HyperGas 1.0: A Python package for analyzing hyperspectral data for greenhouse gases from retrieval to emission rate quantification

Xin Zhang[1], Joannes D. Maasakkers[1], Tobias A. de Jong[1], Paul Tol[1], Frances Reuland[2], Adam R. Brandt[2], Eric A. Kort[3], Taylor J. Adams[3], and Ilse Aben[1,4]

[1]SRON Space Research Organisation Netherlands, Leiden, 2333 CA, The Netherlands
[2]Department of Energy Science & Engineering, Stanford University, Stanford, California 94305, United States
[3]Department of Climate & Space Sciences & Engineering, University of Michigan, Ann Arbor, MI 48109, USA
[4]Department of Earth Sciences, Vrije Universiteit Amsterdam, Amsterdam, 1081 HV, The Netherlands

Correspondence: Xin Zhang (xin.zhang@sron.nl)

We thank the reviewers for their positive comments and very careful reading of our article. The individual corrections suggested are addressed below. The reviewers' comments will be shown in ***black italic***, our response in **normal font**, and changes made to the paper are shown in **blue** block quotes. Unless otherwise indicated, page and line numbers correspond to the original manuscript, while reference numbers follow the new order.

Contents

1 Reviewer 1

Major Concerns

(1) *Since this framework is intended not only for spaceborne hyperspectral imagers but also potentially for airborne instruments in the future, using unit absorption spectra (k) in units of ppm·m is recommended to improve applicability and inter-instrument comparability (see, e.g., DOI: 10.1016/j.rse.2021.112574).*

Because the current HyperGas package is primarily designed for hyperspectral satellite data, we use ppb for methane and ppm for CO_2 as the default units. HyperGas is fully configurable, and users can easily switch to ppm·m via the configuration file when needed. We have clarified this configuration in the revised manuscript.

> For methane, we evaluate enhancements from 0 to 6400 ppb in geometric progression (doubling from 100 ppb), while CO_2 enhancements range from 0 to 160 ppm (doubling from 2.5 ppm), reflecting the broader dynamic range and higher background concentration of CO_2 compared to methane. **For airborne data, users can modify the configuration file to compute *k* in ppmm units**.

(2) *In addition, unlike the commonly used per-column basis analysis, this work derives reference spectra based on clustering. While this approach is reasonable, as also noted by the authors, detector arrays often exhibit cross-track variability, especially for instruments with strong smile effects. Under such conditions, I do not expect a clustering-based approach (regardless of how surface types are classified) to outperform the per-column approach, at least over homogeneous surfaces. Have the authors conducted any comparison between the clustering-based and per-column-based methods?*

HyperGas supports both per-column and cluster-tuned matched filters. We compared the two methods in Appendix Figure A1 and discussed the possible reasons for the lack of performance improvement. Since the smile effect is not the only cause for cross-track variability, we have now added the cross-track point and noted future work of correcting the smile effect.

> We have also tested the cluster-tuned matched filter in urban areas (Appendix Figure A1), however, the results are still noisy, making it challenging to differentiate plumes from the background. The improved retrieval performance of previous airborne imaging spectrometers is likely related to their higher spatial resolution (3–8 m) and narrower swath width (∼5 km) compared with HSIs, which also leads to longer observed plumes within a single scene. **Another challenge in the cluster analysis is the cross-track variability, which includes the smile effect (Guanter et al., 2021), thereby affecting the cluster-tuned matched filter. A joint application of smile-effect correction and a cluster-tuned matched filter could improve retrieval performance and can be developed for *HyperGas* at a later stage.** Therefore, *HyperGas* does not rely on the k-means method but instead applies land and water masks derived from the OSM and ESA WorldCover datasets.

.....

This initial release (*HyperGas* v1.0) provides robust functionality that can serve as the foundation for several additional features that can be developed in the future, including:

1. Automatic co-registration of PRISMA data (De Luca et al., 2024) to correct spatial misalignment between PRISMA imagery and actual Earth surface features.

2. **Correction of smile effects (Guanter et al., 2021)**.

....

Specific & minor points

(1) *Line 14: It would be better to include the chemical formula (CH4) after methane, consistent with the notation used for CO2.*

For the main text we prefer that the notation of the gases follows pronunciation conventions: "CO_2" for carbon dioxide and "methane" for methane. We do now include the introduction of the chemical formula on the first mention of methane.

Greenhouse gases such as carbon dioxide (CO_2) and **methane (CH_4)** are the primary drivers of anthropogenic climate change, contributing to global warming and altering the Earth's energy balance (Intergovernmental Panel on Climate Change (IPCC), 2023).

(2) *Line 16: It may be more appropriate to replace "for identifying emission sources" with "for quantifying point source emissions" in this context.*

We have included the quantification point.

Monitoring these gases at facility-scale is increasingly important **for identifying and quantifying emission sources**, supporting and verifying mitigation efforts, and informing climate policy.

(3) *Line 19: It may be more appropriate to replace "methane and CO2 emission plumes from individual facilities" with "methane and carbon dioxide point sources." Also, please avoid mixing full names and abbreviations (methane vs. CO2); abbreviations should be used consistently after being defined (see DOI: 10.1126/sciadv.adh2391). Please check and correct this writing issue throughout the entire manuscript (e.g., Line 26, Line 37, etc.).*

As explained in our response to the first minor point, we perfer to keep the "methane and CO_2" format.

We aim to emphasize the detection of plumes from different facilities, such as oil and gas sites (point sources) and landfills (area sources). Therefore, we changed it to "concentrated sources".

Spaceborne hyperspectral imagers (HSI), with their ability to capture hundreds of narrow, contiguous spectral bands, have been demonstrated to enable the detection and quantification of methane and CO_2 **concentrated sources** (Guanter et al., 2021; Jacob et al., 2022; Cusworth et al., 2023; Thorpe et al., 2023; Borger et al., 2025; Zhang et al., 2025).

(4) *Line 116: Please define FWHM (full width at half maximum) at its first occurrence before using the abbreviation.*

Thank you for this correction. We have added the definition of the full term.

