

Министерство науки и высшего образования Российской Федерации  
Санкт-Петербургский политехнический университет Петра Великого

Институт прикладной математики и механики

Высшая школа теоретической механики

Работа допущена к защите

Директор высшей школы

\_\_\_\_\_ А.М. Кривцов

«\_\_»\_\_\_\_\_ 2021 г.

## **ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

Магистерская диссертация

### **«РАЗРАБОТКА МЕТОДИКИ ОЦЕНКИ ЦЕННОСТИ ИНФОРМАЦИИ ДЛЯ ПРОВЕДЕНИЯ ПОЛЕВЫХ ГЕОФИЗИЧЕСКИХ ИССЛЕДОВАНИЙ»**

по направлению подготовки (специальности)  
01.04.03 «Механика и математическое моделирование»

Направленность (профиль)  
01.04.03\_02 «Механика и математическое моделирование  
(международная образовательная программа)»

Выполнил  
студент гр.  
3640103/90201

В.Р. Бадрдинов

Руководитель  
доцент, к.ф.-м.н.

Е.А. Подольская

**САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ  
ПЕТРА ВЕЛИКОГО**

**Институт прикладной математики и механики  
Высшая школа теоретической механики**

УТВЕРЖДАЮ

Директор высшей школы

А.М. Кривцов

«\_\_» \_\_\_\_\_ 2021 г.

**ЗАДАНИЕ**

**на выполнение выпускной квалификационной работы**

студенту Бадртдинову Владиславу Робертовичу, 3640103/90201

1. Тема работы: Разработка методики оценки ценности информации для проведения полевых геофизических исследований
2. Срок сдачи студентом законченной работы: 07.06.2021
3. Исходные данные по работе: Геолого-промысловые данные об условиях и показателях разработки месторождения, а также эксплуатации скважин, информация из технической литературы (включая периодические издания и патенты), месячный эксплуатационный рапорт по месторождению, рапорты геофизических исследований.
4. Содержание работы (перечень подлежащих разработке вопросов): Статистический анализ показателей эксплуатации скважин и геофизических исследований, создание схемы расчёта ценности информации для промыслово-геофизических исследований с целью повышения эффективности ремонтно-изоляционных работ, построение алгоритмов, реализованных в виде программного обеспечения на языке программирования Python.
5. Перечень графического материала (с указанием обязательных чертежей): Заключение геофизических исследований, схемы сценариев расчета, графики статистического анализа, график сравнения методов интерполяции, результаты сравнения моделей.
6. Консультанты по работе: П.М. Кикин, специалист по машинному обучению, ООО «Шторм-технологии»; В.С. Дряхлов, главный специалист, ООО «Морнефтегазпроект».
7. Дата выдачи задания 25.01.21

Руководитель ВКР \_\_\_\_\_ Е.А. Подольская  
(подпись)

Задание принял к исполнению 25.01.21

Студент \_\_\_\_\_ В.Р. Бадртдинов  
(подпись)

## РЕФЕРАТ

На 59 с., 32 рисунка, 5 таблиц, 0 приложений

**КЛЮЧЕВЫЕ СЛОВА:** ЦЕННОСТЬ ИНФОРМАЦИИ, ПРОМЫСЛОВО-ГЕОФИЗИЧЕСКИЕ ИССЛЕДОВАНИЯ, РЕМОНТНО-ИЗОЛЯЦИОННЫЕ РАБОТЫ, ОПТИМАЛЬНОЕ ПЛАНИРОВАНИЕ, ПРОГНОЗИРОВАНИЕ ДЕБИТА НЕФТИ.

Тема выпускной квалификационной работы: «Разработка методики оценки ценности информации для проведения полевых геофизических исследований».

Целью работы является разработка подхода и алгоритмов для оценки экономической целесообразности проведения промышленного геофизических исследований на нефтяных скважинах с предполагаемым наличием негерметичности эксплуатационной колонны и/или заколонной циркуляции. Задачи, которые решались в ходе исследования:

1. Создание схемы расчёта ценности информации для промышленно-геофизических исследований с целью повышения эффективности ремонтно-изоляционных работ;
2. Построение алгоритмов, реализованных в виде программного обеспечения на языке программирования Python.

Реализованный подход призван обеспечить повышение экономической эффективности проводимых исследований, снижение материальных и трудовых затрат на их планирование, автоматизацию оценки экономического эффекта от проведения ремонтно-изоляционных работ, что позволит избежать затрат на ремонт заведомо убыточных скважин.

