Interactive comment on “Ammonia fluxes in relation to cutting and fertilization of an intensively managed grassland derived from an inter-comparison of gradient measurements” by C. Milford et al.

C. Milford et al.

Received and published: 7 April 2009

We thank the three reviewers for their evaluation and useful comments on the manuscript. In the following we address the points raised by each reviewer.

Reviewer 1

General comments: The reviewer notes that it would be interesting if authors could indicate what further measures (besides regular calibration and zero-flux test) could be taken to correct for any possible anomalous behavior of the AMANDA analyzer.

To answer this comment we draw attention to Section 4.1 Paragraph 1 which states:

S3344
"As concluded by Harrison and Kitto (1990), operator differences can induce the same amount of variation in NH3 measurements as different measuring techniques and, although techniques such as AMANDA and the WEDD have been shown to be reliable in measuring NH3, operators have to be vigilant in their running of these systems. A reliable clean deionised water supply, regular changing of pump tubing and regulation of instrument operating temperature are all essential to maintain the reliability of these systems." Additional text has been added to Section 4.1 Paragraph 3 indicating that as well as regular calibration and zero-flux tests, zero-concentration tests are also essential.

Specific questions: Questions about the AMANDA analyser

1) We consider that there are significant uncertainties in the AMANDA temperature correction, for this reason we would recommend that the AMANDA's temperature is controlled and excessively hot or cold temperatures are avoided. The same (exponential) temperature correction was used for all the AMANDA analyzers. Regarding the question of ventilation inside the monitor, the monitor does come installed with a fan, but in our experience, in hot summer temperatures as experienced during the Braunschweig campaign (daily maximums reaching 39 °C, Nemitz et al., 2009) this is not sufficient and extra ventilation should be provided.

2) The AMANDA analyzers were operated with a stabilising time of 120 seconds and an averaging time of 30 seconds.

3) The reviewer asks why we did not choose 100 second intervals; this was due to the need for a reasonably long stabilising time.

4) The fluctuations in the measurement of the air-flow during the measuring period were typically in the range 0 to 1 l min⁻¹, which with a mean value of 25 l min⁻¹ represents 0 to 4%. The air flow rate of each denuder is governed by a critical orifice in the outlet of the denuder which is designed to maintain the flow rate reasonably constant. There can therefore be differences in air flow rates between denuders and analyzers.
depending on the critical orifices. In the case of the analyzers used in this experiment, the three denuders of the FRI, FAL-D and CEH Amanda had mean air flow rates of 27.8 to 30.1 l min⁻¹, 23.0 to 24.1 l min⁻¹ and 25.1 to 26.6 l min⁻¹, respectively. The air flows were measured every two days or so throughout the campaign regarding the CEH AMANDA analyzer and slightly less frequently for the other analyzers.

5) The reviewer asks if three standards are enough to ensure a stable (linear) calibration curve. Yes. In fact, the calibration curve is not linear, so fitting with a curve is appropriate. The calibrations of the AMANDA analyzers were conducted approximately once every five days, or more frequently if a problem was suspected or any maintenance operation was carried out in the analyzer.

6) The AMANDA analyzer has a measurement range of 0.02 to 100 µg m⁻³, as indicated in Section 1 Paragraph 2. We have changed text in Section 4.1 Paragraph 1 to: "There was also close agreement achieved across the concentration range, except for some days where overestimation of concentration is suspected to have occurred in some of the systems, possibly due to high operating temperatures and inaccuracies in the measurement which could be exacerbated in the large concentration range, where the calibration is no longer as robust."

7) Aerosol capture is potentially possible, due to the horizontal geometry of the denuder. However, several tests, analysing the collection solution for Na which only originates from the aerosol phase, have shown that the aerosol collection efficiency is small for Na and completely negligible for submicron particles, which is the fraction that contains NH₄⁺.

