Detecting ship plumes using NO_2, SO_2 and HCHO

Ruben Ahrens | Leiden

27-01-2025

Universiteit Leiden
The Netherlands

Supervisor 1: Prof. Fons Verbeek
Supervisor 2: Dr. Cor Veenman

Introduction: Emissions from ships

- Health issues
  - Respiratory issues: NO_2, SO_2
  - Cancerous: HCHO
- Exacerbate climate change
  - Greenhouse gases: CO_2, CH_4

Most affected: Marine life, People living in coastal regions

Photo by Ian Taylor on Unsplash

Introduction: Regulations

- IMO2020
  - Limits SO_x content in fuel
- MARPOL ANNEX VI
  - Limits NO_x emissions
- Compliance monitoring
  1. ~~Self reporting~~
  2. ~~Sensors under bridges or aircraft (sniffers)~~
  3. **Remote sensing**

Photo by Ian Taylor on Unsplash

Sentinel-5P: TROPOMI

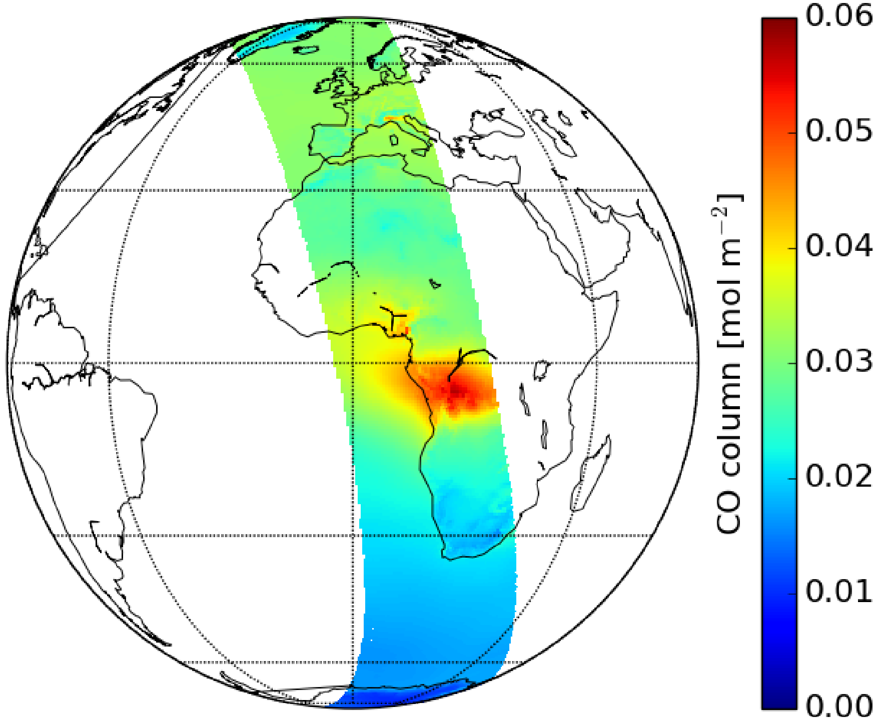
- Sentinel-5P = satellite
- TROPOMI = instrument

- Measures gases
  - **Nitrogen dioxide (NO_2)**
  - **Sulphur dioxide (SO_2)**
  - **Formaldehyde (HCHO)**
  - Carbon monoxide (CO)
  - Ozone (O_3)
  - Methane (CH_4)

(image from KNMI)

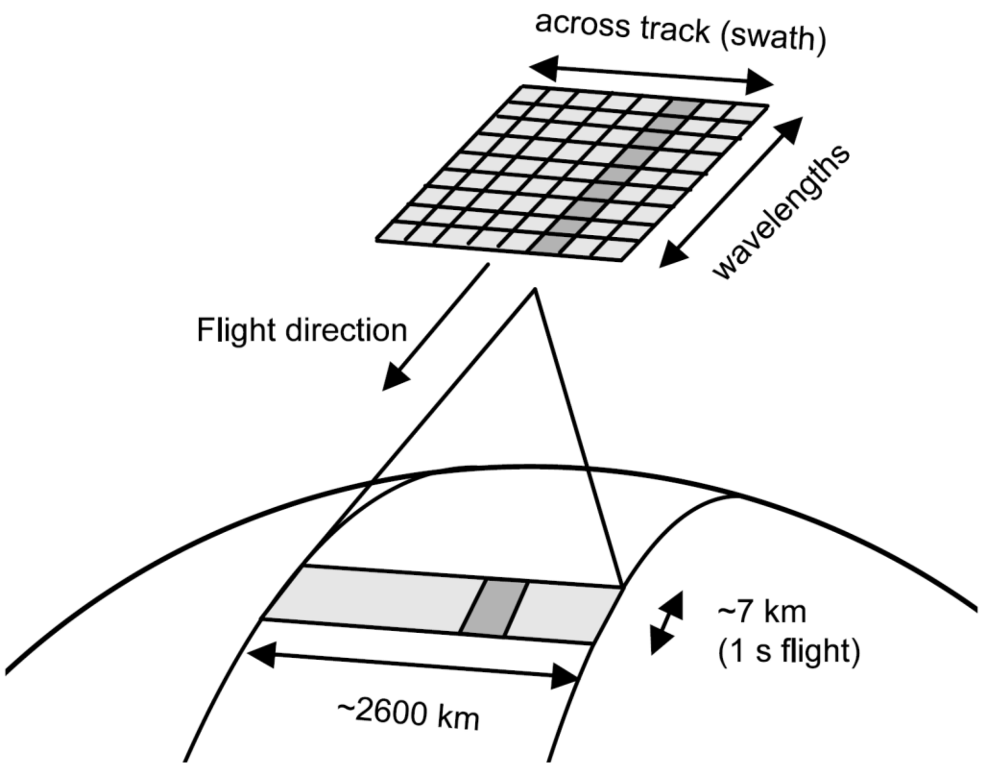
TROPOMI Data

CO measurements of one TROPOMI orbit



(image from Landgraf, 2016)

Resolution: 5.5km by 3.5km



(image from NO_2 ATBD, 2023)

Ship locations: AIS

- Database of (historic) ship locations
- Used to locate ship's emissions in TROPOMI data
- Mandatory collection
  - Passenger ships
  - Ships over 300GT
- Ship location
  - Latitude x_{ship}
  - Longitude y_{ship}
  - Timestamp t_{ship}

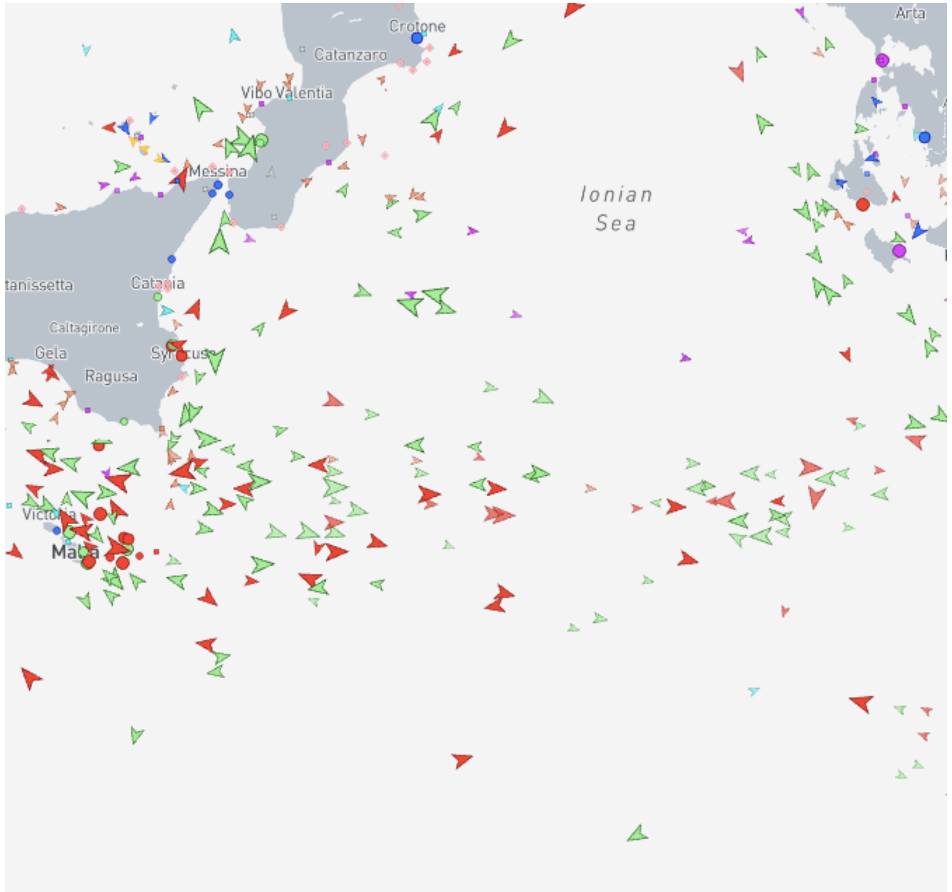
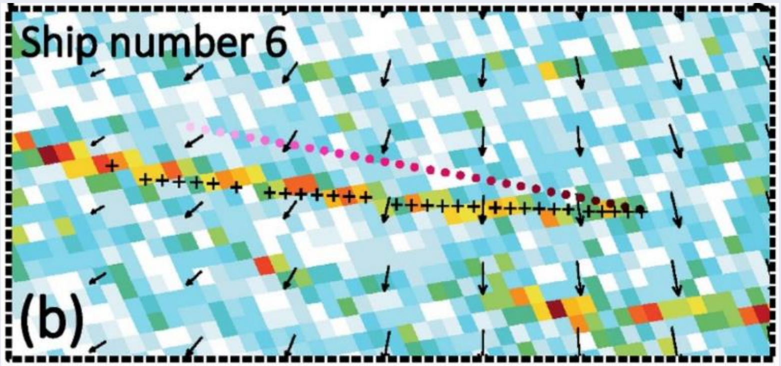


Image from marinetraffic.com

Related work

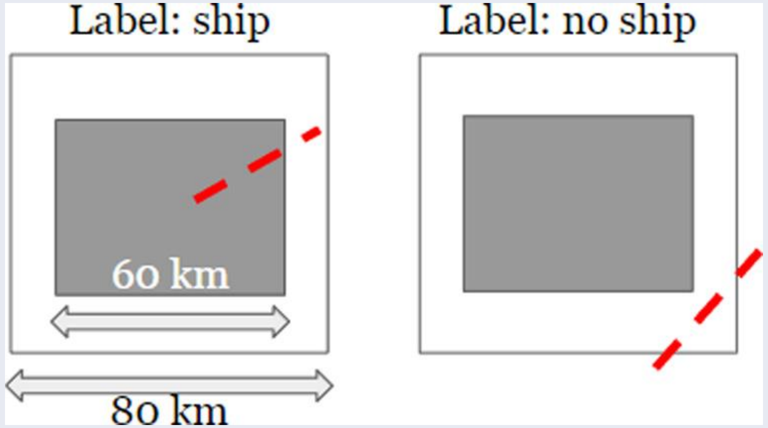
(Georgoulias, 2020)

