To new heights by flying low: Comparison of aircraft vertical NO$_2$ profiles to model simulations and implications for TROPOMI NO$_2$ retrievals

Tobias Christoph Valentin Werner Riess[1], Klaas Folkert Boersma[1,2], Ward Van Roy[3], Jos de Laat[2], Enrico Dammers[4], and Jasper van Vliet[5]

[1]Department of Meteorology and Air Quality, Wageningen University, Wageningen, the Netherlands
[2]Climate Observations Department, Royal Netherlands Meteorological Institute, De Bilt, the Netherlands
[3]Royal Belgian Institute of Natural Sciences, Brussels, Belgium
[4]Climate, Air and Sustainability (CAS), Netherlands Organisation for Applied Scientific Research (TNO), Utrecht, the Netherlands
[5]Human Environment and Transport Inspectorate, the Netherlands

Correspondence: T. Christoph V.W. Riess (christoph.riess@wur.nl)

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, ~~which was already~~ equipped with a sniffer sensor system, measuring NO$_x$ ($= $ NO $+$ 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. ~~We observe~~ 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 ~~improve~~ improves the surface level pollution estimates, ~~but~~. 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.

Keywords. NO$_2$, TROPOMI, validation, shipping, profile, aircraft

1 Introduction

Satellite data of air pollutants is increasingly used for policy making, which requires reliable retrievals. This paper evaluates TROPOMI tropospheric NO_2 columns by comparing aircraft measurements of NO_2 profiles over the polluted North Sea to chemical transport models, and studying uncertainty and bias in the TROPOMI NO_2 retrieval from modeled profile shapes.

Nitrogen oxides (NO_x = NO+NO_2) decrease air quality, having negative impact on human health and environment. NO_2 is known to cause cardiovascular and respiratory diseases (Luo et al., 2016). Depending on chemical regime, nitrogen oxides also lead to surface O_3 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 CO_2 emission (IMO, 2020; European Comission, 2022). While NO_x emissions from most anthropogenic sectors have been decreasing in recent years in western countries (e.g ~~Zara et al. (2021); For~~ Zara et al., 2021; Fortems-Cheiney et al., 2021; Jiang et al., 2022 and references therein), intensity of ~~ocean going~~ ocean-going ships has been and is expected to keep rising (IMO, 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 ~~emission~~ reductions in land-based ~~sources~~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 (IMO, 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~~and~~, 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 NO_2 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., 2014), the higher spatial resolution and lower noise of TROPOMI make single ship plume detection possible (Georgoulias et al., 2020). Recent studies succeeded to discriminate NO_2 ship plume signatures from the background using TROPOMI tropospheric NO_2 columns~~(Kurchaba et al. (2021); Finch et al. (2022))~~ (Kurchaba et al., 2021; Finch et al., 2022). However, the validity of TROPOMI NO_2 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 NO_2 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 ~~therefor~~ therefore given to improve these aspects: recent updates in the cloud retrieval used for the TROPOMI

NO$_2$ ~~columns~~ column retrieval lead to better agreement with independent data and reduce the known negative bias in tropo-
spheric NO$_2$ columns (Van Geffen et al., 2022a; 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 NO$_2$. 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 opera-
tional TROPOMI NO$_2$ product are very coarse compared to the ground pixel size of the measurements (3.5x5.5 km^2 at nadir)
while NO$_2$ profiles close to spatially confined emission sources such as ships are expected to vary significantly within kilome-
ters ~~(?Griffin et al., 2019; Ialongo et al., 2020; Chen et al., 2005)~~(Douros et al., 2023; Griffin et al., 2019; Ialongo et al., 2020; Chen et al.,
. 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 NO$_2$
profiles.

The TROPOMI NO$_2$ product allows the user to replace the a priori profiles with their own modelled or measured profiles
(e.g. ~~Visser et al. (2019); ?~~Visser et al., 2019; Douros et al., 2023). ~~?~~Douros et al. (2023) used the high-resolution CAMS en-
semble mean NO$_2$ profile to replace the TM5-MP a priori NO$_2$ profiles in the calculation of the air mass factor (AMF) and
to create an improved European TROPOMI NO$_2$ product. They found significant changes in resulting tropospheric columns
with increases at hot-spot regions of ~~up to 30%~~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 ~~a polluted area~~ the polluted North Sea
with major shipping routes and nearby industrial and densely populated centres~~: the North Sea~~. We combine ten spiral flights
with three horizontal scans to obtain vertical NO$_2$ 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$_2$ ~~profile~~
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 ~~of~~satellite trace gas ~~retrieval~~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.

2 Materials

The following section gives an overview of the data used and their sources, starting with the TROPOMI instrument in
90 Sect. 2.1~~and~~, followed by the aircraft, LOTOS-EUROS model data & ship location data in subsections 2.2, 2.3 & 2.4, respectively.

Table 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°x1°	No
	v.2.2	July 2021 - ~~present~~ November 2021		
Reprocessed PAL	v2.3.1	April 2018 - ~~present~~ November 2021	TM5 1°x1°	Yes
TROPOMI~~$_{CAMS}$~~ $_{CAMS}$	v2.3.1	April 2018 - ~~present~~ November 2021	CAMS 0.1°x0.1°	Yes

2.1 TROPOMI NO$_2$ satellite data

Table 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 ~~the Sentinel-5 Precursor~~ (S5P~~) satellite~~, which was launched in October 2017, and provides
95 ~~scientific quality~~ 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~~, and at~~. At the latitude of the North Sea (52°N) ~~the instrument~~ S5P frequently overpasses the same ground scene ~~multiple times~~ twice per day. The spatial resolution is 5.5 x 3.5 km^2 for nadir pixels, and 5.5 x 14 km^2 for pixels at the edge of TROPOMI's swath.
100 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) ~~(Platt and Stutz, 2008)~~, 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 ($0.5*10^{15}$ molec/cm^2) is determined by the uncertainty in the spectral fitting procedure and has been validated in Tack
105 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 NO$_2$ in the troposphere (Martin et al., 2002; Boersma et al., 2004). For the TROPOMI NO$_2$ 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,
110 surface albedo, ~~residual~~ cloud fraction and height, and the a priori vertical NO$_2$ 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 NO$_2$ close to the sea surface (e.g. ~~Eskes and Boersma (2003); Vinken et al. (2014)~~ Eskes and Boersma, 2003; Vinken et al., 2014). Good knowledge of ~~these~~ 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., 2022b). While the cloud algorithm used in the TROPOMI operational NO$_2$ retrieval has recently been ~~improved to better account for residual cloud scattering~~ modified to provide a more accurate cloud pressure estimate for partially cloudy scenes (FRESCO+ wide) (Riess et al., 2022; Van Geffen et al., 2022a), the a priori vertical NO$_2$ profiles remain a major source of AMF uncertainty, especially over sea.

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 ~~to the lower troposphere (±1500~~(<30 m) to 1500 m.

