CAMS Service Evolution

D3.3 Report on OSSE on IRS NH3 , O3/CO and Sentinel-4 NO2 capabilities and assimilation

Due date of deliverable	31/12/2025
Submission date	17/12/2025
File Name	CAMEO-D3-3-V1.1
Work Package /Task	Task 3.3
Organisation Responsible of Deliverable	TNO
Author name(s)	Tyler Wizenberg, Enrico Dammers, Beatriz Hererra (TNO), Arjo Segers, Gunnar Lange, Peter Wind (Met Norway), Joanna Struzewska, Jacek W. Kaminski Lech Lobocki, Tomasz Przybyla, Damian Mochocki, Wojciech Baginski (IOS-PIB), Gaël Descombes, Alexandre Simeon, Augustin Colette, Anthony Ung (INERIS), Vincent Guidard, Olivier Coopmann (MF), Guillaume Monteil, Jeronimo Escribano, Oriol Jorba (BSC)
Revision number	1
Status	ISSUED
Dissemination Level	PUBLIC

Funded by the European Union

The CAMEO project (grant agreement No 101082125) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

This report describes the activities conducted in the Task 3.3 entitled "*OSSEs on IRS and Sentinel-4*" performed during the CAMEO project.

Section 3, which covers Subtask 3.3.1, presents the assimilation of synthetic ammonia (NH_3) observations from the upcoming Meteosat Third Generation – InfraRed Sounder (MTG-IRS) satellite instrument into the LOTOS-EUROS and EMEP regional air quality models. Using advanced data assimilation techniques, an Observing System Simulation Experiment (OSSE) was performed that demonstrates that integrating high-resolution geostationary NH_3 measurements can significantly improve the accuracy of simulated emissions, concentrations, and deposition fields across Europe. The work also identifies areas for further improvement in the data assimilation systems, such as optimizing vertical profile representation and temporal variability. The results show substantial reductions in model biases and highlight the potential of MTG-IRS data to enhance reactive nitrogen fields in air quality analyses and forecasts.

Section 4, which covers Subtask 3.3.2, presents an OSSE-based evaluation of the benefits of assimilating synthetic MTG-IRS radiances for ozone (O_3) and carbon monoxide (CO) in the regional MOCAGE model. A robust OSSE framework was used, with a Nature Run representing the true atmospheric state to generate and verify synthetic IRS radiances. These were assimilated using a 3D-Var system over a summer period, with realistic coverage, cloud screening, and channel selection based on sensitivity to O_3 and CO. Assimilation of the synthetic radiances led to significant reductions in forecast errors for both species, up to 50% in some cases, particularly in the troposphere and lower stratosphere where IRS channels are most sensitive, with the greatest improvements in regions of higher observation density and clearer skies. These results confirm that high-frequency, high-resolution IRS radiances can provide valuable constraints for regional chemical composition analysis, support more accurate air quality forecasts, and demonstrate the feasibility of handling large radiance data volumes in an operational context.

Section 5, which covers Subtask 3.3.3, presents an OSSE-based assessment of how geostationary NO_2 observations from the Copernicus Sentinel-4 mission could improve CAMS regional air quality analyses and forecasts. Synthetic Sentinel-4 NO_2 retrievals were generated from GEM-AC Nature Runs and assimilated into four CAMS regional models: CHIMERE, GEM-AQ, LOTOS-EUROS, and MONARCH, to evaluate their ability to handle high-frequency geostationary data. All models successfully incorporated the observations within operational timeframes, enhancing near-surface NO_2 fields in the analysis. While geostationary column data add value to reanalysis, the impact on forecasts is limited, typically lasting only a few hours after assimilation. Optimizing emissions, rather than directly adjusting concentrations, may be a more effective approach for incorporating Sentinel-4 observations. These findings provide a strong foundation for further development of assimilation strategies within CAMS.

In summary, the results from these three tasks demonstrate that the CAMS service is well positioned to benefit from the next generation of geostationary satellite observations. The successful assimilation of synthetic MTG-IRS and Sentinel-4 data across multiple models confirms both the technical feasibility and the scientific value of integrating high-frequency, high-resolution satellite products into regional air quality analyses and forecasts. These advances will enable CAMS to deliver more accurate, timely, and policy-relevant information to users, supporting improved monitoring and management of atmospheric composition across Europe. A summary of the overall conclusions from the work performed in this deliverable and the implications for the CAMS ecosystem is provided in Section 6.

Table of Contents

2 Introduction

2.1 Background

Monitoring the composition of the atmosphere is a key objective of the European Union's flagship Space programme Copernicus, with the Copernicus Atmosphere Monitoring Service (CAMS) providing free and continuous data and information on atmospheric composition.

The CAMS Service Evolution (CAMEO) project will enhance the quality and efficiency of the CAMS service and help CAMS to better respond to policy needs such as air pollution and greenhouse gases monitoring, the fulfilment of sustainable development goals, and sustainable and clean energy.

CAMEO will help prepare CAMS for the uptake of forthcoming satellite data, including Sentinel-4, -5 and 3MI, and advance the aerosol and trace gas data assimilation methods and inversion capacity of the global and regional CAMS production systems.

CAMEO will develop methods to provide uncertainty information about CAMS products, in particular for emissions, policy, solar radiation and deposition products in response to prominent requests from current CAMS users.

CAMEO will contribute to the medium- to long-term evolution of the CAMS production systems and products.

The transfer of developments from CAMEO into subsequent improvements of CAMS operational service elements is a main driver for the project and is the main pathway to impact for CAMEO.

The CAMEO consortium, led by ECMWF, the entity entrusted to operate CAMS, includes several CAMS partners thus allowing CAMEO developments to be carried out directly within the CAMS production systems and facilitating the transition of CAMEO results to future upgrades of the CAMS service.

This will maximise the impact and outcomes of CAMEO as it can make full use of the existing CAMS infrastructure for data sharing, data delivery and communication, thus supporting policymakers, business and citizens with enhanced atmospheric environmental information.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

The primary aim of this deliverable is to enhance the readiness of models within the CAMS regional ensemble for the integration of forthcoming geostationary instruments, such as Sentinel 4 and MTG-IRS, and to assess the potential advantages for the CAMS regional service. To achieve this, three Observing System Simulation Experiments (OSSEs) were carried out.

An OSSE was conducted for the NH_3 retrievals using MTG-IRS, which is set to deliver geostationary NH_3 observations at an unprecedented spatial and temporal resolution (approximately 4x4km^2 and half-hourly, respectively). In this experiment, a suite of synthetic observations was generated. The data assimilation systems within the EMEP and LOTOS-EUROS models were further developed to accommodate the assimilation of these synthetic

observations, enabling an evaluation of their benefits for improving concentration, emission and deposition fields.

An OSSE framework was established to simulate radiances in the ozone (O_3) and carbon monoxide (CO) bands of the MTG-IRS spectrum. These radiances were assimilated using the RTTOV radiative transfer model, ensuring realistic spatial coverage by excluding areas affected by cloud contamination. Assimilation trials were performed to assess the potential of utilising IRS data over a summer period, particularly during ozone pollution events. Both the Nature Run and Control Run made use of MOCAGE, configured in two distinct manners to ensure robust results.

Additionally, an OSSE was undertaken for the NO_2 retrievals from Sentinel 4. This work initially utilised Nature Runs conducted in previous projects, such as CAMS61. However an updated Nature Run was also produced with GEM-AC, from which NO_2 observations, replicating Sentinel 4 characteristics, were generated and assimilated into the CHIMERE, GEM-AC, LOTOS-EUROS, and MONARCH models. By involving multiple models, the project ensures the development of assimilation capacity within CAMS-regional for Sentinel 4 data as soon as it becomes available, thereby providing a more comprehensive understanding of the added value offered by this assimilation.

2.2.2 Work performed in this deliverable

In this deliverable the work as planned in the Description of Action (DoA, WP3 T3.3.1, T3.3.2 and T3.3.3) was performed. The efforts within this work package were focused on the following aspects:

- T3.3.1 The creation of synthetic MTG-IRS NH_3 observations and the assimilation of these observations in the data assimilation systems of the LOTOS-EUROS and EMEP models to evaluate the benefits for improving NH_3 concentration, emission and deposition fields (section 3).
- T.3.3.2 The development of an OSSE framework to simulate IRS radiances in the ozone and CO bands of the IRS spectrum (section 4).
- T.3.3.3 Conducting a multi-model OSSE for the NO_2 retrievals of Sentinel 4 with the CHIMERE, GEM-AC, LOTOS-EUROS, and MONARCH models (section 5).

2.2.3 Deviations and counter measures

For T3.3.1, the GEM-AC model from IOS-PIB was intended to be used as a nature run for the creation of the synthetic MTG-IRS NH_3 observations to be assimilated in the LOTOS-EUROS and EMEP models. Some significant differences in the chemistry schemes and partitioning of NH_3/$NH4^+$ in GEM-AC were discovered relative to LOTOS-EUROS that were likely to impact the OSSE. It was thus decided that the synthetic observations to be assimilated in LOTOS-EUROS would be generated from the EMEP Meteorological Synthesizing Centre – West (MSC-W) model, while those to be assimilated into the EMEP MSC-W model would be generated from LOTOS-EUROS. This change did not alter the scope or objectives of the task, but ensured internal consistency between the nature run and the assimilation models.

2.2.4 CAMEO Project Partners:

ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
Met Norway	METEOROLOGISK INSTITUTT
BSC	BARCELONA SUPERCOMPUTING CENTER-CENTRO NACIONAL DE SUPERCOMPUTACION
KNMI	KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT-KNMI
SMHI	SVERIGES METEOROLOGISKA OCH HYDROLOGISKA INSTITUT
BIRA-IASB	INSTITUT ROYAL D'AERONOMIE SPATIALEDE BELGIQUE
HYGEOS	HYGEOS SARL
FMI	ILMATIETEEN LAITOS
DLR	DEUTSCHES ZENTRUM FUR LUFT - UND RAUMFAHRT EV
ARMINES	ASSOCIATION POUR LA RECHERCHE ET LE DEVELOPPEMENT DES METHODES ET PROCESSUS INDUSTRIELS
CNRS	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS
GRASP-SAS	GENERALIZED RETRIEVAL OF ATMOSPHERE AND SURFACE PROPERTIES EN ABREGE GRASP
CU	UNIVERZITA KARLOVA
CEA	COMMISSARIAT A L'ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES
MF	METEO-FRANCE
TNO	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO
INERIS	INSTITUT NATIONAL DE L'ENVIRONNEMENT INDUSTRIEL ET DES RISQUES - INERIS
IOS-PIB	INSTYTUT OCHRONY SRODOWISKA - PANSTWOWY INSTYTUT BADAWCZY
FZJ	FORSCHUNGSZENTRUM JULICH GMBH
AU	AARHUS UNIVERSITET
ENEA	AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

3 Assimilation of MTG-IRS NH_3 Synthetic Observations in LOTOS-EUROS and EMEP

3.1 Introduction

For Task T3.3.1, Observing System Simulation Experiments (OSSEs) were conducted to evaluate the potential impact of future geostationary ammonia (NH_3) observations from the Meteosat Third Generation Infrared Sounder (MTG-IRS) within the CAMS regional air quality service. In these experiments, synthetic MTG-IRS NH_3 retrievals were assimilated into two established data assimilation frameworks: the LOTOS-EUROS Local Ensemble Transform Kalman Filter (LETKF) and the EMEP Local Fractions Variational Inversion System (ELVIS). The primary objective was to assess how these high-frequency, spatially resolved observations could improve the representation of NH_3 emissions, atmospheric concentrations, and deposition processes in the LOTOS-EUROS and EMEP MSC-W models. A diagram of the OSSE setup that was used is shown in Figure 1.

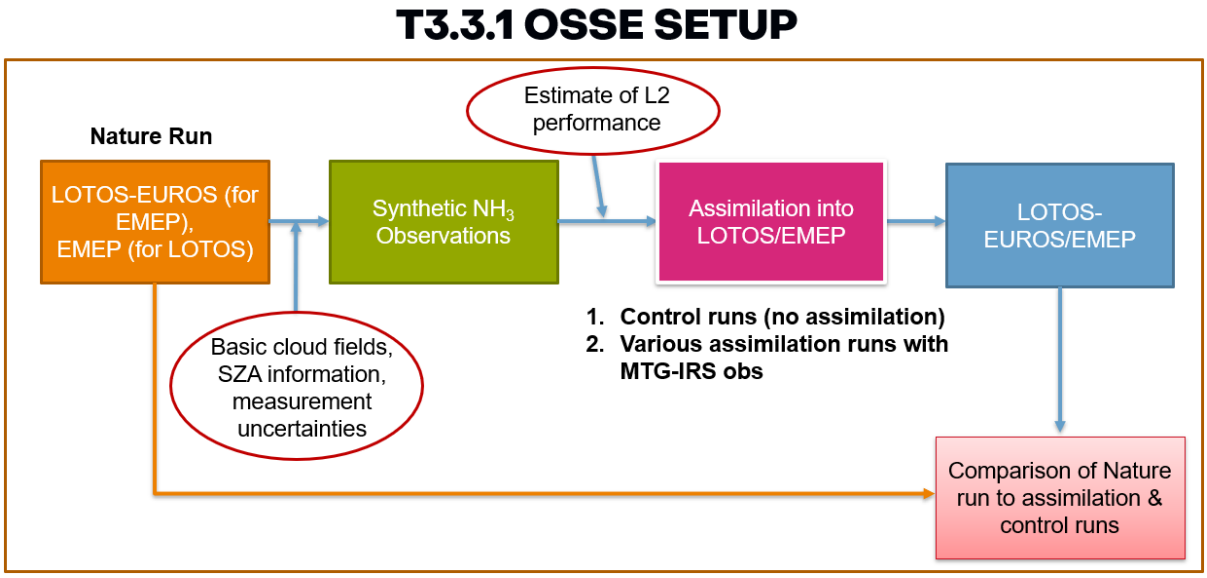

Figure 1 – Diagram of the OSSE setup that was used for the assimilation of synthetic MTG-IRS NH_3 observations in LOTOS-EUROS and EMEP MSC-W in Task 3.3.1.

The synthetic MTG-IRS dataset for assimilation was generated from an EMEP MSC-W nature run provided by Met Norway, designed to mimic the characteristics (e.g., footprint geometry, vertical sensitivity, and errors) of the upcoming MTG-IRS instrument. By assimilating these simulated observations, we quantify the added value of geostationary NH_3 measurements and identify practical challenges in integrating such large datasets into operational systems. This OSSE framework allows us to systematically compare assimilated and non-assimilated simulations using key performance metrics and synthetic ground-based observations, providing a controlled environment to evaluate improvements in model skill.

Beyond assessing observational impact, these experiments serve as a testbed for refining both the LOTOS-EUROS LETKF system and the EMEP variational inversion system to handle high-volume satellite data streams. This dual-model setup strengthens the robustness of the results and facilitates uptake in different assimilation systems within the CAMS regional ensemble. The insights gained will guide the development of operational methodologies for MTG-IRS data assimilation, ultimately supporting enhanced air quality analyses and forecasts across Europe.

For the EMEP modelling framework, the inversion system ELVIS was developed to estimate emissions from satellite observations, relying on Local Fraction outputs optionally written by EMEP simulations. These Local Fractions describe the sensitivity of simulated concentrations to emissions from nearby locations and recent hours. The system was evaluated using the same synthetic IRS observations as LOTOS-EUROS, with a different emission configuration to ensure independence from the nature run. This evaluation provided valuable input for further development toward using real MTG-IRS observations.

3.2 Methods and Datasets

3.2.1 The LOTOS-EUROS model and the Local Ensemble Transform Kalman Filter (LETKF)

Model description

In this study, we utilized the LOTOS-EUROS (LOng Term Ozone Simulation-EURopean Operational Smog) v2.3.0 chemical transport model to simulate atmospheric concentrations over the study region (Manders et al., 2017). LOTOS-EUROS is a three-dimensional Eulerian model designed for regional air quality assessments and operational forecasting in Europe. It effectively simulates the dispersion, chemical transformation, and deposition of atmospheric pollutants, including gases and aerosols. LOTOS-EUROS is part of the Copernicus Atmospheric Monitoring Service (CAMS) European air quality ensemble (Colette et al., 2025). 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 (Peuch et al., 2022; Colette et al., 2025). LOTOS-EUROS has also participated in numerous model inter-comparisons, typically showing a strong performance (Bessagnet et al., 2016; Colette et al., 2017; Vivanco et al., 2017).

The model incorporates detailed representations of atmospheric processes such as advection, diffusion, and chemical reactions. It uses the Carbon Bond Mechanism IV (CBM-IV) for gas-phase chemistry (with an option to use the more complex CBM-VI scheme), which includes a comprehensive set of reactions relevant to ozone formation and other photochemical oxidants. Aerosol dynamics are modelled using size-resolved modules that account for primary emissions, secondary formation, and processes like coagulation and deposition.

Model outputs, including concentrations of key pollutants such as ammonia (NH_3), ozone (O_3), nitrogen dioxide (NO_2), and particulate matter (PM_{10} and $PM_{2.5}$), are regularly validated against observational data from the Dutch LML air quality monitoring network, measurements from the German environmental agency (the Umweltbundesamt; UBA) and the EBAS network throughout the European Union.

The LETKF system

The Ensemble Kalman Filter (EnKF; Evensen, 2003) is a sequential data assimilation technique that combines observations, here from satellite instruments, with a chemical transport model (CTM) to optimize the model's concentration and emission fields. The LETKF applies this principle by integrating model simulations with observations to produce an improved estimate of the system state.

In this study, we use LOTOS-EUROS LETKF v3.0.7, which has previously been applied to particulate matter (Lopez-Restrepo et al., 2020) and NH_3 (van der Graaf et al., 2022). The implementation follows the formulation of Hunt et al. (2007) and the approach described by

Shin et al. (2016). Compared to the standard EnKF, the LETKF is more computationally efficient because it performs the analysis locally for each grid cell, using only nearby observations determined by a specified spatial correlation length. Temporal variability in emissions is represented through a temporal correlation length τ (set to 3 days for NH_3) and a correlation coefficient α defined as (Lopez-Restrepo et al., 2020; van der Graaf et al., 2022):

$$\alpha_k = e^{-|t_k - t_{k-1}|/\tau},$$

where t_k and t_{k-1} are consecutive time points. The state vector includes three-dimensional trace-gas concentrations and two-dimensional emission perturbation factors, represented by an ensemble of *N* members to capture uncertainties in both the model and observations. Here, *N*=12, as recommended by van der Graaf et al. (2022) for balancing statistical robustness and computational efficiency. Ensemble initialization samples perturbation factors from a normal distribution with mean zero and standard deviation one, which can be adjusted if needed.

To ensure efficiency and avoid spurious correlations, the LETKF applies spatial and temporal localization in a per grid cell approach following the method described by Shin et al. (2016). For spatial localization, the simulated observations are first computed for all ensemble members, and then for the given grid cell to be analysed, all observations (both simulated and real) within 3.5ρ distance are gathered, with ρ being the specified correlation length. The analysis is then performed using the observations collected, with the weight of the additional observations being limited by a Gaussian decay function (Shin et al., 2016):

$$\mathrm{w}(\Delta \mathrm{d}) = \exp\left(-\frac{\Delta \mathrm{d}^2}{2\rho^2}\right),$$

where Δd is the distance between the observation and the point of the model grid. The spatial correlation length was estimated based on the pixel sizes for MTG-IRS, which is expected to be roughly 4×4 km^2 at nadir. To account for larger off-nadir pixel sizes, a slightly wider correlation length of ρ = 8 km was used in the assimilation.

For a full description of the LOTOS-EUROS LETKF and its theoretical formulation, we refer the reader to van der Graaf et al. (2022).

Configuration and Inputs

Here a short summary is provided of the most important model inputs and configuration parameters that were used in the LOTOS-EUROS simulations. For a more detailed description of the LOTOS-EUROS model we direct the reader to Manders et al. (2017).

The LETKF v3.0.7 is coupled to version 2.3.0 of the LOTOS-EUROS model. The model domain covers the entirety of Europe (25°W to 45°E; 35°N to 70°N) at a spatial resolution of 0.1°. The ECMWF EAC4 boundary conditions are used for the simulation, and 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 the vertical layer structure of the ECMWF meteorology dataset. The base simulation and assimilation were performed for the period of March 2019, with a 5-day model spin-up to sufficiently initialize the chemistry and transport beginning on 25 February 2019.

The anthropogenic emissions applied in this study are taken from the CAMS REG inventory (Kuenen et al., 2021), version 6.1.1 Ref2 v2.1, which provides emissions at the European

scale. For ammonia emissions from agricultural sources, a meteorologically dependent parameterization is used that accounts for weather-driven shifts in fertilizer application timing. Emissions are vertically distributed according to sector-specific injection heights, which is particularly important for sources such as industry and public power plants where average stack heights strongly influence dispersion.

The purpose of assimilation in this context is to optimize the model state to obtain improved estimates of key atmospheric components that strongly influence air quality predictions. This includes atmospheric NH_3 concentrations across the vertical column and near the surface, which affect transport and chemistry; emission estimates, particularly for agriculture, which are uncertain and highly variable; and deposition fluxes (wet and dry), which depend on both concentrations and meteorology. By adjusting the model state through perturbations to the emissions to better match observed total columns, the LETKF indirectly improves these processes and outputs, leading to more accurate simulations of column and surface concentrations, as well as deposition.

Matching the model to satellite footprints and enabling assimilation is carried out using the CAMS Satellite Operator (CSO; https://ci.tno.nl/gitlab/cams/cso). CSO is an open-access tool developed at TNO to support efficient intercomparisons between modelled and satellite concentrations. It consists of (i) a pre-processor that downloads, selects, and reformats satellite observations into a common structure, along with post-processing routines for aggregation and visualization, and (ii) a source-code module designed for use within regional air quality models and data assimilation systems such as LOTOS-EUROS. The CSO module can read pre-processed files, simulate satellite observations from model variables, and apply the appropriate observational operators.

3.2.2 The EMEP model and the Local Fractions method

Model description

The EMEP model (Simpson et al., 2012) is a 3D Eulerian Chemical Transport Model developed by the Norwegian Meteorological Institute. It simulates air pollution by considering processes like emissions, transport, chemical reactions, and deposition to calculate air concentrations and deposition fields for pollutants such as acidifying and eutrophying substances, photo-oxidants, and particulate matter. The model is part of the ensemble of regional air qualities that provides daily forecasts (Peuch et al., 2022) under the Copernicus Atmospheric Modelling Service (CAMS). In this study, revision 5.4 was used, with processes modelled similar as used operationally in 2024. The model version used for the assimilation experiments differed in some aspects from the version used to create the nature run for the LOTOS-EUROS experiments, for example in using a limited domain (as shown in Figure 2). Also the emission inventory is slightly different; it is similar as used in the EMEP-Reporting (Fagerli et al., 2023) which are also based on the total emissions reported by countries, but uses a different spatial allocation.

Local fraction output

A particular feature that is used here is the option to write out Local Fraction information (Wind et al., 2020). If enabled, the model keeps track of the source cell where selected tracer gases were emitted up to 24 hours ago. The source region that is kept track of is limited to a selected region of cells around the grid cell for which the local fractions should be calculated, which are

in practice all grid cells of the model. This is illustrated in Figure 2. For the selected location in the middle panel (marked with a purple star) and for the selected time, the model has kept track of the NH_3 emitted during the last 24 hours in the purple box of, in this configuration, the 11x11 cells around the selected cell. The output then contains the fraction of the NH_3 concentration at (in this case) the surface that was emitted from each of the cells in the box for each of the past 24 hours. As illustration, the panels on the left show the fractions for the past 3 hours; in total 24 past records are available. As NH_3 in the atmosphere is short-lived, the fractions quickly decrease when originating from grid cells further away and released more hours ago. In this example, about 82% of the NH_3 in the selected cell is emitted in the past hour, and 9% one hour ago, and 4% two hours ago. The figure on the right shows for the selected moment the total fraction of NH_3 at the surface could be traced back to the surrounding emissions during the past 24 hours. In locations with high NH_3 emissions the local sources are dominant, and can explain almost 100% of the concentrations. However, at other locations where NH_3 is for example chemically created out of ammonium, only 40-50% of the concentrations could be explained in this way.

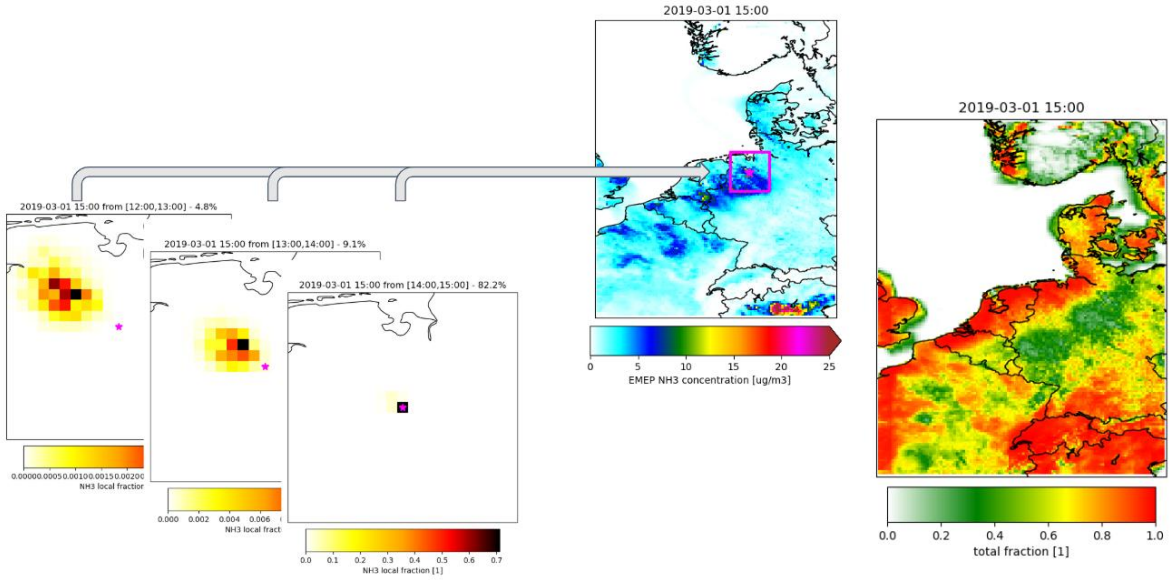

Figure 2 – Illustration of the *Local Fraction* output of the EMEP model, in this case for simulated NH_3 concentrations. The 3 panels on the left show the computed Local Fractions for the grid cell marked in the middle panel, for the past 3 hours. The panel on the right shows the total fraction of NH_3 surface concentrations that can be explained by the local emissions.

Although computationally efficiently implemented, the output of the Local Fractions makes a simulation more expensive, and also requires a huge storage facility. However, when available, the Local Fractions are a powerful method to relate emissions with (observed) concentrations. This is exploited in this study to develop a method that could estimate emissions of NH_3 based on satellite observations.

ELVIS – EMEP Local-fraction Variational Inversion System

The EMEP Local-fraction Variational Inversion System (ELVIS) has been developed in this project as a tool to employ the Local Fraction output to estimate emissions based on observations. It employs a variational approach of minimizing a cost function J over the elements of a state vector $\boldsymbol{x}$:

$$J(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_b)^T\boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{x}_b) + \frac{1}{2}(\boldsymbol{H}(\boldsymbol{x})-\boldsymbol{y})^T\boldsymbol{R}^{-1}(\boldsymbol{H}(\boldsymbol{x})-\boldsymbol{y})$$

In here, the following elements are present.

- The vector ***x*** contains emission deviation factors *x(i,j,t)* per grid cell *(i,j)* and deviation time step *t*. The factors are values around zero, and expand the background emissions e_b to deviated emission:
$$e(i,j,t) = e_b(i,jt) \times (1 + x(i,j,t))$$
The minimum value of *x* is -1 for an emission equal to zero. Note that the emissions itself will actually not be used in the inversion, as their amount is implicitly included in the Local Fractions; the only thing used is relative change. In this application, the emissions consists of hourly gridded NH_3 emissions, but the emission deviations do not necessarily have the same temporal resolution, for example because there are not enough observations to constrain hourly fields. A suitable deviation resolution is daily, as NH_3 emissions are often strong on just a few days in spring.
- The *prior* state $\boldsymbol{x}_b$ is set to zero, meaning no deviations of the emissions.
- The *background covariance* matrix ***B*** should describe the uncertainty in the deviation factors and how these are correlated. A standard deviation (square root of main diagonal of ***B***) of 50% is assumed initially, as in our experiments synthetic observations are used that were created using model simulations that are not too different from the model version used for inversions. A spatial correlation for the emission deviations is parameterized using a Gaussian decay. A rather short length scale of 25 km is assumed, as the NH_3 emissions are assumed to vary strongly over short distance given differences in agricultural practices and weather conditions. Currently no temporal correlation between the daily deviation patterns is assumed, although the system is able to include that.
- The observation vector ***y*** contains the (in this case) synthetic IRS/NH_3 observations derived from an EMEP nature run, as will be described in section 3.2.3.
- The *observation operator* ***H(x)*** simulates the observations given the emission deviation factors in the state ***x***. The operator is itself also a series of operations applied after each other:
$$\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{G}\,(\,\boldsymbol{c} + \boldsymbol{I}(\boldsymbol{c})\,\boldsymbol{F}\,\boldsymbol{x}\,)$$
where:
    - ***c*** is the 3D concentration array simulated by the EMEP model in the run that produced the local fraction output;
    - ***F*** is the local fraction operator that for each cell in the 3D grid multiplies the corresponding local fractions with the emission deviation factors x and sums the result;
    - the diagonal matrix ***I(c)*** with ***c*** on its main diagonal transform from a concentration fraction to concentration value;
    - satellite operator ***G*** samples a concentration array at the location of the footprints and applies the averaging kernel convolution that will be described in section 3.2.3.
- The *observation representation error covariance* ***R*** describes the expected error between the observations in ***y*** and the simulations $\boldsymbol{H}(\boldsymbol{x}^t)$ given the (unknown) *true* deviations $\boldsymbol{x}^t$ .In this study, the errors are assumed be uncorrelated with a standard deviation equal to the assumed retrieval error that is supplied with the synthetic product.

The value of ***x*** for which *J(**x**)* reaches a minimum is a balance between perturbing the emissions within the assumed uncertainty, and providing simulations close to the observations. The ‘ratio’ between the covariances B and R defines the optimal value: for a large background covariance B the proposed deviations in emissions will be large too but the

simulations will be close to the observations, while for a large observation representation error R the emission deviations will be smaller but the simulations will also be further off from the observations.

The optimal value of ***x*** is determined in an iterative minimization procedure following the M1QN3 algorithm (Gilbert & Lemaréchal, 1989). The procedure requires computation of the gradient of *J(**x**)* towards the elements of ***x***. This gradient could be efficiently computed using the transpose of the local fraction operator ***F***, which could be done with almost any addition costs when evaluating ***H(x)***. The majority of the computational costs are actually in reading the local fraction arrays from the storage, and when this is done, the actual operations on it are rather cheap.

Note that using the Local Fractions for emission inversion assumes that the response in concentrations (or actually, observed concentrations or columns) is linear with respect to changes in emissions. For short-lived tracers such NH_3 this is a valid assumption, as there is not much time for higher order chemical interaction with other tracers.

