для эффектов по кросс- секциям	для эффектов по периодам и кросс-секциям	
Тест Вальда	47,44 (0,00)	0,05 (1,00)	20,99 (0,00)	Модель с фиксированными эффектами по периодам, а также по периодам и кросс- секциям предпочтительнее модели сквозной регрессии
Тест Хаусмана	97,88 (0,00)	0,04 (0,98)	169,68 (0,00)	Гипотеза H <sub>0</sub> о наличии случайных эффектов по периодам и кросс- секциям в модели отклоняется
Тест Бреуша- Пагана	31105,04 (0,00)	41,05 (0,00)	31146,09 (0,00)	Гипотеза H <sub>0</sub> о наличии случайных эффектов по периодам и кросс- секциям в модели отклоняется

Для каждой скважины считалось расстояние до всех остальных скважин. Выбиралось *n* ближайших скважин, которые считались соседями на заданном

расстоянии. В данном анализе выбраны расстояния 750 м и 1000 м. На 500 м не у всех скважин в окружении есть соседи.

В качестве весов использовался метод Radial Distance Weights. В данном методе расстояние является важным критерием пространственного влияния. Выбиралось пороговое расстояние  $d$ , за пределами которого нет прямого пространственного влияния между скважинами, соответствующая матрица весовых коэффициентов радиального расстояния  $W$  имеет пространственные веса между  $i$ -ой и  $j$ -ой скважинами вида:

$$w_{ij} = \begin{cases} 1, & 0 \leq d_{ij} \leq d \\ 0, & d_{ij} > d \end{cases} \quad (3.2)$$

Положительным моментом построения матрицы соседства является учет влияния только близко располагающихся скважин, но при это результаты зависят от выбранного расстояния.

С целью определения «возможности построения пространственных моделей на панельных данных были проведены тесты на пространственную связность (различные варианты тестов Балтаги-Сонга-Коха и тестов Хаусмана)» [60], представленные в Таблице 3.7. Результаты тестов показывают, что в данных панелей имеется пространственная связность, учесть которую можно за счёт построения моделей спецификации пространственного лага.

В проведённых тестах нулевая гипотеза о состоятельности коэффициентов в обеих моделях (с фиксированными и случайными эффектами) отвергается в пользу гипотезы о том, что в модели со случайными эффектами коэффициенты несостоятельны только в случае их построения с пространственным лагом или с пространственной ошибкой. В случае построения моделей и с пространственным лагом, и с ошибкой отмечается состоятельность коэффициентов в обеих моделях по оценке пространственной авторегрессии дебита жидкости.

Таким образом, в исследовании пространственного взаимовлияния дебита жидкости скважин с учётом экзогенного влияния внутрипластового давления целесообразно построение моделей спецификации SAC random (панельная модель пространственной автокорреляции со случайными панельными эффектами), SAR

within (панельная модель пространственной авторегрессии с фиксированными панельными эффектами), SEM within (панельная модель пространственной ошибки со фиксированными панельными эффектами) [69].

Таблица 3.7 – Тесты на наличие пространственных эффектов в панельных моделях регрессии

Наименование теста	Значение	Альтернативная гипотеза
Тест Балтаги-Сонга-Коха	$LM-H = 31454906,$ $p < 0,000$	Присутствуют случайные эффекты по скважинам и пространственная автокорреляция
Тест Балтаги-Сонга-Коха на маржинальный эффект	$LM_1 = 5607,8,$ $p < 0,000$	Присутствуют случайные эффекты
Тест Балтаги-Сонга-Коха на маржинальный эффект	$LM_2 = 84,224,$ $p < 0,000$	Присутствует пространственная автокорреляция
Условный тест Балтаги-Сонга-Коха	$LM^*-\lambda = 13,517,$ $p < 0,000$	Присутствует пространственная автокорреляция
Условный тест Балтаги-Сонга-Коха	$LM^*-\mu = 5151,6,$ $p < 0,000$ $LM^*-\mu = 5572,4,$ $p < 0,000$	Присутствуют эффекты по скважинам
Модифицированный тест Хаусмана для пространственных моделей - с пространственным лагом; - с пространственной ошибкой; - с пространственным лагом и ошибкой	$\chi^2 = 5,0273, df = 1,$ $p = 0,02495$ $\chi^2 = 5,0807, df = 1,$ $p = 0,02419$ $\chi^2 = 0,54213, df = 1,$ $p = 0,4616$	Одна модель не совместима

Для оценки пространственной связности добычи жидкости на скважинах использовали также инструменты пространственной статистики [45].

В качестве исходных данных для проверки гипотез о пространственных корреляциях использованы ежедневные данные по одному месторождению в динамике с января 1997 по октябрь 1999. Всего 115 скважин, из них добывающих – 82. По добывающим скважинам: номер скважины, дата наблюдения, ежедневная добыча нефти (в м<sup>3</sup>), ежедневная добыча жидкости (в м<sup>3</sup>), FormationPressure (Внутрипластовое давление, МПа) и BottomHolePressure (Забойное давление,

МПа). По нагнетательным скважинам: номер скважины, дата наблюдения, объем закачиваемой ежедневно жидкости (м<sup>3</sup>). Для каждой скважины также имеется две координаты ее местоположения.

Для выявления пространственных зависимостей показателей ежедневной добычи жидкости и суммарной добычи жидкости за месяц с учетом расположения скважин на месторождении использовались индексы Морана и Гири.

Индекс Морана измеряет пространственную автокорреляцию, то есть оценивает наличие корреляционной связи между объектами, сдвинутыми на пространственный лаг друг от друга.

$$I_M = \frac{n}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum \sum w_{ij} (H_i - \bar{H})(H_j - \bar{H})}{\sum_i (H_i - \bar{H})^2},$$

где  $n$  – число скважин;

$w_{ij}$  – элемент матрицы пространственных весов, учитывающий соседство между  $i$  и  $j$  скважинами;

$\bar{H}$  – среднее значение показателя добычи;

$H_i$  – показатель добычи на  $i$ -ой скважине.

Статистическая значимость индекса Морана проверяется на основе стандартного  $z$ -критерия.

«Значение индекса Морана изменяется от -1 до 1:

при  $I > 0$  имеется положительная пространственная автокорреляция, т. е. в целом значения наблюдений в соседних территориях являются подобными;

при  $I < 0$  – отрицательная автокорреляция, т. е. в целом значения уровня смертности наблюдений в соседних территориях отличаются;

при  $I = 0$  значения наблюдений в соседних территориях расположены случайным образом» [48].

Важным этапом анализа полученных данных является построение пространственной диаграммы рассеяния Морана. По оси абсцисс откладываются стандартизированные  $z$ -значения исследуемого показателя, а по оси ординат – значения пространственного фактора  $Wz$ . На диаграмме отображается линия

регрессии  $Wz$  на  $z$ , наклон, который равен коэффициенту общей пространственной автокорреляции  $I$  при стандартизированной матрице весов [41].

Коэффициент пространственной автокорреляции показывает степень линейной взаимосвязи между вектором  $z$  центрированных значений добычи нефти и вектором  $Wz$  пространственно взвешенных центрированных значений добычи на соседних скважинах.

Индекс Гири изменяется от 0 до 2 и также служит для определения пространственной зависимости между исследуемыми показателями, при этом если индекс в пределах ноль – единица, то наблюдается положительная пространственная автокорреляция.

При использовании матрицы соседства для учета пространственных зависимостей определяется влияния только близко располагающихся скважин, но при этом результаты зависят от выбранного расстояния для «оценки соседства». В настоящем исследовании для рассматриваемого месторождения из 82 добывающих скважин анализировали соседние скважины, определённые по следующему принципу соседства: расстояние до ближайшей скважины – до 750 м, расстояние до ближайшей скважины – до 1000 м. В зависимости от выбора расстояния определяется число соседних добывающих скважин у каждой из скважин месторождения. Распределение по количеству соседних скважин представлено на Рисунке 3.6 – для «соседей» в близости до 750 м и Рисунке 3.7 – для «соседей» в близости до 1000 м.



Рисунок 3.6 – Распределение по количеству соседних скважин в близости до 750м



Рисунок 3.7 – Распределение по количеству соседних скважин в близости до 1000 м

В результате карта соседства скважин имеет различный вид для скважин в радиусе 750 м (Рисунок 3.8, а)) и в радиусе 1000 м (Рисунок 3.8, б)).



## Карта соседства добывающих скважин на месторождении

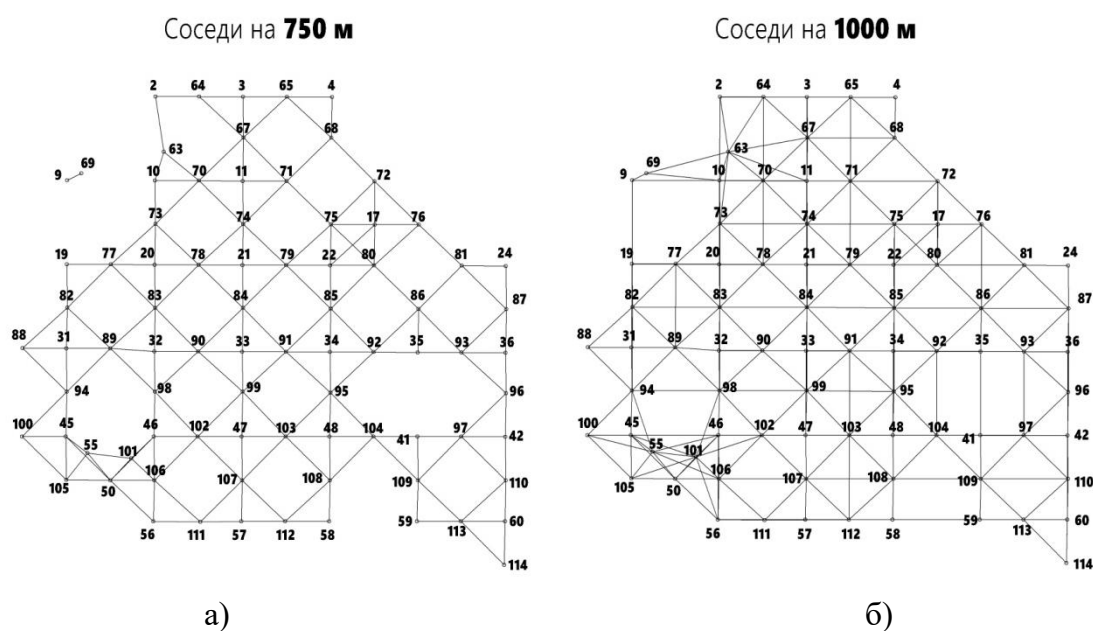


Рисунок 3.8 – Карта соседства скважин а) для скважин в радиусе 750 м, б) скважин в радиусе 1000 м

Так как данные по месторождению были динамические, в частности ежедневные, то индексы пространственной корреляции Морана и Гири добычи дебита жидкости считались за каждый день по 82 скважинам, между которыми определены пространственные связи в виде матрицы соседства. Также провели агрегацию ежедневных данных по дебиту жидкости в ежемесячную динамику. На Рисунках 3.9 и 3.10 изображены значения индекса Морана и соответствующая ему статистическая значимость (считали, что пространственная корреляция статистически значима, если уровень значимости  $p < 0,05$ ) в ежедневной и агрегированной ежемесячной динамике соответственно. На графике индексы Морана считали при определении соседних скважин в радиусе до 750 и до 1000 м соответственно. Как видно из графиков, все индексы статистически значимы. Но коэффициенты пространственных зависимостей проявляются сильнее при определении соседних скважин на расстоянии до 750 м. Что логично: более близко расположенные к друг другу скважины теснее связаны.



Рисунок 3.9 – Коэффициенты корреляции Морана и их статистическая значимость для ежедневного дебита жидкости добывающих скважин в радиусе 750 м, для добывающих скважин в радиусе 1000 м

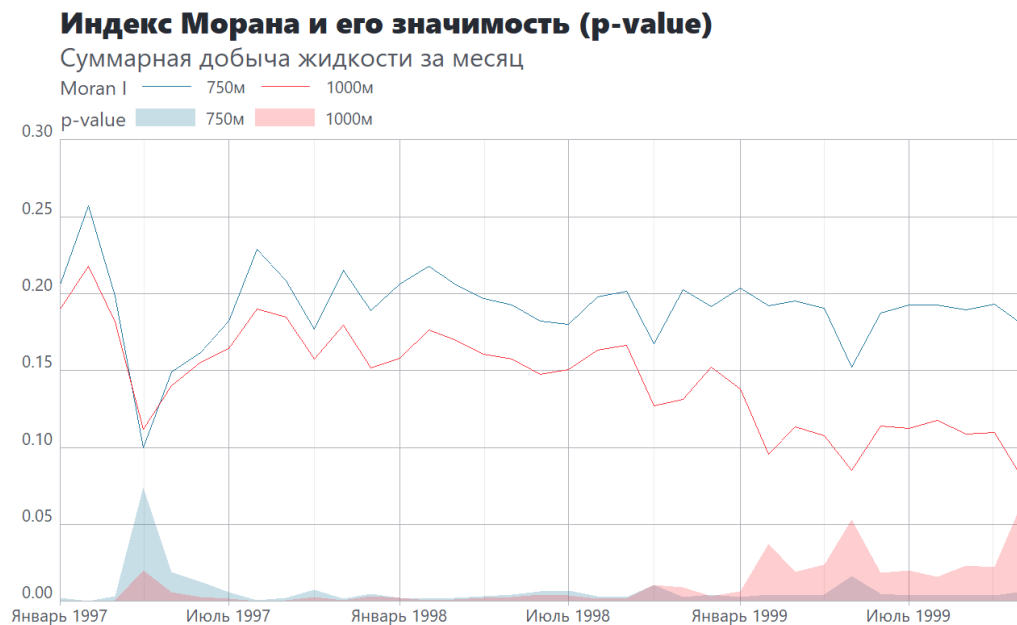


Рисунок 3.10 – Коэффициенты корреляции Морана и их статистическая значимость для ежемесячного дебита жидкости добывающих скважин в радиусе 750 м, для добывающих скважин в радиусе 1000 м

### Индекс Geary и его значимость (p-value)

Ежедневные данные Добычи Жидкости

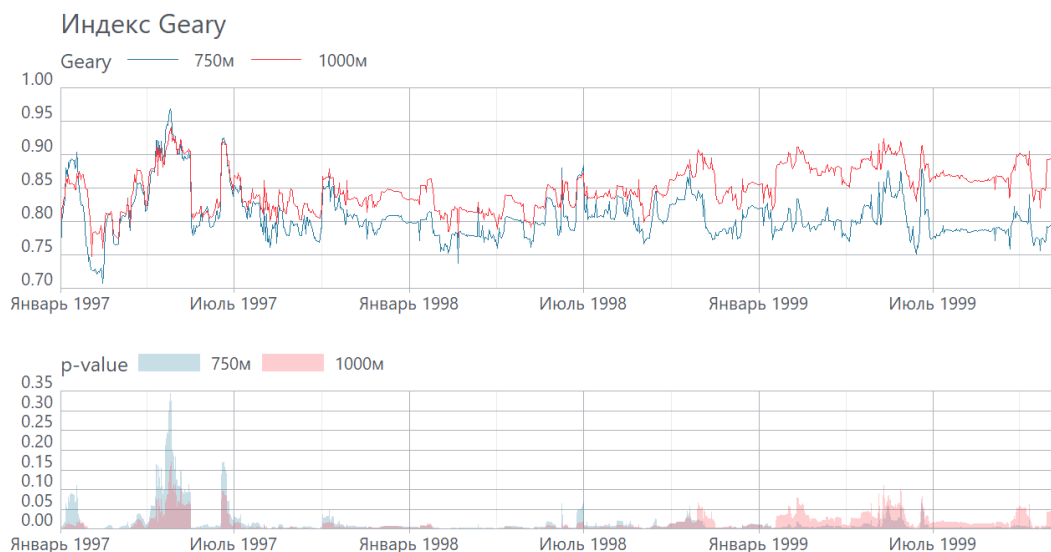


Рисунок 3.11 – Коэффициенты корреляции Гири и их статистическая значимость для ежедневного дебита жидкости добывающих скважин в радиусе 750 м, для добывающих скважин в радиусе 1000 м

### Индекс Geary и его значимость (p-value)

Суммарная добыча жидкости за месяц

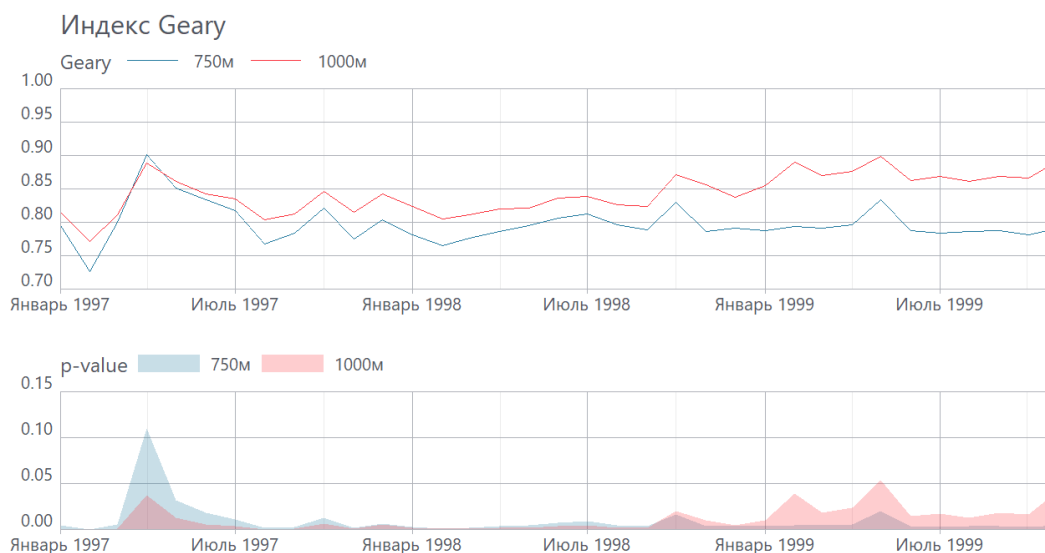


Рисунок 3.12 – Коэффициенты корреляции Морана и их статистическая значимость для ежедневного дебита жидкости добывающих скважин в радиусе 750 м, для добывающих скважин в радиусе 1000 м

По такому же принципу анализа для индексов Морана были проанализированы индексы пространственной корреляции Гири. Результаты представлены на графиках 3.11 и 3.12 для ежедневной и ежемесячной динамики соответственно. Как видно из графиков индексы пространственной корреляции Гири для соседства между скважинами в пределах 1000 м имеют более высокие значения, чем индексы пространственной корреляции Гири для соседства между скважинами в пределах 750 м.

Анализ рассчитанных значений индексов корреляции Морана и Гири показали, что существует положительная пространственная авторегрессионная зависимость средней добычи нефти скважины с соседними скважинами. Для оценки количественно выраженной такой связи целесообразно построение моделей с пространственным лагом [21].

### 3.3 Методика построения динамических моделей пространственной статистики с учетом отсроченного эффекта как по времени, так и по пространству

Так как предварительно проведённые тесты подтвердили наличие пространственных зависимостей между добитом жидкости скважин, то для количественного выражения таких зависимостей были построены модели пространственного лага.

Для учета пространственного влияния использовали модели в следующей спецификации [69]:

1) модели пространственного лага (SAR):

$$Y = \rho WY + X\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2), \quad (3.3)$$

где  $Y$  – матрица добычи нефти по скважинам и годам;

$W$  – весовая пространственная матрица;

$\rho$  – пространственный коэффициент;

$X$  – матрица объясняющих переменных;

$\beta$  – вектор оцениваемых коэффициентов при объясняющих переменных;

$\varepsilon$  – случайная компонента.

2) пространственная модель Дарбина (SDM) [97]:

$$Y = \rho WY + X\beta + WX\beta_n + \varepsilon, \quad (3.4)$$

где  $Y$  – матрица добычи нефти по скважинам и годам;  
 $W$  – весовая пространственная матрица;  
 $\rho$  – пространственный коэффициент  
 $X$  – матрица объясняющих переменных  
 $\beta$  – вектор оцениваемых коэффициентов при объясняющих переменных;  
 $\varepsilon$  – случайная компонента.

3) модель пространственной автокорреляции (SAC):

$$Y = \rho WY + X\beta + u, u = \lambda Wu + \varepsilon, \quad (3.5)$$

где  $Y$  – матрица добычи нефти по скважинам и годам;  
 $W$  – весовая пространственная матрица;  
 $\rho$  – пространственный коэффициент;  
 $X$  – матрица объясняющих переменных;  
 $\beta$  – вектор оцениваемых коэффициентов при объясняющих переменных;  
 $\varepsilon$  – случайная компонента,  $\lambda$  – пространственный коэффициент ошибки;  
 $u$  – случайная компонента.

4) модель с пространственным взаимодействием в ошибках (SEM):

$$Y = X\beta + \varepsilon, \varepsilon = \lambda W\varepsilon + u, u \sim N(0, \sigma^2), \quad (3.6)$$

где  $Y$  – матрица добычи нефти по скважинам и годам  
 $W$  – весовая пространственная матрица;  
 $\rho$  – пространственный коэффициент;  
 $X$  – матрица объясняющих переменных;  
 $\beta$  – вектор оцениваемых коэффициентов при объясняющих переменных;  
 $\varepsilon$  – случайная компонента;  
 $\lambda$  – пространственный коэффициент ошибки;  
 $u$  – случайная компонента.

Была построена модель SDM на конкретную дату 31 мая 1999 г. по 82 скважинам, где в качестве матрицы пространственных весов использовалась матрица соседства скважин в радиусе до 750 м. В качестве экзогенных переменных рассматривались показатель пластового давления (Formation Pressure), показатель забойного давления; в качестве целевой (зависимой) переменной рассматривался дебит жидкости добывающих скважин. Оценку коэффициентов проводили методом максимального правдоподобия. Результаты оценки модели сведены в Таблицу 3.8.

Таблица 3.8 – Результаты оценки модели спецификации SDM

	Оценка	Стд. ошибка	z-статистика	Pr(> z )
Константа	413,21	292,84	1,41	0,15
Внутрипластовое давление	-0,69	0,97	-0,71	0,47
Пространственно взвешенное Внутрипластовое давление	-0,34	1,57	-0,21	0,82
Показатели оценки модели	Пространственный коэффициент $\rho=0,243$	LR статистика 2,826 $p=0,092749$	z-value: 1,7643, $p=0,07768$	Статистика Вальда: 3,1128 $p=0,07768$
	Асимптотическая стандартная ошибка: 0,13813	Логарифм функции правдоподобия -553,7883	Дисперсия остатков 42339, ( $\sigma=205,76$ )	
Показатели качества подгонки модели под реальные данные	AIC=1117,6, (AIC для lm: 1118,4)	Псевдо- $R^2$ : 0,047039	LM для остаточной автокорреляции Значение статистики: 0,50596, $p=0,47689$	

Была также построена модель спецификации SAR на конкретную дату 31 мая 1999 г. по 82 скважинам, где в качестве матрицы пространственных весов использовалась матрица соседства скважин в радиусе до 750 м. В качестве экзогенной переменной рассматривался показатель пластового давления (Formation Pressure), в качестве целевой (зависимой) переменной рассматривался дебит жидкости добывающих скважин. Оценку коэффициентов проводили методом максимального правдоподобия. Результаты оценки модели сведены в Таблицу 3.9.

Построенные модели в спецификации SAR и SDM являются статистически значимыми в целом, однако их объясняющие способности у модели всего около 4% (псевдокоэффициент  $R^2$  не превышает 0,05). Статистические тесты выполнены при уровне значимости только  $p=0,01$ .

Таблица 3.9 – Результаты оценки модели SDM

	Оценка	Стд. ошибка	z-статистика	Pr(> z )
Константа	362,93	178,71	2,03	0,042
Внутрипластовое давление	-0,77	0,89	-0,86	0,387
Показатели оценки модели	Пространственный коэффициент $p= 0,246$	LR статистика 2,913 $p= 0,087891$	z-value: 1,7915, $p= 0,0732$	Статистика Вальда: 3,2094 $p= 0,07321$
	Асимптотическая стандартная ошибка: 0,13754	Логарифм функции правдоподобия -553,8117	Дисперсия остатков 42328, ( $\sigma$ : 205,79)	
Показатели качества подгонки модели под реальные данные	AIC=1115,6, (AIC для lm: 1116,5)	Псевдо- $R^2$ : 0,046493	LM для остаточной автокорреляции Значение статистики: 0,14353, $p= 0,7048$	

Представленные пространственные модели строятся на конкретно выбранную дату, при этом не учитываются предыдущая история наблюдений. Таким образом, выявляется связь только на выбранную дату, при этом остаётся неизвестным какие были связи и тенденции до этого момента. Из-за высокой волатильности в данных по добыче жидкости и давлению появляется вероятность выявления ложной корреляционной пространственной связи.

Таблица 3.10 – Результаты оценки модели SDM

	Оценка	Стд. ошибка	z-статистика	Pr(> z )
Забойное давление	-1,53	0,01	-138,12	<0,00001
Внутрипластовое давление	0,42	0,02	20,37	<0,00001
Пространственный коэффициент $\rho$	-0,15	0,01	-12,24	<0,00001
Коэффициент пространственной ошибки $\lambda$	0,16	0,01	15,57	<0,00001
Показатели качества подгонки модели под реальные данные	Псевдо- $R^2$ : 0,8889			

Так как тесты на спецификацию показали, что целесообразно оценивать модель по панельным данным с включением индивидуальных панельных моделей, то пространственная модель была перестроена в спецификации:

пространственно-панельная модель с пространственной и лаговой компонентой с фиксированными эффектами (SAC - модель с пространственным коэффициентом и пространственной ошибкой). Метод оценки коэффициентов моделей – метод максимального правдоподобия. Начало выборки: 1 января 1997; конец выборки: 31 мая 1999, наблюдений (кросс-секций), скважин – 82, весовая пространственная матрица строилась как матрица соседства скважин в радиусе до 750 м. Результаты оценки представлены в Таблице 3.11. В Таблице 3.12 представлены результаты оценки панельной пространственной модели SAC с включением двух экзогенных переменных: внутрипластового давления и забойного давления. Данная модель оказалась лучшей моделью по спецификации,  $R^2=0,8889$ , Так как обе модели имеют спецификацию с включением фиксированных эффектов как по периодам, так и по объектам (скважинам), то все эффекты были рассчитаны и приводятся в Приложении 2 (по модели SAR Таблица 1 и Таблица 2 соответственно по эффектам по объектам и по периодам). Для модели SAC приводятся только эффекты по скважинам (82 эффекта) с указанием их стандартной ошибки,  $t$ -статистики и уровня значимости  $p$ , (Приложение С, Таблица 1). Для наглядности представления фиксированных эффектов для модели SAC с включением двух экзогенных переменных внутрипластового давления и забойного давления, были построены столбчатые гистограммы соответствующих значений эффектов по каждой из скважин (Рисунок 3.13).

