Answer to Comment to the paper „Evaluation of a new inference method for estimating ammonia volatilisation from multiple agronomic plots "by Benjamin Loubet et al. by Albrecht Neftel (Neftel Research Expertise, CH-3033 Wohlen b. Bern, CH, Switzerland) and Christoph Häni (School of Agricultural, Forest and Food Sciences, Bern University of Applied Sciences, CH-3052, Zollikofen, CH, Switzerland)

This paper presents an appealing and low-cost approach to determine NH3 losses from adjacent multiple agronomical plots by a combination of concentration measurements using passive sampler devices and a dispersion model that is driven by turbulent parameters inferred from standard 30 minutes meteorological data.

The aim of the paper is described as: "Can inverse dispersion modelling approaches be used for inferring NH3 emissions from multiple small plots (agronomic trials) using passive samplers, and to which degree of accuracy?"

The overall answer is encouraging with the statement in the conclusions "In this study we have demonstrated that it is possible to infer with reasonable biases ammonia emissions from multiple small fields located near each other using a combination of a dispersion model and a set of passive diffusion sensors which integrate over a few hours to weekly periods".

According to our judgement the accuracy will be mainly determined by two aspects addressed below.

a) Bias related to the applied dispersion modelling

Dispersion models are a simplified mathematical representation of the turbulent motion in the surface layer and will always deviate from reality. Systematic biases can be expected when modelling the lower heights of the measurements that are discussed in this paper. For concentration sensors place close to the ground (e.g. 25cm above ground) transfer functions are likely to be biased due to e.g. the needed simplifications that must be made to describe the exchange process at the ground, the natural heterogeneity on a small scale at the surface or the violation of model assumptions such as the failure of K-theory close to the canopy (Raupach and Legg, 1984). Furthermore, the translation of the sensor height in a model framework is challenging for very low heights since the sensor height value (and with that the resulting value of D at that location) becomes very sensitive to sensor height measurement errors as well as to the absolute values of z0 and d. To us, a sensor height of 25 cm seems too close to the surfaces.

We do agree with A. Neftel and C. Häni concern that model representations may be biased close to the ground, and especially Gaussian like models, since they do not intrinsically account for near-field dispersion as shown by Raupach and Legg (1984) and subsequent publications from M.R. Raupach (Raupach, 1987; Raupach, 1989b). This is not the case of the Langevin models that account for near-field dispersion (Raupach, 1989b; Thomson, 1987). However, as exposed by Raupach (1989a), the height at which the near-field effect is sensible would be equal to $\sigma_w T_L \sim 0.3\ \sigma_w h / u_* \sim 0.3 \times 1.25 \times h \sim 0.4\ h$ where σ_w is the vertical air velocity standard deviation, h is canopy height and T_L is the Lagrangian time scale. Numerical values were derived from Raupach (1989a), Figure 1. Hence we would expect this far field effect to be small on situations with small canopies (or by extension with

small roughness height for a bare soil). Typically this would correspond to about 5 cm above a 10 cm canopy and ~1 cm above a 3 cm roughness. Hence we agree with A. Neftel and C. Häni that this would represent a quite important fraction of the sensor height if this sensor would be placed lower than 50 cm (20% for 25 cm). This would especially be critical for canopies that are taller than 10 cm.

Regarding the uncertainty in determining the height of the sensor close to the ground, this is a very sound remark. We however see in Figures 6 and 10 of the manuscript that the method gives similar biases for sensors placed at 25 cm and 50 cm above ground (and also close biases for h = 1 m), for plots of 25 m x 25 m. We hence agree with the concern of A. Neftel and C. Häni that 25 cm would be too low and we should rather target heights of 50 cm. We propose to add this statement in the conclusions: *"Although the lowest sensors have the best condition number, we would rather recommend in practice using heights of 50 cm above the canopy in order to reduce uncertainty in positioning the sensors close to the ground as well as avoid being too close to the roughness layer close to the canopy which is characterised by non-diffusive transfer."*

The authors are using their FIDES-3D model that is based on an analytical solution of the advection-diffusion equation. This model is compared with the backward Lagrangian Stochastic dispersion (bLS) model described in Flesch et al. (2004) (the "WindTrax" software, Thunder Beach Scientific, Nanaimo, Canada). For the presented analysis the FIDES model Kz was adopted to match the far field approximation of Kz of the bLS model. We are missing an explanation, why this was done.

