

---

## Supplementary Information for

### High-resolution maps of above- and belowground woody biomass in China from 2003 to 2020

Yongzhe Chen<sup>1,2</sup>, Xiaoming Feng<sup>1,2,\*</sup>, Bojie Fu<sup>1,2</sup>, Haozhi Ma<sup>3</sup>, Constantin M. Zohner<sup>3</sup>, Thomas

5 W. Crowther<sup>3</sup>, Yuanyuan Huang<sup>4,5</sup>, Xutong Wu<sup>6</sup>, Fangli Wei<sup>1</sup>

1 State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-  
Environmental Sciences, Chinese Academy of Sciences, Beijing, PR China.

2 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing,  
PR China.

10 3 Institute of Integrative Biology, ETH Zurich (Swiss Federal Institute of Technology), Zurich,  
Switzerland.

4 Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ,  
Université Paris-Saclay, Gif-sur-Yvette, France

5 Commonwealth Scientific and Industrial Research Organisation, Aspendale, Victoria,  
15 Australia

6 State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of  
Geographical Science, Beijing Normal University, Beijing, PR China.

\*Correspondence authors: [fengxm@rcees.ac.cn](mailto:fengxm@rcees.ac.cn)

---

## Supplementary text

### 20 Text S1. Reasonability of the selected predictors in the random forest for BGB estimation

When forest type (FOR\_T) is excluded, the BGB's predicting  $R^2$  according to the ten-fold cross validation will be slightly reduced to  $0.85\pm 0.02$ , while the RMSE will be slightly elevated to  $7.3\pm 0.6 \text{ t ha}^{-1}$ . On the other hand, if stand age information is not added as a supplementary predictor, the simulation  $R^2$  and RMSE of BGB in China will be  $0.86\pm 0.01$  and  $7.1\pm 0.5 \text{ t ha}^{-1}$ ,  
25 respectively. The seasonality of temperature and precipitation (Tsea and Psea) both have significant impacts on BGB. Without Tsea and Psea, the predicting efficiency of BGB in Chinese forests will be significantly reduced to  $0.84\pm 0.01$  in  $R^2$  and  $7.5\pm 0.4 \text{ t ha}^{-1}$  in RMSE. These analyses proved the reasonability of adding these predictors in the RF designed for BGB estimation.

30 Soil properties, e.g., soil sand content are reported to have an impact on root biomass allocation (Ma et al., 2021). However, compared to the climatic impacts on woody RSR, the direct edaphic effect is usually much weaker (Mokany et al., 2006; Luo et al., 2012). To prove this assumption, here we chose the 'soil database of China for land surface modeling' which was derived from 8979 soil profiles and the Soil Map of China using polygon linkage method (Shangguan et al.,  
35 2013). The basic edaphic variables with large quantities of soil profile records are selected, including soil texture (sand and clay fractions), pH value, soil organic matter, total nitrogen and soil total phosphorus. After adding these six edaphic predictors, the predicting  $R^2$  and RMSE for BGB according to ten-fold cross validations are  $0.88\pm 0.01$  and  $6.5\pm 0.5 \text{ t ha}^{-1}$ , respectively, compared to an  $R^2$  of  $0.89\pm 0.02$  and an RMSE value of  $6.3\pm 0.5 \text{ t ha}^{-1}$  when edaphic factors are  
40 excluded. Accordingly, the incorporation of soil property maps as predictors will not improve the simulation efficiency of BGB in China. As shown in Figure S13b, the contributions of edaphic factors are quite limited. By comparison between the contribution fractions of all predictors when edaphic factors are added and not added (see Figure S13), we may conclude that although soil conditions can slightly affect BGB, they are highly correlated with the plots'  
45 climatic and biotic conditions, making the edaphic factors no longer an essential input of the

---

random forest.

In fact, the currently available large-scale soil property maps are interpolated from site-scale measurements, or estimated through machine learning with climatic variables applied as key predictors, rather than directly observed. So, it is supposed to have much lower quality than climatic background data. Hence, to avoid over-fitting by the RF model, as well as introducing errors embedded in the soil property maps, we did not incorporate edaphic factors as predictors of the RF model for BGB mapping in China.

### **Text S2. Simplified uncertainty quantification methods**

Uncertainties of AGB calculated in this study came from four sources. First is the error within the improved reference AGB map for China. We quantified it as the uncertainty of the random forest intended for integrating three different AGB datasets and LiDAR-derived timber volume with the reference of plot-level measurements. During the ten-fold cross validation process, ten RF models were developed, while the AGB predicting uncertainty could be estimated as the relative error (the standard deviation divided by the mean value) from the ten-fold predictions. The extension of AGB's time series based on the long-term integrated VOD and VCF datasets brought the second source of uncertainty, which can be estimated from the ten-fold predictions of the 1/12° resolution AGB over 2003~2020 mapped through the random forest approach. This source of uncertainty was determined annually. The third source of uncertainty arose from the downscaling process of annual coarse resolution AGB. Apart from the AGBs per tree cover (TC) and per short vegetation cover (SC), we also calculated the 95% confidence intervals of these two regression coefficients, which were subsequently calibrated using the algorithm described in Section 2.3. Based on the upper and lower limits of the AGB per TC and AGB per SC, we derived two alternatives of the downscaled AGB products. Therefore, the uncertainty related to the downscaling process could be estimated as the ratio of the standard deviation of these three downscaled AGB to the standard high-resolution AGB map for each year. The 4<sup>th</sup> uncertainty was induced by the AGB decomposition process. AGB decomposition was achieved based on binary linear regression as well. Therefore, we calculated the 95% confidence intervals of per-

---

area forestland AGB and those of per-area shrubland AGB. The lower limits and the upper limits of these two variables were processed, which ultimately led to two alternative estimates of the sum of forest' AGB and shrublands' AGB in every woody pixel. Then, similarly, we calculated the standard deviation of the three estimates, and the ratio of which to the mean value can generally stand for the uncertainty induced by the AGB decomposition.

The uncertainty of BGB was mainly made up of the error within the AGB time series and the predicting uncertainty of the models which transform AGB to BGB. For forestland, the latter was calculated from ten-fold RF model predictions. In shrublands, to keep the consistency, we divided the whole dataset into ten parts, and then select nine of them each time to derive a set of regression coefficients. Ten sets of regression coefficients led to ten estimates of shrublands' BGB, which can produce the relative error induced by the conversion from shrublands' AGB to BGB.

### 85 **Text S3. Open data adopted in this study and the related references**

AGB data in (Liu et al., 2015) is available at: [http://wald.anu.edu.au/data\\_services/data/global-above-ground-biomass-carbon-v1-0/](http://wald.anu.edu.au/data_services/data/global-above-ground-biomass-carbon-v1-0/);

Biomass dataset in (Xu et al., 2021) is available at: <https://doi.org/10.5281/zenodo.4161694>;

GlobBiomass product is from: [https://globbiomass.org/wp-content/uploads/GB\\_Maps](https://globbiomass.org/wp-content/uploads/GB_Maps);

90 CCI-Biomass product is from: <https://climate.esa.int/en/odp/#/project/biomass>; (Santoro and Cartus, 2019)

GLASS-Biomass v2 product is available from: <http://glass.umd.edu/AGB/V02/>;

ATLAS/ICESat-2 vegetation height data is from: <https://nsidc.org/data/atl08/versions/4>;

95 GLAS/ICESAT-derived global forest canopy height map (i.e., Simard et al.'s map) is available from: [https://webmap.ornl.gov/wcsdown/dataset.jsp?ds\\_id=10023](https://webmap.ornl.gov/wcsdown/dataset.jsp?ds_id=10023);

Global Land Analysis and Discover (GLAD)'s Global Forest Canopy Height 2019 data is from: <https://glad.umd.edu/dataset/gedi>;

