https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

Random Error and Averaging Strategies for XCO_2 Observations from Spaceborne IPDA Lidar Onboard DQ-1 Satellite

Zengchang Fan[1,2], Chuncan Fan[3], Lu Zhang[4], Lingbing Bu[1,2], Jiyu Wang[1,2], Yingxue Chen[1,2], Jinrui Guo[1,2], Qingting Zhu[1,2], Jiqiao Liu[3], Weibiao Chen[3]

[1]State Key Laboratory of Environment Characteristics and Effects for Near-space, Nanjing University of Information Science and Technology, Nanjing, 210044, China
[2]China Meteorological Administration Aerosol-Cloud and Precipitation Key Laboratory/Collaborative Innovation Centre on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing, 210044, China
[3]Wangzhijiang Innovation Center for Laser, Aerospace Laser Technology and System Department, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai, 201800, China
[4]The Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, National Satellite Meteorological Center (National Center for Space Weather) and Innovation Center for FengYun Meteorological Satellite (FYSIC), China Meteorological Administration (CMA), Beijing, 100081, China

Correspondence to: Lu Zhang (zhanglu_nsmc@cma.gov.cn) and Lingbing Bu (lingbingbu@nuist.edu.cn)

Abstract.

Carbon dioxide (CO_2) is one of the most important anthropogenic greenhouse gases in the atmosphere, and accurate monitoring of CO_2 concentration is essential for carbon flux inversion and emission reduction policies. Spaceborne integrated path differential absorption (IPDA) lidar enables day-and-night observations, high-latitude coverage, and reduced sensitivity to clouds and aerosols, showing great potential for monitoring the column-averaged dry-air mole fraction of CO_2 (XCO_2). However, single-shot XCO_2 retrievals from spaceborne IPDA lidar generally have relatively large random errors, and their along-track averaging is affected by surface conditions and data gaps. Based on 2023 DQ-1 spaceborne IPDA lidar observations, this study calculated global single-shot XCO_2 random errors using the return-signal SNRs at the online and offline wavelengths. The results show pronounced spatial heterogeneity in single-shot random errors. Larger errors occur over water bodies, permanent snow and ice, and some complex land cover types, whereas day–night differences are relatively weak. Combining MCD12C1 land cover data with random-error statistics, the global observation scenes were classified into four surface random-error categories. Allan variance analysis of continuous observation segments suggests averaging approximately 160, 200, 232, and 296 observations for the four categories, corresponding to typical along-track scales of about 54, 67, 77, and 100 km. To address data gaps and along-track discontinuities, fixed-spatial-resolution and random-error-threshold-controlled averaging strategies were compared. The fixed-spatial-resolution strategy maintains more stable spatial representativeness and is more suitable as the default averaging option, whereas the threshold-controlled strategy improves error consistency but may substantially expand the averaging window. For land–sea boundaries and other rapid surface-transition regions, a boundary mixed dynamic averaging strategy was further proposed, with surface-category

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

fractions recommended as quality indicators. These results support random-error evaluation and averaging-strategy design for spaceborne IPDA lidar XCO_2 products.

1 Introduction

Carbon dioxide (CO_2) is one of the most important anthropogenic greenhouse gases in the atmosphere, accounting for approximately 64 % of the total radiative forcing from long-lived greenhouse gases (World Meteorological Organization, 2023). The additional radiative forcing caused by CO_2 in the Earth system has not only contributed to an increase in the global mean surface temperature by approximately 0.8 °C but has also affected many weather and climate extremes worldwide, posing adverse impacts on human activities and the Earth system (IPCC, 2023). The international community has promoted actions to reduce anthropogenic CO_2 emissions through treaties such as the Kyoto Protocol and the Paris Agreement (United Nations, 1998; UNFCCC, 2015). Nevertheless, the formulation and implementation of emission reduction policies still require accurate global CO_2 monitoring data as support (United Nations, 1998; UNFCCC, 2015; Nisbet and Weiss, 2010).

In recent years, satellite observations of the column-averaged dry-air mole fraction of CO_2 (XCO_2) have been widely used for carbon monitoring and carbon emission inversion, especially in regions where ground-based observations are sparse or unavailable (Duren and Miller, 2012). Passive remote sensing instruments, represented by OCO-2/3, GOSAT, and TanSat, achieved on-orbit observations relatively early and have promoted the continuous development of XCO_2 retrieval algorithms (O'Dell et al., 2018; Noël et al., 2022; Liu et al., 2018). These instruments retrieve XCO_2 by using solar radiation reflected by the surface and have demonstrated high single-sounding retrieval precision under typical observing conditions, providing an important data basis for global carbon emission monitoring (Taylor et al., 2023; Fang et al., 2023). However, owing to their dependence on sunlight, passive remote sensing instruments have limited observing capability under conditions such as high latitudes, aerosol pollution, cloud contamination, and nighttime, making it difficult to provide high-precision XCO_2 observations in these scenarios.

Compared with passive remote sensing, spaceborne Integrated Path Differential Absorption (IPDA) lidar is an active remote sensing technique that can provide stable XCO_2 observations under both daytime and nighttime conditions. It is also relatively less sensitive to solar radiation and aerosol effects, and therefore provides an important complementary capability for XCO_2 monitoring (Fan et al., 2024). However, previous airborne and spaceborne IPDA lidar studies have shown that single-shot XCO_2 retrievals generally have relatively large random errors and require along-track averaging to reduce the observational uncertainty to a usable level (Tellier et al., 2018). For example, airborne observations related to ASCENDS showed that a multi-wavelength IPDA lidar system for CO_2 column measurements required a pulse accumulation time of 1 s to achieve an XCO_2 precision of approximately 1 ppm (Abshire et al., 2018). Airborne payload experiments related to A-SCOPE also showed that the random error of averaged XCO_2 retrievals could be reduced to below 1 ppm with an averaging window of approximately 10 s (Amediek et al., 2009, 2017). Existing validation studies of the DQ-1 Aerosol and Carbon

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

Detection Lidar (ACDL), which includes a spaceborne IPDA lidar for XCO_2 measurements, used a sliding averaging window of 296 along-track observations for comparison with TCCON measurements (Fan et al., 2024). It should be noted that averaging is not a simple accumulation of observations; its performance depends not only on the random error of individual retrievals but also on the distribution of valid observations and the spatial scale of the averaging window. Therefore, for spaceborne IPDA lidar that has already obtained real on-orbit observations, quantitatively evaluating the random error of single-shot XCO_2 retrievals and further developing an averaging strategy suitable for data gaps and transitions between surface types remain key issues affecting the accuracy and usability of XCO_2 products.

Previous studies have investigated the evaluation of random errors and the averaging of spaceborne IPDA lidar XCO_2 retrievals from different perspectives. Ehret et al. (2008) proposed a method for estimating the random error of spaceborne IPDA lidar XCO_2 retrievals from the signal-to-noise ratio of the observations. By applying this method in simulation experiments, they estimated that the XCO_2 error for a spaceborne IPDA lidar operating near 1.6 µm is approximately 0.2 %, corresponding to an absolute random error of 0.84 ppm at an along-track spatial resolution of 50 km and a mean CO_2 concentration of 420 ppm. Tellier et al. (2018) investigated in detail the bias introduced by the averaging of spaceborne IPDA lidar observations and pointed out that the AVX scheme, i.e. direct averaging of XCO_2, is affected by statistical bias induced by noise nonlinearity and by concentration averaging bias. Cao et al. (2024) analyzed the number of averaged ACDL observations over typical land regions using simulation experiments and Allan variance analysis, and indicated that the optimal averaging window length for ACDL is 50 km, corresponding to approximately 150 observations. Yang et al. (2019) and Zhu et al. (2021) simulated the random errors of ACDL observations over land and ocean using MODIS surface reflectance and aerosol optical depth products. In their averaging scheme, 148 observations were averaged over land and 296 observations were averaged over ocean, and the results showed that the ACDL random error was largest in boreal spring, reaching 0.88 ppm. Wang et al. (2020) further estimated the global random error of XCO_2 retrievals using CALIOP cloud and aerosol optical depth data, together with sea-surface wind speed data to improve the simulation of sea-surface reflectance. After averaging 148 observations over land and 296 observations over ocean, 61.24 % of the global area had a relative random error below 0.25 %, corresponding to an absolute random error of approximately 1 ppm.

Overall, previous studies have laid the foundation for the evaluation of random errors and the averaging of spaceborne IPDA lidar XCO_2 retrievals, but systematic analyses based on real on-orbit ACDL observations remain insufficient. On the one hand, most existing studies are based on simulation experiments or case studies over typical regions, and the dependence of the random error of single-shot retrievals on surface cover type and daytime or nighttime observing conditions has not been sufficiently characterized. On the other hand, existing recommendations for the number of averaged observations are usually based on relatively continuous observations and relatively homogeneous underlying surfaces. They do not fully account for data gaps caused by cloud screening, weak signals, and the removal of abnormal returns. In particular, for regions with rapid transitions in surface type, such as land–sea boundaries, a stable averaging strategy suitable for XCO_2 data product applications is still lacking. Therefore, it is important to evaluate random errors and investigate averaging strategies based on real ACDL observations.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

To support the scientific application of spaceborne IPDA lidar XCO_2 data products, this study focuses on the evaluation of random errors in XCO_2 retrievals and the optimization of averaging strategies. Section 2 introduces the observation and retrieval principles of XCO_2, the method for estimating random errors, and the Allan variance analysis approach. Section 3 evaluates the random errors of single-shot XCO_2 retrievals based on observations, analyzes their dependence on surface type and daytime or nighttime observing conditions, and determines the number of averaged observations using Allan variance analysis. Section 4 compares different averaging strategies and proposes an optimized method for regions with rapid transitions in surface type, such as land–sea boundaries. Section 5 discusses the differences between observation-based and simulation-based random error studies, as well as the discrepancy between the Allan standard deviation and the theoretical random error.

2 Principle and methods

2.1 Spaceborne IPDA lidar XCO_2 observation and retrieval principle

Spaceborne IPDA lidar alternately transmits two laser pulses with nearby wavelengths in the vicinity of a CO_2 absorption line. By exploiting the difference in CO_2 absorption between the online and offline wavelengths, the CO_2 differential absorption optical depth is obtained from return signals reflected by hard targets, such as the Earth's surface or cloud tops, and is then used to retrieve XCO_2. For a stable spaceborne IPDA lidar system, the time separation between a pair of online and offline laser pulses is very short. Therefore, their footprints on the hard target nearly overlap, with a center-to-center distance of only approximately 1.5 m. Assuming that the difference in surface reflectance between the two pulses can be neglected, the one-way CO_2 differential absorption optical depth (DAOD) can be expressed as follows:

$$DAOD = \frac{1}{2} \cdot ln\left(\frac{P_{off} \cdot E_{on}}{P_{on} \cdot E_{off}}\right) \tag{1}$$

where P is the received return energy, E is the transmitted pulse energy, and the subscripts on and off denote the online and offline wavelengths, respectively.

The integrated weighting function (IWF) is calculated by assuming that all dry-air molecules along the laser propagation path are CO_2 molecules, and represents the one-way optical depth corresponding to this assumption:

$$IWF = \int_{R_T}^{R_{TOA}} \frac{N_A \cdot p(r) \cdot \Delta\sigma_{CO_2}\big(p(r), T(r)\big)}{RT(r)\left(1 + X_{H_2O}(r)\right)} dr \tag{2}$$

where N_A denotes the Avogadro constant, p the atmospheric pressure, T the temperature, $\Delta\sigma_{CO_2}$ the CO_2 differential absorption cross section, R the ideal gas constant, and X_{H_2O} the water vapor mixing ratio.

The dry-air mixing ratio of CO_2 along the laser propagation path, XCO_2, can then be obtained as the ratio of the one-way CO_2 differential absorption optical depth (DAOD) to the integrated weighting function (IWF):

$$XCO_2 = \frac{DAOD}{IWF} \tag{3}$$

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

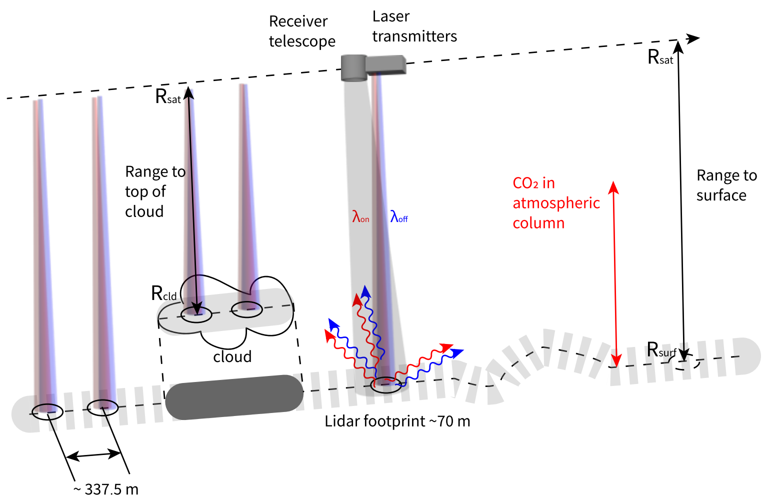


Figure 1. Schematic illustration of the observation principle of spaceborne IPDA lidar.

