

Response to Reviewer 1

General

We thank both reviewers for their thorough reviews and their wealth of suggestions. We greatly appreciate their input. We recognize in their reactions that the progress that we have made in the described research now opens up a large number of new questions: our research team felt exactly the same. With the reviewer's help we now harvested a number of good ideas to further optimize the data evaluation of this experiment, but more important: of experiments to come. Because facing the editor's request to reduce the length of the paper substantially does not allow expanding the evaluation on all aspects highlighted by the reviewers. We have tried to find a balance between which parts of the review comments we take on board now, and which parts we will take with us to our next campaign.

A main decision is how to deal with the footprint issues raised when evaluating the data. In the past, the Cabauw site has worked very well as a location for many flux intercomparison experiments with other gases (e.g. Peltola et al., 2014 <https://bg.copernicus.org/articles/11/3163/2014/>). However, this campaign has shown that it is probably somewhat too inhomogeneous for detailed intercomparison of ammonia fluxes. This was aggravated by the wind directions during most of the campaign, which were different from the normally prevailing southwesterly winds. Considering that (1) we did not collect the detail of activity data in the direct surroundings for a proper evaluation of the impact of different footprints on the fluxes of EC & AGM, and (2) that we are urged by the editor to reduce the paper length considerably, we have decided to limit our analysis in this paper to the standard, 3D-homogeneous flux approach. We are aware that with this approach some of the observed differences will originate from footprint issues. These issues will get more emphasis and attention in our next campaign, that will be located in a more homogeneous area.

In the remainder of this document we will follow the reviewers text, and our pointwise reactions and answers can be found in *italic blue text* at the applicable position.

On behalf of the author team,

Daan Swart.

General comments

This paper is reporting a comparison of the measurement techniques, flux gradient and eddy covariance, in measuring NH₃ deposition fluxes. The benefits of this study would be great for the gas measurement community, especially for the peoples who are interested in NH₃ emissions and deposition because the nature of NH₃ “sticky” character and having practical friendly large-scale field equipment make the field measurements difficult. However, there are some drawbacks in the current stage of the manuscript, I recommend authors to address the issues before considering to be published in this journal.

- (1) In the abstract, I would expect to see some clear messages including the measured NH₃ deposition fluxes measured by both techniques and the difference in the NH₃ deposition fluxes between these two techniques, especially during the periods when winds came from no-objects direction (green sector).

Our primary aim of the campaign was to test if both novel instruments were indeed capable of measuring the dry exchange flux of ammonia at high temporal resolution, to determine under which conditions, and for how long. Instrumental errors were found and corrected, and a large part of the data analysis was developed further during the campaign and its analysis afterwards. A thorough intercomparison of the results, focusing on the observed differences between the measurements of both instruments and the underlying causes is a large step further, would require considerably more effort, and was outside our scope for this first campaign. Actually, this first campaign was needed to demonstrate that such a study would now be feasible and meaningful, and to give us -at least to some extent- the clues on how to set it up and what additional information would be needed.

However, we agree with the reviewer that where possible additional indicative numbers are relevant to illustrate the flux comparison results. Hence, we have given those comparative numbers between line 25 and 27. To further clarify the two methods' differences in the green sector, we will provide an extra sentence in the abstract following Line 27, related to the green sector net flux:

“While differences in cumulative fluxes were small (~10%) as long as the upwind terrain was homogeneous and free of nearby obstacles.”

- (2) Secondly, in the methods and materials section, there is lack of information on the soil physical and chemical properties along the footprint or the transects from 100-200 m away to the tower, particularly within the green wind sector. The soil properties, including N and C contents, moisture contents, pH, soil texture, and grass types, crop age etc, could help us better understand the upwind footprint areas and how they could contribute to the measurements in this study.

We agree that such information would be needed to better understand the measurement results, and especially the differences between the techniques, given the fact that the effective footprints of both instruments are different. However, as stated above, this was not our primary aim for this campaign, and for this campaign we did not gather the information needed for this analysis. For this paper, we limit our analysis, assuming a 3D-homogeneous flux field and no footprint issues (knowing that this is not completely in line with reality). In fact, we think that the Cabauw site may be too complex an environment for such a more elaborate analysis. We would like to address this issue first in a simpler environment in our next campaign.

