living planet symposium BONN 2022

Biomass Estimation from Tree Heights on Individual-Level with Gaussian Process Regressor

Qian Song (1), Xiao Xiang Zhu (1, 2)
1. German Aerospace Center; 2. Technical University of Munich

Abstract

We propose a tree-level biomass estimation model approximating allometric equations by LiDAR data. Since tree crown diameters estimation is challenging from spaceborne LiDAR measurements, we develop a model to correlate tree height with biomass on the individual tree level employing a Gaussian process regressor. In order to validate the proposed model, a set of 8,342 data points on tree height, trunk diameter, and biomass has been assembled. It covers seven biomes globally present. We reference our model to four other models based on both, the Jucker data and our own dataset. Although our approach deviates from standard biomass–height–diameter models, we demonstrate the Gaussian process regression model as a viable alternative. In addition, we decompose the uncertainty of tree biomass estimates into model- and fitting-based contributions. We verify the Gaussian process regressor has capacity to reduce the fitting uncertainty down to below 5%. Exploiting airborne LiDAR measurements and a field inventory survey on the ground, a stand-level (or plot-level) study confirms a low relative error of 1% for our model.

Motivation

Tree-level biomass product:
- Field inventory
- Allometry equations

Coarse-resolution biomass product:
- Using remote sensing images
- Low accuracy, but large coverage

Middle-level biomass product:
- Remote sensing data + allometry equations

Fig. 1. Illustration of biomass products with fine (individual), middle, and coarse-grained (stand) level.

A trade-off between time-consuming in-situ measurements and coarse-resolution biomass estimation is the estimation of tree-level parameters (such as height, crown diameter, etc.) from high-resolution remote sensing data as input to allometric equations. Although highly correlated with biomass, parameters such as wood density and diameter cannot reliably get estimated by aerial imagery. Jucker et al. confirmed that height and crown diameter of trees are sufficient to estimate the trunk diameter by a single equation [2]. Crown diameter and height can be derived from airborne laser scanning (ALS) data. However, the crown diameter estimation is a source of significant model error. In this paper, we proposed to estimate biomass from tree height, alone.

Results

We compare the single-input Gaussian process regressor (biomass-height) model [3] with a random forest (RF) biomass-height model, and three allometric equations, specifically: biomass–height–crown diameter (LR), biomass–height (LR2), and biomass–height–diameter (LR3). The form of the three allometric equations read:

LR: $\text{In}\, B = a\, \text{In}\, (H \times CD) + b + \epsilon$ (1) or LR: $\text{In}\, B = a\, \text{In}\, (D) + b + \epsilon$ (2)

LR2: $\text{In}\, B = a\, \text{In}\, (H) + b + \epsilon$ (3)

LR3: $\text{In}\, B = a\, \text{In}\, (H) + b\text{In}(D) + c + \epsilon$ (4)

where a, b, c are the coefficients and bias terms determined by the training data; CD refers to the crown diameter; and ϵ is model residuals. Since no crown diameter measurements in our collected data, we utilize an alternative biomass-diameter model in Eq. (2).

To evaluate model accuracy, three indices get derived: R–squared (R2), root mean square error (RMSE), and model bias. R–square refers to the coefficient of determination, and is defined according to

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (5)$$

where y_i and $\hat{y}_i$ are the ith ground truth and predicted values. $\bar{y}$ amounts for the average mean of ground truth. According to these definitions, the R–squared score may receive impact by a single, strongly biased estimation. Thus, calculating R2, we exclude outliers when the corresponding absolute error exceed the mean absolute error by at least three times, cf. red circles in Figure 3.

$$RMSE(y_i, \hat{y}_i) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (6)$$

Bias relates to relative systematic error. It is defined as

$$Bias(y_i, \hat{y}_i) = \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{y}_i - y_i}{y_i} \quad (7)$$

The negative or positive value of the bias indicates biomass under- or overestimation.

