Response Letter

Title: Soil nutrient competitive traits of plants, microbes, and mineral surfaces explain nutrient acquisition in tropical experimental manipulations

General Response:

We would like to thank the two anonymous referees for their constructive comments. In this revision, we used a standard function (gelman.diag) in the R package "coda" to calculate the Gelman-Rubin criterion. As suggested by the reviewers, we also added more discussion about model imperfection, parameter edge effects, and soil heterogeneity.

The response letter is organized by (1) reviewers' major comments in black and (2) authors' responses in blue.

Reviewer #1

1. Unfortunately, the authors seem to have used a wrong equation for the Gelman-Rubin diagnostic aka the "potential scale reduction factor", psrf. In Figure A2 the psrf is considerably below 1 for 6 parameters. However, the limit value is 1 and the psrf should be $\geq 1!$ Please check this equation again.

Response:

Thanks for pointing out the error. We coded the Gelman-Rubin in MATLAB with an incorrect equation. Instead, now we used the R function gelman.diag to calculate the Potential Scale Reduction Factor (PSRF). The new Gelman-Rubin convergence criterions are updated in Table 4 and Figure A2.

Figure A2. Gelman-rubin convergence criterion (solid lines). Baseline value is set to 1.1 (dash lines). When Gelmen-rubin criterion is smaller than or equal to 1.1, the chains are thought to converge.



2. Clearly, convergence is not reached, and I would advise the authors to refrain from using the term posterior. Instead they could state that they used the best parameter set obtained from MCMC sampling, while it was not possible to obtain a proper sample from the posterior.

Response:

Thanks for the suggestion. In the revised manuscript, we have replaced the term "posterior parameters" with "calibrated parameters".

Reviewer #2

1. The authors have tried to address the concerns about the non-perfect model calibration. Although I am not satisfied with their response, I recommend following the suggestion of reviewer 2 to clearly acknowledge the imperfections of the calibration and go on with conclusions that can be drawn despite of these calibration imperfections.

Response:

Thanks for the suggestion. We added more discussion in results and conclusion sections to clearly acknowledge the model imperfections.

2. Thanks for implementing the comments regarding prediction uncertainty, and using quantiles instead of fitting Gaussians for inferring uncertainty reduction.

Response:

Thanks for the positive comments.

3. Fig.4: The sampling yields one multivariate distribution. The fact that a several of the marginals of this multivariate distribution do not differ (inferred by the Gelman criterion) does not help, that the chains sample different regions and failed to converge to the same limiting distribution. I agree with reviewer 2 that data paucity should not be the reason for this. A good sampler would sample the prior. Instead of using the Gelman criterion for the marginal, also the multivariate Gelman criterion can be used (Brooks, SP. and Gelman, A. (1998) General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics, 7, 434-455). It is implemented with R function gelman.diag of the coda package.

Response:

Thanks for the suggestion of using "gelman.diag". In this revision, we used R function gelman.diag to calculate Potential Scale Reduction Factor (PSRF). The MCMC chains are univariate. Therefore, we did not calculate multivariate Gelman criterion, which assumes multivariate chains.

We argue that parameter non-convergence could be due to the equifinality issue. We did not mean to say that the non-convergence was totally due to data paucity. However, we argue that the sparse sampling of data could be one of the reasons for the equifinality and non-convergence. Starting from different initial values, MCMC chains ended up with different posterior values, however, they may have similar model-data misfit.

As is also illustrated in our previous response letter, again here we conceptually showed how sparse data result in parameter non-convergence. For example, only three measurements of soil free phosphate were made during 1999. Many detailed dynamics are therefore missing. MCMC sampling may end up with three different posterior models (blue lines versus red line) that have similar model-data misfits. We argue that more data across the year, which better represent seasonal dynamics, would lead to better constrained and converged posteriors. In this case, more POx observations in January, February, August, and September would be extremely helpful to constrain POx associated parameters.



4. Fig 2, parameter at edges: I do not agree with the replies on this issue. Model discrepancy becomes apparent for the phosphorous processes, probably because those processes are constrained by relatively few observations. If the results with such a high parameter are reasonable, the prior density maybe of wrong magnitude. You could try a wider prior, but the parameter might still be at the edge. This edge effect should be at least stated. Better add a little discussion.

