Interactive comment on “Modelling land atmosphere daily exchanges of NO, $NH_3$ , and $CO_2$ in a semi-arid grazed ecosystem in Senegal” by Claire Delon et al.
Anonymous Referee #1
Received and published: 13 January 2019

We thank both referees for their careful consideration and comments on the manuscript. We bring answers to every comment hereafter, and indicate corresponding changes that will be made in our manuscript.

The authors investigate daily exchanges of NO, $NH_3$, and $CO_2$ in a semi-arid grazed ecosystem in Senegal. Three different models (STEP-GENDEC-NOflux, Zhang2010 and Surfatm) are used to simulate daily fluxes during the years 2012 and 2013. Model results are evaluated against experimental results acquired during three field campaigns. Despite the vast extent and importance for global C and N cycling, studies from semi-arid regions are underrepresented in the literature mainly due to challenging conditions for acquiring robust field data. Hence, this study tackles an important topic by testing the suitability of different models to study the land surface-atmosphere exchange of NO, $NH_3$, and $CO_2$.
The manuscript is within the scope of BG, it is mostly well-written, has a relatively clear structure, and it presents new and important data. In principle, it has the potential to be a good contribution; however, the authors have shown little care in the description of the methods, and I unfortunately fail to recognize any interest in ensuring reproducibility of the results. Attention to detail and scientific rigor is rather underwhelming and not up to the standards of the journal. Additionally, some conclusions about $NH_3$ exchange are drawn on a temporal scale that is not warranted by concentrations measured with passive samplers on a monthly basis. I recommend addressing these issues and expanding the discussion in a major revision.

General remarks and major issues:
1) Typesetting is very sloppy. Subscripts are missing, there are periods in units where spaces should be, variables are not slanted and therefore indistinguishable from descriptive subscripts, captions are missing periods, etc.
The typesetting will be corrected in the whole manuscript.

2) The manuscript unfortunately suffers from a lack of units, written variable descriptions, and necessary information in general, both in the main text and the appendix. E.g. Table A5, while generally important and potentially useful, is entirely useless in its current state. It is downright impossible to extract any kind of meaningful information from it unless the reader already knows the model anyway. Time scales are often missing from figures.
A careful read of the entire manuscript will allow the correction of these errors. Equations and variables used will be gathered in a single table (Table A4) to make the reading easier, this will clarify between input data (table A1), initialization parameters (table A2), numerical values of parameters used in the equations (Table A3) and equations (table A4) with explanation of variables, constants and parameters used in them.
Units and variable descriptions will be brought everywhere it is necessary throughout the manuscript. Time scales will be added in the figures.

3) The authors need to carefully address the consequences of using monthly concentration data as input data for a model that is being executed at a 3 hour time step. There needs to be an effort at convincing the reader that the conclusions they draw about the exchange of $NH_3$ are valid, given that the flux is directly driven by Ca-Xcp, and Ca is only available at a 1 month resolution, whereas Xcp is calculated every 3 hours and bound to be variable throughout the day and over the course of a month, due to its exponential dependence on temperature. Part of their first modelling goal is to investigate daily $NH_3$ fluxes, so this is crucial to their objectives.

To address this remark we investigated the relevance of passive samplers for concentration measurements to be used in modeling at a shorter time scale.

In the discussion section 4.1, a paragraph will be added:

"4.1.1 Relevance of monthly $NH_3$ concentration input vs daily $NH_3$ flux outputs

In the two models, $C_{NH3}$ used as input data arises from passive sampler measurements, integrated at the monthly scale (see section 2.2.2). Outputs fluxes are provided at a 3h timescale, averaged at the daily scale for the purpose of this study. The relevance of using monthly $NH_3$ concentrations instead of concentrations with finer resolution in time has been already approached in the literature. Riddick et al. (2014, 2016) have used ALPHA samplers to measure $NH_3$ concentrations at the scale of the week and/or the month. They have noticed that time averaged $NH_3$ fluxes from these samplers provided similar estimated fluxes to those calculated from on line sampling. In the case of passive sampling concentration measurements, meteorological and area sources of uncertainty can still be accounted for in the flux calculation. Riddick et al. (2014) conclude that active and passive sampling strategies give similar results, which support the use of low cost passive sampling measurements at remote locations where it is often logistically hard to deploy expensive active sampling methods for flux measurements. These statements have been confirmed in Loubet et al. (2018), and provide a valuable reason to use monthly concentrations as inputs in the present study."

