



Evaluating metre-scale methane plume simulations with MicroHH against AVIRIS-4 remote sensing observations

Sandro Meier^{1,2}, Dominik Brunner¹, and Gerrit Kuhlmann¹

¹Empa, Laboratory for Air Pollution / Environmental Technology, Ueberlandstrasse 129, 8600 Duebendorf, Switzerland

²Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

Correspondence: Sandro Meier (sandro.meier@empa.ch), Gerrit Kuhlmann (gerrit.kuhlmann@empa.ch)

Abstract. High-resolution imaging spectrometers resolve anthropogenic plumes at metre to sub-metre scales, creating opportunities for emission quantification but requiring models that reproduce turbulent plume morphology at comparable resolution. We evaluate the large-eddy simulation model MicroHH against Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-4) observations from 2024 controlled releases in France. Using site-specific wind, surface, and source parameters, we simulated the four releases with the longest plumes at 1 m horizontal resolution. Synthetic CH₄ column enhancements were evaluated using plume structure, turbulence spectra, Gaussian cross-section statistics, along-plume flux variability, and plume power spectral density. For one release with a pronounced double plume, we integrated the three-dimensional model field along the slant illumination and viewing paths to investigate parallax effects. MicroHH reproduced the expected inertial-subrange behaviour of turbulent kinetic energy spectra and observed plume width, intermittency, and spatial power spectra for most releases. However, modelled vertical columns were more Gaussian than the observations and did not reproduce the double plumes. The exploratory slant-path calculation generated a comparable double structure, indicating that observational geometry can explain part of this discrepancy. Other differences likely reflect rapid near-source dilution at finite grid resolution, the absence of online plume-rise and buoyancy treatments, and omitted obstacle effects on near-surface flow. Overall, MicroHH can generate realistic high-resolution synthetic plume imagery for retrieval evaluation and machine-learning applications when inputs are well constrained and observational geometry is considered. This first systematic multi-metric comparison of LES plume morphology with metre-scale imaging-spectrometer observations provides a validation framework for other high-resolution dispersion models.

1 Introduction

Recent years have seen a rapid increase in high-resolution airborne and spaceborne observations of anthropogenic point-source emissions of methane (CH₄) and other gases. Imaging spectrometers on satellites can now resolve plume structures at spatial scales of a few tens of metres. Similar spectrometers on airborne platforms can resolve plume structures at about ten times higher resolution down to metre and sub-metre scales. This makes it possible not only to infer source strengths from observed column enhancements, but also to attribute leaks to specific infrastructure. Together, these capabilities create new opportunities for quantifying emissions from localised sources using top-down methods. At the same time, higher spatial resolution makes



25 emission estimation more demanding. Processes such as turbulent mixing, wind shear, surface heterogeneity, and radiative transfer effects, including object shadows and parallax effects depending on viewing and illumination geometry, become more visible in the observations (e.g. Meier et al., 2026; Gorroño et al., 2026). Understanding and correcting for these effects requires physically based simulation tools that can reproduce plume evolution at the same scales that are now being observed. The increasing number of plume observations creates a second need for model simulations: Large observational datasets
30 require automated approaches that can identify emission sources, segment plumes, and estimate emissions with minimal human intervention. Machine-learning (ML) methods are increasingly used for this purpose (e.g. Jongaramrungruang et al., 2022; Joyce et al., 2023; Dumont Le Brazidec et al., 2023; Rouet-Leduc and Hulbert, 2024; Bruno et al., 2024; Ouerghi et al., 2025; Plewa et al., 2025), but they require large labelled training datasets that are difficult to obtain from observations alone because source strengths and plume masks are often subjective, uncertain or unavailable. Many recent ML models are therefore
35 trained on synthetic data generated by computational fluid dynamics (CFD) simulations. Their usefulness depends directly on the realism of the synthetic plumes used for training. If the CFD model does not represent the source characteristics, plume morphology, and variability well, the resulting retrieval can inherit systematic biases. Such differences between synthetic training data and real observations create a so-called domain gap that can reduce model performance on real imagery. This gap may be reduced through more realistic simulations or addressed explicitly using domain-adaptation methods, but both
40 approaches require characterising the discrepancies between synthetic and observed plumes (Ganin and Lempitsky, 2015; Tzeng et al., 2017; Liu and Zhang, 2022). Therefore, validating synthetic plume imagery against observations is crucial. Despite this need, systematic validation of synthetic plume realism against real high-resolution observations remains rare, and no established set of reference metrics or acceptance criteria currently exists for this purpose. Existing synthetic-plume studies have often tested downstream retrieval or detection workflows rather than validating the synthetic plume field itself.
45 Varon et al. (2018) used WRF-LES CH₄ plumes at 50 m resolution to evaluate GHGSat-style source-rate retrieval methods, and Gorroño et al. (2023) inserted WRF-LES CH₄ plumes into Sentinel-2 scenes to benchmark detection limits and flux-rate uncertainty across surface types. These studies demonstrated the value of synthetic plume imagery for downstream applications, but their evaluation metrics were detection thresholds, flux errors, wind sensitivity, and plume-mask uncertainty, rather than independent comparisons of plume morphology and spatial statistics against observed plumes. Direct model–observation
50 comparisons are also limited in scope. Krol et al. (2024) compared chemistry-enabled MicroHH NO₂ plumes with Sentinel-5P/TROPOMI observations, including comparisons after mapping simulated fields to TROPOMI pixels. However, their focus was on NO_x chemistry, NO₂ lifetimes, and NO_x:NO₂ conversion, with pixel-wise comparisons used mainly to support that chemical and emission analysis. This study therefore does not eliminate the need for a systematic evaluation of the two-dimensional morphology and spatial statistics resolved by imaging spectrometers. This is a recognised gap in the field: Bue
55 et al. (2026) showed that ML plume detectors trained solely on WRF-LES-generated plumes generalise poorly to real observations, attributing this to unquantified structural differences between simulated and observed plumes. Establishing whether, and how well, high-resolution LES output reproduces observed plume structure is therefore a necessary prerequisite for using such output with confidence, independent of the specific downstream application.



Previous studies have used a range of atmospheric and dispersion models to generate synthetic plume data, including Lagrangian particle dispersion models such as ARTM (Hanfland et al., 2024) and mesoscale Eulerian models such as WRF (Joyce et al., 2023; Wolff et al., 2021; Brunner et al., 2023), COSMO-GHG (Brunner et al., 2019; Kuhlmann et al., 2019, 2021; Brunner et al., 2023), ICON (Brunner et al., 2023), and LOTOS-EUROS (Koene and Brunner, 2023). These models are valuable for large-scale transport and inversion studies (e.g. Brunner et al., 2023), but they are typically designed and configured for grid resolutions from kilometres to hundreds of metres, and their dynamical cores, parameterisations, and forcing strategies are therefore not optimised for resolving plume morphology at metre to sub-metre scales.

Large-eddy simulation (LES) models provide a suitable framework for the simulation of turbulent flows because they explicitly resolve the dominant turbulent eddies that control plume dispersion while parameterising only the smaller scales. This allows plume morphology to be represented at high spatial resolution while remaining computationally more tractable than direct numerical simulation (DNS). Among LES models, MicroHH has proven to be a suitable candidate because of its computational efficiency and scalability. The model was designed for modern high-performance computing architectures, including efficient parallelisation on CPUs and GPUs, making sub-metre-scale simulations over larger domains feasible (van Heerwaarden et al., 2017). MicroHH was introduced as a computational fluid dynamics model for DNS and LES of atmospheric boundary-layer flows. Its dynamical core was validated against benchmark cases covering neutral, stable, and convective regimes, as well as dry and moist conditions (van Heerwaarden et al., 2017). Since then, the model has been validated against mobile in situ trace-gas measurements (Ražnjević et al., 2022b), used to evaluate emission estimation methods from channel-flow simulations (Ražnjević et al., 2022a), adapted with an immersed-boundary method to resolve a debris-covered glacier (Bonekamp et al., 2020), applied to cloud–radiation interactions with 3D radiative transfer (Tijhuis et al., 2024), used to analyse convective systems (Tijhuis et al., 2025), and applied to quantify NO_x emissions from power plants (Meier et al., 2024; Hakkarainen et al., 2024), cities (Mols et al., 2026), and ships (Riess et al., 2025). Together, these studies show that MicroHH can represent atmospheric convection, transport, chemistry, radiative processes, and heterogeneous surface forcing at spatial and temporal scales relevant to high-resolution plume observations.

Although MicroHH has been compared with observed plumes, for example from Sentinel-5P or in situ field measurements, these comparisons have not established whether it can reproduce the spatial statistics and morphology revealed by high-resolution plume imagery. This distinction is important because kilometre-scale observations smooth turbulent fluctuations and primarily constrain bulk plume properties, whereas metre-scale imagery resolves fine-scale structures shaped by turbulence, wind shear, surface heterogeneity, and obstacles. Agreement at coarse resolution therefore does not demonstrate that a model can reproduce realistic metre-scale plume morphology.

To address this gap, we use MicroHH to generate and evaluate synthetic plume imagery at 1 m resolution. We perform simulations of a controlled CH₄ release experiment in France in 2024 and compare the resulting synthetic plumes with observed CH₄ column enhancements acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-4). We evaluate the simulations qualitatively and quantitatively by analysing plume statistics such as width, intermittency, spatial variability, and power spectral density. This combined approach allows us to assess MicroHH's capability to reproduce metre-scale plume structures, which is important for improving emission estimation methods and for training realistic ML models.



2 Data and methods

95 This section describes the observational data, modelling framework, and analyses used to evaluate the turbulent plume simulations. Section 2.1 presents the high-resolution CH₄ observations from AVIRIS-4. Section 2.2 describes the MicroHH model and its configuration for the controlled-release experiment, and Section 2.3 outlines the initial and boundary conditions, including the treatment of plume rise. Section 2.5 describes the derivation of synthetic CH₄ slant column densities used to assess viewing-geometry effects. Finally, Section 2.4 introduces the complementary metrics used for the quantitative comparison:
100 spectral analyses of the turbulence and CH₄ fields and cross-sectional statistics of plume morphology.

2.1 AVIRIS-4 CH₄ observations

The CH₄ column enhancements used in this study were taken from a controlled CH₄ release experiment conducted between 16 and 20 September 2024 at the TotalEnergies Anomalies Detection Initiatives (TADI) site in Lacq, France (McManemin et al., 2026b, a). Release rates ranged from 0.02 to 350 kg CH₄ h⁻¹, and releases were performed at different locations and heights
105 (0.01–6.5 m above ground level), with each release lasting 45 minutes. Wind conditions were measured with a ZX 300 Doppler wind lidar, which provided vertical profiles of wind speed and direction up to 300 m above ground at 20 s resolution. A wind anemometer mounted at 1.5 m next to the lidar system was used to correct wind directions. High-resolution plume observations were acquired with the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-4), flown at multiple altitudes between 1000 and 3700 m above ground, which yielded across-track spatial resolutions of 0.5–2.0 m. The along-track resolution was
110 0.4 m, independent of flight altitude. CH₄ column enhancements were retrieved with an iterative matched-filter approach, with additional corrections for surface albedo effects (Foote et al., 2020; Kuhlmann et al., 2025). A detailed description of the data processing and emission estimation is provided by Meier et al. (2026). The retrieved CH₄ column enhancements were remapped to the MicroHH output grid using Climate Data Operators (CDO) conservative remapping, enabling a direct comparison between observed and modelled plumes. In total, 12 releases between 1.45 and 290 kg CH₄ h⁻¹ were observed with
115 AVIRIS-4. However, only the four releases with the highest emission rates (52.9, 56.7, 80.1, and 290 kg CH₄ h⁻¹) had plume lengths exceeding 10 m. Because plume morphology and statistics are more meaningful for longer plumes, the analysis focuses on these four releases.