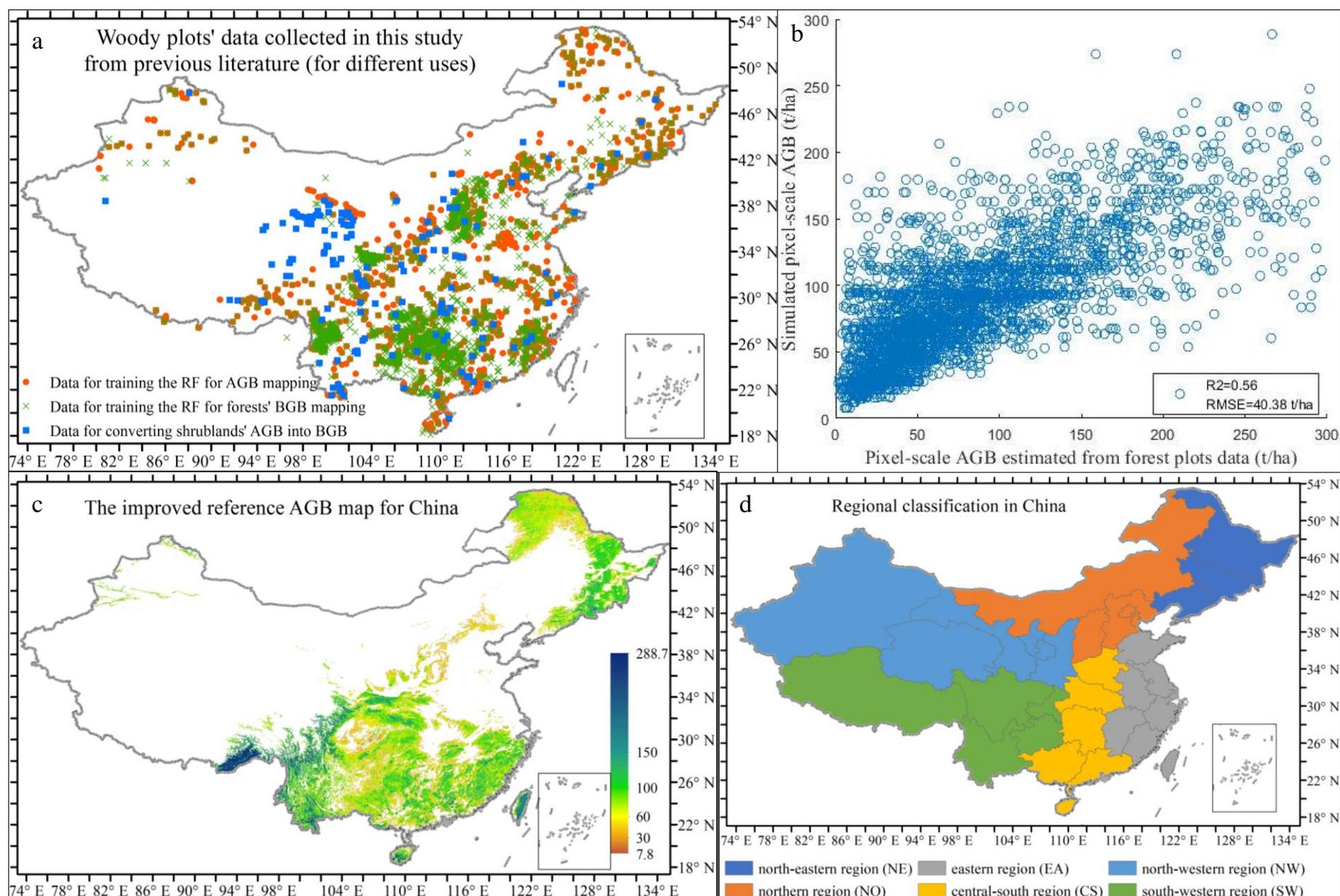
- 
- MODIS Vegetation Continuous Fields is from: <https://lpdaac.usgs.gov/products/mod44bv006/>;  
(Dimiceli et al., 2015)
- 100 ESA's 100 m global land cover data is from: <https://land.copernicus.eu/global/products/lc>;  
(Buchhorn et al., 2020)
- AMSR2/GCOM-W1's vegetation optical depth (VOD) acquired in ascending and descending paths are: [https://disc.gsfc.nasa.gov/datasets/LPRM\\_AMSR2\\_DS\\_A\\_SOILM3\\_001/summary](https://disc.gsfc.nasa.gov/datasets/LPRM_AMSR2_DS_A_SOILM3_001/summary) and [https://disc.gsfc.nasa.gov/datasets/LPRM\\_AMSR2\\_DS\\_D\\_SOILM3\\_001/summary](https://disc.gsfc.nasa.gov/datasets/LPRM_AMSR2_DS_D_SOILM3_001/summary);
- 105 VODCA VOD dataset is available from: <https://zenodo.org/record/2575599#.YVKEz44zaUk>;  
SMAP 9 km VOD data are from: [https://nsidc.org/data/SPL3SMP\\_E/versions/5](https://nsidc.org/data/SPL3SMP_E/versions/5); (O'Neill et al., 2021)
- ESA's 300 m annual land cover dataset, including the major forest type are obtained from: <http://maps.elie.ucl.ac.be/CCI/viewer/download.php>;
- 110 MODIS LAI data is from: <https://lpdaac.usgs.gov/products/mod15a2hv006/>; (Myneni et al., 2015)
- GEOV2 LAI data is available from: <https://land.copernicus.eu/global/products/lai>;
- National Forest Inventory (NFI) dataset can be found at: <http://www.forestdata.cn/index.html>;
- WorldClim bioclimatic background is from: <https://www.worldclim.org/data/worldclim21.html>;
- 115 Soil property maps for China is available at: <http://globalchange.bnu.edu.cn/research/soil2>;
- Remote sensing-based surface soil moisture (RSSSM) dataset is available from: <https://data.tpdc.ac.cn/zh-hans/data/d4e12be6-7dbe-453d-8833-079c5c73b0ff/> or <https://doi.org/10.1594/PANGAEA.940004>;
- Forest age map in China is provided by Yuan Zhang et al. (Mapping spatial distribution of forest age in China. Earth Space Sci., 4, 108-116)
- 120

