

# Новые способы обработки временных рядов на примере анализа влияния эпидемиологической обстановки на электроэнергетику в России

Павел Литвинов, РТСофт  
[litvinov\\_pv@rtsoft.ru](mailto:litvinov_pv@rtsoft.ru)



**cigre**  
SC D2

Ежегодный семинар НИК D2 РНК СИГРЭ  
Июль 2020 г., Москва

# Содержание

1. **Контекст:** AI и Data Science в энергетике. А-Платформа.
2. **Предмет:** Анализ временных рядов является краеугольным камнем решения большого класса задач:
  - от наблюдения состояния системы и управления технологическими процессами
  - до прогнозирования и долгосрочного планирования.
3. **Смена парадигмы:** АСУ все чаще рассматриваются как киберфизические системы, предъявляют новые требования к способам обработки временных рядов.
4. **Методология:** Новизна предлагаемой технологии в архитектуре, основанной на использовании только открытого ПО; подготовке, преобразовании и обработке данных исключительно в оперативной памяти; наглядных и информативных способах визуализации



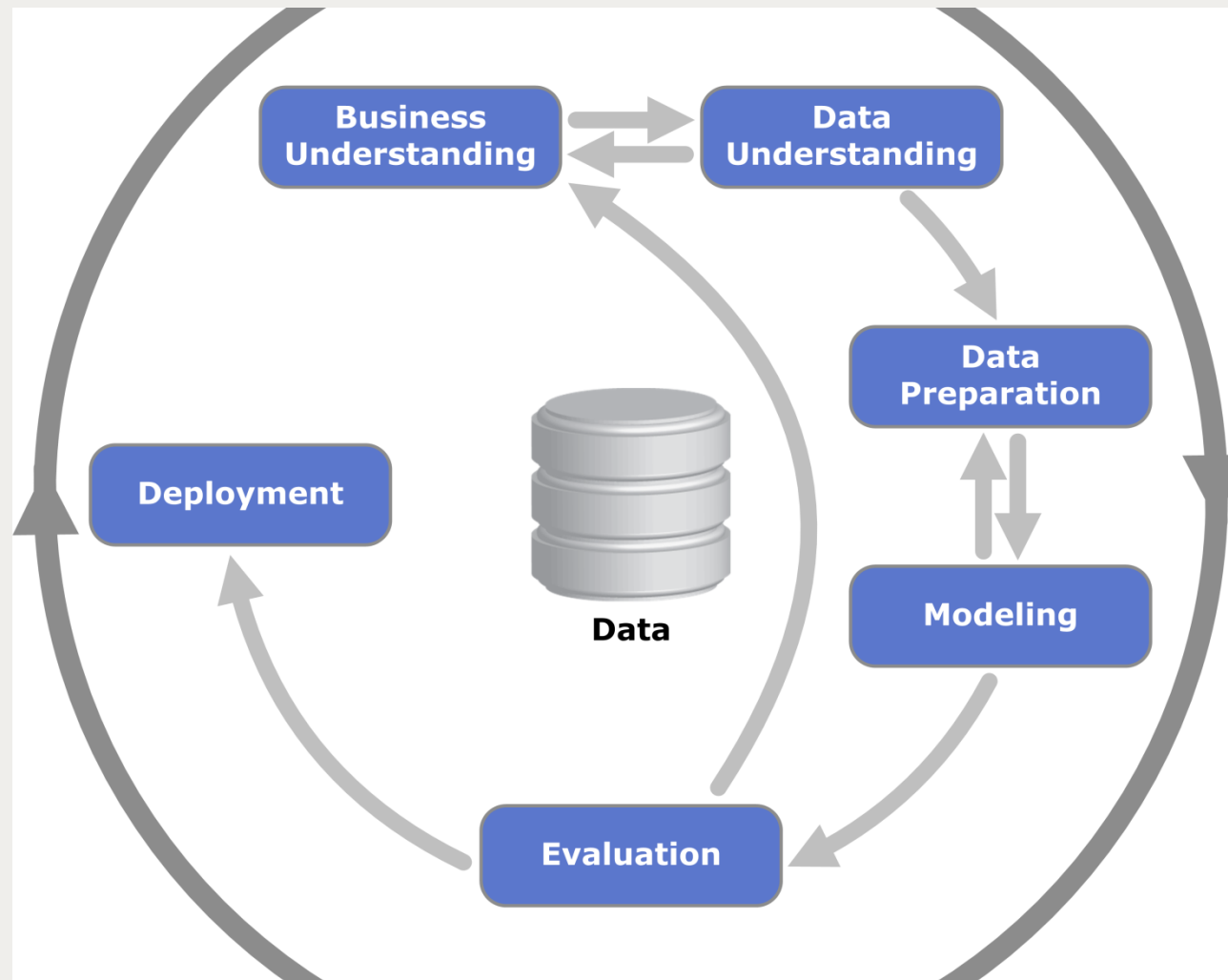
## CRISP-DM

### Cross-Industry Standard Process for Data Mining

1. Осмысление бизнеса (Business understanding).
2. Осмысление данных (Data understanding).
3. Подготовка данных (Data preparation).
4. Моделирование (Modeling).
5. Оценка результатов (Evaluation).
6. Внедрение (Deployment).