The aim of matching the far-field diffusivity of the two models was to make the FIDES model consistent with the bLS approach, which is a commonly used method nowadays for estimating ammonia emissions with inversion techniques. As exposed in the manuscript, one major difference arises from the fact that the Phillip (1959) approximation of the advection-diffusion equation (which is identical to the approach of Korman and Meixner (2001)) has a Schmidt number equal to 1, while bLS approaches have a Schmidt number equal to 0.64. Wilson (2015) showed that the choice of the Schmidt number has a great effect on footprint, and hence on concentration above a small source and should therefore be explicitly given. Wilson further showed that the difference in footprints predicted by diffusion and Langevin models (like bLS) are small under neutral and stable conditions provided they have similar Schmidt numbers, although the difference in footprints remains large under unstable conditions even with identical Schmidt number (with the Langevin models diffusing less than the Eulerian ones). We hence chose to use an approach that was as close as possible to the bLS approach. To do so, we matched the far field diffusivities of the two models, as this would ensure that the two models would provide similar concentrations at heights larger than a few decimetres. Moreover there was no point in matching near-field dispersion, as FIDES does not account for near-field dispersion.

In the supplement, a detailed investigation is presented how the two models differ in their formulation of the vertical diffusivity Kz. The assumed far field vertical diffusivity in the bLS model is approximated by parametrizations provided in Flesch et al. (1995). We would like to remark that WindTrax uses slightly different default parametrizations of σ_w than provided in Flesch et al. (1995) (see e.g. the manual on the WindTrax homepage). This is resulting in vertical diffusivities given as:

$$Kz\,(Z) = 0.5 * 1.25 * u*Z/(1 + 5Z/L) \text{ for } z/L \geq 0 \qquad \text{(WT1)}$$

$$Kz\,(Z) = 0.5 * 1.25 * u*Z \times (1 - 6Z/L)^{0.25}\,(1 - 3Z/L)^{(1/3)} \text{ for } z/L < 0 \qquad \text{(WT2)}$$

with a Schmidt number value of $Sc \cong 0.64$ for near-neutral stabilities with a smooth transition from L = ∞ to L = -∞. These equations differ from the equations S7 and S8, and imply a different interpretation of the differences between FIDES and WindTrax, though without changing the numeric results of the comparison.

We would like to thank C. Häni and A. Neftel for providing the exact expression used in Windtrax. We indeed only referred to the work of Flesch et al. (1995). These equations WT1 and WT2 are more consistent than those we reported since, as opposed to equations S7 and S8, they insure continuity in $Kz(Z)$ when $z/L \to 0$. We see from the set of equations WT1, WT2, S7 and S8 that $Kz(Z)$ is similar under non-neutral conditions in S7 and WT1 though 4% smaller in Windtrax but that $Kz(Z)$ is 16% smaller in Windtrax (WT2) than in the "tuned FIDES" (S8) under unstable conditions. This therefore explains better Figure S3 which shows a good fit between the "tuned FIDES" and WindTrax under stable and neutral conditions but a lower concentration modelled with the "tuned FIDES" at 2 m above ground under unstable conditions. Indeed, since the far-field diffusivity is larger in the tuned FIDES, this model lead to larger diffusion and hence lower concentrations away from the source. However, another difference comes from constitutive differences between Eulerian and Langevin models under unstable conditions as shown by Wilson (2015).

We have checked that the results reported in this manuscript remain mostly unchanged since already in line with the most important feature of $Sc \sim 0.64$. We however quantified a difference of around -18% ± 10% in the concentration modelled using equations S7 and S8 (Flesch et al., 1995) compared to equations WT1 and WT2 (Flesch et al., 2004) for a single source of 25 m x 25 m. Since this difference is systematic and since we use the same model for forward and backward modelling we do not expect any impact on the conclusions we have drawn from this study. Indeed, we checked for a single source of 25 m x 25 m that the biased inferred using eq. S7 and S8 and WT1 and WT2 were similar within less than 1% for most cases. Noticeably, the biases for the highest sensors were diminished with WT1 and WT2. We hence propose to leave equations s7 and S8 in section S4.2 as they are but to stipulate explicitly that these correspond to Flesch et al. (1995) and not to WindTrax. We also propose to modify Figure S3 to explicitly mention this (see new Figure 3 below).