To derive $\boldsymbol{k}$, we apply a radiative transfer model (Gloudemans et al., 2008), incorporating the instrument's spectral response function characterized by its central wavelength and **full width at half maximum (FWHM; Thompson et al., 2015)**.

(5) *Equations (3) and (5) appear to be inconsistent with the corresponding equations in the cited reference. Please check and revise them accordingly.*

Equation (3) is the same as Thompson et al. (2015), while Equation (5) represents its lognormal form.

In the lognormal matched filter, the logarithm is applied to both the target and the observation so that the data are modeled as Gaussian in log space (Schaum, 2021). Thus, the covariance should also be computed in the same transformed domain, i.e., $\boldsymbol{\Sigma} \sim \mathrm{Cov}(log\boldsymbol{L})$. Using the covariance from the original linear domain would mix incompatible statistical spaces. Moreover, the quadratic form should remain dimensionless. This is only valid when both the mean and covariance are defined consistently in log space. This is not correctly represented in equation (8) in Pei et al. (2023).

(6) *Line 165: It would be better to replace "in (for example) urban areas" with "in heterogeneous areas."*

We have rephrased the sentence as you suggested.

For plume detection, we perform the matched filter over the entire near-infrared window (1300–2500 nm, Roger et al., 2024), instead of only over the strong methane (2100–2450 nm) and CO_2 (1928–2200 nm) absorption windows, to mitigate the background noise **in heterogeneous areas**.

(7) *Regarding Figure 5, it is unclear what the 30° azimuth difference refers to and how it is calculated. Please clarify what angle is being compared and how the orientation of the rectangular masks is defined. In Figure 5(b), when compared with Figure 5(a), the gray rectangle on the right also appears to contain a methane plume, but in Figure 5(c) this potential plume is excluded. Could the authors clarify why this candidate was removed and which specific criterion was responsible for the exclusion? Moreover, the two orange rectangles in Figure 5(b) appear more likely to originate from the same point source. Visual discontinuities in plumes are common, especially in high-spatial-resolution but relatively low-precision (compared to*

TROPOMI) methane plume detection. Please justify why these are treated as two separate candidates rather than merged into a single plume.

Thank you for these valuable comments. In this case, we focus on the western plume and explain how other candidate plumes in the scene are filtered.

As shown in Fig. 5b, we first manually place a white star marker at the identified emission source location. The watershed algorithm produces multiple candidate masks, and we calculate the minimum rotated rectangle (oriented bounding box) for each mask. The rectangle containing the source marker is designated as the reference plume. We then calculate the azimuth difference between the reference rectangle and all other candidate rectangles. Masks whose rectangles have azimuth differences less than 30° from the reference are merged together as a single plume (shown in orange in Fig. 5b), assuming minimal wind direction changes near the source. This is why the western plume appears in orange while the eastern candidate remains gray.

Regarding the gray rectangle on the right in Figure 5b: this candidate was excluded from being part of the western plume because its azimuth differs by more than 30° from the reference plume, suggesting it likely originates from a different source or represents a spurious detection rather than being part of the same emission event. The 30° threshold is a conservative criterion designed to ensure merged masks represent the same wind-advected plume. Concerning the two orange rectangles: while we acknowledge that visual discontinuities are common in high-resolution plume detection, our algorithm initially treats them as separate candidates through the watershed segmentation. However, because both rectangles satisfy the azimuth criterion (< 30° difference from the reference), they are automatically merged into a single plume mask in the final output (Fig. 5c). This merging process, combined with the dilation step, ensures that visually discontinuous but physically connected plume segments are treated as one emission source.

We have revised the figure caption and main text to clarify these steps.

> Emission sources are manually identified based on wind data and ESRI imagery. **To determine whether multiple candidate masks belong to the same emission source, we calculate the minimum rotated rectangle (oriented bounding box) for each mask. The mask containing the manually identified source is designated as the reference, and we compute the azimuth difference between its rectangle and those of all other candidates. Only masks with azimuth differences less than 30° relative to the reference are merged together as a single plume (e.g., orange rectangles in Fig. 5b), assuming minimal wind direction changes near the source. This azimuth-based filtering helps exclude spurious detections or plumes from different sources (e.g., gray rectangles in Fig. 5b). Figure 5c demonstrates the final mask determined for the western methane emission plume in the scene.**

(8) *Line 185: The description of how the mask is determined based on the angle is confusing. Please reorganize and clarify this part.*

As mentioned above, we have modified the paragraph to make it clearer.

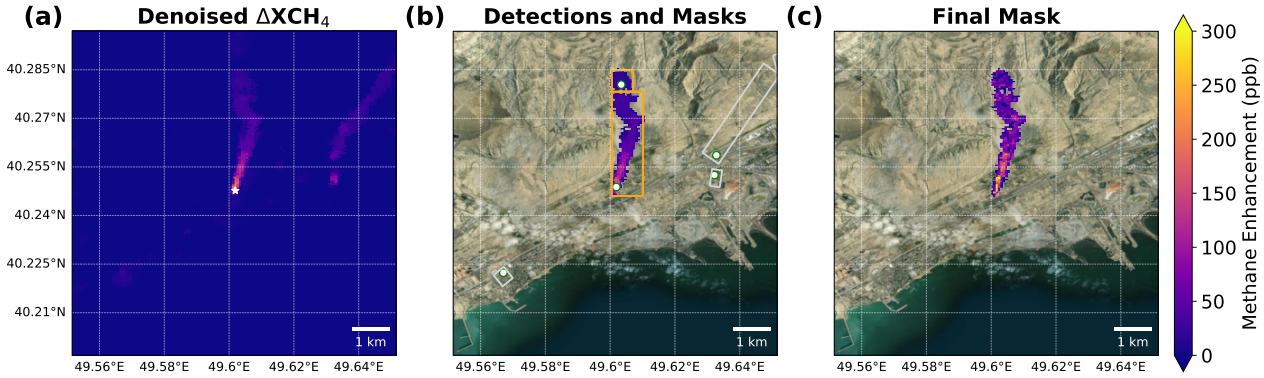

Figure 1. Plume mask creation process for the western plume from Fig. 3a. (a) The denoised methane enhancement (ΔXCH_4) field obtained by applying the Chambolle total variation (TV) denoising filter to ΔXCH_4 within the 1300∼2500 nm window. The white star marks the manually identified source location for the western plume. (b) Initial candidate masks derived from the watershed algorithm. White dots indicate high-ΔXCH_4 locations, and colored rectangles represent the minimum rotated rectangles (oriented bounding boxes) for each mask. Orange rectangles denote masks with azimuth differences less than 30° relative to the reference mask (containing the source marker of the western source) and are merged into a single plume. The gray rectangles are candidates excluded by the azimuth criterion. (c) The final ΔXCH_4 plume mask for the western emission source after merging. Background imagery source: Esri | Powered by Esri.

