Responses to referee 1:

We would like to thank the referee for the useful comments and constructive suggestions. In the following, we address the referee's comments and describe the corresponding changes we have made to the manuscript. The referee's comments are listed in *italics*, followed by our response in blue. New/modified text in the manuscript is in **bold**.

This paper discusses the design of both airborne and spaceborne instruments covering both the SWIR and TIR wavelength regions, and then discusses the achievable precision of N_2O observations based on detailed instrument specifications and retrieval conditions. The paper presents basic retrieval performances, such as measurement sensitivities to surface-layer's concentrations, conducts a quantitative assessment of the conditions required to detect actual N_2O variations in the atmosphere, and discusses the detection limits of surface N_2O emissions varying depending on the emission amounts as well as practical methods for spatially aggregation of observed data. I was impressed by the comprehensiveness of this research. However, as the paper contains many equations and conditions/assumptions to elucidate overall measurement performances, it is sometimes hard to understand because of lack of explanations. The paper is suitable for the purpose of AMT journal and I recommend it to be published after adding necessary explanations and considering some revisions.

We are grateful for the reviewer's insights. We are especially encouraged by the reviewer's recognition of the comprehensiveness of the study. Below, we provide point-by-point responses and detail the corresponding changes made to the manuscript.

Major comments.

1) A priori covariance matrix for N_2O retrieval: *Judging from Figure 4(a), the N_2O standard deviation profile seems to be overestimated compared to realistic variabilities in N_2O concentrations in the atmosphere. If a priori covariance matrix is set by combinations with N_2O standard deviation profile (Panel a) and its error correlation (Panel b), a priori error would be overestimated even if smaller γ values are applied, and consequently, measurement information introduced to retrieval results would be also overrated.*

We thank the reviewer for this important comment. We adopted the MethaneAIR CH_4 a priori error covariance matrix for N_2O because no mature N_2O covariance structure is currently available. We scaled down the established CH_4 standard deviation profile by using the atmospheric abundance ratio between CH_4 and N_2O (1900 ppb/330 ppb=5.76), while retaining the same correlation structure because both species are long-lived trace gases, relatively well-mixed in the troposphere and exhibit decreasing concentrations in the stratosphere. The intention behind adopting the scaled MethaneAIR a priori covariance was not to represent the exact climatological variability of atmospheric N_2O, but rather to provide a reasonable regularization suitable for evaluating instrument performance and retrieval sensitivity under controlled assumptions. We agree that the standard deviation profile shown in Fig. 4a appears to be larger than the realistic variability of atmospheric N_2O. However, using a realistic but small prior error causes the retrieval to become strongly prior dominated, resulting in

weak averaging kernel sensitivity and limited observational information content. As shown in Fig. 7(a,c) a very small prior covariance indeed reduces the measurement error but simultaneously causes the averaging kernel to decrease substantially, indicating a significant loss of meaningful information from the observations and stronger prior dominance in the retrieval. The γ factor was introduced to investigate this tradeoff and $\gamma = 0.5$ was selected as a practical balance between a priori and observational constraint. To implement this response, we have modified the lines 212-216 in the manuscript as follows:

"Since no mature operational N_2O covariance structure is currently available, we adopt the MethaneAIR CH_4 a priori error covariance matrix (Chan Miller et al., 2024), which originates from the GOSAT CH_4 algorithm developed by the University of Leicester (Parker et al., 2020), for both N_2O and CH_4 profiles. Similar to Chan Miller et al. (2024), we decompose the error covariance matrix into a standard deviation profile and an error correlation matrix. The N_2O standard deviation profile, shown in Fig. 4a, is scaled down from the original CH_4 standard deviation profile by a factor of 5.76 based on the ratio between typical atmospheric abundance of CH_4 (∼1900 ppb) and N_2O (∼330 ppb), reflecting the relatively lower abundance of N_2O. The purpose of adopting this scaled CH_4 prior covariance structure here is not to reproduce the exact climatological variability of atmospheric N_2O, which is expected to be smaller than the prior standard deviation shown in Fig. 4a. Instead, the intention was to provide a reasonable regularization for evaluating retrieval sensitivity to the a priori constraint. Using the realistic N_2O variability as prior error would strongly constrain the retrieval toward the prior state, substantially reducing averaging kernel sensitivity and limiting the information from observations. Section 4.1 further evaluates this tradeoff by continuously scaling the prior strength."

