Dear Editors and Reviewers,

We would like to thank the editor for giving us an extension to resubmit the paper, and also thank the reviewers for the constructive suggestions which help us to improve the quality of the paper.

Here we submit a new version of the manuscript with the title "UAV-based method for measuring CO_2 emissions in forest ecosystems", modified from the original version according to the reviewers' suggestions. We have addressed the comments carefully as detailed below. The original comments are in black italic and our replies in black normal font. The revised paragraphs are also shown in blue after each reply to show the changes.

Sincerely yours,

Shao-Meng Li, Professor

The following is a point-to-point response to the two reviewers' comments.

Reviewer #1:

General comments

(1) *Unfortunately, the abstract lacks clarity what sort of data is collected and what is happening to it. Please can it be explicitly stated that this work is dedicated to flux calculation? What sort of sensor is used? Perhaps this is not an important detail but then what is? At least some details need to be provided in this abstract to appreciate the novelty of this work. Otherwise, the abstract simply tells us that a UAV was used to measure forest emissions in this work, which is vague.*

Response: We agree that the original abstract did not describe clearly enough what data were collected, what sensors were used, and how these data were used in the analysis. We have therefore revised the abstract to make the scope and novelty of the study more explicit.

Specifically, the revised abstract now clarifies that this work is dedicated to three-dimensional CO_2 emission calculation, rather than general UAV-based detection of forest emissions. We now explicitly state that the UAV platform measures CO_2 mixing ratios using a closed-path CO_2 analyzer and wind vectors using a calibrated ultrasonic anemometer, together with temperature and pressure information. We also clarify how these observations are used: Within a mass-balance framework, the UAV measurement system is flown along horizontal loops at multiple heights, together with vertical ascent flight tracks, acquiring data along the flight tracks that are used to simultaneously quantify horizontal advective and vertical turbulent carbon dioxide emission components. We appreciate this comment, which helped us improve the clarity and specificity of the abstract. The revised sentence is as follows:

"This study addresses the difficulty of accurately quantifying CO_2 emissions in forest ecosystems due to spatial heterogeneity, complex terrain, and the combined effects of horizontal and vertical transport. A UAV-based flux quantification methodology, capable of three-dimensional CO_2 emission computation, was developed, evaluated and applied to a forest ecosystem. The UAV system is integrated with a high-precision closed-path CO_2 analyzer to measure CO_2 mixing ratios and a calibrated ultrasonic anemometer to measure wind vectors, together with profile measurements of temperature and pressure. Within a mass-balance framework, the UAV measurement system is flown along horizontal loops at multiple heights together with vertical ascent flight tracks, acquiring data along the flight tracks that are used to simultaneously quantify horizontal advective and vertical turbulent carbon dioxide emission components. The method was applied at a subtropical plantation forest and evaluated against a collocated eddy-covariance (EC) flux tower. Validation against a long-term flux tower shows that vertical CO_2 fluxes derived using the gradient method agree well with EC observations across seasons, with a correlation coefficient r of 0.9. Horizontal measurements further reveal pronounced local canopy air diurnal CO_2 mass flux variation: lateral advection dominates in the early morning, whereas vertical uptake prevails under well-mixed midday conditions. Sensitivity analysis of emission intensities across multi-scale grid boxes ($50\times50\times50$ m^3, $100\times100\times50$ m^3, and $150\times150\times50$ m^3) revealed stronger scale dependence during autumn mornings, indicating significant spatial heterogeneity in CO_2 emissions. Uncertainty assessment suggests that the total relative uncertainty of UAV-derived emissions is typically ~22%, with wind-field measurement as the main error source. Overall, this UAV approach provides a flexible and reliable method for analyzing near-surface three-dimensional CO_2 transport over complex and heterogeneous forests, helping to overcome the limited spatial coverage of traditional flux towers and remote sensing."

The specific revisions are shown in lines 29 to 54 of the revised manuscript.

(2) *The introduction lacks citations to more recent work. Many citations are old. This is fine, but they should be used in addition to more recent work.*

Response: We agree that, although several older references are important because they provide the foundational theory and methodological background for eddy covariance, boundary-layer transport, and mass-balance approaches, the Introduction should also cite more recent studies to better reflect the current state of the field.

In the revised manuscript, we have therefore updated the Introduction by adding recent literature alongside the classical references. In particular, we strengthened the discussion of:

(1) recent progress in UAV-based greenhouse gas and flux observations:

"Accurate quantification of forest-atmosphere carbon exchange remains essential for constraining regional and global carbon budgets and for evaluating ecosystem feedbacks from climate change. Recent synthesis studies and observation-network analyses continue to highlight the importance of improving carbon flux measurements across heterogeneous ecosystems (Fang et al., 2024; Burman et al., 2025)."

The specific revisions are shown in lines 65 to 70 of the revised manuscript.

(2) Regarding the introduction to EC:

"Eddy covariance (EC) flux measurement technique is currently the most widely used method (Aubinet et al., 2012; Baldocchi, 2003; Burman et al., 2025; Pastorello et al., 2020; Fang et al., 2024)"

The specific revisions are shown in lines 83 to 84 of the revised manuscript.

(3) Newer studies on eddy-covariance datasets and flux-network applications:

for quantifying ecosystem-scale vertical CO_2 fluxes and has been implemented globally through networks such as the FLUXNET (Aubinet et al., 2012; Baldocchi et al., 2001; Fang et al., 2024; Pastorello et al., 2020; Wilson et al., 2002).

The specific revisions are shown in line 86 of the revised manuscript.

(4) Recent work addressing CO_2 flux measurements over complex and heterogeneous ecosystems

"Collectively, these limitations highlight a persistent observational gap in the characterization of CO_2 three-dimensional transport in forest ecosystems (Fang et al., 2024; Bolek et al., 2024)."

The specific revisions are shown in line 112 of the revised manuscript.

(5) Recent advances in rotating wing unmanned aerial vehicle:

"Recent advances in rotating wing unmanned aerial vehicle (UAV) platforms and lightweight sensor technologies have created new opportunities for atmospheric measurements over complex environments (Bolek et al., 2024)."

The specific revisions are shown in line 115 of the revised manuscript.

(6) UAV wind speed correction:

"To address these challenges, various sensors, including multi-hole probes, pitot tubes, and sonic anemometers, have been deployed on UAVs for wind measurements and correction algorithms have been developed to mitigate measurement disturbances associated with rotor downwash, platform-induced airflow, and rapid changes in UAV motion and attitude (Langelaan et al., 2011; Niedzielski et al., 2017; Nolan et al., 2018; Rogers and Finn, 2013; Soddell et al., 2004; Soltaninezhad et al., 2025; Spiess et al., 2007; Yang et al., 2024)"

The specific revisions are shown in lines 131 to 135 of the revised manuscript.

At the same time, we retained the earlier references where they are foundational and still directly relevant. We believe this revision improves both the currency and the scholarly balance of the Introduction.

(3) *The manuscript conflates measurement platforms and flux quantification methods. This occurs throughout the manuscript and particularly in the introduction. The authors should make a clear distinction between platforms (such as stationary towers, aircraft, satellites and UAV) and flux quantification methods (such as the Eddy covariance method and mass balance box modelling). This is particularly problematic when referring to Eddy covariance towers. It is true that many of these towers are sold to process data using the Eddy covariance technique. However, any tower with high frequency concentration and vertical wind sampling can be used for Eddy covariance fluxes. Similarly, measurements from a system "sold" as an Eddy covariance tower do not have to be used for Eddy covariance flux calculation; the data can be used in other ways.*

Response: We agree that the previous version of the manuscript did not always make a sufficiently clear distinction between measurement platforms and flux quantification methods, especially in the Introduction. We also agree that this distinction is important for scientific clarity.

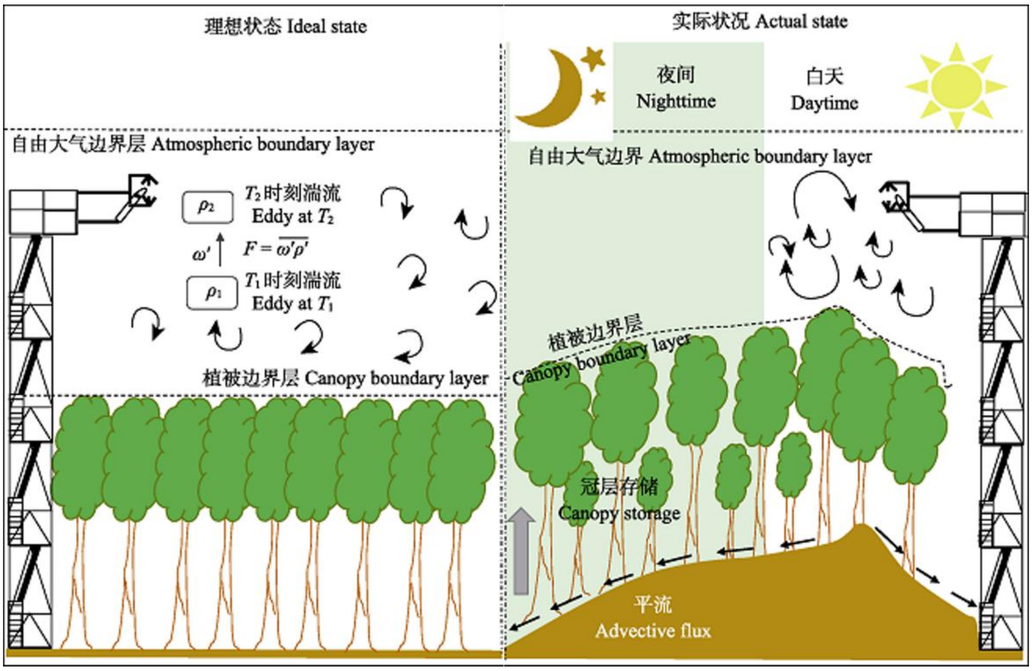
In the revised manuscript, we have carefully revised the relevant text throughout the paper to distinguish between platforms—such as stationary towers, aircraft, satellites, and UAVs—and methods used to derive fluxes from observations, such as the eddy covariance (EC) method, the mass-balance approach, and the gradient method. In particular, we have avoided wording that could imply that a tower itself is equivalent to the EC method. Instead, we now describe EC tower as flux tower.

We have made corresponding revisions especially in the Introduction, where the previous wording may have conflated "EC towers" with the EC method itself. The text now clarifies that the EC method is a flux calculation technique, whereas a tower is a flux tower. We thank the reviewer for highlighting this issue, which has helped us improve the conceptual precision of the manuscript.

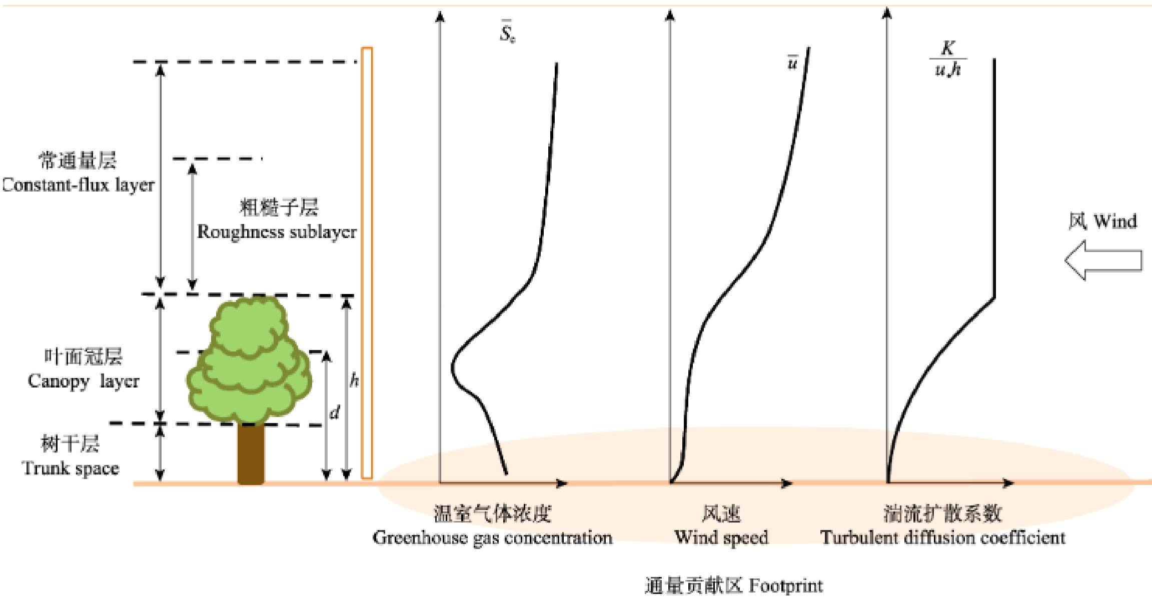
(4) *The flights are conducted at a minimum height of 35 m above ground level. Yet the canopy height is 18 m above ground level. Therefore, any mass-balance box modelling in the horizontal direction is fundamentally flawed as the source is not included within the box and nowhere near the top of the tree canopy. Any horizontal component to the overall flux will miss most of the emission or absorption. If the vertical wind speed is negligible (as stated by the authors), then it will take a long time for activity below the forest canopy to influence air at a higher elevation. Perhaps the horizontal box sampling is not a very useful component of this work.*

Response: Regarding the reviewer's concern that the lowest altitude of the UAV box-model measurements was 35 m while the canopy height was only 18 m, which may result in missing a large fraction of the emissions or uptake, we would like to provide further explanation from both technical and observational perspectives.

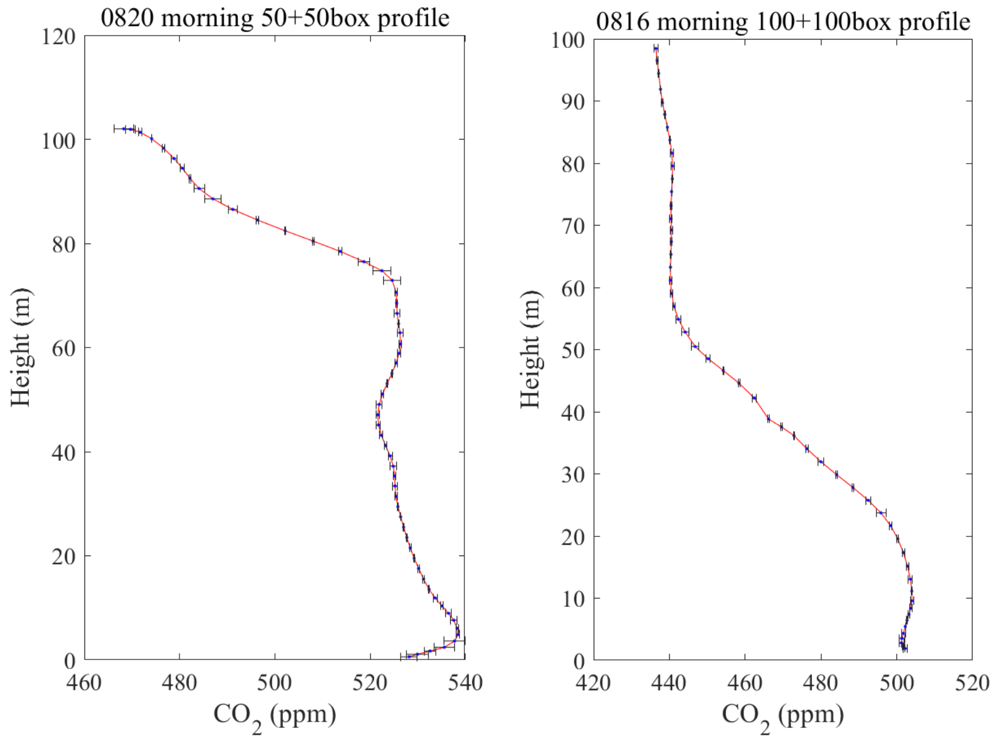
First, from a technical perspective, our measurements were not conducted over ideal flat terrain (see the schematic on the left). The altitude used in this study was already the lowest that could be safely achieved by the UAV, because terrain relief also had to be taken into account during the measurements (see the schematic on the right). Although we did not carry out a detailed topographic survey, a rough estimate based on the 15 m elevation difference between the UAV take-off location and the flux-tower site suggests that the terrain relief within the flight area was at least about 10 m. We had previously attempted flights at 30 m, but the UAV rotor blades struck tree branches and were damaged. The EC sensors on the Qianyanzhou flux tower are mounted at a height of 31 m, and we therefore consider our minimum flight altitude of 35 m to be the closest practical altitude to the tower measurement height that could be safely achieved in this forest environment.



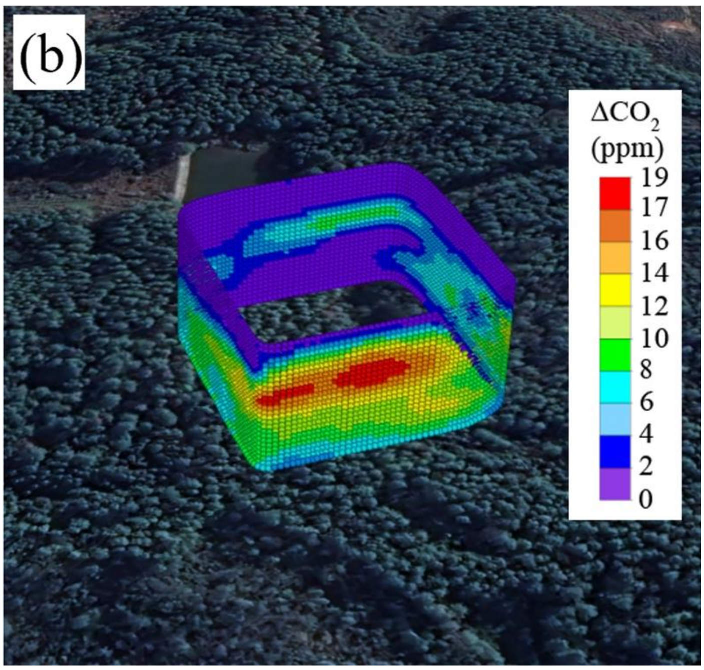
Second, from an observational perspective, the reviewer suggests that the box is located too far above the canopy top. However, in practice, we observed CO_2 variations extending well above the traditional canopy height. Under ideal conditions, the variation in CO_2 concentration above the canopy would be expected to be relatively small (as illustrated in the figure below).



