NO2 ship-plume segmentation with supervised learning on TROPOMI/S5P satellite data

Solomiia Kurchaba, Jasper van Vliet, Fons J. Verbeek, Jacqueline J. Meulman, Cor J. Veenman

27-05-2022

Universiteit Leiden
The Netherlands

Human Environment and Transport Inspectorate
Ministry of Infrastructure and Water Management

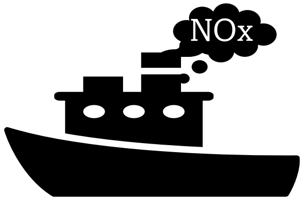
This presentation participates in OSPP

EGU

Outstanding Student Poster & PICO Contest

#EGU22

Motivation

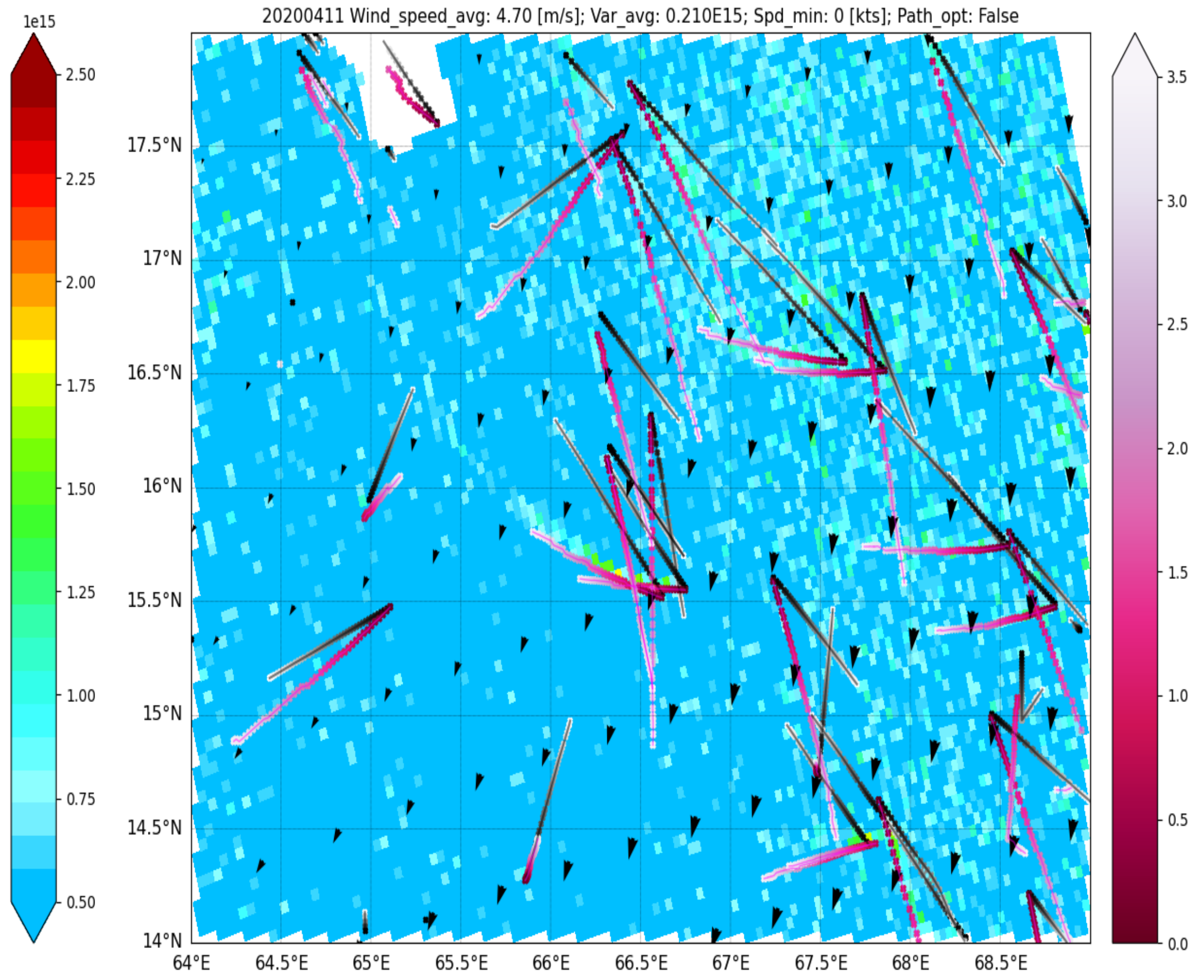


- Shipping industry - one of the biggest (15-35% worldwide) NOx/SOx emitters
- International Maritime Organization (IMO) tightens the restrictions regarding NOx/SOx presence in ships' fuel
- Lack of possibilities to perform a fuel quality monitoring above the open sea