Считывание входных данных реализовано для файлов формата MS-Excel имеющих типовую структуру представления исторических данных дебита скважин.

Проведено тестирование алгоритма на предоставленных для анализа данных по скважинам нефтегазоконденсатного месторождения N.

## THE ABSTRACT

59 pages, 32 figures, 5 tables, 0 appendices

KEY WORDS: VALUE OF INFORMATION, FIELD AND GEOPHYSICAL RESEARCH, REPAIR AND INSULATION WORKS, OPTIMAL PLANNING, OIL FLOW RATE FORECASTING.

The subject of the graduate qualification work is «Development of a methodology for assessing the value of information for field geophysical surveys».

The aim of the project is to develop an approach and algorithms for assessing the economic viability of field geophysical surveys on oil wells with suspected production casing leakage and/or behind-the-casing circulation. The objectives of the work were:

1. Creation of a scheme for calculating the value of information for field geophysical surveys in order to improve the efficiency of repair and insulation works;
2. Construction of algorithms implemented as software in Python programming language.

The implemented approach is designed to improve the economic efficiency of surveys, reduce material and labour costs for their planning, and automate the assessment of the economic effect of repair and insulation work, which will avoid the cost of repairing knowingly unprofitable wells.

The input data reading is implemented for MS-Excel format files with typical structure of historical well flow rate data representation.

The algorithm was tested on the data on the wells of oil-gas-condensate field N provided for the analysis.

## СОДЕРЖАНИЕ

Введение .....	6
Глава 1. Ценность информации для ПГИ .....	8
1.1. Описание исходных данных .....	8
1.2. Общее описание применения методики VOI для ПГИ .....	9
1.3. Схема расчета VOI для ПГИ .....	11
Глава 2. Статистический анализ исходных данных .....	17
2.1. Анализ данных ПГИ .....	17
2.2. Анализ данных РИР .....	20
2.3. Алгоритм фильтрации выбросов .....	26
2.4. Алгоритм формирования синтетического датасета .....	30
Глава 3. Предсказательные модели .....	32
3.1. Модель идентификации скважин с дефектами .....	32
3.1.1. Логистическая регрессия .....	36
3.2. Модель прогнозирования дебита нефти и уровня обводненности без ПГИ/РИР .....	37
3.3. Модель прогнозирования уровня обводненности и дебита нефти после ПГИ/РИР .....	44
Глава 4. Алгоритм оценки ценности информации .....	47
4.1. Алгоритм оценки $EMV_{woi}$ или $EMV_{wi}$ .....	47
4.2. Ранжирование скважин на основе прогнозируемого эффекта от РИР (VOI) .....	49
Глава 5. Результаты тестирования алгоритмов расчета VOI на реальных данных .....	52
Заключение .....	56
Список использованных источников .....	58

## Введение

Одним из основных элементов контроля разработки нефтегазовых залежей являются промыслово-геофизические исследования (ПГИ) скважин. Проведение промыслово-геофизических исследований скважин обеспечивает оценку технического состояния скважин, герметичности колонн, глубины текущего искусственного забоя, контроль качества и целостности цементного камня и др. Результаты интерпретации ПГИ позволяют обосновывать ремонтно-изоляционные работы (РИР), которые проводят с целью изоляции пластовых вод, поступающих через цементный стакан и по заколонному пространству, подошвенных и контурных вод, поступающих по наиболее проницаемым интервалам и трещинам пласта, т.е. обеспечивают оптимальные условия работы продуктивного пласта, для достижения запланированной (максимальной) выборки запасов нефти.

Выбор и реализация РИР должны основываться на данных ПГИ. Тем не менее, полный охват действующего фонда скважин промыслово-геофизическими исследованиями в большинстве случаев затруднён из-за технологических ограничений и затрат на проведение исследований, в том числе за счёт потерь добычи нефти при остановке скважин на время проведения ПГИ.

Существует обширный список существующих технологий ПГИ, выбор которых зависит от различных факторов: взаиморасположение насосно-компрессорных труб (НКТ) и интервалов перфорации, число совместно эксплуатируемых пластов, дебит / расход закачиваемой жидкости, газосодержание продукции и др. В зависимости от технологий и методов проведения ПГИ будут зависеть затраты на их проведение, в том числе за счёт потерь добычи нефти при остановке скважин на время проведения ПГИ.

