

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

«\_\_\_»\_\_\_\_\_2021г.

## **ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

магистерская диссертация

### **АЛГОРИТМ АВТОМАТИЧЕСКОГО ПРИНЯТИЯ РЕШЕНИЯ ПО СПОСОБУ РАСЧЕТА ТР НА ОСНОВАНИИ МЕТОДОВ СТАТИСТИЧЕСКОГО АНАЛИЗА**

по направлению подготовки

01.04.03 Механика и цифровое математическое моделирование

профиль

01.04.03\_03 Механика и цифровое производство

Выполнил  
студент гр. 3640103/90301

Н.В. Огуречников

Руководитель  
к.ф.-м.н., ст. преп.

Е.В. Юдин

Санкт-Петербург – 2021

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ  
ПЕТРА ВЕЛИКОГО

Институт прикладной математики и механики

УТВЕРЖДАЮ  
Директор  
Высшей школы  
теоретической механики  
А.М. Кривцов  
«\_\_» \_\_\_\_\_ 2021г.

**ЗАДАНИЕ**

**на выполнение выпускной квалификационной работы**

студенту Огуречников Николай Владимирович, гр.3640103/90301

- 1. Тема работы:** Алгоритм автоматического принятия решения по способу расчета ТР на основании методов статистического анализа.
- 2. Срок сдачи студентом законченной работы:** 07.06.2021
- 3. Исходные данные по работе:** имеющаяся проблематика с данными для формирования Технологического Режима.
- 4. Содержание работы (перечень подлежащих разработке вопросов):** анализ текущей методики ТР на ДО, изучение методов статистического анализа данных, реализация методики автоматического принятия решения для расчета ТР.
- 5. Перечень графического материала (с указанием обязательных чертежей):** отсутствуют. График разброса значений по дебиту жидкости. Алгоритм определения анализируемого периода. Метод статистического анализа. Основные этапы алгоритма.
- 6. Дата выдачи задания:** 25.01.2021

Руководитель ВКР \_\_\_\_\_ Е.В. Юдин  
(подпись)

Задание принял к исполнению 25.01.2021

Студент \_\_\_\_\_ Н.В. Огуречников  
(подпись)

## РЕФЕРАТ

На стр. 46, 27 рисунков, 2 таблиц, 3 приложений, 12 литературных источника.

**КЛЮЧЕВЫЕ СЛОВА:** ТЕХНОЛОГИЧЕСКИЙ РЕЖИМ, СТАТИСТИЧЕСКИЙ АНАЛИЗ, ЛИНЕЙНАЯ АППРОКСИМАЦИЯ, АВТОМАТИЗАЦИЯ.

Тема выпускной квалификационной работы: «Алгоритм автоматического принятия решения по способу расчета ТР на основании методов статистического анализа».

Данная работа посвящена стандартизации алгоритмов по автоматическому выбору методики расчёта фактического Технологического режима (ТР). Задачи, которые решались в ходе исследования:

1. Изучение методики формирования ТР на ДО.
2. Выявление проблематики при формировании ТР.
3. Изучение методики по автоматическому выбору способа расчета на основе статистического анализа.
4. Реализация методики автоматического принятия решения для расчета ТР.

Данная работа посвящена реализации автоматизированной методики принятия решения при формировании месячного ТР с целью повышения качества данных и ускорения процесса формирования ТР. Рассмотрен метод статистического анализа. Описан подход проведения линейной аппроксимации функции зависимости дебита жидкости от времени работы скважины. Реализован прототип автоматического принятия решения по способу расчета ТР на основании методов статистического анализа. Были проведены сравнительные расчеты, показывающие возможность применения данного подхода.

## ABSTRACT

On page 39, 27 figures, 2 tables, 3 appendices, 12 literary source.

**KEY WORDS:** TECHNOLOGICAL MODE, STATISTICAL ANALYSIS, LINEAR APPROXIMATION, AUTOMATION.

The theme of the final qualification work: "Algorithm for automatic decision-making on the method of calculating TR based on methods of statistical analysis." This work is devoted to the standardization of algorithms for the automatic selection of the methodology for calculating the actual Technological Mode (TR). Tasks that were solved during the study:

1. Study of the methodology for the formation of TR at the DO.
2. Identification of problems in the formation of TR.

3. Study of the methodology for the automatic selection of the calculation method based on statistical analysis.
4. Implementation of the method of automatic decision-making for calculating TR. This work is devoted to the implementation of an automated decision-making technique in the formation of monthly TR in order to improve the quality of data and accelerate the process of TR formation. The method of statistical analysis is considered. An approach to linear approximation of the function of fluid flow rate dependence on the well operation time is described. A prototype of automatic decision-making on the method of calculating TR based on the methods of statistical analysis has been implemented. Comparative calculations were carried out showing the possibility of applying this approach.

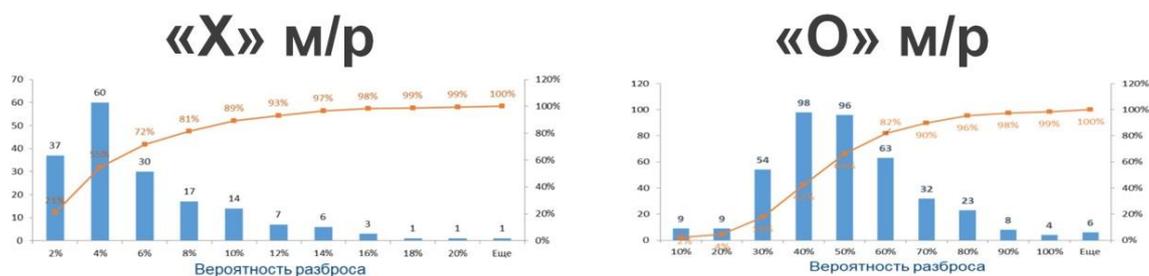
## СОДЕРЖАНИЕ

Введение.....	7
Глава 1. Жизненный цикл месторождений нефти и способы её добычи.....	8
1.1. Жизненный цикл месторождений.....	8
1.2. Способы эксплуатаций нефтяных скважин.....	9
1.3. Технологический режим .....	14
Глава 2. Методология и прототип Автоматизации ТР.....	15
2.1. Общее описание методики.....	15
2.2. Определение периода расчета по замерам дебита жидкости .....	16
2.3. Скважина в режиме накопления.....	17
2.4. Метод статистического анализа.....	19
2.5 Алгоритм расчета.....	19
2.6 Результаты анализа и выбор алгоритма расчета.....	21
2.7 Получение граничных коэффициентов.....	23
2.8 Расчет остальных параметров на основании решения $Q_{ж}$ .....	23
2.9 Алгоритм расчета средних значений параметров Буферное давление и Давление в линии.....	24
2.10 Q-тест Диксона для выявления выбросов для небольших выборок.....	25
2.11 Алгоритм расчета среднего значения параметра затрубное давление ( $P_{затр}$ ).....	27
2.12 Алгоритм расчета средних значений параметров Давление газлифтного газа ( $P_{г/г}$ ).....	28
2.13 Информационные потоки модуля.....	29
2.14 Блок-схема прототипа автоматического принятия решения по способу расчета ТР на основании методов статистического анализа.	30
2.15 Реализация блок автоматического принятия решения по способу расчета ТР на основании методов статистического анализа.....	36
2.16 Практический результат.....	39

Заключение.....	41
Список использованных источников.....	42
Приложение 1. Основные этапы алгоритма.....	44
Приложение 2. Метод Диксона.....	45
Приложение 3. Таблица выгрузки данных.....	46

## ВВЕДЕНИЕ

По умолчанию на всех объектах где проводится добыча нефти при формировании ТР берут последние значения, которые были получены в ходе добычи. Но есть объекты, где данные по добыче очень разнятся с предыдущими периодами и принятый метод для расчета ТР не работает.



### «X» месторождения

- разброс значений по Qж составляет всего лишь 4%

### «O» месторождение

- разброс значений по Qж составляет более чем в 50%

Рисунок 1. График разброса значений по дебиту жидкости(Qж).

На месторождении, где разброс значений по дебиту жидкости (Qж) составляет более чем в 50%, выбирая срез определенного дня для расчета ТР, можно ошибиться на 50% (Рисунок 1). Поэтому на некоторых месторождениях не используется метод среза на определенную дату. А берут среднее значение за период, что бы нивелировать разброс ошибок в данных. Или срез на определённую дату. Пользователи каждую скважину рассматривает индивидуально, экспертно принимают решение по методу выбора оценки ТР, и вручную формируют отчет в Excel.

Главная проблематика, которую мы будем решать в данной работе, заключается в том, чтобы автоматизировать выбор метода расчета на месторождениях.

## ГЛАВА 1. Жизненный цикл месторождений нефти и способы её добычи.

### 1.1 Жизненный цикл месторождений.

Нефтяное месторождение при разработке проходит определенный жизненный цикл[2]:

- 0 стадия – поиск и разведка запасов;
- 1 стадия – интенсивное освоение месторождения;
- 2 стадия – стабилизация и выход на максимальный уровень добычи;
- 3 стадия – равномерное падение добычи;
- 4 стадия – завершающая стадия разработки, заметное сокращение добычи.

Наглядно увидеть этапы разработки мы можем на Рисунке 2.

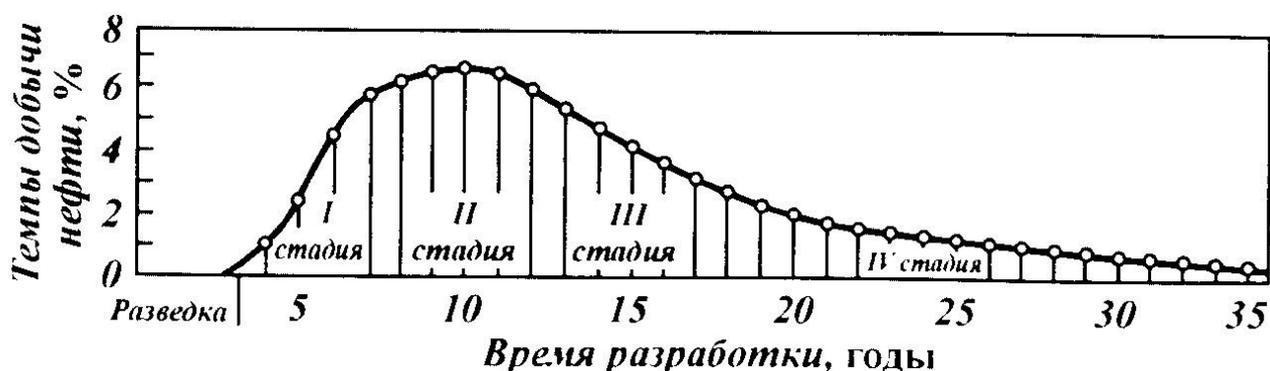


Рисунок 2. Этапы разработки месторождения.

0 стадия – Перед началом добычи нефти её необходимо найти, провести оценку запасов, определить их потенциал и доступность. При определении достаточного количества сырья переходят к разработке месторождения. На нулевом этапе добыча не производится, идет подготовка. Производится исследование пласта, бурение поисково-разведочных скважин.

1 стадия - На ней производится интенсивное бурение и освоение месторождения. В этот период непрерывно увеличивается темп разработки, что приводит к максимальному значению к концу стадии. Обычно первой стадии добывают безводную нефть.

2 стадия – Характеризуется стабильной добычей. На второй стадии достигается максимальная добыча нефти и в планах указывается год, когда это произойдет. Основная задача стадии заключается в регулировании режимов добычи нефти и в полной мере использовать систему заводнения. Если скважина к концу периода перестает фонтанировать, её переводят на механизированный способ эксплуатации.

3 стадия – На данной стадии происходит заметное снижение добычи на фоне обводненности при водонапорном режиме или повышенного газового

фактора при газонапорном режиме. Большая часть скважин переведена на механизированный способ добычи. Значительная часть скважин к концу периода выбывает из эксплуатации.

4 стадия - характеризуется низкими темпами разработки. Наблюдаются высокая обводненность продукции и медленное уменьшение добычи нефти.

Завершающий комплекс работ, который необходимо совершить, когда скважина перестала быть рентабельной:

- прекращение добычи нефти
- скважина ликвидируется или консервируется
- лицензия на разработку возвращается государству.

## 1.2 Способы эксплуатаций нефтяных скважин.

В современном мире при решении задач связанных с добычей нефти и газа применяется способ бурения скважин. Он позволяет добраться до извлекаемых запасов ископаемых и залежей в пластах. Перед эксплуатацией нефтяных скважин происходит выбор способа добычи. Способ добычи зависит от ряда факторов связанных с величиной пластового давления и глубиной залегания пласта. На способ эксплуатации так же влияет степень обводненности, напор жидкости в стволе, состав нефти и газа. Важной составляющей в выборе способа эксплуатации скважины является пластовая энергия. Исходя, из данных по ней происходит выбор между фонтанным насосным или газлифтным способом эксплуатации скважины. Данные способы относятся к механизированной добыче полезных ископаемых[2].

### *Фонтанный способ эксплуатации скважин*

Способ, при котором нефть по насосно-компрессорным трубам поднимается на поверхность за счёт гидростатического напора или энергии расширяющего газа. При данном способе эксплуатации особое оборудование не требуется (Рисунок 3). Состав наземного оборудования для фонтанной скважины:

- колонная головка
- фонтанная арматура
- выкидная линия

Подземное оборудование: - Колонны насосно-компрессорных труб (НКТ) (они спускаются до верхних дыр перфорации).