Idea: *Application of TROPOMI measurements for automatic estimation of the amount of NO2 produced by individual ships*

Data example:

- **Background – S5P-TROPOMI** NO2 tropospheric column;
- **Black lines** – original ships' traces from **AIS** data;
- **Magenta lines** – ships' traces shifted according to the wind speed/direction;
- **Colour intensity** decreases with the time passed before the TROPOMI overpass (scale in hours);
- **Black arrows** – wind direction;



Data pre-processing

a) **Ship Plume Image** – Topomi NO2 *VCD_trop* around the ship

b) **Image Enhancement** – local Moran's *I*

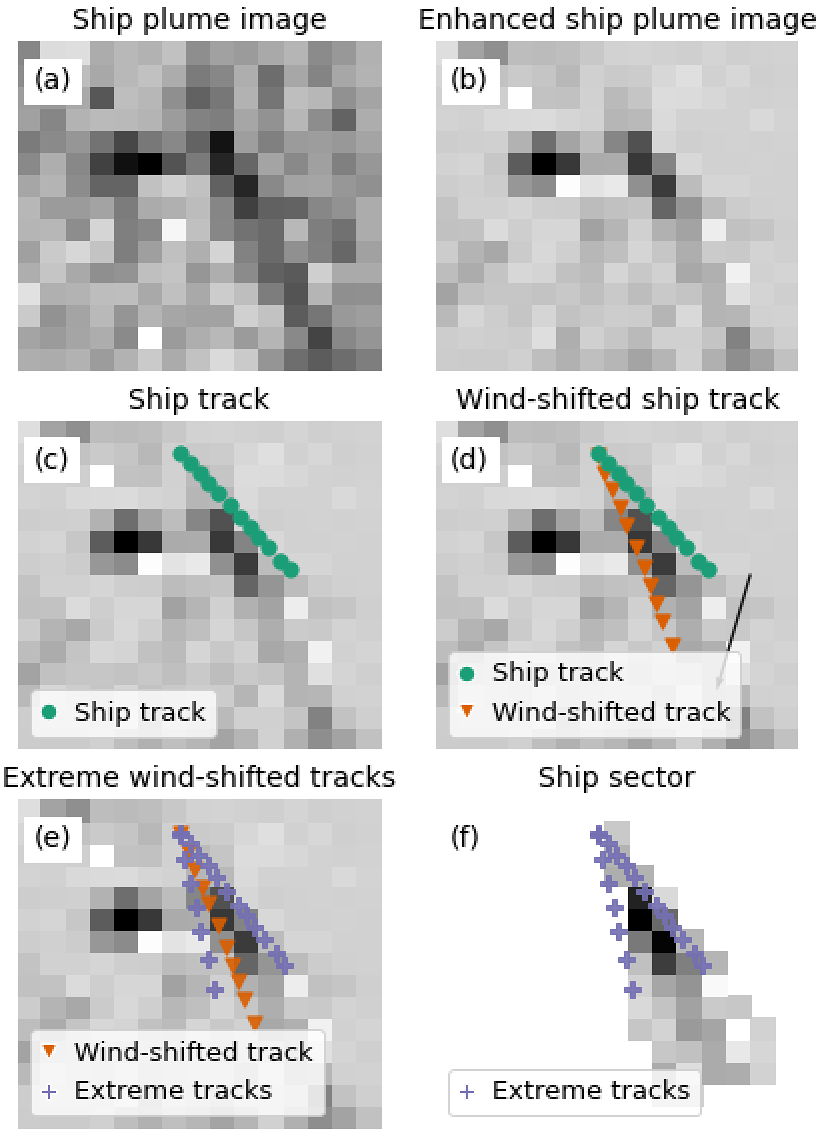
$$I_i = \frac{x_i - \bar{x}}{s^2} \sum_{j=1, j \neq i}^{n} w_{ij} (x_j - \bar{x})$$