Таблица 3.11 – Результаты оценки модели SAR

	Оценка	Стд. ошибка	$z$ -статистика	$Pr(> z )$
Внутрипластовое давление	0,333	0,022	15,286	<0,00001
Коэффициент пространственной автокорреляции $\rho$	-0,406	0,013	-30,801	<0,00001
Коэффициент пространственной ошибки $\lambda$	0,412	0,011	39,492	<0,00001



Таблица 3.12 – Результаты оценки модели SAC с включением экзогенных переменных

	Оценка	Стд, ошибка	z-статистика	Pr(> z )
Забойное давление	-1,53	0,01	-138,12	<0,00001
Внутрипластовое давление	0,42	0,02	20,37	<0,00001
Пространственный коэффициент $\rho$	-0,15	0,01	-12,24	<0,00001
Коэффициент пространственной ошибки $\lambda$	0,16	0,01	15,57	<0,00001
Показатели качества подгонки модели под реальные данные	Псевдо- $R^2$ : 0,8889			

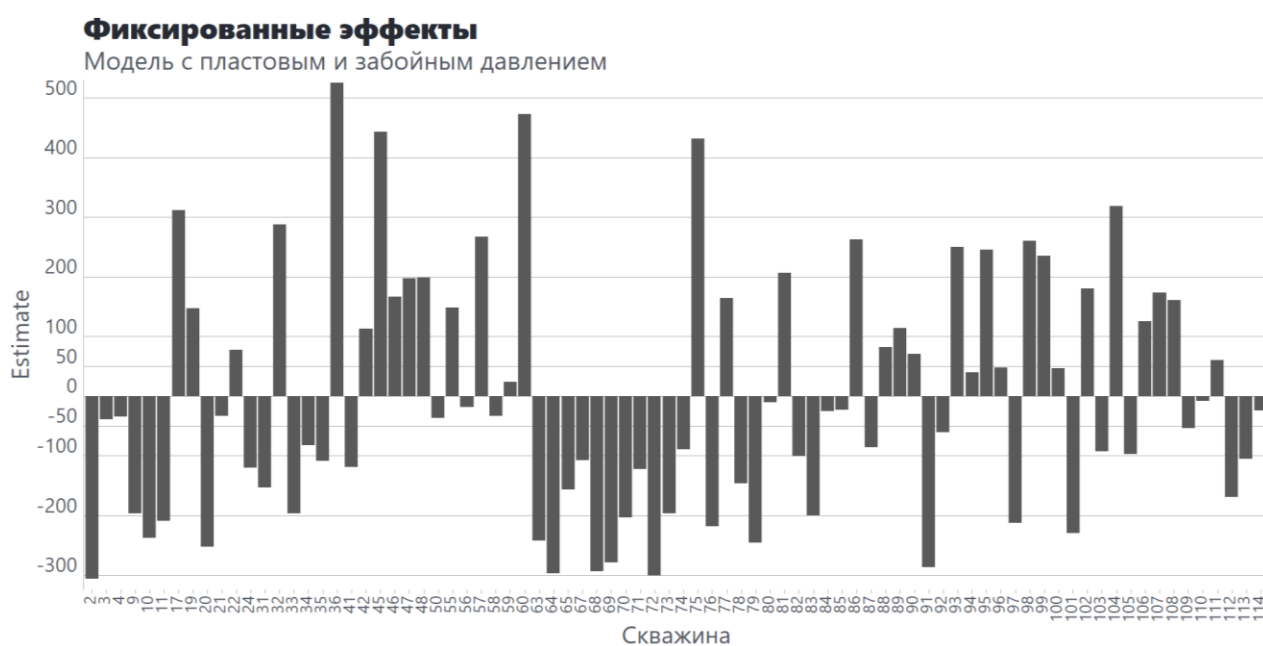


Рисунок 3.13 – Индивидуальные эффекты по скважинам для модели SAC с включением двух экзогенных переменных внутрипластового давления и забойного давления

Были также построены панельные модели следующих спецификаций, приведённые в Приложении D:

1. Модели обобщённой панели без индивидуальных эффектов (Приложение D, Таблица 1):

- Обобщённая панельная модель с пространственным лагом.
- Обобщённая панельная модель с пространственной ошибкой.
- Обобщённая панельная модель с пространственной ошибкой (метод ML).
- Обобщённая панельная модель с независимыми ошибками.
- Обобщённая панельная модель с пространственным лагом и с пространственной ошибкой.
- Обобщённая панельная модель с пространственным лагом и с пространственной ошибкой (метод ML).
- Обобщённая панельная модель с пространственным лагом и с независимыми ошибками.

2. Модели с индивидуальными фиксированными эффектами по кросс-секциям и периодам (Приложение D, Таблица 2):

- Панельная модель с пространственным лагом и фиксированными эффектами по периодам и кросс-секциям.
- Панельная модель с пространственной ошибкой и фиксированными эффектами по периодам и кросс-секциям.
- Панельная модель с пространственным лагом и с пространственной ошибкой и фиксированными эффектами по периодам и кросс-секциям.

3. Модели с индивидуальными фиксированными эффектами по кросс-секциям (Приложение D, Таблица 3):

- Панельная модель с пространственным лагом и фиксированными эффектами по кросс-секциям.
- Панельная модель с пространственной ошибкой и фиксированными эффектами по кросс-секциям.
- Панельная модель с пространственным лагом и с пространственной ошибкой и фиксированными эффектами по кросс-секциям.
- Модели с индивидуальными фиксированными эффектами по периодам (Приложение D, Таблица 4):

- Панельная модель с пространственным лагом и фиксированными эффектами по периодам.
- Панельная модель с пространственной ошибкой и фиксированными эффектами по периодам.
- Панельная модель с пространственным лагом и с пространственной ошибкой и фиксированными эффектами по периодам.

Среди всех построенных статистически значимых моделей наибольшее значение псевдокоэффициента детерминации  $R^2$  и наименьшие значения информационных критериев  $AIC$ ,  $BIC$  имеет модель спецификации: модель с фиксированными эффектами по периодам и кросс-секциям (скважинам) в спецификации  $SAC$  (с пространственной ошибкой  $\lambda$  и пространственным коэффициентом  $\rho$ ) с включением экзогенных переменных внутрислоевого давления и забойного давления.

Код построения моделей SAR и SAC в Приложении Е.

Модель полученной спецификации позволяет оценить взаимное влияние добычи скважин с учётом внешнего влияния внутрислоевого давления, забойного давления и пространственного и временного лага запаздывания с учётом индивидуального эффекта, присущего конкретной скважине и временному периоду.

### 3.4 Вероятностное прогнозирование прерывистой динамики добычи жидкости обусловленными нормализующими потоками

Для прогнозирования по историческим данным всего набора временных рядов месторождения апробирована была также модель DeepAR - алгоритм технологии рекуррентных и LSTM нейронных сетей [89]. Также, как и в LSTM сетях, внутреннее состояние сети изменяется по мере прохождения уже известной части, что позволяет переключаться между различными режимами поведения, чтобы выбрать лучшее описание её прогнозной части. Чтобы агрегировать информацию сеть должна сформировать достаточное большое количество

прогнозов на участке. Из вероятностных оценок в рамках рассматриваемой нами задачи как лучшая выбрана медиана. Во вход подавался весь набор временных рядов, регулированный с учетом типа процесса ( $DS(I(0)), DS(I(1))$ ).

Для обучения DeepAR сети, использовались части исторических данных по скважинам 2, 4, 11, 17, 22, 67, 71, 96, 24, 41, 81, 85, 109 и 110. Сеть делает прогноз на один день вперед по предшествующим значениям добычи жидкости и по предшествующим значениям разниц между пластовым и забойным давлениями. Чтобы усреднить результат, проводится 1000 запусков сети, и по полученному результату берётся медиана [10].

Так как производится прогноз разницы второго порядка от добычи жидкости, то для того, чтобы получить прогноз добычи необходимо знать значение исходной разницы первого порядка. Из-за того, что временной ряд добычи жидкости волатилен, что приводит к сильным скачкам разниц первого порядка, необходимо её усреднить. Для этой цели берутся три последние разницы первого порядка и рассчитывается среднее арифметическое из них. Зная прогнозную разницу первого порядка и последнее известное значение добычи, расчёт прогнозных значений проводится автоматически. Из-за особенностей поведения разниц второго порядка конечный прогноз в целом будет представлять продолжение текущей тенденции с очень редкими изгибами.

Сравнение с апробированными методиками BVAR модель DeepAR показала лучшее качество прогноза. Так, для прогноза на 30 дней DeepAR смог улучшить качество прогноза для 54 скважин из 82 (67%). В Приложении F (Таблицы F1-F2) представлены таблицы с оценкой ошибки MAPE и RMSE, соответственно, для моделей, прогнозирующих объем добычи нефти на 30 дней горизонта прогноза с введением следующих экзогенных переменных – пластовое давление, разница пластового и забойного давлений, пластовое давление на новом кусте. Также менялся период, на котором проводили тестирование свойства прогнозировать модели обученной DeepAR. с 1 апреля и с 1 мая.

Использование результатов синхронного анализа для регуляризации входной информации позволяет существенно повысить качество прогноза по

модели DeepAR и провести сравнительный анализ полученных результатов. Однако существенный недостаток данного алгоритма – не интерпретируемость результатов.

## Выводы по третьей главе

1. Полученные в результате параметризации на данных месторождения байесовские модели векторной авторегрессии показали свою состоятельность и применимость для прогнозирования: в среднем для каждой скважины из куста показатель объяснимости вариации дисперсии модели ( $R^2$ ) составляет от 0,88 до 0,98, что соответствует достаточному уровню прогнозных возможностей. Модель BVAR позволяет предсказать объем дебита жидкости на добывающей скважине с учетом влияния дебита жидкости других добывающих скважин, соседних по кусту с учетом имеющегося лага запаздывания (до 28 дня включительно), что повышает ценность данных моделей.

2. Проведённый комплекс статистических тестов для выбора спецификации модели и определения целесообразности перехода к моделям панельной структуры с учетом влияния пространственного лага позволил выявить: в модель целесообразно ввести панельные фиксированные эффекты (во всех тестах на спецификацию  $p < 0,05$ ), индивидуальные для каждой скважины месторождения, и учесть пространственную автокорреляцию между добычей соседних скважин (во всех тестах на спецификацию и во всех тестах на статистическую значимость индексов пространственной автокорреляции Морана и Гири  $p < 0,05$ ). Причём тесты показали обоснованность выбора в качестве радиуса оценки максимального воздействия на дебит скважины добывающих скважин в радиусе 1000 м.

3. Процедура селекции моделей пространственного лага на панельных данных позволила определиться с оптимальной спецификацией модели: модель с фиксированными эффектами по периодам и кросс-секциям (скважинам) в спецификации SAC (с пространственной ошибкой  $\lambda$  и пространственным коэффициентом  $\rho$ ) с включением экзогенных переменных внутрислоевого давления и забойного давления. Модель полученной спецификации позволяет оценить взаимное влияние добычи скважин с учётом внешнего влияния внутрислоевого давления, забойного давления и пространственного и

временного лага запаздывания с учётом индивидуального эффекта, присущего конкретной скважине и временному периоду.

4. В качестве альтернативы построенным моделям применялась модель, основанная на гибридизации инструментов анализа временных рядов и искусственных нейронных сетей – DeepAR с введением экзогенных переменных (внутрипластового давления и забойного давления). Несмотря на то, что модель показала сопоставимое качество прогнозирования с моделью BVAR на тестовой выборке, применение ее в практике ограничено из-за отсутствия возможности интерпретации результатов и отсутствия выявленного влияния окружения добывающей скважины.

## 4 ВАРИАНТНЫЕ ЭКСПЕРИМЕНТЫ ПОСТРОЕНИЯ ПРОГНОЗНЫХ МОДЕЛЕЙ VAR ПО ПРОГРАММНОМУ КОДУ

### 4.1 Методика формирования прогнозов добычи на скважинах месторождения

Оценки качества построенных моделей в целях прогнозирования проводились на основе метрик: корня из среднеквадратичная ошибки (RMSE, root mean square error) и средней ошибки (MAPE, mean absolute percentage error) [2]:

$$MAPE_{series} = \frac{1}{T} \cdot \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\% \quad (4.1)$$

$$RMSE_{series} = \sqrt{\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{T}}, \quad (4.2)$$

где  $y_t$  – фактическое значение в  $t$ -ый период времени;

$\hat{y}_t$  – прогнозное значение в  $t$ -ый период времени;

$T$  – временной интервал, на котором тестируется или определяется прогноз.

Для оценки доверительного интервала прогнозов использовали стандартный подход, где в качестве уровня доверия, который влияет на ширину доверительного интервала прогноза, рассматривали 95%.

Нижняя и верхняя границы доверительных интервалов определяются соответственно путем вычитания из прогнозной оценки модели  $x_T(\tau)$  и сложения с ней величины  $R(\tau)$  [17]:

$$R(\tau) = Z_k(\tau) \cdot \sigma_\varepsilon \cdot t_n^p, \quad (4.3)$$

где  $\sigma_\varepsilon = RMSE$  – среднеквадратичное отклонение ошибки прогноза;

$t_n^p$  –  $t$ -статистика распределения Стьюдента с вероятностью  $p$  и степенью свободы  $n$  (равная числу наблюдений в интервале горизонта прогнозирования  $\tau - 1$ );

$Z_k(\tau)$  – точечный прогноз на  $\tau$  шагов вперед.

Ширина доверительного интервала зависит от периода упреждения прогноза  $\tau$  и от средней квадратической ошибки  $\sigma_\varepsilon$ , характеризующей качество



модели на ретроспективном участке прогнозирования  $\tau=1$ . То есть доверительный интервал прогноза имеет вид  $(x_i(\tau)-R(\tau); x_i(\tau)+R(\tau))$ .

Для определения прогнозных качеств оценённых моделей и оценки возможности их использования в эксплуатации нефтяного месторождения использовали тройной подход проверки предиктивных свойств модели, включая этап слепой валидации.

На первом этапе для корректировки моделей оценивали RMSE и MAPE по всей ретроспективе ряда, на котором обучали модели байесовской векторной авторегрессии (BVAR). В результате средняя ретроспективная ошибка MAPE (найденная по ретроспективе всего временного ряда дебита жидкости) по всем скважинам, за исключением скважин №№ 63, 64, 73; 81, 87, 42, 97, 110, составляет 3,84%, максимальная ошибка MAPE не превышает для этих скважин 12%. Для скважин №№ 73, 64, 110 и 87 ошибка MAPE составила 14-18% по всей истории ряда, для скважин №№ 63, 81, 42 и 97 ошибка MAPE составила 36-51%. «Плохие» скважины территориально локализованы (Рисунок 4.1), имеют значительные периоды остановки.

### Карта месторождения

С учетом режима работы скважин

а Оба режима    а Только добыча    а Только закачка

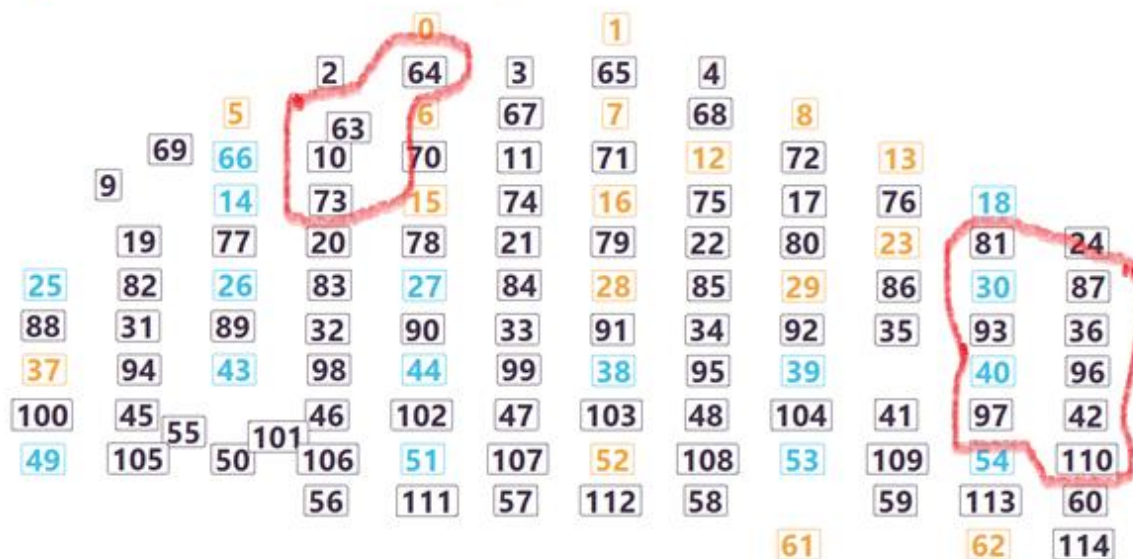


Рисунок 4. 1 – Территориальная локализация скважин модели, для которых модель дает прогнозную ошибку более 10%

На втором этапе использовали валидационные выборки в период 30 дней, но с различным началом их отсчёта: с начала месяца и с середины месяца. На данном этапе проверяли свойства робастности полученных оценок параметров моделей (в данном случае модели BVAR). В случае если при смещении периода начала валидационного тестового набора, прогнозная ошибка практически не меняется, а отклонение ошибок не является статистически значимым (согласно критерию Стьюдента [26] при  $p > 0,2$ ), то считаем, что полученные прогнозные оценки являются устойчивыми и прогнозные свойства моделей не зависят от того какой интервал времени для прогноза рассматривается (имеется в виду не по горизонту прогнозирования, а по локации в динамике временного ряда).

На третьем этапе использовали данные месторождения №2, которые никак не чувствовали в обучении модели. Но рассматривались из них только 35 скважин, объединённых в 8 кустов. Выбор только такого набора скважин был обусловлен тем, что по этим данным у нефтяной компании, которая обслуживает их имелись фактические и прогнозные данные, полученные по другим моделям (модель CRM (готовое программное решение компании) и модель экстремального градиентного бустинга [23] и модели, основанные на применении искусственных нейронных сетей).

#### 4.2 Подбор спецификации прогнозных моделей добычи на скважинах месторождения

Следует отметить, что одним из основных этапов любого прогнозного моделирования является подбор оптимальной спецификации модели, то есть такой модели, которая дает наименьшую прогнозную ошибку на периоде упреждения прогноза. При этом в исследовании подбирались экзогенные факторы влияния, длина лага запаздывания, учитываемая для показателей соседних по кусту скважин, период обучения модели. В исследовании рассматривались 16 следующих спецификаций моделей для месторождения с ежедневными данными

(при этом делалась попытка поиска оптимального начала обучающей выборки за счёт его смещения и смещения начала тестируемой выборки):

1) Целевая переменная – ежедневная добыча жидкости на скважине:

- Независимые переменные: добыча жидкости на других скважинах на кусте.

- Экзогенные переменные: отсутствуют.

- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).

- Период обучающей выборки: с 1 января 1990 по 1/15 июня 1999.

- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.

- Горизонт прогноза: 30 дней.

- Количество моделей: 24 для каждого из кустов

2) Смещения начала обучающей выборки на 1 июня 1997:

- Целевая переменная: Ежедневная добыча жидкости на скважине.

- Независимые переменные: добыча жидкости на других скважинах на кусте.

- Экзогенные переменные: отсутствуют.

- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).

- Период обучающей выборки: с 1 июня 1997 по 1/15 июня 1999.

- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.

- Горизонт прогноза: 30 дней.

- Количество моделей: 24 для каждого из кустов.

3) Смещения начала обучающей выборки на 1 января 1997

- Целевая переменная: Ежедневная добыча жидкости на скважине.

- Независимые переменные: добыча нефти на других скважинах на кусте.

- Экзогенные переменные: отсутствуют.

- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).

- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.

- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.

- Горизонт прогноза: 30 дней.

- Количество моделей: 24 для каждого из кустов.

4) Целевая переменная: Ежедневная добыча жидкости на скважине:

- Независимые переменные: добыча жидкости на других скважинах на кусте.

- Экзогенные переменные: забойное давление для каждой скважины куста (текущее).

- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).

- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.

- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.

- Горизонт прогноза: 30 дней.

- Количество моделей: 24 для каждого из кустов.

5) Целевая переменная: Ежедневная добыча жидкости на скважине:

- Независимые переменные: добыча жидкости на других скважинах на кусте.

- Экзогенные переменные: пластовое давление для каждой скважины куста (текущее).

- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).

- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.

- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.

- Горизонт прогноза: 30 дней.

- Количество моделей: 24 для каждого из кустов.

6) Добавление разницы давлений (столбец: разница давлений):

- Целевая переменная: Ежедневная добыча жидкости на скважине.

- Независимые переменные: добыча нефти на других скважинах на кусте.

- Экзогенные переменные: разница давлений на каждой скважине куста.

- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).

- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.

- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.

- Горизонт прогноза: 30 дней.

- Количество моделей: 24 для каждого из кустов.

7) Изменение формата формирования кустов:

- Выбран радиус куста в первое ближайшее окружение каждой скважины.  
Целевая переменная: Ежедневная добыча нефти на скважине.
- Независимые переменные: добыча нефти ближайшего окружения главной скважины (целевая переменная добычи).
- Экзогенные переменные: пластовое давление главной (целевой) скважины.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 30 дней.
- Количество моделей: 82 отдельно для каждой скважины.

8) Отсечение начала обучающей выборки для каждой скважины согласно ДСММ (дискретным скрытым марковским моделям):

- Индивидуальное отсечение начала обучающей выборки (Новые отсечения начала обучающей выборки (максимум 1 мая)).
- Целевая переменная: Ежедневная добыча нефти на скважине.
- Независимые переменные: добыча нефти ближайшего окружения главной скважины (целевая переменная добычи).
- Экзогенные переменные: разница давлений (внутрипластового и забойного) для каждой скважины куста.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: у каждой скважины своя дата начала обучающей выборки (примерно с 1 января 1990 по 1 мая 1999) и до 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 30 дней.
- Количество моделей: 82 отдельно для каждой скважины.

9) Индивидуальное отсечение начала обучающей выборки (Новые отсечения начала обучающей выборки (максимум 1 апреля)):

- Целевая переменная: ежедневная добыча нефти на скважине.

- Независимые переменные: добыча нефти ближайшего окружения главной скважины (целевая переменная добычи).
- Экзогенные переменные: разница давлений (внутрипластового и забойного) для каждой скважины куста.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: у каждой скважины своя дата начала обучающей выборки (примерно с 1 января 1990 по 1 апреля 1999) и до 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 30 дней.
- Количество моделей: 82 отдельно для каждой скважины.

10) Ежедневная добыча жидкости на скважине:

- Независимые переменные: другие скважины на кусте.
- Экзогенные переменные: пластовое давлений на каждой скважине куста.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 30 дней.
- Количество моделей: 24.

11) Пересчёт с обновлёнными данными на пластовом давлении, проверка прогноза на 30, 60 и 90 днях:

- Целевая переменная: Ежедневная добыча жидкости на скважине.
- Независимые переменные: дебит жидкости других скважин в кусте.
- Экзогенные переменные: пластовое давлений на каждой скважине куста.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 30, 60 и 90 дней.
- Количество моделей: 24.

12) Пересчёт с обновлёнными данными на разнице давлений, проверка прогноза на 30, 60 и 90 днях:

- Целевая переменная: Ежедневная добыча жидкости на скважине.
- Независимые переменные: дебит жидкости других скважин в кусте 49.
- Экзогенные переменные: пластовое давлений на каждой скважине куста.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 30, 60 и 90 дней.
- Количество моделей: 24.

13) Добавление нагнетательных скважин в модель:

- Целевая переменная: ежедневная добыча жидкости на скважине.
- Независимые переменные: дебит жидкости других скважин в кусте.
- Экзогенные переменные: пластовое давление на каждой скважине куста и суммарная закачка нагнетательных скважин.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 90 дней.
- Количество моделей: 24.

14) Добавление нагнетательных скважин (столбец: Модели на новых данных с разницей давлений (разница давлений)):

- Целевая переменная: ежедневная добыча жидкости на скважине.
- Независимые переменные: дебит жидкости других скважин в кусте.
- Экзогенные переменные: разница давлений на каждой скважине куста и суммарная закачка нагнетательных скважин.
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 90 дней.

- Количество моделей: 24.

15) Проверка учёта 28 лагов запаздывания в показателях соседних скважин:

- Целевая переменная: Ежедневная добыча жидкости на скважине
- Независимые переменные: дебит жидкости других скважин в кусте 50
- Экзогенные переменные: разница давлений на каждой скважине куста и суммарная закачка нагнетательных скважин.
- Лаговые переменные: 28 дней (для каждого дня своя лаговая переменная).
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999.
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999.
- Горизонт прогноза: 90 дней.
- Количество моделей: 24.

16) Проверка дифференцирования основной переменной (не были использованы дополнительные переменные):

- Целевая переменная: дифференцированная ежедневная добыча жидкости на скважине
- Независимые переменные: дебит жидкости других скважин в кусте
- Экзогенные переменные: отсутствует
- Лаговые переменные: 7 дней (для каждого дня своя лаговая переменная)
- Период обучающей выборки: с 1 января 1997 по 1/15 июня 1999
- Дата начала тестовой выборки: 1 (Начало) или 15 (Середина) июня 1999
- Горизонт прогноза: 90 дней
- Количество моделей: 24

*Примечание:* в моделях с нагнетательными скважинами учитывалась суммарная закачка жидкости всех нагнетательных скважин, относящихся к кусту в виде одной экзогенной переменной. Одна нагнетательная скважина может относиться к разным кустам. На каждый куст приходится примерно 2-3 нагнетательные скважины (в некоторых случаях и больше).