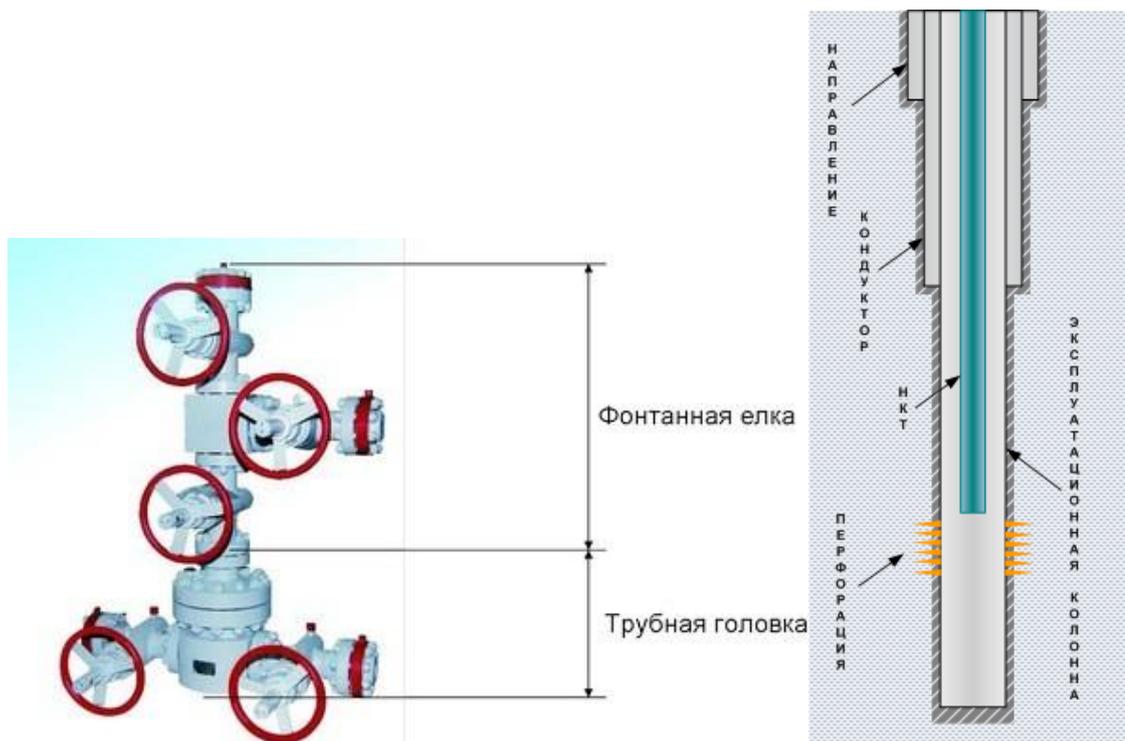
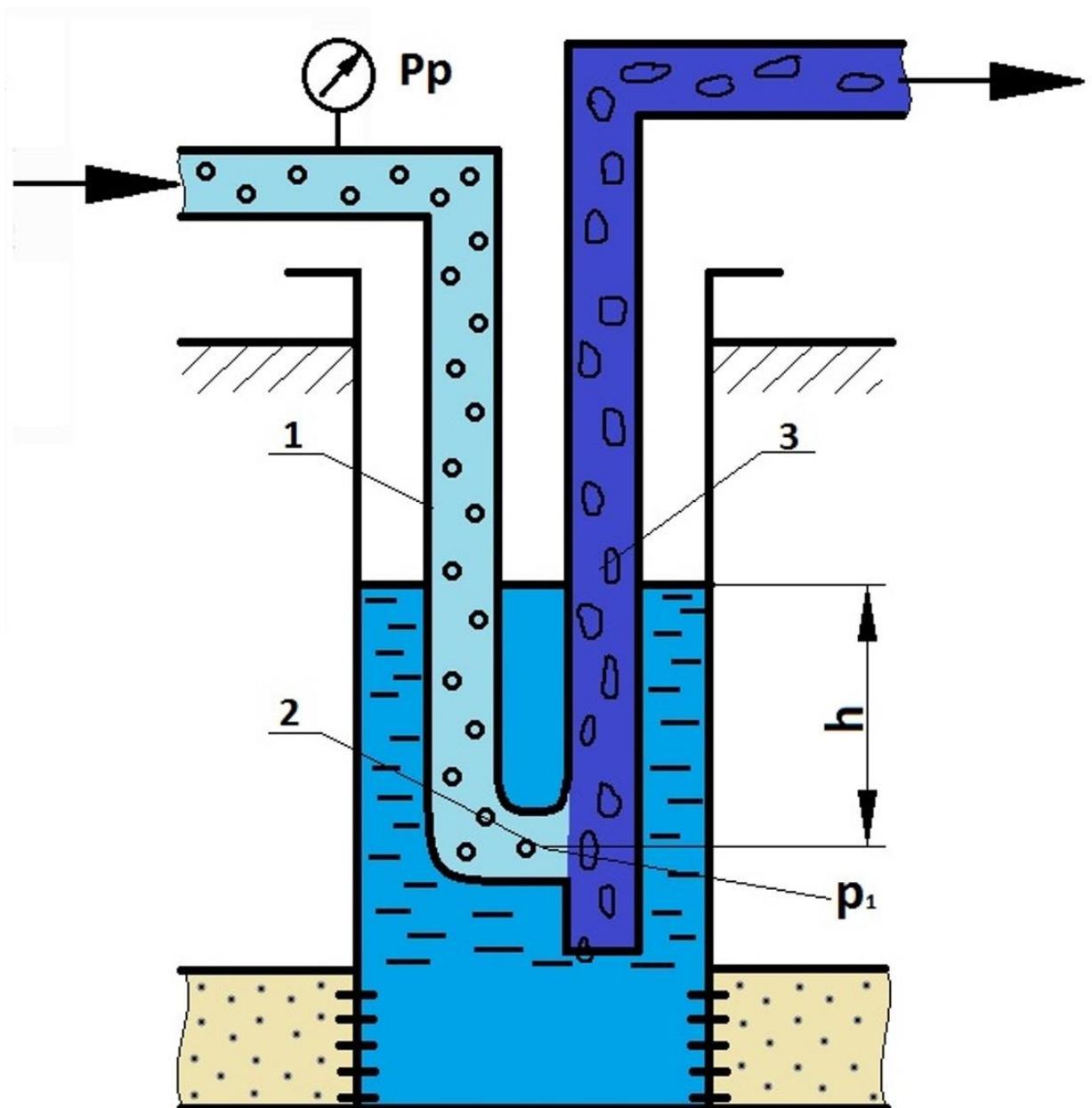


Рисунок 3. Фонтанная крестовая арматура, схема скважины.

### *Газлифтный способ эксплуатации скважин*

Газлифтный способ эксплуатации является логически возможным продолжением фонтанного способа эксплуатации. Когда подъем жидкости перестает проходить из-за недостаточного количества газа в пласте, газ начинают закачивать извне. В нефтяную скважину газ можно закачать из газового пласта без его дополнительной компрессии. Газлифтный способ эксплуатации характеризуется высокой эффективностью, отсутствием трущихся деталей, простотой в обслуживании и регулировании режимов добычи. При устройстве газлифтной скважины используют дополнительно два ряда насосных труб, которые опускают в затрубное пространство и подают туда газ или воздух. Есть труба воздушная, а есть труба подъемная. При подаче газа плотность нефти изменяется, и газированная жидкость будет подниматься в систему сбора (Рисунок 4).



**1-колонна труб; 2-башмак; 3-подъемная труба**

Рисунок 4. Схема Газлифтный способ эксплуатации скважин.

### *Насосный способ эксплуатации скважин*

При насосном способе эксплуатации скважины применяются различные типы насосов:

- Штанговые (Рисунок 5)
- погружные
- электроцентробежные
- винтовые и др. (Рисунок 6).

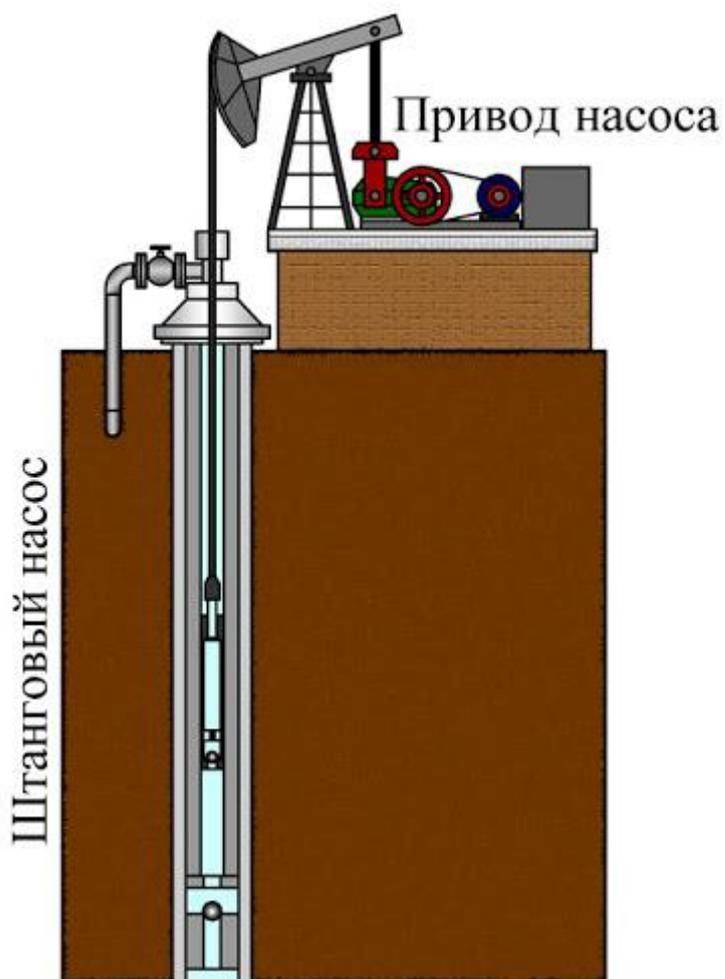


Рисунок 5. Штанговый насос

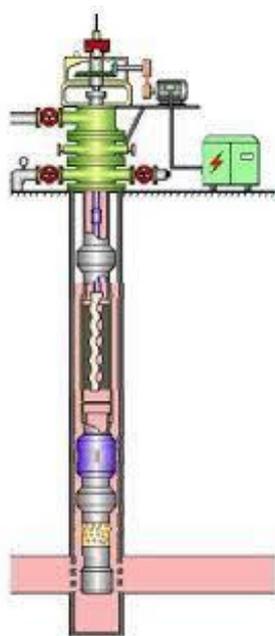


Рисунок 6. Винтовой насос.

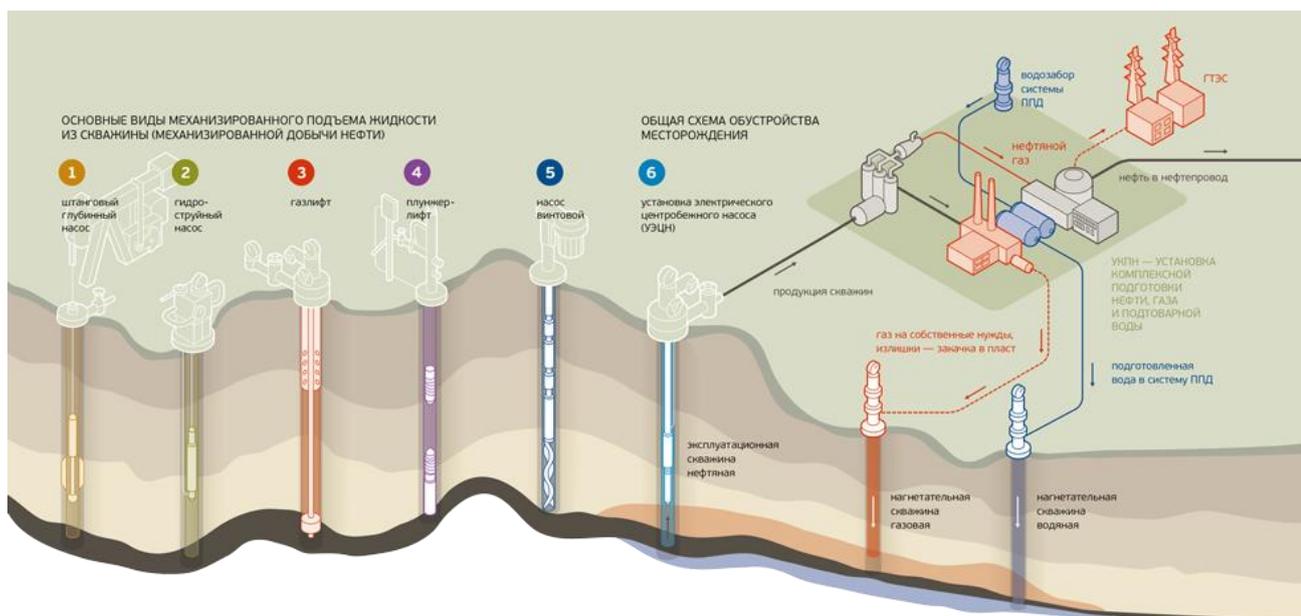


Рисунок 7. Основные виды механизированного подъема жидкости из скважины (механизированной добычи нефти).

### 1.3 Технологический режим

Совокупность основных параметров работы скважины обеспечивающих плановый отбор нефти, жидкости, газа и конденсата предусмотренные технологическим проектом при соблюдении условий надежности и безопасности эксплуатации скважины и является Технологическим режимом (ТР). [10]

Технологический режим работы добывающих скважин характеризуется следующими основными параметрами:

а) пластовым, забойным и устьевым давлениями, а для месторождений, содержащих свободный газ, также устьевой температурой;

б) дебитом жидкости (газа), обводненностью продукции, газовым фактором (выходом конденсата) и количеством механических примесей в продукции;

в) типоразмерами установленного внутрискважинного оборудования, режимами и временем его работы.[10]

ТР скважины назначается не реже одного раза в квартал и утверждается недр пользователем на основе проектных показателей. Исходя из утвержденного и проведенного плана ГТМ, составляется ТР.

При эксплуатации скважин важно проводить регулярный контроль технического состояния эксплуатационной колонны, работы оборудования, получения данных для корректировки и оптимизации ТР. Этапы формирования ТР представлены на Рисунке 8.

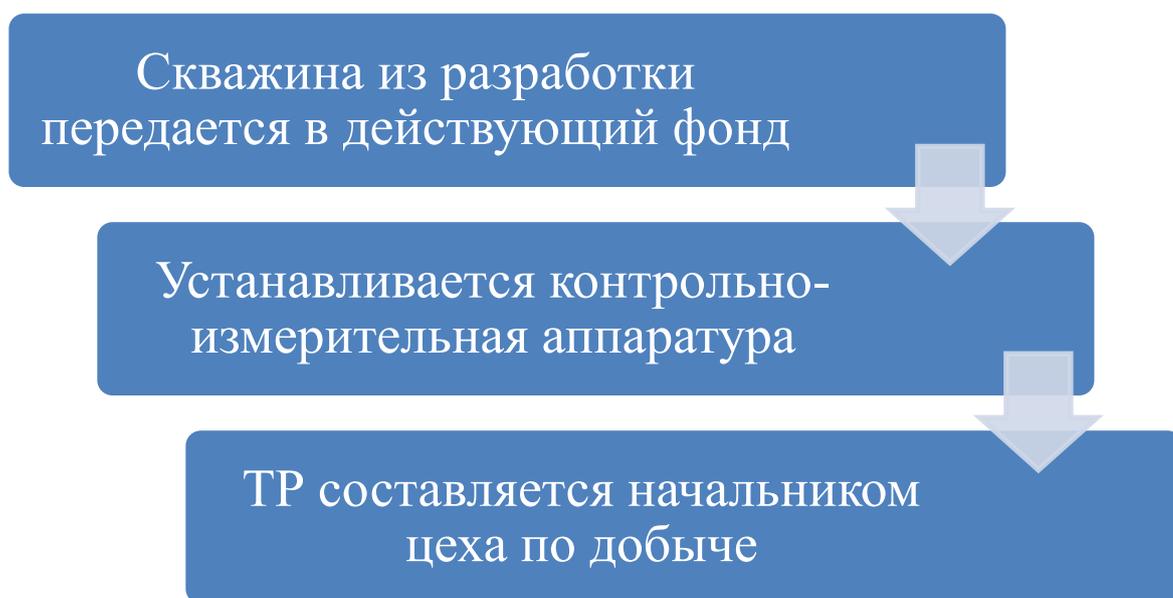


Рисунок 8. Схема по формированию ТР.

## ГЛАВА 2. Методология и прототип Автоматизации ТР.

### 2.1 Общее описание методики.

При формировании методики необходимо взять величину, которая более точно отражала чистоту данных. Дебит жидкости(Qж) замеряется через групповую замерную установку (ГЗУ) и является величиной, от которой мы можем отталкиваться при построении методики. Ниже описаны основные этапы алгоритма (Приложение 1):

1. Определяем период расчета на основании Qж (Рисунок 9). За основу берется скользящий месяц или квартал в зависимости от количества замеров и проведенных мероприятий (Рисунок 10).
2. Статистический анализ значений Qж за выбранный период.
3. Выбор последнего или среднего значения по Qж на основании граничных значений по параметрам статистического анализа (Рисунок 12).
4. Расчет остальных параметров на основании решения по Qж (Рисунок 13).
5. Для исключения выбросов по параметрам буферного и линейного давления дополнительно использует метод Диксона[1] (Приложение 2).