The aircraft is equipped with a sniffer sensor system measuring NO$_2$, SO$_2$, and CO$_2$. This system is developed for the purpose of monitoring ~~of~~ 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 NO$_2$*, which is produced when the air sample passes through a reaction cell filled with O$_3$ and proceeds through NO+O$_3$ -> NO$_2$*+ 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 (NO + 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 ~~make sure~~ 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)~~ 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
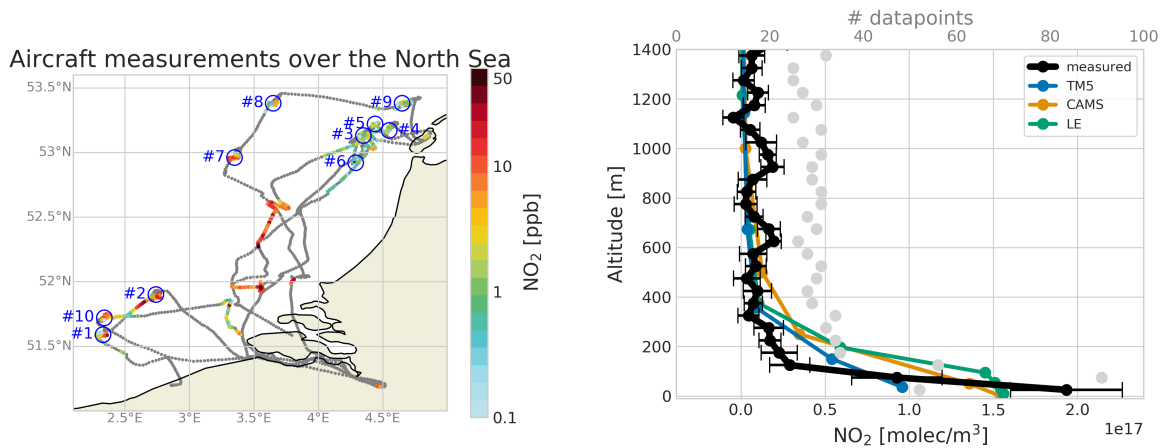
Figure 1. Left: Routes of all aircraft flights during the campaign. The 30 second mean NO$_2$ mixing ratio is shown as color for ~~the~~ measurements in flight heights below 200 m. Blue circles indicate the locations of the ~~aircraft profiles~~spiral flights. Right: Mean vertical NO$_2$ profiles for the aircraft data (black), co-sampled TM5 (blue, ~~Eskes and van Geffen (2021); Huijnen et al. (2010)~~ Williams et al., 2017; Eskes and van Geffen, 2021; Huijnen et al., 2010), CAMS (yellow, ~~METEO FRANCE et al. (2022)~~ METEO FRANCE et al., 2022; Marécal et al., 2015) and LOTOS-EUROS (green, Manders et al., 2017). The light gray dots indicate the number of 10 second NO$_2$ 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).

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. 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.

2.3 LOTOS-EUROS model simulations

We use LOTOS-EUROS version 2.2.002 (LE, ~~Manders et al. (2017); Thürkow et al. (2021)~~Manders et al., 2017; Thürkow et al., 2021) at 2x2 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 50.5-54.5°N and 1.5-5.0°E with a spin up time 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.
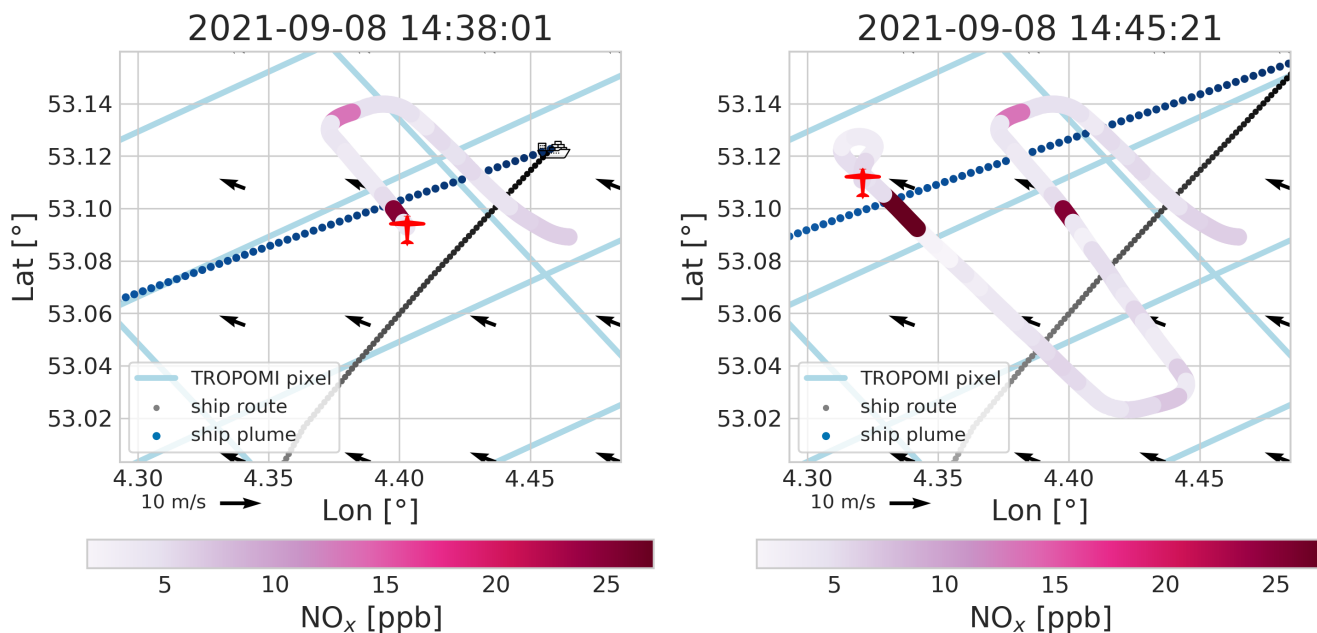
Figure 2. Two snapshots of one of the horizontal scans: Black and blue dots show ship path and plume center location at the moment indicated by the timestamp, respectively, with lighter colors indicated older locations. In pink we see the flight path with the color indicating the measured NO_x concentration. 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.

Key characteristics of LOTOS-EUROS and other model data used in this study can be found in Table 2.

Table 2. Main characteristics of the model products used.

model	LOTOS-EUROS	CAMS	TM5
horizontal resolution	2x2 km²	0.1°x0.1°	1°x1°
emissions	CAMS-REG-AP_v5.1	CAMS-REG-AP	see William
meteorology	~~IFS~~ ECMWF Integrated Forecasting System (IFS)	IFS	ERA-Interim
vertical mixing scheme	see ECMWF (2015) with Monin–Obukhov length calculated as in Golder (1972)	model dependent	see Holtslag
full description	Manders et al. (2017)	~~METEO FRANCE et al. (2022)~~ Marécal et al. (2015); METEO FRANCE et al. (2022)	Williams et a Eskes and va Huijnen et al

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 (IMO, 2014). The historic AIS data set used here was made available to the Dutch Human Environment and Transport.

3 Aircraft NO_2 interpretation and representation at the scale of a TROPOMI pixel

Table 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	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

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 NO_2 concentrations reveal substantial variability within the spatial extent of a TROPOMI pixel~~The~~, see Fig. 2. The observed spatial heterogeneity of NO_2 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). In general, aircraft spatial sampling characteristics are not uniform across a TROPOMI pixel as evident from Fig. 2.~~ 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.

3.1 ~~Representative~~ Representativeness of NO$_2$ vertical profile measurements

Pixel-scale aircraft NO$_2$ profiles

We first take care to ~~represent the~~ ensure the representativeness of the aircraft NO$_2$ ~~measurements~~ 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 NO$_2$ 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 NO$_2$ concentration in a pixel. For each vertical profile flight listed in Table 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~~and with~~. No ship plume signatures where 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 disproportionally much time in the plume. We apply these as multiplicative correction factors to the in-plume and out-of-plume NO$_2$ values to improve the spatial representativeness of the vertical NO$_2$ profile for the TROPOMI pixel.

