Machine learning-based NO2 estimation from seagoing ships using TROPOMI/S5P satellite data

Kurchaba, S.

Chapter 4

Automated assignation of a ship Region of Interest for estimation of NO_2 emission from individual ships using satellite data

Extended from: Kurchaba, S., van Vliet, J., Meulman, J.J., Verbeek, F.J., Veenman, C.J., 2021. Improving evaluation of NO_2 emission from ships using spatial association on TROPOMI satellite data, in: 29th International Conference on Advances in Geographic Information Systems, pp. 454–457. doi:10.1145/3474717.3484213.

Abstract As of 2021, more demanding NO_x emission requirements entered into force for newly built ships operating on the North and Baltic Sea. Even though various methods are used to assess the emission from ships in ports and off the coastal areas, monitoring over the open sea has been infeasible until now. In this Chapter, we present an automated method for the evaluation of NO_2 emissions produced by individual seagoing ships. We use the spatial association statistic local Moran's I in order to improve the distinguishability between the plume and the background. Using the Automatic Identification Signal (AIS) data of ship locations as well as incorporated uncertainties in wind speed and wind direction, we present a method for automatic association of the detected plumes with individual ships. We evaluate the quality of ship-plume matching by calculating the Pearson correlation coefficient between the values of a model-based emission proxy and the estimated NO_2 concentrations. For five of the six analyzed areas, our method yields improved results against the baseline approach used in a previous study.

4.1 Introduction

Once the sensitivity limits of the detection system using TROPOMI data are estimated, we focus our attention on using the TROPOMI data for the quantification of NO_2 emission from individual ships. In [41], the authors introduced the first attempt to quantify isolated ship plumes that can be identified by visual inspection of daily data. However, the NO_2 traces from the majority of the ships in the area are not sufficiently stronger than the NO_2 background concentration. As a result, only plumes of larger ships were assessed in that study. In addition, the authors acknowledged that their approach requires multiple manual steps. In order to be able to apply the TROPOMI instrument data for the global and continuous monitoring of the NO_2 emissions from seagoing ships, an automated method for ships' NO_2 estimation is needed.

In this Chapter, we present a heuristic for automated evaluation of NO_2 concentrations resulting from NO_x emissions produced by individual seagoing ships. We start with the enhancement of the contrast between NO_2 plumes and the background. We then introduce a method for automated assignation of the Region of Interest (RoI) to a studied ship. Finally, we apply a thresholding method for the separation of ship plumes from the background and estimation of ships' NO_2. The presented approach allows the quantification of the ships' emission, even if the produced plume cannot be distinguished visually so that the performance of more and smaller ships can be assessed in a single satellite overpass. The obtained results are benchmarked against the method proposed in [41].

With this study, we address the following research questions:

- **RQ4**: How to assign a TROPOMI signal associated with a certain plume to a potential emitting ship?
- **RQ5**: To what extent can the NO_2 plumes be segmented in the TROPOMI data using a simple thresholding method?

The rest of the Chapter is organized as follows: In Section 4.2, we present our methodology. We start with the introduction of the applied image enhancement method in Section 4.2.1. We then present a developed approach to ship-plume assignment in Section 4.2.2. We further explain how we estimate and evaluate the quality of estimation of ship's NO_2 in Section 4.2.3. The results of the study are presented in Section 4.3, and conclusions can be found in Section 4.4.

4.2 Method

In this Section, we describe all steps of the proposed approach for automated evaluation of NO_2 emission from individual ships. We start with the introduction of the technique used for the enhancement of the TROPOMI data. We then present a proposed approach for the delineation of the region of interest of the studied ship. Finally, we explain how we evaluate the obtained results of ship NO_2 estimation.

4.2.1 Image enhancement

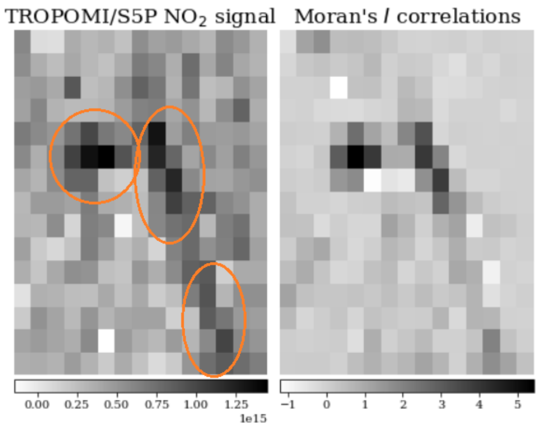

Figure 4.1: Reduction of the background noise as a result of application of the local Moran's I. Orange circles indicate ship plumes that can be distinguished on the plot. NO_2 is in [molec/cm^2].

