This paper describes the integration of agricultural practices into the TRIPLEX-GHG Model and the evaluation and tuning of the model against measurements.

It is an enormous project to integrate the great heterogeneity of global agriculture into a global model and to evaluate it. It is clearly a task that needs to be published in steps. The current paper indeed makes some headway in this task. Looking at the paper as a whole, two questions stand out. (i) Does this paper really go far enough at this stage in development to warrant publication? To answer this question the paper should improve its explanation of how it is modified from previous versions. (ii) Secondly, the authors should further explore the underlying physical meaning of the model tuning. Details are given below.

1. On my initial reading of the paper it was difficult to discern the new modifications to the TRIPLEX-GHG model from what has already been published. The model description discusses both the older published model and the newer added processes. It would be helpful at the outset to clearly describe exactly what new processes have been added. From my reading they are the following: addition of agricultural fertilizer as either synthetic fertilizer or manure, some global rules for the incorporation of tillage, the addition of flood irrigation, new rules for plant uptake and return of harvested plant residues to soils. As I understand it the nitrification sub-model (and Table S1) is not new but is part of what has already published. If already published as part of a coherent modeling system it is unclear why the paper needs to repeat the model equations here, except to highlight the 13 tuning parameters examined.

Some of the newly added processes should be described in greater detail. Flood irrigation is not well described. Is this done continuously or only when soils dry sufficiently, or ...? Is it done everywhere or are rainfed croplands separated from those irrigated. Harvest is
also not well described. What are the carbon and nitrogen losses during harvesting? It is not just litterfall that is lost during harvesting but presumably a good amount of the harvested plant is lost to the food system.

When building a new model the sensitivity to added processes is usually evaluated. While the paper describes the sensitivity to various parameters it is not clear how the newly added agricultural processes impact the nitrogen cycle. To what extent are nitrification and denitrification processes altered by these processes as claimed in the manuscript? (For example, what is the impact of tilling on the nitrogen cycle?)

The model description does not adequately describe some of the shortcomings of the implementation from the onset, although the conclusion adds more detail. It seems this new version of the model does not really have a crop model, but the crops are somehow woven into the existing pfts. For example, the paper states over Australia that PFTs with tropical forest and shrub instead of agricultural lands were used. Moreover, while the emissions of N2O from specific types of crops are measured, only C3, C4 and rice are used as crop types in the model. How does this comparison work? It seems probable that crops do not have their own soil column although this is not clear. The paper needs to state these considerations at the beginning. Far more detail is needed into how crops are integrated into the pft structure of the model. Is the added nitrogen from fertilizer apportioned to the crops or is it apportioned to the whole grid square? How do the nitrogen demands of the crops interface with the other vegetation within each model grid?

Related Text:

l223-224: How were data transformed to spatial resolution of the model? Does this mean each 0.5x0.5 grid only includes one crop or are other pfts mixed in? How does this fit into the vegetation dataset? Is there a separate agriculture model? How were specific crop measurements evaluated within the model structure?

l163. What is the cropland ecosystem? Have crops been added specially?

2. Equations (2) and (3). Is the plant N demand for all plants changed or just for crops alone? The extent to which crops and plants prefer taking up nitrate is quite a strong assumption, not universally true, and contrary to many models. The references included to justify this assumption are rather old. This warrants more discussion as the results are likely to be very dependent on this assumption.

Despite the rather simplified representation of agricultural practices the tuned model reproduces the measurements with remarkable fidelity. While the authors list many aspects of the agriculture that could be improved, it is difficult to see how they could improve on their present results with the metrics used (R2=87%)! It seems like there
might be two possibilities: (i) the precise representation of agriculture is not important or (ii) there is something about the tuning procedure that allows the model to get the right answer. My guess it is the second possibility. If that is true, the model won’t show much sensitivity to the parameterization of agriculture or the parameterization is irrelevant as the tuning parameterization will easily compensate (see (1) above). This would be easy to check.

One hypothesis is that to a large extent the model solution is dictated by the timing of the fertilizer input. As the plants do not take up ammonia preferentially, most ammonia added is quickly nitrified to nitrate. With this assumption nitrate acts as the crucial pivot point controlling N2O production making it rather easy to tune. Real model skill could best be assessed by looking at other aspects of the simulation: for example, the emissions per fertilizer added (the emission factor), the interannual variability in the emission factor at a particular station or the difference in emission factors between stations. The D value and RMSE with which the model is evaluated would seem to emphasize getting the maximum values correct which seems highly dependent on the amount of fertilizer added. More interesting would be to assess skill in predicting the emission factor or other aspects of the simulation.

3. The authors use their present results to justify the physics in their model. Thus the authors state that (see L 420-441) the fidelity of their model is “…. derived from three features of our model” (line 423). Because the simulation gets the N2O peak following fertilization with tuned parameters does not imply these three features are important. To make this statement the authors need to do considerably more work and to use other metrics to justify their model. I don’t find these conclusions justified based on the current paper.

4. The paper title seems to imply the authors are simulating global cropland N2O emissions. Please give global emission estimates from the model and evaluate the distribution against other published estimates. Alternatively the authors could rephrase to emphasize the fact that only select locations have been evaluated (albeit on most continents) and tuned against.

5. Also, if I understand correctly the tuning parameters used are different on every continent. It is admittedly rather strange for a process-based model to use a different tuning parameter on every continent. This is not based on any environmental variable or other physical aspect. This feels rather unsatisfactory to use in a global model and needs further justification. Otherwise it seems like the introduction of an arbitrary tuning parameter.

6. It would be helpful to understand how the 13 sensitivity parameters control denitrification. Some of these parameters don’t seem to be included in S1. This is where a clear link between the model equations and these parameters would be helpful.
7. Figures 3-7. It is not clear what is being shown here (unless I missed it). Please state clearly in the figure caption and in the text. Are these showing measurements against the fitted model solution at each fitted site? Are these showing measurements and the model solution with the continental average of the fitted parameter for the fitted sites? Are these showing the unfitted sites with the continental fitted parameter. It is really the latter which should be shown – otherwise the authors are just showing the tuned results.

Some minor comments:

l88 "agricultural practice modules“ – it is not clear what is meant here.

Table S1. If retained, please make sure that all the terms in S1 are given. I found a number of symbols and parameters that were not specified.

It is unclear from the model description how soil nitrogen loss to N2 is handled. Could the authors clarify. 150, N2 should be a major gaseous N loss from agriculture.

Are atmospheric resistances used to parameterize the flux to the atmosphere?

Please make sure all abbreviations are spelled out: e.g. NOx and DOC are not given.

l161. It is stated that the supplemental figure referred to proves the effectiveness of the design. First, it is very difficult to evaluate this figure with the information given. What does this show? What are the blue dots and red lines? What are the conditions simulated? Was fertilizer added? It is difficult to see how this figure alone really proves anything.

How is biofixation handled?

What is assumed for fertilization timing, harvest timing, tillage timing, irrigation timing?

It is not clear what governs the crop demand for nitrate versus the demand from denitrifiers? How are these partitioned?

l267: how were sites chosen for model calibration? At random?
Are the fitting parameters significantly different between the continents? Can you give them in the supplement?

It would be helpful if each of the panels in Figures 3-7 show D, RMSE and the correlation. Also it would be helpful if the fertilization time is shown.