c) **Ship Track** based on AIS data

d) **Shifted Ship Track** in accordance to wind speed/direction

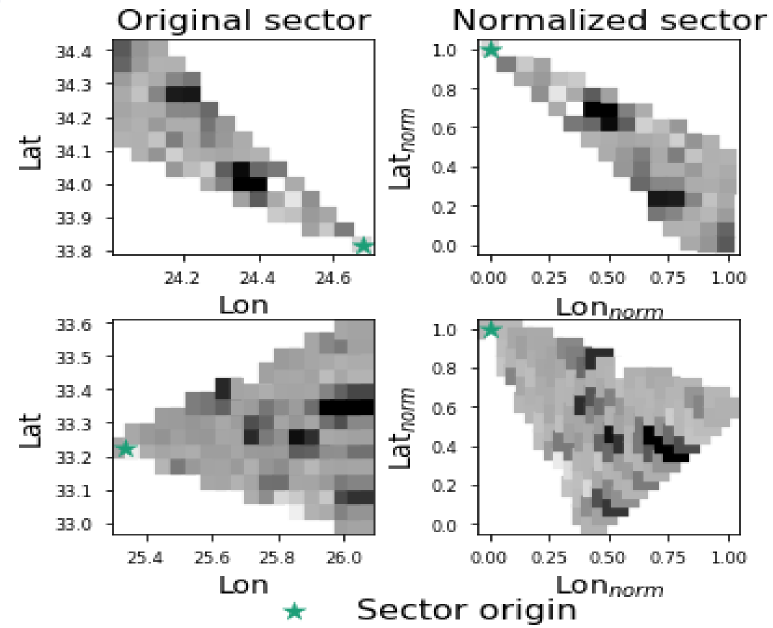
e) **Extreme Ship Tracks** – based on assumed wind uncertainty

f) **Ship Sector** – the ships' Region of Interest. We assume that the plume of analyzed ship is located within **Ship Sector.**

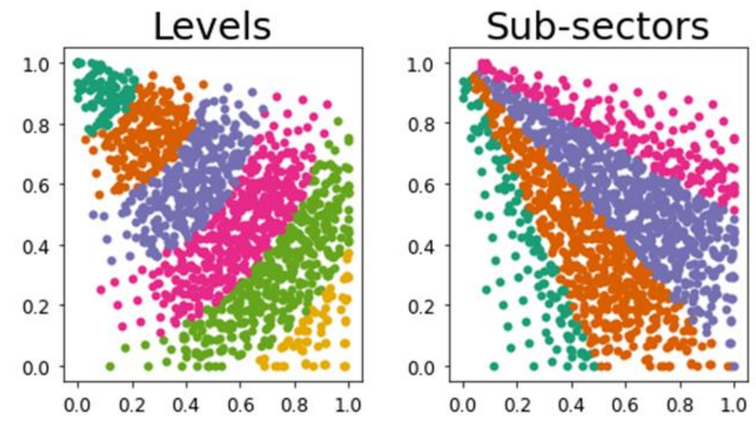


Sector standarization

Sector rotation
standardization of the position of the ship's plume with respect to the position of the ship



Sub- regions
the division allows to characterize each pixel within the normalized sector with two values: ***Level*** and ***Sub-sector***