Plume NO$_x$-to-NO$_2$ conversion

The NO$_2$ 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 NO$_2$ concentrations due to small time-delay between the NO$_x$ and NO sampling in the Ecotech instrument, as mentioned in Sec. 2.2, and illustrated in Fig. 3 (right panel). When the aircraft samples background air, the NO$_2$ values inferred 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 NO$_2$ concentrations via local NO$_2$: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 NO$_2$: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 NO$_2$ portion typically increases to 0.45 within some 15-30 minutes after emission following entrainment of O$_3$, and subsequent NO$_2$ formation via the NO + O$_3$ 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 NO$_2$ enhancements of several ppbs, but since near-zero, background air NO$_2$ levels differed by a few ppb between flights on different days, we re-calibrated the aircraft NO$_2$ concentrations to ensure that the measured near-zero NO$_2$ levels at altitudes above 250 m are on average consistent with NO$_2$ 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
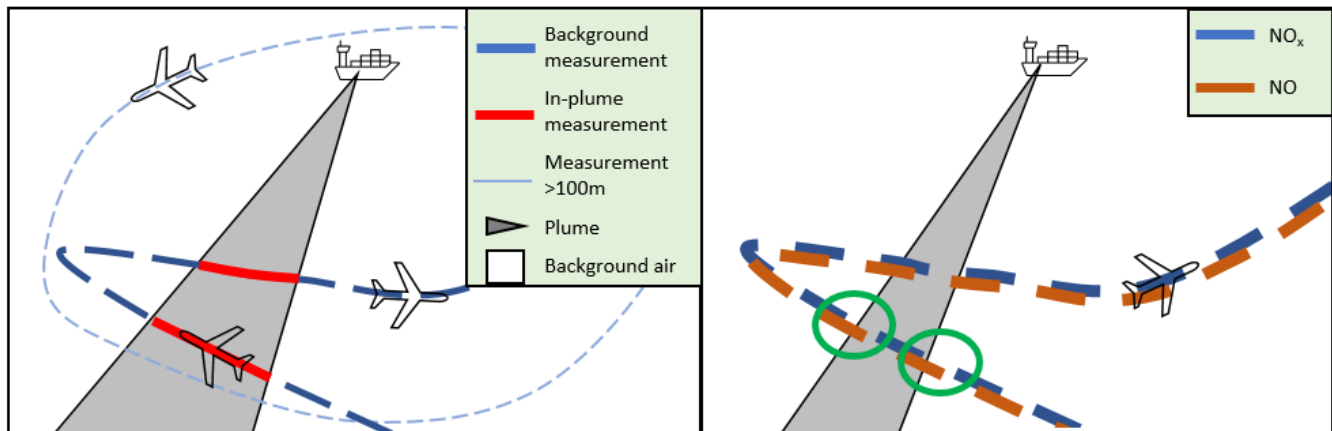
Figure 3. Sketches of profile flights visualizing the corrections. Left: The grey area indicates the part of the 2D-plane covered by a plume and the thick line the aircraft measurements in the polluted layer, with red showing in-plume measurements and blue indicating background sampling. The mismatch 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 NO_2 values, respectively.

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
215   uncertainty range.

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

3.2   Observed vertical NO_2 profiles

We now present the vertical NO_2 profiles obtained from the BN2 aircraft measurements over the North Sea following the procedure sketched in Sect. 3.1. Each of these vertical NO_2 profiles is spatially representative for the spatial scale of a TROPOMI
220   pixel. For time and location of the profiles taken see Table 3. Aircraft NO_2 measurements were aggregated in 50 m altitude bins, where the reported altitude is the mean of the lower and upper boundary of each bin.

The aircraft data shows the highest NO_2 concentrations close to the sea surface, ~~and a sharp drop above~~ strongly decreasing within the lowest 100 m (Fig. 1)~~,~~. This is in agreement with the CO_2 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
225   period from the TM5-MP, CAMS, and LOTOS-EUROS models (see Sect. 2.3). The mean simulated NO_2 profiles coinciding with the aircraft flights show NO_2 pollution up to 200 m and above (Fig. 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 NO_2 profiles according to their distinct meteorolog-
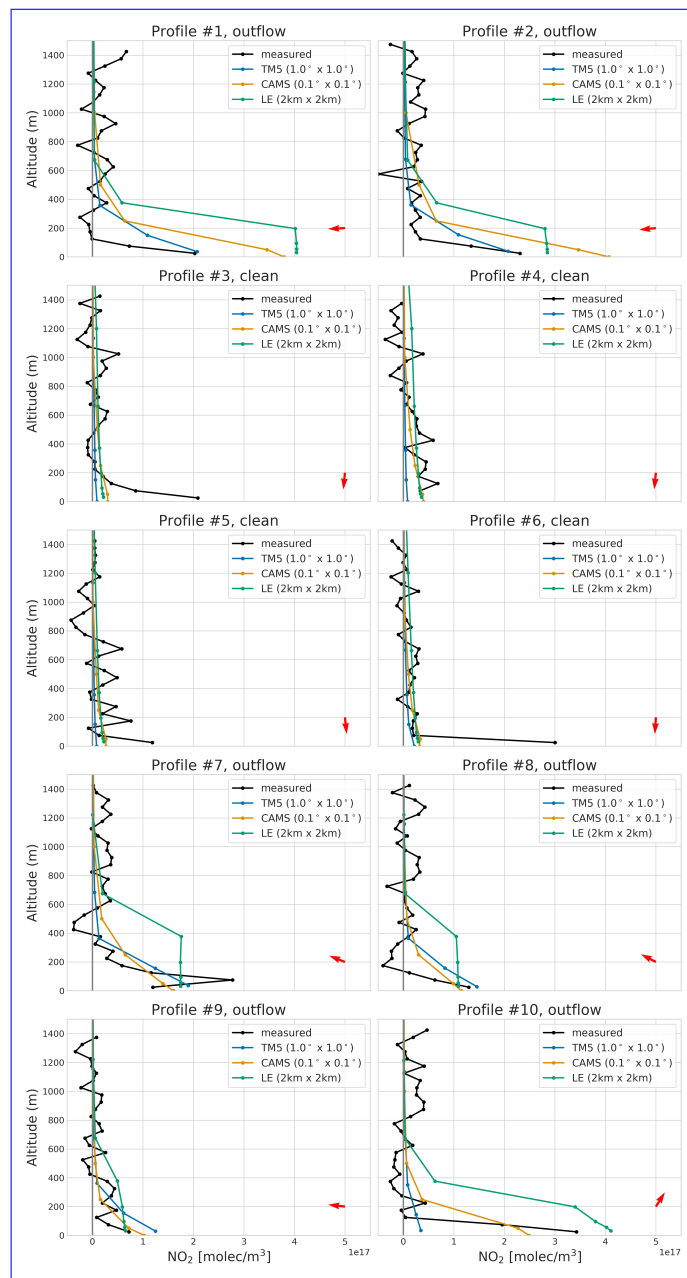
Figure 4. Profiles of all flights as well as coinciding TM5, CAMS ensemble mean and LOTOS-EUROS profiles. The red arrows indicate the mean measured wind direction during the profile flights. The indicaters 'outflow' and 'clean' in the subtitles follow the classification in Sec. 3.2.