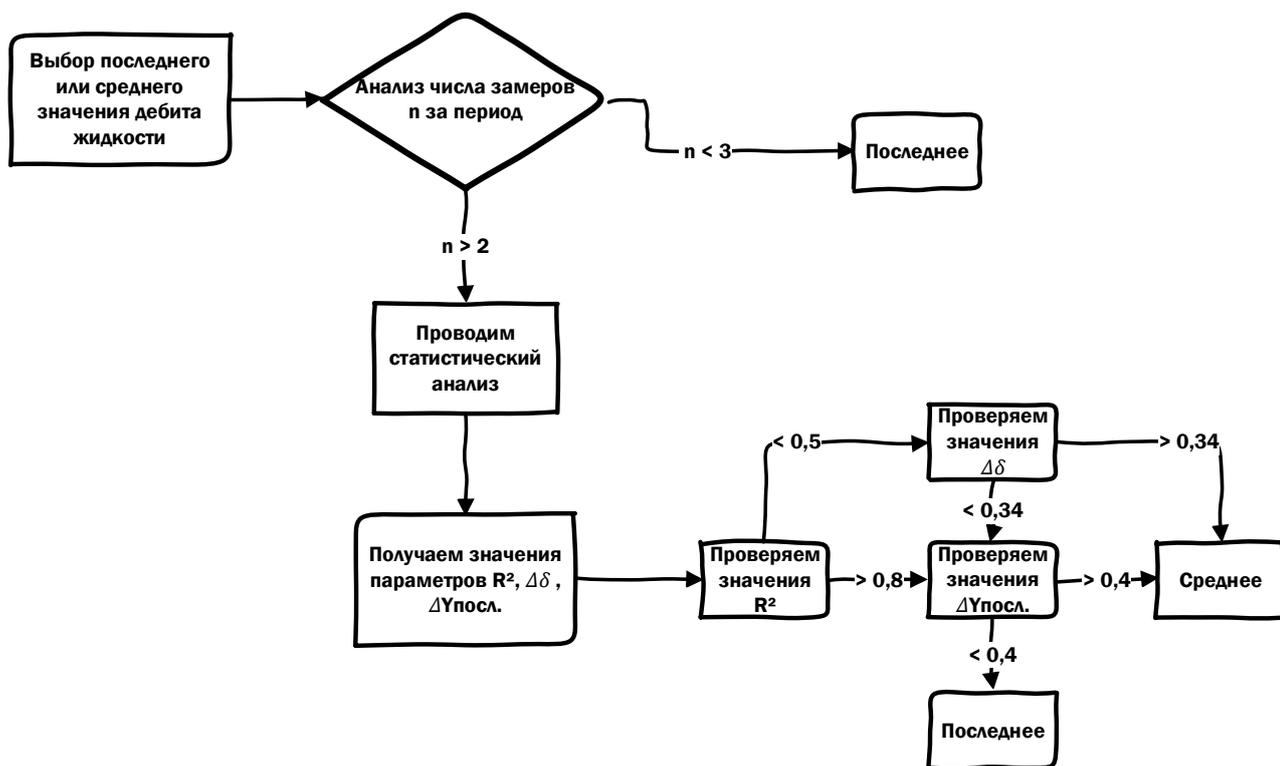


Рисунок 9. Логическая схема выбора значения дебита жидкости.

## 2.2 Определение периода расчета по замерам дебита жидкости

За анализируемый период принимается скользящий месяц на дату формирования ТР. Предположим ТР на май месяц формируется 17 апреля, тогда анализируются замеры  $Q_{ж}$  с 17 марта по 17 апреля.

Начало анализируемого периода сдвигается в зависимости от того, проводилось мероприятие на скважине, и от количества замеров в отчетный период:

- Если в скользящем месяце проводился ТКРС, то начало анализируемого периода смещается на дату запуска скважины в работу. (Если мероприятие было 20 марта, то анализируемый период будет считаться с 20 марта по 17 апреля)

- Если за отчетный период ТКРС не проводился, то оценивается количество замеров по дебиту жидкости:

- Если количество замеров больше 3 и не было мероприятий, то рассчитывает статистический анализ.
- Если число замеров меньше трех, то анализируемый период увеличивается до скользящего квартала. (Период увеличивается с 17 января по 17 апреля)

- В скользящем квартале также проверяется, наличие мероприятия, и анализируется количество замеров:
  - Если было мероприятие, то начало анализируемого периода смещается на дату запуска скважины в работу.
  - Если мероприятий ГТМ, ТКРС не было, то проверяем количество замеров
    - ✓ Если количество замеров больше 8 замеров, то дата начала анализируемого периода смещается, чтобы в расчет попадали только последние 8 замеров.
    - ✓ Если количество замеров меньше 3 замеров, то берется последнее значение.
    - ✓ Если от 3 до 9, рассчитываем статистический анализ.

Выбранный период по Qж используется для расчет средних значений остальных параметром ТР.

### 2.3.Скважина в режиме накопления.

При условии если скважина работает в накоплении больше 24 часов, мы вводим поправку с учетом времени накопления.  $k_{\text{накоп}}=t_{\text{раб}}/(t_{\text{раб}}+t_{\text{накоп}})$ . Для этого каждое из значений замера дебита жидкости умножаем на полученный коэффициент накопления. И статистический анализ проводится для полученных (уточненных) значений ( $Q_{\text{жнакоп}}$ ).

Формулы расчета:

$$k_{\text{накоп}}=t_{\text{раб}}/(t_{\text{раб}}+t_{\text{накоп}}),$$

Где  $k_{\text{накоп}}$  – поправочный коэффициент (час/час),

$t_{\text{раб}}$  – период работы скважины (час) в анализируемом периоде (час),

$t_{\text{накоп}}$  – период работы скважины в режиме накопления (час).

$$Q_{\text{жнакоп}}=Q_{\text{ж}} \cdot k_{\text{накоп}},$$

Где  $Q_{\text{жнакоп}}$  – уточненное значение дебита жидкости (т/сут),

$Q_{\text{ж}}$  – дебит жидкости (т/сут).

Если за анализируемый период (после корректировки на ТКРС и количество замеров) количество замеров  $n$  меньше трех ( $n < 3$ ), то в ТР берется последнее значение дебита жидкости. Данный результат передается дальше для расчета фактического технологического режима скважины и в расчет прочих параметров работы скважины.

Если замеров больше двух, то для выбора метода расчета параметра ТР проводится статистическая обработка значений замеров Дебита жидкости.

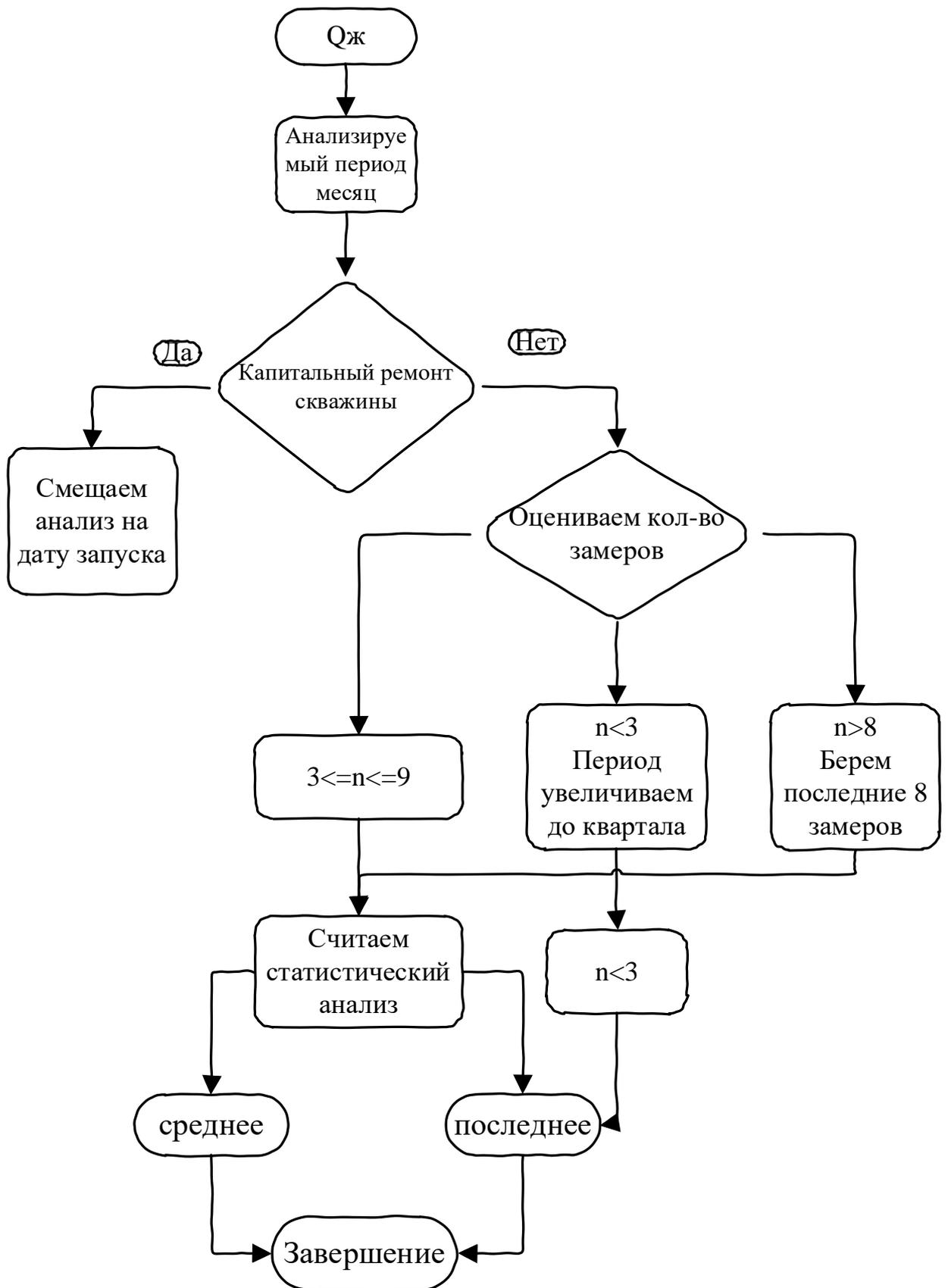


Рисунок 10. Алгоритм определения анализируемого периода.

## 2.4. Метод статистического анализа

Для проведения статистического анализа Дебита жидкости вычисляются следующие 3 параметра:

$R^2$  - величина достоверности аппроксимации. Показывает, как данные отклоняются от тренда.

$\Delta\delta$  - Процент отношения стандартного отклонения к среднему значению. Показывает, как данные отклоняются от среднего значения в выборке.  $\Delta\delta = \delta / \bar{Y}$  (Рисунок 11).

$\Delta Y_{\text{посл}}$  - Отношение отклонения значения от линии тренда для последнего значения в месяце. Показывает на сколько последнее значение отклоняется от тренда.  $\Delta Y_{\text{посл}} = |Y_{\text{посл}} - Y_{\text{тренда}}| / Y_{\text{посл}}$ .

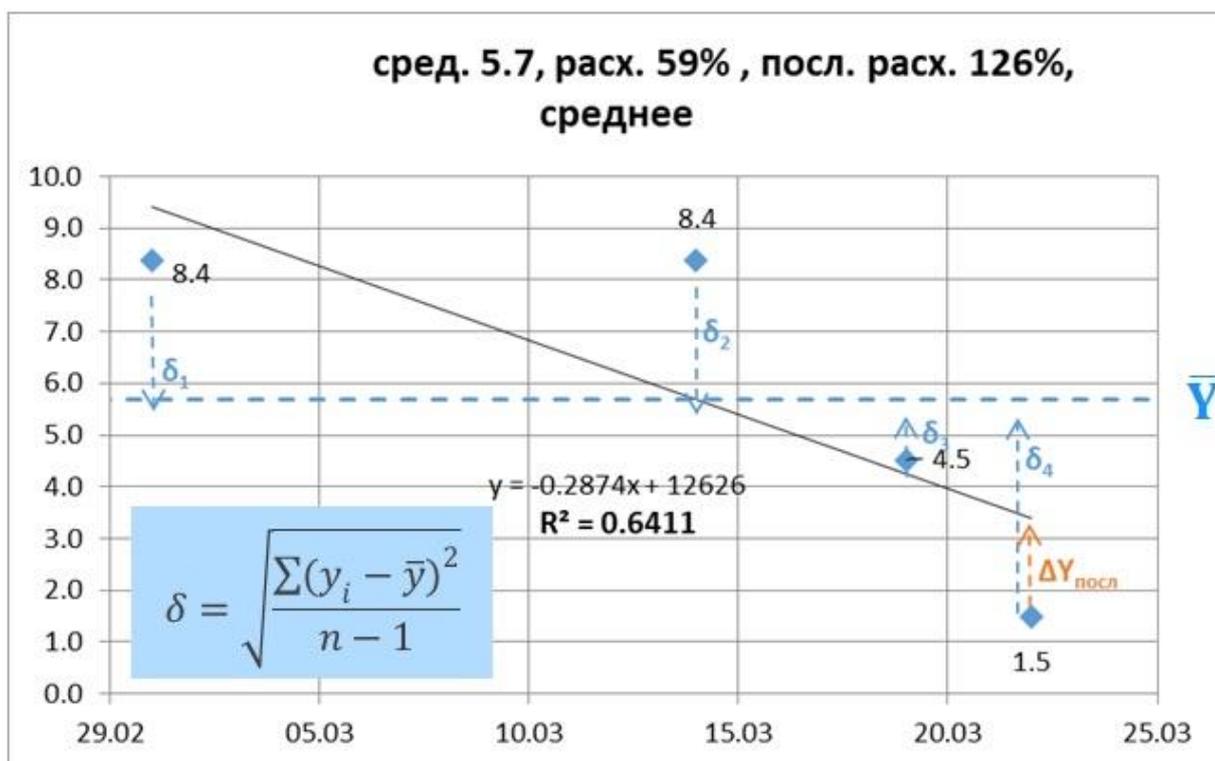


Рисунок 11. Метод статистического анализа.

## 2.5. Алгоритм расчета.

2.5.1 Проводится линейная аппроксимация функции зависимости Дебита жидкости от времени работы скважины (на дату замера). При линейной аппроксимации связи между двумя параметрами для нахождения эмпирических коэффициентов линейной функции используется метод наименьших квадратов. Суть метода состоит в том, что линейная функция (вида  $y = a + b \cdot x$ ) «наилучшего соответствия» проходит через точки графика,

соответствующие минимуму суммы квадратов отклонений измеряемого параметра. Такое условие имеет вид[4]:

$$s = \sum_{i=1}^{i=n} (y_i - a - b \cdot x_i)^2 \rightarrow \min$$

$$n \cdot a + b \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$a \cdot \sum_{i=1}^n x_i + b \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \cdot y_i$$

$$b = \frac{\sum_{i=1}^n x_i \cdot y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2}$$

$$a = \bar{y} - b \cdot \bar{x}$$

Где  $a$  и  $b$  – коэффициенты линейной функции,

$x_i$  – значение времени работы скважины (дата замера минус дата начала анализируемого периода).

$y_i$  – значение Дебита жидкости скважины (с учетом коэффициента накопления, если он есть для данной скважины).

$n$  – число замеров (согласно ранее описанным алгоритмам).

С помощью коэффициентов  $a$  и  $b$  строится линия тренда.

Структура данных для построения линии тренда имеет следующий вид:

Где  $Y_{\text{тренда}}$  – значение зависимой переменной, соответствующее полученной линии тренда.

$\bar{y}$  - среднее значение дебита жидкости ( $Q_{\text{ж ср.}}$ )

2.5.2 Рассчитывается параметр  $\Delta Y_{\text{посл}}$  - отношение отклонения значения от линии тренда для последнего значения в месяце. Показывает на сколько последнее значение отклоняются от тренда.  $\Delta Y_{\text{посл}} = |Y_{\text{посл}} - Y_{\text{тренда}}| / Y_{\text{посл}}$ .

$Y_{\text{посл}}$  – последнее значение дебита нефти в анализируемом периоде,

$Y_{\text{тренда}}$  – значение зависимой переменной линии тренда, соответствующее по дате значению  $Y_{\text{посл}}$ .

2.5.3 Рассчитывается значение параметра  $R^2$  - величина достоверности аппроксимации.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

$$SS_{res} = \sum_{i=1}^n (y_i - Y_{i \text{ тренда}})^2$$

$$SS_{tot} = \sum_{i=1}^n (y_i - \bar{y})^2$$

2.5.4 Рассчитывается стандартное отклонение  $\delta$ .

$$\delta = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n - 1}}$$

2.5.5 Рассчитывается параметр  $\Delta\delta$  - отношение стандартного отклонения к среднему значению.

$$\Delta\delta = \frac{\delta}{Q_{ж\text{ ср.}}}$$

2.6 Результаты анализа и выбор алгоритма расчета.

