



Summary Data Gathering Report

Ali Abdollahian Noghabi

CS student at AmirKabir University

a.noghabi2002@gmail.com

Experiment 1) [Full Report](#)

I. 28 October until 4 November Fruit Pictures

double sensor in a box gathering data from a green banana

on 28 October banana was almost green maybe nearly 20% yellow - [Link](#)

on 29 October banana was 50% green and 50% yellow - [Link](#)

on 30 October banana was almost yellow and the head started turning black - No Picture Unfortunately 😞

on 31 October banana was nearly 10% black - [Link](#)

on 1 November banana was nearly 40% black - [Link](#)

on 2 November the banana was nearly 70% black - [Link](#)

on 4 November banana was fully spoiled that I had to throw it away (it smelled very bad)

II. 28 October until 4 November Charts

28 October Charts - [Link](#)

29 October Charts - [Link](#)

30 October Charts - [Link](#)

31 October Charts - [Link](#)

1 November Charts - [Link](#)

2 November Charts - [Link](#)

Hourly Charts (28oct-4nov) - [Link](#)

Experiment 2) Sensor 1 - isolated [FullReport](#)

I. 11 November until 16 November Fruit Pictures

one sensor in an isolated box gathering data from a 100% green banana
on 11 November banana was 100% green - [Link](#)

on 12 November banana was almost green maybe 20% yellow - [Link](#)

on 13 November banana was almost yellow maybe 20% green - [Link](#)

on 14 November banana was almost yellow maybe 10% green - [Link](#)

on 15 November banana was in perfect condition yellow and started turning black - [Link](#)

on 16 November banana still was in perfect condition yellow with some little black scratches - [Link](#)

II. 11 November until 16 November Charts

11 November Charts - [Link](#)

12 November Charts - [Link](#)

13 November Charts - [Link](#)

14 November Charts - [Link](#)

15 November Charts - [Link](#)

16 November Charts - [Link](#)

Hourly Charts (28oct-4nov) - [Link](#)

Experiment 2) Sensor 2

I. 11 November until 16 November Fruit Pictures

one sensor in a box with air conditioning gathering data from a 100% green banana
on 11 November banana was 100% green - [Link](#)

on 12 November banana was almost green maybe 20% yellow - [Link](#)

on 13 November banana was almost yellow maybe 20% green - [Link](#)

on 14 November banana was still 10% green but had some big black scratches - [Link](#)

on 15 November banana was 50% black - [Link](#)

on 16 November banana was 70% black - [Link](#)

II. 11 November until 16 November Charts

11 November Charts - [Link](#)

12 November Charts - [Link](#)

13 November Charts - [Link](#)

14 November Charts - [Link](#)

15 November Charts - [Link](#)

16 November Charts - [Link](#)

Hourly Charts (28oct-4nov) - [Link](#)



The results obtained from the experiments are as follows:

I. Concentration of VOC gas:

- The level of VOC gas doesn't act linearly in a non-isolated environment. It doesn't follow a steady and predictable pattern that allows us to predict it based on the fruit's ripening time. Instead, the total amount of this gas goes up as the fruit ripens. [Chart](#)
- However, in an isolated environment in another experiment, I approached a more linear response of this gas. [Chart](#)

II. The reaction of the VOC gas and other factors

The reaction of the VOC gas has a relationship with the concentration of certain other gases.

- **Relationship with CO₂:** It has a direct relationship with this gas, and its reactions to environmental factors are precisely correlated. [Chart](#)
- **Relationship with RawH₂ and RawEthanol:** It has an inverse relationship with these gases, and their reactions to environmental factors are exactly opposite. [RawH2Chart](#) [RawEthanolChart](#)
- **Relationship with temperature and humidity:** there is no readily observable specific correlation between these two factors and VOC gas. However, it may be possible to arrive at some conclusions with reliable data and machine learning. [TemperatureChart](#) [HumidityChart](#)



Challenges and Tasks:

I. Gathering Reliable Data:

- I ran into some issues during the second experiment because of power outages, which messed up the sensor calibration when it was reset. To fix this, I'm using a continuous power source like a power bank for the sensor. However, I'm still trying to figure out a way to calibrate the sensor consistently so that it can continue collecting data from where it left off after a reset.
Up to this point, I've researched and found the correct method for calibrating the sensor, and it's scheduled to be tested in the next experiment. [powerOutages](#)
- The impact of the type and brand of sensors on data collection
As you can see, I've used two different brands of sensors for humidity and temperature, and it's clear that they didn't provide similar data, even though they were in the same environment.
- I've been looking into getting CO2 / RawH2 / RawEthanol gas sensors to make sure SGP30 sensor data is trustworthy. And planning to compare the data from other sensors that measure gases to validate the results.
- I'm not sure whether we should do experiments in an isolated environment or a regular one. There were noticeable differences in the results between these two environments in the second experiment, as seen in the pictures and graphs.
- I'm also working on setting up a notification system for any weird data points from the sensor. Sometimes, the sensor might give us incorrect data under certain conditions, like when it measures too quickly or during electrical fluctuations, and we need a way to deal with these strange data points.



II. Running Experiments in Different Conditions:

- experiments in both isolated and non-isolated environments.
- experiments with and without fruit to compare and validate the dataset.
- testing in various environmental conditions, like a relatively stable setting to mimic a supermarket environment and less stable conditions to simulate transportation stages.

