

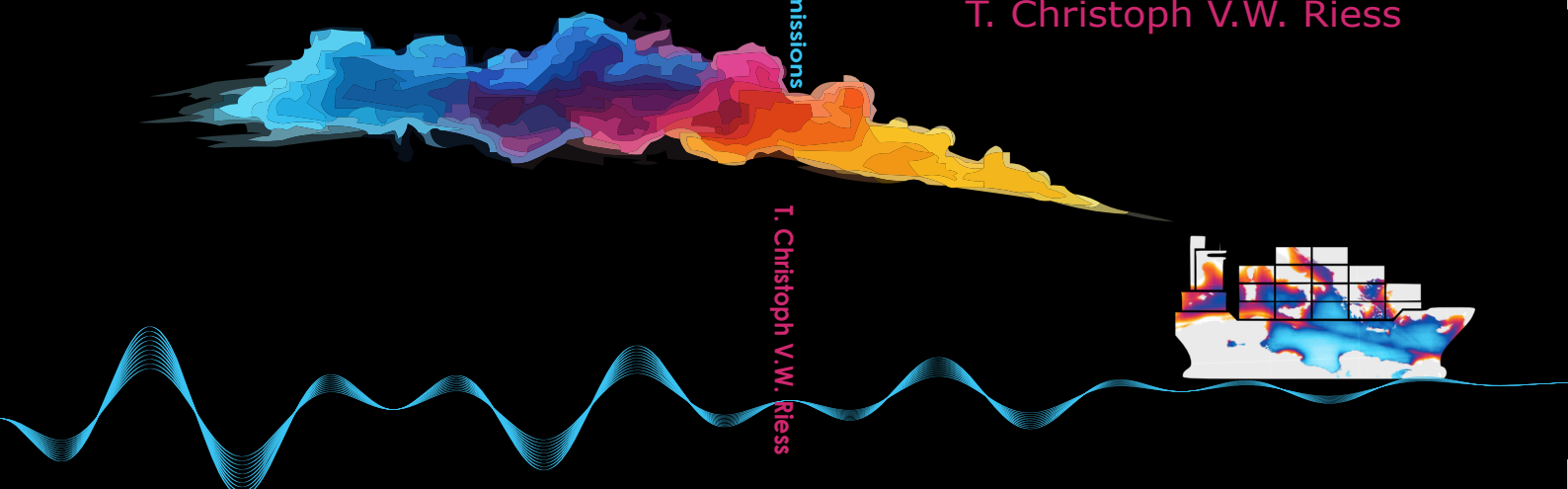
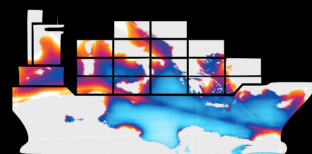


Spaceborne monitoring of nitrogen oxides emissions from ships with TROPOMI

T. Christoph V.W. Riess

Spaceborne monitoring of nitrogen oxides emissions
from ships with TROPOMI

T. Christoph V.W. Riess



Propositions

1. Current IMO regulations for ship NO_x emissions are inherently ineffective.
(this thesis)
2. Ship NO_x plume detection by satellites is not primarily limited by cloudiness.
(this thesis)
3. Propositions are the academic equivalent of clickbait.
4. Supporting Bayer 04 Leverkusen builds the patience and resilience essential for a PhD.
5. Granting voting rights to non-citizen residents offers more benefits than drawbacks.
6. Any form of climate activism benefits the climate more than the criticism thereof.

Propositions belonging to the thesis, entitled

Spaceborne monitoring of nitrogen oxides emissions from ships with TROPOMI

T. Christoph V. W. Riess

Wageningen, 28 October 2024 (Date defence ceremony)

**Spaceborne monitoring of
nitrogen oxides emissions
from ships with TROPOMI**

T. Christoph V. W. Riess

Thesis committee

Promotors:

Dr K.F. Boersma

Associate Professor, Meteorology and Air Quality

Wageningen University & Research

Research Scientist, R&D Satellite Observations

Royal Netherlands Meteorological Institute (KNMI), de Bilt

Prof. Dr Wouter Peters

Personal chair, Meteorology and Air Quality

Wageningen University & Research

Professor of Atmospheric Composition Modelling

University of Groningen

Other members:

Prof. Dr K.M. de Beurs, Wageningen University & Research

Prof. Dr I. Aben, SRON Netherlands Institute for Space Research, Leiden

Prof. Dr G.J.M. Velders, Utrecht University

Prof. Dr M. Sofiev, Finnish Meteorological Institute, Helsinki, Finland

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE)

Spaceborne monitoring of nitrogen oxides emissions from ships with TROPOMI

T. Christoph V. W. Riess

Thesis

submitted in fulfillment of the requirements for the degree of doctor at

Wageningen University

by the authority of the Rector Magnificus

Prof. Dr C. Kroeze,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Monday 28 October 2024

at 10:30 a.m. in the Omnia Auditorium.

T. Christoph V. W. Riess

Spaceborne monitoring of nitrogen oxides emissions from ships with TROPOMI,
218 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2024)

With references, with summary in English

DOI <https://doi.org/10.18174/671443>

Summary

Nitrogen oxides ($\text{NO} + \text{NO}_2 = \text{NO}_x$) are toxic gases and lead to the formation of ozone and secondary pollutants, affecting human and environmental health. Besides small emissions from natural processes such as lightning and soil bacterial activity, they are emitted mainly in the form of NO in combustion processes, e.g. in power plants or ship engines. International shipping is one of the major contributors to NO_x pollution. With a large fraction of these emissions happening close to the coast, an estimated 800,000 premature deaths per year are attributed to ship emissions. The International Maritime Organization (IMO) has therefore established regulations to reduce NO_x emissions from ships. For these regulations to be effective, monitoring of NO_x emissions is necessary. Traditional approaches to monitor NO_x emissions are based on in-situ or ground-based remote sensing measurements, which are time-consuming, expensive and typically limited to coastal proximity. Therefore, satellite remote sensing of NO_2 has been proposed as a complementary tool to monitor NO_x emissions from ships.

In this thesis we study **if** and **how** data from the state-of-the-art Tropospheric Monitoring Instrument (TROPOMI) can be used to monitor ship NO_x emissions from space as part of the *Algorithms for the Verification of Emissions from Shipping with Satellites* (AVES-oculus) project. The project is a collaboration between groups at Wageningen University, Leiden University and the Dutch Inspectorate for Environment and Transport (ILT). AVES-oculus aims to develop a methodology to monitor ship NO_x emissions from space using TROPOMI data, bringing together expertise of atmospheric sciences, remote sensing, data science and inspection of shipping emissions. This thesis approaches the topic from the atmospheric science and remote sensing perspective.

After an introduction to the research topic and some key aspects in the TROPOMI retrieval in **Chapters 1 and 2**, respectively, we start with an evaluation of the TROPOMI NO_2 product under specific measurement conditions in **Chapter 3**: partly cloudy and sun glint (the reflection of sunlight via the ocean surface) scenes. While sun glint scenes are typically flagged as lower quality retrievals, we show that sun glint enhances TROPOMI's capabilities to detect NO_2 close to the sea surface and results in reliable data. During the time of writing, a new version of the TROPOMI NO_2 product

was released, which includes an updated algorithm for the cloud height retrieval, relevant in the retrieval of NO_2 columns. We show that the updated retrieval indeed improves the agreement of cloud height with independent data and substantially increases the NO_2 columns in partly cloudy scenes, reducing the known bias in the previous TROPOMI version. After having dealt with the data quality of TROPOMI NO_2 columns, we use the data to detect NO_2 pollution over several major shipping lanes around Europe, of which several had not been detected from space before. Lastly, we study the changes in the NO_2 above some of these shipping lanes during the COVID-19 lockdown in 2020. We find that the NO_x emission in the shipping lanes decreased by 20% during the lockdown, which is less than the decrease observed over land, but agrees reasonably well with changes in independent ship activity data.

Chapter 4 tackles another major source of uncertainty in TROPOMI NO_2 data: the vertical NO_2 concentration profiles. For this chapter, 10 vertical profiles of NO_2 and other trace gases were sampled with an aircraft above the North Sea in summer 2021. We find that compared to models used in the TROPOMI retrieval, the lowest 150 m exhibit more NO_2 , while models overestimate the NO_2 in the layers above. We show that when we adjust the TROPOMI data for the correct vertical distribution of NO_2 , the resulting columns increase by 20% on average. Over land, the models agree better with observed profiles. Together with the dark sea surface, this highlights the challenge of space-based NO_2 retrievals over sea. We conclude that the reported negative bias in TROPOMI columns over sea has implications for the interpretation of TROPOMI NO_2 data, for example for estimations of emissions from shipping.

Chapter 5 forms the core of this thesis. Here, we move from the evaluation of TROPOMI NO_2 data to the estimation of NO_x emission fluxes and factors of individual ships. We match observed plumes in TROPOMI NO_2 over the eastern Mediterranean in 2019 with ship location and identity data, identifying 130 plumes of 119 different ships. Using a Gaussian plume model with atmospheric chemistry, which is capable of simulating the evolution of a plume, we create a large library of NO_2 , capturing a wide range of meteorological and chemical conditions. We then use the members of this library to train a machine learning XGBoost model to learn the relationship between emissions of NO_x and observable NO_2 . By applying this model to the observed TROPOMI data, we create the first-ever space-based estimates of NO_x emissions from individual ships. Most of the observed ships are very large container ships, and the high ozone values in summer increase the chance of detection as it facilitates the conversion of emitted NO to observable NO_2 . We also find that - contrary to expectation - newer ships that have to follow more strict emission regulations, emit more NO_x per power used than older ships in real world conditions. This is especially true when running at low engine loads which is the most common mode of operation observed. This finding has severe implications for the effectiveness of the IMO regulations.

Lastly, in **Chapter 6** we use our plume model to predict how diurnal and seasonal cycles of meteorological and chemical parameters in the atmosphere change the detectability of ship plumes. This is of interest, as recent and upcoming satellites enable the monitoring of NO_2 at hourly temporal resolution during daytime versus TROPOMI's daily overpass. We find that TROPOMI's overpass is already ideal for the detection of ship plumes, mostly caused by the chemical conditions. We also find that summer months are more favorable for the detection of ship plumes, as the higher ozone values increase the conversion of emitted NO to observable NO_2 , even when ignoring the higher frequency of cloud free condition. Even though detectability is best in the early afternoon, hourly overpasses will produce a significant increase in the number of detected plumes, especially in summer.

Zusammenfassung

Stickstoffoxide (NO_x) - eine Sammelbezeichnung für Stickstoffmonoxid (NO) und Stickstoffdioxid (NO_2) - sind giftige Gase, die zur Bildung von Ozon und sekundären Schadstoffen führen und dadurch die menschliche Gesundheit sowie die Umwelt beeinträchtigen. Neben kleineren Emissionen in natürlichen Prozessen wie Blitzschlag und bodenbakterieller Aktivität werden sie hauptsächlich bei Verbrennungsprozessen ausgestoßen. Bei der Verbrennung in Kraftwerken oder Schiffsmotoren entsteht hauptsächlich Stickstoffmonoxid (NO). Dabei ist der internationale Schiffsverkehr einer der Hauptverursacher von NO_x -Verschmutzung. Da der Schiffsverkehr hauptsächlich in der Nähe zum Festland stattfindet, hat dies auch direkte Folgen auf uns Menschen. Es wird geschätzt, dass etwa 800.000 vorzeitige Todesfälle auf die Stickoxiverschmutzung aufgrund der Seefahrt zurückgeführt werden können. Die Internationale Seeschiffahrts-Organisation (IMO) hat daher Vorschriften zur Reduzierung der NO_x -Emissionen von Schiffen erlassen. Um die Wirksamkeit dieser Vorschriften zu garantieren und beurteilen zu können, ist die Überwachung der NO_x -Emissionen erforderlich. Traditionelle Ansätze dazu basieren auf in-situ- oder bodengestützten Fernerkundungsmessungen. Diese Methoden sind allerdings sehr zeitaufwändig, teuer und daher typischerweise auf küstennahe Bereiche beschränkt. Daher liegt es nahe die satellitengestützte Fernerkundung von NO_2 als ergänzendes Instrument zur Überwachung von NO_x -Emissionen aus Schiffen zu nutzen.

In dieser Doktorarbeit untersuchen wir im Rahmen des Projekts *Algorithms for the Verification of Emissions from Shipping with Satellites* (AVES-oculuS), **ob** und **wie** Daten des hochmodernen Tropospheric Monitoring Instrument (TROPOMI) verwendet werden können, um NO_x -Emissionen von Schiffen aus dem Weltraum zu überwachen. Das Projekt ist eine Zusammenarbeit zwischen Gruppen der Universität Wageningen, der Universität Leiden und der niederländischen Inspektion für menschliche Umwelt und Verkehr (ILT). Ziel des Projekts ist es, eine Methodik zur Überwachung von Schiffsemissionen von NO_x aus dem Weltraum unter Verwendung von TROPOMI-Daten zu entwickeln, indem Fachwissen aus den Bereichen Atmosphärenwissenschaften, Fernerkundung, Datenwissenschaft und Inspektion von Schadstoffemissionen der Schifffahrt zusammengeführt wird. Diese Arbeit betrachtet das Thema aus der Perspektive der Atmosphärenwissenschaften und der atmosphärischen Fernerkundung.

Nach einer Einleitung in das Forschungsfeld sowie wichtigen Einzelheiten des TRPOMI-Algorithmus in den **Kapiteln 1 und 2** beginnen wir in **Kapitel 3** mit einer Untersuchung der TROPOMI-NO₂-Daten unter spezifischen Messbedingungen: teilweise bewölkte Pixel und Szenen unter *sun glint* (der Reflexion von Sonnenlicht durch die Meeresoberfläche). Während *sun glint* typischerweise als qualitativ minderwertige Messungen markiert sind, zeigen wir, dass es in der Praxis die Fähigkeiten von TROPOMI zur Erkennung von NO₂ in Meeresnähe verbessert und zuverlässige Daten liefert. Zudem wurde zum Zeitpunkt des Verfassens eine neue Version des TROPOMI-NO₂-Produkts veröffentlicht, die einen aktualisierten Algorithmus für die Wolkenhöhenbestimmung enthält. Die Wolkenhöhe ist für die Erfassung von NO₂-Säulen in teils bewölkten Pixeln sehr relevant. Wir zeigen, dass die aktualisierte Version tatsächlich die Übereinstimmung der Wolkenhöhe mit unabhängigen Daten verbessert und die NO₂-Säulen in teilweise bewölkten Szenen deutlich nach oben korrigiert, wodurch der bekannte Bias im Vergleich zur vorherigen TROPOMI-Version verringert wird. Nachdem wir uns mit der Datenqualität der TROPOMI-NO₂-Säulen befasst haben, verwenden wir die Daten, um NO₂-Verschmutzung über mehreren wichtigen Schiffsrouten rund um Europa zu detektieren, von denen einige erstmals aus dem Weltraum erfasst wurden. Schließlich untersuchen wir die Veränderungen des NO₂ über einigen dieser Schiffsrouten während des COVID-19-Lockdowns im Jahr 2020. Wir stellen einen Rückgang der NO_x-Emissionen in den Schiffsrouten während des Lockdowns um 20% fest, der damit geringer ausfiel als über Land, aber relativ gut mit den unabhängig ermittelten Änderungen in den Schiffsaktivitätsdaten übereinstimmt.

In **Kapitel 4** behandeln wir eine weitere bedeutende Quelle von Unsicherheiten in den TROPOMI-NO₂-Daten: die vertikalen NO₂-Profile. Für dieses Kapitel wurden 10 vertikale Profile von NO₂ und anderen Spurengasen mit einem Flugzeug über der Nordsee im Sommer 2021 gemessen. Wir stellen fest, dass im Vergleich zu den Modellen, die im TROPOMI-Algorithmus verwendet werden, die unteren 150 m in unseren Messungen mehr NO₂ aufweisen, während die Modelle das NO₂ in den darüber liegenden Schichten überschätzen. Wir zeigen darüber hinaus, dass es bei Anpassung der genutzten Profile an die direkt gemessene vertikale Verteilung von NO₂ zu einer durchschnittlichen Erhöhung der TROPOMI-NO₂-Säulen um 20% kommt. Über Land stimmen die Modelle besser mit beobachteten Profilen überein. Dies verdeutlicht die besondere Herausforderung der satellitengestützten NO₂-Erfassung über dem Meer, die zudem durch die dunkle Meeresoberfläche erschwert wird. Der gefundene negative Bias in den TROPOMI-Säulen über dem Meer hat Auswirkungen auf die Interpretation der TROPOMI-NO₂-Daten, insbesondere für Schätzungen von Schiffsemissionen.

Kapitel 5 bildet den Kern dieser Arbeit. Hier wechseln wir von der Bewertung der TROPOMI-NO₂-Daten und Änderungen auf langen Zeitskalen zur Schätzung der NO_x-Emissionsflüsse und -faktoren einzelner Schiffe. Wir gleichen beobachtete

Abgasfahnen in TROPOMI-NO₂ über dem östlichen Mittelmeer im Jahr 2019 mit Standort- und Identitätsdaten von Schiffen ab und ordnen 130 Abgasfahnen 119 verschiedenen Schiffen zu. Unter Verwendung eines Gauss'schen Abgasfahnenmodells mit atmosphärischer Chemie erstellen wir eine große Bibliothek von NO₂-Fahnen, die eine breite Palette meteorologischer und chemischer Bedingungen abbildet. Anschließend verwenden wir die Mitglieder dieser Bibliothek, um ein XGBoost-Modell - ein Algorithmus maschinellen Lernens - zu trainieren, um die Beziehung zwischen NO_x-Emissionen und beobachtbarem NO₂ zu erlernen. Durch Anwendung dieses Modells auf die beobachteten TROPOMI-Daten erstellen wir die weltweit ersten satellitengestützten Schätzungen der NO_x-Emissionen einzelner Schiffe. Die meisten beobachteten Schiffe sind große Containerschiffe, und die hohen Ozonwerte im Sommer erhöhen die Chance der Detektion, da sie die Umwandlung von ausgestoßenem NO in beobachtbares NO₂ erleichtern. Wir stellen auch fest, dass - im Gegensatz zur Erwartung - neuere Schiffe, die strengeren Emissionsvorschriften unterliegen, mehr NO₂ pro genutzter Leistung ausstoßen als ältere Schiffe. Dies trifft besonders dann zu, wenn sie bei niedriger Motorenlast betrieben werden, was der häufigste beobachtete Betriebsmodus ist. Diese Erkenntnis hat schwerwiegende Auswirkungen auf die Wirksamkeit der IMO-Vorschriften in Bezug auf die Luftqualität.

Abschließend verwenden wir in **Kapitel 6** unser Abgasfahnenmodell, um vorherzusagen, wie sich tageszeitliche und saisonale Zyklen meteorologischer und chemischer Parameter in der Atmosphäre auf die Erkennbarkeit von Abgasfahnen von Schiffen auswirken. Dies ist von Interesse, da neue und kommende Satelliten die Überwachung von NO₂ mit stündlicher Auflösung während des Tages ermöglichen, im Gegensatz zu TROPOMI, das einen täglichen Überflug hat. Wir stellen fest, dass der Zeitpunkt des TROPOMI-Überflugs im frühen Nachmittag bereits ideal für die Erkennung von Abgasfahnen von Schiffen ist, hauptsächlich aufgrund der chemischen Bedingungen. Darüber hinaus zeigen wir, dass die Sommermonate günstiger für die Erkennung von Abgasfahnen von Schiffen sind, da die höheren Ozonwerte die Umwandlung von ausgestoßenem NO in beobachtbares NO₂ erhöhen, selbst wenn man davon absieht, dass im Sommer seltener Wolken die Beobachtungen stören. Obwohl die Erkennbarkeit am frühen Nachmittag am besten ist, werden stündliche Überflüge zu einer signifikanten Zunahme der Detektion von Abgasfahnen führen - und damit präzisere Emissionsbestimmungen.

Contents

	Page
Summary	v
Zusammenfassung	ix
Contents	xiii
Chapter 1 Introduction	1
1.1 Atmospheric composition and structure	2
1.2 Nitrogen oxides	3
1.3 Shipping industry	7
1.4 Satellite remote sensing	10
1.5 Modelling of NO _x plumes	14
1.6 Thesis overview	16
Chapter 2 Key aspects in the TROPOMI NO₂ retrieval	19
2.1 Retrieval versions	20
2.2 The importance of clouds	20
2.3 The importance of a-priori profiles	23
2.4 The importance of surface albedos	25
Chapter 3 Improved monitoring of shipping NO₂ with TROPOMI	27
3.1 Introduction	29
3.2 Methods and Materials	31
3.3 Results	36
3.4 Discussion and Conclusion	49
3.A Appendix	51
3.S Supplement	59
Chapter 4 To new heights by flying low	65
4.1 Introduction	67
4.2 Materials	69

4.3	Aircraft NO ₂ interpretation and representation at the scale of a TROPOMI pixel	73
4.4	Validation of TROPOMI NO ₂ over the North Sea	81
4.5	Discussion	83
4.6	Conclusion	86
4.S	Supplement	88
Chapter 5	Estimating NO_x emissions of individual ships	101
5.1	Introduction	103
5.2	Methods & Materials	104
5.3	Results	110
5.4	Discussion	116
5.S	Supplement	119
Chapter 6	Temporal variability in ship NO₂ plume detectability	139
6.1	Introduction	141
6.2	Material and methods	142
6.3	Results	148
6.4	Discussion & Conclusion	154
Chapter 7	Synthesis	157
7.1	Introduction	158
7.2	Satellite remote sensing of tropospheric NO ₂ and other trace gases	158
7.3	Methodological advances	163
7.4	NO _x emissions from international shipping	167
References		173
About the author		197
Acknowledgements		199

Chapter 1

Introduction

Air pollution, a byproduct of global industrialization, poses one of the most significant threats to human health and the environment worldwide. Each year, it is associated with 4.2 million premature deaths (WHO, 2024). The United Nations acknowledges its gravity by integrating the pursuit of clean air into three of its Sustainable Development Goals, namely Goals 3 (*good health and well-being*), 7 (*affordable and clean energy*), and 11 (*sustainable cities and communities*) (United Nations, 2015).

Air pollution comes in the form of particulate matter (PM) and many trace gases. Among those polluting gases are nitrogen oxides ($\text{NO}_x = \text{NO}_2 + \text{NO}$), sulfur dioxide (SO_2), carbon monoxide (CO) and ozone (O_3). Sources of air pollutions are as diverse as the pollutants themselves. They range from natural sources such as volcanoes and wildfires to anthropogenic sources including traffic, industry and shipping. In contrast to other sectors like road transport and industry, regulation of air pollutions from ocean-going vessels (OGVs) was long missing, and the current regulation lacks a global monitoring. This thesis therefore focuses on using space-based data of air pollution over sea, more concretely on NO_x emitted by OGVs.

In this chapter, the context of the thesis is described, starting with a brief description of the atmosphere in Section 1.1 and a detailed description of NO_x and its chemistry in Section 1.2. This is followed by a description of the shipping industry and the emission regulations that apply to it in Section 1.3. Finally, the use of satellite remote sensing for monitoring tropospheric nitrogen dioxide and its history in ship emission monitoring is described in Section 1.4. In Section 1.6, the remainder of the thesis is summarized and the structure of the thesis is described.

1.1 Atmospheric composition and structure

The Earth is surrounded by a relatively thin layer of air, called the atmosphere. The atmosphere is composed of several layers, each with distinct characteristics. The lowest layer, the troposphere, extends from the Earth's surface to an altitude of about 8-15 km, depending on the latitude and season. The troposphere contains about 80% of the atmosphere's mass. It is characterized by a typically decreasing temperature with altitude. The layer above the troposphere is the stratosphere, which extends from the tropopause - which marks the upper boundary of the troposphere - to about 50 km. The stratosphere contains the ozone layer, which absorbs most of the sun's ultraviolet radiation. Above the stratosphere lie the mesosphere and thermosphere. The exosphere is the outermost layer of the atmosphere and extends to about 10,000 km above the Earth's surface. It is where the atmosphere gradually transitions into outer space.

Our focus will be the troposphere. Within the troposphere, one often distinguishes between the boundary layer and the free troposphere. The boundary layer is the lowest part of the troposphere and is characterized by strong vertical mixing, i.e. in sufficient distance from a source the composition of the air within the boundary layer does not

change with altitude. One of the drivers of this mixing during daytime is convection, caused by the sun, which heats up the Earth's surface, subsequently warming the air above, which starts to rise. The boundary layer is where most of the Earth's weather occurs and where we humans live. It is typically a few hundred meters to a few kilometers deep, depending on the time of day and the weather. The depth of the boundary layer determines the volume over which pollution is spread, meaning that shallow boundary layers lead to higher concentration of pollutants. The marine boundary layer - which is discussed in this thesis - is typically shallow, as water surfaces warm more slowly compared to land surfaces and therefore suppress convection of the air above.

The (dry) air in the troposphere consists of several gases of which nitrogen (N_2) and oxygen (O_2) are the most abundant. Nitrogen makes up about 78% of the atmosphere, while oxygen makes up about 21%. The remaining 1% consists of trace gases such as the well-known Greenhouse Gases (GHGs) carbon dioxide (CO_2) and methane (CH_4), as well as ozone (O_3), nitrogen oxides (NO_x), and others. The abundance of trace gases is typically reported as mixing ratios of the trace gas relative to the dry air in *parts per million (ppm)*, *parts per billion (ppb)*, or *parts per trillion (ppt)*. The local mixing ratio of trace gases is not constant and can vary with location and time. In the case of nitrogen oxides, mixing ratios can range from a few hundred ppt in clean air e.g. over seas, to up to some hundred ppb in cities, and even higher values in exhaust plumes. Even though NO_x abundances are typically very low, they play an important role as air pollutant. In the following section we will have a closer look into its chemistry and sources.

1.2 Nitrogen oxides

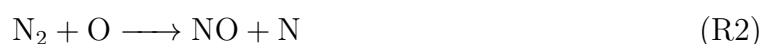
Nitrogen dioxide (NO_2) and nitric oxide (NO) are commonly referred to as nitrogen oxides (NO_x). NO_x is a key air pollutant in itself (Huang et al., 2021), and - in the presence of Volatile Organic Compounds (VOCs) and sunlight - contributes to the formation of ground-level ozone and fine particulate matter, which are harmful to human health and the environment (Scheel et al., 1959; Brown et al., 1950; Todd and Garber, 1958). NO_x is also a precursor to acid rain and eutrophication on land and water bodies. This also makes it - although with a smaller contribution compared to agricultural ammonia - one of the causes of the nitrogen excess in the environment. In Europe - and especially in the Netherlands - the large nitrogen excess has been recognized and mitigation regulations are one of the dominant political and regulatory issues in recent years (de Vries, 2021). Here, we will discuss the atmospheric chemistry involving NO_x and its sources.

1.2.1 NO_x formation and chemistry

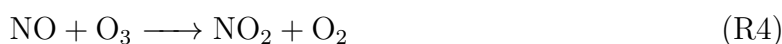
NO_x is emitted during combustion processes, such as in power plants, industrial facilities and vehicles, mainly in the form of NO. NO_x is also formed naturally during lightning and in soils. The formation in combustion processes is caused by the heat causing molecular oxygen (O₂) to form atomic oxygen:



The resulting oxygen atoms (O) then react with molecular nitrogen (N₂) - the most abundant gas in the atmosphere - to form NO:



The produced NO can then be oxidized to NO₂ by the reaction with ozone (O₃):



Close to strong sources such as in cities or plumes of smoke stacks, reaction R4 can be limited by the amount of ozone, when there is too much NO present for the available ozone to form NO₂. In that case, ozone levels can locally be reduced significantly, a process called ozone titration.

In the presence of sunlight, NO₂ can be photolyzed to form NO and atomic oxygen, which reacts to ozone, yielding the following net reaction:



Reactions R4 and R5 are called the null cycle, as they do neither produce nor consume NO_x. This cycling happens at a timescale of seconds, explaining why we are usually refer to NO and NO₂ commonly as NO_x.

In the presence of the hydroperoxy radical HO₂ (formed in the oxidation of carbon monoxide) or an organic peroxy radical RO₂ (produced during the oxidation of hydrocarbons such as methane or VOCs, where R represents a more complex molecule of mostly carbon and hydrogen atoms), alternative pathways for reaction R4 exist:



If followed by reaction R5, these alternative pathways to R4 lead to net ozone production. At ground-level, the produced ozone is harmful for humans and the biosphere.

The typical tropospheric lifetime¹ of NO_x is in the order of several hours, but it can be as

¹The tropospheric lifetime τ_X of a substance X is defined as the average time a molecule of X remains in the troposphere and can be calculated as the total mass of X divided by the removal rates, e.g. chemical loss or deposition.

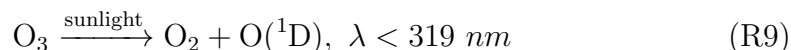
short as a few minutes in urban areas and as long as a few days in remote areas (Kenagy et al., 2018). Besides the spectral characteristics of NO_2 (see Section 1.4) it is this lifetime that makes it ideal for satellite remote sensing as it is short enough to guarantee a low background concentrations but long enough for large (i.e. the size of several satellite pixels) emission plumes to be visible.

The lifetime of NO_x during the day is determined by the rate of reaction with the hydroxyl radical (OH), which is the main oxidant in the atmosphere and often called the *detergent* of the atmosphere as it *cleans* the air of many GHGs and air pollutants. The reaction of NO_2 with OH including a third body M is the main sink of NO_x in the atmosphere:



The resulting nitric acid (HNO_3) is soluble in water and can be removed by wet deposition, potentially leading to nitrogen excess in soils. It is also a precursor to fine particulate matter. Reaction R8 terminates the catalytic ozone production.

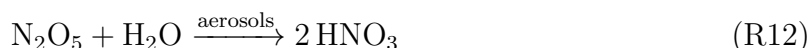
The hydroxyl radical is formed by the photolysis of ozone by UV radiation and the reaction of water vapor with excited oxygen atoms:



We can now understand that the tropospheric chemistry involving ozone and nitrogen oxides is highly nonlinear: Reactions R9 and R10 show that the availability of OH - the main day-time sink of NO_x - in the troposphere is depending on the availability of ozone, which in turn is influenced by NO_x emissions. The lifetime of NO_x can thus partially be dependent on the abundance of NO_x itself (Valin et al., 2014).

Besides the direct human and environmental health impacts, nitrogen oxides via their interaction with ozone and OH thus also play an important role for the oxidizing capacity of the atmospheric boundary layer, and, in the context of shipping emissions, the marine boundary layer. For example, Vinken et al. (2011) have found that ship NO_x emissions lead to an increase in ozone of 4 ppb (15%) over the North Atlantic in summer.

During nighttime, in the absence of OH, NO_2 can be removed from the atmosphere by the reaction with the nitrate radical (NO_3), followed by hydrolysis of N_2O_5 on aerosols:



Lastly, NO_2 can react with organic radicals to form peroxyacetyl nitrate (PAN), which is a reservoir for NO_x and can be transported over long distances:



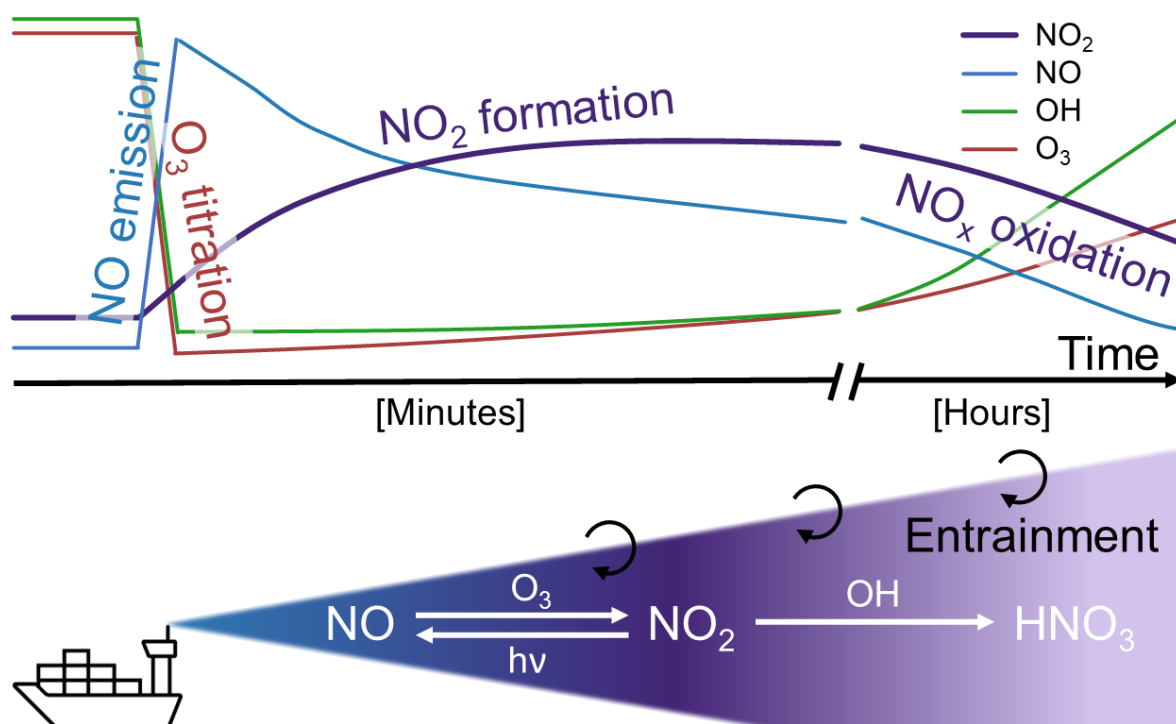


Figure 1.1: Abundance of chemical species in an aging NO_x plume as a function of time after emission. The dashed line show the background concentration before the emission. The graph shows a few hours of plume aging.

Figure 1.1 shows the behavior of chemical species inside a plume with time after the emission of NO_x . Most NO_x is emitted as NO, leading to an instant peak in NO. Inside the fresh plume, all O_3 and OH is titrated by reactions R4 and R8, respectively, and other pollutants that are oxidized by OH and O_3 . This titration leads to an initially long NO_x lifetime inside the plume. Only with the growing of the plume and successive entrainment of fresh, ozone-rich air, O_3 is recovered and NO_2 formed. Nitrogen dioxide therefore peaks only several kilometers downwind of the emissions. NO_x is oxidized faster in the aging plume when OH is recovered. The NO_x chemistry discussed in this section is crucial for interpreting NO_2 plumes observed at the resolution of TROPOMI pixels (see Section 1.4.1 below). This is often done with the help of chemical models as introduced in Section 1.5.

1.2.2 NO_x sources

Globally, anthropogenic emissions of nitrogen oxides are estimated to be around $120 \text{ Tg NO}_2 \text{ yr}^{-1}$ per year ($36.5 \text{ Tg N yr}^{-1}$) (McDuffie et al., 2020) in 2017. Around 60%

¹Here and for the remainder of this chapter, NO_x emission are given as NO_2 mass unless detailed otherwise.

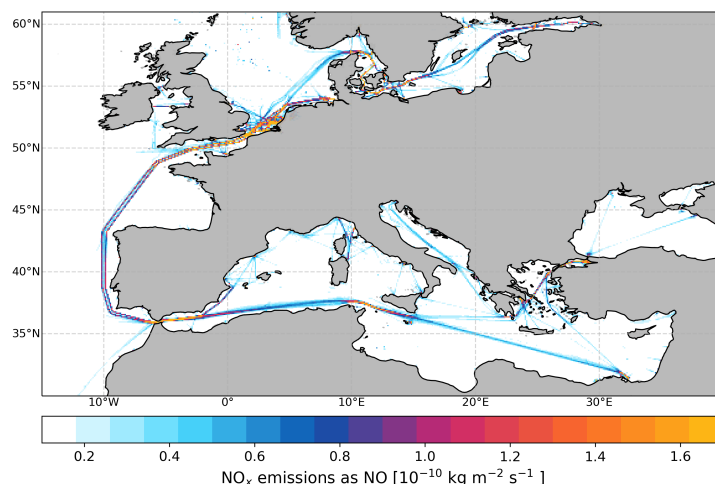


Figure 1.2: Map of shipping related NO_x emissions in April-November 2019 as reported in the CAMS inventory (Granier et al., 2019; Johansson et al., 2017).

of these emissions come from energy generation, industry and road transport. Additionally, NO_x is emitted by natural sources such as lightning, biomass burning and soil microbes. Overall, anthropogenic sources are dominant, but natural sources can play a significant role depending on the region and the time of year (Zhang et al., 2003). Another major source is international shipping, which was estimated to account for 20% (24.4 Tg NO₂ yr⁻¹) of global anthropogenic NO_x emissions. Other studies estimated that shipping emissions are responsible for 7-34 Tg NO₂ yr⁻¹ (Crippa et al., 2018; Eyring et al., 2010; Johansson et al., 2017). There is therefore a large uncertainty connected to shipping emissions and more accurate estimates are needed to assess the impact of shipping emissions on the complex chemistry in the marine boundary layer, human health and the environment. This will become even more relevant in the future, as the shipping industry and its emissions are expected to grow, as we will discuss in the following Section.

1.3 Shipping industry

Globalization and international trade have led to a rapid increase in the number of ocean-going vessels (OGVs) and their emissions. The shipping industry is responsible for a significant fraction of global NO_x emissions as detailed above. OGVs also emit other air pollutants such as SO₂, VOCs and particulate matter including black carbon. In step with increases in global trade volume, the intensity of international shipping is expected to increase in the coming years (UNCTAD, 2019). While many sectors (at least in developed countries) have reduced their emissions in the past decades (e.g. Goldberg et al., 2021; Krecl et al., 2021; Curier et al., 2014; Hassler et al., 2016), the shipping industry has not (De Ruyter de Wildt et al., 2012; Boersma et al., 2015). This is at least partially due to the fact that shipping emissions often happen outside national borders and are

therefore harder to regulate by national policies. Figure 1.2 shows a map of shipping NO_x emissions around Europe in the CAMS inventory (Granier et al., 2019; Johansson et al., 2017), showing busy shipping lanes in the North Sea, Bay of Biscay and crossing the Mediterranean Sea. Additionally, the shipping industry is exempt from many environmental regulations that apply to other sectors. Furthermore, OGVs are typically used for 20-30 years, meaning that the fleet is slow to adapt to new regulations, which usually apply to newly-build ships only.

70% of ship emissions are within 400 km of coastlines (Eyring et al., 2010). This means that the emissions are close to human populations. In fact, shipping-related pollution degrades air quality in many coastal areas with adverse consequences for human and ecosystem health, causing up to 800,000 premature deaths per year (Sofiev et al., 2018). This health penalty motivates a significant societal interest to ensure lower emissions.

1.3.1 Emission regulations

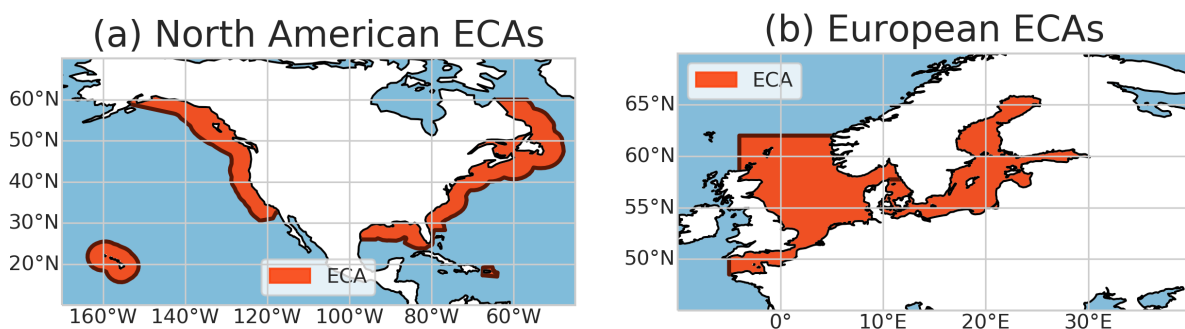


Figure 1.3: Emission Control Areas (ECAs) in (a) North America and (b) Europe.

Anticipating above-mentioned issues, the International Maritime Organization (IMO) has implemented several regulations to reduce the environmental impact of the shipping industry. The most important of these are the MARPOL Annex VI and the designation of Emission Control Areas (ECAs). The MARPOL Annex VI, which came into force in 2005, sets limits on the sulfur content of marine fuels and NO_x emissions from OGVs. The ECAs are designated areas in which stricter emission limits apply than globally. Currently, ECAs are in place in the North Sea and Baltic Sea, the North American coast, and part of the Caribbean, see Fig. 1.3. The current North American and European ECAs in North and Baltic Sea include both nitrogen dioxide and sulfur limits. From 2025 on, the Mediterranean will also be an ECA for sulfur. However, nitrogen limits for the new Mediterranean ECA are still under discussion. Besides the MARPOL Annex VI regulations, the IMO has also set goals on Greenhouse Gases (GHGs) emissions, striving for a reduction of 20% until 2030 and net-zero *close to* 2050 (International Maritime Organization, 2023).

Emission limits for NO_x depend on the year of construction - in three tiers - and engine rated speed of the ship. For sulfur emissions, the IMO has also set a global limit on

Table 1.1: Emission limits for NO_x for different Tiers of the IMO emission regulations. Ships have to follow the emission limit on a weighted average for several engine load points. Adapted from (International Maritime Organization, 2013).

Tier	Ship construction date on or after	Total weighted cycle emission limit (g/kWh)		
		n = engine's rated speed (rpm)		
		n < 130	130 ≤ n ≤ 1999	n ≥ 2000
I	1 January 2000	17	45 n ^{-0.2}	9.8
II	1 January 2011	14.4	44 n ^{-0.23}	7.7
III	1 January 2016	3.4	9 n ^{-0.2}	2

the sulfur content of marine fuels, which came into force in 2020. Table 1.1 shows the emission limits for NO_x for the different emission Tiers at the time of writing. Tier III limits only hold in the nitrogen oxide ECAs, with the European ECA only applying to ships with keel laying date after January 1st 2021. Tier I & II limits do not include a not-to-exceed value. Ship main engines have to comply with the regulations on a weighted average emission factor EF_{wa} (in g/kWh) that is calculated from the emission factors at engine loads of 25, 50, 75 and 100 % with weighting factors of 0.15, 0.15, 0.5 and 0.2, respectively (International Maritime Organization, 2008):

$$EF_{wa} = \sum_{i=1}^4 w_i \cdot EF_i \quad (1)$$

$$= 0.15 \cdot EF_{25\%} + 0.15 \cdot EF_{50\%} + 0.5 \cdot EF_{75\%} + 0.2 \cdot EF_{100\%} \quad (2)$$

This nature of weighted average emissions makes compliance monitoring in the real world difficult. Additionally, the dependency of emission Tiers on keel laying date led to a surge in ships being officially recorded with build dates in 2020 (Van Roy et al., 2023).

To achieve compliance with nitrogen Tier III limits, ships can for example use exhaust gas recirculation (EGR, commonly used in car engines) or selective catalytic reduction (SCR). SCR is the most common method for large ships. In SCR, a catalyst is used to chemically convert NO_x to nitrogen and water, abating 80-90% of NO_x emissions. As these technologies are associated with additional costs for installation, maintenance and a (small) increase in fuel consumption, shipowners might try to avoid compliance. To ensure the envisioned effectiveness and cleaner air, compliance monitoring is necessary. We will discuss monitoring approaches in the following subsection.

1.3.2 Monitoring and enforcement

The regulations are enforced by national authorities, which are responsible for monitoring and verification of ship emissions. Monitoring can be done in various ways, including

the more traditional onboard checks of engine certificates and fuel samples (the latter only for fuel sulfur content). Other methods include ground-based remote sensing of ship plumes (Cheng et al., 2019), and airborne sniffer techniques using helicopters, small aircraft (Mellqvist and Conde, 2021; Chen et al., 2005), or drones (Van Roy and Scheldeman, 2016). Van Roy et al. (2023) find that while sulfur compliance is generally good and growing, nitrogen emissions from ships are actually going up. They also show that compliance to sulfur limits decreases with the distance from the coast. This is likely due to the fact that the majority of the checks are done in port, and that the regulations are not enforced in international waters. Additionally, all the methods mentioned above are expensive and time-consuming, making it impossible to monitor the entire fleet. Because of their cost, these monitoring tools are mainly applied in wealthy countries, possibly leading to further disparity in global air quality. Here, satellite remote sensing with its global coverage and high temporal resolution offers a promising alternative for a globally uniform and potentially near real-time monitoring of ship emissions. We will therefore now turn to satellites and their retrieval of trace gases from space.

1.4 Satellite remote sensing

Satellite remote sensing offers a unique opportunity to monitor the earth in a near real-time and globally uniform manner. There are two main classes of satellites used in remote sensing: low earth orbit (LEO) and geostationary satellites. LEO satellites are closer to the earth and orbit it several times a day flying over the poles. Geostationary satellites are further away from the earth and stay above the same location on earth and can therefore measure the same spot on earth at different times of day. This makes them ideal for monitoring diurnal cycles of surface properties or atmospheric composition. LEO satellites on the other hand cover the full Earth surface over several days and therefore are better suited to map spatial patterns, and changes in them. The applications of satellite remote sensing are countless and range from monitoring destruction from natural disasters like earthquakes or flooding to creating land-use maps and air-pollution monitoring such as the abundance of trace gases.

To retrieve trace gas composition in the atmosphere, satellites use optical absorption features of the constituent at question: Every molecule has a unique absorption spectrum, which can be used to identify and quantify the abundance of the molecule. Differences between solar irradiance and earth radiance spectra are attributed to absorptions by molecules in the air in the different light paths. With good knowledge of the absorption cross section of different molecules (e.g. NO_2) and of radiative transfer it is possible to determine the abundance of the molecule along the effective light path². In the case of NO_2 , the Ultra-Violet(UV)/Visible(VIS) spectrum is used, see Fig. 1.4. The right side of

²The light path is the path traveled from the source to the detector. For visible light this path starts at the sun, reflects off the Earth's surface or atmosphere, and then reaches a satellite in space.

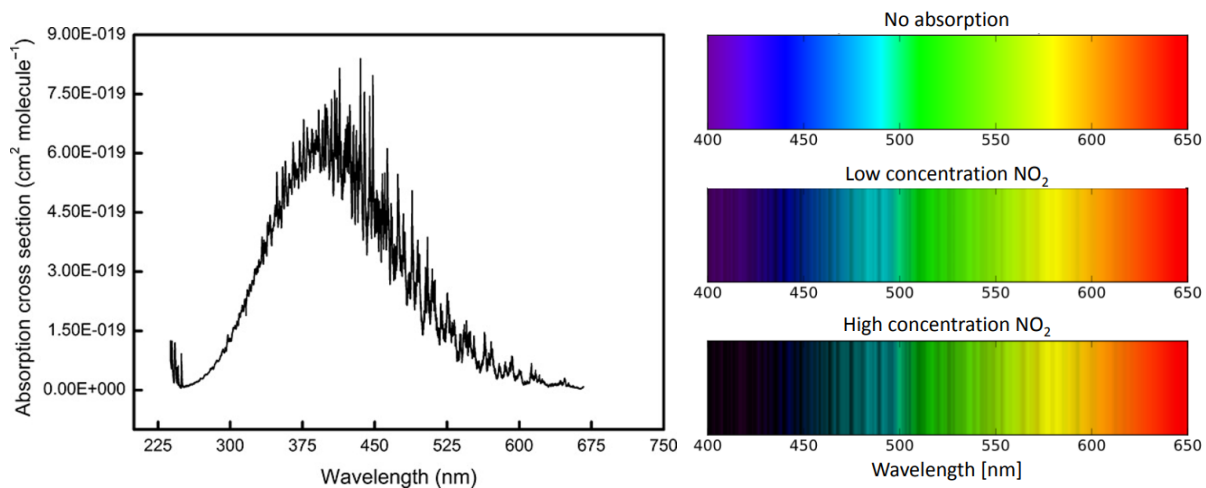


Figure 1.4: Absorption spectrum of NO₂ molecule from (Vandaele et al., 1998) (left) and schematic of satellite retrieval spectral analysis (right), courtesy of Pepijn Veefkind.

the figure is an illustration of the spectrum of light for different concentrations of NO₂ in the light path: with higher concentrations, a bar-code-like structure becomes visible, caused by the absorption lines of NO₂ shown in the left. By measuring the intensity of absorption - here visible as the 'blackness' of the barcode - in the NO₂-specific wavelengths in the Earth-reflected sunlight, the amount of NO₂ in the light path can be determined. The satellite therefore retrieves the integrated amount of NO₂ along the atmospheric light path, referred to as *column density*.

1.4.1 TROPOMI

The TROPospheric Monitoring Instrument (TROPOMI) is a state-of-the-art satellite sensor for monitoring trace gases in the atmosphere. TROPOMI is on board the Sentinel-5-Precursor (S5P) launched in October 2017 and part of the Copernicus program. It is a LEO satellite in a sun-synchronous orbit, meaning that it overpasses the equator everywhere at roughly the same local time of the day. S5P crosses the equator in an ascending fashion at around 13:30 local time at an altitude of 824 km. TROPOMI employs a push-broom design and a 2-D detector to capture back-scattered radiation from the Earth's atmosphere, accommodating viewing zenith angles up to 57°. It operates in the spectral region ranging from UV to short wave infrared. The device is outfitted with a polarization scrambler, which simplifies the radiative transfer analysis. With a swath width of approximately 2600 km, TROPOMI ensures daily coverage along the equator; at higher latitudes, the orbits overlap, typically resulting in two data retrievals per day over the Netherlands.

NO₂ is retrieved in band 4. Here, TROPOMI provides 450 measurements across-track, resulting in pixels with a minimal width of 3.5 km. Along track, the resolution of TROPOMI is 5.5 km (it was 7 km before August 2019). TROPOMI offers a significant improvement in

spatial resolution compared to its predecessor OMI: The area of the smallest TROPOMI pixel is approximately 19 km^2 , which is a factor of 17 smaller compared to its predecessor OMI.

1.4.2 NO₂ retrieval

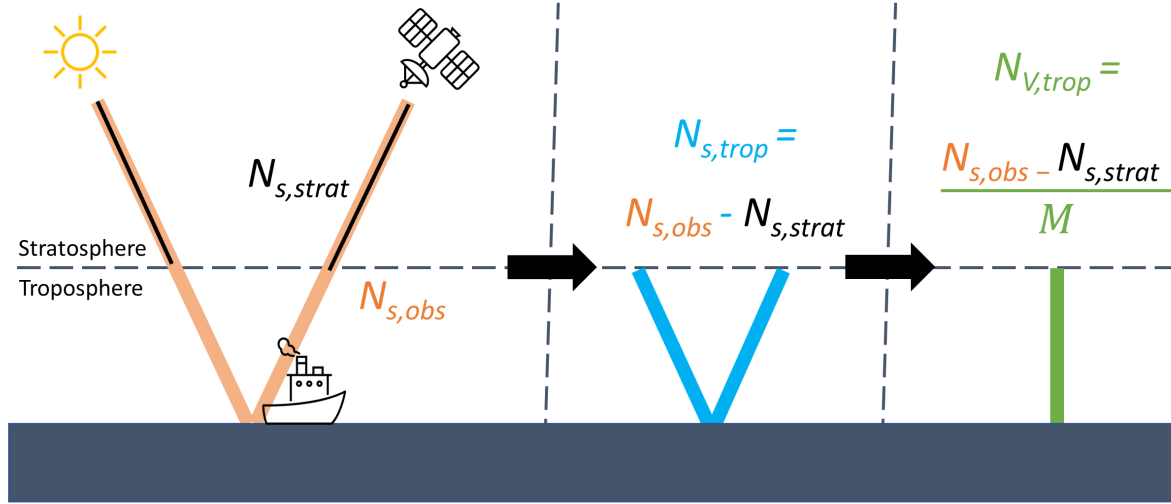


Figure 1.5: Schematic of the 3-step retrieval DOAS retrieval used in TROPOMI. The retrieval uses the spectrum of reflected and scattered sunlight that reaches the satellite instrument after passing through troposphere and stratosphere. In the first step, the slant column density $N_{s,obs}$ is retrieved. In step 2, the stratospheric contribution to the slant column $N_{s,strat}$ is subtracted from $N_{s,obs}$ to calculate the tropospheric slant column $N_{s,trop}$. In the last step, the air mass factor M is used to convert the slant column into a vertical column density $N_{v,trop}$.

To retrieve vertical NO₂ columns TROPOMI and other satellites use the Differential Optical Absorption Spectrometry (DOAS) (Platt and Stutz, 2008) in a 3-step technique, see Fig. 1.5: first the slant column density (SCD, N_s) is obtained by fitting the differential absorption cross sections to the reflectance spectrum $R(\lambda)$, the ratio of the Earth's radiance and solar irradiance spectrum of the measured Earth radiance $I(\lambda)$ to the solar irradiance spectrum $E_0(\lambda)$:

$$R(\lambda) = \frac{\pi I(\lambda)}{\mu_0 E_0(\lambda)} \quad (3)$$

in the 405-465 nm window (Van Geffen et al., 2021, 2020; Zara et al., 2018), with μ_0 being the cosine of the sun solar zenith angle. The SCD describes the amount of trace gas in the average photon path from the sun to the satellite, reflected by the Earth and its atmosphere. In the DOAS fit a polynomial function is used as high-pass filter to account for spectrally smooth scattering, absorption and surface effects. Additionally, the differential absorption cross sections of different molecules are included in the fit. In the NO₂ retrieval, part of the UV/VIS spectrum is used for the DOAS fit. For TROPOMI

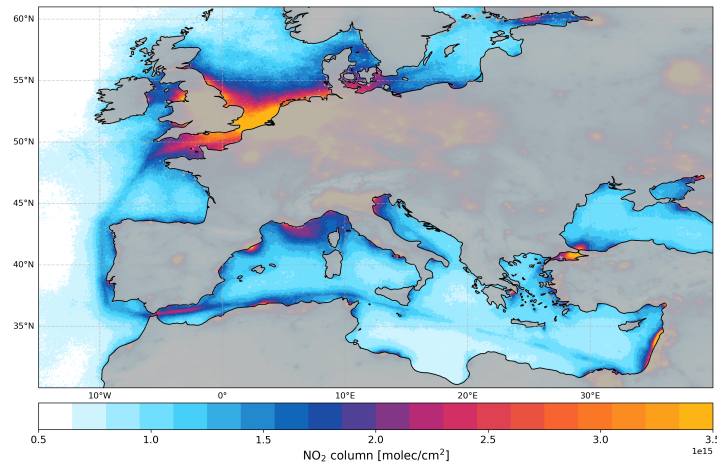


Figure 1.6: TROPOMI NO₂ columns over European seas averaged for April-November 2019. The land area is masked out.

the following function is fitted to the reflectance spectrum:

$$R(\lambda) = P(\lambda) \cdot \exp \left[- \sum_{k=1}^{N_k} \sigma_k(\lambda) \cdot N_{s,k} \right] \cdot \left(1 + C_{\text{ring}} \frac{I_{\text{ring}}(\lambda)}{E_0(\lambda)} \right) \quad (4)$$

where $\sigma_k(\lambda)$ is the absolute cross section and $N_{s,k}$ the slant column amount of molecule k , C_{ring} is describing the Ring effect for inelastic Raman scattering, and $I_{\text{ring}}(\lambda)/E_0(\lambda)$ the sun-normalized synthetic Ring spectrum. $P(\lambda)$ is the polynomial described above.

As we are mostly interested in atmospheric composition in the troposphere (the lowest atmospheric layer), we need to separate the contribution of higher atmospheric layers. For this, in the second step, data assimilation of the retrieved slant column density in the global chemistry Transport Model 5 (TM5-MP) is used to create vertical NO₂ profiles. These can facilitate the separation of the stratospheric and tropospheric contribution to the slant columns, $N_{s,trop}$ and $N_{s,stat}$, respectively (Van Geffen et al., 2020; Dirksen et al., 2011).

In the last step, Air Mass Factors (AMFs, M) are calculated (Lorente et al., 2017) to translate the N_s into vertical column densities (N_v) as $N_v = N_s/M$. The AMF accounts for the viewing and solar geometry as well as surface properties, cloud effects and scattering. The AMF is calculated using the DAK radiative transfer model (de Haan et al., 1987; Stammes, 2001). For this purpose, Look-Up-Tables (LUTs) are created where the altitude-dependent AMFs for different viewing geometries, surface albedos and surface pressures are stored. Some independent information is needed in this step: cloud height and fraction information as well as an a-priori NO₂ profiles and surface properties. Cloud properties are retrieved in separate parts of the spectrum while a-priori NO₂ profiles are simulated with a chemical transport model. We discuss the

different retrieval versions used in this thesis and the effects of clouds properties and a-priori profile on TROPOMI NO_2 data in more detail in Chapter 2. Figure 1.6 shows the average tropospheric NO_2 columns over European seas in 2019 with clearly visible shipping lanes.

1.4.3 History of satellite NO_2 and ships

There is a long history of satellite remote sensing of shipping related NO_2 , starting with GOME (Beirle et al., 2004) and SCIAMACHY (Richter et al., 2004), and followed by using OMI (Vinken et al., 2014b; Marmer et al., 2009). Those studies focussed on the study of major shipping lanes in time-averaged satellite data. Besides identification of these lanes, long term trends could be analyzed (Boersma et al., 2015). TROPOMI's higher spatial resolution and new machine learning techniques have sparked a lot of research into the (automated) detection of (ship) plumes (Kurchaba et al., 2022; Finch et al., 2022; Yuan et al., 2023) and finding outliers among them (Kurchaba et al., 2023) in the past four years. However, these studies focussed on finding plumes. Fundamental understanding of why and under which conditions plumes are visible is still lacking. As discussed in Section 1.2.1, the relation between NO_x emissions and observed NO_2 is not linear but depends on complex chemistry. Therefore, chemical models are needed to estimate the emissions that caused the observed plumes, as has been done for NO_x plumes of power plants (Meier et al., 2024; Krol et al., 2024). After having introduced NO_x and its chemistry (Section 1.2), the shipping industry as a source (Section 1.3) and the TROPOMI retrieval (Section 1.4), we will therefore end this introduction with a short Section on atmospheric (plume) modelling as the final ingredient for emission estimation from TROPOMI observations.

1.5 Modelling of NO_x plumes

In Section 1.2 we have already alluded to the difficulties in NO_x emission estimation caused by the complex, non-linear chemistry involved. To make the step from observations of vertically integrated NO_2 to the underlying NO_x emissions we need to account for the physical (e.g. plume dispersion and advection) and chemical processes that lie between the emissions and observations. Atmospheric models can help us overcome this problem by simulating these processes in the atmosphere. In this thesis we make explicit use of two such models, namely PARANOX and MicroHH, which have different advantages and use-cases as detailed below.

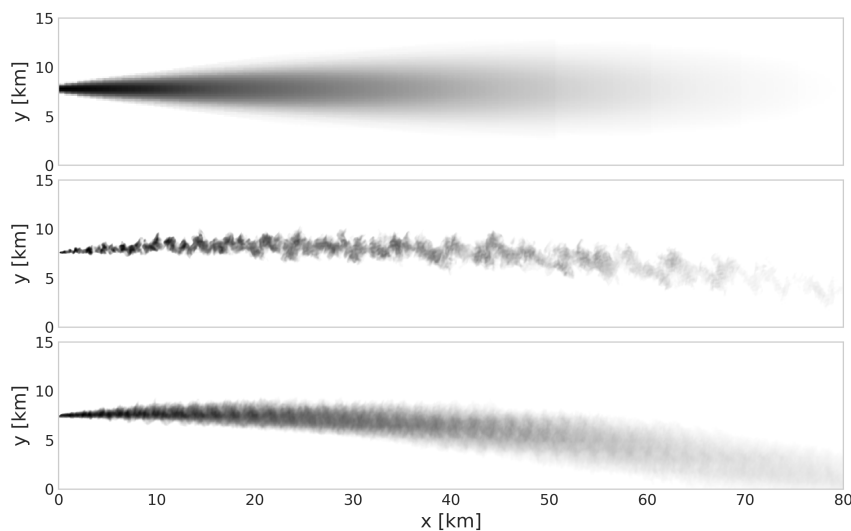


Figure 1.7: Comparison of Gaussian Plume Model (PARANOX, top) and Large Eddy Simulation (MicroHH, middle and bottom). The middle panel shows a MicroHH snapshot in time; the bottom shows the average of 7 snapshots taken 5 minutes apart. All figures display vertically integrated NO₂ on different colorscales under similar meteorological and chemical conditions. The wind direction changes over time in the MicroHH simulation, causing the plume curvature.

1.5.1 PARANOX

PARANOX (PARAmetrization of emitted NO_x) model is a Gaussian plume model (Vinken et al., 2011). In PARANOX, the plume is described as 10 concentric elliptical half-rings that expand over time. Each ring exchanges mass with the two neighboring rings. The model is used to simulate the dispersion of the plume and the chemical processes within the plume. It includes detailed O₃-NO_x chemistry, incorporating 43 species and 98 reactions (Vinken et al., 2011).

The advantage of PARANOX is its low computational cost. By design, it is not able to handle multiple sources or crossing plumes, and the parametrization of turbulent diffusion is a rather crude simplification of reality. However, it is a useful tool when looking at the dispersion of plumes in the stable marine boundary layer and interpreting observations of coarse spatial resolution such as from TROPOMI. In this thesis, PARANOX is used in Chapter 4 to simulate the NO₂-NO_x-ratio in the plumes observed over the North Sea and in Chapter 5 to create a large library of plumes for a large set of atmospheric conditions as well as in Chapter 6.

1.5.2 MicroHH

MicroHH (Van Heerwaarden et al., 2017) is a Large-Eddy-Simulation (LES) model. These models resolve part of the turbulence explicitly and are therefore a more realistic repre-

sentation of a plume than a Gaussian Plume Model. They are run at a high spatial resolution (typically 10-100 m) at which large eddies occur. Accordingly, the computational costs are high. Krol et al. (2024) implemented a simplified chemistry scheme based on the scheme used by ECMWF in their Integrated Forecasting System and Copernicus Atmosphere Monitoring System (CAMS) reanalysis (Inness et al., 2021). The simplified scheme contains 14 species in 33 reactions with an emphasis on NO_x for computational efficiency. MicroHH is a good tool to study the chemistry and dispersion of a single plume as it captures realistic gradients in concentrations, driven by chaotic turbulence. In this thesis, MicroHH is used in Chapter 5 to simulate the dispersion of a single ship plume and the chemical processes in it and serves for the validation of the PARANOX model. Figure 1.7 shows the difference in the representation of a plume in PARANOX (top) and MicroHH (middle). The LES model MicroHH resolves the turbulent structures in the plume explicitly, while the Gaussian Plume Model only parametrizes the dispersion of the plume. When averaging over time (bottom panel), the MicroHH plume becomes more Gaussian. At the resolution of TROPOMI pixels ($3.5 \times 5.5 \text{ km}^2$ at nadir), the fine spatial structures simulated by MicroHH are not resolved and the Gaussian Plume Model might therefore be a good approximation.

1.6 Thesis overview

TROPOMI offers a previously unknown level of detail in the observation of tropospheric NO_2 . With its high spatial resolution, it has made the detection of individual ship plumes possible (Georgoulias et al., 2020). This capability raises the key question addressed in this thesis: **(How) Can we estimate NO_x emissions from individual ships using satellite data?**

To answer this central question, several sub-questions need to be addressed. Firstly, we must assess the quality of the TROPOMI data over sea under various retrieval conditions and understand its uncertainties and biases caused by retrieval geometry and a-priori NO_2 profiles. Only with this understanding can we trust the data and use it for emission estimation. Furthermore, we need to understand the main meteorological and chemical drivers that influence the evolution and detectability of NO_x plumes. And lastly, we must determine how to combine all this knowledge to estimate NO_x emissions from individual ships and evaluate the effectiveness of the IMO NO_x emission regulations described above.

After a more detailed description of crucial aspects in the TROPOMI NO_2 retrieval and its versions in Chapter 2, we will address these questions in this thesis. In Chapter 3 we will look into the retrieval of NO_2 under sun glint conditions, a special retrieval geometry over water under which the first individual ship plumes observed with TROPOMI were reported. Retrievals under this geometry are considered of lower quality, but seem beneficial for detecting ship plumes. Furthermore, in the same chapter

we look into the effect of a new cloud algorithm on the retrieved NO_2 and end with a study on how shipping related NO_x emissions changed during 2020, the first year affected by COVID-19. In Chapter 4 we will focus on another important aspect for the quality of NO_2 retrievals mentioned above: the a-priori vertical profiles, which are modelled with a chemical transport model at a coarse spatial resolution. Here, we present results and implications of an aircraft campaign in the North Sea, measuring vertical NO_2 profiles. In Chapter 5 we will then use the knowledge gained in the previous chapters to estimate NO_x emissions from individual ships in the Eastern Mediterranean, combining atmospheric chemistry models with a machine learning approach, taking a big step towards automated, near-real-time and global ship emission monitoring. Finally, in Chapter 6 we will dare a view into the future of geostationary satellites and their potential for monitoring ship emissions by studying the diurnal and seasonal cycles of NO_2 plumes in the Mediterranean.

Chapter 2

Key aspects in the TROPOMI NO₂ retrieval

In this chapter the interested reader will find more technical details on the TROPOMI NO₂ retrieval. We will start with an overview of the different versions of the retrieval used in this thesis and their main differences in Section 2.1. We then discuss the effect of three key aspects in the TROPOMI NO₂ retrieval: clouds, a-priori profiles and surface albedos in Sections 2.2, 2.3 and 2.4, respectively. These have been identified as main sources of uncertainties in the NO₂ retrieval (van Geffen et al., 2021; Lorente et al., 2017) and are evaluated in Chapters 3 and 4 of this thesis.

2.1 Retrieval versions

Since the start of the TROPOMI mission, several updates to the retrieval algorithm have been made. Table 2.1 gives an overview of the different TROPOMI NO₂ retrieval versions used in this thesis. In the early versions, surface albedos for TROPOMI were based on the surface albedo climatology from OMI (Kleipool et al. (2008), 0.5° × 0.5°), and cloud properties were based on the FRESCO+ algorithm (Koelemeijer et al., 2001; Wang et al., 2008). A-priori NO₂ profiles for all standard TROPOMI versions are simulated with TM5-MP (Williams et al. (2017), 1° × 1°) and terrain height comes from Global 3 km Digital Elevation Model (DEM_3KM). Main updates in the past years were the introduction of a better cloud scheme with version 1.4 and the use of a TROPOMI-based directional Lambertian-equivalent reflectivity (directional LER, DLER, (Tilstra et al., 2017)) albedo database at a resolution of 0.125° × 0.125° in v2.4. Douros et al. (2023) presented a data version, which uses the regional CAMS data at a resolution of 0.1° × 0.1° for the calculation of AMF over Europe.

2.2 The importance of clouds

The air mass factor can be written as

$$M = \frac{\sum_l m_l n_l c_l}{\sum_l n_l} \quad (1)$$

where l is the vertical model layer, m_l the altitude-dependent air mass factor, n_l the a-priori column density and c_l a temperature dependent correction factor in that layer. The altitude-dependent air mass factor m_l is closely related to the averaging kernel elements a_l via $a_l = \frac{m_l}{M_{trop}}$ with M_{trop} being the tropospheric air mass factor (Eskes and Boersma, 2003; Boersma et al., 2016). The averaging kernel elements are a measure for the sensitivity of the satellite sensor to NO₂ in that layer. They typically decrease towards the surface as light is scattered and absorbed in the atmosphere and is especially low above dark surfaces such as water bodies, see Fig. 2.1.

