

## **Referee #2**

*The work by Sun et al. is impressive in the sense that many analyses are taken to understand the Orchidee-CNP model. But on the other hand, the paper lacks a clear flow of arguments.*

**[Comment 1]** *The reason why every time the models need to be more detailed is that we are not satisfied by the performance of the old models. If we only focus on carbon and water (WUE) then we clearly see problems in the dynamics, sinks and sources, which was the reason to include Nitrogen and now also the Phosphorus cycle. **However, if one the main conclusions is that the current version of this model is unable to simulate carbon sinks, then the choices of expanding the model need to N and P should be much more discussed in detail. As long as we cannot model the carbon cycle, what kind of implications has this on the N and P cycle? If there is a large problem in land carbon sink estimates, then I would like to see the consequences to all other coupled processes, including water. If this is large, then this should be solved first or we should simply conclude that there is too less understanding to couple the models as proposed in a global model. An evaluation as proposed by the authors of this version of the model doesn't help us in answering this problem.***

**[Response to #1]** Thanks for this comment. We agree that including more details (e.g. nutrients) into models are done with the aim to reach a better model performance (this was described on Page 2 Lines 69-71). One of the aims of LSM is the quantification of the response of the land C balance to man-made environmental changes in the past, present and future. Here we show that the inclusion of nutrient cycles in ORCHIDEE-CNP did deteriorate the simulated land C sink in the Northern Hemisphere (NH) for recent decades compared to ORCHIDEE. However, we do show that ORCHIDEE-CNP performs better in simulating the underlying mechanisms than ORCHIDEE: e.g. sensitivities of ecosystem productivity to increasing CO<sub>2</sub> and to variation in water availability. Moreover, ORCHIDEE-CNP tends to better reproduce observed vegetation resource use efficiencies (see further details below). This suggests that ORCHIDEE 'got the right result for the wrong reason' or, in other words, both models 'cannot model the C cycle'. However, ORCHIDEE and ORCHIDEE-CNP perform well compared to other LSM as indicated by the iLAMB benchmarking tool (Friedlingstein et al., 2019). Thus, they reflect our current capabilities in modelling the carbon cycle.

Compared to ORCHIDEE, ORCHIDEE-CNP tends to improve the performance of resource use efficiencies, the sensitivity of plant productivity to increasing CO<sub>2</sub> (CO<sub>2</sub> fertilization effect), inter-annual variation of GPP in the northern hemisphere (NH). Besides, ORCHIDEE-CNP is able to better reproduce the variation of NBP in tropical regions. For NH, ORCHIDEE-CNP simulated a more realistic CO<sub>2</sub> fertilization effect, which can be explained by nutrient effects on plant carbon uptake in line with theory and observations (e.g. Jiang et al., 2020). The underestimated carbon sink in NH points

toward a driver of the NH sink which is not included in ORCHIDEE and ORCHIDEE-CNP, e.g. forest regrowth (Pugh et al., 2019). We also showed that ORCHIDEE-CNP underestimates P availability in the NH, thus another explanation is that the NH sink in this study is too low because of too strong P limitations in this region. The detailed explanation for the underestimated NH C sink can be found in Sect. 5.5.3. More developments are needed to improve ORCHIDEE-CNP in the NH.

We emphasized here that our evaluation for N and P together with C cycling goes well beyond the evaluation of other global and site scale CNP models (Wang et al., 2010; Yang et al., 2014; Goll et al., 2012; Fleischer et al., 2019). We argue that our study provides insights on the strengths and weaknesses of ORCHIDEE-CNP and thus allows us (1) to define model applications for which realistic predictions could be expected (e.g. tropics) and (2) to interpret the model behavior.

We realized that by our efforts in identifying the underlying reasons for the model biases, the focus of our original manuscript was shifted towards simulating the land C sink, away from the intended focus on the evaluation of key aspects governing the coupling of the C cycle to nutrient cycles. In the revised manuscript, we strengthen the focus on the evaluation for 4 key emerging model properties related to nutrients: (1) vegetation resource use efficiencies, (2) the response of GPP to increasing CO<sub>2</sub>, (3) ecosystem N and P turnover and openness, and (4) large-scale pattern of ecosystem stoichiometries (see details in **Response to #2 of referee #2**). The evaluation of land C sink serves as an example of the implication including nutrients in a major area of LSM application (i.e. dynamics of land C balance). In the revised paper, we reconstructed the result and discussion sections using this storyline to clarify the statements for pros and cons of ORCHIDEE-CNP as well as the ways to address the model biases.

