



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Выпускная квалификационная работа

## На тему:

### Прогнозирование технологических остановок экструдера в производстве полипропилена

Выполнил: Есаков Н.С.

Научный руководитель:

Абдуракипов С.С., PhD

Руководитель лаборатории искусственного интеллекта  
ПАО «НК «Роснефть»

# Описание предметной области

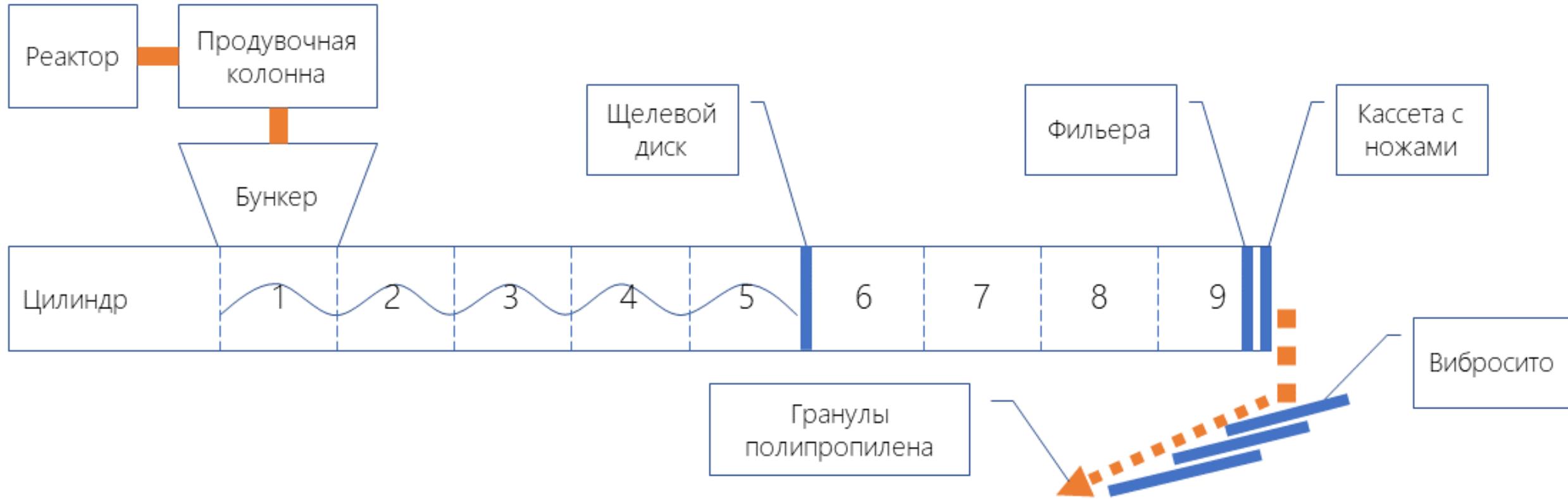
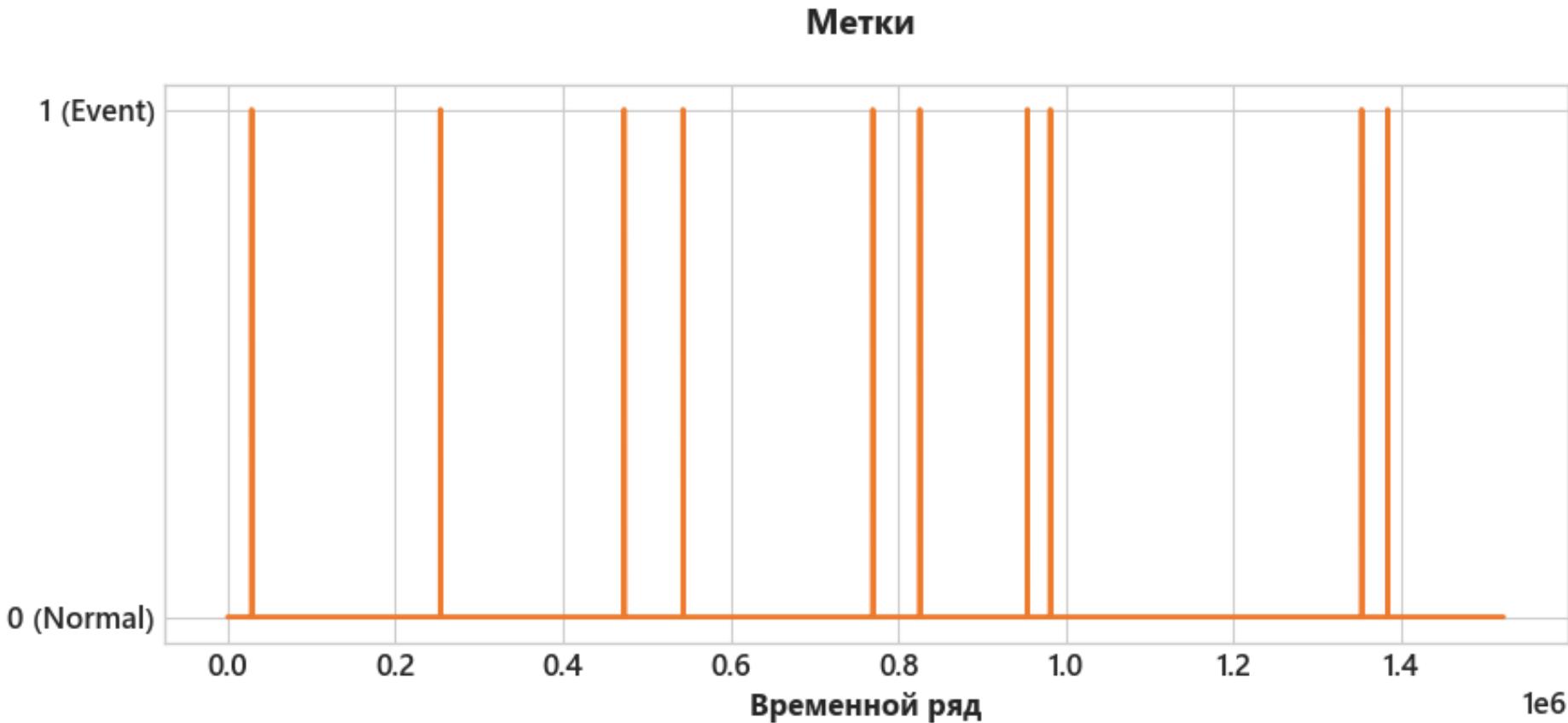


Схема установки производства гранул полипропилена

# Постановка задачи проекта



Задача: Создание предиктивной модели для предупреждения остановок технологической линии гранулирования полипропилена.