3.2.3 Synthetic MTG-IRS NH_3 Observations

The synthetic observations for MTG-IRS were created using the proposed dwellings over Europe provided by EUMETSAT with the sample L1B data[1], as shown in Figure 3. The original samples were converted into a clean (empty) dataset with this footprint mapping in the CSO[2] file format. For experiments of this type, it is important that the selected “nature” run is sufficiently different from the model state. For the cross-OSSE approach performed here, two “nature” runs were used: one with the EMEP model for use in LOTOS-EUROS assimilations, and one with LOTOS-EUROS for use in EMEP inversions. In the description below, results are shown for the EMEP-based nature run, as this was used for the initial tests. When the procedure was defined, the final creation of synthetic observations was also applied to the LOTOS-EUROS “nature” run.

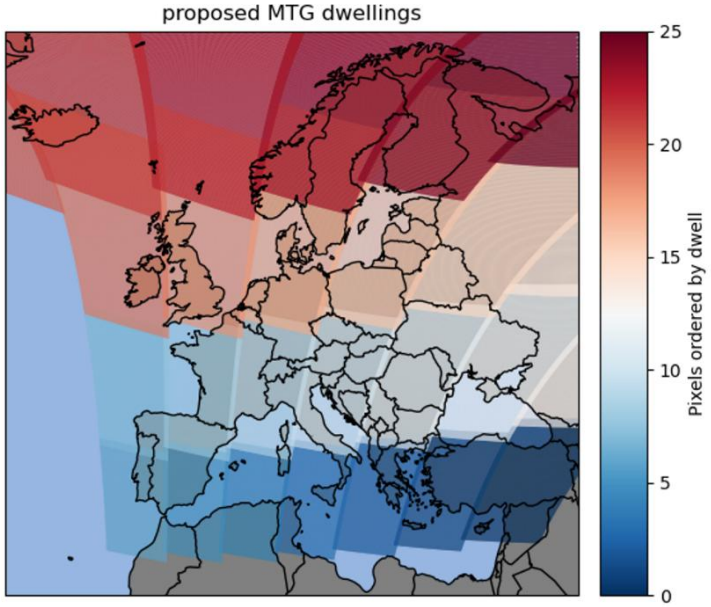

Figure 3 – Proposed MTG-IRS footprint dwellings with the pixels coloured by dwell.

[1] https://user.eumetsat.int/resources/user-guides/mtg-test-data
[2] https://ci.tno.nl/gitlab/cams/cso

To support the OSSE framework, two distinct sets of synthetic MTG-IRS observations were generated, each serving a specific purpose. The first set represents an idealized scenario, designed as a best-case benchmark and used primarily to verify that the assimilation systems were functioning correctly under optimal conditions. These observations assume perfect vertical sensitivity by applying a unit averaging kernel (value of 1 at all levels), meaning there is no influence from an a priori vertical profile. This configuration allows us to isolate the assimilation system's performance without introducing complexities related to retrieval characteristics. For measurement errors, we adopted estimates from the Cross-track Infrared Sounder (CrIS) Fast Physical Retrieval (CFPR) NH_3 product described by Shephard et al. (2020), rounding the reported total relative errors for simplicity (see Table 1). This idealized dataset was created for March 2019 by sampling EMEP vertical profiles at MTG-IRS footprint locations.

Table 1 - Total uncertainties assigned to synthetic NH_3 observations binned by total column concentration range derived from the CrIS CFPR product (Shephard et al., 2020).

NH_3 total column range (molec. cm^{-2})	Total relative uncertainty (%)
$<0.5x10^{16}$	40.0
$0.5 – 1.0x10^{16}$	34.0
$1.0 – 2.5x10^{16}$	31.5
$>2.5x10^{16}$	30.0

In addition to the idealized set, a second realistic scenario was developed to approximate the expected performance of MTG-IRS under normal operating conditions. These observations incorporate retrieval characteristics such as vertical sensitivity, diurnal variability, and thermal contrast, which are factors that are likely to influence real MTG-IRS retrievals. Specifically, the mean CrIS averaging kernel and a priori profiles from the CFPR v1.6.4 product (Shephard et al., 2015; 2020) calculated over the entire model domain for March 2019 (shown in Figure 4) were applied to the EMEP model profiles at the footprint locations. The mean averaging kernel (AVK) (Figure 4a) indicates sensitivity between roughly 1000 and 500 hPa, covering much of the surface to mid-troposphere where elevated NH_3 concentrations typically occur. This is consistent with the mean a priori profile (Figure 4b), which shows the highest NH_3 concentrations near the surface, decreasing steadily up to approximately 600 hPa.

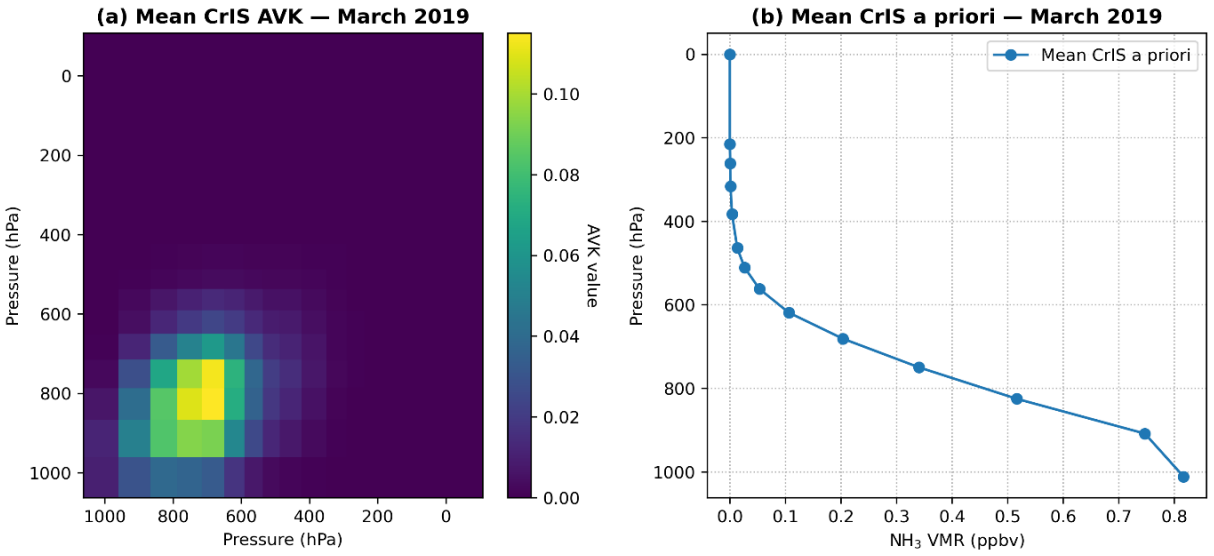

Figure 4 – (a) Mean CrIS NH_3 volume mixing ratio AVK on the retrieval pressure layers for March 2019, and (b) mean CrIS NH_3 a priori profile.

Both datasets account for cloud screening based on ECMWF cloud fraction outputs from LOTOS-EUROS: any simulated MTG-IRS pixels with cloud coverage greater than 30% were excluded to mimic operational filtering. By comparing assimilation results using these two datasets, we can distinguish between improvements achievable under ideal conditions and those likely under real-world constraints. This dual approach not only validates the assimilation systems but also provides critical insight into the robustness of the methodology when faced with realistic observational uncertainties.

The averaging kernel and a priori profile are applied to the EMEP volume mixing ratio (VMR) profiles sampled at the footprints to produce the simulated retrieval $\hat{x}$ following the approach of Rodgers & Connor (2003):

$$\hat{x} = x_a + A(x_{EMEP} - x_a),$$

with x_a the a priori profile, A the VMR averaging kernel, and x_{EMEP} the EMEP NH_3 profile. To account for the expected poor observational conditions and higher incidences of failed retrievals during the dawn and dusk hours of each day, we removed these observations based on the solar zenith angle (SZA). Here dawn and dusk observations were selected using the criterion of -18° < SZA < 0°. In addition, to account expected higher uncertainties for retrievals during the nighttime, we inflated the relative uncertainties by a factor of 2 for any observations with SZA less than -18°. A plot of the mean simulated retrieved column and the mean unsmoothed (i.e., no averaging kernel or a priori applied) EMEP columns for the month of March 2019 are shown in Figure 5a and Figure 5b, while the mean absolute and relative differences are shown in Figure 5c and Figure 5d, respectively.

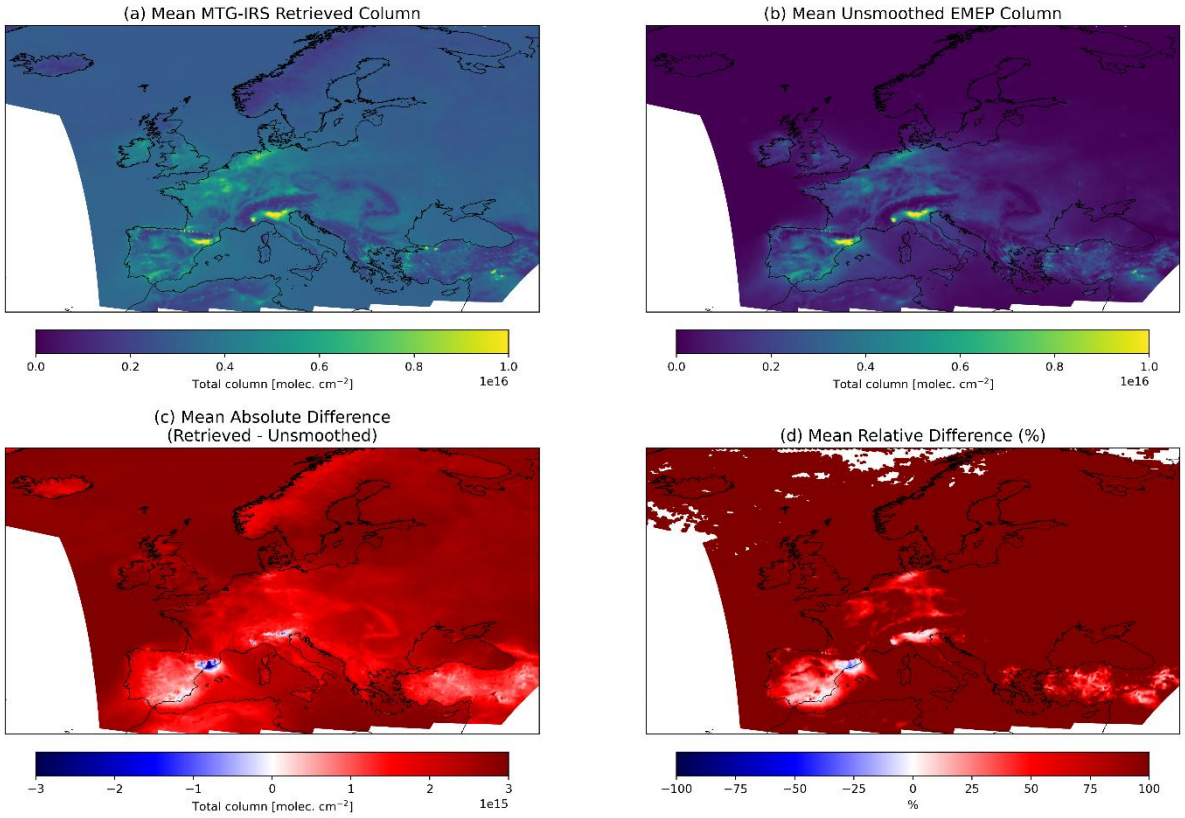

Figure 5 – Means calculated over the period of March 2019 of (a) simulated MTG-IRS retrieved NH_3 total column, (b) the mean EMEP NH_3 total columns at the MTG-IRS footprints used to simulate the retrievals, (c) the mean absolute differences (retrieved – EMEP), and (d) the mean relative differences ((retrieved – EMEP)/ EMEP) in %.

In general, the application of the averaging kernel led to increased total columns, particularly in areas where concentrations were low in EMEP, such as over the ocean. This effect can primarily be attributed to the mean CrIS a priori profile having higher concentrations on average than the corresponding EMEP profiles. In a few regions such as Northern Spain and the Po Valley in northern Italy, the relative differences in the MTG-IRS simulated and un-smoothed EMEP columns are lower or negative. In general, the simulated observations capture the spatial patterns of the NH_3 concentrations shown in the EMEP simulation.

In the following subsections, we will discuss the results of the assimilation tests on the modelled emissions, deposition and concentration fields using these synthetic observations.

3.3 Results and Discussion

3.3.1 Assimilation of "idealized" observations in LOTOS-EUROS

The purpose of this first assimilation experiment using the "idealized" set of MTG-IRS synthetic observations was to verify that the LETKF framework can successfully adjust the model state under simplified and well-controlled conditions. In this setup, the observations were assumed to be unbiased and uniformly distributed throughout the day, with a consistent level of uncertainty. This allows us to test the system's ability to respond dynamically to a steady and information-rich stream of geostationary measurements before introducing additional real-world complexities such as the application of averaging kernels and a priori profiles, enhanced nighttime uncertainties, and the omission of dawn and dusk observations.

The assimilation was conducted for a shorter five-day period (1-5 March 2019), enabling a detailed examination of the model's response over consecutive assimilation cycles. The analysis focused on three main aspects: the information content of the assimilated observations, the resulting emission adjustments, and the corresponding impacts on surface and column concentrations. These diagnostics provide a first assessment of whether the LETKF correctly propagates observational information into emission updates and concentration changes under idealized conditions.

Information content

To evaluate the impact of the number of available satellite observations on the assimilation and how this might correspond with adjustments to the emission and concentration fields, we can estimate the number of independent pieces of information provided by the satellite measurements through the Degrees of Freedom for Signal (DOFS) in the model domain. For reference, a DOFS of 1 in a given model cell at a given time point implies that there is 1 full independent piece of information from the observations that can be used to adjust the emissions in that cell, while a DOFS of 0.5 or less implies that the observations only partially inform the emissions adjustment in that grid-cell. Here, the DOFS are calculated on the model grid using the method described in Chen et al. (2023):

$$DOFS = 1 - \hat{S}' \mathrm{S_a^{-1}}$$

where $\hat{S}'$ is the a posteriori error covariance matrix, and $\mathrm{S_a}$ is the a priori error covariance matrix, both of which are directly outputted per time-step from the LETKF. The resulting DOFS can be a useful diagnostic for understanding and interpreting the underlying emissions and concentration changes.

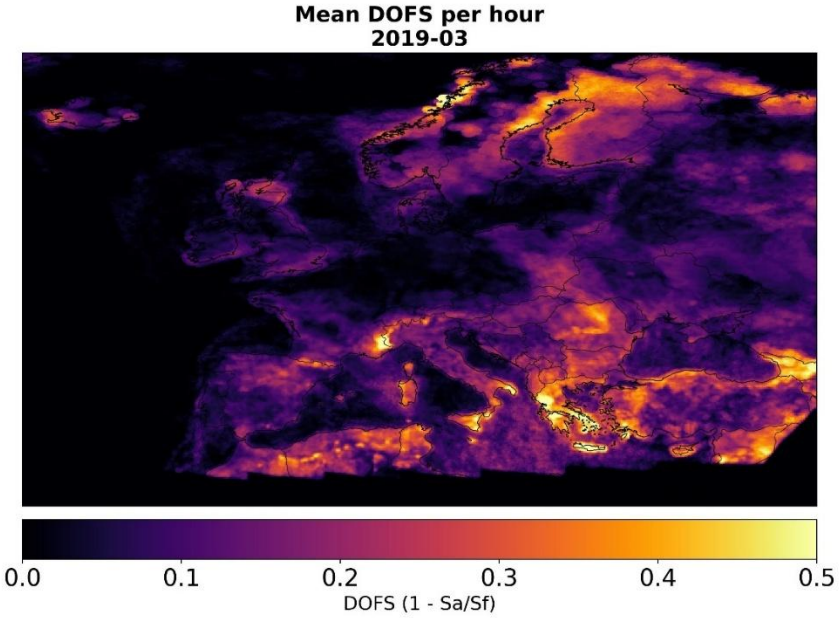

Figure 6 – The mean DOFS per hour calculated in the model domain for the period of 1-5 March 2019 for the "idealized" observations case.

A map of the mean DOFS per hour from the "idealized" observation test is shown in Figure 6. In the simulation, the total DOFS are mostly driven by the observations available in the given grid-cell which is (in this case of a geostationary sounder) a product of the total percentage of cloud-cover affecting that grid-cell. From this, we can see that there is a pattern of higher DOFS over much of southern Europe, while there is a region of lower DOFS over Germany and the Benelux (Belgium-Netherlands-Luxembourg) region which was affected by higher cloud-coverage during the first half of the month. The mean DOFS in the "idealized" case shown here is expected to be higher than the "realistic" observations set (shown later) since the dawn and dusk observations were not removed.

In general, the synthetic geostationary MTG-IRS observations provide a significant amount of information to the assimilation system (measured here with the DOFS) which is used to inform the emissions changes, in particular when compared to current polar-orbiting sensors such as CrIS that only provide (at best) two overpasses per day over an area but also have poorer spatial resolution (footprints of roughly 15 km^2 at nadir).

Total column concentrations

Because the LETKF system adjusts the model state via perturbations to the emissions to achieve the closest match to the observed total columns (accounting for measurement uncertainties) our evaluation begins with these quantities. Figure 7 shows the mean total columns at MTG-IRS footprint locations from the synthetic data, base run and assimilation run for 1–5 March 2019, along with the absolute differences between both model runs and with the synthetic MTG-IRS observations. As seen in Figure 7a to Figure 7c, extensive cloud cover over much of northern Europe during this period limited observation availability in that region, resulting in lower total column values (a pattern also reflected in the mean DOFS shown in Figure 6).

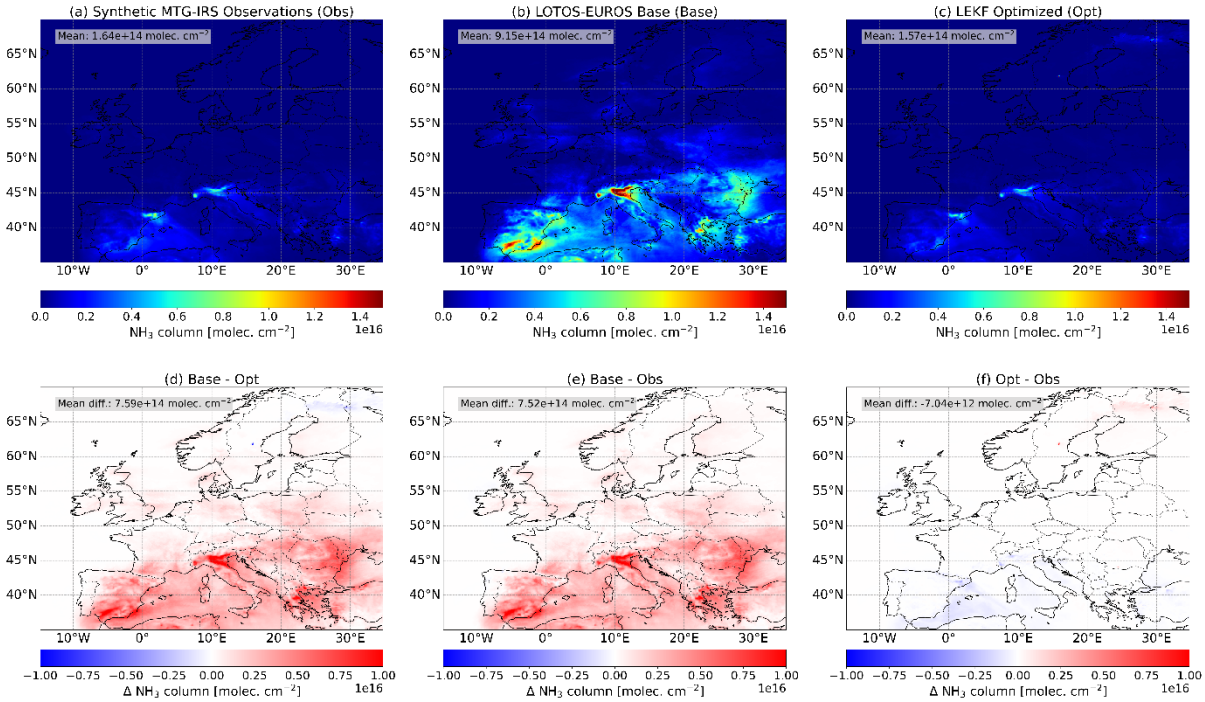

Figure 7 – The mean NH_3 column concentrations at the satellite footprints for 1-5 March 2019 from (a) the synthetic MTG-IRS observations, (b) the base LOTOS-EUROS simulation, (c) the LETKF optimized simulation, and the differences for (d) LETKF optimized simulation minus the base simulation, (e) base simulation minus the observations, and (f) LETKF optimized simulation minus the observations .

The difference plot for the base simulation versus synthetic observations (Figure 7e) reveals that LOTOS-EUROS exhibits a positive bias, particularly across the southern half of the domain, with a mean bias of $+7.52\times10^{14}$ molec. cm^{-2}. After assimilation, these differences are substantially reduced across the domain, with the mean difference dropping to $+7\times10^{14}$ molec. cm^{-2}. Small residual negative biases remain in northern Spain and northern Italy, consistent with patterns seen in surface concentrations (Figure 9f). Overall, the marked reduction in total column bias demonstrates that the LETKF analysis effectively adjusted the model toward the atmospheric state represented by the synthetic MTG-IRS observations, providing confidence in the system's capability to assimilate high-density geostationary NH_3 measurements.

Emissions

As emissions are the primary control variable directly adjusted during the assimilation, we begin by examining the resulting emission fields.

The total NH_3 emissions from EMEP, the LOTOS-EUROS base simulation, and the LETKF optimized simulation and the absolute differences between each case are shown in Figure 8. The total EMEP emissions (mean over the entire domain of 25.72 mg/m^2) are lower than the emissions from the base LOTOS-EUROS simulation (mean over the domain of 54.72 mg/m^2) across much of the model domain, with some exceptions that can be seen in of Figure 8e. Most notably, EMEP shows higher emissions in parts of northern and southern Germany, Northern Spain and Turkey. The higher emissions in LOTOS-EUROS compared to EMEP can primarily be attributed to the difference in the CAMS REG emissions datasets used.

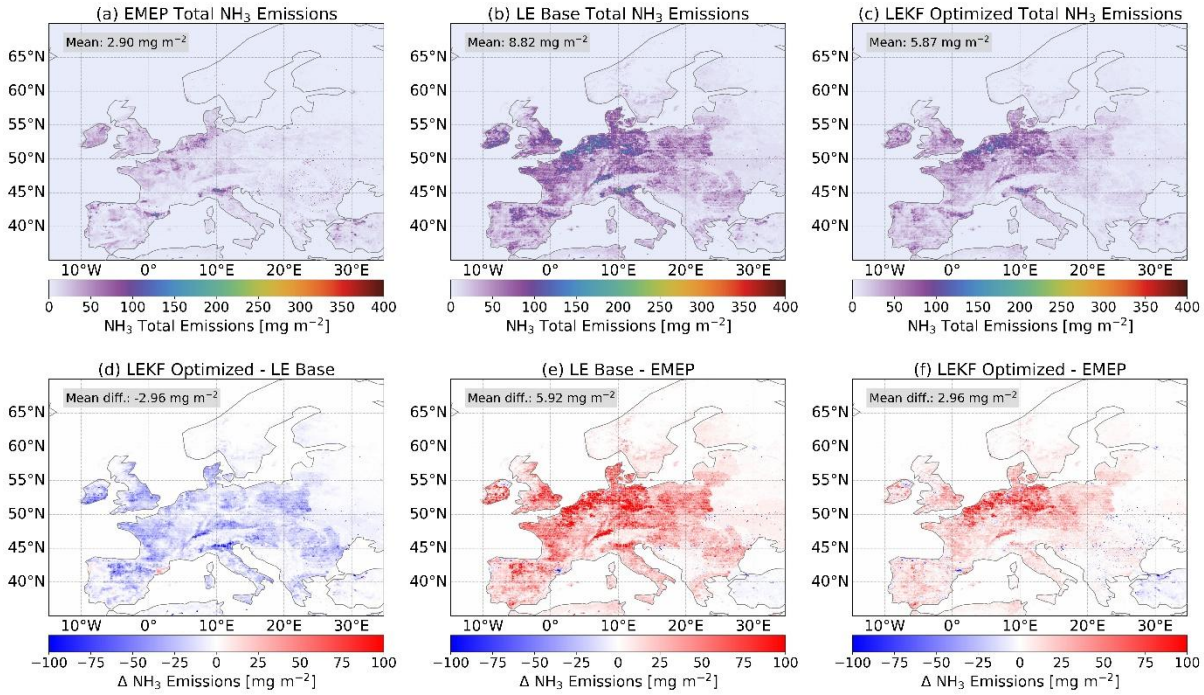

Figure 8 – The total NH_3 emissions for 1-5 March 2019 from (a) EMEP, (b) the base LOTOS-EUROS simulation, (c) the LETKF optimized simulation, and the differences for (d) LETKF optimized simulation minus the base simulation, (e) base simulation minus EMEP, and (f) LETKF optimized simulation minus EMEP.

After the assimilation of the synthetic MTG-IRS observations, the emissions in LOTOS-EUROS are broadly reduced and the spatial pattern of the EMEP emissions is better captured, which is visible in Figure 8f. The mean difference was reduced from +5.92 mg/m^2 to +2.96 mg/m^2, with noticeable decreases in the biases across much of southern Europe, indicating that the assimilation of "idealized" MTG-IRS synthetic observations is effective at bringing the model state closer to that of the EMEP nature run. Some of the remaining biases between the optimized and EMEP emissions can be attributed to differing spatial patterns in the two underlying emissions inventories, as the LETKF is only able to scale the base a priori emissions (via the application of a multiplicative mask) but has difficulties redistributing the broader spatial patterns of emissions to and from new regions. Furthermore, the pattern of emission changes largely reflects the distribution of the mean DOFS that was shown in Figure 6, which is logical because in the regions of lower DOFS there is less information from the observations of the "true" state that can be used to adjust the emissions and the a priori state (the original emissions) is favoured more heavily.

To evaluate the impact of the assimilation on deposition, an additional forward model run would normally be required in which the optimized multiplication factor fields are applied to the emissions before the forecast is propagated in time. However, since the purpose of this initial test was primarily to verify the performance of the LETKF system with the synthetic observations, this additional step was omitted. The analysis of optimized deposition will therefore be deferred to the final assimilation experiment using the "realistic" observation setup, discussed later in this section.

Surface concentrations

In addition to the total column concentrations, we also compare the surface level NH_3 concentrations pre- and post-assimilation. The surface NH_3 concentration fields from EMEP, the base LOTOS-EUROS simulation, and the LETKF optimized simulation, as well as the differences, are shown in Figure 9. Corresponding to the higher emissions in the base LOTOS-

EUROS simulation compared to EMEP, a similar pattern is evident in the surface concentrations. From Figure 9a and Figure 9b, it can be seen that the surface concentrations are significantly higher in the base LOTOS-EUROS simulation than those from EMEP throughout the domain, with overall domain-wide means of 1.98 ppb and 0.66 ppb, respectively, and a mean absolute difference of +1.31 ppb. Post-assimilation (Figure 9c), the surface concentrations are reduced across much of the model domain, particularly in southern Europe. The biases between the LETKF optimized surface concentrations and EMEP (Figure 9f) are substantially reduced in Spain, Italy, France and much of the Balkans, with the mean difference decreasing from +1.31 ppb to +0.34 ppb.

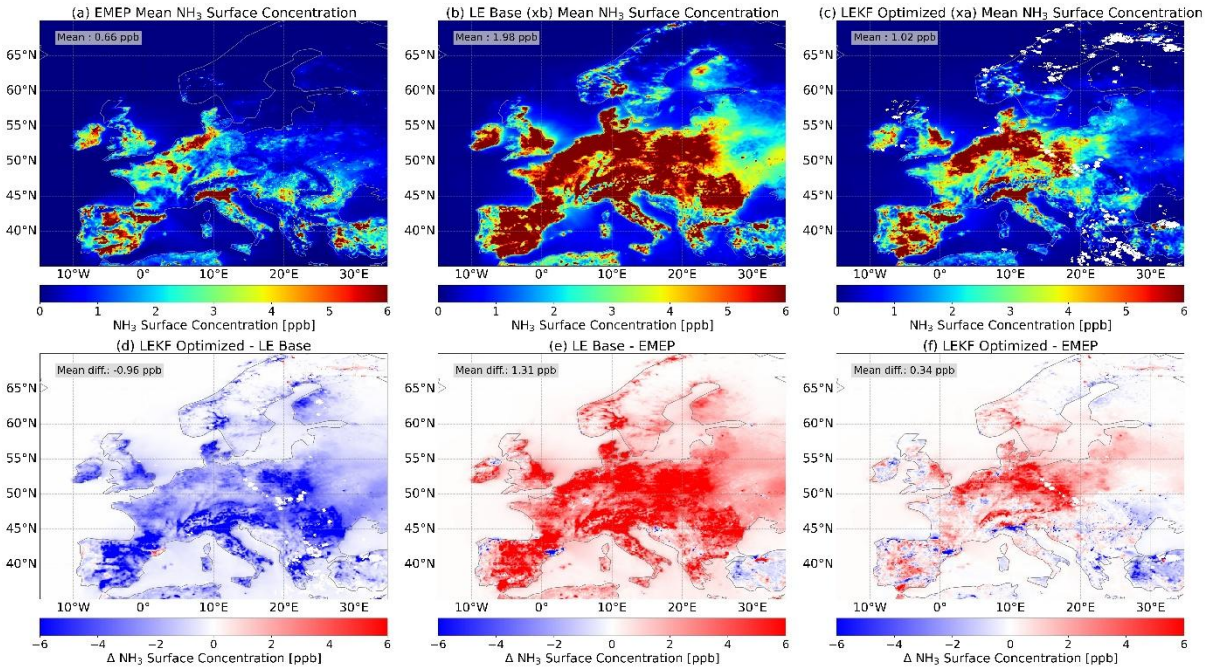

Figure 9 – The mean NH_3 surface concentrations for 1-5 March 2019 from (a) EMEP, (b) the base LOTOS-EUROS simulation, (c) the LETKF optimized simulation, and the differences for (d) LETKF optimized simulation minus the base simulation, (e) base simulation minus EMEP, and (f) LETKF optimized simulation minus EMEP.

In parts of Turkey, northern Spain, and northern Italy, the LETKF yields slightly lower column NH_3 than the EMEP nature run, producing a small negative bias. This likely reflects mismatches in vertical allocation and diurnal emission profiles, as well as differences in how the two models translate emissions into column concentrations through meteorology and vertical transport. In regions with complex terrain, such as Northern Spain and the Po Valley, these discrepancies are further amplified by differences in boundary-layer dynamics and local circulation patterns, which can shift plume position or vertical mixing depth. Moreover, when the observation operator assumes a vertical NH_3 profile that is more surface-weighted than in the EMEP nature run, a given negative innovation can translate into an overly strong near-surface adjustment, effectively driving the surface concentrations below the true value. Together, these factors can cause the LETKF gain to over-correct in isolated areas. A similar effect was observed in a previous OSSE experiment with Sentinel 5p and Sentinel 4 synthetic observations by Timmermans et al. (2019). Nevertheless, these effects remain spatially limited; over the full domain, the assimilation markedly reduces the median bias and narrows the overall spread.

