The manuscript has clearly been improved in some parts. Particularly the separation of results and discussion helped a lot. After reading the new version of the manuscript with a much extended discussion, the problematic points in the manuscript show up even more clearly – this is because now one can more easily separate the single experiments and the logics behind them. I have three major points:

1. Which regression method and closure time should be used to get the most accurate estimate of the real flux?

In my opinion, the most apparent pitfall is still the fact that the tests and experiments shown in the paper and the conclusions drawn from them are not well justified, but seem to be set up rather arbitrarily. For example, the question of linear vs non-linear regression is tackled by comparing fluxes calculated with both methods. Fine, but unfortunately you are not able to disentangle what is the effect of the used regression method only, and what part of the difference comes from the different closure times (Fig. 4A). Another example: I do not see where does the recommendation in abstract and conclusion of using chamber closure times of maximum 10 minutes come from? Why not 12 minutes? Or 20 minutes? Where is the basis for that conclusion? I do not find it in results or discussion. Is that coming from the "higher curvature during long chamber closure"? (see point 2 below)

I repeat myself by asking why not to study the effect of regression method separately, by using the same data and the same length of data? And why not to study the effect of closure time separately? Or why not to study the impact of both by gradually increasing the length of the data, separately for the linear and exponential regression? You do not have plenty of data, which is one of the weaknesses of the paper, but there should be enough to examine the questions of the best regression method and closure time properly.

One more example: the comparison of GC and QCL (Figs. 6 and 7). It is clear that the GC method needs the full 60 minutes, as all the four samples are needed. But what is your argument for selecting exactly the 3-minute length for the QCL data and comparing these two? In other words, can you show that 3 minutes is the best option? And why do you consider the linear fitting most optimal for the 3-minute data?

Same is valid for the point that you remove first 2 minutes. Can you show how much you affect your flux estimate by doing that? I know that in some papers the authors just shortly report that they removed a certain amount of data and used a certain regression method, without justifying anything (e.g. Savage et al. 2014), but as you are already comparing different closure times and regression methods in your paper, why not to do it in a proper way?

So to conclude this point, how do you justify that you use exactly the pre-defined closure times of 3, 10 and 60 minutes? This could be somehow explained by stating that these were the times which you have used and are going to use, and here you just want to show how the fluxes calculated with these closure times relate to each other. But this does not remove the problem that when comparing 3-minute lin fluxes to 60 (10) min exp fluxes you are not able to say whether the difference is due to the closure time or regression method. All you are now showing is that 3-min lin flux produces about the same values as 60-min exp flux, but it is not clear if any of these is the even close to the best estimate of the real flux.

If you, however, want to stick on the figures and results you are currently presenting, you have to give arguments for the relevance of examining such things, for example: "we decided to use 3-min linear calculation (because..?), and as for the GC 60-minute closures are typically used, we wanted to show that 3-min linear is as good as the 60-min exp". But you cannot say based on your analysis shown here that these are the best estimates of the fluxes, or that in other chamber systems it would be a good choice to use 10-minute closures

(as you now do in conclusions). Also, your 3 vs 60 min analysis is not very convincing: you show four flux values >200 µg (Fig. 4B and discussion on p. 12, lines 8-15), three of them destroy your nice correlation, and two out of these three are labelled with different color to indicate that these measurements were showing "steady linear start, followed by a sudden relatively sharp bend..". This emphasizes 1) the small number of flux measurements on which your analysis is based, and 2) there is something strange in your concentration development with high fluxes, if such distortions occur in all high flux values. A question arises, why you want to use such measurements and make generalizations from them, and, last, what is the reason behind them?

2. Curvature question

What comes to the questions of curvature and kappa, the conclusion that the closure time had an impact on curvature is definitely not justified by your data (Fig. 3B) – if I have understood right the type of data used for that experiment. You used different measurements (if I understand right the legend text of Fig. 3) for the two box-plots in 3B. Those with 10-min closure were done in different days ("before DOY 105.5 and after DOY 108.5") with highly different flux levels than those with 60-min closure. I do not see, how you can conclude that the differences in kappa were due to closure time, and not for example different flux levels or soil moisture (possibly affecting the curvature) or some other reason. The kappa question should definitely be addressed by using the 60 minute QCL measurements only, and by calculating the kappa for different closure times. E.g., increasing the closure time a few minutes at a time. If you then see that kappa increases (or actually decreases), then you can argue that using longer closure times you get more distorted concentration curves.

However, the question will still remain: what is the reason behind the changing kappa when data gets longer? It might be that it is an inherent tendency of the exp function to result in a steeper increase (or decrease) in c'(t) and c''(t) when t=0, with a smaller amount of data used, because of the higher relative noise in the data due to the shorter closure time. It does not mean, however, that using a shorter closure time is automatically the best choice, but vice versa it could mean that the exponential function causes more uncertainty when using too short data. Can you show which closure time is too short, in terms of kappa?

Abstract line 18: Why it was expected that the curvature was higher with long chamber closure than with 10 minute closure? In chapter 4.1 you state that "This was expected as the rate of transport …declines …because …decline in the vertical concentration gradient". Of course it is expected that the concentration change is non-linear! But this sentence does not justify, why kappa would change with longer closure. Please explain, why it is expected.

In general, the wordings and terms describing curvature in chapter 4.1 are used in a bit sloppy way. Please define clearly what you mean by curvature: that the rate of change in concentration is changing, i.e. c'(t) is changing (which is self-evident), or that the second derivative at t=0 is changing when you use longer closure time.

3. Conclusions chapter

Your conclusions section is a summary now, not really concluding anything. There are only two conclusions on lines 6-8 now, related to closure time (unjustified!) and the frequency of raw data (nicely justified). Please rewrite. Also, it is still not clear for me, which fitting method you prefer based on the results of this study? Linear or exponential? Could you indicate this in abstract and conclusions?

I recommend that the authors carefully consider the points presented in (1) above, and if they decide not to change much, at least they should add a justification and clarification what are the limitations in their analysis. The points (2) and (3) should be taken into account and corrected in the revised version.

Some detailed comments:

p. 3 lines 31-32: remove this sentence, not needed. A colon within a list followed by a colon is strange, remove.

p. 4 line 1 “density”, should it rather be sealing or tightness?

p. 10 line 17 onwards: what is ”increase characteristics”? Of what?

p. 13 lines 16-18: This sentence is difficult to understand. Perhaps a word is missing?

Figures 3, 4, 6 and 7: It is still challenging to read and understand the figures and the differences between them, since different closure times and regression methods have been used. I hope this could be solved by plotting different figures as suggested above.

In Fig. 4B: indicate what are the red circles here. Now the reader assumes they are “60 min and 10 min closure”, although it has been told in discussion that these are measurements with astrange shape in concentration increase. Are these same data as in 4A? If yes, add 10-min to y-axis legend. If not, why not? Same to 4D, which seems to consist mostly of 10-minute fluxes. (?)

It seems to me that figures 4B and 7 B are showing contrasting results: in 4B the 3-min linear flux is only half of the 60-min exp flux, while in 7B the 3-min lin flux is twice that of 60-min exp (mostly, since HMR was used). What is this telling us? Impossible to say, because so many variables are changing from figure to figure. This pair of figures crystallizes the confusion the reader have when spending time with the manuscript!!