Принятие решения о выборе скважины для проведения ПГИ, предполагает экономическое обоснование, что является нетривиальной задачей в условиях неопределенности. При планировании ПГИ необходимо определять баланс

между затратами на проведение исследования и ценностью информации, которую оно даёт.

Для решения этой задачи в рамках данной выпускной квалификационной работы (ВКР) были проведены исследования с целью изучения возможностей применения подхода к оценке ценности информации (VOI) в отношении ПГИ. В рамках ВКР был разработан подход к оптимальному планированию проведения ПГИ на основе математических методов машинного обучения. Реализованный подход призван обеспечить повышение эффективности исследований, снижение трудозатрат на их планирование за счет автоматизации оценки достижимого экономического эффекта от проведения ПГИ/РИР и повышение прозрачности самих процессов планирования и обоснованности выбранных для исследования скважин и методов исследования.

В ходе выполнения выпускной квалификационной работы предполагается решить следующие задачи:

1. Статистический анализ исходных данных;
2. Создание предсказательных моделей;
3. Расчет ценности информации;
4. Анализ полученных результатов.

Исполнитель ПГИ	ТЦД	Заказчик ПГИ		
<b>Заключение ПГИ</b>				
Месторождение	Скважина	Буст		
		Дата исследования		
		20/03/2017		
Определение технического состояния скважины				
Способ исследования при исследовании - АСОПРОБАНЬ				
<b>Приборы</b>				
Тип прибора	Номер	Пороговое значение		
КСАТ-6	190			
<b>Конструкция скважины</b>				
Диаметр эксплуатационной колонны:	178 мм			
Диаметр НКТ:	73 мм			
Диаметр хвостовика:	114 мм			
Искусственный лавой:	3334 м			
Глубина лавой:	2441 м			
Интервал мультислойного павара:	2314.4-2316.6м 2389.3-2390.6 м			
Глубина спуска НКТ по данным промысла:	1700 м			
<b>Результаты измерения давления в трубах</b>				
Дата измерения	Время	Давление, атм	Г.Д.Р., м	В.Н.Р., м
20/03/2017	16:06	917	917	966
	16:17	900	900	960
	16:58	872	872	925
	17:02	851	851	901
	17:27	827	827	880
	20:47	778	778	830
<b>Данные по перфорации</b>				
Пласт	Интервал промывки, м	Интервал по ПГИ, м	Тип перфорации	Дата перфорации
ВВВ-1	3002.38 - 3103.08			
ВВВ-1	3228.11 - 3228.81			
ВВВ-1	3323.5 - 3333.68			
<b>Техническое состояние скважины</b>				
Кривая, м	Подопла, м	Состояние скважины		
2314.4	2316.4	Павар мультислойной герметизации		
2389.2	2390.6	Павар мультислойной герметизации		
<b>Барометрия</b>				
Риски	Дата	Время	Давление, атм	Глубина, м
Фоновый замер	20/03/2017	02:00	160.1	2300
Фоновый замер	20/03/2017	06:00	157.35	2300
При компрессировании	20/03/2017	07:00	161.68	2300
При компрессировании	20/03/2017	08:40	167.35	2300
При компрессировании	20/03/2017	09:10	165.48	2300
После промывки пусковой муфты	20/03/2017	09:30	176.1	2300
После промывки пусковой муфты	20/03/2017	09:50	172.22	2300
После промывки пусковой муфты	20/03/2017	10:20	157.64	2300
После промывки пусковой муфты	20/03/2017	11:00	145.7	2300
После промывки пусковой муфты	20/03/2017	12:00	129.67	2300
После промывки пусковой муфты	20/03/2017	13:00	138.94	2300
После промывки пусковой муфты	20/03/2017	14:06	136.57	2300
После промывки пусковой муфты	20/03/2017	14:30	136.4	2300
При стабилизации	20/03/2017	14:35	82.47	2300
При стабилизации	20/03/2017	14:39	83.3	2300
При стабилизации	20/03/2017	14:40	84.46	2300
При стабилизации	20/03/2017	14:41	85.51	2300
При стабилизации	20/03/2017	14:42	86.51	2300
При стабилизации	20/03/2017	14:43	87.5	2300
При стабилизации	20/03/2017	14:44	88.46	2300
При стабилизации	20/03/2017	14:45	89.42	2300
При стабилизации	20/03/2017	14:53	96.45	2300
При стабилизации	20/03/2017	15:28	119.03	2300
При стабилизации	20/03/2017	15:33	121.19	2300
При стабилизации	20/03/2017	15:50	127.7	2300
При стабилизации	20/03/2017	16:00	141.38	2300
При стабилизации	20/03/2017	16:20	142.06	2300
При стабилизации	20/03/2017	16:40	142.72	2300
При стабилизации	20/03/2017	16:00	143.28	2300
При стабилизации	20/03/2017	16:20	143.77	2300
При стабилизации	20/03/2017	16:40	144.28	2300
При стабилизации	20/03/2017	20:05	144.73	2300
При стабилизации	20/03/2017	20:28	145.11	2300
Дополнение:				
Выполнены детальные исследования в интервале 2140-2441,6м. Статистический уровень на уровне 5-6%. На фоне искусственного воздействия потока с глубины 2441,6м "тепла"				

