

Реферат

Выпускная квалификационная работа 103с., 27 рис., 23 таб., 20 источников, 1 прил.

Ключевые слова: сейсморазведка, месторождение, атрибутивный анализ, коэффициент пористости, коэффициент песчаности, эффективная толщина, регрессионный анализ, нейронные сети.

Объектом исследования являются совокупность скважин и пласт нефтяного месторождения.

Цель работы – характеристика трехмерного строения нефтяного месторождения на основе методов математической статистики.

В процессе исследования проводились статистический анализ, регрессионный анализ, нейросетевой анализ.

В результате исследования были прокоррелированы отражающие горизонты, рассчитаны сейсмические атрибуты, построены регрессионные и нейросетевые модели для коэффициента пористости, коэффициента песчаности и эффективной толщины, проведен анализ значимости и адекватности полученных моделей, а также построены модельные поверхности для рассмотренных характеристик пласта.

Экономическая эффективность/значимость работы – работа эффективна с позиции финансовой и ресурсной эффективности.

Оглавление

Введение

Обзор литературы

Объект и методы исследования

1 Теоретические основы

1.1 Интерпретация отражающих горизонтов

1.2 Атрибутный анализ

1.3 Последовательность атрибутного анализа при прогнозе ФЕС

1.4 Сейсмические атрибуты

1.5 Регрессионный анализ

1.5.1 Корреляционный анализ

1.5.2 Регрессионные модели

1.5.3 Оценивание значимости регрессионной модели

1.6 Основы нейронных сетей

2 Результаты проведенного исследования

2.1 Корреляция отражающих горизонтов и разломов

2.2 Регрессионные модели

2.2.1 Модель для коэффициента пористости

2.2.2 Модель для коэффициента песчаности

2.2.3 Модель для эффективной толщины пласта

2.3 Модели нейронных сетей

2.3.1 Модель для коэффициента пористости

2.3.2 Модель для коэффициента песчаности

2.3.3 Модель для эффективной толщины пласта

2.4 Поверхности характеристик пласта

3 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

3.1 Предпроектный анализ

3.1.1 Потенциальные потребители результатов исследования

3.1.2 Анализ конкурентных технических решений

- 3.1.3 SWOT-анализ
- 3.2 Планирование управления научно-техническим проектом
 - 3.2.1 Иерархическая структура работ проекта
 - 3.2.2 Контрольные события проекта
- 3.3 Бюджет научного исследования
- 3.4 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования
 - Оценка сравнительной эффективности исследования
- 4 Социальная ответственность
 - 4.1 Анализ опасных и вредных факторов
 - 4.2 Обоснование и разработка мероприятий по снижению уровней опасного и вредного воздействия и устранению их влияния на работающих
 - 4.2.1 Требования к ПЭВМ и организация работы
 - 4.3 Электробезопасность
 - 4.4 Пожарная и взрывная безопасность
- Заключение
- Список публикаций студента
- Список используемой литературы
- Приложение А

Введение

Благодаря высокой энергоемкости и транспортабельности, с середины XX века нефть служит самым важным источником энергии в мире. На производство топлива идет до 84% добываемого сегодня объема. Остальные 16% служат сырьем для переработки в пластмассы, растворители, удобрения, лекарственные средства и прочую продукцию, без которой современная цивилизация просто невозможна. В связи с этим добыча нефти является важной задачей современного мира.

Одним из широко применяемых геофизических методом исследования земной коры, который позволяет выделить месторождение полезных ископаемых (в основном нефти и газа) и выявить перспективные структуры и дальнейшие рекомендации по геологоразведочным работам, является сейсморазведка.

При формировании геологических представлений о строении месторождения используются данные разной точности, размерности, физической природы. Максимально надежные результаты совместного использования всей геологической информации могут быть достигнуты путем применения различных статистических методик, связывающих различные данные между собой. В работе предполагается сопоставление двух- и трехмерной информации, полученной по результатам сейсмических исследований, и одномерной скважинной информацией, полученной геофизическими исследованиями в пробуренных скважинах.

В связи с этим цель практики заключается в характеристике трехмерного строения нефтяного месторождения, используя методы математической статистики.

Для достижения поставленной цели необходимо решить следующие задачи:

- прокоррелировать необходимые отражающие горизонты и разломы;
- рассчитать сейсмические атрибуты;

- оценить степень влияния атрибутов на исследуемый параметр пласта и выявить наиболее значимые;
- получить модели для прогноза распределения параметров пласта в зависимости от сейсмических атрибутов;
- оценить достоверность полученных моделей и выбрать оптимальные;
- построить поверхности рассматриваемых параметров нефтенасыщенного пласта.

На сегодняшний день для проведения атрибутивного анализа сейсмических данных используются регрессионный анализ и метод нейронных сетей.

Объектом исследования выбраны совокупность скважин и пласт нефтяного месторождения.

В качестве предмета исследования выступает оценка параметров нефтенасыщенного пласта.

Обзор литературы

Для геологоразведочного процесса сейсмические атрибуты универсальны, так как используются на всех стадиях геологоразведочного процесса. Они помогают при трассировании горизонтов и разломов, а также необходимы при подсчете запасов на месторождении. Такие работы, как Kurt J. Marfurt, Satinder Chopra «Seismic Attribute for Prospect Identification and Reservoir Characterization» и Kurt J. Marfurt «Seismic Attribute Mapping of Structure and Stratigraphy (SEG, EAGE)», подробно описывают не только сами атрибуты, но и их использование.

Работы таких авторов, как Н.Н.Боровко, Ф.Грейбилла, У.Крамбейна, Р.Миллера, Д.Кана, содержат применение математических методов в качестве вспомогательных средств в процессе построения структурных и фациальных карт. Кроме того, И.П.Шарапова, А.Б.Вистелиуса, Д.Н.Родионова, В.В.Бондаренко, Дж.С.Дэвиса и многие другие в своих исследованиях освещают статистические методы обработки геологической информации.

Д.П.Криге и Ж.Матерон создали теорию геостатистики, работая над получением пространственных переменных на месторождении полезных ископаемых. Далее теория геостатистики была развита у таких авторов, как А.Карлье, В.И.Щеглов, М.Давид и Ю.Е.Капутин.

Условно принято выделять четыре периода в использовании математических методов в геологии. С начала XIX в. до 30-х годов XX в. – первый период, характеризующийся началом развития данного исследования. Выделяются единичными исследованиями.

1930-1965 гг. являются вторым периодом, в течение которого математические и статистические методы начали широко применяться в геологии.

Третий период приурочен к появлению ЭВМ после 1965 г., так как ЭВМ увеличили число решаемых с помощью математических методов задач.

С 1990 г. каждому геологу стали доступны ПК, так как к этому времени они уже получили широкое распространение, тем самым обеспечили оперативную обработку геологической информации. В связи с этим выделяется четвертый период.

В настоящее время в геологии математические методы, в том числе статистические, применяются для:

1) накопление, хранение и систематизация (сортировка, получение выборок и пр.) геологической информации с целью более полного и быстрого ее использования;

2) обработка геологической информации преимущественно на базе методов теории вероятностей и математической статистики для описания, сравнения, классификации геологических объектов и прогнозирования их свойств;

3) математическое моделирование геологических объектов и явлений для решения научных и прикладных задач;

4) автоматизация технологических операций, распространенных в геологии и горном деле, таких как построение геологических карт и разрезов, подсчет запасов и ресурсов, проектирование разведочных и эксплуатационных работ и др.

Объект и методы исследования

При формировании геологических представлений о строении месторождения используются данные разной точности, размерности, физической природы. Максимально надежные результаты совместного использования всей геологической информации могут быть достигнуты путем применения различных статистических методик, связывающих различные данные между собой. В работе предполагается сопоставление двух- и трехмерной информации, полученной по результатам сейсмических исследований, и одномерной скважинной информацией, полученной геофизическими исследованиями в пробуренных скважинах.

В качестве основных методов для достижения поставленной цели были использованы такие статистические методы исследования зависимостей, как регрессионный анализ и метод нейронных сетей.

Объектом исследования выбраны совокупность скважин и пласт нефтяного месторождения.

В качестве предмета исследования выступает оценка параметров нефтенасыщенного пласта, такие как средняя пористость по коллектору, коэффициент песчаности и эффективная толщина.

Корреляция отражающих горизонтов, расчет сейсмических атрибутов и построение поверхностей по полученным моделям осуществлялись в ПО Petrel; статистическая обработка, анализ данных и построение моделей производились с использованием системы Statistica.

1 Теоретические основы

1.1 Интерпретация отражающих горизонтов

Важность корреляции отражающих горизонтов по сейсмическим данным переоценить достаточно трудно. С одной стороны, результаты корреляции разломов и горизонтов (совместно с созданием скоростной и структурной моделей) определяют структурный каркас месторождения. С другой – они являются базой для динамического анализа волнового поля, включая инверсию. Соответственно, ошибки в корреляции отражающих горизонтов критическим образом влияют не только на точность построения структуры резервуара, но и на точность оценки изменения динамических характеристик волнового поля.

Под корреляцией в сейсморазведке понимается процесс выделения и прослеживания отражающих горизонтов, различных сейсмофациальных комплексов (риффы и др) во времени/глубине и в пространстве, на сейсмограммах и суммарных временных и глубинных сейсмических данных.

В процессе прослеживания отражающих горизонтов используют совокупность кинематических и динамических сейсмических атрибутов. При их комплексном анализе осуществляется корреляция отражающих границ волнового поля в пространстве путем прослеживания наиболее выраженных экстремумов (или перехода через 0) волнового поля, при этом главным образом учитывают признак подобия соседних сейсмических трасс.

Одновременно принимают во внимание свойство плавности изменения времени регистрации прихода волны. Линию, соединяющую характерные особенности (экстремумы) одной и той же волны на разных трассах, принято называться осью синфазности. Отраженные волны обычно коррелируются по наиболее четким экстремумам (фазам). При этом интерпретаторы обычно придерживаются принципа – от более надежного к менее надежному.

Корреляция сейсмических данных разделяется на две части: корреляция отражающих горизонтов и тектонических нарушений. Эти две задачи могут выполняться одновременно. Однако для более наглядного примера мы рассмотрим их по порядку. Вначале проследим горизонты. Сначала те, которые в изучаемом районе работ уверенно прослеживаются на значительной площади и имеют надежную геологическую привязку. Такие отражающие горизонты принято называть опорными или реперными. Они являются региональными маркерами. Их прослеживание и интерпретация позволяют существенно повысить понимание всего сейсмического материала, тектонической истории, условий осадконакопления.

Затем выполним корреляцию менее устойчивых отражений, при этом эффективно используя выравнивание разреза на близко расположенное к исследуемому горизонту опорное отражение, о котором известно, что оно было поверхностью выравнивания (т.е. уровнем максимального затопления).

При наличии в разрезе сбросов, взбросов, надвигов и других нарушений на разрезах появляются характерные особенности записи – разрыв сплошности фаз (его называют нарушением оси синфазности), значительное затухание амплитуды без разрыва сплошности фазы, появление сходных по форме отраженных волн с заметным сдвигом по времени.

В зависимости от характера тектонического нарушения оси синфазности, прослеживаемые по обе стороны от него. Могут либо перекрываться, либо образовывать зону отсутствия отражений волнового поля. Иногда зоны нарушений выделяются по затуханию колебаний. Часто в подобных зонах вообще не удается проследить отражающий горизонт.

Целесообразно выполнять корреляцию, следуя от центральных (ненарушенных) частей тектонических блоков, подходя одноименным отражением к плоскости тектонического нарушения с разных сторон. Вообще процесс корреляции отражений и картирования тектонических нарушений – циклический: уточняя корреляцию горизонтов, мы уточняем положение тектонических нарушений, и наоборот.

Современные программные средства позволяют выполнять автоматического прослеживание отражающих горизонтов, основанное на вычислении степени подобия трасс в определенном окне с учетом сдвига трасс относительно друг друга. Естественно, что в областях неустойчивости отражений и сложной интерференционной картины автоматическое прослеживание затрудняется, тогда корреляция отражающих горизонтов выполняется вручную.

Иногда по результатам стратиграфической привязки одной и той же геологической границе в разных скважинах соответствуют различные части сейсмического колебания. Тогда рекомендуется прослеживать наиболее устойчивую часть отражения, а затем переходить от сейсмической границы к геологической, учитывая невязку по площади в разных скважинах.

В отдельных случаях корреляция сейсмических отражений по экстремумам фаз невозможна, и приходится выполнять корреляцию, используя динамические особенности волнового поля. Например, при картировании границы раздела осадочного чехла (на нем мы наблюдаем наличие отражений) и фундамента, глинистого или соляного диапира (хаотичное волновое поле, абсолютно не когерентные друг другу сейсмические отражения, очень слабые амплитуды отражений).

В сложных геологических разрезах сейсмическая корреляция позволяет уточнить положение стратиграфических отбивок (маркеров) пластов на скважинах (т.е. скорректировать начальный вариант положения отбивок).

1.2 Атрибутный анализ

Под сейсмическими атрибутами понимают как конкретные измерения геометрических, кинематических и динамических характеристик волнового

поля, так и различные статистические оценки, полученные с помощью сейсмических данных, т.е. некоторые производные от сейсмических данных.

Сейсмические атрибуты используются на всех стадия геологоразведочного процесса, а также на стадии разработки месторождений – для контроля разработки.

В 60-х годах прошлого столетия с появлением цифровой записи полевых сейсмических данных возникла технология «яркого пятна» («brightspot»), которая связывала наличие изолированных сильных отражений в сейсмическом поле с локальными изменениями литологии или флюидонасыщения пластов, в первую очередь – с наличием свободного газа.

Широкое внедрение цифровой записи метода ОГТ резко увеличило точность оценок сейсмических атрибутов и их выделения на фоне шумовой компоненты сейсмической записи, и, как следствие, увеличило их информативность для использования для интерпретации геологических результатов работ. В 70-80-х годах появилась возможность визуализации сейсмической записи в цвете, что резко повысило информативность сейсмических данных. Была изучена и показана практическая возможность использования преобразования Гильберта для расчета комплексных трасс и мгновенных атрибутов сейсмических трасс.

Другим важным решением было предложение конвертации сейсмических атрибутов в акустический импеданс, более тесно связанный с физическими свойствами пород и насыщающих их флюидов (сейсмической инверсии).

Получило свое развитие и геологическое направление в интерпретации сейсмических данных – ассоциация различных видов рисунка сейсмической записи на сейсмических разрезах с разными условиями осадконакопления: сейсмическая стратиграфия, связь геометрических атрибутов (угол, азимут) с нарушениями.

Анализ рисунка сейсмической записи и изменчивости сейсмических атрибутов на погоризонтных (стратиграфических) слайсах и увязка его с

разными условиями осадконакопления дали развитие сейсмофациальному анализу. Важно было также создание метода AVO анализа и его применение как для поисков, так и для разработки.

В 90-х годах появились анализ когерентности, многомерный статистический анализ, методы классификации сейсмических трасс, появилась возможность оценки сейсмического импульса, началось применение нейронных сетей.

В поздних 90-х годах XX и в начале XXI века возникла технология спектральной декомпозиции, продолжила развитие технология сейсмической инверсии и анализа скоростей. В настоящее время появившиеся возможности широко и полно азимутальной съемки позволяют оценивать сейсмические атрибуты не только скалярно (в точке), но и векторно (изменения по разным направлениям).

Существует три основных результата использования сейсмических атрибутов.

Первый – это разумная экстраполяция и интерполяция в областях с низкой плотностью скважин скважинных измерений ФЕС пород на основе гораздо более плотных сейсмических наблюдений. Инструменты прогноза включают в себя технологии многомерной статистики, геостатистики и нейронных сетей.

Отметим, что при использовании статистических методик прогноза (регрессионный и нейросетевой анализы) необходимо помнить, что результаты прогноза корректны, строго говоря, только в тех областях изменчивости сейсмических атрибутов, с которых была получена обучающая выборка. Так, связь акустического импеданса и пористости, полученная по данным ГИС на разбуренном поднятии, будет некорректна для применения в существенно погруженной области (хотя может быть скорректирована с использованием нормальных трендов уплотнения).

Второй – использование сейсмических атрибутов для реконструкции тектонической, аккумулятивной и диагенетической истории резервуара, что позволяет сделать вывод о закономерностях изменений литологии и ФЭС.

Третий – применение атрибутов для решения собственно традиционных сейсмических задач корреляции горизонтов, трассирования нарушений. Сюда же можно отнести прогноз наличия залежей УВ на этапе поисков.

Эти три варианта использования сейсмических атрибутов относятся к созданию статической модели резервуара. Многократные во времени сейсмические наблюдения (4D seismic) позволяют использовать атрибуты и для контроля разработки и уточнения динамической модели резервуара.

Пространственная экстраполяция и интерполяция в областях с низкой плотностью скважин скважинных измерений ФЭС пород с использованием сейсмических данных основывается на атрибутом анализе.

Атрибутный анализ при прогнозе ФЭС – анализ динамики (изменений) атрибутов волнового поля в межскважинном пространстве и прогноз свойств резервуара на основании связей, установленных между скважинными характеристиками (пористость, эффективная толщина, песчанность и т.д.) и свойствами акустических волн (амплитуда, частота, фаза) в точках скважин при двумерном анализе (по картам атрибутов) и вдоль траектории скважин (по кубам).

Физическая основа метода заключается в том, что атрибуты сейсмических волн связаны со свойствами пород: амплитуды отражений связаны с контрастами акустических жесткостей на границе слоев, частоты связаны с мощностями пластов и их литологическим составом, фазы реагируют на характер переслаивания (экстремумы волны, а не фазы, связаны с контрастными границами, в то время как смещение отражения по фазе обусловлено интерференцией от толщи переслаивания).

Необходимо помнить, что сейсмические атрибуты в силу ограниченной разрешающей способности сейсморазведки всегда носят

интегральную оценку и характеризую некий интерференционный объем пород со средними эффективными характеристиками.

Отметим, что современные методики детерминистской сейсмической инверсии предоставляют в распоряжение интерпретатора поля продольных импедансов (при акустической инверсии) или поля продольных, поперечных импедансов и плотности (при одновременной AVA инверсии), в значительной мере освобожденные от влияния интерференции сейсмических волн. В результате инверсии мы получаем упругие характеристики среды, которые гораздо теснее связаны с ФЕС, чем результаты атрибутного анализа.

1.3 Последовательность атрибутного анализа при прогнозе ФЕС

Работа начинается с тщательной увязки между собой синтетической скважинной сейсмограммы и реальной сейсмической трассы.

На следующем этапе проводится сейсмическое моделирование (одно- или двумерное), в процессе которого определяются факторы (пористость, толщина и т.д.), влияющие на изменение формы сейсмической трассы, и атрибуты, откликающиеся на изменение этих факторов (амплитуда, частота и т.д.), а также чувствительность различных атрибутов к изменениям ФЕС.

Здесь же определяется величина окна или интервала, в котором проявляется это влияние. Затем производится расчет выбранных атрибутов в выбранном окне. Набор рассчитанных атрибутов сопоставляется со скважинными параметрами при помощи построения графиков, где по оси X выставляются значения сейсмических атрибутов в точках скважин, а по оси Y – значения скважинных параметров.

Выявляются атрибуты, изменения величин которых коррелируются с изменением скважинных параметров. Эти атрибуты также анализируются визуально для определения их геологической обоснованности и соответствия

принципиальной геологической модели, априори, построенной по априорным данным.

В более упрощенном варианте сейсмическое моделирование не выполняется, а рассчитываются сейсмические атрибуты, имеющие физический смысл и априорную связь с определяемыми параметрами ФЕС.

При расчете атрибутов окна выбираются исходя из геологической информации о мощности и размерах искомым объектов. Далее ищутся зависимости атрибутов со скважинными данными, и из всего массива рассчитанных атрибутов выбираются те, которые имеют достаточно высокую степень корреляции с прогнозируемыми параметрами ФЕС.

При отборе сейсмических атрибутов для анализа нужно отбирать атрибуты, по возможности не коррелируемые между собой.

На завершающем этапе производится прогноз ФЕС пластов в межскважинном пространстве с использованием информативных сейсмических атрибутов в качестве тренда.

1.4 Сейсмические атрибуты

Сейсморазведка - геофизические методы исследования земной коры, основанные на регистрации искусственно и естественно возбуждаемых упругих волн, которые вызываются взрывом или ударом, опускаются на глубину земной коры, где происходит их преломление или отражение. Волны частично возвращаются на поверхность Земли, где их регистрируют специальные приборы, показания которых позволяют сделать вывод о составе горных пород, через которые прошла волна.

Сейсморазведка используется для:

– выделения месторождения полезных ископаемых (в основном нефти и газа);

- выявления перспективных структур и дальнейших рекомендаций по геологоразведочным работам;
- уточнения объема запасов и ресурсов углеводородов;
- построения комплексных геологических и гидродинамических моделей с целью уточнения плана разработки месторождений и повышения общей эффективности эксплуатации.

