Response to reviewers comments

ACPD manuscript doi: 10.5194/egusphere-2025-5926

Title: 'Improved constraints on ammonia emissions and deposition from co-assimilating NH_3 and NO_2 satellite observations over the Netherlands' by Wizenberg et al.

We would like to thank both reviewers for their helpful comments and suggestions for the manuscript. The reviewer comments are in blue, author responses are in black, and any additions to the text are underlined. The page and line numbers correspond to the version of the manuscript that is available on ACPD.

Responses to reviewer #2

Comment C2.1: I recommend add in the introduction about why the modeled region including Netherland and German is an area of interest to study. Why NH3 emission is important for Netherland?

Reply: We thank the reviewer for this suggestion. We agree that the motivation for focusing on the Netherlands and the surrounding northwestern European region can be stated more clearly in the introduction. We have revised the text in the final paragraph of the Introduction section to explain that the Netherlands is one of the major reactive nitrogen hotspots in Europe due to its high density of intensive agriculture, especially livestock production and fertilizer use, and that adjacent regions of Germany are also important because of transboundary transport and the continuity of agricultural source regions across the border.

Page 3, Line 58 – 61: "In this paper, we perform a co-assimilation of measurements of NH_3 from IASI, CrIS and NO_2 from TROPOMI in the LOTOS-EUROS local ensemble transform Kalman filter (LETKF) over a model domain encompassing the Netherlands and adjacent parts of northwestern Germany. This is a particularly relevant region for studying atmospheric NH_3, as it forms a major reactive nitrogen hot-spot in Europe due to intensive agriculture, especially livestock production and fertilizer use. At the same time, accurate quantification of NH_3 is particularly important for the Netherlands given the ongoing nitrogen crisis and the associated pressures of nitrogen deposition on sensitive ecosystems. Including the neighboring German source regions is also necessary because NH_3 and secondary inorganic nitrogen are influenced by cross-border transport, such that concentrations and deposition over the Netherlands cannot be interpreted from domestic emissions alone. We evaluate resulting optimized emissions and deposition fields, with a focus on NH_3, and we compare the results against independent observations from ground-based measurement networks."

Comment C2.2: Figure 2 shows that the summer emission peak is much lower than the spring peak, even after optimization. In contrast, Lieven Van Damme et al. (2022) reported more comparable peaks between spring and summer. Do you have any interpretation for the cause of the double peaks in the seasonal cycle?

Reply: We thank the reviewer for this observation. We assume that the reviewer is referring to Figure 3 (the time-series of emissions) instead of Figure 2 (the annual emissions and difference maps). The double-peaked seasonal cycle reflects the combination of the prescribed agricultural emission timing in the base inventory (i.e., fertilization events during the spring months) and meteorologically driven variability in NH_3 volatilization (i.e., warmer temperatures during the summer). In the optimized simulation, the assimilation reduces emissions in the early spring and increases them during summer, but the summer peak remains weaker than suggested by independent observational studies (e.g., that of Van Damme et al. (2022)). We therefore interpret the remaining discrepancy as evidence that the current temporal emission parameterization in LOTOS-EUROS still underestimates summertime NH_3 emissions under favorable warm and dry conditions.

In the current version of the manuscript, we discuss our seasonal cycle and the emission changes after assimilation in the context of recent studies, including that by Van Damme et al. (2022), in the text in the paragraph at Lines 347-360. Additionally, in the study by Van Damme et al. (2022), they include a comparison of the seasonal cycle derived from IASI with that from the LML network in the Netherlands. The seasonal cycle they derive from LML appears more consistent with that in our simulations (i.e., a larger springtime peak and a smaller summertime peak). We have added an additional sentence to this section to explain more clearly to the reader what the cause of the double-peaked emission seasonal cycle is:

Page 14, Line 338 – 339: "Figure 3 presents the time series of monthly NH_3 emission totals, aggregated over the same region shown in Figure 2, providing further insight into how the assimilation influences variability across individual months. <u>The base emissions show a similar seasonal cycle in all years, characterized by a pronounced spring peak and a smaller secondary peak in summer. This double-peaked seasonal cycle likely reflects the combination of the prescribed seasonal timing in the agricultural NH_3 emission parameterization and meteorologically driven variability in NH_3 volatilization under warmer conditions.</u>"

Comment C2.3: For the remain discrepancy in the emissions, I was look for analysis like Figures B2 and B4, when reading this part of discussion. However, they appears quite late. These plots demonstrate the temporal performance of the optimized run against independent observations in the concentration field. I feel these figures worth to show in the main manuscript.

