

# DS-prof-2022-08

# Меня хорошо видно && слышно?



# Защита проекта

## Тема: Прогнозирование электропотребления предприятия с применением искусственных нейронных сетей

Здесь  
могла  
бы  
быть  
ваша  
реклам  
а

Здесь  
могла  
бы  
быть  
ваша  
реклам  
а

Дьяконов Семен

Начальник управления балансов и энергоресурсов  
ООО «НОВОГОР-Прикамье» г. Пермь

# План защиты

Цели проекта

Что планировалось

Используемые  
технологии

Что получилось

Выводы



**Здесь  
могла бы  
быть ваша  
реклама**

# Цели проекта

---

1. Исследование применения искусственных нейронных сетей для краткосрочного прогнозирования потребления электроэнергии предприятием;
- 

# Методы проекта

---

1. Применение Рекуррентных нейронных сетей (РНС, англ. Recurrent neural network, RNN), Длинные цепи элементов краткосрочной памяти (англ. Long short-term memory; LSTM)
  2. Сравнить базовый прогноз и прогноз построенный с применением искусственных нейронных сетей;
-

# Что планировалось



1. Сбор и анализ данных

---



2. Анализ временного ряда

---



3. Моделирование

---



4. Оценка результатов

---



5. Выводы

---

# Используемые технологии

1. Файлы в формате \*.xlsx
2. Excel, python стандартные библиотеки, pytorch, RNN, LSTM
3. Ноутбук MSI GF63 Thin

# Что получилось. Репозиторий на github.

1. Ссылка на репозиторий с исходным кодом:

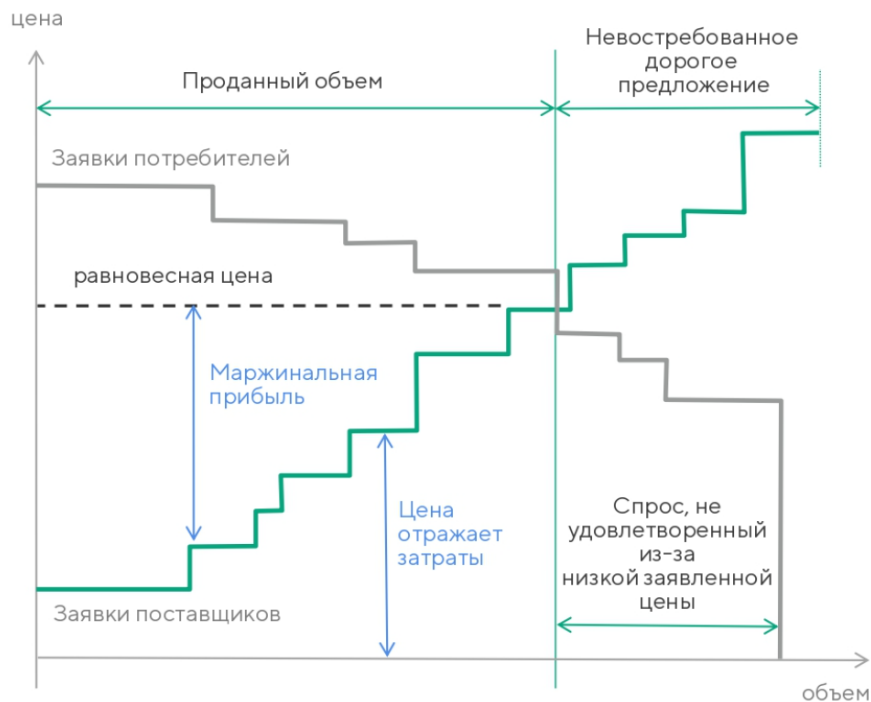
[https://github.com/smddyakonov/otus\\_ds\\_dyakonov/tree/main/hw-19](https://github.com/smddyakonov/otus_ds_dyakonov/tree/main/hw-19)



# Функционирование почасового планирования



# Ценообразование РСВ



РСВ за сутки до поставки определяет плановые объемы производства и потребления, однако фактическое потребление неизбежно отличается от планового. Торговля отклонениями фактических объемов производства/потребления от плановых осуществляется в режиме реального времени на балансирующем рынке.