- Если  $R^2 < R^2_{\text{граничное}}$  и  $\Delta\delta > \Delta\delta_{\text{граничное}}$ , то результатом расчета является **среднее** значение дебита жидкости за отчетный период. Формула расчета среднего значения дебита жидкости

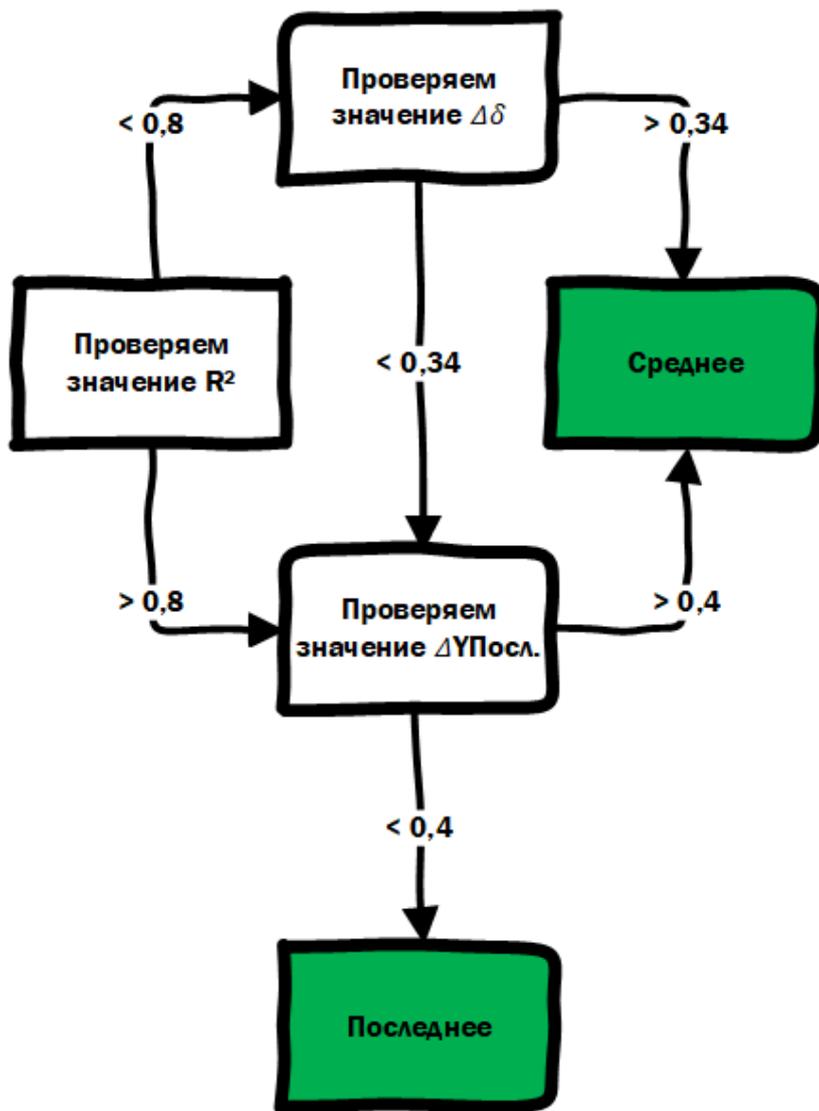
$$Q_{ж\text{ ср.}} = \frac{\sum_1^n Q_{ж}}{n},$$

Где  $Q_{ж\text{ ср.}}$  – среднее значение дебита жидкости (т/сут),

$Q_{ж}$  - значения замеров дебита жидкости (т/сут),

Среднее значение дебита жидкости является конечным результатом расчета.

- Если  $R^2 < R^2_{\text{граничное}}$  и  $\Delta\delta < \Delta\delta_{\text{граничное}}$ , то проверяем значение  $\Delta Y_{\text{посл}}$ :
  - Если  $\Delta Y_{\text{посл}} > \Delta Y_{\text{посл, граничное}}$ , то берем **среднее** значение. Расчет среднего значения описан выше. Среднее значение является конечным результатом. Значение передается в форму «Факт» технологического режима ШТР и для расчета прочих показателей.
  - Если  $\Delta Y_{\text{посл}} < \Delta Y_{\text{посл, граничное}}$ , то берем **последнее** значение. Последнее значение замера дебита жидкости является конечным результатом.
- Если  $R^2 > R^2_{\text{граничное}}$ , то проверяем значение  $\Delta Y_{\text{посл}}$ :
  - Если  $\Delta Y_{\text{посл}} > \Delta Y_{\text{посл, граничное}}$ , то берем **среднее** значение. Расчет среднего значения описан выше. Среднее значение является конечным результатом. Значение передается в форму «Факт» технологического режима ШТР и для расчета прочих показателей.
  - Если  $\Delta Y_{\text{посл}} < \Delta Y_{\text{посл, граничное}}$ , то берем **последнее** значение. Последнее значение замера дебита жидкости является конечным результатом. Значение передается в форму «Факт» технологического режима ШТР и для расчета прочих показателей.



Показатель	Граничное значение
$R^2$	0,8
$\Delta\delta$	0,34
$\Delta УПосл$	0,4

Рисунок 12. Результаты анализа и выбор алгоритма расчета.

## 2.7 Получение граничных коэффициентов

Для принятия решения, какое значение брать среднее или последнее не малую роль играют граничные коэффициенты. Для того чтобы их найти нам необходимо выполнить определенный алгоритм.

- Для каждой скважины считается достоверной аппроксимации( $R^2$ ), отношение стандартного отклонения к среднему( $\Delta\delta$ ) и отношению отклонения значения от линии тренда для последнего значения в месяце( $\Delta Y_{\text{посл}}$ ).
- Задаются случайные граничные коэффициенты
- По случайным коэффициентам определяем среднее или последнее значение  $Q_{\text{ж}}$
- Из полученных значений вычитаем фактические значения  $Q_{\text{ж}}$ .
- Суммируем значения по каждой скважине. Находим суммы для каждого варианта допустимых коридоров.
- Определяем минимальное значение суммы
- Берем полученные коэффициенты в работу.

## 2.8 Расчет остальных параметров на основании решения $Q_{\text{ж}}$

По следующим показателям всегда выбираем в расчете последнее значение: Диаметр штуцера, Диаметр штуцера для подачи газлифтного газа, забойное давление. Расчет параметров Обводненности, Расход газлифтного газа ( $q$  г/г), Дебит газа попутного ( $Q$  пг), Дебит газа на устье скважины с учетом газлифта, Давление на приеме, Динамический уровень, Буферное давление, Давление в линии, Затрубное давление и Давление газлифтного газа ( $P_{\text{г/г}}$ ) зависит от результата расчета статистического анализа по дебиту жидкости.



Рисунок 13. Расчет остальных параметров.

## 2.9 Алгоритм расчета средних значений параметров Буферное давление и Давление в линии.

В случае среднего значения параметров «Буферное давление» и «Давление в линии» дополнительно производится обработка данных выборки:

1. Замеры берутся с момента, когда диаметр штуцера постоянный
2. Исключаются замеры, когда скважина не работала.
3. Исключаются максимальные и минимальные значения по методу Диксона[1] (Приложение 2)

Схема расчета средних значений параметров представлена на Рисунке 14.

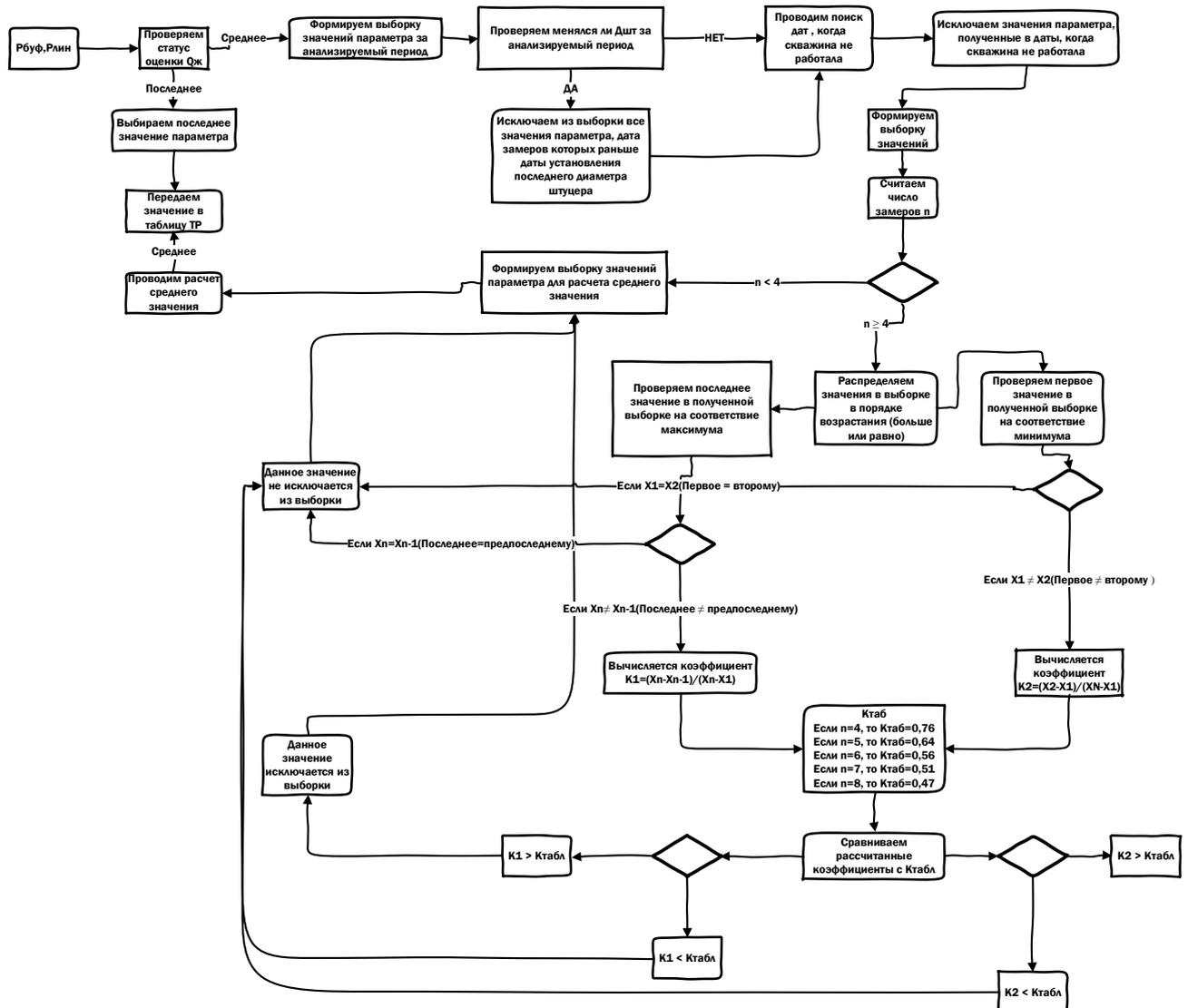


Рисунок 14. Алгоритм расчета средних значений параметров Буферное давление и Давление в линии.

## 2.10 Q-тест Диксона для выявления выбросов для небольших выборок

Q-тест Диксона [11] - удобный способ быстро идентифицировать выбросы в наборах данных, которые содержат только небольшое количество наблюдений: обычно  $3 \leq n \leq 10$ .

Мы используем этот тест для законного удаления (потенциальных) выбросов из набора данных.

- наши данные нормально распределены,
- мы не используем этот тест более одного раза для одного и того же набора данных.

Данный тест мы проводим для того чтобы отсечь выбросы по буферному давлению, давлению в линии, затрубному давлению и давлению газлифтного газа.

Сначала мы упорядочиваем данные для нашей выборки в порядке возрастания (от наименьшего к наибольшему значению):

$$X_1 < X_2 < \dots < X_n$$

Затем мы вычисляем экспериментальное значение Q ( $Q_{exp}$ ). Обратите внимание, что в более поздней статье в 1953 году Диксон и Дин [12] пересмотрели расчет Q-значения и сообщили о различных уравнениях для разных сценариев [13]:

$$r_{10} = \frac{x_n - x_{n-1}}{x_n - x_1}$$

$r_{10}$  для  $3 \leq n \leq 7$

$r_{11}$  для  $8 \leq n \leq 10$

$r_{21}$  для  $11 \leq n \leq 13$

$r_{22}$  для  $n \leq 14$

Однако, согласно утверждению (наблюдению) в более поздней статье [13]: «Коэффициент  $r_{10}$  обычно обозначается как «Q» и обычно считается наиболее удобным, законным, статистическим тестом. Так же он доступен для отклонения девиантных значений из небольшой выборки, соответствующей распределению Гаусса. (Он одинаково хорошо подходит для больших наборов данных, если присутствует только один выброс)».

Поэтому я буду использовать  $r_{10}$  для следующей реализации Q-теста Диксона:

$$Q_{exp} = \frac{x_2 - x_1}{x_n - x_1}$$

где предполагается, что данные расположены в порядке возрастания:  
 $x_1 < x_2 < \dots < x_n$

Сравниваем рассчитанные  $Q_{exp}$  к табличному критическому значению  $Q_{crit}$  (Таблица 1).

На следующем этапе мы сравним рассчитанные  $Q_{exp}$  значение к табличному критическому значению  $Q_{crit}$  для выбранного доверительного интервала.

Если рассчитанное значение  $Q$  для конкретного наблюдения больше критического значения  $Q$  ( $Q_{exp} > Q_{crit}$ ), это наблюдение считается выбросом в соответствии с Q-тестом.

N	Q90%	Q95%	Q99%
3	0,941	0,97	0,994
4	0,765	0,829	0,926
5	0,642	0,71	0,821
6	0,56	0,625	0,74
7	0,507	0,568	0,68
8	0,468	0,526	0,634
9	0,437	0,493	0,598
10	0,412	0,466	0,568
11	0,392	0,444	0,542
12	0,376	0,426	0,522
13	0,361	0,41	0,503
14	0,349	0,396	0,488
15	0,338	0,384	0,475
16	0,329	0,374	0,463
17	0,32	0,365	0,452
18	0,313	0,356	0,442
19	0,306	0,349	0,433
20	0,3	0,342	0,425
21	0,295	0,337	0,418
22	0,29	0,331	0,411
23	0,285	0,326	0,404
24	0,281	0,321	0,399
25	0,277	0,317	0,393
26	0,273	0,312	0,388
27	0,269	0,308	0,384
28	0,266	0,305	0,38
29	0,263	0,301	0,376
30	0,26	0,29	0,372

Таблица 1. значения двустороннего Q-теста Диксона.

Мы в своей работе используем первую достоверность составляющую 90%.

## 2.11 Алгоритм расчета среднего значения параметра затрубное давление ( $P_{затр}$ )

При расчете среднего значения параметра затрубное давление из анализируемой выборки исключаются значения параметра, полученные, когда скважина не работала, и максимальные и минимальные значения в соответствии с методом Диксона (Приложение 2).

Схема расчета среднего и последнего значений параметра затрубное давление представлена на Рисунке 15.