Временной интервал содержит 9 меток отказов

Частота дискретизации - 10 секунд

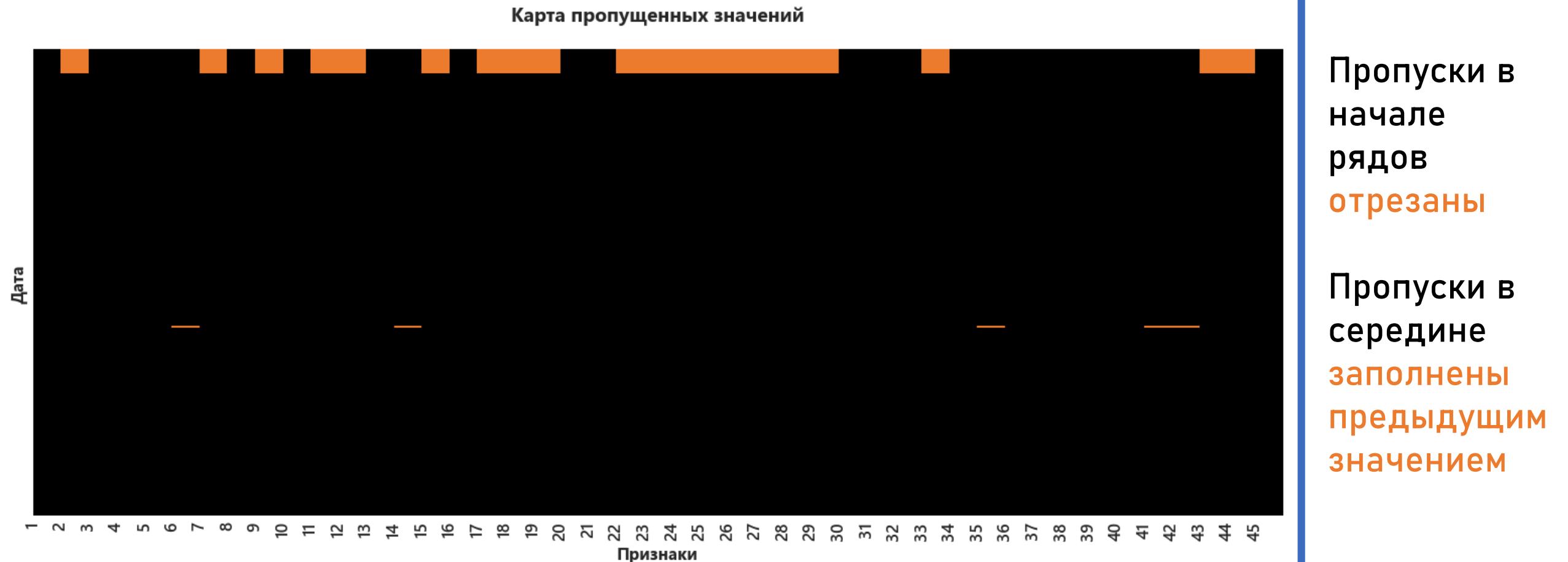
# Источник и структура датасета

Исходный датасет представляет собой временные ряды, сформированные системой сбора данных SCADA за период с 01.01.2017 по 26.06.2017. Общий объем данных — 1,5 млн. записей с временной меткой, **43 признака** (технологических параметра).

## Функциональные группы признаков:

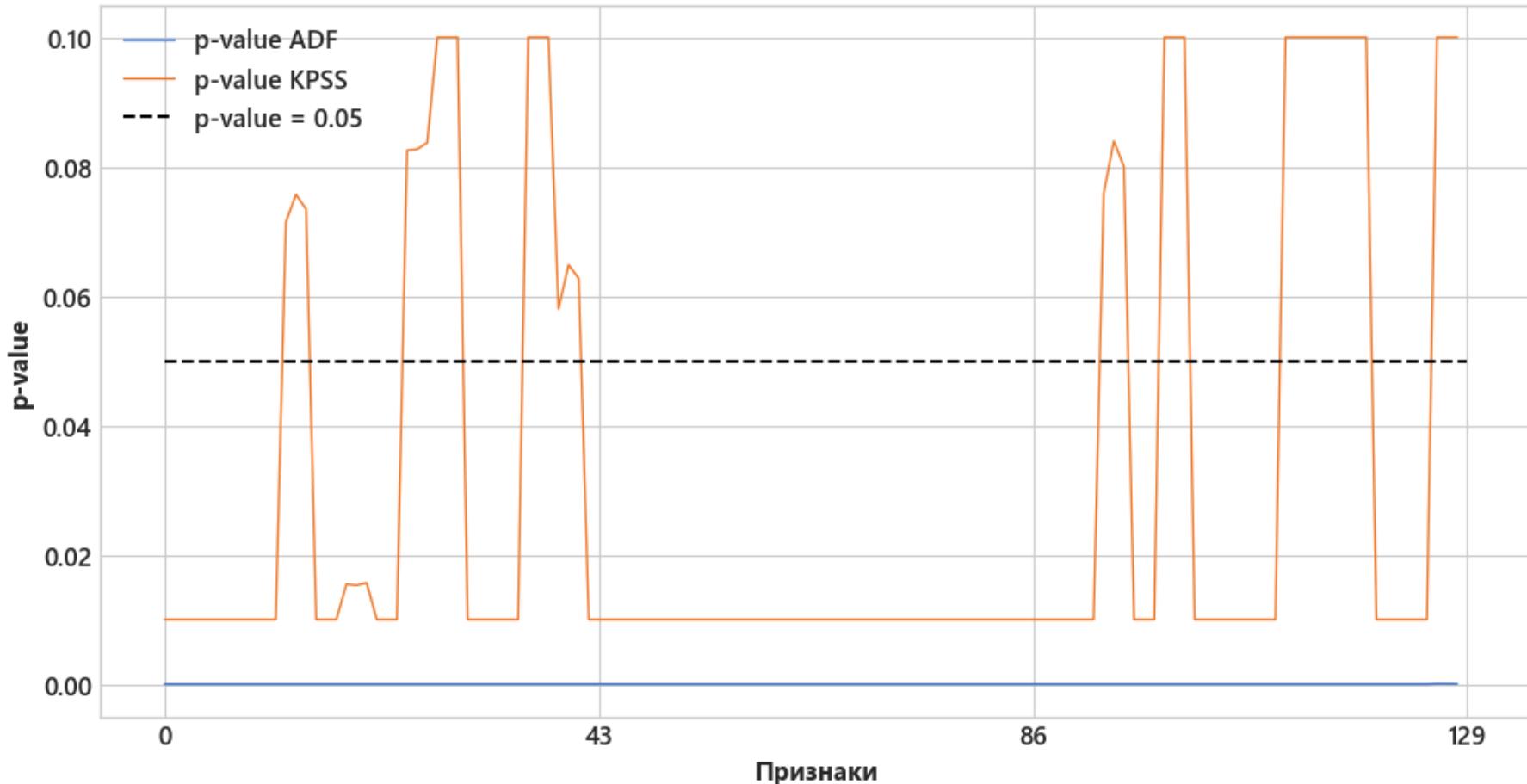
1. Сила тока электродвигателей
2. Частота вращения электродвигателей
3. Температура полипропилена на разных участках цилиндра
4. Внешние температуры цилиндра
5. Температура щелевого диска
6. Температура охлаждающей воды
7. Положение ножей
8. Расход подачи порошка полипропилена
9. Давление расплава до и после сит, до фильтры.