Response:

Thanks for pointing out the edge effect issue, which is only for one parameter (k^{plant}_{P}) . In the revised manuscript, we added more discussion on this edge effect issue.

The prior ranges of k^{plant}_{P} are what we believe to be physically reasonable. We argue that model calibration tried to minimize the model-data misfit conditioned that the parameter must be within the prior knowledge. If a parameter is out of the prior ranges, the system is physically unreasonable, even though in that condition the model-data misfit is smaller. Having a calibrated parameter at the edge (high end) of prior implies that plant is highly efficient in P uptake, which is reasonable given that tropical plants are likely highly adapted to low phosphorus environments (Begum et al., 2005; Foshe et al., 1988).

5. Soil heterogeneity: I understand that the model focuses on ecosystem scale. Still, I am interested in your expectations or hypotheses what would change if you could account for soil heterogeneity.

Response:

We agree that soil heterogeneity is an important issue in modeling soil nutrient cycles. By assuming a well-mixed soil, large-scale models completely ignore the heterogeneity. We acknowledge in the manuscript that this is an inevitable flaw, because of large computational demands and a lack of scale-aware parameters and model structures for large-scale models to run fine scale simulations.

The ECA nutrient competition framework here considered for large-scale model integration is readily applicable to fine scale models that explicitly consider soil heterogeneity. Accounting for soil heterogeneity, a fine-scale model simulates nitrogen movement towards root surface due to diffusion and mass flow (Leadley et al., 1997; Nye and Marriot 1969). Plant nutrient uptake creates a nitrogen-depleted zone (named rhizosphere), where strong competition between roots and microbes is expected. The ECA competition framework could be directly applied to the calculation of plant nitrogen uptake at root surface.

Accordingly, we have to change the assumption that roots compete with soil microbes everywhere in the soil. In stead, we hypothesize that in bulk soil nutrient competition only occurs among different microbes because they are ubiquitous in the soil (*e.g.*, nitrifier versus microbial decomposer); while in the rhizosphere competition occurs among plants and microbes (*e.g.*, nitrifier versus microbial decomposer versus roots).

We have added more discussion about the soil heterogeneity issue in the revised manuscript.

6. K^plant_NO3: the statement "may result from random effects of MCMC sampling, but not inferred by the calibration data" Suggests not trusting the sampling too much. But then also take caution in drawing conclusions on importance of parameters.

Response:

In the revised manuscript, we have removed the results about the parameter importance. In the conclusion section, we acknowledge that the calibrated model is the best we can get based on limited data, but it is not the best posterior model we want. Future work on data collection, model calibration, and uncertainty reduction are definitely needed.

Reference:

- Begum, H., and M. Islam. "Role of synthesis and exudation of organic acids in phosphorus nutrition in plants in tropical soils."Biotechnology 4.4 (2005): 333-340.
- Fohse, D., N. Claassen, and A. Jungk. "Phosphorus efficiency of plants: I. External and internal P requirement and P uptake efficiency of different plant species." Plant and Soil (1988): 101-109.
- Leadley, Paul W., James F. Reynolds, and F. S. Chapin III. "A model of nitrogen uptake by Eriophorum vaginatum roots in the field: ecological implications." Ecological Monographs 67.1 (1997): 1-22.
- Nye P., and F. Marriot 1969. A theoretical study of the distribution of substances around roots resulting from simultaneous diffusion and mass flow. Plant and Soil 30: 459–472.

Appendix A. CNP fluxes

The fluxes of carbon, nitrogen, and phosphorus coming from the upstream pool (*i*) to the downstream pool (*j*) due to SOM decomposition are calculated as:

$$F_{C,ij}^{move} = f_{ij} F_{C,i}^{dec} g_i$$
(A1)

$$F_{N,j}^{move} = f_{ij} F_{C,i}^{dec} \min(\frac{1}{CN_i}, \frac{g_i}{CN_j})$$
(A2)

$$F_{P,jj}^{move} = f_{ij} F_{C,i}^{dec} \min(\frac{1}{CP_i}, \frac{g_i}{CP_j})$$
(A3)

where g_i is the percentage of carbon remaining in the soil after decomposition of the *i*th SOM pool (*i.e.*, CUE, with the rest being released as CO₂); f_{ij} is the fraction of SOM leaving the *i*th pool and entering the *j*th pool; and $F_{C,i}^{dec}$ is the first order decay of the *i*th SOM pool. CN and CP are soil C:N and C:P ratios, respectively.