The following references will be added:

Loubet B. , M. Carozzi, P. Voylokov , J.-P. Cohan, R. Trochard, and S. Génermont, Evaluation of a new inference method for estimating ammonia volatilisation from multiple agronomic plots, Biogeosciences, 15, 3439–3460, 2018.

Riddick, S. N., Blackall, T. D., Dragosits, U., Daunt, F., Braban, C. F., Tang, Y. S., MacFarlane, W., Taylor, S., Wanless, S., and Sutton, M. A.: Measurement of ammonia emissions from tropical seabird colonies, Atmos. Environ., 89, 35–42, 2014.

Riddick, S. N., Blackall, T. D., Dragosits, U., Daunt, F., Newell, M., Braban, C. F., Tang, Y. S., Schmale, J., Hill, P. W., Wanless, S., Trathan, P., and Sutton, M. A.: Measurement of ammonia emissions from temperate and sub-polar seabird colonies, Atmos. Environ., 134, 40–50, 2016.

4) Another one of the three modelling goals is to compare the two NH3 models; however, only for one of them the component fluxes (Fveg and Fsoil) are analysed separately. Even then, there is no further differentiation into stomatal and cuticular fluxes. Why? This is where you learn the most about when exactly the models behave differently.

Indeed, the difference between Fcut and Fstom is available from Surfatm and Zhang2010. Fig. 10 will be modified to show Fstom and Fcut behavior, for both models. Outputs will be compared, added in Table 2 (now table 4 because of inclusion of 2 new tables asked by reviewer 2) and in a new figure 10.

As asked by reviewer 2, Fig. 10 will be mentioned in the result section in paragraph 3.4, with the comparison of Fcut and Fstom from Zhang2010 and Surfatm.

In the discussion section, the figures will be interpreted.

New figure 10:

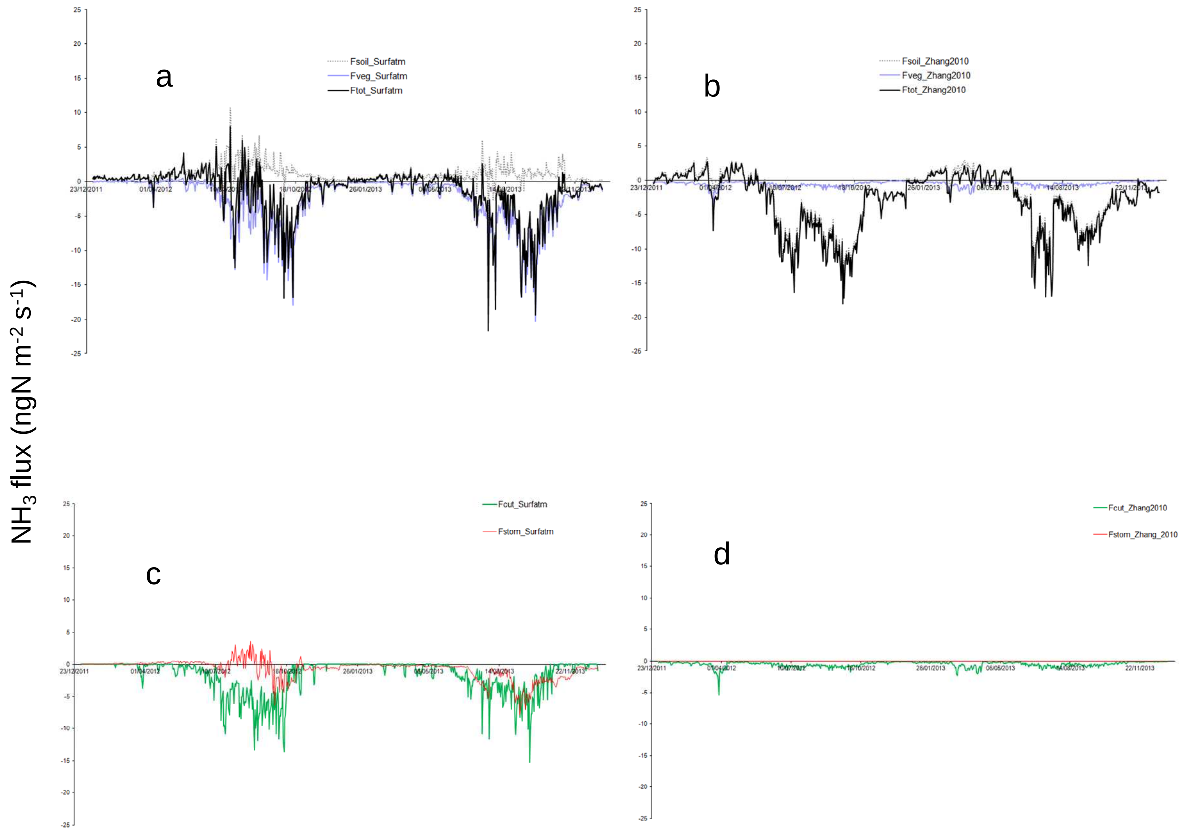


Figure 10: Figure 10: Daily $NH_3$ flux (in ngN $m^{-2}$ $s^{-1}$) partitioned between soil and vegetation. Black line is for total net flux (Ftot), grey dashed line is for soil flux (Fsol) and blue line is for vegetation flux (Fveg) for Surfatm in (a), and for Zhang2010 in (b). Red line is for stomatal flux (Fstom) and green line is for cuticular flux (Fcut) for Surfatm in (c) and for Zhang2010 in (d).

New table 4 (ancient table 2)

| Average flux and standard deviation | Ftotal (net flux) (ngN.$m^{-2}$.$s^{-1}$) | Fsoil (ngN.$m^{-2}$.$s^{-1}$) | Fvegetation (=Fstom + Fcut) (ngN.$m^{-2}$.$s^{-1}$) | Fstom (ngN.$m^{-2}$.$s^{-1}$) | Fcut (ngN.$m^{-2}$.$s^{-1}$) |
|---|---|---|---|---|---|
| Dry seasons Surfatm | -0.2±1.6 | 0.7±0.6 | -0.9±1.7 | -0.4±0.8 | -0.5±1.2 |
| Wet seasons Surfatm | -4.3±4.8 | 2.0±1.9 | -6.3±3.7 | -1.5±2.2 | -4.8±2.7 |
| 2012-2013 Surfatm | -1.4±3.5 | 1.1±1.3 | -2.5±3.5 | -0.7±1.5 | -1.8±2.7 |
| Dry seasons Zhang | -0.9±2.3 | -0.5±2.3 | -0.4±0.5 | -0.02±0.01 | -0.4±0.5 |
| Wet seasons Zhang | -8.1±3.2 | -7.3±3.0 | -0.8±0.3 | -0.03±0.01 | -0.7±0.3 |
| 2012-2013 Zhang | -3.1±4.2 | -2.6±4.0 | -0.5±0.4 | -0.02±0.01 | -0.5±0.4 |

Specific comments:

- P3L17: 2nd objective is unclear, more detail needed.
"Formalisms" will be replaced by "outputs". Actually the behavior of the two models is investigated.

- P3L31-32: Please check sentence. How can rainfall be on average 356 mm for 2013? Or does "average" relate to the period 1951-2013?
The sentence is replaced by: "At Dahra, the annual rainfall was 515mm in 2012 and 356mm in 2013 with an average of 416mm for the period 1951-2013"

- P4L11p., P8L14p., P8L28, and other parts of the manuscript: What is the reason for the use of 3 h averages instead of a higher resolution, if all the forcing variables except NH3 concentrations are available every 15 min?
The meteorological data were available every 15 minutes but the forcing in Zhang2010 was designed to be every 3h. To keep consistency between the two models, we chose to force Surfatm every 3h as well.

- P4L15: Which Gill sonic model exactly?
Gill R3 Ultrasonic anemometer. The specification will be added in the text.

- P4-5: (How) were the different measurement heights (meteorology, sonic and IRGA, passive sampler concentrations) considered in the modelling studies?
Meteorology sensors were located 2m above ground level (AGL), and concentrations were measured 1.5m AGL. IRGA and sonic were located 9m AGL but not used in modeling forcing. Fluxes measured by IRGA and sonic give surface fluxes and hence independent on measurement height (as long as it is recorded within the inertial sublayer). All data were used as measured without any correction for height differences in the models.

- P4L16: Why was an outdated version of EddyPro used? There have been lots of bug fixes since 2013.
These are published data, and the data treatment was originally done for the study of Tagesson et al. (2015b). The EddyPro version used was the best available at the time of the data analysis.

