

ПРОБЛЕМА ФИЛЬТРАЦИИ ВАРИАЦИЙ ПЭС

БШФФ-2019, г. Иркутск , 16-21 сентября 2019 г.



Малецкий Б.М.
Ясюкевич Ю.В.
Веснин А.М.



Раньше

Детрендрование



Селекция
Вариаций

Полосовой
фильтр
скользящего
среднего

Предлагается

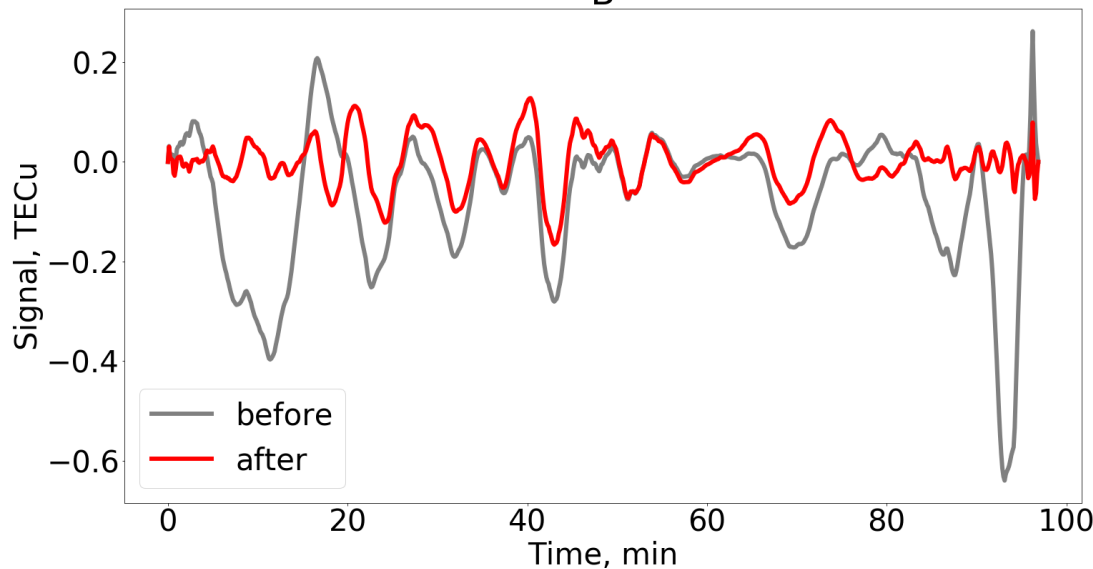
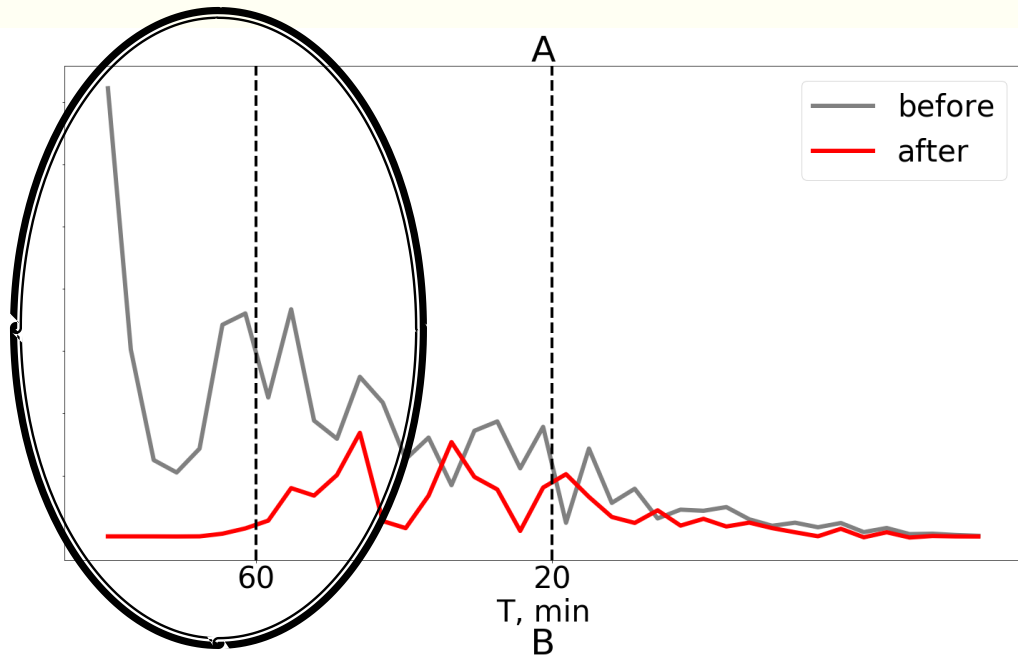
Детрендрование

Свой метод

Селекция
Вариаций

Свой метод

Пример артефактов



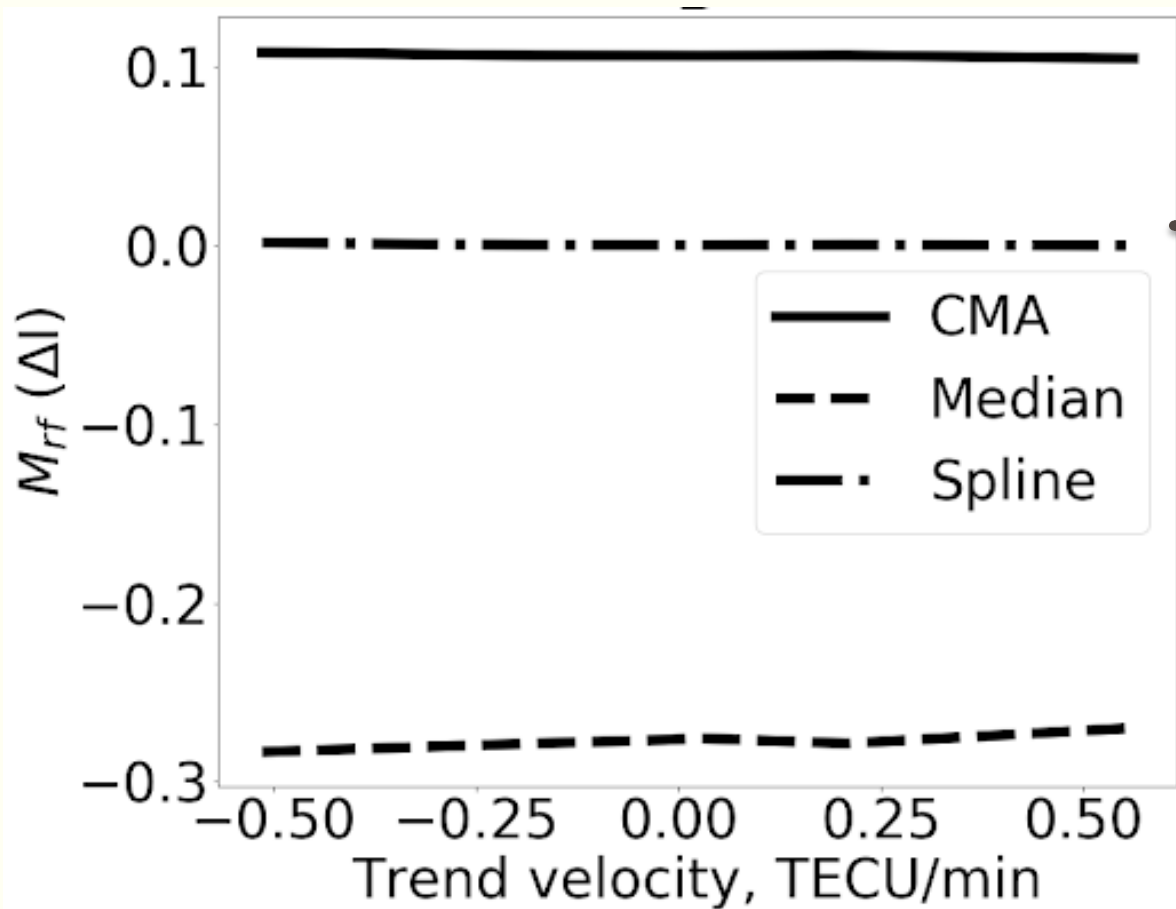
Пример артефактов
для диапазона 20-60
минут на основе
экспериментальных
данных

