We would like to thank both reviewers for taking the time to review our work and for providing constructive comments to improve the manuscript. Please find our responses to reviewer comments below in italics, with the original comments in regular text.

Note that our references to line numbers refer to those of the tracked changes document.

Reviewer 1

'Line 276: I'm not sure why the flexible stoichiometry here is described as representing changes in species only rather than ecosystem level changes which can be either plastic changes or shifts in species compositions, since the model (as most other models) cannot differentiate between the two?'

The text here has been amended to reflect that the changes in stoichiometry do not just occur at the species level and extend to the ecosystem scale as the reviewer points out. The updated paragraph is on lines 201 – 206:

'Stoichiometry of coarse tissue is constant but the fine tissue of each plant functional type has two stoichiometric end members. This allows the model to represent transitions from N-poor to N-rich plant communities or an enrichment of the fine tissues within plants (or a combination of both) [Davies et al., 2016b], dependent on available N. '

'Line 290: Is BNF regulated by anthropogenic deposition only or more generally by soil N content? This might help the reader understand the model-data discrepancy under P addition in the acidic grassland.'

BNF in the model is downregulated by N deposition only, not soil N content more generally as it is assumed that atmospherically deposited N is readily available to N fixers, whereas soil N is not necessarily available. We have clarified this in the methods section lines 212 – 217:

'The initial rate of N fixation is based on literature values for a given plant functional type and is downregulated by anthropogenic N deposition, but not soil N content more generally, as it is assumed that atmospherically deposited N is readily available to N-fixers. Nitrogen fixation in the model is also related to P availability. The degree to which P availability limits this maximum rate of fixation is determined by a constant; K_{Nfix} *[Davies et al. 2016b]'*

'Line 301: Why coarse litter and not all litter?'

Within the model, coarse litter forms the litter pool and decomposes on the soil surface to contribute N and P to the bioavailable pool. Fine litter is not included in this pool and due to its more rapid turnover time, it is incorporated directly into the SOM pool, where it can then decompose to contribute nutrients to the bioavailable N and P pool. A sentence to clarify this has been added on lines 228 – 231:

'Plant available N is derived from biological fixation, the decomposition of coarse litter and SOM, atmospheric deposition and direct N application. Fine plant litter enters the SOM pool directly due to its rapid rate of turnover whereas coarse litter contributes N and P through decomposition and does not join the SOM pool.'

'Line 606: Onwards check the precision of your reported values, there are some inconsistencies'

Thank you for pointing this out, we have amended the precision of the reported $P_{CleaveMax}$ values to two decimal places so they are consistent with the reporting of the R^2 values later in the text. These values are on lines 416 and 417. The reported precision for the percentage difference between empirical and modelled data and their standard errors remain to one decimal place, which we felt was more appropriate for the magnitude of the data.

Reviewer 2

'The manuscript is generally very lengthy'

We agree with the reviewer that the manuscript is fairly long, though its length is comparable to other Biogeosciences manuscripts, particularly those detailing modelling results and methodology. Through our revisions we have cut much content, particularly from the discussion, though this has largely been replaced by additional text in the methodology to more clearly describe the model as requested by the reviewers, to ensure the reader can effectively interpret our results.

'Abstract sentences 1 and 2 are contradicting. Suggest to delete the first sentence'

The first sentence of the abstract has been revised so that it is no longer contradictory and to better integrate the theme of P limitation in C, N and P cycling. The new sentence on lines 29 – 31 now reads:

'Phosphorus (P) limited ecosystems are widespread, yet there is limited understanding of how these ecosystems may respond to anthropogenic deposition of nitrogen (N), and the interconnected effects on the biogeochemical cycling of carbon (C), N and P.'

'Above ground biomass carbon, soil organic carbon, and total N were newly measured for this study. I was confused that the methods were not described and the results not shown in the main text, and only later found a description of the methods in the supplement. However, soil analyses are not defined appropriately. They are at once way too lengthy and on the other hand lacking any citation. Such basic measurements as above ground biomass, soil organic carbon and total N should be done according to standard protocols. This should be cited appropriately. I would suggest to move the methods section (in a more concise) form to the main text. Also, I would suggest to start the results section with these observational results (leaving out the modelling in a first table or figure). This will help the reader get to know these sites and how they respond to the different treatment, which are then the basis for interpreting the model.'

Thank you for these suggestions, we have implemented most of the recommended revisions. We have cut unnecessary content from the description of the empirical methods, and provided relevant citations for standard protocols, and detailed if and how these were adapted. In addition, and as proposed, we have put all empirical data into a table separate from any modelled data, to aid the reader in interpreting empirical responses to nutrient treatments. This also necessitated the removal of some redundant information from later data tables.

However, we believe that keeping these empirical methods and data in the supplementary material is preferable, given our intended focus of the manuscript. Firstly, and as identified by the reviewer, the manuscript is already fairly long, hence we would be hesitant to add detail to the methods, results and potentially discussion section when all the information is accessible in the supplementary material (in a now more succinct format). We have added a couple of sentences in the methods (lines 143 - 144) and results (lines 403 – 404) to direct readers who are particularly interested in the empirical data.