III. Figuring Out Where This Project Is Going:

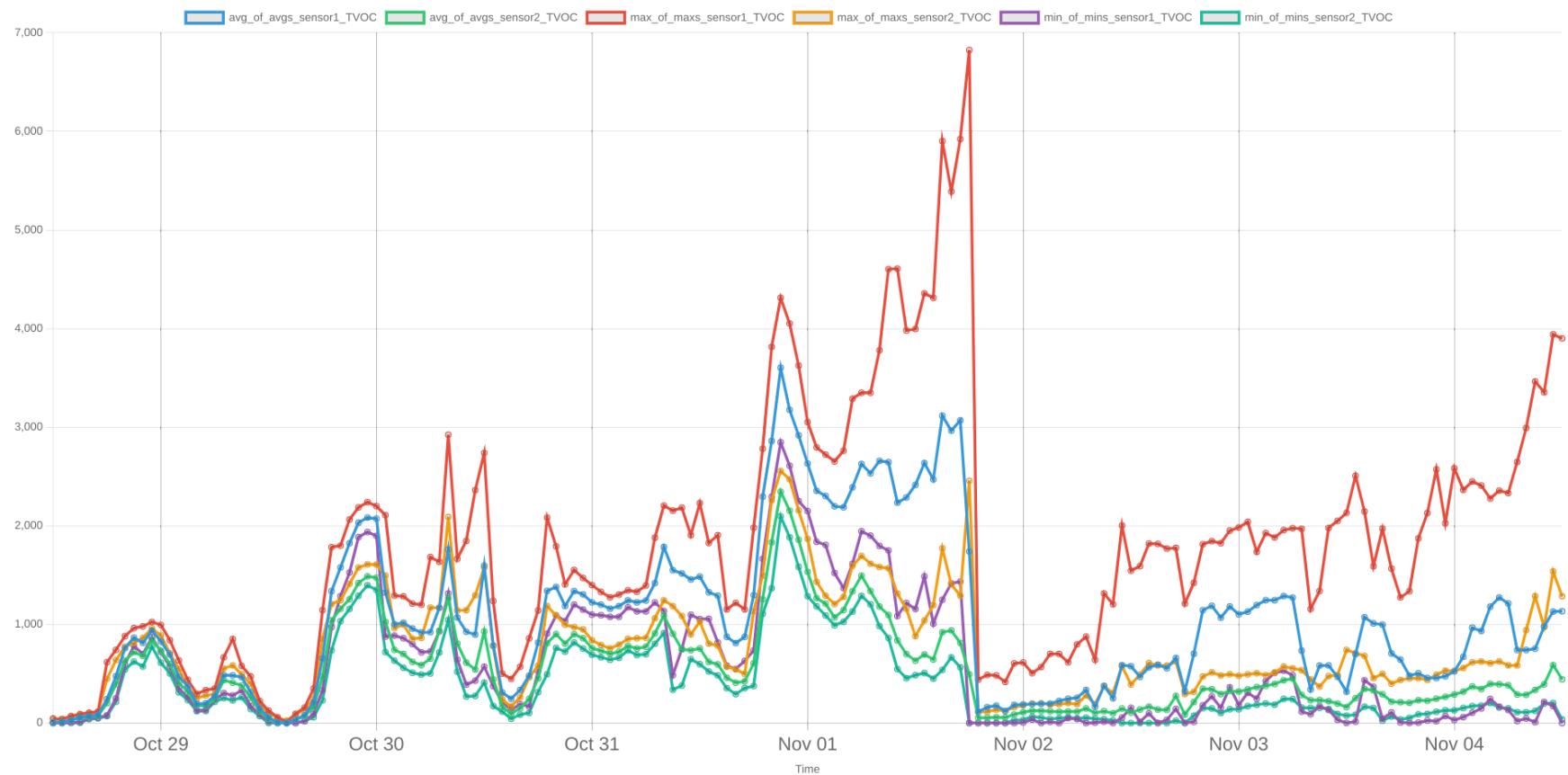
- What are the main questions we're trying to answer?
- checking how reliable our collected datasets are and exploring ways to validate them.
- are the datasets we've gathered suitable for machine learning applications?