However, in practice, the CO_2 concentration profiles measured in the profile pattern, as shown in the figure below, indicate that the CO_2 concentration changes between 18 and 35 m were approximately 4 ppm and 12 ppm in the two examples, respectively, whereas the corresponding changes within the 35–85 m range examined in this study were approximately 50 ppm and 40 ppm, respectively. This indicates that the minimum flight height of 35 m still captured the majority of the vertical emission component.



In addition, as shown in the figure below, the CO_2 concentrations measured in the box-pattern flights indicate that the high-concentration zone occurred above 50 m, with a concentration difference of 19 ppm. This suggests that our measurement approach is capable of capturing not only the vertical emission or uptake component, but also the horizontal emission or uptake component.



Finally, we would like to further emphasize that our measurements were conducted over heterogeneous underlying surfaces in complex terrain. In such complex ecosystems, local circulation and horizontal transport may occur, leading to spatial differences in CO_2 emissions. This can be seen directly from the CO_2 concentration distributions derived from both the profile-pattern and box-pattern measurements, which also indirectly demonstrates the advantages of the UAV-based measurement approach established in this study.

(5) *The equations are unfortunately incredibly difficult to follow. Many terms are not defined and they are written incorrectly, using inconsistent terminology. The authors must revise all of the equations and ensure that each term is correctly defined in the main text. The style of identical terms must remain consistent. For example, if a lower-case "u" represents wind speed, it cannot be changed into an upper-case "U" in a different equation, unless it has been defined as a different term in the main text. There are countless examples like this. It is really important that the authors correct this and make everything consistent as the typesetters will not necessarily know that "U" and "u" mean the same thing.*

Response: Thank you for this important comment. We agree that, in the previous version of the manuscript, the equations in Sect. 3.2 were not sufficiently clear and that inconsistent notation made the derivation difficult to follow. In the revised manuscript, we have carefully revised Eqs. (3.1)–(3.17) and the corresponding text throughout Sect. 3.2 to improve clarity, consistency, and completeness of symbol definitions. The revised sentence is as follows:

"According to the divergence theorem, the rate of change of mass of a stable compound within a control volume equals the surface integral of the mass flux through its enclosing surfaces. Consequently, the total CO_2 emission rate within the volume equals the net outward flux integrated across all box wall surfaces, where positive flux contributions represent CO_2 mass leaving the box and negative flux contributions represent CO_2 mass entering the box. Based on the integral form of the continuity equation, the emission rate within the control volume enclosed by the box pattern can be expressed as

$$E_C = E_{C,H} + E_{C,HT} + E_{C,V} + E_{C,VT} - E_{C,M} \tag{3.1}$$

where E_C denotes the total CO_2 emission rate within the control volume, $E_{C,H}$ and $E_{C,V}$ denote the integrals of horizontal and vertical advective fluxes through the box walls, respectively; $E_{C,HT}$ and $E_{C,VT}$ represent the integrated turbulent fluxes in the horizontal and vertical directions, and $E_{C,M}$ accounts for changes in CO_2 mass associated with air-density variations. Eq. (3.1) forms the basis of the TERRA emission inversion model (Gordon et al., 2015).

To investigate the contributions of each component to the emissions, the TERRA model was updated here, in particular the treatment of the vertical flux algorithm, to directly compute vertical turbulent fluxes based on the UAV measurements instead of relying on parametrization assuming entrainment. Here, the mean wind vectors ($\bar{u}$, $\bar{v}$, $\bar{w}$) measured with the UAV sonic anemometer over a 30-minute period and the corresponding wind fluctuations (u', v', w'), obtained by subtracting the 30-minute mean wind components from the instantaneous wind measurements, are used to determine fluxes in the three directions. For the horizontal terms, the mean horizontal wind speed $\bar{u}$ can be regarded as the main contributor to the $E_{C,H}$ term, typically 2-3 m/s. The horizontal turbulent velocity (u') is the main contributor to $E_{C,HT}$. Using the 95th percentile of u', the turbulent velocity is less than 0.1 m/s. The horizontal advection term ($E_{C,H}$) is larger than the turbulence term ($E_{C,HT}$) by more than a factor of 20, so that the $E_{C,HT}$ term can be neglected from Eq. (3.1). For the vertical flux terms, the reverse is true; the mean vertical wind speed $\bar{w}$, at typically <0.02 m/s, makes $E_{C,V}$ a minute contribution compared to the turbulent flow contribution $E_{C,VT}$, which depends on the vertical turbulent velocity (w'). Using the 95th percentile of w', the turbulent velocity can reach 0.3 m/s. As such, vertical advection term can also be neglected from Eq. (3.1).

The density-related term can be expressed as

$$E_{C,M} = M_R \iiint \chi_C \frac{d\rho_{air}}{dt} dxdydz \tag{3.2}$$

Because the maximum vertical (≈50 m), horizontal (≈150 m) extents, and duration of the UAV flights are limited, air-density variation over time $\frac{d\rho_{air}}{dt}$ within the control volume can be assumed to be negligible, and $E_{C,M}$ was therefore neglected. Accurate estimation of the remaining flux terms ($E_{C,H}$ and $E_{C,VT}$) enables computation of the net CO_2 emission rate. The reliability of the box-pattern approach for resolving CO_2 emission is further supported by Han et al. (2024), who conducted a series of controlled UAV-based CO_2 validation experiments over an industrial coking facility.

3.2.1 Computation of CO_2 horizontal emission component

The horizontal emission component was derived from box-pattern measurements by integrating the advective CO_2 flux across the four vertical walls of the virtual box. The horizontal CO_2 emission rate can be written as

$$E_{C,H} = M_R \oiint \chi_C(s,z)\rho_{air}(s,z)u_n(s,z)dsdz \tag{3.3}$$

where $E_{C,H}$ is the horizontal CO_2 emission, M_R is the ratio of the molar mass of CO_2 to that of air, $\chi_C(s,z)$ is the CO_2 mole fraction, $\rho_{\mathrm{air}}(s,z)$ is the air density, and $u_n(s,z)$ is the wind speed component normal to the flight path at point (s,z). Here, z denotes the flight altitude, s represents the path position along the flight trajectory starting from a predefined fix position on the trajectory path, and (s,z) indicates the point located at the path position s at flight altitude z.

For each altitude level z_i, the line flux $F(z_i)$ was first calculated from the box-pattern data as

$$F(z_i) = M_R \sum_{j=1}^{N_i} \chi_C(s_j,z_i) \cdot \rho_{air}(s_j,z_i) \cdot u_n(s_j,z_i) \cdot \Delta s_j \tag{3.4}$$

where j denotes the sampling point index along the flight path at altitude z_i, N_i is the total number of sampling points at altitude z_i, and Δs_j is the flight distance represented by the j-th sample interval.

To suppress small-scale variability, a moving average was applied to consecutive line-flux values, yielding layer-averaged fluxes for discrete altitude intervals. These values were then vertically integrated to obtain the total horizontal CO_2 emission rate

$$E_{C,H} = \sum_{i=1}^{n} F(z_i)\Delta z_i = M_R \sum_{i=1}^{n} \left(\sum_{j=1}^{N_i} \chi_C\,(s_j, z_i)\rho_{\text{air}}(s_j, z_i)u_n(s_j, z_i)\Delta s_j \right) \Delta z_i \tag{3.5}$$

where i denotes the vertical layer index from the first layer to the n-th layer, n is the total number of vertical layers, and Δz_i is the thickness of the i-th altitude interval. In this study, the first observed altitude layer z_1 corresponds to the lowest UAV sampling height (35 m above ground level).

3.2.2 Computation of CO_2 vertical emission component

The vertical CO_2 emission component was estimated from profile measurements using the gradient method, which relates turbulent scalar fluxes to vertical gradients of the scalar concentration. In practice, vertical profiling over forested terrain is operationally challenging, and UAV measurements were conducted at a single representative location. The resulting vertical fluxes were assumed to be representative of the broader domain, an assumption evaluated through direct comparison with eddy covariance (EC) flux tower measurements (see Section 4.2.1).

The vertical CO_2 emission component was obtained from integrating the vertical flux over the horizontal area enclosed by the box pattern and can be approximated by the product of vertical flux and the horizontal area, expressed as

$$E_{C,VT} = \iint F_v dxdy = A \cdot F_v \tag{3.6}$$

where F_v denotes the vertical flux and A is the horizontal area of the control volume.

According to the flux-gradient relationship (Kaimal and Finnigan, 1994), turbulent scalar fluxes are proportional to the vertical gradient of the scalar quantity:

$$F_v = -\rho_{air} K_C \frac{\partial \bar{c}}{\partial z} \tag{3.7}$$

where K_C denotes turbulent exchange coefficient, $\bar{c}$ was obtained by converting the measured CO_2 mixing ratio to concentration using the concurrently measured temperature and pressure during the UAV profile flight.

A mixing-length formulation was adopted, in which the eddy diffusivity is expressed as the product of a mixing length and a velocity scale represented here by the friction velocity (Lee and Mahrt, 2005)

$$K_M = l u_* \tag{3.8}$$

where l is the mixing length and K_M is the exchange coefficient for momentum. Under neutral stratification conditions, K_C can be assumed to be equal to K_M, and the friction velocity is defined as

$$u_* = l \frac{\partial \bar{u}}{\partial z} = \kappa z \frac{\partial \bar{u}}{\partial z} \tag{3.9}$$

where κ is the von Karman constant (κ=0.4), $\frac{\partial \bar{u}}{\partial z}$ is the vertical gradient of the mean wind speed.

To account for atmospheric stratification, a stability correction (Budyko et al.,

1962) was introduced as

$$l = m\kappa z \tag{3.10}$$

where m is the stratification correction parameter, determined as a function of the gradient Richardson number (R_i) according to Budyko's study

$$m = (1 - R_i)^{1/2} \tag{3.11}$$

Under neutral stratification conditions ($R_i = 0$), m = 1. For computational convenience, the gradient Richardson number R_i is approximated by the bulk Richardson number R_B in finite-difference form

$$R_i \approx R_B = \frac{\frac{g\Delta\bar{\theta}}{\bar{\theta}\Delta z}}{\left(\frac{\Delta\bar{u}}{\Delta z}\right)^2} = \frac{g}{\overline{\theta_1}+\overline{\theta_2}} \frac{(\overline{\theta_2}-\overline{\theta_1})(z_2-z_1)}{(\overline{u_2}-\overline{u_1})^2} \tag{3.12}$$

where g is the gravitational acceleration constant; $\Delta\bar{\theta}$ and $\Delta\bar{u}$ represent the changes in the temperature and wind speed between the layers (Δz), after smoothing with a seven-point moving window. In applying Eq. (3.12), the smoothed $\bar{\theta}$ and $\bar{u}$ vertical profiles are directly used to compute R_B, taking respective data points from two adjacent heights, z_1 and z_2, using the smoothed temperature and wind speed profiles at the two heights for $\overline{u_1}$, $\overline{u_2}$, $\overline{\theta_1}$, and $\overline{\theta_2}$. When applying the gradient method to estimate the flux, the vertical separation, $\Delta z = z_2 - z_1$, is typically chosen within a range of 1-10 m (Meredith et al., 2014). In this study, $\overline{u_1}$, $\overline{u_2}$ and $\overline{\theta_1}$, $\overline{\theta_2}$ were taken 5 m below and 5 m above any target height ($\Delta z = 10$ m), respectively, during the UAV ascent in the vertical profile measurements.

Because forest canopies represent aerodynamically rough and structurally complex surfaces, wind profiles above the canopy may deviate from the ideal logarithmic form. Following the neutral logarithmic wind-profile formulation from Monin–Obukhov similarity theory (Raupach, 1994), the surface-layer mean wind profile ($\bar{u}$) was expressed as

$$\bar{u} = \frac{u_*}{\kappa} ln\frac{z-d}{z_0} \tag{3.13}$$

where d is the zero-plane displacement height, which for this type of forest is typically taken as 70%-80% of the vegetation height, and z_0 is the surface roughness length.

Assuming neutral atmospheric stratification and treating u_*, d, and z_0 as constants over the same time period, u_* can be derived from wind measurements at two heights thereby eliminating z_0 from Eq. (3.13) and rearranging the equation so that

$$u_* = \frac{\kappa(\overline{u_2}-\overline{u_1})}{\ln[(z_2-d)/(z_1-d)]} \tag{3.14}$$

Substituting Eq. (3.10), Eq. (3.11) and Eq. (3.14) into Eq. (3.8) yields

$$K_c = \kappa^2(1 - R_i)^{1/2} \frac{(\overline{u_2}-\overline{u_1})}{\ln[(z_2-d)/(z_1-d)]}(z - d) \tag{3.15}$$

Further substituting Eq. (3.15) into Eq. (3.7) and integrating the latter result in the vertical flux

$$F_v = \rho_{air}\kappa^2(1 - R_i)^{1/2} \frac{(\overline{u_2}-\overline{u_1})(\overline{c_1}-\overline{c_2})}{\{\ln[(z_2-d)/(z_1-d)]\}^2} \tag{3.16}$$

where $\overline{c_1}$ and $\overline{c_2}$ represent the mean concentrations of CO_2 at heights z_1 and z_2, respectively, converted from measured CO_2 mixing ratios at using the measured temperature and pressure at both heights. In this study, z_1 and z_2 correspond to 5 m below and 5 m above the target height, respectively, during the UAV ascent.

The vertical CO_2 emission ($E_{C,VT}$) is finally expressed as

$$E_{C,VT} = A\rho_{air}\kappa^2(1 - R_i)^{1/2}\frac{(\overline{u_2}-\overline{u_1})(\overline{c_1}-\overline{c_2})}{\{\ln[(z_2-d)/(z_1-d)]\}^2} \tag{3.17}$$

In summary, horizontal CO_2 emission components were derived from box-pattern measurements using the divergence theorem, whereas vertical emission components were obtained from profile-pattern measurements using the gradient method. Finally, the three-dimensional net CO_2 emission was obtained using the original observations. The reliability of the vertical flux estimates, computed using Eq (3.16), was evaluated through direct comparisons with results from flux tower observations (see Section 4.2.1)."

The specific revisions are shown in lines 333 to 482 of the revised manuscript.

(6) *Section 3.2.2 provides extensive details on how the vertical turbulent flux component is derived with many complex steps. However, there are absolutely no citations in this subsection to justify why such steps are taken. For example, m is defined as the stratification correction parameter which can be calculated using equation 3.11. But where does this equation come from? Did the authors derive these equations themselves from first principles?*

Response: We agree that the previous version of Sect. 3.2.2 did not provide sufficient citations to support the derivation steps used in the vertical turbulent flux calculation. In particular, the theoretical basis of the gradient-flux formulation (Kaimal and Finnigan, 1994), the mixing-length treatment (Lee and Mahrt, 2005), and the stability correction expressed through the Richardson number (Budyko et al., 1962) were not cited clearly enough. And the revised sentence please refer to our response to Comment 5 in the General Comments section.

(7) *Section 4.1 provides a useful inter-comparison between measurements from the two co-located gas analysers. It would be good if some correlation statistics could be provided, with a linear regression fit. The 10 Hz data can be averaged to every second for this comparison.*

Response: We agree that a quantitative intercomparison between the two co-located gas analyzers would improve Section 4.1. Following the reviewer's suggestion, the 10 Hz data were averaged to 1 s, and linear regression statistics were added to evaluate the consistency between the two instruments. The revised manuscript now includes the corresponding correlation statistics and regression fit：

"Further, the scatterplot of the CO_2 measurement results from both instruments, shown in Fig. 5, reveals a strong linear relationship. The data are closely distributed around a narrow band of the fitted linear regression line, with a correlation coefficient of r = 0.986. Because both analyzers are subject to measurement uncertainties, reduced major axis (RMA) regression was applied. The RMA regression yielded a slope of 1.19 and an intercept of 0.64 ppm. The close clustering of the data points around the fitted line indicates low random error and high precision between the two instruments. This regression result indicates that the UAV observations were able to capture temporal variability in CO_2 in the real atmosphere, such as minute-scale increases, decreases, and multi-peak structures. In addition, the calculations of the horizontal line flux divergence and the vertical flux both require subtraction of the background CO_2 concentration and estimation of the CO_2 gradient, respectively. Therefore, the systematic bias between the two instruments is expected to have a negligible impact on the flux calculations.”

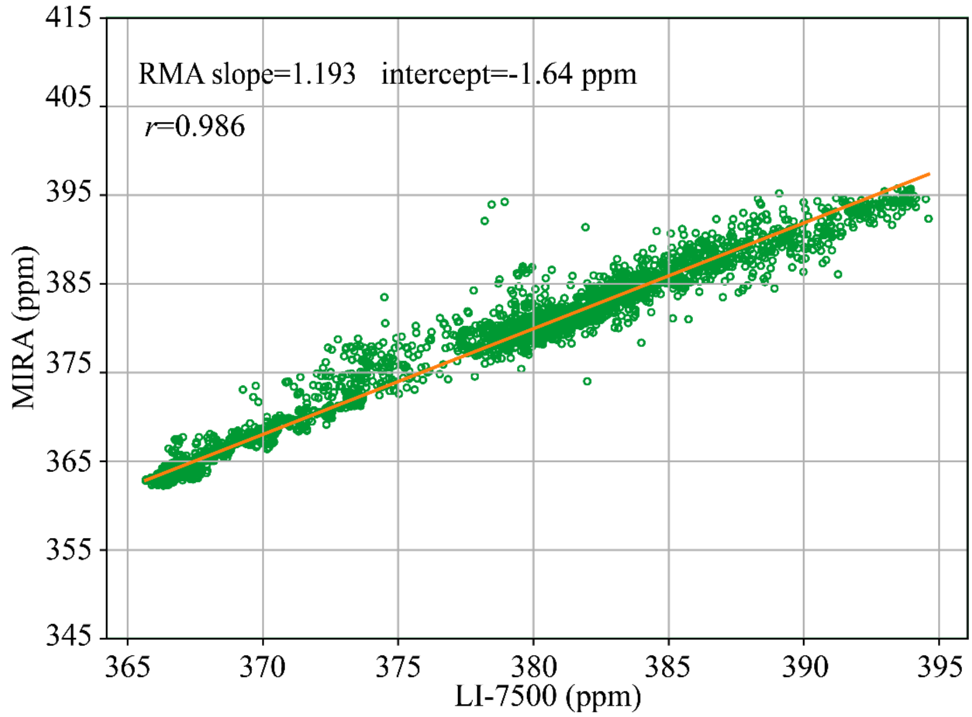


Figure 5. Scatterplot comparison of CO_2 mole fractions measured by the UAV-mounted MIRA analyzer and the flux-tower LI-7500 analyzer. The orange line denotes the reduced major axis (RMA) regression. The regression yielded a slope of 1.193 and an intercept of -1.64 ppm, with a correlation coefficient of r = 0.986, indicating strong agreement between the two instruments under field conditions.