3.3.2 Assimilation of “realistic” synthetic observations in LOTOS-EUROS

The second assimilation experiment was carried out using the “realistic” set of MTG-IRS synthetic observations to assess the performance of the LETKF under more representative observational conditions. In contrast to the idealized case, this configuration incorporated several sources of realism that are expected in an operational retrieval: the application of averaging kernels and a priori profiles, the use of realistic retrieval uncertainties that vary with time of day, and the exclusion of observations around dawn and dusk where thermal contrast is limited (see section 3.3.1). Together, these factors provide a more realistic representation of the observation sampling and sensitivity expected from the MTG-IRS mission.

The assimilation was performed for the full month of March 2019, allowing the evaluation of the system’s behaviour over a longer sequence of assimilation cycles and under varying meteorological conditions. Here the analysis focuses on the same key aspects as before; the information content of the assimilated observations, the emission adjustments produced by the LETKF, and the resulting changes in surface and column NH_3 concentrations—but now within a more realistic observational framework. This experiment therefore serves as the main test of how effectively the MTG-IRS observations can constrain NH_3 emissions and concentrations in a practical assimilation scenario.

Information content

A map of the mean DOFS per hour over the model domain for the month of March 2019 is shown in Figure 10. In general, the mean information content in the system from the observations is higher in the south of Europe, in particular over Spain, Portugal and Italy, and lower in north-western Europe and Scandinavia, reflecting the presence of clouds In comparison to the DOFS from the “idealized” case shown in Figure 6, the mean values shown here are lower which is driven by the lower number of observations overall (i.e., due to the removal of the dawn and dusk observations each day), however, this is likely more reflective of the true information content that would be provided by MTG-IRS as the observations during this period of the day are likely to be of low quality due to unfavourable observational conditions including poor thermal contrast and rapidly changing vertical temperature gradients.

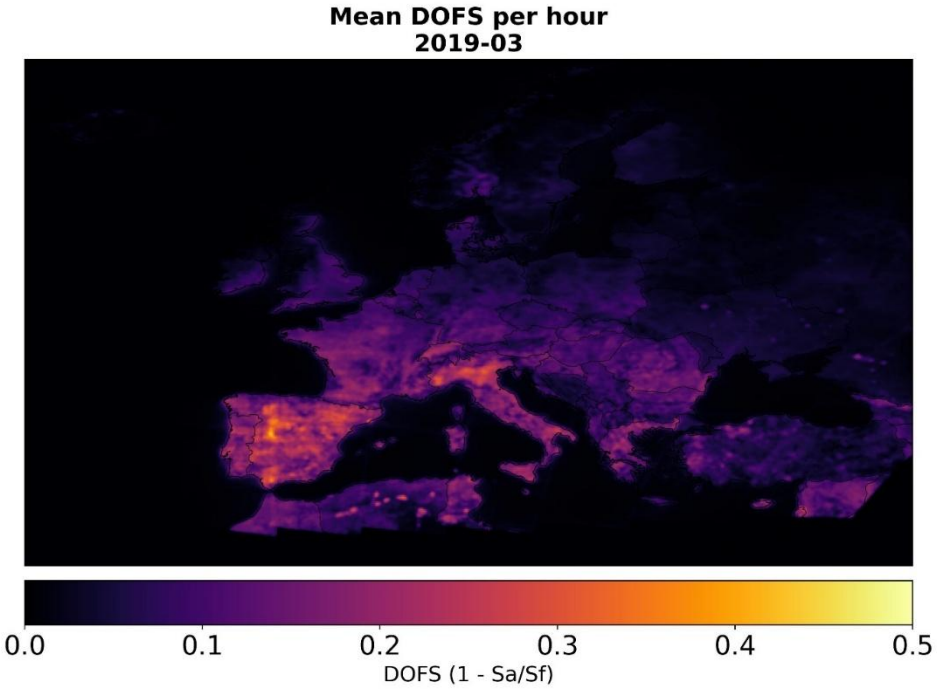

Figure 10 – The mean DOFS per hour calculated in the model domain over the month of March 2019 for the "realistic" observations case.

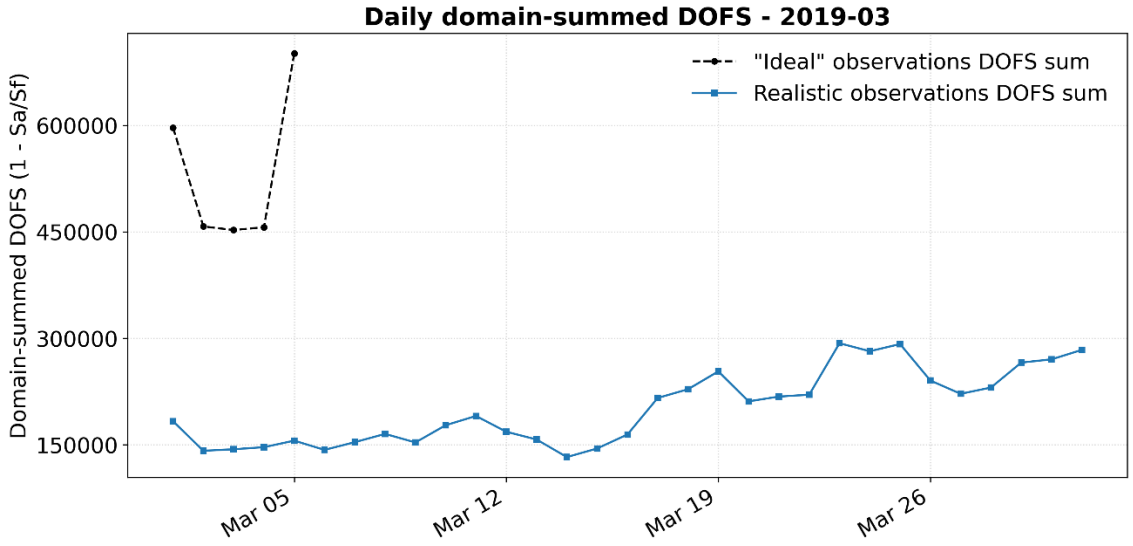

Figure 11 – Total DOFS per day summed over the entire model domain for the month of March 2019. For comparison, the daily total DOFS per day for the "idealized" observations case is shown by the black dashed line.

A time-series of the DOFS for the month from the realistic observations dataset and the "ideal" observations case for the first five days is provided in Figure 11. It can firstly be seen that, as expected, the idealized observations provide higher total DOFS to the system compared to the realistic observations set. Additionally, it can be seen that the total DOFS within the model domain increases towards the end of the month, and this is a result of the lower overall cloud-cover during this period in comparison to the beginning of the month. In particular, the clearer conditions over Belgium, the Netherlands, and Germany during the final ten days of March contribute markedly to the higher DOFS values. These results highlight that cloud cover is the primary factor governing the availability and information content of MTG-IRS observations for

data assimilation. Nevertheless, even accounting for this limitation, a geostationary platform such as MTG-IRS is expected to provide substantially more information than current polar-orbiting instruments such as CrIS and the Infrared Atmospheric Sounding Interferometer (IASI), which typically offer only two clear-sky overpasses per day. The half-hourly temporal sampling and smaller ground footprints of MTG-IRS will further enhance the likelihood of capturing cloud-free observations between intermittent cloud fields, making it far more effective under partly cloudy conditions.

Total Columns

We again first examine the total column NH_3 concentrations at the MTG-IRS footprints before and after assimilation in the "realistic" observations case. Figure 12 presents the monthly mean total columns for March 2019, along with the corresponding absolute differences. Comparing Figure 12a and Figure 12b along with the differences shown in Figure 12e reveals that mean total columns from base LOTOS-EUROS simulation are biased high throughout most of the domain relative to the synthetic MTG-IRS observations (mean column difference of +5.11x1015 molec. cm^{-2}), with the exception of over the sea and the northern boundary of the model domain. It is important to note that the differences observed here between the total columns from the synthetic observations and the base model are broadly representative of the expected magnitude of the differences between the model and real NH_3 satellite observations. Earlier comparisons of IASI and CrIS with LOTOS-EUROS over western Europe in the period of 2014-2022 showed mean differences of -52.7% and -59.9%, respectively. The difference plot of the LETKF optimized total columns versus the MTG-IRS synthetic observations in Figure 12f show that the bias over land is almost entirely gone (small mean difference of -0.59x10^{15} molec. cm^{-2}).

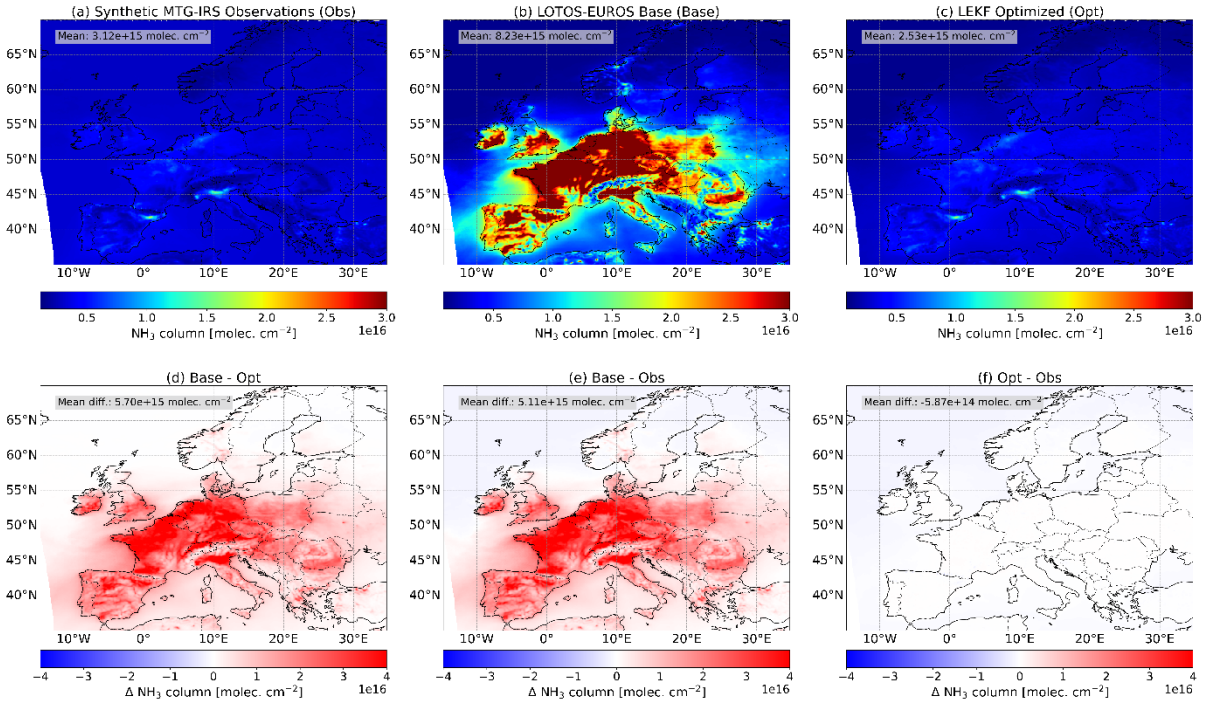

Figure 12 – The mean NH_3 column concentrations at the satellite footprints for March 2019 from (a) the synthetic MTG-IRS observations, (b) the base LOTOS-EUROS simulation, (c) the LETKF optimized simulation, and the differences for (d) LETKF optimized simulation minus the base simulation, (e) base simulation minus the observations, and (f) LETKF optimized simulation minus the observations.

After assimilation, the total column NH_3 concentrations at the MTG-IRS footprints closely reflect the values from the synthetic observations, indicating that the assimilation framework effectively aligns the modelled columns with the observational constraints.

Emissions

Figure 13 compares total NH_3 emissions from the base EMEP nature run, the base LOTOS-EUROS simulation, and the LETKF-optimized simulation using the realistic observations, as well as their absolute differences. The base EMEP emissions for March 2019 exhibit a domain mean of 25.72 mg m^{-2} (Figure 13a), while the base LOTOS-EUROS simulation produces more than double that amount, with a mean of 54.68 mg m^{-2} (Figure 13b). After assimilation of the realistic MTG-IRS synthetic observations, the gap between the EMEP nature run and the LOTOS-EUROS emissions is substantially reduced, leaving a mean difference (LEKF optimized – EMEP) of only +8.95 mg m^{-2}, compared to a mean difference of +28.96 mg m^{-2} pre-assimilation (Figure 13e).

The differences between the LETKF-optimized and base LOTOS-EUROS simulations (Figure 13d) are negative throughout the domain, indicating that emissions were consistently reduced after assimilation. Notably, regions that already showed negative differences in Figure 13e became even more negative following assimilation, reflecting a stronger adjustment in those areas. The spatial pattern of emissions in the optimized LOTOS-EUROS simulation (Figure 12c) more closely matches that of the EMEP nature run (Figure 12a), especially in regions with higher observation density and less cloud cover, where the assimilation system has more information to constrain the emissions.

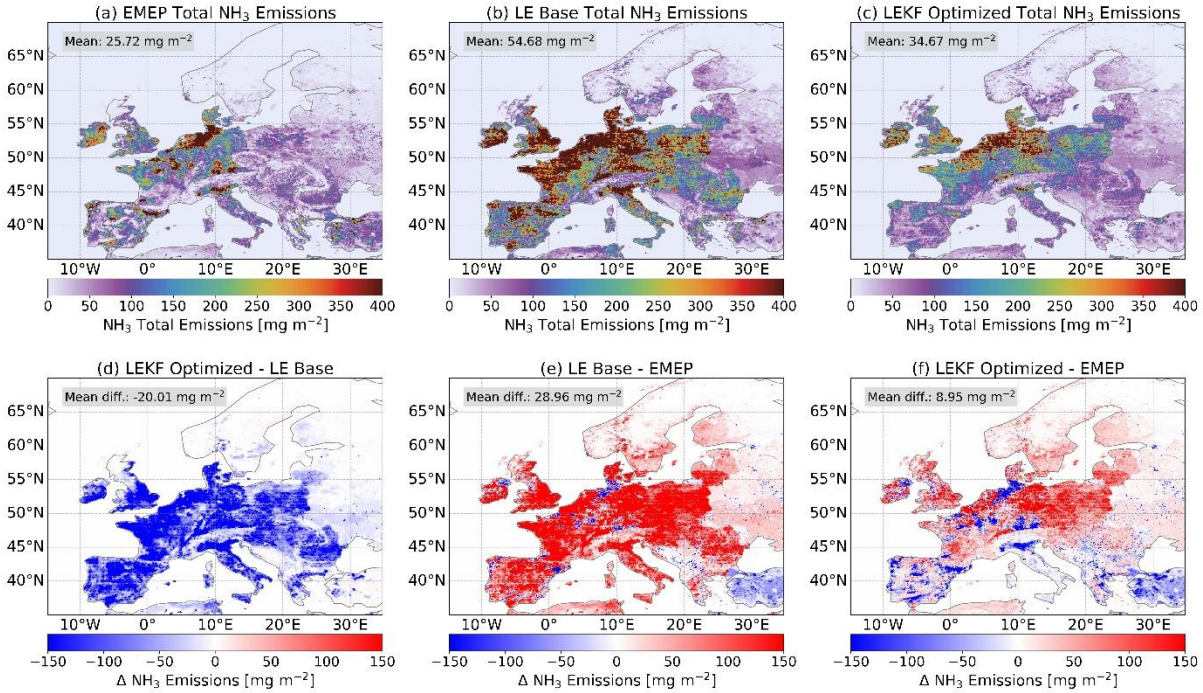

Figure 13 – The total NH_3 emissions for March 2019 from (a) EMEP, (b) the base LOTOS-EUROS simulation, (c) the LETKF optimized simulation, and the differences for (d) LETKF optimized simulation minus the base simulation, (e) base simulation minus EMEP, and (f) LETKF optimized simulation minus EMEP.

The assimilation process adjusts emissions such that the best possible match is achieved between the simulated and observed NH_3 total columns, explicitly accounting for the measurement uncertainties associated with the synthetic MTG-IRS observations. This means

that regions with lower observational uncertainty and higher information content (as quantified by the DOFS) experience more substantial emission corrections. Conversely, areas with persistent cloud cover or higher retrieval uncertainty see less adjustment, as the assimilation system relies more heavily on the prior model state. It is important to note that differences in emissions and deposition still remain between the LETKF-optimized simulation and the EMEP nature run. These residual discrepancies are likely attributable to differences in model parameterizations, such as chemistry, vertical mixing and transport processes, which influence how emissions are distributed and deposited within the atmosphere. As a result, while the assimilation brings the total columns into agreement with the observations, underlying variations in emissions and deposition highlight the impact of model-specific dynamics that might vary between EMEP and LOTOS-EUROS which are not fully resolved by the assimilation process alone.

Despite these advances, as previously discussed, a key limitation of the LETKF approach remains: emissions are adjusted via multiplicative factors per grid cell, restricting the ability to redistribute emissions spatially. As a result, some regional discrepancies persist where the underlying emissions inventories differ substantially.

Surface concentration fields

In addition to the emissions and deposition, we evaluated NH_3 concentration fields for the full month of March 2019. This analysis compares the base EMEP nature run, the base LOTOS-EUROS simulation, and the LETKF-optimized simulation to assess how assimilation of "realistic" MTG-IRS synthetic observations influences modelled concentrations across the domain.

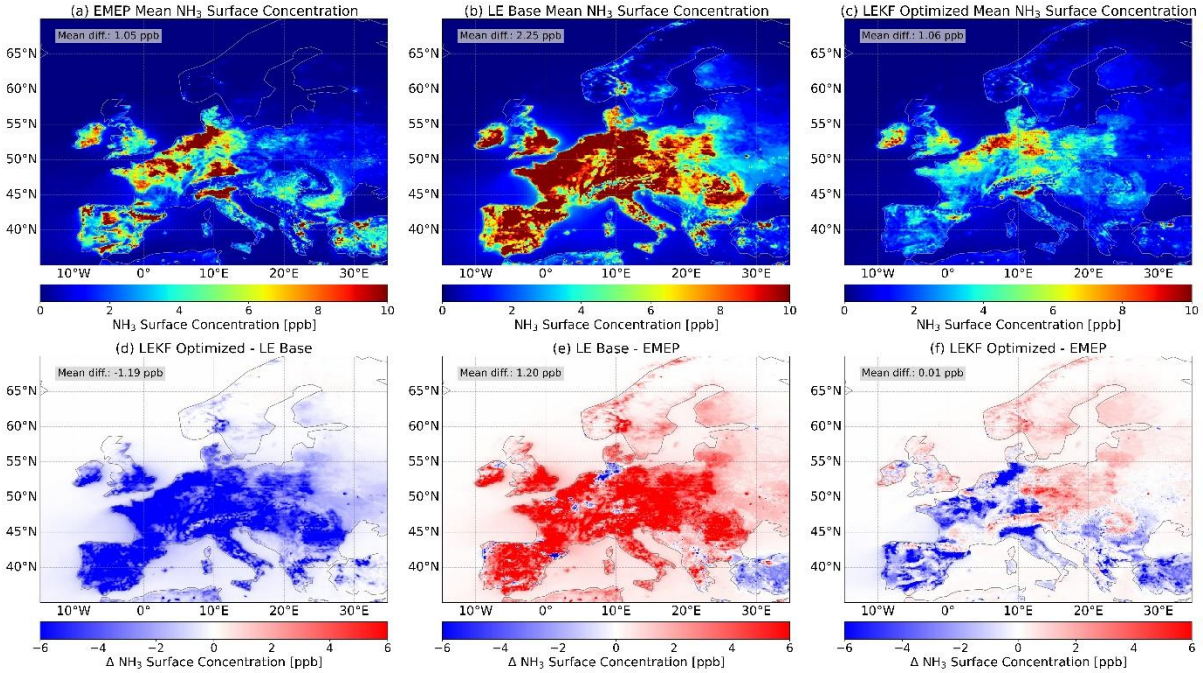

Figure 14 – The mean NH_3 surface concentrations for March 2019 from (a) EMEP, (b) the base LOTOS-EUROS simulation, (c) the LETKF optimized simulation, and the differences for (d) LETKF optimized simulation minus the base simulation, (e) base simulation minus EMEP, and (f) LETKF optimized simulation minus EMEP.

Plots of the monthly mean surface NH_3 concentrations from the EMEP nature run, the base LOTOS-EUROS simulation, and the LETKF-optimized run, along with their respective differences, are presented in Figure 14. For March 2019, the domain mean surface

concentration in the base LOTOS-EUROS simulation (Figure 14b) is 2.25 ppb, more than double the 1.05 ppb observed in the EMEP nature run (Figure 14a). The concentrations and differences shown in Figure 14 directly reflect the emissions and emission differences shown in Figure 13. An overestimation is generally widespread across the domain, although certain regions, where EMEP emissions are higher (see Figure 13e), show the opposite pattern, with base LOTOS-EUROS concentrations falling below those of EMEP (as indicated by negative differences in Figure 14e). This localized underestimation of the surface NH_3 concentrations relative to the EMEP nature run again reflects the strong corrective effect of the assimilation process, which in some regions may slightly overshoot, particularly where observational constraints are robust or where the underlying emissions inventories differ. This is especially likely where the prior model state differs substantially from the observations.

Assimilation of the realistic MTG-IRS observations leads to a pronounced reduction in bias: the mean difference between LOTOS-EUROS and EMEP drops from 1.20 ppb before assimilation to just 0.01 ppb after (pre- and post-assimilation, respectively). The spatial distribution of concentrations in the LETKF-optimized simulation now closely mirrors the EMEP reference, especially in regions with dense observational coverage, highlighting the effectiveness of the assimilation process in constraining surface NH_3 concentrations.

Deposition

Figure 15 compares total NH_3 deposition from the base EMEP nature run, the base LOTOS-EUROS simulation, and the LETKF-optimized simulation, as well as their absolute differences. The base EMEP simulation provides a reference for deposition fluxes, while the base LOTOS-EUROS simulation generally shows higher deposition across the domain. After assimilating the realistic MTG-IRS synthetic observations, the LETKF-optimized simulation shows a substantial reduction in deposition bias with mean differences of +19.35 mg m^{-2} and +5.50 mg m^{-2} pre- and post-assimilation, respectively, and the spatial pattern of deposition more closely resembling that of the EMEP nature run. The differences between the LETKF-optimized and base LOTOS-EUROS simulations (Figure 15d) are predominantly negative, indicating that deposition rates were reduced following assimilation, consistent with the reductions observed in emissions and concentrations. Regions with higher observation density and lower cloud cover exhibit the strongest adjustments, reflecting the greater constraint provided by the assimilated observations. In some areas, particularly those with persistent cloud cover or higher retrieval uncertainty, the adjustments are less pronounced, and the assimilation system relies more heavily on the prior model state.

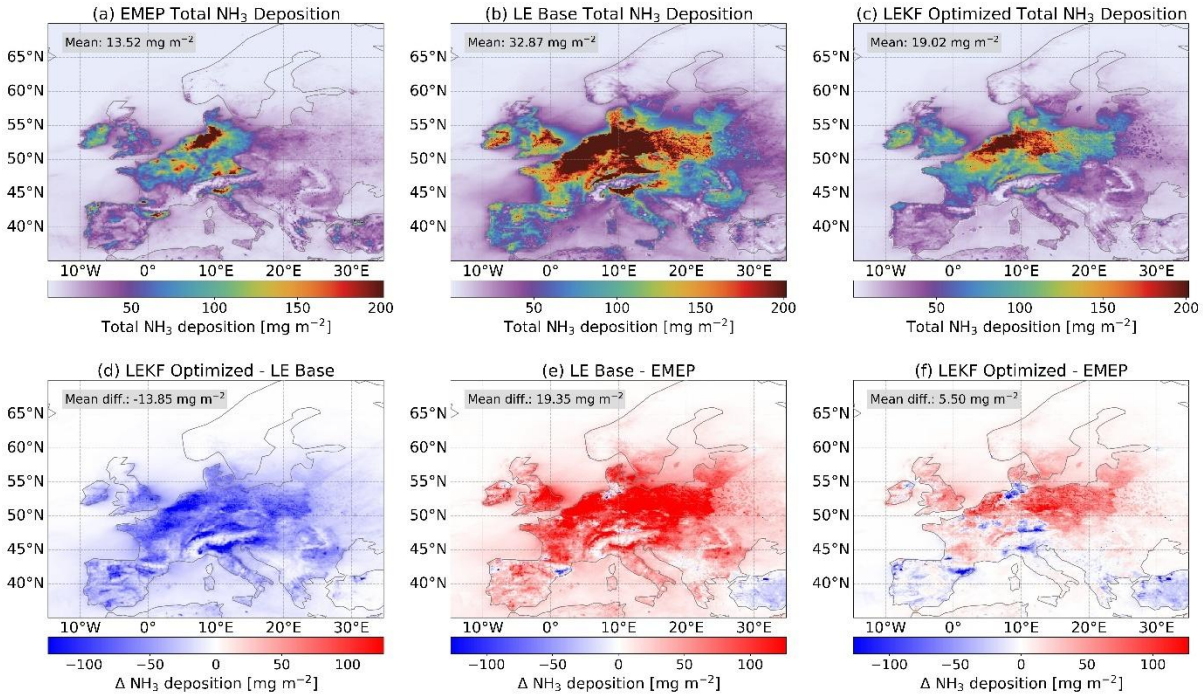

Figure 15 – The total NH_3 deposition (wet + dry) for March 2019 from (a) EMEP, (b) the base LOTOS-EUROS simulation, (c) the LETKF optimized simulation, and the differences for (d) LETKF optimized simulation minus the base simulation, (e) base simulation minus EMEP, and (f) LETKF optimized simulation minus EMEP.

These improvements in deposition result from the assimilation process, which adjusts emissions to achieve optimal agreement with the observed NH_3 total columns, while accounting for the uncertainties inherent in the MTG-IRS retrievals. Nevertheless, some discrepancies in deposition remain, reflecting differences in how LOTOS-EUROS and EMEP represent atmospheric transport and deposition processes. Overall, the assimilation of realistic MTG-IRS NH_3 observations leads to improved agreement in deposition fields between LOTOS-EUROS and the EMEP nature run, particularly in regions where observational constraints are strongest.

3.3.3 Validation of LOTOS-EUROS assimilation against synthetic ground-based observations

To further assess the effectiveness of the assimilation and the improvements in modelled NH_3 concentrations and deposition, we next perform a validation using synthetic ground-based observations generated from the EMEP nature run. This approach allows for an independent evaluation of the LETKF-optimized simulation by comparing its surface concentration and deposition fields against a set of reference observations that are broadly representative of existing ground-based networks, providing additional insight into the performance and reliability of the assimilation framework, and whether the assimilation of the synthetic observations brings the LOTOS-EUROS model closer to the “true” state represented by the EMEP nature run. In this section, we solely focus on the “realistic” observations case.

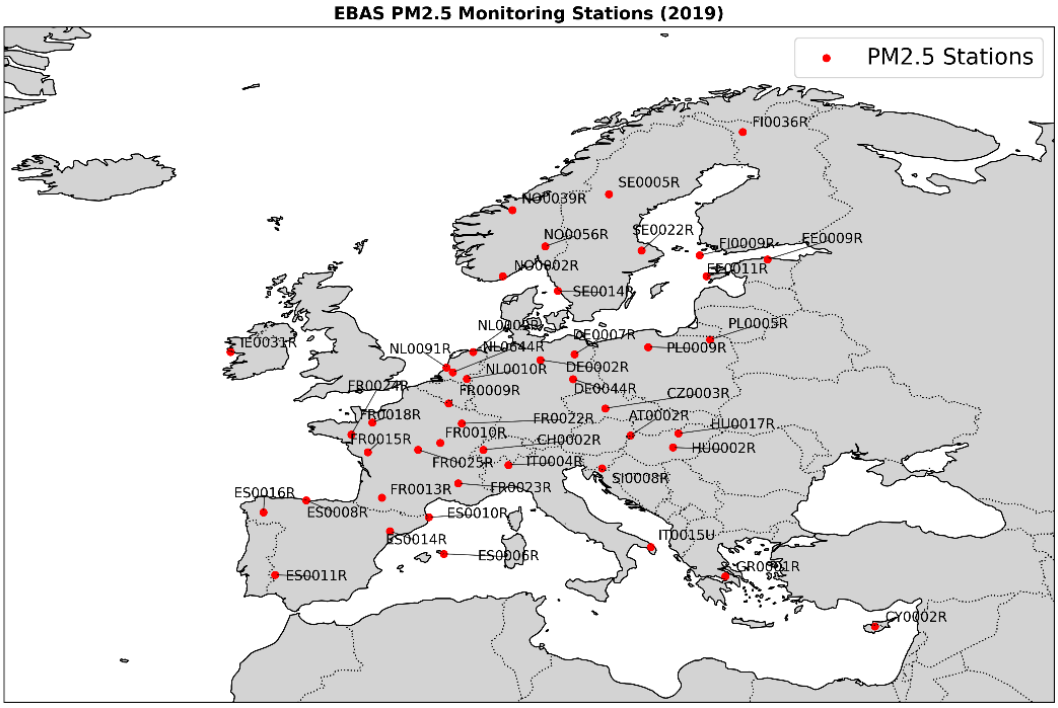

Figure 16 – A map of the 45 EBAS PM2.5 site locations that were selected for the synthetic ground-based comparisons.

To generate the synthetic ground-based observations, the locations of all European PM2.5 surface measurement sites (45 in total) that were active during the year of 2019 within the EBAS network were selected. The selection of these 45 EBAS PM2.5 sites was made to ensure a representative spatial coverage across the domain, encompassing both rural and urban environments as well as areas with varying emission intensities. A map of the 45 sites is shown in Figure 16. The modelled EMEP NH_3 surface concentrations and wet and dry deposition fluxes were then sampled hourly at the locations of these surface measurement sites, and are then treated as representative of the "true" state of the nature run at the surface level.

Comparisons of surface concentrations

First, we compare the modelled surface concentrations pre- and post-assimilation with the synthetic ground-based observations. A plot of the mean surface NH_3 biases calculated per-site for the month of March 2019 is shown in Figure 17. Across the 45 EBAS sites, the base LOTOS-EUROS simulation (left panel of Figure 17) exhibits a consistent positive bias relative to the EMEP-derived "true" values, with the majority of sites showing overestimations in surface NH_3 concentrations prior to assimilation, with one exception being the ES0016R site in north-western Spain that shows a negative bias. The domain-averaged bias before assimilation is substantial (+4.7 ppb across all sites), reflecting the tendency of the base simulation to over-predict surface ammonia, especially in regions with lower emissions or where model parameterisations differ from the EMEP reference.