---

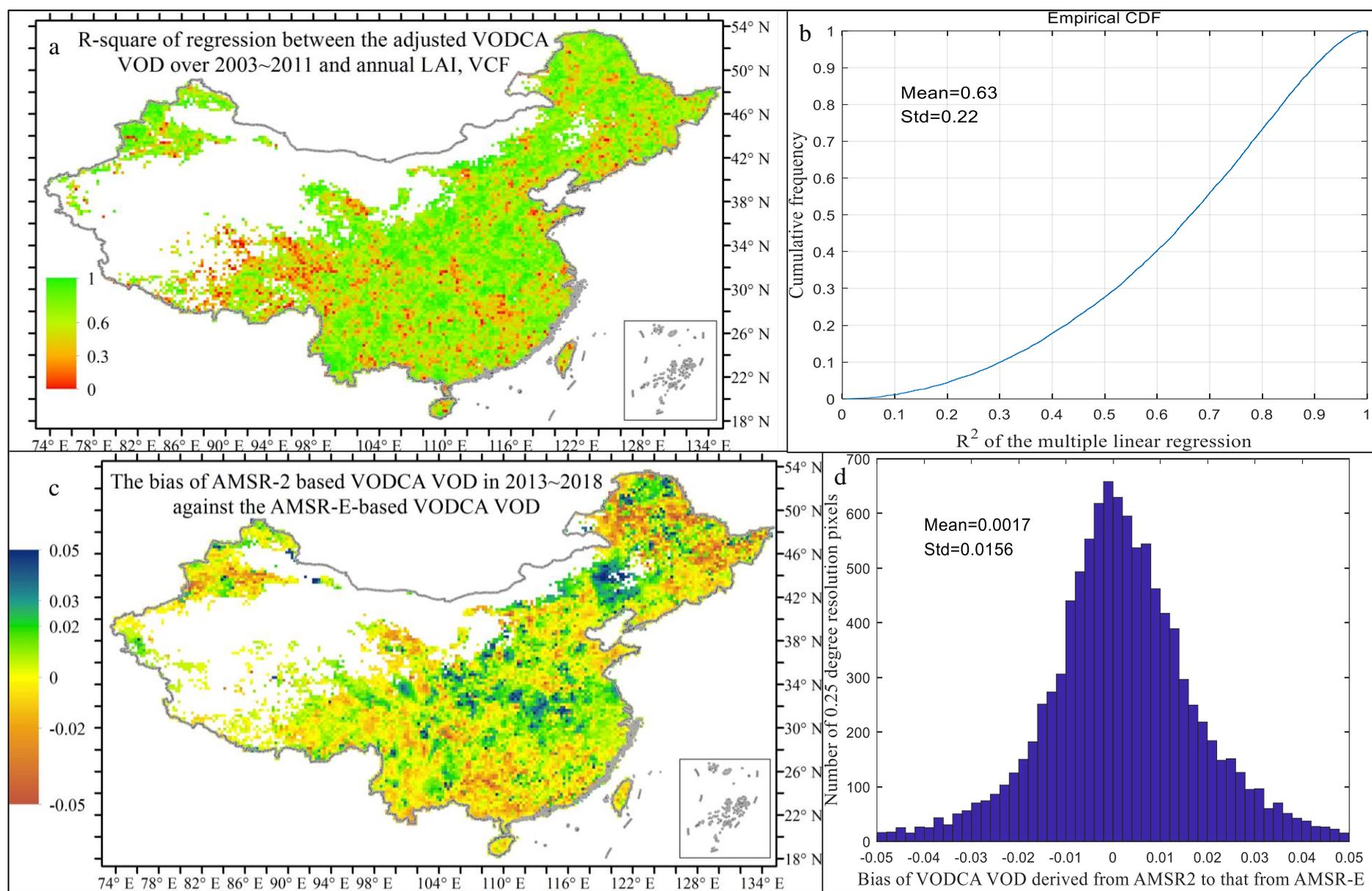
## Additional references

- Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N.-E., Herold, M., and Fritz, S.: Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2015: Globe (V3.0.1), Zenodo [dataset], 125 <http://doi.org/10.5281/zenodo.3939038>, 2020.
- Dimiceli, C., Carroll, M., Sohlberg, R., Kim, D. H., Kelly, M., and Townshend, J. R. G.: MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006, NASA EOSDIS Land Processes DAAC [dataset], 2015.
- Liu, Y. Y., van Dijk, A. I. J. M., de Jeu, R. A. M., Canadell, J. G., McCabe, M. F., Evans, J. P., and Wang, G.: 130 Recent reversal in loss of global terrestrial biomass, *Nat. Clim. Change*, 5, 470-474, <https://doi.org/10.1029/2018EF00089010.1038/nclimate2581>, 2015.
- Luo, Y., Wang, X., Zhang, X., Booth, T. H., and Lu, F.: Root:shoot ratios across China's forests: Forest type and climatic effects, *Forest Ecol. Manag.*, 269, 19-25, <https://doi.org/10.1016/j.foreco.2012.01.005>, 2012.
- Ma, H., Mo, L., Crowther, T. W., Maynard, D. S., van den Hoogen, J., Stocker, B. D., Terrer, C., and Zohner, C. 135 M.: The global distribution and environmental drivers of aboveground versus belowground plant biomass, *Nat. Ecol. Evol.*, <https://doi.org/10.1038/s41559-021-01485-1>, 2021.
- Mokany, K., Raison, R. J., and Prokushkin, A. S.: Critical analysis of root: shoot ratios in terrestrial biomes, *Global Change Biol.*, 12, 84-96, <https://doi.org/10.1111/j.1365-2486.2005.001043.x>, 2006.
- Myneni, R., Knyazikhin, Y., and Park, T.: MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 140 Global 500m SIN Grid V006 NASA EOSDIS Land Processes DAAC [dataset], <https://doi.org/10.5067/MODIS/MCD15A2H.006>, 2015.
- O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., Bindlish, R., and Chaubell, J.: SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 5., NASA National Snow and Ice Data Center Distributed Active Archive Center. [dataset], <https://doi.org/10.5067/4DQ54OUIJ9DL>, 2021.
- 145 Santoro, M. and Cartus, O.: ESA Biomass Climate Change Initiative (Biomass\_cci): Global datasets of forest above-ground biomass for the year 2017, v1. [dataset], <https://doi.org/10.5285/bedc59f37c9545c981a839eb552e4084>, 2019.
- Shangguan, W., Dai, Y., Liu, B., Zhu, A., Duan, Q., Wu, L., Ji, D., Ye, A., Yuan, H., Zhang, Q., Chen, D., Chen, M., Chu, J., Dou, Y., Guo, J., Li, H., Li, J., Liang, L., Liang, X., Liu, H., Liu, S., Miao, C., and Zhang, Y.: A China 150 data set of soil properties for land surface modeling, *J. Adv. Model. Earth Sy.*, 5, 212-224, <https://doi.org/10.1002/jame.20026>, 2013.
- Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., Pongratz, J., Bloom, A. A., Bowman, K., Worden, J., Liu, J., Yin, Y., Domke, G., McRoberts, R. E., Woodall, C., Nabuurs, G.-J., de-Miguel, S., Keller, M., Harris, N., Maxwell, S., and Schimel, D.: Changes in global terrestrial live biomass over the 21st century, *Sci. Adv.*, 7, eabe9829, 155 <https://doi.org/10.1126/sciadv.abe9829>, 2021.