**Отклонения** фактического потребления от планового квалифицируются собственными или внешними инициативами. **Отклонение по собственной инициативе возникает из-за действий участника рынка (потребителя или поставщика)**, по внешней – в результате команд Системного оператора. Покупатель, потребивший больше планового объема РСВ, покупает электроэнергию на БР, а потребивший меньше – продает. Стоимость отклонений формируется так, чтобы стимулировать участников к более точному исполнению планового потребления и производства электроэнергии, определенного на РСВ, и к выполнению команд Системного оператора. Таким образом, на БР «штрафуются» участники рынка, допускающие наибольшие отклонения фактических объемов потребления и выработки от плановых по собственной инициативе, и «премируются» участники, придерживающиеся планового потребления и максимально точно выполняющие команды Системного оператора. На основе стоимости отклонений определяются предварительные требования и предварительные обязательства БР, разница между которыми формирует небаланс балансирующего рынка. Отрицательный небаланс распределяется между участниками пропорционально их собственным инициативам. Положительный небаланс распределяется между поставщиками, пропорционально величине исполнения внешних инициатив, и потребителями, максимально точно придерживающимися планового потребления. Такая система распределения небаланса является дополнительным стимулирующим и дисциплинирующим фактором для участников рынка.

# Водоснабжение и водоотведение в г.Пермь



## Водоснабжение

Вода в городскую сеть Перми с Чусовских очистных сооружений (Чусовской водозабор (ЧОС)) подается в среднем около 240 тысяч кубометров воды в сутки, 8-9 тысяч кубометров ежесуточно принимает сеть г. Краснокамска. Общая протяженность водопроводных сетей в Перми – 1349.70 км.

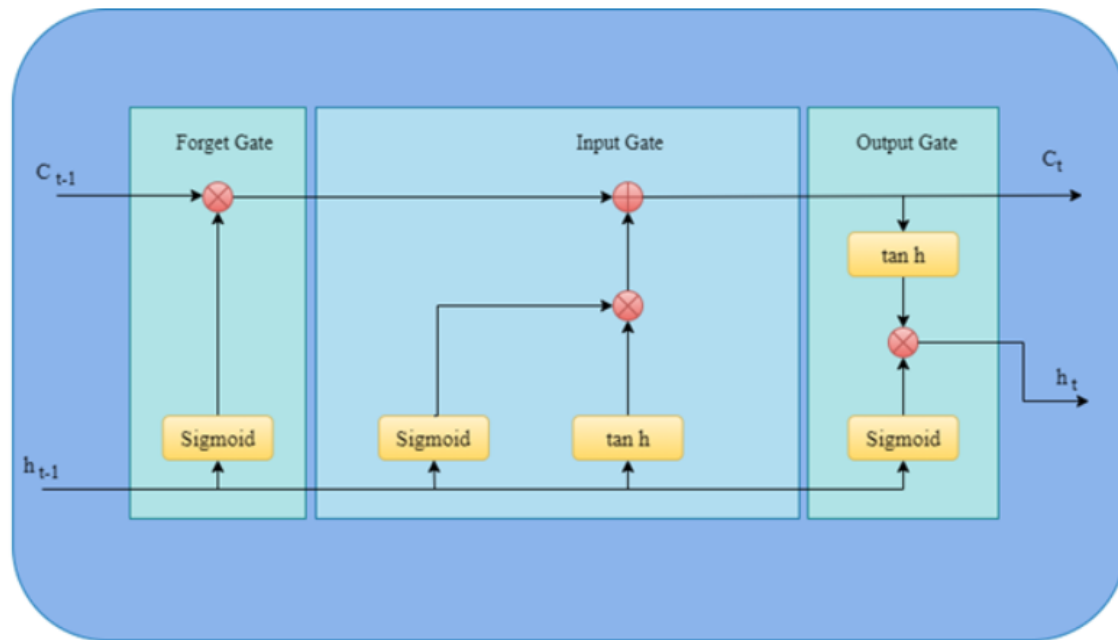
**Транспортировка сточных вод** Общая протяженность сетей водоотведения составляет 1259 км.

• **БОС Гладеново** – очистка городских канализационных сточных вод. Пропускная способность БОС составляет 440 тыс. м<sup>3</sup>/сут.

• **БОС пос. Новые Ляды** – очистка хозяйственно-бытовых сточных вод. Фактическая мощность составляет 2,5-3,5 тыс. м<sup>3</sup>/сут. при максимальной потребности в 5 тыс. м<sup>3</sup>/сут.

