This manuscript attempts to improve AGB maps by combining the machine learning model (SVM, RF, and RBF-ANN) and spatial statistical model (P-BSHADE). Overall the manuscript seems technically sound and, in most cases, is well written. The experiment is designed for one type of forest (Eucalyptus forest), thus I'm afraid the influence is limited. The results are reasonable, I question a number of aspects of the source data. Based on the comments below I suggest major revisions.

Response:

Thank you for your consideration of our manuscript.

We agree that the articles published in *Biogeosciences* should be universal and representative, and the research results can help to solve the hot issues in ecology and geosciences. Therefore, we understand your concern about the influence of our manuscript. We are sorry that we did not clearly explain the innovation and influence of our manuscript. We explain this as follows:

1. The innovation of our manuscript is to integrate machine learning and a spatial statistical model. The integration of these two can help to complement each other 's advantages and improve the accuracy of the AGB estimation model. Machine learning has the advantage of being able to handle complex and potentially nonlinear relationships between forest AGB and other variables. However, the initial samples of machine learning were randomly selected, which may lead to differences in the results of each operation of the model. Additionally, specific machine learning algorithms have their own disadvantages, such as RF uses the average value of all regression trees in the calculation, which may result in the overestimation of the lower value and the underestimation of the higher value. As opposed to machine learning, the P-BSHADE model (a spatial statistical model) takes into account the spatial autocorrelation and spatial heterogeneity of forest AGB and of environmental covariates and remedies the bias of the observed values of the sampling plots in theory, which corresponds more to actual situations. A combined model takes the result of machine learning as the reference data (input data) of P-BSHADE so that the fitting process of the combined model accounts more for spatial relationships than is the case for the single machine learning model. In addition to the theoretical advantages of these methods, case studies presented in this study also demonstrate the empirical superiority of the combined

model.

2. The allometric model is a simple, fast, and universal equation that is used in many studies. However, selection error in plot-level allometric modeling still leads to over 40% uncertainty (Djomo et al., 2016; Fayolle et al., 2013; Chave et al., 2014), and simple or complex forms of the allometric model account for 20% – 60% of the uncertainty (Picard et al., 2015). In our manuscript, we propose an improved method of AGB estimation which combines machine learning which is good at prediction with spatial statistical model which is good at reflecting spatial relationship to improve the estimation accuracy of the AGB model at the plot level.

3. Over the past 20 years, with the growing area of *Eucalyptus* plantations around the world, Brazil, India, China, Chile, Spain, DR Congo, Australia, South Africa, and other countries have established contiguous *Eucalyptus* planting areas. There have been many studies and reports on the study of the biomass estimation of *Eucalyptus* plantations combined with ecological process models at different temporal and spatial scales. China is the country with the largest area of planted forest in the world. However, China's planted forest suffers from the three practical problems of low productivity, unsustainability, and incongruous production function and ecological function, which urgently need to be solved by appropriate management measures. On the one hand, with the rapid growth of population, the timber demand is also increasing rapidly. On the other hand, there is the severe reality that total global forest resources have declined sharply in recent years. Thus, many countries and regions are vigorously developing fast-growing non-native trees to alleviate the contradiction between the supply and demand of timber and forest products in order to maintain economic and social development. *Eucalyptus* is one of the fastest-growing trees, and is controversial in the development of planted forest. A special study on *Eucalyptus* is of great significance.

4. Although *Eucalyptus* is taken as a case study in our manuscript, the model we proposed can also be applied to other tree species or mixtures of tree species. There is no particular relationship between our model settings and tree species. No unique characteristics of *Eucalyptus* are added to the model, and other forest types also can provide the input data such as tree height, DBH, longitude, and other variables. Therefore, the model we proposed can be expanded and its

influence is not limited to *Eucalyptus* forests. In fact, this idea of improving the accuracy of the model also has the potential to solve other ecological problems.

To sum up, we did not clearly express our meaning at first, however, in view of the influence of the four aspects of the manuscript we described above, we believe it can attract attention and citation. I hope it has been explained clearly now.

General questions:

Q1: Your number of sample plots (N=30) is too small for machine learning models. How could the sample plots represent the region? The range of biomass (1.02 – 135.79 Mg/ha). Did this range cover the whole range of your target species? The age of harvest trees ranged from 1 to 10 yr. What's the age range in the study area? What's the DBH range of your 90 harvested trees? Same questions as biomass and age.

