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 #2:

General comments

(1) *This is an interesting paper that takes a few steps in the right direction of where surface-atmosphere exchange work can go with the rapidly increasing availability of UAVs that can carry the relevant sensors. The paper presents data from field work that compared CO2 flux data from a stationary eddy covariance flux tower in a southern Chinese coniferous forest with fluxes estimated from vertical profile and box flights above this tower with a UAV carrying CO2 and turbulence instrumentation. 167 flights over a year-and-a-half were analyzed, split into seasons and time of day. Boxes of different horizontal sizes were flown in an attempt to evaluate heterogeneity and advective effects. The derived vertical fluxes appear to agree well, but the analysis of horizontal advection seems a bit more inconclusive.*

Response: We sincerely thank the reviewer for this positive and encouraging assessment of our work. We appreciate the reviewer's recognition that UAV-based measurements may offer an important pathway for advancing future surface-atmosphere exchange studies. We also appreciate the reviewer's summary of the main contributions of this study, particularly the finding that the UAV-derived vertical fluxes agree well with the flux tower observations.

At the same time, we agree with the reviewer that the analysis of horizontal advection is less conclusive. In the original version, 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 796 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 3.18, 0.944, and 1.32 for the 50×50×50, 100×100×50, 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 6.41 kg h^{-1}, while the median $E_{C,VT}$ values range from 0.033 to 4.85 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.

Overall, we are grateful for the reviewer's constructive evaluation, which helped us improve both the balance and the clarity of the manuscript.

(2) *Unless this is made clearer in the text, many readers may assume that part of the motivation was that the UAV approach can do what EC can't, and that is to quantify horizontal advection. The problem with that is that you are quantifying the advection between z=35m to z=100m, not the advection most flux tower operators are worried about, i.e. below the EC measurement level. I think the paper would be strengthened by a careful explanation / conceptualization early on that what is being compared when doing the tower – UAV comparison are the turbulent fluxes across the horizontal plane at z = 31m (which is what the tower gives you) and the mass budget of the box sitting on top of this plane, from z = 35m to 100m, typically 100x100m horizontally), which contains no CO2 sources or sinks, i.e. if all the fluxes across the 6 surfaces of the box are accounted for, the numbers should be the same in the limit of insignificant flux divergence/advection. And it seems to work, more or less! By explaining this, you can alleviate any worries a reader may have about the complexities below the z=31m level.*

Response: We sincerely thank the reviewer for this very insightful and constructive comment.

We fully agree that this conceptual distinction needs to be stated much more clearly in the manuscript. In the original version, readers could indeed have gained the impression that the UAV framework was intended to resolve the type of horizontal advection below the EC measurement height that is of primary concern in conventional flux-tower applications. This is not the case.

Following the reviewer's suggestion, we have revised the manuscript to clarify early on that the tower-UAV comparison concerns two different but physically related quantities: (i) the turbulent CO_2 flux across the horizontal plane at the EC measurement height (~31 m), as measured by the flux tower, and (ii) the mass budget of the UAV-defined control volume above that level, extending approximately from 35 to 100 m. We now explicitly state that this upper control volume contains no major internal CO_2 sources or sinks, so that if the fluxes across all six surfaces are properly accounted for and internal flux divergence is small, the two estimates should be consistent. We agree that making this conceptualization explicit helps avoid unnecessary concern about the complexity of transport below the EC level.

Accordingly, we have added this clarification in both the Introduction, Section 4.2.1 and Section 4.4 of the revised manuscript.

The revised sentence of Introduction is as follows:

"As presented further below, the results demonstrate a highly versatile methodology for determining net carbon fluxes over complex ecosystem terrains that complements existing carbon flux tower measurements by providing spatially explicit information on three-dimensional CO_2 transport and, with sufficient UAV observations, an independent estimate of forest carbon fluxes."

The specific revisions are shown in lines 161 to 163 of the revised manuscript.

The revised sentence of Section 4.2.1 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.

The revised sentence of Section 4.4 is as follows:

"To verify whether net emissions (E_C) calculated using the mass-balance approach differ across spatial scales, emission computations were made 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.

(3) *In its current form, the manuscript has a few weaknesses that need to be fixed. The theory section is somewhat sloppy, with many undefined terms, missing references/explanations, etc. – see specific comments for details. The interpretation / discussion of the discrepancies in 4.2 and 4.3 is weak and needs to be revised. One big problem is that the uncertainties are vastly underestimated (as they were in Gordon et al., 2015). There are errors due to non-stationarity and the non-instantaneous nature of the UAV flights that should be quantified and included. Perhaps running an LES simulation might shed some light on differences between actual fluxes and trying to estimate them with a single point measurement moving through space to map a box over ~ 20 minutes or so.*

Response: We have rewritten Section 3.2. And details please refer to our response to Comment 1 in the Specific Comments section.

We have revised both Sections 4.2 and 4.3. In particular, for Section 4.2, we replaced the original coefficient of determination (R^2) with the correlation coefficient (r) and moved the nighttime flux comparison analysis to the Supplement. For Section 4.3, we reanalyzed the horizontal flux component.

With regard to the questions on uncertainties, we agree that non-stationarity is an issue that is hard to address. The work of Gordon et al. (2015) showed uncertainties determined from direct comparison between aircraft measurements and point sources measurements on a minute-by-minute basis, and it is hard to argue against such comparison results. It is possible, that other meteorological conditions may produce different uncertainties. In fact, a follow up paper by Fathi et al. (2021) showed that non-stationariy can indeed cause large uncertainties in the Gordon et al. (2015) algorithm, but these would still be on the order of -25% to 24%, only slightly larger than Gordon et al. (2015) reported. In a third study, the TERRA algorithm was further compared with a separate methodology (Erland et al., 2022), and the agreement between the separate emission computation algorithms was within 3-25%.