Reply: We agree that Appendix Figs. B2 and B4 provide useful supporting information on the temporal behavior of the optimized NH_3 concentration fields relative to independent MAN observations. However, these figures diagnose the modeled concentration response rather than the emissions directly, and we therefore consider them more appropriate in the later section dedicated to ground-based validation. To improve the narrative flow, we have added a sentence with an earlier forward reference to these appendix figures at the end of the concentration discussion (Section 3.3), while retaining the full MAN analysis in Section 3.5.4:

Page 16, Line 407: Additional support for the temporal behavior of the optimized NH_3 concentration fields is provided by the monthly MAN time series shown in Appendix Figs. B2 and B4, although these comparisons are discussed in detail later in Section 3.5.4.

Comment C2.4: In addition, the legends for Figures B2 and B4 introduce two new terms. Does “LE Background Run” correspond to the “base run,” and does “Analysis” refer to the “optimized run”? Please clarify these definitions and ensure consistent terminology throughout.

Reply: We thank the reviewer for noting this inconsistency. The “LE background run” and “Analysis” do indeed refer to the “base” and “optimized” runs, respectively. We have corrected the corresponding legend labels in Figures B2 and B4. Additionally, we have modified the legends of Figures 10 and A1 which also had similar labels that were inconsistent with the intended terminology of “Base” and “Optimized”.

Comment C2.5: Similarly, in Figure 7, the labels begin to use “background” and “analysis.” Please clarify these terms in the manuscript or revise them to maintain consistency.

Reply: This was unintended and we agree with the reviewer. To ensure consistency and clarity throughout the manuscript, we have modified the axis labels on several figures to be consistent with “Base” and “Optimized” terminology. This includes Figures 7, 8, 11, 12, 13, 14, 16, B1 and B3.

Comment C2.6: Finally, maps of emissions by source type (e.g., agriculture, anthropogenic) would help illustrate the spatial distribution of agricultural fields versus anthropogenic sources.

Reply: We thank the reviewer for this suggestion. We agree that source-resolved emission maps would provide useful additional context for interpreting the spatial patterns in the optimized emissions. However, in the present study the LETKF updates the total NH_3 emission field rather than sector-specific emissions, so a robust attribution of the optimized changes to individual source categories is beyond the scope of the current analysis. We have therefore not added new source-type emission maps here, in order to avoid over-interpreting the emission adjustments. We have added a sentence to the discussion/conclusion section to more explicitly highlight that this is a limitation of the current LETKF system and that sector-specific attribution would require an extension of the current framework, for example through a label-based Kalman filtering approach that allows sector-specific emission optimization.

Page 31, line 633: “Because the present LETKF setup optimizes total NH_3 emissions rather than sector-resolved source contributions, the spatial emission adjustments shown here cannot be attributed robustly to individual source types. Further progress may be enabled by adopting a label-based Kalman filtering approach. The labeling functionality introduced in LOTOS-EUROS v2.3 could be extended to the LETKF, allowing sector-specific emission optimization and supporting finer-scale improvements.”

Comment C2.7: Section 2.1.1 Page 4 line 108, you mention time step t_{k-1} and t_k but in equation 2 only used k and k-1.

Reply: We thank the reviewer for pointing this out. We agree that the notation was not fully consistent, in particular since the temporal correlation coefficient (Eq. 1) was defined using the explicit times $t_k$and t_{k-1}, while Eq. 2 used only the corresponding indices. To improve clarity, we have revised the notation in Eq. 2 to use the explicit time notation consistently and we have modified the corresponding text following Eq. 2 that describe the variables.

Comment C2.8: Section 2.1.2 How is the weight of the additional observations applied to observations for spatial localization?