Example chart of not linear behavior

Hourly Charts 28 October until 4 November

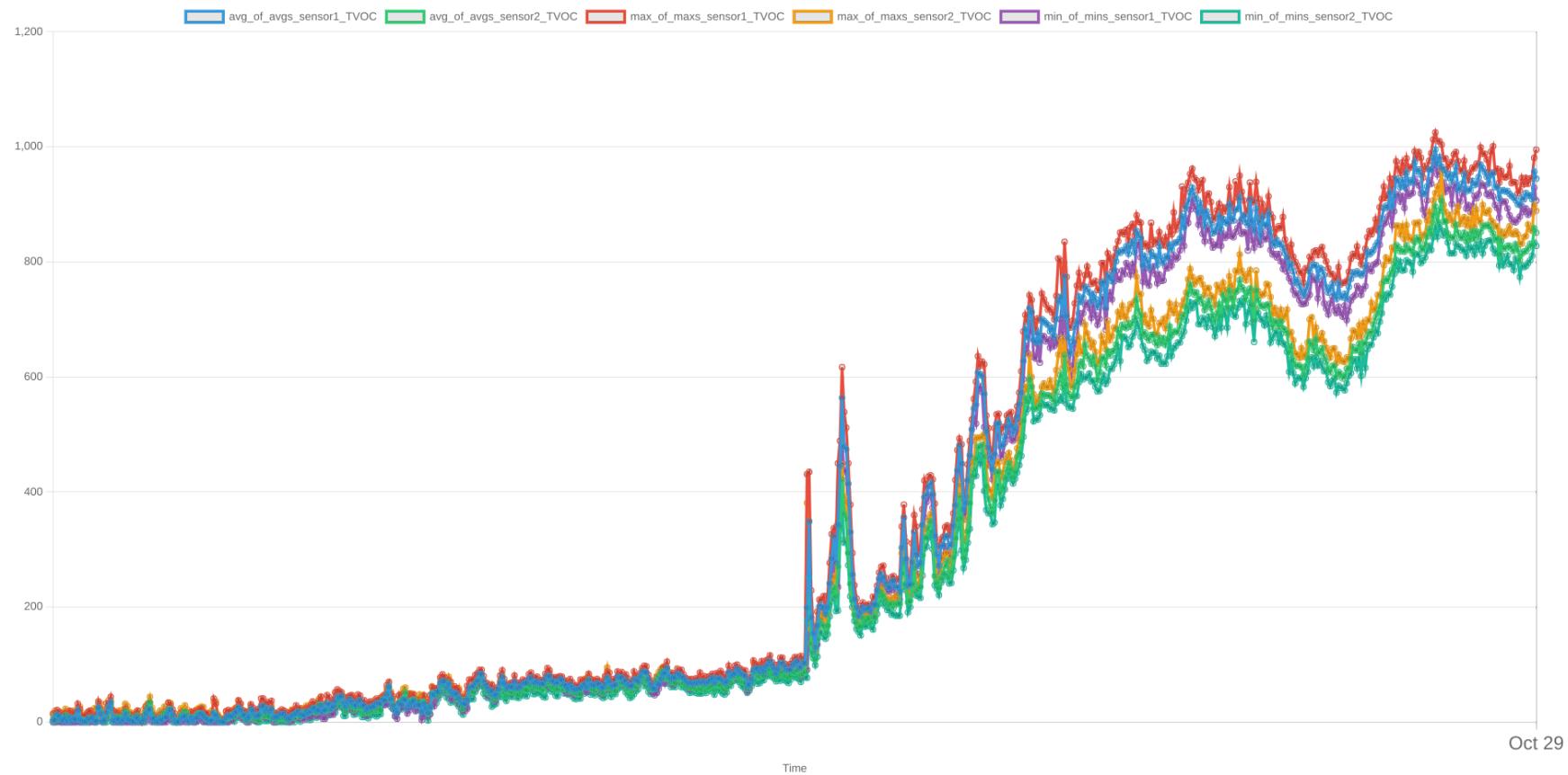
Avg/Max/Min of VOC sensor1 and sensor2





28 October Charts

Avg/Max/Min of voc sensor1 and sensor2

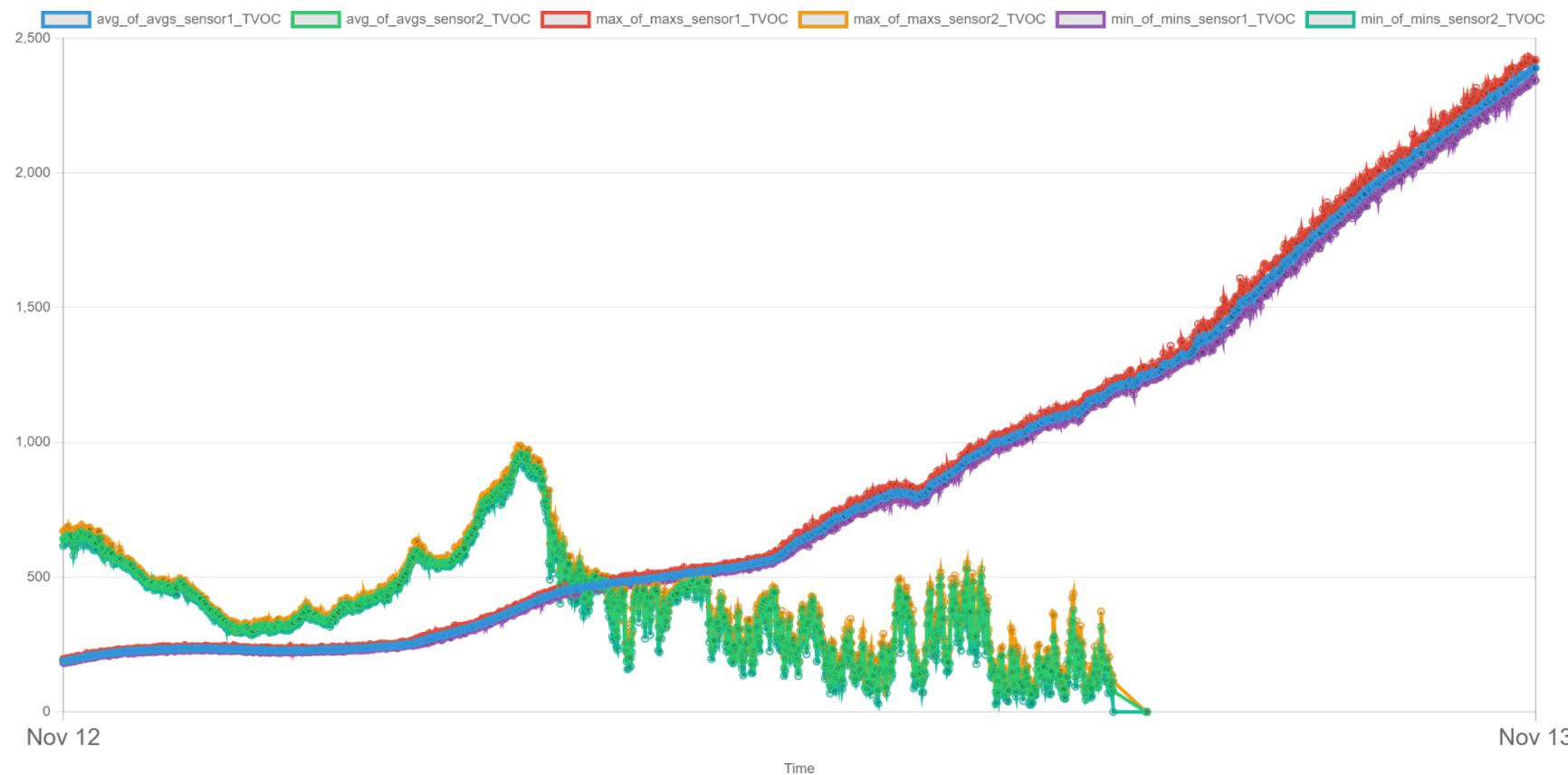




Example chart of linear behavior