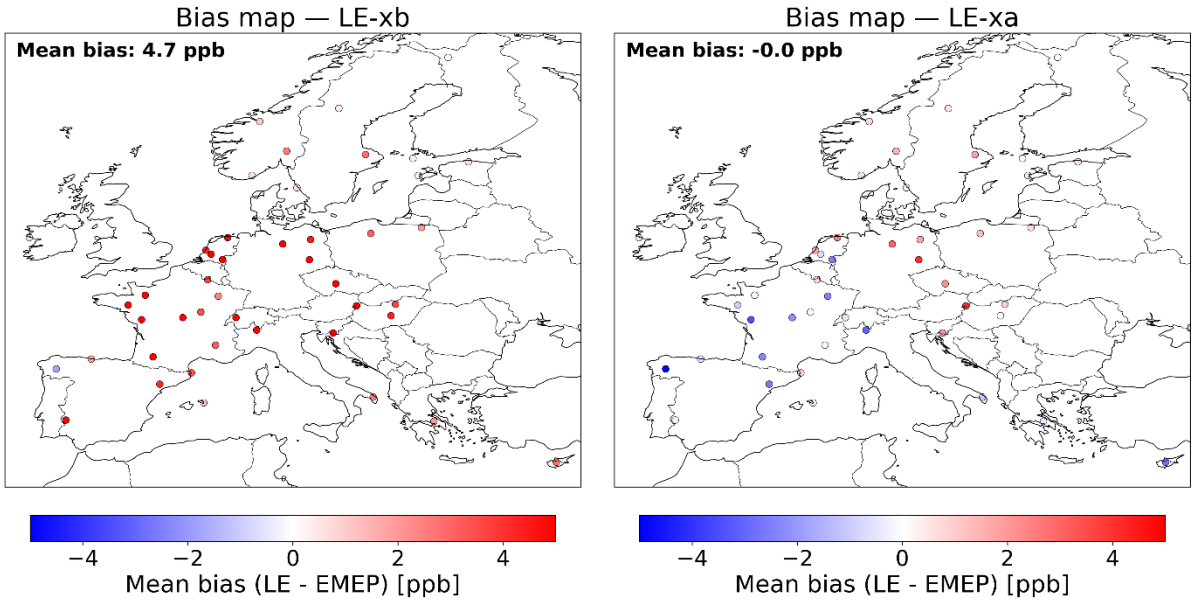

Figure 17 – Maps showing the mean biases (in ppb) in the NH_3 surface concentrations at each of the 45 EBAS ground-based sites for (left) the base LOTOS-EUROS run (xb), and (right) the LETKF optimized simulation (xa). The overall mean bias calculated across all sites is shown in the top left corner of each panel.

Following assimilation of the synthetic MTG-IRS observations (right panel of Figure 17), the LETKF-optimised simulation demonstrates a marked reduction in mean bias at many sites, with the total mean bias over all sites being reduced to 0 ppb. Post-assimilation, the bias map shows that while the overall mean bias is nearly eliminated, a number of sites now exhibit negative biases. Nevertheless, the overall magnitude of these residual biases is much smaller than the pre-assimilation overestimations, and the overall spatial agreement with the EMEP reference is improved.

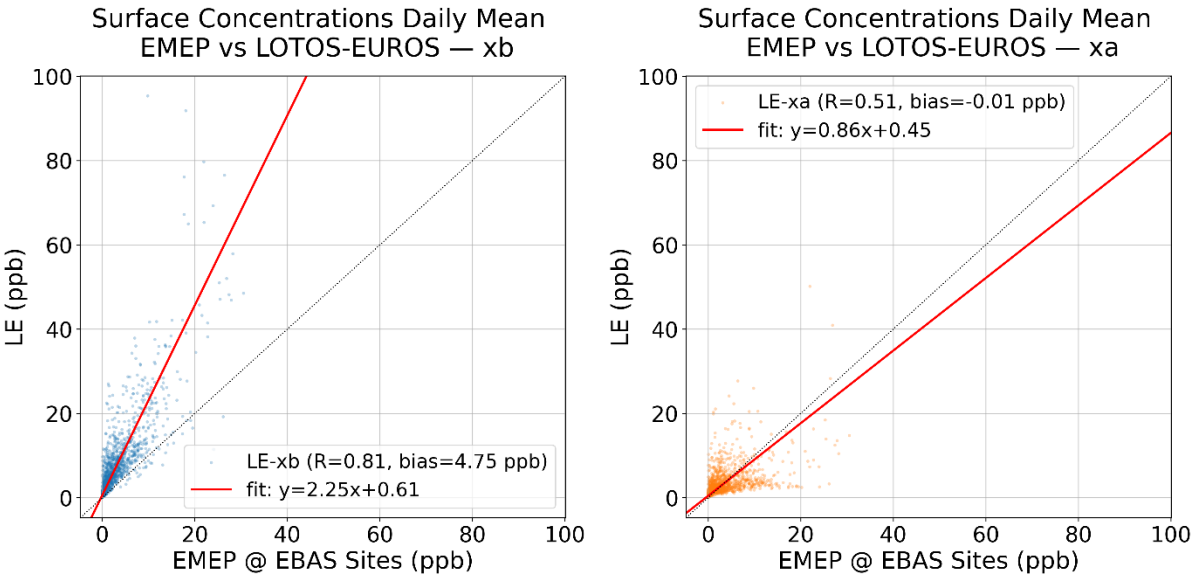

Figure 18 – Scatter plots comparing daily mean NH_3 surface concentrations from the 45 EBAS ground-based sites to (left) the base LOTOS-EUROS simulation and (right) the LETKF optimized simulation.

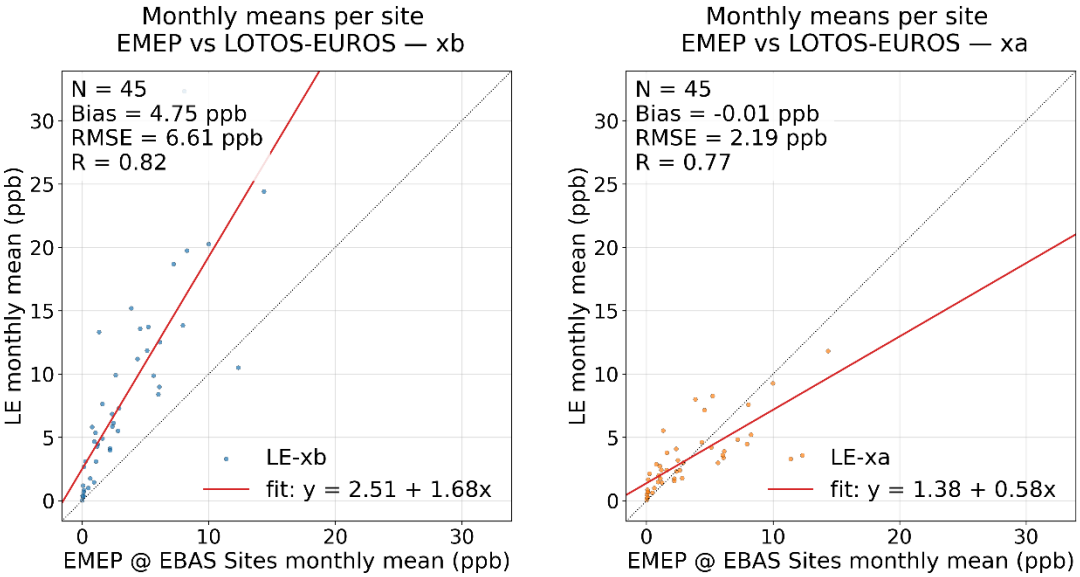

Figure 19 – Scatter plots comparing the monthly mean NH_3 surface concentrations at each of the 45 EBAS ground-based sites to (left) the base LOTOS-EUROS simulation and (right) the LETKF optimized simulation.

Figure 18 compares the daily mean surface NH_3 concentrations at all ground-based sites for the base and LETKF-optimized LOTOS-EUROS simulations. In the base run (left panel), LOTOS-EUROS shows a substantial positive bias relative to the synthetic ground observations sampled from the EMEP nature run, as indicated by the clustering above the 1:1 line and the domain-averaged bias of +4.75 ppb. This does not reflect the model's bias against real-world measurements but rather the structural differences between LOTOS-EUROS and EMEP. After assimilation (right panel), the agreement improves markedly, with the mean bias reduced to -0.01 ppb and the regression slope improving from 2.25 to 0.86. This reduction in bias is accompanied by a notable decrease in correlation (R = 0.81 to 0.51). Figure 19 shows a scatter plot of the monthly means calculated at each individual site which is representative of the temporal correlations before and after assimilation. Again the mean bias can be seen to be reduced from 4.75 ppb to -0.01 ppb after assimilation, which is also accompanied by an improvement in the RMSE from 6.61 ppb to 2.19 ppb. The correlation is decreased slightly from R = 0.82 to R = 0.77. Taken together, the results from Figure 18 and Figure 19 indicate broad improvements post-assimilation, particularly to the biases and the temporal agreement.

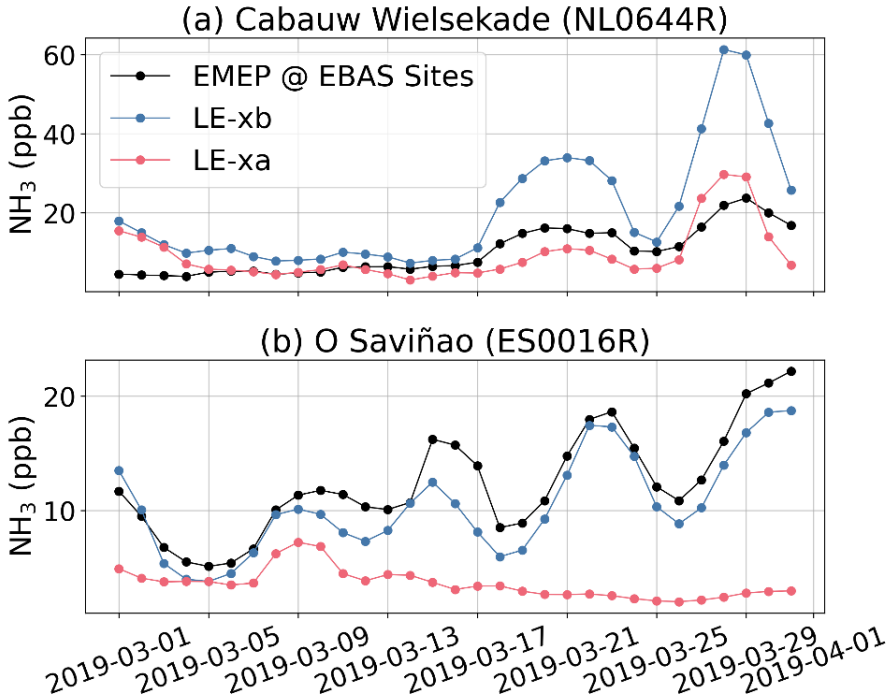

Figure 20 – Time-series of daily-mean surface concentrations at two selected sites (a) Cabauw Wielsekade in the Netherlands, and (b) O Saviñao in Spain.

It should be noted that some sites show a degree of underestimation after assimilation. These effects arise from differences in vertical profiles, surface exchange parameterizations, deposition schemes, and local meteorology between the two models. Because the LETKF adjusts the model state to match observed total columns, mismatches in vertical distribution can lead to disproportionate corrections at the surface. Example time-series for two individual sites, Cabauw Wielsekade in the Netherlands (NL0644R), and O Saviñao in Spain (ES0016R) are shown in Figure 20. Here, Cabauw serves as an example of a site where the assimilation adjusts the surface concentrations well, while O Saviñao is a site where the "overshoot" can be observed and the surface concentrations are underestimated post-assimilation. While this issue is amplified in OSSEs due to model-to-model inconsistencies, it is expected to be less pronounced when assimilating real MTG-IRS observations, which are independent of any single model's structure. Nonetheless, retrieval averaging kernels and a priori assumptions may still interact with model profiles, so future assimilation exercises using the real MTG-IRS data should ensure correct treatment of these factors.

It is also instructive to evaluate how the assimilation of MTG-IRS synthetic observations affects the modelled diurnal variability of NH_3. Figure 21 shows the mean diurnal cycle of surface NH_3 concentrations averaged across all ground-based sites from the EMEP nature run, the base LOTOS-EUROS run, and the LETKF-optimized run. The base LOTOS-EUROS simulation (LE-xb) exhibits a more pronounced diurnal cycle than EMEP, with greater variability and an exaggerated morning peak occurring one hour earlier (07:00 versus 08:00). It also shows a secondary late-afternoon maximum around 19:00 that is absent in the nature run.

After assimilation, the LETKF-optimized simulation (LE-xa) closely matches the magnitude of the EMEP concentrations and substantially reduces the overall bias. However, some discrepancies remain in the diurnal pattern, particularly during the late afternoon and evening. While LE-xa reproduces the morning and midday evolution well, it overestimates the secondary evening increase that is not seen in the EMEP reference, resulting in a mismatch in the late-day temporal profile. This behaviour likely reflects the intrinsic lag in the LETKF

system, which assimilates total column observations but adjusts emissions. The influence of these emission updates on surface concentrations occurs indirectly and with some delay through atmospheric transport processes, limiting the system's ability to capture short-term diurnal dynamics.

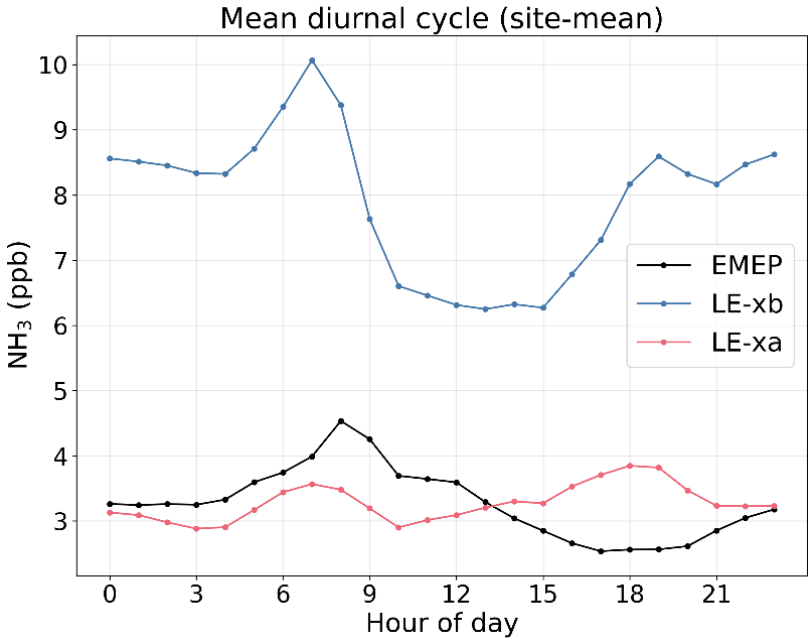

Figure 21 – Mean diurnal NH_3 surface concentration variation calculated over all EBAS synthetic ground-based sites from (black) synthetic ground based observations sampled from the EMEP nature run, (blue) the LOTOS-EUROS base simulation, and (red) the LETKF optimized simulation.

This residual mismatch suggests that, although the LETKF effectively corrects the overall bias and the exaggerated morning peak of the base model, it does not fully resolve the weaker late-day temporal dynamics of NH_3 emissions in the EMEP nature run. This highlights areas for improvement in the LOTOS-EUROS LETKF system, such as implementing a temporal smoothing window or adopting a 4D-LETKF approach (e.g., Bisht et al., 2023), where all observations within a short assimilation window (e.g., 3–6 hours) are considered collectively rather than at a single analysis time. Such enhancements could help the system better capture sub-daily variability and improve consistency with diurnal emission patterns.

Comparisons of wet and dry deposition

To further evaluate the impact of assimilating MTG-IRS synthetic observations, we analyzed the monthly-sum wet NH_3 deposition at the locations of the synthetic ground-based EBAS sites. Figure 22 presents the spatial distribution of the monthly wet deposition bias for the base (LE-xb) and LETKF-optimized (LE-xa) LOTOS-EUROS simulations relative to the EMEP nature run.

In the base simulation (Figure 22a), LOTOS-EUROS exhibits a widespread positive bias across most of central and western Europe, with the largest overestimations occurring over Germany, the Netherlands, and northern France. The mean domain-wide bias amounts to +9.3 mg m^{-2} $month^{-1}$, indicating that the model systematically overpredicts wet NH_3 deposition. After assimilation (Figure 22b), the overall bias is considerably reduced to +2.3 mg m^{-2} $month^{-1}$, with fewer locations showing large positive deviations. The improvement is particularly evident over northwestern Europe, where the strongest corrections were made to

surface concentrations and total columns. Because precipitation originates from relatively high altitudes, wet deposition is largely representative of the total column NH_3 burden, making it a useful indicator of overall atmospheric ammonia levels.

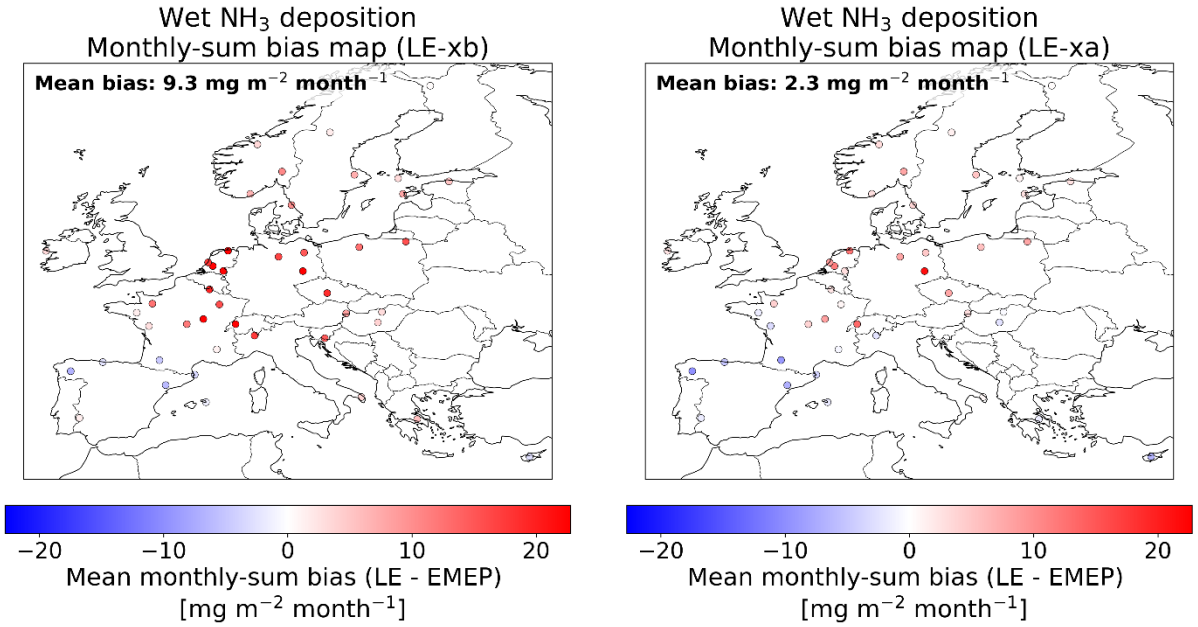

Figure 22 – Maps showing the biases (in mg m^{-2}) in the total monthly NH_3 wet deposition amounts at each of the 45 EBAS ground-based sites for (left) the base LOTOS-EUROS run (xb), and (right) the LETKF optimized simulation (xa). The overall mean bias calculated across all sites is shown in the top left corner of each panel.

The corresponding scatter plots in Figure 23 quantify these improvements. Prior to assimilation, the base run shows a regression slope of 1.86 and a strong correlation (R = 0.93) with the reference data, confirming that the (model captures spatial variability well but with a significant positive offset. Following assimilation, the slope decreases to 1.27, consistent with a substantial reduction in the overall bias to +2.3 mg m^{-2} $month^{-1}$, while the correlation remains unchanged (R = 0.93).

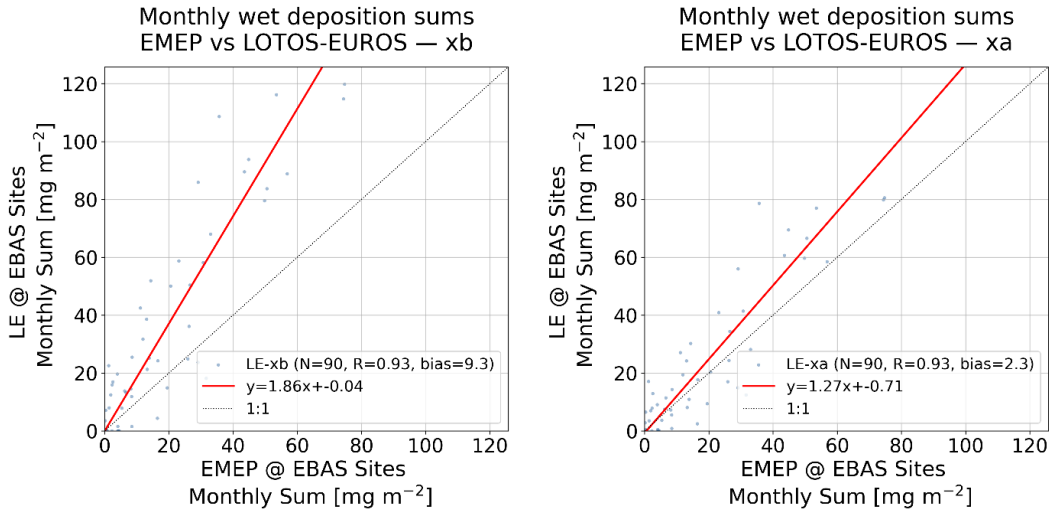

Figure 23 – Scatter plots comparing monthly total NH_3 wet deposition amounts (in mg m^{-2}) from the 45 EBAS ground-based sites for (left) the base LOTOS-EUROS simulation and (right) the LETKF optimized simulation.

These results demonstrate that the LETKF assimilation effectively corrects the excessive NH_3 wet deposition present in the base simulation and brings it closer to the "true" state represented by the EMEP nature run, primarily by lowering surface and column concentrations in regions with the largest prior overestimations. The preserved correlation indicates that the assimilation adjusts the magnitude of deposition without degrading the spatial consistency of the model's deposition pattern, and is also related to the fact that wet deposition is more directly linked to column values in comparison with surface concentrations and dry deposition.

The monthly-sum dry NH_3 deposition simulated by LOTOS-EUROS was also evaluated against the synthetic ground-based observations from the EMEP nature run. Figure 24 shows the spatial distribution of the mean monthly deposition bias for the base (LE-xb) and LETKF-optimised (LE-xa) simulations.

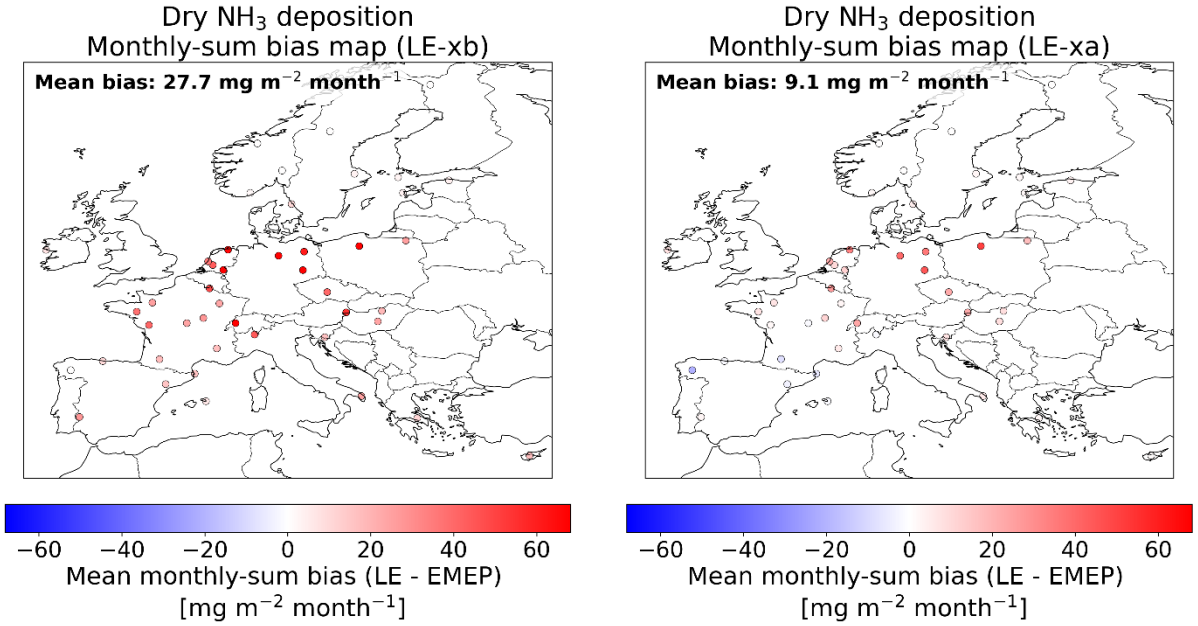

Figure 24 – Maps showing the biases (in mg m^{-2}) in the total monthly NH_3 dry deposition amounts at each of the 45 EBAS ground-based sites for (left) the base LOTOS-EUROS run (xb), and (right) the LETKF optimized simulation (xa). The overall mean bias calculated across all sites is shown in the top left corner of each panel.

In the base simulation (Figure 24a), LOTOS-EUROS substantially overestimates dry deposition across most of central and western Europe, particularly over Germany, Belgium, and the Netherlands. The mean bias amounts to +27.7 mg m^{-2} $month^{-1}$, indicating that the model tends to deposit excessive NH_3 to the surface prior to assimilation. After assimilation (Figure 24b), this bias is reduced to +9.1 mg m^{-2} $month^{-1}$, with fewer sites showing large positive deviations. The strongest improvements occur in regions where surface and column NH_3 concentrations were previously overpredicted, consistent with the spatial corrections introduced by the LETKF analysis.

The scatter plots in Figure 25 provide a quantitative summary of these changes. In the base run, the regression slope of 2.83 confirms a strong overestimation of dry deposition, though the model still captures site-to-site variability reasonably well (R = 0.86). Following assimilation, the slope decreases to 1.65, and the mean bias is reduced by two-thirds to +9.1 mg m^{-2} $month^{-1}$, while the correlation remains comparable (R = 0.82). This demonstrates that the LETKF assimilation effectively scales down the excessive dry deposition fluxes without significantly degrading spatial consistency across sites.

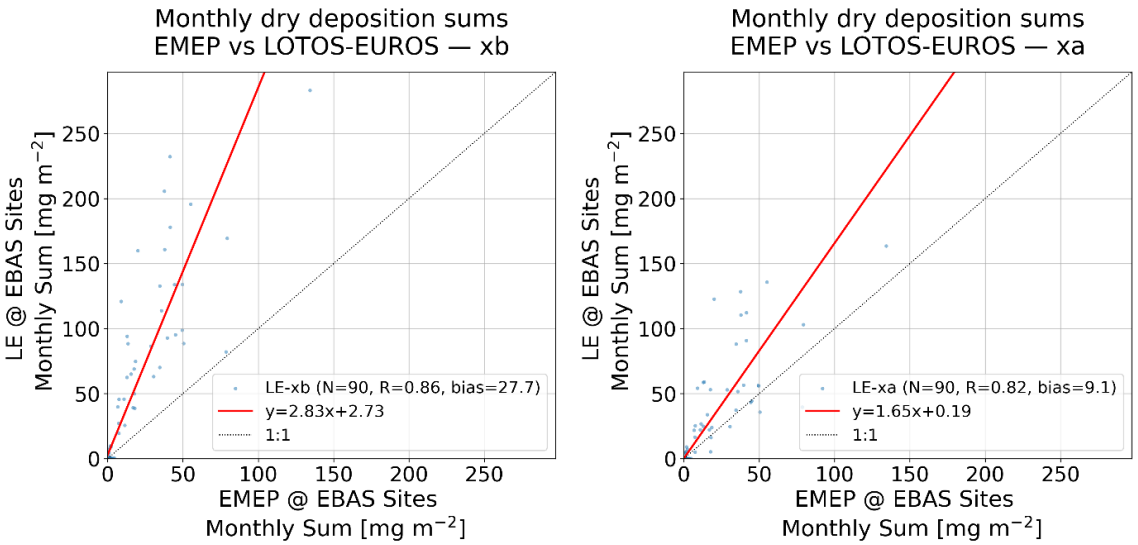

Figure 25 – Scatter plots comparing monthly total NH_3 dry deposition amounts (in mg m^{-2}) from the 45 EBAS ground-based sites for (left) the base LOTOS-EUROS simulation and (right) the LETKF optimized simulation.

Taken together, the wet and dry deposition comparisons demonstrate that assimilating the synthetic MTG-IRS synthetic observations yields a consistent and physically coherent reduction in both components of the total NH_3 deposition budget. Across both pathways, the LETKF analysis reduces large positive biases present in the base LOTOS-EUROS simulation in comparison to the nature run by approximately a factor of three, while maintaining high spatial correlations with the reference data. These results suggest that geostationary NH_3 observations from MTG-IRS have the potential to significantly improve the representation of NH_3 deposition processes in regional models by constraining both atmospheric concentrations and associated surface fluxes.

3.3.4 Assimilation of synthetic observations in EMEP

Experimental setup

The newly developed ELVIS emission inversion system described in section 3.2.2 has been evaluated using synthetic IRS/NH_3 observations.

The set of synthetic observations was created in the same way as described in section 3.2.3 with the 'realistic' settings for pixel availability and uncertainty. For the inversions with ELVIS the synthetic data was based on a nature run with the LOTOS-EUROS model to ensure that the synthetic observations are not already very close to the underlying EMEP simulations.

The usually computed Local Fractions output from the EMEP model describe only the fraction of surface concentrations that originate from local sources. These have been used to perform twin experiments with synthetic NH_3 surface observations to test the new ELVIS system. However, to compute the sensitivity of simulated IRS NH_3 columns it is necessary to have full 3D sensitivities. Therefore, the Local Fraction processing has been adapted to provide output on multiple model levels. In this application, 14 model layers were defined between the surface and a top of 200 hPa, which is sufficient to simulate the columns observed by IRS. The 3D Local Fraction output is available for the full month of March 2019.

Simulated IRS/NH_3 observations

Figure 26 shows an example of the synthetic IRS/NH_3 observations and retrieval errors, and the corresponding ELVIS simulations before and after inversion, valid for 19:00 UTC at March 22. The simulation domain covers north-west Europe, including the strong source areas in Netherlands, north Germany, and Denmark. To the south, also a part of the Po-valley is included. For this particular moment, a large part of the domain is cloud-free and has synthetic observations (left). The retrieval error std.dev. (2nd panel) is roughly 20-30% of the observed value.

For the locations where synthetic observations are available, the standard model simulation (3rd panel) shows high values at the expected locations, but substantial differences with the observations are present. For example, the EMEP simulations show high values over a part of Netherlands only, while in the synthetic observations almost the entire country is covered by a thick layer of ammonia. The model shows however higher values over northern Germany than what is seen in the observations. After inversion, the simulated columns (right panel) is in better agreement with the observations for these, with for example increased columns over the Netherlands and lower columns over north Germany. However, differences are left, for example in the plume over the North Sea just north of Germany.

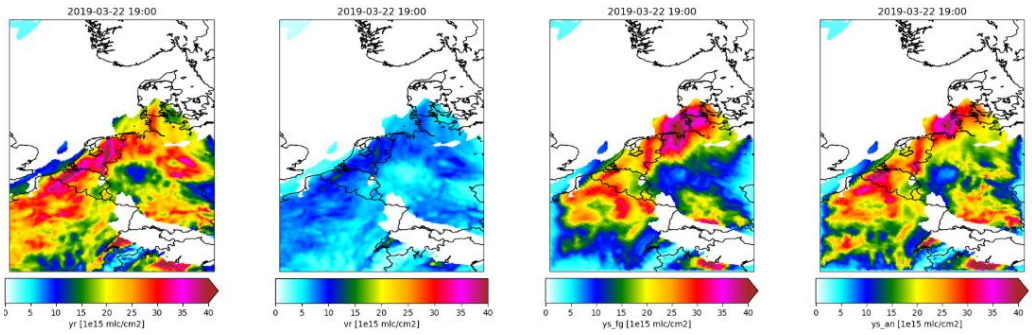

Figure 26 – Snapshot of synthetic IRS/NH_3 columns for March 22, 2019, 19:00: observations (left), observation representation std.dev. (2nd), prior model simulation (3rd), and posterior simulation (right).