The main focus of this paper is the nutrients-related emerging model properties and an implication on C cycle. Considering the paper is very comprehensive, we choose to follow the storyline mentioned above and not include additional evaluation for water or energy fluxes. We will focus in more detail on the effects of nutrients on water and energy fluxes in our follow-up studies.

**[Comment 2]** *A second major concern is therefore that the evaluation is far too broad while missing the in-depth analyses. The number of figures are too many and jumping from one type of comparisons to another: a. On one hand, you are showing the dynamics, but then **I would like to have much more information on understanding the drivers**. For instance, how much dependent is the dynamics of P and N on the P and N deposition? There are studies who have shown in other models that this N Deposition is one of the main drivers. b. Then you make some snap shots of global patterns, while later on you focus on different ecoregions and then different soil types and then on vegetation with different photosynthetic pathways. It would be very helpful **to structure this far better and to integrate those aspects**.*

**[Response to #2]** Thanks for this comment. An issue we face with the evaluation of global models is the limited availability of data for evaluation covering only a small subset of simulated variables. In addition, the temporal and spatial coverage of the scarce data varies among the datasets. To improve the structure of the presentation of our analysis, we focus in the revised manuscript on key aspects for the coupling of the cycles of C, N and P (see previous answers) and moved a large part of the analysis to the SI.

To better ‘integrate the various information’, we focus now on 4 key aspects related to the effects of nutrients on the simulated response of ecosystem productivity and carbon storage to (changes in) climate and CO<sub>2</sub>: (1) vegetation resource use efficiencies, (2) the response of GPP to increasing CO<sub>2</sub>, (3) ecosystem N and P turnover and openness, and (4) large-scale pattern of ecosystem stoichiometries. Point (1) and (2) control the response of vegetation carbon uptake which operates on timescales of years to decades, while point (3) and (4) control the response of the ecosystem carbon storage potential which operates on timescales of centuries and longer. This has been stated in the revised introduction section (Page 3; Lines 101-109). The choice was further based on the availability of observational data.

We agree that the depth of our analysis has been obscured in the original manuscript due to the various aspects we have analyzed and which have not been linked well enough in the discussion. We hope that by the narrower focus of the revised manuscript on key factors, we are now able to provide a ‘depth analyses’ of such effects.

Our analysis does not aim at disentangling the drivers of the land carbon sink. The land sink part is merely an example for the implications of including nutrient cycling in ORCHIDEE (as stated now in the introduction Lines 108-109). We argue that the identification of underlying drivers (e.g. CO<sub>2</sub>, nutrient deposition etc.) is out of the scope, in particular as there exists large uncertainty with respect to reconstructions of phosphorus deposition and the sensitivity of ecosystem carbon storage to nutrient deposition (Wang et al., 2017) which requires a study on its own.

**[Comment 3]** *A third major concern is the for me **random way of comparing the results**. I found the **comparison with only ORHIDEE-C not very convincing**. Why not comparing to the average performance of the land models as done in TRENDY? There are other global model results as well on C and N. For instance, LPJ guess.*

**[Response to #3]** Thanks for pointing this out. We want to stress that the results of ORCHIDEE-CNP are not only compared to ORHIDEE-C, but also compared to data-driven products and observations. The comparison to ORCHIDEE-C serves the purpose to separate the effect of including nutrient cycles on the simulated C cycle, as ORCHIDEE-CNP differs primarily from ORCHIDEE-C with respect of having cycles of N and P. This is not possible using results from other models as there are various

additional differences making it nearly impossible to explain differences among models with certainty (e.g. Rogers et al., 2017). Instead of comparing our simulated NP flows to results from other land surface models as suggested by the referee, we compared them to results from the model-data-synthesis tool GOLUM-CNP which provides more robust estimates of CNP flows than a LSM. The structure of GOLUM-CNP allows the assimilation of observational data from various sources, which is not possible in LSM due to their complexity. Nonetheless, we use results from the Trendy model ensemble (i.e. the ensemble average which has been demonstrated to perform substantially better than any single model for various aspects) for the land C balance (as for example done in Global Carbon Project; Friedlingstein et al., 2019).

