REMARK: Please find in the following our response to the referees comments. In addition to this official review we revised the manuscript in response to a personal communication reaching use after the discussion period was closed. Our response to the latter can be found afterwards.

All line numbers refer to the manuscript with marked changes.

Referee #1:

This manuscript describes the implementation of a terrestrial phosphorus cycle into the land surface model ORCHIDEE. The authors used the fertilization experiments at two sites along the Hawaii chronosequence to evaluate the nutrient representation in the NP-enabled ORCHIDEE. With the increasing realization of the important role of phosphorus cycle in affecting global carbon cycle, this work is timely. Overall the manuscript is well written as a modelling paper. I just have some suggestions and edits as shown below.
DSG: We thank the referee very much for his constructive comments on our work.

General comments:
1) Introduction a bit too short – maybe it is sufficient for a modeling paper on GMDD but I think a bit more on how this representation of P cycle is different from other P models will position this work better
DSG: We fear a fruitful discussion of existing model concepts would make the introduction very long and distract from the the main scope. Nonetheless, we added a sentence to the introduction to stress the novel concepts we developed here: "It is the first global phosphorus models which explicitly simulates root uptake of dissolved phosphorus accounting for soil moisture effects on soil phosphorus mobility."

2) I would like to see a model diagram that shows the phosphorus pools and fluxes and maybe with major phosphorus processes. Although those have been described here and there in the manuscript, a diagram will help the reader to better understand the model and link them with many equations in the text.
DSG: We added a Figure (Figure 1) with pools and fluxes of the phosphorus cycle.
3) Figure 2 seems a repetition of a subcomponent of Figure 1 and can be emoved,
DSG: We removed Figure 2.
4) Labile phosphorus was used in the text to describe both labile phosphorus in plant and soils and it can be confusing sometimes.

DSG: we now consistently use 'soil labile phosphorus' / 'dissolved labile phosphorus' for labile P in soils and 'plant labile phosphorus' for labile P in plants throughout the text.

Minor comments:
Page 2, line 10: change "rise" to "rises"
DSG: corrected
Page 2, lines 22-25: acronyms were used (Ama, Euc, IMB, AFE…….), not sure what they are and no references for them
DSG: these are abbreviations are due to the latex citation style file. This will be corrected with the help of the typsetter at a later stage.
Page 3, line 20: should be "except for"
DSG: corrected
Page 3, eq. 2: is this eq. complete?
DSG: equation was incorrect. we corrected the equation
Page 4, lines 12-14: this is not clear to me
DSG: we rephrased the text to
"As the concentration of phosphorus in roots is orders of magnitudes larger than the concentration in the soil solution, passive uptake of phosphorus via diffusion is negligible (Schachtman et al., 1998). Thus, only active uptake via specialized transporters on the root surface is accounted for in the model."
Page 5, eq. 3: umax – maximum root uptake capacity and vmax ->maximum uptake capacity of roots. Seems the same thing to me – need to be better defined. Also, I could not get the units on the two sides of the eq. consistent
DSG: equation was incorrect. we corrected the equation
Page 6, line 22: " the actual value of fPNplant may be higher than 1" – what is the implication for plant P uptake? Is it realistic?
DSG: This is the justification for the min() in the equation which was missing. we corrected the equation
Page 8, eq 12: how C growth is scaled? Photosynthesis or NPP? Equation will be helpful here
DSG: we added the equation EQ13
Page 9, section 2.1.3: This section is very general – not P specific. Could be removed or make it more focusing dynamics in litter and soil organic matter
DSG: The phosphorus dynamics in litter and soil organic matter follow the original model. We report the general concepts here. We would like to the keep this section for completeness, but can remove it to the appendix if the editor thinks this is needed.
Page 9, line 27: the turnover time of phosphorus is set to half the turnover times used for biological mineralization of organic matter – what is based on? Or Any reference?