Clouds have several relevant effects on satellite trace gas retrievals: they shield the lower part of the atmosphere - the part of the atmosphere were most anthropogenic emissions

Table 2.1: TROPOMI NO₂ retrieval versions used in this thesis with their main differences

	v1.1/1.3	v1.4/2.1	v2.2/2.3	v2.4	EU
Data period	30 April 2018 - 29 Nov 2020	29 Nov 2020 - 01 Jul 2021	30 April 2018 - Mid November 2021	30 April 2018 - recent	30 April 2018 - recent
Spatial resolution	3.5 km × 5.5 km (3.5 km × 7 km)	3.5 km × 5.5 km	3.5 km × 5.5 km	3.5 km × 5.5 km	3.5 km × 5.5 km
fitting window	405-465 nm	405-465 nm	405-465 nm	405-465 nm	405-465 nm
surface albedo for AMF	OMI minimum LER at 440 nm (0.5°)	OMI minimum LER at 440 nm (0.5°)	OMI minimum LER at 440 nm (0.5°), adjusted to avoid negative cloud fractions	TROPOMI DLER at 0.125degrees	OMI minimum LER at 440 nm (0.5°), adjusted to avoid negative cloud fractions
a-priori cloud retrieval	TM5-MP at 1° × 1° FRESCO+	TM5-MP at 1° × 1° FRESCO+wide	TM5-MP at 1° × 1° FRESCO+wide	TM5-MP at 1° × 1° FRESCO+wide	CAMS 0.1° × 0.1° FRESCO+wide
cloud fraction	Retrieved from 405- 465 nm continuum	Retrieved from 405- 465 nm continuum	Retrieved from 405- 465 nm continuum	Retrieved from 405- 465 nm continuum	Retrieved from 405- 465 nm continuum
cloud pressure	Narrow O2-A band (758, 761 and 765 nm)	Wide O2-A band (758, 761 nm and 765-770 nm)	Wide O2-A band (758, 761 nm and 765-770 nm)	Wide O2-A band (758, 761 nm and 765-770 nm)	Wide O2-A band (758, 761 nm and 765-770 nm)
surface albedo for cloud pressure	GOME-2 minimum LER at 758 & 772 nm (0.25° × 0.25°)	GOME-2 minimum LER at 758 & 772 nm (0.25° × 0.25°)	GOME-2 minimum LER at 758 & 772 nm (0.25° × 0.25°)	TROPOMI DLER at 0.125degrees	GOME-2 minimum LER at 758 & 772 nm (0.25° × 0.25°)
used in Chapters	3	3 & 4	4	5 & 6	4

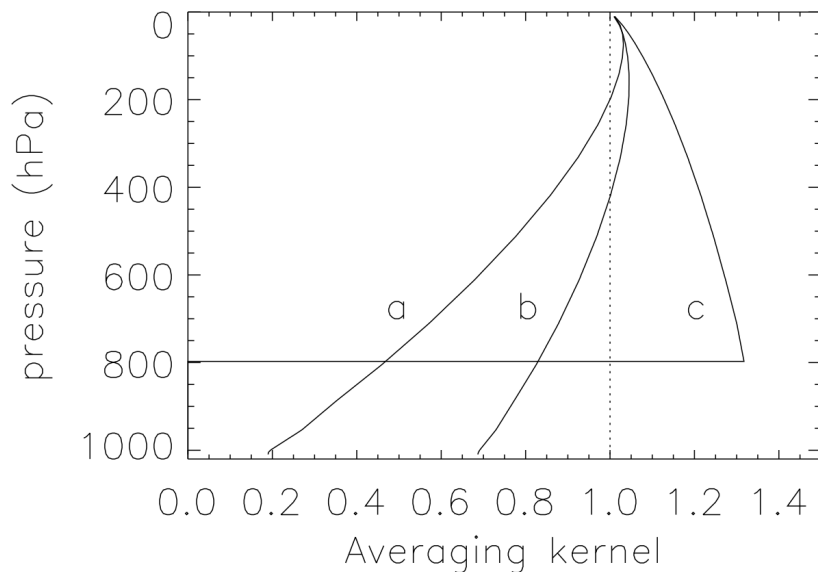


Figure 2.1: Example of averaging kernel profiles for the NO₂ retrieval: (a) clear-sky pixel with a low surface albedo, (b) clear pixel with a higher surface albedo,; (c) pixel with an optically thick cloud at 800 hPa. Taken from Eskes and Boersma (2003).

including those from shipping happen, effectively reducing m_l to zero below the cloud. For good sensitivity, users are advised to select scenes with a small cloud fraction. In TROPOMI L2 data, this is taken care of by the quality assurance (qa) value. It is advised to select scenes with $qa \geq 0.75$ only, which corresponds to an effective cloud fraction smaller than 0.2 (van Geffen et al., 2021). But even a low cloud fraction increases the scene albedo significantly, especially over dark surfaces as the deep ocean. Also, a partial cloud cover decreases the average photon path, which has to be taken into account in calculating AMFs. Partially cloudy scenes have a larger averaging kernel values above the cloud and smaller values below the cloud compared to cloud free scenes. Good knowledge of cloud fraction and height is therefore key for high quality trace gas column retrievals. If a too low cloud height is used in the retrieval, the averaging kernel for that layer is too high, leading to a too large AMF and successively too low tropospheric column densities. Cloud fraction and cloud pressure are retrieved simultaneously from the O₂-A band measurements, based on the Fast Retrieval Scheme for Clouds from the Oxygen A band (FRESCO+) (Koelemeijer et al., 2001; Wang et al., 2008) using three spectral windows at 758-759 nm (continuum, no absorption), 760-761 nm (strong absorption) and 765-766 nm (moderate absorption). In FRESCO+, the cloud is assumed to be a Lambertian reflector with a fixed albedo of 0.8, representing an optically thick cloud in line with the approach in the trace gas retrieval (Stammes et al., 2008). In the so-called independent pixel approximation, the measured reflectance R is thought to be the weighted sum of a

clear and a cloudy part:

$$\begin{aligned} R &= R_{cloud} + R_{clear} \\ &= 0.8 \cdot T_{cloud} \cdot c + A_{surface} \cdot T_{surface} \cdot (1 - c) \end{aligned} \quad (2)$$

where c is the effective cloud fraction, and T_{cloud} and $T_{surface}$ the transmittance of the atmosphere to the cloud layer and surface, respectively. $A_{surface}$ is the surface albedo. If the initial retrieval results in a cloud fraction outside the range $[0,1]$, the surface albedo is adjusted, so that the modelled and observed reflectances agree.

Likewise, the air mass factor M of a partially cloudy scene can be written as

$$M = w \cdot M_{cloud} + (1 - w) \cdot M_{clear} \quad (3)$$

where w is the radiance-weighted cloud fraction

$$w = \frac{c \cdot I_{cloud}}{c \cdot I_{cloud} + (1 - c) \cdot I_{clear}} \quad (4)$$

with I_{cloud} and I_{clear} the radiances of fully cloudy and fully clear scenes, respectively.

Previous studies have found that FRESCO+ retrieves cloud pressures close to the middle of a single layer cloud, in cases of a two layer cloud the retrieved cloud pressure is between those layers (Wang et al., 2008).

Since the version 1.4 NO₂ retrieval of the TROPOMI (November 2020), one of the spectral windows for the cloud retrieval is widened to 5 nm (765-770 nm), hereafter called FRESCO+wide, which is thought to increase the sensitivity to low clouds and solve the low bias in cloud altitudes in previous versions (Compernelle et al., 2021). A comparison to NO₂ and cloud data from OMI suggest that the bias to low clouds in early TROPOMI versions produces a systematic low bias in TROPOMI NO₂ columns of around 10% (Verhoelst et al., 2021). The FRESCO+wide retrieval is studied as part of this thesis in Chapter 3.

2.3 The importance of a-priori profiles

Equation 1 shows the importance of the a-priori profile in the retrieval, as it is one of the defining factors in the AMF calculation (van Geffen et al., 2021). For the operational TROPOMI NO₂ products, TM5 profiles at $1^\circ \times 1^\circ$ are used, which is significantly coarser than the spatial resolution of TROPOMI. This is a source of error in the retrieval, especially in regions with strong gradients in NO₂ concentration, such as close to point sources. Its emissions are smeared out in the coarse model resolution, resulting in too small subcolumns in the lowest, most polluted layers and therefore too large AMFs and too low tropospheric NO₂ columns. Likewise, tropospheric NO₂ columns will be overestimated in regions next to large NO₂ sources, as the coarse TM5 profiles will overestimate

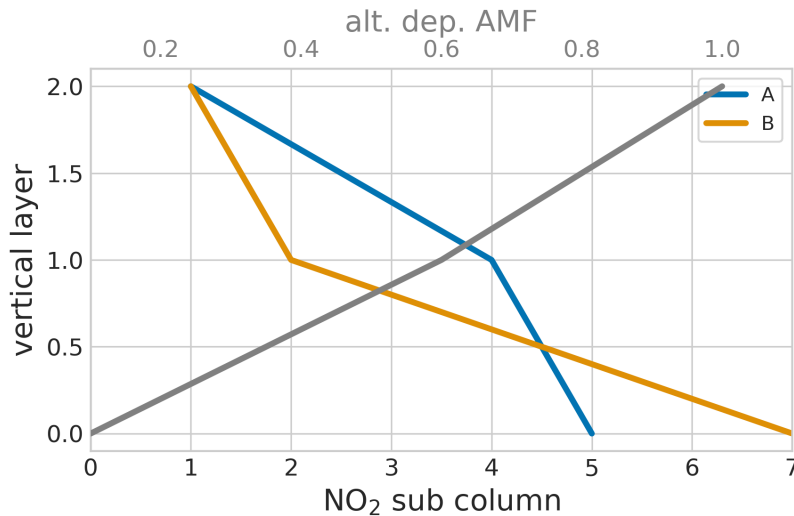


Figure 2.2: Visualization of the thought experiment on the influence of the vertical NO₂ distribution. The two colored lines show profiles A (blue) and B (orange), the gray line shows the altitude-dependent air mass factors with very low values in the lowest layer, as typical for dark surfaces such as the sea.

the NO₂ there. Douros et al. (2023) find an increase in tropospheric NO₂ columns of up to 30% close to hotspots when replacing TM5 profiles with high-resolution CAMS profiles at $0.1^\circ \times 0.1^\circ$.

Next to the effect of the coarse horizontal resolution, the correct vertical distribution of NO₂ in the a-priori profiles is important too, as can be easily seen by a simplified thought experiment: Assuming an atmosphere with 3 layers, arbitrary units, altitude-dependent air mass factor elements of $m_1 = 0.1$, $m_2 = 0.4$ and $m_3 = 1.0$, profile A with $n_1 = 5$, $n_2 = 4$ and $n_3 = 1$ and profile B with $n_1 = 7$, $n_2 = 2$ and $n_3 = 1$, visualized in Fig. 2.2. Both profiles thus contain the same amount of the trace gas. Ignoring the correction factor c_i in 1, we get the AMFs:

$$\begin{aligned}
 M_A &= \frac{0.1 \cdot 5 + 0.4 \cdot 4 + 1 \cdot 1}{5 + 4 + 1} = 0.3 \\
 M_B &= \frac{0.1 \cdot 7 + 0.4 \cdot 2 + 1 \cdot 1}{7 + 2 + 1} = 0.25
 \end{aligned} \tag{5}$$

Using $N_v = N_s/M$ we see that the vertical column in case B is 24% larger than in case A, caused by the vertical distribution of NO₂ alone (the retrieved slant column and the a-priori *column* were the same). This is a simplified example, but it shows the importance of the a-priori profile in the retrieval, which is further studied in Chapter 4 of this thesis.

2.4 The importance of surface albedos

The surface albedo is another important factor in the NO_2 column retrieval as it is used in the AMF calculation and the cloud retrieval discussed in Section 2.2. For the cloud retrieval, the FRESCO+ algorithm has originally used a surface albedo climatology derived from OMI (spatial resolution of $0.5^\circ \times 0.5^\circ$, Kleipool et al. (2008)) for the cloud fraction in the NO_2 window. This is also used in the AMF calculation. The GOME-2 minimum LER at $0.25^\circ \times 0.25^\circ$ at wavelength of 758 & 772 nm is used for the cloud pressure retrieval (Tilstra et al., 2017). As resolution of OMI (and GOME-2) and TROPOMI differ significantly, the surface albedo database is thought of as one of the largest sources for errors in cloud property retrieval of FRESCO+. Since version 2.4, a TROPOMI-derived directional LER (DLER) albedo database is used at a resolution of $0.125^\circ \times 0.125^\circ$ for the retrieval of both the cloud pressure and cloud fraction in the NO_2 window.

The surface albedo is typically low over water compared to over land. As shown in Fig. 2.1 the lower surface albedo leads to decreased averaging kernels in lower atmospheric levels, i.e. lower sensitivity of the satellite sensor to NO_2 in the lower atmosphere, making NO_2 retrievals over sea especially challenging. The figure also indicates that a too low surface albedo can lead to too low AMFs and too high NO_2 columns, and the inverse. This is relevant for the retrieval of NO_2 over the ocean, where the surface albedo is generally low but enhanced under sun glint, as studied in Chapter 3, and is generally influenced by waves, and wind speed and direction.

Chapter 3

Improved monitoring of shipping NO₂ with TROPOMI: decreasing NO_x emissions in European seas during the COVID-19 pandemic

This chapter is based on:

Riess, T. C. V. W., Boersma, K. F., Van Vliet, J., Peters, W., Sneep, M., Eskes, H., and Van Geffen, J. (2022). Improved monitoring of shipping NO₂ with TROPOMI: Decreasing NO_x emissions in European seas during the COVID-19 pandemic. *Atmospheric Measurement Techniques*, 15(5):1415–1438. Publisher: Copernicus GmbH

Abstract

TROPOMI measurements of tropospheric NO₂ columns provide powerful information on emissions of air pollution by ships on open sea. This information is potentially useful for authorities to help determine the (non-)compliance of ships with increasingly stringent NO_x emission regulations. We find that the information quality is improved further by recent upgrades in the TROPOMI cloud retrieval and an optimal data selection. We show that the superior spatial resolution of TROPOMI allows the detection of several lanes of NO₂ pollution ranging from the Aegean Sea near Greece to the Skagerrak in Scandinavia, which have not been detected with other satellite instruments before. Additionally, we demonstrate that under conditions of sun glint TROPOMI's vertical sensitivity to NO₂ in the marine boundary layer increases by up to 60%. The benefits of sun glint are most prominent under clear-sky situations when sea surface winds are low, but slightly above zero (± 2 m/s). Beyond spatial resolution and sun glint, we examine for the first time the impact of the recently improved cloud algorithm on the TROPOMI NO₂ retrieval quality, both over sea and over land. We find that the new FRESCO+wide algorithm leads to 50 hPa lower cloud pressures, correcting a known high bias, and produces $1\text{-}4\cdot 10^{15}$ molec \cdot cm⁻² higher retrieved NO₂ columns, thereby at least partially correcting for the previously reported low bias in the TROPOMI NO₂ product. By training an artificial neural network on the 4 available periods with standard and FRESCO+wide test-retrievals, we develop a historic, consistent TROPOMI NO₂ data set spanning the years 2019 and 2020. This improved data set shows stronger (35-75%) and sharper (10-35%) shipping NO₂ signals compared to co-sampled measurements from OMI. We apply our improved data set to investigate the impact of the COVID-19 pandemic on ship NO₂ pollution over European seas and find indications that NO_x emissions from ships reduced by 10-20% during the beginning of the COVID-19 pandemic in 2020. The reductions in ship NO₂ pollution start in March-April 2020, in line with changes in shipping activity inferred from Automatic Identification System (AIS) data on ship location, speed and engine.

3.1 Introduction

Emissions of nitrogen oxides ($\text{NO}_x = \text{NO} + \text{NO}_2$) have several primary and secondary effects on air quality, human health and the environment. NO_x is a toxic gas itself (WHO, 2003) and contributes to the formation of secondary pollutants and ozone. Ozone close to the Earth's surface is a toxic pollutant which can lead to respiratory problems and has negative effects on plant growth and crop yield (e.g. Wang and Mauzerall (2004)). NO_x also contributes to acid deposition and eutrophication, harming sensitive ecosystems (European Environment Agency, 2019b).

The international shipping sector is a strong source of NO_x and other air pollutants to the atmosphere (eg. Eyring et al. (2010)). Previous studies suggest that international shipping makes up for annual emissions of 2.0-10.4 TgN (Crippa et al., 2018; Eyring et al., 2010; Johansson et al., 2017), or 15-35% of total anthropogenic NO_x emissions worldwide. While cars, the power sector and industry have shown substantial reductions in their emissions over the last 10-20 years in Europe and the United States (Curier et al., 2014; Hassler et al., 2016), NO_x emissions from shipping activity have increased (De Ruyter de Wildt et al., 2012; Boersma et al., 2015) and the number of ship movements and ship size is expected to keep increasing in the future (Eyring et al., 2005; UNCTAD, 2019). Shipping-related air pollution emissions are estimated to lead to 60,000 premature deaths annually, especially in coastal regions (Corbett et al., 2007; Marais et al., 2015).

To mitigate these and other harmful impacts, more stringent regulations on NO_x emissions for ships have been implemented in coastal regions and on the open ocean (International Maritime Organization, 2013). For example, ships built in 2011 or later have to follow Tier II nitrogen emission regulation as defined in MARPOL Annex VI. In so-called Emission Control Areas (ECAs) even more stringent rules apply. From 1 January 2021 onwards, the new MARPOL Annex VI Regulation 13 determines that newly built ship engines should be compliant with Tier III in the new ECA in the Baltic and North Sea, which should result in 75% lower NO_x emissions from new ships. The exact limits depend on ship engine speed (International Maritime Organization, 2013), see Supplement 1.

For new regulations to be effective, monitoring and verification of ship emissions is required. Traditional compliance monitoring includes national authorities conducting on-board checks of engine certificates and keel-laying date. This is not a direct verification of emissions and can only be done for a limited number of vessels. Other methods are on-board measurements at the ship's exhaust pipe (e.g. Agrawal et al. (2008)) or downwind measurements of emission plumes using sniffer techniques or DOAS (Differential Optical Absorption Spectroscopy) measurements (e.g. Lack et al. (2009); Berg et al. (2012); McLaren et al. (2012); Pirjola et al. (2014); Seyler et al. (2019)). Modern techniques also include airborne platforms such as helicopters, small aircrafts (Mellqvist and Conde, 2021; Chen et al., 2005) or drones (Van Roy and Scheldeman, 2016). While these methods do not require inspectors to board the vessel, they require proximity to the ships monitored and are thus less fit-for-purpose when a large number of ships is to be checked, or on open

sea away from land. For the above reasons monitoring by satellite remote sensing offers a promising alternative.

Satellite instruments have observed enhancements of NO₂ column densities over major shipping routes, e.g. from GOME (Beirle et al., 2004), SCIAMACHY (Richter et al., 2004) and OMI (Vinken et al., 2014b; Marmer et al., 2009). These satellite measurements have recently been continued with new observations from the TROPOMI (TROPOspheric Monitoring Instrument) sensor. With a pixel size of 3.5x5.5 km² TROPOMI provides a spatially more resolved evaluation of NO₂ pollution patterns compared to its predecessors GOME (40x320 km²), SCIAMACHY (30x60 km²) and OMI (13x24 km²). Indeed, previous studies demonstrated TROPOMI’s capability to pinpoint emissions from the mining industry (Griffin et al., 2019), emissions patterns within cities (Beirle et al., 2019; Lorente et al., 2019), emissions along a gas pipeline in Siberia (van der A et al., 2020) and even from individual ships in the Mediterranean Sea (Georgoulas et al., 2020). Ding et al. (2020) used TROPOMI NO₂ columns and inverse modelling to show NO_x emission reductions during the COVID-19 lockdown over urban centers and regions with strong maritime transport.

While the aforementioned studies demonstrate the large potential of TROPOMI and its high resolution, retrieval problems remain. Validation studies (e.g. Griffin et al. (2019); Verhoelst et al. (2021)) suggest a 15%-40% low bias in TROPOMI tropospheric vertical NO₂ ($N_{v,trop}$) columns relative to independent in-situ and MAX-DOAS measurements. Cloud properties present one of the leading sources of uncertainty in trace gas retrieval from space (Boersma et al., 2004; Lorente et al., 2017) and cloud heights used until (and including) v1.3 of the operational TROPOMI retrieval algorithm have been suggested to be biased low (Compernelle et al., 2021). To address this bias in cloud heights, the Royal Dutch Meteorological Institute (KNMI) recently updated the FRESCO+ cloud retrieval by widening the spectral window, which is supposed to improve the sensitivity to low clouds.

The here presented study presents and assesses the impact of steps towards an improved monitoring of shipping NO₂ with TROPOMI. First, we demonstrate TROPOMI’s capability to detect ship emissions applying a typical data selection and compare it to OMI’s. We examine previous suggestions of improved retrieval sensitivity over sun glint scenes (Georgoulas et al., 2020). Additionally, we evaluate the new FRESCO+wide cloud pressure retrieval in and its impact on the TROPOMI NO₂ columns in v1.4/2.1 of the operational TROPOMI NO₂ algorithm. Based on our findings, we create a data set of historical TROPOMI NO₂ columns consistent with the v1.4 data allowing for otherwise challenging trend analysis. We conclude with an application of our findings to quantify the effects of the COVID-19 pandemic on ship pollution, an unique opportunity to assess the relationship between the anticipated emission reductions and observed NO₂ columns.

3.2 Methods and Materials

3.2.1 TROPOMI and OMI NO₂ column measurements

The European TROPOMI (Veefkind et al., 2012) is on board the Sentinel-5-Precursor launched in October 2017. TROPOMI has a push-broom design with a 2-D detector, which measures back-scattered radiation from the Earth's atmosphere for viewing zenith angles up to 57°, in the spectral region from UV to short wave infrared. The instrument is equipped with a polarization scrambler, simplifying the radiative transfer analysis. The width of the TROPOMI swath is about 2600 km, which results in daily (near-)global coverage with about 25 million measurement points. In band 4, where NO₂ is retrieved, TROPOMI provides 450 measurements across-track, with a minimal width of 3.5 km.

The design of OMI is similar to that of TROPOMI, but OMI measures in a smaller spectral range (270-500 nm) (Levelt et al., 2006, 2018). Another important difference is that OMI has only 60 across-track measurements, with the smallest pixels having a width of 25 km. Along track, the resolution of TROPOMI is 7 km (5.5 km since August 2019), compared to 13 km for OMI. Combined, the area of the smallest TROPOMI pixel is 19 km², while it is 325 km² for OMI, a factor of 17 improvement in spatial resolution. Both instruments are in a sun-synchronous ascending orbit and have an equator overpass time of about 13:30 hrs local time.

To retrieve tropospheric NO₂ columns, TROPOMI uses a 3-step retrieval approach based on the DOAS (Differential Optical Absorption Spectroscopy, Platt and Stutz (2008)) technique: first the slant column density (N_s) is retrieved by spectral fitting of a modeled reflectance spectrum to the observed reflectances in the 405-465 nm window (van Geffen et al., 2021; Van Geffen et al., 2020; Zara et al., 2018). In the second step, data assimilation in the global chemistry Transport Model 5 (TM5-MP) results in vertical NO₂ profiles that are then used to separate the stratospheric and tropospheric contribution to the slant columns (Van Geffen et al., 2020; Dirksen et al., 2011). In the last step, Air Mass Factors (AMFs) are calculated (Lorente et al., 2017) to translate the N_s into vertical column densities (N_v). The AMF is calculated using the DAK radiative transfer model (de Haan et al., 1987; Stammes, 2001), and accounts for the viewing and solar geometry as well as surface properties and cloud effects. Cloud height information is retrieved with TROPOMI's FRESCO+ cloud algorithm (driven by the 761 and 765 nm O₂ absorption depth), and cloud fraction from the reflectance levels within the 405-465 nm NO₂ fitting window. Other input parameters to the TROPOMI AMF calculation are the surface albedo climatology (Kleipool et al. (2008), 0.5°x0.5°), a priori NO₂ profiles simulated with TM5-MP (Williams et al. (2017), 1°x1°) and terrain height from Global 3 km Digital Elevation Model (DEM_3KM).

The retrieval of tropospheric NO₂ columns ($N_{v,trop}$) from OMI (Boersma et al., 2018) proceeds along the same lines, and is therefore similar in many aspects. On the other hand, especially spatial resolution, signal-to-noise and the retrieval of cloud properties

differ as highlighted in Table 3.1.

Clouds have several relevant effects on NO₂ retrieval. Clouds shield the lower part of the atmosphere which is most influenced by anthropogenic emissions including those from shipping. Therefore, data users are typically advised to consider scenes with cloud radiance fractions below 50% (Eskes et al., 2019). Initial validation of TROPOMI NO₂ v1.2/1.3 pointed out that the FRESCO+ algorithm retrieves cloud heights close to the surface heights, leading to overestimations in the TROPOMI NO₂ AMFs, and, consequently underestimations of the tropospheric NO₂ columns (Verhoelst et al., 2021). Accurate knowledge of cloud fraction and height is key for high quality trace gas column retrievals (e.g. Boersma et al. (2004); Van Geffen et al. (2021)). A detailed description of the TROPOMI and OMI cloud algorithms and recent updates therein is given in the following subsection.

Table 3.1: Retrieval settings for the TROPOMI and OMI NO₂ retrievals used in this work.

		TROPOMI v1.2/1.3	TROPOMI v1.4/2.1 ¹	OMI QA4ECV
Data availability		ESA science hub	ESA science hub	qa4ecv.eu
Public data period		30 Apr 2018 – 29 Nov 2020	29 Nov 2020 - 01 Jul 2021	Oct 2004 -
Spatial resolution at nadir		3.5 km × 5.5 km (3.5 km × 7 km)	3.5 km × 5.5 km	13 km × 25 km
N_s	Fitting window	405-465 nm	405-465 nm	405-465 nm
	Signal-to-noise ratio	~1500	~1500	~500
Solar reference spectrum		Daily	Daily	2005 average
DOAS polynomial degree		5	5	4
Intensity offset correction		no	no	yes
Destriping		yes (since v1.2)	yes	yes
AMF	Surface albedo	OMI minimum LER at 440 nm (0.5°)	OMI minimum LER at 440 nm (0.5°)	OMI minimum LER at 440 nm (0.5°)
	A priori NO ₂ profiles	TM5-MP at 1° × 1°	TM5-MP at 1° × 1°	TM5-MP at 1° × 1°
Cloud retrieval		FRESCO+	FRESCO+wide	OMCLDO2
Cloud fraction		Retrieved from 405-465 nm continuum	Retrieved from 405-465 nm continuum	Retrieved from 470-490 nm continuum
Cloud pressure		Narrow O ₂ -A band (758, 761 and 765 nm)	Wide O ₂ -A band (758, 761 nm and 765-770 nm)	O ₂ -O ₂ absorption feature (477 nm)
Surface albedo in cloud pressure retrieval		GOME-2 minimum LER at 758 & 772 nm (0.25° × 0.25°)	GOME-2 minimum LER at 758 & 772 nm (0.25° × 0.25°)	OMI minimum LER at 758 & 772 nm (0.5° × 0.5°)

3.2.2 Improved TROPOMI FRESCO+, OMI and VIIRS cloud retrievals

FRESCO+ (Fast Retrieval Scheme for Clouds from the Oxygen A band) retrieves cloud pressures from the relative depth of O₂-A band measurements (Koelemeijer et al., 2001; Wang et al., 2008) using three spectral windows at 758-759 nm (continuum, no absorption), 760-761 nm (strong absorption) and 765-766 nm (moderate absorption). In the algorithm, clouds are assumed to be Lambertian reflectors with a fixed albedo of 0.8,

consistent with assumptions for the NO₂ AMF calculation. The surface albedo assumed in the cloud pressure retrieval is from the GOME-2 minimum LER climatology at 758 & 772 nm (Tilstra et al., 2017), which is a potential source of uncertainty in the cloud pressure retrieval as the resolution and overpass time of GOME-2 is different from TROPOMI. FRESCO+ has been compared to other cloud data sets by Compernelle et al. (2021), who reported on tendencies in FRESCO+ to overestimate cloud pressures.

To address the high-bias in TROPOMI FRESCO+ cloud pressures, a new version of the FRESCO+ algorithm was introduced and implemented in the operational NO₂ retrieval with the introduction of TROPOMI v1.4 in December 2020. This version, called FRESCO+wide, uses a wider spectral window for the cloud retrieval (765-770 nm, see Table 3.1), which includes the flank of the absorption band, where oxygen absorption is weaker than in the center of the O₂-A band (761 and 765 nm). Adding weaker O₂ absorption features improves the sensitivity to clouds low in the atmosphere. This is not possible from the strong O₂ absorption at 761 nm, which is so close to saturation that it becomes difficult to use its absorption depth in order to distinguish between bright reflecting layers at the Earth's surface and reflecting surfaces in the lower atmosphere.

Prior to the implementation of FRESCO+wide in the operational TROPOMI NO₂ retrieval in December 2020, KNMI produced 4 periods with TROPOMI NO₂ test data based on FRESCO+wide, the so-called diagnostic data set 2B (DDS-2B). DDS-2B contains data from four v1.2/v1.3 periods during 2018-2019 additionally processed with v2.1 of the TROPOMI algorithm. The most significant difference between the two is that v2.1 (and v1.4) uses cloud fractions and AMFs determined from the FRESCO+wide cloud pressure instead of the FRESCO+ cloud pressures used in v1.2/v1.3 data.

Additionally, we use co-sampled cloud information from the Ozone Monitoring Instrument (OMI) on board of EOS-Aura. The OMI OMCLDO2-retrieval uses the relative depth of the O₂-O₂ absorption feature at 477 nm to retrieve cloud pressures (Acarreta et al., 2004; Veeffkind et al., 2016). The general approach of using Lambertian reflectors is similar to the FRESCO+ algorithm, but an important difference is that the OMCLDO2-algorithm needs to account for Raman scattering and O₃ absorption, and that the absorption strength of the O₂-O₂ features is proportional to the square of the O₂ concentration, making it more sensitive to low clouds compared to FRESCO+.

We also use cloud information from VIIRS (Visible Infrared Imager Radiometer Suite) on board of the SUOMI National Polar-orbiting Partnership (SNPP) as a completely independent means of verification. SNPP orbits the Earth in a sun-synchronous, ascending node with full daily global coverage and observes the same scenes as TROPOMI within three minutes. We use NASA's CLDPROP L2 VIIRS-SNPP cloud product (Platnick

¹In addition to improved cloud parameters, TROPOMI v2.1 data has improved further through a better calibration of level-1 spectra, especially in the treatment of outliers and saturation (Ludewig et al., 2020), and through improvements in the NO₂ algorithm itself (Van Geffen et al., 2021). Version v2.1 is only used for production of the DDS-2B test data, not for publicly released data. Version v2.2, available publicly as of July 2021, is essentially the same as v2.1.

et al., 2017) with a resolution of 750 m at nadir, which provides a cloud mask, cloud (top) pressure, cloud optical thickness (COT) and cloud water phase for each pixel retrieved. The VIIRS retrieval derives a cloud top temperature using an optimal estimation approach in the thermal infrared spectral bands M14-M16 (8.5-12.3 μm). In a subsequent step, these cloud top temperatures are converted to cloud pressures using Numerical Weather Prediction temperature profiles (Heidinger and Li, 2017). In addition to the cloud top pressure, we use the VIIRS cloud optical thickness (COT) to generate (effective) cloud fractions that can be compared directly to the TROPOMI cloud fractions. First, we derive a geometrical cloud fraction by calculating the share of cloudy VIIRS pixels per grid cell. Then, we translate this geometrical cloud fraction $f_{c,geo}$ into a effective cloud fraction $f_{c,eff}$ using:

$$f_{c,eff} = f_{c,geo} * a_c / 0.8 \quad (1)$$

with a_c the cloud albedo. The cloud albedo is calculated from the VIIRS COT and a previously established empirical relationship between cloud optical thickness and cloud albedo for liquid water clouds (Buriel, 2005; Boersma et al., 2016)².

To evaluate the improvements in the FRESCO+wide retrieval, we compare daily gridded, co-sampled cloud data from (partly) cloudy pixels seen by TROPOMI (FRESCO+ and FRESCO+wide), OMI and VIIRS over parts of the Mediterranean Sea (37.0°N-41.25°N, 2.0°W-8.0°W), the Bay of Biscay (43.5°N-47.5°N, 10.0°E-3.0°E) and Northwestern Europe (50.0°N-53.0°N, 4.0°W-9.0°W). These regions represent different surface types (land and ocean), climatological conditions and pollution levels. We define partly cloudy pixels as all pixels with an effective cloud fraction $f_c \geq 0.05$. For TROPOMI we additionally apply sufficient quality of retrieval ($qa \geq 0.5$) and a pressure difference between surface pressure and cloud pressure of at least 7 hPa. The last filter is applied to filter out 'ghost' clouds coming from sun glint viewing geometries (see Sec. 3.2.3 below). For OMI, we use the OMCLDO2 cloud properties and take only pixels with solar and viewing zenith angle smaller than 80° into account. As Eq. 1 is valid for liquid water clouds only, we select liquid water clouds, and reject ice clouds, as indicated by the VIIRS cloud water phase. Around 25-30% of VIIRS pixels are missed due to this filter.

3.2.3 Sun glint in the TROPOMI NO₂ retrieval

The term *sun glint* refers to particular satellite viewing geometries, under which the ocean acts as a mirror by reflecting sun light directly to the satellite instrument. In

²The cloud albedo a_c for liquid water clouds can be expressed as a 6th order polynomial of the VIIRS cloud optical thickness (τ) as $a_c = b_0 + b_1 * \tau + b_2 * \tau^2 + b_3 * \tau^3 + b_4 * \tau^4 + b_5 * \tau^5 + b_6 * \tau^6$ with the coefficients $b_0 = 0.0153$, $b_1 = 0.0967$, $b_2 = -0.00605$, $b_3 = 0.000212$, $b_4 = -0.00000405$, $b_5 = 0.0000000392$, and $b_6 = -0.000000000150$.

the TROPOMI data product pixels that are potentially in sun glint mode are identified based on the combination of their solar and viewing zenith and azimuth angles. The sun glint condition is fulfilled when the scattering angle Θ is smaller than a threshold angle Θ_{\max} :

$$\Theta = \arccos [\cos \theta \cos \theta_0 - \sin \theta \sin \theta_0 \cos (\phi_0 - \phi)] \leq \Theta_{\max} \quad (2)$$

with θ and θ_0 the solar and viewing zenith angles and ϕ and ϕ_0 the solar and viewing azimuth angles, respectively (see Supplementary Figure S1). For the TROPOMI data products the maximum threshold angle has been set at 30° . Smaller angles are used before, e.g. for SCIAMACHY and GOME-2 (Loots et al., 2017). The TROPOMI algorithm treats the enhanced albedo as a partially cloudy scene with the cloud pressure located at or close to the sea surface.

3.2.4 Relationship between NO_x emissions and columns

When studying NO_2 columns to investigate emission trends, the non-linearity of NO_x chemistry needs to be taken into account. For example, the lifetime of NO_x depends on the background O_3 level, the available sun light and NO_x concentrations themselves (Jacob, 1999). We use a (modeled) β factor to express the sensitivity of relative NO_2 column changes to changes in the relative emission strength following the approach in Vinken et al. (2014a) with

$$\beta = \frac{\Delta E/E}{\Delta N/N} \quad (3)$$

where $\Delta E/E$ represents the imposed relative change in NO_x emission flux and $\Delta N/N$ the relative change in subsequently simulated tropospheric NO_2 columns. Here we use beta values from (Vinken et al., 2014b) modeled with GEOS-Chem at $0.5^\circ \cdot 0.67^\circ$ and accounting for plume-in-grid chemistry. These beta values have a similar spatial resolution as the spatially averaged TROPOMI NO_2 signals from ships (see Fig. 3.9). As we are interested in European Seas only, we average β in the area 35°N - 40°N and 5°E - 10°W for Gibraltar and 30°N - 37°N and 15°W - 35°W for the Eastern Mediterranean. We use the resulting β value to estimate relative changes in NO_x emissions $(E_{2020} - E_{2019})/E_{2019}$ as

$$\frac{E_{2020} - E_{2019}}{E_{2019}} = \beta \cdot \frac{N_{obs,2020} - N_{obs,2019}}{N_{obs,2019}} \quad (4)$$

where $(N_{obs,2020} - N_{obs,2019})/N_{obs,2019}$ is the observed relative change in NO_2 columns.

3.2.5 AIS data and ship specific data

To relate the TROPOMI NO_2 columns to shipping activity, we use data from the Automatic Identification System (AIS) for shipping. Since 2005, the International Maritime Organization (IMO) requires all ships with a gross tonnage over 300 and all passenger ships to carry an AIS transponder. These transponders broadcast static (e.g. identity,

size) and dynamic (e.g. position, speed, course) information of the ship, which can be received by other ships, shore stations, and satellites (International Maritime Organization, 2014). Here we use historical AIS data available to the Dutch Human Environment and Transport Inspectorate (ILT) to assess changes in shipping activity over densely travelled European shipping lanes in 2019 and 2020. We use AIS data of ships in a part of the shipping lane in the Eastern Mediterranean (31.91°N-34.53°N and 25.91°E-27.67°E) and close to the Strait of Gibraltar (35.0°N-37.0°N and 4.0°W-2.5°W). Furthermore, we use information on ship dimensions from the official ship registrations (gisis.imo.org) to calculate a ship emission proxy E from ship length L and ship speed v as $E \propto L^2 \cdot v^3$ as used e.g. in Georgoulias et al. (2020). For the areas and times under study, ship specific data was available only for 50% (Gibraltar) and 70% (Eastern Mediterranean) of the ships.

3.3 Results

We start with demonstrating TROPOMI’s capabilities to detect shipping NO₂ applying established data selection criteria. Next, we show steps to optimize monitoring of ship emissions making use of sun glint (Sec. 3.3.2) and recent improvements in the cloud retrieval (Sec. 3.3.3) and compare the improved TROPOMI data to OMI data in Sec. 3.3.4. We end with an application of our findings to quantify NO₂ emissions reductions from shipping due to COVID in 2020 in Sec. 3.3.5.

3.3.1 Detection of NO₂ pollution over European shipping lanes

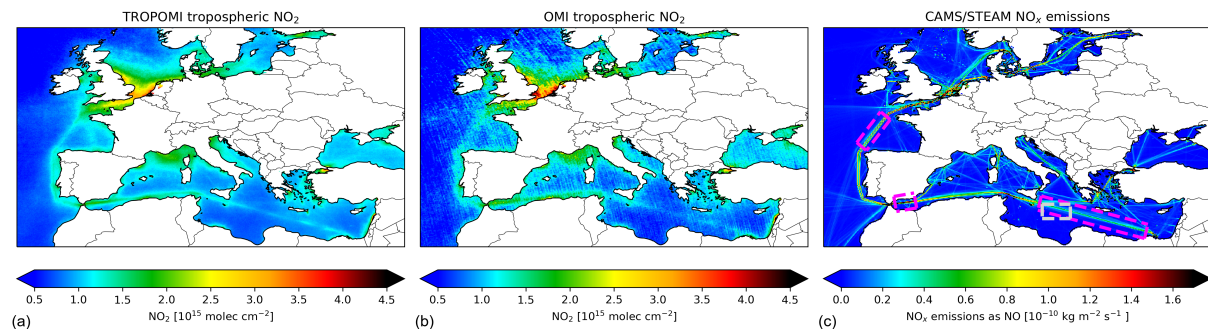


Figure 3.1: Summertime (May-September) mean tropospheric NO₂ columns from TROPOMI (a) and OMI (b) over European seas in 2019. The right panel shows the summertime mean NO_x emissions from the CAMS/STEAM emission inventory (Granier et al., 2019; Johansson et al., 2017). The gray and pink rectangles in the center panel indicate areas used in Sec. 3.3.2 and Sec. 3.3.4, respectively.

TROPOMI detects unprecedented spatial detail in shipping NO₂ over busy shipping routes. Fig. 3.1 shows the summertime mean (May-September 2019) NO₂ columns from TROPOMI and OMI averaged to a common 0.0625°x0.0625° grid as well as CAMS/STEAM NO₂ emissions (Granier et al., 2019; Johansson et al., 2017) for the

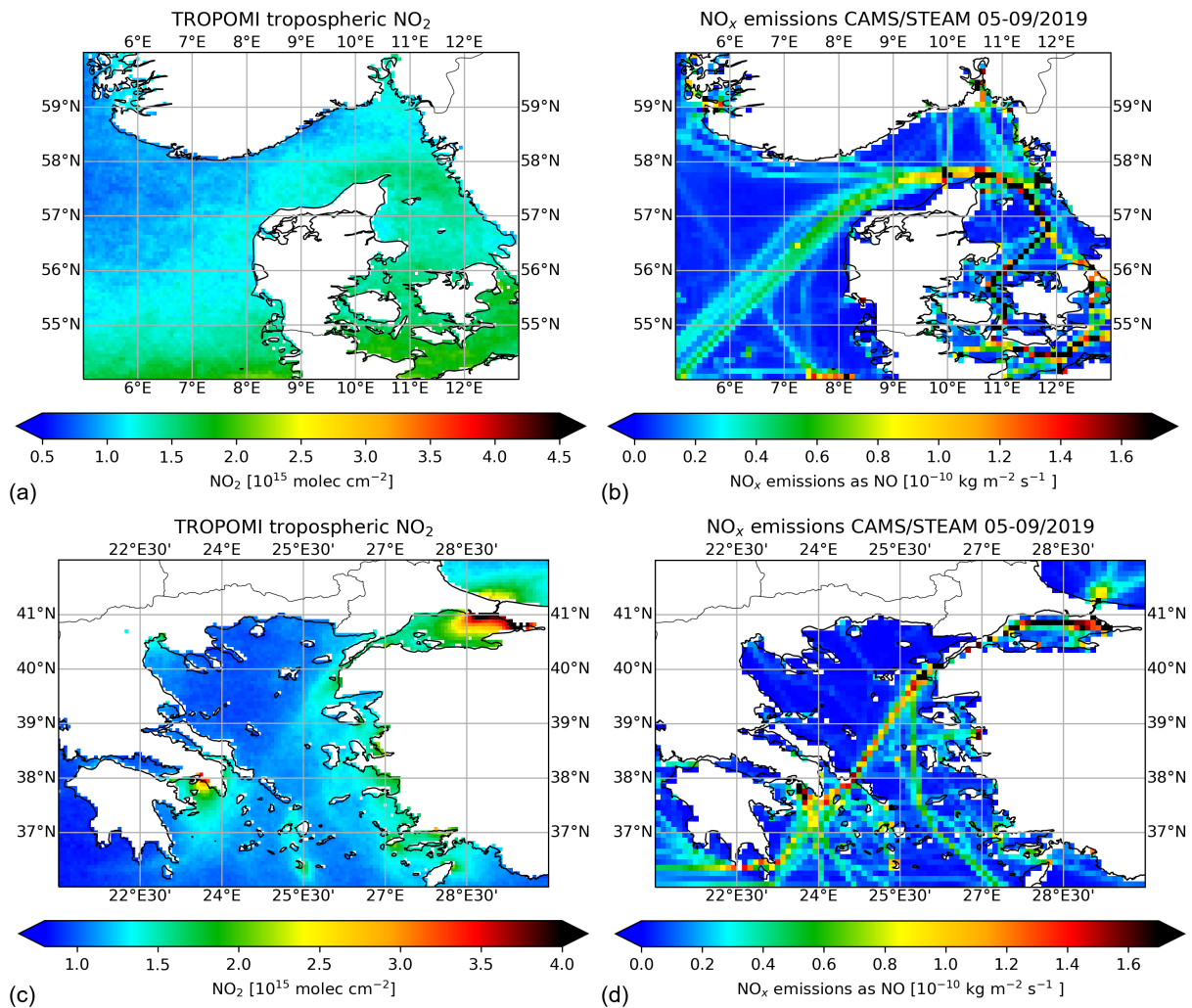


Figure 3.2: 2019 summertime mean (May-September) tropospheric NO_2 columns from TROPOMI (a & c) and summertime mean NO_x emissions from the CAMS/STEAM emission inventory (b & d, Granier et al., 2019; Johansson et al., 2017) of shipping lanes around Denmark (a & b) and Eastern Aegean Sea (c & d) for the first time detected with satellites.

same period. We find a clear signal of shipping NO_2 in the TROPOMI data west of Portugal and from the Strait of Gibraltar to the East. There are further indications of enhanced NO_2 related to shipping in the Bay of Biscay from the tip of Brittany towards the North-West of Spain, and in the Eastern Mediterranean from South of Sicily towards the Suez Canal. Previous studies reported NO_2 enhancements over these shipping lanes with other satellites (e.g. by OMI (Vinken et al., 2014b)). Additionally, we see a clear NO_2 enhancement in the Aegean Sea between Istanbul and the Greek Islands as well as around Denmark as shown in Fig. 3.2, which to our knowledge have not been observed by satellite instruments previously. Furthermore, (clear) hints of shipping activity can be seen in the Baltic Sea, the Eastern Aegean Sea, the Adrian Sea, north-east of Corsica, the British Channel, and several forks in the Eastern Mediterranean and south east of

Sicily, which are all present as shipping lanes in CAMS/STEAM emissions. Corresponding zoomed in maps of TROPOMI tropospheric NO₂ and CAMS/STEAM emissions are shown in the Appendix 3.A.1. For the analysis, we selected mostly clear-sky pixels with a quality assurance value (qa) of 0.75 or higher as recommended in the TROPOMI (Eskes et al., 2019) and (equivalent settings) OMI user manuals (Boersma et al., 2017). These enhancements are not an artefact in the retrieval coming from the AMF calculation (see Section 3.2.1) as they are visible in tropospheric vertical column densities $N_{trop,geo}$ using a geometric AMF (see Appendix 3.A.2) shown in Fig. 3.A.2.

TROPOMI and OMI show a comparable high spatial correlation to CAMS/STEAM emission data of $R = 0.93$ and $R = 0.91$, respectively. For the calculation, we brought OMI and TROPOMI tropospheric data to the CAMS resolution ($0.1^\circ \times 0.1^\circ$) and selected only grid cells over the Mediterranean Sea. This was done to ensure comparable meteorological and chemical conditions. Next, we binned the data by emission strength in bins of $0.05 \cdot 10^{-10} \text{ kg m}^{-2} \text{ s}^{-1}$. A reduced major axis regression of all bins with more than 10 entries lead to the correlation coefficients given above. Corresponding scatter plots can be found in Fig. 3.A.3. The y-axis intercept of $1.07 (1.05) \cdot 10^{15} \text{ molec.cm}^{-2}$ for TROPOMI (OMI) represents the mean background NO₂ column over the summertime Mediterranean. Other emission bin sizes lead to slightly different but comparable regression results.

Besides the higher resolution of the TROPOMI instrument, TROPOMI $N_{v,trop}$ thus have a comparable spatial correlation with emission inventories when compared to OMI's. The distinct shipping lanes visible in Fig. 3.1 and 3.A.2 visualize TROPOMI's unprecedented capabilities to detect shipping NO₂.

3.3.2 Sun glint

For situations of sun glint (see Sec. 3.2.3) the usually dark ocean appears bright in the TROPOMI data, leading to a strong increase in the effective scene albedo with decreasing scattering angle as shown in Fig. 3.3(a). Figure 3.3(b) shows that the increase in scene albedo leads to substantially higher vertical sensitivities, as diagnosed by the averaging kernels (AK) in the operational TROPOMI NO₂ product. The sensitivity increased most in the lowest vertical layer, where the kernel values are on average $\approx 60\%$ higher for sun glint compared to non sun glint circumstances (0.44 vs 0.28). Increased albedo generally enhances a satellite sensor's sensitivity to NO₂ concentrations in the lower atmosphere (e.g. Eskes and Boersma (2003)), and sun glint scenes have been tentatively used previously to attribute shipping plumes to individual ships in the Mediterranean Sea (Georgoulas et al., 2020).

The scene albedo and vertical sensitivity can be further increased by focusing on scenes with low-moderate wind speeds ($\approx 2 \text{ m/s}$) as wind-induced waves are expected to change the reflectivity. Fig. 3.4(a) shows the relationship between effective scene albedo and wind speed for scenes with small scattering angles $\Theta \leq 15^\circ$. For very low wind speeds the mean scene albedo is almost as small as for non sun glint scenes and smaller than for all

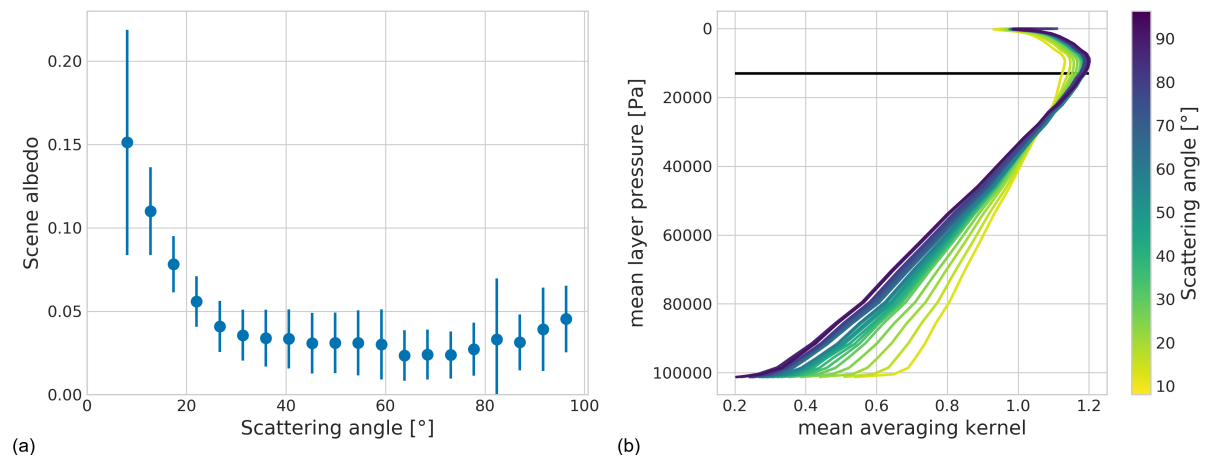


Figure 3.3: (a) Change of effective scene albedo with scattering angle over the Central Mediterranean north of Libya in June-July-August 2018 (see gray rectangle in Fig. 3.1(c), $\approx 200,000$ data points in total). The error bars indicate the standard deviation of each bin. (b) Mean averaging kernel (profiles) for different scattering angles, sampled as for (a). Only pixels with cloud radiance fractions < 0.25 or $p_{surf} - p_{cloud} \leq 300$ Pa were selected. The black line in (b) indicates the average tropopause altitude.

other wind speeds. For wind speeds between 1.5 and 2.0 m/s we find an effective scene albedo of almost 0.25, which is approximately double compared to the average for these scattering angles and more than 5 times as high as for non sun glint scenes. For higher wind speeds the scene albedo decreases to around 0.10. In Fig. 3.4(b) the effect on the averaging kernel profile is shown. As expected low wind speeds lead to the smallest AK in the lower atmosphere, whereas wind speeds between 1.5 and 2.0 m/s show the largest AKs close to the sea surface. This relationship can be understood in terms of wind-induced sea surface roughness (Cox and Munk, 1956). Both very low and strong winds limit the probability that a scattering angle $\Theta \leq \Theta_{max}$ leads to sun glint effects at the sensor: For very low wind speeds, the sea surface is effectively flat, leading to sun glint only for very small scattering angles $\Theta \ll \Theta_{max}$, whereas for strong winds the sea surface is so rough that the sun light is reflected in all directions, making the reflections towards the satellite instrument unlikely.

Additionally, we find that sun glint scenes can be used with confidence for detecting ship pollution signals from UV/Vis spectrometers such as TROPOMI and the usage of sun glint data should be encouraged. The (normalized) tropospheric slant columns ($N_{trop,geo} = N_{s,trop}/M_{geo}$, see Appendix B) observed under sun glint conditions are 20-25% higher than under non sun glint conditions as shown in Fig. 3.5(a). Vertical profiles of NO_2 over oceans typically feature enhancements from ships within the marine boundary layer, and small background levels above (e.g. Chen et al. (2005); Boersma et al. (2008), see Fig. 3.1). Therefore, it is no surprise that the AK increases in the lower atmosphere lead to small but detectable increases in (tropospheric) slant columns over the study re-

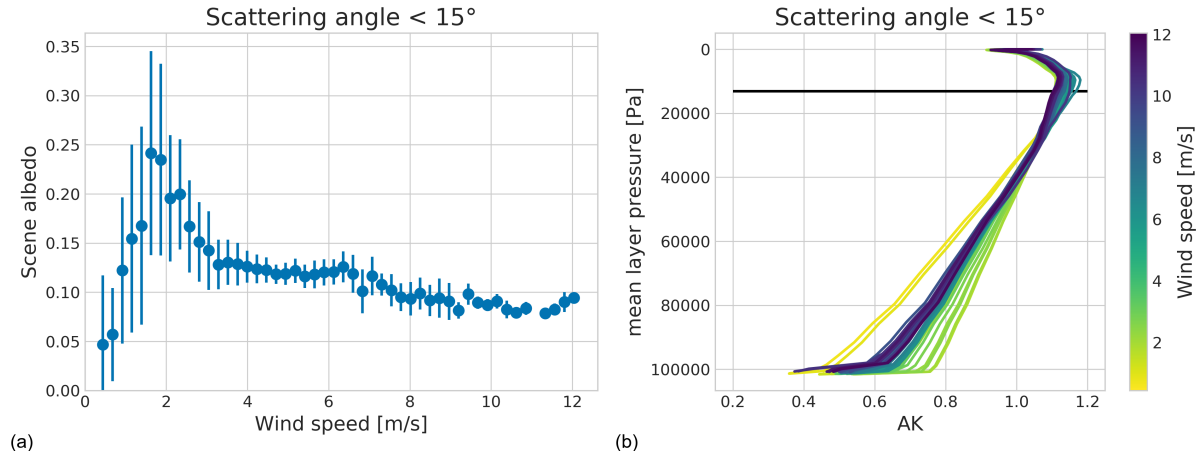


Figure 3.4: (a) Change of effective scene albedo with wind speed over the Central Mediterranean north of Libya in June-July-August 2018 (see gray rectangle in Fig. 3.1(c)) for scenes with scattering angles smaller than 15° ($\approx 22,000$ data points in total). The error bars indicate the standard deviation of each bin. (b) Mean averaging kernel (profiles) for different wind speeds, sampled as for (a). Only pixels with cloud radiance fractions < 0.25 or $p_{surf} - p_{cloud} \leq 300$ Pa were selected. The horizontal black line in (b) indicates the average tropopause altitude.

Table 3.2: Summary of mean effective scene albedo, normalized tropospheric slant column and $N_{v,trop}$ for non sun glint, sun glint ($\Theta \leq 30^\circ$) and super sun glint ($\Theta \leq 20^\circ$) over the Central Mediterranean north of Libya in June-July-August 2018 (see gray rectangle in Fig. 3.1(c)).

	non sun glint	sun glint	super sun glint
Effective scene albedo	0.03 ± 0.02	0.08 ± 0.05	0.11 ± 0.05
$N_{trop,geo}$ (molec·cm ⁻²)	$(0.65 \pm 0.28) \cdot 10^{15}$	$(0.80 \pm 0.30) \cdot 10^{15}$	$(0.83 \pm 0.30) \cdot 10^{15}$
$N_{v,trop}$ (molec·cm ⁻²)	$(1.06 \pm 0.42) \cdot 10^{15}$	$(1.05 \pm 0.40) \cdot 10^{15}$	$(1.06 \pm 0.38) \cdot 10^{15}$

gion covering a frequently travelled shipping lane.

The enhanced slant columns are correctly accounted for by increased AKs leading to reliable retrievals under sun glint. Fig. 3.5(b) compares the tropospheric vertical columns reported in the official TROPOMI NO₂ product sampled under sun glint compared to non sun glint conditions. The differences between the distributions are only small. Mean values for scene albedo, (normalized) tropospheric slant columns and tropospheric vertical columns reported in the official TROPOMI NO₂ product for different scattering angles are summarized in Table 3.2.

3.3.3 Cloud properties

Here we evaluate TROPOMI's capability to retrieve realistic cloud parameters retrieved from the 405-465 nm continuum reflectances and effective cloud pressures from the O₂-A band (Table 3.1), addressing recent improvements in the FRESCO+ algorithm to avoid

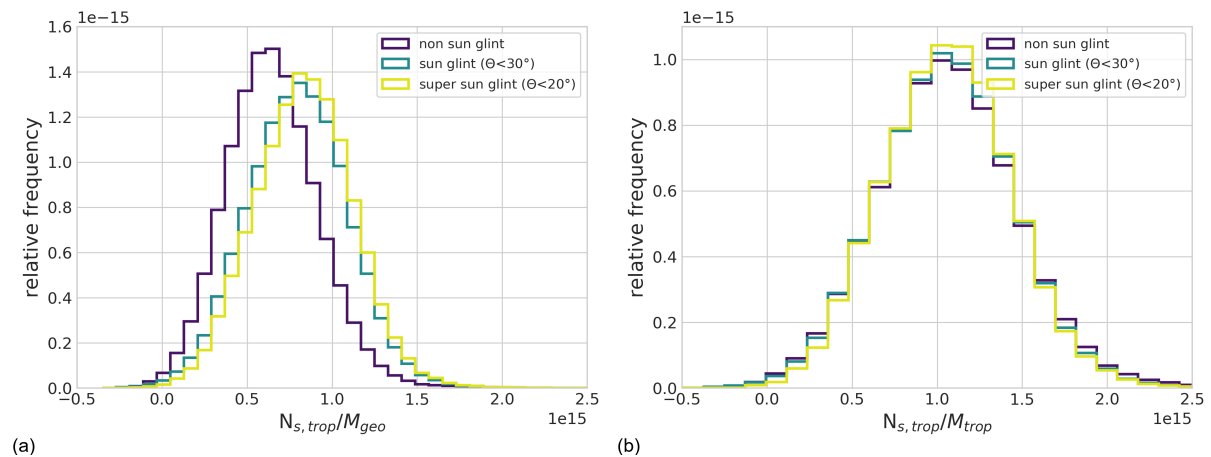


Figure 3.5: (a) Probability distribution of tropospheric NO₂ columns ($N_{s,trop}/M_{geo}$) over the Central Mediterranean north of Libya in June-July-August 2018 (see grey rectangle in Fig. 3.1(c)) taken under non sun glint, sun glint ($\Theta \leq 30^\circ$) and super sun glint ($\Theta \leq 20^\circ$). (b) Probability distribution of tropospheric NO₂ columns in the official TROPOMI NO₂ product ($N_{s,trop}/M_{trop}$) for the same data selection.

overestimated cloud pressures (Compernolle et al., 2021). These improvements in cloud retrievals lead to an inconsistency in the tropospheric NO₂ column record.

3.3.3.1 Cloud fractions

We find that improved TROPOMI cloud fractions are of sufficient quality to support the TROPOMI NO₂ AMF calculation. They show good correlation to independent data such as from OMI and VIIRS. TROPOMI v1.2 and v2.1 cloud fractions are very similar with the new v2.1 cloud fractions being slightly smaller. More details can be found in Appendix 3.A.4.

3.3.3.2 Cloud pressure

FRESCO+wide cloud pressures are a clear improvement over the FRESCO+ data used in v1.2/1.3. Figure 3.6 shows a comparison of gridded, co-sampled cloud pressure distributions from TROPOMI v1.2 (FRESCO+), TROPOMI v2.1 (FRESCO+wide), OMI QA4ECV and VIIRS over the Bay of Biscay between 1 and 7 July 2018. As expected, the improved TROPOMI v2.1 cloud pressures are ≈ 40 hPa lower than for v1.2, in line with their enhanced sensitivity, and show more realistic, elevated clouds. It is apparent that OMI cloud pressures are generally lower and show a flatter distribution than the other products. TROPOMI v2.1 and v1.2 show similar distributions as VIIRS, with v1.2 pressures higher by 50 hPa in the median, and v2.1 moving closer to VIIRS with a difference of 2 hPa relative to VIIRS. We find similar agreement between TROPOMI and independent data over the Mediterranean Sea and northwestern Europe as shown in Ta-

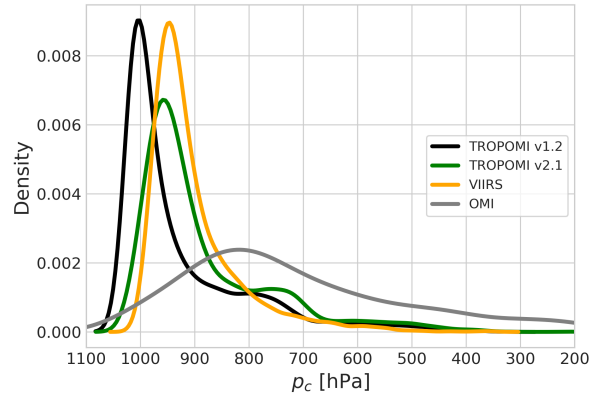


Figure 3.6: Probability distribution function of effective cloud pressures from TROPOMI v1.2, TROPOMI v2.1, OMI, and VIIRS for 1-6 July 2018 over the Bay of Biscay. Only cloud pressures for cloud fractions between 0.05 and 0.20 were selected, as these are most relevant for AMF calculations for mostly clear-sky pixels.

Table 3.3: Evaluation of TROPOMI v2.1 cloud pressures against reference data for the Bay of Biscay.

	Median cloud pressure [hPa]	10th/90th percentile [hPa]	Geometric mean [hPa]
TROPOMI v1.2	979	753/1017	925
TROPOMI v2.1	930	717/988	884
OMI QA4ECV	769	426/934	720
VIIRS	928	794/973	901

ble 3.3. FRESCO+wide cloud pressures agree best but remain higher than VIIRS in the median (both FRESCO cloud pressure distributions show a larger tail towards low pressures compared to VIIRS, possibly caused by filtering for liquid water clouds in VIIRS). This is in line with expectations as VIIRS’s infrared cloud retrieval is mostly sensitive to the cloud top (Platnick et al., 2017), whereas FRESCO’s O₂-A band retrieval is more sensitive to the center of a cloud (e.g. Sneep et al. (2008)). Around 25-30% of VIIRS cloud retrievals in the areas studied here are ice water clouds and therefore not included in the analysis. As these clouds appear at higher altitudes, improved cloud pressures have only little influence on the NO₂ columns (see Sec. 3.3.3.3).

3.3.3.3 Effect of improved cloud pressure on TROPOMI NO₂ columns

The improved cloud pressures lead to increases of NO₂ columns of up to 40% depending on area and season. The left panel of Fig. 3.7 shows the change in tropospheric NO₂ columns as a function of cloud pressure over the Bay of Biscay and northwestern Europe in Summer. We see that NO₂ columns increase most for locations that had the highest

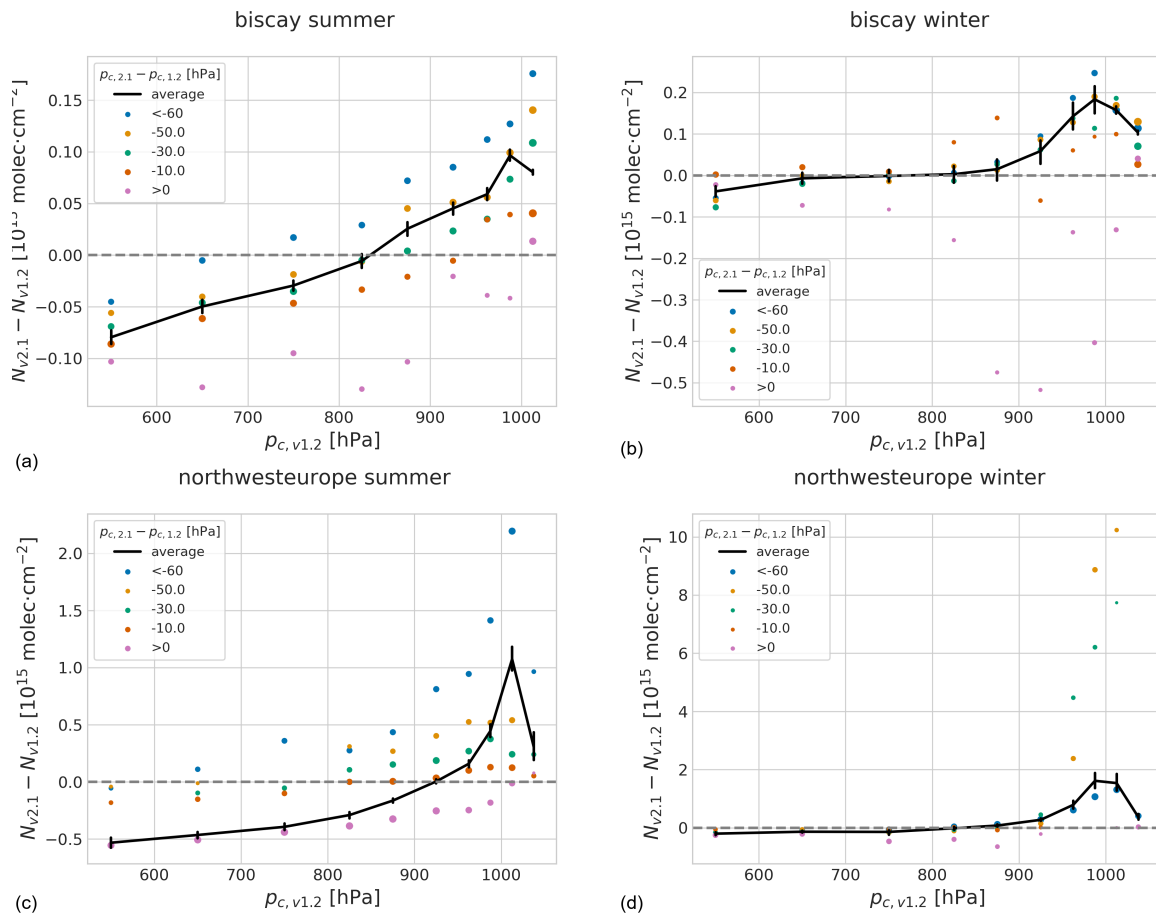


Figure 3.7: Difference between tropospheric NO_2 columns retrieved with TROPOMI v1.2 and v2.1 as a function of the v1.2 cloud pressure for 27 June – 6 July 2018 over the mildly polluted Bay of Biscay in summer (a) and winter (b) as well as and the highly polluted Northwest Europe (c & d). Colors indicate the difference in cloud pressures between the two versions. The marker size is proportional to the logarithm of the sample size and the black line shows the effective average.

original v1.2 cloud pressures, and that the improvements are strongest when cloud pressures are reduced most (light blue dots). The increase over the Bay of Biscay is smaller (up to $0.1 \cdot 10^{15} \text{ molec-cm}^{-2}$) than over northwestern Europe (up to $1.0 \cdot 10^{15} \text{ molec-cm}^{-2}$), reflecting the higher pollution levels over the mainland. We see similar patterns with stronger improvements in Winter, as shown in the right panel of Fig. 3.7. The increased v2.1 NO_2 columns indicate that the v1.2 TROPOMI NO_2 product suffers from a ‘cloud shielding’ effect: NO_2 columns are underestimated due to too low clouds situated within the polluted boundary layer and that improved v2.1 cloud pressures (at least partly) resolve the low bias in v1.2 NO_2 columns. For this analysis, we compared the TROPOMI v2.1 columns retrieved with improved cloud information, to the TROPOMI v1.2 NO_2 columns. We used 10 days in 4 different seasons (27 June – 6 July 2018, 28 December 2018 – 5 January 2019, 25 March – 5 April 2019, and 13 - 23 September 2019)

for which both v2.1 test data and v1.2 operational data were available to us as part of the DDS-2B. Our comparison focused on mostly clear-sky situations ($f_c < 0.2$), which are most relevant for detection of near-surface pollution sources.

