

# Предпрогнозный анализ временного ряда тока экструдера (тег 14)

## Период анализа

- **Общие данные:** 05.03.2024 – 05.03.2025
- **Выборки:** месячные, суточные, 4-часовые интервалы

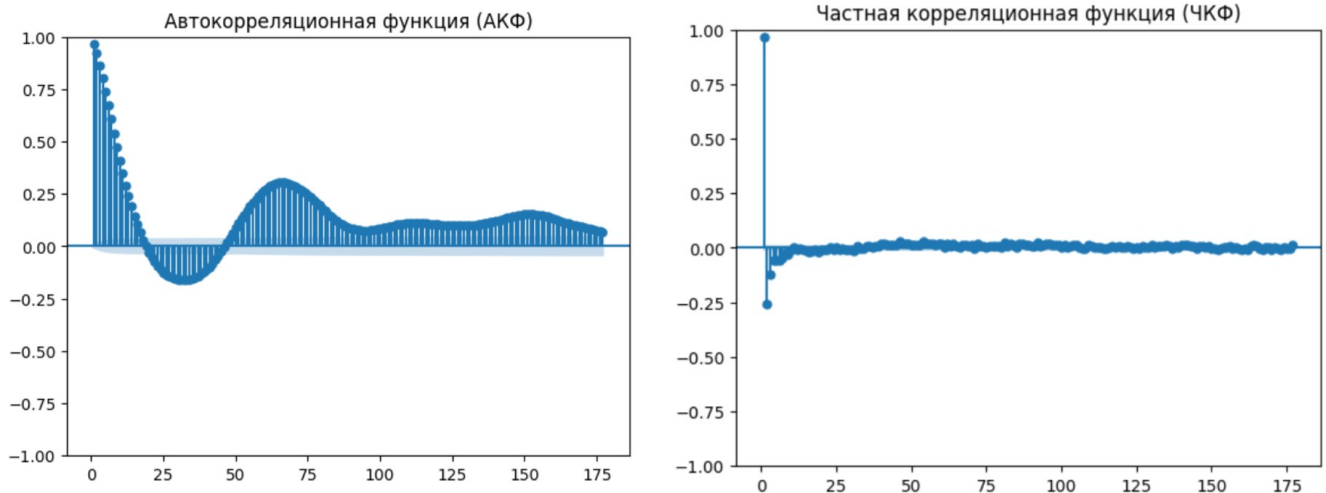
## Ключевые статистики

Период	Среднее	СКО	Асимметрия	Экссесс
Апрель 2024	292	61.6	-1.07	0.72
Октябрь 2024	325	99.3	-0.45	-1.21
Февраль 2025	308	79.2	-0.38	-0.79

- ◆ Распределение **не нормальное**, с левосторонней асимметрией.
- ◆ Высокая дисперсия в октябре — признак **нестабильности процесса**.