In the supplement Figure S3 presents a comparison between evaluated concentrations with FIDES and WindTrax respectively using the prescribed emission sources with the SVAT model. This figure is hiding the apparent differences as a double logarithmic representation is used and the concentrations are shown using an emission source that shows a positive correlation between the meteo input parameters of the models and the source strength. E.g. for neutral conditions the regression of ratio of the concentrations calculated with FIDES and WindTrax at a height of 0.25m is indicated as $cFIDES = 0.97 \cdot cWindtrax^{0.87}$. For a concentration of 1 the ratio is 0.97 and for a concentration of 100 the ratio becomes 0.53. As the transfer function D in FIDES and WindTrax are only depending on the prevailing turbulent parameters it would be more instructive to use a constant unit emission of 1 and show the ratio on a linear scale as function of u_* and L in a similar way as the authors have done in a previous paper (Carozzi et al., 2013).

We thank A. Neftel and C. Häni for this very useful comment. We propose to modify Figure S3 and section S4.3. Below is given the proposed updated Figure S3 showing the comparison between the two models using linear regressions forced to zero and graphs with linear scales. We propose to change the text in the supplementary material to the following: *In Figure 3, we notice that the concentration modelled with the tuned FIDES at 2 m height was lower by roughly 15% compared to WindTrax under unstable conditions but is comparable under stable and neutral conditions. Lower down at 0.25 m height, the tuned FIDES systematically underestimated the concentration by 15-22% whatever the stability.*

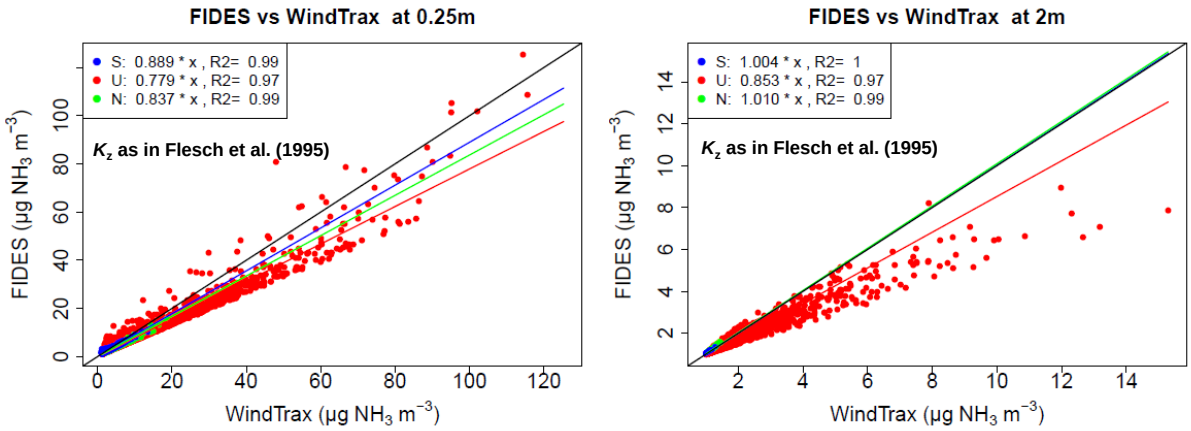

Figure S3. FIDES versus Windtrax concentration modelled above an ammonia source of 25 × 25 m at 0.25 and 2 m heights. The FIDES vertical diffusivity K_z fits the WindTrax K_z. The comparison is made over the entire year of 2008 in the FR-Gri ICOS site. S, U and N stand for stable, unstable and neutral atmospheric conditions. The linear regression equation is given for each condition together with the R^2 of that regression. The black line is the 1:1 line.

Below we also computed the graphs showing the concentration residual as a function of 1/L, as suggested by C. Häni and A. Neftel (Figure S4). We propose to include Figure 4 and the following text in the supplementary material: *Figure 4 shows that under unstable conditions FIDES underestimated the concentrations up to 30% at 0.5 m compared to WindTrax, while this gap was reduced and more scattered at 2 m height. Moving towards neutral conditions the two models tend to agree notwithstanding an overestimation of 10% by WindTrax at 2 m height concurrently with an underestimation of the same magnitude by FIDES at 0.5 m. Under stable conditions there was a good agreement at 2 m height, while this agreement remains poorer at lower heights.*