Рисунок 1.1.3 – Образец заключения ПГИ

## 1.2. Общее описание применения методики VOI для ПГИ

В технической документации, регламентирующей способы контроля и мониторинга эксплуатационного фонда нефтяных скважин, указано, что исследования, направленные на установление факта негерметичности скважин и определение параметров заколонных перетоков должны проводиться при возникновении подозрений на негерметичность или в процессе ГИС-контроля.

Автоматизация контроля за параметрами работы скважин и выявлением аномалий обводненности позволит упростить работу технических специалистов компании. А предварительная оценка достижимого эффекта от проведения ПГИ/РИР позволит избежать проведения исследований на скважинах, которые заведомо нецелесообразно поддерживать в работоспособном состоянии.

Именно эту задачу призвана решить предлагаемая методика ранжирования скважин-кандидатов на ПГИ/РИР на основе ценности информации (VOI).

Для повышения эффективности расходов на исследования, предшествующие ремонтным работам, может быть применена методика оценки ценности информации.

Метод VOI (Value of information) в нефтегазовой промышленности впервые введён в 1960 году Джексоном Гейсоном (Gayson, C.J. 1960. Decisions under uncertainty: drilling decisions by oil and gas operators.). Ценность



информации (VOI) – это разница между ожидаемой ценностью проекта с учетом и без учета полученной информации.

В ситуации, когда принимаемое решение может привести к нескольким исходам, применяется термин ожидаемая ценность проекта (expected monetary value - EMV). Ожидаемая ценность проекта — это сумма произведений вероятности на чистую приведённую стоимость и предполагает учет различных вариантов исходов, возникающих в результате принятия конкретного решения и выраженного в денежных единицах (NPV - накопленный дисконтированный денежный доход), с учетом вероятности наступления этих исходов (P).

В рассматриваемой области методика VOI применима для определения целесообразности затрат на проведение ПГИ скважины, имеющей косвенные признаки негерметичности, перед принятием решения о проведении РИР.

Таким образом, для оценки ожидаемой ценности проекта и, соответственно, целесообразности проведения ПГИ необходимо сравнить экономический эффект от работы скважины в текущем состоянии (без проведения ПГИ и РИР) и в случае проведения исследований и возможного ремонта. Сопоставление ожидаемой ценности проекта (EMV – Expected Monetary Value) в прогнозный период по этим двум вариантам дерева решений и разность между ними (VOI – Value of Information) позволит принять обоснованное решение о проведении ПГИ или об отказе от него.

Расчет VOI позволяет экономически оценить целесообразность проведения исследования на конкретном объекте и, соответственно, выбрать в качестве целей исследования скважины с наибольшими показателями ценности информации.



Рисунок 1.2.1 – Дерево решений VOI

### 1.3. Схема расчета VOI для ПГИ

Применительно к поставленной задаче необходимо рассчитать ценность проекта в состоянии “как есть” ( $EMW_{woi}$ ), без проведения ПГИ и последующего ремонта и ценность проекта при проведении ПГИ/РИР ( $EMW_{wi}$ ).

$$VOI = EMV_{wi} - EMV_{woi}$$

$EMW_{woi}$  складывается из доходов от реализации нефти за вычетом затрат на утилизацию воды (в прогнозируемых объемах дебита за рассматриваемый период). Поскольку этот вариант предусматривает один сценарий, в данном случае применяется термин NPV (Net Present Value или текущая приведенная стоимость проекта).

$$EMV_{woi} = \sum P_i \cdot NPV_{woi}$$

$EMW_{wi}$  складывается из доходов от реализации нефти за вычетом затрат на утилизацию воды (в прогнозируемых объемах дебита) и затрат на проведение ПГИ и РИР. Длительность продуктивной работы скважины в прогнозный период в этом случае сокращается на количество дней, отводимых на ПГИ и РИР.

