Response to Reviewer #1

Manuscript: "Spatial resolution versus precision in CO_2 point-source emission retrievals from the DQ-1 spaceborne lidar: a plume-width-normalized interpretation"

Journal: Atmospheric Measurement Techniques (AMT)

We thank the reviewer for the careful and constructive reading of the manuscript. The comments helped us distinguish more clearly between a universal averaging factor, which the data do not support, and two scene-specific quantities that can be determined reproducibly: the precision-detectability balance scale N_{det} and the emission-qualification scale N_Q. N_Q is the qualification optimum under prespecified satellite-only diagnostics; it is not a truth-based emission-accuracy optimum. Reviewer comments are reproduced below in grey italics, our responses are shown in blue, and manuscript changes are shown in red. Because pagination may change during production, locations are given primarily by section, equation, figure, and table number.

General comment

Reviewer comment

Active lidar remote sensing can complement the current GHG satellite observing system by contributing measurements at high latitudes, at night, and in between cloud gaps. This technical paper analyzes the relationship between along-track averaging length of DQ-1 IPDA lidar observations and the CO_2 emission retrieval skill for strong overflown point sources. The authors conclude that the optimum averaging length depends on case-to-case methodical and geophysical conditions and therefore cannot be determined automatically. The manuscript has a fair presentation quality yet lacks substance. Only few DQ-1 measurements are described, and many simulation results are trivial, such as small averaging length - too much noise - wrong CO_2 emission retrieval, or, large averaging length - CO_2 emission plume too coarse - retrieval fails. Similar but passive remote sensing studies can be found in the literature. I would still recommend its publication in AMT, after consideration of my comments below, because the present study is a first necessary step towards a future more comprehensive assessment of the highly innovative DQ-1 lidar mission.

Author response

We thank the reviewer for recognizing both the value of active-lidar observations and the role of this study as an initial quantitative assessment of DQ-1. The revised manuscript moves beyond the qualitative statement that fine averaging is noisy and coarse averaging is over-smoothed in three ways. First, controlled calculations express the detection-balance support relative to plume FWHM and report the actual point-source dispersion parameterization. Second, the sample structure is explicit: 156 overpasses enter detection analysis, 48 are complete at all six detection scales, 44 cases enter the separate flux pre-screen, and 38 retain at least one eligible emission scale. The complete 38-case set is now reported in Supplement Table S1; Table 3 remains a 20-case illustration. Third, N_{det} and N_Q are selected independently from local observations. In the 35 cases with both scales, N_Q is coarser in 10, equal in 13, and finer in 12. The result is therefore a reproducible two-objective workflow, not a universal N or a universal ordering.

The revised manuscript now makes the reporting rule explicit: N_Q identifies the best-qualified support from satellite-only gates and ranking, whereas adjacent-scale agreement assigns a separate stability label. Eleven of 38 eligible cases meet the factor-of-two stability test. Score and gate sensitivity is now reported in Supplement Table S2: 37–38 of 38 choices are unchanged under score-component, weight, and tie-rule perturbations, and strict gates preserve 35 of 36 retained choices.

Changes in the revised manuscript

The Abstract, Introduction, Methods Sects. 2.4-2.5, Results Sects. 3.2-3.3, Tables 1-3, Figs. 13-18, Discussion Sect. 4, and Conclusions were rewritten or expanded. The manuscript now states explicitly that $L_{det}/W \approx 1.2$ is a controlled-model scaling rather than a universal field law, and that external inventories are joined only after the satellite-only scale decisions are frozen.

Point-by-point response

Comment 1: Missing Amediek et al. (2017) reference

Reviewer comment

Line 55 (and line 493): the reference Amediek et al., 2017 is missing.

Author response

Amediek et al. (2017) is cited in the Introduction to support the demonstrated capability of airborne IPDA lidar for strong point-source measurements, and the complete entry is now included in the revised reference list.

Changes in the revised manuscript

Revised manuscript, References: Amediek, A., Ehret, G., Fix, A., Wirth, M., Büdenbender, C., Quatrevalet, M., Kiemle, C., and Gerbig, C.: CHARM-F—a new airborne integrated-path differential-absorption lidar for carbon dioxide and methane observations: measurement performance and quantification of strong point source emissions, Appl. Opt., 56, 5182–5197, https://doi.org/10.1364/AO.56.005182, 2017.

Comment 2: Gaussian-plume dispersion term in Eq. (1)

Reviewer comment

Eq. 1: after Eq. A4 in Bovensmann et al, 2010, the Gaussian plume standard deviation is defined as $a(x+x_0)^{0.894}$, instead of $a(x/x_0)^{0.894}$, with unit m. Please correct.