2.2 MicroHH model description and setup

We used the CFD model MicroHH in large-eddy simulation (LES) mode, which is able to resolve turbulent atmospheric flows
120 at high spatial resolution (van Heerwaarden et al., 2017). The model solves the filtered Navier–Stokes equations using finite-difference methods and represents subgrid-scale turbulence with a Smagorinsky–Lilly model (van Heerwaarden et al., 2017).

For the simulations presented here, advection was computed using a second-order scheme with sixth-order interpolation, and the Smagorinsky coefficient was set to 0.16 according to field measurements in Kleissl et al. (2006). A flux limiter was applied to the passive CH₄ scalar to ensure monotonic transport. Surface-atmosphere exchange was represented using a Monin-
125 Obukhov similarity theory (MOST) surface model, with no-slip boundary conditions at the lower boundary and free-slip



conditions at the domain top. Boundary conditions for scalar quantities such as temperature and CH_4 were prescribed as surface fluxes and zero-gradient (Neumann) conditions at the upper boundary. Lateral boundaries were periodic for momentum and thermodynamic variables to minimise the influence of the boundary conditions on the simulated turbulence within the domain, whereas outflow conditions were applied for CH_4 to prevent recirculation of the plume.

130 CH_4 emissions were represented as time-dependent Gaussian point sources with a standard deviation of $0.25 \text{ m} \times 0.25 \text{ m} \times 0.125 \text{ m}$, located at the release heights listed in Table 1.

The modelled domain covered a horizontal area of $1024 \text{ m} \times 1024 \text{ m}$ with a resolution of $1 \text{ m} \times 1 \text{ m}$. The vertical extent was discretised with 512 grid levels up to 2038 m, using a vertically stretched grid. This resulted in a resolution of 0.5 m near the surface, increasing to 13.7 m at the top of the domain. Surface roughness lengths were set to 0.2 m for momentum and heat,
 135 representative of grassland conditions. Dry thermodynamics with an anelastic approximation was used, with time-dependent surface pressure and temperature profiles prescribed using ECMWF’s ERA5 reanalysis data (Hersbach et al., 2018).

To mimic the conditions during the controlled release experiment, we applied time-dependent large-scale forcing and nudging to the horizontal wind velocity and temperature fields, using the data described in Section 2.3. Time-dependent nudging was applied with a relaxation timescale of 15 minutes.

140 Each simulation covered 2 hours and used a fourth-order Runge–Kutta time-integration scheme with adaptive time-stepping. This included a 75-minute spin-up phase followed by a 45-minute release period. Model output was sampled every 20 s, and synthetic CH_4 column enhancements were calculated as density-weighted vertical integrals of the CH_4 mixing ratios. Gaussian noise with a standard deviation of 280 ppm m, or $1.84 \times 10^{-4} \text{ kg m}^{-2}$, was added to mimic the noise characteristics of the AVIRIS-4 retrievals coarsened to the MicroHH grid (Meier et al., 2026).

145 **2.3 MicroHH initial and boundary conditions**

The analysis of wind speed and direction in Meier et al. (2026) showed that ERA5 reanalysis winds (Hersbach et al., 2018) agreed poorly with the on-site wind lidar data. The ERA5 wind direction was generally inconsistent with the observed plume orientation. This may be related to the coarse spatial resolution of ERA5, which does not capture the fine details of the topography at the site that influence local wind conditions. We therefore used the on-site wind lidar data to prescribe the
 150 initial and boundary conditions for the MicroHH simulations. We applied a 15-minute rolling mean to the raw wind data to remove high-frequency variability, then resampled the resulting data to 1 min resolution. We also smoothed the wind lidar data vertically using a rolling average over three height levels to reduce rapid vertical fluctuations that could introduce artificial turbulence in the model. The lidar-derived wind direction showed poor agreement with the plume orientation, likely because the data processing did not properly account for direction ambiguity when averaging the raw data to 20 s averages (Knoop et al., 2021). Therefore, we applied the wind direction measured by the anemometer to the full vertical profile while keeping the absolute wind speed measured by the lidar. Finally, we extrapolated the wind data from the highest lidar level to the top of the model domain, assuming constant wind speed and direction. The resulting wind speed at the top of the domain closely agreed with ERA5 wind speeds (Hersbach et al., 2018). Temperature profiles and surface pressure were taken from ERA5 using the $(LS)^2D$ Python package (van Stratum et al., 2023).



Table 1. Parameters of the controlled CH₄ releases used in this study, sorted by date and time.

Date	Start time [UTC]	End time [UTC]	CH ₄ flux [kg h ⁻¹]	Release height [m a.g.l.]	Internal diameter [m]	Outlet ejection speed [m s ⁻¹]	Plume rise [m]
2024-09-16	09:59	10:47	56.70	1.09	0.0221	57.4	1.55
2024-09-18	09:44	10:33	52.94	1.82	0.0221	53.6	0.00
2024-09-18	13:30	14:17	290.00	1.28	0.0779	23.6	2.50
2024-09-19	12:44	13:31	80.10	1.28	0.0525	14.4	0.63

160 To reproduce the controlled release experiment, we prescribed the emissions in the model using the provided source locations, CH₄ fluxes, and release heights. Because the releases had considerable ejection speeds, momentum-driven plume rise was not negligible. We estimated the plume rise Δh using the Briggs plume-rise model with separate formulations for the near field and the far field. The model describes plume rise as a function of downwind distance x , wind speed u , source diameter d , and gas ejection speed v_g (Briggs, 1969).

165
$$\Delta h_{\text{near}} = 1.44 \cdot \frac{(dv_g)^{2/3} \cdot x^{1/3}}{u^{2/3}} \quad \text{and} \quad \Delta h_{\text{far}} = 3 \cdot \frac{dv_g}{u}. \quad (1)$$

The final plume rise was calculated as $\Delta h_{\text{final}} = \min(\Delta h_{\text{near}}, \Delta h_{\text{far}})$ and added to the release heights. The release characteristics are listed in Table 1.

2.4 Quantitative metrics of plume similarity

170 Atmospheric turbulence inherently limits direct comparisons between models and observations because of its stochastic behaviour (Lorenz, 1969). A single aircraft overpass captures only one instantaneous plume realisation, as does a single model output, and the two therefore represent different samples from a stochastic distribution. Even when the model captures the governing transport processes well, the exact plume structures are not expected to agree point by point. It is therefore more meaningful to compare statistical characteristics and scale-dependent variability than the precise locations of individual plume features (Brunner et al., 2023). Temporal averaging, as used for example by Ražnjević et al. (2022b), was not suitable here

175 because the simulated plume direction varied substantially over time.

We therefore evaluate plume similarity using complementary spectral and morphology-based metrics. Turbulent kinetic energy (TKE) spectra test whether the model reproduces the turbulent energy distribution and temporal wind variability that drive plume dispersion, independently of the CH₄ field. CH₄ power spectra quantify how plume variance is distributed across spatial scales and test the balance between large-scale structure and fine-scale fluctuations without requiring individual features to be

180 spatially aligned. Finally, cross-sectional statistics in a plume-following coordinate system provide physically interpretable measures of plume shape, width, displacement, amplitude, and intermittency (Gałkowski et al., 2025). Together, these metrics assess both the turbulent forcing and how its scale-dependent variability is expressed in plume geometry and concentration. Because no standard set of validation metrics for high-resolution plume simulations has previously been established in the



literature, we designed this combination of metrics as a general-purpose evaluation framework that can be applied beyond the
 185 present case to other LES or CFD models and future high-resolution observational datasets.

2.4.1 Spectral analysis of turbulence

TKE spectra were computed from the three wind components (u, v, w) at each vertical level. The staggered MicroHH velocity
 components were first collocated onto the scalar grid. For each horizontal plane and wind component, the horizontal mean was
 removed before applying a two-dimensional fast Fourier transform (FFT). The resulting Fourier coefficients were converted
 190 to a discrete, variance-preserving spectrum using $|\hat{u}'|^2 / (N_x N_y)^2$, where N_x and N_y are the numbers of grid points in the
 two horizontal directions. This normalisation follows from Parseval's theorem for the unnormalised forward FFT convention:
 summing the spectrum over all two-dimensional Fourier modes recovers the variance of the corresponding mean-removed
 velocity fluctuation field (Durran et al., 2017). The TKE spectrum was then computed as

$$\Phi_{\text{TKE}}(k_x, k_y) = \frac{1}{2} (\Phi_{uu} + \Phi_{vv} + \Phi_{ww}).$$

195 Here, k_x and k_y denote spatial frequencies in cycles per metre and Φ the component-wise spatial spectra. The two-dimensional
 spectrum was reduced to a radial spectrum by binning the Fourier modes into annuli in (k_x, k_y) space and summing the modal
 variance contributions within each annulus.

Temporal TKE spectra were computed from wind-component time series at each vertical level. For the model output, the
 time series were extracted at the central grid point used for the analysis, $(x, y) = (512, 512)$. For the lidar data, the observed
 200 vertical wind profiles were used directly over the corresponding case period. For each component, a one-dimensional Welch
 power spectral density was estimated with 128-sample segments and linear detrending of each segment. The temporal TKE
 spectrum was calculated as

$$\Phi_{\text{TKE}}(f) = \frac{1}{2} (\Phi_{uu}(f) + \Phi_{vv}(f) + \Phi_{ww}(f)).$$

With the Welch density scaling, integration over frequency approximates the variance of the velocity fluctuations (Welch,
 205 1967).

For both the spatial and temporal wind spectra, only model levels between 0 and 300 m above ground were retained, matching
 the altitude range sampled by the wind lidar.

2.4.2 Spectral analysis of CH₄ plumes

The spatial power spectral density (PSD) of MicroHH and AVIRIS-4 CH₄ plumes was computed from column density maps
 210 within the detected plume mask, which was cut to a plume length of 50 m and a width of 75 m (see Section 2.4.3). The
 spatial mean within the plume mask was removed to isolate spatial fluctuations and prevent the zero-wavenumber component
 associated with the mean plume enhancement from dominating the spectrum, and missing values in AVIRIS-4 images were
 replaced with zero. A two-dimensional Hann window $h(x, y)$ was applied before the FFT to reduce spectral leakage. The



two-dimensional PSD density was calculated as

$$215 \quad \Phi(k_x, k_y) = \frac{|\mathcal{F}\{c'(x, y)h(x, y)\}|^2 \Delta x \Delta y}{N_x N_y \langle h^2 \rangle},$$

where c' denotes the mean-removed plume field, N_x and N_y denote the numbers of pixels along the x and y directions, respectively, $\Delta x = \Delta y = 1$ m is the grid spacing, and $\langle h^2 \rangle$ is the mean squared amplitude of the Hann window. The corresponding Fourier-mode spacings are $\Delta k_x = 1/(N_x \Delta x)$ and $\Delta k_y = 1/(N_y \Delta y)$. Each Fourier-grid cell therefore covers an area $\Delta k_x \Delta k_y$ in wavenumber space, and

$$220 \quad \Phi(k_x, k_y) \Delta k_x \Delta k_y$$

is the contribution of that mode to the spatial variance. Consequently, summing over all modes approximately recovers the variance of the mean-removed plume field:

$$\text{Var}(c') \approx \sum_{k_x, k_y} \Phi(k_x, k_y) \Delta k_x \Delta k_y.$$

To obtain a one-dimensional spectrum, the Fourier modes were grouped into annular bins. For an annulus A_j of width Δk_j centred on k_j , the radial spectrum was defined as

$$\Phi_{\text{rad}}(k_j) = \frac{1}{\Delta k_j} \sum_{(k_x, k_y) \in A_j} \Phi(k_x, k_y) \Delta k_x \Delta k_y.$$

Thus, $\Phi_{\text{rad}}(k_j) \Delta k_j$ is the total variance contributed by all Fourier modes in annulus A_j . Summing these contributions over the radial bins, or equivalently integrating $\Phi_{\text{rad}}(k)$ over k , recovers the variance represented by the two-dimensional PSD.