По каждой из перечисленных моделей рассчитывались ошибки MAPE и RMSE на соответствующий горизонт прогнозирования. По серии численных экспериментов сделаны следующие выводы: модели BVAR, построенные для



прогнозирования объема добычи нефти с учетом влияния соседних по кусту скважин, давали приемлемую ошибку прогноза (при горизонте прогнозирования в 30 дней) только для 60% скважин месторождения № 1. Следует заметить, что для скважин 11, 91, 3, 89 33, 42, 56, 59, 64, 65, 69, 71, 75 и 76 изменение начала тестовой выборки (начало месяца или середина) существенно изменяли ошибку прогноза. Проведены варианты численные эксперименты модели BVAR с целевой переменной – дебитом жидкости на скважинах с горизонтом прогнозирования 30, 60, 90 дней.

Итак, модели BVAR, построенные для прогнозирования объема добычи жидкости с учетом влияния соседних по кусту скважин, в различных спецификациях давали приемлемую ошибку прогноза MAPE (до 10%) (при горизонте прогнозирования в 90 дней) почти для всех скважин месторождения № 1. Ошибки MAPE RMSE для всех спецификаций моделей с горизонтом прогноза 30-60-90 дней представлены в приложении F (табл. F3-F5). Исключение составили 8 следующих скважин: 22, 42, 75, 79, 89, 91, 97, 98. Следует отметить, что динамика дебита жидкости на этих скважинах была нестандартной: либо имелись большие перерывы в работе скважины (остановки более года, как у скважины 97), либо резкие перепады в добыче (как у скважины 79, для которой добыча первый период была на уровне 400-500 т, в день, затем остановка, и восстановление добычи до 50-60), либо динамика дебита относилась к резко взрывному процессу DS(I(2)), как например, скважина 42, либо в динамике ряда имелось множество скачков в виде остановок, как для скважин 22, 75, 91. Следует отметить, что скважины 97 и 42 относятся к одному кусту - №23.

В Приложении F в Таблицах F6 представлены результаты расчета ошибки MAPE по построенной модели BVAR по 461 кусту месторождения с ежемесячными данными при горизонте прогнозирования дебита жидкости на 5 месяцев в следующих спецификациях:

1. Модель, в которой прогноз строился протяжкой последнего значения ряда дебита жидкости скважины.

2. Модель, в которой учитывается влияние только добывающих скважин (с лагом запаздывания 3 месяца).

3. Модель, в которой учитывается влияние добывающих скважин (с лагом запаздывания 3 месяца) и в качестве экзогенных переменных – разница давлений внутрипластового и забойного.

4. Модель, в которой учитывается влияние добывающих скважин (с лагом запаздывания 3 месяца) и в качестве экзогенных переменных – разница давлений внутрипластового и забойного, суммарный объем закачки ближайших (в радиусе 500 м) нагнетательных скважин.

Лучшей оказалась спецификация под номером 3, с разбросом ошибки MAPE от 3,74% до 14,21%.

В результате было отобрано три спецификации модели BVAR для тестирования прогнозных свойств модели в режиме реальной эксплуатации месторождения:

1) target\_m1 (bvar\_nonph) – модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте;

2) target\_m2 (bvar\_ph) – модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте и пластового давления на каждой скважине куста;

3) target\_m3 (bvar\_ph\_inj) – модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте, пластового давления на каждой скважине куста и суммарной закачки нагнетательных скважин.

На Рисунке 4.2 приведено изменение ежедневной ошибки при горизонте прогнозирования 90 дней суммарно на всех отобранных скважинах для моделей target\_m1, target\_m2 и target\_m3. Как видно модель без введения экзогенных переменных показала лучший результат (наименьшую ошибку), возможно, это связано с тем, что экзогенные переменные также прогнозируются, и ошибка при увеличении горизонта прогнозирования накапливается и увеличивается.

На Рисунке 4.3 дана MAPE, рассчитанная при горизонте прогноза 90 дней, для дебита жидкости каждой скважины. Как видно 20% ошибку для прогноза дебита скважин реже превышала модель с введением в качестве экзогенных

переменных пластового давления на каждой скважине куста и суммарной закачки нагнетательных скважин. На Рисунке 4.4 представлены построенные интерактивные карты ошибки прогноза для дебита каждой скважины при горизонте прогнозирования 30 дней в сравнении с прогнозами, построенными компанией: *trans* – прогноз на основе уравнений трансмиссий, *ela* – прогноз на основе искусственной нейронной сети, *xgb* – прогноз на основе алгоритма экстремального градиентного бустинга, *svm* – прогноз на основе модели CRM. Как видно модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте даёт сопоставимое качество с моделью CRM.

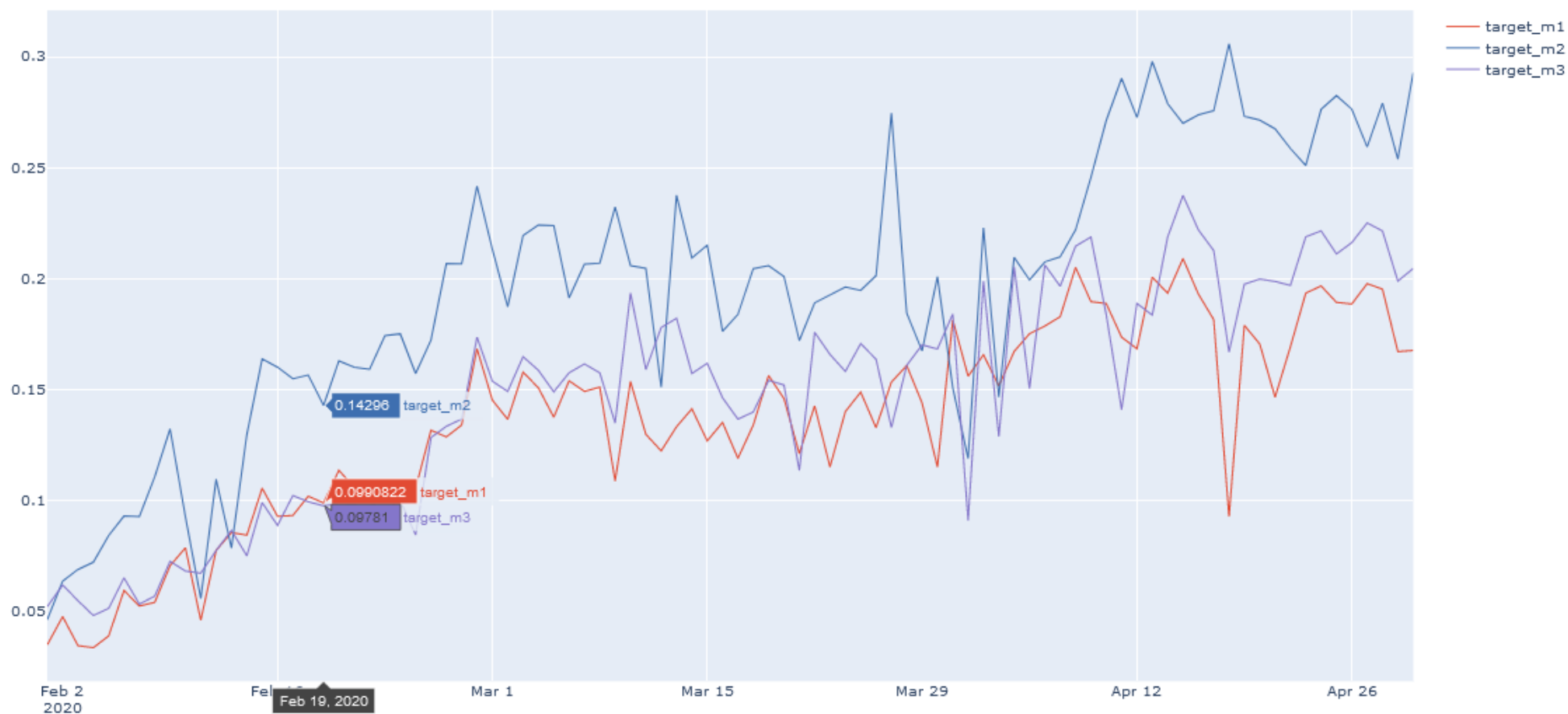


Рисунок 4.2 – Динамика изменения ежедневной ошибки при горизонте прогнозирования 90 дней суммарно на всех отобранных скважинах для моделей target\_m1 (модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте), target\_m2 (модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте и пластового давления на каждой скважине куста) и target\_m3 (модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте, пластового давления на каждой скважине куста и суммарной закачки нагнетательных скважин)

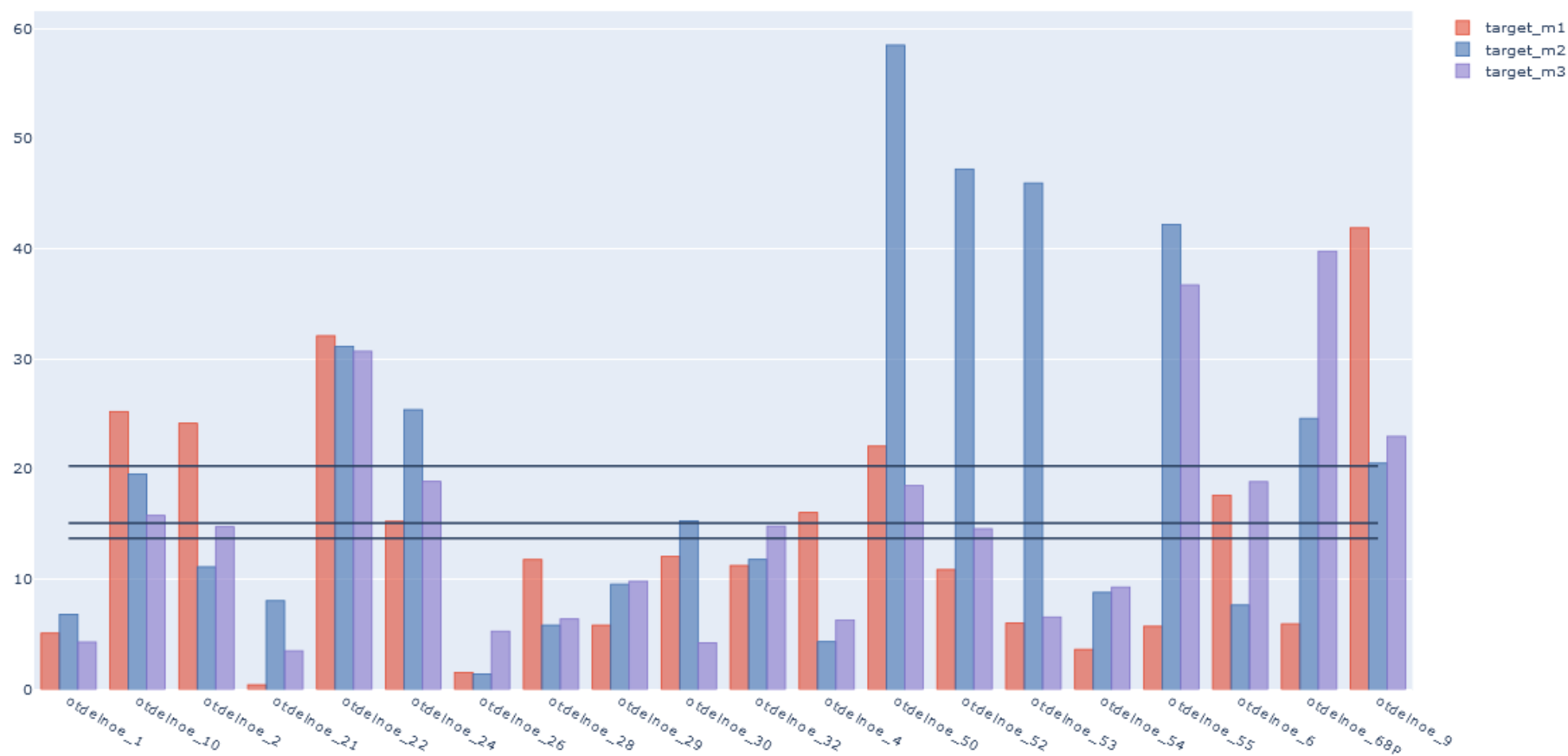


Рисунок 4.3– MAPE, рассчитанная при горизонте прогноза 90 дней, для дебита жидкости каждой скважины для моделей target\_m1 (модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте), target\_m2 (модель с учетом влияния лагового до 28 дня дебита соседних скважин на кусте и пластового давления на каждой скважине куста) и target\_m3 (модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте, пластового давления на каждой скважине куста и суммарной закачки нагнетательных скважин)

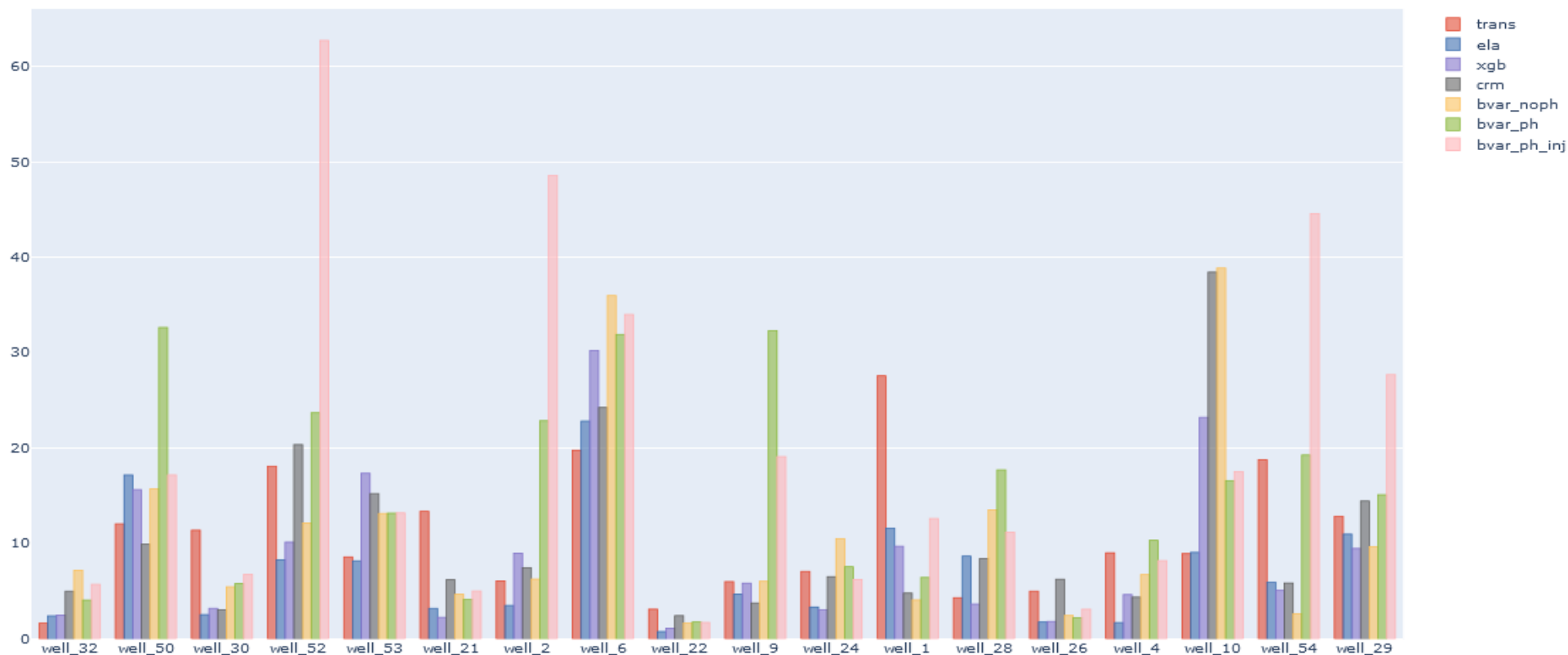


Рисунок 4.4 – Ошибки прогноза для дебита каждой скважины при горизонте прогнозирования 30 дней: trans – прогноз на основе уравнений трансмиссий, ela – прогноз на основе искусственной нейронной сети, xgb – прогноз на основе алгоритма экстремального градиентного бустинга, crm – прогноз на основе модели CRM, bvar\_noph – модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте; bvar\_ph – модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте и пластового давления на каждой скважине куста; bvar\_ph\_inj – модель с учётом влияния лагового до 28 дня дебита соседних скважин на кусте, пластового давления на каждой скважине куста и суммарной закачки нагнетательных скважин

#### 4.3 Программная реализации методики получения прогнозов добычи на скважинах месторождения с учётом их взаимного влияния

Для реализации использования на практике разработанной методики построения прогнозных моделей с помощью обученной модели байесовской векторной авторегрессии (BVAR) было разработано специальное программное обеспечение на языках R и Python. Программный код для самой модели BVAR реализован был в R [12], а подготовка обучающих и тестовых наборов производится в Python [48], далее эти наборы передаются в R, где строятся модели и получается окончательный прогноз. Кроме того, соответствующий код компилировался в виде DLL-библиотеки для лёгкой бесшовной интеграции [15] с информационной системой компании для анализа и прогнозирования добычи нефти.

Следует отметить, что библиотека для построения модели BVAR `bvarr` находится в GitHub, следовательно необходимо будет установить дополнительную библиотеку `devtools`, позволяющую установить пакет из гитхаба. (то есть: `devtools::install_github("bdemeshev/bvarr")`).

Запуск Python производился с помощью среды разработки `rcharm`, запуск также возможен и в других средах, совместимых с python, или через командную строку. Для работы необходимы библиотеки: `Pandas`, которая запускается по команде `Pip install pandas`.

Модели BVAR не реализованы в Python, поэтому для построения моделей и получения прогноза по скважинам используется язык R. Подготовка обучающих и тестовых наборов производится в Python, далее эти наборы передаются в R, где строятся модели и прогноз. Прогноз по каждому кусту сохраняется в отдельный файл формата `.csv`. Для работы необходимы библиотеки: `devtools`, `bvarr`, `stringi`. В отдельном файле прописывается стандартный путь к R. Требования к данным: при наличии пропусков в «сырых» данных важно полностью их восполнить (использовать можно методы восполнения нулем, средним по ряду и методом протяжки последнего значения по времени до пропуска). Для обучающей

выборки необходимо наличие истории наблюдения добычи нефти на месторождении не менее года. В случае если на кусте присутствуют скважины, у которых обучающая выборка меньше остальных скважин, то данные по неработающей истории нужно восполнить нулем. Для кода необходимы 4 набора данных: дата наблюдения, номер скважины, дебит/добыча жидкости/нефти, дополнительные переменные (забойное или пластовое давления). Данные по истории закачки: дата наблюдения, номер скважины, уровень закачки, 2 набора с координатами скважин (x,y) (координаты добывающих скважин, координаты закачивающих скважин). Возможные экзогенные переменные, которые были проверены в ходе данной работы с использованием предварительного моделирования (забойное давление, пластовое давление, уровень закачки на соседних нагнетательных скважинах).

Для удобства работы программного модуля прогнозирования целесообразна следующая организация хранения и доступа информации [27]: папка, хранящая начальные данные `Data`; папка, хранящая результат прогнозирования `intermediate_data`; Python-скрипт для создания обучающей выборки и взаимодействия с R-скриптом; R-скрипт для построения модели и прогноза (Таблица 4.1). За расположение папки, хранящей начальные данные `Data` в скрипте, отвечает строка: `begin_data = os.getcwd() + "/Data/"`.

На вход для построения прогнозов подаются данные формата `csv`. Для изменения расположения данных необходимо указать новый путь в этих строках.

Для модели `BVAR` с дополнительными переменными выполняется:

- 1) настройка кустов: `m_length` – максимальная длина между скважинами-соседями, `max_length` – максимальная длина между скважинами внутри куста (не только соседями);
- 2) настройки наборов и моделей: `sd_train` – начало обучающей выборки (`sd_train = datetime,date ()`), `sd_test` – начало тестовой выборки (`sd_test = datetime,date()`), `fcast_step` – горизонт прогноза, `Pressure` – название дополнительной переменной в файле `debit.csv`.



Таблица 4.1 – Описание файлов с входными данными и переменных (основных и дополнительных) для модели BVAR

Файл	Описание файла	Название переменной	Описание переменной	Наличие в моделях дополнительных переменных	
				БЕЗ	С
debit.csv	Данные по добывающим скважинам	Well	Номер скважины	✓ <input type="checkbox"/>	✓ <input type="checkbox"/>
		Date	Дата наблюдения	✓ <input type="checkbox"/>	✓ <input type="checkbox"/>
		Дополнительные переменные, Например: Pressure_RP R или Pressure_N D	Данные по давлению (забойное, пластовое итд), В скрипте с дополнительными переменными необходимо указать корректное название данной переменной,	✗	✓ <input type="checkbox"/>
		Liquid	Дебит жидкости	✓ <input type="checkbox"/>	✓ <input type="checkbox"/>
coords_debit.csv	Координаты добывающих скважин	NODE_X	Координата X	✓ <input type="checkbox"/>	✓ <input type="checkbox"/>
		NODE_Y	Координата Y	✓ <input type="checkbox"/>	✓ <input type="checkbox"/>
		UWI	Номер скважины	✓ <input type="checkbox"/>	✓ <input type="checkbox"/>
coords_inj.csv	Координаты нагнетательных скважин	NODE_X	Координата X	✗	✓ <input type="checkbox"/>
		NODE_Y	Координата Y	✗	✓ <input type="checkbox"/>
		UWI	Номер скважины	✗	✓ <input type="checkbox"/>
Inject.csv	Данные по закачке нагнетательных скважин	Well	Номер скважины	✗	✓ <input type="checkbox"/>
		Date	Дата наблюдения	✗	✓ <input type="checkbox"/>
		Injectivity	Закачка жидкости	✗	✓ <input type="checkbox"/>

Для модели BVAR без дополнительных переменных выполняется:

- 1) настройка кустов: m\_length – максимальная длина между скважинами-соседями, max\_length – максимальная длина между скважинами внутри куста (не только соседями);
- 2) настройки наборов и моделей: fcast\_step – горизонт прогноза, sd data\_begin – начало обучающей выборки, data\_end – конец обучающей выборки.

Папка, хранящая результат: `intermediate_data`, то есть результат прогнозирования сохраняется в папку `intermediate_data`, и туда записываются csv файлы с прогнозами по каждому кусту. При этом каждый файл содержит следующие переменные: `variable` – номер скважины, `h` – шаг прогноза (1 – первый день, 2 – второй день, и т.д.), `mean` – основной прогноз, дополнительные переменные (`sd` – среднеквадратическое отклонение, `data_end` – конец обучающей выборки, `median` – медианный прогноз, `lower_80`, `lower_95`, `upper_80`, `upper_95` – доверительные интервалы).

Также был организован цикл по каждой добывающей скважине  $i$ . Под неё формируется список кустов  $c\_1$ , куда эту скважину можно присоединить. Для этого организуется цикл по сформированным на данный момент кустам (переменная  $c$ ), а внутри него ещё один по каждой скважине  $j$ , входящий в список данного куста  $c$ . Если расстояние до скважины  $j$  меньше `m_length`, то делается ещё один цикл по скважинам текущего куста  $c$ , для проверки того, чтобы не было превышения между парой скважин расстояния `max_length`. Это нужно для того, чтобы избежать "вытянутых" через всё месторождение кустов. Если скважина  $i$  подходит по данному условию, то добавляется данный куст  $c$  в список  $c\_1$ . После проверки всех кустов проверяется список кустов  $c\_1$ . Если  $c\_1$  пустой, то создаётся новый куст  $c$  единственной скважиной  $i$ . Если список  $c\_1$  не пустой, то добавляется скважина в первый по списку куст  $c\_1[0]$ . После этого проводится проверка, есть ли кусты, состоящие только из одной скважины. Если такие кусты есть, то они автоматически присоединяются к первому по списку кусту, если у него есть хотя бы одна скважина в пределах расстояния `m_length`. Условие на вытянутость (с расстоянием `max_length`) при этом игнорируется.

Также в скрипте Python обучающая выборка формируется с добывающими скважинами (деление скважин по кустам, создание сводной таблицы (по признаку «добыча жидкости»), заполнение пропусков, приведение всех скважин к одной длине (к началу периода обучения (`train`)), сохранение обучающих выборок по каждому кусту в формате `csv`). Проводится формирование части обучающей выборки с учетом факторов давления (деление скважин по кустам, создание

сводной таблицы (по признаку «Давление»), заполнение пропусков, сохранение данных в отдельный словарь).

Модель без дополнительных переменных не использует данные по Давлениям и Закачке. Часть кода, отвечающая за формирование кустов и основной обучающей выборки (**train**), идентична коду с дополнительными переменными. При этом следующие изменения в коде R: не загружаются дополнительные наборы (**train\_xreg**, **test\_xreg**), модели и прогноз строятся без этих дополнительных наборов.

Несмотря на высокую точность прогнозирования предложенная методика имеет ряд ограничений:

- 1) невозможность ее применения на впервые вводимых в эксплуатацию скважинах и на скважинах после длительного простоя (более 3-х месяцев) за счет отсутствия возможности учета исторических данных;
- 2) сложность ее применения на практике для скважин, на которых недавно (менее чем 10 дней) проводился гидроразрыв пласта.

#### 4.4 Оценка экономической эффективности моделей прогнозирования с учетом взаимной связности

При проведении любой научно-исследовательской работы (НИР) или научно-исследовательских и опытно-конструкторских работ (НИОКР) всегда предусматривается этап по определению экономической эффективности проекта. Существуют различные подходы к оценке экономической эффективности от простого затратного метода, учитывающего экономический анализ затрат и результатов [17], до усложнённых маржинальных методов, учитывающих предельную добавленную стоимость за счёт реализации проекта [46].

Проблеме оценки экономической эффективности добывающих скважин посвящено достаточно большое количество исследований, например, [42]. Один из самых важных вопросов, решаемых при этом, является вопрос: до какого

периода можно эксплуатировать скважину, так чтобы полные затраты на ее обслуживание, сопровождение и применяемые мероприятия по интенсификации добычи нефти, полностью компенсировались за счет ее реализации. Естественно, что ответ на этот вопрос можно получить, зная прогноз дебита на данной скважине, причём с учетом прогноза ее дебита на скважинах ее окружения. То есть прогноз позволяет определиться с необходимым «объемом добычи нефти с целью продления периода рентабельной разработки для максимально возможного отбора имеющихся запасов и достижения безубыточности работы нефтедобывающей компании» [12].