If the upstream-decomposed soil organic nitrogen (phosphorus) is more than enough to sustain the downstream C:N (C:P) ratio, then the excess nitrogen (phosphorus) enters the soil NH_4^+ (PO_x) pool. PO_x represents the sum of PO₄³⁻, HPO₄²⁻, and H₂PO₄⁻ that could be utilized by plants and microorganisms, and adsorbed by mineral surfaces.

$$F_{N,ij}^{mob} = f_{ij} F_{C,i}^{dec} \max(\frac{1}{CN_i} - \frac{g_i}{CN_j}, 0)$$
(A4)

$$F_{P,ij}^{mob} = f_{ij} F_{C,i}^{dec} \max(\frac{1}{CP_i} - \frac{g_i}{CP_j}, 0)$$
(A5)

where $F_{N,ij}^{mob}$ and $F_{P,ij}^{mob}$ are the nitrogen and phosphorus gross mineralization rates. Eqn. A4 - A5 ensure that gross mineralization is not less than zero. In contrast, if nitrogen (phosphorus) is insufficient, soil microbes immobilize free NH₄⁺ and NO₃⁻ (PO_x):

$$F_{N,ij}^{immob,pot} = f_{ij} F_{C,i}^{dec} \max(\frac{g_i}{CN_j} - \frac{1}{CN_i}, 0)$$
(A6)

$$F_{NH4,ij}^{immob,pot} = F_{N,ij}^{immob,pot} \cdot \frac{[NH4]}{[NH4] + [NO3]}$$
(A7)

$$F_{NO3,jj}^{immob,pot} = F_{N,jj}^{immob,pot} \cdot \frac{[NO3]}{[NH4] + [NO3]}$$
(A8)

$$F_{P,ij}^{inmob,pot} = f_{ij} F_{C,i}^{dec} \max(\frac{g_i}{CP_j} - \frac{1}{CP_i}, 0)$$
(A9)

where $F_N^{immob,pot}$, $F_{NH4}^{immob,pot}$, $F_{NO3}^{immob,pot}$, and $F_P^{immob,pot}$ are microbial N, NH₄⁺, NO₃⁻, and PO_x immobilization rates. [NH4] and [NO3] are the free NH₄⁺ and NO₃⁻ pools, respectively. We assume that microbes have no preference for NH₄⁺ or NO₃⁻ (Eqn. A7-A8). If soil nutrients are limited, a limitation factor will be applied to those potential soil decomposition CNP fluxes (Eqn. A1 – A9) to maintain the soil organic matter CNP stoichiometry.

Besides decomposing microbe nutrient immobilization, other potential nutrient uptakes are:

$$F_{NH4}^{nit,pot} = [NH4] \cdot k_{nit} \cdot r_{\theta} \cdot r_{T} \cdot (1 - f^{anox})$$
(A10)

$$F_{NO3}^{den,pol} = \min(f(decomp), f([NO3])) \cdot f^{anox}$$
(A11)

$$F_{p}^{surf,pot} = \frac{VMAX_{p}^{surf,P} \cdot K_{M}^{surf,P}}{(K_{M}^{surf,P} + [PO_{x}])^{2}} \cdot \frac{d[PO_{x}]}{dt} \cdot$$
(A12)

where $F_{NH4}^{nit,pot}$, $F_{NO3}^{den,pot}$, and $F_{P}^{surf,pot}$ are potential rates for NH₄⁺ nitrification, NO₃⁻

denitrification, and mineral surface PO_x adsorption. k_{nit} is the maximum fraction of free NH₄⁺ pool that could be utilized by nitrifiers. The potential nitrification rate is controlled by soil temperature (r_{T}), soil moisture (r_{θ}), and soil oxygen status (*1-f^{anox}*). The potential

denitrificaiton rate $(F_{NO3}^{den,pot})$ is either constrained by substrate availability (f(decomp)) or NO₃⁻ availability (f([NO3])) [Del Grosso et al., 2000], taking into account the soil anaerobic condition (f^{anox}) . $F_p^{surf,pot}$ is derived from the Langmuir adsorption model [Barrow, 1978], where adsorbed P is equal to $VMAX_p^{surf} \cdot \frac{[PO_x]}{K_M^{surf,P} + [PO_x]}$. Taking the time

derivative leads to the adsorption rate [Wang et al., 2010].

