Tanner responses to comments by W Wieder.

1) **Comment.**

“Was soil mineral mass measured in each plot, in each treatment, or in a single pit (like bulk density)”

Response

Soil mineral matter was calculated for each plot and soil depth from soil carbon concentration (mineral matter is total soil mass minus twice soil carbon content). Bulk density was measured for every plot for 0-5 cm depth; below 5 cm we used the bulk density from one soil pit (lines 146-151 in the manuscript).

2) **Comment.**

“More broadly, the emphasis placed on soil mineral mass to extrapolate findings seems somewhat surprising,”

Response

The emphasis on expressing soil carbon per mineral mass is to deal with the (general) problem that as soil organic matter changes the bulk density changes, so sampling to the same depth will not be comparing like with like. This is well known problem - Powlson et al 2011 say “The principle is that an equal mass of organic-matter-free mineral soil should be sampled between the treatments or times being compared.” For this reason in our study in Panama we expressed carbon relative to an unchanging mineral mass. It is also an easy calculation to make and can often be made retrospectively on published data. It was not done to get round a problem of non-significant results.

3) **Comment.**

“If data are available to make an extrapolation of Fig 3 with depth on the X axis it would be much more valuable for studies trying to quantify or model changes in soil C stocks, as information about mineral mass is typically lacking or not considered.”

Response

Fig 4 shows the cumulative (with depth) mineral matter and soil depth in the control plots, down to about 93 cm. An e mail exchange with the referee clarified that he wanted a second axis In Fig 3 showing the soil depth in the control plots – we have done this. We disagree with the comment that “information about mineral matter is typically lacking”, because if samples have data on soil carbon per dry soil mass, then the mineral matter is easily calculated (as total mass minus twice soil carbon - there will be a small error because soil organic matter is not exactly twice soil carbon, but the effect will be trivial.)

4) **Comment.**

“I recall publications from some of the temperate DIRT plots (e.g., Lajtha references in the paper) showed changes in different soil C fractions. I assume similar data are not available for this study, but I wonder if consideration of C stabilization mechanisms and soil mineralogical conditions could

help explain some of the differences between temperate and tropical sites. Is it worth a brief discussion on this point (e.g. expanding / developing the paragraph that begins on line 202)?"

Response

Other researchers are working on this in the experiment. As we present no data on carbon fractions in this paper we think it best to leave discussion of that subsequent manuscripts.

5) **Comment.**

"The authors (justifiably) seem keen on their soil P results, which are interesting and relevant (line 262). Is it possible to extrapolate findings for P, similar to the soil C figure 3, making this a multi-panel figure?"

Response

It is not sensible to express cumulative Mehlich P per cumulative mineral mater (in an analogous way to cumulative carbon per cumulative mineral matter in Fig. 3) because a substantial (but unknown) amount of Mehlich P comes from organic matter. Soil matter is either organic or mineral and we plot one against the other in Fig. 3; Mehlich P is different - it comes from both mineral and organic matter.

6) **Comment.**

"The discussion starts off with the introduction of new results. I appreciate the authors wanting to focus readers' attention on these findings, but feel like results (Figs 3 & 4) are best introduced in the results, not discussion section of a paper"

Response

We disagree. The 'results' are concentrations of carbon per mass. We then use those results to calculate concentrations of carbon per mineral matter.

7) **Comment.**

"Finally, calling out the small plots from the Costa Rican study seems a bit unjustified in a single paragraph subsection of the discussion. Granted the authors make a good point about the appropriate size of experimental plots, but I think Leff and co-authors (2012, cited in the paper) acknowledge the limitation of their small plots. If the authors want this section to remain they should more broadly discuss other litter manipulation studies, not just the Costa Rican site."

Response

We are not making any personal points here, but we do think that there is a real issue about the size of experimental plots affecting the qualitative patterns of results. Specifically, small (3 x 3 m) litter removal and addition plots might be local cold spots and hot spots that will affect the responses. The pattern of results from small plots might be the OPPOSITE of those from large plots. For example, small litter addition plots might cause extra root growth into local patches of soil with extra nutrients, but large litter addition plots (45 x 45 m) might cause reduced root growth because the whole tree is receiving extra nutrients and 'can afford' to reduce root growth and put more into shoot growth, in other words, a completely opposite pattern of results caused by differences in experimental design. We simply want to point out that the design of these experiments might well affect the pattern of results. If there were lots of experiments like this we could look for patterns, but there aren't many.

To address the reviewer's comment, we have changed the last line to "small hot and cold spots may not represent what would happen in plots on the scale of the large trees - as pointed out by Leff et al 2012."

8) Technical corrections:
Comment. Introduction: specific values for C pools, turnover times, and fractions seem unnecessarily detailed (lines 33, 36). More broadly the introductions reads a bit like a bullet point of disconnected ideas. This is a stylistic concern, not a scientific one.

Response. As Wieder says this is stylist – we think this is clear and informative

Comment. Throughout, check that abbreviations are defined before they are used in the text (eg. LR and LA line 55, GFP line 251).

Response. We have changed all 'LR' to 'litter removal' and all 'LA' to 'litter addition'. We have reworded the text so that GFP is no longer used.

Comment. Line 66-68, This is unclear P mineralization (0-2 cm) in LR plots met 20% of NPP needs, or the decline in P mineralization would have met this demand?

Changed to "mineralization of organic phosphorus (P) (inferred from the decrease in the concentration of organic P) in the top 2 cm of soil during three years of litter removal was calculated to be sufficient to supply 20% of the P needed to sustain forest growth"

Comment. Line 76. This study looked at net nitrification and should be Wieder et al. 2013 (i before e).

Response. Added 'net' and corrected spelling of Wieder.

Comment. Line 89. Awkward. Forest productivity isn't mitigated, but increases in terrestrial C storage can mitigate atmospheric CO2 accumulation.

Response. Changed to "can thus be considered as partial mitigation of atmospheric CO_2 accumulation through increased forest productivity"

Comment. Line 210. Awkward, maybe insert 'a' here: In a deciduous forest in MA. . .

Response. Changed to "In a deciduous forest in"

Comment. Line 307. What is meant by 'polluted' sites? Is this sites receiving large amounts of N, P or micronutrient deposition (is the later actually a real a thing)? Is this just to say that litter manipulations aren't identical to CO2 enrichment alone, because they also serve as nutrient manipulations that modify ecosystem dynamics?

Response. This site is not receiving large amounts of N or P (though N input is increasing Hietz et al 2011 Science 334, 664). Our comparisons are based on N & P inputs in polluted sites in USA and Europe. We have added ‘temperate’.
We don’t mention micronutrients in the Conclusions – so we ignore that part of the comment.