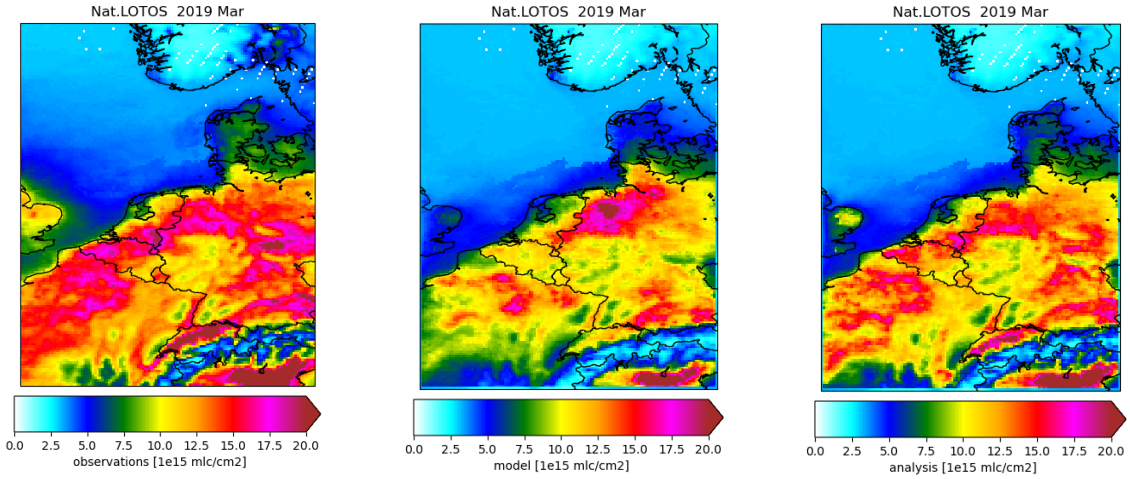

Figure 27 – Monthly averaged IRS/NH_3 columns: synthetic retrievals (left), prior simulations (middle), and posterior simulations after emission inversion (right).

Similar differences and changes are seen at other hours. Figure 27 shows the averaged IRS/NH_3 columns over the entire month, after averaging the (simulated) observations on the model grid. The synthetic observations (left) show high values over south Netherlands, Flanders, northern France, and eastern Germany that the EMEP model simulation (middle) does not fully reproduce. Contrary, the EMEP simulations persistently exceed the synthetic observations in north-west Germany. After inversion, the posterior simulations (right) have at all these locations changed in the direction of the observations. Although overall the simulations are still lower than the observations, the difference is decreased substantially.

To illustrate the changes during the month, Figure 28 shows a time series of the daily domain averaged IRS/NH_3 columns. The daily averages are rather constant during the month, but clear peaks are seen around days 6, 20-24, and 30. The prior simulations (blue) roughly follow the same pattern, but are about 20-30% lower. The posterior simulations (red) are closer to the observations, especially during the peak days. This could be explained from the relative small retrieval error for high NH_3 columns, which have therefore a higher impact in the inversion.

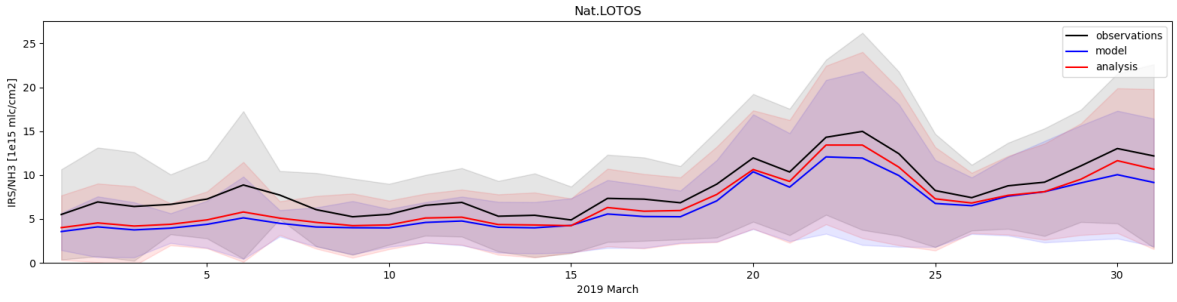

Figure 28 – Time series of mean (lines) and std.dev. range (shaded) over daily and domain averaged IRS/NH_3 columns: observations (black), prior simulations (blue), and posterior simulations (red).

Maps of emissions

The NH_3 emission maps before and after inversion are shown in Figure 29. The *prior* emissions (left panel) show high emissions especially in north-west Germany, but also in the Netherlands, northern France, and Po valley. However, also some large isolated but strong sources are visible, for example at the south-west corner of Norway. The middle panel shows the monthly averaged emission deviation factors estimated by the inversion to decrease the difference between the synthetic IRS observations and prior model simulations. The deviation factors show for example that the emissions over north-west Germany should decrease, while on average an increase is necessary over most other area. The posterior emissions are computed by multiplying the *prior* emissions with one plus the emission deviation factors (right panel). The result shows a pattern that is comparable with the left panel of Figure 13, which shows the prior emissions used by the LOTOS-EUROS model. As these were also used for the nature run from which the synthetic observations were simulated, the ELVIS inversion is indeed bringing the emissions closer to the (synthetic) reality.

The emission deviation factors show a smooth pattern that is related to the chosen length scale of the spatial correlations in the background covariance matrix. This was set to 25 km, with means that spatial structures of about 100 km diameter could be present in the deviations. A smaller length scale might be possible given the fine structures in the emissions; however, this has the disadvantage that emissions are left unchanged if the most nearby observations are further away than 2-3 times the length scale.

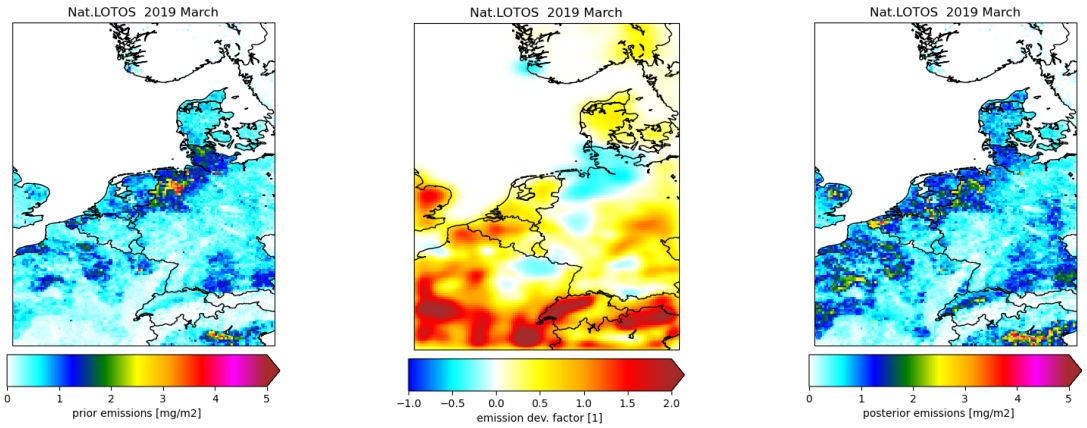

Figure 29 – Monthly averaged prior emissions (left), optimized emission deviation factors (middle), and posterior emissions (right) in ELVIS inversion.

Recommendations

The experiments with synthetic IRS/NH_3 observations showed that the ELVIS system is technically capable of using them for emission estimations. The Local Fractions that are pre-computed by the EMEP model efficiently relate emissions to the satellite observations.

The limited atmospheric lifetime of NH_3 ensures that even using a small source area and limited history, the Local Fractions are able to trace back almost the entire concentration to nearby emissions, at least in area with high emissions. The hourly frequency on which IRS will provide observations is a clear advantage here, as this will provide information on all hours of the day and one does not have to extrapolate a single overpass to the remaining hours of the day.

For the area outside the main source regions, the Local Fraction data shows that a substantial part (up to 50%) of the NH_3 does not originate from nearby sources. Here it is for example chemically produced out of ammonium aerosols transported over longer distances. Total NH_3 columns in these areas are small however, and observations by IRS will therefore probably have a (relative) larger retrieval error. The value of IRS for emission estimation is there expected to be more limited over these 'remote' area, although inversion including longer time windows might still lead to improved estimates over here.

The simulation over areas further away from the sources might also benefit from integration of full EMEP model simulations in the inversion chain. At the moment, updated emission estimates could not be fed back into the model run, which is also not needed to obtain the emission update. However, it would be useful to implement this feature in order to analyse the impact of the updated emissions on for example deposition, and on the effect on other nitrogen compounds that are transported further away from the source area.

3.4 Summary and conclusions

In the first part of this work package, we evaluated the potential of assimilating synthetic MTG-IRS NH_3 observations into the LOTOS-EUROS model using the Local Ensemble Transform Kalman Filter (LETKF) framework, and with the EMEP model using a variational emission inversion approach. Two sets of synthetic observations were generated from either an EMEP or a LOTOS-EUROS nature run: an "idealized" set with full temporal coverage and perfect vertical sensitivity, and a more "realistic" set incorporating vertical sensitivity, variable uncertainties, and realistic temporal sampling.

The two observations sets based on the EMEP nature run were assimilated with the LOTOS-EUROS LETKF system. This led to substantial improvements in the representation of NH_3 emissions, surface and column concentrations, and in deposition fields across Europe. Importantly, the LETKF system proved highly effective at correcting the overall positive biases present in the base LOTOS-EUROS simulation, bringing modelled concentrations and deposition fluxes into much closer agreement with the EMEP nature run. For example, the assimilation of realistic MTG-IRS observations reduced the mean bias in daily averaged surface NH_3 concentrations from +4.75 ppb to -0.01 ppb at ground-based sites. The assimilation also worked exceptionally well for both wet and dry deposition: it greatly reduced the biases in deposition fields while maintaining high spatial correlations with the reference data. This demonstrates that the LETKF approach not only corrects the magnitude of deposition but also preserves the spatial consistency of the model's deposition patterns.

A set of "realistic" synthetic observations was also assimilated in the ELVIS emission inversion framework around the EMEP model. This system uses a variational approach to provide optimized fields of scale factors to adjust the prior emissions. In these experiments, the nature run that provided the synthetic observations was a LOTOS-EUROS simulation. The simulated NH_3 columns strongly improve, with better results for locations where concentrations are mostly due to local emissions. Emissions are adjusted in the direction of the emissions used for the nature run, where an exact match cannot be reached as observation-representation-errors need to be taken into account too. Analysis of the impact of inverted emissions on for example deposition is left for future, as currently the optimized emissions could not be fed back into a model simulation.

The results also highlight important limitations of the current LETKF approach used with LOTOS-EUROS. Because the LETKF system assimilates total column observations but adjusts emissions, mismatches in the a priori and Nature Run vertical profiles can in some cases lead to worse agreement at the surface level. Additionally, there is an inherent lag in the assimilation response. This lag arises from the time required for emission adjustments to

propagate through atmospheric transport and influence surface concentrations. As a result, while the LETKF framework is highly effective at correcting overall biases and aligning modelled columns with observations, it is less able to capture short-term temporal variability, particularly the diurnal cycle of NH_3. The analysis showed that, although the LETKF-optimized simulation closely matches the magnitude of the EMEP nature run during morning and midday hours, it overestimates the late-day increase in surface concentrations, and it has difficulties capturing the flatter diurnal profile from the EMEP nature run. This residual mismatch suggests that the assimilation framework does not fully resolve the temporal dynamics of NH_3 emissions and boundary-layer processes that drive late-day enhancements. For the variational approach used with the EMEP model, this is less of an issue, as for every hour the sensitivity of concentrations for emissions from all previous 24 hours is available. To some extent a similar approach might be used for improvement in the LOTOS-EUROS LETKF system. This could include for example the implementation of a temporal smoothing window or a 4D-LETKF approach such as that described by Bisht et al. (2023) where the assimilation of all observations over a longer time window are considered instead at a single analysis time point.

To conclude, the LOTOS-EUROS LETKF and EMEP ELVIS system have been evaluated and prepared for the assimilation of real MTG-IRS NH_3 data when it ultimately becomes available. Our results demonstrate that geostationary NH_3 observations from MTG-IRS have significant potential to improve the representation of NH_3 emissions, concentrations, and deposition in regional air quality model, and that they are likely to provide significant higher amounts of information to our assimilation systems than existing sensors. The assimilation framework developed here provides a robust foundation for the uptake of future satellite data in CAMS production systems, supporting enhanced monitoring and management of atmospheric ammonia across Europe.

4 OSSE on IRS O_3/CO radiance potential benefits

The results presented in this section have been compiled as a manuscript titled '*Preparation for assimilation of MTG-IRS radiances into an atmospheric composition model for ozone and carbon monoxide forecasts.*' which is currently under review at the Quarterly Journal of the Royal Meteorological Society (Coopmann et al., 2025).

4.1 Introduction

Atmospheric chemical composition forecasting is a relatively recent concept, derived from numerical weather prediction (NWP) in meteorology. The principle is to forecast chemical composition over several days in the troposphere and stratosphere using a CTM. These models use theoretical concepts of physico-chemical transformations in the atmosphere and are forced by meteorological forecasts and fed by emission inventories. These models are capable of providing forecasts of UV index, pollution, air quality and 3D gas and aerosol composition. In a context of climate disruption, it has never been more crucial to provide scientists, policy-makers and the general public with increasingly accurate forecasts. Research in this field has led to ever-greater improvements in these models, increasing the accuracy of processes, inventories and vertical and horizontal resolution. The CTM developed and operational at Météo-France is the Modèle de Chimie Atmosphèrique de Grande Échelle (MOCAGE).

To further improve the forecasts produced by these models, one of the strengths is the use of data assimilation systems, which allow a compromise to be reached between a background (short-term model forecast) and observation (conventional or satellite), to estimate the most probable state of the atmosphere, known as an "analysis," which is used as the initial condition for the forecast. Currently, most forecasting centres only assimilate geophysical products (L2 products) constructed from Level 1 (L1) satellite observations. In this study, we have chosen to take a different approach, as the objective is to prepare for the assimilation of IRS radiances (L1 products) in the regional domain of MOCAGE CTM for ozone and carbon monoxide, and to evaluate the impact on the quality of the subsequent forecasts, as part of a demonstration for the evaluation of the CAMS regional service.

In 2025, a new generation of infrared sounders was launched on board the European MTG series. MTG-IRS will be able to measure the radiance emitted by the Earth at the top of the atmosphere with a high spatial and temporal frequency (respectively 7 km and every 30 minutes over Europe) thus providing an accurate 4D picture of the atmospheric state in terms of temperature, water vapour, winds, clouds, surfaces and chemical composition. IRS will therefore be of particular interest for monitoring and forecasting atmospheric composition specifically in Europe.

4.2 OSSE framework

Preparing for the arrival of future satellite observations is a challenge for meteorological, space, and research centres. This requires cutting-edge techniques and methods to prepare systems for assimilating this new data. For several years now, OSSEs have been providing a strong, tried-and-tested method for the assessing the impact of new observations on models' analyses and forecasts. This strategy is one of the most commonly used for meteorology and atmospheric composition.

The typical OSSE concept is often the same, and can be summed up as follows: simulate the new observing system on the basis of an atmospheric model considered as reality, then assimilate these generated observations, and finally assess the impact of these data on

analyses and forecasts, compared with a control experiment that does not yet assimilate these new observations. These two settings have to be different enough not to be considered identical. The OSSE specific for the present work has included the following steps:

- The simulation of an atmospheric state considered as reality, called Nature Run (NR) and generated from MOCAGE CTM, which is used both to simulate the new observation system (in our case IRS) and as verification data.
- Accurate creation of the observation system, here with a radiative transfer model capable of simulating IRS infrared radiances, and estimation of the associated errors.
- A CTM model (MOCAGE) and data assimilation system (3D-Var) to generate analyses and forecasts.

4.2.1 Atmospheric simulation

One of the first things to be implemented in an OSSE is the creation of the Nature Run (NR), which consists of a long, uninterrupted forecast lasting several months, most often. For this study, we use the OSSE framework constructed for the study of Vittorioso et al. (2025). We decided to focus on the exploitation of IRS for the regional model (named MACC01) over the summer period from May 21, 2019 to August 31, 2019, including a spin-up on the first 10 days. This NR uses the regional domain limited to 28°N, 26°W and 72°N, 46°E, with a horizontal resolution of 0.1°x0.1° centred over Europe and 60 vertical levels down to 0.1hPa. This MACC01 NR must be coupled to the boundary conditions by a global NR over the same period, also derived from the MOCAGE global model (named GLOB11) with a horizontal resolution of 1°x1°. Finally, meteorological forcings are taken from the ARPEGE global NWP model. This NR MACC01 is used as a real atmospheric field for the creation of synthetic IRS observations, and as comparison data for assimilation experiments.

For a valid OSSE, it is important that the coupled NRs (GLOB11-MACC01), used to simulate the observations, are not the same as those used as Control Run (CR) or Assimilation Run (AR), which would lead to an overestimated benefit from the observations (i.e. the identical twin problem). To avoid this difficulty, other coupled simulations (GLOB11-MACC01) have been set up over the same period but with different conditions such as: biogenic emissions, anthropogenic emissions, biomass fires for example. Note that meteorological forcings are the same in both NR and CR. These different coupled simulations, known as (CR), have different ozone and carbon monoxide behaviours to those of the NRs, thereby limiting the identical twin problem. The MACC01 (CR) is therefore used as a basis for the (AR) which makes use of the synthetic IRS observations in the 3D-Var data assimilation system. Figure 30 shows a scheme of the OSSE framework used in this study and technical details are summarized in Table 2. As a reminder, this study focuses solely on the regional part of the OSSE.

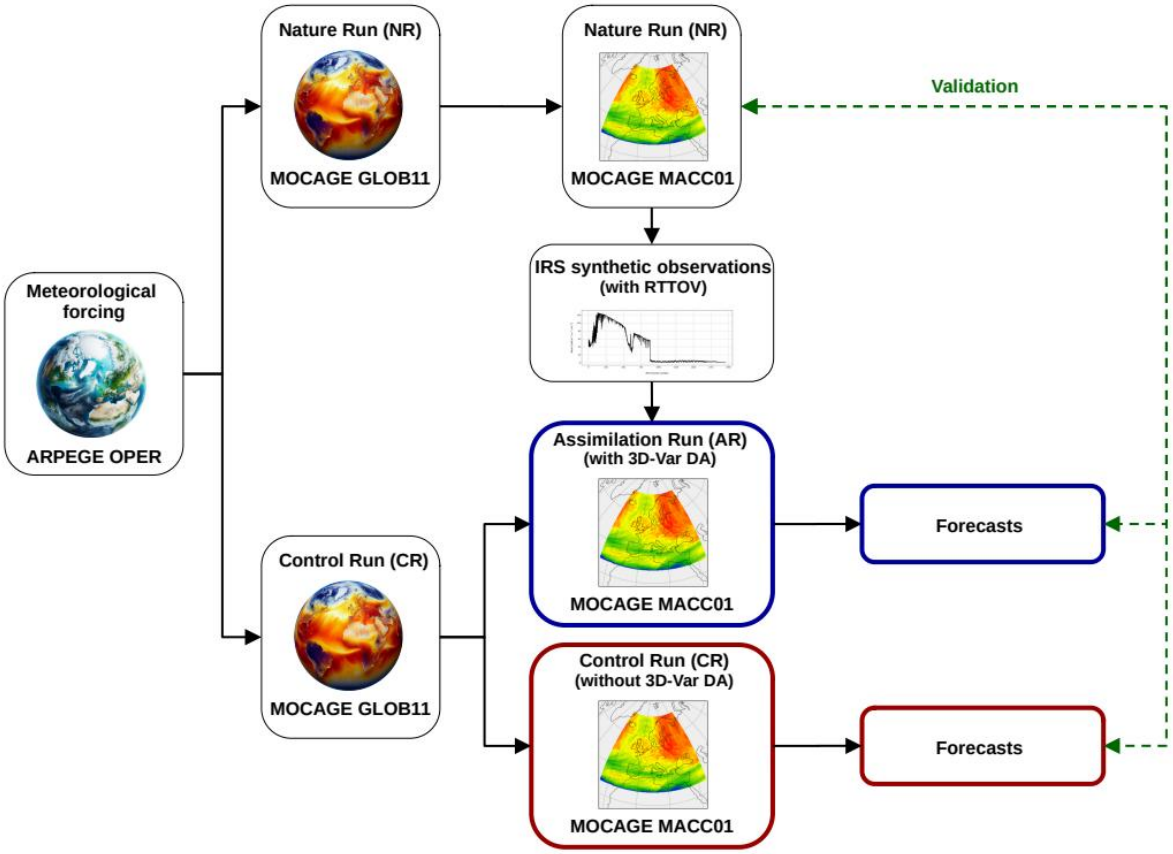

Figure 30 – The OSSE scheme used for task 3.3.2.

Table 2 – Summary of the configurations of the OSSE.

OSSE	Nature Run (NR)		Control Run (CR)		Assim Run (AR)	
Domain	**Global**	**Regional**	**Global**	**Regional**	**Global**	**Regional**
Period	2019/05/22 to 2019/05/31 (spin-up) 2019/06/01 to 2019/08/31 (data assimilation experiment)					
Meteorological forcing	ARPEGE (operational)					
Biogenic emissions	MEGAN-MACC (2010) & GEIA for NOx (1990)		CAMS-GLOB-BIO (2000) & CAMS-GLOB-SOIL for NOx (2000)			
Anthropogenic emissions	MACCity (2016)	CAMS-REG-AP (2017)	CAMS-GLOB-ANT (2000)	CAMS-REG-ANT (2000)	CAMS-GLOB-ANT (2000)	CAMS-REG-ANT (2000)
Biomass burning	GFAS (2019)		MACCity (2000)			
Data assimilation	IASI L1c (15 O3 channels)	None	None	None	None	IRS L1c (76 O3 channels) (35 CO channels)

4.2.2 IRS synthetic observations

For an OSSE, precise creation and calibration of the new observation system to be evaluated is crucial. At this stage, we do not need to calibrate any other observation systems, as no other observing system had been assimilated in the regional domain at the time of this operational study. The aim is to obtain synthetic IRS observations as close to reality as possible. As a first step, EUMETSAT provided us with information on the geometry of the IRS instrument (angles, longitudes, latitudes, etc.) and its instrumental noise. More about the technical characteristics of IRS can be found in the paper Coopmann et al. (2022). Then, to simulate perfect observations, we used the RTTOV v12 radiative transfer model. The IRS coefficient file was generated by CEMS (Centre d'Etudes en Météorologie Satellitaire) with pseudo-Hamming apodization and spectral sampling of 0.6 cm^{-1} for 1960 channels between (680-1210 cm^{-1}) and (1600-2250 cm^{-1}).

IRS will provide observations with high spatial and temporal resolution, however, our systems are not capable of simulating and assimilating such a dense amount of data. Thus, the atmospheric profiles from NR-MACC01 (including ozone and carbon monoxide) given to RTTOV to simulate perfect observations have been filtered in the following ways:

- We choose to assimilate only in clear sky conditions, so for a given pixel, only the atmospheric profiles considered to be fully cloud-free are used to simulate the observations. The cloud information comes from the ARPEGE-OPER meteorological forcing fields. Thus, the cloud fields taken into account are the same for all RUNs (NR, CR, AR).
- To limit errors in the radiative transfer simulation, we simulated synthetic IRS observations with a zenith angle of less than 63°. To avoid the influence of the coupler model, we simulated IRS pixels on a domain slightly smaller than the MOCAGE regional domain: 30.5°N, 23.5°W and 69.5°N, 43.5°E.
- Taking into account horizontal correlations between observation points is still complex. So, to further reduce the volume of observations, we used thinning to simulate observations for 1 pixel per 0.4° box.

Finally, unlike the previous study, we have chosen to simulate all channels (1960) and to construct synthetic observations using inter-channel correlated errors. In the initial work by Vittorioso et al. (2024), the IRS observations were derived by adding to the simulations only the uncorrelated instrumental noise multiplied by a random Gaussian function with mean zero and standard deviation equal to 1. However, a study has shown that including correlated observation errors in the creation of synthetic observations increases both the realism of these observations and the way in which they are ingested by the data assimilation system. To achieve this, we replaced simple uncorrelated instrumental noise with a full error correlation matrix in the process of generating synthetic observations. This matrix was derived from a previous study on IRS assimilation in the AROME regional model at Météo-France (Coopmann et al., 2023).

Synthetic observations were generated hourly for each pixel (as selected above) for the period May 21 to August 31, 2019 for the 1960 IRS channels. Figure 31a shows that the IRS spectrum averaged over 3 months is similar to the look of spectra from other infrared observations currently in orbit, such as IASI. Figure 31b shows a map of the number of pixels used over the study period. Unsurprisingly, the largest amount of observations is obtained over the eastern Mediterranean, as this region is the least cloudy at this time of year, compared with the Atlantic and northern Europe.

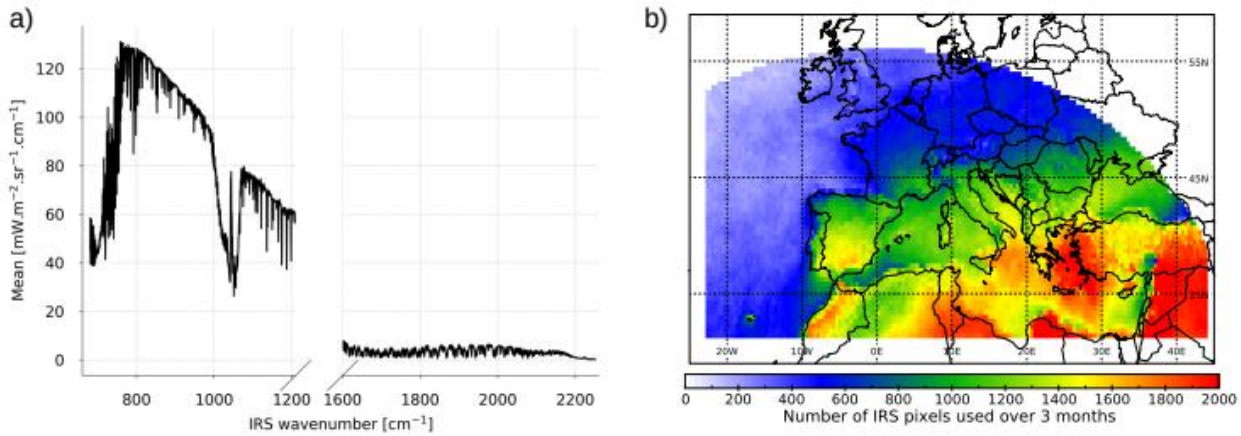

Figure 31 – (a) Mean of synthetic IRS observations (radiance spectrum) according to 1960 channels over the 3 months of study (JJA) and (b) map of the number of IRS pixels used over the same period.

4.3 Experiment design

4.3.1 IRS sensitivity analysis and channel selection

The aim of this work is to assess the benefits of assimilating IRS radiances to improve ozone and carbon monoxide forecasts in the MOCAGE regional model. One of the first things to be done is to analyse the sensitivity of IRS to these gaseous compounds. A commonly used method, which implies evaluating the brightness temperature response to a perturbation (ΔBT) for each atmospheric constituent separately (in our case O_3 and CO). For this, we again use the RTTOV radiative transfer model. We use the database of atmospheric profiles including O_3 and CO available with RTTOV and representative of the variability of these constituents. The analysis consisted in simulating IRS observations from these raw profiles, then running new simulations, this time using ozone and carbon monoxide profiles perturbed by 10 % and 1 % respectively. The different statistics between the two simulation sets are then computed. The results are shown in Figure 32a for O_3 and Figure 32b for CO. One can observe a very high sensitivity between (980-1100 cm^{-1}) representative of the ozone absorption band and also sensitivity of some IRS channels between (660-760 cm^{-1}) in the CO_2 absorption band and also between (2050-2100 cm^{-1}) identified as the CO absorption band. Finally, the IRS channels show a sensitivity to CO on a single absorption band between (2050-2250 cm^{-1}).

One of the principles of data assimilation for infrared radiances is the compromise between useful information and data volume. Indeed, the sheer volume of these observations, with their number of channels and ever finer spatial and temporal representations, represents a challenge for data assimilation systems, which are limited by their computational cost. The main methods used to reduce this amount of data are spatial thinning (as we have already done) and information selection. To select the most informative ozone and carbon monoxide channels, we use a method known as physical selection, based on the results obtained previously with sensitivity analysis. Channel by channel, we evaluate which are most sensitive to the variable we're trying to assimilate, and which are capable of providing the best information content. In the case of ozone, we focus on the ozone absorption band, since the other sensitivities identified also apply to other compounds. The aim here is to obtain pure information. However, we are well aware that with infrared radiances, this kind of pure information does not exist, since in the ozone absorption band, channels are also sensitive to atmospheric temperature and humidity, as well as to skin temperature. This method allowed to select 76 ozone-sensitive channels. As previously mentioned, ozone-sensitive channels are also present in the carbon monoxide absorption band. Because of this overlapping, these

channels are excluded from the selection, leaving a total of 35 CO-sensitive channels selected. Figure 32c illustrates these selected channels on a typical IRS brightness temperature spectrum.

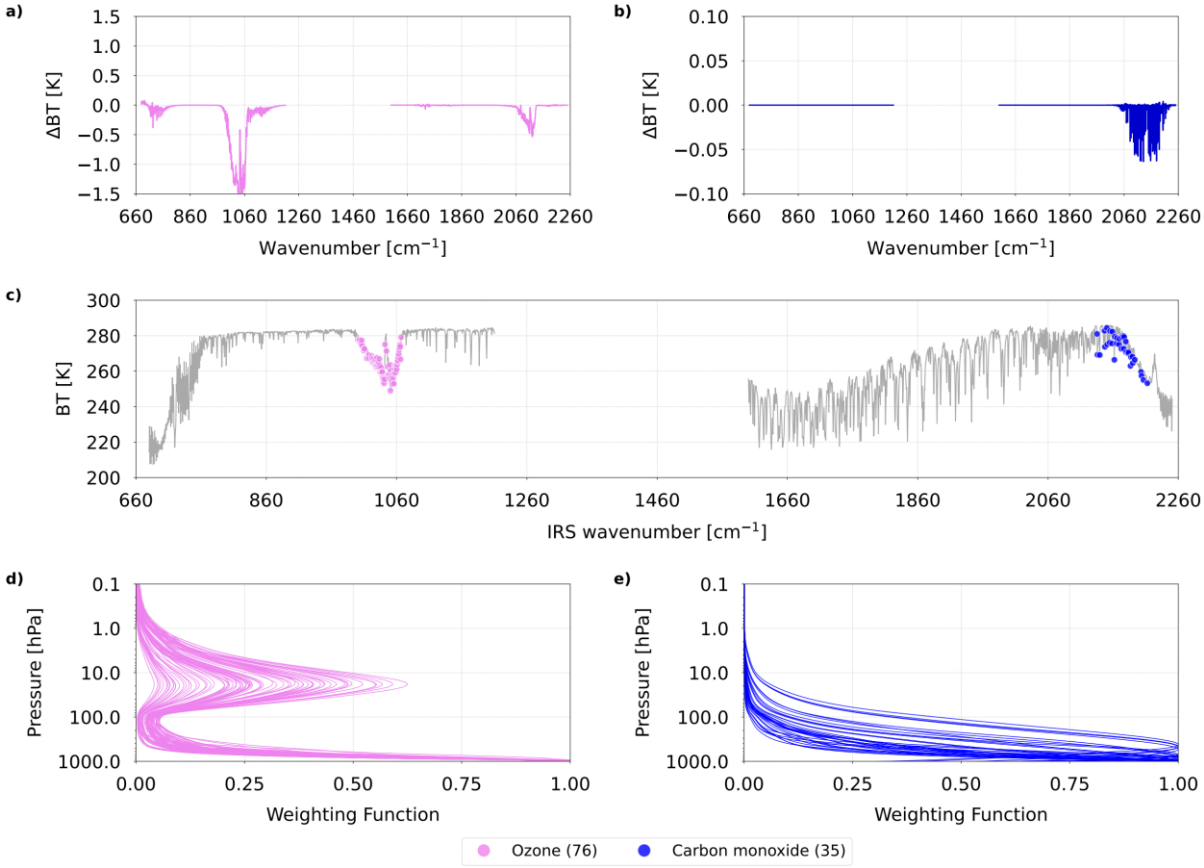

Figure 32 – Differences in IRS Brightness Temperature (BT) between raw and perturbed simulations for (a) O_3 and (b) CO; illustration of selected channels on (c) a typical spectrum and weighting function associated with these selected channels for (d) O_3 and (e) CO.