The specific revisions are shown in lines 536 to 555 of the revised manuscript.

(8) *Equation 3.16 proposes that a vertical flux can be derived from mean concentrations at two heights. However, the heights are not provided in the manuscript. Section 4 presents vertical fluxes without describing which two specific heights these fluxes are derived from.*

Response: We agree that the previous version of the manuscript did not describe clearly enough which two heights were used in the calculation of the vertical flux in Eq. (3.16). This information is essential for understanding both the method and the interpretation of the results presented in Sect. 4.

In the revised manuscript, we have clarified this point explicitly in Sect. 3.2.2. The vertical flux at a given target height is derived using UAV profile measurements taken 5 m below and 5 m above that height. Thus, the two heights used in Eq. (3.16) are the adjacent levels bracketing the target level within the UAV ascent profile. The revised sentence is as follows:

"where g is the gravitational acceleration constant; $\Delta\bar{\theta}$ and $\Delta\bar{u}$ represent the changes in the temperature and wind speed between the layers (Δz), after smoothing with a seven-point moving window. In applying Eq. (3.12), the smoothed $\bar{\theta}$ and $\bar{u}$ vertical profiles are directly used to compute R_B, taking respective data points from two adjacent heights, z_1 and z_2, using the smoothed temperature and wind speed profiles at the two heights for $\overline{u_1}$, $\overline{u_2}$, $\overline{\theta_1}$, and $\overline{\theta_2}$. When applying the gradient method to estimate the flux, the vertical separation, $\Delta z = z_2 - z_1$, is typically chosen within a range of 1-10 m (Meredith et al., 2014). In this study, $\overline{u_1}$, $\overline{u_2}$ and $\overline{\theta_1}$, $\overline{\theta_2}$ were taken 5 m below and 5 m above any target height ($\Delta z = 10$ m), respectively, during the UAV ascent in the vertical profile measurements."

The specific revisions are shown in lines 442 to 451 of the revised manuscript.

For the comparison with the eddy-covariance system installed at 31 m above ground level, the UAV-derived vertical flux was calculated using measurements at 26 m and 36 m. We have now added this information explicitly in both Sect. 3.2.2 and the relevant part of Sect. 4.2.1 to ensure that the derivation and interpretation of the vertical fluxes are clear. The revised sentence is as follows:

"For the comparison with the EC system installed at 31 m above ground, the UAV-derived vertical flux was calculated using measurements at z_2=26 m and z_1=36 m."

The specific revisions are shown in lines 567 to 569 of the revised manuscript.

(9) *Many of the conclusions of this work are based on the coefficient of determination of the linear fit between Eddy covariance and UAV fluxes, with each of these two types of fluxes calculated independently of each other. The authors should instead present the Pearson correlation coefficient where the coefficient of determination is currently used throughout this manuscript. A coefficient of determination is used to evaluate the ability of a model to fit a data input to a data output. Yet, in this work, the purpose of the linear fit is not to use one flux to derive the other flux but rather to compare two flux values expected to have the same value. A coefficient of determination is therefore better is it evaluates how two sets of results, expected to have the same value, are correlated together.*

Response: Thank you for this important and constructive comment. We agree with the reviewer that, in the present study, the purpose of the regression analysis is not to use one flux measurement to predict the other, but rather to evaluate the degree of agreement between two independently derived flux estimates. In this context, the Pearson correlation coefficient (r) is more appropriate than the coefficient of determination (R^2) for describing the strength of the linear association between the UAV-derived fluxes and the eddy-covariance (EC) measurements.

In the revised manuscript, we have therefore replaced the coefficient of determination with the Pearson correlation coefficient (r) throughout the manuscript wherever the comparison is intended to assess the correlation between two independently measured or derived quantities. We have also revised the corresponding text to avoid implying a predictive input–output relationship between the EC and UAV fluxes. The regression analysis is now described more clearly as a comparison of two independent estimates of the same physical quantity. We thank the reviewer for pointing out this important issue, which has helped us improve the statistical interpretation and presentation of the results.

(10) *Section 4.3 compares different total flux estimates, including different boxes covering different sized areas. However, different sets of UAV flights are used for each of the different area sizes (this is clear because the number in Table 5 is not the same for each square size). These different datasets are compared directly in Figure 9 and Figure 10. I do not think direct comparisons can be made unless the same flights are used for all three distances. Averages from the same set of flights should be compared. If one flight is rejected at 50 m, it should also be rejected at all distances*

Response: In fact, the measurements for the three different control volumes were obtained within a relatively short time window (20 min). This design was intended to ensure that the resulting values were as comparable as possible. Because we did not explain this clearly in the original manuscript, the reviewer may have been led to assume that the comparison was not based on the same set of observations. We sincerely apologize for this lack of clarity and have now added an explicit explanation in the revised manuscript.

In addition, regarding the sample size of n =7 or 8 reported in Table 5, we would like to clarify that, on only one of those days, we were unable to obtain data for the 50 × 50 × 50 m^3 control volume, whereas the other measurements were acquired nearly simultaneously within a short period. We fully agree with the reviewer that samples with inconsistent coverage should not be directly compared. Accordingly, we have removed this incomplete case from the revised manuscript and redrawn Figs. 9 and 10. Moreover, in the original manuscript we only stated that the selected datasets corresponded to autumn mornings and summer noons, which we now realize was insufficient. To better demonstrate that these cases were obtained under highly comparable observational conditions, we have additionally compared their observation periods, wind conditions, and EC fluxes. The revised sentence is as follows:

"To ensure the validity of the comparison, measurements from the three virtual control volumes must be obtained within 20 minutes of each other. Under such a requirement, data from seven days of observations were used for the spatial heterogeneity analysis. Further examination of the measurement conditions during these periods showed that, for the summer noon cases, observations were conducted between 13:00 and 14:00 under predominantly southeasterly winds, while the fluxes measured by the fixed tower varied by less than 8% during the comparison period. For the autumn morning cases, observations were conducted between 06:00 and 07:00 under predominantly southwesterly winds, while the fluxes measured by the fixed tower varied by less than 5%. Overall, the observational conditions were considered sufficiently similar to support further comparative analysis."

The specific revisions are shown in lines 771 to 781 of the revised manuscript.

The sample size reported in the table has also been revised; please refer to the table below.

	Summer Noon			Autumn Morning		
Box Size (m^3)	50×50×50	100×100×50	150×150×50	50×50×50	100×100×50	150×150×50
n	7	7	7	7	7	7
E_C Mean (kg h^{-1})	-2.24	-10.6	-18.6	0.166	7.64	9.35
E_C Median (kg h^{-1})	-2.79	-13.7	-12.8	0.138	7.25	11.3
IQR (kg h^{-1})	2.65	7.24	20.4	1.85	3.26	15.2
$E_{C,H}$ Median (kg h^{-1})	0.0687	0.0817	0.273	0.105	3.52	6.41
$E_{C,H}$ SD (kg h^{-1})	0.0162	0.0321	0.0732	0.0235	1.25	1.58
$E_{C,VT}$ Mdian (kg h^{-1})	-2.77	-13.6	-12.5	0.033	3.73	4.85
$E_{C,VT}$ SD (kg h^{-1})	0.437	2.35	3.54	0.012	1.15	1.43
Median $E_{C,H}/E_{C,VT}$	-0.025	-0.006	-0.022	3.18	0.944	1.32

The specific revisions are shown in lines 815 to 820 of the revised manuscript.

Finally, the reviewer raised the important concern that conclusions based on only seven or eight data points may not be sufficiently robust. We would like to explain that the limited sample size was a direct consequence of our effort to ensure that the comparison was made only under highly similar observational conditions. Although we collected a total of 168 UAV flights, we considered the influence of boundary-layer conditions on flux measurements and therefore restricted the analysis to morning, noon, and evening periods. We also considered the seasonal dependence of CO_2 exchange in the forest ecosystem and therefore further separated the data into spring, summer, autumn, and winter. Ultimately, the datasets suitable for comparison were limited to summer noon and autumn morning cases. In addition, only those cases in which the three different box sizes were measured within a short time window could be retained, which resulted in seven valid datasets.

We believe that ensuring comparable observational conditions is the most important prerequisite for a meaningful comparison. At the same time, we were fully aware of the limitation imposed by the small sample size. To address the issue of the limited number of data points, we further performed ANOVA to help ensure the reliability of the results as much as possible.

(11) *The authors assert that this horizontal flux component can be as important as the vertical turbulent component. This was evaluated by taking the ratio between the two flux components. While this may be the case, it is not clear whether this is due to an especially large horizontal flux component or if the vertical turbulent flux component*

is naturally small at certain periods. Figure 6 suggests that the vertical flux component is naturally small in the morning, which might explain why the horizontal flux component appears to be so dominant in the autumn morning. In reality, if the flux magnitudes are compared, perhaps it is not so important.

Response: We agree that the ratio $E_{C,H}/E_{C,VT}$ alone is not sufficient to demonstrate the relative importance of the horizontal-transport term, because a large ratio may arise either from an enhanced horizontal component or from a naturally weak vertical turbulent-exchange component. Following the reviewer's suggestion, we have revised the discussion to consider not only the ratio $E_{C,H}/E_{C,VT}$, but also the absolute magnitudes of $E_{C,H}$ and $E_{C,VT}$ reported in Table 5.

Specifically, during summer noon, the median $E_{C,H}$ values are very small (0.0687-0.273 kg h^{-1}) compared with the much larger median $E_{C,VT}$ values (-13.6 to -2.77 kg h^{-1}), indicating that the net CO_2 budget is clearly dominated by vertical turbulent exchange during this period. In contrast, during autumn morning, the median $E_{C,VT}$ values are much smaller (0.033-4.85 kg h^{-1}), while the median $E_{C,H}$ values range from 0.105 to 6.41 kg h^{-1}. Therefore, the relatively large $E_{C,H}/E_{C,VT}$ ratios observed in autumn morning reflect not only the presence of horizontal CO_2 gradients, but also the comparatively weak vertical turbulent-exchange component under weak-turbulence conditions.

We have revised the manuscript accordingly to make this interpretation clearer and more balanced. The rewritten text is as follows:

"In addition, during summer at noon, the median values of $E_{C,H}/E_{C,VT}$ in Table 5 remain very small (from -0.025 to -0.006), indicating that the net CO_2 budget is dominated by the vertical turbulent-exchange term. This conclusion is further supported by the absolute magnitudes of the two components: the median $E_{C,H}$ values range only from 0.0687 to 0.273 kg h^{-1}, whereas the corresponding median $E_{C,VT}$ values range from -13.6 to -2.77 kg h^{-1}. Under summer noon conditions, stronger turbulence and more efficient vertical mixing likely suppress the relative influence of horizontal CO_2 gradients, so that horizontal transport contributes only a minor fraction to the virtual control volume budget. Although the negative values of $E_{C,H}/E_{C,VT}$ indicate that the horizontal and vertical components often have opposite signs, the magnitude of $E_{C,H}$ remains much smaller than that of $E_{C,VT}$."

Table 5. Summary statistics of net CO_2 emission estimates (E_C) derived from three box sizes during summer noon and autumn morning. For each box size, n denotes the number of valid samples. Reported metrics include the mean and median of E_C (kg h^{-1}), the interquartile range (*IQR*), the median and standard deviation (*SD*) of the horizontal-transport contribution ($E_{C,H}$) and the vertical-exchange contribution ($E_{C,VT}$), as well as the median ratio $E_{C,H}/E_{C,VT}$. All emission estimates are expressed in kg h^{-1}.

	Summer Noon			Autumn Morning		
Box Size (m^3)	50×50×50	100×100×50	150×150×50	50×50×50	100×100×50	150×150×50
n	7	7	7	7	7	7
E_C Mean (kg h^{-1})	-2.24	-10.6	-18.6	0.166	7.64	9.35

E_C Median (kg h^{-1})	-2.79	-13.7	-12.8	0.138	7.25	11.3
IQR (kg h^{-1})	2.65	7.24	20.4	1.85	3.26	15.2
$E_{C,H}$ Median (kg h^{-1})	0.0687	0.0817	0.273	0.105	3.52	6.41
$E_{C,H}$ SD (kg h^{-1})	0.0162	0.0321	0.0732	0.0235	1.25	1.58
$E_{C,VT}$ Mdian (kg h^{-1})	-2.77	-13.6	-12.5	0.033	3.73	4.85
$E_{C,VT}$ *SD* (kg h^{-1})	0.437	2.35	3.54	0.012	1.15	1.43
Median $E_{C,H}/E_{C,VT}$	-0.025	-0.006	-0.022	3.18	0.944	1.32

The specific revisions are shown in lines 815 to 820 of the revised manuscript.

"In contrast, during autumn morning, the role of horizontal transport becomes much more pronounced. The median values of $E_{C,H}/E_{C,VT}$ increase to 3.181, 0.944, and 1.321 for the 50×50×50 m^3, 100×100×50 m^3, and 150×150×50 m^3 control volumes, respectively. Importantly, these elevated ratios reflect not only the presence of horizontal CO_2 gradients, but also the fact that the vertical turbulent-exchange component is comparatively weak during this period. Specifically, the median $E_{C,H}$ values range from 0.105 to 6.406 kg h^{-1}, while the median $E_{C,VT}$ values range from 0.033 to 4.848 kg h^{-1}. This indicates that, under weak-turbulence morning conditions, horizontal CO_2 transport can become comparable to, or even exceed, the vertical turbulent component. Such behavior is consistent with enhanced spatial heterogeneity in CO_2 exchange across the different control volumes. As the control-volume scale increases, small-scale horizontal gradient structures are progressively averaged out, reducing the relative contribution of $E_{C,H}$ and leading to the observed scale dependence of E_C."

The specific revisions are shown in lines 831 to 843 of the revised manuscript.

Specific comments

(1) *Line 32: "evaluated"*

• *How can a platform be evaluated? I think this refers to a UAV flux quantification methodology, rather than the UAV platform itself.*

Response: We also agree that using the term UAV platform is indeed not very appropriate. Therefore, in accordance with your suggestion, we have revised UAV-based platform to "A UAV-based flux quantification methodology, capable of three-dimensional CO_2 emission computation, was developed, evaluated and applied to a forest ecosystem."

The specific revisions are shown in lines 31 to 32 of the revised manuscript.

(2) *Line 32: "applied to a forest ecosystem"*

• *How can a UAV-based platform be "applied" to a forest ecosystem? This doesn't* *really make sense (applying a hardware system to an environment). Again, perhaps this refers to the overall flux quantification methodology.*

Response: We did not carefully refine the wording, and we sincerely appreciate your meticulous identification of the problems in our phrasing. According to your comment (1), we have revised UAV platform to UAV flux quantification methodology.

The specific revisions are shown in lines 31 to 32 of the revised manuscript.

(3) *Line 34: "box-pattern and profile-pattern"*

• It isn't clear to me what this means at this stage. What is a box-pattern and profile-pattern flight? Perhaps a few details can be given here.

Response: You are right that directly mentioning box-pattern and profile-pattern in the abstract may be confusing to readers. We have therefore revised this part to a more straightforward and accessible expression:

"the UAV measurement system is flown along horizontal loops at multiple heights together with vertical ascent flight tracks, acquiring data along the flight tracks that are used to simultaneously quantify horizontal advective and vertical turbulent carbon dioxide emission components."

The specific revisions are shown in lines 36 to 40 of the revised manuscript.

(4) *Line 36: "resolve horizontal and vertical exchange"*

• It is really not clear what this refers to in the abstract. Is this referring to the exchange of carbon dioxide (i.e. carbon dioxide flux) or some sort of study on atmospheric dynamics? I think the authors can be clearer here that this manuscript quantifies horizontal advective and vertical turbulent carbon dioxide flux components.

• Based on the manuscript, the horizontal turbulent component is ignored and the vertical non-turbulent component is ignored. This is not reflected here in the abstract, where it seems like all vertical and horizontal components are considered.

Response: Regarding your first point, the sentence we used was indeed not sufficiently clear, and your suggested wording is very appropriate for our abstract. In addition, although our calculation method requires fluxes as inputs, the quantity we actually calculate is the emission rate, with units of kg/h, which is distinct from flux. Therefore, we have replaced "resolve horizontal and vertical exchange" in the revised manuscript with "quantify horizontal advective and vertical turbulent carbon dioxide emission components"

Regarding your second point, we also find your comment very reasonable. Therefore, to present our work clearly in the abstract, we have adopted the revision you suggested in your first point.

Finally, considering sentence fluency and overall expression, the text has been revised as follows: The system integrates a high-precision closed-path CO_2 analyzer with a calibrated ultrasonic anemometer. Within a mass-balance framework, the UAV measurement system is flown along horizontal loops at multiple heights together with vertical ascent flight tracks, acquiring data along the flight tracks that are used to simultaneously quantify horizontal advective and vertical turbulent carbon dioxide emission components.

The specific revisions are shown in lines 36 to 40 of the revised manuscript.

(5) *Line 38: "R2 ≈ 0.76–0.77"*

• *Why is the coefficient of determination used here? When comparing like-for-like flux estimates it would be better to provide the correlation coefficient. The coefficient of determination is more apt when evaluating the goodness of a model fit, which is not taking place here.*

• *Why is an approximately equal sign used for a range? If a range is provided, it is simpler to just precisely give the upper and lower bounds of the range.*

Response: Regarding your first point, we agree that the correlation coefficient is more appropriate than the coefficient of determination for a direct comparison of like-for-like flux estimates, because no model fitting is involved in this case. Therefore, we have replaced all R^2 values with the correlation coefficient r in the revised manuscript.

Regarding your second point, we also agree that using an approximate sign for a reported range is unnecessary. We have replaced "R^2 ≈ 0.76–0.77" in the revised manuscript with ", with a correlation coefficient r of 0.9".

The specific revisions are shown in line 43 of the revised manuscript.

(6) *Line 38: "Box-pattern flights"*

• *The UAV flights themselves cannot be used to infer diurnal variation but rather the measurements from flights are used to makes such evaluations.*

Response: We have replaced "Box-pattern flights" in the revised manuscript with "Horizontal measurements further reveal pronounced local canopy air diurnal CO_2 mass flux variation:".

The specific revisions are shown in lines 44 to 45 of the revised manuscript.

