

## ***Interactive comment on “Carbon storage in phosphorus limited grasslands may decline in response to elevated nitrogen deposition: a long-term field manipulation and modelling study” by Christopher R. Taylor et al.***

**Anonymous Referee #3**

Received and published: 6 December 2020

The authors presented a modelling application of the N14CP model to investigate the role of N deposition on the soil C storage of P limited grasslands. The N14CP model seems to be a simple but heavily calibrated model, but it is not adequately described in the paper for the readers to fully understand the long discussion about the pattern of different model outputs.

However, the main deficiency that I find is the model performance against measurements in figure 2. First of all, I don't think the 1-to-1 point plot is the best way to display the results, since each point is representing a different experiment and to me it is more

C1

interesting to see the different model performance of varying scenarios rather than looking at a overall  $r^2$  of eight very different scenarios. One example as the authors have already noticed, is that AGB carbon and soil C are noticeably overestimated in acidic grassland but soil N is not, and more surprisingly, total soil P is underestimated. This pattern really indicates that the model is not capturing the SOM stoichiometry, and it actually worries me about the main focus of the paper is on effects of N and P on soil C storage. Secondly, it is unclear to me if all the eight experiments are calibrated or only the two unfertilized ones are calibrated. Also, the initial soil pool sizes are not clear to me either. I find it really difficult to understand how to spin up the model for 10000 years and compare to the present day soil measurement. From figure 3 it seems that the model is still far from equilibrium in both ABG C and soil C, particularly in the acidic grassland. It actually confuses me about the poor soil C correlation between modelled and measured soil C in the acidic soil. Why do you choose to spin up the model for 10000 years, and how does the spin up time affect your results?

A final comment, the discussion need to focus much less on the speculation of model outputs, but include some discussion about the possible caveats of model or study design and uncertainties caused by these limitations.

---

Interactive comment on Biogeosciences Discuss., <https://doi.org/10.5194/bg-2020-392>, 2020.

C2