2.4.3 Plume statistics

230 For the morphology-based analysis, we used the *ddeq* Python package to calculate a plume-following coordinate system for each plume. We detected plumes with a threshold of 0.3 g m^{-2} and fitted a centreline to the detected plume pixels using a Bézier curve with two control points. This centreline defined a local coordinate system for the detected pixels, with the along-plume direction given by the tangent to the centreline and the across-plume direction by the normal to the centreline (Kuhlmann et al., 2024). Following centreline fitting, we selected the first 50 m of each plume for the analysis. To ensure consistency, the
 235 same plume-detection and centreline-fitting parameters, including the detection method and threshold, number of centreline nodes, and maximum plume length and width, were applied to the observed and modelled plumes.

Next, we divided the plumes into 3 m-wide cross-sections and fitted a Gaussian profile to each cross-section (Kuhlmann et al., 2024):

$$g(y) = \frac{q}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) + my + b. \quad (2)$$

240 where y is the across-plume coordinate, σ is the standard deviation, μ is the mean of the Gaussian, and $my + b$ is a linear background term. To assess whether the modelled plumes reproduce the observed plume morphology, we compared the plume

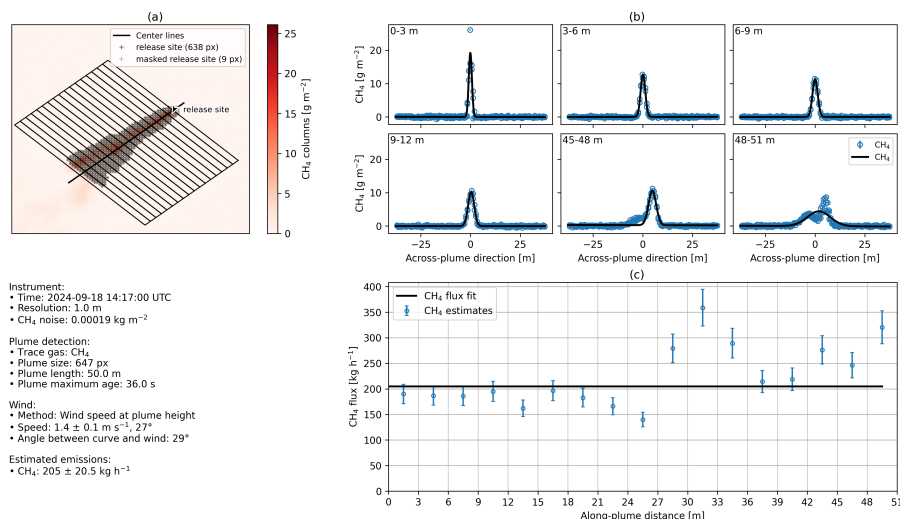


Figure 1. (a) Simulated CH₄ columns for one snapshot of the 290 kg CH₄ h⁻¹ release, (b) cross-sections of the plume along the fitted centreline and Gaussian fits for each cross-section, and (c) along-plume CH₄ flux calculated from the Gaussian fits and wind speed at the source height.

width σ , plume meandering μ , R^2 values of the Gaussian fits, and plume amplitude $A = q/(\sqrt{2\pi}\sigma)$, where q is the integral of the Gaussian fit. The coefficient of determination, R^2 , was computed from the residual and total sums of squares as $1 - SS_{\text{res}}/SS_{\text{tot}}$. We also calculated the CH₄ flux along the plume to compare plume intermittency. For this calculation, q was multiplied by the horizontal-domain-averaged wind speed at the prescribed source height (see Table 1). These statistics were calculated only for cross-sections with $R^2 \geq 0.5$ to ensure that the Gaussian fit was meaningful. Figure 1 illustrates the cross-section fitting and flux calculation.

Finally, we compared the spatial variability of the observed and modelled plumes by calculating the two-dimensional power spectral density (PSD) of the CH₄ column enhancements as described in Section 2.4.2.

For each release, the simulation covered the full 45-minute release period at 20 s intervals, resulting in 135 model snapshots, whereas the observations consisted of seven to eight AVIRIS-4 overpasses. After quality filtering, the number of valid AVIRIS-4 observations additionally varied among along-plume cross-sections. We first compared the mean and variability calculated from all available modelled and observed data. To assess the influence of the unequal sample sizes, we performed a sample-size-matching sensitivity analysis. At each along-plume position, we randomly sampled valid MicroHH snapshots without replacement to match the number of valid AVIRIS-4 observations. We then recalculated the mean and sample standard deviation using the matched datasets. The resulting comparison is shown in Figures A1 and A2.



2.5 Accounting for 3D parallax effects in AVIRIS-4 CH₄ column densities

As an exploratory sensitivity analysis, we derived geometry-projected CH₄ column enhancements (GPCDs) for the AVIRIS-4 acquisition on 18 September 2024 at 13:33 UTC. This calculation was motivated by the double plumes observed during the controlled-release experiment, which were attributed to parallax effects (Meier et al., 2026). Vertical column densities (VCDs), obtained by vertical integration, cannot capture this effect because they neglect the horizontal displacement of the plume signal along the illumination and viewing paths (Schwaerzel et al., 2020). Accounting for the slant geometry was therefore necessary to assess whether it could reproduce the observed double-plume morphology.

We restricted this analysis to the 13:33 UTC acquisition because it exhibited a clear double-plume structure and the modelled and observed plumes were closely aligned. During the other acquisitions, rapidly changing wind directions often caused the modelled and observed plume directions to diverge. Close alignment is essential because it exhibited a clear double-plume structure and the modelled and observed plumes were closely aligned. This alignment matters because the morphology produced by the slant-path calculation is highly sensitive to differences in plume position and orientation. Misalignment may generate an apparent double plume where none was observed or obscure a double plume that was present. We therefore treated the GPCD comparison as an exploratory case study and excluded it from the systematic plume evaluation.

The three-dimensional CH₄ mass-mixing-ratio fields were converted to CH₄ mass densities using the MicroHH reference air-density profile. The incoming light path was defined by the solar zenith angle (SZA = 47°) and solar azimuth angle (SAA = 214°), whereas the outgoing path was assumed to be nadir viewing. For each horizontal grid cell, we integrated the CH₄ mass density along the horizontally displaced incoming and outgoing rays to obtain the solar- and viewing-path slant columns, L_s and L_v , respectively. Normally, unbiased path-wise column contributions would be obtained by dividing each path integral by its respective geometric air-mass factor:

$$C_{\text{unbiased}} = \frac{L_s}{\sec(\text{SZA})} + \frac{L_v}{\sec(\text{VZA})}. \quad (3)$$

The AVIRIS-4 matched-filter retrieval, however, assumes that light traverses the plume along both the incoming and outgoing paths (Meier et al., 2026). We reproduced this retrieval assumption by calculating

$$C_{\text{MF}} = \frac{L_s + L_v}{\sec(\text{SZA}) + \sec(\text{VZA})}. \quad (4)$$

to generate a synthetic observation that is consistent with AVIRIS-4 observations.

3 Results

This section evaluates the MicroHH simulations against the AVIRIS-4 observations. Section 3.1 first compares the observed and modelled CH₄ plume morphology visually, and Section 3.2 assesses the resolved turbulence using spatial and temporal TKE spectra. Sections 3.3–3.6 then quantify plume structure using Gaussian cross-section fits, along-plume flux, meandering, and plume power spectral density. Finally, Section 3.7 assesses the influence of viewing geometry by comparing the observed column enhancements with the modelled vertical and slant column densities and their Gaussian-fit statistics.

3.1 Visual comparison

Figure 2 compares CH_4 column enhancements for one of the AVIRIS-4 overpasses during the $290 \text{ kg CH}_4 \text{ h}^{-1}$ release (top row, red) with six consecutive MicroHH snapshots around the overpass time (lower rows, blue). The observed plume is elongated near the source and fans out farther downwind. This structure is reproduced well by the model snapshots at 13:55:20 UTC and 13:56:20 UTC, which show a similar plume width and a compact initial plume. The other snapshots show stronger turbulent variability, illustrating how plume morphology can change over short time periods. This example highlights the difficulty of comparing a single model snapshot with a single observed plume. Similar plots for all AVIRIS-4 overpasses and corresponding MicroHH snapshots are shown in the Supplementary Material (Figures S1–S60).

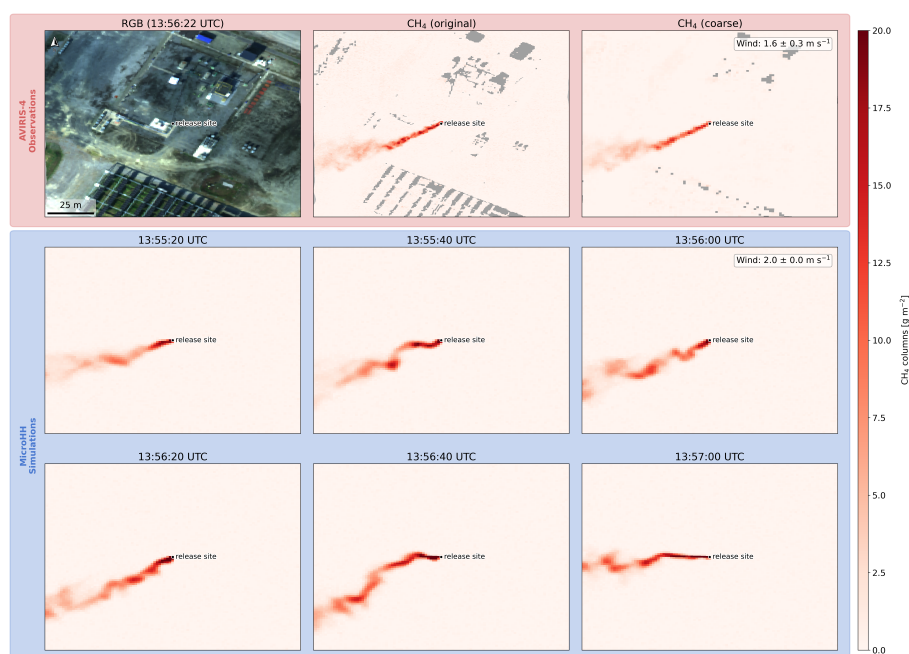


Figure 2. The top row in red shows the RGB scene at the AVIRIS-4 overpass time, the retrieved CH_4 column enhancement at native resolution, and the same retrieval coarsened to the MicroHH grid. The lower panels in blue show MicroHH-simulated CH_4 column enhancements at 20 s intervals around the overpass, illustrating the evolving plume morphology and transport relative to the observed plume.

3.2 Turbulence statistics

The upper row in Figure 3 shows that the mean spatial TKE spectra of the modelled wind fields at the final simulated time step closely follow the theoretical $-5/3$ slope within the inertial subrange (Kolmogorov, 1941; Kaimal et al., 1972). This indicates that MicroHH reproduces the expected turbulent energy cascade over the resolved inertial scales. At the lowest wavenumbers, the spectra roll off at a wavelength of approximately 500 m before reaching the energy-containing range. At the highest wavenumbers, the spectra roll off near 0.25 m^{-1} , which indicates that turbulent motions are only reliably resolved

down to wavelengths of about 4 m ($\lambda = 1/k$), while the formal Nyquist limit of the 1 m grid is 0.5 m^{-1} , corresponding to a two-grid-cell wavelength. The modelled temporal TKE spectra in the second row of Figure 3 agree well with the wind-lidar spectra for the 57, 53, and 80 kg $\text{CH}_4 \text{ h}^{-1}$ releases. For the 290 kg $\text{CH}_4 \text{ h}^{-1}$ release, the modelled spectra have slightly lower energy than the observed spectra at lower frequencies.

