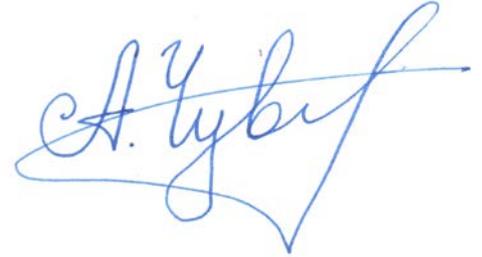


На правах рукописи



Чувашов Артур Александрович

**МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
В ТОПЛИВНО-ЭНЕРГЕТИЧЕСКОМ СЕКТОРЕ ПРОМЫШЛЕННОСТИ
ДЛЯ ОЦЕНКИ ДОСТОВЕРНОСТИ ИСХОДНЫХ ДАННЫХ
НА ОСНОВЕ ТЕОРИИ НЕЧЕТКИХ МНОЖЕСТВ**

1.2.2. Математическое моделирование, численные методы
и комплексы программ

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Ухта, 2025

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Ухтинский государственный технический университет» на кафедре вычислительной техники, информационных систем и технологий.

Научный руководитель: **Кожевникова Полина Валерьевна**, кандидат технических наук, доцент, доцент кафедры вычислительной техники, информационных систем и технологий федерального государственного бюджетного образовательного учреждения высшего образования «Ухтинский государственный технический университет».

Официальные оппоненты: **Антонов Вячеслав Викторович**, доктор технических наук, доцент, заведующий кафедрой автоматизированных систем управления федерального государственного бюджетного образовательного учреждения высшего образования «Уфимский университет науки и технологий»;
Волков Антон Владимирович, кандидат технических наук, доцент кафедры информационной безопасности и сервиса федерального государственного бюджетного образовательного учреждения высшего образования «Национальный исследовательский Мордовский государственный университет им. Н.П. Огарёва».

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего образования «Сыктывкарский государственный университет имени Питирима Сорокина», г. Сыктывкар.

Защита диссертации состоится «25» ноября 2025 г. в 15⁰⁰ часов на заседании диссертационного совета 24.2.350.05, созданного на базе федерального государственного автономного образовательного учреждения высшего образования «Омский государственный технический университет» по адресу: 644050, г. Омск, пр. Мира, д. 11, ауд. Г-226.

С диссертацией можно ознакомиться в библиотеке Омского государственного технического университета и на официальном сайте <http://www.omgtu.ru>.

Автореферат разослан «___» _____ 202_ г.

Ученый секретарь
диссертационного совета 24.2.350.05,
доктор технических наук, доцент



Варепо Л. Г.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность диссертационной работы. Одним из важнейших условий получения достоверных математических (и не только) моделей является их базирование на достоверных исходных данных. Человеческий фактор, погрешность приборов, методики измерений – все это оказывает влияние на качество моделей. К примеру, от качества исходных данных напрямую зависит результат оценки количества запасов углеводородов и прогнозирования обводненности продукции нефтедобывающих скважин (топливно-энергетический сектор промышленности), корректность диагностики технического состояния агрегатов транспортных средств (эксплуатация автотранспортных средств), точность определения состава минеральных вод большой залежи (гидрогеология). Основываясь на оптимальности соотношения «универсальность-необходимое количество исходных данных», чаще всего для прогнозирования параметров используются методы математической статистики (в частности, объемный метод и его модификации). Однако, существенным недостатком применения статистических методов является отсутствие заложенной в них оценки достоверности входных данных. Компенсировать данный недостаток способно многократное проведение экспериментальных замеров, что не представляется возможным при решении задач в приведенных выше областях науки.

В условиях, не предполагающих повторные эксперименты, нужна обработка данных технологией, которая позволяет учитывать их особенности (неоднородную структуру), ранжируя исходные данные по достоверности. Создание технологии возможно с использованием нечеткого моделирования. Однако, это возможно при выполнении следующих условий:

- адаптации этой технологии к определенным задачам (для решения конкретных проблем);
- развития математического аппарата.

С помощью данной технологии возможна оценка достоверности значений параметров, что позволит снижать риски технико-экономического характера.

Зачастую, для построения математических моделей параметры имеют следующие особенности, которые связаны с неоднородным строением среды и переносятся в конечные построения:

- фрагментарность (разное количество измеренных значений);
- неоднородность (большой разброс значений одновременно измеренных параметров);
- неопределенность (отсутствие измеренных значений).

Анализ литературы показывает, что, вопрос изучения нечеткого моделирования является актуальным в последнее десятилетие, и теоретические основы его использования в решении производственных задач описаны в трудах ряда мировых и российских ученых, таких, как: В. В. Борисов, А. Е. Алтунин, И. П. Жабрев, Xu Dong Jing, A. Bardossy и др. В научных работах А. И. Кобрунова, А. С. Могутова, А. Н. Дорогобед и П. В. Кожевниковой приведены результаты адаптации теории нечетких множеств к прикладным задачам и апробации на тестовых данных. Однако, остался ряд задач, которые требуют создания новых методов. Конкретно, отсутствуют работы, свидетельствующие:

- об использовании производной Фреше для решения оптимизационной задачи при разработке итерационного алгоритма оценки данных;

- об определении оптимального размера шага расчетной сетки при оценке достоверности подсчетов запасов углеводородов;
- о возможном применении большинства созданных математических методов и комплексов программ в различных областях науки.

Объектом исследования являются фильтрационно-емкостные свойства (пористость, проницаемость, насыщенность) и геолого-геофизические параметры (акустический импеданс, интервальное время) пород-коллекторов в объеме данных об исследовании скважин конкретных месторождений нефти и газа; данные определения состава газа в водах природных источников; результаты анализа образцов отработанного моторного масла; данные по обводненности флюида нефтедобывающих скважин.

Предметом исследования - неоднородность взаимозависимостей геолого-геофизических и фильтрационно-емкостных параметров и их возможные способы моделирования; неоднородность исходных данных в процессе прогнозирования геолого-промысловых характеристик месторождений по геофизическим и лабораторным исследованиям; неоднородность исходных данных в процессе определения состава газовой смеси в водах природных источников; неоднородность исходных данных в процессе прогнозирования потери эксплуатационных свойств смазочных материалов для автотранспорта; неоднородность исходных данных при прогнозировании обводненности флюида нефтедобывающих скважин.

Целью диссертационного исследования являются разработка численных методов, математической модели и программного комплекса, предназначенных для оценки достоверности значений неоднородных исходных данных в задачах топливно-энергетического сектора промышленности (с демонстрацией возможности применения результатов разработки в различных областях науки, где исходные данные обладают неоднородностью).

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

1. Предложить математическую модель, представляющую зависимости между исходными данными в топливно-энергетическом секторе промышленности (обладающими неоднородностью, фрагментарностью, неопределенностью), в форме нечетких отношений.
2. Разработать численный метод для подбора оптимального вектора источников информации с целью оценки достоверности исходных данных, используемых для решения задач топливно-энергетического сектора.
3. Разработать комплекс программ для анализа влияния размера расчетной сетки (при проведении процедуры оценки достоверности исходных данных) на точность построения моделей для решения задач топливно-энергетического сектора промышленности и, соответственно, на достоверность оценки исходных данных.
4. Продемонстрировать экспериментальное обоснование возможности применения созданного численного метода и комплекса программ в различных областях науки и техники, где исходные данные обладают аналогичными особенностями.

