

The 9th International Conference in Deep Learning in Computational Physics,
July, 2-4, 2025, Moscow, Russia

USING MACHINE LEARNING METHODS FOR JOINT PROCESSING OF DATA FROM MULTIPLE SEMICONDUCTOR GAS SENSORS*

Isaev I.V.^{1,2,3}, Chernov K.N.⁴, Dolenko S.A.¹, Krivetskiy V.V.^{2, 5}

¹ D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University,

² Scientific-Manufacturing Complex Technological Centre,

³ MIREA – Russian Technological University,

⁴ Physics Department, M.V. Lomonosov Moscow State University

⁵ Chemistry Department, M.V. Lomonosov Moscow State University

** The study was carried out at the expense of the grant No. 22-19-00703-P from the Russian Science Foundation.
<https://rscf.ru/en/project/22-19-00703/>.*

Определение состава газовых смесей

Области применения

❑ Промышленность

- Безопасность технологических процессов
- Средства индивидуальной защиты персонала

❑ Экологический мониторинг

- Оценка качества воздуха городов, промышленных зон, транспорта, жилых и рабочих помещений

❑ Медицина

- Неинвазивная медицинская диагностика и контроль течения болезней
- Контроль эффективности спортивных тренировок

❑ Экспресс - анализ качества, оригинальности, свежести продуктов питания и напитков на основе их запаха.



Определение состава газовых смесей

Существующие технические решения

Принципы работы газоаналитических датчиков:

☐ Термокаталитический

- Измерение количества тепла выделяемого при сгорании газа на катализаторе

☐ Термокондуктивный

- Измерение разницы теплопроводности

☐ Электрохимический и гальванический

- Измерение силы тока, возникающего при химических процессах в электролите

☐ Оптический

- Инфракрасный (ИК-поглощение), интерферометрический, фотоионизационный (УФ)

☐ Полупроводниковый

- Измерение изменения сопротивления полупроводника при абсорбции газов

Определение состава газовых смесей

Существующие технические решения

Преимущества полупроводниковых сенсоров:

- ☐ Возможность длительной непрерывной работы
 - Не имеют расходующихся частей в отличие от электрохимических сенсоров
- ☐ Более низкий предел обнаружения газов
 - По сравнению с оптическими и термокаталитическими сенсорами
- ☐ Широкий спектр детектируемых веществ
- ☐ Низкая стоимость
 - Возможность использования масштабируемых технологий изготовления
- ☐ Низкое энергопотребление
- ☐ Миниатюрность

Определение состава газовых смесей

Направление работ

Общее направление работ: разработка мультidetекторного прибора, содержащего как один, так и массив сенсорных элементов, для одновременного выявления нескольких веществ в воздухе, а также определения качественного и количественного состава газовых смесей.

Определение состава газовых смесей

Возможные постановки задач ММО

Общее направление работ: разработка мультidetекторного прибора, содержащего как один, так и массив сенсорных элементов, для одновременного выявления нескольких веществ в воздухе, а также определения качественного и количественного состава газовых смесей.

Возможные постановки задач, при использовании методов машинного обучения:

☐ Регрессия

- Определение концентрации конкретного газа в воздухе,

☐ Бинарная классификация, многоклассовая классификация

- Определение наличия/отсутствия конкретного газа в воздухе, содержащим примесь одного газа

☐ Многометочная классификация

- Определение качественного состава газовой смеси из нескольких газов

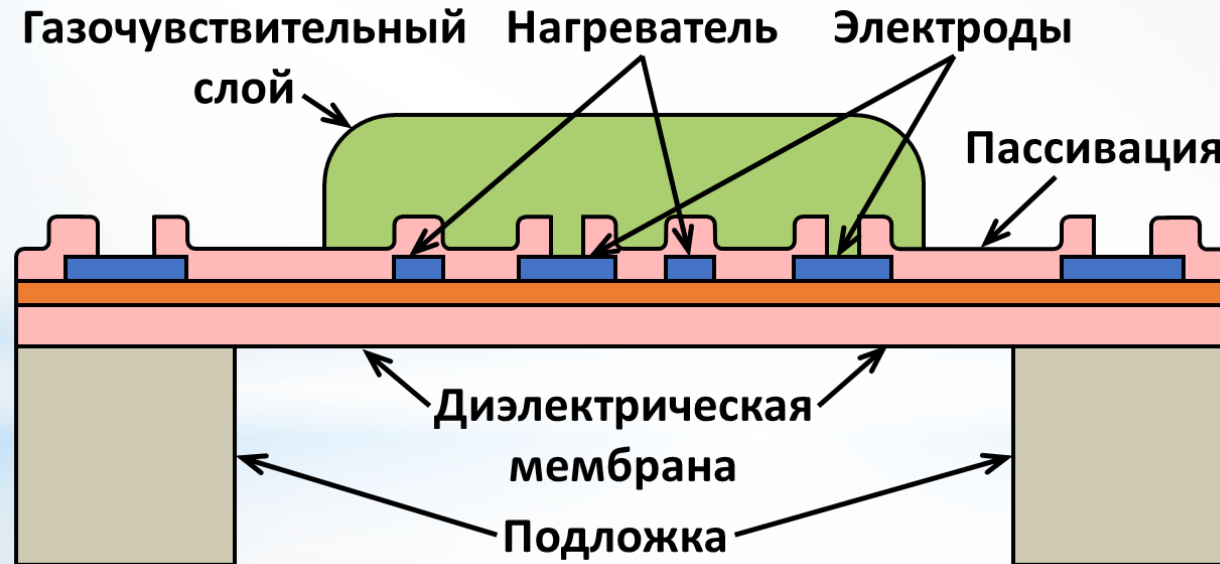
На этапе
набора
данных

Планируется

Физический эксперимент

Полупроводниковый газовый сенсор

Схематическое устройство полупроводникового газового сенсора

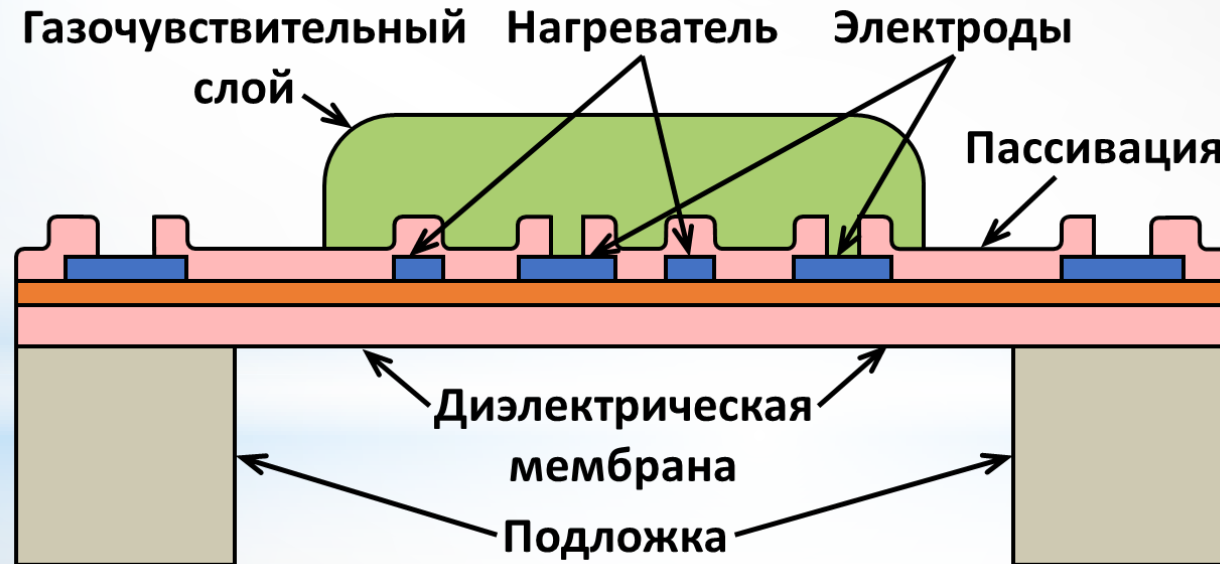


Принцип работы полупроводникового газового сенсора основан на том, что молекулы газа, связываясь с материалом сенсора, влияют на его проводимость (электрическое сопротивление).