2) Hardware configurations for TIR band: *The spaceborne instrument design for TIR band is presented in Table 1. The wavelength resolution (spectral sampling) is set to 0.25 nm at around 8 μm N_2O absorption lines, which is rather high compared to current spaceborne-sensors in operation. Is such high-spectral resolution measurement feasible?*

We thank the reviewer for raising question regarding an important parameter of TIR instrument. We would like to clarify that the value of 0.25 nm reported in Table 1 refers to the spectral sampling interval rather than the effective spectral resolution of the spaceborne TIR sensor. In our configuration, the instrument spectral response function (ISRF) is assumed to have a full width at half maximum (FWHM) spanning three sampling intervals (Table 1), corresponding to a spectral resolution of approximately 0.75 nm, or equivalently about 0.12 cm^{-1} near 8 μm. Therefore, the actual resolving capability of the instrument is substantially coarser than the sampling interval itself.

Although this spectral resolution is relatively high for spaceborne TIR measurements, it remains within the range achieved or proposed for trace-gas remote sensing instruments. For example, the TES instrument operated with a spectral resolution of approximately 0.1 cm^{-1}

in the thermal infrared (Bowman et al., 2006), which is comparable to our effective FWHM resolution in this study. Similarly, the TANSO-FTS-2 instrument onboard GOSAT-2 operates at spectral resolutions of approximately 0.2 cm^{-1} (Nakajima et al., 2012), while the proposed MIN2OS mission concept considered a TIR spectral resolution of approximately 0.25 cm^{-1} (Ricaud et al., 2021). These examples demonstrate that high spectral resolution thermal infrared measurements are technically feasible, particularly for instruments specifically optimized for trace-gas observations. In addition, a preliminary optical design assessment has been performed for the proposed TIR instrument, including the grating requirements and optical train. The grating is the component most strongly affected by the high spectral sampling, and a manufacturability assessment indicates high confidence that such a grating can be produced. We have added information about spectral resolution in line 228 as follows:

"The key instrument design parameters, including spectral coverage and spectral resolution, were selected through iterative assessment of N_2O absorption strength and instrument performance using linear sensitivity analysis in concert with the current technology and industry standards for both airborne and spaceborne instruments. A Gaussian instrument spectral response function (ISRF) is assumed with a full width at half maximum (FWHM) spanning three spectral sampling intervals ($d\lambda$). This corresponds to spectral resolutions of approximately 0.1725 nm (0.33 cm^{-1}) in the SWIR band for both airborne and spaceborne instruments. In the TIR band, the spectral resolutions are 0.90 nm (0.14 cm^{-1}) and 0.75 nm (0.12 cm^{-1}) for the airborne and spaceborne instruments, respectively. These spectral resolutions are within the range of existing and proposed trace-gas remote sensing instruments operating in the SWIR and TIR spectral regions (Bowman et al., 2006; Ricaud et al., 2021; Nakajima et al., 2012). In addition, a preliminary optical design assessment has been performed for the proposed TIR instrument, including the grating requirements and optical train. The grating is the component most strongly affected by the high spectral sampling, and a manufacturability assessment indicates high confidence that such a grating can be produced."