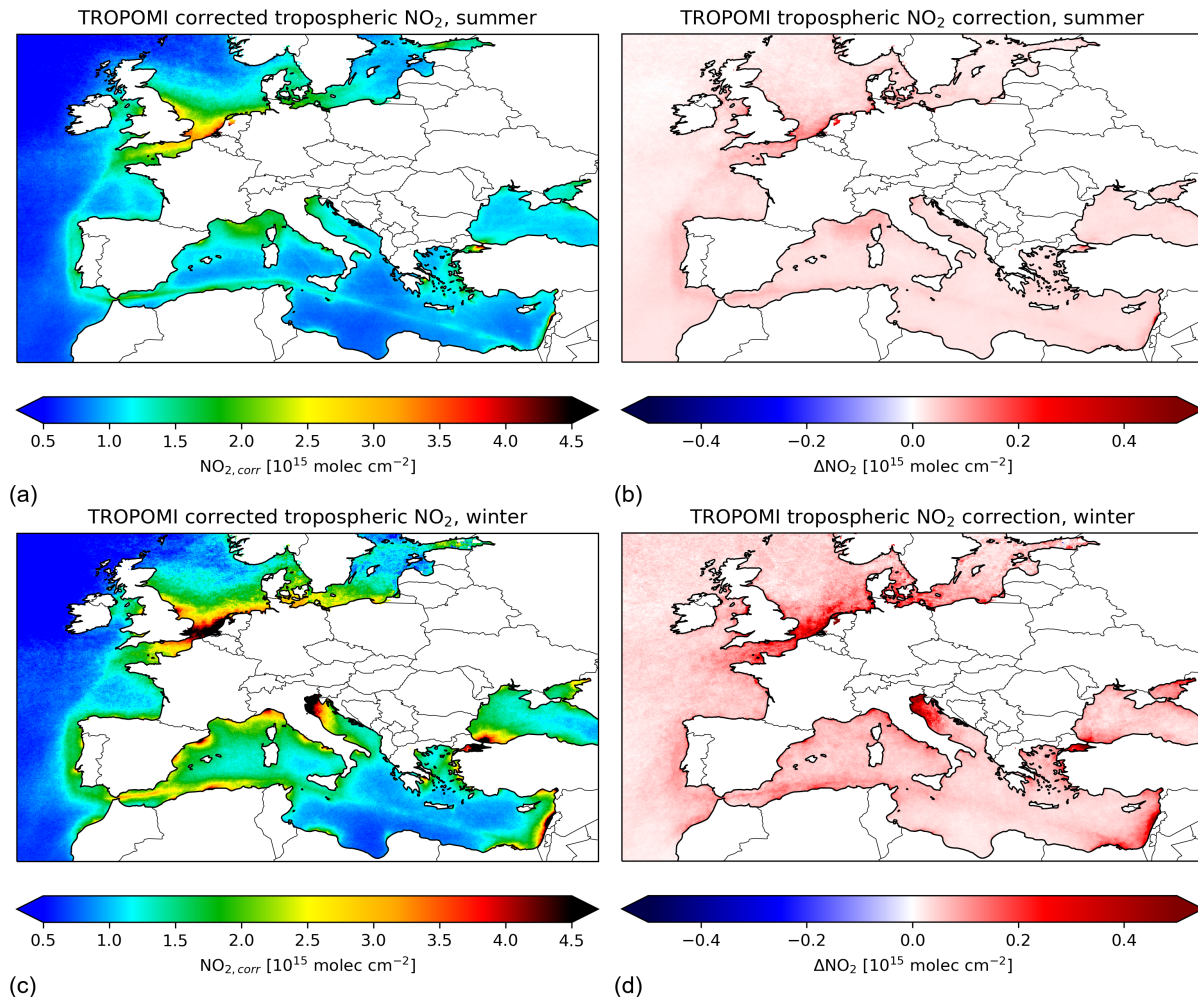


Figure 3.8: Effect of DNN correction: (a) Corrected TROPOMI data for summer (May–September) 2019, (b) change in NO₂ columns by the correction for the same period (v2.1p–v1.3), (c) Corrected TROPOMI data for winter (November–April) 2019/2020, (d) change in NO₂ columns by the correction for the same period (v2.1p–v1.3). Land areas are whitened out for clarity.

We trained a Deep Neural Network (DNN) to predict v2.1 columns for the full TROPOMI mission period up to December 2020 and thereby created a consistent data set. The DNN-predicted v2.1 (hereafter v2.1p) reduces the mean difference to the retrieved v2.1 NO₂ columns to $< 0.01 \cdot 10^{15}$ molec·cm⁻² (original v2.1 – v1.2 mean difference was $0.12 \cdot 10^{15}$ molec·cm⁻²) over the 3 areas (see Sec. 3.2.2) of study during the 4 periods, suggesting considerable skill in the DNN approach. Details can be found in Appendix 3.A.6.

Figure 3.8 shows the averaged NO₂ columns from v2.1p over the Summer of 2019 and

Winter of 2019/20. The difference map in the right panel indicates that predicted v2.1 NO_2 columns are higher by up to $0.5 \cdot 10^{15} \text{ molec. cm}^{-2}$, especially over the most polluted seas such as the English Channel and shipping lanes. We find a stronger impact of the improved cloud pressures in the winter season, reflecting that NO_2 pollution is confined in a thinner marine boundary layer in that season.

3.3.4 Comparison of TROPOMI and OMI NO_2 columns in shipping lanes

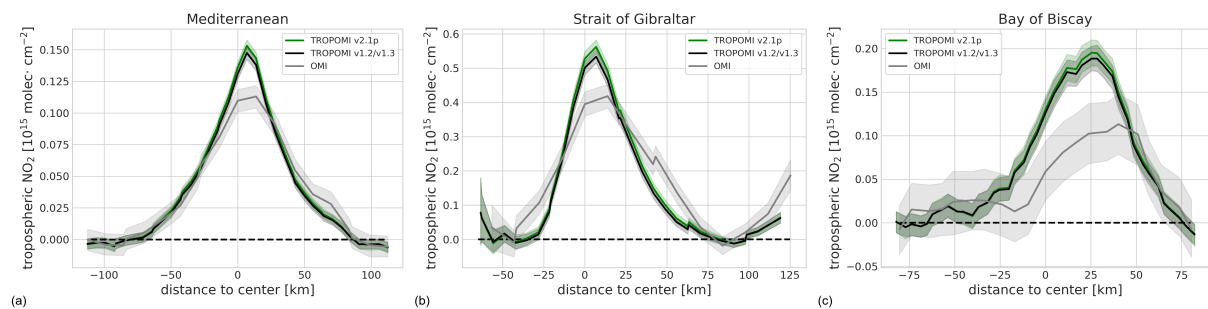


Figure 3.9: Mean enhancement cross sections in June-August 2019. TROPOMI v1.2/v1.3 in black, the improved TROPOMI v2.1p in green and OMI in grey. Shaded areas indicate the 95% confidence interval.

TROPOMI detects a more pronounced and narrower region of ship NO_2 pollution than OMI. On average, TROPOMI v2.1p detects 45% higher peak NO_2 values than OMI. TROPOMI data allow the attribution of 14% more NO_2 to shipping lane enhancements, over 23% narrower shipping lanes. To quantitatively compare TROPOMI's capability to detect NO_2 over shipping lanes under different measurements conditions and compare it to OMI's, we created average NO_2 cross sections over busy shipping lanes. We studied NO_2 enhancements in summer 2019 (June-August) over shipping lanes in the Bay of Biscay, from Sicily to the Suez Canal, and East of Gibraltar, the regions visually defined in Fig. 3.1(c). First, we defined the location of the shipping lanes according to the emission data shown in Fig. 3.1(c). Then, we calculated the average NO_2 columns along the shipping lane and parallel to it, taking care to exclude NO_2 columns measured over land. In that way we created an average cross section of NO_2 over shipping lanes. In the last step, we performed a background correction by subtracting a linear NO_2 background to isolate the NO_2 enhancements caused by shipping. The orbital data was gridded to regular grids of $0.0625^\circ \times 0.0625^\circ$ and $0.125^\circ \times 0.125^\circ$ resolution for TROPOMI and OMI, respectively. For TROPOMI only pixels with $qa > 0.75$ were taken into account. For OMI, a consistent filtering was applied, including maximal solar and viewing zenith angles of 80° and maximal cloud radiance fractions of 0.5. The resulting cross sections are shown in Fig. 3.9. Table 3.4 summarizes the peak value, the area under the curve (i.e. the total NO_2 attributed to shipping) and the full width at half maximum (FWHM) for the three

Table 3.4: Statistics for the NO₂ enhancement cross section over the Mediterranean shipping lane.

Shipping lane	Product	max [10 ¹⁵ molec/cm ²]	area under curve [10 ¹⁵ km·molec/cm ²]	FWHM [km]
Bay of Biscay	TROPOMI v1.2/v1.3	0.189	11.34	54.6
	TROPOMI v2.1p	0.195	11.72	54.6
	OMI	0.113	7.75	59.2
Strait of Gibraltar	TROPOMI v1.2/v1.3	0.534	24.47	41.3
	TROPOMI v2.1p	0.562	26.18	41.7
	OMI	0.418	28.83	63.8
Mediterranean	TROPOMI v1.2/v1.3	0.148	8.80	52.3
	TROPOMI v2.1p	0.153	9.19	52.6
	OMI	0.113	8.56	70.2

shipping lanes. It should be noted that the grid used for OMI is 2x coarser than the one used for TROPOMI. Gridding TROPOMI to the coarser grid used for OMI only changes the results slightly, indicating that the improved spatial resolution of TROPOMI indeed improves the detection of NO₂ from narrow ship lanes and is in line with the finding of new shipping lanes shown in Fig. 3.2.

As already seen in Fig. 3.8, the v2.1p data set shows slightly higher NO₂ compared to the TROPOMI v1.2/v1.3 data, especially in the center of the lane while background NO₂ is less affected by the correction. The impact of the DNN is larger in winter than in summer as discussed before.

For the Bay of Biscay it is also apparent that the NO₂ peak is shifted to the East for all data sets. As the location is defined by an emission inventory based on AIS data (and therefore real ship location), this is likely an effect of dominant westerly winds.

We conclude that TROPOMI provides a significant improvement for the detection of shipping NO₂ with sharper and more pronounced shipping lanes in seasonal averages. The improved v2.1p TROPOMI data increase the signal further.

3.3.5 Reductions of ship NO_x emissions during the COVID-19 pandemic

Emissions proxies derived from AIS data and from TROPOMI NO₂ suggest emission reductions from shipping in 2020 compared to 2019 as depicted in Fig. 3.10 (c & f). While in the first three months of 2020 the ship emissions were generally higher compared to 2019, both emission proxies show reductions starting in April and lasting until the end of the year. This reduction can be linked to the COVID-19 pandemic, which led to economic lockdowns in many countries of the world. Europe had its most stringent measures in Spring and Autumn 2020.

We created daily 0.0625·0.0625°-maps of TROPOMI data, using v2.1p NO₂ columns as

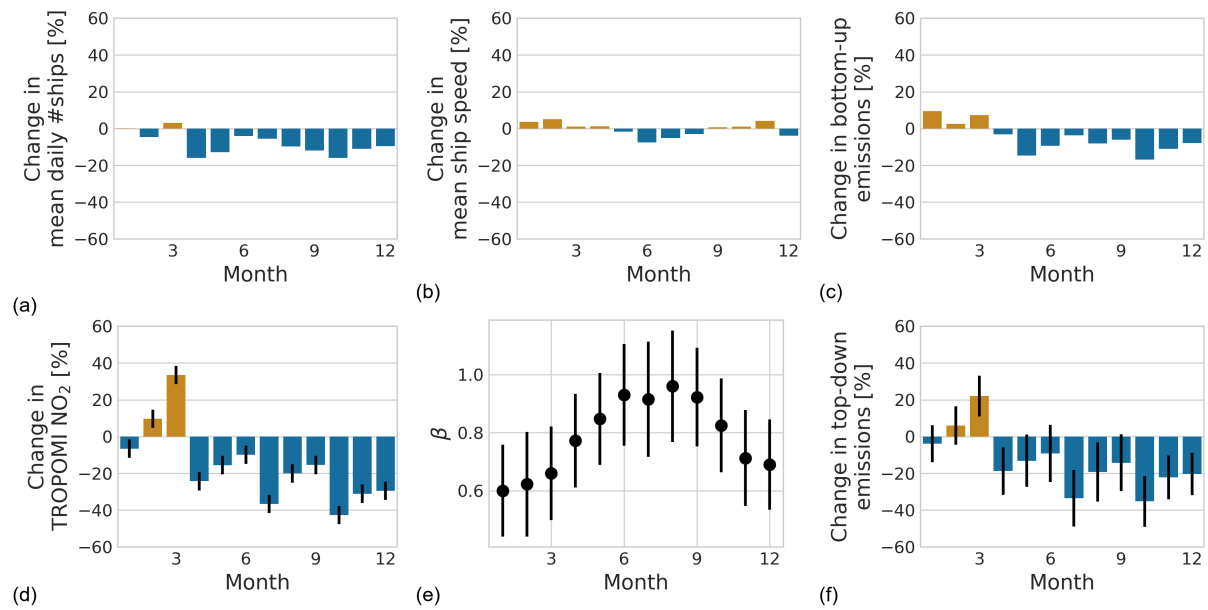


Figure 3.10: (a) Relative change in monthly mean of daily number of ships passing the Strait of Gibraltar between 2019 and 2020 inferred from AIS data. (b) Same but with average ship speed. (c) Relative change in emission proxy ($v^3 \cdot L^2$). (d) Relative change in TROPOMI shipping NO₂. (e) Monthly β values from Vinken et al. (2014b) with month-to-month variability imposed by monthly values from Verstraeten et al. (2015). Uncertainty intervals represent the temporal and spatial variability of β (see discussion in the text). (f) Relative change in top-down emissions from shipping. Error bars represent the propagated uncertainties in TROPOMI shipping NO₂, beta and differences in meteorological conditions between 2019 and 2020 (see discussion in the text).

described in Section 3.3.3.3 with $qa \geq 0.75$. We calculate the area under the cross section as a measure for shipping NO₂ for monthly mean NO₂ columns for the shipping lanes of Gibraltar and Mediterranean defined in Fig. 3.1(c). Monthly TROPOMI shipping NO₂ for 2019 and 2020 can be seen in Supplementary Figure S2(c). Figure 3.10(d) shows the relative change in shipping NO₂ from 2019 to 2020 in the Strait of Gibraltar. Using β values and the approach described Section 3.2.4 and shown in Fig. 3.10(e), we arrive at the TROPOMI based relative change in emission changes shown in Fig. 3.10(f).

The uncertainty in our top-down NO_x emission changes follows from : (i) The sensitivity of TROPOMI shipping NO₂ to the area of study ($\sigma_{area} = 5\%$), (ii) the inter-year differences on monthly averaged NO₂ columns over the areas of study caused by meteorology and (iii) the combined spatial and temporal spread of β caused by differences in the chemical regime caused e.g. by differences in atmospheric composition and radiation ($\sigma_{\beta} = 0.15$). Fig. 3.10(d) and Fig. 3.A.5(d) show σ_{area} , in panels (e) uncertainties (ii) and (iii) are used while for panels (e) a full error propagation of all uncertainties listed above was performed. A full discussion on the uncertainty estimates can be found in Supplement 4.

Additionally, we used AIS data to calculate an AIS based emission proxy as described in Sec. 3.2.5. We filtered for days with TROPOMI coverage of at least 50% of each study area. AIS data indicates that the number of ships passing per month through the Strait of Gibraltar has reduced from March 2020 onwards relative to 2019 (Fig. 3.10(a) and Supplementary Figure S2(a)). The average speed of the ships passing through the shipping lanes is lower between May-September 2020 compared to the same period in 2019 as well (Fig. 3.10(b) and Supplementary Figure S2(b)). This is in agreement with a study by (Millefiori et al., 2021) who found an increase in container ship speed in May and June in 2019 which is absent in 2020 leading to a relative decrease. Finally, Fig. 3.10(c) shows the relative change in AIS deduced emission proxy from 2019 to 2020. Similar results for the shipping lane in the Mediterranean can be found in Fig. 3.A.5 and Supplementary Fig. S3.

Several studies report changes in ship activity in 2020 using AIS data. Additional to the 5% decrease in ship speed in the Mediterranean between March and April 2020 compared to 2019 mentioned above, Millefiori et al. (2021) reported global mobility of container ships to have decreased by 10% between March and June 2020 compared to the previous year. March et al. (2020) find increases in traffic density for January and February 2020 with decreases in March-June, with Western Europe showing very strong reductions. Both studies show strong variations by vessel category and geographical distribution. Doumbia et al. (2021) find a global decrease of container ship port calls in 2020 of 7%, with a monthly local reduction of 20% for Europe in June. The timing and magnitude of reductions reported in these studies agree with our findings: On average, AIS data indicates a reduction of 9% (2%) in April-December for Gibraltar (Mediterranean). The TROPOMI based emission estimates agrees in timing, but shows larger magnitude in reduction (20% and 10%) in April-December for the shipping lanes in the Strait of Gibraltar and the Eastern Mediterranean, respectively.

While the top-down emission reductions show a larger magnitude compared to the bottom-up emissions, they largely agree within the margin of uncertainty on a month-to-month basis. The difference in the mean reduction magnitude might be due to chemistry. The β values used here are calculated on a coarse grid ($0.5^\circ \cdot 0.67^\circ$). Additionally, we assume the chemical conditions in 2019 and 2020 to be similar to 2006 for when the β values were calculated. Furthermore, lateral transport complicates the choice of NO₂ background as a discrimination between land and ship emissions is not possible. The NO₂ background in turn has a large influence on the top-down emission estimate. These factors are all considered in our uncertainty estimates of the top-down emission changes. Other possible sources of uncertainty lie in the different temporal and spatial sampling of AIS and TROPOMI data and the simplified emission proxy. However, this is not expected to lead to a systematic bias.

3.4 Discussion and Conclusion

We used tropospheric NO₂ column observations from the TROPOMI sensor to optimally monitor ship NO₂ pollution and study the changes in ship NO_x emissions over European seas in 2019-2020. Satellite observations of tropospheric NO₂ columns provide valuable information on ship air pollution over open seas, which can be used to inform compliance monitoring by flag states and national authorities. We evaluated the high-resolution TROPOMI NO₂ retrievals for its potential to better detect ship NO₂ pollution. In European waters alone, TROPOMI finds 6 new lanes with enhanced NO₂ ranging from the Aegean Sea to the Skagerrak between Denmark and Norway, which are not detected by OMI, and which have not previously been reported in the literature. These newly found lanes of pollution coincide with busy sailing routes and bottom-up emission proxies.

To better understand the recent detection of an individual ship's NO₂ plume under conditions of sun glint, we examined how sun glint viewing geometries affect subsequent steps in the TROPOMI retrieval procedure. We find that sun glint drives higher apparent scene reflectivity, which enhances the signal strength from spectral fitting of NO₂ columns along the average light path by 20-30% over clear-sky shipping lanes. In such situations, the vertical sensitivity to NO₂ within the marine boundary layer increases by up to 60%. This effect is especially strong when sea surface wind speeds are low, but non-zero. When winds are strong, the wash causes sunlight to be reflected in other directions than directly towards the satellite, leading to little gain in vertical sensitivity. We find that the TROPOMI NO₂ algorithm accounts for these effects, so that data within and outside of sun glint geometries can be used with confidence. Nevertheless, our work clearly indicates that optimal spectral fitting can be accomplished for small scattering angles (<15°) and sea surface wind speeds of 1.5-3 m/s. Although selecting a subset fulfilling these sampling criteria reduces the amount of available data sharply, our findings indicate that sun glint conditions are beneficial for quantifying previously undetectable small NO_x emissions sources over open sea, and holding promise for also detecting other trace gases with UV/Vis satellite instruments over water, where surface reflectivity and vertical sensitivity is generally small.

In November 2020, KNMI implemented an improved FRESCO+ cloud retrieval called FRESCO+wide in the operational TROPOMI NO₂ algorithm. We find here that this new FRESCO+wide cloud retrieval provides some 50 hPa lower cloud pressures which agree better with coinciding cloud top heights from the VIIRS sensor than the standard FRESCO+. We show that the improved cloud pressures lead to a more realistic description of vertical sensitivities in the TROPOMI NO₂ algorithm, and at least partly address the known low bias in the tropospheric NO₂ product prior to November 2020, thus not only solving a known issue in the TROPOMI NO₂ retrieval but also increasing signal strength. We then trained a neural network on a limited data set of simultaneously available standard and improved cloud and NO₂ retrievals. Based on 4 different training sets, the neural network learned the statistical relationship between standard FRESCO+

cloud pressures and other parameters and the new tropospheric NO₂ columns. We used the neural network to predict updated NO₂ columns for the entire 2019-2020 TROPOMI NO₂ record. The neural network predicts a general increase in tropospheric NO₂ columns. Increases are particularly strong (up to $4 \cdot 10^{15}$ molec·cm⁻²) in the most polluted regions of Europe in wintertime. Our predicted (v2.1p) TROPOMI dataset enables the consistent analysis of temporal changes in NO₂ during the COVID-year 2020 and is useful to other data users until the TROPOMI NO₂ reprocessing scheduled for 2022 has been completed. We compared changes in our v2.1p TROPOMI NO₂ columns between 2019 and 2020 to changes in the number of ships, their speed and their size obtained from AIS data in the main European traffic lanes. From April 2020 onwards, TROPOMI observes 25% less NO₂ pollution than in the year before, in step with a 10% reduction in the number of ships and a 5% speed reduction relative to 2019. Accounting for non-linearity in local NO_x chemistry, we infer an average 20% reduction in top-down NO_x emissions in the Strait of Gibraltar from ships during months in which COVID-measures were in force in Europe, and global mobility decreased as a result of the pandemic. For future research, a full chemical transport modelling of AIS-based emissions and strict co-sampling of AIS and TROPOMI data can help understanding the observed differences in top-down and bottom-up emission changes and reduce the error margins.

We showed that TROPOMI is a superior instrument to analyze relatively small enhancements in NO₂ pollution over dark European seas. Its vertical sensitivity to ship pollution is substantially enhanced for small scattering angles under cloud free conditions and low wind speeds. Such sun glint scenes should allow improved detection of other pollutants, such as formaldehyde and SO₂, as well. KNMI's operational TROPOMI NO₂ product is subject to continuous improvement, which causes step changes in the publicly available data record until the official reprocessing has been finalized. Our improved (v2.1p) TROPOMI dataset offers a consistent alternative that can be used over Europe in and after 2019, and may be applied to other regions of the world where consistent NO₂ time series are needed.

3.A Appendix

3.A.1 Zoomed in NO₂ maps

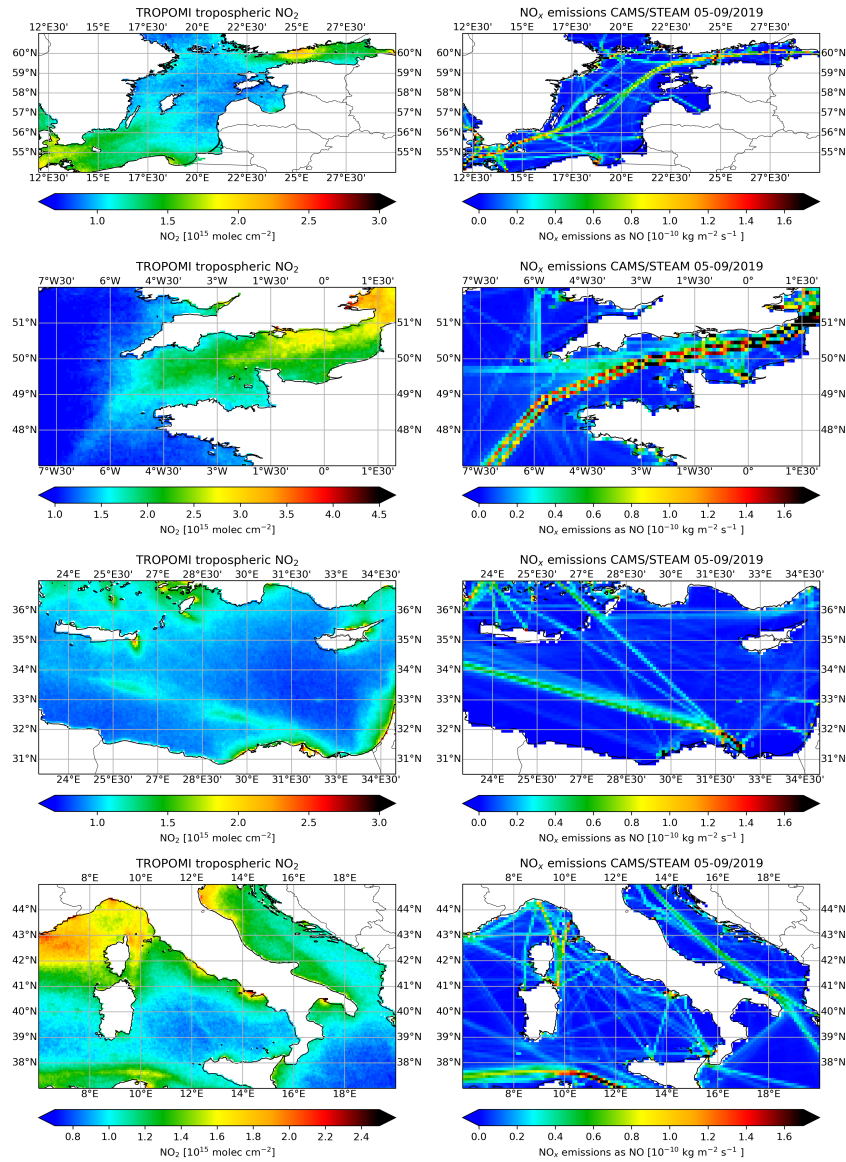


Figure 3.A.1: Summertime mean (May-September) tropospheric NO₂ columns from TROPOMI (left panel) and summertime mean NO_x emissions from the CAMS/STEAM emission inventory (right, Granier et al. (2019); Johansson et al. (2017)).

3.A.2 $N_{TROP,geo}$

We calculate a geometric tropospheric vertical column density $N_{trop,geo}$ using

$$N_{trop,geo} = N_{s,trop}/M_{geo} \quad (5)$$

where $N_{s,trop}$ is the tropospheric slant column density which can be calculated from the TROPOMI files using

$$N_{s,trop} = N_{s,tot} - N_{s,strat} = N_{s,tot} - N_{v,strat} * M_{strat} \quad (6)$$

where M , N_s , and N_v mean air mass factor, slant column density, and vertical column density, respectively. The subscripts trop, tot, and strat indicate tropospheric, total, and stratospheric columns, respectively. M_{geo} can be calculated using the solar zenith angle θ and the viewing zenith angle θ_0 as $M_{geo} = 1/\cos(\theta) + 1/\cos(\theta_0)$. The resulting tropospheric column is shown in Fig. 3.A.2.

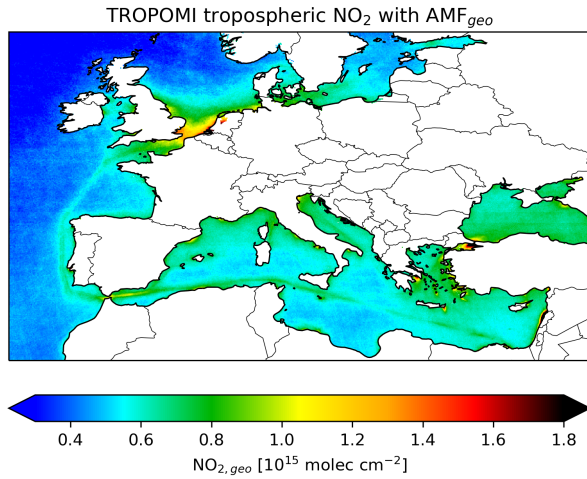


Figure 3.A.2: Mean of NO₂ columns calculated with geometrical AMF for summer 2019 (May-September), land areas have been whitened for clarity.

3.A.3 Spatial correlation to emissions

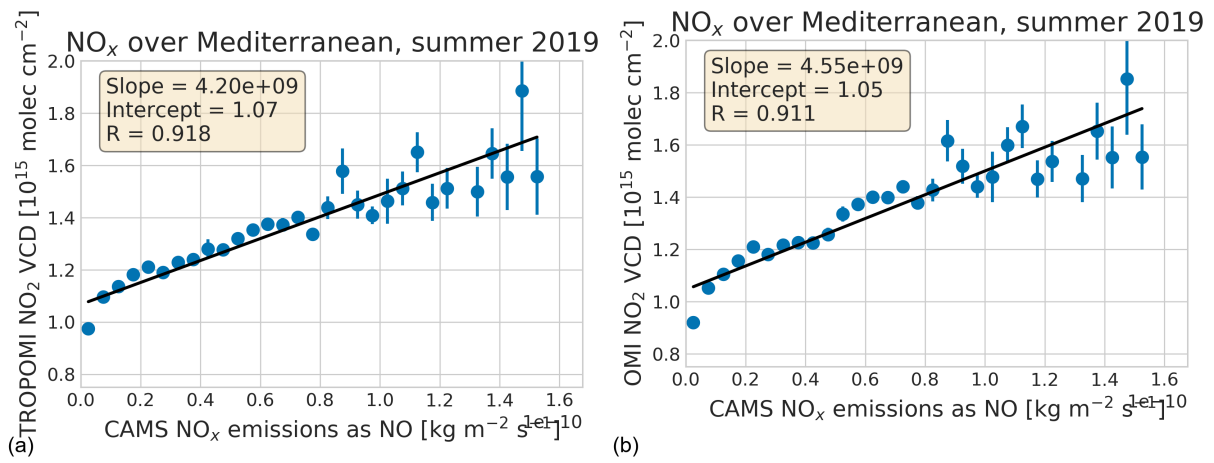


Figure 3.A.3: Scatter of binned summertime (May-September) 2019 tropospheric NO₂ columns vs emissions from CAMS/STEAM for the same period at 0.1°x0.1° in the Mediterranean. Error bars indicate the standard error of the bin. Left panel: TROPOMI, right panel: OMI.

3.A.4 Cloud fractions

The improved (v2.1) and old (v1.2) cloud fractions have strong correlation ($R^2=0.99$), but v2.1 cloud fractions are 5% lower on average, see Table 3.A.1. The spatio-temporal correlation between TROPOMI v2.1 and the well-established OMI QA4ECV cloud fraction product is also very high ($R^2=0.78$), with TROPOMI v2.1 cloud fractions 3% lower than OMI on average. TROPOMI v2.1 shows high correlation ($R^2=0.66$) and somewhat lower cloud fractions (-11%) compared to the co-sampled effective VIIRS cloud fractions. TROPOMI cloud fractions are especially lower for partly cloud-covered scenes, possibly resulting from biased surface albedo's assumed in the TROPOMI retrieval (from the GOME-2 climatology at 0.5° resolution, see Table 3.1). We find similar high correlation and small differences between TROPOMI and independent data over the Mediterranean Sea and Northwestern Europe as shown in Table 3.A.1.

Table 3.A.1: Evaluation of TROPOMI v2.1 cloud fractions over European shipping lanes (1-6 July 2018) against reference data.

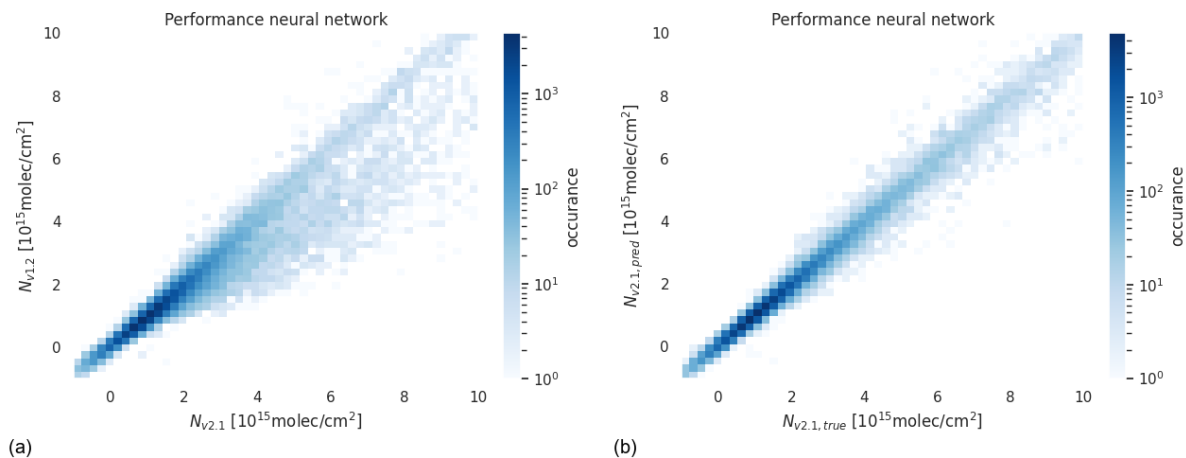
Shipping lane		Mean bias	RMS	R^2	Regression
Biscay	TROPOMI v2.1 vs. 1.2	-0.020	0.036	0.99	0.95·x
	TROPOMI v2.1 vs. OMI QA4ECV	0.002	0.124	0.78	0.01+0.97·x
	TROPOMI v2.1 vs. VIIRS	-0.062	0.181	0.66	-0.01+0.89·x
Mediterranean	TROPOMI v2.1 vs. 1.2	-0.009	0.017	0.99	0.96·x
	TROPOMI v2.1 vs. OMI QA4ECV	0.00005	0.09	0.67	-0.01 + 1.05·x
	TROPOMI v2.1 vs. VIIRS	-0.050	0.147	0.60	0.02+0.65·x
NW Europe	TROPOMI v2.1 vs. 1.2	-0.015	0.046	0.95	0.94·x
	TROPOMI v2.1 vs. OMI QA4ECV	-0.038	0.111	0.76	-0.01+0.91·x
	TROPOMI v2.1 vs. VIIRS	-0.026	0.156	0.64	0.04+0.74·x

3.A.5 Cloud pressures

Table 3.A.2: Evaluation of TROPOMI v2.1 cloud pressures against reference data for European shipping lanes.

		Median cloud pressure [hPa]	10th/90th percentile [hPa]	Geometric mean [hPa]
Mediterranean	TROPOMI v1.2	980	684/1010	920
	TROPOMI v2.1	947	653/978	889
	OMI QA4ECV	781	509/903	739
	VIIRS	935	743/976	896
NW Europe	TROPOMI v1.2	839	504/969	785
	TROPOMI v2.1	861	590/955	812
	OMI QA4ECV	740	474/862	712
	VIIRS	863	702/993	853

3.A.6 DNN



(a)

(b)

Figure 3.A.4: (a) Scatterplot of TROPOMI v1.2 (uncorrected) vs. actually retrieved v2.1 TROPOMI NO₂ columns observed over Europe in the 4 test periods. (b) Scatterplot of DNN-predicted vs. actually retrieved v2.1 TROPOMI NO₂ columns observed over Europe in all areas and periods under study.

An artificial Neural Network allows us to predict v2.1 columns for the full TROPOMI mission period up to December 2020. We find the predicted v2.1 columns to be close to actual retrieved v2.1 in a testing data set. Figure 3.A.4 illustrates the skill of the DNN approach to reliably predict v2.1 data: as a reduced major axis regression shows, the DNN-predicted v2.1 (hereafter v2.1p) NO₂ columns agree substantially better with the retrieved v2.1 NO₂ values ($N_{v2.1,true} = 0.98 \cdot N_{v2.1,pred} + 0.03 \cdot 10^{15} \text{ molec} \cdot \text{cm}^{-2}$, $R^2=0.98$, $n=56219$) compared to the originally retrieved v1.2 NO₂ columns ($N_{v2.1} = 0.87 \cdot N_{v1.2} + 0.09 \cdot 10^{15} \text{ molec} \cdot \text{cm}^{-2}$, $R^2 = 0.91$). The improvement from TROPOMI v1.2 to v2.1 is driven by the improved cloud pressures and associated changes in the tropospheric AMFs.

We trained the artificial Deep Neural Network (DNN) using the Python package Keras (Chollet, 2015) with three hidden layers. We divided the combined v1.2 and v2.1 data sets in 3 random subsets for training (60%), validation (20%), and testing (20%). The input parameters to predict TROPOMI (pseudo) v2.1 NO₂ columns are $N_{v,v1.2}$, M_{trop} , f_{cl} , p_{cl} , all viewing geometry parameters, surface albedo, and the qa value (all from v1.2). The DNN was then trained to minimize the mean absolute difference between the predicted and actually retrieved v2.1 NO₂ columns from the training set. This means our prediction does not use FRESCO+ wide cloud pressures for dates outside the training set period. Rather, the DNN has been trained to predict new NO₂ columns based on the old FRESCO+ cloud pressures and other parameters. Our DNN application succeeds in reducing the mean difference between the predicted and retrieved v2.1 NO₂ columns to $< 0.01 \cdot 10^{15} \text{ molec} \cdot \text{cm}^{-2}$ (original v2.1 – v1.2 mean difference was $0.12 \cdot 10^{15} \text{ molec} \cdot \text{cm}^{-2}$) over the 3 areas of study during the 4 periods, suggesting considerable skill in the DNN

approach. Our improved data set consists of the original L2 TROPOMI NETCDF files with the predicted change in tropospheric NO₂ columns as additional variable.

To show that DNN is capable of capturing seasonal variations in NO₂ corrections and, more broadly, that we can use a generic DNN to correct historic TROPOMI v1.2 data, we train a DNN based on 3 seasons (Summer, Winter, and Spring) and tested its predicted NO₂ columns against actually retrieved v2.1 data in Autumn. This analysis is done for the 3 testing areas defined in 3.3.3. After application of DNN, the mean discrepancy between predicted and retrieved v2.1 NO₂ columns reduces to $< 0.01 \cdot 10^{15} \text{ molec}\cdot\text{cm}^{-2}$ (original mean discrepancy: $0.09 \cdot 10^{15} \text{ molec}\cdot\text{cm}^{-2}$) and R² improved from 0.82 to 0.97.

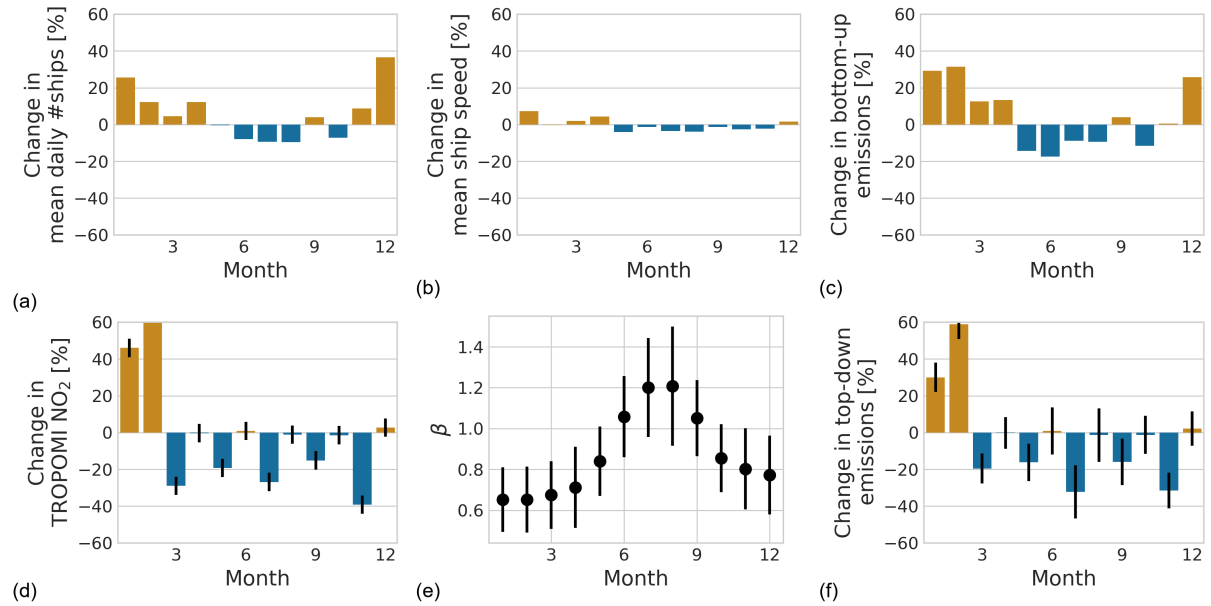
3.A.7 Ship NO_x emission reductions during the COVID-19 pandemic

Figure 3.A.5: (a) Relative change in monthly mean of daily number of ships passing the Mediterranean shipping lane between 2019 and 2020. (b) Same but with average ship speed. (c) Relative change in emission proxy ($v^3 \cdot L^2$). (d) Relative change in TROPOMI shipping NO_2 , the error bars indicate the sensitivity to changes in the area of study. (e) Monthly β values as discussed in Sec. 3.2.4. Error bars represent uncertainty originating from the temporal and spatial variability (see discussion in the text). (f) Relative change in top-down emissions from shipping. Error bars represent the propagated uncertainties in TROPOMI shipping NO_2 , beta and differences in meteorological conditions between 2019 and 2020 (see discussion in the text).

3.S Supplement

3.S.1 NECA and SECA guidelines

Inside the so-called Nitrogen Emission Control Areas (NECAs), the amount of maximum allowed NO_x emitted per engine kWh output is regulated by IMO guidelines. The limit of allowed NO_x emissions per ship is based on the building year of a ship (Table 5.S.2). Globally, all new ships have to comply with IMO's Tier II NO_x standard. Since 1 January 2021, the stricter Tier III standard applies to new ships operating within the European NECA, which is expected to reduce per ship emissions by 75% relative to Tier II. The new Tier III guidelines have been in place for American NECAs since 1 January 2016. The Sulphur Emission Control Area (SECA) has been in place in Europe since May 2005. The amount of sulphur allowed in the ship's fuel is capped by MARPOL Annex VI at 4.5% globally, but with progressively tighter limits after that date. Within SECA zones, the amount of fuel sulphur content has been tightened from 1.5% in 2005 to 0.1% after 2015. On January 1st 2020, the amount of sulphur allowed outside SECAs has been sharply reduced (Table 3.S.2).

Furthermore, the number of SECA areas is being expanded. For example, late 2021 it was decided to assign a Mediterranean SECA starting 1 January 2025.

Table 3.S.1: NECA guidelines according to MARPOL Annex VI. These guidelines apply to newly built ships operating within an emission control area (ECA).

Ships built after		Total weighted cycle emission limit $\text{NO} + \text{NO}_2$ (g/kWh) for engine speed n (revolutions per minute)		
		$n = 130$	$n = 1000$	$n = 2000$
Tier I	1 January 2000	17.0	11.3	9.8
Tier II	1 January 2011	14.4	9.0	7.7
Tier III	1 January 2021	3.4	2.3	2.0

Table 3.S.2: SECA guidelines according to MARPOL Annex VI. These guidelines apply to all ships.

Date of enforcement	Maximum fuel sulphur content allowed outside SECA (in %)	Maximum fuel sulphur content allowed inside SECA (in %)
May 2005	4.50	1.50
July 2010	4.50	1.00
January 2012	3.50	1.00
January 2015	3.50	0.10
January 2020	0.50	0.10

3.S.2 Scattering Angle

The scattering angle Θ can be calculated as

$$\Theta = \arccos [\cos \theta \cos \theta_0 - \sin \theta \sin \theta_0 \cos (\phi_0 - \phi)] \quad (7)$$

where θ and θ_0 are the solar and viewing zenith angles and ϕ and ϕ_0 the solar and viewing azimuth angles, respectively.

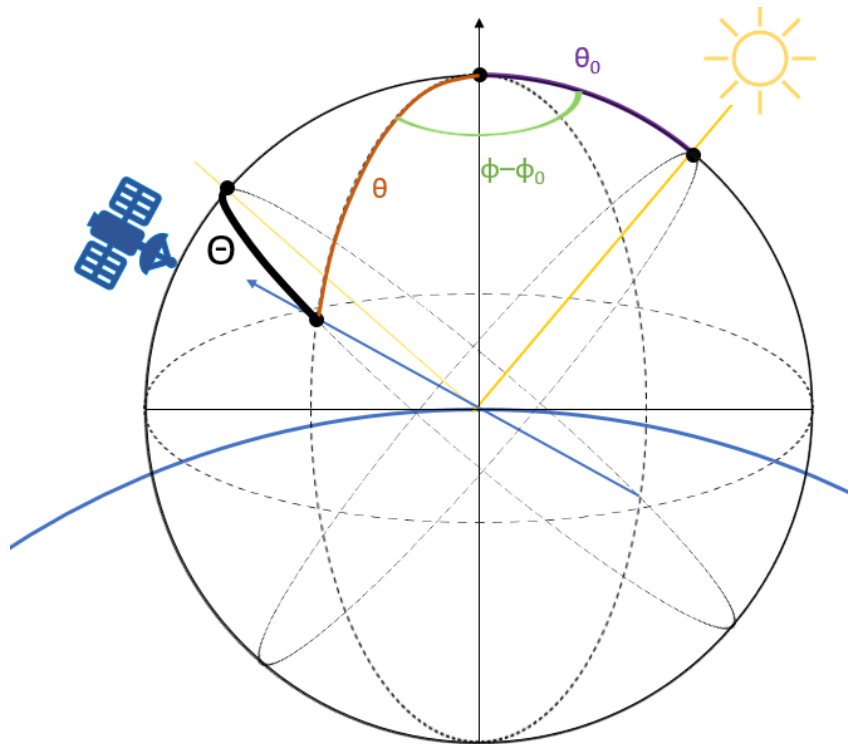


Figure 3.S.1: Sketch of satellite and solar geometry. The thick and thin yellow lines indicate the incoming and directly reflected solar light, respectively. The blue arrow indicates the light scattered in the direction of the satellite. The large blue arc is the Earth's surface. In green we show the difference of the viewing (satellite) and solar azimuth angle, orange and purple show the viewing and solar zenith angles, respectively. The thick black arc is the scattering angle, i.e the angle between the reflected sunlight and the viewing direction.

3.S.3 COVID

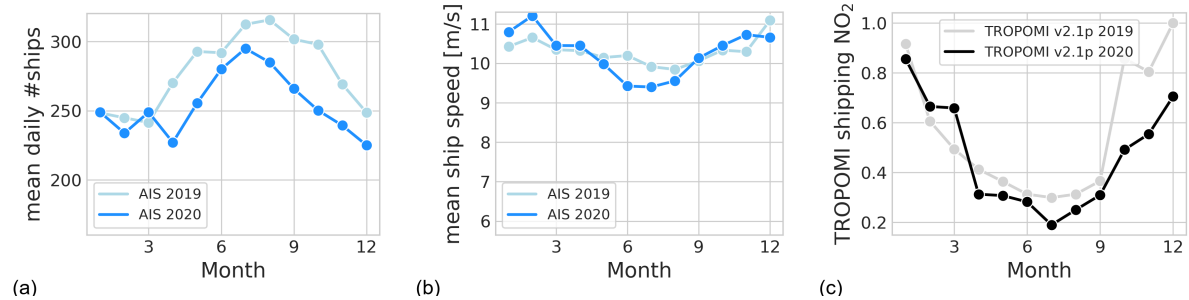


Figure 3.S.2: Left: Monthly mean of daily number of ships passing the Strait of Gibraltar in 2019 (light blue) and 2020 (darker blue) as defined in Section 3.2.4 for the main text. Center: Same but with average ship speed. Right: Monthly shipping NO_2 for 2019 (grey) and 2020 (black) from TROPOMI in arbitrary units for the area shown as purple rectangles in Fig. 1 (c) of the main text.

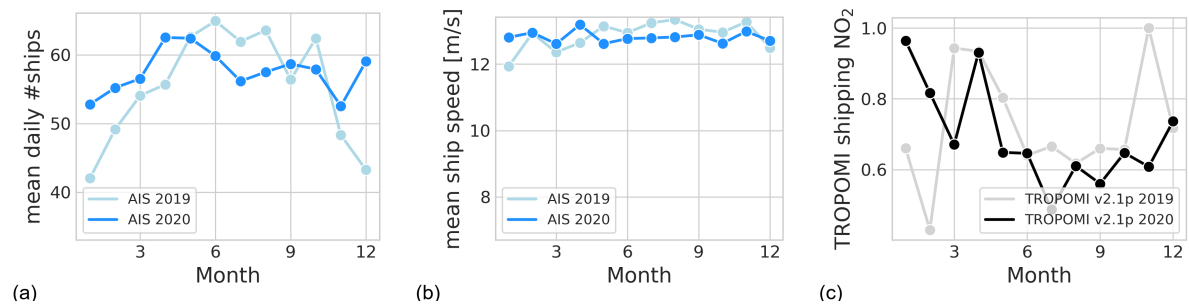


Figure 3.S.3: Left: Monthly mean of daily number of ships passing the Mediterranean shipping lane in 2019 (light blue) and 2020 (darker blue) as defined in Section 3.2.4 for the main text. Center: Same but with average ship speed. Right: Monthly shipping NO_2 for 2019 (grey) and 2020 (black) from TROPOMI in arbitrary units for the area shown as purple rectangles in Fig. 1 (c) of the main text.

3.S.4 Top-down emission uncertainties

For Figures 10 and F1 (e-f) we considered the following sources of uncertainty: $\sigma_{area} = 5\%$, $\sigma_{meteo} = 20\%$ and $\sigma_{\beta} = 0.15$.

To determine σ_{area} we performed a sensitivity analysis, making the cross sections used to calculate the TROPOMI shipping NO_2 20% smaller. As the difference in the relative changes between this and the default calculation was 5%, we use $\sigma_{area} = 5\%$ as the uncertainty of TROPOMI shipping NO_2 .

σ_{meteo} is the Root Mean Square of the differences between Vertical Column Densities with the same emission strength but different meteorology on a monthly average between 2019 and 2020. This is estimated to be 16% for Gibraltar and 11% for the Eastern Mediterranean. For this we used monthly mean NO_2 columns in the CAMS European Air Quality Forecasts ensemble mean (METEO FRANCE et al. (2020)) for 2019 and 2020 for the areas of study. This forecast represents a business-as-usual scenario where the meteorological conditions are taken into account, but emissions are not corrected for possible reductions due to COVID-19.

As we take β values for 2006 for changes between 2019 and 2020, we estimate a temporal variability in β from the 2005-2006 difference found in Vinken et al. (2014b) and combine this with the standard deviation of all β values from Verstraeten et al. (2015) in the area of study as monthly $\sigma_{\beta} = 0.15$.

Other sources of uncertainty are considered to be of minor importance. For example, while single TROPOMI columns have large (random and systematic) uncertainties these cancel out largely when taking the relative differences between 2 years. The large spatial and temporal sampling smoothens the random error while the systematic errors cancel out largely in the relative changes studied here.

Chapter 4

To new heights by flying low: Comparison of aircraft vertical NO₂ profiles to model simulations and implications for TROPOMI NO₂ retrievals

This chapter is based on:

Riess, T. C. V. W., Boersma, K. F., Van Roy, W., de Laat, J., Dammers, E., and van Vliet, J. (2023). To new heights by flying low: comparison of aircraft vertical NO₂ profiles to model simulations and implications for TROPOMI NO₂ retrievals. *Atmospheric Measurement Techniques*, 16(21):5287–5304. Publisher: Copernicus GmbH

Abstract

The sensitivity of satellites to air pollution close to the sea surface is decreased by scattering of light in the atmosphere and low sea surface albedo. To reliably retrieve tropospheric nitrogen dioxide (NO_2) columns using the TROPospheric Monitoring Instrument (TROPOMI), it is therefore necessary to have good a priori knowledge of the vertical distribution of NO_2 . In this study, we use an aircraft of the Royal Belgian Institute of Natural Sciences, equipped with a sniffer sensor system, measuring NO_x ($= \text{NO} + \text{NO}_2$), CO_2 and SO_2 . This instrumentation enables us to evaluate vertical profile shapes from several chemical transport models and to validate TROPOMI tropospheric NO_2 columns over the polluted North Sea in the summer of 2021. The aircraft sensor observes multiple clear signatures of ship plumes from seconds after emission to multiple kilometers downwind. Besides that, our results show that the chemical transport model TM5-MP, which is used in the retrieval of the operational TROPOMI NO_2 data, tends to underestimate surface level pollution - especially under conditions without land outflow - while overestimating NO_2 at higher levels over the study region. The higher horizontal resolution in the regional CAMS ensemble mean and the LOTOS-EUROS model improves the surface level pollution estimates. However, the models still systematically overestimate NO_2 levels at higher altitudes, indicating exaggerated vertical mixing and overall too much NO_2 in the models over the North Sea. When replacing the TM5 a priori NO_2 profiles with the aircraft-measured NO_2 profiles in the air mass factor (AMF) calculation, we find smaller recalculated AMFs. Subsequently, the retrieved NO_2 columns increase by 20%, indicating a significant negative bias in the operational TROPOMI NO_2 data product (up to v2.3.1) over the North Sea. This negative bias has important implications for estimating emissions over the sea. While TROPOMI NO_2 negative biases caused by the TM5 a priori profiles have also been reported over land, the reduced vertical mixing and smaller surface albedo over sea makes this issue especially relevant over sea and coastal regions.

4.1 Introduction

Satellite data of air pollutants is increasingly used for policy making, which requires reliable retrievals. This paper evaluates TROPOMI tropospheric NO₂ columns by comparing aircraft measurements of NO₂ profiles over the polluted North Sea to chemical transport models, and studying uncertainty and bias in the TROPOMI NO₂ retrieval from modeled profile shapes.

Nitrogen oxides (NO_x = NO+NO₂) decrease air quality, having negative impact on human health and environment. NO₂ is known to cause cardiovascular and respiratory diseases (Luo et al., 2016). Depending on chemical regime, nitrogen oxides also lead to surface O₃ formation which in turn harm the human respiratory system and plant growth. The international shipping sector is responsible for at least 15% of anthropogenic nitrogen oxides emissions globally (Crippa et al., 2018; Eyring et al., 2010; Johansson et al., 2017) while causing 3% of anthropogenic CO₂ emission (International Maritime Organization, 2020; European Commission, 2024).

While NO_x emissions from most anthropogenic sectors have been decreasing in recent years in western countries (e.g Zara et al., 2021; Fortems-Cheiney et al., 2021; Jiang et al., 2022 and references therein), intensity of ocean-going ships has been and is expected to keep rising (International Maritime Organization, 2020) and individual ships' NO_x emissions have been observed to increase (Van Roy et al., 2022b). NO_x emissions from shipping can lead to high background pollution levels in often densely populated coastal areas, limiting the impact of reductions in land-based emissions. For all the above reasons, international regulations for (newly build) ships constrain emissions with incremental limits. For example, the NO_x Emission Control Area (NECA) in the North and Baltic Sea came into effect on 1st January 2021, requiring that newly build ships sailing in these seas comply with International Maritime Organization (IMO) Tier III, which should result in 75% lower NO_x emissions compared to ships build since 2011 (International Maritime Organization, 2013). Details in emission limits depend on engine speed. For these regulations to be effective, monitoring of ship emissions is essential. Current monitoring routines include airplanes equipped with sniffer sensors (Van Roy et al., 2022b) or other remote sensing devices. Aircraft monitoring is costly, time consuming and practically feasible in coastal regions only. For a consistent, temporally and spatially complete approach current and upcoming satellite remote sensing missions offer promising options.

TROPOMI (TROPOspheric Measurement Instrument) on the European Sentinel-5 Precursor (S5P) is one of these satellite instruments and has been used to study NO_x emissions patterns within cities (Beirle et al., 2019; Goldberg et al., 2020; Lorente et al., 2019) as well as urban OH concentrations (Lama et al., 2022). While NO₂ over shipping lanes and its trends were previously studied on long-time averages of TROPOMI's predecessors GOME, SCIAMACHY & OMI (Richter et al., 2004; Beirle et al., 2004; Vinken et al., 2014b), the higher spatial resolution and lower noise of TROPOMI make single ship plume detection possible (Georgoulias et al., 2020). Recent studies succeeded to discrimi-

nate NO₂ ship plume signatures from the background using TROPOMI tropospheric NO₂ columns (Kurchaba et al., 2021; Finch et al., 2022). However, the validity of TROPOMI NO₂ and its uncertainties needs to be studied further to be able to reliably determine a ship's emissions and monitor compliance.

Prior knowledge of the state of the atmosphere during satellite remote sensing of trace gases such as NO₂ is key for the retrieval process. This includes surface radiative properties, radiative transfer in the atmosphere and vertical distribution of the trace gas. Much attention is therefore given to improve these aspects: recent updates in the cloud retrieval used for the TROPOMI NO₂ column retrieval lead to better agreement with independent data and reduce the known negative bias in tropospheric NO₂ columns (Van Geffen et al., 2021; Riess et al., 2022). Likewise, Riess et al. (2022) have shown that columns retrieved under sun glint conditions are reliable and enhance the instruments sensitivity to low altitude NO₂. Glint conditions are therefore in principle beneficial for the monitoring of NO_x emissions over sea. On the other hand, a priori profiles remain a source of uncertainty. The profiles from the Transport Model 5 (TM5-MP) with a resolution of 1°x1° used in the operational TROPOMI NO₂ product are very coarse compared to the ground pixel size of the measurements (3.5x5.5 km² at nadir) while NO₂ profiles close to spatially confined emission sources such as ships are expected to vary significantly within kilometers (Douros et al., 2023; Griffin et al., 2019; Ialongo et al., 2020; Chen et al., 2005). Additionally, uncertainties in the vertical mixing and thus in the a priori profile shapes, combined with the satellite's non-linear decreasing sensitivity towards the surface, pose a source of error. Furthermore, the model assumes temporally averaged emissions which does not hold for varying emission sources such as moving ships, adding to uncertainties in the a priori NO₂ profiles.

The TROPOMI NO₂ product allows the user to replace the a priori profiles with their own modelled or measured profiles (e.g. Visser et al., 2019; Douros et al., 2023). Douros et al. (2023) used the high-resolution CAMS ensemble mean NO₂ profile to replace the TM5-MP a priori NO₂ profiles in the calculation of the air mass factor (AMF) and to create an improved European TROPOMI NO₂ product. They found significant changes in resulting tropospheric columns with increases at hot-spot regions of typically 5-30%, depending on location and time. A similar study found a 20% increase in tropospheric columns over Europe when using LOTOS-EUROS profiles as a priori (Pseftogkas et al., 2022). For the above reasons, validation of these modelled a priori profiles is very important. In the past, validation has focused on land (Ialongo et al., 2020) and clean background over sea (Boersma et al., 2008; Shah et al., 2022; Wang et al., 2020). However, evaluation over and near shipping lanes is missing from literature.

In this study, we investigate aircraft-based in-situ measurements of NO_x (and more) over the polluted North Sea with major shipping routes and nearby industrial and densely populated centres. We combine ten spiral flights with three horizontal scans to obtain vertical NO₂ profiles in the lower 1.5 km of the troposphere. The aircraft is routinely used by the Belgian coast guard for compliance monitoring of ship emissions and is equipped

for measuring NO_x over sea. The aircraft measurements of 3-D NO_2 distributions over the North Sea provide a new means for satellite and model NO_2 validation. The aircraft profiles are representative of areas comparable to the TROPOMI ground pixel size. We compare the profiles to (temporally and spatially) coinciding modelled profiles from TM5-MP (as used in the operational TROPOMI NO_2 product), CAMS ensemble mean (as used in the European TROPOMI product by Douros et al., 2023), and LOTOS-EUROS. As a contrasting case, we show co-sampled model profiles over land close to the Cabauw tower in the Netherlands and compare the lowest 200 m to measured NO_2 concentrations, highlighting the special challenge of satellite trace gas retrievals over sea. In the last step, we present re-calculated TROPOMI NO_2 columns replacing the TM5-MP a priori NO_2 profile with the aircraft measured profile, accounting for the vertical sensitivity of the NO_2 retrieval and quantifying the error caused by a priori profiles modelled using coarse spatial resolution and time-averaged emissions.

4.2 Materials

The following section gives an overview of the data used and their sources, starting with the TROPOMI instrument in Sect. 4.2.1, followed by the aircraft, LOTOS-EUROS model data & ship location data in subsections 4.2.2, 4.2.3 & 4.2.4, respectively.

Table 4.1: Overview of the TROPOMI products used and their key differences.

NO_2 Retrieval	Processor version	Period covered	A priori profile	Adjustment of surface albedo
Operational product	v1.4	April 2018 - July 2021	TM5 $1^\circ \times 1^\circ$	No
	v2.2	July 2021 - November 2021		
Reprocessed PAL	v2.3.1	April 2018 - November 2021	TM5 $1^\circ \times 1^\circ$	Yes
TROPOMI _{CAMS}	v2.3.1	April 2018 - November 2021	CAMS $0.1^\circ \times 0.1^\circ$	Yes

4.2.1 TROPOMI NO_2 satellite data

Table 4.1 lists three different TROPOMI tropospheric NO_2 column data products used in this study. TROPOMI (Veefkind et al., 2012) is the single payload of S5P, which was launched in October 2017, and provides retrievals of various trace gases, including NO_2 , since April 2018. S5P is flying in a sun-synchronous, ascending orbit with an equator overpass time of 13:30 local time. With a swath width of approximately 2600 km TROPOMI has near daily coverage at the equator. At the latitude of the North Sea (52°N) S5P frequently overpasses the same ground scene twice per day. The spatial resolution is $5.5 \times 3.5 \text{ km}^2$ for nadir pixels, and $5.5 \times 14 \text{ km}^2$ for pixels at the edge of TROPOMI's swath.

The retrieval of tropospheric NO_2 columns follows a three-step procedure: retrieval of a slant column density (N_s) with the DOAS-method (Platt and Stutz, 2008) in the visible spectrum (405-465 nm), separation of the stratospheric and tropospheric contributions

($N_{s,trop}$), and conversion of the tropospheric slant column into a vertical column ($N_{v,trop}$) by application of the air mass factor (AMF, M): $N_{v,trop} = N_{s,trop}/M$. The single-pixel slant column detection limit ($0.5 \cdot 10^{15}$ molec/cm²) is determined by the uncertainty in the spectral fitting procedure and has been validated in Tack et al. (2021). Of most interest for this study is the calculation of the tropospheric AMFs, which is the dominant error source in the retrieval (Lorente et al., 2017; Boersma et al., 2018). The AMF depends on the solar zenith angle, on the satellite viewing zenith angle, on the scattering properties of the atmosphere and the surface, and on the vertical profile of the NO₂ in the troposphere (Martin et al., 2002; Boersma et al., 2004). For the TROPOMI NO₂ retrievals used here, the AMFs are calculated with the DAK radiative transfer model v3.3 (Lorente et al., 2017), based on pixel-specific input data on viewing geometry, surface albedo, cloud fraction and height, and the a priori vertical NO₂ profile. Scattering of light in the atmosphere together with the low sea surface albedo in the visible part of the spectrum decrease TROPOMI's sensitivity to NO₂ close to the sea surface (e.g. Eskes and Boersma, 2003; Vinken et al., 2014b). Good knowledge of a priori profiles as well as cloud coverage and surface albedo are therefore key for a good quality retrieval. In the recent version, the surface albedo is adjusted for individual scenes where the cloud retrieval gives negative cloud fractions using the original albedo database (van Geffen et al., 2021). While the cloud algorithm used in the TROPOMI operational NO₂ retrieval has recently been modified to provide a more accurate cloud pressure estimate for partially cloudy scenes (FRESCO+ wide) (Riess et al., 2022; Van Geffen et al., 2021), the a priori vertical NO₂ profiles remain a major source of AMF uncertainty, especially over sea.

4.2.2 Aircraft campaign over the North Sea

The Britten Norman Island (BN2) aircraft from the Royal Belgian Institute of Natural Sciences, operating from Antwerp airport, flew six missions over the North Sea between 2 June and 9 September 2021. The missions provided unique sampling of the marine mixed layer, intercepting outflow from land, and vertical profiles within the lower troposphere, from the sea surface (<30 m) to 1500 m.

The aircraft is equipped with a sniffer sensor system measuring NO₂, SO₂, and CO₂. This system is developed for the purpose of monitoring the compliance by ships to emission regulations (Mellqvist et al., 2017), specifically the MARPOL Annex VI regulation 13 on NO_x emission strength and MARPOL Annex VI regulation 14 on sulphur fuel content from ships. The detailed technical setup is described in Van Roy et al. (2022b,a,c). Of interest to our study is the NO_x sensor (Ecotech Serinus 40), which operates with two separate paths to determine the NO and NO_x concentration almost simultaneously and is in use since 2020. In the first path, the concentration of NO in the air sample is determined from the observed chemiluminescent intensity emitted by activated NO₂^{*}, which is produced when the air sample passes through a reaction cell filled with O₃ and proceeds through $\text{NO} + \text{O}_3 \rightarrow \text{NO}_2^* + \text{O}_2$ (Ecotech, 2023). The NO_x concentration in

the air sample is determined by first converting all NO_2 to NO , and then letting the total NO ($\text{NO} + \text{converted NO}_2$) in the second path react with ozone in the reaction cell, resulting in a chemiluminescence signal from activated NO_2^* . The NO_2 is then calculated as the difference between NO_x and NO over the measurement time interval of 10 s. A delay loop is installed between the two loops to ensure they sample the same air mass. A small mismatch can however not be ruled out. With an aircraft ground speed of 30–50 m/s, the horizontal scale at which NO_2 gradients can be detected is on the order of several hundred meters. The reported detection limit of the chemiluminescence analyser is 0.4 ppb (Ecotech, 2023). The sensor is equipped with an optical bandpass filter to avoid the measurement of interfering species and has successfully been used in previous scientific studies (e.g. Wong et al., 2022; Namdar-Khojasteh et al., 2022; Van Roy et al., 2022b).

The aircraft NO_2 campaign served two purposes. The first goal was to obtain vertical profiles of NO_2 in the vicinity of ships sailing the North Sea. The software on board the BN2 aircraft showed the live locations and tracks of ships within AIS range, as well as the expected location of the ship’s exhaust plume based on wind conditions and the speed and course of the ship. After visual detection and approaching of a ship, at least one transect through the ship’s plume was flown, followed by a spiraling climb from < 30 m to 1500 m altitude, continuously measuring NO and NO_x concentrations with a temporal resolution of 10 s. These vertical spirals were executed such that they coincide within 30 minutes of the TROPOMI overpass time on that day. The second goal of the campaign was to sample the horizontal distribution of air pollution within the lower marine boundary layer. On 8 September 2021, three zig-zag patterns were flown through the exhaust plume of ships at a constant altitude of approximately 40 m, where the aircraft would usually find the center of the plumes and the gradient between in-plume and outside-plume are the largest. The measurements of NO_x during these in- and out of plume patterns serve the purpose to better understand the spatial representativeness and distribution of NO_x concentrations in the presence of emitting ships at the scale of a TROPOMI pixel. Fig. 4.1 shows an overview of the campaign: The left panel shows the spatial extend of the flights as well as the NO_2 range measured, the right panel shows the mean measured NO_2 profiles as well as co-sampled model profiles. A detailed description of the weather and chemical conditions during the flights can be found in the supplementary materials S1.

4.2.3 LOTOS-EUROS model simulations

We use LOTOS-EUROS version 2.2.002 (LE, Manders et al., 2017; Thürkow et al., 2021) at $2 \times 2 \text{ km}^2$ resolution with 12 vertical levels (of which 7 are typically below 1500 m altitude) reaching up to around 9 km altitude. This model setup is similar to the model version operated within the CAMS ensemble and typically performs well in intercomparison studies, and is typically near the ensemble mean. The runs were performed over/around the Dutch North Sea for an area between $50.5\text{--}54.5^\circ\text{N}$ and $1.5\text{--}5.0^\circ\text{E}$ with a spin up time

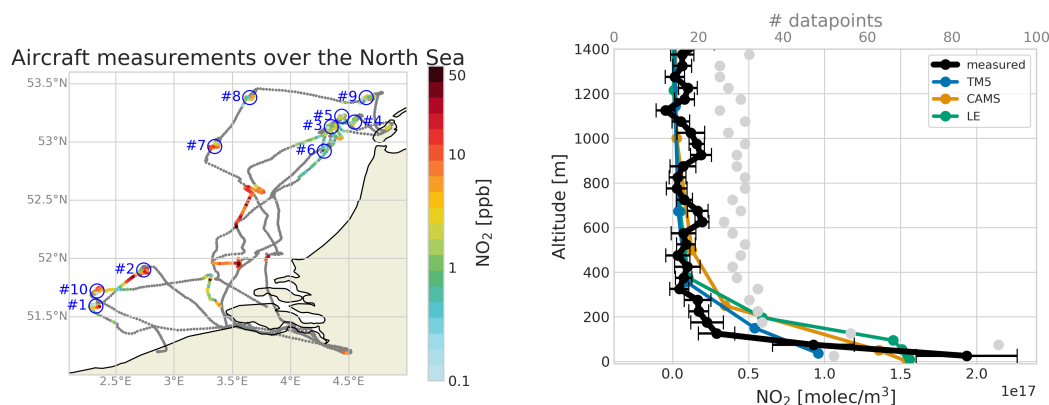


Figure 4.1: Left: Routes of all aircraft flights during the campaign. The 30 second mean NO₂ mixing ratio is shown as color for measurements in flight heights below 200 m. Blue circles indicate the locations of the spiral flights. Right: Mean vertical NO₂ profiles for the aircraft data (black), co-sampled TM5 (blue, Williams et al., 2017; Eskes and van Geffen, 2021; Huijnen et al., 2010), CAMS (yellow, METEO FRANCE et al., 2022; Marmer et al., 2009) and LOTOS-EUROS (green, Manders et al., 2017). The light gray dots indicate the number of 10 second NO₂ measurements at each height in the top x-axis. The aircraft profiles and their mean can be found in the dataset associated with this publication (see below).