ical circumstances. Fig. 4 shows the individual measured and modeled profiles with the numbering consistent to Table 3. For
230 uncorrected profiles and the uncertainty estimates see Fig. S4. Meteorological conditions such as mean wind directions reveal
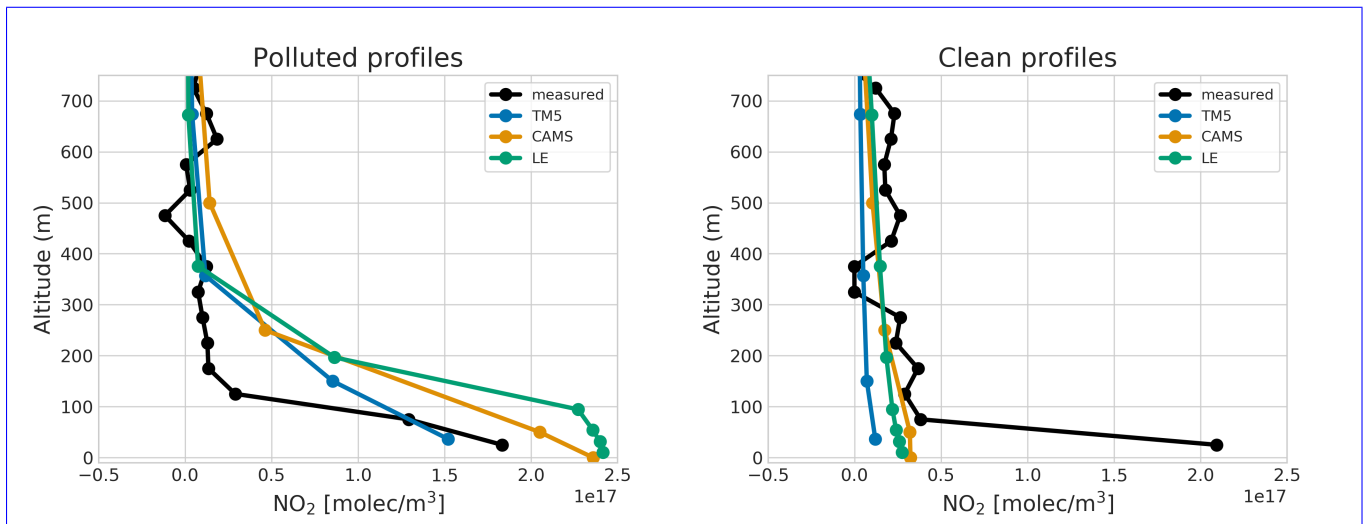
Figure 5. Mean aircraft-measured profiles as well as coinciding TM5, CAMS ensemble mean and LOTOS-EUROS profiles for land outflow (left, profiles 1, 2, 7, 8, 9, 10) and clean conditions/northerly winds (right, profiles 3, 4, 5, 6).

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 NO_2 concentrations. Hereafter we classify these profiles as 'land outflow', and 'clean', see Fig. 5. A more complete description of the general chemical and meteorological conditions during each flight can be found in Supplement S1.

235 **NO_2 profiles during land outflow - profiles 1, 2, 7, 8, 9, 10**

Fig. 6 shows the observed and simulated NO_2 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 show pollution close to the sea surface (see also the left panel of Fig .5). 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

240 - show elevated NO_2 at usually up to 200 m and above, overestimating . This gives an overestimation in the total NO_2 in the column. The measured and modelled potential temperature profiles (Fig. S2) 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. 6. This means that emissions

245 within the cell are originating can originate from land-based sources as well as ships. Likewise, boundary layer dynamics are a mix of sea and land characteristics. Overall Nonetheless, TM5 profiles show too little only slightly less NO_2 in the lowest layer and overestimate the NO_2 above. Nonetheless, than the LOTOS-EUROS, CAMS and the measured profiles for outflow cases (see Fig. 5, left). Overall, the coarse TM5 columns show reasonable agreement with TROPOMI retrieved columns during outflow conditions with the exception of profile 10. 10 (see Fig. 4 and S5).
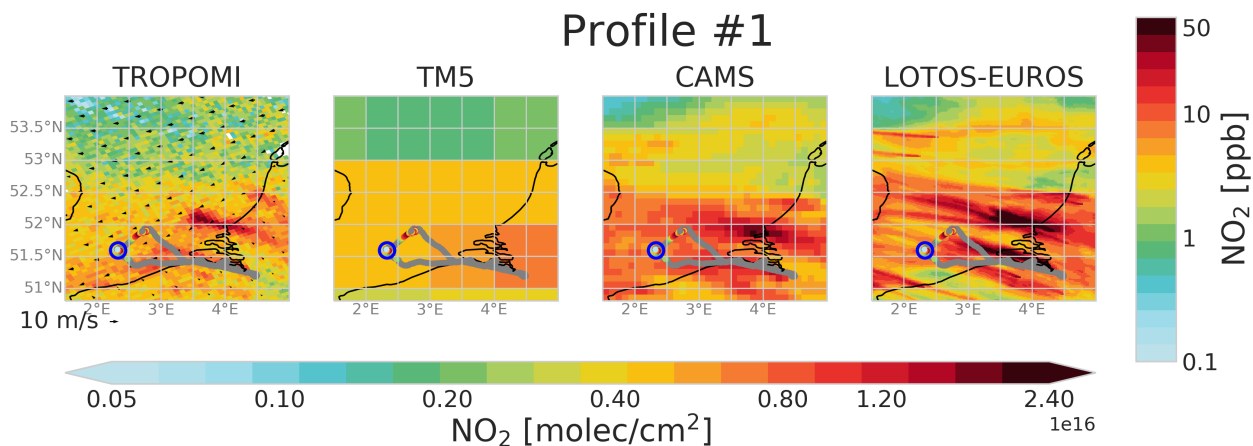
Figure 6. NO$_2$ columns (indicated by the bottom color bar) as seen by TROPOMI and several model products for the time of the first profile measurement~~as indicated by the bottom colorbar~~. ~~Overlayed are the~~ The aircraft measurements are overlayed in grey for flights above 200 m and in colors below as indicated by the colorbar on the right~~as well as wind~~. Wind speed and direction at 10 m from ERA5 are indicated by the arrows in the left panel.

250 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. 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 and 5: While surface NO$_2$ levels in
255 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. Additionally, 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 ($\approx$0.2 kg/s for these two sources).
260 TROPOMI and modelled NO$_2$ columns during the other profile flights can be found in Supplement S4.

A special case is profile ~~7.~~ 7 on September 8th. This is the only profile with clearly enhanced NO$_2$ above 100 m (see also S5 for the CO$_2$ 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 NO$_2$ in the column. This enhanced NO$_2$ 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
265 the stable surface layer. This hypothesis is supported by parts of the flight on June 2nd, when enhanced NO$_2$ 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. 7. The observed NO$_2$ layer at 300 m is also present in the co-sampled LOTOS-EUROS profile (as a thicker NO$_2$ layer around 500 m) but not in CAMS. These findings also demonstrate that the
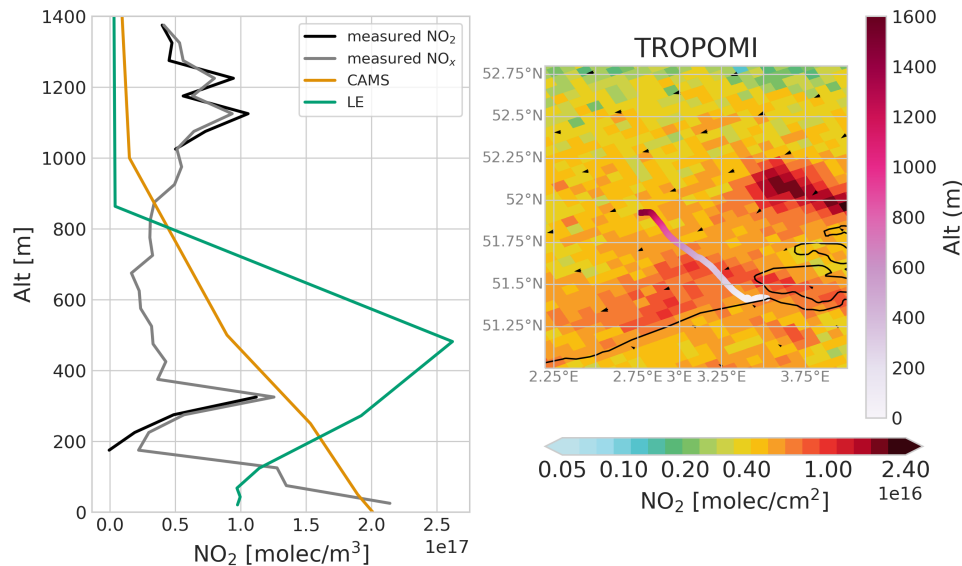
Figure 7. Left: Measured and modeled Vertical distribution of NO_2 along the flight path indicated on the right. This is not a vertical profile in the strict sense, as the sampling took place over $\approx$70 km horizontal extend. During part of the flight the airplane instrumentation was operating in a different mode so that no NO_2 data is available. However, NO_x (gray) was sampled throughout the whole flight and indicates a thin pollution layer between 300 and 400 m.

aircraft instrumentation is able not only to detect high NO_2 values in fresh plumes but also to capture diluted NO_2 pollution

270   from land. Additionally, this suggests that at least for profile 2 (which was sampled right before) enhanced NO_2 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) and thus possibly masking the low-level NO_2 from ships.