3) Discussion on X_{N_2O} measurement errors: *If measurement errors shown in Figure 7 are calculated on the basis of Eq. 6, they are affected variabilities and sensitivities of H_2O, CH_4, and temperature (their Jacobians and a priori covariances) through the Gain matrix, unless spectral channels with purely N_2O absorption are selected. In addition, does a γ factor that is introduced to modulate N_2O a priori covariance matrix also affect the a priori covariance matrices of these other state vector parameters?*

We thank the reviewer for this insightful comment. Yes, the X_{N_2O} measurement errors shown in Fig. 7 are calculated based on Eq. 6 and are influenced by the variabilities and sensitivities of the other retrieved state vector elements through the gain matrix. We evaluated the impact of these non-N_2O state vector elements on the X_{N_2O} retrieval using the interference error

equation (Connor et al., 2016) given below:

$$\hat{\mathbf{S}}_\mathbf{i} = \mathbf{A}_{\mathbf{N_2O-nonN_2O}} \mathbf{S}_{\mathbf{nonN_2O}} \mathbf{A}_{\mathbf{N_2O-nonN_2O}}{}^T. \quad (1)$$

where $\mathbf{S}_{\mathbf{nonN_2O}}$ is the ensemble covariance matrix for non N_2O elements of the state vector (approximated by prior covariance matrix) and $\mathbf{A}_{\mathbf{N_2O-nonN_2O}}$ is the off-diagonal block of the averaging kernel matrix relating non N_2O elements to the retrieved N_2O profile. The resulting interference errors from non-N_2O state vector elements were found to be small compared to the dominant N_2O retrieval uncertainty. The interference errors associated with CH_4, H_2O, temperature and surface temperature are found to be approximately 0.15, 0.40, 0.35 and 0.0001 ppb, respectively. To implement this point in the manuscript we did the following additions in line 194:

"This measurement error, or precision of X_{N_2O} is influenced by interference errors associated with the variabilities and sensitivities of the other retrieved state vector elements through the gain matrix (Connor et al., 2016). However, the interference errors from the non-N_2O state vector elements are at least one order of magnitude smaller than the N_2O measurement error, indicating that their contributions are minimal compared to the dominant N_2O measurement uncertainty."

Regarding the γ factor, we clarify that γ is applied only to the N_2O prior covariance matrix to investigate the sensitivity of the N_2O retrieval to the strength of the a priori regularization. The prior covariance matrices for H_2O and temperature are adopted from the CrIS Level 2 product, and they are not scaled by γ. Similarly, the CH_4 prior covariance is kept fixed during the γ-sensitivity analysis. Therefore, the γ-scaling shown in Fig. 7 only modulates the N_2O a priori covariance, while the prior constraint for all other state vector elements remain unchanged. We have made the following addition to line 358 of the manuscript to clarify this point.

"The applied γ only modulates N_2O a priori covariance matrix and does not affect the a priori covariance matrices of all other state vector elements."

4) X_{N_2O} measurement errors for SWIR band: *Although X_{N_2O} measurement error reductions with respect to a factor γ are understandable for TIR band, the SWIR case has a quite different characteristics both for airborne and spaceborne simulations; the SWIR measurement errors drop abruptly at a certain value of γ. It should be explained in more detail.*

We thank the reviewer for this careful observation. The abrupt decrease in the SWIR-only X_{N_2O} measurement error occurs when the retrieval begins to strongly "feel" the effect of the a priori constraint. At larger values of γ, the retrieval remains relatively insensitive to the prior. As γ keep decreasing, the prior constraint becomes sufficiently strong that it may rapidly suppress the influence of the observational constraint. Consequently, the retrieved X_{N_2O} measurement error decreases abruptly. This transition is synchronized with the simultaneous

reduction in the near-surface sensitivity shown in Fig. 7(b,d). This transition behavior is different in SWIR and TIR cases likely due to inherent properties of shortwave and longwave infrared radiative transfer. Similar abrupt behavior associated with prior covariance tuning has also been reported for SWIR CH_4 retrieval in Chan Miller et al. (2024). We have added the following to line 370 in the manuscript.

"For the SWIR-only case, this transition to a prior-dominated regime is particularly abrupt, resulting in a rapid reduction in the X_{N_2O} measurement error that is synchronized with the simultaneous decrease in near-surface sensitivity. This behavior is also consistent with prior covariance tuning behavior reported for SWIR CH_4 retrievals in Chan Miller et al. (2024) and likely reflects the inherent differences between shortwave and longwave radiative transfer, including both the strength and vertical distribution of the N_2O Jacobians."

