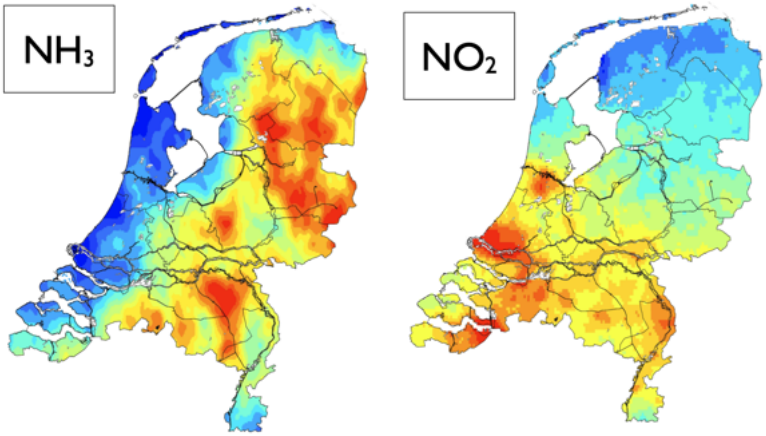
Royal Netherlands
Meteorological Institute
Ministry of Infrastructure
and Water Management

Using Satellite Observations for Assessing the Spatial and Temporal Variation of Nitrogen Emissions and Deposition in the Netherlands

E. Dammers, T. Wizenberg (TNO)
H. Eskes, F. Cifuentes, R. van der A, J. Ding (KNMI)
R. Wichink Kruit, S. van der Graaf (RIVM)
S. Li (CML)
H. Kros (WUR)

De Bilt, 2026| KNMI-publicatie 26-02



Technical Report

Using Satellite Observations for Assessing the Spatial and Temporal Variation of Nitrogen Emissions and Deposition in the Netherlands

Author(s)	E. Dammers, T. Wizenberg (TNO) H. Eskes, F. Cifuentes, R. van der A, J. Ding (KNMI) R. Wichink Kruit, S. van der Graaf (RIVM) S. Li (CML) Hans Kros (WUR)

Date	22 October 2025
Number of pages	126
Project name	NKS-SAGEN

TNO innovation for life

Koninklijk Nederlands Meteorologisch Instituut
Ministerie van Infrastructuur en Waterstaat

Rijksinstituut voor Volksgezondheid en Milieu
Ministerie van Volksgezondheid, Welzijn en Sport

Universiteit Leiden

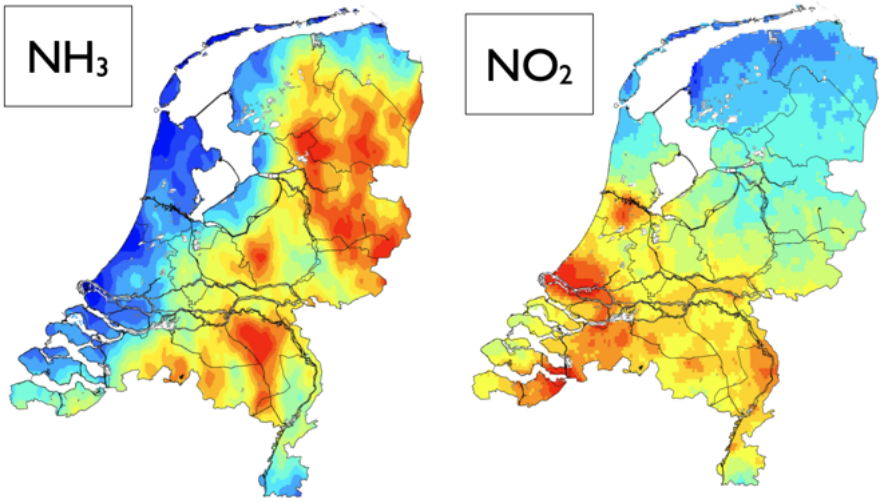
WAGENINGEN UNIVERSITY & RESEARCH

Samenvatting

Het monitoren van de twee belangrijkste spelers in de stikstofproblematiek, de gassen ammoniak (NH_3) en stikstofdioxide (NO_2), is van groot belang om de gevolgen voor natuur, mens en klimaat in kaart te brengen. In Nederland is deze monitoring momenteel gebaseerd op (1) instrumenten die de concentraties en depositie op de grond op verschillende locaties meten, (2) emissieschattingen op basis van gerapporteerde emissiecijfers, in combinatie met (3) modellen om de hoeveelheid depositie te berekenen. Momenteel wordt er geen operationeel gebruik gemaakt van satellietmetingen.

Dit rapport presenteert de resultaten van een onderzoek naar de toegevoegde waarde van satellieten voor het monitoren van reactief stikstof in Nederland. Deze studie maakt deel uit van het Nationaal Kennisprogramma Stikstof (NKS), gefinancierd door het Ministerie van Landbouw, Visserij, Voedselzekerheid en Natuur (LVVN). Het NKS-programma heeft als doel het waarneemnetwerk en de modellering van stikstof in Nederland verder te ontwikkelen en te verbeteren. Dit programma is mede gestart naar aanleiding van de aanbevelingen in het rapport van de commissie Hordijk ("Meer meten, robuuster rekenen", gepubliceerd op 15 juni 2020).

De gassen NO_2 en NH_3 worden beide gemeten met satellietinstrumenten. Figuur S1 toont de verdeling van deze twee gassen over Nederland, waargenomen vanuit de ruimte.

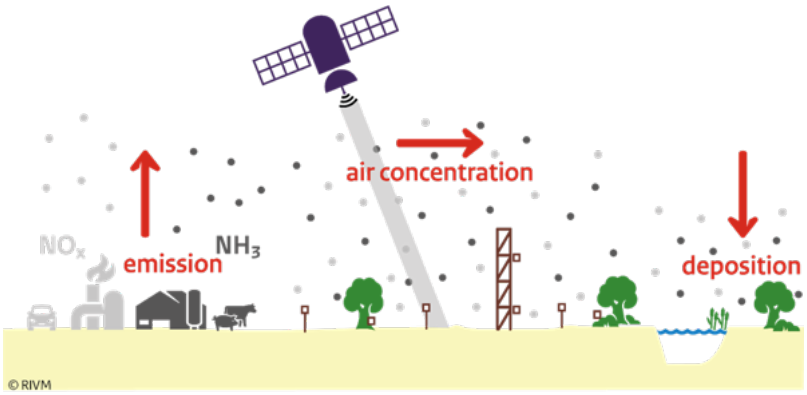


Figuur S1. Jaarlijks gemiddelde ammoniak- (links) en stikstofdioxideverdeling (rechts) in Nederland, zoals waargenomen door de satellietinstrumenten CrIS en TROPOMI. Ammoniak wordt voornamelijk aangetroffen in het oosten van het land, met landbouw en mestaanwending als belangrijkste bronnen. De stikstofdioxideconcentraties zijn hoog in het westen van het land, gekoppeld aan bijdragen van wegverkeer, woningen, industrie, havens, scheepvaart en luchtvaart.

Om de depositie van reactieve stikstof te kwantificeren, is gedetailleerde kennis nodig van de emissies, menging en transport, chemische reacties in de atmosfeer en depositieprocessen. Dit wordt geïllustreerd in figuur S2. Merk op dat het transport van reactieve stikstof grotendeels plaatsvindt in hogere luchtlagen tot ongeveer een kilometer hoogte in de atmosfeer. Satellieten meten de totale kolomhoeveelheid NO_2 en NH_3 (grijze arcering). Dit heeft meerdere voordelen: a) De kolommen kunnen worden omgezet in op waarnemingen gebaseerde emissieschattingen om de gerapporteerde emissies te controleren; b) De satelliet neemt stikstofpluimen waar terwijl ze zich van de bron naar de depositiegebieden verplaatsen, wat informatie bevat over het atmos-

ferische transport en depositie; c) De satellieten meten overal en vullen informatielacunes tussen de meetstations op.

Satellietwaarnemingen bevatten echter beperkte informatie over de concentraties nabij het aardoppervlak. Daarom vormen satellietwaarnemingen een aanvulling op de waarnemingen aan het aardoppervlak en kunnen ze het meetnet op de grond niet vervangen. De toegevoegde waarde van satellietmetingen wordt uitgebreider beschreven in de NKS-whitepaper "Het gebruik van satellietmetingen voor het monitoren van stikstof in Nederland".



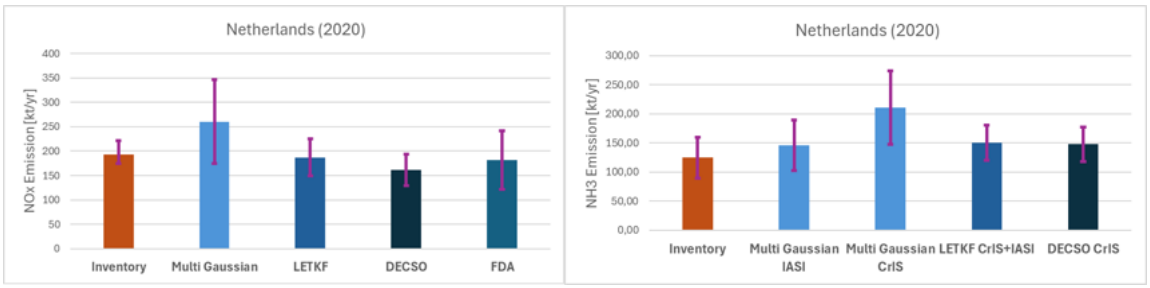
Figuur S2: Schematisch overzicht van stikstofstromen in de atmosfeer. Emissies van NOx en NH_3 in de lucht verspreiden zich door menging en transport in de atmosfeer, worden omgezet door chemische reacties en komen uiteindelijk op het aardoppervlak terechtvia de depositie. De LML-, MAN- en COTAG-netwerken van het RIVM leveren waarnemingen van concentraties en depositie nabij het aardoppervlak. De satelliet levert extra informatie over concentraties boven het aardoppervlak, tot ongeveer 1 km hoogte, en vult hiaten tussen de meetlocaties.

Dit rapport geeft een gedetailleerd overzicht van het NKS satellietonderzoek dat in de jaren 2023-2024 is uitgevoerd. Er is gebruikgemaakt van NH_3-waarnemingen van de CrIS- en IASI-instrumenten, en NO_2-waarnemingen van het TROPOMI-instrument. Het werk is uitgevoerd door wetenschappers van TNO, KNMI, RIVM en de universiteiten van Leiden en Wageningen. Binnen NKS zijn twee promovendi aangesteld. Dit onderzoek bestond uit de volgende studies:

- Het uitvoeren van meerjarige benchmarksimulaties met de drie modellen OPS, EMEP en LOTOS-EUROS, en het evalueren van deze runs met grondmetingen van LML, MAN en COTAG (hoofdstuk 2).
- Het evalueren van de satellietmetingen op basis van grondgebonden kolommen van NH_3 en NO_2, gemeten met FTIR- en MAXDOAS/PANDORA-instrumenten in Nederland en wereldwijd (hoofdstuk 3).
- Het evalueren van de simulaties van LOTOS-EUROS en EMEP met satellietmetingen van NO_2 door TROPOMI (hoofdstuk 4) en NH_3 door CrIS en IASI (hoofdstuk 5).
- Het schatten van de emissies van NH_3 en NO_x op basis van deze satellietmetingen met behulp van meerdere methoden, geïntroduceerd in paragraaf 6.2. De resultaten voor de vier methoden worden besproken in paragraaf 6.3 tot en met 6.7.

De hoofdstukken in dit rapport bevatten beschrijvingen van de gebruikte modellen, metingen en methoden, inclusief relevante verwijzingen naar de wetenschappelijke literatuur.

De evaluatie van de TROPOMI, CrIS en IASI satellietgegevens met grondgebaseerde FTIR, MAXDOAS en Pandora waarnemingen laat over het algemeen hoge temporele correlaties zien. De systematische fouten vallen doorgaans binnen de 25%. Dit toont aan dat de satellietgegevens van voldoende kwaliteit zijn om belangrijke nieuwe informatie over reactief stikstof op te leveren.



Figuur S3: De jaargemiddelde totale emissie in Nederland voor de vier verschillende methodieken om emissies af te leiden uit satellietgegevens, met de bijbehorende onzekerheidsmarges, voor stikstofoxiden (links) en ammoniak (rechts). Deze emissies zijn vergeleken met de gerapporteerde emissies, aangeduid als "inventarisatie".

De vergelijking van de gesimuleerde NO_2-kolommen van LOTOS-EUROS en EMEP4NL met de dagelijkse TROPOMI-gegevens laat over het algemeen een goede ruimtelijke en temporele overeenkomst zien. In beide modellen worden systematische verschillen waargenomen tussen de satelliet en de gemodelleerde velden, in de orde van 10 tot 25%, afhankelijk van de locatie en het model. Een analyse voor het EMEP4NL-model toont aan dat, hoewel het model realistische verdelingen en pluimen van vervuiling simuleert, die pluimen soms langer en verschoven zijn in vergelijking met TROPOMI. Deze bevindingen benadrukken de algehele betrouwbaarheid van beide modellen voor het simuleren van de verdeling van NO_2 boven Nederland, maar wijzen ook op specifieke aspecten die verder verbeterd kunnen worden met vervolgonderzoek.

Ook voor ammoniak komt de gemiddelde ruimtelijke verdeling gesimuleerd door de LOTOS-EUROS en EMEP4NL modellen goed overeen met IASI en CrIS. De modellen vertonen echter een consistente onderschatting van NH_3 boven Nederland ten opzichte van beide satellietproducten voor alle jaren. Deze bias is iets groter voor CrIS in vergelijking met IASI, hoewel de algehele ruimtelijke patronen vergelijkbaar blijven. Het LOTOS-EUROS model volgt de interjaarlijkse variabiliteit zoals waargenomen door zowel IASI als CrIS. Voor IASI vertoont de relatieve afwijking echter een seizoensgang. De ruimtelijke verdeling van NH_3 in het EMEP4NL model komt goed overeen met de satellietmetingen. Bovendien is de seizoensgang vergelijkbaar.

De vergelijking van de vier onafhankelijke emissieschattingen op basis van satellietmetingen (Figuur S3) toont aan dat de geteste methodieken realistische resultaten opleveren op een jaarlijkse, landelijke schaal, zowel voor NH_3 als NO_x. De resulterende emissietotalen liggen dicht bij de gerapporteerde emissies (emissieregistratie.nl) en vallen grotendeels binnen de onzekerheidsgrenzen van de satellietgegevens, inversiemethoden en inventarisatie. Een meer diepgaande vergelijking van de gedetailleerde verdeling in ruimte en de tijdsafhankelijkheid van de emissies is het onderwerp van een vervolgstudie.

Het onderzoek naar het gebruik van satellietwaarnemingen binnen NKS zal door de twee promovendi tot 2027 worden voortgezet. De onderzoeksonderwerpen voor de komende periode staan vermeld in hoofdstuk 7 van de conclusies.

De kwaliteit en dekking van satellietproducten zullen naar verwachting alleen maar verder verbeteren. De recent gelanceerde geostationaire satellietinstrumenten (Copernicus Sentinel 4 en Meteosat Third Generation Infrared Sounder IRS) en de volgende generatie IASI-satellieten zullen de dekking in ruimte en tijd, en de kwaliteit van de metingen verder verbeteren. Over enkele jaren worden deze meetreeksen nog verder uitgebreid met de waarnemingen van de CO2M en TANGO satellietmissies.

Concluderend kunnen we stellen dat de nauwkeurigheid van de huidige satellietwaarnemingen voldoende is om belangrijke nieuwe informatie toe te voegen aan de meetreeksen van de bestaande in-situ netwerken. De combinatie van grond- en satellietgegevens zou de nieuwe standaard kunnen worden voor modelevaluatie. De meetnetwerken blijven essentieel vanwege hun kwaliteit

voor het meten van lokale concentraties aan het oppervlak, terwijl satellietinstrumenten de gaten opvullen tussen de meetinstrumenten en extra informatie verschaffen over de concentraties boven de grond. Opgemerkt moet worden dat satellietgegevens van de atmosferische concentraties alleen direct kunnen worden benut met een model dat driedimensionale uurlijkse concentratieverdelingen kan produceren. Modellen zoals OPS kunnen indirect profiteren, bijvoorbeeld door verbeteringen in de emissies en de gemodelleerde levensduur. Op deze manier kan het operationele stikstofmonitoringsysteem worden verbeterd. Nieuwe satellietmissies en satellietproducten die in de komende jaren beschikbaar komen zullen de detectie, kwantificering en monitoring van NH_3 en NOx bronnen verder verbeteren.

Summary

Monitoring the two main players in the nitrogen problem, the gases ammonia (NH_3) and nitrogen dioxide (NO_2), is of great importance to map the consequences for nature, humans and the climate. In the Netherlands, this monitoring is currently based on (1) instruments that measure the concentrations and deposition on the ground at a number of locations, (2) estimates of emissions based on reported emission figures, in combination with (3) models to calculate the amount of deposition. Currently, no operational use is made of satellite measurements.

This report presents the results of a study on the added value of satellites for the monitoring of reactive nitrogen in the Netherlands. This study is part of the National Nitrogen Knowledge Programme (NKS), funded by the Ministry of Agriculture, Fisheries, Food Security and Nature (LVVN). The NKS programme has the aim to further develop and improve the observation network and modelling of nitrogen in the Netherlands. This programme was initiated partly in response to the recommendations in the report of the Hordijk committee ("Meer meten, robuuster rekenen", published 15 June 2020).

The gases NO_2 and NH_3 are both measured with satellite instruments. Figure S1 shows the distribution of these two gases over the Netherlands as observed from space.

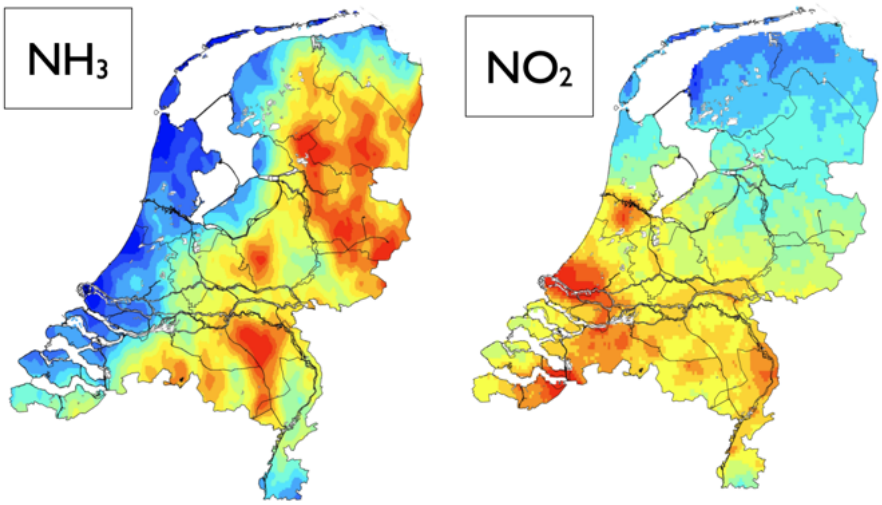


Figure S1. Annually averaged ammonia (left) and nitrogen dioxide (right) in the Netherlands, as observed by the CrIS and TROPOMI satellite instruments. Ammonia is mainly found in the east of the country, with agriculture as the main source. Nitrogen dioxide concentrations are high in the west of the country, linked to contributions from road traffic, residential, industry, harbours, shipping and aviation.

In order to quantify the deposition of reactive nitrogen a detailed knowledge of the emissions, mixing and transport, chemical reactions in the atmophere, and deposition processes is needed. This is illustrated in figure S2. Note that the transport of reactive nitrogen happens largely in elevated layers covering roughly the first kilometer in altitude in the atmosphere. Satellites measure the total column amount of NO_2 and NH_3 (grey shading). This has several advantages: a) The satellite total columns can be converted into observation-based emission estimates to check the reported emissions; b) The satellite observes the plumes of N pollution as they travel from the source to the deposition areas, providing constraints on the atmospheric transport and deposition; c) The satellites measure everywhere, and fill information gaps in between the monitoring stations. Satellite observations, however, contain limited information on the concentrations near the surface.

Therefore satellite observations complement the observations at the surface and can not replace the surface networks. The added value of satellite measurements is described in more detail in the NKS whitepaper "Het gebruik van satellietmetingen voor het monitoren van stikstof in Nederland" (in Dutch).

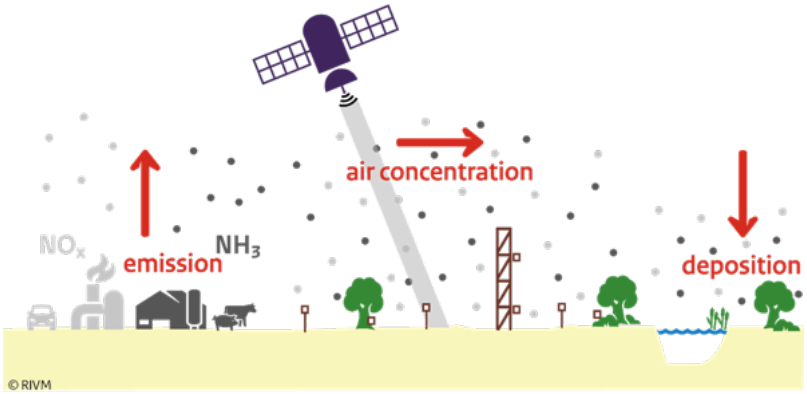


Figure S2: Schematic overview of nitrogen flows in the atmosphere. Emissions of NOx and NH_3 into the air are spread out by mixing and transport in the atmosphere, are converted by chemical reactions and are finally deposited to the surface. The LML, MAN and COTAG networks operated by RIVM provide observations of concentrations and deposition near the surface. The satellite provides extra information on concentrations above the surface, up to about 1 km, and fills gaps in between the measurement locations.

This report provides a detailed overview of the NKS satellite research performed in the years 2023-2024. Use is made of NH_3 observations from the CrIS and IASI instruments, and NO_2 observations from the TROPOMI instrument. The work was carried out by scientists from the institutes TNO, KNMI, RIVM, and the universities of Leiden and Wageningen. Two PhD students have been appointed within NKS for this research. This research consisted of the following studies:

- Performing multi-year benchmark simulations with the three models OPS, EMEP and LOTOS-EUROS, and evaluating these runs with LML, MAN and COTAG ground measurements (chapter 2).
- Evaluating the satellite measurements based on ground-based columns of NH_3 and NO_2 measured with FTIR and MAXDOAS/PANDORA instruments in the Netherlands and world-wide (Chapter 3).
- Evaluating the simulations of LOTOS-EUROS and EMEP with satellite measurements of NO_2 by TROPOMI (Chapter 4) and NH_3 by CrIS and IASI (Chapter 5).
- Estimating the emissions of NH_3 and NO_x based on these satellite measurements with multiple methods, introduced in Section 6.2. The results for the four methods are discussed in Sections 6.3 to 6.7.

The chapters in this report contain descriptions of the models, measurements and methods used, including relevant references to the scientific literature.

The evaluation of the TROPOMI, CrIS and IASI retrievals with ground-based FTIR, MAXDOAS and Pandora remote sensing observations in general show high temporal correlations. Systematic errors are generally within the 25% range. This demonstrates that the satellite retrievals are of good enough quality to bring important new information on reactive nitrogen.

The comparison of LOTOS-EUROS and EMEP4NL simulated NO_2 columns with TROPOMI data overall demonstrated strong spatial and temporal agreement. In both models a bias was observed between the satellite and modelled fields, in the order of 10-25% depending on location and model. A day-to-day analysis of the EMEP4NL model showed that, while the model effectively captures

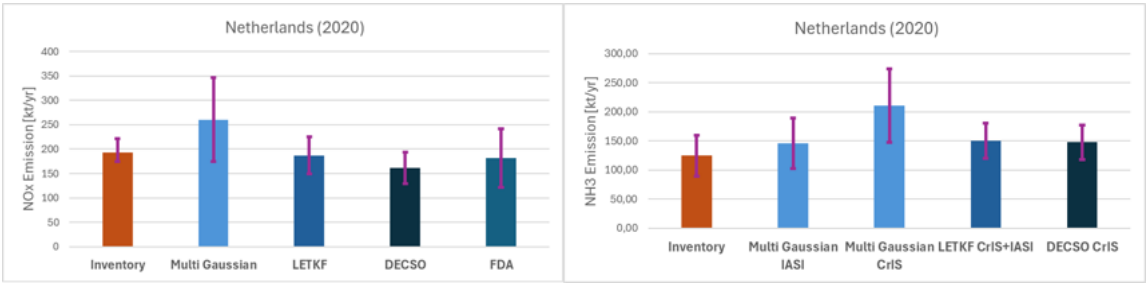


Figure S3: Comparison of NO_x and NH_3 emissions estimated from satellite observations and reported emission, labelled as "inventory". The figure shows yearly-mean total emissions for the four different methodologies to derive emissions, with the corresponding uncertainty ranges, for nitrogen oxide (left) and ammonia emissions (right).

pollution plumes, it occasionally simulates plumes that are longer and slightly shifted in direction compared to TROPOMI, suggesting areas for further investigation.

LOTOS-EUROS and EMEP4NL broadly captured the spatial distribution of NH_3 observed by IASI and CrIS over the Netherlands. However, the models show a consistent low bias relative to both satellite products for all years. The bias is slightly larger for CrIS compared to IASI, although the overall spatial patterns remain similar. LOTOS-EUROS captured the overall inter-annual variability observed by both IASI and CrIS. However, for IASI, the relative difference with the model displays a seasonality. The EMEP4NL model captures the spatial distribution of the satellite retrieved NH_3 total columns over the Netherlands well. Moreover, the modelled and retrieved NH_3 total columns show a similar interannual variability.

The comparison of the various emission estimates (Figure S3) showed that all four tested methodologies produce realistic results at a yearly country-wide scale, both for NH_3 and NO_x. The resulting emission totals are close to those reported by the reported emissions (emissieregistratie.nl), and mostly within the uncertainty limits of the satellite data, inversions and inventory. A more in-depth comparison of the detailed spatio-temporal distribution is the topic of a follow-up study.

The research on the use of satellite observations within NKS will continue through the work of the two PhD students until 2027. Research topics for the coming period are listed in the conclusions, chapter 7.

The quality and coverage of satellite products is expected to only improve further. The recently launched geostationary satellite instruments (Copernicus Sentinel 4 and Meteosat Third Generation Infrared Sounder IRS) as well as the next generation of IASI satellites will further enhance the spatio-temporal coverage and quality of the datasets. In a few years time these capabilities are further extended with the CO2M and TANGO satellite missions.

To conclude, the overal information content and accuracy of the satellite observations can help fill the current gap in the spatio-temporal coverage of the existing in-situ measurement networks. The combination of ground-based and satellite data could become the new standard for model evaluation. Ground-based instruments will remain unsurpassed in their quality for measuring local surface concentrations, whereas the satellite instruments also provide information on the above-surface concentrations, which is not provided by the ground-based networks. It should be noted that these benefits of atmospheric satellite data can only be directly exploited with a model capable of producing three-dimensional hourly concentration distributions. Models like OPS can indirectly benefit, for example from improvements to the emissions and modelled lifetime. In this way, the operational nitrogen monitoring system can be improved. The increased number of satellite instruments and products in the coming years will give insight into the routes for better detection, quantification, and monitoring of NH_3 and NO_2 sources and integration of such observations into the emission inventories.

Contents

Samenvatting 2
1 General introduction 10
2 Long-term simulation 2013-2024 using LOTOS-EUROS, EMEP and OPS 12
2.1 Introduction 12
2.2 Models 12
2.3 Datasets 14
2.4 Model-measurement comparisons 18
3 Validation of satellite data with surface remote sensing observations and models 37
3.1 Introduction 37
3.2 Datasets 37
3.3 Satellite validation results 44
4 Satellite-model inter-comparisons for NO_2 55
4.1 Introduction 55
4.2 Comparison of LOTOS-EUROS and TROPOMI NO_2 55
4.3 Comparison of EMEP with TROPOMI NO_2 56
5 Satellite-model inter-comparisons for NH_3 62
5.1 Introduction 62
5.2 Comparison of LOTOS-EUROS with CrIS and IASI 62
5.3 Comparison of EMEP4NL with CrIS 64
6 Emission estimates using four approaches for NOx and NH3 69
6.1 Introduction 69
6.2 Inverse methods for emission estimates 69
6.3 NH_3 and NO_x emission estimates derived using the Gaussian plume method (MSPM) 72
6.4 Accuracy of the flux-divergence approach to estimate emissions 80
6.5 NO_x emissions derived using DECSO and TROPOMI observations 83
6.6 NH3 emissions derived using DECSO and CrIS observations 87
6.7 Optimized NH_3 and NO_2 Distributions and Emissions using the LOTOS-EUROS LETKF 95
7 Overview of results and conclusions 108
8 List of NKS publications and presentations 115
8.1 NKS publications 115
8.2 NKS presentations 115
8.3 Other 116
References 117
Abbreviations 126

1 General introduction

Nitrogen (N_2) is a colorless and odorless gas. It is all around us. Nitrogen itself is not harmful to humans or the environment. However, when people talk about nitrogen, they often mean the chemical compounds of nitrogen in the air, which can be harmful to humans and the environment: nitrogen oxides and ammonia. Nitrogen oxides (NO_x, a compound of nitrogen and oxygen) are mainly released into the air through exhaust gases from traffic and emissions from industry. NO_x is the sum of nitrogen monoxide (NO) and nitrogen dioxide (NO_2). Ammonia (NH_3, a compound of nitrogen and hydrogen) mainly comes from animals in livestock farming. Manure from animals evaporates as ammonia and thus ends up in the air. This is released in stables and during the fertilization (also by artificial fertilizer) of pastures and fields. A small amount of ammonia comes from other sources such as industry, construction, and traffic. Once in the air, the gases can be transported over long distances, and along the way, the gases can react with each other and form particles (aerosols). Eventually, the nitrogen compounds will settle back on the surface (deposition). Figure 1 schematically shows how different nitrogen components move through the atmosphere.

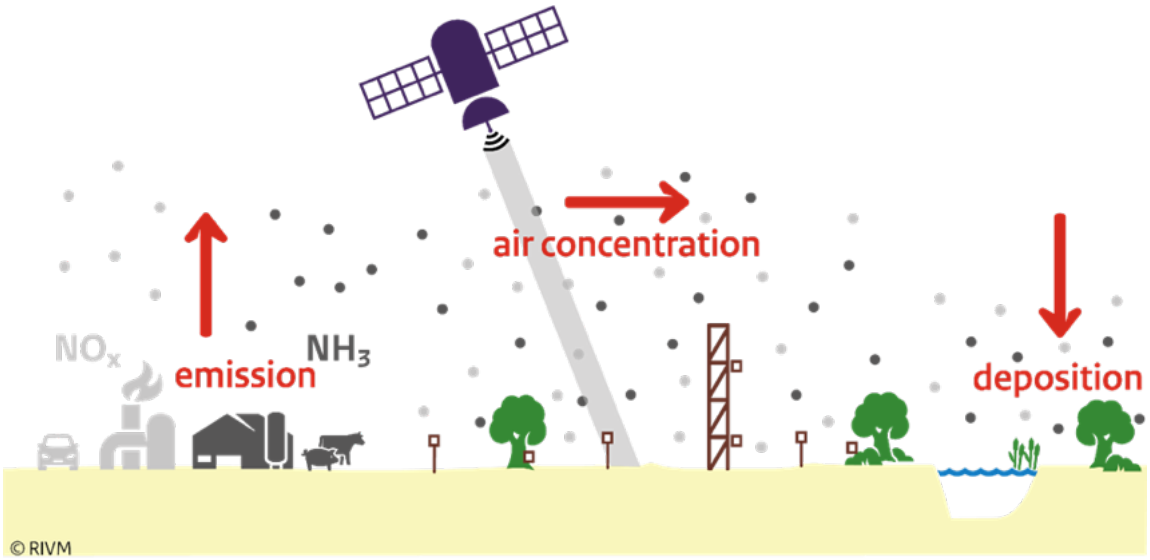


Figure 1: Schematic overview of nitrogen flows in the atmosphere.

In the Netherlands, the emission of nitrogen compounds is compiled by the Emission Registration (emissieregistratie.nl) based on diverse inputs such as reported emissions from large industries, fuel use, number of vehicles and emission per vehicle, heating of houses, agricultural practices, lifestock data. The amount of nitrogen oxides and ammonia in the air (concentration) is measured by RIVM at many different locations in the country through several monitoring networks. In addition, at a limited number of locations, measurements of nitrogen deposition are carried out. Therefore, models are used to calculate a detailed picture of nitrogen deposition. These models can also be used to make projections for the future.

As illustrated in the figure, the RIVM networks measure concentrations and deposition at the surface. However, the transport of reactive nitrogen happens largely in elevated layers up to typically the first kilometer in altitude in the atmosphere. Satellites measure the total amount of NO_2 and NH_3. This has several advantages: a) The satellite total columns can be converted into observation-based emission estimates to check the reported emissions; b) The satellite observes the plumes of N pollution from the source to the deposition areas, providing constraints on the

atmospheric transport and deposition; c) The satellites measure everywhere, and fill information gaps in between the monitoring stations.

In the NKS satellite work package, it was investigated how satellite measurements can contribute to better mapping of nitrogen emissions, atmospheric mixing and flow, and deposition in the Netherlands. The first phase of the activity (2 years) was applying existing techniques and provided an overview of the main applications of satellite observations: a) validation of models, b) trends and diurnal and seasonal variation in concentrations, and c) estimation of emissions using multiple techniques. In addition, a few assimilation platforms were developed in which satellite observations are combined with model calculations. The uncertainties in the satellite observations were studied based on specialised observations measuring the columns from the surface. It is indicated where further improvements can be made for the use of satellites and what future developments are expected.

2 Long-term simulation 2013-2024 using LOTOS-EUROS, EMEP and OPS

2.1 Introduction

The three main models used for Nitrogen modelling, for air-quality forecasting and for policy support in the Netherlands are:

- The OPS model, used by RIVM for official reporting on reactive nitrogen concentrations and deposition, as well as on air quality.
- The Dutch LOTOS-EUROS model, mainly developed by TNO.
- The EMEP4NL model mainly used at RIVM, which is a version of the EMEP model developed in Norway, adapted to the Netherlands.

Both LOTOS-EUROS and EMEP are also used in a European context for daily air quality forecasting and for policy support in the Copernicus CAMS atmosphere service. Long-term runs have been performed with the three models to evaluate them against the long-term observation records available from the satellites, and to study changes in emissions and inter-annual variations caused by changes in the weather.

The models are introduced in section 2.2, the surface measurements in the Netherlands in 2.3 and the comparison of the multi-year model runs with the measured ground concentrations are discussed in section 2.4.

2.2 Models

2.2.1 LOTOS-EUROS model

The LOTOS-EUROS (LOng Term Ozone Simulation—EURopean Operational Smog) chemical transport model is a three-dimensional Eulerian model designed for regional air quality assessments and operational forecasting in Europe (Manders et al., 2017). It allows for the simulation of the dispersion, chemical transformation, and deposition of atmospheric pollutants, including gases and aerosols. The model incorporates detailed representations of atmospheric processes such as advection, diffusion, chemical reactions and deposition. The model is commonly used for operational forecast in the Netherlands and Europe (Manders et al., 2017). It has been used in research applications for regions around the globe, such as China (Petersen et al., 2019; Timmermans et al., 2017) and South America (Yarce Botero et al., 2021).

LOTOS-EUROS is part of the Copernicus Atmospheric Monitoring Service (CAMS) European air quality ensemble. This service provides forecasts for the main air pollutants using an ensemble of state-of-the-science CTMs. Within CAMS, LOTOS-EUROS is regularly validated against in-situ observations and TROPOMI satellite data, as well as evaluation against the other ensemble members (Colette et al., 2024; Peuch et al., 2022). LOTOS-EUROS also has participated in numerous model inter-comparisons, typically showing a strong performance (Bessagnet et al., 2016; Colette et al., 2017; Vivanco et al., 2017).

Here we will give a short summary of the most important model parameters used. For a detailed description on the inner workings of the model we point the reader to Manders et al. (2017). The setup of LOTOS-EUROS (v2.2.009) used here follows a nesting procedure with 3 different

domains. The outside domain covers most of Europe (15°W-35°E; 35-70°N) with a resolution of about 25×25 km^2, followed by a second domain covering most of North-western Europe (2-16°E; 47-56°N) with a resolution of 7×7 km^2. The output of the second domain is then used as a boundary conditions for the final target domain covering the Netherlands (3.1-7.5°E; 50.3-53.7°N) with a horizontal resolution of 2×2 km^2. The model is driven by meteorological fields obtained from the ECMWF short-term forecast model at a 3-hourly temporal resolution which is then interpolated to an hourly frequency within the model. The simulations were conducted using 12 vertical levels, extending from the ground to about 10 km above the earth's surface, matching a combination of the ECMWF meteorology datasets layering.

The emissions used are a combination of a European scale emission inventory, CAMS-REG-v5.1, where the Dutch and German emissions are replaced by the national GrETa and ER emission inventories, produced in NKS-WP2. After 2019 the emission totals for 2019 are used. Temporally the inventory emissions, are distributed using hourly time factors for different aggregated source categories. In case of ammonia emissions from agricultural sources, emissions are described by a meteorologically dependent parameterization, that accounts for shifts in fertilizer application due to weather conditions. Additionally emissions are vertically distributed with specific heights assigned on a sector-by-sector basis. This particularly matters for industrial sources and public power stations which have vertical distributions based on average stack heights.

The model output for all major nitrogen species and total particulate matter (2.5 and 10 microns) consists of simulated concentrations (surface, and all vertical layers) and wet and dry deposition, as well as concentrations matched at the footprints of all satellite products (I.e. IASI, CrIS, TROPOMI). To match the model with the satellite footprints the CAMS Satellite Operator (CSO) is used. CSO (https://ci.tno.nl/gitlab/cams/cso) is an open-access tool developed at TNO and implemented to facilitate fast intercomparisons between modelled and satellite concentrations. The tool consists of two entities: a pre-processor to download, select, and convert satellite observations into a common format, accompanied with a post-processing tools to aggregate and visualize the data; and a source code that can be used within regional air quality modelling and assimilation systems such as LOTOS-EUROS. The CSO module is able to read the files created by the pre-processor, and simulate the observations using model variables, applying observational operators where applicable.

2.2.2 The EMEP4NL model

The EMEP MSC-W model (Simpson et al., 2012) is an Eulerian atmospheric chemistry transport model. In this study, its implementation for the Netherlands, i.e., the EMEP4NL configuration (van der Swaluw et al., 2021, Verweij et al., 2023), based on EMEP MSC-W model version 4.45 is used. EMEP4NL uses one-way nested grids scaling up from a coarse European domain ($0.5° \times 0.5°$) to a high resolution inner domain that covers the Netherlands, the western part of Germany and Flanders on a spatial resolution of 1-2 kilometers. The spatial resolution is increased by a factor 3 in each nest. The EMEP4NL model is driven by meteorological output from the Weather Research Forecast (WRF) model (Skamarock et al., 2008). The vertical layers in EMEP4NL are inherited from the WRF output (which is defined on fixed pressure levels). EMEP4NL computes on 21 layers with pressure coordinates, ranging from 1000 hPa (corresponding to a height of approximately 50m) at the bottom layer to 100 hPa (corresponding to approximately 15km). The official EMEP emissions are used as emissions for the coarse domains. The two innermost domains use high resolution emission data for the Netherlands from the Emission Registration (ER). These emissions are the same ones as used to obtain the large-scale concentration maps of the Netherlands (GCN-round), see for example Hoogerbrugge et al., 2021.

In this project, the EMEP4NL model is run to calculate surface concentrations as well as 3-dimensional concentration fields of NH_3, NO_2 and NO, for instance for comparison with satellite products, between 2013-2023. The emissions used in EMEP4NL come from the GCN-round of 2024. These are consistent with the emissions from the GCN round of 2023, a detailed description

is given in Hoogerbrugge et al., 2023. For the last year of the time series, 2023, emissions from 2022 were used. More details about the emission totals and trends over the Netherlands can be found in Smeets et al., 2023.

2.2.3 The OPS-LT model

The OPS (Operational Priority Substances) model is a source-receptor model that combines Gaussian plume modelling for local applications with a Lagrangian trajectory modelling for long range transport (Sauter et al., 2018). The main purpose of OPS is to calculate air pollution concentrations and deposition (e.g. for NO_x and NH_3) over the Netherlands on a high spatial resolution. OPS is driven by meteorological data provided by the Royal Netherlands Meteorological Institute (KNMI). OPT-LT (Long Term) is the long-term version of OPS that is used to calculate yearly concentration and deposition values. OPS-LT is a statistical model in the sense that its calculations are done for several typical meteorological classes and the long term values are obtained by adding up and weighting these classes with their relative occurrence frequencies. In this project, the same input emissions for the Netherlands as EMEP4NL are used in OPS-LT (i.e., the emissions from the GCN-round of 2024). The OPS-LT model is used for the production of yearly concentration and deposition maps of the Netherlands (e.g., Hoogerbrugge et al., 2021), and is for example integrated in the AERIUS calculator to provide insights on the contributions of new projects and plans on nitrogen deposition on Natura 2000 nature areas (https://calculator.aerius.nl/).

2.3 Datasets

2.3.1 LML

The Dutch National Air Quality Monitoring Network, known as the "Landelijk Meetnet Luchtkwaliteit" (LML) (Elskamp, 1989) (Elzakker & Buijsman, 1999), is a comprehensive ground-based measurement network designed to monitor air quality across the Netherlands. Operated by Rijksinstituut voor Volksgezondheid en Milleu (RIVM), the LML network consists of a large number of monitoring stations distributed across urban, suburban, and rural areas. These stations continuously collect data on various air pollutants, including particulate matter (PM10 and PM2.5), nitrogen dioxide (NO_2), ozone (O_3), sulfur dioxide (SO_2), carbon monoxide (CO), and volatile organic compounds (VOCs).

The network provides real-time data, which is crucial for assessing the air quality in different regions and understanding the impact of pollution on public health and the environment. LML stations employ state-of-the-art sensors and analytical techniques to ensure high data accuracy and consistency, enabling authorities to monitor trends, detect exceedances of air quality standards, and develop policy interventions when necessary.

Data from the LML are publicly accessible (Luchtmeetnet.nl), allowing citizens, researchers, and policymakers to track air quality levels in near real-time. This transparency helps raise awareness about air pollution issues and supports efforts toward improving air quality across the country. The LML network is also integrated with broader European air quality initiatives, contributing to the wider understanding of transboundary pollution and climate change mitigation efforts.

Comparisons were performed between the LML ground-based data with LOTOS-EUROS for NH_3, NO_2, and PM2.5 over the period of 2014-2023. Six sites in particular were selected for these comparisons; De-Zilk Volgelaarsdreef, Valthermond-Noorderdiep, Vredepeel-Vredeweg, Wekerom-Riemterdijk, Wiereingerwerf-Medeblikkerweg, and Zegveld-Oude Meije. These specific sites were selected because they provided hourly NH_3 measurements over a long period (>1 year). Each of these six sites provide hourly NO_2 measurements, while De-Zilk, Vredepeel, Wekerom, and Wieringerwerf also provide hourly PM2.5 measurements. The locations and details of these sites

Site Name	Coordinates	Species Measured	Type
De-Zilk Vogelaarsdreef	52.30° N, 4.51° E	NH_3, NO_2, PM2.5	Coastal
Valthermond-Noorderdiep	52.88° N, 6.93° E	NH_3, NO_2	Rural/Agricultural
Vredepeel-Vredeweg	51.54° N, 5.85° E	NH_3, NO_2, PM2.5	Rural/Agricultural
Wekerom-Riemterdijk	52.11° N, 5.71° E	NH_3, NO_2, PM2.5	Suburban
Wieringerwerf-Medeblikkerweg	52.80° N, 5.05° E	NH_3, NO_2, PM2.5	Coastal/Agricultural
Zegveld-Oude Meije	52.14° N, 4.84° E	NH_3, NO_2	Agricultural

Table 1: Locations and details of LML sites used for comparisons with LOTOS-EUROS/EMEP4NL.

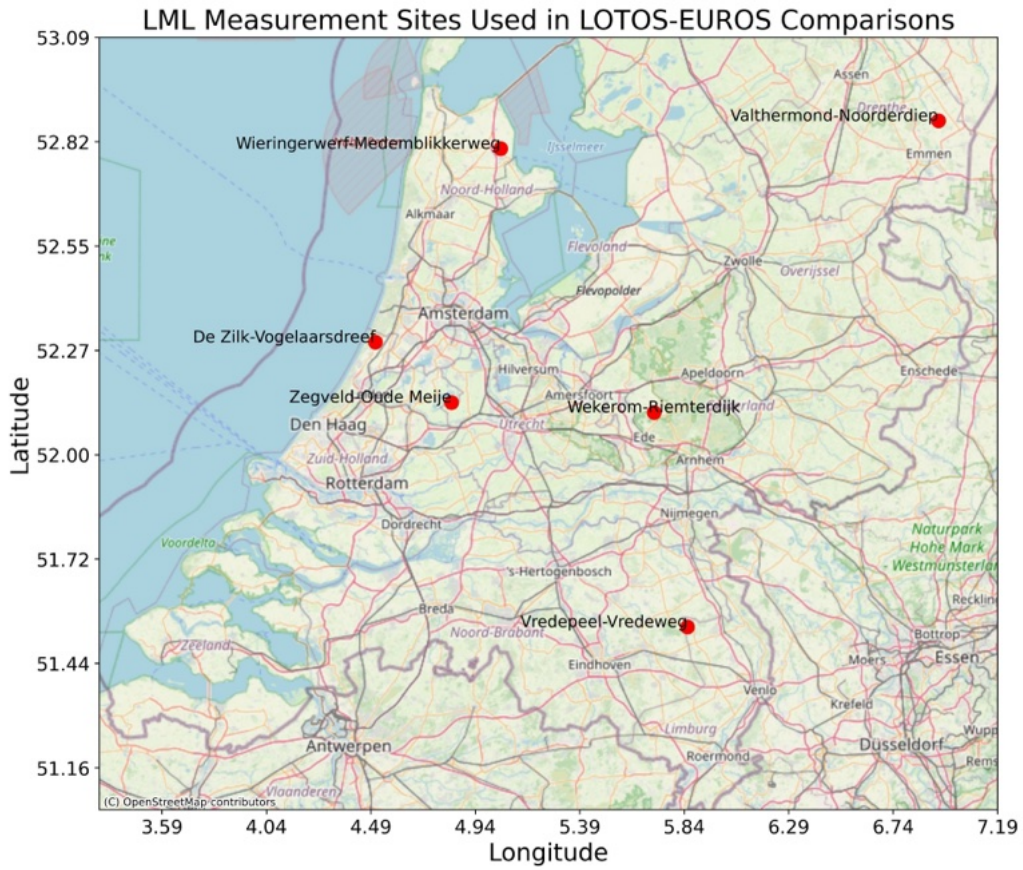


Figure 2: Map of LML sites used for comparisons with LOTOS-EUROS.

are provided in Table 1, and they are also shown on a map in Figure 2.

The LML network has been monitoring ambient NH_3 concentrations using AMOR (Ammonia MOnitoR) instruments until 2016. In 2016, the AMOR instruments were replaced by miniDOAS (Differential Optical Aborption Spectroscopy) instruments. MiniDOAS measurements have an estimated detection limit of 0.25 ug m^{-3} and a precision of 2.25% , with a minimum of 0.25 ug m^{-3}. For hourly averaged concentrations, the precision is estimated to be 0.1 ug m^{-3} (Berkhout et al., 2017).

2.3.1.1 MAN

The "Measuring Ammonia in Nature" (MAN) network was established in 2005 to monitor atmospheric NH_3 concentrations in nature reserve areas in the Netherlands, particularly in Natura2000

areas sensitive to nitrogen deposition (Lolkema et al., 2015). The MAN network provides crucial data for assessing national NH_3 concentration trends, validating air quality models, and analyzing regional variations.

Currently, the MAN network has expanded to include over 300 sampling sites spread throughout the Netherlands. The NH_3 concentration data collected through MAN helps to evaluate the impact of local emission sources, such as agricultural activities, and is used to detect spatial concentration patterns and regional deviations. The data are also essential for validating model predictions, particularly with the OPS (Operational Priority Substances) model (Sauter et al., 2018).

The MAN network's data provides insight into both the effectiveness of environmental policies aimed at reducing nitrogen emissions and the long-term trends in nitrogen deposition. The network has demonstrated the capability to detect trends as low as 3% annually over extended time series, making it a valuable resource for air quality management and biodiversity conservation efforts in nitrogen-sensitive regions (Lolkema et al., 2015).

The network utilizes passive Gradko NH_3 samplers, which are inexpensive, easy to deploy, and suitable for widespread monitoring. Volunteers, often local nature rangers, exchange the samplers monthly, ensuring consistent data collection across a range of habitats. To diminish meteorological influences on the passive samplers, they are calibrated monthly against high-performance reference instruments for NH_3 at 6 locations in the Netherlands (from the LML-network (see Berkhout et al., 2017). At these 6 locations, Gradko measurements are collected in triplicate. To determine the calibration parameters, the ratio between the NH_3 concentration of the reference instruments and the mean of the triplicates is determined for each LML site. A linear regression is performed on the ratios against the triplicate means. Then, the determined calibration parameters (i.e., the linear regression coefficients) are then applied to the all other MAN passive samplers. More details on the calibration procedure can be found in Noordijk et al., 2020.

The uncertainty of the MAN measurements consists of two parts: the measurement uncertainty itself and the uncertainty in the calibration procedure. In the most recent evaluation of the uncertainties in the MAN observations by Noordijk et al., 2020, the total uncertainty of a single monthly observation is given by $\epsilon_{tot} = [0.32^2 + (0.10 \cdot \text{conc})^2]$. For example, for a measurement of 1 ug m^{-3} this equates to a total uncertainty of 0.34 ug m^{-3} (34%), and for a measurement of 5 ug m^{-3} this would correspond to an error of 0.60 ug m^{-3} (12%) (Noordijk et al., 2020).

For comparisons with the LOTOS-EUROS long simulations, all available data was used over the period of 2014-2023, which amounts to a total of 329 sites spread across the country. It should be noted that not many of these sites provide an uninterrupted time-series over the full 2014-2023 period.

2.3.2 COTAG

The Conditional Time-Averaged Gradient (COTAG) network is a specialized system operated by RIVM in the Netherlands for measuring the dry deposition fluxes of various atmospheric pollutants to the Earth's surface (Rutledge-Jonker et al., 2023). The COTAG network utilizes a combination of conditional sampling and the time-averaged gradient technique to achieve accurate measurements. By measuring pollutant concentration differences between two vertical levels above the surface and averaging these over time, the network minimizes short-term atmospheric fluctuations. The "conditional" aspect refers to selectively sampling under specific atmospheric conditions, like certain wind speeds and directions, to ensure that the data accurately represent deposition processes.

Instrumentation within the network includes air inlet systems positioned at different heights, pollutant analyzers, and meteorological sensors that record wind speed, direction, temperature, and humidity. Sites are strategically selected to represent broader regional conditions and are often located near areas of environmental significance, such as industrial zones, forests, or agricultural

lands. Data from the COTAG network is used to support environmental impact assessments, enhance atmospheric models, and inform regulatory policies by providing high-resolution temporal and spatial information on pollutant deposition.

The COTAG data including estimated uncertainties can be downloaded at https://www.rivm.nl/stikstof/meten/drogedepositieNH3. Here, the reported lower and upper ranges represent $\pm 2\sigma$, where σ is the total uncertainty (random and systematic) calculated following the methodology described in Rutledge-Jonker et al., 2023.

2.4 Model-measurement comparisons

2.4.1 Comparisons LOTOS-EUROS and LML

To compare the LOTOS-EUROS simulated concentrations of NH_3, NO_2 and PM2.5, data was extracted at the location of each LML site. The correlations (in terms of the Pearson correlation coefficient R), and mean biases (μ) were evaluated on a site-by-site basis, as well as with spatial and temporal averaging across all sites. The long-term trends were also evaluated in both the model and measurements using a Fourier-series fitting approach. In the following subsections the results of these comparisons are summarized.

2.4.1.1 **NH_3**

An example correlation plot of the measured versus simulated NH_3 daily mean surface concentrations (in units of ug m^{-3}) at the Wekerom-Riemterdijk site is shown in Figure 3, as well as a time-series of the monthly mean NH_3 surface concentrations in Figure 4. The results for the daily mean NH_3 comparisons for all sites are summarized below in Table 2.

The agreement between the measured and modeled NH_3 surface concentrations varies between sites, with correlations ranging between R = 0.61 (De-Zilk Vogelaarsdreef and Wekerom-Riemterdijk) to R = 0.27 (Wieringerwerf-Medeblikkerweg), and with mean biases ranging from 0.1 ug m^{-3} (Valthermond-Noorderdiep) to -3.6 ug m^{-3} (Vredepeel-Vredeweg). The annual and inter-annual variability in the observations is generally well captured in the model, particularly for the site at Wekerom-Riemterdijk (see Figure 4), but for some sites (e.g., Valthermond-Noorderdiep) larger differences were found during the springtime relative to the rest of the year. In addition to the daily mean comparisons, the correlation in the monthly temporal means across all sites were evaluated and is shown for NH_3 in Figure 5. The temporal means show a relatively strong correlation (R = 0.80), and a mean bias of -1.5 ug m^{-3} across all sites. The data shown in Figure 5 displays a moderate degree of spread, with a standard deviation of the differences of 4.2 ug m^{-3}.

The long-term NH_3 trends were derived from the LML measurements and the LOTOS-EUROS simulated concentrations by applying a Fourier-series fitting approach (of order 3) to the time series. The 95% confidence intervals, a measure of the robustness of the fitted trends, were derived using a bootstrap sampling approach (with 5000 samples). Here, we consider a trend to be statistically significant if the 95% confidence intervals do not overlap with zero. An example trend fit for the measured and modeled surface NH_3 data at Wekerom-Riemterdijk is shown in Figure 6. The fitted NH_3 trends for all sites are summarized in Table 3. Statistically significant trends in the observed NH_3 surface concentrations were found at Vredepeel-Vredeweg (-0.43 ug m^{-3} yr^{-1}), Wekerom-Riemterdijk (-0.44 ug m^{-3} yr^{-1}), and Zegveld-Oude Meije (-0.18 ug m^{-3} yr^{-1}), while the only statistically significant trend in LOTOS-EUROS was found at Vredepeel-Vredeweg (-0.31 ug m^{-3} yr^{-1}). All of the statistically significant trends indicate decreasing NH_3 surface concentrations over time.

2.4.1.2 **NO_2**

A similar analysis as for NH_3 was performed for NO_2 at the same six LML sites. The comparisons of the daily mean NO_2 surface concentrations are summarized in Table 4. The measured and modeled daily mean NO_2 surface concentrations are generally very well correlated with Pearson correlation coefficients ranging from R = 0.82 (De-Zilk Vogelaarsdreef) to R = 0.77 (Valthermond-Noorderdiep), and mean biases ranging from 4.7 ug m^{-3} (Zegveld-Oude Meije) to -1.2 ug m^{-3} (Wieringerwerf-Medeblikkerweg). An example time-series of monthly mean observed and LOTOS-EUROS simulated NO_2 surface concentrations at the Wekerom-Riemterdijk LML site is shown in Figure 7. The inter- and intra-annual variabilities in the monthly means that are observed in the LML data at this site are well modeled by LOTOS-EUROS. A mean difference between the

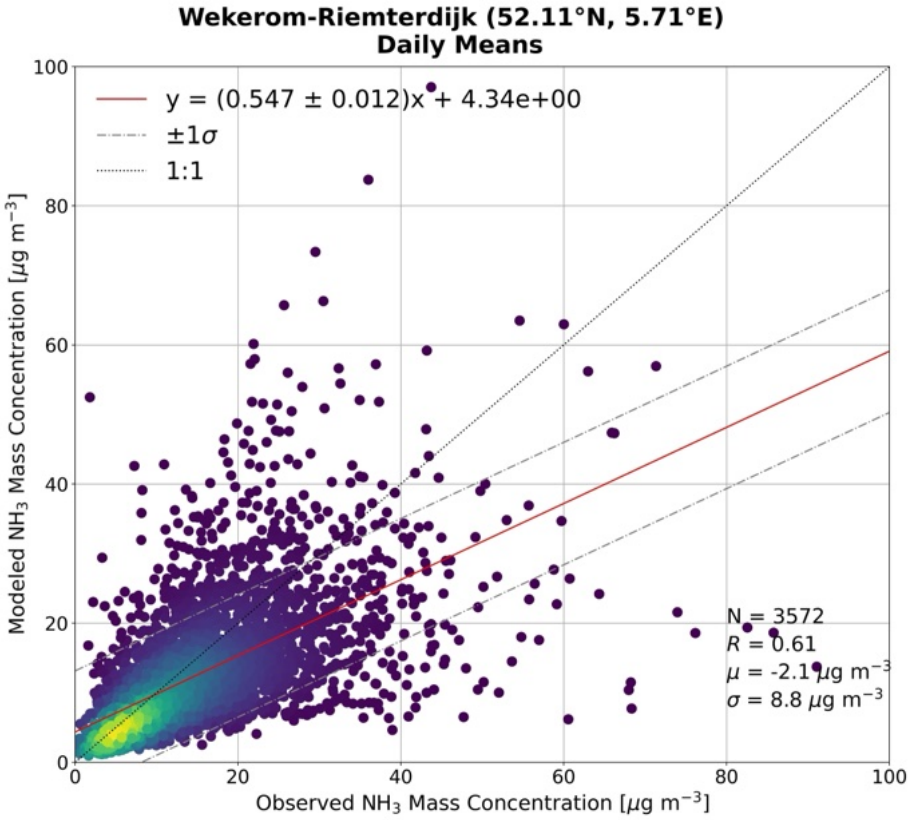


Figure 3: Correlation plot of daily mean NH_3 surface measurements versus simulated LOTOS-EUROS concentrations at the Wekerom-Riemterdijk LML site.

