Response to comments on "Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks" by Réjou-Méchain et al.

Dear Editor,

Please find the revision of our manuscript. We really appreciated the careful assessment of our manuscript by reviewer 1. We have addressed all his/her minor comments. Below we provide answers to all points raised in his/her review.

We hope that the corrected manuscript, in which all changes are highlighted in red, will now be suitable for publication in *Biogeosciences*.

Thank you for your time and consideration.

Sincerely yours,

On behalf of the authors, Maxime Réjou-Méchain

This paper is now more carefully written and is a very useful addition to the literature about the accuracy with which we can learn about landscape scale biomass, especially from remote sensing, using information gathered at sample plots.

Response: Thanks for this positive comment.

The initial remarks in my first review still hold good, and I repeat them: This paper poses and attempts to answer several important questions that are significant in the context of current efforts to infer large scale biomass maps from remote sensing and to make more general inferences on landscape scale biomass from a set of sample plots. **In fact, the paper is not really about remote sensing per se, but about how accurately one can extrapolate measurements at one scale to a larger scale.** In general, it illustrates that the sampling error when small plots are used to represent the average biomass of a larger area can lead to significant errors in the regression relation between the two. This is of special importance when training remote sensing data with plots that are significantly smaller than the resolution of the instrument.

Response: We completely agree that our results are more general than the remote sensing context. We focus specifically on the relevance to remote sensing because of the very topical nature of these particular errors given at these particular scales for planned satellite missions for remote sensing of biomass.

The authors have adequately addressed the points I made in my first review, but there are still some matters and careless use of language that need addressing before publication, as set out below.

129; deforestation should surely also be mentioned
Response: Done

131 REDD+ was set up later than 2008, so this ref is not right
Response: This reference has been replaced by Agrawal et al., 2011

138 delete "still"
Response: Done

139 remove "to" before stratify
Response: Done

140: stratum
Response: Done

141-146: for landscape scale fluxes, if the forest is stratified and we mean the mean right for each stratum, we'll end up with the right overall mean biomass or flux. Also, what has forest classification got to do with degradation or regrowth: these are different things? What is meant buy "insensitive"?

Response: In theory we agree that if each stratum has a precise mean estimate, the overall mean biomass change will be captured correctly. However, the error in the mean increases as variability within a stratum increases. This seems intuitively obvious, and we do not want to belabor this issue in the manuscript, as it is not a central point. We changed the two sentences on forest classification to a single sentence: "This approach inherently overlooks extensive spatial variation in carbon density within strata, including variation related to forest degradation and regrowth, both crucial components of forest carbon fluxes (Harris et al., 2012; Lewis et al., 2009).".

146: what is meant by a continuous measure? Isn't carbon density always continuous?
Response: Rephrased: "Thus, recent studies have moved from classification approaches involving a discrete number of forest types toward approaches encompassing continuous spatial variation in forest structure and carbon density"

152: "mission"; particular
Response: Done

155 & 300: the intrinsic resolution of BIOMASS is around 50-60 m; the scale at which it will map biomass is ~200 m
Response: We rewrote this sentence as: "The products from this instrument will have a relatively coarse resolution (200 m) and will rely on ground data to train their inversion models and to evaluate the results."

167: remove potential
Response: Done

168: often not ellipsoidal; radar is nearly square
Response: Done

170 slant range correction is always applied to radar products as part of the pre-processing for applications such as forestry, so this comment is wrong
Response: Sorry, we indeed used the wrong term. This sentence has been modified as: "Side-looking radar observation is a typical example of such spatial mismatch with field-based tree stem measurements (Villard and Le Toan, in press)"

175: remove potential
Response: Done

179; what is meant by a large footprint; perhaps delete this?
Response: We added "(≥0.5 ha)", the minimum resolution illustrated in figure 6.

182 what is a "standard" plot size?
Response: We replaced "standard plot sizes" with "the most commonly used plot sizes."

181-184; isn't (1) the same as (2)?
Response: No. Question 1 is about spatial variability. Spatial variability can be spatially structured (aggregated, etc.), but it need not be, it could just be random. Question 2 is about spatial structure, that is whether there is a spatial pattern distinguishable from complete spatial randomness, which can be detected by autocorrelation analyses, wavelet analyses, etc. We recognize that other readers unfamiliar with the term "spatial structure" might have the same question as the reviewer; thus we rewrote question 2 as: "Does local AGBD variability exhibit significant spatial structure (e.g., aggregation), and if so, what is that structure (strength, spatial scales)?".