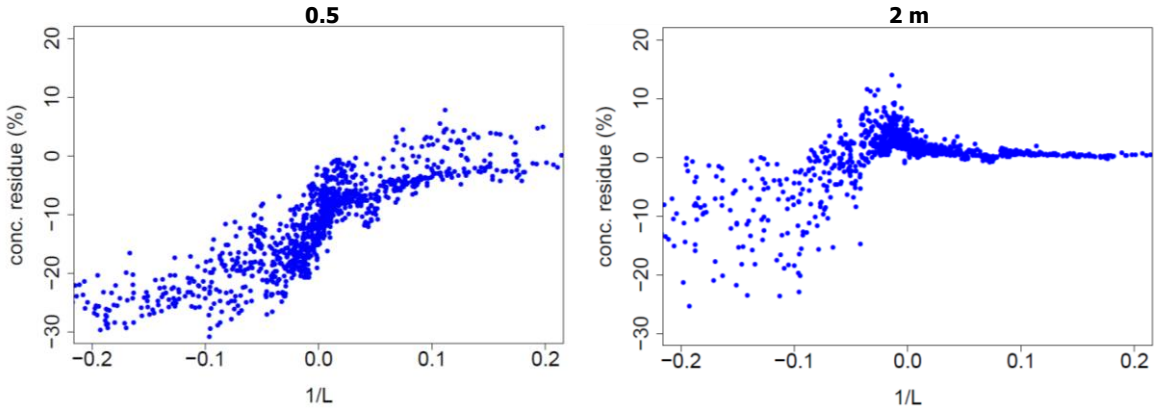

Figure S4. Relative difference between FIDES and WindTrax concentrations as a function of the stability parameter (1/L). Data refer to the same conditions reported in Figure S3.

b) Bias related to the concentration measurements

The use of passive diffusive samplers is a challenging business. Within different networks the reliability of PS such as ALPHA samplers or Radiellos have been proven, but the use of them close to emitting source showed major deviations compared to other measurements. E.g. Misselbrook et al. (2005) found severe overestimations of passive diffusive samplers. The latest investigation stems from the Dronten experiment and is discussed in a paper by Michael Bell et al. (submitted to AFM). In this experiment the ALPHA samplers were affected by a positive bias in the order of 50% relative to the other devices. We speculated that the exposure of the PS with the protection hat above them cached eddies from below loaded with higher NH3 concentrations but shielded eddies with lower concentrations from above. Figure 1 illustrates the NH3 dynamic that occur over an emitting surface. The concentration was measured with a fast device described in Sintermann et al. (2011). Immediately after application of slurry with a splash plate the NH3 concentration was measured at a height of 1m above ground with an ionization technique and a strongly heated inlet line to avoid as much as possible damping effects.

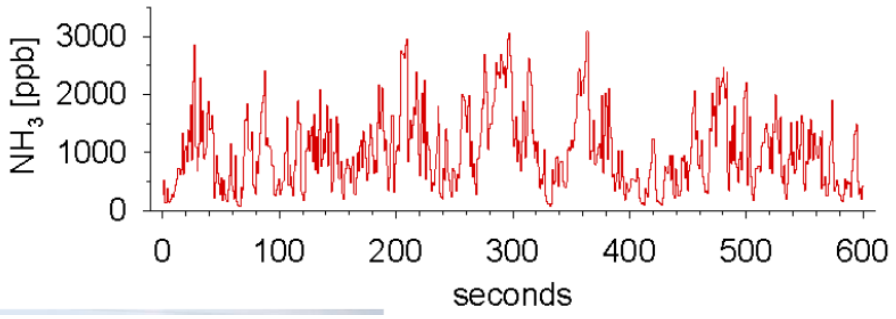

Figure 1: NH3 concentration time series measured 1m above ground over a manured surface with splash plate.

First we notice as A. Neftel and C. Häni that ALPHA badges are very reliable for network measurements concentrations, as for instance showed by the recent Met-NH3 project (http://www.metnh3.eu). We are however very much aware that the use of ALPHA badges close to emitting sources may be biased. The reason for that bias is however unclear and would need comparisons with fast and unbiased sensors. The assessment by Misselbrook et al. (2005) shows that with high concentrations diffusion samplers may lead to overestimation of up to 70% of the concentration. They suggest potential issues related to the deformation of the Teflon membrane which would modify the distance between coated filters and the Teflon membrane that could cause sampler saturation.