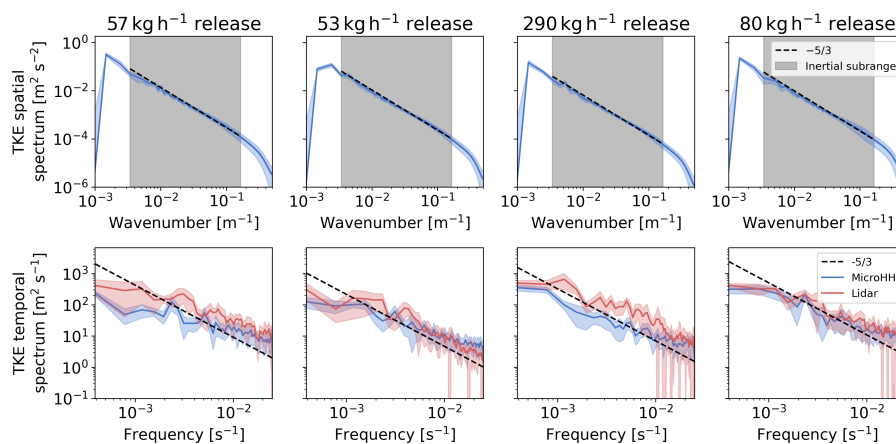


Figure 3. First row: vertically averaged modelled three-dimensional TKE power spectral density between 0 and 300 m above ground and the theoretical $-5/3$ slope (black dashed line) within the inertial subrange for the last simulated time step. Second row: comparison of the vertically averaged modelled temporal TKE spectrum (blue) with the observed TKE spectrum derived from wind lidar data over the same altitude range (red), including the theoretical $-5/3$ slope (black dashed line). Shading indicates the vertical standard deviation. Each column corresponds to one simulated release.

3.3 Gaussian fitting of across-plume cross-sections

Building on the visual comparison in Section 3.1, the Gaussian-fit statistics provide a quantitative comparison of the modelled and observed plume cross-sections.

The top row of Figure 4 shows that the modelled cross-sections generally have higher R^2 values and are therefore closer to Gaussian in shape than the observed cross-sections. The modelled R^2 values are highest near the source and decrease with along-plume distance, whereas the observed values are more variable, are often lowest near the source for the 290 and 80 kg $\text{CH}_4 \text{ h}^{-1}$ releases, and approach the modelled values farther downwind. The middle row shows that the fitted plume widths agree close to the source, but the observed plumes are generally wider and more variable along the plume. The bottom row shows that the fitted maxima decrease with along-plume distance in both datasets. The maxima agree well for the 57 and 53 kg $\text{CH}_4 \text{ h}^{-1}$ releases, whereas the modelled maxima are generally higher for the 80 and 290 kg $\text{CH}_4 \text{ h}^{-1}$ releases. When all available snapshots are included, the modelled statistics also appear less variable than the observed statistics, particularly for these latter two releases.



The systematic differences among all three metrics can largely be attributed to the parallax effect (Section 3.7). Viewing and illumination geometry can split the observed enhancement into two branches, producing a double-peaked cross-section with a lower R^2 , a greater fitted width, and a lower fitted maximum than the corresponding modelled vertical column. Farther downwind, the branches spread and overlap, making the observed cross-sections more Gaussian and causing their R^2 values to approach those of the modelled plumes. Thus, several apparent model–observation differences reflect both the observed double-plume morphology and the limitations of representing this structure with a single Gaussian.

The apparent lower variability and intermittency of the modelled plumes is largely a sampling effect. The comparison using all available data includes 135 model snapshots but only seven to eight observed overpasses, with as few as one valid observation at some along-plume positions after quality filtering. When MicroHH was randomly subsampled at each position to match the number of valid AVIRIS-4 observations, the means and variability of the Gaussian-fit statistics became much more comparable (Figure A1). The results therefore do not provide robust evidence that MicroHH systematically underestimates plume intermittency.

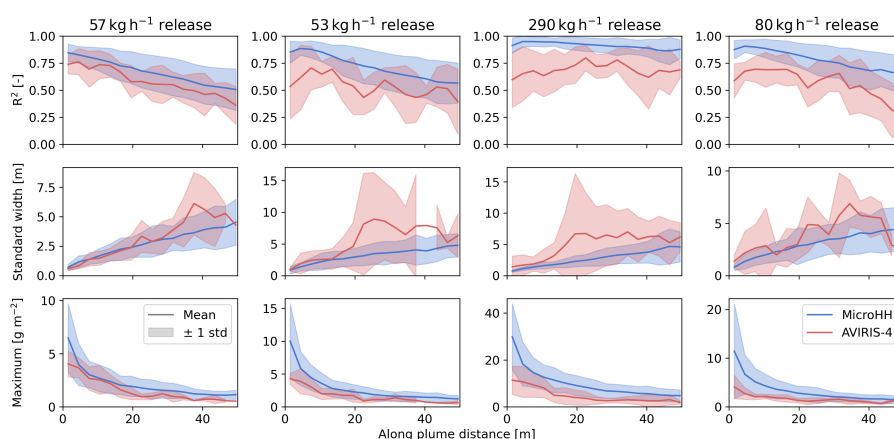


Figure 4. Gaussian-fit statistics for all across-plume cross-sections as a function of along-plume distance for modelled MicroHH (blue) and observed AVIRIS-4 (red) plumes. Rows show the R^2 values (top), standard widths (middle), and maxima (bottom) of the fitted Gaussian curves; columns correspond to the four controlled releases. Lines and shading indicate the mean and ± 1 standard deviation, respectively. This comparison uses all available snapshots. A sample-size-matched comparison is shown in Figure A1.

3.4 Along-plume flux

Figure 5 shows that the mean modelled flux remains nearly constant for all four releases, whereas the observed flux shows large variations along the plume. The apparent difference in flux variability is, however, strongly influenced by the unequal sample sizes. As mentioned in Section 3.3, the comparison between MicroHH and AVIRIS-4 observations includes 135 model snapshots, but only seven to eight observed overpasses. After MicroHH was subsampled at each position to match the number of valid AVIRIS-4 observations, the mean fluxes and their variability became much more comparable (Figure A2). Thus, the



full-sample comparison does not indicate a systematic underestimation of plume intermittency by MicroHH. The remaining offsets in mean flux, particularly for the $290 \text{ kg CH}_4 \text{ h}^{-1}$ release, may partly reflect differences in the wind speeds used to convert the integrated Gaussian enhancements to fluxes, as we used the wind speed at source height, neglecting plume rise.

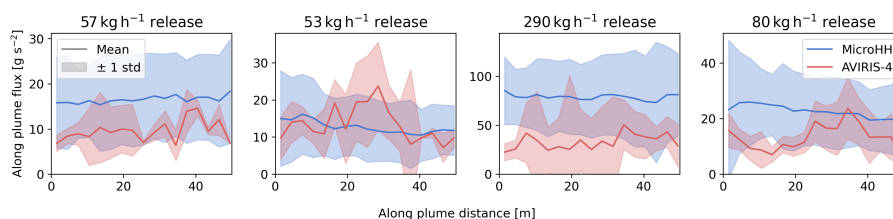


Figure 5. Computed CH_4 flux for all cross-sections along the plume for modelled (blue) and observed (red) plumes, using all available snapshots. A sample-size-matched comparison is shown in Figure A2.

3.5 Meandering

340 The lateral displacement μ of the centre of each Gaussian fit from the fitted plume centreline provides a measure of plume meandering and curvature. Figure 6 shows that the meandering distribution is generally less Gaussian for the observed plumes (upper row) than for the modelled plumes (lower row). This partly reflects the limited number of AVIRIS-4 overpasses, with seven to eight overpasses per release and 16 cross-sections per overpass. It also reflects the more complex morphology of the observed plumes, which leads to larger displacements from the fitted centreline. The fitted Gaussian distributions show that the

345 observed plumes have a larger absolute mean displacement, suggesting that they are more curved than the modelled plumes. This may partly reflect the smaller observational sample size: with only a few overpasses, the observed plumes can by chance be curved preferentially in one direction, whereas the larger set of model snapshots is more likely to include plumes curving in both directions, which compensates for the mean displacement. The difference may also be related to surface heterogeneity and obstacles near the release site (see Section 4). Because these effects are not explicitly represented in the current model

350 setup, they may increase the lateral variability of the observed plumes and contribute to the larger mean displacement. Overall, the comparison suggests that the observed plumes contain more irregular lateral structure than the modelled plumes.

3.6 Plume power spectral density

Figure 7 shows that the two-dimensional power spectral densities (PSDs) of the observed and modelled plumes have similar spectral shapes, including similar slopes in the inertial subrange between $2\Delta x$ and the plume length L of 50 m. For the 53, 290,

355 and $80 \text{ kg CH}_4 \text{ h}^{-1}$ releases, the modelled PSD values are slightly higher than the observed values, with the largest difference occurring for the $80 \text{ kg CH}_4 \text{ h}^{-1}$ release. These higher PSD values primarily reflect the larger amplitudes and spatial variance of the modelled CH_4 enhancements because spectral power is proportional to the squared Fourier amplitude. As the prescribed emission rates were identical to the measured release rates, the larger modelled amplitudes can be explained by the lower wind speeds in the simulations for these three releases (see Figures S45–S60). For the 290 and $80 \text{ kg CH}_4 \text{ h}^{-1}$ releases, parallax

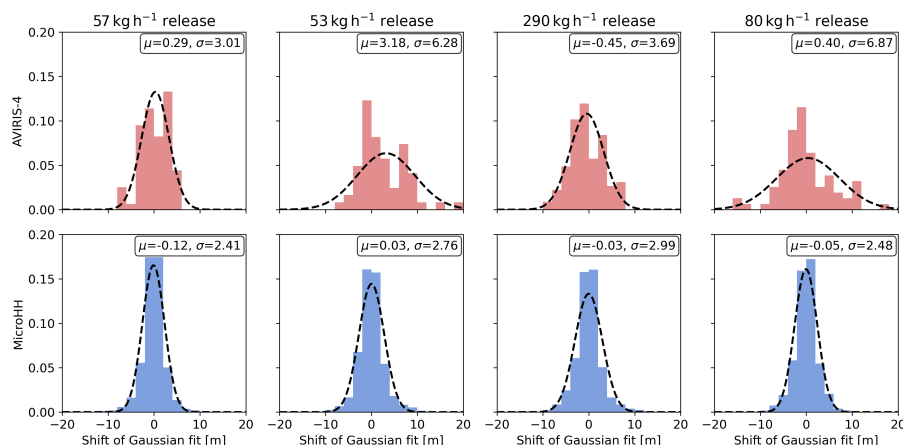


Figure 6. Distributions of the lateral displacement of the Gaussian-fit centres from the fitted plume centreline for observed (top row) and modelled (bottom row) plumes, calculated using all 3 m-wide cross-sections along the entire 50 m plume length. The dashed lines show a fitted Gaussian distribution for comparison.

effects in the observations additionally contributed to the differences in plume amplitude and PSD. Overall, the similar spectral shapes indicate that MicroHH reproduces the scale-dependent distribution of turbulent plume variability well, despite small differences in absolute variance.

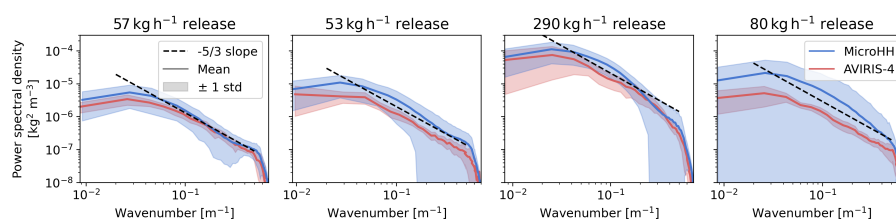


Figure 7. Power spectral density of detected observed and modelled plumes. The dashed line shows the theoretical $-5/3$ slope for the inertial subrange between $2\Delta x$ and the plume length L .