Сейсмические волны относятся к классу акустических волн и могут распространяться в упругих и относительно изотропных средах. [1]

Выделяют два главных типа волн:

- продольные – волны, колебание которых происходит вдоль линии распространения волны;
- поперечные – волны, частицы которых колеблются поперек направления распространения волны.

С понятием волн тесно связана амплитуда. Амплитуда – величина отклонения частиц среды от положения равновесия при прохождении сейсмической волны. С целью уменьшения шумовой высокочастотной составляющей также используется куб амплитуд с обрезанными верхними частотами и усиленными нижними.

Набор усредненных амплитуд с указанными отметками времени прихода волны, попавшими на границу раздела сред в точке с фиксированными латеральными координатами, называется сейсмотрассой.

Вся информация, полученная из сейсмических данных, прямыми измерениями, логическим или опытным путями называется сейсмическими атрибутами. Атрибуты используются интерпретатором в процессе прослеживания разломов, при описании резервуара и прогнозировании коллекторских свойств в межскважинном пространстве. [1]

Для проведения анализа по исходному кубу сейсмических данных были построены кубы значений для следующих атрибутов:

– Эквалайзер (EQ) –отфильтрованные амплитуды, где усилены нижние частоты (до 30 Гц), убраны высокие частоты (от 40 Гц). Используется для поиска зависимости с пористостью.

– Мгновенная фаза (Phase) - характеризует угол фазы в любой точке вдоль сейсмической трассы. Используется для поиска зависимости с эффективной толщиной и коэффициентом песчаности.

– Мгновенная частота (Freq) - описывает скорость изменения мгновенной фазы комплексной сейсмотрассы. Используется для выделения тонких глинистых пропластков и коэффициента песчаности.

– Жесткость отражения (Env) - количественно характеризует полную энергию её вещественной и мнимой части. Используется для нахождения зависимости с эффективной толщиной и коэффициентом песчаности.

Мгновенные атрибуты были вычислены из комплексной трассы, $q(t)$, которая получена из сейсмической трассы, $r(t)$, и преобразования Гильберта, $h(t)$, которое похоже на сдвинутую на 90° по фазе трассу (рисунок 1). Написание комплексной трассы в полярной форме дает нам два основных признака: мгновенную амплитуду, $A(t)$, и мгновенную фазу, $f(t)$. [1]

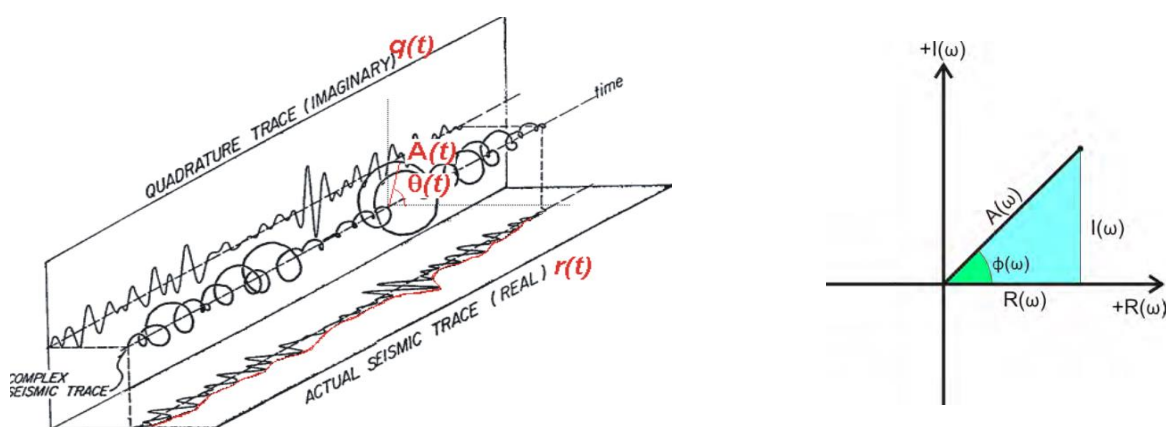


Рисунок 1 – Мгновенные сейсмические атрибуты

Анализ мгновенных параметров основан на преобразовании Гильберта сейсмических записей. Данное преобразование имеет важное преимущество, которое заключается в том, что с помощью простой и доступной вычислительной схемы преобразования, функция, оставаясь во временной области, переводится в комплексный вид, что позволяет оценить ее мгновенные параметры: текущих амплитуд, фаз и частот. Не следует искать физического смысла в самом преобразовании Гильберта. Это всего лишь формальный способ трансформации сигнала, который обеспечивает доступность к удобному для анализа новому представлению функции.[1]

Мгновенная фаза – это характеристика угла фазы в любой точке вдоль сейсмической трассы, рассчитанная для каждого дискрета горизонта. [1]

Фазовая составляющая является основным фактором в определении формы отражающих границ, их геометрической конфигурации и т.д.

Математически, мгновенная фаза определяется как:

$$Q(t) = \operatorname{arctg} \frac{q(t)}{r(t)}$$

где:

$Q(t)$ – мгновенная фаза ,

$r(t)$ – реальная трасса,

$q(t)$ – мнимая часть сейсмической трассы, Гильберт трансформанта от $f(x,t)$.

Мгновенная фаза подчеркивает пространственную неоднородность отражений, не зависит от амплитуды и поэтому на разрезах и срезах мгновенных фаз слабые сигналы становятся более отчетливыми.

Мгновенная фаза наиболее информативна при картировании зон:

- Выклиниваний;
- угловых несогласий;
- разломов;
- литологических ловушек;
- границ пластов и др.

Мгновенная частота - это первая производная мгновенной фазы, которая описывает скорость изменения фазы комплексной трассы. [1]

$$W(t) = \frac{dQ(t)}{dt}$$

Мгновенная Частота наиболее информативна при картировании зон:

- Индикатор углеводорода – низкочастотная аномалия.
- Индикатор зоны трещиноватости, проявляется как низкочастотные зоны.
- Индикатор толщины пласта. Более высокие частоты указывают четкие границы или тонкие глинистые пропластки, низкие частоты указывают богатые песчаные пропластки.
- Индикатор отношения песок/глина.

1.5 Регрессионный анализ

1.5.1 Корреляционный анализ

Корреляционный анализ – это совокупность методов устанавливающих наличие связи между случайными величинами и оценивающих степень этой связи. [2]

Составить представление о наличии и форме зависимости между случайными величинами ξ и η можно по характеру расположения точек корреляционного поля. Предположение о наличии связи между ξ и η можно проверить, например, используя критерий χ^2 Пирсона.[2]

В качестве характеристики силы линейной связи между двумя случайными величинами используется коэффициент корреляции [2]:

$$\rho_{\xi,\eta} = \rho(x, y) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}} = \frac{M((\xi - M(\xi))(\eta - M(\eta)))}{\sigma_{\xi}\sigma_{\eta}}. \quad (1)$$

В качестве оценки коэффициента корреляции между величинами η и ξ , можно взять выборочный коэффициент корреляции :

$$\bar{\rho} = \bar{\rho}_{\eta, \xi} = \frac{\bar{K}_{\eta, \xi}}{\sqrt{\bar{D}_{\eta} \bar{D}_{\xi}}} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}. \quad (2)$$

Частным коэффициентом корреляции между величинами ξ_i и ξ_j , входящими в систему величин $\{\xi_1, \xi_2, \dots, \xi_k\}$, называется величина равная :

$$\rho_{i,j(J(i,j))} = \frac{M(\xi_i - \eta_i)(\xi_j - \eta_j)}{\sqrt{D(\xi_i - \eta_i)D(\xi_j - \eta_j)}}, \quad (3)$$

где $J(i, j) = \{1, 2, \dots, k\} \setminus \{i, j\}$, $\eta_1 = \alpha_0 + \sum_{p \in J(i, j)} \alpha_p \xi_p$, $\eta_2 = \beta_0 + \sum_{p \in J(i, j)} \beta_p \xi_p$, а

коэффициенты $\alpha_0, \alpha_p, \beta_0, \beta_p, p \in J(i, j)$ определяются соответственно из условия минимума функций: $q_i = M(\xi_i - \eta_i)^2$ и $q_j = M(\xi_j - \eta_j)^2$. При этом индексы i, j называют первичными, а индексы $J(i, j) = \{1, 2, \dots, k\} \setminus \{i, j\}$ - вторичными.[2]

Частные коэффициенты корреляции могут быть выражены через ковариации или полные коэффициенты:

$$\rho_{i,j(J(i,j))} = -\frac{A_{ij}}{\sqrt{A_{ii}A_{jj}}} = -\frac{R_{ij}}{\sqrt{R_{ii}R_{jj}}}, \quad (4)$$

где A_{ij}, A_{ii}, A_{jj} - алгебраические дополнения к соответствующим элементам матрицы ковариаций $A(\bar{\xi})$, R_{ij}, R_{ii}, R_{jj} - алгебраические дополнения к соответствующим элементам матрицы корреляций $R(\bar{\xi})$. [2]

В качестве оценок частных коэффициентов корреляции используют выборочные частные коэффициенты корреляции:

$$\bar{\rho}_{i,j(J(i,j))} = -\frac{\bar{A}_{ij}}{\sqrt{\bar{A}_{ii}\bar{A}_{jj}}} = -\frac{\bar{R}_{ij}}{\sqrt{\bar{R}_{ii}\bar{R}_{jj}}}. \quad (5)$$

Если связь имеет выраженный нелинейный характер, то в качестве характеристики связи используется так называемое корреляционное отношение. Пусть $M(\eta/\xi = x) = \varphi(x)$. Функцию $\varphi(x) = M(\eta/\xi = x)$ называют регрессией величины η на величину ξ , а уравнение $y = \varphi(x)$ - уравнением

регрессии. Таким образом, уравнение регрессии характеризует зависимость среднего значения величины η от значений величины ξ . Заметим, что среди всех функций $\tilde{\varphi}(x)$, функция $\varphi(x)$ является наилучшим приближением для значений величины η в смысле среднеквадратичной ошибки, то есть:

$$\min M(\eta - \tilde{\varphi}(\xi))^2 = M(\eta - \varphi(\xi))^2. [2]$$

Если $y = \varphi(x)$ уравнение регрессии величины η на величину ξ , то уравнение связи между η и ξ можно представить в виде: $\eta = \varphi(\xi) + \varepsilon$, где ε - некоторая случайная величина с $M(\varepsilon) = 0$ и $D(\varepsilon) = \sigma_\varepsilon^2$. При этом дисперсию величины η можно представить в виде суммы двух слагаемых: среднего квадрата отклонений значений величины η от значений функции регрессии и среднего квадрата отклонений значений функции регрессии от математического ожидания величины η :

$$D(\eta) = M((\eta - \varphi(\xi))^2) + M((\varphi(\xi) - M(\eta))^2), \quad (6)$$

или:

$$\sigma_\eta^2 = \sigma_\varphi^2 + \sigma_\varepsilon^2. \quad (7)$$

Если $\sigma_\varepsilon^2 = 0$, $\sigma_\eta^2 = \sigma_\varphi^2$, это означает, что вся изменчивость величины η объясняется уравнением регрессии, то есть, между средним значением η и ξ существует функциональная связь. Если $\sigma_\varphi^2 = 0$, то функциональная связь между ξ и средним значением величины η отсутствует. [2]

Корреляционным отношением величины η по величине ξ называется величина, равная

$$r_{\eta,\xi}^2 = \frac{\sigma_\varphi^2}{\sigma_\eta^2} = 1 - \frac{\sigma_\varepsilon^2}{\sigma_\eta^2}. \quad (8)$$

Непосредственно из этого соотношения вытекает, что $0 \leq r_{\eta,\xi}^2 \leq 1$. Причем, при $r_{\eta,\xi}^2 = 0$ вариация среднего значения η при изменении ξ отсутствует, при этом линия регрессии – горизонтальная прямая. Равенство

$r_{\eta, \xi}^2 = 1$ означает, что η и ξ связаны функциональной зависимостью: $\eta = \varphi(\xi)$.

[2]

Корреляционное отношение не что иное, как квадрат коэффициента корреляции между η и $\varphi(\xi)$, а следовательно статистический анализ корреляционного отношения при известной функции регрессии можно свести к анализу коэффициента корреляции.

В качестве характеристики степени линейной зависимости величины ξ_0 от нескольких величин $\xi_1, \xi_2, \dots, \xi_k$ используется множественный коэффициент корреляции. Множественным коэффициентом корреляции между величиной ξ_0 и величинами $\xi_1, \xi_2, \dots, \xi_k$, входящими в систему величин $\{\xi_0, \xi_1, \xi_2, \dots, \xi_k\}$, называется величина равная:

$$\rho_{0(1,2,\dots,k)} = \frac{M(\xi_0 - M(\xi_0))(\eta - M(\eta))}{\sqrt{D(\xi_0)D(\eta)}}, \quad (9)$$

где $\eta = \beta_0 + \sum_{i=1}^k \beta_i \xi_i$, а коэффициенты $\beta_i, i = \overline{0, k}$ определяются из условия минимума функции: $q = M(\xi_0 - \eta)^2$, то есть, величина η является наилучшим линейным приближением величины ξ_0 величинами $\xi_1, \xi_2, \dots, \xi_k$ в среднеквадратичном смысле. Множественный коэффициент корреляции можно определить через элементы корреляционной матрицы R :

$$\rho_{0(1,2,\dots,k)} = \sqrt{1 - \frac{\det R}{R_{00}}}, \quad (10)$$

где R_{00} - алгебраическое дополнение к элементу ρ_{00} матрицы R . [9]

В случае нормального распределения вектора $(\xi_0, \xi_1, \xi_2, \dots, \xi_k)$, функция $\varphi(x_1, x_2, \dots, x_k) = \beta_0 + \sum_{i=1}^k \beta_i x_i$ является функцией регрессии величины ξ_0 на величины $\xi_1, \xi_2, \dots, \xi_k$, то есть $M(\xi_0 / \xi_1 = x_1, \xi_2 = x_2, \dots, \xi_k = x_k) = \varphi(x_1, x_2, \dots, x_k)$. Но тогда, для множественного коэффициента корреляции справедливо:

$$\rho_{0(1,2,\dots,k)}^2 = \frac{D(\varphi(\xi_1, \xi_2, \dots, \xi_k))}{D(\xi_0)} = \frac{\sigma_\varphi^2}{\sigma_{\xi_0}^2}. \quad (11)$$

Величину $\rho_{0(1,2,\dots,k)}^2$ принято называть коэффициентом детерминации. Коэффициент детерминации можно определить аналогичным образом и для любой другой, отличной от линейной, зависимости $\varphi(x_1, x_2, \dots, x_k) = M(\xi_0 / \xi_1 = x_1, \xi_2 = x_2, \dots, \xi_k = x_k)$. Так как коэффициент детерминации не что иное, как квадрат коэффициента корреляции между η и $\varphi(\xi_1, \xi_2, \dots, \xi_k)$, то статистический анализ коэффициента детерминации при известной функции регрессии можно также свести к анализу коэффициента корреляции. [2]

1.5.2 Регрессионные модели

Регрессионный анализ - раздел математической статистики, предназначенный для изучения по выборочным данным статистической зависимости ряда величин; некоторые из которых являются случайными. [2]

Исследование зависимости случайных величин приводит к моделям регрессии и регрессионному анализу на базе выборочных данных. Теория вероятностей и математическая статистика представляют лишь инструмент для изучения статистической зависимости, но не ставят своей целью установление причинной связи.

Числовые данные, содержащиеся в таблицах, обычно имеют между собой явные (известные) или неявные (скрытые) связи.

Явно связаны показатели, которые получены методами прямого счета, т. е. вычислены по заранее известным формулам.

Связи же второго типа (неявные) заранее неизвестны. Однако необходимо уметь объяснять и предсказывать (прогнозировать) сложные явления для того, чтобы управлять ими. Поэтому специалисты с помощью наблюдений стремятся выявить скрытые зависимости и выразить их в виде

формул, т. е. математически смоделировать явления или процессы. Одну из таких возможностей предоставляет регрессионный анализ.

Пользуясь методами регрессионного анализа, аналитики измеряют тесноту связей показателей с помощью коэффициента корреляции. При этом обнаруживаются связи, различные по силе (сильные, слабые, умеренные и др.). Если связи окажутся существенными, то целесообразно будет найти их математическое выражение в виде регрессионной модели и оценить статистическую значимость модели[2].

Регрессионный анализ называют основным методом современной математической статистики для выявления неявных и завуалированных связей между данными наблюдений.

Регрессионный анализ – это статистический метод исследования функциональной связи величины η от переменных ξ_j , $j = \overline{1, k}$, рассматриваемых в регрессионном анализе как неслучайные величины, независимо от истинного распределения этих величин. Регрессионный анализ включает в себя: определение (подбор) класса функций, связывающих результативный показатель (среднее значение величины η и переменные ξ_j); оценку неизвестных параметров уравнения связи; анализ значимости и точности полученного уравнения.

Обычно в регрессионном анализе исходят из следующей модели, устанавливающей связь между входными переменными $\xi_1, \xi_2, \dots, \xi_k$ и выходной переменной η :

$$\eta = f(\xi_1, \xi_2, \dots, \xi_k) + \varepsilon, \quad (12)$$

где $f(x_1, x_2, \dots, x_k) = M(\eta / \xi_1 = x_1, \xi_2 = x_2, \dots, \xi_k = x_k)$ – функция регрессии величины η на величины $\xi_1, \xi_2, \dots, \xi_k$, ε – случайная величина с математическим ожиданием равным нулю и дисперсией σ^2 , в общем случае зависящей от значений величин $\xi_1, \xi_2, \dots, \xi_k$. Уравнение $y = f(x_1, x_2, \dots, x_k)$ называют уравнением регрессии. Величину σ^2 называют остаточной дисперсией или дисперсией предсказания [2].

Пусть функция $f(\bar{x})$ представима в виде линейной комбинации некоторого набора линейно независимых базисных функций $\phi_i, i = \overline{0, m}$:

$$f(\bar{x}, \vec{\beta}) = \sum_{i=0}^m \beta_i \phi_i(\bar{x}), \quad (13)$$

где функции $\phi_i, i = \overline{0, m}$ предполагаются заданными. Поскольку данное уравнение является линейным относительно неизвестного вектора параметров $\vec{\beta}$, то в этом случае говорят о модели регрессии, линейной по параметрам.

Наряду с линейными моделями вида используют и нелинейные, например, полиномиальную, логарифмическую, инверсионную, тригонометрическую, степенную и другие модели. Их применяют, если в результате анализа данных получают статистически ненадежную регрессионную модель. Чаще всего нелинейный регрессионный анализ используют для анализа трендов сезонных факторов[2].

Задача регрессионного анализа состоит в том, чтобы на основе выборочных данных найти вектор $\vec{\beta}^*$ оценок коэффициентов уравнения регрессии, исходя из заданного критерия качества. Обычно используется для нахождения оценок коэффициентов МНК. Оценки МНК состоятельные, несмещенные и являются наилучшими линейными оценками, то есть имеют минимальную дисперсию среди всех линейных оценок.

Пусть для оценки неизвестных параметров $\beta_j, j = \overline{0, k}$ уравнения регрессии взята случайная выборка объемом n из $k+1$ - мерной случайной величины $(\eta, \xi_1, \xi_2, \dots, \xi_k)$. Тогда

$$Y = XB + \varepsilon, \quad (4)$$

где $Y = (y_1, y_2, \dots, y_n)^T$ - вектор значений переменной η , $B = (\beta_0, \beta_1, \dots, \beta_k)^T$ - вектор параметров модели, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$ - вектор ошибок, а X - матрица значений переменных ξ_j размерами $n \times (k+1)$. Первый столбец матрицы X

содержит единицы (значения фиктивной переменной ξ_0), остальные столбцы значения переменных $\xi_1 - \xi_k$:

$$X = \begin{pmatrix} 1 & x_1^1 & \cdots & x_k^1 \\ 1 & x_1^2 & \cdots & x_k^2 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_1^n & \cdots & x_k^n \end{pmatrix}.$$

Тогда, оценка вектора параметров МНК имеет вид:

$$B^* = (X^T X)^{-1} X^T Y.$$

Для нахождения дисперсии оценок найдем матрицу ковариаций коэффициентов K :

$$\begin{aligned} K &= M[(B^* - M(B^*))(B^* - M(B^*))^T] = M[(B^* - B)(B^* - B)^T] = M[(X^T X)^{-1} X^T \varepsilon](X^T X)^{-1} X^T \varepsilon \\ &= M[(X^T X)^{-1} X^T \varepsilon^T X (X^T X)^{-1}] = (X^T X)^{-1} X^T M[\varepsilon^T] X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1} X^T E X (X^T X)^{-1} = \\ &= \sigma^2 (X^T X)^{-1}. \end{aligned}$$

Так как диагональные элементы матрицы K есть дисперсии оценок коэффициентов β_j^* : $k_{jj} = M[(\beta_j^* - M(\beta_j^*))(\beta_j^* - M(\beta_j^*))] = D(\beta_j^*)$, то:

$$D(\beta_j^*) = \sigma^2 [(X^T X)^{-1}]_{jj}, \quad j = \overline{0, k}. \quad (15)$$

Несмещенная оценка дисперсии предсказания σ^2 есть:

$$s^2 = \frac{1}{n - k - 1} \sum_{i=1}^n e_i^2, \quad (16)$$

где $e = Y - XB^*$.