Found individual ship plumes in TROPOMI NO_2 data

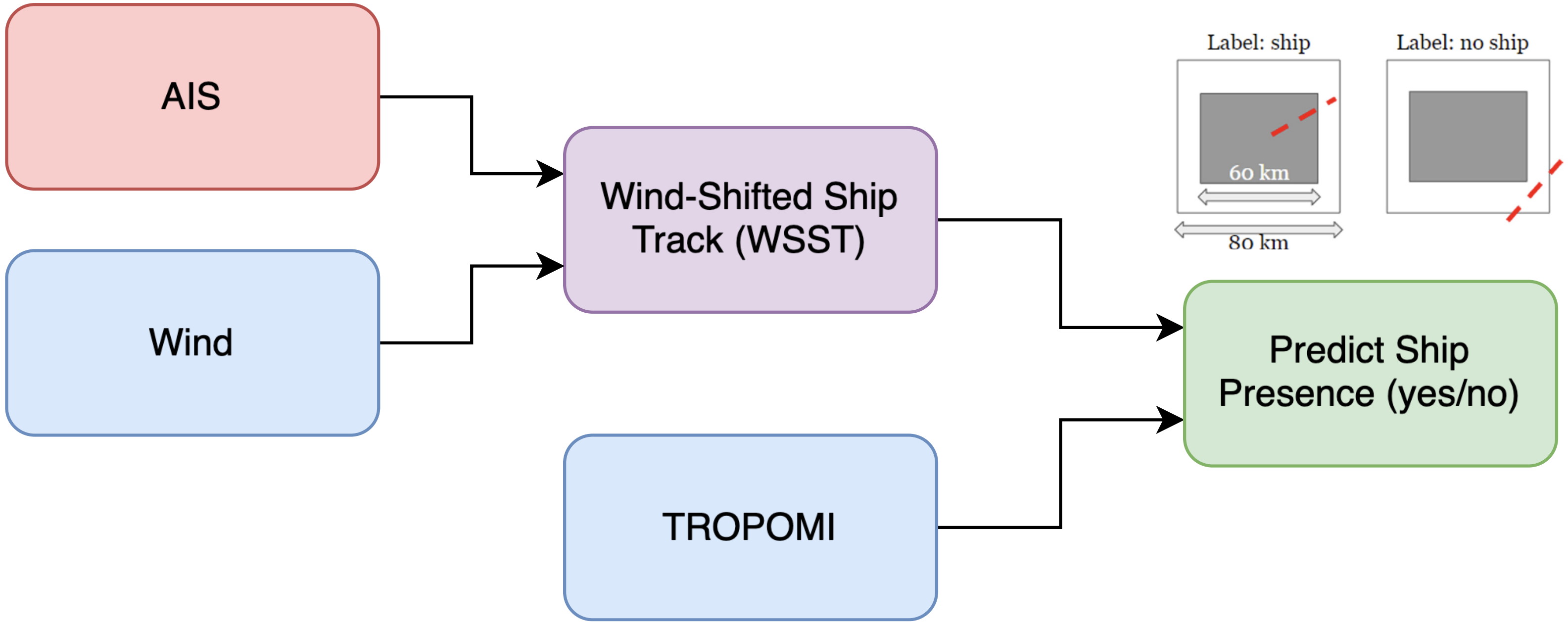


(Kurchaba, 2024)

Automatic ship plume detection using TROPOMI NO_2 data

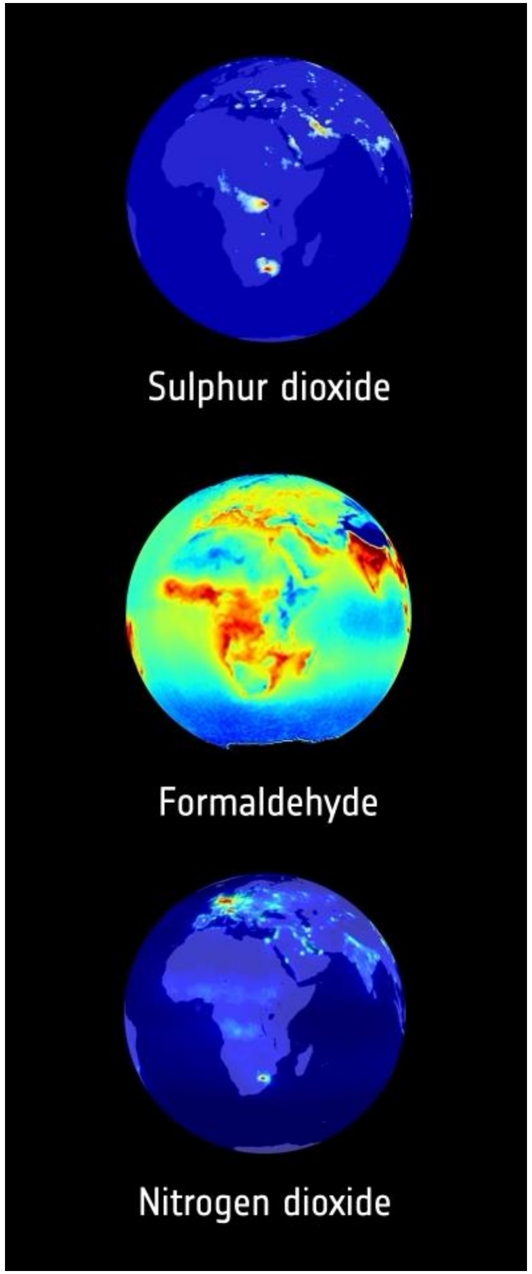


Simplified overview



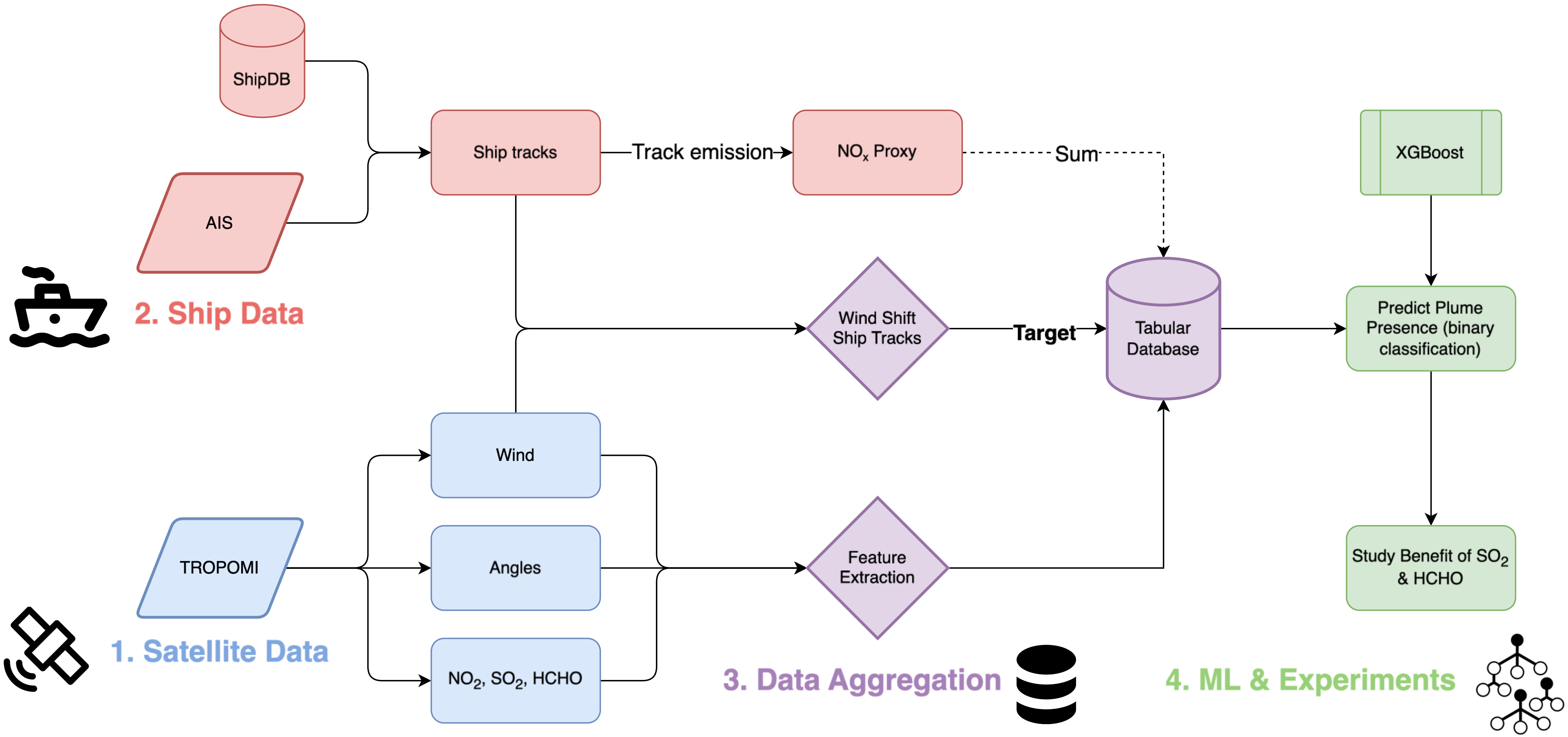
Research Questions

1. How does the inclusion of **SO_2** and **HCHO** data, alongside **NO_2** data, improve the detection accuracy of ship plumes using data from a single TROPOMI orbit?
2. Can ship plumes be detected based on **SO_2** or **HCHO** data from a single TROPOMI orbit?



(image from ESA)

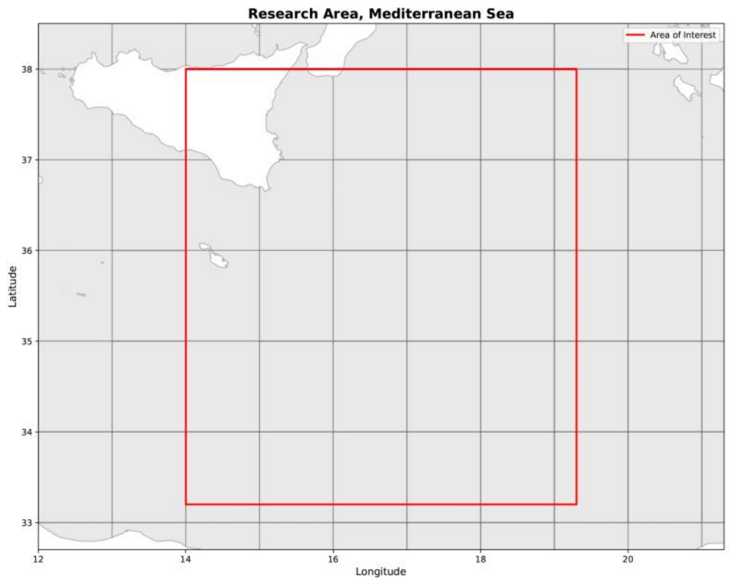
Method

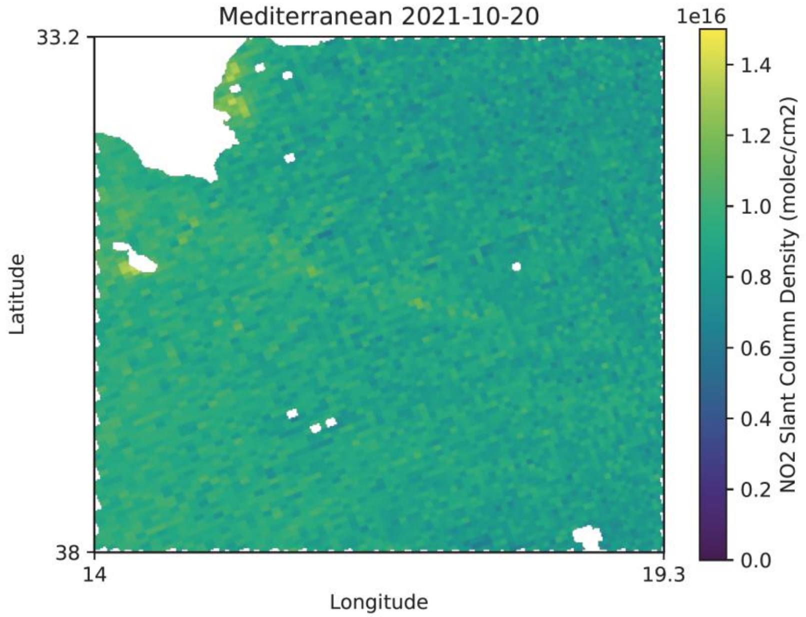


ShipDB
AIS
Ship tracks
Track emission
NO_x Proxy
Sum
XGBoost
2. Ship Data
Wind Shift Ship Tracks
Target
Tabular Database
Predict Plume Presence (binary classification)
Wind
TROPOMI
Angles
Feature Extraction
Study Benefit of SO_2 & HCHO
1. Satellite Data
NO_2, SO_2, HCHO
3. Data Aggregation
4. ML & Experiments