Response to Q1:

1. For machine learning, the data volume of 30 plots is small. In practice, it is very difficult to obtain 30 plots in forestry, in which three trees were cut down for each plot, tree height and DBH of each tree were measured for each plot to obtain the most real plot data. If conditions permit, we attempt to obtain more representative plot data. However, we must also consider the cost–benefit problem in practical applications. Therefore, we took three measures to make up for the possible problems caused by the small sample size. The first was to use the spatial block cross-validation to make full use of the existing samples, and also to avoid overfitting; the second was to combine the P-BSHADE spatial statistical model to remedy the fitting results biased in machine learning; and the third was to investigate another 22 sample plots to test our proposed model again.

2. The setup of 30 plots takes into account the spatial heterogeneity of the whole study area and was carried out strictly according to the selection criteria of a set of representative forestry plots. We have included a detailed introduction of how the 30 sample plots were determined in the Supplementary Material as follows:

"The purpose of plot selection was to establish fixed and permanent plots representing regional Eucalyptus growing conditions and to provide harvested tree data on the single-tree scale with adequate consideration of spatial heterogeneity. Patches were selected first and met the following six conditions: (1) patch records were available from FMPI data for 2009; (2) forest stands were classified as timber or commercial forest; (3) forest patches were disturbance-free during the previous seven years, including but not limited to logging, fire, and pests; (4) forest patches were not replanted; (5) patches contained closed canopy forests; and (6) patches were monocultures, not mixed stands. Based on these six conditions, 2,980 Eucalyptus patches were selected from the FMPI data and fixed and permanent plots were established. The 2,980 selected patches were divided into ten groups based on forest age. Each stand group had been planted at the same time. We calculated the mean basal area for each group and used this as the basis for fixed plot selection, which was obtained from specified plot design and sampling procedures. In parallel, we considered site conditions, forest use, and forest origin (natural vs. man-made), and subsequently established 30 permanent square plots (20 m × 20 m)."

3. The AGB of the 30 sample plots was found to be 1.02–135.79 mg/ha, which covers the regional AGB of 4.70–119.78 mg/ha calculated with the allometric model and the regional AGB of 7.54–89.93 mg/ha calculated with the optimal model (RF and P-BSHADE method).

4. The age range of *Eucalyptus* forest in the study area is 1–51 years, however only 24 out of the 2980 patches are over 10 years old. According to the growth habits and characteristics of *Eucalyptus*, its growth after the age of 10 is very slow compared with that in the previous 10 years. In China, in order to maximize economic benefits, the common practice is to cut down *Eucalyptus* trees that have been growing for 10 years for commercial use and replant new seedlings. The 24 of these small patches may not be cut down for special reasons.

5. The DBH in the 30 sample plots ranged from 2.1–18.4 cm, and that in the study area ranged from 5.0–60.0 cm. Of the 2980 patches, only 53 had DBH values larger than 18.4 cm, 22 had values larger than 20 cm, and only two had values larger than 40 cm. It should be noted that 90 trees were selected and cut down based on the average DBH of each plot, so the DBH of the 90 trees were respectively at the average level of their forest age.