Источники: <https://www.novogor.perm.ru/watersnab/7614>  
[https://www.novogor.perm.ru/water\\_o/7616](https://www.novogor.perm.ru/water_o/7616)

# LSTM (долговременная краткосрочная память)



Данные временных рядов носят временной или последовательный характер и могут обрабатываться рекуррентной нейронной сетью (RNN). RNN запоминают прошлую информацию и передают ее в текущее состояние ввода. Однако обычные RNN не способны запоминать долгосрочные зависимости из-за исчезающих градиентов. LSTM — это модернизированный тип RNN, решающий эту проблему. LSTM состоит из блоков, каждый из которых состоит из четырех нейронных сетей, они используются для обновления состояния своей ячейки с использованием информации из новых входных данных и прошлых выходных данных.

Источники: <https://pro.arcgis.com/ru/pro-app/latest/tool-reference/geoai/how-time-series-forecasting-models-work.htm>

# Что получилось. Описание исходных данных

Почасовое потребление электрической энергии представлено

- за период с 2018 по июнь 2023 г.;
- в разрезе точек поставки;
- в разрезе дата-время

## Предварительные преобразования для моделирования

### Сделано в отдельном ноутбуке

- сформирован список исходных файлов
- проверено, что во всех исходных файлах есть лист "Данные эл.эн.". В листе "Данные эл.эн." хранятся почасовые расходы электроэнергии по точкам поставки
- объединить точки поставки (ТП) по годам и найти перечень всех ТП за период с 2019 - 2023 гг
- добавить привязки ТП - Объект - группа точек поставки (ГТП)

# Что получилось.

Ввод [7]:

1 df\_ee

Out[7]:

|       | Дата время (мск)    | ГТП_ВС   | ГТП_ВО  |
|-------|---------------------|----------|---------|
| 0     | 2019-01-01 00:00:00 | 6344.080 | 8562.00 |
| 1     | 2019-01-01 01:00:00 | 6311.008 | 8170.24 |
| 2     | 2019-01-01 02:00:00 | 6345.960 | 7837.36 |
| 3     | 2019-01-01 03:00:00 | 6190.032 | 7835.24 |
| 4     | 2019-01-01 04:00:00 | 6354.032 | 7821.12 |
| ...   | ...                 | ...      | ...     |
| 39379 | 2023-06-29 19:00:00 | 7132.000 | 6772.00 |
| 39380 | 2023-06-29 20:00:00 | 7058.000 | 6860.00 |
| 39381 | 2023-06-29 21:00:00 | 6967.000 | 7041.00 |
| 39382 | 2023-06-29 22:00:00 | 6265.000 | 6772.00 |
| 39383 | 2023-06-29 23:00:00 | 5696.000 | 6664.00 |

39384 rows × 3 columns

В итоге получил дата-фрейм:

Колонка «Дата время (мск)» – дата-время, часовой пояс – «МСК»

Колонка «ГТП\_ВС» - потребление электрической энергии за каждый час, по группе точек поставок (ГТП) водоснабжение (ВС);

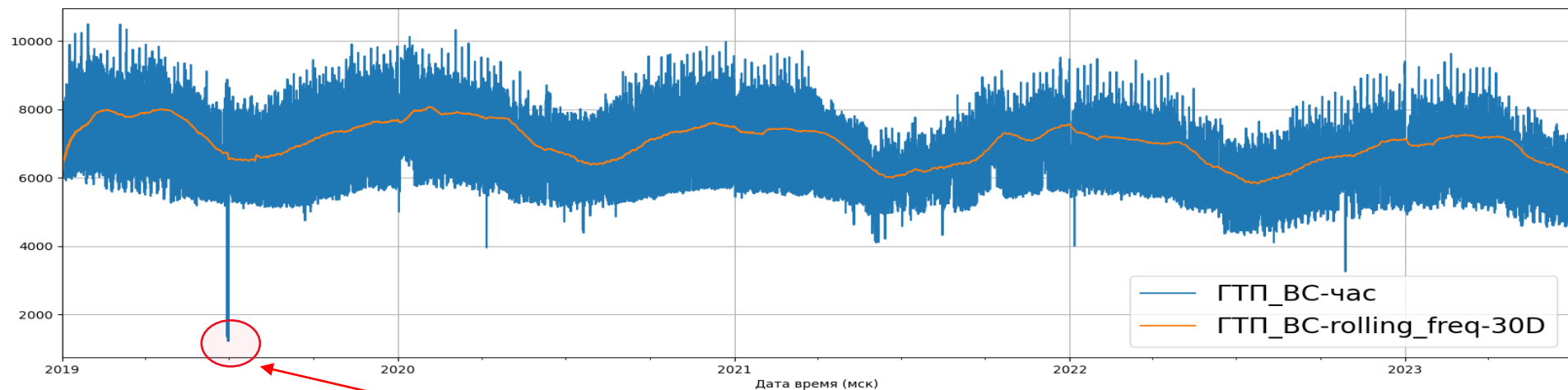
Колонка «ГТП\_ВО» - потребление электрической энергии за каждый час, по группе точек поставок (ГТП) водоотведение (ВО);