The first step is an enhancement of the TROPOMI data. To increase the contrast between the ships' plumes and the background, we used spatial association statistic local Moran's I [5] — one of the most often used methods for hot spot detection [78]. The local Moran's I spatial auto-correlation statistic allows the enhancement of the intensity of high-value pixels located in a cluster while suppressing isolated concentration peaks randomly occurring in the background. We characterize a ship plume as a cluster of pixels adjacent to each other with a concentration higher than the background average. This way, calculating the spatial auto-correlation on TROPOMI image, we combine image denoising with the enhancement of the relevant part of the image. Figure 4.1 illustrates an example where the applied image enhancement procedure notably reduced the background noise and increased the contrast between the plume and the background, improving the detectability of the ships' traces.

Formally, the local Moran's I spatial auto-correlation statistic is defined as follows:

$$I_i = \frac{(x_i - \mu)}{\sigma^2} \sum_{j=1, j \neq i}^{N} w_{ij}(x_j - \mu), \tag{4.1}$$

where i is the pixel of an image, x_i is the value of the respective pixel, N is the number of analyzed pixels of a *ship plume image* (in our case 18×18), μ is a mean value of all N pixels, σ^2 their variance, and w_{ij} is the value of an element in a binary spatial contiguity weight matrix W at location j with regards to the analyzed pixel i. The value of an element of the binary spatial contiguity matrix w_{ij} is 1 for pixels that are considered to be the neighbors of the analyzed pixel i, and 0 otherwise. For the study, the queen spatial contiguity [43], which is the 3×3 8-connected neighborhood of the analyzed (central) pixel was applied. The value I_i becomes the value of the corresponding pixel of the resulting enhanced image.

4.2.2 Ship-plume assignment

Determining the spatial correspondence between the TROPOMI signal and the location of the ships is a challenging task. The emitted plume is displaced by the prevailing winds so that observed NO_2 concentrations no longer coincide with the tracks of the ships. At the same time, linear transformation of the ship trajectory that is solely based on available wind data (e.g. [41]) might be inaccurate due to the wind-related uncertainties (see Figure 4.2b). Such a method is further used for the benchmarking of our approach. We refer to it as *ship track shift*. To overcome the problem of inaccurate ship-plume matching, we propose to assign to each ship a Region of Interest (RoI), the so-called *ship sector* (Figure 4.2d). The RoI of a ship defines the region within which the produced NO_2 is expected based on ship speed, wind speed, direction, and uncertainties.

To define a *ship sector*, we start with estimating the trajectory of the ship – a *ship track* – using AIS ship data, from some time before, until the moment of the satellite overpass (c.f. Figure 4.2a). For each of the analyzed regions, we experimentally determined the optimal time prior to the satellite overpass during which the AIS data was used for plume trace localizing. For the Mediterranean Sea area, this time was equal to 1 hour, and for the Arabian Sea to 40 minutes. The difference in the considered time can be explained by the fact that the time needed for $NO_x \rightarrow NO_2$ transformation is mostly determined by the atmospheric conditions [94], which in the area of the Arabian and Mediterranean Sea are different.

We now move to the introduction of the second step of *ship sector* definition. Following [41], we assume that the plume emitted by a ship has moved in accordance with wind direction by a distance $d = v \times |\Delta t|$, where v is the local wind speed for a coinciding time, and $|\Delta t|$ is a time difference between the time of the satellite overpass and the time of a given AIS ship position. In this way, we obtain a trajectory that we call a *wind-shifted ship track*. An illustration of a *wind-shifted ship track* is depicted in Figure 4.2b. Both wind speed and wind direction are assumed to be constant for the whole time during which we study the plume.

