Response to Anonymous Referee #1

Overall response: We would like to thank reviewer 1 for these insightful and useful comments. These help improve the manuscript. Here we outline the point-by-point responses below in blue, and the relevant figures are attached.

Comment: The manuscript reports a new process-based model for the emission of ammonia following fertiliser application, and then applies this on a global scale. The topic of agricultural ammonia emissions is an important one, and there is clearly a need for well-calibrated, process-based models as an aid to understanding the basic processes involved in ammonia volatilisation and as a means of upscaling to regional/global scales.

Reply: We appreciate that the reviewer recognizes the value of our study. We thank the reviewer for spending time reviewing the manuscript and the development of the AMCLIM model.

Comment: My understanding is that the AMCLIM model is a lightweight model that focuses on ammonia emissions in a period of 1-2 weeks following fertilisation, when the vast majority of ammonia is emitted. It uses a relatively simple Ohm's law like structure, which leads to most of the equations being linear. Non-linear contributions to N cycling, for example due to microbial activity, coupling between C and N and mineralisation of organic N compounds are ignored as being unimportant for predicting ammonia emissions shortly after fertilisation. Soil moisture and soil temperature are not explicitly modelled, but rely on measurement data. The relative simplicity makes the model a potentially valuable tool for experimental groups carrying out ammonia emission measurements. As compared to more detailed process based models (e.g. Daycent or DNDC type models) I assume the model is considerably easier to set-up, faster to run and doesn't require long spin up periods. Furthermore, the comparison to the GRAMINAE site shows that the model does a good job of capturing the variation in ammonia emissions in the days following fertilisation.

On the other hand, I think the application of the AMCLIM model to global ammonia emissions from croplands is premature. Insufficient evidence is provided to show that the model is well calibrated. The comparison to a single grassland site in Germany suggests that the model shows promise in capturing the diurnal cycle of ammonia emissions following fertilisation. However, before applying the model to global croplands I would like to see (see also calibration comments below):

Reply: The calibration process of a model aims to improve the model results by tuning the model parameters to better represent particular conditions. In this study, we performed a detailed timeseries comparison between measured and modelled NH_3

emissions at a high temporal resolution of 15 mins (Fig. 4), and adequate amount of parameter tuning and testing of model complexity have been done and discussed (as shown in Fig. 5). Meanwhile, we have done a multi-site model-measurement comparison (Fig. 12 and 13). More detailed responses are given at the reviewer's calibration comments section.

Comment: 1. Improved evidence that the pH dependence of ammonia emissions is well represented. Figure 13a,b provides some information in this direction, but it is hard to conclude from this that the model is well calibrated (for example a factor 2 difference in the y axis scale is needed to show the modelled Pv values as compared to the measured values).

Reply: As shown in Fig. 13a, b, volatilization rates were plotted against soil pH. The volatilization rates generally increase towards higher pH, which is reflected by both measurements (Fig. 13a) and the model (Fig. 13b). We kindly remind the reviewer that the comparisons between measured and modelled volatilization rates are shown in Fig. 12 (not Fig. 13). Overall (in the original Fig. 12), 18 out of 26 and 3 out of 4 modelled P_V are within a factor of 2 for simulations of urea application and ammonium application, respectively.

Comment: 2. Evidence that the pH change and the impact this has on ammonia emissions following urea application is well calibrated.

Reply: We agree that pH is a critical factor that influences the NH_3 volatilization, especially for urea application. We are aware of the soil pH increase after urea application, which has been discussed in Section 2.2.1 under *Soil pH scheme in AMCLIM-Land* (line 284 to 299). Our soil pH dynamics follow a simple scheme, which is developed based on Chantigny et al. (2004), and Móring et al. (2016), as now shown in Fig. R1-1.