(7) *Line 41: "Sensitivity analyses across scales of 50 m, 100 m, 150 m indicate that CO2 emission intensity is sensitive to control-volume dimensions and shows spatial heterogeneity above the forest."*

• *This sentence is currently unhelpful in the abstract, without any details on the nature of the sensitivity analysis.*

• *What does it mean to say that emission intensity is sensitive to control-volume dimensions? What are control-volume dimensions? I do not follow this.*

• *A forest ecosystem being heterogeneous is not a new discovery. Again, without any details on the nature of the sensitivity analysis, I don't know the novelty what is being revealed here and its significance.*

• *Having read the manuscript, I now understand what the three different distances mean. However, at this point in the abstract, this is vague. It is not clear that these refer to different sized boxes.*

Response: Thank the reviewer for the comment.

Regarding your first point, we rewrote this section and provided a detailed description.

Regarding your second and forth point, we agree that the original sentence was too vague to convey the meaning of the scale analysis, particularly with regard to the dimensions of the control volume. We have replaced “scales of 50 m, 100 m, 150 m” in the revised manuscript with “multi-scale grid boxes (50×50×50 m^3, 100×100×50 m^3, and 150×150×50 m^3)”.

Regarding your third point. We intend to emphasize that the key innovation of this study lies in the ability of our newly developed three-dimensional measurement approach to detect spatial variability in CO_2 exchange above the forest canopy. Because the manuscript was submitted to *Atmospheric Measurement Techniques* under the title *UAV-based method for measuring CO_2 exchange in forest ecosystems*, the present paper focuses primarily on the methodological innovation of the three-dimensional measurement framework, rather than on an in-depth analysis of spatial heterogeneity. In fact, our method makes it possible to quantify this spatial heterogeneity, which cannot be achieved through stationary towers measurements, as such measurements can only provide an estimated footprint area. Specifically, emission intensities were calculated using three different control-volume sizes, from which the standard deviation and mean were derived; the coefficient of variation was then obtained as the ratio of the standard deviation to the mean. Furthermore, the coefficients of variation were calculated for different periods across the four seasons, as shown in the table below. Finally, a statistical analysis was applied to derive the mean coefficient of variation (28.3% in our study), which can be used as an indicator of the effect of spatial heterogeneity.

Table 1. Coefficients of variation of emission factors for different area sizes across time periods in different seasons.

Period	Spring	Summer	Autumn	Winter
Morning	38.3%	15.6%	68.4%	13.8%
Noon	12.5%	10%	20.3%	14.7%
Evening	56.6%	20.1%	54.8%	36.2%

In summary, we wish to emphasize the innovation of the three-dimensional measurement approach in the present manuscript, as the main focus of this study is on the measurement methodology itself. Accordingly, we have not expanded the discussion of spatial heterogeneity at this stage. Nevertheless, if the reviewer considers a more detailed discussion necessary, we would be happy to incorporate further revisions.

Finally, the text has been revised as follows:

“Sensitivity analysis of emission intensities across multi-scale grid boxes (50×50×50 m^3, 100×100×50 m^3, and 150×150×50 m^3) revealed stronger scale dependence during autumn mornings, indicating significant spatial heterogeneity in CO_2 emissions.”

The specific revisions are shown in lines 46 to 49 of the revised manuscript.

(8) *Line 62: “estimate CO2 exchange”*

• The list provided in this sentence is not a list of methods to estimate carbon dioxide exchange rates but rather a list of sampling platforms. Please can the authors clarify this? Measurements from different sampling platforms can be treated in different ways to evaluate gas exchange rates and fluxes.

Response: We sincerely thank the reviewer for the exceptionally careful and detailed examination of each sentence. This statement was indeed not sufficiently precise, and we have revised it accordingly following your suggestion.

The text has been revised as follows:

"despite decades of carbon-cycle research aimed at estimating CO_2 exchange across spatial scales using a variety of observational platforms, including stationary towers, satellite-based remote sensing, and manned aircraft measurements"

The specific revisions are shown in lines 70 to 71 of the revised manuscript.

(9) *Line 62: "eddy covariance (EC) towers"*

• I would recommend replacing this with "stationary towers" because that is what all systems from which Eddy covariance fluxes are derived actually are, as a sampling platform. Fluxes from towers are often evaluated using the Eddy covariance methodology, but not necessarily.

• The Eddy covariance is a method of flux calculation whereas a tower is a platform (like a satellite).

Response: We thank the reviewer for this comment.

We have replaced "eddy covariance (EC) towers" in the revised manuscript with "stationary towers".

The specific revisions are shown in line 72 of the revised manuscript.

(10) *Line 73: "most widely used method"*

• Please provide some citations here to support this assertion.

Response: We have added citations to the relevant literature.

(Aubinet et al., 2012; Baldocchi, 2003; Burman et al., 2025; Pastorello et al., 2020; Fang et al., 2024)

The specific revisions are shown in lines 83 to 84 of the revised manuscript.

(11) *Line 75: "(Aubinet et al., 2012; Baldocchi et al., 2001; Baldocchi, 2003; Wilson et al., 2002)"*

• Are there any recent references? The youngest citation in this list is from 14 years ago. If there is no recent work with FLUXNET, please can the authors state that this is a historical initiative that is no longer operational?

Response: We have added two recent references on FLUXNET-related research.

(Aubinet et al., 2012; Baldocchi et al., 2001; Fang et al., 2024; Pastorello et al., 2020; Wilson et al., 2002)

The specific revisions are shown in line 86 of the revised manuscript.

(12) *Line 104: "complex"*

• In what way are these environments complex?

We replaced "forested or topographically complex terrains" with "complex terrain and heterogeneous surfaces". The latter is a common and technical expression in atmospheric boundary layer and flux research. The same expression can be found in the following references:

[1] AMS conference call-for-papers title: https://www.ametsoc.org/ams/meetings-events/ams-meetings/32agforest-22blt-3biogeo/call-for-papers/?utm_source=chatgpt.com

	issues and new perspectives in eddy covariance measurements
Understanding and uncertainties in measuring and modeling ecosystem-atmosphere exchanges	Boundary layer processes over complex terrain and heterogeneous surfaces
	Novel instruments and systems for measuring ecosystem - atmosphere exchanges
	New theory, ideas and approaches in understanding ecosystem -atmosphere exchanges

[2] Wang et al., 2021, ACP: The impact of inhomogeneous emissions and topography on ozone photochemistry in the vicinity of Hong Kong Island. https://doi.org/10.5194/acp-21-3531-2021

> the terrain on the segregation was not considered. Previous studies showed that complex terrain has an important impact on the turbulence structure in the boundary layer (e.g., Cao et al., 2012; Rotach et al., 2015; Liang et al., 2020) and on the evolution with time of the boundary layer height (De Wekker and Kossmann, 2015). Therefore, the segregation intensity is

[3] Li et al., 2017, Atmospheric Research: Effect of a cold, dry air incursion on atmospheric boundary layer processes over a high-altitude lake in the Tibetan Plateau. https://doi.org/10.1016/j.atmosres.2016.10.024

> In recent years, numerous efforts, ranging from the field observation (Barlow et al., 2011; Han et al., 2010; Laiti et al., 2013) to the numerical modeling (Brunsell et al., 2011; Gerken et al., 2013; Lü et al., 2004; Reen et al., 2014), have been made to explore the atmospheric boundary layer (ABL) processes over heterogeneous surfaces (e.g. lake, oasis and urban). A radiosonde study suggested that the convective boundary layer (CBL) can grow up to 1750 m in the Nam Co Lake basin (Lü et al., 2008). However, it is difficult to inspect in detail the ABL character-

The specific revisions are shown in lines 91 to 92 of the revised manuscript.

(13) *Line 107: "fixed EC towers"*

• *I would recommend replacing with this with "stationary towers" as not all measurements from towers are used for Eddy covariance fluxes.*

Response: We replaced "fixed EC towers" with "stationary towers"

The specific revisions are shown in line 119 of the revised manuscript.

(14) *Line 114: "flight attitude"*

• *What does flight attitude mean? Maybe this is a technical UAV term, but it must be defined in an atmospheric science journal.*

• *Why does this effect carbon dioxide flux estimates?*

Response: Regarding your first point, we have revised "flight attitude" to "changing UAV attitude during flight". The UAV attitude refers to its pitch, roll, and yaw angles. Its changes during flight lead to false wind signals measured from the UAV (Yang et al., 2024). We added "(e.g., variations in the pitch, roll, and yaw angles) (Yang et al., 2024)" in the revised manuscript.

Regarding your second point, we have revised the sentence to clarify that changes in flight attitude may affect the measurements of wind speed and greenhouse gas concentration, thereby influencing the accuracy of CO_2 flux estimates. The revised sentence is as follows:

"Nonetheless, these UAV applications have not demonstrated the feasibility of CO_2 flux measurements owing to their inability to conduct accurate wind speed measurements, both of which can be affected by propeller-induced airflow disturbances (Soltaninezhad et al., 2025; Thielicke et al., 2021) and rapidly changing UAV attitude during flight (e.g., variations in pitch, roll, and yaw angles) (Yang et al., 2024), thereby introducing additional uncertainty into the flux calculation."

The specific revisions are shown in lines 125 to 128 of the revised manuscript.

(15) *Line 118: "disturbances"*

• *What does this mean? What sort of disturbances does this refer to?*

Response: We agree that the term "disturbances" in the original sentence was too vague. To improve clarity, we have revised the sentence to explicitly specify the types of disturbances involved, including rotor downwash, platform-induced airflow, and rapid changes in UAV motion and attitude, all of which may affect wind measurements made from UAV platforms. The revised sentence is as follows:

"To address these challenges, various sensors, including multi-hole probes, pitot tubes, and sonic anemometers, have been deployed on UAVs for wind measurements and correction algorithms have been developed to mitigate measurement disturbances associated with rotor downwash, re-orientation of airflow by platform geometry, and rapid changes in UAV motion and attitude."

The specific revisions are shown in lines 131 to 133 of the revised manuscript.

(16) *Line 126: "GHG"*

• *This term has not been defined and is not used elsewhere in the manuscript.*

Response: Indeed we agree that the abbreviation "GHG" was not yet defined here and was not used consistently elsewhere in the manuscript. To improve clarity and consistency, we have revised the sentence by replacing "GHG" with "greenhouse gas concentrations". The revised sentence is as follows:

Building on our previous work on the development of an UAV platform for integrated high sensitivity, high precision, fast response measurements of greenhouse gas concentrations…

The specific revisions are shown in lines 141 to 142 of the revised manuscript.

(17) *Line 127: "3-D"*

• *This term has not been defined.*

Response: We agree that "3-D" was not defined clearly. Accordingly, we have

replaced "3-D wind speed" with "three-dimensional wind speed" in the revised manuscript. The revised sentence is as follows:
"Building on our previous work on the development of a UAV platform for integrated high-sensitivity, high-precision, and fast-response measurements of greenhouse gas concentrations and three-dimensional wind speed, ..."

The specific revisions are shown in line 142 of the revised manuscript.

(18) *Line 130: "UAV platform"*

• *It is not clear what "evaluating" a "UAV platform" means. Does this mean testing the sensor on the platform? Does it mean evaluating a flux quantification algorithm using data acquired from on-board the UAV? Perhaps it means something else.*

Response: We agree that the phrase "evaluate the UAV platform" was ambiguous. Our intention was not to imply an evaluation of the hardware platform alone, but rather an assessment of how the UAV-based measurement system and the wind correction algorithm perform in quantifying three-dimensional CO_2 transport and deriving emission estimates. We have therefore revised the sentence to make this point explicit. The revised sentence is as follows:

"we use observations collected with the UAV platform to evaluate the performance of the wind correction algorithm, assess the capability of the integrated UAV measurement system to quantify three-dimensional CO_2 transport, including canopy-scale vertical CO_2 fluxes, and characterize the robustness and uncertainty of the UAV-derived emission/uptake estimates"

The specific revisions are shown in lines 145 to 150 of the revised manuscript.

(19) *Line 133: "emission"*

• *Maybe "emission" is not the best word here and throughout the manuscript as "emission" implies that carbon dioxide is generally released into the atmosphere. Yet, during photosynthesis, it is absorbed resulting in a negative flux. I would recommend evaluating the use of this word throughout the manuscript.*

Response: We agree that the term "emission" alone may be misleading, as ecosystem–atmosphere CO_2 exchange includes not only release to the atmosphere but also CO_2 uptake during photosynthesis. To address this concern while retaining the original terminology framework of the manuscript, we have revised the wording to "CO_2 emission/removal estimates" where appropriate, so as to explicitly include both positive and negative fluxes. The revised sentence is as follows:

"we use observations collected with the UAV platform to evaluate the performance of the wind correction algorithm, assess the capability of the integrated UAV measurement system to quantify three-dimensional CO_2 transport, including canopy-scale vertical CO_2 fluxes, and characterize the robustness and uncertainty of the UAV-derived emission/uptake estimates."

The specific revisions are shown in line 150 of the revised manuscript.

In addition, we believe that it is preferable to retain the term "emission" throughout the manuscript for two main reasons. First, although our measurement approach requires flux calculations, the final quantity derived in this study is the emission rate (kg h^{-1}), which differs in units from flux. Second, in our subsequent work, we intend to use these emission estimates to further derive emission factors. According to the *2006 IPCC Guidelines for National Greenhouse Gas Inventories*, Chapter 4 (“Forest Land”), the terms “emissions/removals” for forests are also collectively referred to as “emissions” in many contexts (please see the figure below). We recognize that terminology may vary across disciplines; however, our work is more closely aligned with the field of atmospheric pollutant emissions, and research on natural-source emissions represents only one part of our broader research scope. We also study industrial emissions, and our future work may require integrating anthropogenic and natural-source emission results to develop a more complete emission inventory. For these reasons, we believe that retaining the term “emission” is more appropriate. Nevertheless, if the reviewer still considers a revision necessary, we would be very willing to make further modifications in accordance with the reviewer’s suggestion.

Chapter 4: Forest Land

4.2 FOREST LAND REMAINING FOREST LAND

This section deals with managed forests that have been under Forest Land for over 20 years (default), or for over a country-specific transition period. Greenhouse gas inventory for *Forest Land Remaining Forest Land* (FF) involves estimation of changes in carbon stock from five carbon pools (i.e., above-ground biomass, below-ground biomass, dead wood, litter, and soil organic matter), as well as emissions of non-CO_2 gases. Methods for estimating greenhouse gas emissions and removals for lands converted to Forest Land in the past 20 years (e.g., from Cropland and Grassland) are presented in Section 4.3. The set of general equations to estimate the annual carbon stock changes on Forest Land are given in Chapter 2.

(20) *Line 137: “mass flux measurements”*

• The flux is not itself measured, but rather computed from other measurements. So I don’t think “measurements” is the right word here.

Response: We agree that flux is not measured directly, but rather calculated from the underlying observations. To improve accuracy, we have revised the wording from “mass flux measurements” to “mass flux estimates derived from this platform.” The revised sentence is as follows:

“While horizontal mass flux estimates derived from this platform have been previously evaluated and validated through field measurements, ...”

The specific revisions are shown in line 154 of the revised manuscript.

(21) *Line 141: “tower-based EC flux”*

• This is the correct way to describe EC fluxes derived from a stationary tower. This phrase makes it clear that measurements from a tower (the platform) are used within the Eddy covariance method (the flux quantification technique). This statement correctly does not imply that all towers are used for Eddy covariance fluxes. Please can the authors apply this correct approach to the rest of the manuscript to avoid ambiguity?

Response: We agree that the expression “tower-based EC flux” is a clearer and more precise way to distinguish the observational platform from the flux quantification technique. Following the reviewer’s suggestion, we have carefully revised the relevant expressions throughout the manuscript to improve consistency and avoid ambiguity. In particular, we now distinguish more explicitly between the measurement platform (e.g., tower-based, UAV-based) and the methodological approach used to derive the fluxes (e.g., eddy covariance, mass balance, or related calculations). We believe these revisions have improved the clarity and precision of the manuscript.

(22) *Line 156: "the height from the body stands to the body top is 565 mm"*
• *Does this include the height of the air inlet? Please clarify this.*

Response: We agree that the original description was not sufficiently clear regarding whether the reported UAV height included the air inlet. To avoid ambiguity, we have revised the sentence to explicitly state that the reported height refers to the distance from the landing gear to the top of the main body and does not include the height of the air inlet. The revised sentence is as follows:

"The UAV has a maximum horizontal span, including the arms and rotor blades, of 2570 mm, and the height from the landing gear to the top of the main body is 565 mm, excluding the air inlet."

The specific revisions are shown in lines 172 to 175 of the revised manuscript.

(23) *Line 168: "global positioning system (GPS)"*
• *The UAV probably contains a "GPS sensor" or "GPS tracker". The GPS itself is in space.*

Response: We agree that the original wording was not sufficiently precise, since the GPS itself refers to the satellite-based positioning system rather than the onboard device installed on the UAV. Following the reviewer's suggestion, we have revised the sentence by replacing "global positioning system (GPS)" with "global positioning system (GPS) sensor" to improve technical accuracy. The revised sentence is as follows:

"The airframe is constructed from carbon-fiber composite materials and houses an integrated flight control system, a power supply module, and a communication unit. The flight control system includes an inertial measurement unit (IMU), accelerometers, and a global positioning system (GPS) sensor capable of real-time kinematics (RTK)."

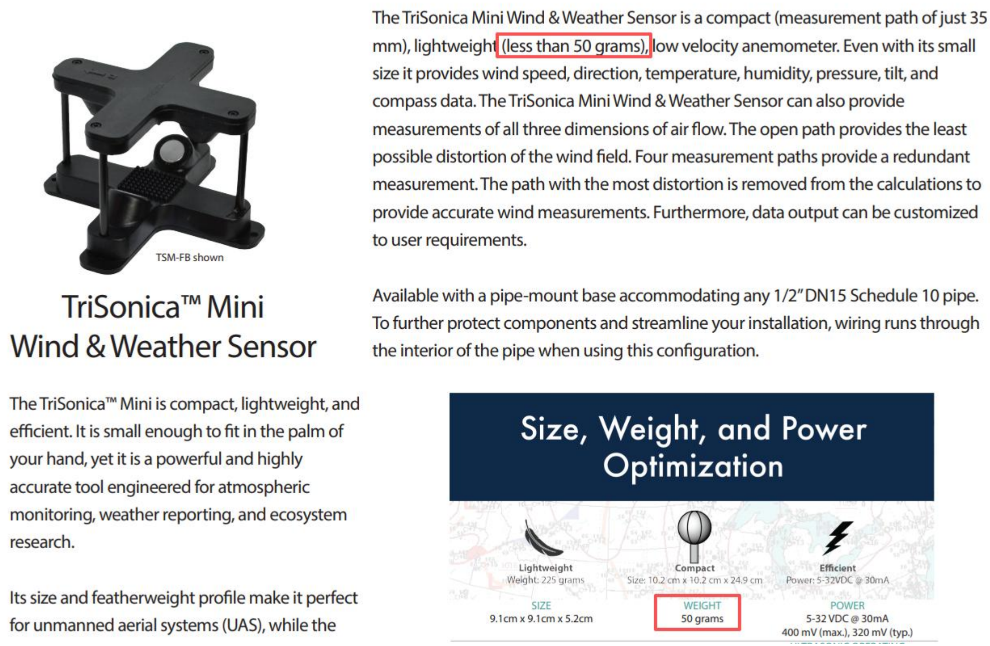
The specific revisions are shown in line 187 of the revised manuscript.