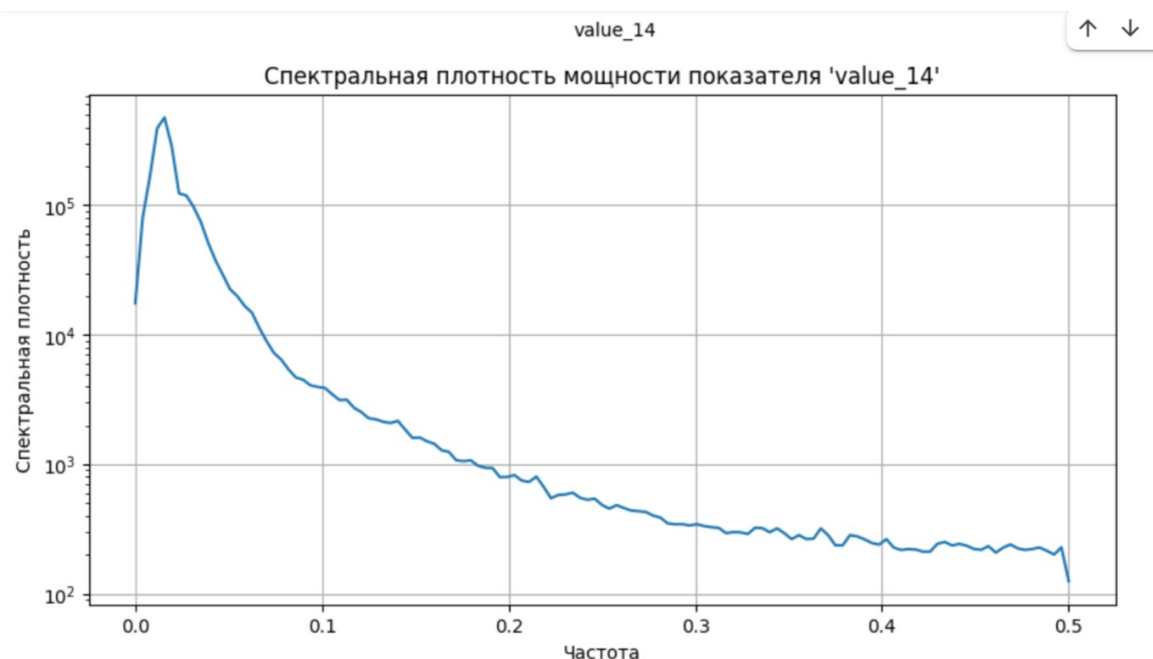


## 🔍 Стационарность и автокорреляция



- **ADF-тест:**
  - $p\text{-value} = 0$  на месячных/суточных интервалах → **ряд близок к стационарному**
  - На 4-часовых интервалах — **нестационарен** → требует дифференцирования
- **ACF (автокорреляция):**
  - Очень высокая на лагах 1–2 → **сильная краткосрочная зависимость**
  - Волнообразное поведение на лагах 75–125 → **признак цикличности** (возможно, сменная или суточная)
- **PACF (частичная автокорреляция):**
  - Резкий спад после лага 2 → признак **AR(2)** компоненты

## 📈 Спектральный анализ



- **Пик на низких частотах** → преобладание долгосрочных трендов
- **Флуктуации на высоких частотах** → возможен шум или кратковременные всплески
- **Энергия сосредоточена в низкочастотной области** → основная динамика — медленные изменения

## Обновлённые выводы и обоснование выбора ML-моделей

### 1. Ряд имеет смешанную структуру

- **Краткосрочная зависимость** (AR-эффект, лаги 1–2)
- **Долгосрочные циклы** (волны в ACF на лагах 75–125)
- **Нестационарность на коротких интервалах**
- **Разные режимы работы** (пуск, работа, простой)

Это не классический стационарный ряд, а сложный промышленный процесс с несколькими масштабами зависимости.

### 2. Почему классические модели (ARIMA) — недостаточны

<b>Смешанные режимы</b>	ARIMA не умеет адаптироваться к режимам (пуск/работа/простой)
<b>Циклы разной длины</b>	SARIMA требует фиксированного периода — а у нас он варьируется
<b>Нестационарность</b>	Частое дифференцирование теряет смысл при переключении режимов
<b>Нелинейность</b>	ARIMA — линейная модель, а процесс содержит нелинейные переходы
✗ ARIMA может дать базовый прогноз, но <b>не уловит сложную динамику</b> .	

### 3. Почему выбраны нейросети: CNN, LSTM, GRU, Attention

<b>LSTM / GRU</b>	Могут запоминать <b>долгосрочные зависимости</b> (лаги 75–125), адаптироваться к <b>режимам работы</b>
<b>CNN</b>	Обнаруживает <b>локальные паттерны</b> (всплески, падения тока) — как "сигнатуры" отказов
<b>Слой внимания (Attention)</b>	Позволяет модели <b>фокусироваться на ключевых участках</b> (например, перед отказом)
<b>Гибридные архитектуры</b>	CNN + LSTM + Attention — <b>максимальная гибкость</b> для сложного ряда

✓ Такая модель может:

- Учитывать **и краткосрочные, и долгосрочные зависимости**
- Различать **режимы работы**
- Обработать **нестационарность без дифференцирования**
- Обучаться на **сырых данных** (с нулями, всплесками и шумом)

---

#### 4. Как анализ поддерживает ML-подход

- **Высокая автокорреляция на малых лагах** → оправдывает LSTM/GRU для краткосрочного прогноза
  - **Цикличность на лагах 75–125** → требует долгой памяти → поддерживает использование **внимания**
  - **Нестационарность** → делает **ручную предобработку рискованной** → лучше, чтобы **сеть училась сама**
  - **Разные режимы** → идеальный кейс для **обучения на размеченных данных** (как ты и сделал)
- 

#### ✓ Общие выводы

1. **Временной ряд тока экструдера — сложный, нестационарный, с несколькими масштабами зависимости.**
  2. **Классические модели (ARIMA, SARIMA) — ограничены** в описании таких процессов.
  3. **Нейросети (LSTM, GRU, CNN, Attention) — обоснованный выбор**, так как:
    - Могут моделировать **нелинейные и многомасштабные зависимости**
    - Не требуют жёсткой предобработки
    - Адаптивны к режимам работы
  4. **Анализ ACF, PACF, спектра — не противоречит использованию ML, а обосновывает его:** он показывает, **почему нужна мощная модель.**
- 

#### 📁 Рекомендации

- Использовать **глубокие модели с механизмом внимания**
- Включить **CNN-слои** для обнаружения локальных аномалий
- Разделить данные на **режимы** (пуск, работа, простой) — для лучшего обучения
- Продолжить анализ **корреляций с другими тегами** (давление, температура) — для мультивариантного прогноза