

Biogeosciences Discuss., author comment AC2
<https://doi.org/10.5194/bg-2022-114-AC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Lin Yu et al.

Author comment on "Improved representation of phosphorus exchange on soil mineral surfaces reduces estimates of phosphorus limitation in temperate forest ecosystems" by Lin Yu et al., Biogeosciences Discuss., <https://doi.org/10.5194/bg-2022-114-AC2>, 2022

Reply to "Comment on bg-2022-114 RC2"

This paper by Yu et al. describes a new algorithm to better represent soil phosphorus sorption dynamics in a terrestrial biosphere model – QUINCY. The authors proposed the use of a double-surface Langmuir isotherm to better capture the non-linear relationships between solution P and labile P pools in the soil. They performed a review on both published data and model assumptions on P sorption. They then compared their simulation against data at a range of P availability sites, performed sensitivity on

parameters and compared simulations for CO₂ and P enrichment scenarios. They argued that the double-surface Langmuir isotherm is a better modeling scheme because it simulated observed pattern of soil organic pools well, it maintained a relatively stable solution P pool to act as a buffer against instability, which then led to less P limitation at the P-poor site, and it led to improved simulation of folia N and P concentration.

Authors: *we thank the referee for acknowledging our work.*

Overall, this is a clearly-written manuscript. The rationale and objectives are crystally clear. The discussion is also well written. My comments mostly focus on two aspects of the results that I want to discuss with the authors and receive their clarifications:

- Did it indeed lead to improved estimate relative to the conventional single surface approach? All models performed well for reproducing the measured SOC etc. as reported in figure 3. The novelty of the double-surface scheme, as the authors argued, is that it better reproduced the ratio between Plab and exchangeable Pi (L190-191; Table 3). But looking at Table 3, the statistical significance is relatively weak ($p = 0.014$ for lab-to-exchangeable P ratio, and 0.044 for SIP). At the same time, I wonder if the new scheme actually increase model complexity or not. May be the authors should make a paragraph discussing whether the gained benefits in terms of improved simulation accuracy is worth the added complexity, if there's indeed additional complexity associated with the new scheme. In particular, does it require additional parameters relative to the conventional approach? And, if we want to constrain the parameters in the new model scheme, what data collection should we make? If it doesn't involve additional complexity, I think it's very useful to highlight.

Authors: As we specifically calibrated the SOM profile for each site at each depth to avoid the side-effects of organic cycling, the good model fit of SOM profile in Fig.3 is the outcome of model calibration. The main improvement in Fig. 3 and Table 3, is the improved soil inorganic P simulation.

The new double Langmuir isotherm in QUINCY does require more parameters compared to other TBMs, but in this study, the siLang and dbLang shared the same set of parameters of QUINCY, i.e. the siLang also takes account the effect of soil texture, pH and extractable Al/Fe in the P sorption parameterization. Therefore, the dbLang complexity is the same as siLang in this paper, but unfortunately, we did not compare the QUINCY parametrizations with other TBMs in this study, as shown in Table 2, the original version of QUINCY (Thum et al. 2019) used Qmax and Km values which are very different from other TBMs. It is mostly because we already use soil texture and SOM content to parameterize the Qmax and Km in the original QUINCY, as many other TBMs directly use prescribed values.

The new parameterizations are dependent on soil texture, pH, and extractable Al/Fe. As far as we know, the biggest difficulty for upscaling is the extractable Al/Fe.

- What does it mean for the land C sink estimates under future rising CO₂ if the model simulated a less P limitation at the P-limited site. As the authors introduced, there has been a lot of model development to add P-cycle into models. The relative magnitude of the P limitation is obviously different, but one of the crucial argument for the inclusion of P-cycle in models is that they would impose additional processes to constrain ecosystem productivity for P-poor regions of the world. The new scheme seemed to alleviate the extent of P limitation, and therefore I wonder how does it compare to a simulation without the P-cycle turned on. Do you obtain similar CO₂ responses for the Plimited site? Obviously the CN-only simulation does not have the capacity to accurately reflect the processes limiting CO₂ responses at the P-poor site, but it would be interesting to see if there's indeed difference between the two approaches.

Authors: thanks for the very intriguing point. The initial motivation for this study is that, the field observations of our five study sites (along a soil P gradient) do not show a clear trend in biomass, productivity, or foliar P content, as most of the observed differences were seen in soils, such as the SOM stocks and quality, the root biomass, the microbial biomass, and the root/litter P contents. It indicates that the temperate ecosystem has mechanisms to adapt to the varying soil P availability, and such mechanisms were not represented in our TBMs yet. In other words, the simulated high P stress by siLang model does not exist in reality.

The double Langmuir isotherm alleviate the P stress, but only to a certain extent. Because it only releases the buffering capacity of soil mineral surfaces, but not yet changes the organic P cycling part, which is believed to be the dominant processes in the low P system. We specifically discuss it in Sect. 4.4 to point out that this new scheme is not enough to resolve the P cycling features in extremely low P ecosystems, such as Amazon or Australian eucalyptus.

We believe the main processes regulating the P limitation are not yet implemented in the TBMs yet, such as the regulation of microbial cycling, the phosphatase production, the resorption within plant and etc. We have been working towards this direction, but as the referee mentioned, the control

simulation in our study only reflects the CO2 responses induced by the current model processes, and it is a promising future direction to include and quantify the effects of these processes.

Lastly, one question I have, which isn't a criticism per se, is that why the authors didn't use the Jena Soil Model (that they developed) to investigate the effects of different soil P sorption functions. It appears to be a soil question and therefore I wonder is there any particular reason that deemed JSM unsuitable for this work?

Authors: at the time when this study started, the JSM is not yet fully functional with QUINCY vegetation processes. Now it is possible to test different sorption functions with JSM enabled, and it is for sure something we would like to test in future that what is the role of inorganic versus organic cycling processes given different soil P availabilities, as we believe more mechanisms of organic P cycling processes are included in JSM which would facilitate such tests.

Specific comments:

Title: replace P with phosphorus.

Authors: thanks, will revise

Figures: Figure presentation needs to improve quite significantly. Figures are currently in low quality resolution. Units and variable names in the figure legend/axis need to be properly labeled.

For Figure 3, the authors may need to think of a better way to show the results, as it's not very clear to see differences in panel a and b (but maybe because they are similar and therefore it's not important to show the differences?). But still, the 4pool and control color are too hard to differentiate from each other.

Authors: thanks, we will remake the figures to better present the results.

Table 3. One could argue the statistical significance is very weak.

Authors: yes, agreed. since we only have 5 sites, it is not very likely to yield very low p values in the paired-t-test. Fig. 3 presents the visual difference between different model performances, but we think it might still be necessary to conduct a simple statistic analysis to confirm the pattern. We will be careful in interpreting the results.