(9) *Line 235: Why not directly resample the original 25 m data to 30 m or 60 m, instead of using the current resampling strategy?*

Since the original WRF-LES simulation has a grid spacing of 25 m, resampling to 50 m was initially straightforward. We have improved the algorithm to conservatively resample the data to 30 m and 60 m resolutions, and the resulting calibration coefficients remain similar. All estimates and figures have been updated using these revised wind calibrations.

To match sensor spatial resolution, the WRF-LES data are conservatively resampled to 30 m for EnMAP and PRISMA and to 60 m for EMIT. The resampling weights are based on the ratio of source cell area overlapped with the corresponding destination cell. .

2 Reviewer 2

Major Concerns

(1) *My primary concern relates to the limited algorithmic novelty and insufficient differentiation from existing tools. The core retrieval algorithm implemented in HyperGas is the standard linear matched filter (Thompson et al., 2015), supplemented by the lognormal variant (Schaum, 2021; Pei et al., 2023). The emission quantification relies on the well-established IME and CSF frameworks (Varon et al., 2018). While the authors mention existing tools such as mag1c and emit-ghg, the differentia-*

tion remains superficial. In particular, the ddeq library (Kuhlmann et al., 2024) already provides a modular, sensor-agnostic framework for point-source quantification from remote sensing images, including CSF implementation. The authors should provide a more rigorous and systematic comparison with ddeq and other existing packages (e.g., the Carbon Mapper pipeline described in Duren et al., 2025), clearly articulating what HyperGas contributes beyond integration of known methods into a single workflow. Does HyperGas offer measurable improvements in retrieval accuracy, computational efficiency, or detection sensitivity compared to these tools?

Thank you for the comment. We agree that the manuscript should more clearly distinguish HyperGas from existing retrieval and quantification tools. Our intention was not to claim new methods in the underlying matched filter or IME/CSF formulations, which are indeed established approaches, but rather to address a practical gap in the availability of an open, reproducible, user-friendly, and end-to-end framework that integrates retrieval, detection, plume delineation, and emission quantification across multiple instrument data within a single workflow. We have revised the manuscript to make this distinction explicit.

> Existing open-source tools, such as mag1c (https://github.com/markusfoote/mag1c/) and emit-ghg (https://github.com/emit-sds/emit-ghg), are restricted to a single HSI or data format, **while frameworks such as ddeq (Kuhlmann et al., 2024) focus mainly on source quantification rather than a complete end-to-end workflow. To address these limitations, we introduce *HyperGas*, an open-source Python package designed to streamline greenhouse gas analysis across different HSI datasets. Instead of proposing a new retrieval algorithm, *HyperGas* combines established matched-filter techniques for gas enhancement retrieval into a unified workflow. It also applies standard integrated mass enhancement (IME) and cross-sectional flux (CSF) methods for emission quantification, enabling consistent application across multiple sensors. The package supports enhancement retrieval, plume detection, masking, uncertainty handling, and emission estimation through both automated batch processing and an interactive graphical user interface. Its modular design also allows users to add additional instruments, incorporate and contribute new algorithms, as well as develop customized processing strategies.** In this paper, we present the design principles behind *HyperGas*, demonstrate its capabilities through a number of real-world case studies involving methane and CO_2, and outline how it can serve as a foundation for reproducible, scalable greenhouse gas analysis using HSI data.

(2) *The controlled-release validation relies on only 19 experiments, all with plume lengths under 1 km and all conducted in limited geographic and meteorological conditions. This is a relatively small sample, and the absence of EMIT from the controlled-release validation is a notable gap, particularly since EMIT has a fundamentally different spatial resolution that could affect detection and quantification performance. Furthermore, the CSF method is not validated against controlled releases at all, the authors note that plume lengths were too short, but this leaves the CSF implementation essentially unvalidated. The CO2 quantification is demonstrated only qualitatively (three power plant cases) with no independent validation data. The*

authors acknowledge this but it weakens the paper's claims about CO2 capabilities. If possible, expand the validation dataset or at least compare their CO2 estimates against reported emission data from the demonstrated power plants.

Since submitting our preprint, additional Stanford University controlled-release data from 2025 have become available, these form the state-of-the-art controlled release dataset. We have summarized a total of 75 cases, including 8 EMIT cases, and have updated the main text. We have also validated our CSF results for controlled releases involving more than five CSF cross-section lines. Future releases at higher emission rates will further strengthen validation of the CSF method. For CO_2 estimates, we have added comparisons with Continuous Emission Monitoring Systems (CEMS).