We have discussed including the empirical data in the main manuscript and we feel that its inclusion may detract from what is at its core a modelling paper, and perhaps lead to interpretations that we did not intend. For instance, the narrative has been revised in the current version to clarify that we were not attempting to replicate the empirical data and instead used it to inform model development and interpret our model outputs. Adding the empirical data back into the main text may be counter-productive to this aim. Instead, we opted to describe the empirical results in context of the modelling in a more qualitative way and provide the empirical data in the supplementary information for interested readers.

'Lines 233 – 234: P deposition is assumed to be negligible in this model. Actually, more and more evidence is showing that P deposition is just as important as rock weathering for P inputs to terrestrial ecosystems (see e.g. Aciego et al. 2017). This should be considered for further model development in the future'

Thank you for highlighting this interesting study. We assume that P deposition is negligible here as we cannot account for losses of P that also occur through landscape redistribution, which is an issue discussed in a meta-analysis some co-authors were involved with (Tipping et al. 2014). We do agree though, that in combination with improved empirical data on P deposition and redistribution, this would make an interesting modelling study for the future. We have elaborated on our assumption in the methods lines 242 – 246:

'In principle, P can be added in small quantities by atmospheric deposition [Ridame and Guieu, 2002] but for the purpose of this work, P deposition is set to zero in the model. While the contribution of P through atmospheric deposition is increasingly realised [Aciego et al. 2017], we cannot account for the losses of P that may also occur through landscape redistribution [Tipping et al. 2014].'

'To improve readability, I suggest to reduce abbreviations. Specifically, no need to abbreviate PFT'

Plant functional type is no longer abbreviated throughout the text.

'Line 349, replace comma with a period'

Amended.

'Line 628: Authors state organic P release from SOM and immobilization are poorly represented in models and that they encourage further study to quantify these processes. I agree with these statements; however, from reading the manuscript I wondered if the authors were aware of the state of the art P flux measurements since the results are not discussed in light of measurement data? Several studies have actually measured organic P mineralization and microbial immobilization with radioisotopes, and would be relevant for interpreting the modelling results presented here. For example, Bünemann et al. 2012 looked at mineralization fluxes in grasslands under NPK treatments and Schneider et al. 2017 calculated organic P fluxes in calcareous soil'

Thank you for these suggestions. We were not aware of these measurements and have incorporated some of their findings into the discussion. In particular, we have added a brief paragraph to contextualise some of our modelling results with the findings of the recommended papers (lines 578 – 585):

'In addition to physico-chemical processes reducing P availability, in P-limited grassland soils, microbial processes may be dominant drivers of ecosystem P fluxes [Bünemann et al. 2012]. For instance, while mineralisation of organic P may increase inorganic P in solution [Schneider et al. 2017], this can be rapidly and almost completely immobilised by microbes, particularly when soil P availability is low [Bünemann et al. 2012]. As the model lacks a mechanism for increasing access to secondary mineral P forms comparable to organic P-cleaving, and microbial P immobilisation is incompletely represented for P-limited conditions, it is possible that the uptake of organic P by the acidic grassland in the model is exaggerated.'

'General remark on over-selling: authors should be careful not to over interpret their results stemming from modelling two grassland soils, especially given the limitations as discussed.'

We have amended our interpretation of the results so as not to extend them beyond what is reasonable (see related comments below for specifics).

'Line 46: Wardlow is not a globally important C sink. Please delete this sentence, since it is not appropriate to extrapolate from two sites simulated here onto a global level'

The concluding sentence of the abstract (lines 46 – 48) has been revised so it no longer implies Wardlow is a globally important carbon sink. This was not our initial intention but the wording was problematic as the reviewer identifies. The sentence now reads:

'We conclude that grasslands differing in their access to organic P may respond to N deposition in contrasting ways and where access is limited, soil organic carbon stocks could decline.'

'Line 641: Same here. It is inappropriate to generalize from the two grassland sites about ecosystems in general all over the world.'

We have changed lines 664 – 667 so they do not generalise to ecosystems around the globe and instead make explicit reference to ecosystems such as Wardlow:

'This suggests that in P-limited limestone grasslands such as at Wardlow, where access to organic P forms may be comparatively limited, N deposition may worsen pre-existing P limitation and reduce ecosystem C stocks [Goll et al. 2012, Li et al. 2018].'

‘Line 648: I don’t consider N14CP to be “one of the first” CNP models. Many other models come to mind, some of which much older or much more developed: JSBACH, CABLE-CNP, CLM-CNP, ORCHIDEE-CNP, QUINCY, ForSAFE’

This is a fair comment and we have amended the text accordingly on lines 673 – 675. We highlight that N14CP is one of few but not one of the first:

‘While N14CP is a fairly simple ecosystem model by design, it is one of few models to integrate the C, N and P cycles for semi-natural ecosystems and has been extensively tested against empirical NPP and soil C, N and P data.’