Comment on gmd-2021-93
Anonymous Referee #2

Referee comment on "Constraining a land cover map with satellite-based aboveground biomass estimates over Africa" by Guillaume Marie et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-93-RC2, 2021

This study attempts to improve the representation of plant function types (PFTs) over the African continent using a Bayesian approach informed by a map of aboveground biomass. Although the topic of the manuscript is important, I recommend rejection based on both technical and ethical reasons.

First, I think it is ethically questionable that a group of European scientists publish an analysis on Africa's natural resources without any involvement of African scientists. My motivation for this comment is the recent debate generated by the article of Misnany et al. (2020, https://doi.org/10.1016/j.geoderma.2020.114299), where they describe the concept of Helicopter Research as a form of neo-colonialism. Although the situation here is slightly different than that of collecting samples for scientific analysis, I still believe that similar questions can be asked in this manuscript. Why are African scientists not involved in this manuscript given the importance of defining PFTs and biomass for their own ecosystems? In their article, Misnany et al. highlight four negative aspects of this type of studies that do not involve local scientists: 1) Ignoring land ownership and disrespecting sovereignty. 2) Having little contribution to local science and development. 3) Promoting exclusivity—potential benefits to the studied country are often neglected, and further widens the gap between developed and developing countries. 4) Creating negative sentiments in local scientists towards international research. I think these issues should be addressed by the authors before this manuscript is considered again for publication.

Second, the use of the Bayesian approach is poorly developed. In particular, the choice of prior distributions is not consistent with formal theory for the specification of conjugate priors. For instance, the use of an uniform distribution U ~ (0, 200) for the prior distribution in equation 8 has no theoretical support; it leads to a distribution of biomass that extents to the negative side. In general, the formal Bayesian concepts for specifying hyperparametes are not used in this analysis. Therefore, I question the theoretical validity of the results presented in this study.

Minor comments
- Line 21 and thereafter. You use comma to separate decimal places. This is not standard notation in the English language.
- Ln 145. ABG -> AGB.
- Ln 204. The correct spelling is 'confidence interval'. However, notice that in Bayesian statistics the correct term to use is 'credible interval' (see https://en.wikipedia.org/wiki/Credible_interval).
- Equation 5. Why do you assume a normal distribution? Biomass at the landscape level usually has a few sites with very large biomass. A distribution with a longer right tail would be more appropriate. Please provide a rationale for the selection of the gaussian distribution.
- Equation 8. This choice of prior for the standard deviation is unreasonable. It inevitably leads to negative biomass values.
- Equation 9. I also see a problem with this choice of distribution. Once you pick one random value for one of the fractions, the other values are not independent. The Beta distribution alone cannot deal with this situation. The classical way to address this problem is with Dirichlet priors (see https://en.wikipedia.org/wiki/Dirichlet_distribution).
- Ln 270. Why a reference to a study in preparation? There are hundreds of papers using Orchidee, and it has been described extensively everywhere else.
- Ln 287. Revise sentence.
- Ln 306. Three?
- Ln 630. Without the African AGB map being publicly available this study would not meet reproducibility standards.