# Анализ пропусков



# Проверка стационарности

Результаты тестов на стационарность для агрегированного ВР  
Количество стационарных признаков (ADF) - 129  
Количество стационарных признаков (KPSS) - 33



Исходный ряд  
ресемплируется с  
интервалом 10 минут  
Данный downsampling  
позволяет уменьшить  
размерность данных и  
шум, а также убрать  
некоторые пропуски

Агрегация  
статистиками  
`max()` – максимальное  
значение за 10 минут  
`min()` – минимальное  
значение за 10 минут  
`mean()` – среднее  
значение за 10 минут

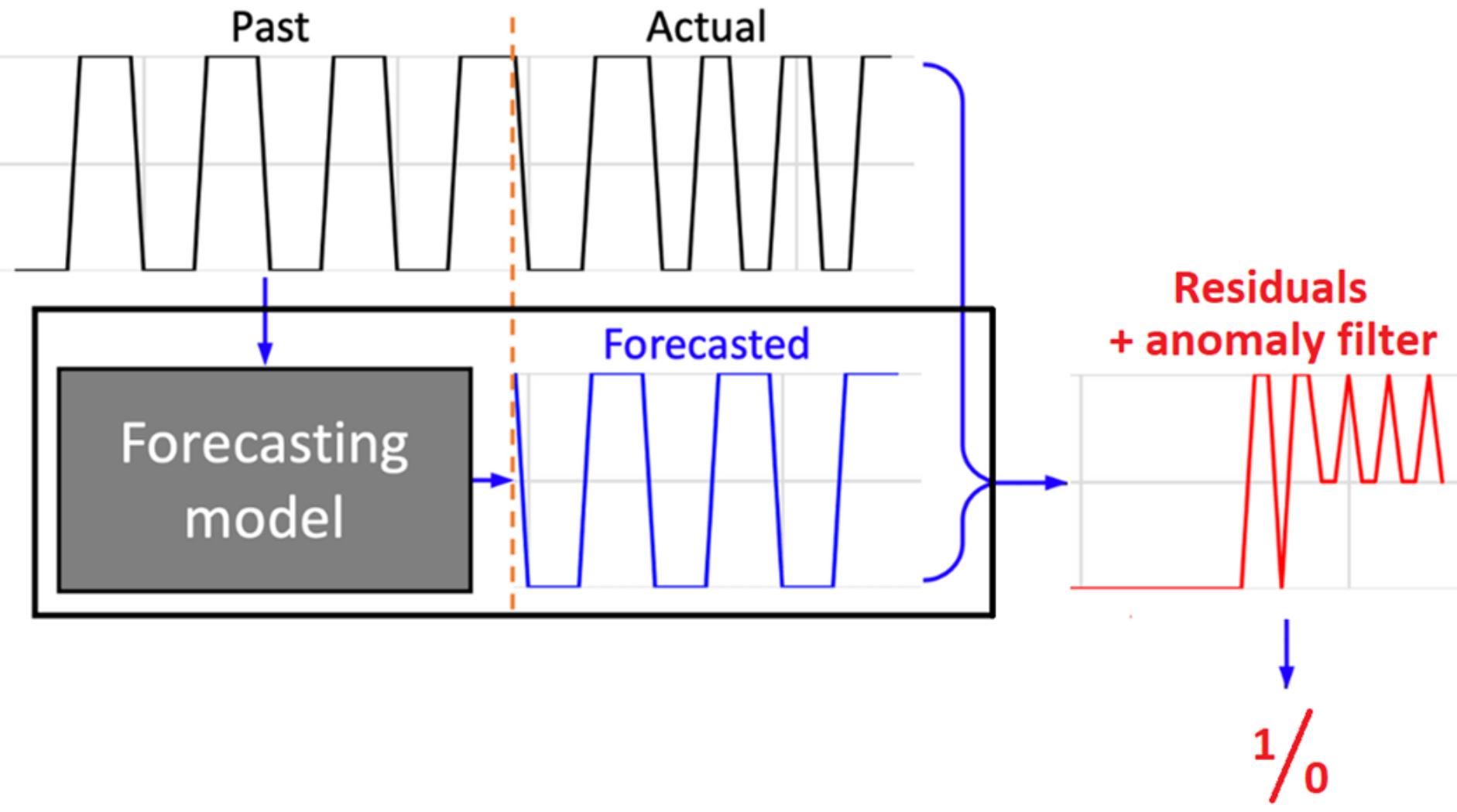
# Проверка сезонности



Обнаружена  
месячная  
сезонность по  
всем признакам.

Для улучшения  
прогнозных  
качеств модели  
месяцы в  
формате One-  
Hot Encoding  
были  
использованы в  
качестве  
ковариат при  
обучении.