(24) *Lin 186: "< 50 g"*
• *Is it less than 50 g or equal to 50 g? In Table 2, it is given as equal to 50 g, not less than. In any case, it is better to give the actual mass, rather than saying it is less than something.*

Response: We agree that the original expression "< 50 g" was ambiguous and inconsistent with Table 2. We re-examined the TriSonica manual and found that the description in the manual itself is also not sufficiently accurate (see the figure below). Following the reviewer's suggestion, we have replaced it with the actual value, "50 g," in the revised manuscript. The revised sentence now reads:

"with a measurement path of 35 mm and a mass of 50 g"

The specific revisions are shown in line 205 of the revised manuscript.


TSM-FB shown

TriSonica™ Mini
Wind & Weather Sensor

The TriSonica™ Mini is compact, lightweight, and efficient. It is small enough to fit in the palm of your hand, yet it is a powerful and highly accurate tool engineered for atmospheric monitoring, weather reporting, and ecosystem research.

Its size and featherweight profile make it perfect for unmanned aerial systems (UAS), while the

The TriSonica Mini Wind & Weather Sensor is a compact (measurement path of just 35 mm), lightweight (less than 50 grams), low velocity anemometer. Even with its small size it provides wind speed, direction, temperature, humidity, pressure, tilt, and compass data. The TriSonica Mini Wind & Weather Sensor can also provide measurements of all three dimensions of air flow. The open path provides the least possible distortion of the wind field. Four measurement paths provide a redundant measurement. The path with the most distortion is removed from the calculations to provide accurate wind measurements. Furthermore, data output can be customized to user requirements.

Available with a pipe-mount base accommodating any 1/2" DN15 Schedule 10 pipe. To further protect components and streamline your installation, wiring runs through the interior of the pipe when using this configuration.

(25) *Line 188: "propeller blade downwash"*

• This is a very good idea and often lacking in other studies. I am glad that the authors have taken this into account.

Response: We sincerely thank the reviewer for this positive and encouraging comment. We are pleased that the reviewer recognizes the importance of accounting for propeller blade downwash, which is indeed often overlooked in related studies. As noted in the manuscript, we carefully considered this effect in the design of the measurement system and in the data processing procedure in order to improve the reliability of the UAV-based observations. We appreciate the reviewer's recognition of this aspect of our work.

(26) *Line 206: "Table 2"*

• Where are these accuracy values from? Are these values provided by the manufacturer or were the sensors laboratory tested for the authors to determine these values independently?

• Why is the carbon dioxide accuracy given in ppb per second? Accuracy is generally independent of time (at short time-scales).

• The terms "T" and "P" have not been defined.

Response: We thank the reviewer for this helpful comment.

Regarding your first point, we agree that the source of the accuracy values in Table 2 should be clearly indicated. In the revised manuscript, we have added a clarification in the table heading to specify that these accuracy specifications are provided by the manufacturers. The revised table title is as follows:

"Table 2. Detailed information of scientific payloads (provided by the manufacturers)."

Regarding your second point, we agree that the original unit given for the CO_2

specification in Table 2 was inappropriate. Following the reviewer's suggestion, we have revised it. The revised text in the manuscript:

"CO_2: < 200 ppb/s (Sensitivity)"

Regarding your third point, we have replaced the abbreviations "T" and "P" with their full terms, "Temperature" and "Pressure".

The specific revisions are shown in line 228 of the revised manuscript.

(27) *Lin 209: "Figure 1"*

• *Which pump does this diagram refer to? Did the carbon dioxide sensor have an in-built pump or was an external plump used? Please provide details on this pump in the main text. If an external pump was used, was it tested to ensure a constant flow rate?*

Response: In our system, the CO_2 analyzer was operated using the analyzer's built-in sampling ports rather than an external pump. We re-examined the manufacturer documentation and found that the model-specific MIRA Ultra N_2O/CO_2 datasheet explicitly states that two programmable sampling ports are built into the analyzer for calibration, re-zeroing, and differential measurements. In addition, the manufacturer's MIRA Ultra series documentation indicates that the series includes a built-in pump. However, the publicly available documentation does not provide a detailed pump flow specification, so we have avoided reporting a flow value not explicitly documented by the manufacturer.

We have therefore clarified in the revised manuscript. The revised sentence now reads:

"CO_2 sampling was performed through the analyzer's built-in sampling ports and internal pump".

The specific revisions are shown in lines 224 to 225 of the revised manuscript.

(28) *Line 242: "40 flights in spring, 55 in summer, 53 in autumn, and 20 in winter"*

• *Do these refer to astronomical or meteorological seasons? It is also worth stating that these refer to Northern Hemispheric seasons.*

Response: We agree that the seasonal definition should be clarified. In the revised manuscript, we have specified that the seasons referred to in this study are based on the meteorological seasons of the Northern Hemisphere. The revised sentence is as follows:

"These included 40 flights in spring, 55 in summer, 53 in autumn, and 20 in winter, defined according to the Northern Hemispheric seasons."

The specific revisions are shown in lines 266 to 267 of the revised manuscript.

(29) *Line 246: "Figure 2"*

• *Are the flight tracks from a single flight (if so, please provide flight details)? Are all UAV flights identical?*

Response: Figure 2a shows the flight tracks from a single representative flight. All UAV flights in this study were conducted following the same predefined flight routes. To avoid ambiguity, we have revised the figure caption accordingly. The revised caption is as follows:

"**Figure 2.** (a) A top-down view of the UAV's flight tracks from a representative flight. The green flight tracks represent the box pattern, categorized into three box sizes: 50×50×50 m^3, 100×100×50 m^3, and 150×150×50 m^3. Site A marks the location of the flux tower where EC measurements of vertical CO_2 fluxes are continuously carried out, while Site B denotes the take-off point for the profile pattern. The straight-line distance between the two sites is 150 m. All UAV flights were conducted following the same predefined flight routes."

The specific revisions are shown in lines 270 to 276 of the revised manuscript.

(30) *Line 257: "Using flight-path fitting and screen interpolation (Han et al., 2024), the three-dimensional spatial distributions of CO2 volume mixing ratio, wind vectors, and air density were reconstructed within the box domain (Fig. 3b)."*

• *Was this interpolated data used during flux calculation or is this purely for visualising. Please make this very clear here and in Section 3. It is incredibly important to know if fluxes are derived from raw measurements or from interpolated measurements.*

Response: We agree that the original wording was not sufficiently clear and could give the misleading impression that the fluxes were calculated from interpolated data. In fact, the interpolated three-dimensional fields were used only for visualization of the spatial distributions of CO_2, wind vectors, and air density within the box domain. They were not involved in the flux calculations. All fluxes reported in this study were derived from the original observations. To avoid ambiguity, we have revised the relevant text here and in Section 3 to make this distinction explicit. The revised text is as follows:

"In the present study, we reconstructed the spatial distributed data for visualization purposes only (Fig. 3b), but computed the horizontal fluxes using the original observations directly. Flight-path fitting refers to the construction of a smooth, closed loop constructed from the discrete UAV trajectories at all altitudes, thereby defining the four virtual vertical surfaces of the box. Screen interpolation refers to the mapping of CO_2 mixing ratios and corrected wind speed measurements observed along the flight path onto a regular two-dimensional grid on the vertical surfaces for visualization of spatial patterns."

The specific revisions are shown in lines 285 to 292 of the revised manuscript.

At the same time, we also emphasized in Section 3 that the flux results were derived from the original observations. The revised text is as follows:

"Finally, the three-dimensional net CO_2 emission was obtained. It should be emphasized that this value was derived from the original observations."

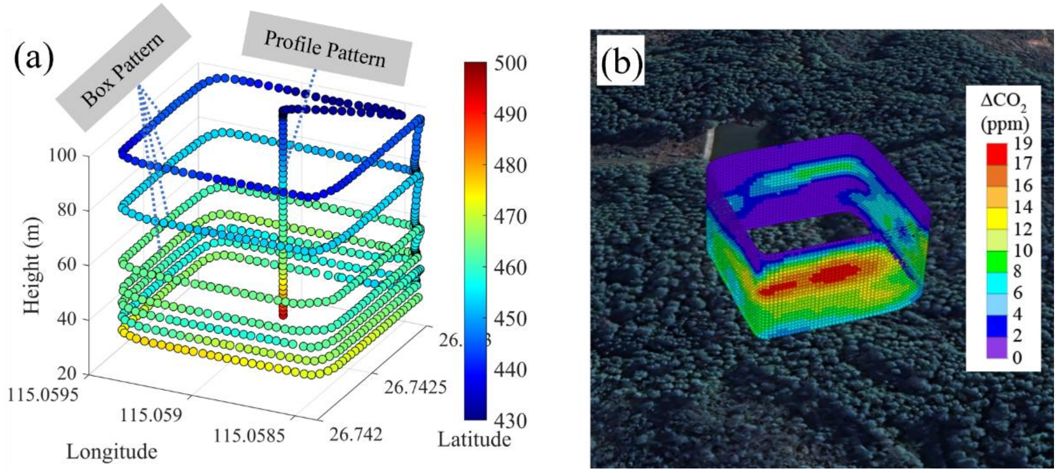
The specific revisions are shown in lines 478 to 479 of the revised manuscript.

(31) *Line 278: "Figure 3"*

• *It is difficult to tell that there are two flight strategies here. After reading the text, I now understand that there is a vertical profile in the middle of the box. Perhaps the two flight strategies can be labelled in this figure. Otherwise it is easy to miss.*

Response: We agree that, in the original figure, the distinction between the two flight strategies was not sufficiently clear and could be easily overlooked. To improve clarity, we have revised Fig. 3a by explicitly labelling "Box Pattern" and "Profile

Pattern” to indicate the corresponding flight paths. We believe this revision makes the two flight strategies much easier to identify in the figure.



The specific revisions are shown in line 303 of the revised manuscript.

(32) *Line 286: “characterize the flow-field distortion by the UAV”*

• Was this characterisation conducted for the specific UAV used in this work, with the finalised UAV configuration including the mounted carbon dioxide sensor and air inlet?

Response: We confirm that the wind speed calibration algorithm was specifically developed for this UAV platform and the mounted instrumentation used in this study. In other words, the characterization of the UAV-induced flow-field distortion was carried out for the specific UAV configuration employed in this work. To make this clear, we have revised the sentence accordingly. The revised sentence is as follows:

“. This algorithm, developed by Yang et al. (2025), was specifically developed for the UAV platform and the mounted instrumentation used in this study, and employs computational fluid dynamics (CFD) simulations to characterize the flow-field distortion by the UAV rotor-induced airflows.”

The specific revisions are shown in lines 311 to 312 of the revised manuscript.

(33) *Line 309: “equals the net outward flux”*

• It may be worth stating here that this integration includes negative flux components. If the wind is blowing into the box horizontally, the flux contribution is negative and if the wind blowing out of the box horizontally, the flux contribution is positive.

Response: We agree that the sign convention for the horizontal flux components should be stated more explicitly. In the revised manuscript, we have clarified that outward flux contributions are defined as positive, whereas inward flux contributions are defined as negative. The revised sentence is as follows:

“Consequently, the total CO_2 emission rate within the volume equals the net outward flux integrated across all box wall surfaces, where positive flux contributions represent CO_2 mass leaving the box and negative flux contributions represent CO_2 mass entering the box.”

The specific revisions are shown in lines 336 to 338 of the revised manuscript.

(34) *Line 324: "corresponding wind fluctuations"*

• *What do these wind fluctuations represent and how were they derived? Do they represent the variance in wind speed over a certain integrating time? Please explain this.*

Response: We agree that the meaning and derivation of the wind fluctuations should be clarified. In the revised manuscript, we now explicitly state that the wind fluctuations (u', v', w') are obtained by subtracting the mean wind components from the corresponding instantaneous wind measurements. The revised sentence is as follows:

"Here, the mean wind vectors ($\bar{u}$, $\bar{v}$, $\bar{w}$) measured with the UAV sonic anemometer over a 30-minute period and the corresponding wind fluctuations (u', v', w'), obtained by subtracting the 30-minute mean wind components from the instantaneous wind measurements, are used to determine fluxes in the three directions."

The specific revisions are shown in lines 353 to 354 of the revised manuscript.

(35) *Line 348: "χ"*

• *Where is this term defined in the manuscript? I can't find it.*

(36) *Line 348 "s"*

• *What is "s"? I can see that "Δs" has been defined, but not "s".*

(37) *Line 348: "$U^{\perp}$"*

• *Where is this term defined in the manuscript? I can't find it.*

Response: We agree that the definitions of χ, s, and the wind component normal to the flight path were not sufficiently explicit in the original manuscript. To improve clarity, we have revised the relevant text to define these quantities directly after the equation. Specifically, $\chi_C(s,z)$ is now defined as the CO_2 mole fraction, s as the path along the flight trajectory, and $u_n(s,z)$ [previously denoted as $U_\perp$] as the wind speed component normal to the flight path at point (s,z). The revised text is as follows:

"where $E_{C,H}$ is the horizontal CO_2 emission, M_R is the ratio of the molar mass of CO_2 to that of air, $\chi_C(s,z)$ is the CO_2 mole fraction, $\rho_{\text{air}}(s,z)$ is the air density, and $u_n(s,z)$ is the wind speed component normal to the flight path at point (s,z). Here, z denotes the flight altitude, s represents the path position along the flight trajectory starting from a predefined fix position on the trajectory path, and (s,z) indicates the point located at the path position s at flight altitude z."

The specific revisions are shown in lines 381 to 386 of the revised manuscript.

(38) *Line 352: "$u^{\perp}$"*

• *Is this lower-case u different to the upper-case U in equation 3.3? In any case, neither has been defined.*

Response: We agree that the notation was inconsistent in the original manuscript. The lower-case $u_\perp$ and upper-case $U_\perp$ were intended to represent the same quantity, but we did not notice this inconsistency during manuscript preparation. To avoid confusion, we have now redefined both symbols uniformly as u_n throughout the manuscript. Here, $u_n(s,z)$ denotes the wind speed component normal to the flight path at point (s, z). We have revised the relevant equations and accompanying text accordingly.

$$F(z_i) = M_R \sum_{j=1}^{N_i} \chi_C(s_j, z_i) \cdot \rho_{air}(s_j, z_i) \cdot u_n(s_j, z_i) \cdot \Delta s_j$$

The specific revisions are shown in line 389 of the revised manuscript.

(39) *Line 359: "Σ"*

• What is the index of this summation? I can see that it spans from 0 to N, but what is summing over? Is it z? This should be placed underneath the sigma symbol.

• Assuming the index of this summation is z, this is apparently summed from a height of 0 m. However, the lowest UAV sampling height is 35 m. So it is actually a summation from 35 m. It is not possible to sum data from a data point that does not exist. Please clarify this.

Response: Regarding your first point, we agree that the summation index and the vertical integration range were not sufficiently clear in the original manuscript. In the revised version, Eq. (3.4) has been reformulated so that the summation is explicitly written over the sampling point index j, where $j = 1, \ldots, N_i$ denotes the discrete sampling points along the flight path at altitude z_i. Likewise, in Eq. (3.5), the vertical integration is now expressed as a summation over the layer index i, where $i = 1, \ldots, n$ represents the discrete observed altitude layers.

Regarding your second point, we also agree that the previous formulation could incorrectly suggest integration starting from 0 m. This has now been corrected. The revised equation does not imply summation from ground level; instead, it integrates only across the actual sampled altitude intervals, beginning at the first observed flight level and continuing to the n-th layer. Therefore, no contribution is assumed below the lowest UAV sampling height. The revised text and equations are now given as follows:

"For each altitude level z_i, the line flux $F(z_i)$ was first calculated from the box-pattern data as

$$F(z_i) = M_R \sum_{j=1}^{N_i} \chi_C(s_j, z_i) \cdot \rho_{air}(s_j, z_i) \cdot u_n(s_j, z_i) \cdot \Delta s_j \tag{3.4}$$

where j denotes the sampling point index along the flight path at altitude z_i, N_i is the total number of sampling points at altitude z_i, and Δs_j is the flight distance represented by the j-th sample interval.

To suppress small-scale variability, a moving average was applied to consecutive line-flux values, yielding layer-averaged fluxes for discrete altitude intervals. These values were then vertically integrated to obtain the total horizontal CO_2 emission rate

$$E_{C,H} = \sum_{i=1}^{n} F(z_i)\Delta z_i = M_R \sum_{i=1}^{n} \left(\sum_{j=1}^{N_i} \chi_C(s_j, z_i)\rho_{\mathrm{air}}(s_j, z_i)u_n(s_j, z_i)\Delta s_j \right) \Delta z_i$$

(3.5)

where i denotes the vertical layer index from the first layer to the n-th layer, n is the total number of vertical layers, and Δz_i is the thickness of the i-th altitude interval. In this study, the first observed altitude layer z_1 corresponds to the lowest UAV sampling height (35 m above ground level)."

The specific revisions are shown in lines 387 to 402 of the revised manuscript.

(40) *Line 359: "un"*

• *What is un? Where is it defined? Is it related to the upper-case U in equation 3.3 and the lower-case u in equation 3.4?*

Response: We agree that the notation in the original manuscript was not sufficiently clear. The upper-case U in Eq. (3.3) and the lower-case u in Eq. (3.4) were intended to represent the same quantity. To avoid confusion, we have now unified the notation throughout the manuscript and consistently denote this term as $u_n(s,z)$, which represents the wind speed component normal to the flight path at point (s,z). The relevant equations and accompanying text have been revised accordingly.

The specific revisions are shown in lines 377 to 402 of the revised manuscript.