#### + Контроль

Оставляем место для Fail Fast



Data-Driven подход в организации требует наличия: **DevOps, DataOps, ModelOps** => А-Платформа

## Качество данных

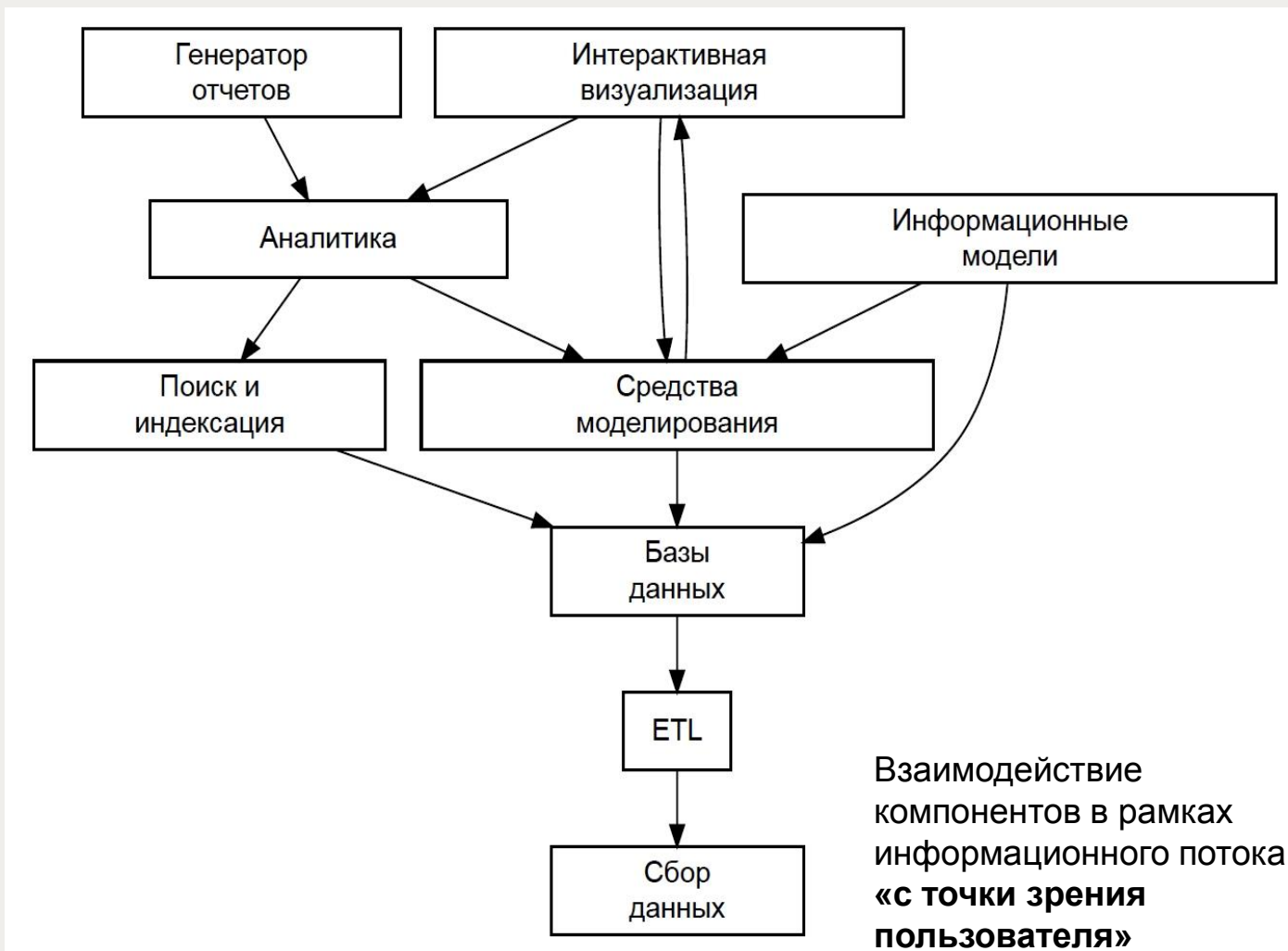
- Качество данных, содержащихся во временных рядах, непосредственно определяет возможность и полезность их применения для решения практических задач
- Хорошо известны основные факторы снижения качества:
  - «шум» (посторонние значения)
  - пропущенные значения
  - дублирующиеся значения
  - аномальные значения и противоречия
  - значения, не соответствующие формату
- Есть известная шутка: «Данные бывают плохие и очень плохие : )»
- Лучшая практика – коды качества, худшая – «сжатие» данных.

Парадигма больших данных (BigData) «Собирайте и храните столько данных, сколько возможно, потому что сможете извлечь из них пользу»,



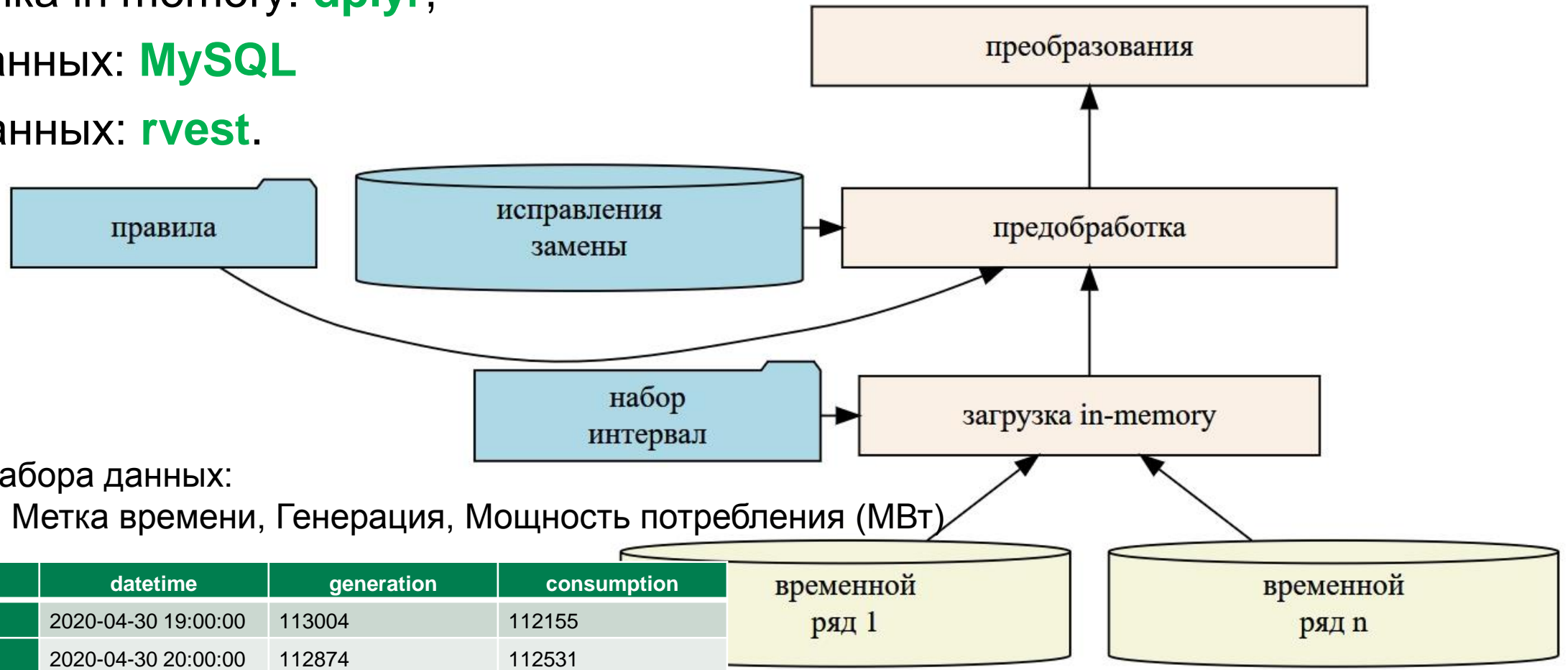
# Архитектура аналитического приложения (вариант)