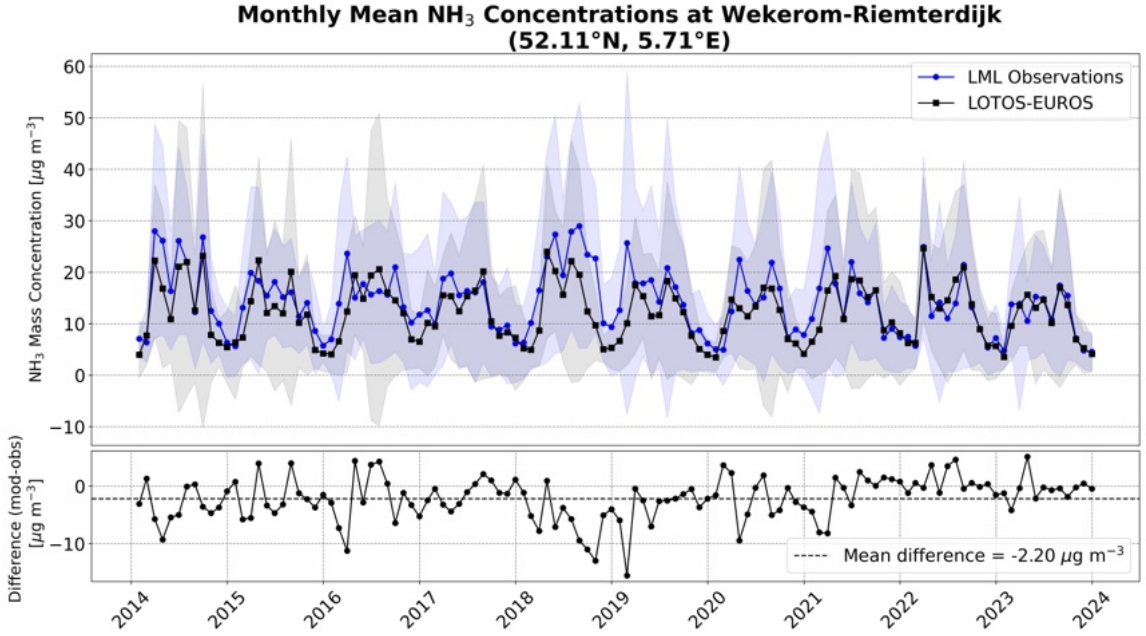


Figure 4: Monthly mean time-series of surface measurements and LOTOS-EUROS simulated NH_3 at the Wekerom-Riemterdijk LML site.

Site Name	N_{days}	Correlation (R)	Mean bias (ug m^{-3})	Slope	Intercept
De-Zilk Vogelaarsdreef	3548	0.61	-0.1	0.81	0.42
Valthermond-Noorderdiep	3109	0.49	0.1	0.83	1.13
Vredepeel-Vredeweg	3541	0.28	-3.6	0.18	1.20
Wekerom-Riemterdijk	3572	0.61	-2.1	0.55	4.34
Wieringerwerf-Medeblikkerweg	3293	0.27	-2.8	0.15	2.79
Zegveld-Oude Meije	3384	0.53	-2.3	0.62	1.06

Table 2: Summary of daily mean NH_3 surface concentration comparisons between LML and LOTOS-EUROS for the period of 2014–2023.

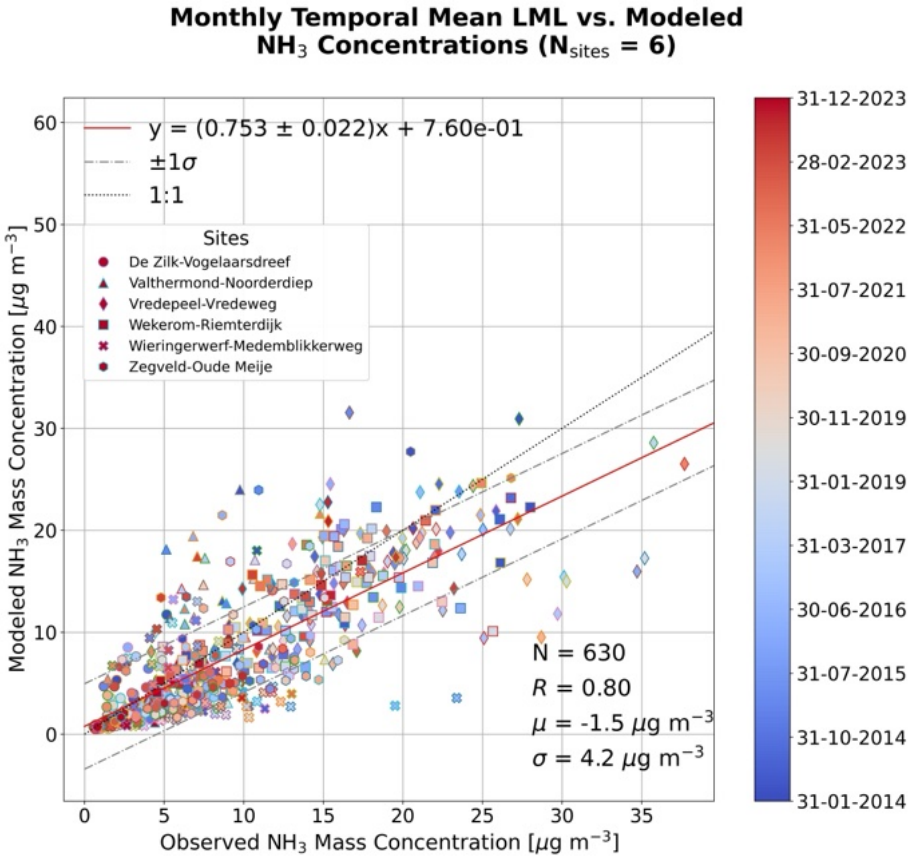


Figure 5: Correlation plot of the monthly temporal NH_3 surface concentration means of all LML sites. Each individual data point represents a monthly mean at a given site, and the data points are colored according to time.

observations and the simulated surface concentrations of 1.80 ug m^{-3} is found at this site over the 2014–2023 time-series. Additionally, a correlation plot of the monthly temporal means in the NO_2 surface concentrations for all sites are shown in Figure 8. Over all LML sites, a Pearson correlation coefficient of R = 0.80 is found, with a relatively small mean bias of 1.1 ug m^{-3}. In comparison to the NH_3 monthly temporal means (Figure 5), the spread in the NO_2 monthly means is smaller and there are fewer notable outliers, with a standard deviation in the measurement-model differences of 3.4 ug m^{-3}.

The trends in the observed and measured NO_2 surface concentrations were evaluated, and are

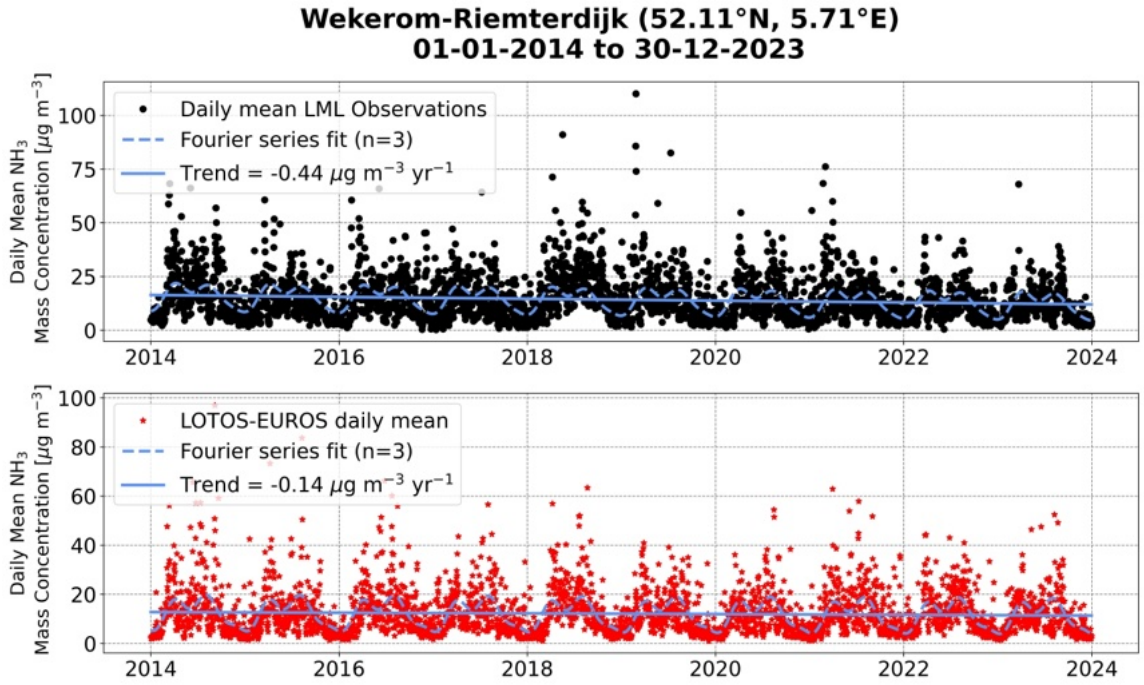


Figure 6: Monthly mean time-series of LML measured and LOTOS-EUROS simulated NH_3 mass concentrations at the Wekerom-Riemterdijk LML site.

Site Name	Trend (ug m^{-3} yr^{-1})		Significant?	
	Observed	Model	Observed	Model
De-Zilk Vogelaarsdreef	+0.04 ± [-0.03, 0.10]	-0.01 ± [-0.07, 0.06]	No	No
Valthermond-Noorderdiep	+0.04 ± [-0.05, 0.12]	-0.06 ± [-0.19, 0.07]	No	No
Vredepeel-Vredeweg	-0.43 ± [-0.73, -0.13]	-0.31 ± [-0.50, -0.12]	Yes	Yes
Wekerom-Riemterdijk	-0.44 ± [-0.63, -0.25]	-0.14 ± [-0.30, 0.02]	Yes	No
Wieringerwerf-Medeblikkerweg	-0.15 ± [-0.29, 0.0]	-0.05 ± [-0.14, 0.03]	No	No
Zegveld-Oude Meije	-0.18 ± [-0.33, -0.04]	-0.13 ± [-0.28, 0.03]	Yes	No

Table 3: Trends derived from the LML observations and LOTOS-EUROS modeled NH_3 surface concentrations. The values provided in the square brackets are the 95% confidence intervals calculated using a bootstrapping approach. A trend is considered significant if the confidence intervals do not overlap with zero.

summarized in Table 5. Statistically significant negative trends were found in both the observations and the simulated surface concentrations at all sites, with largest trend of -0.82 ug m^{-3} yr^{-1} derived from the observations at Zegveld-Oude Meije. From Table 5, it can be noted that the observed trends from the LML measurements are broadly larger in magnitude than those estimated from the LOTOS-EUROS simulated surface concentrations. This difference with the observed trends is expected since the final year of the emissions inventory used in LOTOS-EUROS is 2019, and thus the decrease in NO_2 between 2020–2023 is likely being underestimated in the model.

Site Name	N_{days}	Correlation (R)	Mean bias (ug m^{-3})	Slope	Intercept
De-Zilk Vogelaarsdreef	3615	0.82	1.9	0.95	2.63
Valthermond-Noorderdiep	3233	0.77	0.6	0.87	1.92
Vredepeel-Vredeweg	3611	0.80	0.3	0.91	1.63
Wekerom-Riemterdijk	3625	0.80	1.8	0.92	2.92
Wieringerwerf-Medeblikkerweg	3616	0.80	-1.2	0.74	2.05
Zegveld-Oude Meije	3578	0.80	4.7	0.97	5.12

Table 4: Summary of daily mean NO_2 surface concentration comparisons between LML and LOTOS-EUROS for the period of 2014–2023.

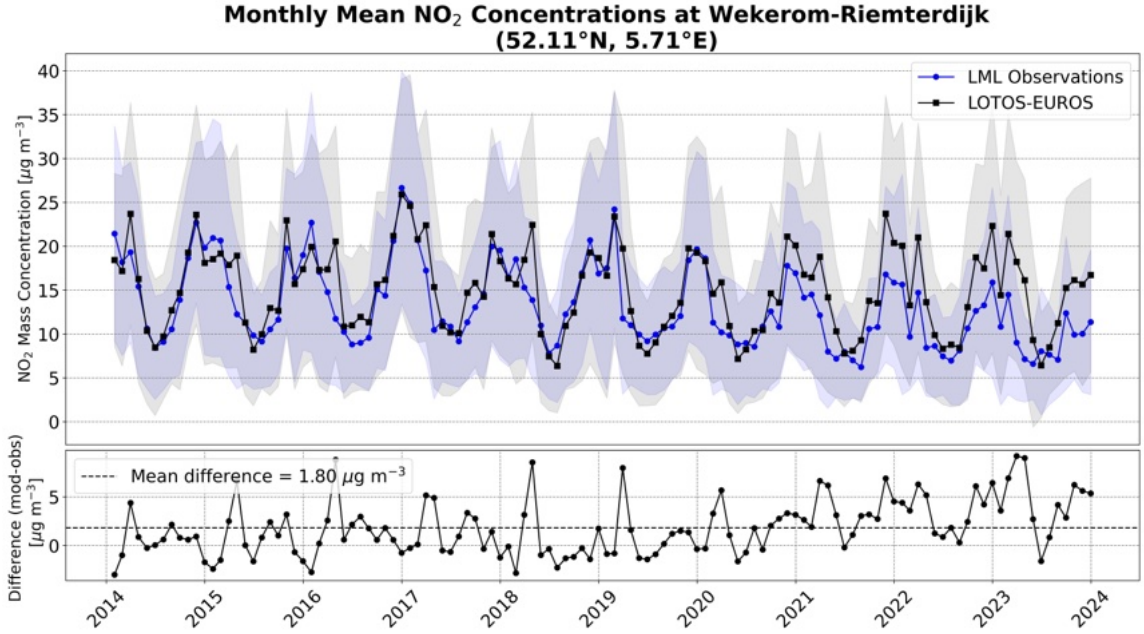


Figure 7: Monthly mean time-series of LML measured and LOTOS-EUROS simulated NO_2 mass concentrations at the Wekerom-Riemterdijk LML site.

2.4.1.3 **PM2.5**

Four LML sites provide hourly PM2.5 measurements; De-Zilk, Vredepeel, Wekerom, and Wieringerwerf. The results of the daily mean comparisons of the LML-measured and LOTOS-EUROS simulated PM2.5 surface concentrations are summarized in Table 6. The correlations between the measurements and the modeled concentrations are strong and range between R = 0.82 (Vredepeel-Vredeweg and Wieringerwerf-Medeblikkerweg) to R = 0.78 (De-Zilk Vogelaarsdreef), with very small mean biases ranging from -0.4 ug m^{-3} (Wieringerwerf-Medeblikkerweg) to 0.0 ug m^{-3} (Vredepeel-Vredeweg and Wekerom-Riemterdijk). An example time-series of the monthly mean measured and LOTOS-EUROS simulated PM2.5 surface concentrations at Wekerom-Riemterdijk is shown in Figure 9. Variations in the differences between the LML measurements and modelled surface concentrations can be observed with respect to time, however the mean bias calculated over the full time-series for this site is 0.0 ug m^{-3}. A correlation plot of the monthly means calculated across all four LML sites is provided in Figure 10. A strong correlation of R = 0.84 is found between the simulated and observed monthly mean PM2.5 mass concentrations, with a small mean bias of -0.2 ug m^{-3}. The measured and modeled PM2.5 surface concentrations sit quite closely along the 1:1 line, and the spread in the data is relatively compact with a standard

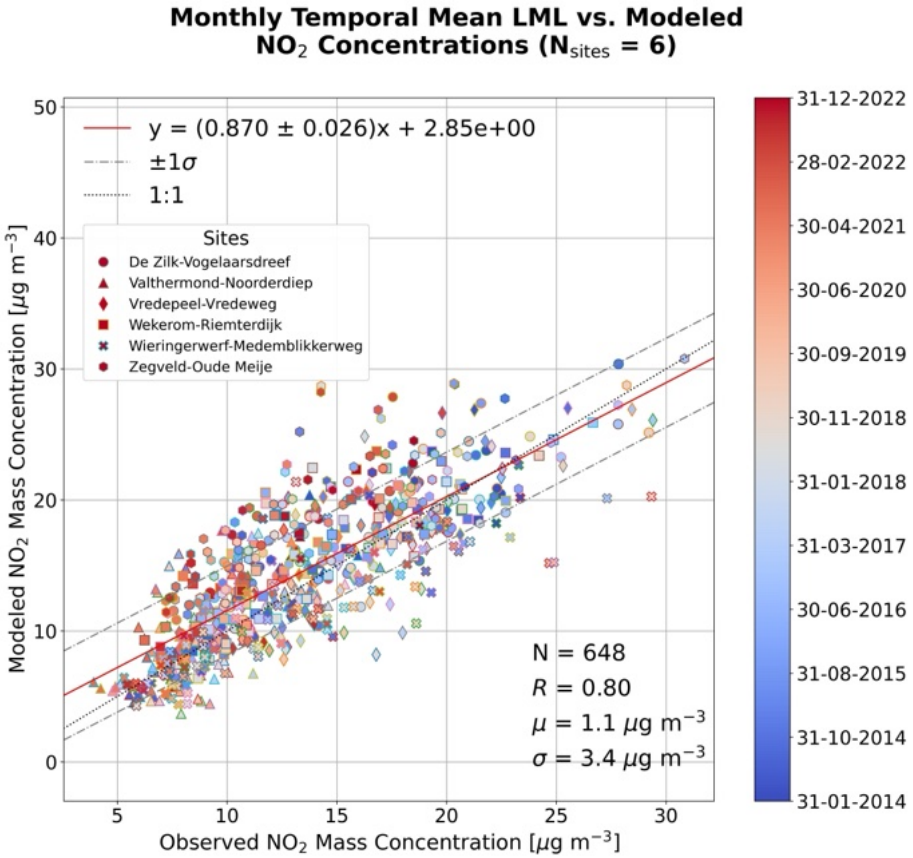


Figure 8: Correlation plot of the monthly temporal NO_2 surface concentration means of all LML sites. Each individual data point represents a monthly mean at a given site, and the data points are colored according to time.

Site Name	Trend (ug m^{-3} yr^{-1})		Significant?	
	Observed	Model	Observed	Model
De-Zilk Vogelaarsdreef	-0.65 ± [-0.80, -0.49]	-0.27 ± [-0.46, -0.08]	Yes	Yes
Valthermond-Noorderdiep	-0.42 ± [-0.54, -0.30]	-0.22 ± [-0.36, -0.09]	Yes	Yes
Vredepeel-Vredeweg	-0.67 ± [-0.82, -0.53]	-0.34 ± [-0.49, -0.19]	Yes	Yes
Wekerom-Riemterdijk	-0.68 ± [-0.80, -0.56]	-0.26 ± [-0.40, -0.14]	Yes	Yes
Wieringerwerf-Medeblikkerweg	-0.41 ± [-0.55, -0.27]	-0.25 ± [-0.39, -0.11]	Yes	Yes
Zegveld-Oude Meije	-0.82 ± [-0.95, -0.69]	-0.36 ± [-0.52, -0.20]	Yes	Yes

Table 5: Trends derived from the LML observations and LOTOS-EUROS modeled $N0_2$ surface concentrations. The values provided in the square brackets are the 95% confidence intervals calculated using a bootstrapping approach. A trend is considered significant if the confidence intervals do not overlap with zero.

deviation of the differences of 2.1 ug m^{-3}.

The trends in the measured and simulated PM2.5 surface concentrations were evaluated and are summarized in Table 7. No statistically significant trends in the PM2.5 surface concentrations were found in the LML measurements at De-Zilk Vogelaarsdreef and Vredepeel-Vredeweg, while slight

Site Name	N_{days}	Correlation (R)	Mean bias (ug m^{-3})	Slope	Intercept
De-Zilk Vogelaarsdreef	3224	0.78	-0.1	0.77	1.95
Vredepeel-Vredeweg	3041	0.82	0.0	0.81	2.12
Wekerom-Riemterdijk	3062	0.80	0.0	0.72	3.02
Wieringerwerf-Medeblikkerweg	3201	0.82	-0.4	0.73	1.98

Table 6: Summary of daily mean PM2.5 surface concentration comparisons between LML and LOTOS-EUROS for the period of 2014–2023.

negative trends of -0.25 ug m^{-3} yr^{-1} and -0.33 ug m^{-3} yr^{-1} are seen in the model at these two sites, respectively. At Wekerom-Riemterdijk and Wieringerwerf-Medeblikkerweg identical negative trends of -0.36 ug m^{-3} yr^{-1} were estimated from the measured PM2.5 surface concentrations, while in the model trends of -0.34 ug m^{-3} yr^{-1} and -0.26 ug m^{-3} yr^{-1} were found, respectively.

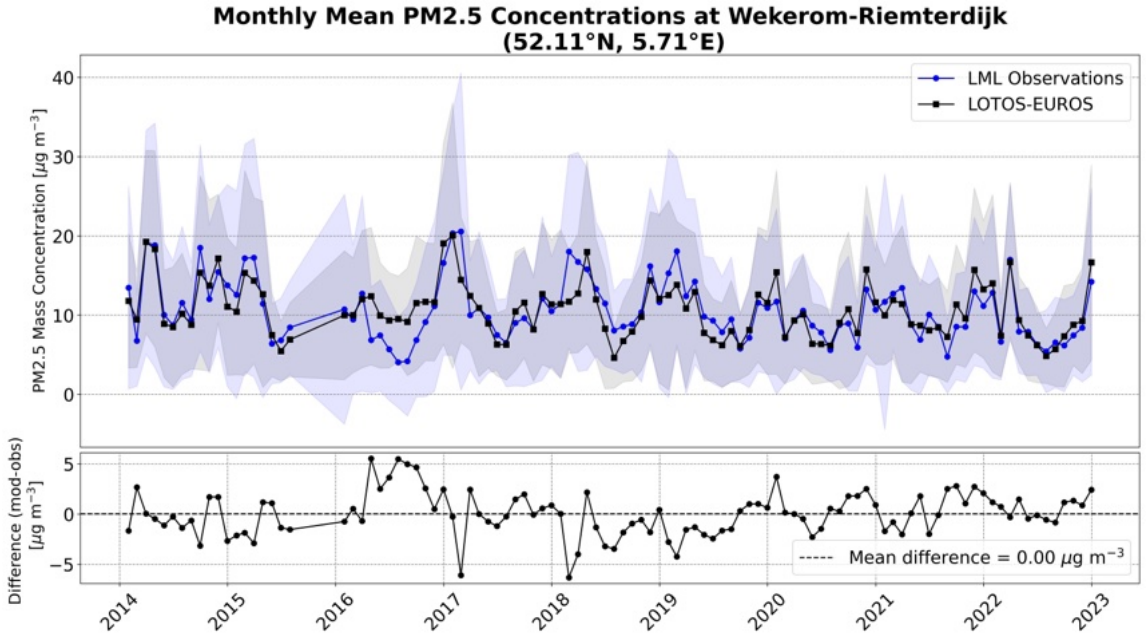


Figure 9: Monthly mean time-series of LML measured and LOTOS-EUROS simulated PM2.5 mass concentrations at the Wekerom-Riemterdijk LML site.

Site Name	Trend (ug m^{-3} yr^{-1})		Significant?	
	Observed	Model	Observed	Model
De-Zilk Vogelaarsdreef	-0.15 ± [-0.31, 0.02]	-0.25 ± [-0.41, -0.08]	No	Yes
Vredepeel-Vredeweg	-0.03 ± [-0.22, 0.17]	-0.33 ± [-0.51, -0.16]	No	Yes
Wekerom-Riemterdijk	-0.36 ± [-0.55, -0.16]	-0.34 ± [-0.51, -0.17]	Yes	Yes
Wieringerwerf-Medeblikkerweg	-0.36 ± [-0.54, -0.18]	-0.26 ± [-0.41, -0.10]	Yes	Yes

Table 7: Trends derived from the LML observations and LOTOS-EUROS modeled PM2.5 surface concentrations. The values provided in the square brackets are the 95% confidence intervals calculated using a bootstrapping approach. A trend is considered significant if the confidence intervals do not overlap with zero.

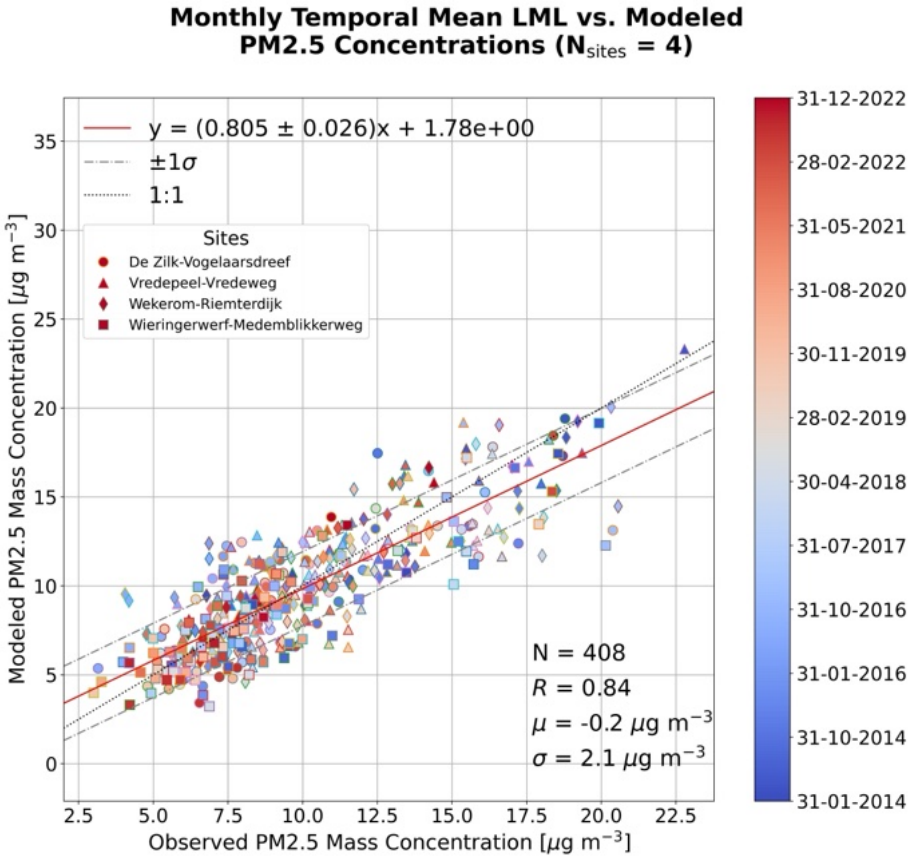


Figure 10: Correlation plot of the monthly temporal PM2.5 surface concentration means of all LML sites. Each individual data point represents a monthly mean at a given site, and the data points are colored according to time.

2.4.2 Comparisons LOTOS-EUROS with MAN NH_3

Monthly NH_3 surface concentration data from a total of 329 MAN sites was available within the period of 2014 to 2022. The mean observed and simulated surface concentrations at each MAN site over the full 2014–2022 time period is shown in Figure 11. LOTOS-EUROS is broadly able to capture the east to west gradient in the MAN observed NH_3 surface concentrations. Lower surface concentrations are found near the western coast of the Netherlands and higher concentrations are found in the east and south-east. A distinct east-to-west gradient in the measurement-model difference can also be seen, where a low bias of around -20% in the model relative to the measurements is found at the westernmost sites, and a high bias on the order of +40% is seen at the easternmost MAN sites. A similar east-to-west bias is not discernible in the LML comparisons.

A scatter plot with a linear regression of all MAN-observed monthly mean NH_3 surface concentrations from 329 sites against the LOTOS-EUROS simulated equivalent for the period of 2014-2022 is shown in Figure 12. A moderate correlation of R = 0.66 between the measured and modeled monthly means is found, with a relatively small mean bias of 1.3 ug m^{-3}, and a slope of 1.095. The cluster of higher values in the model to the top left of the 1:1 line is primarily from the higher values in the simulated surface concentrations at the easternmost MAN sites (as is visible in Figure 11).

Lastly, a time-series of the observed and simulated monthly mean surface concentrations calculated over all MAN sites is shown in Figure 13. The modeled and observed surface NH_3 concentrations broadly capture the same temporal variability for much of time-series, however the model appears to consistently capture higher concentrations in the springtime in comparison with the

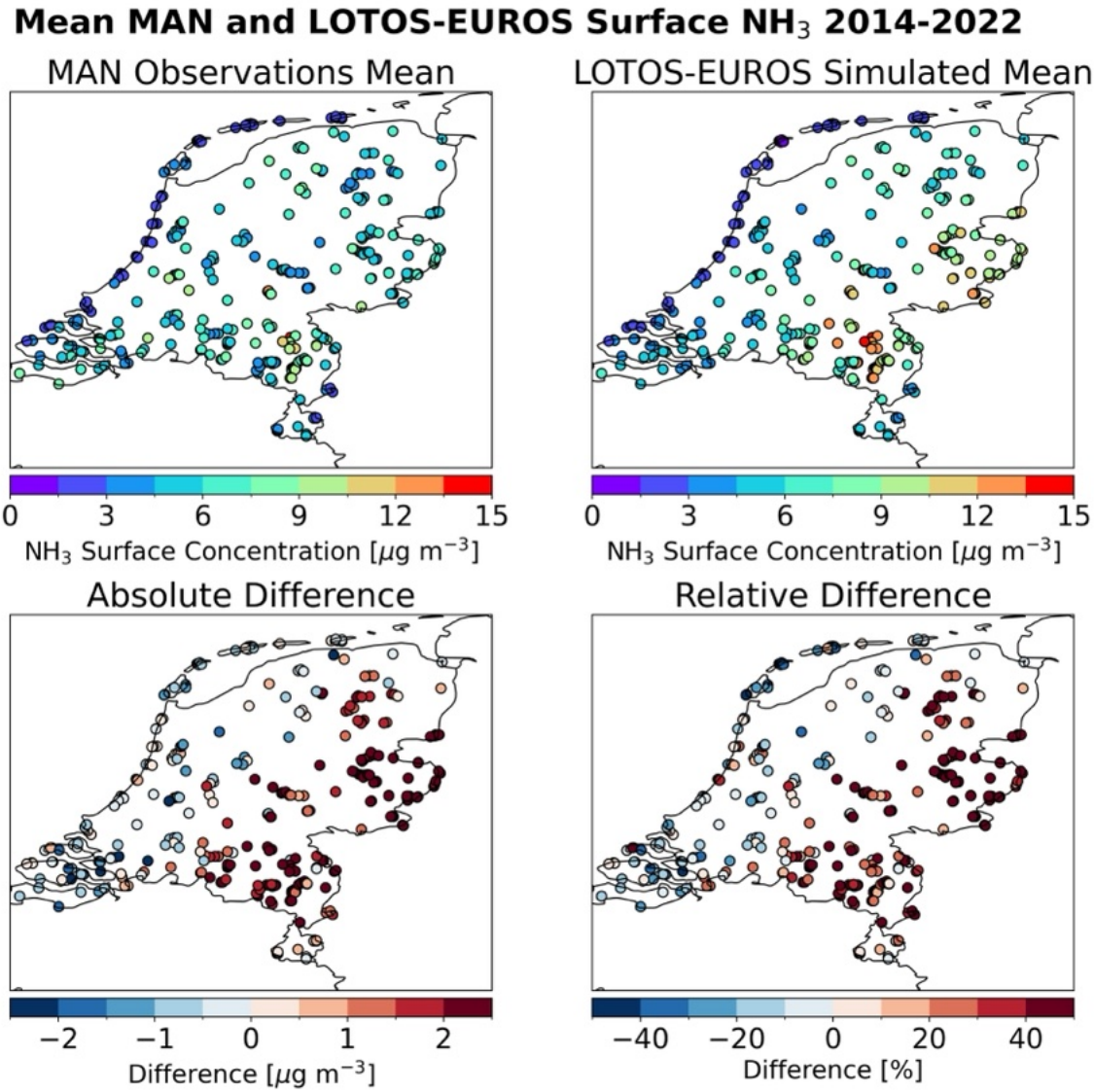


Figure 11: Mean measured and LOTOS-EUROS simulated NH_3 surface concentrations at MAN sites calculated over the period of 2014–2022.

MAN measurements.

Although part of the observed bias between the model and MAN measurements may stem from sensor-to-sensor differences and possible calibration issues with the Gradko samplers, the remaining bias is likely related to representativeness differences between the point-scale MAN measurements and the coarser model grid. Additional factors contributing to the spatial and temporal mismatch may include overestimated springtime emissions, slower-than-expected vertical mixing and underestimated deposition in forested areas within the model. Unlike LML sites, which are typically located in open fields close to emission sources, most MAN sites are situated in Natura2000 areas, often within or adjacent to forests, where canopy effects and fine-scale variability can substantially lower NH_3 concentrations relative to surrounding agricultural landscapes. Capturing this sub-grid heterogeneity remains highly challenging at the spatial resolution of LOTOS-EUROS simulations.

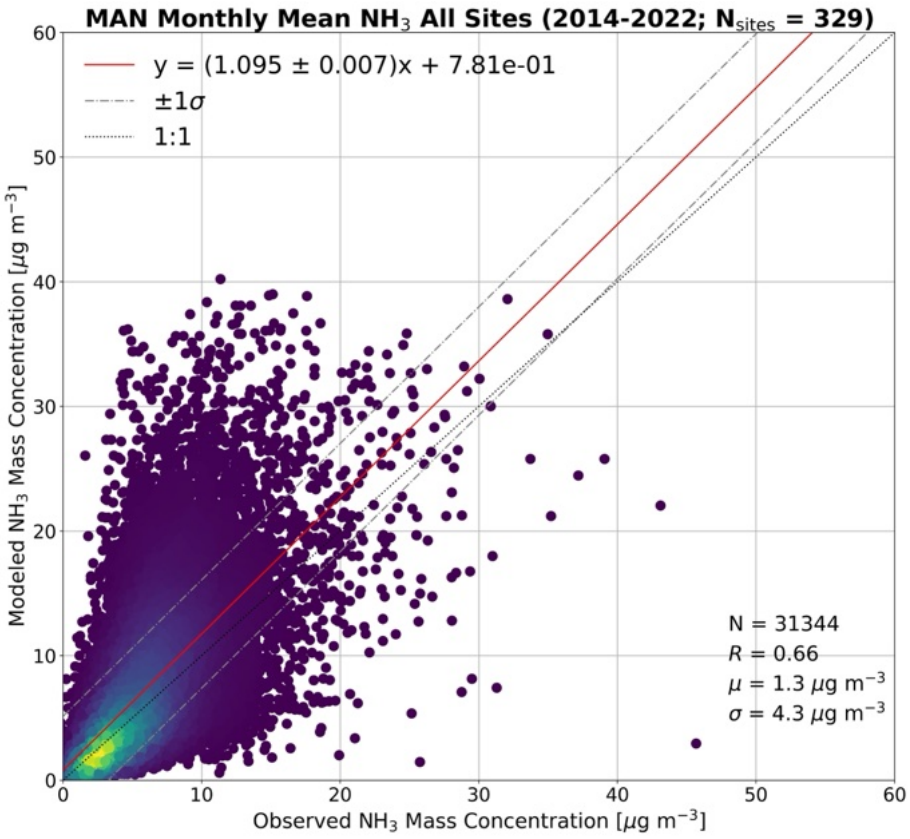


Figure 12: Correlation plot of the monthly temporal NH_3 surface concentration means of all MAN sites (x-axis) versus LOTOS-EUROS (y-axis). Each individual data point represents a monthly mean at a single site.

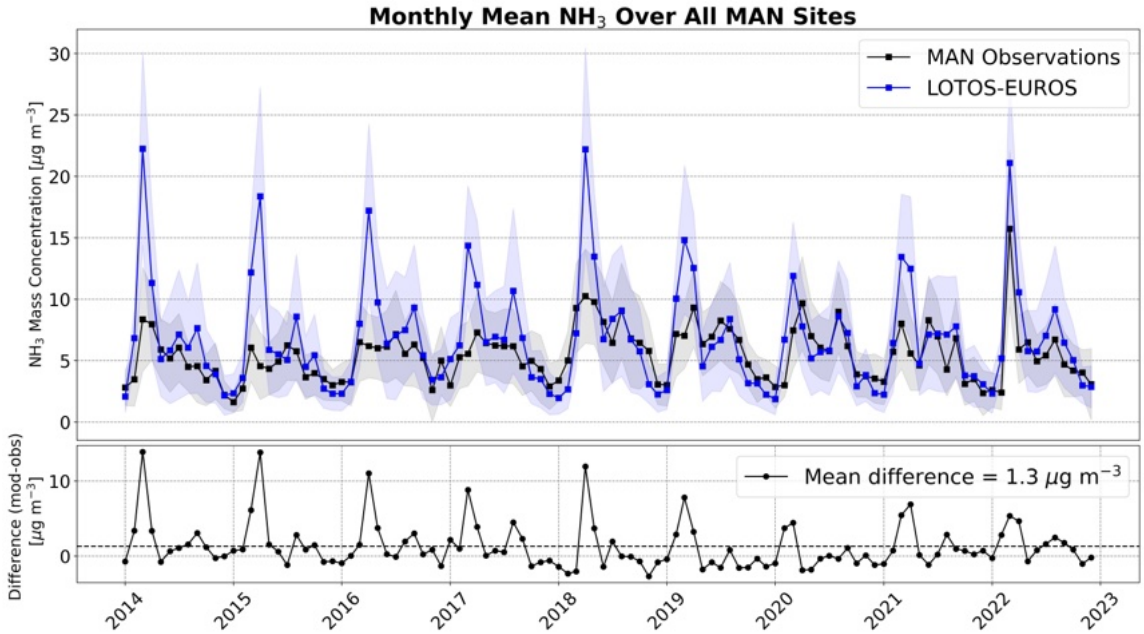


Figure 13: Monthly mean time-series of MAN observed and LOTOS-EUROS simulated NH_3 monthly mean mass concentrations calculated across all sites.

2.4.3 Comparisons of EMEP4NL and OPS-LT with MAN NH_3

In this section the measured NH_3 concentrations from the MAN network are compared to the modelled NH_3 concentrations from EMEP4NL and OPS-LT. The comparison is done on a yearly basis from 2013 to 2023. Here, the OPS-LT ("Long Term") model is used to model yearly NH_3 concentrations at the MAN measurement locations as receptor points. Furthermore, the yearly mean NH_3 concentrations from EMEP4NL at the MAN sites were extracted and compared. Only MAN sites with an uninterrupted timeseries of measurements were considered here. For this time period, that means that a total of 235 sites were used. The yearly MAN NH_3 concentrations were obtained by averaging the monthly observed values, after gap filling for non-detects following the methodology described in (Hoogerbrugge and Liem, 2000).

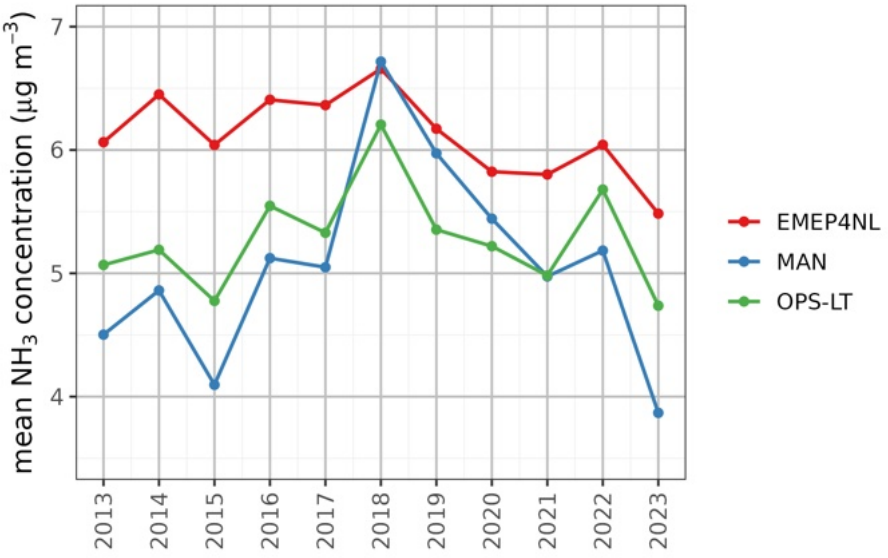


Figure 14: Mean measured and modelled NH_3 concentrations (from EMEP4NL and OPS-LT) from 2013-2023 at measurement locations (N=235).

The mean NH_3 concentrations from 2013 to 2023 measured at the MAN sites, and modelled with OPS-LT and EMEP4NL are shown in Figure 14. The measured NH_3 concentrations over the Netherlands generally increase from 2013 towards 2018, and decrease after 2018. Overall, this pattern is captured in both EMEP4NL and OPS-LT. The changes in NH_3 concentrations in EMEP4NL is, however, less steep and the NH_3 concentrations are generally overestimated compared to the measurements (with the exception of 2018). The modelled NH_3 concentrations from OPS-LT are generally closer to the observed ones, but seem to be underestimated from 2018 to 2020, and overestimated in the remainder of the years.

The yearly mean modelled NH_3 concentrations from 2013 to 2023 were compared to the measured concentrations from the MAN network. The results are shown in Figure 15. The modelled NH_3 concentrations from OPS-LT are generally closer (smaller RMSE, bias and slope) to the observed values than the modelled NH_3 concentrations from EMEP4NL. Overall, both models have a strong (EMEP4NL) to very strong (OPS-LT) correlation to the observed values, and seem to generally overestimate the NH_3 concentrations. The coloring of the points depict the corresponding longitude of each MAN site. In both models a similar spatial pattern is observed: an overall underestimation of modelled NH_3 sites towards the west parts of the Netherlands (the coast), and an overestimation of the modelled NH_3 values towards the east.

The measured and modelled yearly mean NH_3 concentrations were compared for each year individually. An orthogonal regression is fitted through the data points and the root-mean-square errors (RMSE), bias and Pearsons correlation coefficients were computed. The results are given

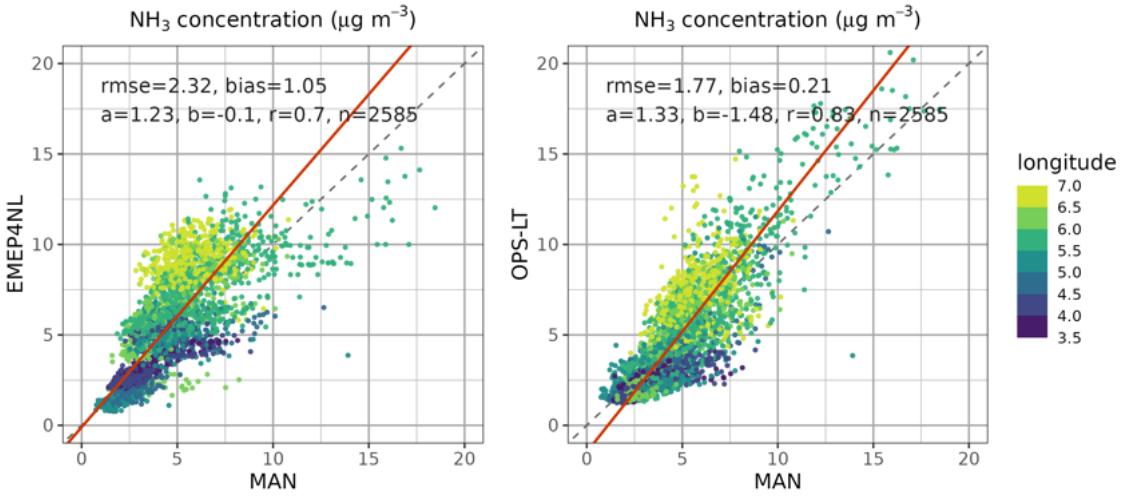


Figure 15: Scatter plot of mean yearly measured versus modelled NH_3 concentrations from 2013-2023 at measurement locations (N=235). The colors depict the corresponding longitude of each site.

in Table 8 (EMEP4NL versus MAN) and Table 9 (OPS-LT versus MAN). Furthermore, the fitted regression lines per year are shown in Figure 16.

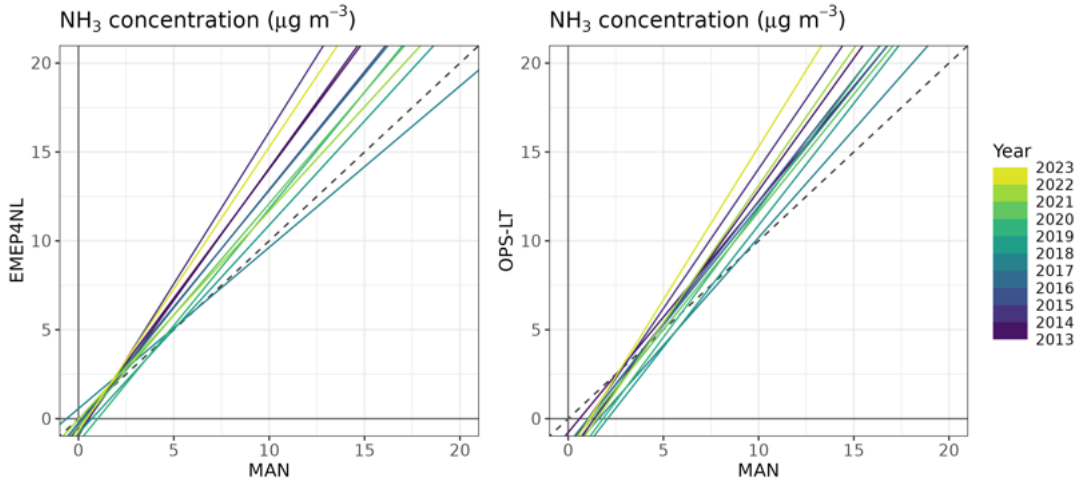


Figure 16: Fitted orthogonal regression lines for the comparison of the yearly measured and modelled NH_3 concentrations from 2013-2023 at measurement locations (N=235). The left plot shows the comparison with the modelled NH_3 concentration from EMEP4NL, the right plot with the modelled NH_3 concentration from OPS-LT.

2.4.4 Summary

The yearly mean modelled NH_3 concentrations between 2013 and 2023 from EMEP4NL and OPS-LT were compared to the measured concentrations from the MAN network. The modelled NH_3 concentrations from OPS-LT are generally closer (smaller RMSE, bias and slope) to the observed values than the modelled NH3 concentrations from EMEP4NL. Both EMEP4NL and OPS-LT show a strong (R=0.70, EMEP4NL) to very strong (R=0.83, OPS-LT) correlation to the MAN measurements, but generally overestimate the NH_3 concentrations.

	RMSE [μg m^{-3}]	Bias [μg m^{-3}]	r [-]	slope [-]	intercept [μg m^{-3}]
All years	2.32	1.05	0.70	1.23	-0.10
2013	2.80	1.56	0.67	1.45	-0.46
2014	2.67	1.59	0.71	1.49	-0.77
2015	2.76	1.94	0.74	1.71	-0.96
2016	2.61	1.28	0.68	1.33	-0.40
2017	2.43	1.31	0.74	1.31	-0.24
2018	2.15	-0.06	0.74	0.91	0.56
2019	1.79	0.20	0.75	1.17	-0.82
2020	1.77	0.38	0.74	1.31	-1.32
2021	1.96	0.83	0.77	1.25	-0.43
2022	1.95	0.86	0.72	1.17	-0.03
2023	2.30	1.62	0.73	1.60	-0.69

Table 8: Computed statistics for comparison yearly mean NH_3 concentrations from MAN and EMEP4NL.

	RMSE [μg m^{-3}]	Bias [μg m^{-3}]	r [-]	slope [-]	intercept [μg m^{-3}]
All years	1.77	0.21	0.83	1.33	-1.48
2013	1.85	0.57	0.80	1.30	-0.78
2014	1.81	0.33	0.84	1.49	-2.04
2015	1.69	0.68	0.86	1.57	-1.66
2016	1.93	0.42	0.82	1.37	-1.48
2017	1.71	0.28	0.85	1.34	-1.42
2018	1.91	-0.51	0.86	1.21	-1.95
2019	1.73	-0.62	0.85	1.37	-2.84
2020	1.60	-0.22	0.85	1.45	-2.65
2021	1.52	0.01	0.86	1.32	-1.60
2022	1.87	0.49	0.83	1.54	-2.32
2023	1.81	0.87	0.82	1.72	-1.92

Table 9: Computed statistics for comparison yearly mean NH_3 concentrations from MAN and OPS-LT.

2.4.5 Comparisons of EMEP4NL and LML

The modelled NH_3, NO_2 and NO concentrations from EMEP4NL were extracted at the locations of the LML sites. The modelled and measured concentrations were compared on hourly, daily and monthly basis by averaging across all sites (N=77) for NO_2 and NO, and across the 6 selected sites for NH_3. The comparison was done for 10 years between 2014 and 2023. The results are

summarized in the following sections.

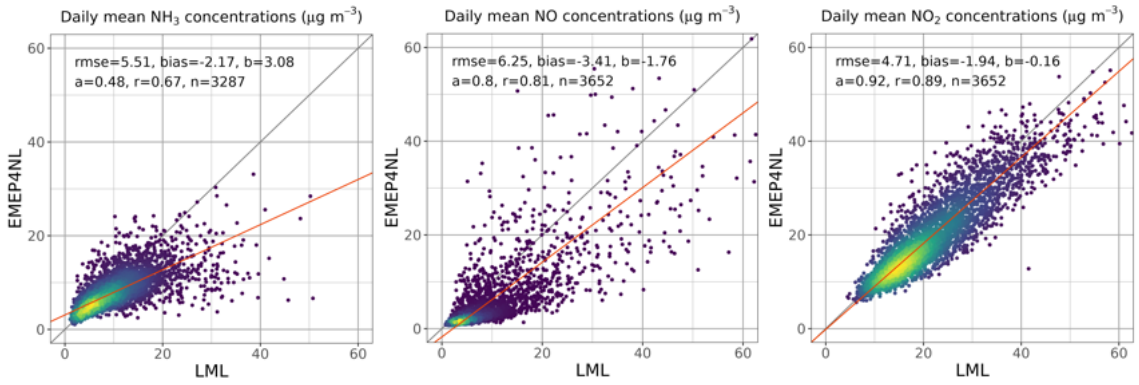


Figure 17: Scatter plots of daily mean observed (LML) and simulated (EMEP4NL) NH_3 (left), NO (middle) and NO_2 (right) surface concentrations from 2014-2024. The shown values indicate the computed root-mean-squared-errores (RMSE), bias, the slope (a) and offset (b) of the fitted orthogonal regression line and the Pearson's correlation coefficient (r).

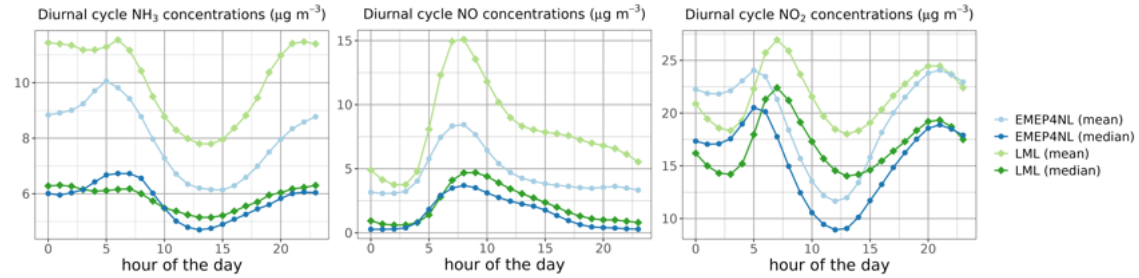


Figure 18: Average and median diurnal cycles of the observed (LML) and simulated (EMEP4NL) NH_3 (left), NO (middle) and NO_2 (right) surface concentrations from 2014-2024.

2.4.5.1 **NH_3**

The comparison of the daily observed and modelled NH_3 surface concentrations between 2014 and 2024 is shown in Figure 17. Here, all 6 LML sites are included. Overall, there is a high correlation (r=0.67) between the observed and modelled NH_3 surface concentrations. The mean bias between the two is relatively small, namely -2.17 ug m^{-3}. The root-mean-squared error is RMSE = 5.51 ug m^{-3}. The coefficients of the fitted orthogonal regression between the two (slope = 0.48, intercept = 3.08 ug m^{-3}) show an overall underestimation of the modelled NH_3 concentrations. However, closely examining the discrepancy between the two, we see that the modelled and observed NH_3 concentrations mainly diverge for the higher concentration range (>20 ug m^{-3}). For the lower range of the concentrations (<20 ug m^{-3}), on the other hand, the comparison between the two is better and closer to the 1:1 line.

Figure 18 shows the modelled and observed diurnal variation of the NH_3 surface concentrations from LML and EMEP4NL at the 6 LML sites. To get more insight into the statistical distribution of the modelled and observed NH_3 concentrations, both the mean and median values are plotted. The median diurnal cycles from LML and EMEP4NL compare well and are close to one another. The mean diurnal cycles, on the contrary, are quite different in magnitude, but still show an overall good comparison in terms of patterns.

The daily mean NH_3 surface concentrations from EMEP4NL and LML between 2014 and 2024 at the 6 LML locations are shown in Figure 19. Also, the absolute differences (LML-EMEP4NL)

between the two are plotted. Figure 20 shows the same comparison, but now averaged on monthly basis. Overall, the time variation of the modelled NH_3 surface concentrations from EMEP4NL corresponds well with the observations. The magnitude of the spring peak in NH_3 concentrations is captured well in EMEP4NL. Also, the modelled and observed NH_3 concentrations in winter correspond well. Larger differences are generally found in summer and autumn. The largest differences between the modelled and observed NH_3 concentrations can be seen in 2018. Here, the monthly mean NH_3 at the 6 LML sites were up to 10 ug m^{-3} higher than the modelled ones.

Furthermore, a trend line was fitted through the daily mean observed and modelled NH_3 surface concentrations between 2014 and 2023. The results are shown in Table 10. For this period, we found no statistically significant trend in the observed NH_3 surface concentrations at the LML sites. The modelled NH_3 surface concentrations, on the other hand, show a small but significant trend of -0.14 $\pm$ (-0.19,-0.08) ug m^{-3} yr^{-1}.

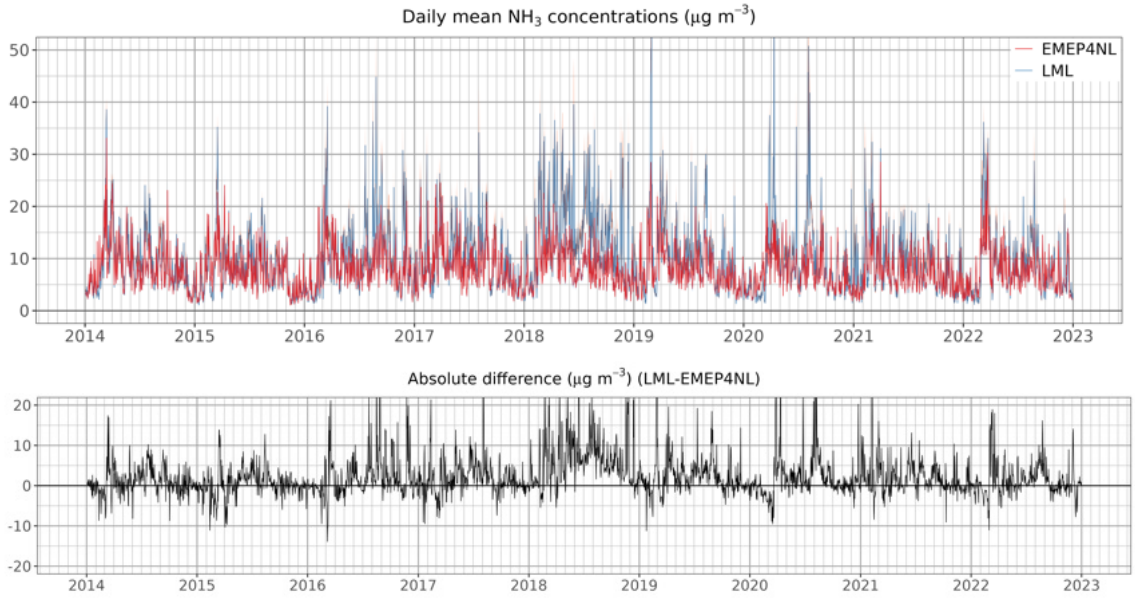


Figure 19: Timeseries of the daily mean LML measured (blue) and EMEP4NL (red) simulated NH_3 surface concentrations, and the absolute differences between the two (LML-EMEP4NL).

2.4.5.2 **NO_2**

The comparison of the daily observed and modelled NO_2 surface concentrations between 2014 and 2024 is shown in Figure 17. Here, all available LML measurements are included. The modelled and observed NO_2 surface concentrations compare well. Overall, there is a very high correlation (r=0.89) between the two and relatively small differences (RMSE = 4.71 ug m^{-3}, bias = -1.94 ug m^{-3}). Also, the coefficients of the fitted orthogonal regression are close to the 1:1 line (slope = 0.92, intercept = -0.16 ug m^{-3}). The observed NO_2 surface concentrations at the LML sites are slightly higher than the modelled values from EMEP4NL.

The diurnal variation of the observed and modelled NO_2 surface concentrations between 2014 and 2024 is shown in Figure 18. To get more insight into the statistical distribution of the concentrations, both the mean and median values are plotted. For NO_2, the mean and median diurnal cycles are closer to one another than for NH_3. There appears to be a shift in the modelled versus observed daily variation. The observed morning peak in NO_2 surface concentrations at the LML sites, in particular, is 2 hours later than the modelled one from EMEP4NL. Furthermore, the subsequent decrease in NO_2 surface concentrations is much steeper in EMEP4NL than seen at the LML sites.

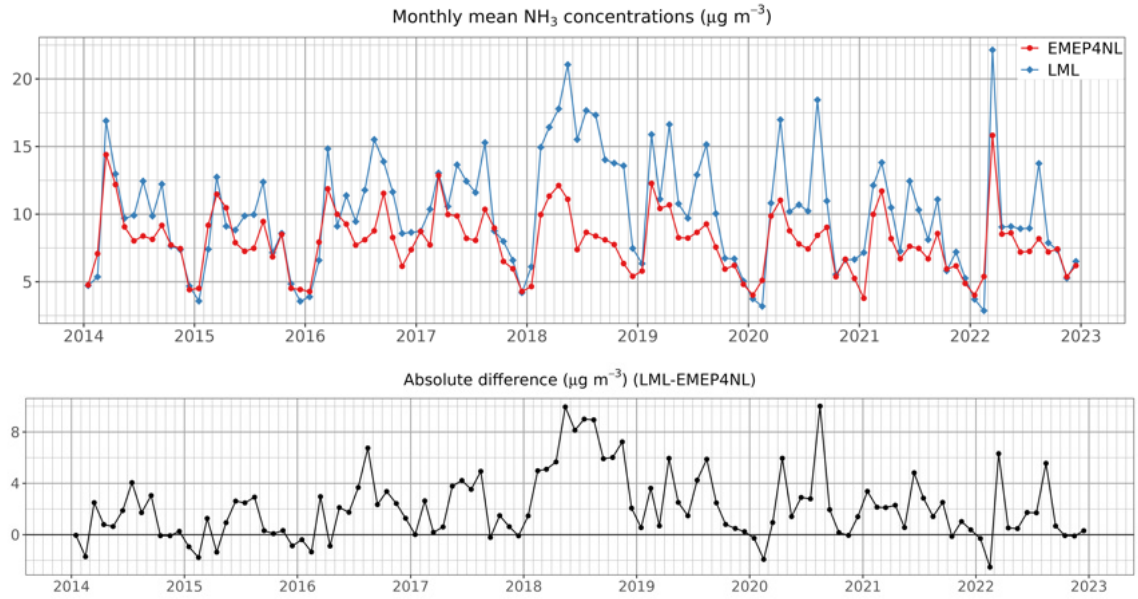


Figure 20: Timeseries of the monthly mean LML measured (blue) and EMEP4NL (red) simulated NH_3 surface concentrations, and the absolute differences between the two (LML-EMEP4NL).