2.2 Error evaluation method for spaceborne IPDA lidar XCO_2 retrievals

To quantitatively evaluate the random errors of spaceborne IPDA lidar XCO_2 retrievals, this study estimates the random error of single-shot retrievals from the signal-to-noise ratio of the return signals through error propagation, and calculates the theoretical random error for different averaging windows. Allan variance analysis is further applied to continuous observation segments to characterize the variation of XCO_2 retrieval errors with the number of averaged observations and to determine the optimal averaging scale.

2.2.1 Random error evaluation method for XCO_2 retrievals based on the signal-to-noise ratio

Under ideal conditions, the random error of single-shot XCO_2 retrievals can be characterized by the dispersion of a sufficiently large number of independent repeated observations. In practical spaceborne IPDA lidar observations, however, it is difficult to obtain sufficient repeated measurements over the same surface target under identical atmospheric conditions. Therefore, the random error of each single-shot XCO_2 retrieval cannot be directly estimated from repeated observations at a single location. Based on existing error propagation theory for IPDA lidar (Ehret et al., 2008), this study estimates the random error of each single-shot XCO_2 retrieval using the signal-to-noise ratios of the online and offline return signals.

According to Eqs. (1) and (3), the XCO_2 and DAOD obtained from the ith observation can be expressed as

$$XCO_{2,i} = \frac{DAOD_i}{IWF_i} \tag{4}$$

$$DAOD_i = \frac{1}{2} \cdot ln\left(\frac{P_{off,i} \cdot E_{on,i}}{P_{on,i} \cdot E_{off,i}}\right) \tag{5}$$

Assuming that the noise terms in the online and offline return signals and in the transmitted pulse energy measurements are independent, the random error of XCO_2 can be derived by error propagation as

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

$$\sigma_{XCO_{2,i}} = \frac{1}{2IWF_i}\sqrt{Var\left(\ln P_{off,i}\right) + Var\left(\ln P_{on,i}\right) + Var\left(\ln E_{off,i}\right) + Var\left(\ln E_{on,i}\right)} \tag{6}$$

Because the monitored transmitted pulses usually have high signal-to-noise ratios, their contribution to the random error of XCO_2 is neglected. For return signals with sufficiently high signal-to-noise ratios, the first-order Taylor expansion gives

$$Var(\ln P) \approx \frac{1}{SNR(P)^2} \tag{7}$$

Thus, the random error of the ith XCO_2 retrieval can be approximated as

$$\sigma_{XCO_{2,i}} \approx \frac{1}{2IWF_i}\sqrt{\frac{1}{SNR_{off,i}^2} + \frac{1}{SNR_{on,i}^2}} \tag{8}$$

For an averaging window containing N valid observations, if the random errors of individual observations are independent, the random error of the averaged XCO_2 retrieval is

$$\sigma_{XCO_2-ave} = \frac{1}{N}\sqrt{\sum_{i=1}^{N} \sigma_{XCO_{2,i}}^2} \tag{9}$$

When the random errors of the N observations within the averaging window are approximately equal, Eq. (9) can be simplified as

$$\sigma_{XCO_2-ave} \approx \frac{\sigma_{XCO_2}}{\sqrt{N}} \tag{10}$$

In the processing of spaceborne IPDA lidar observations, the signal-to-noise ratio is calculated as the ratio between the pulse signal intensity P and the standard deviation of the noise segment. As shown in Fig. 2, after multi-peak identification, the noise segment is selected from the part of the waveform before the return signal. Specifically, if X denotes the index of the earliest detected pulse peak, the noise segment is taken from the initial waveform index to $(X - 50)$.

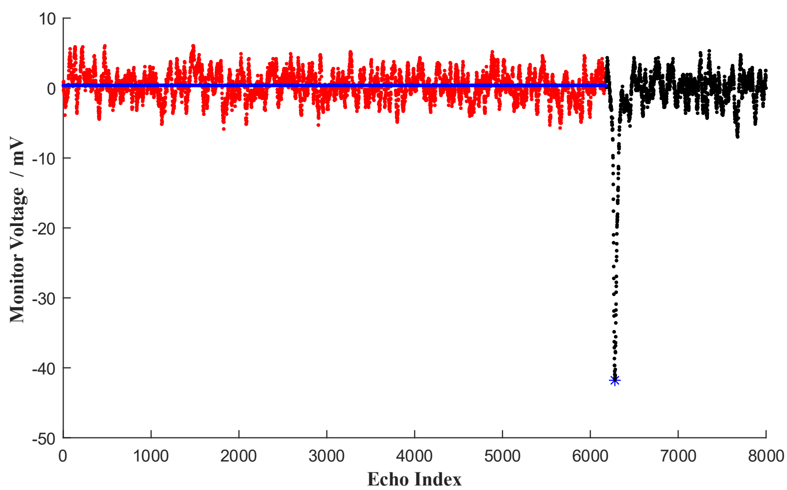


Figure 2. Schematic illustration of signal extraction and noise-segment identification for spaceborne IPDA lidar observations. The red region denotes the noise segment, and the blue line denotes the mean value of the noise segment.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

2.2.2 Allan variance analysis for determining the number of averaged XCO_2 observations

According to Eq. (10), the random error of XCO_2 retrievals can be effectively reduced by averaging multiple observations, mainly because independent random noise tends to cancel out during the averaging process. However, real spaceborne IPDA lidar observations are affected not only by random errors but may also include correlated error components. Therefore, the total error of averaged XCO_2 retrievals can be expressed as (Kulawik et al., 2019):

$$\sigma_{\mathrm{total}} = \sqrt{\frac{\sigma_{\mathrm{random}}^2}{\mathrm{N}} + \sigma_{\mathrm{corr}}^2} \tag{11}$$

where $\sigma_{\mathrm{random}}^2$ represents the random error variance, and σ_{corr}^2 represents the correlated error variance, including spatially correlated systematic biases. As the number of averaged observations (N) increases, the term $(\frac{\sigma_{\mathrm{random}}^2}{\mathrm{N}})$ gradually decreases, whereas the correlated error does not decrease according to the $(\frac{1}{\sqrt{N}})$ relationship. In addition, an excessively large averaging window reduces the along-track spatial resolution of XCO_2 retrievals. Therefore, the number of averaged observations should be selected by balancing error reduction against along-track spatial resolution.

Based on this consideration, Allan variance analysis is introduced to evaluate the overall error level of the XCO_2 observation sequence under different numbers of averaged observations and to provide a basis for determining the averaging number (Wagner and Plusquellic, 2016; Fix et al., 2017; Kawa et al., 2018). It should be noted that Eq. (11) usually requires comparisons with high-precision ground-based XCO_2 observations to separate random and correlated error components. By contrast, Allan variance analysis based on spaceborne IPDA lidar along-track observations does not directly separate these error components. Therefore, the difference between the Allan standard deviation and the theoretical random error after averaging should not be simply interpreted as the correlated error, because it may also include real along-track spatial variations in XCO_2.

For a continuous XCO_2 observation sequence $(XCO_2^1, XCO_2^2, \ldots, XCO_2^N)$, let the averaging window length be (L). The mean XCO_2 of the kth group is given by

$$\overline{XCO_2^k}(L) = \frac{1}{L}\sum_{i=(k-1)L+1}^{kL} XCO_2^i \tag{12}$$

where $(k = 1,2, \ldots, K)$, and $(K = \frac{N}{L})$ is the number of available groups. The Allan variance is defined as one half of the mean squared difference between two adjacent group averages:

$$\sigma_A^2(L) = \frac{1}{2(K-1)}\sum_{k=1}^{K-1}[\overline{XCO_2^{k+1}}(L) - \overline{XCO_2^k}(L)]^2 \tag{13}$$

The corresponding Allan standard deviation is

$$\sigma_A(L) = \sqrt{\frac{1}{2(K-1)}\sum_{k=1}^{K-1}[\overline{XCO_2^{k+1}}(L) - \overline{XCO_2^k}(L)]^2} \tag{14}$$

Here, ($\sigma_A(L)$) describes the overall difference between adjacent averaged results when every (L) observations are averaged. If ($\sigma_A(L)$) decreases approximately following ($L^{-\frac{1}{2}}$) as (L) increases, the variability of the observation sequence is mainly controlled by independent random errors, and increasing the averaging number can still effectively improve the stability of the retrievals. When the decreasing trend of ($\sigma_A(L)$) becomes weaker, independent random errors are no longer the dominant factor, and correlated errors or real spatial variations in XCO_2 begin to make a more significant contribution to the sequence variability.

3 Random error and averaging-number analysis of spaceborne IPDA lidar XCO_2 retrievals

3.1 Random error of single-shot XCO_2 retrievals from spaceborne IPDA lidar

This section estimates the random error of single-shot XCO_2 retrievals from spaceborne IPDA lidar based on Eq. (8), and analyzes its overall level and spatial distribution. To systematically characterize its variability and provide a basis for determining the number of averaged observations and evaluating averaging strategies, the differences in single-shot random errors among different surface cover types and between daytime and nighttime observations are further analyzed.

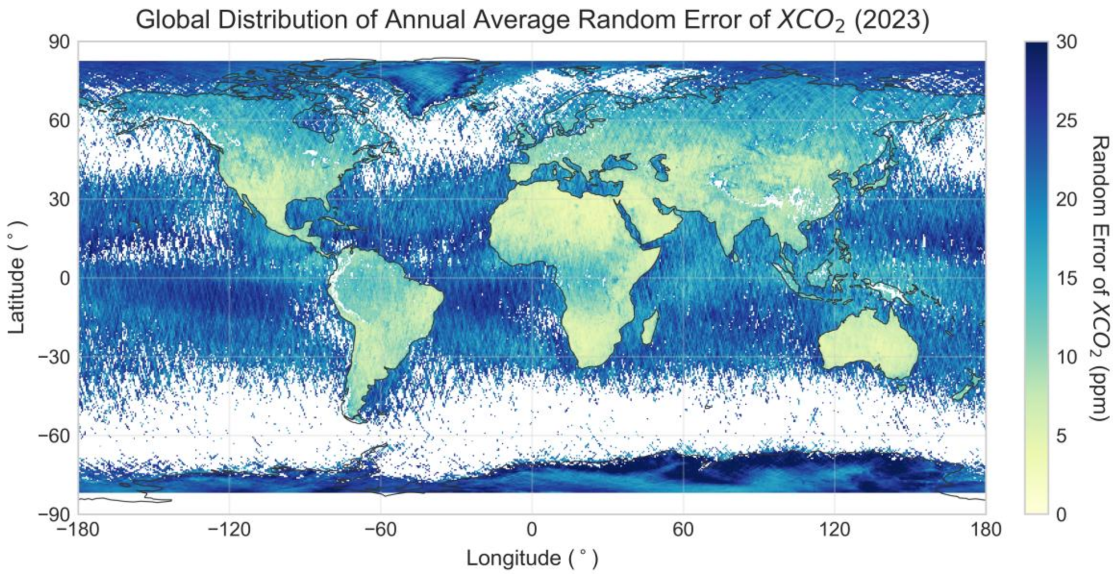
3.1.1 Global distribution of random errors in single-shot XCO_2 retrievals from spaceborne IPDA lidar

Based on the 2023 spaceborne IPDA lidar L2B data product, the signal-to-noise ratios of the online and offline return signals were first extracted, and the random error of each single-shot XCO_2 retrieval was calculated using Eq. (8). To reduce the influence of nonlinear statistical bias introduced by noise in the AVX scheme, an additional quality control step was applied using a combined signal-to-noise ratio threshold for the online and offline returns. Cloud and surface observations were then classified according to the difference between the footprint elevation from the spaceborne IPDA lidar and a high-resolution digital elevation model, and cloud-top observations were excluded. To reduce the influence of a small number of extreme values on the statistical results, the monthly random-error samples were filtered using the 5th–95th percentile range, and only observations within this range were retained. Finally, all single-shot retrievals were projected onto a 0.5° × 0.5° global grid to analyze the spatial distribution of random errors. To ensure basic representativeness of the gridded statistics, grid cells with fewer than 300 valid observations during the year were excluded and set to NaN.