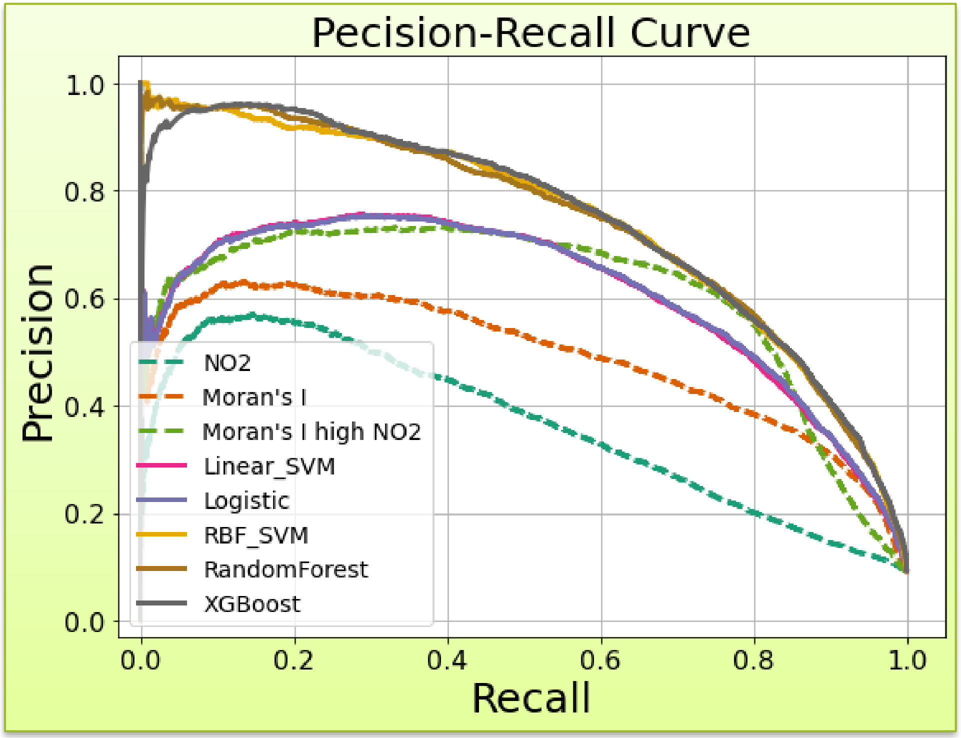
Experiment

- **Studied Area:** Mediterranean Sea [lon: 19.5 – 29.5; lat: 31.3 – 34.2]
- **Studied Period:** 04.2019 – 12.2019
- **Dataset:** 754 images manually labeled pixel-wise
- **Feature set:** 17 features including:
  - Pixel position in terms of Levels and Sub-sectors
  - Pixel intensity
  - Wind and ship related features
- **Models:** 5 linear and non-linear supervised classifiers
- **Benchmarks:** pixel intensity based thresholding methods (NO_2 based or Moran's I based)

Data Example

Results

	Model	Average Precision
Benchmarks	NO2	0.375±0.062
	Moran's I	0.493±0.063
	Moran's I on high NO2	0.607±0.056
Linear	Linear SVM	0.609±0.063
	Logistic	0.610±0.064
Non-linear	RBF SVM	0.742±0.031
	Random Forest	0.743±0.030
	XGBoost	**0.745±0.030**