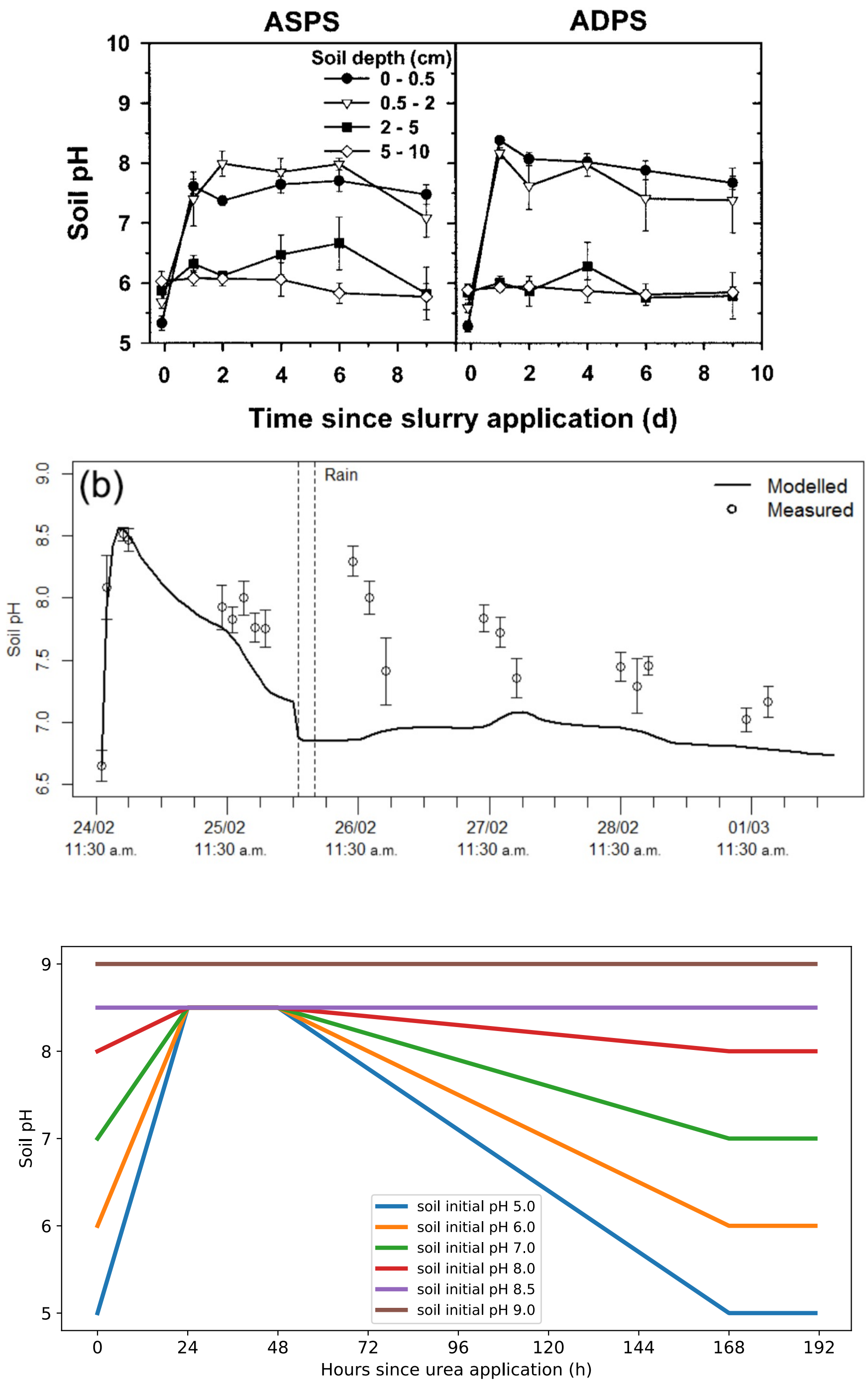
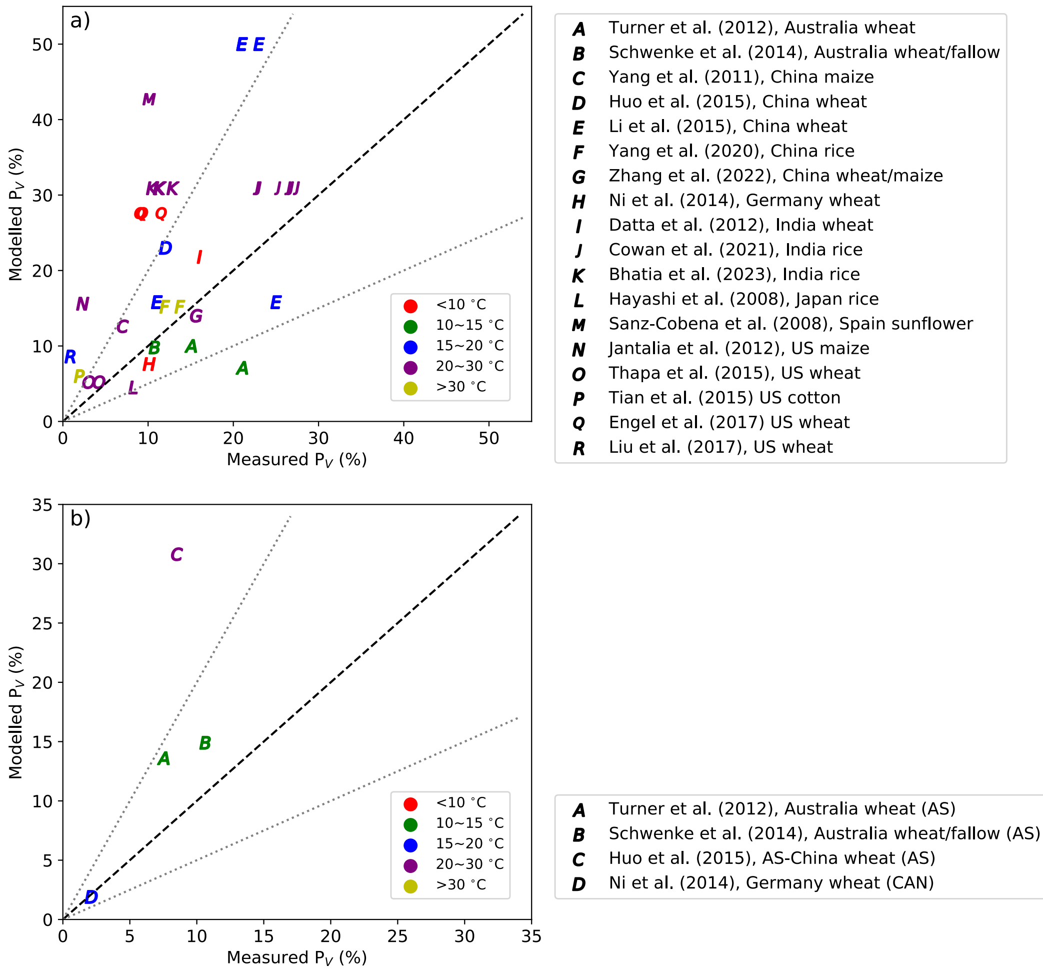


Figure R1-1 Top panel: soil pH change after slurry application (Chantigny et al. (2004)) Mid panel: soil pH change after cattle urine deposition (Móring et al. (2016)) Bottom panel: Soil pH scheme used in AMCLIM–Land. Changes of soil pH for 192 hours (8 d) after urea application for soils with initial pH of six different values.

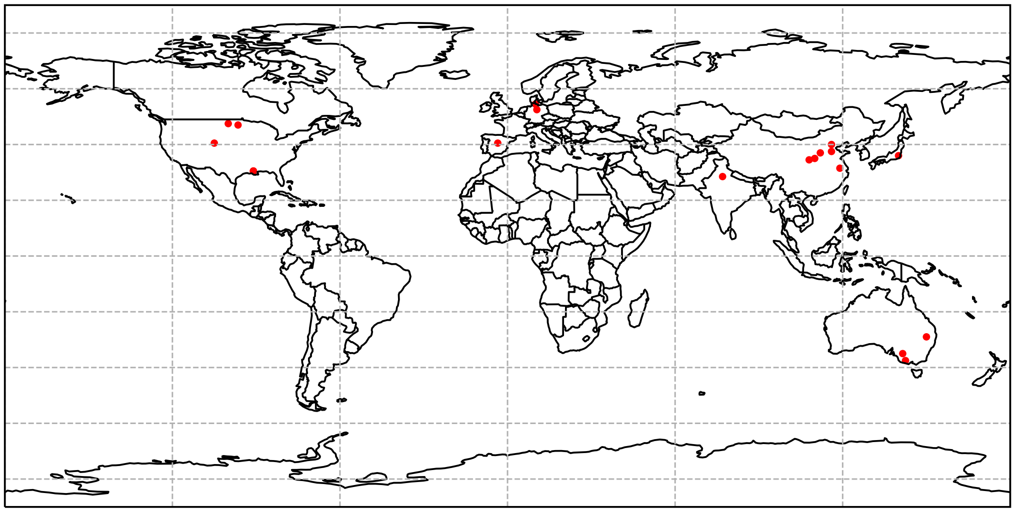
Since there are no available datasets of NH_3 emission measurement following urea application that have sufficient input to drive AMCLIM, we conducted a multi-site comparison between model results and measurements across the globe, as shown in Fig. 12a.