Научная новизна диссертационной работы состоит в следующем:

- предложена математическая модель, позволяющая представлять исходные данные в форме нечетких отношений, отличающаяся тем, что исходные данные ранжируются по степени достоверности;

- предложен математический метод, позволяющий рассчитывать оптимальный вектор входных параметров при построении физических зависимостей, отличающийся применением теории нечетких множеств;
- разработан программный комплекс, позволяющий осуществлять подбор шага расчетной сетки для построения моделей с большой степенью достоверности при допустимом расчетном времени, апробированный на предприятиях ТПП «Усинскнефтегаз» и ООО «АГМ-Сервис»;
- разработан программный комплекс, позволяющий выполнять прогноз обводненности флюида нефтедобывающих скважин, апробированный на предприятии ТПП «Усинскнефтегаз»;
- продемонстрирована возможность применения разработанной математической модели, позволяющей ранжировать по достоверности значения параметров, в различных технических отраслях науки.

Теоретическая и практическая значимость работы. Созданные алгоритмы, методы и комплекс программ на основе элементов теории нечеткого моделирования могут применяться при оценке достоверности данных, обладающих рядом особенностей: прогнозный подсчет запасов углеводородов; определение состава газовой смеси в водах природных источников; прогнозирование потери эксплуатационных свойств смазочных материалов; прогнозирование обводненности флюида нефтедобывающих скважин.

Результаты диссертационного исследования апробированы и внедрены в прикладную деятельность следующих предприятий: учебный процесс ФГБОУ ВО «Ухтинский государственный технический университет»; ООО «АГМ-Сервис» (г. Новосибирск); ТПП «Усинскнефтегаз» (г. Усинск).

Методы исследования. При выполнении диссертационной работы использовались методы математического моделирования; элементы нечеткой логики; методы оптимизации функции принадлежности; методы решения дифференциальных и интегральных уравнений.

Основные положения, выносимые на защиту:

1. Математическая модель представления исходных данных в форме нечетких отношений, основанная на итерационном алгоритме подбора оптимального вектора источников информации, позволяет ранжировать данные по достоверности.
2. Программный продукт, основанный на подборе оптимального шага расчетной сетки при формировании исходных данных в форме нечетких отношений, позволяет повысить точность результатов вычислений.
3. Разработанный метод, основанный на нечетком моделировании, позволяет выполнять оценку достоверности значений параметров в различных отраслях науки.

Соответствие паспорту научной специальности

Диссертационная работа соответствует паспорту научной специальности 1.2.2 «Математическое моделирование, численные методы и комплексы программ» по следующим пунктам: п.2 – Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий; п.3 – Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента; п.8 – Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента.

Степень достоверности результатов. Достоверность результатов подтверждается математически корректными выводами из анализа результатов обработки реальных исходных данных: фактических геолого-геофизические характеристики реальных нефтегазовых месторождений Тимано-Печорской нефтегазоносной провинции и Северо-Устюртской нефтегазоносной области; данных определения состава газа в водах природных источников Эльбрусской и Казбекской вулканических областей Большого Кавказа; результатов анализа образцов отработанного моторного масла. Результаты, полученные при апробации, позволяют сделать вывод о возможности использования разработанных методов для оценки достоверности производственных моделей. Результаты исследований, отраженные в четвертой главе, позволяют сделать вывод о возможности использования разработанных методов в различных областях науки, где присутствуют аналогичные проблемы исходных данных.

Апробация результатов работы. Ключевые положения диссертационного исследования были представлены и обсуждены на следующих научных конференциях: Международная молодежная научно-практическая конференция «Рассохинские чтения» (Ухта, 2021-2025 гг.); Международная молодежная научно-практическая конференция «СЕВЕРГЕОЭКОТЕХ» (Ухта, 2021-2025 гг.); Всероссийская научно-практическая конференция «Управление устойчивым развитием в топливно-энергетическом комплексе» (Ухта, 2021-2025 гг.); XIV Всероссийская конференция молодых ученых, специалистов и студентов «Новые технологии в газовой промышленности» (Москва, РГУНГ им. И. М. Губкина, 2022 г.); Всероссийская научно-практическая конференция «Цифровые технологии в образовании, науке, обществе» (Петрозаводск, ПетрГУ, 2022-2024 гг.).

Публикации. Основные положения диссертации напечатаны в 21 научной работе, в том числе 6 – в журналах, рекомендованных ВАК РФ, включенных в «Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук»; 15 – в научных журналах, сборниках материалов конференций и сборниках тезисов конференций различного уровня. Получено 2 свидетельства о регистрации программы для ЭВМ.

Личный вклад автора. Разработанные в диссертационном исследовании и выносимые на защиту численные (вычислительные) методы и комплекс программ принадлежат лично автору.

Структура и объем диссертации.

Диссертация состоит из введения, четырех глав, заключения, списка литературы и трех приложений. Объем диссертации с приложениями составляет 115 страниц, без приложений - 111 страниц. Диссертация содержит 10 таблиц и 38 рисунков. Список литературы включает 81 наименование.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационной работы, сформулирована цель исследования, выделены задачи, отражена научная новизна и практическая значимость выполненных исследований, сформулированы выносимые на защиту научные положения, описана структура диссертации.

В первой главе описаны основные особенности исходных данных, используемых при построении петрофизических моделей, представлены краткие теоретические сведения о нечетком моделировании.

К отличительным чертам зависимостей петрофизических параметров типа «кern-кern» и «кern-ГИС», описанных в главе 1 и составляющих основу математических моделей в области нефтегазодобычи, можно отнести неоднородность, фрагментарность и неопределенность.

Нечеткость исходных данных, и объективная нечеткость петрофизических моделей в задаче нефтегазовой геологии приводит к введению понятия нечеткого моделирования при создании математических моделей нефтегазоперспективных объектов и оценке его параметров.

Метод прогнозирования параметров геологической модели на основе нечетких данных и нечетких отношений между переменными основан на принципах нечеткого логического вывода и конструирования композиций отношений с целью исключения промежуточных параметров. Первый этап состоит в том, что экспериментальные данные о параметрах среды и петрофизические модели представляются в форме нечетких величин и нечетких отношений. Этот этап называется фазификацией данных. Он выполняется для каждого объекта – залежи или пропластка. Петрофизические модели состоят из данных, определяющих нечеткие зависимости между параметрами, организованными в цепочки от начальных параметров, к конечным - прогнозным. Эти нечеткие зависимости имеют характер отношений между нечеткими величинами. Начальные параметры - это те, которые измеряются. Это, как правило, данные ГИС. Конечные - это те, распределение которых составляет предмет моделирования.

Метод нечеткого прогнозирования основан на теоретических принципах нечеткого логического вывода, выраженного в построении композиций Мамдани и нечетких композиций в целом. Он включает в себя следующие этапы:

1. *Формализация и анализ.* Построение структурно-логической связи между измеряемыми, промежуточными и прогнозными параметрами для каждого из пропластков изучаемой залежи, которые представляют собой самостоятельные единицы исследования. Итогом анализа служит: выявление структурно-логической связи между измеряемыми, промежуточными и прогнозными параметрами для каждого из пропластков изучаемой залежи.

2. *Фазификация.* Для каждой цепочки структурно-логической схемы фазификация состоит в представлении исходных данных, относящихся к изучаемой единице месторождения, в виде нечетких величин и нечетких отношений. Этап включает в себя:

2.1. *Фазификация отношений.* Состоит в построении функций принадлежности для отношений между исходными \mathbf{X} , промежуточными $\mathbf{y}, \mathbf{h}, \dots, \xi$, и прогнозными параметрами $\mathbf{Z} : \mu(\mathbf{x}, \mathbf{y}), \mu(\mathbf{y}, \mathbf{h}), \dots, \mu(\xi, \mathbf{z})$, для каждой из изучаемых единиц. Реализация этапа, основана на тех же данных, которые служат основой для построения регрессионных уравнений. Этот этап является пересмотром и обобщением материалов, служащих основанием для создания петрофизической модели.

