Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2019-202-RC2, 2019

Biogeosciences
Discussions

Interactive comment on "Improving non-representative-sample prediction of forest aboveground biomass maps: A combined machine learning and spatial statistical approach" *by* Shaoqing Dai et al.

Anonymous Referee #2

Received and published: 4 July 2019

General comments

This is an excellent study providing most accurate AGB maps based on fusion of forest inventory data and machine learning techniques. The approach combines the advantages of machine-learning algorithms and accounts for spatial correlation among data points, as well as, non-linear relationships between environmental covariates by using the P-BSHADE model. Research questions investigating (1) differences among the presented methods, and (2) how to improve the accuracy of AGB maps are evaluated by providing statistical error metrics, such as mean absolute error (MAE), mean rela-

tive error (MRE) and root means squarer error (RMSE). Based on evaluation of these metrics results suggest that among the different methods random forest (RF) in combination with the spatially explicit model produce the highest prediction accuracy for AGB maps and show that in comparison to traditionally applied allometric models estimates are congruent, but differ in local spatial distribution of AGB. Hence, these results indicate that the proposed method based on non-representative sample prediction of AGB maps should be capable of accounting for spatial heterogeneity of AGB and thus could enhance prediction accuracy of AGB maps.

Specific comments

Although this study introduces a very promising methodology – which due to a combination of machine-learning methods with spatial explicit statistical models, should be capable of resolving problems, such as non-linearity, complexity and spatial heterogeneity commonly comprised in available datasets based due to non-representative sampling of forest inventory plots – there are some minor issues that could be addressed:

(1) Given the fact that environmental parameters (i.e. meteorological variables are missing from the analysis the suitability of the proposed technique for forecasting AGB under future climate scenarios cannot be evaluated.

(2) Therefore, this does not allow to infer conclusions about the behavior of non-stationary systems, such as the response of global forest AGB to climatic signals.

(3) However, under steady-state assumptions, the presented approach can be used to derive management plans based on more accurate assessment of AGB from non-representative sampling plots, which can be compared among different geographic regions.

Nevertheless, after accounting for these minor issues this study should represent a valuable asset to the available literature focusing on improving prediction accuracy of

currently available AGB maps.

L41: design(s).

L86: please explain “stability of the second steps”.

L94: remove “tantamount” and focus on concise description of the research questions to be investigated in the discussion section, i.e. by (1) comparing the RMSE among different methods and (2) to interpret the accuracy of AGB maps.

L439-453: maybe start of this section based on the advantage of your method over the other studies presented here?

L479-486: this seems to be your main results, put up front and discuss according to the points presented here!

L519 (and throughout the text): please explain “single analytic trees” and “forest resource inventory data”.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2019-202, 2019.