Физический эксперимент

Полупроводниковый газовый сенсор

Схематическое устройство полупроводникового газового сенсора



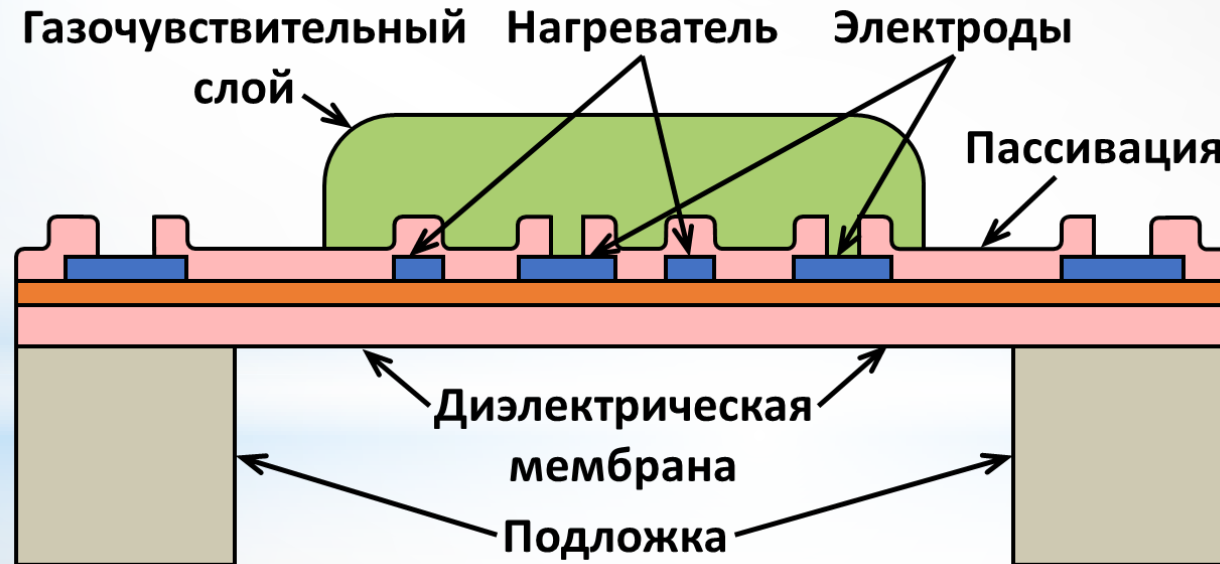
Факторы, влияющие на степень связывания:

- ☐ Тип газа и его концентрация
- ☐ Материал датчика
- ☐ Температура

Физический эксперимент

Полупроводниковый газовый сенсор

Схематическое устройство полупроводникового газового сенсора



Факторы, влияющие на степень связывания:

- ❑ Тип газа и его концентрация
- ❑ Материал датчика
- ❑ Температура

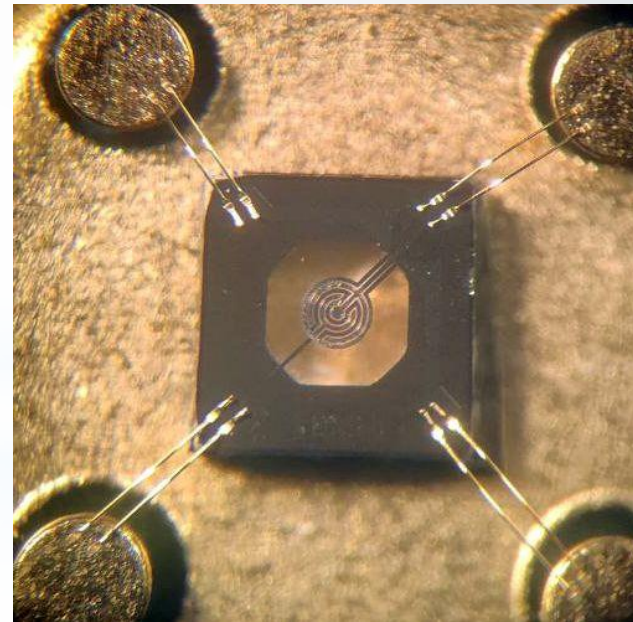
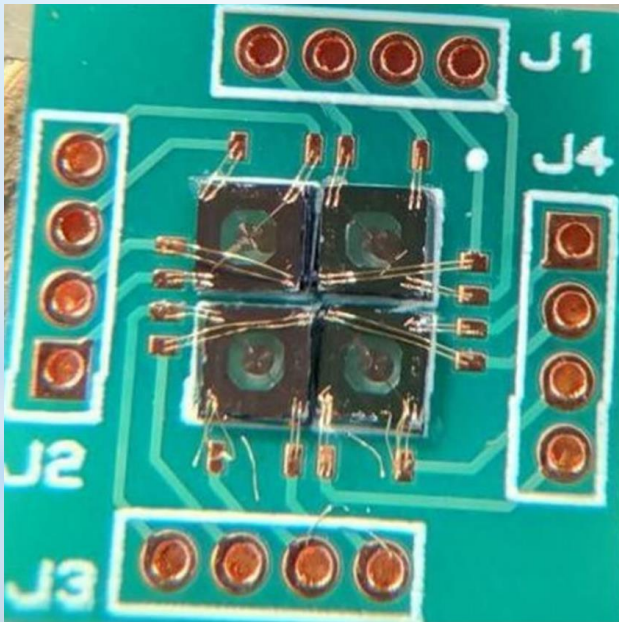


Для достижения **селективности** при определении конкретных газов используется **комплект сенсоров** с различными **легирующими добавками**

Физический эксперимент

Полупроводниковый газовый сенсор

Комплект сенсорных элементов



Факторы, влияющие на степень связывания:

- ☐ Тип газа и его концентрация
- ☐ Материал датчика
- ☐ Температура

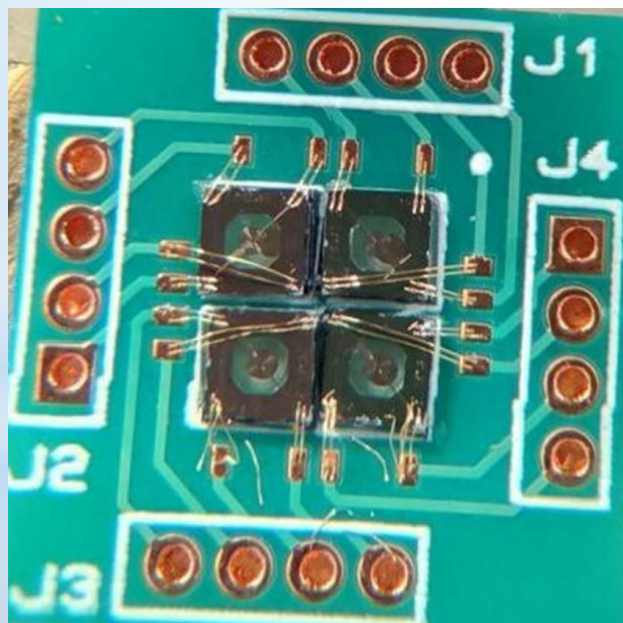


Для достижения **селективности** при определении конкретных газов используется **комплект сенсоров** с различными **легирующими добавками**

Физический эксперимент

Полупроводниковый газовый сенсор

Комплект сенсорных элементов



SnO_2

1. SnO_2 ;
2. $\text{SnO}_2 - \text{Ru}$;
3. $\text{SnO}_2 - \text{Au}$;
4. $\text{SnO}_2 - \text{Pt}$;
5. $\text{SnO}_2 - \text{Pd}$;
6. $\text{SnO}_2 - \text{Cr} - \text{Nb}$;
7. $\text{SnO}_2 - \text{Si}$;
8. $\text{SnO}_2 - \text{Si} - \text{Au}$;

TiO_2

9. $\text{TiO}_2 - \text{Cr}$;
10. $\text{TiO}_2 - \text{Cr} - \text{Au}$;
11. $\text{TiO}_2 - \text{Nb} - \text{Au}$;
12. $\text{TiO}_2 - \text{Nb}$

Факторы, влияющие на степень связывания:

- ☐ Тип газа и его концентрация
- ☐ Материал датчика
- ☐ Температура

Для достижения **селективности** при определении конкретных газов используется **комплект сенсоров** с различными **легирующими добавками**

Физический эксперимент

Полупроводниковый газовый сенсор

Динамики нагрева

Факторы, влияющие на степень связывания:

- ☐ Тип газа и его концентрация
- ☐ Материал датчика
- ☐ Температура



- Степень связывания молекул газа уменьшается с повышением температуры.
- **Необходим** нагрев для приведения датчика в **исходное состояние**.
- Для обеспечения **временного разрешения** необходимо использовать **циклический нагрев и охлаждение**.

Физический эксперимент

Полупроводниковый газовый сенсор

Динамики нагрева

Факторы, влияющие на степень связывания:

- ☐ Тип газа и его концентрация
- ☐ Материал датчика
- ☒ Температура



- Электрическое сопротивление полупроводниковых материалов также зависит от температуры.



Для **повышения селективности** определения газов использовались различные температурные режимы работы:
т. н. **динамики нагрева**.



