Response to reviewer #1

Major concerns:

I appreciate that this is largely a model development and documentation paper, but found some of the discussion surrounding main display items rather hasty and lacking appropriate depth. For example, the spatial patterns of N limitation (Fig. 8) suggest that both tropical and boreal forests are not N limited (indeed, forests globally may have a low sensitivity to N availability)? This is just one example, but on revisions I would encourage the authors to unpack and explore their findings a bit more. Take the opportunity to call out strengths and weaknesses in the current approach and discuss particular model assumptions that are responsible for these features. This doesn't have to be exhaustive, but will help add depth to the results and discussion.

Page 1, Line 6. Significant improvements in crop yields are not apparent in the updated Fig. 10 and SI material uploaded by the authors. Would a plot of global crop yields vs. observations make this point more clearly? If crop model improvements are the big advancement in the current model development, I'd encourage more attention be given to establishing these improvements in the main text. That said, the estimates of N losses (through leaching and N2O emissions) also have important regional and global consequences and seem to be done well in this version of the model. Should these accomplishments be highlighted in the abstract too?

Answer: The results section has been expanded. We have now structured results into 3 sections: Carbon pools and fluxes, Nitrogen pools and fluxes, Land use and nitrogen dynamics. In particular, Taylor diagrams of regional crop productivities have been added to the manuscript and supplement. They clearly indicate that significant parts of the calibration previously needed are now no longer necessary, as N limitation already reduces simulated crop yields in low-input countries. The strong influence of land use on nitrogen losses is now mentioned in the abstract.

Minor comments

P1, L 8. This suggest the is still scaling occurring in regions with favorable climate and N inputs? Is this true?

Answer: Scaling of yields has still to be done for regions where climate and nitrogen supply are not the main limiting factor, but, e.g. phosphorus, poor pest management or other management aspects not explicitly considered in the model as now mentioned in section 4.4.3.

P1, L 18 Zaehle and co-authors (2015) made similar findings, seems worth citing here? Throughout section 2 is it worth briefly distinguishing the similarities and differences in the approach taken with LPJmL vs. other land models, especially LPJguess (Smith et al 2014) which is referenced throughout?

Answer: In fact, nitrogen dynamics have been incorporated in many other dynamical vegetation models (as described on P1, L22). As LPJ-GUESS, our model considers not only natural vegetation but

also takes into account managed crops. In contrast to LPJ-GUESS nitrogen transformation in soils are simulated in a more sophisticated way in LPJmL5, including immobilization of nitrogen. This is now specified in the introduction and model description.

P4, L 6. It seems odd to introduce table 2 before table 1. Similarly, table 4 is introduced before table 3 (page 7).

Answer: The order of tables has been changed following the suggestions of the referee.

Section 2.2. Is the soil biogeochemistry for this version of the model vertically resolved, as implied with e.q. 7? If so should this be mentioned in section 2: model description?

Answer: Soil processes are vertically resolved in 6 soil layers. A sentence has been added to the model description.

Eq. 13-14, could pft specific root distribution parameters be easily described in the current tables or included elsewhere?

Answer: The parameterization of PFT-specific root distribution is now described in the paper and the corresponding parameters included in Table 2.

P5, L 27. As written this sounds like rates of GPP are reduced by respiration rates? This strikes me as strange. Wouldn't autotrophic respiration be subtracted from GPP to calculate NPP rates (eq. 18). This also isn't clear in Fig 1

Answer: Indeed, for calculating NPP, respiration has to be subtracted from GPP but GPP itself is not directly reduced. This has been clarified in section 2.4. In figure 1, GPP (flux) is not shown, but the process (photosynthesis).

Figure 3: check to see the colors for each arrow are labeled correctly and/or defined. Alternatively, the approach here seems pretty standard, I wonder if the distracting rainbow of flux arrows are really necessary?

Answer: Indeed, the approach is based on standard implementations. We find a colored scheme more helpful than a black-and-white figure but acknowledge that others may have other color preferences. The boxes, however, are now colored in black.

Section 2.6 How are all the competing fates of inorganic N solved (e.g. sequentially, simultaneously, etc)?

Answer: The transformation of nitrogen is calculated in sequential order as now stated in section 2.6.

Section 2.6. Are litter and SOM pools have a fixed C:N ratio or are they flexible (w/in bounds).