12 November Charts

Avg/Max/Min of voc sensor1 and sensor2

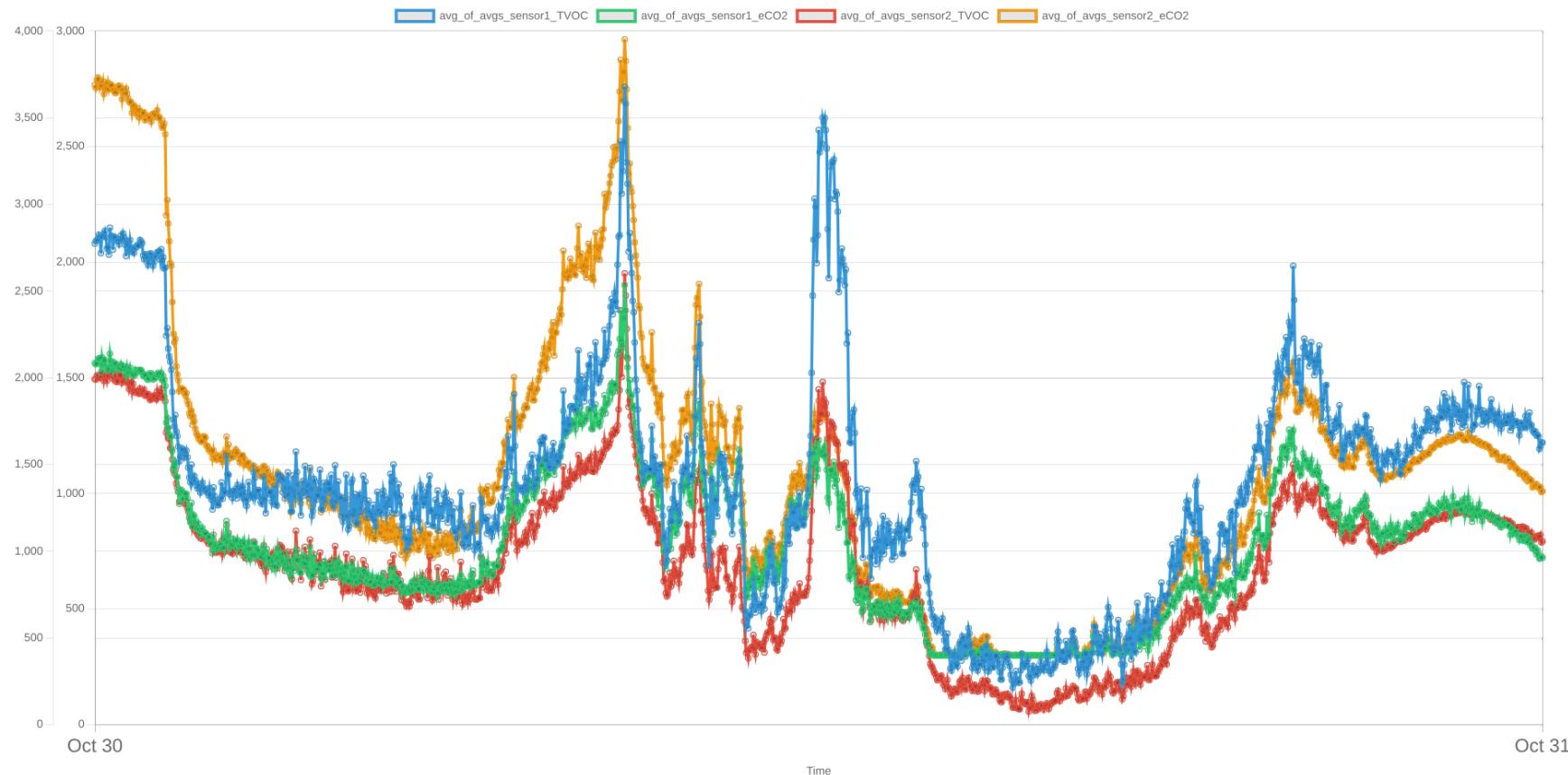




Example of VOC and CO₂ behavior

30 October Charts / CO₂

Avg of voc & Avg of co2 sensor1 and sensor2

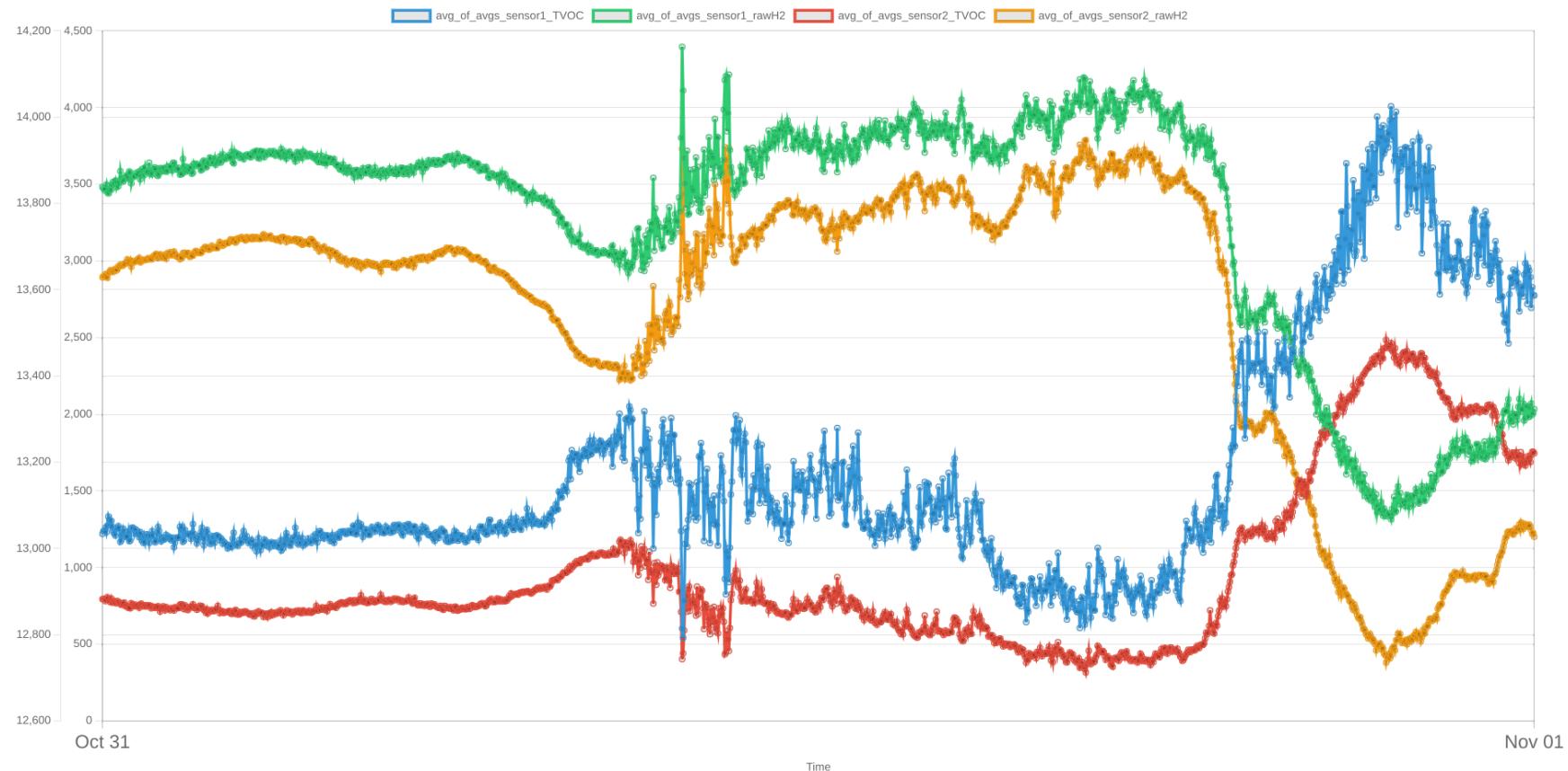




Example chart of VOC and RawH2