DSG: We added "are set arbitrarily, due to the lack of observational constraints,"
Page 10, line 2: from primary minerals
DSG: corrected
Page 11, eq 27: should it be the other way around? When diffusion flux is greater than uptake, there is no change in the difference in labile P between root surface and the surrounding
DSG: corrected
Page 12, section 2.1.7: this section only deals with N fixation – can be put in appendix if needed
DSG: moved to appendix
Page 12, section 2.1.8:
DSG: ?
Page 19, line 6: should be "due to"
DSG: corrected
Page 20, Table 4: I feel the comparison between simulated and observed can be better shown with a chart instead of a table
DSG: we prefer to show the results in a table. no changes done.
Page 20, lines 8-10: I thought the PFT used here is tropical evergreen instead of tropical deciduous
DSG: this is just the phenology, there is still climate(drought) related mortality
Page 24, Line 14-15: foliar P concentration is much more variable than N concentration, could it be due to that the pre-defined foliage P concentration range is too narrow?
DSG: The leaf N:P range is prescribed from data from a global plant trait database with extensive data coming from nearby location. Thus the confidence is high, however the concepts applied to adjust stoichiometry itself are problematic and a cause of the mismatch between observed and simulated adjustments. Thus we added: "However, the stoichiometric adjustments in ORCHIDEE are not process-based and might be itself the cause for the bias in the simulated response of leaf stoichiometry."

<u>Personal communication / Referee #2:</u>
Comment on "A representation of the phosphorus cycle for ORCHIDEE" by D. S. Goll et al.

The manuscript by Goll et al. describes a new representation of phosphorus (P) cycle in theORCHIDEE model. The authors use observed data from two hronosequence, fertilization experiments in Hawaii to calibrate and evaluate the model. The impact of nutrients has been shown to be very important to global carbon cycle under climate change and increasing CO2 in many independent studies. The introduction of P cycle in land carbon models helps to improve our understanding on

how P interacts with carbon (C) and nitrogen (N) cycles and thus associated future climate change. Therefore, the manuscript is in the scope of Geoscientific Model Development. Nevertheless, there are several issues which need to be addressed.

DSG: We thank the referee very much for is constructive and in depth comments on our work. We believe they substantially improved the quality of our work.
All line numbers refer to the manuscript with marked changes.

Major comments:

1. The manuscript presents a new P cycle parameterization added in a C-N cycle coupled land surface model, but it is not clear what the difference is in C/N cycle when P cycle is implemented in comparison with the original C-N model. A quantity analysis on C, N pools and fluxes with and without P cycle would be easier for readers to follow the impacts of P inclusion in the model. The authors only present results from the P model simulation. It would be important to show that including P cycle at least does not degrade the simulation by model with just N limitation.

DSG: The structure of the model does not allow to turn on/off the nutrient cycles in a straightforward way like it is possible with simpler models like CLM (Yang et al., 2014) or JSBACH (Goll et al., 2012) in which phosphorus affects carbon fluxes only via a minimum law. Technically we can run the model assuming an optimal nutrient availability, but these simulations deviate strongly from the observation as indicated by the strong simulated response to nutrient addition at the two sites presented here.

The rationale behind the nutrient manipulation experiment presented here is to evaluate the nutrient cycle and their effect on plant biomass and productivity. As stated in the introduction (Page2, L15-21), it is shown and now widely accepted in the community that manipulation experiment are best suited to evaluate the structure of a model. We therefore do not see much advantage of presenting simulation in a carbon only mode (assuming this would be technically feasible which it isn't).

No changes to the manuscript.

2. The length corrected spin-up seems problematic to me because it assumes that the model is able to reproduce the evolution of land ecosystem before its equilibrium state, but usually it is not the case. Most commonly spin-up consist of initializing both physical and biogeochemical variables before running the model to equilibrium. The

integration time needed by the model to converge towards its own attractor is different from the true state of the natural ecosystem (Séférian et al. 2016). So the 300yr model period during spin-up should not be treated as the 300yr actual time of natural system. I may drop the way in Appendix B and (1) run the model until the C, N and P pools comparable to observed values (Table2, Crews et al. 1995), (2) ensure that the carbon accumulation rate in vegetation and soil fall in the range of the ~10yr observation. This ensures that the model state is comparable to the true state at the end of spin-up, which is required by starting the fertilization experiment.