Компонент	Возможные реализации
Интерактивная визуализация	Панель (dashboard), композиция (mashup), диаграммы, таблицы, инфографика, текстовая аналитика
Генератор отчетов	Отчеты: автоматические (по-расписанию), ситуационные (по-событиям), ad hoc (по-запросу)
Аналитика	Системы поддержки принятия решений (СППР, англ. DSS); аналитическая обработка (OLAP), анализ данных (Data Mining)
Информационные модели	Обогащение общей информационной модели электроэнергетики (Common Information Model, CIM), цифровые двойники
Поиск и индексация	Elasticsearch
Средства моделирования	AnyLogic, R, Python; JupyterLab, Open Source
Базы данных	Реляционные БД, нереляционные БД, файловые хранилища, распределенная файловая система HDFS (для больших данных), специализированные БД класса Time-series databases (TSD)
(Extract, Transform, Load – ETL)	Адекватные источникам и базам данных решения
Сбор данных	Существующие протоколы передачи данных



# Технология обработки данных

- Генератор отчетов: **Jupyter Notebook, R Markdown**;
- Визуализация: **ggplot2**;
- Аналитика in memory: **dplyr**;
- Базы данных: **MySQL**
- Сбор данных: **rvest**.



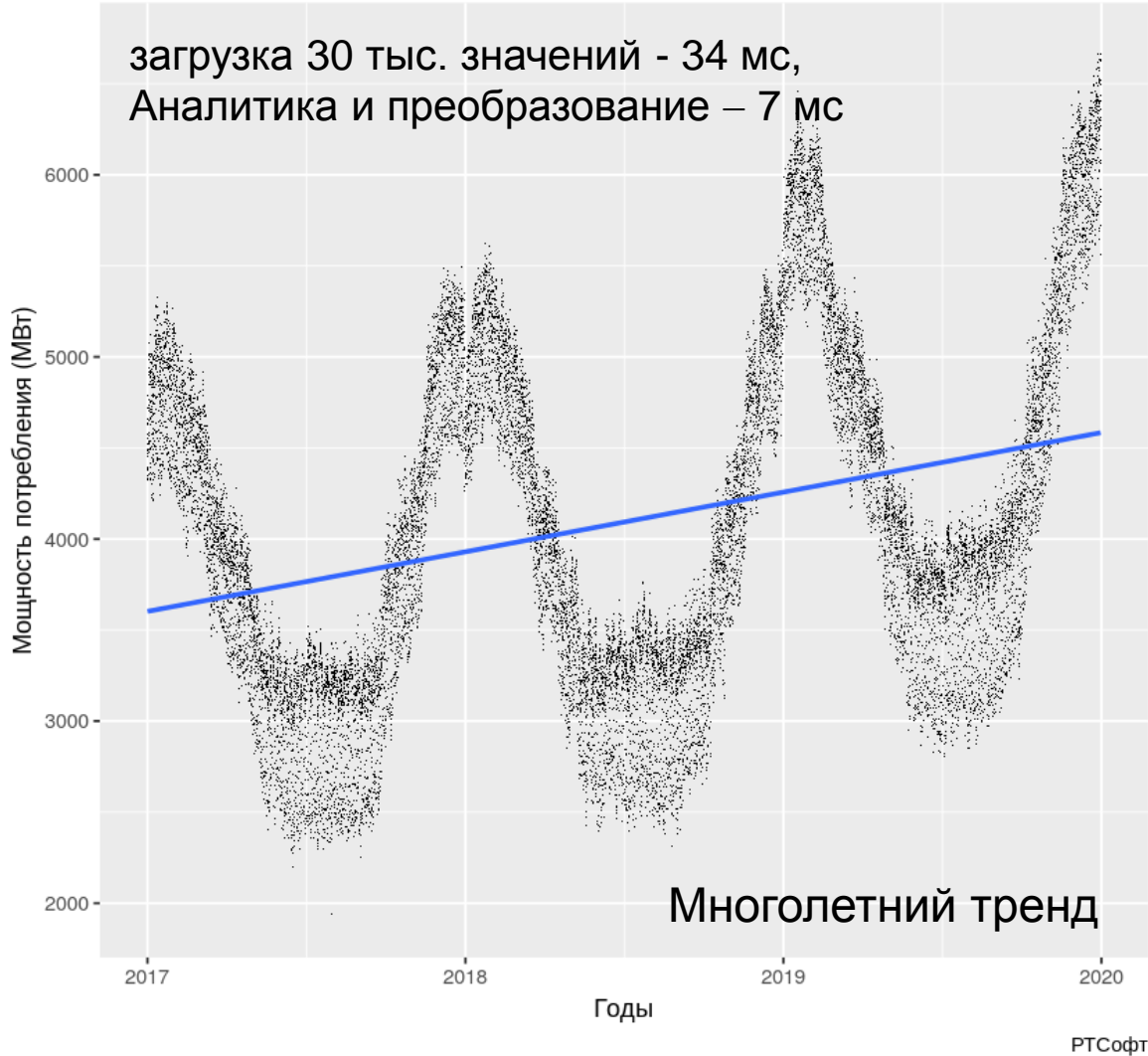
Фрагмент набора данных:

Код записи, Метка времени, Генерация, Мощность потребления (МВт)

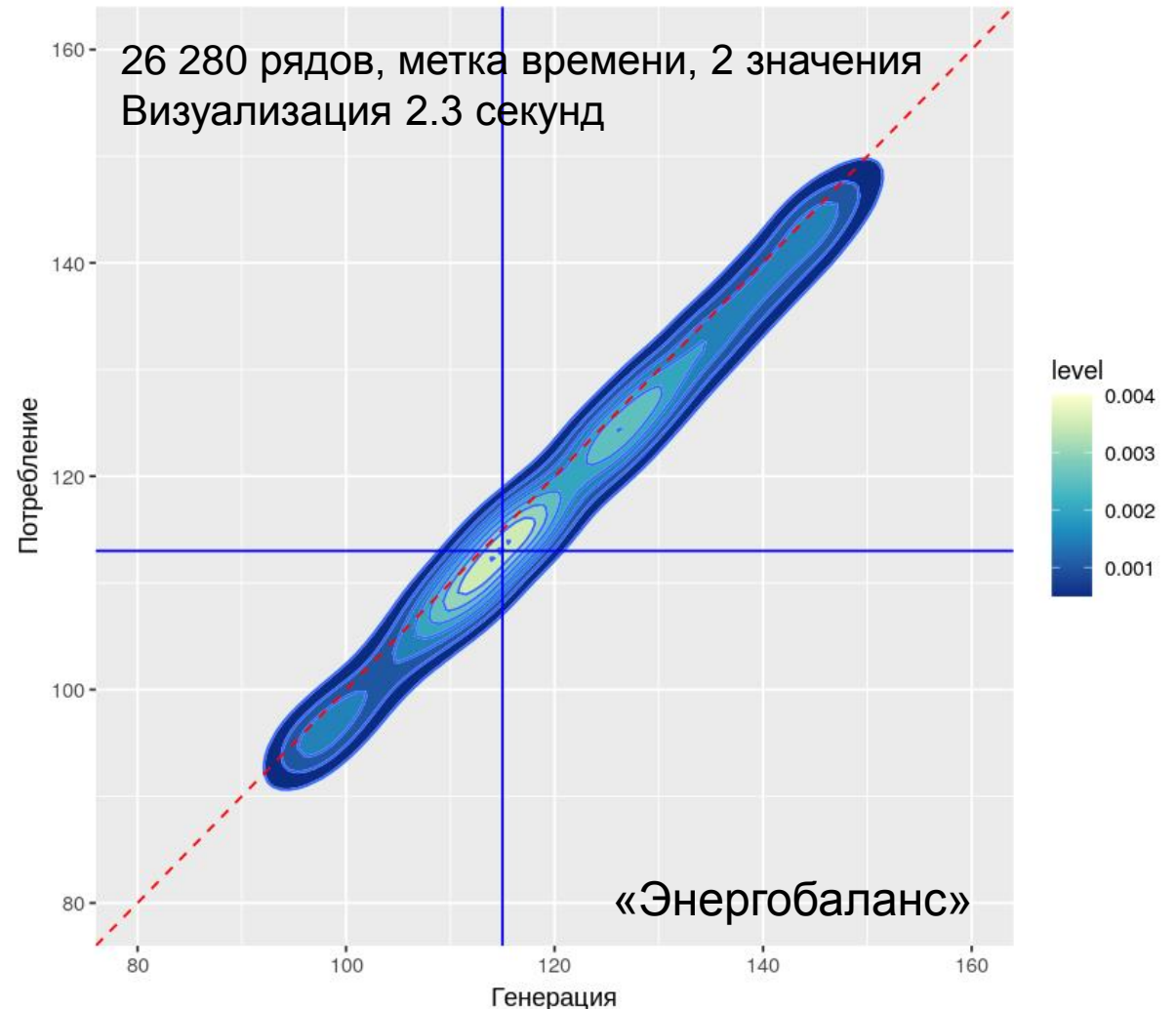
id	datetime	generation	consumption
29180	2020-04-30 19:00:00	113004	112155
29181	2020-04-30 20:00:00	112874	112531

# Показатели производительности

Почасовое потребление ОЭС Востока  
(Осцилограмма + линейная интерполяция)