- P7L6: Add parentheses around Eq. reference.
Corrected.

- P7L7pp. / Eq. (1): Units are missing. Move number to right margin.
Units will be added and number moved to right margin.

- P7L29: Fix typesetting of Eq. Remove "equation" in "(equation 3)". Move number to right margin.
Corrected.

- P8L19: Typo in "Surface".
Corrected.

- P8L27: Typo in "Hansen"
Corrected.

- P8L30: (How) was LAI measured?
LAI was measured according to the methodology developed in Mougin et al., Estimation of LAI, fAPAR and fCover of Sahel rangelands (Gourma, Mali), Agricultural and Forest Meteorology 198–199 (2014) 155–167. Data from Dahra were measured monthly during the wet season and were not published (Mougin, personal communication). Linear interpolation was performed between these

monthly estimations, and values for the dry season were found in Adon et al. (2013), for an equivalent semi arid ecosystem in Mali, derived from MODIS measurements. These explanations will be added in the text P8L30.
An error was found in the text in the sentence P8L30: "Forcing also includes constant values of roughness length Z0, Leaf Area Index (LAI), displacement height D, canopy height Zh, measurement height Zref, stomatal emission potential, and ground emission potential."
This sentence will be modified: "Forcing also includes values of Leaf Area Index (LAI, measured), canopy height Zh (estimated), roughness length Z0 (0.13Zh), displacement height D (0.7Zh), stomatal emission potential (constant), ground emission potential (derived from measurements during field campaigns, constant the rest of the time), and measurement height Zref (2m)".

- P9L3pp.: Are surface values of T and RH used for the calculation of Rc and compensation points, and if yes, were the parameterisations adapted to it in any way?
According to Personne et al., (2009), Rc depends on RH, and the compensation points depend on T. These parameterizations were not adapted specifically for semi arid conditions.

I assume most of them were originally developed using ambient values at a certain reference height?
Always according to Personne et al. (2009), yes, these parameterizations were developed using ambient values at a certain height, RH at Zref=2m. T is the temperature of the leaves, calculated by the energy budget model using air temperature at Zref=2m.

- P9L6pp.: This section needs to be significantly expanded. It is completely unclear what was done. p-values need a null hypothesis.
This section is an indication of the tools used to calculate simple and multiple regressions between matrices. The text will be modified:
"The R software (http://www.R-project.org) was used to provide results of simple and multiple linear regression analysis. The cor.test() function was used to test a single correlation coefficient R, i.e. a test for association between paired samples, using one of Pearson's product moment correlation coefficient. The p-value is used to determine the significance of the correlation. If p-value is less than 0.05, the correlation is considered as non significant. The lm() test was used for stepwise multiple regression analysis. The adjusted R-Squared (i.e. normalized multiple R-squared $R^2$), determines how well the model fits to the data. Again, the p-value is calculated, and has to be less than 0.05 to give confidence in the significance of the determination coefficient $R^2$."

- P9L11pp.: Correlations alone are not really helpful in determining the accuracy of the models, please report offsets and slopes of the regressions as well.
Thanks for this suggestion, offsets and slopes will be added on the figures or in the text when necessary.

- P9L26: Decimal point missing.
Decimal point added.

- P10L1pp.: Sign convention needs to be mentioned somewhere.
Positive fluxes correspond to emission. The sign convention will be specified.

- P10L28: Was the ANN trained on data from similar ecosystems?
The ANN was trained with data from both temperate and semi arid ecosystems (Gourma region, Sahel, Mali). This will be specified line 28.

- P12L6-8: p=0.2 is not "weakly correlated", it is simply not significantly correlated. There is no such thing as "almost significant" in null-hypothesis significance testing.
This will be changed by "not significantly correlated."

- P12L13-15: "Indeed, canopy compensation point and ambient concentration values are quite similar" How do you know if you compare 3 h modelled compensation points with 1 month ambient concentrations? This needs to be discussed.
The sentence will be modified as follows: "Monthly averaged compensation point and ambient concentration values are quite similar during the dry seasons. Compensation point concentration averaged during the 2012 and 2013 dry seasons is 3.8±1.5ppb, and averaged ambient concentration is 4.3±1.5ppb for the same period. If the 2012 and 2013 dry seasons are considered separately, the values of the means remain the same."