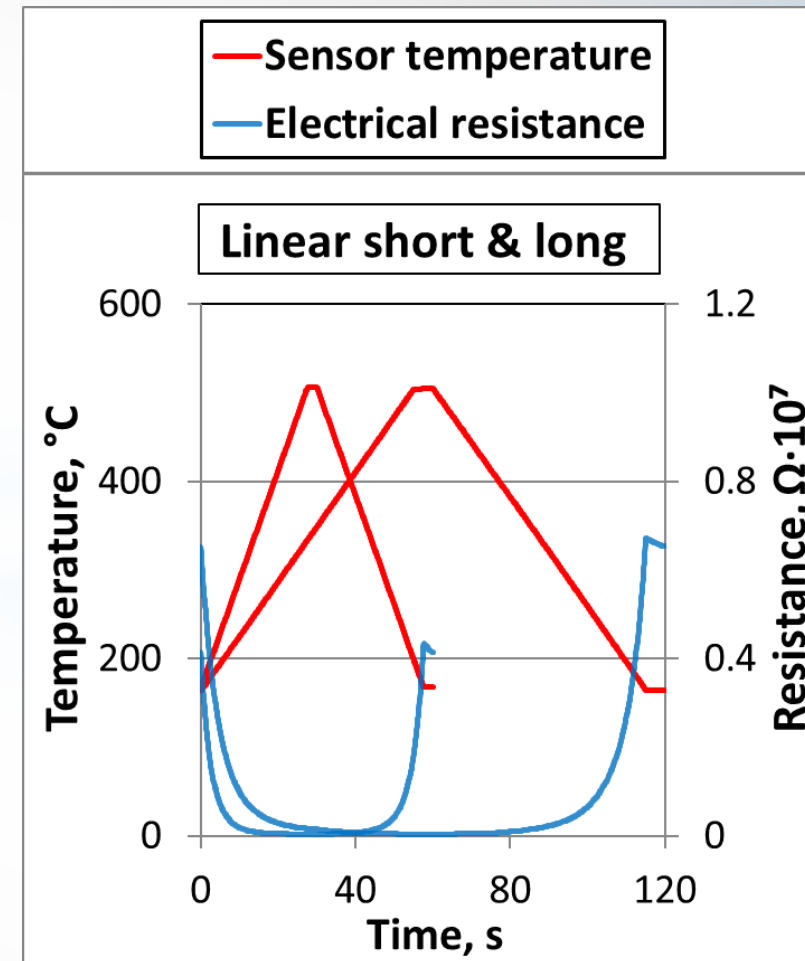
- Степень связывания молекул газа уменьшается с повышением температуры.
- Необходим нагрев для приведения датчика в исходное состояние.
- Для обеспечения временного разрешения необходимо использовать **циклический нагрев и охлаждение**.



Физический эксперимент

Динамики нагрева

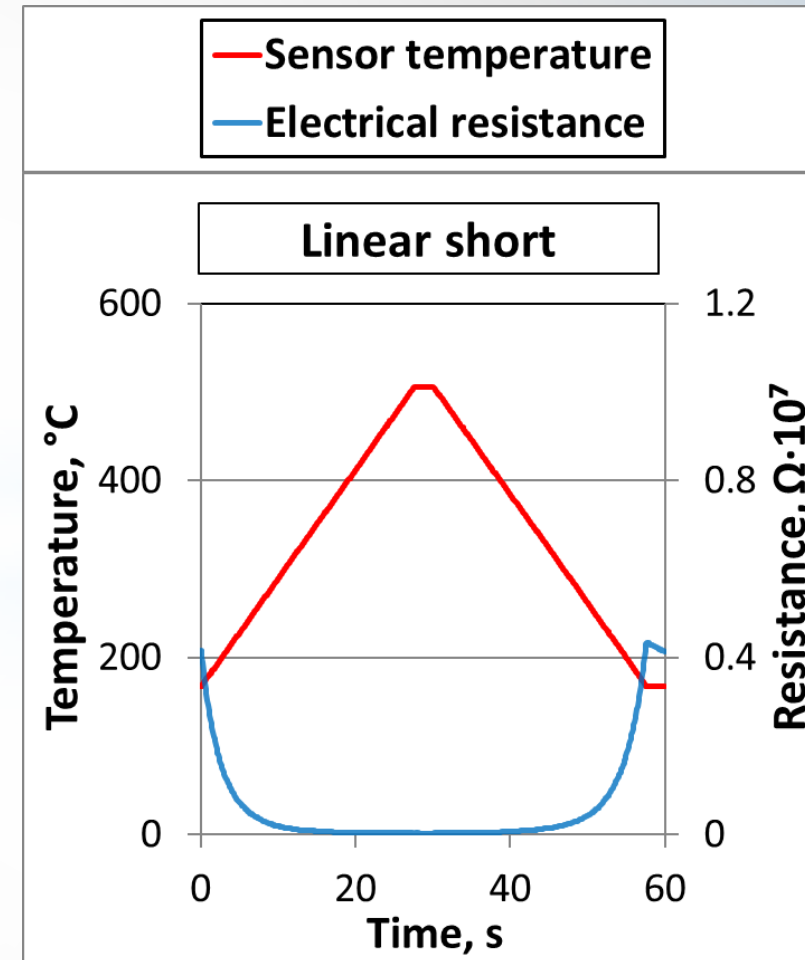
- ❑ Линейный нагрев и охлаждение
 - Линейный длинный
 - Линейный короткий
- ❑ Ступенчатое плавное повышение или понижение температуры
 - Пошаговый вверх
 - Пошаговый вниз
- ❑ Ступенчатое плавное повышение или понижение температуры с кратковременными импульсными скачками температуры до максимума.
 - Импульсный вверх
 - Импульсный вниз



Физический эксперимент

Динамики нагрева

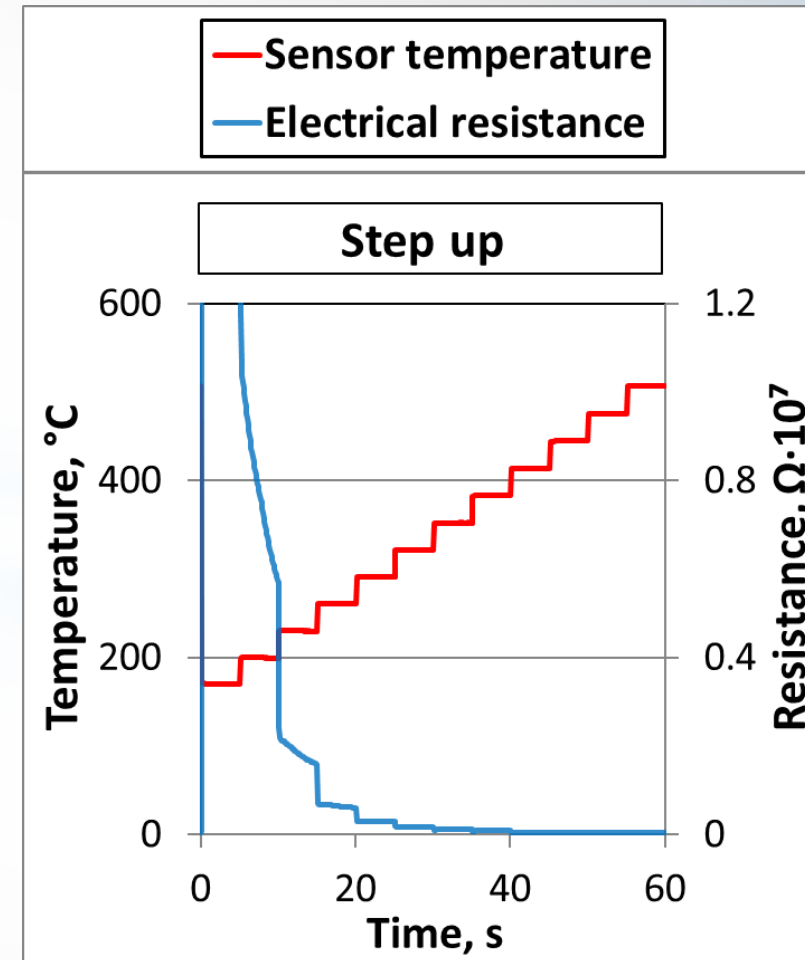
- ❑ Линейный нагрев и охлаждение
 - Линейный длинный
 - **Линейный короткий**
- ❑ Ступенчатое плавное повышение или понижение температуры
 - Пошаговый вверх
 - Пошаговый вниз
- ❑ Ступенчатое плавное повышение или понижение температуры с кратковременными импульсными скачками температуры до максимума.
 - Импульсный вверх
 - Импульсный вниз



Физический эксперимент

Динамики нагрева

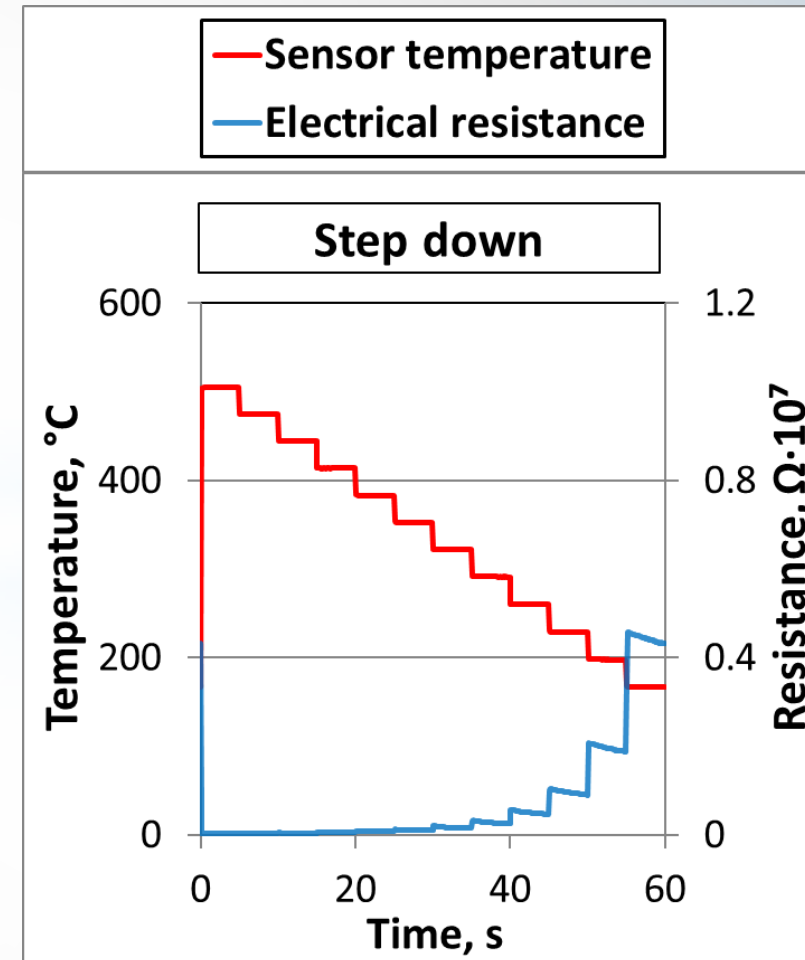
- ❑ Линейный нагрев и охлаждение
 - Линейный длинный
 - Линейный короткий
- ❑ Ступенчатое плавное повышение или понижение температуры
 - Пошаговый вверх
 - Пошаговый вниз
- ❑ Ступенчатое плавное повышение или понижение температуры с кратковременными импульсными скачками температуры до максимума.
 - Импульсный вверх
 - Импульсный вниз