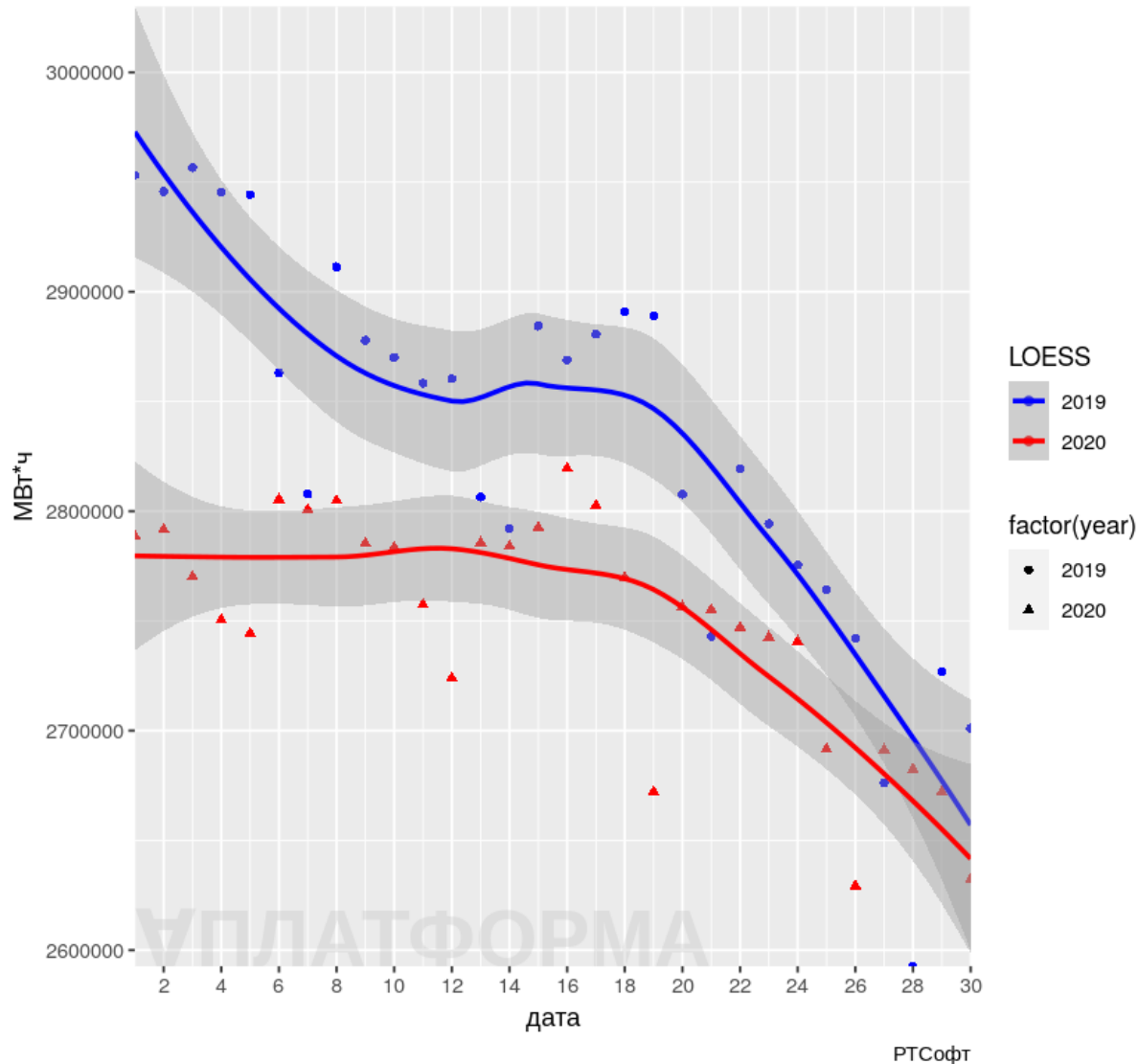


Генерация & Потребление (ГВт)  
(Температурная карта набора данных)



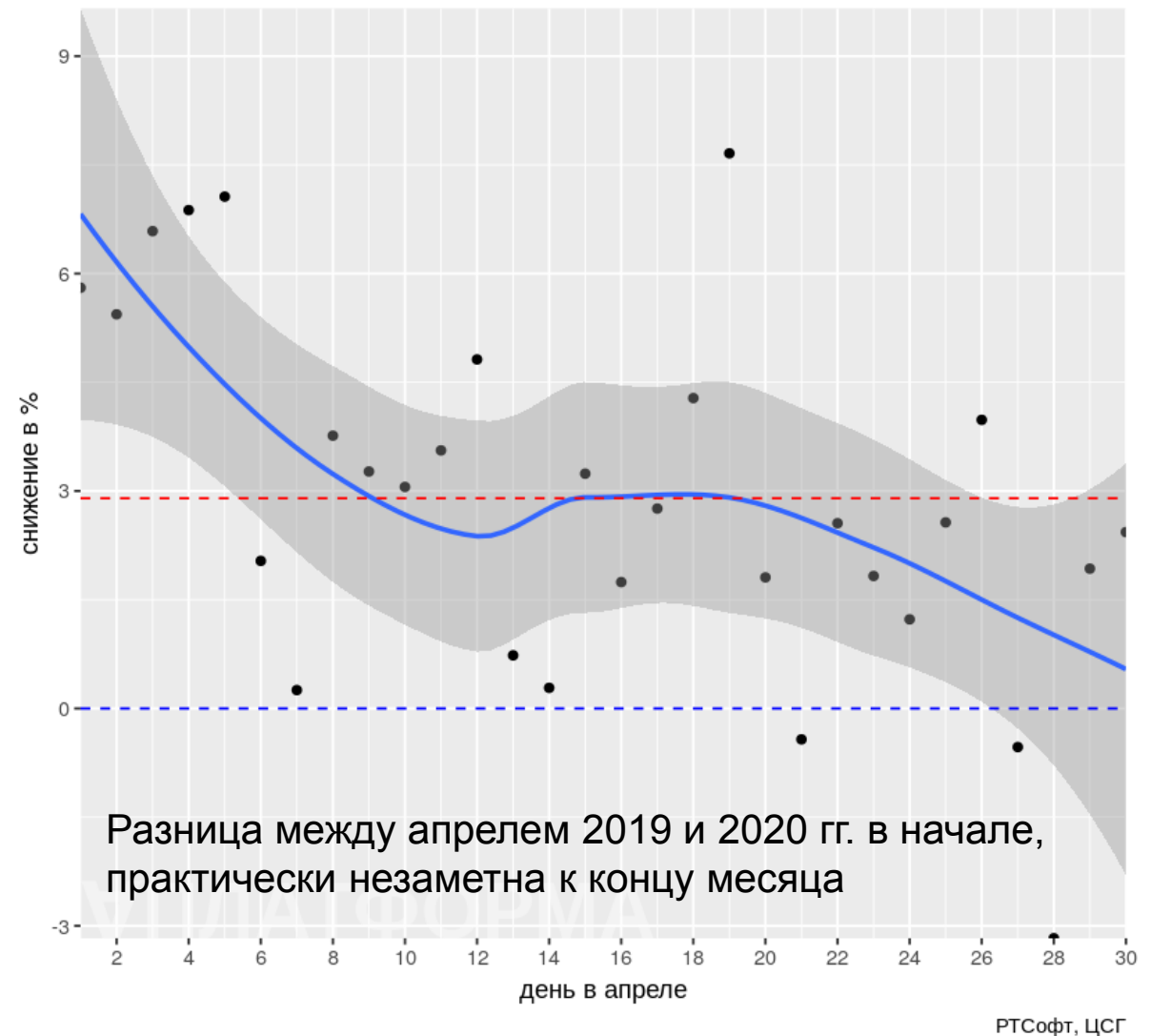
# Влияние карантинных мероприятий

Суточное потребление ЭЭ в ЕЭС  
(апрель месяц)



