TROPOMI Project Goal

# Intro

23-04-20 smart systems engineering master students and Ronald van Elburg discussed goals for the TROPOMI project.

# Topics described by students:

Zulkad:

Using algorithms to detect gas leakages and verify with current knowledge.

Mei Tao:

Use historical data to validate our model. Use this model to predict future leakages

Samir:

Apply algorithm on tropomi images with known higher concentration of methane, on chinese pipelines. Create a report of found results

Natanael

Start small and see how well that works. After that go analyze data from China

Joep

Reproducing their work and try to improve the accuracy

Dirk

Using a ML classification algorithm to detect a leakage.

Predict gas leakage increase early on by analyzing the trends in historical data from identified gas leakages.

Charmy

Start small and see how well that works. After that go analyze data from China and possibly India.

Akshit

Combine TROPOMI data with visual? data to flag areas that might have leakage. Check for spikes by comparing historical data with new data.

Abhirup

Reduce the need of using drones by improving the analyzation by satellites.

# Goal

Our goal is to create python software that automates the detection of CH4 leakages and provides (daily?) reports with new leakages and concerning trends. Historical TROPOMI data (with labels) is used to train a (classification) model in identifying/detecting gas leakages and possible developing leakages. The software of the user interface has an interface for sensor system management (choosing functions, changing points of views). System requirements are analyzed to identify applicable sensor fusion architectures. The data acquisition is analyzed to construct a representational format (PPM? Size of leakage?).

# Sub-goals

Find suitable algorithm for detection of gas leakage

Reading in the TROPOMI data

Choose to focus on leakages or premature leakages

Choose to identify or predict

Train the model

Evaluate the model

Develop output of results (API, UI, CSV, interpretation of data in UI)

[tip: swamps are good sources of methane?]

6. We apply data sensor fusion by fusing geographical information such as object data (oil/ gas plants or industrial plants, factories, swamps etc.) with methane concentrations.

7. We create a set of requirements for the above assessment aspect;

8. Construction of a common way of representing the output of data processing: which maps shall we use (Map of China, but decide on lay-out of map), how do we represent the areas with higher methane concentrations (i.e. how to draw border lines), and link it with real geographical locations such as a gas field or industrial plant (if TROPOMI data resolution allows this);

9. Fusion across time is done by comparing periods of time (i.e. TROPOMI data from 10-05-2020 with 15-05-2020), while spatial alignment is considered, since we will investigate relation between identified areas with methane concentration;

10. Analyze how the TROPOMI data is stored, and how certain area or time interval of interest can be chosen. Look at how the interface works, and look at standards for data representation. People can understand how to use it. Wants to see we considered applying standards. Can also lead to negative advice for applying standard.

11. Looking at geographical locations which cause higher concentrations of methane. Match the location with the possible cause, i.e. an industrial plant causing higher concentrations, or some other phenomena or cause. How to make sure correct interpretation gets to users: User interface shows and produces correct interpretations for end user. These are easy to understand by most of humans.

12. Application of mathematical data fusion model, for higher concentrations of methane + geographical location and type of region (i.e. industrial plant or factory). Currently unknown which mathematical data fusion model. Proof from other work can be used.

13. Device virtualization: small report for portfolio…

Also propagation error + unscented Kalman filter?

Should we provide portfolio per group or individually?

If the portfolio is individual, should we only insert the parts that were done individually?

When we need to validate the data, we need to have labeled data. IS this provided (i.e. from professor Abe/Yu)?

# Notes

Notes Zulkad: I recommend that we first read the s5p mission performance center methane readme

There is algorithm of getting CH4 information from raw data

Their algorithm is not realtime process

I think its a nice definition, I would add somewhere "programming in Python"

Wala’a comment: Rubric is not clear. It looks like it doesn’t fit with the project. Which points are covered by the practicals and which by the project?