- P12L25-26: Again, how do you know that the concentration decreases within a single month if you only have one data point?!
Actually, the sentence is not correctly written. We meant that the concentration decreases in August compared to the month of July. This specification will be added in the text.

- P13L6p. and P14L29pp.: Soil temperature at which depth?
Soil surface temperature was specified in the text.

- P13L9pp.: How exactly was the model selection done? Have you thought about using an information criterion, such as AIC? Also note that most of these variables are inherently correlated through overlapping diurnal cycles.
The model selection was done by adding each variable step by step, i.e. the best combination was chosen with the best associated significant $R^2$ (p-value < 0.05). This technique was preferred to the AIC to avoid any statistical over interpretation instead of physical interpretation. In other words, the inclusion of indispensable variables (usually used in parameterizations) in the combination could have been dismissed by the AIC, whereas its physical impact on fluxes is incontestable.
The variables we consider are: $NH_3$ concentrations, air humidity, wind speed, soil surface temperature and moisture. Correlations were calculated between each variable (except $NH_3$ concentration). The results are shown in the table below. The only significant correlation was found between air humidity and soil surface moisture. Furthermore, the stepwise multiple linear regression analysis was performed with daily means for all variables, and the diurnal cycle is therefore hidden.

| | Air humidity | Wind speed | Surface temp | Surface moisture |
|---|---|---|---|---|
| Air Humidity | | 0.0045 | 0.01 | 0.5 |
| Wind speed | | | 0.2 | 0.06 |
| Surface temp | | | | 0.02 |
| surface moisture | | | | |

Only the following sentence will be added P13L9: "The model selection was done by adding each variable step by step, i.e. the best combination was chosen with the best associated significant $R^2$ (p-value < 0.05)."

- P13L15p.: I can't really follow this, please elaborate.
The sentence beginning line 15 will be modified as follows:
"As for Zhang2010 fluxes, a stepwise multiple linear regression analysis is run between Surfatm $NH_3$ fluxes and $NH_3$ concentrations, air humidity, wind speed, soil surface temperature and latent heat fluxes since. $R^2$ is 0.6 (with p<0.001)."

- P13L21: This should also be possible with Zhang2010 since both Zhang2010 andSurfatm follow a similar structure after Nemitz (2001). I don't understand why this was only done for Surfatm.
As mentioned in point 4 of the general remarks, the comparison of Fstom and Fcut in Zhang2010 and Surfatm will be added in the manuscript, along with a new figure 10 in the result section.

The following will be added in paragraph 3.4:
"In Fig. 10a, the total net flux above the canopy in Surfatm results from an emission flux from the soil and a deposition flux onto the vegetation via stomata and cuticles, especially during the wet season. On the contrary, the total flux in Zhang2010 in Fig. 10b results from a strong deposition flux on the soil and a very low deposition flux onto the vegetation. This is explained by a strong contribution of deposition on cuticles in Surfatm (Fig. 10c) whereas it is close to zero in Zhang2010 (Fig. 10d).
In Surfatm, emission from stomata also occurs but it is largely offset by the deposition on leaf surfaces which leads to a deposition flux onto vegetation (Sutton et al., 1995). In Surfatm, the deposition on cuticles is effective until the end of the wet season, whereas deposition through stomata lasts until the vegetation is completely dry, i.e. approximately 2 months after the end of the wet season. On the basis of the different averages for each contributing flux in table 4, we estimate that the soil is a net source of $NH_3$ during the wet season, while the vegetation is a net sink in Surfatm, and the soil is a net sink in Zhang2010."

Paragraph 4.1.4 will be modified as follows:
"4.1.4 Contribution of soil and vegetation to the net $NH_3$ flux:
In Surfatm, during the wet season, deposition on the vegetation through stomata and cuticles dominates the exchange. Indeed, during rain events, the cuticular resistance becomes small and cuticular deposition dominates despite an increase of soil emission. This increase is due to an increase of the deposition velocity of $NH_3$, consecutive to the humidity response of the surface, and a decrease of the canopy compensation point, sensitive to the surface wetness (Wichink-Kruit et al., 2007). In Zhang2010, despite the difference in magnitude, cuticular deposition increases as well during the wet season, but is dominated by deposition on the soil.
During the dry season, aboveground herbaceous dry biomass stands for a few months after the end of the wet season when the soil becomes bare, and the vegetation effect negligible in both models. At the end of wet season 2013, the soil contribution to the total flux increases significantly in Surfatm due to the increase of the ground emission potential prescribed at 2000 (instead of 400 for the rest of the year, to be consistent with measurements noted in Delon et al., (2017))."