Всего строк: 39 384

Тип данных: временной ряд с частотой 1 час

# Что получилось. Предобработка данных (Data Preprocessing)

Профиль потребления электрической энергии за период наблюдения



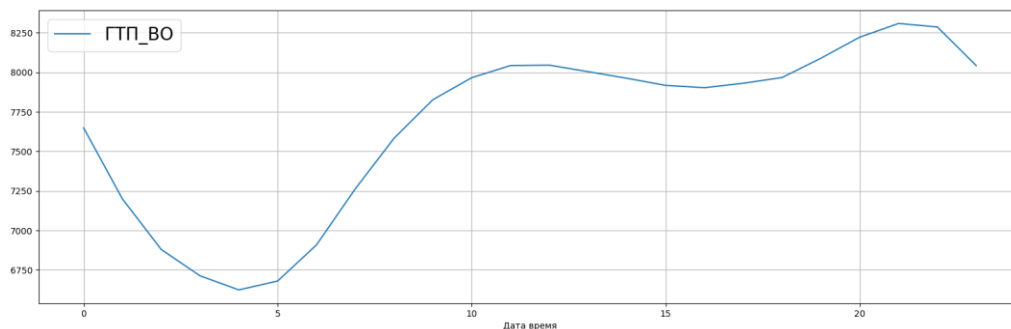
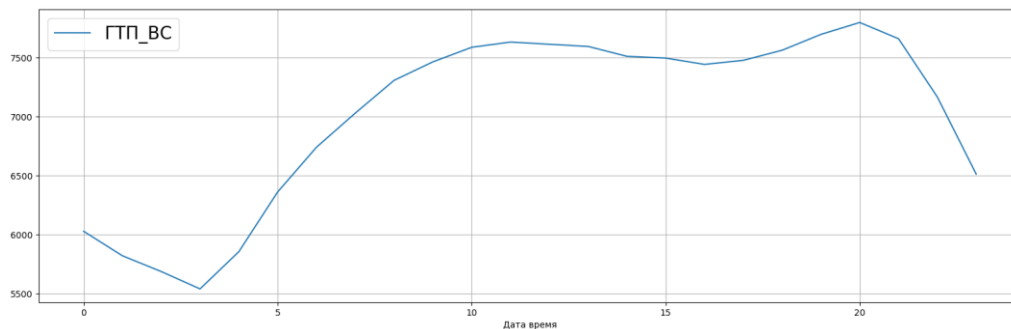
Проверка на выбросы, обнаружено не характерно поведение данных: технологический останов сооружений. Замену на среднее значение аналогичного года, месяца, часа.  
По ГТП\_ВО выброс не обнаружено

# Что получилось. EDA

Усредненный за час профиль потребления электрической энергии за период наблюдения по группам ГТП-Вс, ГТП-ВО

```
1 df_ee_eda.describe()
```