Следует отметить, что существующие методики оценки рентабельности добывающей скважины, как правило, рассматривают каждую в отдельности, тогда как предложенная методика прогнозирования дебита скважины позволяет рассматривать куст целиком с учетом взаимного влияния скважин друг на друга.

$$\sum_{i=1}^k \sum_{t=t_0}^T C_{it} = KC \quad (4.4)$$

В формуле (4.4) кумулятивная рентабельность  $KC$  по  $k$  скважинам куста за период упреждения прогноза  $T$  определяется как сумма накопленной ежедневной рентабельности по каждой из скважин с учетом их взаимного влияния.

Соответственно, рентабельность по  $i$ -ой скважине куста в  $t$ -ый день определяется как разница между ежедневной выручкой  $Rec_{it}$  и суммой прямых  $DC_{it}$  и косвенных затрат  $IC_{it}$  по обсаживанию и эксплуатации данной скважины:

$$C_{it} = Rec_{it} - (DC_{it} + IC_{it}) \quad (4.5)$$

Расчет выручки по  $i$ -ой скважине  $Rec_{it}$  зависит от прогнозируемой суточной добычи  $D_{it}$  по данной скважине и текущей цены на нефть в рассматриваемый период  $t PO_t$ . Следует отметить, что в цене на нефть учитываются налоги на добычу полезных ископаемых и расходы, связанные с транспортировкой и реализацией нефти.

$$Rec_{it} = D_{it} * PO_{it} \quad (4.6)$$

В формуле (4.6) используется прогноз дебита (тонн/сутки), полученный по моделям, разработанным в главе 3. Соответственно, можно рассчитать валовую выручку от добычи нефти по конкретному кусту при выборе различного горизонта прогнозирования – 30, 60 и 90 дней. Прямые  $DC_{it}$  и косвенные затраты  $IC_{it}$ , связанные с добычей нефти, в свою очередь сами зависят от прогноза дебита, так как данные расходы считаются на 1 тонну добытой нефти.

В данной методике есть ограничение: в формуле (4.6) используется текущая цена на нефть, хотя корректно использовать также её прогноз, который в свою очередь зависит от нестабильного прогноза рыночной цены на углеводороды [11].

#### Выводы по 4 главе

1. При проведении численных экспериментов для корректировки прогнозных моделей оценивали среднюю квадратичную ошибку и ошибку аппроксимации по всей ретроспективе ряда, на котором обучали модели байесовской векторной авторегрессии (BVAR). В результате средняя ретроспективная ошибка аппроксимации по всем скважинам, за исключением территориально локализованных скважин, на которых имеются значительные периоды остановки работы, составляет 3,84%, максимальная ошибка – не превышает для этих скважин 12%.

2. Был проведён подбор оптимальной спецификации модели прогнозирования, исходя из минимума ошибки прогнозирования на тестовом наборе данных в 30, 60 и 90 дней в качестве периода упреждения прогноза. При этом подбирались экзогенные факторы влияния, длина лага запаздывания, учитываемая для показателей соседних по кусту скважин, период обучения модели. В результате было рассмотрено 16 спецификаций моделей для месторождения с 24 моделями в каждой. Минимальный разброс средней ошибки прогноза (от 3,74% до 14,21%) был получен при рассмотрении в качестве целевой переменной – ежедневной добыча жидкости на скважине; в качестве независимых переменных – добычи нефти на других скважинах на кусте, при отсутствии экзогенных переменных, учете лаговых переменных с 7 днями запаздывания (для

каждого дня своя лаговая переменная), периодом обучающей выборки с 1 января 1997 по 1/15 июня 1999, периода тестовой выборки с 1 (Начало) или 15 (Середина) июня 1999, с горизонтом прогноза в 30 дней.

3. Для реализации использования на практике разработанной методики построения прогнозных моделей с помощью обученной модели байесовской векторной авторегрессии было разработано специальное программное обеспечение на языках R и Python, где предобработка данных о добычи проводилась с использованием библиотек Python, а непосредственное моделирование с использованием библиотек в R. Соответствующий код компилировался в виде DLL-библиотеки для бесшовной интеграции с информационной системой компании для анализа и прогнозирования добычи нефти.

4. Выявлены ограничения использования методики прогнозирования: невозможность ее применения на впервые вводимых в эксплуатацию скважинах и на скважинах после длительного простоя (более 3-х месяцев), ограничение применения для скважин, на которых недавно (менее чем 10 дней) проводился гидроразрыв пласта.

5. Предложена методика оценки эффективности добычи в виде накопленной ежедневной рентабельности по каждой из скважин куста с учетом их взаимного влияния.

## ОСНОВНЫЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ

1. Анализ научных исследований проблем прогнозирования процессов разработки месторождений позволил идентифицировать предпочтительность широкого применения в указанной области методов статистического моделирования и машинного обучения как более упрощённых подходов в сравнении с физическими, предназначенными для описания внутренних механизмов объекта исследования. Основные преимущества указанных методов заключаются в разнообразии инструментов, в существенном сокращении размерности задач и объёма требуемых для реализации ресурсов, в содержательных алгоритмах последовательного перебора и возможности усложнения спецификации в целях повышения качества прогноза. Данный анализ может быть полезен всем исследователям, планирующим прогнозировать дебит скважин с учетом влияния нагнетательных и добывающих скважин в окружении.

2. На исторических данных двух месторождений апробирован алгоритм синхронного анализа временных рядов, который заключается в поэтапном исследовании закономерностей индивидуального и коллективного поведения измеряемых характеристик для получения вероятностных оценок разработки каждой скважин, куста и месторождения в целом. Анализ показал, что подавляющее большинство динамических показателей эксплуатации скважин месторождений относятся к стационарным случайным процессам в широком смысле даже при наличии структурных скачков. Незначительным является число скважин месторождений, динамические показатели которых не стационарны – это ряды со структурными изменениями. Для оценки взаимного влияния динамических показателей по добыче на скважинах месторождения был проведен коинтеграционный анализ, позволивший определить длину лага запаздывания между связными показателями (от 7 до 28 дней) и определиться с предварительной спецификацией моделей. Разработка метода формирования кустов, в том числе с учетом расстояния от нагнетательной скважины, позволил сформировать сценарии рассмотрения кустообразования для обучения в

дальнейшем модели связности пластовых систем на основе оценки взаимовлияния скважин друг на друга. Основанный на применении метода дискретных скрытых марковских моделей подход позволил выделить однородные участки в рядах данных и как следствие определиться с точкой разбиения истории ряда на обучающую и тестовую выборку, что позволило существенно улучшить прогнозные качества разрабатываемых моделей. Разработанный комплексный подход по предварительному анализу исторических данных можно рекомендовать использовать при построении прогнозов для получения их высокой достоверности. Апробирование разработанной методики синхронизированного анализа временных рядов продемонстрировали свою эффективность при разработке: вероятностной модели прогнозирования на прерывистых данных всего месторождения на основе рекуррентных нейронных сетей DeepAR; структуры и спецификации модели BVAR, а также прогнозирования взаимосвязанности пластовых систем на примере двух месторождений.

3. Для оценки взаимовлияния скважин и определения связанности пластовых систем разработаны модели байесовской векторной авторегрессии BVAR. Модель BVAR продемонстрировала высокое прогнозное качество на основе геологических и технологических факторов двух месторождений. Например, модели по 24-м кустам месторождения 1 обнаруживают статистическую значимость зависимости объема дебета жидкости от соседних добывающих скважин с лагом до 4 дней включительно, влияния разницы забойного и внутривластового давлений, суммарного дебита закачки по всем нагнетательным скважинам в кустах. Средняя ретроспективная ошибка аппроксимации составляет 3,84%, максимальная ошибка для 5% скважин – 36-51%. Главное преимущество BVAR в интерпретируемости результатов, что позволяет количественно оценить зоны с высокой остаточной нефтенасыщенностью для планирования стратегии уплотняющего бурения. Соответственно, использование данной модели может быть полезно при принятии решений о выборе скважин-кандидатов в кусте к консервации.



4. Вероятностная модель по технологии рекуррентных и LSTM нейронных сетей по историческим данным DeepAR с введением экзогенных переменных (внутрипластового давления и забойного давления) показала высокую точность получаемых прогнозов на период упреждения 30 дней. Сравнение результатов эксперимента модели DeepAR с методикой BVAR показало, что предложенный подход может улучшить качество прогноза по ретроспективе данных на 30 дней для 54 скважин из 82 (67%). Перспективность данного алгоритма заключается в том, что предсказание оценивается исходя из множества исторических данных всего месторождения без разбивки его на кусты. Требуемое качество прогноза достигается после регуляризации временного ряда с учетом результатов предобработки по методологии синхронного анализа временных рядов. Однако у подхода DeepAR есть недостатки: во-первых, применение модели на практике ограничено из-за отсутствия возможности интерпретации результатов и отсутствия выявленного влияния окружения добывающей скважины; во-вторых, высокая точность прогноза достигается на горизонте прогнозирования в 30 дней, а на горизонте в 90 дней точность уступает прогнозам, полученным по моделям BVAR. Соответственно, модель DeepAR может быть рекомендована для прогноза дебита нефти на скважине в краткосрочный период (до 30 дней включительно).

5. Разработанная модель пространственного лага на панельных данных в спецификации SAC (с пространственной ошибкой  $\lambda$  и пространственным коэффициентом  $\rho$ ) с включением экзогенных переменных, позволила оценить взаимное влияние добычи скважин с учётом внешнего влияния внутрипластового давления, забойного давления, пространственного и временного лага запаздывания с учётом индивидуального эффекта, присущего конкретной скважине и временному периоду. Подобная модель может быть полезна при принятии решений относительно бурения новой скважины в кусте и прогноза её дебита, так как учитывает пространственное распределение в нем скважин.

6. Вариантные численные эксперименты модели BVAR при варьировании начала тестовой выборки, при наличии или отсутствии экзогенных переменных, с изменением формата формирования куста из ближайшего окружения скважины,

отсечение начала обучающей выборки по дискретным скрытым марковским моделям при горизонте прогноза соответственно 30-60-90 дней показали приемлемую ошибку для 65% скважин первого месторождения. Лучшей оказалась спецификация модели без экзогенных переменных. Лучшая модель месторождения 2 с еженедельными данными по горизонту прогноза дебита жидкости на 3 месяца в спецификации, где учитывается влияние добывающих скважин (с лагом запаздывания 28 дней) с разницей давлений внутрислоевого и забойного давала ошибку аппроксимации от 3,74% до 14,21%.

7. Для реализации использования на практике разработанной методики построения прогнозных моделей с помощью обученной модели байесовской векторной авторегрессии было разработано специальное программное обеспечение на языках R и Python. Соответствующий код компилировался в виде DLL-библиотеки, которую удобно интегрировать в используемые нефтедобывающей компанией информационные системы.

8. Предложенная методика оценки эффективности добычи в виде накопленной ежедневной рентабельности по каждой из скважин куста с учётом их взаимного влияния на основе полученных прогнозов по разработанным моделям может использоваться при принятии решений относительно заморозки или вывода из эксплуатации активов месторождения.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Абидов, Д.Г. Метод материального баланса как первичный инструмент оценки показателей разработки участка месторождения при заводнении / Д.Г. Абидов, М.Р. Камартдинов // Известия Томского политехнического университета. – 2013. – Т. 322. № 1. – С.90-96.
2. Авдеенко, Т.В. Компьютерные методы анализа временных рядов и прогнозирования: учебное пособие / Т.В. Авдеенко. – Новосибирск: Изд-во Новосибирского гос. технического университета, 2008. – 270 с.
3. Бабешко, Л.О. Эконометрические исследования: инструменты и методы: монография / Л.О. Бабешко, И.В. Орлова, Н.В. Концевая // М. : Сер. Научный фонд, 2021. – 240 с.
4. Бабешко, Л.О. Инструментарий современного эконометрического моделирования: монография / Л.О. Бабешко, И.В. Орлова // М. : Сер. Научный фонд, 2020. – 336 с.
5. Бакшевская, В.А. Методы моделирования геофильтрационной неоднородности осадочных отложений / В.А. Бакшевская, С.П. Поздняков // Геоэкология. Инженерная геология, гидрогеология, геокриология. – 2012. – № 6. – С. 560-570.
6. Бахитов, Р.Р. К проблеме выбора оптимальных технических решений для разработки рифейских карбонатных трещиноватых коллекторов / Р.Р. Бахитов // Сб. ст.: Научные исследования: проблемы и перспективы. XXXIV Междунар. науч.-практ. конф. – 2021. – С. 29-36.
7. Бахитов, Р.Р. Предобработка входной информации на основе геофизических критериев и экспертных гипотез относительно процессов разработки месторождения / Р.Р. Бахитов // Сб. науч. тр. EUROPEAN SCIENTIFIC CONFERENCE (по матер. XLII междунар.науч.конф.). – М., 2021. – С. 12-22.
8. Бахитов, Р.Р. Применение алгоритмов машинного обучения в задачах прогноза коэффициента продуктивности скважин карбонатных месторождений / Р.Р.Бахитов // Нефтяное хозяйство. – 2019. – №5. – С. 52-56.

9. Бахитов, Р.Р. Методология создания концепции разработки Восточно-Мессояхского месторождения / Р. Р. Бахитов, Р.Р. Исмагилов, Б. В. Белозеров и др. // Нефтяное хозяйство. – 2013. – №12. – С. 54-57.
10. Бахитов, Р.Р. Опыт применения модели DEEPAR для прогноза дебита жидкости добывающей скважины / Бахитов Р.Р., Гиндуллин Р.В., Лакман И.А. // Сб. ст.: Актуальные проблемы нефтегазовой отрасли (по матер. трех науч.-практ. конф. журнала Нефтяное хозяйство). – М. – 2022. – С. 283-286.
11. Бахитов, Р.Р. Методика оценки экономической эффективности моделей прогнозирования добычи с учетом взаимного влияния добывающих скважин / Бахитов Р.Р. // Сб. науч. тр. : Цифровые технологии в бизнесе. (по матер. II междунар. науч.-практ. конф.). – Уфа, 2023. – 0,16 печ. л.
12. Бахитов, Р.Р. Применение концептуального геологического моделирования в ОАО «Газпром нефть» / А. В. Билинчук, Р. Р. Бахитов, А. Н. Ситников и др. // Нефтяное хозяйство. – 2013. – № 5. – С. 44-48.
13. Баюк, О.А. Практикум по анализу данных на языках Python и R: учебное пособие / О.А. Баюк, М.Р. Исаева, М.О. Самсонкин. – М.: Прометей, 2023. – 100 с.
14. Бикташев, Э.К. Определение нерентабельных скважин с помощью прогноза критического дебита / Э.К. Бикташев // Сб. ст.: Рыночная трансформация экономики России: проблемы, перспективы, пути развития. Междунар. науч.-практ. конф. – 2018. – С. 41-45
15. Вичугова, А.А. Методы и средства интеграции информационных систем в рамках единого информационного пространства / А.А. Вичугова, В.Н. Вичугов, С.Г. Цапко, Е.А. Дмитриева // Труды 12-й Междунар. науч.-практ. конф. – 2012. – № 5(6) – С. 61-64.
16. Гадильшина, В.Р. Численное решение обратной задачи неизотермической фильтрации в нефтяном пласте. / В.Р. Гадильшина, И.Т. Салимьянов // Вестник Технологического университета. – 2015. – Т. 18. № 1. – С. 323-325.
17. Галявиева, М.С. Математика. Статистические методы прогнозирования. Краткий конспект лекций / М.С. Галявиева. – Казань: Изд-во: Медицина, 2011. – 55 с.

18. Гафарова, Е.А. Практикум по анализу панельных данных. / Е.А. Гафарова, И.А. Лакман, Р.Х. Бахитова. – Уфа: Изд-во РИО БашГУ, 2017. – 112 с.
19. Данько, М.Ю. Применение метода динамического материального баланса и CRM-метода (capacitance-resistive models) к подсчету запасов ачимовских и баженовских *коллекторов* / М.Ю. Данько, Л.С. Бриллиант, А.С. Завьялов // Недропользование XXI век. – 2019. – № 4. – С. 76–85.
20. Дейк, Л.П. Практический инжиниринг резервуаров. / Л.П. Дейк // – М.: Институт компьютерных исследований. – М.-Ижевск, Изд-во «ИКИ». – 2008. – 668 с.
21. Демидова, О.А. Методы пространственной эконометрики и оценка эффективности государственных программ. / О.А. Демидова // Прикладная эконометрика. – 2021. – № 4 (64). – С. 107-134.
22. Демьянов, В.В. Последовательный геостатистический анализ данных: примеры исследования. / В.В. Демьянов, Е.А. Савельева // Геостатистика: теория и практика. – М.: Наука, 2010. – С. 226-248.
23. Дьяконов, И.Д. Решение задачи прогнозирования при помощи градиентного бустинга над решающими деревьями / И.Д. Дьяконов, С.В. Новикова // Научный форум: Технические и физико-математические науки: сб. науч. тр. (по матер. XV Междунар. науч.-практ. конф.). – М., 2018. – Т.5. – С. 9-12.
24. Зубарева, В.Д. Методы оценки экономической эффективности НИОКР и стоимости объектов интеллектуальной собственности. / В.Д. Зубарева, О.А. Чистова // Нефть, газ и бизнес. – 2007. – № 7. – С. 14-18
25. Канторович, Г.Г. Лекции: Анализ временных рядов. Экономический журнал Высшей Школы экономики. / Г.Г. Канторович. – 2002. – Т. 6. № 3. – С. 379–401.
26. Лакман, И.А. Пространственная модель воспроизводства на панельных данных. Экономика и математические методы / И.А. Лакман, В.М. Тимирьянова. – Уфа, 2021. – Т. 57. № 2. – С. 34-44.
27. Левчук Е.А. Технологии организации, хранения и обработки данных: учеб. пособие / Е.А. Левчук. – Минск: «Вышэйшая школа», 2007. – 240 с.
28. Леонтьев, Н.Е. Основы теории фильтрации. / Н.Е. Леонтьев. – М.: Изд-во ЦПИ при механико-математическом факультете, МГУ, 2009. – С. 24–29. – 88 с.

- 29.Магнус, Я.Р. Эконометрика. Начальный курс. / Я.Р. Магнус, П.К. Катыхев, А.А. Пересецкий. – М.: «Дело», 2007. –504 с.
- 30.Мальцев В.А. Оценка анизотропии различными интерполяторами при блочном моделировании месторождений. Недропользование XXI век. / В.А. Мальцев. – 2010. – № 3. – С. 54-61.
- 31.Мееров, М. В. Оптимизация систем многосвязного управления / М.В. Мееров, Б. Л. Литвак. – М.: «Наука»,1972. – 344 с.
- 32.Носко, В.П. Эконометрика. Введение в регрессионный анализ временных рядов / В.П. Носко. – М., 2002. – 273 с.
- 33.Орлов, А.И. Вероятность и прикладная статистика: основные факты / А.И. Орлов. – М.: Изд-во «КноРус». – 2015. – 180 с.
- 34.Ошмарин, Р. А. Особенности формирования рейтинга изученности зон и программы доизучения трещинного карбонатного коллектора Куюмбинского месторождения / Б. В. Белозеров, Р.Р. Бахитов // Нефтяное хозяйство. – 2015. – № 5. – С. 52-56.
- 35.Пергамент, А.Х. Система тестов для алгоритмов геологического моделирования / И.Р Минниахметов, А.Р. Ахметсафина, А.Д. Балашов // Вестник ЦКР Роснедра. – 2009. – № 5. – С. 16-24.
- 36.Поспелова, Т.А. Исследование прогностической способности численной и аналитической моделей на примере оценки взаимовлияния скважин / Н.О. Шевцов, С.В. Степанов, Т.А. Поспелова // Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. – 2020. – Т. 6, № 3 (23). – С. 131-142.
- 37.Рапаков, Г.Г. Геоинформационное картографирование на основе машинного моделирования вариограммы / Г.Г. Рапаков, Е.А. Лебедева, В.А. Горбунов, А.Л. Кузьминов, К.А. Абдалов // Вестник Череповецкого государственного университета. – 2018. – № 3 (84). – С. 46-54
- 38.Рочев, А.Н. Повышение информативности гидродинамических исследований скважин: дис. ...канд. техн. наук: 25.00.17 / Рочев А. Н. – Ухта, 2004. – 145 с.

- 39.Скроботов, А.А. Тестирование наличия единичных корней в панельных данных при однородной альтернативе. / А.А. Скроботов // Научный вестник ИЭП им. Гайдара.ру. – 2016. – № 11. – С. 39-46.
- 40.Софронова, В.А. Метод материального баланса при оценке эффективности системы поддержания пластового давления на Верх-Тарском месторождении / В.А Софронова //М. Вестник науки и образования. – 2018. – № 12(48). – С. 18-19.
- 41.Степанов, С. В. Проблематика оценки взаимовлияния добывающих и нагнетательных скважин на основе математического моделирования / С. В. Степанов, С. В. Соколов, А. А. Ручкин, А. В. Степанов, А. В. Князев, А. В. Корытов // Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. – 2018. – Том 4. № 3. С. 146-164. DOI: 10.21684/2411-7978-2018-4-3-146-164
- 42.Тарасюк, В.М. Проблемы бизнес-прогнозирования добычи нефти на поздней стадии разработки месторождений и инструменты их решения: автореф. дис. ...д-ра экон. наук: 08.00.05 / Тарасюк Василий Михайлович. – М., 2007. – 39 с.
- 43.Хасанов, М.М. Обзор исследований по моделированию геологического строения и процессов разработки месторождения / М.М. Хасанов, Р.Р. Бахитов, И.А. Лакман // Нефтяное хозяйство. – 2021. – № 10. – С. 46-51.
- 44.Хасанов, М.М. Методика количественной оценки взаимовлияния скважин с учетом отсроченного эффекта / М.М. Хасанов, Р.Р. Бахитов, И.А. Лакман, Т.Ф. Манапов // Нефтяное хозяйство. – 2021. – №12. – С. 51-55
- 45.Хасанов, М.М. Пространственное моделирование связности добывающих скважин / М.М. Хасанов, Р.Р. Бахитов, И.А. Лакман, В.М. Тимирьянова // Нефтяное хозяйство. – 2023. – № 12. – С. 51-55
- 46.Хасанов, М.М. Развитие кост-инжиниринга в ОАО «Газпром нефть» / М. М. Хасанов, Д. А. Сугаипов и др. // Нефтяное хозяйство – 2013 – №12 – С.14-16
- 47.Хатмуллин, И.Ф. Полуаналитические модели расчета интерференции скважин на базе класса моделей CRM. / И.Ф. Хатмуллин, А.М. Андрианова, А.С. Маргарит, М.В. Симонов, Д.С. Перец, А.П. Цанда, С.А. Буденный, В.А. Лушпеев // Нефтяное хозяйство. – 2018. – № 12. – С. 38–41

- 48.Шувалова, И.С. Алгоритмизация и программирование базовых алгоритмов на VBA и Python: учеб. пособие / И.С. Шувалова, Л.Ф. Макаренко. – Москва: МАДИ, 2020. – 88 с.
- 49.Энгл, Р.Ф., Грэнджер, К.У.Дж. Коинтеграция и коррекция ошибок: представление, оценивание и тестирование / Р.Ф. Энгл, К.У.Дж. Грэнджер // Прикладная эконометрика: сб. научн.тр. – 2015. – № 3 (39). – С. 106-135.
- 50.Юдин, Е.В. Метод оценки интерференции скважин с использованием данных технологических режимов их эксплуатации / Е.В. Юдин, А.Е. Губанова, В.А. Краснов // Нефтяное хозяйство. – 2018. – № 8. – С. 64-69.
- 51.Akande, K.O. Comparative Analysis of Feature Selection-Based Machine Learning Techniques in Reservoir Characterization / K.O. Akande, O. Olatunji, T.O. Owolabi, A.A.A. Raheem // SPE Saudi Arabia Section Annual Technical Symposium and Exhibition. – 2015.
- 52.Albertoni, A. Inferring Interwell Connectivity Only From Well-Rate Fluctuations in Waterfloods / A. Albertoni, L. Lake // SPE Reservoir Evaluation & Engineering. – 2003. – 6(1). – P.6–16. – DOI:10.2118/83381-PA.
53. Anselin, L. Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL / L. Anselin, J.R. Sergio// – Chicago: GeoDa Press LLC. – 2014.
- 54.Anselin, L. Visualizing Multivariate Spatial Correlation with Dynamically Linked Windows. / L. Anselin, I. Syabri, O. Smirnov // Computing Science and Statistics. – 2002. – P.33. – <https://www.semanticscholar.org/paper/Visualizing-Multivariate-Spatial-Correlation-with-Mirnov/4e34bd70317377971ba8df7259288b972ad6a239>.
- 55.Apergis, N., A time series analysis of oil production, rig count and crude oil price: Evidence from six U.S. oil producing regions / – 2016. – Proceedings of the ICE – Energy 97:339-349. – DOI:10.1016/j.energy.2015.12.028
- 56.Artun, E. Characterizing interwell connectivity in waterflooded reservoirs using data-driven and reduced-physics models: a comparative study / E. Artun // Neural Computing and Applications. – 2017. – 28(7). – P.1729–1743. – DOI:10.1007/s00521-015-2152-0



57. Attanasi, E. D. Growth Drivers of Bakken Oil Well Productivity. / E. D. Attanasi, P. A. Freeman // *Natural Resources Research*. – 2019. – DOI:10.1007/s11053-019-09559-5
58. Bakari, H.R. Time Series Analysis Model for Production and Utilization of Gas (A Case Study of Nigeria National Petroleum Corporation “Nnpc”) / H.R. Bakari, H.A. Chamalwa, A.D. Mohammed // *IOSR Journal of Mathematics*. – 2013. – Vol. 9. – P. 17-23. – DOI:10.9790/5728-0911723.
59. Baltagi, B.H. *Econometric Analysis of Panel Data* / B.H. Baltagi // 3rd Edition, John Wiley & Sons Inc., New York. – 2005.
60. Baltagi, B.H. Testing Panel Data Regression Models with Spatial Error Correlation / B.H. Baltagi, S.H. Song, W. Koh // *Journal of Econometrics*. – 2003. – 117(1). – P.123–150. – DOI:10.1016/S0304-4076(03)00120-9.
61. Bansal, Y. Forecasting Well Performance in a Discontinuous Tight Oil Reservoir Using Artificial Neural Networks / Y. Bansal, T. Ertekin, Z. Karpyn, L. Ayala, A. Nejad, Fnu Suleen, O. Balogun, D. Liebmann, Q. Sun. // The Pennsylvania State University, Society of Petroleum Engineers This paper was prepared for presentation at the Unconventional Resources Conference-USA held in The Woodlands, Texas, USA. – 10–12 April 2013.
62. Brown, P. C. A quick look kriging technique for reservoir characterization / P. C. Brown, G. K. Falade // SPE 85679, 27th Annual SPE Technical Conference and Exhibition, Abuja, Nigeria. – 2003.
63. Chugh, S.H. Revisiting reservoir flood-surveillance methods using streamlines / S.H. Chugh, R.O. Baker // *SPE Reservoir Evaluation & Engineering*. – 2008. – V. 11. – № 2. – P. 387–394.
64. Cressie, N. Fixed rank kriging for very large spatial data sets / N. Cressie, G. Johannesson // *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. – 2008. – 70(1). – PP. 209–226. – DOI:10.1111/j.1467-9868.2007.00633.x.
65. Damsleth, E. Geostatistical approach in reservoir revaluation / E. Damsleth, H. Omre // *J. Pet. Technol.* – 1997. – 49:498–501.