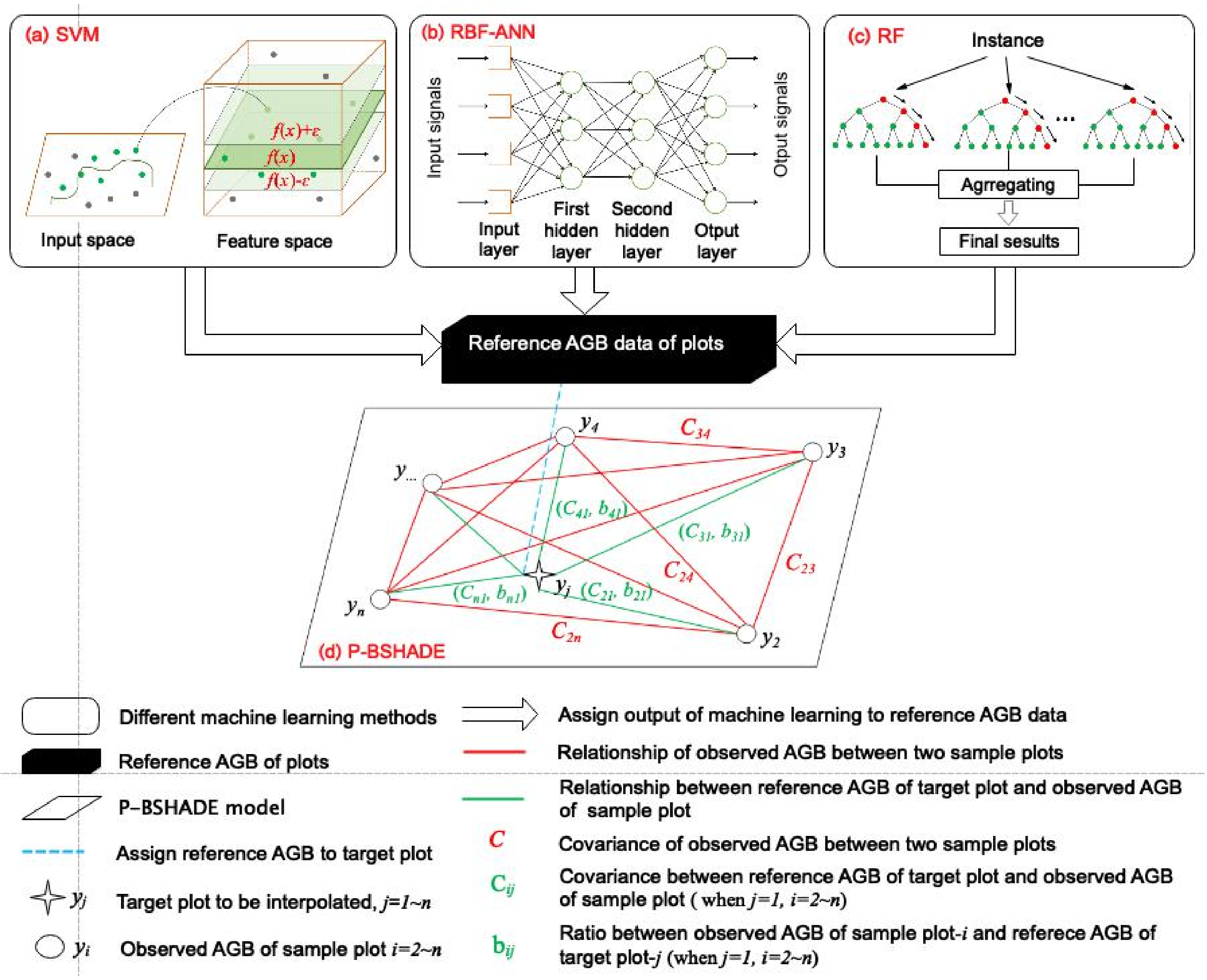
Q2: How did you combined machine learning and spatial models? I got confused after reading your descriptions. Comparing to other methods, your description of the P-BSHADE model a little bit lengthy, suggest moving details into the supplement.

Response to Q2:

P-BSHADE is a spatial statistical inference model. It requires the reference value of the target object (which is forest AGB in our manuscript) as input data when estimating. Determining how to obtain the reference value is the key to combining the P-BSHADE model with other methods. In our manuscript, we first used a machine learning method to estimate AGB and then used this estimate as a reference value to input into the P-BSHADE model to estimate AGB again.

The revised figure is as follows:

Also, thank you for your suggestion, we have moved details of P-BSHADE into the Supplementary Material (Section S1).



Q3: <u>Your results suggest that plot-level biomass models need to be built per species and per ecoregion? The problem of not using allometric models is, how to quantify the AGB of not-so-common species?</u>

Response to Q3:

We suggest to determine whether it is necessary to establish a plot-level model for per species and per ecoregion according to service objectives, cost-effectiveness, etc.

Allometric models has been widely used to estimate forest biomass, however biogeographic variation in allometric relations, wood density, and soil fertility introduce sources of uncertainty, which will be expanded and propagated in the next step when the model is applied to the region. The combination of machine learning and spatial statistics proposed in our study can reduce the error caused by the plot-level model. In our manuscript, although we take pure *Eucalyptus* forest as a case study, it does not mean that our approach can only be used to establish a plot-level model for *Eucalyptus*. Rather, our model can be applied to both pure species forests and mixed species forests, and furthermore can be applied to a single ecoregion or a larger-scale ecoregion, as long as enough sample plots representing the research region can be collected. When a high-precision biomass map with clear spatial distribution is required, our model can help to reduce the error of the plot-level model. This method is attractive as the plot-level model allows the subsequent construction of large-scale regional biomass maps.

Q4: Did you consider the influence of plot size? Say could your model build using a 20-m plot applied to 40-m or 100-m scale? This important when considering the need to apply models at larger geographical domains via the combined use of remote sensing datasets.

