Supplementary Material

A

Figures

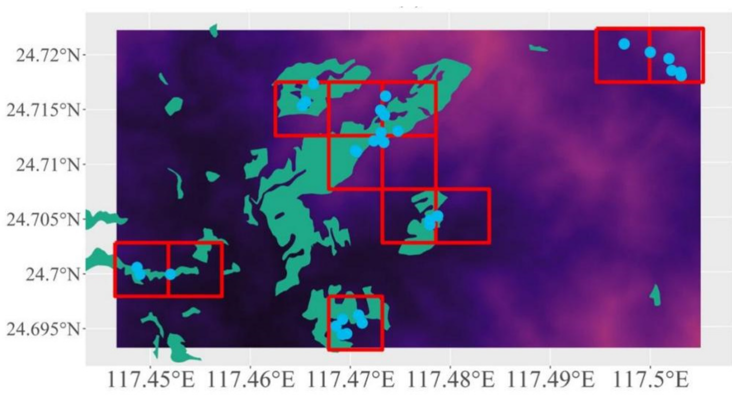

Figure S1. Results of spatial blocking. The 30 blue dots represent 30 sample plots, green patches represent forest patches, and 12 red rectangles represent the results of spatial blocking.

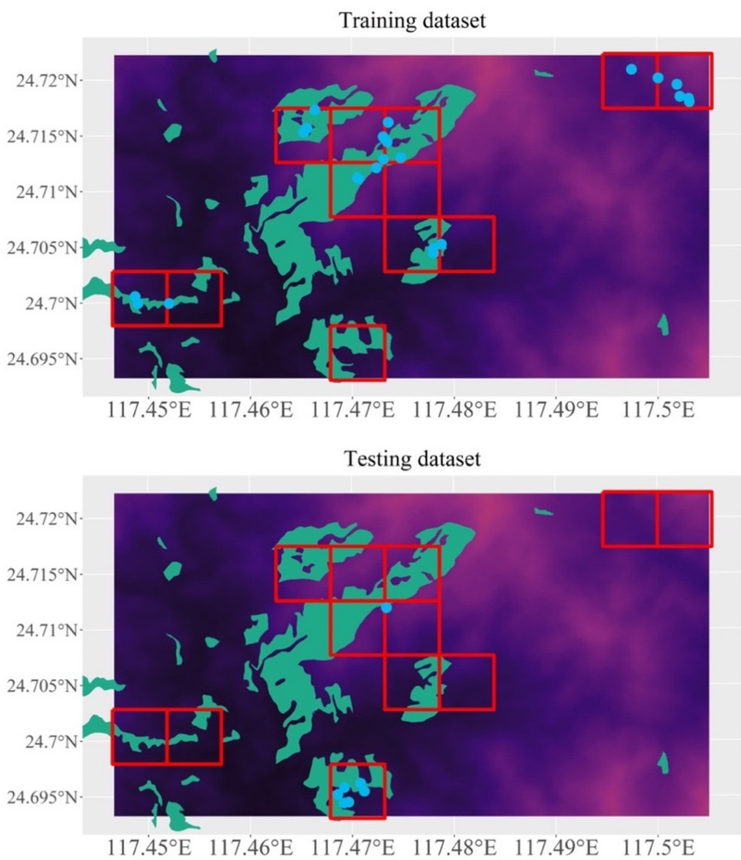

Figure S2. The one situation of spatial block cross-validation.

