General comments

This paper attempts to produce representative values of 3 disturbance parameters and background mortality for forests within 25 km x 25 km tiles and 0.25° x 0.25° cells using forward simulations from a model that is massaged to give comparable values across 17 measures to those found using the 2010 GlobBiomass AGB map. It is very much a paper produced by a computer scientist in which the physical reality of disturbance is hardly considered. As a symptom of this, no definition of forest disturbance is given in the paper. Is it deforestation, degradation (in its many forms); does forest growth come into it? Similarly, no definition is given of the 3 disturbance parameters. For example, what is "extent" of disturbance in a possible mix of forest clearance and degradation? How is severity defined? And what is frequency (is this temporal or spatial frequency)? Bizarrely, the three parameters (μ, α, β) have different meanings in the main text and supplement. How is disturbance distinguished from natural variation in space?

However, the major problem with the paper is that it relies uncritically on the GlobBiomass map as being a faithful representation of disturbed forest regions without paying any attention to the limitations of the map, including biases, saturation effects, large dispersion at 25 m and possible spatial correlation induced by the data processing (which may have nothing to do with spatial correlation in AGB). There is no discussion of the qualities of the map and how they might affect its ability to truly represent the properties of disturbed forests. Hence the "disturbance" they are representing is entirely a description of pixel properties of the map, which may, inter alia, include false evidence of disturbance or fail to capture true disturbance.

Specific comments

The AGB map is for 2010 but the landcover is for 2019. How does this affect the results?

The phrase "potential extrapolation" is used in several places without ever being properly explained. The idea seems to be that if the simulation disagrees with the properties of the map then it's because the training set isn't large enough, but there's no evidence to support this. Why can't it be because the trained data just makes mistakes?

In a couple of places, biomass products are described as "globally consistent" but it's not clear what this means. Consistent with ground data (this isn't true)? Consistent with each other (very much not true)? Internally consistent error properties (not true)? So what is it?

Figure 2 suggests significant biases for some parameters over some ranges.

3.1.2 talks about aggregating 25 km tiles onto a 0.25° grid, but 0.25° corresponds to ~27.75 km (at all latitudes) so what is meant by aggregation?

There is no discussion of Fig. 4 and whether the patterns shown are consistent with other estimates, such as of turnover time, fire statistics, etc.

l. 289 - 295 This section first says uncertainty is high for the parameters, then it says it is high for β. Both can't be true. This section seems oblivious of the fact that GlobBiomass relies on ALOS-1 data which is completely insensitive to high biomass, so what information is it supplying in these forests and how does this affect the procedure in this paper?

In Fig. 6 the high and low end of values are indistinguishable so it's nearly impossible to interpret what is being shown. A better colour range is needed.

Fig. 7: what do the colours mean? The text is too small to be legible. I have serious doubts about the authors' interpretation of what these images say about disturbance for all three parameters, especially the high β example

Technical corrections

Is it Kb or K_b (many discrepancies)?

"Where" after an equation should be "where" in several cases.

Terminology is inconsistent: we have "landscape", "tile", "landscape tile", "landscape domain" and possibly other phrases all representing the same thing, as far as I can tell.

The paper in several places talks about EO data but the AGB map is not EO data. It is a product derived from EO data (with other inputs); this is an important distinction.

The authors use the phrase "scientific plausibility" in a few places. There's nothing scientific about this; it's just plausibility.

l.68 Le Toan

l. 145 superior skill compared to what?

In the supplement, table S2: what is pathes? Below, "integer"

S1.3. L_d is given 2 different definitions

Next para: what are the initial conditions and what effect do they have?

How can a suite of metrics be methodogically identical to anything?

Fig S2.2 Labelling illegible

S2.2 This talks about influential features but the meaning of this is never explained. Influential on what? How is influence measured? Do we really need 17 metrics?

Fig S2.3 What's a simulation-observation gap?

S3 is full of confusing labelling, sometimes f, sometimes j, and no statement that x is a vector

S3.2 talks about Section 2.5.1; there is no such section

Fig S4.1 conveys almost no information as it all just looks bluey-grey; it needs a much better choice of display format.