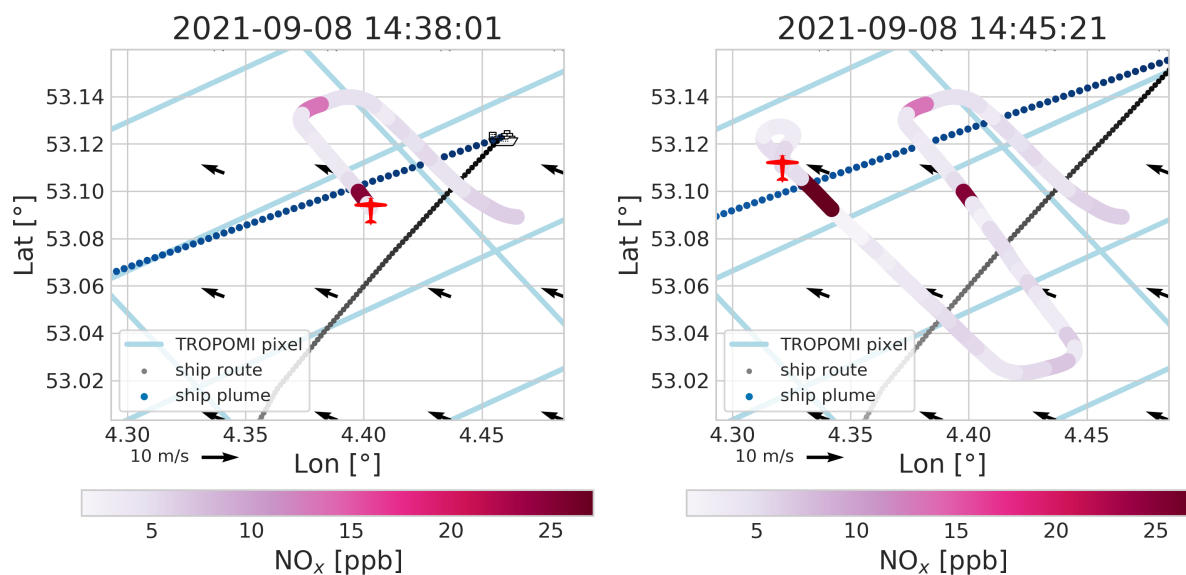


Figure 4.2: Two snapshots of one of the horizontal scans: Black and blue dots show ship path and plume center location at the moment indicated by the timestamp, respectively, with lighter colors indicated older locations. In pink we see the flight path with the color indicating the measured NO_x concentration. The light blue lines show the edges of TROPOMI pixels for the coinciding orbit. An animated version - illustrating the dynamics and highlighting the match between expected and observed plume location - is available in the supplement.

of one month. To ensure appropriate boundary conditions the model was nested within a LOTOS-EUROS run covering a part of north-western Europe (1-16°E, 47-56°N), which itself was nested within an European domain (15°W-35°E, 35-70°N), both run for a similar period and spin-up time.

Key characteristics of LOTOS-EUROS and other model data used in this study can be found in Table 4.2.

Table 4.2: Main characteristics of the model products used.

Model	LOTOS-EUROS	CAMS	TM5
Horizontal resolution	2x2 km ²	0.1°x0.1°	1°x1°
Emissions	CAMS-REG-AP_v5.1	CAMS-REG-AP	see Williams et al. (2017)
Meteorology	ECMWF Integrated Forecasting System (IFS)	IFS	ERA-Interim re-analysis
Vertical mixing scheme	See ECMWF (2015) with Monin–Obukhov length calculated as in Golder (1972)	Model-dependent	See Holtslag and Boville (1993)
Full description	Manders et al. (2017)	Marécal et al. (2015), METEO FRANCE et al. (2022)	Williams et al. (2017), Eskes and van Geffen (2021); Huijnen et al. (2010)

4.2.4 Ship location and course

To interpret the measured data we use AIS (Automatic Identification System) data on ship location, speed and heading together with the aircraft-measured wind data to predict the location of pollution plumes. The IMO requires all large ships (> 300 tonnes) to broadcast static (e.g. identity) as well as dynamic (position, speed) data, which can be received by other ships, shore stations, and satellites (International Maritime Organization, 2014). The historic AIS data set used here was made available to the Dutch Human Environment and Transport.

4.3 Aircraft NO₂ interpretation and representation at the scale of a TROPOMI pixel

The comparison of satellite retrievals with aircraft measurements requires that differences in sampling characteristics are reconciled first. Individual flights were not uniformly stretched out over a TROPOMI pixel, and the measured horizontal patterns in NO₂ concentrations reveal substantial variability within the spatial extent of a TROPOMI pixel, see Fig. 4.2. The observed spatial heterogeneity of NO₂ within a pixel is driven by the fraction of time the aircraft spent within ship plumes, and by the age of the plume at the moment of intercept (e.g. Chen et al., 2005). Additionally, the chosen aircraft operation and instrumentation requires post-processing of the measured data as detailed in the following section and Supplement S3.

Table 4.3: Overview of vertical profile flights taken during this campaign. Times are in UTC. Latitude and Longitude columns indicate the center of the profile.

Profile number	Date (dd.mm.yyyy)	Time	TROPOMI orbit	TROPOMI overpass	Latitude [°N]	Longitude [°E]
#1	02.06.2021	11:03-11:18	18842	12:00:15	51.59	2.33
#2	02.06.2021	11:36-11:50	18842	12:00:15	51.90	2.74
#3	22.07.2021	10:42-11:01	19551	11:23:04	53.13	4.35
#4	22.07.2021	11:16-11:33	19551	11:23:04	53.17	4.55
#5	22.07.2021	13:00-13:19	19552	13:02:56	53.22	4.44
#6	22.07.2021	13:36-13:54	19552	13:02:56	52.92	4.29
#7	08.09.2021	11:13-11:34	20232	11:23:15	52.96	3.35
#8	08.09.2021	11:51-12:12	20233	13:03:07	53.38	3.65
#9	08.09.2021	12:44-12:59	20233	13:03:07	53.38	4.65
#10	09.09.2021	15:56-16:10	20247	12:44:11	51.72	2.34

4.3.1 Representativeness of NO₂ vertical profile measurements

Pixel-scale aircraft NO₂ profiles

We first take care to ensure the representativeness of the aircraft NO₂ profiles at the scale of a TROPOMI pixel. The coastguard flights approached ships and their plumes in order to measure the composition of the exhaust. The measurements are therefore not necessarily representative of the mean NO₂ concentrations over the pixel: the aircraft may have spent a relatively large fraction of its measurement time within ship plumes compared to the fraction of the pixel filled with those plumes. Such a situation would lead to an overestimation of mean NO₂ concentration in a pixel. For each vertical profile flight listed in Table 4.3, we therefore calculated the ratio of the predicted fraction of the pixel covered by ship pollution plumes to the proportion of in-plume to overall time spent by the aircraft in a pixel. Fig. 4(a) illustrates the approach: the predicted plume-covered area is taken as the ratio of the grey area to the overall (grey and white) area, and the in-plume aircraft proportion is taken as the ratio of the time spent in the plume (red) to the total time spent below 100 m (all solid lines). Ideally, the two ratios would be identical, and a correction would not be needed. Using the AIS data we can calculate the expected presence of ship plumes in the lowest 100 m for all profile flights. No ship plume signatures were observed at higher altitudes. With the help of the three horizontal scans we predict the plume-covered area. On average, we over-sample plumes by a factor of 1.9 (0.0-5.7, median 1.1), meaning we spend disproportionately much time in the plume. We apply these as multiplicative correction factors to the in-plume and out-of-plume NO₂ values to improve the spatial representativeness of the vertical NO₂ profile for the TROPOMI pixel.

Plume NO_x-to-NO₂ conversion

The NO₂ measurement values are taken from the differences between the Ecotech-sensor's NO_x and NO concentrations. However, near the edges of plumes, we find unrealistically high or even negative NO₂ concentrations due to small time-delay between the NO_x and NO sampling in the Ecotech instrument, as mentioned in Sec. 4.2.2, and illustrated in Fig. 4.3 (right panel). When the aircraft samples background air, the NO₂ values inferred

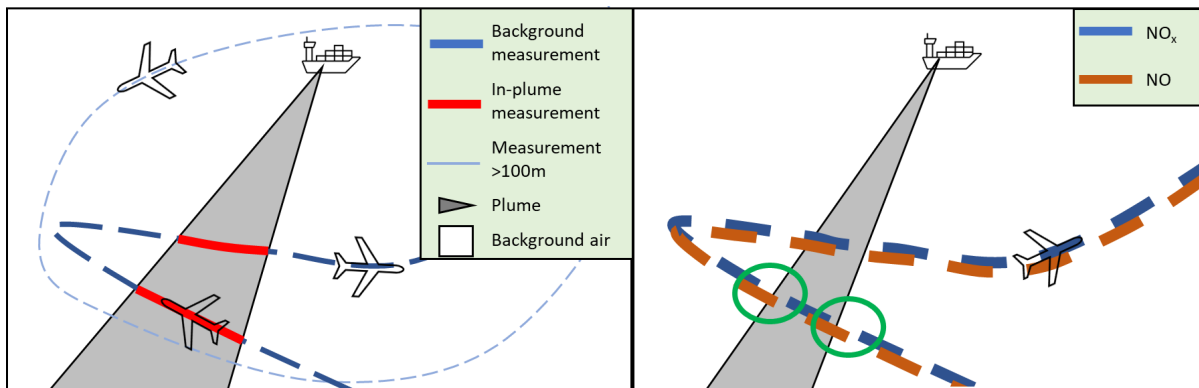


Figure 4.3: Sketches of profile flights visualizing the corrections. Left: The grey area indicates the part of the 2D-plane covered by a plume and the thick line the aircraft measurements in the polluted layer, with red showing in-plume measurements and blue indicating background sampling. The mismatch between the fraction of time spent in-plume and the fraction of the area covered by the plume is apparent. Right: The blue dashes indicate intervals of measuring NO_x, while the orange dashes indicate NO-intervals. For the situations highlighted by the green circles NO is measured partly in-plume while NO_x is measured fully in-plume (left circle) or out-of-plume (right). This will lead to negative or extremely high NO₂ values, respectively.

from NO_x - NO are still reliable in spite of the small delay. But when the aircraft samples the plume, we can not necessarily rely on NO_x - NO and instead convert the NO_x concentration measurements into NO₂ concentrations via local NO₂:NO_x ratios simulated with the PARANOX plume chemistry model which has been used before by Vinken et al. (2011) for ship plume modelling. PARANOX NO₂:NO_x ratio's depend strongly on the age of the plume, as NO_x in the early stages after emissions is mostly present as NO, but the NO₂ portion typically increases to 0.45 within some 15-30 minutes after emission following entrainment of O₃, and subsequent NO₂ formation via the NO + O₃ reaction in the plume. More details on PARANOX can be found in Supplement S2.

Zero-level offset calibration

The Ecotech sensor is capable of detecting clear in-plume NO₂ enhancements of several ppbs, but since near-zero, background air NO₂ levels differed by a few ppb between flights on different days, we re-calibrated the aircraft NO₂ concentrations to ensure that the measured near-zero NO₂ levels at altitudes above 250 m are on average consistent with NO₂ values from the CAMS simulations. The calibration offset is applied as an additive correction to the entire profile, and its value is consistent for multiple profiles measured on the same day, as anticipated from the daily calibration routine executed prior to flight. The calibration offsets vary between 0 and 4 ppb between the different days, and we assume a uncertainty of the bias correction of 0.5 ppb. Using only values above 500 m for the offset calculation leads to slightly different offsets that fall within the assumed uncertainty range.

For a more detailed description of the three corrections, see supplement S3.

4.3.2 Observed vertical NO₂ profiles

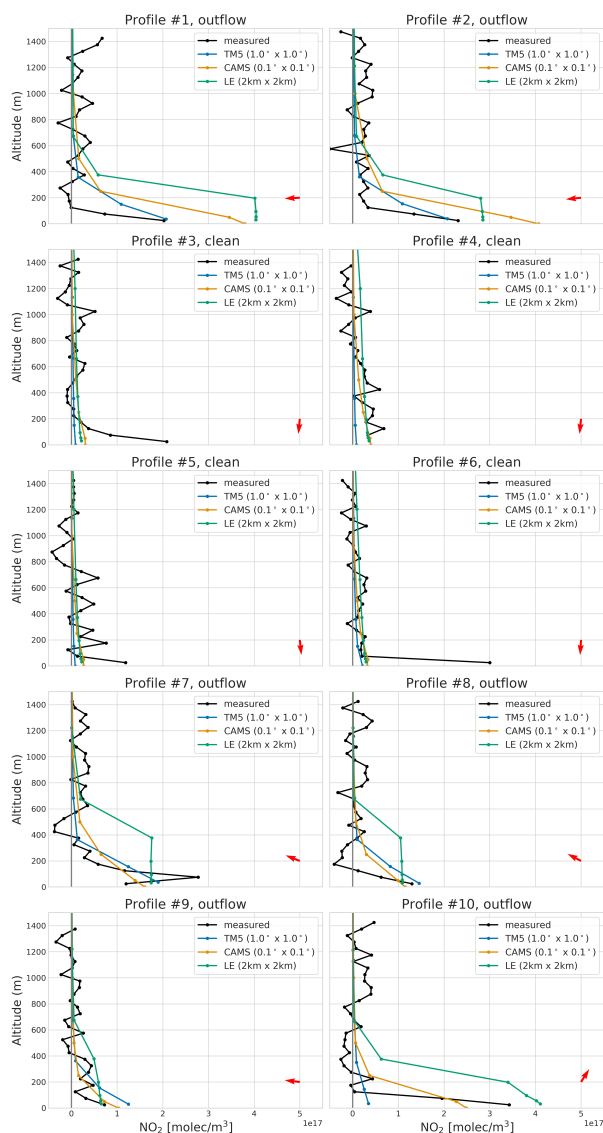


Figure 4.4: Profiles of all flights as well as coinciding TM5, CAMS ensemble mean and LOTOS-EUROS profiles. The red arrows indicate the mean measured wind direction during the profile flights. The indicators 'outflow' and 'clean' in the subtitles follow the classification in Sec. 4.3.2.

We now present the vertical NO₂ profiles obtained from the BN2 aircraft measurements over the North Sea following the procedure sketched in Sect. 4.3.1. Each of these vertical NO₂ profiles is spatially representative for the spatial scale of a TROPOMI pixel. For time and location of the profiles taken see Table 4.3. Aircraft NO₂ measurements were aggregated in 50 m altitude bins, where the reported altitude is the mean of the lower

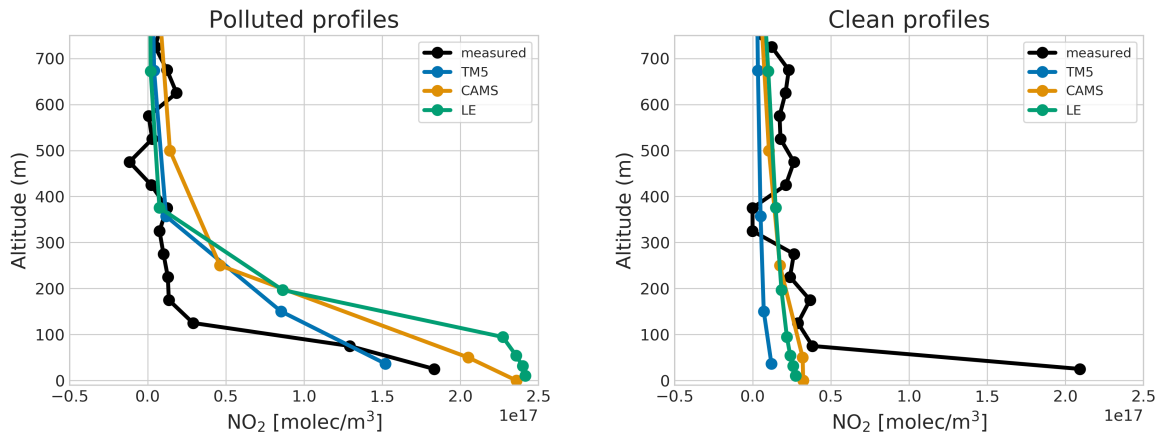


Figure 4.5: Mean aircraft-measured profiles as well as coinciding TM5, CAMS ensemble mean and LOTOS-EUROS profiles for land outflow (left, profiles 1, 2, 7, 8, 9, 10) and clean conditions/northerly winds (right, profiles 3, 4, 5, 6).

and upper boundary of each bin.

The aircraft data shows the highest NO₂ concentrations close to the sea surface, strongly decreasing within the lowest 100 m (Fig. 4.1). This is in agreement with the CO₂ profiles shown in S5. To better understand the emissions sources and physical transport processes leading to the observed profile shapes, we analyse simulations over the campaign period from the TM5-MP, CAMS, and LOTOS-EUROS models (see Sect. 4.2.3). The mean simulated NO₂ profiles coinciding with the aircraft flights show NO₂ pollution up to 200 m and above (Fig. 4.1). In the following, we will investigate the roles of model vertical mixing, emission strength, and transport of pollution from elsewhere as possible explanations for the mismatch between the simulations and observations. For that we need to study the NO₂ profiles according to their distinct meteorological circumstances. Fig. 4.4 shows the individual measured and modeled profiles with the numbering consistent to Table 4.3. For uncorrected profiles and the uncertainty estimates see Fig. 4. Meteorological conditions such as mean wind directions reveal that vertical profiles have been collected for two distinctly different types of situations over the North Sea: one with outflow of possibly polluted air from the Low Countries over the North Sea, and one under pristine conditions with wind from the North and low background NO₂ concentrations. Hereafter we classify these profiles as 'land outflow', and 'clean', see Fig. 4.5. A more complete description of the general chemical and meteorological conditions during each flight can be found in Supplement S1. **NO₂ profiles during land outflow - profiles 1, 2, 7, 8, 9, 10**

Fig. 4.6 shows the observed and simulated NO₂ in a situation of outflow from continental Europe. We see that the profile (indicated by the blue circle) was indeed sampled under conditions of pollution outflow from land. The corresponding profiles for all outflow cases in Fig. 4.4 show pollution close to the sea surface (see also the left panel of Fig. 4.5).

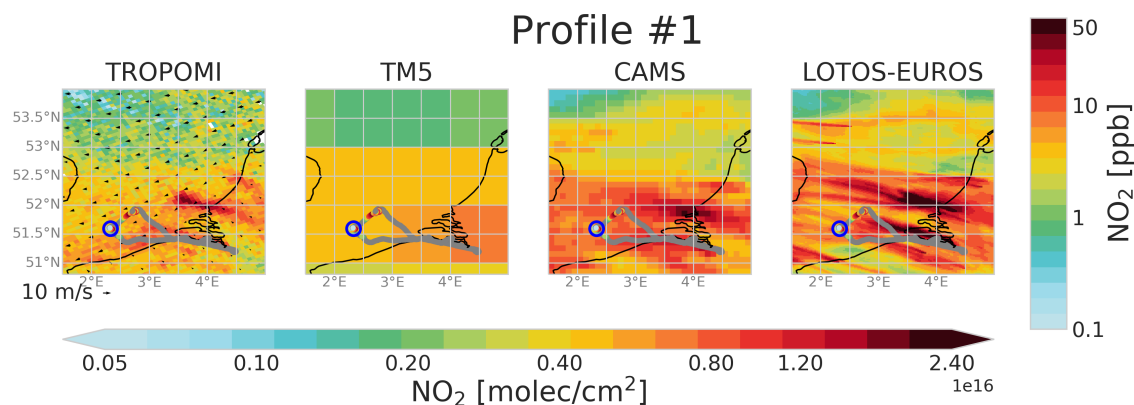


Figure 4.6: NO_2 columns (indicated by the bottom color bar) as seen by TROPOMI and several model products for the time of the first profile measurement. The aircraft measurements are overlaid in grey for flights above 200 m and in colors below as indicated by the colorbar on the right. Wind speed and direction at 10 m from ERA5 are indicated by the arrows in the left panel.

While the aircraft measured NO_2 is enhanced only in the lowest 100 m (for the exception of profile 7 see below), the models - especially LOTOS-EUROS - show elevated NO_2 usually up to 200 m and above. This gives an overestimation in the total NO_2 in the column. The measured and modelled potential temperature profiles (Fig. 2) show a cold sea surface with a strong gradient in the lowest 400 m, hinting at a strong stratification. Together with moderate wind speeds this indicates stable conditions with limited vertical mixing.

TM5 grid cells are very large and contain a mixture of land and sea surface as can be seen in Fig. 4.6. This means that emissions within the cell can originate from land-based sources as well as ships. Likewise, boundary layer dynamics are a mix of sea and land characteristics. Nonetheless, TM5 profiles show only slightly less NO_2 in the lowest layer than the LOTOS-EUROS, CAMS and the measured profiles for outflow cases (see Fig. 4.5, left). Overall, the coarse TM5 columns show reasonable agreement with TROPOMI retrieved columns during outflow conditions with the exception of profile 10 (see Fig. 4.4 and S5).

On the other hand, the higher horizontal resolution in CAMS and LOTOS-EUROS allows the separation of sea and land NO_x contributions. The resulting columns show massive outflow of NO_2 from land, we see plume-like structures from the region of Antwerp and Rotterdam in CAMS, LOTOS-EUROS and TROPOMI. The aircraft profile 1 shown in Fig. 4.6 was taken within the outflow of Antwerp pollution. LOTOS-EUROS, and to a lesser degree also CAMS, show overestimated NO_2 columns compared to TM5 and TROPOMI. This is in line with the observed profiles shown in Fig. 4.4 and 4.5: While surface NO_2 levels in LOTOS-EUROS and CAMS are in reasonable agreement with observations overall, the polluted layer is significantly deeper than in the observations, leading to a high bias in LOTOS-EUROS and CAMS NO_2 columns in these outflow cases. Addi-

tionally, CAMS and LOTOS-EUROS show two strong emission plumes in the North Sea (e.g. around 53.3°N, 2.5°E), which are not visible in TROPOMI or TM5. These likely originate from gas platforms, but the missing plumes in the TROPOMI observations point at large overestimations of the emission strength in the CAMS inventory (≈ 0.2 kg/s for these two sources). TROPOMI and modelled NO₂ columns during the other profile flights can be found in Supplement S4.

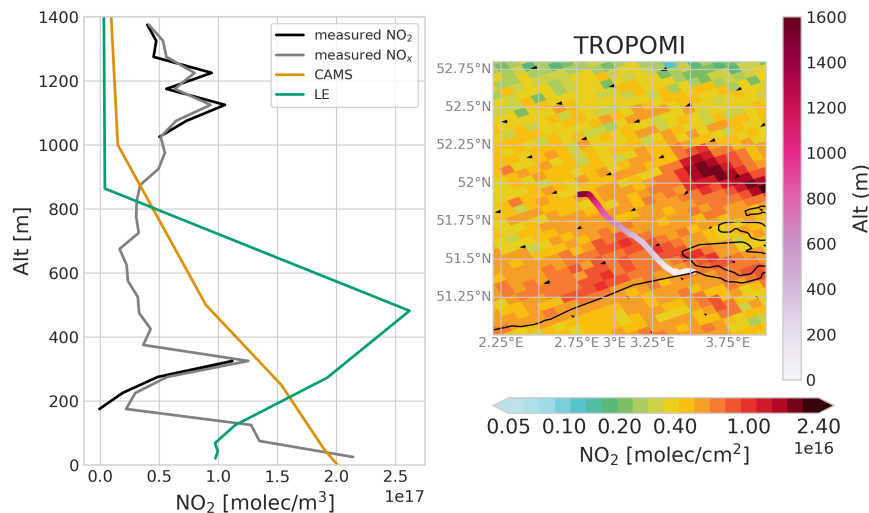


Figure 4.7: Left: Measured and modeled Vertical distribution of NO₂ along the flight path indicated on the right. This is not a vertical profile in the strict sense, as the sampling took place over ≈ 70 km horizontal extend. During part of the flight the airplane instrumentation was operating in a different mode so that no NO₂ data is available. However, NO_x (gray) was sampled throughout the whole flight and indicates a thin pollution layer between 300 and 400 m.

A special case is profile 7 on September 8th. This is the only profile with clearly enhanced NO₂ above 100 m (see also S5 for the CO₂ profile). In fact, the profile agrees reasonably well with TM5 and CAMS data, whereas LOTOS-EUROS again shows a too deep mixing layer and too much NO₂ in the column. This enhanced NO₂ observed between 100 and 300 m altitude might be caused by polluted air masses originating from the Netherlands and transported over Sea while rising above the stable surface layer. This hypothesis is supported by parts of the flight on June 2nd, when enhanced NO₂ was observed at an altitude of 300 m descending towards Antwerp airport into the land outflow after taking profile 2. A vertical profile for this part of the flight and the flight path can be seen in Fig. 4.7. The observed NO₂ layer at 300 m is also present in the co-sampled LOTOS-EUROS profile (as a thicker NO₂ layer around 500 m) but not in CAMS. These findings also demonstrate that the aircraft instrumentation is able not only to detect high NO₂ values in fresh plumes but also to capture diluted NO₂ pollution from land. Additionally, this suggests that at least for profile 2 (which was sampled right before) enhanced NO₂

seen at 200 m in the models are unlikely to be caused by land emissions, as pollution originating from land would be expected higher in the atmosphere. Finally, this indicates that land outflow often observed by TROPOMI over the North Sea can be located in higher atmospheric layers, where TROPOMI has a higher sensitivity (see Sect. 4.4) and thus possibly masking the low-level NO_2 from ships.

In summary, all models successfully simulate the occurrence of outflow and match the observed surface pollution reasonably well, but especially CAMS and LOTOS-EUROS overestimate the (vertically integrated) amount of NO_2 . From our observations it remains unclear whether the high NO_2 in LOTOS-EUROS and CAMS is caused by overestimations in land-based emissions, timing of the emissions in the models, advection, too long NO_2 lifetimes or vertical mixing. Similar to the other models, TM5 shows too high NO_2 at 200 m and above, hinting at uncertainties in the vertical mixing. The low surface pollution of TM5 in profile 10 likely showcasts the limitations of a coarse resolution. The very shallow pollution layer visible in the NO_2 measurements is also visible in the uncorrected and simultaneously measured CO_2 data (see S5) and therefore unlikely to result from the non-simultaneous measurement of NO_x species and our corrections.

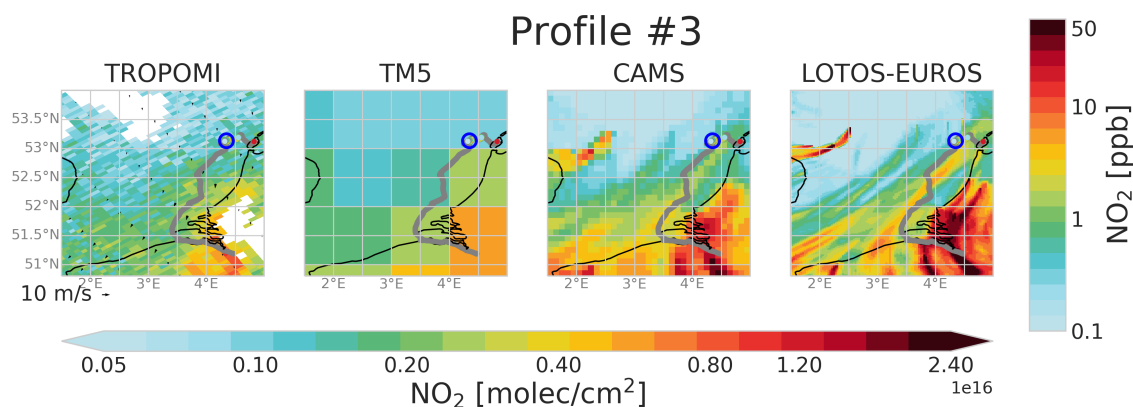


Figure 4.8: As Fig. 4.6 but now for the third profile. Profile 4 is taken in collocation with the same TROPOMI orbit and its location is shown in Fig. 5.

NO_2 profiles during clean conditions - profiles 3, 4, 5, 6

Fig. 4.8 shows the observed and simulated NO_2 in a situation without outflow from continental Europe. Profiles 3 to 6 have all been taken on the same day, 22 July 2021. On this day northern winds were prevailing, transporting clean air into the North Sea, resulting in low NO_2 columns as observed by TROPOMI in Fig. 4.8 and Fig. 1. The potential temperature profile on 22 July 2021 (see Fig. 2) indicates a well mixed marine boundary layer of 800 m depth. All modelled NO_2 profiles show little pollution at the surface and NO_2 concentrations are slightly decreasing towards higher altitudes. While the profiles were taken right above the shipping lane, marked by the blue circle in Fig. 4.8, in CAMS and LOTOS-EUROS the shipping pollution can be seen south of the profile, caused by the northerly winds. Again, TM5 shows less NO_2 compared to the other models (see Fig. 4.5, 4.8 & S5.).

The observed profiles 4 and 5 (see Fig. 4.4) agree reasonably well with the models, showing little NO₂ enhancement close to the sea surface. On the other hand, profiles 3 and 6 show strong NO₂ enhancements in the lowest 50 m, in contrast to the models. This is driven by exceptionally high NO_x concentration measured in ship plumes (>250 ppb NO_x for profile 3). In fact, a Monte Carlo approach (see Supplement S3 & Fig. 4, leading to a more multi-profile-average 'in plume' NO₂ concentration) shows very similar surface NO₂ values of $\approx 1.5 \cdot 10^{17}$ molec/m³ for all 4 flights on that day. This shows the presence of ship plumes in all 4 profiles, while in two cases the plume was either not captured well due to the temporal sampling of the Ecotech sensor or the ships in profiles 4 & 5 were emitting significantly less.

The mean clean profile in the right panel of Fig. 4.5 shows that none of the models captures the clear enhancement in the lowest 50 m due to NO_x emissions from ships. The ship NO_x emissions - while captured by the aircraft - are spatially diluted over the area of the model grid cell, especially for the coarse TM5 model, and throughout the well-mixed boundary layer and advected with the prevailing wind. Additionally, the models represent ships with averaged, constant emission fluxes in the model grid cells along the ship tracks, whereas in reality a ship might be in a given model grid cell for a short time with a higher emission flux. Therefore, in reality strongly localized emission levels are observed as sharply defined plumes, not resolved by the CTMs. These observations indicate the weakness of temporally and spatially averaged emissions in the models which fail to capture high pollution levels in the vicinity of strong and moving emitters. Overall, the models seem to underestimate the influence of ship emissions, likely due to temporal and spatial averaging of emissions and instant dilution thereof in the grid cell.

4.4 Validation of TROPOMI NO₂ over the North Sea

4.4.1 Recalculate AMFs

With the observed vertical NO₂ profiles we can calculate a modified TROPOMI NO₂ column, replacing the coarse TM5 a priori in the retrieval with aircraft-measurement based vertical profiles. As the measured NO₂ profiles only extend to 1400 m, we use TM5 profiles to fill the gap to the tropopause. The combined aircraft-TM5 profiles have then been interpolated and sampled according to the TM5-MP vertical levels. The adjusted tropospheric AMF $M_{\text{trop,ADJ}}$ can be calculated using the AMF from the a priori $M_{\text{trop,TM5}}$, the averaging kernels of layer l $A_{\text{trop},l}$ provided in the TROPOMI files as well as the NO₂ column density $x_{l,\text{meas}}$ of layer l from the aircraft data as

$$M_{\text{trop,ADJ}} = M_{\text{trop,TM5}} * \frac{\sum_{l=1}^L A_{\text{trop},l} x_{l,\text{meas}}}{\sum_{l=1}^L x_{l,\text{meas}}}$$

where L is the highest TM5 layer below the tropopause. Replacing the a priori with the measured NO_2 profiles and recalculating the AMFs is explicitly advised in the TROPOMI NO_2 documentation (Eskes and van Geffen, 2021) and has been done to improve satellite observations and validations previously (Visser et al., 2019; Douros et al., 2023). The adjusted vertical, tropospheric column can then be calculated as $N_{v,\text{trop,ADJ}} = N_s/M_{\text{trop,ADJ}}$. Too low NO_2 concentrations in TM5 close to the surface are expected to lead to a negative bias in the TROPOMI NO_2 retrievals, since the sensitivity to NO_2 close to sea surface is generally small as indicated by the averaging kernel (see Fig. 4.9). The shallow boundary layer depth over sea in combination with the low surface albedo values (≈ 0.04) emphasize the difficulty to detect air pollution over sea with satellite remote sensing, despite the high signal-to-noise ratio and resolution of TROPOMI NO_2 .

4.4.2 Tropospheric columns

We compare vertical tropospheric columns of NO_2 retrieved by TROPOMI (operational, PAL & CAMS) as well as measured columns. Lastly, we add the new product TROPOMI_{ADJ} which includes a re-calculation of the AMFs and vertical tropospheric NO_2 columns using the measured profiles following Sect. 4.4.1.

Table 4.4 shows the mean columns densities of all datasets mentioned above as well as their Pearson correlation coefficient and Root Mean Squared Error (RMSE) against the aircraft data. The ten aircraft measured NO_2 column densities averaged at $3.37 \cdot 10^{15}$ molec/cm². This is significantly higher than the coinciding operational TROPOMI ($2.42 \cdot 10^{15}$ molec/cm²) and TROPOMI_{PAL} ($2.47 \cdot 10^{15}$ molec/cm²) data. Using the re-calculated AMFs an average column density $N_{v,\text{trop,adj}}$ of 2.89 (2.71–3.23) $\cdot 10^{15}$ molec/cm² is determined. This is ≈ 20 (12 – 33)% higher than the TROPOMI products and brings the satellite retrievals closer to the columns determined from the aircraft measurements, showing a significant negative bias in operational TROPOMI NO_2 columns. The TROPOMI_{CAMS} dataset (see Sect. 4.2.1) is closer to the measured columns at mean columns of $3.03 \cdot 10^{15}$ molec/cm². It should be noted that CAMS NO_2 columns (see Figs. 4.6, 4.8 & S5) are systematically higher compared to measurements and TM5. TROPOMI_{CAMS} and TROPOMI_{ADJ} also show an increased Pearson correlation coefficient to the aircraft columns of 0.87 and 0.91, respectively, compared to 0.82 of the operational product. Lastly, the RMSE of the TROPOMI columns towards the aircraft columns is reducing going from the operational ($1.26 \cdot 10^{15}$ molec/cm²) to TROPOMI_{CAMS} ($0.99 \cdot 10^{15}$ molec/cm²) data and smallest for the aircraft-adjusted columns at $0.77 \cdot 10^{15}$ molec/cm².

Given the large uncertainty and corrections involved at the lowest level NO_2 concentration, the sensitivity of the recalculated AMFs to that value was tested. A 20% change in the NO_2 number density leads to a change in AMF of less than 5%, and even a change of 50% in surface level NO_2 changes the AMF only by 10%. This supports the finding of a negative bias caused by the a priori profile as the differences in AMFs can not be

explained by the surface level NO₂ alone.

Table 4.4: Tropospheric NO₂ columns measured by the aircraft and different TROPOMI products. For TROPOMI_{ADJ}, the values in the bracket give the average of the lower and upper estimates based on the uncertainties shown in Fig. 3.

Product	Mean tropospheric NO ₂ column [10 ¹⁵ molec/cm ²]	Correlation to aircraft column	RMSE to aircraft column [10 ¹⁵ molec/cm ²]	relative difference to aircraft column [%]
Aircraft	3.37	-	-	-
TROPOMI	2.42	0.82	1.26	-28
TROPOMI _{PAL}	2.47	0.83	1.24	-27
TROPOMI _{CAMS}	3.03	0.87	0.99	-10
TROPOMI _{ADJ}	2.89 (2.71-3.23)	0.91	0.77	-14

4.4.3 The land-sea contrast in TROPOMI NO₂ retrieval

As a contrasting case, Fig. 4.9 compares the sea NO₂ profiles to NO₂ profiles during the TROPOMI validation experiment (TROLIX) in 2019 (Sullivan et al., 2022) over the Netherlands (51.97°N, 4.93°E). The left panel shows mean TM5 NO₂ and averaging kernel profiles over land and sea at the time of the aircraft measurements as well as the mean aircraft-measured profiles. While modeled surface pollution levels over land are on average close to those over sea, the boundary layer is significantly more evolved with elevated pollution levels in the models reaching 400 m and above. At the same time, the averaging kernel over sea is smaller compared to land throughout the entire boundary layer. The right part of the same figure shows midday NO₂ concentrations measured at Cabauw tower as well as coinciding TM5 and CAMS profiles co-sampled during the TROLIX campaign which took place at a different time than the aircraft measurements, but under similar meteorological conditions. No measured profile data are available at Cabauw for the days of the aircraft campaign. The measurements confirm a well mixed lowest 200 m, in contrast to the presented profiles over sea. Even if the models would overestimate vertical mixing over land, the higher mixed layer over land would lead to a smaller relative difference between modeled NO₂ concentration and observations compared to over Sea. This - together with the lower surface albedo (<0.04 for the North Sea vs 0.05 for land) causing a lower sensitivity to NO₂ close to the surface - emphasizes the challenge of accurate satellite retrieval of NO₂ over sea compared to over land. For more details, see Supplement S6. Overall, we find on average 20% lower tropospheric AMFs over the North Sea compared to land given similar overall retrieval conditions.

4.5 Discussion

We evaluated the TROPOMI tropospheric vertical NO₂ column retrieval over the North Sea. For this, we measured ten vertical NO₂ profiles in the immediate vicinity of ships emitting air pollutants coinciding with the TROPOMI overpass, compared them to modeled profiles and studied the impact of a priori profiles on the TROPOMI NO₂ column

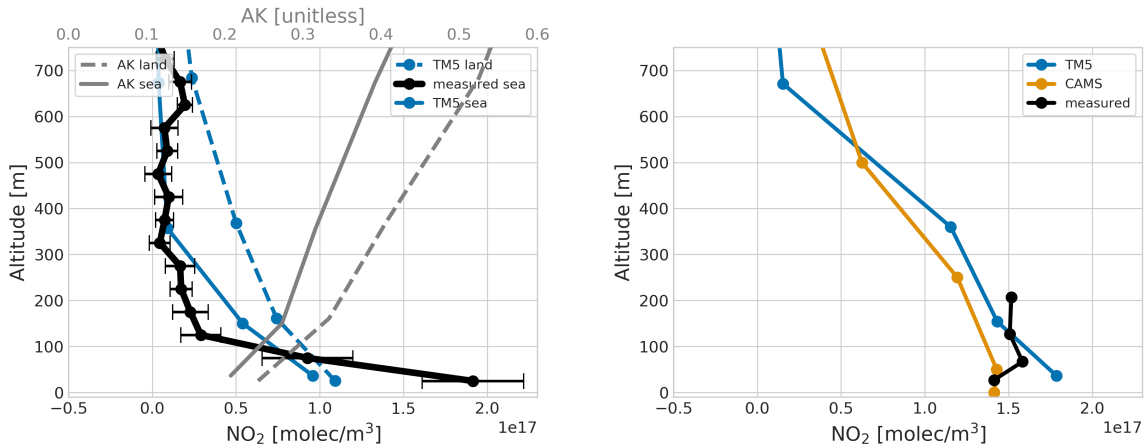


Figure 4.9: Left: The solid blue line shows mean TM5 profiles coinciding with the aircraft profiles (black). The dashed blue line shows simultaneous TM5 NO_2 profiles at the Cabauw tower in the Netherlands. Additionally, the mean TROPOMI averaging kernel profiles for land (sampled for all TROPOMI pixels within 51.90°N - 52.04°N and 4.86°E - 5.00°E) and sea (co-sampled with the aircraft profile measurements) are shown. The figure on the right shows mean measured (black) and modeled (TM5 in blue and CAMS in yellow) profiles at the Cabauw tower in the Netherlands for 6 cloud free days in September/October 2019 during the TROLIX-19 campaign (Sullivan et al., 2022).

retrieval.

Flying down to below 30 m above the sea surface allowed us to fully capture ship plumes and NO_2 pollution over the North Sea. While our measurements suffer from the indirect measurement of NO_2 , the horizontal zig-zag patterns and applied corrections lead to profiles that are truly representative at the time and scale of a TROPOMI pixel.

Our measurements strongly hint at systematic negative bias in TROPOMI NO_2 columns over the polluted North Sea. Using the aircraft profiles to recalculate the AMFs and tropospheric NO_2 columns, the TROPOMI columns are $\approx 20(12 - 33)\%$ larger on average compared to $\text{TROPOMI}_{\text{PAL}}$ data using TM5 for a priori profiles. This is in agreement with earlier studies (Douros et al., 2023) for point sources. The vertical profile measurements over the North Sea reveal a very shallow boundary layer of 100-150 m above sea level, where the averaging kernel is the smallest. With one exception our measurements show no significant pollution above 150 m. This finding is supported by co-sampled CO_2 profiles presented in S5. The low pollution layer is in contrast to model profiles and could be attributed to an overestimated vertical mixing in the models compared to observations on four summer days in 2021. The mixing schemes for vertical transport in the boundary layer used in TM5 (Williams et al., 2017; Holtslag and Boville, 1993) are known to overestimate vertical mixing for stable conditions (Köhler et al., 2011) which prevailed during several of the campaign days (see Sect. 4.3.2). The updated K-diffusion based on Monin-Obukhov length used in LOTOS-EUROS (ECMWF, 2015) is expected to result in more shallow stable boundary layers. However, we still find a high bias in LOTOS-EUROS in

the mixed layer height. Hints towards uncertainties in the vertical mixing of the LOTOS-EUROS can also be found in Escudero et al. (2019), who show a positive bias in boundary layer height (BLH) over Madrid in summer as well as overestimated vertical mixing in the boundary layer using the LOTOS-EUROS mixed-layer scheme. Additionally, they find more gradual vertical mixing and a better correlation of ozone surface measurements when increasing the number of vertical layers. Likewise, Skoulidou et al. (2021) connect underestimated surface NO_2 levels in Athens to problems in the temporal evolution of the BLH in LOTOS-EUROS, which is taken from the ECMWF operational weather analysis. The very shallow mixed layer observed during the flights is in agreement with the observed strong gradient in potential temperature and indicates stable conditions. The reasons the models fail to reproduce the shallow mixed layer over the North Sea remain unclear and need further studies.

Next to the overestimated mixing, the TM5 profiles during clean conditions show less pollution close to the surface than the aircraft data and the other model simulations. This is likely an effect of the coarse TM5 resolution of $1^\circ \times 1^\circ$ where ship emissions are smeared out over a larger area and time. The exaggerated vertical mixing and underestimation of the lowest part of the profile in TM5 leads to high-biased AMFs which in turn decreases the vertical column density via $N_v = N_s/M$. While the higher spatial resolutions of CAMS and LOTOS-EUROS increase the surface level NO_2 (in fact, for 8 out of 10 profiles, the surface pollution in these model product agrees reasonably well with observations), the overestimated pollution layer height, giving a substantial overestimation of the total NO_2 in the columns. This may be caused by overestimated NO_x emissions, their timing in the models, exaggerated advection or too long NO_x lifetimes, and shows that increased horizontal resolution does not necessarily give more accurate profile shapes. While TROPOMI columns using CAMS profiles as a priori are higher and show better correlation and lower RMSE to the aircraft columns than using TM5, this is caused rather by the higher NO_2 column than by a correct profile shape. The TROPOMI_{CAMS} product, essentially, demonstrates improved agreement with the aircraft column compared to the operational product. However, using the aircraft profiles in the AMF calculation exhibit the highest correlation and lowest RMSE.

Furthermore, we conclude that TM5, CAMS and LOTOS-EUROS are unable to fully capturing the spatially and temporally confined ship emissions over sea and that the pollution levels as a result of land outflow dominate the model results. This is supported by profiles 3-6, which were measured in clean conditions without land outflow. Observed and modeled temperature profiles indicate a well mixed atmosphere up to ≈ 800 m and we see little NO_2 enhancement in all model products while we observe strong enhancements in profiles 3 & 6 as discussed before. The observed enhancements can be directly linked to fresh ship plumes that show to be vertically confined to the lowest 50 m and are not present in the models. Better results can be expected with plume resolving models, incorporating ship plumes using AIS and ship specific data for their location and emission strength (e.g. from Jalkanen et al., 2016), or from a climatology of representative NO_2

profiles observed over shipping routes. The presented profiles can be the starting point for such a climatology.

More validation flights over polluted sea are desirable, especially spanning different locations, seasons and meteorological conditions as this study was limited to 4 days over the North Sea in summer. Six out of the ten profiles (on three of the four days) were taken under land outflow conditions. Being close to major polluting areas in the British Islands and North-Western Europe, land outflow happens frequently and we therefore expect these sampling conditions to be representative for the North Sea. While this study presents a cost-efficient way of measuring NO₂ profiles utilizing an aircraft already equipped for emissions monitoring, direct NO₂ measurements with a temporal resolution of 1 Hz or higher and higher accuracy could have reduced post-processing and uncertainties. Better calibration, a more sensitive sensor and expanding the flights to higher altitudes can further reduce the dependence on model simulations.

Overall, this study shows the bias arising from using modelled and uncertain a priori profiles. This is true especially over sea where the boundary layer is less developed than over land and the surface is darker. The observed negative bias in TROPOMI has important implications for the application of TROPOMI NO₂ columns for ship emission monitoring. As advised in Eskes and van Geffen (2021) the recalculation of AMFs using more realistic a priori profiles is beneficial.

4.6 Conclusion

This study clearly shows the need for additional evaluation of vertical NO₂ profiles over sea for both model and TROPOMI validation while providing a recipe for such an analysis. We present ten vertical profiles of NO₂ over the North Sea in Summer, which - due to the low-altitude sampling (≈ 30 m) and the location over busy shipping routes - present a unique opportunity to evaluate TROPOMI vertical NO₂ columns and model profiles (TM5, CAMS & LOTOS-EUROS) that was previously missing from literature.

We find that on average the coarse resolution of TM5 leads to too low NO₂ concentrations near the surface while overestimating NO₂ above 100 m. The higher model resolution of CAMS and LOTOS-EUROS results in more accurate surface NO₂ values, while at the same time vertical mixing is exaggerated compared to our observations. Additionally, CAMS and LOTOS-EUROS vertical NO₂ columns are too high compared to aircraft and TROPOMI data.

Furthermore, the comparison between observed and modeled vertical NO₂ profiles, along with the examination of TROPOMI averaging kernels over land and sea, stresses the significant challenges involved in accurately retrieving satellite NO₂ columns over sea, where vertical sensitivity to NO₂ is 20% lower than over land, because of lower surface albedo and confinement of NO₂ pollution in a thin marine boundary layer.

When replacing the TM5 a priori profiles with the aircraft-measured NO₂ profiles in the

TROPOMI AMF calculation, we find a significant increase of the retrieved vertical NO_2 columns of ≈ 20 (12 – 33)%, showing substantially improved agreement with aircraft-measured columns. Our findings align with previous studies (e.g. by Douros et al., 2023; Pseftogkas et al., 2022; Lorente et al., 2017), highlighting the importance of precise vertical a priori profiles for satellite-based trace gas retrieval.

4.S Supplement

4.S.1 General weather and chemical conditions

Figure S4.S.1 shows the flight paths for the 10 profiles and 3 horizontal scans overlaid over the coinciding TROPOMI NO₂ columns. Fig. 4.S.2 shows the measured and modeled potential Temperature profiles for the 10 vertical profiles.

Below we summarize the general weather conditions relevant for interpreting the aircraft data during the four flight days.

2 June: Moderate winds from the east/southeast, over the North Sea around 10 m/s, wind gusts similar, and mostly cloud free. Warm weather, temperatures over the Netherlands >25°C. Air pollution from the Randstad is blown onto the southern North Sea with a east-southeast to west-north-west orientation.

22/23 July: Light, mostly northerly winds mostly, with wind speeds of 3-7 m/s, wind gusts similar, mostly cloud free with some shallow cumulus and stratocumulus clouds over the North Sea. Pleasant weather, maximum temperatures over the Netherlands 20°-25°C. Chemical conditions are clean and northerly flow causes air pollution from the Netherlands to be blown away from the North Sea with a east-southeast to west-north-west orientation.

8 September: Light to moderate winds east/southeasterly, wind speeds 5-10 m/s, wind gusts similar. Nearly cloud free warm weather, maximum temperatures over the Netherlands >25°C. Air pollution from the Randstad is advected onto the southern North Sea.

9 September: Light to moderate winds from the south, wind speeds 5-10 m/s, wind gusts similar. Scattered low clouds and some stratocumulus fields, indicative of increasing atmospheric instability. Warm weather, maximum temperatures over the Netherlands >25°C. Air pollution from the south-western Netherlands is blown onto the southern North Sea with a north-south orientation. The Southernmost North Sea is relatively clean with a southwesterly flow.

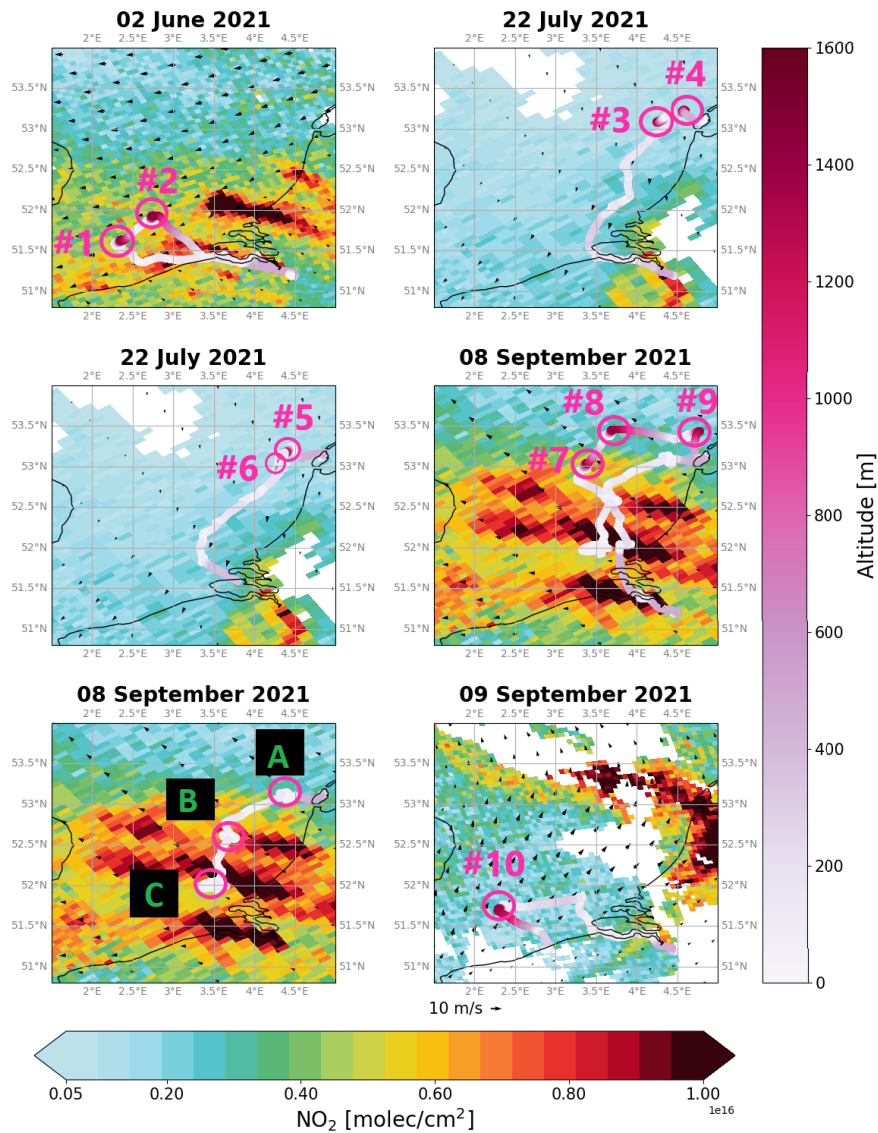


Figure 4.S.1: Overview of flight paths for ten vertical profiles (1-10) and three horizontal scans (A-C). Shown are the NO_2 column densities over the North Sea measured by TROPOMI at time of flight and the route of the aircraft as well as ECMWF 10 m wind indicated by the arrows.

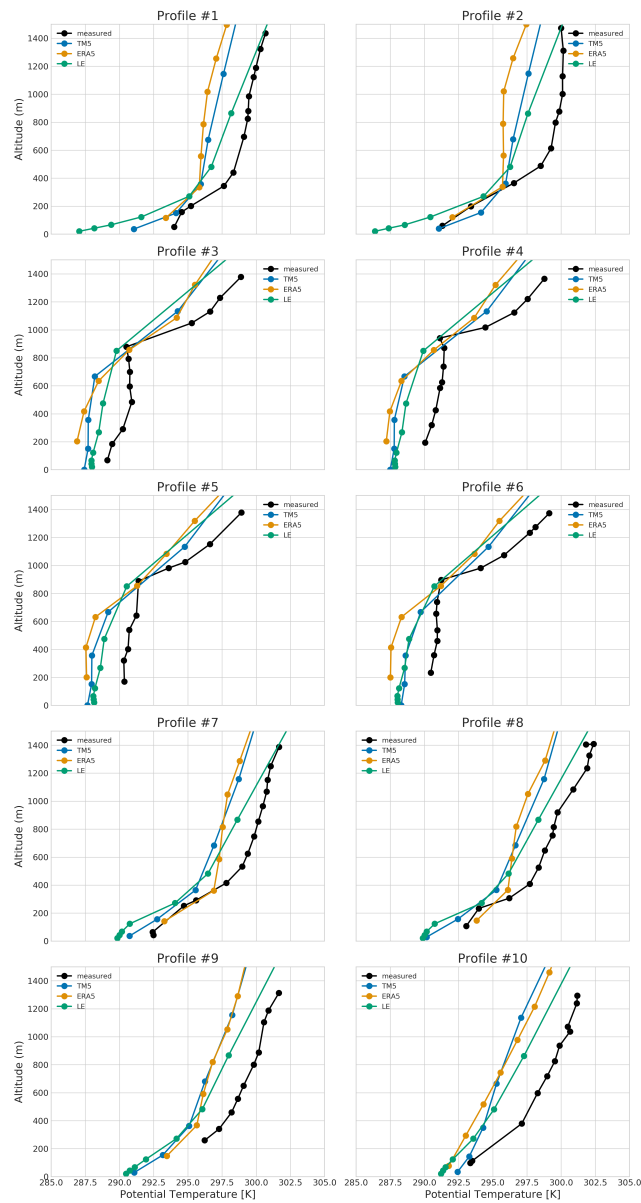


Figure 4.S.2: Profiles of potential temperature of all flights as well as coinciding, TM5, CAMS ensemble mean and LE 2x2 km profiles.

4.S.2 PARANOX

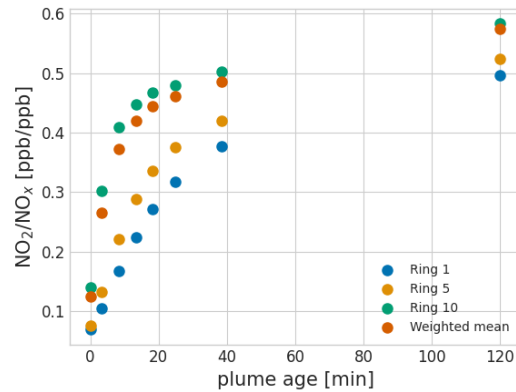


Figure 4.S.3: Relationship NO_2/NO_x ratio to binned plume age for PARANOX modeled plume for conditions during profile flight 1.

The PARANOX (PARAmetrization of Aircraft emitted NO_x) model is a Gaussian plume model based on work from Vila-Guerau de Arellano et al. (1990) which was successfully used for the parameterization of aircraft emissions and later for ship emissions Vinken et al. (2011). PARANOX simulates the dispersion of a plume in 10 concentric elliptical rings and the chemical evolution of the concentrations of several atmospheric trace gases inside it.

In this study we use the weighted average NO_2/NO_x ratio of the simulated aging plumes for neutral stability. For each of the 10 profiles a separate simulation was performed, using coinciding wind speeds from ERA5 reanalysis data, and background NO_x and O_3 values from the CAMS ensemble mean. For all simulations, the NO_x emission strength was set at 60 g/s with 94% of the NO_x emitted as NO in line with Eyring et al. (2005). Fig. 4.S.3 shows the simulated NO_2/NO_x ratio for profile 1.

4.S.3 Spatial representativeness of profile flights and other corrections

Table 4.S.1: Post-hoc calibration of the measured profiles.

Profile number	MUMM flight	date	Post-hoc NO ₂ calibration [ppb]
#1	21082	02.06.2021	3.28
#2	21082	02.06.2021	4.44
#3	21116	22.07.2021	-0.04
#4	21116	22.07.2021	-0.02
#5	21117	22.07.2021	-0.80
#6	21117	22.07.2021	0.02
#7	21133	08.09.2021	-3.96
#8	21133	08.09.2021	-3.28
#9	21133	08.09.2021	-3.60
#10	21135	09.09.2021	-2.68

1. **Calibration.** We apply a post-hoc calibration of NO₂ mixing ratio C to fit the CAMS ensemble mean in the clean troposphere (above 250 m). The corrected volume mixing ratio C_{mumm} is the sum of the measured mixing ratio C_{meas} and the post-hoc calibration b : $C'_{mumm} = C_{meas} + b$ with b such that $\overline{X(C_{meas} + b)}_{>250m} = \overline{X(C_{CAMS})}_{>250m}$ where X is the mass density. The table above shows our post-hoc calibration values b for every flight. The biases are largely consistent during each day, consistent with the daily calibration routine executed prior to flight. Given the spread of calibrations during individual days, we assume a uncertainty of the bias correction of 0.5 ppb (mean of standard deviations for days 02.06., 22.07., 08.09.).

2. **In-plume bias.** We expect any given area (e.g. a TROPOMI pixel) to be covered partially by a plume and partially by NO₂ background values. The coastguard flights actively approached ships and their plumes (see Sect. 2.2 in the main text) in order to measure the composition of the exhaust. Our measurements are therefore not necessarily representative of the mean concentrations over the entire area. Therefore, we use the representative random sampling of the transect flights to calculate an expected fraction of time in the plume for the lowest part of each profile (if it would have been a random sampling) and correct the observed measurements to fit this.

To tackle this, we use 20-sec interpolated AIS ship location and the wind speed measured by the aircraft to calculate plume locations. We define an area of interest (AOI) which spans a rectangle from the minimal to maximal longitude/latitude. We use plume locations based on the AIS information on ship position and speed for all ships sailing within a distance of 0.5° margin on both latitude and longitude within 2 hours before the end of an aircraft profile. We use the mean wind direction and wind speed measured by the aircraft below 100 m to predict the presence of the plumes from any ship within the vicinity. We apply the projection method that was introduced by Georgoulias et al. (2020): the location of a plume at time t is

simply the projected ship location from time t_0 multiplied with the wind speed in the direction of the wind: $x_{new} = x_{AIS} + u \cdot (t-t_0)$. We then count the number N of expected 20-second plume locations in the AOI and divide it by the size of the AOI A to calculate a so-called in-plume fraction $f_{AOI} = N_{AOI}/A_{AOI}$.

Taking the three transect flights (where the whole horizontal plane was sampled and therefore have a representative image of the area), we look at the NO_x measurements and define a measurement to be ‘in plume’ if the mixing ratio is at least 8 ppb above a background given by the 25 percentile of all NO_x measurements in said transects. Similar values for the threshold have been tested and yield similar results. From that we calculate the fraction of in plume measurements r as: $r_{AOI} = \frac{N_{inplume}}{N_{total}}$. Next, we assume a linear relationship between f and r and perform a linear fit through the three transect flights, yielding $r = c_1 * f + c_0$. This is our in-plume bias correction fit.

Using c_0 and c_1 from this fit, we calculate the expected r_{exp} for the lowest 100 m of each profile based on f_{AOI} . Additionally, we calculate the actual r_{actual} from the measured data. To correct each profile with the in-plume bias correction fit obtained from the transects we then calculate corrected mixing ratios $C''_{mumm,inplume} = C'_{mumm,inplume} * \frac{r_{exp}}{r_{actual}}$ for in-plume measurements and $C''_{mumm,noplume} = C'_{mumm,noplume} * \frac{1-r_{exp}}{1-r_{actual}}$ for measurements out of the plume. Since the linear fit is based on three data points only, we assume an uncertainty of 20% on r_{exp} .

To test the reliability, we additionally use a Monte Carlo simulation to calculate the NO_2 concentration in the lowest 50 m. Here, we first calculate the expected ratio of in/out of plume measurements r_{exp} as above and then draw random values of in/out of plume measurements of all profile flights, using the same in/out of plume definition as above. For every profile, we run 100 such simulations, using 20 measurements each (of which $20 * r_{exp}$ are drawn from the ‘in plume’ pool). The 100 results give a range of expected values for the lowest 50 m and are largely in line with the first approach. The Monte Carlo approach gives a more multi-flight representation of the in- and out-of-plume measurements and might thus overestimate or underestimate NO_2 concentrations in plumes of weak or strong emitting ships, respectively.

3. **NO_2/NO_x ratio correction.** The NO_2 values are not directly measured but taken from the difference between NO_x and NO . If the two are not measured completely simultaneously but slightly shifted, a (strong) gradient in NO_x concentrations will lead to strongly negative NO_2 if the aircraft is measuring NO while in the plume, and NO_x when outside the plume. Similarly, when NO_x is measured in-plume and NO out-of-plume, the NO_2 values will be overestimated in respect to the measurement interval. To avoid such extreme NO_2 values, we run the Gaussian

plume model PARANOX (see Supplemente S2) with NO_x and O_3 background levels similar to those during measurements. The modeled NO_2/NO_x ratios for rings 1 (center of plume), 5, 10 (edge of plume) and a volume weighted average for conditions during profile 1 are shown in Fig. 4.S.3.

All NO_2 measurements in-plume with negative NO_2 or NO_2/NO_x ratios above 0.8 are then corrected to fit the weighted modeled ratio according to the plumes age. We assume an uncertainty for the modeled NO_2/NO_x ratio of 0.1.

Fig. 4.S.4 shows the uncorrected and corrected profiles including the uncertainty estimate for the corrected data as well as the Monte Carlo simulations for the lowest level. Generally, the calibration bias dominates as can be seen from the offset between the gray and black bold lines. Below 150 m, the other corrections come into play, making the surface pollution more or less pronounced.

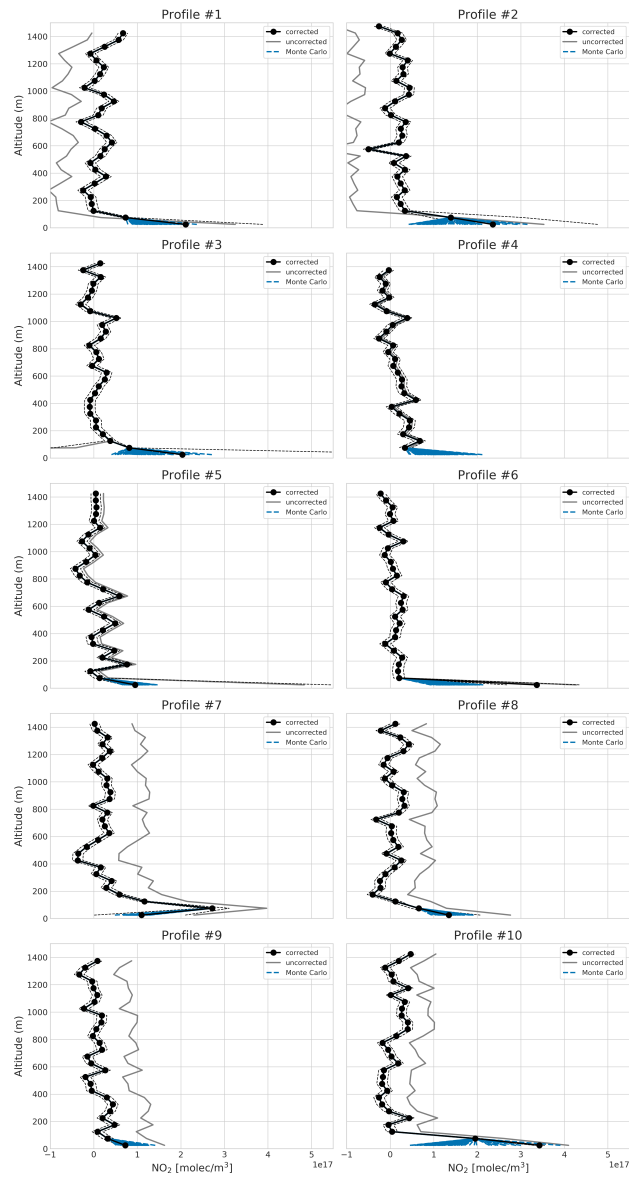


Figure 4.S.4: Aircraft measured profiles. The bold gray line shows the raw measured data, the bold black line the corrected profile, the dashed black lines the uncertainty estimate and the blue lines show the Monte Carlo simulation for the lowest 50 m.

4.S.4 Additional NO₂ maps

Fig. 4.S.5 shows the TROPOMI and modeled NO₂ column during the profile flights not presented in the main manuscript as well as the flight path and aircraft NO₂ measurements below 200 m.

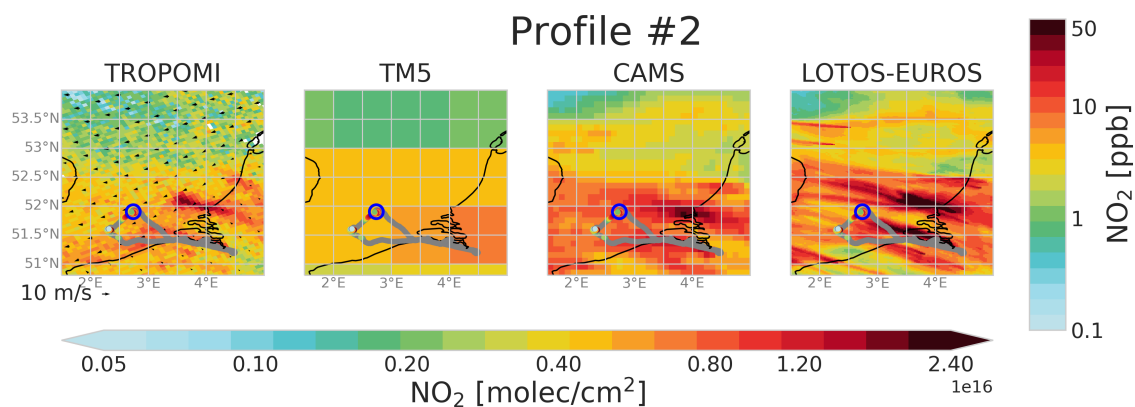
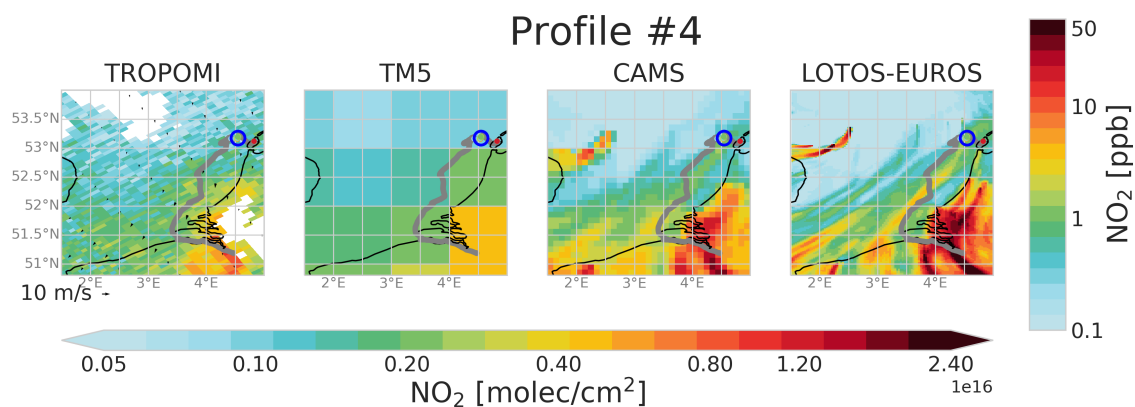
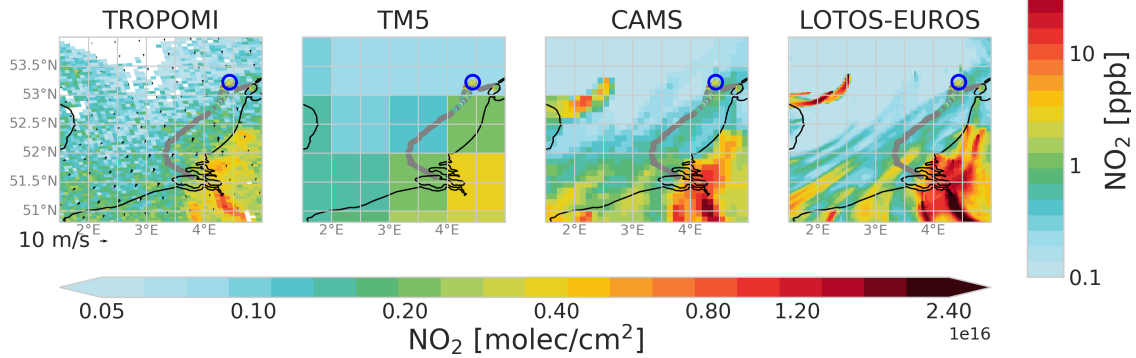


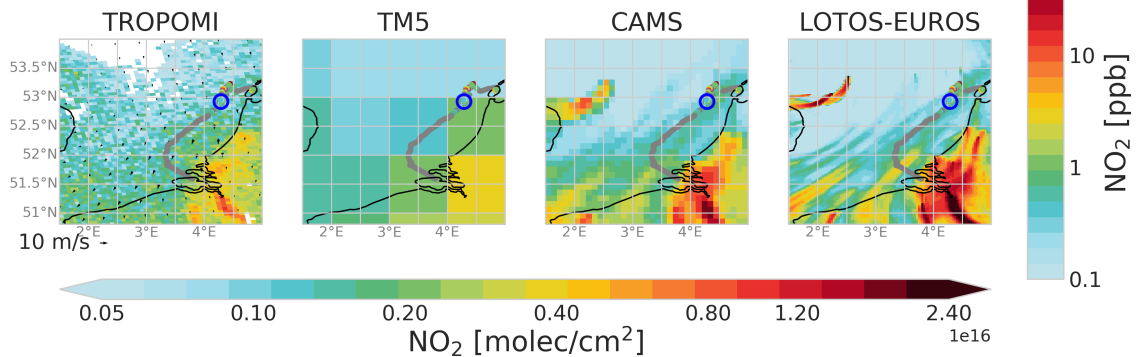
Figure 4.S.5: NO₂ columns as seen by TROPOMI and several model products for the time of the second profile measurement as indicated by the bottom colorbar. Overlaid are the aircraft measurements in grey for flights above 200 m and in colors below as indicated by the colorbar on the right as well as wind speed and direction by the arrows in the left panel.