8) Regarding the data coverage, the values of data coverage for each individual system before and after gap-filling are indicated in Table 2 ii) and iii) showing data coverage to range between 64 to 92% for the individual systems after gap-filling. It is only with the advantage of the four systems that the high data coverage of 98% is achieved. The authors do not think it would be possible to achieve a data coverage of 98% when mea-
suring over a full year with a single AMANDA analyzer due to the necessary downtime for maintenance and calibrations.

9) The liquid flow rate (approximately 1.0 ml min⁻¹, Section 2.2) in the AMANDA analyzer is tuned to the optimal performance of the instrument given the detector sensitivity and typical ambient ammonia concentrations.

Questions about the applied statistics

10) The robustness of the regressions is clearly indicated by our reporting of confidence intervals of the regression terms a, and b, which is more than is typically provided in such comparisons. Indeed, in linear regression more weight will be given in the slope to points at the ends of the distribution. This is not a problem, for such a simple descriptive comparison as we make, particularly since there is little evidence for log-normality in the distribution of the data themselves (about the regression line) as opposed to the log normal distribution in the frequency of the data availability.

Nevertheless, to satisfy the referee, we have also added into Table 4 the regression results calculated using log transformed data as: \( \log_{10}[\chi(1m)\text{individ}] = e \log_{10}[\chi(1m)\text{mg}] + d \). The values calculated are of course slightly different from using the linear regression approach, but as we expected, demonstrate the same qualitative picture that we already described. For the purpose of a simple descriptive comparison of the four systems, we consider that this should be sufficient.

Such a log transformation for the fluxes data is less suited, due to the existence of negative fluxes. The most negative flux, recorded for FAL-D was -315 ng m⁻² s⁻¹ and a transformation of \((\log_{10}[Fz(zo)\text{individ} + 320] = e \log_{10}[Fz(zo)\text{mg} + 320] + d)\) was tested. However, this did not succeed in normalizing the data, for this reason, only the linear relationships are shown for the fluxes.

The following text has been added to the paper accordingly. Section 3.4.2 Paragraph 1: "A comparison of \( \chi(1m)\text{mg} \) versus the individual systems was conducted (Fig. 4a,
b, c and d). The regression results are presented in Table 4, these include results calculated from the simple linear function and also using log transformed data."

Section 3.4.2 Paragraph 4 and 5: "The regression results for the log transformed data show lower values of r² with respect to the linear results. This is due to increased scatter of small concentrations near the detection limit in the log transformed dataset, resulting in a greater variance in the dataset for low concentrations. By contrast, in the simple linear plot (Fig 4) the variance in y is similar for the full range of concentrations, and therefore this is more appropriate for the comparison of the data than the log transformed data. Such a log transformation for the fluxes data is less suited, due to the existence of negative fluxes. The most negative flux, recorded for FAL-D was -315 ng m⁻² s⁻¹ and a transformation of (log₁₀[Fz(zo)individ + 320] = e log₁₀[Fz(zo)mg + 320] + d) was tested. However, this did not succeed in normalizing the data, while the variance was again larger at smaller values (due to scatter at smaller fluxes). For this reason, only the linear relationships are shown for the fluxes."

11) The reviewer comments that more departure is noticed at the lower concentration interval for the concentration measurements from the FAL-D analyzer. We consider that this comment reflects on the importance of keeping the statistical analysis relatively straightforward, since it simply reflects the fact that the performance of the ammonia measurement systems was not constant, but varied through the campaign. We already discussed this in Section 3.4.1 of our paper. The higher values of concentrations (and also fluxes) shown by the FAL-D system at low concentrations are already clearly indicated by the larger values of the intercept in Table 4 and 5. We accept however, that our presentation of Figs. 4 and 5 may not have been as clear as it could be, since the blue symbols were plotted over and hide the pink symbols. We have solved this by simply re-plotting Figs. 4 and 5 to show the results in 4 frames, with one dataset per frame.