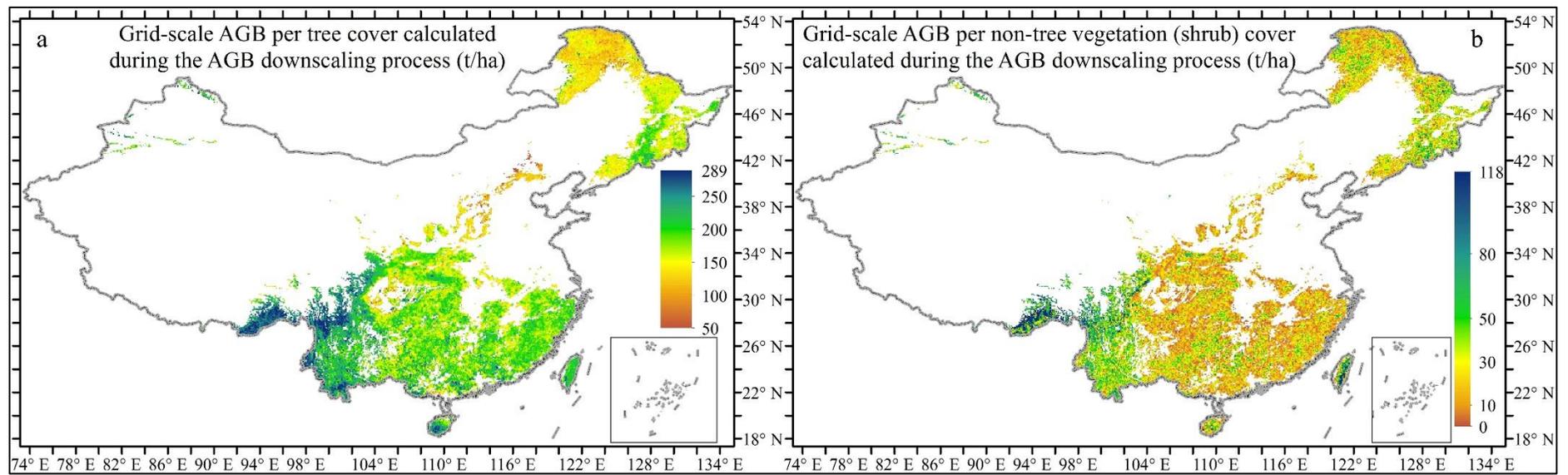
## Supplementary Figures



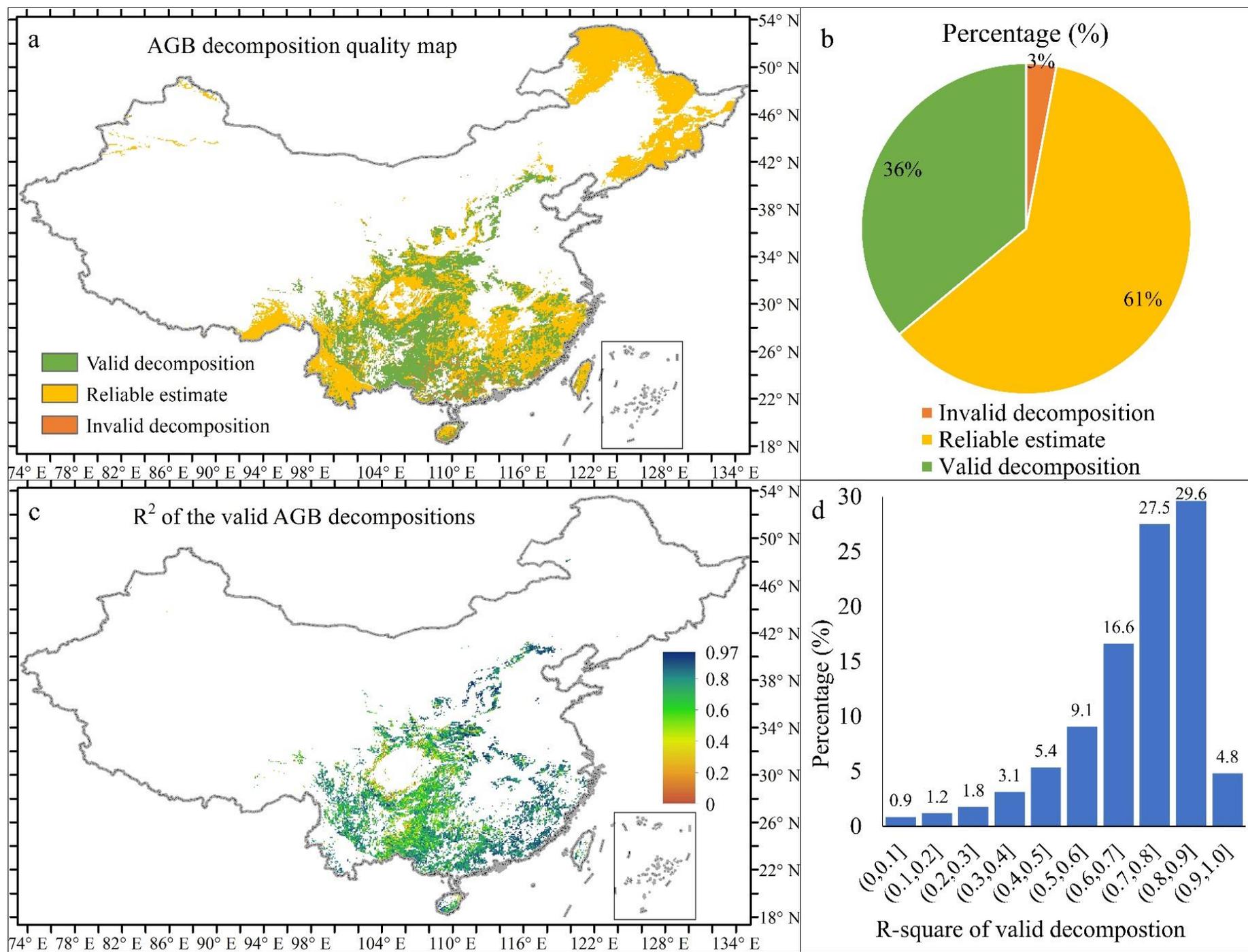
**Figure S1.** Improved aboveground biomass (AGB) mapping in China based on in-situ measurements: (a) the spatial distribution of the collated woodland plots with AGB (or both AGB and BGB) calculated using allometric equations or clear-cutting method. The different uses of data are distinguished by different symbols. ‘RF’ represents random forest model. (b) the overall quality of the improved reference AGB map in China when compared with 1 km pixel AGB upscaled from the data records at forest plots, indicated by R-square ( $R^2$ ) and root mean square error (RMSE) of the linear regression; (c) The improved reference AGB map in China; (d) regional classification map for China.



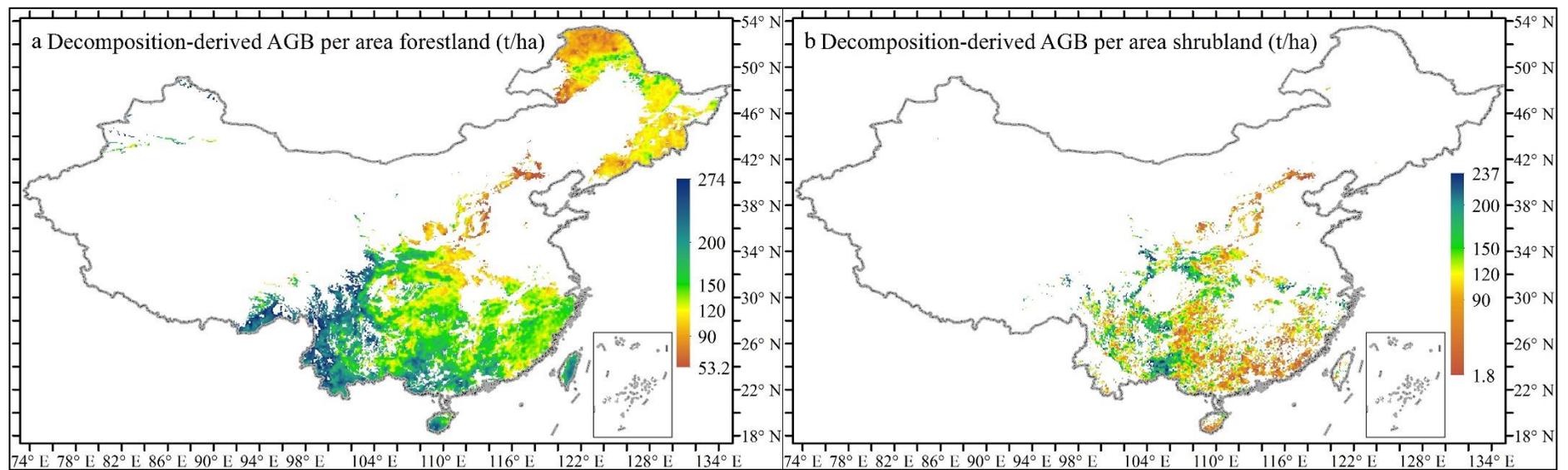
**Figure S2.** Correction of the bias of AMSR2 based VODCA VOD against AMSR-E based VODCA VOD. (a) spatial pattern of the  $R^2$  of regression between AMSR-E based VODCA VOD during 2003~2011 and annual LAI as well as VCF data; (b) the cumulative frequency curve for (a); (c) the map of the calculated bias of AMSR2 based VODCA VOD in 2013~2018 against the AMSR-E based VOD; (d) the histogram for (c).