Author response

We agree that the original finite-source transcription was incorrect. The controlled model represents the plants as point sources, so the offset x_0 from Eq. (A4) of Bovensmann et al. (2010) is not used. The revised manuscript now states the actual three-dimensional dispersion implementation explicitly: $\sigma_k(x) = \exp[A_{k,0} + A_{k,1} \ln x + A_{k,2}(\ln x)^2]$ for $k \in \{y, z\}$, with x, σ_y, and σ_z in metres. Vertical integration yields the column form in Eq. (1), and Supplement Table S3 lists both coefficient sets. This replaces the erroneous $a(x/x_0)^{0.894}$ expression rather than substituting another finite-source formula that is not used by our point-source solver. Equations (1)–(4) describe the controlled forward model; Eqs. (5a)–(5b) separately define the real-case cross-sectional-flux estimator.

Changes in the revised manuscript

Revised manuscript: Sect. 2.2 and Supplement Table S3. The paragraph following Eq. (1) now gives the implemented point-source $\sigma_y(x)$ and $\sigma_z(x)$ relations, their units, the coefficient source, and the vertical-integration link to Eq. (1); no finite-source offset is used.

Comment 3: Definition and unit of wind speed

Reviewer comment

Line 116: ...and surface wind speed u (m/s).

Author response

We agree that the wind variables and their units required clearer definition. The revised manuscript distinguishes the prescribed transport wind-speed magnitude U used in the controlled simulations, the effective wind component U_{eff} normal to the real satellite track, and the near-surface wind used with cloud cover and solar elevation to assign Pasquill-Gifford stability. All wind speeds are reported in m s−1.

Changes in the revised manuscript

Revised manuscript: Sects. 2.1.3 and 2.2, especially the definitions following Eq. (1) and the paragraph defining $U_{eff} = |U \cdot n_{trk}|$. The simulation captions and text now state when $U_{eff} = U$ by construction.

Comment 4: Grammatical correction

Reviewer comment

Line 118: ... is converted (instead of 130 converted).

Author response

We agree and corrected this typographical error.

Changes in the revised manuscript

Revised manuscript: Sect. 2.2, text introducing Eq. (2): "The modelled column-mass enhancement is converted to a dry-air column-averaged mole-fraction enhancement ...".

Comment 5: Terminology for averaging interval

Reviewer comment

Line 144: better "time interval" instead of "averaging factors".

Author response

We agree that the original term was ambiguous. Because the DQ-1 products are analysed as an along-track spatial sequence, we now use "averaging count", "averaging scale", or "N-sample support" rather than treating N as a dimensionless processing factor. We did not replace every occurrence with "time interval", because the physical quantity relevant to plume resolution is the along-track spatial support. With the corrected 337.5 m sample-centre spacing, each N is converted directly to kilometres.

Changes in the revised manuscript

Revised manuscript: Sects. 2.3-2.5 and all relevant figure captions. The tested N = 5, 10, 15, 20, 25, and 30 correspond to nominal supports of 1.6875, 3.375, 5.0625, 6.75, 8.4375, and 10.125 km, respectively.

Comment 6: Figure 2 presentation and white-noise reference

Reviewer comment

Figure 2: the different colors cannot be distinguished; include the $\tau^{-1/2}$ relationship for pure white noise; add a km-axis for the averaging length.

Author response

The revised figure separates desert and forest results into two panels, emphasizes the group mean with a black line and markers, and uses the individual coloured curves only as secondary track-level context, so interpretation no longer depends solely on distinguishing every colour. A white-noise $N^{-1/2}$ reference is plotted explicitly and is described as a reference rather than a fitted instrument-noise

model. Because the data are indexed by sample count rather than elapsed averaging time, the reference is expressed as $N^{-1/2}$. The physical conversion is 0.3375N km.

Changes in the revised manuscript

Revised manuscript: Sect. 2.3 and Fig. 2. The white-noise reference and its interpretation are included, and the upper x-axis now shows L = 0.3375N km so that averaging count and physical support can be read together.

Comment 7: Definition of the coefficient of determination

Reviewer comment

Section 3.1: explain which data you use to calculate your "coefficient of determination", R^2.

Author response

The original manuscript used R^2 ambiguously for two different fits and did not clearly separate the two emission estimators. The revised Methods now define them separately. For the controlled simulations in Figs. 5–12, R^2_{flux} is calculated from the same retained cross-plume enhancement samples used in the Eq. (4) forward fit. It compares the observed noisy/block-averaged enhancements with the Gaussian-plume forward-model values evaluated at Q̂, using the sample-mean enhancement as the reference. Consequently, R^2_{flux} may be negative when the constrained plume template performs worse than the sample mean. For real overpasses, Q̂ is obtained instead from the cross-sectional-flux estimator in Eqs. (5a)–(5b), based on Â$\hat{\sigma}U_{eff}$, and R^2_{prof} is calculated separately from the Gaussian-plus-linear-background profile fit. R^2_{prof} is used only as a profile-agreement diagnostic, not as emission accuracy.