5) Calculation of Eq. 17 and Figure 10: *Two different X_{N_2O} variabilities, for spaceborne and airborne cases, are calculated by using Eq. 17? In the calculation, which parameters come from the designed instruments and from MAIZE observations are not clear, so the authors should explain in more detain in the section 3.3.*

Yes, two different variabilities are calculated for spaceborne and airborne case using Eq. 17. The primary difference between both X_{N_2O} variabilities calculated using Eq. 17 arises from the observational height h_t. Due to different h_t, airborne and spaceborne instrument sample different portions of the atmospheric column. Most of the N_2O enhancement and spatial variability associated with surface emissions occurs within the planetary boundary layer (PBL). An airborne instrument observes the atmosphere from much closer to the PBL and therefore measures a shorter atmospheric column that is more strongly influenced by these enhancements. As a result, the observed column variability is relatively large. In contrast, a spaceborne instrument observes the entire atmospheric column from orbit. Although the same N_2O enhancements are still present, they are diluted by the much larger background atmospheric column above the PBL. Consequently, the resulting column-averaged variability becomes smaller for the spaceborne.

In Eq. 17, the PBL height h_b and the N_2O mixing ratio (μ_{N_2O}) are taken from the MAIZE observations. The observational height h_t is from instrument design and differs between the airborne and spaceborne. The atmospheric scale height H is assumed to be 7.5 km. We have done the following revision in line 307:

"h_t is the observational height determined by the instrument design parameters listed in Table 1."

The sentence in line 313-314 is revised as follows:

"Here, the semivariogram is directly calculated from extensively measured N_2O mixing ratios (μ_{N_2O}) along horizontal flight legs within the well-mixed PBL during the MAIZE campaign, and the PBL height is inferred from collocated spiral profiles by identifying the sharp vertical transition of trace gases, temperature,

and relative humidity between the well-mixed boundary layer and the free troposphere (Zhang et al., 2020). The X_{N_2O} variabilities calculated using Eq 17 differ between the airborne and spaceborne instruments because they sample different portions of the atmospheric column due to their different observational height (h_t).”

Minor comments.

1) Introduction, page 2, line 42 and reference Waldmann et al., 2026. This paper has been already accepted and published.

We thank the reviewer for noting this. We have updated the Waldmann et al. reference to cite the published article instead of the earlier EGUsphere discussion preprint.

2) Introduction, page 3, line 57 It may be better to write X_{CO_2} and X_{CH_4} in this context.

We thank the reviewer for pointing it out. We agree that X_{CO_2} and X_{CH_4} are more appropriate and consistent with the standard terminology used in remote sensing domain. We have revised the text as suggested.

3) 2. Data, Figure 2b How can we see the values in the figure? It has two different vertical axes with different numbers [ppb].

We appreciate the reviewer's careful observation. In Fig. 2b, both vertical axes represent the same quantity, namely in situ N_2O mixing ratio. The dual-axis labeling is used to improve visual clarity by reducing overlap among the time series from different flight days. The values shown on the left and right axes correspond to alternating flight days and have the same units, ppb. We have added the following in line 140 to explicitly clarify this presentation.

“Figure 2b has two vertical axes, with the left and right axes showing in situ N_2O mixing ratio in ppb for alternating flights to improve readability.”

We also added this information in Figure 2 caption:

“Both vertical axes represent the in situ N_2O mixing ratio in ppb. The left and right axes are used for alternating flight days to reduce overlap and improve readability.”

4) 3. Measurement Methodology, page 8, lines 190-191 “.. so the weighting vector h . . . and zero otherwise.” What does “zero” mean? When the elements of the weighing vector are zero, there is completely no N_2O molecule?

We thank reviewer for the important question which requires some clarity. The state vector $\mathbf{x}$ contains multiple retrieved quantities, including 19-layer profiles of N_2O, CH_4, H_2O, and temperature, together with a scalar surface temperature term. Therefore, the full state vector contains $19 \times 4 + 1 = 77$ elements, and the weighting vector $\mathbf{h}$ has the same dimension.