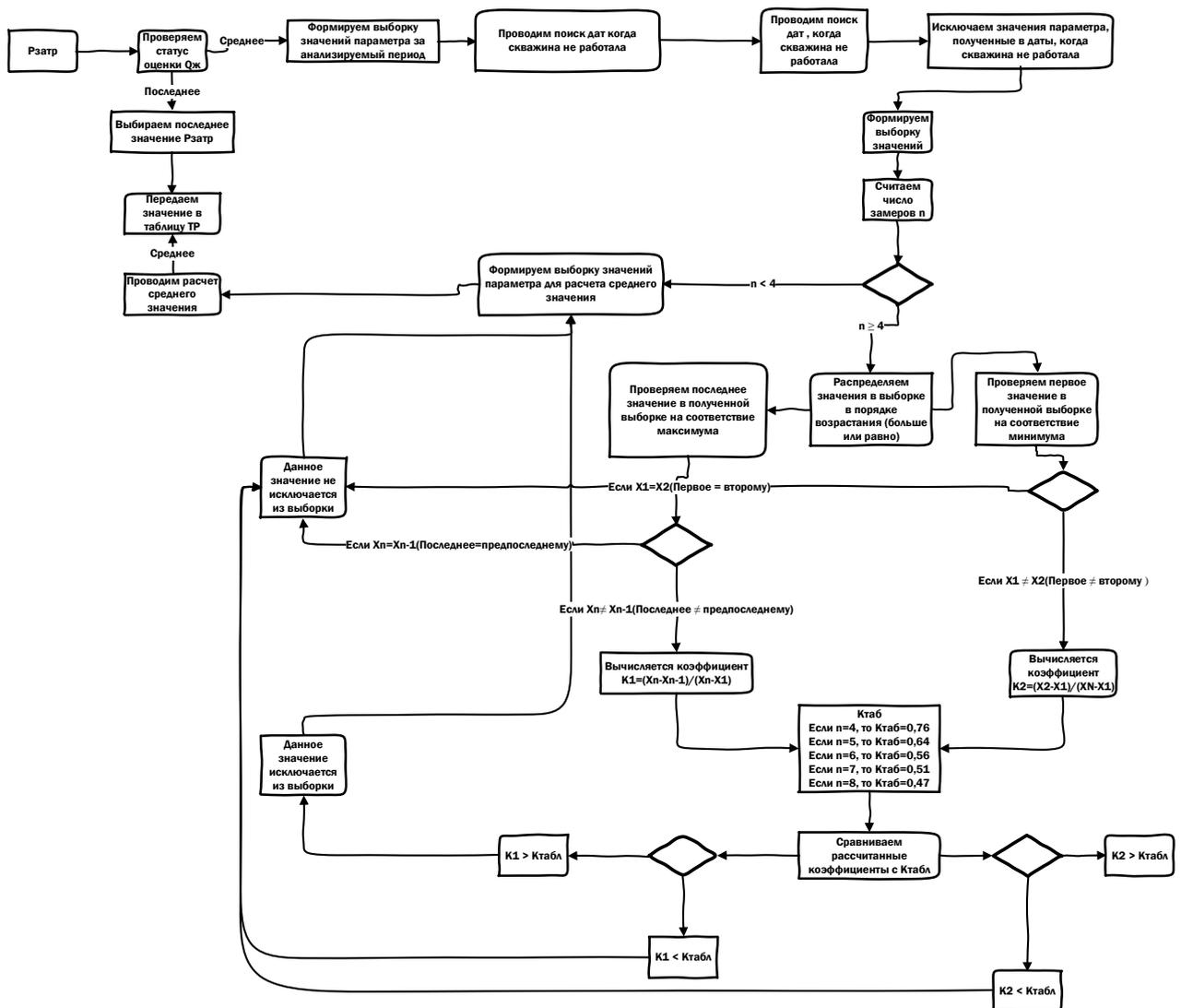


Рисунок 15. Алгоритм расчета среднего значения параметра затрубное давление.

## 2.12. Алгоритм расчета средних значений параметров Давление газлифтного газа (Рг/г).

Для параметра Давление газлифтного газа (Рг/г) значения берутся с момента, когда Диаметр штуцера для подачи газлифтного газа постоянный, также исключаются максимальные и минимальные значения и значения, когда скважина не работала.

Схема расчета средних значений параметров представлена на Рисунок 16.

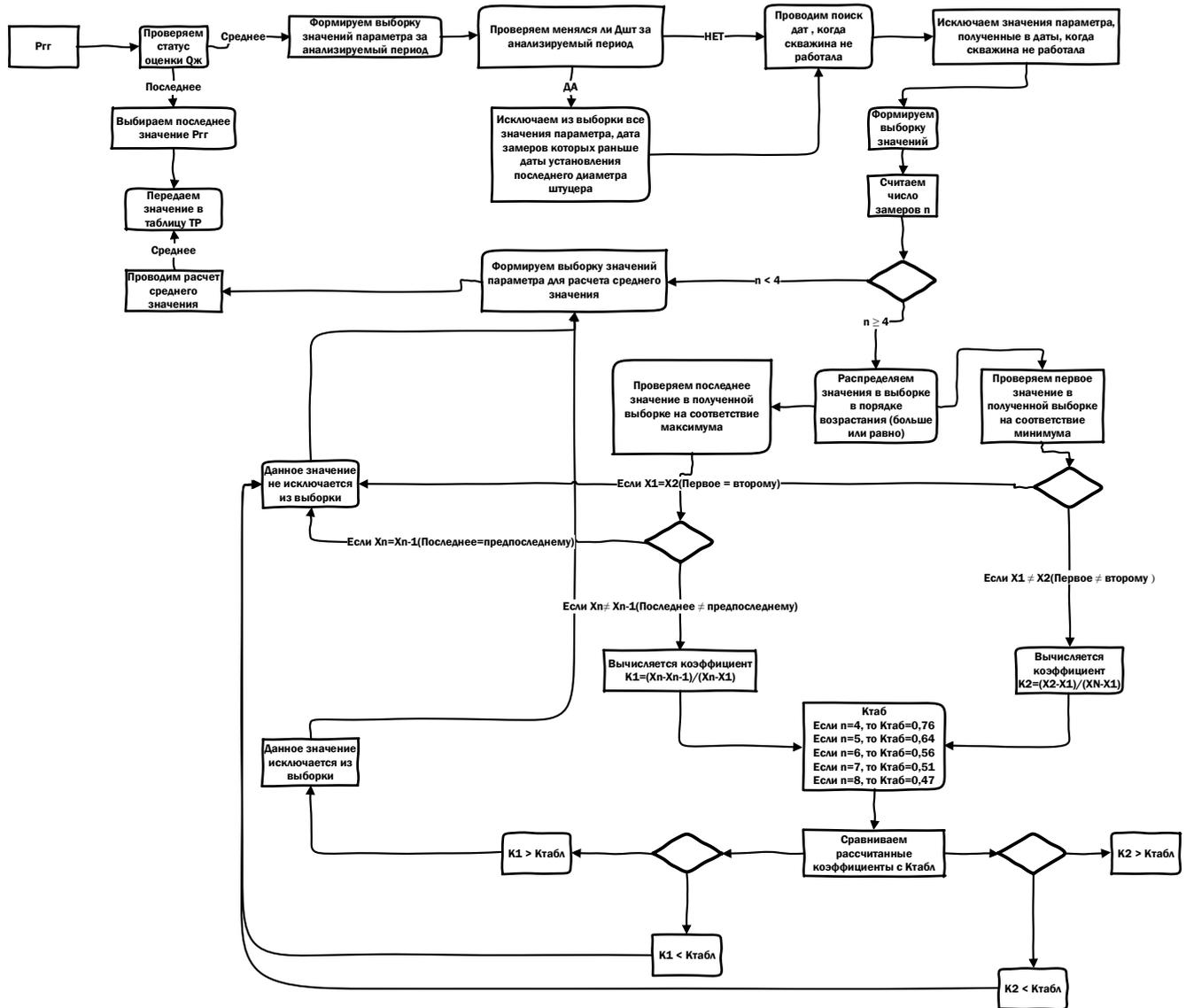


Рисунок 16. Алгоритм расчета средних значений параметров Давление газлифтного газа (Рг/г).

Модуль Автоматизации ТР работает по принципу, приведенному на Рисунке 17. С описанием компонентов схемы информационных потоков можно в Таблице 2.

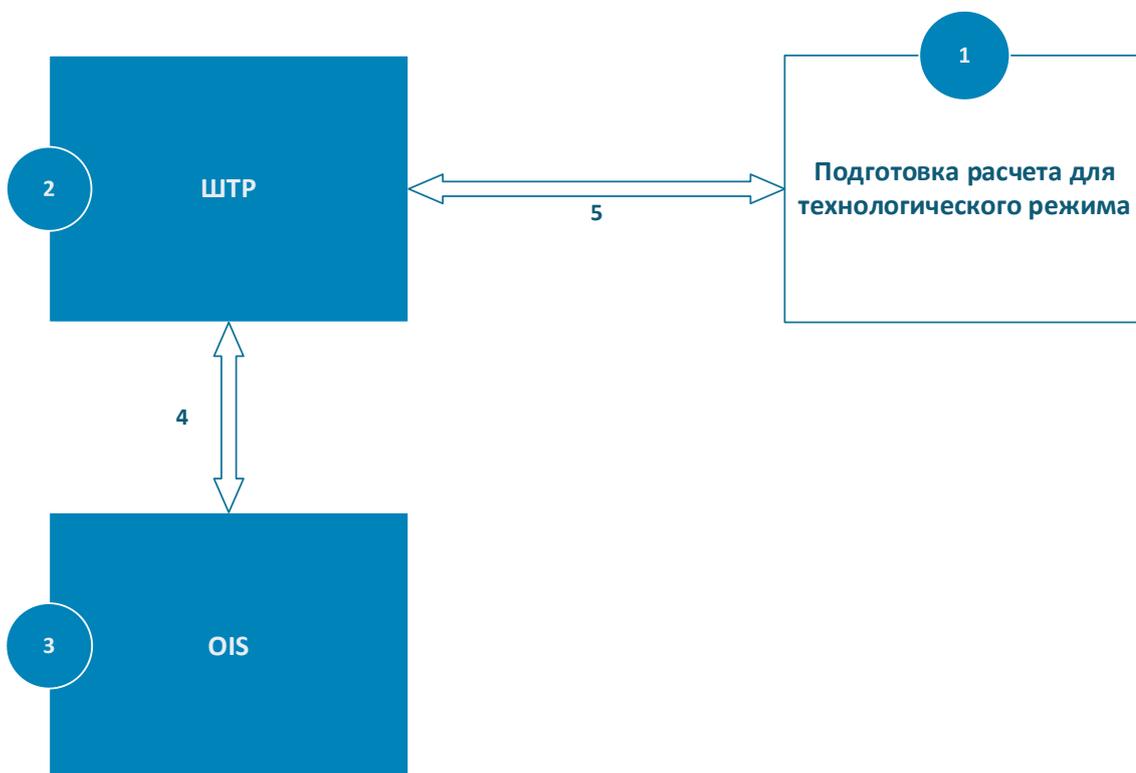


Рисунок 17. Схема информационных потоков модуля «Подготовка расчета для технологического режима»

№ п/п	Название компонента	Описание компонента
1	Модуль «Подготовка расчета для технологического режима»	Модуль, предназначенный для подготовки параметров и определения методики расчета технологического режима.
2	ШТР	Информационная система «Шахматка и Техрежим». Предназначена для сбора промышленной информации
3	ОИС	Информационная система OIS Production Предназначена для сбора и обработки нефтепромысловой информации
4	Информационный поток ОИС - ШТР	Двухсторонний информационный поток между ШТР и ОИС. Передаваемые данные: показатели скважин, виды эксплуатации и др.(Разработка потока не требуется)
5	Информационный поток ШТР - Модуль «Подготовка расчета для технологического режима»	Двухсторонний информационный поток между базой данных ШТР и модулем «Подготовка расчета для технологического режима».

	режима»	<p>Передаваемые данные из ШТР: данные по работе и показатели скважин (вид эксплуатации, режим работы, замеры показателей, даты замеров).</p> <p>Передаваемые данные в ШТР: подготовленные параметры (Дебит жидкости, Обводненность, Расход газлифтного газа (<math>q</math> г/г), Дебит газа попутного (<math>Q</math> пг), Дебит газа на устье скважины с учетом газлифта, Давление на приеме, Динамический уровень, Буферное давление, Давление в линии, Давление газлифтного газа (<math>P_{г/г}</math>), Забойное давление по исследованию, Диаметр штуцера, Диаметр штуцера для подачи газлифтного газа).</p>
--	---------	--

Таблица 2. Описание компонентов схемы информационных потоков

#### 2.14 Блок-схема прототипа автоматического принятия решения по способу расчета ТР на основании методов статистического анализа.

При создании прототипа был выбран язык программирования python из-за возможности быстрого обучения и возможности использовать специализированные библиотеки, которые позволяют работать с большим объемом данных и быстро проводить оптимизацию данных. Так же python позволяет быстро визуализировать таблицы и выводить данные в удобный формат csv, xls.

В работе мы использовали библиотеки numpy и pandas. Данные библиотеки позволяют просчитывать большие массивы данных и строить таблицы для их визуального отображения. Так же в работе мы использовали библиотеку plotly для создания наглядной анимации данных.

В создании прототипа мы следовали определенным этапам (Рисунок 18).

**OIS(oil information system)-**  
**официальная**  
**база данных**  
**по добычи**  
**нефти по**  
**предприятию**  
**TR-Данные по**  
**состоянию**  
**скважин**

**Мероп-данные**  
**мероприятий**  
**по скважинам**

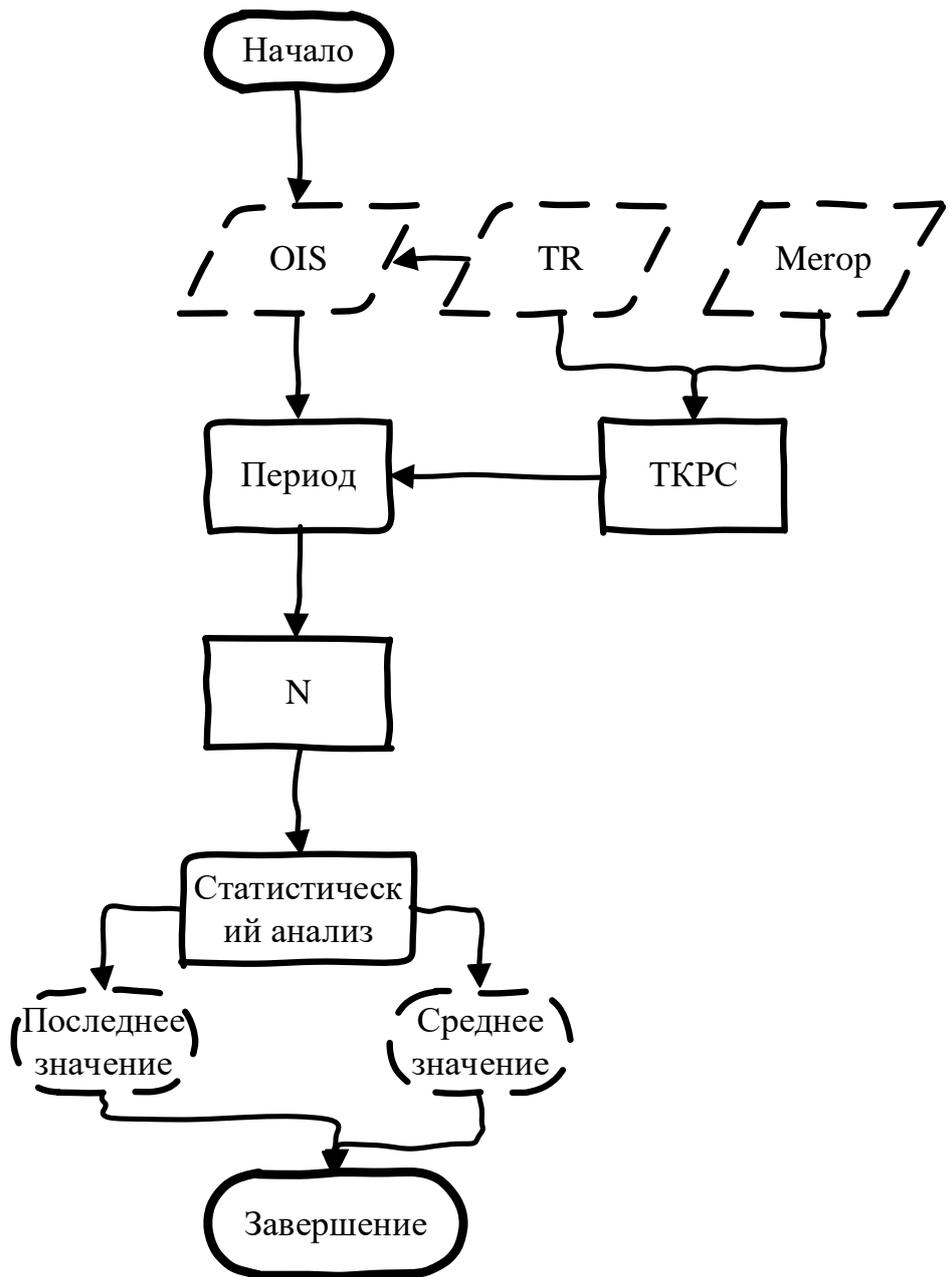


Рисунок 18. Блок-схема прототипа автоматического принятия решения по способу расчета TR на основании методов статистического анализа.