The assumption of constant speed and direction of the wind may create uncertainties in the expected position of the plume of the ship. Therefore, in the third step, we calculate the *extreme wind-shifted tracks*, by adding the margin of wind-related uncertainty to each side of the *wind-shifted ship track* – c.f. Figure 4.2c. The applied values of wind speed and wind direction uncertainties are sub-optimal and are equal to 2 m/s and 20° respectively. The *extreme wind-shifted tracks* define the borders of the RoI of the analyzed ship that we refer to as a *ship sector*.

The radius of the *ship sector* is determined as a maximum distance from the position of the ship at the moment of the satellite overpass to the position of the ship at the earliest moment of the observation (1 hour for Mediterranean Sea, 40 minutes for Arabian Sea) in accordance to *ship track*, *wind-shifted ship track*, or *extreme wind-shifted tracks* (the furthest point is taken into consideration). The *ship sector* delineates the area within which we study the plume produced by the analyzed ship. In Figure 4.2d an example of a resulting RoI that we call a *ship sector* is presented.

4.2.3 NO_2 estimation and model performance evaluation

Within each *ship sector*, we separated pixels that were determined as a plume by thresholding the enhanced image. The applied threshold is equal to the 25^{th} percentile of the values of the pixels lying inside the sector. The average value of NO_2 pixels determined as a plume held as the evaluation value of NO_2 for the corresponding ship and was used for the comparison with the emission proxy defined in Section 2.4.

The quality of ship-plume assignment was reported in terms of the Pearson correlation between the assigned value of NO_2 and the emission proxy E_s (see Chapter 2) for all ships in an analyzed area. Note that the used simple model of emission proxy does not reflect all factors that are needed for a precise estimation of ship emission potential; as a result, it is not expected to lead to a perfect correlation with experimentally determined values of NO_2.

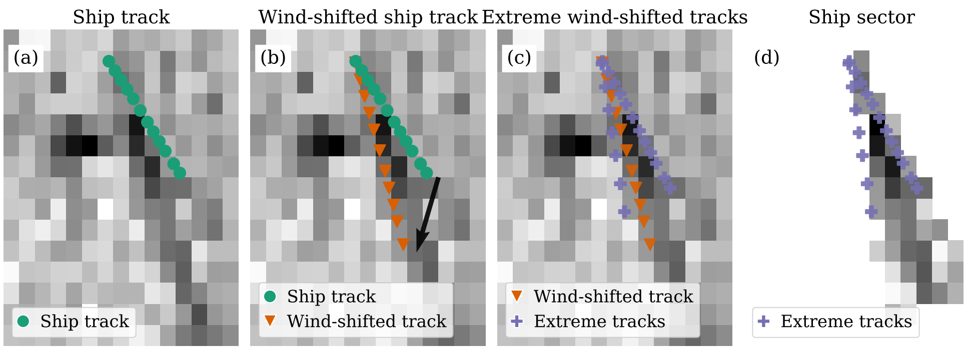

Figure 4.2: Ship sector definition pipeline. Background – the TROPOMI NO_2 signal around the analyzed ship. Two ship plumes can be distinguished at the central part of the image, and one, less intense – in the bottom right of the image. Only one is of interest here. (a) *Ship track* – estimated, based on AIS data records. (b) *Wind-shifted ship track* – a ship track shifted in accordance with the speed and direction of the wind. It indicates the expected position of the ship plume. A black arrow indicates the wind direction. (c) *Extreme wind-shifted ship tracks* – calculated, based on wind information with assumed uncertainties; define the borders of the *ship sector*. (d) A resulting *ship sector* – an ROI of an analyzed ship.