Acquiring Satellite Data

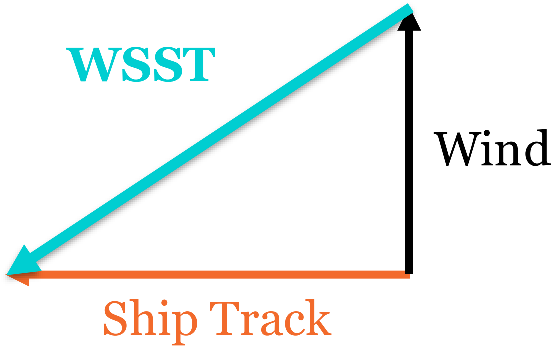
- Research Area: Mediterranean
- Time period: Jan 2020 - Jul 2022
- GUI
  - Copernicus Browser
  - EO (Earth Observation) Browser
- Automation
  - Catalog API (for searching for orbit files and their metadata)
  - S3 API (for bulk downloading orbit files)
- 740 orbits





Processing Ship Data

- Ship track
  - Ship locations from $t_{overpass}$ - 2h -> $t_{overpass}$
- WSST (Wind-Shifted Ship Tracks) using:
  - Wind speed (x & y axis)
  - Δt: ship location time and overpass time
- Filter out
  - Minimum length: 90m
  - No private or military ships



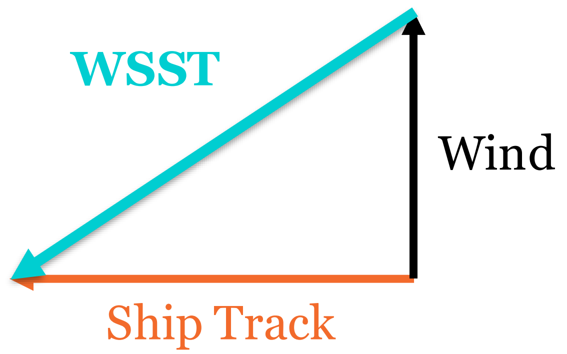
$$\Delta t = t_{\text{S5P}} - t_{\text{ship}}$$
$$x_{\text{plume}} = x_{\text{ship}} + \frac{V_{\text{zonal}} \cdot \Delta t}{D}$$
$$y_{\text{plume}} = y_{\text{ship}} + \frac{V_{\text{meridional}} \cdot \Delta t}{D}$$

(WSST calculation)

NO_x Proxy

$$E_i = l_i^2 \cdot v_i^3$$

$\boldsymbol{l_i}$ is length (m)
$\boldsymbol{v_i}$ is velocity (m/s)
of ship i

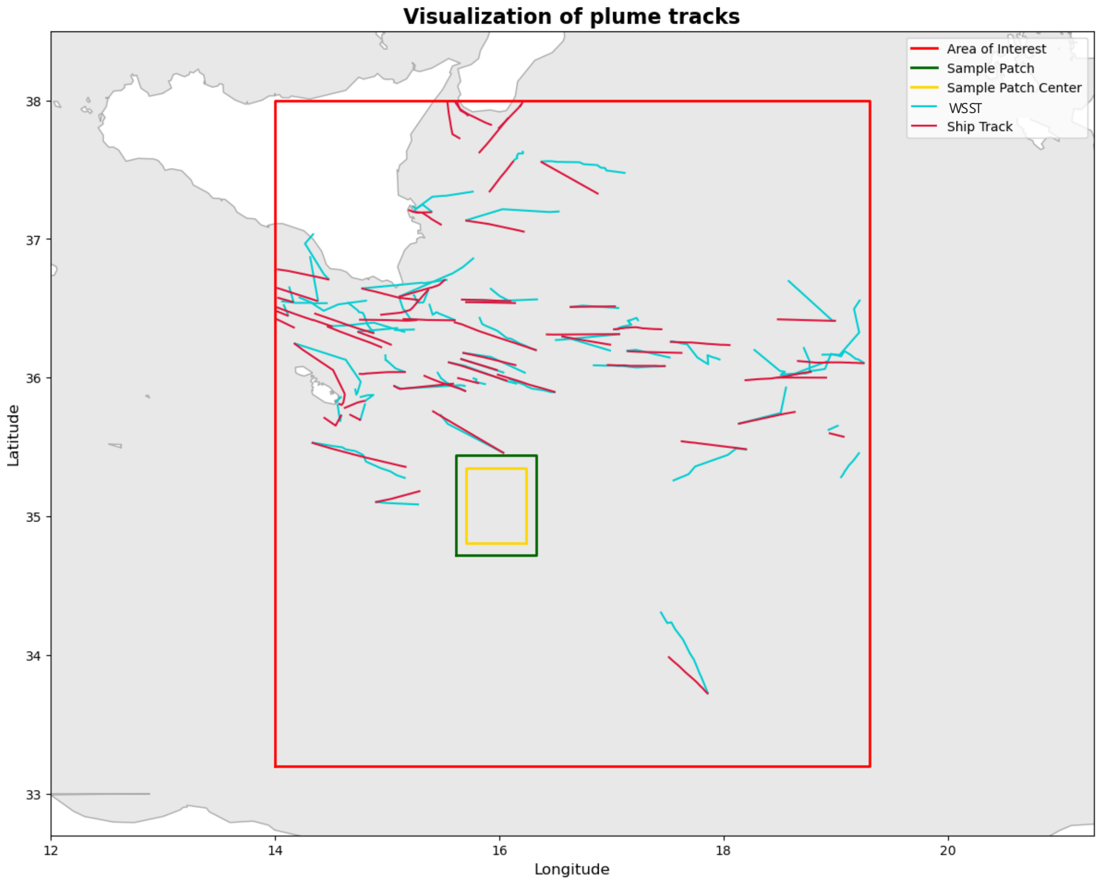


- Introduced by (Georgoulias, 2020)
- Used in other works e.g. (Kurchaba, 2024)

- Estimation for the amount of NO_x emitted by a ship.
  - Needs only AIS data

- Test detection performance for different expected emissions

Building Tabular Database



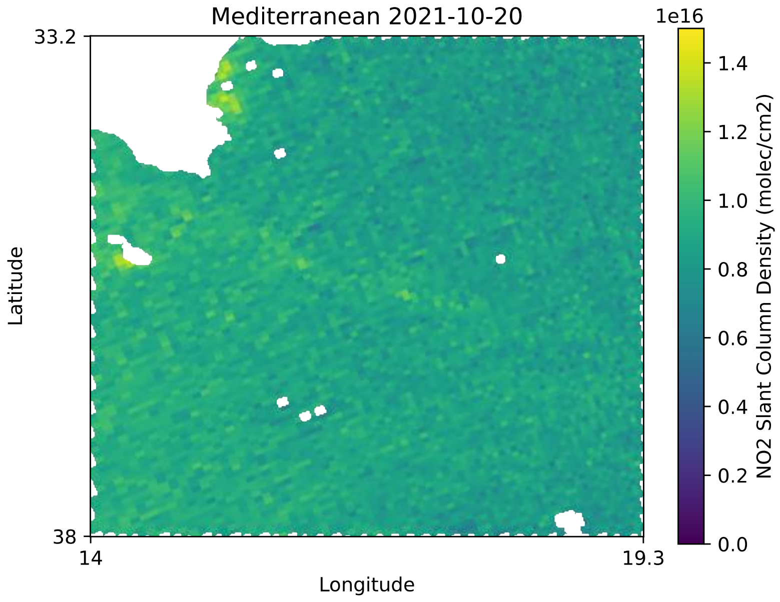
- 80 x 80 km sample patches
- 50 per TROPOMI orbit

NO_2	SO_2	HCHO	Supp.	Ship	Σ proxy
1	1	1	…	1	0,7e8

Building Tabular Database

- 740 orbits
- 50*740 = 37,000 samples
- Filter measurements on:
  - cloud cover < 0.5
  - wind speed > 10 m/s,
  - QA value < 0.5
  - Land area

  Over 50% NaN values -> skip sample
- After filtering: 25,232 samples



Count	**0 (No Ship)**	**1 (Ship)**
Frequency	11,374	13,858
Percentage	45.07%	54.93%

Features, target & other variables

Features

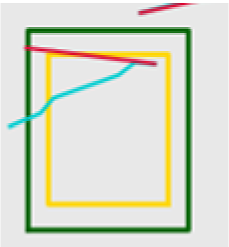
- NO_2 SO_2 HCHO
  - Mean, Min, Max, Std, Median

Support features

- Angles & Wind
  - Mean

Target

- Is a ship plume present (1) or not (0) in an 80km by 80km area
  - WSST estimates ship plume location

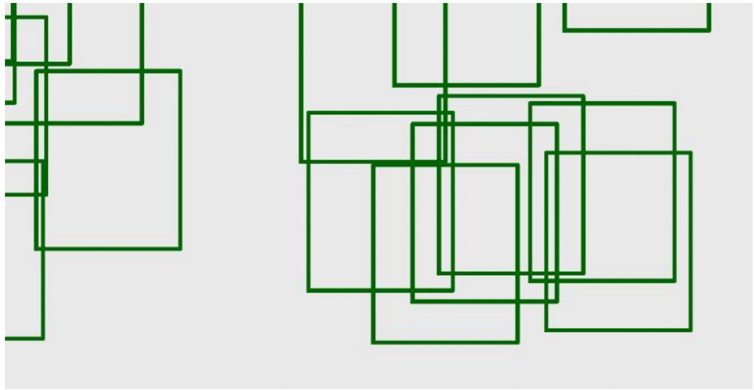


Other variables

- NO_x proxy: sum
- Orbit date

Machine learning

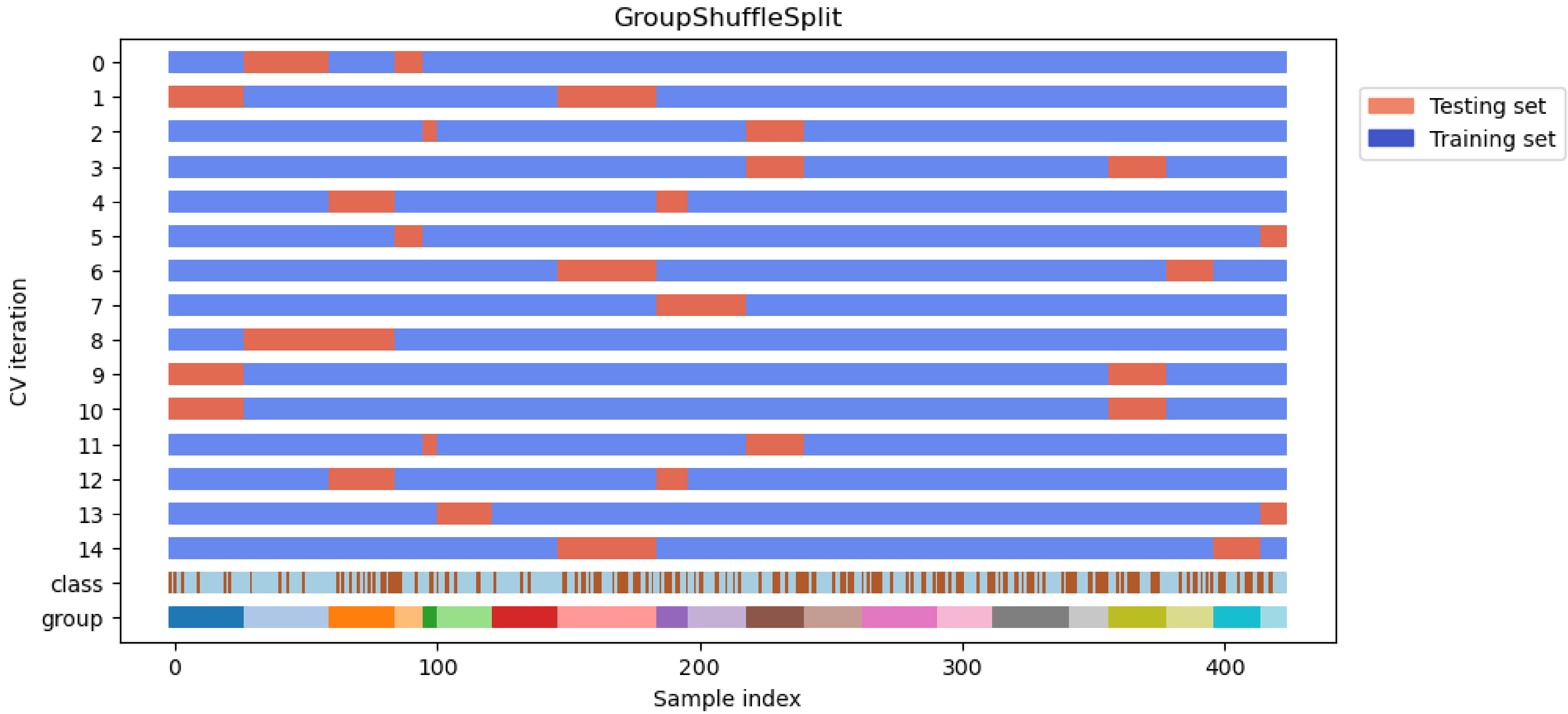
- xgboost-python
- Hyperparameter optimization (HPO)
  - Nested cross-validation
  - 15 outer folds: GroupShuffleSplit
  - 5 inner folds: StratifiedGroupKfold
  - 30 iterations Bayesian search
- Split on orbit date
  - Overlapping sample patches