Соответственно несмещенные оценки дисперсий оценок параметров модели:

$$s_j^2 = \frac{1}{n - k - 1} [(X^T X)^{-1}]_{jj} \sum_{i=1}^n e_i^2. \quad (17)$$

Имеем $\hat{y}(X_0) = X_0 B^*$, тогда $M(\hat{y}) = M(X_0 B^*) = X_0 M(B^*) = X_0 B = \tilde{y}$.

Дисперсию $\hat{y}(X_0)$ найдем, учитывая, что дисперсия линейной комбинации

$\alpha^T \xi = \alpha_1 \xi_1 + \alpha_2 \xi_2 + \dots + \alpha_k \xi_k$ случайного вектора ξ равна $D(\alpha^T \xi) = \alpha^T K_\xi \alpha$, где

K_ξ - матрица ковариаций вектора ξ . Тогда

$$D(\hat{y}) = D(X_0 B^*) = X_0^T K(B^*) X_0 = \sigma^2 X_0^T (X^T X)^{-1} X_0.$$

Для построения доверительного интервала для \tilde{y} в точке $X_0 = (1, x_1^0, x_2^0, \dots, x_k^0)$ используем статистику $t = \frac{\hat{y} - \tilde{y}}{\sqrt{s^2 X_0^T (X^T X)^{-1} X_0}}$, где $\hat{y}(X_0) = X_0 B^*$.

Тогда доверительный интервал уровня α будет иметь вид: $\left[\hat{y}(X_0) \pm \tau_{1-\alpha/2} \sqrt{s^2 X_0^T (X^T X)^{-1} X_0} \right]$, где $\tau_{1-\alpha/2}$ - квантиль распределения Стьюдента.

Для оценки качества линейной модели можно использовать значение коэффициента детерминации:

$$R^2 = 1 - \frac{D(\varepsilon)}{D(y)}, \quad (18)$$

он показывает, какую долю общей дисперсии объясняет регрессионная модель. Если $R^2 = 1$, то $D(\varepsilon) = 0$, т.е. значения y однозначно определяются уравнением регрессии. Если $R^2 = 0$, то $D(y) = D(\varepsilon)$, то есть дисперсия отклика равна дисперсии неучтенных составляющих и линейная связь между y и x_1, x_2, \dots, x_k отсутствует. Считается что, если $0,01 \leq R^2 < 0,1$, то связь между $\{x_1, x_2, \dots, x_n\}$ и y слабая; если $0,1 \leq R^2 < 0,5$, то говорят о наличии средней связи, при $R^2 \geq 0,5$ связь сильная [2].

1.5.3 Оценивание значимости регрессионной модели

Оценивание регрессионной модели состоит из статистического оценивания регрессионной модели и статистического оценивания надежности коэффициентов регрессии [2].

1. Для проверки значимости модели используется критерий Фишера-Снедекора. Для этого находим оценки коэффициента детерминации

$$R^{2*} = \frac{(XB^*)^T (XB^*)}{Y^T Y} \quad \text{и} \quad \text{остаточной дисперсии}$$

$s^2 = \frac{1}{n-k-1} \sum_{i=1}^n e_i^2 = \frac{1}{n-k-1} (Y - XB^*)^T (Y - XB^*)$ и находим наблюдаемое значение

статистики критерия $F = \frac{R^{2*}}{1-R^{2*}} \frac{n-k-1}{k}$. Если значение статистики больше, чем

критическое $F_{кр}$, где $F_{кр}$ - критическая точка распределения Фишера уровня α , то уравнение признается значимым. Качество модели оценивается также по значениям R^{2*} и s^2 .

2. Если модель значима, то проверяется значимость каждого из коэффициентов β_j . Для этого находятся дисперсии оценок

$D(\beta_j^*) = \sigma^2 \left[(X^T X)^{-1} \right]_{jj}$, и вычисляется наблюдаемое значение статистики

$t_j = \frac{\beta_j^*}{\sqrt{s^2 \left[(X^T X)^{-1} \right]_{jj}}}$, которое сравнивается с критическим $t_{кр}$, где $t_{кр}$ -

критическая точка распределения Стьюдента уровня $\alpha/2$. Если наблюдаемое значение больше, чем критическое, коэффициент признается значимым.

1.6 Основы нейронных сетей

Нейронные сети – математическая модель, построенная на подобию того, как организованы и функционируют сети нервных клеток живого организма – биологические нейронные сети. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. [3]

Общая модель нейрона состоит в следующем: нейрон состоит из дендритов (каналы ввода информации) и аксона (канал вывода информации).

Синапсы соединяют аксон одного нейрона с дендритами других. При возбуждении нейрон посылает сигнал по своему аксону, который передается по синапсам другим нейронам, которые могут как возбуждаться, так и перейти в состояние торможения.

Существует порог возбуждения или активации (порог b), при превышении которого суммарным уровнем сигналов, происходит возбуждение нейрона. Активность синапсов определяет интенсивность получаемого нейроном сигнала.

На рисунке 2 показана модель нейрона. Из него видно, что на вход нейрона поступают сигналы x , каждый из которых умножается на вес w (для каждого сигнала имеется собственный вес); производится сложение преобразованных сигналов и добавляется порог (блок сумматор на диаграмме); результат преобразуется с помощью функции активации f и подается на выход нейрона. В модели вес моделирует силу синапса. [3]

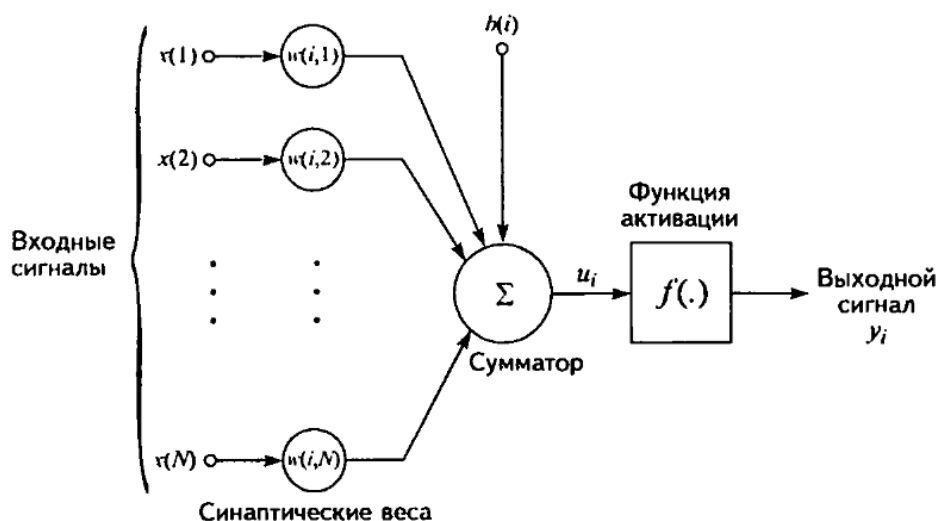


Рисунок 2 - Модель нейрона

Нейрон получает сигналы через несколько входных каналов, они показаны на рис. 2 слева.

Каждый сигнал проходит через соединение – синапс, который имеет определенную силу или вес $w(i,j)$ соответствующий синаптической активности нейрона.

Коэффициенты $w(i,j)$ называются весами синаптических связей, положительное значение которых соответствует возбуждающим синапсам, отрицательные значения – тормозящим синапсам; если $w(i,j)=0$, то говорят, что связь между нейроном i и нейроном j отсутствует. [3]

Текущее состояние нейрона описывается формулой

$$u_i = \sum_{j=1}^N w(i, j)x(j) + b(i),$$

где $x(j), j = 1, 2, \dots, N$ – входные сигналы;

$w(i, j)$ – веса синаптических связей;

$b(i)$ – пороговое значение (порог активации);

Полученный нейроном сигнал преобразуется с помощью нелинейной функции активации или передаточной функции f в выходной сигнал

$$y_i = f(u_i),$$

где f – функция активации.

Индекс i относится к номеру рассматриваемого нейрона в сети, индекс j указывает номер синаптической связи. [3]

При работе сети на входные элементы подаются значения входных данных (входной сигнал), затем возбуждаются нейроны первого промежуточного слоя, второго промежуточного слоя, и т.д.; в итоге преобразованный сигнал поступает на выходной слой.

Преобразование сигнала нейроном сети проводится следующим естественным образом [3]:

- Находится разница взвешенной суммы выходов элементов предыдущего слоя и порогового значения;
- Полученные значения преобразуются с помощью передаточной функции;
- В результате получается выходной сигнал нейрона, который поступает на вход следующего нейрона или подается на выход сети, если нейрон является конечным.

Входными данными при нейросетевом анализе является обучающая выборка (примеры входных значений и соответствующие им выходные значения) и тестирующая выборка. Выходными данными является

информация о весовых коэффициентах, активационных функциях, структуре сети и ошибках при тестировании и обучении.

Обучение сети – это подгонка свободных параметров с целью адаптации к внешним воздействиям.

Тестирование – это проверка полученной сети на примерах, которые не участвовали в обучении. При этом число тестовых примеров должно быть больше, чем выше качество обучения.

2 Результаты проведенного исследования

2.1 Корреляция отражающих горизонтов и разломов

Первоначальной задачей являлась корреляция отражающих горизонтов Bentiu 1, Bentiu 2 и Bentiu 3 в ПО Petrel по сейсмическому кубу данных, представленному на Рисунке 3. Под корреляцией в сейсморазведке понимается процесс выделения и прослеживания отражающих горизонтов, различных сейсмофациальных комплексов (рифты и т.д.) во времени/глубине и в пространстве, на сейсмограммах и суммарных временных глубинах сейсмических данных.

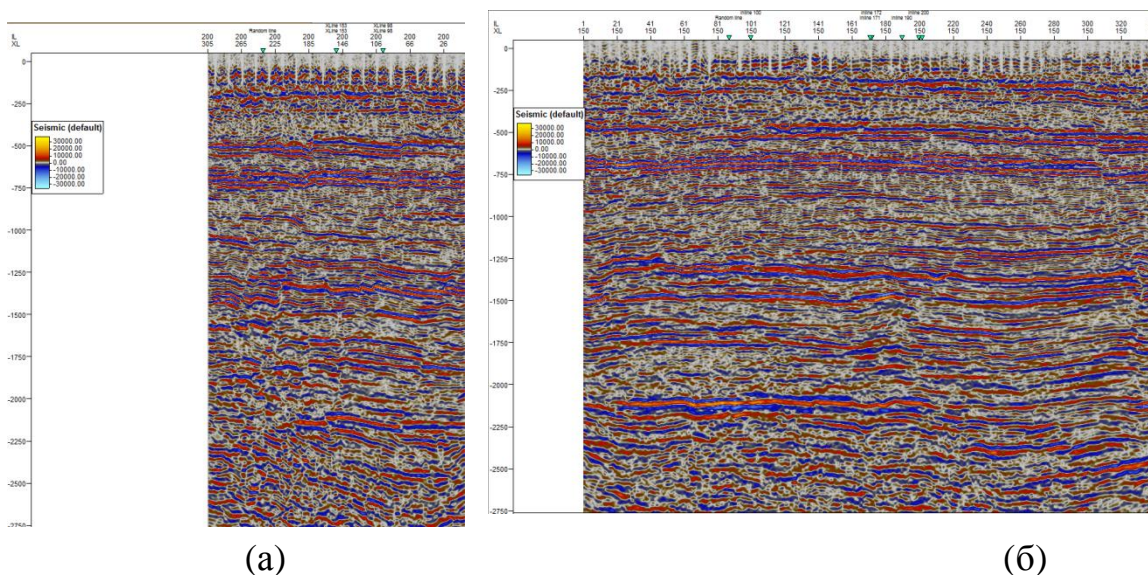


Рисунок 3 – профили сейсмического куба: (а) – Inline 200, (б) – Xline 150

В процессе прослеживания отражающих горизонтов используют совокупность кинематических и динамических сейсмических атрибутов. При их комплексном анализе осуществляется корреляция отражающих границ волнового поля в пространстве путем прослеживания наиболее ярко выраженных экстремумов (или перехода через 0) волнового поля, при этом главным образом учитывают признак подобия соседних сейсмических трасс. Прослеженные горизонты для Inline 200 представлены на Рисунке 4.

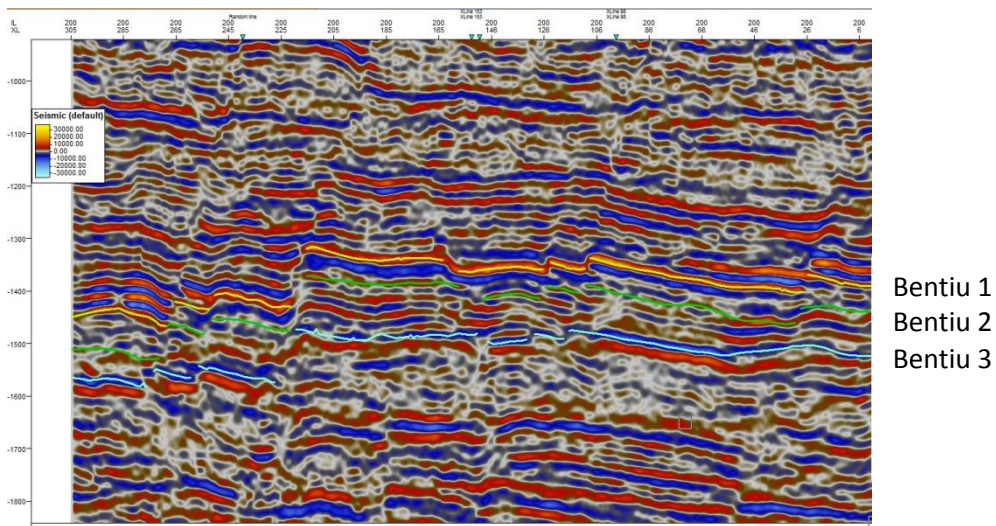


Рисунок 4 – горизонты Bentiu 1, Bentiu 2, Bentiu 3

Был выбран шаг в 2 профиля (то есть через каждые 2 профиля осуществлялось выделение горизонтов). В качестве примера выполнения данной операции на Рисунке 5 представлен «прослеженный» горизонт Bentiu 1. Подобный вид «сетки» имели и остальные горизонты.

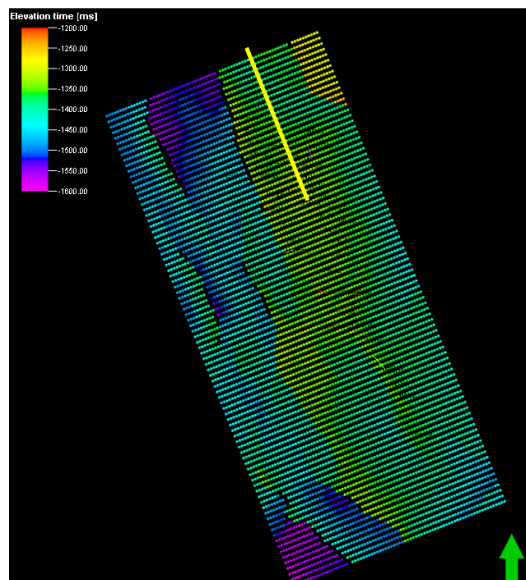


Рисунок 5 – "Прослеженный" горизонт Bentiu 1

Получение непрерывной поверхности горизонтов было осуществлено возможностями ПО Petrel. В качестве примеров на Рисунке 6 представлены Bentiu 2 и Bentiu3.

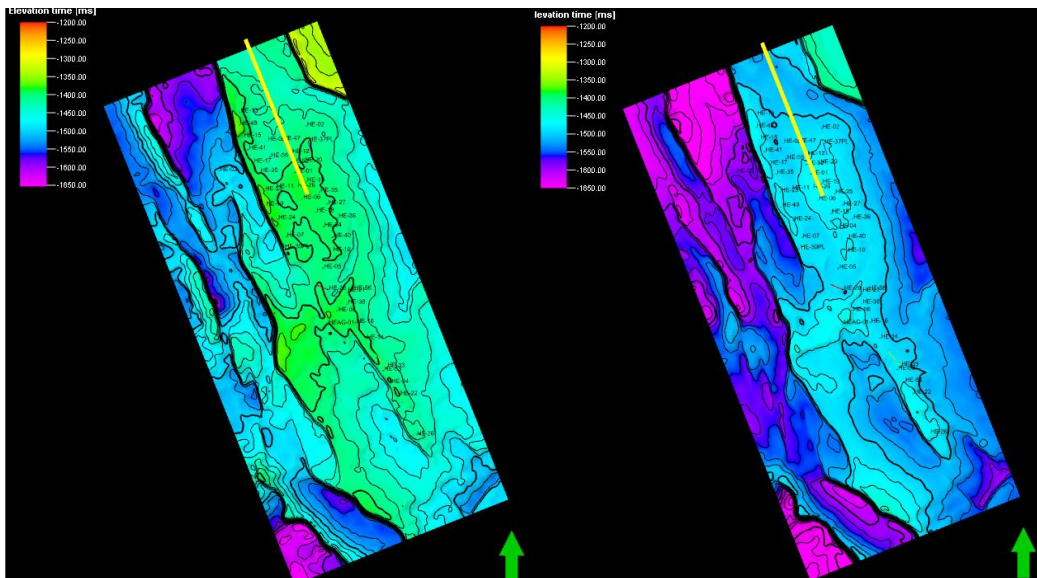


Рисунок 6 – Поверхности горизонтов Bentiu 2 и Bentiu 3

Кроме корреляции отражающих горизонтов, необходимо производить выделение разломов. На Рисунке 7 показано отображение прокоррелированных разломов на Inline 200.

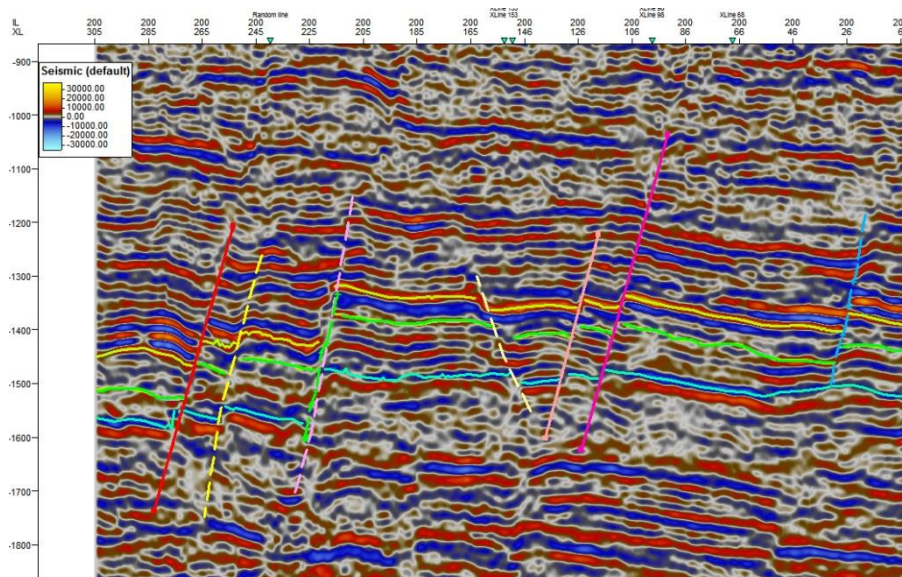


Рисунок 7 – Разломы, Inline 200

Результат корреляции отражающих горизонтов и разломов на следующем этапе служит основой для корректного проведения атрибутивного анализа.

После завершения корреляции, были рассчитаны сейсмические атрибуты для статистического исследования их зависимостей с параметрами пласта.