(41) *Line 360: "Δz denotes altitude intervals"*

• *I guess that the lowest interval spans from 0 m to 35 m (the lowest sampling height). Therefore, a single measurement at 35 m represents the entire forest plane (with a canopy height of 18 m), with no measurement at ground level. So the UAV sampling misses the entire forest. Please discuss this.*

Response: We apologize for the confusion caused by the non-uniform mathematical notation in the original manuscript. We have also clarified that Δz_i is the thickness of the i-th altitude interval, and added the following statement to specify the lowest sampled layer:

"In this study, the first observed altitude layer z_1 corresponds to the lowest UAV sampling height (35 m above ground level)."

Please see our response to Specific Comment 39 for the detailed revisions.

The specific revisions are shown in lines 400 to 402 of the revised manuscript.

(42) *Line 379 "T and P"*

• *These terms have not been defined.*

Response: We agree that the terms "T" and "P" were not explicitly defined in the original manuscript. To improve clarity, we have replaced "T" and "P" with their full terms, "temperature" and "pressure," in the revised manuscript. The revised text is as follows:

"where K_C denotes turbulent exchange coefficient, $\bar{c}$ was obtained by converting the measured CO_2 mixing ratio to concentration using the concurrently measured temperature and pressure during the UAV profile flight."

The specific revisions are shown in line 421 of the revised manuscript.

(43) *Line 382: "K"*

• *What is "K"? Has it been defined?*

(47) *Line 414: "K"*

• *Are "K" and "KC" the same thing? "K" has not been defined.*

Response: We agree that the notation was inconsistent in the original manuscript and may have caused confusion. Under neutral stratification conditions, the turbulent exchange coefficient (K) and the exchange coefficient for momentum (K_C) were intended to represent the same quantity.

The relevant equations and accompanying definitions have been updated accordingly:

$$\text{“}F_v = -\rho_{air} K_C \frac{\partial \bar{c}}{\partial z} \tag{3.7}$$

where K_C denotes turbulent exchange coefficient, $\bar{c}$ was obtained by converting the measured CO_2 mixing ratio to concentration using the concurrently measured temperature and pressure during the UAV profile flight.

A mixing-length formulation was adopted, in which the eddy diffusivity is expressed as the product of a mixing length and a velocity scale represented here by the friction velocity (Lee and Mahrt, 2005)

$$K_M = l u_* \tag{3.8}$$

where l is the mixing length and K_M is the exchange coefficient for momentum. Under neutral stratification conditions, K_C can be assumed to be equal to K_M, and the friction velocity is defined as"

The specific revisions are shown in lines 418 to 428 of the revised manuscript.

(44) *Line 399: "surface-layer wind profile"*

• *How was this wind profile devised? What is this equation based on? A citation could be useful here.*

Response: We agree that the theoretical basis of Eq. (3.13) should be clarified and supported with an appropriate reference. In the revised manuscript, we now state explicitly that this equation is based on the neutral logarithmic wind-profile formulation from Monin-Obukhov similarity theory, with the inclusion of the zero-plane displacement height and aerodynamic roughness length for rough vegetated surfaces. We have also added citations to support this formulation and the commonly used assumption that the displacement height for forest canopies is on the order of 0.7-0.8 times canopy height. The revised text is as follows:

"Because forest canopies represent aerodynamically rough and structurally complex surfaces, wind profiles above the canopy may deviate from the ideal logarithmic form. Following the neutral logarithmic wind-profile formulation from Monin–Obukhov similarity theory (Raupach, 1994), the surface-layer mean wind profile ($\bar{u}$) was expressed as"

The specific revisions are shown in lines 452 to 456 of the revised manuscript.

The references cited are as follows:

Raupach, M. R.: Simplified expressions for vegetation roughness length and zero-plane displacement as functions of canopy height and area index. Boundary-Layer Meteorology, 71, 211–216, https://doi.org/ 10.1007/BF00709229, 1994.

(45) *Line 409: "z1 and z2"*

• *According to this equation, any two heights can be chosen. So which heights were chosen in this work and why? This important detail is missing from the manuscript. Was this an arbitrary choice?*

(48) *Line 417: "z1 and z2"*

• *So which heights were chosen in this work and why? I can't find this vital information in the manuscript.*

Response: We combined responses to Comments (45) and (48) here.

We agree that the choice of z_1 and z_2 was not clearly described in the original manuscript. When applying the gradient method to estimate fluxes, the vertical separation is typically selected within a range of 1-10 m (Meredith et al., 2014). In this study, we used data collected 5 m below and 5 m above the target height for the calculation. We have revised the manuscript accordingly to make this point explicit. The revised text is as follows:

"When applying the gradient method to estimate the flux, the vertical separation, $\Delta z = z_2 - z_1$, is typically chosen within a range of 1-10 m (Meredith et al., 2014). In this study, $\overline{u_1}$, $\overline{u_2}$ and $\overline{\theta_1}$, $\overline{\theta_2}$ were taken 5 m below and 5 m above any target height ($\Delta z = 10$ m), respectively, during the UAV ascent in the vertical profile measurements."

The specific revisions are shown in lines 447 to 451 of the revised manuscript.

"where $\overline{c_1}$ and $\overline{c_2}$ represent the mean concentrations of CO_2 at heights z_1 and z_2, respectively, converted from measured CO_2 mixing ratios at using the measured temperature and pressure at both heights. In this study, z_1 and z_2 correspond to 5 m below and 5 m above the target height, respectively, during the UAV ascent."

The specific revisions are shown in lines 470 to 473 of the revised manuscript.

The references cited are as follows:

Meredith L. K., Commane R., Munger J. W., Dunn, A., Tang, J., Wofsy, S. C., and Prinn R. G.: Ecosystem fluxes of hydrogen: a comparison of flux-gradient methods, Atmospheric Measurement Techniques, 7(9): 2787-2805, https://doi.org/10.5194/amt-7-2787-2014, 2014.

(46) *Line 413: "Eq. (3.8) yields"*

• *Why not substitute equation 3.12 into 3.15, to eliminate Ri?*

Response: We thank the reviewer for this valuable suggestion.

Although Eq. (3.12) can be substituted into Eq. (3.15) to eliminate R_i, we chose to retain R_i explicitly in the final expression. This is because R_i serves as a physically meaningful indicator of atmospheric stability, and keeping it in the equation makes the role of stability effects in determining K more transparent. In addition, full substitution would lead to a more cumbersome expression without improving the practical implementation of the method.

(49) *Line 436: "The friction velocity (u*) and scalar fluxes (Fv) were computed"*

• *Please provide citations for the source of these equations.*

Response: We agree that the source of these equations should be explicitly provided. In the revised manuscript, we have added citations to standard eddy covariance references to support the formulations used for the friction velocity and scalar flux calculations. The revised text is as follows:

"Fluxes were computed over 30 min averaging intervals, and coordinate rotation, virtual temperature correction, and frequency-response correction were applied to ensure high-quality vertical flux estimates. Following standard eddy covariance formulations (Webb et al., 1980; Baldocchi, 2003), the friction velocity (u_*) and scalar fluxes (F_v) were computed as"

The specific revisions are shown in lines 494 to 495 of the revised manuscript.

The references cited are as follows:

Webb, E. K., Pearman, G. I., & Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer. Quarterly Journal of the Royal Meteorological Society, 106, 85–100, https://doi.org/10.1002/qj.49710644707, 1980.

(50) *Line 445: "open-path eddy covariance analyzer"*

• It is not an Eddy covariance analyser. It is a carbon dioxide gas analyser used by the Eddy covariance system.

Response: We thank the reviewer for taking the time to review and spot the sloppiness in the original manuscript. This is very helpful.

We agree that the original wording was not sufficiently precise. The LI-7500 is not itself an "eddy covariance analyzer," but rather an open-path CO_2/H_2O gas analyzer used as part of the eddy covariance system. To improve accuracy and avoid ambiguity, we have revised the sentence accordingly. The revised text is as follows:

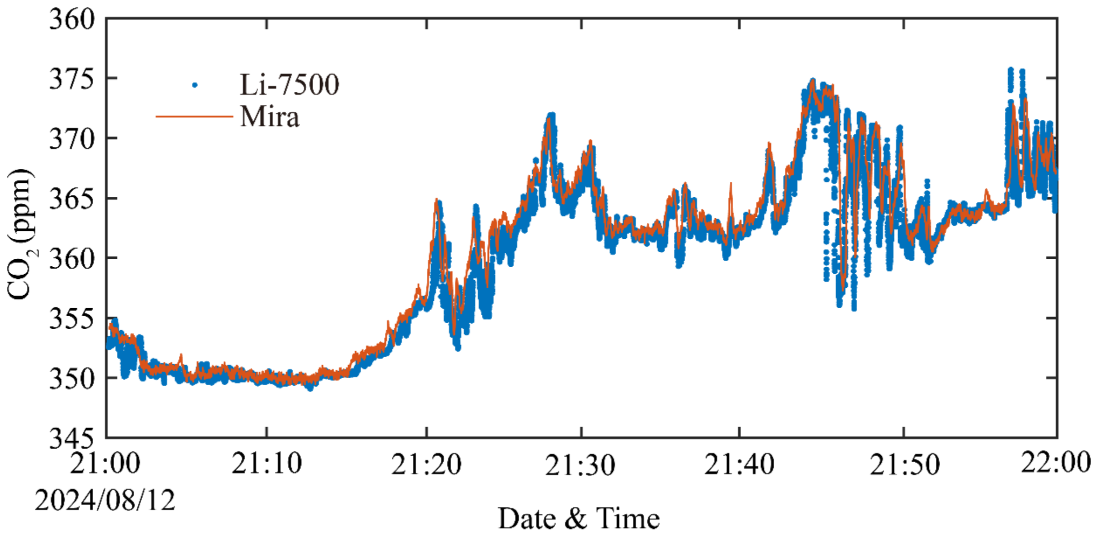
"To evaluate the consistency between different CO_2 analyzers, ground-based intercomparison experiments were conducted using an open-path CO_2/H_2O gas analyzer (LI-7500), which was part of the eddy covariance system, and the closed-path analyzer (MIRA) used on the UAV."

The specific revisions are shown in lines 503 to 506 of the revised manuscript.

(51) *Line 457: "Figure 4"*

• What is the horizontal axis of this plot? What is "No."? It cannot be the number of measurements as the two gas analysers operate at sampling at different frequencies. I would recommend plotting this data against time.

Response: We agree that the original horizontal axis labelled as "No." was not appropriate and could be misleading, especially because the two gas analyzers operated at different sampling frequencies. Following the reviewer's suggestion, we have revised Fig. 4 and now plot the comparison data against date and time on the horizontal axis. This revision provides a clearer and more physically meaningful comparison between the LI-7500 and MIRA measurements. The revised figure is shown below:



The specific revisions are shown in line 513 of the revised manuscript.

(52) *Line 468: "interactions with the sampling line material"*

• *Which material was used for the sampling line?*

• *Was the effect of this sampling line on carbon dioxide measurements tested in the laboratory?*

Response: In this study, the sampling line was a 1/8-inch outer-diameter polytetrafluoroethylene (PTFE) tube. We agree that the original wording was not sufficiently specific, and we have revised the manuscript accordingly to identify the tubing material explicitly. PTFE tubing is commonly used in gas-sampling applications because of its chemical inertness and compatibility with trace-gas measurements, and recent atmospheric measurement studies also report the use of PTFE sampling tubes in closed-path or active gas-sampling systems (Burba et al., 2012).

We did not perform a dedicated laboratory test to isolate the effect of the sampling line on CO_2 measurements in our UAV system. This is mainly because the sampling tube used in this study was very short, approximately 0.5 m. The literature indicates that long intake tubes can introduce substantial low-pass filtering, whereas extremely short intake tubes on the order of 0-0.5 m have a much smaller influence and may behave more similarly to open-path configurations than conventional long-tube closed-path systems (Burba et al., 2012). In addition, previous studies have shown that tubing, filters, and intake assemblies in enclosed systems can attenuate high-frequency CO_2 fluctuations, but the magnitude of this effect increases with sampling-system complexity and tube length (Metzger, et al., 2016).

Accordingly, we have revised the manuscript to avoid overstating the role of tubing-material interactions and now describe the attenuation more cautiously. The revised sentence is as follows:

"This attenuation is primarily attributed to high-frequency losses associated with the closed-path sampling system, including viscous damping and signal smoothing during sample transport through the 1/8-inch outer-diameter polytetrafluoroethylene (PTFE) sampling tube. Because the sampling tube used in this study was relatively short (approximately 0.5 m) and the sample residence time inside the tube was 0.5 s, its influence on CO_2 measurements is expected to be limited."

The specific revisions are shown in lines 525 to 529 of the revised manuscript.

The references cited are as follows:

Burba, G. G., Schmidt, A., Scott, R. L., Nakai, T., Kathilankal, J. C., Fratini, G., Hanson, C.,: Calculating CO2 and H2O eddy covariance fluxes from an enclosed gas analyzer using an instantaneous mixing ratio approach. Global Change Biology, 18, 385–399, https://doi.org/10.1111/j.1365-2486.2011.02536.x, 2012.

Metzger, S., Durden, D., Sturtevant, C., Desai, A. R., Li, J. H.; Luo, H. Y., and Zulueta, R. C.: Optimization of an enclosed gas analyzer sampling system for measuring eddy covariance fluxes of H2O and CO2. Atmospheric Measurement Techniques, 9, 1341–1359, https://doi.org/10.5194/amt-9-1341-2016, 2016.

(53) *Line 485: "coefficient of determination"*

• *Please provide the Pearson correlation coefficient, which is a better metric to evaluate the linear correlation between two datasets. The coefficient of determination is usually better-suited to evaluate the goodness of a model fit.*

(55) *Line 498: "coefficient of determination"*

• *As stated previously, the Pearson correlation coefficient is really required here and not the coefficient of determination. This is not an evaluation of a model fit but rather how one set of data compares against another.*

• *Here the coefficient of determination is a lower-case r2 whereas in the main text it is an upper-case R2 . There are many inconsistencies in this manuscript which are unacceptable in a scientific publication.*

(57) *Line 506: "coefficient of determination"*

• *The correlation coefficient is really required here to draw such conclusions.*

Response: We agree that the Pearson correlation coefficient is more appropriate than the coefficient of determination for evaluating the linear relationship between two observational datasets. Accordingly, we have replaced "coefficient of determination" with "correlation coefficient" throughout the manuscript. In addition, to improve consistency and avoid ambiguity, we have standardized the notation by replacing both R^2 and r^2 with r throughout the revised manuscript. The revised sentence is as follows:

"which is more appropriate than ordinary least squares when neither dataset can be treated as error-free. The season-specific RMA slope, intercept, and Pearson correlation coefficient (r) are summarized in Table 3."

The specific revisions are shown in lines 567 of the revised manuscript.

(54) *Line 489: "Figure 5"*

• *What does SMA mean? This has not been defined.*

• *It is impossible to tell which dataset is plotted on which axis. Please make this clear and make the axis labels bigger.*

• *Why is a linear equation given with "y" as the subject of the equation? "y" has not been defined. It is better to give this equation as one vertical flux in terms of the* *other vertical flux.*

Response: Regarding your first point, SMA and RMA were intended to refer to the same regression method. To avoid confusion, we have removed "SMA" throughout the manuscript and retained only "RMA"

Regarding your second point, we agree that the axis assignments were not sufficiently clear in the original figure. In the revised figure, we have clarified that the x axis represents the flux tower measurements and the y axis represents the UAV-based measurements, and we have enlarged the axis labels accordingly.

Regarding your third point, we agree that using an undefined "y" in the regression equation was inappropriate. We have therefore revised the figure so that the equations are now written using the defined physical variables shown in each panel rather than generic symbols. The revised caption and figure are as follows:

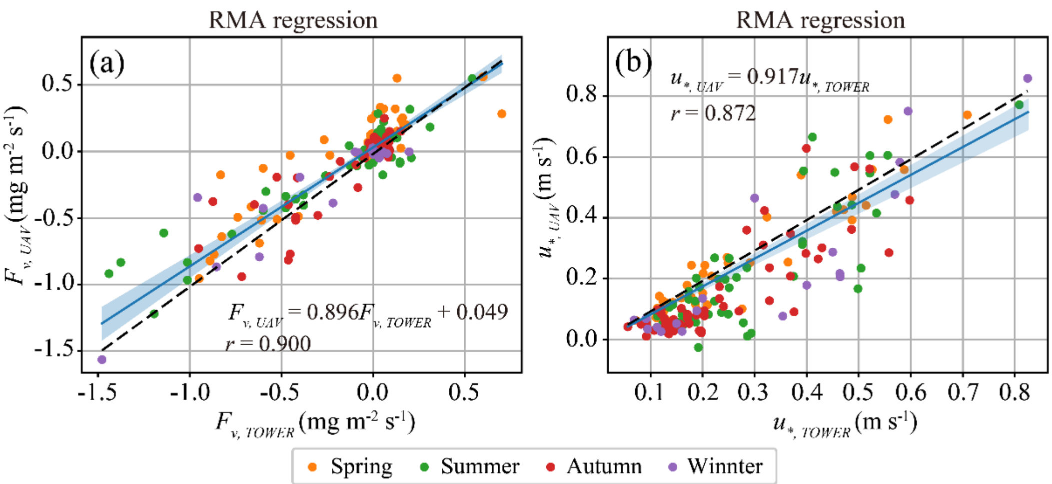


Figure 6. Reduced major axis (RMA) regressions comparing UAV-based measurements with flux tower observations across all seasons. Panel (a) shows CO_2 flux (F_v) and panel (b) shows friction velocity (u_*). In both panels, the x axis represents the flux tower measurements and the y axis represents the UAV-based measurements. Points are colored by season (Spring, Summer, Autumn, Winter). The solid blue line denotes the RMA fit, the shaded band indicates the 95% confidence interval (CI) of the regression line, and the dashed black line represents the 1:1 reference. The regression equations are expressed using the corresponding physical variables in each panel, and the Pearson correlation coefficient (r) is also reported."

The specific revisions are shown in lines 571 to 579 of the revised manuscript.

(56) *Line 500: "yUAV = slope xTOWER + intercept*

• *As stated above, the equation should not be given in terms of undefined "y" and "x" but rather, in terms of vertical flux values.*

Response: We agree that the regression relationship should not be expressed in terms of undefined "x" and "y". In the revised manuscript, we have already modified Fig. 6 so that the regression equations are written using the defined physical variables. For Table 3, because the table includes two different quantities, F_v and u_*, we considered it inappropriate to retain a single generic expression of the form "y_{UAV} =

slope x_{TOWER} + intercept". We have therefore removed this wording from the table description. The revised text is as follows:

"**Table 3.** Seasonal reduced major axis (RMA) regression results comparing UAV-derived measurements with flux tower observations for CO_2 flux (F_v) and friction velocity (u_*). Seasons are defined according to the three-month meteorological seasons: spring (March-May), summer (June-August), autumn (September-November), and winter (December-February). For each season, the table reports the sample size (n), correlation coefficient (r), RMA slope and its 95% confidence interval (CI), and the intercept. The u_* regressions are constrained to pass through the origin (intercept fixed at 0)."