In Gordon et al.'s case, the measurement periods were much longer, on the order of 1-2 hours, thus making the stationarity assumption relatively weak. But even so, these studies have shown uncertainties that are not far off from their original conclusions.

Now, in the case of using the UAV, the stationarity assumption can be expected to be much more applicable compared with those situations faced by Gordon et al. (2015) simply because of the much shorter time, at about 20 minutes each flight, spent in completing the measurements, even though it will not completely eliminate the issue. Our validation experiments, using measurements of coking factory smoke stack emissions, indicate an uncertainty of approximately 10% (Han et al., 2024). Our unpublished validation work (paper in preparation), in experiments using point source tracer release, showed uncertainties of 4.8% and precision of 6.7%. Overall, all results indicate that a 22% uncertain (all the data were used for recalculation) estimate for the present UAV methodology is reasonable.

If the UAV methodology were to be applied under rapidly changing boundary layer wind field, the stated uncertainties could be biased low; however, our UAV flights were performed under normal meteorological conditions with no strong gusts or other observable changing conditions such as strong shears or layering as indicated by temperature/relative humidity changes, and we believe the present results are reasonably good estimates of the uncertainty in the methodology.

Regarding the reviewer's suggestion to use large-eddy simulation (LES) for further analysis, we would like to point out that LES itself also involves uncertainties. At the moment, we have no plan to carry out such an analysis.

References:

Fathi, S., Gordon, M., Makar, P. A., Akingunola, A., Darlington, A., Liggio, J., Hayden, K., and Li, S.-M.: Evaluating the impact of storage-and-release on aircraft-based mass-balance methodology using a regional air-quality model, Atmospheric Chemistry and Physics, 21, 15461–15491, https://doi.org/10.5194/acp-21-15461-2021, 2021.

Erland, B. M., Adams, C., Darlington, A., Smith, M. L., Thorpe, A. K., Wentworth, G. R., Conley, S., Liggio, J., Li, S.-M., Miller, C. E., and Gamon, J. A.: Comparing airborne algorithms for greenhouse gas flux measurements over the Alberta oil sands, Atmospheric Measurement Techniques, 15, 5841–5859, https://doi.org/10.5194/amt-

15-5841-2022, 2022.

Han, T. R., Xie, C. H., Liu, Y. Y., Yang, Y. R., Zhang, Y. H., Huang, Y. F., Gao, X. Y., Zhang, X. H., Bao, F. M., and Li, S. M.: Development of a continuous UAV-mounted air sampler and application to the quantification of CO_2 and CH_4 emissions from a major coking plant, Atmospheric Measurement Techniques, 17, 677-691, https://doi.org/10.5194/amt-17-677-2024, 2024.

(4) *This kind of field work would ideally be done with several fully automated UAVs flying simultaneously, and as continuously as possible – food for thought! Also, a flatter site, or one with several flux towers that can provide an independent estimate of horizontal advection, may be worth considering for future studies.*

Response: We agree that simultaneous measurements using multiple UAVs would represent an ideal approach for capturing the full spatial and temporal variability of the control volume, and could significantly reduce uncertainties associated with sequential sampling. We also agree that conducting experiments over flatter terrain or at sites equipped with multiple flux towers would provide additional constraints on horizontal advection and improve the robustness of the analysis.

Following the reviewer’s suggestion, we will consider conducting multi-UAV flights and experiments over flatter terrain in our future work.

(5) *Given the lengthy list of improvements suggested below, I will categorize this as “major revisions”.*

Response: Thanks. We have carefully addressed all comments and have revised the manuscript extensively in response. We believe that these revisions have improved the clarity, robustness, and overall presentation of the work. Detailed responses to each comment are provided below.

Specific comments

(1) *All equations throughout: make sure ALL terms are defined explicitly.*

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 CO2 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

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

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 CO2 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)

$$K_M = lu_* \tag{3.8}$$

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

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

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

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

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

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

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

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

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.

"

(2) *Line 38: simplify to r2 = 0.77. With such a tight range, no need for the "approximately" or the range.*

Response: We have revised this statement as suggested. Moreover, the goodness-of-fit metric has been consistently changed to the Pearson correlation coefficient (r) throughout the manuscript. The revised sentence now reads:

"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."

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

(3) *Line 144: the UAV method complements the EC fluxes, but not in the way most might expect. The current approach, which looks only at the box above 35m, says nothing about advection below the eddy covariance level and therefore cannot help making the*

EC estimate (which relies on assuming horizontal homogeneity below the measurement height) more accurate. But, with a significant amount of UAV data, it can provide an independent measure of the forest flux.

Response: We agree that, under the current flight design, the UAV measurements are limited to the box above 35 m and therefore do not constrain advection below the eddy-covariance (EC) measurement height. As a result, the present UAV framework cannot be used to directly improve the EC estimate by addressing the assumption of horizontal homogeneity below the EC level. We have revised the manuscript accordingly to avoid overstating this capability. The revised text now clarifies that the UAV method complements existing carbon flux tower measurements by providing spatially explicit information on three-dimensional CO_2 transport above the canopy and, with sufficient UAV observations, can provide an independent estimate of forest carbon fluxes. The revised sentence now reads:

"As presented further below, the results demonstrate a highly versatile methodology for determining net carbon fluxes over complex ecosystem terrains that complements existing carbon flux tower measurements by providing spatially explicit information on three-dimensional CO_2 transport and, with sufficient UAV observations, an independent estimate of forest carbon fluxes."

The specific revisions are shown in lines 158 to 163 of the revised manuscript.