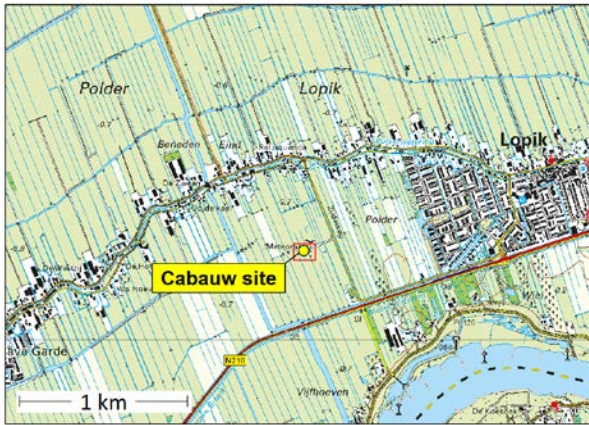


Figure 1. The area surrounding the Cabauw measurement site. Cabauw is in a flat area at -1m, being in the delta of the river Lek shown in the south east. The line with housing going east-west, running north of the tower has a series of farms. Map from www.pdok.nl/ (downloaded 07-02-2021).

To illustrate the different land cover classes, we will also add to the Supplementary Materials a simple land cover classification based on a Google Map (RGB) image, derived using Flux Footprint Predictions (<https://geography.swansea.ac.uk/nkljun/ffp/www/>). The Flux Footprint Predictions tool derives an unsupervised land cover classification based on a Bing or a Google (RGB) map for the footprint area and overlays the footprint climatology with the classification, providing a simple estimate of what land cover contributes most to the measured fluxes. The unsupervised classification will derive five land cover classes. This is only a very simple approach for land cover classification that cannot substitute a detailed analysis.