2.2. *Фазификация исходных данных.* Состоит в построение функций принадлежности $\mu(\mathbf{x}, \mathbf{R}_i)$, для исходных величин параметра \mathbf{X} в локальных интервалах \mathbf{R}_i , в которых будет выполняться прогнозирование.

3. *Расчёт композиций нечетких отношений.* Выполняется в случае, когда начальный и итоговый параметр связаны цепочкой нечеткой петрофизических моделей. Служит для установления отношений $\mu(\mathbf{x}, \mathbf{z})$ между начальными и конечными параметрами в цепочке для каждой изучаемой единицы.

4. Прогноз функции принадлежности $\mu(z, R_i)$, для параметра Z , в локальной точке пространства R_i , на основе принципа нечеткого логического вывода в реализации композиции Мамдани $\mu(x, z)$ и $\mu(x, R_i)$.

Схематично технология нечеткого моделирования, разработанная научной школой Кобрунова А. И., представлена на рисунке 1.

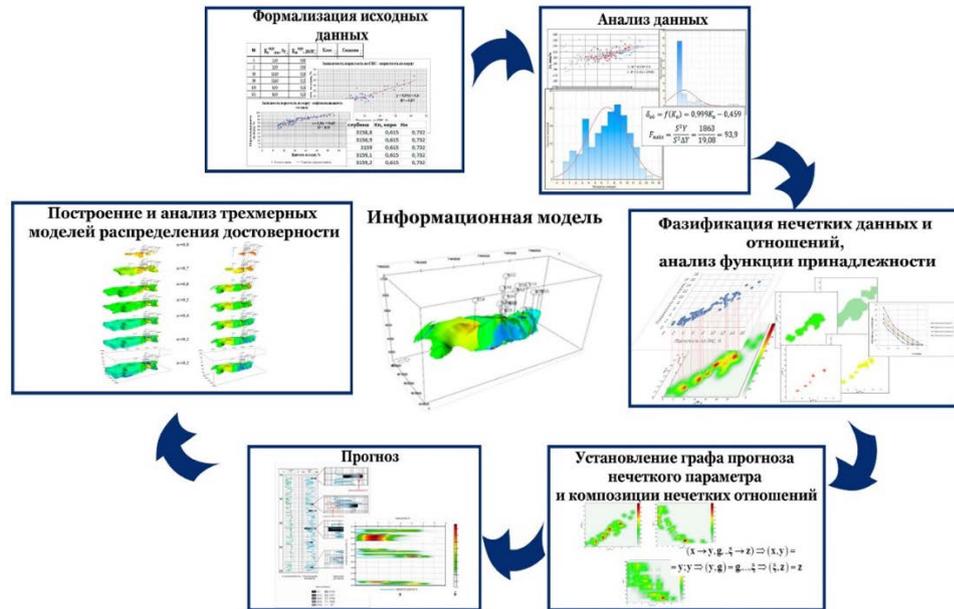


Рисунок 1 – Технология нечеткого моделирования и построения информационной модели

Этап *Фашификация* является самым трудоемким процессом технологии нечеткого моделирования, качество выполнения которого, сильно влияет на результат построения геологической модели, ранжированной по достоверности, поэтому все научные положения данной работы выполнялись в рамках этапа Фашификация.

Во второй главе представлено описание построения математической модели, основанной на итерационном алгоритме подбора оптимального вектора источников информации, позволяющей представлять исходные данные в форме нечетких отношений, ранжированных по достоверности.

Для оценки достоверности неоднородные данные (исходные данные взаимосвязи между параметрами) следует рассматривать как нечеткие отношения и конструировать функции принадлежности с помощью аппарата, основанного на алгоритме сжатия информации, заключающегося в поиске наименьшего количества источников, при которых невязка между исходными данными и рассчитанным нечетким отношением удовлетворяет заданному условию. Алгоритм сжатия подразумевает, что данные об одновременно измеренных значениях параметров исходят от источников, что в свою очередь позволяет существенно снизить размерность измеряемых величин.

Пусть \mathcal{X} – исходные данные, представляющие собой матрицу размерностью $M \times L$, где M – количество одновременно измеренных значений параметров, а L – количество параметров. Строка матрицы есть вектор значений одновременно измеренных параметров. Матрица исходных данных имеет следующий вид:

$$\mathfrak{A} = \begin{pmatrix} a_{11} & \cdots & a_{1L} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{ML} \end{pmatrix}$$

Далее рассмотрим задачу в условиях, когда количество параметров в исходных данных $L = 2$.

Исходные данные покрываются сеткой S , размерностью $N_l, l = 1 \div L$, где N_l – количество ячеек в l измерении, и рассчитывается карта плотности данных $\mathfrak{A}(\mathbf{s})$, где \mathbf{s} – ячейки сетки.

Размер ячейки $r = r_1 \times r_2 \times \dots \times r_L$ вычисляется по формуле:

$$r_l = \frac{\max\{a_{ml}\} - \min\{a_{ml}\}}{N_l}, m = 1 \div M.$$

Карта плотности данных $\mathfrak{A}(\mathbf{s})$ отображает количество значений параметров, принадлежащих ячейкам сетки, относительно общего числа одновременно измеренных значений параметров.

Основанием для представления данных \mathfrak{A} в форме нечетких отношений служит понятие поле рассеяния, представляющее линейную комбинацию функций экспоненциальной модели:

$$\mathfrak{A}^\varepsilon(\mathbf{s}) = \sum_{k=1}^K \omega_{s^k} \frac{1}{\sqrt{\pi} \zeta_{s^k}} e\left(-\frac{|\mathbf{s}-s^k|^2}{\zeta_{s^k}^2}\right),$$

где \mathbf{s} – набор ячеек сетки, для которых выполняется расчет,

ε – погрешность, с которой выполняется расчет поля рассеяния,

ζ_{s^k} – эффективный параметр источника, влияющий на рассеяние данных,

K – количество источников информации,

s^k – ячейка, в которой расположен k -ый источник,

ω_{s^k} – вес k -го источника.

Поле рассеяния должно удовлетворять следующему условию:

$$\mathfrak{A}^\varepsilon(\mathbf{s}) - \mathfrak{A}(\mathbf{s}) \leq \varepsilon, \quad (1)$$

где ε – допустимая погрешность при построении поля рассеяния.

Предложенная модель имеет принцип максимальной энтропии, т.е. информация, полученная от источников – это итог диффузии, которая длилась определенное время. Задача заключается в нахождении минимального количества источников, от которых, как результат диффузии, мы получаем исходные данные, при которых поле рассеяния не превышает заданную погрешность.

Решение задачи заключается в распределении источников s^k , нахождении веса ω_{s^k} и эффективного параметра ζ_{s^k} для каждого источника, которое основано на представлении о том, что рассчитанное значение поля рассеяния, следующее из экспериментальных данных, это результат диффузии под действием эффективного параметра ζ_{s^k} , который привел к сглаживанию более точной зависимости ω_{s^k} :

$$\begin{cases} A[\mathbf{v}] = \mathfrak{A}^\varepsilon(\mathbf{s}) \\ \|\mathfrak{A}^\varepsilon(\mathbf{s}) - \mathfrak{A}(\mathbf{s})\| \rightarrow \min' \end{cases} \quad (2)$$

где A – оператор для расчета поля рассеяния,

\mathbf{v} – вектор входных параметров: $\mathbf{v} = \{\omega, \zeta\}$.

Алгоритм решения задачи состоит из двух этапов:

- на первом этапе решается задача расположения источников;
- на втором – подбор веса и эффективного параметра для каждого источника.