In the definition of X_{N_2O}, only the N_2O profile elements are intended to contribute to the column-average calculation. The phrase “zero otherwise” does not mean that the abundance

of N_2O or other molecules is zero. Rather, it means that the weighting coefficients corresponding to all non-N_2O elements of the state vector are set to zero in the weighting vector **h**. Consequently, the CH_4, H_2O, temperature profile, and surface temperature elements do not contribute to the calculation of X_{N_2O}, and only the 19 N_2O profile layers are included in the weighted column average. We have revised the line 190 to clarify this point as following:

"Currently, we assume X_{N_2O} only depends on the retrieved N_2O mixing ratio profile, so the weighting vector h is the fractional dry air column at the corresponding layers of the N_2O profile and zero for all other elements of the state vector, such that only N_2O profile contributes to the calculation of X_{N_2O}."

5) 3. Measurement Methodology, page 9, lines 215-216 "The N_2O standard deviation profile, . . . , by a factor of 5.76, reflecting the lower atmospheric abundance of N_2O." Related question to one of the main comments; how do you determine such specific number, 5.76, for a factor for the N_2O standard deviation profile.

The scaling factor of 5.76 is derived from the ratio of typical atmospheric abundances of CH_4 and N_2O and is used as a first-order approximation to adapt the CH_4 variability profile for N_2O. In the absence of well-constrained N_2O prior covariance models for retrieval studies, scaling the CH_4 standard deviation profile based on relative abundance of gases provides a practical approach for constructing the prior regularization. CH_4 is used as the reference gas because it is also a long-lived and relatively well-mixed trace gas, and its prior covariance structure is already well characterized and commonly used in atmospheric retrieval applications. We have revised the line 215 in the manuscript to justify the assumed scaling factor. This revision is a part of major comment 1 as well.

"The N_2O standard deviation profile, shown in Fig. 4a, is scaled down from the original CH_4 standard deviation profile by a factor of 5.76 based on the ratio between typical atmospheric abundance of CH_4 (∼1900 ppb) and N_2O (∼330 ppb)."

6) 3. Measurement Methodology, Table 1 Why is wavelength range for TIR band different between airborne and spaceborne instruments? Just a limitation of the hardware configurations?

We thank the reviewer for this thoughtful question. The difference in the TIR wavelength ranges between the airborne and spaceborne instruments is primarily constrained by the size of focal plane selected and the sensor cut-off wavelength. The airborne instrument is designed with a smaller focal plane array and therefore requires reduced instrument size and power consumption. We have added the following information in line 230 of the manuscript for clarity.

"Due to the substantially smaller TIR focal plane array assumed for the airborne instrument relative to the spaceborne, a narrower TIR wavelength range is selected for the airborne case while maintaining the desired spectral performance."

7) 3. Measurement Methodology, page 14, lines 293-295 (Eq. 14) Can we simply assume the measurement error linearly decrease with respect to the distance that is adopted for data aggregation?

We thank the reviewer for this important question. The scaling relationship in Eq. 14 is applied specifically to the measurement error which is a linear projection of detector noise. The detector noise between independent soundings is assumed to be uncorrelated, such that spatial aggregation reduces the effective measurement error with increasing aggregation scale. We agree that other sources of retrieval uncertainty may exhibit spatial correlations and therefore may not decrease at the same rate with aggregation distance. However, those error components are not considered in the scope of present study. This work focuses on the propagation of measurement noise under the assumption of independent soundings. We have revised the line 293-294 accordingly.

"When spatially aggregating X_{N_2O} observations, the measurement error component propagated linearly from detector noise decreases with the target length scale d. This relationship assumes that the detector noise between individual soundings is independent and uncorrelated."

8) 3. Measurement Methodology, page 15, line 308 "... $\Delta_d \mu_{N_2O}$ is the spatial differentiation of PBL N_2O mixing ratio, denoted as μ_{N_2O}." How to calculate PBL N_2O mixing ratio from aircraft data? Just the average of all data under PBL height?