(4) *Table 1: define the "D" in the GPS accuracy*

Response: We agree that the symbol D in the GPS accuracy specification was not clearly defined in Table 1 and could cause confusion. We have revised Table 1 to explicitly define D as the baseline distance between the UAV receiver and the RTK base station, expressed in kilometers. To improve clarity and consistency with standard GNSS/RTK accuracy specifications, we also revised the notation using the conventional "ppm × D" format. The revised sentence now reads:

"**Table 1.** Specifications and capabilities of the UAV.

hover endurance	≥30 min (for 15 kg payload)
Peak flight velocity	18 m s^{-1}
Peak climb velocity	4 m s^{-1}
Control radio range	≥7 km
Operating ambient temperature	-20 °C~55 °C
Maximum operating altitude	≥5000 m
Wind resistance capability	14.4 m s^{-1}
Payload capacity	15 kg (without battery)
Rotor blade length	765±5 mm
Symmetric motor center-to-center length	1635±10 mm
Body stands to the body top length	565 ± 5 mm
Wing span	2570 ±10 mm
Battery capacity/unit	30000 mAh (2 units)
Battery weight/unit	6 kg ± 0.1 kg
Radio modem	902 MHz ~928 MHz
GPS accuracy	Horizontal direction: $\pm (8+1\times10^{-6}D)$ mm;

	Vertical direction: ± (15+$1\times10^{-6}D$) mm
Navigation	GPS (RTK system)

Note: In the GPS accuracy specification, *D* is the baseline distance between the UAV receiver and the RTK base station, expressed in kilometers.”

The specific revisions are shown in lines 180 to 182 of the revised manuscript.

(5) *Table 2: I presume you simply mean “ppb” for the CO2 accuracy*

Response: We agree that the original unit given for the CO_2 specification in Table 2 was inappropriate. The revised text in the manuscript:

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

Instrument	Sample Frequency (Hz)	Weight (kg)	Measurement	Sensitivity or Accuracy
MIRA Ultra	1	6	CO_2	< 200 ppb/s (Sensitivity)
Trisonica Mini	5	0.05	Wind speed	(0 - 10 m/s): ±0.2 m/s (11 - 30 m/s): ±2%
			Wind direction	± 1.0°
			Temperature	± 2.0 ℃
			Pressure	± 1.0 kPa

”

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

(6) *Fig. 1: what is the LED for?*

Response: The LED shown in Fig. 1 is used as a visual status indicator for the main controller and the UAV measurement system. It helps display the operating state of the system during flight and ground testing. We have revised the figure caption/text accordingly to clarify its function. The revised text in the manuscript now reads:

“The LED is used as a status indicator to display the operating state of the UAV measurement system during flight and ground testing.”

The specific revisions are shown in lines 231 to 233 of the revised manuscript.

(7) *Section 3.1: very nice. This is very important & useful to future studies trying to do this.*

Response: We fully agree that wind-speed correction is a crucial component for airborne atmospheric measurements. In fact, this correction framework is the foundation that enables our group to conduct airborne flux measurements and is fundamental to the present study. At present, many airborne observation studies can provide only the spatial distribution of pollutants, but cannot quantitatively determine their emission rates or fluxes. A key reason is the difficulty in obtaining accurate wind measurements under UAV flight conditions. We therefore appreciate the reviewer’s recognition of the importance of Section 3.1.

(8) *Line 328: a horizontal u’ of 0.02 m/s seems very low – it should be at least on the order of w’.*

Response: Following the reviewer's comment, we recalculated the horizontal turbulent velocity fluctuation u'. The updated results show that the 95th percentile of u' is less than 0.1 m/s, rather than 0.02 m/s. We have revised the manuscript accordingly. The revised text in the manuscript now reads:

"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 specific revisions are shown in line 358 of the revised manuscript.

(9) *Line 329: by more than a factor of 100.*

Response: We agree that our original wording was not sufficiently precise. Following the reviewer's suggestion, we have revised the sentence to use a clearer and more standard expression.

Revised text in the manuscript:

"The horizontal advection term ($E_{C,H}$) is larger than the horizontal turbulence term ($E_{C,HT}$) by more than a factor of 20."

The specific revisions are shown in lines 359 to 360 of the revised manuscript.

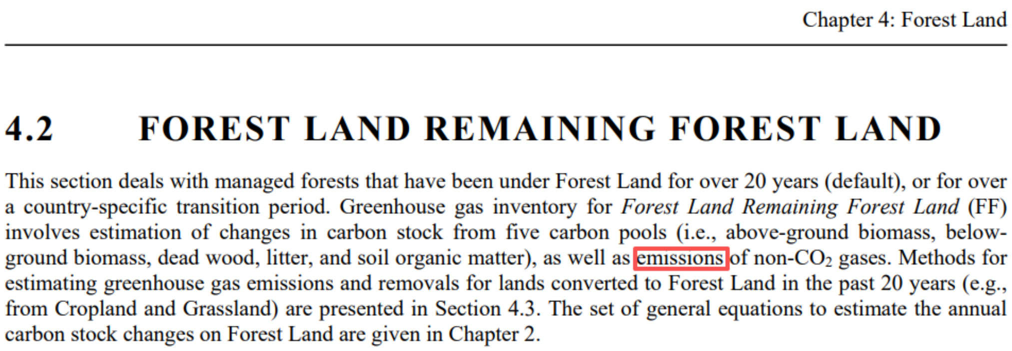
(10) *Line 345, 362 & elsewhere: the term "emission" is misplaced here. What you are really talking about is advection. Emission should be reserved for processes related to the actual creation of CO_2.*

Response: 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. In addition, to avoid potential misunderstanding, we have introduced the term "emission/removal" in the Introduction, while retaining "emission" as the primary term in the remainder of the manuscript. 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.

Revised text in the manuscript:

"In the present study, 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.



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.