**[Comment 4]** *L402: why do you have a smaller natural land cover? Is it a problem from ORCHIDEE or from GOLUM-CNP. Is it then still useful to compare?*

**[Response to #4]** Thanks for this question. ORCHIDEE-CNP uses a different land cover than the one used to upscale results from the biome-scale model GOLUM-CNP. The main difference in land cover originates from the omission of managed land area in GOLUM-CNP. As the biome-scale model GOLUM-CNP does not resolve spatial variation within a biome, we compared the results on biome scale for nutrients use efficiencies and ecosystem N and P residence time. When comparing the global spatial pattern of ecosystem nutrients openness, we masked the areas with managed lands (agricultural and pasture lands) >50% in ORCHIDEE-CNP. Therefore, this problem does not exist in our study.

For the comparison of C, N and P flows and storages with GOLUM-CNP, we only compared the values per area (unit: g C/N/P m<sup>-2</sup>) and removed the panel with global values (unit: Pg C/N/P) in the revised manuscript (Sect. 4.1). We admit that the comparisons on a per area basis for some highly sensitive variables to agricultural activities (e.g. P leaching) are not real valid. Thus, we stated this issue (Page 5; Lines 328-335) and did additional comparisons for those variables using more valid datasets that considering agricultural activities (Sect. S5 in the supplement).

**[Comment 5]** *Comparing global news N-leaching: other forcings → this doesn't help us in understanding the role of the different mechanisms → can you also compare it with similar forcings? If not, is it still valid to include this comparison?*

**[Response to #5]** Thanks for this question. This study is the first among all current P-enabled LSMs studies to evaluate the N and P loads from land to rivers on both basin and global scale. Despite the different forcings used, ORCHIDEE-CNP simulated N and P loads from land to rivers **in the same order of magnitude** as the GlobalNEWS2 model. After we reconstructed the storyline in the revised paper (see Response to #1 and #2), all of the information for N and P leaching went into the SI (Sect. S5 in the supplement) and serves as the explanation for biased nutrients turnover and openness (Sect.

4.4 and 5.3). Although we agree that the investigations for N and P leaching drivers are valuable, we argue that this is out of the scope of this study. The main purpose of our evaluation is to provide a whole picture of the current states of C, N and P dynamics by ORCHIDEE-CNP rather than to understand the drivers of changes the different mechanisms for specific process (also see **Response to #2**).

**[Comment 6]** *How did you downscale from HYDE3.2 to 1x1 km? Did you use the same allocation rules as done by Klein Goldewijk for the 30 minute resolution?*

**[Response to #6]** We did not downscale HYDE3.2 to finer resolution. In contrast, we aggregated it to coarser resolution of  $2^{\circ} \times 2^{\circ}$  to constrain our land-cover maps. We corrected the description of the historical land-cover maps for ORCHIDEE-CNP as follows (also in the Sect. 3.1.3 of the revised manuscript; Page 5; Lines 169-183).

“The historic land-cover change maps were based on the European Space Agency Climate Change Initiative (ESA-CCI) land-cover data (Bontemps et al., 2013). To be used by global vegetation models ORCHIDEE-CNP, ESA-CCI land-cover data were aggregated to  $2^{\circ} \times 2^{\circ}$ , and grouped into PFTs using the reclassification method from Poulter et al. (2011, 2015). The fraction of cropland and pasture in the PFT map was further constrained by the cropland area and the sum of pasture and rangeland area of the year 2010 in the History Database of the Global Environment land use data set (HYDE 3.2; Klein Goldewijk et al., 2017a, b) respectively, which were also aggregated to  $2^{\circ} \times 2^{\circ}$ . The above processes produced a reference ESA-CCI-based PFT map for the year 2010. The land-use changes derived from and Land-Use Harmonization (LUH) v2 (<http://luh.umd.edu/data.shtml>; an update release of Hurtt et al., 2011) were aggregated to  $2^{\circ} \times 2^{\circ}$  and then were applied to this reference PFT map to constrain the land-cover changes of forest, grassland, pasture and rangeland, and cropland during the period 1700-2017 using the backward natural land cover reconstruction method of Peng et al. (2017). As a result, a set of historic PFT maps suitable for global vegetation models were established distinguishing global land-cover changes for the period of 1700-2017 at  $2^{\circ} \times 2^{\circ}$  resolution.”