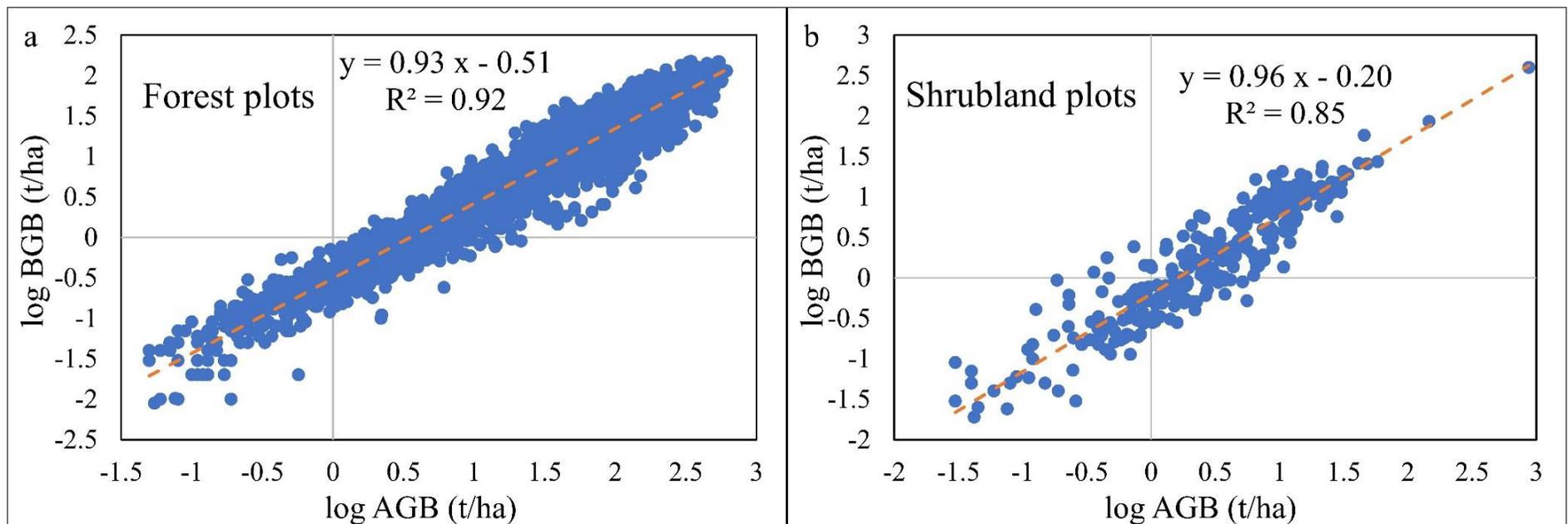
**Figure S3.** The 1/12° grid resolution aboveground biomass (a) per tree cover and (b) AGB per non-tree vegetation (short vegetation) cover derived from the binary linear regression in the AGB downscaling process.



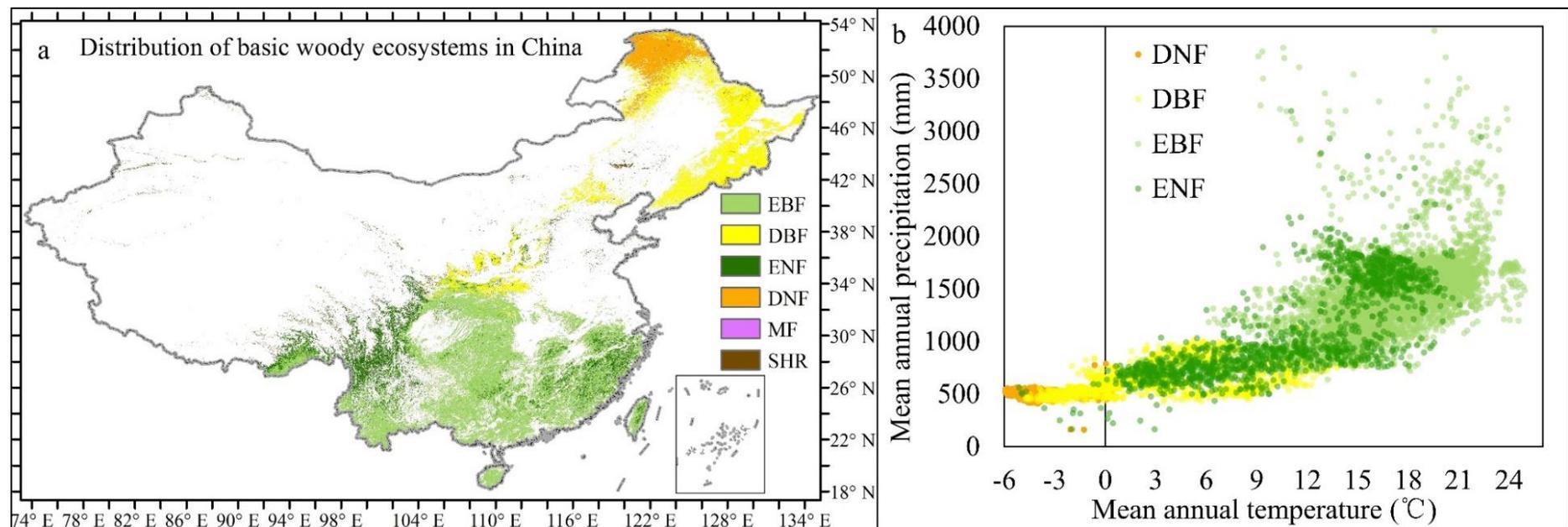
**Figure S4.** Reliability of AGB decomposition in China: (a) the spatial distribution of grid cells where AGB decomposition is invalid, those with reliable estimates of AGB per area forestland or shrubland and those with complete AGB decomposition through valid multiple linear regressions; (b) the pie chart showing the percentages of three types of grid cells in (a); (c) the spatial pattern of the  $R^2$  of the valid multiple linear regressions intended for AGB decomposition; (d) the histogram of the  $R^2$  of the valid binary linear regressions.



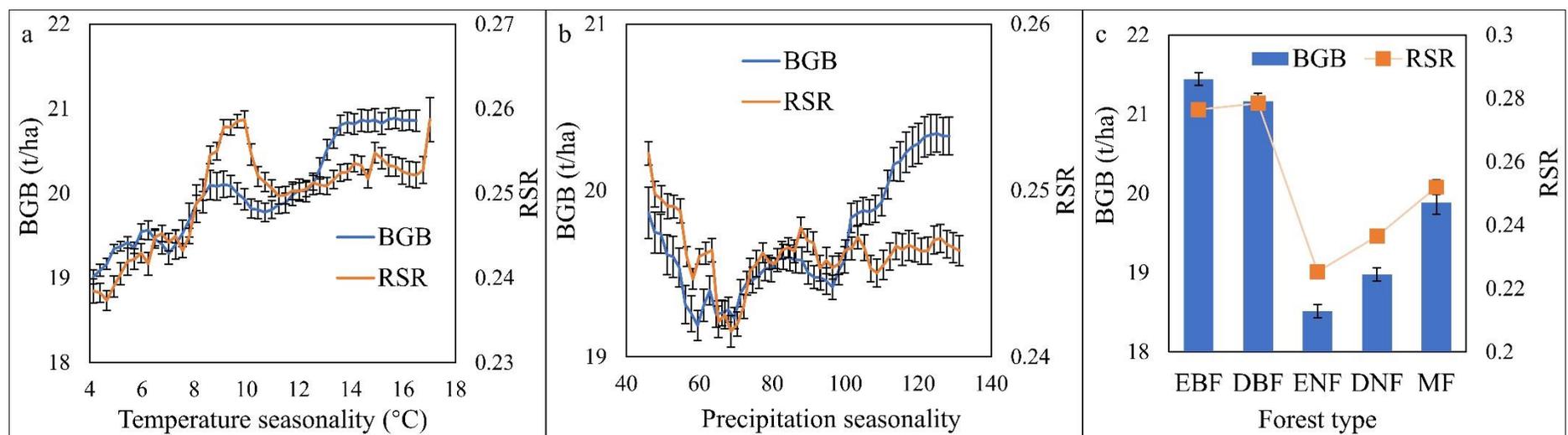
**Figure S5.** Mixed pixel AGB decomposition result in China: (a~b) maps of the decomposition-derived (a) AGB per-area forestland and (b) AGB per-area shrubland in China in around 2017, the nominal year of the Copernicus Global Land Service: Land Cover 100m (CGLS-LC): version 3.



