We thank the reviewer for their valuable and justified comments. All of them have been taken into account during the final revision of the manuscript.

Specific comments:

Reliable quantification of NH_3 remains challenging due to the inlet interactions. However, the NH3 losses in the inlet were not evaluated or calibrated. Though the NH_3 loss in both inlet lines was similar, it would bias the accurate quantification of NH_3 fluxes. The uncertainties induced by the inlet interactions should be evaluated. This can be done by using standard gases or a permeation tube of NH_3.

We thank the reviewer for this important comment. We agree that reliable quantification of NH_3 is challenging because of possible inlet wall interactions, including adsorption, desorption, and memory effects. Although the use of Teflon inlet tubing is intended to minimize wall interactions, NH_3 adsorption/desorption may still occur even on such surfaces and may therefore introduce a potential bias.

In the present study, the calibration of the photoacoustic NH_3 analyzer was performed using the same inlet configuration as during the field measurements, i.e. 4 m Teflon inlet lines, the same flow rate of 800 cm^3 min^{-1}, and the same relevant NH_3 concentration range. Therefore, after equilibration of the inlet walls, any constant recovery loss or systematic inlet-related response is included in the calibration function. In addition, the two inlet lines used for the aerodynamic gradient measurements were identical and operated alternately using the same sampling protocol. Consequently, a constant inlet effect would not substantially affect the measured concentration difference between the two heights. The potentially critical issue is whether the inlet caused a dynamic, time-dependent adsorption/desorption bias during the short alternating measurements.

A separate laboratory recovery experiment using an NH_3 standard gas or a permeation source would indeed be valuable. However, because the instrument was operated continuously in a long-term field campaign, it was not feasible to remove the measurement system from the field for laboratory testing without interrupting the measurements. We therefore performed an additional in-situ test under the same field operating conditions in order to evaluate the magnitude of possible dynamic inlet effects.

This short additional inlet test was carried out between 6 and 9 July 2026. The two inlet lines were placed at the same height, so that both sampled the same ambient air. One inlet was kept unchanged at its original length of 4 m, while the other was shortened to 0.5 m. This represents an eightfold difference in inlet length and wall contact surface and therefore provides a sensitive test of possible inlet-induced adsorption/desorption effects. The measurement protocol was similar to that used during the field campaign: alternating two-minute measurements between the two channels, with one minute of flushing followed by one minute of averaging. At the applied flow rate of 800 cm^3 min^{-1}, the residence time in the original 4 m inlet was approximately 3.8 s, which is short compared with the one-minute flushing period.

During the test, the measured NH_3 mixing ratios ranged from 33 to 75 ppb. In total, 969 two-minute measurements were obtained for each inlet configuration. For the analysis of dynamic inlet effects, we evaluated the change in NH_3 mixing ratio relative to the preceding measurement for the same inlet; this gave 968 concentration-change values for each inlet length. If the longer 4 m inlet had introduced a significant adsorption/desorption bias or memory effect, this should have appeared as a systematic difference in the mean

concentration change and/or as a clearly different variability compared with the shortened 0.5 m inlet.

This was not observed. The mean concentration change was −0.0202 ppb for the 4 m inlet and −0.0182 ppb for the 0.5 m inlet. The difference between the two mean changes was therefore only 0.0020 ppb. This is negligible compared with the observed short-term variability and with the measurement uncertainty reported in the manuscript. The corresponding standard deviations were also very similar: 1.93 ppb for the 4 m inlet and 1.84 ppb for the 0.5 m inlet. Thus, shortening the inlet by a factor of eight did not produce any meaningful change in either the average dynamic response or the variability of the measured NH_3 mixing-ratio changes.

As an additional check, we also examined the distributions of positive and negative NH_3 mixing-ratio changes separately for the 4 m and 0.5 m inlet configurations. This analysis was used to test whether the longer inlet introduced an asymmetric dynamic response, for example through delayed adsorption/desorption or memory effects. If such an inlet-induced bias had been important, the positive and/or negative concentration changes measured with the 4 m inlet would be expected to show a different distribution from those measured with the shortened inlet.