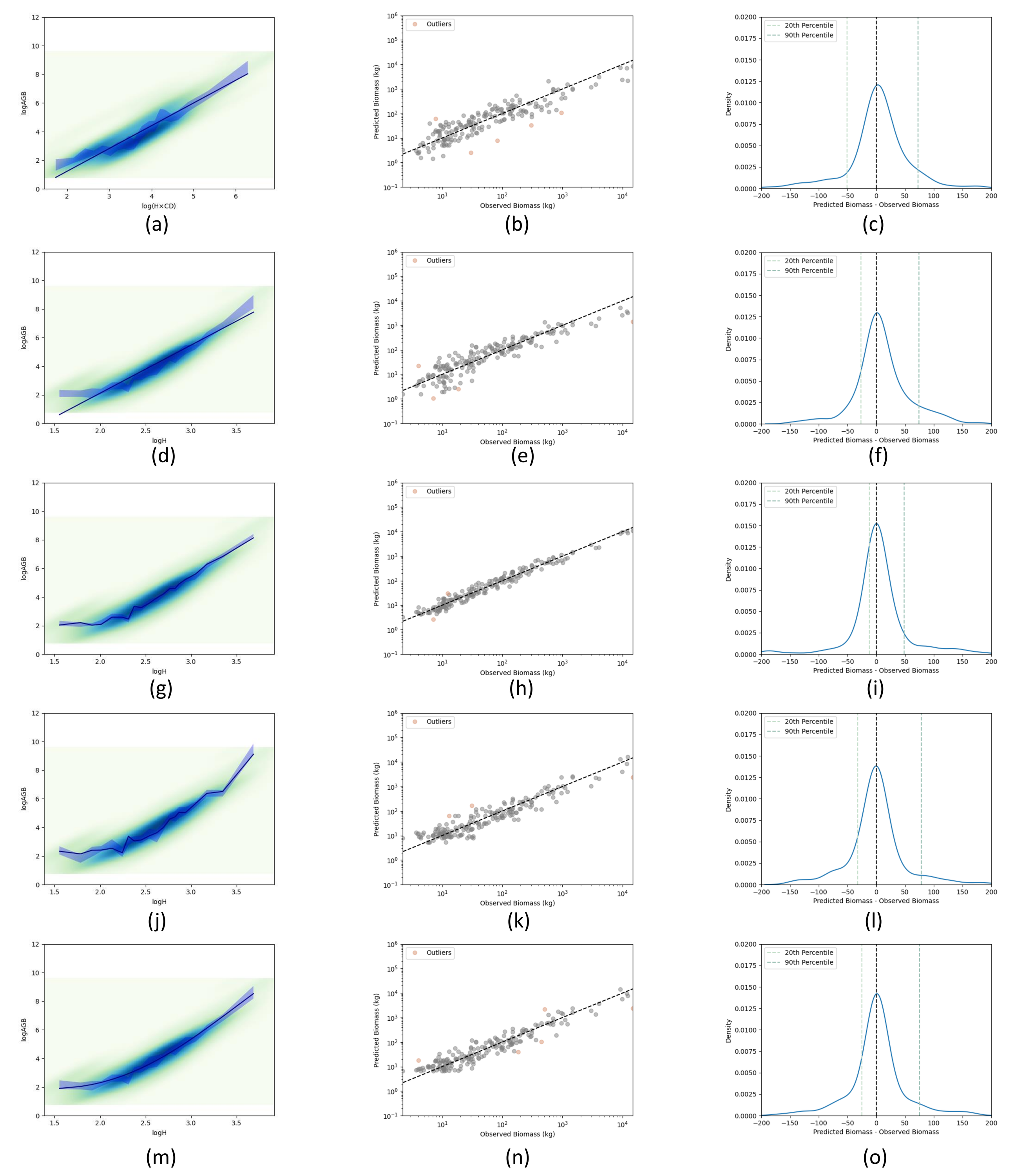

Fig. 3. Plots of the fitted curves with corresponding prediction errors (left column), scatters of predicted and observed are biomass shown in the middle column, and the distributions of errors is depicted by the contents of the right column. The evaluation is based on the Jucker data [2]. Each row corresponds to one of the five models—from top to down: LR (a)-(c), LR2 (d)-(f), LR3 (g)-(i), RF (j)-(l), GPR (m)-(o).