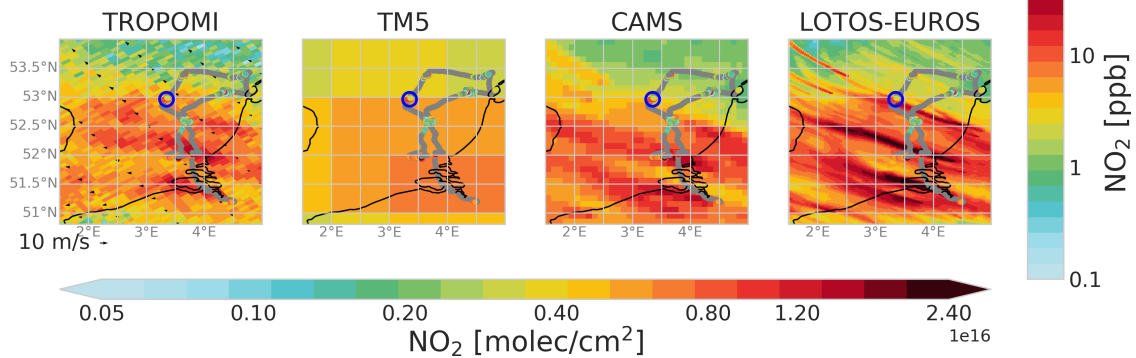
Profile #5



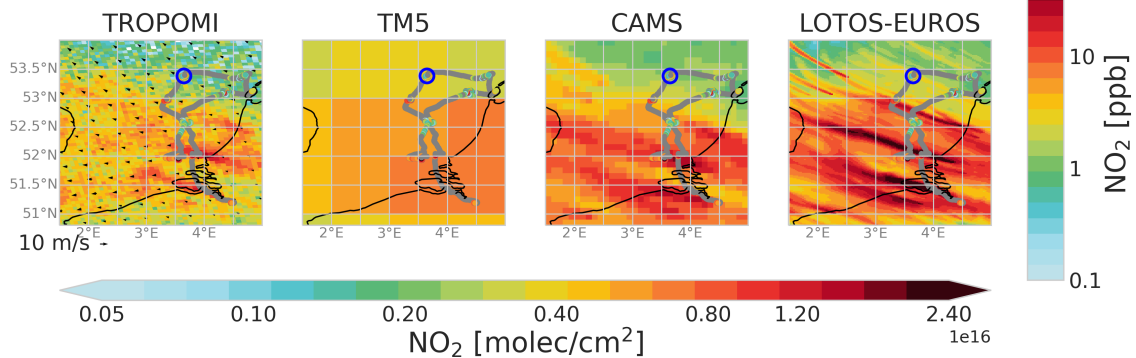
Profile #6

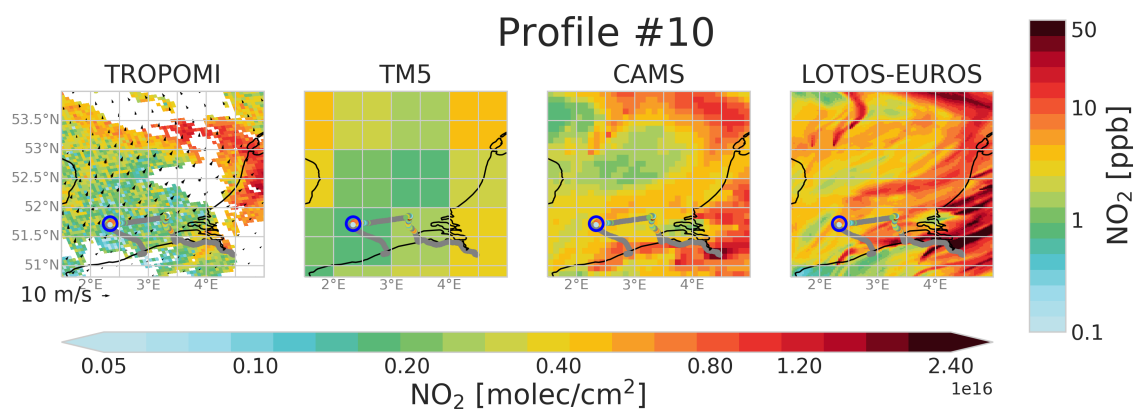
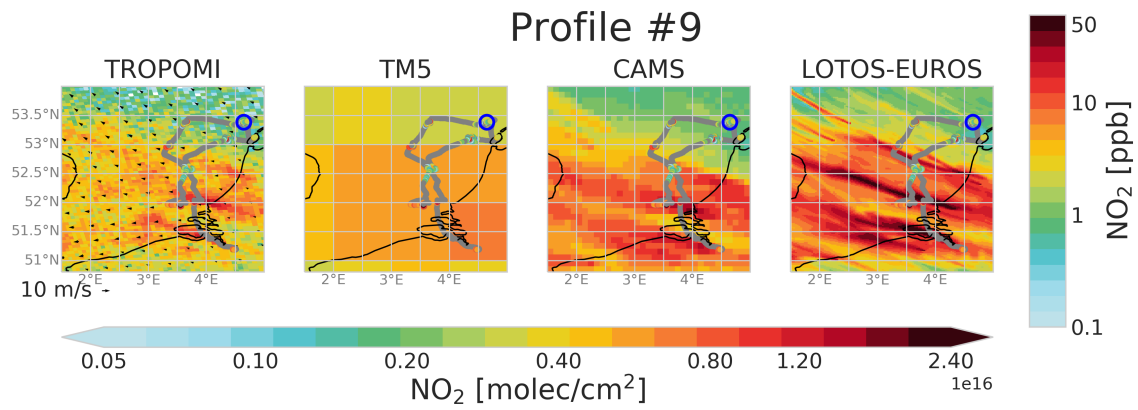


Profile #7



Profile #8





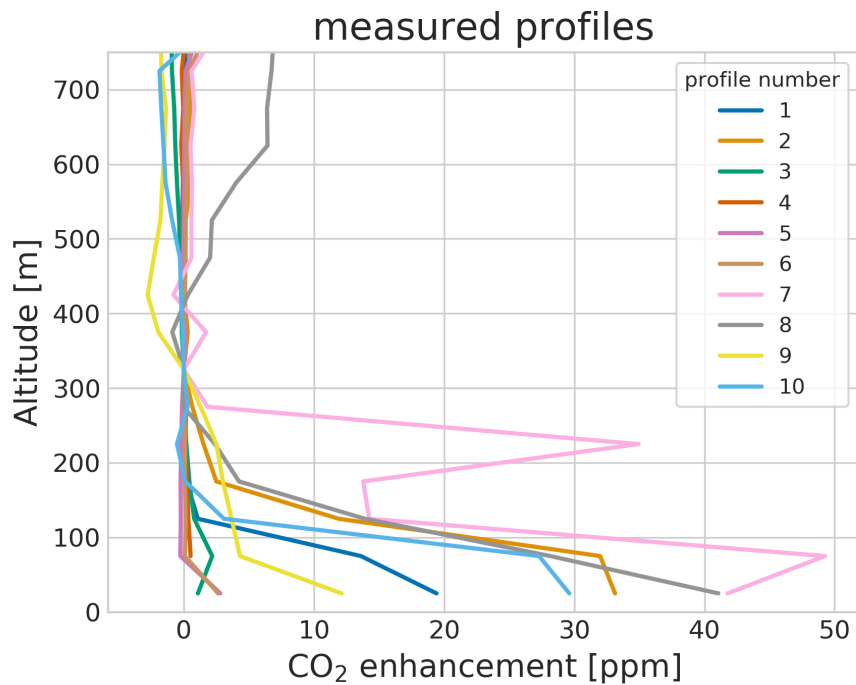
4.S.5 CO₂ profiles

Figure 4.S.6: CO₂ enhancement profiles for the ten spiral flights. As the CO₂ sensor was not calibrated, the 300-350 m value is subtracted to visualize the enhancement in the mixed layer.

Fig. 4.S.6 shows the CO₂ enhancements in the mixed layers measured for the ten profiles. The observed mixing layer height of ≈ 150 m agrees well with the observations in the NO₂ profiles. The CO₂ concentrations are measured directly with a temporal resolution of 1 Hz and no corrections are applied here.

4.S.6 Sea-land contrast

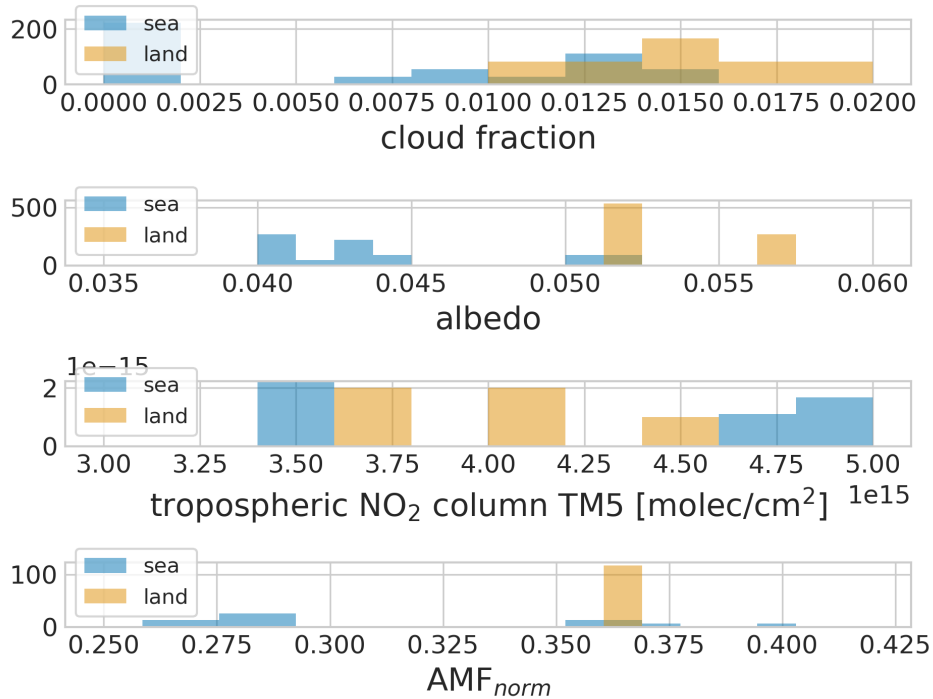


Figure 4.S.7: Normalized distributions of cloud fraction, surface albedo, TM5 a priori columns and normalized tropospheric AMF over land (yellow) and sea (blue).

We compare tropospheric AMFs over sea at the time and location of the aircraft measurements to AMFs over land at the Cabauw tower in the same TROPOMI orbits. For this analysis, we normalize tropospheric AMFs for viewing and solar geometry by dividing the topospheric AMFs provided in the TROPOMI files by the geometric AMFs. To make sure the observed differences are caused by the differences in surface properties, we include only scenes with similar a priori columns and cloudiness levels. Fig. 4.S.7 shows the distribution of cloud fractions, albedo, a priori columns and AMFs for the land and sea scenes. We find on average 20% lower tropospheric AMFs over the North Sea compared to land given similar overall retrieval conditions.

Chapter 5

Estimating NO_x emissions of individual ships from TROPOMI NO_2 plumes

This chapter is based on:

Riess, T. C. V. W., Boersma, K. F., Prummel, A., van Stratum, B. J. H., de Laat, J., and van Vliet, J. (2024). Estimating NO_x Emission of Individual Ships from TROPOMI NO_2 Plumes. <http://dx.doi.org/10.2139/ssrn.4858709>. under review

Abstract

Maritime transportation is a substantial contributor to anthropogenic NO_x emissions and coastal air pollution. Recognizing this, the International Maritime Organization (IMO) has steadily implemented stepwise stricter emission standards for ships in recent years. However, monitoring emissions from sea-bound vessels poses inherent challenges, prompting the exploration of satellite observations as a promising solution. Here we use TROPOMI measurements of NO_2 plumes together with information on ship position and identity, and atmospheric models to quantify the NO_x emissions of 130 plumes from individual ships in the eastern Mediterranean Sea in 2019. Because most of the emitted NO_x is in the form of NO , which is not immediately converted into detectable NO_2 , plumes show their NO_2 maximum some 15-30 km downwind of the ship's stack. Further downwind NO_2 decreases because of plume dispersion and photochemical oxidation. Background ozone and wind speed play a significant role both in detectability of the NO_2 plume and the relationship between NO_x emissions and observed NO_2 , explaining the good detection conditions in the eastern Mediterranean summertime, where ozone levels are high. Taking such effects of emissions, dispersion, entrainment, and in-plume chemistry in full account, we find emission strengths of 10-317 $\text{g}(\text{NO}_2) \text{ s}^{-1}$. We then calculate emission factors of the detected ship plumes using AIS and ship specific data and find that newer Tier II ships have higher emission factors compared to older Tier I ships. This is especially the case when running at lower engine loads, which is the most frequently observed mode of operation in our ensemble. Additionally, at the time of detection around half of the emission factors detected for Tier II ships lie above the IMO weighted average limits. The presented method sets the stage for automated ship emission monitoring at sea, contributing to better air quality management.

5.1 Introduction

International shipping is responsible for the transport of 90% of traded goods worldwide (OECD, 2024) and for contributing 13 % of anthropogenic nitrogen oxides ($\text{NO}_x = \text{NO} + \text{NO}_2$) emissions (International Maritime Organization, 2015). NO_x plays an important role in atmospheric chemistry by influencing the formation of tropospheric ozone, a potent greenhouse gas and air pollutant, and producing hydroxyl radical (OH), the main oxidant in the atmosphere. It therefore perturbs the oxidative capacity of the atmosphere, impacting the lifetimes of many pollutants and greenhouse gases. In the marine boundary layer, nitrogen oxides interact with halogen chemistry (e.g. Elshorbany et al. (2022)) which is not included in most current models. Accurate knowledge of anthropogenic NO_x emissions in maritime regions is crucial to better understand atmospheric chemistry in the marine boundary layer. NO_2 is a toxic gas itself, and it leads to the formation of aerosols, causing and worsening human respiratory diseases and damaging plant growth. Nitrogen deposition associated with ship emissions further worsens the issue of eutrophication, particularly in offshore regions (Neumann et al., 2020). Overall, NO_x emitted by ships degrade air quality in many coastal areas with serious consequences for human and ecosystem health, causing up to 800000 premature deaths per year (Sofiev et al., 2018). To mitigate the health impact of the growing maritime transport, the International Maritime Organization (IMO) has implemented standards for NO_x emissions by ocean-going vessels (OGV) in the MARPOL Annex VI, grouped in three engine tiers: emissions from ships build in or after the year 2000 have to follow Tier I limits, stricter Tier II limits apply to ships build after 1 January 2011 and the most stringent NO_x regulations are required from ships with a keel-laying date of 2016 or later sailing in so-called Emission Control Areas (ECAs). Ships build prior to 2000 are unregulated. However, there is insufficient understanding and verification on how ship characteristics and operations such as size, age, speed, and engine load influence the emissions. Recent studies indicate that while SO_x emissions of ships are decreasing, NO_x emissions of OGVs are not decreasing as anticipated (Van Roy et al., 2023; Comer et al., 2023), highlighting the problem at hand. Those studies require the availability of in-situ or ground-based NO_x measurements and are therefore spatially limited. Satellite remote sensing of NO_2 , on the other hand, offers a global perspective.

Nitrogen oxides are emitted in combustion processes such as in ship engines, mainly as nitrogen monoxide (NO). Conversion of NO to nitrogen dioxide (NO_2) requires the presence of sufficient O_3 in the plume. This conversion takes some time, especially when O_3 is titrated in the early stages of the plume (Vinken et al., 2011), so that NO_2 columns are expected to have their highest magnitudes several kilometers downwind of the ship source. Whether NO_2 columns originating from individual ships are actually detectable from space depends on the strength of the NO_x emissions, and on the $\text{NO}_x:\text{NO}_2$ ratio in the plume. This ratio is driven by the chemical regime, i.e. the abundance of O_3 , radiation and atmospheric dynamics within and near the plume. Therefore, accurate knowledge of

these parameters is needed to interpret satellite observations of NO_2 column plumes in terms of the underlying NO_x emissions from individual ships.

The TROPOMI instrument on board of the Sentinel-5p satellite for the first time allowed detection of NO_2 plumes from individual ships with satellites (Georgoulias et al., 2020). While previous studies with the GOME, SCIAMACHY, and OMI sensors (Beirle et al., 2004; Richter et al., 2004; Vinken et al., 2014a) identified major shipping lanes and studied trends in maritime NO_2 , those satellite instruments were incapable of detecting single ship NO_2 plumes. The higher spatial resolution ($3.5 \times 5.5 \text{ km}^2$) NO_2 measurements of the European TROPOMI-sensor therefore sparked research into automated (ship) plume detection (Kurchaba et al., 2022; Finch et al., 2022; Kurchaba et al., 2023).

Here we combine tropospheric NO_2 columns from TROPOMI and coinciding data on ship position and identity from the automatic identification system (AIS) to attribute NO_2 plumes to individual large ships. We investigate the main chemical and meteorological drivers of the relationship between a ship's NO_x emission strength and its observable NO_2 plume. Using the observed NO_2 plumes we then quantify the NO_x emission strength of 130 OGVs in the eastern Mediterranean Sea in 2019 via an inversion framework accounting for atmospheric dynamics and in-plume NO_x -VOC- O_3 -chemistry. Finally, we derive emission factors and evaluate these across ship's engine tiers and IMO's regulatory standards.

5.2 Methods & Materials

5.2.1 Data

5.2.1.1 TROPOMI NO_2 data

TROPOMI (Veefkind et al., 2012) is a polar sun-synchronous spectrometer on board of Sentinel-5Precursor (S5P), launched in 2017. S5P crosses the equator at 13:30 local time, and covers the earth daily, with overlapping orbits at the Mediterranean. The spatial resolution is $5.5 \times 3.5 \text{ km}^2$ for nadir pixels, and $5.5 \times 14 \text{ km}^2$ for pixels at the edge of TROPOMI's swath. Here we use TROPOMI L2 tropospheric NO_2 vertical columns densities of version 2.4 over the eastern Mediterranean (31°N - 36°N , 18°E - 30°E , see Fig. 5.1) between April and November 2019, freely available via the Copernicus Data Space (<https://dataspace.copernicus.eu>). Only pixels with $qa \leq 0.75$ are used throughout the manuscript, therefore excluding cloudy pixels. The retrieval of tropospheric NO_2 consists of 3 steps: (1) the spectral fitting of modelled to observed reflectance spectra, providing the NO_2 slant column density, (2) the subtraction of the stratospheric contribution from the NO_2 slant column, resulting in the tropospheric NO_2 slant column estimate, and (3) the application of the tropospheric air mass factor (AMF) to correct for the effects of satellite viewing geometry and reduced vertical sensitivity to NO_2 , which provides the tropospheric vertical column. For the calculation of the AMF, a priori information of the

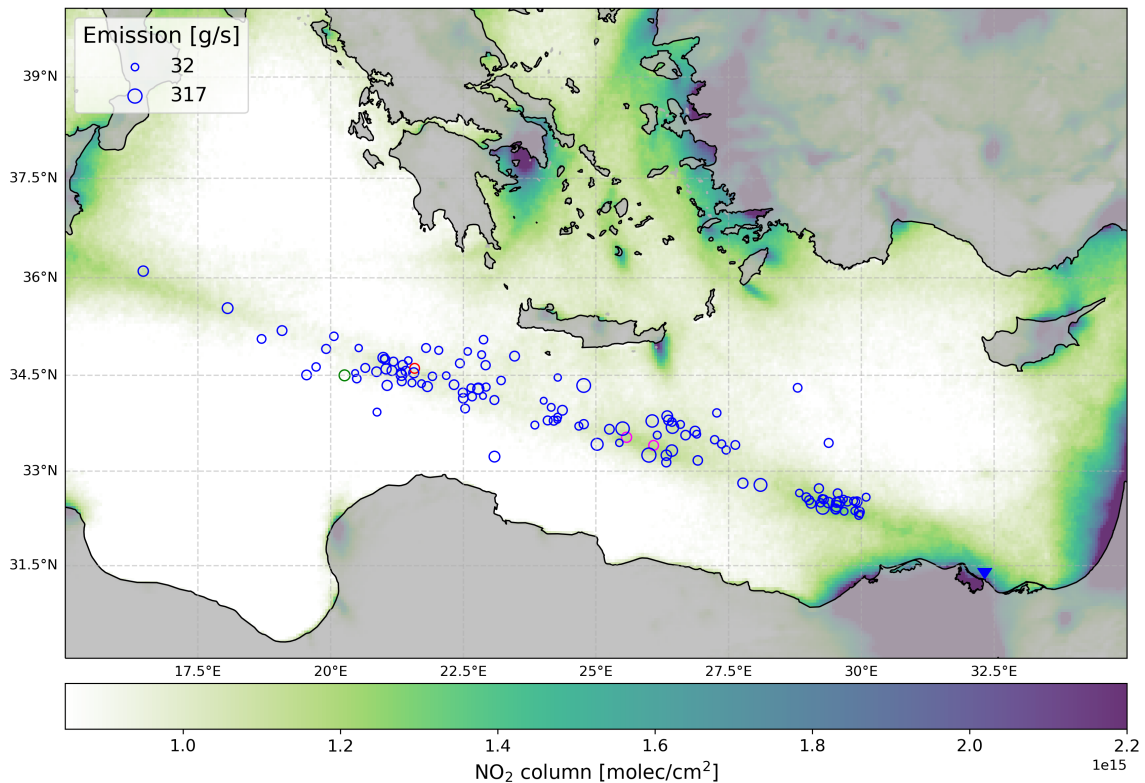


Figure 5.1: Map of vertical tropospheric NO_2 columns over the eastern Mediterranean in April–November 2019. We averaged cloud-free TROPOMI v2.4 data with $qa \leq 0.75$. The blue circles indicate the position of the ships of which the plume was detected and their size is proportional to the estimated emission flux causing the plume. The blue triangle indicates the entrance of the Suez Canal. Land surface is indicated by the darker color. Circles in other colors indicate plumes shown throughout the manuscript (red: Fig. 5.3, green: Fig. 5.5, magenta: Fig. S10)

vertical distribution of NO_2 is needed. The a priori NO_2 profile in the TROPOMI retrieval is modelled with TM5, at a resolution of $1 \times 1^\circ$, smearing out ship plumes and leading to an underestimation of in-plume NO_2 columns. We therefore multiply all in-plume pixels with a factor of 1.2 to correct for low bias caused by a priori vertical profiles, motivated by forward model simulations of ship plumes presented in Supplement A, and in line with findings in (Riess et al., 2023). For the emission estimates (see subsection 5.2.3.1), we use these adjusted TROPOMI NO_2 columns to calculate the NO_2 enhancement of the 3h-long ship plume over the background as presented in detail in supplements B and C.

5.2.1.2 AIS and ship specific data

To relate the TROPOMI NO_2 plumes to shipping activity, we use data from the automatic identification system (AIS) for shipping. Since 2005, the International Maritime Organization (IMO) requires all ships with a gross tonnage over 300 and all passenger

ships to carry an AIS transponder. These transponders broadcast static (e.g., identity) and dynamic (e.g., position, speed, course) information of the ship, which can be received by other ships, shore stations, and satellites (International Maritime Organization, 2014). Here we use historical AIS data available to the Dutch Human Environment and Transport Inspectorate (ILT). Furthermore, we retrieve information on ship dimensions, design speed and engine power from Lloyds List Intelligence and the official ship registrations (<https://gisis.imo.org/Public/Default>). We use the IMO number as primary key, and MMSI when IMO number is not available in AIS data. For 7 ships, information on the maximum engine power was missing. For these 7 vessels, we used a regression to predict engine power based on design speed, length, keel laying year and ship type, using the registration of 40550 ships as training data.

5.2.1.3 Atmospheric composition and meteorological data

To interpret the observed NO_2 observations from TROPOMI we need to know the chemical and meteorological conditions under which the NO_x emissions occurred. For that, we use NO , NO_2 and O_3 data for levels 0 m, 50 m and 250 m from the European Air Quality Reanalysis ensemble median data set provided by the Copernicus Atmospheric Monitoring System (CAMS) to determine the NO_x and O_3 background for observed ship plumes and use those for model forward runs and as features in the emission estimations. The data is available via the atmospheric data store (<https://ads.atmosphere.copernicus.eu>). Furthermore, we use wind information at 10 and 100 m from ERA5 reanalysis data as well as temperature data from the lowest 3 levels to determine the stability class via the temperature gradient, all available via the CAMS climate data store (<https://cds.atmosphere.copernicus.eu>).

5.2.2 Models

To estimate NO_x emissions from individual ship plumes visible in the TROPOMI data, we use two atmospheric models - PARANOX and microHH - and the machine learning tool XGBoost introduced in the following.

5.2.2.1 PARANOX

PARANOX (PARAMetrisation of NOX) is a Gaussian Plume model (GPM) based on work from Vila-Guerau de Arellano et al. (1990) first used to model aircraft emissions (Meijer, 2001) and later adjusted and used for ship plumes (Vinken et al., 2011; Riess et al., 2023). PARANOX accounts for entrainment and detailed chemistry (43 species in 98 reactions) in 10 concentric, elliptical half-rings. The rings expand over time with plume growth - enabling entrainment of fresh air- depending on wind speed and Pasquille Stability Classes, using the dispersion parameters for over sea as proposed by Hanna et al. (1985). PARANOX is capable of simulating ship NO_2 plumes to reasonable accuracy as has been

validated using aircraft observations from Chen et al. (2005) (Vinken et al., 2011). Here, we use PARANOX to create a library of plumes, varying the following forward parameters: NO_x emission strength in g/s (95% of NO_x are released as NO), NO_x and O_3 background concentrations, Pasquille stability class, wind speed, ship speed, the angle between ship and wind direction, coordinates of the ship and date of emission. For each library member we compute several simulations with PARANOX, emitted every 15 minutes and create a superposition of these plumes, representing the NO_2 at 13:30 local time as it would be observed by TROPOMI. In supplement D the PARANOX model is introduced in more detail and supplement E shows the plume library.

5.2.2.2 MicroHH

We use MicroHH for a validation of the PARANOX model. MicroHH is an open-source, fluid dynamics model, supporting Large Eddy Simulations (LES) (Van Heerwaarden et al., 2017). LES are used to model turbulent flow which is characterized by a wide range of length and time scales. In LES, the goal is to capture the larger, energy-containing eddies explicitly, while parameterizing the effects of the smaller, dissipative eddies. Recently, non-linear atmospheric chemistry has been added to MicroHH, enabling the modeling of NO_x plumes (Krol et al., 2024). The simulations used in this research have been performed at a horizontal resolution of $50 \times 50 \text{ m}^2$ with 128 vertical layers. The boundary conditions for the MicroHH simulations are from ERA5 and CAMS data as in van Stratum et al. (2023). Emissions are simulated as moving Gaussian blob with $\sigma_x = \sigma_y = \sigma_z = 25 \text{ m}$ at $z=50 \text{ m}$ (Ražnjević et al., 2022). In the comparison to PARANOX we find a good agreement for the dominantly stable conditions observed in the marine boundary layer, especially when compared at TROPOMI resolution. The NO_2 plumes simulated by MicroHH are slightly wider and show a higher NO_2 peak but are shorter lived. Using a plume simulated with MicroHH as input in the PARANOX-based inversion framework presented in Section 5.2.3.1 results in an emission estimate that is 10 % higher than the MicroHH input. More details on the model comparison can be found in supplement F.

5.2.2.3 XGBoost model

To infer NO_2 emission estimates of individual ships based on TROPOMI NO_2 data, and meteorological and chemical conditions from ERA5 and CAMS (see 5.2.1), we use three XGBoost models fed with the PARANOX plume library: For Pasquille stability classes neutral (D), slightly stable (E) and stable (F) a separate XGBoost model is trained to learn the relationship between NO_x emission and TROPOMI NO_2 observation given the forward parameters using the python package xgboost (Chen and Guestrin, 2016). To that end we calculate the 3h plume enhancements E_{3h} as the NO_2 enhancement caused by the ship in the three hours before observations relative to the background NO_2 (see supplement B for details) for all library members and detected plumes. The other fea-

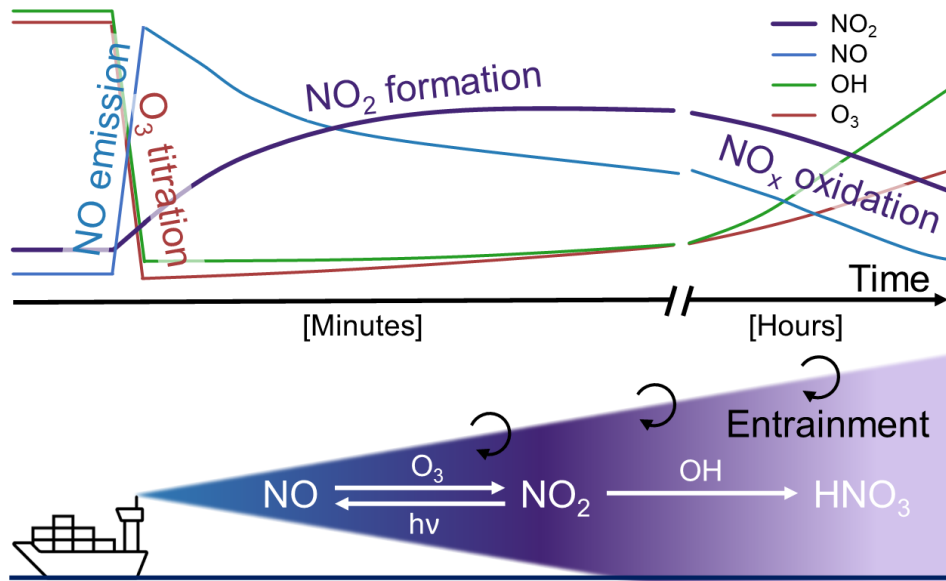


Figure 5.2: Schematic of ship plume chemistry: The vast majority of NO_x is emitted by the ship in the form of NO . The fresh plume is therefore dominated by NO . NO_2 formation requires ozone, which is titrated in the early stages of the plume. The peak in NO_2 becomes visible only a few kilometers downwind, when ozone increases through the entrainment of ozone-rich background air. NO_2 can be oxidized by OH , the major sink of NO_x . Like ozone, OH is depleted in the fresh plume but entrained in the aging plume.

tures used in the XGBoost model are the forward parameters introduced in 5.2.2.1, of which ozone and wind speed turn out to have the highest impact (see supplement G). We use 60% of the plume library members for training and 20% for tuning of the hyperparameters. Using a grid search scored on the negative mean squared error, we find the optimal hyperparameter settings for the models, as `n_estimator=3000`, `max_depth=5`, `learning_rate=0.02`, `subsample=0.8`, `colsample_bytree=0.9` for all three stability classes.

The remaining 20% of the data are used to evaluate the performance of the models. The average relative error of the models are 10%, 9% and 8% for classes D, E and F, respectively. For emission fluxes above 80 g/s, the average relative error reduces to 4%, 4% and 3%, respectively. We see that the model is successful in reproducing accurate predictions of the emission flux for our data set. For more details, see supplement B, C, and G.

5.2.3 Data processing

5.2.3.1 Emission estimates

We quantify NO_x emissions from the observed NO_2 plumes using a three-step approach: first, we generate a library of 40,000 virtual NO_2 plumes for a wide variety of NO_x emission strengths and forward model parameters with the Gaussian Plume model PARANOX (see

above). PARANOX agrees well with MicroHH (see supplement F), and is computationally fast. The forward parameters included in our inversion are the relevant chemical and meteorological conditions at the time of measurement such as background ozone, wind speed, stability class (i.e. entrainment strength) and radiation in the marine boundary layer. The model explicitly accounts for the effects of O₃ titration and entrainment, as well as NO₂ formation and oxidation as depicted in Figure 5.2. Thus, the PARANOX-library shows that NO₂ plume magnitudes increase with NO_x emission strength, and with boundary layer O₃ concentrations (see supplement E).

In step 2, the plume library is fed to an XGBoost model (see subsection above) to learn the relationship between NO₂ plume magnitude and underlying NO_x emission strength given the environmental conditions under which the observations were done. Our XGBoost model is capable of reproducing strong NO_x emission strengths from simulated NO₂ plumes with an average relative error below 5%.

In step 3, we apply our XGBoost model on the observed TROPOMI NO₂ plumes in combination with the actual environmental conditions during emission (obtained from CAMS and ERA5 data, see 5.2.1, and supplement B) and quantify the NO_x emissions.

5.2.3.2 Uncertainty estimation

Several uncertainties of the estimated emission fluxes need to be considered: (a) uncertainties introduced by the model choice, (b) the XGBoost model, and (c) the uncertainty of features of the XGBoost models. The uncertainty (a) introduced by the use of a Gaussian Plume model can be estimated by inverting a MicroHH simulation. This has been done using the MicroHH simulation of the plume discussed in supplement F, resulting in an emission flux that was 10% higher than the input. (b) has been determined using part of the training data as discussed in 5.2.2.3. Uncertainty (c) is harder to assess. The plume enhancement is by far the most important feature of the XGBoost model (see supplement G). Within its calculation the determination of the background level, the ship sector (mainly due to wind information) and lastly the TROPOMI retrieval itself are the largest uncertainty factors. We use three different definitions of the background and ship sector, using five different methods to calculate plume enhancements (see supplement B). The standard deviation of inversions of these 5 enhancements combined with the presented PARANOX and XGBoost uncertainties make up the total uncertainties. Uncertainty (c) is on average about 20% of the estimated emission flux and therefore the dominating uncertainty for most plumes.

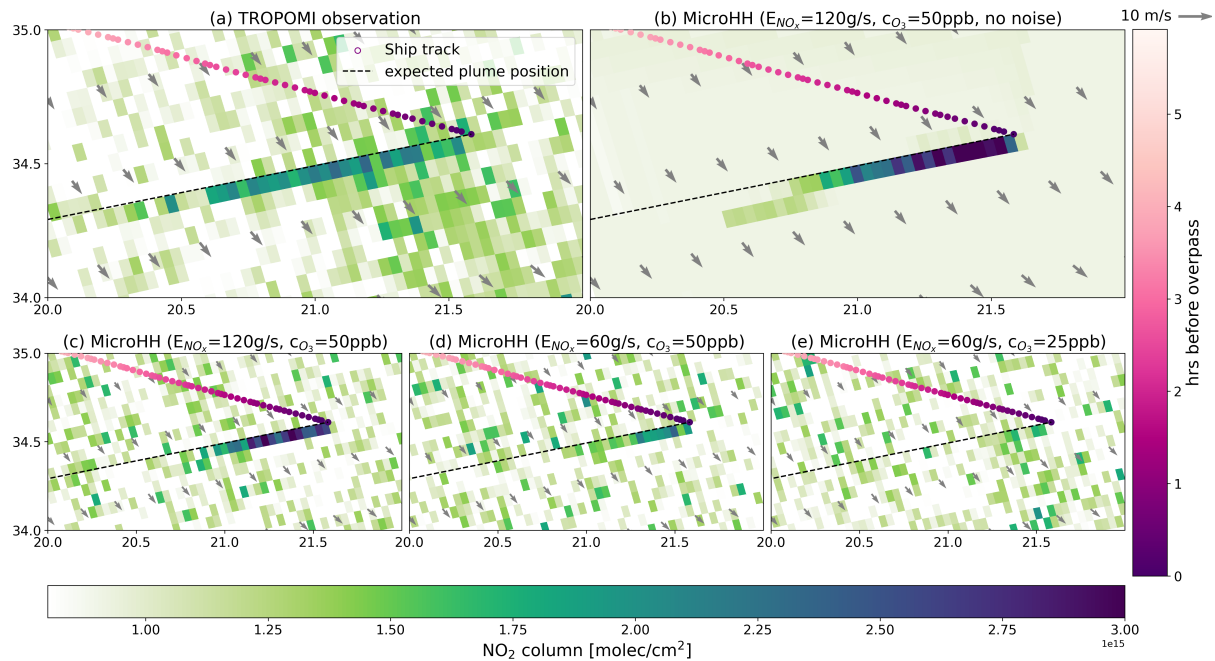


Figure 5.3: NO_2 plume of a ship sailing in the Mediterranean on 5 June 2019. (a) Observed TROPOMI data, (b) MicroHH simulation resampled at TROPOMI grid using an emission strength of 120 g s^{-1} , (c) like (b) but with realistic instrument noise, (d) like (c) but with half the emission flux and (e) like (d) but with half the background ozone. The pink dots indicate the AIS position of the ships in the three hours before overpass at 11:42 UTC as indicated by the right color bar, the black dashed line shows the wind-shifted track, which is the ship location displaced by the distance $d = \vec{v}_{wind} * \delta t$ as in Georgoulas et al. (2020). \vec{v}_{wind} is the 10 m wind speed and δt the time between AIS time of the ship position and TROPOMI overpass. The choice for halving the emission is motivated by the emission limit of Tier III ships (see supplement I), the choice for halving the background ozone is motivated by the difference in ozone levels between the eastern Mediterranean and the North Sea. The NO_x emissions given here and reported elsewhere in this manuscript are given as mass of NO_2 , meaning that an emission of 46 g s^{-1} equals 1 mol s^{-1} .

5.3 Results

5.3.1 Nitrogen dioxide hotspots over the eastern Mediterranean

We identified 130 clear ship plumes in the Eastern Mediterranean ($31.^\circ\text{N}$ - $36.^\circ\text{N}$, $18.^\circ\text{E}$ - $30.^\circ\text{E}$) in daily TROPOMI NO_2 columns (see methods section for data selection) between April and November 2019. Figure 5.1 shows the 9-month average NO_2 map with regionally strong NO_2 pollution around Athens, Alexandria and Israel, originating from urban activity and industrial emissions with average columns of up to $2.2 \times 10^{15} \text{ molec cm}^{-2}$. The map also reveals the effect of NO_x emissions from the Atherinolakkos power plant at the south-east of Crete. Figure 5.1 shows enhanced NO_2 columns of up to $1.2 \times 10^{15} \text{ molec}$

cm^{-2} between the Suez Canal and Sicily, representing dense traffic in this shipping lane. Most NO_2 plumes detected between April and November 2019 are found near this shipping lane.

5.3.2 Nitrogen dioxide plume attribution to single ship

Figure 5.3 (a) shows an NO_2 plume originating from a container ship sailing in south-easterly direction on 5 June 2019. The prevailing northwesterly winds predict the pollution from the container ship to be displaced to the southeast, which is exactly where TROPOMI imagery shows an NO_2 column plume. Such close-up analysis has been done for all ships, boundary layer wind and TROPOMI NO_2 data available to us. Whenever the predicted location of the pollution enhancement coincided with a single TROPOMI NO_2 plume, the observed NO_2 enhancement was attributed to that particular ship. Through careful manual inspection, we thus identified a total of 130 plumes between April and November 2019 caused by 119 different vessels, with 9 vessels being observed twice and one three times. Occurrences of plumes in the vicinity of multiple wind-shifted ship-tracks were excluded to select unperturbed single ship plumes.

5.3.3 Characteristics of environmental conditions and observed ships

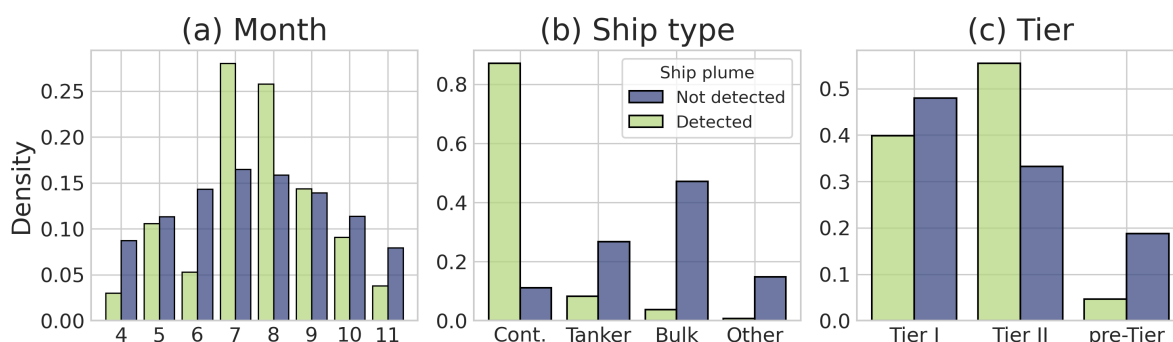


Figure 5.4: (a) Percentage of plumes in each month, (b) type and (c) tier of ships. Plumes that were detected are shown in light green and not detected ships in the studied area and time are colored in purple.

We compare characteristics of ships whose plumes were detected and the environmental conditions in which they sailed to those that were not detected in Figure 5.4 and Table 5.1. 85% of all ships detected are container ships that are some two times longer and six times heavier than the fleet average. Additionally, detection is most likely for ships sailing at speeds above 5 m s^{-1} . This is in line with the 10 knots limit found in (Kurchaba et al., 2024), and the ships exposed here represent the fast-travelling end of the fleet. Additionally, ozone abundance in the marine boundary layer is positively correlated with the detectability of ship plumes, which is further discussed in Section 5.3.4. Furthermore, we found that ship NO_2 plume detection is most likely in summer months over the eastern

Table 5.1: Median and interquartile range (in brackets) of conditions and ship characteristics of detection and non-detection of ship plumes in the eastern Mediterranean during TROPOMI overpasses in April–November 2019. O_3 and NO_x come from coinciding CAMS data, wind from ERA5 and ship characteristics from ship registrations matched with AIS data.

	Not detected	Detected
O_3 [ppb]	48.7 (8.4)	53.1 (5.8)
NO_x [ppt]	246 (185)	401 (125)
Ship length [m]	183 (110)	366 (76)
Ship speed [m s^{-1}]	5.9 (2.8)	9.0 (2.2)
Gross tonnage	24800 (40800)	145600 (81200)
Wind speed [m s^{-1}]	4.9 (3.2)	5.6 (2.2)

Mediterranean, with most plumes being detected in July and August. In summer, fewer clouds are present over the Mediterranean than in other seasons, which favours satellite detection of NO_2 plumes in the marine boundary layer. The detected ships have more recent keel laying years and larger engines than typical for the Mediterranean fleet. This shows that our ship ensemble consists mostly of Tier II rather than Tier I ships compared to the Mediterranean fleet as a whole (Fig. 5.4 (c)). Lastly, we find that ship plume detection is more likely with tailwind than with headwind (see Supplement H).

5.3.4 Emissions and ozone levels drive observable plumes

While NO_2 plume detection is most likely when a ship’s NO_x emission is strong and skies are clear, there are other factors that influence the detectability of a ship NO_2 plume. Figure 5.3(a) shows the NO_2 plume from a large container ship (370 m) sailing at high speed (10 m/s) in southeasterly direction on 5 June 2019 over the Mediterranean, as observed by TROPOMI. We performed a high-resolution simulation of the pollution plume from this large container ship with the turbulence and chemistry-resolving model MicroHH (van Stratum et al., 2023; Krol et al., 2024). Assuming that the ship emitted NO_x at a rate of 120 g s^{-1} , MicroHH simulates the ship’s NO_2 plume at the same location and of similar magnitude as TROPOMI, providing confidence in our attribution method. We then evaluated to what extent emission strength and ozone entrainment drive the detectability of the ship’s NO_2 plume. Figures 5.3(c)–(e) illustrate that reducing emissions by 50% and reducing boundary layer O_3 concentrations by 50% in the MicroHH simulations leads to lower detection probability, as evident from the much less prominent NO_2 column enhancement (see Methods) in the simulations.

Figure 5.2 shows the main processes that determine the evolution of the NO_x plume, explaining the TROPOMI NO_2 plumes. Ships emit NO_x mainly as NO. However, because of ozone titration in the fresh plume, the conversion of NO to NO_2 is delayed. Once sufficient ozone is entrained into the plume, the reaction of NO with ozone forms NO_2 , making the NO_2 peak detectable several kilometers downwind of the ship’s stack. During the day,

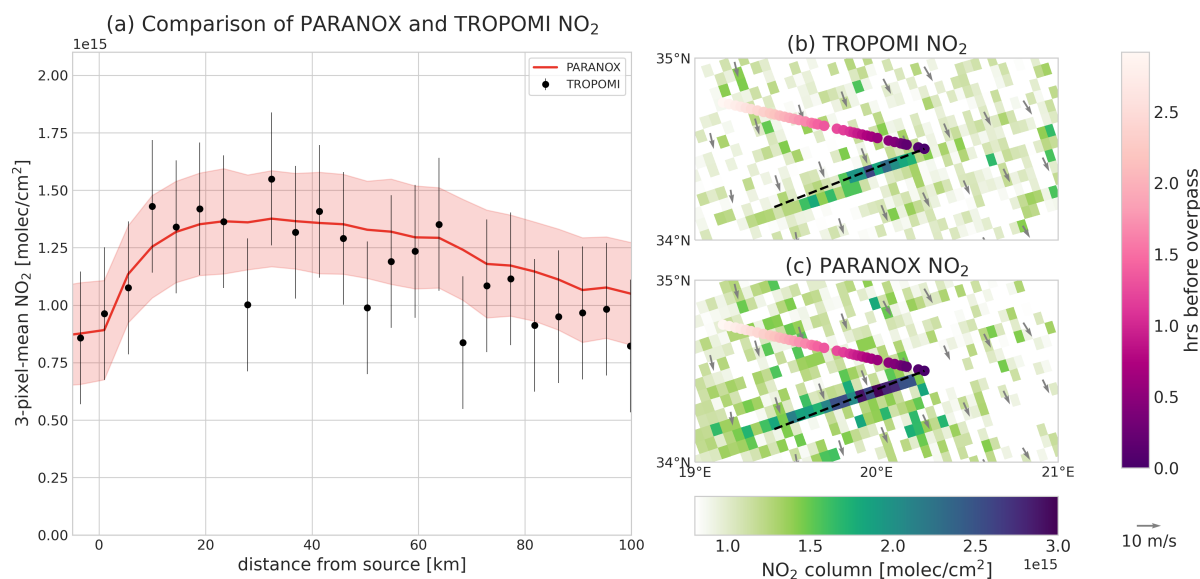


Figure 5.5: Comparison of observed and modeled ship plume on 11 July 2019. (b) The ship plume observed by TROPOMI, which our inversion framework estimates to be caused by an emission flux of 130 g s^{-1} . (c) The plume simulated by PARANOX, using this inverted flux and forward parameters from CAMS and ERA5. The pink dots in both plots indicate the ship position in the 3 hours before overpass, the black dashed line show the expected plume position. (a) NO₂ columns of the ship plume as a function of distance from the stack. In black, the 3-pixel-wide mean of TROPOMI data is shown. In red, the 3-pixel-wide mean of mean simulated observation is shown. The uncertainty band shows the standard deviation of 500 applications of realistic TROPOMI noise, resulting from the single-pixel uncertainty of $0.3 \cdot 10^{15} \text{ molec cm}^{-2}$ as well as spatially correlated uncertainty of $0.15 \cdot 10^{15} \text{ molec cm}^{-2}$ caused by the albedo database.

NO₂ can undergo photolysis and form back NO and O₃. NO_x cycling is terminated by reaction of NO₂ with the hydroxyl radical (OH) to form HNO₃. Both the rate at which NO₂ is formed as well as the ratio between NO and NO₂ thus depend on the abundance of ozone in the plume. With its high summertime ozone levels in the marine boundary layer, the eastern Mediterranean provides optimal conditions for the detection of ship NO₂ plumes compared to other marine regions with generally lower ozone levels.

5.3.5 NO_x emission estimates

NO_x emissions have been quantified for 130 ships sailing in the Eastern Mediterranean in 2019. Ships in our ensemble emit between 10 and 317 g s^{-1} in the hours right before the TROPOMI overpass. Krause et al. (2021) reported NO_x emission strengths of some 5 g s^{-1} for sea-going ships at the Elbe River, where ships are generally smaller and sail at slower speed. The larger ships in their ensemble emitted up to 115 g s^{-1} , more in line

with our findings here. (Chen et al., 2005) observed emission strengths of 13 and 20 g (N) kg^{-1} fuel for ships off the Californian coast, corresponding to 100-150 g s^{-1} for large ships. Our study's emission fluxes are likely at the higher end of these values because only a fraction of plumes can be detected and attributed from inherently noisy TROPOMI NO_2 imagery. Therefore, we do not expect the TROPOMI-based emission estimates to be representative for the ship fleet ensemble but rather its upper end of emitters.

Figure 5.5 shows the comparison of a forward run of PARANOX using the estimated emission NO_x flux of 130 g s^{-1} and the TROPOMI NO_2 plume for one of the investigated ships observed on 11 July 2019. The forward simulation agrees with TROPOMI to within their uncertainties in terms of plume enhancement (0.6×10^{15} molec cm^{-2}) and shape (maximum some 30 km downwind of stack), indicating that our inversion resulted in an accurate emission strength of 130 g s^{-1} .

Plumes from 10 vessels have been detected more than once from separate orbits. These offer the possibility to study the consistency of emissions by one and the same ship. This is especially true for two NO_2 plumes from the same ship observed on 14 November 2019 in two consecutive orbits, as discussed in supplement J. The overpasses are ≈ 1 h 40 min apart, and the ship sailed at the same speed so we expect comparable emission strengths. We find emissions of 126 ± 26 and 112 ± 19 g s^{-1} for this ship, well within the estimated range of uncertainty. Emission estimates for other plumes originating from one and the same ship also show a higher degree of consistency with one another than estimates from plumes originating from different ships, even after accounting for differences in speed and size.

5.3.6 Estimates of NO_x emission factors

From an atmospheric science standpoint, emissions are conventionally quantified as the mass emitted per unit time, such as our reported emission strengths in grams per second (g s^{-1}). However, from an engineering perspective, it is more appropriate to express emissions relative to the amount of power (or fuel) used. Thus, IMO emission limits are formulated in grams per kilowatt-hour (g kWh^{-1}). Here, we focus on emissions originating from individual ships operating their engines across a spectrum of real-world power levels. We derive emission factors from our emission strength estimates and knowledge about the identity of the ship and the power used by its engine at the time of observation. Figure 5.6 shows the distribution of emission factors (in g kWh^{-1}) grouped by tiers and estimated engine loads as well as the MARPOL Annex VI limits. We find that several Tier I and around half of Tier II ships emit above the limit set for weighted average emissions by the IMO¹. In the median, the emission factors are 11.5 g kWh^{-1} . This is comparable to the 12.6 g kWh^{-1} found by Van Roy et al. (2023) in the North and Baltic

¹A ship's emission factor, calculated as the weighted average of emission factors at 25%, 50%, 75% and 100% engine load, weighted with a weighting factor of 0.15, 0.15, 0.50 and 0.20, respectively, are subject to MARPOL Annex VI limits.

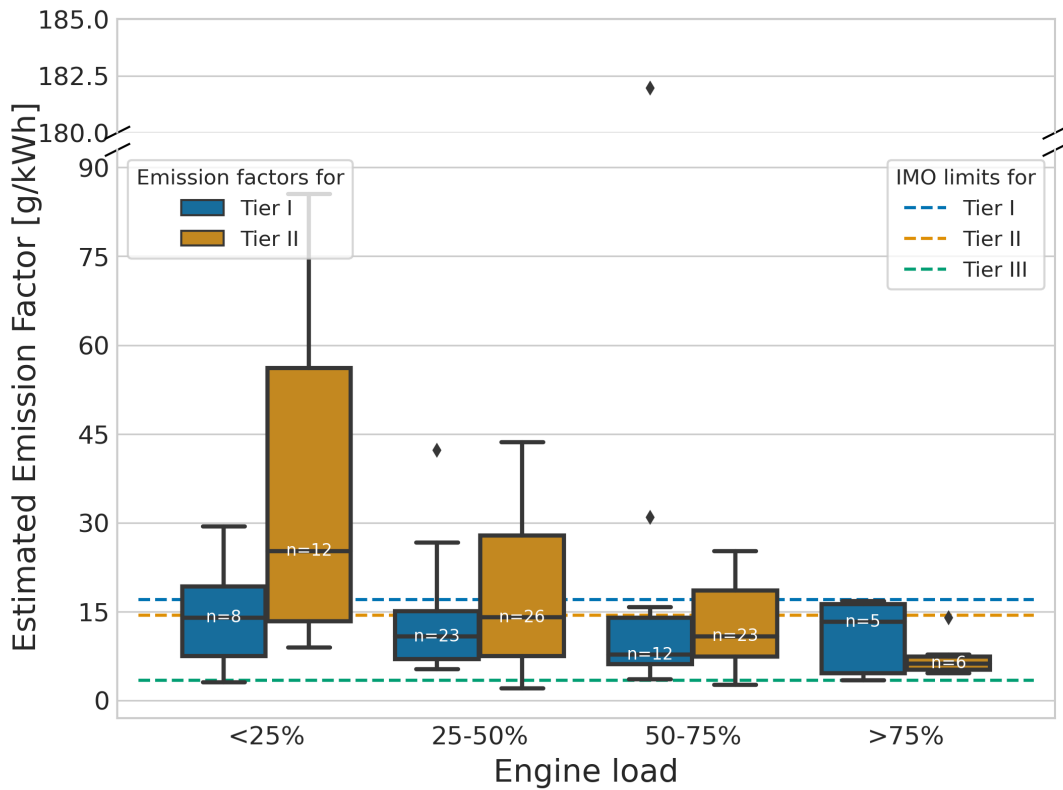


Figure 5.6: Estimated NO_x emission factors grouped by engine tier and engine load. The dashed lines show the IMO/MARPOL emission limit for each tier assuming engine rated speed ≤ 130 . We calculate the emission factor as the ratio of emissions and power used (P_{used}), with $P_{used} = P_{max} * (\frac{v_{current}}{v_{design}})^3$ where P_{max} is the maximum engine power, $v_{current}$ the current ship speed from AIS data and v_{design} the design speed of the vessel. The white numbers written in the Tier I and Tier II boxes indicate the sample size of that subgroup, black diamond show outliers.

Sea in 2020. Surprisingly, newer Tier II ships also have higher emission factors than Tier I ships, a finding in line with recent reports by Van Roy et al. (2023) and Comer et al. (2023). This is true especially for low engine loads. With increasing engine load, the emission factors of Tier II ships reduce significantly. Only at $\geq 75\%$ engine load are the detected Tier II ships cleaner than Tier I ships on average. Ship engines are likely optimized for fuel consumption, sacrificing NO_x emission reductions. Simulations of ship engines show that they have longer combustion times and lower engine efficiencies when running at low engine loads, which can lead to higher NO_x emission factors. This effect is especially pronounced with large Tier II engines running in slow-steaming mode (Andreadis et al., 2011; Cheng et al., 2017). We also find that the small number of detected older, unregulated pre-Tier ships have significantly higher NO_x emission factors than Tier

I and Tier II ships (see Supplement K), in line with expectations.

5.4 Discussion

We demonstrated that it is possible to quantify NO_x emissions from individual large ships using satellite measurements of NO_2 plumes over the Mediterranean Sea. We combined the AIS-information on ship position and speed with wind speed and direction to trace the plumes back to their stacks of origin and infer the corresponding emission rates. The detected plumes are mostly from large container ships that sail at firm speed under clear-sky circumstances.

The interplay of atmospheric chemistry and plume dispersion is crucial for understanding the complex relationship between observable NO_2 enhancements and the underlying NO_x emissions. The plumes show maximum NO_2 at 15-30 km downwind of the ship engine's stack because the chemical conversion of emitted NO into detectable NO_2 requires the entrainment of O_3 into the plume, which takes time. Further downwind, the NO_2 plume recedes into the background due to dispersion and in-plume oxidation of NO_2 . The models MicroHH and PARANOX as well as the conditions under which we detect ship NO_2 plumes from space point out the special importance of ozone. This explains the good detection conditions in the Eastern Mediterranean in summer compared to other busy waterways, as in the Mediterranean a significant ozone background is present alongside the frequent clear-sky conditions.

Taking plume dispersion, entrainment, and chemistry into full account, we infer emission strengths of 10-317 g s^{-1} for the large ships in our ensemble. From our knowledge of each ship's identity and inferred engine power supplied at the time of observation, we estimate the NO_x emission rates in g kWh^{-1} and find that half of these Tier II ships exceed the limits set for weighted average emissions set by the International Maritime Organisation at the time of detection, and that the presumably cleaner Tier II engines emit more NO_x per kWh than Tier I engines, except at high engine loads. Our remote sensing-based emission factors are consistent with those from in-plume analysis (Van Roy et al., 2023; Comer et al., 2023), and both imply that MARPOL ANNEX VI regulations are lacking real-world effect. While those studies were conducted in the North and Baltic Sea, this study shows that higher Tier II emissions are observed in the Mediterranean, too, and are caused by ships running at low engine loads. A possible explanation is that new ships are optimized for high engine loads and minimized fuel consumption, sacrificing NO_x emissions when sailing slower (Cheng et al., 2017). While these ships could nonetheless be compliant with the weighted average emission limit, in which weighting factors for emissions at high engine loads are highest, real world operating conditions lead to higher emissions from these ships. With the long turnover times of the maritime fleet, the international shipping sector might therefore become the largest contributor to atmospheric NO_x emissions in the future.

Our work also show an inherent limitation of the current MARPOL Annex VI limits for NO_x emissions of Tier II vessels. Even with our accurate emission estimates, compliance can not be monitored by atmospheric measurements during the voyage alone: The limits depend on engine rated speed and are formulated in g kWh^{-1} which required us to estimate the power used from the engine capacity and speed. Furthermore, the limits depend on the engine rated speed, which we have assumed for all detected ships to be at or below 130 rpm, where emission limits are the least strict. To be compliant, a vessel has to be below the limit not at every moment but on a weighted average along different engine loads. Knowing the emission strength of a vessel at a given time is therefore not enough, but three additional prerequisites need to be met: The engine load at time of emission needs to be known, the engine's rated speed needs to be known, and the vessel needs to be observed over a range of different engine loads to calculate the weighted average emission. In practice, this is impossible outside a laboratory. One way to solve this shortcoming is an IMO-defined not-to-exceed limit for the emission flux of a Tier II vessel, as exists for Tier III and for SO_x emissions. Additionally, requiring vessels to send their current and past engine loads as well as their engine's rated speed along with the AIS data should be considered. However, comparing our estimated emission factors to the respective average limits can inform a pre-selection by inspection agencies of ships to monitor more closely.

Some of the limitations of our method lie in the coarse spatial resolution and NO_2 detection limit of the TROPOMI sensor. These could be addressed in the future by using higher resolution NO_2 satellite data from the upcoming CO2M or TANGO missions. The detection limit of the satellite retrievals could be improved by better accounting for Raman scattering in liquid water interfering with NO_2 absorption features in the satellite spectra, by correcting for directional reflectance effects over oceans and by higher resolution of a priori profiles. A lower detection limit would increase the number of ship plumes that can be distinguished from the background and bring also smaller ships under scrutiny.

With the availability of hourly NO_2 observations from the geostationary GEMS and TEMPO sensors (over Asian and North American waters), or, soon from Sentinel-4 (over European seas), it will become possible to track a ship's emissions regionally from one hour to the next, when cloud cover remains low. Hourly observations of the same ship will allow for more robust emission estimates. Currently, because TROPOMI's local overpass time of 13:30 hrs, our estimates are restricted to ship emission strengths around noontime. It will remain challenging to discriminate between overlapping plumes from multiple ships sailing close together at the TROPOMI overpass time. Improved disentangling methods, including knowledge of stack height and precise information on wind speed, or higher spatial resolution of future satellites are required to resolve this. Overall, this study takes a step towards near real-time ship NO_x emission monitoring. Satellite remote sensing can be used to identify ships with high emission rates as candidates for on-board inspection. Our method can be used for other areas and time

periods as well and become part of a more effective effort to ensure global measures to reduce air pollution deliver on expectations.

5.S Supplement

A Impact of a-priori profiles

The TROPOMI NO₂ columns are dependent on modelled vertical NO₂ profiles, for which validation over sea are sparse and have been shown to cause a 20% negative bias in NO₂ columns over the North Sea (Riess et al., 2023). We therefore revise the air mass factor calculation in the TROPOMI NO₂ retrieval by replacing the original TM5-MP (1°×1°) NO₂ profile with plume-specific PARANOX vertical NO₂ profiles in the plume of one of the observed ships (as presented in Fig. 3 in the main manuscript). These are simulated using the initial emission estimate as input parameters. The updated data is then used to arrive at a 2nd emission estimate. This step can therefore be conducted in an iterative manner with increasing emission fluxes. We find lower AMFs, increasing the in-plume NO₂ columns by 20%, leading to an emission increase from 90 g/s in the initial inversion to 120 g/s after 3 iterations. A fourth iteration gave no significant further increase. Figure 5.S.1 shows the change of the NO₂ columns by using the final recalculated AMFs for the ship plume presented in Fig. 3 in the main manuscript.

The extensive validation of marine NO₂ profiles in Riess et al. (2023) indicated that TROPOMI NO₂ columns in the marine boundary layer need to be revised upwards by 20%, and that, together with the findings from Fig. 5.S.1, this suggests a generic multiplication factor of 1.2 for in-plume retrievals, thus avoiding forward simulations in the inversion. In-plume pixels are defined as the 90-percentile of all pixels in the ship sector. All emissions given in the manuscript are based on these bias-corrected column values, with the exception of Fig. 5 where we compare TROPOMI data to a forward using the inverted emission strength.

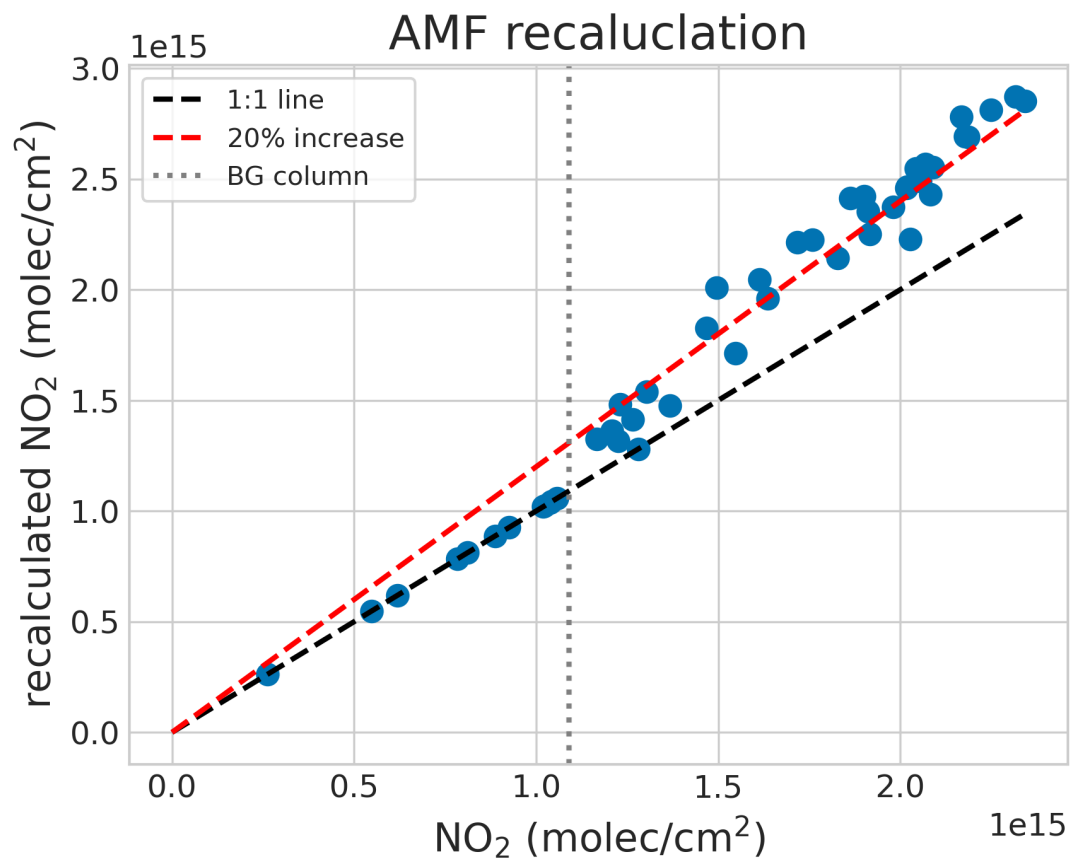


Figure 5.S.1: Recalculated NO_2 columns (with plume-resolved AMFs) versus the original TROPOMI v2.4 NO_2 columns (with TM5-MP $1^\circ \times 1^\circ$ AMFs) for a plume on 5 June 2019. The black dashed line shows the 1:1 line, the red dashed line a line of slope 1.2 and the vertical gray dotted line the background column.

B Determining XGBOOST features for the inversion

While all features are well known in the training data of the PARANOX library, their calculation in real plumes is not trivial.

- **Wind speed** This is derived from ERA5 data as the 3-hour average in 10 m wind speed around the ship track. We also calculate this for the 100 m wind in ERA5 and manually adjusted wind values if the discrepancy between the two layers was large to better fit the observed plume locations.
- **Wind direction** As wind speed.
- **Ship speed** As the average in 3h of AIS data.
- **Ship direction** Same as ship speed.
- **3h enhancement**, i.e. the NO_2 column enhancement caused by emissions from a ship in the 3 hours before and up to the observation relative to the background NO_2 . Calculated from TROPOMI data. For the calculation we need:
 - **Ship sector** This is defined as a circle sector around the net plume direction of length y_{max} (see C). We assume a 25 degree uncertainty in wind direction, the sector is therefore 50 degrees wide.
 - **NO_2 background** Is defined in an iterative process as the median of all pixels in the ship sector excluding pixels that are considered part of the plume (defined as $0.5 \cdot 10^{15}$ molecules/cm² above the initial median). As this is a somewhat arbitrary choice, we also calculate two additionally background levels, one as above but with a threshold of $0.4 \cdot 10^{15}$ molecules/cm² and one with as the median of all pixels in a 50 km radius around the last ship position.