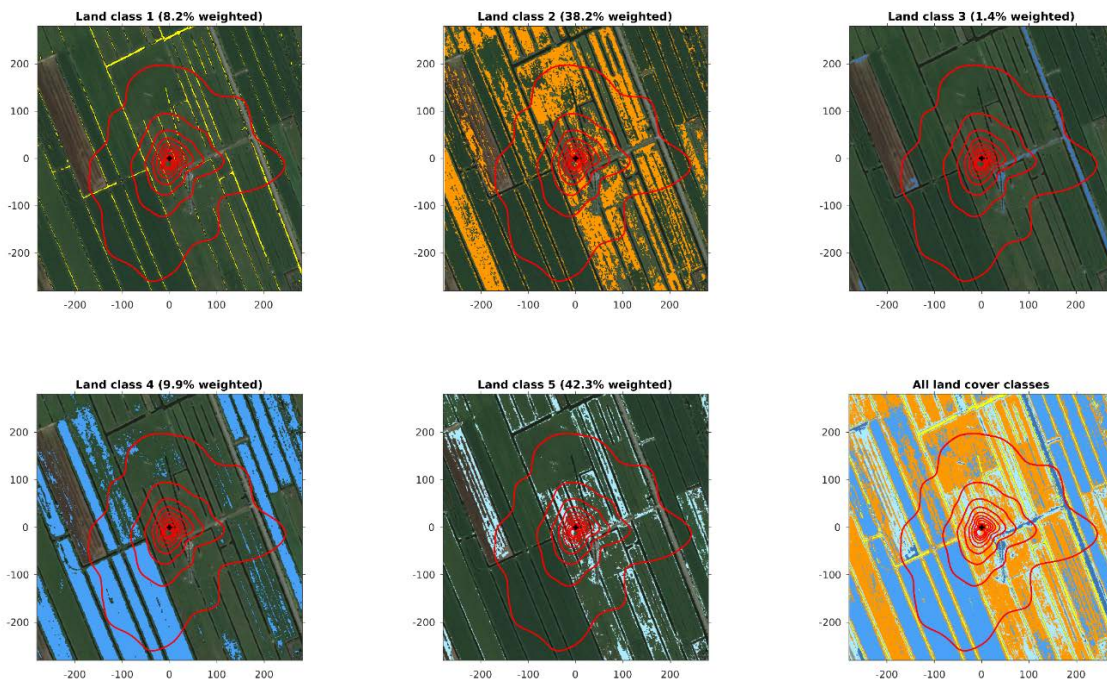


Figure 2. Five land cover classes and their distribution ratios under the EC footprint area, and the overall land cover classes. Percentages of each land class are shown in the individual land class diagrams. A comparison of the Google Earth image with the land cover diagrams suggests the following classification: Land class 5: wet grassland, in light blue (count 42.3%); Land class 2: less wet grassland, in orange (count 38.2%); Land class 4: hay land (harvested grassland), in bright blue (count 9.9%); Land class 1: ditches and drainage lines, in light yellow (count 8.2%); Land class 3: concrete road surface, in dark blue (count 1.4%); Land class 5 and 2 have a similar roughness height, while land class 4 has lower roughness.

The plot below also gives an illustration of the difference in management of individual paddocks in the flux footprint through time.

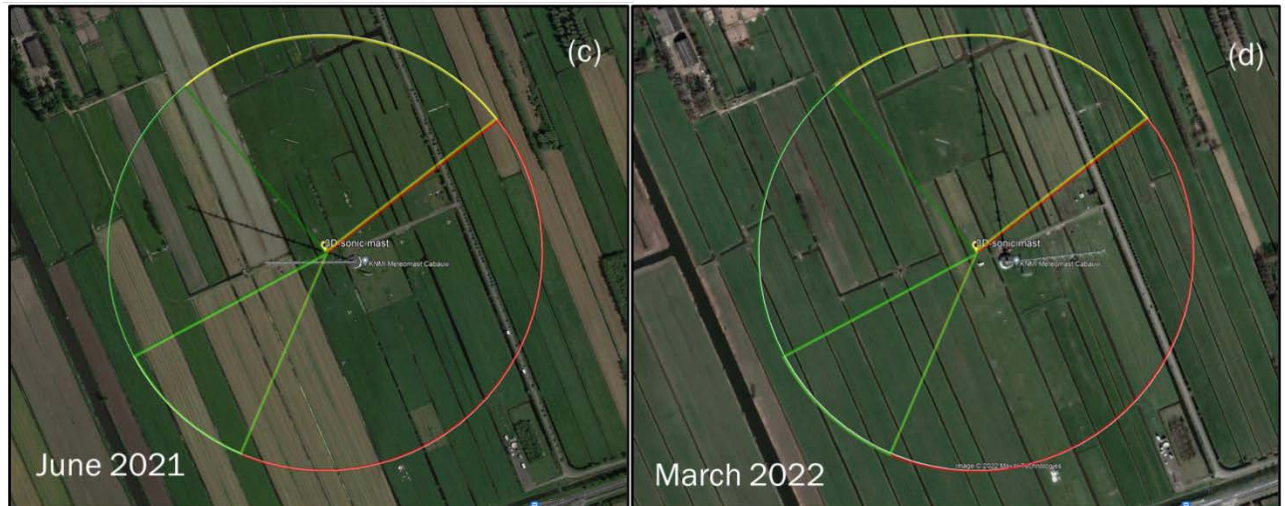


Figure 3. (c) land cover within a 300-meter radius in June 2021 (google earth image); (d) idem in March 2022.

- (3) Thirdly, in the results and discussion sections, authors spent a length to discuss the footprint. My main concern is that the two techniques have different footprints due to the height of the sensors, and different footprint area could contribute to the different NH_3 emissions and deposition (due to the land use and farm management practices). EC is often used at a larger area and but can't be deployed at many filed studies due to the limited size of the paddocks, which requires a difficult "footprint" analysis. I suspect there is lack of an accurate footprint modelling to correct EC measurements in this study. see some studies from Coates et al., 2018, 2021. Coates, T. W., Benvenuti, M. A., Flesch, T. K., Charmley, E., McGinn, S. M., and Chen, D.: Applicability of Eddy Covariance to Estimate Methane Emissions from Grazing Cattle, *J Environ Qual.*, 47, 54-61, 10.2134/jeq2017.02.0084, 2018.