Dataset

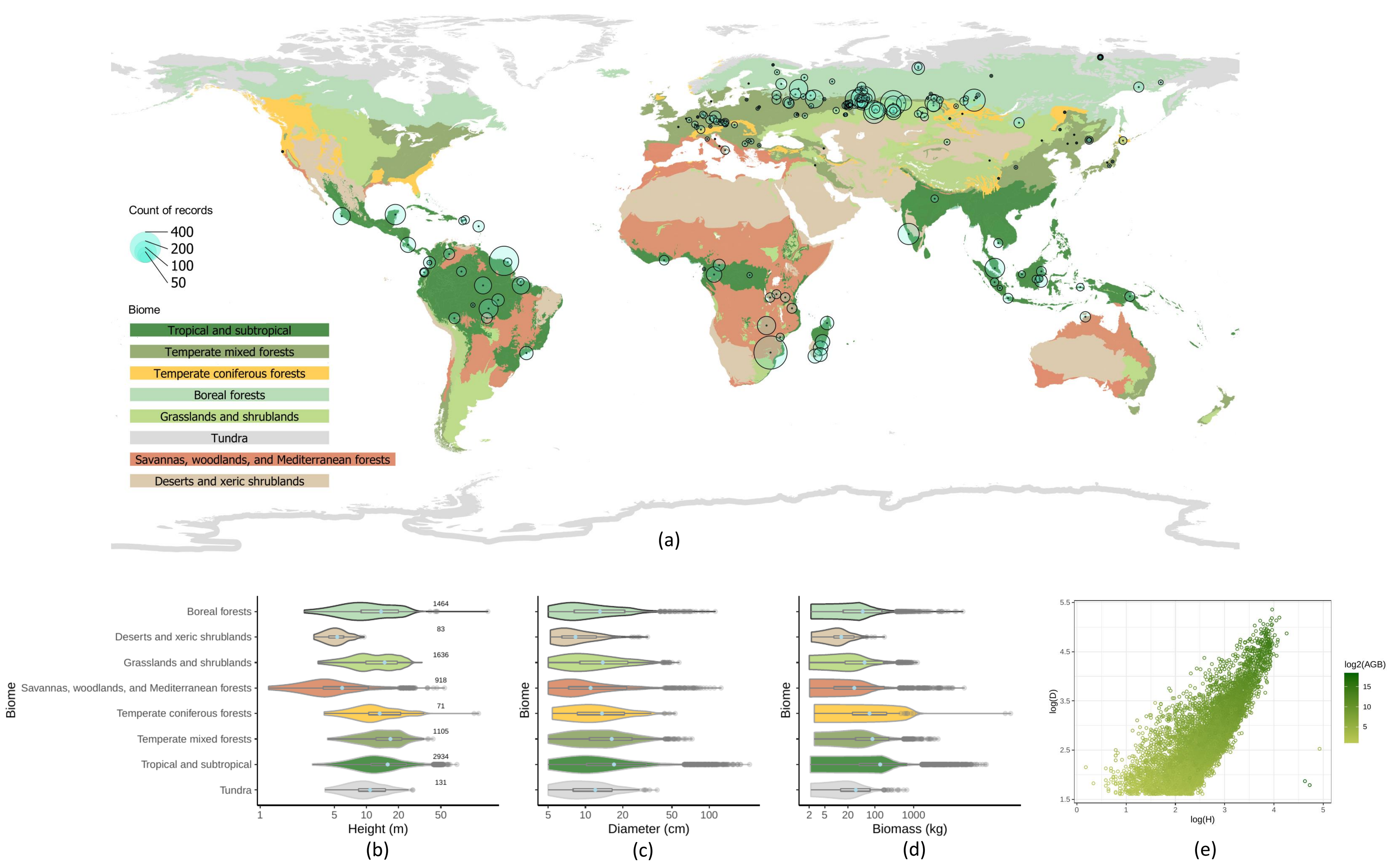

Fig. 2. Summary of data collected: Geospatial distribution of the measurements plotted on top of the biome classification map. Circle diameters represent the number of records at each geo-location (a); violin plots of the distributions of tree heights in meters (b), tree diameters in centimeters (c), and above-ground biomass in kilograms (d) for various biomes. The number of records for each biome is shown as absolute number to the right.

SUMMARY OF R-SQUARE SCORES, RMSE AND BIAS OF A SERIES OF REGRESSION MODELS FOR BIOMASS ESTIMATION BENCHMARKED ON THE JUCKER DATA.

	R2	RMSE	Bias
LR	0.664463	1108.855	0.26314
LR2	0.53256	1466.553	0.2931
LR3	0.950643	424.6752	0.079669
RF	0.803928	1147.066	0.20911
GPR	0.837668	1117.9	0.218732

SUMMARY OF R-SQUARED, RMSE, AND BIAS FOR FIVE REGRESSION MODELS ESTIMATING BIOMASS FROM THE DATASET CURATED.

	R2	RMSE	Bias