Using the method described above, the global distribution of annual mean random errors for single-shot XCO_2 retrievals from spaceborne IPDA lidar in 2023 was obtained, as shown in Fig. 3. The single-shot random error exhibits pronounced spatial heterogeneity, indicating that the retrieval precision is strongly affected by real observing conditions. In general, random errors over most land regions are clearly lower than those over ocean, mainly because land surfaces usually provide higher signal-to-noise ratios for the return signals. By contrast, ocean observations generally show larger random errors because of the lower sea-surface reflectance, and their regional differences may be related to variations in sea-surface reflectance caused by changes in wind speed (Wang et al., 2020). Large random errors are also found in high-latitude regions,

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

especially over Antarctica and Greenland, which may be associated with the low surface reflectance of ice- and snow-covered surfaces. The blank or sparse areas in Fig. 3 mainly reflect the effects of cloud contamination, weak-signal filtering, and insufficient valid observations under strict quality-control criteria.



215 **Figure 3. Global distribution of annual mean random errors for single-shot XCO_2 retrievals from spaceborne IPDA lidar in 2023.**

3.1.2 Surface-type dependence of random errors in single-shot XCO_2 retrievals from spaceborne IPDA lidar

To examine the effect of surface cover on the random errors of single-shot XCO_2 retrievals, the MCD12C1 land cover product was used with the IGBP classification scheme. To match the 0.5° × 0.5° grid used for the random-error statistics, the original 0.05° land cover data were resampled to 0.5° using the majority method, and the resulting global surface cover distribution is shown in Fig. 4. The 2023 single-shot random-error estimates were then spatially matched with the resampled land cover data and grouped by IGBP surface type. For each surface type, the mean random error was calculated to quantify the influence of different underlying surfaces on spaceborne IPDA lidar XCO_2 retrieval precision, as shown in Fig. 5.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

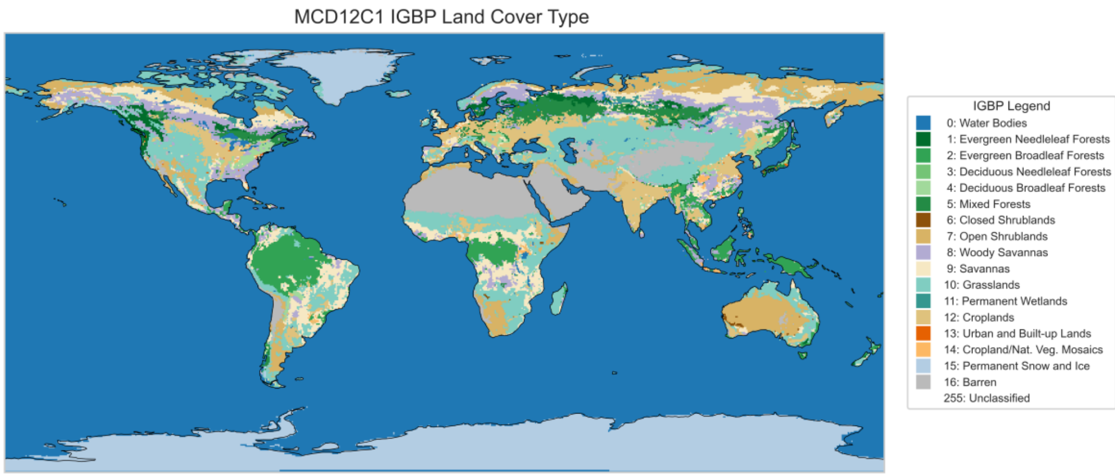


Figure 4. Global distribution of surface cover types derived from the IGBP classification scheme.

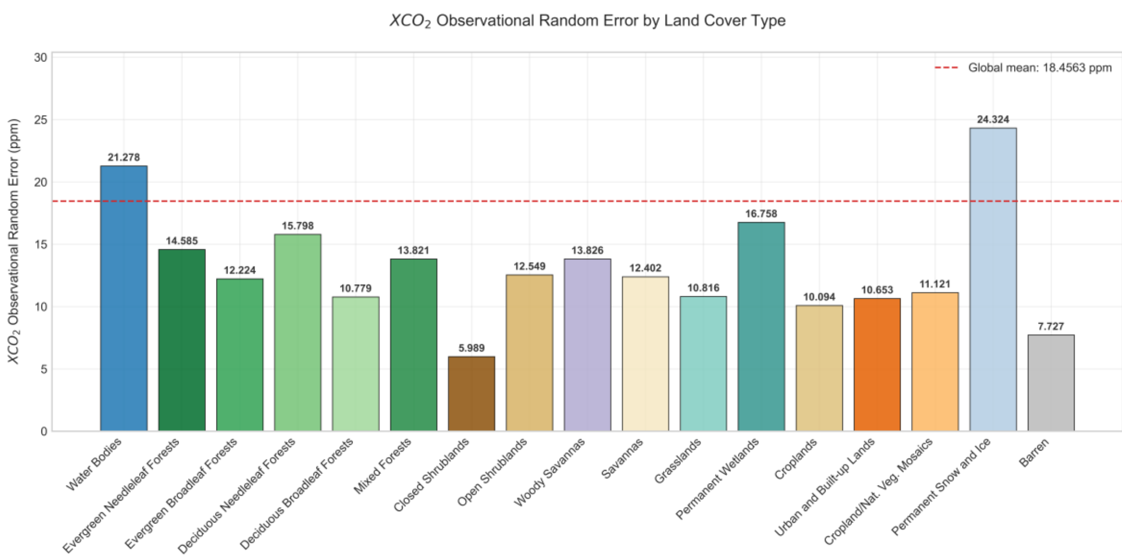


Figure 5. Random errors of single-shot XCO_2 retrievals over different surface cover types.

The random errors of single-shot XCO_2 retrievals from spaceborne IPDA lidar show pronounced differences among surface cover types. The global mean single-shot random error is approximately 18.46 ppm. Most land cover types have random errors below this mean level, whereas water bodies and permanent snow- and ice-covered regions show substantially larger errors. The random error over water bodies is approximately 21.28 ppm, which is higher than that over most land surfaces, mainly because water surfaces have much lower reflectance than land surfaces at the 1572 nm laser wavelength. Permanent snow- and ice-covered regions have the largest random error, approximately 24.32 ppm. This behavior differs from simulation-based results and is discussed further in Sect. 5.1.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

In contrast, most vegetation-covered surfaces and barren land show relatively low single-shot random errors. The random errors for evergreen broadleaf forests, deciduous broadleaf forests, grasslands, croplands, urban and built-up lands, and cropland/natural vegetation mosaics are mostly within approximately 10–14 ppm, clearly below the global mean. Closed shrublands and barren land have the lowest random errors, approximately 5.99 and 7.73 ppm, respectively, possibly because of stronger and more stable surface returns in these regions. Forest-covered regions also show distinct differences among forest types. In general, evergreen forests have larger random errors than deciduous forests, and needleleaf forests have larger random errors than broadleaf forests. For example, the random error over evergreen needleleaf forests is approximately 14.59 ppm, higher than the 12.22 ppm over evergreen broadleaf forests; similarly, the random error over deciduous needleleaf forests is approximately 15.80 ppm, higher than the 10.78 ppm over deciduous broadleaf forests. These results suggest that, even within forest-covered regions, differences in canopy structure and optical reflectance characteristics can affect the signal-to-noise ratio of the return signals and thus the random error of XCO_2 retrievals.

3.1.3 Day–night differences in the random errors of single-shot XCO_2 retrievals from spaceborne IPDA lidar

Considering that solar background noise reflected by the surface may affect return-signal extraction and random-error estimation for spaceborne IPDA lidar, this study further calculates the random errors of single-shot XCO_2 retrievals separately for daytime and nighttime observing conditions. The results are shown in Figs. 6. In terms of spatial distribution, the regional patterns of daytime and nighttime random errors are generally consistent. Single-shot random errors are generally lower over land regions and higher over ocean and permanent snow- and ice-covered regions, indicating that daytime or nighttime conditions do not substantially change the main spatial pattern of single-shot random errors.

In terms of the overall error level, the difference between daytime and nighttime single-shot random errors is not obvious. The mean single-shot random error is 20.23 ppm during daytime and 20.63 ppm during nighttime, with a difference of only approximately 0.40 ppm. It should be noted that the additional quality control has a very weak influence on the day–night difference in single-shot random errors. Without the additional quality control, the mean single-shot random error is 23.51 ppm during daytime and 23.92 ppm during nighttime, with a difference of only approximately 0.41 ppm. Although daytime observations may theoretically be affected by solar background noise reflected by the surface, the statistical results show that this influence is not significant in the current spaceborne IPDA lidar observations. This may be related to the good narrow-band filtering performance of the lidar system, the stability of temperature control, and the background subtraction process in return-signal processing, which effectively suppress the contribution of solar background noise to the final random errors of single-shot XCO_2 retrievals.

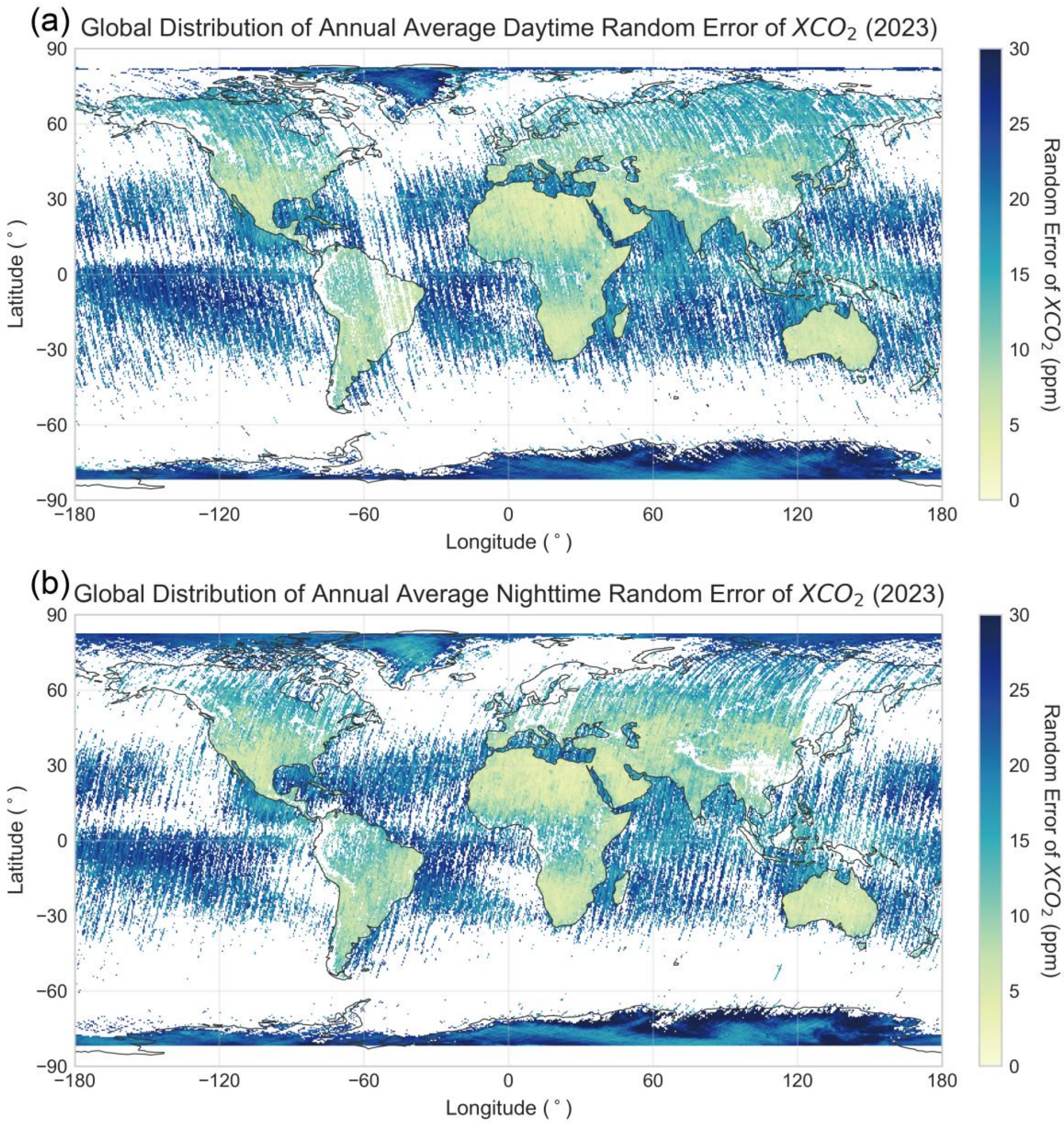


Figure 6. Global distribution of annual mean random errors for single-shot XCO_2 retrievals from spaceborne IPDA lidar in 2023 under (a) daytime and (b) nighttime conditions.