Результатом атрибутивного анализа сейсмических данных является прогноз петрофизических параметров интересующего интервала недр. Основой анализа служит предположение о наличии взаимосвязей между атрибутами волнового поля куба сейсмических данных и петрофизическими свойствами соответствующих пород. Однако признается, что не существует универсальных связей этих величин. В каждом отдельном случае один и тот же атрибут может оказываться полезным при прогнозировании разных свойств пласта. Наличие зависимостей или их отсутствие может обосновываться множеством причин, как техногенного характера, так и геологического. В тоже время при идеальных условиях предполагается, что ряд атрибутов сейсмической записи с высокой степенью надежности коррелирует с эффективными толщинами, неоднородностями залегания пород, пористостью коллекторов и, даже, трещиноватостью и проницаемостью. Большинство таких идеальных случаев относится к атрибутам, описывающим комплексную составляющую сейсмической трассы, например, ее мгновенные параметры – частоту, фазу. Это связано с тем, что комплексное представление сейсмической трассы и атрибуты, построенные на ее основе, имеют непрерывный характер, как и свойства пород вдоль разреза. А амплитудная составляющая и атрибуты, построенные на ее основе, характеризуют коэффициенты отражения на границах раздела пород с разными акустическими и плотностными свойствами.

Для исследования были выбраны 43 скважины и следующие объемные атрибуты:

– Эквалайзер (EQ) –отфильтрованные амплитуды, где усилены нижние частоты (до 30 Гц), убраны высокие частоты (от 40 Гц). В построенном кубе данных содержится исходная сейсмическая информация, при этом выделение низких частот позволяет выделить основные тренды и

отсеять помехи. Выбор пороговых значений частот согласован с частотой импульса, который использован при создании синтетических сеймотрасс и инверсии куба сейсмических данных. Куб амплитуд является традиционным параметром, используемым в атрибутивном анализе, поскольку несет в себе исходную сейсмическую информацию и существенно зависит от изменения плотности залегающих пород и скорости распространения упругих волн в них, которые, в свою очередь, неразрывно связаны с такими петрофизическими характеристиками пласта, как пористость, песчаность и т.д. Термин песчаность будет иметь смысл отношения мощности продуктивной части пласта к общей мощности пласта, и не соотносится с породами, слагающими пласт.

– Мгновенная фаза (Phase) –еще один параметр, рассчитываемый по комплексному представлению сеймотрассы, характеризующий угол фазы в любой точке вдоль сейсмической трассы. Мгновенная фаза подчеркивает пространственную неоднородность отражений, поэтому может использоваться для поиска зависимостей с такими параметрами пласта, как эффективная толщина и коэффициент песчаности.

– Мгновенная частота (Freq) - описывает скорость изменения мгновенной фазы комплексной сеймотрассы. Используется для выделения тонких глинистых пропластков и коэффициента песчаности.

– Жесткость отражения (Env) –один из параметров, рассчитываемых по комплексному представлению сеймитрассы, количественно характеризующий полную энергию ее вещественной и мнимой части. В различных исследованиях предполагается, что значение данного атрибута может оказаться полезным при поиске зависимостей с эффективными толщинами и коэффициентом песчаности.

Предполагалось, что будет исследоваться пласт, кровлей которого является отражающий горизонт Ventiu 1, а подошвой – Ventiu 2.

Для каждого объемного атрибута осуществлялся расчет семи поверхностных значений для выбранного пласта:

- Максимальное значение атрибута между кровлей и подошвой (Max);
- Среднее значение атрибута между кровлей и подошвой (Mean);
- Минимальное значение атрибута между кровлей и подошвой (Min);
- Корень квадратный отношения суммы квадратов амплитуд к количеству значений между кровлей и подошвой (“RootMeanSquare” - RMS);
- Арифметическое среднее положительных амплитуд, умноженное на число значений в между кровлей и подошвой («Sumofpositiveamplitude» - Sum_pos);
- Половина энергии, которая отображает время или глубину, необходимую для достижения половины всей энергии для атрибута между кровлей и подошвой выбранного пласта(half)
- Длина дуги, которая измеряет неоднородность в коэффициенте отражения, и может быть использована для количественной оценки латеральных изменений в модели отражения(arclength).

В поиске зависимостей участвовали следующие характеристики пласта:

- Эффективная толщина, H_{ef} ;
- Коэффициент пористости (по коллектору), RHI ;
- Коэффициент песчанистости, NTG .

Задача о поиске уравнения не имеет единственного решения, полученные уравнения будут отличаться друг от друга в зависимости от того, какие характеристики приняты при построении регрессионной модели и от конкретных значений выбранных параметров. При этом из всех полученных моделей необходимо выбрать ту, которая позволит наиболее достоверно рассчитать прогнозную карту заданного параметра.

2.2 Регрессионные модели

Для оценки достоверности полученных регрессионных моделей в данной работе использовались следующие критерии:

- Коэффициент детерминации должен быть больше 0.50 (в д.ед.);
- Остатки (разность фактического и расчётного значений) должны иметь распределение, близкое к нормальному (не допускаются всплески на краях построенной гистограммы, максимальная плотность должна достигаться вблизи нуля);
- Уровень значимости (p-level) – показывает степень значимости конкретной входной переменной на результат – должен находиться на уровне 0.05-0.15 (в д.ед.);
- Число параметров в модели не больше трех.

2.2.1 Модель для коэффициента пористости

Наилучшая полученная регрессионная модель для коэффициента пористости имеет следующий вид:

$$PHIE = 0,22939 + 0,00007 * Env_{acrlength} + 0,00040 * Phase_{RMS} - 0,00237 * Phase_{arclength}$$

где $PHIE$ – коэффициент пористости;

$Env_{acrlength}$ – длина дуги жесткости отражения;

$Phase_{RMS}$ – среднекв. значение мгновенной фазы;

$Phase_{arclength}$ – длина дуги мгновенной фазы.

В таблице 1 приведены P-значения (достигнутый уровень значимости) коэффициентов регрессии, с помощью которых можно сказать, что все полученные коэффициенты значимы.

Таблица 1 – р-значения регрессионных коэффициентов

	Значение к-та	р-уровень
Св. член	0,22939	0,00
Env_arc length	0,00007	0,00
Phase_RMS	0,00040	0,01
Phase_arc length	-0,00237	0,00

В таблице 2 отображены основные параметры, используемые для проверки достоверности модели. Коэффициент Фишера определяет данную модель как значимую; коэффициент детерминации, межфакторный коэффициент корреляции и средняя ошибка принимают допустимые значения, что указывает на то, что модель является достоверной.

Таблица 2– Параметры достоверности модели

R^2	0,54	
Фишер/р-уровень	19,91	0,00
Корр. коэф., абс. макс.	0,27	
Средняя ошибка:д.ед/%	0,01	4

Гистограмма, представленная на Рисунке 8, необходима для оценки распределение остатков. Остатки максимальны в нуле и не имеют "всплесков" на краях, что удовлетворяет установленным критериям достоверности модели.

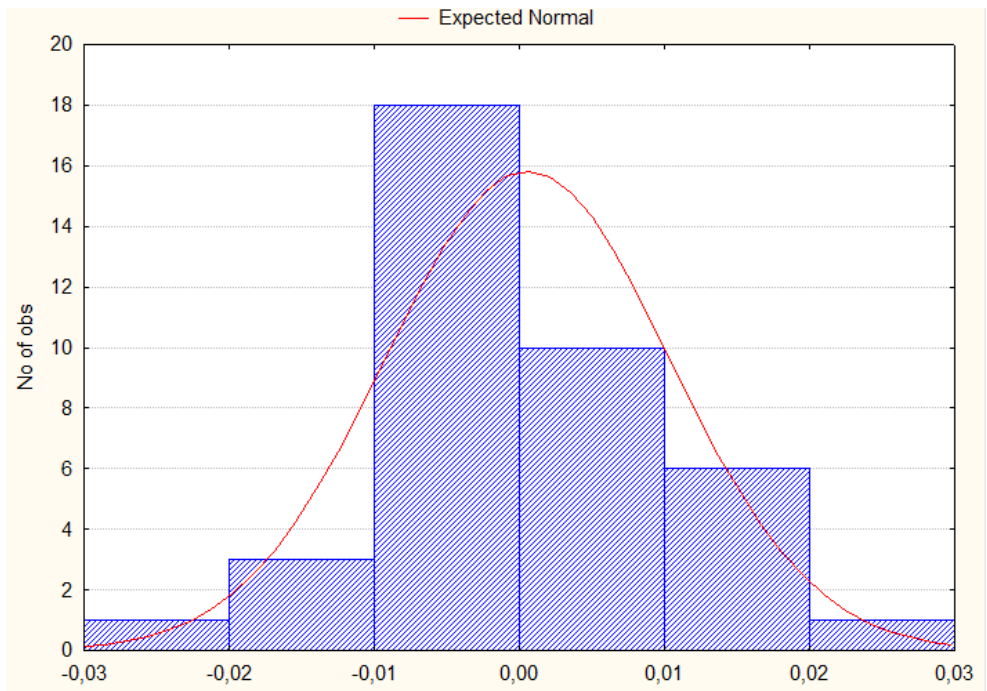


Рисунок 8 – Гистограмма распределения остатков коэффициента пористости

На Рисунке 9 представлены результаты моделирования модели для коэффициента пористости.

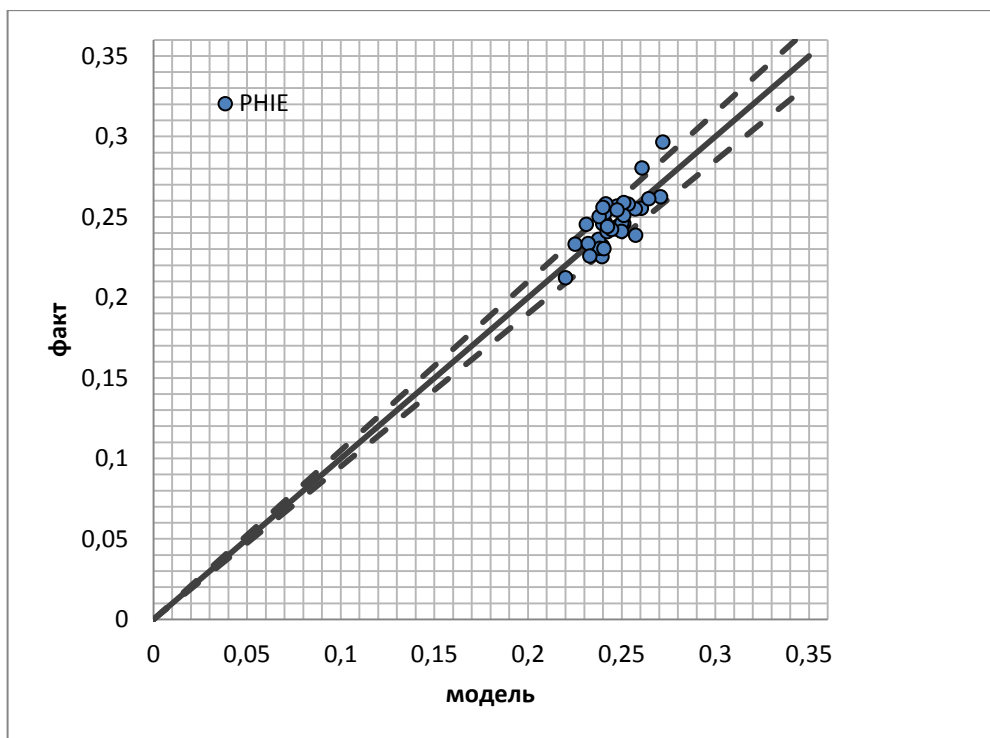


Рисунок 9 – Результат моделирования для коэффициента пористости

Таким образом, из совокупности параметров достоверности модели можно сделать вывод, что данная модель является достоверной и ее можно использовать для прогнозирования.

2.2.2 Модель для коэффициента песчанности

Наилучшая полученная регрессионная модель для коэффициента песчанности принимает следующий вид:

$$NTG = 1,08249 - 0,00004 * Env_{RMS} - 0,11683 * Freq_{arc\ length} - 0.00203 * Phase_{mean}$$

где NTG – коэффициент песчанности;

Env_{RMS} – среднекв. знач. жесткости отражения;

$Freq_{arc\ length}$ – длина дуги мгновенной частоты;

$Phase_{mean}$ – среднее значение мгновенной фазы.

В Таблице 3 приведены Р-значения (достигнутый уровень значимости) коэффициентов регрессии. Так как Р-уровни для всех коэффициентов уравнения меньше, чем 0,05, принимаем их как значимые.

Таблица 3 – р-значения регрессионных коэффициентов коэффициента песчанности

	Значение к-та	р-уровень
Св. член	1,08249	0,00
Env_RMS	-0,00004	0,00
Freq_arc length	-0,11683	0,00
Phase_mean	-0,00203	0,00

В таблице 4 отображены основные параметры, используемые для проверки достоверности модели. Коэффициент Фишера указывает на

значимость самого уравнения, средняя ошибка, коэффициент межфакторной корреляции и коэффициент детерминации принимают допустимые значения.

Таблица 4 – параметры достоверности модели

R^2	0,50	
Фишер/р-уровень	11,40	0,00
Корр. коэф., абс. макс.	0,36	
Средняя ошибка:д.ед/%	0,07	10

На Рисунке 10 представлена гистограмма, отображающая распределение остатков. Из нее видно, что распределение остатков близко к нормальному, удовлетворяя установленные критерии достоверности модели.

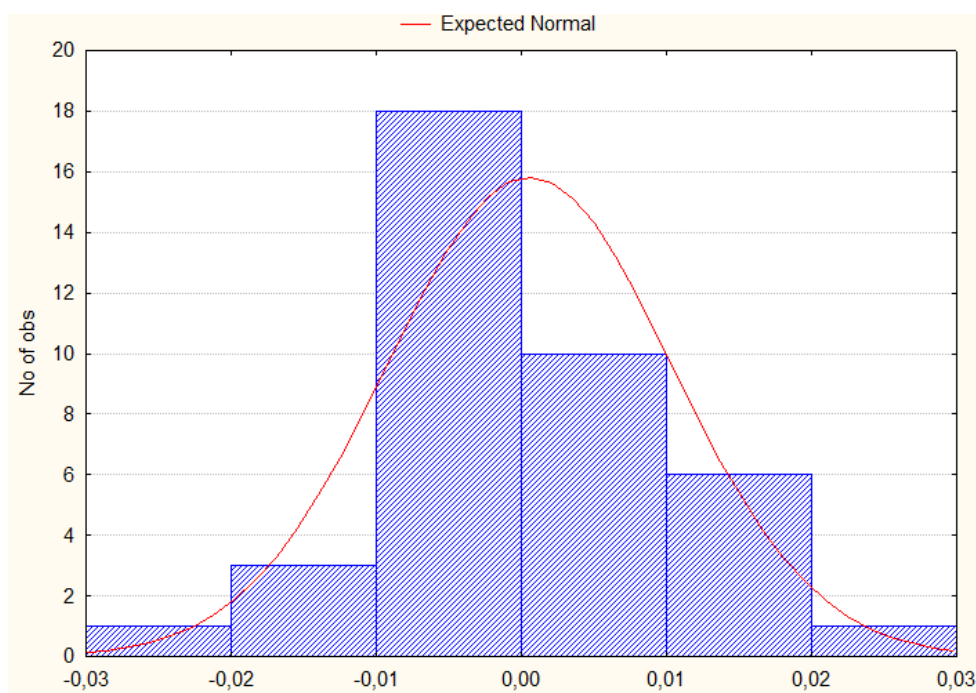


Рисунок 10 – Гистограмма распределения остатков для коэффициента песчанности

Результаты моделирования модели для коэффициента песчанности представлены на Рисунке 11.

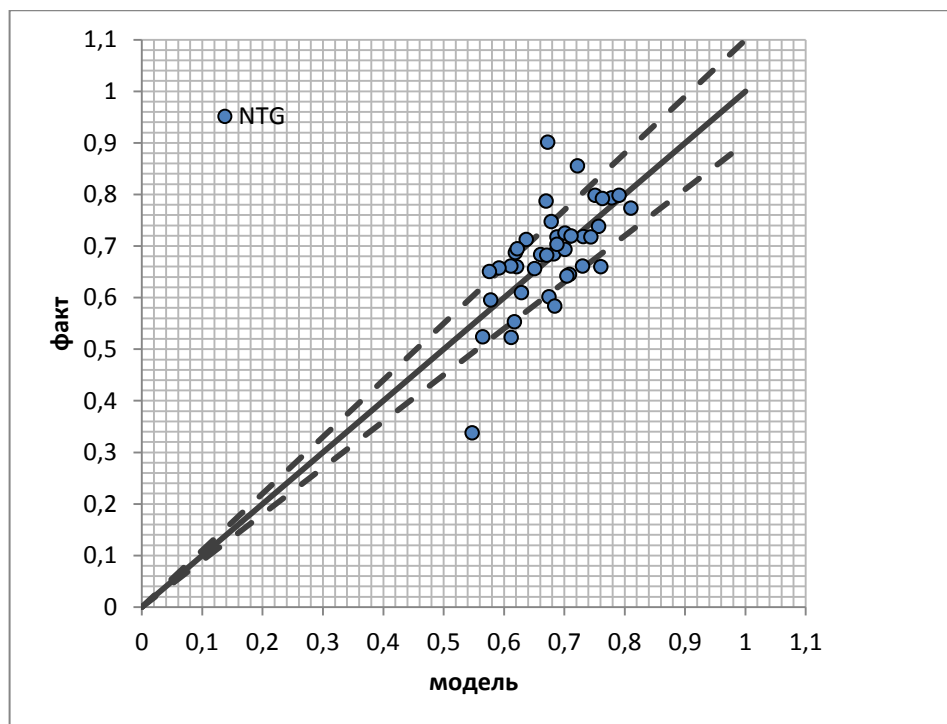


Рисунок 11 – Результат моделирования для коэффициента песчанистости

Таким образом, из совокупности параметров достоверности модели можно сделать вывод, что данная модель является достоверной.

2.2.3 Модель для эффективной толщины пласта

Наилучшая полученная регрессионная модель для эффективной толщины имеет следующий вид:

$$H_{ef} = 60,8583 + 0,5176 * Env_{half} - 0,5000 * Phase_{half} - 0,0004 * EQ_{sum}$$

где H_{ef} – эффективная толщина;

Env_{half} – половина энергии жесткости отражения;

$Phase_{half}$ – половина энергии мгновенной фазы;

EQ_{sum} – сумма амплитуд эквалайзера.

В Таблице 5 приведены р-уровни (достигнутый уровень значимости) коэффициентов регрессии, с помощью которых можно сказать, что все полученные коэффициенты значимы.

Таблица 5 – р-уровни регрессионных коэффициентов для эффективной толщины

	Значение к-та	р-уровень
Св. член	60,8583	0,00
Env_half	0,5176	0,00
Phase_half	-0,5000	0,00
Phase_sum	-0,0004	0,00

Из Таблицы 6 можно сделать вывод, что уравнение удовлетворяет всем вышеуказанным нормам: является значимым, коэффициент детерминации больше 0,5, а средняя ошибка и коэффициент корреляции между факторами не превышают установленные значения.

Таблица 6 – параметры достоверности модели

R^2	0,59	
Фишер/р-уровень	16,50	0,00
Корр. коэф., абс. макс.	0,12	
Средняя ошибка:м/%	8,12	15

На Рисунке 12 представлена гистограмма, отображающая распределение остатков. Из нее видно, что остатки распределены близко к нормальному.

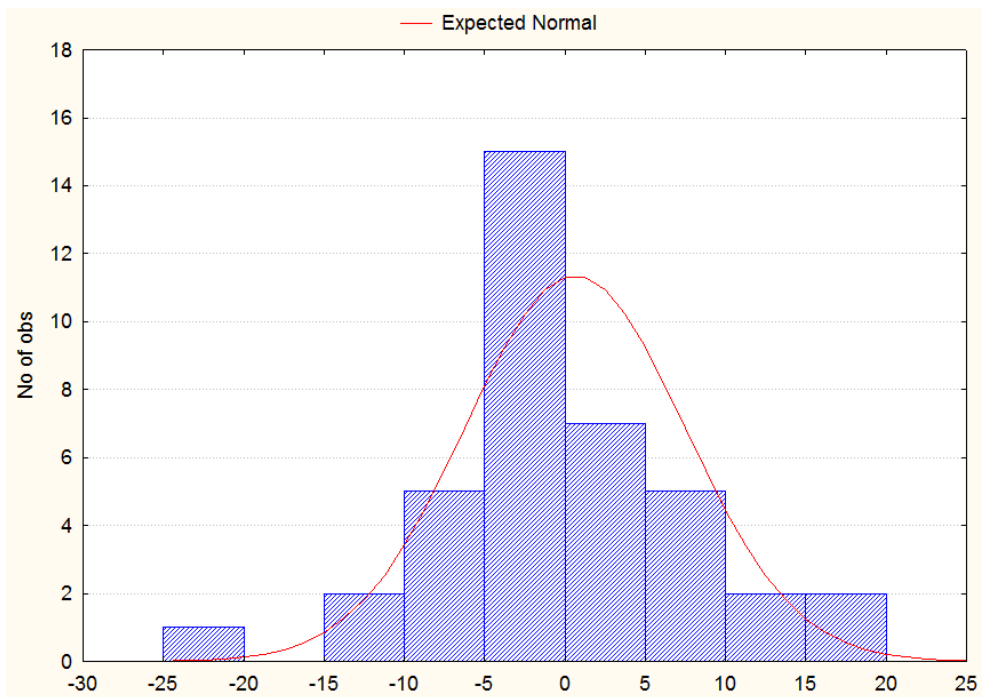


Рисунок 12 – Гистограмма распределения остатков эффективной толщины

На Рисунке 13 представлены результаты моделирования модели для эффективной толщины.