> We also validated our default IME and CSF estimates using **75** Stanford controlled releases conducted in 2024 and 2025 (Reuland et al., 2025; Reuland and Brandt, 2025; Brandt et al., 2026). **The results show good agreement between the IME estimates and the controlled releases, with slopes ranging from 0.92 to 1.22 (Fig. 8).** Since all plume lengths were under 1 km, any differences between the IME and IME fetch estimates are purely caused by different wind calibration coefficients and not by different sampling of the plumes. Figure 8 shows that the results do not differ significantly across the various IME fetch limits. *HyperGas* currently defaults to wind calibration without any fetch limitation.
>
> **Because short plume lengths may be insufficient for reliable CSF estimation, we perform CSF validation only for plumes with more than five CSF cross-section lines. The fitted slope between estimates and controlled release rates decreases from 1.47 to 1.16 when wind calibration is based on LES cases with emission rates below 2 t/h, while it remains unchanged for the IME method. Thus, we recommend using IME results by default and relying on CSF primarily for sensitivity analysis of the estimates.** In the future, controlled release experiments involving elongated plumes may support calibration in two key ways: (1) evaluating the accuracy of different fetch limits, and (2) enabling wind calibration directly using measured U_{10} data, which can then be applied to other cases with available wind observations.
>
> In addition to methane, *HyperGas* also supports quantifying CO_2 emission rates. Figure 11a–c show CO_2 plumes from three power plants: James H. Miller Jr. (AL, USA), Rockport (IN, USA), and Stanwell Power Station (Stanwell, Australia). The estimated emission rates range from 0.5 to 2 kt h^{-1}. On average, CSF estimates tend to be higher, particularly for the compact plume observed at the Stanwell Power Station and James H. Miller Jr (Fig. 11d). **For U.S. power plants, we compare our estimates with stack-level estimates from Continuous Emission Monitoring Systems (CEMS, Fig. 11e–f; MEDUSA team D9.1, 2025). Although the CEMS reports generally fall within the uncertainty range of the CSF estimates, they tend to lie at the upper end of, or exceed, the IME results.**

(3) *The wind calibration that underpins both the IME and CSF quantification methods is derived from only five 3-hour WRF-LES simulations. These simulations represent a limited range of atmospheric stability conditions and boundary-layer*

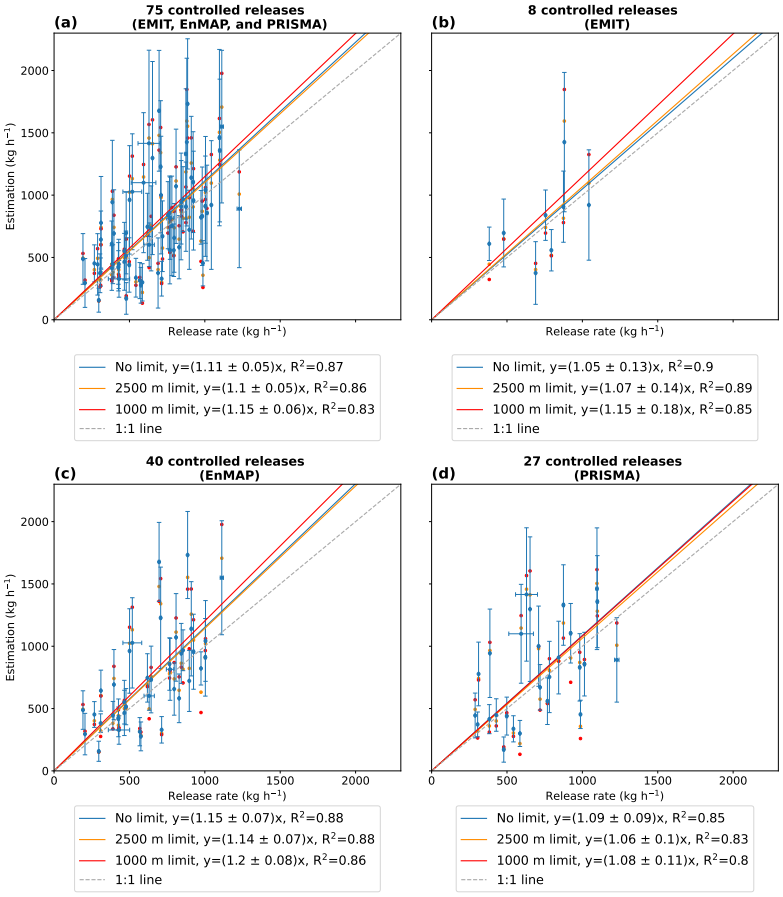

Figure 2. Quantification performance of methane-emission estimates using the IME method under different fetch limits (none, 2500 m, and 1000 m) with calibrations shown in Fig. 7a–c. (a) Controlled releases observed by EMIT, EnMAP, and PRISMA. (b) EMIT only. (c) EnMAP only. (d) PRISMA only. Regression lines are fitted with Ordinary Least Squares (OLS). Estimation errors are derived from *HyperGas* outputs. The release-rate errors are estimated as the standard deviation of mean emission rates released over consecutive time windows starting 1-5 minutes before and ending at the overpass time (i.e., [T–1 min, T], [T–2 min, T], . . . , [T–5 min, T]). For clarity, scatter points for the fetch-limited results are shown without error bars.

regimes. The study does not discuss how the calibration might perform under strongly stable or convective conditions, or how sensitive the results are to the turbulence parameterization in WRF-LES. Given that wind speed uncertainty typically dominates the error budget in point-source quantification (Jacob et al., 2022), the robustness of these calibrations to a wider range of meteorological scenarios should be addressed. Additionally, the 50 m resampling used for EMIT does not exactly match EMIT's 60 m pixel size, and this discrepancy should be justified.

