This study uses fine root data from a pine plantation in China to illustrate the relationship between sampling effort and uncertainty in the resulting estimates.

My first impression was that the paper would be useful as a tutorial to researchers interested in applying uncertainty analysis in support of experimental design. I am now concerned that the data set selected to illustrate the approach is not suited to it—the sampling design of the plantation geometry would be fine for detecting treatment effects of change over time, but not for characterizing roots per unit area (as is needed for carbon budgets). I also have doubts about the assumptions of independence in the statistical calculations. I would recommend that the root data be published for the purpose for which they were collected, and that the statistical approach presented here be applied to data for which it is more appropriate. Or the statistical approach could be offered as a means to calculated minimum detectable change, but not to anything relevant to carbon budgeting (IPCC).

The case study focuses on fine root surface area density, which surprised me a little; isn't there more interest in biomass? I understand that it's the approach that's important, and it could be applied to any problem with this structure. But if your basis for aiming for 10% uncertainty is something from the IPCC, I'm guessing that it would be about carbon, not root acquisition of soil resources.

I am concerned about the geometry of soil sampling in a plantation setting, and what relationship the results you show have to the average root biomass or RAD of the area. If roots were randomly distributed with respect to tree stems, this would not be an issue, but they are not. If, as is the case for many studies in natural forests, the root samples were randomly distributed with respect to the stems (or systematically distributed with the stems being random) then the sample would be representative and unbiased. I was not able to understand the sampling design, and this is important. You describe a trench that extends from a tree halfway to the next row of trees. It is not clear whether the entire length of this trench was sampled. If it was, which seems unlikely, then there is no bias in sampling from one row to the next. But there is a huge bias in sampling from one tree to the next, within a row. There would be much less root mass if the trench were placed between trees instead of starting from a tree. If I am understanding this correctly, your paper does a great job of addressing the precision in measurements that are fatally biased. You could be more in error due to the placement of your samples than to your sampling intensity. If this is the case, it would be better to illustrate your statistical ideas using an unbiased sample. There are plenty of unbiased (random) data sets to choose from.

10% uncertainty seems unreasonable as a target—and your results show that. Has there already been a review paper describing confidence in detecting root carbon (or RAD)? I wrote one on forest floor carbon (Yanai et al. 2003) (the other Yanai et al. 2003, not the more highly cited one) and nobody should expect to have 10% confidence, or detect a 10% change with confidence, in forest floor biomass. I contributed to a paper on detectable differences in root biomass (Fahey et al. 2017), but we didn't attempt to review the statistical power of other

studies. If you can't cite such a study, add it to a list of papers that you or future students could write.

I am not an expert reviewer of the statistical approach; I might have missed some problems. I have one nagging question, which is whether the observations should really be considered as independent. I know that depth increments from a single soil pit are not as independent as the same depth samples each collected from a different soil pit. It would be inefficient to collect them independently, I'm not saying anyone should do that, but treating them correctly in terms of statistical power is important. I have been told to use a repeated measures approach in these cases, for the repetition in space. I'm also vaguely aware of nesting. It's possible that the statistics you present do not depend on the independence of the observations. Perhaps you could add some clarification of this point.

A practical question is whether the effort required to describe depth distributions is warranted, if in fact the usual goal is to estimate carbon stocks. Pooling samples across depth increments, when root density is low, would save time and thus cost. This would eliminate the danger of incorrectly treating depth increments from the same core or pit as independent!

311 My hackles rise at the suggestion that an exponential relationship has an inflection point. I have often seen researchers point to a value above which they claim the relationship increases more steeply. This is equally true at every point in an exponential curve. Perhaps this is not what you meant; I certainly hope not.

Specific comments

The paper is very well written. I will attempt to scan and return a pdf with my scribble, because there were very few points where I stumbled and the careful writer among you (I bet it's one author, or else an editor) will be interested in them.

I would not use the term "segment" for a soil depth increment. The term is often applied to roots but never to soils. Does it apply only to things that are 1-dimensional?

119 these two questions are not mutually exclusive. Eliminate the "or".

166-171 This is background material that sounds like it belongs in the Introduction. The methods are usually dry and rarely read.

179-188 These are results.

It's conventional in reporting observations with depth in the soil (or sediment cores, etc.) to put depth on the y axis, with zero at the top, and the response variable on the x axis.

"i.i.d" If I'm in your target audience, you should perhaps not use this abbreviation, as I had to look it up. Or perhaps you can assume that your equations will be rarely read.

When I looked up “i.i.d.” I recognized this as a potential problem—the samples collected from a single tree are not as independent as those from different trees, as discussed above. The fact that you chose 2 trees from a row is also troubling. This should be reflected in your statistical model, but I don’t know how and I doubt it makes much difference.

239 “10, 10, 30, 50” would be easier to understand and have the same effect.

249 This is worrying, how could it be irrelevant which samples came from which tree? This relates to my question of whether you are treating them (unfairly) as independent.

I’m sorry, I got this message, saw the title of the paper, and I want to add another review comment.

The title of the paper should not emphasize “fine root trait parameters.” What “parameters”? I think even “traits” is a bit much. I believe that the results would be the same for any root trait. If the authors are truly interested in different traits, they could easily make the calculation and see if they get different results. It would be surprising if they differed much; most of the variation is in the amount of roots, not the variation in the relationships among the traits. But they report just one.

"Sample sizes required to achieve desired uncertainties in the amounts of fine roots in forests” or how about “confidence” instead of “desired uncertainties”?

Except now I doubt whether they have done that. Sigh.