В связи с тем, что реализация ПГИ и РИР может иметь множество возможных исходов, связанных с рисками при проведении ПГИ и РИР, возможным отсутствием НЭК/ЗКЦ на скважине и с различной реализацией работ для НЭК и ЗКЦ – итоговая величина  $EMW_{wi}$  рассчитывается как сумма произведений NPV по рассматриваемым исходам на вероятность этих сценариев.

$$EMV_{wi} = \sum P_i \cdot NPV_{wi}$$

Таким образом, основными источниками неопределенности при расчете VOI являются:

1. Неопределенность технического состояния скважины, а именно наличия НЭК или ЗКЦ. Косвенные признаки негерметичности скважины, фиксируемые на основе анализа статистических показателей дебита и обводненности, могут быть результатом появления дефекта или проявляться под воздействием иных, внешних факторов.

2. Неопределенность исхода ПГИ. Проведение исследования технического состояния скважины не во всех случаях позволяет однозначно установить наличие или отсутствие негерметичности скважины и причину наблюдаемых изменений дебита и обводненности. Могут потребоваться повторные или расширенные исследования.

3. Неопределенность исхода РИР. Проведение ремонтно-изоляционных работ не во всех случаях устраняет выявленный дефект скважины и может потребовать дополнительных затрат, точно оценить которые до начала ремонта невозможно.

4. Неопределенность дебитов нефти, воды и уровня обводненности скважины на последующий период её работы без проведения РИР и после их проведения.

5. Нестабильность показателей, влияющих на экономическую эффективность РИР (колебания цен на нефть, увеличение материальных затрат на проведение исследовательских и ремонтных работ и оборудование).

6. Вероятности наступления исходов.

Для использования VOI применительно к ПГИ было построено дерево решений, включающее в себя различные варианты развития событий в случае проведения и не проведения ПГИ на скважине.

Изначально было выявлено 6 возможных сценариев для  $EMWwi$ , которые в рамках общего подхода к расчету VOI схематично отражены на рисунке 1.3.1.

NPV1 предполагает исход, при котором был проведен стандартный комплекс ПГИ, который не выявил наличия НЭК или ЗКЦ, в результате чего, дальнейших РИР не проводилось.

NPV2 предполагает исход, в результате которого был проведен стандартный комплекс ПГИ, результаты которого оказались недостаточно информативными, и в следствии чего было проведено дополнительное ПГИ, которое не выявило наличия НЭК или ЗКЦ, в результате чего, дальнейших РИР не проводилось.