Физический эксперимент

Динамики нагрева

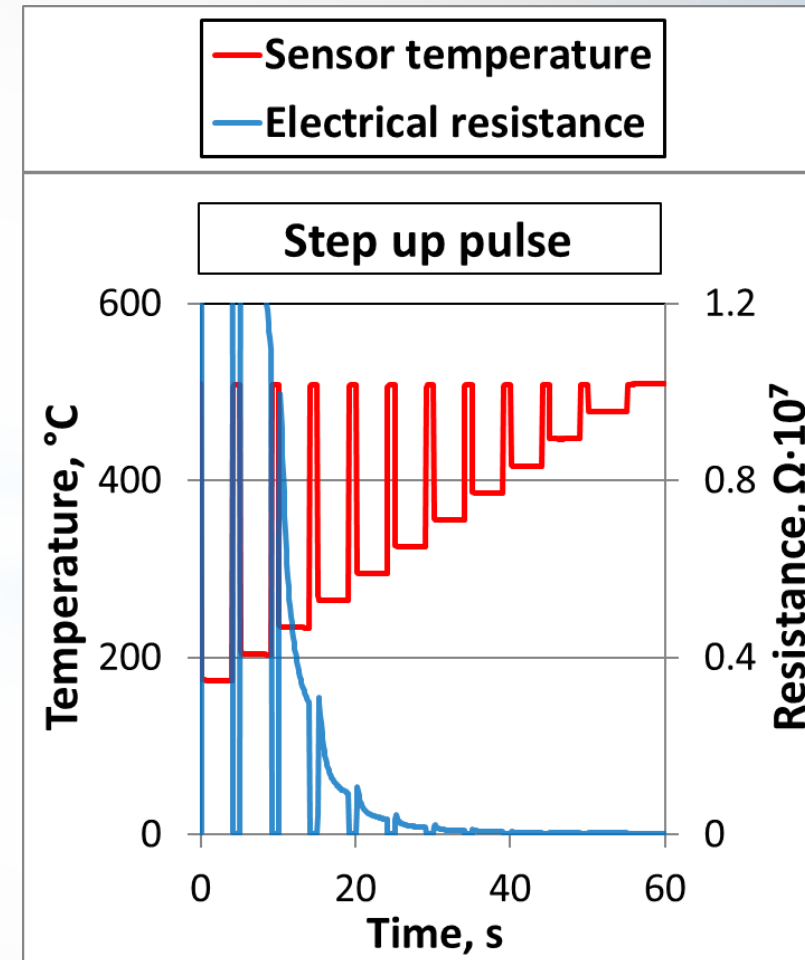
- ❑ Линейный нагрев и охлаждение
 - Линейный длинный
 - Линейный короткий
- ❑ Ступенчатое плавное повышение или понижение температуры
 - Пошаговый вверх
 - Пошаговый вниз
- ❑ Ступенчатое плавное повышение или понижение температуры с кратковременными импульсными скачками температуры до максимума.
 - Импульсный вверх
 - Импульсный вниз



Физический эксперимент

Динамики нагрева

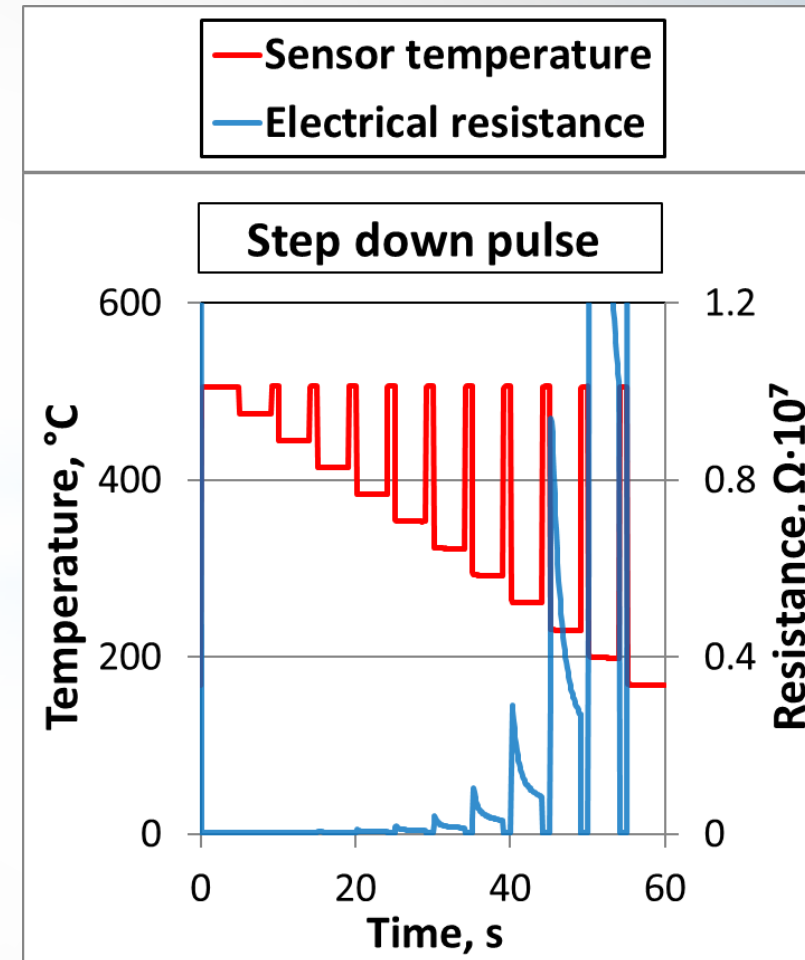
- ❑ Линейный нагрев и охлаждение
 - Линейный длинный
 - Линейный короткий
- ❑ Ступенчатое плавное повышение или понижение температуры
 - Пошаговый вверх
 - Пошаговый вниз
- ❑ Ступенчатое плавное повышение или понижение температуры с кратковременными импульсными скачками температуры до максимума.
 - Импульсный вверх
 - Импульсный вниз



