

Appendix 1: Large-scale Maps Prediction Performance

We performed a secondary analysis to determine the best map for our study. This validation exercise is important for users that do not have UAV-LiDAR data and intent to validate canopy height global maps. Therefore, instead of analyzing the correlation between the large-scale map information and the ground truth data, we evaluate their ability to predict carbon stock. For this analysis, we used three different models: Linear Regression Model (Lm), Random Forest (RF), and Generalized Additive Model (GAM), where we only used the map information as covariate and carbon stock as response variable. We analyzed the performance of these maps to predict carbon stock using a 10-fold cross validation, where in each iteration one fold was held out for validation while the model was trained on the remaining data. Following each iteration of the cross-validation process, we calculated MAE as a metric of predictive performance.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \text{ [Eqn.1]}$$

where y_i is the observed carbon stock from the field plot and $\hat{y}$ is the predicted carbon stock by the model.

Our results showed that the Lang et al. (2023) map showed the lowest MAE mean values for the three models ($\overline{Lm} = 141.6$; $\overline{RF} = 163.8$; $\overline{GAM} = 140.5$), followed by the Meta's global canopy height map ($\overline{Lm} = 171.4$; $\overline{RF} = 191.9$; $\overline{GAM} = 162.9$), Ometto et al. (2023) Amazon AGB map ($\overline{Lm} = 168.3$; $\overline{RF} = 203.1$; $\overline{GAM} = 168.4$), and CCI global AGB map ($\overline{Lm} = 194.6$; $\overline{RF} = 228.8$; $\overline{GAM} = 195.6$). The similar pattern between the three models highlights the robustness of our results.

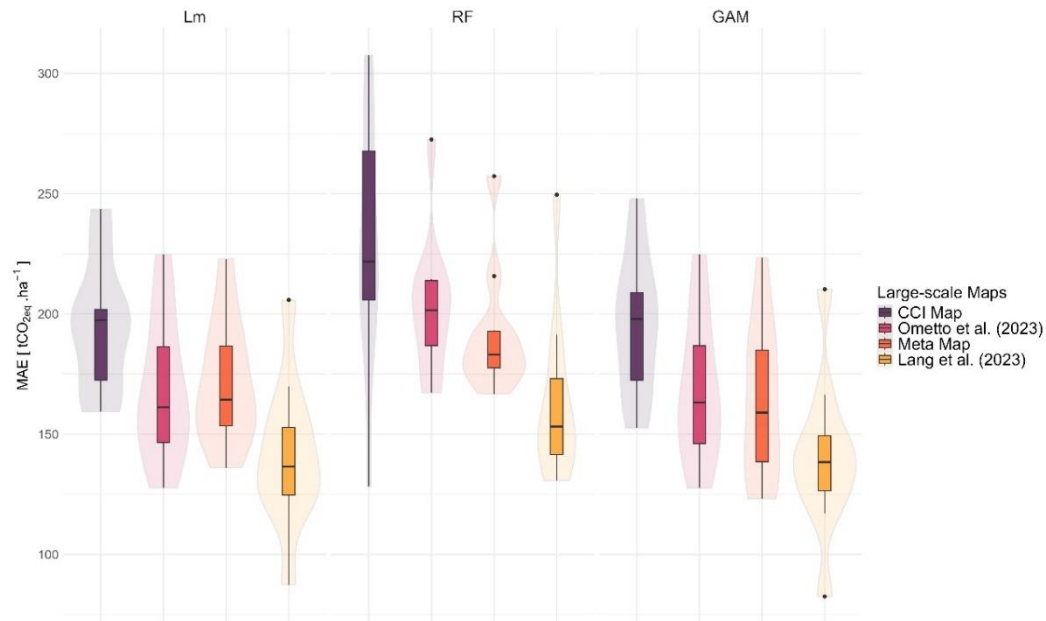


Figure 1. Boxplots of MAE values from the cross-validation exercise comparing the predictive ability of the different large-scale maps using three different models (Linear Regression Model (Lm), Random Forest (RF), Generalized Additive Model (GAM)). The models are used to estimate carbon stock, and the estimated values were compared to the observed values to calculate MAE.