Overall, the modelled daytime NO_2 surface concentrations from EMEP4NL are lower than the observed ones. The timing of the evening peak in NO_2 surface concentrations, on the other hand, compares well.

The daily mean NO_2 surface concentrations from EMEP4NL and LML between 2014 and 2024 at all available LML locations are shown in Figure 21. Also, the absolute differences (LML-EMEP4NL) between the two are plotted. Figure 22 shows the same comparison, but now averaged on monthly basis. The smallest differences between the observed and modelled NO_2 surface concentrations are found in winter. Notably, the observed and modelled NO_2 surface concentrations from the LML and EMEP4NL seem to converge towards one another, with a significantly decrease in the absolute differences between the two. The best comparison between the two is found during the last years of the comparison period. The daily mean NO_2 surface concentrations from both EMEP4NL and LML show a statistically significant trend between 2014 and 2023 (Table 10). We found a significant decrease in observed NO_2 surface concentrations at the LML sites of -1.15 $\pm$ (-1.24,-1.05) ug m^{-3} yr^{-1}. The corresponding decrease in modelled NO_2 surface concentrations from EMEP4NL is almost twice as small. Here, we found a trend of -0.67 ± (-0.77,-0.58) ug m^{-3} yr^{-1}.

2.4.5.3 **NO**

The comparison of the daily observed and modelled NO surface concentrations between 2014 and 2024 is shown in Figure 17. Again, all available LML measurements are included. We found a very strong correlation (r=0.81) between the observed and modelled NO concentrations. Overall, the measured NO concentrations were higher than the modelled ones, with a mean bias of -3.41 ug m^{-3} and a fitted slope of 0.80 (intercept = -1.76 ug m^{-3}). There is a relatively larger spread in observed versus measured NO concentrations than found for NH_3 and NO_2, which corresponds to a slightly higher root-mean-squared error of 6.25 ug m^{-3}.

The diurnal variation of the observed and modelled NO surface concentrations between 2014 and 2024 is shown in Figure 18. Again, both the mean and median values are plotted to get insight into the statistical distributions of the observed and modelled concentrations. The median observed

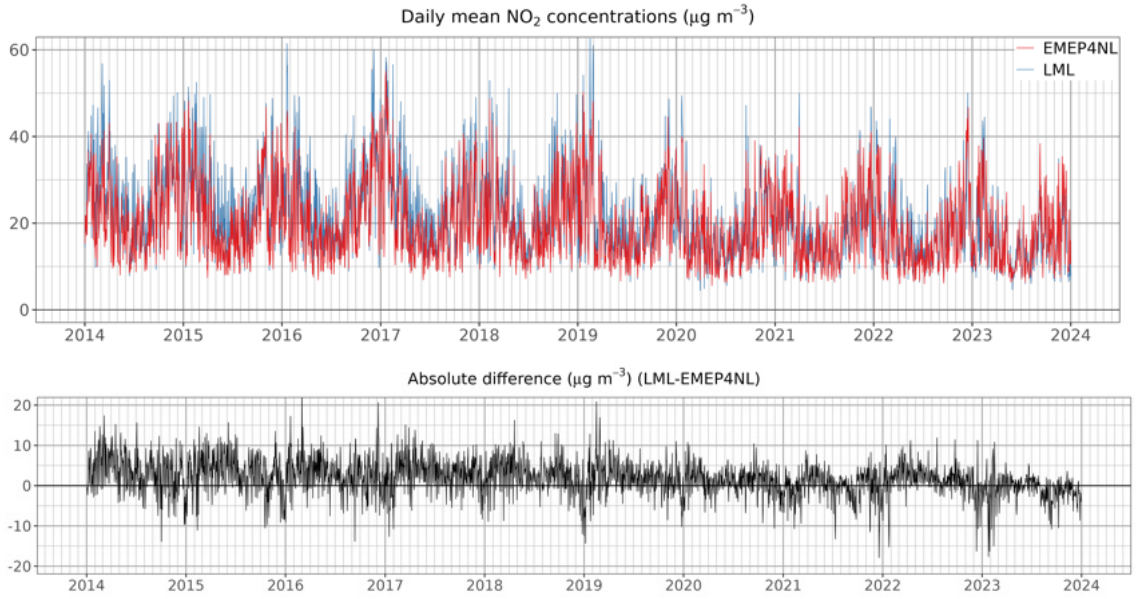


Figure 21: Timeseries of the daily mean LML measured (blue) and EMEP4NL (red) simulated NO_2 surface concentrations, and the absolute differences between the two (LML-EMEP4NL).

and modelled NO diurnal cycles are very similar in both pattern and magnitude. The mean diurnal cycle of the LML observations is much higher than the modelled one from EMEP4NL (up to almost a factor 2), which is indicative of fewer outliers in the modelled NO concentrations compared to the observed ones.

The daily mean NO surface concentrations from EMEP4NL and LML between 2014 and 2024 at all available LML locations are shown in Figure 23. Also, the absolute differences (LML-EMEP4NL) between the two are plotted. Figure 24 shows the same comparison, but now averaged on monthly basis. The differences between the observed and modelled NO concentrations are generally smallest in summer, and largest in spring and autumn. Moreover, we observe a similar pattern as for NO_2; the differences between the observed and modelled NO concentrations become smaller over time. This corresponds to the discrepancy in computed trends between 2014 and 2023 in daily mean observed versus modelled NO concentrations 10. Both trends are statistically significant, but where the observed NO surface concentrations at the LML sites show a trend of -0.81 $\pm$ (-0.90,-0.71) ug m^{-3} yr^{-1}, the corresponding trend in modelled NO_2 surface concentrations from EMEP4NL is only -0.27 $\pm$ (-0.35,-0.19) ug m^{-3} yr^{-1}.

	Trend (ug m^{-3} yr^{-1})	
	observations	model
NH_3 (daily means)	-0.04 $\pm$ (-0.13,0.05)	-0.14 $\pm$ (-0.19,-0.08)
NO_2 (daily means)	-1.15 $\pm$ (-1.24,-1.05)	-0.67 $\pm$ (-0.77,-0.58)
NO (daily means)	-0.81 $\pm$ (-0.90,-0.71)	-0.27 $\pm$ (-0.35,-0.19)

Table 10: Trends derived from the daily mean measured and modelled surface concentrations of NH_3 NO_2, and NO. The measured concentrations are from the LML sites, the modelled from EMEP4NL. The values the square brackets represent the 95% confidence intervals of the fitted slopes.

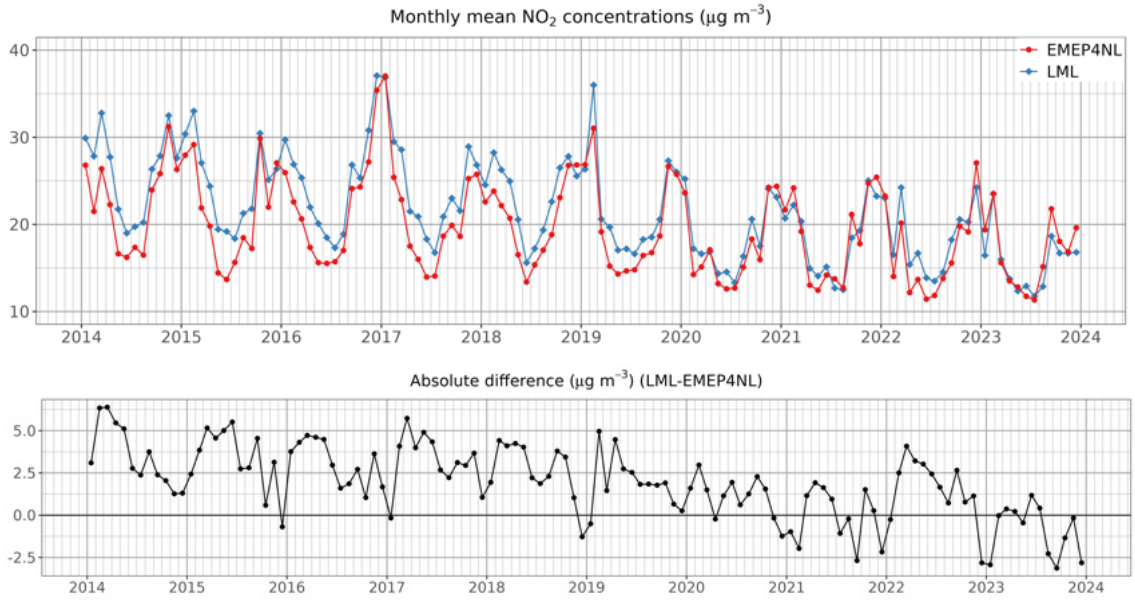


Figure 22: Timeseries of the monthly mean LML measured (blue) and EMEP4NL (red) simulated NO_2 surface concentrations, and the absolute differences between the two (LML-EME4NL).

2.4.6 Summary

The modelled NH_3, NO_2 and NO concentrations from EMEP4NL were compared to the measured concentrations at the LML sites between 2014 and 2023.

- The comparison for NH_3 at 6 selected LML sites showed an overall high correlation (r=0.67) between the observed and modelled daily NH_3 surface concentrations. The modelled and measured concentrations compare well in terms of temporal distribution. However, the measured concentrations are generally higher. For NH_3, no statistically significant trend in the mean measured concentrations at the LML sites was found. The modelled NH_3 surface concentrations from EMEP4NL showed a small but significant trend of -0.14 $\pm$ (-0.19,-0.08) ug m^{-3} yr^{-1}.
- The comparison for NO_2 showed a very high correlation (r=0.89) between the daily observed and modelled concentrations and relatively small differences (RMSE = 4.71 ug m^{-3}, bias = -1.94 ug m^{-3}, slope = 0.92, intercept = -0.16 ug m^{-3}). Overall, the temporal patterns between the two compare well, with the smallest differences in wintertime. However, a shift between the modelled and observed diurnal cycles of the NO_2 concentrations was found. We found statistically significant trends in the daily mean NO_2 surface concentrations from EMEP4NL and LML between 2014 and 2023. The decrease in observed NO_2 surface concentrations at the LML sites amounts to -1.15 $\pm$ (-1.24,-1.05) ug m^{-3} yr^{-1}, where the corresponding decrease in modelled NO_2 surface concentrations from EMEP4NL is almost twice as small and amounts to -0.67 ± (-0.77,-0.58) ug m^{-3} yr^{-1}.
- The comparison for NO showed a very strong correlation (r=0.81) between the observed and modelled concentrations. The measured NO concentrations are generally higher than the modelled ones but compare well in terms of temporal distribution. We found statistically significant trends in the daily mean NO surface concentrations from EMEP4NL and LML between 2014 and 2023. Like NO_2, the trend in the observed NO concentrations from the LML (-0.81 $\pm$ (-0.90,-0.71) ug m^{-3} yr^{-1}), is significantly higher than the trend in modelled NO concentrations from EMEP4NL (-0.27 $\pm$ (-0.35,-0.19) ug m^{-3} yr^{-1}).

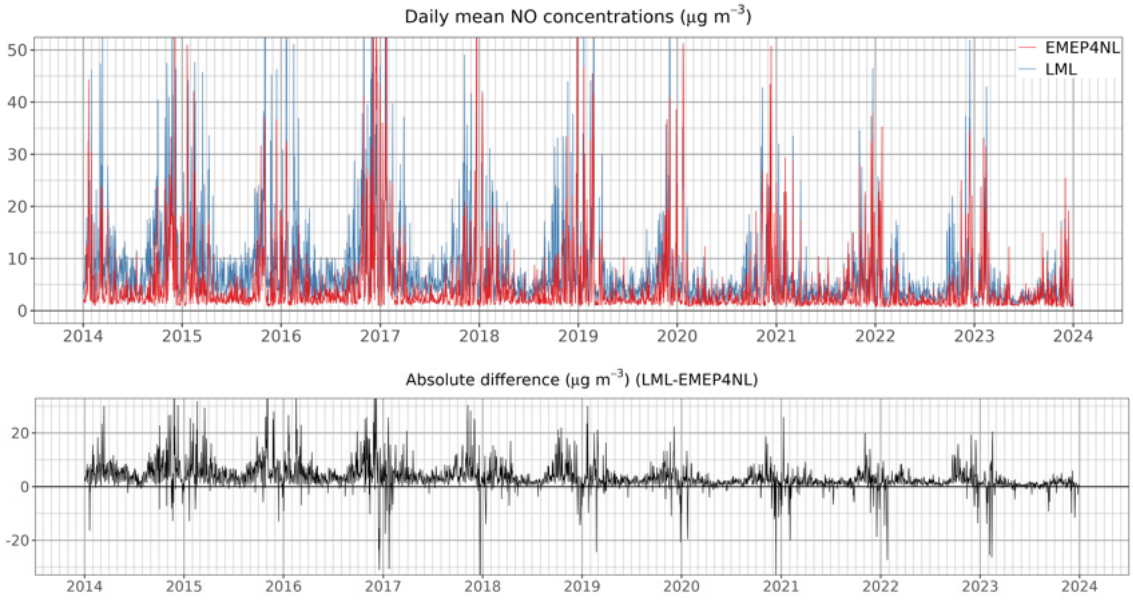


Figure 23: Timeseries of the daily mean LML measured (blue) and EMEP4NL (red) simulated NO surface concentrations, and the absolute differences between the two (LML-EMEP4NL).

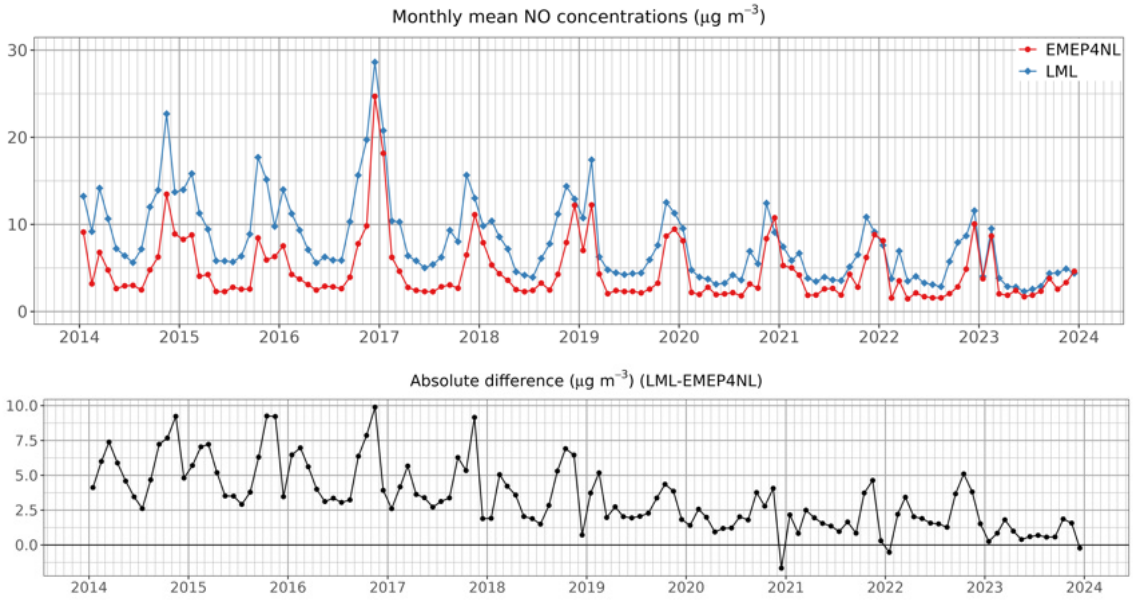


Figure 24: Timeseries of the monthly mean LML measured (blue) and EMEP4NL (red) simulated NO surface concentrations, and the absolute differences between the two (LML-EMEP4NL).

3 Validation of satellite data with surface remote sensing observations and models

3.1 Introduction

In this chapter we discuss the available satellite datasets for NH_3 and NO_2, the ground-based remote sensing datasets used to validate the satellite data, and then discuss the results of the comparisons between the ground-based and satellite products for the two species.

3.2 Datasets

3.2.1 Introduction to satellite validation

The best way to validate a satellite product is through comparison of measurements made with any instrument capable of making an observation that resembles the same state of the atmosphere as observed by the satellite as closely as possible. In the case of TROPOMI-NO_2 and CrIS and IASI-NH_3 satellite observations the amount of available vertical information of the atmosphere is fairly limited. This means that direct comparison with in-situ surface concentrations is of limited value, as the conversion to a total column concentration strongly depends on the knowledge of the shape of the vertical profile, which is not observed. Direct comparison with airborne observations measuring the full tropospheric vertical profile (and stratospheric in the case of NO_2) would be ideal. Such observations however are cost and manpower intensive and not regularly repeated around the globe. This means that the representativity of such validation sets is fairly limited, as there are not enough airborne campaigns to create a representative set of observations for typical atmospheric states around the globe. An alternative are instruments that continuously observe the complete vertical atmospheric state. There are several instrument capable of making such observations of NO_2 and NH_3 concentrations, with varying capability to resolve the vertical profile. In the case of NO_2 there are the MAX-DOAS and Pandora instruments of which there are now over a 100 in continuous operation around the globe. For NH3 there is the option in using Fourier Transform InfraRed instruments (FTIR) of which there are 10s in continuous operation within the NDACC-IRWG network. In the next two subsections more details will be given on the measurement methodologies of both sets of instruments.

3.2.2 FTIR-NH3

3.2.2.1 FTIR retrieval strategy and data product

The Network for the Detection of Atmospheric Composition Change InfraRed Working Group (NDACC-IRWG) is a global network of universities and institutes operating FTIR instruments and related projects. The working group collaborates closely to ensure a high common standard in the measurements of infrared solar spectra. Such spectra can be used to derive atmospheric concentrations of trace gases in the atmosphere. One trace gas that can be observed with FTIR instruments is ammonia. At the start of the NKS project, contacts within the NDACC-IRWG network were requested to join in a voluntary network-wide update of the FTIR NH_3 data product. Participants were asked to perform retrievals using a common retrieval methodology following a general protocol. Most groups respected the protocol with some deviations based on site requirements and time limitations. Several groups participated, resulting in an update of the measurement records at a total of 19 sites. A focus was given to the data records covering our period of inter-

est (2008–2023) but in some cases spectral records go back much further. An overview of the measurement sites is given in Figure 25. Because of the network's global coverage, a variety of atmospheric ammonia regimes are covered by instruments, ranging from arctic to tropical, and from unpolluted to highly polluted conditions.

Figure 25: Map of the NDACC-IRWG FTIR sites used in the validation of the CrIS and IASI NH_3 products.

FTIR instruments are used to measure solar absorption with the use of a Fourier transform spectrometer. These solar spectra can be used to analyze the absorption of molecular species in the atmosphere. With the use of a line-by-line radiative transfer model, the absorption lines can be modeled using pre-determined absorption parameters from the HITRAN 2012 spectral database (Rothman et al., 2013). Dammers et al. (2015) developed a retrieval strategy to optimally determine atmospheric NH_3 concentrations. The retrieval recipe is based on a combination of two spectral microwindows covering two strong ammonia absorption lines at [930.32 — 931.32 cm-1, MW1] and [966.97 — 967.68 cm-1, MW2]. In addition to NH_3, there are a number of interfering species with absorption lines that are taken into account (e.g., H_2O, O_3, CO_2, N_2O, HNO_3, SF_6, CFC-12 and solar lines). For some of the locations this common strategy was not feasible due to very low ambient NH_3 concentrations, and small adjustments were made to widen the windows and increase the information content [929.40 — 931.40 cm-1, MW1] and [962.10 — 970.00 cm-1, MW2]. Spectra at locations with enhanced ammonia concentrations also allowed for the retrieval of concentrations without any further constraints to the profile. In some cases with lower concentrations, however, further constraints were needed for the profile shape, and Tikhonov regularization was added to the retrieval protocol. To ensure that only the best observations are used for the satellite product validation, posterior data filters were applied to remove observations with negative column totals and those with an information content (Degrees of Freedom) less than 0.5. All sites (except Lauder) used a common set of modeled concentrations (GEOS-CHEMv11, 2013-2017 mean) for the a priori ammonia concentration profiles, details can be found in (Herrera et al., 2022). The FTIR instruments are not always located in regions with high ammonia concentrations, with most sites purposely built away from common anthropogenic sources. This means that even with global coverage only a few sites provide data under enhanced ammonia concentrations. For the purposes of the NKS project, all high altitude and arctic sites were removed from the dataset, to ensure that the validation of the data-product represents typical Dutch conditions, which are elevated concentrations. Table 11 lists the remaining sites after the removal of the high altitude (Izana – Tenerife/Spain, Maido - Reunion, Jungfraujoch – Switzerland), tropical (St. Denis – Reunion/France, Paramaribo - Suriname, Porto Velho – Brazil) and arctic sites (Eureka – Canada, Thule – Greenland/Denmark, Ny Alesund – Spitsbergen/Norway).

Site	Country	Latitude	Longitude	Period
St. Petersburg	Russia	59.88	29.83	2010 – 2023
Bremen	Germany	53.10	8.85	2006 – 2023
Toronto	Canada	43.66	-79.40	2002 – 2023
Hefei	China	31.91	117.17	2017 – 2023
Rikubetsu	Japan	43.46	143.77	2014 – 2023
Boulder	United States	40.04	-105.24	2010 – 2023
Tsukuba	Japan	36.05	140.12	2014 – 2023
Mexico City	Mexico	19.33	-99.18	2014 – 2023
Wollongong	Australia	-34.41	150.88	2007 – 2023
Lauder	New Zealand	-45.04	169.68	1996 – 2023

Table 11: FTIR sites with datasets used in this study, ordered by latitudinal location. Period indicates the length of the dataset, cut-off at the start of 2023.

3.2.2.2 Further data-criteria

In this study we only use observations of the CrIS instrument passing the recommended Quality_Flag of $\geq$3 and those that do not have a Cloud_Flag equal to 1. Furthermore, to optimize the comparison of similar airmasses and information content of the FTIR and satellite observations, additional filter criteria have been applied in this validation exercise. All FTIR observations with a DOF smaller than 0.5 as well as those with negative total column concentrations have been excluded. Similarly, all CrIS observations with low information content (DOF$\leq$0.5), non-detects (Cloud_Flag equal to 3), low signal-to-noise ratios (SNR$\leq$2.0) and high measurement errors (relative measurements error $\geq$1.0) and low negative thermal contrast or thermal inversions (TC$\leq$2) have been removed. Also all CrIS observations with suspected signatures of biomass burning plumes, e.g. those with a Cloud_Flag equal to 2, a DOF$\geq$1.5 and those with a concentration profile peaking at least 2 layers beyond the mean are excluded. Lastly only daytime observations are used for both CrIS and IASI.

3.2.2.3 Co-location criteria and the observational operator application

Satellite and FTIR observations are co-located in time and space by matching each individual FTIR observation with all satellite observations within 25 kilometers and 45 minutes. Ammonia column concentrations can vary strongly spatially and throughout the day (Dammers et al., 2017; Tournadre et al., 2020). The tight criteria reduces variations in observed sources while ensuring enough observations remain for a validation study. Impact of topography is reduced by removing all observations with a surface altitude difference of more than 300 meters. For each combination of FTIR and satellite observation the FTIR observational operator is applied. To account for instrument sensitivity and to ensure comparability between the IASI and CrIS results, the FTIR operator is applied to the satellite data and not in reverse. The satellite profiles are mapped to the FTIR retrieval layers after which the operator is applied following equation 3.1;

$$\hat{x}_{\text{sat}} = x^{a}_{ftir} + A_{ftir}(x_{sat} - x^{a}_{ftir}) \tag{3.1}$$

where x^{a}_{ftir} is the FTIR apriori, x_{sat} is the satellite-retrieved concentration profiles, A_{ftir} is the FTIR averaging kernel, and $\hat{x}_{sat}$ is the adjusted satellite concentration profile. The IASI concentration profiles are derived based on the formula given in Clarisse et al. (2023). At the final step, all

FTIR-satellite combinations matching a unique satellite overpass are averaged together on an overpass-to-overpass basis.

3.2.3 MAX-DOAS and PANDORA NO2

3.2.3.1 Introduction

Satellite observations provide a comprehensive global perspective of NO_2 (Goldberg et al., 2019), complementing localized data from sparsely distributed ground-based monitoring stations. Among these, the TROPOspheric Monitoring Instrument (TROPOMI) on the Sentinel-5 Precursor satellite has become a cornerstone of atmospheric monitoring. With its high spatial resolution (up to 3.5 × 5.5 km^2) and daily global coverage, TROPOMI offers valuable data that supports various aspects of air quality management, including analyzing pollution dynamics, tracking emission sources, and validating air quality models.

The accuracy of satellite-based NO_2 observations, and the reliability of any applications built upon them, depends on rigorous validation against independent ground-based and airborne measurements. Traditionally, this validation is performed using UV-VIS Differential Optical Absorption Spectroscopy (DOAS) instruments. For tropospheric column measurements, Multi-Axis DOAS (MAX-DOAS) is widely used, while Pandora direct Sun instruments are employed for assessing total columns (Verhoelst et al., 2021). The validation thus helps to identify and correct biases, assess uncertainties, and refine retrieval algorithms.

TROPOMI Level 2 (L2) NO_2 measurements undergo operational validation through the S5P Mission Performance Centre – Validation Data Analysis Facility (S5P-MPC-VDAF) and several independent studies. These studies generally indicate a low bias in TROPOMI observations over polluted regions, with underestimations of up to 30%, primarily attributed to the coarse vertical a priori profile used in the retrieval process. In this study, we focus on validating TROPOMI NO_2 data specifically for the Netherlands to assess how these biases manifest under the unique characteristics of the region. Factors such as seasonal variations, latitude, weather conditions, aerosol presence, source distribution, and regional pollution patterns are expected to influence the biases, making region-specific analysis essential.

3.2.3.2 Pandora

The Pandora instrument is a ground-based stationary spectrometer designed to measure sunlight in the UV-VIS spectral range (280-525 nm) with a high resolution of 0.6 nm (Herman et al., 2009). It delivers high-quality, spectrally resolved radiance measurements from direct-sun observations or sky scans, facilitating the retrieval of total and tropospheric column densities, as well as vertical profile information, for trace gases such as NO_2, O_3, SO_2, and CH_2O (Herman et al., 2019; Tzortziou et al., 2023).

NO_2 total vertical column densities are obtained through direct-sun measurements, during which the instrument actively tracks the sun to capture direct sunlight. Slant column densities are derived using a Differential Optical Absorption Spectroscopy (DOAS) technique within the 400-440 nm spectral range, and subsequently converted to total column densities by applying direct-sun geometry AMF (Herman et al., 2009). Direct-sun measurements are highly valuable for validation and evaluation because they yield low uncertainties in the AMF.

The Pandora instrument 118s1, located in Cabauw, Netherlands (51.97° N, 4.93° E), is a member of the Pandonia Global Network. For this study, we exclusively use direct-sun observations from this instrument, processed via the nvs3 retrieval algorithm (Cede et al., 2023). The dataset is publicly accessible at https://data.ovh.pandonia-global-network.org/Cabauw/Pandora118s1/L2/Pandora118s1_Cabauw_L2_rnvs3p1-8.txt.

3.2.3.3 MAX-DOAS

The Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) technique is a remote sensing method designed to measure the vertical and horizontal distributions of atmospheric trace gases and aerosols (Dimitropoulou et al., 2020; Honninger et al., 2004). This technique utilizes spectroscopic observations of scattered sunlight in the UV-VIS spectral range at various viewing angles, and a DOAS technique to derive slant column densities. By observing sunlight scattered across different angles, MAX-DOAS instruments achieve greater sensitivity to near-surface absorbers and can retrieve vertical profile information, which enhances the accuracy of measurements within the lower atmosphere (Honninger et al., 2004).

Publicly accessible MAX-DOAS measurements of NO_2 are provided by the Network for the Detection of Atmospheric Composition Change (NDACC) for two observation sites in the Netherlands: De Bilt (52.10°N, 5.18°E) and Cabauw (51.97°N, 4.93°E). These instruments are operated by the KNMI. For our validation, we used only data that was centrally processed using the FRM4DOAS processor, applying the Mexican MAX-DOAS Fit (MMF) method for NO_2 retrievals (Friedrich et al., 2019).

3.2.4 CrIS-NH3

The CrIS-1 instrument on board the Suomi-NPP and CrIS-2 on NOAA-20 were launched in October 2011 and November 2017 with data products starting in May 2012 and March 2019. Both instruments are in a sun-synchronous orbit, providing twice daily global coverage at around 13:30 and 1:30 local solar time and overpass within 45 minutes of one another. The instruments provide observations with circular pixel footprints of around 14 km at nadir over a 2200km wide swath. Within this study we use the CrIS-Fast Physical Retrieval (CFPR) NH_3 product version 1.6.4 (M. W. Shephard & Cady-Pereira, 2015; M. W. Shephard et al., 2020). In short the CFPR method consists of a physical retrieval based on the Rodgers (2000) optimal estimation method in combination with a fast optimal spectral sampling forward model (Moncet et al., 2008), minimizing the residual between the measured and simulated spectra. The CFPR product has a reported detection limit at around 0.9 ppbv (M. W. Shephard & Cady-Pereira, 2015), which under ideal conditions improves to 0.3 ppbv (Kharol et al., 2018). The product went through several iterations since the last major validation study (Dammers et al., 2017) with the major changes being the addition of a cloud flag based on VIIRS data, the addition of non-detects (White et al., 2023), and a quality flag. The validation study by Dammers et al. (2017) reported an overall good correlation between FTIR and satellite observations ($R \approx 0.8$) with a slight high bias between both datasets (slope = 1.02). For column concentrations in the higher range the satellite observations showed a small positive difference with a standard observation around $25 - 50\%$, while in the lower range the bias increased to 2.5×10^{15} molecules cm^{-2} with a standard deviation of around 50-100%. In this study we only use observations passing the recommended quality_flag of $\geq$3 and exclude all observations with a cloud_flag equal to 1. Only daytime observations are used, as nighttime FTIR observations are not available. The CrIS CFPR NH_3 data product is freely available via a formal request to Environment & Climate Change Canada.

3.2.5 IASI-NH3

The IASI instruments aboard the MetOp-A, -B and -C satellites are also in Sun-synchronous orbits and pass locations twice daily with Equator crossing times at 09:30 and 21:30 local time with a difference of about 45 min to one another (Clerbaux et al., 2009). As of writing the data products of IASI-A, B and C cover the periods of October 2007 – October 2021 (IASI-A), March 2013 – June 2022 (IASI-B, with the latter date indicating the availability of data at the start of the project) and September 2019 – June 2022 (IASI-C, with the latter date indicating the availability of data at the start of the project) with the product of IASI-B and -C still in operation. All three IASI

instruments have an observational swath width of over 2000 km and have a pixel footprint of around 12 km diameter at nadir increasing to up to $\tilde{2}0\times40$ km^2 at the edges of the swath. In this study we evaluate the most recent IASI product, the Artificial Neural Network for IASI (ANNI)v4 (Clarisse et al., 2023). It is an updated version of the earlier IASI-NNv2 and LUT products (Van Damme et al., 2014; Van Damme et al., 2017). Similar to the earlier versions, the IASI-ANNIv4 retrieval consists of two steps. The first is the calculation of the Hyperspectral Range Index (HRI), characterizing the signal strength of NH_3 within each spectrum. The second step consists of a neural network, which was trained on a large set of modeled data, linking the HRI to an NH_3 total column while accounting for various parameters that can impact the spectrum around the NH_3 lines. The main change is the additional of a column averaging kernel to the product, to ease the comparison with in-situ and model data, allowing to account for the effect of the a-priori profile shape. Some of the earlier products were evaluated by Dammers et al. (2016), Herrera et al. (2022), and Kutzner et al. (2021) in earlier studies, in most cases with a limited set of (in-situ) observations. Here we use the recommended data quality filters; pre-filter and post-filter set to 1, cloud_fraction <25% and we only use observations of the morning overpass, as the available late evening FTIR observations are very limited. The IASI NH_3 data is freely available through the AERIS infrastructure, https://iasi.aeris-data.fr/nh3/.

3.2.6 TROPOMI NO_2

The TROPOMI instrument, on board of the Sentinel-5 Precursor (S5P) polar satellite, is a nadir-viewing spectrometer launched in October 2017. The equator overpass occurs around 13:30 local time. It measures radiation across the ultraviolet, visible, and infrared spectral regions and is utilized for monitoring atmospheric trace gases and aerosols (Veefkind et al., 2012). The measured NO_2 tropospheric columns have a spatial resolution in nadir of 7 x 3.5 km before, and improved to 5.5 x 3.5 km since 6 August 2019.

The NO_2 columns are retrieved following a three-step procedure. First, the NO_2 slant column density is derived from the L1b spectra measured by TROPOMI using a DOAS fit. The slant column is divided into a stratospheric and tropospheric fraction using data assimilation within the TM5-MP model at a 1°x1° horizontal resolution (Williams et al., 2017). Lastly, the slant columns are converted into VCD using total and altitude-dependent air mass factors (AMFs). The AMFs are dependent on the NO_2 vertical profiles derived from TM5-MP, the viewing geometry of the satellite, the surface albedo, surface pressure, and clouds and aerosols characteristics. Further description of the retrieval can be consulted at (J. van Geffen et al., 2022; J. H. G. M. van Geffen et al., 2024).

The development of the NO_2 retrieval is described in detail in the Sentinel-5P NO_2 Algorithm Baseline Document (J. H. G. M. van Geffen et al., 2024). An overview of the processor versions and main retrieval updates is provided in the Product Readme File (Eskes et al., 2025). We are using the latest versions of the retrieval, which is v2.4 for the reprocessed data from 30 April 2018 until 25 July 2022, followed by offline v2.4 (until 2023-03-12), v2.5 (until 2023-11-26), v2.6 (until 2024-09-07), v2.7.1 (until 2024-11-20) and v2.8 TROPOMI NO_2 observations. The version 2.4 reprocessed data, v2.4 offline data and v2.5, from 30 April 2018 up to 26 November 2023 (5 1/2 years of data) were produced with the same retrieval code, and may be combined into one consistent dataset. The averaged measurements over the Netherlands for this period are shown in Fig. 26.

We note that v2.6 introduced a change in the cloud retrieval with repercussions for NO_2, generally leading to lower values, and it is advised to discard data from this version, in particular for trend studies (Eskes et al., 2025). Version 2.7.1 introduced a new albedo dataset, with in general small, but significant, increases between 0% and 9%, mainly during winter months at mid-latutudes. The v2.8 upgrade introduced an improved cloud retrieval, with an impact especially over snow-covered land.

Routine validation of TROPOMI NO_2 observations against ground-based MAX-DOAS measure-

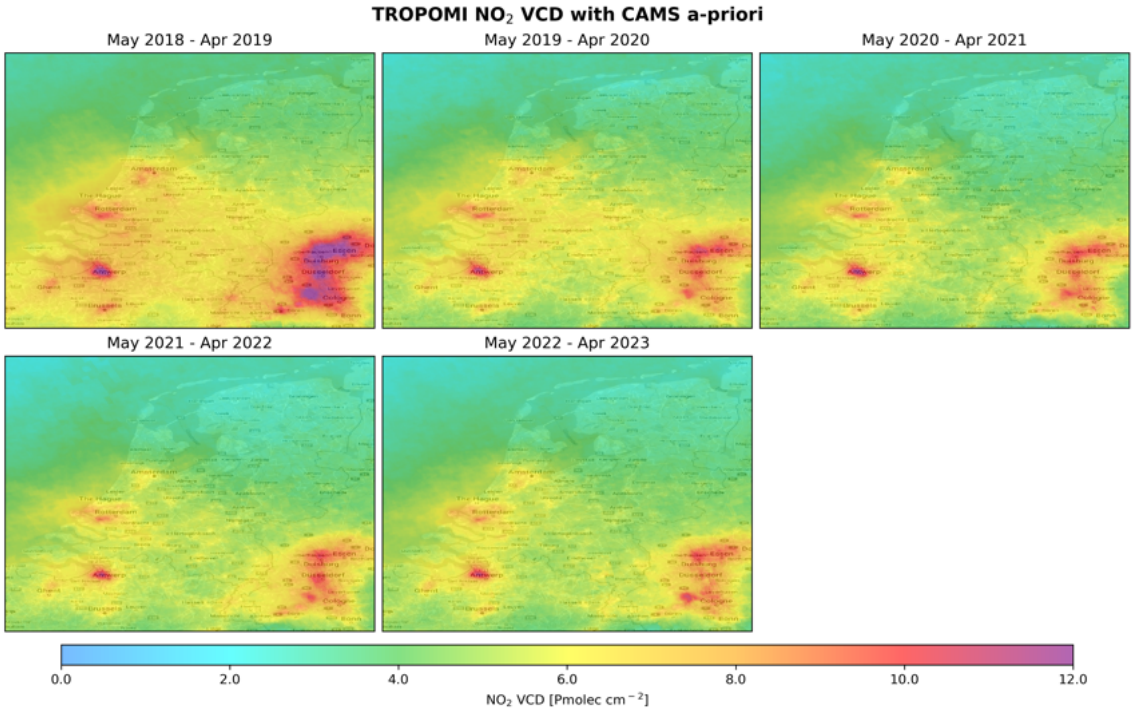


Figure 26: Yearly averaged TROPOMI NO_2 tropospheric vertical column measurements over the Netherlands, based on the v2.4 and v2.5 retrievals from May 2018 to May 2023.

ments at 29 stations revealed a mean bias of -28%, escalating to -40% over heavily polluted regions (Chan et al., 2020; Lambert & et al., 2024; Verhoelst et al., 2021).

The TROPOMI NO_2 retrieval developments in versions 2.x have reduced the negative bias compared to ground remote sensing data, but differences remain. These remaining discrepancies in TROPOMI observations can be partially attributed (about half) to the a-priori vertical NO_2 profile shape, and can be mitigated by updating the a-priori with one derived from a higher-resolution air quality model (Douros et al., 2023; Griffin et al., 2019; Judd et al., 2020; Zhao et al., 2020). This procedure is implemented using the TROPOMI averaging kernels, as explained in the TROPOMI NO_2 Product User Manual (Eskes et al., 2024).

This study used the TROPOMI NO_2 version 2.4.0 reprocessed product in combination with later offline products. To enhance the data reliability, pixels with a quality assurance value below 0.75 were excluded, effectively removing pixels with cloud radiance fractions higher than 0.5 and minimizing the impact of uncertain retrievals (Eskes et al., 2024; J. van Geffen et al., 2022; J. H. G. M. van Geffen et al., 2024). For Europe, it means that about 45% of the observations are used.

3.3 Satellite validation results

3.3.1 Satellite evaluation using FTIR-NH3

The resulting comparison between satellite and FTIR products varies quite strongly from site to site and from satellite to satellite. The results of each of the individual stations for two of the sensors, CrIS-2 and IASI-A are shown in Figure 27 and Figure 28. Correlations for each of the individual sites are quite high, especially for the sites with higher total column concentrations (Bremen, Hefei, Toronto, Boulder). On the weaker side are the sites with lower total column concentrations, Wollongong, Lauder and St. Petersburg, with each of these showing strong variations in the satellite column while the FTIR columns barely vary. The comparisons at each site indicate a high bias for the CrIS product for the low range of the concentrations while IASI varies closer around zero. Outside of these stations, the results for Mexico City also show an especially strong variation, which is also reflected in the higher FTIR error metrics for this site. While the tight co-location criteria help the comparison on a whole, some of the individual sites end up with a small number of matches, which is especially true for the Bremen, Rikubetsu and Tsukuba (not shown in Figure 28) sites for CrIS-2, with most observation having a mismatch in the sampling time.

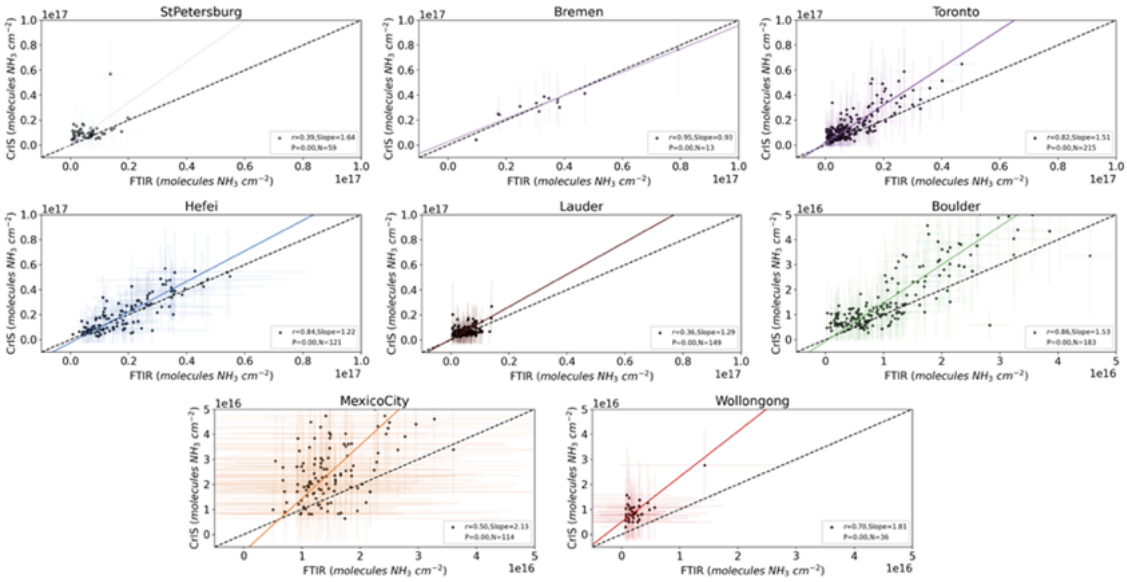


Figure 27: FTIR to CrIS-2 comparison showing the each of the individual stations. The error bars indicate the combined error of each observation combination. The colored fit lines indicate the results of the individual regression analysis. Note that the horizontal and vertical ranges are different between the stations.

Figure 29 and Table 12 provide the results for the individual satellite products when all site data are combined into a single comparison. When batched together, the results are similar as Figure 27 and Figure 28, with the CrIS satellite products having a slight high bias and/or offset, especially for the low range of concentrations. For total columns in the higher range, the comparisons are more favorable but an offset in the CrIS columns remains visible. This offset is also visible when the observations are batched by monthly total, as illustrated in Figure 30 for CrIS-1. For almost all locations, outside of Hefei, an almost consistent offset of around 5×10^{15} molecules cm^{-2} is observed. The lower correlation of CrIS-1 is strongly influenced by the results of the Mexico City site for which especially the summer observations show very large differences. Most of the locations have a larger difference between CrIS and the FTIR observations in wintertime. This coincides with lower total column concentrations, due to lower emissions, as well as less favorable measurement conditions, with lower thermal contrast especially being a factor. The IASI results show similarly high correlations in the range of 0.75 to 0.78, with a negative intercept consistent

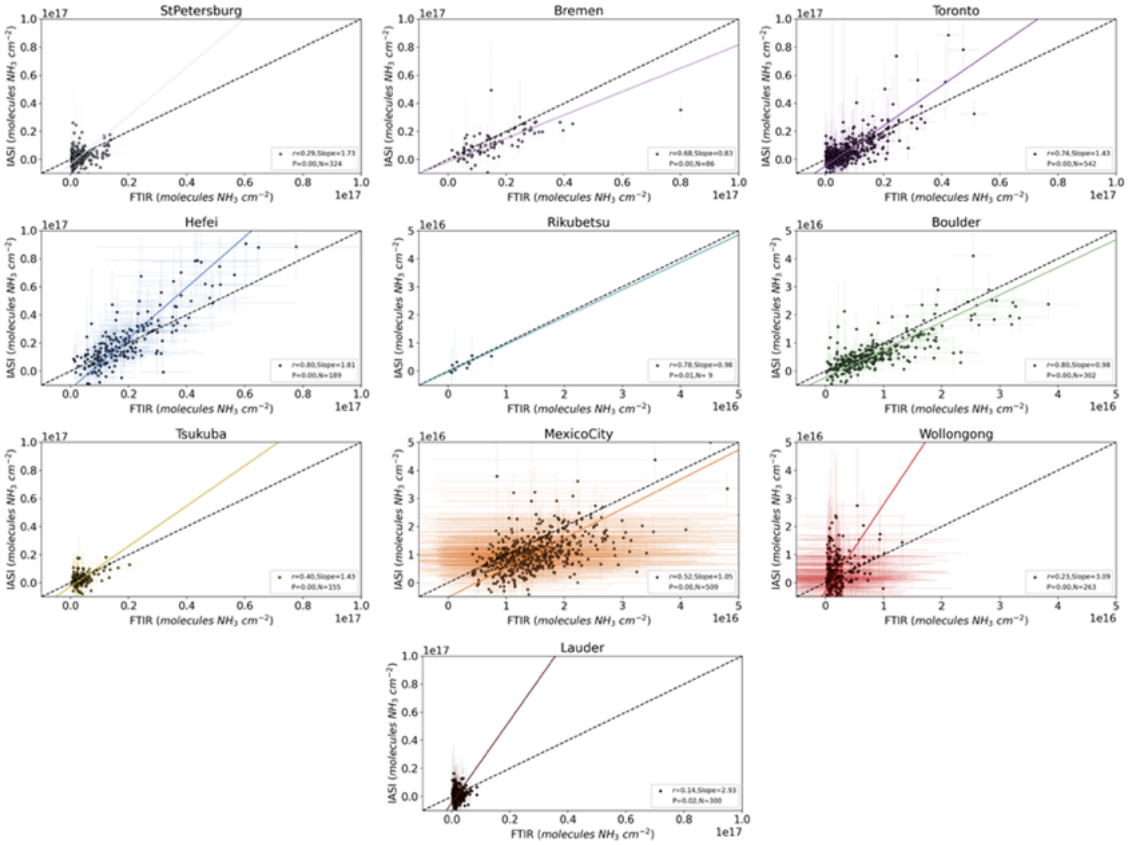


Figure 28: FTIR to IASI-A comparison showing the each of the individual stations. The error bars indicate the combined error of each observation combination. The colored fit lines indicate the results of the individual regression analysis. Note that the horizontal and vertical ranges are different between the stations.

among the instruments (-3.55 to -4.32×10^{15} molecules cm^{-2}) and a slope of 1.25 (IASI-A) to 1.35 (IASI-C). The high slope mostly seems the result of a few of the observations from the Hefei site which is the only site for IASI showing a clear slope above 1.0. This is confirmed when the observations are clustered by month as illustrated in Figure 31. Almost consistently the IASI observations, on average, show a slight negative offset with the exception of the Hefei site. Another anomaly seems the results for the Lauder site, where both CrIS and IASI show an almost counterphase in the total column concentrations. Surprisingly both satellite instruments (CrIS, Figure 30 and IASI, Figure 31) show a peak in the southern hemisphere wintertime, which requires further investigation but is outside the scope of the NKS project.

Satellite	Period	Number	Correlation	Slope	Intercept
CrIS-1	2013 – 2021	871	0.60	1.11	2.08×10^{15}
CrIS-2	2019 – 2023	889	0.81	1.35	1.40×10^{15}
IASI-A	2008 – 2022	2676	0.75	1.25	-3.55×10^{15}
IASI-B	2013 – 2023	2755	0.77	1.32	-4.10×10^{15}
IASI-C	2019 – 2023	1217	0.78	1.35	-4.32×10^{15}

Table 12: Statistics of FTIR to satellite product comparisons.

The varying results can further be investigated by calculating the mean absolute difference (MAD)

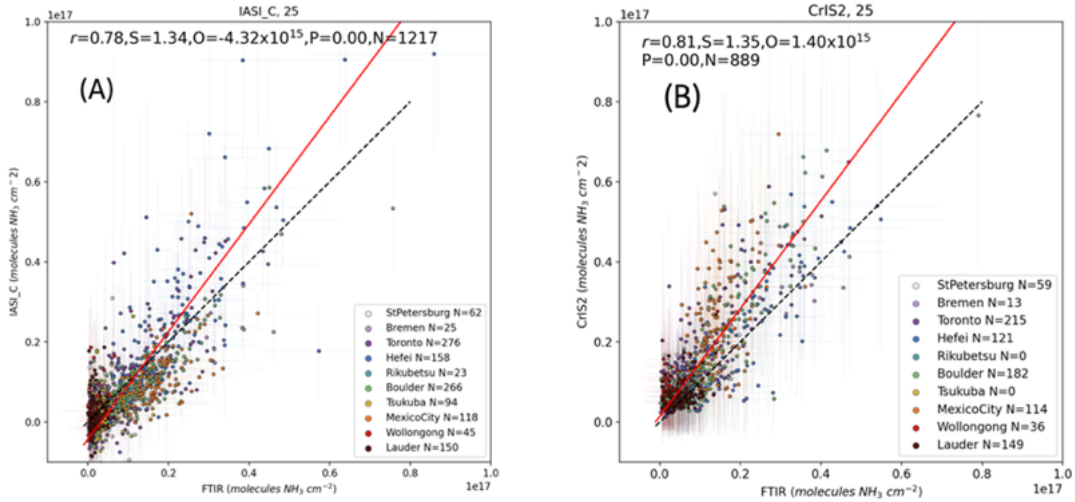


Figure 29: FTIR to IASI-C (A) and CrIS-2 (B) total column comparisons. The coloring of the dots indicate the different measurement sites. The error bars indicate the combined error of each observation. The fit lines indicate the results of the regression analysis.

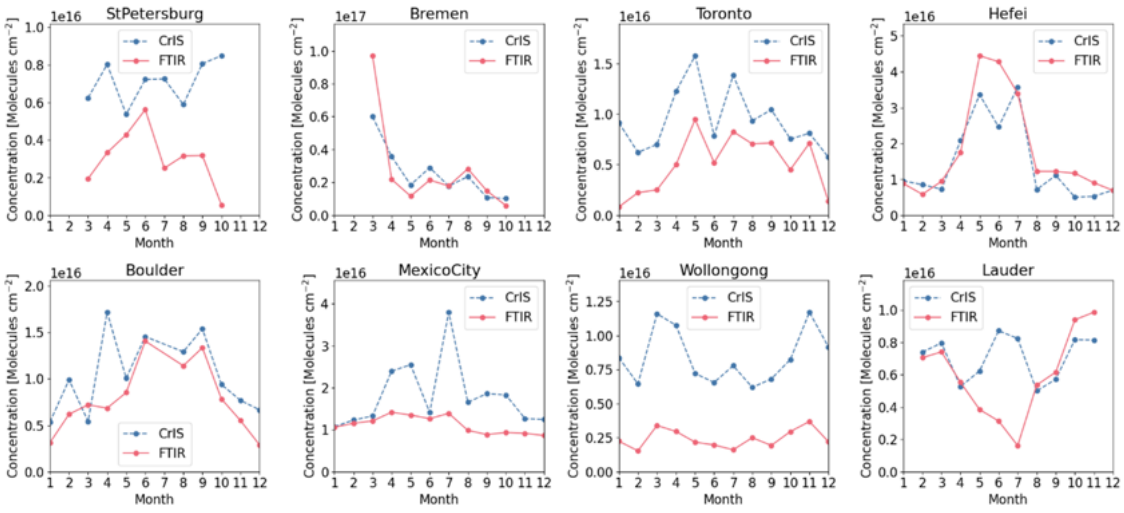


Figure 30: NH_3 monthly averages for CrIS-1 and FTIR datasets for the 8 sites with enough observations in both the CrIS and IASI datasets.

and mean relative differences (MRD). The results are clustered for IASI-A and CrIS by both sites and more explicitly by concentration ranges in Figure 32 and Figure 33. Figure 32 shows the results when clustered by site, and ordered by the mean measured FTIR concentration at each site. Even though the results are not directly clustered by concentration range, a clear pattern emerges, with the sites with lower concentrations showing a distribution of the absolute differences around zero for IASI-A and around 5×10^{15} molecules cm^{-2} for CrIS-1, while the high concentration range showing an underestimation that increases with column concentration for IASI-A while CrIS shows a closer distribution around zero. In relative terms this results in a difference of on average about 25% (Bremen, Mexico City and Boulder). The outlier in the IASI set is the results of the Hefei site,

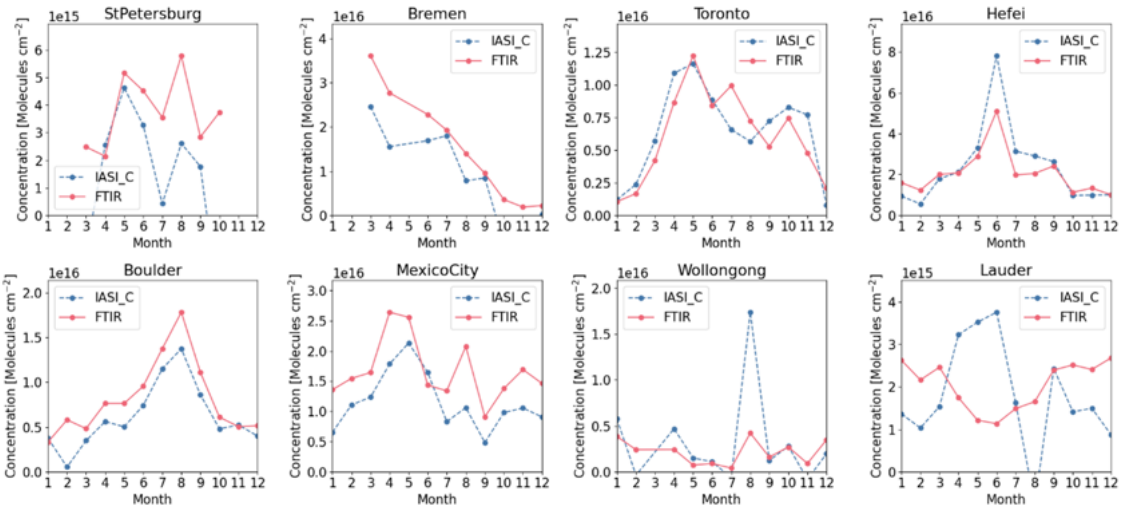


Figure 31: NH_3 monthly averages for IASI-C and FTIR datasets for the 8 sites with enough observations in both the CrIS and IASI datasets.

which show a clear high bias. For CrIS this seems the case for the site at Bremen, but there it is simply a result of very few collocated observations in the high concentration range.

3.3.1.1 Conclusions and outlook

To summarize the satellite instruments, IASI-A, -B and -C and CrIS-1 and CrIS-2 have been evaluated with FTIR observations from nine sites spread around the globe. On a site to site basis the satellites compare well, with especially the measurement sites in the higher concentration range showing favorable results with high correlations. Merging the sites together into a single comparison also showed high correlations for each of the sensors with slopes generally above 1 in the $1.25 - 1.35$ range. The exception was CrIS-1 due to a strong spread in the column concentrations around Mexico City. Intercepts ranged from positive for CrIS-1 and CrIS-2 ($1.5 - 2.0 \times 10^{15}$ molecules cm^{-2}) to negative for the IASI instruments (consistently around -4×10^{15} molecules cm^{-2}. All instruments show absolute and relative differences that strongly vary with the mean concentration around each measurement site. When split into concentration regimes the results reveals a clear pattern, with CrIS having a positive offset in the lower concentration range while showing a consistent results in the higher concentration range. The IASI instruments meanwhile show a consistent negative mean relative difference of around 25% for the higher concentration range, while on average the instruments match the FTIR columns for the lower concentration range.

The results of this project are quite consistent with previous studies (Dammers et al., 2017; Herrera et al., 2022; Tournadre et al., 2020), even though there have been several updates to both the IASI and CrIS satellite products. Some improvements are observed in case of the IASI product, where previously relative differences increased up to 50% it is now limited to 25% whereas its removed completely for lower concentration regimes. While in the previous FTIR-CrIS validation study by Dammers et al. (2017) the number of co-locations were somewhat limited, a much broader set of observations is now available. The earlier validation study indicated that for total columns greater than 10×10^{15} molecules cm^{-2}, CrIS-FPR values have a small 0%–5% difference with a standard deviation (SD) of 25% up to 50%, which is consistent with the results found here, and within range of the retrieval uncertainties (Dammers et al., 2015; Dammers et al., 2016, 2017). For the lower concentration range (10×10^{15} molecules cm^{-2}), the difference is larger, with past results showing

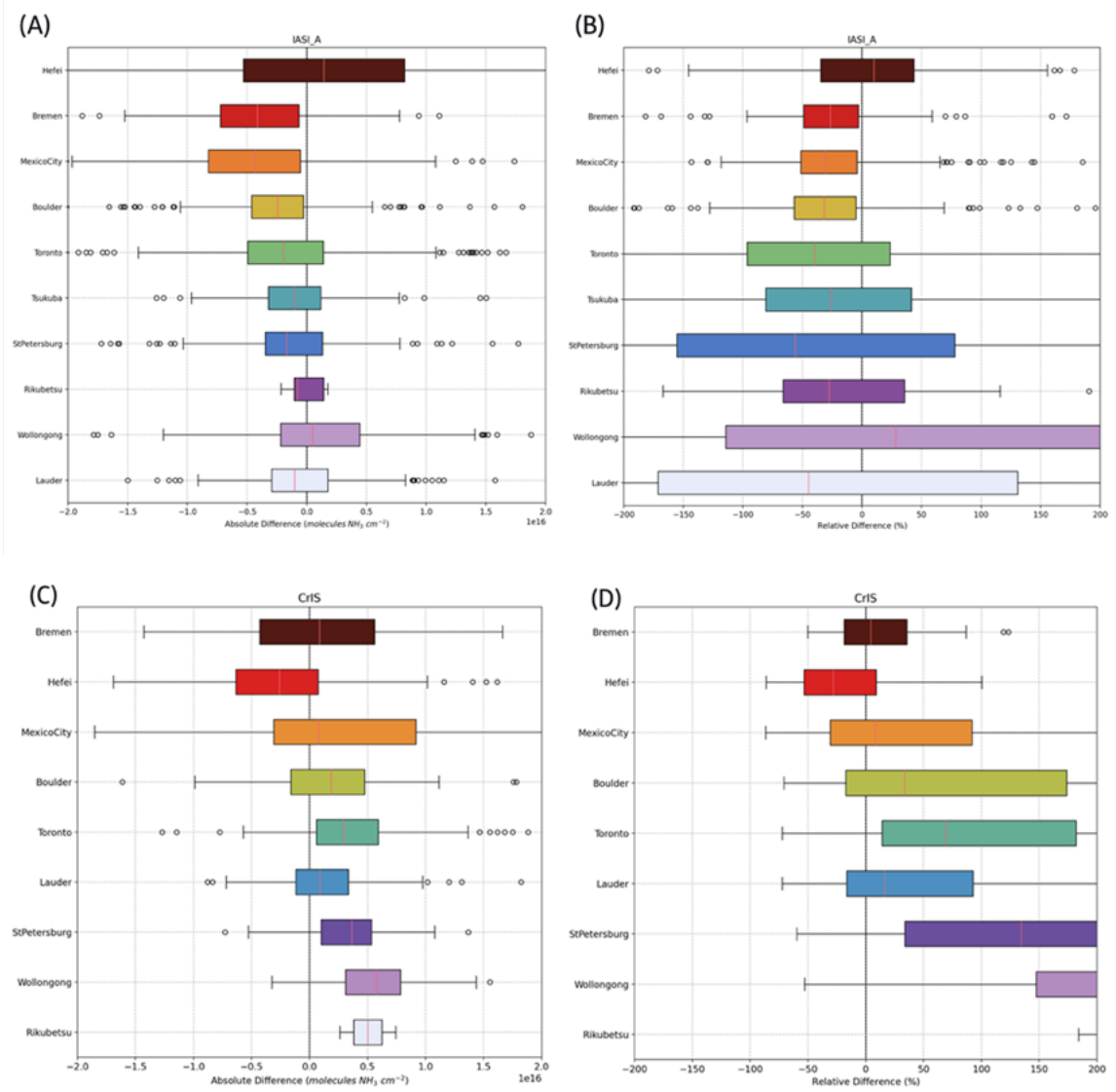


Figure 32: Absolute and relative difference between satellites IASI-A, CrIS-1 and FTIR total columns for each of the sites. Subplot (A), shows the absolute difference defined as IASI-FTIR. Subplot (B) shows the relative difference defined as (IASI-FTIR)/FTIR. Subplot (C), shows the absolute difference defined as CrIS-FTIR. Subplot (D) shows the relative difference defined as (CrIS-FTIR)/FTIR.

a positive bias of 2.4×10^{15} which now increased to about 5×10^{15} molecules cm^{-2}, equivalent to a relative difference of around 50% ($\pm$100%). A potential cause for the high bias can be found in the log-domain of the retrieval, which restricts retrievals to positive values. Furthermore, observations that fail to converge are not in the dataset at all. The introduction of the non-detect observations to the dataset (White et al., 2023) is meant to counteract this shortage and is expected to reduce the high bias in the lower concentration range. In this study however, the non-detects were removed from the dataset, as these rarely occur in the Netherlands. The inclusion of such observations would have reduced the representativeness of the results for the Netherlands.