DSG: We agree with the referee that these model might not be able to reproduce the pace at which land ecosystem converge towards its equilibrated state.

As a response to the referees concerns, we show the evolution of NPP, total biomass, total soil organic matter during the spinup simulation in Figure R1. The simulated variables for the year 234 (which correspond to the first year of climatic forcing period) are closest to the observation (Figure R1). This year is very close to the length of 230 yr we deduced from nitrogen accumulation estimates.

Due to the small differences, we prefer to keep the 230 yr length. The good agreement of simulated and observed carbon fluxes and stocks shows that the model is, despite we cannot assume such behaviour per se as the referee stated, able to capture the state of the ecosystem (Figure 3 & Table 4).

We state this more clearly now (P30,L19):
"As the simulated NPP , biomass, and soil organic matter match the observation (Table 4) and the dynamically simulated BNF rates are with 2.25 g(N)m−2 yr−1 well within the observed range, the reduction of the simulation length seems appropriate."

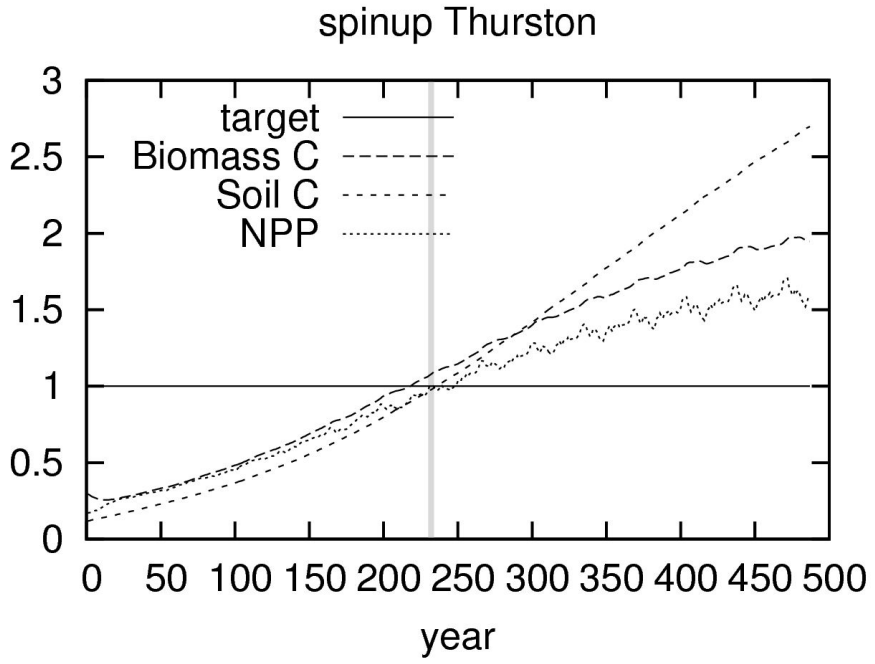

Figure R1: Net primary productivity, total biomass, and soil organic matter during the spinup simulation. Shown are 30yr running averages of simulated values divided by the observation flux of pool for the site. If the variables equals 1 it means simulated and observed variables are equal.

3. It is worthy of looking at why there is so large interannual variability of NPP as simulated by the model. Figure 3 shows that compared to the observation, the interannual variability of NPP is much higher than observations at both the N-limited and P-limited sites. This implies there must be some parameters or processes too sensitive to model forcing. A sensitivity analysis would help improve the understanding of the model dynamics. The authors attribute the larger interannual variability of the 4.1Myr site to an amplification of the drought induced growth reduction under low P availability. It is not very convincing because it only shows monthly values within a single year (Figure A1). However fluxes with large monthly variability do not necessarily combine to fluxes with strong interannual variability. I am wondering why the authors not present the annual results to sddress this issue directly?