А – для частотной
области

В – для временной
области

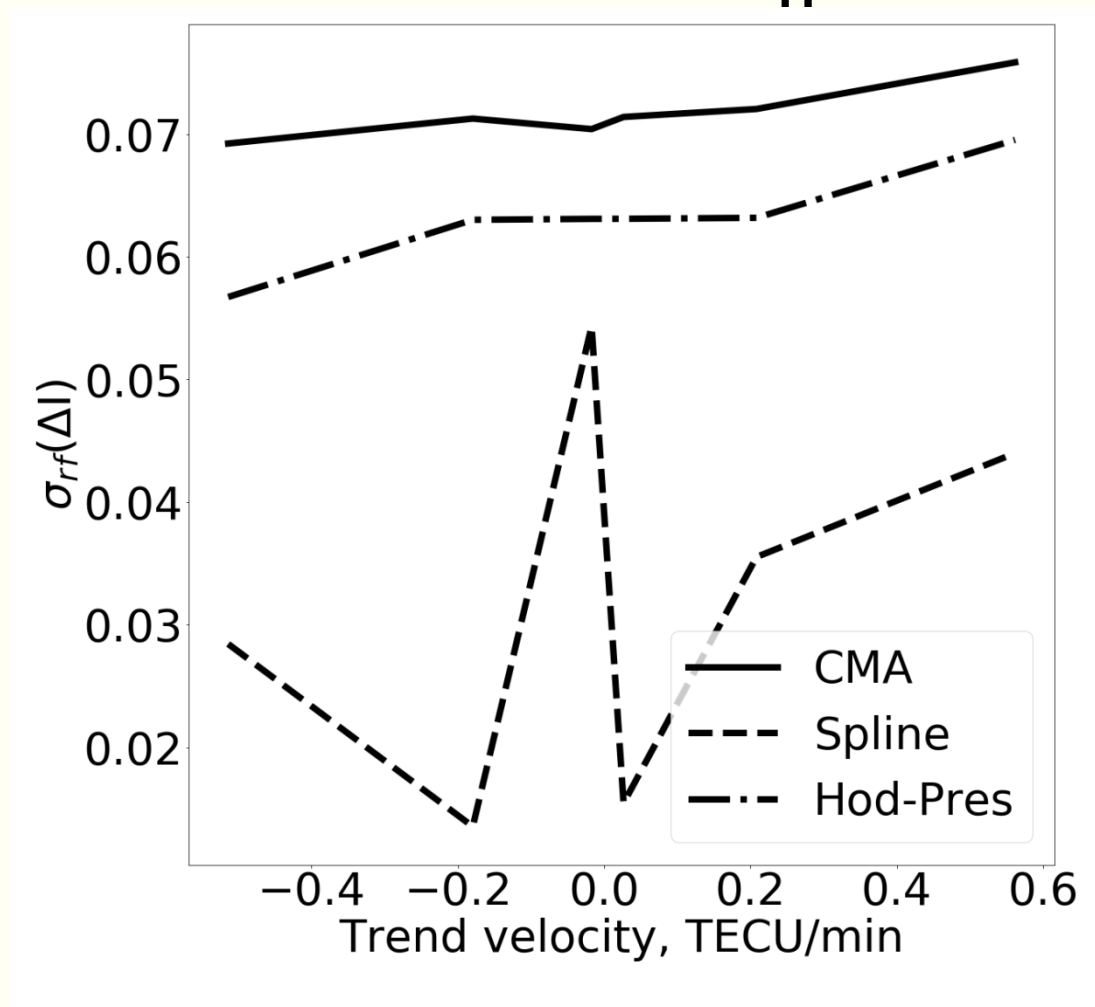
УДАЛЕНИЕ ТРЕНДА

Скорость тренда vs. Математическое ожидание ошибки m_{rf}

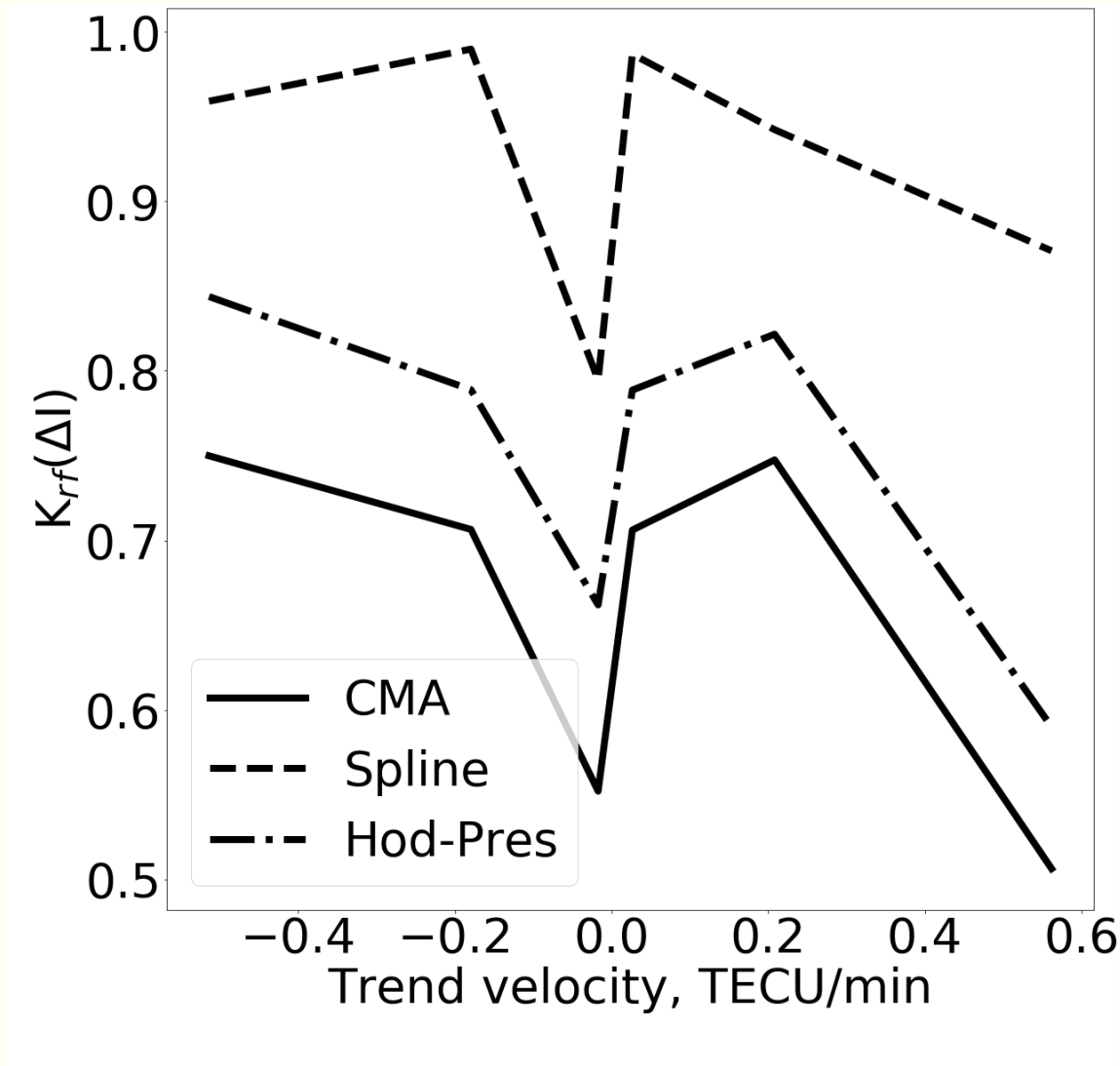


- Полином 6-ой степени
- Фильтр Ходрика-Прескотта
- Фильтр *l1*
- Кубический сглаживающий сплайн.
- Двойное применение скользящего среднего