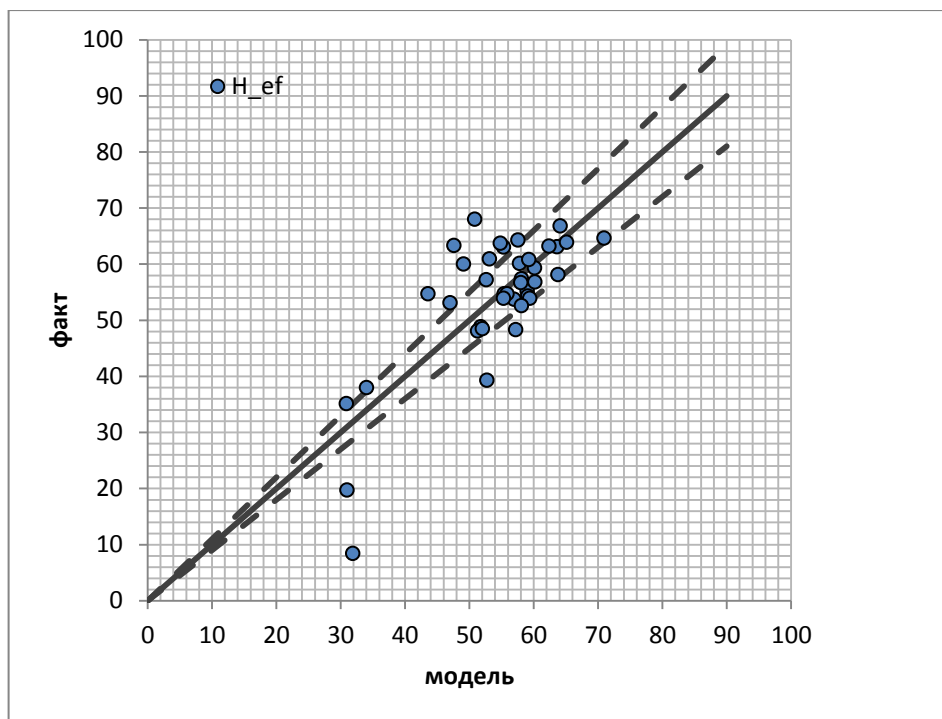


Рисунок 13 – Результаты моделирования для эффективной толщины

Таким образом, из совокупности параметров достоверности модели можно сделать вывод, что данная модель является достоверной и ее можно использовать для прогнозирования.

2.3 Модели нейронных сетей

Для полученных данных также были построены нейросетевые модели. Но из-за малого объема выборки, полученные модели не применимы для прогнозирования.

Основными критериями для выбора лучшей из полученных моделей были:

- Корреляционные коэффициенты обучения и теста должны превышать 0,7.
- Межфакторный коэффициент корреляции должен быть меньше 0,5.
- Средняя ошибка построенной модели должна быть меньше 0,15.
- Распределение остатков должно быть близко к нормальному.

2.3.1 Модель для коэффициента пористости

Наилучшая модель, полученная для коэффициента пористости, имеет следующий вид:

$$PHIE = PHIE(Env_{arc\ length}, Freq_{mean})$$

где $PHIE$ – коэффициент пористости;

$Env_{arc\ length}$ – длина дуги жесткости отражения;

$Freq_{mean}$ – среднее значение мгновенной частоты.

В таблице 7 отражены основные показатели по полученной модели.

Таблица 7 – Показатели нейронной сети для коэффициента пористости

Количество весовых коэффициентов	13	
Схема нейронной сети	2-3-1,Exp+ Logistic	
Корреляционный коэффициент обучение/тест	0,70	0,93
Межфакторный корреляционный коэффициент	0,03	
Средняя ошибка: д.ед./%	0,008	4

Из данной таблицы можно сделать вывод, что корреляционные коэффициенты обучения и теста, межфакторная корреляция и средняя ошибка модели принимают допустимые значения.

На рисунке 14 представлены гистограммы распределения остатков по обучающей выборке (рис 14(а)) и по всему объему выборки (рис. 14(б)), исходя из которых видно, что остатки как обучающей, так и всей выборки имеют максимум в нуле и не имеют "вылеты", но данные не симметричны.

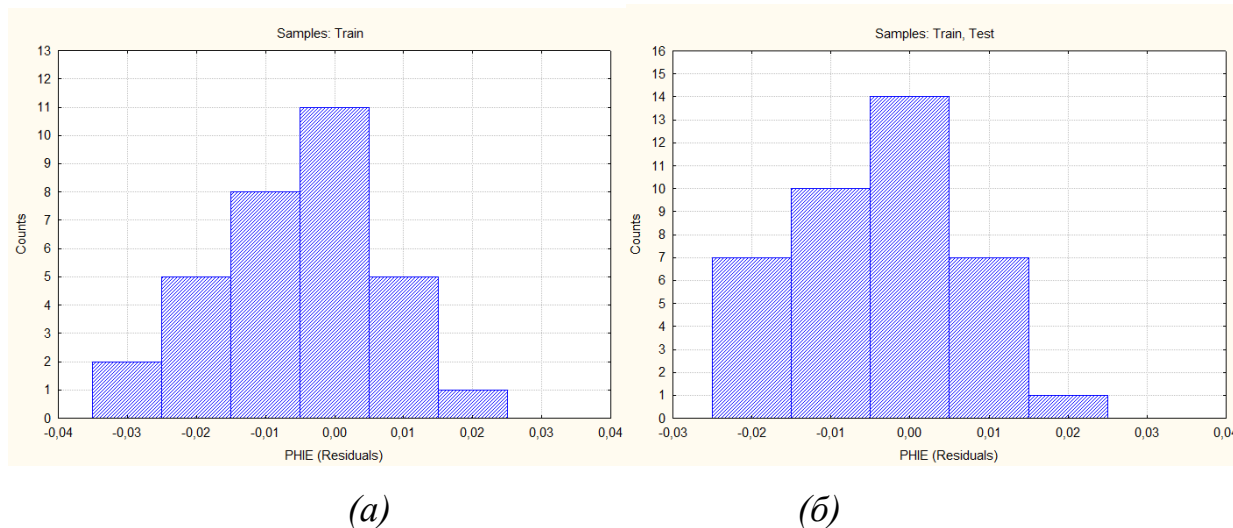


Рисунок 14 – Гистограммы распределения остатков по обучающей выборке (а) и по всей выборке (б) для коэффициента пористости

На Рисунке 15 представлены результаты построения модели нейронных сетей для коэффициента пористости.

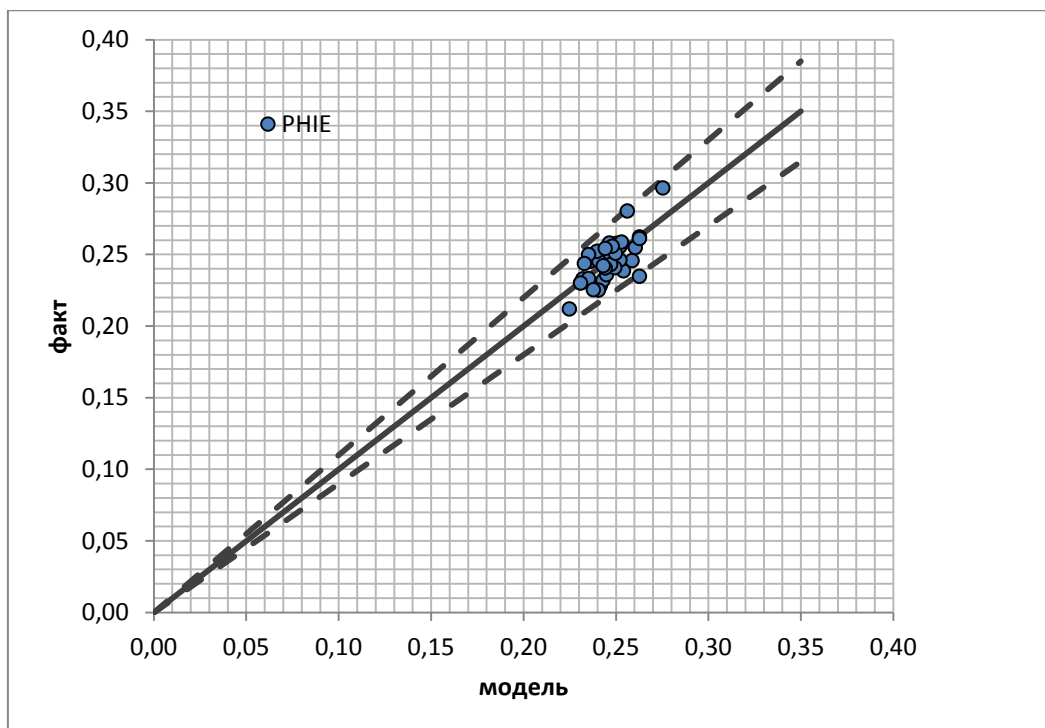


Рисунок 15 – Результат моделирования для коэффициента пористости

2.3.2 Модель для коэффициента песчанистости

Наилучшая модель, полученная для коэффициента песчанистости, имеет следующий вид:

$$NTG = NTG(Phase_{sum\ pos}, Phase_{acr\ length})$$

где NTG – коэффициента песчанистости;

$Phase_{sum\ pos}$ – сумма положительных амплитуд мгновенной фазы;

$Phase_{acr\ length}$ – длина дуги мгновенной фазы.

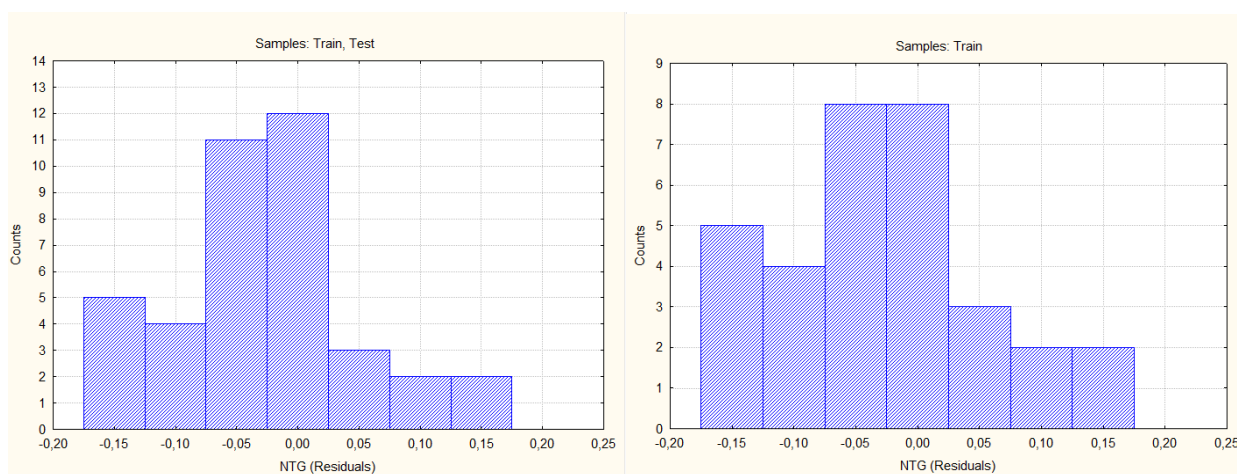
В таблице 8 отражены основные показатели по полученной модели.

Таблица 8 – Показатели нейронной сети для коэффициента песчанности

Количество весовых коэффициентов	13	
Схема нейронной сети	2-3-1, Tanh+ Identity	
Корреляционный коэффициент обучение/тест	0,70	0,78
Межфакторный корреляционный коэффициент	0,28	
Средняя ошибка: д.ед./%	0,04	6

Из данной таблицы можно сделать вывод, что корреляционные коэффициенты обучения и теста, межфакторная корреляция и средняя ошибка модели принимают допустимые значения.

На рисунке 16 представлены гистограммы распределения остатков по обучающей выборке (рис 16(а)) и по всему объему выборки (рис. 16(б)), исходя из которых видно, что остатки не имеют нормальное распределение.



(а)

(б)

Рисунок 16 – Гистограммы распределения остатков по обучающей выборке (а) и по всей выборке (б) для коэффициента песчанности

На Рисунке 17 представлены результаты построения модели нейронных сетей для коэффициента песчанности.

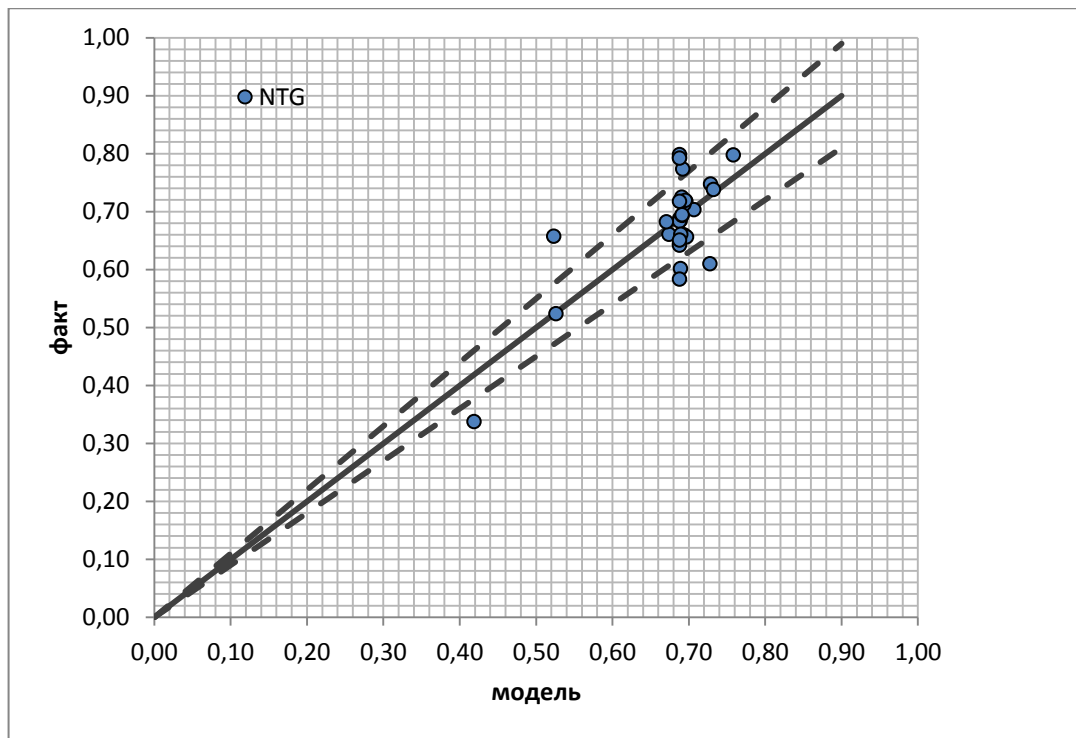


Рисунок 17 – Результат моделирования для коэффициента песчаности

2.3.3 Модель для эффективной толщины пласта

Наилучшая модель, полученная для эффективной толщины, имеет следующий вид:

$$H_{ef} = H_{ef}(Freq_{mean}, EQ_{mean})$$

где H_{ef} – эффективная толщина пласта;

$Freq_{mean}$ - среднее значение мгновенной частоты;

EQ_{mean} – среднее значение эквалайзера.

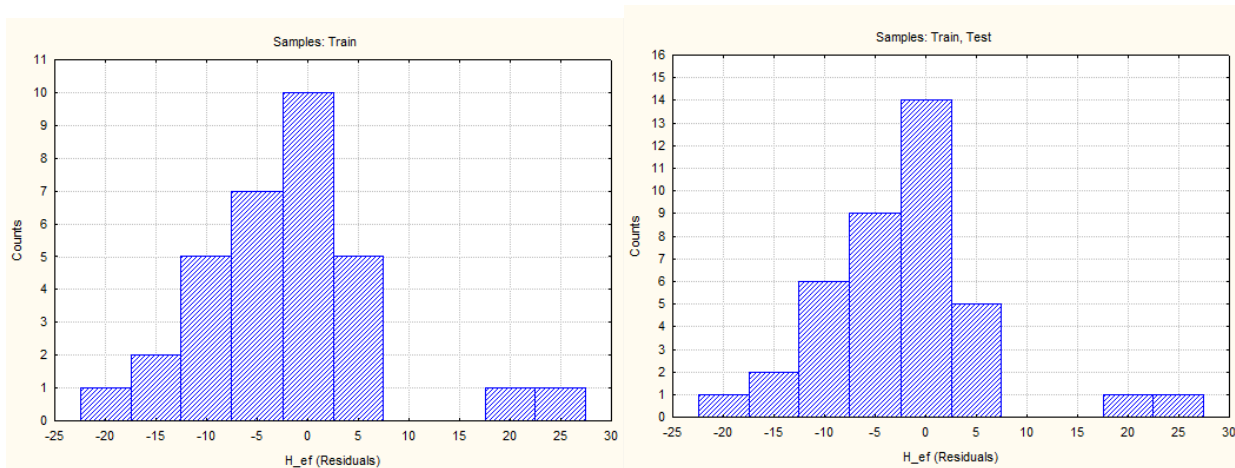
В таблице 9 отражены основные показатели по полученной модели.

Таблица 9 – Показатели нейронной сети для эффективной толщины

Количество весовых коэффициентов	13	
Схема нейронной сети	2-3-1, tanh + Exp	
Корреляционный коэффициент обучение/тест	0,75	0,62
Межфакторный корреляционный коэффициент	0,18	
Средняя ошибка: м/%	6	10

Из данной таблицы можно сделать вывод, что корреляционные коэффициенты обучения, межфакторная корреляция и средняя ошибка модели принимают допустимые значения, но коэффициент корреляции тестовой выборки меньше 0,7, тем самым не удовлетворяет заданным показателям.

На рисунке 18 представлены гистограммы распределения остатков по обучающей выборке (рис 18(а)) и по всему объему выборки (рис. 18(б)), исходя из которых видно, что распределения остатков, как по обучающей выборке, так и по всему объему данных, отличны от нормального.



(a)

(б)

Рисунок 18 – гистограммы распределения остатков по обучающей выборке (а) и по всей выборке (б) для эффективной толщины

На Рисунке 19 представлены результаты построения модели нейронных сетей для эффективной толщины пласта.

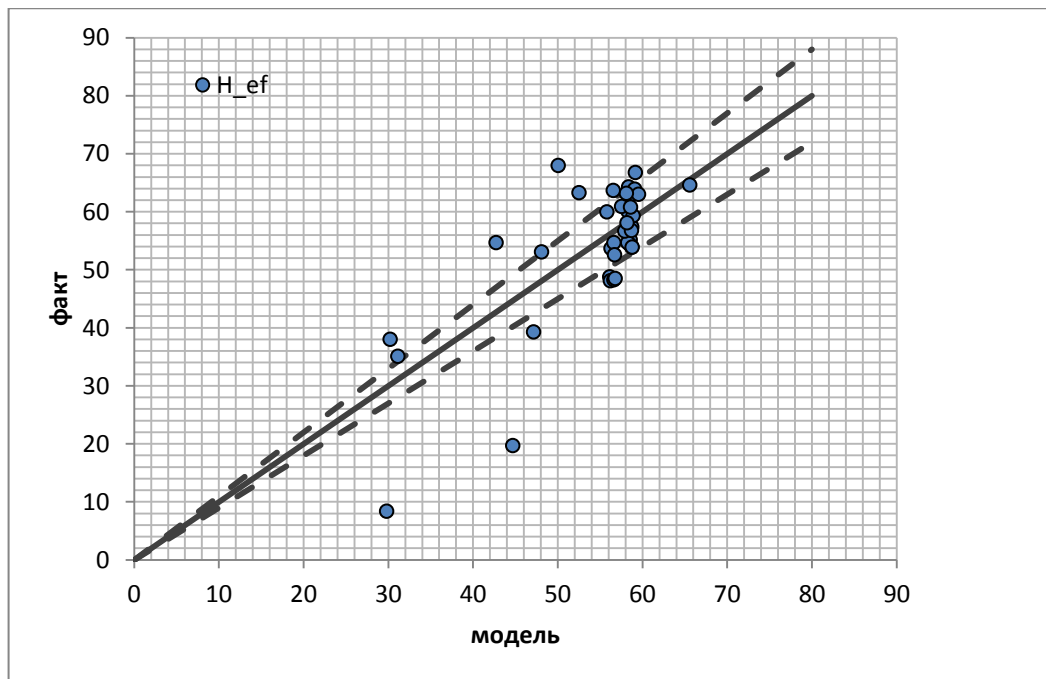


Рисунок 19 – Результат моделирования для эффективной толщины

2.4 Поверхности характеристик пласта

На основе полученных регрессионных моделей можно построить карты распределения исходного параметра пласта по рассчитанным картам значений атрибута. Таким образом, были рассчитаны прогнозные поверхности для всех рассматриваемых характеристик нефтенасыщенного пласта: эффективной толщины, среднего коэффициента пористости по коллектору и коэффициента песчаности пласта.