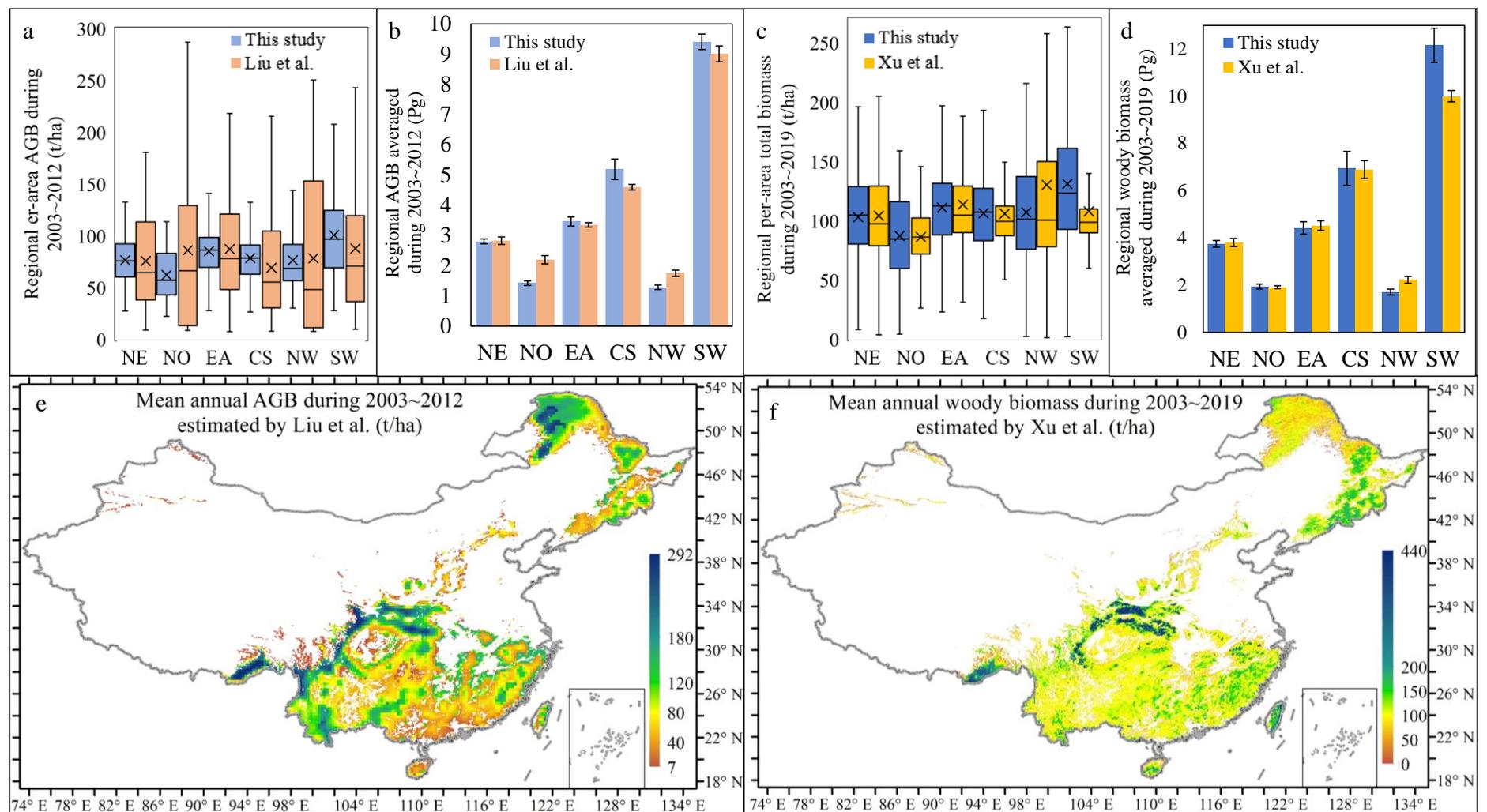
**Figure S6.** Relationship between woody vegetation's BGB and AGB in China: (a) the scatter plot between log (BGB) and log (AGB) in the collected forest plots; (b) the relationship between log (BGB) and log (AGB) in the collected shrubland plots.



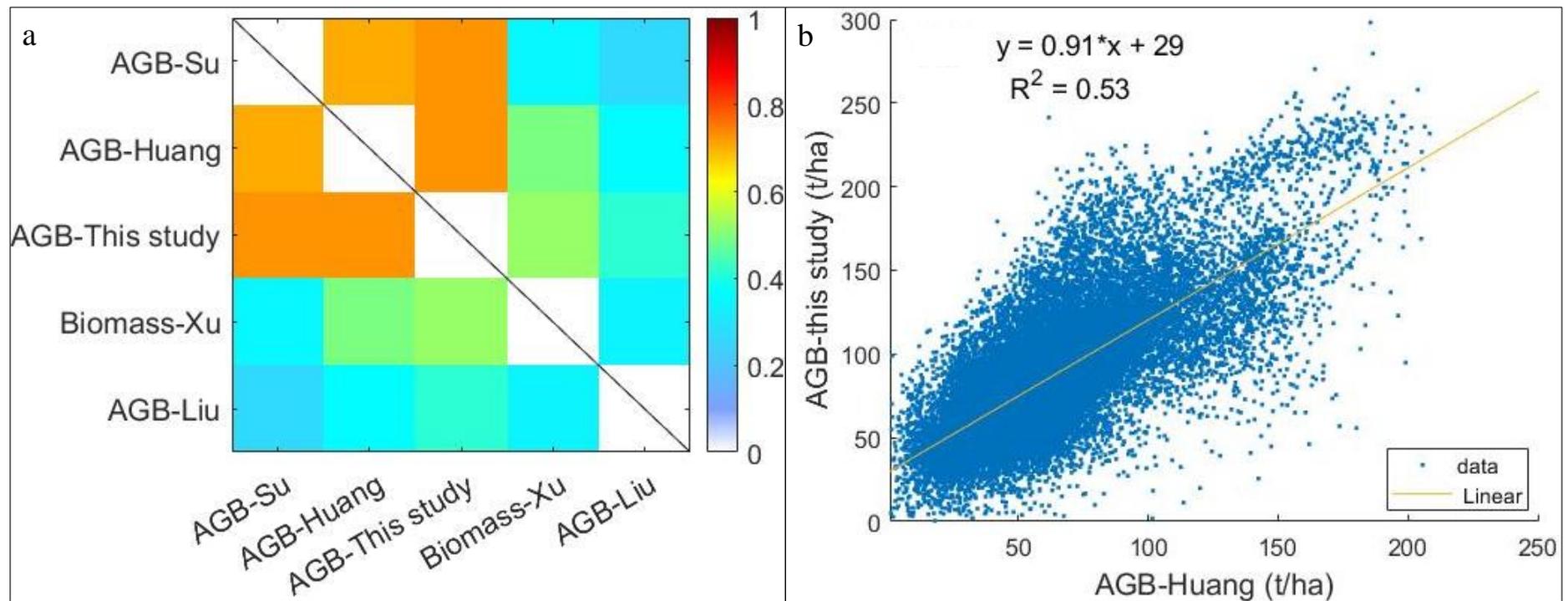
**Figure S7.** Classification of woody vegetation ecosystems in China. (a) the map of the distribution of six different woody ecosystems in China (EBF: evergreen broadleaf forest ecosystem; DBF: deciduous broadleaf forest ecosystem; ENF: evergreen needleleaf forest ecosystem; DNF: deciduous needleleaf forest ecosystem; MF: mixed forest ecosystem; SHR: shrubland ecosystem); (b) the basic climatic backgrounds of four major woody ecosystems in China, indicated by mean annual temperature and mean annual precipitation of 10000 1/120° pixels that are randomly chosen.



