

# **Применение методов машинного обучения для прогнозирования сигналов СМНР**

Бутин К.П.\* , Попов А.И.\* , Угрюмов И.А.\* , Родионов А.В.\*

\* ООО «Инженерный центр «Энергосервис», Архангельск, Россия

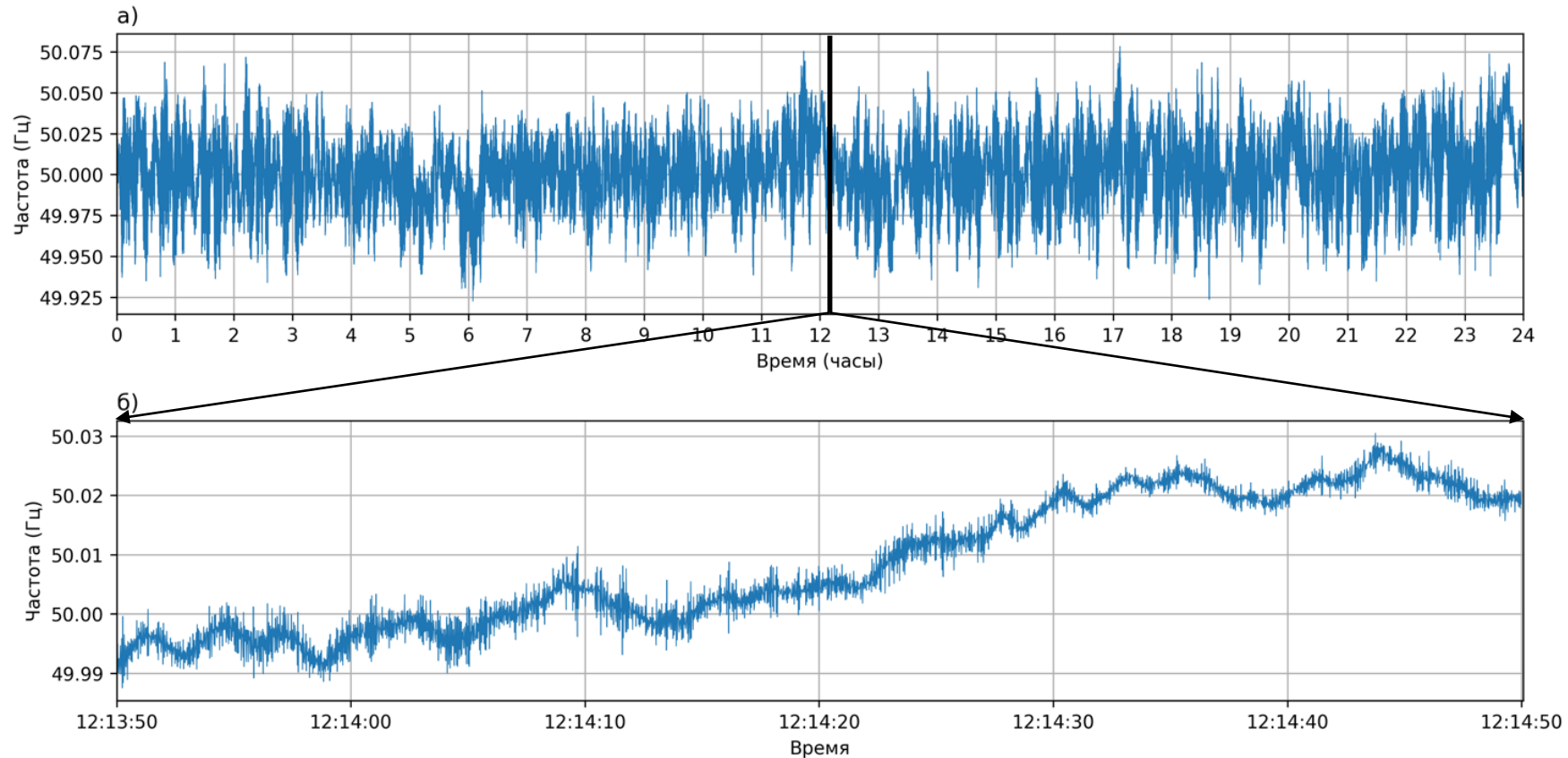
# Введение

- Развитие и распространение технологии синхронизированных векторных измерений (СВИ)
- Большой объем данных
- Трудности при построении и актуализации адекватных математических моделей реальных энергосистем
- Активное внедрение методов машинного обучения (МО) в различные прикладные сферы
- Одним из важных приложений МО является прогнозирование временных рядов
- В докладе предлагается провести оценку степени обусловленности сигналов своими предыдущими значениями

# Прогнозирование по данным СВИ

- Направления исследований:
  - **Прогнозирование временных рядов**
  - Прогнозирование состояний
- **Обнаружение событий и аномалий** в массивах данных СМПР
- **Восстановление** пропущенных и коррекция искаженных **измерений**
- Разнообразие методов прогнозирования:
  - Без МО (метод ближайших соседей, ARIMA, продолжение гармоник и др.)
  - С МО (линейная регрессия; **нейронные сети**: полносвязные, свёрточные, LSTM, трансформеры и др.; метод опорных векторов; решающие деревья и др.)
- Требуется обеспечить расчет в **режиме реального времени**