Скорость тренда vs. Среднеквадратическое отклонение ошибки σ_{rf}

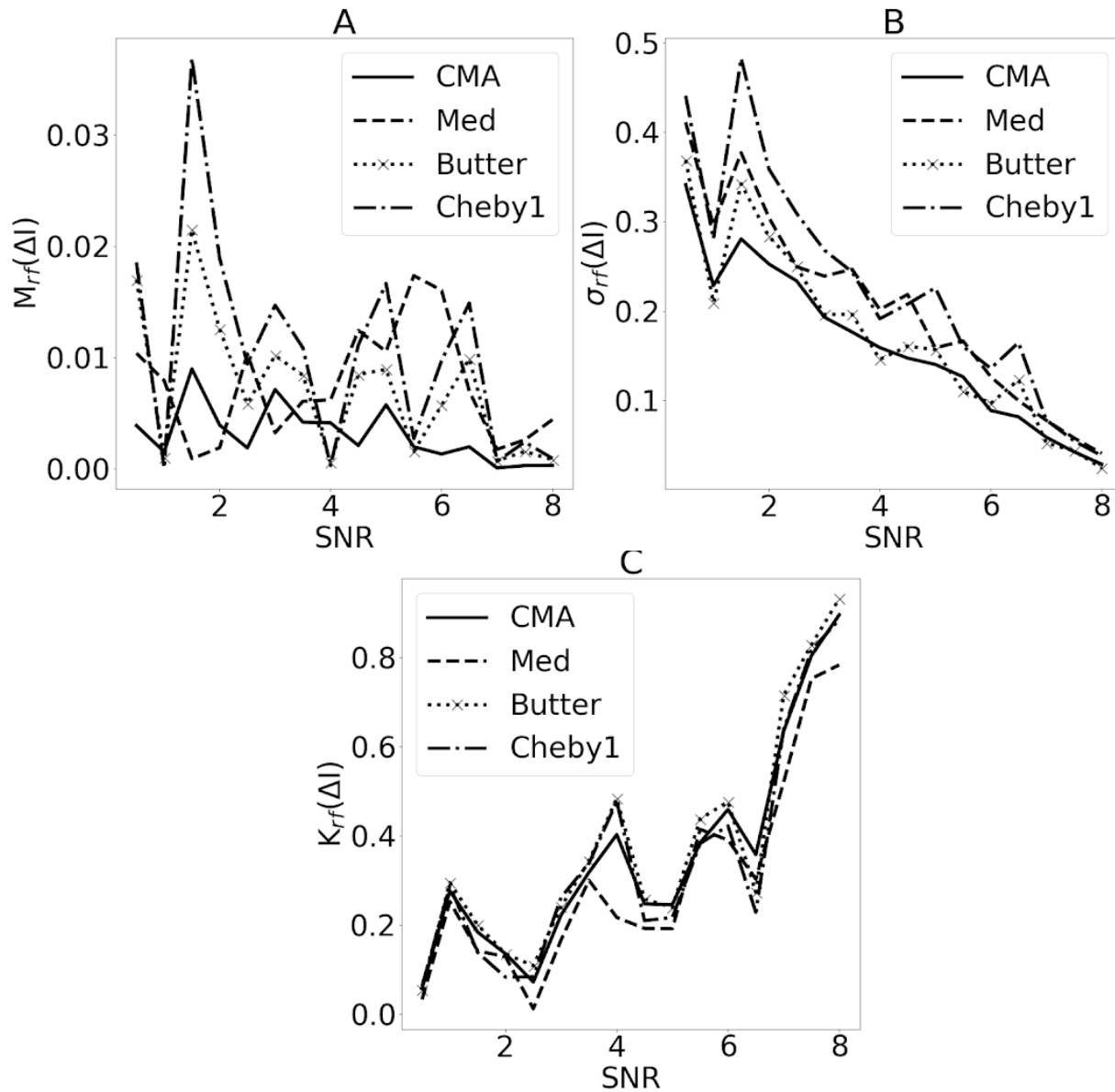


Скорость тренда vs. Коэффициент корреляции



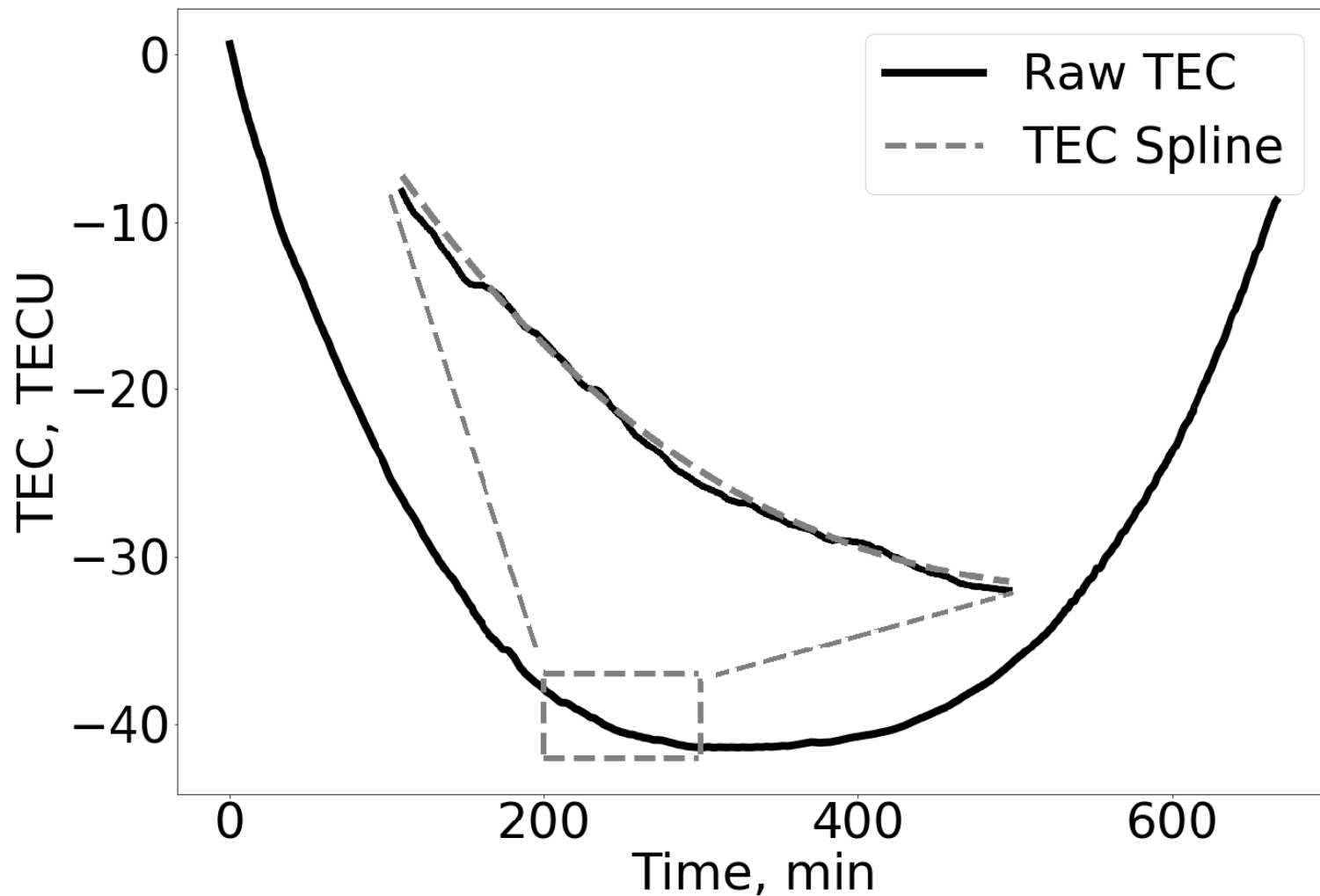
СЕЛЕКЦИЯ ВАРИАЦИЙ

Качество селекции

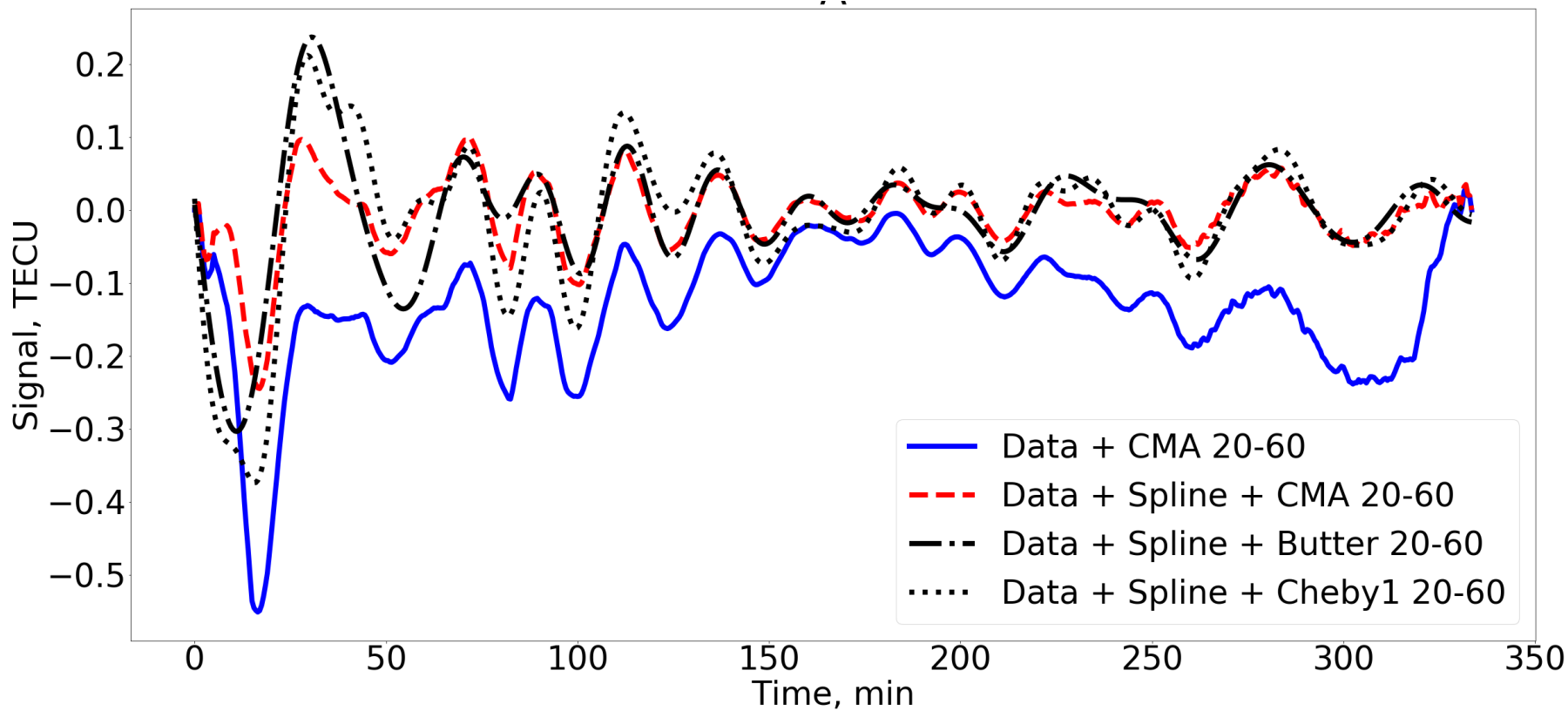


ЭКСПЕРИМЕНТАЛЬНЫЕ ДАННЫЕ

Экспериментальный ряд ПЭС



Вариации ПЭС