Comment: 3. Evidence that the total ammonia emissions are well represented across multiple crop, soil and climate conditions. For example, that the model can capture ammonia emissions from paddy rice fields in South East Asia. Figure 12 goes some way in this direction, but I find it hard to conclude from this figure that the model is performing well across multiple conditions.

Reply: We would like to once again address that there is lack of suitable measurement datasets that can be used for a detailed model-measurement comparison like what we did with the GRAMINAE data. Therefore, we conducted the multi-site comparison (Fig.12). These studies used for comparison were from 12 sites in seven different countries across the globe. We have now further improved the figure to show their climatic conditions and the geographic locations, with several additional studies from the US added. The updated figure shows that the selected sites for comparison have a reasonable spatial distribution for various climatic conditions (representing by average temperature from 8.5 to 31.7 degC) and soil conditions (representation by soil pH between a range from 5.7 to 8.5), with five major crops being examined (wheat, rice, maize, sunflower and cotton). Overall (in the updated Fig. 12), 20 out of 33 (originally 18 out of 26) and 3 out of 4 modelled P_V are within a factor of 2 for simulations of urea application and ammonium application, respectively. One aspect that needs to be remembered is that measurements also have uncertainties, especially when using enclosure methods to measure NH_3 emissions (Kamp et al., 2024), with the finding that chambers may substantially underestimate emissions in some circumstances (e.g. when dealing with short vegetation and limited mixing in chambers). This points to the need for the future for more exact information being provided by experimentalists than has often been the case in the existing literature.



Updated Figure 12. Modelled percentage volatilization rates (P_V, %) compared with experimental studies (Hayashi et al., 2008; Sanz665 Cobena et al., 2008; Yang et al., 2011; Datta et al., 2012; Jantalia et al., 2012; Turner et al., 2012; Ni et al., 2014; Schwenke et al., 2014; Huo et al., 2015; Li et al., 2015; Thapa et al., 2015; Tian et al., 2015; Engel et al., 2017; Liu et al., 2017; Yang et al., 2020; Cowan et al., 2021; Zhang et al., 2022; Bhatia et al., 2023). Measurement data were from literature that studied NH_3 volatilization from (a) urea application and (b) ammonium fertilizer application to field. Black dashed line is the 1:1 line, and grey dotted lines indicates the values within a factor of 2. Colours represent average temperature during the measurement periods.



Geographical distributions of the measurement sites used for model comparison.

At the same time, the development of the AMCLIM model is based on the understanding at process level, and we tried not to "invent" parameters, of which the only purpose is to improve model performance. We think it is debatable that a model should be intensively calibrated (or this should be the only method), especially when the uncertain parameters do not physically (chemically and biologically) make sense. This hampers the identification of limitations in current models.

Comment: My opinion is that the model description and application to the GRAMINAE site would already make a valuable paper. I would suggest to leave out the application to global croplands and publish this at a later date, once more extensive calibration can be performed using site-scale data.

Reply: We thank the review for recognizing the value of the development of AMCLIM and the site application of the model. We try our best to address reviewer's comments and improve the manuscript, and wish to include the global simulations in this manuscript and its revised version.

It worth mentioning that one of the main goals of this study (and the following papers/manuscripts) is to provide global estimates of agricultural NH_3. Ideally, we would like to have more good quality measurement datasets that can be used to evaluate the AMCLIM model. Unfortunately, the GRAMINAE study appears to be the only one that is suitable. We admit that it is a limitation that AMCLIM was only intensively evaluated against one single measurement dataset, therefore the multi-site comparison is provided as complimentary evidence to demonstrate the capability of the model. We think the global application is useful, and not having a lot of measurements for intensive comparison/evaluation should not be a reason not to apply the model to larger scales. It could take decades until there are multiple comparable datasets of the quality of the GRAMINAE experiment, and it would be unreasonable to prevent publication until such time.

We would like to also use this manuscript to address the importance of good quality measurements, and raise the awareness that it would be more helpful for process interpretation, model development and evaluation if future studies on measuring NH_3 emissions can have well-documented measured variables. Since nowadays "code availability" is a requirement for modelling studies, sharing the full sets of measurements should also be encouraged. In the revised manuscript we propose to outline a list of requirements for reporting of future measurements of agricultural NH_3 emissions:

1) Information of the field site (coordinates, basic climatic and soil conditions)
2) Meteorological variables that are measured at high frequency and reported with high temporal resolution (ideally sub-hourly, e.g., 15 mins), including air temperature and wind speed at a reference height, atmospheric pressure, precipitation and humidity. Radiation and heat flux measurements are also very useful.
3) Soil temperature and soil moisture measured at a specified depth (better to have measurements at multiple depths) with the same measured frequency as the meteorological variables.
4) Soil textures, bulk density and pH. If possible, the soil pH should be measured continuously or periodically following urea application.
5) Above-canopy fluxes of ammonia and atmospheric concentration of NH_3 at a reference height above the canopy (e.g. 1 m), together with reporting of uncertainties (especially where uncertainties vary over time), with clear information on the NH_3, measurement method and the flux method used, including any assumptions.
6) Description of the field site, including estimates of surface displacement height, roughness height and single sided leaf area index (LAI, if vegetation presented).
7) Record of human management practices, such as fertilization information (date, time, amount and technique) and irrigation.
8) (Optional) Hydraulic conductivity and cation exchange capacity.