3.7 Influence of viewing geometry on plume statistics

Figure 8 compares the observed CH₄ column enhancements with the MicroHH-simulated CH₄ VCDs and GPCDs for the AVIRIS-4 overpass at 13:33 UTC on 18 September 2024. The observed plume shows a double-plume structure, which is not reproduced by the modelled VCDs. However, the derived GPCDs show a double-plume structure very similar to the observations, demonstrating that viewing geometry can account for much of the observed morphology. Moreover, the cross-sections along the fitted centreline in panels (b)–(e) show that the GPCDs reproduce the observed cross-sectional plume structure much better than the VCDs, which are too narrow and exhibit a higher peak than the observations.

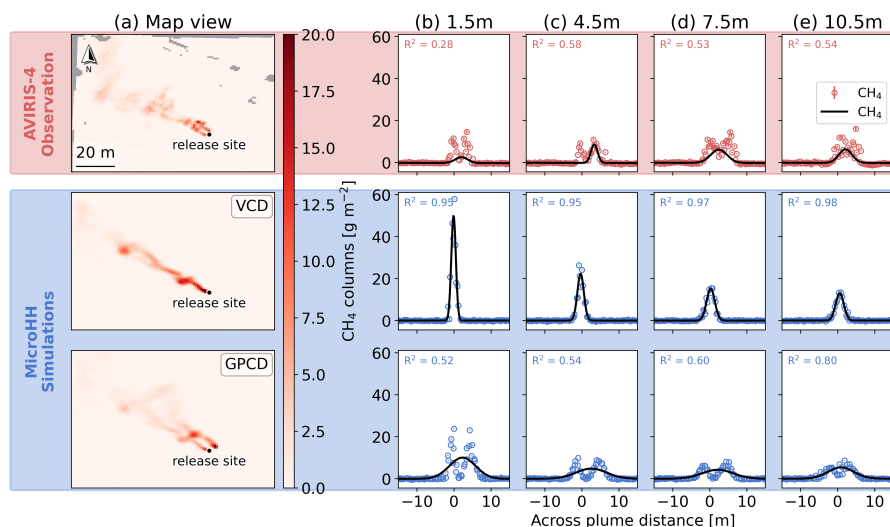


Figure 8. (a) Map view of the AVIRIS-4 overpass at 13:33 UTC on 18 September 2024, showing the retrieved CH₄ column enhancement (first row), the corresponding MicroHH-simulated CH₄ VCDs (second row), and the GPCDs (third row). (b)–(e) Cross-sections of the observed and modelled plumes along the fitted centreline for the first four cross-sections.

370 This is also evident in Figure 9, which displays the R^2 values, standard widths, and maxima of the fitted Gaussian curves for all cross-sections along the plume. In comparison to the VCDs, the GPCDs exhibit much better agreement with observations across all three metrics, highlighting the importance of viewing geometry in accurately reproducing plume morphology and statistics. Notably, the lower R^2 values of the GPCDs near the source support the hypothesis that the low observed R^2 values result from the double-plume structure. These values increase along the plume as it fans out and the two branches merge. Only
 375 the standard width of the GPCDs is overestimated. This may partly result from the simplified treatment of plume rise: the source height was prescribed as the physical release height plus the calculated final plume rise (Table 1), thereby representing plume rise as an instantaneous vertical displacement of the entire plume. This approximation is inaccurate close to the source, where the plume is still rising, and may bias the parallax-induced horizontal displacement and thus the apparent width of the GPCD plume.

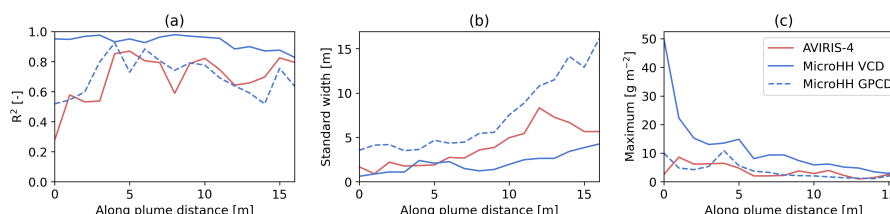


Figure 9. (a) R^2 values of Gaussian fits for all cross-sections along the observed (red) and modelled (blue) plumes for the AVIRIS-4 overpass at 13:33 UTC on 18 September 2024. (b) Standard widths of the Gaussian fits. (c) Maxima of the fitted Gaussian curves.



380 4 Discussion

The results highlight several important aspects of MicroHH’s ability to simulate metre-scale CH₄ plumes observed by AVIRIS-4. Although MicroHH has already been used to simulate turbulent plumes in previous studies (e.g., Ražnjević et al., 2022b), this study moves to a much finer spatial resolution and, to our knowledge, is the first to systematically evaluate LES-simulated plume morphology and spatial statistics against high-resolution two-dimensional imaging-spectrometer observations using a
 385 complementary suite of spectral and morphological metrics. This complements previous comparisons that validated single aggregate quantities, such as flux accuracy (Jongaramrungruang et al., 2019) or in situ concentrations measured during flights (Koene and Brunner, 2023; Brunner et al., 2023; Yuvaraj et al., 2026b). Overall, the model reproduced the main characteristics of turbulent plume dispersion, including spatial and temporal energy spectra and the overall plume morphology. The spatial TKE spectra agree with the theoretical $-5/3$ slope, indicating that MicroHH reproduces the expected turbulent energy cascade
 390 over the resolved inertial scales, while the temporal TKE spectra show that the model captures wind variability at the relevant scales, although slightly underestimating it for two of the releases. Taken together, these results suggest that, when driven by realistic boundary conditions and source parameters, MicroHH can generate synthetic datasets that resemble real observations at high spatial resolution.

The comparison using all available snapshots initially suggested that the modelled plumes had less variability and intermit-
 395 tency than the observed plumes. However, the sample-size-matched sensitivity analysis produced much more similar Gaussian-fit and along-plume flux statistics for MicroHH and AVIRIS-4 (Figures A1 and A2). The apparent difference in variability is therefore largely attributable to the much smaller observational sample rather than clear evidence that MicroHH systematically underestimates plume intermittency. This result highlights the importance of matching effective sample sizes when comparing statistics from large simulation ensembles with a limited number of observed plume realisations.

400 Achieving this level of realism, however, required careful attention to the input data. MicroHH can be driven with ERA5 reanalysis data, but Meier et al. (2026) showed that ERA5 wind fields were biased at this site and did not represent the local wind conditions well. On-site wind lidar data were therefore crucial for realistic plume simulations that could be compared to AVIRIS-4 observations. The controlled-release experiment also provided detailed source information, including release height, ejection speed, and source diameter, which helped to represent plume rise and the effective wind speed at the source.

405 The analysis of GPCDs shows that accounting for radiative and geometric effects can bring the MicroHH simulations even closer to the AVIRIS-4 observations. Integrating the three-dimensional CH₄ field along the solar illumination and sensor view-
 ing paths reproduced aspects of the observed double-plume structure that were absent from the vertical columns. An elevated plume intersects the incoming and outgoing rays at different horizontal locations such that their contributions can appear as two branches: a displaced plume-shadow branch along the incoming path and a source-plume branch along the reflected path to
 410 the sensor. When these branches are spatially separated, each is dominated by only one path contribution but is normalised by the total two-way air-mass factor (see Equation 4). Its apparent CH₄ enhancement is therefore underestimated, consistent with the geometric retrieval effect identified in the AVIRIS-4 observations (Meier et al., 2026) and illustrated in Figure B1. Consequently, differences between modelled vertical columns and retrieved images cannot be attributed to atmospheric transport



alone. These geometry-induced biases are difficult to remove from observations because doing so would require the unknown
415 plume height at each pixel. They can, however, be reproduced in synthetic imagery. For comparisons with AVIRIS-4 or other
imaging spectrometers, three-dimensional model fields should therefore be converted to GPCDs using the acquisition-specific
illumination and viewing geometry rather than compared as VCDs. Depending on the application, the forward modelling should
also represent surface reflectance, shadows, instrument response, and retrieval effects. Synthetic data generated this way can
expose downstream methods, including ML-based emission retrievals, to the same geometric biases as real observations and
420 thereby improve their evaluation and robustness.

Even after this post-processing, retrieved AVIRIS-4 images can contain systematic, scene-dependent biases that are unrelated
to atmospheric plume transport. For example, cast shadows can produce retrieval artefacts, while spectral features of some
surface materials can overlap or interfere with the CH₄ absorption signal and generate spatially structured biases (Meier
et al., 2026; Gorroño et al., 2026). Such effects are difficult to reproduce using fully synthetic backgrounds. For downstream
425 applications such as ML training and retrieval evaluation, post-processed MicroHH plumes should therefore be embedded into
real AVIRIS-4 background observations. This has also been done for ML applications with AVIRIS-NG, PRISMA, Sentinel-2,
and GHGSat imagery (Jongaramrungruang et al., 2022; Joyce et al., 2023; Rouet-Leduc and Hulbert, 2024; Bruno et al., 2024;
Ouerghi et al., 2025). Such hybrid images preserve realistic surface variability, noise, and retrieval artefacts while retaining
known plume properties, thereby helping to reduce the domain gap.

430 Importantly, however, realistic backgrounds alone do not guarantee that synthetic training data are representative of real
plumes. Bue et al. (2026), for example, trained convolutional neural network plume detectors exclusively on synthetic CH₄
plumes generated by WRF-LES. Although instance-detection F_1 scores approached 1.0 when evaluated on further WRF-LES-
generated plumes, they fell below 0.3 for real AVIRIS-NG observations, despite the synthetic plumes having been injected
into observed background scenes. The authors attributed this generalisation gap partly to unresolved temporal variability in
435 wind speed and direction during plume evolution and to the absence of diffuse area-source emissions in the training data.
They also noted that synthetic plume datasets used for ML are commonly evaluated against further synthetic data rather than
independently assessing whether the simulated plume structures themselves resemble real observations.

The multi-metric evaluation performed here addresses this issue directly for MicroHH. Rather than assuming that LES-
generated plumes are sufficiently realistic for downstream use, the comparison tests whether their spatial and spectral structure
440 is consistent with independent observations. This type of validation can identify deficiencies in processes such as plume-rise
dynamics, source representation, or turbulent dispersion before synthetic data are used for retrieval development, emission
quantification, or ML training, thereby reducing the risk that systematic model deficiencies are propagated into these applica-
tions.

Because no published reference values or acceptance thresholds currently exist for such systematic, multi-metric compar-
445 isons at metre scale, the metrics and results presented in Section 3 also provide an initial reference point for future intercom-
parisons of LES and CFD models and for assessing their suitability for specific downstream applications.



4.1 Limitations

The results demonstrate that MicroHH can already produce realistic metre-scale plume images, while also identifying several aspects of the present configuration that can be refined further. The modelled plumes are generally more Gaussian-shaped than the observed plumes. This is partly due to the source representation used here, which is at least one grid cell wide, whereas the actual source diameter was only 2.2–7.8 cm. Numerical diffusion can also contribute to plume widening. Resolving this source-scale structure directly would require a much finer grid, ideally closer to the size of the release equipment, and would therefore strongly increase both computational cost and model output size. More generally, a grid spacing of 1 m does not imply an effective resolution of 1 m, as indicated by the TKE spectrum, because numerical models resolve structures reliably only when they span multiple grid cells. Ideally, the simulations would therefore oversample the observational resolution by a factor of 2–4 and subsequently be coarsened to the observed resolution before comparison. However, AVIRIS-4 is commonly operated at flight altitudes of approximately 7000 ft above ground during operational campaigns, yielding an across-track resolution of about 1 m. The 1 m simulations used here are therefore already at the high-resolution end of what is required to represent typical operational AVIRIS-4 observations, and a moderately coarser model grid may be sufficient for many applications.