Answer: Litter and SOM pools have flexible C:N ratios now explicitly stated in section 2.6.

P 15 L 23 What is a “ nodulating leguminous crops pulses”, the phrase seems redundant? Maybe just use “soybeans and pulses”.

Answer: There are non-nodulating legumes (doi:10.14719/pst.2015.2.2.97) as well as non-nodulating soybean varieties (doi::10.1007/978-94-009-4482-4_19) and we wanted to make clear that we're considering these leguminous crops as nodulating varieties with capacities to fix nitrogen.

P17, line 13. What happens to the other 40% of the manure? Is it not really applied, or does it go into SON pools?

Answer: No, the data from Elliott et al. (2015) only provide mineral reactive N forms. In their data set, only the plant available fraction (60% of manure) is specified, the rest is ignored. This is now clarified in the text.

Section 3. What spatial resolution are these simulations? Does each grid cell have a single pft, or is there subgrid variability of vegetation? If there is subgrid PFT variation, do all plants share a soil column, or to they each have individual columns (that is, does manure and fertilizer applied stay on crop only soils, or is it available to plants throughout the grid)?

Answer: Simulations are carried out for a resolution of 0.5 x 0.5°. PFTs can be established concurrently within a cell competing for light, water and nitrogen sharing the same soil resources. This has been explained in more detail in the paper.

Table 4: Although they are described in the main text (section 3) the abbreviations for experiments used in column headings are non-intuitive enough to prevent the information contained in the display item from standing on its own. Consider adding text to the table heading or columns to make these data more understandable.

Answer: The table captions now include an explanation of the different suffixes for the abbreviations of the experiments.

Table 4. I’m surprised global NO3 pools are an order of magnitude larger than the NH4 pools. I wonder if there is spatial structure to these patterns (e.g. accumulation of NO3 is warm or arid regions), or if the patterns is relatively globally distributed. Regardless, it seems surprising given the relatively high mobility and multiple loss pathways of NO3, compared with NH4, and suggests that nitrification rates may be too high in the model? Alternatively, decomposition rates may be high, supplying excess inorganic N, or plant NO3 uptake may be underestimated? This may be worth mentioning in the discussion (section 4.1.1 or 5)?

Answer: This is indeed discussed in section 4.1.1 and we have now expanded this discussion. NH4 is constantly converted to NO3 by nitrification, a microbe-mediated process under aerobic conditions, which are often prevailing. NO3 is lost via leaching or denitrification, again a microbe-mediated process but under anaerobic conditions, which are not dominant conditions in most soils. As a consequence, NO3 concentrations are typically much higher than NH4 concentrations, unless N is limiting plant uptake, see e.g. Kabala et al. (2017). Also, NO3 is accumulated in lower soil layers, despite leaching losses. This phenomenon is reported by Walvoord et al. (2003) and Ascott et al.

(2017) and reproduced by the model. Still, this nitrogen is mostly unavailable to plants, as these have little root access to the lower soil layers at 3m depth. This is now discussed in the paper.

P 18 L 8 Why are agricultural lands (that are presumably being fertilized) becoming increasingly N limited? Is there some metric of N limitation that can illustrate this point more directly, as it's not intuitive from Fig 4a,b. Also, it seems odd to increase carbon use efficiency (NPP:GPP) if the system is becoming more N limited? I see how it occurs in the model, because of higher tissue C:N ration and lower RA costs, but is it ecologically realistic?

Answer: In the simulations, plant growth is enhanced through CO2 fertilization, whereas fertilizer inputs are static. Also, cropland expansion over the 20th century dominantly expands in areas where fertilizer inputs per hectare are low. As such, the global average fertilizer amount (kg N per hectare) declines. This is now also briefly described in the text (section 4.4.3).

The relationship of RA costs on C:N ratios is well established and was already implemented in the carbon-only model version. There plant organs had a static prescribed C:N ratio that determined respiration costs for maintenance.

Fig 4. How were relative GPP changes calculated, I didn't see this described in the text? Also, consider adding information about line colors to the figure caption, as the legend insets are very small and hard to read.

Answer: Relative changes are calculated by dividing the values by their 1901-1910 average as now stated in the figure caption. The meaning of the different colors in the figures is now explained in the figure caption.