It is critical to avoid large gaps of measurement data, especially at key periods if measurement data are to be used for detailed comparison with a model application.

Comment: More detailed questions/comments follow below:

** Model details **

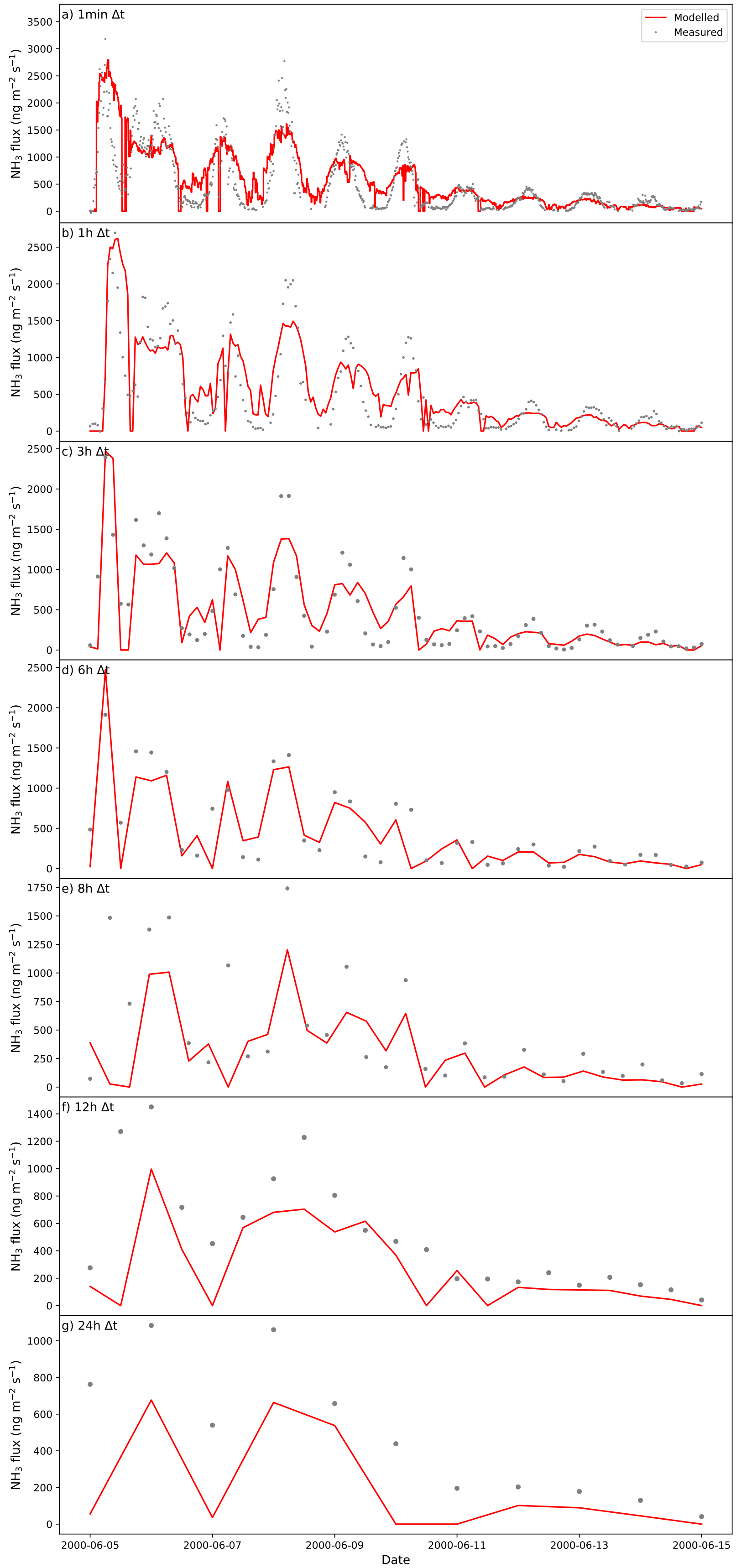
The AMCLIM model includes N uptake by plants, but ignores N uptake by microbes (immobilisation). This may be an important process, especially for fertilisation events at or before planting, when plant N uptake is low. Please discuss why/under what circumstances it is reasonable to ignore microbial N uptake.

Reply: Immobilisation or microbial N uptake is a competing process against plant N uptake and is considered to be primarily regulated by available C in soils and gross ammonification ((Butterbach-Bahl et al., 2011). It remains uncertain that how microbial N uptake affects NH_3 emissions. Perhaps the clearest indication is provided by manipulation of ecosystems, such as by using nitrification inhibitors, the use of which has tended to lead to a small increase in NH_3 emissions, by increasing the lifetime of surface NH_4^+ pools (Snyder et al. 2009).

AMCLIM does not simulate soil C dynamics, and the explicit incorporation of microbial activities is beyond the scope of this study. In contrast, we used a simple N uptake scheme to estimate plant N uptake, which is used as an indicator to evaluate the nitrogen use efficiency. We propose to modify the manuscript to clarify that microbial N uptake is not simulated.

Comment: To what extent has the sensitivity of the model to changes in temporal and spatial resolution been tested? Ideally the spatial and temporal resolution is reduced until the ammonia emissions become relatively independent of further decreases (and I see no reason why the resolution cannot be made finer than the meteorological and soil inputs, e.g. by splitting each soil layer into sub-layers). Figure 5 suggests that changing the spatial resolution of the top soil layer leads to large changes in the model behaviour (comparing circles for z1=1,2,3 cm), and thus that the model behaviour has not yet converged.