Вначале необходимо выгрузить все необходимые базы данных (Рисунок 19).

```
In [1]: 1 #подгружаем библиотеки
2 import numpy as np
3 import pandas as pd
4 from datetime import datetime
5 from datetime import timedelta
6 import math
7 import time
8 from itertools import groupby
9 from operator import itemgetter

In [2]: 1 #доп. библиотеки
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4 sns.set_style('whitegrid')
5 %matplotlib inline

In [3]: 1 #загружаем датафрей, определяем кол-во столбцов с которыми будем работать.
2 pathin = 'C:/Users/onvt1/OneDrive/Рабочий стол/Газпром-нефть/Материалы по проекту Автоматизация ТР/состояние_оренбург_кварт
3 pathtr = pathin + 'ORB_exel.xlsx'
4 pathtr1 = pathin + 'мероп_akbkb2w3.csv'
5 #мероприятия
6 pathtr_2 = pathin + 'tr_m2v9dk4f.csv'
7 #зачищенные скважины
8 pathtr_3 = pathin + 'b2.xlsx'
9 #файл посчитанного тр Оренбург
10 file = 'C:/Users/onvt1/OneDrive/Рабочий стол/тех.режим_Оренбурга/1.xlsx'
11 file1 = 'C:/Users/onvt1/OneDrive/Рабочий стол/тех.режим_Оренбурга/2.xlsx'

In [4]: 1 x1 = pd.ExcelFile(pathtr)
2 print(x1.sheet_names)
3 mass_data = x1.parse('Лист1')
4
5 # Сбор данных
6 #mass_data = np.genfromtxt(pathtr, usecols = (1, 2, 3, 4, 14, 7,12, 13, 15, 5, 6, 9, 11),delimiter=';', dtype=None, skip_he
7 мероп_data = np.genfromtxt(pathtr1, usecols = (0,2,3), delimiter=';', dtype=None, skip_header=True, encoding=None)
8 df_data=pd.DataFrame(data=mass_data)
9 #field'
10 df_data.columns=['0','field','well','date','Q', 'W','Qpng','Qgg','Pzab_isl', 'Pzab_Pr', 'Ppr','Pzab_Hdin', 'Hdin', 'Pbuf', '
11 #df_data = df_data.drop(0)
12 df_data.replace('', np.nan, inplace=True)
13
14 x11 = pd.ExcelFile(pathtr_3)
15 print(x11.sheet_names)
16 tr_clean = x11.parse('Лист1')
17
18 x12 = pd.ExcelFile(file)
19 print(x12.sheet_names)
20 TR_Or1 = x12.parse('Лист1', usecols = (4,9,15,78,79,80,81,82,85,75,87,88,89,90,91,92,93,95,96,97), header=1)
21 TR_Or1 = TR_Or1.iloc[2:]
22
23 x13 = pd.ExcelFile(file1)
24 print(x13.sheet_names)
25 TR_Or2 = x13.parse('Лист1', usecols = (5, 9, 14, 71, 72, 77, 78, 79, 82, 68, 84, 87, 88, 89, 90, 91, 93, 95), header=1)
26 TR_Or2 = TR_Or2.iloc[2:]

['Лист1']
['Лист1']
['Лист1']
['Лист1']
```

Рисунок 19. Загрузка библиотек и баз данных.

1. OIS(oil information system)- официальная база данных по добычи нефти по предприятию.
2. База данных по ремонту скважин
3. База данных по проведенным мероприятиям и статусу скважин на месторождениях.

После того как все данные были выгружены нами была проведена чистка от нежелательных знаков и пропусков. Так как все таблицы с данными имели разный формат и вид, необходимо было отфильтровать и грамотно отобразить данные (Рисунок 20).

```
In [10]: 1 df_data['well']=df_data['well'].str.replace(' ', '')
2 df_data['well']=df_data['well'].str.replace('а', '')
3 df_data['well']=df_data['well'].str.replace('р', '')
4 df_data['well']=df_data['well'].str.replace('Р', '')
5 df_data['well']=df_data['well'].str.replace('кл', '')
6 df_data['well']=df_data['well'].str.replace('Д', '')
7 df_data['well']=df_data['well'].str.replace('Тг', '')
8 df_data['well']=df_data['well'].str.replace('сл', '')
```

```
In [11]: 1 tr['well']=tr['well'].str.replace('G5', '')
2 tr['well']=tr['well'].str.replace('G4', '')
3 tr['well']=tr['well'].str.replace('G3', '')
4 tr['well']=tr['well'].str.replace('G2', '')
5 tr['well']=tr['well'].str.replace('G', '')
6 tr['well']=tr['well'].str.replace('A', '')
7 tr['well']=tr['well'].str.replace('R', '')
8 tr['well']=tr['well'].str.replace('P', '')
9 tr['well']=tr['well'].str.replace('D', '')
10 tr['well']=tr['well'].str.replace('ST2', '')
11 tr['well']=tr['well'].str.replace('O', '')
12 tr['well']=tr['well'].str.replace('T', '')
13 tr['well']=tr['well'].str.replace('SLO', '')
14 tr['well']=tr['well'].str.replace('SL', '')
```

```
In [12]: 1 merop['well']=merop['well'].str.replace('G5', '')
2 merop['well']=merop['well'].str.replace('G4', '')
3 merop['well']=merop['well'].str.replace('G3', '')
4 merop['well']=merop['well'].str.replace('G2', '')
5 merop['well']=merop['well'].str.replace('G', '')
6 merop['well']=merop['well'].str.replace('A', '')
7 merop['well']=merop['well'].str.replace('R', '')
8 merop['well']=merop['well'].str.replace('P', '')
9 merop['well']=merop['well'].str.replace('D', '')
10 merop['well']=merop['well'].str.replace('ST2', '')
11 merop['well']=merop['well'].str.replace('O', '')
12 merop['well']=merop['well'].str.replace('T', '')
13 merop['well']=merop['well'].str.replace('SLO', '')
14 merop['well']=merop['well'].str.replace('SL', '')
15 merop['well'] = merop['well'].map(lambda x: str(x)[4:])
```

Рисунок 20. Чистка от нежелательных пропусков и символов, вводящих в заблуждение.

Далее данные по скважинам отфильтровываются на возможность ремонта или других мероприятий плановых или неплановых.

Следующий этап заключался в создании периодов с количеством замеров. В нашей методике замеры берутся не в один период, а в несколько из-за этого нам и потребовался дополнительный блок в нашей программе (Рисунок 21).

```

In [16]: 1 # Определяем период для расчета
2 merop_m=merop[(merop['St'] > date_st_m) & (merop['St'] < date_end)]
3 merop_q=merop[(merop['St'] > date_st_q) & (merop['St'] < date_end)]
4
5 # массив с месячными данными
6 df_Q_m=df_data[['well', 'field', 'date', 'Q']].copy().dropna()
7 df_Q_m=df_Q_m[(df_Q_m['date'] > date_st_m) & (df_Q_m['date'] < date_end)]
8 Coll_mass = df_Q_m.groupby(['well', 'field'])[['Q']].count()
9 Coll_mass=Coll_mass.reset_index()
10 Coll_mass.columns=['well', 'field', 'n']
11
12 Coll_mass= Coll_mass.merge(merop_m, how='left')
13 Coll_mass=Coll_mass.drop('date', axis=1)
14 Coll_mass.columns=['well', 'field', 'n', 'merop']
15 Coll_mass['merop_bl'] = False
16 Coll_mass['merop_bl'][Coll_mass['merop'] < date_end]=True
17 Coll_mass['min_n'] = True
18 Coll_mass['min_n'][(Coll_mass['merop_bl'] == False) & (Coll_mass['n'] < 3) ]=False
19 Coll_mass['Stp'] = date_st_m
20 Coll_mass['Stp'][Coll_mass['merop_bl'] == True] = Coll_mass['merop']
21
22
23 # массив с квартальными данными
24 df_Q_q=df_data[['well', 'field', 'date', 'Q']].copy().dropna()
25 df_Q_q=df_Q_q[(df_Q_q['date'] > date_st_q) & (df_Q_q['date'] < date_end)]
26 Coll_mass_q = df_Q_q.groupby(['well', 'field'])[['Q']].count()
27
28 Coll_mass_q=Coll_mass_q.reset_index()
29 Coll_mass_q.columns=['well', 'field', 'nq']
30 Coll_mass_q= Coll_mass_q.merge(merop_q, how='left')
31 Coll_mass_q=Coll_mass_q.drop('date', axis=1)
32 Coll_mass_q.columns=['well', 'field', 'nq', 'meropq']
33 Coll_mass_q['merop_blq'] = False
34 Coll_mass_q['merop_blq'][Coll_mass_q['meropq'] < date_end]=True
35 Coll_mass_q['stq'] = date_st_q
36 Coll_mass_q['stq'][Coll_mass_q['merop_blq'] == True] = Coll_mass_q['meropq']
37 Coll_mass_q = Coll_mass_q[['well', 'nq', 'stq']]
38
39 Coll_mass = Coll_mass.merge(Coll_mass_q, how='left')
40
41 # если мало значений то берм кварталные данные
42 Coll_mass['Stp'][Coll_mass['min_n']==False] = Coll_mass['stq']
43 Coll_mass['n'][Coll_mass['min_n']==False] = Coll_mass['nq']

```

Рисунок 21. Определение периода расчета.

После того как все отфильтровано, посчитано и определена периодичность мы переходим к проведению статистического анализа (Рисунок 22).

```

In [17]: 1 # Фильтрация массива данных по вычисленному периоду
2 #1. Из Coll_mass деберем St (начала расчета)
3 df_st=Coll_mass[['well', 'field', 'Stp']]
4
5 #2. Фильтр df_data на st по отдельным скважинам в общем массиве
6 df_data = df_data.merge(df_st, how='left')
7 df_data = df_data [(df_data ['Stp']<=df_data ['date'])& (date_end >=df_data ['date'])]

In [18]: 1 #3. Выделения максима Q и проведение стат анализа
2 df_Q=df_data[['well', 'field', 'indate', 'date', 'Q']].dropna()
3 df_Q = df_Q.groupby(['field', 'well', 'date'])[['indate', 'Q']].last()
4 df_Q = df_Q.reset_index()

In [19]: 1 #округление. Нужно для того чтобы избежать ошибки(Length mismatch: Expected axis has 2 elements,
2 df_Q = df_Q.apply(pd.to_numeric, errors='ignore')
3 df_Q['Q'] = np.round(df_Q['Q'], 5)
4 # определение коэффициентов a и b
5 df_Q['XY'] = df_Q['Q']* df_Q['indate']
6 df_Q['X2'] = df_Q['indate']* df_Q['indate']
7
8 stat1 = df_Q.groupby(['well', 'field'])[['Q', 'indate']].mean()
9 nstat = df_Q.groupby(['well', 'field'])[['Q']].count()
10 sumstat = df_Q.groupby(['well', 'field'])[['X2', 'XY']].sum()
11 stat1 = stat1.reset_index()
12 nstat = nstat.reset_index()
13 sumstat= sumstat.reset_index()
14 stat1.columns=['well', 'field', 'meanQ', 'meanD']
15 nstat.columns=['well', 'field', 'n']
16 sumstat.columns=['well', 'field', 'sumX2', 'sumXY']
17 stat1 = stat1.merge(nstat)
18 stat1 = stat1.merge(sumstat)
19 stat1['a1'] = stat1['sumXY'] - stat1['meanQ']*stat1['meanD']*stat1['n']
20 stat1['a2'] = stat1['sumX2'] - stat1['meanD']*stat1['meanD']*stat1['n']
21 stat1['a'] = stat1['a1']/stat1['a2']
22 stat1['b'] = stat1['meanQ'] - stat1['meanD'] *stat1['a']
23 statout= stat1[['well', 'field', 'meanQ', 'a', 'b']]
24
25 # определение дисперсии
26 df_Q = df_Q.merge(statout)
27 df_Q['dfun']=df_Q['a']*df_Q['indate'] + df_Q['b']
28 df_Q['disf'] = (df_Q['Q'] -df_Q['dfun'])
29 df_Q['disf']=df_Q['disf']*df_Q['disf']
30 df_Q['dist']=(df_Q['Q'] -df_Q['meanQ'])
31 df_Q['dist']=df_Q['dist']*df_Q['dist']
32
33 maxdat = df_Q.groupby(['well', 'field'])[['indate']].max()
34 maxdat= maxdat.reset_index()
35 maxdat.columns=['well', 'field', 'maxindate']
36 liqmaxdis = df_Q.merge(maxdat)
37 liqmaxdis = liqmaxdis[liqmaxdis['indate']==liqmaxdis['maxindate']] [['well', 'field', 'Q', 'disf']]
38 liqmaxdis.columns=['well', 'field', 'maxliq', 'maxindis']
39
40 # расчет R2
41 sumdis = df_Q.groupby(['well', 'field'])[['disf', 'dist']].sum()
42 sumdis= sumdis.reset_index()
43 sumdis['R2'] = 1 - sumdis['disf']/(sumdis['dist'])
44 # Вывод данных
45 stat1 = stat1.merge(sumdis)
46 stat1['dist'] =stat1['dist'] /(stat1['n']-1)
47 stat1 = stat1.merge(liqmaxdis)
48 statout= stat1[['well', 'field', 'n', 'meanQ', 'dist', 'a', 'b', 'R2', 'maxliq', 'maxindis']]

```

Рисунок 22. Проведение статистического анализа.

В заключительной части прототипа мы проводим определения последнего или среднего значения исходя из количества замеров, работы скважины и возможности определения данных которые мы считаем исходя из дебита жидкости (Рисунок 23).

Out[22]:

	well	field	n	meanQ	dist	a	b	R2	maxliq	maxindis	sig	siglast	solve	outliq
0	1	Новосамарское	13	50.335769	0.439597	-0.035129	1598.720439	0.267698	50.937	1.212525e+00	0.013172	2.161783e-02	last	50.937
1	1	Оренбургское	6	8.386667	3.775027	0.182740	-8046.689087	0.452030	10.606	2.335846e-01	0.231670	4.556911e-02	last	10.606
2	1	Царичанское+Филатовское	7	17.856429	0.177655	0.009434	-397.956052	0.049237	18.560	3.327999e-01	0.023604	3.108233e-02	last	18.560
3	10	Калитоновское	7	139.620000	33.550400	0.327556	-14297.347119	0.136903	143.073	1.706946e-02	0.041486	9.131712e-04	last	143.073
4	10-1	Оренбургское	11	11.837364	7.773245	-0.059096	2614.884644	0.205894	9.699	4.554909e-04	0.235530	2.200457e-03	last	9.699
..	...	...	...	...	...	...	...	...	...	...	...	...	...	...
0	80	Новозаринское	4	33.451000	37.222791	-0.372553	16455.134866	0.597539	25.728	1.247344e+01	0.182388	1.372736e-01	last	25.728
1	82	Новозаринское	7	27.732714	72.120316	2.380589	-104871.216983	0.568768	46.861	6.249708e+01	0.306222	1.687013e-01	last	46.861
2	86	Оренбургское	2	4.903000	0.506018	-0.143714	6339.038286	1.000000	4.400	1.326075e-25	0.145085	8.276208e-14	last	4.400
3	95-1	Оренбургское	7	22.971143	31.922484	-0.120061	5315.120716	0.053433	21.155	6.520645e-02	0.245961	1.207069e-02	last	21.155
4	95-2	Оренбургское	5	18.834800	7.843362	0.280197	-12330.791694	0.632618	20.764	8.626868e-01	0.148693	4.473172e-02	last	20.764

Рисунок 23. Получение среднего, последнего при проведении статистического анализа.

Остальные данные просчитываются по отдельным методикам, которые были описаны выше(п7,8,9,10,11). Полученная таблица данных выгружается в удобный для оператора формат: csv, xls. Пример в Приложении 3.

## 2.15 Реализация блок автоматического принятия решения по способу расчета ТР на основании методов статистического анализа.