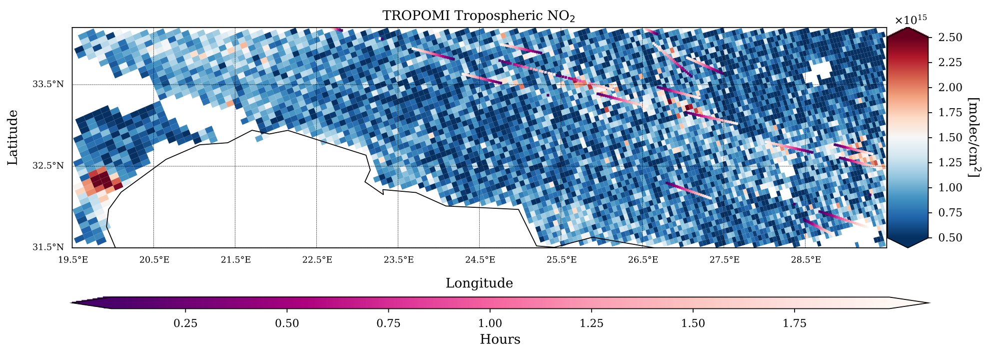

Figure 4.3: The NO_2 tropospheric vertical column density. Region: Mediterranean Sea, restricted by the Northern coasts of Libya and Egypt from the south and the South coast of Crete from the north. Date: April 2nd, 2019. Magenta lines indicate ships' tracks based on information from AIS data. The native local size of the TROPOMI pixels is presented in the Figure.

Date	Region	Long min	Long max	Lat min	Lat max
02-04-19	Mediterranean	21.5	29.5	32	34.5
07-08-19	Mediterranean	19.5	27	33	35
27-09-19	Mediterranean	20	24.74	32.5	35
11-04-20	Arabian	64	69	14	18
01-05-20	Arabian	64	71	14	18
03-05-20	Arabian	64	70	12	18

Table 4.1: The boundaries and dates of observation of six analyzed scenes in the Mediterranean and the Arabian Sea.

4.2.4 Data selection

For the experiments, we chose six days with suitable weather conditions (wind < 6 m/s, low level of background pollution) in the Arabian and Mediterranean Sea (see Table 4.2). An example of an analyzed area can be found in Figure 4.3. The boundaries of the analyzed areas are provided in Table 4.1. The differences in the coordinates of the boundary boxes can be explained by shifts in the TROPOMI orbit coverages. To ensure the high quality of the satellite signal, the following filtering criteria were applied to TROPOMI data: $qa\_value > 0.5$, $cloud\ fraction < 0.5$. For a detailed description of the variables see [93]. Some pre-processing steps were applied to the AIS ship data. First, only ships with an overall length exceeding 150 m and speed above 12 kt[1] were taken into account. As was shown in the previous Chapter, the emission levels of smaller or slower ships are expected to be below the detection level of the satellite. Second, ships of which the location was known for less than 25 minutes prior to the TROPOMI overpass were excluded from the study. Finally, if the plume of a ship undergoes intersection with any other plume of sufficient concentration, the ship was excluded from the study.

4.3 Results

The results of the experiments are summarized in Table 4.2. We compared the results obtained with our approach with the baseline method of ship track shift [41]. For five out of the six analyzed regions, the method proposed in this study led to an improvement in the quality of the linear correlation between the estimated NO_2 values and the emission proxy of the corresponding ship. The weighted mean value of the

[1] kt - knot, a unit of speed equal to a nautical mile per hour. 12 kt $\approx$ 6.2 m/s.