186: remove the comma
Response: Done

190: what is meant by "standardized"? Is it necessary to say this, since below it says what was actually done?
Response: Standardized has been removed

201-207: I agree with the other reviewer; I see no reason to talk about topographic variability when what is actually used is elevation range. These are two different things and it is misleading to conflate them. It is not surprising that elevation changes can affect biomass, and this seems the more interesting variable. This affects several parts of text, including, e.g, the caption for Fig. 3, Fig.5, section round 348, 355, etc.

Response: The elevation ranges within all of these plots are quite small. Previous site-specific studies show that biomass variation within plots is strongly associated with topographically defined habitat types such as valleys, ridges, and slopes – not generally with elevation per se. That is, it is not absolute elevation that drives biomass variation at these spatial scales, but elevation relative to surroundings (valley vs. ridge). Across our plots, elevation ranges are correlated with this kind of topographical variability; that is, the plots with higher ranges have more pronounced valleys vs. ridges, while those with lower ranges have much shorter ridges and shallower valleys. Consistent with this, as shown in Fig. S1, elevation range and elevation variability are highly correlated among plots. We thus decided to use elevation range as a proxy for topographic variability and we prefer to keep the term "topographic variability" as this term is highly familiar in forest related works and speaks to the true drivers. Speaking exclusively in terms of elevation would bring to mind the sorts of changes

in biomass that one sees along large altitudinal gradients on mountains, which is not an important mechanism within the scale of our plots. The fact that in practice our analyses concern elevation range is explicitly stated everywhere in the manuscript.

204: delete publicly
Response: Done

218-220; relative to what? This is not really correct as written. The CV may be large but the sampling error may be small (all to do with that word relative)
Response: Relative to the mean AGBD. We could have compared variability in terms of the SD; that would be absolute variability. We chose to compare variability in terms of the CV, which gets at relative variability. We have adjusted the wording to clarify: A higher CV value indicates a higher relative spatial variability of AGBD (relative to the mean), and therefore greater random sampling error relative to the mean estimate when small subplots are used as samples to represent the full plot area."

226: rm "qualitatively with", and put "graphically"
Response: Done

229-234: rm "in the … autocorrelation" and re-order, to put the equation for CV(s) first, before the descriptive text (which may not be needed). They may be true, but where is the justification for the statements about the effects of correlation? The paper lacks is a clear quantitative relation between correlation and variance, and this weakens several parts of the paper, e.g 436-440
Response: We have consulted with several colleagues and they are of the opinion that the statements are correct and clear as written. The qualitative relationship between spatial correlation and the slope is intuitively obvious, and this qualitative relationship is what is relevant for our questions and interpretation. We consider a quantitative treatment to be unnecessary given the aims of this paper. In our view, this paper has enough analyses and figures in it already without adding additional material, especially technical material.

244 remove "a class"
Response: We disagree; d is a class of distance.

247 rm "more usefully"
Response: Done

251: rm "The"
Response: Done

260, 264, 374 (maybe elsewhere): what does "complete .. randomness" mean? Independent, independent identically distributed?
Response: Yes this means spatially independent. This is now defined the first time that this term is used. We keep this term as it is the appropriate and commonly used technical term (>3000 occurrences in Google scholar for "complete spatial randomness").

267 why "repeated"
Response: removed

275, 276: I think it should be pointed out here that this is NOT just a statement about remote sensing data but any attempt to extrapolate sample measurements to a larger scale. In fact, given the idealised assumptions about perfect EO measurements, it is quite unrealistic. This relates to 293; it's not the measured AGBD but the exact AGBD.
Response: We added the following sentence: "Note that this approach more generally attempts to assess the errors generated when sample measurements are extrapolated to a larger scale."
We also removed the word "measured"

304 , rm "regression"
Response: Done

343 What does "local" refer to?
Response: We believe that the term "Local" is necessary to illustrate that we work at a relatively small spatial scale (<500 m).

348: why italic rho?
Response: Because this is a Greek word.

349; what does considerable mean? quantify this.
Response: This is quantified in the previous sentence by the Spearman's rho.