Первый этап

Для определения местоположения источников на сетке воспользуемся алгоритмом кластеризации. Кластеризация – это разбиение исходных данных на кластеры. Кластеры представляют из себя группы со схожими характеристиками, которыми, в нашем случае, являются значения одновременно измеренных параметров. Внутри каждой группы должны оказаться «похожие» объекты, а объекты разных группы должны быть как можно более отличны. В работе были изучены два метода кластеризации: *k*-средних и иерархическая кластеризация.

Алгоритм *k*-средних представляет собой версию EM-алгоритма, применяемого также для разделения смеси гауссиан. Он разбивает множество элементов векторного пространства на заранее известное число кластеров *k*.

Основная идея заключается в том, что на каждой итерации перевычисляется центр масс для каждого кластера, полученного на предыдущем шаге, затем векторы разбиваются на кластеры вновь в соответствии с тем, какой из новых центров оказался ближе по выбранной метрике.

Алгоритм завершается, когда на какой-то итерации не происходит изменения внутрикластерного расстояния. Это происходит за конечное число итераций, так как количество возможных разбиений конечного множества конечно, а на каждом шаге суммарное квадратичное отклонение уменьшается, поэтому заикливание невозможно.

В математическом виде метод *k*-средних представляется в виде:

$$\min_C \sum_{k=1}^K \sum_{i=1}^M \rho(\mathbf{a}_i, \mathbf{s}^k),$$

здесь $C = \{\mathbf{s}^k; k = 1 \div K\}$ – набор кластеров.

Входными данными для решения задачи кластеризации являются вектор значений параметров и количество источников. Данный метод удобен тем, что рассчитанные центры кластеров будут использованы в качестве координат источников данных.

Иерархическая кластеризация – совокупность алгоритмов упорядочивания данных, направленных на создание иерархии (дерева) вложенных кластеров.

Алгоритмы иерархической кластеризации предполагают, что анализируемое множество объектов характеризуется определённой степенью связности. Чаще всего используются агломеративные методы.

В данной диссертационной работе рассматривается такой вид иерархической кластеризации, как агломеративная дендрограмма.

В основе описанных выше алгоритмов кластеризации лежит критерий сравнения объектов x и x' , которым, как правило, является расстояние ρ ; за n обозначим количество характеристик объекта. Для расчета расстояния необходимо экспериментальным путем выбрать оптимальную метрику. В работе рассматриваются следующие метрики:

а) метрика евклидова расстояния – геометрическое расстояние между объектами в многомерном пространстве:

$$\rho(x, x') = \sqrt{\sum_i^n (x_i - x'_i)^2};$$

b) метрика квадрата евклидова расстояния - позволяет придавать больше веса более отдаленным друг от друга элементам в кластере:

$$\rho(x, x') = \sum_i^n (x_i - x'_i)^2;$$

c) метрика Манхэттенского расстояния – средняя разность по координатам:

$$\rho(x, x') = \sum_i^n |x_i - y_i|;$$

d) метрика расстояния Чебышева – вычисленное как максимум абсолютного значения разности между элементами расстояния:

$$\rho(x, x') = \max |x_i - x'_i|$$

Второй этап

Вычислительной основой для решения задачи (2) в рамках второго этапа для определения веса источников в точках, рассчитанных на этапе 1, и эффективного параметра рассеяния для каждого источника служит итерационный процесс, на $z + 1$ итерации имеющий вид:

$$\begin{cases} \mathbf{v}^{z+1} = \mathbf{v}^z + \alpha^z A'^*[\mathbf{v}] \boldsymbol{\varphi}^z \\ \|\mathbf{v}^{z+1} - \mathbf{v}^z\| \leq \theta \\ \boldsymbol{\varphi}^z = \mathfrak{A}(\mathbf{s}) - \mathfrak{A}^\varepsilon(\mathbf{s}) \end{cases} \quad (3)$$

В рамках итерационного процесса необходимо адаптировать модель (подобрать параметры \mathbf{v}) к реальным данным.

Здесь:

$\boldsymbol{\varphi}^z$ – невязка между рассчитанным полем рассеяния на итерации z и картой плотности исходных данных;

θ – необходимая разница между векторами входных параметров на соседних итерациях;

α^z – параметр релаксации, подбираемый на каждом шаге так, чтобы итерационный процесс сходился:

$$\alpha^z = \frac{\|A'^*[\mathbf{v}] \boldsymbol{\varphi}^z\|^2}{\|A'[\mathbf{v}] A'^*[\mathbf{v}] \boldsymbol{\varphi}^z\|^2},$$

где $A'^*[\mathbf{v}]$ – сопряженный к $A'[\mathbf{v}]$ оператор.

Оператор $A'[\mathbf{v}]$ есть производная Фреше к оператору $A[\mathbf{v}]$ в «точке» \mathbf{v} , матрица $\{K \times K * 2\}$ обратного оператора A в частных производных по параметрам $\{\boldsymbol{\omega}, \boldsymbol{\zeta}\}$. $A'[\mathbf{v}]$ – линейный оператор, действующий на \mathbf{v} :

$$A'[\mathbf{v}] = \begin{pmatrix} \frac{\partial A^1[\mathbf{v}]}{\partial v_1} & \frac{\partial A^1[\mathbf{v}]}{\partial v_2} & \dots & \frac{\partial A^1[\mathbf{v}]}{\partial v_{K*2}} \\ \frac{\partial A^2[\mathbf{v}]}{\partial v_1} & \frac{\partial A^2[\mathbf{v}]}{\partial v_2} & \dots & \frac{\partial A^2[\mathbf{v}]}{\partial v_{K*2}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial A^K[\mathbf{v}]}{\partial v_1} & \frac{\partial A^K[\mathbf{v}]}{\partial v_2} & \dots & \frac{\partial A^K[\mathbf{v}]}{\partial v_{K*2}} \end{pmatrix}$$

Например, если количество источников на первом этапе было задано 3, то оператор $A'[\mathbf{v}]$ будет иметь размерность $\{3 \times 6\}$:

– 3 – количество источников (K);

– 6 – значения параметров $\boldsymbol{\omega}$ и $\boldsymbol{\zeta}$ для каждого источника.

Оператор A имеет вид:

$$\left\{ \begin{array}{l} A[\mathbf{v}] = 0, \text{ где } s \neq s^k \\ A[\mathbf{v}] = \sum_{k=1}^3 \omega_{s^k} \frac{1}{\sqrt{\pi} \zeta_{s^k}} e^{\left(-\frac{|s-s^k|^2}{\zeta_{s^k}^2}\right)}, \text{ где } s = s^k \end{array} \right.$$

Частные производные оператора A по параметрам $\mathbf{v} = \{\omega, \zeta\}$:

1. Частная производная по параметру ω :

$$\frac{\partial A[\mathbf{v}]}{\partial \omega} = \frac{1}{\sqrt{\pi} \zeta_{s^k}} e^{\left(-\frac{|s-s^k|^2}{\zeta_{s^k}^2}\right)}.$$

2. Частная производная по параметру ζ :

$$\frac{\partial A[\mathbf{v}]}{\partial \zeta} = -\frac{\omega_{s^k}}{\sqrt{\pi} * \zeta_{s^k}} * e^{\left(-\frac{|s-s^k|^2}{\zeta_{s^k}^2}\right)} * \left(\frac{1}{\zeta_{s^k}} - \frac{2 * (s - s^k)^2}{\zeta_{s^k}^3} \right).$$

Вектор параметров v для трех источников имеет вид:

$$v = \{\omega_{s^{(1)}}, \zeta_{s^{(1)}}, \omega_{s^{(2)}}, \zeta_{s^{(2)}}, \omega_{s^{(3)}}, \zeta_{s^{(3)}}\}.$$