The three-hour enhancement S_{3h} is then calculated as the sum of all pixels in the ship sector minus the background level multiplied with the area of the pixels. Additionally, we assume an uncertainty in y_{max} of 10.8 km (corresponding to an uncertainty in wind speed of 1 m/s), and calculate two additional ship sectors. In total, we thus end up with 5 different E_{3h} values. Typical values for E_{3h} are of the order of $3 \cdot 10^{27}$ molecules with a standard deviation of the five different values of $0.6 \cdot 10^{27}$ molecules. The uncertainties given in the main manuscript is the standard deviation of the inversion of these 5 values, combined with PARANOX and XGBoost uncertainties.

- **Ozone background** From CAMS as the 3h-average in the lowest 250 m of the atmosphere in the vicinity of the ship track.
- **NO_x background** From CAMS as for ozone.
- **Stability** Derived from the vertical temperature gradient in 3h-averaged ERA5

data.

- **Latitude** The latitude of the ship location at TROPOMI overpass time.
- **Month** The month in which the observation was done.

C Simple metrics to quantify plumes

To characterize the modeled plumes and compare to TROPOMI observation, we coarsen the PARANOX output for all the simulations in the plume library to a grid of 5 km x 5 km, in the same order as typical TROPOMI pixel sizes. We define 5 metrics to quantify the plume: (i) The 3h-plume sum S_{3h} , (ii) the 3h-plume enhancement E_{3h} , (iii) the plume length l , (iv) the plume peak N_{max} and (v) the location of the plume peak relative to the ship d_{max} and test their performances to describe the plume as possible features used in the XGBoost inversion framework.

S_{3h} is defines as the sum of all columns N in the center of the plume from the ship location to a distance $y_{max} = 3h * \vec{v}_{eff}$, which marks the distance of the 3-hour old plume from the stack. Here, \vec{v}_{eff} is the effective wind speed $\vec{v}_{eff} = \vec{v}_{wind} - \vec{v}_{ship}$, with \vec{v}_{wind} and \vec{v}_{ship} being wind and ship speed, respectively. S_{3h} can be calculated as

$$S_{3h} = \sum_{i=0}^{y_{max}} N_i \quad (1)$$

and thus all the NO_2 in the 5 km-wide centerline of the freshest 3 hours of the plume. Likewise,

$$E_{3h} = \sum_{i=0}^{y_{max}} (N_i - N_{bg}), \quad (2)$$

where N_{bg} is the background column. The plume length is defined as

$$l = \max(i; N_i > 1.25N_{bg}) * 5km \quad (3)$$

marking the length of the plume that is expected to be significantly enhanced over the background and thus visible in TROPOMI observations. Finally,

$$N_{max} = \max(N_i) \quad (4)$$

and

$$d_{max} = (i, \text{where } N_i = N_{max}) * 5km \quad (5)$$

These 5 metrics have in common that they are easy to compute on both modelled and observed plumes. We study the influence of model forward parameters on the metrics. As our goal is a NO_x emission inversion, we are looking for a metric strongly correlated to the emission flux, while only slightly influenced by the other parameters. Additionally, the metric should be insensitive to TROPOMI noise and exact pixel size. In the following section, the metrics are compared regarding there sensitivity to model forward parameters.

D PARANOX

PARANOX is a Gaussian Plume Model (GPM). A GPM is a plume model, following the evolution of an air parcel (here the plume) over time. For the comparison to TROPOMI, we are interested in a snapshot of a plume emitted at different points in time, experiencing different radiation intensities and therefore different photochemistry. We compute several simulations with PARANOX, emitted every 15 minutes and create a superposition of these plumes. The 15-minute-interval was chosen to keep the computational costs low without sacrificing the superposition-character of the plume. For our simulations, we emit 95% of NO_x as NO and 5% as NO_2 .

Using the effective wind speed

$$\vec{v}_{eff} = \vec{v}_{wind} - \vec{v}_{ship} \quad (6)$$

where \vec{v}_{wind} and \vec{v}_{ship} are wind and ship speed, respectively, we can translate the output into an Eulerian space, creating a typical plume shape. The Eulerian Space outside of the rings (and therefore the plume) is filled with a background profile from CAMS. We then integrate the 3D NO_2 field, creating a TROPOMI-esque column view.

To understand driving factors in plume shape, we create a large library of simulated plumes, capturing a wide range of atmospheric conditions. To that end, we span a wide space of input parameters, changing both plume dispersion (wind speed, ship speed, their relative angle, Pasquille stability class, Boundary Layer Height (BLH)) and plume chemistry (NO_x background, O_3 background, radiation (latitude, month)), and the NO_x emission strength. As TROPOMI overpasses always occur at the same local time, the time of day was not needed to vary. The range of the models forward parameters can be found in the Fig. 5.S.2.

E The PARANOX plume library and important model parameters

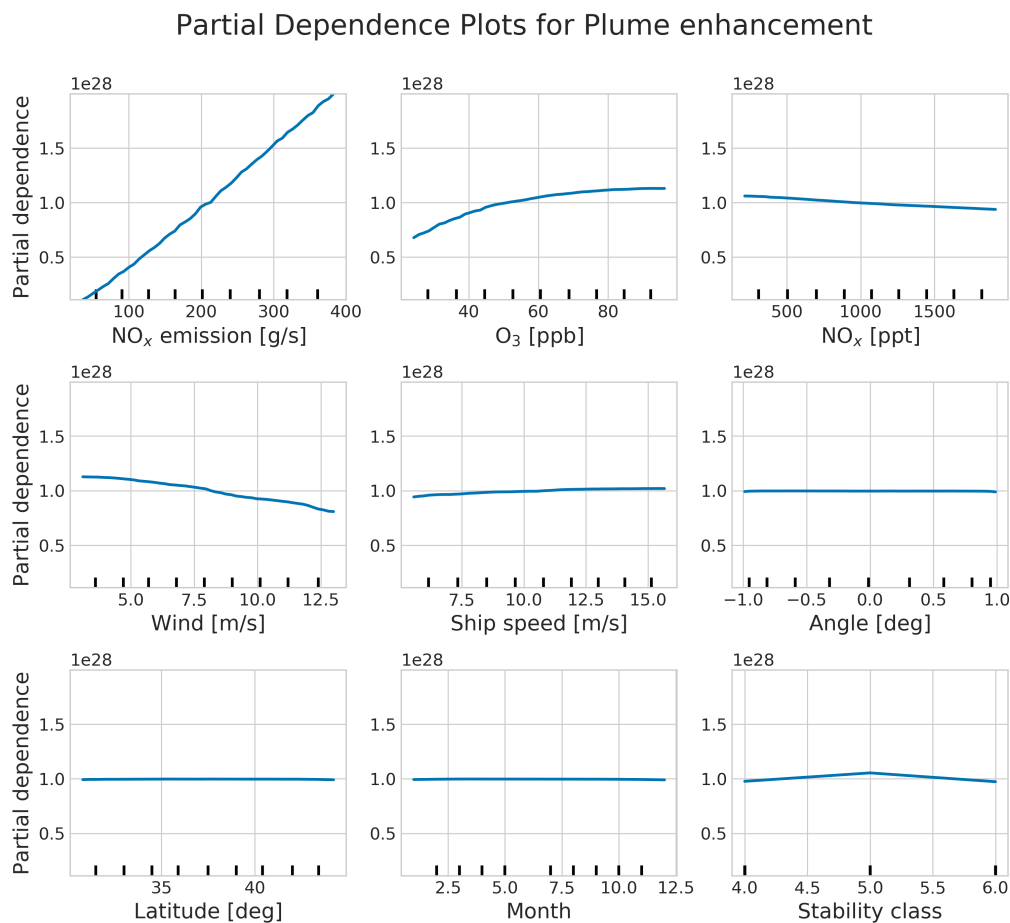


Figure 5.S.2: Partial dependence of E_{3h} . All parameters were randomly co-varied in the model runs.

In total 40880 model runs with PARANOX were performed. We find that the 3h plume enhancement E_{3h} (see above) to be the most promising metric for characterizing plumes for the purpose of emission inversions. Fig. 5.S.2 shows the partial dependence plot of the resulting E_{3h} on different forward model parameters.

We see a strong, almost linear dependence on the emission flux. Other important parameters are the background ozone levels and the wind speed. For ozone, we observe an elevated plume enhancement with increasing ozone. This can be understood from the NO_x null cycle. Here, the ozone level defines both the rate by which the emitted NO is reacting to the NO_2 observable from satellites and the steady-state balance of the two subspecies. At very high background ozone levels the dependence curve flattens. This can be interpreted as a shorter but more pronounced plume, caused by the fast NO-to- NO_2 conversion but shorter lifetimes at very high ozone levels. Higher wind speed leads

a more dispersed plume, accelerating the entrainment of background air, including the hydroxyl radical OH and ozone, thus leading to a faster destruction of NO_x . The plume enhancement is remarkably robust with regard to the other forward parameters, only slight dependencies can be observed. It should be noted, that due to the non-linear chemistry the dependencies co-vary with other parameters. E.g. for small emission fluxes, the ozone dependency will flatten earlier and the background NO_x will become more important. Furthermore, we ignored that many of the parameters have a seasonal dependence (e.g. winds are higher in winter), and studied only the effect of radiation changes in different months.

Other tested metrics such as the 3h plume sum S_{3h} , peak NO_2 N_{max} , location of peak

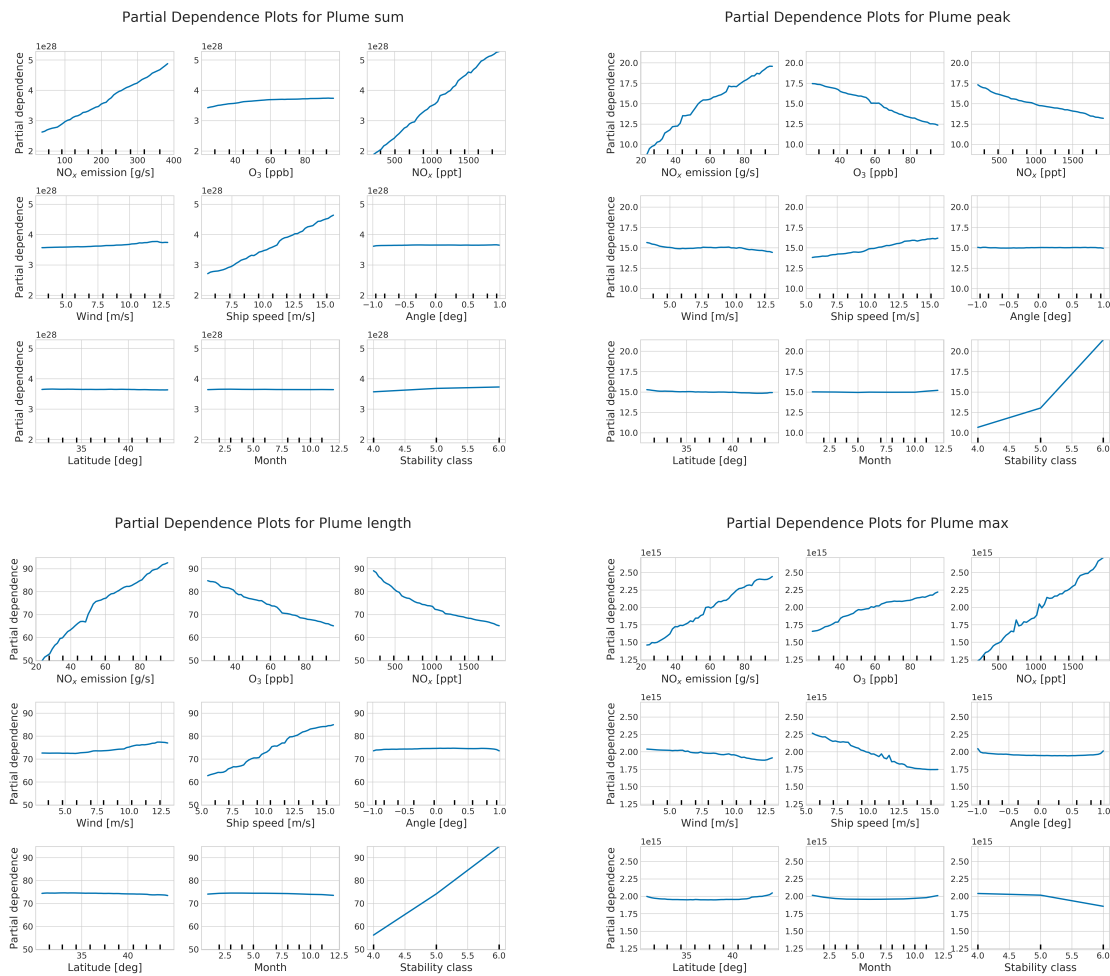


Figure 5.S.3: Partial dependence of plume sum, distance of the peak location to the stack, length of the visible plume and maximum column density. All parameters were randomly co-varied in the model runs.

NO_2 relative to the stack d_{max} , and plume length l also show a strong dependence on the emission flux but are more sensitive to other model parameters, as shown in Fig. 5.S.3.

For example, the distance between emission stack and plume NO_2 peak decreases with higher ambient ozone, as the conversion of emitted NO to visible NO_2 is sped up with higher entrainment of ozone. This is also visible with the increase of the distance with more stable conditions. We conclude, that - besides for emission flux, ozone background and wind speed - model parameters affect the shape of the plume, but not the integrated NO_2 amount.

For the inversions to be the most robust, we are looking for a metric that is strongly dependent on the emission flux but ideally insensitive to other parameters, which are retrieved from reanalysis data. For example, we retrieve the stability class from the temperature gradient in the lowest three levels in ERA5 data. As stability classes are discrete, a slight inaccuracy in ERA5 could lead to a large bias in an emission estimate if based on the plume sum, which is strongly dependent on stability class. Overall, the 3h-plume-enhancement E_{3h} is the least sensitive to these parameters and therefore our choice of metric for this study.

F Comparing PARANOX and MicroHH simulations

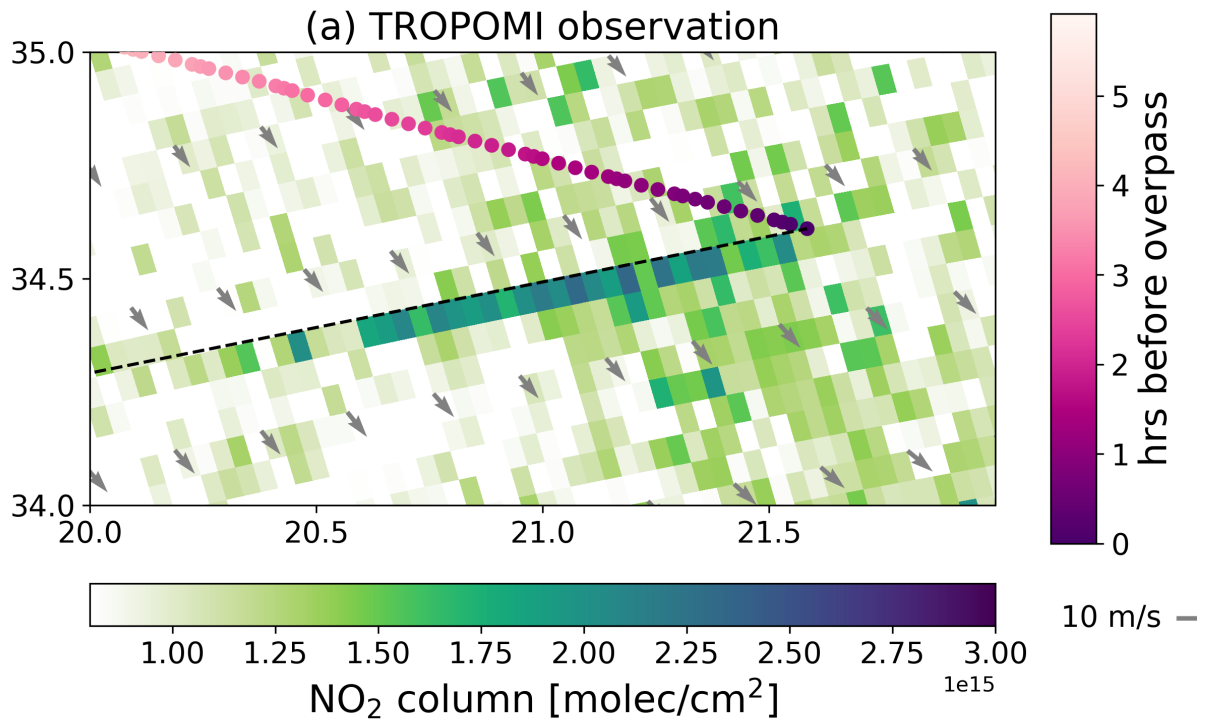


Figure 5.S.4: TROPOMI scene with clear ship plume on 5 June 2019. Overlaid in circles are the ship locations in the hours before overpass taken from AIS data. The black dashed line is the wind-shifted ship track, indicating the expected location of the plume.

We take a closer look at a ship plume observed on 5 June 2019, shown in Fig. 3 and Fig. 5.S.4. This plume was selected as it is aligned with the TROPOMI swath, simplifying a comparison of model to satellite data. We start with a comparison of PARANOX runs of different stability classes and a MicroHH simulation with the TROPOMI observation. For an analysis of the impact of a-priori profiles see supplement A.

In the presented inversion method, the XGBoost model is trained with PARANOX runs. As GPMs rely on parametrizations, we test here how realistic the resulting plume is. As we use the model runs to analyze TROPOMI retrievals at ≈ 5 km resolution, the comparison of the models at TROPOMI resolution is essential. However, to understand possible differences in the model runs, we also compared the models at their native resolution.

We compare the modelled plume of a ship sailing in the Eastern Mediterranean, with a constant emission flux of 120 g/s on 5 June 2019. The resulting plume is oriented mainly in east-west direction as shown in Fig 5.S.5. Fig 5.S.6 shows the NO_2 enhancement as a function of distance from the ship. The turbulent structure of the MicroHH becomes immediately clear, whereas PARANOX gives a very smooth curve as is expected for a Gaussian model. We also see that situations with more stability in the PARANOX simulations result in longer in-plume NO_2 lifetimes, implying longer NO_2 plumes. This is

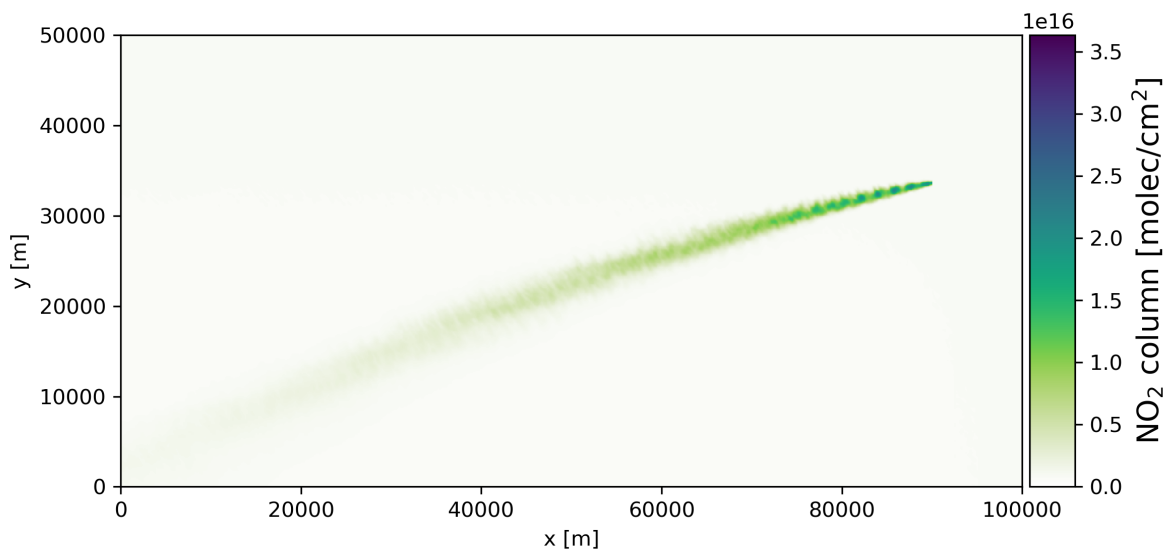


Figure 5.S.5: Ship plume modelled with MicroHH on 5 June 2019. We show the NO_2 column as the averaged output of 7 timesteps, each 5 minutes apart.

caused by less mixing in of fresh air, thus less O_3 to form NO_2 in the early stages of the plume and less OH, which cleans the air of NO_x . Overall, the agreement is reasonably good, especially the neutral case (D) matches the MicroHH run reasonably well. Comparing the cross-sections shown in Fig. 5.S.8 we see that all PARANOX simulations are exhibiting a slightly wider plume than MicroHH. This is probably due to different averaging times: a turbulent plume will always exhibit meandering, which leads to a wider plume in long-time averages. A Gaussian Plume model is designed to catch such a time-averaged plume. As we are interested in using TROPOMI data, these small spatial scales are of less interest: The cross-sections clearly show that even at an older stage (e.g. after 40 km), most of the NO_2 is still within a width of 5 km and thus the typical size of a TROPOMI pixel. Overall, the PARANOX simulation of stability F comes closest to the observation, while the run with stability D is closest to MicroHH. ERA5 data on the other hand indicates a temperature gradient belonging to stability E. Using the MicroHH simulation - regridded to TROPOMI resolution - as an input for the XGBoost inversion results in an estimated emissions of 133 g/s, some 10% higher than the actual emissions used for the MicroHH runs of 120 g/s.

Figure 5.S.7 shows the evolution of the NO_2 plume for the observed data as well as model simulations regridded to TROPOMI resolution. The PARANOX run agrees well with TROPOMI data (see Table 5.S.1). Furthermore PARANOX agrees reasonably well with MicroHH. We conclude that PARANOX is a suited tool to simulate ship plumes evaluated at a resolution of several kilometers and use PARANOX to create a plume library used in our inversion.

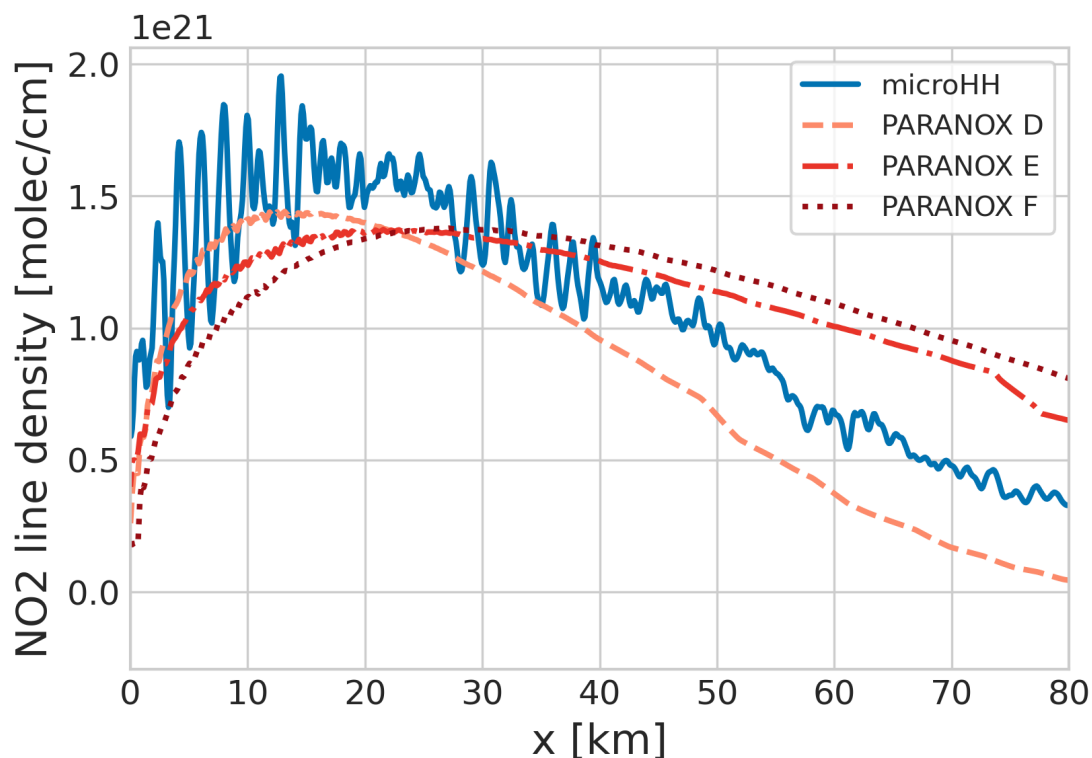


Figure 5.S.6: NO_2 enhancement of the plumes with distance from the ship. Blue shows the MicroHH plume, PARANOX simulations of stability class D, E and F are shown in different shades of red.

Table 5.S.1: Comparison of PARANOX and MicroHH simulation to the recalculated TROPOMI columns.

simulation	R	RMSE	residuals
MicroHH	0.14	$7.07\text{e}+14$	$2.78\text{e}+15$
PARANOX D	0.27	$7.39\text{e}+14$	$8.20\text{e}+15$
PARANOX E	0.47	$4.80\text{e}+14$	$3.41\text{e}+15$
PARANOX F	0.64	$4.04\text{e}+14$	$3.64\text{e}+15$

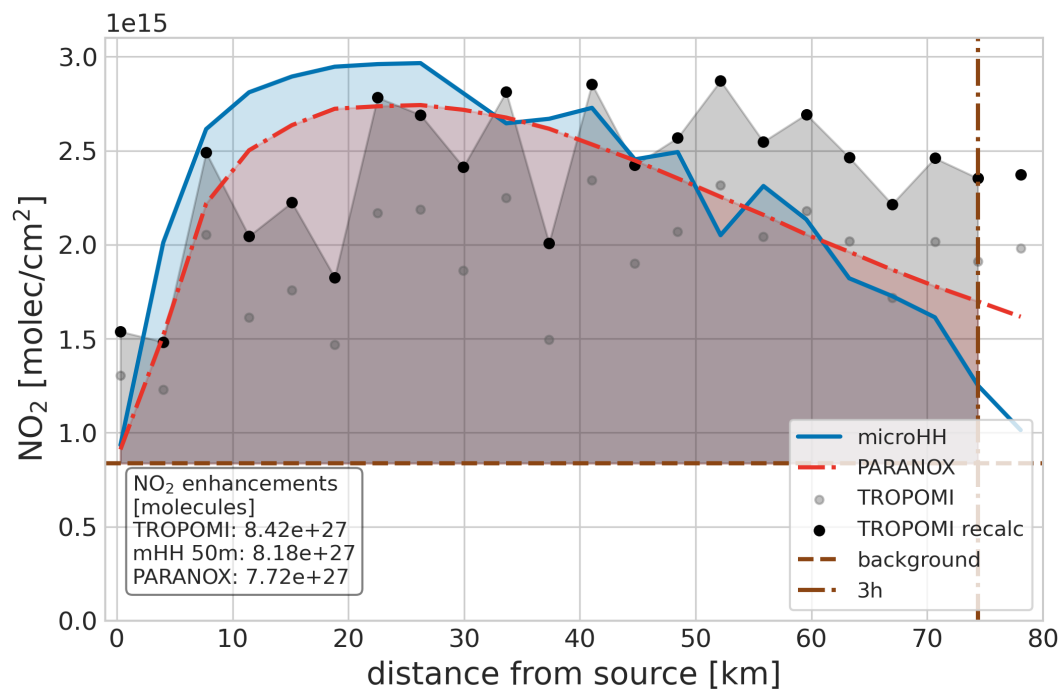


Figure 5.S.7: Similar to Fig. 5.S.6 but the model output is gridded to the TROPOMI pixels, corresponding to Fig. 2 in the main text. MicroHH is shown in blue again, red shows the PARANOX run of stability class E, grey dots are the TROPOMI columns and in black the recalculated TROPOMI columns using the PARANOX a priori profiles. The shaded areas indicate the 3-hour-plume-enhancement E_{3h} used in the inversion.

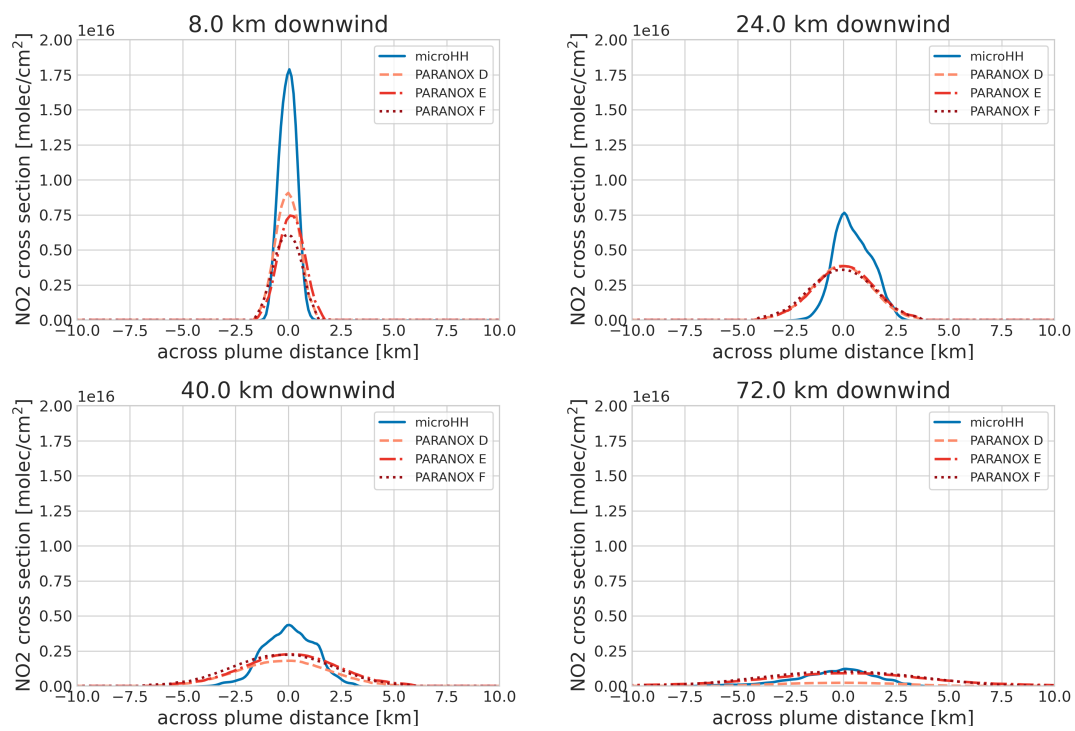


Figure 5.S.8: Cross section of the plume simulated with MicroHH (full line) and three different stability classes in PARANOX at different distances from the source.

G XGBOOST model

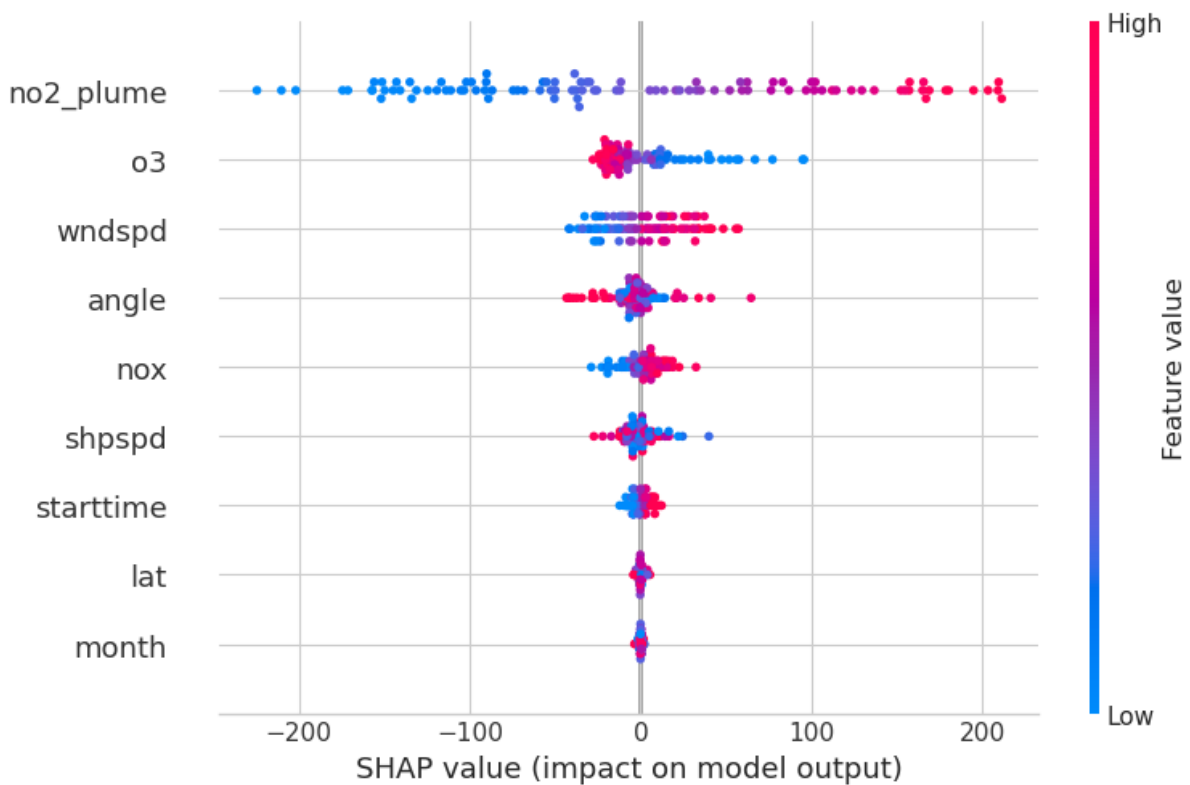


Figure 5.S.9: SHAP value for the XGBoost model for stability class D. The absolute SHAP value is an indication of the importance of the model features on the output, a feature with a positive SHAP value has a positive impact on the output, a negative value indicates negative impact. We see that the plume enhancement (no2_plume), ozone background and wind speed are the most relevant parameters when determining the emission flux.

In Fig. 5.S.9 we show the SHAP (SHapley Additive exPlanations) value of the model for stability class D, indicating the importance and influence of the forward parameters on the XGBOOST output (i.e. the emission flux). As expected, the NO₂ in the plume has the largest influence, followed by the ozone background and the wind speed. High ozone values lead to lower emission estimates as the NO₂ is formed more efficiently, compare also Fig. 5.S.2.

H The role of the head wind strength

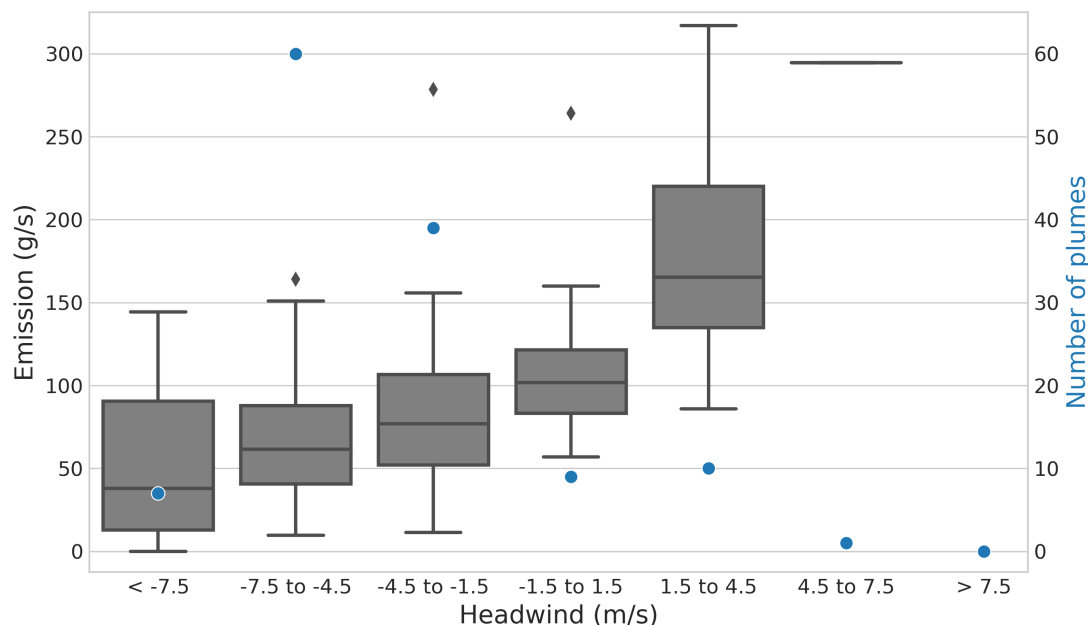


Figure 5.S.10: Box plot of estimate emissions binned for strength of the headwind (the negative tangential part of the wind).

We find that the detected plumes of ships sailing with the winds are caused by lower emissions on average (see main text). We argue this is an artefact of biased selection, caused by more elongated plumes in case of headwind, making the enhancement per pixel smaller and thus harder to detect given the noise level of TROPOMI. To test this we use a plume observed on 6 October 2019 with a strong tailwind, leading to a very compact plume almost staying directly above the ship. We run PARANOX simulation for this plume with the estimated emission flux of 40 g/s with the real wind as well as a headwind. The two simulations are shown in Fig. 5.S.11. In the headwind figure, the plume becomes invisible, while in the tailwind case, the plume of the size of only two pixels is clearly visible, confirming our hypothesis.

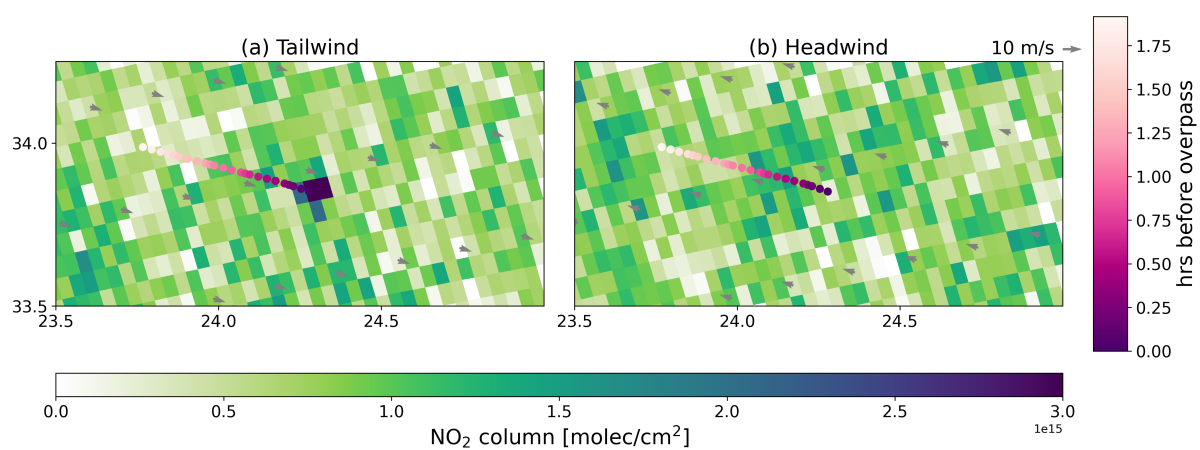


Figure 5.S.11: PARANOX simulation of a ship with an emissions flux of 40 g/s with tailwind (a) and headwind (b). The plume is much more compact in the tailwind case but invisible in the headwind case.

I IMO MARPOL Annex VI nitrogen emission regulations

Table 5.S.2: NECA guidelines according to MARPOL Annex VI.

	Keel laying date after	Total weighted cycle emission limit of NO+NO ₂ (g/kWh) for engine speed n (revolutions per minute)		
		$n = 130$	$n=1000$	$n = 2000$
Tier I	1 January 2000	17.0	11.3	9.8
Tier II	1 January 2011	14.4	9.0	7.7
Tier III	1 January 2016	3.4	2.3	2.0

The amount of allowed NO_x emitted per engine output in g NO₂/kWh is regulated by International Maritime Organization (IMO) MARPOL Annex VI. The limits depend on the keel laying date of a ship (Table 5.S.2) and must be adhered to based on a weighted cycle average. Globally, all ships build since 1 January 2000 have to comply with Tier I NO_x standards, while ships build in or after 2011 have to comply with Tier II standards. Since 1 January 2021, the stricter Tier III standard applies to ships build after 1 January 2016 operating within the European Nitrogen Emission Control Areas (NECAs) in the North and Baltic Sea, which is expected to reduce per ship emissions by 75% relative to Tier II. The new Tier III guidelines have been in place for American NECAs since 1 January 2016.

While the Mediterranean Sea will become a Sulphur Emission Control Area (SECA) in 2025, the introduction of a Mediterranean NECA is still under discussion.

J Plumes from the same ship in consecutive orbits

On 14 November 2019 a ship's plume was observed in two consecutive TROPOMI orbits, as shown in Figure 5.S.12. In the main manuscript, we discuss the emission estimates, agreeing within the range of uncertainty.

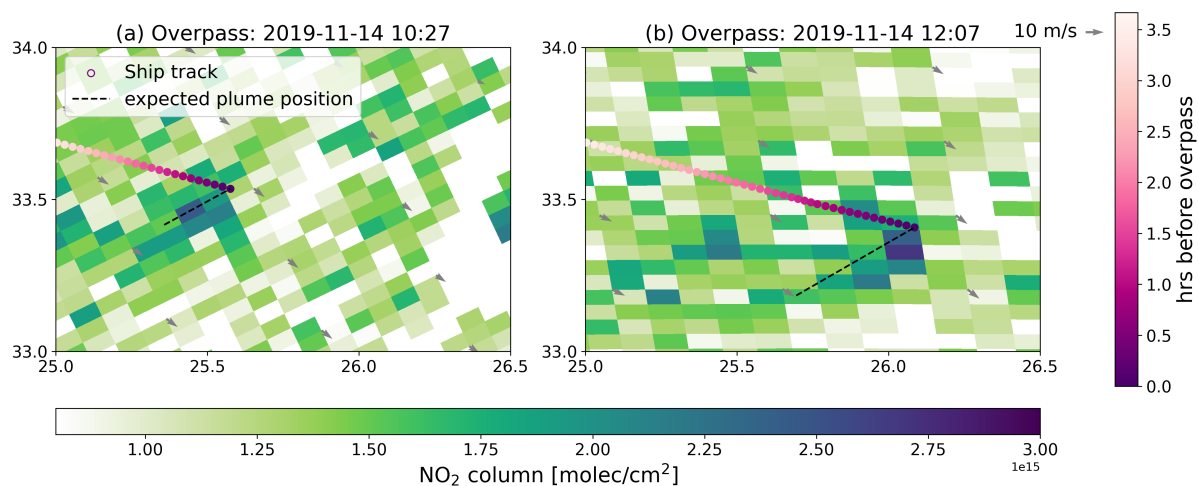


Figure 5.S.12: Observed plume 14-11-2019 of the same ship's plume in two consecutive orbits. Between the two overpasses, 1 h 40 m have passed.

K Emission factors per Tier

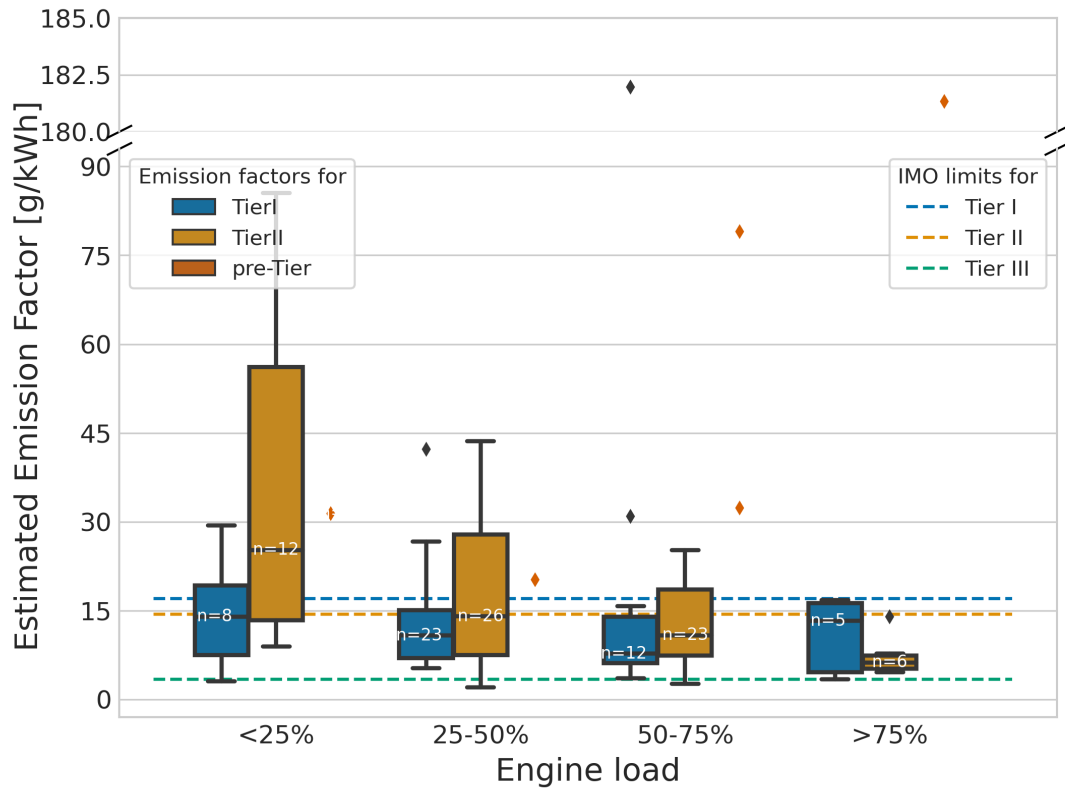


Figure 5.S.13: Similar to Figure 6 in the main text but including pre-Tier ships. Estimated emission factors grouped by engine Tier. The dashed lines show the IMO/MARPOL emission limit for each Tier assuming engine rated speed ≤ 130 . We calculate the emission factor as the ratio of emissions and power used (P_{used}), with $P_{used} = P_{max} * (\frac{v_{current}}{v_{design}})^3$ where P_{max} is the maximum engine power, $v_{current}$ the current ship speed and v_{design} the design speed of the vessel. The white numbers written in the Tier I and Tier II boxes are the number of plumes in that subgroup.

Chapter 6

Temporal variability in ship NO₂
plume detectability from space

Abstract

Ship NO₂ plume detection with satellites will enter a new era with the coming of geostationary and high-resolution satellites. As atmospheric chemistry and meteorological conditions have a strong diurnal and seasonal cycle, it is not trivial what these new data sources will reveal. In this chapter, we use the Gaussian Plume Model PARANOX to study the effects of radiation, ozone and NO_x background, seasonal cycles of wind and Pasquille stability class, boundary layer height, and temperature in the marine boundary layer on plume detectability. We identify the key factors influencing the detectability of ship plumes over the eastern Mediterranean Sea, namely radiations, ozone abundance and wind speed. We find that the TROPOMI overpass is already at the optimal time of day for ship plume detection over the Eastern Mediterranean, when NO_x background is low and ozone levels are high, favoring the conversion of the emitted NO to observable NO₂. Additionally, the short NO_x lifetime in the early afternoon make overlapping of plumes less likely as they are less elongated than for example in the early morning. We find that the detectability of ship plumes is highest in summer, when the chemical and meteorological conditions are most favorable, even when ignoring the lower cloudiness in summer. The new geostationary satellites are therefore expected to provide more captures of ship NO₂ plumes and lead to better estimates of ship NO_x emissions.

6.1 Introduction

In the previous chapter we have demonstrated that satellites can not only detect NO_2 plumes but that we understand their evolution and can estimate the emissions causing the observed plumes by utilizing the Gaussian plume model PARANOX. Furthermore, we have shown that in the Eastern Mediterranean plumes are most often detected in summer, likely linked to higher ozone background and smaller scattering angles compared to other seasons. However, a quantification of the effects leading to this much better detectability in summer, shown in the previous chapter is missing.

As previous work presented in this thesis was steered by TROPOMI observations, we focused on plumes around noon local time, when TROPOMI flies over. With the new geostationary satellites such as GEMS (over Asia), TEMPO (Northern America) and the upcoming Sentinel-4 (over Europe in 2024), ship plumes might soon also be observable in other times of the day. Furthermore, future orbiting satellites such as CO2M (expected launch in 2026, resolution of $2 \times 2 \text{ km}^2$) and TANGO (expected launch in 2027, resolution $300 \times 300 \text{ m}^2$) will provide NO_2 observations at higher spatial resolution and different overpass times. Additionally, the polar orbiting Sentinel5 (to be launched in 2025) will provide similar data to TROPOMI but at 9:30 am local time. All this is especially interesting because it will likely allow to observe ship plumes at different times of day. Additionally, chemical compositions and reaction rates in the atmosphere change over the course of the day. For example, NO_2 is photolyzed by sunlight. This means, that at nighttime in the absence of sunlight NO_x is present almost exclusively as NO_2 while during the day NO_2 is photolyzed to NO . Furthermore, OH - the main daytime sink of NO_x - is produced by the photolysis of O_3 . We therefore expect higher NO_2 -to- NO ratios and longer lifetime of NO_x during the early morning and late afternoon when the sun is low and so are the photolysis rates. This is expected to increase detectability of ship NO_x plumes. However, the likely more shallow and stable marine boundary layer in the morning as well as lower ozone levels might slow down the conversion of the emitted NO to detectable NO_2 , thus decreasing detectability. Which of these two processes is dominant needs to be studied.

In this chapter, we therefore focus on the implications of the diurnal and seasonal cycles of chemistry and meteorology on (ship) plume detectability. We will study when the chemical and meteorological conditions are optimal for ship plume detection and attribution in the eastern Mediterranean Sea and which of the processes detailed above are dominant. To this end, we will use the same Gaussian Plume Model as in the previous chapter, PARANOX, and include the temporal variability of the background concentration of NO_2 , NO and O_3 , additionally to the radiation, which was already changing with daytime in the previous chapters. This is important as we expect more pronounced changes in the background concentration of these species close to sunset and sunrise. We then study the diurnal cycle of ship NO_2 plumes and discuss the implications for the detectability of ship NO_2 plumes, anticipating the new and upcoming

satellite missions. More concretely, we study the effects of radiation, ozone and NO_x background, seasonal cycles of wind and Pasquille stability class, boundary layer height, and temperature in the marine boundary layer on detectability and identify the key factors influencing the detectability of ship plumes over the eastern Mediterranean Sea.

6.2 Material and methods

6.2.1 PARANOX

We use PARANOX as introduced in Chapters 4 and 5, simulating a plume caused by 4 hours of emission as a superposition of 15-minute-long snippets from in total 16 simulations. Each of these 16 simulations has been computed with different sun angles, causing different photolysis rates. We will now extend this model to include the temporal variability of the background concentration of NO₂, NO and O₃ from CAMS regional reanalysis (see below), instead of using a constant value as input. For wind speed, stability class and temperature we will use the ERA5 reanalysis and provide the value during simulated observation time.

To produce NO₂ columns we use the resulting in-plume NO₂ fields together with background CAMS data (see Sec. 6.2.2) and regridded to a resolution of 5x5 km². To simulate the in-plume low bias caused by the a priori profiles in the TROPOMI retrieval (see Chapter 4), we lower all simulated in-plume pixels - identified by being 0.1·10¹⁵ molec/cm² above the background - by 20% to provide a realistic observation simulation. To judge the detectability we then count the 5x5km² pixels with a NO₂ column above the noise level, which we assumed to be 0.35·10¹⁵ molec/cm². Considering the explicit treatment of the 20% low bias this is in line with the 0.5·10¹⁵ molec/cm² given in the TROPOMI NO₂ ATBD (van Geffen et al., 2021). All plumes were simulated with a headwind to facilitate a fair comparison.

6.2.2 CAMS

We use mean hourly data from the Copernicus Atmosphere Monitoring Service (CAMS, METEO FRANCE et al., 2022) regional reanalysis for the year 2019 at 34°N, 24°E. This dataset provides hourly data of NO₂, NO and O₃ at 0.1x0.1° resolution. We interpolate this data to 15-minute resolution and use it as background concentration in PARANOX. Figure 6.1 shows the resulting diurnal cycles of NO_x and O₃ for different months in 2019. We see a clear diurnal cycle in both species, with lowest NO_x concentration in the middle of the day, when photolysis rates - producing OH - are highest. We also see that during nighttime virtually all NO_x is present as NO₂ as there is no photolysis of NO₂. The diurnal cycle is strongest in summer, when daytime NO_x reaches the lowest values. The O₃ concentrations peak in the afternoon, with the highest values found in

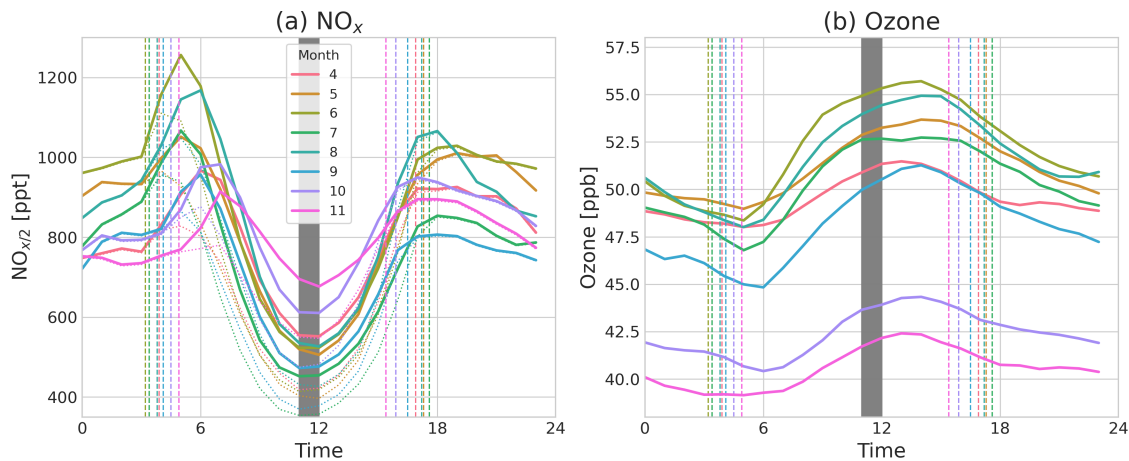


Figure 6.1: CAMS NO_x (a) and O_3 (b) background concentration in 2019 at 34°N , 24°E . The colors show diurnal cycles for different months (April - November). The vertical dashed lines indicate sunrise and sunset for each month. The gray area indicates typical TROPOMI overpasses. For panel (a), the thinner, dotted lines indicate the NO_2 concentration. Times are giving in UTC.

June and July. There are two chemical mechanisms that influence ship NO_2 plume detectability in opposite directions: On the one hand, ship NO_2 plume detection could be easiest in summer afternoons, when ozone is high leading to a fast NO to NO_2 conversion at higher NO_2 -to- NO ratios. On the other hand, NO_x lifetimes are shortest in the early afternoon, leading to a fast removal of NO_x from the plume, making the plume shorter and more difficult to detect. Additionally, early morning or late afternoon are of interest as photolysis rates are low, providing potentially high NO_2 -to- NO ratios, and low OH levels lead to long NO_x lifetimes, potentially leading to longer visible plumes.

6.2.3 ERA5

We use mean hourly data from the European Center for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis for the year 2019 at 34°N , 24°E . This dataset provides hourly data of temperature, wind speed, wind direction, cloud cover and boundary layer height $0.25 \times 0.25^\circ$ resolution. We interpolate this data to 15-minute resolution and use it as input for PARANOX. To define the Pasquille stability class, we use the temperature gradient in the lowest 50 hPa layers. We use the definition following NOAA (<https://www.ready.noaa.gov/READYpgclass.php>), defining temperature gradients between -1.5 and -0.5 K/100 m as stability D (neutral), and gradients between -0.5 and 1.5 as stability E (slightly stable).

Figure 6.2 shows the diurnal cycles of wind speed, Pasquille stability class, boundary layer height and near-surface temperature for different months in 2019. We see a seasonal (and less pronounced diurnal) cycle on 10 m wind speed. Winds were highest in

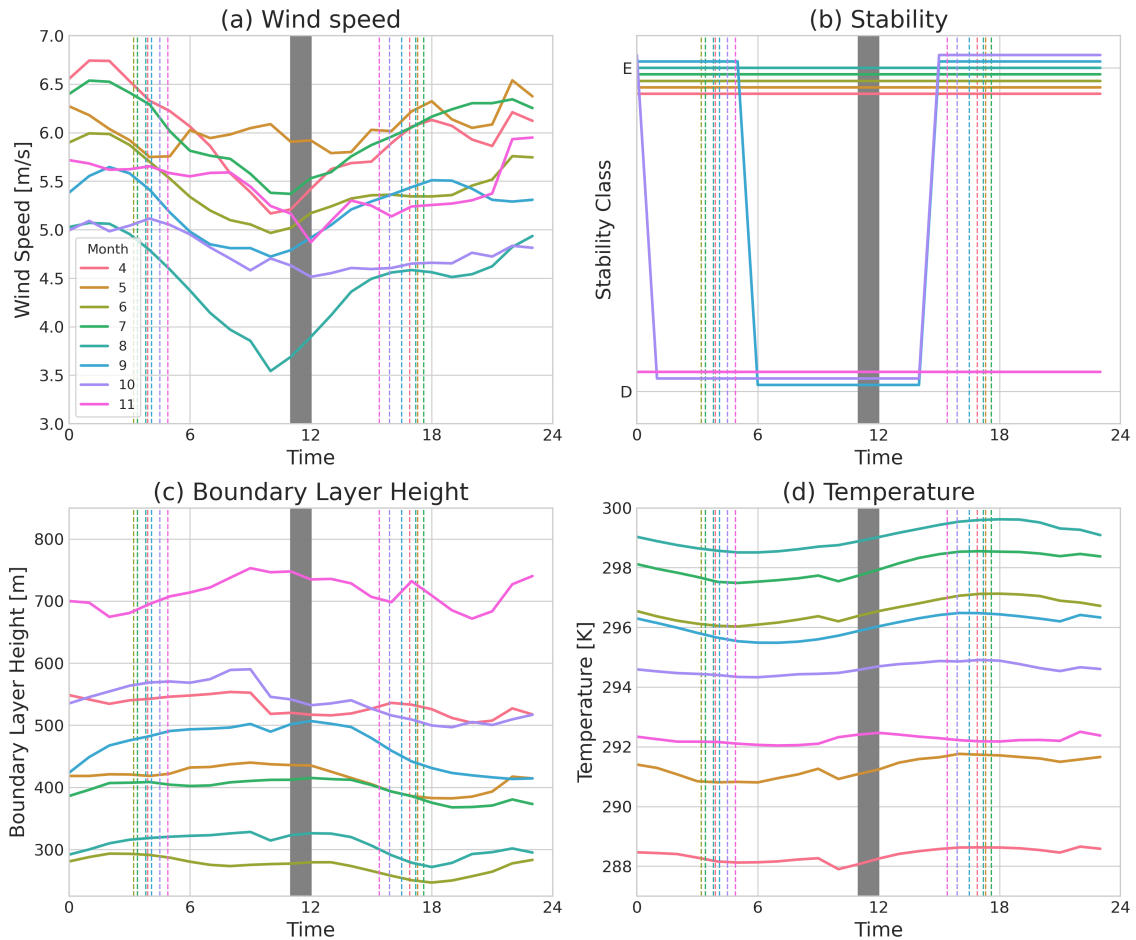


Figure 6.2: Wind speed (a), Pasquille stability class (b, months are displaced on the y-axis for clarity), boundary layer height (c) and near-surface temperature (d) in 2019 at 34°N, 24°E. The colors show diurnal cycles for different months (April - November). The vertical dashed lines indicate sunrise and sunset for each month. The gray area indicates typical TROPOMI overpasses.

May and July but lowest in August. The dominant Pasquille stability class is E, with D present in autumn. The marine boundary layer in the eastern Mediterranean is thus considerably more stable than over land, at least during the day. Boundary layers in the Mediterranean are usually shallow, with a very weak diurnal cycle. This can be explained with the water surface, which - unlike land surface - is not warmed up much by solar radiation, thus suppressing the convection observed over land. Highest boundary layers were observed in November, when the air is cold and the sea surface warm, and lowest in June when the opposite is true. The near-surface temperature shows a clear seasonal and weak diurnal cycle, too. The highest temperatures were observed in July and August, the lowest in November. The effects of these parameters on the detectability of ship plumes is not trivial. For example, higher wind speeds will lead to faster dilution of the plume, making detection more difficult. At the same time, more dispersion also leads to faster

ozone entrainment and subsequently to higher NO_2 -to- NO ratios. Higher temperatures will lead to higher reaction rates, leading both to a fast conversion of NO to NO_2 but also a shorter lifetime. As seasonal and diurnal temperature gradients in the study location are low, this effect is likely to be small.

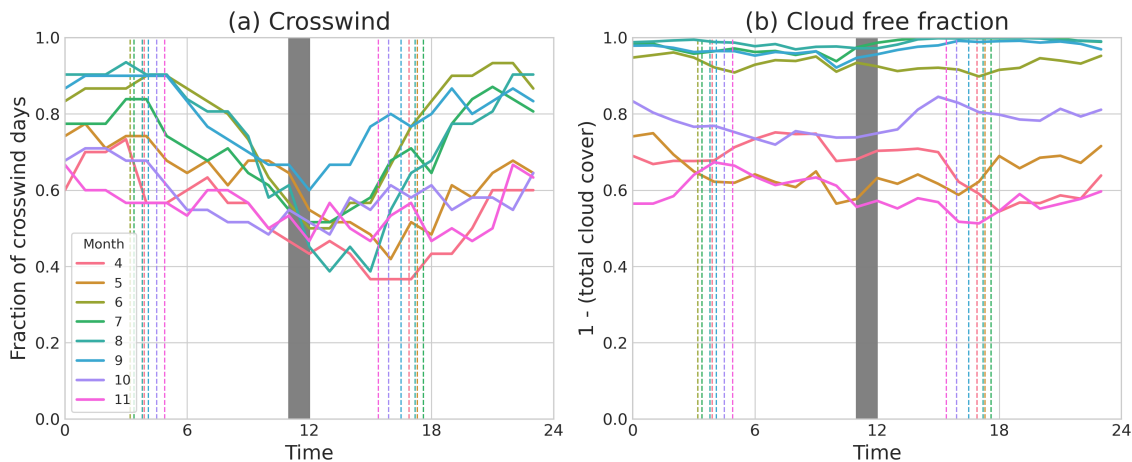


Figure 6.3: Frequency of beneficial wind directions (a) and cloud free scenes (b) in 2019 at 34°N , 24°E from ERA5. Fraction of occurrences of winds perpendicular to the main ship track. The colors show diurnal cycles for different months (April - November). The vertical dashed lines indicate sunrise and sunset for each month. The gray area indicates typical TROPOMI overpasses.

We also investigate how frequent wind speed perpendicular to the ship lanes occur. These wind directions make an attribution of plumes to an individual ship easier as they make an overlap of plumes from several ships in the busy ship track less likely. Figure 6.4 shows TROPOMI observations above a busy ship track for two distinct wind directions. For the along-track-wind on the left, the clearly visible plume cannot be assigned to a single ship but might be caused by several ships sailing (anti)parallel. For the perpendicular wind direction, four distinct ship plumes are visible and can be assigned to individual ships, even though several ships sail in (anti)parallel. As the ship track in the eastern Mediterranean is roughly oriented along 105° we define perpendicular wind direction as $>330^\circ$, $<60^\circ$ or between 150° and 240° . Figure 6.3 shows the fraction of time winds were perpendicular to the ship track for different months and times of day in 2019. We see that these wind directions are most frequent in the morning and evening and less frequent during TROPOMI overpass. We also see that wind directions are more favorable for plume detection in summer months. Furthermore, the figure shows the frequency of cloud free conditions. We see that cloud free conditions are most frequent in June - September when there are almost no clouds, while only 60-80% of the days are cloud free in April, May, October and November. The diurnal cycle of cloud cover is weak.

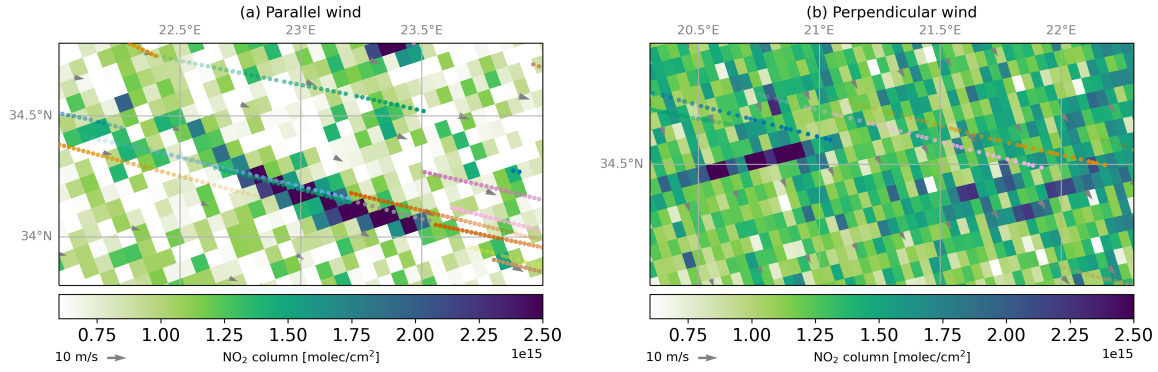


Figure 6.4: TROPOMI observations above the busy ship track for winds along the ship track (a, 9 May 2019) and perpendicular to the ship track (b, 1 August 2019). The colors show the NO₂ column in molec/cm². The colored circles indicate past ship locations, which each color representing a distinct ship and the more transparent circles indicating earlier locations.

6.2.4 TROPOMI

We use TROPOMI L2 data of tropospheric NO₂, freely available on the Copernicus website (<https://dataspace.copernicus.eu>) of version 2.4. From the data we extract qa values, averaging kernel profiles as well as the scattering angle Θ :

$$\Theta = \arccos [\cos \theta \cos \theta_0 - \sin \theta \sin \theta_0 \cos (\phi_0 - \phi)] \leq \Theta_{\max} \quad (1)$$

where θ and θ_0 are the solar and viewing zenith angles and ϕ and ϕ_0 the solar and viewing azimuth angles, respectively. The scattering angle is the angle between satellite viewing direction and sunlight reflected by the earth surface. For a scattering angle of 0°, a smooth sea surface reflects the sunlight directly into the satellite sensor, leading to a higher signal-to-noise ratio. As the water surface is not typically completely flat, this so-called sun glint is also observed for scattering angles up to 20°. For more details on the scattering angle, see Chapter 3.

Figure 6.5 shows the monthly mean ratio of pixels with $qa \geq 0.75$, the scattering angle and the averaging kernel profiles for TROPOMI overpasses in 2019 in the Eastern Mediterranean. We see that clear sky conditions are dominant from June to September, with the fraction of pixels with $qa \geq 0.75$ dropping below 70% in May and November. This is in line with the cloud cover data from ERA5 presented in Fig. 6.3. We also see that the lower sun makes detection of near-surface NO₂ more difficult in April, October and November with very low averaging kernels - meaning low sensitivity of the satellite instrument to NO₂ close to the surface- in the lowest layers. This is likely because of longer atmospheric light paths for low sun angles, leading to more scattering and absorption of the sunlight above the near-surface layers.

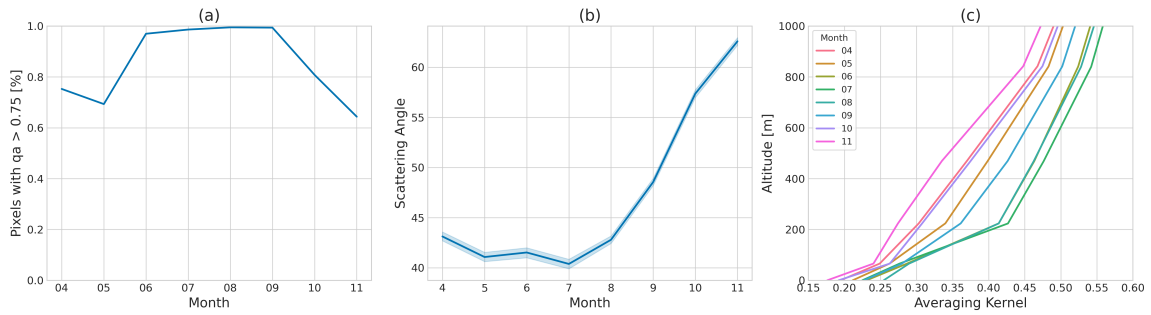


Figure 6.5: Monthly mean ratio of pixels with $q \geq 0.75$ (a), scattering angle (b) and averaging kernel profiles (c) for TROPOMI overpass in 2019 at $33.5\text{-}34.5^\circ\text{N}$, $23.5\text{-}24.5^\circ\text{E}$. For the averaging kernel profiles we used a generic pressure profile to translate the provided level pressures to altitudes.