Changes in the revised manuscript

Revised manuscript: Sect. 2.2, Eqs. (6)-(7), and the definitions immediately following them; corresponding notation is used consistently in Results Sect. 3.1, Figs. 5-12, Sect. 3.3, and Fig. 16.

Comment 8: Non-monotonic wind-speed response in Figure 11

Reviewer comment

Figure 11: following Eq. 1, increasing windspeed u decreases the plume enhancement. Therefore, the skill to detect the plume, expressed by your "coefficient of determination", R^2, green lines, should decrease as well. This is observed in Figure 11 at 2, 6 and 8 m/s, but not at 4 m/s. The R^2 values at 4 m/s are lower than at 6 m/s. Please explain. I think this is an error.

Author response

We appreciate this careful observation. The expected monotonic decrease applies when wind speed varies at fixed dispersion. Our plotted experiment is instead a coupled wind–stability scenario: class A is used at 1 m s^{-1}, B at 2–4 m s^{-1}, C at 5 m s^{-1}, and D at 6–10 m s^{-1}. Figure 11 shows the representative 2, 4, 6, and 8 m s^{-1} subset, so 2 and 4 m s^{-1} are class B and 6 and 8 m s^{-1} are class D. The 4-to-6 m s^{-1} transition therefore changes both amplitude and dispersion. We retain the calculated curves but no longer describe them as a monotonic wind-only response; the mapping and coefficients are now given explicitly.

Changes in the revised manuscript

Revised manuscript: Results Sect. 3.1 and the captions of Figs. 11–12 now list the wind/stability assignments explicitly: 2 and 4 m s^{-1} use class B, and 6 and 8 m s^{-1} use class D. The text therefore identifies the 4-to-6 m s^{-1} class transition rather than describing the non-monotonic ordering as a wind-only effect.

Comment 9: Corresponding anomaly in Figure 12b

Reviewer comment

Figure 12b shows the same anomaly at 4 m/s. Please explain. I think this is an error.

Author response

The same explanation applies to Fig. 12b, but the revised reply now distinguishes it from Fig. 11. Figure 12 contains the complete 1–10 m s^{-1} sweep: A at 1, B at 2–4, C at 5, and D at 6–10 m s^{-1}. Thus, the apparent non-monotonicity near 4–6 m s^{-1} coincides with two stability-class transitions (B → C → D). The panel is interpreted as a set of coupled wind–stability scenarios. At fixed stability and geometry, changing U rescales enhancement amplitude; the plotted cross-class ordering is not used to claim a pure wind-speed dependence.

Changes in the revised manuscript

Revised manuscript: Results Sect. 3.1 and the revised Fig. 12 caption now state the assignments B, B, D, and D for 2, 4, 6, and 8 m s^{-1}, respectively, and identify the class change between 4 and 6 m s^{-1} as the reason that the panel cannot be read as a pure wind-speed sequence.

Comment 10: Wind direction and ERA5 vectors in the former Figure 13

Reviewer comment

Figure 13: was the wind direction parallel to the orbit track? To show the wind variability, please add the ERA5 wind vectors of the pressure level you used to estimate the emissions.

Author response

In the controlled simulations, the prescribed 270° wind is perpendicular to the idealized orbit track, so U_{eff} = U by construction and no ERA5 field is used. For the real overpasses, the satellite track can intersect the plume obliquely. Revised Fig. 15 now overlays a sparse set of native 0.25° ERA5 grid vectors from the nearest hourly analysis. The hourly 10 and 100 m winds and pressure-level winds were combined and interpolated to 490 m above ground, representing an assumed 240 m stack height plus 250 m plume rise. The emission calculation uses U_{eff}, the component of this wind normal to the local track, rather than requiring the wind to be parallel or perpendicular to the observed track. Figure 15 reports U_{eff} = 0.58, 1.50, 2.74, 2.14, and 7.81 m s^{-1} for Luohuang, Berezovskaya, Zouxian, Mary, and Anpara, respectively.

Changes in the revised manuscript

Revised manuscript: Sects. 2.1.3 and 2.2 now document the ERA5 vertical interpolation and define U_{eff}; the controlled-simulation geometry is stated in Sects. 2.4 and 3.1. Revised Fig. 15 retains the plant, native track, and selected non-overlapping AVS blocks, overlays sparse ERA5 grid wind vectors at 490 m above ground, and reports the case-specific U_{eff} used for emission inversion.

Comment 11: Selection criteria for the cases in the former Table 2

Reviewer comment

Table 2: after which criteria were these cases selected: wind direction parallel to the orbit track, other criteria?