The following sentence will be added in paragraph 4.1.5:
"Again, the flexibility of this parameter is more adapted than fixed values for 1D modeling, and this may lead to completely different repartitions of the fluxes between the soil and the vegetation, as shown in Fig. 10. This difference in flux repartition highlights the importance of the choice in the type of soil and/or vegetation for the simulations. However, the close correlation between both models ($R2=0.5$, $p<0.01$, slope=0.6, offset=0.4) indicates a similar representation of the net flux in each model and emphasizes clear changes at the transition between seasons."

- P14L14p.: "Indeed, Zhang2010 model was specifically designed to address [. . .] average temporal scales [. . .]" See above, I don't think you can predict more than average temporal scales from your input data.
As discussed in point 3 of the general remarks, the objective of our study is to estimate $NH_3$ exchange fluxes on an annual timescale rather than exploring processes in detail. In that purpose, the use of passive methods for measuring $NH_3$ concentration is particularly suited, especially for field campaigns in remote places.
We agree that for the reasons evoked here, even Surfatm results comport an uncertainty associated with monthly averaged concentrations forcing. This sentence will be canceled, and the rest of the paragraph will be more moderated about temporal rapid changes. We propose to modify the paragraph as follows:
"The lack of variability of the ground emission potential in Zhang2010 highlights the sensitivity of fluxes to this specific parameter for 1D modelling in semi arid soils. The abrupt transitions between

seasons needs a certain flexibility of the ground emission potential to represent the changes in flux direction."

- Appendix A: Typesetting of Tables is wildly inconsistent. A1-A4 look completely different from A5. A5 is absolutely impossible to follow, because not a single variable is explained.
Table A5 will be merged with Table A4 to explain each variable in due place and avoid referring to another table for understanding the equations.

- Appendix C / P28L11: Typo "Penman"
Corrected.

- Table 1: Numbers come out of the blue. Please add sources. Add period at the end of caption.
Numbers will be explained. A column will be added in Table 1. Some parameters are measured and refer to Delon et al., 2017. They are noted "measured" in the Table. Some parameters are estimated from the ranges given in Hansen et al., 2017. They are noted "Estimated" in Table 1. Some parameters are given out of these ranges, and are clearly adapted to semi arid ecosystems. They are noted "estimated specifically for semi arid ecosystems". A sentence will be added P9L1 to specify that these parameters were adjusted.

- Table 2: See above, why only Ftotal for Zhang2010?
Fveg, Fstom and Fcut from Zhang2010 outputs were not generated at first because not directly available from the outputs. They will be generated, and comparisons with Surfatm outputs will be made and available in Table 4 and in the results and discussion parts.

- Figure 1: Questionable use of Comic Sans in a professional setting.
The font will be changed for a more adapted one (Calibri).

- Figure 3: Remove white space (put the subplots next to each other). What is the temporal scale (I assume 3 hour averages)?. 1:1 line and regression are hard to distinguish, I advise plotting one as a dashed line. The systematic mismatch for LE in the 20- 60 W m-2 region is a little suspicious, do you have an idea what is happening there?
As all the manuscript discussion is based on daily averages for all variables, the figures contain daily averages. Figure 3 will be corrected with dashed line for 1:1 line, white spaces will be removed. As mentioned in the caption, "Available measured EC data are more numerous for H than for LE due to the criteria applied by the postprocessing (see supplementary material of Tagesson et al. (2015b))". These criteria lead to numerous missing data for intermediate periods between dry and wet seasons for medium flux values.
We found errors in equations of the linear regressions written on the figures. The equations will be replaced by correct ones.

- Figure 4: See above re: whitespace. 4a is also a good example why I asked for slopes and offsets in section 3.
Figure 4 will be corrected. Daily will be added in the caption. 1:1 line will be dashed.

- Figure 6: NH4+ is not ammonia. Same error appears in the text when the figure is referenced.
Ammonia will be replaced by ammonium both in the text and in the caption.

- Figure 9: Caption mentions error bars, I don't see any.
Error bars are actually black for measured values. We will specify "error bars for measurements" in the caption and change the colors as asked by reviewer 2.