DSG: We agree with the referee about the shortcomings of our analysis. However, a detailed analysis of the inter-annual variability would blow up the manuscript to an extent we fear the main points we make with this work would be diluted. Hereby, the lack of any information about the interannual variability of the two Hawaii site poses a big problem. We thus remove this analysis and will address the interactions of phosphorus limitation and interannual variability of NPP using an appropriate observational dataset in a follow up study.

4. It is unclear what processes are responsible for the occurrence of the N limitation at the 300yr site and the P limitation at the 4Myr site. Further analysis and discussion is missing to clarify why the model simulates contrasting responses to the same fertilization treatment at the N-limited and P-limited sites as shown in Figure 4.
DSG: We adjusted our simulation setup so that the only differences between the two sites are a) organic matter accumulation rate, b) phosphorus inputs, c) labile phosphorus sorption. Thereby, the changes in the N vs P limitation is solely an outcome of the different soil development stages following the classical concept by Walker and Syres (1976).

We modified the method section accordingly (P15,L25-30): ***"We account for changes in the sorption characteristics of volcanic soils are they develop. For the 300\unit{yr} site, we use the average value of k_s for Andisols from \citep{Yang2011}. For the 4.1\unit{Myr} site k_s was scaled with the relative difference in soil phosphorus sorption capacity between the two sites as computed dynamically in the P-enabled version (Violette in prep) of the mechanistic weathering model WITCH \citep{Godderis2006}."***

All values in text, tables and figure were updated with the simulated value from the new simulation. We took advantage of our recent code developments and use a newer revision of ORCHIDEE (r4520).

We now explicitly state in the discussion the mechanism resulting in the contrasting nutrient availabilities (which are now reported in Table 4):

(P21,L7-14): ***"The differences in stoichiometry mirror differences in the respective availabilities of mineral nitrogen and soil labile phosphorus (Table~\ref{tab:results}). While the concentration of mineral nitrogen is extremely low at the young site due to a high immobilization demand of accumulating soil organic matter, the concentration is high at the old site where immobilization demand is met by the mineralisation of nitrogen from organic matter. In case of phosphorus, the high phosphorus input of 434. \unit{mg m^{-2} yr^{-1}} at the 300\unit{yr} site keep soil labile phosphorus concentration high despite the high immobilization demand. At the the old site, the extremely low phosphorus inputs of 0.27 \unit{mg m^{-2} yr^{-1}} result in low soil labile phosphorus concentration as the ecosystem relies primarily on the mineralisation of phosphorus from soil organic matter. "***

and

(P22/23): ***"The 300\unit{yr} site is accumulating organic material, in particular soil organic matter, and the accompanied immobilization of soil nutrients is the major driver of nutrient scarcity. This leads to extremely low mineral nitrogen concentration. whereas the high weathering release of phosphorus from minerals is sufficient to keep soil labile phosphorus concentration relatively high (Table~\ref{tab:results}). Therefore, vegetation reacts strongly to the addition of nitrogen at the young site. The lack of any stimulation of plant productivity in the model to phosphorus addition at the young site indicates an overestimation of plant available phosphorus likely due to the omission of differences in the occlusion rate of soil labile phosphorus among sites which tends to be much higher at the young site (Violette in prep). At the 4.1\unit{Myr} site the remobilization of phosphorus from soil organic matter is the major source of phosphorus for vegetation, as the minerals are phosphorus depleted leading to low soil labile phosphorus concentration Table~\ref{tab:results}). Compared to the young site, also a higher fraction of soil labile phosphorus is adsorbed to soil particles and thus not available to plants. Therefore, vegetation reacts strongly to the addition of phosphorus at the old site, but not to nitrogen addition."***

5. Section 2.2.3: I would prefer to compare Xuptake and Xcontent with observations instead of using the derived values XUE, and Xprod and XMRT. Simpler analysis would be more straightforward and easier to understand the model's P cycle parameterization.
DSG: we agree that this variables are interesting and added the simulated values to Table 4.