Response to Q4:

Yes, the influence of plot size is very important. In order to avoid this influence, the model established in our manuscript can only be applied to 20-m plots. Considering the universality of the model and the convenience of direct comparison between different research results, a 20-m plot size is usually selected in subtropical regions. If it is necessary to apply the model to plots with a size of 40-m or 100-m, it is better to establish plot-level models optimized for these sizes. That is, the setting of the model plot size should be adjusted according to the resolution of the data to be combined in the next step when the model is to be applied to the regional scale.

The resolution of remote sensing data and plot size are very important for up-scaling in the estimation of forest biomass. In particular, the sample plots not only provide validation data for large-scale prediction but also provide multi-source environmental variable information that can be used to calibrate models. Therefore, it is very important to set the plot size to be consistent with the resolution of the remote sensing data to be combined in the next step. However, the actual situation is that the larger the plot size, on the one hand, the required manpower and financial investment will increase, and on the other hand, the increase of workload and difficulty may affect the accuracy of the sample data. There have been many studies on the impact of sampling schemes (such as different sample sizes and plot sizes), model selection (such as different model types), and data selection (such as different types of remote sensing data) on the accuracy of forest biomass estimation and on how to organically combine the above three factors to obtain the optimal scheme. We think that this is another important issue, however it is beyond the scope of our manuscript.

Q5: This study constructs local AGB allometric models, for a small Eucalyptus forest in Nanjing county. However, how should we apply your method in other places over a large geographical domain?

Response to Q5:

We want to share and promote our model which combines machine learning and spatial statistics and thus combines the advantages of both in order to improve the accuracy of AGB estimation. Although in our study we took pure *Eucalyptus* forest as a case study, this model is not limited to the study of *Eucalyptus* forest, and furthermore can also be applied to other regions, including larger geographical areas (such as China, Asia, etc.). The application of this method to a larger geographical area should also follow the up-scaling steps from tree-level to plot-level to regional level. The most significant contribution of our model is to reduce the estimation uncertainty at the plot level, reduce the extent of further propagation or up-scaling of the uncertainty, and prepare for the combination of plot-level results and remote sensing data to realize the up-scaling of estimation and obtain high-precision distribution maps in the future.

Q6: Did you compare your models and existing allometric models within the region? What's the influence of excluding small stems (living stem <8 cm) in your estimation of AGB?

Response to Q6:

1. Our team has previously established plot allometric models of *Eucalyptus* forest biomass within this region (Qiu et al., 2018), however, we did not produce AGB allometric models for this region then. But in this manuscript, we have already established an AGB allometric model of the sample plots within this region. When using the P-BSHADE model, we used the results of the allometric model as reference AGB data. An accuracy comparison with other models established in our manuscript is shown in the table below:

	Method	MAE	MRE	RMSE
1	SVM	11.17	0.25	10.39
2	ANN-RBF	12.15	0.27	10.39
3	RF	10.16	0.26	9.43
4	P-BSHADE	18.37	0.39	14.08
5	SVM&P-BSHADE	6.88	0.12	6.30
6	ANN-RBF&P-BSHADE	10.14	0.20	9.63
7	RF&P-BSHADE	5.68	0.13	5.30
8	Allometric Model	14.90	0.53	23.04

2. After re-examining the data processing, we found that trees with DBH values of less than 8 cm were also included. At the beginning, the data collecting plan excluded trees with DBH values of less than 5 cm according to the "Detailed rules for the Eighth National Forest Resources Inventory of Fujian Province" standard. However, in the practical investigation, we found that most *Eucalyptus* forests aged 1–3 years had DBH values of less than 8 cm. If we had followed the aforementioned standard, our experiment would not have been carried out. Therefore, the actual operation later included all the trees, no matter what their DBH were. We are sorry for the mistake in the manuscript.

Specific comments:

L49: “the use of inadequate sampling data to construct the plot level prediction models” did you solve this issue?

Response:

We did not resolve this issue, since it involves sampling design. Our manuscript points out two primary sources of uncertainty in regional biomass maps, and we also mentioned that the second source of uncertainty (L52-53) was the main issue we were attempting to solve in this manuscript, as follows. Please check this. Although the first source falls outside the scope of this manuscript, we will study it in the future.