The specific revisions are shown in lines 584 to 587 of the revised manuscript.

(58) *Line 533: "UAV platform reproduces"*

• *The UAV platform does not reproduce anything, but rather the flux derived from the measurements made on-board the UAV platform.*

Response: We agree that the original wording was not sufficiently precise. The UAV platform itself does not "reproduce" the observed variability; rather, it is the fluxes and turbulence metrics derived from measurements collected on board the UAV platform that reproduce the variability observed by the tower. We have revised the sentence accordingly. The revised text is as follows:

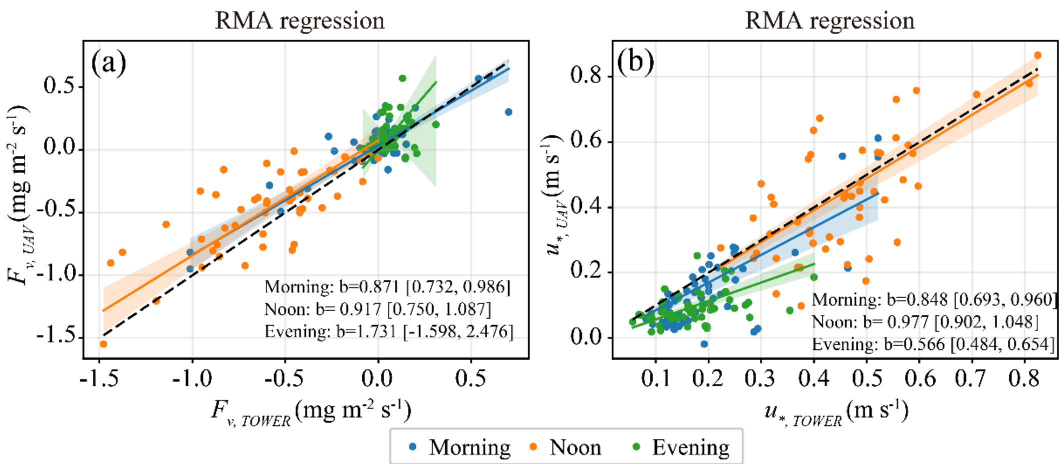
"Taken together, Fig. 6 and Table 3 demonstrate that the fluxes and turbulence metrics derived from the UAV measurements reproduce well the tower-observed variability of both canopy-scale CO_2 exchange and key turbulence characteristics across all seasons (r=0.900 for F_v and r=0.872 for u_*)"

The specific revisions are shown in lines 619 to 620 of the revised manuscript.

(59) *Line 551: "Figure 6"*

• *As for Figure 5, the axis labels here are very difficult to read.*

Response: We agree that the axis labels in the original version of Fig. 6 were difficult to read. In the revised manuscript, we have enlarged the font size of the axis labels in Fig. 7 to improve the clarity and readability of the figure.



The specific revisions are shown in line 637 of the revised manuscript.

(60) *Line 580: "low coefficient of determination"*

• *I would be interested to know what the correlation coefficient is and if the vertical turbulent flux correlation is really that bad.*

Response: We agree that the correlation coefficient is the more appropriate metric in this context. Accordingly, we have replaced the coefficient of determination with the correlation coefficient in the revised manuscript. The revised sentence is as follows:

"In contrast, the agreement for F_v deteriorates substantially during the evening period, with a low correlation coefficient (r=0.235, n=54) and a highly uncertain slope (1.730, with a 95% confidence interval spanning both positive and negative values)."

The specific revisions are shown in line 667 of the revised manuscript.

(61) *Line 589: "as radiative forcing weakens and a stable stratification develops, turbulence above the canopy diminishes"*

• *Please provide some citations to support this assertion.*

Response: We agree that this statement should be supported by appropriate references. In the revised manuscript, we have added citations to studies showing that stable stratification suppresses turbulent mixing and reduces vertical scalar transport above and within forest canopies, particularly during the late-afternoon-to-evening transition. The revised sentence is as follows:

"In the late afternoon to evening, as radiative forcing weakens and a stable stratification develops, turbulence above the canopy diminishes, leading to reduced vertical fluxes (Wharton et al., 2017; Kiefer and Zhong, 2013)."

The specific revisions are shown in *SI Appendix* S3, line 97 of the revised manuscript.

The references cited are as follows:

Wharton, S., Ma, S., Baldocchi, D. D., Falk, M., Newman, J. F., Osuna, J. L., and Bible, K.: Influence of regional nighttime atmospheric regimes on canopy turbulence and gradients at a closed and open forest in mountain-valley terrain, Agricultural and Forest Meteorology, 237–238, 18–29, https://doi.org/10.1016/j.agrformet.2017.01.020, 2017.

Kiefer, M. T., and Zhong, S.: The effect of sidewall forest canopies on the formation of cold-air pools: A numerical study, Journal of Geophysical Research: Atmospheres, 118(12), 5965–5978, https://doi.org/10.1002/jgrd.50509, 2013.

(62) *Line 597: "seasonal distributions of nighttime CO2 flux differences"*

• *How many data points are there for each season during the night?*

• *What is the definition of night in the manuscript? What hours of the day does this correspond to?*

• *What is the motivation for looking at night-time fluxes independently? Please provide some justification here.*

Response: Regarding your first point, the data used in this analysis correspond to the "evening" period reported in Table 4. The number of data points for each season is

12 in spring, 15 in summer, 20 in autumn, and 7 in winter. We are very sorry that this was not explained in the original manuscript. We have included this information in the caption of Figure 7. The revised sentence is as follows:

Figure S1. Boxplots of evening differences in CO_2 fluxes (ΔF_v) between UAV and flux tower observations across four seasons. The boxes indicate the interquartile range with median values shown as horizontal lines, small circles represent outliers and whiskers represent 1.5 times the interquartile range. The number of data points in this figure is 12 for spring, 15 for summer, 20 for autumn, and 7 for winter.

The specific revisions are shown in *SI Appendix* S3, lines 134 to 138 of the revised manuscript.

Regarding your second point, we agree that the term "nighttime" was incorrectly used in this part of the manuscript and may have caused confusion. In fact, the analysis was based on the "evening" data reported in Table 4, corresponding to the 19:00-21:00 time period (see also the legend in Fig. 7). To avoid ambiguity, we have replaced "nighttime" with "evening" throughout the relevant text.

Regarding your third point, the motivation for analyzing the evening fluxes separately is that the correlation between UAV-derived fluxes and tower-based fluxes was relatively low during this period. However, a low correlation coefficient alone does not fully characterize the relationship between the two datasets. Therefore, we further examined their differences, defined as $\Delta F_v = F_{v,\text{UAV}} - F_{v,\text{Tower}}$, and performed an ANOVA analysis to assess whether these differences varied significantly among seasons. This approach is also suitable given the relatively limited sample sizes. Overall, the seasonal analysis of ΔF_v shows that the differences are relatively small and do not exhibit a clear seasonal bias. This indicates that, despite the low correlation during the evening period, the CO_2 fluxes derived from the UAV platform are not statistically different from the tower-based measurements.

The revised text is as follows:

"Due to the correlation coefficient between the fluxes calculated by the flux tower and those derived from the UAV was very low during the evening, and a low correlation coefficient alone cannot fully explain the relationship between the two datasets, an additional analysis of the flux differences was therefore carried out. Figure S1 presents the seasonal distributions of evening CO_2 flux differences"

The specific revisions are shown in *SI Appendix* S3, lines 104 to 108 of the revised manuscript.

(63) *Line 600: "0.1"*

• *What are the units for this value?*

Response: We agree that the unit should be explicitly provided here. In the revised manuscript, we have clarified that the value refers to CO_2 flux differences expressed in units of mg m^{-2} s^{-1}. The revised text is as follows:

"The mean values are small, with absolute values generally smaller than 0.1 mg m^{-2} s^{-1}, indicating that…"

The specific revisions are shown in *SI Appendix* S3, line 111 of the revised manuscript.

(64) *Line 604: "F=2.50"*

• *What is F? Has it been defined? Maybe it is related to ANOVA; if so, some background should be provided here.*

(65) *Line 606: "p"*

• *What is p?*

Response: To clarify, the F-statistic in ANOVA measures the ratio of variability between the groups (seasons) to the variability within the groups (random fluctuation). A higher F-value indicates more variability between groups compared to within groups. In our analysis, the F =2.50 suggests that the variability between seasons is 2.50 times larger than the variability within seasons.

The p-value represents the probability that the observed differences are due to random variation. A p-value greater than 0.05, as in our case ($p = 0.0699$), indicates that the observed seasonal differences are not statistically significant, meaning they could be due to random fluctuation rather than a true seasonal effect.

We have revised the manuscript to include this explanation for clarity, and the revised sentence is as follows:

"Specifically, the ANOVA yielded an F-statistic of 2.50, which indicates that the variability between seasons is 2.50 times greater than the variability within seasons. However, the p-value of 0.0699, which is greater than 0.05, indicates that these seasonal differences are not statistically significant, meaning that the observed differences could be due to random variation rather than a true seasonal effect."

The specific revisions are shown in *SI Appendix* S3, lines 115 to 119 of the revised manuscript.

(66) *Line 632: "reflecting reduced turbulent mixing"*

• *Why does a positive flux indicate reduced turbulent mixing?*

Response: We agree that a positive flux does not by itself demonstrate reduced turbulent mixing, and that the original wording was therefore not sufficiently precise. In the revised manuscript, we have clarified that morning and evening fluxes being close to neutral or weakly positive primarily reflect reduced net ecosystem CO_2 uptake/emissions under lower photosynthetic/respirational activity. The revised text is as follows:

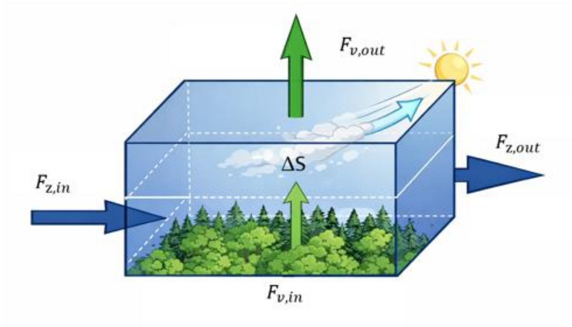
"In contrast, morning and evening fluxes were much reduced from mid-day peaks, probably resulting from neutral or weakly positive turbulent mixing as well as from combined effect of weaker uptake/emissions of CO_2 under lower photosynthetic/respirational activities."

The specific revisions are shown in lines 691 to 694 of the revised manuscript.

(67) *Line 659: "height=50m"*

• *According to equation 3.5, the boxes are integrated across all heights. So I don't understand why a fixed height of 50 m is being used here. This is a very important point and perhaps represents some missing information in the methodology.*

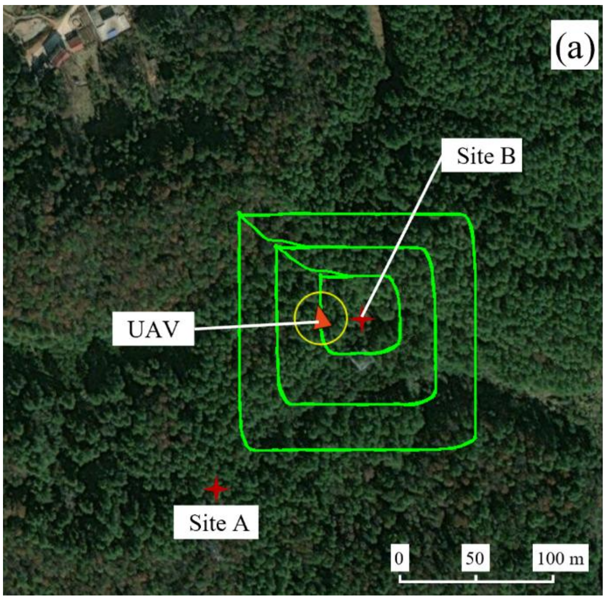
Response: We are very sorry that our wording in the original manuscript caused a misunderstanding. Our intention was not to indicate that the calculation was performed at a fixed height of 50 m. Rather, we intended to describe three different box-shaped control volumes with dimensions of 50×50×50 m^3, 100×100×50 m^3, and 150×150×50 m^3. The ideal shape of the control volume is shown in the figure below.



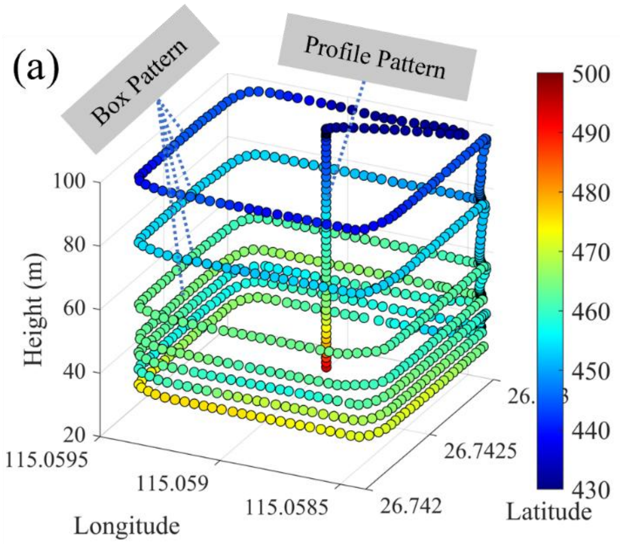
The three control volumes differ in their horizontal dimensions, but all have the same vertical extent of 50 m.

Below, we provide a more detailed explanation to help the reviewer better understand this point.

Within a given time window (20 min), we flew three three-dimensional control volumes of different sizes. As shown in the top-view figure below, three different horizontal dimensions were covered: 50×50 m^2, 100×100 m^2, and150×150 m^2.

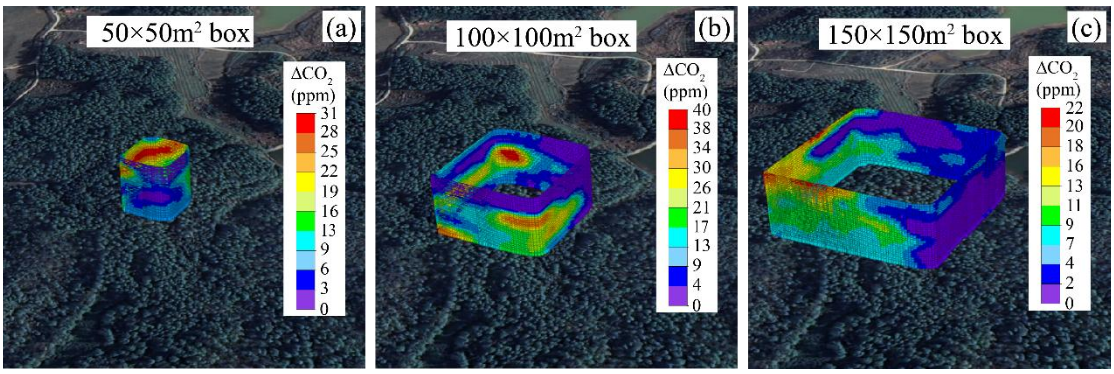


For each control volume, measurements were conducted at heights of 35, 40, 45, 50, 60, and 85 m (as shown in the figure below). Therefore, the height of the control volume was defined as 50 m, obtained by subtracting the lowest measurement height from the highest one.



This resulted in three control volumes with dimensions of 50×50×50 m³, 100×100×50 m³, and 150×150×50 m³, which were then used for the subsequent analysis of spatial heterogeneity.

In addition, we attempted to present the three-dimensional distribution of CO_2 more intuitively. However, conducting multiple vertical profile flights in a forested area was extremely difficult because no suitable forest opening was available for UAV take-off. As a result, we were unable to fully characterize the vertical distribution of CO_2 using multiple profile measurements, and only one vertical profile was available. Therefore, we could only present the horizontal distribution of CO_2, and the CO_2 distributions within the three control volumes of different sizes are shown in the figure below.



In summary, our spatial heterogeneity analysis was conducted for these three different control volumes and included both the horizontal and vertical emission components.

To avoid ambiguity, we have revised the sentence accordingly. The revised text is as follows:

"To verify whether net emissions (E_C) calculated using the mass-balance approach differ across spatial scales, emission computations were performed for three box-shaped control volumes constructed from the UAV flight tracks, with horizontal dimensions of 50×50 m^2, 100×100 m^2, and 150×150 m^2, respectively, and a common vertical extent of 50 m."

The specific revisions are shown in lines 761 to 763 of the revised manuscript.

(68) *Line 676: "This result may suggest relatively weak spatial heterogeneity during the summer-noon period."*

• *This is probably because sampling at a fixed height of 50 m is used, when the tree canopy only reaches up to 18 m. The spatial heterogeneity of the forest may be lost at this height. So this entire analysis seems a bit pointless to me.*

Response: According to our response to Specific Comment 67, the comparison in this section was not conducted at a fixed height of 50 m. Rather, we compared whether the net CO_2 emissions differed among three control volumes of different sizes, namely 50×50×50 m^3, 100×100×50 m^3, and 150×150×50 m^3, in order to evaluate the spatial heterogeneity of CO_2 emissions. We are very sorry that our previous wording may have caused the misunderstanding that this analysis was based only on measurements at 50 m height.

Regarding the reviewer's concern that this entire analysis seems a bit pointless. we believe that this type of spatial heterogeneity analysis is both interesting and necessary. The currently dominant eddy covariance method measures CO_2 flux at a single point and cannot directly determine the true spatial extent represented by that measurement. In practice, this representativeness is often approximated using footprint analysis, and one common rule of thumb is that the representative distance is on the order of 5-10 times the tower height, typically about 200-400 m. However, fluxes outside that representative range generally remain unknown. At the same time, it is widely recognized that CO_2 exchange in complex ecosystems such as forests is spatially heterogeneous, yet this heterogeneity has rarely been quantified directly. The UAV-based method developed in this study provides a way to quantify the variability in CO_2 emissions associated with spatial heterogeneity. Because quantifying spatial heterogeneity is not the main focus of the present paper, which is primarily intended to establish and demonstrate the measurement methodology, we have not expanded this analysis in great detail here. A brief outline of the quantification approach is provided in our response to Comment 7. Nevertheless, we consider it necessary to include this spatial heterogeneity analysis, because otherwise an important advantage of the UAV platform—its flexibility and mobility for spatially distributed measurements—would not be adequately demonstrated.