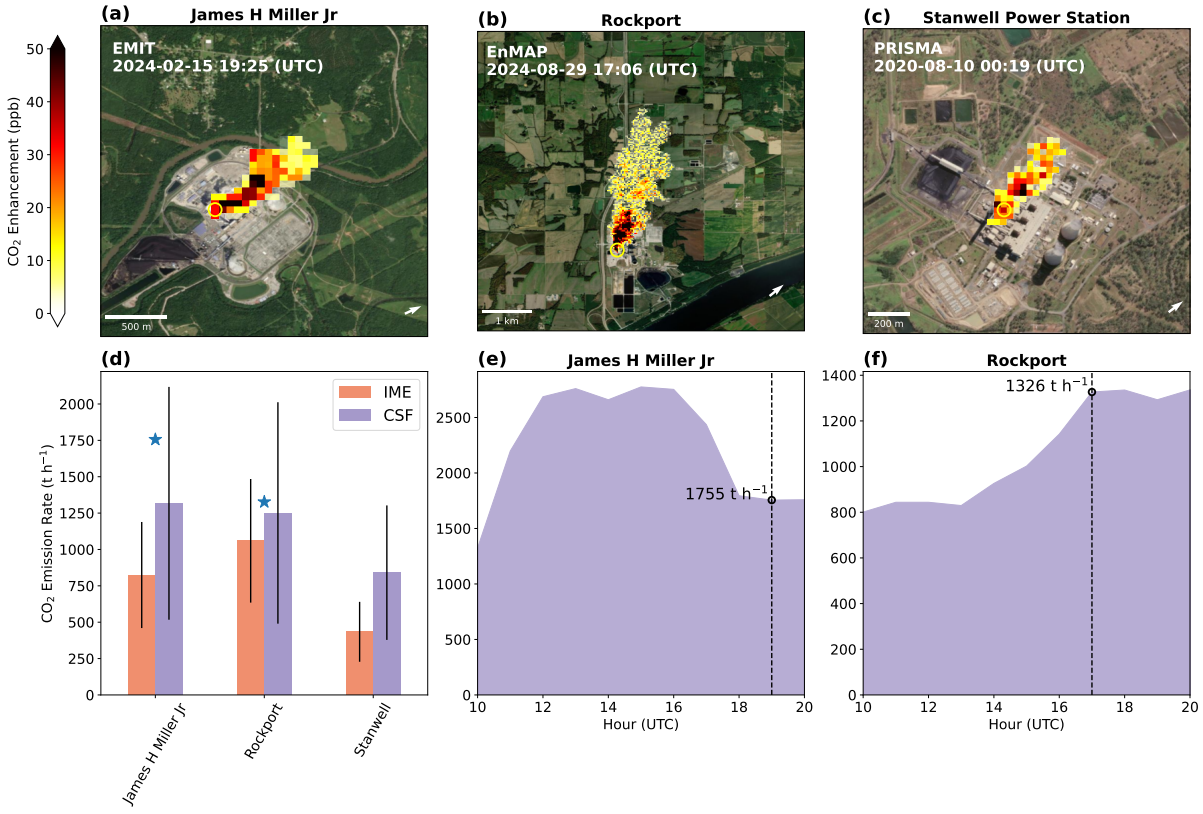

Figure 3. CO_2 plumes from power plants detected by (a) EMIT, (b) EnMAP, and (c) PRISMA. Instrument names and satellite overpass times are labeled in white text at the top of each panel. Yellow circles are identified plume source locations. Background ESRI imagery reproduced with permission as granted on Esri's website for noncommercial scholarly use (Esri et al., 2022). Background imagery source: Esri | Powered by Esri. (d) Emission rate estimates derived using the IME and CSF methods. The blue stars are the CEMS-measured CO_2 emissions from the EPA database, interpolated at the time of the satellite overpass. (e–f) Time series of CEMS-measured CO_2 emissions from the EPA database. The dotted line indicates the satellite overpass time.

We thank the reviewer for raising this point. We clarify that the WRF-LES simulations span boundary-layer depths of 500–2000 m and sensible heat fluxes of 100–300 W m^{-2}. These conditions are intentionally chosen to represent daytime, cloud-free convective boundary layers, which are the primary regimes under which SWIR satellite instruments operate. Accordingly, the calibration is designed for meteorological conditions most relevant to satellite-based methane plume observations, rather than for all possible atmospheric states.

We agree that strongly stable and more extreme convective regimes are not explicitly included. Finally, consistent with Jacob et al. (2022), wind uncertainty is included in the estimate: a relative error of 50% for wind speeds below 3 m/s and a fixed error of 1.5 m s^{-1} for wind speeds exceeding 3 m/s (Varon et al., 2018; Zhang et al., 2025). We have also conservatively resampled the data to 60 m resolution for EMIT, as also suggested by Reviewer #1.

To calibrate $U_{\mathrm{eff,IME}}$ and $U_{\mathrm{eff,CSF}}$ against the domain-averaged 10 m wind speeds (U_{10}), we have used five 3-hour large-eddy simulations (LES) performed using the Weather Research and Forecasting model (WRF-LES) at 25 m spatial resolution and 30 s temporal resolution (Varon et al., 2018; Maasakkers et al., 2022). **The simulation accounts for varying meteorological conditions, including a range of mixed layer depths (500–2000 m) and sensible heat fluxes (100–300 W m^{-2}), representing typical conditions observed with SWIR satellite instruments.**

(4) *The authors state that HyperGas defaults to the standard matched filter and switches to the lognormal variant "for large methane emissions (e.g., > 10 t h-1)" (Pei et al., 2023). However, the nonlinearity issue in matched-filter retrievals is fundamentally a function of the column enhancement magnitude (ΔXCH4), not directly of the emission rate. A 10 t h-1 source under strong winds may produce lower per-pixel enhancements than a 2 t h-1 source under calm conditions. The switching criterion should be based on retrieved ΔXCH4 values rather than on emission rates, which are not known a priori. This design choice needs either better justification or revision.*

We agree that switching between MF and LMF based on the enhancement value is a better approach. According to the sensitivity analysis by Pei et al. (2023), a high bias (>20%) can occur when methane enhancement exceeds 1200 ppb. To mitigate potential albedo effects, we first examine the number of pixels within the plume mask that exceed 1200 ppb. If more than five such pixels are identified, the data are reprocessed using the LMF method.

Since measurement noise can produce negative radiance values in some spectral bands, the lognormal matched filter, which assumes strictly positive inputs, may amplify background noise. Therefore, *HyperGas* defaults to the standard matched filter and switches to the lognormal matched filter when strong methane enhancements are detected. **Specifically, if more than five pixels within the plume mask exceed 1200 ppb, the data are reprocessed using the lognormal matched filter, following the sensitivity analysis in Pei et al. (2023).** Because the CO_2 data are generally noisier than the methane data, *HyperGas* does not automatically switch to the lognormal matched filter for CO_2.