Fig. 4 It looks like the two control models (3.5 and 5) lose vegetation and soil C throughout the 20th century, but GCP data suggests the land surface should be a C sink, at least over then end of the 20th century (e.g., LeQuere et al, 2015). Given increases in plant productivity in Fig 4a,b- this suggests the land use C change flux must be pretty large?

Answer: Yes, see also Table 3 where the NBP (net biome productivity) is listed. For natural vegetation (LPJmL3.5_PNV and LPJmL5_PNV), the terrestrial biosphere is a net carbon sink of about 1.5 and 1.7 PgC/yr, whereas when accounting for current land-use patterns (LPJmL3.5 and LPJmL5), the terrestrial biosphere is a net source.

P 18 L 15. Where is the data showing that N limitation increases by a factor of two? How is N limitation being assessed in this statement?

Answer: In fact, the statement is misleading and has been omitted from the manuscript.

P 18 L 16. What's causing the higher leaching losses with the control model? Does it have to do with vegetation demand for N, rooting profiles of managed vegetation, or other factors?

Answer: The higher leaching under land use is caused by the additional fertilizer input and the irrigation water application under land use, which do not occur in natural vegetation. This is now explained in the text (section 4.3).

Table 4. It's it worth discussing limitations (or uncertainties) of some of the 'observational' estimates presented here

Answer: This discussion is placed in section 4.1.1 and we have extended this now. Certainly, estimates based on only one other model have no value for a model evaluation and only serve as a point of reference.

Fig 6. Agreement on crop C:N ratios doesn't seem that surprising, given the ranges for leaf C:N and allocation that are proscribed in the model (Tables 1,2) . I'm assuming the values for R3 in Table 2 were tuned to provide the spread shown in Fig 6? This is fine, but should be acknowledged. In the text.

Answer: In fact, R3 values were calibrated to the observed values. The approximated relationships between leaf C:N ratios and storage organ C:N ratios based on Bodirsky (2012} lead to consistent but variable C:N ratios in harvested crop organs, as stated in the manuscript.

Fig 7a, it strikes me as odd to have low values (<1) indicative of high leaching losses, especially when points above the 1:1 line show areas of high leaching under present vegetation (fig 7b)

Answer: Values <1 in Fig7a indicate regions where leaching of the potential natural vegetation is less than with land use. This corresponds to points Fig 7b above the 1:1 line. It is not useful to define the ratio the other way round, as leaching under natural vegetation can be very small but very large under land use (irrigation, fertilizer input). We have swapped the x and y axis in Fig7b. Now points are below the 1:1 in panel b for values<1 in panel a. This has been also applied for Fig8.

Fig 9, where do the obs come from- especially for NPP and Veg C. Is each point supposed to represent an individual sites? This are from the same FluxNet sites as in the SI figures? Also, where do the observational error bars come from & how were they calculated? Finally, should correlation coefficients & significance be reported?

Answer: The missing reference to the observational data has been added to the manuscript. The horizontal bars denote the minima and maxima of observed data belonging to the same LPJ grid cell.

Section 4.1.3 Where are the LAI data shown that the addition of N biogeochemistry supposedly fixes? If this is the big advancement with the model presented here should these data also be shown? Is it just the addition of N biogeochemistry that's responsible for the proported improvements, or were other parametric or structural changes made?

Answer: In previous versions without accounting for nitrogen dynamics, the intensity of crop production had been calibrated by scaling the maximum LAI value, the harvest index and the factor for scaling leaf-level photosynthesis to stand level (alpha_a) of each crop at the national scale as described in Fader et al. (2010). Without this calibration, yield levels are typically overestimated, especially in low-input regions (blue lines in Figures 10, S21-23). Significant parts of this calibration

are now no longer necessary, as N limitation already reduces simulated crop yields in low-input countries. However, the N limitation is not the only mechanism that leads to lower crop yields, so that a further calibration could be performed. However, we did not calibrate yields here to show that N limitation already leads to a better representation of low-input crop production. This is now better discussed in the text (now section 4.4.3).

Fig. 10 A bunch of questions: What are the residuals and how are they calculated? What are units for the y-axis (what is the 'FM')? What are the little numbers in the top of each panel showing? Finally, it looks like global maize production increases over the period shown but the models are all flat. What's driving the increase in yields that the model is apparently missing? Is this true for other major crops?