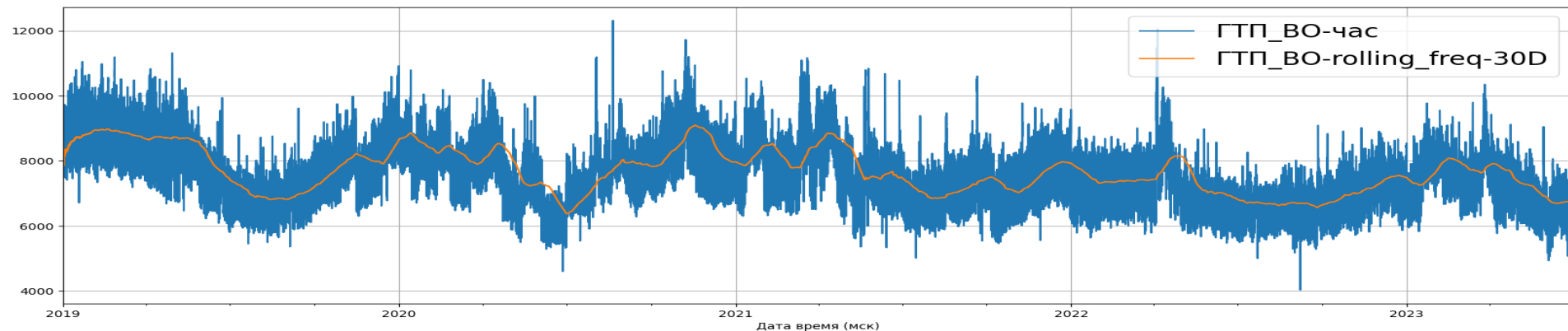
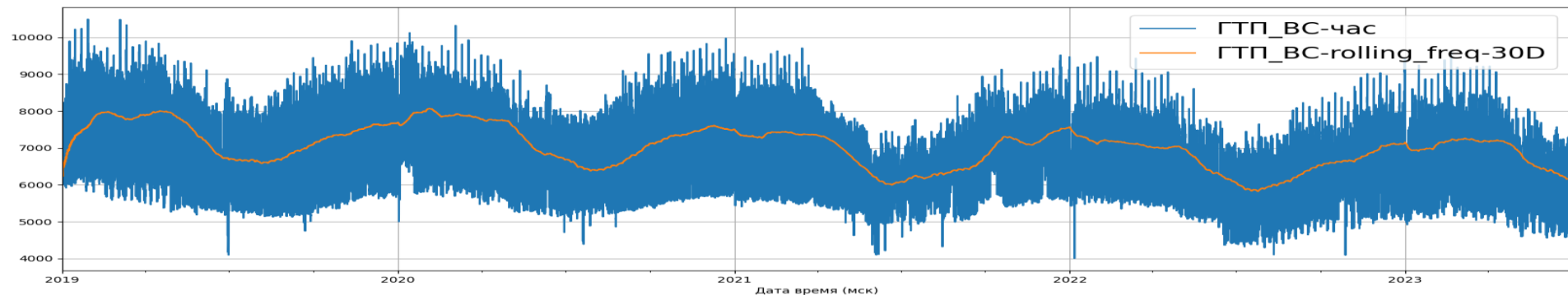
|       | ГТП_Вс       | ГТП_ВО       |
|-------|--------------|--------------|
| count | 39384.000000 | 39384.000000 |
| mean  | 7025.898740  | 7667.255423  |
| std   | 1067.284958  | 989.238861   |
| min   | 4000.000000  | 4030.000000  |
| 25%   | 6208.000000  | 6969.000000  |
| 50%   | 7004.000000  | 7612.000000  |
| 75%   | 7844.898000  | 8291.000000  |
| max   | 10493.112000 | 12312.840000 |