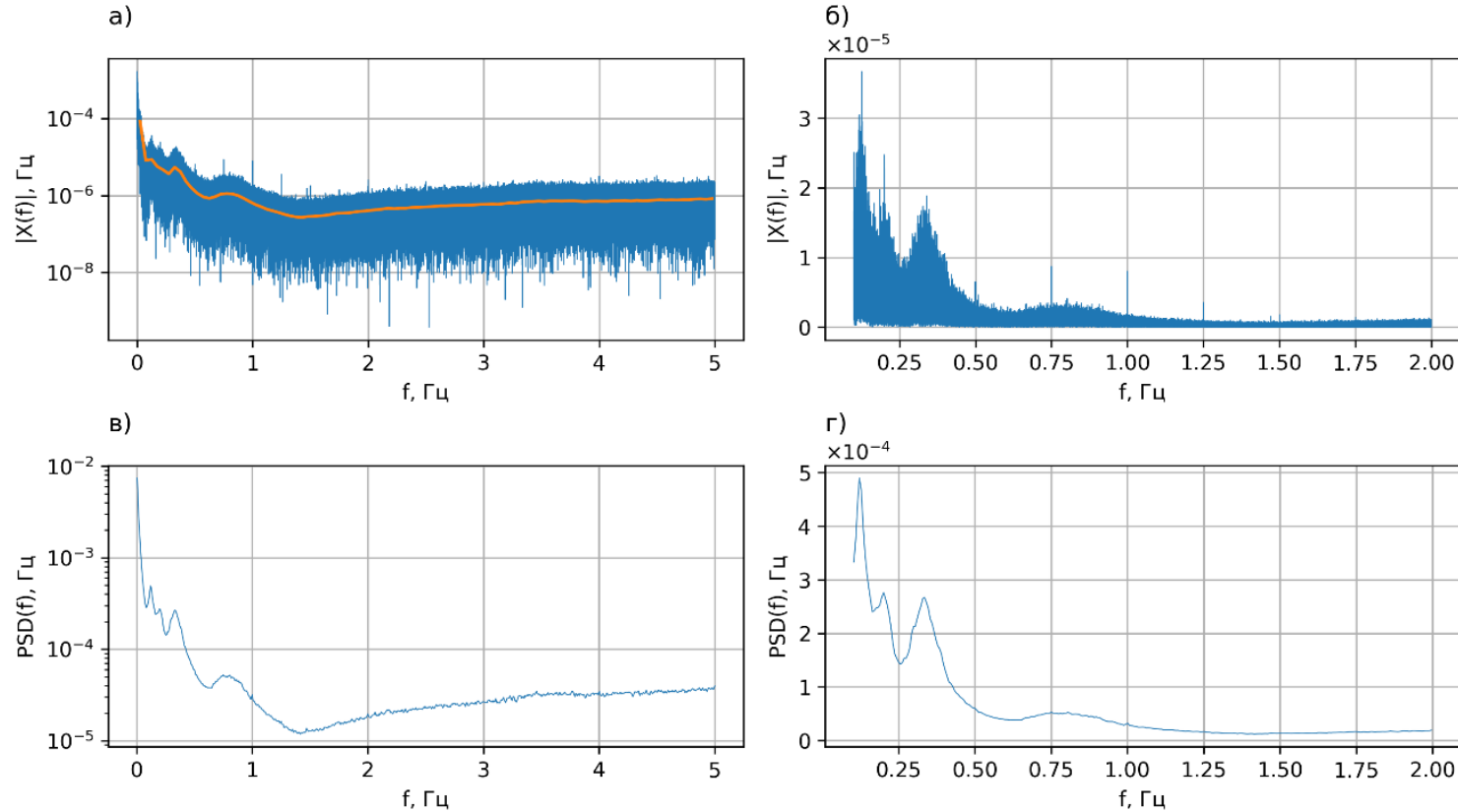
# Исходные данные



Измеренный сигнал частоты с присоединения ЛЭП ЕЭС РФ 500 кВ

а) за сутки б) 1 минута; Нормальный режим работы; Данные без пропусков и аномалий