**[Comment 7]** *L890: N and P leaching: the current problem of understanding N and P leaching is large: there are all kind of confounding factors that determine these leaching rates which are in the end extremely important to understand water quality and functioning of the system. In the current paper I cannot find this sensitivity back. Leaching is highly dynamic due to fire, soil water fluxes by extremes, different season lengths depending on ecoregion and latitude etc.*

**[Response to #7]** As we argued before, understanding the role of the underlying mechanisms for N and P leaching is out of the scope of this study (see **Response to #2** and **#5**). Besides, we cannot explore the sensitivity of N and P leaching to fire, extreme events or erosion as they are currently not resolved in ORCHIDEE-CNP.

## References

- Friedlingstein, P., Jones, M., O'Sullivan, M., Andrew, R., Hauck, J., Peters, G., Peters, W., Pongratz, J., Sitch, S., and Le Quéré, C.: Global carbon budget 2019, *Earth Syst. Sci. Data*, 11, 1783-1838, <http://doi.org/10.5194/essd-11-1783-2019>, 2019.
- Goll, D. S., Brovkin, V., Parida, B., Reick, C. H., Kattge, J., Reich, P. B., Van Bodegom, P., and Niinemets, Ü.: Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling, *Biogeosciences*, 9, 3547-3569, <http://doi.org/10.5194/bg-9-3547-2012>, 2012.
- Jiang, M., Medlyn, B. E., Drake, J. E., Duursma, R. A., Anderson, I. C., Barton, C. V., Boer, M. M., Carrillo, Y., Castañeda-Gómez, L., Collins, L., Crous, K. Y., De Kauwe, M. G., dos Santos, B. M., Emmerson, K. M., Facey, S. L., Gherlenda, A. N., Gimeno, T. E., Hasegawa, S., Johnson, S. N., Kännaste, A., Macdonald, C. A., Mahmud, K., Moore, B. D., Nazaries, L., Neilson, E. H. J., Nielsen, U. N., Niinemets, Ü., Noh, N. J., Ochoa-Hueso, R., Pathare, V. S., Pendall, E., Pihlblad, J., Piñeiro, J., Powell, J. R., Power, S. A., Reich, P. B., Renchon, A. A., Riegler, M., Rinnan, R., Rymer, P. D., Salomón, R. L., Singh, B. K., Smith, B., Tjoelker, M. G., Walker, J. K. M., Wujeska-Klaue, A., Yang, J., Zaehle, S. and Ellsworth, D. S.: The fate of carbon in a mature forest under carbon dioxide enrichment, *Nature*, 580(7802), 227–231, <http://doi.org/10.1038/s41586-020-2128-9>, 2020.
- Pugh, T. A., Lindeskog, M., Smith, B., Poulter, B., Arneeth, A., Haverd, V., and Calle, L.: Role of forest regrowth in global carbon sink dynamics, *Proc. Natl. Acad. Sci. U. S. A.*, 116, 4382-4387, <http://doi.org/10.1073/pnas.1810512116>, 2019.
- Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., von Caemmerer, S., Dietze, M. C., Kattge, J., Leakey, A. D. B., Mercado, L. M., Niinemets, Ü., Prentice, I. C., Serbin, S. P., Sitch, S., Way, D. A., and Zaehle, S.: A roadmap for improving the representation of photosynthesis in earth system models, *New Phytol.*, 213, 22– 42, 2017.
- Wang, R., Goll, D., Balkanski, Y., Hauglustaine, D., Boucher, O., Ciais, P., Janssens, I., Penuelas, J., Guenet, B., and Sardans, J.: Global forest carbon uptake due to nitrogen and phosphorus deposition from 1850 to 2100, *Glob. Change Biol.*, 23, 4854-4872, <http://doi.org/10.1111/gcb.13766>, 2017.
- Wang, Y., Law, R., and Pak, B.: A global model of carbon, nitrogen and phosphorus cycles for the 1610 terrestrial biosphere, *Biogeosciences*, 7, <http://doi.org/10.5194/bg-7-2261-2010>, 2010.
- Yang, X., Thornton, P., Ricciuto, D., and Post, W.: The role of phosphorus dynamics in tropical forests- a modeling study using CLM-CNP, *Biogeosciences*, 11, <http://doi.org/10.5194/bg-11-1667-2014>, 2014.