3.2 Analysis of the averaging number for spaceborne IPDA lidar XCO_2 retrievals

3.2.1 Land–ocean-based analysis of the averaging number

Based on the 2023 spaceborne IPDA lidar L2B data product, continuous observation segments were further selected after data quality control, and Allan variance analysis was applied to examine how XCO_2 retrieval errors vary with the number of averaged observations. Because Allan variance analysis requires sufficiently long and continuous observation sequences,

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

different segment-length thresholds were applied to land and ocean observations. The continuous segments used for the averaging-number analysis are summarized in Table 1. For land observations, valid measurements were relatively abundant, and 11 008 continuous segments longer than 148 × 7 observations were obtained. In contrast, ocean observations were more strongly affected by quality screening, resulting in fewer long continuous segments. To retain a sufficient sample size for Allan variance analysis, the segment-length threshold for ocean observations was relaxed, yielding 348 continuous segments longer than 296 × 5 observations.

Table 1. Continuous observation segments used for the land–ocean-based analysis of the averaging number.

Land Cover	Filter Criteria	Number of Consecutive Observation Segments
Land	N > 148*7	11 008
Ocean	N > 296*5	348

Based on these continuous segments, Allan variance analysis was performed separately for land and ocean observations, and the results are shown in Fig. 7. For both land and ocean continuous segments, the Allan standard deviation was calculated over an averaging range of 2–400 observations with a step of two observations. The results were then statistically averaged over 11 008 land segments and 348 ocean segments, respectively.

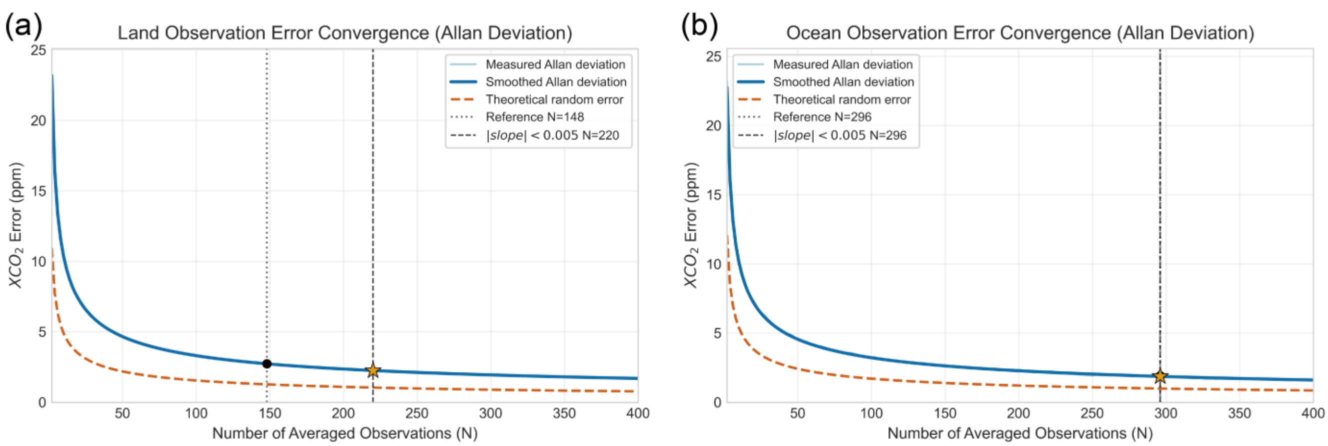


Figure 7. Variations in the Allan standard deviation and theoretical averaged random error with the number of averaged observations for (a) land and (b) ocean observations. The solid blue line represents the smoothed Allan standard deviation calculated from real observations, the light-blue line represents the unsmoothed Allan standard deviation, and the orange dashed line represents the theoretical averaged random error estimated under the assumption of independent random errors.

The Allan standard deviations for both land and ocean generally decrease as the number of averaged observations increases, but the decreasing rate gradually becomes weaker. In this study, a slope threshold of 0.005 was used to identify the point at which the decrease in the Allan standard deviation begins to flatten. This threshold is half of that used by Cao et al. (2024), and therefore provides a relatively stricter criterion for identifying the flattening point. Based on this criterion, the averaging numbers for land and ocean observations are approximately 220 and 296, respectively. The ocean result is generally consistent with previous simulation-based recommendations, indicating that ocean observations, as a relatively homogeneous

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

but weak-return observing scenario, usually require more observations for averaging. By contrast, the land averaging number is clearly higher than the previously suggested value of 148 observations from simulation-based studies. This indicates that treating all land surfaces as a single category is affected by complex surface types, such as permanent snow and ice and permanent wetlands, which increases the overall error level and thus the required averaging number. In addition, the theoretical averaged random error is lower than the observed Allan standard deviation, suggesting that the real observation sequence may also contain non-independent error components, such as along-track correlated errors and real spatial variations in XCO_2.

Overall, the land–ocean-based results indicate that using only a simple land–ocean separation to determine the averaging number has certain limitations. Applying 220 observations as a uniform averaging number over all land surfaces would help control errors over low-reflectance or complex land regions, but it would also over-average many low-error land regions and reduce the along-track spatial resolution of the XCO_2 product. Therefore, the subsequent averaging strategy should not rely only on a simple land–ocean classification. Instead, representative averaging numbers should be assigned to regions with similar error levels by further considering surface type or the single-shot random-error level.

3.2.2 Surface-type-based analysis of the averaging number

Considering the pronounced differences in single-shot XCO_2 random errors among surface cover types and the influence of surface-type continuity on the subsequent design of averaging windows, the global surface was further regrouped according to the single-shot random-error level and the spatial continuity of surface cover types. The classification criteria are listed in Table 2. The regrouped surface categories include low-random-error land surfaces, moderate-random-error land surfaces, relatively high-random-error land surfaces, and high-random-error regions, with mean single-shot XCO_2 random errors of 9.597, 13.235, 16.278, and 22.801 ppm, respectively. Permanent snow and ice and water bodies show the largest random errors and are therefore grouped together as high-random-error regions, which are discussed separately in the following averaging-number analysis. This classification avoids neglecting the internal heterogeneity of land surfaces in a simple land–ocean separation and provides a basis for assigning representative averaging numbers to regions with different error levels.

Table 2. Grouping of surface cover types based on random errors in XCO_2 retrievals.

Category	Mean random error of individual XCO_2 observations	IGBP Code	Land Cover
Low random error land surfaces	9.597 ppm	4	Deciduous Broadleaf Forests
		6	Closed Shrublands
		10	Grasslands
		12	Croplands
		13	Urban and Built-up Lands
		14	Cropland/Nat. Veg. Mosaics

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

		16	Barren
Moderate random error land surfaces	13.235 ppm	1	Evergreen Needleleaf Forests
		2	Evergreen Broadleaf Forests
		5	Mixed Forests
		7	Open Shrublands
		8	Woody Savannas
		9	Savannas
Relatively high random error land surfaces	16.278 ppm	3	Deciduous Needleleaf Forests
		11	Permanent Wetlands
High random error regions	22.801 ppm	15	Permanent Snow and Ice
		0	Water Bodies

The global distribution of the regrouped surface categories is shown in Fig. 8 and is used in the following Allan variance analysis to compare the decay characteristics of XCO_2 observation errors among different random-error categories and to determine representative averaging numbers under different surface conditions.

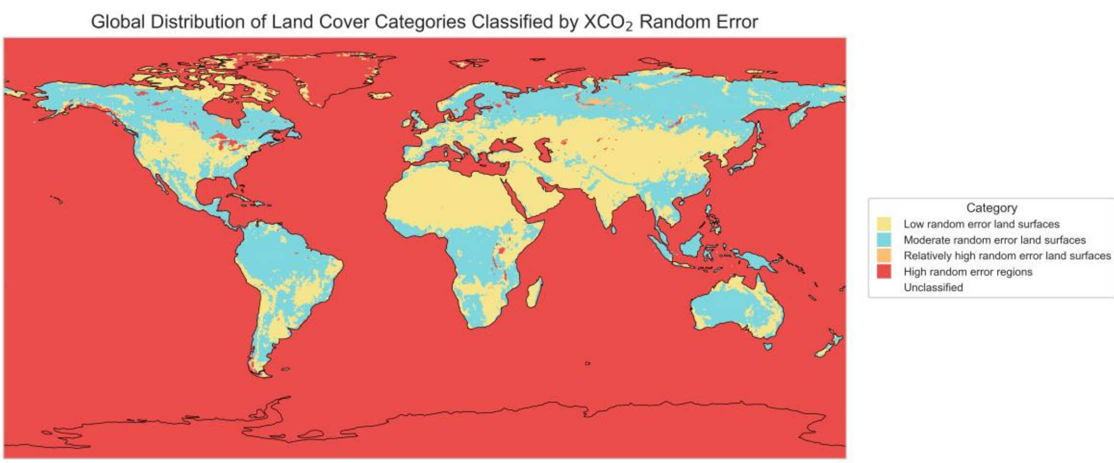


Figure 8. Global distribution of land cover categories regrouped according to random errors in XCO_2 retrievals.

Table 3 summarizes the number of valid continuous segments satisfying the segment-length requirement ($N > 1200$) under the regrouped surface categories, which were used to support the Allan variance analysis for different surface types. The low- and moderate-random-error land surfaces provide 2367 and 777 continuous segments, respectively, indicating that their spatial continuity is sufficient for statistical analysis. No valid continuous segment is obtained for the relatively high-random-error land surfaces, suggesting that these surfaces are spatially fragmented and are not suitable for an independent

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

averaging-number analysis. The high-random-error regions provide 6931 continuous segments, mainly because water bodies and permanent snow and ice have large-scale continuous distributions.

Table 3. Information on the continuous observation segments used for the averaging-number analysis under the regrouped surface classification scheme.

Category	Filter Criteria	Number of Consecutive Observation Segments
Low random error land surfaces	N > 1200	2367
Moderate random error land surfaces	N > 1200	777
Relatively high random error land surfaces	N > 1200	0
High random error regions	N > 1200	6931

Figure 9 shows the variation in XCO_2 observation errors with the number of averaged observations for different surface random-error categories. Overall, the Allan standard deviations for the three surface categories decrease as the averaging number increases, but their error levels and flattening points differ substantially. The low-random-error land surfaces show the lowest Allan standard deviation, with an averaging number of approximately 160 identified using the slope criterion. At this point, the theoretical random error is approximately 0.4 ppm. The moderate-random-error land surfaces show slightly larger Allan standard deviations, with the flattening point occurring near N = 200, corresponding to an Allan standard deviation of approximately 1.5 ppm and a theoretical random error of approximately 0.5 ppm. The high-random-error regions have Allan standard deviations clearly higher than those of the first two categories, with an averaging number of approximately 296 identified using the slope criterion. However, the Allan standard deviation remains approximately 2.4 ppm at this point, while the theoretical random error is approximately 1.2 ppm. The theoretical random error remains slightly higher than 1 ppm after averaging 296 observations, mainly because the high-random-error regions includes permanent snow- and ice-covered surfaces with large single-shot random errors. Although the random error could be further reduced, even to below 0.5 ppm, by increasing the averaging number, this would require a substantially larger averaging window and would inevitably reduce the along-track spatial resolution. This indicates that observation errors in these regions remain relatively large even after averaging over a large window, and that further increasing the averaging number provides only limited improvement in precision.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

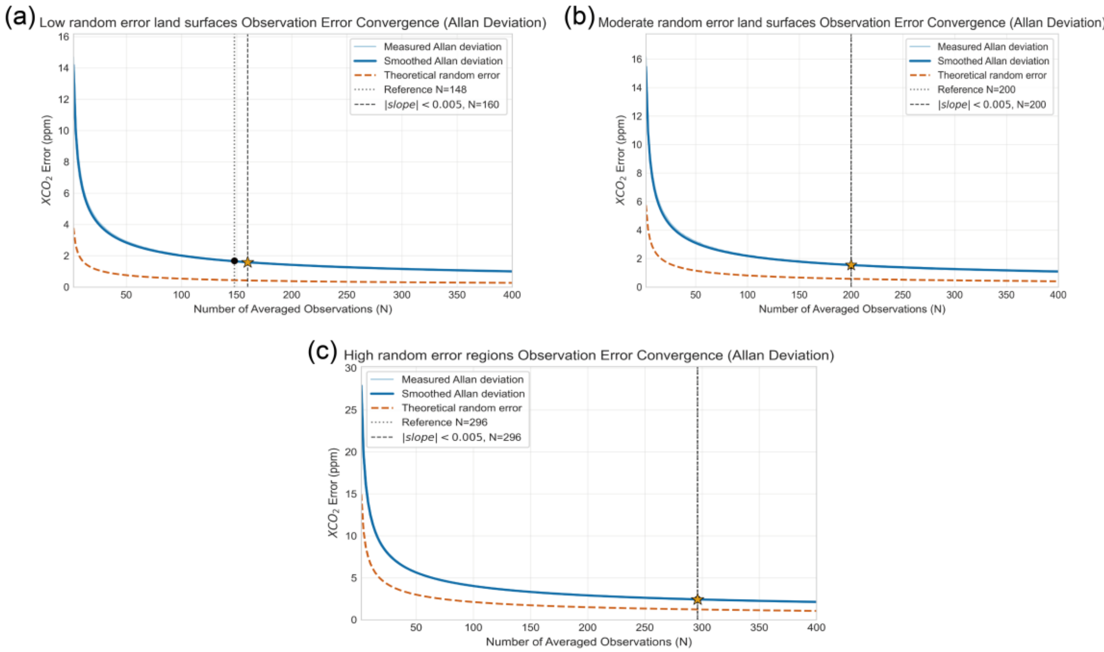


Figure 9. Allan variance analysis of XCO_2 observation error convergence with increasing averaging number for different surface random-error categories. (a) Low-random-error land surfaces, (b) moderate-random-error land surfaces, and (c) high-random-error regions . The blue lines represent the Allan standard deviations derived from real observations, the orange dashed lines represent the theoretical averaged random errors under the assumption of independent random errors.