Reply: In the LETKF implementation, all observations within the localization radius are included in the local analysis, but their influence is weighted according to a Gaussian distance-decay function (Eq. 11 on Page 6). In practice, this is implemented by reducing the contribution of more distant observations in the local analysis through the localization weights, so that observations nearest to the analyzed grid cell contribute most strongly. We have clarified this in Section 2.1.2.

Page 6, line 142: "The analysis is then performed using the observations collected, with the weight of the additional observations being limited by a Gaussian decay function:"

was re-written to:

"The analysis is then performed using the collected observations, with the contribution of each observation to the local analysis weighted according to a Gaussian distance-decay function".

We have also added an additional sentence following Eq. 11 to make this clearer and more explicit:

Page 6, line 145: "where Δd is the distance between the observation and the model grid point. <u>As a result, observations closest to the analyzed grid cell have the largest influence, while the contribution of more distant observations decreases smoothly with distance.</u>"

Comment C2.9: Section 2.1.3 Page 6 line 166, you only mentioned "For years after 2019, the 2019 emission totals are used as a baseline but are adjusted dynamically according to meteorological conditions." The analyzed simulation covers 2018 to 2022. What about the emission for 2018? Why is 2019 but not 2018 used emission as baseline?

Reply: The year 2018 was simulated using the corresponding inventory year. For years after 2019, the 2019 emission totals were used as the baseline because 2019 was the most recent year available in the harmonized CAMS GRETA-ER emissions dataset, and these emissions were then adjusted dynamically according to meteorological conditions. We have clarified this in the manuscript.

The sentence starting on page 6, line 166: "For years after 2019, the 2019 emission totals are used as a baseline but are adjusted dynamically according to meteorological conditions."

was replaced for clarity with the following:
"The base emission dataset was compiled using the corresponding inventory year where available; for years after 2019, the 2019 emission totals were used as the baseline because 2019 was the most recent year available in the harmonized CAMS GRETA-ER emissions dataset, and these emissions were then adjusted dynamically according to meteorological conditions."

Comment C2.10: Section 2.1.3 Page 7 line 176 please provide a full name of TNO.

Reply: We have added the full name of TNO at first mention in the revised manuscript.

Comment C2.11: Section 2.2.1 page 7 line 185 "The IASI instruments onboard the MetOp-A, -B, and -C satellites are in Sun-synchronous orbits, passing locations twice daily with Equator crossing times at 09:30 and 21:30 local time, and with a time difference of approximately 45 minutes between them (Clerbaux et al., 2009)." It is unclear what does it mean for "45 minutes between them"? Do you mean since there are three IASI instruments onboard, around each day and night overpass, there are three IASI observations with 45 minutes apart? So each day, there are 6 observations? Maybe explain clearly.

Reply: We agree that this wording was ambiguous. The MetOp platforms carrying IASI share the same nominal local solar time orbit, but they are phased along that orbit and therefore do not acquire measurements at the same location on the ground simultaneously. We have revised the text to clarify that the approximate 45-minute value refers to the temporal separation between the platforms, rather than implying repeated observations of the same ground location. We have adjusted the text in Section 2.2.1 to state this explicitly.

Page 7, lines 184-186: "The IASI instruments onboard the MetOp-A, -B, and -C satellites are in sun-synchronous orbits, ~~passing locations twice daily~~ with nominal Equator crossing times at approximately 09:30 and 21:30 local time~~, and with a time difference of approximately 45 minutes between them~~ (Clerbaux et al. 2009). Although the three platforms fly in the same local-time orbit, they are phased along that orbit and therefore do not acquire measurements simultaneously over the same ground location; the temporal separation between the platforms is on the order of 45 minutes."

Comment C2.12: Section 2.2.1 Page 7 line 195, please add reference for 'HRI'.

Reply: We have added the appropriate references, Van Damme et al. (2017) and Clarisse et al. (2023) to this line when first mentioning the hyperspectral range index (HRI).