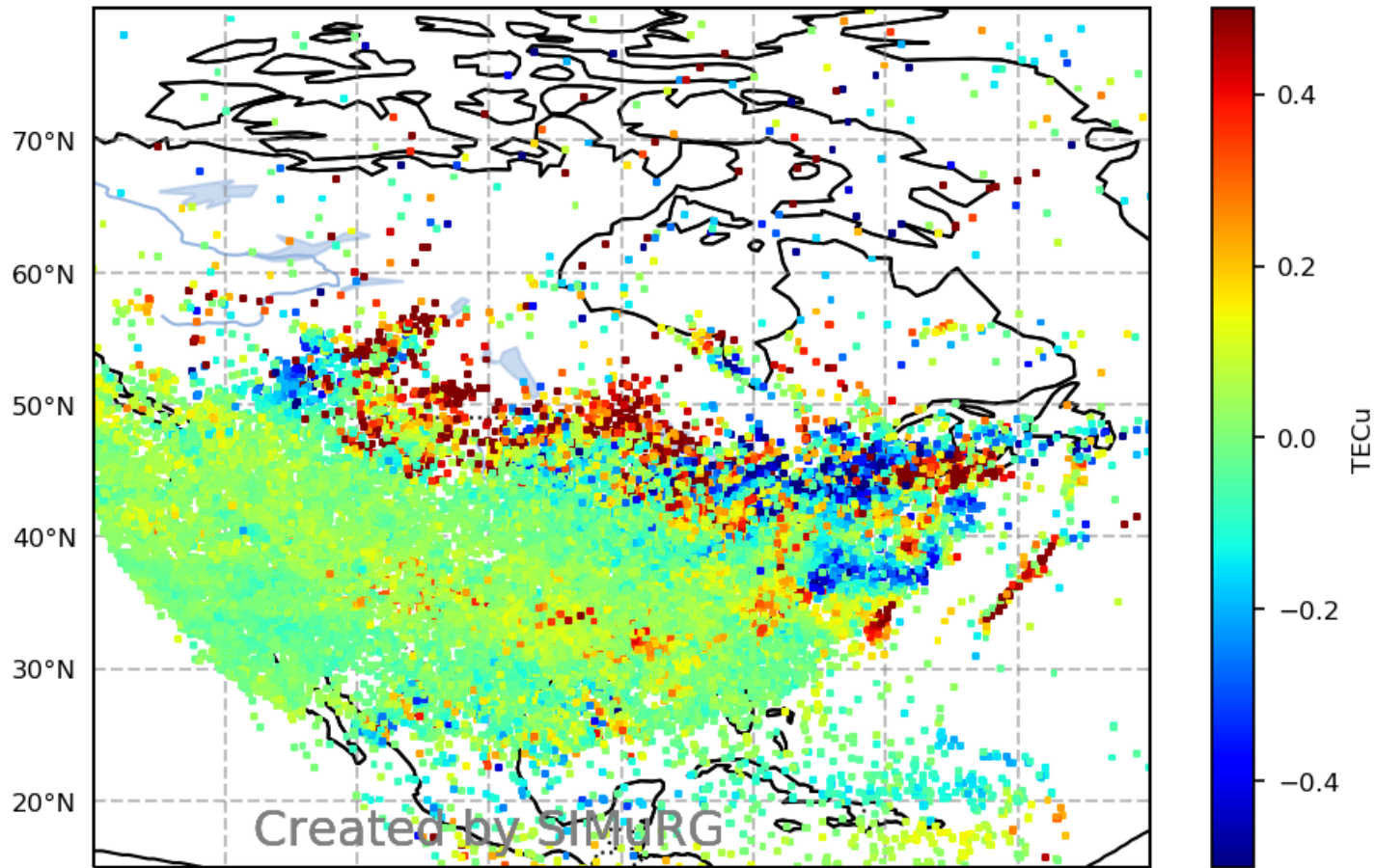


Выводы

✓ Для задачи удаления тренда, среди представленных в данной работе вариантов, лучшие результаты дает сглаживающий кубический сплайн. Он характеризуется минимальным значением математического ожидания ошибки и среднеквадратического отклонения ошибки, а также максимумом коэффициента корреляции.