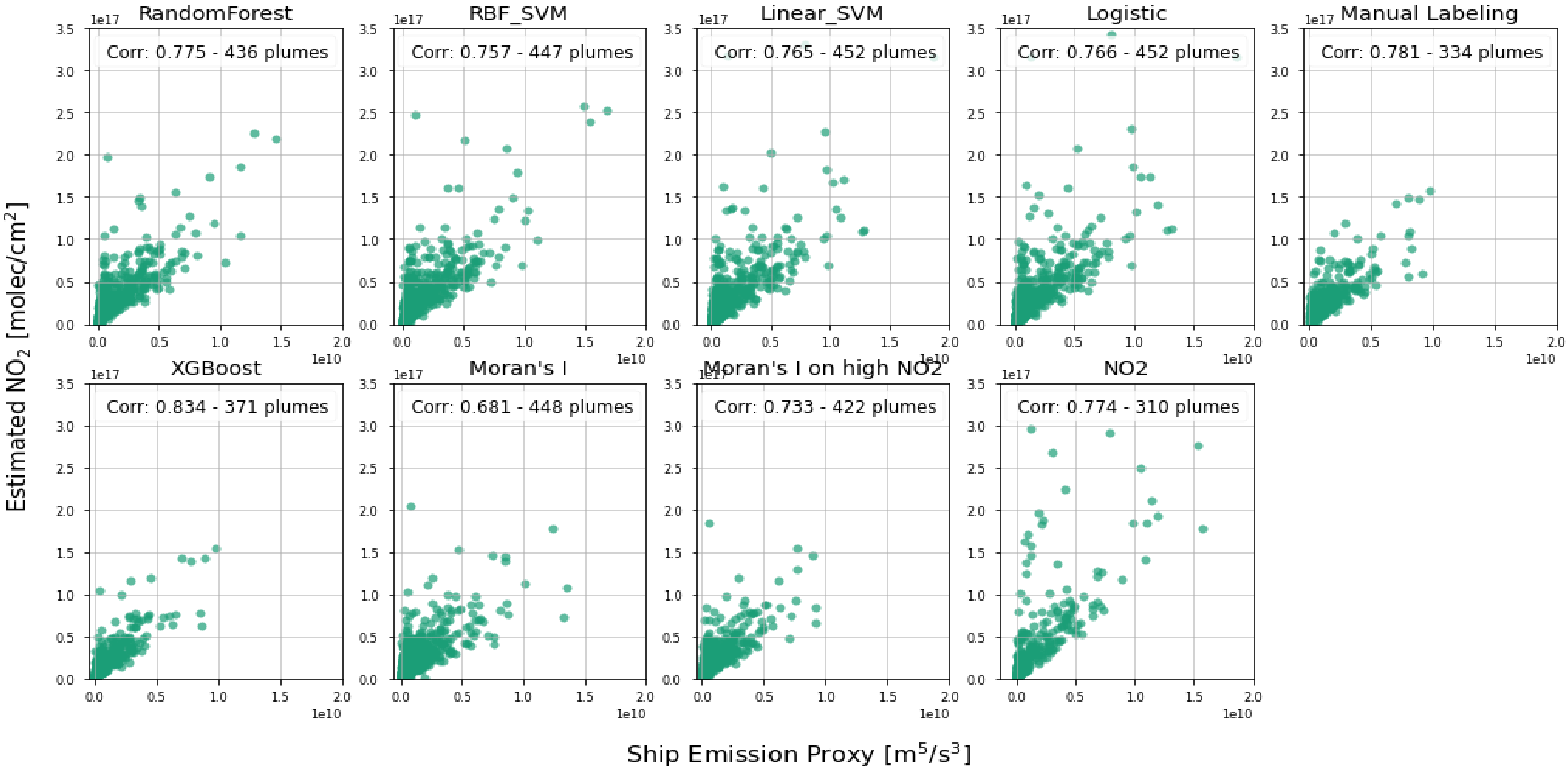
Conclusions

- All models outperform the benchmarks
- Non-linear models allow to achieve very high precision with acceptable level of recall

Validation of NO2 estimation

Ship emission proxy Es *allows to estimate the expected amount of produced NOx emission for a ship of a certain size and speed*

$$E_s = L_s^2 \cdot U_s^3$$



Thank you for attention!

Solomiia Kurchaba, PhD candidate

LIACS, Leiden University

Contact details

- **Email:** s.kurchaba@liacs.leidenuniv.nl
- **LinkedIn:** https://www.linkedin.com/in/solomiia-kurchaba-4341a9118/

#EGU22

Supervisors:

Dr. Cor Veenman
LIACS, Leiden University

Jasper van Vliet
ILT, project Team Lead

Prof. Jacqueline Meulman
MI, Leiden University

Prof. Fons Verbeek
LIACS, Leiden University

**This work is funded by the Netherlands Human Environment and Transport Inspectorate, the Dutch Ministry of Infrastructure and Water Management, and the SCIPPER project which receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nr.814893.*

Read more about the study

Pre-print on arXiv

https://arxiv.org/abs/2203.06993

Full Results Supplement

$$\textbf{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$\textbf{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$\textbf{F1} = \frac{Precision * Recall}{Precision + Recall}$$

Average Precision (AP) – area under Precision-Recall curve

Model	Precision	Recall	F1	AP	
NO2	0.40±0.051	0.48±0.03	0.44±0.04	0.38±0.06	*Benchmarks*
Moran's I	0.44±0.04	0.71±0.03	0.54±0.04	0.49±0.06	
Moran's I on high NO2	0.65±0.03	0.69±0.03	0.67±0.03	0.61±0.06	
Linear SVM	0.44±0.03	0.83±0.02	0.57±0.03	0.61±0.06	*Linear models*
Logistic	0.44±0.03	0.84±0.02	0.58±0.03	0.61±0.06	
RBF SVM	0.47±0.03	0.86±0.02	0.61±0.03	0.74±0.03	*Non-linear models*
Random Forest	0.55±0.02	0.81±0.02	0.65±0.02	0.74±0.03	
XGBoost	0.77±0.03	0.58±0.04	0.66±0.03	**0.75±0.03**	