Сопряженный обратный оператор $A'^*[\mathbf{v}]$:

$$A'^*[\mathbf{v}] = \begin{pmatrix} \frac{\partial A^1[v]}{\partial \omega_{s^{(1)}}} & 0 & 0 \\ \frac{\partial A^1[v]}{\partial \zeta_{s^{(1)}}} & \frac{\partial A^2[v]}{\partial \omega_{s^{(2)}}} & \frac{\partial A^3[v]}{\partial \omega_{s^{(3)}}} \\ 0 & \frac{\partial A^2[v]}{\partial \zeta_{s^{(2)}}} & \frac{\partial A^3[v]}{\partial \zeta_{s^{(3)}}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Оператор $A'^*[\mathbf{v}]$ представляет собой матрицу частных производных по параметрам, где количество столбцов определяется количеством источников, а количество строк в два раза больше количества источников. Процесс завершается, когда разница между реальными данными и полем рассеяния φ^z принимает удовлетворительно малое значение, либо не удалось получить новый вектор \mathbf{v}^z . При этом в качестве нулевого приближения значений параметров принимаем:

$$\begin{cases} \omega^0 = 1, & \text{если } s \in s^k, \text{ иначе } 0 \\ \zeta^0 = 1 \end{cases}$$

После того как итерационный процесс завершился согласно условию в (3) и найдено решение $v(s)$, служащее новым приближением для распределения веса источников, задается новое значение количества источников, меньше предыдущего на единицу, на основе которых весь процесс повторяется с 1-го этапа.

Далее рассмотрим эксперимент по формированию функции принадлежности, основанного на алгоритме сжатия.

Исходные данные – взаимосвязь между одновременно измеренными параметрами «Пористость» и «Проницаемость» (данные по Бавлинскому месторождению взяты из открытых источников). Данные фильтрационно-емкостные свойства получены в результате совместной интерпретации данных акустического и гамма-гамма-плотностного каротажа (ГГКп) скважин, информация об оборудовании отсутствует.

Исходные данные и карта плотности данных представлены на рисунке 2 (А и Б соответственно).

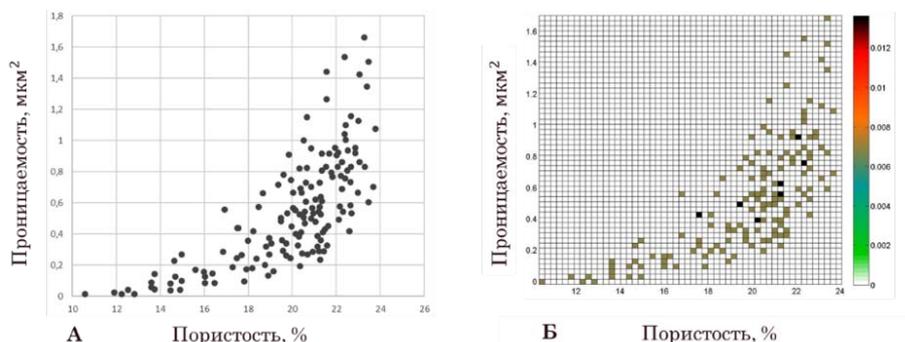


Рисунок 2 – Исходные данные (А) – взаимосвязь между параметрами «Пористость по ГИС» и «Пористость по керну»; карта плотности исходных данных $\mathcal{X}(s)$ (Б)

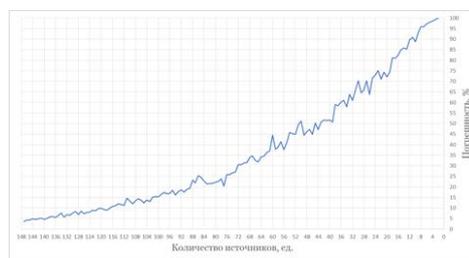
Максимально возможное количество кластеров равно количеству измеренных значений, а именно 147. Шаг количества источников был принят равным 1. Допустимая погрешность была принята $\leq 10\%$. Задача стояла в нахождении наименьшего количества источников, которыми являются кластеры.

Основные результаты вычислительных экспериментов с использованием метода кластеризации k -средних и метрик, описанных выше, представлены в таблицах 2.1 – 2.4 полной версии диссертации. Для краткой демонстрации в таблице 1 данного автореферата приведены результаты вычислительных экспериментов по методу k -средних с использованием метрики евклидова расстояния).

На рисунке 3 представлены графики динамики погрешности для рассмотренных метрик.



Метрика евклидова расстояния



Метрика квадрата евклидова расстояния



Метрика Манхэттенского расстояния



Метрика расстояния Чебышева

Рисунок 3 – Динамика погрешности при использовании разных метрик

Таблица 1 - Результаты вычислительных экспериментов (метод k-средних, метрика евклидова расстояния)

К-во ист.	Результаты первого этапа (раздел 2.2.2)	Результаты второго этапа (раздел 2.2.3)	Расчитанное поле рассеяния	Погрешность
147				3,6%
120				9,9%
100				15,2%
70				31,3%
30				70,3%
5				97,9%

При проведении сопоставительного анализа графиков можно сделать вывод о том, что минимальное количество источников, удовлетворяющее заданной погрешности в 10%, составило 117 для метрики евклидова расстояния, и 118 для трех остав-

шихся метрик (квадрата евклидова расстояния, Манхэттенское расстояние и расстояния Чебышева). Таким образом, было установлено, что в условиях данного диссертационного исследования разница между метриками минимальна, и для проведения дальнейших вычислительных экспериментов с использованием иерархического метода кластеризации достаточно выбрать одну из исследованных метрик. Автором диссертации была выбрана метрика евклидова расстояния. Затем были проведены аналогичные вычислительные эксперименты с использованием метода иерархической кластеризации, результаты которых кратко отображены в таблице 2.

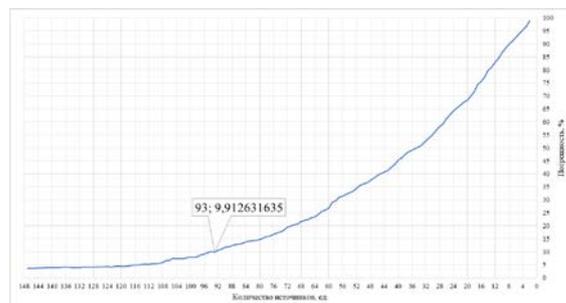
Таблица 2 - Результаты вычислительных экспериментов (метод иерархической кластеризации, метрика евклидова расстояния)

К-во ист.	Результаты первого этапа (раздел 2.2.2)	Результаты второго этапа (раздел 2.2.3)	Рассчитанное поле рассеяния	Погрешность
147				3,6%
120				4,4%
100				7,9%
70				20,5%
30				55,0%
5				94,0%

Для окончательного выбора алгоритма кластеризации необходимо сопоставить графики динамики погрешности результатов. Данные графики отражены на рисунке 4.



Метод кластеризации k-средних, метрика евклидова расстояния



Метод иерархической кластеризации, метрика евклидова расстояния

Рисунок 4 – Динамика погрешности при использовании разных метрик

Анализируя динамику погрешности при использовании иерархической кластеризации, можно сделать вывод о том, что минимальное количество источников, удовлетворяющее заданной погрешности в 10%, составило 93 (63% от общего количества), что существенно меньше, чем при использовании кластеризации по методу k-средних (117, что составляет 80% от общего количества). Таким образом, установлено, что для построения функций принадлежности необходимо использовать метод иерархической кластеризации.

Разработанный метод фазификации отношений между одновременно измеренными параметрами, основанный на алгоритме сжатия информации, позволяет снизить размерность величин, а также избавиться от «случайной» информации, которая может являться причиной неверной интерпретации данных

В третьей главе описан процесс разработки комплекса программ для подбора оптимального шага расчетной сетки с целью повышения точности результата вычислений.

