Dear Editor,

The authors have made significant improvements to the manuscript. However, there remain a few outstanding issues, especially with regard to the definition of variables and acronyms. These issues should be resolvable with extensive minor revisions. But the manuscript is unpublishable in its current state.

General comments

The authors have failed to use consistently defined terms in this manuscript. u(bar) changed from the "mean east-west wind speed" to "mean wind-speed magnitude" within subsection 3.2. This makes it very difficult to review. I already raised this point in two previous rounds of review. If the authors are unable to rectify their terminology in any further revision, I suggest that this manuscript be rejected. This is totally unacceptable in a scientific publication, regardless of quality scientific content.

In addition, as with the previous versions of this manuscript, terms are not defined at their first point of use, leading to confusion. It is generally conventional to define a term once at its first point of use and then to use the abbreviated term for the remainder of the manuscript (except in the abstract and maybe the conclusion), rather than interchanging between the two.

I hope the authors can appreciate that a reviewer is tasked with evaluating the scientific content of a publication and not to spend hours trying to check whether the equations and the terms are consistent. Please can the authors make an effort to resolve these issues before submission, to make reviewing the scientific content of this interesting work far easier.

The end of the first paragraph explains that horizonal advection can "affect the interpretation of vertical flux measurements" in "spatially heterogeneous or weakly mixed conditions". This is a major improvement from the previous version of the manuscript. However, a detailed reason for this lacking. This point should be expanded on clearly here as this underpins the motivation for a key component of this work. I appreciate that the penultimate sentence of the first paragraph adds some reasoning, but it is a bit vague with no citations.

The method with which turbulent wind fluctuations are calculated is concerning and not adequately described. Apparently, each wind component is subtracted from its 30-minute average. This is fine if all 30 minutes of measurements are made at a single location. However, if the UAV is moving, this assumes that the three-dimensional wind field is independent of height and lateral position (which is unlikely).

Subsections 4.2 and 4.3 are nicely written.

Subsection 4.4 needs some work. It is not clear which flights and which data are being used and in which way. There is confusion between UAV flights and "days of observations". Does one UAV flight occur on each day of observation for each time period? Or can multiple flights take place in a single autumn morning, for example?

Specific comments

Line 37: "is flown"

- Perhaps this should be in the past tense

Line 38: "simultaneously"

- Is it simultaneous? As far as I understand, the vertical transect is used for vertical fluxes and the horizontal transect is used for horizontal fluxes. This means the measurements are not used simultaneously. They are close in time, but not simultaneous.

Line 42: “a nearby”

- The definite article should be used here, as this is referring to the EC system referred to in the previous sentence, rather than some other permanently installed system.
- I think the adjective “nearby” is unnecessary here as it is already used in the previous sentence to describe the EC system.

Line 42: “eddy-covariance”

- Replace this with “EC”. It has been defined in the previous sentence.

Line 44: “turbulence-based exchange coefficient”

- Perhaps make it clearer that this refers to a UAV coefficient, as opposed to the aforementioned carbon dioxide concentration gradient, which is from the EC tower. It should be clearer what is being compared with what and from which sampling platform each of these terms refers to.

Line 79: “For example, eddy-covariance methods primarily quantify vertical turbulent exchange, whereas aircraft-based mass-balance approaches commonly emphasize horizontal advective transport.”

- Please provide some citations to support these assertions.
- As far as I am aware, all EC fluxes are vertical emissions to the atmosphere. I am not aware of any horizontal eddy covariance work. The word “primarily” implies that EC is mostly for vertical fluxes rather than entirely for vertical fluxes.

Line 81: “Consequently, simultaneous quantification of both horizontal and vertical CO_2 transport remains challenging (Leuning et al., 2008)”

- From this opening paragraph, it is still not clear why this is necessary. It seems like a challenge with no clear purpose.
- What is the point of evaluating the horizontal flux from the overall forest to a small horizontal area within the forest?

Line 88: “affect the interpretation of vertical flux measurements under spatially heterogeneous or weakly mixed conditions.”

- Why? Please explain.
- Why do weakly mixed or spatially heterogeneous conditions effect the vertical flux.
- This is a key motivation for this work, so it demands adequate explanation.
- Please provide some citations to support this statement and the following sentence.

Line 90: “Therefore”

- Is this word correct? This is currently written as the previous sentence being a reason for the current sentence (*i.e.* use of the word “therefore”). Instead, this sentence could be used to explain the previous sentence in support of it.

Line 126: “UAV”

- What does this acronym mean? Please define all acronyms used in this manuscript at their first point of use.

Line 131: “rotating wing”

- Replace this with “rotary-wing”.

Line 137: “characterize”

- This sentence does not make sense. Perhaps add “UAVs” before the word “characterize”.