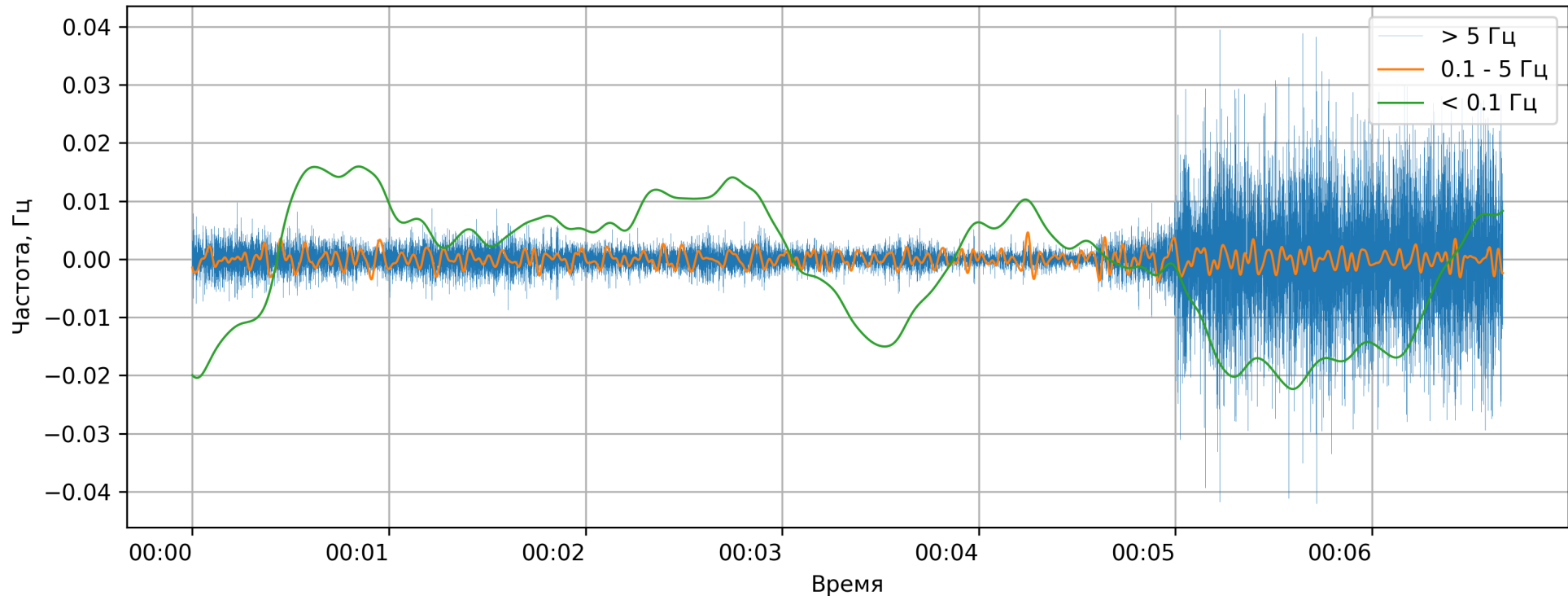
# Спектральное представление



Амплитудный спектр на участке а) от 0 до 5 Гц б) от 0.1 до 2 Гц  
Спектральная плотность мощности на участке в) от 0 до 5 Гц г) от 0.1 до 2 Гц. Метод Уэлча  
а, в – логарифмическая шкала; б, г – линейная шкала

# Разложение исходного сигнала по частотным интервалам\*

\* Фильтр Баттерворта 4 порядка



- $< 0.1$  Гц – область сверхнизких частот, в основном тренд
- $0.1 - 5$  Гц – область низкочастотных колебаний (НЧК)
- $> 5$  Гц – субсинхронные колебания (**в данной работе не рассматривается**)

# Подготовка данных

- Эксперименты с прогнозированием проводятся для трёх частотных интервалов:
  - от 0 до 5 Гц (**тренд + НЧК**)
  - от 0 до 0.1 Гц (преимущественно тренд)
  - от 0.1 до 5 Гц (преимущественно НЧК).
- Для ускорения расчётов после **фильтрации** выполняется **децимация** сигнала:

$$d = \left\lfloor \frac{r}{2f_c} \right\rfloor,$$

где  $f_c$  - верхняя частота полосы пропускания;  $r$  – частота дискретизации исходного сигнала

- Дальность прогноза в секундах выражается следующей формулой:

$$t_p = \frac{dn}{r} = \frac{n}{2f_c},$$

где  $n$  – дальность прогноза, выражаемая в количестве сэмплов

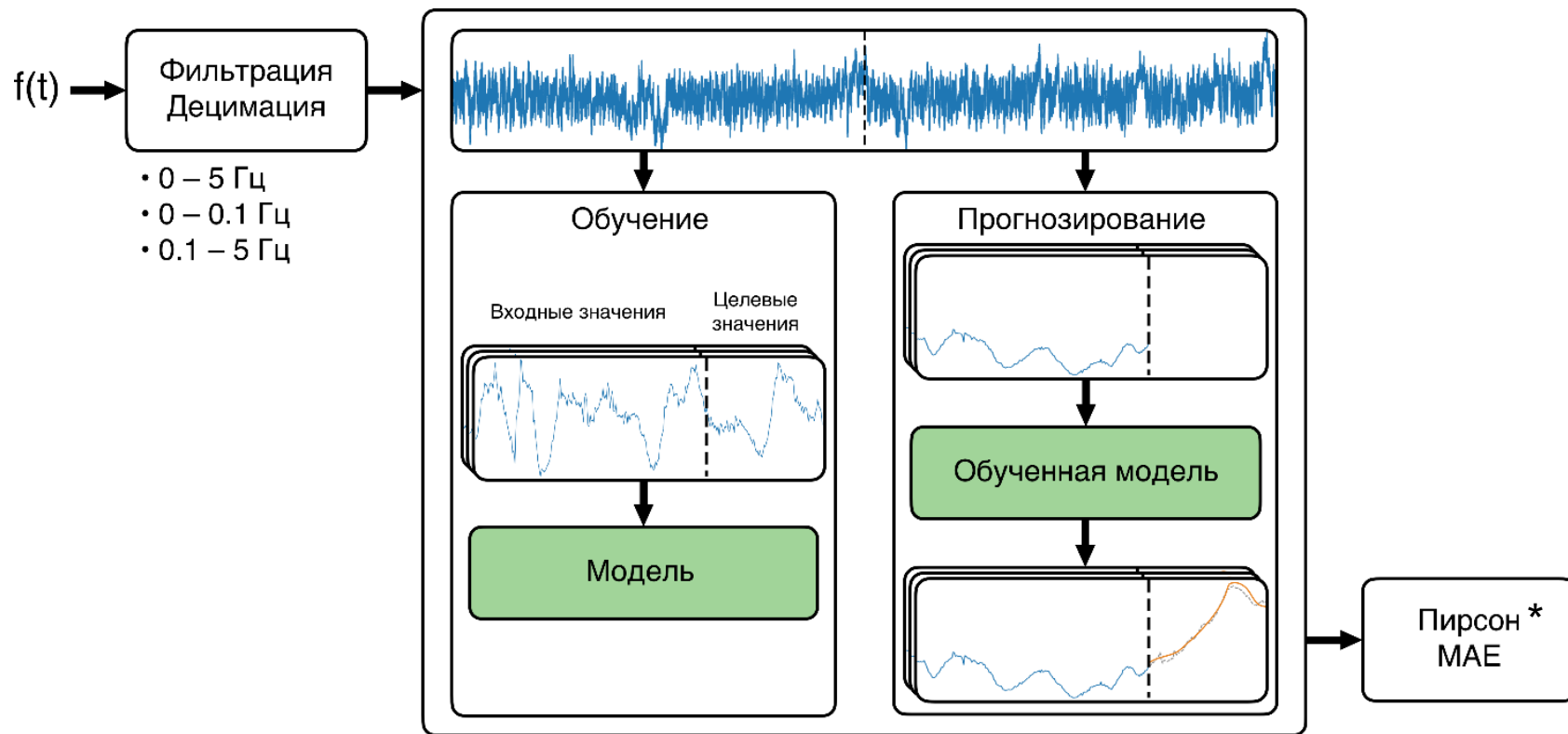
# Методы прогнозирования

| Обозначение |          | Краткая характеристика  |  |
|-------------|----------|---|--|
| Nominal     |          | В качестве прогнозного значения выдаётся 50 Гц.   |  |
| Prev        |          | В качестве прогнозного значения выдаётся значение последнего известного сэмпла.           |  |
| Average     |          | В качестве прогнозного значения выдаётся среднее арифметическое значение на окне анализа. |  |
| Harm        |          | Выдаётся сумма максимальных гармоник, выявленных на окне анализа.                         |  |
| Методы МО*  |          |   |  |
| LR          |          | Линейная регрессия.   |  |
| нейросети   | FCNN     | Полносвязная нейронная сеть (один скрытый слой 128 нейронов).                             | Оптимизатор: Adam  |
|             | LSTM     | Рекуррентная нейросеть со 128 нейронами и одним полносвязным слоем.                       | Функция потерь: MSE                                      |
|             | LSTM-ENC | Seq2Seq модель с двумя слоями LSTM по 64 нейрона.   | Коэффициент скорости обучения: 0.001                     |
|             | BiLSTM   | Двунаправленный LSTM со 128 нейронами.  | Валидационная метрика: MAE<br>Валидационная выборка: 10% |

\* К **МО** относятся методы, в которых модель **на основе данных** (примеров) оптимизирует свои параметры посредством минимизации функции потерь (**обучается**), с явным разделением на обучающую и тестовую выборки и **итеративным обновлением весов**.