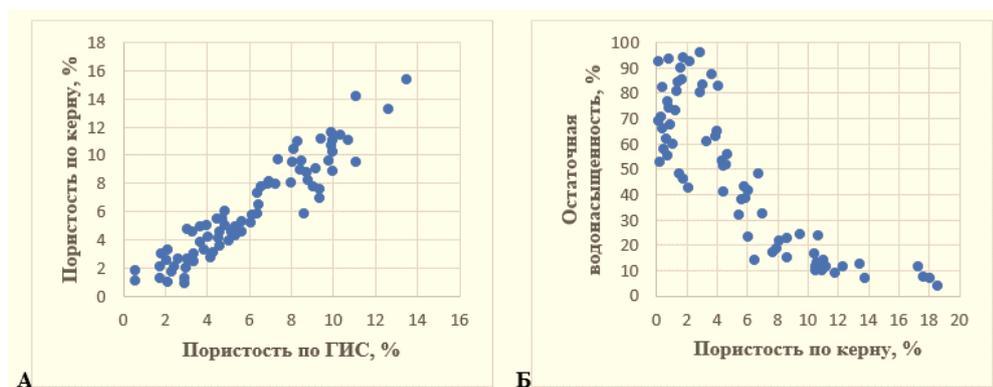
Адекватность значений полученных экспериментальным способом петрофизических моделей описывает функция принадлежности. Построение данной функции осуществляется с помощью покрытия измеренных параметров $s^j \in S, j = 1 \div K$ сеткой прямоугольной формы S размером $N \times M$.

Оптимально подобранный размер шага расчетной сетки позволяет повысить точность производимых вычислений, а также увеличить разрешающую способность сетки, что, в конечном итоге, повысит качество построения функций принадлежности отношений между рядом петрофизических параметров.

Объектами исследования были выбраны два отношения между петрофизическими параметрами, относящимися к ключевым и характеризующие среду по нефтегазоперспективности (Арыстановское месторождение): определенные одновременно численные значения пористости по керну и пористости по ГИС (далее $Kп^{ГИС}$ - $Kп^{кern}$), а также определенные одновременно численные значения пористости по керну и остаточной водонасыщенности (далее $Kп^{кern}$ -Ков). Отражено на рисунке 5.

Функции принадлежности $\mu_A(s)$ рассчитывается как нормированное к единице поле рассеяния:

$$\mu_A(s) = \frac{\mathcal{U}^\varepsilon(s)}{\max_s [\mathcal{U}^\varepsilon(s)]}$$



а – определенные одновременно численные значения «пористость по керну» и «пористость по ГИС»; б – определенные одновременно численные значения «пористость по керну» и «остаточная водонасыщенность»

Рисунок 5 – Экспериментальные данные

Функция принадлежности – это матрица с размерами N на M, которая содержит значения достоверности от 0 до 1. 0 – достоверность отсутствует, параметры не примут рассматриваемые значения, 1 – высокая степень достоверности, параметры примут рассматриваемые значения.

Композиция Мамдани позволяет выполнять свертку нечетких отношений через промежуточный параметр «пористость по керну», получая в результате отношение «пористость по ГИС – остаточная водонасыщенность»:

$$\mu_{\text{ГИС}*\text{Я}}(x, z) = \max_y [\min(\mu_{\text{ГИС}}(x, y), \mu_{\text{Я}}(y, z))].$$

где $\mu_{\text{ГИС}}(x, y)$ – функция принадлежности «пористость по ГИС- пористость по керну», $\mu_{\text{Я}}(y, z)$ – функция принадлежности «пористость по керну – остаточная водонасыщенность», y – параметр «пористость по керну».

Далее приведены результаты экспериментов по подбору оптимального размера расчетной сетки.

Расчеты, представленные ниже, проводились в программе, созданной на языке С# и зарегистрированной в установленном законодательством РФ порядке (свидетельство о регистрации программы для ЭВМ №2022662182 от 30.06.2022 г).

Вычисление промежуточных ($K_{\text{П}}^{\text{ГИС}}-K_{\text{П}}^{\text{КЕРН}}$ и $K_{\text{П}}^{\text{КЕРН}}-K_{\text{ОВ}}$) и итоговой ($K_{\text{П}}^{\text{ГИС}}-K_{\text{ОВ}}$) функций принадлежности началось с размера сетки 5x5, затем размер сетки с шагом 1 увеличивался до размера 1000x1000.

По каждому эксперименту фиксировалась максимальная достоверность итоговой ($K_{\text{П}}^{\text{ГИС}}-K_{\text{ОВ}}$) функции принадлежности. В таблице 3 представлены полученные функции принадлежности при размерах сетки 5x5, 50x50, 300x300 и 1000x1000.

Изменение максимального значения итоговой функции принадлежности в зависимости от размера сетки представлено на графике (рисунок 6).

Как видно из графика, максимальное значение функции принадлежности имеет характер бесконечного числового ряда и стремится к 0,75 (среднее значение по всем наборам расчетных сеток, по которым выполнялись эксперименты). Выбираем допустимой погрешность значения функции принадлежности в 3%, тогда, сетка для расчетов должна быть размером 300x300.

Таблица 3 - Результат построения функций принадлежности $K_{П}^{ГИС}-K_{П}^{кern}$, $K_{П}^{кern}-K_{ОВ}$ и $K_{П}^{ГИС}-K_{ОВ}$ при размерах сетки 5x5, 50x50, 300x300 и 1000x1000