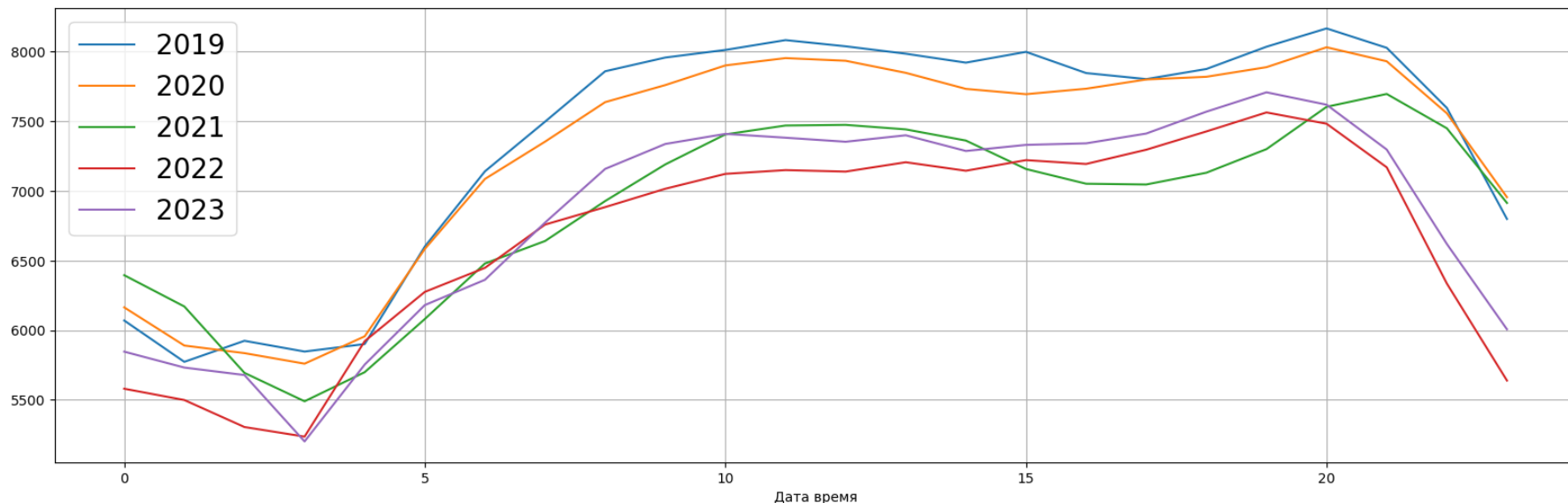
# Что получилось. «Чистые данные»

## Профиль потребления электроэнергии ГТП\_ВС, ГТП\_ВО



# Что получилось. «Чистые данные» ГТП\_ВС

## Усредненный профиль потребления электроэнергии ГТП\_ВС, по годам

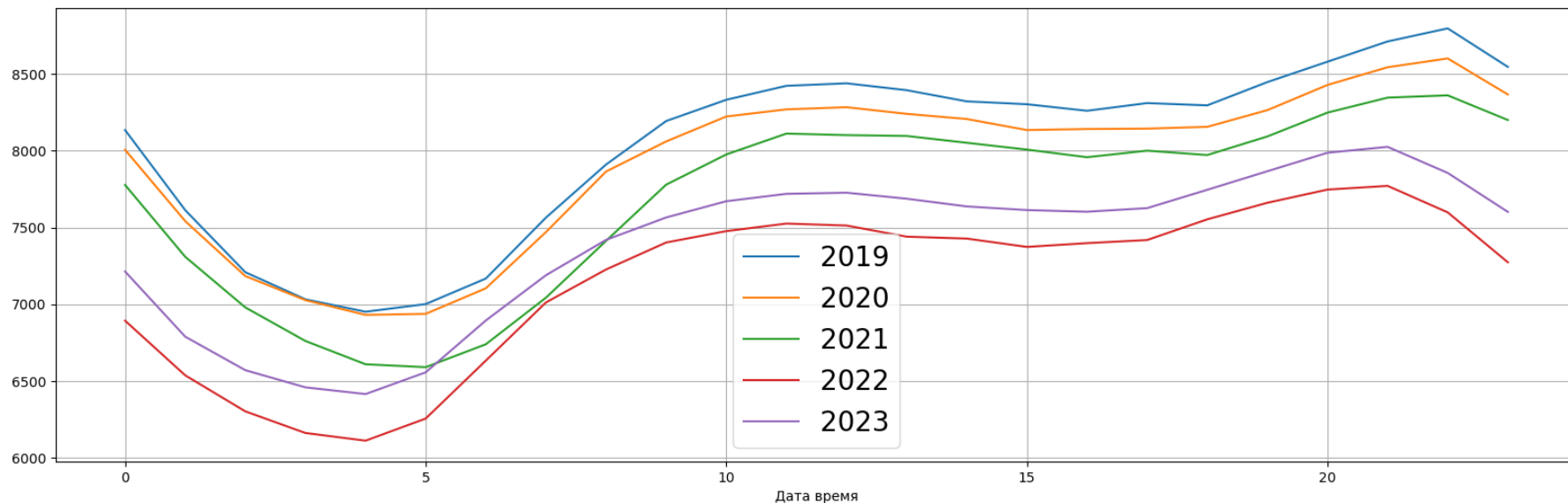


Наблюдается суточная неравномерность

Снижение объема подаваемой воды в сеть – снижение нагрузки

# Что получилось. «Чистые данные» ГТП\_ВО

## Усредненный профиль потребления электроэнергии ГТП\_ВО, по годам



Наблюдается суточная неравномерность

Снижение объема перекачки стоков – снижение нагрузки

## Что получилось. Моделирование

Исследование применения искусственных нейронных сетей для краткосрочного прогнозирования потребления электроэнергии предприятием;