The speculations from A. Neftel and C. Häni are interesting. One could indeed speculate that sweeps, which dominate turbulent transport near the top of the canopy and are characterised by lower wind speed with positive vertical velocity (Poggi and Katul, 2007), could lead to an artificial build-up of the concentration underneath the protecting caps. We could speculate that ejections would not be efficient in "purging" the volume underneath the cap and hence letting over time the concentration being higher in this area. One could also speculate on adsorption-desorption of ammonia on the walls of the ALPHA badges that would be non-linear in response to NH_3 concentrations and lead to possible over estimations under highly fluctuating concentrations as shown in Figure 1 above. This issue necessitates experimental validation of the methodology anyway. We hence propose to add the following text in the discussions section 3.6: *Misselbrook et al. (2005) compared different methodologies and showed that under high concentrations diffusion samplers may lead to overestimation of up to 70% of the concentration. They suggest potential issues related to the deformation of the Teflon membrane which would modify the distance between coated filters and the Teflon membrane that could cause sampler saturation. There is hence some concern on the quality of diffusion samplers to measure concentrations close to large sources which would necessitate field validations.*

We also propose to add a sentence at the end of the conclusion to emphasize this issue for future work: *"Special care should be taken in validating the use of ALPHA samplers near very strong sources"*

Concluding comments:

We judge that the most important potential biases of the proposed multiplot approach are related to biases of the concentration measurements and the used dispersion coefficient. It would be instructive to calculate probability density functions of the estimated emissions with a dataset that reflect the distributions of the measurements and the turbulence parameters that drive the dispersion model.

This would indeed be very instructive but we feel that this issue is rather a work to be for a next study. Indeed, in this study we have explored the first order variability which is driven by the change in meteorological conditions observed in the 13 periods over a year in typical western European climate. The next step could be to extend this assessment to other datasets part of Fluxnet network to incorporate more continental climate conditions. The script we developed for this study actually incorporates measurement noise but we disabled this feature for calculation time reasons.

The authors have tested their setup in field trial in April 2011 applying slurry with a DM content of 6% and an application rate of 41 kg N-NH3/ha. According to the details given in the text, we assume that broadband application was used and was compared to fast incorporation and no application. The cumulated loss amounted to 8 to 10% of the applied NH3. For broadband application, this is a loss on the low side (see e.g. Häni et al.,2016). We would not be astonished if the real emissions would be double as high.

We would like to thank very much A. Neftel and C. Häni for their question and to grant them for their guess. Indeed, we double checked the calculation script used for the real test case and we found one bug in the calculation of the cumulated NH_3 emissions: the multiplicative constant to account for the number of second per time step was set to 30 min while the time step of that particular test case was 60 min. Since a 60 min time step is unusual we did not spot this error in the first place, but we have since then changed the script to calculate the time step from the meteorological dataset. We used this new script and found this bug. This change results in a doubled emission compared to what was given in the discussion manuscript. We hence find, as guessed by A. Neftel and C. Häni that the cumulated losses represented around 20% of the nitrogen applied. We propose to change Figures 16 and 17 as below:

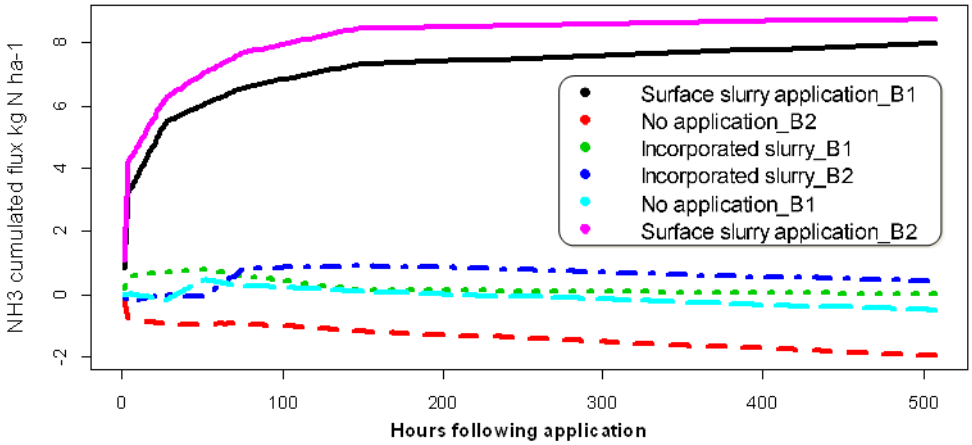

Figure 16. Cumulated fluxes estimated with the inference method on the real test case with strategy C7. Three treatments with two repetitions are compared.