A point of concern is the validity of the results for the typical Dutch conditions. Earlier studies as well as the results in this NKS report indicate that the typical column concentration range in the Netherlands is between $10 - 40 \times 10^{15}$ molecules cm^{-2}, while concentrations at most of the FTIR

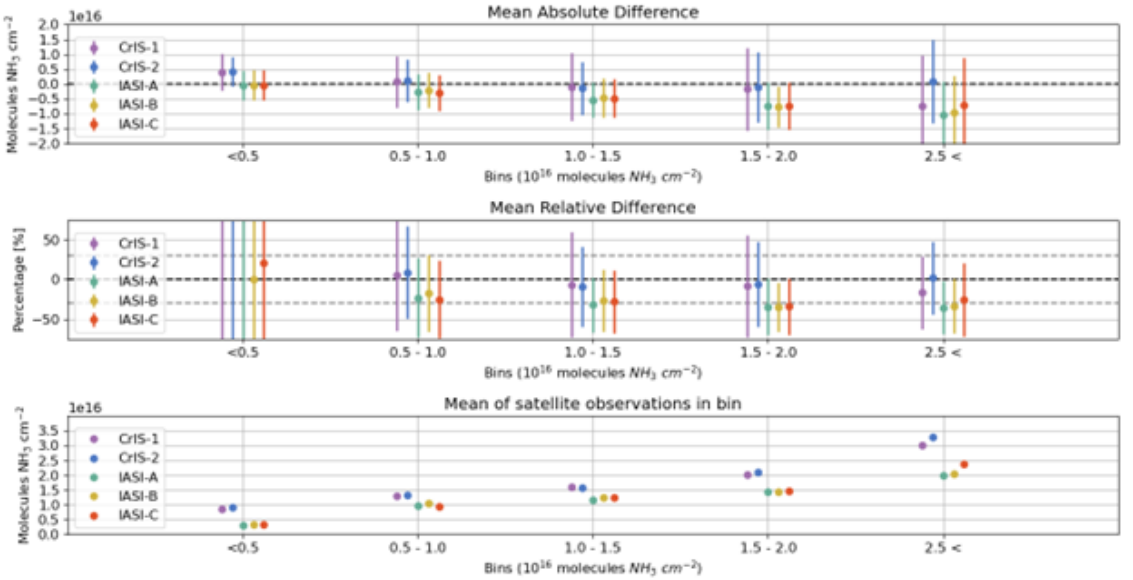


Figure 33: The mean absolute and relative differences between the CrIS, IASI and FTIR instruments as a function of FTIR NH_3 total column. The observations are separated into bins of total columns. Panel (a) shows the mean absolute difference (MAD, satellite-FTIR). Panel (b) shows the mean relative difference (MRD, (satellite-FTIR)/FTIR). The bars in these top two panels show the standard deviation of the values in the bin. Panel (c) shows the mean of the satellite observations in each bin.

sites outside of Hefei are on average much lower. This is also visible in the scatter plots (Figure 29) where the majority of the points are below 20×10^{15} and Figure 33 where the largest concentration bin is open ended but still ends up with the lowest number of observations. To resolve this point of concern an effort should be made to increase the number of satellite-FTIR co-locations in this concentration range. There are several options but most evident case is to deploy one or several FTIR instruments in the Netherlands for a field-campaign (or a dedicated FTIR site) under typical Dutch concentration regimes.

3.3.2 TROPOMI-NO2 validation

3.3.2.1 Methods

TROPOMI pixels with quality assurance values below 0.75 were excluded from the analysis to effectively filter out pixels with cloud radiance fractions above 0.5, minimizing the impact of uncertain retrievals linked to clouds (J. van Geffen et al., 2022). For validation, we used both the NO_2 tropospheric column and the NO_2 summed total column products. The summed total column, which represents the sum of the stratospheric and tropospheric vertical column densities, is preferred over the standard NO_2 total column, as it decreases the dependence on prior stratospheric and tropospheric assumptions, helping to mitigate potential systematic retrieval errors (Ialongo et al., 2020; J. van Geffen et al., 2022).

PANDORA total NO_2 column measurements were filtered by retaining only data with quality flags 0, 1, 10, and 11, which correspond to assured and not-assured high- or medium-quality data. Similarly, MAX-DOAS tropospheric NO_2 observations containing fill values were excluded from the dataset.

Spatially, TROPOMI measurements are matched to the corresponding PANDORA and MAX-DOAS

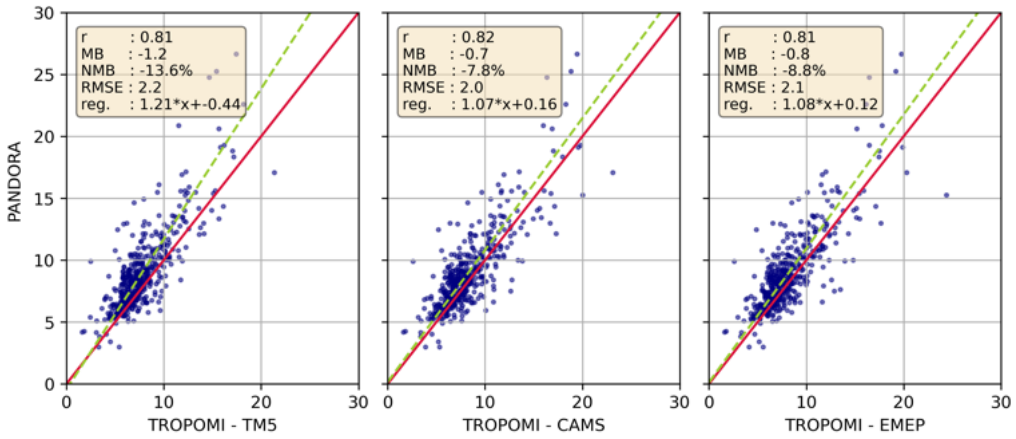


Figure 34: Scatter plot of NO_2 total vertical column densities between Pandora observations and different TROPOMI products at the Cabauw tower location. Left panel: comparison with the original summed column TROPOMI product using TM5-MP a-priori profiles; middle panel: using a-priori profiles from CAMS-Europe with about 10 km resolution; right: using EMEP4NL profiles with about 2 km resolution. Unit: Peta molecules cm^{-2}.

instruments if and when a TROPOMI footprint area with valid values includes the coordinates of the ground-based equipment. Temporally, MAX-DOAS and PANDORA observations are averaged over a 1-hour window (± 30 minutes) centered around the TROPOMI satellite overpass time.

We investigated the impact of using high-resolution NO_2 a-priori profiles instead of the standard TM5-MP profiles. To do this, we incorporated two additional products: one using CAMS regional European a-priori profiles with a spatial resolution of 0.1 x 0.1 degrees (Douros et al., 2023), which are publicly available at https://www.temis.nl/airpollution/no2_cams.php, and another using the EMEP4NL simulations with an approximate 2 x 2 km horizontal resolution. The latter product was generated as part of the NKS-SAGEN project based on the long-term simulations previously described.

For a more accurate comparison between the standard TM5-MP TROPOMI product and the MAX-DOAS tropospheric observations, we also included a MAX-DOAS product smoothed with TROPOMI averaging kernels. This adjustment ensures that the relative comparisons are independent of the a-priori profile assumptions.

3.3.2.2 Validation of NO_2 total vertical columns (Pandora vs TROPOMI)

Figure 34 illustrates a strong correlation (R > 0.8) between TROPOMI total NO_2 VCD and measurements from the PANDORA instrument located in Cabauw. However, TROPOMI shows a negative bias of approximately 14% and a Pandora-TROPOMI slope of 1.21. when using the default TM5-MP a priori data. This bias decreases to 8% when the CAMS a-priori data is used instead, showing the importance of a better resolution a-priori profile.

As shown in Figure 35, TROPOMI on average shows lower columns than PANDORA, and this persists throughout the year, with particularly large biases observed in November, December, and January when using the default TM5-MP product. The alternative versions, which use higher-resolution a priori data from the CAMS European air quality models, help narrow the gap between TROPOMI and PANDORA.

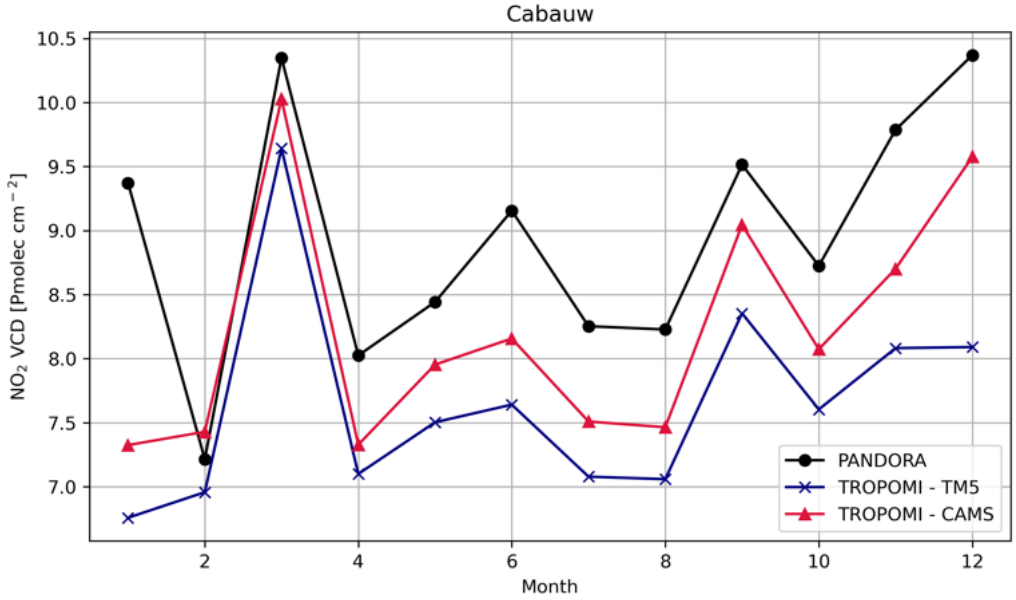


Figure 35: Monthly averages of NO_2 total vertical column densities for PANDORA and TROPOMI observations (blue) at Cabauw location. Replacing the a-priori by the CAMS European forecasts (red curve) brings the TROPOMI columns closer to the observed ones (black). Unit: Peta molecules cm^{-2}.

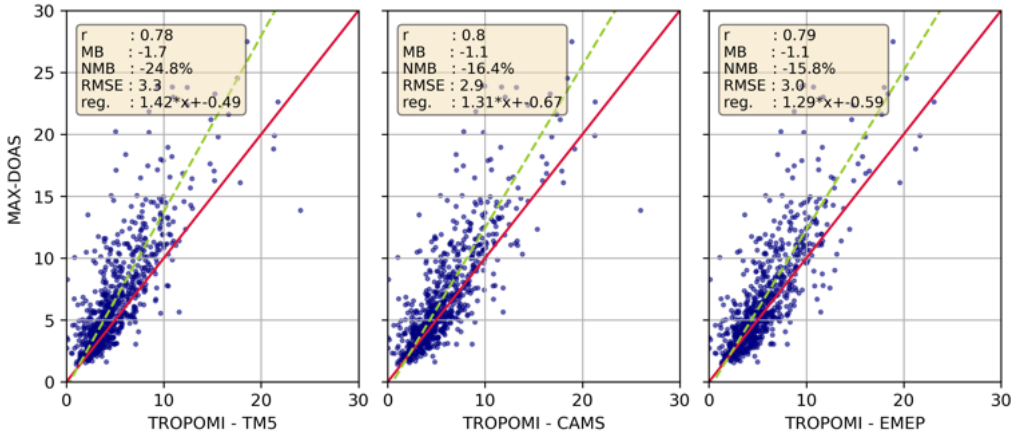


Figure 36: Scatter plot of NO_2 tropospheric vertical column densities between MAX-DOAS observations and different TROPOMI products at De Bilt location. Left panel: comparison with the original tropospheric column TROPOMI product using TM5-MP a-priori profiles; middle panel: using a-priori profiles from CAMS-Europe with about 10 km resolution; right: using EMEP4NL profiles with about 2 km resolution. Unit: Peta molecules cm^{-2}.

3.3.2.3 Validation of NO_2 tropospheric columns (MAX-DOAS vs TROPOMI)

TROPOMI tropospheric NO_2 observations exhibit a moderate correlation with MAX-DOAS measurements ($R \approx 0.7$). The default TM5 product displays significant lower values compared to MAX-DOAS, approximately 25% at De Bilt and 22% at Cabauw (see Figures 36 and 37). Replacing the default a priori with CAMS reduces these differences to about 16% at De Bilt and 14% at Cabauw. The high-resolution a-priori profile provides a more precise representation of the local-scale variability in the meteorological and chemical fields, improving the capacity to capture the NO_2 spatial and temporal variability (Laughner et al., 2016).

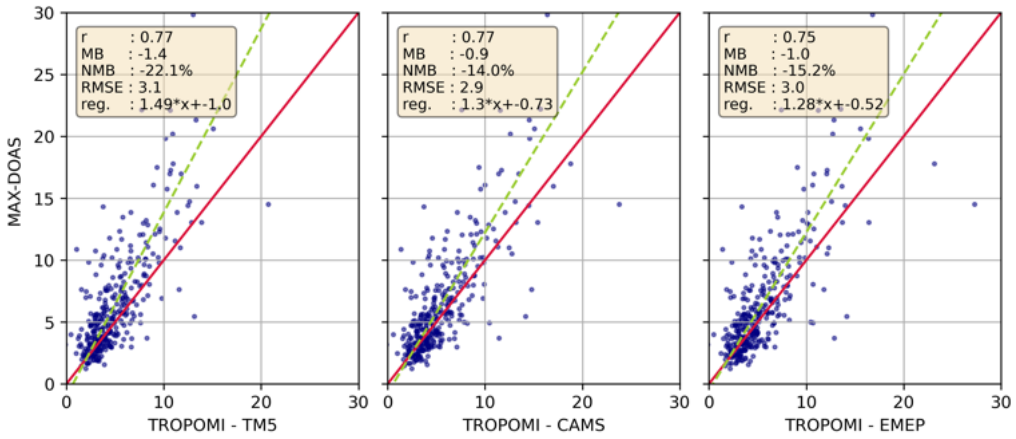


Figure 37: Scatter plot of NO_2 tropospheric vertical column densities between MAX-DOAS observations and different TROPOMI products at Cabauw location. Left panel: comparison with the original tropospheric column TROPOMI product using TM5-MP a-priori profiles; middle panel: using a-priori profiles from CAMS-Europe with about 10 km resolution; right: using EMEP4NL profiles with about 2 km resolution. Unit: Peta molecules cm^{-2}.

Figures 38 and 39 present the monthly average tropospheric NO_2. From April to August, all TROPOMI products align closely with MAX-DOAS measurements, demonstrating strong agreement. However, in the remaining months, significant differences are evident across all TROPOMI retrievals, regardless of the a priori profile used. The increasing differences between the two instruments can be linked to several factors. First, the vertical sensitivity of MAX-DOAS and TROPOMI measurements varies significantly, and more so in winter. MAX-DOAS is more sensitive to the lower troposphere, whereas TROPOMI has limited sensitivity in this region. During winter months, when the PBL is shallow, NO_2 tends to be concentrated in the lower troposphere, allowing MAX-DOAS to capture these concentrations with a higher sensitivity than TROPOMI. Additionally, biases in the AMF calculations in TROPOMI observations could contribute to the discrepancies. Another possible factor is the partitioning of stratospheric and tropospheric NO_2 columns in TROPOMI retrievals, which may introduce biases. Further investigation is ongoing to identify the primary cause of these differences.

3.3.2.4 Conclusions: TROPOMI NO_2 comparison with PANDORA and MAX-DOAS

Overall, both PANDORA and MAX-DOAS show higher columns than the TROPOMI retrieval.

TROPOMI NO_2 total column measurements at the Cabauw location show a strong correlation with Pandora observations (R > 0.8), with on average 14% lower column values that are roughly constant throughout the year. The error reduces to about 7% when high-resolution a-priori data is used from CAMS or EMEP. The slope is 0.97, very close to 1 for this station.

TROPOMI tropospheric NO_2 columns exhibit moderate correlations with MAX-DOAS observations at both validation sites in the Netherlands ($R \approx 0.7$). Between April and August, the mean values from the TROPOMI and MAX-DOAS instruments are closely aligned, while discrepancies are larger during the months of November, December, and January, reaching up to 25%, with TROPOMI lower than the ground-based observations. The increased biases observed during the winter months may be linked to differences in the sensitivity of the instruments to NO_2 vertical air masses, as well as potential inaccuracies in the AMF calculations for both MAX-DOAS and TROPOMI, or the partitioning between stratospheric and tropospheric NO_2 in TROPOMI retrieval.

Replacing the TROPOMI a priori TM5 profile with a higher-resolution model, such as CAMS or EMEP, significantly improves the comparison of the satellite and ground-based observations, re-

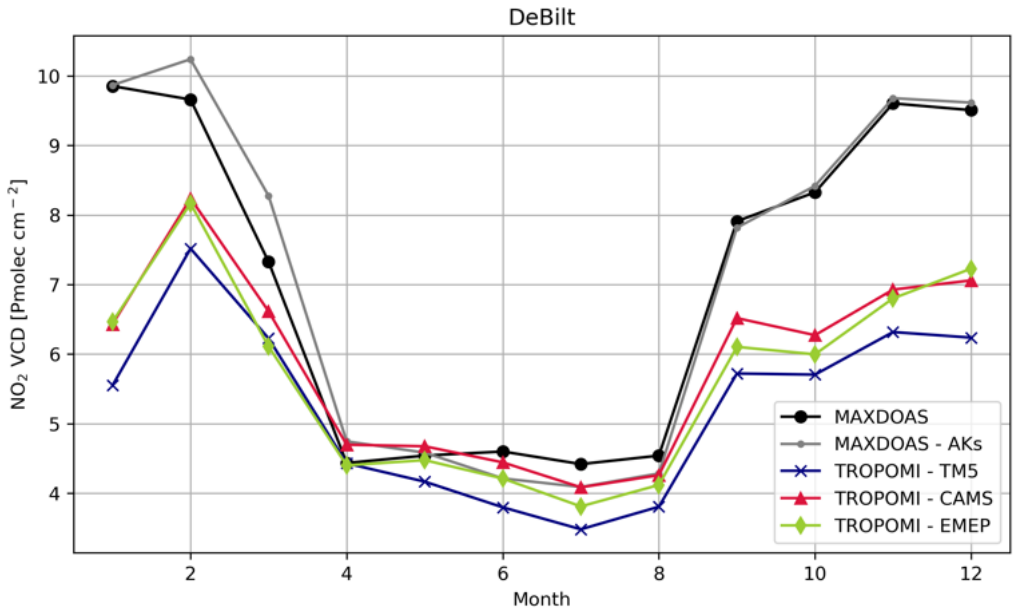


Figure 38: Monthly averages of NO_2 tropospheric vertical column densities for MAX-DOAS and TROPOMI observations at De Bilt location. Black: MAX-DOAS; Grey: MAXDOAS with TROPOMI averaging kernels applied; Blue: original TROPOMI NO_2; Red: TROPOMI with CAMS a-priori; Green: TROPOMI with EMEP a-priori. Unit: Peta molecules cm^{-2}.

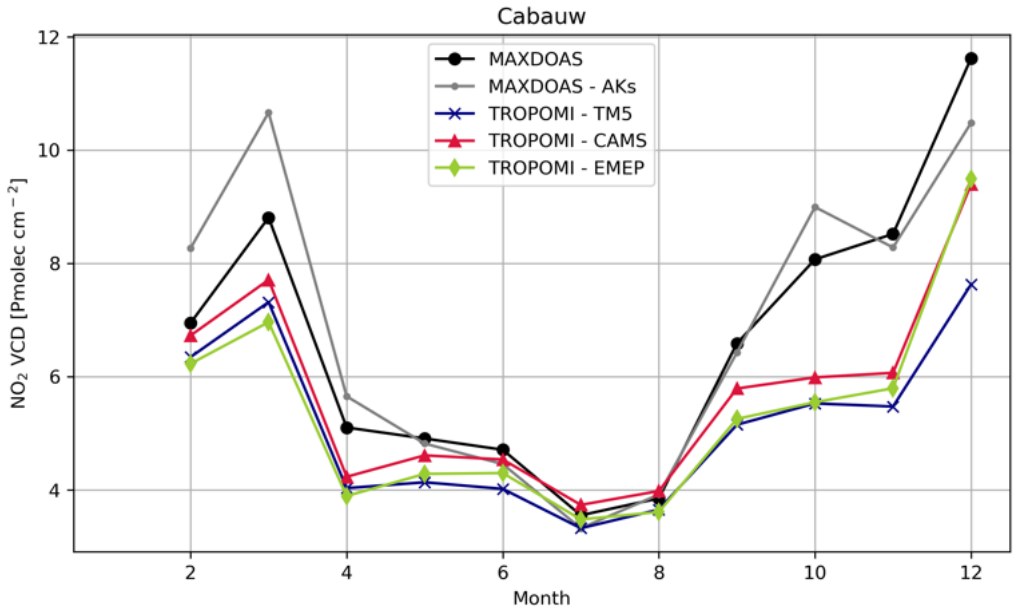


Figure 39: Monthly averages of NO_2 tropospheric vertical column densities for MAX-DOAS and TROPOMI observations at Cabauw location. Black: MAX-DOAS; Grey: MAXDOAS with TROPOMI averaging kernels applied; Blue: original TROPOMI NO_2; Red: TROPOMI with CAMS a-priori; Green: TROPOMI with EMEP a-priori. Unit: Peta molecules cm^{-2}.

ducing biases in both total and tropospheric column validations (Douros et al., 2023). This improvement stems from the higher resolution of the a priori profiles, which enable a more precise representation of the local variability of the NO_2 spatial patterns and profiles.

The sensitivity profiles, or averaging kernels are very different for the two ground-based instruments than for TROPOMI. This implies that the TROPOMI versus ground-based observations comparison is NO_2 profile shape dependent. However, the profile is not measured by these instruments (MAX-DOAS provides some limited vertical information).

The seasonal variability is significantly different for Pandora compared to MAX-DOAS, with high columns in winter for MAX-DOAS. These biases between the surface instruments needs to be understood better. We note that the Pandora direct Sun total column retrieval is characterised by much smaller air-mass factor uncertainties and may be more robust as reference for the validation of satellite retrievals. This is why we use the direct sun Pandora to report biases.

The results presented here are the starting point for a more detailed study to be conducted in 2025 (a publication is in preparation). Aspect that will be studied include:

- The detailed assumptions and inputs for the ground-based and satellite retrievals will be compared, e.g. the impact of the temperature profile.
- Different retrievals for MAX-DOAS seem to produce different results, which will be clarified.
- Filtering for clouds will be revisited.
- Gradients in NO_2 around the sites of Cabauw and De Bilt, and possible impact on the retrieval and comparison will be studied.
- The stratospheric estimate and possible biases in winter will be investigated.

This work was performed by the NKS PhD student Felipe Cifuentes Castaño and supervisors, with additional support from student Julian Gomez Tibaquira from Wageningen university. We acknowledge discussions with Ankie Piters, who is in charge of the measurements at Cabauw and De Bilt.

4 Satellite-model inter-comparisons for NO_2

4.1 Introduction

Chemical transport models (CTMs) are critical tools for understanding pollution dynamics, offering forecasts and interpretations of air quality based on emission inventories and atmospheric processes (Skoulidou et al., 2021). To ensure their reliability, these models must undergo rigorous evaluation through comparisons with observational data from both ground-based monitoring networks and satellite platforms. Satellite-derived NO_2 data, such as that provided by TROPOMI, is particularly valuable for model validation due to its near-global coverage and city-scale resolution. This capability enables CTM evaluations even in regions with sparse or nonexistent ground-based stations. Discrepancies between satellite observations and model outputs can pinpoint areas requiring model improvements, such as adjustments to emission inventories, transport dynamics, or chemical reaction rates, ultimately enhancing model accuracy and predictive capacity.

Both models and observations have uncertainties. In the comparisons in this section use is made of the averaging kernels, removing the a-priori profile shape uncertainty as source of error. Compared to both MAX-DOAS and Pandora ground instruments the TROPOMI retrieved columns are low. Taking the Pandora direct sun comparison at Cabauw as reference suggests a remaining negative bias in TROPOMI of -7% and a slope of 0.97. It is good to keep this in mind when evaluating the model-TROPOMI comparisons discussed below.

In this study, we evaluate the performance of the EMEP and LOTOS-EUROS models for forecasting high-resolution NO_2 tropospheric VCD over the Netherlands by comparing its predictions with TROPOMI observations.

4.2 Comparison of LOTOS-EUROS and TROPOMI NO_2

4.2.1 Methods

The CSO tool (described earlier in section 2.2.1) was used to extract the model concentration profiles at the locations of the TROPOMI footprints. The TROPOMI NO_2 data product was detailed in section 3.2.6. The TROPOMI NO_2 averaging kernel was then applied to the model profiles to account for the vertical sensitivity of the satellite instrument. Comparisons between the TROPOMI NO_2 product and LOTOS-EUROS were then performed over the Netherlands for the period of May 2018 to December 2022.

4.2.2 Results

The LOTOS-EUROS simulated NO_2 columns from the long simulations were compared with the total columns retrieved by TROPOMI for the period of May 2018 to December 2022. A spatial map of the mean columns LOTOS-EUROS and TROPOMI NO_2 calculated throughout the period from 1 May 2018 to 31 December 2022 is shown in Figure 40. LOTOS-EUROS captures a very similar spatial distribution of NO_2 total columns as to that as observed by TROPOMI, which is evident from the generally homogeneous distribution of the differences in the bottom right panel of Figure 40. The unsmoothed LOTOS-EUROS simulated columns show a higher bias relative to TROPOMI than the smoothed columns; however, the smoothed LOTOS-EUROS columns still display a mean high bias relative to TROPOMI on the order of 25%. Figure 41 provides a time series plot of the monthly mean NO_2 total columns retrieved from TROPOMI versus the smoothed LOTOS-EUROS columns. The model is able to capture the temporal variations and intra-annual variability in the

observed NO_2 columns well over the selected period. The relative difference between the model and the TROPOMI measurements is relatively steady throughout the time series (mean relative bias of 17.8% in the monthly means), with the notable exception of January 2021 which displays a higher mean relative bias. This mean relative bias between TROPOMI and LOTOS-EUROS falls within the general range of the biases (-7.8% to -24.8%) observed between TROPOMI and the MAX-DOAS and PANDORA instruments in the Netherlands, as discussed in Section 3.3.2. As such, the differences seen here can likely be attributed in part to the biases in the TROPOMI NO_2 retrievals themselves.

Figure 42 is a correlation plot of the daily-mean simulated and TROPOMI-measured NO_2 columns. The scatter plot shows the smoothed LOTOS-EUROS columns (with TROPOMI averaging kernels applied) versus the observed TROPOMI daily mean. The smoothing of the model columns by the TROPOMI averaging kernels leads to a significant improvement in the agreement between the model and the measurements (un-smoothed comparison not shown), with a reduction in the mean bias from 71% to 23.8%, and a slight improvement in the Pearson correlation coefficients from R = 0.76 to R = 0.80. In general, LOTOS-EUROS and TROPOMI NO_2 correlate well over the period of 2018 to 2022.

4.2.3 Summary and conclusions

The comparison of LOTOS-EUROS simulated NO_2 columns with TROPOMI-retrieved total columns for the period of May 2018 to December 2022 demonstrates strong spatial and temporal agreement.

- LOTOS-EUROS captures the spatial distribution of NO_2 total columns observed by TROPOMI, with generally homogeneous differences. However, the unsmoothed LOTOS-EUROS columns exhibit a higher bias compared to TROPOMI, while the smoothed columns (with TROPOMI averaging kernels applied) reduce this bias but still show a mean high bias of approximately 25%.
- The time series of monthly mean NO_2 columns reveals that LOTOS-EUROS successfully reproduces the temporal variations and intra-annual variability observed by TROPOMI, with a relatively steady mean relative bias of 17.8%. An exception occurs in January 2021, where a higher bias is noted.
- The scatter plot of daily mean NO_2 columns shows that applying TROPOMI averaging kernels to the simulated LOTOS-EUROS columns improves the agreement with TROPOMI measurements. The mean bias reduces significantly from 71% to 23.8%, and the Pearson correlation coefficient increases from R = 0.76 to R = 0.80.

Overall, the LOTOS-EUROS model performs well in capturing both the spatial patterns and temporal variations of TROPOMI-observed NO_2 columns, with smoothing by TROPOMI averaging kernels improving the agreement significantly.

4.3 Comparison of EMEP with TROPOMI NO_2

4.3.1 Methods

To validate the EMEP model against TROPOMI observations, a series of collocation procedures were applied to ensure consistency between the datasets in temporal, horizontal, and vertical dimensions. Temporal collocation was achieved by linearly interpolating the hourly model outputs to match the TROPOMI overpass time, using the two closest simulated hours. Horizontal collocation involved interpolating the original model data, with a spatial resolution of approximately 2 × 2 km,

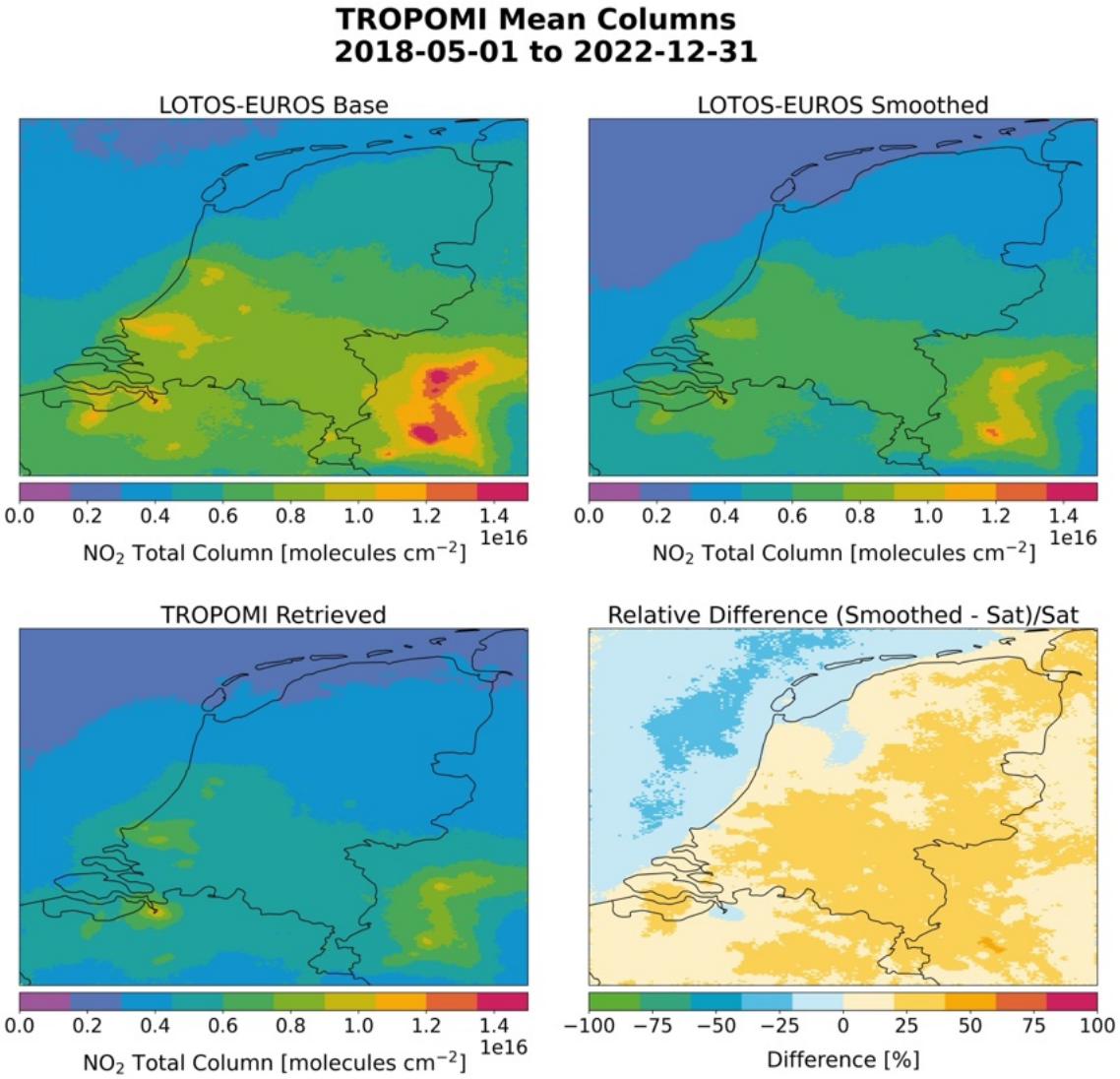


Figure 40: Spatial mean maps calculated of (top left) the base LOTOS-EUROS simulated NO_2 total columns, (top right) the LOTOS-EUROS total columns smoothed by the TROPOMI averaging kernels, (bottom left) the TROPOMI retrieved NO_2 total columns, and (bottom right) the relative difference between the smoothed model columns and the TROPOMI retrieved columns.

onto the TROPOMI footprint using a conservative interpolation method. Then, NO_2 concentrations were integrated into VCD and adjusted using TROPOMI averaging kernels, ensuring that the comparison was independent of the satellite product a-priori assumptions. To reduce uncertainties, we excluded TROPOMI pixels with quality assurance values below 0.75, effectively removing observations with cloud radiance fractions exceeding 0.5 (J. van Geffen et al., 2022). Modeling data corresponding to invalid TROPOMI pixels were also discarded to ensure consistent sampling between datasets.

Two versions of the modeling data were analyzed, one based solely on EMEP forecasted concentrations and another combining EMEP data up to 2500 meters above the surface with CAMS global reanalysis predictions above this threshold. Comparisons were conducted at multiple temporal scales, including daily single-orbit validations, monthly means, and overall averages, providing a comprehensive evaluation of model performance.

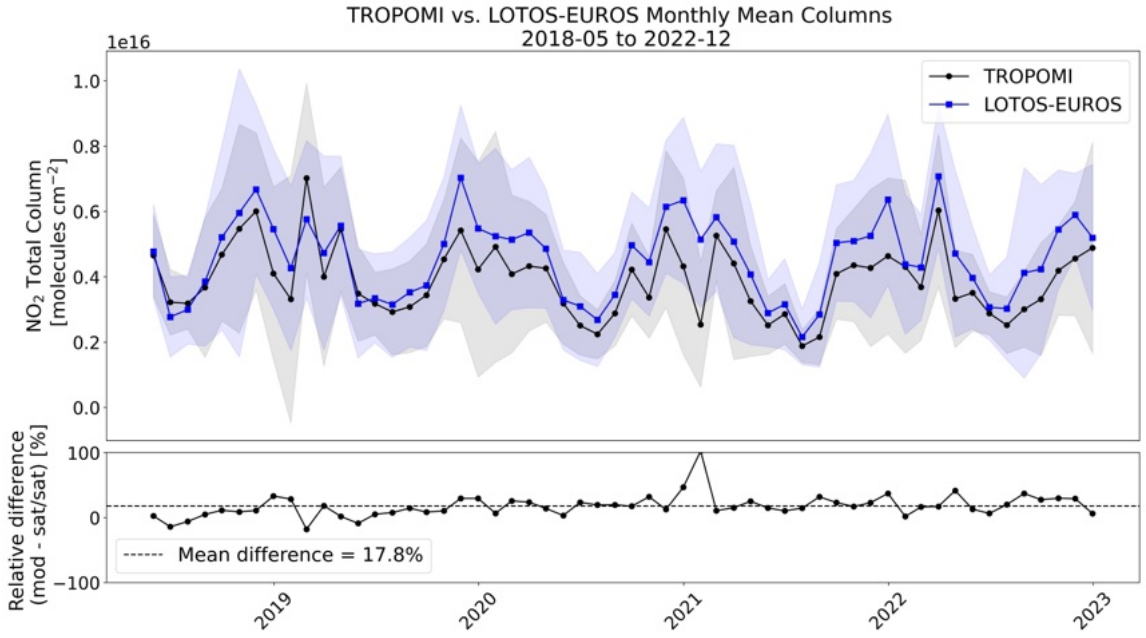


Figure 41: (Top) a time-series of the monthly mean NO_2 total columns from TROPOMI and the smoothed LOTOS-EUROS columns, and (bottom) the time-series of the relative differences.

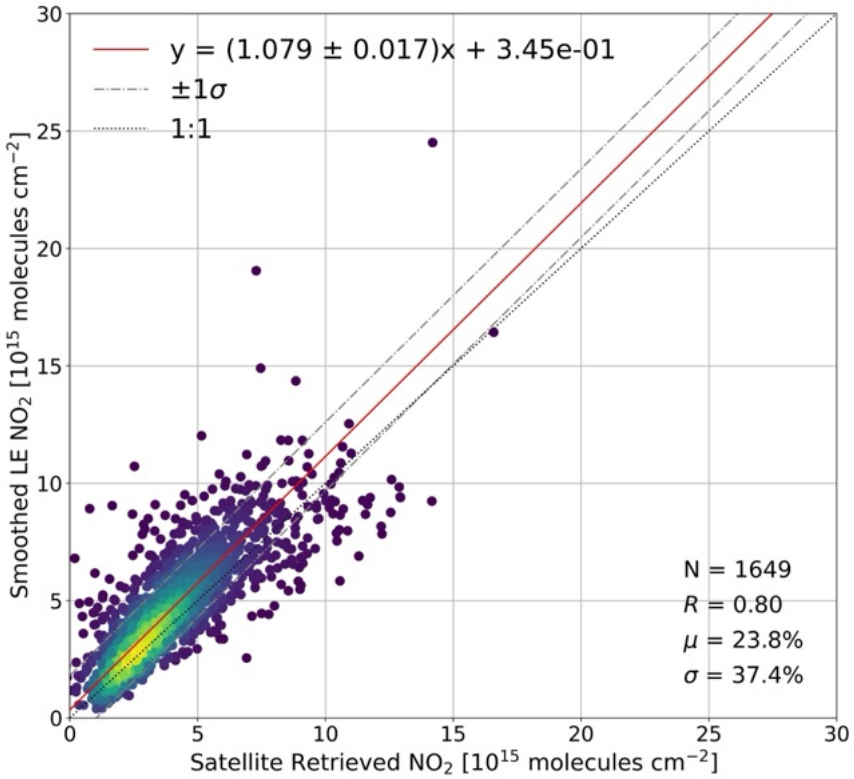


Figure 42: Correlation plot for the period of May 2018 to December 2022 of the TROPOMI retrieved NO_2 columns versus (left) daily mean base LOTOS-EUROS simulated, and (right) daily mean LOTOS-EUROS columns with the TROPOMI averaging kernel smoothing applied.

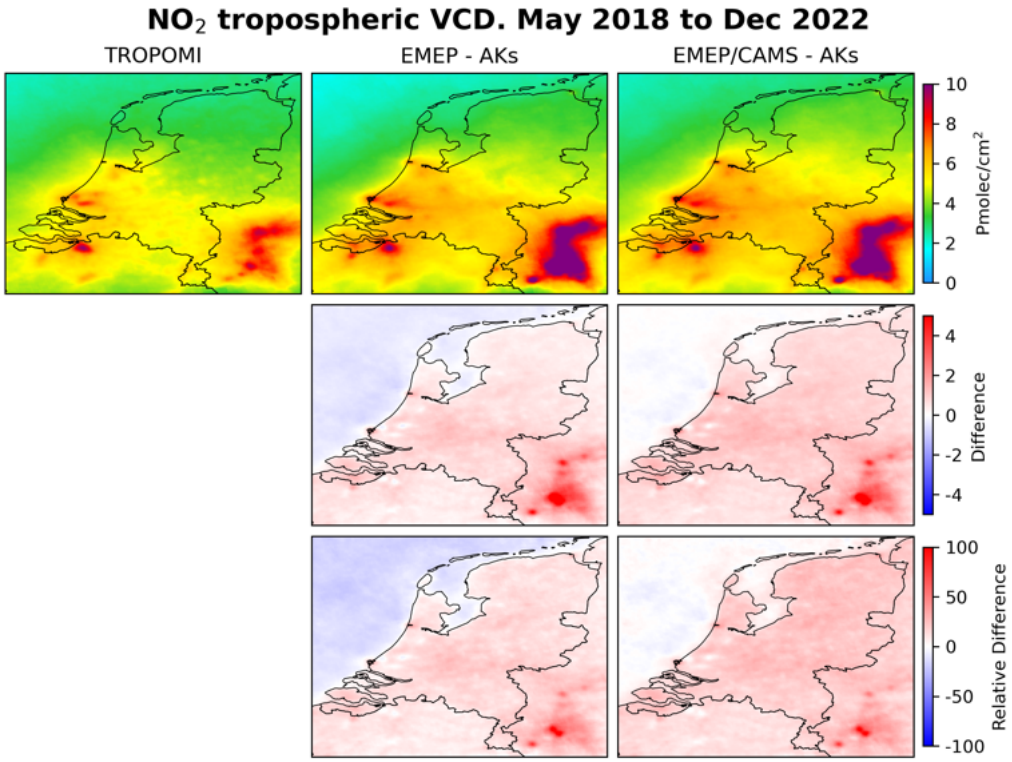


Figure 43: Mean NO_2 VCD between May 1, 2018, and December 31, 2022 according to TROPOMI (left) compared to EMEP (middle) and EMEP-CAMS (right) model simulations.

4.3.2 Results

Figure 43 illustrates the mean NO_2 VCDs from TROPOMI and model simulations for the period between May 1, 2018, and December 31, 2022. Overall, the model effectively captures the spatial distribution of NO_2, identifying key hotspots such as the Ruhr area, Antwerp, Rotterdam, and Amsterdam. However, the EMEP model tends to overestimate NO_2 concentrations over land while underestimating them over the North Sea. The latter bias is partially mitigated by replacing EMEP concentrations above 2500 meters with CAMS global simulations, which provide a better representation of background concentrations, leading to improved model performance.

As illustrated in Figure 44, there is a strong correlation (R = 0.79) between TROPOMI and EMEP. On average, EMEP concentrations are 9% lower than those of TROPOMI when relying solely on EMEP outputs. However, the combined modeling product with CAMS results in columns that are 11% higher.

Figure 45 presents a monthly time series of NO_2 concentrations for several cities in the Netherlands. The model successfully captures the general trend and seasonality throughout the entire simulation period. However, biases increase significantly during the winter months. Ongoing research involving MAX-DOAS and PANDORA is focused on identifying the underlying causes of these differences during the winter period considering both model and satellite biases.

A closer examination of the model's performance on a day-to-day basis reveals that, in general, the major NO_2 plumes are well represented in both magnitude and direction, as illustrated in Figure 46 for a combined plume from Tata Steel, Schiphol and Amsterdam, and a plume from the Rotterdam harbour area. However, on certain days, the simulated plumes appear more elongated, and their orientation may differ slightly from the TROPOMI observations, as shown in Figure 47 for plumne from the Ruhr area. These discrepancies are likely due to biases in factors such as NO_2 lifetimes,

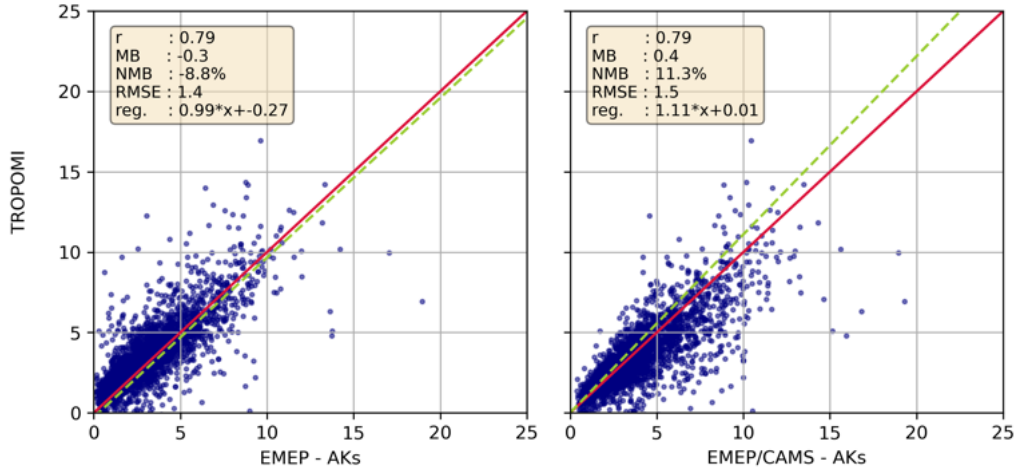


Figure 44: Scatter plot of daily mean NO_2 VCD between May 1, 2018, and December 31, 2022 between TROPOMI observations and EMEP simulations convoluted with the TROPOMI averaging kernels. Left: EMEP-only; right: EMEP-CAMS.

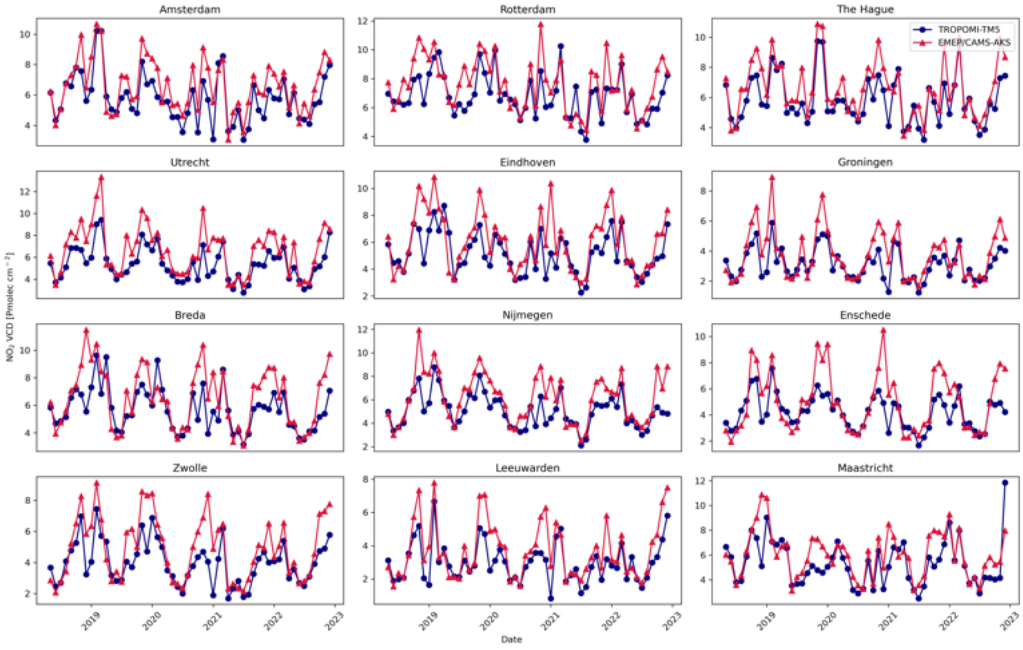


Figure 45: Time series of monthly-mean NO_2 VCD over 12 Dutch cities, between May 1, 2018, and December 31, 2022, for TROPOMI observations (blue) and EMEP simulations (red).

vertical mixing, the vertical distribution and strength of emissions, and errors in the wind fields. Despite these day-to-day variations, the model demonstrates strong overall performance when assessed over the long time series, as shown previously in Figure 43.

4.3.3 Conclusions

The EMEP simulations of NO_2 show strong agreement with TROPOMI observations, as demonstrated by a high correlation coefficient (R = 0.79) and biases based on daily averages between

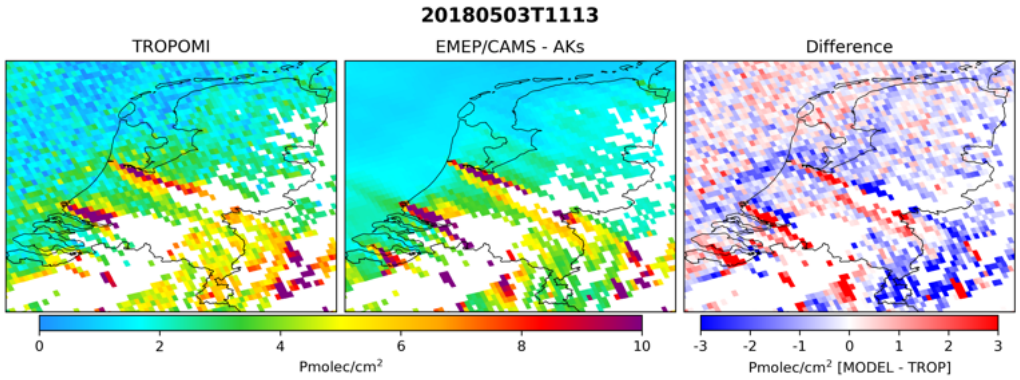


Figure 46: EMEP NO_2 column simulations for 3 May 2018 (middle) compared to the single TROPOMI overpass for the same day (left). The wind is coming from the north-west on this day. White areas indicate missing data due to clouds.

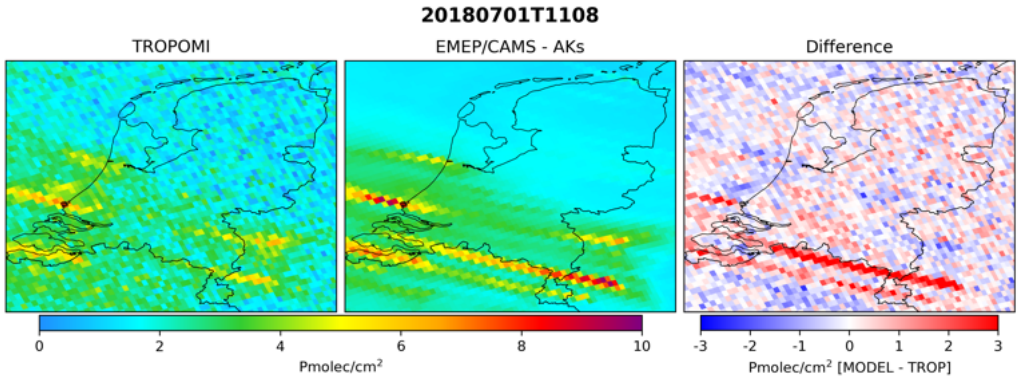


Figure 47: EMEP NO_2 column simulations for 1 July 2018 (middle) compared to the single TROPOMI overpass for the same day (left). The wind is coming from the east on this day. White areas indicate missing data due to clouds.

-9% and 11%. A day-to-day analysis indicates that while the model effectively captures pollution plumes, it occasionally simulates plumes that are longer and slightly shifted in direction compared to TROPOMI, suggesting areas for further investigation. These findings highlight the overall reliability of the EMEP model and the emissions used by the model, while also pointing to specific differences that require further study.

5 Satellite-model inter-comparisons for NH_3

5.1 Introduction

Satellite observations provide an invaluable dataset for the validation of CTMs such as LOTOS-EUROS and EMEP4NL due to their global coverage and ability to capture atmospheric NH_3 distributions over a wide range of temporal and spatial scales. Satellite instruments like IASI and CrIS offer consistent and long-term measurements of NH_3 in the atmosphere, complementing ground-based monitoring networks that often lack spatial coverage. By comparing CTM outputs with satellite observations, it is possible to assess model performance, identify spatial and temporal biases, and pinpoint areas where model assumptions, such as emission estimates or meteorological parameters, may require refinement.

In this context, inter-comparing CTMs against satellite observations serves multiple purposes. It provides a measure of the model's accuracy, helping to ensure its reliability for applications such as emission scenario analysis and regulatory policy support. Additionally, this exercise can highlight discrepancies in satellite retrievals, which are subject to uncertainties related to retrieval algorithms. The combined use of CTMs and satellite data thus facilitates a more comprehensive understanding of atmospheric NH_3 dynamics, enhancing both observational and modeling frameworks.

This section focuses on the comparison of modeled NH_3 total column concentrations from LOTOS-EUROS and EMEP4NL with satellite-derived observations to evaluate the ability of the models to reproduce observed spatial and temporal patterns and magnitudes. Such inter-comparisons are essential for broadly improving confidence in the models' predictions and identifying pathways for further adjustments.

5.2 Comparison of LOTOS-EUROS with CrIS and IASI

The NH_3 total columns from the LOTOS-EUROS long simulations were compared against the measurements from the IASI and CrIS instruments over the period of January 2014 to December 2022 for the regional domain encompassing the Netherlands. A spatial map of the mean NH_3 total columns at the time of the satellite overpasses for IASI and LOTOS-EUROS are shown in Figure 48. Although the general spatial distribution of NH_3 that is observed by the satellite is broadly captured, relative to the satellite-retrieved total columns the model displays a clear low bias. A similar map of the NH_3 columns but for CrIS is shown in Figure 49. While the model again broadly captures the spatial distribution of the NH_3 concentrations, a similar low bias in the modeled NH_3 total columns relative to the CrIS observations can be seen, however the magnitude of the bias between CrIS and LOTOS-EUROS is slightly higher than for IASI.

A correlation plot of all IASI (A, B, and C) and CrIS (1 and 2) daily means versus the corresponding averaging kernel-smoothed LOTOS-EUROS total columns is provided in Figure 50. Comparing the daily mean total columns we find that IASI is quite well correlated with LOTOS-EUROS (R = 0.74) over the full period, however the model biased low relative to the satellite observations on average by 53%, and the regression shows a slope of approximately 0.37. For CrIS, the correlation is slightly poorer with R = 0.59. A time-series of the monthly mean NH_3 total columns calculated across all IASI and CrIS measurements, as well as for the corresponding smoothed LOTOS-EUROS columns is shown in Figure 51. In the case of IASI verus LOTOS-EUROS, the time-series of monthly means and the relative differences displays some seasonality with generally smaller (even sometimes positive) biases during the winter and spring months, and larger negative differences during the summer. A mean bias of -51% in LOTOS-EUROS relative to IASI is seen in

the monthly means across the whole time-series. For the CrIS observations, it can be seen that the satellite and LOTOS-EUROS NH_3 monthly means generally capture the same inter-annual variability in the column concentrations, but LOTOS-EUROS is biased more consistently low relative to CrIS than was the case for IASI. Over the full 2014-2022 time-series a mean relative difference of 63.2% between LOTOS-EUROS and CrIS was found. As discussed in Section 3.3.1, the satellite products themselves contain systematic biases potentially as artifacts of the retrievals, and this may contribute to some extent to the observed model-measurement bias we find here. Other potential sources could be underestimated emissions and/or overestimated deposition processes in the model.

Both satellite products compare broadly well with LOTOS-EUROS in terms of the spatio-temporal representation of NH_3, but the model appears to be broadly biased low for NH_3 column concentrations relative to the satellite observations. At present, the set of NH_3 emissions used in LOTOS-EUROS only extends up to 2019, and it is possible that the inclusion of emissions for the period of 2020–2022 would further improve the agreement between the model and satellite observations.

5.2.1 Summary

The comparison of NH_3 total columns from LOTOS-EUROS long simulations with observations from the IASI and CrIS satellite instruments over the period January 2014 to December 2022 highlights the following key findings:

- LOTOS-EUROS broadly captures the spatial distribution of NH_3 observed by IASI and CrIS over the Netherlands. However, the model shows a consistent low bias relative to both satellite products. The bias is slightly larger for CrIS compared to IASI, although the overall spatial patterns remain similar.
- For IASI, the daily mean NH_3 total columns are well correlated with LOTOS-EUROS (R = 0.74), but the model exhibits a low bias of 53% and a regression slope of approximately 0.37. For CrIS, the correlation is weaker (R = 0.59), and the model displays an even larger low bias compared to CrIS observations.
- The time series of monthly mean NH_3 total columns shows that LOTOS-EUROS captures the overall inter-annual variability observed by both IASI and CrIS. However, for IASI, the relative bias displays seasonality, with smaller (sometimes positive) biases in winter and spring, and larger negative biases during summer months. The overall mean bias is -51% relative to IASI. For CrIS, the low bias is more consistent across the year, with a mean relative difference of 63.2% over the full 2014–2022 period
- Both IASI and CrIS satellite products show broad agreement with LOTOS-EUROS in terms of the spatio-temporal representation of NH_3. However, the model exhibits a significant low bias in NH_3 total column concentrations relative to the satellite observations, particularly for CrIS.

5.2.2 Outlook

Future work could focus on identifying and addressing the causes of the observed low biases in LOTOS-EUROS. Potential sources of these biases could include systematic biases from the satellite products themselves, underestimated NH_3 emissions, and/or overestimated deposition processes in the model. Some suggestions for improvements include:

- Refining the NH_3 emissions inventory, particularly for summer months when biases were found to be the largest.