Specific & minor points

(1) *Section 2.1.1 (L1 radiance data): The statement "No pre-treatment is applied to the L1 data" should be qualified. PRISMA L1 data is known to exhibit spectral smile effects and across-track radiometric non-uniformity (Guanter et al., 2021). If no destriping or spectral recalibration is applied, the authors should discuss the expected impact on retrieval quality.*

We agree that PRISMA L1 data are indeed affected by instrumental effects such as spectral smile, as discussed by Guanter et al. (2021). In particular, spectral calibration errors (e.g., central wavelength shifts and FWHM variations) can influence methane retrievals due to the non-symmetric structure of methane absorption features in the fitting window. Depending on the

sign and magnitude of these shifts, systematic biases in retrieved methane enhancement may arise. Compared to PRISMA, other sensors such as EnMAP have been shown to be less sensitive to spectral smile, resulting in improved across-track uniformity (Roger et al., 2024; Ferrari et al., 2026). For EnMAP, the residual smile is sufficiently small that column-dependent SRF convolution can often be neglected, which is not necessarily the case for PRISMA.

In this study, no explicit destriping or spectral recalibration has been applied. This choice may introduce additional uncertainty and potential across-track biases in the retrievals, particularly in scenes where spectral smile effects are pronounced. We have revised the manuscript to explicitly acknowledge these limitations and to discuss potential solutions.

> For PRISMA, the L1 data are obtained from the PRISMA Portal (https://prisma.asi.it/); for EnMAP, these come from the EOWEB GeoPortal (https://eoweb.dlr.de/); and for EMIT, users can download L1 data from NASA Earthdata (https://search.earthdata.nasa.gov/). **Since no pre-treatment is applied to the L1 data, spectral calibration errors (e.g., central wavelength shifts and full-width at half-maximum variations) can propagate into retrievals due to the absorption structure of trace gases in the fitting window. Compared to PRISMA, sensors such as EnMAP have shown a smaller spectral smile, resulting in improved across-track uniformity (Roger et al., 2024; Ferrari et al., 2026). A potential mitigation strategy is to fit modeled spectra against reference spectra to correct for spectral calibration errors and improve retrieval accuracy (Guanter et al., 2021). This approach can be incorporated into *HyperGas* in future work.**

(2) *Section 2.2.1 (Matched filter): The manuscript uses the SPy (Spectral Python) library for the matched filter implementation. Given that MAG1C (Foote et al., 2020) introduces sparsity priors and per-pixel albedo corrections that demonstrably improve retrievals, the authors should discuss why they chose the simpler SPy implementation over MAG1C-style enhancements.*

In this work, methane enhancements are retrieved using a classical matched filter from the SPy (Spectral Python) library. More advanced approaches, such as MAG1C (Foote et al., 2020), extend the matched filter framework by incorporating a sparse prior and albedo correction, which have been shown to improve detection sensitivity and robustness. However, because the focus of this study on baseline performance assessment, we adopt the simpler matched filter, which is more straightforward to interpret. In addition, MAG1C is computationally less efficient due to the iterative optimization procedure. We now mention the potential expansion of the framework to MAG1C-like approaches in the manuscript.

> This initial release (*HyperGas* v1.0) provides robust functionality that can serve as the foundation for several additional features that can be developed in the future, including:
>
> 1. Automatic co-registration of PRISMA data (De Luca et al., 2024) to correct spatial misalignment between PRISMA imagery and actual Earth surface features.
> 2. Correction of smile effects (Guanter et al., 2021).

3. **Support for a matched filter optimized with a sparse prior (Foote et al., 2020).**

4. ...

(3) *Section 2.2.2 (Lognormal matched filter): The claim that "the lognormal matched filter only supports positive radiances" is somewhat misleading. All physical radiances are positive; the issue is that noise can produce negative values in certain bands. This should be clarified.*

Thanks for the correction. We have revised the sentences.

> **Since measurement noise can produce negative radiance values in some spectral bands, the lognormal matched filter, which assumes strictly positive inputs, may amplify background noise. Therefore, *HyperGas* defaults to the standard matched filter and switches to the lognormal matched filter when strong methane enhancements are detected.** Specifically, if more than five pixels within the plume mask exceed 1200 ppb, the data are reprocessed using the lognormal matched filter, following the sensitivity analysis in Pei et al. (2023).

(4) *Section 2.3 (Plume detection): The use of the full NIR window (1300–2500 nm) for plume detection following Roger et al. (2024) is a sensible choice. However, the combined use of Chambolle TV denoising with J-Invariance calibration and watershed segmentation from tobac introduces multiple tunable parameters. The sensitivity of the final emission estimate to these parameter choices is not systematically assessed. A sensitivity analysis, at least for the most critical parameters, would strengthen confidence in the method's robustness.*

Thank you for highlighting the sensitivity test. We assessed the uncertainties associated with denoising and plume masking using WRF-LES simulations, focusing on two key parameters: the TV filter weight and watershed threshold.

For the TV filter weight, we examined two additional settings, 0.75 times and 1.5 times the default optimized value, alongside the baseline configuration. The default watershed thresholds are defined as mean + 2σ and mean + 3σ. In addition, we tested two alternative configurations, [mean + 1.5σ, mean + 2.5σ] and [mean + 2.5σ, mean + 3.5σ]. For each plume, we generated an ensemble comprising all parameter combinations and quantified the associated uncertainty as the standard deviation across the 9 ensemble members. The resulting mean relative emission uncertainties attributable to parameter selection are 8%, 10%, and 15% for EMIT, EnMAP, and PRISMA, respectively. This component is now incorporated into the final uncertainty estimate.

> **The plume masking procedure depends on two factors, the TV filter weight and watershed segmentation thresholds. For the TV filter weight, we examined two additional settings, 0.75 times and 1.5 times the default optimized value, alongside the baseline configuration. The default watershed thresholds are defined as mean + 2σ and mean + 3σ. In addition, we tested two alternative configurations, [mean + 1.5σ, mean + 2.5σ] and [mean + 2.5σ, mean + 3.5σ]. For each plume, we generated an ensemble comprising all parameter combinations and quantified the associated uncertainty as the**

standard deviation across the 9 ensemble members. The resulting mean relative emission uncertainties attributable to parameter selection are 8%, 10%, and 15% for EMIT, EnMAP, and PRISMA, respectively.