We agree with the reviewer that the analysis of the EC and AGM techniques becomes more difficult if the terrain is complex in spatial structure and land use varies with time. We also agree that the two instruments have different footprints, which also vary in time. This will definitely be a source of differences between the results of the two instruments. In this study, we analyse our data assuming a 3D-homogeneous flux field, as is common practice in many flux studies, and -we feel- a good approach given the aim of our campaign. A next campaign focusing on a thorough intercomparison of the techniques can best be undertaken at a more homogeneous, simpler terrain without local emissions. We feel the Cabauw location is too complex for that.

- (4) Reviewer suggests that there is a need of adding a footprint analysis of AGM measurements as well. In addition to the factor of u^* , we may also want to look at the correlation between AGM footprint and stability length L , which also can tell us the footprint variations during night-time stable and day-time unstable conditions.

We agree that further work on the footprint analysis is needed for both instruments, and especially for the AGM (miniDOAS), as here we have data taken over two paths (not points) and at two different heights. As stated above, in the current study we assume a spatially homogeneous flux field, so in this analysis the footprint issues may be ignored (but not forgotten).

- (5) Further information on different paddocks soil properties and the history of farm management in the last couple of weeks (grazing, fertiliser application, irrigation, ploughing etc) are needed. Majority of N losses as NH₃ occurred at the first 2-3 weeks following N fertiliser application to the soils.

See our earlier remarks under (1) and (2).

Specific comments

- (6) 1, be aware of self-citation. There are too many times using two references Wang et al., 2021, 2022. I'm sure there are many other studies in this area.

The HT is a relatively new, commercially available instrument, and the Wang et al (2021, 2022) papers are currently the only peer-reviewed sources relating to this instrument. As testing the HT instrument is a core subject of this paper, we refer to these publications regularly, when relevant.

In addition, it should be noted that Wang et al. are not co-authors of our paper. Nor are we co-authors in their paper. So this is not a matter of self-citation.

- (7) 2, the manuscript is long. Please remove some repeated parts and shorten unnecessary contents, for example, line 134 to 137.

The editor and the other reviewer have also indicated that the paper needs to be shortened. We aim to reduce the size by eliminating unnecessary information and repetition, and by moving the footprint analysis largely to the supplemental material.

- (8) 3, add a detail map of experimental site including the surrounding terrains (200 m radius) and indicates equipment locations, heights, and the dimension of each paddock.

In the "Campaign setup and Site" section, the combination of the current Figure 1, with the addition of the figure mentioned in our answer under point (2) under cover both the locations of the instruments and the surrounding area.

- (9) Line 37. What does "right circumstance" mean here.

We will adapt the text to: "under relatively dry, low-dust conditions".

- (10) Figure 2. perhaps need number each instrument.

We have tried this, but the photo is already quite busy and we feel adding these numbers did not help to get a clearer overview of what is what.

- (11) Line 212. please explain why the path-length was set at 22.1 m (between miniDOAS and retro- reflector). Was it recommended by the manufactory? Is there any specific reason that

the distance between two miniDOAS paths (upper and lower path) was 1.53 m? would the measurements be better if the distance between the two paths is larger, such as 2.5 m?

*In the paper, we refer to (Berkhout et al., 2017; <https://amt.copernicus.org/articles/10/4099/2017/>) and (Volten et al., 2012; <https://www.doi.org/10.5194/amt-5-413-2012>) for a description and more in-depth info on the miniDOAS. The answers to the reviewer's questions are: 22 meter provides a proper balance between (1) loss of light (mainly by Rayleigh extinction, and geometrically) over the 2 * 22 meter path, and (2) sensitivity to ammonia (as determined by the amount of differential absorption in the fingerprint absorption lines). A shorter path would have more light but be less sensitive, a longer path would have less light but be more sensitive. Both parameters together determine the precision of the ammonia measurement. Path lengths between about 10 and 25 meters can be used in practice. In this case 22 meters was chosen as this is the default separation in the Dutch monitoring network, and fits best to our calibration facilities. The bottom path needs to be low, but well above the vegetation. The top path should be as high up as practically possible, as more vertical distance leads to better sensitivity for the flux-induced gradient. In our case the practical limit was the height of the measurement container.*

(12) Line 290 please provide the details of the functions for stable and unstable conditions.