LR	0.749947	8230.41	0.150352
LR2	0.245381	7722.323	0.504575
LR3	0.780377	8204.319	0.109171
RF	0.812044	3679.73	0.366505
GPR	0.656328	4950.192	0.34472

COMPARISON OF MODEL PERFORMANCE IN TERMS OF RELATIVE ERROR AND RELATIVE RMSE FOR CANDIDATE MODELS: LR, RF, AND GPR

	LR	RF	GPR
RE	0.088971	-0.15184	0.011134
%RMSE	0.074465	0.227719	0.123556

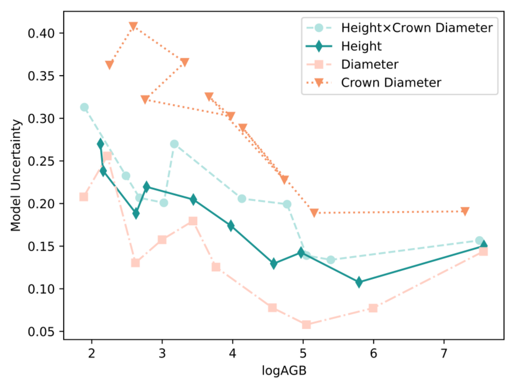

Fig. 5. Study of model uncertainties when working with tree height H, tree diameter D, crown diameter CD, and the product H × CD as input parameter of the allometric equation. The overall model uncertainties read: 18.25%, 14.13%, 29.81%, and 20.57% respectively.

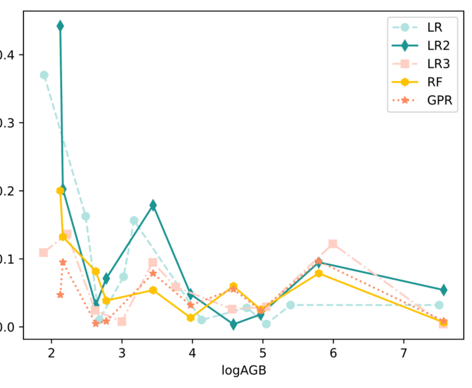

Fig. 6. Biomass-dependent fitting uncertainties for the five candidate models trained on the Jucker data. The overall fitting uncertainties of the five candidate models reduce to 8.80%, 11.45%, 6.13%, 6.90%, and 4.50% respectively.