Finally, it's important to identify the atmospheric level to which the selected channels are sensitive. Again, RTTOV is also able to compute weighing functions for each IRS channel. For a given radiance, the weighing function represents the derivative of transmittance as a function of altitude, and thus provides an estimate of the atmospheric levels best detected from space by a given channel. Figure 32d and Figure 32e represent the normalized weighting functions of the selected channels for ozone and carbon monoxide respectively. It can be seen that the 76 channels selected for ozone are mainly sensitive to the mid- stratospheric region and to low troposphere. The 35 channels selected for CO, on the other hand, present a specific sensitivity to the tropospheric levels. Consequently, the assimilation of these set of observations can potentially provide information only on the just highlighted atmospheric regions.

4.3.2 Experiments

In order to assess the impact of using IRS observations on ozone and carbon monoxide forecasts, we performed two 3D-Var data assimilation experiments for the regional model MOCAGE as part of our OSSE. The experiments are started on May 22, while the evaluation is carried out over a 3-month period from June 01 to August 31, 2019 (allowing for a 10-day spin-up period). For each run, the MACC01 CR is coupled to its boundary conditions to the GLOB11 CR. The first experiment consists of a forecast run without any data assimilation (**EXP Control)**, while the second experiment assimilates hourly IRS synthetic observations,

including 76 ozone-sensitive and 35 carbon monoxide-sensitive channels (**EXP Assim**). For the latter, the experiment uses the previously presented observation and background errors. We note that EXP Assim assimilates IRS observations with a thinning of 0.4°, over land and sea and only in clear-sky conditions. Table 3 summarizes the configuration of these two experiments.

Table 3 – Summary of the configurations of the two experiments in MOCAGE without IRS and with IRS assimilation.

Experiments	EXP Control	EXP Assim
Period	2019-05-22 to 2019-05-31 (spin-up) 2019-06-01 to 2019-08-31 (evaluated)	
Obs. assimilated	None	111 IRS channels
Control variable	None	O_3 and CO
R error (IRS channels)	None	Full and diagnosed R matrix
B error (O_3 & CO)	None	Estimated σ^b by NMC
Verification	Nature Run	

4.4 Results

4.4.1 Forecast impacts

In addition to pave the way for the assimilation of IRS data for atmospheric composition purposes, the objective of this study is to assess the impact that these observations would have on operational forecasts and, more specifically, on ozone and carbon monoxide fields. To examine this, the OSSE framework provides a true atmospheric state (Nature Run) that allows the evaluation of forecasts through the two experiments EXP Control and EXP Assim. A comparison of these two experiments is here presented through the plots in Figure 33. This represent the forecast error in RMSE of EXP Control (resp EXP Assim) compared to Nature Run, and the relative difference between both, integrated over the atmospheric column and averaged for the forecast ranges (in steps of 3 hours up to 24 hours) for the 3 months of study. Ozone forecasts are in the top panels and carbon monoxide in the bottom ones.

In Figure 4a, we can see that the RMSE values of ozone forecast error over the regional domain for EXP Control range between 10^{-7} and 10^{-6} ppv. Assimilation of IRS observations in EXP Assim significantly reduces these values, as shown in Figure 33b, this time with RMSE values between 10^{-8} and 10^{-7} ppv. Although most of observations are located in the Mediterranean region (see Figure 33b), the improvement in forecasts is found to be spread across the full domain, except for the lateral limits, where forecasts are influenced by the MOCAGE GLOB11 Coupling Run. Figure 33c shows the relative difference between the two experiments for ozone forecast errors. IRS assimilation reduces the total forecast error (1%) over a large part of the domain of interest, even though the observation density used is located in the southern part of the domain. It should also be noted that the influence of the coupling domain has no negative impact in terms of error.

Then, Figure 33d**,** shows that for EXP Control, the RMSE values of the carbon monoxide forecast error are between 10^{-8} and 10^{-7} ppv. While in EXP Assim (Figure 33e), the values are rather between 10^{-9} and 10^{-8} ppv. Figure 33f shows the relative difference between the two experiments for carbon monoxide forecast errors. As with ozone, IRS assimilation reduces the total forecast error (1%) over a large part of the domain, with positive impacts extending

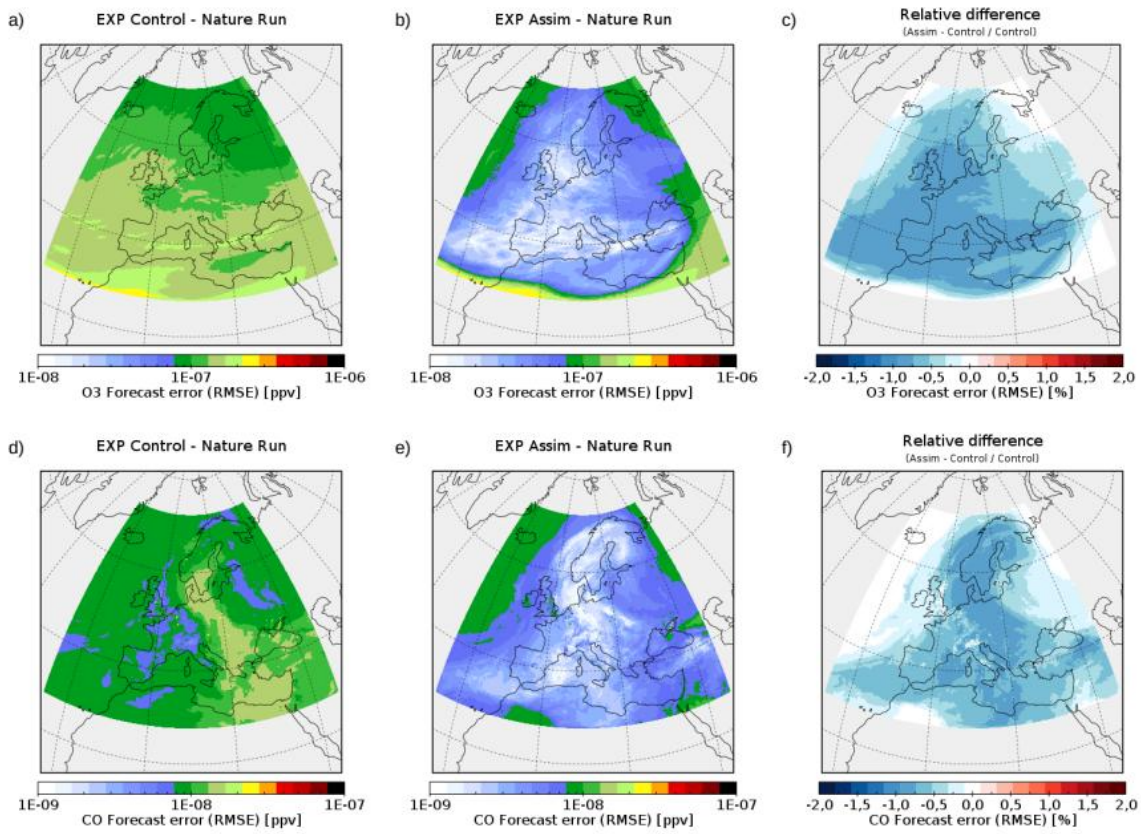

Figure 33 – Forecast error (RMSE) of (left column) the EXP Control, (middle column) EXP Assim compared to Nature Run and (right column) the relative differences between the two experiments, averaged over the atmospheric vertical and forecast ranges (in 3-hour steps up to 24 hours) over the 3-month study period (JJA) for (top row) ozone and (bottom row) carbon monoxide.

northward. There is a slightly more pronounced influence of the coupling domain on carbon monoxide forecast errors, with a neutral impact. These global statistics (averaged vertically and over ranges) illustrate the general positive influence of IRS assimilation on improving ozone and carbon monoxide forecasts.

Another comparison was performed on the atmospheric column by computing a zonal average RMSE of forecast error on the vertical from 1000 to 0.5 hPa. Figure 34a reports the values of RMSE for ozone forecast error of EXP Control. The most important errors are concentrated between 100 and 10 hPa. Non-negligible errors can be noticed in the troposphere. Even if tropospheric ozone represents only 10% of the total atmospheric ozone, the latter is toxic for humans and vegetation, that is why it is important to improve its forecast. Fortunately, as already illustrated in Figure 32d, the selected IRS channels present a double sensitivity to ozone, both between 100 and 10 hPa and in the troposphere. Figure 34b shows significant reductions in the forecast error RMSE of EXP Assim, mainly on these atmospheric layers. The most significant reductions are then located at the bottom of the domain, which is consistent with the higher concentration of observations in this area. However, we note that in EXP Assim the forecast error RMSE increases slightly above 10 hPa compared to EXP Control. This can be seen in the relative differences in Figure 34c, with a reduction in errors (1%) at the surface to 100 hPa thanks to IRS assimilation and a degradation of errors above 100 hPa (2%). Even though the degradation in the upper stratosphere seems greater than the improvement, these results must be put into perspective given the very low error values at these altitudes. This is why Figure 33c shows an overall reduction in errors.

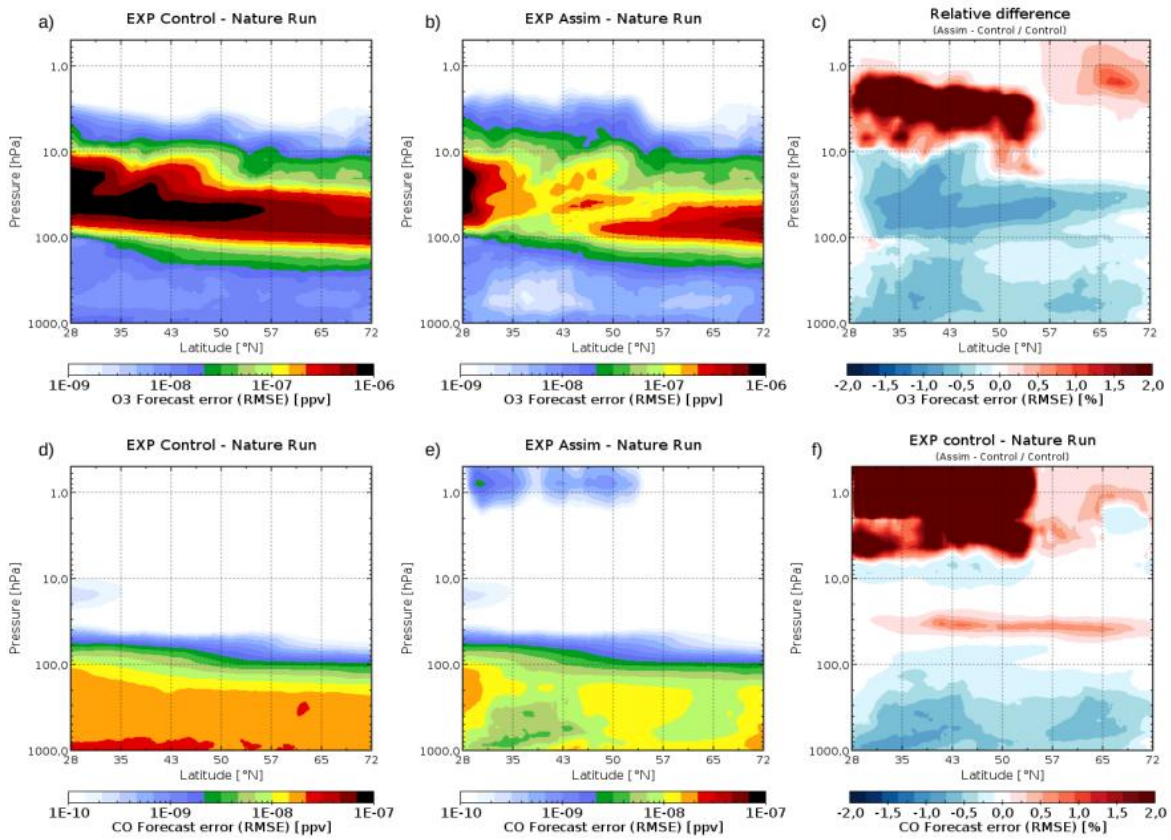

Figure 34 – Zonal mean forecast errors (RMSE) of (left) the EXP Control (left), EXP Assim (middle) compared to Nature Run and the relative differences between the two experiments (right), averaged over forecast periods (in 3-hour increments up to 24 hours) over the 3-month study period (JJA) for ozone (top) and carbon monoxide (bottom).

Finally, the highest RMSE values of forecast error for carbon monoxide are mainly found in the troposphere between 1000 and 100 hPa for EXP Control (Figure 34d). As already mentioned when presenting Figure 32e, this is the atmospheric region to which the 35 CO-sensitive IRS selected channels are most likely to provide information. Consistently, this is what we observe in Figure 34f with a reduction in the RMSE values of forecast error between 1000 and 100 hPa. As for ozone, we note that the strongest reductions are located in the south of the domain. No significant change is observed over the column between 100 and 10 hPa. A slight increase in the RMSE values of forecast error is, instead, recognized in the upper stratosphere above 10 hPa. We also find a limitation of the error reduction at the southern and northern lateral borders of the domain for EXP Assim for the same reason previously mentioned. The relative difference in Figure 34c shows a general reduction in errors between the surface and 50 hPa, followed by alternating improvement, deterioration, and neutral impact between 50 and 5 hPa, and deterioration in the upper stratosphere above 5 hPa. As with ozone, these degradations at these altitudes should be viewed in relation to the very low forecast errors. Thus, this vertical evaluation shows that the assimilation of our 111 channel selection allows to significantly improve the forecasts of the ozone and carbon monoxide fields in the atmospheric layers of interest. It can be noticed on both Figure 33 and Figure 34 that the behaviour of EXP Assim and EXP Control are very similar at the southern border of the domain, which is due to the influence of the information provided by the global domain (i.e. the coupling information) at the borders of the regional domain.

4.5 Summary and conclusions

Two experiments were performed: a first one without any observation assimilation (EXP Control) and a second one with assimilation of IRS observations (EXP Assim) over a 3-month summer period (JJA) and a 10-day spin-up. For EXP Assim, 111 IRS channels are assimilated, the control variables are O_3 and CO, we use the diagnostic R matrix and the vertical profiles of O_3 and CO background error. The statistics on the first-guess (OmB) and analysis (OmA) departures show a significant reduction in bias for the 76 ozone-sensitive channels and a very slight reduction for the 35 CO-sensitive channels. We also note a very small decrease in the standard deviations of OmA compared to OmB for the 111 channels. The analysis of the temporal evolution of the O_3 580 channel [1028.90 cm^{-1}] also shows a decrease in the bias and standard deviations of OmA compared to OmB stable over the period. The same behaviour is observed for the CO 1795 channel [2150.93 cm^{-1}] in a less clear manner. Overall, these results demonstrate an efficient use of IRS observations in the 3D-Var data assimilation system with statistics of OmA values closer to the observations than OmB.

We then compared the ozone and carbon monoxide forecasts of the EXP Control and EXP Assim experiments with respect to the Nature Run. The RMSE statistics of the differences (EXP Control - Nature run) and (EXP Assim - Nature Run) were calculated for O_3 and CO for the MACC01 domain integrated on the vertical over the 3 months of study. The results show a significant decrease in the forecast error in EXP Assim compared to EXP Control over the entire domain for O_3 and CO up to more than one order of magnitude, except on the lateral limits due to the influence of the Coupling Run. These statistics were also calculated as a zonal average in order to evaluate the impact of the IRS assimilation on the atmospheric vertical. The results show that EXP Assim allows to reduce the ozone forecast errors mainly in the atmospheric layers to which the 76-channel O_3 present the strongest sensitivity, namely in the troposphere and the lower stratosphere. Similarly, in EXP Assim, CO forecast errors are reduced in the troposphere. These reductions are more pronounced in the southern part of the domain, where the amount of observation is the largest (more frequent clear sky pixels compared to the rest of the domain). These 3-month results demonstrate the positive impact of IRS assimilation in improving ozone and carbon monoxide forecasts over a large part of the regional (MACC01) domain.

In this study, we chose to simulate only completely clear pixels, so as not to overestimate the impact. In operational use, we will obviously be dependent on the quality of the cloud flag that will be provided with the real IRS data. Furthermore, horizontal thinning may differ in our first operational implementation depending on the limitations on data flow and data storage that we may encounter. Finally, as with any OSSE, the impacts differ from those that will be obtained with real observations. Thus, the results will probably vary with real IRS data. However, the results are very encouraging and demonstrate the significant benefit that IRS observations will generate for the understanding and improvement of our atmospheric composition forecasts. Despite the theoretical and idealized context of the OSSE framework, this can still be considered a robust and proven method in the preparation and evaluation of the influence of a future observation system on forecasts.

The imminent arrival of the new IRS observations represents a great opportunity. Indeed, all the studies carried out on the impact of IRS on numerical weather prediction models and chemistry-transport models demonstrate a major potential for evolution and improvement in our understanding of the atmosphere, climate systems and our associated forecasts. Added to this is the launch of the new IASI-NG sounder in the same year, which will represent a challenge in terms of exploring observational synergies. The number of infrared sounders in polar (IASI, CrIS, HIRAS, IASI-NG) and geostationary orbits (GIIRS, IRS), and their ever-increasing sensitivity to atmospheric composition, opens up a new way of representing

chemical and aerosol fields for our models. In a context of increasing model evolution towards earth-integrated systems capable of representing the different components of the climate system (meteorology - atmospheric composition - surfaces - oceans, etc.), these new observations will allow to initialize these different components thanks to coupled data assimilation. Hence the interest in using level 1 data for IRS assimilation, leaving control and flexibility in the use of observations. The scientific outlook at Météo-France fully embraces the evolution of our models towards this new generation of systems, in particular for the in-line use of atmospheric composition in numerical weather prediction models. Previous studies have demonstrated the benefits of using atmospheric composition and meteorology together. Indeed, realistic chemistry or aerosol fields enable more accurate simulation of infrared observations, leading to more efficient assimilation of observations and hence improved weather forecasts. This is also the case for the improvement of radiation schemes through a more accurate consideration of ozone and aerosol fields, for example. These positive feedbacks are numerous and offer unlimited scope for research on these subjects.

5 OSSE on Sentinel 4 NO_2 assimilation

5.1 Introduction

The assimilation of synthetic NO_2 observations in Task 3.3.3 of the CAMEO project supports the Copernicus Atmosphere Monitoring Service (CAMS) in preparing its regional modelling systems for the upcoming generation of spaceborne air-quality observations. A primary goal is to enhance the assimilation of satellite retrievals by integrating new data streams from the geostationary Copernicus Sentinel-4 mission, which will provide hourly measurements of key air-quality trace gases and aerosols at high spatial resolution over Europe. These observations are expected to enhance near-real-time air quality analysis and forecasting capabilities significantly.

Since Sentinel-4 observations were unavailable during the project, CAMEO focused on developing and testing the necessary assimilation methods using synthetic data created with Observation System Simulation Experiments (OSSEs). The work depended on Nature Runs generated by the GEM-AC model to offer realistic variability of NO_2 mixing ratios in the troposphere. From these Nature Runs, synthetic NO_2 retrievals, reflecting the characteristics of Sentinel-4, were produced and assimilated into a subset of CAMS regional models: CHIMERE, GEM-AQ, LOTOS-EUROS, and MONARCH.

The goal was to assess how integrating Sentinel-4–like NO_2 observations could influence regional air quality forecast performance, thereby supporting CAMS in preparing for the operational use of the Sentinel-4 data.

The OSSE experiment diagram is shown in Fig. 5.1. The Nature Run and the generation of the synthetic observation with related observational errors are central to the modelling system. The Assimilation run uses synthetic observations, while the Free run (or Control run) helps evaluate the impact of the assimilation cycle. The analysis of the modelling streams is outlined in the following sections.

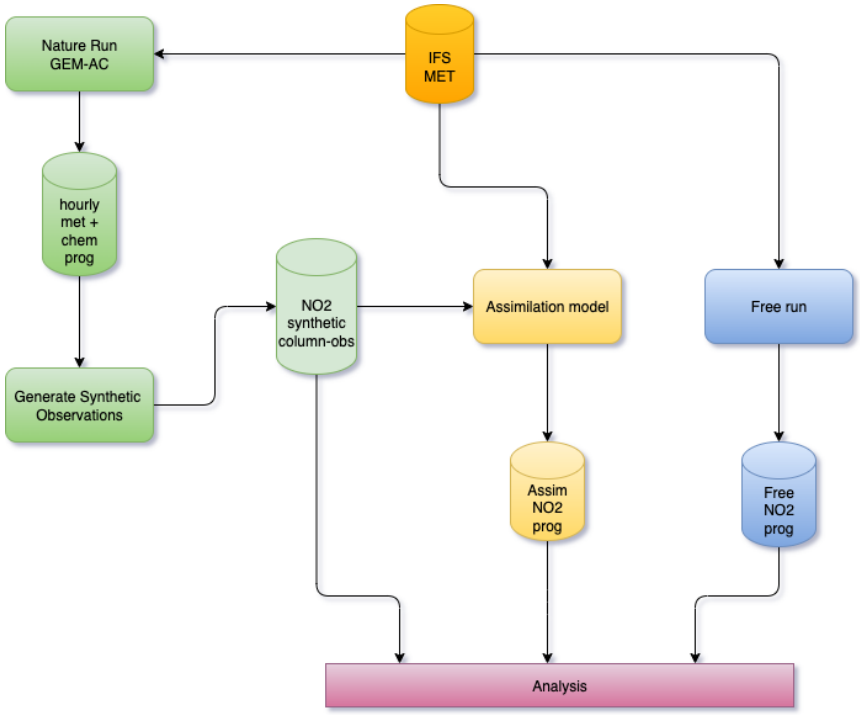

Figure 35 – T3.3.3 OSSE experiment diagram, where the assimilation models are CHIMERE, GEM-AQ, LOTOS-EUROS, and MONARCH.

5.2 Nature run – GEM-AC model configuration

The nature run was produced with the Global Environmental Multiscale Atmospheric Chemistry model (GEM–AC), which is based on the global general circulation and weather forecast model, with comprehensive tropospheric–stratospheric chemistry. The GEM (Global Environmental Multiscale) model was developed by the Meteorological Service of Canada for operational weather forecasting (Côté et al., 1998).

Current simulations for the nature run were performed on a global variable-resolution grid with a uniform horizontal resolution of 0.15° × 0.15° over Europe, extending to a top altitude of 0.1 hPa and comprising 77 vertical levels. The gas-phase chemistry includes a comprehensive reaction scheme covering Ox, NOx, HOx, CO, CH_4, NMVOCs (non-methane volatile organic compounds), halocarbons, ClOx, and BrOx, suitable for stratospheric chemistry, as well as general tropospheric and air quality (AQ) chemistry (de Grandpré et al., 1997; Lupu et al., 2013, Kossakowska and Kaminski, 2020).

Tracers are advected using the semi-Lagrangian scheme native to GEM. The vertical transport includes parameterized subgrid-scale turbulence and deep convection. Dry deposition is implemented as a flux boundary condition in the vertical diffusion equation. The aerosol modal model M7 handles aerosol microphysics (Vignati et al., 2004). Gas–aerosol partitioning is calculated by the thermodynamic equilibrium model ISORROPIA (Nenes et al., 1998).

The emissions used for this run include yearly-averaged anthropogenic and monthly-averaged biogenic, ocean, soil, and biomass burning emission fluxes, as well as NOx from lightning. CH_4, N_2O, and halogenated species are given as lower boundary conditions. Aerosol and precursor (NH_3 and dimethyl sulphide) emissions are from AeroCom (Dentener et al., 2006). Anthropogenic emission fluxes over Europe were year-specific and defined as CAMS-REG-ANT_v7_0 (Figure 36) and CAMS_TEMPO_v4_1.

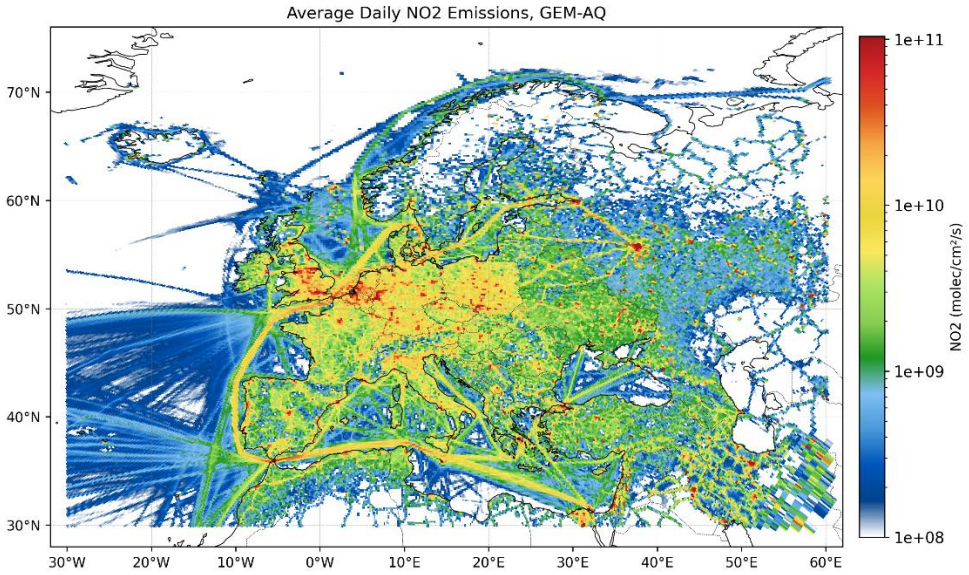

Figure 36 – Daily average NO_2 emission flux in molec./cm^2/s for August 2022.

In order to select the most suitable period for generating synthetic Sentinel-4 observations, Sentinel-5P (TROPOMI) data were analysed. It was assumed that a period with high coverage of high-quality pixels from Sentinel-5P would be representative.

The analysis of TROPOMI NO_2 data showed clear seasonal variations in data completeness, with the highest average number of pixels with QA > 0.75 occurring in August. The nature run was integrated for July and August 2022.

5.3 Synthetic observations

Synthetic observations of NO_2 tropospheric column densities were calculated from the results of the Nature Run. Since the synthetic columns are designed to reflect the observational geometry of a geostationary satellite, it was necessary to carry out the appropriate geometric transformations from the GEM-AC spherical coordinate system.

After creating the grid that reflects the coordinates of the pixel centres and corners, the intersection of the field-of-view (FOV) with the model domain is identified. This process results in a set of tuples of these coordinates with an irregular structure, which is then mapped to the rotated-pole spherical coordinates of the GEM-AC model. To perform interpolation in this setup, the xESMF package was employed, as it facilitates the use of grid-remapping functions available within the Earth System Modelling Framework (ESMF) project.

A conservative interpolation method was used for the irregularly distributed data. This ensures that the resulting columns contain the correct mass fragments originating from the intersecting grid cells of the GEM-AC model.

Next, model-predicted mixing ratios within individual model layers were converted to column units using the predicted pressure values and hydrostatic relationships. This allows for determining the contributions of individual layers to the total columns. Finally, the random forest generated during the course of the CAMS-61 project (provided by BSC team to ensure

sustainability and impact of earlier CAMs projects) was used to generate the averaging kernel and the observation error matrices, and the simulated retrievals were computed through convolution with the kernel.

5.4 Impact of assimilation on forecast performance

The assimilation of synthetic NO_2 observations was performed using a subset of CAMS regional models: CHIMERE, GEM-AQ, LOTOS-EUROS, and MONARCH.

5.4.1 GEM-AQ data-assimilation results

GEM-AQ data assimilation setup

The GEM-AQ model was run at a horizontal resolution of 0.1 degrees and 28 hybrid vertical levels, with the model top at 10 hPa. Meteorological forcings were taken from the ECMWF IFS forecasts. The overall model setup follows the CAMS regional production specifications (Colette, et al., 2025).

For August 2022, it was assumed that daylight hours extend approximately from 05:00 to 21:00 local time, defining the assimilation window. Within this window, the hourly synthetic Sentinel-4 tropospheric column retrievals are expected to influence the full 3-D structure of NO_2 concentrations.

The GEM-AQ model was coupled at observation times with the assimilation system. The assimilation method used is a variational approach based on Sasaki (1969). The cost function is a weighted sum of the squares of the background error and the analysis increment in the model grid's state space. The weighting factors express the estimated accuracy of measurements and model results, in relative terms. Minimization of the functional occurs under a dynamical constraint of linear advection (Burgers equation). The resulting solution is an elliptic (non-homogeneous Helmholtz-type) equation, which is then solved iteratively using a relaxation method.

The experiment design involved three consecutive days of assimilation, followed by a one-day forecast. A synthetic Sentinel-4 assimilation was carried out for 7, 8, and 9 August, followed by a one-day free forecast. A control run (BASE) was initially performed as a free forecast starting on 7 August at 00 UTC.

Evaluation approach

To assess the impact of synthetic Sentinel-4 tropospheric NO_2 column assimilation, we selected five regions with varying levels of anthropogenic emissions (Figure 37). The GEM-AQ model output from both the assimilation and control experiments was compared against synthetic surface observations sampled from the GEM-AC nature run for selected sub-regions.

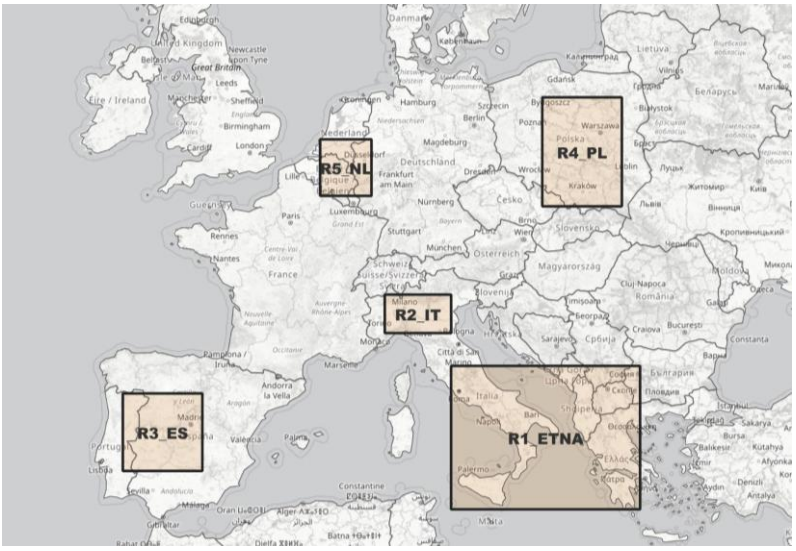

Figure 37 – Map of selected subregions: R1_ETNA, R2_IT (Po Valley), R3_ES (Spain), R4_PL (Poland), R5_NL (Belgium/The Netherlands).