Hyperparameter	Type	Range	Prior
`n_estimators`	Integer	[10, 500]	-
`gamma`	Real	[1e-8, 0.5]	Log-uniform
`max_depth`	Integer	[2, 10]	-
`min_child_weight`	Integer	[1, 12]	-
`subsample`	Real	[0.6, 1.0]	Uniform
`learning_rate`	Real	[1e-3, 1.0]	Log-uniform
`reg_alpha`	Real	[1e-8, 1.0]	Log-uniform

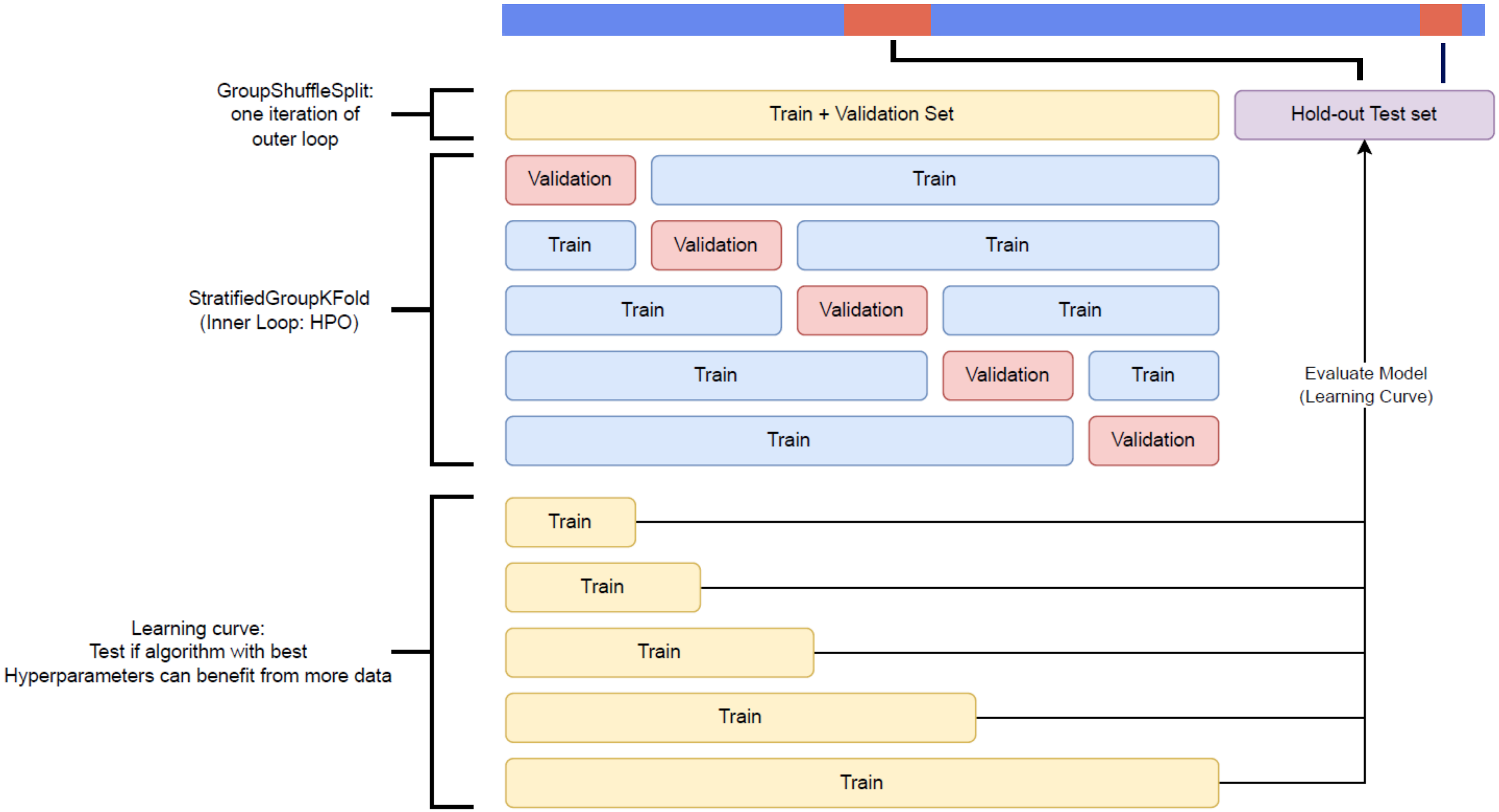
Nested cross-validation

Outer folds



Nested cross-validation

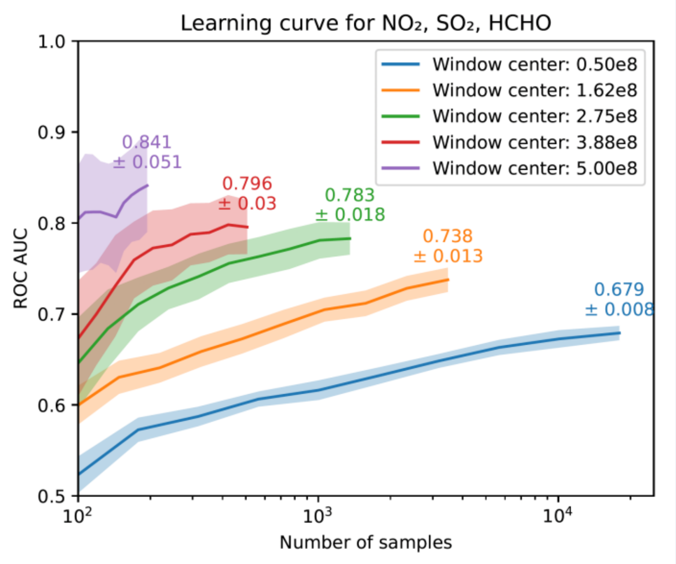
Inner folds



Experiments

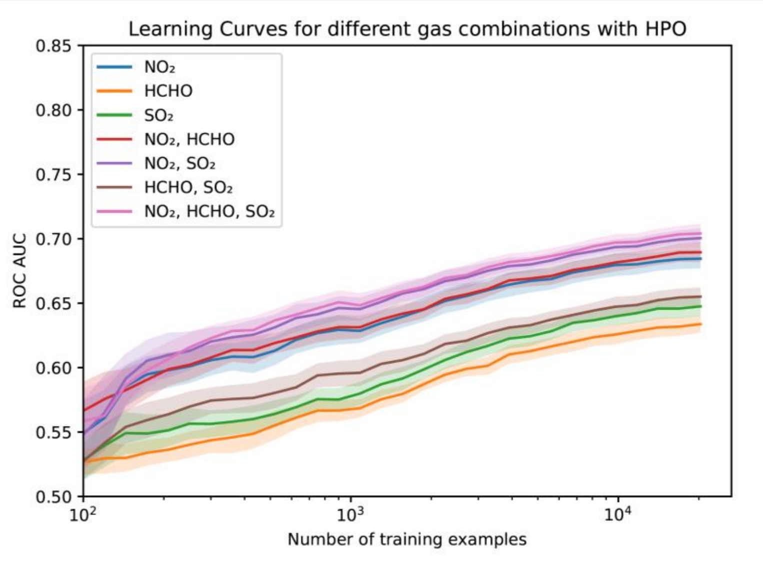
Experiment 1

different proxy bins



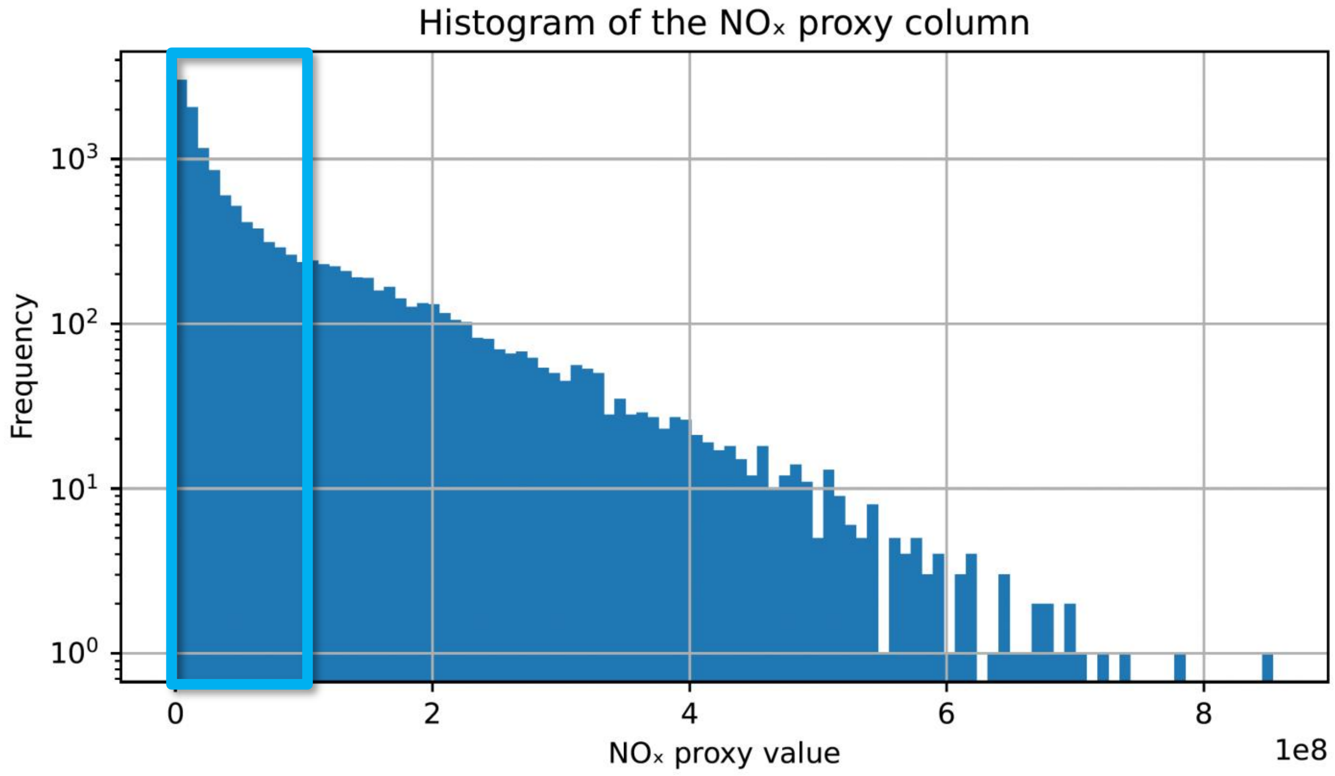
Experiment 2

all combinations of NO_2, SO_2, HCHO



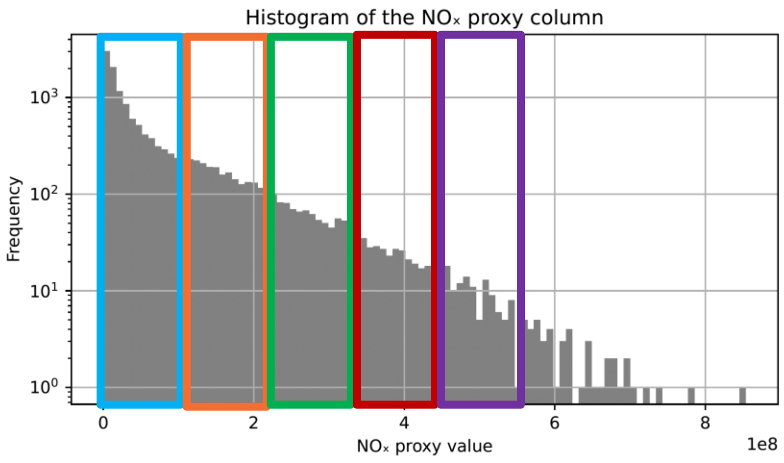
Experiment 1: Proxy bins

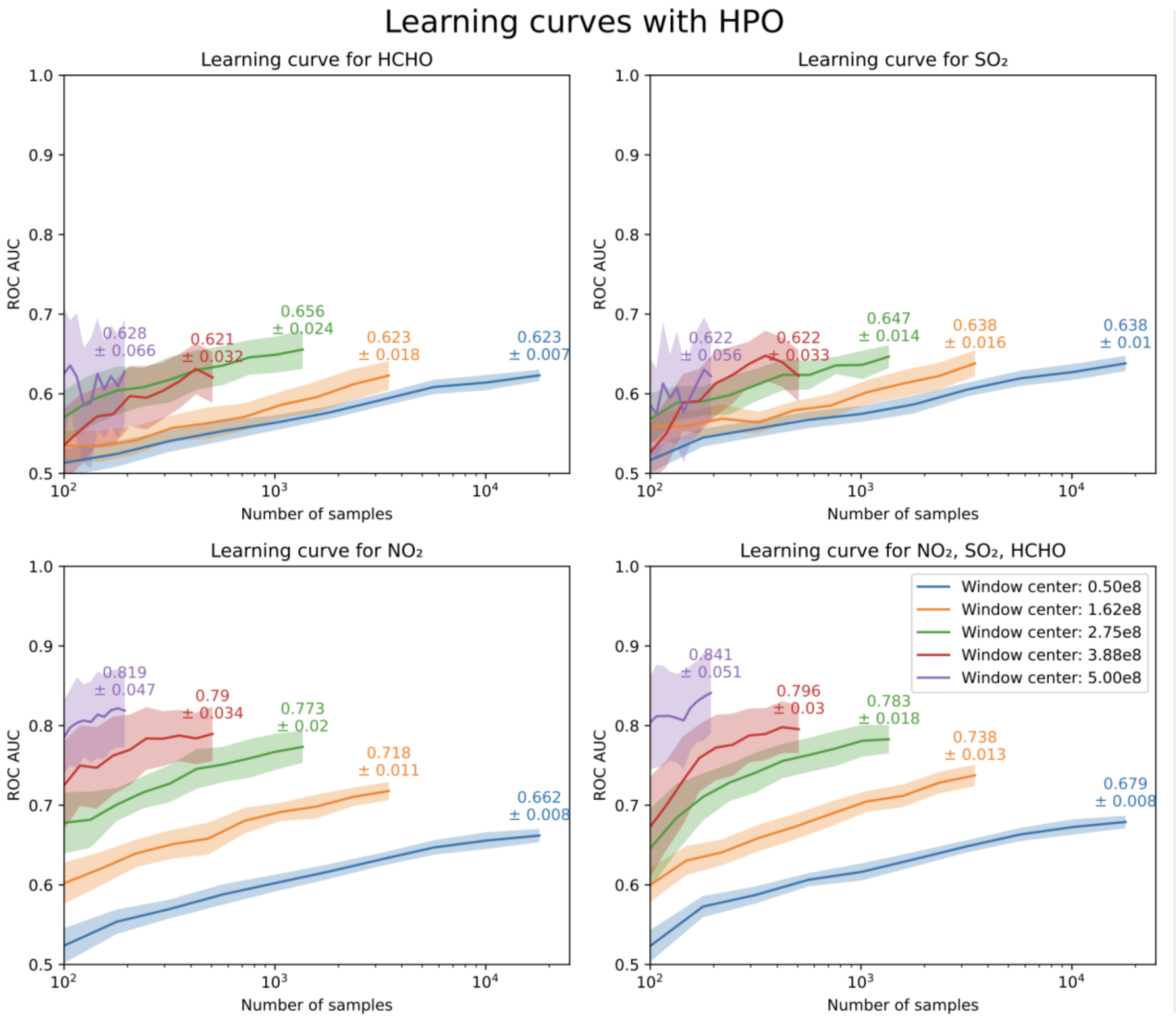
- Use lower and upper threshold to create subset/bin
- Test detection performance for different expected emissions
- 5 bins
  - Window size: 1e8
  - Centers: 0.5e8 - 5e8



Experiment 1: Proxy bins

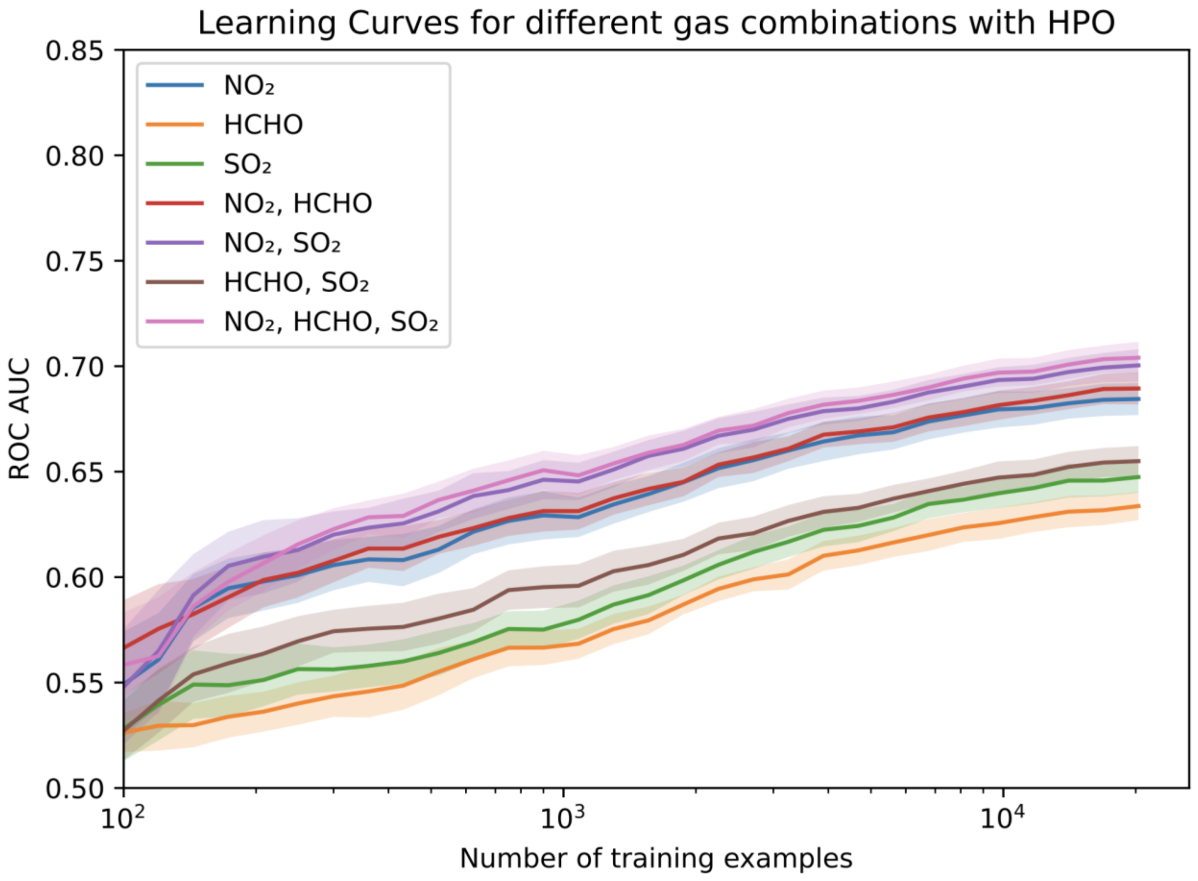
- Learning curve
  - Restricting training set size
- Performance: ROC AUC±$CI_{95\%}$
- Gases tested
  - NO_2, SO_2, HCHO, NO_2+SO_2+HCHO





Experiment 2: Learning curves w/o bins

- More data available due to no proxy binning
- Learning curve
  - Restricting training set size
- Performance: ROC AUC±$CI_{95\%}$
- Gases tested
  - $P(\{NO_2, SO_2, HCHO\})$
- Table: HPO vs no HPO