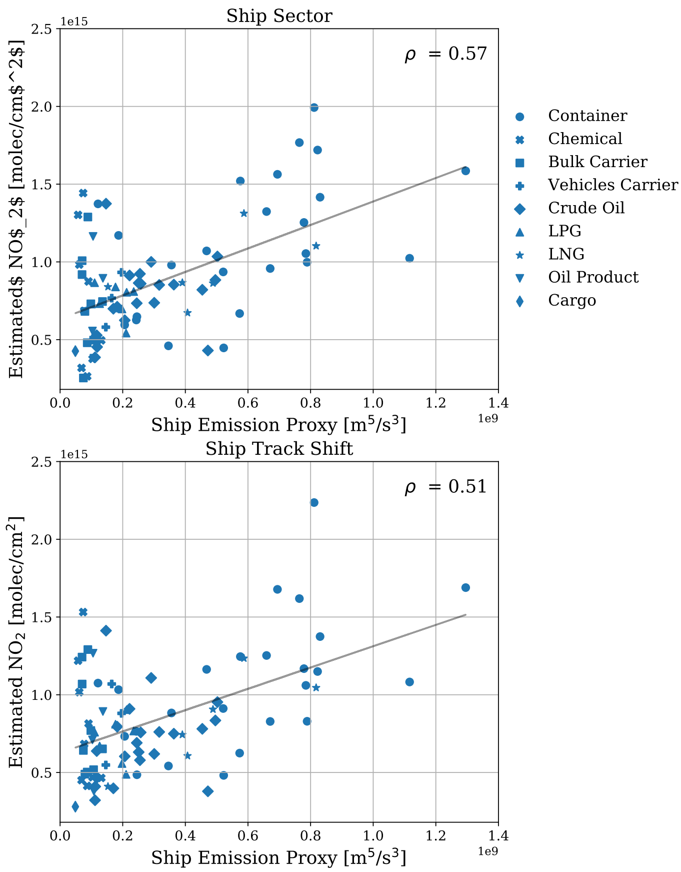

Figure 4.4: Aggregated correlation between the assigned NO_2 values and the corresponding values of the emission proxy for the method proposed in this study (top) and the baseline approach (bottom). Markers indicate the types of ships.

Date	Region	Ship sector	Ship track shift	Detected	Undetected
02-04-19	Med.Sea	**0.81**	0.24	8	1
07-08-19	Med.Sea	**0.61**	0.60	17	1
27-09-19	Med.Sea	**0.78**	0.77	9	1
11-04-20	Arabian	**0.83**	0.67	15	1
01-05-20	Arabian	0.65	**0.71**	20	0
03-05-20	Arabian	**0.80**	0.63	16	2
	AVG/SUM:	**0.73**	0.63	85	6
	STD:	**0.09**	0.13		

Table 4.2: The Pearson correlation coefficients of estimated NO_2 with the corresponding value of the ship's emission proxy for each of the analyzed scenes. The **ship sector** column presents the results achieved with the method introduced in this study. The **ship track shift** column shows the results obtained with the baseline method [41].

Pearson coefficient (the weights are the number of ships analyzed within each area) increased from 0.63 to 0.73, whereas the weighted standard deviation decreased from 0.13 to 0.09. Finally, we calculated an aggregated correlation of NO_2 estimations for all six days analyzed with the associated emission proxies. The results are presented in Figure 4.4. This experiment allowed us to assess the generalization properties of the proposed approach. We can see that the level of achieved correlation decreased for both methods. This was expected due to the presence of smaller ships, whose detectability is highly dependent on the instrument's sensitivity at a given moment of time. Nevertheless, the approach proposed in our study still assures higher quality of the results.

4.4 Conclusions

In this Chapter, we introduced the first approach for the automated evaluation of NO_x emissions of individual seagoing ships with TROPOMI data. We applied a spatial association statistic, local Moran's I, to enhance the separability between the ships' plumes and the background of comparable concentration. We then proposed our method for automated assignation of the RoI to the analyzed ship. We call the proposed RoI a *ship sector*. It is defined based on the uncertainties in the speed and direction of the wind data. We assume that the emission plume produced by a studied ship will be located within the assigned *ship sector*. With this, we addressed the **RQ4** of the thesis: How to assign a TROPOMI signal associated with a certain plume to a potential emitting ship? We then performed a separation of plume-related pixels from the background and estimated the amount of NO_2 associated with a given ship. Since this is the first attempt of NO_2 estimation from individual ships using satellite data, the plume of the ship within the *ship sector* was segmented using the local threshold (**RQ5**). To validate the proposed approach, we used a theoretically derived ship emission proxy. We compared the results obtained using our method with the approach proposed in a previous study. The comparison showed that our method leads to the increment of the linear correlation between estimated values of NO_2 and model-based emission proxy for five of the six analyzed areas, as well as for an aggregated experiment, where the Pearson correlation was calculated for all six analyzed areas at once. We, however, see that the achieved levels of Pearson correlations cannot be characterized as very high. Partially, this can be explained by the fact that the emission proxy used for the validation of the approach does not account for all factors influencing the emission level of the ship. Nevertheless, this also suggests the need for a more complex method of ship plume segmentation.

To conclude, the proposed approach assures more precise quantification of local NO_2 concentrations caused by NO_x emissions of individual ships than the previous (and the only existing study). Moreover, the method introduced in this study does not require any manual steps which is a significant improvement over the current state-of-the-art. As a future research direction, we would like to propose the application of more complex techniques (e.g. machine learning) for the task of ship plume segmentation within the presented here ship's RoI – *ship sector*. Such a methodology will be discussed in the next chapter.