The stability correction functions are described in the references given on the AGM theory. As mentioned in the manuscript, we use Paulson (1970; https://journals.ametsoc.org/view/journals/apme/9/6/1520-0450_1970_009_0857_tmrows_2_0_co_2.xml) for unstable conditions, and Beljaars and Holtlag (1991; [https://www.doi.org/10.1175/1520-0450\(1991\)030<0327:Fpolsf>2.0.Co;2](https://www.doi.org/10.1175/1520-0450(1991)030<0327:Fpolsf>2.0.Co;2)) for stable conditions.

(13) Line 307. It seems to me that the signal to noise ratio corresponds to the detection limit of NH₃ flux. What caused the signal to noise ratio higher in the study by Wang et al., 2022?

There is a change from 0.30 ± 0.05 ppbv in the first study, to 0.41 ± 0.06 ppbv in the second study. We cannot speak for Wang et al., but to us these figures do not appear too different, for different campaigns under different conditions.

(14) Line 311. was there any drifting of the release rate during the measurement period? if yes, what was it?

We consider the way the factory calibration was performed to be outside the scope of this paper. It was not performed by us, so we do not know the finer details. We will remove this part altogether and instead refer to their paper for this.

(15) Line 321-322. It is a concern that weather the HT8700 sensor can be used in a wet season as more NH₃ deposition will be expected in wet season than dry season.

We agree. Therefore we provided our feedback on this undesired limitation to the manufacturer. They are currently working to adapt the HT instrument such that operability under more humid conditions improves. We hope to test their next version soon.

- (16) Line 372. Provide the value for A, B, and C parameters. Are these values the same as that reported in Wang et al., 2021? Would these values be consistent or variable at different environmental conditions?

A, B, C are variables that change with air temperature, pressure and water vapour density. Values change every half hour. The procedure for this is described in Wang et al, 2021)

- (17) Line 379. More details about the stationarity and integral turbulence tests used in this study are needed.

In the text, we refer to Mauder and Foken (2006; <https://doi.org/10.1127/0941-2948/2006/0167>). The procedure is standard in EddyPro.

- (18) Figure 8. Many studies have shown that NH₃ emissions positively correlated with ambient temperature, higher emissions in the middle of the day (higher ambient temperature) than that at night-time, however from the figure 8, both miniDOAS bottom and miniDOAS top show the opposite pattern, higher concentrations at low ambient temperature and lower ones at higher temperature. Why? Perhaps add one or two sentences to mention that at night-time with low wind conditions, the concentrations can be build up and higher concentrations at night-time than day-time were observed. Suggest adding wind speed (or u^*) in the figure.

*We will correct an obvious error in line 488: the highest concentrations are observed during the night, not at noon. We will add some lines to explain the observed behaviour as requested. It should be well noted that higher **emissions** during daytime do not necessarily imply higher **concentrations** during daytime, as the mixing volume is completely different. See also our reply on (21).*

- (19) Line 524. I don't agree, as the better agreement was between 21-24 Sep, while much higher AGM before 18 Sep. Again, suggest comparing the concurrent fluxes from both techniques. Furthermore, should consider using the footprint model to correct the EC fluxes.

Effects of the spreading of manure are to be expected at least several days after the application, so if manuring stops at September 15, effects may be visible up till September 18 or even 20. That is, assuming all farmers obey the rules strictly. The green periods after September 20 (effectively 4 days) compare very well. We will change the line to clarify the exact period intended to: "In the green and light-green wind directions the NH₃ fluxes from the two methods compared very well after September 20. In this period little or no effects of manure application should be present."

- (20) Line 531. Is it necessary to calculate the cumulative flux on this particulate day when the winds were from the SE, large disturbance from the objects?