(11) *Eq 3.8: this is the exchange coefficient for momentum. In what follows you assume that Kc = Km, which is only approximately true, some of the time. This needs to be stated explicitly.*

Response: We agree that Eq. (3.8) formally represents the exchange coefficient for momentum. In the subsequent derivation, we assume that the eddy diffusivity for CO_2 is equal to that for momentum ($K_C = K_M$) under neutral stratification conditions. We have revised the manuscript accordingly to make this assumption explicit.

Revised text in the manuscript:

"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)

$$K_M = lu_* \tag{3.8}$$

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."

The specific revisions are shown in lines 422 to 427 of the revised manuscript.

(12) *Eq. 3.10: where does this stratification correction parameter come from? Please provide references or a derivation. This is not part of the Monin-Obukhov "canon" and needs to be explained.*

Response: 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 1 in the Specific Comments section.

(13) *Line 399: wind profiles "close to the forest"*

Response: We agree that the original wording was not sufficiently precise. We have revised the description associated with Eq. (3.13), added the relevant reference, and clarified in the text that this discussion refers specifically to wind profiles above the canopy.

Revised text in the manuscript:

"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.

(14) *Line 402: the displacement height depends on the type of forest, so state "for this type of forest"*

Response: We agree that the zero-plane displacement height depends on the vegetation/forest type, and that our original wording was too general. We have revised the sentence accordingly to clarify that the value given here applies to this type of forest.

Revised text in the manuscript:

"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."

The specific revisions are shown in lines 458 to 4459 of the revised manuscript.

(15) *Line 425: Eq. (3.17)*

Response: We thank the reviewer for pointing out this error. It is corrected now but using the new equation sequence number.

Revised text in the manuscript:

"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 line 480 of the revised manuscript.

(16) *Line 426: specify what z1 and z2 you used in your calculations*

Response: We have now stated what altitudes were used in the calculation. The revised text is given below. 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 $\bar{c_1}$ and $\bar{c_2}$ 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.

(17) *Line 427: make it explicit that this section is about the (stationary) tower EC system*

Response: We agree that the original section title was not sufficiently explicit. We have revised the title to clarify that this section refers specifically to the stationary tower-based eddy covariance (EC) system.

Revised text in the manuscript:

"3.3 Stationary tower-based data processing"

The specific revisions are shown in line 483 of the revised manuscript.

(18) *Line 445: remove "eddy covariance"*

Response: We have rewritten the sentence accordingly by removing "eddy covariance" and clarifying the description of the instruments. We also added the sensor models for completeness.

Revised text in the manuscript:

"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.

(19) *Lines 450-452 are misplaced and should be moved to 2.1.2.*

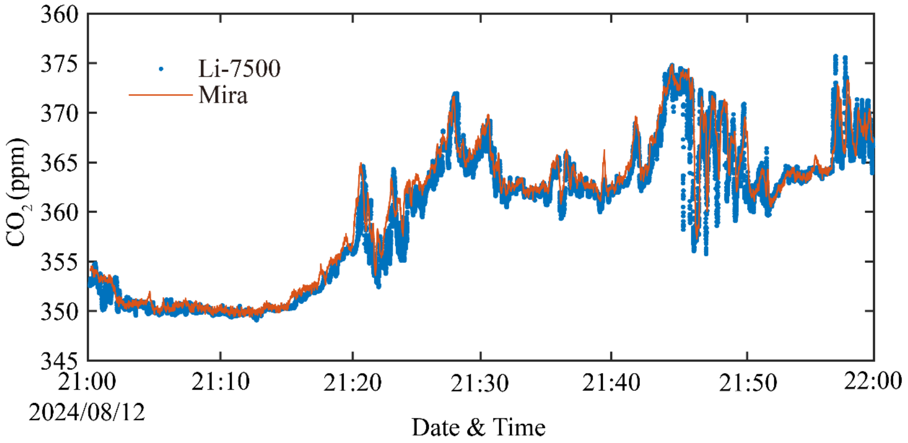
Response: This portion of text is now moved to Section 2.1.2, where the scientific payloads are introduced.

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

(20) *Fig. 4: nice! However, a meaningful x-axis would improve this (time units), and a 1:1 plot along with slope, r2 etc. would also be useful.*

Response: We thank the reviewer for this helpful comment.

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.

(21) *Line 481: missing a verb… "are compared against…"*

Response: We thank the reviewer for pointing this out. We agree that the original sentence was incomplete because the verb was missing. We have revised the sentence accordingly.

Revised text in the manuscript:

"To quantitatively evaluate the agreement between UAV-based profile retrievals and flux tower measurements, the UAV-derived vertical CO_2 flux (F_v) and friction velocity (u_*) are compared against their EC counterparts (Fig. 6)."

The specific revisions are shown in lines 562 to 563 of the revised manuscript.

(22) *Line 483: again; "is used, which…"*

Response: We agree that the original wording was awkward. We have revised the sentence accordingly to improve the grammar and readability.

Revised text in the manuscript:

"Because both datasets contain measurement uncertainties, Reduced Major Axis (RMA) regression (Smith, 2009; Warton et al., 2006) was used, which is more appropriate than ordinary least squares when neither dataset can be treated as error-free."

The specific revisions are shown in line 565 of the revised manuscript.

(23) *Fig. 5: what does SMA stand for*

Response: 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"

(24) *Table 3: state how you define the seasons. For ecological studies, growing seasons are often used, but I presume you are using 3-month calendar seasons.*

Response: We agree that the definition of the seasons should be stated explicitly. In this study, the seasons are defined according to the three-month meteorological seasons rather than ecological growing seasons. We have revised the caption of Table 3 accordingly.

Revised text in the manuscript:

"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 582 to 584 of the revised manuscript.