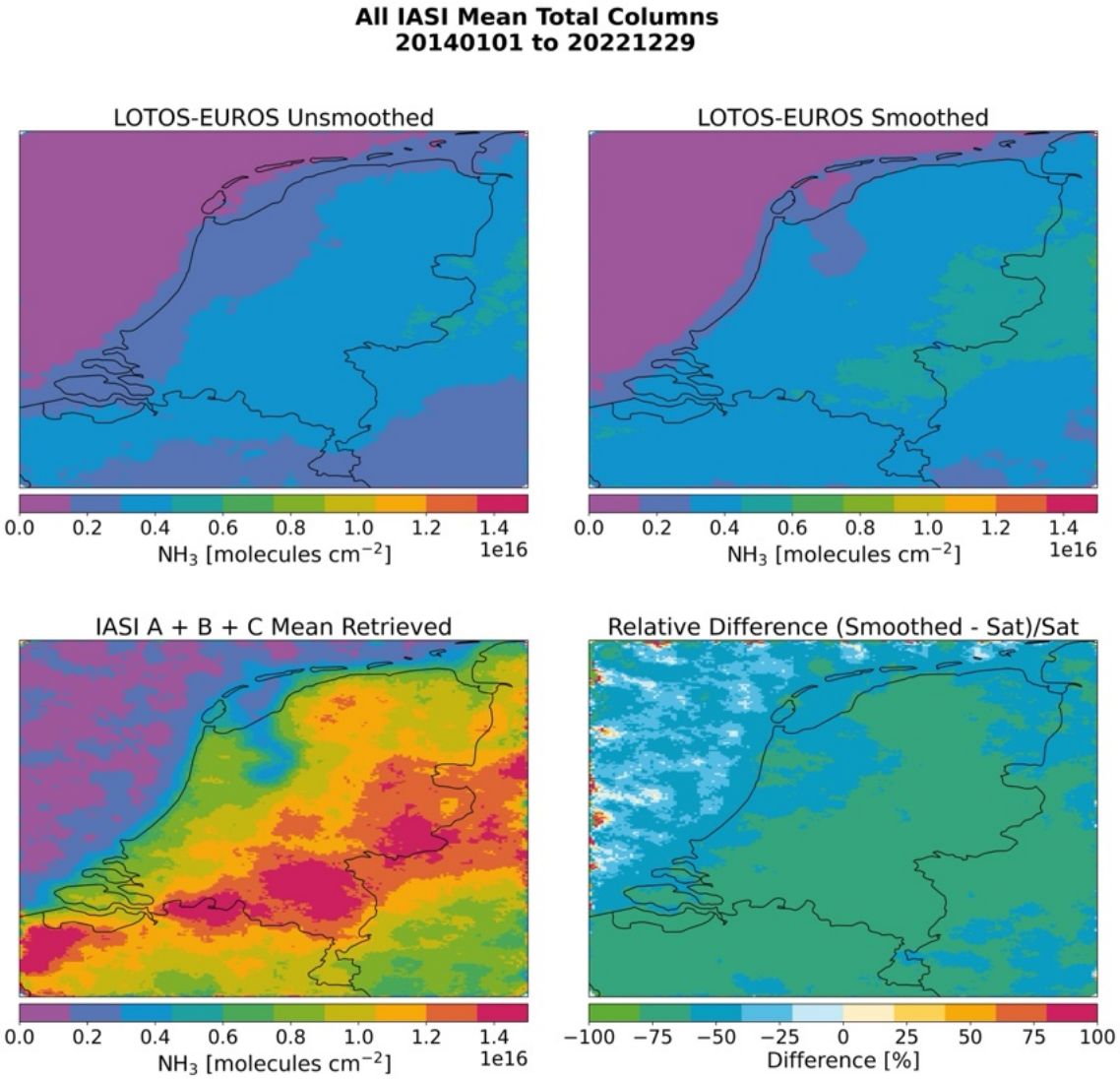


Figure 48: Spatial mean maps calculated of (top left) the base LOTOS-EUROS simulated NH_3 total columns at IASI overpass times, (top right) the LOTOS-EUROS total columns smoothed by the IASI averaging kernels, (bottom left) the IASI-A, B and C retrieved NH_3 total columns, and (bottom right) the relative difference between the smoothed model columns and the IASI retrieved columns.

- Extending the emissions dataset beyond 2019, followed by a re-evaluation of the model-satellite differences for these later years.
- Improving model parameterizations related to NH_3 deposition and chemical processes.

Continued comparisons with high-quality satellite observations such as IASI and CrIS will help evaluate these improvements and ensure a more accurate representation of NH_3 in the Netherlands in LOTOS-EUROS.

5.3 Comparison of EMEP4NL with CrIS

The NH_3 total columns from the EMEP4NL long term simulations were compared against the measurements from the CrIS instruments throughout January 2018 to December 2023. A map

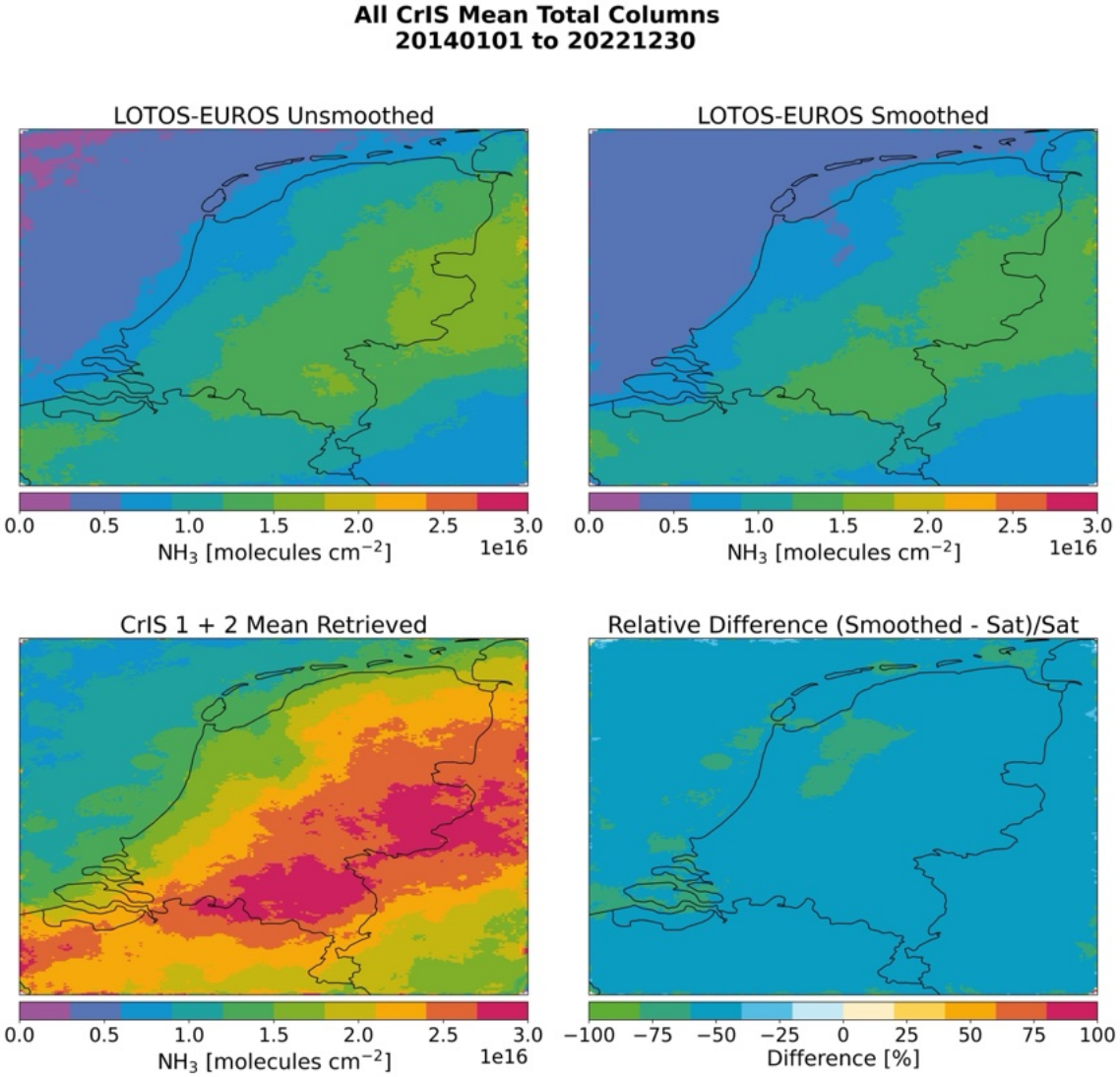


Figure 49: Spatial mean maps calculated of (top left) the base LOTOS-EUROS simulated NH_3 total columns at CrIS overpass times, (top right) the LOTOS-EUROS total columns smoothed by the CrIS averaging kernels, (bottom left) the CrIS 1 and 2 retrieved NH_3 total columns, and (bottom right) the relative difference between the smoothed model columns and the CrIS retrieved columns.

of the mean NH_3 total columns at the time of the satellite overpasses for CrIS 1 and 2 and the matching EMEP4NL total column concentrations is shown in Figure 52. Similar to the comparison with LOTOS-EUROS, the spatial distribution of the NH_3 total columns retrieved by CrIS 1 and 2 is generally captured by EMEP4NL. However, a clear negative bias in the modelled NH_3 total columns compared to the satellite-retrieved ones can be seen.

Figure 53 shows the time series of the monthly mean NH_3 total column from both the CrIS 1 and 2 observations between 2018 and 2023, and the corresponding EMEP4NL values smoothed by the CrIS averaging kernels. The modelled and retrieved NH_3 total columns show the same interannual variability. Over the full period between 2018 and 2023 a mean relative difference of -58.5% between EMEP4NL and CrIS was found. Overall, we can thus conclude that EMEP4NL and CrIS compare well in both the spatial- and temporal representation of NH_3. However, they differ significantly in magnitude, with the modelled NH_3 total columns from EMEP4NL being consistently lower than the satellite-retrieved total columns from CrIS 1 and 2.

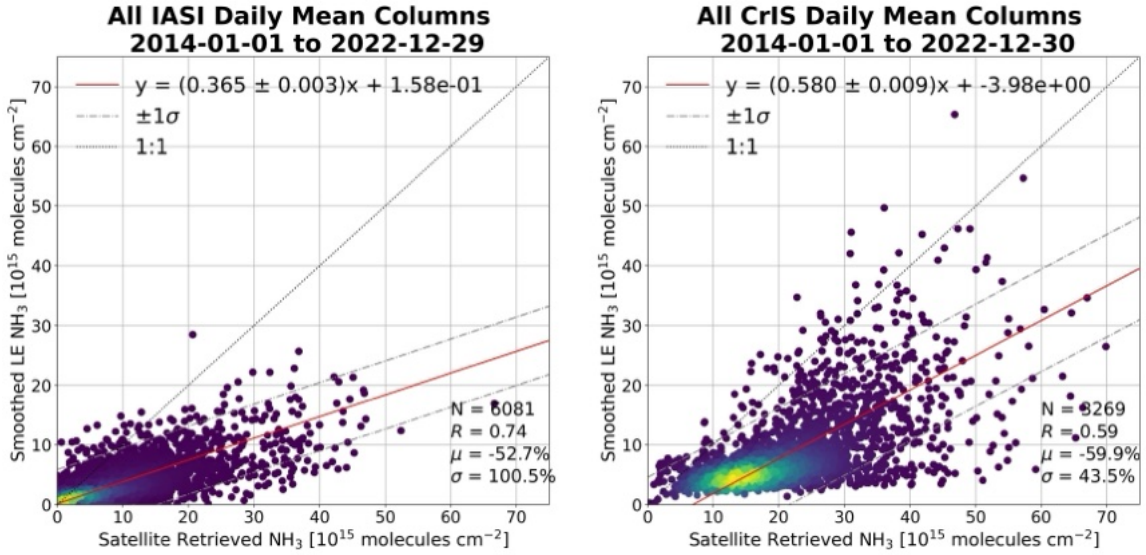


Figure 50: Correlation plots of (left) smoothed LOTOS-EUROS NH_3 total columns versus all IASI daily means, (right) smoothed LOTOS-EUROS NH_3 total columns versus all CrIS daily means.

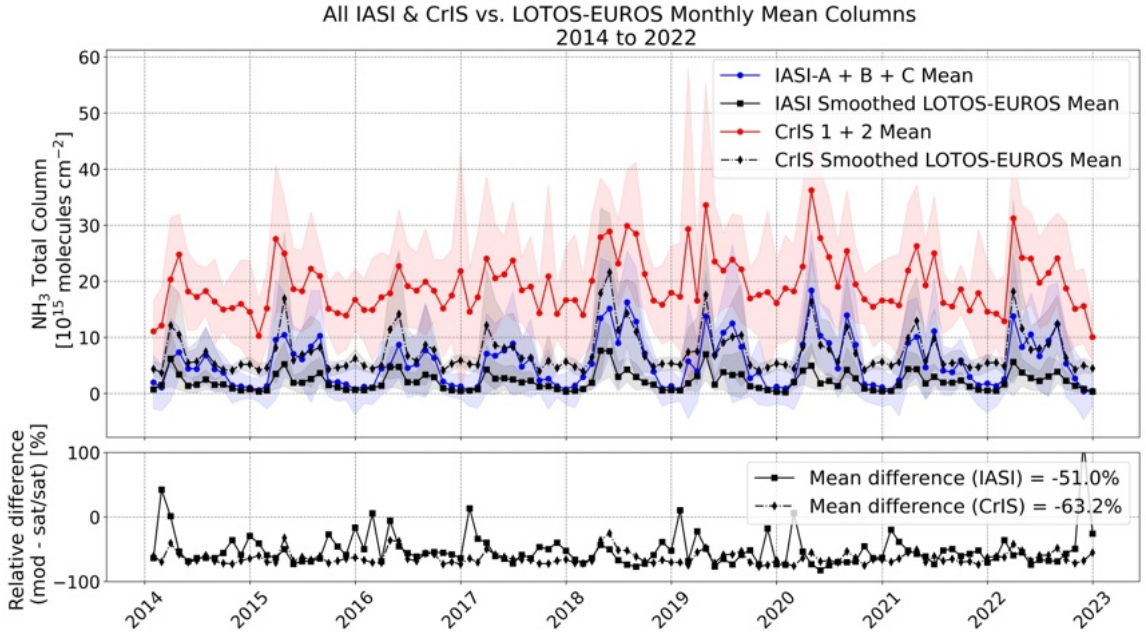


Figure 51: (Top) a time-series of monthly mean columns calculated across all IASI and CrIS observations, and the corresponding smoothed LOTOS-EUROS monthly means (i.e., smoothed by the IASI and CrIS averaging kernels), and (bottom) a time-series of the relative differences.

5.3.1 Summary

The NH_3 total columns from the EMEP4NL long term simulations were compared against the measurements from the CrIS-1 and 2 instruments throughout January 2018 to December 2023.

- The EMEP4NL model captures the spatial distribution of the satellite retrieved NH_3 total columns over the Netherlands well. Moreover, the modelled and retrieved NH_3 total columns show the same interannual variability.

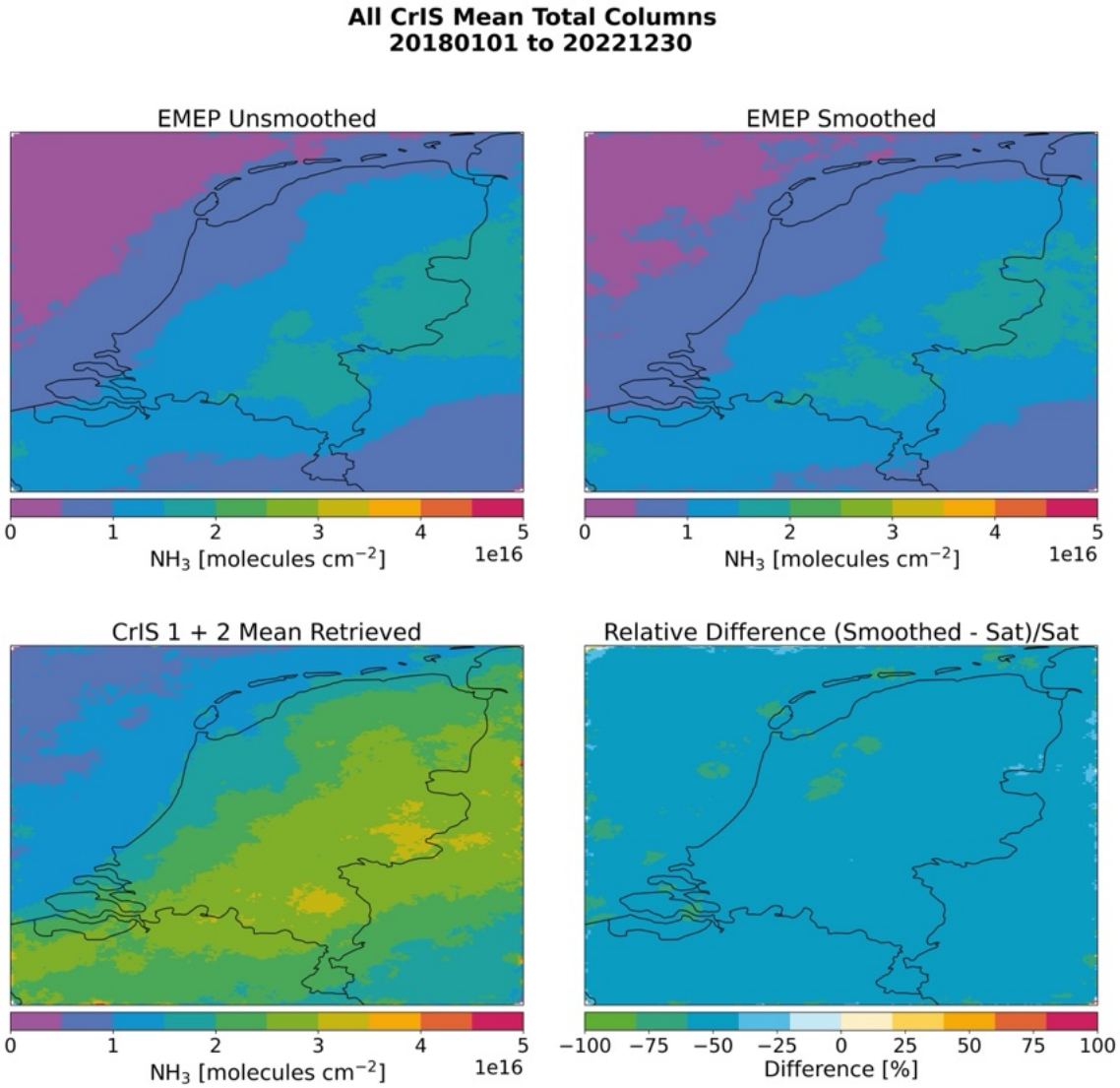


Figure 52: Spatial mean maps calculated of (top left) the base EMEP4NL simulated NH_3 total columns at CrIS overpass times, (top right) the EMEP4NL total columns smoothed by the CrIS averaging kernels, (bottom left) the CrIS 1 and 2 retrieved NH_3 total columns, and (bottom right) the relative difference between the smoothed model columns and the CrIS retrieved columns.

- The EMEP4NL and CrIS-1 and 2 satellite observations thus compare well in both the spatial- and temporal representation of NH_3. However, they differ significantly in magnitude, with the modelled NH_3 total columns from EMEP4NL being consistently lower (on average - 58.5% on monthly means) than the satellite-retrieved total columns from CrIS 1 and 2.

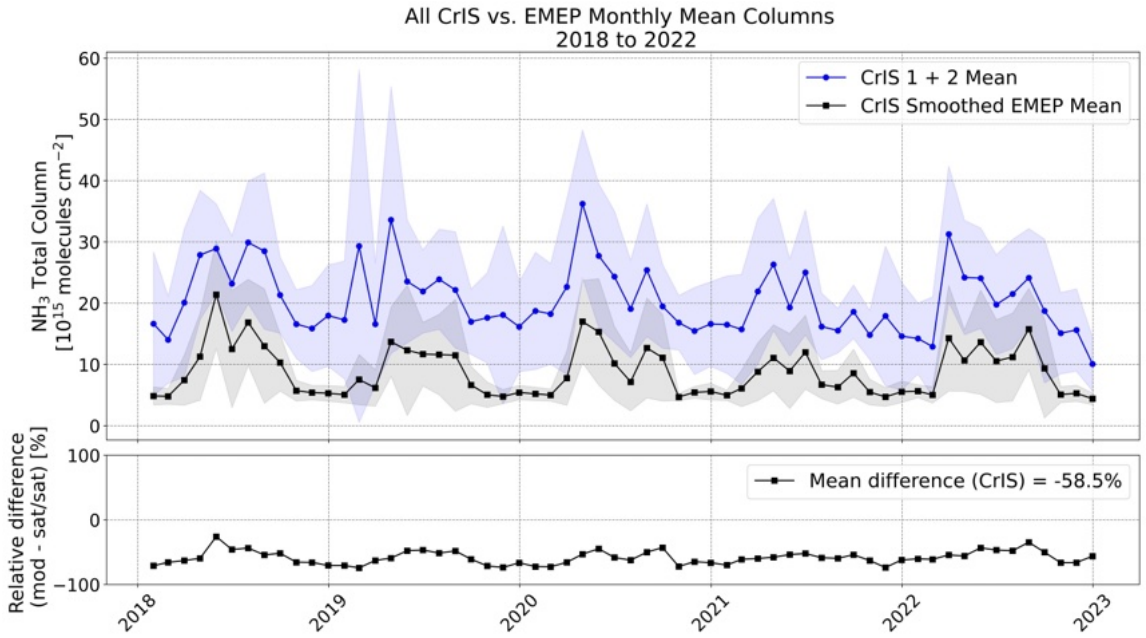


Figure 53: (Top) a time-series of monthly mean columns calculated across all CrIS 1 and 2 observations, and the corresponding smoothed EMEP4NL monthly means (i.e., smoothed by the CrIS averaging kernels), and (bottom) a time-series of the relative differences.

6 Emission estimates using four approaches for NOx and NH3

6.1 Introduction

In this chapter we will present results for emission estimates for NH_3 and NOx performed using the satellite observations of CrIS, IASI and TROPOMI. The four independent methodologies to estimate the emissions will be introduced, after which the results from the research done in NKS-SAGEN will be presented.

6.2 Inverse methods for emission estimates

In this section we introduce the four different methods used in the NKS study to derive emissions from satellite observations.

The first two, the flux-divergence approach and the Gaussian plume method require (in their basic form) minimal inputs.

The third and fourth approach, LETKF and DECSO, are inversion techniques which require runs from a CTM, where the emissions are optimised by minimising the mismatch between the model concentrations and observed satellite measurements.

6.2.1 Flux divergence approach

The Flux Divergence Approach - FDA (Beirle et al., 2019, 2021, 2023), is a method for estimating emissions from satellite data. It was originally used for NOx emissions derived from NO_2 satellite observations, but more recently it is also applied to derive emissions for other gases like methane (CH_4) and sulfur dioxide (SO_2).

The FDA method is based on the steady-state continuity equation and combines spatial patterns observed in satellite measurements with information of horizontal wind components to determine emission rates. The steady-state continuity equation can be written as

$$E = D + S = \nabla(LV\mathbf{w}) + \frac{LV}{\tau} \tag{6.1}$$

where the NO_x emissions (E) are computed as the sum of the divergence of the NO_x flux (D) and a sink term (S). The NO_x flux can be expressed as $LV\mathbf{w}$, where V is the tropospheric NO_2 VCD observation, L is a conversion factor from NO_2 to NO_x, and $\mathbf{w}$ represents the wind field. The sink term S can be represented as LV/τ, where τ is the lifetime of NO_2 at overpass time. Lastly, The divergence term $\nabla(LVw)$ can be estimated on a grid using (for instance) a fourth-order central-finite difference,

$$\nabla\mathbf{f}(x) = \frac{f(x-2h) - 8f(x-h) + 8f(x+h) - f(x+2h)}{12h}, \tag{6.2}$$

where *f(x)* represents LVw and h is the spacing between observations. Note that other finite difference methods can be used.

Since its inception, the method has been utilized in various studies to estimate emissions on a global scale (Beirle et al., 2023) and in specific regions, including the United States (Dix et al., 2022), North India (Misra et al., 2021), and South Asia (de Foy et al., 2022), demonstrating the capability of the FDA to derive emissions, particularly for point sources. However, a quantitative

assessment of the accuracy of these satellite-derived emissions is still lacking. Additional details on this topic are provided in Section 6.4.

6.2.2 Multi-source Gaussian plume method

The multi-gaussian plume method, developed by Fioletov et al., is a plume-based fitting routine that establishes source-receptor relations between point sources and satellite observations. The relatively simple method allows for the derivation of source strength estimates directly from satellite column observations. For an in-depth description of the method we point you to the original publication by Fioletov et al. (2017). To summarize, a total column distribution can be represented by a combination of functions describing the source contributions of several nearby sources (Fioletov et al., 2017). Each of these functions describes the advection, diffusion and decay of the concentrations down and cross-wind from a source. In a system with N sources, the vertical column concentration of each satellite observations can be described as;

$$Column_j(lon_j, lat_j, \mathbf{s_j}) = \sum_i a_i f(x_{i,j}, y_{i,j}) g(y_{i,j}, s_j) + a_0 + \epsilon \tag{6.3}$$

where lon and lat are the longitudinal and latitudinal position of observation j, w the wind vector at observation j. a_0 is the background concentration (i.e., from long-range transport) which for shorter lived species reduces to zero. Lastly ϵ reflects a potential bias in the satellite observations and/or a misrepresentation of the plume. Combining all observations and sources into a single system, we can describe the system as a simple linear equation,

$$\mathbf{Ax} = \mathbf{B} \tag{6.4}$$

where matrix A represents the system of source-receptor relations, vector x the source enhancements, and vector B the satellite observations. Solving this system leads to an estimate of the source enhancements x, which in combination with the lifetime results in a total emission for each of the sources. Since the method was initially developed by Fioletov et al. (2017), it has been applied in several studies for various atmospheric species such as NH_3 (Dammers et al. (2022)), SO_2 (Fioletov et al., 2017), and NOx(Dammers et al., 2024; Fioletov et al., 2022). The plume-based routine is available as part of an open-access tool developed for the Umweltbundesambt in Germany (UBA, https://gitlab.opencode.de/uba-emsit/dev/space-emissions). The open-access tool also contains the divergence methodology developed by Beirle et al. (2019).

6.2.3 LETKF

The LOTOS-EUROS LETKF (Van Der Graaf et al., 2022) is a state-of-the-art ensemble-based data assimilation method that provides an efficient way to update model states by incorporating observational data. The LETKF is coupled directly to the LOTOS-EUROS model, a three-dimensional regional chemical transport model used to simulate the distribution of atmospheric trace gases and aerosols. The localized approach of the LETKF allows for the assimilation of high-dimensional satellite data while maintaining computational feasibility, making it well-suited for applications in high resolution chemical transport models. By assimilating NH_3 and NO_2 satellite measurements, the LETKF system adjusts the model states to produce optimized concentration fields that reflect observations more accurately. The LETKF system also outputs perturbation factor fields that can then be fed back into the forward model to allow for the adjustment of emission and deposition fields, offering improved estimates of pollutant sources and sinks. Furthermore, through a comparison of the optimized and un-optimized ("background") simulations with ground-based and satellite measurements, the improvements can easily be quantified.

The simulation domain used for this study is the PINETI domain covering the Germany, the Netherlands, and Belgium between [3°, 9°] x [49°, 55°] with a spatial resolution of $0.1° \times 0.05°$. The model

was configured with 12 vertical layers, which includes the planetary boundary layer where most emissions and chemical processes occur. LOTOS-EUROS includes modules for emissions, advection, vertical mixing, dry and wet deposition, and gas-phase and aerosol chemistry. Emission data were taken from CAMS REG v5.1 emission inventory, which include anthropogenic, biogenic, and agricultural sources relevant to NH_3 and NO_2. The model was configured to run with hourly meteorological input from ECMWF, providing key parameters such as wind speed, temperature, and boundary layer height. Chemical mechanisms in the model, including gas-phase reactions and aerosol dynamics, were used to simulate the complex interactions between emitted gases and the atmosphere. Model output provides hourly concentration fields of NH_3 and NO_2, which are then used as a baseline for data assimilation with satellite observations. For further details on the LOTOS-EUROS Kalman Filter we point the reader to Van Der Graaf et al. (2022).

6.2.4 DECSO version 6.4: inversion of TROPOMI and CrIS observations

The inversion algorithm DECSO (Daily Emissions Constrained by Satellite Observations) has been developed at KNMI for the purpose of deriving emissions for short-lived gases (Mijling & van der A, 2012). DECSO is using a Kalman Filter implementation for assimilating emissions. The emission forecast model is based on persistency from the analysis, while the concentrations are calculated from the emissions by a chemical transfer model (CTM) and compared to satellite observations. The sensitivity of concentrations to emissions is calculated from multiple forward trajectories to account for the transport of the short-lived gas, but only a single CTM forward run is needed. Recent developments of the algorithm to improve its resolution and quality have led to the release of version 6.4. The most important recent updates are the use of a recent version of the chemical transport model, improved use of TROPOMI observations and changes in the sensitivity matrix calculations.

The chemical transport model in DECSO has been upgraded to the latest version of the Eulerian regional off-line CTM CHIMERE v2020r3 (Menut et al., 2021). The implementation of CHIMERE in DECSO was described in Ding et al. (2017). CHIMERE is combined with the Copernicus Land-cover 2019 data (Buchhorn et al., 2020) and HTAP v3 (Hemispheric Transport of Air Pollution, Crippa et al. (2023) of 2018 for the source sector split of the emissions. The meteorological input data for CHIMERE are the operational European Centre for Medium-Range Weather Forecasts (ECMWF) weather forecasts.

The error parametrizations for the emission model and observations are based on the Observation-minus-Forecast (OmF) and the Observation-minus-Analysis (OmA) statistics of previous runs. The latest version of DECSO can be applied to simultaneous optimisation of emissions of NO_x and NH_3 (Ding et al., 2024).

Although HTAP v3 has been used for the sector distribution of NO_x and NH_3 emissions and the other species in CHIMERE, no use is made of a-priori (bottom-up) emissions in DECSO. DECSO is using a persistency forward model in which the emissions of the current day are equal to the emissions of the previous day. In addition, there is a strong dependency of the calculated emissions on the observations as shown in (Ding et al., 2020). Since the derived emissions are updated by addition and not by multiplication factors, unknown sources or emission changes are detected fast.

In a post-processing step, the total monthly NO_x emissions are split into anthropogenic and (biogenic) soil emission contributions. The soil emissions show a strong seasonal cycle with low emissions in winter, while the anthropogenic emissions are more constant over the year. The soil NO_x emissions are derived by fitting the monthly emissions in a selection of grid-cells without any significant anthropogenic contribution according to land-use data. In this way the monthly averaged soil NO_x emissions in the categories for forest, agricultural and shrub-land are derived. These monthly soil NO_x emissions are weighted with the land-use type of these 3 categories in each grid cell and subtracted from the total derived NO_x emissions to end up with the anthropogenic NO_x emissions

discussed in this study. This splitting method is described in detail in Lin et al. (2023).

For the monthly emissions also the precision of the emission in each grid cell has been calculated. Each daily NO_x emission per grid cell derived by DECSO is accompanied by a standard deviation calculated according the Kalman Filter equations (the standard deviation is part of the emission data product of DECSO). As the starting point of each daily step in the calculation by DECSO is the emissions of the previous day, the resulting emissions will show an autocorrelation in their errors. For each grid cell the autocorrelation function ρ_k (for time lag k) has been calculated for each month. We see typically that the autocorrelation effects in the errors have disappeared completely after about 1 week.

When calculating the variance of the monthly mean values, we must take this autocorrelation function into account. The variance S of the monthly mean NO_x emissions per grid cell is calculated following Box et al. (2008) as

$$S = \frac{\sigma^2}{n}\left[1 + 2\sum_{k=1}^{n-1}(1 - \frac{k}{n})\rho_k\right], \tag{6.5}$$

where σ is the mean standard deviation of the emissions over the month and n is the number of days in the month. We assume here that σ is not varying a lot over the month. This precision σ is calculated in the Kalman equations of the inverse modelling and it depends on the precision of the TROPOMI NO_2 superobservations. The precision depends on the location and emission magnitude, but on average the precision is estimated as 8% for annual emissions, 25% for monthly emissions and between 10 and 60% for the daily emissions.

DECSO has been applied to many regions in the world, usually using a spatial resolution of 0.2° by 0.2°. In this study we focus only on the domain (50° - 54° N, 2° - 9° E) on spatial resolution of 0.1°. The temporal resolution of our inversion is daily, usually averaged to monthly or yearly mean values, for the period of 2019 to 2023.

More detailed information on the method and recent updates can be found in van der A et al. (2024), Ding et al. (2024) and Lin et al. (2023).

6.3 NH_3 and NO_x emission estimates derived using the Gaussian plume method (MSPM)

The multi-source Gaussian plume (MSPM) methodology has been used in several studies to estimate emissions from satellite observations such as NH_3 observed by CrIS (Dammers et al., 2022) and NO_2 and SO_2 observed by the OMI and TROPOMI sensors (**Fioletov_2022**; Dammers et al., 2024; Fioletov et al., 2017). Within this study, the method is applied to observations from both the CrIS and IASI instruments to derive NH_3 emissions, as well as to NO_2 observations from the TROPOMI instrument to derive NO_x emissions over the Netherlands. The accuracy of the method, the lifetime, and the footprint assumptions will be evaluated in a follow-up NKS work package.

6.3.1 Methods for the NH_3 estimates

Past studies using the MSPM methodology were focused on either different regions or used coarser resolutions, making this the first study to appy this method to evaluate the emissions in the Netherlands in-depth. We follow the methodology as described in section 6.2.2 with a target grid of 0.1° x 0.1°. Emissions are derived for various periods, ranging from datasets with the full temporal coverage of the instruments (at maximum from 2014 up to 2020), three-yearly averages as well as monthly estimates. Too few observations are available to look at individual years at the

spatial resolution of the source grid that we target here. The lifetime and plume-spread parameters are dependent on the species and the satellite instrument footprints. The lifetime of ammonia varies strongly both spatially and temporally, is highly dependent on the various sinks, and can range from a few hours up to a day (Dammers et al., 2019; Van Damme et al., 2018). While the lifetime of NH_3 can be as short as a few hours in areas far from sources, it is often primarily limited by the deposition velocity in regions with high NH_3 emissions, as chemical sinks like HNO_3 and H_2SO_4 tend to be largely saturated. Across the Netherlands, this is generally the case, as much of the country lies within the 'blanket' of NH_3 from upwind sources. While the lifetime can be derived within emission plumes of individual sources, this becomes infeasible with the multi-source method, as the lifetime parameter introduces a non-linearity that is difficult to resolve. Instead, the lifetime was derived by tracking the mass of individual sources with the help of the LOTOS-EUROS labeling routine. With the help of this approach, an average yearly lifetime of approximately 8 hours is derived with a standard deviation of 2 hours (for a set of monthly averaged lifetimes). The lifetime can be expected to be shorter in some of the coastal regions, while it is predicted to be slightly longer within the strongest source regions. This will result in a limited underestimation of the emission at the coast while for inland regions, the reverse will be true. The resulting uncertainty in the emissions has been incorporated in the final error estimates. For the plume spread parameter σ a single value of 15 km is chosen for both IASI and CrIS. At a target source grid resolution of approximately 10×10 km^2, the plume spread is largely determined by the satellite pixel footprint, which is (at best) 15 km for CrIS and 12 km for IASI (Dammers et al., 2019).

Satellite observed concentrations of short-lived species are only an indications of emissions near and at the overpass time. A correction factor needs to be applied to the estimated emissions to account for the diurnal variability to get to a representative value for the entire day. The simplest approach is to use a basic box model to approximate the mass over time and then apply a posterior correction. Assuming a mass $m(t)$, lifetime (λ) and the emission E at time t, the mass can be calculated with

$$m(t) = m(t-1)e^{-\tau} + E(t) \tag{6.6}$$

where τ equals $1/\lambda$.

This equation is applied to the Dutch emission set, with the hourly emissions injected into the box model for a whole year. The temporal emission allocation is based on the the average NH_3 emission profile for all NH_3 sources within the Dutch domain (Manders et al., 2017). A lifetime of about 8 hours and an overpass time of around 13:00 LST results in a correction factor of 1.20, meaning that the estimated emissions can be expected to be overestimated by around 20%. Using the same approach, the correction factor for IASI is around 1.0. A similar correction factor can be derived for the seasonal variability to account for the larger number of observations in spring and summer compared to autumn and winter (due to a decrease in cloud cover and better measurement conditions). The resulting correction parameter is based on the weighted mean of the number of observations per month and the correction factor derived for each month. Based on this approach, a value of 1.11 is found for both instruments. When combined, the correction factor becomes 1.34 for CrIS and 1.11 for IASI for any yearly estimate and 1.2 for the individual months. Note that in both cases the values are based on a-priori temporal allocation as used in the model, if the diurnal or seasonal variation is stronger or weaker than expected this will result in a change in the correction factor. Lastly, as the emission estimates were derived early in the project the version of the IASI product used in this study is IASI-ANNIv3.1, which based on Clarisse et al. (2023) is low biased by at least 10% compared to the current product. To correct for the difference we apply the individual concentrations are multiplied by a factor 1.1. The satellite-derived emissions can be directly compared to the inventory emissions, which are a combination of the CAMS-REGv5.1 inventory, the German GreTa inventory and the Dutch emission registration as described in section 2.2. There is a limit to the spatial extent at which satellite instruments can resolve individual emission locations. In the case of ammonia, with an effective lifetime of 8 hours and satellite footprints on the order of 15 km, the limit to detect two separate sources is approximately 20 km (McLinden et al., 2024). To ensure a fair comparison between the satellite-derived emissions and the inventory, a Gaussian smoothing filter is applied to the inventory emissions with

a correlation length of 20 km. Finally, uncertainty estimates of the NH_3 inventory values are based on a report by (Wever et al., 2021) . The report gives an uncertainty value of 28% for the NH_3 emissions in the Netherlands as a whole. The uncertainties of the satellite emission estimates are a mix between the uncertainty in the lifetime ($\pm$2 hours), the meteorology, and the uncertainties in the satellite products. On average this results in a conservative estimate of uncertainty on the order of 30–40%.

6.3.2 Methods for the NO_x estimates

Several studies have used the MSPM method to estimate emissions of NO_x (Dammers et al., 2024; Fioletov et al., 2022). The closest to our application here is a study by Dammers et al. (2024) where the method was applied Germany. That publication was part of a larger study where several simple inversion methods were applied to TROPOMI data (Dammers et al., 2023). Here, we use a similar setup as used in that study. The target source-grid is $0.1^\circ \times 0.1^\circ$, which is the same that was used for the NH_3 MSPM emissions estimates. Emissions are estimated at monthly level and aggregated to yearly values. Due to the higher spatial resolution of the TROPOMI footprints, many more observations are available to constrain the emissions, allowing for shorter periods of data to be used for the emission estimates. Where both setups diverge further is the lifetime. Based on Dammers et al. (2024) a lifetime of 4 hours is assumed which is based on previous studies such as by Beirle et al. (2019). See Dammers et al. (2024) for a more detailed discussion on the choice of lifetime including the potential drawbacks. Like for the NH_3 use case, the plume spread is largely determined by the satellite footprint, therefore the parameter is set to 7 km (Dammers et al., 2024; Fioletov et al., 2022; Griffin et al., 2021). Similar to CrIS and IASI a correction factor needs to be applied to correct for the overpass time of TROPOMI as well as the higher number of observations in summertime. Based on the same method a diurnal correction factor of 1.24 and a seasonal correction of 1.05 are found, which combine into a factor of approximately 1.30. Finally, TROPOMI only observes NO_2 whereas the comparison will be made for NO_x emission estimates. The NO_x:NO_2 concentration ratio depends on chemistry which is influenced by several factors such as O_3 concentrations and the photolysis of NO_2. A commonly used NO_x:NO_2 concentration ratio is 1.32 ± 0.26, which was also used by (Beirle et al., 2019) and (Dammers et al., 2022). Within section 6.4 the accuracy of that assumption under summertime conditions is further explored. The results are again directly compared to the inventory emissions. In the case of NO_2, with an assumed effective lifetime of 4 hours and satellite footprints of approximately 5–7 km, the limit to detect two separate sources is about 10 km McLinden et al. (2024). Here, a Gaussian smoothing filter with a correlation length of 10 km is applied to the inventory emissions (in comparison to the correlation length of 20 km that was used for NH_3). Finally, uncertainty estimates of the NO_x inventory values are based on a report by (Wever et al., 2021) . The report gives an uncertainty value of 17% for the NO_x emissions in the Netherlands as a whole. The uncertainties of the satellite emission estimates are a mix between the uncertainty in the lifetime ($\pm$1 hours), the meteorology and the uncertainties in the satellite products. On average this results in an conservative estimate of uncertainty in the order of 30-35% (Dammers et al., 2024).

6.3.3 Results for the NH_3 estimates

The results of the CrIS and IASI NH_3 emissions estimates are summarized in figure 54 and table 13. Using the country, province and municipal administration level borders both the inventory emissions and the satellite estimates are aggregated to country, provincial and municipality emission totals. The top row in figure 54 shows the emission results for the 2014–2019 period at municipal level in kg per hectare. As per inventory data several large NH_3 emission clusters can be observed throughout the Netherlands, for example in the east of North-Brabant (livestock), and the area around Barneveld (e.g., poultry). Similar hot spots are observed in the CrIS and IASI estimates, although emissions are more smeared out over the neighboring municipalities. This is to be ex-

Region	Inventory [kt]	IASI [kt]	CrIS [kt]
Drenthe	8.5±2.4	9.3±2.8	15.8±4.7
Flevoland	3.7±1.0	5.4±1.6	7.2±2.2
Friesland	12.1±3.4	14.4±4.3	20.8±6.2
Gelderland	20.8±5.8	23.6±7.1	37.1±11.1
Groningen	8.4±2.4	9.0±2.7	12.2±3.7
Limburg	8.9±2.5	8.4±2.5	12.9±3.9
Noord-Brabant	23.3±6.5	28.1±8.4	39.4±11.8
Noord-Holland	6.0±1.7	8.4±2.5	11.3±3.4
Overijssel	15.3±4.3	16.5±4.9	25.2±7.5
Utrecht	5.6±1.6	6.2±1.8	8.8±2.6
Zeeland	4.3±1.2	6.0±1.8	7.3±2.2
Zuid-Holland	7.8±2.2	10.7±3.2	12.7±3.8
Whole Netherlands	124.7±35.0	145.8±43.7	210.8±63.1

Table 13: 2014–2019 NH_3 emission totals from inventory and satellite emission estimates calculated for the individual provinces and the whole Netherlands. Values are reported in terms of NH_3 weight.

pected as the footprint of the satellite as well as the lifetime of ammonia limits the extent of what can be spatially resolved (McLinden et al., 2024). By shifting to province totals the impact of the sensor limitations can be somewhat averted, resulting in a more balanced comparison of the three datasets. In absolute terms, at the province level (middle left panel) the inventory and satellite estimated emissions correlate strongly, indicating that spatially the inventory and satellite estimates align closely. In particular, IASI compares exceptionally well to the inventory totals, deviating at most by a few percent, whereas CrIS gives higher total emissions that are about 30% above the inventory values. When the result is evaluated temporally, a different picture emerges. The middle right panel shows the total emissions per month at country level. The inventory monthly emission cycle peaks in spring and summer. This roughly matches the spreading of fertilizer at the start of spring and the increased volatilization of ammonia (both from spreading and livestock) in summer. Both the CrIS and IASI estimates, however, show that the spring and summertime peaks should be much more in balance instead of one strong peak in spring and a much smaller peak in summer. At an inter-annual level the CrIS and IASI estimates show a very different trend compared to the inventory. Both instruments display a positive trend between 2014 and 2019 whereas the inventory is trending slightly down. Similar to the monthly values, CrIS shows an offset compared to the IASI values but both datasets, relatively, show the same inter-annual pattern.

Discrepancies between the inventory and satellite estimates can have several origins within the inventory, satellite product and method. On the inventory side the first thing to take into account is the uncertainty in the underlying emission factors and proxy maps used to sum up and spatially distribute the emissions. As activity data is accurately tracked in the Netherlands, the main parameter of concern are the emission factors. A potential error arises when soil temperature is not taken into account enough when calculating the total yearly emissions. When temperatures increase, ammonia has a much higher potential to volatilize, increasing the emissions. Most of the recent years are much warmer compared to the typical climatologies, so if inventory emissions are corrected towards a standard climatology the final emissions could end up low. Alternatively, the temporal emission distribution is at fault. On the seasonal timescales it seems that the there is too much ammonia being attributed to springtime fertilizer application, whereas summertime

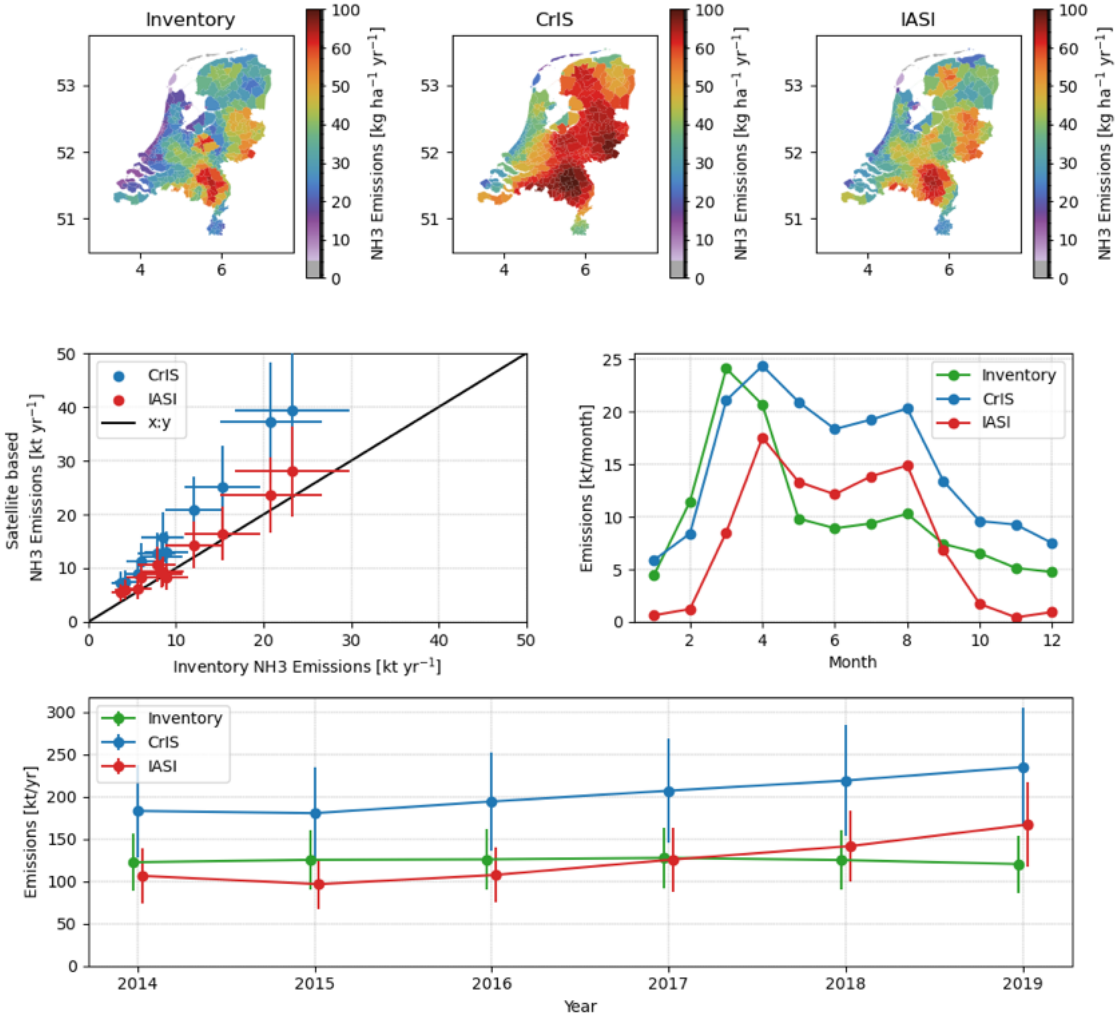


Figure 54: Summary of the results for NH3. The top panels, from left to right, shows the NH_3 inventory emissions, the CrIS based emission estimates, and the IASI based emission estimate, all using 2014-2020 data. The left middle panel, shows the 2014-2020 emission inventory values directly compared to emission estimates based on CrIS and IASI at provincial aggregation level. The right middle panel shows the 2014-2020 averaged monthly inventory and estimated emissions of CrIS and IASI. Lastly, the bottom panel shows the 3-yearly averaged emissions of the emission inventory, and the CrIS and IASI emission estimates.

emissions, when volatilization conditions are favorable, are biased low. Striking is that the overall seasonal pattern seen for CrIS and IASI are very similar, with a difference in the absolute levels. Similar results, with much weaker springtime peaks, were obtained with more complex inversion methods (Cao et al., 2022; Ding et al., 2024; Van Der Graaf et al., 2022). Similarly the annual emissions seem to be increasing in the satellite derived emission estimates. While the lifetime is not adjusted from year to year, the changes cannot solely be explained by changes in chemistry and deposition, as that would only result in several % changes in the column totals. Another possibility it a mismatch in the assumed and real-life diurnal emission cycle. The assumed cycle shows peak emissions in the middle of the day, but if temperatures are on average higher, this day:night ratio could change further. Meaning that daytime emissions could be stronger, and a larger correction

is needed, which would bring the CrIS estimates closer towards the IASI estimates.

On the side of the satellite instruments, the difference between CrIS and IASI can arise from several factors. The most obvious one are the differences in the instrument sensitivities and retrieval products. The CrIS instrument has a much lower detection limit which allows for the detection of lower concentrations whereas the IASI instrument would only observe noise. On the product side however, we have already seen in the validation study in section 3.3.1 that the CrIS product has a slight high bias of about 5×10^{15} molecules cm^{-2}, especially over regions with low to intermediate concentrations. To put this into kg/ha terms, one would end up with a value on the order of 10 kg/ha/yr more. A consistent high bias will naturally result in an offset which translates into a consistent offset. Similarly, the IASI product showed a multiplicative low bias of about 25% for medium to higher concentration regimes, this in effect would result in a multiplicative scaling of the final emissions of around 25%.

Finally there is the methodology itself. The most important factor is the lifetime estimate. Any deviation from the lifetime itself will directly affect the final emission estimate. The assumed lifetime of 8 hours can, for example, be deemed high for regions where emissions and concentrations are much lower (Dammers et al., 2022). This means that for such areas the real lifetime is likely much shorter (i.e., 4 hours), and the emissions are biased low by about 100% of the current value. Vice versa for regions with stronger emissions the lifetime might be too short, which means emissions can be biased high. If the true lifetime is 12 hours instead of the assumed 8 hours, the emissions would end up 33% lower than current estimates. This rough estimate shows that the most crucial parameter to reduce the uncertainties is a more accurate estimate of the lifetimes. The two main parameters affecting lifetime are deposition and chemistry. In the case of the Netherlands, the more dominant term will be deposition as it is a region with limited chemical sinks. It is thus essential to obtain an accurate estimate of the deposition, and especially dry deposition, since it is the largest fraction of the total deposition.

6.3.4 Results for the NO_x estimates

The results of the TROPOMI NO_2 emissions estimates are summarized in figure 55 and table 14. Similar to the NH_3 results, the borders of administration levels are used to aggregate both the inventory emissions and the satellite emission estimates. The top row in figure 55 shows the emission results for the 2019–2021 period at municipal level in kg per hectare. Both the satellite and inventory emissions show the largest emission hotspot around the Rotterdam harbour, similar peaks are observed in other major industrial areas around Amsterdam, Ijmuiden, Eindhoven and Terneuzen (some addition of the Belgian harbours leaking into the picture due to smoothing). A clear difference can be found in the regions with less industry such as Friesland and Drenthe where the satellite estimates show much higher values. A similar offset can be observed throughout the rest of the Netherlands away from the major industrial and population hotspots. On a province to province level (middle left panel) a strong correlation is observed between the inventory and satellite based estimates. Especially the provinces with higher emissions show a much larger spread and uncertainty. The monthly averaged emissions show a clear seasonal cycle in the satellite based estimates, whereas the inventory emissions show a peak at the start of the year while the rest of the year has a fairly constant value. On a whole the year to year satellite based emission estimates vary quite similarly, with the largest emission totals found for 2019, pre-covid, while 2020 shows the expected decrease in emissions, while 2021 ends up with a similar value. The inventory emissions similarly show a high point in 2019 while 2020 and 2021 have similar values. Comparable results were found for the german emission estimates (Dammers et al., 2024). The total satellite based emissions end up about 100kt higher than the emission inventory but are mostly within the uncertainty limits.

As with the ammonia emission estimates, the differences can originate from inventory errors, product biases as well as uncertainties in the methodology. Dammers et al. (2022) gave an extensive overview of potential causes for the mismatch between the inventory and satellite estimates. While

with NH_3 there is a potential for a temperature dependent bias in the inventory emissions, for NO_x there is no obvious candidate. There is some potential source such as differences in fuel sold in other countries while usage occurs inside the Netherlands as well as smearing between the borders, but the largest variations are found in the western provinces with major industry and the larger harbours. The offset found in the background regions is quite similar to the one found for the german agricultural and forest regions, and could be related to an underestimate of soil NO_x emissions.

A more probable cause for the mismatch can be found in the methodology. The current uncertainty range of roughly 35% is dominated by the uncertainty related to the lifetime estimate. By switching to a modeled NO_2 lifetime that uncertainty term could be greatly reduced. It can be expected that the additional of a seasonally dependent lifetime will greatly help the monthly emission estimates. Summertime conditions, with an increase in the photolysis of NO_2, would result in much lower lifetimes compared to wintertime conditions, something that is currently not reflected in the monthly emission estimates. Similarly year to year variations in sunhours would cause variations in the effective lifetime. Another option is to also model the NO_x:NO_2 ratio. Both approaches were tested within section6.4 where the lifetime and NO_x:NO_2 ratios were modelled based on LOTOS-EUROS estimates and showed a clear improvement in accuracy when using both parameters. The study showed that especially lifetime seemed to be longer, even in summer, for the western part of the Netherlands. This naturally would result in a lower estimate of the satellite based emission. It is important to keep in mind that currently the method is mostly independent of other model parameters, and the addition of a modeled lifetime or ratio can also limit the methodologies capabilities to for example find new sources.

Lastly the TROPOMI product itself can be a source of uncertainty. Previous studies (e.g., Verhoelst et al., 2021) have shown a high bias of TROPOMI columns over background regions, and a low bias over industrial and population hotspots. The negative difference found for the two provinces with major industry, north and south holland, could potentially be related to that low bias. Similarly, the positive difference over background regions can be explained by the high bias in those regions.

table 13

Region	Inventory [kt]	TROPOMI [kt]
Drenthe	8.2±1.3	13.0±3.9
Flevoland	5.6±0.9	9.0±2.7
Friesland	11.9±1.9	15.2±4.5
Gelderland	31.9±5.0	36.7±11.0
Groningen	11.3±1.8	12.4±3.7
Limburg	15.0±2.4	20.8±6.3
Noord-Brabant	35.5±5.6	48.9±14.7
Noord-Holland	31.9±5.0	26.2±7.9
Overijssel	14.3±2.3	18.7±5.6
Utrecht	14.6±2.3	12.3±3.7
Zeeland	±14.5±2.3	20.8±6.2
Zuid-Holland	50.2±7.9	42.3±12.7
Whole Netherlands	244.9±38.7	276.2±82.9

Table 14: 2019–2021 averaged NO_x emission totals from the inventory and satellite estimes calculated for the individual provinces and the whole Netherlands. Values are reported in terms of NO_2 weight.

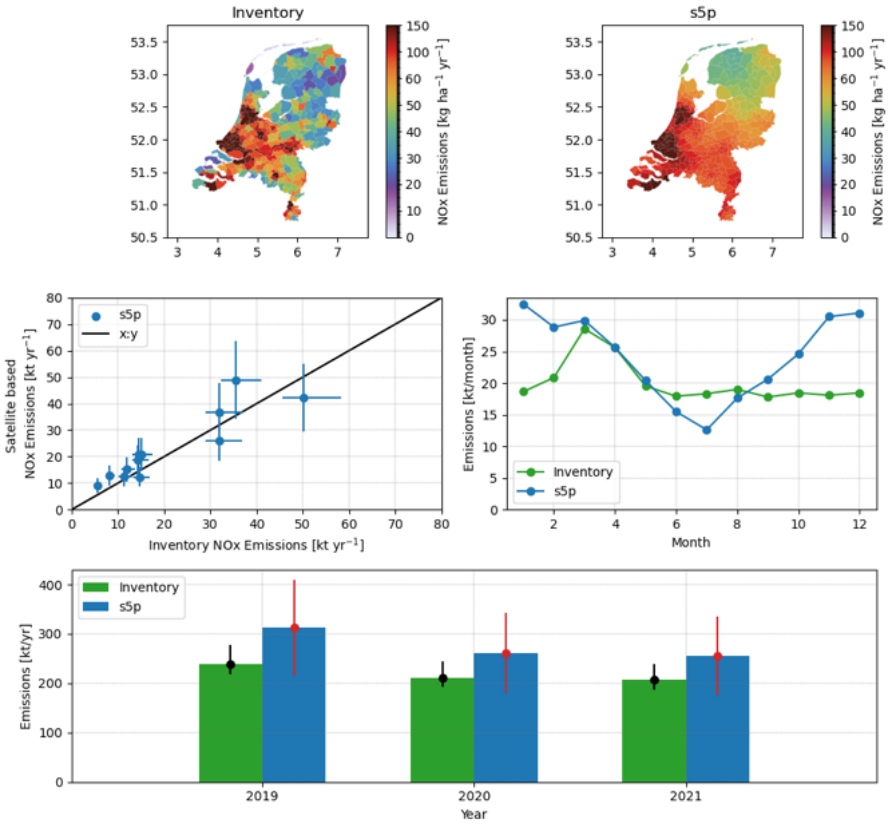


Figure 55: Summary of the results for NO_x. The top panels, from left to right, shows the NO_x inventory emissions and the TROPOMI based emission estimate, all using 2019–2021 data. Values are reported in terms of NO_2 weight. The left middle panel, shows the 2019–2021 emission inventory values directly compared to emission estimates based on TROPOMI at provincial aggregation level. The right middle panel shows the 2019–2021 averaged monthly inventory and estimated emissions based on TROPOMI. Lastly, the bottom panel shows the yearly averaged emissions of the emission inventory, and the TROPOMI emission estimates.