# Архитектура системы обнаружения аномалий



В качестве прогнозной модели был выбран **Transformer Model**

Критерии выбора:

- Учёт кросс-корреляций многомерного ряда
- Работа с длинными временными зависимостями
- Эффективность обучения на большом объеме данных

# Обучение модели трансформера

Основные параметры модели:

Входной сегмент: 288 точек

Прогнозный горизонт: 1 точка

Размер эмбеддингов: 128

Feed-forward сеть: 512

Dropout: 0.1

Активация: ReLU

Количество attention heads: 4

Энкодер: 2 слоя,

Декодер: 1 слой

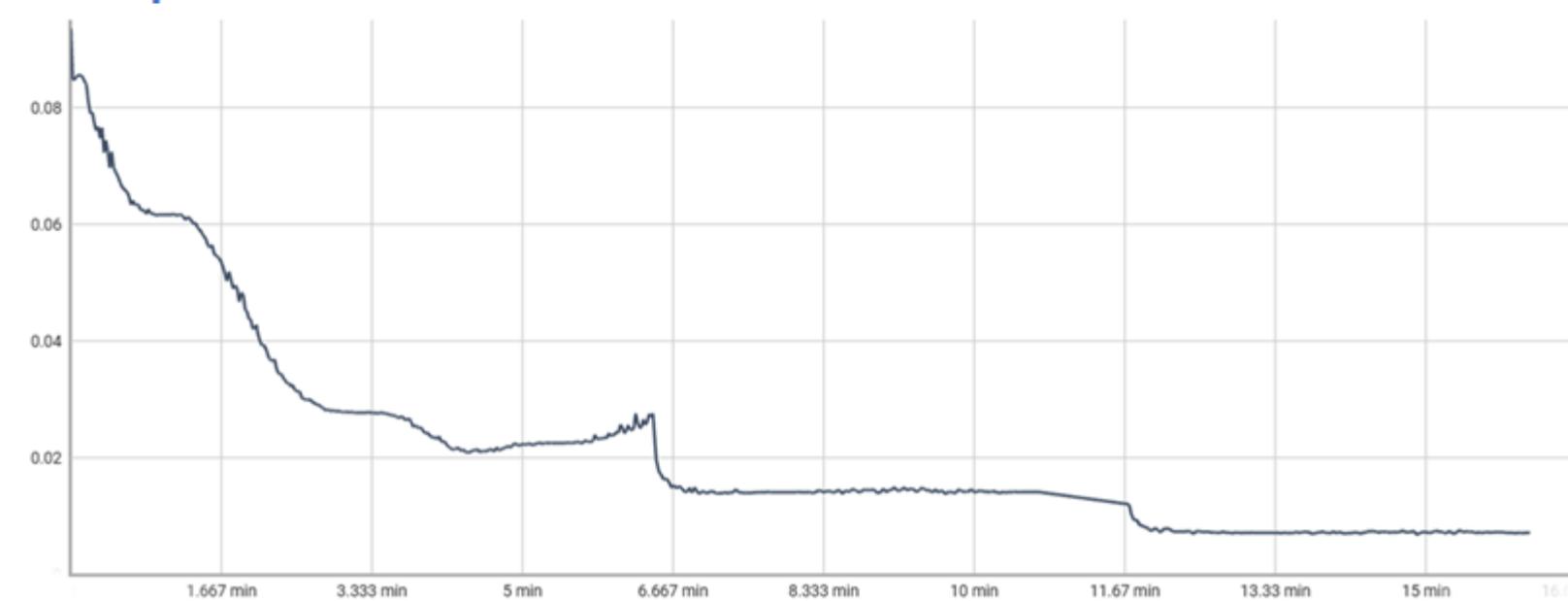
Оптимизатор: Adam

Scheduler: CosineAnnealingLR

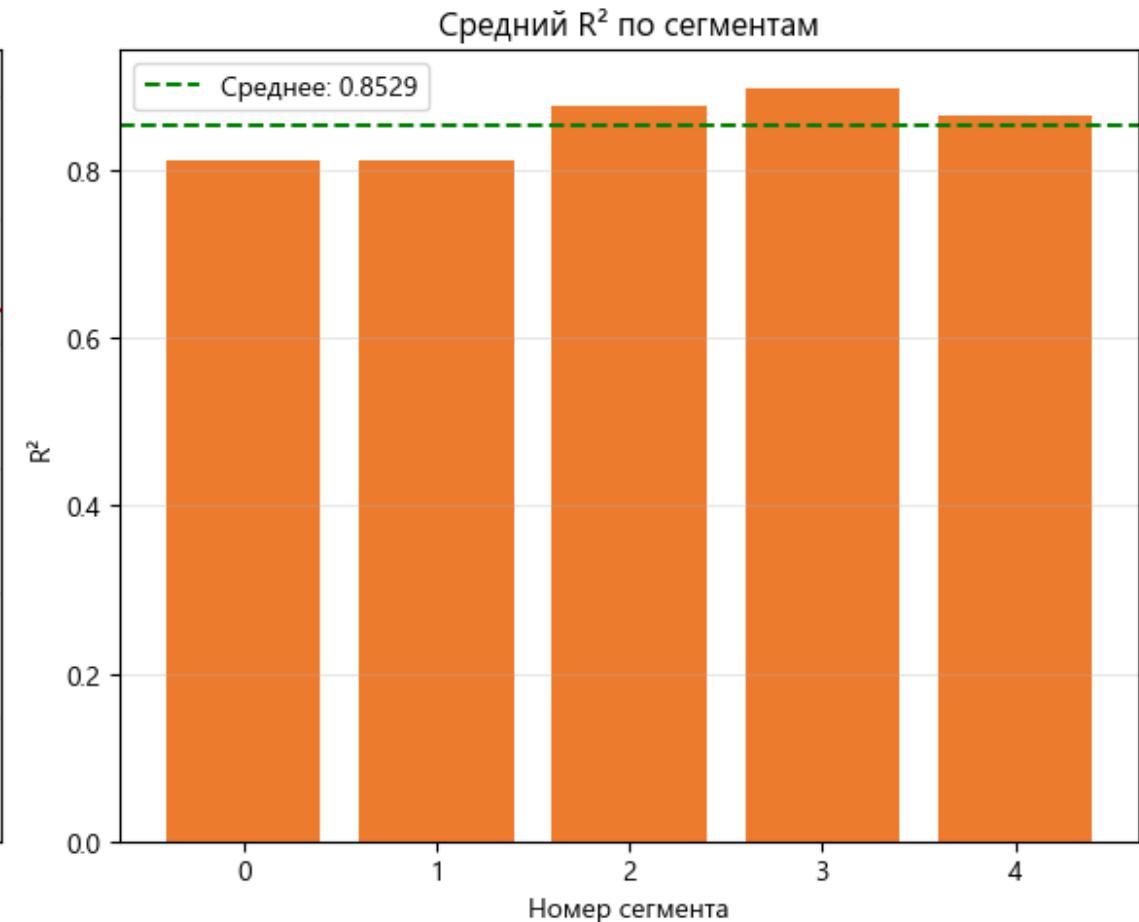
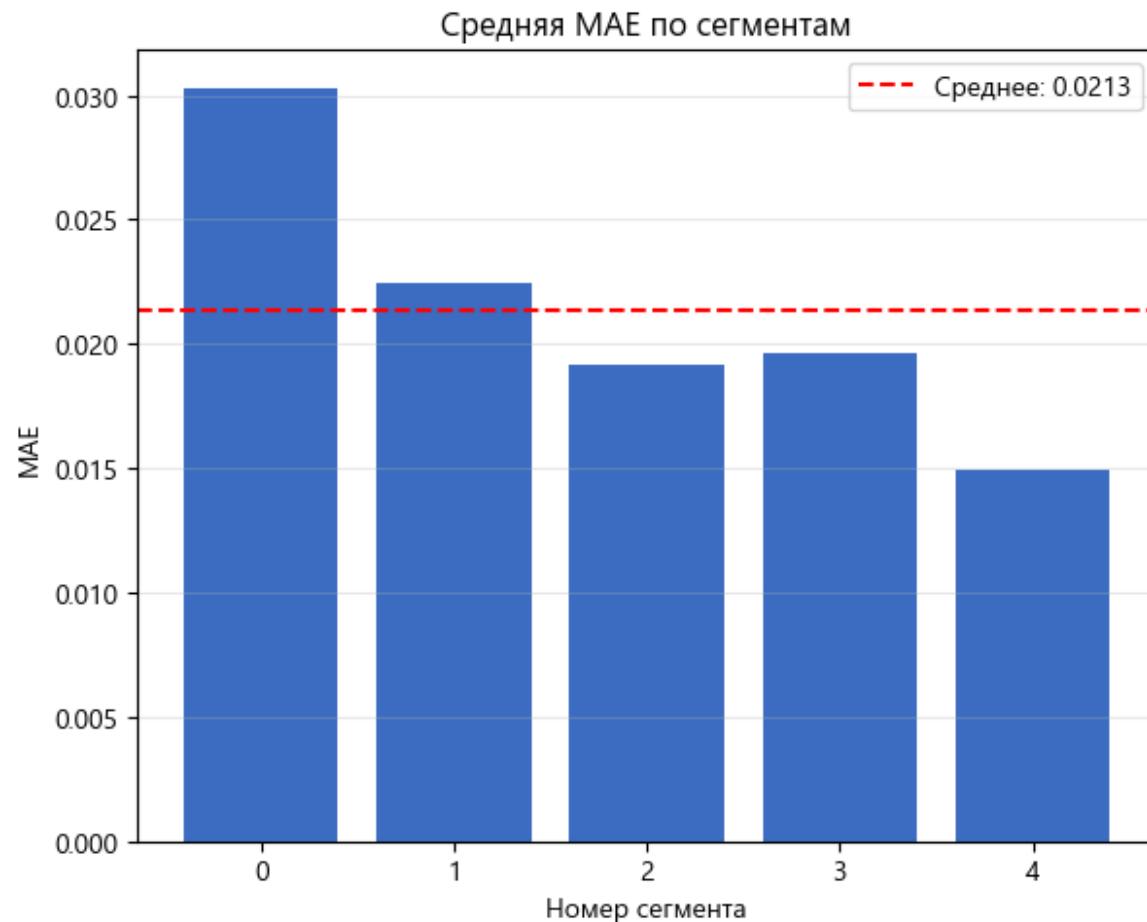
Потери на train



