We thank the reviewers for their constructive comments and suggestions to improve the quality and clarity of our manuscript. We have made major and careful modifications to the original manuscript according to all the comments and suggestions from the reviewers.

Item-by-item responses to the specific comments are provided below, in which the reviews' comments are in blue, our responses in **black**, and modifications of the original manuscript are indicated by red in the revised manuscript.

Reviewer #2:

This manuscript presents a potentially valuable study on global and diurnal tropospheric NH3 monitoring using a constellation of polar-orbiting hyperspectral infrared sounders: FY-3E/HIRAS-II, FY-3F/HIRAS-II, and CrIS. The idea of combining complementary local overpass times to achieve quasi-geostationary-like global sampling is interesting and important, especially for NH3, whose short lifetime and strong diurnal variability make sparse temporal sampling a major limitation.

The manuscript shows promising results, including global seasonal maps, regional diurnal cycles over major source regions, and comparison with FY-4B/GIIRS. But I still have some suggestions for the authors to consider.

We thank the referee for the positive evaluation and constructive comments. We have revised the manuscript substantially to improve the descriptions of the methodology and strengthen the robustness of the results. The main revisions include: (1) expanding the retrieval-method section to better document the profile-scaling approach, the selection of the a priori profile, the state vector, and the cross-sensor treatment; (2) adding a priori profile sensitivity analysis; (3) quantifying the impact of quality filters on the number of observations for further analysis; (4) expanding the FY-4B/GIIRS cross-comparison to representative months/seasons with additional statistics, including sample size, intercept, mean bias, and ODR fitting information; and (5) adding comparisons with ground-based FTIR measurements.

Major Comments:

1) The manuscript states that the retrieval is based on the FY-LeoAIR optimal estimation framework and that a profile scaling approach is used, but the actual NH3, retrieval setup is not sufficiently documented in this paper. Here are some of the contents that are suggested to include into the manuscript: the exact state vector elements retrieved jointly with NH3, whether surface temperature and emissivity are simultaneously fitted or fixed, whether the three sensors are spectrally harmonized before inversion, any sensor-specific bias correction or radiometric adjustment. Since this paper’s novelty rests on a multi-sensor integrated retrieval, the retrieval description must be more self-contained. Referring readers to previous papers is not enough.

We thank the referee for pointing out that the NH_3 retrieval setup should be documented more explicitly. We agree that a self-contained description is necessary because the novelty of this work relies on applying a harmonized retrieval framework to FY-3E/HIRAS-II, FY-3F/HIRAS-II, and CrIS.

We have revised Section 2.2 to describe the FY-LeoAIR NH_3 retrieval configuration in detail. The revised text now specifies the state-vector elements retrieved jointly with NH_3, including one multiplicative scaling factor for the NH_3 a priori profile, one multiplicative scaling factor for the H_2O a priori profile, scaling factors for weakly interfering trace gases, surface skin temperature, atmospheric temperature scaling, and four surface-emissivity adjustment coefficients based on first- to fourth-order Legendre polynomial terms. We also summarize the parameters in the retrieval state vector in the supplementary material.

We also clarify that surface skin temperature is initialized from ERA5 and retrieved simultaneously with the gas-scaling parameters. The forward-model emissivity prior is taken from the UOW-M global infrared land surface emissivity database, following Zeng et al. (2023). The a priori emissivity spectrum is adjusted within the retrieval window using four Legendre-polynomial coefficients.

Regarding the multi-sensor treatment, we now clarify that the same 955–975 cm^{-1} retrieval window is used for FY-3E/HIRAS-II, FY-3F/HIRAS-II, and CrIS. HIRAS-II and CrIS have the same nominal unapodized spectral resolution of 0.625 cm^{-1}, allowing the same retrieval window to be used consistently across the three sensors. Sensor-specific spectral noise estimates are used to construct the diagonal measurement-error covariance matrix. No empirical sensor-specific radiometric bias correction or additional spectral harmonization is applied. The same forward model, retrieval window, auxiliary datasets, and quality-control procedure are used for all three sensors to maintain internal consistency.

We further clarify that the fixed NH_3 a priori profile follows Zeng et al. (2023), based on GEOS-CF NH_3 simulations over representative polluted land regions in East Asia and South Asia. The main modification relative to Zeng et al. (2023) is that NH_3 is retrieved as a single multiplicative profile-scaling factor rather than as independent layer-by-layer NH_3 state-vector elements. This approach is adopted because individual thermal infrared spectra generally provide limited vertical information for NH_3 and because it improves computational efficiency for global multi-sensor processing.

2) The main consistency check is a comparison with FY-4B/GIIRS over the Indo-Gangetic Plain and North China Plain. This is useful, but I think it is not sufficient to support the broader conclusion that the constellation can robustly monitor global and diurnal NH3 variability.

We thank the referee for this important comment. We agree that the original consistency evaluation based only on the annual FY-4B/GIIRS comparison over the Indo-Gangetic Plain and North China Plain is not sufficient to support a broad statement about global retrieval robustness. In the revised manuscript, we have strengthened the evaluation in two ways.

First, we add seasonal comparisons between FY-3E/HIRAS-II, FY-3F/HIRAS-II, CrIS, and FY-4B/GIIRS, shown in Fig. A2. The results show that the seasonal slopes mostly fall within the range of 0.7–1.1, supporting broad consistency between the polar-orbiting retrievals and FY-4B/GIIRS.

Second, we add an independent comparison with ground-based FTIR NH_3 column measurements at Hefei, China, shown in Fig. A3. This provides an additional evaluation against ground-based measurements. We acknowledge that this comparison is still subject to representativeness differences and sensitivity differences, but it strengthens the evidence that the constellation retrievals capture the main temporal and spatial variability of NH_3.