(25) *Table 4: that the evening is so much worse than the morning is a bit surprising. A figure (perhaps in the SI) showing the typical diurnal cycle of u* and/or stability (Ri) (based on the tower system) might be helpful to explain the asymmetry. Fig. 8 suggests morning and evening should behave similarly.*

Response: We agree that the stronger deterioration in the evening compared to the morning requires further explanation. Although Fig. 8 suggests that the flux magnitudes in the morning and evening are similar, this does not necessarily imply a similar level of agreement between UAV-derived and tower-based fluxes.

To address this point, we have added a new figure (Fig. S1) in the Supplement, showing the typical diurnal variation of friction velocity (u_*) for different seasons based on the tower system (see figure below). The results clearly demonstrate a pronounced asymmetry between the morning and evening periods. After sunrise, turbulence develops gradually, leading to increasing u_* values and more stable flux conditions. In contrast, during the late afternoon and evening, turbulence decays rapidly and the atmosphere transitions toward stable stratification, resulting in lower u_* values and increased variability.

This asymmetry is reflected in the larger scatter and lower turbulence intensity during the evening period, which increases the relative contribution of storage and advective processes and leads to greater uncertainties in both UAV-derived and tower-based flux estimates. As a result, the agreement between the two methods is substantially poorer in the evening compared to the morning.

We have revised the manuscript accordingly to include this explanation and the additional figure in *SI Appendix* S3.

Revised text in the manuscript:

"The diurnal variation of friction velocity (u_*) derived from the tower system (Fig. S1) shows a clear asymmetry between morning and evening transitions, with turbulence developing gradually after sunrise but decaying rapidly in the evening, which helps explain the poorer agreement under evening conditions."

The specific revisions are shown in lines 678 to 682 of the revised manuscript.

Revised text in the Supplementary Material:

"**S2. Comparative analysis of fluxes during the morning and evening periods**

The diurnal variation of friction velocity (u_*) exhibits a clear asymmetry between the morning and evening transitions, with turbulence developing gradually after sunrise but decaying rapidly in the late afternoon. The increased variability under evening conditions reflects enhanced instability and contributes to the larger uncertainties observed in flux estimates. The results clearly demonstrate a pronounced asymmetry between the morning and evening periods. After sunrise, turbulence develops gradually, leading to increasing u_* values and more stable flux conditions. In contrast, during the late afternoon and evening, turbulence decays rapidly and the atmosphere transitions toward stable stratification, resulting in lower u_* values and increased variability.

This asymmetry is reflected in the larger scatter and lower turbulence intensity during the evening period, which increases the relative contribution of storage and advective processes and leads to greater uncertainties in both UAV-derived and tower-based flux estimates. As a result, the agreement between the two methods is substantially poorer in the evening compared to the morning.

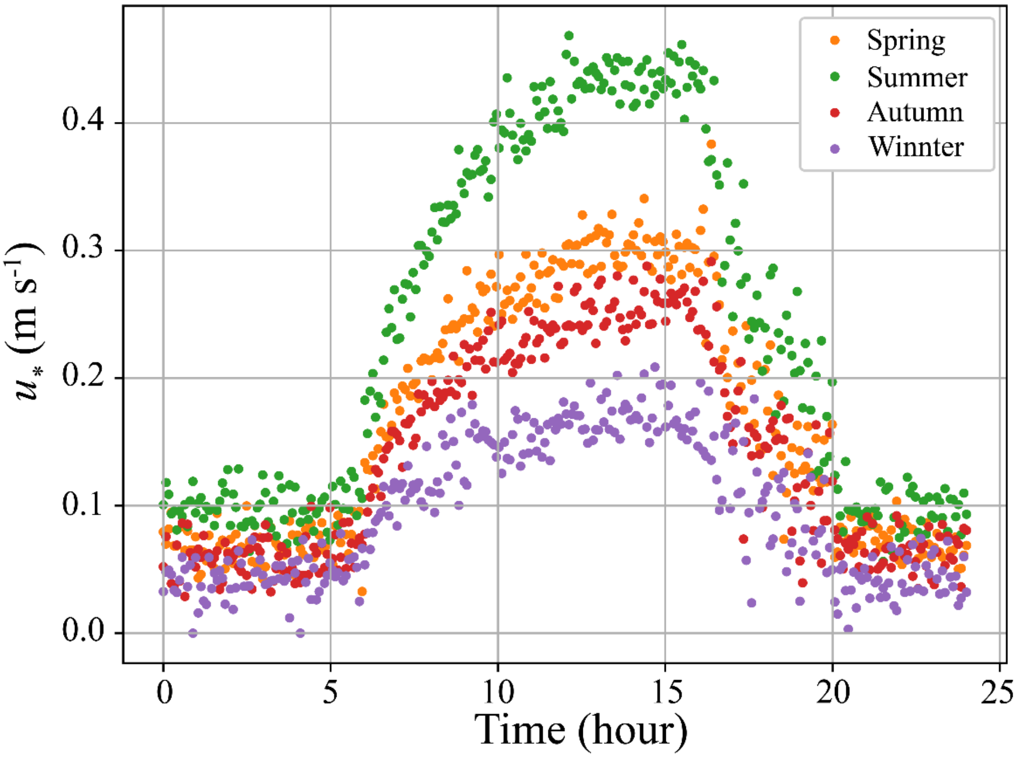


Fig. S1. Diurnal variation of friction velocity (u_*) for four seasons."

The specific revisions are shown in *SI Appendix* S2, lines 76 to 92 of the revised manuscript.

(26) *Also, spelling of "morning*

Response: We thank the reviewer for pointing this out. We have corrected the spelling in Table 4.

The specific revisions are shown in line 664 of the revised manuscript.

(27) *Line 582: Rephrase this sentence. Why would low nighttime turbulence lead to vertical fluxes? I presume you're trying to talk about the relative magnitudes of the various flux terms.*

Response: We agree that our original wording was inaccurate and could be misleading.