Для построения модели эффективной толщины использовались такие атрибуты как половина энергии жесткости отражения, половина энергии мгновенной фазы и сумма амплитуд эквалайзера. В результате построения модели была получена поверхность эффективной толщины, представленная на Рисунке 20.

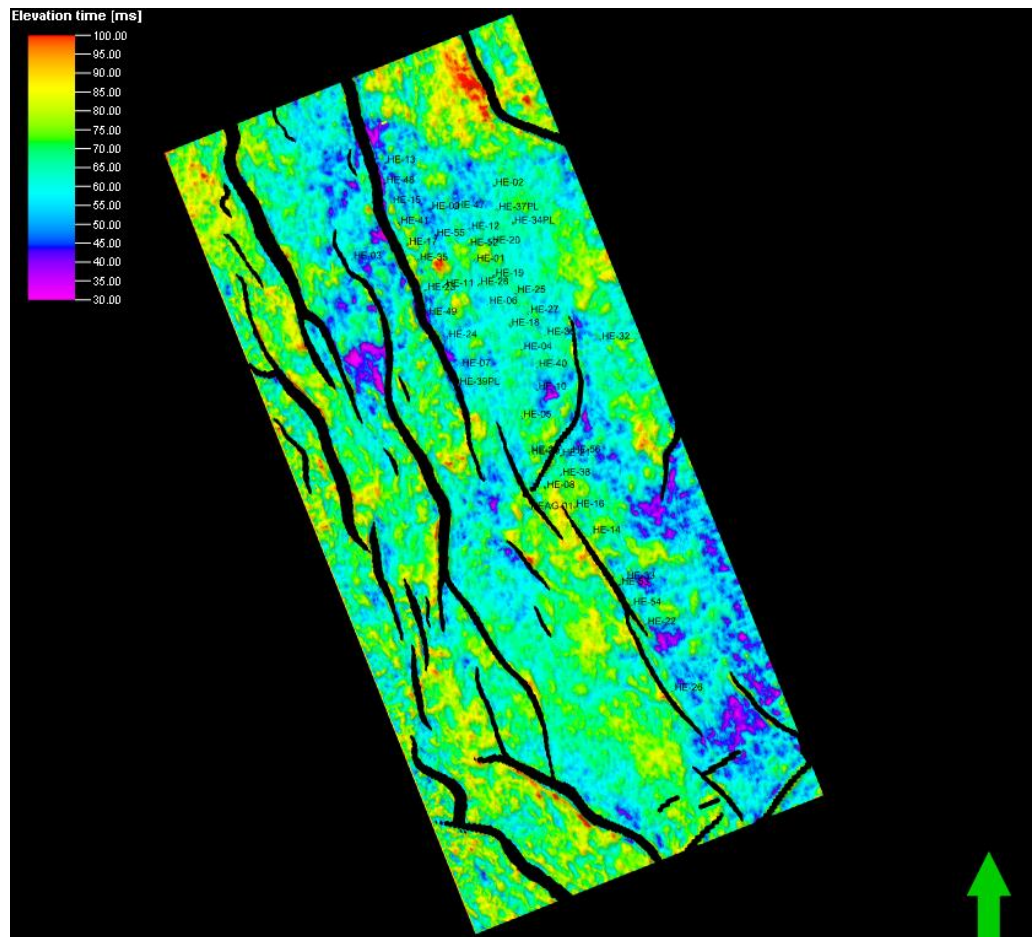


Рисунок 20 – Поверхность эффективной толщины пласта

При построении поверхности было выявлено, что среднее модельное значение эффективной толщины пласта на данном месторождении равно 66,22м. Гистограмма значений модели эффективной толщины представлена на Рисунке 21.

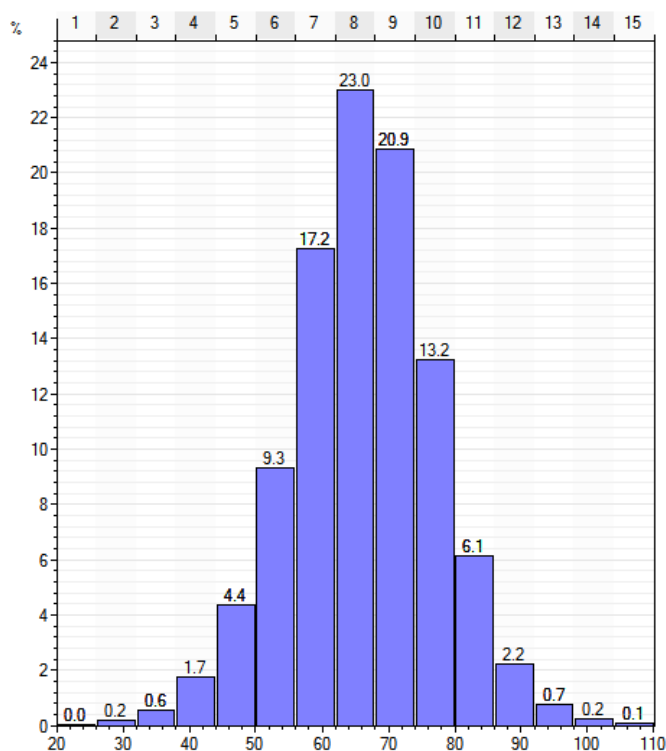


Рисунок 21 – Гистограмма модельных значений эффективной толщины пласта

По результатам статистического исследования для построения модели коэффициента песчанистости пласта используются такие сейсмические атрибуты как среднеквадратичное значение жесткости отражения, длина дуги мгновенной частоты и среднее значений мгновенной фазы. На основе них была построена модельная поверхность, представленная на Рисунке 22.

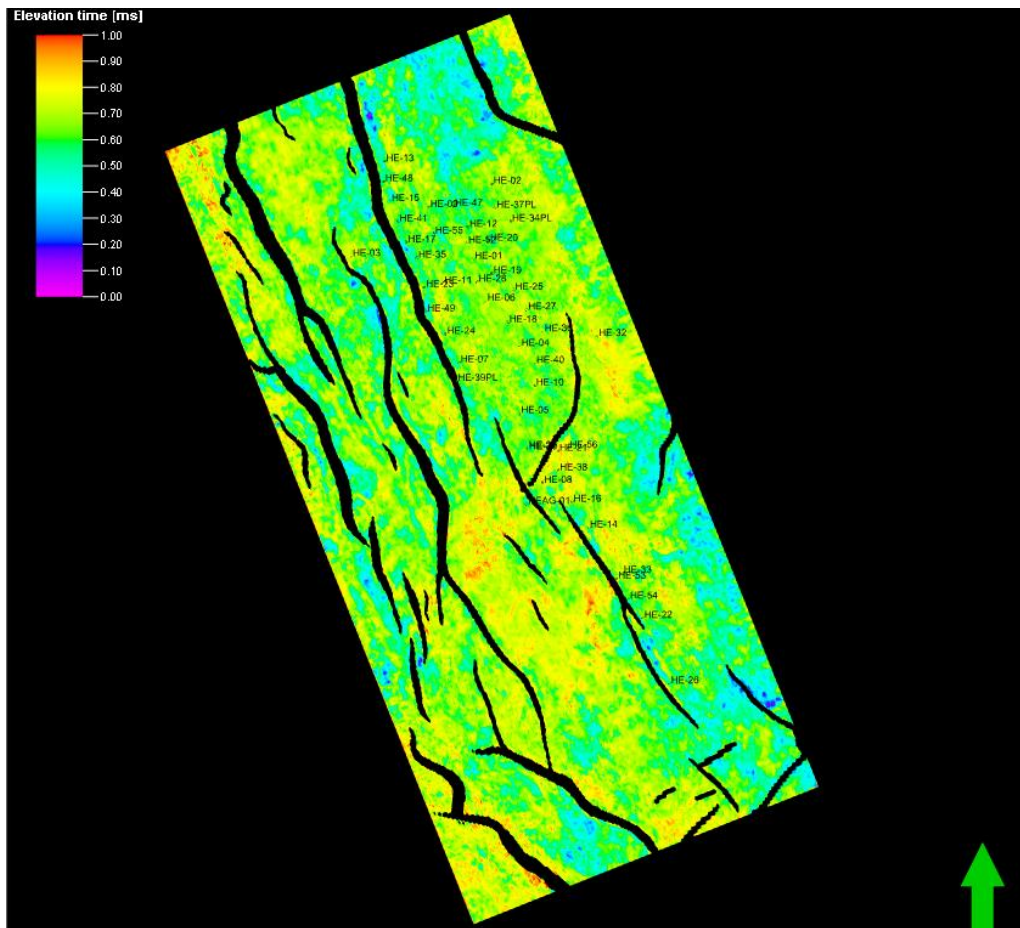


Рисунок 22 – Поверхность коэффициента песчаности пласта

Среднее модельное значение для коэффициента песчаности равно 0,65 д.ед. Гистограмма полученных модельных значений коэффициента песчаности представлена на Рисунке 23.

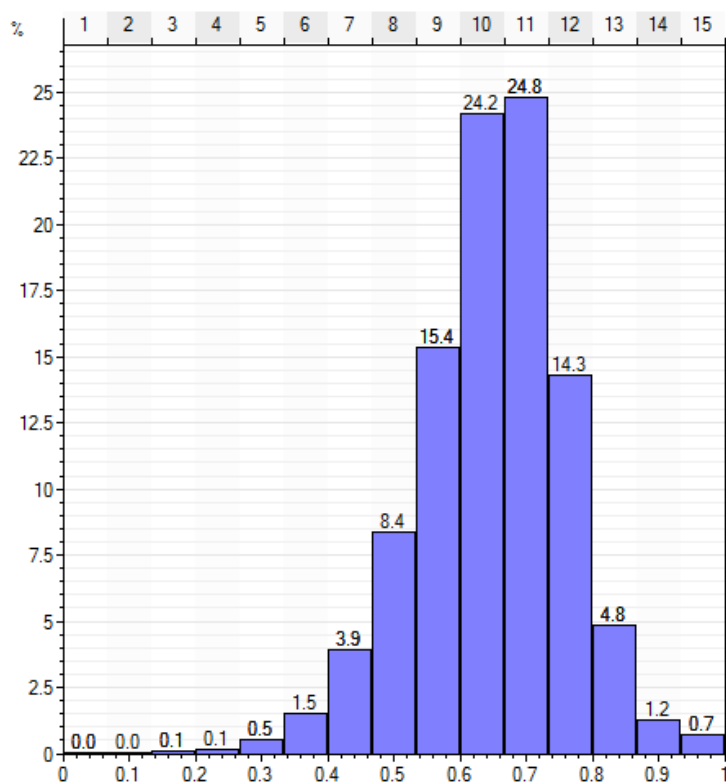


Рисунок 23 – Гистограмма модельных значений коэффициента песчаности пласта

Для построения модели среднего коэффициента пористости по коллектору использовались такие атрибуты, как длина дуги жесткости отражения, среднеквадратичное значение мгновенной фазы и длина дуги мгновенной фазы. В результате расчета модельных значений была получена поверхность коэффициента пористости, представленная на Рисунке 24.

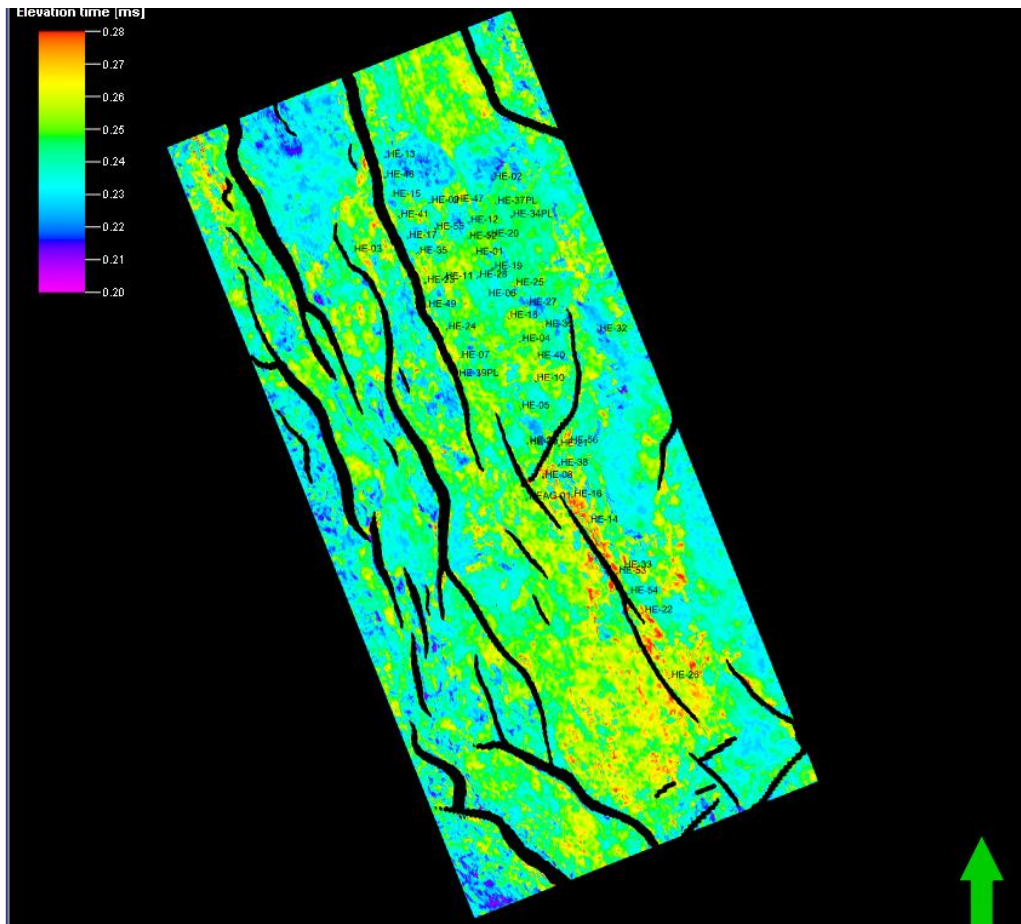


Рисунок 24 – Поверхность среднего коэффициента пористости по коллектору

Из полученной поверхности можно сделать вывод, что среднее значение среднего коэффициента пористости по коллектору равен 0,24д.ед. Гистограмма значений данной модели представлены на Рисунке 25.

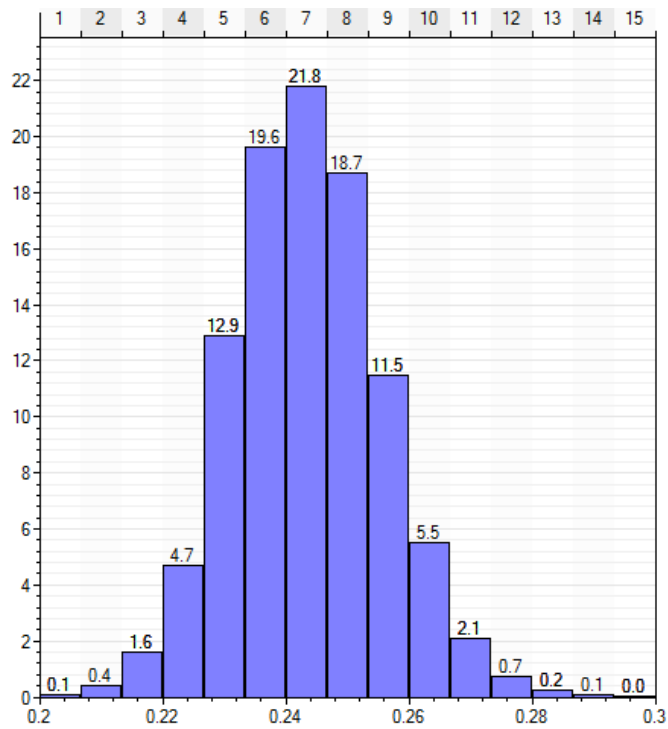


Рисунок 25 – Гистограмма модельных значений среднего коэффициента пористости по коллектору

Представленные характеризуют основное расположение песчаных тел – коллекторов и их петрофизические характеристики и в дальнейшем могут использоваться в качестве тренда для построения геологической модели и расчета нефтяных запасов на данном месторождении.

Заключение

По итогам проведенного комплексного анализа были получены следующие результаты:

- Были прокоррелированы отражающие горизонты Bentiu 1, Bentiu 2 и Bentiu 3 и разломы, рассчитаны сейсмические атрибуты между горизонтами Bentiu 1 и Bentiu 2.

- Было выявлено, что пористость пласта наиболее взаимосвязана с жесткостью отражения и мгновенной фазы; эффективная толщина – с эквалайзером; коэффициент песчаности – с мгновенной фазой.

- Построены регрессионные модели оценки для всех параметров пласта, которые удовлетворяют всем принятым критериям достоверности.

- Построены нейросетевые модели для оценки пористости, эффективной толщины и коэффициента песчаности пласта. Но существенным ограничением данных моделей является малый объем исходной выборки, поэтому они не применимы для прогнозирования.

- На основе полученных регрессионных моделей были построены поверхности рассматриваемых параметров пласта, характеризующие нефтяное месторождение.

Список публикаций студента

1) Галиева Г.Ф. Прогнозирование сбыва воды населению предприятием ООО «Томскводоканал» // Наука и образование в XXI веке / Сборник научных трудов по материалам Международной научно-практической конференции 30 сентября 2013 г.: в 34 частях. Часть 31 :М-во обр. и науки РФ. Тамбов: Изд-во ТРОО «Бизнес-Наука-Общество», 2013. – с. 34-35.

2) Галиева Г.Ф. Пространственно-параметрическая модель рынка жилой недвижимости г.Томска // Перспективы развития фундаментальных наук / Сборник научных трудов XI Международной конференции студентов и молодых учёных. Россия, Томск, 22–25 апреля 2014 г./под ред. Е.А. Вайтулевич. – Национальный Исследовательский Томский политехнический университет, 2014. – с. 558-561.

3) Галиева Г.Ф. Пространственно-параметрическая модель рынка жилой недвижимости г.Томска// Наука и образование в жизни современного общества/сборник научных трудов по материалам Международной научно-практической конференции 30 апреля 2015г.: в 14 томах. Том 11.Тамбов: ООО «Консалтинговая компания Юком», 2015. – с. 21-22.

4) Галиева Г.Ф. Статистическое исследование связи параметров нефтенасыщенных пластов и атрибутов сейсмических// Актуальные вопросы образования и науки /Вестник научных конференций. 2015. №3-3(3). Актуальные вопросы образования и науки: по материалам международной научно- практической конференции 30 ноября 2015 г. Часть3. Тамбов: ООО «Консалтинговая компания Юком», 2015. – с. 47-48.

Список используемой литературы

1. Разин А.В., Меркулов В.П., Чернов С.А. Применение геофизики при изучении месторождений нефти и газа. – Томск: Центр профессиональной переподготовки специалистов нефтегазового дела ТПУ, 2004. – 322 с.
2. Кремер Н.Ш., Пцтко Б.А. Эконометрика: Учебник для вузов/ Под ред. проф. Н.Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2002. 311 с.
3. Нейронные сети. STATISTICANEuralNetworks: Методология и технологии современного анализа данных / Под редакцией В.П. Боровикова. – 2-е изд., перераб. И доп. – М.: Горячая линия – Телеком, 2008. – 392 с.
4. А.А.Халафян STATISTICA 6. Статистический анализ данных. 3-е изд. Учебник – М.: ООО «Бином-Пресс», 2007 г. – 512 с.
5. Кирилов А.С., Закревский К.Е. Практикум по сейсмической интерпретации в RETREL. –М.: Издательство МАИ-ПРИНТ, 2014. – 288 с.
6. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учеб. – 3-е изд., перераб. И доп. М.: Дело, 2000. – 400с.
7. Шерифф Р., Гелдарт Л. Сейсморазведка. Обработка и интерпретация данных / Под редакцией д-ра физ.-мат. Наук, проф. А.В. Калинина. – М.: МИР, 1987. – 400 с.
8. Михалевич И.М., Примина С.П. Применение математических методов при анализе геологической информации (с использованием компьютерных технологий): учебное пособие. – Иркутск: Иркут. гос. ун-т, 2006. – 115 с.
9. Корягин В.В., Сахаров Ю.П. Математическое моделирование в сейсморазведке. М: Наука, 1988. – 160 с.
10. Бондарев В.И., Крылатков С.М. Основы обработки и интерпретации данных сейсморазведки. – Екатеринбург: УГГГА, 2001. – 193 с.