357-366: How do the authors square what is said in the text about low measured autocorrelations, e.g. in Section 3.2? There is a contradiction here that needs explaining.
Response: We believe that there is no contradiction, as illustrated in the sentence from section 3.2:
"Though these increases with distance were generally very small, they were statistically significant in half of the plots at 20 and 50-m resolution (Fig. S6-8), consistent with the results of the analysis of the slope of spatial variability with plot scale (see above)."

We nevertheless add a new clause explaining the connection: "showing that even weak spatial aggregation may have an influence on the scaling of variability in AGBD".

362 gentler -> smaller
Response: Because slopes are negative, we have replaced gentler by shallower.

364-366: what is the justification for this statement?
Response: Because the slope is shallower in presence of spatial autocorrelation. This phenomena is well illustrated by the Fig. S3 referenced here.

377: "might explain"; how does it do this?
Response: We removed this from the result section as it is covered in more detail in the discussion.

378: "gap of variability" means what?
Response: Replaced by "low variability".

384: rm comma
Response: Done

390; rm "(…error)"
Response: Done

395: I think the jargon phrase "dilution bias" should be removed everywhere from the paper. A clear example of why it is jargon is that the authors need to explain what it means at 395: it is slope estimation, so why not just use this simpler, easily understood phrase?
Response: We believe that it is important to retain the term "dilution bias" because this term has been widely used in the statistical literature and it is important that readers can easily link our study to this literature.

4 Discussion: large parts of the discussion just repeat what is above without adding new insight or information and I think it should be considerably shortened (by about 50%).
Response: We carefully checked the discussion and removed some parts that were indeed too repetitive.

424 -428: this relates to a remark made above. The authors are confusing elevation change (which can clearly have an inpact on biomass) with variability, where the connection is less obvious.
Response: See response above

436 et seq: this repeats what was said above without resolving the contradiction. Does significant in 439 mean statistically significant? If it is statistically significant but small how does this translate into AGBD
Response: See response above. We now mention that this is statistically significant.

440: what is spatial aggregation? How does topography "explain" this aggregation
Response: Because elevation range explains the cross-site variation in wavelet variance at large spatial scales (>100 m), spatial aggregation of AGB can be partly explain by topography (Fig. 5b).

444 rm larger
Response: Done

445: what is wavelet coherence?
Response: We rephrased the sentence as: "Detto et al. (2013) likewise found strong wavelet coherence between canopy height (a proxy for AGBD) and topography at scales of 100-800 m".

449: this is another example of where lack of any theory renders the txt obscure; what is actually meant by "more uniformity than expected"? And shouldn't this, if it exists, be visible in autocorrelation?
Response: We rewrote this sentence as "the wavelet analyses also revealed that spatial variability specific to scales of 25-75 m was lower (i.e., more uniformly distributed) than expected by chance."

475: In general is wrong; as noted, radar does not obey this
Response: This part has been deleted.

477: this is a very dubious statement, since it contradicts on the sampling theorem from signal processing: an adequately sampled signal will not lose information by this geometric conversion (note that sampling is here used in a signal processing sense)
Response: This part has been deleted.

481 differences
Response: Done

492 Saatchi's map is not global
Response: Done, removed

513: rm careful: nothing in this paper discusses spatial structure as far as I could see.
Response: Both wavelet analyses and spatial variograms provide important information about the spatial structure of AGBD.

532: I think the correct variable is altitude – how do you stratify by topography?
Response: We rephrased this: "stratified by topographic position (e.g. ridges, valleys and slopes)"

547 rm "approach"
Response: Done

577: products -> estimates of biomass density
Response: Done

582: "- plots" -> "that are"
Response: Done

Figures: Many of the captions should be shortened to simply explain what the figure or plot actually represents. Several of them contain too much detail about methods and/or text that discusses the figure and is better placed in the main text; e.g. Fig 3.
Response: The legends of Fig. 3, 4, 5 and 7 have been shortened.

Fig 4 is not a plot of spatial autocorrelation, or of AGBD. A subplot has a size, not a resolution.
Response: The legend has been modified consistently.

Fig 5;"topo hetero" means nothing and should be removed
Response: "topographic heterogeneity " has been replaced by "elevation range" in the legend.

Fig.6 what does the colour bar refer to; are colours necessary?

Response: The color bar reflects the values reported within the figure (now stated in the legend). We agree that colours are not completely necessary but they help to visually convey the structure of the error at a glance, which we consider sufficiently useful to justify their use.