Потери на val



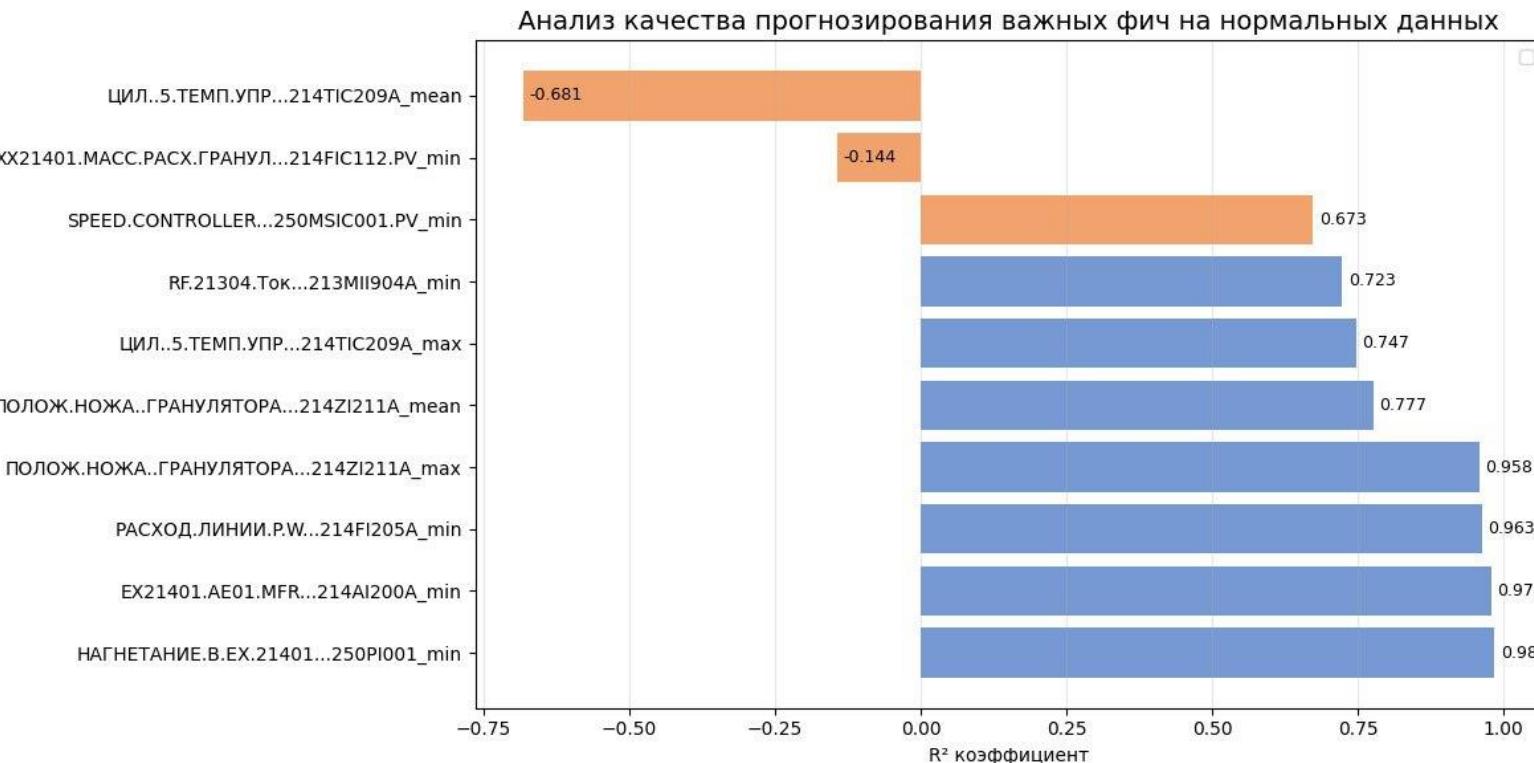
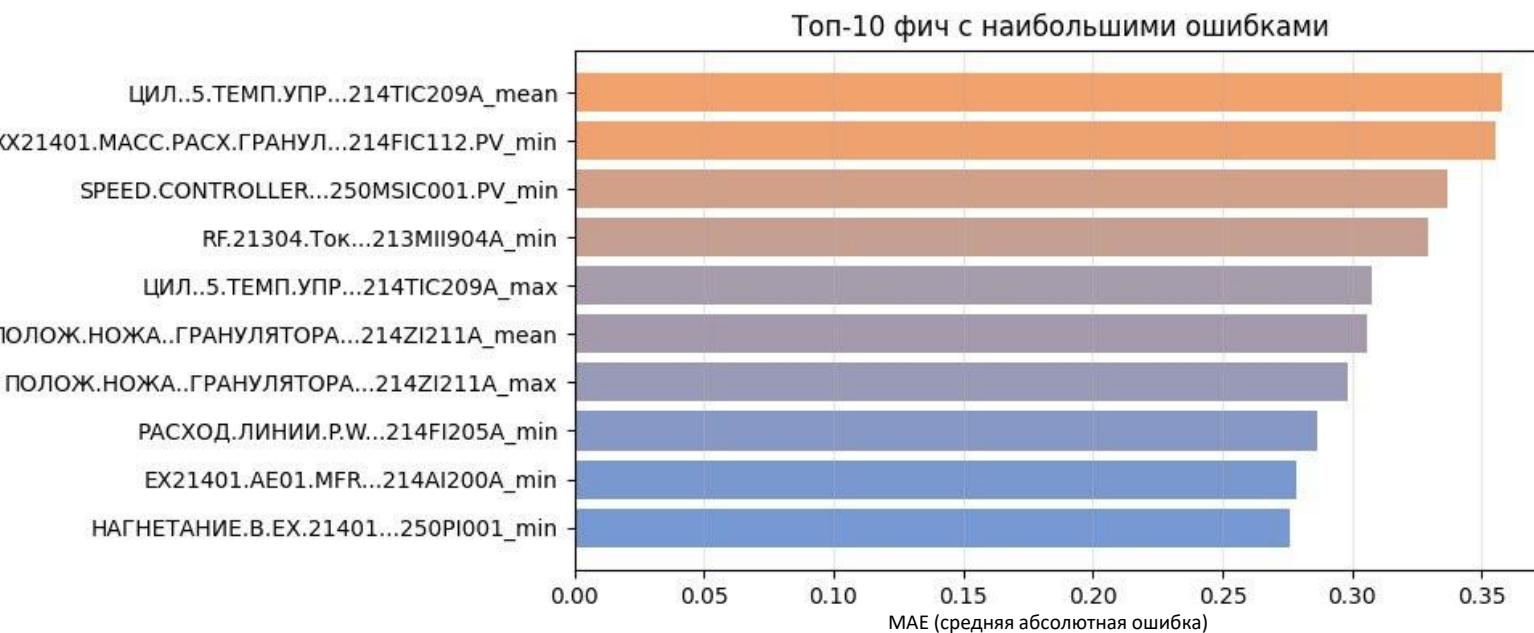
# Тестирование модели на нормальных данных