Revised text in the manuscript:

"Under such conditions, the UAV methodology is subject to larger uncertainties in the derived vertical fluxes and shows an apparent low bias relative to the EC measurements."

The specific revisions are shown in lines 672 to 674 of the revised manuscript.

(28) *Line 597: how is nighttime defined? Same as "evening" before?*

Response: 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. 6). To avoid ambiguity, we have replaced "nighttime" with "evening" throughout the relevant text.

The revised text is as follows:

"Figure S1 presents the seasonal distributions of evening CO_2 flux differences, defined as $\Delta F_v = F_{v,UAV} - F_{v,Tower}$, for spring, summer, autumn, and winter."

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

(29) *Line 604: the F term needs to be explicitly explained – not everyone knows ANOVA intimately. The sentence in 605 can be shortened or eliminated. F=2.50 actually means that "among" is exactly 2.50 times larger than "within", not approximately.*

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.

(30) *Fig. 7 is really not needed, or could be moved to the SI.*

(31) *In fact, this whole section on evening/nighttime fluxes could be shortened to a single paragraph. Obviously the relative errors will become larger at night, but since*

the fluxes are much smaller than during the day, they should be assigned a lower importance – if what you are after is the net ecosystem exchange

Response: We agree that the original discussion of the evening fluxes was too long relative to their importance for interpreting net ecosystem exchange. As the reviewer notes, although the relative uncertainties become larger under weak evening turbulence, the absolute magnitudes of the evening fluxes are much smaller than those during the daytime and therefore contribute less to the net ecosystem exchange. We have therefore substantially shortened this section and the analysis of the nighttime flux differences has been moved to *SI Appendix* S2.

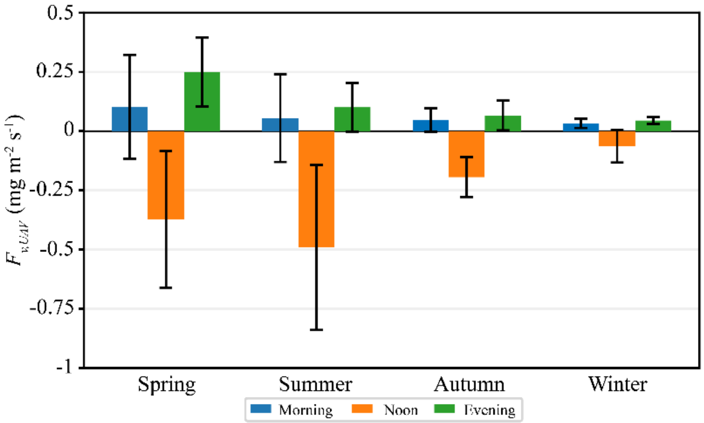
Revised text in the manuscript:

"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 primary reason for such a comparison result is that, under low evening turbulence, the vertical turbulent flux becomes small, while the relative contributions of other terms increase. Under such conditions, the UAV methodology is subject to larger uncertainties in the derived vertical fluxes and shows an apparent low bias relative to the EC measurements. It should be noted, however, that EC vertical flux measurements under such low turbulence conditions also bear large uncertainties, and hence this comparison only implies a low bias in the UAV methodology relative to the EC methodology, and does not indicate the absolute errors in the UAV-determined F_v. Furthermore, an analysis of the evening flux differences indicates that both UAV- and tower-derived vertical CO_2 fluxes are associated with greater uncertainties at evening, whereas the overall differences remain small and show no significant seasonal variations. A more detailed analysis of the evening flux differences is provided in the *SI Appendix* S3."

The specific revisions are shown in lines 682 to 686 of the revised manuscript.

(32) *Fig. 8: ensure color consistency with Fig. 6.*

Response: Following the reviewer's suggestion, we have revised Figure 8 so that its colors are now consistent with those used in Figure 7.



The specific revisions are shown in line 698 of the revised manuscript.

(33) *Line 643: this is a coniferous forest, so I presume the seasonal variation in LAI is*

small. Therefore, LAI is not the explanation. The reason really is simple growing season dynamics related primarily to solar radiative input and temperature.

Response: We agree that, for this coniferous forest, seasonal variation in LAI is likely small and should not be emphasized as the primary explanation for the stronger net CO_2 uptake in spring and summer. We have revised the sentence accordingly.

Revised text in the manuscript:

"Summer and spring display the strongest net CO_2 uptake, consistent with favorable moisture conditions and strong radiative forcing in this subtropical forest."

The specific revisions are shown in lines 704 to 706 of the revised manuscript.

(34) *Section 4.3. When doing these comparisons of different box sizes, you need to be very careful to compare "similar situations". Ideally, the 3 boxes would be captured simultaneously (and instantaneously), but that of course is not possible. So it is very important to bin together very similar conditions (time of day, winds, turbulence, EC flux magnitudes etc.), otherwise you are comparing apples with oranges. Please add an explanation of what exact datasets are compared.*

Response: We fully agree that the comparison among different box sizes is meaningful only if the datasets represent sufficiently similar observational conditions. Since simultaneous measurements of the three control volumes are not possible with the present UAV framework, we have revised Section 4.3 to explicitly describe the dataset selection criteria and the exact cases included in the comparison. In the revised manuscript, we clarify that only UAV measurements obtained under comparable time-of-day and seasonal conditions were used, specifically the summer-noon and autumn-morning cases, which also showed better agreement with the flux tower comparisons in Section 4.2. We further state that measurements from the three control volumes were required to be completed within a short time window of 20 min, and that the corresponding wind direction and tower flux variability were examined to ensure comparability. These additions are intended to make clear that the comparisons were made only under sufficiently similar conditions.

Revised text in the manuscript:

"To minimize the influence of different time periods and seasons on the statistics, and according to the analysis in Section 4.2, the UAV measurements of CO_2 emissions collected in summer noon and autumn morning, which show better agreement with the flux tower comparisons, were selected for the spatial heterogeneity analysis. 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 770 to 781 of the revised manuscript.