Reply: We have tested the model performance at various temporal resolutions, including time-steps at 15 mins, 1 hour, 3 hours, 6 hours, 12 hours and 24 hours (proposed new Fig. R1-2). By decreasing the temporal resolution from 15 mins to 6 hours, the model was still able to capture the main temporal variations in fluxes and to reproduce the peak emissions of each day, while giving a reasonable estimate of cumulative NH_3 emissions. However, when the temporal resolution decreased to 8 hours and even less, the model started to underestimate NH_3 emissions and was not capable of reproducing the emission peaks. This is also the reason why global simulations were performed at an hourly time step.



a) 1min Δt
Modelled
Measured
b) 1h Δt
c) 3h Δt
d) 6h Δt
e) 8h Δt
f) 12h Δt
g) 24h Δt
NH$_3$ flux (ng m^{-2} s^{-1})
2000-06-05
2000-06-07
2000-06-09
2000-06-11
2000-06-13
2000-06-15
Date

Figure R1-2 Comparisons between measured and modelled NH_3 emissions with modelling time-steps varied from 1min to 24h. Simulated cumulative NH_3 fluxes were 0.51 g m^{-2} for 1min time-step, 0.36 g m^{-2} for 3h time-step, 0.34 g m^{-2} for 6h time-step, 0.25 g m^{-2} for 8h time-step, 0.24 g m^{-2} for 12h time-step and 0.18 g m^{-2} for 24h time-step.

The thickness of the top soil layer refers to the assumption of where the fertilizer is applied (by broadcasting, or surface spreading) for technical feasibility, and we found out that a 2cm layer is a reasonable assumption, which also gave a good model result when comparing model and measurements. Figure 5 shows that both increasing the thickness (to 3cm) and decreasing the thickness (to 1cm) of the top soil layer resulted in poorer model performances compared with a 2cm top soil layer setting. We tried not to make the soil layering over complicated for AMCLIM. Pursuing a model convergence behaviour is not practical given the fact that the overall performance of current model settings is reasonable.

Comment: I would expect the time-step to be important as the underlying processes have very different response times. In particular the chemical equilibrium reaction between NH3 and NH4+ is much faster than plant N uptake or nitrification. As such I think it is important to have some time-step control, especially in the minutes following broadcast fertilisation (is 15mins / 1hour short enough?). Similarly, the spatial scale will control the interaction between the concentration of NH3 and NH4+ in the top soil layer following broadcast fertilisation and the transport processes (is 4 soil layers enough?).

Reply: The GRAMINAE campaign has demonstrated that a 15 min time interval is short enough as the sub-hourly variations in NH_3 fluxes have been well captured by the measurements. Meanwhile, the meteorological inputs that drive the AMCLIM model have a temporal resolution of 15 mins. By simply increasing the temporal resolution of simulations (reduce the time-step) without higher resolution input, the model results are insignificantly different, but the computational costs will increase enormously (running the model at 1 min time-step but use the same meteorology for a 15 mins window leads to only a 4.1% difference, which does not justify the substantial increase in computational costs).

Regarding the soil profile, the most critical layer is the top soil layer where NH_3 volatilization takes place. As the main goal of this study is to model NH_3 emissions rather than simulating the soil C dynamics or N_2O fluxes, having a more detailed multi-layering of the vertical soil profile is not as important as getting a suitable value for the top model layer thickness.

Comment: As a related point, it would be useful to briefly mention how the coupled differential equations are solved. Is this by a Euler method or is a higher order method used?

Reply: The prognostics in AMCLIM are solved by the Euler method at each time step. We have added this point to the manuscript.

Comment: I think the section 'volatilisation of NH3' could be improved, in particular the description of how the surface NH3 concentration is calculated.

Reply: The calculation of the surface NH_3 concentration was explained in detail in the Supplementary material see Section S6. We have improved the manuscript accordingly.

Comment: It would also be useful to provide the recovery function for soil pH following urea fertilisation.

Reply: We have now added the equation for soil pH following urea application to land and a corresponding figure (Fig. R1-1).

Comment: For the plant N uptake, what values are used for W_r,i (SM14 and SM17)? I couldn't find this in the supplementary information. Also it would be useful to mention how perennial crops such as grass are treated, especially since this is relevant for the GRAMINAE site (the stages in Table S1 seem to be for annual crops).

Reply: The values of *W_r,i* were taken from Thornley et al. (1991), which are equivalent to 20, 40, 60, 80 g m^{-2}. While the results from the site testing are appropriate for cut grasslands and other agricultural crops, we focus the global upscaling in this study on arable crops. (Note that global upscaling of grazed grassland emissions is included in the second part the AMCLIM model for livestock, to be presented in a forthcoming manuscript). Considering cut/fertilized grassland, we note that there is no information from the GGCMI3 dataset on the spatial distribution of fertilizer amounts, which is much less well constrained than for arable crops. In the revised manuscript, we therefore highlight this gap in knowledge as an important research need.

Comment: ** Calibration **

As far as I understand, model parameters are taken from the literature, and are mostly not calibrated by comparing the AMCLIM model to measurements (a small number of model variations are shown in Figure 5, but are only compared to a single 10-day measurement at 1 site). Compare, for example, to Gurung et al. Nutr Cycl Agroecosyst

(2021) 119:259–273, where Bayesian techniques are used to perform a joint calibration of the 18 parameters relevant to ammonia volatilisation, comparing different levels of model complexity and by taking into account 8 different experimental sites with 42 site-year treatments. I understand that the authors cannot do everything in one paper, and am not expecting them to perform a full Bayesian calibration in this manuscript. However, I think that a comparable level of calibration is necessary before applying the model at a regional/global scale.

Reply: We thank the reviewer for pointing out the study by Gurung et al. 2021, which is an interesting study.

In the ACLIM model, there are three/four types of parameters.

1. Parameters that are measured or derived from lab and field experiments, e.g., Henry's law constant and dissociation constant.
2. Parameters from well-established theory and are widely used in sophisticated models, e.g., aerodynamic and boundary layer resistance for constraining the fluxes.
3. Parameters that are empirically derived or taken from other process-based models, e.g., coefficients for nitrification, coefficients for tortuosity correction that affect diffusion, coefficients for plant N uptake.
4. Parameters for AMCLIM setup, e.g., vertical soil layering, background NH_3 concentrations.

Tuning and testing was done for the last two categories. The model performances were then evaluated based on three statistics: the correlation coefficient, the error and deviation, as shown in the Taylor diagram Fig. 5. We have tested the model performance with different levels of complexity (Fig. 5), e.g., excluding drainage, runoff or nitrification. The current version of the model was found to be the optimal.