275   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. ~~The low surface pollution of~~ Similar to the other models, TM5 ~~showcasts the limitations of a coarse resolution whereas the~~ shows too high NO_2 at 200 m

280   ~~hints~~ 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.

NO$_2$ profiles during clean conditions - profiles 3, 4, 5, 6
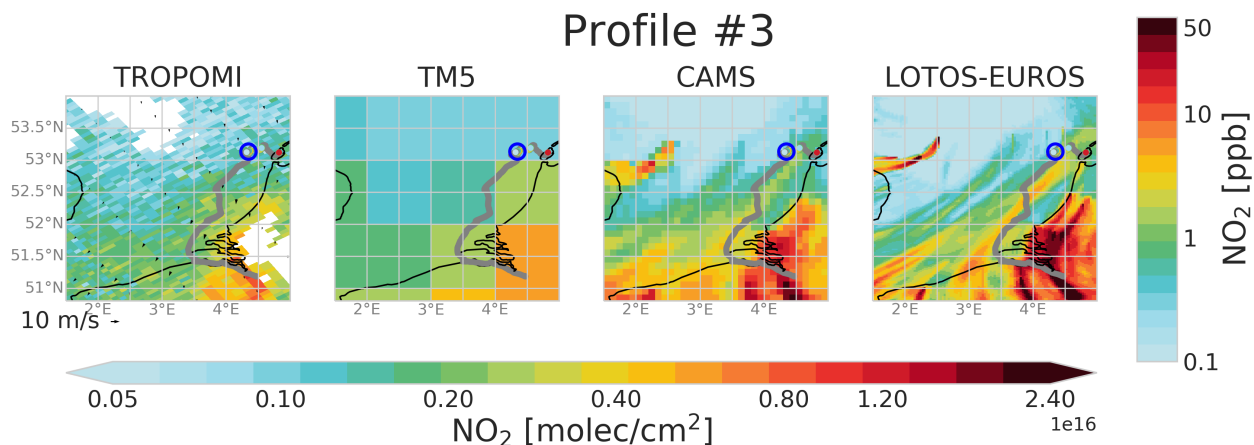
Figure 8. As Fig~~.~~ 6 but now for the third profile. Profile 4 is taken in collocation with the same TROPOMI orbit and its location is shown in Fig. S5.

285   Fig. 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. 8 and Fig. S1. The potential temperature profile on 22 July 2021 (see Fig. S2) 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

290   were taken right above the shipping lane, ~~in CAMS and LOTOES-EUROS~~ marked by the blue circle in Fig. 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. 5, 8 & S5.).

The observed profiles 4 and 5 (see Fig. 4) agree reasonably well with the models, showing little NO$_2$ enhancement close to the sea surface. On the other hand, profiles 3 and 6 show strong NO$_2$ enhancements in the lowest 50 m, in contrast to the models.

295   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. S4, leading to a more multi-profile-average 'in plume' NO$_2$ concentration) shows very similar surface NO$_2$ values of $\approx 1.5*10^{17}$ molec/m^3 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.

300   The mean clean profile in the right panel of Fig. 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 Validation of TROPOMI NO$_2$ over the North Sea

4.1 Recalculate AMFs

With the observed vertical NO$_2$ profiles we can calculate a ~~more accurate~~ modified TROPOMI NO$_2$ column, replacing the coarse TM5 a priori in the retrieval with aircraft-measurement based vertical profiles. ~~The~~ As the measured NO$_2$ 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 ~~$A_{\text{trop,l}}$~~ $A_{\text{trop},l}$ provided in the TROPOMI files as well as the NO$_2$ column density ~~$x_{\text{l,meas}}$ of layer 1~~ $x_{l,\text{meas}}$ 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; ?)~~ (Visser et al., 2019; Douros et al., 2023). The adjusted vertical, tropospheric column can then be calculated as $N_{v,\text{trop,ADJ}} = N_s/M_{\text{trop,ADJ}}$. ~~As the measured NO$_2$ profiles only extend to 1400 m, we use TM5 profiles to fill the gap to the tropopause.~~

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. 9). The shallow boundary layer depth over sea in combination with the low surface albedo values ($\approx$0.04) emphasize the difficulty to detect air pollution over sea with satellite remote sensing, despite the high signal-to-noise ratio and resolution of TROPOMI NO$_2$.

4.2 Tropospheric columns

We compare ~~total~~ vertical tropospheric columns of NO$_2$ retrieved by TROPOMI (operational, PAL & CAMS) as well as measured columns. Lastly, we add the new product TROPOMI$_{\text{ADJ}}$ which includes a re-calculation of the AMFs and vertical tropospheric NO$_2$ columns using the measured profiles following Sect. 4.1.

Table 4 shows the mean columns densities of all datasets mentioned above as well as their Pearson ~~R and Roote~~ correlation coefficient and Root Mean Squared Error (RMSE) against the aircraft data. The ten aircraft measured NO$_2$ column densities averaged at 3.37*10^{15} molec/cm^2. This is significantly higher than the coinciding operational TROPOMI (2.42*10^{15}

molec/cm^2) and TROPOMI$_{\sim\!PAL}$ ~~PAL~~ (2.47*10^{15} molec/cm^2) data. Using the re-calculated AMFs an average column density ~~$N_{v,trop,adj}$~~ $N_{v,trop,adj}$ of 2.89 (2.71-3.23)*10^{15} molec/cm^2 is determined. This is $\approx 20\ (12-33)\%$ higher than the TROPOMI

products and brings the satellite retrievals closer to the ~~aircraft-measured columns~~ columns determined from the aircraft measurements, showing a significant negative bias in operational TROPOMI NO$_2$ columns. The TROPOMI$_{\sim\!CAMS}$ ~~CAMS~~ dataset (see Sect. 2.1) is closer to the measured columns at mean columns of 3.03*10^{15} molec/cm^2. It should be noted that CAMS ~~shows systematically higher~~ NO$_2$ columns (see Figs. 6, 8 & S5) are systematically higher compared to measurements and TM5. TROPOMI$_{CAMS}$ ~~and TROPOMI$_{ADJ}$~~ 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 ~~root mean square error (RMSE)~~ RMSE of the TROPOMI columns towards the aircraft columns is reducing going from the operational (1.26*10^{15} molec/cm^2) to TROPOMI$_{CAMS}$ CAMS (0.99*10^{15} molec/cm^2) data and smallest for the aircraft-adjusted columns at 0.77*10^{15} molec/cm^2.

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 NO$_2$ alone.

