

Санкт-Петербургский политехнический университет Петра  
Великого  
Институт прикладной математики и механики  
Кафедра «Прикладная математика»

## Отчёт по курсовой работе по дисциплине «Математическая статистика»

Выполнил студент:  
Карасев Александр Андреевич  
группа: 3630102/70201

Проверил:  
к.ф.-м.н., доцент  
Баженов Александр Николаевич

Санкт-Петербург  
2020г.

## Содержание

<b>1</b>	<b>Постановка задачи</b>	<b>4</b>
<b>2</b>	<b>Реализация</b>	<b>4</b>
2.1	Получение и обратка данных . . . . .	4
2.2	Вычисление коэффициентов . . . . .	5
2.3	Выведение результатов . . . . .	5
<b>3</b>	<b>Результаты</b>	<b>5</b>
3.1	Все коэффициенты близки к единице . . . . .	5
3.2	Некоторые коэффициенты малы . . . . .	6
3.3	Все коэффициенты малы . . . . .	7
<b>4</b>	<b>Обсуждения</b>	<b>8</b>
<b>5</b>	<b>Ссылки</b>	<b>9</b>

## Список иллюстраций

1	Температура SHT38880 . . . . .	6
2	Температура SHT38882 . . . . .	7
3	Температура SHT38890 . . . . .	8

## Список таблиц

1	результаты SHT38880 . . . . .	5
2	результаты SHT38882 . . . . .	6
3	результаты SHT38882 . . . . .	7

## 1 Постановка задачи

Имеется набор данных, содержащий информацию о температуре плазмы. Из этих данных по некоторым каналам можно получить дискретные сигналы, из которых затем составляется температура плазмы. Необходимо рассмотреть зависимости обработанных сигналов из разных каналов, то есть найти коэффициенты корреляции.

## 2 Реализация

### 2.1 Получение и обработка данных

За основу была взята функция `def mainProcess(shtNum, nums, ROI=True, borders=False, left=LEFT, right=RIGHT, graphics=False)` Принимающая аргументы:

- `shtNum` - номер sht-файла
- `nums` - массив с номерами (могут быть 80, 50, 27, 15) из названий сигналов, например, "SXR 80 mkm". Результаты собраны для всех значений
- `ROI` - bool-переменная, отвечающая за выделение области для анализа (region of interest) у сигналов. Уставлена в `True` для сбора данных
- `borders` - булевская переменная, отвечающая за выставление границ промежутка, указанных в следующих двух переменных. Установлена в `False`
- `left` - значение левой границы промежутка. Не используется
- `right` - значение правой границы промежутка. Не используется
- `graphics` - bool-переменная, отвечающая за вывод графика на экран

Обработка сигнала сигнала проходит в несколько этапов:

1. Считывание данных из sht-файла
2. Выделение области для анализа
3. Применение фильтров к сигналу

После проведенных процедур находить коэффициенты корреляции еще нельзя, так как данные не приведены к единому промежутку.

## 2.2 Вычисление коэффициентов

Обработанные сигналы приводятся к необходимому виду в функции *process*. В ней уже можно применять метод по нахождению коэффициента корреляции.

Для этого был выбран метод Пирсона. Данные имеют большой размер, поэтому метод применим и дает правильный результат. Реализация метода была взята из библиотеки *scipy*. *Замечание.* Реализация считывания сигналов и фильтров для работы с ними, так же для определения интересующей области, находится в библиотеке *pyglobus*.

## 2.3 Выведение результатов

Коэффициенты корреляции для различных пар каналов выводятся в отдельный txt-файл. Подпись "r-80\_50" соответствует коэффициенту корреляции между данными из каналов "SXR 80 mkm" и "SXR 50 mkm".

# 3 Результаты

Был проверен весь исходный набор данных. Рассмотрим все шаблоны в полученных коэффициентах корреляции.

## 3.1 Все коэффициенты близки к единице

В таблице приведены коэффициенты корреляции, полученные для файла "SHT38880.sht"

Пара каналов	Коэффициент корреляции
80, 50	0.9979403374732547
80, 27	0.9896770974701469
80, 15	0.9745135204114533
50, 27	0.9955636793717355
50, 15	0.9841912121588487
27, 15	0.9948662857705999

Таблица 1: результаты SHT38880

Видно, что все коэффициенты близки к единице. На графике температуры это отображается - все промежутки выровнены, пики температур согласованы во времени

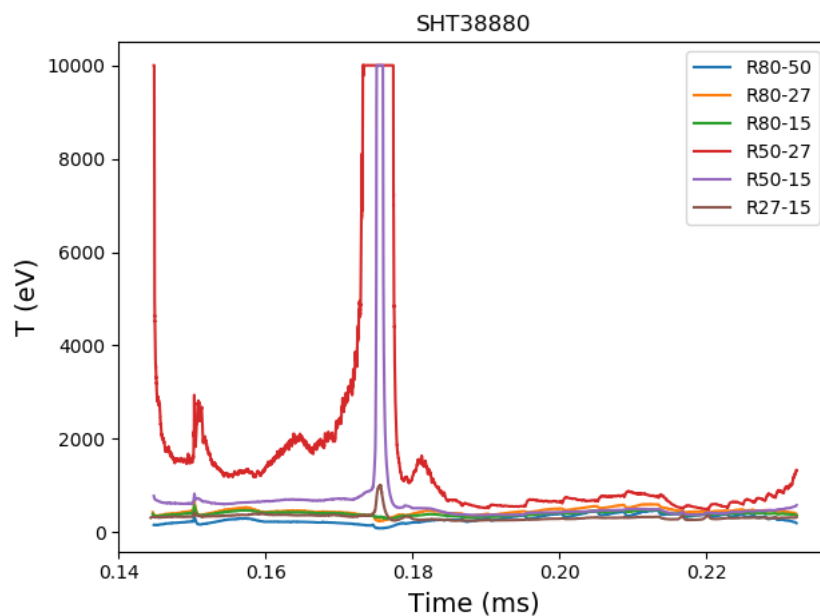


Рис. 1: Температура SHT38880

### 3.2 Некоторые коэффициенты малы

В таблице приведены коэффициенты корреляции, полученные для файла "SHT38882.sht"