12) The reviewer asks if there are not other statistical tools available. There are other tools, and more could doubtless be done for another paper. However, such a detailed
analysis of that kind is not the main purpose of our present paper. Rather, here we focus on using rather simple statistics to highlight the relative performance of the systems, noting that good agreement can be obtained, but that the system performance (and hence the agreement) varies with time. However, our target is not to provide a detailed statistical analysis, but to move forward to show how the multiple systems applied can be used to provide a more robust estimate of fluxes through the whole experiment.

13) It would certainly be of interest to conduct an error propagation study but this was considered beyond the scope of this paper.

14) The values in Table 8 were calculated from the mean concentrations and fluxes from the different analyzers for the different measurement periods.

Minor comments and technical corrections:

15) Reviewer 1: Page 4710, line 18: $u$ is mean horizontal wind speed: it is never mentioned in a formula.

Reply: This text is now removed.

16) Reviewer 1: Page 4705, line 16: the friction velocity is measured by several institutes: all derived from similar sonic anemometers (same brands, open-path or close-path)?

Reply: Different groups used different models of ultrasonic anemometers. A detailed comparison has been provided in a companion paper (Nemitz et al., 2009). This paper showed that differences in the friction velocity were not systematic but linked to spatial statistical variability in turbulence. Thus the use of an average friction velocity value from several instruments should greatly improve the robustness of the flux calculation.

17) Reviewer 1: Page 4713, line 1: Why was FAL-CH used as reference given the fact that data were not available during the pre-cutting period?

Reply: FAL-CH was used as the reference in this analysis because it was present
at Site 1 and also because it helps to illustrate the variation in the response of the FAL-D analyzer. Text has now been added to Section 3.4.1 Paragraph 1 to this effect. Although data was not available from FAL-CH during the pre-cutting period, the data coverage of the FAL-CH analyzer was an improvement on the FRI analyzer during the post-cutting and post-fertilizing period. Given the variation in response in the FAL-D analyzer and the lack of data availability during crucial post-cut and post-fertilizer periods of the FRI analyzer, FAL-CH was chosen as the reference in this analysis.

18) Reviewer 1: Page 4712, line 6: what is the average gap length?
Reply: Data regarding the gap-filling are given in Table 2 iii) for the individual systems. This shows that the median gap length varied from 15 min (FAL-D) to 90 min (CEH) as stated in the text (Section 3.1, page 4711, line 6).

19) Reviewer 1: Page 4714, line 13: a reference to table 5 is lacking. It is also not mentioned that CEH underestimates the flux by 32%.
Reply: A reference to Table 5 has now been added to Page 4714, line 12. It is stated in this paragraph (4714, line 15) that CEH underestimates the flux. It is stated in the Discussion (Page 4717, line 24) and Conclusions (Page 4720, line 20) that one of the flux measurements is within 32% of the mean estimate and reference is made to Table 5 where the values are given for each system.

20) Reviewer 1: Figure 4. Data from 3, 8, 9 and 10 June 2000 were not included. There was also substantial departure at 2 June 2000. Why data from 2 June were not discarded?
Reply: It was felt that data from 3, 8, 9 and 10 June 2000 were the most uncertain in terms of the differences between systems, other days also demonstrated variation but were kept in the dataset for the analysis.

21) Reviewer 1: Page 4719, line 12: nocturnal ammonia emissions. Is there a possibility of nighttime stomatal emission of ammonia due to stomatal aperture at night?
Reply: Although there is evidence that some plants (e.g. potato) do not fully close their stomata during night, we are not aware that significant nighttime stomatal conductances have been reported for the grasses that make up the agricultural field in this study.

Reviewer 2

Specific questions: Corrections on advection etc.

1) Regarding the first comment, uncorrected and corrected flux data are actually presented for the pre-cut period in the synthesis paper of the experiment (Sutton et al. 2009a), for brevity it was decided not to present the data in this form in this paper.

2) Regarding the second comment, we would refer the reader to the papers of Loubet et al. (2009) and Hensen et al. (2009) where these issues are discussed in more detail. In brief, it should be noted that while indeed there will be some random variation in emissions from the farm with time (e.g. over 20 minute periods related to individual gusting), the main effects can be clearly quantified, also with a diurnal variation in emissions from the farm demonstrated by Hensen et al. (2009). We consider that these estimates of emissions are fully sufficient to quantify the effects of advection on the vertical fluxes.