[1]

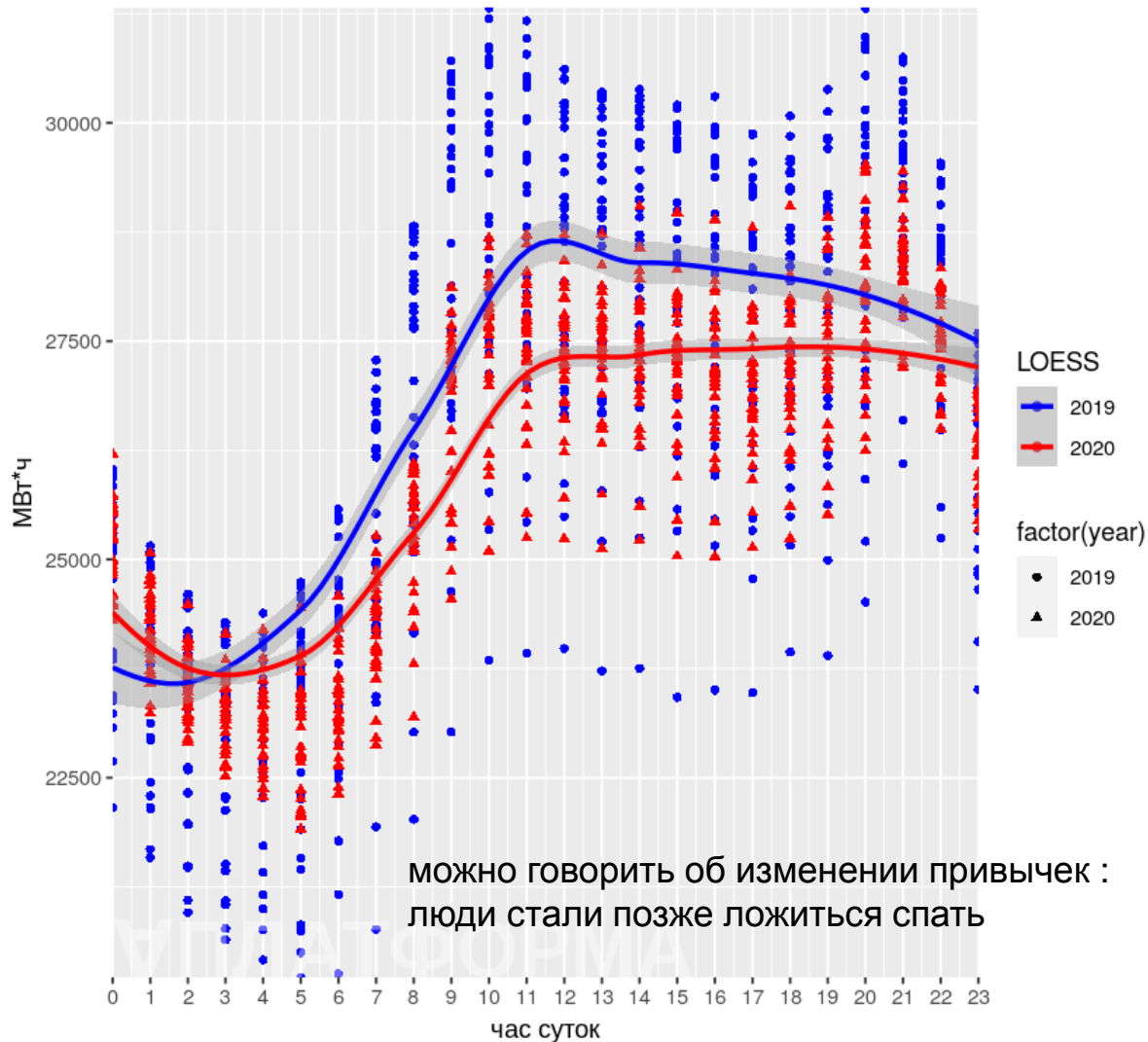
Влияние карантина на потребление ЭЭ в России  
(сравнение апреля 2019 и 2020 )