6.3.5 Conclusions for the NH3 and NOx estimates

To conclude, the multi-source gaussian plume method allows for fast and independent emission estimates to evaluate emission inventories. As satellite products are typically produced near-real-time, emission changes can be rapidly detected and incorporated into air quality modelling. Another strength of the method is that it can detect new emission sources, something that is not always possible with other data-assimilation and inversion methods. In this study we showed that the method is able to estimate both NH_3 and NO_x emissions which compare well with the current inventory totals. While spatial differences are observed, the potential is there, with some tweaks to lifetime, to get fast and accurate estimates. The results showed, based on our current best knowledge, that the current emission inventories are mostly accurate while there are signs that in some regions, especially for ammonia, emissions are too low. Especially for recent years the satellite derived estimates for NH_3 are much higher than the inventory. Similarly, the monthy NH3 emission estimates show large differences with the satellite based estimates showing a much smaller spring:summer peak ratio, indicating that a larger fraction of the emissions should be attributed to summer time. Variations in effective lifetime can be expected to have some effect, but not enough to cause such a large difference. In the case of NO_x the expected changes in lifetime

do have a clear effect, with the estimated emissions following the solar cycle, with higher photolysis rates in summer much reducing the effective lifetime of NO_2. In the next phase of NKS this is an obvious candidate for improvement. Similarly, improvements to the lifetime assumptions will allow for a more accurate tracking of the year-to-year variations in the emissions. While the uncertainties are large on single years, most of the errors components are consistent between the years. For example, a common difference in lifetime, NO_x:NO_2 ratios, and product bias can be expected, and only small year-to-year variations are predicted for each of those parameters. This ensure that tracking of relative emission changes is possible by adding a light constraint of modelled NO_x:NO_2 ratios and NO_2 lifetimes, with a similar case to be made for NH_3. In the past, the multi-gaussian plume method has already been compared to the divergence method, with both methods providing similar results (Dammers et al., 2023). While for TROPOMI-NO_2 it could favorable to use the divergence method, as it is computationally cheaper and the results are similar, in the case of NH_3 the divergence method is not easily applied. TROPOMI has a continuous observational field (mostly) without gaps allowing for the derivation of the divergence term. In case of the ammonia sounders however, most of the observations are circular with small gaps between the individual measurements. While gaps can be filled and pre-processed, the current application was prone to artifacts, originating from the divergence of the flux term. Future sensors such as the MTG-IRS instrument will however provide the more continous observation field, to which the divergence method could be applied.

6.3.6 Outlook for the gaussian plume method

Within the NKS project the next step is to assess the accuracy of the method through a end-to-end system test, both for NO_x and NH_3. The satellite measurements will be simulated with the LOTOS-EUROS model after which the methodology is applied to the simulated total columns. The benefit of such a method is that all potential causes of uncertainty can be modelled and approximated, linking each to an uncertainty in the final emission estimates. Further complexity can thus be added and understood in a stepwise manner. The main parameters of interest are an improved detailing of the lifetime based on model estimates (both for NO_2 and NH_3) as well as the NO_x:NO_2 ratios for the NO_2 estimates. Another improvement could be a more detailed transport of the plume. The current method only uses a basic gaussian plume description, but more detailed source-receptor models could be implemented to, for example, account for downwind changes of direction. Another point should be the improvement of the CPU efficiency of the method. Currently the Ax=B matrix can become quite large while the effective information is quite limited in the system, as are the links between the soures and receptors. Options are the implementation of super observations or a stepwise local solver, that only optimizes a local region instead of the whole region at once. This also will allow for higher source grid detailing when sensors with better spatial resolution are launched.

6.4 Accuracy of the flux-divergence approach to estimate emissions

The NKS work summarized in this section is described in more detail in a publication submitted to the journal Geoscientific Model Development (Cifuentes et al., 2024).

The Flux Divergence Approach (FDA) is widely used for estimating NO_x emissions from tropospheric NO_2 columns observed by the TROPOMI satellite sensor. However, the accuracy of satellite-derived emissions from this method has been insufficiently investigated, largely due to the scarcity of direct stack emission measurements and the fact that comparisons with traditional bottom-up inventories only offer a general sense of the method's effectiveness and accuracy.

We conducted a study to test the FDA's capability to reproduce known NO_X emissions using syn-

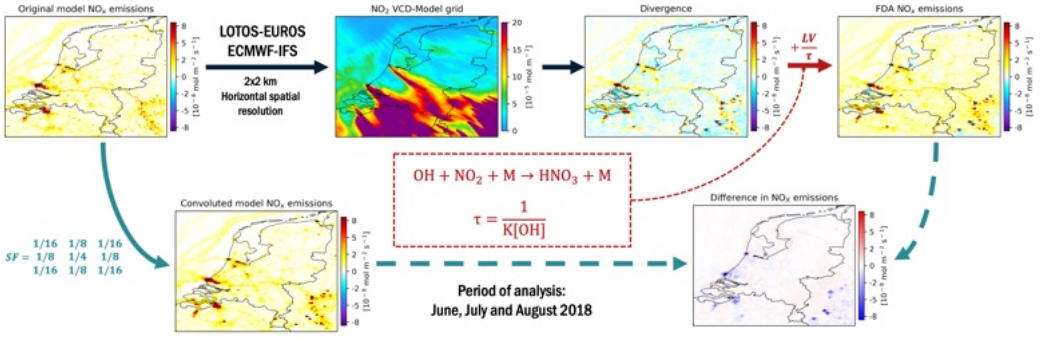


Figure 56: Schematic representation of the evaluation system design to test the FDA. Using the emission inventory as input, the LOTOS-EUROS CTM models the three-dimensional NO_2 concentration in the atmosphere above the Netherlands. Tropospheric columns are generated from the model results resembling an ideal satellite instrument (second panel). The flux divergence is computed (third panel) and the emissions are estimated from this (fourth panel). These emission estimates derived from the columns and the FDA are then compared with the emissions used as input for the model to test the accuracy of the FDA methodology.

thetic (model-predicted) NO_2 satellite column retrievals derived from high-resolution model simulations. A detailed explanation of the methodology and findings can be found in Cifuentes et al. (2024). Here, we provide only a brief summary.

6.4.1 Method

We performed an end-to-end test, schematically illustrated in Figure 56. The process began with known spatially and temporally distributed emissions, which were input into the LOTOS-EUROS model to generate forecasted 3D NO_2 concentrations (latitude, longitude, and vertical distribution). These simulations were produced at high spatial resolution, 0.0125 x 0.025 degree, approximately 1.8 by 1.4 km over the Netherlands. From these concentrations, we derived NO_2 VCDs analogous to those observed by a satellite. We applied the FDA to derive emissions from these synthetic observations and compared them with the original emissions input into the model. This comparison tested the FDA method's ability to accurately reconstruct the input emissions.

We conducted several sensitivity analyses to evaluate the FDA's response to varying input data and configurations. Here, we focus on the performance achieved with the optimal settings, which include using NO_2 columns integrated up to the PBL height, horizontal winds extracted at half the PBL height, and NO_2 lifetimes and NO_x/NO_2 ratios derived from the LOTOS-EUROS simulations.

6.4.2 Results

Overall, the FDA successfully reconstructs the spatial distribution of the original emissions, accurately capturing the distribution of emissions over the Netherlands, including major industrial and urban sources such as the Port of Rotterdam, the city of Amsterdam, and Schiphol Airport, as illustrated in Figure 57. Notably, the correlation coefficient between the predicted and original concentrations at these hotspots is 0.96. Additionally, the method demonstrates the ability to detect emissions from smaller sources, including the road network, inland shipping along the Rhine River, and minor hotspots in Den Helder, Leeuwarden, and Groningen.

The FDA, however, shows an underestimation of emissions at hotspot locations, with a bias of approximately -9%. This discrepancy may arise from several factors, including numerical diffusion introduced by the method used to estimate the divergence term, inaccuracies in the NO_2 lifetime estimations, and potential leakage of NO_2 into the free troposphere, which was not accounted for in

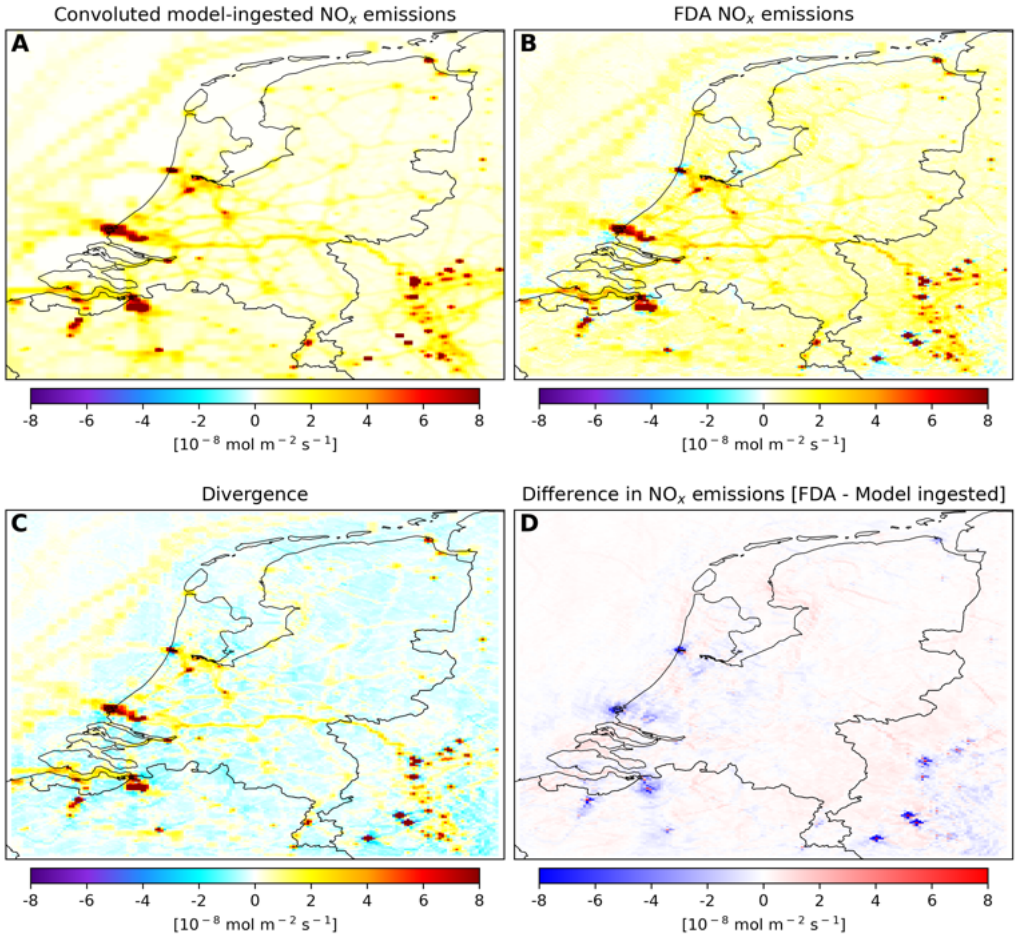


Figure 57: Results from the FDA test using optimal settings. Panel A: The original emission inventory input used by the LOTOS-EUROS model. The input has been smoothed in space, distributing the information over two grid cells in both directions to reflect the resolution of the FDA method. Panel C: The computed divergence using the model-simulated tropospheric columns (synthetic satellite observations). Panel B: The FDA method emissions derived from the divergence and the estimate of the lifetime. Panel D: The difference between the FDA-derived emissions and the emission inventory used as input for the model.

our inversions using PBL-integrated columns. In contrast, the FDA exhibits a slight overestimation in background regions, around 3%, potentially due to biases in the estimated lifetimes.

6.4.3 Conclusions for the FDA emission estimation method

The FDA is a method to estimate NOx emissions (surface fluxes) based on NO_2 column observations. The method simplifies the full three-dimensionality of the transport and chemical reactions in the atmosphere into a two-dimensional continuity equation for the column-integrated amount. A key input for the method is a two-dimensional wind field which advects the column.

In our study we used a high-resolution model (LOTOS-EUROS) to generate a set of synthetic, idealised satellite observations. Subsequently we applied the FDA approach and then compared the FDA emissions with the emissions used as input in the model.

Our study showed that the FDA works well to reproduce the locations and distribution of the emissions in the Netherlands. The spatial resolution that can be achieved with the FDA is about 2x

coarser than the resolution of the measurements, e.g. from the TROPOMI observations with a footprint size of about 5 km we can construct an emission map with a resolution of about 10 km. This factor of 2 may be explained by the fact that the grid cell plus the 4 or 8 nearest neighbor gridcells are needed to compute the divergence (from numerical differences). The FDA method is surprisingly insensitive to the choice of the height level chosen for the two-dimensional wind field, as long as it is representative for the flow in the PBL. Restricting the observations to the PBL column improves the accuracy of the FDA. For more details we refer to Cifuentes et al. (2024).

The FDA is able to produce quantitatively accurate NOx emissions. In the synthetic data setup the FDA captures the magnitude and spatial distribution of the NOx emissions to high accuracy (absolute bias <9 %). However, this is only achieved with accurate information from a model run at high resolution. Neglecting this information can lead to background emission biases of up to 60% and absolute (mainly positive) biases at the hotspots of 20%. The model is needed to produce:

- The variability of NO_2 lifetime along the pollution plume, linked to the OH concentration.
- The NO_X:NO_2 ratio, which also shows a strong variability
- The NO_2 profile shape, which is used to correct the satellite retrievals to account for the fine-scale variability of the NO_2 in three dimensions.
- When steady state is not a good assumption (e.g. in the morning or late afternoon) we also need estimates of the NO_2 concentration increase at a given time.

All three aspects were shown to vary strongly in space and time at the km scale, hence the need for high-resolution model data.

In its original formulation the FDA only needed a wind field as input, which is readily available from weather forecasts. This FDA produced detailed maps of the locations of the emissions, but quantitatively the derived emissions could have considerable biases.

We propose to extend the FDA for NOx emissions from TROPOMI by running in parallel a state-of-the-art CTM, at a resolution of 2 km or better, producing the quantities listed above. Our study showed that using this extra information in the FDA produces more accurate emission estimates.

Note that such a single CTM model run will increase the computational costs. However, it is still much faster than running an ensemble data assimilation or a 4D-Var emission inversion system, which requires multiple (tens of) runs.

6.5 NO_x emissions derived using DECSO and TROPOMI observations

6.5.1 Emission inventories used for DECSO comparisons

For comparison of the emission results in the Netherlands we will use several inventories, often based on official emissions reported to the European Environmental Agency (EEA). The first one is the inventory of national emissions per source category reported under the National Emission reductions Commitments (NEC) Directive of the European Union. Another similar inventory is the Emission inventory reported under the Convention on Long-range Transboundary Air Pollution (LRTAP), which give the country totals of emissions in various source categories. The last one we will use is the European Pollutant Release and Transfer Register (E-PRTR) (EEA, 2012), which is a database of the individual emissions of the biggest industrial facilities (above 0.1 Mg/year) in Europe. The E-PRTR emissions data are reported on an annual basis. From here on we will call those databases simply NEC, LRTAP and E-PRTR.

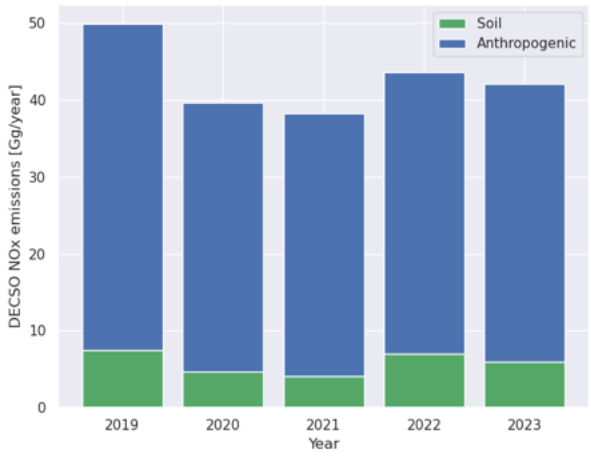


Figure 58: Total NO_x emissions of the Netherlands as derived with the DECSO algorithm.

Besides comparison with these officially reported emissions, we will also compare our emissions to the regional anthropogenic emission inventory CAMS-REG-ANT v5.1 for air quality in Europe (Kuenen et al., 2022) developed for the Copernicus Atmospheric Monitoring Service (CAMS), hereafter called CAMS-REG. For these annual CAMS-REG emissions we use the total emissions regridded from 0.1° x 0.05° to 0.1° x 0.1° and exclude the soil emissions (i.e. agricultural categories), since soil emissions are also excluded in DECSO.

Temporal profiles are also derived in CAMS, which allow us to compare timeseries for monthly averaged values. We will use the Copernicus Atmosphere Monitoring Service TEMPOral profiles (CAMS-GLOB-TEMPO, Guevara et al. (2021, 2023)) for comparison of monthly variations in anthropogenic NO_x emissions. The global emission data version 5.3, called CAMS-GLOB-TEMPO, on a resolution of 0.1° x 0.1° and is hereafter referred to as CAMS-TEMPO.

6.5.2 TROPOMI NO_2 preprocessing for use in DECSO

For the DECSO inversions the measurements are converted to super-observations as described in Ding et al. (2020). To avoid the influence of NO_2 in the free troposphere, governed by processes like lightning, deep convection, aircraft emissions or long-range transport, we adapted the TROPOMI NO_2 retrieval by calculating a partial column up to the 700 hPa level instead of the tropopause level. The stratosphere + free troposphere NO_2 column from the TM5-MP (Tracer Model 5, https://tm5.site.pro/, Williams et al. (2017)) assimilation system are now subtracted from the satellite-observed total column, and new retrieved layer column amounts, air-mass factors and kernels are computed for the surface to 700 hPa layer in the same way as they are computed for the tropospheric column (J. van Geffen et al., 2022).

6.5.3 Results of DECSO for NOx

With the DECSO algorithm on 0.1° we calculate a total of 42.5 (N) Gg per year for the anthropogenic emissions and 7.4 (N) Gg/year for the soil NO emissions in the Netherlands (land only) for the year 2019. In Figure 58 the total emissions for the other years are shown. The country totals reported by LRTAP for anthropogenic emissions in the Netherlands in 2019 is 55.5 (N) Gg/year and by NEC it is 54.2 (N) Gg/year.

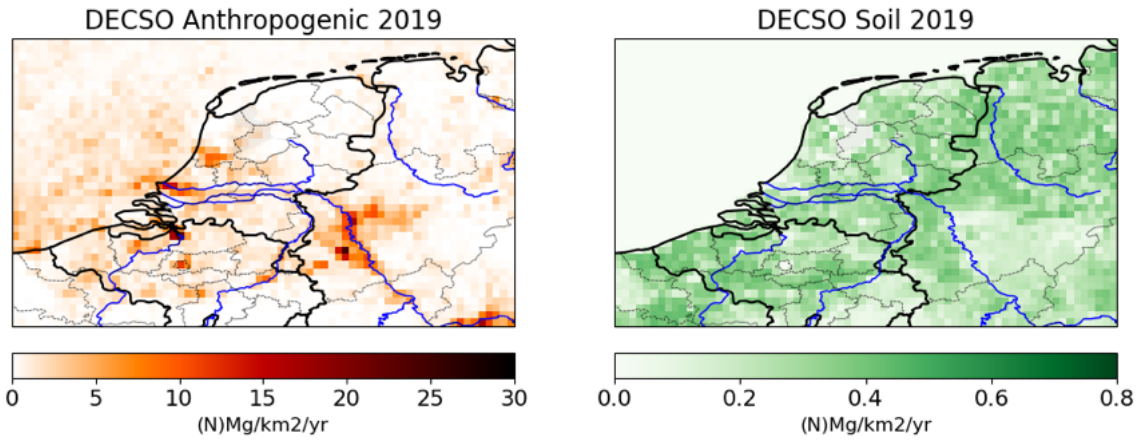


Figure 59: Spatial distributions of the anthropogenic NOx emissions (left plot) and soil NO emissions (right plot) derived by DECSO for the year 2019 from TROPOMI observations.

The spatial distribution of NOx emissions over the Netherlands are shown in Figure 59. The soil NOx emissions are highest in the agricultural regions. The highest anthropogenic NOx emissions are found in the Rotterdam region and regions including Schiphol, Tata steel and Amsterdam. Ship emissions are more spread over large regions of the North Sea with higher emissions close to the port of Rotterdam.

Some (isolated) industrial big emitters in our domain are selected to derive timeseries: Tata Steel factory in IJmuiden, a lignite coal power plant in Weisweiler (Germany), the petrochemical industry in Moerdijk, and the industrial facility Chemelot near Geleen. To evaluate the performance of monitoring emissions from large point sources (LPS), we compare the DECSO emissions with emissions registered in the E-PRTR data base and the CAMS-TEMPO data. Emissions from DECSO are slightly spread to adjacent grid cells because the spatial resolution of the emission field is less than the sampling of the grid cells. To correct for this, we sum the anthropogenic emissions in the 3x3 grid cells around the point source to make sure all emissions are accounted for. For the four selected points, no other big sources exist in these 3x3 grid cell boxes, and soil emissions are excluded. The timeseries are shown in Figure 60. Chemelot and Weisweiler show an overall good match with CAMS-TEMPO, but with DECSO indicating enhancements of emissions in 2022. Moerdijk matches well in 2019 and 2022, but shows lower emissions in 2020. Tata Steel emissions derived by DECSO are generally lower than CAMS-TEMPO and Emissieregistratie (not shown). The inventories E-PRTR and CAMS-TEMPO show differences up to a factor 2. The comparison in general shows that the emissions of big facilities have large uncertainties and high temporal variability.

6.5.4 Summary paper about DECSO v6.3 for Europe on 0.2 degree resolution

In a recent paper (van der A et al., 2024) we made a comparison of DECSO v6.3 vs. CAMS for the European domain on a 0.2 degree resolution. We concluded the following:

- More spatial details became visible as a result of the higher resolution of TROPOMI observations compared to earlier satellite observations.
- In the comparison with CAMS-REG over Europe (where emissions are usually well-known) the deviations are small (within 10%) when looking at country scale. For point sources the spread in the differences is much higher, but no systematic effect is yet found. For cities DECSO show higher emissions, while CAMS-REG is higher for rural regions. Uncertainties in both satellite observations and bottom-up emissions are in general high.

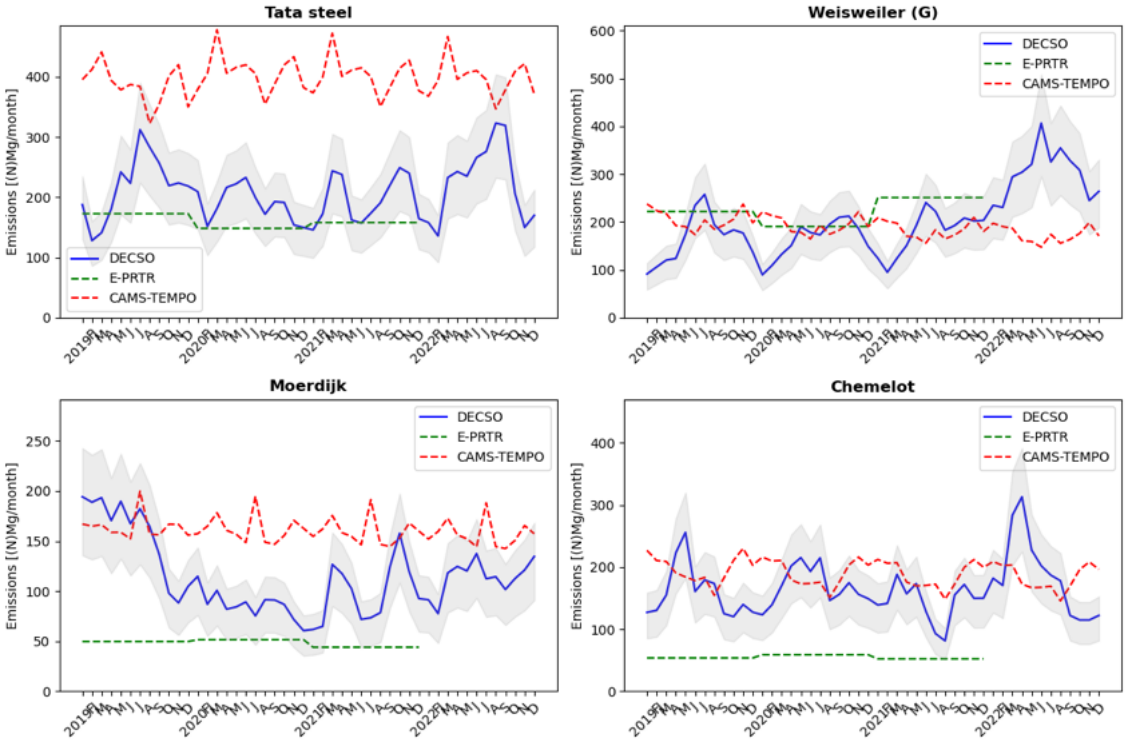


Figure 60: Time series of NO_x emissions derived by DECSO (blue) including the uncertainty range (grey), and reported by CAMS-TEMPO (red), and E-PRTR (green). Shown are results for industrial major emittors TataSteel, the Moerdijk and Chemelot regions, and the Weisweiler power plant.

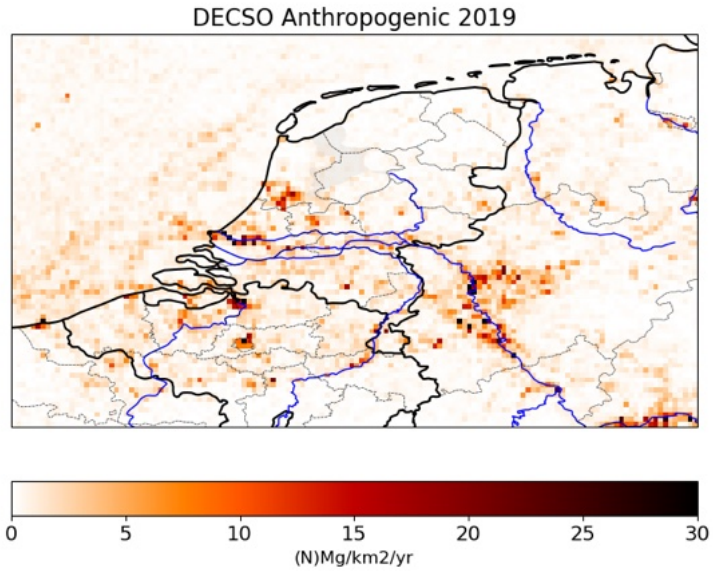


Figure 61: The annual-averaged anthropogenic NO_x emissions for 2019 derived from TROPOMI NO_2 observations using an experimental version of the DECSO algorithm on 0.05 degree resolution.

- In general, we can say that the precision of NO_x emissions given per grid cell (0.2x0.2 degree) is about 8% for annual emissions, 25% for monthly emissions and between 10 and 60% for the daily emissions. When averaging over a larger domain the precision will of course become higher by the square root of the number of grid cells.
- The comparison between CAMS-REG and DECSO emissions showed that DECSO is very similar to CAMS-REG for the spatial distribution and the country totals. While compared to the reported emissions in NEC or LRTAP, DECSO is 7% higher.

6.5.5 Outlook for DECSO NOx

More detailed analyses of the Dutch emissions derived by DECSO will be performed by comparison to the emissions of the Dutch Emission Registration. This currently is work in progress by a small working group of RIVM and KNMI and will be continued by a partnership of TNO/RIVM/KNMI in a new CAMS-NCP project (if the proposal will be approved).

Another development is the increase of the spatial resolution of DECSO to 0.05° x 0.05° (about 3x5 km). Preliminary results are shown in 61 for the year 2019. For this version no super-observations are used since the grid cells are of the same size as the TROPOMI observations and the sensitivity matrices are calculated by using a higher temporal resolution (2 minutes) and a simplified NO_2 lifetime calculation.

6.6 NH3 emissions derived using DECSO and CrIS observations

The work summarized in this section is described in more detail in a publication published in the journal Atmospheric Chemistry and Physics (Ding et al., 2024), and was supported by the EU SEEDS project and NKS-SAGEN.

We applied the existing Daily Emissions Constrained by Satellite Observations (DECSO) algorithm (see section 6.2.4) to NH_3 observations from the Cross-track Infrared Sounder (CrIS) (see section 3.2.4) to estimate NH_3 emissions. Because NH_3 in the atmosphere is influenced by nitrogen oxides (NO_x), we implemented DECSO to estimate NO_x and NH_3 emissions simultaneously. The emissions are derived over Europe for 2020 on a spatial resolution of $0.2° \times 0.2°$ using daily observations from both CrIS and the TROPOspheric Monitoring Instrument (TROPOMI; see section 3.2.6). Due to the limited number of daily satellite observations of NH_3, monthly emissions of NH_3 are reported.

6.6.1 Results

We have run the multi-species DECSO version with NH_3 observations from CrIS-NOAA-20 and CrIS-SNPP respectively to estimate NH_3 emissions over the selected domain of Europe in 2020 (Fig. 62), which is the only year with a full year overlap of NH_3 observations for these two satellites. The total NH_3 emissions derived from observations are about 8 Tg yr^{-1}, with a precision of about 5% - 17% per grid cell per year over the European domain (35-55° N, 10° W-30° E). The spatial distribution of the NH_3 emissions derived from the two satellites agrees well, with small differences (with a relative root-mean-square difference of 1.2%) resulting from deviations in the observed NH_3 columns.

The DECSO emissions are compared with bottom-up emission inventories, see section 6.5.1. spatial distribution of high NH_3 emissions derived from DECSO is similar to that of HTAP, CAMS-REG-ANT, and CAMS-GLOB-ANT but with more local-scale variability and hotspots. The total emissions of DECSO over the European domain are higher than HTAP (4.2 Tg yr^{-1}), CAMS-REG-ANT (4.0 Tg yr^{-1}), and CAMS-GLOB-ANT (5.9 Tg yr^{-1}). We see that emissions from the Netherlands are high in DECSO and in the bottom-up inventories but are missing in the database of E-PRTR. For the countries in eastern Europe (e.g. Poland, Hungary, Romania), the NH_3 emissions derived with DECSO are much higher than those derived from bottom-up inventories.

Fig. 63 shows the monthly NH_3 emissions from DECSO, HTAP, CAMS-REG-TEMPO, and CAMS-GLOB-ANT of the Netherlands, Spain, France, and Poland. We see that the seasonal cycle of NH_3 emissions derived with DECSO is closer to CAMS-GLOB-ANT. HTAP shows the exact same monthly variability for each country. CAMS-REG-TEMPO shows very similar monthly patterns to the ones reported by CAMS-GLOB-ANT, as they both use the same method to derive the temporal profiles for livestock and agricultural soil emissions

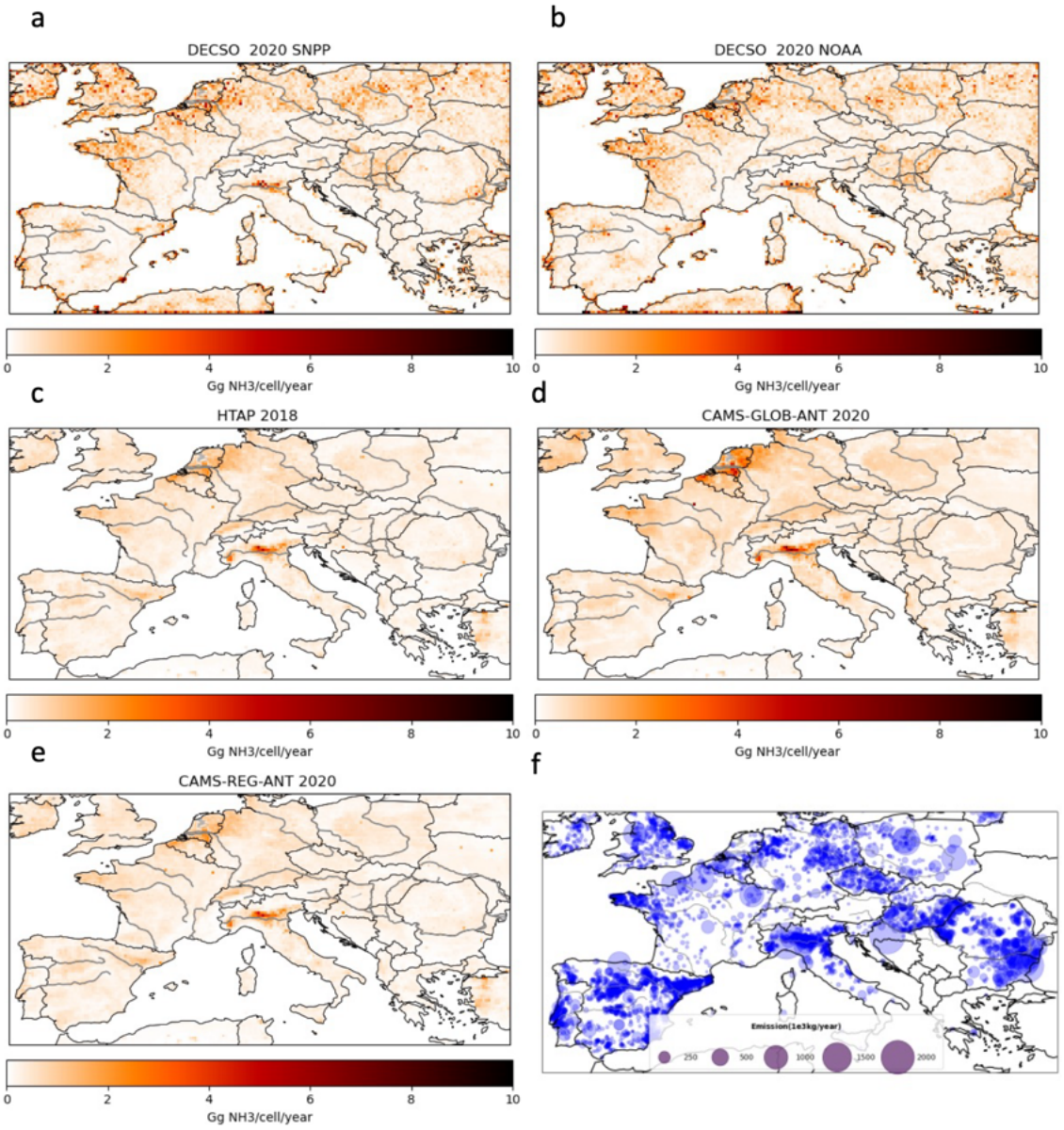


Figure 62: NH_3 emission maps. NH_3 emissions derived with DECSO from (a) SNPP and (b) NOAA-20 in 2020. NH_3 emissions of (c) HTAP in 2018, (d) CAMS-GLOB-ANT in 2020 (e) CAMS-REG-ANT in 2020. (f) The registered point sources of E-PRTR in 2017.

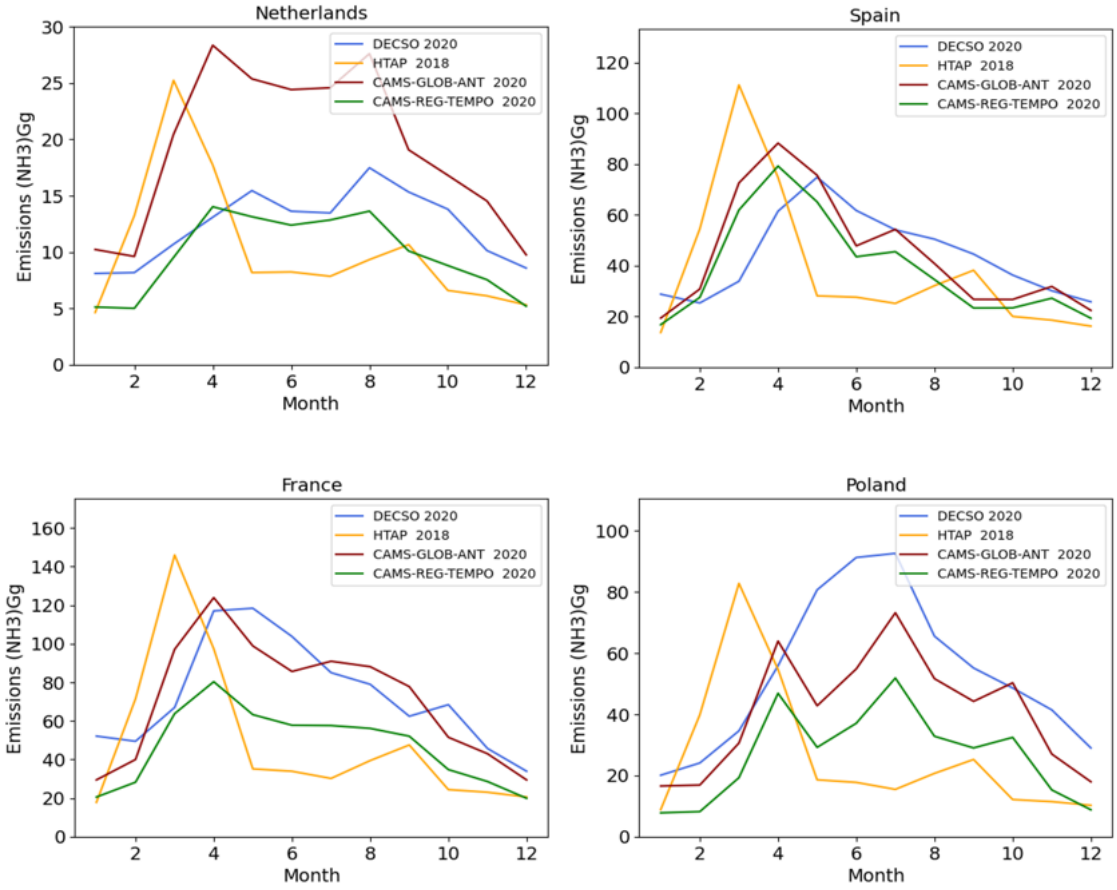


Figure 63: Monthly NH_3 emissions (Gg/month) of DECSO in 2020, HTAP in 2018, CAMS-REG-TEMPO in 2020 and CAMS-GLOB-TEMPO in 2020 for (a) the Netherlands, (b) Spain, (c) France and (d) Poland.

We compare the total NH_3 emissions of DECSO with CAMS-GLOB-ANT, HTAP, and the official national NH_3 emissions of the Netherlands, which are 148, 230, 122, and 123 Gg yr^{-1} respectively. DECSO is lower than CAMS-GLOB-ANT but higher than HTAP and the official NH3 emissions of the Netherlands. Fig. 64 shows the spatial distribution of each inventory in the Netherlands. We see that DECSO captures the high-emission areas and regional distribution over the country.

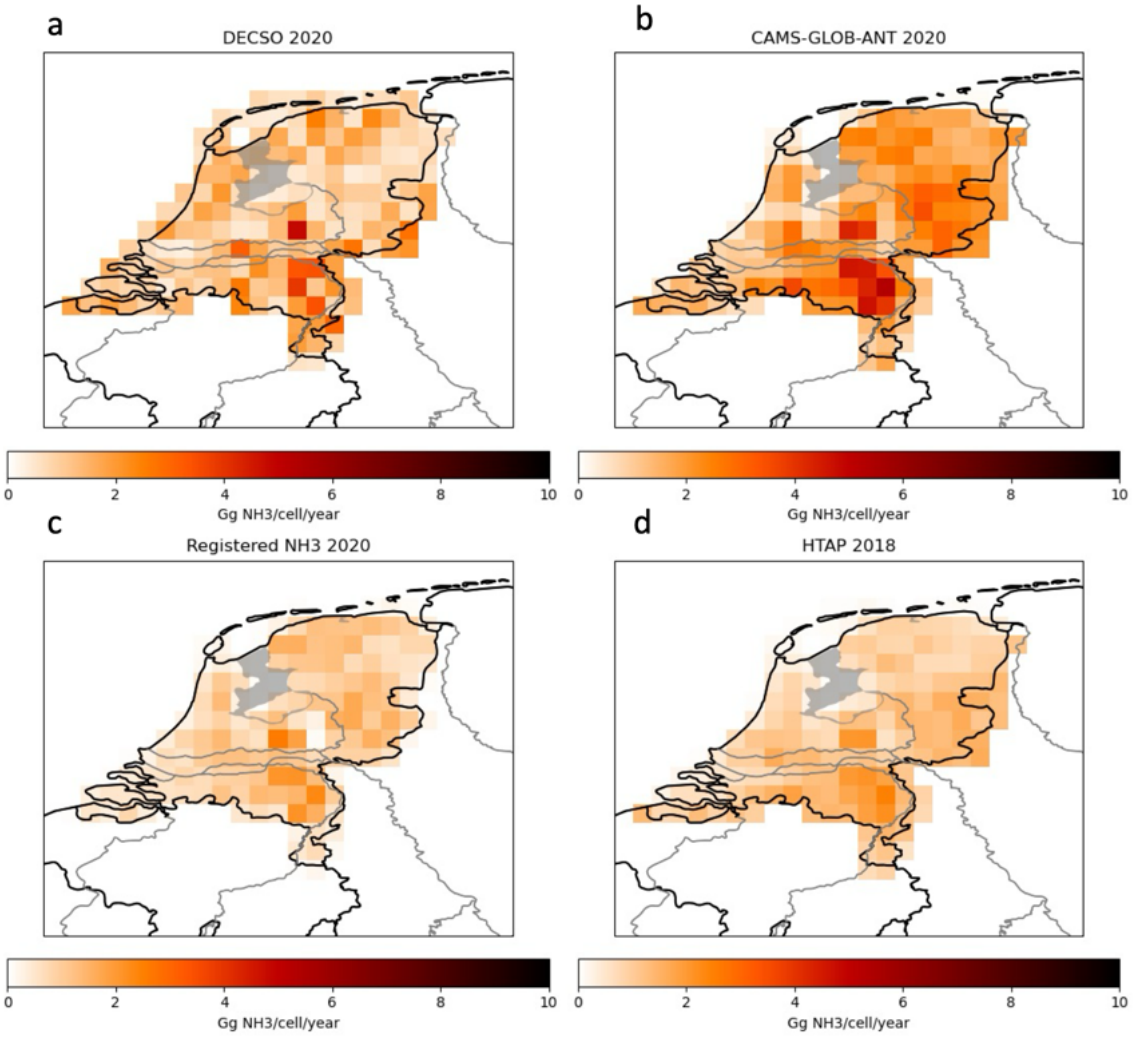


Figure 64: NH_3 emissions in the Netherlands. (a) The averaged NH_3 emissions derived with DECSO from SNPP and NOAA-20. (b) NH_3 emissions of CAMS-GLOB-ANT in 2020. (c) The official national NH_3 emissions for the Netherlands in 2020 (from emissieregistratie.nl). (d) NH_3 emissions of HTAP in 2018.

To further assess the DECSO results using in situ observations from the LML and MAN networks in the Netherlands, we conducted three runs of CHIMERE for the year 2020 using NH_3 emissions from DECSO in 2020, HTAP in 2018, and CAMS-GLOB-ANT in 2020 over the European domain (same as the setup of DECSO). To compare this to the surface NH_3 measurement from the MAN network, we calculated the monthly average of surface NH_3 concentrations from the model simulations. Fig. 65 a–c show the scatter plots of monthly NH_3 concentrations of model simulations against observations for the whole year. We see that modelled NH_3 concentrations with the HTAP emissions are underestimated and that those with the CAMS-GLOB-ANT emissions are overestimated compared to in situ observations. The modelled NH_3 concentrations with DECSO emissions have the lowest absolute bias (modelled concentration minus in situ observations of the MAN network) (Fig. 66). The performance of model simulations is better in summer months (April–September) than in winter months (October–March) (Fig. 65 d-i).

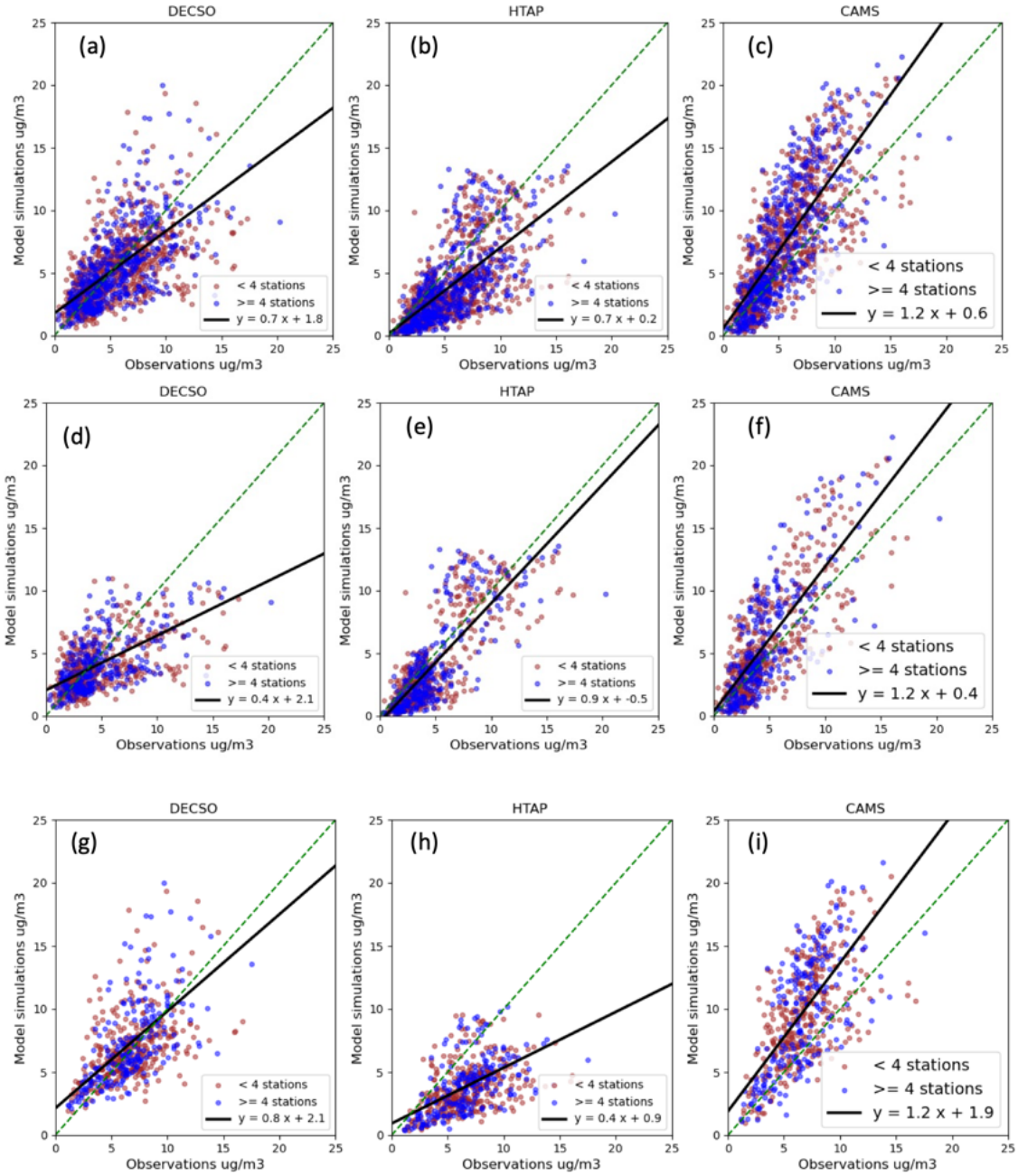


Figure 65: Scatter plots of observations from the MAN network with NH_3 surface concentrations from model simulations with NH_3 emissions from DECSO (left column), HTAP (middle column) and CAMS-GLOB-ANT (right column). (a-c) The scatter plot of data for the whole year for all sites. (d-f) The scatter plot of the data in winter months (October to March). (g-i) The scatter plot of the data in summer months (April to September). Each point presents the model grid cells having the in-situ observations. The red dots mean there are less than four in-situ sites in the grid cells. The blue dots mean there are at least four in-situ sites in the grid cell. The fitted black line is for grid cells with at least four in-situ sites.

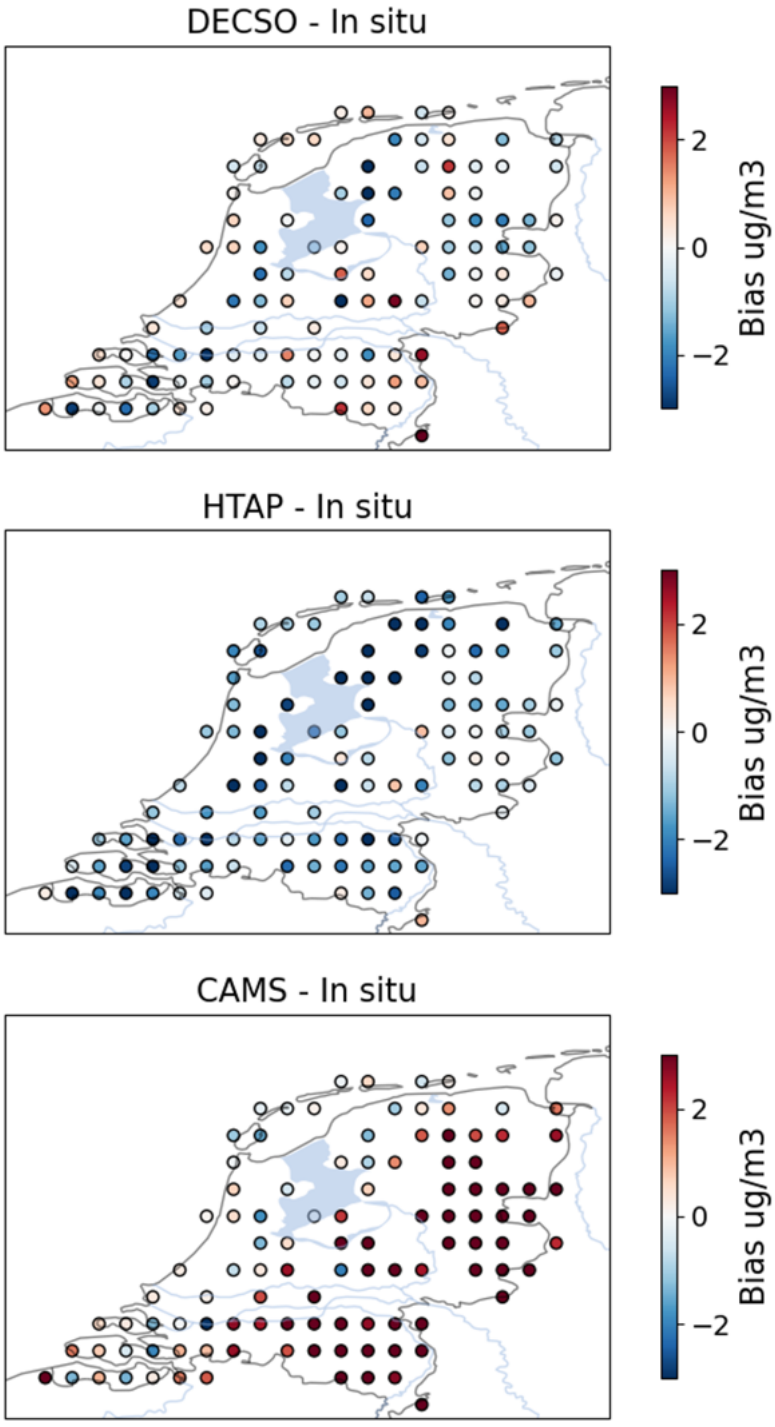


Figure 66: Bias of the model simulated surface concentrations with NH_3 emissions from DECSO (top), HTAP (middle) and CAMS-GLOB-ANT (bottom) compared to the in-situ observations from the MAN network.

6.6.2 Summary: NH3 emissions derived using DECSO and CrIS

We applied the existing DECSO algorithm to NH_3 observations from the CrIS instruments to estimate NH_3 emissions over the European domain on a 0.2 degree resolution. The work is described

in more detail in a recent publication (Ding et al., 2024). We conclude the following:

- In the inversion calculation of NH_3 emissions over the European domain, the influence of changes in NOx emissions needs to be considered in the inversion of NH_3 emissions in DECSO.
- The comparison of the satellite-derived NH_3 emissions from DECSO with independent bottom-up inventories indicates a consistency in terms of magnitude on the country totals, the results also being comparable regarding the temporal and spatial distributions. The spatial distribution of derived NH_3 emissions is similar to the bottom-up inventories, but the annual total emissions derived by DECSO for the whole domain are in general larger than those derived by the bottom-up inventories, especially over Eastern Europe.
- The validation of DECSO over Europe implies that we can use DECSO to quickly derive fairly good monthly emissions of NH_3 over regions with limited local information of NH_3 emissions. The comparison of model simulations using different NH_3 emissions with the MAN in situ observations shows that the simulation using DECSO emissions has a lower bias than simulations based on existing (bottom-up) emission inventories and the performance of model simulations with DECSO is better in summer than in winter.

6.6.3 Outlook

- More detailed analyses of the Dutch emissions derived by DECSO will be performed by comparison to the emissions of the Dutch EmissieRegistratie emission inventory. This currently is work in progress by a small working group of RIVM and KNMI and will be continued by a partnership of TNO/RIVM/KNMI in a new CAMS-NCP project (if the proposal will be approved).
- We are going to use NH_3 observations of IASI in DECSO. The different overpass times of CRIS and IASI will be exploited to evaluate the emission time profiles.
- The DECSO algorithm will be combined with the LOTOS-EUROS model (the current version is based on the CHIMERE model).
- Study the contributions to the error budget of the selected chemical transport model (CTM) used in DECSO and the selected satellite instruments and retrieval approaches for the observations.
- Post-processing using land use data and high spatial resolution bottom-up inventories to derive information of NH_3 emissions on high spatial resolution

6.7 Optimized NH_3 and NO_2 Distributions and Emissions using the LOTOS-EUROS LETKF

6.7.1 LETKF Results for NH_3

An assimilation run using CrIS and IASI NH_3 observations along with TROPOMI NO_2 measurements was performed using the LOTOS-EUROS LETKF for the full year of 2020 on the domain primarily covering Germany, the Netherlands, and Belgium (described earlier in Section 6.2.3). First, a main assimilation run was carried out to produce the hourly perturbation fields (the DC fields), and then a second run was done using only the forward model (but no assimilation) which read in the perturbation fields so that the impact of the optimized emissions on other parameters such as deposition could be evaluated. DC fields are produced by the LETKF on a by-species basis, so separate perturbation fields are produced for NH_3 and NO_2.

We first examine the NH_3 concentration fields pre- and post-assimilation to evaluate the impact of ingesting satellite observations on the model simulation. Plots comparing the mean NH_3 total columns from the background (un-optimized) simulation and the analysis run after the assimilation of the IASI, CrIS and TROPOMI observations are shown in Figure 67. In general, the assimilation of the satellite observations leads to an overall increase in the NH_3 total columns throughout

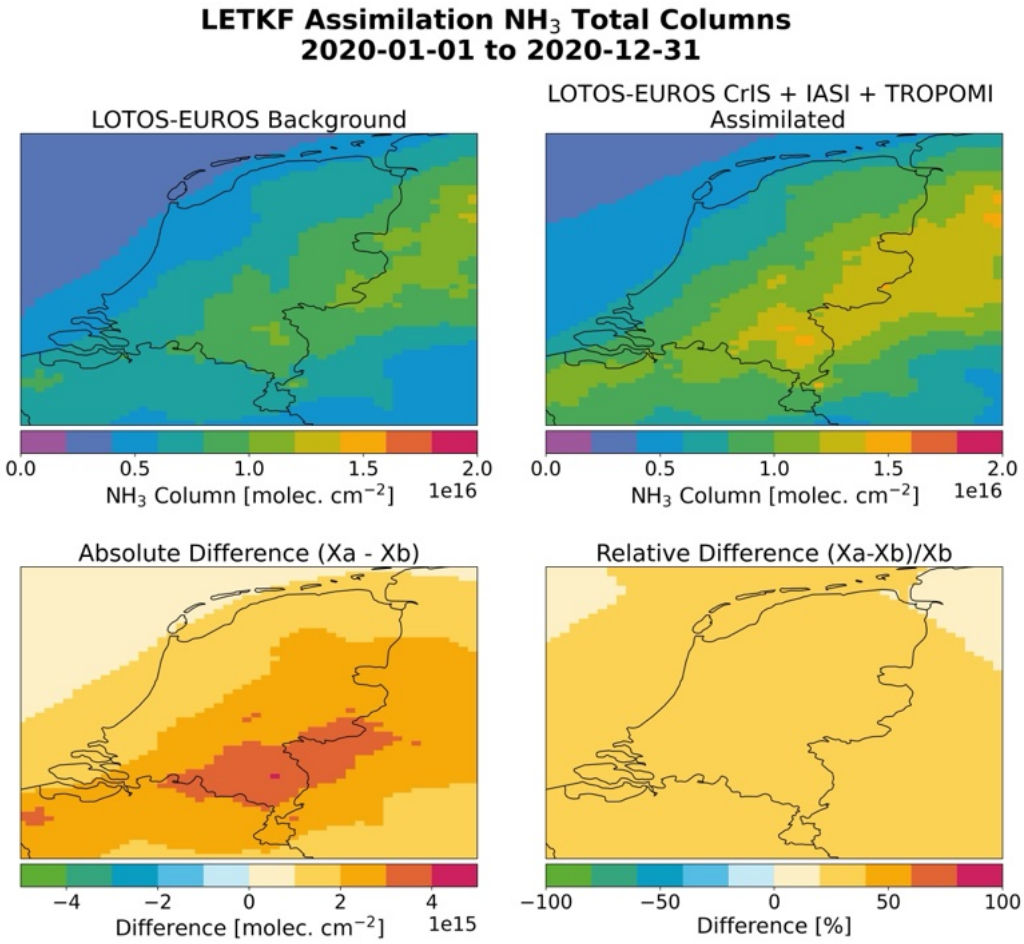


Figure 67: Means calculated over the year of 2020 for (top left) NH_3 total columns from the LOTOS-EUROS background run, (top right) NH_3 total columns from the assimilation run, (bottom left) the absolute differences, and (bottom right) the relative differences (in %).