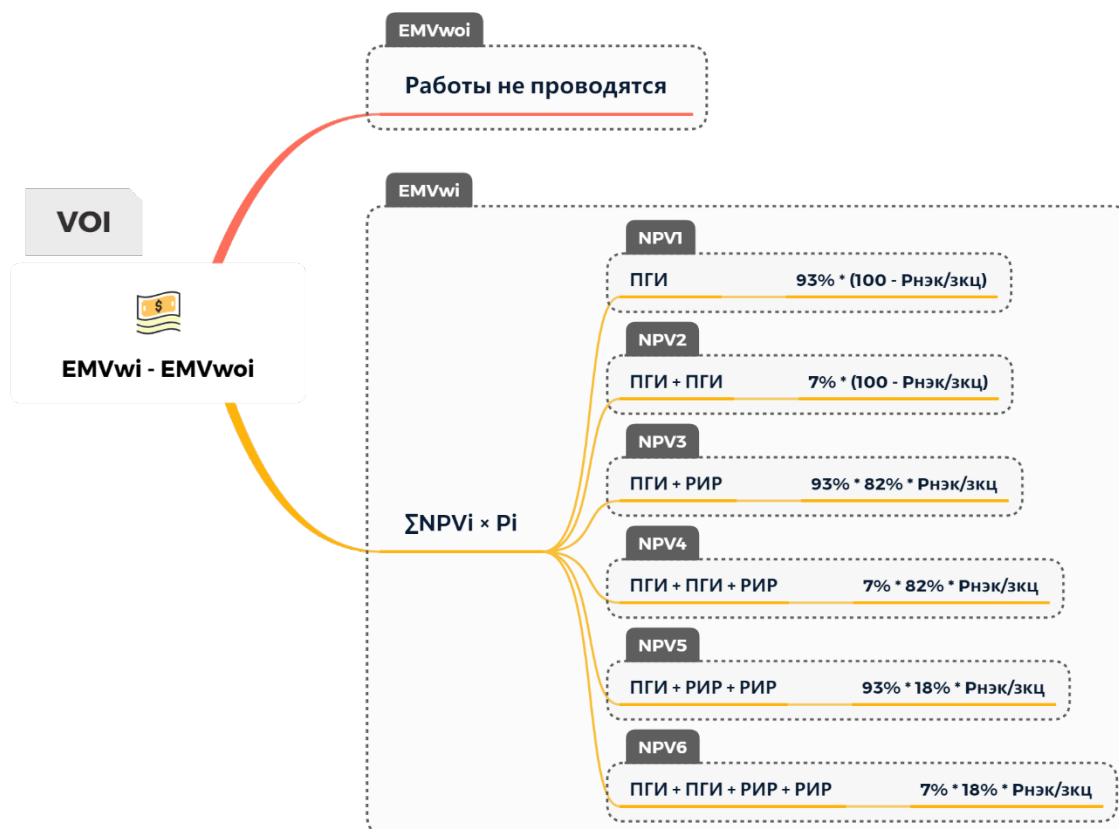


Рисунок 1.3.1 – Упрощенная схема сценариев расчета VOI

NPV3 предполагает исход, при котором был проведен стандартный комплекс ПГИ, после которого были проведены РИР.

NPV4 предполагает исход, при котором был проведен стандартный комплекс ПГИ, результаты которого оказались недостаточно информативными, и в следствии чего было проведено дополнительное ПГИ, после которого были проведены РИР.

NPV5 предполагает исход, при котором был проведен стандартный комплекс ПГИ, после которого были проведены не успешные РИР, вследствие чего были проведены повторные РИР.

NPV6 предполагает исход, при котором был проведен стандартный комплекс ПГИ. Первичное ПГИ оказалось не информативным, вследствие чего было проведено дополнительное ПГИ, после которого были проведены два комплекса РИР.

Сценарии NPV1 и NPV2 являются заведомо убыточными (затраты на исследования не приводят к улучшению показателей работы скважины, РИР не проводится).

Сценарии NPV3 – NPV6 предполагают, как затраты, так и увеличение доходов за счет устранения дефектов и снижения уровня обводненности скважины, а также потенциального увеличения дебита нефти.

В следствии того, что при различных дефектах скважины (НЭК или ЗКЦ) после РИР могут проводиться или не проводиться повторные ПГИ, для определения успешности РИР было принято решение расширить данную схему расчета вои добавлением дополнительных сценариев по типу дефекта. Результирующее дерево решений представлено на рисунке 1.3.2.

В данном варианте расчета предполагаются дополнительные сценарии при обнаружении НЭК, либо ЗКЦ, связанные с тем, что выявленный НЭК подразумевает варианты ремонта, после которых не требуется проведение ПГИ (двухпакерная компоновка).

Расчет ожидаемой ценности проекта без информации ( $EMV_{woi}$ ), когда не проводятся ПГИ и ремонтные работы, сводится к прогнозу дебита нефти и воды и оценке разницы между ожидаемыми доходами от нефти и затратами на утилизацию воды. Прогнозы дебита и обводненности строятся на основе предшествующей динамики для каждой скважины.

Ожидаемая ценность проекта при проведении исследований и/или ремонтных работ ( $EMV_{wi}$ ) оценивается на основе предшествующей динамики показателей с учетом их улучшения после ПГИ и ремонтных работ и за вычетом затрат на проведение работ и потерь от вынужденных простоев.

Ключевым фактором, позволяющим сделать  $VOI$  положительным, является снижение показателя обводненности скважины, которое можно ожидать в результате проведения ремонтных работ, следующих за ПГИ, и потенциально возможное повышение дебита нефти. На первом этапе исследования выдвигалась гипотеза, что оба этих показателя улучшаются в результате проведения ПГИ/РИР.