In addition, we would like to highlight various points: i) Although the advection errors are estimated with the FIDES model based indeed on an "unknown source", the paper from Hensen et al. (2009) shows that the FIDES model gives rather good estimate (within 30

ii) The FIDES model is used in a two-step way where first the source strength from the farm is evaluated every 30 minutes using the 30-min concentration at Site 3 (220 m downwind from the source) and then used as an input to evaluate the advection error in field. Hence the estimated farm source is constrained by the observed concentration difference between Site 3 and the background. This is a major constraint, which, even
though the source location is not well known, ensures that the quantity advected is correct.

iii) One should distinguish the pre-cut period, during which the farm is the major source of local advection errors (together with deposition on the field), and the post-cut and post-fertilisation periods, during which the field itself is the major source of local advection (the concentration increase above the field is due to the field source itself). During this period the source strength (at the ground) is estimated by the FIDES model, and it correlates very well with the source estimated by the gradient method, hence indicating that the source is well determined in this case also.

iv) Although dispersion is indeed harder to model during night time (with low wind speeds), NH3 emissions either from the farm or from the field are often also small during night time due to both lower temperatures (affecting the source potential) and lower wind speed (affecting the transfer resistance). Figure 7 from Hensen et al. (2009) and Figure 5 from Loubet et al. (2009) demonstrate this quite clearly.

v) Yes, the four systems were corrected individually and the corrections did not account for the agreement or disagreement of concentrations and fluxes.

3) Reviewer 2: Considering the uncertainties in the advection calculations, are these corrections relevant considering the uncertainties in the aerodynamic gradient method and the measurements of ammonia concentrations?

Reply: Our results, as pointed out in our paper, and recognized by the referee, do indeed highlight that the variation in instrument performance in the gradient methods often gives more uncertainty than the magnitude in the advection correction (and its associated uncertainties). This does not make it irrelevant to calculate the advection term, but rather puts the advection corrections into the context of the overall uncertainties.

4) Reviewer 2: It is mentioned that measurements at site 1 and 2 are used for advection
considerations Is this reasonable when the concentrations measurements sometimes deviate significant between the measuring systems?

Reply: The "measured" local advection error was rather sensitive to the difference between Site 1 and Site 3 where the concentration difference was the largest, and the modelled advection error was rather sensitive to the difference between the concentration at either Site 3 minus background (advection due the farm emissions) or Site 1 minus background (emission due to field emissions), a difference which was significant.

5) Reviewer 2: Could there be an inhomogeneity in the emission from the field (up wind emissions suppressing the down wind emission area), leading to "true" differences between site 1 and 2?

Reply: In Loubet et al. (2009), the advection error estimated with a constant surface flux hypothesis and a constant surface concentration hypothesis did not differ significantly at Site 1, which indicates that the source strength did not evolve with downwind distance at Site 1, and henceforth at Site 2.

6) Reviewer 2: I assume, that all these corrections are made to correct for the lack of constant flux layer. Why is the flux then calculated to the reference height $z_0$?

Reply: Due to the lack of constant flux layer under situations of advection, the true flux is the flux at the surface. Hence it is appropriate to report the fluxes at reference height $z_0$, rather than e.g. $z(1m)$ which does not necessarily reflect the flux at the surface.

7) Reviewer 2: Average time for fluxes What are the considerations choosing an average time for the fluxes to 15 minutes and is that period sufficient long time during stable and unstable conditions?

Reply: An average time of 30 minutes is typically used for flux measurements above forest. This is similar to an average time of 15 minutes used in this study, at a much reduced measurement height, as the eddy frequency scales with $(z-d)$. As with the forest studies, it is possible that some flux was lost due to the relative short averaging
time. However, shorter averaging times minimise the effect of non-stationarities and provide highly time resolved information to study the processes, a key objective of this study.