Table 6.1: Key concepts used in this chapter.

Concept	Definition
Background NO_2 column	the NO_2 column from CAMS data
Detection limit	$0.35 \cdot 10^{15}$ molec/ cm^2 above the background column
max NO_2 enhancement	the maximum NO_2 column in the simulated NO_2 field minus the background column
#significant pixels	the number of pixels in the simulated NO_2 field with NO_2 columns above the detection limit

6.2.5 Key concepts

To facilitate the discussion of ship NO_2 plume detectability, let us define some key concepts. First of all, we discuss simulated in-plume NO_2 columns relative to the out-of-plume columns. These *background columns* are taken from a climatology of CAMS data. We then define the *detection limit* as NO_2 columns with values $>0.35 \cdot 10^{15}$ molec/ cm^2 above the background column. Together with the 20% low bias introduced in the simulations, this is in line with the $0.5 \cdot 10^{15}$ molec/ cm^2 given in the TROPOMI NO_2 ATBD (van Geffen et al., 2021). The *max NO_2 enhancement* is the maximum NO_2 column in the simulated NO_2 field minus the background column. The *#significant pixels* are the number of pixels in the simulated NO_2 field with NO_2 columns above the detection limit. Table 6.1 summarizes these key concepts.

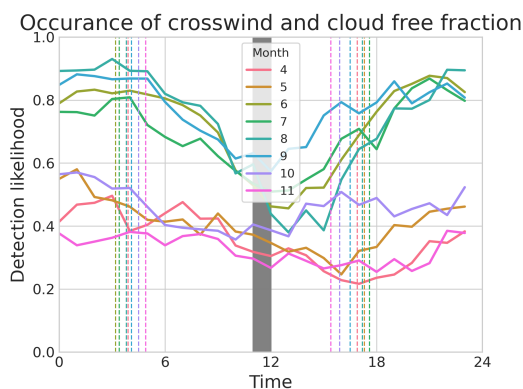


Figure 6.6: Frequency of conditions with winds perpendicular to the major ship lane and cloud free, calculated as the product of the two subfigures in Fig. 6.3. The dashed vertical lines indicate sunset, the gray area indicates typical TROPOMI overpasses.

6.3 Results

6.3.1 Frequency of favorable wind direction and cloud-free conditions

Figure 6.6 shows the frequency of conditions with winds perpendicular to the major ship lane and cloud free conditions. Assuming that plumes can only be detected in cloud-free and perpendicular wind conditions, this gives a baseline for plume detection likelihood, ignoring chemical and other meteorological conditions. We see that in the early morning in June, August and September the baseline conditions are best, with more than 80% of the ERA5 data suggesting favorable conditions. Worst conditions are found in April May and September in the afternoon when only around 30% of the time conditions are good for plume detection in the eastern Mediterranean. During TROPOMI overpass, August and September show the most promising conditions of around 60%, and below 40% are found in April, May, October and November. Overall, ship plume detection is roughly twice as likely in summer than in spring or autumn, taking only wind direction and cloud cover into account. These two factors are important for the retrieval (satellites can not detect NO₂ below clouds) and plume attribution (winds aligned with the busy ship lane are difficult to be attributed to an individual ship). In the following, we will investigate how chemical and additional meteorological parameters influence the plume detectability.

6.3.2 Simulations with varying chemical conditions

Figure 6.7 shows the diurnal and seasonal cycles of the background NO₂ column from CAMS as well as the maximum NO₂ column in the plume and the maximum enhancement of NO₂ column in the plume from PARANOX, using the 2019 diurnal and seasonal cycles in the chemical conditions shown in Section 6.2.2 as model parameters. We see that the

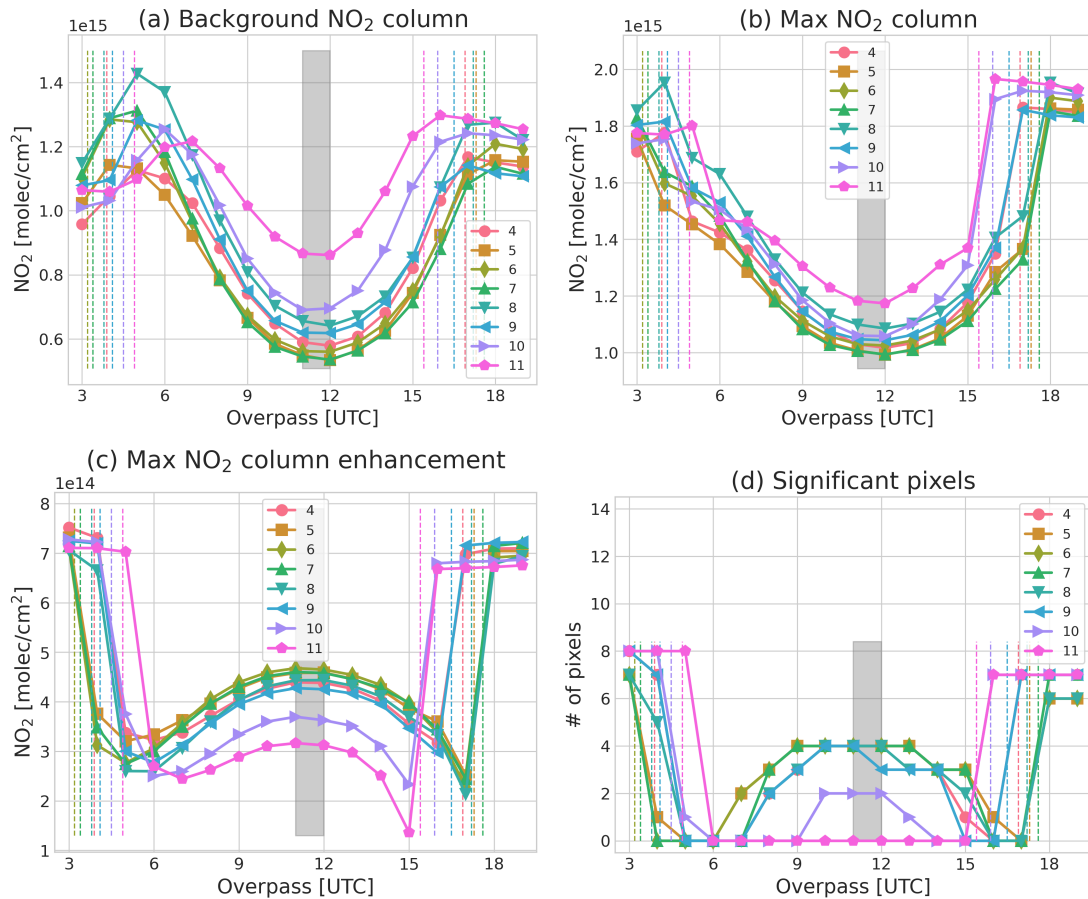


Figure 6.7: Diurnal and seasonal cycles of (simulated) vertical NO₂ columns. (a) background NO₂ column (from CAMS), (b) maximum NO₂ column in the plume, (c) maximum enhancement of NO₂ column in the plume, (d) number of pixels elevated more than $0.35 \cdot 10^{15}$ molec/cm² (the noise level) above the background. Together with the low bias of 20% introduced to the simulations (see Section 6.2.1) this noise level is similar to the $0.5 \cdot 10^{15}$ molec/cm² given in the TROPOMI NO₂ ATBD (van Geffen et al., 2021). The simulations were created with constant meteorological conditions ($T = 295$ K, stability class = E, wind speed = 5 m/s, BLH = 500 m) but seasonal and diurnal cycles of NO_x, O₃ and radiation. The emission strength was set to 60 g NO₂/s.

background NO₂ column shows a clear diurnal cycle, with the lowest values in the middle of the day, around TROPOMI overpass. As discussed before, day-time NO₂ values are lowest in summer, reflecting the low lifetime against oxidation by OH. The maximum NO₂ column in the plume shows a similar diurnal cycle, but less pronounced. In contrast, the plume enhancement is largest in the early afternoon, showing that the NO_x in the background is removed faster than in the plume, in line with findings in Vinken et al. (2011). This is due to ozone titration in the fresh plume, delaying NO₂ and suppressing the availability of OH, which is needed to oxidize the NO₂. Likewise, the number of significant pixels is highest in summer months around local noon time and no pixels are above the noise level in October and November.

6.3.3 Simulations with varying chemical and meteorological conditions

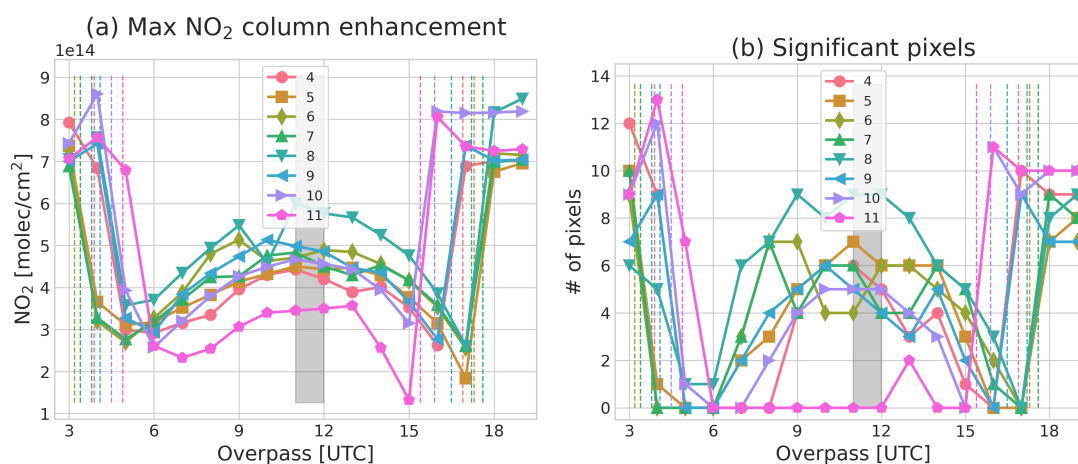


Figure 6.8: As bottom row in Fig. 6.7, but simulated with varying meteorological conditions.

Figure 6.8 shows the seasonal and diurnal cycles of maximum enhancements and number of significant pixels as Fig. 6.7 but with varying meteorological conditions. We see the same tendency to better detectability around midday/early afternoon in summer months as above. The NO₂ enhancements are larger, and we see more significant pixels compared to the simulations with constant meteorology. Most of this is caused by the influence of wind speeds: Figure 6.9 shows simulations with changing meteorology but constant wind speed and is very similar to the simulations with constant meteorological parameters, thus marking wind speed as the most important meteorological parameter, a finding in line with the previous chapter.

Figure 6.10 shows the pseudo-observations of a ship plume over the course of a day in August 2019 used in Fig. 6.9, as well as the chemical conditions within the plume. We see that the plume appears longest in the morning at 8 and 10 UTC, and significantly shorter in late afternoon. This is in line with Fig. 6.9. We also see that NO_x lifetimes in the fresh plume are longest, when all OH is consumed by the high NO_x levels. After about 2 hours, the lifetime reaches a minimum. This can be explained as the NO_x emission

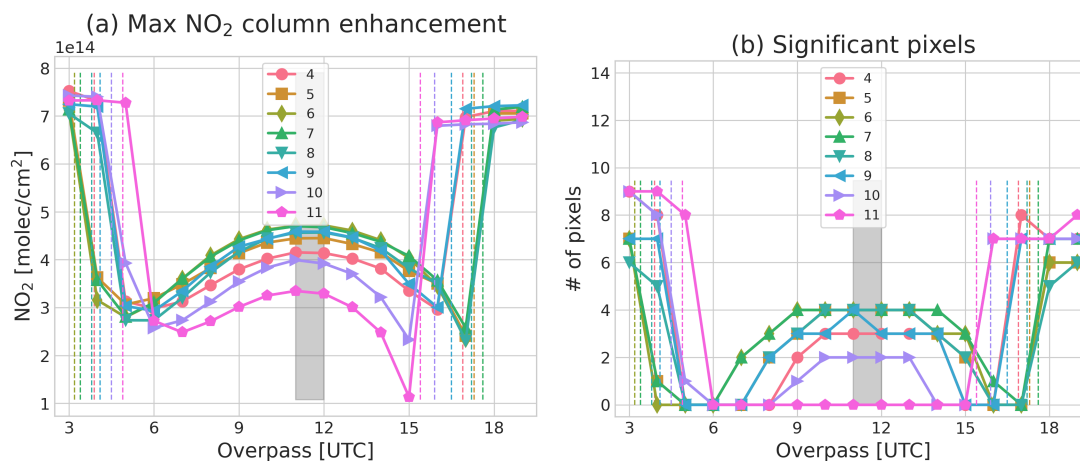


Figure 6.9: As Figure 6.8, but simulated with a constant wind speed of 5 m/s.

lead to ozone production in the aging plume, while being titrated in the fresh plume, together with entrainment of O_3 from outside the plume. Ozone is the most important precursor for OH. The lifetime minimum is shortest at 2 h for afternoon plumes when the high ambient O_3 and available radiation lead to the highest OH values. We also see that lifetimes in the plume center are longer for the fresh plume but reach a lower minimum as the plume ages. In the plume center, concentrations of NO_x are largest, and it takes longer for O_3 and OH to recover. The shorter lifetimes in the plume edges lead to the observed lower NO_2 columns at the edges of the aged plumes in early morning and late afternoon. The NO_2 to NO_x ratio is lowest in the fresh plume, as 95% of the NO_x is emitted as NO, but quickly approaches around 75% during the day (during nighttime, almost all NO_x is present as NO_2 , not shown). The reduction in NO_x concentration and approach of the equilibrium ratio is influenced by the dilution of the plume, entraining cleaner background air.

Figure 6.11 shows the seasonal and diurnal cycles the number of pixels above the noise level ($0.35 \cdot 10^{15}$ molec/cm²) for different emission rates. We see that the number of pixels above the noise level is highest in the early afternoon, when the plume enhancement is highest. This is a strong indication that TROPOMI overpass times are optimal for ship plume detection in the eastern Mediterranean. Trivially, the number of pixels above the noise level is highest for the highest emission rate, and the diurnal cycle is similar for all emission rates. For an emission flux of 60 g/s we find at least 4 pixels above the noise for most of the daytime for all studied months besides November and up to 9 Pixels in August. For emissions of 50 g/s only June and August reach more than 4 pixels. For emissions of 40 g/s only at 11 am in August 2 pixels are enhanced significantly above the background. Not only are more pixels significantly enhanced in summer months, but these enhancements are also visible longer throughout the day. For example, with an emission rate of 50 g/s, the number of pixels above the noise level is at least 4 for 5 hours in August, but only for 1 hour in September.

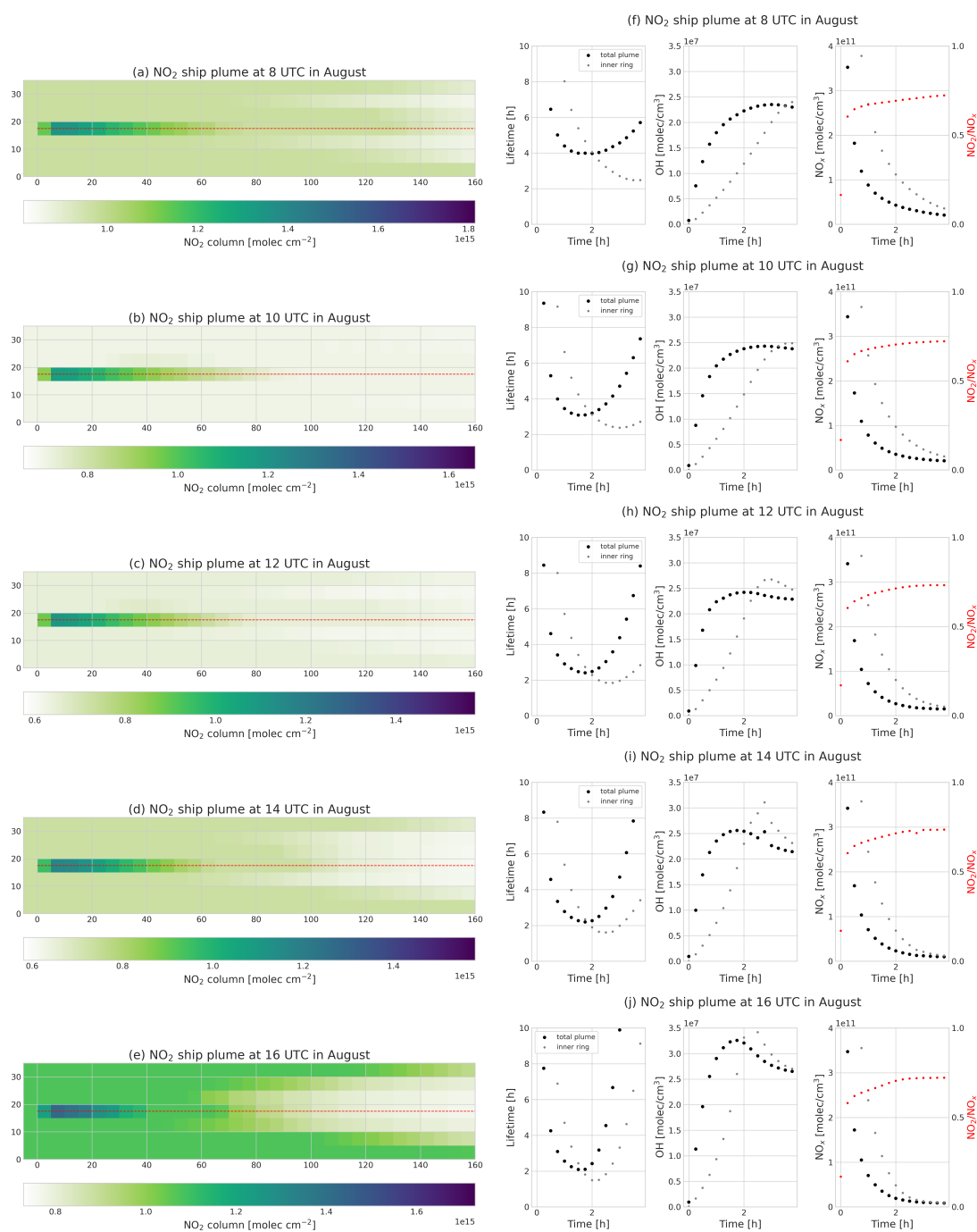


Figure 6.10: (a)-(e): Pseudo-observations of a ship plume over the course of a day in August 2019 (without noise). The colors show the NO_2 column in molec/cm^2 . The dashed red line indicates the ship track. All simulations are created with a constant headwind of 5 m/s, all other parameters are taken from ERA5 and CAMS data as detailed above. To facilitate an easier comparison, the color scale is adjusted for the background NO_2 column. The range is $1.0 \cdot 10^{15}$ molec/cm^2 for all plots. (f)-(j): Plume chemistry for the same simulations: NO_x lifetime as the NO_x concentration divided by the production term of HNO_3 as in (Vinken et al., 2011), middle panel: OH concentration, right panel: NO_x concentration. In black, we show the plume average, in the small gray dots the innermost ring. The red dots in the right panel show the NO_2 to NO_x ratio.

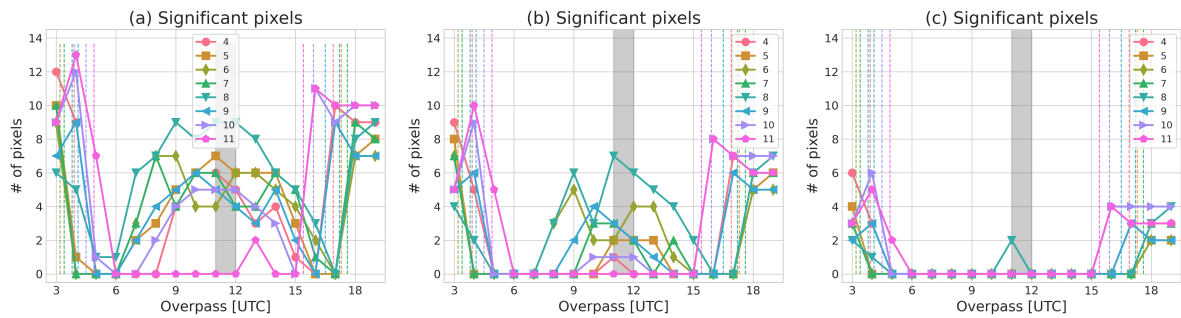


Figure 6.11: Seasonal and diurnal cycles of the number of pixels above the noise level for different emission rates. (a) 60 g NO₂/s, (b) 50 g NO₂/s, (c) 40 g NO₂/s. The dashed vertical lines indicate sunset, the grey area indicates typical TROPOMI overpasses.

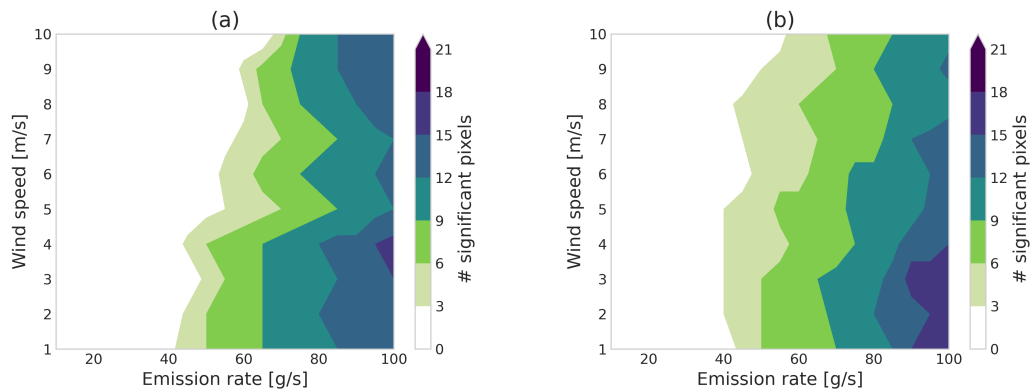


Figure 6.12: Number of pixels more than $0.35 \cdot 10^{-15} \text{ molec/cm}^2$ above the background columns for different emission rates and wind speeds for headwind (a) and wind from the side (b). Simulations were performed for 12am UTC in August, with steps of 1 m/s in wind speed and 5 g/s in emission rate.

6.3.4 Detection dependence on wind speed and emission rate for different wind directions

Figure 6.12 shows the number of pixels above the noise level for different emission rates and wind speeds for headwind and crosswind. We see that plume detection is harder with stronger headwinds: A ship emitting 40 g/s might just be detectable with a slight headwind of 1 m/s, while strong headwind of 10 m/s requires emissions of almost 70 g/s to be detectable. This can be explained by stronger headwinds leading to a more elongated plume, lower enhancements above $0.35 \cdot 10^{15} \text{ molec/cm}^2$, and therefore lower SNR. With crosswind, this dependency is less pronounced: An emission rate of 50 g/s gives the same number of significant pixels for all but very high wind speeds. This is explainable as the plume is less stretched for moderate crosswinds compared to headwinds. Overall, smaller

emissions can be detected for crosswinds than for headwinds of the same strength. This is an additional factor to the easier attribution of the plume to the emitting ship discussed in sec. 6.3.1. For high emissions above 70 g/s, the number of significant pixels is larger for headwinds, likely because the headwinds leads to a longer plume. It should be noted here that we only simulated the plume for 4 hours, giving each plume a finite length and marking an upper limit for the significant pixels, which is higher in the headwind cases. The dip in detectability at 5 m/s compared to 4 and 6 m/s in the headwind case is likely caused by the gridding to 5x5 km² resolution, that - depending on the wind speed - might lead to a favorable or unfavorable alignment of the plume with the grid.

6.4 Discussion & Conclusion

Overall this study shows that from a chemical perspective, the TROPOMI overpass is already at an optimal time of day for NO₂ plume detection over the Eastern Mediterranean, when NO_x background is low and ozone levels are high, favoring the conversion of the emitted NO to observable NO₂. Additionally, the short NO_x lifetime in the early afternoon make overlapping of plumes less likely as they are less long than for example in the early morning. Moreover, measurement geometry is best when the sun is high. However, as Fig. 6.6 shows the wind direction and cloud cover make morning detections & plume attribution more likely: as most ships travel along the same ship lane, situations with perpendicular wind directions make it more easy to avoid overlap of ship plumes and these conditions are most frequent in the morning.

NO_x lifetime in the plume strongly depends on daytime and age of the plume. Overall, lifetimes are the shortest around 2 h after emission in the (late) afternoon. The center of the plume experiences shorter minimal lifetimes, while having longer lifetimes in the fresh plume. This is guided by the titration of O₃ and OH that takes time to recover, especially in the center of the plume where NO_x concentration are highest. In the morning, radiation and ambient O₃ are low, leading to longer lifetimes.

We see that even when ignoring the more cloudy conditions occurring outside June-September (Fig 6.3), chemical conditions are such that with emission fluxes of 50 g/s or below, detection of ship plumes should not be expected in April, October or November. For emissions of 60 g/s, detection is possible between April and October, but not in November. Overall, August shows the best detection likelihood for a sensor with similar resolution and signal-to-noise ratio as TROPOMI.

Besides the influence of chemistry, the wind speed (and direction) plays a significant role: Inside the plume ozone is typically titrated, limiting the NO₂-to-NO ratio in the early stages of the plume. However, stronger wind also leads to more ozone entrainment and subsequently to higher NO₂-to-NO ratios. On the other hand, higher wind speeds lead to a more dispersed plume, both in plume direction as well as perpendicular. For emission fluxes close to the detection limit, this can lead to too low NO₂ columns to be detectable.

For higher emissions, high wind speeds produce a longer visible plume, leading to more significant pixels. This is especially true for crosswinds, that facilitate ozone entrainment while elongating the plume less than headwinds.

The (last) missing piece in this study is a detailed analysis of radiative transfer for the diurnal and seasonal cycle for sun zenith angles and how they influence the respective satellite's sensitivity to NO_2 close to the sea surface. However, it is known (and included in the AMFs) that the lower sun angles especially in the early morning and late afternoon introduce higher uncertainty and increase noise levels (Lorente et al., 2017), making plume detection significantly harder and negating the benefits of favorable wind direction in the morning hours.

While the choice of emission fluxes is in line with those observed in the previous chapter, the wind direction as headwind is somewhat arbitrary and largely influences the amount of significant pixels and the presented results should be seen as a lower end estimate for detectability. However, keeping the wind direction constant facilitates a comparison of how chemical and (other) meteorological conditions change plume detectability between different months.

Some new satellites such as CO2M and TANGO will have a higher spatial resolution, making the disentanglement of plumes easier in theory. How this plays out with the lower signal-to-noise ratio (SNR) is to be seen¹.

Planned upcoming satellite missions will greatly enhance the number of ship plumes observed from space. From a plume-chemistry and dynamics perspective the overpass time of TROPOMI is already optimal for NO_2 plume detection. However, the additional retrieval times will increase the number of samples of plumes from the same ship and thereby reduce the uncertainty when using satellite retrievals of NO_2 for NO_x emission estimates. This is especially true for large emitters, whose plumes are expected to be observed for a large part of the day. For example, a plume caused by an emission of 60 g/s is expected to be observable for up to 10 consecutive hours with a geostationary satellite in summer months. However, a polar orbiting satellite with overpasses 4 hours prior to TROPOMI is expected to capture significantly fewer plumes than TROPOMI. While morning hours exhibit favorable wind directions for plume detection in the eastern Mediterranean, the low sun zenith angles and chemical conditions are expected to negate that benefit. Overall, although one should not expect to detect a ship plume in each scan of a geostationary satellite but mainly close to local noon, they will greatly contribute to better accuracy in our estimates of ship emissions, especially for large emitters.

¹TANGO will have a spectral resolution of 0.6 nm and an SNR of 400 for NO_2 , CO2M will have a spectral resolution of 0.6 nm and aims for an SNR of 1000 vs TROPOMI's spectral resolution of 0.55 nm and SNR of 1500.

Chapter 7

Synthesis

7.1 Introduction

In this last chapter of the thesis we will take a step back and reflect on the presented work in a broader context. To that end we decided to focus on three recurring themes in the thesis: the satellite retrieval of tropospheric NO_2 columns, the use of improved or new methods in ship emission monitoring and atmospheric sciences, and the NO_x emissions from international shipping. There are certainly more themes that could be discussed, such as atmospheric (plume) modelling, however, we decided to focus on these three themes as they are likely to be of interest to a broader audience or of potential special relevance in the future. We will discuss the main findings of the thesis in the context of these themes and draw implications for the future.

7.2 Satellite remote sensing of tropospheric NO_2 and other trace gases

Satellite remote sensing of atmospheric composition with its global coverage and consistent retrieval offers a complementary view on atmospheric composition to in-situ or ground based remote sensing data. In-situ measurements are time-consuming and expensive, creating a spatial bias towards industrialized, wealthy regions. For similar reasons, measurements and monitoring of ship NO_x emissions on open sea are sparse and limited to coastal regions close to developed countries. In this thesis, we therefore used TROPOMI satellite data to study ship emissions in the open sea. However, the use of satellite data comes with its own set of challenges, which we will discuss in this section. Furthermore, we will discuss the implication of the findings of this thesis with regard to ship NO_x emission monitoring and the future of satellite remote sensing of atmospheric composition.

7.2.1 This thesis

The TROPOMI retrieval of tropospheric NO_2 columns

In this thesis we addressed some of the most important error and bias sources in satellite retrieval of tropospheric NO_2 columns. In Chapter 3 we investigated retrievals under sun glint, which have previously been flagged with a lower quality assurance value, but at the same time have proven promising in detecting ship plumes, as shown by Georgoulias et al. (2020) and in Chapter 3. We found that retrievals under sun glint have a higher scene albedo and elevated averaging kernels, meaning a better sensitivity to near-surface NO_2 . This benefit is strongest at moderate wind speeds, when the sea surface is sufficiently wavy to create a large sun glint area, but not so rough that it ceases to act like a mirror. We found that sun glint gives higher (geometry-normalized) slant columns while the AMF does successfully account for the increased photon path length.

In the same chapter, we looked at the impact of a new cloud algorithm, that was being

tested at the time of writing. The improved, or rather repaired cloud retrieval algorithm was expected to solve the bias in the previous TROPOMI versions v1.2 and v1.3, where the cloud pressure was too high, causing too low NO₂ column densities in the retrieval. We found that the updated algorithm indeed increases cloud height and subsequently the NO₂ column densities, which was later confirmed in Latsch et al. (2022) and van Geffen et al. (2022).

While the retrieval method for satellite remote sensing of tropospheric composition is consistent globally, the accuracy of the retrieval results are not. In Chapter 4 we investigate the role of vertical profiles of NO₂ that are needed as a priori information in the TROPOMI NO₂ column retrieval. We find that over the North Sea, models overestimate the vertical mixing and marine boundary layer height, giving too low NO₂ concentrations near the surface and too high concentrations at 150-400 m altitude. This leads to a systematic negative bias in NO₂ columns of 20%. We also stress the specific difficulty with trace gas retrieval over sea, as the sea surface is dark compared to typical land surfaces, leading to very low sensitivities to NO₂ in the near-surface layer. Additionally, in-situ data for model and satellite validation are sparse over sea. We show that using CAMS profiles as a priori as in Douros et al. (2023) gives higher TROPOMI NO₂ columns and therefore a better agreement with the aircraft-based data over the North Sea, but likely because of higher a priori columns instead of the correct profile shape.

On the global consistency of satellite data

We conclude that a priori profiles remain an issue for trace gas retrievals and subsequently emission estimates based on remote sensing data: while the retrieval algorithm is the same globally, the data quality is not. It depends on atmospheric transport models that might have larger biases in some regions (e.g., over sea as shown in Chapter 4 and over mountains (Rotach et al., 2022)) than in others. Additionally, the models are run with emission estimates that usually come from bottom-up emission inventories with spatially varying quality (Elguindi et al., 2020; Flerlage et al., 2021). In the case of ship emissions, these are time averaged and therefore not a satisfactory solution to monitor emissions of individual ships. Furthermore, the radiative transfer - and thus the satellite's sensitivity to near-surface NO₂ - depends on the surface reflectance.

The TROPOMI NO₂ product therefore has spatially varying uncertainties. But in an inversion, even if these factors are fully taken into account, the inverted fluxes will have less information gain in regions with high uncertainties. It is therefore unlikely that inversion systems based on TROPOMI or other satellite sensors produce emission estimates of globally consistent quality. More research efforts should therefore go into regions with sparse in-situ data and large uncertainties in satellite retrievals. Despite this, satellite remote sensing data might still be the best tool for global atmospheric composition and inversion studies.

The detectability of ship plumes

In Chapter 5 we study NO₂ plumes of individual ships. We find that the observable NO₂ - besides the underlying emissions - mainly depend on ozone levels in the background air and the wind speed. Therefore, the detection limit of ship emission varies with space and time and the assessment of air quality or underlying emission will, too. This is further studied in Chapter 6 where we study the diurnal and seasonal cycle of ship plume detectability. Even when ignoring the more cloudy conditions that prevent high-quality NO₂ column retrieval, seasons outside of summer show unfavorable conditions to detect ship plumes in the eastern Mediterranean.

7.2.2 Outlook

The retrieval of tropospheric NO₂ columns

In addition to the retrieval aspects touched upon in this thesis, there are further potential improvements to the TROPOMI algorithm. One example is the impact of vibrational Raman scattering in liquid water on the slant column retrieval (Lampel et al., 2015; Richter et al., 2011). In the OMI QA4ECV retrieval this was implicitly taken care of by the intensity offset (Boersma et al., 2018), which is not included in the TROPOMI retrieval (Van Geffen et al., 2020). While this offset does not have a clear physical meaning, it has been shown to reduce the residual in the NO₂ retrieval connected to vibrational Raman scattering, especially over sea (Oldeman, 2018). Furthermore, while the latest TROPOMI NO₂ version v2.4 does include directional surface reflectance over land, the complex reflectance of water surfaces, dependent on wind-induced waves (see Cox and Munk (1954)), is not treated. Including a more realistic Bidirectional Reflectance Distribution Function (BRDF) over water that includes such wind and wave effects would improve the surface albedo and therefore the AMF calculation (Zhou et al., 2010), also beyond sun glint effects which are already treated appropriately by the current retrieval, as shown in Chapter 3.

Future mission of satellite remote sensing instruments for atmospheric composition

Besides these potential improvements in the TROPOMI algorithm, this decade will see a lot of exciting changes in the field of space based monitoring of atmospheric composition and underlying emissions. In 2020, GEMS, the first geostationary satellite to retrieve NO₂ columns, has been launched and has operational hourly data over Asia available since November 2022 (Kim et al., 2020). TEMPO, a comparable satellite launched in April 2023, will provide hourly NO₂ over North America (Zoogman et al., 2017). In 2025, Sentinel4 - a geostationary satellite over Europe - and Sentinel5 - a polar orbiting satellite with morning overpasses - are scheduled to be launched (Ingmann et al., 2012). These satellites

will significantly enhance the space-based monitoring capabilities by the increased spatial sampling. This might result in a larger number of ship plumes being captured and a significant reduction of uncertainties if the signal-to-noise ratios are sufficiently high. Missions like TANGO - which aim at using a pair of satellites to measure both NO₂ and CO₂ - at very high spatial resolution (200x200 m) will take further steps, using the good detection capabilities for NO₂ sensors to detect CO₂ plumes (Landgraf et al., 2020). The Japanese GOSAT-GW satellite, set to launch in 2024, has a similar setup and will measure NO₂ columns alongside GHGs, alternating between a wide (10 km resolution) and focus (3 km) mode (Matsunaga et al., 2022). This will be a significant step towards monitoring the emissions of greenhouse gases from space as demonstrated in Zhang et al. (2023), using coinciding TROPOMI NO₂ and OCO-2 CO₂ satellite data. For methane, an operational GHG monitoring system is already in place (Schuit et al., 2023; Maasakkers et al., 2022). Those instruments and the emission monitoring systems based on them can play a large role in the global stock take of GHG emissions as part of the Paris Agreement. However, one of the downsides of satellite-based atmospheric composition and emission monitoring will remain: these UV/Vis backscatter satellite sensors require daylight and can therefore not be used to monitor emissions at night. This is less of less importance for sources like shipping that we do not expect to have a strong diurnal pattern, but more relevant for urban or biogenic emissions with a clear diurnal profile (Squires et al., 2020; Visser et al., 2021).

The challenge of higher spatial resolution

If higher resolution satellites have a sufficient signal-to-noise ratio, detection of plumes will become easier as the contrast between the plume and the background air will increase. This will make it easier to detect smaller sources and their emissions or distinguish plumes of two nearby sources, for example ships sailing close to one another. TROPOMI and the work presented in this thesis represented some steps in the transition from using averaged satellite data to study long-term emissions and their trends as done in Chapter 3 for major shipping lanes, to daily - or soon hourly - emissions of (moving) point sources. This will also make it easier to pinpoint sources and potentially distinguish between sectors of emission.

However, the higher resolution will make the radiative transport analysis more difficult, as a spatial smearing out to neighboring pixels will increase uncertainties caused by 3D radiative effects such as cloud shadows and horizontal light scattering across pixels (Wagner et al., 2023). This inherent problem sets a practical limit to the spatial resolution of independent pixels, which depends on the wavelength used and therefore the molecule studied.

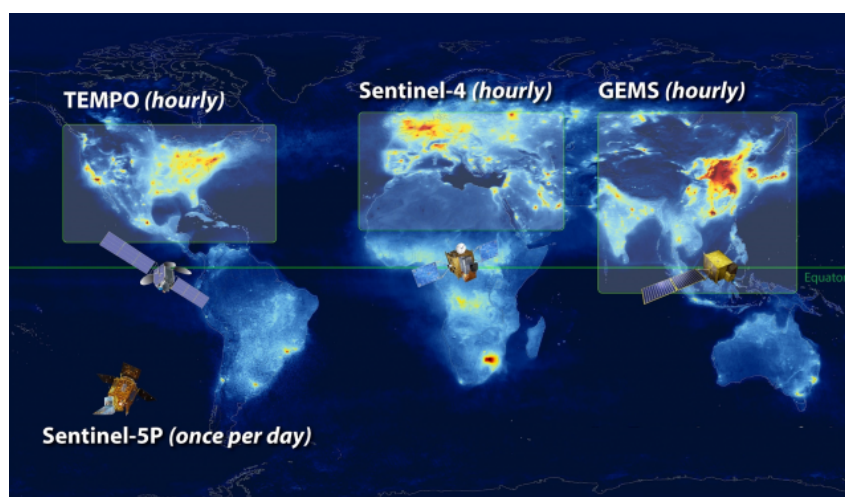


Figure 7.1: Coverage of the geostationary satellites GEMS, TEMPO and Sentinel-4 as well as Sentinel-5P/TROPOMI. The latter is a polar orbiting satellite and therefore has a global coverage. The geostationary satellites are limited to the Northern Hemisphere. Credits: NASA

Atmospheric chemistry and transport models

Furthermore, observations at higher resolution will also increase the need for better models at higher spatial resolution for the AMF calculation and the interpretation of observations. Only at higher spatial resolution can we properly capture the interplay of turbulence and atmospheric chemistry at smaller scales than we addressed in this thesis. LES models such as MicroHH used in Chapter 5 are a promising tool for this (Ražnjević et al., 2022). For the time being, computational costs of running global or even regional simulations using LES are too high to make them feasible for operational use. However, a look-up-table approach could be used to derive generic corrections for NO_2 columns close to point sources. But even that approach produces high computational costs. Additionally, atmospheric chemistry models should move towards including halogen chemistry in the marine boundary layer, which can reduce NO_x pollution by up to 20% (Li et al., 2021), leading to underestimations of emission when not included.

The lack of geostationary satellites in the Southern Hemisphere

Looking at these new and upcoming missions in Figure 7.1, the lack of geostationary satellites and therefore hourly satellite data of trace gases in the Southern Hemisphere becomes obvious. Its global coverage and consistent retrieval methods are some major advantages and selling points for the use of satellite data compared to in-situ data. While polar orbiting satellites will remain as part of the fleet, the Northern Hemisphere will have a significantly higher temporal sampling of air quality data. New satellite missions will thus partially re-create a bias towards wealthy countries that earlier missions closed.

General outlook on satellite remote sensing of atmospheric composition

The amount of data becoming available with the new geostationary or high-resolution polar orbiting satellites will also require new methods of data processing. Machine Learning and Artificial Intelligence will likely play a large role in this, as we will discuss in the following subsection.

All this opens up the question of how the wealth of data from numerous new satellites can be utilized to significantly improve global emission estimates. While the increased spatial and temporal resolution will reduce uncertainties of emission estimates, the inherent challenges that come with the use of satellite data - such as the need for accurate a priori profiles, the treatment of clouds, and the radiative transfer analysis - will remain. Additionally, satellite remote sensing needs to go hand in hand with ground-based and in-situ measurements for validation purposes and process understanding at smaller scales. The increasing number of satellites will certainly enhance the need for cross-validation of the satellite products, too. If these steps are not forgotten, the future of satellite-based emission monitoring is bright, but the challenges are manifold and will require a concerted effort from the scientific community to overcome.

7.3 Methodological advances

Having discussed the uncertainties in and future of satellite remote sensing of atmospheric composition, we will now turn to methodological innovations made in this thesis and foreseen for the future. These include the use of Machine Learning and Artificial Intelligence - arguably *the* buzzwords of the last years - as well as the use of aircraft data and plume resolving models for the inversion of NO₂ columns to NO_x emissions. We will end this subsection with a discussion of future validation and improved AMF calculations for ship plume monitoring.

7.3.1 Machine Learning and Artificial Intelligence

The last decade has seen a surge in the use of Machine Learning and Artificial Intelligence in atmospheric and climate sciences (e.g. Ladi et al., 2022; Schneider et al., 2022). More specifically in the field of remote sensing of atmospheric composition, the large amounts of data have spurred research not only into automated detection of plumes (Kurchaba et al., 2022; Schuit et al., 2023; Finch et al., 2022), but also inferring surface concentrations using NO₂ column data (Kang et al., 2021; Kuhn et al., 2024).

In this thesis, two Machine Learning algorithms are used: firstly, in Chapter 3 we use a deep neural network to correct historical TROPOMI NO₂ data for the updated cloud retrieval and therefore create a consistent data record, anticipating the scheduled re-processing of all orbits. Secondly, in Chapter 5 we used an XGBoost model to learn the

relationship between modelled NO₂ columns and underlying NO_x emissions in order to invert TROPOMI NO₂ columns to NO_x emissions. To this end, we trained the model with a large library of simulated NO₂ plumes, which represent a large set of chemical and meteorological conditions that influence the shape and magnitude of the observable plume.

Computational efficiency and generalization

The biggest advantage of Machine Learning in the context of this thesis is the computational efficiency. Instead of having to wait for a re-processing of the TROPOMI data, we generalized the effect of the updated retrieval algorithm by identifying the key features of the cloud retrieval that influence the NO₂ columns. This allowed us to correct the data in a fraction of the time it would take to re-process the data, and we could analyze trends in shipping NO₂ over a period that was not covered by the same operational TROPOMI product at the time of publication. In the case of the inversion, we also used a Machine Learning tool for generalization, now in the relationship between NO₂ columns and NO_x emissions. We used a 10-dimensional parameter space for the creation of the plume library, meaning that a classical look-up-table approach would have been computationally infeasible: even with only 5 values per parameter, the library would have contained $5^{10} = 9.8 * 10^6$ entries. With a simulation time of 2 minutes per run, the look-up-table would have taken about 37 years to create. The XGBoost model however was able to correctly interpolate between the nodes in the parameter space with only 40,000 simulations and estimate the emissions for any plume observed by TROPOMI, as shown by the good performance of the model on the test data.

The reliance on physical models and domain knowledge

Both our use-cases have in common that we trained the Machine Learning models with data that was produced with a physical model: the TROPOMI NO₂ retrieval using the updated cloud algorithm FRESKO+wide and PARANOX inputs and outputs, respectively. This shows that the use of Machine Learning tools will not overcome the need for extensive validation and physics-based models. While it can speed up the computation immensely, it can confirm biases inherent to the physical models in the results as a ground-truth is often missing. We have tried to avoid this pitfall by using MicroHH for validation of both PARANOX and the trained XGBoost model. As will be discussed below, an extensive ground-based validation would be beneficial.

Looking at the use of Machine Learning in atmospheric sciences outside this work, models are often trained with data labeled by humans, which is time extensive, often requires experts, and is not always objective. Experts are also required to choose the right features for the model. Another point of attention in the use of Machine Learning in atmospheric science - and in science in general - is, whether a black box model can actually contribute to process understanding or if the benefit remains at the computational efficiency. In this

thesis, we have tried to understand and explain the behavior of Machine Learning models with our existing knowledge on radiative transfer, satellite retrievals and atmospheric plume chemistry and dynamics. This shows that when guided by domain knowledge, Machine Learning can be a very powerful tool in every scientist's toolbox.

Overall, the use of Machine Learning in the field of atmospheric science is - while growing rapidly - still a new field, and many questions remain open: are the models interpretable? Do they give physically consistent results? Can models trained on historical data be used for the future, how accurate will ML-based weather predications be in a new climate? Can it predict not yet seen extremes? These questions are not unique to atmospheric science, but are especially relevant in a field where the consequences of wrong predictions can be severe. However, with the growing data volumes and need to computations efficiency, they offer unique opportunities.

7.3.2 Using aircraft data to evaluate TROPOMI data over sea

In Chapter 4 we used aircraft data of NO_2 to create vertical NO_2 profiles. These profiles are then used to evaluate the modelled a-priori profiles used in the retrievals as discussed in Section 7.2.1. To arrive at vertical profiles that are representative at the spatial scale of a TROPOMI pixel, we combined the raw measured data with information on ship (plume) density and modelled $\text{NO}-\text{NO}_2$ ratios for plume of various ages. Only then could we compare the aircraft data to the modelled profiles and find a systematic bias in the modelled profiles. This method could be used to validate other satellite data over sea, where in-situ data is sparse.

7.3.3 Plume resolving models

In this thesis we used two plume resolving models: the dedicated (Gaussian) plume model PARANOX (Chapters 4, 5 & 6) and the LES model MicroHH (Chapter 5). While PARANOX has been used to model (ship) NO_x plumes in the past (e.g. Vinken et al., 2011), the NO_x chemistry in MicroHH has only been implemented recently (Krol et al., 2024) and is used here for ship plumes for the first time. While we mainly rely on PARANOX for the inversion of TROPOMI NO_2 columns to NO_x emissions in Chapter 5, MicroHH is used as a validation for the Gaussian model. Only with the help of these models and their NO_x chemistry could we estimate the NO_x emissions of individual ships.

Validation of satellite data and models for ship NO_x emission monitoring

To better estimate NO_x emissions of ships using current and upcoming satellite data, a measurement campaign combining several methodologies should be done: (a) installing NO_x measurements at the stack of selected ships, (b) using drones or airplanes to measure NO and NO_2 in the aging plume and outside up to several hours downwind and

taking vertical profiles, (c) MAX-DOAS instruments measuring vertical columns in and around the ship plumes. These measurements should be done coinciding with satellite overpasses of TROPOMI and other sensors. (a) can serve as a ground-truth to compare satellite-based emission fluxes with. (b) is similar to the campaign used in Chapter 4 which could serve as a recipe for the analysis, but should be conducted with higher temporal sampling and measuring NO and NO₂ simultaneously. The data can then be used to validate (plume) models such as PARANOX or MicroHH as well as chemical transport models used for a priori profiles in the retrieval. Furthermore, the campaign should take place in the Eastern Mediterranean or another region where ship plumes are frequently observed in TROPOMI NO₂ data. (c) can be used as an additional validation of the satellite columns. Combining such a campaign with bottom-up estimations of ship emissions based on engine and ship specifics would provide a rich data set. It would allow for a real evaluation not only of TROPOMI NO₂ columns but also of the inferred emission flux. The proposed campaign would be in addition to ongoing and planned satellite validation efforts with operational ground based networks. For example, TROPOMI column data has been heavily validated against networks of MAX-DOAS and TCCON measurements (Verhoelst et al., 2021; Lorente et al., 2021; De Smedt et al., 2021). For the North Sea, a satellite validation platform is planned, designed to measure vertical distribution and columns of several trace gases among which NO₂ and SO₂ (KNMI, 2024). Compared to these validations of satellite products, the proposed campaign would have the advantage of being targeted at ship emissions and could additionally be used to validate the (plume) models and inverted emission fluxes.

Dynamical AMF calculation for ship plumes

Additionally, TROPOMI-based monitoring of shipping NO₂ would largely benefit from a dynamical AMF approach: using plume-resolving a priori NO₂ profiles instead of the coarse resolution TM5 (or CAMS) data. In these models ship emissions are represented as time averages along the ship track while in reality emissions are highly variable. Furthermore, the plumes are immediately diluted over the large grid cell. In Chapter 5 of this thesis we use a PARANOX-simulated NO₂ profile to recalculate the AMF for in-plume pixels, arriving at a generic correction factor of 1.2 in line with findings of Chapter 4. This approach should be extended with a library of ship plumes of varying emission strength and meteorological and chemical conditions. For each ship plume visible in the slant columns and coinciding with AIS data, the most appropriate library member could be chosen based on reanalysis data of the atmosphere and a ship emission model such as STEAM (Johansson et al., 2017). For in-plume pixels the vertical NO₂ columns could then be adjusted using these simulated profiles. Alternatively, one could use AIS data and STEAM to simulate ship plumes in a high resolution chemical transport model. However, this has the risk that observed and simulated plume locations do not match, introducing new biases. To circumvent this problem, wind speed and direction in the transport model

could be adjusted to match the plume location. Either of these approaches would allow for a more accurate AMF calculation and therefore a more accurate NO₂ column retrieval over sea and subsequently more accurate NO_x emission estimates.

7.4 NO_x emissions from international shipping

7.4.1 This thesis

The shipping sector is a major source of NO_x emissions, especially in coastal regions and near major ports. In this thesis, we started in Chapter 3 with the detection of major shipping lanes around Europe using the - at that time - novel TROPOMI data. Some of these lanes have not been shown in satellite data before, showcasing the abilities of TROPOMI for ship emission monitoring. In the same chapter we grabbed the opportunity that COVID handed us in the form of a natural experiment and studied the impact of the pandemic on ship NO_x emissions around Europe. The inferred reduction of 20% in a single year was sizeable but smaller compared to reduction in other anthropogenic emissions, such as car traffic (Guevara et al., 2021). This is not surprising as international trade decreased less than for example car commutes to work with the strict contact restrictions that made home office the norm in many jobs, at least temporally. We then compared these top-down reductions to bottom-up trends, making use of AIS data. We found that the timing of the reduction was consistent, while the magnitude in the top-down approach was larger, possibly caused by our simplified treatment of NO_x chemistry.

In the computation of the bottom-up emission changes we assumed constant emission factors of Ocean Going Vessels (OGVs), which is reasonable for the study period of 2 years. On longer timescales and if IMO regulations have the intended effect, we expect these emission factors to change. The presented methods can therefore be used to study the effectiveness of IMO regulations on a fleet level: if newer ships are indeed cleaner, we expect a decrease in emission factors and therefore a decrease in top-down emission compared to our bottom-up emission assuming constant emission factors.

With Chapter 4 we shifted the focus to individual NO₂ plumes and the emissions that caused them. After identifying and matching 130 plumes to individual ships (mainly container ships of 300-400 m length), we estimated their emissions with values of up to 300 g/s. Additionally, we find that, contrary to expectations (and legislature), newer Tier II ships have higher NO_x emission factors compared to older Tier I ship, and also higher emissions per distance travelled. Furthermore, the presented methods set the stage for a new, potentially global and automated way of monitoring ship emissions from space, combining satellite data with plume models and Machine Learning.

7.4.2 Implication and outlook

Emission trends

Many sectors have been decreasing their nitrogen oxide emissions in the past decades (Krecl et al., 2021; European Environment Agency, 2019a). In fact, nitrogen oxide pollution is decreasing over Europe by up to 2% annually (Aas et al., 2024). This trend is likely to continue with the phase out of fossil fuel combustion and the introduction of renewable energy sources or electric vehicles. In contrast, the shipping sector has had rather constant emissions (Boersma et al., 2015) and is expected to rise with increasing trade volumes (Müller-Casseres et al., 2021). This will make the shipping sector an even more important source of nitrogen oxide pollution in the future.

As a result, the shipping sector will experience increased pressure to play its part in cleaning up the air. For example, since the beginning of 2024, GHG emissions from international shipping are included in the European Union's Emission Trading System (ETC) (European Commission, 2024), likely leading to the use of alternative fuels with potentially lower NO_x emissions as discussed below. This is reflected in the IMO's GHG reduction plans (International Maritime Organization, 2020).

New fuels and abatement technologies

There are several possible alternative fuels to replace the currently common heavy fuel oils. Al-Enazi et al. (2021) have identified LNG, hydrogen and ammonia as the most promising candidates. While natural gas is generally cleaner and produces lower emissions of CO₂, SO_x and NO_x, it is a fossil fuel and therefore not a long-term solution. Hydrogen and ammonia are both carbon-free and can be produced using renewable energy sources. However, the infrastructure for these fuels is not yet in place and the production of hydrogen and ammonia is energy-demanding and currently not carbon-free. Additionally, hydrogen comes with other obstacles, such as suitable storage. Ammonia has the advantage that engines can easily be made compatible to it without major modifications (Al-Enazi et al., 2021). On the other hand, using the traditional Haber-Bosch process in the production of ammonia creates large amounts of GHG emission, and NO_x emissions during the use of ammonia fuels are high (Gubbi et al., 2023), making the use of abatement technologies such as exhaust gas recirculation (EGR) or selective catalytic reduction (SCR) necessary. Furthermore, ammonia is toxic, adding further complication to the transport and storage (EMSA, 2023). Overall, the low supply - especially for hydrogen - combined with long lifespans of vessels (typically 20 to 30 years (Marshallsay, 2023)), the long fleet turnover times of the shipping industry, as well as storage and transportation challenges will make a shift towards clean fuels unrealistic in the short term. On a longer timescale, NO_x plumes seen by satellites might become a less relevant topic if energy production (and ship propulsion) shifts towards (non-ammonia) renewables and anthropogenic NO_x emissions decrease.

Ship pollution in ports and coastal areas

Besides new fuels or abatement technologies there are also low-tech solutions to reduce air pollution caused by ships: when at berth, cruise ships often have their engines running to produce electricity to power the ship, producing avoidable air pollution (Lansø et al., 2023; Hulskotte and Denier van der Gon, 2010; Dragović et al., 2018). This is especially problematic when in the harbor of large cities or inland waterways such as Rotterdam or Venice and can be avoided by using on-shore power supply. The port of Rotterdam therefore recently presented plans, aiming at 90% of ships using on shore power by 2030 (Buitendijk, 2024). This is a simple and effective way to reduce air pollution in densely populated areas.

Even the increasing spatial and temporal resolution of new and upcoming satellites will not be able to detect all emission sources. In the case of shipping, pollution in ports is of special interest, as it affects human health and air quality in densely populated areas. However, in ports typically many ships are close to one another, making it impossible to distinguish individual plumes using satellite imagery. Additionally, other emission sources such as road traffic or industry are often close by, further complicating the detection of ship plumes. For that reason, a combination of satellite data with in-situ or ground based remote sensing measurements (e.g. Van Roy et al., 2022c; Ripperger-Lukosiunaite et al., 2023; Mettepenningen et al., 2024; Krause et al., 2021) will be necessary to monitor emissions from shipping and air quality in these areas. This is true not only for shipping but for all sectors emitting in industrial areas.

Other coastal areas might be more promising, if emission from other sources are sufficiently far away. Even in busy ship tracks, repeated measurements could decrease the issue of overlapping or crossing plumes. Firstly, with more observations it becomes more likely that the plumes of two ships sailing antiparallel to one another are distinguishable. Additionally, inversions of crossing plumes - while not attempted in this thesis - are not impossible, but connected to higher uncertainty. With repeated observations provided by geostationary satellites, this uncertainty would decrease.

Sulfur emissions

The shipping industry is also a prime example on how anthropogenic interference can have unforeseen and unwanted consequences. With the introduction of the global sulfur cap in 2020, the sulfur content in marine exhaust had to decrease from 3.50% to 0.50%. As sulfur oxides are a major source of aerosols, this led to a decrease in the number of aerosols in the atmosphere, which in turn is observed to have decreased the cloudiness over the Northern Atlantic and possibly caused record-high surface ocean temperature anomalies (Diamond, 2023; Yoshioka et al., 2024). While this should certainly not be used as an argument against air quality regulations, it shows that human influence on the environment has become so large that even small changes in our behavior can have large consequences. It also shows that atmospheric processes and climate feedbacks are not yet

fully understood and emphasizes the care we should use when discussing manipulating the climate system further.

The (in-)effectiveness of IMO NO_x regulations

Our finding that newer ships have higher NO_x emission factors than older ships is concerning. This finding is especially true for ships running at lower engine loads, which was the main mode of operation found in our dataset. While we cannot draw conclusions from our high-emitting dataset to the whole international fleet, it is a worrying finding: the IMO regulation was thought to lead to less air pollution and our study provides indications that regulations are not effective in reducing NO_x emissions from ships. This trend was predicted in Cheng et al. (2017), who show that the optimization of ship engines in regard to fuel consumption might have the unwanted side effect of higher NO_x emissions. They criticize the high weighting factor of 0.5 for engine loads of 75% in the IMO regulation, which does not represent the real-world sailing behavior of ships with the trend to slow-steaming and subsequent lower engine loads for economic reasons (see Fig. 7.2). This aligns with our finding and shows that the real-world emissions of ships are not well represented in the current MARPOL Annex VI. The regulation might therefore fail its goal to limit air pollution.

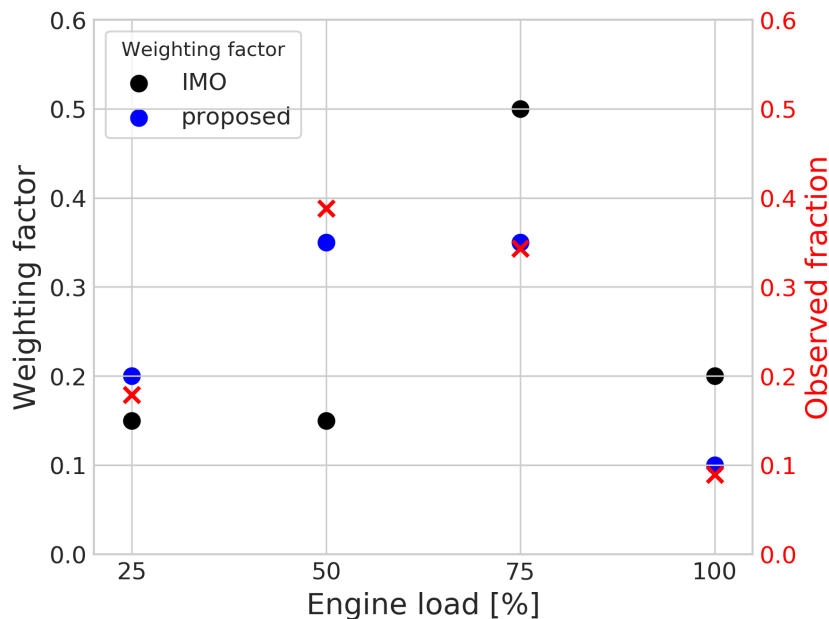


Figure 7.2: Weighting factors for NO_x emissions of ship main engines in the IMO Tier II regulation (black (International Maritime Organization, 2008)) and operational behavior observed in the Mediterranean in Chapter 5 (red). The highest weight is given to 75% engine load, while we observe most ships to sail at lower engine loads. The blue dots are proposed weighting factors that would better represent the real-world operations of ships.

7.4.3 How to improve the IMO regulation

After having identified the shortcomings of the IMO NO_x regulation, we want to end this section (and thesis) with suggestions on how to improve it.

The easiest and probably most effective change would be to add a not-to-exceed limit on NO_x emission factors for Tier II ships to the existing regulation on a weighted average, taking a step towards the envisioned lower emissions of these newer vessels as also recommended in Verbeek et al. (2022). This should be accompanied by the requirement to send the current power usage along with the AIS data, so that the emission factors (in g/kWh) can be calculated from inferred emission strength (in g/s), instead of estimating the power used from design and current vessel speed. This would make it easier to monitor the compliance with the regulation from space, as we have shown in this thesis, or using ground-based remote sensing or in-situ data. Currently, the regulation is almost impossible to enforce outside the laboratory and allows for non-compliance to go undetected. Furthermore, a regulation that limits effective emissions, for example per kilogram freight and kilometer travelled would be more sensible than the current formulation.

If the weighted average and emission factor focussed nature of the regulation is to be kept, the IMO should at least increase the weighting factors for low engine loads to fit the real-world operations of ships while reducing the weights for 75% and 100% engine loads as proposed in Figure 7.2. At the moment, the highest weight is given to the high engine loads even though this does not reflect real-world operations. This would make the regulation more effective in reducing NO_x emissions, and in this thesis we have shown that satellite remote sensing is equipped to play its role in this effort.

References

- Aas, W., Fagerli, H., Alastuey, A., Cavalli, F., Degorska, A., Feigenspan, S., Brenna, H., Glib, J., Heinesen, D., Hueglin, C., Holubová, A., Jaffrezo, J.-L., Mortier, A., Murovec, M., Putaud, J.-P., Rüdiger, J., Simpson, D., Solberg, S., Tsyro, S., Tørseth, K., and Yttri, K. E. (2024). Trends in Air Pollution in Europe, 2000–2019. *Aerosol and Air Quality Research*, 24(4):230237. Publisher: Taiwan Association for Aerosol Research.
- Acarreta, J. R., De Haan, J. F., and Stammes, P. (2004). Cloud pressure retrieval using the O₂-O₂ absorption band at 477 nm. *Journal of Geophysical Research: Atmospheres*, 109(5):1–11.
- Agrawal, H., Welch, W. A., Miller, J. W., and Cocker, D. R. (2008). Emission measurements from a crude oil tanker at sea. *Environmental Science and Technology*, 42(19):7098–7103.
- Al-Enazi, A., Okonkwo, E. C., Bicer, Y., and Al-Ansari, T. (2021). A review of cleaner alternative fuels for maritime transportation. *Energy Reports*, 7:1962–1985.
- Andreadis, P., Zompanakis, A., Chryssakis, C., and Kaiktsis, L. (2011). Effects of the Fuel Injection Parameters on the Performance and Emissions Formation in a Large-Bore Marine Diesel Engine. *International Journal of Engine Research - INT J ENGINE RES*, 12:14–29.
- Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., and Wagner, T. (2019). Pinpointing nitrogen oxide emissions from space. *Science Advances*, 5(11). Publisher: American Association for the Advancement of Science.
- Beirle, S., Platt, U., von Glasow, R., Wenig, M., and Wagner, T. (2004). Estimate of nitrogen oxide emissions from shipping by satellite remote sensing. *Geophysical Research Letters*, 31(18):4–7.
- Berg, N., Mellqvist, J., Jalkanen, J. P., and Balzani, J. (2012). Ship emissions of SO₂ and NO₂: DOAS measurements from airborne platforms. *Atmospheric Measurement Techniques*, 5(5):1085–1098.
- Boersma, K. F., Eskes, H. J., and Brinkma, E. J. (2004). Error analysis for tropospheric NO₂ retrieval from space. *Journal of Geophysical Research D: Atmospheres*, 109(4).
- Boersma, K. F., Eskes, H. J., Richter, A., De Smedt, I., Lorente, A., Beirle, S., van Geffen,

- J. H. G. M., Zara, M., Peters, E., Van Roozendael, M., Wagner, T., Maasakkers, J. D., van der A, R. J., Nightingale, J., De Rudder, A., Irie, H., Pinardi, G., Lambert, J.-C., and Compernolle, S. C. (2018). Improving algorithms and uncertainty estimates for satellite NO₂ retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project. *Atmospheric Measurement Techniques*, 11(12):6651–6678. Publisher: Copernicus GmbH.
- Boersma, K. F., Jacob, D. J., Bucsela, E. J., Perring, A. E., Dirksen, R., van der A, R. J., Yantosca, R. M., Park, R. J., Wenig, M. O., Bertram, T. H., and Cohen, R. C. (2008). Validation of OMI tropospheric NO₂ observations during INTEX-B and application to constrain NO_x emissions over the eastern United States and Mexico. *Atmospheric Environment*, 42(19):4480–4497. Publisher: Pergamon.
- Boersma, K. F., Van Geffen, J. H., Eskes, H. J., van der A, R. J., De Smedt, I., Van Roozendael, M., Yu, H., Richter, A., Peters, E., Beirle, S., Wagner, T., Lorente, A., Scanlon, T., Compernolle, S. C., and Lambert, J.-C. (2017). QA4ECV Product Specification Document for the QA4ECV NO₂ ECV precursor product. Technical report, KNMI.
- Boersma, K. F., Vinken, G. C., and Eskes, H. J. (2016). Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV-Vis tropospheric column retrievals. *Geoscientific Model Development*, 9(2):875–898.
- Boersma, K. F., Vinken, G. C., and Tournadre, J. (2015). Ships going slow in reducing their NO_x emissions: Changes in 2005-2012 ship exhaust inferred from satellite measurements over Europe. *Environmental Research Letters*, 10(7):74007. Publisher: IOP Publishing.
- Brown, J. H., Cook, K. M., Ney, F. G., and Hatch, T. (1950). Influence of Particle Size upon the Retention of Particulate Matter in the Human Lung. *American Journal of Public Health and the Nations Health*, 40(4):450–480.
- Buitendijk, M. (2024). All ships at RWG container terminal to use shore power | SWZ. <https://swzmaritime.nl/news/2024/03/01/ships-at-rwg-container-terminal-to-use-shore-power-from-2026/>. Accessed: 2024-06-05.
- Buriez, J.-C. (2005). An improved derivation of the top-of-atmosphere albedo from POLDER/ADEOS-2: Narrowband albedos. *Journal of Geophysical Research*, 110(D5):D05202. Publisher: John Wiley & Sons, Ltd.
- Chen, G., Huey, L. G., Trainer, M., Nicks, D., Corbett, J., Ryerson, T., Parrish, D., Neuman, J. A., Nowak, J., Tanner, D., Holloway, J., Brock, C., Crawford, J., Olson, J. R., Sullivan, A., Weber, R., Schaubler, S., Donnelly, S., Atlas, E., Roberts, J., Flocke, F., Hübler, G., and Fehsenfeld, F. (2005). An investigation of the chemistry of ship emission plumes during ITCT 2002. *Journal of Geophysical Research D: Atmospheres*,

- 110(10):1–15.
- Chen, T. and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA. ACM. event-place: San Francisco, California, USA.
- Cheng, C.-W., Hua, J., and Hwang, D.-S. (2017). NO_x emission calculations for bulk carriers by using engine power probabilities as weighting factors. *Journal of the Air & Waste Management Association*, 67(10):1146–1157. Publisher: Taylor & Francis .eprint: <https://doi.org/10.1080/10962247.2017.1356763>.
- Cheng, Y., Wang, S., Zhu, J., Guo, Y., Zhang, R., Liu, Y., Zhang, Y., Yu, Q., Ma, W., and Zhou, B. (2019). Surveillance of SO₂ and NO₂ from ship emissions by MAX-DOAS measurements and the implications regarding fuel sulfur content compliance. *Atmospheric Chemistry and Physics*, 19(21):13611–13626. Publisher: Copernicus GmbH.
- Chollet, F. (2015). keras. <https://github.com/fchollet/keras>. Accessed: 2022-07-05.
- Comer, B., McCabe, S., Carr, W., Elling, M., Sturup, E., Knudsen, B., and Beecken, J. (2023). Real-World NO_x Emissions from Ships and Implications for Future Regulations. Technical report, INTERNATIONAL COUNCIL ON CLEAN TRANSPORTATION.
- Compernelle, S., Argyrouli, A., Lutz, R., Sneep, M., Lambert, J. C., Mari Fjæraa, A., Hubert, D., Keppens, A., Loyola, D., O'Connor, E., Romahn, F., Stammes, P., Verhoelst, T., and Wang, P. (2021). Validation of the Sentinel-5 Precursor TROPOMI cloud data with Cloudnet, Aura OMI O₂-O₂, MODIS, and Suomi-NPP VIIRS. *Atmospheric Measurement Techniques*, 14(3):2451–2476. Publisher: Copernicus GmbH.
- Corbett, J. J., Winebrake, J. J., Green, E. H., Kasibhatla, P., Eyring, V., and Lauer, A. (2007). Mortality from ship emissions: A global assessment. *Environmental Science and Technology*, 41(24):8512–8518. Publisher: American Chemical Society.
- Cox, C. and Munk, W. (1954). Statistics of the sea surface derived from sun glitter. *Journal of Marine Research*, 13(2).
- Cox, C. and Munk, W. (1956). Slopes of the sea surface deduced from photographs of sun glitter. *Bulletin of the scripps institution of oceanography of the University of California*, 6(9):401–488. arXiv: physics/0608246v3 ISBN: 9780980200454.
- Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., Monni, S., Doering, U., Olivier, J. G. J., Pagliari, V., and others (2018). Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4. 3.2. *Earth Syst. Sci. Data*, 10(4):1987–2013.
- Curier, R. L., Kranenburg, R., Segers, A. J., Timmermans, R. M., and Schaap, M. (2014). Synergistic use of OMI NO₂ tropospheric columns and LOTOS-EUROS to evaluate the NO_x emission trends across Europe. *Remote Sensing of Environment*, 149(2):58–69.