(35) *Line 662-666: remove; superfluous*

Response: We agree that this part was superfluous and have removed it from the revised manuscript accordingly.

(36) *Line 672: not "based on". Simply say "the means (see boxes in Fig. 9) were ..."*

Response: We agree that the original wording was awkward and have revised the sentence accordingly.

Revised text in the manuscript:

"The means for the three box scales (see boxes in Fig. 10) were -0.25, -0.30, and -0.24 mg m^{-2} s^{-1}, respectively."

The specific revisions are shown in lines 784 to 785 of the revised manuscript.

(37) *Line 674: summer at noon*

Response: Following the reviewer's suggestion, we have revised this wording for clarity and consistency. Here, "at summer noon" has been changed to "during the summer-noon period".

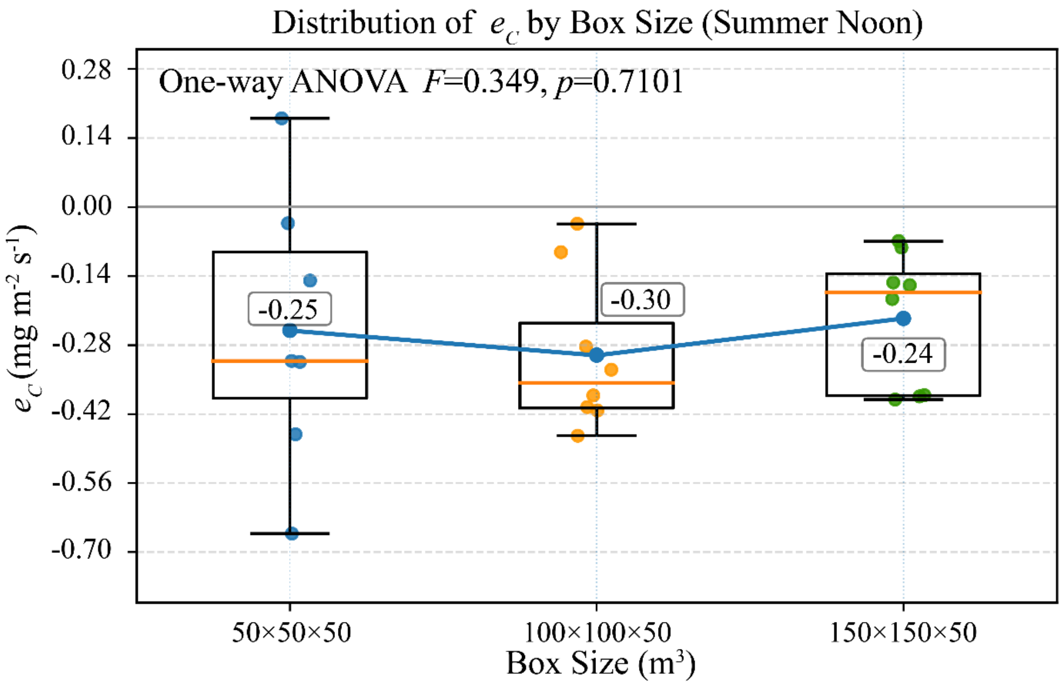
Revised text in the manuscript:

"This indicates a pronounced net CO_2 uptake (i.e., negative emissions) of the system as a whole during summer at noon."

The specific revisions are shown in lines 786 to 787 of the revised manuscript.

(38) *Fig. 9: move central annotation to make it legible*

Response: We have revised the figure accordingly by repositioning the central annotation to improve its legibility. The revised figure is shown below.



The specific revisions are shown in line 808 of the revised manuscript.

(39) *Table 5: normalize the emission rates by area for easier comparability*

Response: We agree that area-normalized emission intensity is useful for comparing the overall net exchange among different box sizes. In fact, this normalization has already been adopted in the main text when discussing the scale dependence of e_C. However, the purpose of Table 5 is different: it is intended to further examine the relative contributions of the horizontal transport component and the vertical exchange component to the net emission estimate. If only the area-normalized emission intensity were reported in Table 5, it would not be possible to directly evaluate the respective magnitudes of $E_{C,H}$ and $E_{C,VT}$.

At the same time, we agree that the original statistics in Table 5 were not sufficiently informative. Therefore, in the revised Table 5, we have expanded the summary by adding the median values and standard deviations of $E_{C,H}$ and $E_{C,VT}$, so that the relative roles of the horizontal and vertical components can be evaluated more clearly. The revised table is shown 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})	-2.235	-10.581	-18.596	0.166	7.641	9.354
E_C Median (kg h^{-1})	-2.787	-13.652	-12.758	0.138	7.253	11.254
IQR (kg h^{-1})	2.648	7.235	20.437	1.849	3.257	15.241
$E_{C,H}$ Median (kg h^{-1})	0.0687	0.0817	0.273	0.105	3.522	6.406
$E_{C,H}$ SD (kg h^{-1})	0.0162	0.0321	0.0732	0.0235	1.252	1.583
$E_{C,VT}$ Mdian (kg h^{-1})	-2.77183	-13.570	-12.485	0.033	3.731	4.848
$E_{C,VT}$ *SD* (kg h^{-1})	0.437	2.353	3.539	0.012	1.153	1.428
Median $E_{C,H}/E_{C,VT}$	-0.025	-0.006	-0.022	3.181	0.944	1.321

The specific revisions are shown in line 820 of the revised manuscript.

(40) *Line 706: units*

Response: We thank the reviewer for pointing this out. We have added the unit to the reported emission intensities for clarity. The unit is mg m^{-2} s^{-1}.