66. Delfiner, P. Application of geostatistical analysis to the evaluation of petroleum reservoirs with well logs / P. Delfiner, J. P. Delhomme, J. Pelissier // SPE 1983-WW, SPWLA 24th Annual Logging Symposium, New Orleans, LA, June 27–30
67. Demiryurek, U. Neural-Network based Sensitivity Analysis for Injector-Producer Relationship Identification / U. Demiryurek, F. Banaei-Kashani, C. Shahabi, F. G. Wilkinson // Intelligent Energy Conference and Exhibition. – 2008. – DOI:10.2118/112124-MS.
68. Diggle, P. J. Model-based geostatistics / P. J. Diggle, J. A. Tawn, R. A. Moyeed // Journal of the Royal Statistical Society: Series C (Applied Statistics). – 2002. – 47(3), 299–350. DOI:10.1111/1467-9876.00113.
69. Dubé J. Spatial Econometrics Using Microdata / J. Dubé Legros // 2014. – John Wiley & Sons, Inc. DOI:10.1002/9781119008651.
70. Dumitrescu, E.-I. Testing for Granger non-causality in heterogeneous panels / E.-I. Dumitrescu, Ch. Hurlin // – Vol. 29, Issue 4, July. – 2012. – PP. 1450-1460.
71. Ediger, V. S. Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model / V.S. Ediger, S. Akar, B. Uğurlu // Energy Policy, 2006. – 34(18), 3836–3846.
72. Frausto-Solis, J. Forecasting Oil Production Time Series with a Population-Based Simulated Annealing Method / J. Frausto-Solis, M. Chi-Chim, L. Sheremetov // ARABIAN JOURNAL FOR SCIENCE AND ENGINEERING, 40(4). – 2015. – DOI: 10.1007/s13369-015-1587-z
73. Gentil, P.H. The use of multilinear regression models in patterned water floods: physical meaning of the regression coefficients / P.H. Gentil // MS thesis, The University of Texas at Austin, Texas. – 2005.
74. Giraldo, R. Hierarchical clustering of spatially correlated functional data / Giraldo R., Delicado P., Mateu J. // Statistica Neerlandica. – 2012. 66(4), P.403–421. DOI:10.1111/j.1467-9574.2012.00522.x.
75. Hassani Pak, A. A. Geostatistics (1st ed.) / A. A. Hassani Pak // Tehran, Iran: Tehran University Press. – 1998.

- 76.Hrafnkelsson, B. Environmental and Ecological Statistics. / B. Hrafnkelsson // – 2003. – 10(2), 179–200. DOI:10.1023/a:1023674107629.
- 77.Hu, H. Forecasting Method of Crude Oil Output Based on Optimization of LSSVM by Particle Swarm Algorithm / H. Hu, X. Huand, X. Guan // 4th International Conference on Information Science and Control Engineering (ICISCE), Changsha. – 2017. P. 334-338.
- 78.Hu, H. There search on modeling and simulation of crude oil output prediction based on KPCA-DE-SVM / H. Hu, L. Fanand, X. Guan // 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), Beijing. – 2017. – P. 93-97.
- 79.Johannesson, G. (). Finding large-scale spatial trends in massive, global, environmental datasets / G. Johannesson, N. Cressie // Environmetrics. 2003. – 15(1), 1–44. DOI:10.1002/env.624.
- 80.Kammann, E. E. Geoadditive models. Journal of the Royal Statistical Society / E. E. Kammann, M. P. Wand // Series C, Appl. Statist. 2003. – 52(1). – P. 1–18. DOI:10.1111/1467-9876.00385.
- 81.Kapoor, M. Panel Data Model with Spatially Correlated Error Components / M. Kapoor, H.H. Kelejian, I.R. Prucha // Journal of Econometrics. – 2007. – 140(1). – P.97–130. DOI:10.1016/j.jeconom.2006.09.004
- 82.Kaviani, D. (). Interwell connectivity evaluation from wellrate fluctuations: a waterflooding management tool / D. Kaviani // Doctoral dissertation, Amirkabir University of Technology, Tehran Polytechnic. – 2009. – Retrieved from oaktrust.library.tamu.edu
- 83.Kelkar, M. Application of geostatistics for reservoir characterization accomplishments and challenges / M. Kelkar // J. Can. Pet. Technol. – 2000. – 39(7). 2000:25–29
- 84.Khasanov, M. M., "Optimization of Production Capacity for Oil Field in the Russian Arctic. / R. R. Bakhitov, A. N. Sitnikov, O. S. Ushmaev, D. N. Dmitruk, S. A. Nekhaev. // Paper presented at the SPE Arctic and Extreme Environments Technical Conference and Exhibition, Moscow, Russia, October 2013. DOI: <https://doi.org/10.2118/166905-MS>

85. Khatmullin, I. F. Semi-analytical models for calculating well interference: Limitations and applications / I. F. Khatmullin, A. P. Tsanda, A. M. Andrianova, S. A. Budenny, A. S. Margarit, V. A. Lushpeev, M. Simonov, D. S. Perets // *Neftyanoe khozyaystvo - Oil Industry*. – 2018. – 12, P. 38-41. <https://DOI.org/10.24887/0028-2448-2018-12-38-41>.
86. Kim, M.S. One Method of the Reserves Calculation by Using the Gaussian Disjunctive Kriging / M.S. Kim, G.Ch. Hwang, O Ch.N. // *Наука, техника и образование*. – 2017. – Т. 1. – № 5 (35). – С. 101-104.
87. Klenner, R. Characterization of Fracture-Driven Interference and the Application of Machine Learning to Improve Operational Efficiency / R. Klenner, G. Liu, H. Stephenson, G. Murrell, N. Iyer, N. Virani, A. Charuvaka // – 2018. – SPE Liquids-Rich Basins Conference - North America. – DOI:10.2118/191789-MS
88. Lake, L. Optimization Of Oil Production Based On A Capacitance Model Of Production And Injection Rates / L. Lake, X. Liang, T. F. Edgar, A. Al-Yousef, M. Sayarpour, D. Weber // *Hydrocarbon Economics and Evaluation Symposium*. – 2007. – DOI:10.2118/107713-MS
89. Lee, K.-H. (2010). A Multivariate Autoregressive Model for Characterizing Producer-producer Relationships in Waterfloods from Injection/Production Rate Fluctuations / K.-H. Lee, A. Ortega, N. Jafroodi, I. Ershaghi // *SPE Western Regional Meeting*. – 2010. – DOI:10.2118/132625-MS
90. Levin, A. Unit root tests in panel data: asymptotic and finite-sample properties. / A. Levin, C.-F. Lin, C.-S. J. Chu // *Journal of Econometric*. – 2002. – 108(1). – P. 1–24. – DOI:10.1016/S0304-4076(01)00098-7
91. Liang, J. S. Well Group Connectivity Relations Discriminate Based on CART Algorithm. / J. S. Liang, W. Zheng, Z. B. Yuan // *Applied Mechanics and Materials*. – 2014. – PP. 513-517, 1252–1255. DOI:10.4028/www.scientific.net/amm.513-517.1252
92. Liu, W. Reservoir Inter-Well Connectivity Analysis Based on a Data Driven Method / W. Liu, W. D. Liu, J. Gu // *Abu Dhabi International Petroleum Exhibition & Conference*. – 2019. – DOI:10.2118/197654-MS
93. Lütkepohl, H. Новое введение в анализ множественных временных рядов / H. Lütkepohl // Берлин: Springer, 2007. – С. 222–229.

94. Maojun, C. Study on inferring interwell connectivity of injection-production system based on decision tree / C. Maojun, Sh. Fuhua // 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). – 2013. – DOI:10.1109/FSKD.2013.6816343
95. Marsily, Gh. Dealing with spatial heterogeneity / Gh. Marsily, F. Delay, J. Gonalvs, Ph. Renard, V. Teles, S. Violette // Hydrogeol Journal / – 2005. – 13(1). – P.161–183
96. Montgomery, J. B. Spatial variability of tight oil well productivity and the impact of technology / J. B. Montgomery, F. M. O’Sullivan // Applied Energy. – 2017. — 195. – P.344–355. DOI:10.1016/j.apenergy.2017.03.03
97. Mur, J. The Spatial Durbin Model and the Common Factor Tests / J. Mur, A. Angulo // Spatial Economic Analysis. – 2006. – Vol. 1 (2). – P. 207-226
98. Nazarpour, A. Geostatistical Modeling of Spatial Distribution of Porosity in the Asmari Reservoir of Mansuri Oil Field in Iran / A. Nazarpour , S. R. Shadizadeh, G. Zargar // Petroleum Science and Technology. – 2014. – 32(11). – P.1274-1282. – DOI:10.1080/10916466.2011.594835
99. Nwachukwu, A. Machine Learning-Based Optimization of Well Locations and WAG Parameters under Geologic Uncertainty / A. Nwachukwu, H. Jeong, A. Sun, M. Pyrcz, L. W. Lake // SPE Improved Oil Recovery Conference. – 2018. – DOI:10.2118/190239-MS
100. Nwachukwu, A., (2018). Fast evaluation of well placements in heterogeneous reservoir models using machine learning / A. Nwachukwu, H. Jeong, M. Pyrcz, L. W. Lake // Journal of Petroleum Science and Engineering – 2018. – 163. – P.463–475. – DOI:10.1016/j.petrol.2018.01.019
101. Olenchikov, D. M. Povysheniyye tochnosti otsenki produktivnosti plasta pri pomoshchi ucheta statisticheskikh dannyykh o ego svoystvakh [Improving the Accuracy of Reservoir Productivity Evaluation by using Reservoir Properties Statistical Data] / Olenchikov D. M., Sapozhnikov A. E. et al // Nauchno-technicheskiy vestnik OAO «NK «Rosneft». — 2011. – № 23. — P. 39-41.

102. Salinas, D. DeepAR: Probabilistic forecasting with autoregressive recurrent networks / D. Salinas, V. Flunkert, J. Gasthaus, T. Januschowski // International Journal of Forecasting. – Vol.36(3). – 2020. – P. 1181–1191
103. Sayarpour, M. Development and application of capacitance-resistive models to water/CO<sub>2</sub> floods (Thesis for: Ph.D, The University of Texas at Austin). Retrieved from utexas.edu. – 2008. – DOI:10.13140/RG.2.1.1798.3847
104. Sayarpour, M. The use of capacitance-resistance models for rapid estimation of waterflood performance and optimization / M. Sayarpour, E. Zuluaga, Sh. Kabir, L. Lake // Journal of Petroleum Science and Engineering. – 2009. – 69(3-4). – P. 227–238. – DOI:10.1016/j.petrol.2009.09.006
105. Smith, R. Correlating geologic and seismic data with unconventional resource production curves using machine learning / R. Smith, T. Mukerji, T. Lupo // Geophysics. – 84 (2). – P. 1-37. – O39–O47. DOI: <https://DOI.org/10.1190/geo2018-0202.1>
106. Suhartono, D.D. Comparison between VAR, GSTAR, FFNN-VAR and FFNN-GSTAR Models for Forecasting Oil Production / D.D. Suhartono, H. Prastyo, M. Kuswanto, M.H. Lee // MATEMATIKA. – 2018. – Vol. 34(1). – P. 103–111
107. Tao, L. A New Productivity Prediction Hybrid Model for Multi-Fractured Horizontal Wells in Tight Oil Reservoirs / Tao, L. Guo, J. Guo, X. Zhou, A. Kitaeva, J. Zeng // 2018. – 10.2118/191714-18RPTC-MS
108. Valko, P.P. Development and Application of the Multiwell Productivity Index (MPI) / P.P. Valko, L.E. Doublet, T.A. Blasingame // 2000. – SPE Journal 5(1):21-31. – DOI:10.2118/51793-PA.
109. Wang, O. A novel hybridization of nonlinear grey model and linear ARIMA residual correction for forecasting U.S. shale oil production / O. Wang, S. Xiaoxing, R. Li. // 2018. – Energy. 165(6339):1320-1331. – DOI:10.1016/j.energy.2018.10.032
110. Weber, D. Improvements in Capacitance-Resistive Modeling and Optimization of Large Scale Reservoirs / T. F. Edgar, L. W. Lake, L. S. Lasdon, S. Kawas, M.Sayarpour // SPE Western Regional Meeting. – 2009. – DOI: 10.2118/121299-MS

111. Wigwe, M. E. Application of Geographically Weighted Regression to Model the Effect of Completion Parameters on Oil Production. Case Study on Unconventional Wells. / M. E. Wigwe, M. C Watson, A. Giussani, E. Nasir, S. Dambani // Paper presented at the SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, August 2019. DOI: <https://DOI.org/10.2118/198847-MS>
112. Wiliantoro, O. Production decline analysis of oil and gas resources with robustfit and timeseries analysis / O. Wiliantoro, D. Niu, H. Li // *Int. J. Oil, Gas and Coal Technology*. – 2019. – Vol. 20, No. 1. – P.1–30
113. Yousef, A. A. A Capacitance Model to Infer Interwell Connectivity from Production and Injection Rate Fluctuations. / A. A. Yousef, P. H. Gentil, J. L. Jensen, L. W. Lake // *SPE Reservoir Evaluation & Engineering*. – 2006. – 9(06). – P. 630–646. DOI:10.2118/95322-pa.
114. Yousef, A. A. Analysis and Interpretation of Interwell Connectivity from Production and Injection Rate Fluctuations Using a Capacitance Model / A. A. Yousef, L. W. Lake, J. L. Jensen // *SPE/DOE Symposium on Improved Oil Recovery*. – 2006. – DOI:10.2118/99998-MS
115. Yousef, A.A. Investigating statistical techniques to infer interwell connectivity from production and injection rate fluctuations / Yousef A.A // (Doctoral dissertation, The University of Texas at Austin). Retrieved from [utexas.edu](http://utexas.edu). – 2006.
116. Yusof, N.M., Malaysia crude oil production estimation: an application of ARIMA model / N.M. Yusof, R.S.A. Rashid, Zamzulani, M. // *International Conference on Science and Social Research*. – 2010. – Kuala Lumpur, Malaysia. – P. 1255–1259
117. Zimmerman, D.A. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. / D.A. Zimmerman, G. de Marsily, C.A. Gotaway, M.G. Marietta, C.L. Axness, R. Beauheim, R. Bras, J. Carrera, G. Dagan, P.B. Davies, D. Gallegos, A. Galli, J. Gomez-Hernandez, P. Grindrod, A.L. Gutjahr, P. Kitanidis, A.M. Lavenue, D. McLaughlin, S.P. Neuman, B.S. Ramarao, C. Ravenne, Y. Rubin // *Water Resour Res*. – 1998. – 34(6). – P. 1373–1413

## Приложение А

### Добывающие скважины на месторождении: со 2 по 73 скважину

Ежедневные данные, на этих скважинах велась только добыча нефти

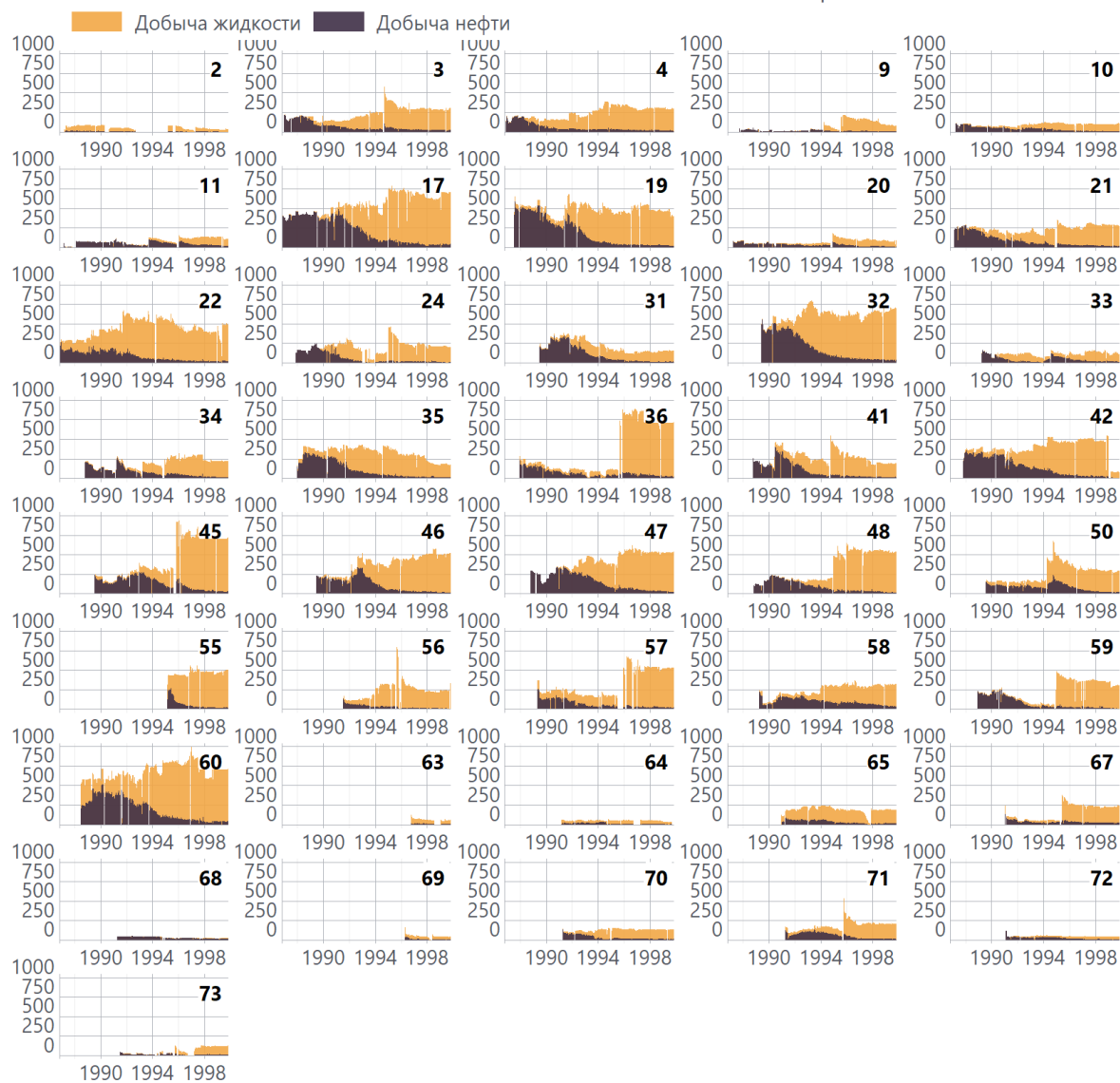


Рисунок А.1 – Графики объёмов добычи нефти и жидкости со 2 по 73 скважинам на в период 1990-98 гг.



## Добывающие скважины на месторождении: С 74 по 114

Ежедневные данные, на этих скважинах велась только добыча нефти

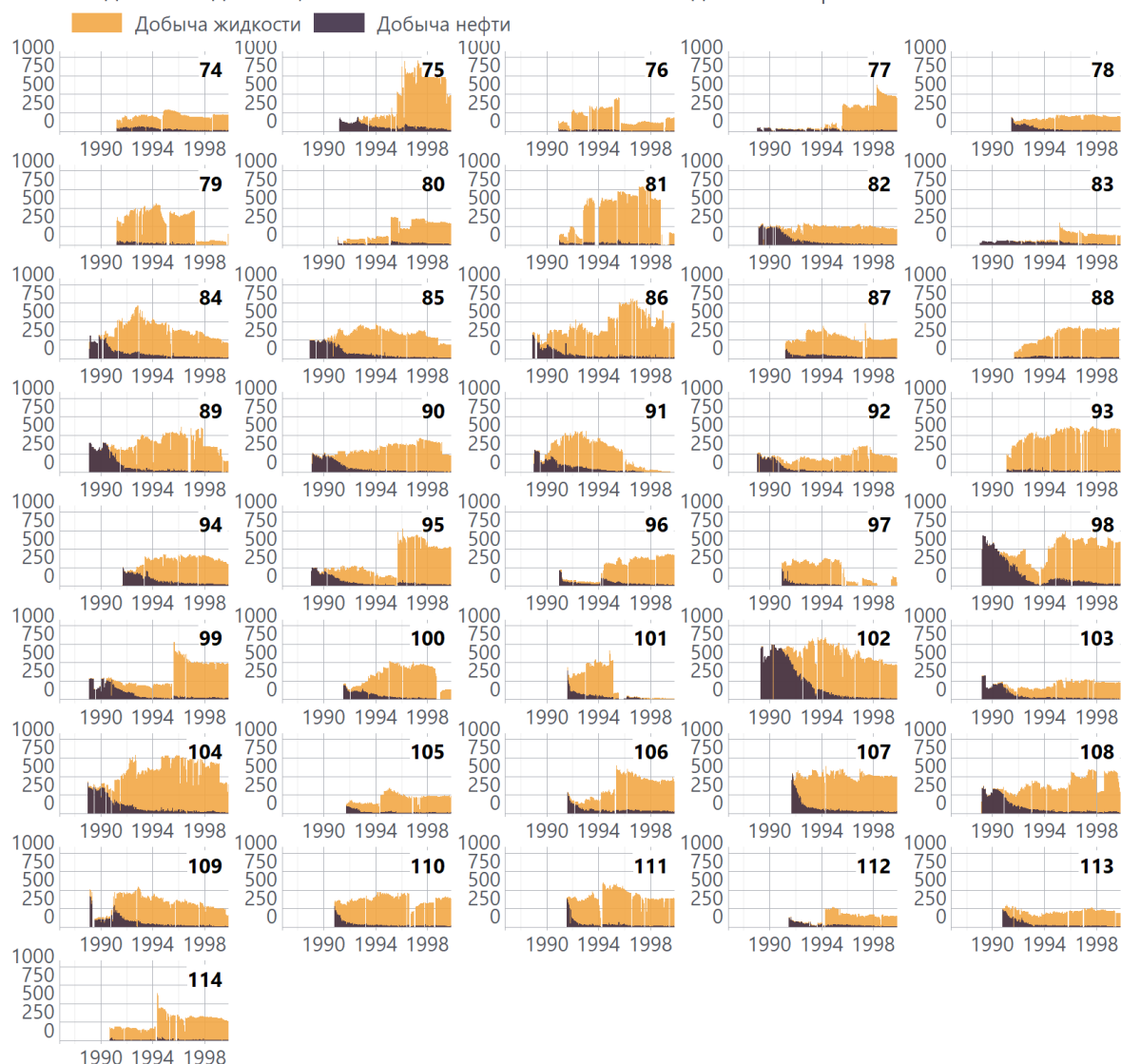


Рисунок А.2 – Графики объемов добычи нефти и жидкости с 74 по 114 скважинам на в период 1990-98 гг.