To keep the simulations computationally feasible at 1 m resolution, we used a domain size of $1024\text{ m} \times 1024\text{ m}$, which is smaller than that used in several previous MicroHH studies (e.g., Ražnjević et al., 2022b; Ražnjević, 2023; Krol et al., 2024). This choice represents a practical trade-off between domain size, spatial resolution, and output volume. Very large boundary-layer eddies can span a substantial fraction of the planetary boundary layer (PBL), so a larger or nested domain would be preferable to fully represent the largest scales of atmospheric variability. However, sensitivity tests using domains of $640\text{ m} \times 640\text{ m}$, $1280\text{ m} \times 1280\text{ m}$, $2560\text{ m} \times 2560\text{ m}$, and $5120\text{ m} \times 5120\text{ m}$ at 10 m resolution produced comparable turbulence statistics at the lowest wavenumbers for all but the smallest domain (see Figure C1). This suggests that the domain used here was sufficiently large to capture the main energy-containing part of the turbulence spectrum. Moreover, during the controlled-release experiment, the horizontal domain size was comparable to the vertical extent of the PBL (see Figure C2). Any larger-scale wind variability not explicitly resolved within the domain was partly represented through the time-dependent boundary conditions derived from smoothed wind lidar data. Because the largest atmospheric eddies are much larger than the observed plume extent, they are expected to have only a limited influence on the short-range plume morphology considered here.

The influence of both grid resolution and domain size can be mitigated through open-boundary and two-way nesting capabilities, which are currently under development for MicroHH. These extensions will complement the model's existing high-resolution capabilities by making it easier to combine larger domains with fine resolution in the region of interest without a proportional increase in computational cost.

Beyond adequate grid resolution and domain size, representing site-specific objects and topography is another important requirement for generating realistic turbulence and plume morphology at metre scales, because these features may contribute to specific morphological differences, including the non-Gaussian structure, branching, and meandering observed in individual plumes. For example, Yuvaraj et al. (2026b) showed that including terrain in high-resolution CH_4 plume simulations had a



strong effect on plume height, vertical extent, and modelled TKE. They also found that, at higher wind speeds, turbulence generated by obstacles or the release structure itself can substantially alter plume dynamics. MicroHH already includes an immersed-boundary feature in DNS mode (Ražnjević, 2023; Bonekamp et al., 2020). Extending this capability to computationally efficient LES configurations will enable more detailed representation of obstacles in applications such as the present one.

The treatment of source dynamics provides another opportunity for refinement. Accurately representing the source flow is important not only for momentum-driven plume rise, but also because high exit velocities or large volumetric flow rates can generate source-induced turbulence that modifies near-source mixing and plume intermittency. In the present setup, momentum-driven plume rise was represented indirectly by adjusting the effective source height. This captures the initial vertical displacement of the release, although not its detailed downwind evolution or source-induced turbulence. For the controlled releases considered here, this simplification is unlikely to dominate the results because the computed momentum rise was modest (0–2.5 m) and comparable to the release-height scale (1.1–1.8 m). Buoyancy effects were also omitted because release temperature information was not available. Their importance depends strongly on source type: CH₄ from pressurised leaks can cool during adiabatic expansion, reducing buoyancy, whereas emissions from coal mine ventilation shafts can be warmer than ambient air. Yuvaraj et al. (2026b) showed that for small-pipe, low-momentum sources similar to the controlled releases studied here, terrain topography dominates plume shape and remains important up to 40 m from the source. Exit temperature has a smaller but still relevant influence, particularly at high exit velocities, where cold or hot releases can alter plume height and morphology. In contrast, Yuvaraj et al. (2026a) found that, for high-momentum, large-area sources such as coal mine ventilation shafts, exit velocity and the volumetric flow rate of the vented air are the main controls on plume behaviour, while high-resolution terrain and building data are less critical. An online treatment of buoyancy-driven plume rise is currently under development and will extend the range of buoyant sources that can be represented directly in MicroHH. Because plume-rise calculations are highly sensitive to model resolution (Yuvaraj et al., 2026b), the current development of nesting capabilities will be an important asset for making such simulations computationally feasible.

4.2 Recommended applications

The results of this study, together with previous work, suggest three main use cases for generating synthetic plume images with MicroHH.

First, the model is well suited for plumes over homogeneous or moderately heterogeneous terrain, where surface features can be reasonably approximated and where boundary conditions, such as wind profiles, surface roughness, and source parameters, are well constrained. Under these conditions, MicroHH can simulate atmospheric plumes and turbulence at spatial scales ranging from several kilometres down to the metre level, provided that sufficient computational resources are available (e.g. Ražnjević et al., 2022b; Krol et al., 2024; Koene and Brunner, 2023). The current main version of MicroHH is already highly capable in these settings, as demonstrated by the realistic plume morphology obtained here. Complex terrain, large obstacles, subgrid-scale source diameters, and strong momentum or buoyancy effects require additional parameterisation or finer-scale



treatment. The obstacle, nesting, open-boundary, and online plume-rise developments described above are intended to extend
515 MicroHH's applicability to these more demanding cases.

Second, MicroHH is useful when synthetic imagery is needed for reactive trace gases rather than passive tracers alone. The model includes atmospheric chemistry, including a condensed scheme for the coupled evolution of NO, NO₂, O₃, and related species in turbulent plumes (Krol et al., 2024). This capability is particularly relevant for NO_x emission studies because observed NO₂ columns do not scale linearly with emitted NO_x: the relationship depends on chemical conversion, photo-
520 chemical loss, entrainment, background O₃, radiation, plume age, and the spatial scale of the observation (Meier et al., 2024; Hakkarainen et al., 2024; Krol et al., 2024). MicroHH can therefore be used to generate process-resolving synthetic NO₂ and NO_x plume images for testing satellite emission retrievals, deriving NO₂-to-NO_x conversion factors, and evaluating applications for power plants, cities, and ships (Krol et al., 2024; Meier et al., 2024; Hakkarainen et al., 2024; Mols et al., 2026; Riess et al., 2025).

Third, MicroHH is especially attractive when many plume realisations and long simulation periods are needed. A key practical advantage is its high computational efficiency, particularly on GPU architectures (van Heerwaarden et al., 2017). In the setup used here, the GPU implementation achieved an average utilisation of 95%, demonstrating that the code pushed the hardware (in this case an extremely powerful NVIDIA GH200 Grace Hopper Superchip) to its limits. This makes the model well suited for generating large ensembles of synthetic plume data for ML applications, such as plume segmentation
530 and emission retrieval, where many realistic but not necessarily site-perfect realisations are needed. By contrast, if the main objective is to reproduce a single plume as realistically as possible at a specific site, models that explicitly resolve local terrain, obstacles, and detailed source-scale dynamics may be more appropriate despite their higher computational cost.

Beyond these specific use cases, the validation approach applied here also addresses a broader need: because this appears to be the first systematic, multi-metric comparison of LES-simulated plume structure against metre-scale imaging-spectrometer
535 observations, no published reference values or acceptance thresholds currently exist against which such simulations can be judged. The metrics and results presented in Section 3 therefore provide an initial reference point that can support future intercomparison of CFD models.

5 Conclusions

Synthetic plume images from numerical models are increasingly used to develop and test emission-retrieval methods, but
540 their value depends on how well they reproduce observed plume morphology. Here, we evaluated the ability of the LES model MicroHH to simulate metre-scale CH₄ plume images by comparing simulations with AVIRIS-4 observations from a controlled-release experiment in France. To our knowledge, this is the first study to systematically evaluate LES-simulated plume morphology against metre-scale imaging-spectrometer observations using a combination of spectral and morphological metrics, providing both a specific assessment of MicroHH and a transferable framework for validating high-resolution plume
545 simulations more generally.



MicroHH reproduced several key plume characteristics when boundary conditions and source parameters were well constrained, including width, mean concentration structure, turbulence spectra, and spatial CH₄ power spectra. Although the comparison using all 135 model snapshots initially suggested that the simulated plumes were less intermittent than the observations, this difference was strongly reduced when the model output was subsampled to match the number of valid AVIRIS-4 observations. The sample-size-matched Gaussian-fit and along-plume flux statistics showed comparable variability, indicating that the apparent lack of model intermittency is largely a sampling effect. The observed plumes nevertheless exhibited more complex and non-Gaussian structures, some of which were related to viewing and illumination geometry. The remaining differences likely stem from finite grid resolution, simplified surface and obstacle representation, simplified treatment of momentum-driven plume rise, and omitted buoyancy effects.

Overall, MicroHH is already a highly capable and computationally efficient model for generating realistic synthetic plume images, including large ensembles for retrieval benchmarking and ML training, as well as simulations of trace gases such as CH₄. The present evaluation shows strong agreement with metre-scale observations across complementary morphological and spectral metrics. Ongoing developments in open-boundary conditions, nesting, and online buoyancy-driven plume-rise treatment, together with future work on surface and obstacle representation, would further broaden the range of environments and source types that can be represented and improve the capture of the finest-scale plume variability.

Code availability. The code and packages used in this study are publicly available: the MicroHH model (version 2.0.1) is hosted on GitHub (<https://github.com/microhh/microhh>); the $(LS)^2D$ module (version 1.0), used to generate large-scale forcing from ERA5, is available on GitHub (<https://github.com/LS2D/LS2D>); and the *ddeq* package (version 1.1) for plume morphology analysis can be found on GitLab (<https://gitlab.com/empa503/remote-sensing/ddeq>).

Data availability. The initial conditions and initialisation files for running MicroHH, the simulated CH₄ column data, and the processed results are available in a Zenodo data repository (<https://doi.org/10.5281/zenodo.22029716>). The retrieved AVIRIS-4 CH₄ maps, wind and source data, and estimated emissions are also available on Zenodo (<https://doi.org/10.5281/zenodo.16410532>) (Meier et al., 2026).



Appendix A: Influence of unequal sample sizes

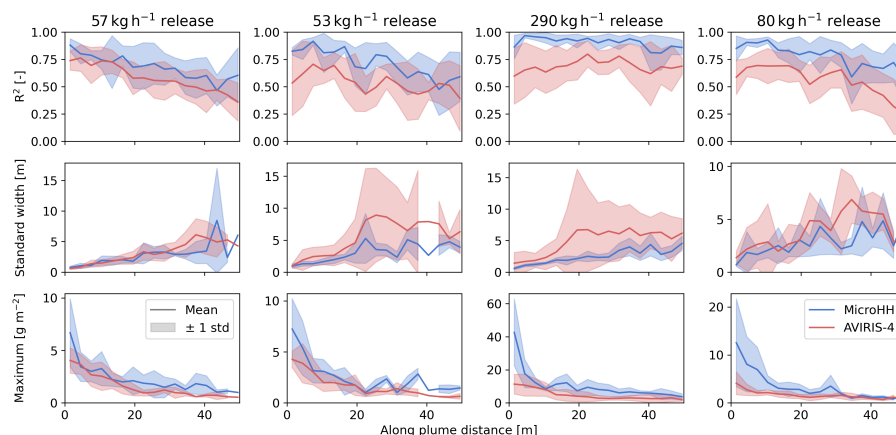


Figure A1. Sample-size-matched comparison of Gaussian-fit statistics for MicroHH (blue) and AVIRIS-4 (red) for the four controlled releases. Rows show the Gaussian-fit R^2 values (top), standard widths (middle), and maxima (bottom). At each along-plume position, valid MicroHH snapshots were randomly sampled without replacement to match the number of valid AVIRIS-4 observations remaining after quality filtering. Lines show the mean, and shading indicates ± 1 sample standard deviation. Compared with the unequal-sample comparison in Figure 4, the sample-size-matched datasets exhibit substantially more similar variability.