Revised text in the manuscript:

"The mean values in Fig. 11 indicate that the emission intensities ($\boldsymbol{e_C}$) derived for different box sizes (0.02, 0.21, and 0.11 mg m^{-2} s^{-1}, respectively) differ to some extent."

The specific revisions are shown in line 824 of the revised manuscript.

(41) *Line 711: In which upper layer? And are you saying that the footprint is homogeneous in the summer but heterogeneous in the fall? That seems unlikely. It may simply be a matter of the relative magnitudes of ECV and ECH, and the mornings (regardless of season) may be more strongly affected by advective issues such as katabatic flows.*

Response: We agree that our original wording was not sufficiently precise. What we intended to express is that, above the canopy, the net CO_2 emission intensities calculated for the three different control volumes show spatial heterogeneity. We did not intend to imply that the footprint is homogeneous in summer but heterogeneous in autumn. Rather, the comparison was carried out separately using two subsets of data, namely summer-noon and autumn-morning cases.

As the reviewer suggests, the morning period, regardless of season, is indeed more likely to be affected by advection. In our analysis of the three box sizes for the autumn-morning cases, we found scale-dependent differences in the calculated net emission intensities. It should be emphasized that these emission intensities were derived from the combined three-dimensional budget, including both the horizontal and vertical components. Under weak early-morning photosynthetic conditions, the vertical component is expected to vary less across the different control volumes; therefore, the observed differences in net emission intensity are likely associated mainly with horizontal transport. This interpretation is consistent with the reviewer's suggestion that morning conditions are more strongly affected by advective processes. We have revised the manuscript accordingly to clarify this point. This also highlights an important advantage of the present UAV-based method, namely its ability to resolve spatial differences in CO_2 exchange that cannot be directly diagnosed by a traditional fixed flux tower.

Revised text in the manuscript:

"This provides direct evidence that the net CO_2 emission intensity derived for the three virtual control volumes above the canopy is spatially heterogeneous."

The specific revisions are shown in lines 828 to 830 of the revised manuscript.

(42) *Lin 721: it's not that simple: the lowest ratio is at the 100×100m scale.*

Response: We agree with the reviewer that the interpretation here was oversimplified and not sufficiently supported by the results, especially given that the lowest ratio occurs at the 100×100 m scale. We have therefore removed this paragraph from the revised manuscript.

(43) *Line 799: this uncertainty is unrealistically low. Significant noise is almost certainly introduced by the assumption of stationarity (partly, assuming the ~20-minute flight represents an instantaneous snapshot of a static field, and partly, that even the instantaneous snapshot would be useless unless many snapshots are averaged to capture all relevant eddies), by interpolation, etc. Look at Fig. 11 and imagine all the potential variability inside the box that you can't see!*

Response: Please refer to our response to Comment 3 in the General Comments section.

we have added a statement in the revised manuscript explicitly noting that the reported uncertainty does not include these additional sources of variability.

"It should be noted that this uncertainty estimate primarily reflects measurement and propagation errors within the flux calculation framework, and does not explicitly account for additional uncertainties associated with assumptions of stationarity, spatial interpolation, or unresolved sub-grid variability."

The specific revisions are shown in lines 936 to 940 of the revised manuscript.

In addition, 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.

Finally, after recalculation, we obtained a maximum uncertainty of 22%.

(44) *Section 4.5. I think moving this section forward, e.g. right after 4.2, would improve the flow.*

Response: We have revised the manuscript accordingly by moving Section 4.5 to immediately follow Section 4.2, which improves the logical flow and readability of the Results and Discussion section.

The specific revisions are shown in lines 717 to 757 of the revised manuscript.

(45) *Line 865: spatial heterogeneity, sure, but also advection of concentration "blobs", coupled with the finite integration time, sequential vs. instantaneous measurements etc. In 20 minutes, a lot can change. The 200m box can be traversed by a 1m/s wind in 200s; if that is your round-trip time, you may consistently hit the maximum of a hypothetical sinusoidal wave with a wavelength of 400m on the upwind side of the box and the minimum on the downwind side (as a simplified example).*

Response: We agree that the observed scale dependence cannot be attributed solely to spatial heterogeneity. As the reviewer points out, horizontal advection of concentration structures, the finite integration time, and the sequential rather than instantaneous nature of the UAV measurements may also contribute to the observed variability. These effects are difficult to fully avoid within the present UAV measurement framework, because the three-dimensional control volumes cannot be sampled instantaneously.

Therefore, we introduced several assumptions at the beginning of Section 3.2: "CO_2 emissions from within the control volume defined by the virtual box were quantified using a mass-balance framework that integrates measurements obtained from box-pattern and profile-pattern UAV flights. This approach assumes that CO_2 emitted within the control volume is transported by the mean wind field and becomes sufficiently mixed, such that the net emission rate can be inferred from the fluxes crossing the boundaries of the volume. Under steady or quasi-steady wind conditions, the mass flux of a chemical compound through a plane perpendicular to the mean wind direction can be considered approximately constant."

In fact, the issues raised by the reviewer are highly professional and well founded. At present, the main way to address such issues is through a set of necessary assumptions. For example, eddy-covariance flux measurements require assumptions of horizontally homogeneous underlying surfaces, flat terrain, and stationarity; atmospheric dispersion is often treated under the assumption of Gaussian diffusion; pollutant emission calculations commonly assume neutral atmospheric stratification; and in wind engineering, wind resource assessments are also often carried out under the assumption of neutral stratification.

Accordingly, we have not expanded the conclusion of spatial heterogeneity at this stage. Nevertheless, if the reviewer considers a more detailed discussion necessary, we would be happy to incorporate further revisions.

(46) *Refs: Yang et al., 2025 link does not work.*

Response: We checked the DOI link on our side and it opens normally. We kindly suggest trying the following link again: https://doi.org/10.5194/amt-18-3035-2025