We thank the editor for his useful comments, which we addressed point by point below (our replies written in blue).

Line numbers refer to the PDF with trackchanges. Line numbers were removed from the final preprint PDF.

Revision round #2

Decision for round #2 : *Revision needed*

Thanks for your revised manuscript. Despite changes made were not so profound, I found the ms clearly improved in terms of readability and ecological background related to movement, dispersal and migration.

I have a few minor suggestions before I can recommend the ms:

· The 5 min limitation recording justified by the UAV software might be perceived as a stronger limitation than what it is in reality. Indeed, if the video files could be put end-to end without frames being dropped in-between, analysis would be nearly as easy as if a single video file was produced. I suggest indicating that the key point is not the potential splitting of the video file but the ability to avoid some non-recorded period in-between consecutive videos, so readers can consider this when choosing an UAV to record longer movements.

We re-checked our UAV test files, and there was indeed a single frame drop between splitted files. We reformulated on L197-199.

· Track straightness is also known in part of the movement literature as Net To Gross Displacement Ratio (NGDR); maybe worth mentioning it.

Thank you for pointing this out. We added the reference on L360.

· Fig 4: at first read, it was unclear what the small vertical bars are on the bottom track. I had to zoom quite a lot to see these are + signs indicating individual positions. I'd advise to make that explicit. Also, I do not really see the benefit of duplicating the track; I'd keep only the bottom depiction, showing both the general track, individual positions (hence relative speed) and the comparison with the position that could be recorded using SHR.

We modified Fig 4 and caption accordingly. (Please note that the trackchange PDF resolution was reduced to 220 ppi to comply with the 10 MB limit. The BiorXiv PDF is 330 ppi, so presents better quality figures).

· Uncertainty about flight altitude: is the video resolution small enough to use the butterfly (bob) size as a measure of flight altitude, assuming individual real size is not too variable? I guess not because blob size is even not precise enough to discriminate the two species. But the reader might wonder themselves about that

before reading the section on blob size for species discrimination, so maybe a sentence on that in the flight altitude section is beneficial. Also discussing this in relation to the benefits of using a higher camera resolution could be interesting, especially for cases where flight altitude would not be useful only to correct for recorded distance but also as a response variable of biological interest.

We added a sentence on L478-480 in the flight height section of the results.

Regarding resolution, at first glance, for the blob area to be usable for measuring height, a gain of ~1 order of magnitude in resolution would probably be needed, as well as a model to correct for the wide variability in blob area linked to wing beats and the other factors listed on L696-702 about specific identification. We added a mention to flight height on L705, but we really believe that at the moment, if flight height is a variable of interest, one should better turn to natively 3D methods, as explained on L651-655.

· Fig 9, caption: it would be clearer to remind there that the Raleigh test null hypothesis is a uniform distribution, i.e. that a red vector indicates that there is a significant directional preference. Also specify exactly what arrow length represents.

We added explanations in Fig 9 caption.

· Discussion about “repeat the measurement of the reference scale on every video frame”: what about using reference marks that can be automatically thresholded on every frame (e.g. use a specific colour channel), allowing for an automatized measure of UAV drift?

The verb « repeat » might let think that this was a manual operation, but it was not, as we used the cross-correlation-based auto-tracking in DLTdv8. We rephrased on L415 and L668-672 to be clearer on this topic.

That being said, I still believe that the ms would be clearer and more focussed with less details on the biological interpretation (and the many associated tests) of the Pieris example. I fear that trying to reach two goals (methodological advance in tracking using UAV and biological interpretation of Pieris tracks) may decrease the impact of the methodological advancements. Some results are of secondary importance to me and contribute to make a longer ms with many figures. This might also be a constraint if you want to try and publish this ms into a journal. To be clear, I let you to decide how you want to deal with that; I’ll recommend the ms whatever the choice you make.

Thank you for this advice on impact optimization. We are well aware of the « hybrid » nature of the manuscript (methodology + results). We agree that this makes the ms a bit long, but prefer to keep this self-contained format rather than e.g. splitting the study in two shorter articles. Also, this is part of the reason (aside community-based, diamond OA) why we submitted it to PCIecology (for publication in PCJ), because it seemed to us that this publication medium was probably more flexible in terms of format and content than traditional journals. In our experience, most traditional journals push a bit too hard towards format reduction and standardization, which

ends up being – in our view – somewhat detrimental to creativity. So thank you very much for allowing us to choose the final format of our work.