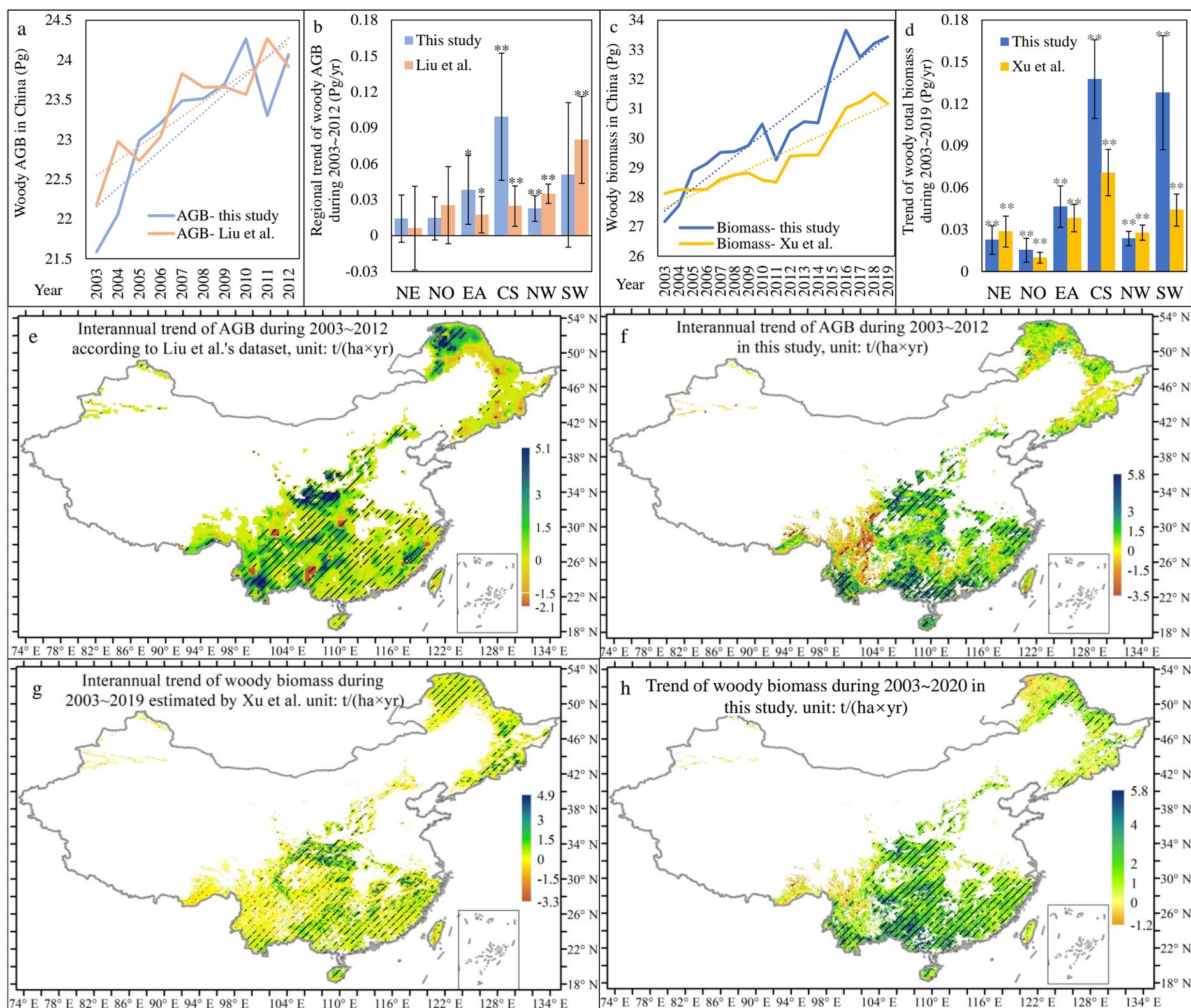
**Figure S8.** The partial dependence plots (PDP) between BGB or RSR and some supplementary predictors: (a) the PDP for temperature seasonality; (b) the PDP for precipitation seasonality; (c) the PDP for forest type (the partial impact of forest type on BGB or RSR). The lines are the mean PDP values, while the error bars are the standard deviation values of the ten-fold trainings.



**Figure S9.** Spatial pattern comparison of the woody biomass calculated in this study against the existing long-term biomass datasets. (a) Comparison of per-area AGB calculated in different regions of China against those estimated by Liu et al. during the common period of 2003~2012. (b) Comparison of total AGB calculated in different regions of China against those estimated by Liu et al. during 2003~2012. NE: north-eastern region, NO: northern region, EA: eastern region, CS: central-south region, NW: north-western region and SW: south-western region. See Figure S1d for the locations of these geographical regions. (c) Comparison of per-area total biomass calculated in different regions of China against those estimated by Xu et al. during the common period of 2003~2019. (d) Comparison of total AGB calculated in different regions against those estimated by Xu et al. during 2003~2019. (e) Average annual woody AGB map in China according to Liu et al. (f) The map of average annual woody biomass according to Xu et al. In subfigures (a) and (c), the bottom, middle and top band of the boxes represent the 25th, 50th (median) and 75th percentiles, and the ends of the whiskers represent the 5th and 95th percentile for all corresponding grid cells, while the mean values are labelled by ‘×’. In subfigures (b) and (d), the error bars represent the standard deviations of annual AGBs and total biomass.