В данный момент модуль реализован на тестовом сервере компании, является частью Информационная система «Шахматка и Техрежим» и проходит коррекцию и тестирование (Рисунок 24).

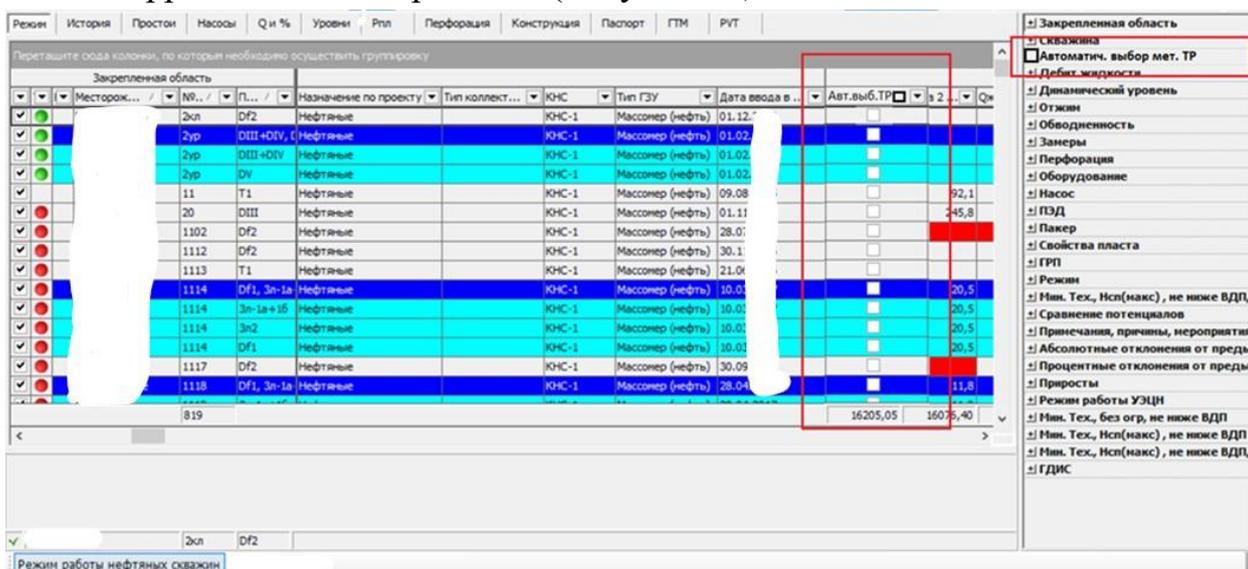


Рисунок 24. Рабочее пространство информационной системы «Шахматка».

Помимо функции автоматического расчета ТР есть и функция отключения данного режима (Рисунок 25).

Мероприятия		Закрепленные св-ва пласта				Ограничения Рзаб	
Скважина	Оборудование	Пласт	Факт	Потенциал	Резерв		
<input type="checkbox"/> Автоматический выбор методики расчета Техрежима							
Замеры							
Над, м	За 1 нед.	За 2 нед.	За 3 нед.	За 4 нед.	Способ оценки	Итого	
					Последнее знач		
Qж, м3/сут	За 1 нед.	За 2 нед.	За 3 нед.	За 4 нед.	Способ оценки	Итого	
					Последнее знач		
Обвод-ть	За 1 нед.	За 2 нед.	За 3 нед.	За 4 нед.	Способ оценки	Итого	
					Последнее знач		
Рбуф	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Рлен	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/> Руч. корр. Рбуф					<input type="checkbox"/> Руч. корр. Рлен	
Рпл	173,24	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Р на приеме	<input type="checkbox"/>	
	<input type="checkbox"/> Руч. корр. Рпл						
Рзатр	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Рзаб зачер.	<input type="checkbox"/>	
	<input type="checkbox"/> Руч. корр. Рзатр						
Нстат	22,5	КВЧ	215,9		Р заб расчетное	<input type="checkbox"/>	
Qж уплот.	<input type="checkbox"/>	Qж, т/сут	<input type="checkbox"/>		Дата замера Рпл	01.04.2020	
Дебит ПНГ	<input type="checkbox"/>	Ргг	<input type="checkbox"/>		Депрессия, атн	<input type="checkbox"/>	
Дебит общ газа	<input type="checkbox"/>						
Добыча							
Время работы, час	744	Добыча жидкости	<input type="checkbox"/>	Добыча нефти (т)	<input type="checkbox"/>		

Рисунок 25. Окно ввода начальных данных.

В процессе работы была изменена форма заполнения, добавлен параметр диаметр штуцера для подачи газлифтного газа, дебит газа попутного ( $Q_{пг}$ ), дебит газа на устье скважины с учетом газлифта, давление газлифтного газа ( $P_{г/г}$ ), расход газлифтного газа ( $q_{г/г}$ ). необходимый для последующего расчета ТР, а также опцию ручной корректировки данного параметра (Рисунок 26).

Режим История Простои Насосы Q и % Уровни Рпл Перфорация Конструкция Паспорт ГТМ РVT

Перетащите сюда колонки, по которым необходимо осуществить группировку

Закрепленная область				Рт/л	Руч. корр. Ргг	Деб.	Руч. корр. Деб	Деб. П	Руч. корр. Деб	Расх	Руч. корр. Расх
2кл	DF2		20	299		69		68		68	
2ур	ОПШ+ОПУ, с	1	346,6		4193,984			133		133	
2ур	ОПШ+ОПУ	1	335,7		4193,984			266		266	
2ур	Dv	1	327,5		4193,984			266		266	
11	T1	13	250		3561,078			60		60	
20	ОПШ	25	320		17338,061			81		81	
1102	DF2	0	306		591,19						
1112	DF2	10	306		110,644			61		61	
1113	T1	12	306		563,686			59		59	
1114	DF1, 3л-1а	14	289,6		939,591			62		62	
1114	3л-1а+1б	14	285		939,591			308		308	
1114	3л2	14	285		939,591			308		308	
1114	DF1	14	291		939,591			103		103	
1117	DF2	12	306		270,95			9		9	
1118	DF1, 3л-1а	12	288,6		438,822			61		61	
				819							

Режим работы нефтяных скважин

**Закрепленная область**

- Скважина
- Дебит жидкости
- Динамический уровень
- Отжим
- Обводненность
- Замеры
  - Рпл, атн 268,98
  - Дата замера Рпл 01.07.2020
  - Рпл на ВДП, атн 290,7
  - Дата замера Рпл н 01.07.2020
  - Рзаб (замер), атн
  - Рбуф, атн 19
  - Руч. корр. Рбуф
  - Рлен, атн 12
  - Руч. корр. Рлен
  - Рзатр, атн 20
  - Руч. корр. Рзатр
  - Рприем, атн 40
  - Нстат, н
  - Рзатр при Нстат, н
  - Рпл нач. 299
  - Q пл 69
- Ргг
  - Руч. корр. Ргг
  - Дебит ПНГ
  - Руч. корр. деб. пнг
  - Дебит общ. газа
  - Руч. корр. деб. о.г.
  - Расход акт. газа
  - Руч. корр. Расх.

Мероприятия | Закрепленные св-ва пласта | Ограничения Рзаб

Скважина | Оборудование | Пласт | Факт | Потенциал | Режим

Автоматический выбор методики расчета Техрежима

Замеры

	За 1 нед.	За 2 нед.	За 3 нед.	За 4 нед.	Способ оценки	Итого
Нд, м					Последнее знач.	
Qж, м3/сут					Последнее знач.	
Обвод-ть					Последнее знач.	

Рбуф   Руч. корр. Рбуф Рлен   Руч. корр. Рлен

Рпл   Руч. корр. Рпл Р на приеме

Рзатр   Руч. корр. Рзатр Рзаб замер.

Нстат  КВЧ  Р заб расчетное

Qж уплот.  Qн, т/сут  Дата замера Рпл

Дебит ПНГ   Руч. корр. Депрессия, атн

Дебит общ. газа   Руч. корр. Расход акт. газа   Руч. корр. Ргг   Руч. корр.

Добыча

Время работы, час  Добыча жидкости  Добыча нефти (т)

OK Отмена

Рисунок 26.Дополнительные формы для ввода данных.

Была создана новая экранная форма для отображения статистического анализа значений замеров и граничных условий по статистическому анализу (Рисунок 27).

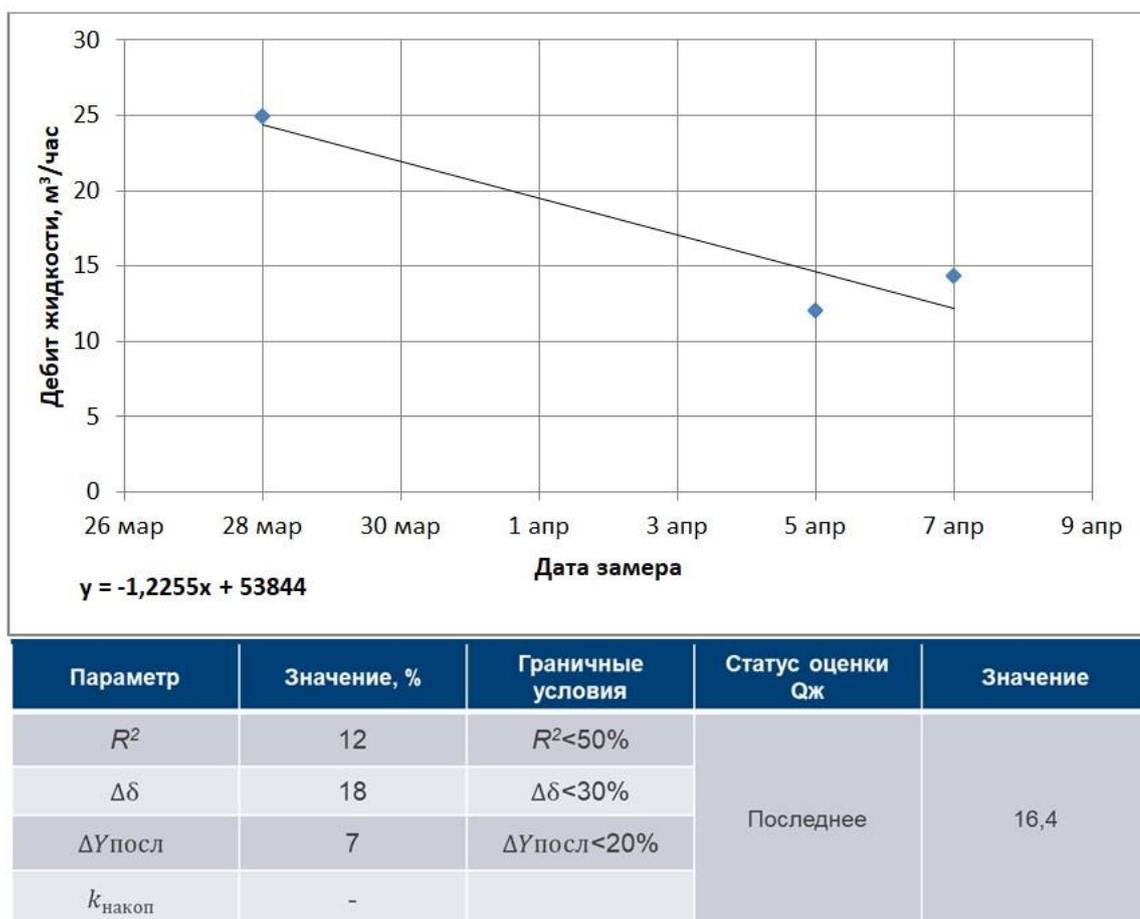


Рисунок 27. Форма отображения статистического анализа и граничных условий.

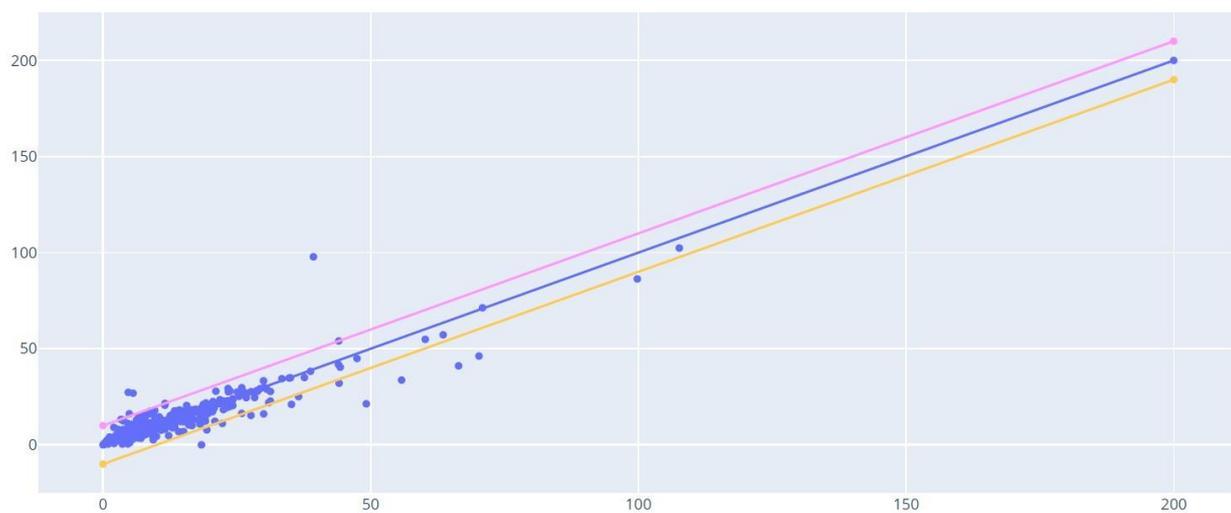
## 2.16 Практический результат

Был сформирован и стандартизирован алгоритм по автоматическому выбору метода расчета фактического ТР. Данный алгоритм был реализован и введен в промышленную эксплуатацию в информационной системе «Шахматка и Техрежим». Результат показал сходимость данных в 98,5% (Рисунок 28) с методом экспертной оценки.

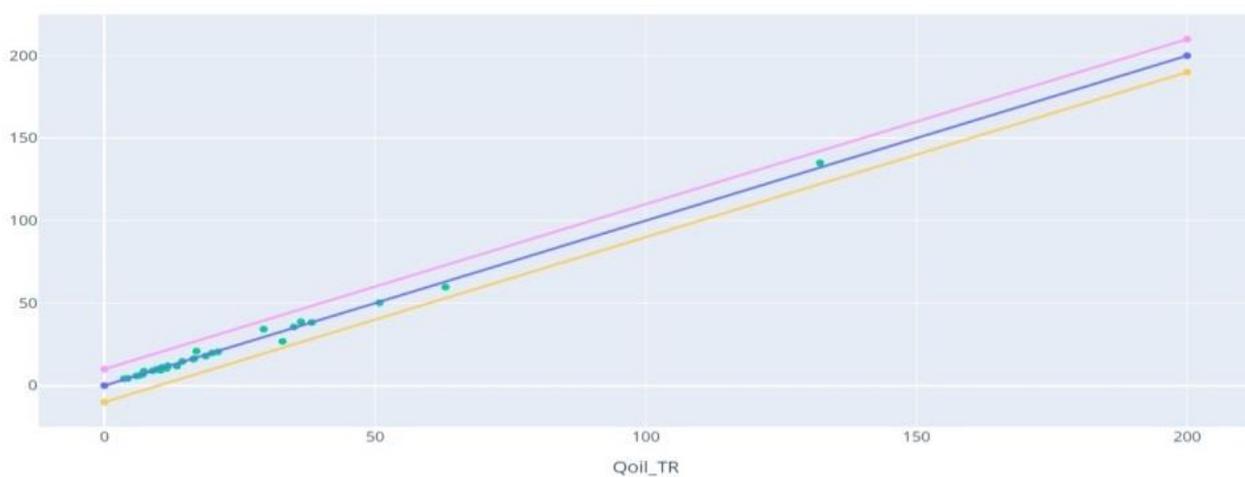
В результате проведенной работы повысили качество данных и ускорили процесс формирования ТР, т.к. отказались от ручного ввода и устранили ошибки, связанные с «человеческим фактором».