The PBL N_2O mixing ratio, μ_{N_2O}, is directly taken from the in situ measured mixing ratios during horizontal flights in the PBL. We have revised line 308 to incorporate response as follows:

"$\Delta_d \mu_{N_2O}$ is the spatial differentiation of PBL N_2O mixing ratio between observation pairs separated by distance d, where μ_{N_2O} represents the in situ measurements collected along horizontal flight legs within the well-mixed PBL."

9) 3. Measurement Methodology, page 15, lines 313-314 I think accurate PBL height determination is a key for the analysis of this paper. It is better to explain how to estimate PBL height from aircraft data.

We agree that accurate PBL height determination is important for this analysis. The PBL height is estimated from the aircraft spiral profiles by identifying the sharp vertical gradient of trace gases, temperature and relative humidity that separates the well-mixed boundary layer from the free troposphere.We have revised line 313-314 to include this explanation as follows:

"Here, the semivariogram is directly calculated from extensively measured N_2O mixing ratios (μ_{N_2O}) along horizontal flight legs within the well-mixed PBL during the MAIZE campaign, and the PBL height is inferred from collocated spiral profiles by identifying the sharp vertical transition of trace gases, temperature, and relative humidity between the well-mixed boundary layer and the free tro-

posphere (Zhang et al., 2020).”

10) 3.3 Detectability of N_2O emissions at different spatial scales. If I understand correctly, overall, this section seems to mix discussions/explanations of horizontal and vertical distances. It might be better to clearly distinguish between the two.

In Section 3.3, the horizontal and vertical dimensions play different roles in the detectability analysis. The spatial scale d represents the horizontal aggregation length scale over which emissions accumulate and observations are averaged. The only vertical parameter used in this section is the observation altitude h_t, which determines the vertical extent of the atmospheric column sampled by the instrument. Since the airborne and spaceborne instruments observe different vertical atmospheric columns due to their different observation altitudes, the resulting column amount errors and detectability characteristics differ between the two configurations. We have made the following chnages to the manuscript:

Addition to line 334:

“...., g is gravity and H is the assumed atmospheric scale height.”

Addition to line 336:

“h_t is the only vertical parameter in this formulation and is used to determine the vertical extent of the atmospheric column observed by the instrument.”

11) 4. Results, page 16, line 356. Why and how do you set a factor γ to 0.03 to 10?

The range of γ from 0.03 to 10 was intentionally selected to span a broad range of retrieval regimes, from strongly prior-dominated cases at small γ values to observation-driven cases at large γ values. The sensitivity analysis was performed by systematically looping over logarithmically spaced γ values across this range to examine the transition between these retrieval regimes. We have revised line 355-357 to mention this reason.

“In Figure 7 the a priori constraint is adjusted by scaling the GOSAT/MethaneSAT-based N_2O prior standard deviation profile (see Fig. 4a) by a factor γ. The sensitivity analysis is performed using logarithmically spaced γ values between 0.03 and 10, spanning retrieval regimes from strong domination by the prior regularization (small γ) to those controlled primarily by the observation constraint (large γ).”

12) 4. Results, Figure 7 (a) and (c). X_{N_2O} measurement error shown in the figure is the error defined as Eqs. 6 and 14? If so, the estimated measurement error may be underestimated if the assumption of errors decreasing linearly with average distance is not valid. It is better to show also measurement errors of a “single-shot” observation for comparison.

The X_{N_2O} measurement errors shown in Fig. 7 are solely derived from the linear sensitivity analysis using Eq. 6 and correspond to single-sounding measurement errors at the native instrument footprint size d_0 of 20 m for airborne and 0.7 km for spaceborne instruments. They do not involve the spatial aggregation assumption described in Eq. 14, which is implemented

in semivariogram analysis presented in Section 4.2. Therefore, the results in Fig. 7 already represent the "single-shot" observation values at native footprint size. To clarify this point we have done the following changes in the manuscript.

Addition to line 352:

"The measurement error is calculated using Eq. 6 and vertical sensitivity is calculated using Eq. 8."