Selected final scores from learning curve

HPO Combination	No	Yes
NO_2	0.661 ± 0.009	0.684 ± 0.011
HCHO	0.613 ± 0.011	0.634 ± 0.015
SO_2	0.626 ± 0.013	0.647 ± 0.015
NO_2, HCHO, SO_2	0.677 ± 0.011	**0.704 ± 0.013**

Discussion

- Moderate improvement by adding SO_2 and HCHO to NO_2 data.
  - Additional information
  - Most emitted NO_x is NO (Riess, 2024)
- On SO_2 and HCHO alone: worse than NO_2

Limitations:

- HPO
  - Continuous search space
- Wind data
  - Potential data leakage
  - Simplifications in wind shift
- Support features

Selected final scores from learning curve

HPO Combination	No	Yes
NO_2	0.661 ± 0.009	0.684 ± 0.011
HCHO	0.613 ± 0.011	0.634 ± 0.015
SO_2	0.626 ± 0.013	0.647 ± 0.015
NO_2, HCHO, SO_2	0.677 ± 0.011	$\mathbf{0.704 \pm 0.013}$

Future works

This study

- More training data
  - Get more AIS data
  - Expand research area
  - Use orbit files (partially) overlapping research area
- Wind data
  - Historic wind data
  - Higher-resolution wind data

Other directions

- Deep learning
  - Retain spatial information in TROPOMI measurements
  - Improve plume segmentation
- Investigate SO_2 feature importance using IMO2020

Conclusion

Research Questions

1. How does the inclusion of **SO_2** and **HCHO** data, alongside **NO_2** data, improve the detection accuracy of ship plumes using data from a single TROPOMI orbit?