✓ Для задачи селекции вариаций, среди представленных в данной работе вариантов, лучшие результаты получены для фильтра скользящего среднего. Селекция вариаций играет важную роль в интерпретации данных, особенно при автоматической обработке.

Что дальше?



1. Зайти на simurg.iszf.irk.ru и увидеть то, как работает новый способ фильтрации
2. Помочь проекту в поиске артефактов, как связанных с обработкой так и физическими явлениями
3. Задать вопросы, потому что доклад закончен, спасибо за внимание!

- Математическое ожидание ошибки

$$M_{RF}(\Delta I) = \frac{1}{N} \sum_{i=1}^N (I_R - I_F) \quad (1)$$

- Среднеквадратическое отклонение ошибки

$$\sigma_{RF}(\Delta I) = \sqrt{\frac{1}{N} \sum_{i=1}^N (I_R - I_F)^2} \quad (2)$$

- Коэффициент корреляции между I_R и I_F

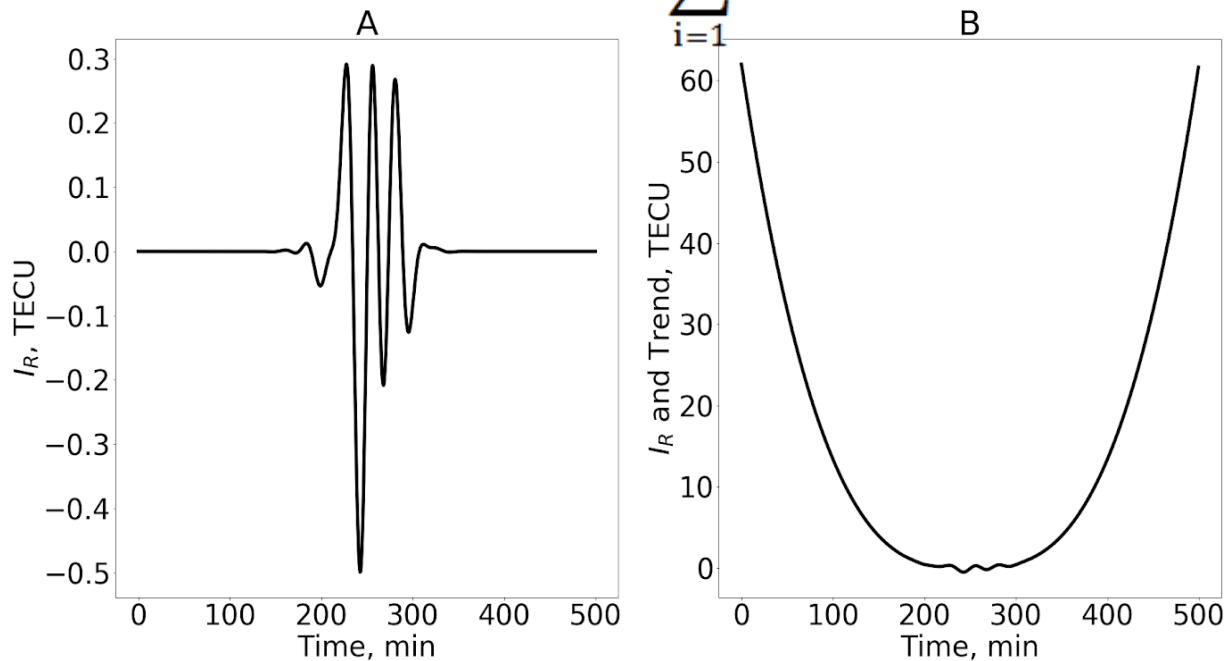
$$K_{RF}(\Delta I) = \frac{\sum_{i=1}^N (I_R - \langle I_R \rangle)_i (I_F - \langle I_F \rangle)_i}{\sqrt{\sum_{i=1}^N (I_R - \langle I_R \rangle)_i^2} \sqrt{\sum_{i=1}^N (I_F - \langle I_F \rangle)_i^2}} \quad (3)$$

Где I_R , I_F - опорный сигнал и сигнал полученный после осуществления процедур

Модель сигнала и модель тренда

В составе волнового пакета присутствуют три гармоники с частотами колебаний 20, 40 и 60 минут. Амплитуда гармоник - 0,2 TECU. Модель опорного сигнала имеет вид:

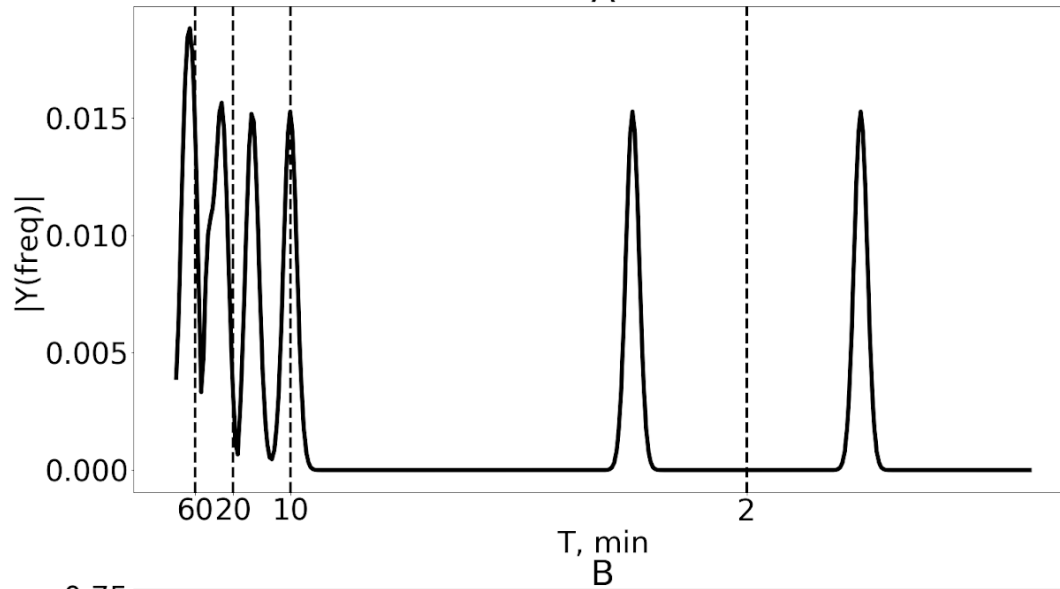
$$I_R = e^{\frac{-(t-t_m)^2}{2dt^2}} \sum_{i=1}^n \sin(\omega_i t)$$



Модель сигнала I_F (слева) и сигнала с трендом (справа)