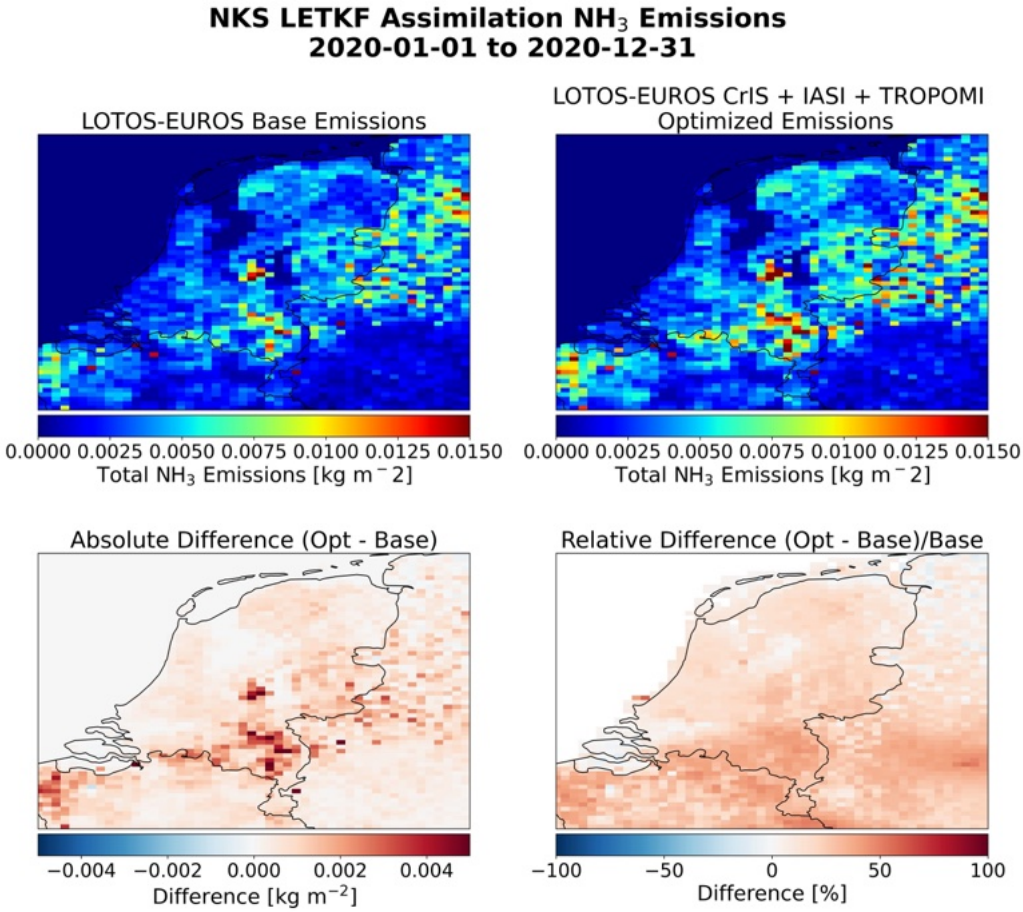


Figure 68: (Top left) 2020 total NH_3 emissions from the LOTOS-EUROS background run, (top right) 2020 total NH_3 emissions from the LOTOS-EUROS assimilation run, (bottom left) the absolute differences, and (bottom right) the mean relative differences (in %).

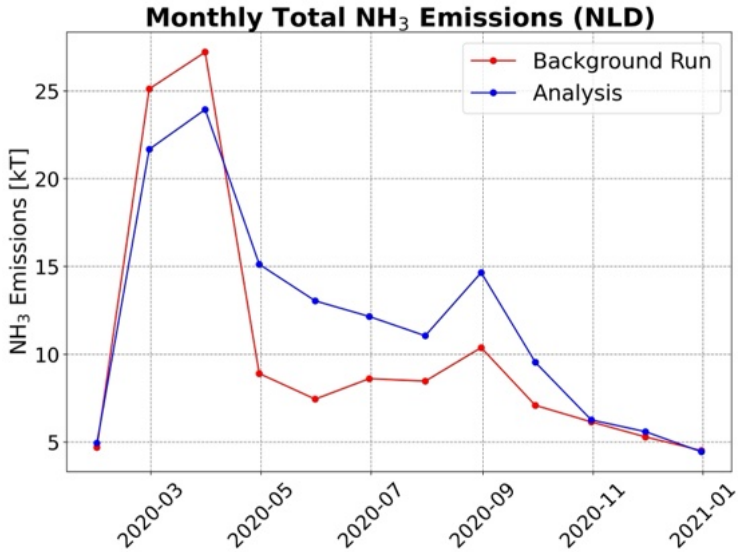


Figure 69: Time-series of the monthly summed NH_3 emissions (in kt) for the Netherlands from the background and analysis simulations.

the model domain on the order of 30-50%. The largest increase in the total columns (in absolute terms), can be seen in the Northwestern region of Germany and throughout the south and eastern part of the Netherlands. These changes reflect the fact that satellite retrievals generally report higher NH_3 columns than the prior model simulation, particularly in regions with intensive agricultural activity. While this adjustment brings the model fields into closer agreement with the satellite observations, it does not necessarily imply that the analysis is closer to the true state, since satellite retrievals themselves can be biased under certain conditions (e.g., cloud contamination, poor thermal contrast). In this sense, the assimilation balances information from both the model and the observations: it corrects for systematic low biases in the background, but the extent to which it approaches "truth" can only be established by comparison with independent reference datasets such as LML and MAN, which will be the focus later in this section.

The background and optimized NH_3 emissions are presented in Figure 68. The spatial patterns of the changes in the emission fields closely follow those seen in the total column concentration fields. The largest absolute increases occur in the southeastern Netherlands, particularly in Brabant and Limburg. The observed increase aligns with the findings of Ge et al. (2020), who showed that incorporating detailed agricultural activity data into the MACC inventory led to higher emissions in this region due to improved spatial allocation, more accurate representation of manure management and application timing, and region-specific regulatory constraints. Their study also highlighted that emissions in this area are disproportionately driven by intensive pig and poultry farming, in contrast to the dominance of dairy cattle across much of the rest of the country. Conversely, the decreases we observe in the north are consistent with their conclusion that refined allocation reduces emissions where earlier inventories overestimated dairy activity. Together, these parallels suggest that the positive adjustments in our assimilation reflect underlying structural biases in the base inventory, both in the spatial distribution and in the livestock-sector partitioning of emissions, rather than being solely artifacts of the assimilation process.

In addition to the annual emission changes, it is also informative to examine the temporal evolution of emissions at finer timescales. A time-series of the NH_3 emissions from the background and assimilation runs are shown in Figure 69. A re-distribution of the emissions from the springtime peak

Region	Emissions (inventory) [kt]	Emissions (analysis) [kt]	Difference [kt]	Relative Difference (%)
Drenthe	7.54	8.70	+1.15	15.24
Flevoland	4.14	4.49	+0.35	8.56
Friesland	13.20	15.44	+2.24	16.94
Gelderland	20.28	25.39	+5.11	25.20
Groningen	8.82	9.68	+0.86	9.79
Limburg	7.68	10.39	+2.71	35.28
Noord-Brabant	22.39	29.88	+7.49	33.47
Noord-Holland	7.06	8.15	+1.09	15.41
Overijssel	14.40	17.35	+2.95	20.49
Utrecht	5.36	6.53	+1.17	21.77
Zeeland	4.98	5.18	+0.20	3.99
Zuid-Holland	8.04	9.18	+1.13	14.09
Whole Netherlands	123.90	150.40	+26.46	21.35

Table 15: 2020 NH_3 emission totals from the background and analysis runs calculated for the individual provinces and the whole Netherlands.

to the summer months is seen. Warm, dry conditions can substantially enhance emissions above climatological norms, while precipitation events may suppress them or trigger post-event peaks (Ge et al., 2023). Such variability may contribute to the year-to-year differences in our assimilation adjustments, and could explain the large emissions increase in 2020, when the Netherlands experienced a very sunny spring and an unusually warm, dry summer.

Additionally, the base and optimized emission totals for the year of 2020 for the individual provinces and the whole country are summarized for NH_3 in Table 15. The NH_3 emissions analysis reveals a nationwide increase of 26.46 kt (+21.4%). Significant regional increases were observed, including Noord-Brabant (+33.5%), Limburg (+35.3%), and Gelderland (+25.2%). All provinces showed positive differences, with Zeeland experiencing the smallest relative increase (+4.0%).

To accurately assess the impact of the assimilation and to evaluate whether the optimized fields actually provide an improvement, it is important to compare the results against an independent dataset. Here, the base and optimized runs were compared against ground-based surface observations from six sites in the LML network in the Netherlands that provide hourly NH_3 measurements. An example time-series plot comparing the monthly mean observed surface concentrations with the simulated background and analysis runs for the sites at Zegveld-Oude Meije and De Zilk-Vogelaarsdreef are shown in Figure 70. At both sites, the temporal representativeness of the NH_3 surface concentrations improved significantly between the background and analysis runs relative

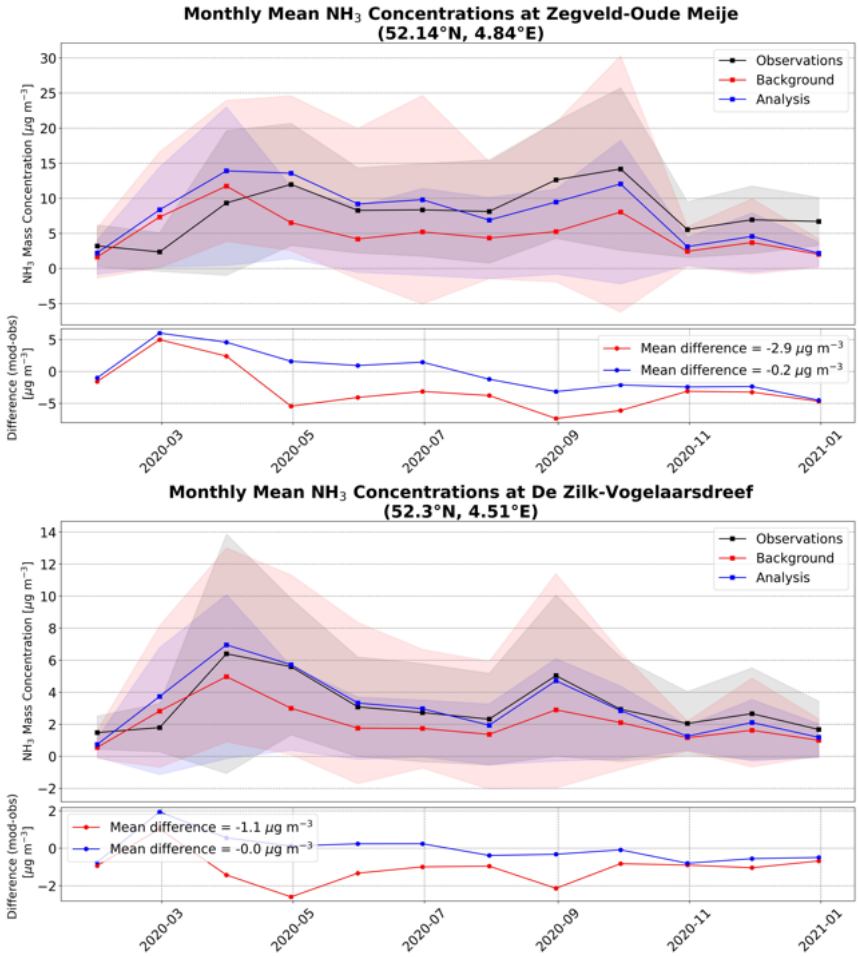


Figure 70: Time-series of monthly mean NH_3 concentrations and differences at (top) the LML site at Zegveld-Oude Meije, and (bottom) the LML site at De Zilk-Vogelaarsdreef.

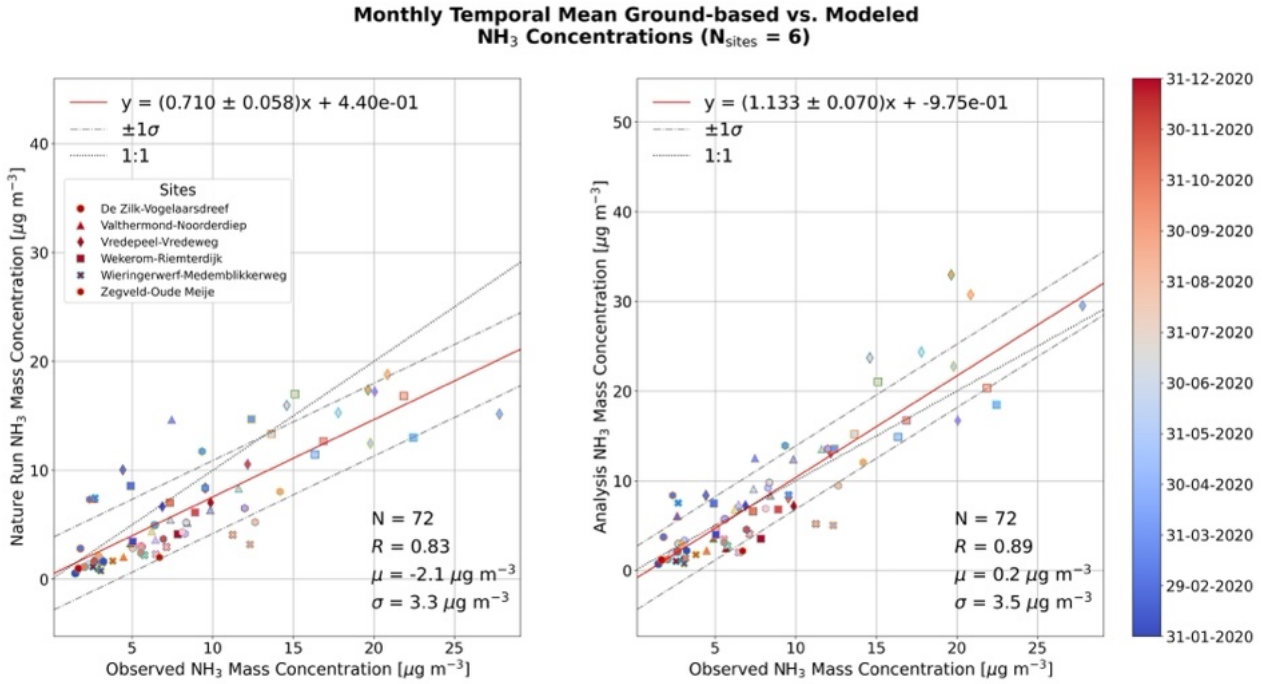


Figure 71: Scatter plot of monthly temporal means of (left) LOTOS-EUROS background NH_3, and (right) LOTOS-EUROS LETKF optimized NH_3 versus LML observed NH_3 surface concentrations.

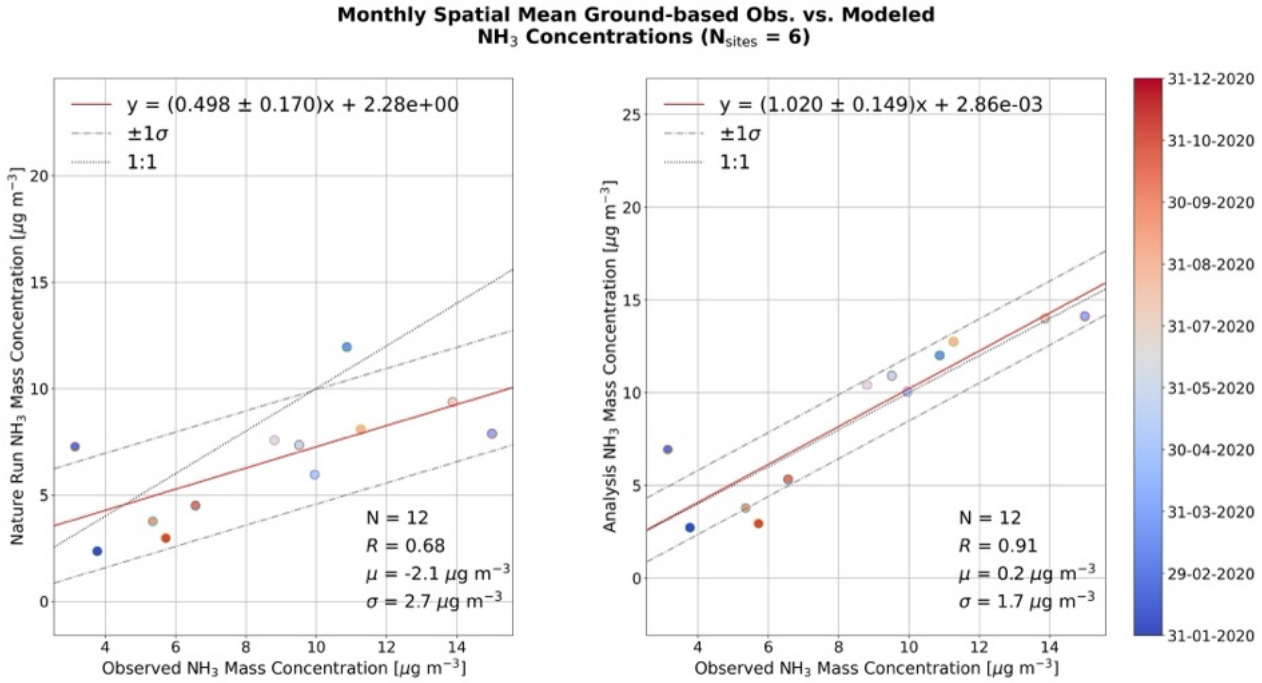


Figure 72: Scatter plot of monthly spatial means of (left) LOTOS-EUROS background NH_3, and (right) LOTOS-EUROS LETKF optimized NH_3 versus LML observed NH_3 surface concentrations. Each data-point represents the mean calculated across all LML sites for a given month.

to the LML observations. The mean bias at Zegveld-Oude Meije reduced from -2.9 ug m^{-3} to -0.2 ug m^{-3}, and at De Zilk-Vogelaarsdreef from -1.1 ug m^{-3} to -0.0 ug m^{-3} in the background and analysis runs, respectively.

A correlation plot of the monthly temporal means is shown in Figure 71 and a similar plot of the spatial means (i.e., all sites averaged for a given month) is provided as Figure 72. A significant improvement in the mean bias from -2.1 ug m^{-3} to 0.2 ug m^{-3} relative to the ground-based measurements can be seen between the background and analysis runs in Figure 71, respectively. The correlation also improved slightly from R = 0.83 to R = 0.89, and the slope of the regression also improved from 0.71 to 1.13. A large improvement in the agreement of the spatial means is observed in Figure 72, with an increase in the Pearson correlation coefficient from R = 0.68 to R = 0.91 and a significant improvement of the regression slope from 0.50 to 1.02 from the background simulation to the analysis run, respectively. The diurnal cycle on NH_3 also showed improvement in most cases in the analysis run relative to the background simulation even though only morning and afternoon satellite overpasses were used. An example of the observed versus modeled diurnal cycles for the LML site at Zegveld-Oude Meije is provided in Figure 73. It can be seen that in the background simulation, the diurnal cycle was largely underestimated relative to the surface observations, however after the assimilation of the satellite observations a much closer agreement between the measurements and the model is found.

Furthermore, to investigate the influence of the assimilation of the satellite observations on the deposition of NH_x, the model runs were compared with dissolved NH_4^+ wet deposition measurements from 8 LML sites. A correlation plot of the spatial means of the background and analysis runs versus the measured dissolved NH_4^+ concentrations is shown in Figure 74. A strong increase in the correlation from R = 0.58 to R = 0.82 was found in the analysis run relative to the background simulation, and a decrease in the spread of the model data relative to the observations is also seen. Broadly, the assimilation led to a decrease in the wet deposition throughout the domain.

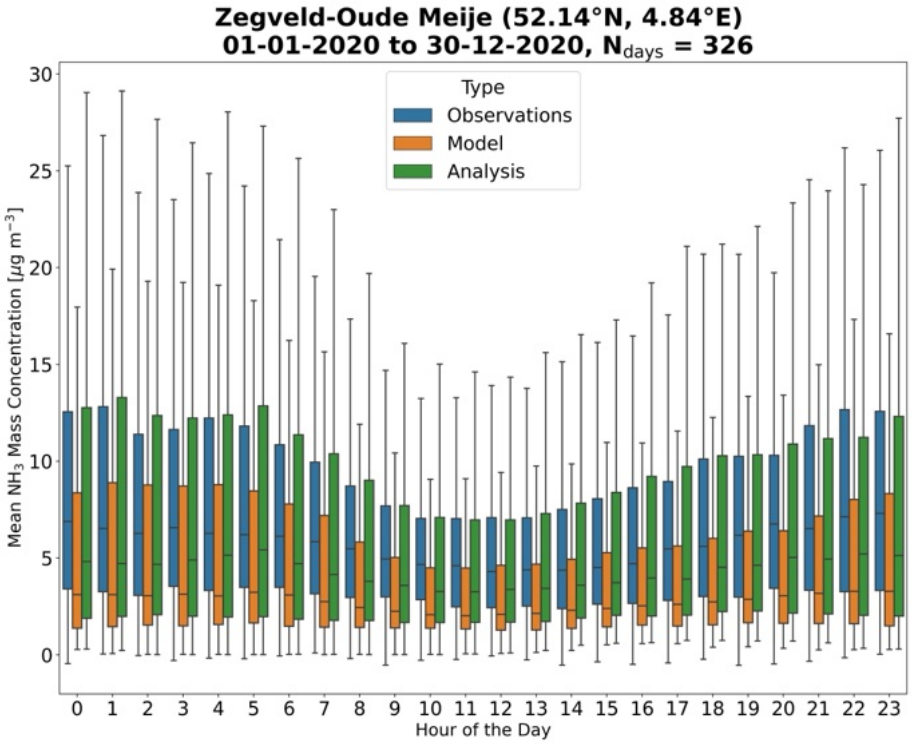


Figure 73: Diurnal cycle of NH_3 for the year of 2020 at the LML site at Zegveld-Oude Meije from the observations (blue), the background model run (orange), and the analysis run (green).

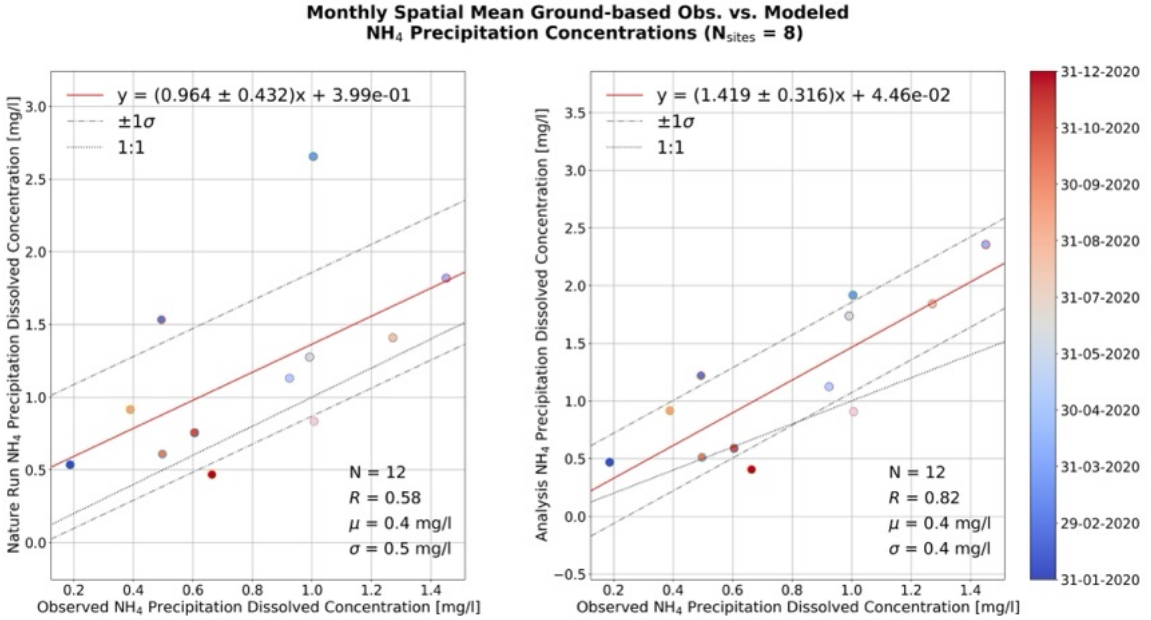


Figure 74: Scatter plot of monthly spatial means of (left) LOTOS-EUROS background dissolved NH_4+ concentration, and (right) LOTOS-EUROS LETKF optimized dissolved NH_4+ concentration versus LML observed dissolved NH_4+ concentration. Each data-point represents a mean across all sites at a given time point.

The background and analysis runs were also compared with the monthly MAN NH_3 surface observations. Observations from a total of 301 individual MAN sites were available during 2020. A scatter plot comparing the monthly mean NH_3 surface mass concentrations from the background and analysis runs with the MAN observations is provided as Figure 75. An improvement in the correlation between the model and the MAN measurements can be seen between the background and analysis runs (R = 0.66 to R = 0.74), however, the bias also increased from 0.1 ug m^{-3} to 3.0 ug m^{-3} in the background and analysis simulations, respectively. This is in contrast with the com-

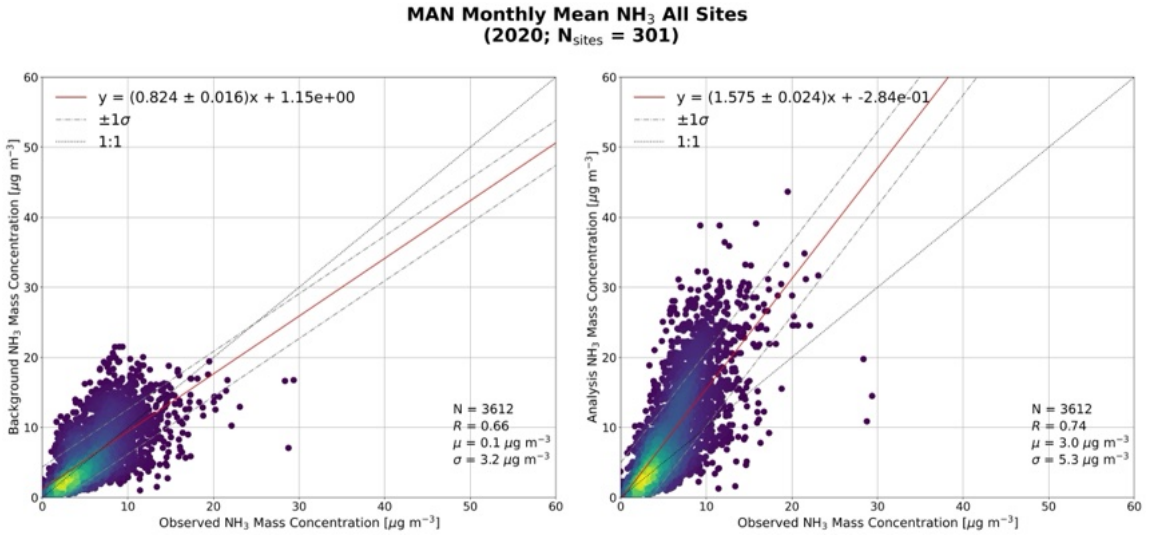


Figure 75: Scatter plot comparing monthly means of observed MAN NH_3 mass concentrations with (left) LOTOS-EUROS background surface NH_3 mass concentration, and (right) LOTOS-EUROS LETKF optimized NH_3 mass concentration. Each data-point represents a mean at a single MAN site.

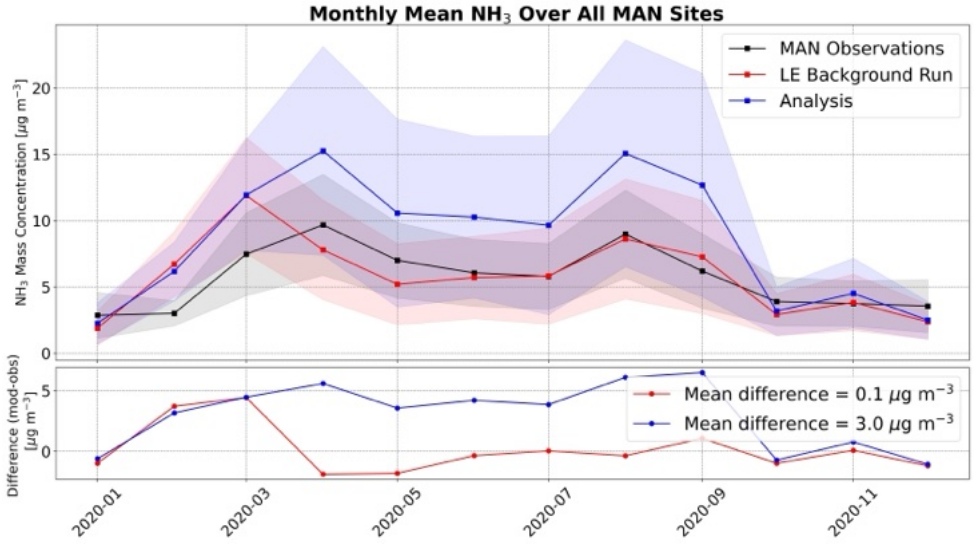


Figure 76: Time-series of monthly mean NH_3 surface mass concentrations and differences calculated across all MAN sites during the year of 2020.

parisons to LML NH_3 observations that showed consistent improvements in both the biases and correlations between the background and analysis runs. This is related to the spatial bias in the MAN data relative to LOTOS-EUROS that was highlighted in Section 2.4.2, but it should be noted that no such spatial bias pattern was found between the model and the LML surface observations, nor in the total column satellite observations from CrIS and IASI.

A time-series plot of the monthly means of MAN and the LETKF simulations calculated across all sites is shown in Figure 76. The shape of the interannual variability of the analysis run more closely matches the shape of the time-series of the mean MAN observations relative to the background run. However, the high bias in the analysis run is apparent, with the largest differences occurring between April and September.

The exact source of the bias between model and the MAN observations is not conclusively known, however some of the bias can likely be attributed to differences in representativeness between the point measurements made by the MAN passive sensors and the coarser model grid. While LML sites are typically located in open fields near source regions, the majority of the MAN sites are situated within Natura2000 areas, often inside forests or near patches of trees, where local canopy effects and fine-scale variability can substantially reduce NH_3 concentrations relative to surrounding agricultural landscapes. It is highly challenging to resolve this sub-grid heterogeneity at the spatial resolution of the model and at the spatial resolution of the NH_3 satellite observations (roughly $\tilde{1}$2-14km at nadir). Further work should be done in the future to better understand the underlying cause of this measurement-model bias.

6.7.2 LETKF Results for NO_2

A similar analysis as was done for NH_3 was repeated for NO_2. A plot comparing the mean NO_2 total columns between the background and analysis runs is provided as Figure 77. A broad decrease in the NO_2 columns on the order of 15-30% can be seen between the background simulation and the analysis run, respectively. Particularly notable decreases in the mean NO_2 columns can be seen in the region around the ports of Rotterdam and Antwerp in Belgium, as well as in the vicinity of the cities of Düsseldorf and Essen in Germany. A map of the background and optimized NO_x emission fields are shown in Figure 78, and the base (inventory) and optimized NO_x emission totals for the year of 2020 for the individual provinces and the whole country are summarized in Table 16.

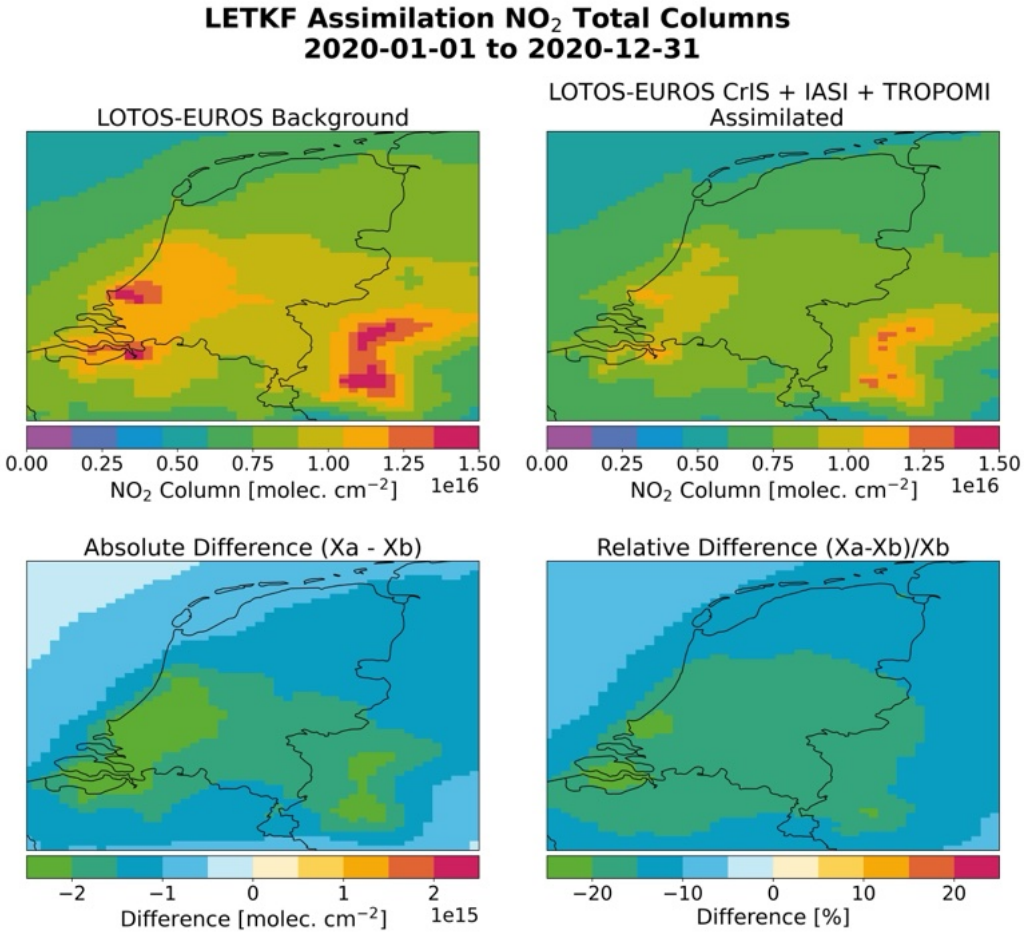


Figure 77: Means calculated over the year of 2020 for (top left) No_2 total columns from the LOTOS-EUROS background run, (top right) NO_2 total columns from the assimilation run, (bottom left) the absolute differences, and (bottom right) the relative differences (in %).

Comparing the base and optimized NO_x emissions indicates a nationwide decrease of 71.28 kt (-27.6%). Substantial reductions were observed in Zuid-Holland (-34.7%), Zeeland (-37.5%), and Noord-Holland (-27.7%). All provinces exhibited negative differences, with Friesland showing a smaller but notable decrease of -23.8%. The spatial distribution of the change in the emissions that can be seen in Figure 77 corresponds directly with the observed decreases in the mean total columns that were shown in Figure 77. Wide-scale decreases in NO_2 concentrations observed from satellites including TROPOMI across Europe due to the 2020 COVID-19 lockdowns have previously been reported (Cooper et al., 2022; Fisher et al., 2024). The decreases in the emissions (and consequently column concentrations) we find here are broadly consistent with the reported COVID-19 related reductions.

The background and assimilation runs for NO_2 were again compared with the same six LML sites that were used for the comparisons with NH_3. The temporal mean correlation plots are shown in Figure 79 and the spatial mean correlations are provided in Figure 80. The temporal mean correlations do not show as drastic of a change between the background and analysis runs as was seen with NH_3 in Figure 71, however a reduction in the mean bias from 1.3 ug m^{-3} to 0.2 ug m^{-3} was observed. The spatial means display a slight improvement in the Pearson correlation coefficients from R = 0.91 to R = 0.95, as well as a reduction in the mean bias from 1.3 ug m^{-3} to 0.2 ug m^{-3}. The diurnal cycle of NO_2 also showed better agreement at most LML sites. An example of the diurnal cycles for the observations, background run, and assimilation run are

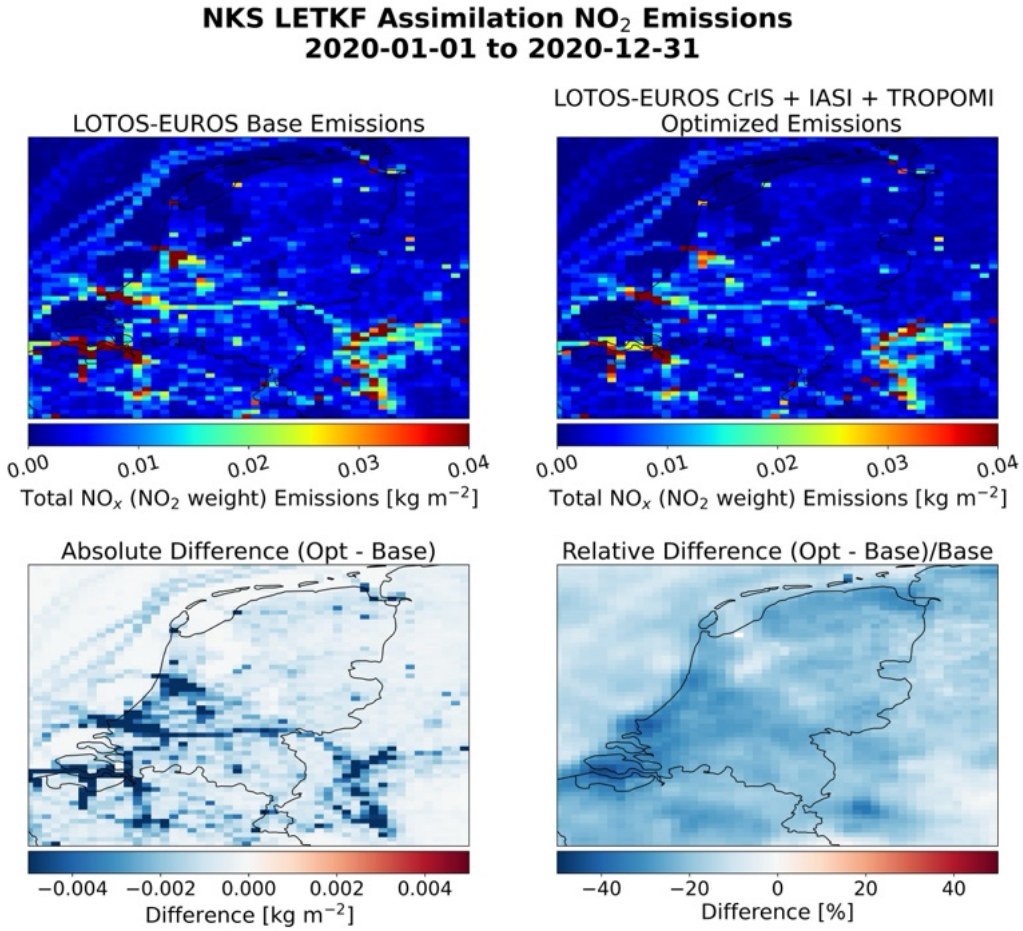


Figure 78: (Top left) 2020 total NO_2 emissions from the LOTOS-EUROS background run, (top right) 2020 total NO_2 emissions from the LOTOS-EUROS assimilation run, (bottom left) the absolute differences, and (bottom right) the mean relative differences (in %).

shown for the site at Zegveld-Oude Meije in Figure 81. The difference between the diurnal cycle of the observations and the analysis run is smaller than for the background run, however the model still appears to be overestimating the NO_2 diurnal cycle at this site particularly at night.

As discussed previously in Section 3.3.2, it should be takin into account that the TROPOMI NO_2 product generally showed low biases in its total column concentrations relative to the MAX-DOAS and PANDORA observations during validation. As such, by assimilating these measurements we may be introducing a low bias in the resulting NO_2 emission and concentration fields. However, in the comparisons against the LML surface measurements, this does not appear to be the case. In general, it can be said that the assimilation broadly improved the agreement with the LML surface NO_2 observations, particularly in the mean spatial and temporal biases, although not to as significant extent as for NH_3.

6.7.3 Summary and Conclusions

An assimilation run using CrIS and IASI NH_3 and TROPOMI NO_2 observations for the year of 2020 was performed using the LOTOS-EUROS LETKF. The LETKF was found to be an effective approach for optimization of the NH_3 and NO_2 concentration, emission, and deposition fields in the LOTOS-EUROS model in comparison to ground-based observations. The NH_3 analysis suggests

Region	Emissions (inventory) [kt]	Emissions (analysis) [kt]	Difference [kt]	Relative Difference (%)
Drenthe	7.18	5.85	-1.34	-18.63
Flevoland	6.28	5.11	-1.17	-18.59
Friesland	12.44	9.48	-2.96	-23.77
Gelderland	29.94	23.12	-6.82	-22.77
Groningen	11.81	8.89	-2.92	-24.71
Limburg	14.12	11.37	-2.74	-19.44
Noord-Brabant	32.40	24.49	-7.92	-24.43
Noord-Holland	31.82	23.00	-8.82	-27.72
Overijssel	13.08	10.57	-2.51	-19.19
Utrecht	12.28	9.00	-3.28	-26.68
Zeeland	19.64	12.28	-7.36	-37.49
Zuid-Holland	67.63	44.18	-23.45	-34.68
Whole Netherlands	258.60	187.30	-71.28	-27.56

Table 16: 2020 NO_x emission totals from the background and analysis runs calculated for the individual provinces and the whole Netherlands. Values are reported in terms of NO_2 weight.

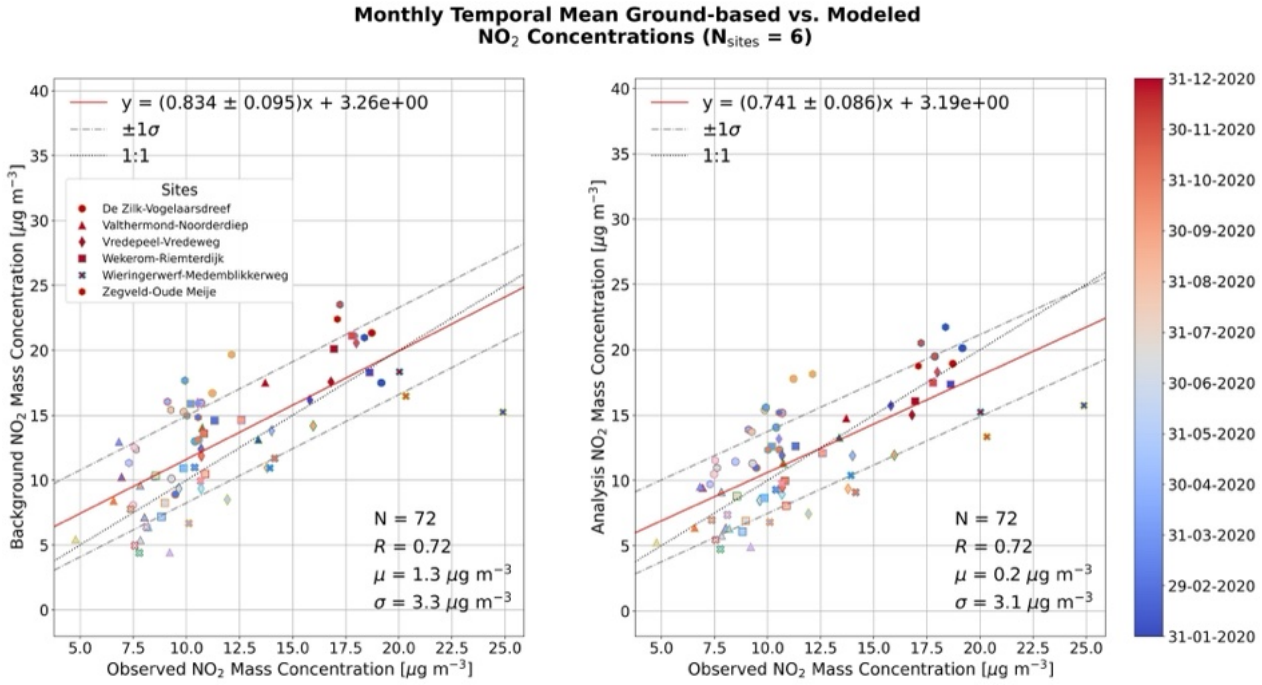


Figure 79: Scatter plot of monthly temporal means of (left) LOTOS-EUROS background NO_2, and (right) LOTOS-EUROS LETKF optimized NO_2 versus LML observed NO_2 surface concentrations.

a nationwide increase of 26.46 kt (+21.4%) in the emissions for the year of 2020 relative to the base inventory. The summer of 2020 was particularly warm in the Netherlands, and it has been well studied that Warmer and drier meteorological conditions can drive higher NH_3 emissions due to increase volatilization and reduced wet deposition. As such, this is likely to be an important

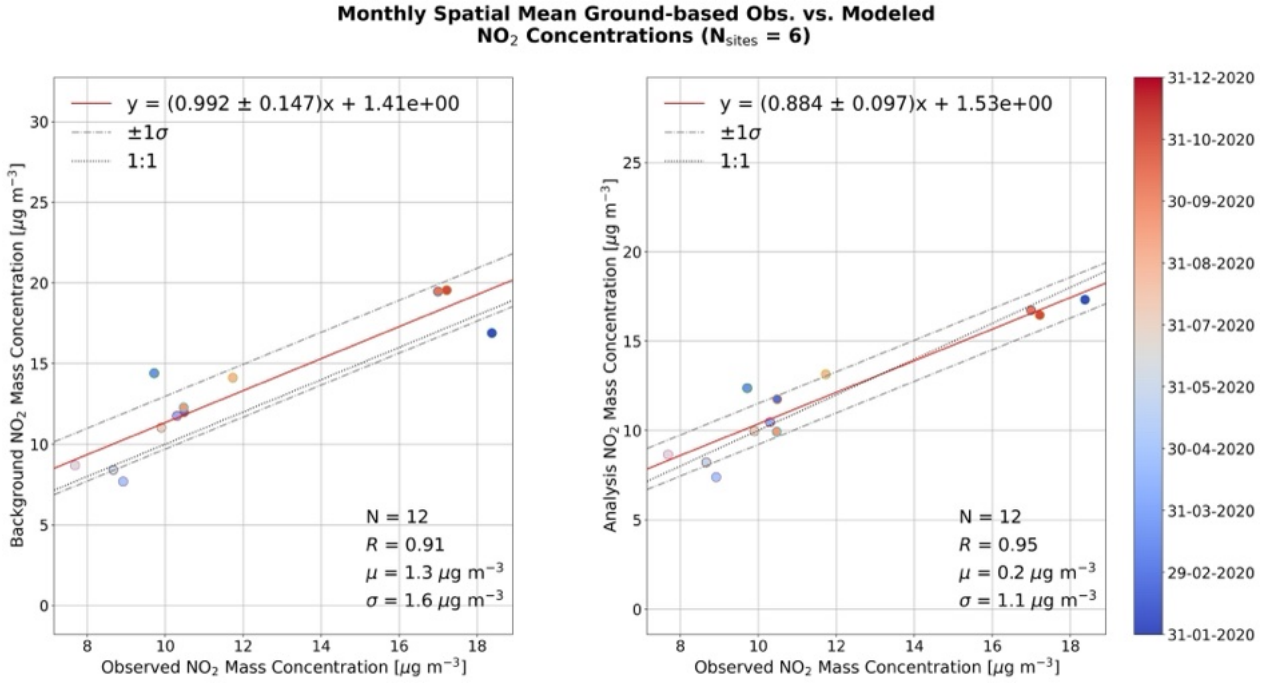


Figure 80: Scatter plot of monthly spatial means of (left) LOTOS-EUROS background NO_2, and (right) LOTOS-EUROS LETKF optimized NO_2 versus LML observed NO_2 surface concentrations. Each data-point represents the mean calculated across all LML sites for a given month.

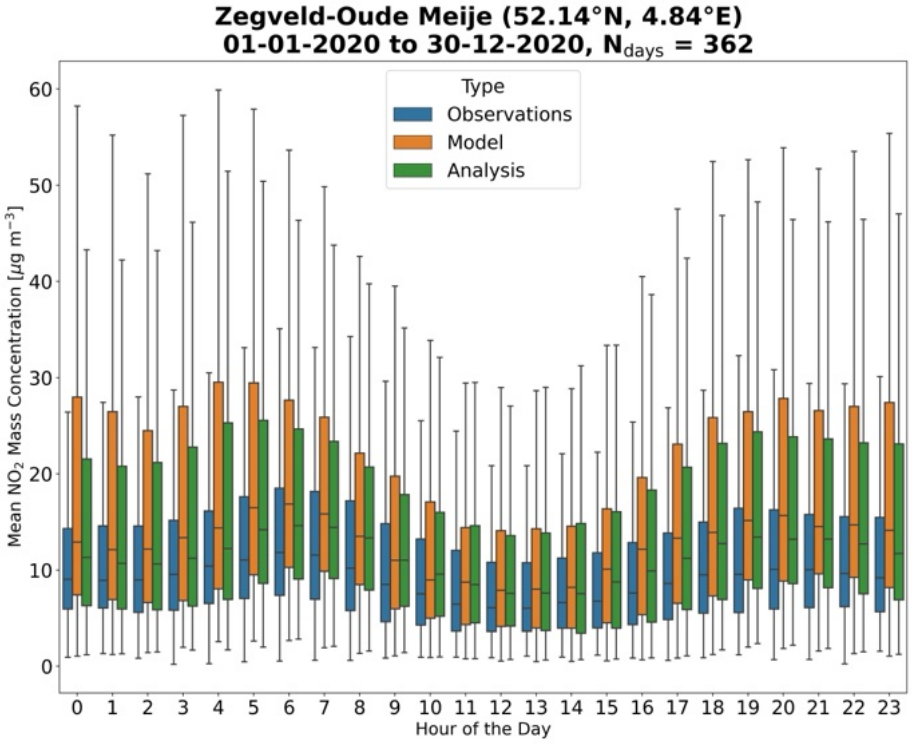


Figure 81: Diurnal cycle of NO_2 for the year of 2020 at the LML site at Zegveld-Oude Meije from the observations (blue), the background model run (orange), and the analysis run (green).

driver of the higher observed emissions in the assimilation run. The model simulations with- and without assimulation were compared against independent ground-based observations from the LML and MAN networks in the Netherlands. For NH_3, strong improvements were observed in the correlations, biases, and diurnal cycles relative to the ground-based LML observations. In the comparisons with the monthly MAN NH_3 surface observations, an improvement in the correlations was seen but a larger bias and slope were observed in the analysis run in comparison to the background model run. Improvements in the spatial representation of NH_4+ wet deposition relative to LML observations was also observed.

For NO_2, broad decreases in the total column concentrations throughout the model domain were found and a corresponding reduction of the NO_x emissions of 71.28 kt (-27.6%) was found. These findings are broadly consistent with previously reported decreases in NO_2 linked to the COVID-19 pandemic and lockdowns in 2020. Comparisons with LML surface NO_2 measurements showed notable decreases in the mean biases post-assimilation, and an increase in the spatial correlation. The diurnal cycle in the model was improved, but was found to still be underestimated relative to the observations, particularly at night, and may need improvement.

6.7.4 Outlook

- Extend the LETKF analysis to additional years (2018–2022) for a broader multi-year comparison.
- We aim to perform a full Observing System Simulation Experiment (OSSE) experiment in the future to more accurately evaluate the performance of the LETKF system.
- Tests on further fine-tuning of the LETKF parameters for improved results.
- Work is currently underway on a publication summarizing the LETKF results.

7 Overview of results and conclusions

In this report, it was investigated how satellite measurements can be used to improve our understanding and monitoring of nitrogen emissions, concentrations, and deposition in the Netherlands. Chemical transport models (CTMs) such as OPS, EMEP4NL and LOTOS-EUROS are critical tools for understanding pollution dynamics, offering forecasts and interpretations of air quality based on emission inventories and atmospheric processes (Skoulidou et al., 2021). To ensure their reliability, these models must undergo rigorous evaluation through comparisons with observational data from both ground-based monitoring networks and satellite platforms. Satellite-derived NO_2 data, such as that provided by TROPOMI, is particularly valuable for model validation due to its city-scale resolution and daily revisit time. Similarly, NH_3 satellite data from the CrIS and IASI instruments is providing useful information at the municipality scale. This capability enables CTM evaluations even in regions with sparse or non-existent ground-based stations. Discrepancies between satellite observations and model outputs can pinpoint areas requiring model improvements, such as adjustments to emission inventories, transport dynamics, or chemical reaction rates, ultimately enhancing model accuracy and predictive capacity. Satellites provide information on concentrations above the surface, which is missing in the ground monitoring networks. This information is important to study the long-range transport of pollution and enables the estimation of emissions.

In the report, an overview of the existing satellite observation datasets for NH_3 and NO_2 was given, as well as a validation of the current products. As there are currently several satellite products with over a decade of data, a long-term model run was initiated for comparison for EMEP4NL, LOTOS-EUROS and OPS (2014-2023). The spatial-temporal distribution as well as the trends in the concentrations fields of all three models were validated with both ground-based data and with satellite data for two models (EMEP4NL and LOTOS-EUROS). Furthermore, NO_2 and NH_3 emissions were derived from satellite data, based on several methods of various degrees of complexity and integration to current AQ models. The OPS model is providing distributions of yearly-mean concentrations at the surface, and as a result can not be directly compared to the satellite vertical column measurements. Therefore the use of satellite data in this study is based on the EMEP4NL and LOTOS-EUROS models, which both model three-dimensional concentration fields with an hourly timestep. Finally, an overview was given where further improvements can be made for the use of satellites and what future developments can be expected.

Long-term model runs:
The long-term model runs were described in detail and compared with in-situ measurements from the MAN and LML networks. All three models showed strong correlations with the NH_3 in-situ measurements but tended to overestimate the NH_3 concentrations in comparison with the MAN observations. The opposite was the case for the measurements of the LML network which tended to be higher than the simulations. Similarly for NO_2 the spatio-temporal distributions compared favorably and high correlations were observed. In the analysis of the NH_3 fields of LOTOS-EUROS, relatively small but significant negative trends were observed at some of the sites in the observations. This negative trend was only matched by the model at Vredepeel. This was not the case in the analysis for OPS and EMEP4NL. In this analysis, no statistically significant trend in the mean measured NH_3 concentrations at the LML sites was found. Contrary to that, the modelled NH_3 surface concentrations from EMEP4NL showed a small but significant trend. The difference originated in the method, in which the analysis for OPS and EMEP4NL merged all 6 datasets instead of directly comparing the modelled and observational data at each location. In case of NO_2, at all sites significant negative trends were observed in both measurements and modelled fields (LOTOS-EUROS, OPS and EMEP4NL). The modelled trends were found to be about half of the observed reductions (independent of the model). In the case of EMEP4NL, a shift between the modelled and measured diurnal cycles of the NO_2 concentrations is observed.

Validation of satellite data

The overall quality of the observations was evaluated by performing a validation study, before the satellite observations were applied for model comparisons and emission estimates. The best way to validate a satellite product is through comparison of measurements made with any instrument capable of making an observation that resembles the same state of the atmosphere as observed by the satellite (a total column amount) as closely as possible. In case of NH_3, the FTIR NH_3 product from the NDACC-IRWG network was used to validate the IASI and CrIS products. In the case of the TROPOMI NO_2 product, data from MAX-DOAS and Pandora instrument networks were used.

Validation of satellite NH_3 data with surface remote sensing observations
On a site-to-site basis the NH_3 satellite products compared well, especially in the higher concentration range, showing favourable results with high correlations. When split into concentration regimes, the results reveal a clear pattern, with CrIS having a positive offset in the lower concentration range while showing consistent results in the higher concentration range. The IASI instruments, meanwhile, show a consistent negative mean relative difference of around 25% for the higher concentration range, while on average the instruments match the FTIR columns for the lower concentration range. The results of this project are broadly consistent with results reported in earlier studies, while the performance of the NH_3 satellite products appears to be improved.

A point of concern is the validity of the results for the typical Dutch conditions. Earlier studies as well as the results in this NKS report indicate that the typical column concentration range in the Netherlands is between $10 - 40 \times 10^{15}$ molecules cm^{-2}, while concentrations at most of the FTIR sites are on average significantly lower. To resolve this point of concern an effort should be made to increase the number of satellite-FTIR co-locations in this concentration range. Several options are possible, but the most direct would be to deploy one or several FTIR instruments in the Netherlands for continuous monitoring and/or field-campaigns.

Validation of satellite NO_2 data with surface remote sensing observations
TROPOMI NO_2 total column measurements at the Cabauw location showed strong correlation with Pandora observations ($R > 0.8$), with on average 7% lower column values that persist throughout the year. These results are obtained when more realistic profiles from high-resolution models (CAMS-Europe, LOTOS-EUROS or EMEP4NL) are used in the TROPOMI retrieval. The scatter plot shows a slope close to 1, and differences that are within the error bars for the instruments. Overall, both Pandora and MAX-DOAS stations in Europe show generally higher columns than those retrieved from TROPOMI. The comparison between Pandora direct sun and MAX-DOAS tropospheric NO_2 columns also reveals systematic differences between the ground measurements. The increased biases observed during the winter months for MAX-DOAS may be linked to differences in the sensitivity of the instruments to NO_2 vertical airmasses, as well as potential inaccuracies in the AMF calculations in the MAX-DOAS and TROPOMI retrievals and/or the partitioning between stratospheric and tropospheric NO_2 in TROPOMI retrieval. A more detailed analysis of these comparisons is the topic of a follow-on NKS-SAGEN study and a new publication to be submitted in 2025.

Evaluation of models with satellite data: NO_2
The TROPOMI satellite product was used for a direct evaluation of EMEP4NL and LOTOS-EUROS NO_2 fields. The comparison of LOTOS-EUROS and EMEP4NL simulated NO_2 columns with TROPOMI data overall demonstrated strong spatial and temporal agreement. In both models a bias was observed between the satellite and modelled fields, in the order of 10-25% depending on location and model. A day-to-day analysis of the EMEP4NL model furthermore, showed that while the model effectively captures pollution plumes, it occasionally simulates plumes that are longer and slightly shifted in direction compared to TROPOMI, suggesting areas for further investigation. These findings highlight the overall reliability of both models, while also pointing to specific aspects that can be further refined.

Evaluation of models with satellite data: NH_3
The modelled NH_3 fields were evaluated with CrIS and IASI instruments. LOTOS-EUROS was

compared to both satellite products, while EMEP4NL was only compared to CrIS satellite data. LOTOS-EUROS and EMEP4NL broadly captured the spatial distribution of NH_3 observed by IASI and CrIS over the Netherlands. However, the models show a consistent low bias relative to both satellite products for all years. The bias is slightly larger for CrIS compared to IASI, although the overall spatial patterns remain similar. Temporally, LOTOS-EUROS captured the overall interannual variability observed by both IASI and CrIS. However, for IASI, the relative bias displays seasonality. The EMEP4NL model captures the spatial distribution of the satellite retrieved NH_3 total columns over the Netherlands well. Moreover, the modelled and retrieved NH_3 total columns show similar interannual variability.

Evaluation of models: synthesis
Based on the results of the surface and column concentration evaluations several things can be concluded. Most variations in concentrations due to emission and deposition processes take place within a timescale of seconds to hours, and thus should be reproduced as close as possible within a model to get the most accurate representation at overpass of the satellite. EMEP4NL and LOTOS-EUROS currently output concentrations (and deposition) at hourly intervals, but under the hood of both models processes run at much shorter time intervals. In case of NO_2 there is a slight mismatch with diurnal and seasonal variations compared to in-situ and satellite measurements, which could be improved by further detailing the temporal emission variations of individual sectors. In the case of NH_3, the base level of the vertical column densities do not fully reflect what is observed by the satellites. The satellite (IASI/CrIS) and ground-based instruments (FTIR) however agree much closer, which points at a problem with the mass in the models. Modelled and measured surface level concentrations however, are relatively close. Several processes can potentially explain the difference in the column total while keeping the surface concentrations relatively constant. The most likely explanation is a combination of an increase in the total emissions (or reduction in deposition) coupled with a change of the diurnal timing of emissions. The diurnal concentration cycle of NH_3 at the LML sites, shows that the overall concentrations can increase throughout the day. A simple multiplier increase of the total emissions without changing the diurnal variation would increase the nighttime concentrations more than daytime due to enhanced vertical mixing (PBL) during the day, so a combination of the two factors is a more logical approach.