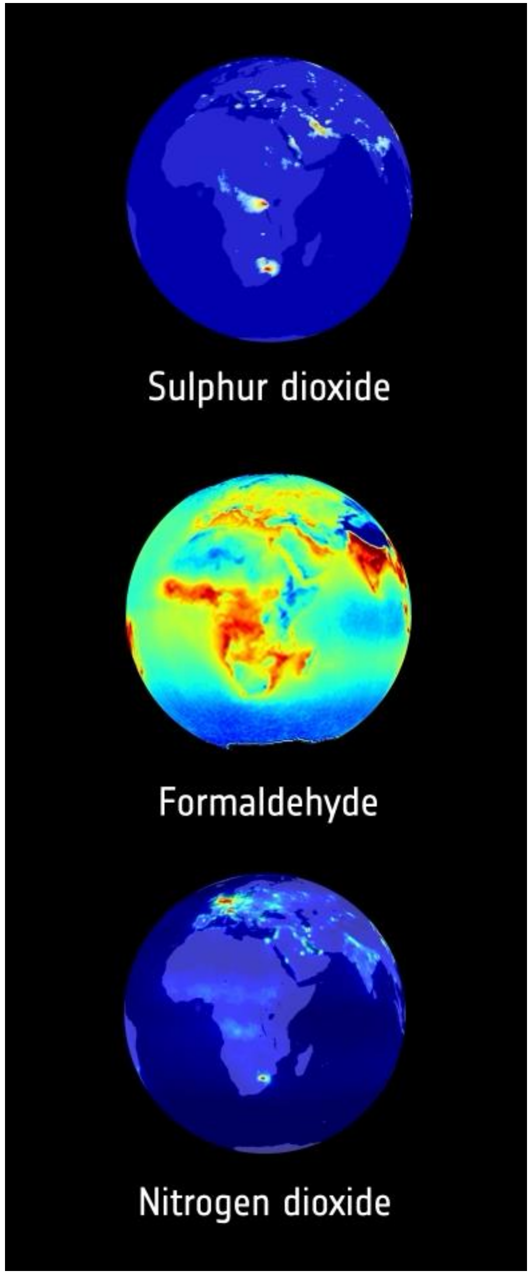
- Inclusion of SO_2 and HCHO data **improves** ship plume detection

2. Can ship plumes be detected based on **SO_2** or **HCHO** data from a single TROPOMI orbit?

- Ships are **challenging** to detect based solely on SO_2 or HCHO

Impact of findings

- **Multi-gas** models could **improve** compliance monitoring methods



(image from ESA)

Thank you

Universiteit Leiden
The Netherlands

Special Thanks To:

- Dr. Cor Veenman
- Jasper van Vliet, ILT
- Prof. Fons J. Verbeek

Data providers

- European Space Agency (ESA)
- Royal Netherlands Meteorological Institute (KNMI)
- ILT for AIS & ShipDB data
  - Data under NDA

References

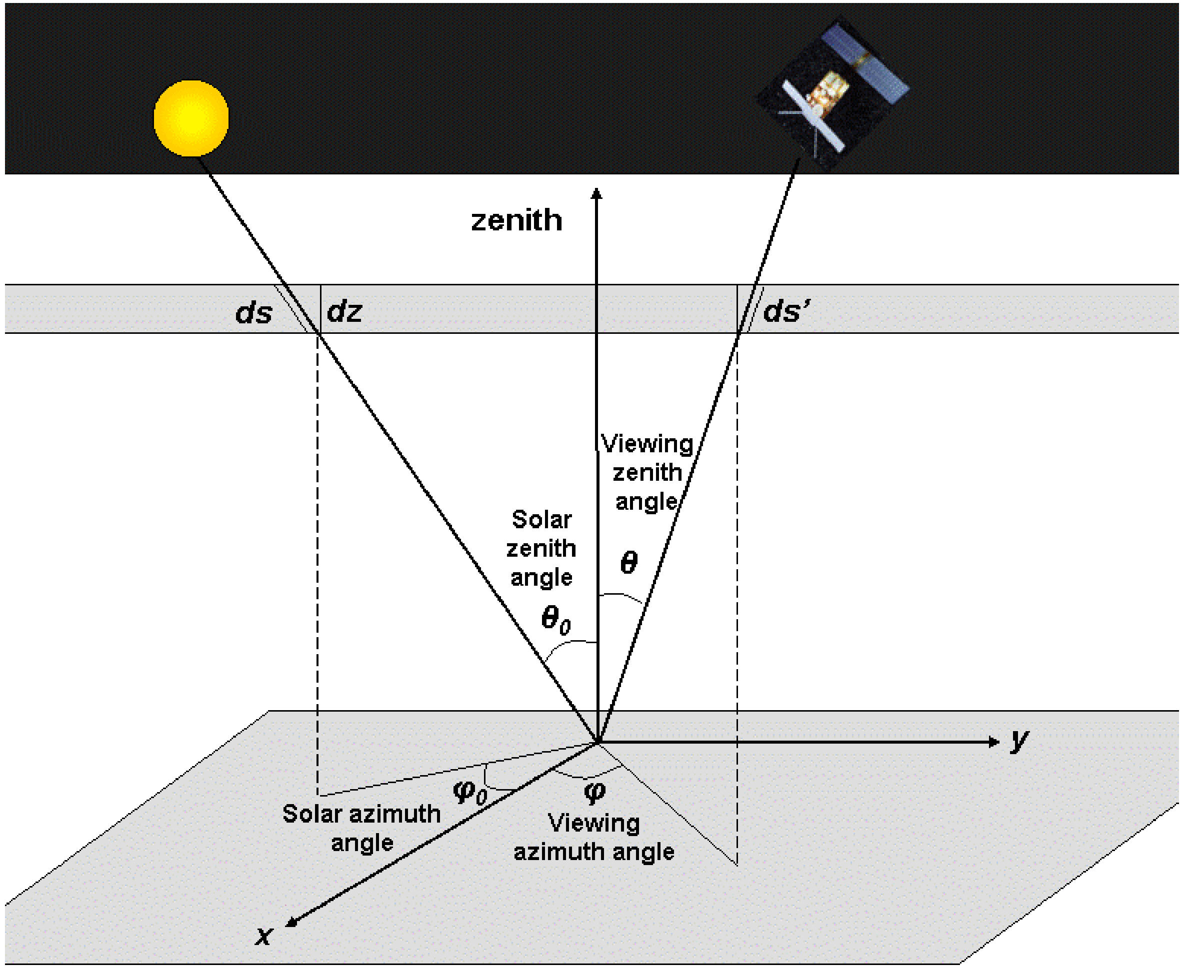
- (Georgoulias, 2020)
  - Georgoulias, A. K., Boersma, K. F., Van Vliet, J., Zhang, X., Zanis, P., & de Laat, J. (2020). Detection of NO2 pollution plumes from individual ships with the TROPOMI/S5P satellite sensor. Environmental Research Letters, 15(12), 124037.
- (Kurchaba, 2024)
  - Kurchaba, S., Sokolovsky, A., van Vliet, J., Verbeek, F. J., & Veenman, C. J. (2024). Sensitivity analysis for the detection of no2 plumes from seagoing ships using tropomi data. Remote Sensing of Environment, 304, 114041.
- (NO2 ATBD, 2023)
  - Geffen, J., Eskes, H., Boersma, K., Maasakkers, J., & Veefkind, J. (2019). TROPOMI ATBD of the total and tropospheric NO2 data products [White paper]. KNMI.
- (SO2 ATBD, 2022)
  - Nicolas Theys, Isabelle De Smedt, Christophe Lerot, Huan Yu, Michel Van Roozendael, & Pascal Hedelt (2022). S5P/TROPOMI SO2 ATBD [White paper]. BIRA-IASB & DLR.
- (Landgraf, 2016)
  - Landgraf, J., Scheepmaker, R., Borsdorff, T., Hu, H., Houweling, S., Butz, A., ... & Hasekamp, O. (2016). Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements. Atmospheric Measurement Techniques, 9(10), 4955-4975.
- (Riess, 2024)
  - Riess, Christoph & Boersma, Klaas & Prummel, Aude & Stratum, Bart & Laat, Jos & Vliet, Jasper. (2024). Estimating NOx Emission of Individual Ships from Tropomi NO2 Plumes. 10.2139/ssrn.4858709.

Backup Slides

Universiteit Leiden
The Netherlands

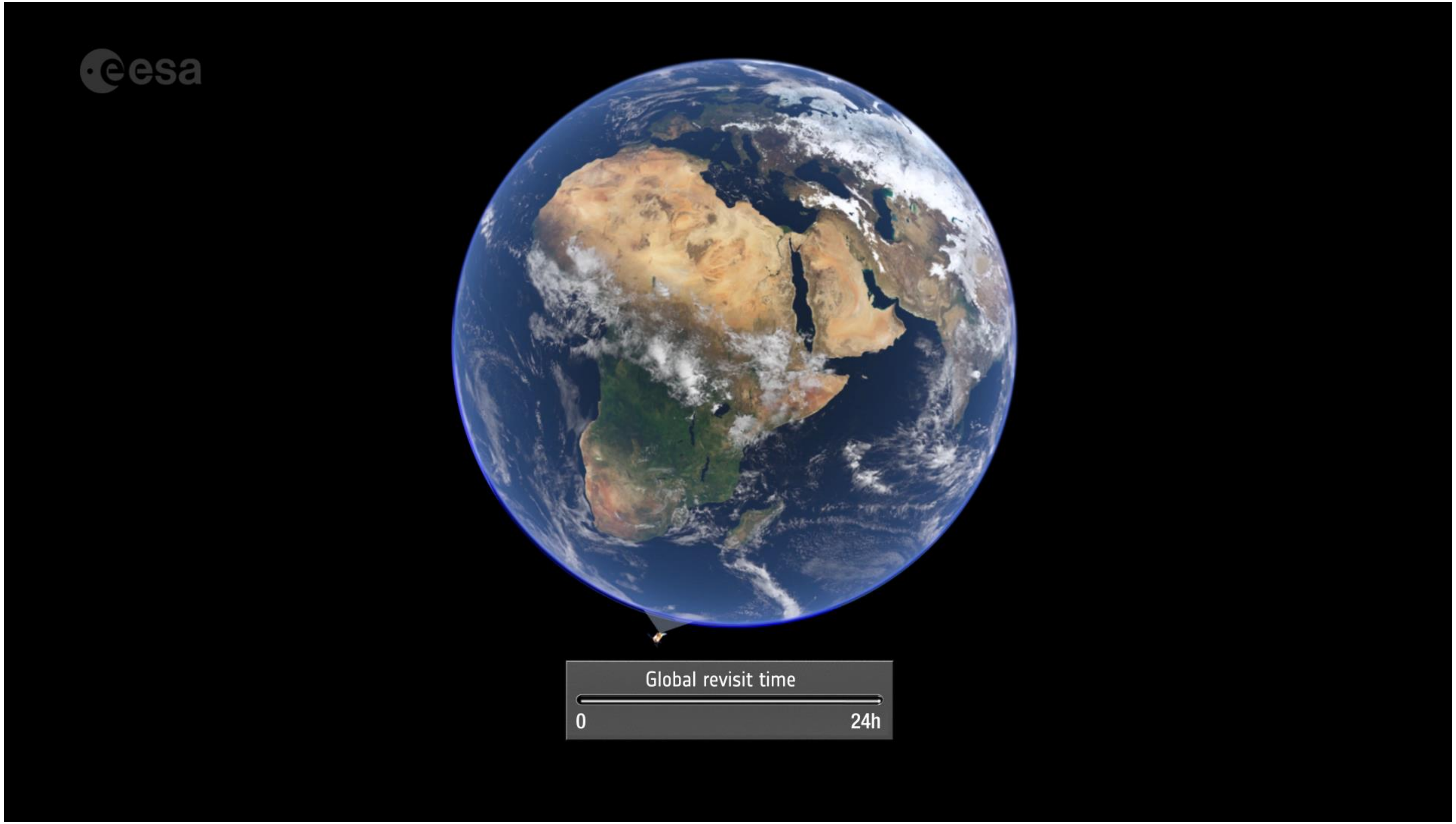
Support features

- Mean over all values in sample patch
- Angles
  - Solar zenith
  - Solar azimuth
  - Sensor zenith
  - Sensor azimuth
- Wind (ECMWF 10m)
  - meridional (vertical)
  - zonal (horizontal)