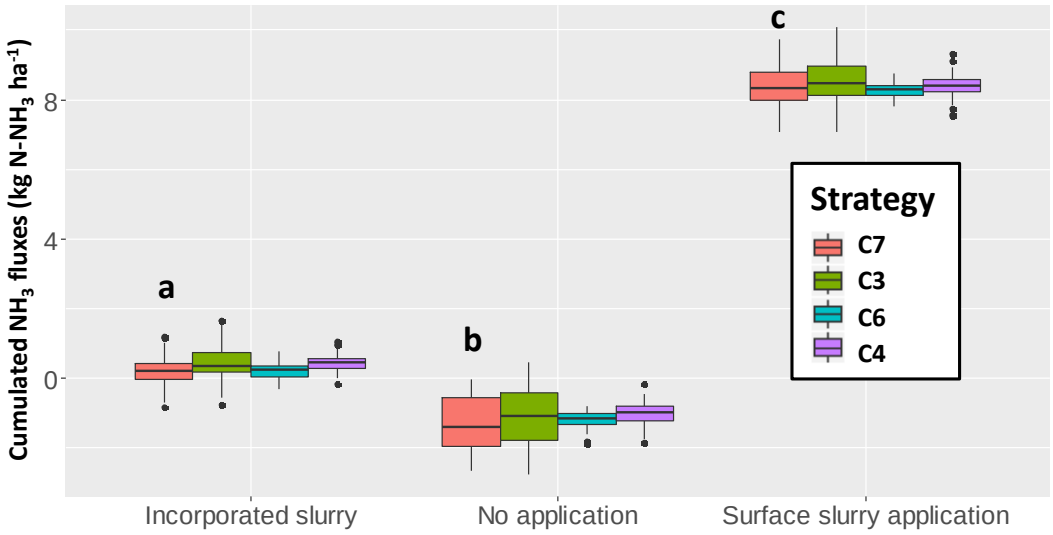

Figure 17. Same as Figure 16 but grouped by treatments and with additional strategies C4 and C6 which consider that replicates have the same surface flux. The variability in the boxplot aggregates the uncertainty on the inference method (the standard deviation on the flux estimate in the least-square model, which accounts for the variability in the replicated concentration measurements), and the variability between the repetitions in each treatment. Letters *a, b* and *c* show significant differences between treatments for the C7 strategy, according to a Tukey test (95% family-wise confidence level).

We also propose to modify the text in section 3.5 (lines 630-636) to: *Surface slurry application showed the largest emissions: 9 ± 0.3 kg N ha-1in B1 and 10 ± 0.2 kg N ha^{-1} in B2 (median and confidence interval). This corresponds to an emission factor around 23% of the N-NH$_4$ applied and 8% of the total N applied, which is in-line with agronomic references (Sintermann et al., 2011a; Sommer et al., 2006). In contrast, the incorporated slurry showed much smaller emissions: 0.3 ± 0.2 kg N ha^{-1} in B1 and 0.6 ± 0.2 kg N ha^{-1} in B2. It is noticeable that the no-application showed slight deposition, especially in B2: -0.26 ± 0.2 kg N ha^{-1} in B1 and -1.7 ± 0.2 kg N ha^{-1} in B2.*

We also propose to change lines 659-663 as follows: *Therefore, we could expect that the real flux is the one measured with C7 times 1.15 (± 0.08), hence would be 10.9 ± 1.3 kg N ha^{-1}. This corresponds to 27 ± 3 % of the N-NH$_4$ applied and ~9 ± 1% of the total N applied. For the incorporated slurry, the emissions are around 20 times smaller than the emissions from the surface applied slurry. Under these conditions, the bias on the emission would be around -20%, which means that the corrected emissions would range from 0.5% to 2.5% of the N-NH$_4$ applied and 0.2 and 0.8% of the total N applied.*

The presented approach to perform NH3 emission measurements in a multiplot arrangement is encouraging and goes in a good direction. To make the approach more robust, the employed ALPHA NH3 sampling systems should be validated under real conditions, i.e. over an emitting source in comparison with e.g. MiniDOAS systems (Sintermann et al., 2016).

We completely agree with A. Neftel and C. Häni about this issue. Although the ALPHA badges have shown to be very precise in comparisons under laboratory and field conditions, it is worth comparing

them with an independent technique in a situation where the source is small and intense and where the sensor is placed near the ground (e.g. 50 cm above ground).