Author response

We agree that the original selection criteria were insufficiently documented. A parallel wind-track geometry was not required. Candidate overpasses were required to contain valid Level-2C retrieval inputs, $300 \leq X_{CO_2} \leq 550$ ppm, cloud flag 2, state flag 1, land flag 1, positive weighting functions, and a continuous valid segment within 75 km of the plant. The fitted enhancement had to lie within 50 km of the source and inside the downwind sector, and cases with a source-to-plume direction differing

from ERA5 wind by more than 25° were rejected. For an emission scale to be eligible, the revised procedure additionally requires $R^2_{prof} \geq 0.20$, relative propagated uncertainty ≤ 2, $U_{eff} \geq 0.30$ m s−1, at least five profile points on each side of the centre, adequate edge coverage, a non-boundary fitted width, a physically bounded positive flux, and the background-cleanliness test. The satellite track can be oblique to the wind; this is handled through U_{eff} rather than by requiring parallelism or perpendicularity.

Changes in the revised manuscript

Revised manuscript: Sects. 2.2 and 2.5.1-2.5.2, especially the candidate-screening, wind-direction, eligibility, and background-cleanliness paragraphs. Current Table 2 summarizes population robustness tests; the plant-level cases and their diagnostics are reported in current Table 3.

Comment 12: Determination of averaging points in the former Table 3

Reviewer comment

Table 3: explain how you determined the averaging points.

Author response

The averaging count is no longer chosen by agreement with an inventory. Each original unaveraged orbit is reprocessed at N = 5, 10, 15, 20, 25, and 30. A fine-scale downwind maximum is identified once and retained as a fixed anchor. Complete non-overlapping N-sample blocks are then constructed from that anchor towards both ends of the valid track, and the calibrated online and offline signals are combined before the nonlinear XCO_2 retrieval. The pooled controlled relation $L_{det} = 1.209W$ supplies only an initial N_{det} trial; the baseline coefficient 1.182 is retained only in the baseline audit. The final N_{det} is selected as the maximum of $D(N) = \hat{A}(N)/\sigma_{local}(N)$. For real-overpass emission qualification, the fitted Gaussian enhancement area $\hat{A}\hat{\sigma}\sqrt{(2\pi)}$ is converted to $\hat{Q}$ with the cross-sectional-flux estimator in Eqs. (5a)–(5b) and U_{eff}; Eq. (4) is confined to the controlled simulations. Only scales passing the fit, wind, geometry, uncertainty, containment, support, and background gates are eligible. N_Q is the eligible scale maximizing the prespecified satellite-only score S_Q; near ties are resolved by smaller relative uncertainty, then a cleaner background, and then the smaller N. This makes N_Q a qualification optimum under the stated decision rule, not a truth-based emission-accuracy optimum. EPA CEMS and Climate TRACE values are attached only after N_Q has been frozen. Table 1 summarizes this decision sequence and its reporting outputs; the width-class boundaries themselves are defined in Sect. 2.5.2.

Changes in the revised manuscript

Revised manuscript: Sects. 2.4 and 2.5.1–2.5.3, unified Eqs. (9a)–(9b) and (12), Table 1, revised Fig. 15, and current Table 3. Table 3 now reports N_{det}/N_Q, retrieved flux and uncertainty, R^2_{prof}/RMSE, U_{eff}, distance, background excursions, and the post-selection external comparison.

Comment 13: Public availability of DQ-1 data

Reviewer comment

Line 410: can you comment when the DQ-1 data will become publicly available?

Author response

We appreciate the importance of this question. The wider DQ-1/ACDL Level-2C release schedule remains subject to mission policy. For this paper, however, the revised submission provides one representative original Level-2C product corresponding to the Anpara example in Fig. 15, the versioned derived-data archive, and a separate AMT supplement. Supplement Table S1 reports all 38 eligible emission cases with the processed case identifier, matched H5 UTC date, exact scale decision, and diagnostics; Table S2 reports N_Q sensitivity; Table S3 reports the controlled dispersion

coefficients. The derived package contains the corresponding scale-level rows and cohort crosswalk. The associated Zenodo record is available at https://doi.org/10.5281/zenodo.23874343; this concept DOI represents all versions and resolves to the latest version.

Changes in the revised manuscript

Revised manuscript: Data availability statement and Supplement Tables S1–S3. The statement distinguishes the restricted wider mission archive from the one representative original Level-2C product and complete derived case table supplied for this paper.

Closing statement

We thank the reviewer again for the constructive comments and for recommending publication after revision. The revised manuscript corrects the equation and notation issues, defines both coefficients of determination, explains the coupled wind-speed/stability experiment, documents the complete case-selection and averaging-scale procedures, and states the current DQ-1 data-access limitation. It now provides a bounded but actionable conclusion: the detection-balance and emission-qualification scales are selected independently from satellite observations, rather than by imposing one universal averaging count or by tuning to an external inventory. N_Q is reported as a qualification optimum, not as a truth-based accuracy optimum.