Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-197-RC2, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "A novel injection technique: using a field-based quantum cascade laser for the analysis of gas samples derived from static chambers" *by* Anne R. Wecking et al.

Anonymous Referee #2

Received and published: 24 July 2020

In their submission "A novel injection technique: using a field-based quantum cascade laser for the analysis of gas samples derived from static chambers" Wecking et al. report on their experiences with a sampling strategy that analyses both eddy covariance data and chamber measurements on-site using the same insrument. They calculated fluxes of N2O from concentrations measured by GC (conventional) and directly injected into a QCL (novel) in three different fertilisation treatments (+ control). Both concentrations anf fluxes were analysed with a variety of statistical methods and found to be practically equivalent, with the exception of near-zero flux conditions.

I have thoroughly enjoyed reviewing this manuscript. It is well-structured, well-written,

and the figures and tables are polished to a degree that is rarely seen at the preprint stage. It fits the scope of AMT well and should be of value for the community after some clarifications.

Major points:

1) I partially agree with reviewer #1 that the overall idea looks a lot like something that could have been achieved in a simple comparison of concentration measurements. If you assume little error in the sampling itself and that the two analysers work with practically identical samples, there would be no reason to do this in the field, to generate the increased N2O concentrations via fertilisation instead of using standards, or to even calculate the fluxes at all (which are of course identical if the concentrations are identical). For an instrument comparison these would all be unwanted potential sources of error and confounding variables in the analysis.

That said, I think I see the authors' reasoning, which is to showcase that their idea actually works well in practice and for its intended purpose (measuring fluxes). It is an unfortunate truth that just because something works well in the laboratory doesn't necessarily mean that it must work well in the field.

I think this misconception is something that can be remedied quite easily by explicitly discussing early in the manuscript how and why this is much more than just comparing two instruments' ability to measure concentrations. It left me quite puzzled throughout half of the manuscript, because it only really becomes clear after reading and thinking about it for a while.

2) I don't really get the data workflow. What does the QCL output (shouldn't it directly be ppb?), what is QCL peak area in mV supposed to be and how was it translated into concentrations? Figure S1 also doesn't make a lot of sense to me due to this. What is "N2O calculated"? If you did a calibration with standards in glass syringes and those result in higher peak areas, how can you then accurately calculate concentrations for samples that were obtained with plastic syringes which apparently result in lower peak

areas...? Obviously it did work in some way or you wouldn't get so similar results to GC, but you have absolutely lost me somewhere on the way there. Expanding section 2.3 would help a lot. I would like to see basically a recipe to get from the QCL output to whatever you did in Fig. S1 (and further)

3) This is in regards to L327-339 - I fully admit I'm not overly familiar with bioequivalence statistics, but I have a strong feeling that you are boldly overstating what it can do. From what you wrote and what I could find in your sources, it's still just frequentist inferential statistics.

Don't get me wrong, I applaud that you are willing to do solid statistics and think outside the old t-tests-and-scatterplots box. But to me this honestly just looks like another type of null-hypothesis significance testing, with even more 100 % arbitrary (but hopefully consensus-based) thresholds and ranges. I.e. there is nothing objective and certainly nothing that justifies calling something a "proof" about it.

I suggest to word pretty much everything about bioequivalence with a bit less praise. It's a good and interesting approach and it makes sense to apply it to fluxes, but that's about it.

4) In section 3.4.2 you write about "using a QCL [...] without much disruption of other measurements". I respectfully disagree with that, consdering that in section 3.3 you say that you already need an initial lag time of 10 to 30 mins, so I assume you lose at least two half-hourly EC measurements for a single sample (How much is it actually? Please state it in the manuscript!). This is something that I, as someone working at an EC station, would not want to sacrifice at least during daytime and/or after significant management events (but I wouldn't want to inject gas samples the whole night either). You mention postponing analysis later in section 3.4.2. I would like to ask you to elaborate on this idea. Would it make sense to collect multiple samples and analyse them in one batch? Is this feasible in the field? In Tab. 1 you state that you can inject 200 samples per hour. For how long can the samples be stored on site (e.g. could it wait

until the next maintenance of the EC system excl. the QCL)? I would like to see some of your ideas on this and maybe an actual example for a sampling plan that minimises EC downtime. This disruption is a core issue for anyone doing EC, so it should play a much more central role in the discussion.

Minor points: * L97: Please give a justification for the very high application rates somewhere around here. * Check typographical rules for formulae, etc. Variables should be cursive (but _descriptive_ indices upright). * L163: "Since the quadratic fit suited lower C_N2O better than a linear fit, quadratic models were preferred [...]" The fit will naturally be better (in terms of R^2) if you throw more parameters at your model. Am I missing something here? * L203: Power depends on (among other things) the sample size. You can't just say a 90 % CI corresponds to 80 % statistical power. * L268: I think here you can replace "might explain" with "explains". At least to my understanding it's somewhat trivial that you calculate a larger flux if you measure higher concentrations with the QCL, no? * L314-315: Have you tested injecting blanks and see what happens? * L520: Typo "TAylor" * Figure 3: Panel c and d should have equal scaling on their respective x and y axes (i.e. the 1:1 line should be the diagonal). * Table S3: Where were the soil samples taken?

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-197, 2020.