11. Шевченко А.А. Скважинная сейсморазведка. М: РГУ нефти и газа, 2002 – 129 с.
12. Левянт В.Б., Ампилов Ю.П., Глоговский В.М., Колесов В.В., Коростышевский М.Б., Птецов С.Н. Методические рекомендации по использованию данных сейсморазведки (2D, 3D) для подсчета запасов нефти и газа. – Москва, 2006. – 39 с.
13. Ампилов Ю.П. От сейсмической интерпретации к моделированию и оценке месторождений нефти и газа. – М.: СПЕКТР, 2008. – 384 с.
14. Аксенов С.В., Новосельцев В.Б. Организация и использование нейронных сетей (методы и технологии) / Под общ. Ред. В.Б. Новосельцева. – Томск: Изд-во НТЛ, 2006. – 128 с.
15. Жуков Л.А., Решетникова Н.В. Приложения нейронных сетей: Учебное пособие для студентов, учащихся лицей и ЗПШНИ. – Красноярск: ИПЦ КГТУ, 2007. – 154 с.
16. Буреева Н.Н. Многомерный статистический анализ с использованием ППП «STATISTICA»: учебно-методический материал по программе повышения квалификации «Применение программных средств в научных исследованиях и преподавании математики и механики». – Нижний Новгород, 2007. – 112 с.
17. Симчера В.М. Методы многомерного анализа статистических данных: учеб. Пособие. – М.: Финансы и статистика, 2008. – 400 с.
18. Мешкалкина Ю.Л., Самсонова В.П. Математическая статистика в почвоведении: Практикум. – М.: МАКС Пресс, 2008. – 84 с.
19. Куприяненко Н.В., Пономарева О.А., Тихонов Д.В. Статистика. Методы анализа распределений. Выборочное наблюдение. 3-е изд.: учебное пособие. – СПб.: Изд-во Политехн. ун-та. 2009. – 138 с.
20. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение: учебно-методическое пособие / Н.А. Гаврикова, Л.Р. Тухватулина, И.Г. Видяев, Г.Н. Серикова, Н.В. Шаповалова: Томский

политехнический университет. – Томск: Издательство Томского политехнического университета, 2014. - 73 с.

Приложение А

Раздел 2 Результаты проведенного исследования

Студент:

Группа	ФИО	Подпись	Дата
0ВМ41	Галиева Галина Фанисовна		

Консультант кафедры _____ ВММФ :

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Шинкеев М.Л.	к.ф.-м.н.		

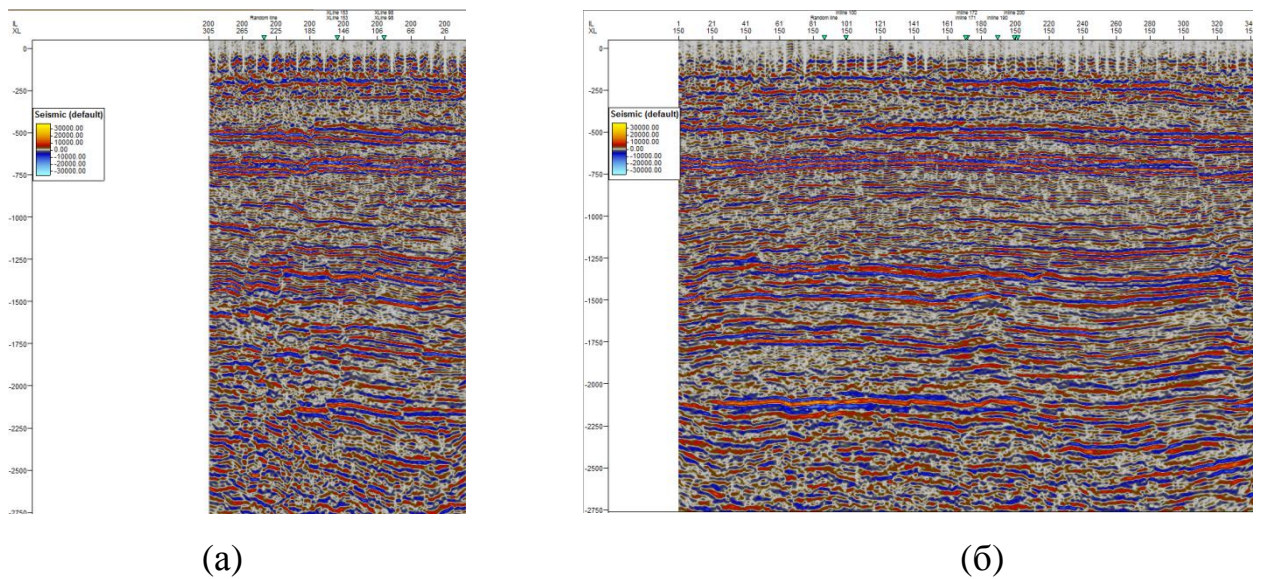
Консультант – лингвист кафедры _____ ВММФ _____ :

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Демьяненко Н.В.			

1 Investigation results

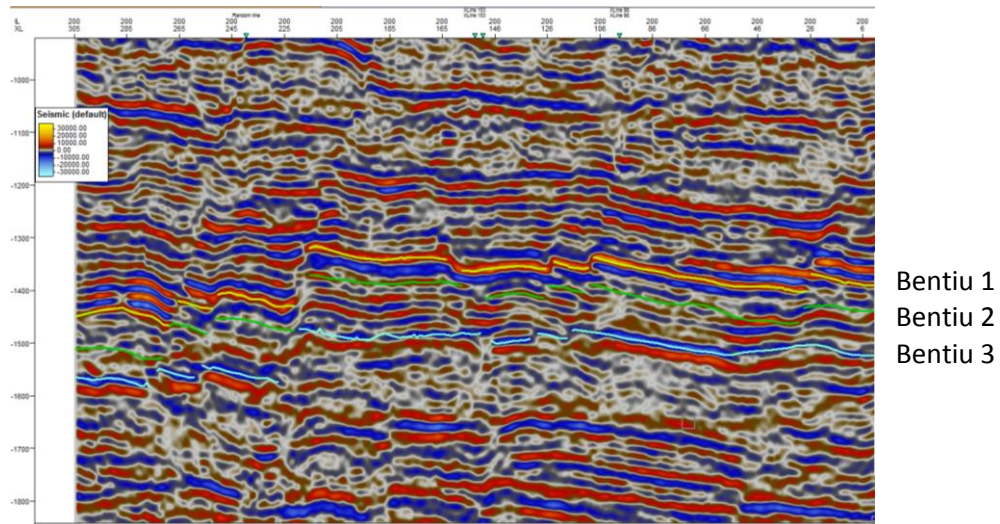
1.1 Correlation of reflection horizons

The first task is to find correlation of reflection horizons Bentiu 1, Bentiu 2 and Bentiu 3 across the seismic survey using the Petrel software. The achieved results are shown in the Picture A.1.



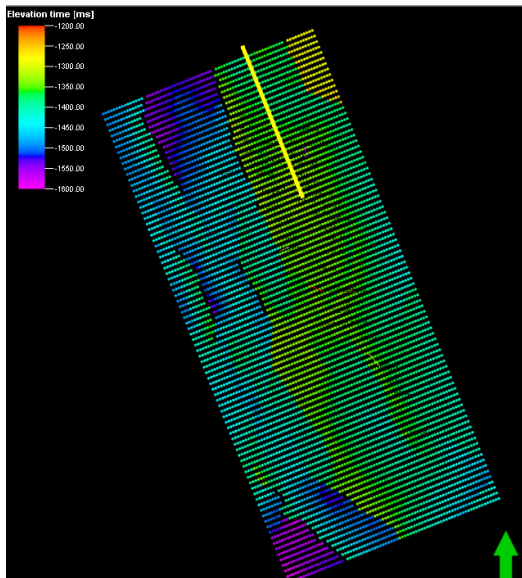
Picture A.1 – Vertical seismic profile: (a) – Inline 200, (b) – Xline 150

Correlation consists of two trace horizons across the whole seismic cube data (both Xline and Inline). Traced horizons for Inline 200 are shown in the Picture A.2.



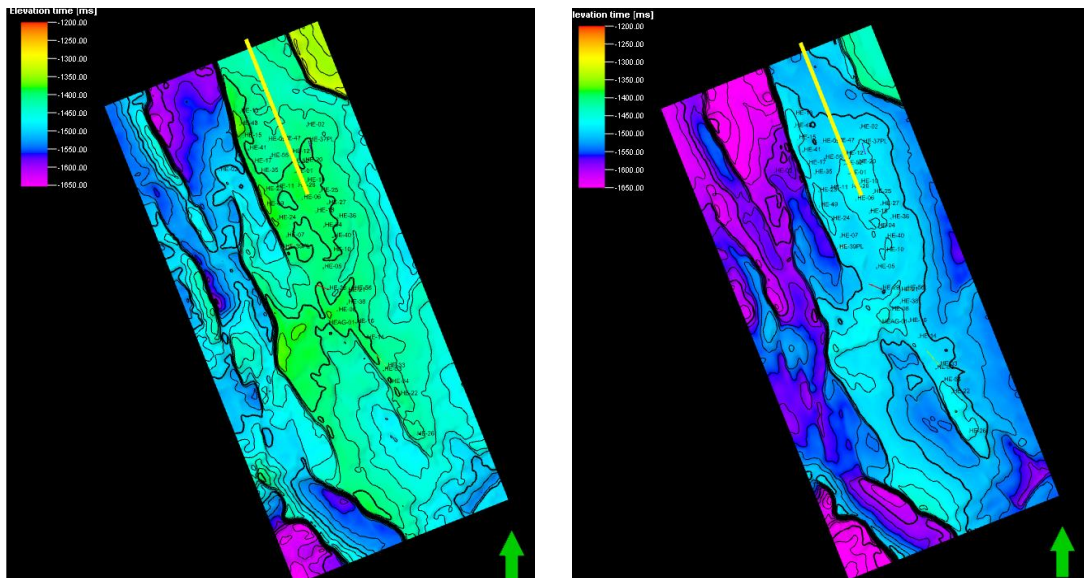
Picture A.2 – The Horizons Bentiu 1, Bentiu 2, Bentiu 3

We used a two-profile step, i.e. horizons were picked every 2 profiles. To illustrate this procedure we constructed traced horizon Bentiu 1 presented in the Picture A.3. The rest of the horizons had similar mesh shape.



Picture A.3 – Trace horizon Bentiu 1

Continuous surface of horizons was constructed by means of the Petrel software. Examples of the achieved results (Bentiu 2 and Bentiu 3) are presented in the Picture A.4.



Picture A.4 – Surface of horizons Bentiu 2 and Bentiu 3

The next stage after finding the correlation was to calculate seismic attributes. We used them for statistical analysis of their dependence with formation characteristics. To achieve this we selected 43 wells with the following volume attributes was selected for the research:

- Graphic Equalizer (EQ) – as a user-defined frequency band-pass filter that can be applied to attenuate or enhance specific frequencies, or interval of frequencies, in order to improve the signal-to-noise ratio and the reflectors continuity. That uses for search dependence with porosity.
- Instantaneous phase (Phase) – describes phase angle in any point along seismic trace. That uses for search of dependences with effective thickness and net-to-gross ratio.
- Instantaneous frequency (frequency) – is calculated from the temporal rate of change of the instantaneous phase.
- Envelope (Env) - the total instantaneous energy of the analytic signal (the complex trace), independent of phase. That uses for search of dependences with effective thickness and net-to-gross ratio.

The investigation object is an oil reservoir with the reflection horizons Bentiu 1 and Bentiu 2 as top and bottom respectively.

For each volume attribute we calculated seven surface values for the chosen oil reservoir:

- Maximum attribute value between top and bottom (Max);
- Minimum attribute value between top and bottom (Min);
- Mean attribute value between top and bottom (Mean);
- RMS Amplitude. RMS Amplitude is the square root of the sum of the squared amplitudes, divided by the number of live samples (RMS)
- Sum of positive amplitudes is the arithmetic mean of the positive amplitudes, multiplied by the number of samples in the window. (Sum_pos)
- The half energy attribute computes the time or depth for energy within a window to reach one-half of the total energy for the attribute of the entire window. (half)
- Arc length measures reflection heterogeneity, and can be used to quantify lateral changes in reflection patterns. (arc length).

WE used the following formation characteristics:

- effective thickness, H_{ef} ;
- porosity factor, PHIE;
- net-to-gross ratio, NTG.

1.2 Regression models

To validate the obtained regression models we used the following indicators:

- coefficient of determination (R-squared) must be greater than 0.5 (in unit fraction);
- residual (the difference between fact and model values) must have distribution similar to the normal one (peaks on the edges of histograms are not allowed, maximum density must be achieved near zero);

- significance level (p-level) which is measure of significance of particular input variable on a result must be smaller 0.05-0.15 (in unit fraction);
- the number of parameters in the model must be less than 3.

1.2.1 Porosity factor model

The most appropriate regression model for porosity factor has the following equation:

$$PHIE = 0,22939 + 0,00007 * Env_{arc\ length} + 0,00040 * Phase_{RMS} - 0,00237 * Phase_{arc\ length}$$

where $PHIE$ – porosity factor;

$Env_{arc\ length}$ – arc length of envelope;

$Phase_{RMS}$ – RMS of instantaneous phase ;

$Phase_{arc\ length}$ – arc length of instantaneous phase.

P-levels of regression coefficients with help of which we can conclude that all of them are significant shown in the Table 1.1.

Table 1.1 – P-levels of regression coefficients

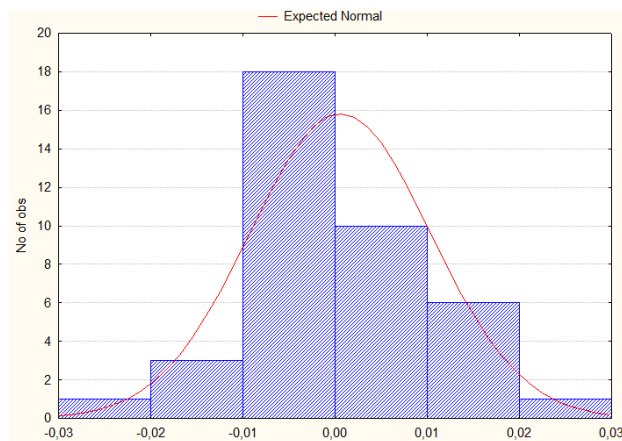
	Coefficient value	p-level
free term	0,22939	0,00
Env_arc length	0,00007	0,00
Phase_RMS	0,00040	0,01
Phase_arc length	-0,00237	0,00

Main parameters are used to validate the model are shown in the Table 1.2. Fisher coefficient determines this model as significant; coefficient of determination, correlation coefficient between factors and average error have an allowable value that determines this as reliable.

Table 1.2– Parameters reliability of model

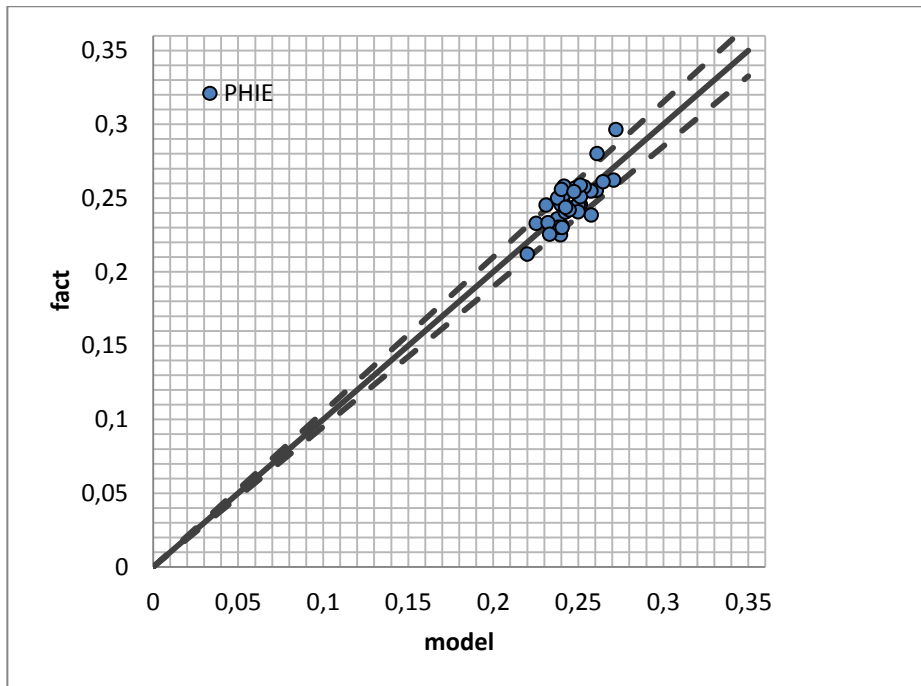
R^2	0,54	
Fisher/p-level	19,91	0,00
Correlation coefficient, absolutely maximum	0,27	
average error: unit fraction /%	0,01	4

The histogram shown in the Picture A.5 is used to assess the distribution of residuals. The residuals are maximum in zero and don't have peaks on the edges. This fact satisfies the model reliability criterion.



Picture A.5 – Histogram of residual distribution of porosity factor

Picture A.6 illustrates the model results for porosity factor.



Picture A.6 – Model results for porosity factor

Thus, according to the parameters set of model validity we can conclude that the model is veridical and can be used as a prediction instrument.

1.2.2 Net-to-gross ratio model

The best regression model for net-to-gross ratio has the following equation:

$$NTG = 1,08249 - 0,00004 * Env_{RMS} - 0,11683 * Freq_{acr\ length} - 0,00203 * Phase_{mean}$$

где NTG – net – to – gross ratio;

Env_{RMS} – RMS of envelope;

$Freq_{acr\ length}$ – arc length of instantaneous frequency;

$Phase_{mean}$ – mean value of instantaneous phase.

P-levels of regression coefficients are shown in Table 1.3. As P-levels for all equation coefficients aren't greater 0.05 we determine this as reliable.

Table 1.3 – P-levels of net-to-gross regression coefficient

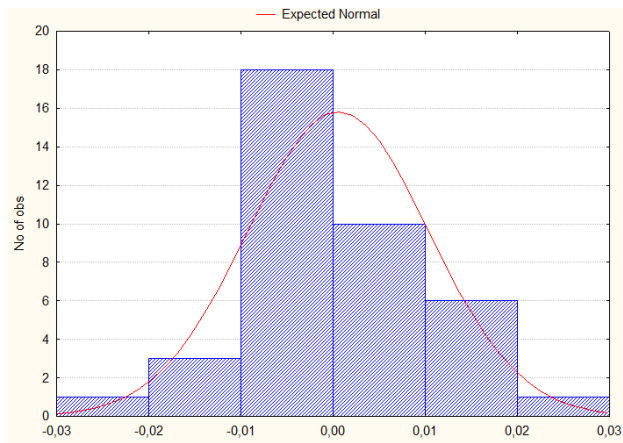
	Coefficient value	P-level
free term	1,08249	0,00
Env_RMS	-0,00004	0,00
Freq_arc length	-0,11683	0,00
Phase_mean	-0,00203	0,00

Main parameters used to validate the model are shown in the Table 1.4. Fisher coefficient determines this model as significant; coefficient of determination, correlation coefficient between factors and average error take an allowable value that determines this as reliable.