Finally, we would like to invite the authors to collaborate with us to compare the FIDES and WindTrax approach. We have an extensive dataset from field trials where we released CH4 or a mixture of NH3 and CH4 from a circular artificial source with a diameter of 20 meters (Häni et al., 2017).

We are honoured by this invitation to collaborate and would be very happy to compare our approach with WindTrax on CH4 and NH3 emissions.

References:

Bell M.W., A. Hensen, A. Neftel, B. Loubet , P. Robin , Y. Fauvel , Y. Hamon , M. Haaima, A.J.C. Berkhout , D.P.J. Swart, W.C.M. van den Bulk, B.F. van Egmond, D. van Dinther, A. Frumau, B. Esnault, C. Decuq, C.R. Flechard Quantifying ammonia emissions from plot-scale and farm-scale sources using integrated mobile measurements and inverse dispersion modelling, (submitted to AFM).

Carozzi, M., Loubet, B., Acutis, M., Rana, G., Ferrara, R.M., 2013. Inverse dispersion modelling highlights the efficiency of slurry injection to reduce ammonia losses by agriculture in the Po Valley (Italy). Agric. For. Meteorol. 171, 306–318. 10.1016/j.agrformet.2012.12.012.

Flesch, T.K., Wilson, J.D., Harper, L.A., Crenna, B.P., Sharpe, R.R., 2004. Deducing ground-to-air emissions from observed trace gas concentrations: A field trial. J. Appl. Meteorol. 43 (3), 487–502.

Flesch, T.K., Wilson, J.D., Yee, E., 1995. Backward-Time Lagrangian Stochastic Dispersion Models and Their Application to Estimate Gaseous Emissions. J. Appl. Meteorol. 34 (6), 1320–1332.

Häni, C., Sintermann, J., Kupper, T., Jocher, M. and Neftel, A., 2016. Ammonia emission after slurry application to grassland in Switzerland. Atmospheric Environment, 125: 92-99.

Häni, C., Voglmeier, K., Jocher, M., Ammann C., 2017. Recovery rates from line-integrated NH3 and CH4 measurements using backward Lagrangian stocahstic dispersion modelling. Geophysical Research Abstracts, Vol. 19, EGU2017-19557, 2017

Kormann, R. and Meixner, F.X., 2001. An analytical footprint model for non-neutral stratification. Boundary Layer Meteorol., 99(2): 207-224.

Misselbrook, T.H., Nicholson, F.A., Chambers, B.J. and Johnson, R.A., 2005. Measuring ammonia emissions from land applied manure: an intercomparison of commonly used samplers and techniques. Environmental Pollution, 135(3): 389-397.

Poggi, D. and Katul, G., 2007. The ejection-sweep cycle over bare and forested gentle hills: a laboratory experiment. Boundary Layer Meteorol., 122(3): 493-515.

Raupach, M.R., 1987. A Lagrangian Analysis of Scalar Transfer in Vegetation Canopies. Q.J.R. Meteorol. Soc., 113(475): 107-120.

Raupach, M.R., 1989a. Applying Lagrangian Fluid-Mechanics to Infer Scalar Source Distributions from Concentration Profiles in Plant Canopies. Agric. For. Meteorol., 47(2-4): 85-108.

Raupach, M.R., 1989b. Stand Overstorey Processes. Philos T Roy Soc B, 324(1223): 175-190.

Raupach, M.R., Legg, B.J., 1984. The uses and limitations of flux-gradient relationships in micrometeorology. Agricultural Water Management 8 (1-3), 119–131. 10.1016/0378-3774(84)90049-0.

Sintermann, J., Dietrich, K., Häni, C., Bell, M., Jocher, M., Neftel, A., 2016. A miniDOAS instrument optimised for ammonia field measurements. Atmos. Meas. Tech. 9 (6), 2721–2734. 10.5194/amt9-2721-2016.

Sintermann, J., Spirig, C., Jordan, A., Kuhn, U., Ammann, C., and Neftel, A.: Eddy covariance flux measurements of ammonia by high temperature chemical ionisation mass spectrometry, Atmos. Meas. Tech., 4, 599–616, doi:10.5194/amt-4-599-2011, 2011.

Thomson, D.J., 1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. Journal of Fluid Mechanics, 180: 529-556.

Wilson, J.D., 2015. Computing the Flux Footprint. Boundary Layer Meteorol., 156(1): 1-14.