We have also revised Section 3.3 and the conclusions to avoid overstatement. The revised manuscript now describes these analyses as consistency evaluations and states that they support the capability of the constellation to capture major NH_3 spatial and temporal patterns.

3) Another major concern is that retrieval sensitivity is explicitly controlled by thermal contrast (TC), and the authors apply thresholds such as TC > 3 K globally and TC > 5 K in hotspot analyses. Since TC itself has a diurnal cycle, the filtering preferentially retains daytime observations and may distort the apparent amplitude and phase of the NH3 diurnal cycle.

We thank the referee for raising this important concern. We agree that TC-based filtering can affect the local-time sampling of the retrievals because TC itself has a strong diurnal cycle. We have revised the manuscript to clarify that the filter is applied to the absolute TC, rather than to positive TC only. The absolute TC > 3 K threshold is used in Section 3.1 for the global spatial-distribution examples, where the aim is to retain sufficient sampling density while excluding retrievals with very weak thermal contrast. For the hotspot diurnal analysis in Section 3.2, we use a stricter threshold of absolute TC > 5 K together with surface AVK > 0.3 to reduce the influence of local-time-dependent retrieval sensitivity on the inferred diurnal variations.

To directly evaluate whether the absolute TC and surface AVK filtering distorts the retrieved diurnal cycles, we add a sensitivity analysis using three filtering configurations: surface AVK > 0.1 and absolute TC > 3 K, surface AVK > 0.3 and absolute TC > 5 K, and surface AVK > 0.4 and absolute TC > 7 K. We compare the resulting spatial distributions over the North China Plain and Indo-Gangetic Plain, the regional diurnal cycles, and the total number of retained retrievals for each satellite. These results are provided in the supplementary material Figs. S7-S10.

The sensitivity test shows that the number of retained retrievals decreases as the filters become stricter, but the spatial distributions remain broadly similar and the main diurnal patterns are not obviously distorted by the choice of absolute TC and surface AVK threshold. We now state that the exact column magnitude and sampling density remain filter-dependent, but the principal diurnal features discussed in the manuscript are not primarily an artifact of the TC-based filtering.

Minor Comments:

1) Section 3 is titled “Results and Discussions,” while Section 4 is said to be “the discussions” in the introduction, but there is no standalone Section 4 in the provided text.

We thank the referee for pointing out this inconsistency. We have revised the section description in the Introduction to match the actual manuscript structure. Section 3 now presents and discusses the results, and the conclusions are given in Section 4.

2) The title and text alternate between “tropospheric ammonia,” “NH3 column,” and “total column.” Since the retrieval is limited to 11 layers from the surface to 200 hPa, please define precisely what column is being reported and use consistent terminology.

We thank the referee for pointing out the need for a precise definition of the reported column quantity. We have revised the manuscript to clarify that the retrieval state vector adjusts NH_3 in 11 layers from the surface to 200 hPa. The reported NH_3 total column is then calculated as the sum of the retrieved surface-to-200 hPa column and the fixed a priori contribution above 200 hPa. Since the NH_3 abundance above 200 hPa is generally small and is not independently constrained by the spectra, the spatial and temporal variability discussed in this study is primarily determined by the retrieved surface-to-200 hPa component. We now define this explicitly in Section 2.2 and use “NH_3 column” consistently throughout the manuscript to refer to this reported total column.

3) A 2.5°×2.5° box is fairly large, especially in regions with heterogeneous sources. Please explain why this spatial extent was selected and how sensitive results are to box size.

We thank the referee for this comment. We have added a discussion explaining that the 2.5° × 2.5° box is chosen to balance spatial representativeness and sample size. Because the analysis is separated by month and overpass time and uses strict quality filters, smaller boxes can result in too few valid retrievals, especially for nighttime observations. We also acknowledge that the selected box size may smooth sub-regional source heterogeneity, and we identify box-size sensitivity as a limitation to be examined in future work.

4）The manuscript states that the emissivity database is “Monthly Global 0.05° V003” and then says it provides data “at a 0.5° spatial resolution.” This should be checked.

We thank the referee for noting this inconsistency. We have corrected the description of the emissivity product. The Combined ASTER and MODIS Emissivity over Land Database Monthly Global V003 product has a spatial resolution of 0.05°.

5）Section 3.3 title says “Comparisons with geostationary NH3 observations,” but the text uses “validate.” I suggest replacing “validate” with “evaluate consistency” or “cross-compare.”

We agree with the referee that FY-4B/GIIRS provides an independent satellite retrieval product rather than an absolute validation reference. We have therefore replaced “validate” with “evaluate consistency” or “cross-compare” throughout Section 3.3 and revised the section title to “Cross-comparison with geostationary FY-4B/GIIRS and ground-based FTIR NH_3 observations”

6）The collocation uses <0.5° in latitude/longitude and <0.5 h in time. Given the gradients and diurnal variability of NH3, please justify these thresholds and discuss representativeness mismatch.

We thank the referee for this suggestion. We have added the collocation criteria to the main text and explained the rationale for the selected thresholds. A polar-orbiting retrieval and a FY-4B/GIIRS retrieval are considered collocated when their longitude and latitude differences were both less than 0.5° and their observation time difference is less than 0.5 h. These criteria are selected as a compromise between retaining enough matched samples and reducing representativeness differences.

7) The manuscript uses both “PBL” and “BLH.” Please make it consistent.

We thank the referee for pointing this out. We have standardized the terminology throughout the manuscript. The term “boundary layer height (BLH)” is defined at its first occurrence, and “BLH” is used consistently thereafter.