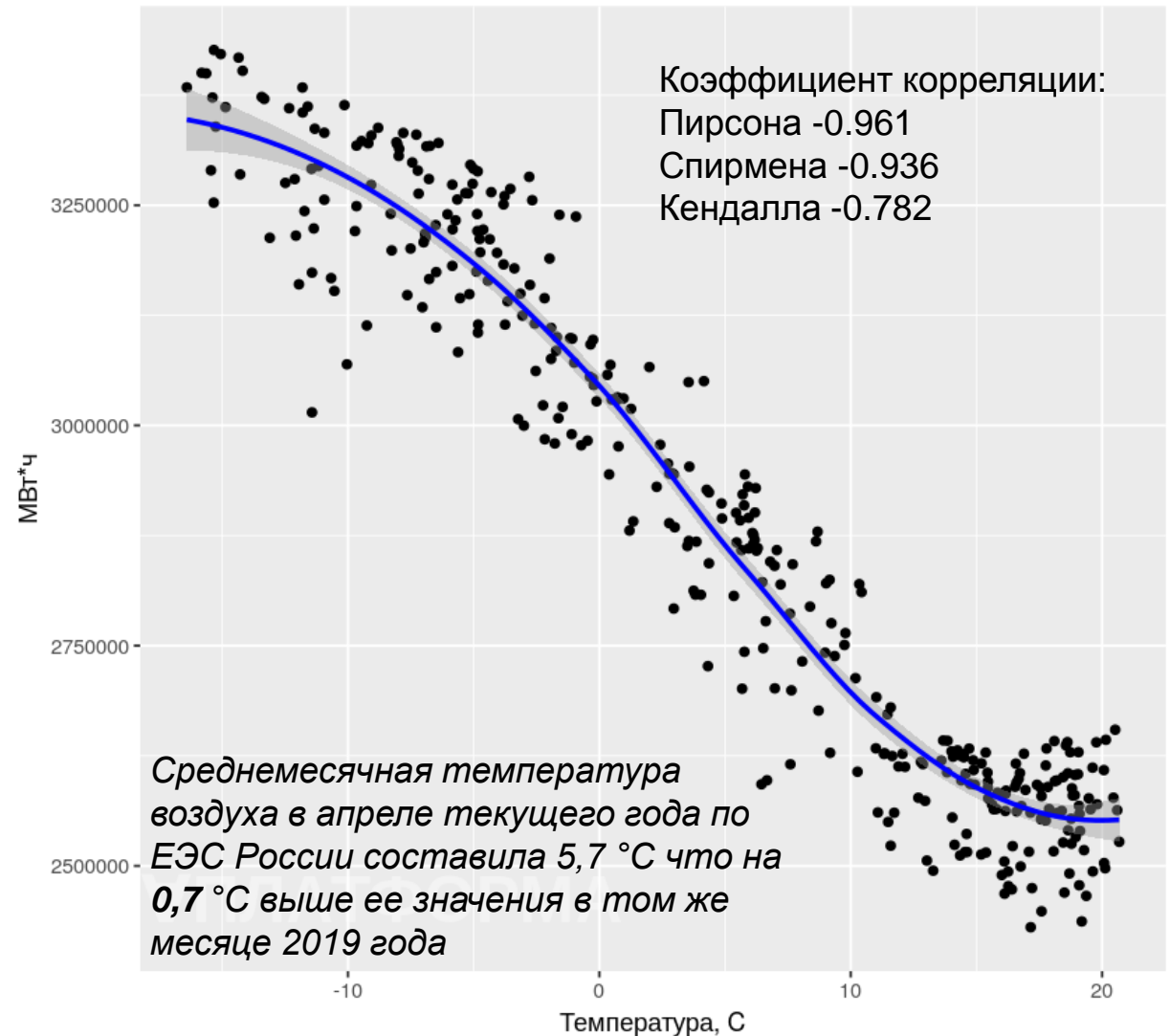


# Суточный профиль потребления

Суточный профиль потребления ЭЭ в ОЭС Центра  
(апрель месяц)

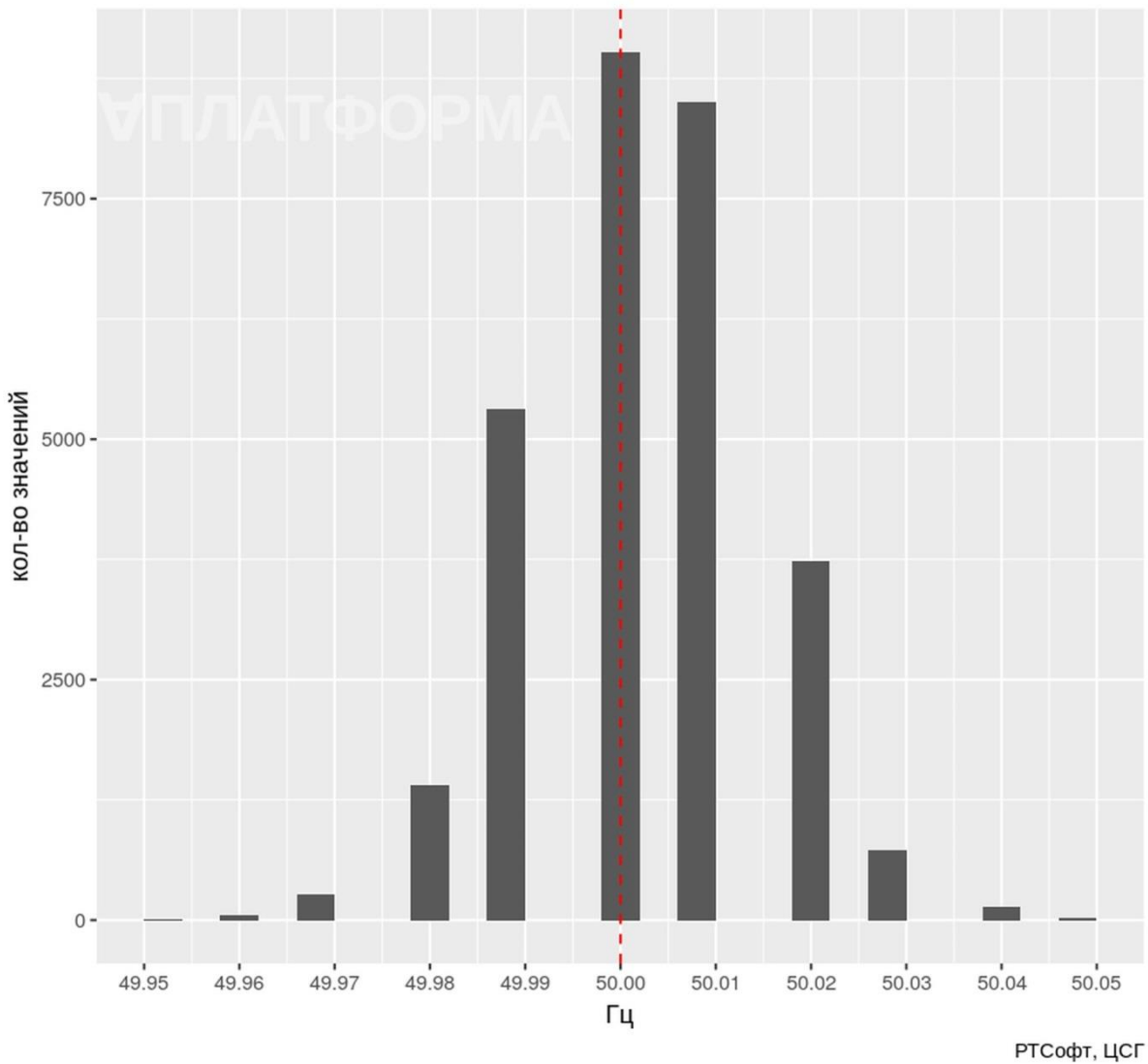


Зависимость суточного потребления от температуры  
(средняя температура воздуха в ЕЭС, 2019 г.)