Revised line 358-359:

"The results presented in Fig. 7 are the mean values calculated separately using the environmental conditions at 100 CrIS sounding locations shown in Fig. 1. These results correspond to single-shot observations at native instrument footprint size ($d_0 = 20$ m for airborne and $d_0 = 0.7$ km for spaceborne)."

13) 4. Results, page 17, lines 364-365 "In contrast, the SWIR and TIR bands alone have fewer observations, . . . " What does this sentence mean?

The phrase "fewer observations" refers to the reduced spectral information content available when using only the SWIR or TIR band compared to the joint SWIR–TIR setting. The combined retrieval utilizes spectral measurements from both bands, providing stronger observational constraints and improved retrieval sensitivity. We have revised lines 364-365 in the manuscript to clarify this point.

"In contrast, the SWIR and TIR bands alone have fewer spectrum data points and thus provide weaker observation constraints than the joint band case, leading to larger errors."

14) 4. Results, Figure 7 (b) and (d) The values are calculated by Eq. 8? Judging from the figures, TIR band has more sensitivity to near-surface layer (atmospheric 1st layer?) than I expected. In larger γ cases, SWIR sensitivity to near-surface is lowest and the combination one is highest (Panel b). The author should explain more about this.

Yes, the near-surface sensitivity shown in Fig. 7(b,d) is calculated using Eq. 8 and is defined here as the mean averaging kernel value of the lowest two atmospheric layers rather than only the first atmospheric layer. These lowest two layers correspond approximately to pressure levels of 962 and 901 hPa, representing the lowest $\sim$ 1 km of the atmosphere. To implement it we have done the following changes.

Addition to line 352:

"The measurement error is calculated using Eq. 6 and vertical sensitivity is calculated using Eq. 8."

Addition to line 355:

"The near-surface sensitivity here is defined as the mean averaging kernel value of the lowest two atmospheric layers, representing the lowest $\sim$ 1 km of the

atmosphere (corresponding pressure values are 962 and 901 hPa)."

The relatively strong TIR near-surface sensitivity at large γ values occurs because the retrieval is weakly regularized by the prior in this regime, such that the observational constraint dominates, allowing the TIR observations to retain stronger sensitivity to the lower atmosphere. In contrast, the SWIR-only retrieval provides weaker observational constraints and therefore exhibits lower near-surface sensitivity. The combined SWIR–TIR retrieval benefits from the complementary information content of both spectral regions and consequently shows the highest near-surface sensitivity. As γ decreases, the increasing prior regularization gradually transitions the retrieval from an observation-dominated regime towards a more balanced contribution between the prior and the observations. The following sentence is added to line 371:

"At a very large γ (rightmost side) the retrieval is primarily observation dominated, resulting in relatively strong TIR near-surface sensitivity due to the stronger TIR observational constraint. Under this regime, the SWIR-only retrieval exhibits the weakest near-surface sensitivity, while the joint SWIR–TIR retrieval benefits from complementary information from both bands and therefore shows the highest sensitivity."

15) 4. Results, Figure 10 It is better to explain a bit more what each of the circular makers indicates.

We thank the reviewer for the suggestion. The circular markers indicate the distance bins used in the semivariogram calculation. Each circular marker represents the calculated X_{N_2O} variability for a given horizontal separation distance between observation pairs. The semivariograms are calculated over spatial scales ranging from 0 to 50 km using a bin width of 0.25 km. We have add the following explanation to line 426.

"The circular markers represent the distance bins used in the semivariogram calculation."

We also added this information to the Figure 10 caption.

"Solid lines with circular markers show X_{N_2O} variability derived from the MAIZE flights, where the circular markers indicate the semivariogram distance bins separated by 0.25 km."

16) 4. Results, Figure 11 It is better to briefly touch on "white area" in the figure. Maybe due to a limitation of reliable calculations?

We thank the reviewer for this helpful suggestion. The white regions in Fig. 11 do not indicate unreliable calculations. They correspond to cases where the calculated detectability metric $q(E, d)$ falls below the minimum contour level displayed in the figure ($q<2^{-8}$). Therefore, the white areas simply reflect the selected plotting cutoff used for visualization. We have mentioned this point explicitly in the manuscript.