Answer: FM depicts "fresh matter" (as used for all data reported by the FAOstat) and tFM stands for ton fresh matter and is now explained in the figure caption. The residuals are the detrended observed and simulated values, by subtracting a moving average as described by Müller et al. (2017). The numbers in the panels show the correlation coefficients between the residuals of the FAO-stat data and the residuals of the simulation (dashed lines). This is now also explained in the text. Indeed, simulated crop yields show no trend, as all management assumptions (including fertilizer input) are static, whereas technological advances are driving the observed yield trend. This is now also discussed in the text.

P 26 L 4, What improvements are necessary? What additional complexity may improve things further? What data are critical to getting terrestrial C-N dynamics less wrong? As presented these are kind of empty / throwaway statements. Can the be flushed out with some more detail, both in the main text and in this summary conclusion?

Answer: Thanks for pointing this out. We now discuss further possible improvements in more detail, mainly incorporating aspects that have been implemented in parallel in the carbon-only version (LPJmL4.0, Schaphoff et al. 2017a,b), such as the improved tree phenology, but also more detail in the representation of agricultural management, such as tillage, or fertilizer type and timing.

References in answers:

Ascott, MJ, DC Gooddy, L Wang, ME Stuart, MA Lewis, RS Ward, and AM Binley. 2017, Global patterns of nitrate storage in the vadose zone, Nat. Commun., 8, 1416, doi: 10.1038/s41467-017-01321-w.

Fader, M, S Rost, C Müller, A Bondeau, and D Gerten. 2010, Virtual water content of temperate cereals and maize: Present and potential future patterns, Journal of Hydrology, 384, 218-231, doi: 10.1016/j.jhydrol.2009.12.011.

Kabala, C, A Karczewska, B Galka, M Cuske, and J Sowinski. 2017, Seasonal dynamics of nitrate and ammonium ion concentrations in soil solutions collected using MacroRhizon suction cups, Environ Monit Assess, 189, 304, doi: 10.1007/s10661-017-6022-3.

Müller, C, J Elliott, J Chryssanthacopoulos, A Arneth, J Balkovic, P Ciais, D Deryng, C Folberth, M Glotter, S Hoek, T Iizumi, RC Izaurralde, C Jones, N Khabarov, P Lawrence, W Liu, S Olin, TAM Pugh, DK Ray, A Reddy, C Rosenzweig, AC Ruane, G Sakurai, E Schmid, R Skalsky, CX Song, X Wang, A de Wit, and H Yang. 2017, Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications, Geosci. Model Dev., 10, 1403-1422, doi: 10.5194/gmd-10-1403-2017.

Schaphoff, S, M Forkel, C Müller, J Knauer, W von Bloh, D Gerten, J Jägermeyr, W Lucht, A Rammig, K Thonicke, and K Waha. 2017, LPJmL4 - a dynamic global vegetation model with managed land: Part II – Model evaluation, Geosci. Model Dev. Discuss., 2017, 1-41, doi: 10.5194/gmd-2017-146.

Schaphoff, S, W von Bloh, A Rammig, K Thonicke, H Biemans, M Forkel, D Gerten, J Heinke, J Jägermeyr, J Knauer, F Langerwisch, W Lucht, C Müller, S Rolinski, and K Waha. 2017, LPJmL4 – a dynamic global vegetation model with managed land: Part I – Model description, Geosci. Model Dev. Discuss., 2017, 1-59, doi: 10.5194/gmd-2017-145.

Walvoord, MA, FM Phillips, DA Stonestrom, RD Evans, PC Hartsough, BD Newman, and RG Striegl. 2003, A Reservoir of Nitrate Beneath Desert Soils, Science, 302, 1021.

References: Le Quéré C, et al. (2015) Global Carbon Budget 2015. Earth System

Science Data 7: 349-396. DOI:10.5194/essd-7-349-2015

Smith et al. (2014) Implications of incorporating N cycling and N limitations on primary

production in an individual-based dynamic vegetation model, Biogeosciences,

11, 2027-2054, doi:10.5194/bg-11-2027-2014.

Zaehle, et al. (2015). Nitrogen availability reduces CMIP5 projections of twenty-firstcentury

land carbon uptake. Journal of Climate 28.6: 2494-2511.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-228,

2017.