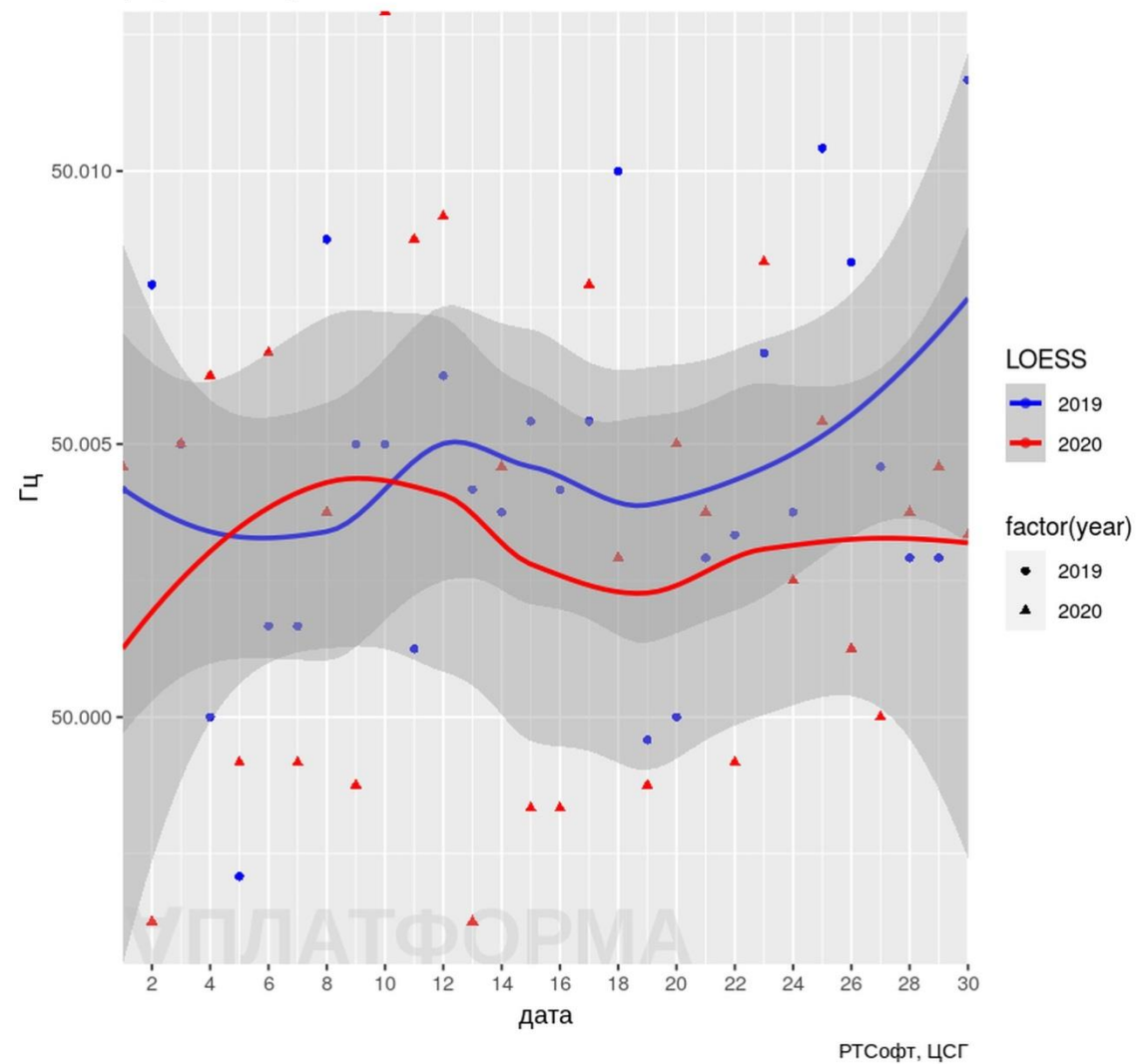


# Влияния карантина на частоту в ЕЭС России

Частота в ЕЭС России  
(с 2017-01-01 по 2020-04-30)



Частота в ЕЭС России  
(апрель месяц)



## Заключение

- Карантинные мероприятия или не явились вызовом для системы управления ЕЭС России, или энергетики с ними успешно справились. Настолько успешно, что не удалось найти даже следы существенных отклонений.
- В результате был продемонстрирован подход к решению аналитических задач на базе перспективной российской программной платформы управления распределенной энергетикой.
- Оценена возможность использования свободного программного обеспечения.
- Подтверждена возможность и целесообразность использования преобразований в оперативной памяти для ускорения вычислений

## Список литературы

- ГОСТ Р ИСО 9000-2015. Национальный стандарт Российской Федерации. Системы менеджмента качества. Основные положения и словарь"
- Указ Президента Российской Федерации от 9 мая 2017 г. № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017 - 2030 годы».
- ЦМАКП: «Энергопотребление в ЦФО падает на 10%, деловая активность – на 20%», <https://peretok.ru/news/distribution/21895/>
- Макоклюев Б.И. Анализ и планирование электропотребления. М.: Энергоатомиздат, 2008
- Пресс-релиз, "Потребление электроэнергии в ЕЭС России в апреле 2020 года уменьшилось на 2,9 % по сравнению с апрелем 2019 года", СО ЕЭС, 07.05.2020
- СТО 59012820.27.100.003-2012 "Регулирование частоты и перетоков активной мощности в ЕЭС России. Нормы и требования", СО ЕЭС

# Спасибо за внимание!



**sigre**

For power system expertise

АО РТСофт

Тел: +7 (495) 967-15-05

Факс: +7 (495) 742-68-29

E-mail: [microgrid@rtsoft.ru](mailto:microgrid@rtsoft.ru)

**Центральный офис:**

г. Москва, ул. Никитинская, д. 3

**Инженерный дом:**

г. Москва, ул. Верхняя Первомайская, д. 51