The candidate averaging numbers estimated from the surface random-error classification are close to those determined by the Allan variance analysis, indicating that grouping surface types according to the single-shot random-error level can effectively reflect differences in error convergence among different surface conditions. This is physically reasonable because the random error of single-shot XCO_2 retrievals is mainly controlled by the signal-to-noise ratios of the online and offline returns, which are further affected by the reflectance characteristics of different surface types. Therefore, assigning representative averaging numbers based on single-shot random-error levels can provide a practical basis for designing the averaging strategy.

For the relatively high-random-error land surfaces, no stable Allan variance analysis could be performed because sufficiently long and continuous valid observation segments were not available after quality control. Therefore, the recommended averaging number for this category was estimated from its single-shot random-error level. Its mean single-shot random error is 16.278 ppm, which lies between that of the moderate-random-error land surfaces (13.235 ppm) and the high-random-error regions (22.801 ppm). Linear interpolation between the corresponding recommended averaging numbers of 200 and 296 gives an estimated averaging number of approximately 232 for the relatively high-random-error land surfaces.

In summary, the convergence characteristics of XCO_2 observation errors differ clearly among surface random-error categories. The recommended averaging numbers are approximately 160 for low-random-error land surfaces, 200 for moderate-random-error land surfaces, 232 for relatively high-random-error land surfaces, and 296 for high-random-error

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

regions. Compared with a simple land–ocean-based strategy, the stratified averaging strategy based on single-shot random errors and surface type can better balance error control and spatial resolution. It avoids unnecessary over-averaging in low-random-error land regions while providing stronger error suppression in high-random-error regions, thereby improving the applicability of spaceborne IPDA lidar XCO_2 data products under different observing conditions.

4 Analysis of averaging strategies for spaceborne IPDA lidar XCO_2 retrievals

In practical spaceborne IPDA lidar XCO_2 retrievals, some observations are missing because of signal quality control, surface-return screening, and other valid-observation constraints. As a result, surface XCO_2 observations are usually not fully continuous along track. Directly averaging discontinuous observations may introduce real spatial differences in XCO_2 between distant locations into the averaged result, thereby weakening the reduction of random errors and degrading the spatial representativeness of the averaged XCO_2 product. Therefore, it is necessary to design appropriate averaging strategies for real observations with data gaps and spatial discontinuities, so as to reduce random errors while maintaining data availability and spatial representativeness.

4.1 Analysis of XCO_2 averaging strategies for homogeneous surface types

4.1.1 Fixed-spatial-resolution averaging strategy

As shown in Fig. 10, the fixed-spatial-resolution strategy uses each XCO_2 observation as the center and averages valid observations from the same surface category within a predefined spatial window before and after the center point, yielding the averaged XCO_2 and its random error. Because cloud contamination, weak-signal filtering, and abnormal-return screening can lead to discontinuities in surface observations, the actual number of observations included within a fixed spatial window may be smaller than the theoretical value. As a result, the random error after averaging still depends on the number of valid observations. Based on the results above, fixed averaging windows of approximately 54, 67, 77, and 100 km are applied to low-random-error land surfaces, moderate-random-error land surfaces, relatively high-random-error land surfaces, and high-random-error regions, respectively, corresponding to approximately 160, 200, 232, and 296 observations. The averaged results are further evaluated according to whether the random error after averaging is lower than 1 ppm.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

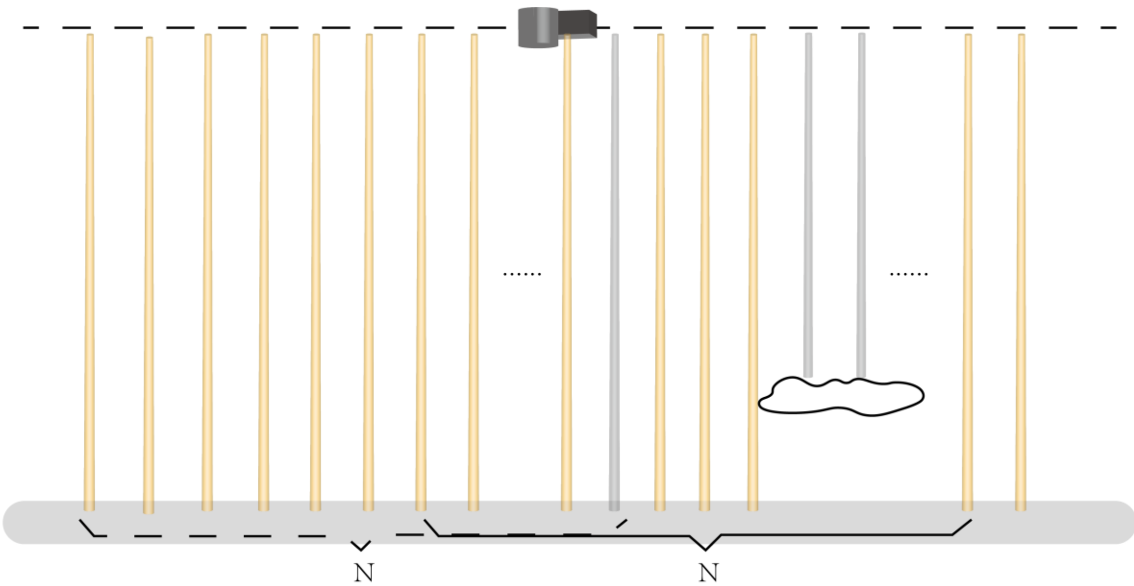


Figure 10. Schematic illustration of the fixed-spatial-resolution averaging strategy.

Figure 11 shows the distribution of random errors in averaged XCO_2 retrievals for different surface random-error categories under the fixed-spatial-resolution strategy. The black dashed line denotes the 1 ppm random-error threshold. Overall, as the surface random-error category increases from low to high, the distribution of averaged random errors gradually shifts to the right, indicating that the error-control performance still differs markedly among surface types even when category-specific fixed averaging windows are used.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

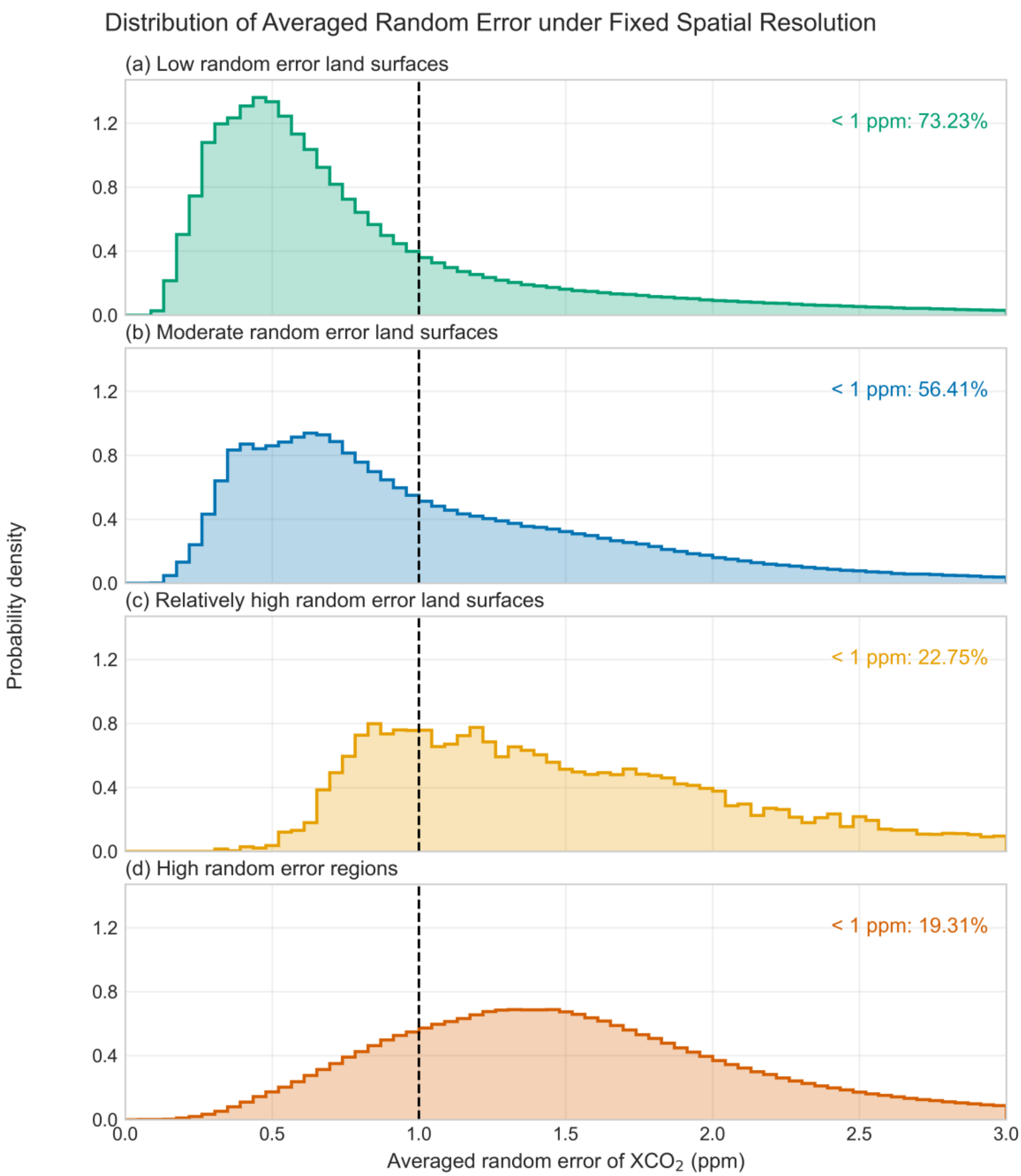


Figure 11. Statistical distribution of random errors in averaged XCO_2 retrievals under the fixed-spatial-resolution averaging strategy.

400 For low-random-error land surfaces, the averaged random errors are mainly below 1 ppm, with 73.23 % of the retrievals satisfying the 1 ppm threshold. This indicates that an averaging window of approximately 54 km, corresponding to 160 observations, can effectively control the random error for this category. For moderate-random-error land surfaces, the

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

distribution shifts toward larger errors, and the fraction below 1 ppm decreases to 56.41 %. This suggests that the 67 km window, corresponding to 200 observations, reduces random errors but is more strongly affected by return-signal quality and observation gaps. For relatively high-random-error land surfaces and high-random-error regions, most averaged random errors remain above 1 ppm, with only 22.75 % and 19.31 % of the retrievals below the threshold, respectively. In particular, the distribution peak for high-random-error regions is clearly located to the right of 1 ppm, indicating that random errors in these regions are difficult to reduce reliably to the target level even with a larger fixed averaging window.