## Приложение Б

На Рисунках В.1 – В.4 представлены h-графики для направлений 0, 45, 90 и 135 градусов и длиной векторов h в интервале от 0 до 5000

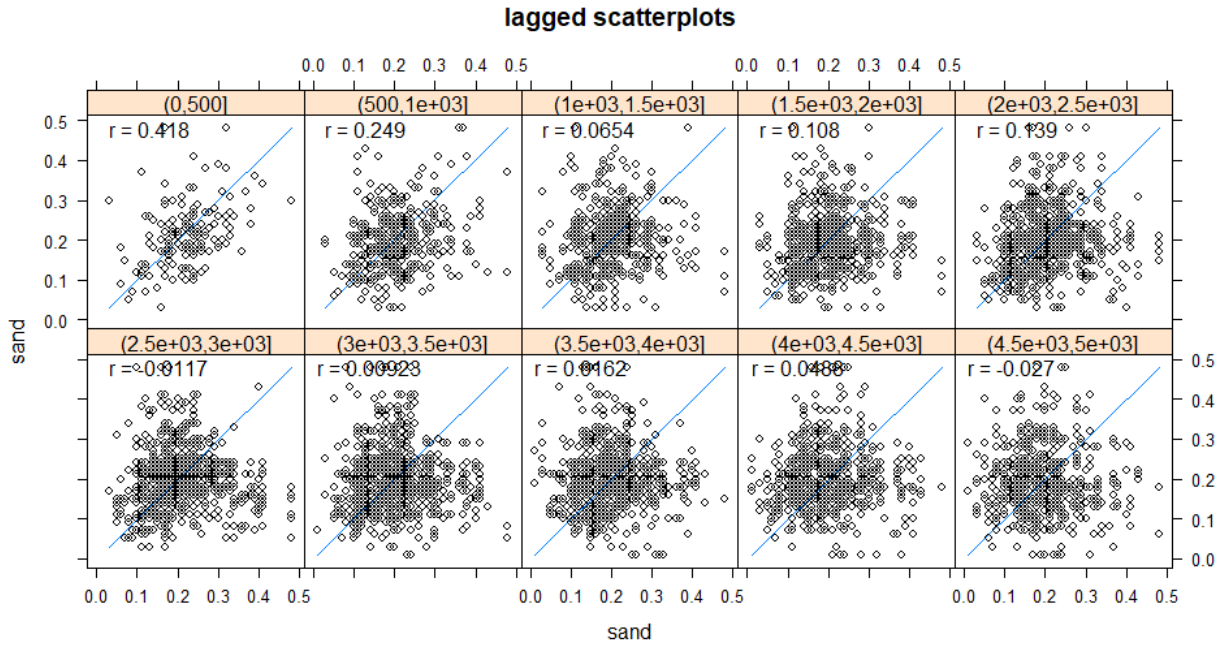


Рисунок В.1 – h-график для направления 0 градусов

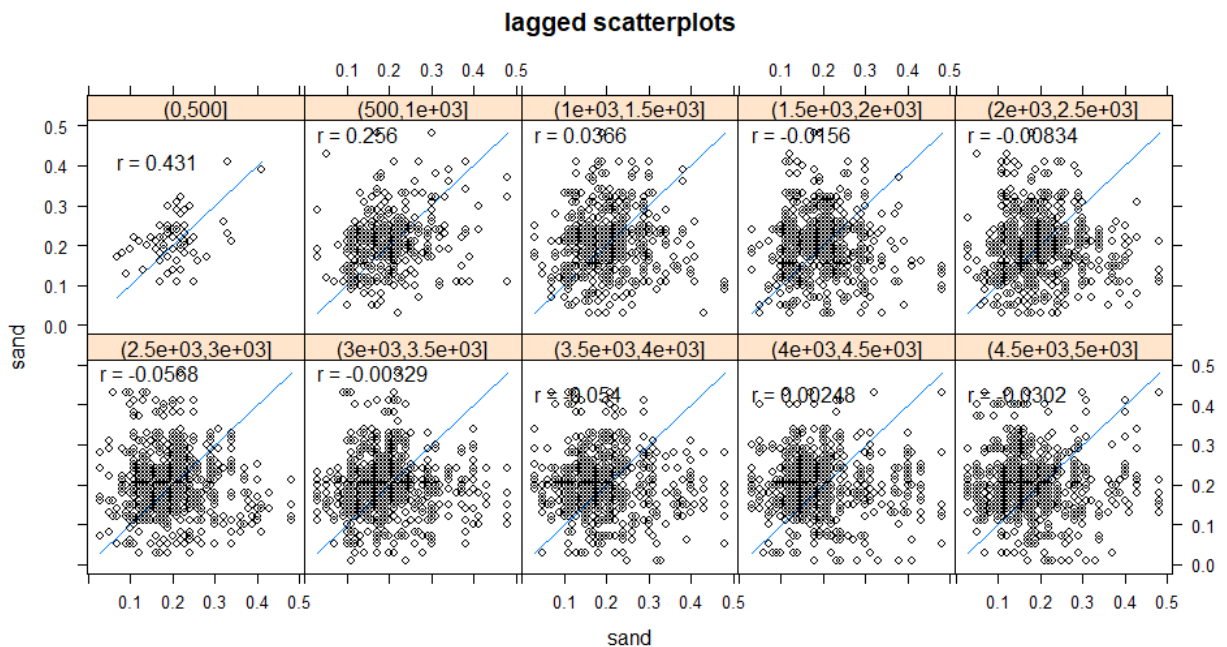


Рисунок В.2 – h-график для направления 45 градусов

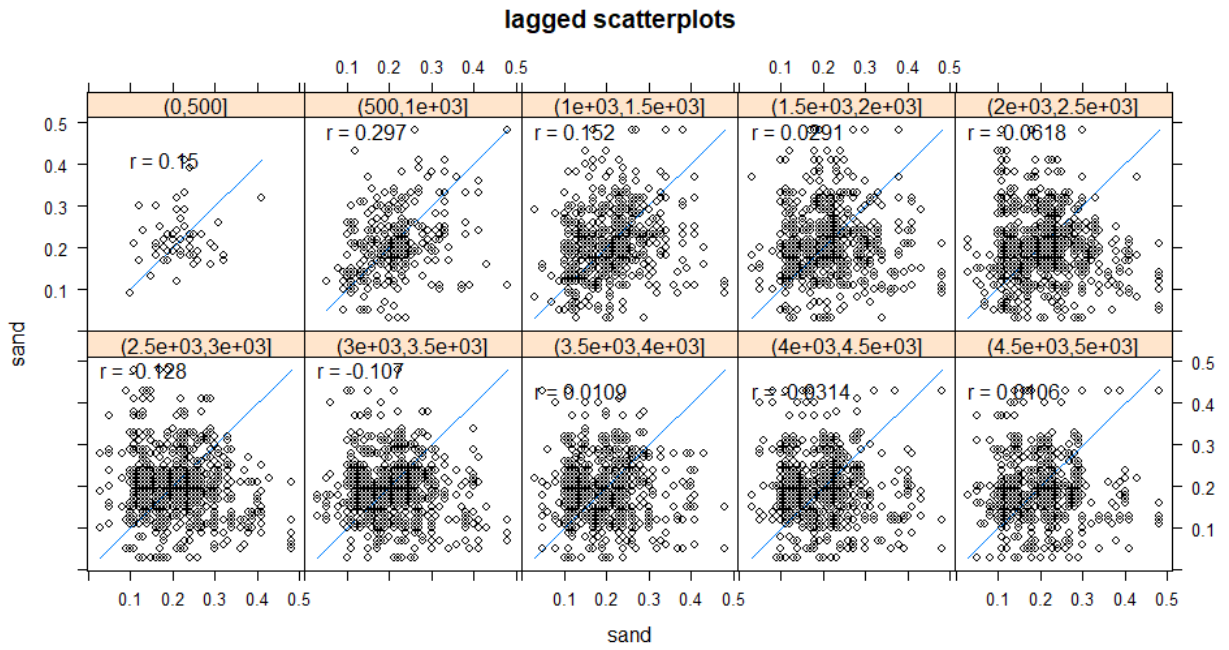


Рисунок В.3 – h-график для направления 90 градусов

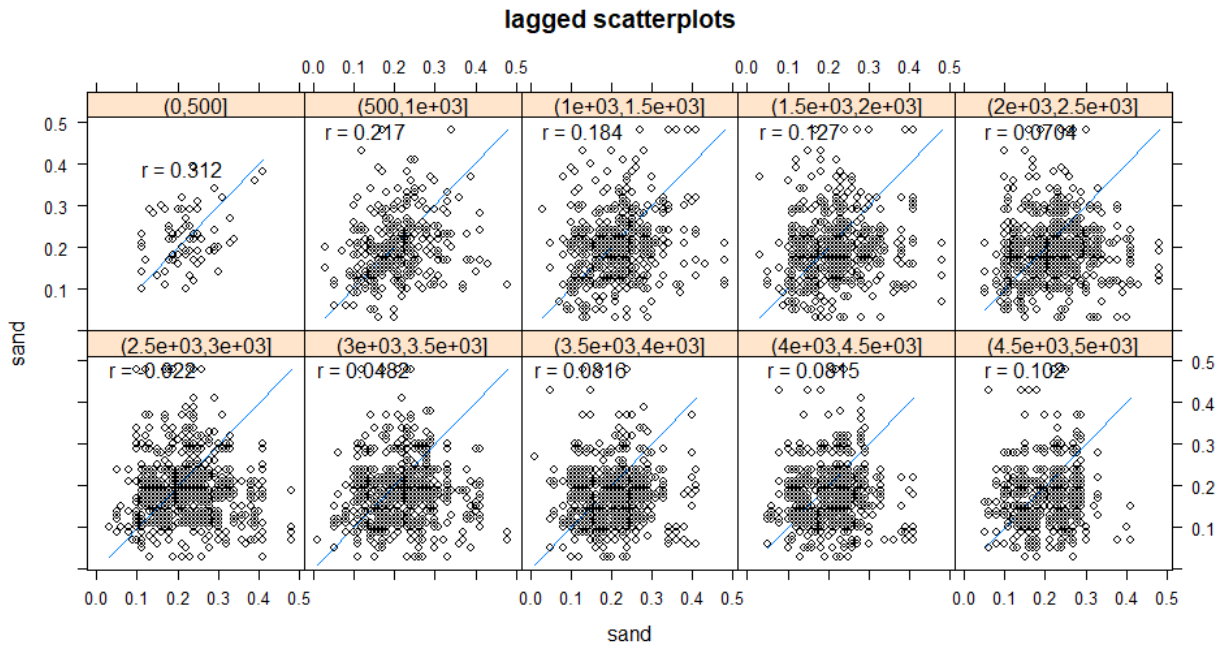


Рисунок В.4 – h-график для направления 135 градусов

## Приложение С

Таблица С.1 – Пространственные фиксированные эффекты для скважин:

	Оценка	Ст.ошибка	$t$	$\Pr(> t )$	
К.	311.8321	77.508	4.0232	< 0.0001	***
1	-296.9852	5.4512	-54.4812	< 0.0001	***
2	-38.7812	4.8074	-8.0669	< 0.0001	***
3	-33.0462	4.9034	-6.7395	< 0.0001	***
4	-191.1963	4.4711	-42.7625	< 0.0001	***
5	-230.9210	4.8318	-47.7922	< 0.0001	***
6	-195.1038	4.6647	-41.8252	< 0.0001	***
7	342.4124	4.3440	78.8237	< 0.0001	***
8	163.6170	4.9222	33.2409	< 0.0001	***
9	-232.6713	5.1348	-45.3127	< 0.0001	***
10	-32.3885	4.6887	-6.9077	< 0.0001	***
11	99.1358	4.3499	22.7902	< 0.0001	***
12	-103.2824	4.7785	-21.6140	< 0.0001	***
13	-166.9449	4.8590	-34.3582	< 0.0001	***
14	325.3947	4.4551	73.0387	< 0.0001	***
15	-180.8681	4.2917	-42.1438	< 0.0001	***
16	-72.6228	4.5894	-15.8241	< 0.0001	***
17	-92.2222	4.2934	-21.4801	< 0.0001	***
18	474.4923	4.5493	104.3009	< 0.0001	***
19	-109.5825	4.4395	-24.6836	< 0.0001	***
20	48.3201	4.9556	9.7505	< 0.0001	***
21	423.2273	4.0436	104.6648	< 0.0001	***
22	182.0899	4.6594	39.0802	< 0.0001	***
23	219.7073	4.9956	43.9798	< 0.0001	***
24	233.5472	4.6761	49.9454	< 0.0001	***
25	-13.4711	3.8111	-3.5347	0.0004	***
26	172.7626	4.0206	42.9697	< 0.0001	***
27	-75.0321	4.1446	-18.1038	< 0.0001	***
28	207.2448	4.7313	43.8029	< 0.0001	***
29	-9.4001	4.5026	-2.0877	0.0368	*
30	26.5243	4.6087	5.7553	< 0.0001	***
31	479.0854	4.3802	109.3750	< 0.0001	***
32	-279.2866	5.2564	-53.1327	< 0.0001	***
33	-284.3658	5.2008	-54.6774	< 0.0001	***
34	-160.3273	5.4853	-29.2287	< 0.0001	***
35	-93.8454	4.5936	-20.4297	< 0.0001	***
36	-300.1067	5.1121	-58.7047	< 0.0001	***
37	-277.7460	4.2934	-64.6921	< 0.0001	***
38	-193.6786	5.0622	-38.2600	< 0.0001	***
39	-108.8148	4.5046	-24.1562	< 0.0001	***
40	-261.7495	4.0864	-64.0531	< 0.0001	***
41	-218.3505	5.1126	-42.7080	< 0.0001	***
42	-118.0403	4.7895	-24.6457	< 0.0001	***
43	405.7557	4.2916	94.5464	< 0.0001	***
44	-180.6804	5.3822	-33.5703	< 0.0001	***
45	102.5083	5.3169	19.2797	< 0.0001	***
46	-117.7070	5.4251	-21.6966	< 0.0001	***
47	-226.7981	5.1740	-43.8340	< 0.0001	***
48	18.9760	4.1301	4.5946	< 0.0001	***
49	113.6603	5.0748	22.3972	< 0.0001	***

продолжение таблицы С.1

50	-72.9887	4.5187	-16.1526	< 0.0001	***
51	-176.4230	4.6828	-37.6750	< 0.0001	***
52	-27.6135	4.5215	-6.1072	< 0.0001	***
53	-37.4330	4.3129	-8.6794	< 0.0001	***
54	207.3539	4.7777	43.4007	< 0.0001	***
55	-57.5699	4.5802	-12.5694	< 0.0001	***
56	71.3868	5.1331	13.9072	< 0.0001	***
57	41.6662	4.6400	8.9798	< 0.0001	***
58	49.2110	5.0172	9.8085	< 0.0001	***
59	-281.5308	4.5061	-62.4780	< 0.0001	***
60	-46.9764	4.0754	-11.5269	< 0.0001	***
61	253.7933	4.3524	58.1234	< 0.0001	***
62	62.8466	4.9928	12.5875	< 0.0001	***
63	247.5542	4.7675	51.9253	< 0.0001	***
64	75.9651	5.4752	13.8743	< 0.0001	***
65	-267.1238	5.2499	-50.8819	< 0.0001	***
66	277.4889	4.8932	56.7093	< 0.0001	***
67	175.7903	5.4500	32.2554	< 0.0001	***
68	8.9404	4.8222	1.8540	0.0637.	
69	-278.4065	4.0714	-68.3803	< 0.0001	***
70	209.8116	5.1063	41.0884	< 0.0001	***
71	-73.4316	4.7164	-15.5694	< 0.0001	***
72	289.5202	4.5155	64.1174	< 0.0001	***
73	-79.3428	4.2947	-18.4747	< 0.0001	***
74	141.3872	4.0221	35.1524	< 0.0001	***
75	204.4798	4.4305	46.1532	< 0.0001	***
76	162.7199	4.8235	33.7350	< 0.0001	***
77	-47.5629	4.6773	-10.1688	< 0.0001	***
78	7.8240	5.2010	1.5043	0.1325	
79	67.7071	4.2045	16.1037	< 0.0001	***
80	-157.3254	4.2971	-36.6122	< 0.0001	***
81	-90.5655	4.6664	-19.4079	< 0.0001	***
82	-3.5973	4.8196	-0.7464	0.4554	

## Приложение D

Таблица D. 1 – Оценка пространственных моделей в различных спецификациях с обобщенными панельными эффектами

Наименование	1	2	3	4	5	6	7
Название модели	W_pool_SAR	W_pool_SEMb	W_pool_SEMk	W_pool_SEMn	W_pool_SACb	W_pool_SACk	W_pool_SACn
Тип модели	Обобщённая панельная модель с пространственным лагом	Обобщённая панельная модель с пространственной ошибкой	Обобщённая панельная модель с пространственной ошибкой (метод ML)	Обобщённая панельная модель с независимыми ошибками	Обобщённая панельная модель с пространственным лагом и с пространственной ошибкой	Обобщённая панельная модель с пространственным лагом и с пространственной ошибкой (метод ML)	Обобщённая панельная модель с пространственным лагом и с независимыми ошибками
Спецификация в R	model="pooling", lag=TRUE, spatial.error=NULL	model="pooling", lag=FALSE, spatial.error="b"	model="pooling", lag=FALSE, spatial.error="kcp"	model="pooling", lag=FALSE, spatial.error="none"	model="pooling", lag=TRUE, spatial.error="b"	model="pooling", lag=TRUE, spatial.error="kcp"	model="pooling", lag=TRUE, spatial.error="none"
Константа	486.848*** (5.24)				486.848*** (5.24)		
Внутрипластовое давление	-1.084*** (0.027)			Просто сквозная панель	-1.084*** (0.027)		
Коэффициент пространственной ошибки	0.0385*** (0.0055)				0.0385*** (0.0055)		
Пространственный автокорреляционный коэффициент	0.1058*** (0.0066)				0.1058*** (0.0066)		

Статистическая значимость: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Таблица D. 2 – Оценка пространственных моделей в различных спецификациях с фиксированными двунаправленными панельными эффектами

Наименование	35	36	37	39	40	41
Тип модели	Панельная модель с пространственным лагом и фиксированными эффектами по периодам и кросс-секциям	Панельная модель с пространственной ошибкой и фиксированными эффектами по периодам и кросс-секциям		Панельная модель с пространственным лагом и с пространственной ошибкой и фиксированными эффектами по периодам и кросс-секциям		
Название модели	W_within_SAR	W_within_SEMb	W_within_SEMk	W_within_SACb	W_within_SACk	W_within_SACn
Спецификация в R	model = "within", effect = "twoways", lag = TRUE, spatial.error = NULL	model = "within", effect = "twoways", lag = FALSE, spatial.error = "b"	model = "within", effect = "twoways", lag = FALSE, spatial.error = "kkp"	model = "within", effect = "twoways", lag = TRUE, spatial.error = "b"		
Внутрипластовое давление	0.1041*** (0.0209)	0.10595*** (0.0209)	Тот же результат что и spatial.error = "b"	0.105341*** (0.0209)		
Коэффициент пространственной ошибки		-0.0671*** (0.00575)		-0.039501 (p=0.9) (0.4218)		
Пространственный автокорреляционный коэффициент	-0.0668*** (0.00575)			-0.02777 (p=0.9) (0.42)		
R <sup>2</sup>	0.8586	0.8583		0.8584		
AIC	989943.7	990139.9		990045		
BIC	989971.8	990168		990082.5		

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Таблица D. 3 – Оценка пространственных моделей в различных спецификациях с фиксированными кросс-секционными панельными эффектами

Наименование	48	49	50	51	52	53
effect = "individual"	Панельная модель с пространственным лагом и фиксированными эффектами по кросс-секциям	Панельная модель с пространственной ошибкой и фиксированными эффектами по кросс-секциям		Панельная модель с пространственным лагом и с пространственной ошибкой и фиксированными эффектами по кросс-секциям		
Название модели	W_within_SAR	W_within_SEMb	W_within_SEMk	W_within_SACb	W_within_SACk	W_within_SACn
Тип модели	model = "within", effect = "individual", lag = TRUE, spatial.error = NULL	model = "within", effect = "individual", lag = FALSE, spatial.error = "b"	model = "within", effect = "individual", lag = FALSE, spatial.error = "kcp"	model = "within", effect = "individual", lag = TRUE, spatial.error = "b"		
Внутрипластовое давление	0,242426*** (0,021186)	0.231425 *** (0,021199)	Тот же результат что и spatial.error = "b"	0,25198*** (0,1918)		
Коэффициент пространственной ошибки		0.0740810 *** (0,00548)		-0.57508 *** (0,01425)		
Пространственный автокорреляционный коэффициент	0,07605*** (0,00547)			0,51935*** (0,00936)		
R <sup>2</sup>	0,8504	0,8499		0,8462		
AIC	994753.5	995034.7		997140.3		
BIC	994781.6	995062.8		997177.8		

signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1



Таблица D. 4 – Оценка пространственных моделей в различных спецификациях с фиксированными панельными эффектами по периодам

Наименование	42	43	44	45	46	47
Тип модели	Панельная модель с пространственным лагом и фиксированными эффектами по периодам	Панельная модель с пространственной ошибкой и фиксированными эффектами по периодам		Панельная модель с пространственным лагом и с пространственной ошибкой и фиксированными эффектами по периодам		
Название модели	W_within_SAR	W_within_SEMb	W_within_SEMk	W_within_SACb	W_within_SACk	W_within_SACn
effect=" time"	model = "within", effect = "time", lag = TRUE, spatial.error = NULL	model = "within", effect = "time", lag = FALSE, spatial.error = "b"	model = "within", effect = "time", lag = FALSE, spatial.error = "kkp"	model = "within", effect = "time", lag = TRUE, spatial.error = "b"		
FormationPressure	-0.77679*** (0.0258)	-0.68383*** (0.0276)	Тот же результат что и spatial.error = "b"	-0.77676*** (0.0264)		
Spatial error parameter:		0.40087*** (0.00448)		-0,0177(p=0,51) (0.0269)		
Spatial autoregressive coefficient:	0.3992*** (0.00447)			0.41216*** (0.021)		
R <sup>2</sup>	0.128	0.02586		0.13		
AIC	1145662	1155151		1145474		
BIC	1145690	1155179		1145511		

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Приложение Е

Листинг программы для построения пространственных моделей SAC и SAR

```
# --- Пакеты -----
library(lubridate)
library(readxl)
library(ape)
library(spdep)
library(rgdal)
library(maptools)
library(rgeos)
library(spatialreg)
library(splm)
library(tidyverse)
options(scipen = 999)
# --- Данные -----
oil <- read_excel("Data/Production_wells_cleaned2.xlsx")
coords <- read_excel("Data/Prod_Well_coordinates.xlsx")
# --- Подготовка данных -----
# Замена всех пропусков нулями
oil[is.na(oil)] <- 0
# Создание наборов данных для построения моделей
# oil_1995_05_31 - добыча на определенный день для SDM/SAR
# oil_panel - Добыча за определенный период для панельной пространственной модели
oil_1995_05_31 <- oil %>% mutate(Date = as.Date(Date)) %>% filter(Date == "1999-05-31")
oil_panel <- oil %>% mutate(Date = as.Date(Date)) %>% filter(Date >= "1997-01-01" & Date <
"1999-06-01")

# --- Матрица расстояния -----
# Матрица с координатами
coords_matrix <- as.matrix(coords[, 2:3])
# Матрица соседей. Задано расстояние 750 метров
# Для каждой скважины находятся соседи на этой дистанции
dist_matrix <- dnearneigh(x = coords_matrix, d1 = 0, d2 = 750, row.names = coords$Well)
# Весовая матрица
dist_matrix_listw <- nb2listw(dist_matrix)
summary(dist_matrix_listw)

# --- Пространственная модель --- SDM, SAR -----
# Обе модели построены на обоих имеющихся давлениях
# Можно оставить только одно необходимое давление:
# Liquid_production_rate ~ FormationPressure
# Liquid_production_rate ~ BottomHolePressure
# --- SAR ---
model_sar <- lagsarlm(formula = Liquid_production_rate ~ FormationPressure + BottomHolePressure,
# Формула
                    data = oil_1995_05_31, # Данные
                    listw = dist_matrix_listw # Весовая матрица)
```

```

summary(model_sar, Nagelkerke = TRUE) # Результаты модели

# --- SDM ---
model_durbin <- lagsarlm(formula = Liquid_production_rate ~ FormationPressure +
BottomHolePressure,
      data = oil_1995_05_31,
      listw = dist_matrix_listw,
      Durbin = TRUE) # Требуется для построения SDM модели
summary(model_durbin, Nagelkerke = TRUE)

# --- Пространственная модель --- Панельная -----
# --- Модель с забойным и пластовым давлением ---
model_within <- spml(formula = Liquid_production_rate ~ FormationPressure + BottomHolePressure,
      data = oil_panel, listw = dist_matrix_listw,
      model = "within", lag = TRUE, effect = "twoways")
summary(model_within) # Результаты модели
summary(model_within)$rsqr # R2

# Фиксированные эффекты по кросс-секциям
model_effects_spatial <- round(effects.splm(model_within)$SETable, 3)
model_effects_spatial
# Фиксированные эффекты по периодам
model_effects_time <- round(effects.splm(model_within)$TETable, 3)
model_effects_time

# --- Модель с пластовым давлением ---
model_within_2 <- spml(formula = Liquid_production_rate ~ FormationPressure,
      data = oil_panel, listw = dist_matrix_listw,
      model = "within", lag = TRUE, effect = "twoways")
summary(model_within_2)
summary(model_within_2)$rsqr
model_2_effects_spatial <- round(effects.splm(model_within_2)$SETable, 3)
model_2_effects_spatial

model_2_effects_time <- round(effects.splm(model_within_2)$TETable, 3)
model_2_effects_time

```

Приложение F

Таблица F.1 – MAPE (%) для моделей, прогнозирующих объем добычи нефти на 30 дней горизонта прогноза

Скважина	Экзогенная переменная - пластовое давление		Экзогенная переменная - разница давлений		Экзогенная переменная - пластовое давление на новом кусте		Изменения отсечения начала обучающей выборки (максимум 1 мая) по ДСММ		Изменения отсечения начала обучающей выборки (максимум 1 апреля) по ДСММ	
	Начало	Середина	Начало	Середина	Начало	Середина	Начало	Середина	Начало	Середина
2	5,4	13,6	8,5	16,0	6,7	12,2	12,0	15,4	12,0	15,4
3	10797,0	34,6	6385,8	58,8	10813,6	35,3	4429,3	81,5	10342,4	97,4
4	15,2	18,0	19,9	21,4	19,3	20,4	32,5	14,3	32,5	14,3
9	14,8	11,4	27,3	4,8	19,4	6,7	12,2	4,9	12,2	4,9
10	10,8	19,0	17,4	22,5	13,2	17,9	29,6	19,0	29,6	19,0
11	1108,9	247,4	3140,9	636,3	2199,5	294,7	1203,3	279,7	3026,2	206,6
17	24,7	15,0	17,4	12,2	21,2	14,0	24,1	17,2	24,1	17,2
19	16,7	26,2	6,9	18,1	7,6	14,7	26,5	13,1	26,5	13,1
20	2,5	18,7	7,2	21,8	2,7	21,3	3,7	16,0	3,7	16,0
21	19,2	16,0	18,9	14,2	22,5	18,8	20,2	21,6	20,2	21,6
22	29,1	30,0	25,8	29,4	30,1	29,9	33,1	29,0	33,1	29,0
24	284,9	210,8	233,4	174,9	243,6	180,9	162,2	167,3	165,3	150,6
31	53,0	42,3	51,6	39,0	52,2	41,7	58,0	39,0	58,0	39,0
32	4,0	8,9	12,5	3,8	11,6	3,4	1,5	4,0	1,5	4,0
33	91,9	32,1	95,3	30,5	105,2	32,8	93,8	35,4	93,8	35,4
34	5,0	6,4	10,9	7,5	10,7	11,7	7,1	3,9	7,1	4,1
35	9,7	10,5	12,2	14,1	4,6	9,6	10,1	7,9	9,0	7,9
36	13,7	1,7	17,5	4,0	26,5	12,7	7,5	5,8	7,5	5,8
41	23,7	16,7	18,1	12,4	28,3	19,2	13,7	12,4	4,0	5,5
42	34,0	90,3	136,7	177,7	47,3	79,6	52,9	34,6	60,2	34,1
45	12,3	28,9	16,4	39,1	11,6	32,9	26,4	15,0	26,4	15,0
46	33,7	11,4	28,7	11,2	17,4	6,8	24,0	20,1	24,0	20,1