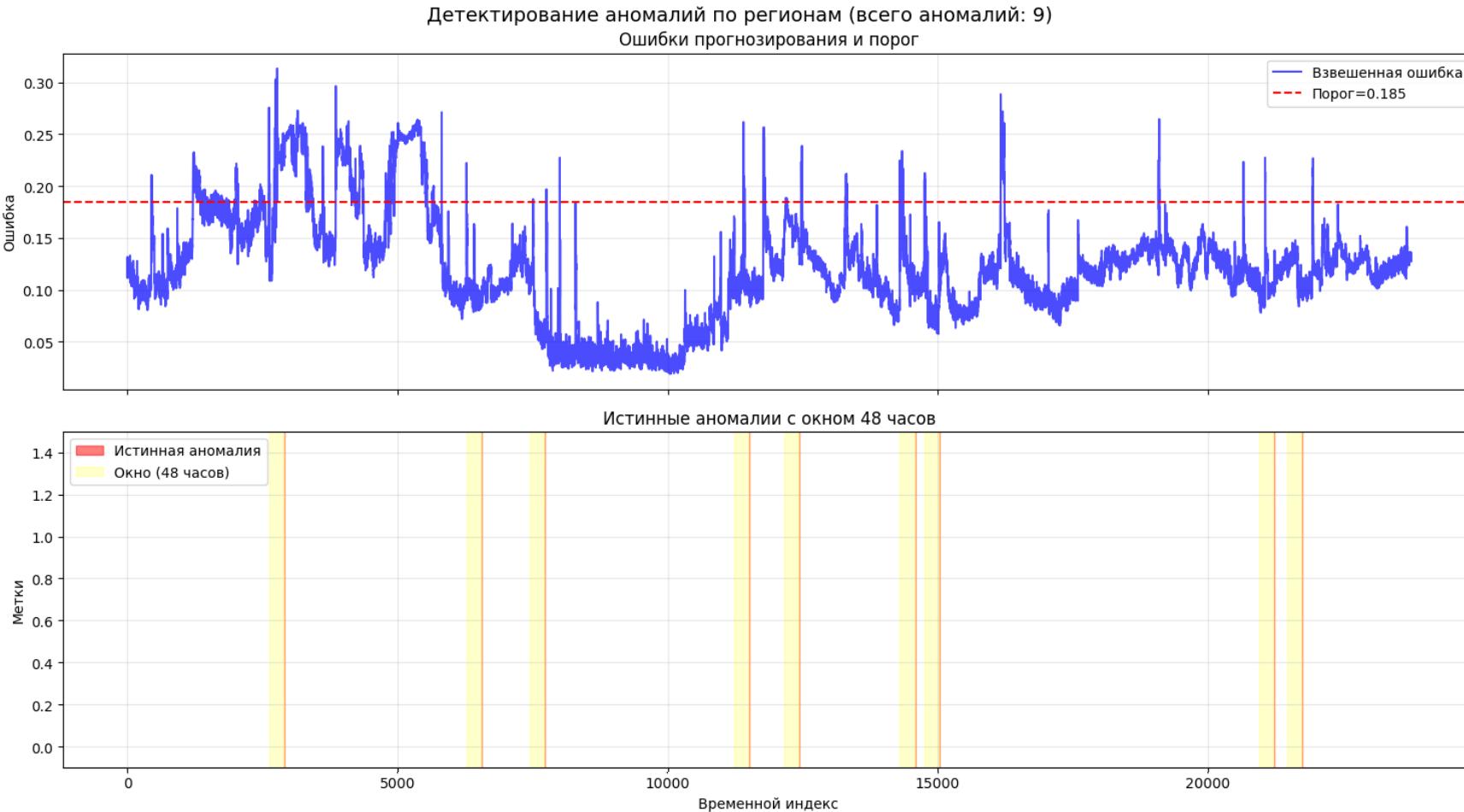
# Первичный отбор признаков для детекции аномалий

Для прогнозирования были выбраны **нормализованные** признаки с наибольшими значениями ошибок **МАЕ**.