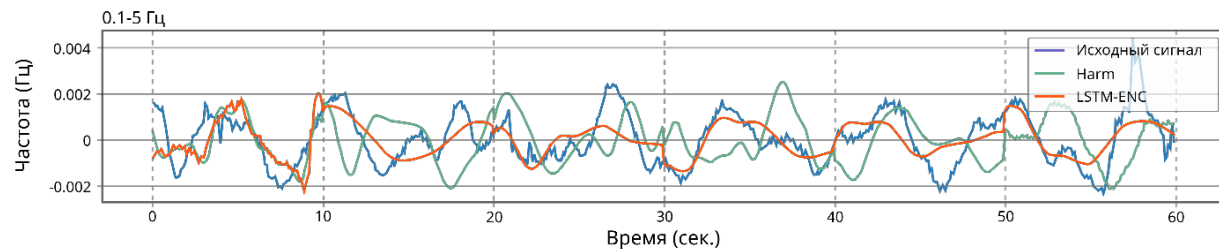
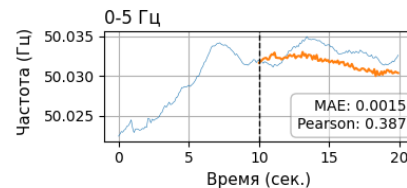
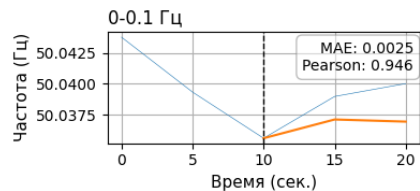
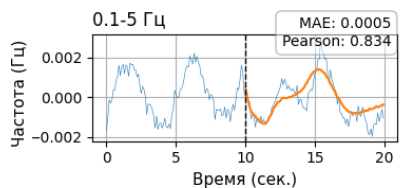
# Схема обработки данных



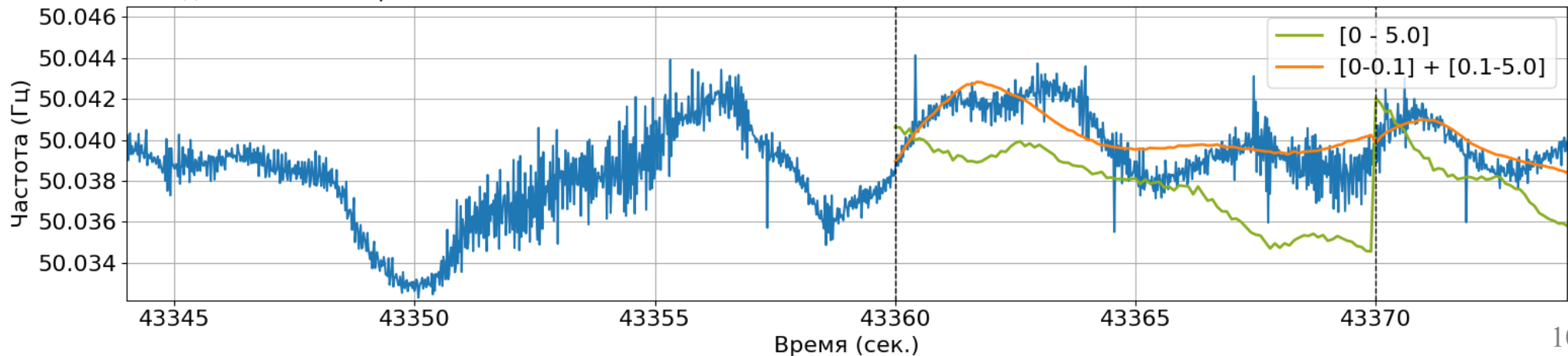
$$* r = \frac{cov(x, y)}{\sigma_x \sigma_y}$$