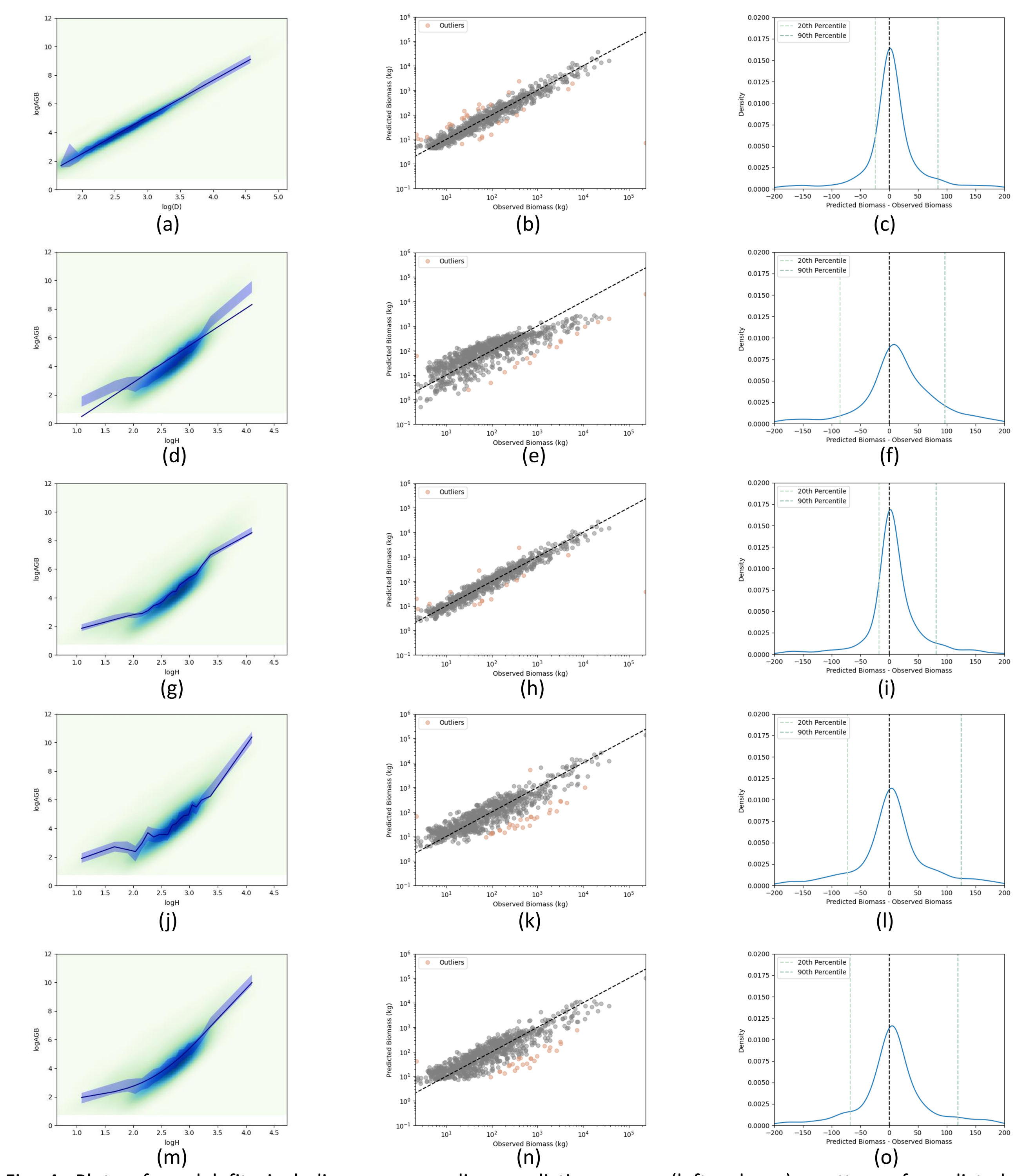

Fig. 4. Plots of model fits including corresponding prediction errors (left column), scatters of predicted and observed biomass (middle column), and the distributions of errors (right column) based on curated data. Each panel corresponds to one of the five models, i.e. LR: (a)-(c), LR2: (d)-(f), LR3: (g)-(i), RF: (j)-(l), GPR: (m)-(o).