Soil NH₄⁺ content is altered by inputs from deposition (F_{NH4}^{dep}) and biological N₂ fixation (F^{BNF}) [*Cleveland et al.*, 1999]:

$$F^{BNF} = 1.8 \cdot \frac{1 - e^{-0.003 \cdot NPP_{annual}}}{365 \cdot 86400}$$
(A13)

where NPP_{annual} is annual net primary production. Controls on biological N₂ fixation are complex and several models have been developed for large-scale land BGC models [*Cleveland et al.*, 1999; *Fisher et al.*, 2010; *Hartwig*, 1998; *Parton et al.*, 1993; *Running et al.*, 1989; *Vitousek and Field*, 1999]. However, the emergent responses predicted across these model structures are inconsistent [*Galloway et al.*, 2004]. Recognizing this important structural uncertainty, we used a simple model where biological N₂ fixation (F^{BNF}) is modeled as a function of annual NPP [*Cleveland et al.*, 1999].

Soil NO3⁻ content is modified by external deposition inputs ($F_{\rm NO3}^{\rm dep}$) and leaching

losses (F_{NO3}^{leach}):

$$F_{NO3}^{leach} = \frac{[NO3]}{W} \cdot Q_{dis}$$
(A14)

where soil nitrate concentration ([NO3]: gN m⁻²) divided by soil water content (W: gH₂O m⁻²) results in the concentration of dissolved nitrate (DIN). The hydrologic discharge $(Q_{dis}: gH_2O m^{-2} s^{-1})$ applied to DIN (gN gH₂O⁻¹) leads to the leaching loss (gN m⁻² s⁻¹).

Soil PO_x content is affected by external inputs from parent material weathering $(F^{weather})$ and leaching losses (F_p^{leach}) . Sorbed P (P_S) could be further strongly occluded and become unavailable for plant and microbial uptake. Parent material stock can be increased by atmospheric dust deposition (F_p^{dep}) [*Mahowald et al.*, 2008]:

$$F^{\text{weather}} = [P_p] \cdot k_{\text{weather}}$$
(A15)

$$F_{P}^{leach} = \frac{[PO_{x}]}{W} \cdot \mathcal{Q}_{dis}$$
(A16)

$$F_{p}^{occl} = [P_{S}] \cdot k_{occl} \tag{A17}$$

where parent material weathering ($F^{weather}$) is calculated using a weather rate ($k_{weather}$) and parent material P content ([P_P]). PO_x leaching loss is modeled with a similar approach to nitrate leaching (Eqn. A16). Phosphorus occlusion rate is modeled as the product of a

constant rate (k_{occl}) and the sorbed P content $([P_S])$.

References:

- Barrow, N. J. (1978), The description of phosphate adsorption curves, *Journal of Soil Science*, 29(4), 447-462.
- Cleveland, C. C., A. R. Townsend, D. S. Schimel, H. Fisher, R. W. Howarth, L. O. Hedin, S. S. Perakis, E. F. Latty, J. C. Von Fischer, and A. Elseroad (1999), Global patterns of terrestrial biological nitrogen (N2) fixation in natural ecosystems, *Global Biogeochemical Cycles*, 13(2), 623-645.
- Del Grosso, S. J., W. J. Parton, A. R. Mosier, D. S. Ojima, A. E. Kulmala, and S. Phongpan (2000), General model for N2O and N2 gas emissions from soils due to dentrification, *Global Biogeochemical Cycles*, 14(4), 1045-1060.
- Fisher, J. B., S. Sitch, Y. Malhi, R. A. Fisher, C. Huntingford, and S. Y. Tan (2010), Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation, *Global Biogeochemical Cycles*, 24(1).