31 October Charts / RawH2

Avg of voc & Avg of rawh2 sensor1 and sensor2

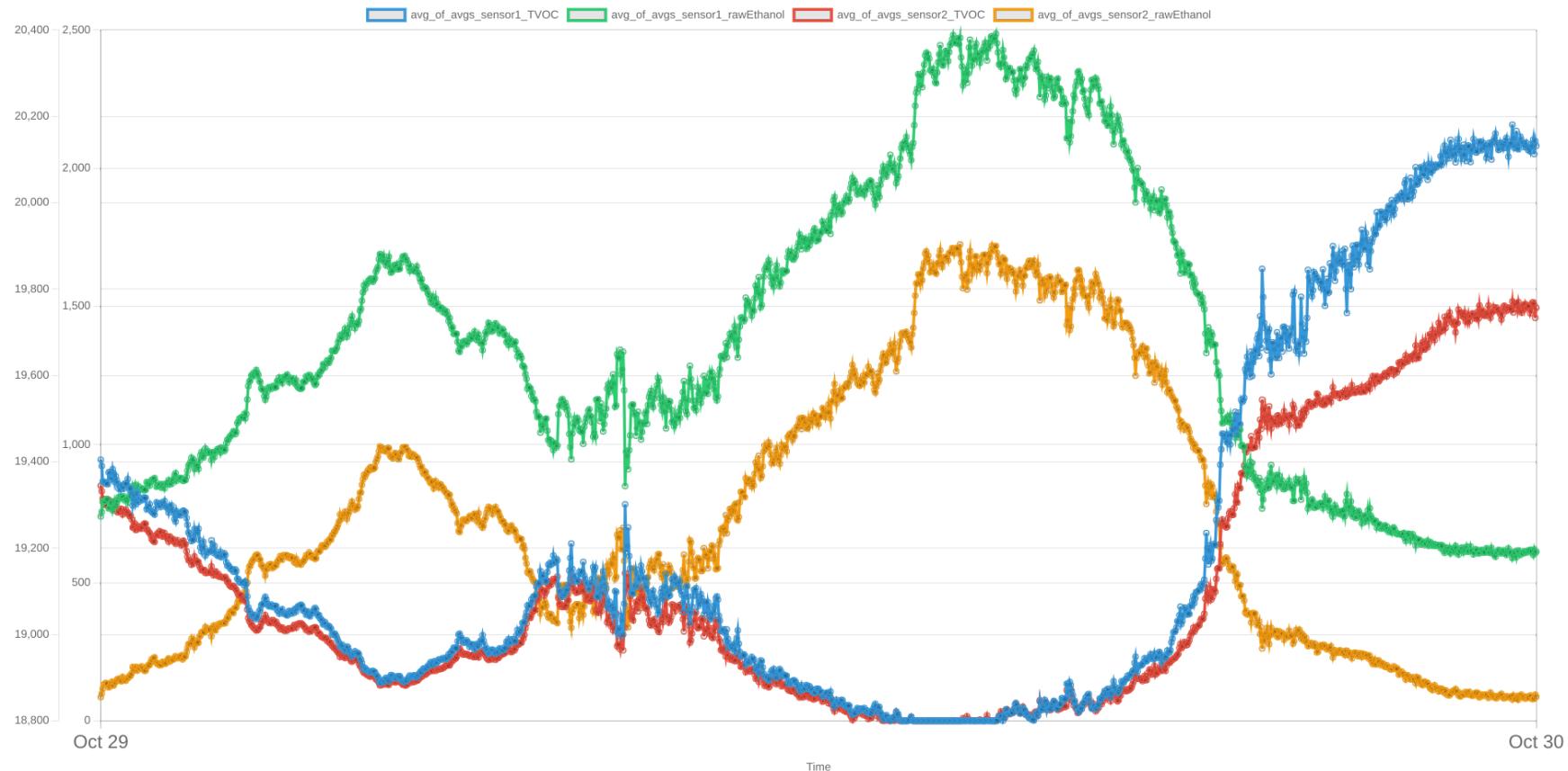




Example chart of VOC and RawEthanol

29 October Charts / RawEthanol

Avg of voc & Avg of rawEthanol sensor1 and sensor2

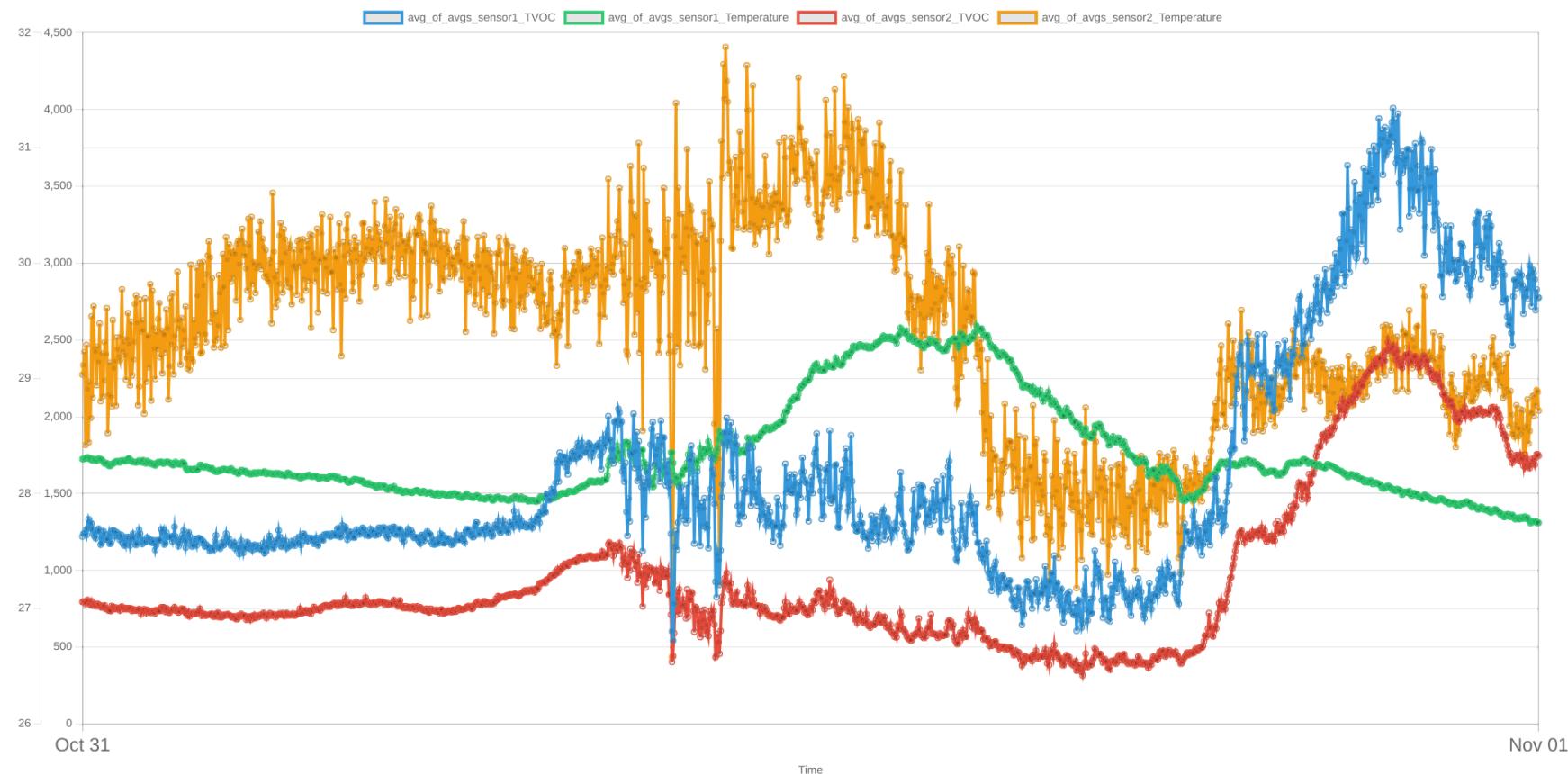