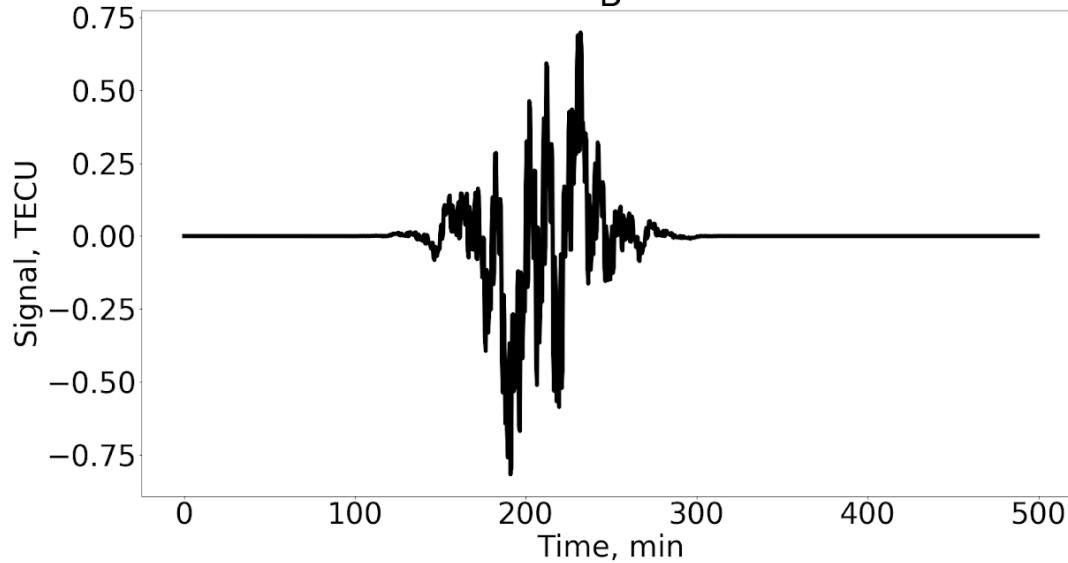
Модель сигнала

A



$$I_R = e^{-\frac{(t-t_m)^2}{2d_t^2}} \sum_{i=1}^n \sin(\omega_i t)$$

B



Фильтр Ходрика-Прескота

Применяется, когда общий сигнал:

Во временных рядах имеется тренд и циклично повторяющийся компонент

От параметра лямбда зависит к каким колебаниям чувствительный фильтр: краткосрочным и долгосрочным

$$\min_{\tau} \left(\sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right).$$

или

Темп роста
составляющей,
отвечающей за тренд

$$\min_{\tau} \left(\sum_{t=1}^n (y_t - x_t)^2 + \lambda \sum_{t=2}^{n-1} (x_{t-1} - 2x_t + x_{t+1}) \right).$$

Фильтр l1

Применяется, когда общий сигнал:

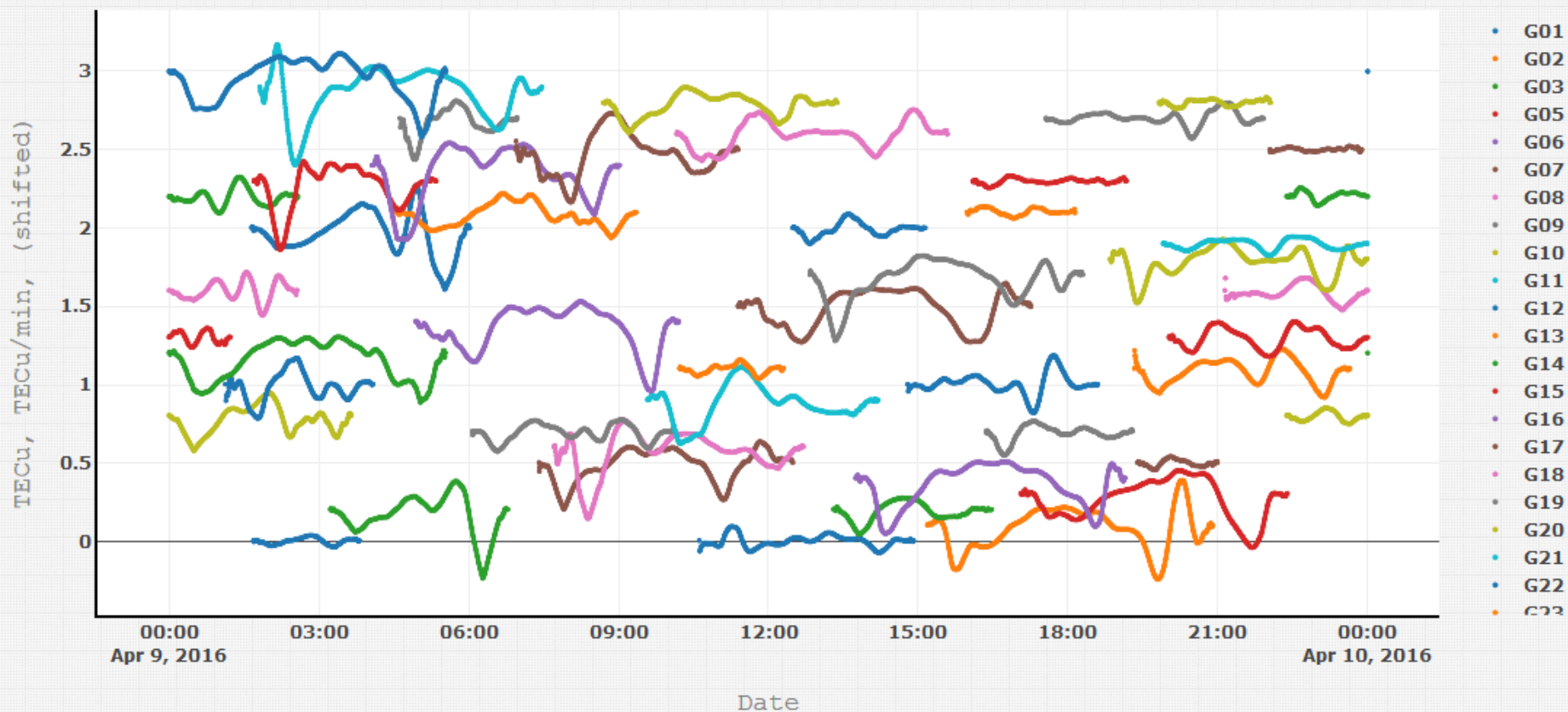
Во временных рядах имеется тренд и циклично повторяющийся компонент

От параметра лямбда зависит к каким колебаниям чувствительный фильтр:
краткосрочным и долгосрочным

$$\min_{\tau} \left(\sum_{i=1}^n (y_i - x_i)^2 + \lambda \sum_{i=2}^{n-1} |x_{i-1} - 2x_i + x_{i+1}| \right) .$$