Regarding the reviewer's concern that the lowest altitude of the UAV box-model measurements was 35 m while the canopy height was only 18 m, which may result in missing a large fraction of the emissions or uptake, we have provided an explanation for this comment, please refer to our response to Comment 4 in the General Comments section.

(69) *Line 678: "F"*

• *What is F? Has it been defined?*

Response: We agree that the meanings of F and p should be explained more explicitly. In this analysis, $F = 0.349$ indicates that the variability among the three scales was smaller than the variability within each scale group, suggesting that the scale effect was weak. The corresponding p-value of 0.710 is much larger than 0.05, indicating that the observed differences among scales were not statistically significant and could be explained by random variation rather than a true difference among the three spatial scales. We have clarified this interpretation in the revised manuscript. The revised sentence is as follows:

"Furthermore, one-way ANOVA shows the mean differences in e_C among the three scales were not significant, with an ANOVA F-statistic of 0.349, indicating low between-group variance relative to within-group variance, and a p-value of 0.710, indicating that the observed differences could be readily explained by random variation rather than a true scale effect."

The specific revisions are shown in lines 790 to 793 of the revised manuscript.

(70) *Line 701: "EC,V"*

• *Should this be EC,VT? I thought the vertical advective flux component is negligible and only the vertical turbulent flux component is considered.*

Response: We agree that, in this context, the vertical term should refer specifically to the vertical turbulent flux component rather than a general vertical exchange term. Accordingly, in the revised manuscript, we have replaced $E_{C,V}$ with $E_{C,VT}$ to make this distinction explicit and to ensure consistency with the methodological description that the vertical advective component is assumed negligible. The revised text is as follows:

"For each box size, n denotes the number of valid samples. Reported metrics include the mean and median of E_C (kg h^{-1}), the interquartile range (IQR), the median and standard deviation (SD) of the horizontal-transport contribution ($E_{C,H}$) and the vertical-exchange contribution ($E_{C,VT}$), as well as the median ratio $E_{C,H}/E_{C,VT}$. All emission estimates are expressed in kg h^{-1}."

The specific revisions are shown in line 819 of the revised manuscript.

(71) *Line 705: "The mean values in Fig.10 indicate that the emission intensities (e_C) derived for different box sizes (0.02, 0.21, and 0.11, respectively) differ to some extent."*

• *This is probably because the exact same set of UAV flights are not used for the different box sizes. It is impossible to compare them directly, if they do not represent identical sampling.*

Response: We have provided an explanation for this comment, please refer to our response to Comment 10 in the General Comments section.

(72) *Line 714: "3.214 for 50×50, 0.847 for 100×100, and 1.213 for 150×150"*

• *The EC,H and EC,VT mean (or median) values and standard deviation values also need to be provided separately in Table 5. Perhaps autumn morning FC,VT fluxes are very close to zero, amplifying this horizontal flux dominance.*

• *As mentioned above, it is difficult to make direct comparisons, if different sets of UAV flights were used for each box size.*

Response: Regarding your first point, we have revised Table 5 and provided an explanation for this comment. Please refer to our response to Comment 11 in the

General Comments section.

Regarding your second point, we have provided an explanation for this comment, please refer to our response to Comment 10 in the General Comments section.

(73) *Line 715: "These results suggest that there is a significant spatial heterogeneity in the CO2 emissions/uptake in the range covered by the different boxes"*

• It is difficult to draw these sorts of conclusions with only seven or eight data points.

Response: We would like to explain that the limited sample size was a direct consequence of our effort to ensure that the comparison was made only under highly similar observational conditions. Although we collected a total of 168 UAV flights, we considered the influence of boundary-layer conditions on flux measurements and therefore restricted the analysis to morning, noon, and evening periods. We also considered the seasonal dependence of CO_2 exchange in the forest ecosystem and therefore further separated the data into spring, summer, autumn, and winter. Ultimately, the datasets suitable for comparison were limited to summer-noon and autumn-morning cases. In addition, only those cases in which the three different box sizes were measured within a short time window could be retained, which resulted in seven valid datasets.

For a more detailed explanation, please refer to our response to Comment 10 in the General Comments section.

(74) *Line 726: "EC,V"*

• According to section 3, EC,V is negligible and only EC,VT needs to be considered. So I am a little lost with the terminology here.

Response: Thank you for pointing this out.

We agree that the terminology in the original text was inconsistent with Section 3. As described in Section 3. To avoid confusion and maintain consistency throughout the manuscript, we have revised "$E_{C,V}$ (vertical transport)" to "$E_{C,VT}$ (vertical turbulent transport)" in the revised manuscript:

"This scale sensitivity further highlights the true advantage of UAV observations for characterizing spatial heterogeneity. UAV measurements can provide multipoint, three-dimensional CO_2 distributions and associated wind speed gradient information within a short time window, enabling the simultaneous estimation of $E_{C,H}$ (horizontal transport) and $E_{C,VT}$ (vertical transport) within a consistent data framework."

The specific revisions are shown in line 856 of the revised manuscript.

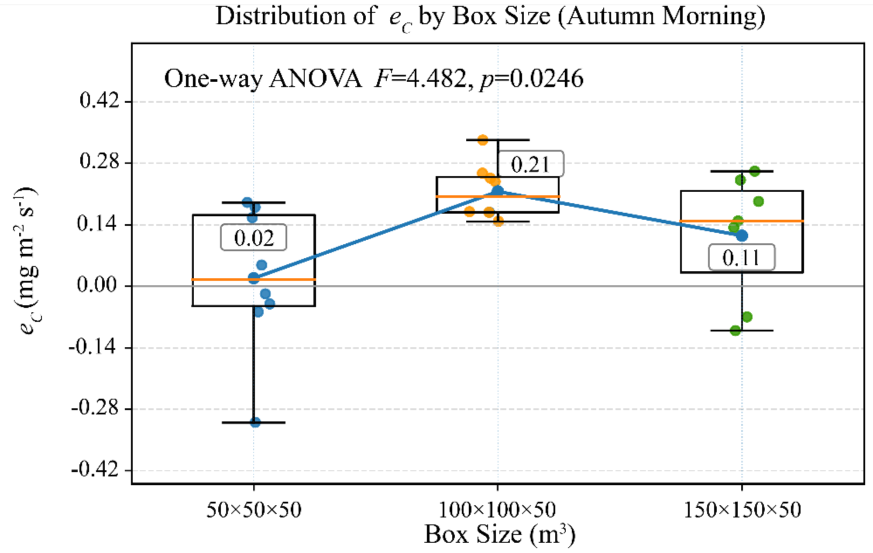
(75) *Line 739: "Figure 10"*

• The closing bracket is missing for the units on the vertical axis.

• What are F and p?

Response: In the revised figure, the missing closing bracket in the unit label on the vertical axis has been corrected. In addition, to improve clarity, we have explicitly clarified that Fdenotes the F-statistic from the one-way ANOVA and $p$denotes the corresponding p-value. These revisions have been incorporated into the revised manuscript and figure.

The revised sentence is as follows:



"**Figure 11.** Distribution of e_C across box sizes (50×50×50 m^3, 100×100×50 m^3, 150×150×50 m^3) of autumn morning. Boxes represent the interquartile range (IQR), the center line is the median, and whiskers extend to 1.5×IQR. Colored points are individual samples. The blue solid polyline connects the mean e_C of each group, with boxed numerical annotations. The reported F and p values denote the F-statistic and corresponding p-value from the one-way ANOVA, respectively."

The specific revisions are shown in lines 872 to 874 of the revised manuscript.

(76) *Line 753: "Table 6"*

• *I am not sure that this is the best way to represent uncertainties. It seems like twelve emission estimates have been cherry-picked out of the full dataset. Why not just give the average percentage uncertainty for each season, rather than picking a single day?*

Response: We agree that the previous presentation of uncertainty in Table 6, based on a limited number of individual emission estimates, may not have been the most representative way to summarize the uncertainty characteristics of the full dataset and could give the impression of selective sampling.

In response, we have revised Table 6 so that it now presents the mean percentage uncertainty for each season instead of values from selected single days. We believe this revised presentation provides a more robust and representative summary of seasonal uncertainty.

Table 6. Mean uncertainties (%) of CO_2 emissions derived from the UAV for four seasons. Columns labeled M, N, and E correspond to morning, noon, and evening periods, respectively. The total mean uncertainty δ is obtained by combining the individual contributions in quadrature for each period.

	Spring			Summer			Autumn			Winter		
	M	N	E	M	N	E	M	N	E	M	N	E
δ_M	2	1	1	2	2	3	2	2	1	2	2	1
δ_V	8	18	11	19	13	12	18	15	18	9	12	14

δ_{dens}	0	1	0	1	2	1	1	2	1	1	1	0
δ	8	18	11	19	13	12	18	15	18	9	12	14

The specific revisions are shown in lines 885 to 888 of the revised manuscript.

(77) *Line 764: "better than 0.2 ppm"*

• *According to Table 2, the accuracy (not the precision) is better than 0.2 ppm. Please clarify whether this refers to accuracy or precision.*

Response: We agree that "better than 0.2 ppm" in Table 2 refers to the accuracy of the closed-path CO_2 analyzer rather than its precision. To avoid confusion, we have revised the sentence accordingly in the manuscript:

"The closed-path CO_2 analyzer has an accuracy better than 0.2 ppm"

The specific revisions are shown in line 895 of the revised manuscript.

(78) *Line 768: "uncertainty on the order of δM ≈5%"*

• *How can the uncertainty be approximately equal to the order of 5%? This doesn't really make sense.*

• *Table 6 shows that the uncertainty is only 5% for one of the twelve fluxes. Most of the time it is much lower than 5%.*

Response: We agree that the original wording, "on the order of δM ≈ 5%," was imprecise and potentially misleading. As shown in Table 6, the instrumental contribution to the emission rate uncertainty is generally below 5% and reaches about 5% only in one case. We have therefore revised the sentence to describe this range more accurately and to avoid implying that 5% is representative of all flux estimates:

"Propagating these errors through the flux equations indicates that the instrumental contribution to the emission rate uncertainty is generally below 5%, with a maximum of about 5% uncertainty shown in Table 6. This contribution is therefore considered minor compared to sampling and methodological effects."

The specific revisions are shown in lines 898 to 900 of the revised manuscript.

(79) *Line 771: "4.4.4 Wind-calibration uncertainties"*

• *It seems strange that wind calibration uncertainties are treated separately to wind measurement uncertainties. It would make more sense to treat wind uncertainties as one single term and carbon dioxide uncertainties as a separate term. Please discuss the rationale for the approach used here.*

Response: We agree that treating wind-calibration uncertainty separately from wind measurement uncertainty may be unclear. Our original intention was to highlight the uncertainty contribution specifically introduced by the wind calibration procedure. However, we agree that it is more logical and clearer to treat this as one component of the broader wind measurement uncertainty, while considering CO_2-related uncertainty separately. Accordingly, we have revised the subsection title and clarified the text in the manuscript:

"4.4.2 Wind measurement uncertainties

Uncertainty associated with wind measurements includes uncertainties arising from sensor performance and interferences from UAV platform geometry-induced flow distortion, UAV attitude changes during flight, as well as UAV flight speeds, and the subsequent wind-correction procedure that was used to correct for these interferences. The wind-correction algorithm, derived from extensive CFD simulations and in-flight validation (Yang et al., 2025), corrects for rotor downwash, UAV attitude changes, and background wind over a wide flight envelope conditions. Residual biases are most likely during periods of strong turbulence or rapid maneuvers, but the close agreement between UAV-derived friction velocity and flux tower estimates (Fig. 6b) suggests that any remaining systematic error in wind speed is modest. Table 6 shows the mean uncertainty for the E_c term associated with wind measurements is estimated at 0.15 kg h^{-1} (8% of E_c, minimal impact) for the spring morning flight and 0.73 kg h^{-1} (22%, maximum impact) for the autumn noon flight. A conservative δ_V ≈8-22% is assigned to this uncertainty term."

The specific revisions are shown in lines 903 to 917 of the revised manuscript.

(80) *Line 792: "can be considered negligible"*

• *But they are not treated as negligible as they are included in the error propagation equation. So I don't understand this point.*

Response: We agree that the phrase "can be considered negligible" was not appropriate, because this uncertainty term is still included in the overall error propagation. Our intention was to indicate that its contribution is small relative to the other uncertainty sources, rather than that it was omitted from the calculation. We have revised the text accordingly to clarify that the uncertainty associated with temperature- and pressure-induced density variations is minor, but still included in the propagated uncertainty estimate:

"The horizontal emission term in Eq. (3.4) depends explicitly on air density, which is determined by air temperature and pressure. In this study, the time-averaged temperature and pressure over the duration of each box-profile pair were used to derive a representative air density for the corresponding control volume. To assess the impact of temporal variability in the atmospheric thermodynamic conditions, the emission calculations were repeated using the minimum and maximum air densities observed during each flight period as alternative scenarios. As shown in Table 6, the resulting uncertainty associated with density variations is small (δ_{dens} ≈1-2%) and therefore represents only a minor contribution relative to the other uncertainty terms."

The specific revisions are shown in lines 920 to 928 of the revised manuscript.

(81) *Line 803: "CO2 enhancement field"*

• *What are the enhancements above? Are they enhanced above some sort of background? How is this background determined?*

• *Enhancements above a background are only useful assuming a positive emission flux. A negative emission flux characteristic of dominant photosynthetic activity would be better represented as mole fraction diminutions.*

Response: We agree that the phrase "CO_2 enhancement field" is potentially misleading because it emphasizes only positive departures from a background concentration. In fact, when photosynthetic uptake dominates, negative departures are more appropriately interpreted as CO_2 mole fraction diminutions relative to the background.

To address this issue, we have replaced "CO_2 enhancement field" with "CO_2 mole fraction anomaly field (ΔCO_2)" in the revised manuscript. In our study, ΔCO_2 is calculated relative to a constant background value defined as the mean of the lowest 5% of the time-series CO_2 observations.

"Figure 9 provides a three-dimensional visualization of the CO_2 mole fraction anomaly field (ΔCO_2), calculated relative to a constant background value defined as the mean of the lowest 5% of the time-series CO_2 observations, along the walls of the $100\times100\times50$ m^3 control volume during the morning flight on 8 August 2024."

The specific revisions are shown in lines 718 to 721 of the revised manuscript.

(82) *Line 815: "Figure 11"*

• Please remind the reader that this interpolated data is derived following the procedure described in Section 2.3 and that this figure does not show raw mole fraction sampling. This data has been processed.

Response: We agree that the caption should more clearly indicate that Figure 12 does not present raw mole fraction sampling, but rather a processed and interpolated field derived from the UAV observations following the procedure described in Section 2.3.

To address this point, we have revised the caption to explicitly state that the figure shows the three-dimensional CO_2 mole fraction anomaly field reconstructed from processed UAV measurements using the interpolation procedure described in Section 2.3. We have also clarified that ΔCO_2 is defined relative to a constant background value, prescribed uniformly for the analyzed flight period:

"**Figure 9.** Three-dimensional distribution of the CO_2 mole fraction anomaly field (ΔCO_2), obtained from processed and interpolated UAV measurements following the procedure described in Section 2.3, within the $100\times100\times50$ m^3 domain during the morning flight on 8 August 2024, relative to a constant background value."

The specific revisions are shown in lines 732 to 735 of the revised manuscript.

(83) *Line 848: "excellent agreement"*

• This is a strong phrase Consider rewording this.

• Pearson correlation coefficient is required to evaluate such agreement.

Response: We agree that the phrase "excellent agreement" is too strong and should be expressed more cautiously. We have therefore revised the sentence to use a more objective description of the relationship between the two datasets.

In addition, following your suggestion, we now report the Pearson correlation coefficient to quantify the level of agreement. The revised text states that the UAV-derived vertical CO_2 fluxes are strongly correlated with the collocated flux tower observations, with $r = 0.900$:

"Vertical CO_2 fluxes derived from UAV vertical profiles are in good agreement with those determined from collocated eddy covariance (EC) flux tower observations across seasons, with a Pearson correlation coefficient of $r= 0.900$, indicating that the

UAV-based gradient method reproduces both the magnitude and variability of vertical turbulent exchange under typical daytime conditions."

The specific revisions are shown in lines 950 to 953 of the revised manuscript.

(84) *Line 854: "are ecologically consistent with long-term site behavior"*

• How is this shown from UAV fluxes? Where is the long-term site behaviour presented in this manuscript? Perhaps a time-series should be shown to make this point.

Response: Thank you for this helpful comment.

We agree that the statement "are ecologically consistent with long-term site behavior" was not sufficiently supported in the manuscript, because long-term site behavior was not explicitly presented with corresponding time-series data.

To address this concern, we have revised the sentence to remove the reference to long-term site behavior and to focus only on the diurnal and seasonal patterns directly shown by the UAV-derived fluxes. The revised sentence now reads:

"The UAV-based vertical fluxes showed clear diurnal and seasonal patterns over the forest ecosystem at Qianyanzhou, with the strongest uptake around midday during the growing season and near-neutral exchange in winter."

The specific revisions are shown in lines 955 to 956 of the revised manuscript.

(85) *Line 859: "dominant term under certain conditions"*

• Horizontal transport may be dominant when vertical turbulent fluxes are naturally small. It is important to evaluate if the horizontal flux magnitude is actually that large in absolute terms, instead of only taking the ratio between the two flux components.

Response: We have added both $E_{C,H}$ and $E_{C,VT}$ to Table 5 and reanalyzed their relative and absolute magnitudes. We also revised the text to avoid overstating horizontal transport as a "dominant term" based solely on the ratio between the two components. The detailed response to this issue is provided in our response to General Comment 11.

(86) *Line 865: "reflecting the spatial heterogeneity of sources/sinks and the changing relative contribution of horizontal transport with control-volume size"*

• This spatial heterogeneity may be amplified if the vertical turbulent flux component is naturally small at certain times of day. It may be totally insignificant if the vertical turbulent flux component is larger. This should be discussed and clarified.

Response: We do not fully understand why spatial heterogeneity may be amplified when the vertical turbulent flux is naturally small; therefore, we have deleted the phrase "and the changing relative contribution of horizontal transport with control-volume size"

The revised text is as follows:

"A multi-scale analysis using different box sizes (50×50×50, 100×100×50, and 150×150×50 m^3) shows that net emission estimates can exhibit scale dependence, reflecting the spatial heterogeneity of sources/sinks."

The specific revisions are shown in line 967 of the revised manuscript.