Example chart of VOC and Temperature

31 October Charts / Temperature

Avg of voc & Avg of Temperature sensor1 and sensor2

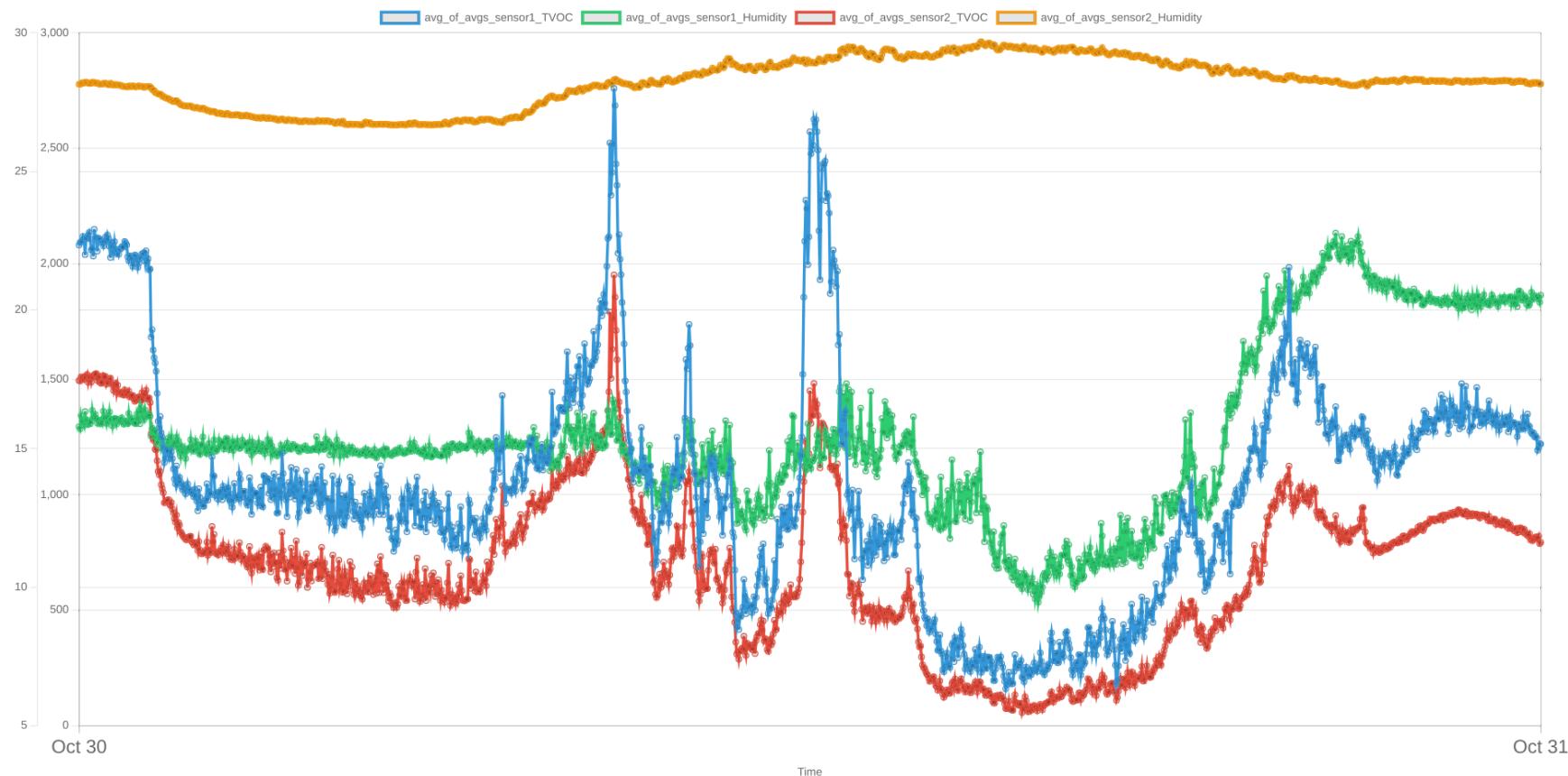




Example chart of VOC and Humidity

30 October Charts/ Humidity

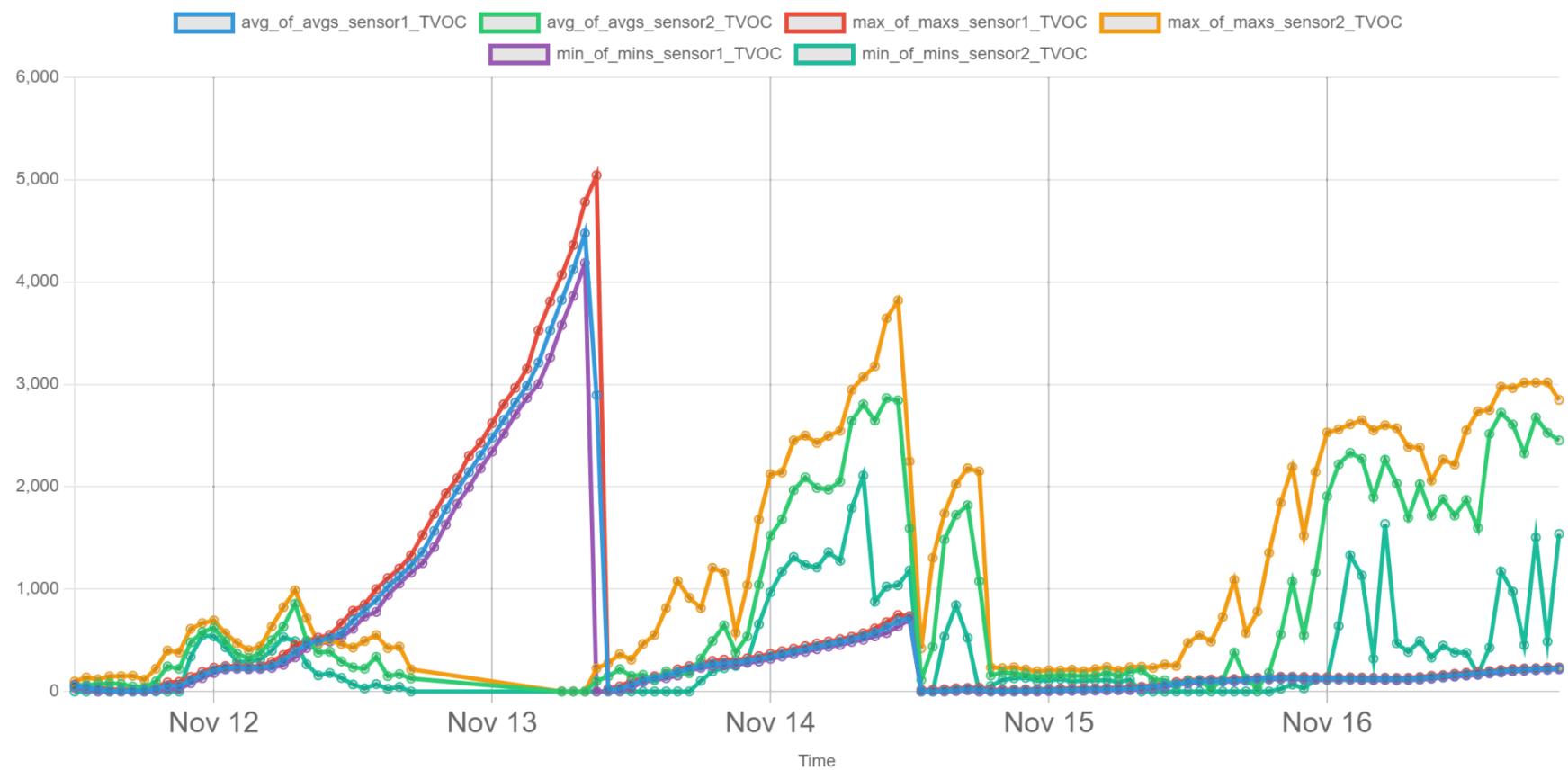
Avg of voc & Avg of Humidity sensor1 and sensor2





Power Outages and Simulate

Hourly chart from 11 November until 16 November





Power Outage Simulated

On November 13th at 13:00 and November 14th at 18:00, there was an electricity outage that affected sensor data gathering. Subsequently, I simulated the experiment by intentionally cutting off the electricity myself. The results indicated that data gathering is indeed affected by an electricity cutoff. Following the cutoff, the sensor data exhibited a significant decline and failed to capture the peak observed before the cutoff.