Gurung et al. (2021) developed an NH_3 emission module and implemented into the DayCent model. The new module is calibrated using a Bayesian method against measurements at several US sites. This is a novel study as it proposed a new NH_3 emission module (for urea application) for the DayCent model, in which the NH_3 volatilization was originally poorly represented. However, we noticed a few limitations:

1) There is no comparison for timeseries between the model results and measurements, given the fact that temporal variation is an important feature of NH_3 fluxes. To what extent the DayCent model is able to reproduce daily and diurnal variabilities in the NH_3 emissions remains unknown. In contrast, AMCLIM aims to replicate the temporal variations in NH_3 fluxes as well as to estimate the cumulative N loss due to NH_3 emissions.

2) DayCent is run at daily timestep. This further raised the question of excluding the temporal variations in NH_3 fluxes. The reviewer previously challenged the time step, questioning whether 15mins/1hour time-step is short enough. However, there is no evidence provided in Gurung et al. (2021), which could "fail" to address the reviewer's concern regarding the temporal solution based on the same reviewing standard.
3) Although Gurung et al. (2021) said input variables were reported in the publication, the reference where the model parameters were taken from are not provided and it is very difficult to trace back (as in Methods and Materials section and in Table 3). Many parameters used are empirically determined to formulate the processes. Since these parameters are not taken from well-established theory or process-based, the range of these parameters that was used as priori for the Bayesian calibration were mostly arbitrarily selected.
4) The comparisons were only for 8 US sites, possibly because the parameters are mainly applicable for US conditions. To what extent this can well represent other places (as the reviewer addressed, "across multiple crops, soil and climate conditions") is uncertain. The comparisons do not differentiate between sites, and it is difficult to assess which simulations perform better or worse.
5) It would be welcome if the underpinning measurement dataset used by Gurung et al. (2021) were to be made publicly available. It would provide a good opportunity for other modellers to benefit from the measurement effort summarized.

We emphasize that the focus of the present study is different from Gurung et al. (2021). We recognize that the two studies reflect two different modelling "philosophy", and it is difficult (and probably not possible) to say which is better. We definitely agree with the reviewer that the authors cannot do everything in one study, so we think it is crucial to fairly acknowledge the advantage and limitations of each study, which also reflects the rigor and inclusiveness of our science community.

Comment: I would find it useful to have a table of all model parameters and their values (e.g. in the supplementary material).

Reply: We agree and will add a table of all model parameters and their values in the supplementary material.

Comment: ** Global simulations **

I mentioned above that I believe the application to ammonia emissions from global croplands is premature, and requires additional calibration of the model. However, if the authors choose to retain the global simulations in the manuscript, it would be useful to address the following points:

As discussed in the manuscript, correct fertiliser timing is important, due to the sensitivity of ammonia emissions to meteorological conditions (especially temperature). As such, the assumption that 50% is applied at planting and 50% midway through the growing season on a global scale seems a very crude approximation. Is no better data available? If not, how much do emissions change when these assumptions are varied?

Reply: As we have explained, we consider it important to retain the global simulations in the manuscript. Concerning the timing of fertilizer application, we agree that different assumptions could have an effect. The simulations rely on a static crop calendar from GGCMI3 dataset that estimates global planting and harvesting seasons for the major crops. There is no such statistical data for specific practices of fertilization time. We have highlighted this as a further research need in the revised manuscript.

In response to this comment, we have performed additional rounds of simulations to test three possible scenarios, 1) 100 % fertilizer N applied at the beginning of planting season, 2) 75 % N applied at planting and 25 % midway, and 3) 40 % N at planting, 30 % at one third and 30 % at two thirds of the growing season. The global NH_3 emissions from synthetic fertilizer use based on the different scenarios were 10.8, 12.9 and 15.8 Tg N yr^{-1}, respectively, as compared with the base assumption of 15.0 Tg N yr^{-1} (when applied at 50%:50%). In general this shows that in AMCLIM adding a larger share of fertilizer later in the growing season is associated with increased emission, which can be linked to warmer temperatures as the growing season progresses. However, it should be noted that further testing of this effect would be warranted given the possible effect of tall crop canopies in reducing emissions, which is not addressed in the present version of AMCLIM reported in this study.

The differences in spatial distribution and seasonal variation between different fertilization scenario are shown by Fig. R1-3 and R1-4.

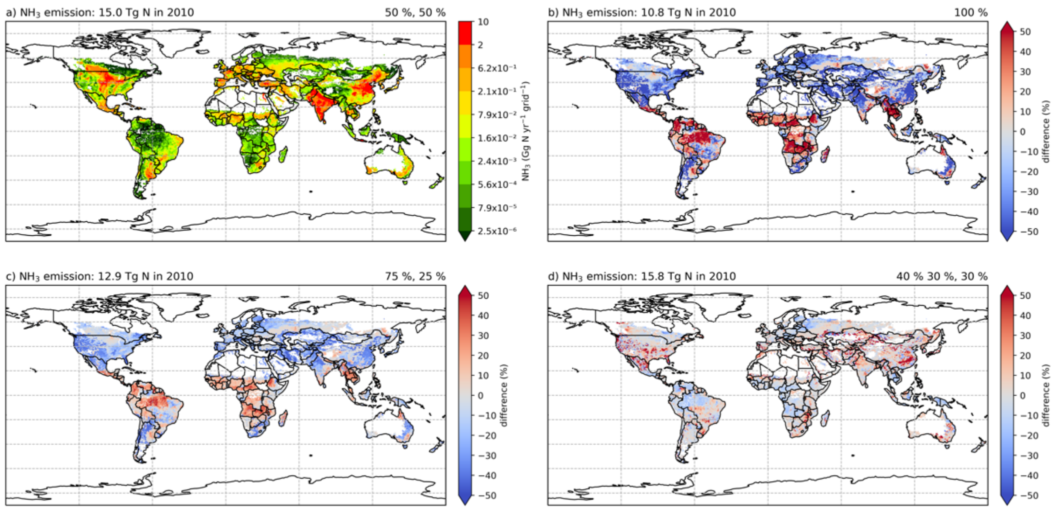


Figure R1-3. Simulated (a) global NH_3 emissions (Gg N yr^{-1} $grid^{-1}$) from synthetic fertilizer use in 2010 using the 50 %, 50 % fertilization scenario, and differences in NH_3 emissions (%) from simulations using (b) 100 % fertilization scenario, (c) 75 %, 25 % fertilization scenario, and (d) 40 %, 30 %, 30 % fertilization scenario.