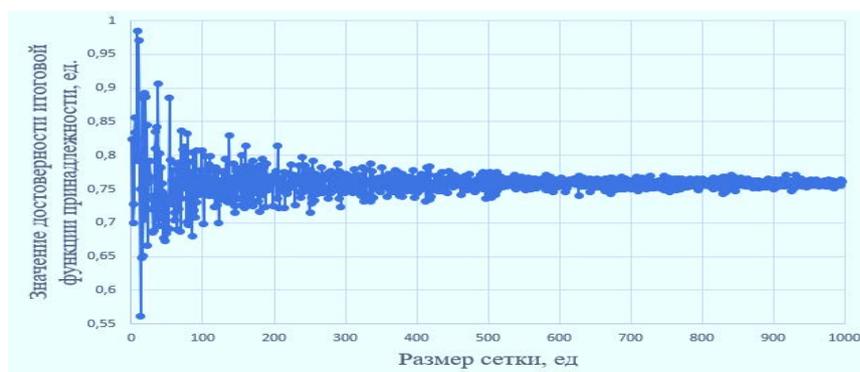
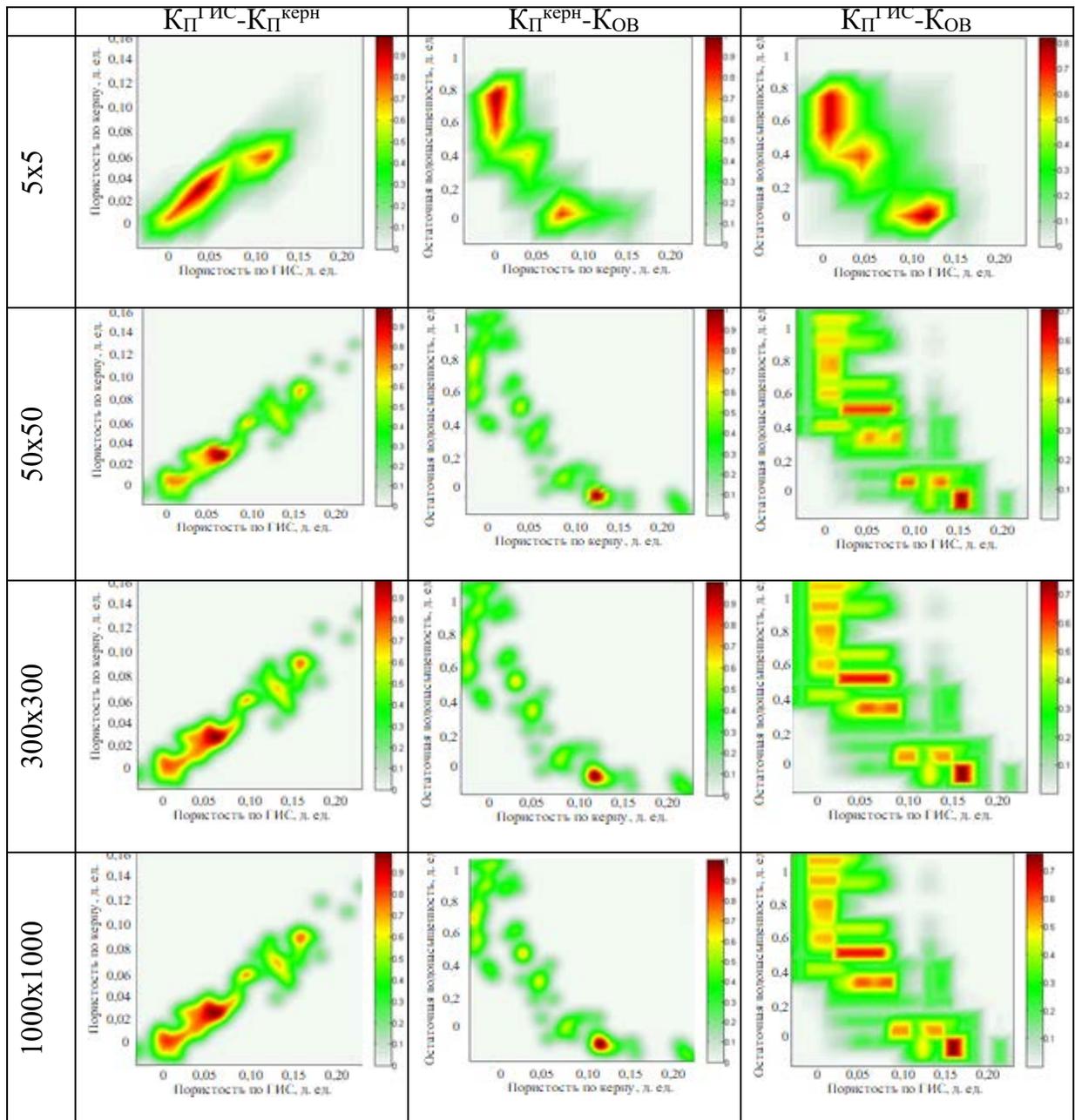


Рисунок 6 – Изменение максимального значения итоговой функции принадлежности «Пористость по ГИС – остаточная водонасыщенность» в зависимости от размера сетки

Однако, применение сетки размером 1000×1000 предполагает существенное (в масштабах объема производственных данных и задач) увеличение времени расчетов (зависимость времени машинного вычисления от размера расчетной сетки представлена на рисунке 7).

Интерфейс программы для подбора оптимального размера сетки и основные этапы работы с ней изображены на рисунке 8 (а-г).

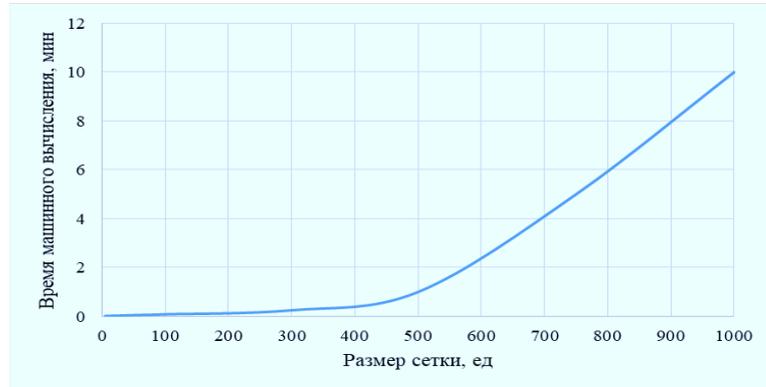
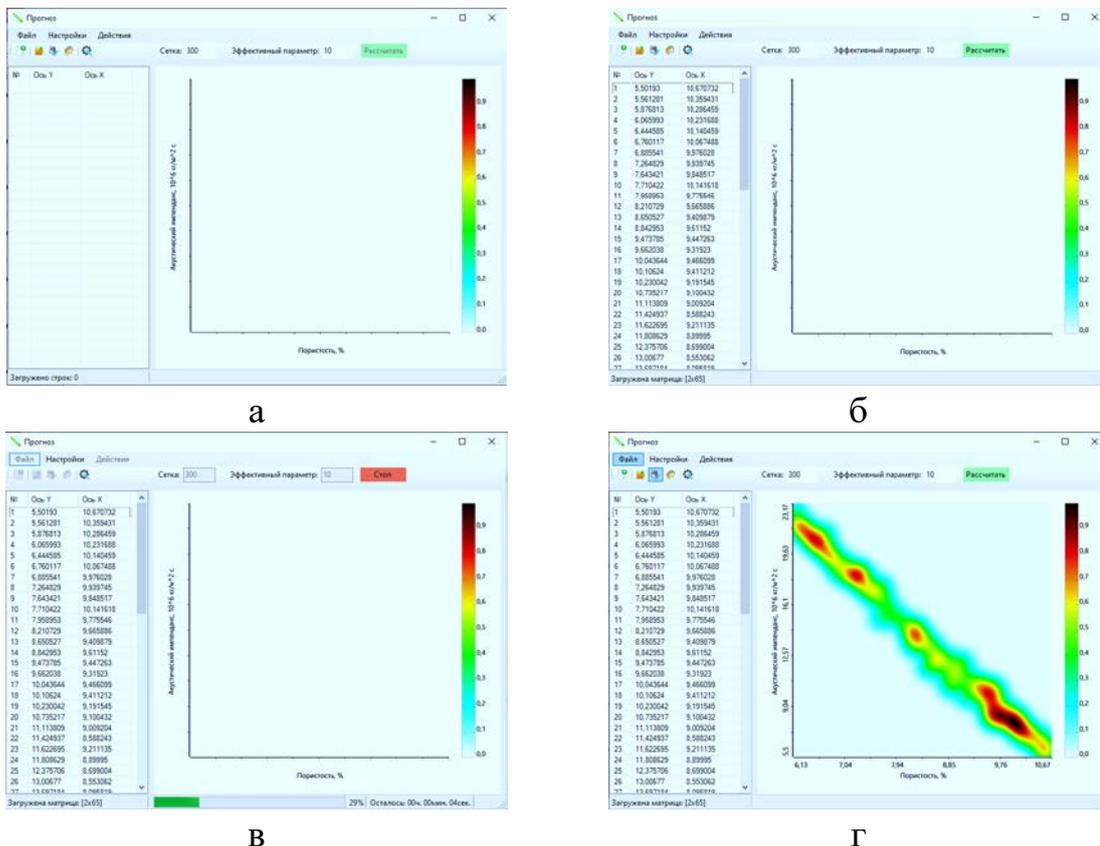


Рисунок 7 - Зависимость времени машинного вычисления от размера расчетной сетки



а – начальная страница; б – загрузка исходных данных;
в – процесс расчета; г – результат расчета

Рисунок 8 – Интерфейс программы для подбора оптимального размера сетки

В четвертой главе экспериментально обосновывается возможность применения разработанного математического аппарата в других областях науки и техники. В автореферате кратко приведены вычислительные эксперименты лишь по одной области, подробные сведения содержатся в полной версии диссертации.