“The uncertainty of such regional maps can be attributed to two primary sources: (1) the use of inadequate sampling data to construct the plot level prediction models, and (2) model-dependent uncertainty, including unreasonable model-parameter assumptions and improper model structure (Chen et al., 2015; Gao et al., 2016; McRoberts et al., 2016). The present study mainly focuses on reducing the second source of uncertainty.”

L59: Selection of the allometric model could account for 20% uncertainty (Duncanson et al. 2017)

Duncanson, L., Huang, W., Johnson, K., Swatantran, A., McRoberts, R., & Dubayah, R. (2017). Implications of allometric model selection for county-level biomass mapping. Carbon Balance and Management, 12

Response:

Thank you for your suggestion. We have revised it.

L82: Some recent studies integrated ground-based plot and remote sensing data for AGB mapping (Sun et al. 2011; Huang et al. 2019; Qi et al. 2019)

Sun, G., Ranson, K.J., Guo, Z., Zhang, Z., Montesano, P., & Kimes, D. (2011). Forest biomass mapping from lidar and radar synergies. Remote Sensing of Environment, 115, 2906-2916

Huang, W., Dolan, K., Swatantran, A., Johnson, K., Tang, H., O'Neil-Dunne, J., Dubayah, R., & Hurtt, G. (2019). High-resolution mapping of aboveground biomass for forest carbon monitoring system in the Tri-State region of Maryland, Pennsylvania and Delaware, USA. Environmental Research Letters, 14, 095002

Qi, W., Saarela, S., Armston, J., Ståhl, G., & Dubayah, R. (2019). Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI lidar data. Remote Sensing of Environment, 232, 111283

Response:

Thank you for your suggestion. We have revised it.

L85: "First, existing studies that used machine learning methods have not considered the spatial heterogeneity of multiple environmental covariates (such as longitude, latitude, and forest structure)" This statement is too arbitrary. What does "structure" refer to? Shouldn't structure information came from lidar or radar?

Response:

Here, the forest structure refers to forest attributes such as stand volume, biomass, mean tree height, mean DBH, etc. It has been revised to "forest attributes like stand volume, biomass, mean height, et al."

L96: "multiple environmental covariates (such as longitude, latitude, and forest structure)" A duplicate statement, modify to be concise;

Response:

Thank you for your suggestion. We have deleted it.

L140: Suggest add equations of the allometric models you used here.

Response:

Thank you for your suggestion. We have added the model equation in the manuscript. Additionally, in the supplementary file, we have also added the specific parameters of the three models we established (Table B.3) as follows:

“$AGB=a[(DBH)^2H]^b$ ”

year	a	b
1 ~ 2	0.1538	0.6993
3 ~ 5	0.0377	0.9244
6 ~ 10	0.0689	0.8489

L175: What software/package did you applied to construct your model?

Response:

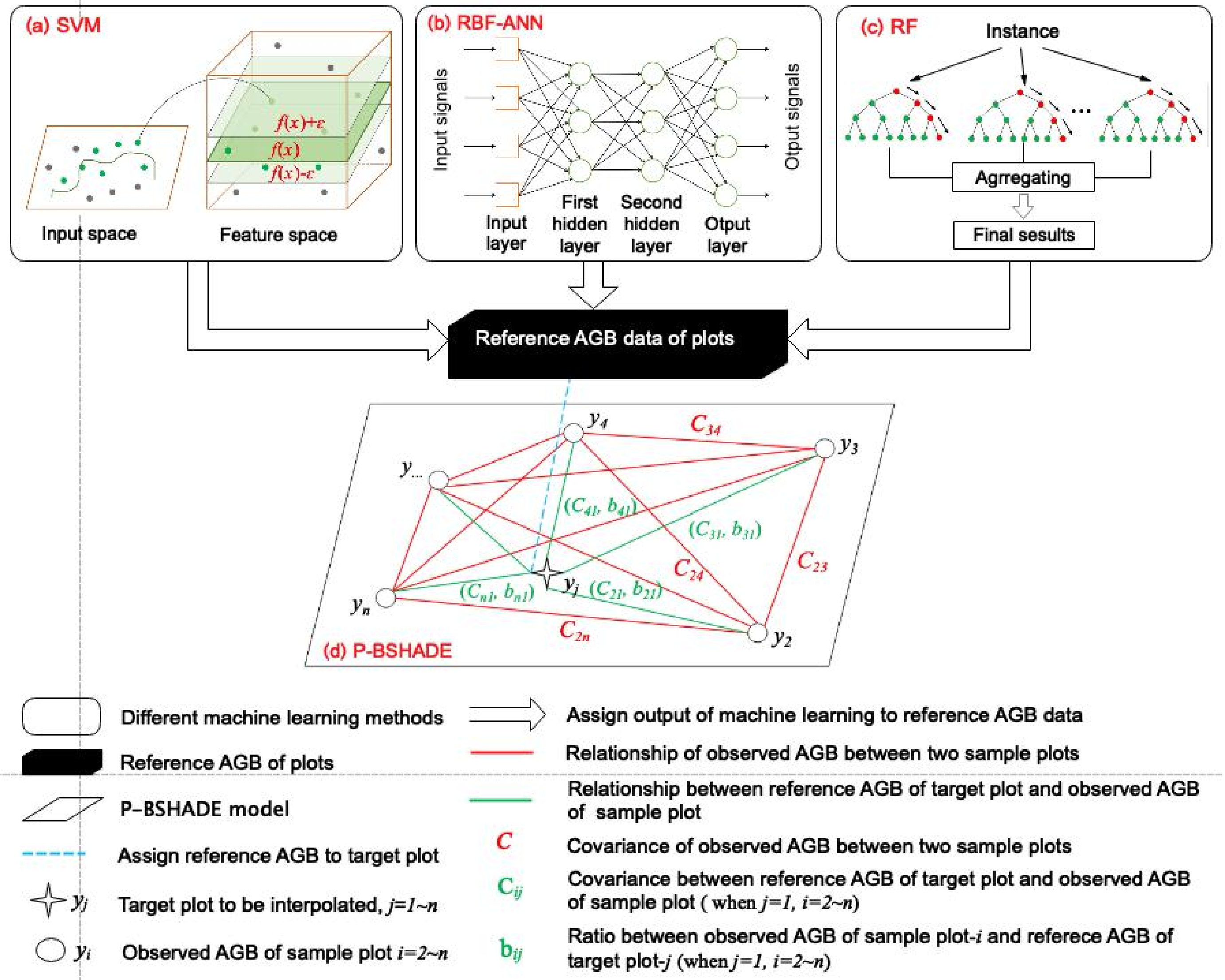
We used R3.5.3 (https://www.r-project.org) to construct our model.

L179: (reference series)? Figure 3. (b) SVM and (c) RF are for classification, not regression;

Response:

Reference series means the corresponding reference series in Figure 3. For better understanding, we have revised it to “ ’reference AGB data of plots’ in Figure 3”.

Thank you for your suggestion. Figure 3 has been revised.



L445: <u>"we" should be "We".</u>

Response:

Thank you for your suggestion. We have revised it.

REFERENCES

Qiu, Q. Y., Yun, G. L.,Zuo, S. D., Yan, J., Hua, L. Z., Ren, Y., Tang, J. F., Li, Y. Y., Chen, Q.: Variations in the biomass of eucalyptus plantations at a regional scale in southern china. Journal of Forestry Research, 29, 1263–1276, https://doi.org/10.1007/s11676-017-0534-0, 2018.

Djomo, A. N., Picard, N., Fayolle, A., Henry, M., Ngomanda, A., Ploton, P., McLellan, J., Saborowski, J., Adamou, I., and Lejeune, P.: Tree allometry for estimation of carbon stocks in African tropical forests, Forestry: An International Journal of Forest Research, 89, 446-455, https://doi.org/10.1093/forestry/cpw025, 2016.

Fayolle, A., Doucet, J.-L., Gillet, J.-F., Bourland, N., and Lejeune, P.: Tree allometry in Central Africa: Testing the validity of pantropical multi-species allometric equations for estimating biomass and carbon stocks, Forest Ecology and Management, 305, 29-37, https://doi.org/10.1016/j.foreco.2013.05.036, 2013.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B. C., Duque, A., Eid, T., Fearnside, P. M., Goodman, R. C., Henry, M., Martínez-Yrízar, A., Mugasha, W. A., Muller-Landau, H. C., Mencuccini, M., Nelson, B. W., Ngomanda, A., Nogueira, E. M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C. M., Saldarriaga, J. G., and Vieilledent, G.: Improved allometric models to estimate the aboveground biomass of tropical trees, Global Change Biology, 20, 3177-3190, https://doi.org/10.1111/gcb.12629, 2014.

Picard, N., Rutishauser, E., Ploton, P., Ngomanda, A., and Henry, M.: Should tree biomass allometry be restricted to power models?, Forest Ecology and Management, 353, 156-163, https://doi.org/10.1016/j.foreco.2015.05.035, 2015.