Пара каналов	Коэффициент корреляции
80, 50	0.43432564703361987
80, 27	0.3512913125981475
80, 15	0.9330225125924979
50, 27	0.05986386772223856
50, 15	0.9667362393544101
27, 15	0.9869440612855251

Таблица 2: результаты SHT38882

Первый (пара 80/50), второй (80/27), четвертый (50/27) коэффициенты меньше единицы. График температур отражает, что для пар с низкими коэффициентами промежуток подобран неверно.

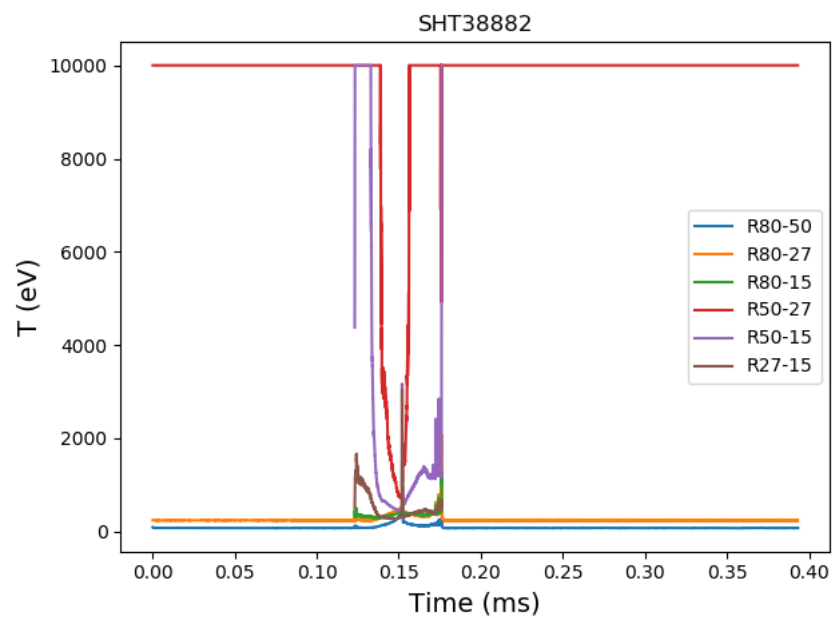


Рис. 2: Температура SHT38882

### 3.3 Все коэффициенты малы

В таблице приведены коэффициенты корреляции, полученные для файла "SHT38890.sht"

Пара каналов	Коэффициент корреляции
80, 50	0.06415572640175918
80, 27	0.018810970169101215
80, 15	-0.5135768913233768
50, 27	-0.48811953595844054
50, 15	0.06822686051128007
27, 15	-0.14243934122518068

Таблица 3: результаты SHT38882

Все коэффициенты меньше единицы по модулю. Данные неправильно коррелируют. Был обнаружен только шум.



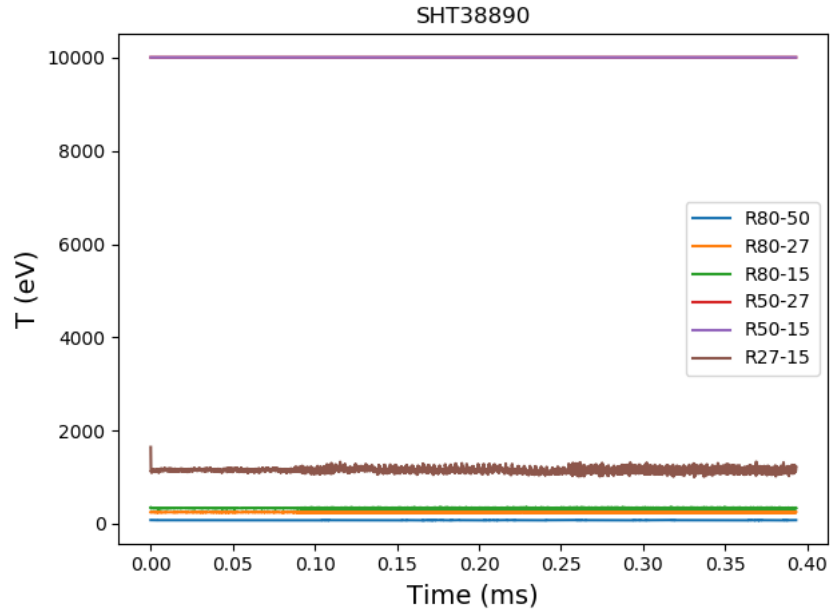


Рис. 3: Температура SHT38890

## 4 Обсуждения

Показано, что коэффициенты корреляции между данными с разных каналов отражают правильность работы функций из библиотеки `pyglobus`.

- Все коэффициенты близкие к единицы означают правильность работы программы и соответствие сигналов
- Низкие коэффициенты на всех парах каналов означают отсутствие информативного сигнала.
- Низкие коэффициенты на части пар каналов соответствуют неправильно определенным границам промежутка.

Заметим, что последний случай составляет основную массу среди всех во-сокупности данных.

## 5 Ссылки

Репозиторий с кодом: <https://github.com/MethaHardworker/MathStats/tree/course>

Реализация pyglobus: <https://github.com/dev0x13/globus-plasma>