Физический эксперимент

Динамики нагрева

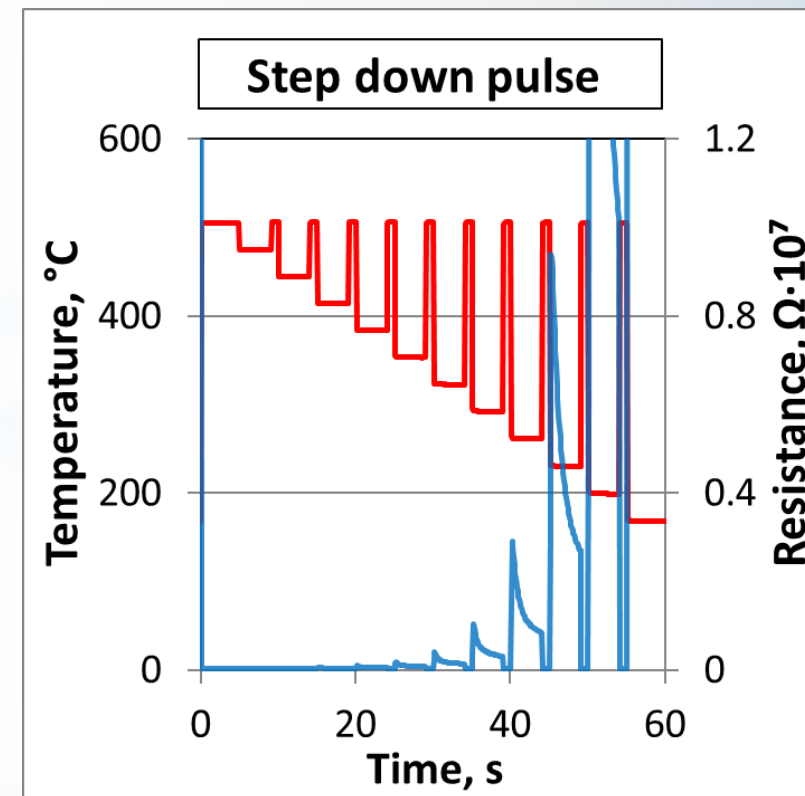
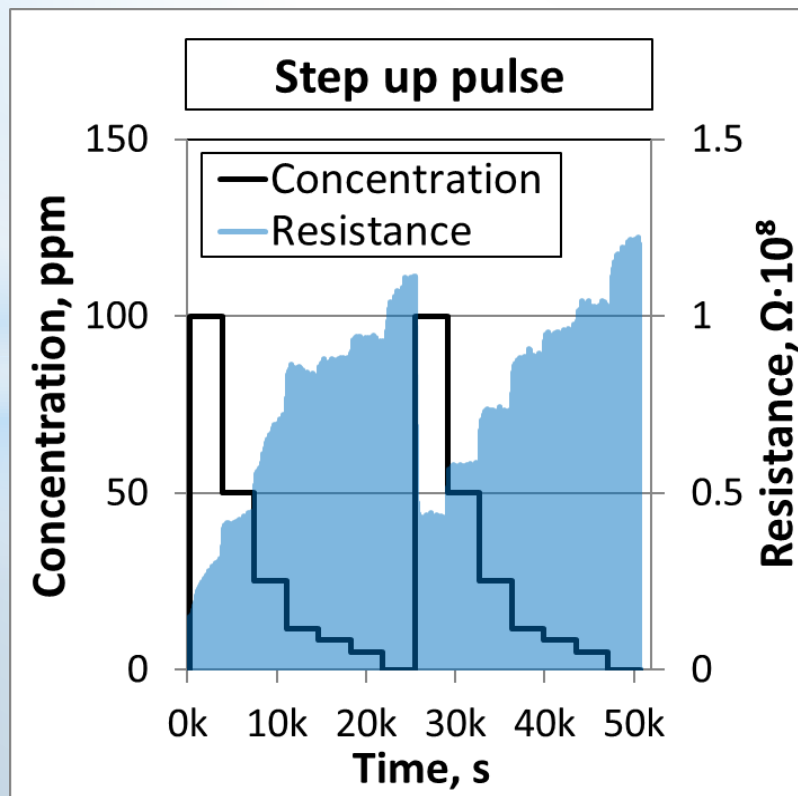
- ❑ Линейный нагрев и охлаждение
 - Линейный длинный
 - Линейный короткий
- ❑ Ступенчатое плавное повышение или понижение температуры
 - Пошаговый вверх
 - Пошаговый вниз
- ❑ Ступенчатое плавное повышение или понижение температуры с кратковременными импульсными скачками температуры до максимума.
 - Импульсный вверх
 - Импульсный вниз



Физический эксперимент

Концентрации газов

В ходе эксперимента были собраны формы сенсорного отклика от **12 датчиков** при **6 различных динамиках нагрева** и **6 различных концентрациях газа**. Несколько серий.



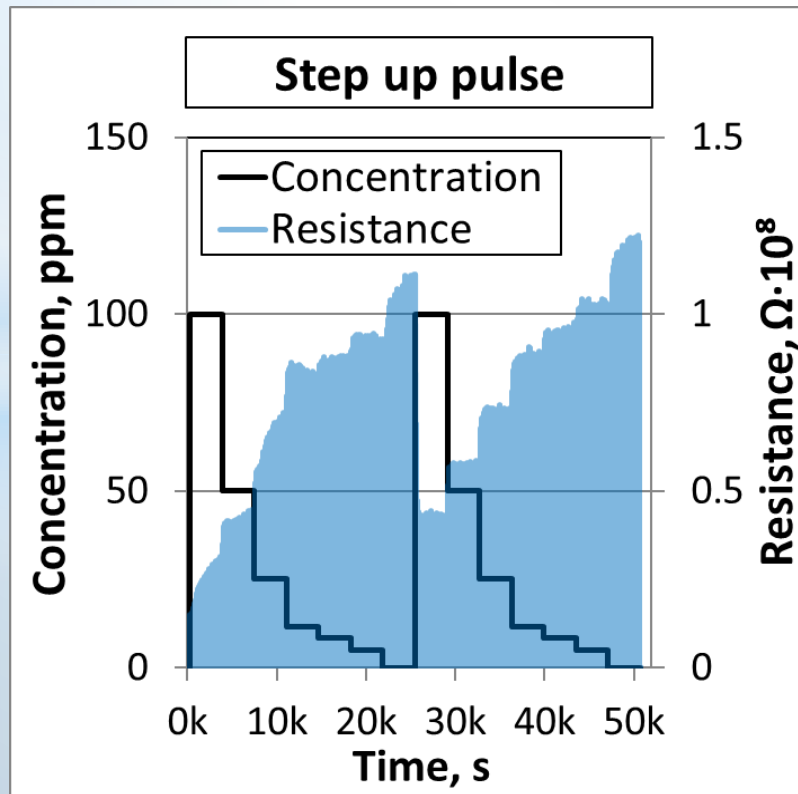
В атмосфере чистого воздуха, а также **в воздухе с примесью газов**:

CO, H₂, CH₄, NH₃, NO, NO₂, H₂S, SO₂, HCON (**только один газ** в каждом эксперименте)

Вычислительный эксперимент

Подготовка и предобработка данных

Этапы подготовки и предобработки данных:

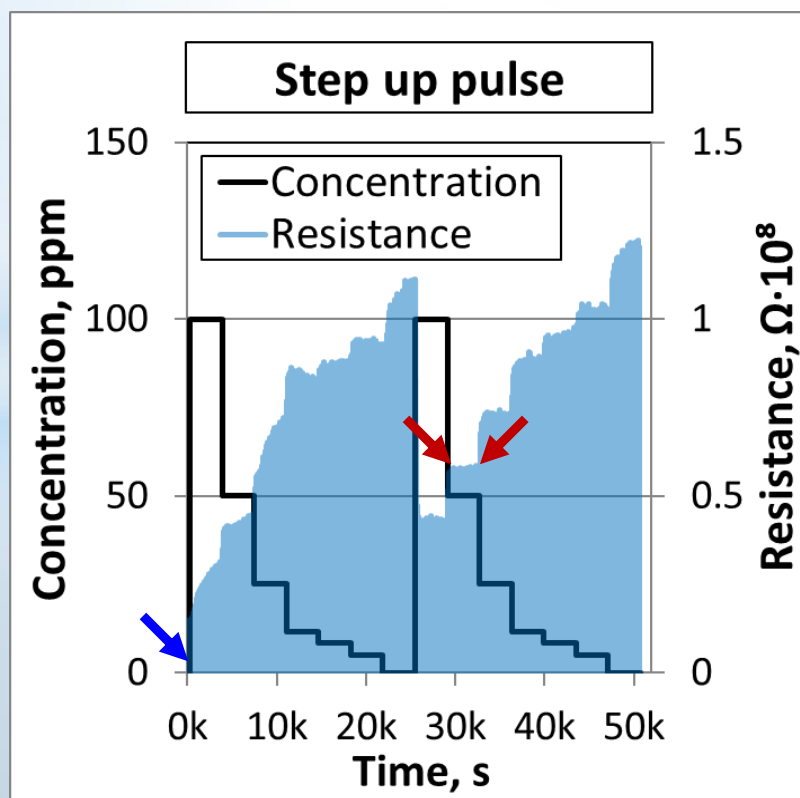


1. Преобразование данных в формат:
1 цикл динамики нагрева –
1 пример набора данных.
2. Замена значений сенсорного отклика
выше 10^{10} Ом
на фиксированное значение 10^{10} Ом,
и ниже 10 Ом
на фиксированное значение 10 Ом.

Вычислительный эксперимент

Подготовка и предобработка данных

Этапы подготовки и предобработки данных:

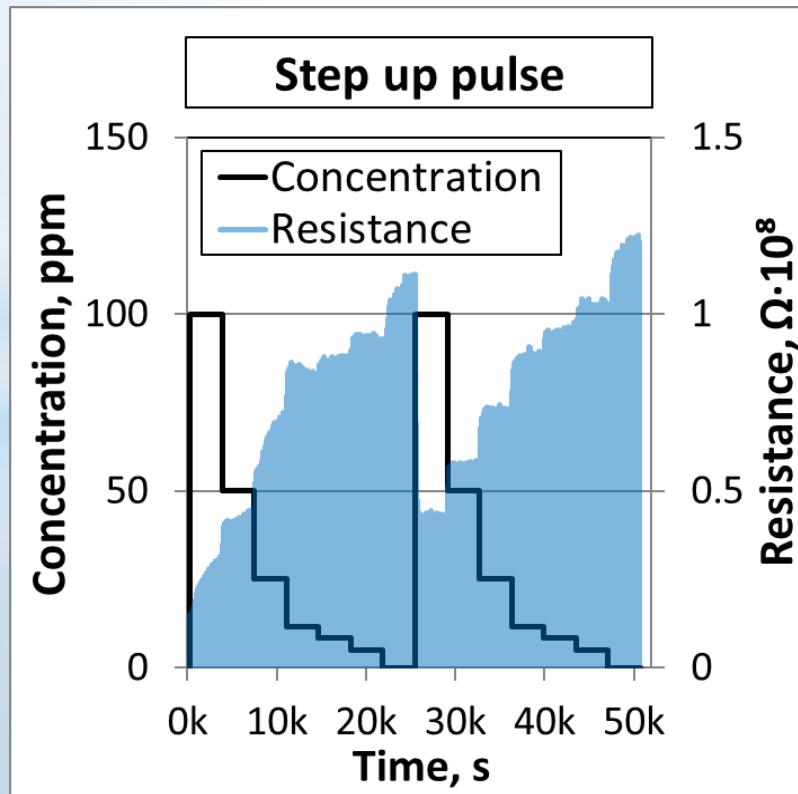


3. **Исключение** участков данных, где производилась продувка экспериментальной установки: **первые несколько циклов** в каждом отдельном эксперименте.
4. Исключение первых нескольких **циклов** **после изменения концентрации** и нескольких циклов непосредственно **перед изменением концентрации**.