Minor:
P2 L23-24: Abbreviations in the text should be written out (for example Ama; Euc; IMB;
AFE).
DSG: these are abbreviations are due to the latex citation style file. This will be corrected with the help of the typsetter at a later stage.
Section 2.1.1: The parameterization of the relationship between leaf nitrogen concentration
and Vcmax needs improvement with mathematical descriptions.
- What is the relationship between Nleaf, Nleaf,h*, Nstr Please clarify using formulas.
DSG: We added equation 2 (P3,L30)
Eq2 seems incomplete, what assumption is used for correction with Nstr?
- Missing is the description of how Jcmax is parameterized.
DSG: We already state in the text (P4,L1-2) ***"The electron transport capacity ($J_{max,h}$) is derived from $ V_{cmax,h}$ using the relationship from***

\citet{Kattge2007} which accounts for acclimation of photosynthesis to monthly temperatures." **See original publications for equations. No changes to manuscript.**

P5, Eq.3: The unit of the function not equal to g(P)g-1 (C)t-1

DSG: We corrected equation (Eq.4, P5)

P7, Eq.8: Representation of function needs recheck. Same parameterization for different
conditions?

DSG: We corrected the equation (Eq.9,P6)

Missing is the function of ftemp in Eq.6, and LF in Eq. 13, and the values of parameter gmax
in Eq.11.

DSG: We added the equation 5 for ftemp and added the value of gmax in Table 2. We already give the reference of how LF is calculated (P8,L3-5) . We think this is sufficient as this variables is independent of the nutrient cycles.

P9, Eq. 17: [...] is missing and should be added in Table A1.

DSG: We added the value to Table 2.

P10, Eq. 20: I assume the sign of [...] in the function should be negative as [...]. Please describe briefly the difference between [...] and [...] To make the text more readable, details of each components in Eq.20 should be described briefly with functions if possible.

DSG: We corrected the equation (Eq21,P10) and description of components were added with references to the respective equations (P10,L5-6).

P21, L11: Edit: Foliage N:P ratios less than 14…

DSG: corrected

P31, L21: typo? Should be "-44.7%" here.

DSG: corrected

Section 2.1.8: How the quality of Litter and SOM affects nutrient availability? A brief description of P immobilization and mineralization with equations would help readability.

DSG: We already refer to the original publications describing the dynamics of decomposition, which were not modified. We prefer to keep the model description focused on the novel aspects. No changes to the manuscript.

Table 1. Could the unit "30minutes-1" be replaced by "t-1
" according to the definition of "t" in Table 2?

DSG: corrected

the unit g(dryweight) should be converted to g(C) for consistency.

DSG: corrected; also units of parameter r_d in Table2.

Table 2. For parameters [...]
represents a unit by volume? Please clarify.

DSG: we corrected equation (Eq4, P5).

Table 3. The 300yr site should be “Kokee” but not “Koke”.
DSG: corrected throughout the manuscript
Fig 1. Adding a schematic representation of model structure as Yang et al. (2014) would be
beneficial for readers to understand the whole picture of the model.
DSG: we added a flowchart of the P cycle (Figure 1)

Fig 2. Duplicated information with Figure 1 but the rational very useful for explaining how nutrients affect plant production in the model. Better to move to the part of discussion and calculate the values of pools and fluxes to explain the different responses to N/P fertilization as shown in Fig4.
DSG: It is not straightforward to quantify the feedbacks from the simulation we performed as the different processes operate on different timescales. To avoid the redundant information we remove this figure.

Reference:
Séférian et al. 2016, Inconsistent strategies to spin up models in CMIP5: implications for
ocean biogeochemical model performance assessment.