Therefore, the fixed-spatial-resolution strategy can maintain a relatively consistent product spatial scale, but its error-control capability depends strongly on surface type. It performs well over low-random-error land regions, whereas for high-random-error regions, the current fixed spatial windows alone are insufficient to ensure that the averaged XCO_2 random error remains below 1 ppm.

4.1.2 Random-error-threshold-controlled averaging strategy

As shown in Fig. 12, the random-error-threshold-controlled strategy uses each valid spaceborne IPDA lidar XCO_2 observation as the center and gradually expands the averaging range to include valid observations from the same surface category before and after the center point. The expansion continues until the random error of the averaged XCO_2 retrieval falls below 1 ppm or the available range of the current continuous observation segment is reached. Unlike the fixed-spatial-resolution strategy, this method does not prescribe a fixed averaging window. Instead, it adaptively determines the number of observations and the spatial range used for averaging according to the single-shot random-error level and the distribution of valid observations. This strategy gives priority to ensuring that the averaged XCO_2 random error meets the specified threshold, but in regions with weak signals, frequent data gaps, or large single-shot random errors, it may require observations over a longer spatial range and thus reduce the spatial representativeness of the averaged results.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

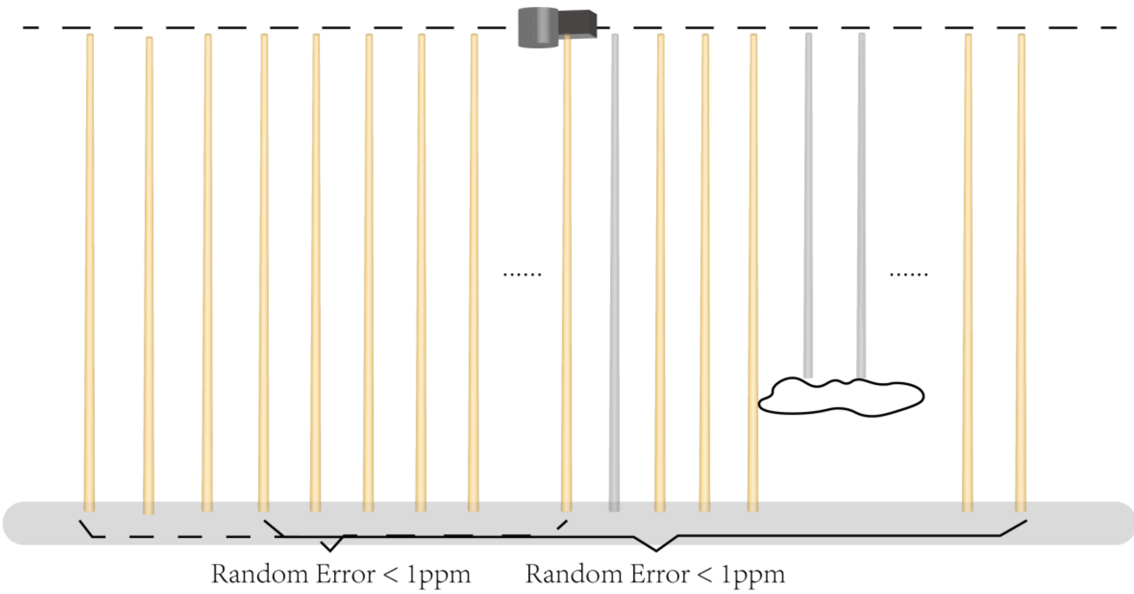
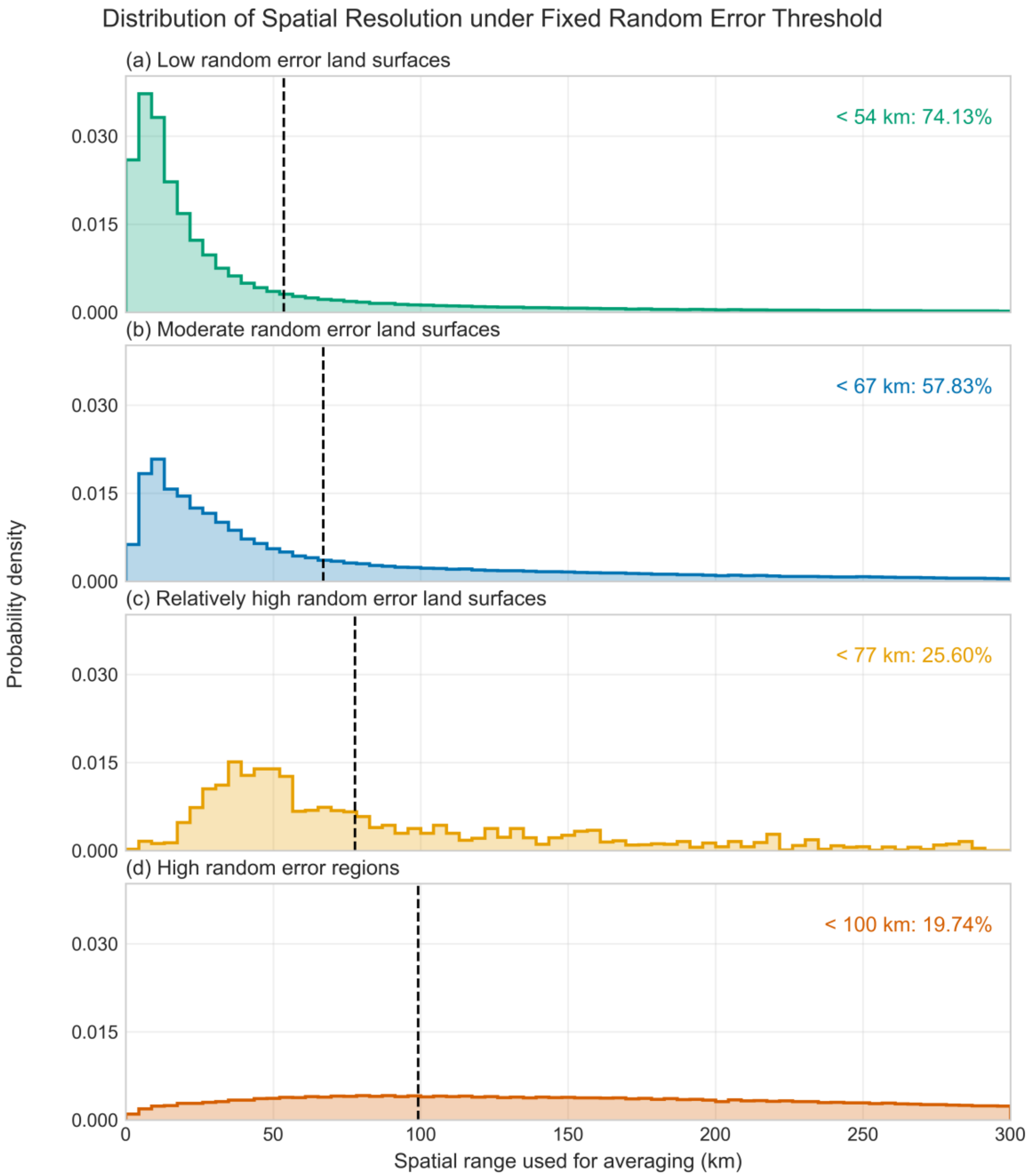


Figure 12. Schematic illustration of the random-error-threshold-controlled averaging strategy.

Figure 13 shows the distribution of the spatial averaging range required to reduce the random error of averaged XCO_2 retrievals below 1 ppm under the random-error-threshold-controlled strategy. Overall, the required averaging range differs markedly among surface random-error categories. Low-random-error land surfaces require the shortest range, with approximately 74.13 % of the observations reaching the 1 ppm target within 50 km. For moderate-random-error land surfaces, the required range increases, and approximately 57.83 % of the observations reach the threshold within 67 km. In contrast, the distributions for relatively high-random-error land surfaces and high-random-error regions extend toward larger spatial ranges, with only 25.33 % and 19.74 % of the observations reaching the 1 ppm target within 77 and 100 km, respectively. This indicates that high-random-error or discontinuous observation regions generally require longer averaging windows to satisfy the error-control requirement.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026



435 **Figure 13. Statistical distribution of the spatial averaging range required under the random-error-threshold-controlled averaging strategy.**

These results show that the random-error-threshold-controlled strategy can improve the consistency of averaged XCO_2 random errors, but its averaging range is strongly scene dependent. Compared with the fixed-spatial-resolution strategy, this strategy is more effective in unifying data quality, but it may introduce stronger spatial smoothing in high-random-error regions and thus reduce the spatial representativeness of the XCO_2 product.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

4.1.3 Comparison of averaging strategies

Overall, the fixed-spatial-resolution strategy is more suitable as the default processing scheme for spaceborne IPDA lidar surface XCO_2 products. Based on the four surface random-error categories, this strategy uses approximately 160, 200, 232, and 296 observations as the default averaging numbers, corresponding to typical along-track scales of approximately 54, 67, 77, and 100 km, respectively. This provides a relatively stable balance between random-error reduction and along-track spatial representativeness, facilitating subsequent applications such as spatial distribution analysis, model comparison, regional statistics, and carbon emission inversion. Although the random-error-threshold-controlled strategy can improve the consistency of averaged random errors, its averaging window varies substantially with single-shot random errors and data gaps. In weak-signal or high-random-error regions, it may introduce stronger spatial smoothing and weaken the ability to characterize local XCO_2 enhancements, land–sea gradients, and urban plumes. Therefore, using a fixed random-error threshold as the sole averaging target is not necessarily more suitable for scientific applications; for applications requiring higher precision, the averaging range can be moderately increased based on the default averaging numbers.

4.2 Analysis of averaging strategies at boundaries between different surface types

In regions with rapid surface-type transitions, such as land–sea boundaries, an averaging window may contain observations from different surface random-error categories. If only observations from the same surface category are used, insufficient valid observations near the boundary may lead to large random errors after averaging. In contrast, expanding the search range for same-category observations to meet a random-error threshold may substantially increase the averaging-window length and reduce spatial representativeness. Therefore, a dedicated averaging strategy is needed for boundary regions to balance random-error control and spatial-window constraints.

The boundary-region analysis was performed using the 2023 spaceborne IPDA lidar XCO_2 observations. After quality control, the observations were first matched with the MCD12C1 land cover product and classified according to the four surface random-error categories defined above. Locations where the surface random-error category changed between adjacent along-track observations were then identified as boundary points. Observation segments within 50 s before and after each boundary point were extracted, corresponding to approximately 2000 observations in the absence of data gaps. To ensure relative completeness of the segments, only cases with more than 1800 valid observations within the window were retained, yielding 18 316 boundary cases.

For each boundary case, three averaging schemes were compared: same-category fixed-spatial-resolution averaging, same-category random-error-threshold-controlled averaging, and boundary mixed dynamic averaging. The first two schemes use only observations from the same surface random-error category as the current observation, corresponding to a fixed averaging number and an expanded window until the random error falls below 1 ppm, respectively. The boundary mixed dynamic scheme allows valid observations from different categories on both sides of the boundary to be averaged together,

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

adding observations in order of increasing distance from the boundary point until the averaged random error falls below 1 ppm or the available observation range is reached.