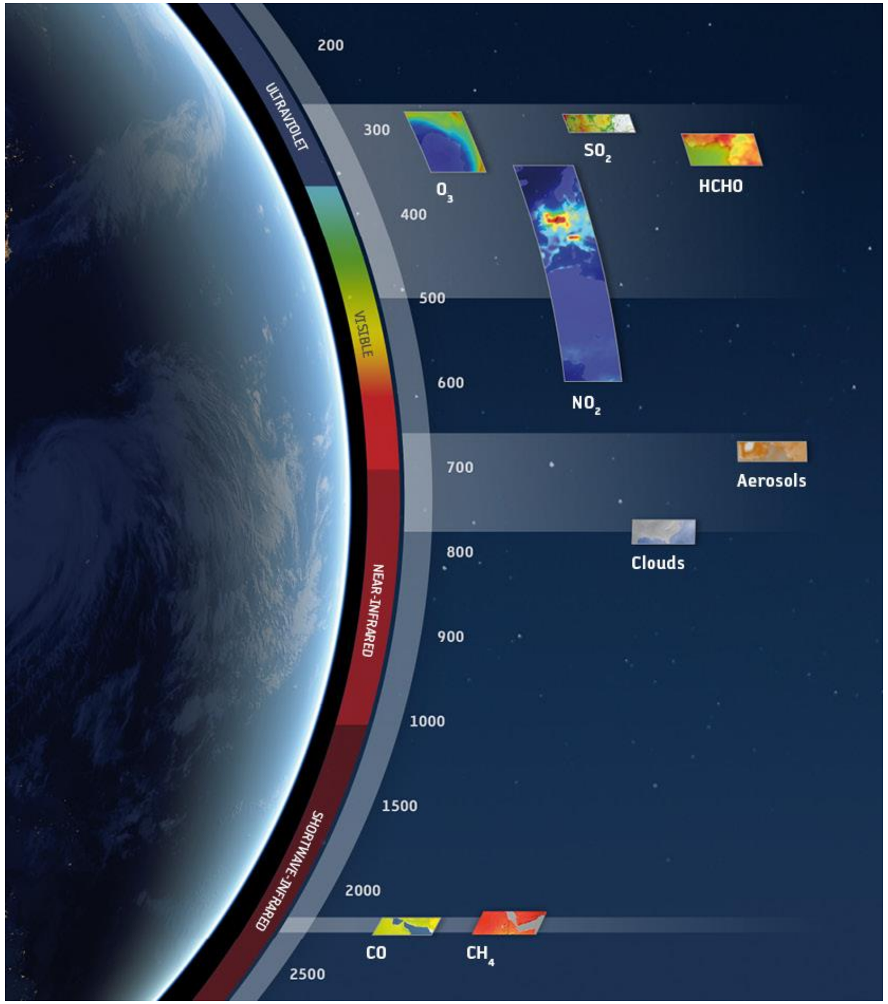
(image from SO_2 ATBD, 2022)

Sentinel-5P Flight: Daily global coverage



How the gases are derived

- From (ir)radiance measured by TROPOMI
- TMP5 and DOAS algorithms for retrieval
- We use: NO2, SO2, HCHO



Exp 2: outtakes

HPO Combination	No	Yes
NO_2	0.661 ± 0.009	0.684 ± 0.011
HCHO	0.613 ± 0.011	0.634 ± 0.015
SO_2	0.626 ± 0.013	0.647 ± 0.015
HCHO, SO_2	0.630 ± 0.012	0.655 ± 0.017
NO_2, HCHO	0.664 ± 0.011	0.689 ± 0.014
NO_2, SO_2	0.674 ± 0.011	0.700 ± 0.012
NO_2, HCHO, SO_2	0.677 ± 0.011	**0.704 ± 0.013**

Table 5.1: Overview of the final values of the learning curve. An ablation study of HPO. Highest performance is highlighted. Values shown as $ROCAUC \pm CI_{95\%}$.

Chapter 5. Experiments & Results

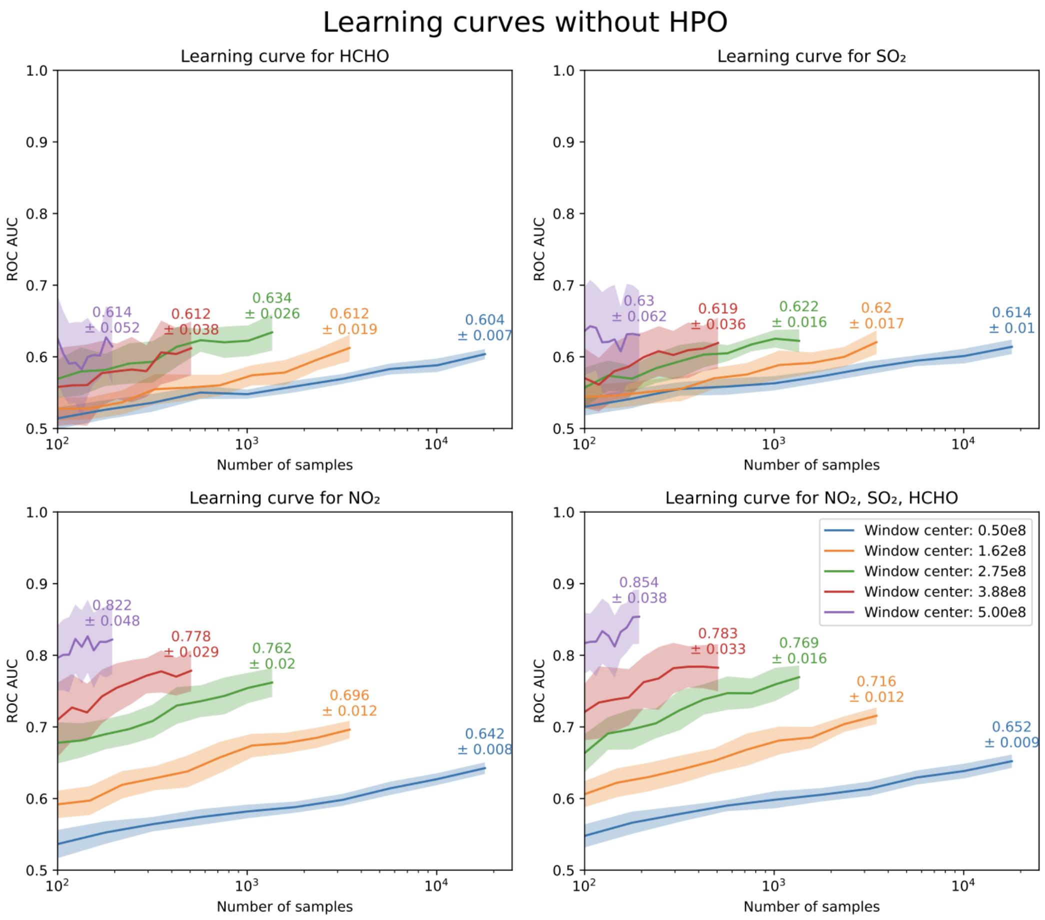
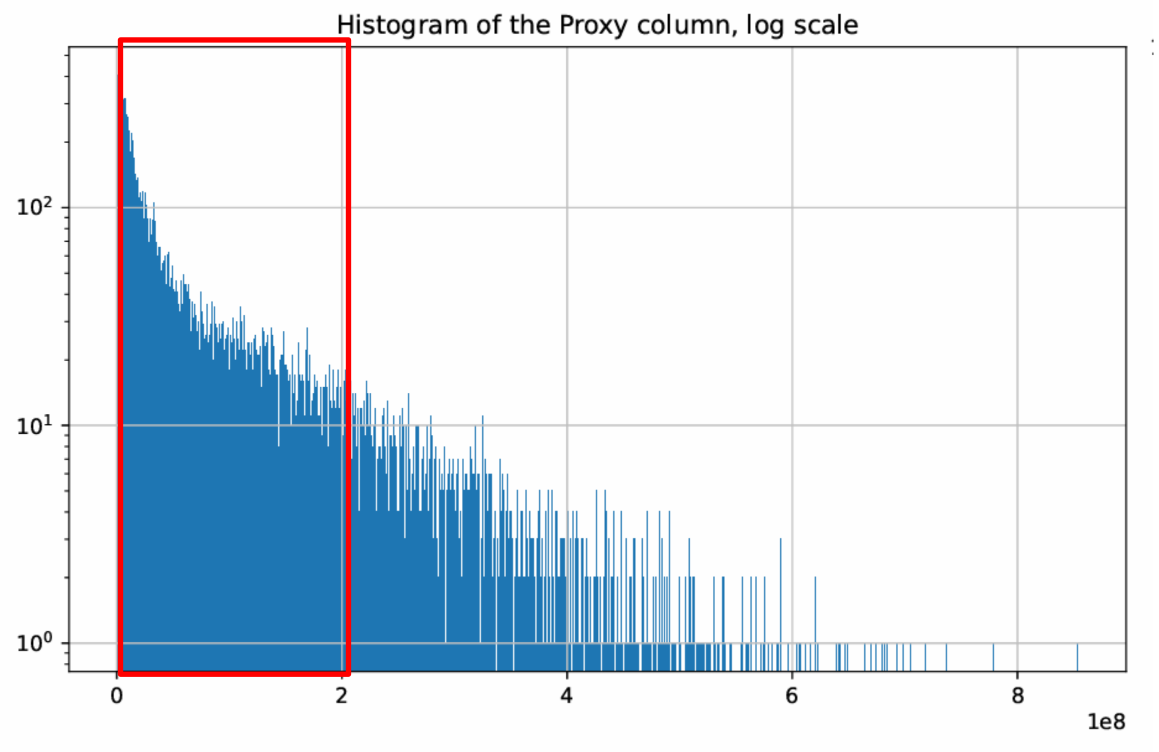


Figure 5.4: Learning curves of ship detection with XGBoost with default HPs. Different proxy windows are compared at the availability of NO_2 or NO_2, SO_2, and HCHO data. A 95% confidence interval is displayed around the lines. The window size is 1e8. The final performance scores are annotated as $ROCAUC \pm CI_{95\%}$.

Experiment 3: Heatmaps

- XGBoost: Default Hyperparameters
- Proxy binning
  - Y-axis
  - Data in lower bins limited by amount of data in higher bins
  - Window size of 2e8
- Learning Curve
  - X-axis
- All gas combinations



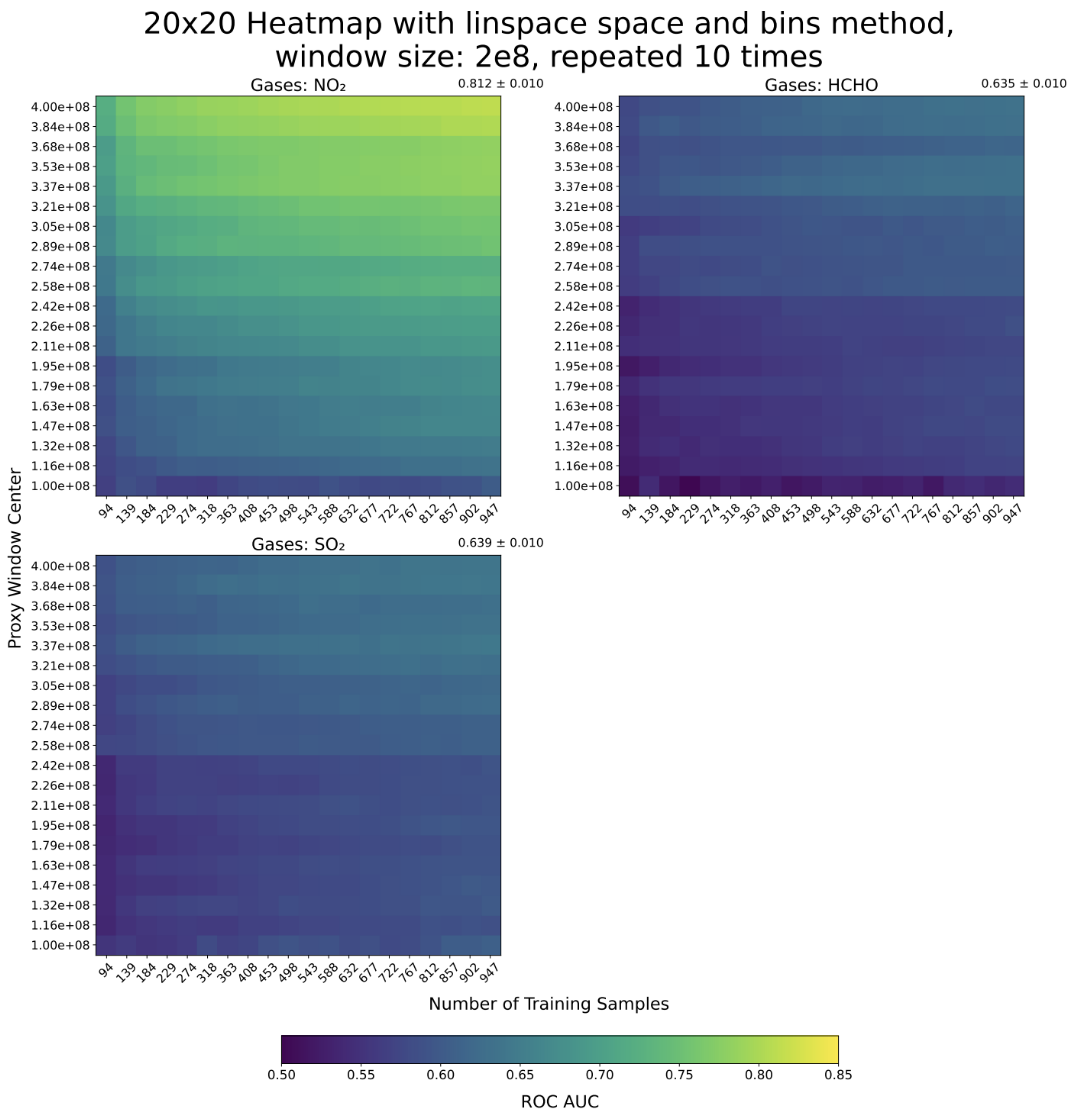


Figure 5.1: Heatmaps for NO_2, SO_2 and HCHO. The heatmaps plot NO_x proxy versus data sets size, where the cell with displayed NOx proxy (p) contains all samples with proxy $p - 1e8$ to $p + 1e8$. The dataset size is increased linearly. The final performance scores for the highest proxies are annotated as $ROCAUC \pm CI_{95\%}$.