Line 173: “horizontal mass flux estimates”

- Please specify that this refers to carbon dioxide fluxes.

Line 221: “3-d”

- Please define this as “three dimensional”. “d” could mean something else in a different context.

Line 231: “with its air inlet set next to the anemometer sonic sensors to ensure spatially consistent wind and CO2 concentration measurements.”

- This is a very good idea.
- Consider replacing “concentration” with “mole fraction” here and throughout the manuscript, for consistency.

Line 258: “eddy covariance (EC)”

- This has already been defined. I am not sure why it is being defined again.

Line 270: “Site B”

- Is it necessary to define Site A and Site B? The word “site” implies that the locations are distant from each other. The authors can keep this if they insist, but it may be better to simply use the “tower” and the “UAV take-off” location.

Line 272: “The UAV-derived fluxes”

- Figure 2a does not show UAV-derived fluxes.

Line 325: “Because the EC system is installed at 31 m, UAV-derived CO2 vertical fluxes used data from 26 and 36 m to enable direct comparison with tower-based EC measurements.”

- It may be useful to remind the reader here that these are all heights above UAV take-off level.

Line 331: “CO2 mole fractions”

- The values in panel b do not appear to be mole fractions but some sort of mole fraction enhancement. It is impossible for carbon dioxide mole fraction to be 0 ppm.

Line 347: “met tower”

- What is a met tower?

Line 350: “CO_2 emission estimation based on UAV measurements”

- Throughout this subsection, terms are defined in a confusing order. For example, it would be better to first present c before presenting c(bar). Similarly, w(bar) is presented before w. It is easier to follow the notation if the vertical wind speed is defined before the mean vertical wind speed.
- There are also problems with the idea of c(bar) being a mean concentration, as it is not clear in this manuscript what this is an average over.

Line 370: “horizontal and vertical directions”

- Add the word “respectively” here.

Line 382: “turbulent fluctuations (u' , v' , w') were calculated by subtracting the 30 min mean value of each component from its instantaneous measurement.”

- Does this 30-minute mean represent sampling from the full duration of each flight? What if a flight was less than 30 minutes. Just above it says that average wind components were derived using sampling from the full duration of each flight. Are these the same average values given above (u(bar), v(bar) and w(bar))? If so, it is easier to say so clearly. If a different 30-minute calculation is used, why?
- This turbulence value is only valid if 30 minutes of sampling has been conducted at a fixed location. The wind vertical wind speed at ground level is not the same as the vertical wind speed at 50 m. Yet if an average of all wind speed, regardless of height, is used to calculate turbulent parameters, how can this be justified?
- A similar issue exists with horizontal wind components with increasing height. The wind speed at 50 m is higher than wind speed at ground level. So, the calculated turbulence will be more positive at the ground and more negative at 50 m. So, it is not really a turbulence value, if it uses measurements from the full duration of each flight.

Line 384: “was”

- Use of the present tense is probably better here, as the comparison is presented just here, rather than this being a comparison done in the past.

Line 416: "u_n"

- It may be better to use a different letter than u, to avoid confusion with the easterly wind component.

Line 462: "3.7"

- Why is air density used in this equation if it is already incorporated into *c*(bar)?

Line 463: "$\overline{c}$ was obtained by converting the measured CO_2 mole fraction (expressed in ppm)"

- Is *c*(bar) an average concentration? It has not been clearly defined here. If so, what duration is it averaged over? If is in not an average, why use a bar?

Line 474: "u_*"

- It would be better to use a different letter to avoid confusion with the easterly wind speed.

Line 475: "the vertical gradient of the mean wind speed."

- This is not the gradient of the mean wind speed. According to the start of this subsection, *u*(bar) is the mean of the easterly wind component. So its derivative cannot be something else.
- It should clearly state which wind speed this refers to. If it is an overall horizontal wind speed magnitude, this should be stated this clearly.
- If an absolute horizontal wind speed magnitude is used here, then a different letter to "u" should definitely be used here. A term cannot be redefined in the middle of a paper.

Line 487: "$\Delta\overline{u}$"

- u(bar) is the easterly wind component, according to the start of this subsection.

Line 488: "wind-speed magnitude"

- Please specify that this refers to the easterly wind component and not the full horizontal wind speed.

Line 489: "$\overline{\theta}$"

- What is this? I can see above that Δθ(bar) is the smoothed temperature change, but θ(bar) has not been defined.

Line 489 "$\overline{u}$"

- Please specify which of the two definitions for *u*(bar) applies here?

Line 500: "mean wind-speed magnitude"

- This is the second time that *u*(bar) has been defined. At the start of the subsection, it is the easterly wind component. Now it is the overall wind speed magnitude.
- It is unacceptable and totally confusing for defined terms to change in the manuscript, especially mid-way through a single subsection.