Figure S6 includes cumulative flux data only from the green and light-green wind sectors. So no winds from the SE are involved. We agree it may be unclear what we want to convey with this sentence. We will change the paragraph to:

"Considering only high-quality measured fluxes during this period, the cumulative daily fluxes of the AGM and EC were in general similar, with typical differences around 10%. When looking at the cumulative flux over the full period however, a larger difference is observed. This difference appears stepwise on a single day, September 24th. On this day, and only

during a few hours around noon, we see a much larger flux observed by EC compared to AGM. Most likely, the discrepancy is caused by footprint issues in combination with very local emissions. Unfortunately, we lack the means to validate this assumption.”

(21) Figure 11.

- a. 1) Figure 11. why is this typical NH₃ diurnal pattern different from the NH₃ concentration plotted in Figure 8.

Flux and concentration are different quantities. During daytime, turbulent mixing is stronger and the boundary layer rises. All surface emissions are diluted and spread over a larger volume, resulting in lower near-surface concentrations during daytime. On the other hand, emissions during daytime increase because (1) temperature is higher and (2) the vegetation can more easily release ammonia when outside concentrations drop. This results in a higher surface flux during daytime.

2) From Figure 9, the top panel shows most of the time EC measurements were higher than AGM, but in the lower panel it shows some time AGM higher than EC, other time EC higher than AGM, and some time they agreed well. However, in Figure 11, obviously AGM (in blue) is higher than EC (in red). Why? It is important for science aspect, it is worth to split the results, to identify in which conditions EC is higher than the red and when AGM is higher than EC.

The reviewer poses an interesting question here. If both techniques are properly calibrated, one would expect random variation in which instrument gives the largest measurement result. If on the contrary the occurrence of e.g. AGM > EC occurs more often related to a specific outside parameter or condition, this definitely points to a need for further study.

The fact that we see AGM to be systematically larger than EC around noon in figure 11 and figure S7 is such a pointer. Yet, it is not a conclusion but just a starting point for further study. The origin may be related to temperature, solar radiation, wind speed, wind direction, but also to instrument-related parameters in one or both instruments, like instrument temperature, straylight, or footprint issues.

We already identified that local manure application is a probable partial cause of the observed discrepancies, as the differences in fig S7 are smaller than in fig 11. But the mechanism of that needs further research. It is currently unclear why these observations are different.

To identify different causes and split them from each other, a much larger dataset of paired observations would be called for, combined with the observation of external parameters as indicated above, including the soil conditions and agricultural management of the footprint area and direct surroundings. Therefore, this calls for a longer intercomparison in a simpler environment. This is one of the aims in our next campaign.

- b. 3) Figure 11. please indicate if the same numbers listed on the top present for both technique measurements? If not, please add different numbers.

In this figure only the hours were used in which data of both instruments was present. So AGM and EC pairs are used. The numbers in the top row indicate the number of pairs used to create the average.

- c. 4) Figure 11. there is lack of explanation on the diurnal pattern. There is a “jump” in the deposition flux in the morning at 5 am, was it really happening or just due to the noise? There are not many available data at night for deposition flux datasets (both only 2 at 5 am and 22 pm).

We present this pattern as a sample result of our observations, as an example of which research area has now been opened by these observations. The exact explanation of the mechanisms behind these curves is a separate research project and not in the scope of this article. Yet, the overall behaviour is plausible, with moderate deposition during the night (humid soil, little convection) and stronger emission during daytime (drier soil, higher temperature, stronger convection, lower concentrations).

- (22) Line 595. I would like to see how the footprint are associated with the surface roughness z_0 and stability length L (stable and unstable conditions).

We expect little effect of z_0 variability on the footprint over the Cabauw site during the campaign. Kljun’s paper (<https://gmd.copernicus.org/articles/8/3695/2015/>) has discussed the sensitivity of the footprint for different values of z_0 . During the campaign, NW & SW patches of grassland were slightly shorter than other patches due to the early June cut. We estimate canopy height differences among grass patches from 5 cm to 10 cm. Assuming z_0 is 13% of canopy height, z_0 difference range from 0.0065 m to 0.013 m, which is relatively small. Ditches were less than 10% of the land cover, with a roughness length of 0 m and evenly distributed among grasslands. In all, z_0 differences and changes were few and are expected to have little impact on the footprint.