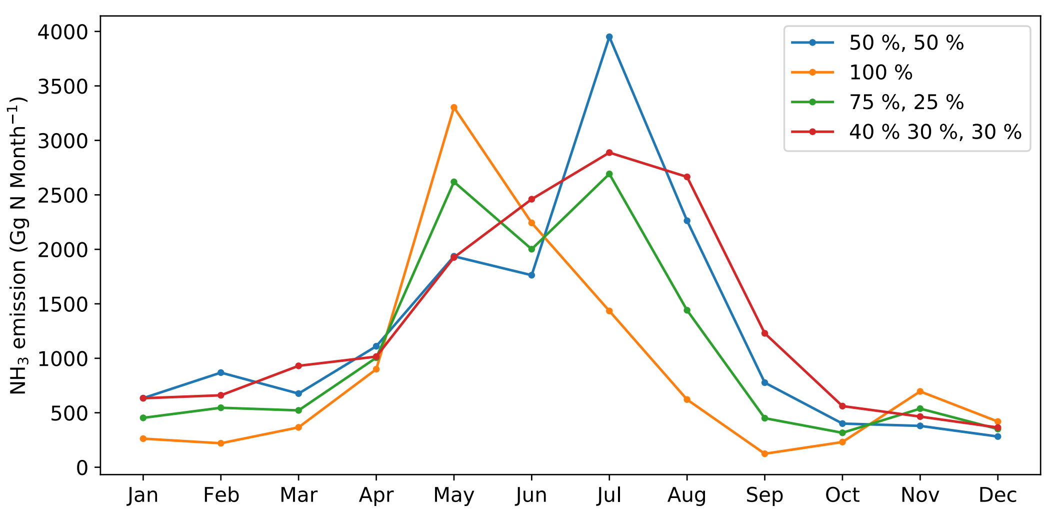


Figure R1-4. Simulated global monthly NH_3 emissions (Gg N $month^{-1}$) from synthetic fertilizer use in 2010 using the four different fertilization scenarios.

Comment: The lack of model adaption for paddy rice systems means it is likely unreliable for these systems. I would suggest to either adapt the model to paddy rice, or to leave rice out of the global simulation.

Reply: On the one hand, there is no data for fertilization of paddy rice systems in GGCMI3 dataset that was used in AMCLIM. On the other hand, we consider simplification is reasonable for global application of the current model version.

Meanwhile, the comparisons in Fig. 12 indicate the simplification provide estimates not out of order of magnitude. We agree future work should include the addressing of individual cropping systems in more detail.

Comment: An uncertainty estimate is given in the discussion section, but no details are provided as to how this was calculated. Please provide details so that the reader can judge how seriously to take this estimate.

Reply: A systematic estimate of uncertainty associated with a process-based model is very complicated. Therefore, we performed a simple analysis for estimating the uncertainty. For ammonium fertilizer, the number was derived from the simulation of GRAMINAE. For urea application, the uncertainty was estimated based on the multi-site comparison (as shown in Fig. 12). It is worth noting that readers should only interpret the estimates and the uncertainty under the context of modelling. We have modified the manuscript to clarify the uncertainty calculations.

Comment: ** Discussion **

I would find it useful to discuss:

How the model differs at a process level from other process-based models such as FAN, DLEM, Daycent or DNDC type models. What are the advantages and disadvantages of the AMCLIM model with respect to these established models and what do the authors see as the future role of the AMCLIM model?

Reply: In response to this comment, we have summarised the features of each model, as listed in Table R1-1. Compared with other models, AMCLIM is a dynamical emission model with an emphasis on the NH_3 volatilization. AMCLIM shows adequate level of complexity in terms of the soil layering construction, and simulations for N processes in soils, the volatilization simulation and soil pH dynamics. AMCLIM has relatively high temporal resolution, which provides implications in the temporal variations of NH_3 fluxes. The highly resolved outputs can be used by atmospheric transport/chemistry models. AMCLIM is considered as a comprehensive emission model rather than a biogeochemical model like the other models shown. However, other advantages reflect in the livestock part of the AMCLIM model, to be published in a forthcoming paper.

Table R1-1 Comparisons of model features between AMCLIM and other models for NH_3 emission simulations. [1*]DayCent does not have an official version that explicitly include NH_3 emissions. The summary is based on Gurung et al. (2021) pointed out by the reviewer. [2*]DLEM is not an open-source model. We are not able to find a model description paper or use manual of the full DLEM model. The summary is based on a model version DLEM-Bi-NH_3 (Xu et al., 2018).

Model	Model type	N processes in soils	Soil pH change/dynamics	NH_3 volatilization process	Vegetation interactions at the surface	Temporal resolution
AMCLIM	Dynamical NH_3 emission model	Four soil layers up to 28 cm depth; A+B+C+D+E+F+G+H+I	Yes; simple generalised scheme but buffering capacity not considered	M1	No	Sub-hourly/hourly
CAMEO	NH_3 Emission module embedded to ORCHIDEE	11 soil layers for hydrology; A+B+E+F+H+I	No	M1	No	Sub-hourly, daily, yearly
DayCent[1*]	Biogeochemical model	14 soil layers up to 210 cm depth; A+C+F+H+I	Yes; empirically derived formula with buffering capacity included	M2	No	Daily
DLEM[2*]	Terrestrial ecosystem model	Unknown soil layering; soil N pools are not explicitly simulated but are derived from fertilizer application rate[2*]	No	M1	Bi-directional exchange scheme	Daily
DNDC	Biogeochemical model	Five soil layers up to 50 cm depth; A+B+C+D+F+H+I	Yes; empirically derived formula with buffering capacity included	M2	No	Daily
FANv2	Process-based N model coupled to CESM	One soil layer 2 cm depth; A+B+C+D+E+F+G+H+K	Yes; pH varies based on age classes	M1	No	Sub-hourly

A–mass balance calculation of N pools
B–NH_3/NH_4^+ equilibrium
C–urea hydrolysis
D–TAN partition
E–surface runoff
F–leaching
G–diffusion in soils
H–nitrification
I–plant N uptake
J–microbial N uptake
K–mechanical N loss
M1–Fluxes are concentration gradient driven and constrained by resistances derived from well-established micrometeorological theory
M2–Empirically derived mass transfer coefficient

Comment: How does the calibration procedure compare to these other models, and what are the consequences for the level of confidence we should have in the AMCLIM model results as compared to established models?