Comparisons of AVK vs. unity kernel (robustness of the methodology)

Initial assimilation tests were carried out using the GEM-AQ model, in order to quantify the sensitivity to the averaging-kernel approach. Two setups were evaluated: (1) assimilation using a unity kernel combined with randomly generated observation errors, and (2) assimilation using averaging kernels and error parameters derived with the CAMS-61 methodology developed by BSC.

ASSIM-NO2-UNITY ASSIM-NO2-AVK

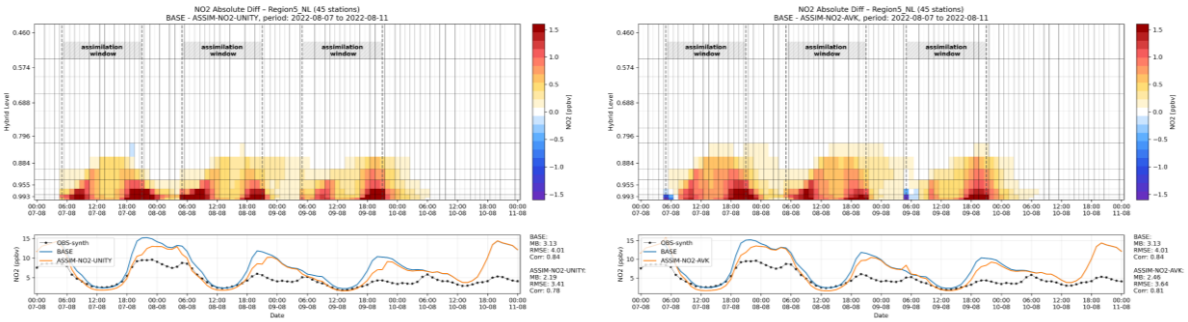

Figure 38 – Comparison of NO_2 concentrations for Region 5, calculated with unity and height-dependent averaging kernels.

Assimilation experiments using the unity kernel and the full averaging-kernel (AVK) approach (Figure 38) produced broadly similar results in terms of the temporal and vertical persistence of the NO_2 assimilation signal. Both configurations effectively introduced the synthetic Sentinel-4 information into the model and maintained its influence for several hours after the assimilation window.

As expected, the unity-kernel configuration produced a stronger impact in the lowest model layers. Because the unity kernel does not redistribute the vertical sensitivity of the retrieval, the analysis increments are applied more directly to the surface and near-surface levels. In contrast, the AVK-based assimilation distributes the increments according to the vertical sensitivity of the retrieval, resulting in a more physically consistent but slightly weaker adjustment near the surface.

Overall, both approaches produced comparable large-scale effects, with a more pronounced response in the lowest layers when using the unity kernel. This indicates that, despite the relatively high uncertainty associated with AVKs derived from Sentinel-5P, the findings of this study remain robust. For the subsequent analysis of the GEM-AQ results, only the outcomes obtained using the AVK-based assimilation approach were used.

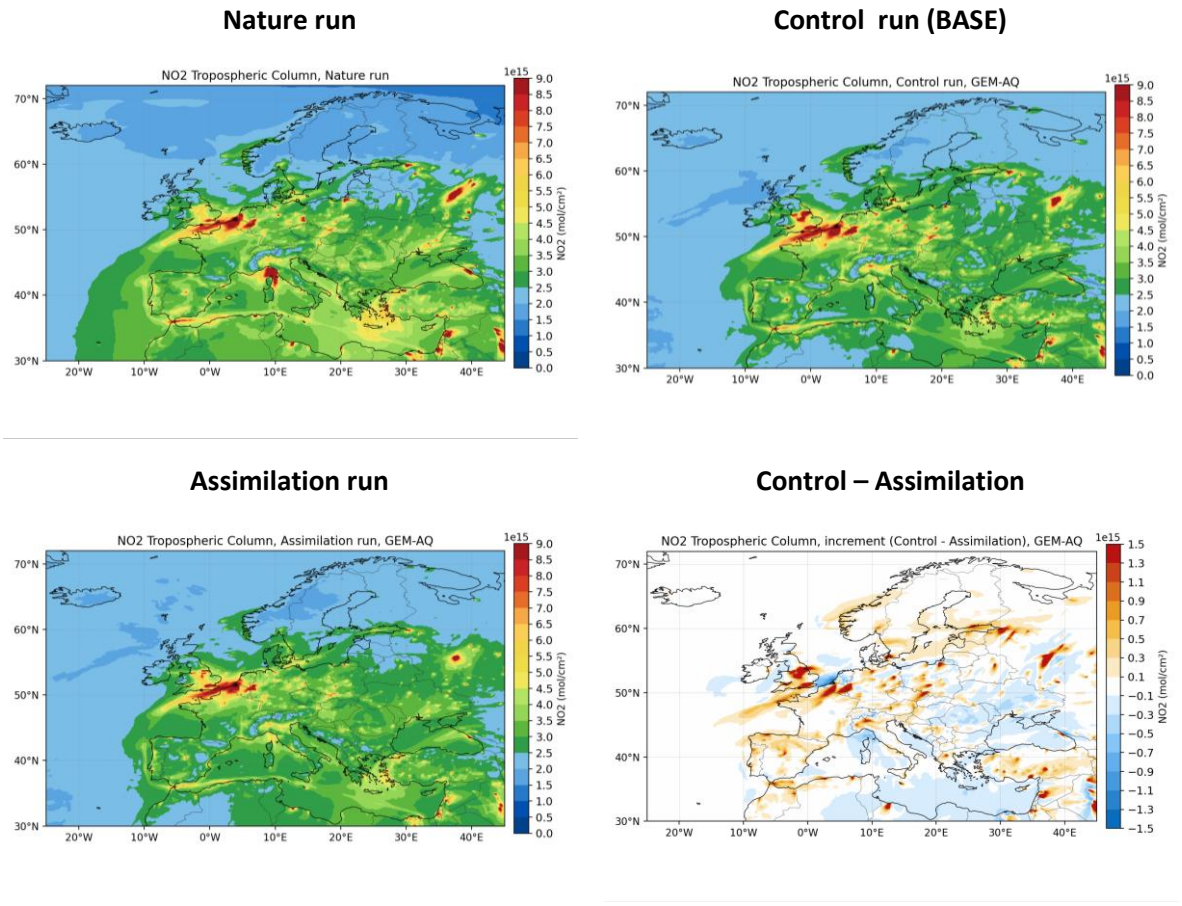

Figure 39 – NO_2 tropospheric column densities averaged for 3 days from August 7 to 9, 2022 for (top left) the Nature run, (top right) the control run with no assimilation, (bottom left) the assimilation run, and (bottom right) the absolute differences between the control and assimilation runs. Units are E+15 molec/cm^2.

Impact of assimilation on the NO_2 tropospheric column

NO_2 tropospheric column densities averaged from August 7 to 9, 2022, for the Nature, Control, and Assimilation runs are shown in Figure 39. Overall, the assimilation process results in a reduction of the tropospheric column density compared to the control run. Significant reductions occur in the vicinity of agglomerations, industrial regions, and selected shipping routes.

Impact of assimilation on surface and upper level concentration per region

The assimilation improved the agreement between the modelled NO_2 concentrations and synthetic observations (Figure 40). During the assimilation window, the NO_2 concentrations from the assimilated run converged toward those of the Nature Run, reducing the bias seen in the control simulation. This effect was consistent across all regions, showing that the assimilation process effectively adjusted near-surface NO_2 levels towards the target values. The persistence of this effect was limited in duration, gradually weakening after the end of the assimilation period (21:00 UTC) as the system evolved freely overnight. On the forecast day without assimilation, the two simulations nearly converged again, indicating that the impact of the assimilation lasted for about 6 to 12 hours after the last update.

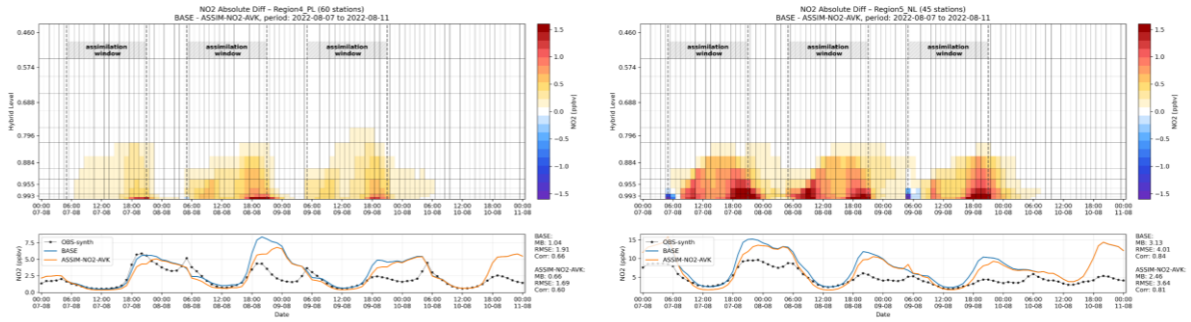

Figure 40 – Vertical and temporal absolute differences in NO_2 for the control and assimilation simulations for Region 4 and 5.

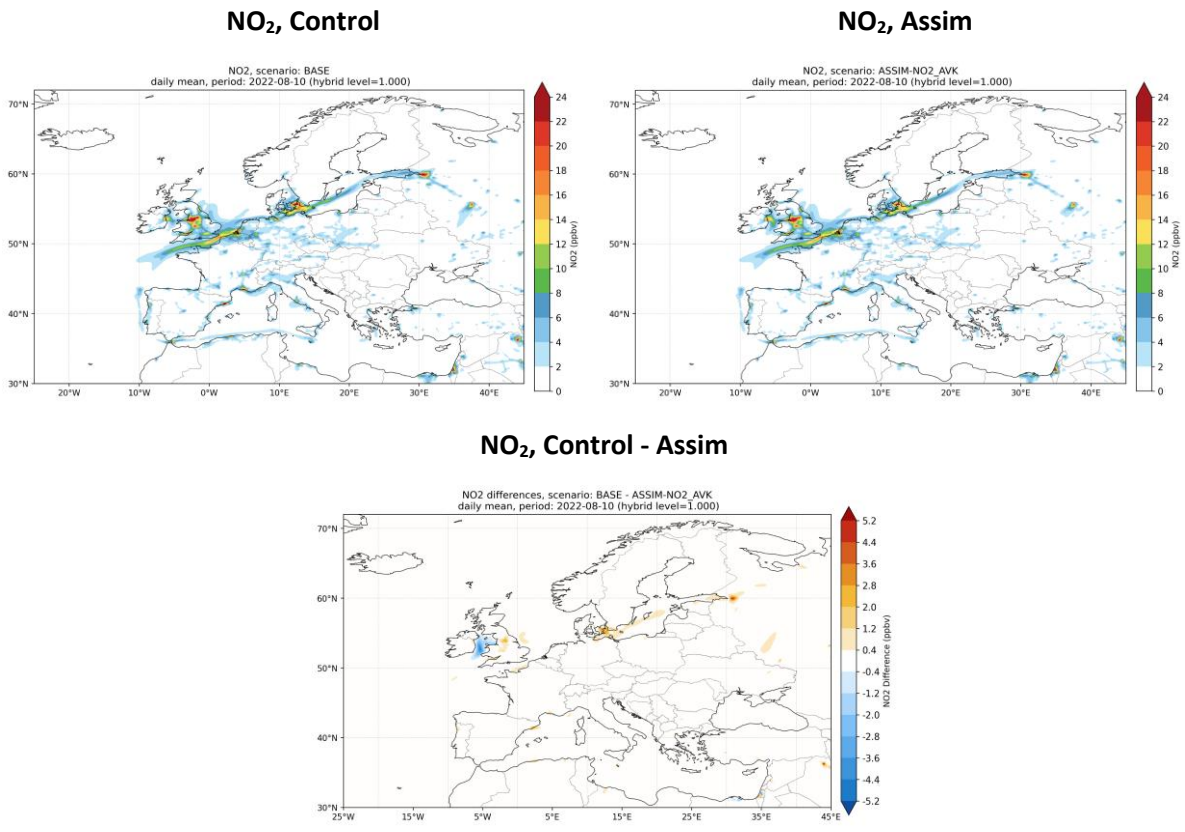

Figure 41 – Maps of the near-surface daily mean NO_2 concentrations for (top left) the control run without assimilation, (top right) the assimilation run, and (bottom) the absolute differences between the control and assimilation runs.

To evaluate the impact of Sentinel-4 assimilation on the forecast, near-surface daily mean NO_2 concentrations were examined for August 10 and presented in Figure 41. For both the free-run forecast (control) and the forecast following three days of assimilation, the spatial distribution remains very similar, with shipping emissions dominating the pattern over the sea and traffic emissions shaping concentrations over major urban areas. A closer examination of the differences between the two forecasts reveals an irregular pattern along the ship tracks over the Baltic Sea, as well as localized signals near cities such as St. Petersburg and Barcelona. An additional interesting feature appears in the strait between Ireland and the British Isles (Figure 41).

5.4.2 CHIMERE data assimilation results

CHIMERE data assimilation system setup

The CHIMERE model is run at a horizontal resolution of 0.2 degrees. Twenty vertical model levels are used with a top of the domain at 200 hPa. The model is coupled with the Data Assimilation Research Testbed (DART) system from the U.S. National Center for Atmospheric Research (NCAR). The Ensemble Adjustment Kalman Filter (EAKF, Anderson, 2001) runs with 20 members. The main perturbations are applied to both anthropogenic and biogenic emissions. Perturbations on boundary conditions are also applied. The Gaspari and Cohn (1999) localization is employed to mitigate the effects of long-distance correlations. The CAMS Satellite Operator (CSO) forward operator has been implemented within the CHIMERE-DART-EAKF framework.

Quality control and pre-processing of observations

Only clear sky observations (pixels with cloud fraction >0 are excluded) with a solar zenith angle lower than 60° are selected. One of the main goals of this experiment is to test the capacity of the system to assimilate the large amount of data provided by the synthetic NO_2 columns of Sentinel-4. Around 40,000 observations are available between 9 am and 14 pm, depending on the cloud presence. No super-observation and no thinning are applied.

Assimilation of NO_2 column retrievals in CHIMERE

Data assimilation of synthetic Sentinel-4 NO_2 tropospheric columns is performed over 3 days from 07/08/2022 to 09/08/2022. A free run (no assimilation) and a run with assimilation are compared. The CHIMERE DA ensemble CPU cost is mostly due to running the forecast of the ensemble. The data assimilation system can run with all available observations.

The comparison of the observations and the free run of CHIMERE (no assimilation) in Figure 42 shows an underestimation of CHIMERE in the Benelux, in the UK, in Po Valley, and in some large agglomerations such as Paris and Madrid. Shipping routes are underestimated more strongly in the west coast of Portugal and across the Channel Sea. The NO_2 total columns observed in the Mediterranean Sea close to Marseille in France and the area from south of Greece to the Türkiye coasts are larger than the Vertical Columns Densities (VCD) modelled by CHIMERE. CHIMERE has larger NO_2 total columns in several towns in Italy. The run with assimilation allows for increasing total columns in the underestimated areas, as shown by the positive increment, except near Marseille. Moreover, no significant decrease is noticed in Italy. It can be expected to have a low impact when the observation and the model prior are highly different. The spread ensemble, definition of the outliers and the value of the observation conditioned the weight in the analysis.

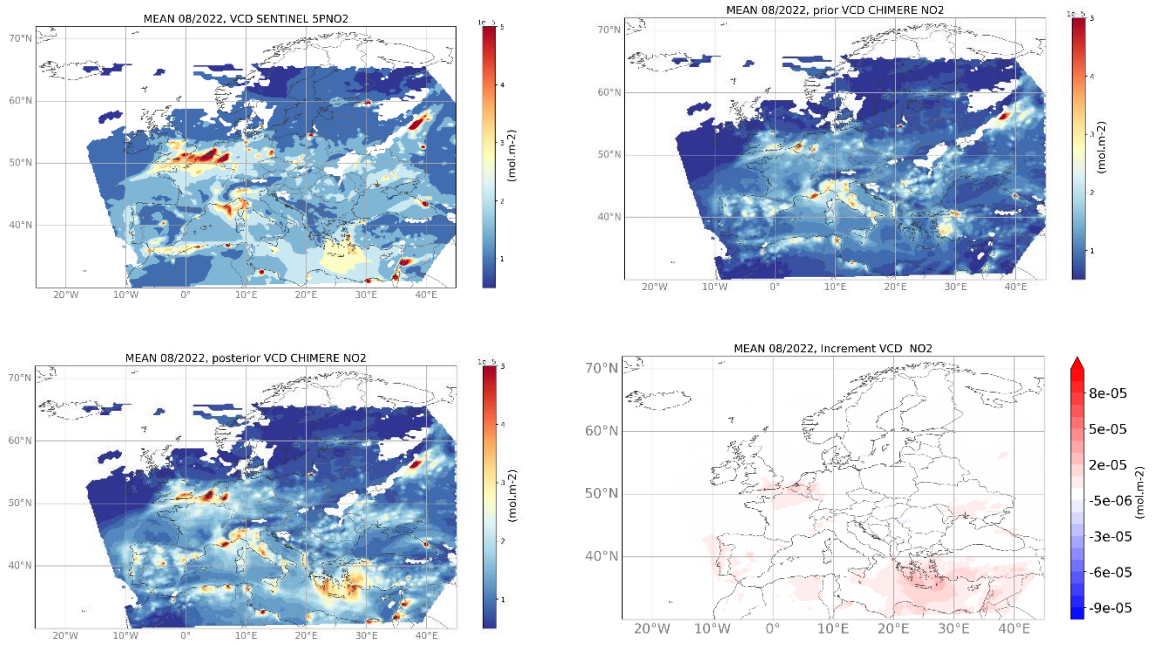

Figure 42 – NO_2 total columns mean of the S4-OSSE satellite observations on the top left, of CHIMERE simulated VCD mean on the top right (no assimilation), CHIMERE simulated VCD mean with assimilation on the bottom left and the difference CHIMERE mean assimilation-no assimilation on the bottom right.

Figure 43 shows that the assimilation of the NO_2 synthetic total columns leads to an increase of NO_2 surface concentration in Benelux, Paris, Madrid, along the shipping routes and the Mediterranean area across Greece to Türkiye coasts.

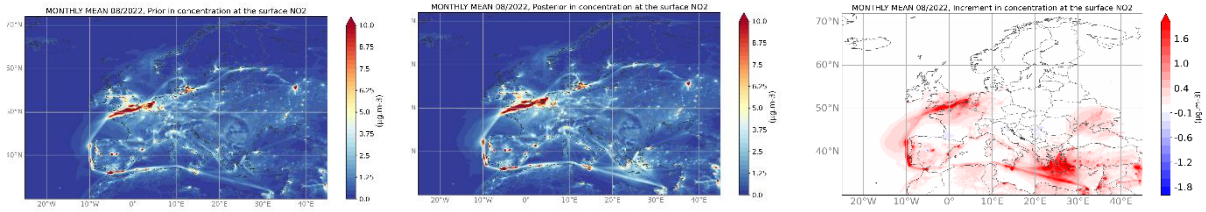

Figure 43 – NO_2 surface concentration for CHIMERE no assimilation (left), with assimilation (middle), and the difference between CHIMERE with and without assimilation (right).

Finally, hourly results in RMSE, correlation and bias for the free run and the run with assimilation are compared to the synthetic NO_2 surface observations in Table 4. Improvement is mainly found for bias, allowing a reduction of the underestimation.

Table 4 – Comparison of CHIMERE hourly results of the free run and the run with assimilation to NO_2 synthetic surface concentrations over the 3-day period from 07/08/2022 to 09/08/2022.

	BIAS (µg/m³)	RMSE (µg/m³)	R
free run (no assimilation)	-1.16	4.74	0.71
run with assimilation	-1.05	4.71	0.71

5.4.3 MONARCH-DA data assimilation results

MONARCH-DA data assimilation methods

The data assimilation (DA) experiments were conducted using the MONARCH-DA data assimilation system, used at BSC to produce the regional CAMS gas and aerosol reanalyses. It relies on the NMMB-MONARCH model to compute the transport and chemical interactions of 34 aerosol and gas tracers. The MONARCH simulations were computed at a horizontal resolution of 0.2°, with 24 vertical levels up to 50 hPa. Anthropogenic emissions were provided by the HERMESv3 model, and meteorological forcings were taken from the ECMWF IFS forecasts. The data assimilation employs a LETKF approach, with six members perturbed in the emission space. The DA setup intentionally replicates the one used for producing the CAMS reanalysis; the only significant differences are the type of observations assimilated: synthetic NO_2 retrievals vs. surface NO_2 concentrations in the operational product, and the use of the analysis at day *n* to initialize the forecast at day *n+1*.

Panel A1 in Figure 44 represents the free-forecast run starting on 7 August, while B1 and B2 correspond to assimilation experiments in which synthetic satellite data were assimilated on 7 August (with forecasts for 8 and 9 August) and on 7 and 8 August (with the forecast for 9 August), respectively.

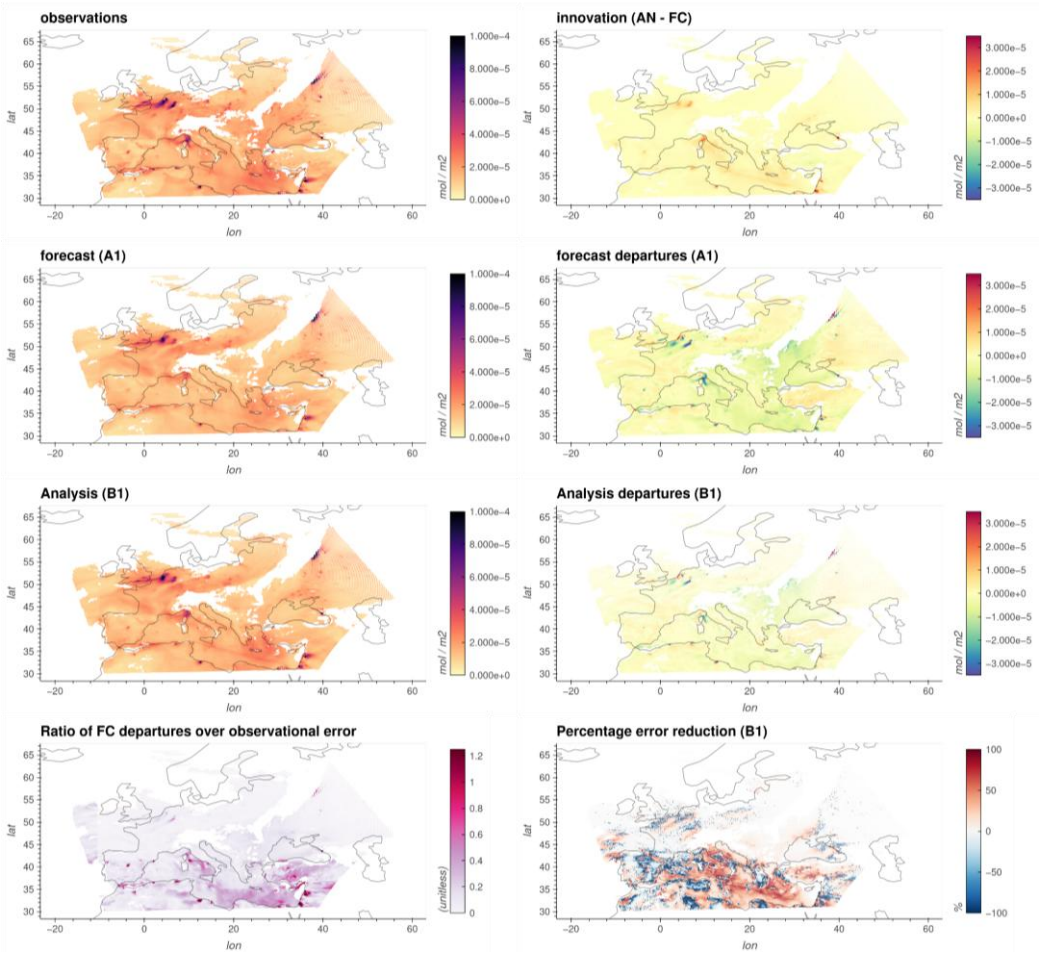

Figure 44 – Left column, from top to bottom: Synthetic observations; model estimate (forecast), model estimate (analysis), and ratio of forecast model-data mismatches over observation uncertainties. Right column: Innovation, in observation space (analysis minus forecast); forecast model data mismatches; analysis model data mismatches; percentage error reduction, defined as 100 * (1 – |error_an| / |error_fc|). All plots are for the first day of analysis.

Observation dataset curation

Clear sky observations (i.e., with a total cloud cover of 0) and a solar zenith angle lower than 60° were selected for assimilation. All synthetic observations were assimilated independently (i.e. no super-observations were constructed), which allowed us to assess the technical performance of the system in the most unfavourable conditions.

Results - Fit to assimilated observations

The forecast shows a negative bias of -12% compared to the synthetic observations, with two hot spots with significant error (both positive and negative): over the Benelux and the Ligurian Sea (Figure 44). In addition, the NO_2 columns are uniformly underestimated over the Eastern Mediterranean Sea and over Eastern Europe (Ukraine, Romania, Moldavia). Finally, MONARCH tends to underestimate the NO_2 columns over large cities (e.g., Moscow, Madrid, Berlin, Hamburg, Tripoli, Beirut) with respect to the synthetic observations computed with the GEM-AC nature run.

The DA reduces the overall bias to -6% and brings the model in much better agreement with the observations around the Mediterranean Sea (see, for instance, Madrid and Beirut). The bias over the Eastern Mediterranean Sea is also largely corrected. However, further north, there are hardly any changes, except over the Netherlands, where the forecast error was very large. This is, however, consistent with the prescribed observation uncertainties, which are very large, in relative terms, over 42°N (bottom left plot in Figure 44).

We also note a tendency of our data assimilation system to degrade the fit to observations in many locations (negative error reductions). This likely results from insufficient variability in our ensemble, which leads to spreading strong error corrections to the nearby grid cells. These relative error increases are, however, rather insignificant in absolute terms as they occur in locations of low NO_2 concentrations.

Results - Fit to assimilated observations (non-assimilated)

The forecast model fits the nature run surface observations with a negative bias of -2.5 ppb (~-51%). The assimilation of satellite data leads to overall negligible fit improvements to the surface observations (see Table 5 and Figure 45). This is expected: as the averaging kernels of the observations peak towards the upper troposphere, the DA tends to localize concentration adjustments there, leaving the surface layer largely unchanged. We, however, note that net error reductions tend to be obtained at sites where the NO_2 values are elevated, and/or where the forecast RMSE was the largest (Figure 46), which is encouraging.

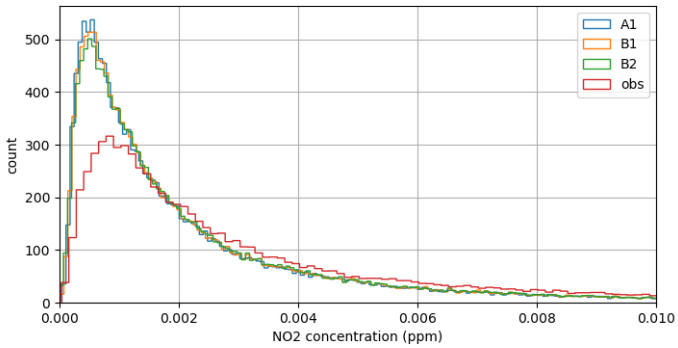

Figure 45 – Distribution of observed (synthetic) and modelled NO_2 concentrations.

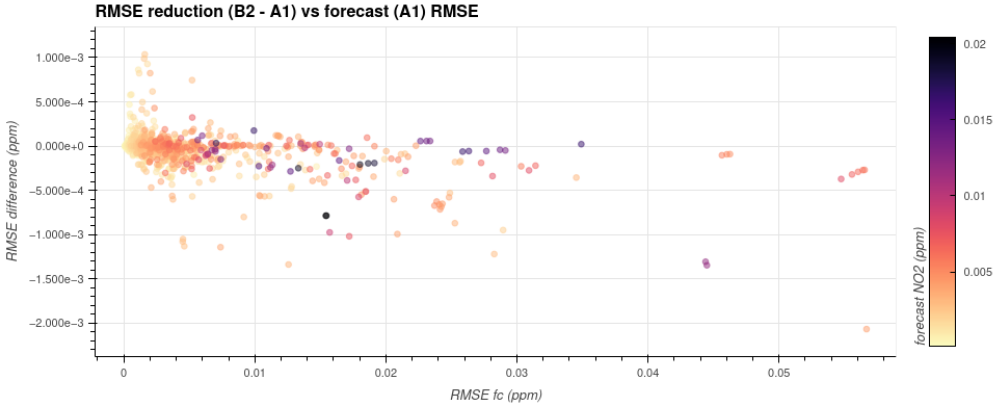

Figure 46 – RMSE reduction (analysis minus forecast), vs. forecast RMSE, for all the surface observation site. The colorbar shows the mean NO_2 concentration at the sites.

Table 5 – Fit statistics to the synthetic surface observations on day 2 of analysis (9 August 2022). The statistics are computed independently for each observation site, and the "Mean" represents the average of all site-specific values. Day 2 is shown as it compares 1) a totally free forecast for "A1", 2) a day with assimilated observations for "B2", and 3) a forecast initialized from the outcome of an analysis on the previous day (B1).

	Bias (ppb)			RMSE (ppb)			Correlation coefficient		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
A1 (fc)	-49.28	-2.59	4.25	0.07	4.14	56.67	-0.16	0.66	0.96
B1 (an)	-49.18	-2.56	4.27	0.07	4.12	56.48	-0.16	**0.67**	0.96
B2 (an)	-48.81	**-2.50**	4.61	0.07	**4.10**	56.26	-0.19	0.65	0.96

Impact on forecast performance

The impact of the assimilation of satellite retrievals on forecast performance was tested through the forward propagation of the analysis state at the end of day 1 to the next 48 hours. It, however yielded no noticeable impact on forecast performance, likely because the emissions remained the same, as well as the boundary and initial conditions for all other species (Figure 47).

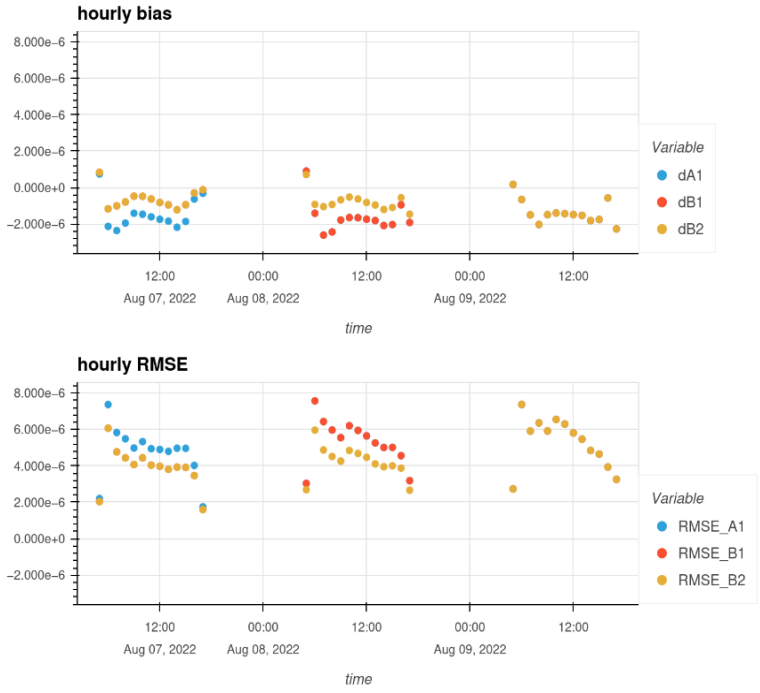

Figure 47 – Hourly evolution of the fit residuals (bias and RMSE) in the forecast (A1) and analyses (B1 and B2). There is no noticeable difference between A1 and B1 on day 2 (when only B2 assimilates observations), and between A1, B1 and B2 on day 3 (when none of the simulations assimilates data).