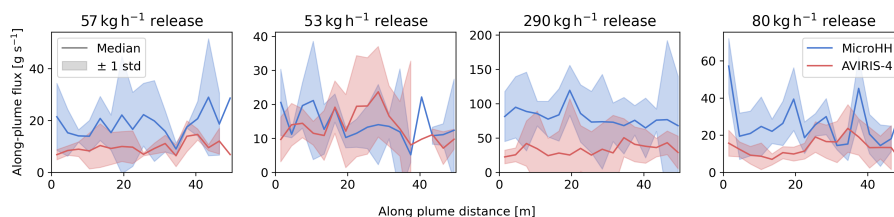


Figure A2. Sample-size-matched comparison of along-plume CH_4 flux for MicroHH (blue) and AVIRIS-4 (red) for the four controlled releases. At each along-plume position, valid MicroHH snapshots were randomly sampled without replacement to match the number of valid AVIRIS-4 observations remaining after quality and Gaussian-fit filtering. Lines show the mean, and shading indicates ± 1 sample standard deviation. The matched comparison shows that much of the apparent difference in flux variability in Figure 5 is sensitive to the unequal sample sizes.



Appendix B: Comparison of modelled VCDs and GPCDs

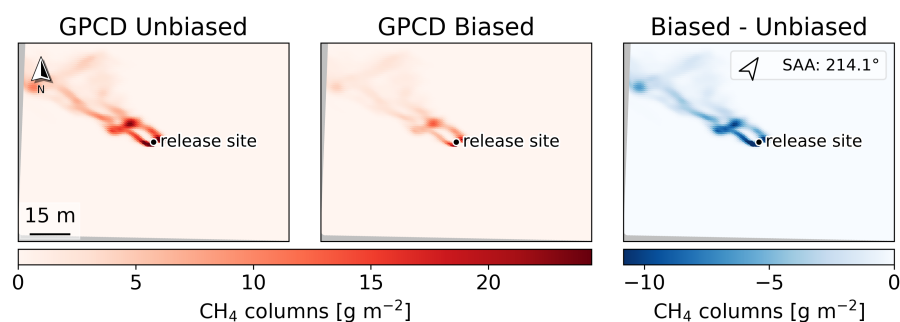


Figure B1. Comparison of unbiased and biased MicroHH geometry-projected CH_4 column densities (GPCDs). The left panel shows the unbiased GPCD calculated according to Equation 3, the middle panel shows the biased GPCD calculated according to Equation 4, and the right panel shows the difference between the two fields.



570 Appendix C: Sensitivity to domain size

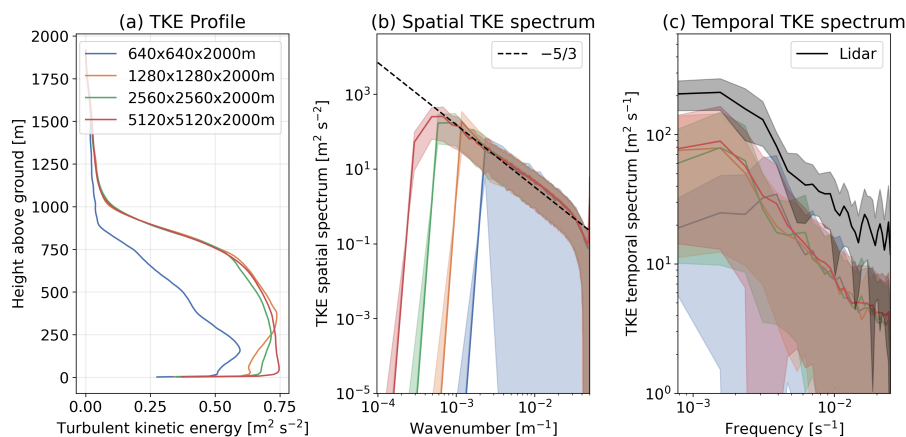


Figure C1. Sensitivity of the simulated turbulence to the horizontal domain size for the $57 \text{ kg CH}_4 \text{ h}^{-1}$ release. The simulations used horizontal domain sizes of 640, 1280, 2560, and 5120 m at a horizontal grid spacing of 10 m and a common vertical extent of 2000 m. Panels show (a) vertical turbulent kinetic energy (TKE) profiles, (b) spatial TKE spectra, and (c) temporal TKE spectra compared with the wind-lidar spectrum. Shading indicates the variability over the dimensions used for averaging, and the black dashed line in panel (b) shows the theoretical $-5/3$ slope.

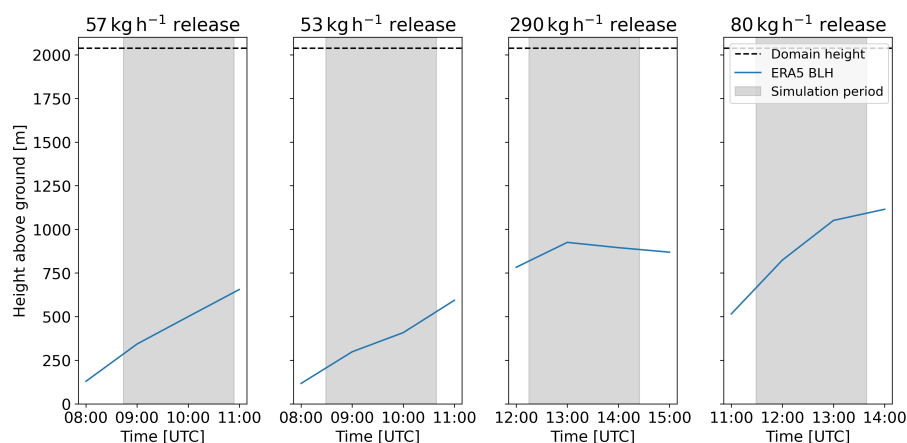


Figure C2. ERA5 planetary boundary-layer height and MicroHH domain height during the four simulated releases. The blue lines show the ERA5 boundary-layer height, the black dashed lines show the vertical extent of the MicroHH domain, and the grey shading indicates the simulation periods.



Author contributions. SM conducted the analysis and authored the manuscript with input from all co-authors; DB offered constructive feedback on the manuscript; GK supervised the project.

Competing interests. The authors declare that they have no competing interests.

575 *Acknowledgements.* We thank Bart van Stratum and Chiel van Heerwaarden from Wageningen University for their support with setting up and parameterising MicroHH. Model simulations, data processing and storage were performed on the Alps supercomputer at the Swiss National Supercomputing Centre (CSCS) under project s1302 and Ip171. The AI tools ChatGPT, Gemini, and Grammarly were used for grammar checking, formatting, copy-editing, and wording improvements.



References

- Bonekamp, P. N. J., van Heerwaarden, C. C., Steiner, J. F., and Immerzeel, W. W.: Using 3D turbulence-resolving simulations to understand the impact of surface properties on the energy balance of a debris-covered glacier, *The Cryosphere*, 14, 1611–1632, <https://doi.org/10.5194/tc-14-1611-2020>, 2020.
- Briggs, G. A.: *Plume Rise*, AEC Critical Review Series TID-25075, U.S. Atomic Energy Commission, Division of Technical Information Extension, Oak Ridge, Tennessee, 1969.
- Brunner, D., Kuhlmann, G., Marshal, J., Clément, V., Fuhrer, O., Broquet, G., Löscher, A., and Meijer, Y.: Accounting for the vertical distribution of emissions in atmospheric CO₂ simulations, *Atmospheric Chemistry and Physics*, 19, 4541–4559, <https://doi.org/10.5194/acp-19-4541-2019>, type: Journal Article, 2019.
- Brunner, D., Kuhlmann, G., Henne, S., Koene, E., Kern, B., Wolff, S., Voigt, C., Jöckel, P., Kiemle, C., Roiger, A., Fiehn, A., Krautwurst, S., Gerilowski, K., Bovensmann, H., Borchardt, J., Galkowski, M., Gerbig, C., Marshall, J., Klonecki, A., Prunet, P., Hanfland, R., Pattantyús-Ábrahám, M., Wyszogrodzki, A., and Fix, A.: Evaluation of simulated CO₂ power plant plumes from six high-resolution atmospheric transport models, *Atmospheric Chemistry and Physics*, 23, 2699–2728, <https://doi.org/10.5194/acp-23-2699-2023>, 2023.
- Bruno, J. H., Jervis, D., Varon, D. J., and Jacob, D. J.: U-Plume: automated algorithm for plume detection and source quantification by satellite point-source imagers, *Atmospheric Measurement Techniques*, 17, 2625–2636, <https://doi.org/10.5194/amt-17-2625-2024>, 2024.
- Bue, B. D., Lee, J. H., Thorpe, A. K., Brodrick, P. G., Cusworth, D., Ayasse, A., Mancoridis, V., Satish, A., Xiong, S., and Duren, R.: Towards operational automated greenhouse gas plume detection and delineation, *Remote Sensing of Environment*, 343, 115 506, <https://doi.org/10.1016/j.rse.2026.115506>, 2026.
- Dumont Le Brazidec, J., Vanderbecken, P., Farchi, A., Bocquet, M., Lian, J., Broquet, G., Kuhlmann, G., Danjou, A., and Lauvaux, T.: Segmentation of XCO₂ images with deep learning: application to synthetic plumes from cities and power plants, *Geoscientific Model Development*, 16, 3997–4016, <https://doi.org/10.5194/gmd-16-3997-2023>, 2023.
- Durran, D., Weyn, J. A., and Menchaca, M. Q.: Practical Considerations for Computing Dimensional Spectra from Gridded Data, *Monthly Weather Review*, 145, 3901 – 3910, <https://doi.org/10.1175/MWR-D-17-0056.1>, 2017.
- 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, 2020.
- Ganin, Y. and Lempitsky, V.: Unsupervised Domain Adaptation by Backpropagation, <https://doi.org/10.48550/arXiv.1409.7495>, arXiv:1409.7495 [stat.ML], 2015.
- Galkowski, M., Marshall, J., Fuentes Andrade, B., and Gerbig, C.: Impact of atmospheric turbulence on the accuracy of point source emission estimates using satellite imagery, *Atmospheric Chemistry and Physics*, 25, 13 831–13 848, <https://doi.org/10.5194/acp-25-13831-2025>, 2025.
- Gorroño, J., Varon, D. J., Irakulis-Loitxate, I., and Guanter, L.: Understanding the potential of Sentinel-2 for monitoring methane point emissions, *Atmospheric Measurement Techniques*, 16, 89–107, <https://doi.org/10.5194/amt-16-89-2023>, 2023.
- Gorroño, J., Pei, Z., Valverde, A., and Guanter, L.: Considering the observation and illumination angular configuration for an improved detection and quantification of methane emissions, *Atmospheric Measurement Techniques*, 19, 1245–1257, <https://doi.org/10.5194/amt-19-1245-2026>, 2026.