продолжение таблицы F.1

47	31,9	28,3	20,2	17,5	64,0	50,4	12,7	13,3	12,7	13,3
48	12,0	14,2	23,7	8,1	40,2	12,5	16,4	14,2	16,4	14,2
50	5,9	9,3	7,6	11,0	13,7	5,3	5,7	9,4	35,3	9,7
55	6,3	7,5	1,6	2,8	6,3	6,9	10,8	13,9	10,8	13,9
56	36,4	20,9	36,0	19,7	35,8	20,3	54,6	134,0	53,8	134,7
57	12,8	15,4	12,9	11,8	12,8	17,4	15,4	13,6	15,4	13,6
58	17,3	3,6	8,0	19,6	4,6	13,0	19,6	18,9	21,0	13,4
59	68,3	28,2	66,8	27,0	72,6	35,1	72,3	37,1	72,3	37,1
60	15,6	17,2	12,0	18,2	9,5	10,8	15,4	7,6	15,4	7,6
63	20,9	33,9	20,1	37,2	13,9	45,2	15,0	8,4	15,0	8,4
64	82,6	44,1	89,5	39,5	66,1	40,2	196,6	31,7	150,4	32,5
65	46,4	65,2	38,6	53,1	40,1	49,0	726,5	59,5	491,8	62,1
67	19,8	23,5	23,6	26,6	25,2	19,5	68,8	20,9	39,6	22,5
68	29,3	27,1	23,1	19,8	28,8	22,5	26,5	21,1	24,6	23,0
69	57,0	78,3	123,5	135,9	60,2	83,7	62,4	78,2	62,4	78,2
70	29,3	36,4	76,3	43,5	44,5	41,1	55,7	52,6	55,7	52,6
71	46,2	32,5	50,7	44,5	39,4	37,0	40,0	69,8	40,0	69,8
72	25,3	22,0	22,0	20,8	24,8	21,4	194,2	63,1	30,7	53,6
73	10,0	19,0	16,5	22,2	9,6	15,9	21,5	23,8	10,8	18,1
74	22,9	22,3	23,8	26,6	22,2	22,7	38,0	22,6	38,0	22,6
75	37,9	68,2	33,7	43,6	42,8	79,1	38,2	56,8	38,2	56,8
76	57,2	110,5	40,8	53,4	38,7	67,1	172,2	124,5	168,5	120,4
77	3,1	0,9	12,0	12,9	2,2	2,7	16,6	6,7	0,8	7,1
78	17,6	11,6	6,6	8,0	5,2	7,5	40,5	7,1	40,5	7,1
79	58,7	28,3	284,8	98,2	47,3	25,1	117,6	24,1	117,6	24,1
80	29,2	24,4	29,8	21,4	27,0	27,8	22,1	20,8	22,1	20,8
81	2073,4	139,1	1933,0	40,0	2717,4	211,1	66,2	588,3	66,2	588,3
82	28,0	30,2	30,5	59,0	24,2	34,2	>10000	46,3	750,6	58,8
83	38,7	25,1	19,0	22,2	15,6	25,4	25,5	41,6	15,4	44,5
84	8,4	5,8	10,2	6,6	5,4	10,9	28,8	6,0	28,8	6,0

продолжение таблицы F.1

85	9,5	16,6	8,1	16,1	4,7	15,9	12,3	11,2	12,3	11,2
86	485,6	486,7	143,3	131,2	572,5	511,5	575,0	522,8	575,0	522,8
87	8,1	5,1	6,2	10,0	0,6	3,4	3,2	14,2	3,2	14,2
88	16,7	7,0	29,4	12,0	23,3	11,0	17,4	6,7	17,4	6,7
89	267,4	39,3	361,5	183,9	267,7	61,5	>10000	28,9	151,5	28,6
90	53,7	26,3	51,7	49,6	47,3	26,8	136,1	19,8	136,1	19,8
91	272,5	970,9	250,9	462,7	57,7	896,4	27,8	36,1	27,8	36,1
92	29,8	16,3	25,2	13,4	18,3	13,9	5,7	10,3	5,7	10,3
93	8,1	18,5	33,0	21,0	32,3	15,7	36,7	20,6	36,7	20,6
94	14,6	18,6	11,9	6,9	12,3	12,0	18,8	8,9	18,8	8,9
95	16,0	20,1	14,1	10,3	17,1	22,5	12,9	18,7	12,9	18,7
96	56,1	20,7	72,6	33,6	64,3	21,4	46,5	13,9	46,5	13,9
97	42,4	42,9	62,6	35,8	46,9	55,2	120,2	35,5	195,6	40,9
98	86,4	40,4	96,7	10,1	107,7	39,7	197,6	5,3	177,3	28,3
99	7,7	6,7	9,8	9,6	12,6	27,0	21,7	18,0	21,7	18,0
100	17,8	50,4	26,0	58,3	29,4	77,8	17,6	29,0	21,3	27,0
101	32,0	23,5	44,6	26,6	35,1	25,8	29,9	25,7	30,1	25,9
102	1,9	10,7	4,2	15,4	1,0	11,5	6,3	11,1	6,7	9,9
103	33,8	19,2	12,6	7,6	7,7	10,4	19,7	11,8	19,7	11,8
104	21,8	37,6	44,1	43,2	68,4	40,2	36,1	34,9	36,1	34,9
105	23,0	27,3	19,6	29,2	22,0	27,9	22,4	47,2	22,4	47,2
106	5,1	10,3	5,2	9,0	6,0	10,0	4,7	10,8	4,7	10,8
107	4,6	12,4	41,6	19,5	45,6	14,2	26,3	33,7	26,5	33,9
108	14,3	13,3	4,3	4,4	8,6	7,5	48,1	9,5	46,8	9,7
109	28,3	29,5	39,5	30,6	24,3	28,9	18,2	33,6	36,7	15,3
110	17,1	16,2	16,8	16,3	21,8	15,2	21,6	15,0	21,6	15,0
111	22,9	21,3	30,7	52,7	22,2	24,7	108,9	50,7	108,9	50,7
112	13,4	12,1	4,4	4,5	7,6	13,2	28,5	21,0	24,7	5,3
113	447,5	484,1	35,2	84,5	451,8	462,2	509,3	525,4	409,9	462,9
114	68,1	27,2	46,8	15,0	93,9	40,5	57,2	17,5	57,2	17,5

Таблица F.2 – RMSE для моделей, прогнозирующих объем добычи нефти на 30 дней горизонта прогноза

Скважина	Экзогенная переменная - пластовое давление		Экзогенная переменная - разница давлений		Экзогенная переменная - пластовое давление на новом кусте		Изменения отсечения начала обучающей выборки (максимум 1 мая) по ДСММ		Изменения отсечения начала обучающей выборки (максимум 1 апреля) по ДСММ	
	Начало	Середина	Начало	Середина	Начало	Середина	Начало	Середина	Начало	Середина
2	0,1	0,3	0,1	0,3	0,1	0,3	0,1	0,3	0,1	0,3
3	12,3	6,8	7,8	7,6	12,5	7,3	73,8	11,8	17,4	13,4
4	2,5	2,8	3,3	3,5	3,2	3,3	5,1	2,4	5,1	2,4
9	0,4	0,3	0,7	0,1	0,5	0,2	0,3	0,1	0,3	0,1
10	0,9	1,6	1,4	1,8	1,1	1,6	2,8	1,5	2,8	1,5
11	3,5	5,2	13,4	2,9	11,9	3,9	4,5	13,6	19,6	14,6
17	11,9	8,3	7,9	6,0	10,1	7,1	11,7	9,4	11,7	9,4
19	3,9	5,6	1,9	3,8	2,1	3,1	6,2	2,7	6,2	2,7
20	0,3	1,7	0,7	2,0	0,3	1,9	0,4	1,4	0,4	1,4
21	6,0	4,7	5,4	4,4	6,8	5,1	6,0	5,5	6,0	5,5
22	7,2	7,3	6,4	7,1	7,4	7,2	7,9	7,0	7,9	7,0
24	6,3	4,5	5,2	3,7	5,4	3,9	3,6	5,2	3,6	4,5
31	4,3	5,9	4,2	5,4	4,3	5,8	4,8	5,4	4,8	5,4
32	1,5	3,5	4,2	2,7	3,9	2,5	0,7	2,6	0,7	2,6
33	9,8	4,1	10,2	3,9	11,2	4,0	10,0	4,3	10,0	4,3
34	0,6	0,8	1,2	1,0	1,3	1,4	0,7	0,7	0,8	0,7
35	1,6	1,8	2,1	2,2	0,7	1,4	1,7	1,5	1,6	1,5
36	3,3	0,5	4,2	1,1	6,5	3,2	1,8	1,4	1,8	1,4
41	2,7	1,9	2,1	1,4	3,3	2,2	1,6	1,7	0,5	0,8
42	2,3	4,8	8,8	8,8	3,2	4,1	3,4	1,8	3,9	1,7
45	3,8	8,4	5,3	11,0	3,6	9,6	9,7	4,3	9,7	4,3
46	4,1	1,9	3,6	2,0	2,6	1,1	2,9	3,3	2,9	3,3
47	7,9	6,5	5,1	4,5	15,5	11,5	3,5	3,1	3,5	3,1

продолжение таблицы F.2

48	2,3	2,9	4,0	1,8	6,8	2,6	2,6	2,8	2,6	2,8
50	1,1	2,0	1,4	2,4	2,5	1,2	1,2	2,1	6,5	1,9
55	1,2	1,8	0,3	0,8	1,2	1,4	2,1	2,7	2,1	2,7
56	5,6	3,6	5,5	3,6	5,5	3,6	8,3	17,7	8,2	17,8
57	2,3	3,2	2,3	2,7	2,6	3,3	2,9	3,5	2,9	3,5
58	6,8	1,8	4,1	7,7	2,2	5,0	7,9	7,3	8,3	5,3
59	2,9	2,1	2,8	2,3	3,0	2,4	3,0	2,5	3,0	2,5
60	7,9	9,0	9,9	12,5	6,6	8,7	10,5	8,3	10,5	8,3
63	1,0	1,3	0,9	1,4	0,7	1,7	0,7	0,4	0,7	0,4
64	0,7	0,3	0,7	0,3	0,5	0,3	3,0	0,2	1,5	0,3
65	4,6	5,5	4,8	4,4	5,0	4,2	111,3	5,3	62,8	5,3
67	4,1	4,1	4,3	4,5	4,8	4,1	19,1	4,2	9,2	4,3
68	4,4	3,8	3,5	3,3	4,1	3,3	4,3	3,5	4,1	3,6
69	1,8	2,1	3,8	3,7	1,9	2,2	1,9	2,1	1,9	2,1
70	2,6	3,0	6,2	3,3	3,4	3,1	4,0	3,9	4,0	3,9
71	6,3	3,2	5,4	4,3	3,9	3,6	5,0	6,7	5,0	6,7
72	1,9	1,7	1,6	1,7	1,8	1,7	25,9	4,8	2,1	3,7
73	1,7	2,7	2,5	3,1	1,5	2,2	3,3	3,4	1,9	2,5
74	3,3	3,3	3,1	3,2	3,1	3,5	6,1	3,5	6,1	3,5
75	7,5	11,6	6,6	8,6	8,5	13,7	7,8	9,4	7,8	9,4
76	2,7	4,4	3,1	2,3	2,8	2,7	9,3	5,1	9,1	4,9
77	0,4	0,1	1,8	1,9	0,3	0,4	2,5	1,0	0,1	1,1
78	1,3	1,3	0,5	0,7	0,5	0,6	3,1	0,7	3,1	0,7
79	0,9	0,3	3,3	1,4	0,5	0,3	1,5	0,3	1,5	0,3
80	3,4	3,0	3,4	2,8	3,2	3,1	2,5	2,5	2,5	2,5
81	3,6	4,0	2,4	1,1	4,6	6,1	3,9	24,1	3,9	24,1
82	1,7	2,7	3,0	5,3	1,7	3,1	>10000	4,8	87,2	5,3
83	4,2	2,0	2,0	1,8	1,7	2,0	2,8	3,1	1,9	3,4
84	1,1	0,8	1,3	0,9	1,0	1,5	3,7	1,0	3,7	1,0
85	0,8	1,6	0,7	1,5	0,5	1,3	1,3	1,3	1,3	1,3



продолжение таблицы F.2

86	4,3	4,2	3,1	1,8	4,9	3,9	5,0	4,2	5,0	4,2
87	1,6	1,0	1,3	2,0	0,2	0,7	0,6	3,6	0,6	3,6
88	2,3	1,1	4,0	1,8	3,3	1,5	2,3	1,3	2,3	1,3
89	2,8	3,0	8,7	11,3	1,6	4,0	>10000	2,0	7,1	2,0
90	5,1	2,0	5,3	3,9	4,6	2,0	11,2	1,5	11,2	1,5
91	0,8	3,3	0,7	1,5	0,2	3,1	0,1	0,1	0,1	0,1
92	1,9	1,2	1,7	1,0	1,2	1,1	0,5	0,9	0,5	0,9
93	1,2	3,8	3,9	2,9	3,7	2,5	4,1	2,7	4,1	2,7
94	3,7	4,8	4,1	1,6	3,5	2,7	4,9	2,4	4,9	2,4
95	3,0	3,7	2,9	2,6	3,3	4,2	2,3	3,5	2,3	3,5
96	6,9	2,7	9,0	4,4	8,0	2,8	5,7	1,9	5,7	1,9
97	0,9	2,8	1,2	2,5	1,0	3,3	2,3	2,6	4,1	2,7
98	9,8	13,1	20,1	3,1	12,0	12,9	30,1	1,7	42,2	8,9
99	1,5	1,3	2,2	1,9	2,6	5,2	4,3	3,5	4,3	3,5
100	1,3	2,9	1,9	3,3	2,2	4,4	1,4	1,8	1,5	1,7
101	0,4	0,5	0,6	0,6	0,5	0,6	0,4	0,7	0,4	0,7
102	0,4	2,5	0,8	3,1	0,2	2,6	1,2	2,6	1,2	2,3
103	3,5	2,0	1,5	1,1	0,9	1,1	2,1	1,4	2,1	1,4
104	2,2	5,8	4,2	5,0	6,7	4,3	3,6	5,7	3,6	5,7
105	1,5	3,5	1,4	3,7	1,4	3,5	2,4	5,7	2,4	5,7
106	1,9	3,5	1,9	3,2	2,1	3,4	1,4	3,7	1,4	3,7
107	1,0	3,2	6,5	3,3	7,0	2,4	4,0	7,5	4,0	7,5
108	3,5	3,2	1,2	1,2	2,1	1,8	11,1	2,4	10,8	2,4
109	3,4	3,4	4,6	3,7	3,1	3,3	2,3	4,4	4,2	2,9
110	2,4	2,2	2,3	2,2	3,2	2,0	3,2	2,0	3,2	2,0
111	2,6	3,2	4,3	8,1	2,5	3,7	12,7	7,6	12,7	7,6
112	1,5	1,3	0,5	0,6	0,9	1,4	3,0	3,1	2,6	0,7
113	3,4	3,8	3,0	4,2	3,4	3,6	4,1	4,2	3,5	3,6
114	3,9	1,9	2,7	2,0	5,4	2,7	3,2	2,2	3,2	2,2

Таблица F.3 – MAPE (%) для моделей, прогнозирующих дебит жидкости на 30-60-90 дней горизонта прогноза

Скважина	Модели месторождения 1 с пластовым давлением						Модели месторождения 1 с разницей давлений (Formation - Bottom Hole pressures)					
	Начало (30 дн)	Середина (30 дн)	Начало (60 дн)	Середина (60 дн)	Начало (90 дн)	Середина (90 дн)	Начало (30 дн)	Середина (30 дн)	Начало (60 дн)	Середина (60 дн)	Начало (90 дн)	Середина (90 дн)
2	5,9	7,8	3,4	6,1	7,0	8,4	4,8	4,7	5,2	6,0	10,1	9,7
3	142,2	35,4	56,5	22,6	35,8	20,0	80,7	18,3	32,2	9,4	20,5	7,5
4	4,4	4,0	5,0	4,8	5,4	5,1	1,8	2,1	2,4	2,8	2,9	3,4
9	9,8	4,6	9,2	5,7	10,0	7,2	36,2	26,6	49,9	40,9	63,1	54,3
10	2,8	1,9	1,9	1,6	2,3	3,8	1,8	3,0	2,8	3,0	3,4	5,4
11	19,4	18,6	11,9	12,1	12,1	10,4	22,2	23,3	25,7	26,8	26,6	24,4
17	1,0	1,2	0,8	2,1	1,6	2,5	4,5	4,4	4,1	3,9	3,7	3,7
19	1,9	2,4	3,0	4,0	3,7	3,9	8,3	7,5	6,7	5,8	5,5	4,9
20	1,7	1,6	2,4	1,8	2,7	4,0	3,2	2,5	4,7	3,1	5,3	5,1
21	2,7	2,2	2,1	1,5	1,6	1,3	1,6	1,8	3,0	2,7	3,5	2,9
22	24,2	23,3	28,3	27,6	30,5	29,0	23,1	24,1	27,3	28,7	30,7	28,8
24	1,1	1,8	2,2	3,6	4,1	5,0	2,3	1,5	1,6	1,9	2,7	3,4
31	3,3	5,5	5,5	5,8	5,9	5,8	4,0	4,3	3,6	3,8	4,4	5,6
32	17,2	15,3	19,2	20,2	19,4	22,3	6,5	8,9	14,5	19,0	15,5	16,8
33	11,6	12,0	12,5	9,6	9,5	7,8	12,9	12,1	14,3	12,3	13,9	11,0
34	4,1	2,9	4,2	5,4	7,5	9,9	4,4	3,2	3,7	3,7	3,9	4,3
35	1,8	2,6	3,3	5,2	6,1	6,6	1,3	2,2	2,5	4,2	4,6	5,4
36	1,3	1,4	1,7	1,6	1,5	1,4	4,3	4,2	4,8	4,3	4,5	4,2
41	8,2	8,8	11,6	12,5	15,4	16,2	0,8	1,1	0,6	1,4	1,3	2,5
42	132,2	117,1	192,1	180,9	220,5	199,0	183,6	147,3	308,5	246,8	391,1	312,9
45	2,1	6,3	6,1	8,9	8,2	11,7	1,2	2,6	1,6	2,3	1,8	2,6
46	3,4	1,1	2,5	1,5	2,2	1,9	3,2	0,5	1,8	0,6	1,3	0,6
47	0,6	1,3	0,8	2,0	1,7	2,7	12,8	10,8	13,7	11,9	14,4	12,9

продолжение таблицы F.3

48	5,2	3,4	4,9	4,5	6,0	4,7	2,0	1,1	1,9	2,1	2,9	3,0
50	2,1	3,3	4,9	6,8	6,7	7,9	6,0	2,8	4,0	1,7	3,0	1,5
55	4,7	3,2	4,2	3,7	3,8	3,9	4,0	3,0	4,2	3,8	4,1	4,4
56	1,6	2,9	1,2	2,3	1,0	2,1	1,6	3,0	1,3	2,7	1,2	2,7
57	0,4	0,7	0,6	0,6	0,5	0,7	3,8	2,9	5,0	4,0	4,5	3,3
58	0,9	0,5	1,0	0,8	1,0	1,0	1,5	1,3	1,8	2,1	2,5	2,8
59	9,4	8,6	12,6	11,5	14,4	13,9	5,0	5,6	7,2	7,4	8,1	7,6
60	8,6	9,5	12,1	12,1	14,3	13,5	8,1	8,8	7,0	7,7	7,3	7,5
63	4,9	2,5	4,2	3,1	3,5	5,8	3,4	3,3	2,7	2,4	2,8	2,5
64	21,5	19,7	11,1	13,4	9,5	15,6	23,1	18,6	12,1	12,2	11,0	14,3
65	65,0	2,5	55,9	4,3	49,1	6,6	27,7	1,8	16,0	2,3	11,6	1,9
67	3,4	4,4	2,9	4,3	2,3	4,4	6,6	5,3	5,1	4,1	4,3	4,0
68	17,6	12,3	15,2	9,8	11,7	7,5	15,2	10,9	11,9	8,2	8,8	6,8
69	9,6	7,2	14,0	11,4	16,6	13,7	12,1	10,4	19,4	17,3	25,0	22,1
70	37,0	29,8	31,5	18,4	28,2	15,7	25,1	27,9	17,2	17,7	14,0	14,6
71	3,2	5,5	3,1	7,4	2,3	7,6	3,3	4,1	3,4	2,5	3,6	2,3
72	3,1	3,8	5,5	5,2	7,8	7,6	1,1	2,0	2,1	3,0	4,0	5,2
73	1,5	1,6	1,9	2,3	2,8	4,3	2,3	1,2	2,5	1,6	2,3	1,8
74	3,3	3,9	5,2	5,7	6,6	6,5	2,7	1,3	2,4	1,4	1,9	1,0
75	49,9	39,0	54,2	44,2	58,4	45,4	37,5	24,2	37,8	30,3	41,9	33,9
76	4,3	2,7	5,4	2,2	4,7	2,2	7,6	3,0	7,7	2,9	7,3	3,3
77	5,3	5,5	7,6	7,5	7,8	6,8	7,4	7,8	8,9	8,0	7,5	6,4
78	1,6	0,8	2,4	1,4	2,3	1,2	0,6	0,9	0,8	1,1	1,2	1,8
79	21,9	21,6	18,1	30,8	65,3	83,4	34,4	33,5	46,7	33,5	66,9	40,2
80	1,0	0,8	1,1	0,7	1,2	0,8	1,0	0,8	1,2	0,9	1,8	1,2
81	513,3	101,3	300,1	121,6	244,8	128,0	432,8	37,0	231,2	43,3	167,9	42,0
82	2,9	1,7	2,6	1,5	2,2	2,0	10,4	9,7	12,6	12,1	17,1	15,5
83	7,9	18,7	15,7	50,9	27,5	32,4	6,7	16,7	14,6	45,6	26,2	26,5
84	6,6	3,7	9,5	5,9	13,0	8,9	7,6	4,8	10,9	7,4	13,8	9,6
85	5,1	5,3	6,8	6,8	8,4	11,0	9,4	9,5	12,9	10,7	10,4	7,5

продолжение таблицы F.3

86	20,2	18,2	15,7	13,3	10,9	9,0	7,5	14,7	12,4	17,5	16,0	18,6
87	12,5	11,9	15,6	13,2	13,5	10,6	2,8	2,3	2,8	2,3	2,8	2,4
88	1,0	1,0	1,6	2,1	1,9	2,1	1,4	1,2	1,4	2,3	2,0	2,0
89	115,4	28,5	87,4	51,7	65,7	41,7	258,9	208,8	256,6	219,8	239,8	204,8
90	34,6	12,3	19,3	15,2	19,6	20,4	29,4	26,7	25,0	23,8	22,8	19,4
91	19,6	149,0	297,2	430,6	291,8	426,0	76,7	92,0	148,4	140,2	140,5	134,5
92	3,8	4,2	5,6	5,2	6,3	5,9	4,0	6,0	5,1	7,8	8,8	11,3
93	10,3	8,6	8,9	6,3	6,3	4,5	11,4	7,1	6,2	4,0	4,9	4,2
94	13,6	9,1	12,6	10,3	14,6	13,1	6,7	2,5	4,5	2,3	4,4	3,4
95	7,1	5,4	8,7	6,3	7,5	5,0	0,7	1,0	0,9	1,3	1,1	1,8
96	0,6	0,7	1,0	0,8	1,0	0,9	0,9	1,0	0,9	1,4	1,5	1,8
97	32,1	37,9	47,5	45,7	53,4	50,6	13,3	27,0	24,7	30,4	28,1	31,6
98	104,4	28,1	105,9	41,1	110,2	48,1	71,1	18,2	62,4	24,7	57,8	25,7
99	2,3	1,4	1,7	1,3	1,9	1,5	2,5	1,3	1,6	1,0	1,3	0,8
100	0,8	3,7	5,1	9,9	11,8	18,1	11,1	11,4	19,6	19,4	26,9	26,7
101	7,2	4,6	7,8	4,0	8,3	5,9	7,6	4,5	8,5	4,0	8,9	6,0
102	2,5	2,6	3,2	2,9	3,2	2,6	5,9	4,5	5,9	4,7	5,6	4,0
103	13,7	7,4	11,5	7,5	11,5	8,2	1,5	5,0	4,1	5,8	5,4	6,7
104	1,1	2,2	1,6	2,9	3,1	5,0	18,5	16,5	20,8	19,6	22,7	21,9
105	0,7	2,6	2,4	3,9	3,6	5,4	1,5	2,2	1,8	2,7	2,7	3,0
106	2,6	1,4	2,2	2,2	1,7	2,8	2,3	1,4	2,1	1,7	1,7	1,8
107	3,1	3,4	4,5	3,9	5,3	4,3	2,1	2,7	3,5	3,2	4,2	3,7
108	6,1	4,1	9,1	6,3	9,3	6,7	2,2	2,8	3,8	4,3	4,9	4,7
109	13,2	10,5	17,7	16,5	22,6	29,0	1,5	0,6	1,3	1,9	4,2	14,4
110	11,1	8,0	13,4	11,1	15,1	13,2	22,1	16,7	26,7	22,1	30,0	24,6
111	13,8	13,3	15,5	14,2	15,6	14,5	23,7	40,4	25,7	24,9	21,0	20,8
112	4,7	3,4	6,7	6,2	8,5	7,5	6,2	7,4	9,5	9,8	10,9	10,6
113	13,2	16,1	12,9	16,9	14,5	16,5	23,0	23,9	15,1	12,7	10,0	8,7
114	7,9	5,4	7,8	5,1	7,3	4,7	4,3	6,2	5,4	7,7	7,3	9,3

Таблица F.4 – MAPE (%) для моделей, прогнозирующих дебит жидкости на 30-60-90 дней горизонта прогноза