Again, this was not observed. The number of positive and negative changes was nearly identical for the two inlet lengths. For the 4 m inlet, 483 positive and 485 negative changes were obtained, while for the 0.5 m inlet the corresponding numbers were 493 and 475. The positive concentration changes had mean values of 1.460 ppb for the 4 m inlet and 1.381 ppb for the 0.5 m inlet, with standard deviations of 1.174 and 1.158 ppb, respectively. The negative concentration changes were also very similar, with mean values of −1.495 ppb and −1.471 ppb, and standard deviations of 1.319 and 1.173 ppb for the 4 m and 0.5 m inlets, respectively.

Non-parametric distribution tests further supported this conclusion. A two-sample Kolmogorov–Smirnov test showed no significant difference between the full distributions of NH_3 mixing-ratio changes for the two inlet lengths ($p = 0.876$). Similarly, no significant difference was found when the positive and negative changes were tested separately ($p = 0.219$ and $p = 0.305$, respectively). The proportion of positive and negative changes was also not significantly different between the two inlet configurations. Therefore, the observed concentration changes were controlled primarily by natural variability in ambient NH_3 rather than by inlet-induced adsorption/desorption effects.

Together, these results indicate that, under the applied flow rate, residence time, concentration range, and flushing/averaging protocol, the dynamic inlet interaction was below the detectable level of the measurement system and negligible relative to the uncertainty of the reported NH_3 fluxes. We have added this additional field test and its interpretation to the revised manuscript. We also acknowledge in the manuscript that a dedicated laboratory recovery test using an NH_3 standard gas or permeation source would be valuable in future work; however, the calibration procedure used here and the additional in-situ inlet-length test both indicate that inlet-induced bias did not significantly affect the reported NH_3 fluxes.

We have added a short description of this in-situ inlet-length test to the main text, in the section describing the measurement system and uncertainty evaluation. The detailed analysis, including the comparison of mean concentration changes, standard deviations, and the distributions of positive and negative NH_3 mixing-ratio changes, has been added as Appendix B.

Lines 38-40: Given that the system has a high detection limit (~ 3 ppb with 2 min average), it may not be applicable to relatively clean conditions. The authors should emphasize the applicability for "agricultural conditions".

We agree with the reviewer that the present detection limit limits the applicability of the system under clean background conditions. We have therefore modified the abstract (lines 40-42) and added to the conclusions a statement (lines 634-639) noting that measurements under cleaner background conditions may require a lower detection limit or longer averaging times.

„The combined QCL–photoacoustic gradient system demonstrated robust field performance and low instrumental bias under post-fertilization agricultural conditions, where NH_3 mixing ratios were generally above the instrumental detection limit. Thus, the system is particularly suitable for long-term field-scale studies of agricultural NH_3 exchange, whereas its applicability under relatively clean background conditions may be limited by the present detection limit."

Lines 161-162: How long is the sampling line? What is the residence time in the sampling line?

We added: Lines 168-171: "The length of each PTFE sampling line was 4 m. Given the inner diameter of 4 mm and the volumetric flow rate of 800 cm^3 min^{-1}, the internal volume of one sampling line was approximately 50 cm^3, corresponding to a residence time of about 3.8 s."

Lines 177-178: Please elaborate on how the sampling-related measurement uncertainty was derived.

We agree that the derivation of the sampling-related uncertainty needed clearer explanation. So we inserted in lines 187-191:

"During this co-located inlet comparison, both sampling channels were positioned at 300 cm above the soil surface, so that no systematic vertical concentration gradient was expected. The comparison therefore reflects the combined uncertainty associated with the two sampling lines, valve switching, and instrumental response. Results are detailed in Section 3.1."

Lines 307-310: Is the intrinsic system memory induced by the measurement system or soil property? More explanation is needed to help understand the mechanistic presence of system memory.

Thank you for this important comment. We agree that the term "system memory" required further clarification. In the revised manuscript, we now explicitly distinguish between possible instrumental memory and the temporal persistence of the soil–surface exchange system. The observed memory was inferred from autoregressive lag terms in the half-hourly NH_3 flux time series, extending up to six half-hourly steps, i.e. approximately 3 h. This timescale is much longer than the residence time of the sampling line and the instrumental equilibration time. The residence time in the 4 m PTFE sampling line was approximately 3.8 s, and each inlet switch was followed by a 1 min flushing period before the 2 min data acquisition. Therefore, instrumental carry-over or sampling-line memory is unlikely to explain the observed multi-hour persistence.