- Hakkarainen, J., Kuhlmann, G., Koene, E., Santaren, D., Meier, S., Krol, M. C., Stratum, B. J. H. v., Ialongo, I., Chevallier, F., Tamminen, J., Brunner, D., and Broquet, G.: Analyzing nitrogen dioxide to nitrogen oxide scaling factors for data-driven satellite-based emission estimation methods: A case study of Matimba/Medupi power stations in South Africa, *Atmospheric Pollution Research*, 15, 102 171, <https://doi.org/https://doi.org/10.1016/j.apr.2024.102171>, 2024.
- Hanfland, R., Brunner, D., Voigt, C., Fiehn, A., Roiger, A., and Pattantyús-Ábrahám, M.: The Lagrangian Atmospheric Radionuclide Transport Model (ARTM) – sensitivity studies and evaluation using airborne measurements of power plant emissions, *Atmospheric Chemistry and Physics*, 24, 2511–2534, <https://doi.org/10.5194/acp-24-2511-2024>, 2024.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), <https://doi.org/10.24381/cds.adbb2d47>, 2018.
- Jongaramrungruang, S., Frankenberg, C., Matheou, G., Thorpe, A. K., Thompson, D. R., Kuai, L., and Duren, R. M.: Towards accurate methane point-source quantification from high-resolution 2-D plume imagery, *Atmospheric Measurement Techniques*, 12, 6667–6681, <https://doi.org/10.5194/amt-12-6667-2019>, 2019.
- Jongaramrungruang, S., Thorpe, A. K., Matheou, G., and Frankenberg, C.: MethaNet – An AI-driven approach to quantifying methane point-source emission from high-resolution 2-D plume imagery, *Remote Sensing of Environment*, 269, 112 809, <https://doi.org/10.1016/j.rse.2021.112809>, 2022.
- Joyce, P., Ruiz Villena, C., Huang, Y., Webb, A., Gloor, M., Wagner, F. H., Chipperfield, M. P., Barrio Guilló, R., Wilson, C., and Boesch, H.: Using a deep neural network to detect methane point sources and quantify emissions from PRISMA hyperspectral satellite images, *Atmospheric Measurement Techniques*, 16, 2627–2640, <https://doi.org/10.5194/amt-16-2627-2023>, 2023.
- Kaimal, J., Wyngaard, J., Izumi, Y., and Coté, O.: Spectral characteristics of surface-layer turbulence, *Quarterly Journal of the Royal Meteorological Society*, 98, 563–589, 1972.
- Kleissl, J., Kumar, V., Meneveau, C., and Parlange, M. B.: Numerical study of dynamic Smagorinsky models in large-eddy simulation of the atmospheric boundary layer: Validation in stable and unstable conditions, *Water Resources Research*, 42, <https://doi.org/https://doi.org/10.1029/2005WR004685>, 2006.
- Knoop, S., Bosveld, F. C., de Haij, M. J., and Apituley, A.: A 2-year intercomparison of continuous-wave focusing wind lidar and tall mast wind measurements at Cabauw, *Atmospheric Measurement Techniques*, 14, 2219–2235, <https://doi.org/10.5194/amt-14-2219-2021>, 2021.
- Koene, E. and Brunner, D.: D4.2 Assessment of Plume Model Performance, Tech. rep., Empa, <https://coco2-project.eu/node/357>, 2023.
- Kolmogorov, A. N.: The local structure of turbulence in incompressible viscous fluid for very large Reynolds number, *Cr Acad. Sci. USSR*, 30, 301, 1941.
- Krol, M., van Stratum, B., Anglou, I., and Boersma, K. F.: Evaluating NO_x stack plume emissions using a high-resolution atmospheric chemistry model and satellite-derived NO₂ columns, *Atmospheric Chemistry and Physics*, 24, 8243–8262, <https://doi.org/10.5194/acp-24-8243-2024>, 2024.
- Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., and Brunner, D.: Detectability of CO₂ emission plumes of cities and power plants with the Copernicus Anthropogenic CO₂ Monitoring (CO2M) mission, *Atmospheric Measurement Techniques*, 12, 6695–6719, <https://doi.org/10.5194/amt-12-6695-2019>, 2019.
- Kuhlmann, G., Henne, S., Brunner, D., Löscher, A., and Meijer, Y.: SMARTCARB2 - Use of satellite measurements of auxiliary reactive trace gases for fossil fuel carbon dioxide emission estimation (Phase 2), Final report of ESA study contract



- n°4000119599/16/NL/FF/mg, Tech. rep., Empa, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland, <https://doi.org/10.5281/zenodo.4674167>, 2021.
- Kuhlmann, G., Koene, E. F. M., 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), accepted in Geo. Mod. Dev., <https://doi.org/10.5194/egusphere-2023-2936>, 2024.
- Kuhlmann, G., Stavropoulou, F., Schwietzke, S., Zavala-Araiza, D., Thorpe, A., Hueni, A., Emmenegger, L., Calcan, A., Röckmann, T., and Brunner, D.: Evidence of successful methane mitigation in one of Europe's most important oil production region, *Atmospheric Chemistry and Physics*, 25, 5371–5385, <https://doi.org/10.5194/acp-25-5371-2025>, 2025.
- Liu, X. and Zhang, S.: Who is closer: A computational method for domain gap evaluation, *Pattern Recognition*, 122, 108 293, <https://doi.org/10.1016/j.patcog.2021.108293>, 2022.
- Lorenz, E. N.: The predictability of a flow which possesses many scales of motion, *Tellus*, 21, 289–307, <https://doi.org/10.1111/j.2153-3490.1969.tb00444.x>, 1969.
- McManemin, A., Juéry, C., Blandin, V., Baker, M., Bauguitte, S., Brunner, D., Burdeau, P., France, J. L., Fredenslund, A. M., Hueni, A., Kuhlmann, G., Lakomic, P., Meier, S., Paglini, R., Papaconstantinou, R., Paris, J.-D., Quehe, P.-Y., Röckmann, T., Scheutz, C., Schmidt, M., Vögtli, M., Wietzel, J. B., and Brandt, A. R.: From research to reality: Academic methane measurement systems tested at the TADI controlled release facility, *Atmospheric Measurement Techniques*, 2026a.
- McManemin, A., Juéry, C., Blandin, V., France, J. L., Burdeau, P., and Brandt, A. R.: Controlled release testing of commercially available methane emission measurement technologies at the TADI facility, *Atmospheric Measurement Techniques*, 19, 923–934, <https://doi.org/10.5194/amt-19-923-2026>, 2026b.
- Meier, S., Koene, E. F. M., Krol, M., Brunner, D., Damm, A., and Kuhlmann, G.: A lightweight NO₂-to-NO_x conversion model for quantifying NO_x emissions of point sources from NO₂ satellite observations, *Atmospheric Chemistry and Physics*, 24, 7667–7686, <https://doi.org/10.5194/acp-24-7667-2024>, 2024.
- Meier, S., Vögtli, M., Hueni, A., McManemin, A., Brandt, A. R., Juéry, C., Blandin, V., Brunner, D., and Kuhlmann, G.: Quantifying CH₄ point source emissions with airborne remote sensing: first results from AVIRIS-4, *Atmospheric Measurement Techniques*, 19, 333–358, <https://doi.org/10.5194/amt-19-333-2026>, 2026.
- Mols, A., Boersma, K. F., Denier van der Gon, H., and Krol, M.: An improved Bayesian inversion to estimate daily NO_x emissions of Paris from TROPOMI NO₂ observations between 2018–2023, *Atmospheric Chemistry and Physics*, 26, 1497–1513, <https://doi.org/10.5194/acp-26-1497-2026>, 2026.
- Ouerghi, E., Ehret, T., Facciolo, G., Meinhardt, E., Marion, R., and Morel, J.-M.: Tightening up methane plume source rate estimation in EnMAP and PRISMA images, *Atmospheric Measurement Techniques*, 18, 4611–4629, <https://doi.org/10.5194/amt-18-4611-2025>, 2025.
- Plewa, T., Butz, A., Frankenberg, C., Thorpe, A. K., and Marshall, J.: Improvements of AI-driven emission estimation for point sources applied to high resolution 2-D methane-plume imagery, *Remote Sensing of Environment*, 331, 115 002, <https://doi.org/https://doi.org/10.1016/j.rse.2025.115002>, 2025.
- Ražnjević, A.: High-resolution modelling of plume dispersion, internal PhD, WU, Wageningen University, Netherlands, ISBN 978-94-6447-787-0, <https://doi.org/10.18174/634327>, 2023.
- Ražnjević, A., van Heerwaarden, C., and Krol, M.: Evaluation of two common source estimation measurement strategies using large-eddy simulation of plume dispersion under neutral atmospheric conditions, *Atmospheric Measurement Techniques*, 15, 3611–3628, <https://doi.org/10.5194/amt-15-3611-2022>, 2022a.



- 690 Ražnjević, A., van Heerwaarden, C., van Stratum, B., Hensen, A., Velzeboer, I., van den Bulk, P., and Krol, M.: Technical note: Interpretation of field observations of point-source methane plume using observation-driven large-eddy simulations, *Atmospheric Chemistry and Physics*, 22, 6489–6505, <https://doi.org/10.5194/acp-22-6489-2022>, 2022b.
- Riess, T. C. V. W., Boersma, K. F., Prummel, A., Stratum, B. J. H. v., Laat, J. d., and Vliet, J. v.: Estimating NO_x emissions of individual ships from TROPOMI NO₂ plumes, *Remote Sensing of Environment*, 324, 114 734, <https://doi.org/https://doi.org/10.1016/j.rse.2025.114734>, 2025.
- 695 Rouet-Leduc, B. and Hulbert, C.: Automatic detection of methane emissions in multispectral satellite imagery using a vision transformer, *Nature Communications*, 15, 3801, <https://doi.org/10.1038/s41467-024-47754-y>, 2024.
- Schwaerzel, M., Emde, C., Brunner, D., Morales, R., Wagner, T., Berne, A., Buchmann, B., and Kuhlmann, G.: Three-dimensional radiative transfer effects on airborne and ground-based trace gas remote sensing, *Atmospheric Measurement Techniques*, 13, 4277–4293, <https://doi.org/10.5194/amt-13-4277-2020>, 2020.
- 700 Tijhuis, M., van Stratum, B. J. H., and van Heerwaarden, C. C.: The impact of coupled 3D shortwave radiative transfer on surface radiation and cumulus clouds over land, *Atmospheric Chemistry and Physics*, 24, 10 567–10 582, <https://doi.org/10.5194/acp-24-10567-2024>, 2024.
- Tijhuis, M., Seifert, A., Lozar, A. d., Stratum, B. J. H. v., and Heerwaarden, C. C. v.: Comparing the dynamics of idealized squall lines between NWP and LES models, <https://arxiv.org/abs/2506.19435>, 2025.
- Tzeng, E., Hoffman, J., Saenko, K., and Darrell, T.: Adversarial Discriminative Domain Adaptation, <https://doi.org/10.48550/arXiv.1702.05464>, arXiv:1702.05464 [cs.CV], 2017.
- 705 van Heerwaarden, C. C., van Stratum, B. J., Heus, T., Gibbs, J. A., Fedorovich, E., and Mellado, J. P.: MicroHH 1.0: A computational fluid dynamics code for direct numerical simulation and large-eddy simulation of atmospheric boundary layer flows, *Geoscientific Model Development*, 10, 3145–3165, 2017.
- van Stratum, B. J. H., van Heerwaarden, C. C., and Vilà-Guerau de Arellano, J.: The Benefits and Challenges of Downscaling a Global Reanalysis With Doubly-Periodic Large-Eddy Simulations, *Journal of Advances in Modeling Earth Systems*, 15, e2023MS003 750, <https://doi.org/https://doi.org/10.1029/2023MS003750>, 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.
- 715 Welch, P.: The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Transactions on Audio and Electroacoustics*, 15, 70–73, <https://doi.org/10.1109/TAU.1967.1161901>, 1967.
- Wolff, S., Ehret, G., Kiemle, C., Amediek, A., Quatrevalet, M., Wirth, M., and Fix, A.: Determination of the emission rates of CO₂ point sources with airborne lidar, *Atmospheric Measurement Techniques*, 14, 2717–2736, <https://doi.org/10.5194/amt-14-2717-2021>, 2021.
- Yuvaraj, R., Knapp, M., Abdallah, C., Butz, A., Swolkien, J., and Lauvaux, T.: High-fidelity methane plume modelling over coal mines: Integrating Large Eddy Simulations (LES) with remote sensing imagery, *ChemRxiv*, 2026, <https://doi.org/10.26434/chemrxiv.15002286/v1>, 2026a.
- 720 Yuvaraj, R., Lauvaux, T., Abdallah, C., Ciais, P., Akani Guery, J., Bonne, J.-L., Groshenry, A., Hoang, N. M., and Joly, L.: High-Resolution Modeling of Methane Plumes: Validation and Sensitivity Experiments to Explore Emission Quantification Approaches, *Environmental Science & Technology*, 60, 4029–4041, <https://doi.org/10.1021/acs.est.5c15941>, 2026b.