Emission estimates, introduction
The validation of the satellite product and evaluation of the current state of the models paved the way for emission estimates based on the satellite observations. Within the NKS project four more commonly used methodologies are applied, with various levels of complexity and model integration.

Two methods can be applied without the direct need of model information. These methods were the divergence and multi-source gaussian plume methods. Both are based on the assumption of a steady state of the estimate, and are more accurate for longer periods of observation.

The other two methods, the DECSO, and LETKF methodologies, are integrated within the CHIMERE and LOTOS-EUROS models, respectively, and are more computationally complex. These methods approximate the state of the atmosphere at the time of overpass and try to derive as much information as possible from single observations.

All four methods were applied to both NH_3 and NOx, with the exception of NH_3, for which the divergence method was found to be incompatible in its current form. Each of the four methods has its advantages and disadvantages, and a preliminary comparison was conducted to highlight the contexts in which they perform well and where shortcomings remain.

Emission estimates, flux divergence approach
The Flux Divergence Approach (FDA) is a method to estimate emissions (surface fluxes) via gradients seen in satellite observations. The method simplifies the atmospheric equations into a two-dimensional mass balance, balancing the sinks and emissions. The results showed that the FDA works well to reproduce the locations and distribution of the emissions in the Netherlands. The ability of the method to produce accurate quantitative result was tested using synthetic obser-

vation obtained from a LOTOS-EUROS model run (Cifuentes et al., 2024). In the synthetic data setup the FDA method captures the magnitude and spatial distribution of the NOx emissions to high accuracy, with an absolute bias < 10% for background values and at the hotspots. However, this is only achieved with accurate information from a model run at high resolution to provide parameters such as atmospheric lifetime, a-priori profile shapes and NO_2/NOx ratios. Neglecting this information can lead to background emission biases of up to 60% and absolute (primarily positive) biases at the hotspots of 20%.

Emission estimates, gaussian plume approach
The multi-source gaussian plume method allows for fast and independent emission estimates to evaluate emission inventories. The method is based around a set of source-receptor calculations, linking individual observations to (potential) source locations. The primary strength of the method is that it is relatively computationally inexpensive, and like the FDA method can detect new emission sources, something that is not always possible with other data-assimilation and inversion methods. In this study we showed that the method was able to estimate both NH_3 and NOx emissions which compared well with the current inventory totals. While spatial differences are observed, there is strong potential that after some tweaks to the chemical lifetimes, fast and accurate emissions estimates can be derived. The results showed, based on our current best knowledge, that the existing emission inventories are mostly accurate in the Netherlands, while there are signs that in some regions, especially for ammonia, emissions are too low. Especially for recent years, the satellite derived estimates for NH_3 are much higher than the inventory, with differences beyond those expected from the uncertainties in input parameters such as lifetime. Similarly, the monthy NH_3 emission estimates show large differences with the satellite-based estimates showing a much smaller spring:summer peak ratio, indicating that a larger fraction of the emissions should be attributed to the summer months. As with the FDA method, a more detailed calculation of a priori input parameters can greatly improve the performance of the gaussian plume method, making it possible to go beyond the province level estimates performed in this study.

In the past, the multi-gaussian plume method has already been compared to the divergence method, with both methods providing similar results (Dammers et al., 2023). While for TROPOMI-NO_2 it could be favorable to use the divergence method, as it is computationally cheaper and the results are similar, in the case of NH_3 the divergence method is not easily applied. TROPOMI has a continuous observational field with relatively few gaps allowing for the derivation of the divergence term. In case of the ammonia sounders however, most of the observations are circular with small gaps between the individual measurements. While gaps can be filled and pre-processed, the current application was prone to artifacts, originating from the divergence of the flux term.

Emission estimates, DECSO inverse modelling, NOx
The DECSO algorithm was applied to both CrIS-NH_3 and TROPOMI-NO_2 observations. In case of NOx the comparison with inventory emissions over Europe showed that the deviations are small (within 10%) when looking at country scale. For point sources the spread in the differences is higher, but no systematic effect is yet found. For cities, DECSO shows higher emissions, while CAMS-REG is higher for rural regions, similar to what was found with the FDA and multi-gaussian plume methods. In general, we can say that the precision of NOx emissions given per grid cell (0.2x0.2 degree) is about 8% for annual emissions, 25% for monthly emissions and between 10 and 60% for the daily emissions.

Emission estimates, DECSO inverse modelling, NH_3
NH_3 emissions were derived using the DECSO-CHIMERE system and CrIS satellite observations. For Europe, the spatial distribution of the derived NH_3 emissions are similar to the bottom-up inventories, but the annual total emissions derived by DECSO for the whole domain are in general larger, especially over Eastern Europe. The validation of DECSO over Europe implies that the method can be used to derive fairly good regional monthly emissions of NH_3. The comparison of model simulations with the MAN station data in the Netherlands shows that a simulation based on DECSO emissions has a lower bias than simulations based on existing (bottom-up) emission inventories and the performance of model simulations with DECSO is better in summer than in

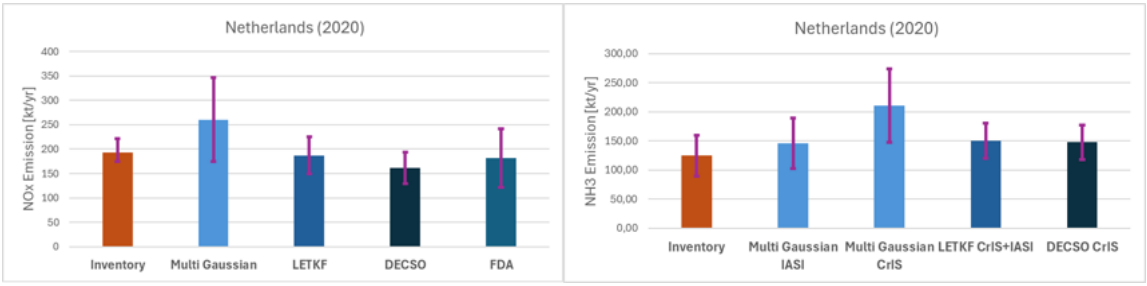


Figure 82: Comparison of satellite based NO_x and NH_3 emission estimates for 2020 with the DECSO, LETKF and multi-gaussian methods and the inventory emissions. As the methodologies are performed at different resolutions, two seperate inventory totals are added, with the left value to be compared with the multi-gaussian and LETKF methods and the right with the DECSO method.

winter. More detailed analyses of the Dutch emissions derived by DECSO will be performed by comparison to the emissions of the Dutch Emission Registration, but this falls outside of scope of this current report.

Emission estimates, assimilation in LOTOS-EUROS

An assimilation run using CrIS and IASI NH_3 and TROPOMI NO_2 observations for the year of 2020 was performed using the LOTOS-EUROS LETKF. The LETKF was found to be an effective approach for optimization of the NH_3 and NO_2 concentration, emission, and deposition fields in the LOTOS-EUROS model in comparison to ground-based observations. In particular for NH_3, strong improvements were observed in the correlations, biases, and diurnal cycles relative to the ground-based LML observations for both NH_3 and NO_2. In the comparisons with the monthly MAN NH_3 surface observations, an improvement in the correlations was seen but a larger bias and slope were observed in the analysis run in comparison to the background model run. Improvements in the spatial representation of NH_4+ wet deposition relative to LML observations was also observed. The NH_3 LETKF analysis suggests a nationwide increase of 26.46 kt (+21.4%) in the emissions for the year of 2020 relative to the base inventory, and a reduction of the NOx emissions of 71.28 kt (-27.6%). These findings for NOx are broadly consistent with previously reported decreases in NO_2 linked to the COVID-19 pandemic and lockdowns in 2020.

Emission estimates, summary

A preliminary side-by-side comparison of the emission estimates (82) shows that all methodologies produce realistic results at a yearly country-wide scale, both for NH_3 and NO_x. The resulting emission estimates are close to those reported by the emission registration, and mostly within the uncertainty limits of the satellite data, inversions and inventory. The amount of spatio-temporal detail derived varies strongly from method to method, due to inherent differences in the underlying models and methodologies. A more in-depth comparison at those scales was outside the scope of this report, but is the recommendation for a follow-up study. Such a comparison could also include a controlled system test, where all methodologies use the same set of pre-constructed observations that are linked to a target set of spatio-temporal varying emissions, which will allow a fair comparison of the strengths and limitations of each method.

Overall summary of results

To summarize, the satellite study indicate that both the NH_3 and NO_2 satellite products have a high enough accuracy to be used more systematically for both model evaluation and emission inversions. The overal information content and accuracy of the satellite observations can help fill the current gap in the spatio-temporal coverage of the existing in-situ measurement networks. The combination of ground-based and satellite data could become the new standard for model evaluation. Ground-based instruments will remain unsurpassed in their quality for measuring local surface concentrations, whereas the satellite instruments fill the spatio-temporal gaps and also provide information on the above-surface concentrations, which is currently completely missed by ground-based networks. In the next few years the quality and coverage of satellite products

is expected to only improve further. The recently launched geostationary satellite instruments (Copernicus Sentinel 4 and Meteosat Third Generation Infrared Sounder IRS) as well as the next generation of IASI satellites will further enhance the spatio-temporal coverage. Current developments in monitoring of CH_4 emission sources by satellite, such as TROPOMI, GHGsat and TANGO, give insight into a potential routes for better detection, quantification, and monitoring of NH_3 and NO_2 sources and integration of such observations into emission inventories.

Short-term research plans.
For the coming two years, there are a number of plans and recommendations for generating products and for conducting further research:

- It is planned to generate emissions for longer periods (e.g. 2018-2024 for TROPOMI).
- Comparing the four different emission estimates is a plan for 2025. The methods and satellite products are relatively independent, and the comparison therefore provides information about uncertainties.
- Evaluate the limit of the inversions based on the current satellite products. Based on an earlier study McLinden et al. (2024) the limit for spatially resolving emission sources should be around 10x10km^2 for TROPOMI-NO_2 and for 20x20km^2 IASI and CrIS-NH_3, which is beyond what was evaluated in this study.
- A start has been made with comparing satellite emissions in detail with the official Emission Registration in the Netherlands. This will be worked on further in 2025. Distinguishing the contributions of different source sectors (such as industry, traffic, shipping, soil emissions) is an important goal here.
- Further optimizing the methods. For example, the use of model data in the flux divergence method and the Gaussian plume method can (will) further improve the quality of the emission satellite products. The application of DECSO for resolutions of 3 by 5 km or better is another aspect where further development is planned.
- The comparison between satellite measurements and ground-based remote-sensing measurements is not yet fully understood. A study has been started to compare TROPOMI NO_2 with MAXDOAS and PANDORA in the Netherlands which we hope to complete in 2025.
- The comparison of the CrIS and IASI products with FTIR will be described in more detail in an upcoming publication that will be completed in 2025.
- In relation to the activities in NKS, a measurement campaign was organized in the summer of 2023 and a mobile FTIR was temporarily placed on Cabauw (near Lopik). The aim of the campaign was to increase the available data under conditions representative for the Netherlands. The results and measurement series of this study will also be completed and published in 2025.
- Further direct evaluation of the processes with satellite data, i.e. emissions, transport and chemical reactions in the models (EMEP4NL, LOTOS-EUROS) by comparing with the NH_3 and NO_2 columns measured by the satellite instruments.
- The coupling between DECSO and LOTOS-EUROS is currently under development. The results of this will provide information on uncertainties related to the chemical transport model used. Emission estimates with DECSO based on IASI NH_3 observations are in preparation.
- In parallel, two studies are ongoing at TNO and WenR to further detail the variability in space and time, and to better map the uncertainties in the ammonia emissions. Model runs with these improved emissions will be compared with the in-situ and satellite data in the future.
- It is planned to prepare the different methodologies for the upcoming new satellite missions launched in 2025-2026.

Potential additions to the Dutch reactive nitrogen monitoring system.
Short term there are two clear advantages of satellite data to the Dutch reactive nitrogen monitoring system which is maintained by RIVM.

- Detection, quantification, and monitoring of emissions derived from satellite observations that can be directly compared to the existing emission inventory (emissieregistratie.nl).
- Quantification of model uncertainties and improvement of parameterizations within the models. This includes the distribution and variability of emissions as a function of time of day, month and year, transport of pollution and realism of pollution plumes, linked to across-border exchange of pollution, and processes that have an impact of lifetime including chemical processes and deposition.

It should be noted that these benefits of atmospheric satellite data can only be directly exploited with a model capable of producing three-dimensional hourly concentration distributions. Models like OPS can indirectly benefit for example from improvements to the emissions and modelled lifetime. In this way, the operational nitrogen monitoring system can be improved.

New planned and launched satellite instruments
In the coming years, a number of new satellite instruments will be launched that will provide even more detail on the distribution of NH_3 and NO_2.

- Sentinel 4 (NO_2) and IRS (NH_3) are geostationary instruments, and will provide hourly observations of column concentrations of NO_2 at 8x8km^2 spatial resolution and half hourly observations of NH_3 at 6x6km^2 pixels.
- The CO2M and TANGO missions will provide NO_2 measurements with improved spatial resolution of 2x2 km^2 and 300x300m^2.
- The IASI-NG satellites will provide NH_3 measurements with improved spectral performance, which will reduce the uncertainty of individual observations and lower the detection limit.

Prospects of longer-term innovation

Several options for innovation for long term that could lead to significant improvement of monitoring and understanding of reactive nitrogen include:

- A dedicated NH_3 satellite instrument, specifically designed for the purpose, with a spatial resolution in the range of 500m - 2km. Earlier studies such as those performed for NITROSAT ESA (2023) indicated the benefits of such a sensor to improve the resolvability of individual sources.
- Alternatively, an imager could also be deployed on a drone or airplane, to further give an insight in detection, quantification and monitoring of individual sources.

8 List of NKS publications and presentations

8.1 NKS publications

The following publications related to work package 3 on use of satellite observations were submitted / accepted with (partial) support from NKS-SAGEN:

- Cifuentes, F., Eskes, H., Boersma, F., Dammers, E., Bryan, C. (2024). Accurate space-based NOx emission estimates with the flux divergence approach require fine-scale model information on local oxidation chemistry and profile shapes. EGUsphere, 2024, 1-40 (accepted for publication). https://doi.org/10.5194/egusphere-2024-2225.
- van der A, R. J., Ding, J., and Eskes, H.: Monitoring European anthropogenic NOx emissions from space, Atmos. Chem. Phys., 24, 7523–7534, https://doi.org/10.5194/acp-24-7523-2024, 2024.
- Ding, J., van der A, R., Eskes, H., Dammers, E., Shephard, M., Wichink Kruit, R., Guevara, M., and Tarrason, L.: Ammonia emission estimates using CrIS satellite observations over Europe, Atmos. Chem. Phys., 24, 10583–10599, https://doi.org/10.5194/acp-24-10583-2024, 2024.

8.2 NKS presentations

The following presentations were held discussing the research results from the NKS SAGEN WP3 satellite activity:

- Henk Eskes & NKS-SAGEN WP3 consortium, "Satellite observations of NO_2: Application to the Netherlands", Symposium Meten en Modelleren van Stikstof Emissies en Depositie, Vrijdag 9 juni 2023, RIVM.
- Enrico Dammers & NKS-SAGEN WP3 consortium, "NH_3 Satellite Observations", Symposium Meten en Modelleren van Stikstof Emissies en Depositie, Vrijdag 9 juni 2023, RIVM.
- Jieying Ding, "NH3 emissions derived from CRIS observations over Europe", Poster presentation, GEIA conference, Brussels, 21-23 June 2023.
- Henk Eskes, "Observing trace gas emissions from space", presentatie op de Emissieregistratie heidag, Ede, 23 januari 2024.
- Henk Eskes & NKS-SAGEN WP3 consortium, "Model validation and emission estimation for NO2/NOx and NH3 over the Netherlands using satellite observations", LOTOS-EUROS Workshop 2024, 25-26 januari, TNO, Utrecht.
- Felipe Cifuentes Castaño, "Performance evaluation of the flux divergence approach for estimating NOx emissions using simulated TROPOMI-like NO2 data", ESA ATMOS 2024 (Poster presentation), Bologna, Italy, 1-5 July 2024.
- Henk Eskes & NKS-SAGEN WP3 consortium, "Gebruik satellietmetingen voor stikstof emissies, concentraties en depositie", Stikstof Depositie Symposium, Driebergen, 6 september 2024.

- Felipe Cifuentes Castaño, "Accurate NOx emission estimates with the flux divergence method and TROPOMI NO2 columns require fine-scale information on local oxidation chemistry", Workshop on high-resolution modeling of urban air quality and greenhouse gas emissions (Oral presentation), Utrecht, Netherlands, 26 September 2024.
- Felipe Cifuentes Castaño, "Accurate space-based NOx emission estimates with the FDA require fine-scale model information on local oxidation chemistry and profile shapes", Plume Inversion seminar from the Data Assimilation, Forecasting Applications and Emissions (DAFAEMIS) group from at the Barcelona Supercomputing Centre (BSC) (Oral), Barcelona, Spain, 24 October 2024.
- Ronald van der A, "Nieuwe ontwikkelingen voor emissiemonitoring met satellietmetingen", DGMI visit to KNMI, 28 November 2024.
- Enrico Dammers, Remote sensing of NH3 and NO2, FONDA Spring School, Aveiro, Portugal , 4-15 March 2024.
- Tyler Wizenberg, "Co-assimilation of NH_3 and NO_2 satellite observations with the LETKF methodology in the LOTOS-EUROS model", IASI Conference 2024, Nancy, France, 2-6 December 2024.
- Dammers, E. and the FTIR-NH3 and satellite NH3 product teams: Validation of the CrIS, IASI, AIRS and GOSAT-NH3 products , EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-9164, https://doi.org/10.5194/egusphere-egu24-9164, 2024.
- Enrico Dammers, "Co-assimilation of NH_3 and NO_2 satellite observations with the LETKF methodology in the LOTOS-EUROS model", ITM 2024, Copenhagen, Denmark, 14-18 October 2024.

8.3 Other

- A whitepaper was written by the NKS-SAGEN WP3 teams: "Het gebruik van satellietmetingen voor het monitoren van stikstof in Nederland, Whitepaper", a publication from KNMI, TNO, CML, WUR and RIVM, collaborating within the national knowledge programme on reactive nitrogen (Nationaal Kennisprogramma Stikstof, NKS). Date: November 2024.
- The following NKS-SAGEN paper will be submitted to BioGeoscience:
  Malte Lessmann, Dennis Walvoort, Jan Cees Voogd, Twan Cals, Leo Renaud, Hans Kros, Uncertainties in nitrogen application from manure and fertilizers and ammonia emissions from agriculture in the Netherlands across different spatial scales, preprint 2024.

References

Beirle, S., Borger, C., Dörner, S., Eskes, H., Kumar, V., Laat, A. D., & Wagner, T. (2021). Catalog of NOx emissions from point sources as derived from the divergence of the NO2 flux for TROPOMI. *Earth System Science Data*, *13*, 2995–3012. https://doi.org/10.5194/ESSD-13-2995-2021

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., & Wagner, T. (2019). Pinpointing nitrogen oxide emissions from space. *Science Advances*, *5*. https://doi.org/10.1126/SCIADV.AAX9800

Beirle, S., Borger, C., Jost, A., & Wagner, W. (2023). Improved catalog of NO_x point source emissions (version 2). *Earth System Science Data*, *15*(7), 3051–3073. https://doi.org/10.5194/essd-15-3051-2023

Berkhout, A. J., Swart, D. P., Volten, H., Gast, L. F., Haaima, M., Verboom, H., Stefess, G., Hafkenscheid, T., & Hoogerbrugge, R. (2017). Replacing the amor with the minidoas in the ammonia monitoring network in the netherlands. *Atmospheric Measurement Techniques*, *10*(11), 4099–4120.

Bessagnet, B., Pirovano, G., Mircea, M., Cuvelier, C., Aulinger, A., Calori, G., Ciarelli, G., Manders, A., Stern, R., Tsyro, S., García Vivanco, M., Thunis, P., Pay, M.-T., Colette, A., Couvidat, F., Meleux, F., Rouïl, L., Ung, A., Aksoyoglu, S., ... White, L. (2016). Presentation of the eurodelta iii intercomparison exercise – evaluation of the chemistry transport models' performance on criteria pollutants and joint analysis with meteorology. *Atmospheric Chemistry and Physics*, *16*(19), 12667–12701. https://doi.org/10.5194/acp-16-12667-2016

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781118619193

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N.-E., Herold, M., & Fritz, S. (2020). *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe (V3.0.1)*. Zenodo. https://doi.org/10.5281/zenodo.3939050

Cao, H., Henze, D. K., Zhu, L., Shephard, M. W., Cady-Pereira, K., Dammers, E., Sitwell, M., Heath, N., Lonsdale, C., Bash, J. O., Miyazaki, K., Flechard, C., Fauvel, Y., Kruit, R. W., Feigenspan, S., Brümmer, C., Schrader, F., Twigg, M. M., Leeson, S., ... Capps, S. L. (2022). 4d-var inversion of european nh3 emissions using cris nh3 measurements and geos-chem adjoint with bi-directional and uni-directional flux schemes [e2021JD035687 2021JD035687]. *Journal of Geophysical Research: Atmospheres*, *127*(9), e2021JD035687. https://doi.org/https://doi.org/10.1029/2021JD035687

Cede, A., Tiefengraber, M., Gebetsberger, M., & Lind, E. S. (2023). *Pandonia Global Network Data Products Readme Document, Version 1.8-6, 31st Dec 2022*. LuftBlick, NASA/GSFC, Virginia Tech. https://www.pandonia-global-network.org/wp-content/uploads/2022/12/PGN_DataProducts_Readme_v1-8-6.pdf

Chan, K. L., Wiegner, M., van Geffen, J., De Smedt, I., Alberti, C., Cheng, Z., Ye, S., & Wenig, M. (2020). MAX-DOAS measurements of tropospheric NO_2 and HCHO in Munich and the comparison to OMI and TROPOMI satellite observations. *Atmospheric Measurement Techniques*, *13*(8), 4499–4520. https://doi.org/10.5194/amt-13-4499-2020

Cifuentes, F., Eskes, H., Boersma, F., Dammers, E., & Bryan, C. (2024). Accurate space-based NO_x emission estimates with the flux divergence approach require fine-scale model information on local oxidation chemistry and profile shapes. *EGUsphere*, *2024*, 1–40. https://doi.org/10.5194/egusphere-2024-2225

Clarisse, L., Franco, B., Van Damme, M., Di Gioacchino, T., Hadji-Lazaro, J., Whitburn, S., Noppen, L., Hurtmans, D., Clerbaux, C., & Coheur, P. (2023). The iasi nh3 version 4 product: Averaging kernels and improved consistency. *Atmospheric Measurement Techniques*, *16*(21), 5009–5028. https://doi.org/10.5194/amt-16-5009-2023

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., & Coheur, P.-F. (2009). Monitoring of atmospheric composition using the thermal infrared iasi/metop sounder. *Atmospheric Chemistry and Physics*, *9*(16), 6041–6054. https://doi.org/10.5194/acp-9-6041-2009

Colette, A., Collin, G., Besson, F., Blot, E., Guidard, V., Meleux, F., Royer, A., Petiot, V., Miller, C., Fermond, O., Jeant, A., Adani, M., Arteta, J., Benedictow, A., Bergström, R., Bowdalo, D., Brandt, J., Briganti, G., Carvalho, A. C., ... Rouıl, L. (2024). Copernicus atmosphere monitoring service – regional air quality production system v1.0. *EGUsphere*, *2024*, 1–92. https://doi.org/10.5194/egusphere-2024-3744

Colette, A., Andersson, C., Manders, A., Mar, K., Mircea, M., Pay, M.-T., Raffort, V., Tsyro, S., Cuvelier, C., Adani, M., Bessagnet, B., Bergström, R., Briganti, G., Butler, T., Cappelletti, A., Couvidat, F., D'Isidoro, M., Doumbia, T., Fagerli, H., ... Wind, P. (2017). Eurodelta-trends, a multi-model experiment of air quality hindcast in europe over 1990–2010. *Geoscientific Model Development*, *10*(9), 3255–3276. https://doi.org/10.5194/gmd-10-3255-2017

Cooper, M. J., Martin, R. V., Hammer, M. S., Levelt, P. F., Veefkind, P., Lamsal, L. N., Krotkov, N. A., Brook, J. R., & McLinden, C. A. (2022). Global fine-scale changes in ambient no_2 during covid-19 lockdowns. *Nature*, *601*, 380–387. https://doi.org/10.1038/s41586-021-04229-0

Crippa, M., Guizzardi, D., Butler, T., Keating, T., Wu, R., Kaminski, J., Kuenen, J., Kurokawa, J., Chatani, S., Morikawa, T., Pouliot, G., Racine, J., Moran, M. D., Klimont, Z., Manseau, P. M., Mashayekhi, R., Henderson, B. H., Smith, S. J., Suchyta, H., ... Foley, K. (2023). The HTAP_v3 emission mosaic: merging regional and global monthly emissions (2000–2018) to support air quality modelling and policies. *Earth System Science Data*, *15*(6), 2667–2694. https://doi.org/10.5194/essd-15-2667-2023

Dammers, E., McLinden, C. A., Griffin, D., Shephard, M. W., Van Der Graaf, S., Lutsch, E., Schaap, M., Gainairu-Matz, Y., Fioletov, V., Van Damme, M., Whitburn, S., Clarisse, L., Cady-Pereira, K., Clerbaux, C., Coheur, P. F., & Erisman, J. W. (2019). NH_3*emissionsfromlargepointsourcesderived* *Atmospheric Chemistry and Physics*, *19*(19), 12261–12293. https://doi.org/10.5194/acp-19-12261-2019

Dammers, E., Tokaya, J., Mielke, C., Hausmann, K., Griffin, D., McLinden, C., Eskes, H., & Timmermans, R. (2024). Can tropomi no_2 satellite data be used to track the drop in and resurgence of no_x emissions in germany between 2019–2021 using the multi-source plume method (mspm)? *Geoscientific Model Development*, *17*(12), 4983–5007. https://doi.org/10.5194/gmd-17-4983-2024

Dammers, E., Tokaya, J., Timmermans, R., Schaap, M., Coenen, P., Mielke, C., & Hausmann, K. (2023). Satellite-based emission verification, pilot study. https://www.umweltbundesamt.de/publikationen/satellite-based-emission-verification

Dammers, E., Vigouroux, C., Palm, M., Mahieu, E., Warneke, T., Smale, D., Langerock, B., Franco, B., Van Damme, M., Schaap, M., Notholt, J., & Erisman, J. W. (2015). Retrieval of ammonia from ground-based ftir solar spectra. *Atmospheric Chemistry and Physics*, *15*(22), 12789–12803. https://doi.org/10.5194/acp-15-12789-2015

Dammers, E., Palm, M., Van Damme, M., Vigouroux, C., Smale, D., Conway, S., Toon, G. C., Jones, N., Nussbaumer, E., Warneke, T., Petri, C., Clarisse, L., Clerbaux, C., Hermans, C., Lutsch, E., Strong, K., Hannigan, J. W., Nakajima, H., Morino, I., ... Erisman, J. W. (2016). An evaluation of iasi-nh_3 with ground-based fourier transform infrared spectroscopy measurements. *Atmospheric Chemistry and Physics*, *16*(16), 10351–10368. https://doi.org/10.5194/acp-16-10351-2016

Dammers, E., Shephard, M., Griffin, D., Chow, E., White, E., Hickman, J., Tokaya, J., Lutsch, E., Kharol, S., van der Graaf, S., et al. (2022). County-level ammonia emissions monitored worldwide. *Nature, preprint, to be resubmitted*.

Dammers, E., Shephard, M. W., Palm, M., Cady-Pereira, K., Capps, S., Lutsch, E., Strong, K., Hannigan, J. W., Ortega, I., Toon, G. C., Stremme, W., Grutter, M., Jones, N., Smale, D., Siemons, J., Hrpcek, K., Tremblay, D., Schaap, M., Notholt, J., & Erisman, J. W. (2017). Validation of the cris fast physical nh_3 retrieval with ground-based ftir. *Atmospheric Measurement Techniques*, *10*(7), 2645–2667. https://doi.org/10.5194/amt-10-2645-2017

de Foy, B., Schauer, J. J., Lorente, A., al -, Tian, Y., Sun, Y., Borsdorff, T., Wu, N., Geng, G., & Yan, L. (2022). An improved understanding of NOx emissions in South Asian megacities using TROPOMI NO2 retrievals. *Environmental Research Letters*, *17*, 024006. https://doi.org/10.1088/1748-9326/AC48B4

Dimitropoulou, E., Hendrick, F., Pinardi, G., Friedrich, M. M., Merlaud, A., Tack, F., De Longueville, H., Fayt, C., Hermans, C., Laffineur, Q., Fierens, F., & Van Roozendael, M. (2020). Validation of TROPOMI tropospheric NO_2 columns using dual-scan multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements in Uccle, Brussels. *Atmospheric Measurement Techniques*, *13*(10), 5165–5191. https://doi.org/10.5194/amt-13-5165-2020

Ding, J., van der A, R., Eskes, H., Dammers, E., Shephard, M., Wichink Kruit, R., Guevara, M., & Tarrason, L. (2024). Ammonia emission estimates using CrIS satellite observations over Europe. *Atmospheric Chemistry and Physics*, *24*(18), 10583–10599. https://doi.org/10.5194/acp-24-10583-2024

Ding, J., van der A, R. J., Eskes, H. J., Mijling, B., Stavrakou, T., van Geffen, J. H. G. M., & Veefkind, J. P. (2020). NOx Emissions Reduction and Rebound in China Due to the COVID-19 Crisis [e2020GL089912 2020GL089912]. *Geophysical Research Letters*, *47*(19), e2020GL089912. https://doi.org/https://doi.org/10.1029/2020GL089912

Ding, J., van der A, R. J., Mijling, B., & Levelt, P. F. (2017). Space-based NO_x emission estimates over remote regions improved in DECSO. *Atmospheric Measurement Techniques*, *10*(3), 925–938. https://doi.org/10.5194/amt-10-925-2017

Dix, B., Francoeur, C., Li, M., Serrano-Calvo, R., Levelt, P., Veefkind, J., McDonald, B., & De Gouw, J. (2022). Quantifying NOx Emissions from U.S. Oil and Gas Production Regions Using TROPOMI NO2. *ACS Earth and Space Chemistry*, *6*(2), 403–414. https://doi.org/10.1021/acsearthspacechem.1c00387

Douros, J., Eskes, H., van Geffen, J., Boersma, K. F., Compernolle, S., Pinardi, G., Blechschmidt, A.-M., Peuch, V.-H., Colette, A., & Veefkind, P. (2023). Comparing Sentinel-5P TROPOMI NO_2 column observations with the CAMS regional air quality ensemble. *Geoscientific Model Development*, *16*(2), 509–534. https://doi.org/10.5194/gmd-16-509-2023

EEA. (2012). *EPRTR: European Pollutant Transfer Register, database version v4.2*. http://prtr.ec.europa.eu/

Elzakker, B., & Buijsman, E. (1999). *Meetactiviteiten in 1999 in het Landelijk Meetnet Luchtkwaliteit, Report 723101032*. RIVM. https://www.rivm.nl/bibliotheek/rapporten/729999002.pdf

ESA. (2023). *Report for assessment: Earth explorer 11 candidate mission nitrosat* (ESA-EOPSM-NITR-RP-4373). European SpaceAgency, Noordwijk, The Netherlands. https://www.esa.int/Applications/Observing_the_Earth/FutureEO/Preparing_for_tomorrow/Scientific_and_technical_mission_documents

Eskes, H., Eichmann, K., Lambert, J. C., Loyola, D., Stein-Zweers, D., Dehn, A., & Zehner, C. (2025). ATM-MPC Mission Performance Cluster Nitrogen Dioxide Readme, Tech. Rep. S5P-MPC-KNMI-PRF-NO2, issue 2.8, processor version 2.8.0, date 19 March 2025. https://sentiwiki.copernicus.eu/web/s5p-products#S5P-Products-L2

Eskes, H., van Geffen, J., Boersma, K., Eichmann, K.-U., Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J. P., & Loyola, D. (2024). *Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Nitrogendioxide, Tech. Rep. S5P-KNMI-L2-0021-MA, Koninklijk Nederlands Meteorologisch Instituut (KNMI), issue 4.3.0, processor version 2.7.1, CI-7570-PUM, date 4 April 2024*. KNMI, ESA. https://sentiwiki.copernicus.eu/web/s5p-products#S5P-Products-L2

Fioletov, V., McLinden, C. A., Griffin, D., Krotkov, N., Liu, F., & Eskes, H. (2022). Quantifying urban, industrial, and background changes in no_2 during the covid-19 lockdown period based on tropomi satellite observations. *Atmospheric Chemistry and Physics*, *22*(6), 4201–4236. https://doi.org/10.5194/acp-22-4201-2022

Fioletov, V., McLinden, C. A., Kharol, S. K., Krotkov, N. A., Li, C., Joiner, J., Moran, M. D., Vet, R., Visschedijk, A. J. H., & Denier Van Der Gon, H. A. C. (2017). Multi-source so_2 emission

retrievals and consistency of satellite and surface measurements with reported emissions. *Atmospheric Chemistry and Physics*, *17*(20), 12597–12616. https://doi.org/10.5194/acp-17-12597-2017

Fisher, B. L., Lamsal, L. N., Fasnacht, Z., Oman, L. D., Joiner, J., Krotkov, N. A., Choi, S., Qin, W., & Yang, E. S. (2024). Revised estimates of no_2 reductions during the covid-19 lockdowns using updated tropomi no2 retrievals and model simulations. *Atmospheric Environment*, *326*. https://doi.org/10.1016/j.atmosenv.2024.120459

Friedrich, M. M., Rivera, C., Stremme, W., Ojeda, Z., Arellano, J., Bezanilla, A., Garcia-Reynoso, J. A., & Grutter, M. (2019). NO_2 vertical profiles and column densities from MAX-DOAS measurements in Mexico City. *Atmospheric Measurement Techniques*, *12*(4), 2545–2565. https://doi.org/10.5194/amt-12-2545-2019

Ge, X., Schaap, M., Dammers, E., Shephard, M., & de Vries, W. (2023). Impact of interannual weather variation on ammonia emissions and concentrations in germany. *Agricultural and Forest Meteorology*, *334*. https://doi.org/10.1016/j.agrformet.2023.109432

Ge, X., Schaap, M., Kranenburg, R., Segers, A., Reinds, G. J., Kros, H., & Vries, W. D. (2020). Modeling atmospheric ammonia using agricultural emissions with improved spatial variability and temporal dynamics. *Atmospheric Chemistry and Physics*, *20*, 16055–16087. https://doi.org/10.5194/acp-20-16055-2020

Goldberg, D. L., Saide, P. E., Lamsal, L. N., de Foy, B., Lu, Z., Woo, J.-H., Kim, Y., Kim, J., Gao, M., Carmichael, G., & Streets, D. G. (2019). A top-down assessment using OMI NO_2 suggests an underestimate in the NO_x emissions inventory in Seoul, South Korea, during KORUS-AQ. *Atmospheric Chemistry and Physics*, *19*(3), 1801–1818. https://doi.org/10.5194/acp-19-1801-2019

Griffin, D., McLinden, C. A., Dammers, E., Adams, C., Stockwell, C. E., Warneke, C., Bourgeois, I., Peischl, J., Ryerson, T. B., Zarzana, K. J., Rowe, J. P., Volkamer, R., Knote, C., Kille, N., Koenig, T. K., Lee, C. F., Rollins, D., Rickly, P. S., Chen, J., … Makar, P. (2021). Biomass burning nitrogen dioxide emissions derived from space with tropomi: Methodology and validation. *Atmospheric Measurement Techniques*, *14*(12), 7929–7957. https://doi.org/10.5194/amt-14-7929-2021

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., … 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. https://doi.org/https://doi.org/10.1029/2018GL081095

Guevara, M., Jorba, O., Tena, C., Denier van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C., & Pérez García-Pando, C. (2021). Copernicus Atmosphere Monitoring Service TEMPOral profiles (CAMS-TEMPO): global and European emission temporal profile maps for atmospheric chemistry modelling. *Earth System Science Data*, *13*(2), 367–404. https://doi.org/10.5194/essd-13-367-2021

Guevara, M., Petetin, H., Jorba, O., Denier van der Gon, H., Kuenen, J., Super, I., Granier, C., Doumbia, T., Ciais, P., Liu, Z., Lamboll, R. D., Schindlbacher, S., Matthews, B., & Pérez García-Pando, C. (2023). Towards near-real-time air pollutant and greenhouse gas emissions: lessons learned from multiple estimates during the COVID-19 pandemic. *Atmospheric Chemistry and Physics*, *23*(14), 8081–8101. https://doi.org/10.5194/acp-23-8081-2023

Herman, J., Abuhassan, N., Kim, J., Kim, J., Dubey, M., Raponi, M., & Tzortziou, M. (2019). Underestimation of column NO_2 amounts from the OMI satellite compared to diurnally varying ground-based retrievals from multiple PANDORA spectrometer instruments. *Atmospheric Measurement Techniques*, *12*(10), 5593–5612. https://doi.org/10.5194/amt-12-5593-2019

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M., & Abuhassan, N. (2009). NO_2 column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation. *Journal of Geo-*

physical Research: Atmospheres, *114*(D13). https://doi.org/https://doi.org/10.1029/2009JD011848

Herrera, B., Bezanilla, A., Blumenstock, T., Dammers, E., Hase, F., Clarisse, L., Magaldi, A., Rivera, C., Stremme, W., Strong, K., Viatte, C., Van Damme, M., & Grutter, M. (2022). Measurement report: Evolution and distribution of nh_3 over mexico city from ground-based and satellite infrared spectroscopic measurements. *Atmospheric Chemistry and Physics*, *22*(21), 14119–14132. https://doi.org/10.5194/acp-22-14119-2022

Honninger, G., von Friedeburg, C., & Platt, U. (2004). Multi axis differential optical absorption spectroscopy (MAX-DOAS). *Atmospheric Chemistry and Physics*, *4*(1), 231–254. https://doi.org/10.5194/acp-4-231-2004

Hoogerbrugge, R., Geilenkirchen, G., den Hollander, H., Siteur, K., Smeets, W., van der Swaluw, E., de Vries, W., & Wichink Kruit, R. (2021). *Large-scale concentration and deposition maps of The Netherlands*. RIVM report. https://doi.org/https://doi.org/10.21945/RIVM-2021-0068

Hoogerbrugge, R., Hazelhorst, S., Huitema, M., Siteur, K., Smeets, W., Soenario, I., Visser, S., de Vries, W., & Kruit, R. W. (2023). Grootschalige concentratiekaarten nederland. rapportage 2023. *RIVM rapport 2023-0113*.

Hoogerbrugge, R., & Liem, A. (2000). How to handle non-detects. *Organohalogen Compounds*, *45*, 13–16.

Ialongo, I., Virta, H., Eskes, H., Hovila, J., & Douros, J. (2020). Comparison of TROPOMI/Sentinel-5 Precursor NO_2 observations with ground-based measurements in Helsinki. *Atmospheric Measurement Techniques*, *13*(1), 205–218. https://doi.org/10.5194/amt-13-205-2020

Judd, L. M., Al-Saadi, J. A., Szykman, J. J., Valin, L. C., Janz, S. J., Kowalewski, M. G., Eskes, H. J., Veefkind, J. P., Cede, A., Mueller, M., Gebetsberger, M., Swap, R., Pierce, R. B., Nowlan, C. R., Abad, G. G., Nehrir, A., & Williams, D. (2020). Evaluating Sentinel-5P TROPOMI tropospheric NO_2 column densities with airborne and Pandora spectrometers near New York City and Long Island Sound. *Atmospheric Measurement Techniques*, *13*(11), 6113–6140. https://doi.org/10.5194/amt-13-6113-2020

Kharol, S. K., Shephard, M. W., McLinden, C. A., Zhang, L., Sioris, C. E., O'Brien, J. M., Vet, R., Cady-Pereira, K. E., Hare, E., Siemons, J., & Krotkov, N. A. (2018). Dry deposition of reactive nitrogen from satellite observations of ammonia and nitrogen dioxide over north america. *Geophysical Research Letters*, *45*(2), 1157–1166. https://doi.org/10.1002/2017GL075832

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., & Denier van der Gon, H. (2022). CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling. *Earth System Science Data*, *14*(2), 491–515. https://doi.org/10.5194/essd-14-491-2022

Kutzner, R. D., Cuesta, J., Chelin, P., Petit, J.-E., Ray, M., Landsheere, X., Tournadre, B., Dupont, J.-C., Rosso, A., Hase, F., Orphal, J., & Beekmann, M. (2021). Diurnal evolution of total column and surface atmospheric ammonia in the megacity of paris, france, during an intense springtime pollution episode. *Atmospheric Chemistry and Physics*, *21*(15), 12091–12111. https://doi.org/10.5194/acp-21-12091-2021

Lambert, J.-C., & et al. (2024). *Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products 24: April 2018 – August 2024, S5P MPC Routine Operations Consolidated Validation Report series, Issue 24, Version 24.00.00, 212 pp., 16 September 2024*. ATM-MPC, ESA.

Laughner, J. L., Zare, A., & Cohen, R. C. (2016). Effects of daily meteorology on the interpretation of space-based remote sensing of NO_2. *Atmospheric Chemistry and Physics*, *16*(23), 15247–15264. https://doi.org/10.5194/acp-16-15247-2016

Lin, X., A, R. J. v. d., Laat, J. d., Huijnen, V., Mijling, B., Ding, J., Eskes, H., Douros, J., Liu, M., Zhang, X., & Liu, Z. (2023). European soil NOx emissions derived from satellite NO2 observations. https://doi.org/10.22541/essoar.170224578.81570487/v1

Lolkema, D. E., Noordijk, H., Stolk, A. P., Hoogerbrugge, R., Zanten, M. C. V., & Pul, W. A. V. (2015). The measuring ammonia in nature (man) network in the netherlands. *Biogeosciences*, *12*, 5133–5142. https://doi.org/10.5194/bg-12-5133-2015

Manders, A. M. M., Builtjes, P. J. H., Curier, L., Denier Van Der Gon, H. A. C., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J. J. P., Segers, A. J., Timmermans, R. M. A., Visschedijk, A. J. H., Wichink Kruit, R. J., Van Pul, W. A. J., Sauter, F. J., Van Der Swaluw, E., Swart, D. P. J., Douros, J., Eskes, H., Van Meijgaard, E., ... Schaap, M. (2017). Curriculum vitae of the lotos–euros (v2.0) chemistry transport model. *Geoscientific Model Development*, *10*(11), 4145–4173. https://doi.org/10.5194/gmd-10-4145-2017

McLinden, C., Griffin, D., Fioletov, V., Zhang, J., Dammers, E., Adams, C., Loria, M., Krotkov, N., & Lamsal, L. (2024). Monitoring of total and off-road no_x emissions from canadian oil sands surface mining using the ozone monitoring instrument. *EGUsphere*, *2024*, 1–44. https://doi.org/10.5194/egusphere-2024-2856

Menut, L., Bessagnet, B., Briant, R., Cholakian, A., Couvidat, F., Mailler, S., Pennel, R., Siour, G., Tuccella, P., Turquety, S., & Valari, M. (2021). The CHIMERE v2020r1 online chemistry-transport model. *Geoscientific Model Development*, *14*(11), 6781–6811. https://doi.org/10.5194/gmd-14-6781-2021

Mijling, B., & van der A, R. J. (2012). Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale. *Journal of Geophysical Research: Atmospheres*, *117*(D17). https://doi.org/https://doi.org/10.1029/2012JD017817

Misra, P., Takigawa, M., Khatri, P., Dhaka, S. K., Dimri, A. P., Yamaji, K., Kajino, M., Takeuchi, W., Imasu, R., Nitta, K., Patra, P. K., & Hayashida, S. (2021). Nitrogen oxides concentration and emission change detection during COVID-19 restrictions in North India. *Scientific Reports 2021 11:1*, *11*, 1–11. https://doi.org/10.1038/s41598-021-87673-2

Moncet, J.-L., Uymin, G., Lipton, A. E., & Snell, H. E. (2008). Infrared radiance modeling by optimal spectral sampling. *Journal of the Atmospheric Sciences*, *65*(12), 3917–3934. https://doi.org/10.1175/2008JAS2711.1

Noordijk, H., Braam, M., Rutledge-Jonker, S., Hoogerbrugge, R., Stolk, A., & van Pul, W. (2020). Performance of the MAN ammonia monitoring network in the Netherlands. *Atmospheric Environment*, *228*, 117400.

Petersen, A. K., Brasseur, G. P., Bouarar, I., Flemming, J., Gauss, M., Jiang, F., Kouznetsov, R., Kranenburg, R., Mijling, B., Peuch, V.-H., Pommier, M., Segers, A., Sofiev, M., Timmermans, R., Van Der A, R., Walters, S., Xie, Y., Xu, J., & Zhou, G. (2019). Ensemble forecasts of air quality in eastern china – part 2: Evaluation of the marcopolo–panda prediction system, version 1. *Geoscientific Model Development*, *12*(3), 1241–1266. https://doi.org/10.5194/gmd-12-1241-2019

Peuch, V.-H., Engelen, R., Rixen, M., Dee, D., Flemming, J., Suttie, M., Ades, M., Agustí-Panareda, A., Ananasso, C., Andersson, E., Armstrong, D., Barré, J., Bousserez, N., Dominguez, J. J., Garrigues, S., Inness, A., Jones, L., Kipling, Z., Letertre-Danczak, J., ... Thépaut, J.-N. (2022). The copernicus atmosphere monitoring service: From research to operations. *Bulletin of the American Meteorological Society*, *103*(12), E2650–E2668. https://doi.org/10.1175/BAMS-D-21-0314.1

Rodgers, C. D. (2000, July). *Inverse methods for atmospheric sounding: Theory and practice* (Vol. 2). WORLD SCIENTIFIC. https://doi.org/10.1142/3171

Rothman, L., Gordon, I., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P., Birk, M., Bizzocchi, L., Boudon, V., Brown, L., Campargue, A., Chance, K., Cohen, E., Coudert, L., Devi, V., Drouin, B., Fayt, A., Flaud, J.-M., Gamache, R., ... Wagner, G. (2013). The hitran2012 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *130*, 4–50. https://doi.org/10.1016/j.jqsrt.2013.07.002

Rutledge-Jonker, S., Braam, M., Hoogerbrugge, R., Wichink Kruit, R., Nemitz, E., & Twigg, M. (2023). *Ammonia deposition measured with Conditional Time-Averaged Gradient (CO-TAG) systems in the Netherlands. Methodological advances and results for 2012 – 2020*. https://doi.org/10.21945/RIVM-2022-0202

Sauter, F., van Zanten, M., van der Swaluw, E., Aben, J., de Leeuw, F., & van Jaarsveld, H. (2018). The OPS-model, Description of OPS 4.5.2. Retrieved December 5, 2024, from https://www.rivm.nl/media/ops/v4.5.2/OPS-model-v4.5.2.pdf

Shephard, M. W., & Cady-Pereira, K. E. (2015). Cross-track infrared sounder (cris) satellite observations of tropospheric ammonia. *8*, 1323–1336. https://doi.org/10.5194/amt-8-1323-2015

Shephard, M. W., Dammers, E., Cady-Pereira, K. E., Kharol, S. K., Thompson, J., Gainariu-Matz, Y., Zhang, J., McLinden, C. A., Kovachik, A., Moran, M., Bittman, S., Sioris, C. E., Griffin, D., Alvarado, M. J., Lonsdale, C., Savic-Jovcic, V., & Zheng, Q. (2020). Ammonia measurements from space with the cross-track infrared sounder: Characteristics and applications. *Atmospheric Chemistry and Physics*, *20*(4), 2277–2302. https://doi.org/10.5194/acp-20-2277-2020

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., & Wind, P. (2012). The EMEP MSC-W chemical transport model - technical description. *Atmospheric Chemistry and Physics*, *12*(16), 7825–7865. https://doi.org/10.5194/acp-12-7825-2012

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., Powers, J. G., et al. (2008). A description of the advanced research WRF version 3. *NCAR technical note*, *475*(125), 10–5065.

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., & Richter, A. (2021). Evaluation of the LOTOS-EUROS NO_2 simulations using ground-based measurements and S5P/TROPOMI observations over Greece. *Atmospheric Chemistry and Physics*, *21*(7), 5269–5288. https://doi.org/10.5194/acp-21-5269-2021

Smeets, W., Geilenkirchen, G., Hammingh, P., Hilbers, H., Nijdam, D., Plomp, A., van Schijndel, M., Traa, M., Vethman, P., Stammes, I., et al. (2023). *Geraamde ontwikkelingen in nationale emissies van luchtverontreinigende stoffen 2023: Rapportage bij de klimaat-en energieverkenning 2022* (tech. rep.). PBL Planbureau voor de Leefomgeving.

Timmermans, R., Kranenburg, R., Manders, A., Hendriks, C., Segers, A., Dammers, E., Zhang, Q., Wang, L., Liu, Z., Zeng, L., Denier Van Der Gon, H., & Schaap, M. (2017). Source apportionment of pm2.5 across china using lotos-euros. *Atmospheric Environment*, *164*, 370–386. https://doi.org/10.1016/j.atmosenv.2017.06.003

Tournadre, B., Chelin, P., Ray, M., Cuesta, J., Kutzner, R. D., Landsheere, X., Fortems-Cheiney, A., Flaud, J.-M., Hase, F., Blumenstock, T., Orphal, J., Viatte, C., & Camy-Peyret, C. (2020). Atmospheric ammonia (nh_3) over the paris megacity: 9 years of total column observations from ground-based infrared remote sensing. *Atmospheric Measurement Techniques*, *13*(7), 3923–3937. https://doi.org/10.5194/amt-13-3923-2020

Tzortziou, M., Loughner, C. P., Goldberg, D. L., Judd, L., Nauth, D., Kwong, C. F., Lin, T., Cede, A., & Abuhassan, N. (2023). Intimately tracking NO2 pollution over the New York City - Long Island Sound land-water continuum: An integration of shipboard, airborne, satellite observations, and models. *Science of The Total Environment*, *897*, 165144. https://doi.org/https://doi.org/10.1016/j.scitotenv.2023.165144

van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., ter Linden, M., Ludewig, A., Boersma, K. F., & Veefkind, J. P. (2022). Sentinel-5P TROPOMI NO_2 retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data. *Atmospheric Measurement Techniques*, *15*(7), 2037–2060. https://doi.org/10.5194/amt-15-2037-2022

van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., Veefkind, J. P., Sneep, M., & ter Linden, M. (2024). *TROPOMI ATBD of the total and tropospheric NO_2 data products, Tech. Rep. S5P-KNMI-L2-0005-RP, Koninklijk Nederlands Meteorologisch Instituut (KNMI), CI-7430-ATBD, issue 2.7.0, processor version 2.7.1, date 18 July 2024*. KNMI, ESA. https://sentiwiki.copernicus.eu/web/s5p-products#S5P-Products-L2

Van Damme, M., Clarisse, L., Heald, C. L., Hurtmans, D., Ngadi, Y., Clerbaux, C., Dolman, A. J., Erisman, J. W., & Coheur, P. F. (2014). Global distributions, time series and error char-

acterization of atmospheric ammonia (nh_3) from iasi satellite observations. *Atmospheric Chemistry and Physics*, *14*(6), 2905–2922. https://doi.org/10.5194/acp-14-2905-2014

Van Damme, M., Clarisse, L., Whitburn, S., Hadji-Lazaro, J., Hurtmans, D., Clerbaux, C., & Coheur, P.-F. (2018). Industrial and agricultural ammonia point sources exposed. *Nature*, *564*(7734), 99–103.

Van Damme, M., Whitburn, S., Clarisse, L., Clerbaux, C., Hurtmans, D., & Coheur, P.-F. (2017). Version 2 of the iasi nh_3 neural network retrieval algorithm: Near-real-time and reanalysed datasets. *Atmospheric Measurement Techniques*, *10*(12), 4905–4914. https://doi.org/10.5194/amt-10-4905-2017

Van Der Graaf, S., Dammers, E., Segers, A., Kranenburg, R., Schaap, M., Shephard, M. W., & Erisman, J. W. (2022). Data assimilation of CrIS NH_3 satellite observations for improving spatiotemporal NH_3 distributions in LOTOS-EUROS. *Atmospheric Chemistry and Physics*, *22*(2), 951–972. https://doi.org/10.5194/acp-22-951-2022

van der Swaluw, E., de Vries, W., Sauter, F., Wichink Kruit, R., Vieno, M., Fagerli, H., & van Pul, A. (2021). Trend analysis of reduced nitrogen components over the Netherlands with the EMEP4NL and OPS model. *Atmospheric Environment*, *248*, 118183. https://doi.org/https://doi.org/10.1016/j.atmosenv.2021.118183

van der A, R. J., Ding, J., & Eskes, H. (2024). Monitoring European anthropogenic NO_x emissions from space. *Atmospheric Chemistry and Physics*, *24*(13), 7523–7534. https://doi.org/10.5194/acp-24-7523-2024

Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., … Levelt, P. (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 [The Sentinel Missions - New Opportunities for Science]. *Remote Sensing of Environment*, *120*, 70–83. https://doi.org/https://doi.org/10.1016/j.rse.2011.09.027

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., Boersma, K. F., Bognar, K., Dehn, A., Donner, S., … Zehner, C. (2021). Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO_2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks. *Atmospheric Measurement Techniques*, *14*(1), 481–510. https://doi.org/10.5194/amt-14-481-2021

Verweij, R. W., Pisoni, E., van Pul, A., Thunis, P., & van der Swaluw, E. (2023). A fast method for ammonia emission reduction scenario studies: Combining EMEP4NL with the SHERPA tool. *Atmospheric Environment*, *306*, 119805.

Vivanco, M., Bessagnet, B., Cuvelier, C., Theobald, M., Tsyro, S., Pirovano, G., Aulinger, A., Bieser, J., Calori, G., Ciarelli, G., Manders, A., Mircea, M., Aksoyoglu, S., Briganti, G., Cappelletti, A., Colette, A., Couvidat, F., D'Isidoro, M., Kranenburg, R., … Ung, A. (2017). Joint analysis of deposition fluxes and atmospheric concentrations of inorganic nitrogen and sulphur compounds predicted by six chemistry transport models in the frame of the eurodeltaiii project. *Atmospheric Environment*, *151*, 152–175. https://doi.org/10.1016/j.atmosenv.2016.11.042

Wever, D., Coenen, P., Dröge, R., Geilenkirchen, G., van Huijstee, J., 't Hoen, M., Honig, E., te Molder, R., Smeets, W., van Zanten, M., & van der Zee, T. (2021). Publication informative inventory report 2021. emissions of transboundary air pollutants in the netherlands 1990–2019. https://doi.org/10.21945/RIVM-2021-0005

White, E., Shephard, M. W., Cady-Pereira, K. E., Kharol, S. K., Ford, S., Dammers, E., Chow, E., Thiessen, N., Tobin, D., Quinn, G., O'Brien, J., & Bash, J. (2023). Accounting for non-detects: Application to satellite ammonia observations. *Remote Sensing*, *15*(10), 2610. https://doi.org/10.3390/rs15102610

Williams, J. E., Boersma, K. F., Le Sager, P., & 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. https://doi.org/10.5194/gmd-10-721-2017

Yarce Botero, A., Lopez-Restrepo, S., Pinel Peláez, N., Quintero, O. L., Segers, A., & Heemink, A. W. (2021). Estimating nox lotos-euros ctm emission parameters over the northwest of south america through 4denvar tropomi no2 assimilation. *Atmosphere*, *12*(12), 1633. https://doi.org/10.3390/atmos12121633

Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Cede, A., Tiefengraber, M., Müller, M., Bognar, K., Strong, K., Boersma, F., Eskes, H., Davies, J., Ogyu, A., & Lee, S. C. (2020). Assessment of the quality of TROPOMI high-spatial-resolution NO_2 *dataproductsintheGreaterTorontoArea*. *Atmospheric Measurement Techniques*, *13*(4), 2131–2159. https://doi.org/10.5194/amt-13-2131-2020

Abbreviations

AMF	Air Mass Factor
CAMS	Copernicus Atmosphere Monitoring Service
CAMS-NCP	National collaboration programme of CAMS
CAMS-GLOB-ANT	CAMS global anthropogenic emissions
CAMS-GLOB-TEMPO	CAMS global anthropogenic emissions including temporal profiles
CAMS-REG-ANT	European anthropogenic emissions inventory developed in CAMS
CFPR	CrIS-Fast Physical Retrieval
CHIMERE	The French multi-scale chemistry-transport model for atmospheric composition analysis and forecast
CrIS	The Cross-track Infrared Sounder (CrIS) instrument
CrIS-SNPP	CrIS flying on the Suomi-NPP satellite
CrIS-NOAA-20	CrIS flying on the NOAA-20 satellite
CTM	Chemistry Transport Model
DECSO	Daily Emissions Constrained by Satellite Observations
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Centre for Medium-Range Weather Forecasts
EDGAR	Emissions Database for Global Atmospheric Research
EEA	European Environmental Agency
EMEP	European Monitoring and Evaluation Programme
EMEP4NL	The EMEP model implemented for the Netherlands
E-PRTR	European Pollutant Release and Transfer Register
ESA	European Space Agency
FDA	Flux Divergence Approach
FTIR	Fourier Transform Infrared Spectroscopy
FTS	Fourier Transform Spectrometer
HRI	Hyperspectral Range Index
HTAP	Hemispheric Transport of Air Pollution (International task force)
IASI	Infrared Atmospheric Sounding Interferometer
KNMI	Royal Netherlands Meteorological Institute
LRTAP	Convention on Long-range Transboundary Air Pollution
MAX-DOAS	Multi-AXis Differential Optical Absorption Spectroscopy
NASA	National Aeronautics and Space Administration
NDACC	Network for the Detec- tion of Atmospheric Composition Change
NEC	National Emission reductions Commitments
NKS-SAGEN	Nationaal Kennisprogramma Stikstof - Satellietdata en Ensemble modellering
NOx	sum of concentrations of NO and NO_2
OMI	Ozone Monitoring Instrument
PBL	Planetary Boundary Layer
PM10	Particulate Matter with a diameter up to 10 micrometers (μm)
PM2.5	Particulate Matter with a diameter up to 2.5 micrometers (μm)
S5P	The Copernicus Sentinel-5P satellite with the TROPOMI instrument
S5P-MPC	S5P Mission Performance Centre
TM5-MP	Tracer Model version 5, massive parallel version
TROPOMI	TROPOspheric Monitoring Instrument
VCD	Vertical Column Density
VDAF	Validation Data Analysis Facility for Sentinel-5P

Royal Netherlands Meteorological Institute

PO Box 201 | NL-3730 AE De Bilt
Netherlands | www.knmi.nl