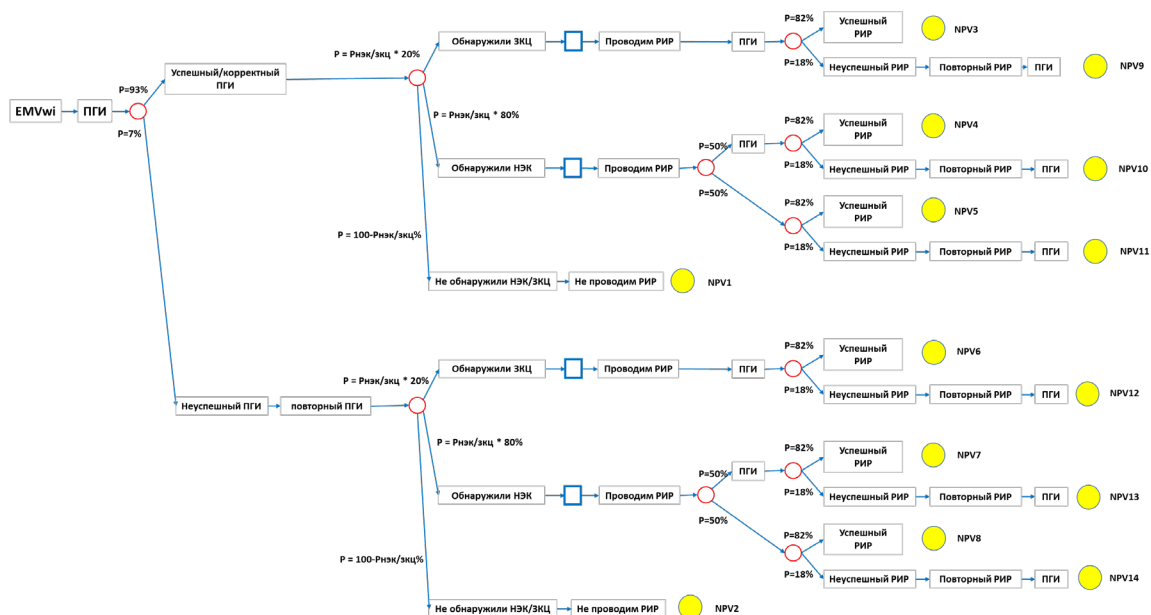


Рисунок 1.3.2 – Расширенная схема сценариев расчета VOI

Анализ фактических данных показал, что для показателей обводненности и дебита жидкости можно с достаточно высокой вероятностью прогнозировать снижение, в то время как для показателя дебита нефти наблюдаемая после РИР динамика менее предсказуема и не позволяет делать однозначных прогнозов.

Для оценки вероятностей сценариев были необходимы следующие данные:

- вероятность успешного исхода ПГИ;
- вероятность успешного исхода РИР;
- вероятность наличия дефекта (НЭК или ЗКЦ) на скважине;
- распределение вероятностей того является ли дефект НЭК, либо ЗКЦ;
- вероятность типа требуемого ремонта в случае НЭК (двухпакерная компоновка или спуск ЭЦН с пакером).

Все вероятности, кроме вероятности наличия дефекта были рассчитаны статистически на основе исходных данных и отчетов по проведенным ранее ПГИ/РИР.

Для расчета вероятности наличия дефекта (НЭК или ЗКЦ) на скважине было принято решение использовать отдельную предсказательную модель.

### 3.1.1. Логистическая регрессия

Логистическая регрессия применяется для прогнозирования вероятности возникновения некоторого события по значениям множества признаков. Для этого вводится так называемая зависимая переменная  $y$ , принимающая лишь одно из двух значений — как правило, это числа 0 (событие не произошло) и 1 (событие произошло), и множество независимых переменных (также называемых признаками, предикторами или регрессорами) — вещественных  $x_1, \dots, x_n$ , на основе значений которых требуется вычислить вероятность принятия того или иного значения зависимой переменной. Как и в случае линейной регрессии, для простоты записи вводится фиктивный признак  $x_0 = 1$ .

Делается предположение о том, что вероятность наступления события  $y = 1$  равна:

$$P\{y = 1 \mid x\} = f(z)$$

$$z = \theta^T x = \theta_0 + \theta_1 x_1 + \dots + \theta_n x_n,$$

где  $x$  и  $\theta$  — векторы-столбцы значений независимых переменных  $1, x_1, \dots, x_n$  и параметров (коэффициентов регрессии) — вещественных чисел  $\theta_0, \dots, \theta_n$ , соответственно, а  $f(z)$  — так называемая логистическая функция:

$$f(z) = \frac{1}{1 + e^{-z}}$$

Так как  $y$  принимает лишь значения 0 и 1, то вероятность принять значение 0 равна:

$$P\{y = 0 \mid x\} = 1 - f(z) = 1 - f(\theta^T x)$$

Для краткости функцию распределения  $y$  при заданном  $x$  можно записать в таком виде:

$$P\{y \mid x\} = f(\theta^T x)^y (1 - f(\theta^T x))^{1-y}, \quad y \in \{0, 1\}$$

Фактически, это есть распределение Бернулли с параметром, равным  $f(\theta^T x)$ .