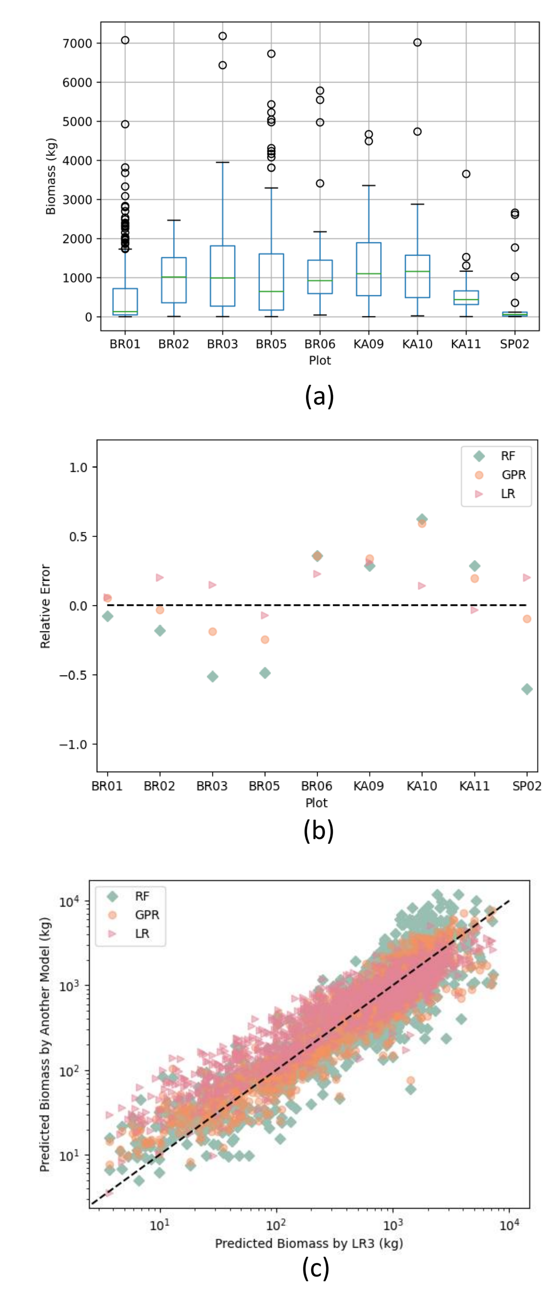

Fig. 7. (a): Box plot of predicted biomass by LR3 model on individual level grouped by plot; (b): the relative error of the three models in each plot where the LR3 model is assumed ground truth; (c): scatter plot of model predicted biomass versus LR3 model reference. The data are from [4].

References

[1] M. Schlund, S. Erasmi, and K. Scipal, "Comparison of aboveground biomass estimation from insar and lidar canopy height models in tropical forests," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 3, pp. 367–371, 2019.
[2] T. Jucker, J. Caspersen, J. Chave, C. Antin, N. Barbier, F. Bongers, M. Dalponte, K. Y. van Ewijk, D. I. Forrester, M. Haeni, et al., "Allometric equations for integrating remote sensing imagery into forest monitoring programmes," Global change biology, vol. 23, no. 1, pp. 177–190, 2017.
[3] G. Camps-Valls, L. Martino, D. H. Svendsen, M. Campos-Taberner, J. Munoz-Mar ˜ ´ı, V. Laparra, D. Luengo, and F. J. Garc´ıa-Haro, "Physicsaware gaussian processes in remote sensing," Applied Soft Computing, vol. 68, pp. 69–82, 2018.
[4] H. Weiser, J. Schafer, L. Winiwarter, et al., "Terrestrial, UAV-borne, and airborne laser scanning point clouds of central European forest plots, Germany, with extracted individual trees and manual forest inventory measurements," 2021.

#LPS22