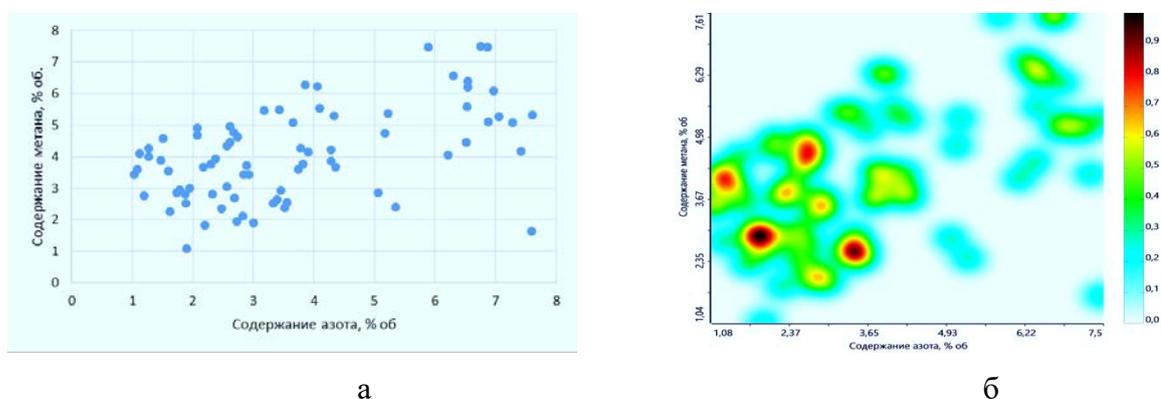
Проблемы исходных данных, описанные в первой главе, присущи не только топливно-энергетическому сектору. К примеру, от достоверности исходных данных

гидрогеологических исследований зависит корректность оценки количества минеральных вод природных источников и отнесение их к тому или иному типу.

Для демонстрации корректности применения разработанного программного продукта из открытых источников были взяты результаты одновременных измерений содержания азота и метана в природных водах Эльбрусской вулканической области Большого Кавказа. Результаты данных измерений являются основой для гидродинамического моделирования характеристик больших площадей залегания минеральных вод, следовательно, своевременная оценка их достоверности положительно скажется на качестве построенных гидрогеологических моделей.

Данные были оцифрованы (рисунок 9, а), после чего загружены в разработанный программный продукт. Размер расчетной сетки был выбран с учетом выводов, приведенных в третьей главе (т.е., 300x300 ед.). По результатам работы программы была получена оценка достоверности данных, представленная на рисунке 9, б.

Исходя из сопоставления исходных данных с результатами работы программы, можно сделать вывод о непротиворечии друг другу и возможности учета данной оценки достоверности при построении гидрогеологических моделей.



а – исходные данные; б – результат расчета

Рисунок 9 – Применение программы для оценки достоверности исходных данных в задачах гидрогеологии

Заключение. Ниже представлены основные результаты проведенных в диссертационной работе исследований.

1. Предложена математическая модель, представляющая зависимости между исходными данными для моделирования в топливно-энергетическом секторе промышленности в форме нечетких отношений. При этом освещены особенности исходных данных, составляющие проблему снижения точности расчетов: неоднородность; фрагментарность; неопределенность.

2. Разработан численный метод для подбора оптимального вектора источников информации с целью оценки достоверности исходных данных, используемых для решения задач топливно-энергетического сектора.

3. Разработан комплекс программ для анализа влияния размера расчетной сетки (при проведении процедуры оценки достоверности исходных данных) на точность построения моделей для решения задач топливно-энергетического сектора промышленности и, соответственно, на достоверность оценки исходных данных. С помощью вычислительных экспериментов установлен и обоснован оптимальный размер шага расчетной сетки и эффективного параметра для решения ряда задач топливно-энергетического сектора промышленности.

4. Продемонстрировано экспериментальное обоснование возможности применения созданного численного метода и комплекса программ в различных областях науки и техники, где исходные данные обладают аналогичными особенностями:

- эксплуатации автотранспортных средств (приведены материалы внедрения результатов диссертационных исследований в деятельность предприятия ООО «АГМ-Сервис»);
- гидрогеологии;
- разработке нефтяных месторождений (приведены материалы внедрения результатов диссертационных исследований в деятельность предприятия ТПП «Усинскнефтегаз»; создан программный комплекс для прогнозирования обводненности продукции нефтедобывающих скважин).
- Перспективы дальнейшего развития диссертационных исследований заключаются в:
 - многостороннем взаимодействии с нефтегазодобывающими компаниями Российской Федерации (внедрение программного комплекса в производственные процессы);
 - создании научных связей с авторами других научных исследований в данной области с целью совместных разработок;
 - совершенствовании созданных численного метода и комплекса программ путем видоизменения вычислительного аппарата (использование интеграла Лебега, интеграла Римана-Стилтьеса, второй производной Фреше).

ОСНОВНЫЕ НАУЧНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИОННОЙ РАБОТЫ

Научные публикации в изданиях из перечня, рекомендованного ВАК:

1. Кожевникова, П.В. Влияние шага расчетной сетки при построении функций принадлежности отношений между петрофизическими параметрами / П. В. Кожевникова, В. Е. Кунцев, **А. А. Чувашов** // Современная наука: актуальные проблемы теории и практики. – 2022. – № 7. – С. 65–70.
2. Кожевникова, П.В. Математическая модель расчета источников информации при построении функции принадлежности в задачах оценки достоверности запасов углеводородов / П. В. Кожевникова, В. Е. Кунцев, **А. А. Чувашов** // Вестник Астраханского государственного технического университета. – 2023. – № 1. – С. 98–104.
3. **Чувашов, А. А.** Применение элементов теории нечеткого моделирования в задачах контроля технического состояния узлов и агрегатов автотранспортных средств / А. А. Чувашов, П. В. Кожевникова, В. Л. Мушинский // Современная наука: актуальные проблемы теории и практики. – 2023. – № 1. – С. 166–171.
4. Шилова, С. В. Оценка технического состояния узлов и агрегатов транспортных средств по результатам исследований образцов отработанных смазочных материалов методом нечеткого моделирования / С. В. Шилова, П. В. Кожевникова, **А. А. Чувашов** // Научно-технический вестник Поволжья. – 2023. – № 2. – С. 24–27.
5. **Чувашов, А. А.** Применение математической модели Ловерье при расчёте технологических показателей теплового метода увеличения нефтеотдачи / А. А. Чувашов, П. В. Кожевникова, А. Н. Гресюк // Научно-технический вестник Поволжья. – 2024. – № 5. – С. 66–69.

6. **Чувашов, А. А.** Оценка достоверности исходных данных для построения прикладных петрофизических и гидрогеологических моделей с применением теории нечетких множеств / А. А. Чувашов // Научно-технический вестник Поволжья. – 2024. – № 12. – С. 99–101.

Свидетельства о государственной регистрации программ для ЭВМ:

1. Свидетельство о государственной регистрации программы для ЭВМ № 2022662182 Российская Федерация. Программа расчета функции принадлежности петрофизических параметров с подбором расчетной сетки № 2022618519: заявл. 05.05.2022, зарег. 30.06.2022 / **Чувашов А. А.**, Кожевникова П. В. – 1 с.

2. Свидетельство о государственной регистрации программы для ЭВМ № 2024665722 Российская Федерация. Программа для прогнозирования обводненности продукции нефтяных скважин с использованием математической модели Ба-клея-Леверетта № 2024664886: заявл. 24.06.2024, зарег. 04.07.2024 / **Чувашов А. А.**, Кожевникова П. В., Мушинский В. Л. – 1 с

В авторской редакции.

Подписано в печать 23.09.2025 г. Формат 60x84 1/16.

Усл. печ.л. 1,05. Тираж 100 экз. Заказ № 55.

Издательство УГТУ

169301, г. Ухта, ул. Первомайская, д. 13

Типография УГТУ