Дополнительно проведена оценка коэффициента детерминации ( $R^2$ ) с целью отбора только тех признаков, которые хорошо аппроксимируются моделью.



# Результаты детекции

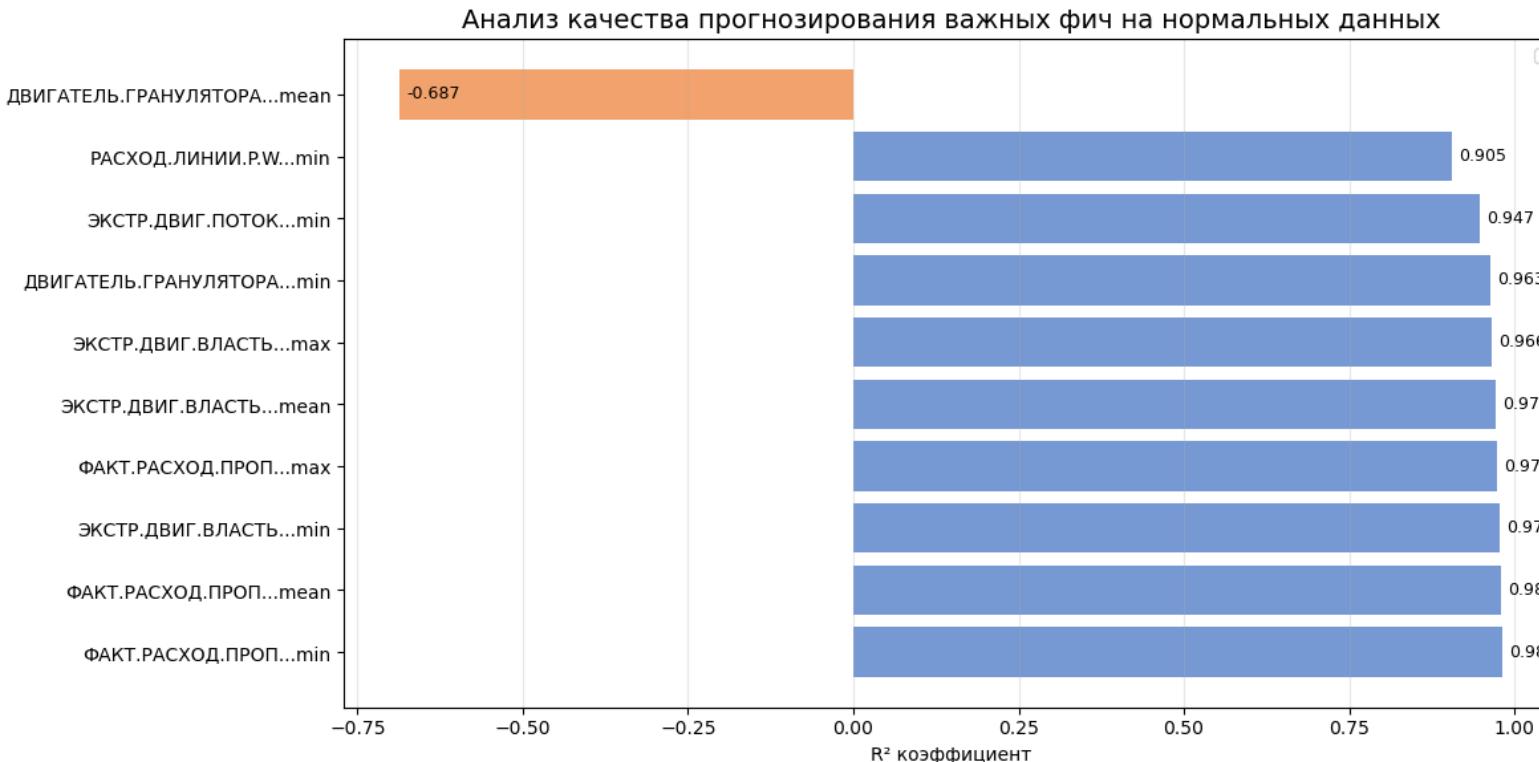
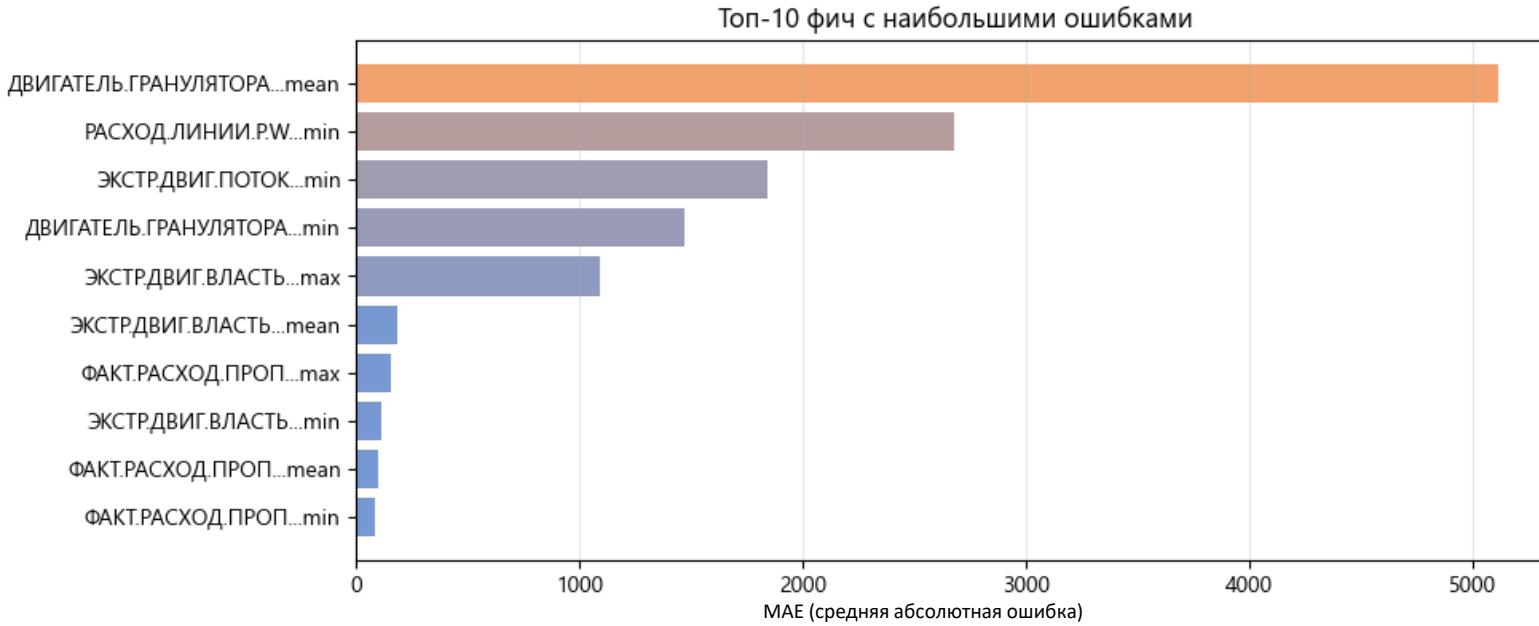