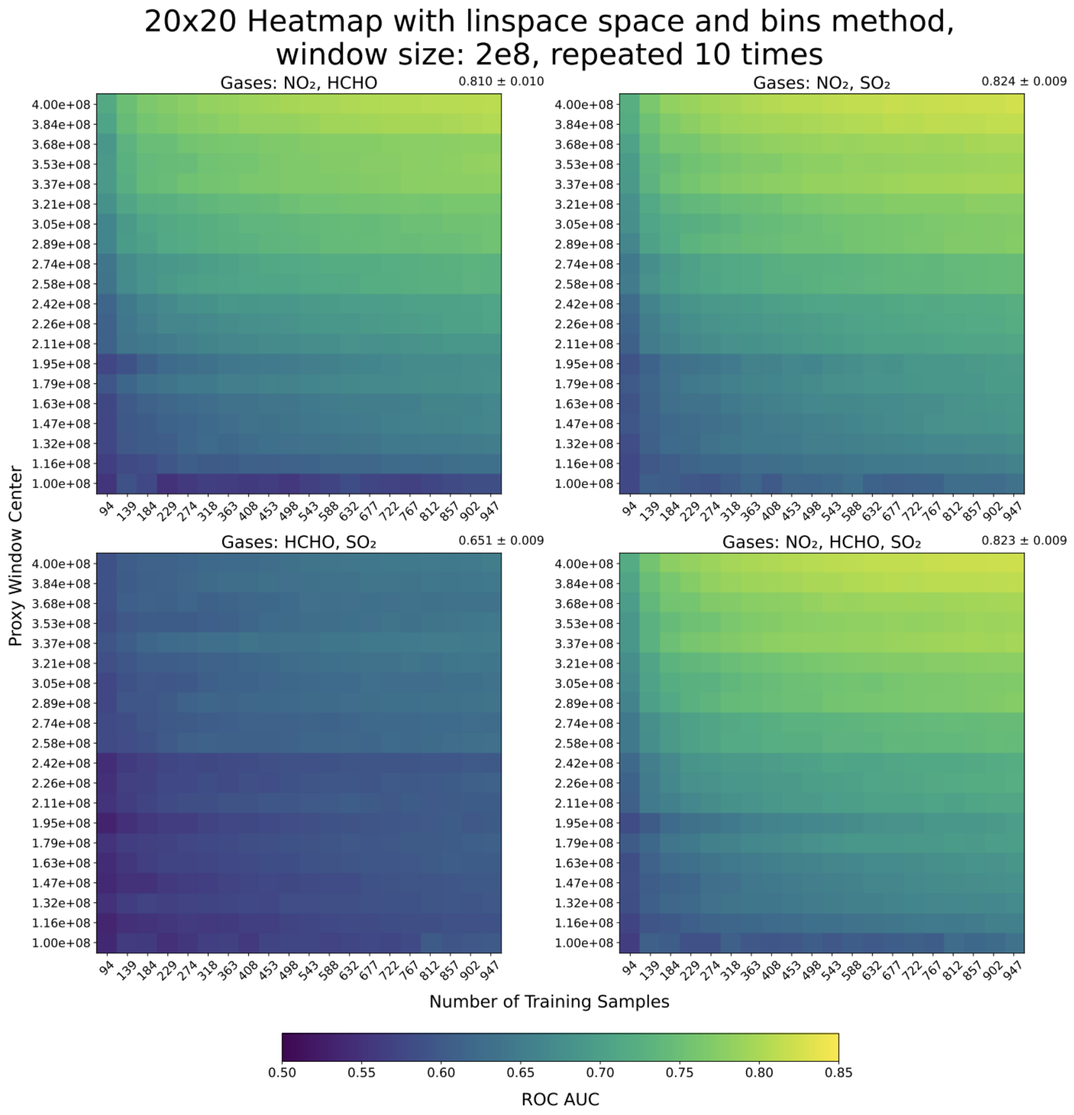


Figure 5.2: Heatmaps for {NO_2, HCHO}, {NO_2, SO_2 }, {HCHO, SO_2 }, and {NO_2, SO_2, HCHO}. The heatmaps plot NO_x proxy versus data sets size, where the cell with displayed NOx proxy (p) contains all samples with proxy $p - 1e8$ to $p + 1e8$. The dataset size is increased linearly. The final performance scores for the highest proxies are annotated as $ROCAUC \pm CI_{95\%}$.

3: Heatmaps

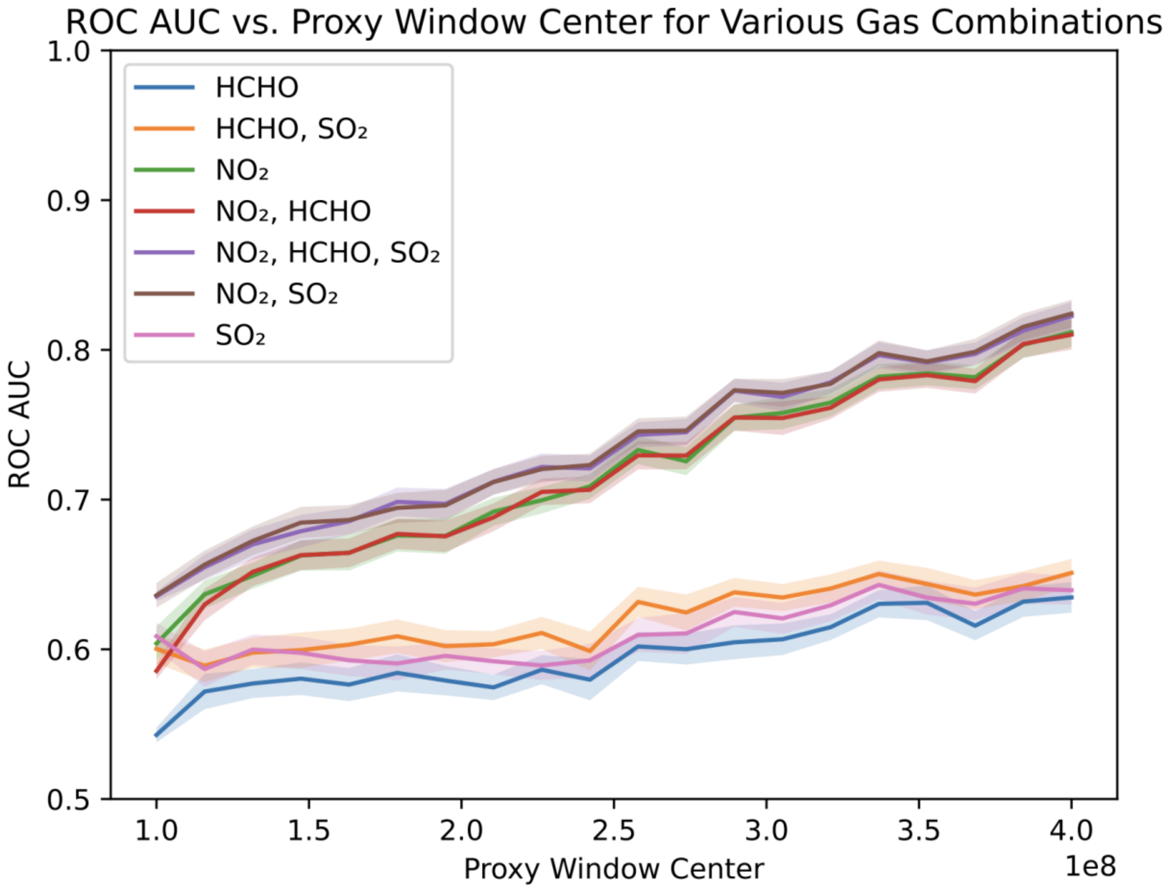
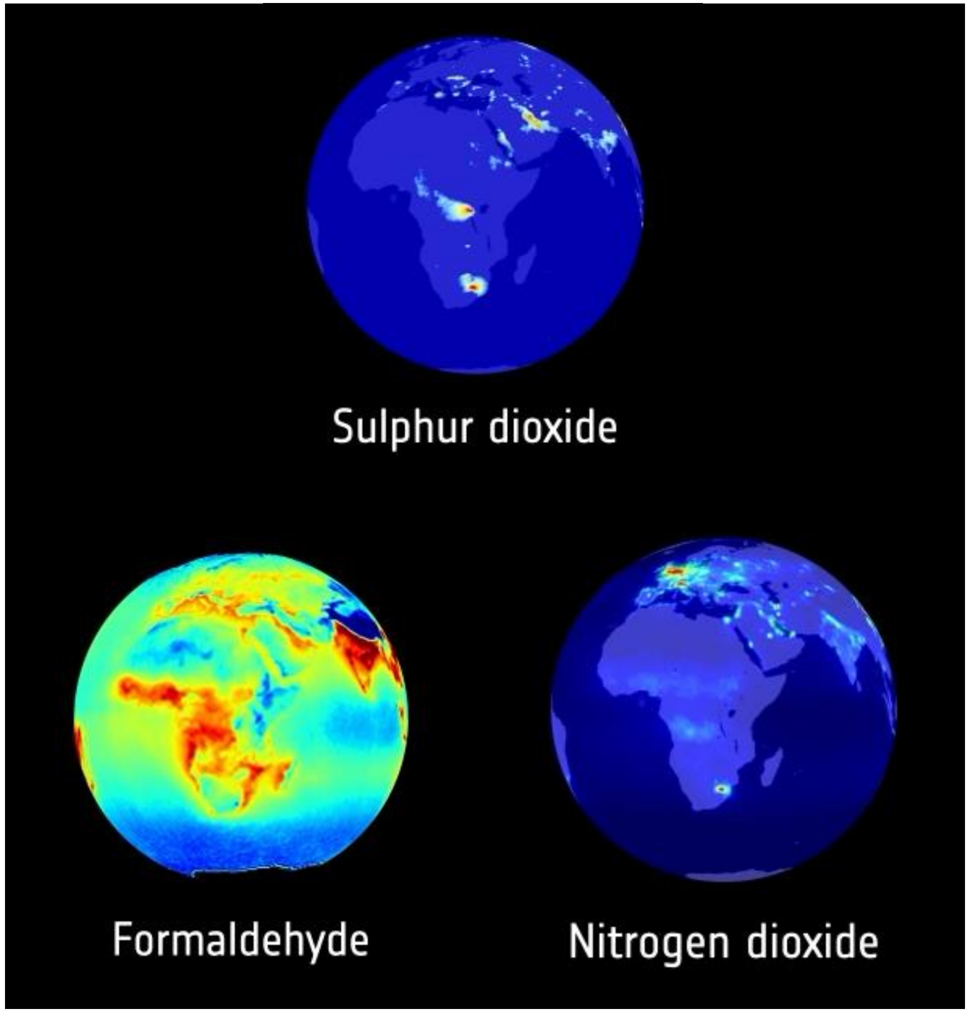


Figure 5.3: This figure shows the relationship between NO_x proxy and the detectability of ships. A window size of 2e8 is used.

Why not CO_2?

- **Higher background levels**
- **No previous work on individual emitters**



(image from ESA)