5.4.4 LOTOS-EUROS data assimilation results

LOTOS-EUROS data assimilation system

Detailed descriptions of the LOTOS-EUROS LETKF data assimilation system and model configuration were provided in section 3.2.1. The key differences for the assimilation exercise performed here versus that for NH_3 described in section 3 was that a temporal correlation τ of 1 day was used for NO_2, and a spatial correlation length was estimated based on the pixel sizes for S4, which is expected to be 8×8 km^2 therefore a correlation length of ρ = 8 km was used in the assimilation.

Data criteria

The covariances, columns and kernels were pre-processed with CSO to match the requirements for its use in LOTOS-EUROS. Besides light pre-processing, the simulated observations were used as provided without further criteria applied to the dataset and assimilated individually.

Results – NO_2 tropospheric column

The assimilation was performed over a short window of 3 days between the 7th and 9th of August 2022. This is done to illustrate the potential for using the S4 observations for near-real time optimization of the emission fields and using the resulting fields for a forecast of the next day. While the system is not optimal in the sense that more of the past information might be available (observations were only available for a short period), it should function as an illustration of the potential capabilities. Optimization of the emission fields is performed for each hour where observations are available, which for S4 is limited to daytime. An overview of the results of the assimilation run is given in Figure 48.

Panel (a) and (b) of Figure 48 show the mean simulated satellite column concentrations and the matching uncertainty of the observations. While S4 covers most of the domain, the availability of the observations is limited by the cloud cover and viewing conditions, meaning there are more observations available towards the south of the domain. The mean simulated error also reduces from north to south, and in its current state seems to be somewhat of an overestimation of the expected uncertainty of the observations. The stepwise decrease in the mean error level from north to south is also reflected in the results of the assimilation. The base LOTOS-EUROS model run (panel c) concentration levels are higher than the simulated observations for most of the domain, which indicates that after assimilation the emission fields can be expected to go down, resulting in lower column concentrations in the assimilated model state. The observation errors will function as a lever that decreases the impact of assimilation of the observations as a function of error total. This is the case as can be observed in the resulting assimilated model state and the absolute and relative differences between the two LOTOS-EUROS states (panel e and f). A near horizontal line can for example be observed from the north of Spain to north-eastern Turkey. Above this line the model states show much less of a change compared to the region below the line. Some exceptions can be observed, especially around the major hotspots, where the significant difference in model and synthetic observations will still result in a reduction of the emissions.

Based on the results it can be concluded that LETKF functions as expected, and the system can match the observations by increasing or decreasing the emissions, modulated by the overall uncertainty level of the ingested observations. Runtime was also acceptable, while still unoptimized, with the LETKF assimilation run for the whole domain taking about 12 hours to complete, including assimilation window of 3 days, and a forecast for the next three days (only one shown in later figures).

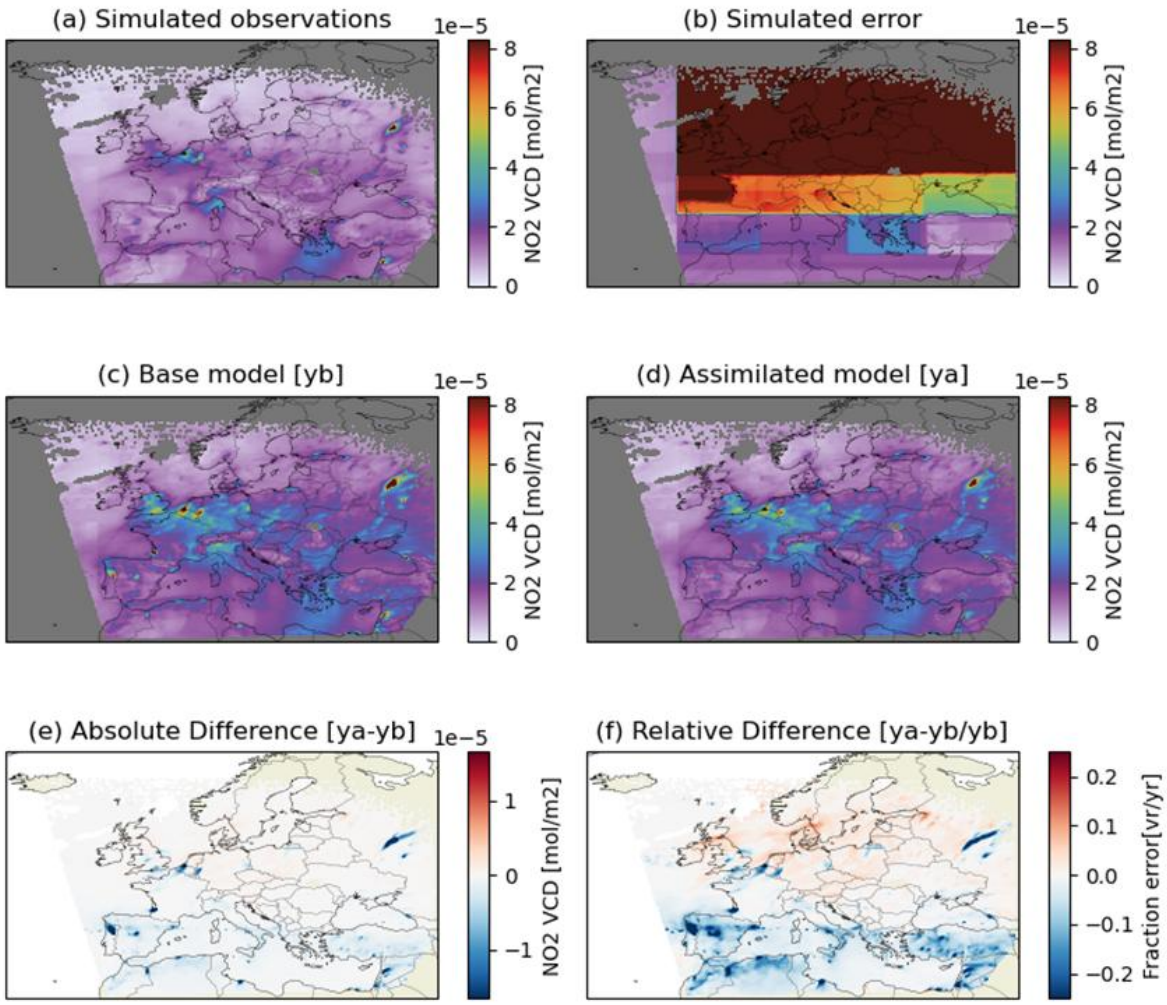

Figure 48 – The mean NO_2 tropospheric column at the satellite footprints, averaged for the period of the 7th-9th of August 2022 from (a) the simulated (synthetic input) observations, (b) the simulated error, (c) the base LOTOS-EUROS simulation, (d) the LOTOS-EUROS simulation after assimilating the observations, (e) the absolute difference between both LOTOS-EUROS runs, an (f) the relative difference.

The LETKF functions by optimizing the underlying emission fields and the resulting perturbed emission fields are illustrated in Figure 49. The figure shows the mean daily emission change for the period of the 7th-9th of August. Like the column concentration fields the emissions show a strong change below the line from north-western Spain to north-eastern Turkey, especially for emissions of smaller or more widely distributed sources. Meanwhile emissions of large clusters and major cities are still shown to change for region north of this line with the uncertainty of the individual observations proving too large to result in changes of smaller emission clusters. Combined with Figure 48 it can be observed that large emissions, from potentially a cluster of fires in Portugal and Spain, are completely removed from the emission fields after assimilation, again showcasing the potential of such a system to pick up period emission changes. Similarly, emission clusters in Iraq, Israel and Syria show increases whereas the surrounding emissions mostly show a decrease, showcasing the potential to pick up localized changes in the emission fields. Finally, emissions of some of the major cities in the northern parts of Europe, with consistently lower column totals, are also reduced over the span of the assimilation.

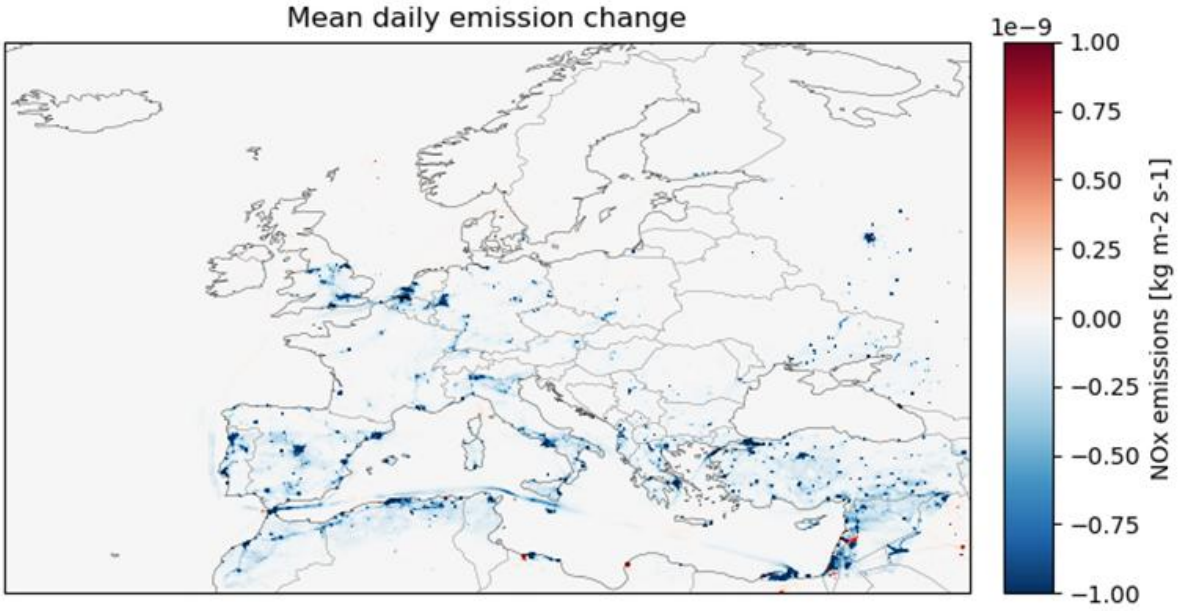

Figure 49 – The mean daily NOx emission change for 7-9 August 2022.

Results - Surface level concentrations

The resulting concentration fields were compared to synthetic surface observations sampled for the location sites within the surface observation network. Site locations were matched with the grid cells within the model domain and directly sampled for each of the hours of the assimilation and forecast periods. The concentrations were averaged for all locations into a single timeseries of which the result is shown in Figure 50. The base LOTOS-EUROS run, and synthetic observations are shown in respectively blue and black. Contrary to the averaged column concentrations at S4 footprints, the mean surface concentrations within the LOTOS-EUROS model are higher than the synthetic observations, displaying the differences between the LOTOS-EUROS and the Nature Run. As the optimized emission fields were lower, the concentration of the assimilated state is also lower for most of the model period, with a slight mean increase observed for parts of the 9th of August. Similarly, the resulting forecast concentrations are also lower than the base LOTOS-EUROS model. Any improvement of performance for the forecast model is hard to quantify due to the intrinsic differences in the column and surface concentration levels.

As the southern part of the model domain showed the largest change between the base and assimilated states, a second set of mean concentrations was derived for only the sites within Portugal, Spain, Italy, Greece, and Turkey (shown in Figure 51). The resulting concentrations were supposed to highlight the performance of those locations. Differences between the base and assimilated model are larger than those observed when using the full dataset. However, again the resulting surface concentrations show a decrease, while there already is a low bias between the base model and synthetic observations, limiting what can be concluded from this exercise.

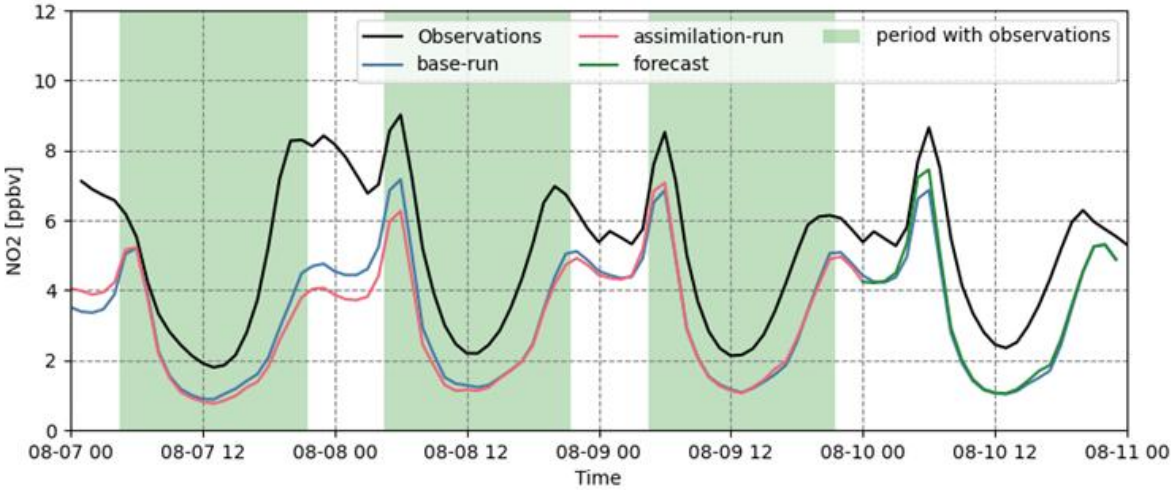

Figure 50 – Time-series of the mean surface concentrations at all sites within the network. Mean concentrations are shown for the synthetic observations, LOTOS-EUROS base, assimilated the forecast concentration fields. The green background illustrates periods when observations were available within the domain.

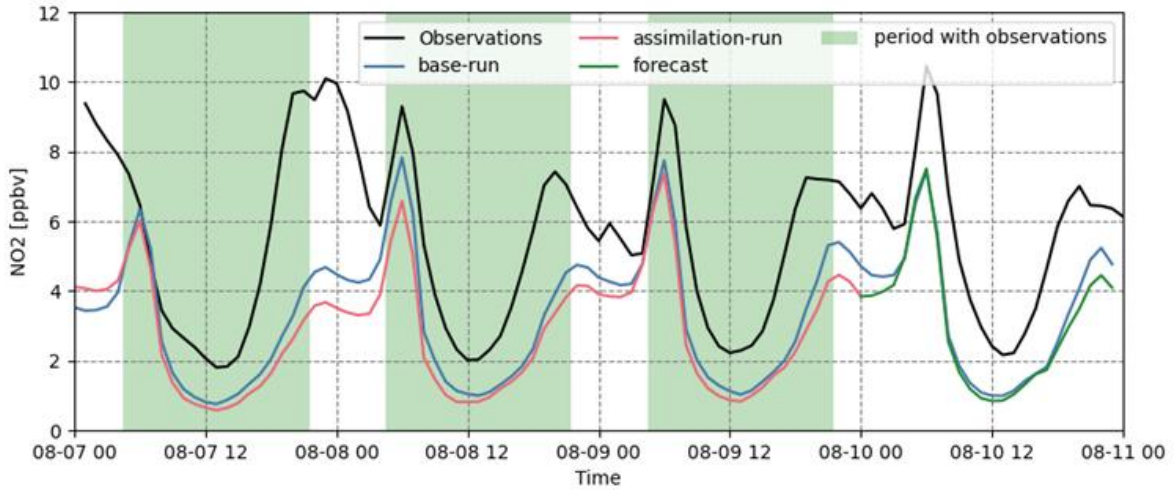

Figure 51 – Time-series of the mean surface concentrations at sites in Portugal, Spain, Italy, Greece and Turkey. Mean concentrations are shown for the synthetic observations, LOTOS-EUROS base, assimilated into the forecast concentration fields. The green background illustrates periods when observations were available within the domain.

5.5 Summary and conclusions

Assimilation of synthetic Sentinel-4–like NO_2 total columns leads to small but measurable improvements in surface-level NO_2 representation across the participating regional models. These improvements are strongest during the assimilation window and tend to diminish in the subsequent free forecast. Nevertheless, the OSSE confirms the feasibility of assimilating high-frequency geostationary column data and provides model-specific insights for future operational use.

For **GEM-AQ**, the assimilation improved the agreement between modelled NO_2 concentrations and the synthetic observations across all analysed regions, demonstrating that the system effectively adjusted near-surface concentrations toward the target values. The influence of assimilation persisted for roughly 6–12 hours after the last update (21:00 UTC), gradually weakening overnight and nearly disappearing by the forecast day, when the assimilated and free-run simulations converged again. The assimilation also introduced changes in NO_2 concentrations up to approximately 3–3.5 km altitude, depending on the region.

For **CHIMERE**, the ensemble DA system successfully assimilated more than 40,000 observations per cycle, although model forecast time remains a limiting factor. The model showed strong sensitivity to shipping corridors and major emission hotspots. Because CHIMERE and the Nature Run share nearly identical NO_2 emissions, differences in chemistry, transport, and other processes become key drivers of variability in total columns, an important consideration for interpretation and for future inversion work. Real Sentinel-4 biases will need careful handling for operational assimilation.

For **MONARCH**, assimilation of NO_2 columns produced very limited improvements at the surface and had a negligible influence on the subsequent forecast. This outcome reflects the instrument's highest sensitivity in the upper troposphere and the DA system's flexibility in adjusting vertical profiles away from the surface. Potential improvements include controlling emissions rather than concentrations, using larger or better-designed ensembles, and refining spatial localization. Despite challenging test conditions (no super-observation aggregation), computational performance remained acceptable for operational use.

For **LOTOS-EUROS** using LETKF, the tests confirmed that Sentinel-4–like observations can be assimilated within operational time limits. The resulting optimized emission fields brought the model state closer to the observations. Although Sentinel-4 provides valuable high-frequency information, optimizing emissions only at the observation time introduces delays. Future improvements should therefore focus on strengthening the temporal link between emissions and observed columns, for example, by widening the assimilation window or restarting the system a few timesteps before each update.

Overall, the OSSE experiment confirmed the feasibility of implementing assimilation of geostationary satellite NO_2 observations in regional air-quality models. Despite the technical challenges associated with handling large volumes of high-frequency data, all modelling systems successfully ingested the synthetic Sentinel-4–like retrievals within operationally realistic time constraints.

The results are promising for future Integrated Regional Analysis (IRA) production, showing that geostationary NO_2 data can effectively adjust model fields during the assimilation window and provide valuable additional constraints on the atmospheric state. At the same time, the impact on the subsequent forecast remains modest, with improvements typically limited to several hours after the last assimilation update.

Finally, the study highlights that, compared with direct modification of concentrations, optimizing emissions may offer a more effective pathway for extracting information from Sentinel-4 observations. Together, these findings provide a solid foundation for further development and refinement of assimilation strategies within the CAMS ecosystem ahead of the operational availability of Sentinel-4 data.

6 Conclusion

This deliverable demonstrates substantial progress in preparing the CAMS regional system for the integration of forthcoming geostationary satellite observations, specifically from MTG-IRS and Sentinel-4, through a series of comprehensive OSSEs. The results across the three main tasks provide clear evidence of the added value these new data streams and assimilation strategies can bring to regional air quality analysis and forecasting.

T3.3.1: Assimilation of MTG-IRS NH_3 Observations (Section 3)

The assimilation of synthetic MTG-IRS NH_3 observations into the LOTOS-EUROS and EMEP models showed that high-frequency, high-resolution geostationary measurements can significantly improve the representation of NH_3 emissions, atmospheric concentrations, and deposition fields across Europe. The LETKF (LOTOS-EUROS) and the newly developed ELVIS (EMEP) variational inversion system both demonstrated strong capability to reduce model biases and align simulated fields with reference data. Notably, the assimilation led to substantial reductions in surface and column concentration biases, as well as improved agreement in wet and dry deposition, particularly in regions with dense observational coverage. These advances highlight the potential for MTG-IRS data to enhance the accuracy of air quality analyses and forecasts within CAMS, especially for short-lived, spatially variable species like NH_3. The work also identified areas for further improvement, such as better handling of vertical profile mismatches and temporal variability, which could be addressed through incremental improvements to the current assimilation techniques (e.g., 4D-LETKF).

T3.3.2: OSSE on IRS O_3/CO Radiance Assimilation (Section 4)

The OSSE conducted for IRS radiances in the ozone (O_3) and carbon monoxide (CO) bands using the MOCAGE model demonstrated that direct assimilation of Level 1 radiances, rather than Level 2 products, is both feasible and beneficial for regional air quality forecasting. Assimilation of IRS radiances led to a significant reduction in forecast errors for both O_3 and CO, with improvements most pronounced in the atmospheric layers and regions where the IRS channels have the highest sensitivity. The positive impact was observed throughout the domain, especially in the troposphere and lower stratosphere, and was robust even under realistic operational constraints such as cloud screening and data thinning. These results confirm that IRS radiance assimilation can provide valuable additional constraints for regional chemical composition, supporting more accurate and timely forecasts in CAMS. The study also underscores the importance of preparing data assimilation systems for the technical and computational challenges associated with high-volume, high-frequency radiance data.

T3.3.3: OSSE on Sentinel-4 NO_2 Assimilation (Section 5)

The multi-model OSSE for Sentinel-4 NO_2 assimilation, involving CHIMERE, GEM-AQ, LOTOS-EUROS, and MONARCH, confirmed the feasibility and operational readiness of the CAMS regional ensemble to ingest high-frequency geostationary column data. All models successfully assimilated synthetic Sentinel-4 NO_2 retrievals within realistic timeframes, demonstrating technical scalability. Assimilation consistently improved the agreement between modelled and observed NO_2 fields during the assimilation window, particularly near the surface and in regions with strong emission sources. However, the impact on subsequent forecasts was generally limited to a few hours after the last assimilation update, indicating that further gains may require a shift from direct concentration adjustment to emission optimization. The experiments also highlighted the importance of model-specific factors, such as chemistry, transport, and vertical sensitivity, in determining the persistence and magnitude of assimilation impacts. These insights provide a strong foundation for the ongoing development of assimilation strategies and support the future operational uptake of Sentinel-4 data in CAMS.

Overall Implications for CAMS Regional System

Collectively, these developments and findings demonstrate that the CAMS regional system is well-positioned to benefit from the next generation of geostationary satellite observations. The tested assimilation frameworks are capable of handling the increased data volume and frequency, and can translate these new observations into tangible improvements in air quality analyses and forecasts. The work also identifies key areas for further enhancement, including:

- Refinement of emission optimization techniques to maximize the impact of assimilated data on forecasts.
- Improved treatment of vertical profile mismatches and temporal variability, especially for short-lived species.
- Continued development of computationally efficient, scalable assimilation systems to manage the operational demands of high-frequency satellite data.

By addressing these challenges, CAMS will be able to deliver more accurate, timely, and policy-relevant information to users, supporting better management of air quality and atmospheric composition across Europe. The methodologies and results presented here provide a robust blueprint for the integration of MTG-IRS, Sentinel-4, and future satellite missions into the CAMS regional production system.

7 References

Anderson, J. L. 2001. An Ensemble Adjustment Kalman Filter for Data Assimilation... Monthly Weather Review. Page(s): 2884–2903, DOI: https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2

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

CAMS61_2019SC1_D61.3.0.3_202106_Final_Report. 2021, TNO (A. Segers)

Chen, Z., Jacob, D. J., Gautam, R., Omara, M., Stavins, R. N., Stowe, R. C., Nesser, H., Sulprizio, M. P., Lorente, A., Varon, D. J., Lu, X., Shen, L., Qu, Z., Pendergrass, D. C., & Hancock, S. (2023). Satellite quantification of methane emissions and oil-gas methane intensities from individual countries in the Middle East and North Africa: implications for climate action. Atmospheric Chemistry and Physics, 23(10), 5945–5967. https://doi.org/10.5194/acp-23-5945-2023

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

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., … Rouïl, L. (2025). Copernicus Atmosphere Monitoring Service – Regional Air Quality Production System v1.0. Geoscientific Model Development, 18(19), 6835–6883. https://doi.org/10.5194/gmd-18-6835-2025

Coopmann, O., Fourrié, N., and Guidard, V. (2022). Analysis of MTG-IRS observations and general channel selection for numerical weather prediction models. Quarterly Journal of the Royal Meteorological Society, 148(745):1864–1885.

Coopmann, Olivier and Fourrié, N and Chambon, P and Vidot, J and Brousseau, P and Martet, M and Birman, C. (2023). Preparing the assimilation of the future MTG-IRS sounder into the mesoscale numerical weather prediction AROME model. Quarterly Journal of the Royal Meteorological Society, 149(757):3110–3134

Coopmann O., Guidard V., Vittorioso F. (2025). Preparation for assimilation of MTG-IRS radiances into an atmospheric composition model for ozone and carbon monoxide forecasts. Currently under review at Quarterly Journal of the Royal Meteorological Society

Côté, J., et al. (1998): The operational CMC–MRB Global Environmental Multiscale (GEM) model, Part I: Design considerations and formulation, Mon Wea Rev, vol. 126, pp. 1373–1395, doi:10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2.

de Grandpré, J., et al. (1997): Canadian middle atmosphere model: Preliminary results from the chemical transport module. Atmosphere-Ocean, vol. 35, pp. 385–431, doi:10.1080/07055900.1997.9649598.

Dentener, F., et al. (2006): Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom. Atmos Chem Phys, vol. 6, pp. 4321–4344, doi:10.5194/acp-6-4321-2006.

Earth System Modeling Framework - ESMF User Guide, Version 8.9.0, 2025 (available from https://earthsystemmodeling.org/)

Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. Ocean Dynamics, 53(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9

Fagerli et al., (2023) Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components. Status Report 1/2023, ISSN 1504- 6192. https://emep.int/publ/reports/2023/EMEP_Status_Report_1_2023.pdf

Gaspari, G., and S. E. Cohn. (1999). Construction of correlation functions in two and three dimensions. Quarterly Journal of the Royal Meteorological Society, 125(554), doi: 10.1256/smsqj.55416.

Gilbert, J.C., Lemaréchal, C. Some numerical experiments with variable-storage quasi-Newton algorithms. Mathematical Programming 45, 407–435 (1989). https://doi.org/10.1007/BF01589113

Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Physica D: Nonlinear Phenomena, 230(1–2), 112–126. https://doi.org/10.1016/j.physd.2006.11.008

Kossakowska, M. and J.W. Kaminski, A note on the potential impact of aviation emissions on jet stream propagation over the northern hemisphere, Acta Geophysica, DOI: 10.1007/s11600-020-00444-x, 2020.

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

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

Lopez-Restrepo, S., Yarce, A., Pinel, N., Quintero, O. L., Segers, A., & Heemink, A. W. (2020). Forecasting pm10 and pm2. 5 in the aburrá valley(Medellín, Colombia) via enkf based data assimilation. Atmospheric Environment, 232, 117507. https://doi.org/10.1016/j.atmosenv.2020.117507

Lupu, A., K. Semeniuk, J. W. Kaminski, J. C. McConnell, GEM-AC: A Stratospheric-Tropospheric Global and Regional Model for Air Quality and Climate Change-Evaluation of Gas-Phase Properties, The Atmospheric Chemistry Experiment ACE at 10: A Solar Occultation Anthology (Peter F. Bernath, editor, A. Deepak Publishing, Hampton, Virginia, U.S.A., 2013), 285-293, 2013.

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

Nenes, A., S. N. Pandis and C. Pilinis (1998): ISORROPIA: A new thermodynamic model for multiphase multicomponent inorganic aerosols. Aquat Geochem, vol. 4, pp. 123–152.

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., & Connor, B. J. (2003). Intercomparison of remote sounding instruments. Journal of Geophysical Research: Atmospheres, 108(3). https://doi.org/10.1029/2002jd002299

Sasaki Y., 1969: Proposed Inclusion of Time, Variation Terms, Observational and Theoretical, in Numerical Variational Objective Analysis. J. Met. Soc. Jpn., 47(2): 115 – 124.

Shephard, M. W., & Cady-Pereira, K. E. (2015). Cross-track Infrared Sounder (CrIS) satellite observations of tropospheric ammonia. Atmospheric Measurement Techniques, 8(3), 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

Shin, S., Kang, J.-S., & Jo, Y. (2016). The local ensemble transform kalman filter (LETKF) with a global NWP model on the cubed sphere. Pure and Applied Geophysics, 173(7), 2555–2570. https://doi.org/10.1007/s00024-016-1269-0

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., and Wind, P.: The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Phys., 12, 7825–7865, https://doi.org/10.5194/acp-12-7825-2012, 2012.

Timmermans, R., Segers, A., Curier, L., Abida, R., Attié, J. L., el Amraoui, L., Eskes, H., de Haan, J., Kujanpää, J., Lahoz, W., Oude Nijhuis, A., Quesada-Ruiz, S., Ricaud, P., Veefkind, P., & Schaap, M. (2019). Impact of synthetic space-borne NO_2 observations from the Sentinel-4 and Sentinel-5P missions on tropospheric NO_2 analyses. Atmospheric Chemistry and Physics, 19(19), 12811–12833. https://doi.org/10.5194/acp-19-12811-2019

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

Vignati, E., et al. (2004): M7: An efficient size-resolved aerosol microphysics module for large-scale aerosol transport models. J Geophys Res, vol. 109, D22202, doi:10.1029/2003JD004485.

Vittorioso, Francesca and Guidard, Vincent and Fourrié, Nadia. (2024). Assessment of the contribution of the Meteosat Third Generation Infrared Sounder (MTG-IRS) for the characterisation of ozone over Europe. Atmospheric Measurement Techniques, 17(17):5279–5299

Vivanco, M. G., Bessagnet, B., Cuvelier, C., Theobald, M. R., 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

Wind, P., Rolstad Denby, B., and Gauss, M.: Local fractions – a method for the calculation of local source contributions to air pollution, illustrated by examples using the EMEP MSC-W

model (rv4_33), Geosci. Model Dev., 13, 1623–1634, https://doi.org/10.5194/gmd-13-1623-2020, 2020.

xESMF: Universal Regridder for Geospatial Data, https://xesmf.readthedocs.io (accessed 2025.11.06)

Document History

Version	Author(s)	Date	Changes
1.0	Tyler Wizenberg (TNO) , Joanna Struzewska, Jacek W. Kaminski (IOS-PIB), Vincent Guidard (MF)	5 Dec 2025	Initial version
1.1	Tyler Wizenberg (TNO) , Joanna Struzewska, Jacek W. Kaminski (IOS-PIB), Vincent Guidard (MF)	16 Dec 2025	Edited version based on feedback from CAMEO internal reviewers

Internal Review History

Internal Reviewers	Date	Comments
Jacek Kaminski (IOS-PIB) and Anne Caroline Lange (FZJ)	15 Dec 2025	Anne Caroline Lange: well written and very interesting deliverable about well designed OSSEs. Only minor corrections/clarifications needed.