Итоговый отчет:  
Всего аномалий в  
тесте: 9  
Обнаружено: 5  
Пропущено: 4  
Ложных срабатываний:  
49

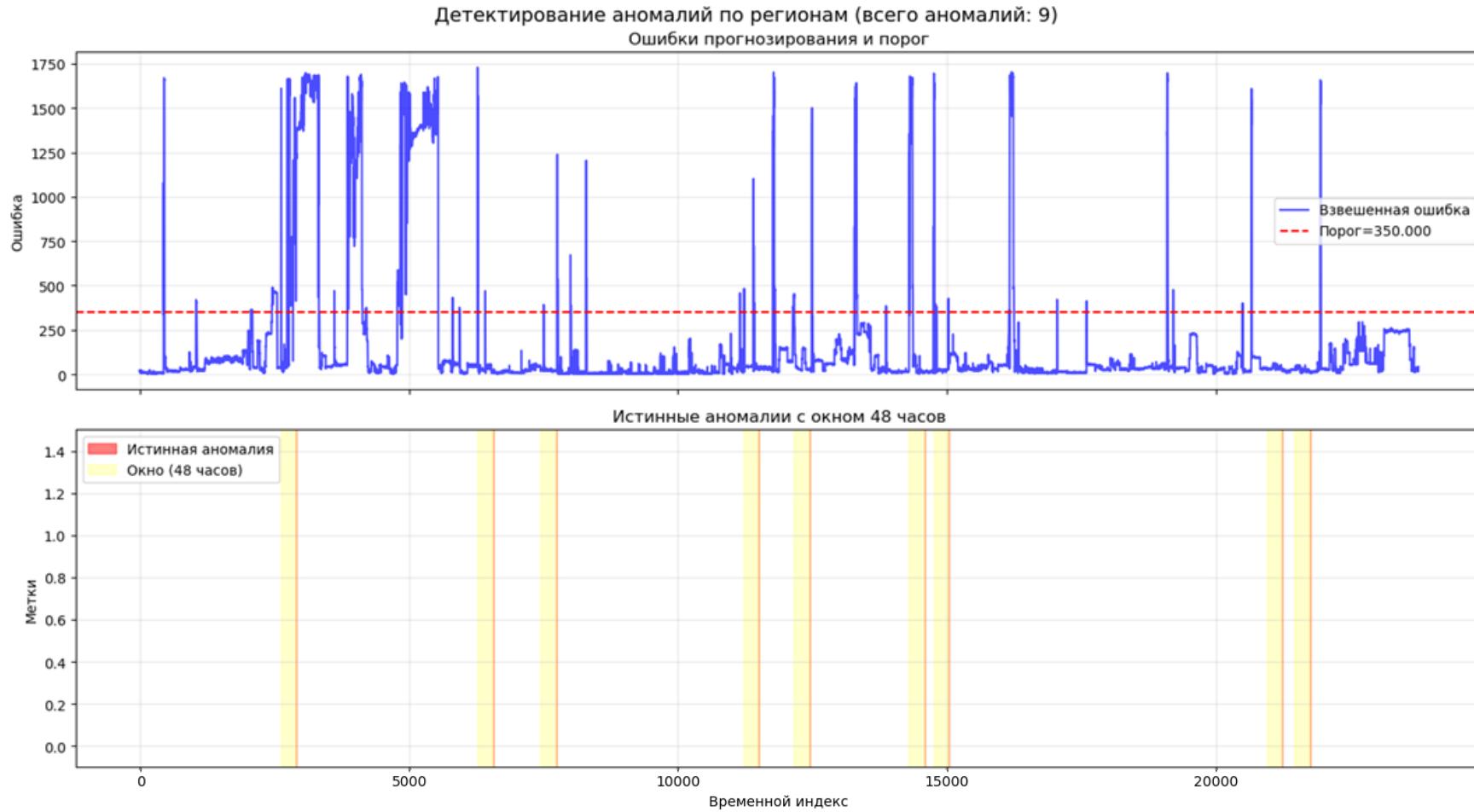
Precision: 0.093  
Recall: 0.556  
F1-score: 0.159

# Альтернативный отбор признаков для детекции аномалий

В качестве альтернативы статистически корректному подходу предлагается метод отбора признаков, основанный на анализе абсолютных ошибок прогнозирования в исходных физических единицах измерения.



# Результаты детекции и выводы



Всего аномалий в  
тесте: 9  
Обнаружено: 7  
Пропущено: 2  
Ложных срабатываний:  
**19** (за полгода)

Precision: 0.269  
Recall: 0.778  
F1-score: 0.400

# **Направления дальнейших исследований**

## **1. Расширение обучающей выборки**

Увеличение объема исторических данных для более качественного обучения модели

## **2. Увеличение прогнозного горизонта**

Расширение выходного окна модели с 1 до N точек прогноза

## **3. Эксперименты с архитектурами трансформеров**

Тестирование специализированных архитектур для многомерных временных рядов (PatchTST, TFT и тд)

## **4. Внедрение контекстуальных Scorer на этапе обработки остатков прогноза.**

## **5. Использование механизма внимания трансформера для интерпретируемого анализа причинно-следственных связей в момент сбоя.**