Table 4. Tropospheric NO$_2$ 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 on the uncertainties shown in Fig. S3

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 [%]
~~height~~ aircraft	3.37	-	-	-
TROPOMI	2.42	0.82	1.26	-28
TROPOMI$_{\sim\!PAL}$ PAL	2.47	0.83	1.24	-27
TROPOMI$_{\sim\!CAMS}$ CAMS	3.03	0.87	0.99	-10
TROPOMI$_{\sim\!ADJ}$ ADJ	2.89 (2.71-3.23)	0.91	0.77	-14

4.3 The land-sea contrast in TROPOMI NO$_2$ retrieval

As a contrasting case, Fig. 9 compares the sea NO$_2$ profiles to NO$_2$ 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 NO$_2$ 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 NO$_2$ concentrations measured ~~during TROLIX~~ 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 NO_2 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 NO_2 close to the surface - emphasizes the challenge of accurate satellite retrieval of NO_2 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.
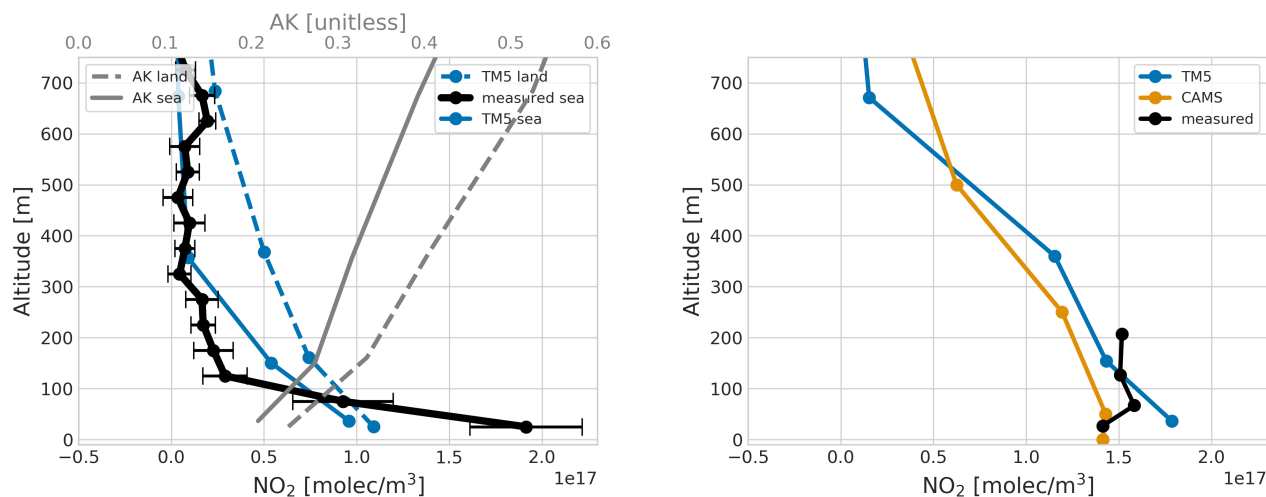
Figure 9. Left: The solid blue line shows mean TM5 profiles coinciding with the aircraft profiles (black). The dashed blue line shows simultaneous TM5 NO_2 profiles at the Cabauw tower in the Netherlands. Additionally, the mean TROPOMI averaging kernel profiles for land (sampled for all TROPOMI pixels 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).

5 Discussion

We evaluated the TROPOMI tropospheric vertical NO_2 column retrieval over the North Sea. For this, we measured ten vertical NO_2 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 ~~retrieved~~ TROPOMI NO_2 ~~columns~~column 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 TROPOMI$_{PAL}$ PAL data using TM5 for a periori 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. 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 $1°x1°$ where ship emissions are smeared out over a larger area and time. The exaggerated vertical mixing and underestimation of he 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 improve 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$_{CAMS}$ 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 ~~show~~ 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)~~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 NO_2 profiles utilizing an aircraft already equipped for emissions monitoring, direct NO_2 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 NO_2 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.

6   Conclusion

This study clearly shows the need for additional evaluation of vertical NO_2 profiles over sea for both model and TROPOMI validation while providing a recipe for such an analysis. We present ten vertical profiles of NO_2 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 NO_2 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 NO_2 concentrations near the surface while overestimating NO_2 above 100 m. The higher model resolution of CAMS and LOTOS-EUROS results in more accurate surface NO_2 values, while at the same time vertical mixing is exaggerated compared to our observations. Additionally, CAMS and LOTOS-EUROS vertical NO_2 columns are too high compared to aircraft and TROPOMI data.

Furthermore, the comparison between observed and modeled vertical NO_2 profiles, along with the examination of TROPOMI averaging kernels over land and sea, stresses the significant challenges involved in accurately retrieving satellite NO_2 columns over sea, where vertical sensitivity to NO_2 is 20% lower than over land, because of lower surface albedo and confinement of NO_2 pollution in a thin marine boundary layer.

When replacing the TM5 a priori profiles with the aircraft-measured NO_2 profiles in the TROPOMI AMF calculation, we find a significant increase of the retrieved vertical NO_2 columns of $\approx 20\,(12-33)\%$, showing substantially improved agreement with aircraft-measured columns. Our findings align with previous studies (e.g. by ~~?Pseftogkas et al. (2022); Lorente et al. (2017)~~ 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.

Data availability. The corrected aircraft NO_2 profiles and co-sampled TM5 profiles are available at $https://zenodo.org/record/$7928291. TROPOMI L2 NO_2 and TM5 data are publicly available via the copernicus open access hub (https://scihub.copernicus.eu). The TROPOMI$_{CAMS}$ data set is available on the temis portal ($https://www.temis.nl/airpollution/no2col/no2_euro_tropomi_cams.php$). CAMS data is available at the Copernicus Atmospheric Data Store ($https://ads.atmosphere.copernicus.eu$). LOTOS-EUROS data can be made available upon reasonable request by contacting the author (christoph.riess@wur.nl).

Author contributions. TCVWR, KFB and JvV designed the study in consultation with WvR and JdL. JdL choose flight dates with forecasted favorable conditions. WvR was the operator of the flights and performed the measurements. TCVWR and KFB led the writing of the manuscript with contributions from all other co-authors. ED ran the LOTOS-EUROS simulations and assisted in their interpretation.

Competing interests. One co-author is a member of the editorial board of Atmospheric Measurement Techniques. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Acknowledgements. The authors want to thank Arnoud Frumau and TNO for measurement and provision of the NO_2 data at Cabauw, Auke Visser for his help with re-calculating the AMFs & Jordi Vila in interpreting the shallow mixing layer. Special thanks also to the MUMM pilots Paul De Prest, Gilbert Witters, Alexander Vermeire, Geert Present, Dries Noppe and Gilbert Witters.

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

References

Beirle, S., Platt, U., von Glasow, R., Wenig, M., and Wagner, T.: Estimate of nitrogen oxide emissions from shipping by satellite remote sensing, Geophysical Research Letters, 31, 4–7, https://doi.org/10.1029/2004GL020312, 2004.

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., and Wagner, T.: Pinpointing nitrogen oxide emissions from space, Science Advances, 5, https://doi.org/10.1126/sciadv.aax9800, 2019.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, Journal of Geophysical Research D: Atmospheres, 109, https://doi.org/10.1029/2003jd003962, 2004.

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.: Validation of OMI tropospheric NO2 observations during INTEX-B and application to constrain NOx emissions over the eastern United States and Mexico, Atmospheric Environment, 42, 4480–4497, https://doi.org/10.1016/j.atmosenv.2008.02.004, 2008.

Boersma, K. F., Eskes, H. J., Richter, A., De Smedt, I., Lorente, A., Beirle, S., Van Geffen, J. H., 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.: 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, 6651–6678, https://doi.org/10.5194/amt-11-6651-2018, 2018.

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., Schauffler, S., Donnelly, S., Atlas, E., Roberts, J., Flocke, F., Hübler, G., and Fehsenfeld, F.: An investigation of the chemistry of ship emission plumes during ITCT 2002, Journal of Geophysical Research D: Atmospheres, 110, 1–15, https://doi.org/10.1029/2004JD005236, 2005.

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: Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4. 3.2, Earth Syst. Sci. Data, 10, 1987–2013, 2018.

Douros, J., Eskes, H., van Geffen, J., Boersma, K. F., Compernolle, S., Pinardi, G., Blechschmidt, A.-M., Peuch, V.-H., Colette, A., and Veefkind, P.: Comparing Sentinel-5P TROPOMI NO 2 column observations with the CAMS regional air quality ensemble, Geoscientific Model Development, 16, 509–534, https://doi.org/10.5194/GMD-16-509-2023, 2023.

ECMWF: IFS Documentation CY41R1 - Part IV: Physical Processes, Tech. rep., ECMWF, https://www.ecmwf.int/en/elibrary/74328-ifs-documentation-cy41r1-part-iv-physical-processes, 2015.