Comment C2.13: Section 2.2.2 page 8 line 219, suggestion give a short sentence about what is "a quality_flag of ≥3"

Reply: We agree with the reviewer and have expanded this sentence to clarify that this threshold retains the higher-quality CrIS retrievals recommended for scientific use and excludes lower-confidence retrievals.

Page 8, line 219-220: "This study only includes observations with a quality_flag of ≥3, thereby excluding failed or lower-confidence retrievals, and with a cloud_flag equal to 0 (clear-sky scenes)."

Comment C2.14: Section 2.2.3 Page 8 line 230 you mentioned "horizontal resolution of 1◦×1◦" and described footprint size for IASI and CrIS. Please also mention the TROPOMI footprint size and swath width here as well.

Reply: We have added an additional sentence in Section 2.2.3 to include information on the swath width and nadir pixel sizes for TROPOMI.

Page 8, line 227: "TROPOMI has a swath width of approximately 2600 km, and the NO_2 product has a nadir spatial resolution of 7.2 km in the along-track direction and 3.6 km in the across-track direction, improving to 5.6 x 3.6 km^2 after 6 August 2019 (van Geffen et al. 2022)."

Comment C2.15: Is the VCD product used in this study?

Reply: Yes, the VCD product was used for the assimilation in this study. To make this clearer to the reader, we have added this to the text at the end of Section 2.2.3:

Page 8, line 241: "In this study, we used the VCDs from the reprocessed TROPOMI NO_2 version 2.4.0 dataset."

Comment C2.16: Page 10, "The locations of the sites within the Netherlands are shown on a map in Figure 1." I recommend move this sentence forward to right after Table 1 was mentioned.

Reply: We agree with the reviewer and have moved this sentence to immediately follow the mention of Table 1.

Comment C2.17: Section 3.1, Page 12 line 323, please define "MACC inventory".

Reply: We have defined MACC at first mention in the revised manuscript.

Comment C2.18: Section 3.1 compares emissions from pre- and post-assimilation and called them base and optimized. While in section 3.5.2, you discussed assimilations with choice of combinations of satellites. Please clarify at the beginning of section 3.1 if optimized run refer to the co-assimilations with all three satellite.

Reply: We have inserted a sentence near the beginning of Section 3.1 to clarify that the "optimized" simulation refers to the co-assimilation of IASI, CrIS and TROPOMI unless otherwise stated.

Page 12, lines 316: "Unless otherwise stated, the LETKF-optimized simulation refers to the main co-assimilation run using NH_3 observations from IASI and CrIS together with NO_2 observations from TROPOMI."

Comment C2.19: Does the optimal update of emission in southern part show in Figure 2 attributable to certain source types we can located with the emission maps?

Reply: At present, we cannot robustly attribute the optimized updates to specific source types on the basis of the current LETKF setup, since the assimilation adjusts the total emission field rather than sector-resolved emissions. We therefore avoid making strong source-type attribution claims from Figure 2 alone. Instead, we interpret the pattern more cautiously in relation to known agricultural source regions and previous inventory studies, and we now clarify this limitation more explicitly in the text. See also our response to C2.6 and the additional discussion added there.

Comment C2.20: Page 14 line 343 "As seen in Figure 2, 2020 shows the most significant emissions changes, with particularly large increases in the emissions (on the order of +70%) in the LETKF-optimized simulation between April and September." Figure 2 can't provide evidence of where increase in the emission up to 70%, while Figure 3 indicate optimal emission higher than base from April to September but maximized at 70%. Do you mean Figure 3 here?

Reply: This was indeed an error and the reviewer is correct that the text should have been referring to Figure 3 instead of Figure 2. This has now been corrected.

Comment C2.21: Section 3.2 Page 14 line 365 "… in 2020 and 2021 with differences of +10.4% and +9.6%," I think here you mean "2020 and 2022".

Reply: This was again an error and the reviewer is correct that it should read "…in 2020 and 2022 with differences of +10.4% and +9.6%.,". The text now reads as such.

Comment C2.22: Section 3.4, what's the max DOFS in ideal case for your assimilation setup?