Addition to line 456:

"White regions in Fig. 11 correspond to the detectability values below the minimum plotted contour level ($q<2^{-8}$)."

Addition to Fig 11 caption:

"White regions indicate values below the minimum detectability threshold displayed in the plot ($q<2^{-8}$)."

17) 4. Results, page 23, lines 478-481 "Using the mean semivariogram ... than using the semivariogram of highest-variability flight (2022-05-27) ..." It may be better to add information such as "week wind and less mixing condition" for this flight.

We thank the reviewer for this helpful suggestion. We examined the meteorological conditions during the MAIZE flights and found that the 20220527 flight was characterized by relatively weak wind and reduced atmospheric mixing conditions. These conditions likely limited horizontal dispersion and preserved localized PBL N_2O enhancements, resulting in stronger spatial heterogeneity and larger semivariogram variability for this flight. We have revised the line 480 in the manuscript to accommodate this information.

"..., which represents a more favorable scenario associated with relatively weak wind and reduced atmospheric mixing conditions."

18) 5. Discussion and Conclusions
This chapter is more suitable as "Summary and Conclusions"?

We have considered and implemented the suggested change.

Line 489:

"Summary and Conclusions"

References

Bowman, K., Rodgers, C., Kulawik, S., Worden, J., Sarkissian, E., Osterman, G., ... Beer, R. (2006). Tropospheric emission spectrometer: retrieval method and error analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(5), 1297-1307. doi: 10.1109/TGRS.2006.871234

Chan Miller, C., Roche, S., Wilzewski, J. S., Liu, X., Chance, K., Souri, A. H., ... Wofsy, S. C. (2024). Methane retrieval from MethaneAIR using the CO_2 proxy approach: a demonstration for the upcoming MethaneSAT mission. *Atmospheric Measurement Techniques*, *17*(18), 5429–5454. Retrieved from `https://amt.copernicus.org/articles/17/5429/2024/` doi: 10.5194/amt-17-5429-2024

Connor, B., Bösch, H., McDuffie, J., Taylor, T., Fu, D., Frankenberg, C., ... Jiang, Y. (2016). Quantification of uncertainties in oco-2 measurements of xco_2: simulations and

linear error analysis. *Atmospheric Measurement Techniques*, *9*(10), 5227–5238. Retrieved from https://amt.copernicus.org/articles/9/5227/2016/ doi: 10.5194/amt-9-5227-2016

Nakajima, M., Kuze, A., & Suto, H. (2012). The current status of GOSAT and the concept of GOSAT-2. In R. Meynart, S. P. Neeck, & H. Shimoda (Eds.), *Sensors, systems, and next-generation satellites xvi* (Vol. 8533, p. 853306). SPIE. Retrieved from https://doi.org/10.1117/12.974954 doi: 10.1117/12.974954

Parker, R. J., Webb, A., Boesch, H., Somkuti, P., Barrio Guillo, R., Di Noia, A., ... Wunch, D. (2020). A decade of GOSAT Proxy satellite CH_4 observations. *Earth System Science Data*, *12*(4), 3383–3412. Retrieved from https://essd.copernicus.org/articles/12/3383/2020/ doi: 10.5194/essd-12-3383-2020

Ricaud, P., Attié, J.-L., Chalinel, R., Pasternak, F., Léonard, J., Pison, I., ... Zwaaftink, C. G. (2021). The monitoring nitrous oxide sources (min2os) satellite project. *Remote Sensing of Environment*, *266*, 112688. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425721004089 doi: https://doi.org/10.1016/j.rse.2021.112688

Zhang, Y., Sun, K., Gao, Z., Pan, Z., Shook, M. A., & Li, D. (2020). Diurnal climatology of planetary boundary layer height over the contiguous united states derived from amdar and reanalysis data. *Journal of Geophysical Research: Atmospheres*, *125*(20), e2020JD032803. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JD032803 (e2020JD032803 2020JD032803) doi: https://doi.org/10.1029/2020JD032803