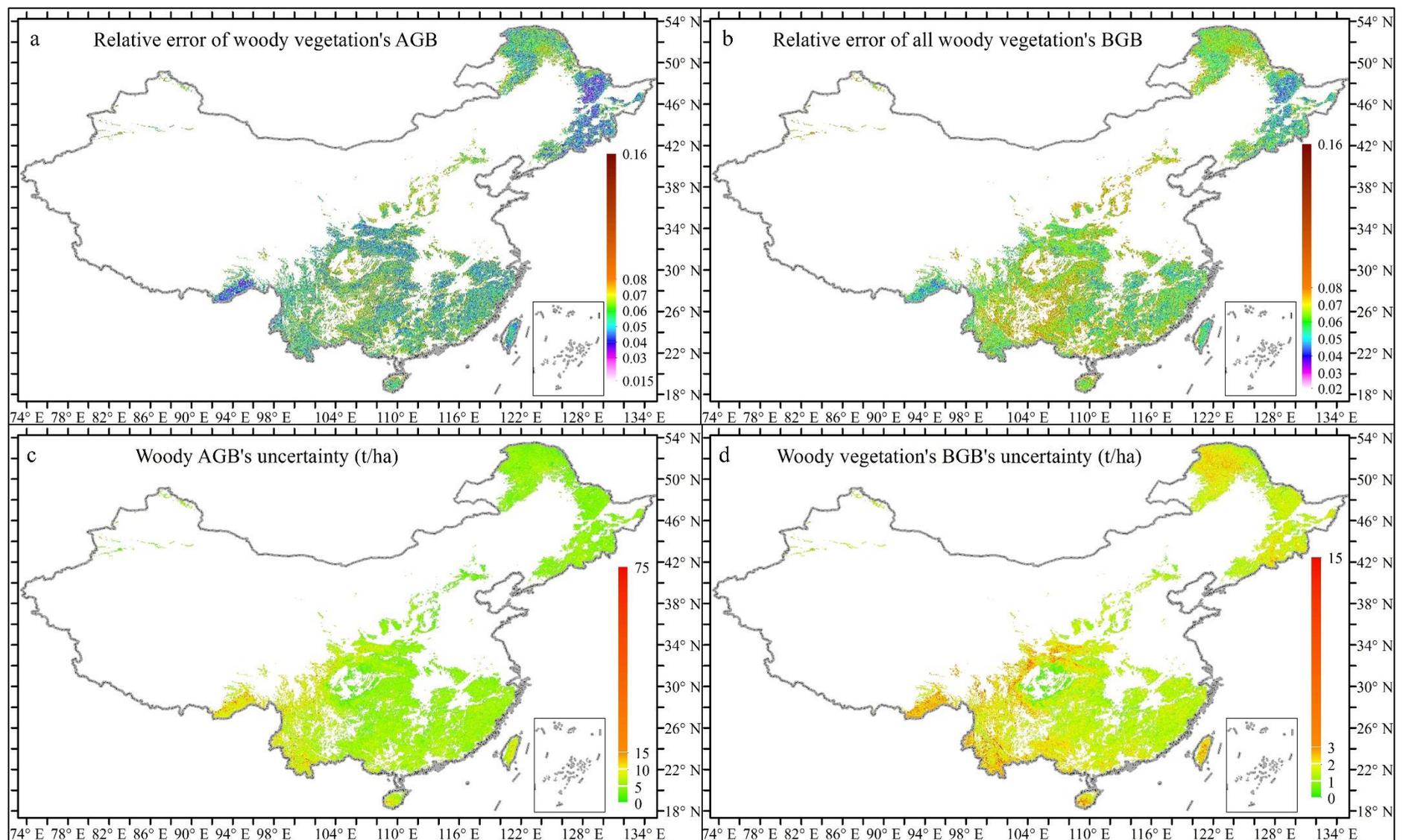


**Figure S10.** Spatial pattern comparison of the mean annual woody AGB in this study (AGB-This study), forest biomass estimated by Xu et al., (Biomass-Xu) and the mean annual AGB estimated by Liu et al. (AGB-Liu) against the high-quality AGB maps for China developed by integrating Lidar, P-band SAR and forest inventory data (i.e., Su et al.'s AGB map and Huang et al.'s AGB map, denoted by AGB-Su and AGB-Huang, respectively). (a) the correlation coefficient matrix among the spatial patterns of five woody biomass datasets; (b) the scatter plot and linear regression between the mean annual AGB calculated in this study and the Chinese forests' AGB dataset developed by Huang et al.

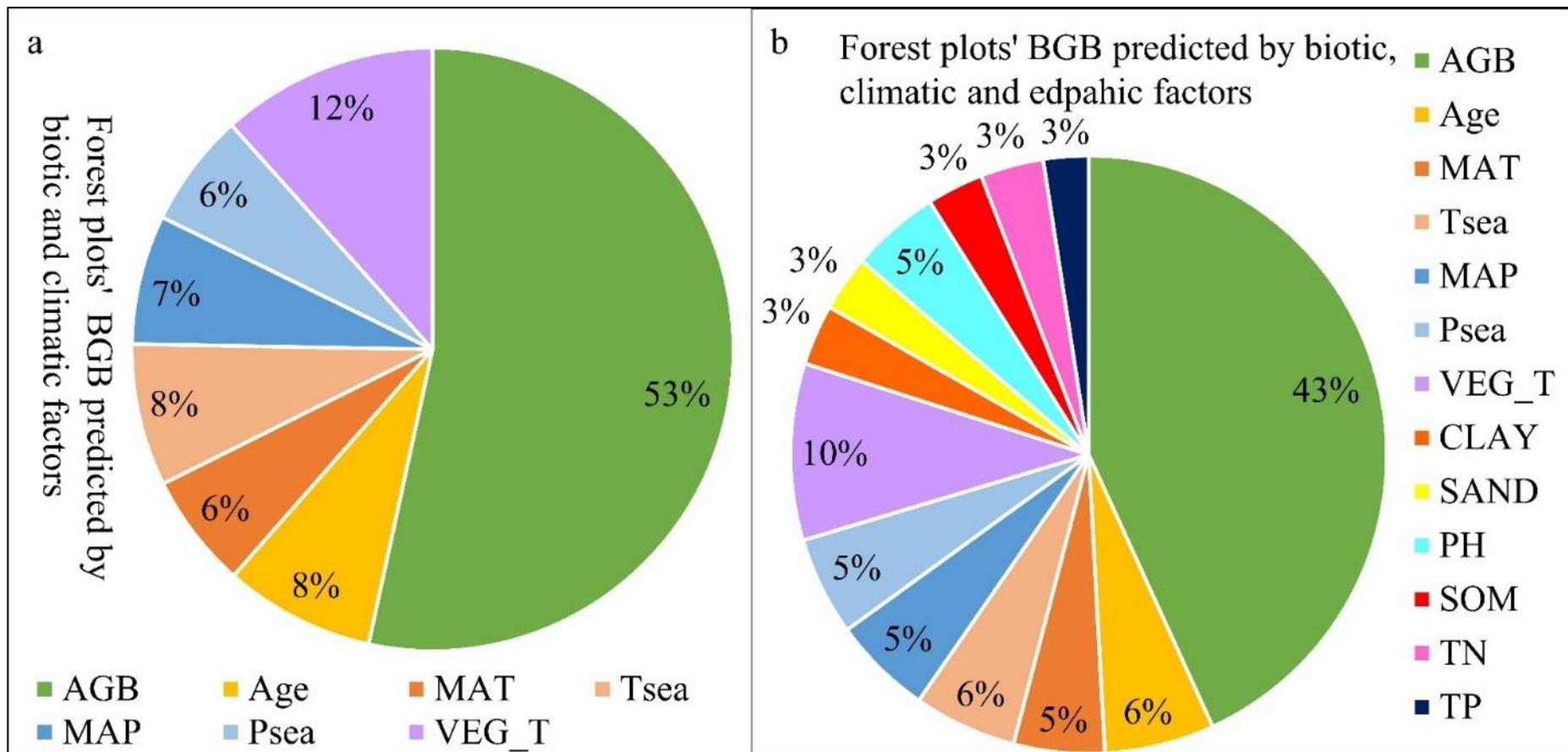


**Figure S11.** Temporal pattern comparison of the woody biomass calculated in this study against the existing long-term biomass datasets: (a) the comparison of the interannual variation of Chinese woody AGB in this study against that estimated by Liu et al. during 2003~2012; (b) the comparison of the interannual trends of woody AGB calculated in different regions of China against those estimated by Liu et al. during the

common period of 2003~2012; (c) the comparison of the interannual variation of Chinese woody biomass in this study against that estimated by Xu et al. during 2003~2019; (d) the comparison of the interannual trends of woody biomass calculated in different regions of China against those estimated by Xu et al. during the common period of 2003~2019; (e) the spatial map of the interannual trends of woody AGB in China during 2003~2012 according to Liu et al.; (f) the map of the trend of woody AGB in China during 2003~2012 according to this study; (g) the map of the interannual trend of woody biomass during 2003~2019 according to Xu et al.; (h) calculated spatial pattern of woody biomass trend in China over 2003~2020. The interannual variations in subfigures (a) and (c) are fitted using linear lines. In subfigures (b) and (d), NE: north-eastern region, NO: northern region, EA: eastern region, CS: central-south region, NW: north-western region and SW: south-western region, see Figure S1d for the regional classification map. The error bars stand for the 95% confidence intervals of the trends, while one and two asterisks indicate statistically significant at the 95% and 99% confidence levels, respectively. The shaded areas in (e~h) represent the regions with significant trends at the 95% confidence level during the corresponding period.



**Figure S12.** The spatial pattern of the uncertainties of woody vegetation's AGB and BGB in China: (a) the relative error of AGB; (b) the relative error of BGB; (c) the uncertainty of AGB and (d) the uncertainty of BGB averaged during 2003~2020.



**Figure S13.** Each predictor's contribution to the simulation efficiency of BGB based on random forest model. (a) The contribution fraction of each climatic or biotic factor when soil property factors are not incorporated; (b) the contribution fractions of all predictors when biotic, climatic, and edaphic factors are all incorporated. Abbreviations to some predictors: Tsea: temperature seasonality; Psea: precipitation seasonality; CLAY: soil clay fraction; SAND: soil sand fraction; PH: soil pH; SOM: soil organic matter content; TN: soil total nitrogen; TP: soil total phosphorus.