Методика получилась гибкой, позволяющая настраивать граничные значения для статистических коэффициентов и выбирать скважины для экспертной оценки.

Дебит нефти



Дебит нефти



Дебит нефти

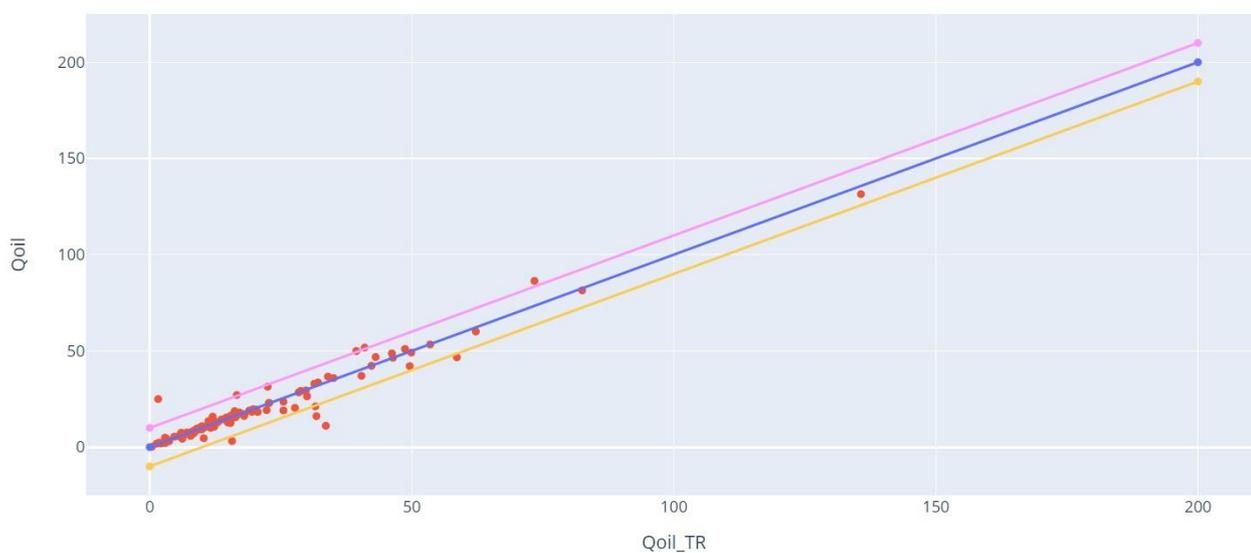


Рисунок 28. Значения дебит нефти полученные нашей методикой.

## **Заключение**

В ходе выполнения работы были выполнены следующие этапы:

- Изучение метода формирования технологического режима на добывающих объединениях.
- Изучение способов добычи нефти на месторождениях.
- Изучение цикла жизнедеятельности месторождения.
- Выявлена проблематика при формировании технологического режима.
- Изучена методика по автоматическому выбору способа расчета на основании статистического анализа.
- Создан прототип по автоматическому принятию решения для расчета технологического режима.
- Реализация методики автоматического принятия решения для расчета технологического режима на тестовом сервере компании.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

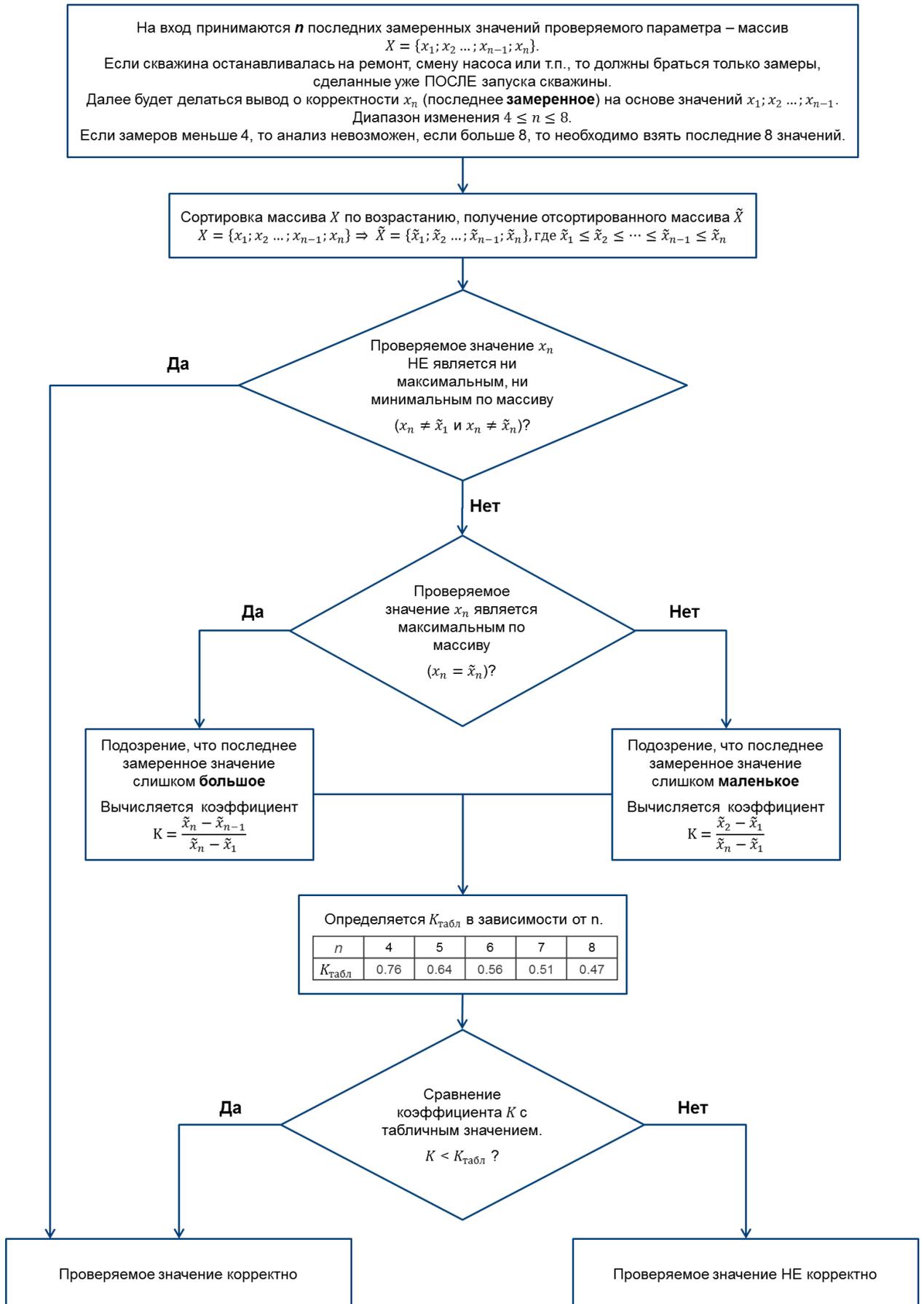
1. Диксон, JD (1981). «Асимптотически быстрая факторизация целых чисел» . Математика. Комп. 36 (153): 255–260. DOI : 10.1090 / S0025-5718-1981-0595059-1 . JSTOR 2007743 . Метод факторизации Диксона
2. Крец В.Г., Шадрина А.В. Основы нефтегазового дела. Учебное пособие. – Томск: Изд-во Томского политехнического университета, 2011. – 200 с
3. Куприенко, Николай Владимирович. Статистические методы изучения связей. Корреляционно-регрессионный анализ: учебное пособие / Н. В. Куприенко, О. А. Пономарева, Д. В. Тихонов; Санкт-Петербургский государственный политехнический университет. — Санкт-Петербург: Изд-во Политехн. ун-та, 2009 Санкт-Петербург 2020 — 1 файл (4,41 Мб)
4. Лоран, П.-Ж. Аппроксимация и оптимизация / П.-Ж. Лоран. - Москва: СИНТЕГ, 2004. - 496 с.
5. Рашка\_С\_, Мирджалили\_В\_Python\_и\_машинное\_обучение\_2020
6. Флетчер Хейслер, Дэвид Амос, Дэн Бейдер, Джоанна Яблонски. Основы Python. Авторские права © Real Python ( realpython.com ), 2012–2020 гг.
7. Хайдт М. Изучаем pandas / пер. с англ. А.В. Груздева. – М.:ДМК Пресс, 2018. – 438 с.: ил.
8. Чернова Н. И. Математическая статистика: Учеб. пособие / Новосиб. гос. ун-т. Новосибирск, 2007. 148 с.
9. Devpractice Team. Pandas. Работа с данными. 2-е изд. - devpractice.ru. 2020. - 170 с.:  
Ил.
10. Приказ Министерства природных ресурсов Российской Федерации от 14 июня 2016 г. N 356  
"Об утверждении Правил разработки месторождений углеводородного сырья"

11. Dean R. B. and Dixon W. J. (1951) "Simplified Statistics for Small Numbers of Observations". *Anal. Chem.*, 1951, 23 (4), 636–638

12. Dixon W.J.: "Processing data for outliers Reference": *J. Biometrics* 9 (1953) 74-89

13. Rorabacher David B. (1991) "Statistical Treatment for Rejection of Deviant Values: Critical Values of Dixon's 'Q' Parameter and Related Subrange Ratios at the 95% Confidence Level." *Analytical Chemistry* 63, no. 2 (1991): 139–46.





# Приложение 3

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
well	lastD_TR	lastD	Phuf_TR	Phuf	Phn_TR	Phn	Prg_TR	Prg	Prab_TR	Prab	Prab Isl	Qoil_TR	Qoil	roll	Q_TR	Q	W_TR	w	Qog_TR	Qog	Qong_TR	Qong	Grng	Gr_TR	K	AA	AB
0	1	18	18	29	29	29	72	72	58	6,598028	8,865729	0,836	8,323	10,606	0	0,01	14,705	5,905	0	848,66	2113,386	1,8	padore	last			
1	2	18	18	27	27	26	73	72	49	12,1638	13,28572	0,836	15	16,549	3	3,97	16,7	4,7	0	386,324	1372,926	1,8	padore	last			
2	3	65	65	27	27	25	74	74	51	5,849074	5,378274	0,836	7	8,828	0,05	5,78	8,7	4,4	0	586,795	752,258	1487,415	1,8	padore	middle		
3	4	65	65	26	26	24,94211	74	77,25	107	1,2122	4,050322	0,836	5	8,998	71	46,156	7	2	0	1649,893	5774,625	1,8	padore	middle			
4	5	65	65	26	26	25	78	72	57	5,016	4,629768	0,836	6	5,338	0	0	6,5	4,3	0	857,258	1295,853	0,5	padore	last			
5	6	65	65	27	27	26	74	70	45	3,59062	7,856728	0,836	4,295	9,398	0	0	11,5	9,03	0	2514,886	3202,789	0,5	padore	last			
6	7	18	18	16	16	19	74	78	39	4,245878	4,410032	0,836	5,09	5,307	0,22	0,6	5,7	-0,3	0	800,766	1342,478	1,8	padore	last			
7	8	18	18	29	29	25	77	72	125	25,39622	25,20477	0,836	43,779	30,935	30,61	2,54	18,5	6,5	18676	255,946	728,455	0,333333	1,8	padore	last		
8	9	65	65	29	29	28	76	71	58	7,524	15,7487	0,836	9	18,8575	0	0	10	5,8	0	770,866	1329,08	0,5	padore	last			
9	10	18	18	32	32	29	80	82	71	2,758248	8,100256	0,836	3,3	9,699	0,02	0,1	6,4	2,4	0	870,1175	2320,313	0,5	padore	last			
10	11	18	18	18	18	17	78	77,5641	33	3,344	6,135639	0,836	4	7,342471	0	0,043529	23,372	20,6	0	6208,715	9222,654	1,8	padore	middle			
11	12	18	18	18	18	19	78	77,8	34	3,317917	4,095252	0,836	4,4	5,778443	9,8	15,22143	30,6	20,6	0	6208,715	9222,654	1,8	padore	middle			
12	13	18	18	19	19	19	80	76,16667	102	6,647872	3,538831	0,836	14,2	10,493	44	59,65833	15,8	3,8	152,9747	571,615	2376,7	1,8	padore	middle			
13	14	18	18	24	24	22	75	80	42	13,25896	13,50976	0,836	15,86	16,16	0	0	34,449	15,749	0	1187,801	2598,168	1,8	padore	last			
14	15	18	18	18	18	21	19	77	50	18,4756	18,83006	0,836	22,1	22,524	0	0	19	12	0	649,503	1028,383	1,8	padore	last			
15	16	18	18	18	18	17	80	76	42	10,868	8,461156	0,836	13	10,121	0	0	24	18	0	1656,238	2208,318	1,8	padore	last			
16	17	65	65	30	30	29	80	83,71795	68	15,048	12,03381	0,836	18	15,91565	0	9,5575	17	6,2	0	412,0149	1129,718	0,5	padore	last			
17	18	18	18	17	17	16	80	79,2	37	5,670625	8,472122	0,836	6,798	10,1435	0,22	0,0925	8,463	2,463	0	434,346	1492,428	1,8	padore	middle			
18	19	18	18	19	19	18	78	78,66667	39	5,323648	10,40363	0,836	6,4	12,596	0,5	1,2025	23	14,6	0	2742,48	4920,346	0,5	padore	middle			
19	20	18	18	18	18	19	80	78	36	4,71504	27,30812	0,836	6	35,192	6	7,18	15	10,5	2992,476	2226,916	3181,309	1,8	padore	middle			
20	21	65	65	18	18	19	78	78	38	9,80628	8,281992	0,836	11,73	10,557	0	6,16	21	11,4	0	1162,52	2141,485	1,8	padore	last			
21	22	18	18	28	28	27	78	77	48	17,3052	18,53078	0,836	20,7	22,166	0	0	21,642	11,242	0	649,6313	1250,607	1,8	padore	last			
22	23	18	18	35	35	34	82	82	54	9,125107	5,089781	0,836	12	10,071	9,04	39,54667	18,1	8,5	4032,233	931,4959	1983,538	1,8	padore	middle			
23	24	65	65	18	18	34	80	77,81818	73	34,77685	34,78033	0,836	41,695	41,695	0,23	0,22	78	78	81114,3	2242,871	2242,871	1,8	padore	last			
24	25	65	65	18	18	42	80	77,81818	66	11,58696	21,56023	0,836	14	25,869	1	0,306333	23,548	11,548	2630,333	996,6376	2032,285	1,8	padore	middle			
25	26	18	18	32	32	25	76	75,33333	66	9,196	9,694152	0,836	11	11,55888	0	0	9	3,2	0	347,9774	978,8684	0,5	padore	middle			
26	27	65	65	27	27	26	78	76,22222	63	2,926	2,214705	0,836	3,5	2,649875	0	0,026667	9	5	0	1708,817	3075,871	0,5	padore	middle			
27	28	18	18	26	26	25	81	76,85714	137	2,421733	5,067832	0,836	2,9	6,062	0,11	0	16,385	9,685	8061,441	3999,202	6705,816	1,8	padore	middle			
28	29	18	18	32	32	31	77	76	70	2,926	4,459112	0,836	3,5	5,336	0	0,04	6,7	6,7	0	2289,815	2289,815	1,8	padore	last			
29	30	18	18	49	49	48	51	51	46	23,06226	22,06498	0,836	27,7	26,412	0,41	0,07	83	83	87206,7	3598,954	3598,954	1,8	padore	last			
30	31	62,1	18	31	31	30	79	73	46	3,340656	5,090875	0,836	4	6,092	0,1	0,04	10,5	10,5	12854,3	3143,095	3143,095	1,8	padore	last			
31	32	62,3	18	31	31	30	71	73	54	5,883768	5,664736	0,836	7,038	6,776	0	0	12,819	5,619	0	955,0003	2178,706	1,8	padore	last			
32	33	69	18	31	31	30	71	73	54	16,06792	16,06959	0,836	19,22	19,222	0	0	94,3	94,3	94302	5868,837	5868,837	1,8	padore	last			
33	34	70,1	65	18	18	31	76	76	54	6,599992	14,16032	0,836	9,277	19,90866	14,9	14,9	32,819	32,819	32819,1	4972,582	4972,582	1,8	padore	middle			