(5) *Section 2.4.2 (Wind calibrations): The three background scenes (Xinjiang, Anna Creek, Madrid) are all over land. No water-adjacent or coastal scenes are included. How representative are the calibrations for plumes occurring near coastlines or over partially water-covered scenes?*

By default, *HyperGas* masks water pixels as described in Section 2.1.2. If users want to analyze plumes over water, such as those from offshore oil and gas extraction basins, a custom wind calibration using LES for comparable regions could improve results. We have noted this in the main text.

Moreover, as the LES-based wind calibrations are derived using simulations over land, specific offshore calibrations could be added in the future to improve results for plumes over water such as from offshore oil and gas platforms.

(6) *Section 3.2 (Carbon dioxide emissions): The discussion of buoyant plume effects on CO2 wind calibration is important but incomplete. The authors note that "actual CO2 emissions are released as hot, buoyant plumes at heights greater than 10 m" but power plant stacks are typically 100–200 m tall, and the effective plume rise can be several hundred meters. This substantially changes the relevant wind speed. Using U10 for elevated sources is a known and significant limitation that deserves more thorough treatment.*

We agree that using wind speed at the appropriate height is important for estimating power plant CO_2 emissions. A more robust approach would involve running LES with representative emission profiles and calibrating the estimates using wind speeds measured representative for the relevant height. As this is beyond the scope of this study, we instead examined the difference between wind speeds at 100 m and 10 m and have noted this limitation in the main text.

This discrepancy may arise from plume simulations that assume surface-level sources and use only U_{10} to represent wind speed, whereas the actual CO_2 emissions are released as hot, buoyant plumes at heights greater than 10 m. Consequently, using U_{10} as a proxy may result in inaccurate wind speed calibrations for our observations, which could potentially affect emission estimates. **As noted in the MEDUSA project (MEDUSA team D9.1, 2025), analyzed wind speeds at 100 m were on average 30% higher than at 10 m. Future refinements could involve CO_2-specific simulations (Brunner et al., 2019) to improve this calibration by accounting for buoyant plumes released at altitude and relating the effective wind speed to a more representative reanalysis metric like the 100-m wind speed.** Nevertheless, both IME and CSF estimates remain within their respective uncertainty bounds.

(7) *Line 361: "HypeGas v1.0" appears to be a typographical error, should be "HyperGas v1.0."*

Thanks for pointing the typo. We have fixed it.

(8) *General Writing: The manuscript is clearly written overall, but there are minor instances of inconsistent notation (e.g., ΔX vs. $\Delta XCH4$) and a few grammatical issues (e.g., Line 269: "We have performed", should be lowercase "we").*

The abbreviation "ΔX" and "ΔXCH_4" actually denote different quantities here. "ΔX" is the general L2 concentration enhancement, while "ΔXCH_4" is the methane enhancement. Therefore, we do not change their notation.

We have re-evaluated the manuscript to correct grammatical and typographical errors.

References

Borger, C., Beirle, S., Butz, A., Scheidweiler, L. O., and Wagner, T.: High-Resolution Observations of NO_2 and CO_2 Emission Plumes from EnMAP Satellite Measurements, Environmental Research Letters, 20, 044 034, https://doi.org/10.1088/1748-9326/adc0b1, 2025.

Brandt, A., Reuland, F., Adams, T., Sherwin, E., Abbadi, S. E., and Kort, E.: Unlocking Credible Space-Based Methane Sensing through a Year-Long Single-Blind Test, https://doi.org/10.21203/rs.3.rs-9110475/v1, 2026.

Brunner, D., Kuhlmann, G., Marshall, J., Clément, V., Fuhrer, O., Broquet, G., Löscher, A., and Meijer, Y.: Accounting for the Vertical Distribution of Emissions in Atmospheric CO_2 Simulations, Atmospheric Chemistry and Physics, 19, 4541–4559, https://doi.org/10.5194/acp-19-4541-2019, 2019.

Cusworth, D. H., Thorpe, A. K., Miller, C. E., Ayasse, A. K., Jiorle, R., Duren, R. M., Nassar, R., Mastrogiacomo, J.-P., and Nelson, R. R.: Two Years of Satellite-Based Carbon Dioxide Emission Quantification at the World's Largest Coal-Fired Power Plants, Atmospheric Chemistry and Physics, 23, 14 577–14 591, https://doi.org/10.5194/acp-23-14577-2023, 2023.

De Luca, G., Carotenuto, F., Genesio, L., Pepe, M., Toscano, P., Boschetti, M., Miglietta, F., and Gioli, B.: Improving PRISMA Hyperspectral Spatial Resolution and Geolocation by Using Sentinel-2: Development and Test of an Operational Procedure in Urban and Rural Areas, ISPRS Journal of Photogrammetry and Remote Sensing, 215, 112–135, https://doi.org/10.1016/j.isprsjprs.2024.07.003, 2024.

Esri, Maxar, Geographics, E., and the GIS User Community: ESRI World Imagery, https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/ 2022.

Ferrari, A., Pampanoni, V., Laneve, G., Tellez, R. A. C., and Saquella, S.: A Multi-Sensor Framework for Methane Detection and Flux Estimation with Scale-Aware Plume Segmentation and Uncertainty Propagation from High-Resolution Spaceborne Imaging Spectrometers, Methane, 5, https://doi.org/10.3390/methane5010010, 2026.

Foote, M. D., Dennison, P. E., Thorpe, A. K., Thompson, D. R., Jongaramrungruang, S., Frankenberg, C., and Joshi, S. C.: Fast and Accurate Retrieval of Methane Concentration From Imaging Spectrometer Data Using Sparsity Prior, IEEE Transactions on Geoscience and Remote Sensing, 58, 6480–6492, https://doi.org/10.1109/TGRS.2020.2976888, 2020.