Ecotech, A. G.: Serinus 40 Oxides of Nitrogen Analyser – Acoem UK, https://www.acoem.co.uk/product/ecotech/serinus-40-oxides-of-nitrogen-analyser/, 2023.

Escudero, M., Segers, A., Kranenburg, R., Querol, X., Alastuey, A., Borge, R., De La Paz, D., Gangoiti, G., and Schaap, M.: Analysis of summer O3 in the Madrid air basin with the LOTOS-EUROS chemical transport model, Atmospheric Chemistry and Physics, 19, 14 211–14 232, https://doi.org/10.5194/acp-19-14211-2019, 2019.

Eskes, H. and van Geffen, J.: Product user manual for the TM5 NO2 , SO2 and HCHO profile auxiliary support product, Tech. Rep. 1.0.0, KNMI, de Bilt, https://sentinel.esa.int/documents/247904/2474726/PUM-for-the-TM5-NO2-SO2-and-HCHO-profile-auxiliary-support-product.pdf/de18a67f-feca-1424-0195-756c5a3df8df, 2021.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmospheric Chemistry and Physics, 3, 1285–1291, https://doi.org/10.5194/acp-3-1285-2003, 2003.

495 European Comission: Reducing emissions from the shipping sector, https://climate.ec.europa.eu/eu-action/transport-emissions/ reducing-emissions-shipping-sector_en, 2022.

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.: Transport impacts on atmosphere and climate: Shipping, Atmospheric Environment, 44, 4735–4771, https://doi.org/10.1016/j.atmosenv.2009.04.059, 2010.

500 Finch, D. P., Palmer, P. I., and Zhang, T.: Automated detection of atmospheric NO2 plumes from satellite data: A tool to help infer anthropogenic combustion emissions, Atmospheric Measurement Techniques, 15, 721–733, https://doi.org/10.5194/AMT-15-721-2022, 2022.

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.: Analysis of the Anthropogenic and Biogenic NOx Emissions Over 2008–2017: Assessment of the Trends in the 30 Most Populated Urban Areas in Europe, Geophysical Research Letters, 48, e2020GL092 206, https://doi.org/10.1029/2020GL092206, 505 2021.

Georgoulias, A. K., Boersma, K. F., Van Vliet, J., Zhang, X., Van Der A, R., Zanis, P., and De Laat, J.: Detection of NO2 pollution plumes from individual ships with the TROPOMI/S5P satellite sensor, Environmental Research Letters, 15, 124 037, https://doi.org/10.1088/1748-9326/abc445, 2020.

Goldberg, D. L., Anenberg, S. C., Griffin, D., McLinden, C. A., Lu, Z., and Streets, D. G.: Disentangling the Impact of 510 the COVID-19 Lockdowns on Urban NO2 From Natural Variability, Geophysical Research Letters, 47, e2020GL089 269, https://doi.org/10.1029/2020GL089269, 2020.

Golder, D.: Relations among stability parameters in the surface layer, Boundary-Layer Meteorology, 3, 47–58, https://doi.org/10.1007/BF00769106/METRICS, 1972.

Griffin, D., Zhao, X., McLinden, C. A., Boersma, F., Bourassa, A., Dammers, E., Degenstein, D., Eskes, H., Fehr, L., Fioletov, V., Hayden, 515 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., Veefkind, P., and Wolde, M.: High-Resolution Mapping of Nitrogen Dioxide With TROPOMI: First Results and Validation Over the Canadian Oil Sands, Geophysical Research Letters, 46, 1049–1060, https://doi.org/10.1029/2018GL081095, 2019.

Holtslag, A. A. M. and Boville, B. A.: Local Versus Nonlocal Boundary-Layer Diffusion in a Global Climate Model, Journal of Climate, 6, 1825–1842, https://doi.org/https://doi.org/10.1175/1520-0442(1993)006<1825:LVNBLD>2.0.CO;2, 1993.

520 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.: 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, https://doi.org/10.5194/gmdd-3-1009-2010, 2010.

525 Ialongo, I., Virta, H., Eskes, H., Hovila, J., and Douros, J.: Comparison of TROPOMI/Sentinel-5 Precursor NO2 observations with ground-based measurements in Helsinki, Atmospheric Measurement Techniques, 13, 205–218, https://doi.org/10.5194/AMT-13-205-2020, 2020.

IMO: Nitrogen oxides (NOx) – Regulation 13, https://www.imo.org/en/OurWork/Environment/Pages/Nitrogen-oxides-(NOx) -–-Regulation-13.aspx, 2013.

IMO: AIS transponders, https://www.imo.org/en/OurWork/Safety/Pages/AIS.aspxhttp://www.imo.org/OurWork/Safety/Navigation/Pages/ 530 AIS.aspx, 2014.

IMO: 4th IMO Greenhouse Gas study, https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/Fourth%20IMO%20GHG%20Study%202020%20-%20Full%20report%20and%20annexes.pdf, 2020.

Jalkanen, J. P., Johansson, L., and Kukkonen, J.: A comprehensive inventory of ship traffic exhaust emissions in the European sea areas in 2011, Atmospheric Chemistry and Physics, 16, 71–84, https://doi.org/10.5194/acp-16-71-2016, 2016.

535 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.: Decadal Variabilities in Tropospheric Nitrogen Oxides Over United States, Europe, and China, Journal of Geophysical Research: Atmospheres, 127, e2021JD035 872, https://doi.org/10.1029/2021JD035872, 2022.

Johansson, L., Jalkanen, J. P., and Kukkonen, J.: Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution, Atmospheric Environment, 167, 403–415, https://doi.org/10.1016/j.atmosenv.2017.08.042, 2017.

540 Köhler, M., Ahlgrimm, M., and Beljaars, A.: Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model, Quarterly Journal of the Royal Meteorological Society, 137, 43–57, https://doi.org/10.1002/QJ.713, 2011.

Kurchaba, S., Van Vliet, J., Meulman, J. J., Verbeek, F. J., and Veenman, C. J.: Improving evaluation of NO_2 emission from ships using spatial association on TROPOMI satellite data, GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems, pp. 454–457, https://doi.org/10.1145/3474717.3484213, 2021.

545 Lama, S., Houweling, S., Boersma, K. F., Aben, I., Denier van der Gon, H. A. C., and Krol, M. C.: Estimation of OH in urban plumes using TROPOMI-inferred NO_2 CO, Atmospheric Chemistry and Physics, 22, 16 053–16 071, https://doi.org/10.5194/ACP-22-16053-2022, 2022.

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., Van Roozendael, M., Wang, Y., Wagner, T., Beirle, S., Lin, J. T., Krotkov, N., Stammes, P., Wang, P., Eskes, H. J., and Krol, M.: Structural un-

550 certainty in air mass factor calculation for NO_2 and HCHO satellite retrievals, Atmospheric Measurement Techniques, 10, 759–782, https://doi.org/10.5194/amt-10-759-2017, 2017.

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.: Quantification of nitrogen oxides emissions from build-up of pollution over Paris with TROPOMI, Scientific Reports, 9, 1–10, https://doi.org/10.1038/s41598-019-56428-5, 2019.

555 Luo, K., Li, R., Li, W., Wang, Z., Ma, X., Zhang, R., Fang, X., Wu, Z., Cao, Y., and Xu, Q.: 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, 1–13, https://doi.org/10.1038/SREP38328, 2016.

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.,

560 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.: Curriculum vitae of the LOTOS-EUROS (v2.0) chemistry transport model, Geoscientific Model Development, 10, 4145–4173, https://doi.org/10.5194/GMD-10-4145-2017, 2017.

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,

565 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., Rouïl, 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.: A regional air

quality forecasting system over Europe: The MACC-II daily ensemble production, Geoscientific Model Development, 8, 2777–2813, https://doi.org/10.5194/GMD-8-2777-2015, 2015.

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 Koelemeijer, R. B.: An improved retrieval of tropospheric nitrogen dioxide from GOME, Journal of Geophysical Research: Atmospheres, 107, 9–1, https://doi.org/10.1029/2001JD001027, 2002.