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n}$$

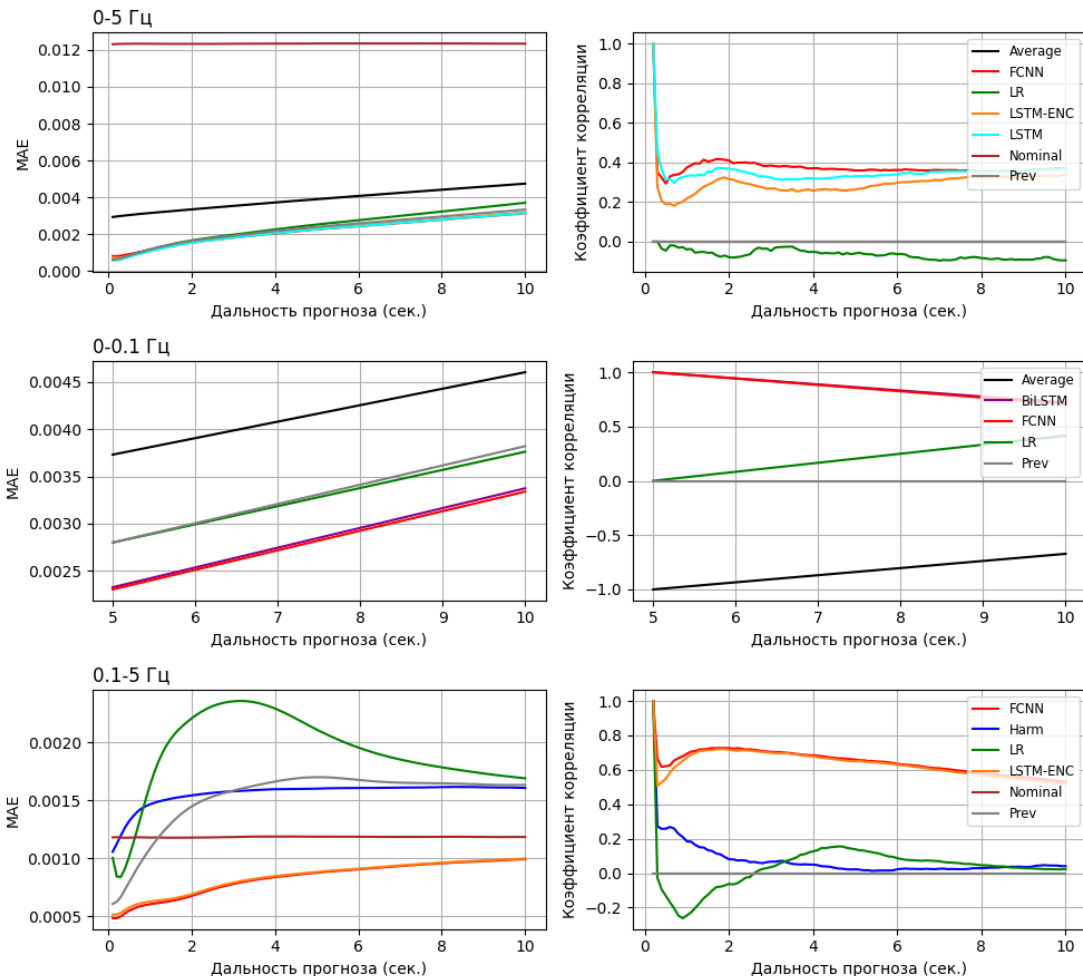
# Примеры спрогнозированных данных



Исходный сигнал и прогноз LSTM



# Зависимость метрик от дальности прогноза



- От 0 до 5 Гц: один из лучших результатов по метрике MAE является метод **Prev**. Но он **не дает никакой информации о динамике сигнала**
- **МО** дает **высокие значения корреляции** при прогнозировании отдельно **в области низкочастотных колебаний и тренда**
- Классические методы выдают низкие показатели по всем метрикам в области НЧК
- Один из лучших результатов показала FCNN – полносвязная нейросеть

# Заключение

- Современные тенденции развития электроэнергетических систем, **нарастание объёмов и повышение качества измерительных данных** о параметрах электрического режима, а также значительные **успехи в области искусственного интеллекта** ставят вопрос о поиске **эффективных применений методов МО** при обработке данных СМПР
- Выполнена **апробация** ряда методов прогнозирования временных рядов **на реальных записях СМПР ЕЭС РФ**
- Проведённые эксперименты позволяют выдвинуть гипотезу о том, что **простые нейронные сети** на интервалах до нескольких секунд **лучше распознают динамическую структуру сигналов**, чем методы продолжения сигналов с окна анализа на окно прогноза
- К возможным **практическим применениям** полученных результатов можно отнести построение на основе нейронных сетей **быстрых** вычислительных процедур **поиска аномалий и событий**

# Спасибо за внимание!

Кирилл Павлович Бутин

ООО «Инженерный центр «Энергосервис», Архангельск, Россия

[k.butin@ens.ru](mailto:k.butin@ens.ru),