- Galloway, J. N., F. J. Dentener, D. G. Capone, E. W. Boyer, R. W. Howarth, S. P. Seitzinger, G. P. Asner, C. C. Cleveland, P. A. Green, and E. A. Holland (2004), Nitrogen cycles: past, present, and future, *Biogeochemistry*, 70(2), 153-226.
- Hartwig, U. A. (1998), The regulation of symbiotic N< sub> 2</sub> fixation: a conceptual model of N feedback from the ecosystem to the gene expression level, *Perspectives in Plant Ecology, Evolution and Systematics*, 1(1), 92-120.
- Mahowald, N., T. D. Jickells, A. R. Baker, P. Artaxo, C. R. Benitez-Nelson, G. Bergametti, T. C. Bond, Y. Chen, D. D. Cohen, and B. Herut (2008), Global distribution of atmospheric phosphorus sources, concentrations and deposition rates, and anthropogenic impacts, *Global Biogeochemical Cycles*, 22(4).
- Parton, W. J., J. M. O. Scurlock, D. S. Ojima, T. G. Gilmanov, R. J. Scholes, D. S. Schimel, T. Kirchner, J. C. Menaut, T. Seastedt, and E. Garcia Moya (1993), Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide, *Global biogeochemical cycles*, 7(4), 785-809.
- Running, S. W., R. R. Nemani, D. L. Peterson, L. E. Band, D. F. Potts, L. L. Pierce, and M. A. Spanner (1989), Mapping regional forest evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation, *Ecology*, 1090-1101.
- Vitousek, P. M., and C. B. Field (1999), Ecosystem constraints to symbiotic nitrogen fixers: a simple model and its implications, *Biogeochemistry*, 46(1-3), 179-202.
- Wang, Y. P., R. M. Law, and B. Pak (2010), A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere, *Biogeosciences*, 7(7), 2261-2282.

Appendix B. Derivation of VMAX

The enzyme substrate reaction is: $S + E \xleftarrow{k_1^+}{k_1^-} C \xrightarrow{k} P + E$, where the enzyme

(*E*) and substrate (*S*) reaction is reversible and forms complex (*C*). The irreversible reaction releases product (*P*) and liberates enzyme (*E*). At steady state, the formation rate of the enzyme substrate complex is equal to the consumption rate:

$$k_1^+[S][E] = k_1^-[C] + k[C]$$
(B1)

To simply the equation, we define an affinity parameter:

$$K_{M} = \frac{k_{1}^{-} + k}{k_{1}^{+}} = \frac{[S] \cdot [E]}{[C]}$$
(B2)

By definition, the total enzymes $[E_{tot}]$ in the system is the sum of free enzymes [E] and enzymes that are bound with the substrate [C]:

$$[E_{tot}] = [E] + [C] \tag{B3}$$

Substituting Eqn. (B3) into (B2), we have:

$$K_{M} = \frac{[S] \cdot ([E_{tot}] - [C])}{[C]}$$
(B4)

Collecting terms containing [C], we have:

$$[C] \cdot (K_M + [S]) = [E_{tot}] \cdot [S]$$
(B5)

The production rate is:

$$\frac{d[P]}{dt} = k \cdot [C] \tag{B6}$$

Substituting Eqn. (B5) into (B6), we have:

$$\frac{d[P]}{dt} = k \cdot [E_{tot}] \cdot \frac{[S]}{K_M + [S]}$$
(B7)

Comparing Eqn. (B7) with the classic Michaelis-Menten equation, it is clear that the definition of maximum production rate is the product of the reaction rate and enzyme abundance in the system:

$$VMAX = k \cdot [E_{i\alpha}] \tag{B8}$$

Figure A1. MCMC chain. Blue line represents the MCMC samples that are used to infer our model posterior parameters. Two other replicated MCMC calibrations (with different random number seeds) were conducted (yellow and red lines), in order to check the convergence of MCMC calibration.





Berkeley Lab 12/9/15 10:00 PM Comment [1]: Figure is updated **Figure A3**. Posterior model parameters (blue bars) fitted to Gaussian distribution (red line).