The "memory" should therefore be interpreted as an apparent temporal persistence of the soil–surface–canopy exchange system rather than as a measurement-system artefact.

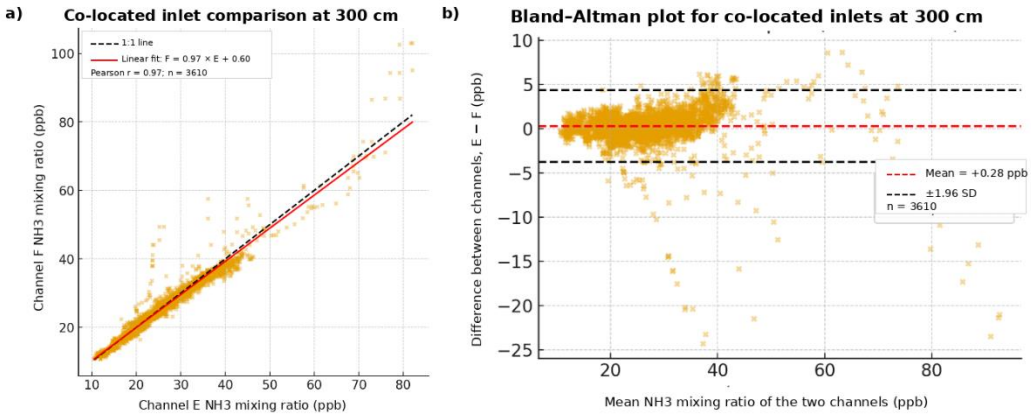
Mechanistically, it likely reflects the gradual evolution of the surface NH_4^+/NH_3 reservoir, soil and surface temperature inertia, moisture-dependent diffusion and adsorption/desorption processes, and persistence in turbulent exchange conditions. We have revised the text accordingly to avoid implying that the lag structure is purely instrumental or purely a fixed soil property.

We inserted into the text in lines 325-332:

"The term "memory" is used here to describe the apparent temporal persistence of the NH_3 flux time series, rather than instrumental carry-over. Since the residence time in the sampling line was only about 3.8 s and each inlet switch was followed by a 1 min flushing period, instrumental memory cannot explain the observed persistence over several half-hourly intervals. The lag dependence is therefore interpreted as a property of the soil–surface–canopy exchange system, reflecting the gradual evolution of the surface NH_4^+/NH_3 reservoir, soil temperature and moisture conditions, adsorption–desorption processes, and persistent turbulence regimes."

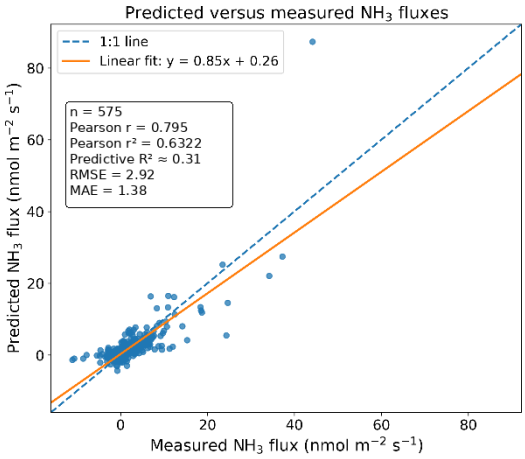
Lines 336-340: This is an important result and should be given as a figure in the main text.

We agree that this is an important result and have therefore added the co-located inlet comparison as Figure 1 in the main text. The figure includes both the scatter comparison and the Bland–Altman analysis, showing the agreement between the two sampling channels at 300 cm. This supports the uncertainty estimate of approximately ±2 ppb derived from the standard deviation of the channel differences.



Lines 477-485: It is better to have figures to show the comparison between predicted and measured NH_3 fluxes.

We agree that a graphical comparison between measured and model-predicted NH_3 fluxes improves the transparency of the machine-learning analysis. We have therefore added Figure 4 to the revised manuscript, showing the predicted-versus-measured comparison with the 1:1 line and linear regression fit. The comparison shows a clear correspondence between measured and predicted values, with Pearson $r = 0.795$ and $r^2 = 0.6322$ for the test subset.



Technical comments:

Line 93: define QCL.
Defined in line 98.