Прогнозирую 24 значения, то за основу возьму RNN интервальное прогнозирование

Базовая модель - прогноз на сутки вперед равен аналогичному дню предыдущей недели

**Гипотеза: прогнозирование с помощью ИИ должно быть лучше, чем базовая модель**

Оценка подхода проводится по бизнес-метрике, БМ:

$$\text{БМ} = \frac{\sum_h^N \frac{|\Phi - \text{Откл\_abs}|}{\Pi}}{N}$$

Φ - фактический расход электрической энергии за час; Π - плановый расход электрической энергии за час; Откл\_abs = |Φ - Π| отклонение фактического расхода за час от планового расхода за час N - число плановых значений за месяц.

**Цель: БМ должна быть не более 4% за месяц.**

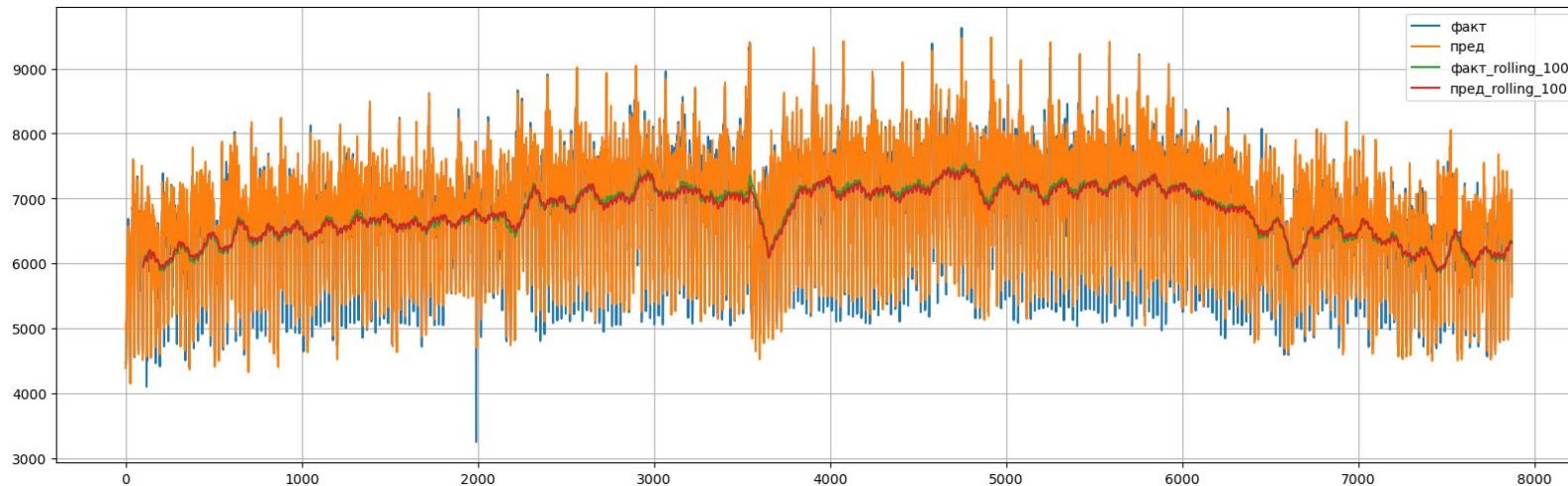
# Что получилось. Моделирование

Определение модели

```
class RNN(nn.Module):  
    def __init__(self, input_size, hidden_size, output_size, num_layers):  
        super(RNN, self).__init__()  
        self.hidden_size = hidden_size  
        self.num_layers = num_layers  
        self.rnn = nn.LSTM(input_size, hidden_size, num_layers, batch_first=True)  
        self.fc = nn.Linear(hidden_size, output_size)  
  
    def forward(self, x, h0, c0):  
        out, (hn, cn) = self.rnn(x, (h0, c0))  
        out = self.fc(out[:, -1, :])  
        return out
```