8) Reviewer 2: Measuring systems. Only the deviations on concentration differences are given in Table 3. To my knowledge the calibration curve of the AMANDA is not linear. Does this influence the deviations on gradients/differences?

Reply: The % differences are given along side the reference values for comparison. Therefore it is in fact possible to relate these to the absolute values. E.g. for Test 1 (25 May 2000) the Delta value of the standards was 84-22 mg/l = 62 mg/l. The FRI system overestimated by 44%, therefore, the Delta derived from the FRI measurement was 89 mg/l. From the data in Table 3 there was no clear evidence that absolute concentration affected the systematic differences between the instruments, although the % differences were smaller on 6 June, associated with a larger ppb signal to be detected.

9) Reviewer 2: What does a non-linear calibration curve have of consequences measuring "out of the calibration range" (p. 4718 line 27)?

Reply: We agree that this statement is confusing and have changed Section 4.1 Paragraph 1 to: "There was also close agreement achieved across the concentration range, except for some days where overestimation of concentration is suspected to have occurred in some of the systems, possibly due to high operating temperatures and inaccuracies in the measurement which could be exacerbated in the large concentration range, where the calibration is no longer as robust."

10) Reviewer 2: What does the QC concentrations correspond to in atmospheric concentrations?

Reply: The QC concentrations correspond to atmospheric concentrations in the range 0-15 µg m⁻³, depending on the exact liquid flows and air sampling rates of the different
instruments.

11) Reviewer 2: Why aren’t the absolute concentration determinations of the QCs in table 3 given since atmospheric concentration measurements are compared?

Reply: The Quality Control standards were prepared as aqueous standards. The atmospheric concentrations corresponding to the QCs depend on the actual air flow and liquid flow rates of each analyzer at that time. As these vary between analyzers, the QC standards are given as µg l-1.

12) Reviewer 2: It is written, that the AMANDA measures 150 s in each of the three heights, having a full profile in 450 s (p. 4707 line 6-7). How is "carry over" in tubes/detector handled when shifting heights/sample? Is "carry over" - if any - equal at the whole concentration range?

Reply: As noted in response to Reviewer 1, the AMANDA analyzers were operated with a stabilising time of 120 seconds and an averaging time of 30 seconds. The stabilising time accounts for this "carry over" when switching between heights/sample. The "carry over" would be greater at greater concentrations, the stabilising time of 120 seconds is considered sufficient to cater across the whole concentration range.

13) Reviewer 2: Concentrations and fluxes Some of the deviation in the fluxes is caused by deviations in concentration determinations by the four systems. Could X*/X (ratio of concentration scaling parameter and concentration) be compared in order to reveal, how much influence the deviations in concentration explain of the flux differences?

Reply: The simple approach to address this is to compare the concentrations and fluxes measured by the different systems. As F =-u*χ*, since the same u* is used for each instrument, the % divergences in F/χ are the same as -χ*/χ. The comparisons in Figure 4 and 5 show that, as expected, in general the difference in F was roughly proportional to the difference in χ, as shown, e.g. for FAL-D and CEH. However, the difference was not precisely proportional (c.f. FAL-CH and FRI in Figures 4 and 5),
highlighting the different performance of the measurement systems, as well as their variation through the experiment.

14) Reviewer 2: Regressions Making regressions, the data need to follow a normal distribution and if not, one should transform them. It is mentioned that the data are log normal distributed (p. 4715 line 7) and therefore, they should be transformed.

Reply: We would refer the reader to the response to reviewer 1 on this topic. To address this point we have also added into Table 4 the regression results calculated using log transformed data as: $\log_{10}[\chi(1m)\text{individ}] = e \log_{10}[\chi(1m)\text{mg}] + d$. The values calculated are of course slightly different from using the linear regression approach, but as we expected, demonstrate the same qualitative picture that we already described. For the purpose of a simple descriptive comparison of the four systems, we consider that this should be sufficient.