Reply: For modelling at global scale, we do not find explicit model calibrations from FANv2 (Vira et al., 2020), CAMEO (Beaudor et al., 2023), DNDC (Yang et al., 2022) and DLEM (Xu et al., 2019) at site scale simulations. We think this is largely because global models tend to provide general representation and try to avoid over calibration. The management practices at the GRAMINAE site were not complicated, which provides a good test situation for the numerical representations of the physical and chemical processes. In contrast, DayCent is widely used for simulating N_2O emissions (not NH_3 emissions) and is intensively calibrated using N_2O measurements from fields. The parameters can be quite different between simulations for different places, e.g., US vs. Switzerland. There are limited studies of applying DayCent at global scale, and we did not find explicit model calibration in the global application of DayCent (e.g., De Grosso et al., 2009).

The global estimates by AMCLIM are broadly consistent with existing models (as shown in Table 2). Combining with the generally close agreement with the GRAMINAE measurements, the AMCLIM model is considered to be robust and capable in estimating agricultural NH_3 emissions.

Comment: Why has urease inhibition not been considered? This is required by many countries when broadcast spreading urea fertiliser and has important consequences for ammonia emissions. To what extent does this limit the usefulness of the model?

Reply: There is no statistical data of the application of urease inhibitor that can be used in the model. To our knowledge, only Germany has regulations on the use of urease inhibitor (or incorporation is required). The use of urea inhibitor can be incorporated into the AMCLIM in the future work, once there is sufficient data.

Reference

Beaudor, M., Vuichard, N., Lathière, J., Evangeliou, N., Van Damme, M., Clarisse, L., and Hauglustaine, D.: Global agricultural ammonia emissions simulated with the ORCHIDEE land surface model, Geosci. Model Dev., 16, 1053–1081, https://doi.org/10.5194/gmd-16-1053-2023, 2023.

Butterbach-Bahl, K., Gundersen, P., Ambus, P., Augustin, J., Beier, C., Boeckx, P., Dannenmann, M., Gimeno, B. S., Ibrom, A., Kiese, R., Kitzler, B., Rees, R. M., Smith, K. A., Stevens, C., Vesala, T., and Zechmeister-Boltenstern, S.: Nitrogen processes in terrestrial ecosystems, in: The European Nitrogen Assessment, vol. Chapter 6, Cambridge University Press, 99– 125, 2011

Chantigny, M. H., Rochette, P., Angers, D. A., Massé, D., and Côté, D.: Ammonia Volatilization and Selected Soil Characteristics Following Application of Anaerobically Digested Pig Slurry, Soil Sci. Soc. Am. J., 68, 306–312, https://doi.org/10.2136/sssaj2004.3060, 2004

Del Grosso, S. J., Ojima, D. S., Parton, W. J., Stehfest, E., Heistemann, M., DeAngelo, B.and Rose, S.: Global scale DAYCENT model analysis of greenhouse gas emissions and mitigation strategies for cropped soils, 67, https://doi.org/10.1016/J.GLOPLACHA.2008.12.006, 2009.

Gurung, R., Ogle, S. M., Breidt, F. J., Williams, S. A., Zhang, Y., Del Grosso, S. J., Del Grosso, S. J., Parton, W. J.and Paustian, K.: Modeling ammonia volatilization from urea application to agricultural soils in the DayCent model, 119, https://doi.org/10.1007/S10705-021-10122-Z, 2021.

Kamp, J. N., Hafner, S. D., Huijsmans, J., van Boheemen, K., Götze, H., Pacholski, A., & Pedersen, J.: Comparison of two micrometeorological and three enclosure methods for measuring ammonia emission after slurry application in two field experiments, Agric. For. Meteorol., 354, 110077, https://doi.org/10.1016/j.agrformet.2024.110077, 2024.

Móring, A., Vieno, M., Doherty, R. M., Laubach, J., Taghizadeh-Toosi, A., and Sutton, M. A.: A process-based model for ammonia emission from urine patches, GAG (Generation of Ammonia from Grazing): description and sensitivity analysis, Biogeosciences, 13, 1837–1861, https://doi.org/10.5194/bg-13-1837-2016, 2016.

Snyder, C. S., Bruulsema, T. W., Jensen, T. L.and Fixen, P. E.: Review of greenhouse gas emissions from crop production systems and fertilizer management effects, 133, https://doi.org/10.1016/J.AGEE.2009.04.021, 2009.

Thornley, J. H. M.: A Transport-resistance Model of Forest Growth and Partitioning, Annals of Botany, 68, 211–226, https://doi.org/10.1093/oxfordjournals.aob.a088246, 1991.

Vira, J., Hess, P., Melkonian, J., and Wieder, W. R.: An improved mechanistic model for ammonia volatilization in Earth system models: Flow of Agricultural Nitrogen version 2 (FANv2), Geosci. Model Dev., 13, 4459–4490, https://doi.org/10.5194/gmd-13-4459-2020, 2020

Xu, R., Tian, H., Pan, S., Prior, S. A., Feng, Y., Batchelor, W. D., Chen, J., and Yang, J.: Global ammonia emissions from synthetic nitrogen fertilizer applications in agricultural systems: Empirical and process-based estimates and uncertainty, Glob Change Biol, 25, 314–326, https://doi.org/10.1111/gcb.14499, 2019.

Yang, Y., Liu, L., Bai, Z., Xu, W., Zhang, F., Zhang, X., Liu, X., and Xie, Y.: Comprehensive quantification of global cropland ammonia emissions and potential abatement, Science of The Total Environment, 812, 151450, https://doi.org/10.1016/j.scitotenv.2021.151450, 2022