Table 1.4 – Parameters reliability of model

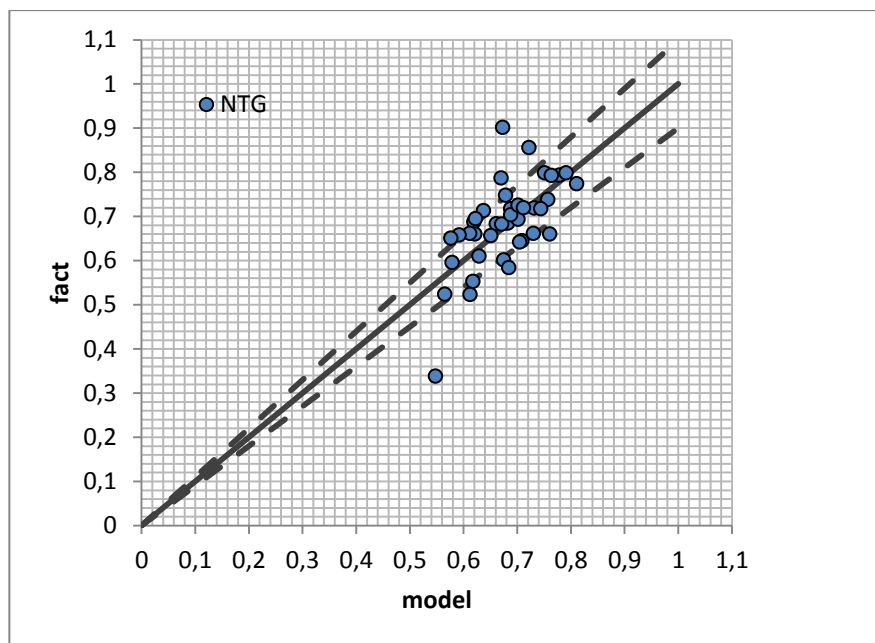
R^2	0,50	
Fisher/p-level	11,40	0,00
Correlation coefficient, absolutely maximum	0,36	
average error: unit fraction /%	0,07	10

A histogram which is shown in the Picture A.7 is used to assess the distribution of residuals. The distribution of residuals is similar to the normal distribution, because it satisfies the indicators of reliability models.



Picture A.7 – Histogram of residual distribution of net-to-gross ratio

Picture A.8 illustrates the model results for net-to-gross ratio.



Picture A.8 - Model results for net-to-gross ratio

Thus, we can conclude from parameters set of model validity that the model is veridical and it can be used to predict.

1.2.3 Effective thickness model

The best regression model for effective thickness has the following equation:

$$H_{ef} = 60,8583 + 0,5176 * Env_{half} - 0,5000 * Phase_{half} - 0,0004 * EQ_{sum}$$

where H_{ef} – effective thickness;

Env_{half} – half energy of envelope;

$Phase_{half}$ – half energy of instantaneous phase;

EQ_{sum} – sum of amplitude of graphic equalizer .

P-levels of regression coefficients with help of which we can conclude that all of them are significant shown in Table 1.5.

Table 1.5 – P-levels of regression coefficients for effective thickness

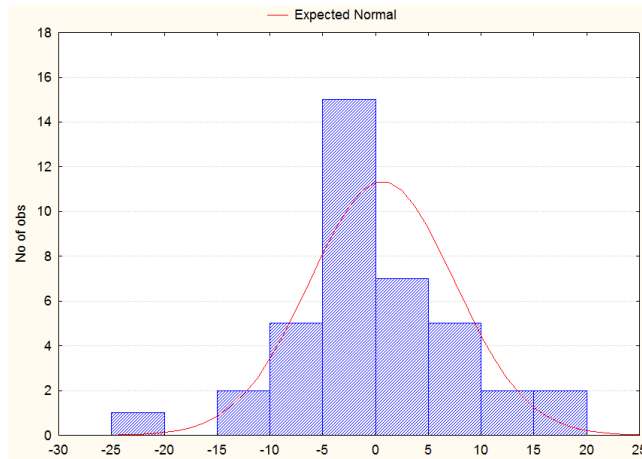
	Coefficient value	P-level
free term	60,8583	0,00
Env_half	0,5176	0,00
Phase_half	-0,5000	0,00
Phase_sum	-0,0004	0,00

According to the Table 1.6 we can conclude that the equation satisfies all indicators mentioned above: it's significant, coefficient of determination is greater than 0.5, average error and correlation coefficient between factors does not greater than specified indicators.

Table 1.6 – Parameters reliability of model

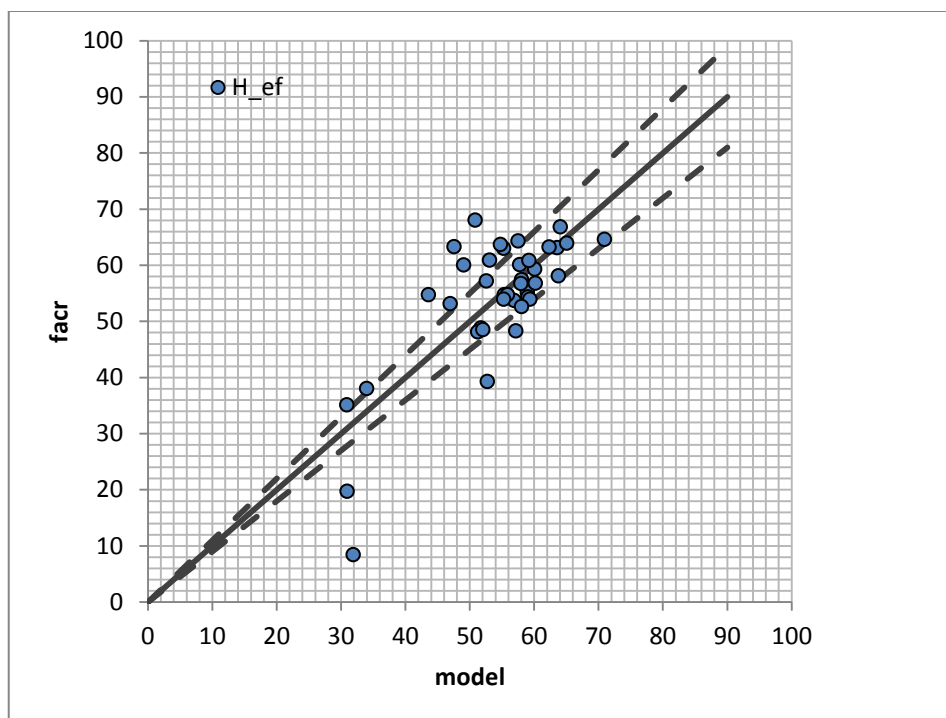
R^2	0,59	
Fisher/p-level	16,50	0,00
Correlation coefficient, absolutely maximum	0,12	
Average error: m /%	8,12	15

The histogram which describes residual distribution is shown in the Picture A.9. We can conclude that the distribution of residuals is similar to the normal distribution.



Picture A.9 – Histogram of residual distribution of effective thickness

Picture A.10 shows the model results for effective thickness.



Picture A.10 - Model results for effective thickness

Thus, according to the parameters set of model validity, we can conclude that the model is veridical and it can be used to predict.

1.3 Neuron net model

The neuron net model was built for derived data. But because of a short data set, the final neuron net model is not applicable for prediction.

We used the following criteria to choose the most appropriate model:

- Correlation coefficient of study and test must be greater than 0.7.
- Coefficient correlation between factors must be less than 0.5.
- Average error must be less than 0.15.
- Residual must have distribution similar to normal.

1.3.1 Porosity factor model

The best regression model for porosity factor has the following equation:

$$PHIE = PHIE(Env_{arc\ length}, Freq_{mean})$$

where $PHIE$ – porosity factor;

$Env_{arc\ length}$ – arc length of envelope;

$Freq_{mean}$ – mean value of instantaneous frequency.

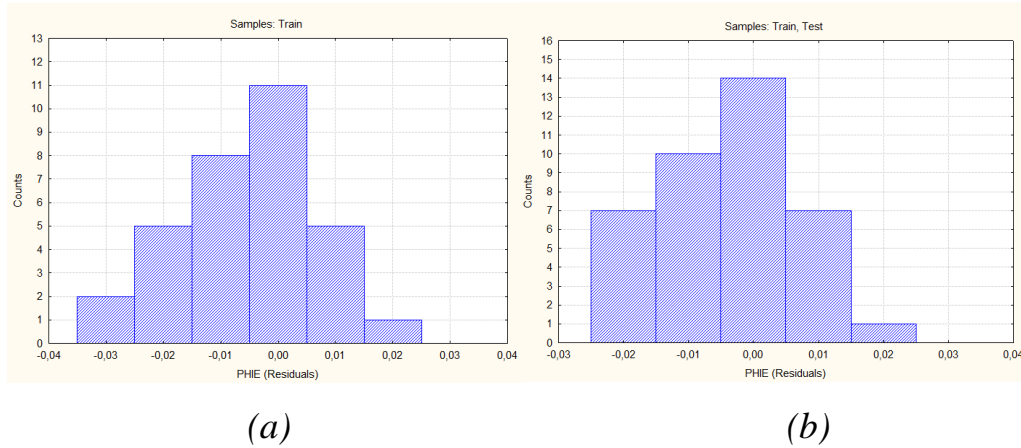
Main indicators of the result model are shown in the Table 1.7.

Table 1.7 – Neuron net indicators for porosity factor

Number of weighting coefficient	13	
Neuron net structure	2-3-1,Exp+ Logistic	
Correlation coefficient of study/test	0,70	0,93
Correlation coefficient between factors	0,03	
Average error : unit fraction/%	0,008	4

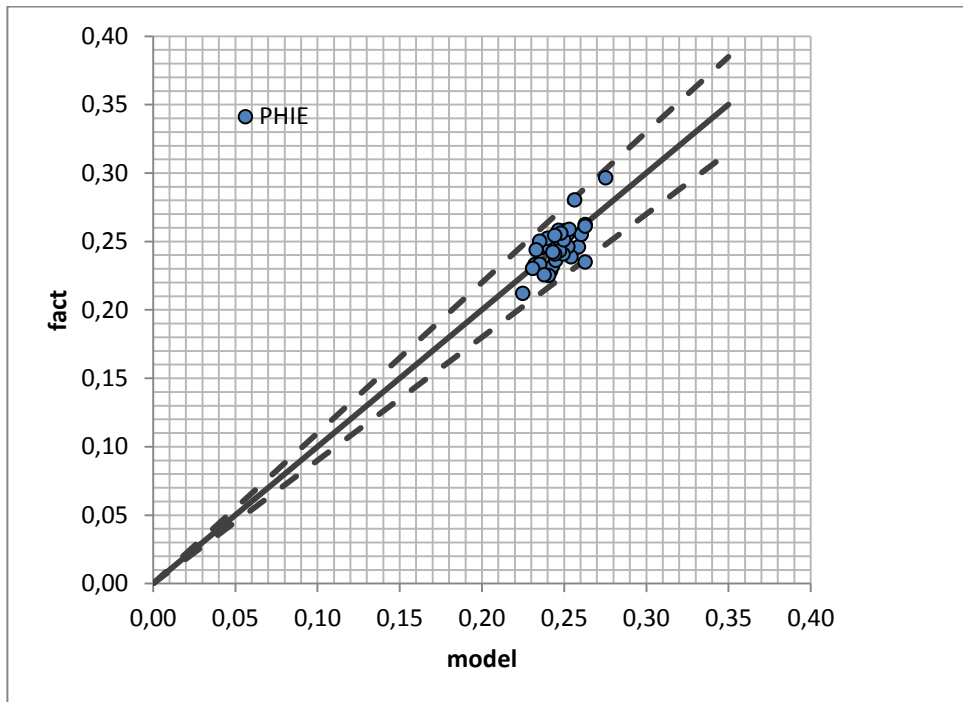
From this table, we can conclude that correlation coefficients of study and test, correlation between factors and average error of model possess the allowed value.

The residual distribution histograms for study data (picture A.11(a)) and for all data (picture A.11(b)) are shown in Picture A.11, on the basis of which it is clear that residuals of both study data and all data have maximum in zero and don't have peak on the edge but it isn't symmetrical.



Picture A.11 – Residual distribution histogram of study data (a) and all data (b) for porosity factor

Picture A.12 illustrates the results of neuron net model building for porosity factor.



Picture A.12 – Model result for porosity factor

1.3.2 Net-to-gross ratio model

The best regression model for net-to-gross ratio has following equation:

$$NTG = NTG(Phase_{sum\ pos}, Phase_{acr\ length})$$

where NTG – net-to-gross ratio;

$Phase_{sum\ pos}$ – sum of positive amplitude of instantaneous phase;

$Phase_{acr\ length}$ – arc length of instantaneous phase.

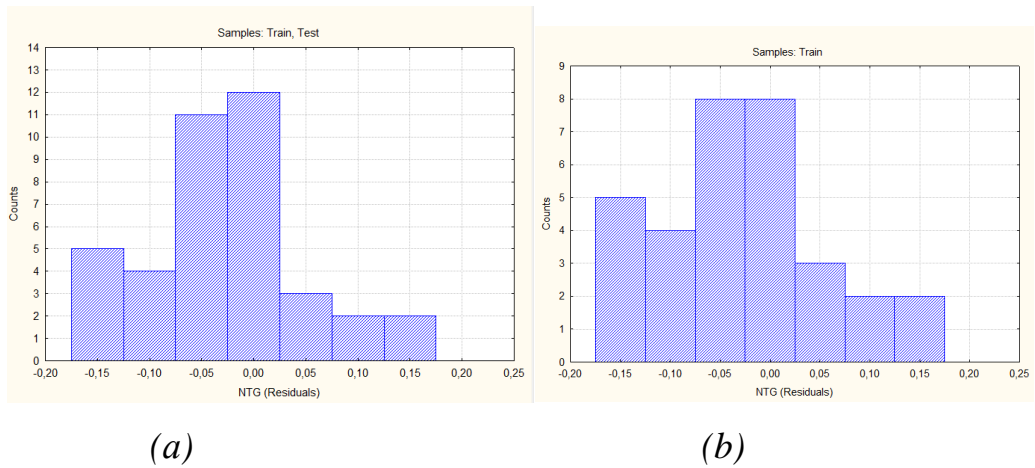
Main indicators of result model are shown in Table 1.8.

Table 1.8 – Neuron net indicators for net-to-gross ratio

Number of weighting coefficient	13	
Neuron net structure	2-3-1, Tanh+ Identity	
Correlation coefficient of study/test	0,70	0,78
Correlation coefficient between factors	0,28	
Average error : unit fraction/%	0,04	6

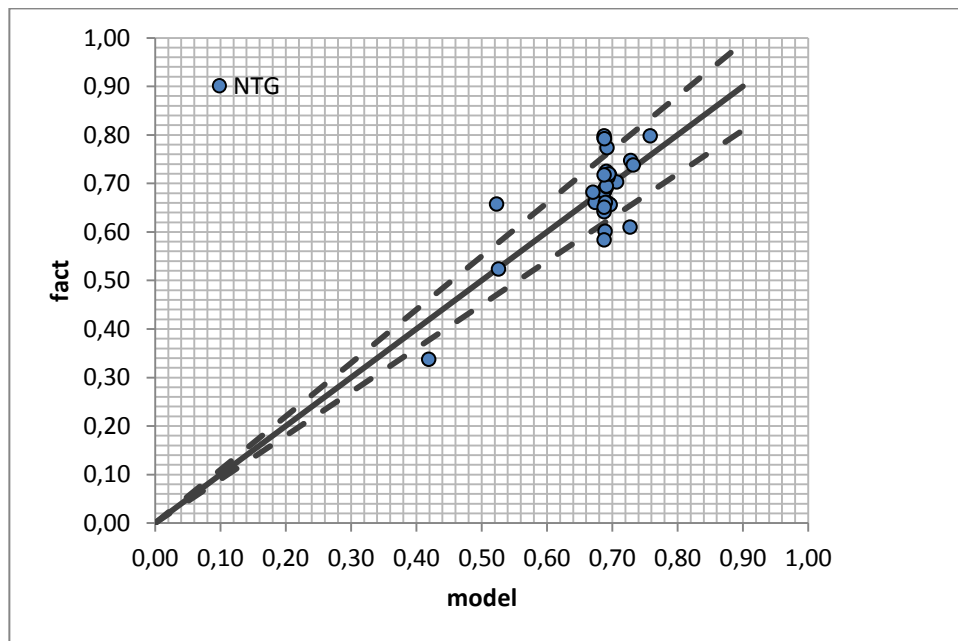
From this table, we can conclude that correlation coefficients of study and test, correlation between factors and average error of model possess the allowed value.

The residual distribution histograms for study data (picture A.13(a)) and for all data (picture A.13(b)) are shown in Picture A.13, on the basis of which it is clear that residuals don't have normal distribution.



Picture A.13 – Residual distribution histogram of study data (a) and all data (b) for net-to-gross ratio

In the Picture A.14 we show the results of the neutron net model built for net-to-gross ratio.



Picture A.14 – Model result for net-to-gross ratio

1.3.3 Effective thickness model

The best regression model for effective thickness has the following equation:

$$H_{ef} = H_{ef}(Freq_{mean}, EQ_{mean})$$

where H_{ef} – effective thickness;

$Freq_{mean}$ – mean value of instantaneous frequency;

EQ_{mean} – mean value of graphic equalizer.

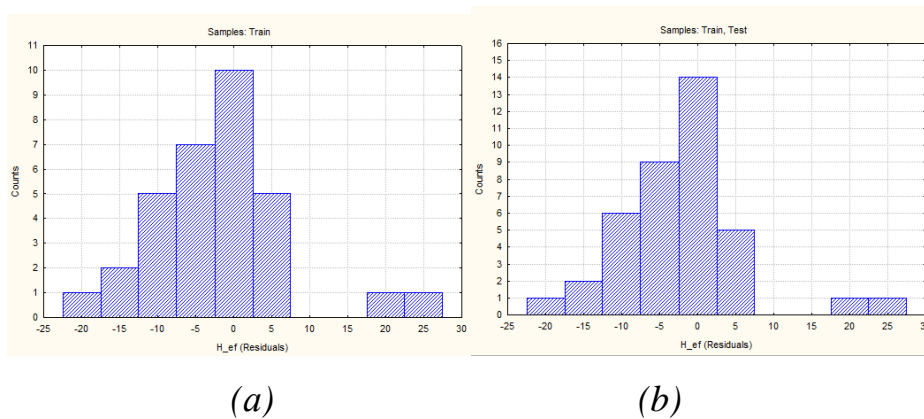
Main indicators of result model are shown in Table 1.9.

Table 1.9 – Neuron net indicators for thickness model

Number of weighting coefficient	13	
Neutron net structure	2-3-1, tanh + Exp	
Correlation coefficient of study/test	0,75	0,62
Correlation coefficient between factors	0,18	
Average error : m/%	6	10

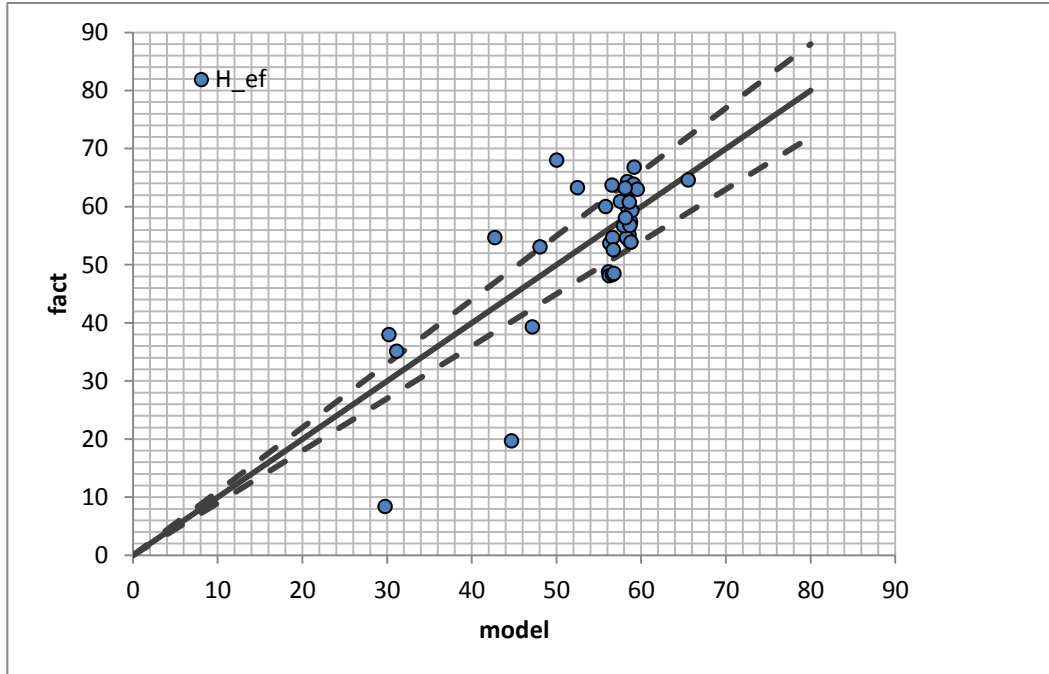
According to this table, we can conclude that correlation coefficients of the studied data, correlation between factors and average error of model possess the allowed value but correlation coefficient of the tested data is less than 0.7, so it doesn't satisfy the indicators.

The residual distribution histograms for the studied data (picture A.15(a)) and for all data (picture A.15(b)) are shown in the Picture A.15. On the basis of these histograms it is clear that residuals of both study data and tested data don't have the normal distribution.



Picture A.15 - Residual distribution histogram of study data (a) and all data (b) for effective thickness

Picture A.16 illustrates the results of the neutron net model built for effective thickness.



Picture A.16 – Model result for effective thickness