Stability length L does impact the extent of the footprint: The footprint is larger in stable conditions (ca. 300 m) and smaller (ca. 200 m) in unstable conditions, see the plots below.

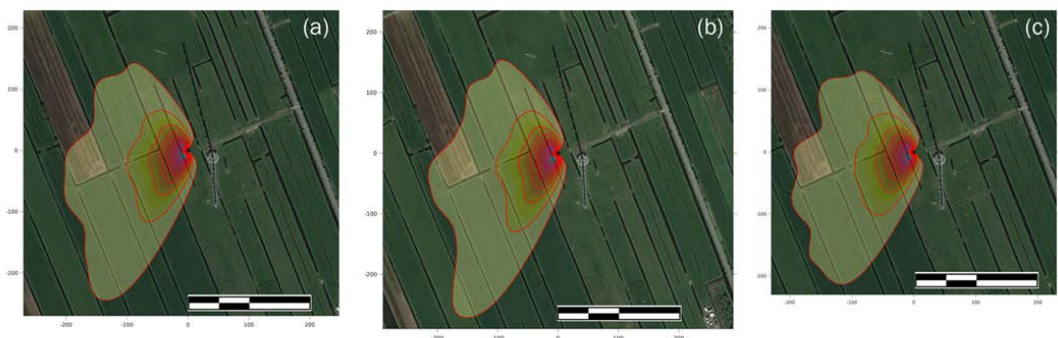


Figure 4. EC footprint within green and light-green wind sectors: (a) all data; (b) under stable conditions; (c) under unstable conditions. Note that the scale bar (50 m per unit) indicates that all figures share the same spatial scale.

(23) Line 660 please indicate it in the Figure 1.

Line 660 describes the DOAS container. This container is already indicated in figure 1, it is the centre of the graph.

(24) Line 802. Please provide a footprint analysis for AGM, the EC measurement should be corrected with a footprint modelling.

AGM footprint analysis: See our answer under (4)

Correction of EC flux measurement using footprint modelling: See also our reply under (1) and (2). Throughout this paper we have assumed the terrain relevant to our observations to be flat, with homogeneous vegetation, and the resulting flux field to be 3D-homogeneous. This is an approach that is commonly used in flux studies, but obviously it is an idealization that in practice does not hold completely. In this approach, footprint issues do not come into play.

During earlier campaigns for methane intercomparing CH₄ EC systems the Cabauw site proved to be really useful for that (Peltola et al., 2014 <https://bg.copernicus.org/articles/11/3163/2014/>) without the need for corrections using footprint modelling. However, specific for NH₃ the terrain turned out to be more complicated than we anticipated.

The exact shape of the footprint and its orientation on the map is determined by meteorological conditions and can vary rapidly with time. Combining and unraveling the mix of the effect of a varying terrain with a varying footprint would be the ultimate challenge for the analysis of our measurements. Correction of the measured flux data with a footprint model is however only possible if all relevant data of each part of the footprint are available throughout the campaign period. This also includes proper knowledge of all agricultural activity in all subplots. We did however not collect these data because we did not anticipate we would need them. So for this campaign we cannot perform the correction requested by the reviewer. Yet, with all data available it would still be questionable if we would succeed in getting a better comparison between the data of AGM and EC. For this, the Cabauw site is possibly too complex and diverse.

In our current analysis footprint issues will result in somewhat different measurement results from the two instruments, and also in somewhat imperfect flux figures. But the results still show that the two techniques can provide half-hourly flux measurements that appear to be consistent. That is the main message of the paper.

We agree with the reviewer that further studies into the footprint issues would take us a step further, especially for the AGM (miniDOAS) measurement. The campaign has shown us the relevance. We therefore will address these issues again in a future campaign in the Ruisdael observatory “de Veenkampen”. Here, the terrain will be much closer to the ideal homogeneous environment, and there will be no sources in the direct vicinity.

Technical comments

(25) Line 410. Remove “respectively”

Will do

(26) Line 413. Add a comma, to be 5%, 5%, and 1%, respectively.

Will do