Publisher: Elsevier Inc.

- de Haan, J., Bosma, P., and Hovenier, J. (1987). The adding method for multiple scattering calculations of polarized light. *Astronomy and astrophysics (Berlin. Print)*, 183(2):371–391.
- De Ruyter de Wildt, M., Eskes, H., and Boersma, K. F. (2012). The global economic cycle and satellite-derived NO₂ trends over shipping lanes. *Geophysical Research Letters*, 39(1):2–7.
- De Smedt, I., Pinardi, G., Vigouroux, C., Compernelle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Pithers, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendaal, M. (2021). Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements. *Atmospheric Chemistry and Physics*, 21(16):12561–12593. Publisher: Copernicus GmbH.
- de Vries, W. (2021). Impacts of nitrogen emissions on ecosystems and human health: A mini review. *Current Opinion in Environmental Science & Health*, 21:100249.
- Diamond, M. S. (2023). Detection of large-scale cloud microphysical changes within a major shipping corridor after implementation of the International Maritime Organization 2020 fuel sulfur regulations. *Atmospheric Chemistry and Physics*, 23(14):8259–8269. Publisher: Copernicus GmbH.
- Ding, J., van der A, R. J., Eskes, H. J., Mijling, B., Stavrou, T., van Geffen, J. H., and Veefkind, J. P. (2020). NO_x Emissions Reduction and Rebound in China Due to the COVID-19 Crisis. *Geophysical Research Letters*, 47(19):1–9.
- Dirksen, R. J., Boersma, K. F., Eskes, H. J., Ionov, D. V., Bucseles, E. J., Levelt, P. F., and Kelder, H. M. (2011). Evaluation of stratospheric NO₂ retrieved from the Ozone Monitoring Instrument: Intercomparison, diurnal cycle, and trending. *Journal of Geophysical Research Atmospheres*, 116(8):1–22.
- Doumbia, T., Granier, C., Elguindi, N., Bouarar, I., Darras, S., Brasseur, G., Gaubert, B., Liu, Y., Shi, X., Stavrou, T., Tilmes, S., Lacey, F., Deroubaix, A., and Wang, T. (2021). Changes in global air pollutant emissions during the COVID-19 pandemic: a dataset for atmospheric modeling. *Earth Syst. Sci. Data*, 13:4191–4206.
- Douros, J., Eskes, H., van Geffen, J., Boersma, K. F., Compernelle, S., Pinardi, G., Blechschmidt, A.-M., Peuch, V.-H., Colette, A., and Veefkind, P. (2023). Comparing Sentinel-5P TROPOMI NO₂ column observations with the CAMS regional air quality ensemble. *Geoscientific Model Development*, 16(2):509–534. Publisher: Copernicus GmbH.
- Dragović, B., Tzannatos, E., Tselentis, V., Meštrović, R., and Škurić, M. (2018). Ship

- emissions and their externalities in cruise ports. *Transportation Research Part D: Transport and Environment*, 61:289–300.
- ECMWF (2015). IFS Documentation CY41R1 - Part IV: Physical Processes. Technical report, ECMWF.
- Ecotech, A. G. (2023). Serinus 40 Oxides of Nitrogen Analyser – Acoem UK. <https://www.acoem.co.uk/product/ecotech/serinus-40-oxides-of-nitrogen-analyser/>. Accessed: 2023-02-03.
- Elguindi, N., Granier, C., Stavrou, T., Darras, S., Bauwens, M., Cao, H., Chen, C., Denier van der Gon, H. a. C., Dubovik, O., Fu, T. M., Henze, D. K., Jiang, Z., Keita, S., Kuenen, J. J. P., Kurokawa, J., Liou, C., Miyazaki, K., Müller, J.-F., Qu, Z., Solomon, F., and Zheng, B. (2020). Intercomparison of Magnitudes and Trends in Anthropogenic Surface Emissions From Bottom-Up Inventories, Top-Down Estimates, and Emission Scenarios. *Earth's Future*, 8(8):e2020EF001520. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020EF001520>.
- Elshorbany, Y., Zhu, Y., Wang, Y., Zhou, X., Sanderfield, S., Ye, C., Hayden, M., and Peters, A. J. (2022). Seasonal dependency of the atmospheric oxidizing capacity of the marine boundary layer of Bermuda. *Atmospheric Environment*, 289. Publisher: Elsevier Ltd.
- EMSA (2023). Alternative Sources of Power. <https://www.emsa.europa.eu/sustainable-shipping/alternative-fuels.html>. Accessed: 2024-06-05.
- Escudero, M., Segers, A., Kranenburg, R., Querol, X., Alastuey, A., Borge, R., De La Paz, D., Gangoiti, G., and Schaap, M. (2019). Analysis of summer O₃ in the Madrid air basin with the LOTOS-EUROS chemical transport model. *Atmospheric Chemistry and Physics*, 19(22):14211–14232. Publisher: Copernicus GmbH.
- Eskes, H. and van Geffen, J. (2021). Product user manual for the TM5 NO₂, SO₂ and HCHO profile auxiliary support product. Technical report, KNMI, de Bilt. Issue: 1.0.0.
- Eskes, H., Van Geffen, J., Boersma, F., Eichmann, K.-U., Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J. P., and Loyola, D. (2019). Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Nitrogen dioxide. *Royal Netherlands Meteorological Institute*, page 147.
- Eskes, H. J. and Boersma, K. F. (2003). Averaging kernels for DOAS total-column satellite retrievals. *Atmospheric Chemistry and Physics*, 3(5):1285–1291.
- European Commission (2024). Reducing emissions from the shipping sector. https://climate.ec.europa.eu/eu-action/transport/reducing-emissions-shipping-sector_en. Accessed: 2024-06-17.
- European Environment Agency (2019a). *NEC Directive reporting status 2019*. Publications Office of the European Union. Accessed: 2024-05-29.

- European Environment Agency (2019b). Nitrogen oxides (NO_x) Emissions European Environment Agency. <https://www.eea.europa.eu/data-and-maps/indicators/eea-32-nitrogen-oxides-nox-emissions-1>. Accessed: 2020-05-13.
- Eyring, V., Isaksen, I. S., Berntsen, T., Collins, W. J., Corbett, J. J., Endresen, O., Grainger, R. G., Moldanova, J., Schlager, H., and Stevenson, D. S. (2010). Transport impacts on atmosphere and climate: Shipping. *Atmospheric Environment*, 44(37):4735–4771. Publisher: Pergamon.
- Eyring, V., Köhler, H. W., Van Aardenne, J., and Lauer, A. (2005). Emissions from international shipping: 1. The last 50 years. *Journal of Geophysical Research D: Atmospheres*, 110(17):171–182. Publisher: John Wiley & Sons, Ltd.
- Finch, D. P., Palmer, P. I., and Zhang, T. (2022). Automated detection of atmospheric NO₂ plumes from satellite data: a tool to help infer anthropogenic combustion emissions. *Atmospheric Measurement Techniques*, 15(3):721–733. Publisher: Copernicus GmbH.
- Flerlage, H., Velders, G. J. M., and de Boer, J. (2021). A review of bottom-up and top-down emission estimates of hydrofluorocarbons (HFCs) in different parts of the world. *Chemosphere*, 283:131208.
- Fortems-Cheiney, A., Broquet, G., Pison, I., Saunois, M., Potier, E., Berchet, A., Dufour, G., Siour, G., Denier van der Gon, H., Dellaert, S. N., and Boersma, K. F. (2021). Analysis of the Anthropogenic and Biogenic NO_x Emissions Over 2008–2017: Assessment of the Trends in the 30 Most Populated Urban Areas in Europe. *Geophysical Research Letters*, 48(11):e2020GL092206. Publisher: John Wiley & Sons, Ltd.
- Georgoulias, A. K., Boersma, K. F., Van Vliet, J., Zhang, X., Van Der A, R., Zanis, P., and De Laat, J. (2020). Detection of NO₂ pollution plumes from individual ships with the TROPOMI/S5P satellite sensor. *Environmental Research Letters*, 15(12):124037. Publisher: IOP Publishing Ltd.
- Goldberg, D. L., Anenberg, S. C., Griffin, D., McLinden, C. A., Lu, Z., and Streets, D. G. (2020). Disentangling the Impact of the COVID-19 Lockdowns on Urban NO₂ From Natural Variability. *Geophysical Research Letters*, 47(17):e2020GL089269. Publisher: John Wiley & Sons, Ltd.
- Goldberg, D. L., Anenberg, S. C., Lu, Z., Streets, D. G., Lamsal, L. N., McDuffie, E. E., and Smith, S. J. (2021). Urban NO_x emissions around the world declined faster than anticipated between 2005 and 2019. *Environmental Research Letters*, 16(11):115004. Publisher: IOP Publishing.
- Golder, D. (1972). Relations among stability parameters in the surface layer. *Boundary-Layer Meteorology*, 3(1):47–58. Publisher: Kluwer Academic Publishers.
- Granier, C., Darras, S., Denier Van Der Gon, H., Jana, D., Elguindi, N., Bo, G., Michael, G., Marc, G., Jalkanen, J.-P., and Kuenen, J. (2019). The Copernicus Atmosphere

- Monitoring Service global and regional emissions (April 2019 version).
- Griffin, D., Zhao, X., McLinden, C. A., Boersma, F., Bourassa, A., Dammers, E., Degenstein, D., Eskes, H., Fehr, L., Fioletov, V., Hayden, K., Kharol, S. K., Li, S. M., Makar, P., Martin, R. V., Mihele, C., Mittermeier, R. L., Krotkov, N., Sneep, M., Lamsal, L. N., Linden, M. t., Geffen, J. v., Veeffkind, P., and Wolde, M. (2019). High-Resolution Mapping of Nitrogen Dioxide With TROPOMI: First Results and Validation Over the Canadian Oil Sands. *Geophysical Research Letters*, 46(2):1049–1060.
- Gubbi, S., Cole, R., Emerson, B., Noble, D., Steele, R., Sun, W., and Lieuwen, T. (2023). Air Quality Implications of Using Ammonia as a Renewable Fuel: How Low Can NOx Emissions Go? *ACS Energy Letters*, 8(10):4421–4426. Publisher: American Chemical Society.
- Guevara, M., Jorba, O., Soret, A., Petetin, H., Bowdalo, D., Serradell, K., Tena, C., Denier van der Gon, H., Kuenen, J., Peuch, V.-H., and Pérez García-Pando, C. (2021). Time-resolved emission reductions for atmospheric chemistry modelling in Europe during the COVID-19 lockdowns. *Atmospheric Chemistry and Physics*, 21(2):773–797. Publisher: Copernicus GmbH.
- Hanna, S. R., Schulman, L. L., Paine, R. J., Pleim, J. E., and Baer, M. (1985). Development and Evaluation of the Offshore and Coastal Dispersion Model. *Journal of the Air Pollution Control Association*, 35(10):1039–1047. Publisher: Taylor & Francis Group ISBN: 35:10391047.
- Hassler, B., McDonald, B. C., Frost, G. J., Borbon, A., Carslaw, D. C., Civerolo, K., Granier, C., Monks, P. S., Monks, S., Parrish, D. D., Pollack, I. B., Rosenlof, K. H., Ryerson, T. B., von Schneidmesser, E., and Trainer, M. (2016). Analysis of long-term observations of NOx and CO in megacities and application to constraining emissions inventories. *Geophysical Research Letters*, 43(18):9920–9930. Publisher: Blackwell Publishing Ltd.
- Heidinger, A. and Li, Y. (2017). Algorithm Theoretical Basis Document AWG Cloud Height Algorithm (ACHA). Technical report, NOAA NESDIS Center for Satellite Applications and Research.
- Holtslag, A. A. M. and Boville, B. A. (1993). Local Versus Nonlocal Boundary-Layer Diffusion in a Global Climate Model. *Journal of Climate*, 6(10):1825–1842.
- Huang, S., Li, H., Wang, M., Qian, Y., Steenland, K., Caudle, W. M., Liu, Y., Sarnat, J., Papatheodorou, S., and Shi, L. (2021). Long-term exposure to nitrogen dioxide and mortality: A systematic review and meta-analysis. *Science of The Total Environment*, 776:145968.
- Huijnen, V., Williams, J. E., Van Weele, M., Van Noije, T. P. C., Krol, M. C., Dentener, F., Segers, A., Houweling, S., Peters, W., De Laat, A. T. J., Boersma, K. F., Bergamaschi, P., Van Velthoven, P. F. J., Le Sager, P., Eskes, H. J., Alkemade, F., and

- Scheele, M. P. (2010). The global chemistry transport model TM5 Geoscientific Model Development Discussions The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0 The global chemistry transport model TM5. *Geosci. Model Dev. Discuss*, 3:1009–1087.
- Hulskotte, J. H. J. and Denier van der Gon, H. A. C. (2010). Fuel consumption and associated emissions from seagoing ships at berth derived from an on-board survey. *Atmospheric Environment*, 44(9):1229–1236.
- Ialongo, I., Virta, H., Eskes, H., Hovila, J., and Douros, J. (2020). Comparison of TROPOMI/Sentinel-5 Precursor NO₂ observations with ground-based measurements in Helsinki. *Atmospheric Measurement Techniques*, 13(1):205–218. Publisher: Copernicus GmbH.
- Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H., and Courrèges-Lacoste, G. B. (2012). Requirements for the GMES Atmosphere Service and ESA’s implementation concept: Sentinels-4/-5 and -5p. *Remote Sensing of Environment*, 120:58–69.
- Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A., Dominguez, J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger M., Remy, S., Schulz, M., and Suttie, M. (2021). CAMS global reanalysis (EAC4) monthly averaged fields.
- International Maritime Organization (2008). NO_x Technical Code (2008). [https://www.imo.org/OurWork/Environment/PollutionPrevention/AirPollution/Pages/Nitrogen-oxides-\(NOx\)---Regulation-13.aspx](https://www.imo.org/OurWork/Environment/PollutionPrevention/AirPollution/Pages/Nitrogen-oxides-(NOx)---Regulation-13.aspx). Accessed: 2024-06-28.
- International Maritime Organization (2013). Nitrogen oxides (NO_x) - Regulation 13. [http://www.imo.org/en/OurWork/Environment/PollutionPrevention/AirPollution/Pages/Nitrogen-oxides-\(NOx\)---Regulation-13.aspx](http://www.imo.org/en/OurWork/Environment/PollutionPrevention/AirPollution/Pages/Nitrogen-oxides-(NOx)---Regulation-13.aspx). Accessed: 2020-04-15.
- International Maritime Organization (2014). AIS transponders, <https://www.imo.org/en/OurWork/Safety/Pages/AIS.aspx> (last access: 24 February 2021). <https://www.imo.org/en/OurWork/Safety/Pages/AIS.aspx>. Accessed: 2021-02-24.
- International Maritime Organization (2015). Third IMO GHG Study 2014 Executive Summary and Final Report. Technical report. Accessed: 2024-01-04.
- International Maritime Organization (2020). 4th IMO Greenhouse Gas study. <https://www.imo.org/en/OurWork/Environment/Pages/Fourth-IMO-Greenhouse-Gas-Study-2020.aspx>. Accessed: 2022-10-10.
- International Maritime Organization (2023). Revised GHG reduction strategy for global shipping. <https://www.imo.org/en/MediaCentre/PressBriefings/pages/Revised-GHG-reduction-strategy-for-global-shipping-adopted-.aspx>. Accessed: 2024-02-28.

- Jacob, D. J. (1999). *Introduction to Atmospheric Chemistry*, by Daniel Jacob (Harvard University). Princeton University Press.
- Jalkanen, J.-P., Johansson, L., and Kukkonen, J. (2016). A comprehensive inventory of ship traffic exhaust emissions in the European sea areas in 2011. *Atmospheric Chemistry and Physics*, 16(1):71–84.
- Jiang, Z., Zhu, R., Miyazaki, K., McDonald, B. C., Klimont, Z., Zheng, B., Boersma, K. F., Zhang, Q., Worden, H., Worden, J. R., Henze, D. K., Jones, D. B., Denier van der Gon, H. A., and Eskes, H. (2022). Decadal Variabilities in Tropospheric Nitrogen Oxides Over United States, Europe, and China. *Journal of Geophysical Research: Atmospheres*, 127(3):e2021JD035872. Publisher: John Wiley & Sons, Ltd ISBN: 10.1029/2021.
- Johansson, L., Jalkanen, J. P., and Kukkonen, J. (2017). Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution. *Atmospheric Environment*, 167:403–415. Publisher: Elsevier Ltd.
- Kang, Y., Choi, H., Im, J., Park, S., Shin, M., Song, C.-K., and Kim, S. (2021). Estimation of surface-level NO₂ and O₃ concentrations using TROPOMI data and machine learning over East Asia. *Environmental Pollution*, 288:117711.
- Kenagy, H. S., Sparks, T. L., Ebben, C. J., Wooldrige, P. J., Lopez-Hilfiker, F. D., Lee, B. H., Thornton, J. A., McDuffie, E. E., Fibiger, D. L., Brown, S. S., Montzka, D. D., Weinheimer, A. J., Schroder, J. C., Campuzano-Jost, P., Day, D. A., Jimenez, J. L., Dibb, J. E., Campos, T., Shah, V., Jaeglé, L., and Cohen, R. C. (2018). NO_x Lifetime and NO_y Partitioning During WINTER. *Journal of Geophysical Research: Atmospheres*, 123(17):9813–9827. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018JD028736>.
- Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S. K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y. J., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Miller, C. C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., Abad, G. G., Haffner, D. P., Ko, D. H., Lee, S. H., Woo, J.-H., Chong, H., Park, S. S., Nicks, D., Choi, W. J., Moon, K.-J., Cho, A., Yoon, J., Kim, S.-k., Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veefkind, P., Levelt, P. F., Edwards, D. P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H.-A., Yang, J., Park, J., Han, K. M., Kim, B.-R., Shin, H.-W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J.-H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J. H., Carmichael, G. R., Newchurch, M. J., Lefer, B. L., Herman, J. R., Swap, R. J., Lau, A. K. H., Kurosu, T. P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C. T., and Choi, Y. (2020). New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS). *Bulletin of the American Meteorological Society*, 101(1):E1–E22. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society.
- Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F. (2008). Earth surface

- reflectance climatology from 3 years of OMI data. *Journal of Geophysical Research Atmospheres*, 113(18):1–22.
- KNMI (2024). North Sea Observatory for Atmospheric and Environmental Research. <https://www.onderzoeksfaciliteiten.nl/node/3827>. Accessed: 2024-06-19.
- Koelemeijer, R. B., Stammes, P., Hovenier, J. W., and De Haan, J. F. (2001). A fast method for retrieval of cloud parameters using oxygen a band measurements from the Global Ozone Monitoring Experiment. *Journal of Geophysical Research Atmospheres*, 106(D4):3475–3490.
- Krause, K., Wittrock, F., Richter, A., Schmitt, S., Pöhler, D., Weigelt, A., and Burrows, J. P. (2021). Estimation of ship emission rates at a major shipping lane by long-path DOAS measurements. *Atmospheric Measurement Techniques*, 14(8):5791–5807. Publisher: Copernicus GmbH.
- Krecl, P., Harrison, R. M., Johansson, C., Targino, A. C., Beddows, D. C., Ellermann, T., Lara, C., and Ketzel, M. (2021). Long-term trends in nitrogen oxides concentrations and on-road vehicle emission factors in Copenhagen, London and Stockholm. *Environmental Pollution*, 290:118105.
- Krol, M., van Stratum, B., Anglou, I., and Boersma, K. F. (2024). Estimating NO_x emissions of stack plumes using a high-resolution atmospheric chemistry model and satellite-derived NO₂ columns. *EGUsphere*, pages 1–32. Publisher: Copernicus GmbH.
- Kuhn, L., Beirle, S., Osipov, S., Pozzer, A., and Wagner, T. (2024). NitroNet – A deep-learning NO₂ profile retrieval prototype for the TROPOMI satellite instrument. *EGUsphere*, pages 1–43. Publisher: Copernicus GmbH.
- Kurchaba, S., Sokolovsky, A., van Vliet, J., Verbeek, F. J., and Veenman, C. J. (2024). Sensitivity analysis for the detection of NO₂ plumes from seagoing ships using TROPOMI data. *Remote Sensing of Environment*, 304:114041.
- Kurchaba, S., Van Vliet, J., Meulman, J. J., Verbeek, F. J., and Veenman, C. J. (2021). Improving evaluation of NO₂ emission from ships using spatial association on TROPOMI satellite data. *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, pages 454–457. Publisher: Association for Computing Machinery ISBN: 9781450386647.
- Kurchaba, S., Van Vliet, J., Verbeek, F. J., Meulman, J. J., and Veenman, C. J. (2022). Supervised Segmentation of NO₂ Plumes from Individual Ships Using TROPOMI Satellite Data. *Remote Sensing 2022, Vol. 14, Page 5809*, 14(22):5809. arXiv: 2203.06993 Publisher: Multidisciplinary Digital Publishing Institute.
- Kurchaba, S., van Vliet, J., Verbeek, F. J., and Veenman, C. J. (2023). Anomalous NO₂ emitting ship detection with TROPOMI satellite data and machine learning. *Remote Sensing of Environment*, 297:113761. Publisher: Elsevier.

- Köhler, M., Ahlgrim, M., and Beljaars, A. (2011). Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model. *Quarterly Journal of the Royal Meteorological Society*, 137(654):43–57. Publisher: John Wiley & Sons, Ltd.
- Lack, D. A., Corbett, J. J., Onasch, T., Lerner, B., Massoli, P., Quinn, P. K., Bates, T. S., Covert, D. S., Coffman, D., Sierau, B., Herndon, S., Allan, J., Baynard, T., Lovejoy, E., Ravishankara, A. R., and Williams, E. (2009). Particulate emissions from commercial shipping: Chemical, physical, and optical properties. *Journal of Geophysical Research Atmospheres*, 114(4). Publisher: Blackwell Publishing Ltd.
- Ladi, T., Jabalameli, S., and Sharifi, A. (2022). Applications of machine learning and deep learning methods for climate change mitigation and adaptation. *Environment and Planning B: Urban Analytics and City Science*, 49(4):1314–1330. Publisher: SAGE Publications Ltd STM.
- Lama, S., Houweling, S., Boersma, K. F., Aben, I., Denier van der Gon, H. A. C., and Krol, M. C. (2022). Estimation of OH in urban plumes using TROPOMI-inferred NO₂ CO. *Atmospheric Chemistry and Physics*, 22(24):16053–16071.
- Lampel, J., Frieß, U., and Platt, U. (2015). The impact of vibrational Raman scattering of air on DOAS measurements of atmospheric trace gases. *Atmospheric Measurement Techniques*, 8(9):3767–3787. Publisher: Copernicus GmbH.
- Landgraf, J., Rusli, S., Cooney, R., Veefkind, P., Vemmix, T., de Groot, Z., Bell, A., Day, J., Leemhuis, A., and Sierk, B. (2020). The TANGO mission: A satellite tandem to measure major sources of anthropogenic greenhouse gas emissions. page 19643. Conference Name: EGU General Assembly Conference Abstracts ADS Bibcode: 2020EGUGA..2219643L.
- Lansø, A. S., Winther, M., Jensen, S. S., and Løfstrøm, P. (2023). Impact on air quality from increasing cruise ship activity in Copenhagen port. *Environmental Research Communications*, 5(2):021003. Publisher: IOP Publishing.
- Latsch, M., Richter, A., Eskes, H., Sneep, M., Wang, P., Veefkind, P., Lutz, R., Loyola, D., Argyrouli, A., Valks, P., Wagner, T., Sihler, H., van Roozendaal, M., Theys, N., Yu, H., Siddans, R., and Burrows, J. P. (2022). Intercomparison of Sentinel-5P TROPOMI cloud products for tropospheric trace gas retrievals. *Atmospheric Measurement Techniques*, 15(21):6257–6283. Publisher: Copernicus GmbH.
- Levelt, P., van den Oord, G., Dobber, M., Malkki, A., Visser, H., Vries, J. d., Stammes, P., Lundell, J., and Saari, H. (2006). The ozone monitoring instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 44(5):1093–1101. Conference Name: IEEE Transactions on Geoscience and Remote Sensing.
- Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Zeevers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., Van Der, R. A., McLinden, C., Fioletov, V.,

- Carn, S., De Laat, J., Deland, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., Abad, G. G., Arola, A., Boersma, F., Miller, C. C., Chance, K., De Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H., and Wargan, K. (2018). The Ozone Monitoring Instrument: Overview of 14 years in space. *Atmospheric Chemistry and Physics*, 18(8):5699–5745. Publisher: Copernicus GmbH.
- Li, Q., Badia, A., Fernandez, R. P., Mahajan, A. S., López-Noreña, A. I., Zhang, Y., Wang, S., Puliafito, E., Cuevas, C. A., and Saiz-Lopez, A. (2021). Chemical Interactions Between Ship-Originated Air Pollutants and Ocean-Emitted Halogens. *Journal of Geophysical Research: Atmospheres*, 126(4):e2020JD034175.
- Loots, E., Rozemeijer, N., Kleipool, Q., and Ludewig, A. (2017). S5P L1B ATDB. *KNMI*.
- Lorente, A., Boersma, K. F., Eskes, H. J., Veefkind, J. P., van Geffen, J. H., de Zeeuw, M. B., Denier van der Gon, H. A., Beirle, S., and Krol, M. C. (2019). Quantification of nitrogen oxides emissions from build-up of pollution over Paris with TROPOMI. *Scientific Reports*, 9(1):1–10.
- Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., aan de Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D., Pollard, D. F., Shiomi, K., Deutscher, N. M., Velasco, V. A., Roehl, C. M., Wennberg, P. O., Warneke, T., and Landgraf, J. (2021). Methane retrieved from TROPOMI: improvement of the data product and validation of the first 2 years of measurements. *Atmospheric Measurement Techniques*, 14(1):665–684. Publisher: Copernicus GmbH.
- Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., Van Roozendaal, M., Wang, Y., Wagner, T., Beirle, S., Lin, J. T., Krotkov, N., Stammes, P., Wang, P., Eskes, H. J., and Krol, M. (2017). Structural uncertainty in air mass factor calculation for NO₂ and HCHO satellite retrievals. *Atmospheric Measurement Techniques*, 10(3):759–782. Publisher: Copernicus GmbH.
- Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., Van Der Plas, E., Rozemeijer, N., Vonk, F., and Veefkind, P. (2020). In-flight calibration results of the TROPOMI payload on board the Sentinel-5 Precursor satellite. *Atmospheric Measurement Techniques*, 13(7):3561–3580.
- Luo, K., Li, R., Li, W., Wang, Z., Ma, X., Zhang, R., Fang, X., Wu, Z., Cao, Y., and Xu, Q. (2016). Acute Effects of Nitrogen Dioxide on Cardiovascular Mortality in Beijing: An Exploration of Spatial Heterogeneity and the District-specific Predictors. *Scientific Reports 2016 6:1*, 6(1):1–13. Publisher: Nature Publishing Group.
- Maasakkers, J. D., Varon, D. J., Elfarsdóttir, A., McKeever, J., Jervis, D., Mahapatra, G., Pandey, S., Lorente, A., Borsdorff, T., Foorthuis, L. R., Schuit, B. J., Tol, P.,

- van Kempen, T. A., van Hees, R., and Aben, I. (2022). Using satellites to uncover large methane emissions from landfills. *Science Advances*, 8(32):eabn9683. Publisher: American Association for the Advancement of Science.
- Manders, A. M., Builtjes, P. J., Curier, L., Gon, H. A. V., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J. J., Segers, A. J., Timmermans, R. M., Visschedijk, A. J., Kruit, R. J., Pul, W. A. J., Sauter, F. J., Van Der Swaluw, E., Swart, D. P., Douros, J., Eskes, H., Van Meijgaard, E., Van Ulft, B., Van Velthoven, P., Banzhaf, S., Mues, A. C., Stern, R., Fu, G., Lu, S., Heemink, A., Van Velzen, N., and Schaap, M. (2017). Curriculum vitae of the LOTOS-EUROS (v2.0) chemistry transport model. *Geoscientific Model Development*, 10(11):4145–4173. Publisher: Copernicus GmbH.
- Marais, E. A., Jacob, D. J., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., Krechmer, J., Zhu, L., Kim, P. S., Miller, C. C., Fisher, J. A., Travis, K., Yu, K., Hanisco, T. F., Wolfe, G. M., Arkinson, H. L., Pye, H. O., Froyd, K. D., Liao, J., and McNeill, V. F. (2015). Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: Application to the Southeast United States and co-benefit of SO₂ emission controls. *Atmospheric Chemistry and Physics Discussions*, 15(21):32005–32047.
- March, D., Metcalfe, K., Tintoré, J., and Godley, B. (2020). Tracking the global reduction of marine traffic during the COVID-19 pandemic. (*Preprint*) DOI: 10.21203/rs.3.rs-47243/v1+.
- Marmer, E., Dentener, F., V Aardenne, J., Cavalli, F., Vignati, E., Velchev, K., Hjorth, J., Boersma, F., Vinken, G., Mihalopoulos, N., and Raes, F. (2009). What can we learn about ship emission inventories from measurements of air pollutants over the Mediterranean Sea? *Atmospheric Chemistry and Physics*, 9(18):6815–6831.
- Marshallsay, H. (2023). How Old is an Elderly Container Ship? <https://www.pfe.express/2023/12/how-old-is-an-elderly-container-ship/>. Accessed: 2024-06-05.
- Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., Fiore, A. M., Li, Q., Yantosca, R. M., and Koелеmeijer, R. B. (2002). An improved retrieval of tropospheric nitrogen dioxide from GOME. *Journal of Geophysical Research: Atmospheres*, 107(D20):ACH 9–1. Publisher: John Wiley & Sons, Ltd.
- Marécal, V., Peuch, V. H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., Van Der Gon, H. A., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouillé, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu,

- M., Poupkou, A., Queguiner, S., Robertson, L., Rouil, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., Valdebenito, Van Velthoven, P., Van Versendaal, R., Vira, J., and Ung, A. (2015). A regional air quality forecasting system over Europe: The MACC-II daily ensemble production. *Geoscientific Model Development*, 8(9):2777–2813. Publisher: Copernicus GmbH.
- Matsunaga, T., Tanimoto, H., Sugita, T., Yashiro, H., Morino, I., Yoshida, Y., Saito, M., Noda, H., Ohyama, H., Inomata, S., Ikeda, K., Someya, Y., Fujinawa, T., Yamashita, Y., Saeki, T., Müller, A., Frey, M. M., Saigusa, N., Kanaya, Y., Sekiya, T., Patra, P. K., Takigawa, M., Kasai, Y., and Sato, T. (2022). Greenhouse Gas and Air Pollutant Observation by TANSO-3 onboard GOSAT-GW: One year before Launch. 2022:A12I-04. Conference Name: AGU Fall Meeting Abstracts ADS Bibcode: 2022AGUFM.A12I..04M.
- McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V. (2020). A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS). *Earth System Science Data*, 12(4):3413–3442. Publisher: Copernicus GmbH.
- McLaren, R., Wojtal, P., Halla, J. D., Mihele, C., and Brook, J. R. (2012). A survey of NO₂:SO₂ emission ratios measured in marine vessel plumes in the Strait of Georgia. *Atmospheric Environment*, 46(2):655–658. Publisher: Elsevier Ltd.
- Meier, S., Koene, E., Krol, M., Brunner, D., Damm, A., and Kuhlmann, G. (2024). A light-weight NO₂ to NO_x conversion model for quantifying NO_x emissions of point sources from NO₂ satellite observations. *EGUsphere*, pages 1–28. Publisher: Copernicus GmbH.
- Meijer, E. (2001). *Modelling the impact of subsonic aviation on the composition of the atmosphere*. PhD thesis, Eindhoven University of Technology.
- Mellqvist, J. and Conde, V. (2021). Best practice report on compliance monitoring of ships with respect to current and future IMO regulation. Technical report, Chalmers University of Technology, Göteborg.
- Mellqvist, J., Conde, V., Beecken, J., and Ekholm, J. (2017). Certification of an aircraft and airborne surveillance of fuel sulfur content in ships at the SECA border Certification of an aircraft and airborne surveillance of fuel sulfur content in ships at the SECA border Compliance monitoring pilot for Marpol Annex VI CompMon. Technical report, Chalmers University of Technology, Göteborg.
- METEO FRANCE, Institut national de l'environnement industriel et des risques (Ineris), Aarhus University, Norwegian Meteorological Institute (MET Norway), Jülich Institut für Energie- und Klimaforschung (IEK), Institute of Environmental Protection – National Research Institute (IEP-NRI), Koninklijk Nederlands Meteorologisch Instituut

- (KNMI), Nederlandse Organisatie voor toegepast-natuurwetenschappelijk onderzoek (TNO), Swedish Meteorological and Hydrological Institute (SMHI), and Finnish Meteorological Institute (FMI) (2020). CAMS European air quality forecasts.
- METEO FRANCE, MET Norway, IEK, IEP-NRI, KNMI, TNO, FMI, ENEA, and BSC (2022). CAMS Regional: European air quality analysis and forecast. Publication Title: METEO FRANCE, Institut national de l'environnement industriel et des risques (Ineris), Aarhus University, Norwegian Meteorological Institute (MET Norway), Jülich Institut für Energie- und Klimaforschung (IEK), Institute of Environmental Protection – Nati. . . .
- Mettepenningen, G., Fayt, C., Tack, F., Doorne, C. V., Bogaert, P., Jacobs, L., Berkenbosch, S., Aubry, A., Desmet, F., Robert, C., Mazière, M. D., and Roozendaal, M. V. (2024). UV-Vis remote sensing of atmospheric pollutants from a wind turbine platform in the North Sea: the SEMPAS project. Conference Name: EGU24.
- Millefiori, L. M., Braca, P., Zissis, D., Spiliopoulos, G., Marano, S., Willett, P. K., and Carniel, S. (2021). COVID-19 impact on global maritime mobility. *Scientific Reports* 2021 11:1, 11(1):1–16. arXiv: 2009.06960 Publisher: Nature Publishing Group ISBN: 0123456789.
- Müller-Casseres, E., Edelenbosch, O. Y., Szklo, A., Schaeffer, R., and van Vuuren, D. P. (2021). Global futures of trade impacting the challenge to decarbonize the international shipping sector. *Energy*, 237:121547.
- Namdar-Khojasteh, D., Yeghaneh, B., Maher, A., Namdar-Khojasteh, F., and Tu, J. (2022). Assessment of the relationship between exposure to air pollutants and COVID-19 pandemic in Tehran city, Iran. *Atmospheric Pollution Research*, 13(7):101474. Publisher: Elsevier.
- Neumann, D., Karl, M., Radtke, H., Matthias, V., Friedland, R., and Neumann, T. (2020). Quantifying the contribution of shipping NO_x emissions to the marine nitrogen inventory – a case study for the western Baltic Sea. *Ocean Science*, 16(1):115–134.
- OECD (2024). Ocean shipping and shipbuilding. <https://www.oecd.org/ocean/topics/ocean-shipping/>. Accessed: 2024-01-04.
- Oldeman, A. (2018). Effect of including an intensity offset in the DOAS NO₂ retrieval of TROPOMI. Technical report, Eindhoven University of Technology/KNMI,.
- Pirjola, L., Pajunoja, A., Walden, J., Jalkanen, J. P., Rönkkö, T., Koussa, A., and Koskentalo, T. (2014). Mobile measurements of ship emissions in two harbour areas in Finland. *Atmospheric Measurement Techniques*, 7(1):149–161.
- Platnick, S., Meyer, K. G., Heidinger, A. K., and Holz, R. (2017). VIIRS Atmosphere L2 Cloud Properties Product.
- Platt, U. and Stutz, J. (2008). *Differential Optical Absorption Spectroscopy*. Springer

- Berlin Heidelberg, Berlin, Heidelberg. Series Title: Physics of Earth and Space Environments.
- Pseftogkas, A., Koukouli, M. E., Segers, A., Manders, A., Geffen, J. v., Balis, D., Meleti, C., Stavrakou, T., and Eskes, H. (2022). Comparison of S5P/TROPOMI Inferred NO₂ Surface Concentrations with In Situ Measurements over Central Europe. *Remote Sensing 2022, Vol. 14, Page 4886*, 14(19):4886. Publisher: Multidisciplinary Digital Publishing Institute.
- Ražnjević, A., van Heerwaarden, C., and Krol, M. (2022). Evaluation of two common source estimation measurement strategies using large-eddy simulation of plume dispersion under neutral atmospheric conditions. *Atmospheric Measurement Techniques*, 15(11):3611–3628. Publisher: Copernicus GmbH.
- Richter, A., Begoin, M., Hilboll, A., and Burrows, J. P. (2011). An improved NO₂ retrieval for the GOME-2 satellite instrument. *Atmospheric Measurement Techniques*, 4(6):1147–1159. Publisher: Copernicus GmbH.
- Richter, A., Eyring, V., Burrows, J. P., Bovensmann, H., Lauer, A., Sierk, B., and Crutzen, P. J. (2004). Satellite measurements of NO₂ from international shipping emissions. *Geophysical Research Letters*, 31(23):1–4.
- Riess, T. C. V. W., Boersma, K. F., Prummel, A., van Stratum, B. J. H., de Laat, J., and van Vliet, J. (2024). Estimating NO_x Emission of Individual Ships from TROPOMI NO₂ Plumes. <http://dx.doi.org/10.2139/ssrn.4858709>. under review.
- Riess, T. C. V. W., Boersma, K. F., Van Roy, W., de Laat, J., Dammers, E., and van Vliet, J. (2023). To new heights by flying low: comparison of aircraft vertical NO₂ profiles to model simulations and implications for TROPOMI NO₂ retrievals. *Atmospheric Measurement Techniques*, 16(21):5287–5304. Publisher: Copernicus GmbH.
- Riess, T. C. V. W., Boersma, K. F., Van Vliet, J., Peters, W., Sneep, M., Eskes, H., and Van Geffen, J. (2022). Improved monitoring of shipping NO₂ with TROPOMI: Decreasing NO_x emissions in European seas during the COVID-19 pandemic. *Atmospheric Measurement Techniques*, 15(5):1415–1438. Publisher: Copernicus GmbH.
- Ripperger-Lukosiunaite, S., Ziegler, S., Donner, S., Kuhn, L., Hoffmann, T., Hoor, P., and Wagner, T. (2023). Estimating Nitrogen Oxides emissions from inland waterway vessels using MAX-DOAS measurements – Results of pioneering measurements and plans to advance the method. Conference Name: EGU23.
- Rotach, M. W., Serafin, S., Ward, H. C., Arpagaus, M., Colfescu, I., Cuxart, J., Wekker, S. F. J. D., Grubišić, V., Kalthoff, N., Karl, T., Kirshbaum, D. J., Lehner, M., Mobbs, S., Paci, A., Palazzi, E., Bailey, A., Schmidli, J., Wittmann, C., Wohlfahrt, G., and Zardi, D. (2022). A Collaborative Effort to Better Understand, Measure, and Model Atmospheric Exchange Processes over Mountains. *Bulletin of the American Meteorological Society*, 103(5):E1282 – E1295. Place: Boston MA, USA Publisher: American

Meteorological Society.

- Scheel, L. D., Dobrogorski, O. J., Mountain, J. T., Svirbely, J. L., and Stokinger, H. E. (1959). Physiologic, biochemical, immunologic and pathologic changes following ozone exposure. *Journal of Applied Physiology*, 14(1):67–80. Publisher: American Physiological Society.
- Schneider, R., Bonavita, M., Geer, A., Arcucci, R., Dueben, P., Vitolo, C., Le Saux, B., Demir, B., and Mathieu, P.-P. (2022). ESA-ECMWF Report on recent progress and research directions in machine learning for Earth System observation and prediction. *npj Climate and Atmospheric Science*, 5(1):1–5. Publisher: Nature Publishing Group.
- Schuit, B. J., Maasackers, J. D., Bijl, P., Mahapatra, G., van den Berg, A.-W., Pandey, S., Lorente, A., Borsdorff, T., Houweling, S., Varon, D. J., McKeever, J., Jervis, D., Girard, M., Irakulis-Loitxate, I., Gorroño, J., Guanter, L., Cusworth, D. H., and Aben, I. (2023). Automated detection and monitoring of methane super-emitters using satellite data. *Atmospheric Chemistry and Physics*, 23(16):9071–9098. Publisher: Copernicus GmbH.
- Seyler, A., Meier, A. C., Wittrock, F., Kattner, L., Mathieu-Üffing, B., Peters, E., Richter, A., Ruhtz, T., Schönhardt, A., Schmolke, S., and Burrows, J. P. (2019). Studies of the horizontal inhomogeneities in NO₂ concentrations above a shipping lane using ground-based multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements and validation with airborne imaging DOAS measurements. *Atmos. Meas. Tech.*, 12:5959–5977.
- Shah, V., Jacob, D. J., Dang, R., Lamsal, L. N., Strode, S. A., Steenrod, S. D., Folkert Boersma, K., Eastham, S. D., Fritz, T. M., Thompson, C., Peischl, J., Bourgeois, I., Pollack, I. B., Nault, B. A., Cohen, R. C., John, H., and Paulson, A. (2022). Nitrogen oxides in the free troposphere: Implications for tropospheric oxidants and the interpretation of satellite NO₂ measurements. *Pedro Campuzano-Jost*, 15:19.
- Skoulidou, I., Koukouli, M.-E., Manders, A., Segers, A., Karagkiozidis, D., Gratsea, M., Balis, D., Bais, A., Gerasopoulos, E., Stavrakou, T., Van Geffen, J., Eskes, H., and Richter, A. (2021). Evaluation of the LOTOS-EUROS NO₂ simulations using ground-based measurements and S5P/TROPOMI observations over Greece. *Atmos. Chem. Phys.*, 21:5269–5288.
- Sneep, M., de Haan, J. F., Stammes, P., Wang, P., Vanbauce, C., Joiner, J., Vasilkov, A. P., and Levelt, P. F. (2008). Three-way comparison between OMI and PARASOL cloud pressure products. *Journal of Geophysical Research*, 113(D15):15–23. Publisher: American Geophysical Union (AGU).
- Sofiev, M., Winebrake, J. J., Johansson, L., Carr, E. W., Prank, M., Soares, J., Vira, J., Kouznetsov, R., Jalkanen, J. P., and Corbett, J. J. (2018). Cleaner fuels for ships provide public health benefits with climate tradeoffs. *Nature Communications 2018*

- 9:1, 9(1):1–12. Publisher: Nature Publishing Group.
- Squires, F. A., Nemitz, E., Langford, B., Wild, O., Drysdale, W. S., Acton, W. J. F., Fu, P., Grimmond, C. S. B., Hamilton, J. F., Hewitt, C. N., Holloway, M., Kotthaus, S., Lee, J., Metzger, S., Pinguha-Durden, N., Shaw, M., Vaughan, A. R., Wang, X., Wu, R., Zhang, Q., and Zhang, Y. (2020). Measurements of traffic dominated pollutant emissions in a Chinese megacity.
- Stammes, P. (2001). Spectral Radiance Modelling in the UV-Visible Range. In *IRS 2000: Current Problems in Atmospheric Radiation*, pages 385–388.
- Stammes, P., Sneep, M., de Haan, J. F., Veefkind, J. P., Wang, P., and Levelt, P. F. (2008). Effective cloud fractions from the Ozone Monitoring Instrument: Theoretical framework and validation. *Journal of Geophysical Research: Atmospheres*, 113(D16). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2007JD008820>.
- Sullivan, J. T., Apituley, A., Mettig, N., Kreher, K., Knowland, K. E., Allaart, M., PETERS, A., Van Roozendaal, M., Veefkind, P., Ziemke, J. R., Kramarova, N., Weber, M., Rozanov, A., Twigg, L., Sumnicht, G., and Mcgee, T. J. (2022). Tropospheric and stratospheric ozone profiles during the 2019 TROPomi vaLIdation eXperiment (TROLIX-19). *Atmospheric Chemistry and Physics*, 22(17):11137–11153. Publisher: Copernicus Publications.
- Tack, F., Merlaud, A., Iordache, M. D., Pinardi, G., Dimitropoulou, E., Eskes, H., Bommans, B., Veefkind, P., and Van Roozendaal, M. (2021). Assessment of the TROPOMI tropospheric NO₂ product based on airborne APEX observations. *Atmospheric Measurement Techniques*, 14(1):615–646. Publisher: Copernicus GmbH.
- Thürkow, M., Kirchner, I., Kranenburg, R., Timmermans, R. M., and Schaap, M. (2021). A multi-meteorological comparison for episodes of PM₁₀ concentrations in the Berlin agglomeration area in Germany with the LOTOS-EUROS CTM. *Atmospheric Environment*, 244:117946. Publisher: Pergamon.
- Tilstra, L. G., Tuinder, O. N. E., Wang, P., and Stammes, P. (2017). Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY. *Journal of Geophysical Research: Atmospheres*, 122(7):4084–4111. Publisher: Wiley-Blackwell.
- Todd, G. W. and Garber, M. J. (1958). Some Effects of Air Pollutants on the Growth and Productivity of Plants. *Botanical Gazette*, 120(2):75–80. Publisher: The University of Chicago Press.
- UNCTAD (2019). Review of Maritime Transport 2019 - Chapter 4. Technical report, UNCTAD. Issue: October ISBN: 978-92-1-112958-8 ISSN: 0566-7682.
- United Nations (2015). Transforming our world: the 2030 Agenda for Sustainable Development. <https://sdgs.un.org/2030agenda>. Accessed: 2024-02-27.

- Valin, L. C., Russell, A. R., and Cohen, R. C. (2014). Chemical feedback effects on the spatial patterns of the NO_x weekend effect: a sensitivity analysis. *Atmospheric Chemistry and Physics*, 14(1):1–9. Publisher: Copernicus GmbH.
- van der A, R. J., de Laat, A. T., Ding, J., and Eskes, H. J. (2020). Connecting the dots: NO_x emissions along a West Siberian natural gas pipeline. *npj Climate and Atmospheric Science*, 3(1):1–7.
- Van Geffen, J., Eskes, H., Compernelle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., Ter Linden, M., Ludewig, A., Boersma, K. F., and Veeffkind, J. P. (2021). Sentinel-5P TROPOMI NO₂ retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data.
- van Geffen, J., Eskes, H., Compernelle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., ter Linden, M., Ludewig, A., Boersma, K. F., and Veeffkind, J. P. (2022). Sentinel-5P TROPOMI NO₂ retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data. *Atmospheric Measurement Techniques*, 15(7):2037–2060. Publisher: Copernicus GmbH.
- Van Geffen, J., Folkert Boersma, K., Eskes, H., Sneep, M., Ter Linden, M., Zara, M., and Pepijn Veeffkind, J. (2020). S5P TROPOMI NO₂ slant column retrieval: Method, stability, uncertainties and comparisons with OMI. *Atmospheric Measurement Techniques*, 13(3):1315–1335.
- van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., and Veeffkind, J. P. (2021). *TROPOMI ATBD of the total and tropospheric NO₂ data products; Report S5P-KNMI-L2-0005-RP, version 2.2.0*. KNMI, De Bilt The Netherlands.
- Van Heerwaarden, C. C., Van Stratum, B. J., Heus, T., Gibbs, J. A., Fedorovich, E., and Mellado, J. P. (2017). MicroHH 1.0: A computational fluid dynamics code for direct numerical simulation and large-eddy simulation of atmospheric boundary layer flows. *Geoscientific Model Development*, 10(8):3145–3165. Publisher: Copernicus GmbH.
- Van Roy, W., Schallier, R., Van Roozendael, B., Scheldeman, K., Van Nieuwenhove, A., and Maes, F. (2022a). Airborne monitoring of compliance to sulfur emission regulations by ocean-going vessels in the Belgian North Sea area. *Atmospheric Pollution Research*, 13(6):101445. Publisher: Elsevier.
- Van Roy, W. and Scheldeman, K. (2016). Best Practices Airborne MARPOL Annex VI Monitoring. Technical report, MUMM.
- Van Roy, W., Scheldeman, K., Van Roozendael, B., Van Nieuwenhove, A., Schallier, R., Vigin, L., and Maes, F. (2022b). Airborne monitoring of compliance to NO_x emission regulations from ocean-going vessels in the Belgian North Sea. *Atmospheric Pollution Research*, 13(9):101518. Publisher: Elsevier.
- Van Roy, W., Van Nieuwenhove, A., Scheldeman, K., Van Roozendael, B., Schallier, R., Mellqvist, J., and Maes, F. (2022c). Measurement of Sulfur-Dioxide Emissions from

- Ocean-Going Vessels in Belgium Using Novel Techniques. *Atmosphere*, 13(11):1756. Publisher: MDPI.
- Van Roy, W., Van Roozendael, B., Vigin, L., Van Nieuwenhove, A., Scheldeman, K., Merveille, J.-B., Weigelt, A., Mellqvist, J., Van Vliet, J., van Dinther, D., Beecken, J., Tack, F., Theys, N., and Maes, F. (2023). International maritime regulation decreases sulfur dioxide but increases nitrogen oxide emissions in the North and Baltic Sea. *Communications Earth & Environment* 2023 4:1, 4(1):1–16. Publisher: Nature Publishing Group.
- van Stratum, B. J. H., van Heerwaarden, C. C., and Vilà-Guerau de Arellano, J. (2023). The Benefits and Challenges of Downscaling a Global Reanalysis With Doubly-Periodic Large-Eddy Simulations. *Journal of Advances in Modeling Earth Systems*, 15(10):e2023MS003750.
- Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M. F., Jenouvrier, A., and Coquart, B. (1998). Measurements of the NO₂ absorption cross-section from 42 000 cm⁻¹ to 10 000 cm⁻¹ (238-1000 nm) at 220 K and 294 K. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 59(3):171–184.
- Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120:70–83. Publisher: Elsevier.
- Veefkind, P. J., De Haan, J. F., Sneep, M., and Levelt, P. F. (2016). Improvements to the OMI O₂-O₂ operational cloud algorithm and comparisons with ground-based radar-lidar observations. *Atmospheric Measurement Techniques*, 9(12):6035–6049.
- Verbeek, R., van Dinther, D., Mamarik, S., Grigoriadis, A., Weigelt, A., van Vliet, J., Smyth, T., Deakin, A., and Irjala, M. (2022). Scipper: Cost-effectiveness of different approaches for compliance monitoring. Technical report.
- Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J. C., Eskes, H. J., Eichmann, K. U., Fjæraa, A. M., Granville, J., Niemeijer, S., Cede, A., Tiefengraber, M., Hendrick, F., Pazmiño, A., Bais, A., Bazureau, A., Folkert Boersma, K., Bognar, K., Dehn, A., Donner, S., Elokhov, A., Gebetsberger, M., Goutail, F., Grutter De La Mora, M., Gruzdev, A., Gratsea, M., Hansen, G. H., Irie, H., Jepsen, N., Kanaya, Y., Karagkiozidis, D., Kivi, R., Kreher, K., Levelt, P. F., Liu, C., Müller, M., Navarro Comas, M., PETERS, A. J., Pommereau, J. P., Portafaix, T., Prados-Roman, C., Puentedura, O., Querel, R., Remmers, J., Richter, A., Rimmer, J., Cárdenas, C. R., De Miguel, L. S., Sinyakov, V. P., Stremme, W., Strong, K., Van Roozendael, M., Pepijn Veefkind, J., Wagner, T., Wittrock, F., Yela González, M., and Zehner, C. (2021). Ground-based validation of the

- Copernicus Sentinel-5P TROPOMI NO₂ measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks. *Atmospheric Measurement Techniques*, 14(1):481–510.
- Verstraeten, W. W., Neu, J. L., Williams, J. E., Bowman, K. W., Worden, J. R., and Boersma, K. F. (2015). Rapid increases in tropospheric ozone production and export from China. *Nature Geoscience*, 8(9):690–695. Publisher: Nature Publishing Group.
- Vila-Guerau de Arellano, J., Talmon, A., Part, P. B. A. E., and 1990, u. (1990). A chemically reactive plume model for the NO-NO₂-O₃ system. *Atmospheric Environment*, 24A(8):2237–2246.
- Vinken, G. C., Boersma, K. F., Jacob, D. J., and Meijer, E. W. (2011). Accounting for non-linear chemistry of ship plumes in the GEOS-Chem global chemistry transport model. *Atmospheric Chemistry and Physics*, 11(22):11707–11722.
- Vinken, G. C., Boersma, K. F., Maasackers, J. D., Adon, M., and Martin, R. V. (2014a). Worldwide biogenic soil NO_x emissions inferred from OMI NO₂ observations. *Atmospheric Chemistry and Physics*, 14(18):10363–10381. Publisher: Copernicus GmbH.
- Vinken, G. C., Boersma, K. F., Van Donkelaar, A., and Zhang, L. (2014b). Constraints on ship NO_x emissions in Europe using GEOS-Chem and OMI satellite NO₂ observations. *Atmospheric Chemistry and Physics*, 14(3):1353–1369.
- Visser, A. J., Boersma, K. F., Ganzeveld, L. N., and Krol, M. C. (2019). European NO_x emissions in WRF-Chem derived from OMI: impacts on summertime surface ozone. *Atmospheric Chemistry and Physics Discussions*, 2000(x):1–36.
- Visser, A. J., Ganzeveld, L. N., Goded, I., Krol, M. C., Mammarella, I., Manca, G., and Boersma, K. F. (2021). Ozone deposition impact assessments for forest canopies require accurate ozone flux partitioning on diurnal timescales. *Atmospheric Chemistry and Physics*, 21(24):18393–18411. Publisher: Copernicus GmbH.
- Wagner, T., Warnach, S., Beirle, S., Bobrowski, N., Jost, A., Pukite, J., and Theys, N. (2023). Investigation of three-dimensional radiative transfer effects for UV-Vis satellite and ground-based observations of volcanic plumes. *Atmospheric Measurement Techniques*, 16(6):1609–1662. Publisher: Copernicus GmbH.
- Wang, P., Piters, A., Van Geffen, J., Tuinder, O., Stammes, P., and Kinne, S. (2020). Shipborne MAX-DOAS measurements for validation of TROPOMI NO₂ products. *Atmospheric Measurement Techniques*, 13(3):1413–1426. Publisher: Copernicus GmbH.
- Wang, P., Stammes, P., Van Der A, R., Pinardi, G., and Van Roozendaal, M. (2008). FRESCO+: An improved O₂ A-band cloud retrieval algorithm for tropospheric trace gas retrievals. *Atmospheric Chemistry and Physics*, 8(21):6565–6576.
- Wang, X. and Mauzerall, D. L. (2004). Characterizing distributions of surface ozone and its impact on grain production in China, Japan and South Korea: 1990 and 2020.

- Atmospheric Environment*, 38(26):4383–4402. Publisher: Pergamon.
- WHO (2003). Health Aspects of Air Pollution with Particulate Matter , Ozone and Nitrogen Dioxide. Technical report, World Health Organization. Publication Title: Report on a WHO Working Group Bonn, Germany 13–15 January 2003 Issue: January.
- WHO (2024). Ambient (outdoor) air pollution. [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health). Accessed: 2024-02-27.
- Williams, J. E., Folkert Boersma, K., Le Sager, P., and Verstraeten, W. W. (2017). The high-resolution version of TM5-MP for optimized satellite retrievals: Description and validation. *Geoscientific Model Development*, 10(2):721–750.
- Wong, Y., Li, Y., Lin, Z., and Kafizas, A. (2022). Studying the effects of processing parameters in the aerosol-assisted chemical vapour deposition of TiO₂ coatings on glass for applications in photocatalytic NO_x remediation. *Applied Catalysis A*.
- Yoshioka, M., Grosvenor, D. P., Booth, B. B. B., Morice, C. P., and Carslaw, K. S. (2024). Warming effects of reduced sulfur emissions from shipping. *EGUsphere*, pages 1–19. Publisher: Copernicus GmbH.
- Yuan, T., Liu, F., Lamsal, L. N., Song, H., Jcet, U. ., and Gsf, N. (2023). Detecting ship-produced NO₂ plumes and shipping routes in TROPOMI data with a deep learning model. *Authorea Preprints*. Publisher: Authorea.
- Zara, M., Boersma, K. F., De Smedt, I., Richter, A., Peters, E., Van Geffen, J. H., Beirle, S., Wagner, T., Van Roozendaal, M., Marchenko, S., Lamsal, L. N., and Eskes, H. J. (2018). Improved slant column density retrieval of nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: Intercomparison, uncertainty characterisation, and trends. *Atmospheric Measurement Techniques*, 11(7):4033–4058. Publisher: Copernicus GmbH.
- Zara, M., Boersma, K. F., Eskes, H., Denier van der Gon, H., Vilà-Guerau de Arellano, J., Krol, M., van der Swaluw, E., Schuch, W., and Velders, G. J. (2021). Reductions in nitrogen oxides over the Netherlands between 2005 and 2018 observed from space and on the ground: Decreasing emissions and increasing O₃ indicate changing NO_x chemistry. *Atmospheric Environment: X*, 9:100104. Publisher: Elsevier.
- Zhang, Q., Boersma, K. F., Zhao, B., Eskes, H., Chen, C., Zheng, H., and Zhang, X. (2023). Quantifying daily NO_x and CO₂ emissions from Wuhan using satellite observations from TROPOMI and OCO-2. *Atmospheric Chemistry and Physics*, 23(1):551–563. Publisher: Copernicus GmbH.
- Zhang, R., Tie, X., and Bond, D. W. (2003). Impacts of anthropogenic and natural NO_x sources over the U.S. on tropospheric chemistry. *Proceedings of the National Academy of Sciences*, 100(4):1505–1509. Publisher: Proceedings of the National Academy of Sciences.

Zhou, Y., Brunner, D., Spurr, R. J. D., Boersma, K. F., Sneep, M., Popp, C., and Buchmann, B. (2010). Accounting for surface reflectance anisotropy in satellite retrievals of tropospheric NO₂. *Atmospheric Measurement Techniques*, 3(5):1185–1203. Publisher: Copernicus GmbH.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, A., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. A., Lamsal, L., Li, C., Lindfors, A., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O’Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M. R., Saiz-Lopez, A., Spurr, R. J. D., Szykman, J. J., Torres, O., Veefkind, J. P., Veihelmann, B., Wang, H., Wang, J., and Chance, K. (2017). Tropospheric emissions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy and Radiative Transfer*, 186:17–39.

About the author

Christoph was born in Kassel, Germany, and grew up in the small town of Grebenstein. He went to primary school and the gymnasium track at a middle school in Grebenstein. Early on he discovered his talent for numbers and choose a maths-physics profile at high school in Kassel. After his graduation he spent a year backpacking through Australia. discovering his love to explore the world. Coming back to Germany he decided to study physics at the University of Heidelberg. During his bachelor he spent two semesters abroad at Uppsala University in Sweden. In 2017 he finished his bachelor degree with a thesis in particle physics.



He stayed in Heidelberg for his master and focused on environmental physics, as he wanted to study a topic with more societal relevance and real-world connection. In his master thesis at the ICOS lab in Heidelberg he worked on measuring and interpreting $^{14}\text{CO}_2$ in the atmosphere to monitor fossil fuel emissions of power plants and cities. During this time he met his dutch partner Sanne on a trip of his student organization to Warsaw.

After finishing his master degree in 2020 he moved to the Netherlands and started a PhD at Wageningen University. In his PhD he worked on monitoring ship emissions from space with satellite data. The results of this project can be read in this thesis.

In his free time Christoph enjoys hiking, playing disc golf and beach volleyball, and playing board games. His love for nature and adventure is reflected in his choice to build out a van and travel through Europe with his partner.

Peer-reviewed journal publications

Riess, T. C. V. W., Boersma, K. F., Van Vliet, J., Peters, W., Sneep, M., Eskes, H., and Van Geffen, J. (2022). Improved monitoring of shipping NO_2 with TROPOMI: Decreasing NO_x emissions in European seas during the COVID-19 pandemic. *Atmospheric Measurement Techniques*, 15(5):1415–1438.

Riess, T. C. V. W., Boersma, K. F., Van Roy, W., de Laat, J., Dammers, E., and van Vliet, J. (2023). To new heights by flying low: comparison of aircraft vertical NO₂ profiles to model simulations and implications for TROPOMI NO₂ retrievals. *Atmospheric Measurement Techniques*, 16(21):5287–5304.

Riess, T. C. V. W., Boersma, K. F., Prummel, A., van Stratum, B. J. H., de Laat, J., and van Vliet, J. (2024). Estimating NO_x Emission of Individual Ships from TROPOMI NO₂ Plumes, *under review, preprint available at*: <http://dx.doi.org/10.2139/ssrn.4858709>

Li, H., Zheng, B., Ciais, P., Boersma, K. F., **Riess, T. C. V. W.**, Martin, R. V., Broquet, G., A. R. v. d., Li, H., Hong, C., Lei, Y., Kong, Y., Zhang, Q., and He, K. (2023). Satellite reveals a steep decline in China's CO₂ emissions in early 2022. *Science Advances*, 9(29):eadg7429. eprint: <https://www.science.org/doi/pdf/10.1126/sciadv.adg7429>

Other scientific publications

Hammer, S., Maier, F., Wang, Y., Van Der Woude, A. M., van der Molen, M. K., Kneuer, T., **Riess, T. C. V.**, Nguyen, L., Meijer, H. A. J., Ramonet, M., Karstens, U., and Levin, I. (2021). Scientific-technical concept for requirements of a fossil fuel observation system. Technical report, available at: <https://www.icos-cp.eu/sites/default/files/2021-01/D1.2.%20Scientific-technical%20concept%20for%20requirements%20of%20a%20fossil%20fuel%20observation%20system.pdf>

Acknowledgements

This is it. A bit more than four years of PhD work are coming to an end with this thesis. Four years of enthusiasm, doubts and hard work. Starting a PhD in a new place during COVID lockdowns is surely far from ideal, but with the support of many people, I believe it became a success. I want to thank everyone who has been part of this journey.

First, my supervisor and promoter Folkert. From the day I came to your office to discuss the position I had a good feeling. The last years have proven my feeling was right: Many calls and personal meetings at your house and my apartment made the start during COVID less difficult. Our exchanges always had a great mix of free time and work related topics which I greatly value. You pushed me when my writing was once again too vague but also supported me whenever I doubted my work. I truly learned a lot from you in these four years, so thank you for everything!

Jasper, this thesis would not exist without you. I always considered you my second supervisor and the academic rules will not change that. Your passion, knowledge and solution-oriented way of working have played a big part in my PhD project, and I am very thankful for that!

I also want to thank Wouter Peters, for being my 2nd promoter and for his valuable input and feedback on my work. I am especially thankful for including me into his carbon group meetings during lockdown, which was a great way to get connected with the group and to get feedback on my work.

Auke and Anne-Wil, thank you for all the coffee breaks, cookies, random chats, banter and enthusiasm for science. And above all, thank you for being my paranymphs! You always made me feel welcome at MAQ.

Wouter, thank you many discussions about camping gear and sports stats, always with the necessary degree of nerdyness. You showed me the beauty in bad weather, although I still prefer blue sky over clouds.

I want to thank all former and current members of the *Temple of Science*: Auke V., Sjoerd, Ruben, Mary Rose, Kim, Esther, Thomas, Sarah, Martin, Raquel and Menno. I am grateful for the shorter and longer exchanges, being able to pick your brains when needed or just talk about our weekends. Coming to Wageningen would have not been the same without you!

Imme, thank you for many bike and train rides to and from WUR, you made the

commute much more enjoyable.

I also want to thank the rest of MAQ for creating a great working atmosphere. I am grateful for the many discussions and shared lunches, beach volleyball sessions and PhD dinners.

A big thanks also to the group at ILT for always welcoming me without any expectations. Thank you all my friends in Germany, the Netherlands and elsewhere, for providing the much needed distraction from my PhD. Special thanks to the crew in Heidelberg, I hope we can continue our tradition of colloquia, and I am looking forward to climbing Kilimanjaro with you.

Thanks to Edith and Sep for the many gezellig evenings in Vinkeveen and Doorwerth, and especially for taking care of Sanne and me in the last stressful weeks of my PhD.

Ganz besonderen Dank an meine Eltern und Schwestern, die mich immer unterstützt haben, egal ob es um meine Studienwahl, meine Reisen oder meine Entscheidung für einen PhD ging. Diese Freiheit in meinen Lebensentscheidungen empfinde ich als besonders. Ohne euch wäre ich nicht da, wo ich jetzt bin. Danke, dass ich mich immer auf euch verlassen kann.

Last but not least, I want to thank my partner Sanne. You have been by my side during the last years, always supporting me or distracting me when I needed it. Building a camper van and traveling through Northern Europe with you was a dream come true. I am looking forward to many more adventures with you!



*Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment*

D I P L O M A

for specialised PhD training

The Netherlands research school for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

Tobias Christoph Valentin
Werner Riess

born on the 26th of October 1992 in Kassel, Germany

has successfully fulfilled all requirements of the
educational PhD programme of SENSE.

Wageningen, 28 October 2024

SENSE coordinator PhD education

Dr Ir Peter Vermeulen

The SENSE Director

Dr Jampel Dell'Angelo



The SENSE Research School declares that **Tobias Christoph Valentin Werner Riess** has successfully fulfilled all requirements of the educational PhD programme of SENSE with a work load of 43.2 EC, including the following activities:

SENSE PhD Courses

- o Environmental research in context (2020)
- o Research in context activity: Science/policy workshop on ship NOx emissions (2023)

Other PhD and Advanced MSc Courses

- o Joint training on atmospheric composition, EUMETSAT/ESA/ECMWF, Online (2020)
- o The Essentials of Scientific Writing and Presenting, Wageningen University (2020)
- o Intensive writing week, Wageningen University (2021)
- o DARE-UK International Summer School , Southhampton, UK (2022)
- o Reviewing a Scientific Manuscript (2022)
- o AMIGO workshop, IGAC, Belgium (2023)
- o Big data exploration and object-oriented programming with Python, WIMEK (2023)

Management and Didactic Skills Training

- o Teaching assistant in the MSc course 'Atmospheric Composition and Air Quality' (2020-2022)
- o Teaching assistant in the BSc course 'Introduction Atmosphere' (2021)
- o Organize internal AVES-OCULUS workshop at WUR (2022)
- o Supervising 2 MSc students with thesis (2023)
- o PhD representative MAQ (2022-2023)

Oral Presentations

- o *Detecting NO₂ from shipping over European Seas with TROPOMI and OMI.* OMI/TROPOMI workshop, 27-29 October 2020, Online
- o *Improving cloud retrieval for better sensitivity to shipping NO₂ with TROPOMI.* EGU, 19-30 April 2021, Online
- o *Aircraft validation reveals a 20% low bias in TROPOMI NO₂ over sea caused by TM5 a priori profiles.* EGU, 23-28 April 2023, Vienna, Austria
- o *TROPOMI-based NO_x emission estimates of individual ships.* EGU, 14-19 April 2024, Vienna, Austria

This research received funding from 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 program under grant agreement Nr.814893.

MicroHH numerical simulations were carried out on the Dutch national e-infrastructure with the support of SURF Cooperative within the financially supported NWO/ENW project on Exact and Physic Science (2023.003).

Financial support from the Meteorology and Air Quality group at Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by TCVW Riess.

Printed by proefschriftmaken.nl