Скважина	Нагнетательные модели (лаг 7)				Нагнетательные модели (лаг 28)		Дифференцирование	
	Начало (Пласт, давление)	Середина (Пласт, давление)	Начало (Разница давлений)	Середина (Разница давлений)	Начало (Пласт, давление)	Середина (Пласт, давление)	Начало (Пласт, давление)	Середина (Пласт, давление)
2	14,3	16,9	20,2	23,7	15,8	17,0	10,2	9,3
3	37,2	21,9	20,3	7,6	37,2	27,2	34,5	35,4
4	4,4	4,6	2,9	3,3	3,9	4,0	1,4	2,7
9	6,3	7,0	63,3	53,7	4,1	2,1	7,4	3,6
10	2,8	4,2	2,1	3,3	2,9	3,9	2,8	4,9
11	11,5	10,2	26,2	26,1	14,8	24,1	29,5	104,4
17	1,5	2,1	1,4	1,9	1,9	3,1	2,3	5,6
19	4,0	3,5	6,7	5,5	3,9	3,2	7,4	4,1
20	4,6	4,6	7,9	6,1	4,7	4,0	1,1	2,1
21	3,2	2,3	3,3	2,7	8,6	1,4	5,5	>10000
22	29,9	29,0	30,6	28,7	31,7	30,3	3,1	107,3
24	5,4	5,9	5,0	5,0	4,2	4,7	3,9	2,8
31	7,0	6,4	4,1	4,9	6,3	5,9	4,8	2,4
32	27,6	15,5	15,8	14,8	42,8	17,7	6,1	7,5
33	12,8	10,9	13,0	10,0	10,9	8,4	9,5	7,8
34	7,1	8,2	4,5	4,0	6,0	6,0	1,2	2,2
35	6,4	7,5	4,9	5,9	7,1	8,4	3,1	4,5
36	2,4	2,5	4,4	4,1	1,4	1,4	1,2	1,8
41	14,5	15,6	1,5	2,0	10,3	11,1	0,7	0,6
42	208,4	185,5	357,6	276,6	193,8	193,4	38,1	25,0
45	8,2	12,4	1,7	2,7	5,5	7,7	4,5	2,5
46	2,6	1,9	1,3	0,6	2,1	1,5	3,0	1,9

продолжение таблицы F.4

47	1,6	2,5	14,7	13,1	2,6	4,3	0,9	0,7
48	8,1	6,2	2,4	2,4	1,6	1,8	1,3	1,7
50	6,5	6,9	2,3	2,0	8,0	8,7	2,9	1,9
55	3,6	3,5	3,9	3,9	4,5	4,8	10,0	5,1
56	1,1	2,2	1,3	2,8	1,4	1,2	1,7	1,7
57	1,8	1,6	4,9	3,6	0,9	0,9	6,2	6,5
58	1,2	1,5	2,4	2,7	1,3	1,5	2,2	0,9
59	15,5	14,6	9,2	8,4	7,0	6,9	2,9	1,6
60	14,0	13,2	7,1	7,2	12,0	7,7	6,0	4,7
63	2,8	4,5	2,8	3,0	13,8	2,7	4,3	27,1
64	10,2	16,2	15,7	20,9	10,0	18,0	33,7	15,3
65	80,5	9,1	17,2	8,7	82,5	12,9	83,9	5,0
67	2,3	3,4	4,5	4,1	4,5	4,5	12,1	1,9
68	12,3	7,8	9,4	6,9	10,3	5,8	15,0	7,5
69	17,7	15,8	24,1	24,9	19,5	17,1	8,3	3,7
70	29,1	17,8	15,2	16,9	34,6	11,7	7,1	106,2
71	2,3	7,4	3,2	2,4	3,3	5,3	2,2	2,9
72	7,5	7,6	3,3	4,5	9,0	10,3	8,5	2,7
73	14,8	10,8	12,6	6,6	12,2	8,5	5,1	5,2
74	5,6	5,5	1,8	1,1	6,4	6,7	1,9	>10000
75	57,7	45,4	42,3	34,0	56,6	43,4	50,1	40,5
76	4,1	2,0	5,7	2,2	4,3	3,0	5,1	5,1
77	9,5	7,7	6,3	5,7	6,2	6,5	3,3	3,5
78	1,7	1,6	1,2	2,0	2,7	1,3	1,2	>10000
79	54,1	77,4	100,5	56,9	70,6	100,9	31,0	>10000
80	0,7	0,8	0,8	0,7	0,5	1,9	1,8	1,4
81	204,3	77,2	170,3	45,2	130,1	22,6	116,9	150,8
82	2,3	1,9	20,4	17,1	2,5	2,6	1,6	1,5
83	27,4	32,1	26,4	26,5	26,8	33,0	28,6	34,1
84	6,9	4,4	11,1	8,1	4,9	4,4	3,3	>10000

продолжение таблицы F.4

85	9,3	11,1	9,6	7,3	3,8	8,1	2,2	2,2
86	10,5	8,1	15,5	18,2	7,6	7,5	13,8	9,7
87	11,8	8,5	2,2	1,6	13,6	10,9	4,8	2,8
88	2,3	2,0	2,5	2,0	2,0	1,8	3,4	3,9
89	39,0	35,4	233,1	198,2	35,0	23,2	121,0	97,6
90	23,0	9,7	23,3	18,1	28,9	18,4	118,3	10,5
91	249,9	348,9	85,6	73,1	95,5	168,5	34,1	29,5
92	6,6	6,7	9,9	12,2	5,7	6,6	3,1	2,9
93	6,2	4,4	5,0	4,3	5,1	3,7	6,7	5,9
94	13,5	12,6	4,5	5,4	17,8	15,7	5,5	5,0
95	5,6	3,6	2,4	2,1	2,5	2,2	0,7	1,7
96	1,6	1,4	1,4	1,8	1,4	1,3	1,7	1,9
97	51,5	48,2	24,7	28,3	57,7	55,7	18,3	39,2
98	78,2	37,1	57,1	27,0	75,3	24,7	96,7	65,9
99	2,1	1,6	1,5	0,9	1,7	1,4	3,1	3,8
100	22,4	15,1	34,1	26,3	4,7	4,1	2,5	4,1
101	6,2	6,6	7,0	6,4	5,9	7,0	11,8	7,6
102	3,3	2,4	6,3	4,5	4,8	4,0	3,8	2,4
103	10,9	8,3	5,5	6,8	13,6	11,0	7,2	7,8
104	3,7	6,1	22,3	21,4	3,5	5,5	2,6	2,1
105	3,5	5,9	2,7	3,3	4,9	8,2	1,1	3,1
106	1,4	3,6	1,0	2,7	1,3	3,7	1,4	1,7
107	5,2	4,3	3,5	3,1	5,3	4,1	3,1	3,3
108	4,3	4,8	5,1	5,0	7,1	5,7	4,2	6,2
109	21,6	28,3	4,5	15,2	22,3	28,5	2,9	4,3
110	19,1	17,0	33,4	27,4	13,5	12,1	6,2	6,0
111	16,4	15,1	23,0	22,6	16,5	14,3	15,7	89,8
112	6,8	6,0	7,4	7,3	4,9	4,1	7,2	3,4
113	13,6	15,7	9,9	8,7	13,6	14,8	22,3	21,2
114	7,4	4,7	8,0	10,3	16,3	6,3	5,9	103,7

Таблица F.5 – RMSE для моделей, прогнозирующих дебит жидкости на 30-60-90 дней горизонта прогноза

Скважина	Модели месторождения 1 с пластовым давлением						Модели месторождения 1 с разницей давлений (Formation - Bottom Hole pressures )					
	Начало (30 дн)	Середина (30 дн)	Начало (60 дн)	Середина (60 дн)	Начало (90 дн)	Середина (90 дн)	Начало (30 дн)	Середина (30 дн)	Начало (60 дн)	Середина (60 дн)	Начало (90 дн)	Середина (90 дн)
2	1,46	2,12	1,01	1,72	2,21	2,81	1,23	1,50	1,34	1,79	2,96	3,26
3	77,35	113,93	48,34	82,91	38,28	72,15	72,85	82,46	45,67	58,31	36,06	49,22
4	13,14	11,93	14,77	13,93	15,75	18,05	5,81	6,24	7,20	8,48	8,71	14,67
9	8,81	4,04	8,14	5,01	8,49	6,30	33,08	24,85	45,05	37,10	55,29	47,18
10	3,06	2,47	2,37	1,98	2,74	6,28	2,15	3,33	3,01	3,32	3,86	7,52
11	21,64	23,84	14,94	17,30	14,21	15,64	24,45	28,32	27,22	29,51	27,83	27,10
17	7,83	10,85	6,40	16,51	13,48	20,30	32,26	31,80	29,51	28,23	27,25	26,97
19	12,65	15,00	18,43	21,10	20,03	20,00	44,25	39,41	36,30	31,64	30,98	27,13
20	1,50	1,33	1,95	1,45	2,21	6,14	2,58	2,05	3,70	2,44	4,13	6,13
21	8,12	7,01	6,58	5,20	5,45	4,41	5,00	6,14	9,64	8,74	10,67	9,09
22	125,75	120,95	142,95	138,98	152,43	145,43	117,06	123,29	136,65	144,31	153,89	145,95
24	2,72	4,42	5,31	8,41	10,35	11,71	5,58	3,78	4,17	4,46	6,61	8,44
31	6,15	9,49	9,77	9,62	10,08	9,65	7,04	7,14	6,58	6,65	7,62	10,20
32	126,65	113,60	136,22	145,72	135,98	159,74	47,76	69,90	127,91	152,09	127,25	133,59
33	19,88	20,66	19,54	16,42	16,23	14,01	20,69	20,83	21,42	18,93	20,32	17,09
34	9,64	6,60	9,59	14,17	19,54	29,01	9,93	7,37	8,51	8,18	8,88	9,70
35	4,34	6,44	7,84	11,09	13,18	13,00	3,65	5,48	6,04	9,14	10,13	10,77
36	12,00	11,10	14,22	12,84	12,34	11,64	35,68	33,59	38,47	34,14	35,57	33,64
41	16,54	17,36	22,85	24,33	30,38	31,86	2,29	2,32	1,78	2,80	3,11	5,51
42	103,22	93,81	147,51	141,77	167,39	154,92	148,77	118,59	244,45	197,01	307,55	253,53
45	17,33	52,35	56,50	71,44	71,25	100,84	10,03	20,40	13,55	17,89	15,37	20,58
46	19,81	6,51	15,24	8,54	13,12	11,68	20,10	2,98	14,29	3,43	11,58	3,69
47	4,87	7,07	5,21	11,84	12,15	15,53	70,89	60,52	74,34	64,33	77,17	69,62



продолжение таблицы F.5

48	27,91	18,76	26,59	25,98	33,22	27,28	11,44	7,43	12,65	14,72	17,63	18,96
50	6,84	12,89	16,52	22,18	21,77	24,66	16,87	7,99	12,53	5,80	10,21	4,90
55	24,42	20,10	22,74	20,01	20,41	22,23	21,51	20,33	22,84	21,78	21,85	24,59
56	4,60	7,62	3,70	6,38	3,28	5,75	4,56	7,82	3,71	6,97	3,27	6,89
57	3,64	4,25	3,78	3,92	3,23	4,21	20,97	17,18	27,91	23,48	24,86	20,04
58	3,98	1,93	4,13	3,34	4,13	4,75	5,36	5,69	6,87	8,57	9,55	11,16
59	29,85	27,80	39,42	36,15	44,74	43,87	17,36	18,82	23,33	23,46	25,85	24,44
60	83,63	90,81	112,64	109,99	129,37	119,96	94,80	99,14	77,05	80,61	73,83	75,31
63	2,81	1,66	2,53	1,90	2,11	3,58	2,27	2,18	1,85	1,67	1,91	1,70
64	6,80	6,55	4,73	4,95	4,13	5,18	6,90	6,54	4,80	4,84	4,35	5,17
65	123,55	5,84	107,91	9,50	96,43	15,19	73,71	4,30	52,73	5,20	43,10	4,66
67	9,40	11,77	7,81	10,77	6,45	10,76	17,15	14,59	13,70	11,20	11,61	10,57
68	5,00	3,46	4,24	2,86	3,44	2,34	4,31	3,13	3,45	2,52	2,82	2,12
69	4,77	3,73	6,69	5,49	7,61	6,32	6,01	5,20	9,32	8,26	11,66	10,20
70	39,02	32,46	36,89	23,78	34,38	21,02	25,59	30,54	20,14	22,69	17,48	19,58
71	13,68	12,26	10,92	16,13	9,06	16,31	11,94	9,69	9,97	7,07	9,37	6,37
72	1,56	2,06	2,81	2,68	4,03	3,82	0,59	1,23	1,19	1,77	2,45	2,88
73	2,20	2,23	2,93	3,23	4,70	6,85	3,19	1,67	3,45	2,08	3,48	2,59
74	8,40	10,12	13,08	14,16	16,35	15,84	6,32	3,67	5,94	3,73	5,10	2,91
75	267,91	208,51	277,52	220,34	287,82	221,00	207,36	129,81	197,74	151,70	209,34	165,45
76	9,53	5,41	11,15	4,77	9,61	4,43	14,55	5,98	14,60	5,67	13,63	6,36
77	28,65	28,66	40,01	38,63	39,55	34,25	36,74	38,95	43,73	40,33	38,70	34,56
78	3,39	1,95	5,05	3,35	5,08	3,05	1,63	2,24	2,20	2,51	3,06	4,55
79	9,57	8,90	8,57	16,79	37,18	42,74	15,28	12,44	21,91	13,81	30,08	17,29
80	3,45	3,00	3,94	2,72	4,03	3,00	3,29	3,05	4,10	3,24	6,19	4,13
81	86,52	147,28	150,02	196,41	179,77	206,17	68,30	49,94	78,35	67,29	79,27	66,52
82	7,57	4,29	6,76	4,00	5,89	5,73	26,16	23,14	30,05	28,82	42,22	38,05
83	12,09	23,47	19,09	31,73	21,26	25,33	10,49	22,23	18,46	29,08	20,73	21,82
84	15,59	8,62	22,22	14,34	30,90	21,73	18,00	11,35	25,48	17,69	32,13	22,50
85	10,29	9,98	13,32	13,32	17,03	26,33	18,76	18,23	25,25	20,70	21,54	16,37

продолжение таблицы F.5

86	84,17	76,94	69,68	62,68	56,87	49,62	41,62	70,95	60,49	79,97	79,42	87,77
87	38,17	35,01	45,10	37,88	41,57	33,51	8,45	6,77	7,94	6,48	8,80	7,30
88	5,16	5,18	7,46	11,16	9,80	10,84	6,69	5,51	6,29	12,34	11,01	11,04
89	37,45	64,10	83,91	96,46	75,79	82,95	298,54	327,43	350,19	340,19	337,06	315,71
90	97,83	32,53	68,91	38,34	62,65	52,41	88,37	74,96	73,50	65,00	67,92	55,64
91	2,21	17,05	20,25	25,13	19,89	24,86	8,41	8,20	10,35	8,37	9,87	8,09
92	8,74	9,65	13,23	11,78	13,99	12,86	11,45	14,43	12,99	18,07	21,69	25,80
93	66,75	55,04	56,94	42,86	46,44	34,78	83,57	58,09	57,72	40,52	47,54	35,93
94	45,68	30,63	42,05	34,62	49,06	44,66	27,13	9,67	19,85	8,56	19,14	15,13
95	39,08	29,71	46,66	34,95	41,23	29,57	5,84	8,24	7,07	9,19	7,81	11,44
96	2,92	3,51	4,43	4,04	4,73	4,41	4,18	4,83	4,34	7,78	7,87	9,25
97	45,92	52,94	64,47	59,16	68,59	63,43	25,25	38,13	36,64	39,67	38,05	39,82
98	585,99	180,98	608,34	258,68	638,85	298,83	398,54	113,19	359,10	152,23	336,41	155,99
99	11,89	7,30	9,48	6,85	10,01	8,09	12,76	7,99	9,77	6,14	8,49	5,80
100	1,47	6,70	9,45	16,65	20,97	30,07	17,17	17,42	29,38	29,23	40,02	40,05
101	1,37	1,02	1,46	0,95	1,75	1,52	1,44	0,92	1,58	0,89	1,84	1,49
102	12,00	16,12	16,56	15,76	16,18	14,12	28,77	22,49	28,56	22,97	26,61	20,51
103	31,05	17,01	26,55	17,11	26,69	18,84	5,56	11,77	11,81	13,47	13,92	15,66
104	5,06	9,67	8,03	15,43	16,22	29,86	76,60	69,77	85,47	81,95	93,00	92,19
105	1,83	7,39	7,33	10,21	10,19	15,15	3,81	5,53	4,55	7,20	7,63	8,07
106	12,29	6,81	10,45	11,09	8,80	13,67	10,84	6,61	9,63	8,54	8,53	8,93
107	18,53	17,63	24,61	19,97	27,96	21,96	13,80	14,06	19,40	16,56	22,58	18,68
108	38,94	27,64	55,34	38,02	54,84	39,95	14,57	18,58	24,53	26,84	30,82	29,58
109	34,35	28,16	45,31	42,81	57,00	69,40	3,95	1,64	4,20	6,50	17,61	48,04
110	47,31	34,40	55,59	47,05	61,97	54,67	91,84	69,77	109,49	92,22	122,25	100,55
111	53,89	53,26	60,49	55,81	60,65	56,56	111,02	168,24	121,12	119,20	101,89	100,89
112	8,05	5,90	11,11	10,44	13,85	12,26	10,95	13,05	15,76	15,72	17,26	16,55
113	35,95	37,40	30,98	34,11	31,26	32,42	54,85	57,30	41,19	37,51	32,74	29,84
114	23,02	15,25	22,04	14,23	20,82	13,28	17,27	20,06	17,79	23,04	23,08	27,42

продолжение таблицы F.5

Скважина	Нагнетательные модели (лаг 7)				Нагнетательные модели (лаг 28)		Дифференцирование	
	Начало (Пласт, давление)	Середина (Пласт, давление)	Начало (Разница давлений)	Середина (Разница давлений)	Начало (Пласт, давление)	Середина (Пласт, давление)	Начало (Пласт, давление)	Середина (Пласт, давление)
2	4,73	5,27	6,40	7,20	5,11	5,48	2,65	2,35
3	39,25	75,25	36,01	49,02	38,93	85,18	36,96	39,31
4	13,07	16,96	8,87	14,58	11,60	15,57	4,37	8,69
9	6,37	6,57	55,68	46,72	4,36	2,12	6,85	3,37
10	3,20	5,80	2,35	5,51	3,24	6,20	3,31	5,02
11	13,72	15,63	27,42	28,74	22,21	31,23	30,57	116,56
17	11,93	16,96	11,97	16,22	14,89	24,18	17,70	41,14
19	21,59	19,54	37,19	30,30	22,66	18,58	39,27	23,56
20	3,78	6,07	6,26	6,30	3,72	6,10	0,97	1,77
21	9,56	7,04	10,19	8,61	27,06	4,98	16,11	>10000
22	148,54	145,68	152,67	145,00	156,95	152,52	15,77	517,28
24	12,05	13,34	11,75	11,82	10,35	12,19	9,43	7,26
31	11,93	10,67	7,79	8,94	10,83	9,90	8,30	4,29
32	195,10	110,94	124,85	118,48	302,36	128,02	44,56	54,56
33	19,27	17,45	18,93	15,83	17,22	14,15	14,85	13,64
34	18,22	22,29	11,23	9,79	14,69	16,09	2,92	5,43
35	13,87	14,94	10,89	11,90	15,57	16,86	7,47	10,80
36	23,70	23,86	41,57	39,34	15,14	15,11	10,22	16,54
41	29,48	31,83	3,35	4,95	20,90	22,87	1,93	1,62
42	160,11	149,35	284,99	234,77	148,50	151,38	30,80	21,93
45	70,14	107,43	14,41	21,22	45,50	61,38	35,45	21,59
46	15,26	11,54	11,42	3,49	13,18	9,88	15,55	14,60

продолжение таблицы F.5

47	10,33	14,37	79,15	70,74	16,37	23,87	5,85	5,11
48	44,03	35,19	15,16	14,79	10,86	12,63	8,29	10,62
50	20,58	21,75	8,71	7,35	25,72	27,29	10,34	6,53
55	19,86	21,39	20,93	23,66	23,07	26,15	53,21	27,70
56	3,31	5,81	3,46	7,16	4,23	4,02	4,90	4,95
57	10,13	9,04	27,33	21,69	6,46	6,07	36,18	38,86
58	4,70	5,68	8,86	10,38	5,01	5,77	8,13	3,79
59	47,86	45,40	28,78	26,66	21,51	21,64	10,81	8,01
60	127,55	118,22	72,53	72,54	108,95	68,73	59,01	48,61
63	1,78	2,78	1,92	2,03	7,47	1,74	2,52	15,00
64	4,23	5,39	5,50	6,70	3,80	5,62	10,32	5,08
65	154,07	22,31	48,34	18,09	158,41	31,74	159,53	13,85
67	6,68	8,70	11,80	10,75	12,35	11,50	28,13	6,18
68	3,64	2,38	2,98	2,15	3,07	1,79	4,14	2,25
69	8,14	7,25	11,71	11,60	9,02	7,92	3,82	1,82
70	35,69	23,14	18,61	21,45	44,77	17,12	14,25	139,24
71	9,16	15,83	8,93	6,45	9,62	13,01	7,86	11,00
72	3,91	3,79	2,08	2,53	4,71	5,00	4,41	1,69
73	19,47	15,51	16,07	9,41	15,92	12,32	6,83	7,37
74	14,04	13,80	4,82	3,00	15,44	16,30	4,95	>10000
75	284,70	221,00	211,18	165,88	279,55	211,47	248,53	225,14
76	8,40	4,04	10,82	4,49	9,11	6,34	10,30	10,35
77	47,59	38,89	33,88	31,33	34,37	36,13	18,25	19,44
78	3,63	3,66	3,37	5,33	5,69	3,21	3,01	>10000
79	30,91	38,89	44,84	24,05	42,14	53,12	13,30	>10000
80	2,61	2,81	2,91	2,40	2,06	6,73	5,98	5,04
81	107,75	120,75	84,83	71,37	58,56	36,94	184,59	171,22
82	6,21	5,41	49,30	41,84	6,36	6,67	3,94	4,49
83	21,22	25,15	20,77	21,82	21,14	25,69	22,13	26,91
84	15,89	10,24	25,87	18,83	12,00	10,41	7,50	>10000

продолжение таблицы F.5

85	20,64	24,35	20,09	16,37	9,31	17,26	4,98	4,83
86	54,83	47,68	77,63	86,69	44,97	43,20	71,02	51,78
87	37,02	28,31	7,15	5,22	42,28	35,27	15,09	9,19
88	10,52	9,77	11,56	9,96	9,49	9,09	17,54	20,06
89	60,34	68,25	323,65	305,90	54,31	42,17	177,07	42,15
90	65,28	24,96	69,91	52,88	77,33	47,24	282,41	27,88
91	17,17	20,47	6,01	4,58	5,66	8,86	2,50	2,57
92	16,18	16,13	24,64	27,87	15,02	16,11	8,54	7,65
93	46,40	35,22	47,83	36,27	40,14	29,69	51,61	49,82
94	46,99	44,92	18,15	21,29	60,17	54,15	19,42	18,18
95	31,07	21,32	16,52	16,08	14,58	16,18	5,08	9,73
96	6,95	6,42	7,28	9,65	6,35	6,04	9,15	8,88
97	66,23	60,61	33,64	35,76	74,50	69,84	25,66	49,03
98	452,68	228,19	331,33	163,75	442,49	170,99	560,86	375,35
99	11,10	8,52	9,99	5,61	9,40	8,00	16,68	21,13
100	32,95	23,71	49,29	39,30	7,12	9,93	4,08	6,15
101	1,43	1,61	1,56	1,55	1,41	1,71	2,35	1,56
102	16,92	14,12	30,33	22,73	24,21	20,98	20,21	15,40
103	25,16	19,32	14,27	16,02	31,10	26,01	17,05	19,53
104	16,82	31,87	91,40	89,35	19,47	29,75	13,58	12,10
105	9,80	16,88	7,86	8,65	14,13	23,44	5,05	7,69
106	7,41	17,39	5,81	12,96	7,30	17,69	7,35	8,78
107	27,55	21,85	18,94	15,96	27,55	20,49	16,82	18,46
108	28,65	32,18	33,50	32,04	45,63	40,46	25,83	38,24
109	55,10	69,12	16,95	46,85	56,28	68,88	13,48	11,66
110	80,09	70,13	136,81	111,08	56,95	50,05	30,48	29,53
111	63,71	58,76	105,79	104,47	64,52	56,15	61,14	59,53
112	11,18	9,59	12,31	11,50	8,16	6,78	12,36	6,57
113	29,40	31,06	32,57	30,10	29,45	29,31	46,20	46,97
114	20,96	13,27	25,09	29,92	46,94	17,53	17,74	292,50

Таблица F6 – RMSE для моделей, прогнозирующих дебит жидкости на 5 месяцев горизонта прогноза по 461 кусту месторождения №2

Куст 461					
Скважина	Дата начала обучающей выборки	MAPE (Baseline)	MAPE (BVar) (учитываются только добывающие скважины)	MAPE (BVar) (добывающие скважины + с учетом разницы давлений)	MAPE (BVar) (добывающие скважины + с учетом разницы давлений + нагнетательные скважины)
503	1995-11	10,18	12,21	12,95	13,71
504	1995-11	4,73	14,80	14,21	15,24
505	1995-11	5,20	03,90	3,95	04,10
507	1995-05	14,11	10,02	10,53	10,54
509	1995-05	32,69	09,86	10,51	17,52
512	1995-04	02,49	3,73	3,74	-
528	1996-03	14,59	10,96	11,45	11,19
544	1995-11	15,83	10,72	10,65	10,66

# РОССИЙСКАЯ ФЕДЕРАЦИЯ



## СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2023683879

**Программа среднесрочного прогнозирования  
показателей продуктивности скважин с учетом влияния  
ее окружения**

Правообладатель: *Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Уфимский государственный нефтяной технический  
университет» (RU)*

Авторы: *Бахитов Ринат Радикович (RU), Попов Денис  
Владимирович (RU)*

Заявка № 2023683079

Дата поступления 31 октября 2023 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 10 ноября 2023 г.



*Руководитель Федеральной службы  
по интеллектуальной собственности*

*Ю.С. Зубов*