Reply: Upon revisiting this section, we realized that our terminology was not fully consistent with Chen et al. (2023). In the original manuscript, we referred to the mapped quantity as the DOFS, whereas, following Chen et al. (2023) the formal DOFS are given by the trace of the averaging kernel matrix, while the gridded quantity shown in our figure corresponds to the local diagonal element of that matrix, i.e. the averaging kernel sensitivity. In our scalar LETKF formulation, this reduces locally to $A_{ii} = 1 - s_a/s_f$, where s_a and s_f are the local analysis and forecast error variances, respectively. We have therefore revised the terminology throughout this section, the Fig. 6 title, and caption to refer to the plotted quantity as the averaging kernel sensitivity rather than the DOFS. For this local scalar quantity, the theoretical upper bound is 1.

Comment C2.23: Section 3.4, Page 17, line 417: "Regions with high observation coverage" refers to areas with good spatial coverage. Moreover, high sampling density is an additional key factor that enhances the observational constraint.

Reply: We agree that the original wording did not clearly distinguish between broad spatial coverage and local sampling density. We have revised the text to make clear that both the spatial availability of observations and the density of successful retrievals contribute to the observational constraint.

Page 17, lines 417-419: "Regions with ~~high~~ broad observation coverage <u>and high sampling density</u>, particularly over areas with higher NH_3 concentrations, where retrieval sensitivity is generally greater and retrieval uncertainties are lower, exhibit elevated averaging kernel sensitivity values, indicating a stronger observational influence."

Comment C2.24: Page 17, line 427: The statement "a positive relationship is observed" between observation density and DOFS is not clearly supported by Figure 6(a) and (b). The figure does not convincingly demonstrate a direct positive relationship; rather, it only appears that the higher DOFS in 2020 may be primarily associated with increased observation counts in that year.

Reply: We agree with the reviewer that the original wording was too strong. We have revised both the terminology and the interpretation in this section. Specifically, we now refer to the plotted quantity as the averaging kernel sensitivity rather than the DOFS, and we no longer state that a general positive relationship is observed. Instead, we now describe the result more cautiously, noting that the higher mean averaging kernel sensitivity in 2020 coincides with the period of greatest overlap in CrIS and IASI availability, suggesting that increased observation availability contributed to stronger observational constraint in that year, while also emphasizing that the relationship is not strictly linear.

Comment C2.25: Figure 8, if I understand correctly, it is the averaging across all sites from Figure 7 so it reduced the points to 60.

Reply: The interpretation of the reviewer is correct, and this is how it is described in the caption of Figure 8: "Each data-point represents the mean calculated across all LML sites for a given month, and are colored corresponding to the month while the marker style indicates the year."

Comment C2.26: Page 22 Line 491 "scale mismatc" should be "scale mismatch"? Please also give estimate of what resolution of the grid-cell does the LOTOS-EUROS model output represents.

Reply: This was indeed a typo and has been corrected from "scale mismatc" to "scale mismatch". The grid resolution of the LOTOS-EUROS simulation is described in the Model configuration section (Sect. 2.1.3), and we have modified the text on Page 22 to reintroduce this information again since it is relevant to the discussion:

Line 499-500: "Second, the model–observation comparison involves a scale mismatch: the LML instruments measure point concentrations, while the LOTOS-EUROS model output represents grid-cell averages <u>at 7x7 km^2 resolution</u>, which can potentially introduce representativity errors."

Comment C2.27: Does the optimized model run in Section 3.5.2 refer to the co-assimilation of IASI, CrIS, and TROPOMI (NO_2)? If so, has the name of this run changed, or is my understanding incorrect? Please clarify and ensure consistent terminology.

Reply: The focus of Section 3.5.2 is to perform a sensitivity test to highlight the impact of assimilating subsets of the satellite products on the comparisons with the ground-based LML network. Here, the optimized model run in panel (a) of Figure 12 refers to the main co-assimilation run of IASI NH_3, CrIS NH_3, and TROPOMI NO_2 which was the focus of the previous sections of the manuscript, but the other panels show results from assimilation runs performed with subsets of the satellites. The terminology was not fully consistent in the previous version, and we have now revised the text and figure descriptions to make this explicit throughout. See additionally our responses and the discussion associated with C2.4, C2.5 and C2.18.