Gloudemans, A. M. S., Schrijver, H., Hasekamp, O. P., and Aben, I.: Error Analysis for CO and CH_4 Total Column Retrievals from SCIAMACHY 2.3 Mm Spectra, Atmospheric Chemistry and Physics, 8, 3999–4017, https://doi.org/10.5194/acp-8-3999-2008, 2008.

Guanter, L., Irakulis-Loitxate, I., Gorroño, J., Sánchez-García, E., Cusworth, D. H., Varon, D. J., Cogliati, S., and Colombo, R.: Mapping Methane Point Emissions with the PRISMA Spaceborne Imaging Spectrometer, Remote Sensing of Environment, 265, 112 671, https://doi.org/10.1016/j.rse.2021.112671, 2021.

Intergovernmental Panel on Climate Change (IPCC): Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, https://doi.org/10.1017/9781009157896, 2023.

Jacob, D. J., Varon, D. J., Cusworth, D. H., Dennison, P. E., Frankenberg, C., Gautam, R., Guanter, L., Kelley, J., McKeever, J., Ott, L. E., Poulter, B., Qu, Z., Thorpe, A. K., Worden, J. R., and Duren, R. M.: Quantifying Methane Emissions from the Global Scale down to Point Sources Using Satellite Observations of Atmospheric Methane, Atmospheric Chemistry and Physics, 22, 9617–9646, https://doi.org/10.5194/acp-22-9617-2022, 2022.

Kuhlmann, G., Koene, E., Meier, S., Santaren, D., Broquet, G., Chevallier, F., Hakkarainen, J., Nurmela, J., Amorós, L., Tamminen, J., and Brunner, D.: The *Ddeq* Python Library for Point Source Quantification from Remote Sensing Images (Version 1.0), Geoscientific Model Development, 17, 4773–4789, https://doi.org/10.5194/gmd-17-4773-2024, 2024.

Maasakkers, J. D., Varon, D. J., Elfarsdóttir, A., McKeever, J., Jervis, D., Mahapatra, G., Pandey, S., Lorente, A., Borsdorff, T., Foorthuis, L. R., Schuit, B. J., Tol, P., van Kempen, T. A., van Hees, R., and Aben, I.: Using Satellites to Uncover Large Methane Emissions from Landfills, Science Advances, 8, eabn9683, https://doi.org/10.1126/sciadv.abn9683, 2022.

MEDUSA team D9.1: D9.1 Intercomparison of High-Resolution CO_2 Satellites, https://climate.esa.int/media/documents/MEDUSA_D9.1_HR_CO2_v2.2 2025.

Pei, Z., Han, G., Mao, H., Chen, C., Shi, T., Yang, K., Ma, X., and Gong, W.: Improving Quantification of Methane Point Source Emissions from Imaging Spectroscopy, Remote Sensing of Environment, 295, 113 652, https://doi.org/10.1016/j.rse.2023.113652, 2023.

Reuland, F. and Brandt, A.: Large-Scale Controlled Methane Releases for Satellite-Based Detection and Emission Quantification of Point-Sources, Ph.D. thesis, 2025.

Reuland, F., Adams, T., Galvin, K., Kort, E. A., and Brandt, A. R.: Large-Scale Controlled Methane Releases for Satellite-Based Detection and Emission Quantification of Point-Sources, https://doi.org/10.25740/qh001qt3946, 2025.

Roger, J., Guanter, L., Gorroño, J., and Irakulis-Loitxate, I.: Exploiting the Entire Near-Infrared Spectral Range to Improve the Detection of Methane Plumes with High-Resolution Imaging Spectrometers, Atmospheric Measurement Techniques, 17, 1333–1346, https://doi.org/10.5194/amt-17-1333-2024, 2024.

Schaum, A.: A Uniformly Most Powerful Detector of Gas Plumes against a Cluttered Background, Remote Sensing of Environment, 260, 112 443, https://doi.org/10.1016/j.rse.2021.112443, 2021.

Thompson, D. R., Leifer, I., Bovensmann, H., Eastwood, M., Fladeland, M., Frankenberg, C., Gerilowski, K., Green, R. O., Kratwurst, S., Krings, T., Luna, B., and Thorpe, A. K.: Real-Time Remote Detection and Measurement for Airborne Imaging Spectroscopy: A Case Study with Methane, Atmospheric Measurement Techniques, 8, 4383–4397, https://doi.org/10.5194/amt-8-4383-2015, 2015.

Thorpe, A. K., Green, R. O., Thompson, D. R., Brodrick, P. G., Chapman, J. W., Elder, C. D., Irakulis-Loitxate, I., Cusworth, D. H., Ayasse, A. K., Duren, R. M., Frankenberg, C., Guanter, L., Worden, J. R., Dennison, P. E., Roberts, D. A., Chadwick, K. D., Eastwood, M. L., Fahlen, J. E., and Miller, C. E.: Attribution of Individual Methane and Carbon Dioxide Emission Sources Using EMIT Observations from Space, Science Advances, 9, eadh2391, https://doi.org/10.1126/sciadv.adh2391, 2023.

Varon, D. J., Jacob, D. J., McKeever, J., Jervis, D., Durak, B. O. A., Xia, Y., and Huang, Y.: Quantifying Methane Point Sources from Fine-Scale Satellite Observations of Atmospheric Methane Plumes, Atmospheric Measurement Techniques, 11, 5673–5686, https://doi.org/10.5194/amt-11-5673-2018, 2018.

Zhang, X., Maasakkers, J. D., Roger, J., Guanter, L., Sharma, S., Lama, S., Tol, P., Varon, D. J., Cusworth, D. H., Howell, K., Thorpe, A. K., Brodrick, P. G., and Aben, I.: Global Identification of Solid Waste Methane Super Emitters Using Hyperspectral Satellites, Environmental Science & Technology, 59, 18 134–18 145, https://doi.org/10.1021/acs.est.4c14196, 2025.