Line 514: "3.16"

- This equation includes air density. However, air density is already incorporated into *c*, which is derived from mole fraction.

Line 526: "at using"

- Should this be "and using"?

Line 515: "mean concentrations"

- Is this a mean concentration or a smoothed concentration from seven datapoints? A mean concentration implies that a mean was obtained at a single fixed height when in reality, the UAV was constantly moving.

Line 520: "3.17"

- This equation includes air density. However, air density is already incorporated into *c*, which is derived from mole fraction.

Line 530: “Finally, the three-dimensional net CO_2 emission was obtained using the original observations.”

- I am lost. I think the word “finally” is confusing here.

Line 549: “318”

- This equation uses u', which is defined as the fluctuation in the east-to-west wind component.

Line 550: “3.19”

- This equation is incorrect as c is defined as being in kg m^{-3}. The density is incorporated twice.

Line 560: “same height”

- Please remind the reader that this refers to height above take-off level for the UAV. This means that the actual height is offset.

Line 620: “Pearson correlation coefficient (r)”

- It would be better to define this properly in the previous subsection, when r is first introduced.

Line 623: “at z_1=26 m and z_2=36 m.”

- This is very confusing, as in the previous section, it says that measurements between 26 m and 36 m were used. Instead, here it says that measurements at 26 m and at 36 m were used.
- Measurements between two locations and at two locations are very different things.
- If the authors have used measurements between these two heights, the simple solution is to replace the word “at” with “between”.
- If only a single measurement at 26 m is used and a single measurement at 36 m is used, then I am totally lost.

Line 706: “This suggests that the UAV-based method is generally capable of reproducing the variability of CO_2 fluxes observed by the EC system, while tending to underestimate flux magnitudes under stronger exchange conditions.”

- Why?

Line 834: “Under such a requirement”

- Which requirement? Why does the aforementioned statement on selecting summer noon and autumn morning measurements result in seven days of data being required? Autumn lasts roughly 90 days.

Line 835: “seven days of observations”

- So, does this mean seven days in autumn and seven days in summer or was the combined data collected over seven days?
- Does each day of observation correspond to a single UAV flight?

Line 839: “fixed tower varied by less than 8% during the comparison period.”

- What exactly is the comparison period? Does this represent all of the EC fluxes for the full duration of each of the UAV flights used in each analysis?
- Is this 8% value a variation between average fluxes collected over seven days? Is this an average variation of each discrete flux measurement collected over the period of each UAV flight over seven days?

Line 847: “-0.25, -0.30, and -0.24”

- I am totally lost. Do these values correspond to autumn morning or summer noon?
- This sentence does not make sense in general and needs to be rewritten.
- Are these values an average of all UAV fluxes over a seven-day period. It is not at all clear what these values represent and where the come from.

Line 851: “ANOVA”

- What is this? This has not been defined.

Line 877: “Table 5”

- The eighth row is called “Mdian”. I think this should be “Median”.

Line 888: “uncertainty of the method used in this study (15%)”

- Where does this 15% uncertainty value come from? I have not seen this before.

Line 893: “In contrast”

- Why is this in contrast to the preceding paragraph?

Line 898: “median $E_{C,H}$ values range”

- This is presented as a range of $E_{C,H}$ values. However, it took me some time to understand what this was a range over. In fact, this is a range over different sized boxes, of which there are only three. It would be better to give all three values, rather than a range over three values.

Line 938: “Computation of uncertainties”

- Consider placing this uncertainty subsection at the end of the methods section. It is not really a flux result so it may be better to know the flux uncertainty before the flux is presented.

Line 945: “three components”

- I only see two components here.

Line 947: “combined in quadrature”

- This approach is only valid for independent standard deviation uncertainties. Perhaps state this assumption here.

Line 992: “Three representative cases”

- Why use three cases? Why not use all of the flights? Is it too computationally demanding? Some reasoning would be welcome here.

Line 995: “uncertainties in E_C calculated using the different methods”

- What about linear interpolation results?

Line 996: “respectively”

- Why is this word here? What are these uncertainties respective to? Does each uncertainty correspond to a single season?
- These values should each use data from each of the three seasons.

Line 1000: “Because the CO2 analyzer uncertainty was small”

- This point should not really matter. When evaluating the vertical methodological uncertainty, the instrumental uncertainty should be irrelevant.

Line 1017: “4.3”

- The gradient uncertainty is based on the wind uncertainty. Therefore, these terms are interdependent and cannot ordinarily be combined in quadrature.

Line 1030: “emission”

- The word “fluxes” may be better here, as “emission” insinuates that a positive outward flux is taking place.