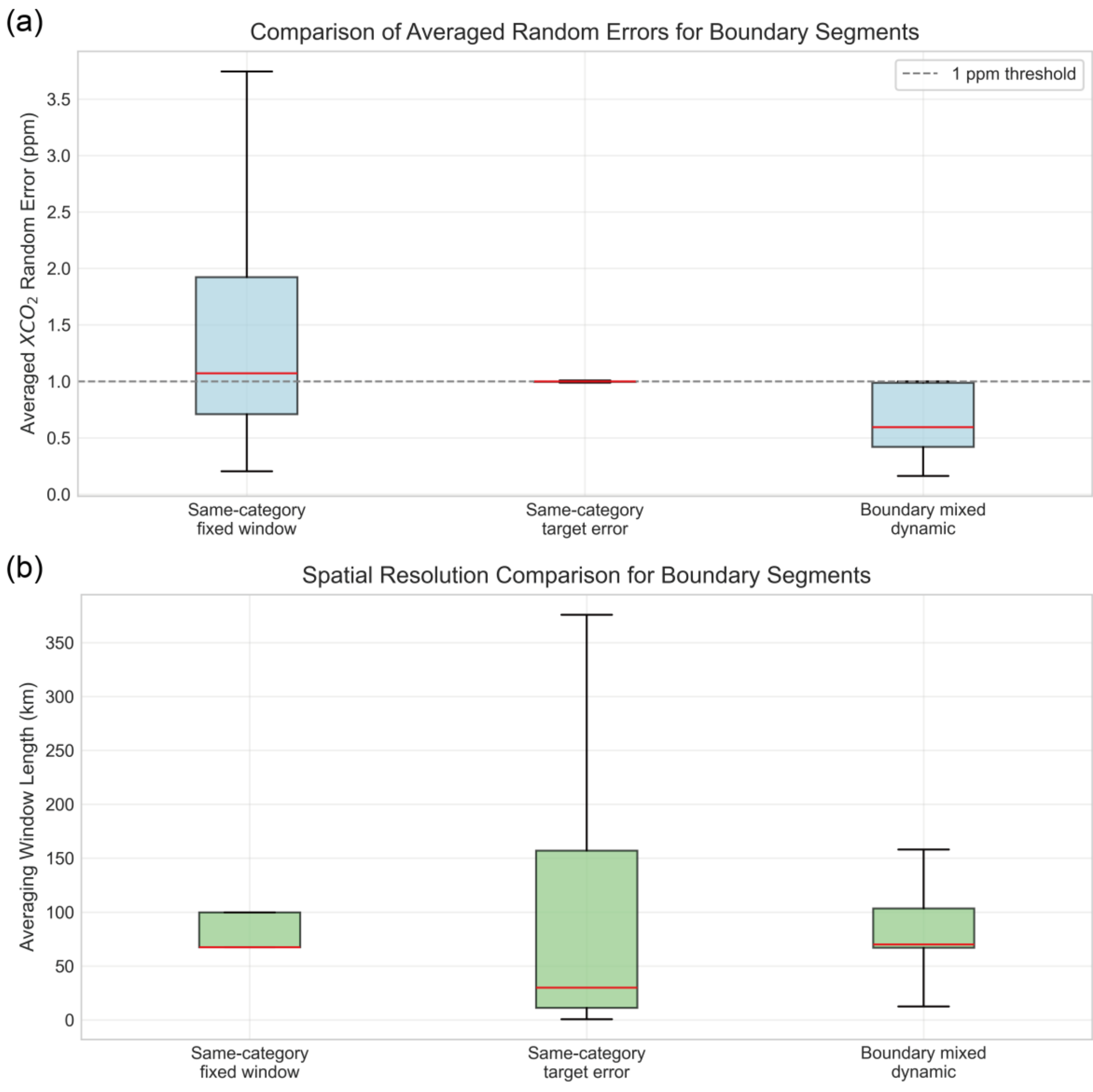


Figure 14. Comparison of averaging strategies for boundary segments between different surface random-error categories. (a) Random errors of averaged XCO_2 retrievals. The grey dashed line denotes the 1 ppm random-error threshold. (b) Averaging-window lengths for the three strategies. The three strategies are same-category fixed-spatial-resolution averaging, same-category random-error-threshold-controlled averaging, and boundary mixed dynamic averaging.

Figures 14 show the averaged random errors and averaging-window lengths for the three strategies, respectively. The same-category fixed-spatial-resolution strategy maintains a relatively stable spatial scale, but some cases show large random errors because of insufficient same-category observations near the boundary. The same-category random-error-threshold-controlled strategy can control the random error near 1 ppm, but some averaging windows extend beyond 300 km, indicating a

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

substantial loss of spatial representativeness. By contrast, the boundary mixed dynamic strategy generally keeps the random error below 1 ppm and produces a more concentrated distribution of window lengths. This indicates that allowing valid observations from both sides of the boundary to participate in averaging can improve error control without excessively expanding the spatial window. Therefore, the boundary mixed dynamic strategy is more suitable for land–sea boundaries and other rapidly changing surface-type regions. However, the fraction of each surface category within the averaging window should be provided as a quality indicator to avoid interpreting mixed-surface averages as observations representative of a single surface type.

5 Discussion

5.1 Comparison between single-shot XCO_2 random errors from spaceborne IPDA lidar and simulation results

Disney et al. (2009) proposed a method for estimating reflectance at 1570 nm using MODIS reflectance at 1640 and 2130 nm, which was later used in ACDL performance simulations to estimate the expected return energies at the online and offline wavelengths (Yang et al., 2019; Zhu et al., 2021). Because surface reflectance, return energy, and return-signal SNR are generally positively correlated, this study compares the simulated global 1572 nm return-energy distribution from Yang et al. (2019) and Zhu et al. (2021) with the 2023 single-shot XCO_2 random errors from real observations. The two datasets show generally consistent large-scale spatial patterns, indicating that the MODIS-based estimation of 1572 nm surface reflectance can reasonably capture the global spatial variability of spaceborne IPDA lidar observation performance. However, a clear discrepancy is found over permanent snow- and ice-covered regions. In the simulations, these regions show strong expected return energies because of their high estimated reflectance, whereas the observed random errors remain relatively large. This suggests that broadband MODIS reflectance may not accurately represent the true spectral reflectance of snow and ice near the laser operating wavelength. Complex polar snow and ice structures, thin clouds, and special atmospheric conditions may further degrade the effective return-signal quality.

The observed single-shot XCO_2 random errors over ocean were also compared with the sea-surface relative random errors simulated by Wang et al. (2020) based on sea-surface wind speed. Except for local differences caused by cloud contamination, quality screening, and observation coverage, the two results show generally consistent spatial patterns over ocean. This indicates that sea-surface-reflectance parameterization based on wind speed can reasonably represent regional differences in ocean return conditions for spaceborne IPDA lidar. In addition, the analysis by Wang et al. (2020) based on CALIPSO cloud and aerosol optical depth can capture observation limitations in regions with frequent cloud contamination, such as the Southern Ocean, but may overestimate the effect in regions with complex cloud and aerosol conditions, such as Southeast Asia. This suggests that cloud and aerosol conditions remain important sources of uncertainty in actual data availability and random-error simulations.

5.2 Sources of the difference between the Allan standard deviation and the theoretical random error

In the averaging of spaceborne IPDA lidar XCO_2 retrievals, the AVX scheme, which directly calculates the arithmetic mean of multiple single-shot XCO_2 retrievals, may be affected by spatial variations in the integrated weighting function (IWF) and thus introduce geophysical bias (Tellier et al., 2018). When the IWF differs among observations within an averaging window, equal-weight averaging may introduce bias, especially over complex terrain, land–sea boundaries, or regions with pronounced elevation variations. To evaluate this effect, the AVX-averaged results and the IWF-weighted averaged results were compared using the continuous observation segments used for the averaging-number analysis. The results show that the IWF-weighted correction does not significantly change the Allan standard deviation at different averaging scales. This indicates that, within the same-category continuous segments selected in this study, the geophysical bias caused by spatial variations in IWF is not pronounced. Therefore, the IWF-related bias in the AVX scheme is not the main reason why the observed Allan standard deviation is higher than the theoretical random error.

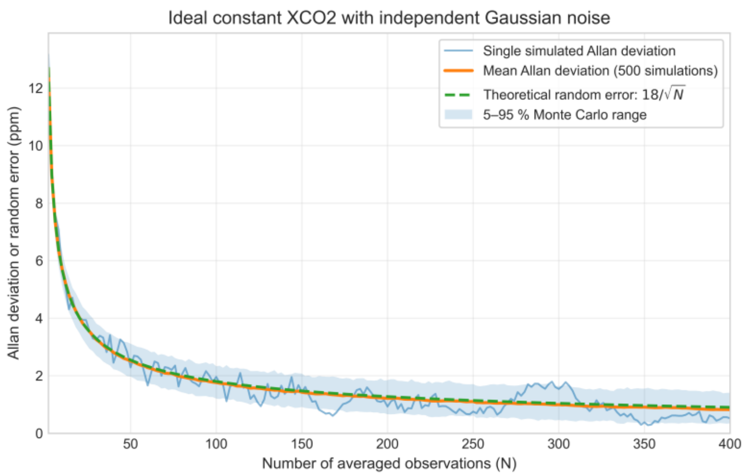


Figure 15. Comparison of Allan deviations derived from idealized XCO_2 simulations and the theoretical random-error curve. The blue line shows the Allan deviation from a single realization, the orange line represents the mean Allan deviation from 500 simulations, and the shaded region denotes the 5–95 % range among the simulations. The dashed line indicates the theoretical random error expected from independent-noise propagation.

To verify the consistency between the Allan variance analysis and the independent-noise propagation theory, an idealized simulation was conducted using a constant XCO_2 sequence of 420 ppm with independent Gaussian noise having a single-shot standard deviation of 18 ppm. As shown in Fig. 15, the Allan deviation obtained from repeated simulations closely follows the theoretical random-error curve. Although the Allan deviation from an individual realization fluctuates at large averaging numbers because of the limited sample size and the reduced number of available averaging windows, the mean curve from 500 simulations agrees well with the theoretical expectation. This result confirms that, under stationary conditions with independent random errors, the Allan variance method provides a consistent description of uncertainty reduction through averaging.

The difference between the Allan standard deviation derived from real observations and the theoretical random error is therefore more likely associated with real differences among single-shot XCO_2 observations and error components with finite correlation scales. In the real atmosphere, XCO_2 can vary along the satellite track. Such variability is included in the Allan variance as differences between adjacent averaged windows, thereby increasing the Allan standard deviation. In addition, correlated errors may arise from instrumental response characteristics, return-signal processing, and other components of the observation chain. Therefore, the observed Allan standard deviation should not be simply interpreted as the averaged result of single-shot random errors. Instead, its deviation from the theoretical random error reflects the combined influence of real spatial variability and non-independent error components. Further separation of these contributions would require high-resolution XCO2 fields, transport model simulations, or independent validation data.

6 Summary and outlook

Based on the 2023 DQ-1 spaceborne IPDA lidar observations, this study evaluated the random errors of single-shot XCO_2 retrievals and their behavior during along-track averaging, and further discussed averaging strategies for different surface random-error categories. The single-shot XCO_2 random errors were calculated from the return-signal SNRs at the online and offline wavelengths. The results show that the random errors of single-shot XCO_2 retrievals from the DQ-1 spaceborne IPDA lidar exhibit pronounced spatial heterogeneity, with clear differences among regions and surface cover types. Water bodies, permanent snow and ice, and some complex land cover types generally show larger random errors, whereas common land cover types usually show lower random errors. The day–night comparison further indicates that the day–night difference in random error is relatively weak.

For the averaging-number analysis, Allan variance analysis was first performed using a simple land–ocean classification. The results show different convergence behaviors between land and ocean retrievals as the averaging number increases. However, this binary classification is insufficient to describe error variations within land areas and over complex surface conditions such as permanent snow and ice and water bodies. Therefore, by combining the MCD12C1 land cover data with the statistics of single-shot random errors, the global observation scenes were further divided into four categories: low-random-error land surfaces, moderate-random-error land surfaces, relatively high-random-error land surfaces, and high-random-error regions. Based on the Allan variance analysis for these four categories, this study recommends averaging approximately 160, 200, 232, and 296 observations, corresponding to typical along-track averaging scales of approximately 54, 67, 77, and 100 km, respectively.

To address missing observations and along-track discontinuities caused by cloud screening, quality control, and abnormal-return filtering, two averaging strategies for homogeneous surface conditions were compared: the fixed-spatial-resolution strategy and the random-error-threshold-controlled strategy. The fixed-spatial-resolution strategy maintains relatively stable spatial representativeness and is therefore more suitable as the default averaging scheme for XCO_2 surface retrieval products. However, in regions with insufficient valid observations or large single-shot random errors, the random error after averaging

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

may still exceed the target threshold. The random-error-threshold-controlled strategy improves the consistency of averaged random errors, but it may substantially expand the averaging window under weak-signal or severely discontinuous conditions, thereby weakening the ability to characterize local XCO_2 variability. Therefore, this study suggests using the fixed-spatial-resolution strategy as the default averaging option, while providing the random error after averaging as a quality indicator.

For regions with rapid surface-type transitions, such as land–sea boundaries, a boundary mixed dynamic averaging strategy was further proposed. This strategy allows valid observations from different surface random-error categories on both sides of a boundary to be averaged together, thereby reducing the risk of large random errors or excessive window expansion caused by insufficient same-category observations. The results show that the boundary mixed dynamic strategy can reduce the random error after averaging without excessively increasing the spatial window. However, because the resulting averaged XCO_2 retrievals have mixed-surface attributes, the fractions of observations from different surface random-error categories within the averaging window should be provided as an important quality indicator for boundary-region products.