**Список использованных источников**

1. Шорохов А.Н., Азаматов М.А., Сохошко С.К. Аналитический метод диагностики негерметичности эксплуатационной колонны для нефтяных добывающих скважин. Статья в журнале "Вестник ЦКР Роснедра", 2014 (4). Издательство "Научно-исследовательский и проектный институт нефти и газа" (Москва). ISSN: 2074-5966.
2. Akoglu L., Tong H., Koutra D. Graph based anomaly detection and description: a survey // *Data Min. Knowl. Discov.* 2015. Т. 29. № 3. С. 626–688.
3. Akouemo H.N., Povinelli R.J. Probabilistic anomaly detection in natural gas time series data // *Int. J. Forecast.* 2016. Т. 32. № 3. С. 948–956.
4. Arps J.J. Analysis of decline curves. SPE-945228-G, *Trans. of the AIME* (1945).
5. Bagnall A. и др. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances // *Data Min. Knowl. Discov.* 2017. Т. 31. № 3. С. 606–660.
6. Bailey, B., J. Tyrie, J. Elphick, F. Kuchuk, C. Romano, L. Roodhart. 2000. Water Control. *Oilfield Review*, Schlumberger 12 (1): 30-51.
7. Bica A.M. Optimizing at the end-points the Akima's interpolation method of smooth curve fitting // *Comput. Aided Geom. Des.* 2014. Т. 31. № 5. С. 245–257.
8. Cao N. и др. TargetVue: Visual Analysis of Anomalous User Behaviors in Online Communication Systems // *IEEE Trans. Vis. Comput. Graph.* 2016. Т. 22. № 1. С. 280–289.
9. Chang I., Tiao G.C., Chen C. Estimation of Time Series Parameters in the Presence of Outliers // *Technometrics.* 1988. Т. 30. № 2. С. 193–204.
10. Chan, K.S. 1995. Water Control Diagnostic Plots. Paper SPE 30775, presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, USA, October 22-25.



11. Clegg, J. D. 2007. Production Operations Engineering. Edited by L. W. Lake. Vol. 4, Petroleum engineering handbook. Texas, USA: Society of Petroleum engineers.
12. Guo, B., W. C. Lyons, and A. Ghalambor. 2007. Petroleum Production Engineering: A Computer-Assisted Approach: Gulf Professional Publishing.
13. Gabdullaevich Akhmadeev R. и др. Assessment of the Tax Base of the Consolidated Group of Taxpayers in Russia using the Method of Polynomial Interpolation.
14. Hyndman R.J., Wang E., Laptev N. Large-Scale Unusual Time Series Detection // 2015 IEEE International Conference on Data Mining Workshop (ICDMW). : IEEE, 2015. С. 1616–1619.
15. Keogh E., Kasetty S. On the Need for Time Series Data Mining Benchmarks: A Survey and Empirical Demonstration // Data Min. Knowl. Discov. 2003. Т. 7. № 4. С. 349–371.
16. Luis Martí, Nayat Sanchez-Pi, José Manuel Molina, Ana Cristina Bicharra Garcia. Anomaly Detection Based on Sensor Data in Petroleum Industry Applications. Article in Sensors 2015,15, 2774-2797; doi:10.3390/s150202774. ISSN 1424-8220.
17. Minou Rabiei. Excess Water Production Diagnosis in Oil Fields Using Ensemble Classifiers. Thesis presented for the Degree of Doctor of Philosophy of Curtin University. October 2011.
18. Pedregosa F. и др. Scikit-learn: Machine Learning in Python // J. Mach. Learn. Res. 2011. Т. 12. № Oct. С. 2825–2830.
19. Reidar B. Bratvold. Value of Information in the Oil and Gas Industry: Past, Present, and Future. Article in SPE Reservoir Evaluation & Engineering. August 2009 DOI: 10.2118/110378-MS.
20. Rizzi S. и др. Comparing non-parametric methods for ungrouping coarsely aggregated age-specific distributions // J. Aviat. Educ. Res. 2016.
21. Vallis O., Hochenbaum J., Kejariwal A. A Novel Technique for Long-Term Anomaly Detection in the Cloud // 2014.