Such a log transformation for the fluxes data is less suited, due to the existence of negative fluxes. The most negative flux, recorded for FAL-D was -315 ng m$^{-2}$ s$^{-1}$ and a transformation of $(\log_{10}[Fz(zo)\text{individ} + 320] = e * \log_{10}[Fz(zo)\text{mg} + 320] + d)$ was tested. However, this did not succeed in normalizing the data, for this reason, only the linear relationships are shown for the fluxes.

15) Reviewer 2: It is a choice to compare FAL-D, CEH to FAL-CH (figure 3), though FAL-CH do also have a measure error. Therefore, a total least square (orthogonal regression) would be more correct in this analysis between systems. Choosing FAL-CH as the reference, the period before cutting is left out. This should be stressed out. Why is FRI not compared to FAL-CH?

Reply: This would be a fair point if indeed we showed the statistics for these individual comparisons in Figure 3. However, in fact we do not report the individual regressions, but simply use Figure 3 to highlight the differences between the datasets. For example, in Figure 3 it is very clear that the performance of FAL-D compared with FAL-CH varies with the data forming widely spread groups. Nevertheless, to address this point, we
have removed the regression lines from Figure 3, and retitled the heading of Section 3.4.1. "Comparisons of $\chi(1m)$ and Fz(zo) estimates against FAL-CH". Text has been added to Section 3.4.1 Paragraph 1 to note that FAL-CH data was not available for the pre-cutting period. FRI was compared to FAL-CH but the data was not shown due to space constraints, see text in Section 3.4.1 Paragraph 1.

16) Reviewer 2: Meteorology More information on meteorological conditions (a figure with i.e. temperature, wind speed/friction velocity, wind direction, precipitation) would be welcomed and more analysis of the concentrations and fluxes in relation to meteorological conditions would be very interesting, giving the data and paper a more general aspect, both with respect to other experiments and the parameterization of model calculations.

Reply: Regarding the comment that more information would be welcomed on meteorological conditions, information on precipitation, wind speed, relative humidity, canopy and air temperature, global radiation, wind direction and friction velocity, is provided in a detailed form in the introductory paper of this special issue (Sutton et al., 2009b) and also in a companion paper of the special issue (Nemitz et al., 2009). The authors did not feel it was necessary to repeat this information in this paper. In addition, measured fluxes were compared with three models: an ecosystem model, a soil vegetation atmosphere transfer model and a dynamic leaf chemistry model. Companion papers in the special issue (Burkhardt et al., 2009; Personne et al., 2009) and the synthesis paper (Sutton et al, 2009a) provide full details of these modelling analyses.

17) Reviewer 2: More explanation p. 4720 line 10-12 is written: "Substantial emission also occurred at night immediately after fertilizer application (5-6 June), demonstrating the importance of surface emission from the soil and litter surface". Could this be elaborated more? As I understood, the cut grass was removed before the field was fertilized?? Why would the soil emission increase after fertilization?

Reply: Increased ammonia emissions after the application of mineral nitrogen fertilizers
are well known and have been demonstrated in many studies. The occurrence of nocturnal emissions (when stomata are largely closed) highlights the role played by direct fertilizer volatilization either from the soil surface or fertilizer adsorbed on to litter and leaf surfaces.

18) Technical corrections. All technical corrections were corrected following the suggestion of the reviewer.

Reviewer 3

In response to the first two comments the following text has been added to Section 2.1 Para 1: "It should be noted that recent research (Flesch et al., 2002; McNaughton, 2006) has questioned the validity and accuracy of the aerodynamic gradient method."

\( \chi(1m) \) was calculated by interpolation of the profile taking into account all three concentration measurements, for each profiles, then averaging the results.

Eq (2) has been modified.

The amount and form of the nitrogen fertilizer is already mentioned in Section 2.3, Page 4708 L20:

References:


All other references refer to companion papers in this special issue: "Processes controlling the exchange of ammonia between grassland and the atmosphere (GRAMINA)."

Interactive comment on Biogeosciences Discuss., 5, 4699, 2008.