Mellqvist, J., Conde, V., Beecken, J., and Ekholm, J.: 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, Tech. rep., Chalmers University of Technology, Göteburg, 2017.

METEO FRANCE, MET Norway, IEK, IEP-NRI, KNMI, TNO, FMI, ENEA, and BSC: CAMS Regional: European air quality analysis and forecast, https://confluence.ecmwf.int/display/CKB/CAMS+Regional%3A+European+air+quality+analysis+and+forecast+data+documentation, 2022.

Namdar-Khojasteh, D., Yeghaneh, B., Maher, A., Namdar-Khojasteh, F., and Tu, J.: Assessment of the relationship between exposure to air pollutants and COVID-19 pandemic in Tehran city, Iran, Atmospheric Pollution Research, 13, 101 474, https://doi.org/10.1016/J.APR.2022.101474, 2022.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy, Physics of Earth and Space Environments, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-75776-4, 2008.

Pseftogkas, A., Koukouli, M. E., Segers, A., Manders, A., Geffen, J. v., Balis, D., Meleti, C., Stavrakou, T., and Eskes, H.: Comparison of S5P/TROPOMI Inferred NO2 Surface Concentrations with In Situ Measurements over Central Europe, Remote Sensing 2022, Vol. 14, Page 4886, 14, 4886, https://doi.org/10.3390/RS14194886, 2022.

Richter, A., Eyring, V., Burrows, J. P., Bovensmann, H., Lauer, A., Sierk, B., and Crutzen, P. J.: Satellite measurements of NO2 from international shipping emissions, Geophysical Research Letters, 31, 1–4, https://doi.org/10.1029/2004GL020822, 2004.

Riess, T. C. V. W., Boersma, K. F., Van Vliet, J., Peters, W., Sneep, M., Eskes, H., and Van Geffen, J.: Improved monitoring of shipping NO2 with TROPOMI: Decreasing NOx emissions in European seas during the COVID-19 pandemic, Atmospheric Measurement Techniques, 15, 1415–1438, https://doi.org/10.5194/AMT-15-1415-2022, 2022.

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.: Nitrogen oxides in the free troposphere: Implications for tropospheric oxidants and the interpretation of satellite NO2 measurements, Pedro Campuzano-Jost, 15, 19, https://doi.org/10.5194/egusphere-2022-656, 2022.

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.: Evaluation of the LOTOS-EUROS NO 2 simulations using ground-based measurements and S5P/TROPOMI observations over Greece, Atmos. Chem. Phys, 21, 5269–5288, https://doi.org/10.5194/acp-21-5269-2021, 2021.

Sullivan, J. T., Apituley, A., Mettig, N., Kreher, K., Knowland, K. E., Allaart, M., Piters, A., Van Roozendael, M., Veefkind, P., Ziemke, J. R., Kramarova, N., Weber, M., Rozanov, A., Twigg, L., Sumnicht, G., and Mcgee, T. J.: Tropospheric and stratospheric ozone profiles during the 2019 TROpomi vaLIdation eXperiment (TROLIX-19), Atmospheric Chemistry and Physics, 22, 11 137–11 153, https://doi.org/10.5194/ACP-22-11137-2022, 2022.

Tack, F., Merlaud, A., Iordache, M. D., Pinardi, G., Dimitropoulou, E., Eskes, H., Bomans, B., Veefkind, P., and Van Roozendael, M.: Assessment of the TROPOMI tropospheric NO2product based on airborne APEX observations, Atmospheric Measurement Techniques, 14, 615–646, https://doi.org/10.5194/AMT-14-615-2021, 2021.

Thürkow, M., Kirchner, I., Kranenburg, R., Timmermans, R. M., and Schaap, M.: A multi-meteorological comparison for episodes of PM10 concentrations in the Berlin agglomeration area in Germany with the LOTOS-EUROS CTM, Atmospheric Environment, 244, 117 946, https://doi.org/10.1016/J.ATMOSENV.2020.117946, 2021.

610 Van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J. C., Sneep, M., Linden, M. T., Ludewig, A., Folkert Boersma, K., and Pepijn Veefkind, J.: Sentinel-5P TROPOMI NO2retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data, Atmospheric Measurement Techniques, 15, 2037–2060, https://doi.org/10.5194/AMT-15-2037-2022, 2022a.

Van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., and Veefkind, J. P.: TROPOMI ATBD of the total and tropospheric NO 2 data products, S5p/TROPOMI, https://sentinel.esa.int/documents/247904/2476257/sentinel-5p-tropomi-atbd-no2-data-products, 2022b.

615 Van Roy, W., Schallier, R., Van Roozendael, B., Scheldeman, K., Van Nieuwenhove, A., and Maes, F.: Airborne monitoring of compliance to sulfur emission regulations by ocean-going vessels in the Belgian North Sea area, Atmospheric Pollution Research, 13, 101 445, https://doi.org/10.1016/J.APR.2022.101445, 2022a.

Van Roy, W., Scheldeman, K., Van Roozendael, B., Van Nieuwenhove, A., Schallier, R., Vigin, L., and Maes, F.: Airborne monitoring of compliance to NOx emission regulations from ocean-going vessels in the Belgian North Sea, Atmospheric Pollution Research, 13, 620 101 518, https://doi.org/10.1016/J.APR.2022.101518, 2022b.

Van Roy, W., Van Nieuwenhove, A., Scheldeman, K., Van Roozendael, B., Schallier, R., Mellqvist, J., and Maes, F.: Measurement of Sulfur-Dioxide Emissions from Ocean-Going Vessels in Belgium Using Novel Techniques, Atmosphere, 13, 1756, https://doi.org/10.3390/ATMOS13111756/S1, 2022c.

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, 625 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.: 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, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Vinken, G. C., Boersma, K. F., Jacob, D. J., and Meijer, E. W.: Accounting for non-linear chemistry of ship plumes in the GEOS-Chem global chemistry transport model, Atmospheric Chemistry and Physics, 11, 11 707–11 722, https://doi.org/10.5194/acp-11-11707-2011, 630 2011.

Vinken, G. C., Boersma, K. F., Van Donkelaar, A., and Zhang, L.: Constraints on ship NOx emissions in Europe using GEOS-Chem and OMI satellite NO2 observations, Atmospheric Chemistry and Physics, 14, 1353–1369, https://doi.org/10.5194/acp-14-1353-2014, 2014.

Visser, A. J., Boersma, K. F., Ganzeveld, L. N., and Krol, M. C.: European NOx emissions in WRF-Chem derived from OMI: impacts on summertime surface ozone, Atmospheric Chemistry and Physics Discussions, 2000, 1–36, https://doi.org/10.5194/acp-2019-295, 2019.

635 Wang, P., Piters, A., Van Geffen, J., Tuinder, O., Stammes, P., and Kinne, S.: Shipborne MAX-DOAS measurements for validation of TROPOMI NO2 products, Atmospheric Measurement Techniques, 13, 1413–1426, https://doi.org/10.5194/amt-13-1413-2020, 2020.

Williams, J. E., Folkert Boersma, K., Le Sager, P., and Verstraeten, W. W.: The high-resolution version of TM5-MP for optimized satellite retrievals: Description and validation, Geoscientific Model Development, 10, 721–750, https://doi.org/10.5194/gmd-10-721-2017, 2017.

Wong, Y., Li, Y., Lin, Z., and Kafizas, A.: Studying the effects of processing parameters in the aerosol-assisted chemi-640 cal vapour deposition of TiO2 coatings on glass for applications in photocatalytic NOx remediation, Applied Catalysis A, https://doi.org/10.1016/j.apcata.2022.118924, 2022.

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.: 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 NOx chemistry, Atmospheric Environment: X, 9, 100 104,

645   https://doi.org/10.1016/J.AEAOA.2021.100104, 2021.