# Что получилось. Моделирование. График сравнения предсказанных и фактических значений

ГТП-ВС



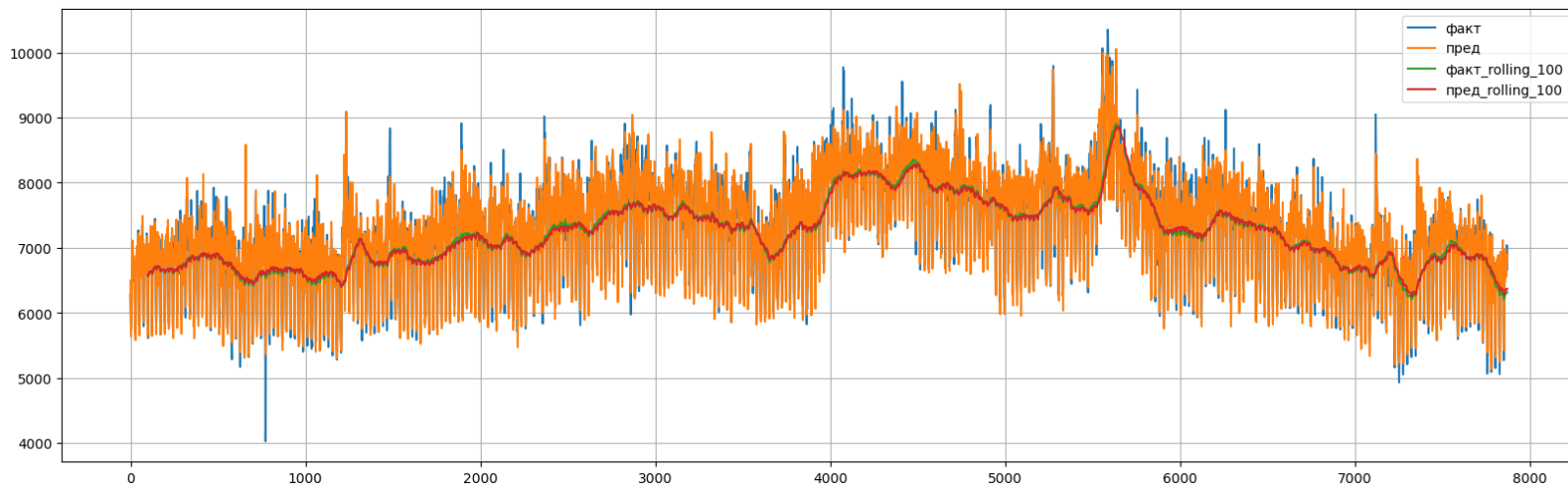
|    | ИИ-модель | base-модель | откл      |
|----|-----------|-------------|-----------|
| ✓1 | 0.036947  | 0.056499    | -0.019552 |
| ✓2 | 0.038417  | 0.038146    | 0.000271  |
| ✓3 | 0.039063  | 0.035011    | 0.004052  |
| 4  | 0.041340  | 0.030568    | 0.010772  |
| 5  | 0.041191  | 0.051436    | -0.010245 |
| 6  | 0.042017  | 0.063498    | -0.021481 |

## Вывод:

1. Для прогнозирования почасового потребления электрической энергии можно использовать искусственные нейросети
2. Для ГТП\_ВС на тесте ИИ-модель показала себя лучше в зимние и летние месяцы.

# Что получилось. Моделирование. График сравнения предсказанных и фактических значений

ГТП-ВО



|    | ИИ-модель | base-модель | откл      |
|----|-----------|-------------|-----------|
| ✓1 | 0.025748  | 0.062967    | -0.037220 |
| ✓2 | 0.023967  | 0.037215    | -0.013248 |
| ✓3 | 0.025988  | 0.061309    | -0.035321 |
| ✓4 | 0.027216  | 0.062264    | -0.035048 |
| ✓5 | 0.024656  | 0.041193    | -0.016537 |
| ✓6 | 0.025323  | 0.064045    | -0.038722 |

## Вывод:

1. Для прогнозирования почасового потребления электрической энергии можно использовать искусственные нейросети
2. Для ГТП\_ВО на тесте ИИ-модель показала себя лучше

# Выводы и планы по развитию

1. Цели достигнуты
2. Для прогнозирования почасового потребления электрической энергии можно использовать искусственные нейросети
3. Применение трансформеров для задачи



# Спасибо за внимание!