Темп роста
составляющей,
отвечающей за тренд

Пример получаемых рядов



2-10 minute TEC variations

10-20 minute TEC variations

20-60 minute TEC variations

ROTI

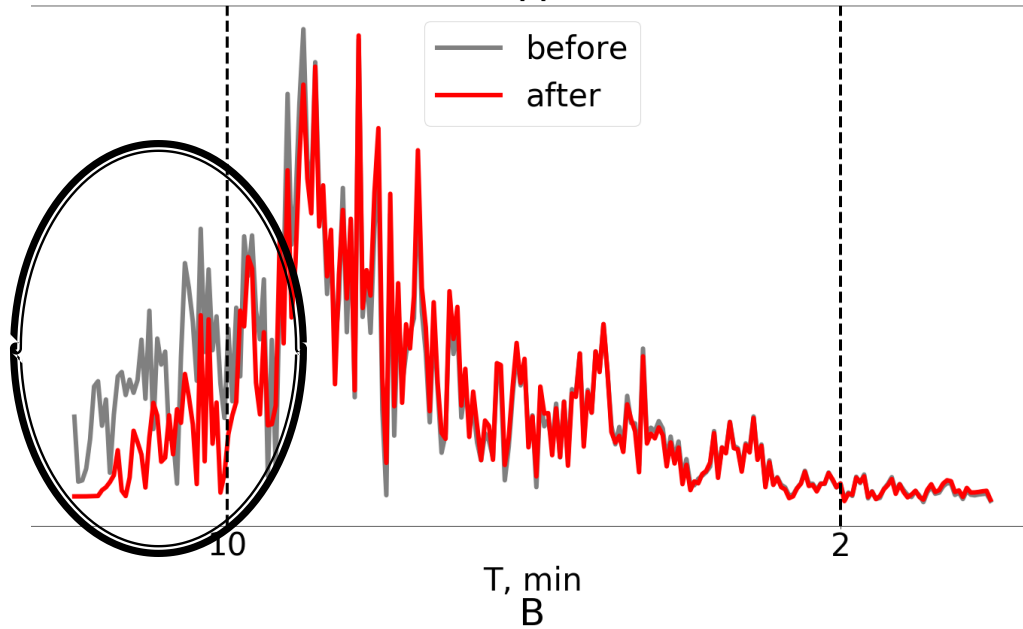
Adjusted TEC

geometry

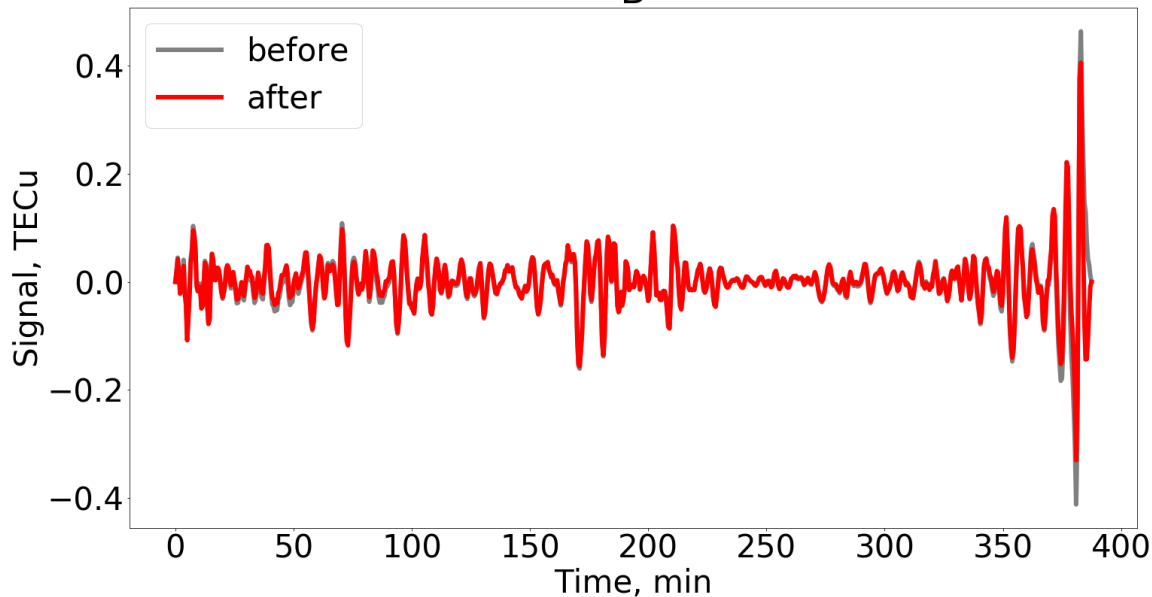
Download ▾

Пример артефактов

A



B

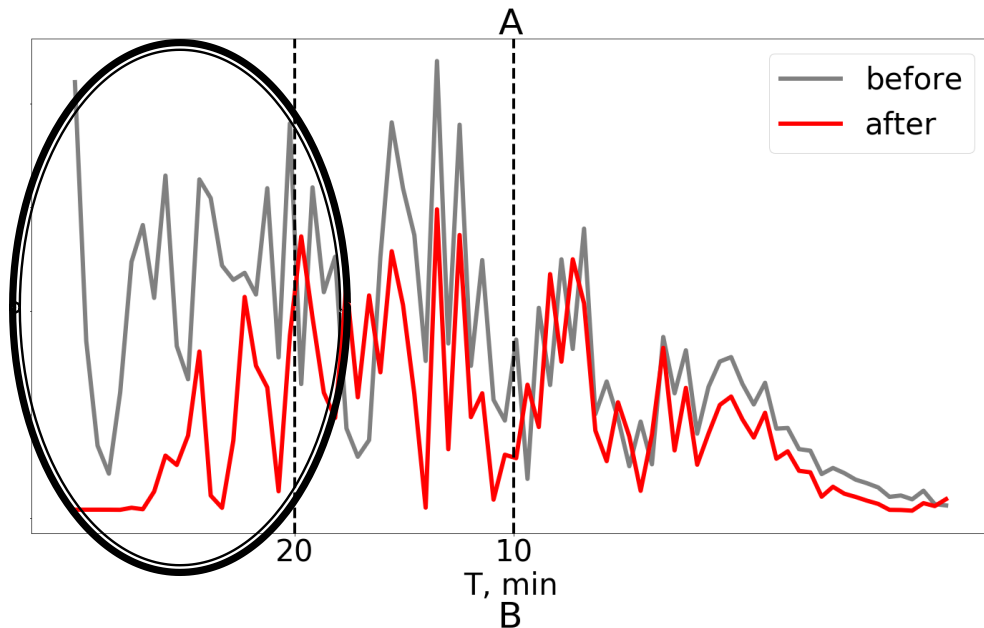


Пример артефактов
для диапазона 2-10
минут на основе
экспериментальных
данных

A – для частотной
области

B – для временной
области

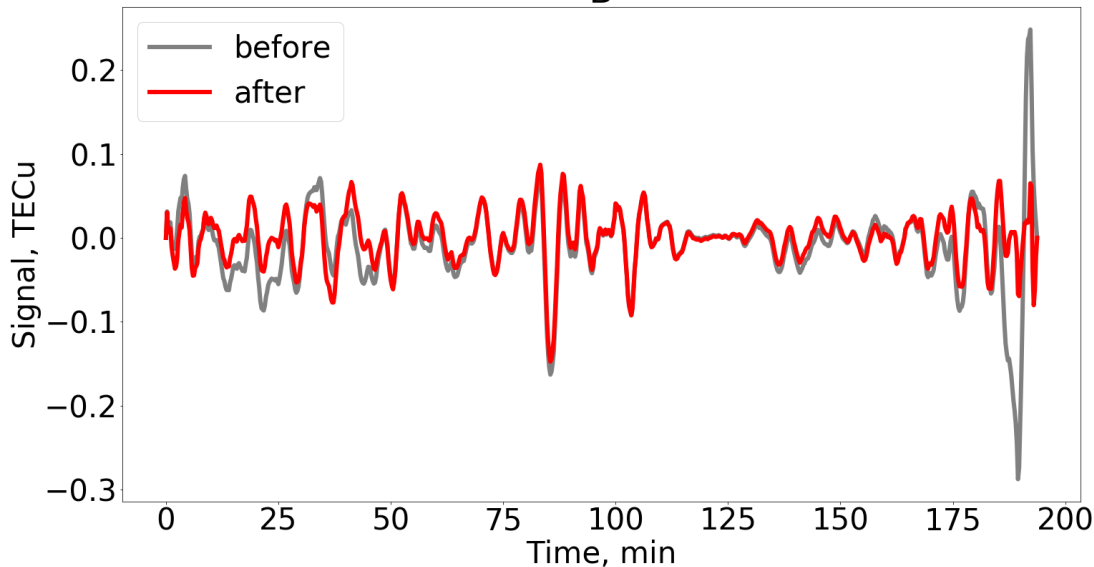
Пример артефактов



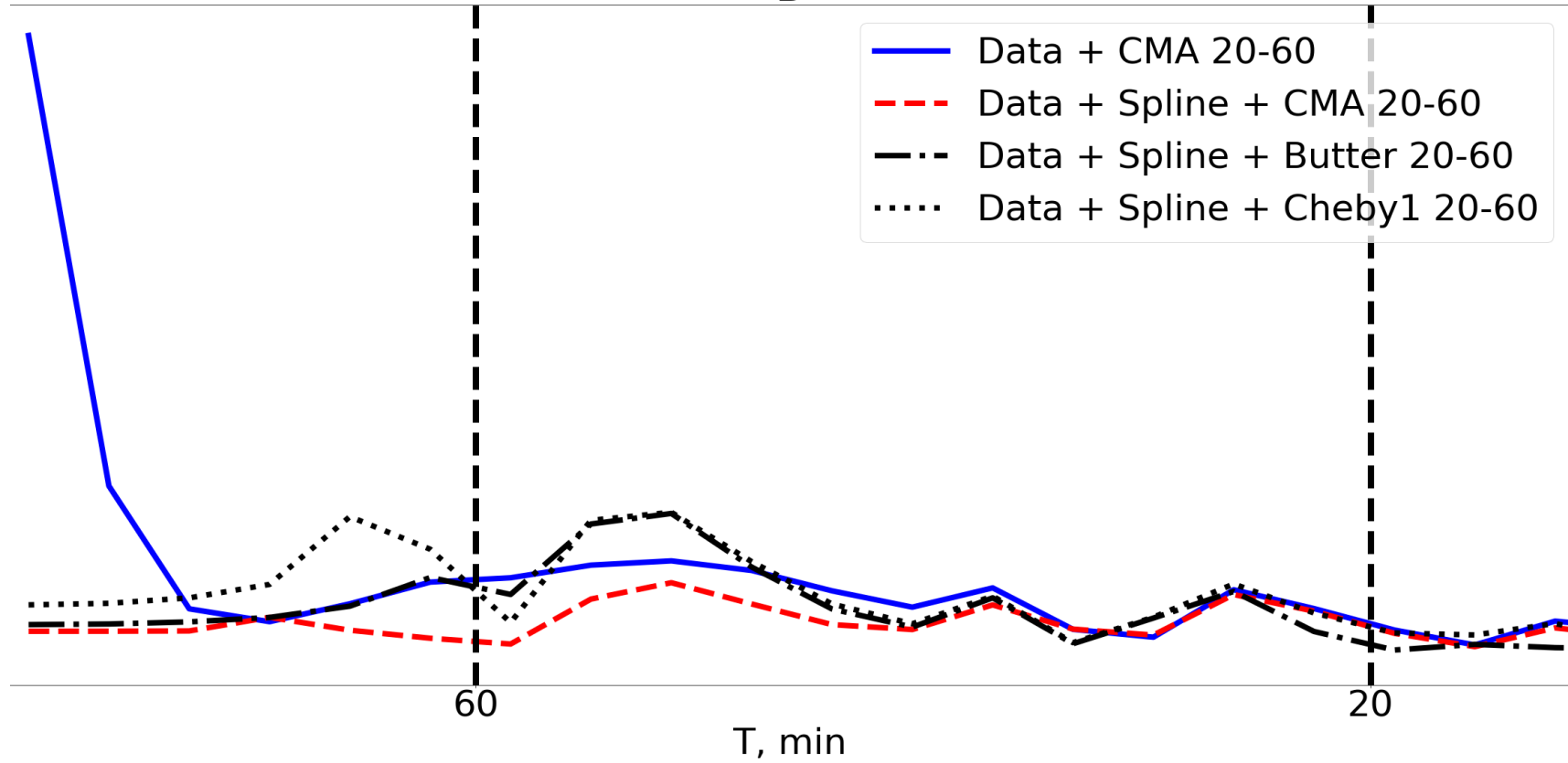
Пример артефактов
для диапазона 10-20
минут на основе