This study also discussed the differences between real observations and simulation results, as well as the possible reasons why the observed Allan standard deviation is higher than the theoretical random error. Compared with existing simulations that estimate 1572 nm surface reflectance from broadband MODIS reflectance, the observed XCO_2 random errors are generally consistent with the simulated return energy at large spatial scales, but clear discrepancies remain over permanent snow- and ice-covered regions. This suggests that the effective reflectance of snow and ice near the laser operating wavelength may not be accurately represented by broadband MODIS reflectance. In contrast, the spatial distribution of observed ocean random errors is generally consistent with simulations based on sea-surface wind speed parameterization, indicating that this method can reasonably represent regional differences in ocean return conditions. Analyses based on cloud and aerosol optical depth can help explain observation limitations in some regions, but uncertainties remain under complex cloud and aerosol conditions. In addition, the IWF-weighted correction does not significantly change the Allan standard deviation within the selected same-category continuous segments, indicating that geophysical bias caused by IWF spatial variations in the AVX scheme is not pronounced and is not the main reason why the observed Allan standard deviation is higher than the theoretical random error. Preliminary simulation tests suggest that this difference is more likely dominated by correlated errors and real along-track spatial variations in XCO_2, with the latter being included in the Allan variance as differences between adjacent averaged windows.

Further improvements are still needed in the evaluation of random errors and averaging methods for spaceborne IPDA lidar XCO_2 retrievals. First, future observation-performance simulations should pay more attention to biases in surface-reflectance parameterization, especially over special surface types such as permanent snow and ice. Second, the difference between the Allan standard deviation and the theoretical random error requires further investigation. High-resolution XCO_2 fields, independent validation data, and related auxiliary information could be used to separate the contributions of real spatial variability, correlated errors, and systematic errors. Finally, for specific scientific applications such as land–ocean carbon flux transport, point-source emission monitoring, and urban emission inversion, the influence of averaging strategies on local

XCO_2 enhancements, spatial gradients, and plume structures should be further evaluated to develop application-oriented averaging schemes.

Data availability

The DQ-1 spaceborne IPDA lidar observation data used in this study are expected to become available through the National Satellite Meteorological Center at https://satellite.nsmc.org.cn/DataPortal/cn/home/index.html. Before public release, the data can be made available for review and research purposes upon reasonable request from Lu Zhang (zhanglu_nsmc@cma.gov.cn). The MCD12C1 land cover product used in this study is publicly available from the NASA Earthdata platform.

Author contributions

ZF designed the study, developed the methodology, performed the data analysis, and prepared the manuscript. CF contributed to the interpretation of the results and manuscript revision. LZ provided the DQ-1 ACDL data, contributed to data interpretation, and supervised the study. LB supervised the study and contributed to the discussion and manuscript revision. JW contributed to data processing and validation. YC contributed to data analysis and visualization. JG contributed to data processing and figure preparation. QZ contributed to the interpretation of the simulation results and manuscript revision. JL contributed to scientific discussion, interpretation of the results, and manuscript revision. WC provided overall scientific guidance, supervised the study, and contributed to manuscript revision. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors acknowledge the National Satellite Meteorological Center of the China Meteorological Administration for providing the DQ-1 ACDL data and related technical support. The authors also thank the teams responsible for the MODIS MCD12C1 land cover product and other auxiliary datasets used in this study.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

Financial support

Supported by the Jing-Jin-Ji Regional Integrated Environmental Improvement-National Science and Technology Major Project of Ministry of Ecology and Environment of China（No. 【2025ZD1200900】).

References

Abshire, J. B., Ramanathan, A. K., Riris, H., Allan, G. R., Sun, X., Hasselbrack, W. E., Mao, J., Wu, S., Chen, J., Numata, K., Kawa, S. R., Yang, M. Y. M., and DiGangi, J.: Airborne measurements of CO_2 column concentrations made with a pulsed IPDA lidar using a multiple-wavelength-locked laser and HgCdTe APD detector, Atmos. Meas. Tech., 11, 2001–2025, https://doi.org/10.5194/amt-11-2001-2018, 2018.

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.

Amediek, A., Fix, A., Ehret, G., Caron, J., and Durand, Y.: Airborne lidar reflectance measurements at 1.57 µm in support of the A-SCOPE mission for atmospheric CO_2, Atmos. Meas. Tech., 2, 755–772, https://doi.org/10.5194/amt-2-755-2009, 2009.

Cao, X., Zhang, X., Zhang, L., Zhou, M., Jiang, Y., Liu, J., Cheng, T., Fan, C., Chen, C., and Chen, L.: Averaging scheme for the aerosol and carbon detection LiDAR onboard DaQi-1 satellite, IEEE Trans. Geosci. Remote Sens., 62, 1–16, https://doi.org/10.1109/TGRS.2024.3380639, 2024.

Disney, M. I., Lewis, P. E., Bouvet, M., Prieto-Blanco, A., and Hancock, S.: Quantifying surface reflectivity for spaceborne lidar via two independent methods, IEEE Trans. Geosci. Remote Sens., 47, 3262–3271, https://doi.org/10.1109/TGRS.2009.2019268, 2009.

Duren, R. M. and Miller, C. E.: Measuring the carbon emissions of megacities, Nat. Clim. Change, 2, 560–562, https://doi.org/10.1038/nclimate1629, 2012.

Ehret, G., Kiemle, C., Wirth, M., Amediek, A., Fix, A., and Houweling, S.: Space-borne remote sensing of CO_2, CH_4, and N_2O by integrated path differential absorption lidar: a sensitivity analysis, Appl. Phys. B, 90, 593–608, https://doi.org/10.1007/s00340-007-2892-3, 2008.

Fan, C., Chen, C., Liu, J., Xie, Y., Li, K., Zhu, X., Zhang, L., Cao, X., Han, G., Huang, Y., Gu, Q., and Chen, W.: Preliminary analysis of global column-averaged CO_2 concentration data from the spaceborne aerosol and carbon dioxide detection lidar onboard AEMS, Opt. Express, 32, 21870–21886, https://doi.org/10.1364/OE.517736, 2024.

Fan, C., Yang, J., Liu, J., Bu, L., Wang, Q., Wei, C., Zhang, Y., Zhu, X., Li, S., Zang, H., and Chen, W.: Airborne atmospheric carbon dioxide measurement using 1.5 µm laser double-pulse IPDA lidar over a desert area, Appl. Opt., 63, 2121–2131, https://doi.org/10.1364/AO.507905, 2024.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

Fang, J., Chen, B., Zhang, H., Dilawar, A., Guo, M., Liu, C., Liu, S., Gemechu, T. M., and Zhang, X.: Global evaluation and intercomparison of XCO_2 retrievals from GOSAT, OCO-2, and TANSAT with TCCON, Remote Sens., 15, 5073, https://doi.org/10.3390/rs15205073, 2023.

Fix, A., Matthey, R., Amediek, A., Ehret, G., Gruet, F., Kiemle, C., Klein, V., Mileti, G., Pereira do Carmo, J., and Quatrevalet, M.: Investigations on frequency and energy references for a space-borne integrated path differential absorption lidar, in: International Conference on Space Optics – ICSO 2014, Tenerife, Canary Islands, Spain, 2014, Proc. SPIE, 10563, 105630A, https://doi.org/10.1117/12.2304145, 2017.

IPCC: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Lee, H. and Romero, J., Intergovernmental Panel on Climate Change, Geneva, Switzerland, 184 pp., https://doi.org/10.59327/IPCC/AR6-9789291691647, 2023.

Kawa, S. R., Abshire, J. B., Baker, D. F., Browell, E. V., Crisp, D., Crowell, S. M. R., Hyon, J. J., Jacob, J. C., Jucks, K. W., Lin, B., Menzies, R. T., Ott, L. E., and Zaccheo, T. S.: Active Sensing of CO_2 Emissions over Nights, Days, and Seasons (ASCENDS): Final Report of the ASCENDS Ad Hoc Science Definition Team, NASA Tech. Publ. NASA/TP-2018-219034, 228 pp., 2018.

Kulawik, S., Wunch, D., O'Dell, C., Frankenberg, C., Reuter, M., Oda, T., Chevallier, F., Sherlock, V., Buchwitz, M., Osterman, G., Miller, C. E., Wennberg, P. O., Griffith, D., Morino, I., Dubey, M. K., Deutscher, N. M., Notholt, J., Hase, F., Warneke, T., Sussmann, R., Robinson, J., Strong, K., Schneider, M., De Mazière, M., Shiomi, K., Feist, D. G., Iraci, L. T., and Wolf, J.: Consistent evaluation of ACOS-GOSAT, BESD-SCIAMACHY, CarbonTracker, and MACC through comparisons to TCCON, Atmos. Meas. Tech., 9, 683–709, https://doi.org/10.5194/amt-9-683-2016, 2016.

Liu, Y., Wang, J., Yao, L., Chen, X., Cai, Z., Yang, D., Yin, Z., Gu, S., Tian, L., Lu, N., and Lyu, D.: The TanSat mission: preliminary global observations, Sci. Bull., 63, 1200–1207, https://doi.org/10.1016/j.scib.2018.08.004, 2018.

Nisbet, E. and Weiss, R.: Top-down versus bottom-up, Science, 328, 1241–1243, https://doi.org/10.1126/science.1189936, 2010.

Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Bovensmann, H., Burrows, J. P., Di Noia, A., Parker, R. J., Somkuti, P., Landgraf, J., Hasekamp, O. P., and Aben, I.: Retrieval of greenhouse gases from GOSAT and greenhouse gases and carbon monoxide from GOSAT-2 using the FOCAL algorithm, Atmos. Meas. Tech., 15, 3401–3437, https://doi.org/10.5194/amt-15-3401-2022, 2022.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Baker, I., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018.

https://doi.org/10.5194/egusphere-2026-3759
Preprint. Discussion started: 14 July 2026

Taylor, T. E., O'Dell, C. W., Baker, D., Bruegge, C., Chang, A., Chapsky, L., Chatterjee, A., Cheng, C., Chevallier, F., Crisp, D., Dang, L., Drouin, B., Eldering, A., Feng, L., Fisher, B., Fu, D., Gunson, M., Haemmerle, V., Keller, G. R., Kiel, M., Kuai, L., Kurosu, T., Lambert, A., Laughner, J., Lee, R., Liu, J., Mandrake, L., Marchetti, Y., McGarragh, G., Merrelli, A., Nelson, R. R., Osterman, G., Oyafuso, F., Palmer, P. I., Payne, V. H., Rosenberg, R., Somkuti, P., Spiers, G., To, C., Weir, B., Wennberg, P. O., Yu, S., and Zong, J.: Evaluating the consistency between OCO-2 and OCO-3 XCO_2 estimates derived from the NASA ACOS version 10 retrieval algorithm, Atmos. Meas. Tech., 16, 3173–3209, https://doi.org/10.5194/amt-16-3173-2023, 2023.

Tellier, Y., Pierangelo, C., Wirth, M., Gibert, F., and Marnas, F.: Averaging bias correction for the future space-borne methane IPDA lidar mission MERLIN, Atmos. Meas. Tech., 11, 5865–5884, https://doi.org/10.5194/amt-11-5865-2018, 2018.

UNFCCC: Adoption of the Paris Agreement, in: Report of the Conference of the Parties on its twenty-first session, Addendum, United Nations Framework Convention on Climate Change, Paris, France, 2 pp., 2015.

United Nations: Kyoto Protocol to the United Nations Framework Convention on Climate Change, United Nations, 1998.

Wagner, G. A. and Plusquellic, D. F.: Ground-based, integrated path differential absorption LIDAR measurement of CO_2, CH_4, and H_2O near 1.6 µm, Appl. Opt., 55, 6292–6310, https://doi.org/10.1364/AO.55.006292, 2016.

Wang, S., Ke, J., Chen, S., Zheng, Z., Cheng, C., Tong, B., Liu, J., Liu, D., and Chen, W.: Performance evaluation of spaceborne integrated path differential absorption lidar for carbon dioxide detection at 1572 nm, Remote Sens., 12, 2570, https://doi.org/10.3390/rs12162570, 2020.

World Meteorological Organization: WMO Greenhouse Gas Bulletin No. 19: The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2022, World Meteorological Organization, Geneva, Switzerland, 11 pp., https://library.wmo.int/idurl/4/68532, 2023.

Yang, J., Zhu, Y., Wang, Q., Bu, L., Liu, J., and Chen, W.: Influence of surface reflectance and aerosol optical depth on performance of spaceborne integral path differential absorption lidar, Chinese J. Lasers, 46, 0910001, https://doi.org/10.3788/CJL201946.0910001, 2019.

Zhu, Y., Yang, J., Zhang, X., Liu, J., Zhu, X., Zang, H., Xia, T., Fan, C., Chen, X., Sun, Y., Hou, X., and Chen, W.: Performance improvement of spaceborne carbon dioxide detection IPDA LIDAR using linearty optimized amplifier of photo-detector, Remote Sens., 13, 2007, https://doi.org/10.3390/rs13102007, 2021.