Вычислительный эксперимент

Подготовка и предобработка данных

Этапы подготовки и предобработки данных:



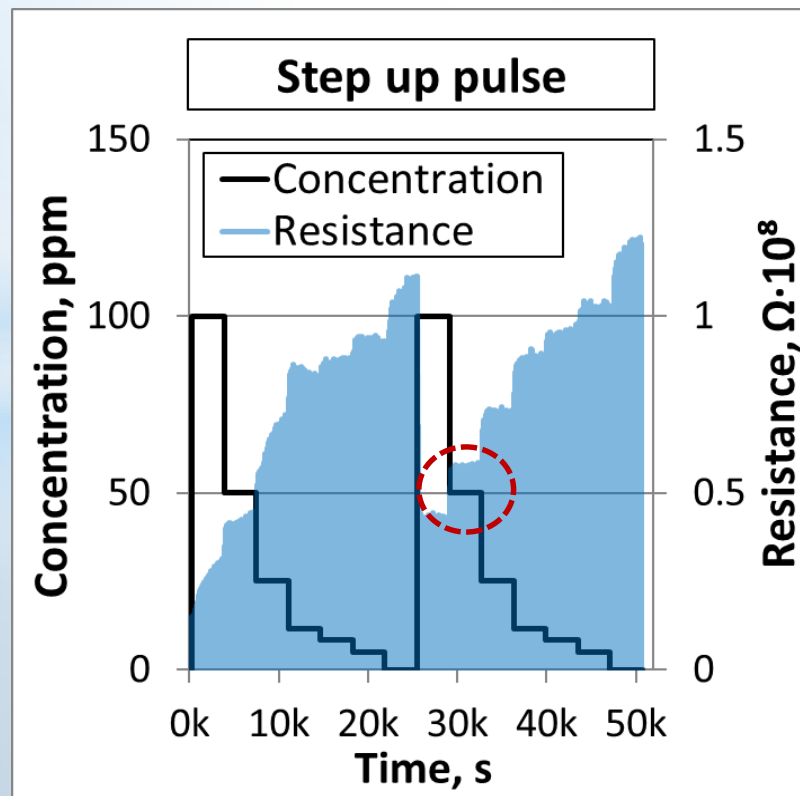
5. Логарифмирование значений сенсорного отклика.

6. Масштабирование значений концентрации газов на диапазон минимума-максимума.

Вычислительный эксперимент

Подготовка и предобработка данных

Этапы подготовки и предобработки данных:



7. Стратифицированное деление на поднаборы

Из каждой **области**
с **фиксированной концентрацией** газа:

- В тренировочный набор отбирались первые n циклов,
- Последующие m циклов – в валидационный набор,
- Последующие k циклов – в тестовый набор

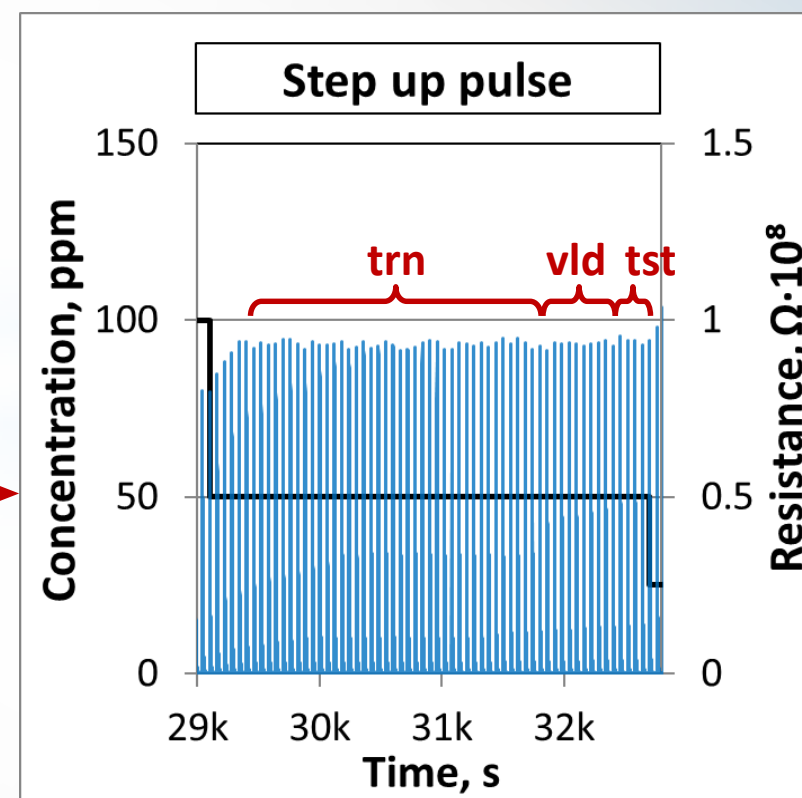
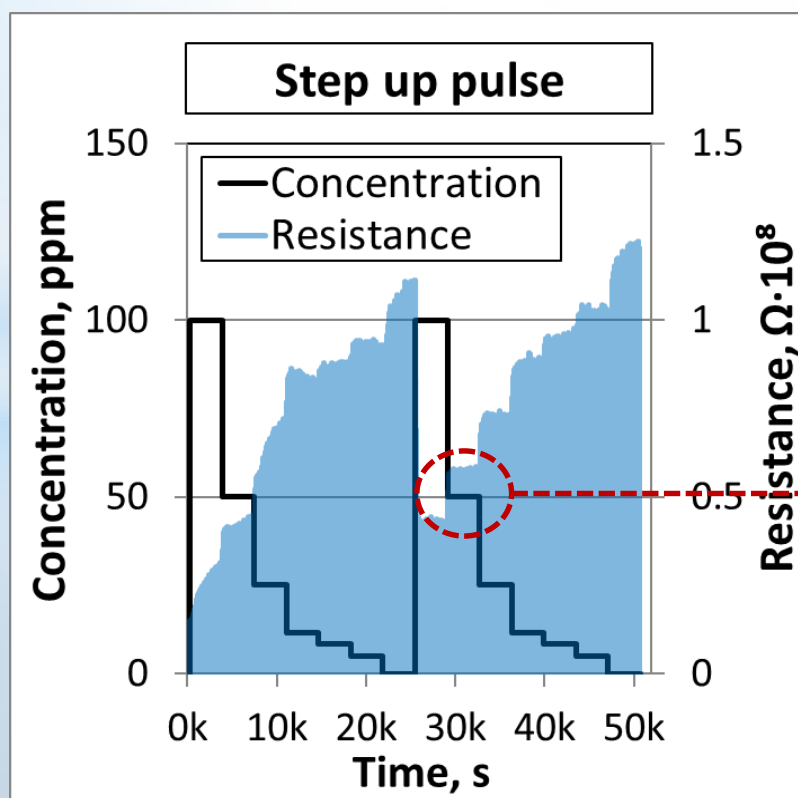
Затем процедура повторялась для остальных участков с фиксированной концентрацией.

Вычислительный эксперимент

Подготовка и предобработка данных

Этапы подготовки и предобработки данных:

7. Стратифицированное деление на поднаборы



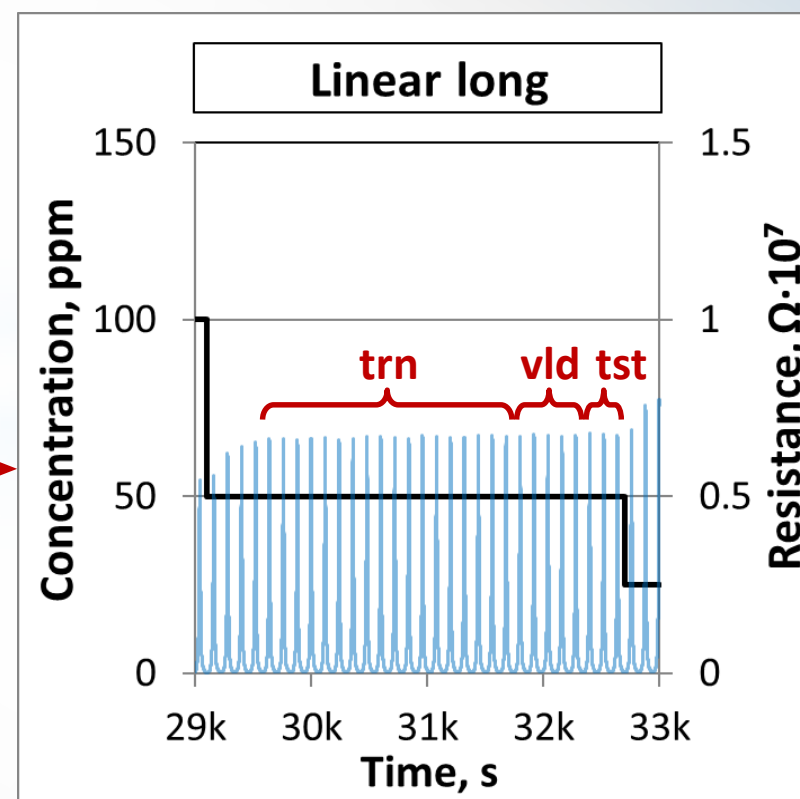
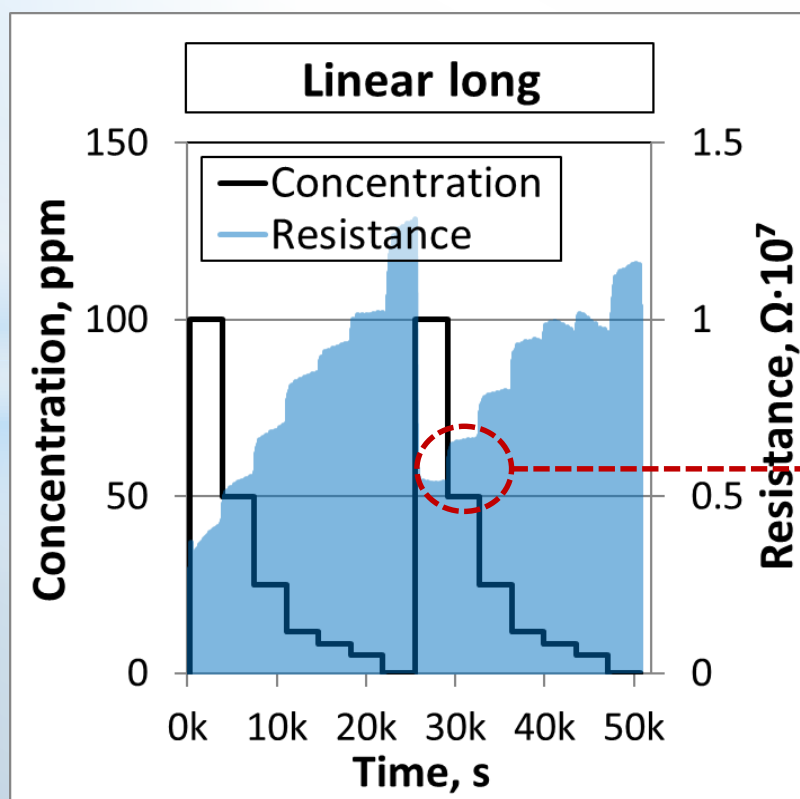
$$n:m:k = 40:10:5$$

Вычислительный эксперимент

Подготовка и предобработка данных

Этапы подготовки и предобработки данных:

7. Стратифицированное деление на поднаборы



$$n:m:k = 18:5:3$$

Вычислительный эксперимент

Методы машинного обучения

- ❑ Линейная и логистическая регрессия
 - Без регуляризации
 - С L1 или L2 регуляризацией
- ❑ Деревья решений
 - Случайный лес
 - Градиентный бустинг
- ❑ Многослойный персептрон

Все методы показали
качественно похожие результаты

Вычислительный эксперимент

Постановка задачи

Эксперимент №1

- ☐ Оценка **чувствительности** сенсоров
- ☐ Использование сенсоров по одному
- ☐ Отдельные наборы данных для каждого газа
 - 1 файл – 1 газ

Эксперимент №2

- ☐ Оценка **селективности** сенсоров
- ☐ Использование комплекта сенсоров
- ☐ Объединенные наборы данных
 - 1 файл – все газы

Результаты. Эксперимент №1.

Задача регрессии. Использование одного сенсора.

Качество решения (R^2) задачи регрессии (тривиальная модель $R^2=0$)

для различных методов машинного обучения на 1-й серии данных

LR-L1

Gas	Sensor No											
	1	2	3	4	5	6	7	8	9	10	11	12
CH ₄	0.922	0.942	0.923	0.907	0.939	0.585	0.921	0.946	0.258	0.357	0.855	0.833
CO	0.961	0.966	0.965	0.959	0.967	0.933	0.951	0.972	0.945	0.956	0.976	0.853
H ₂	0.984	0.978	0.968	0.938	0.943	0.935	0.957	0.938	0.317	0.930	0.917	0.755
H ₂ S	0.987	0.992	0.984	0.972	0.985	0.905	0.969	0.993	0.971	0.985	0.901	0.820
HCOH	0.943	0.949	0.957	0.943	0.963	0.937	0.960	0.962	0.953	0.951	0.811	0.887
NH ₃	0.926	0.967	0.957	0.923	0.933	0.853	0.960	0.962	0.967	0.961	0.656	0.804
NO	0.788	0.795	0.866	0.837	0.837	0.808	0.897	0.890	0.861	0.862	0.529	0.678
NO ₂	0.817	0.840	0.856	0.831	0.801	0.837	0.851	0.857	0.641	0.655	0.410	0.545
SO ₂	0.644	0.658	0.661	0.657	0.730	0.496	0.678	0.681	0.503	0.508	0.545	0.560

RF

Gas	Sensor No											
	1	2	3	4	5	6	7	8	9	10	11	12
CH ₄	1.000	1.000	0.999	0.999	1.000	0.844	1.000	1.000	0.994	0.983	0.966	0.921
CO	1.000	1.000	1.000	1.000	1.000	0.984	1.000	1.000	0.999	1.000	0.996	0.960
H ₂	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993	0.998	1.000	0.972
H ₂ S	1.000	1.000	1.000	1.000	1.000	0.954	1.000	1.000	1.000	1.000	0.998	0.938
HCOH	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999	0.888	0.931
NH ₃	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994
NO	0.886	0.891	0.934	0.886	0.937	0.861	0.870	0.926	0.988	0.967	0.554	0.487
NO ₂	0.876	0.966	0.885	0.946	0.914	0.972	0.934	0.863	0.956	0.955	0.594	0.623
SO ₂	0.822	0.856	0.836	0.894	0.908	0.715	0.812	0.774	0.965	0.959	0.539	0.664

MLP

Gas	Sensor No											
	1	2	3	4	5	6	7	8	9	10	11	12
CH ₄	0.977	0.983	0.978	0.972	0.977	0.647	0.981	0.983	0.408	0.466	0.883	0.793
CO	0.980	0.986	0.989	0.983	0.978	0.964	0.990	0.987	0.953	0.977	0.981	0.803
H ₂	0.994	0.991	0.986	0.984	0.968	0.985	0.988	0.984	0.469	0.919	0.938	0.708
H ₂ S	0.991	0.995	0.995	0.993	0.993	0.955	0.994	0.996	0.982	0.993	0.900	0.685
HCOH	0.980	0.989	0.984	0.985	0.978	0.966	0.990	0.991	0.973	0.974	0.880	0.913
NH ₃	0.982	0.992	0.985	0.984	0.985	0.959	0.987	0.991	0.988	0.989	0.816	0.629
NO	0.929	0.926	0.949	0.948	0.948	0.888	0.964	0.954	0.931	0.933	0.785	0.718
NO ₂	0.909	0.890	0.921	0.878	0.863	0.838	0.903	0.896	0.819	0.824	0.609	0.091
SO ₂	0.794	0.818	0.812	0.782	0.849	0.688	0.833	0.846	0.592	0.621	0.581	0.606

GB

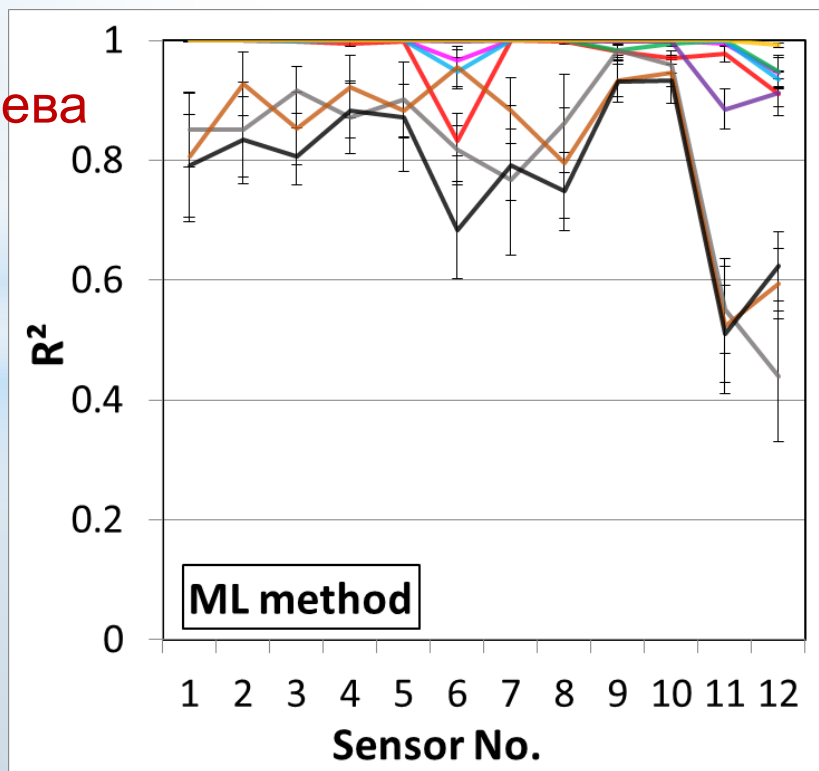
Gas	Sensor No											
	1	2	3	4	5	6	7	8	9	10	11	12
CH ₄	1.000	1.000	0.999	0.995	0.999	0.832	1.000	0.998	0.981	0.971	0.977	0.911
CO	1.000	1.000	1.000	1.000	0.999	0.967	1.000	1.000	0.998	0.999	0.994	0.946
H ₂	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.984	0.995	1.000	0.947
H ₂ S	1.000	1.000	0.999	1.000	1.000	0.948	1.000	1.000	1.000	1.000	0.999	0.935
HCOH	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000	0.999	0.999	0.885	0.913
NH ₃	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992
NO	0.850	0.851	0.917	0.871	0.901	0.818	0.767	0.861	0.984	0.960	0.550	0.439
NO ₂	0.806	0.927	0.853	0.922	0.883	0.955	0.883	0.795	0.933	0.946	0.523	0.594
SO ₂	0.791	0.834	0.806	0.883	0.872	0.683	0.792	0.748	0.931	0.934	0.511	0.623