экспериментальных
данных

А – для частотной
области

В – для временной
области



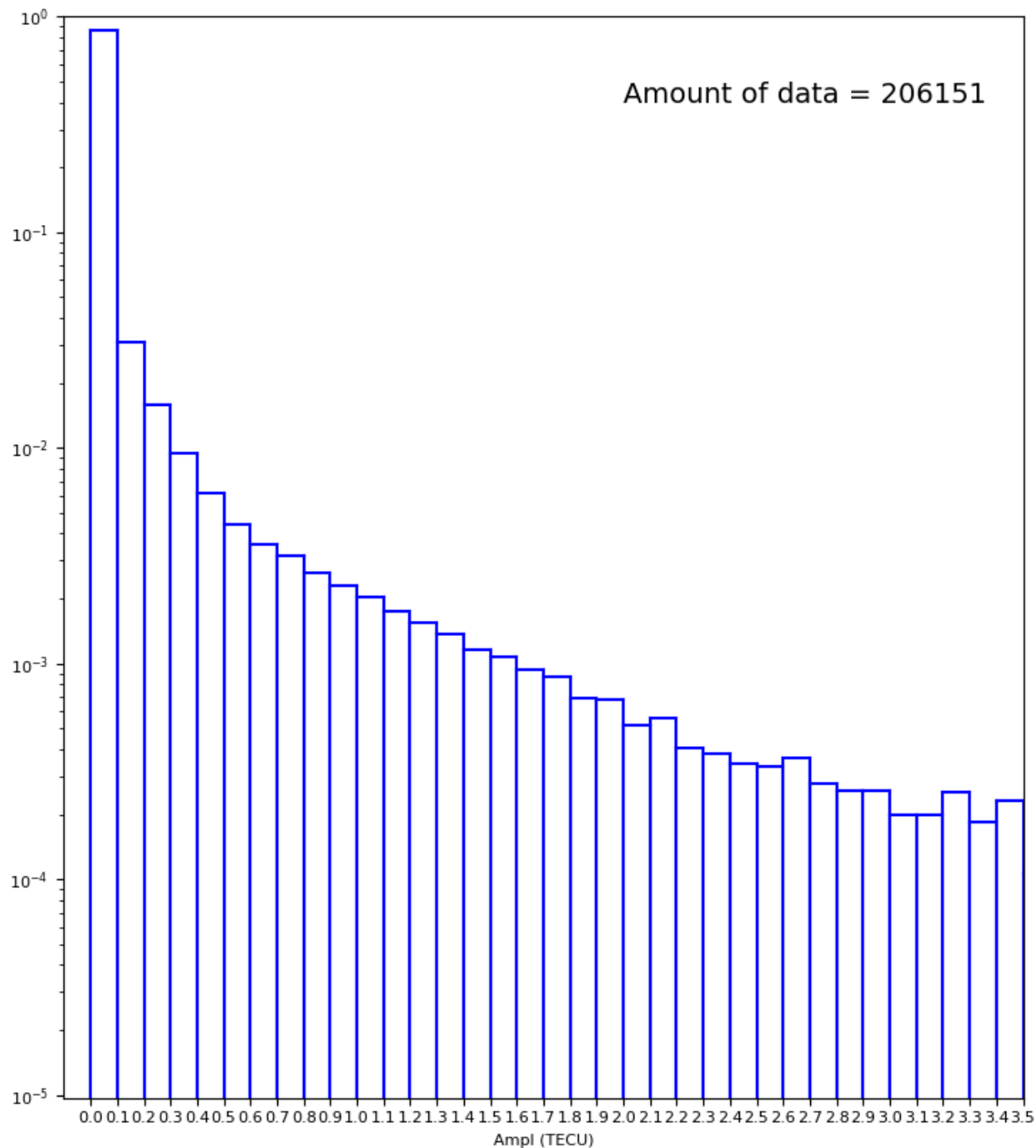
Вариации ПЭС



Скорость расчетов

Табл. 1. Скорость различных методов при удалении тренда

Способ удаления тренда	Время на обработку данных
CMA filter	2 298 seconds
Spline	1 796 seconds
Hodrick–Prescott	610 seconds
Polynom	2 031 seconds
L1 filter	10 965 seconds



Разброс математического ожидания разницы рядов вариаций (20-60 минут) с предварительным детрендированием и без

Amount of data = 206151

Разброс математического ожидания разницы рядов вариаций (2-20 минут) с предварительным детрендрованием и без