Результаты. Эксперимент №1.

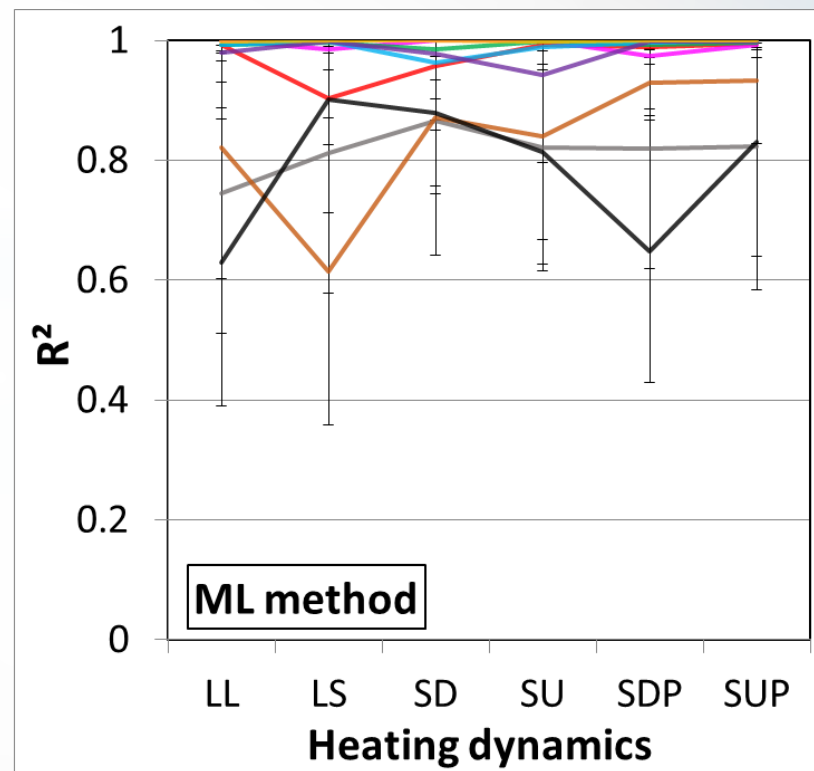
Задача регрессии. Использование одного сенсора.

Усредненное качество решения (R^2) по различным сенсорам и динамикам нагрева для градиентного бустинга на 1-й серии данных

Усреднение по
динамикам нагрева



Усреднение по
сенсорам



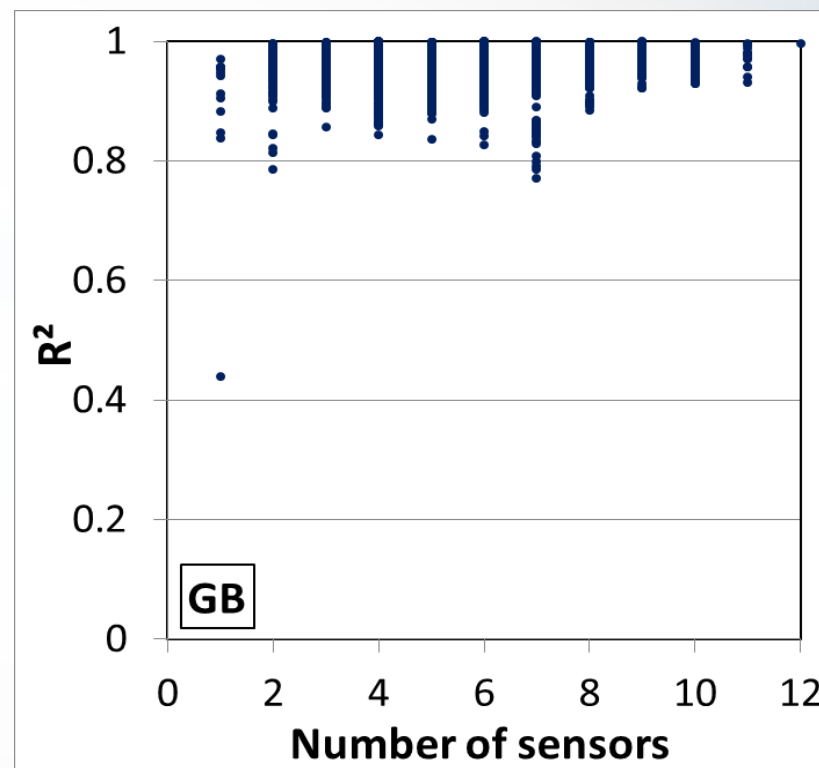
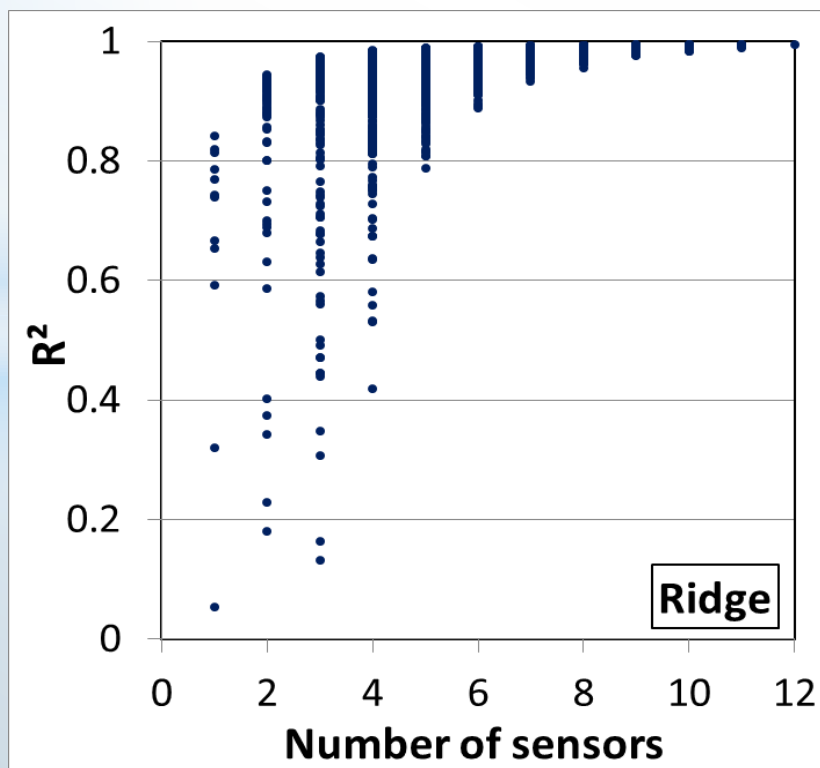
— CH₄ — CO — H₂ — H₂S — HCOH — NH₃ — NO — NO₂ — SO₂

Результаты. Эксперимент №2.

Задача регрессии. Использование комплекта сенсоров .

Качество решения (R^2) для всех сочетаний сенсоров.

Газ – CH_4 , линейная короткая динамика нагрева, 1-я серия данных.



Заключение

Выводы

- ❑ Наилучшая динамика нагрева – линейная длинная
 - ✓ Уменьшение длины (линейная короткая) динамики нагрева приводит к ухудшению качества решения.
- ❑ Динамики нагрева с повышением температуры показывают лучшие результаты, чем с понижением температуры.
- ❑ Сенсоры на основе SnO_2 в среднем показывают лучшие результаты, чем сенсоры на основе TiO_2 .
- ❑ Хуже всего определяются концентрации газов NO , NO_2 , SO_2

Заключение

Выводы

- ❑ При использовании **объединенных наборов** данных, содержащих **все газы**, наблюдается **ухудшение качества** решения регрессионной задачи
- ❑ Использование данных **нескольких сенсоров** позволяет **улучшить качество** решения по сравнению с использованием данных одного сенсора
 - Для **линейных методов** требуется **большее количество сенсоров**, используемых одновременно, для получения хорошего качества решения

Спасибо за внимание !