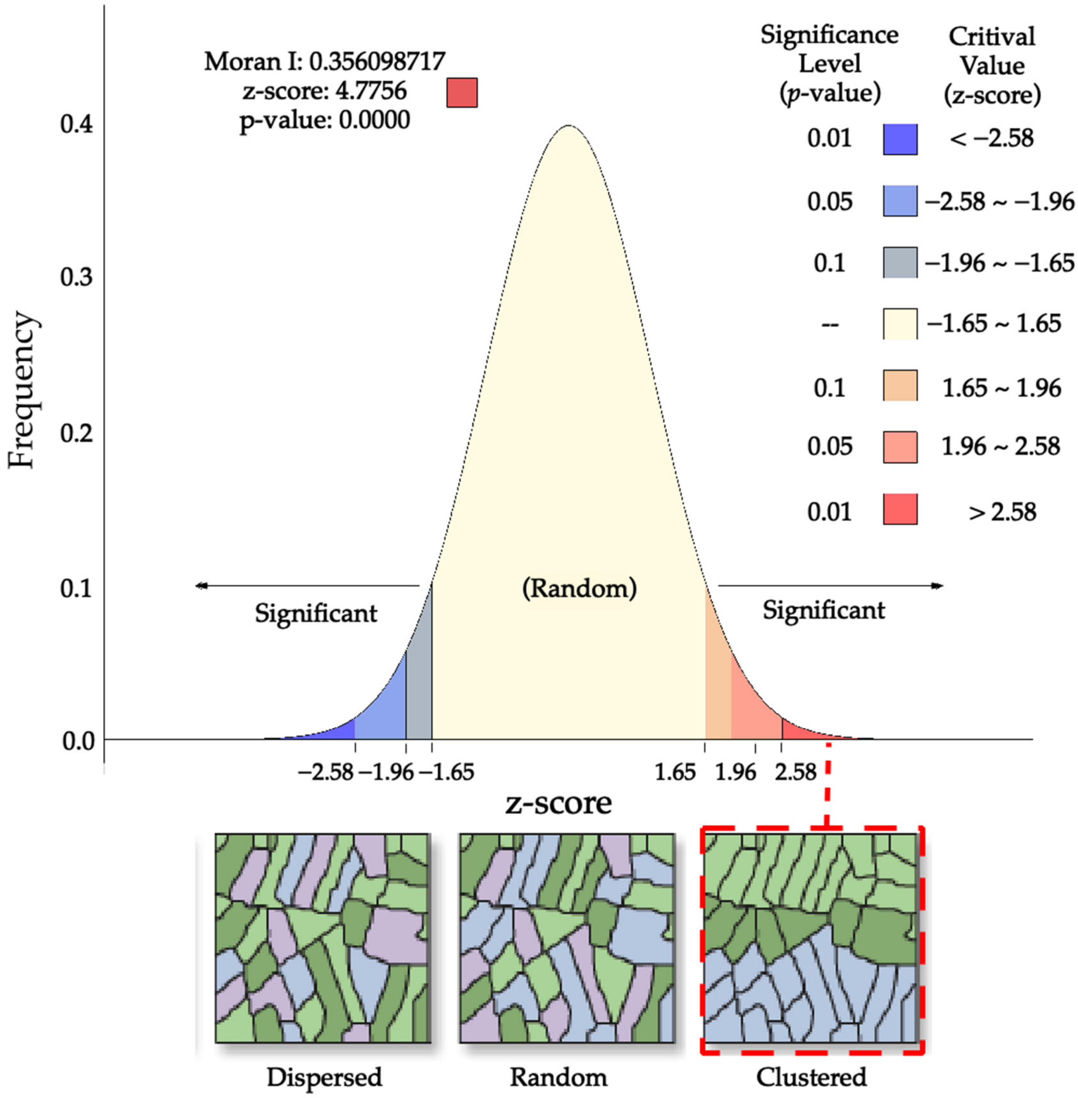

Figure S3. Spatial autocorrelation report.

B

1. Destructive sampling in inventory plots: tree harvest

Trees were harvested from standard woods in the 30 inventory plots. Three trees with a DBH close to the mean DBH of trees in each plot were cut down, for a total of 90 trees harvested from the 30 plots. We then measured the H and DBH of each harvested tree and weighed the biomass of each organ (foliage, stems, and branches) to obtain the AGB of each harvested tree. Tree harvest includes the following three steps: (1.1) plot setup, (1.2) selection and cutting of standard wood, and (1.3) harvested tree measurements.

1.1 Plot setup

The purpose of plot selection is 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. Selected patches met the following six conditions: (1) patch records were available from forest inventory patch 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, 2,977 Eucalyptus patches were selected from the FPI data, and fixed and permanent plots were established. The 2,977 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). We recorded fixed-plot conditions by assigning a code to each fixed plot and recorded environmental conditions, including the following direct and indirect attributes: age, community structure, canopy density, and understory shrub conditions. Finally, a full tree survey was conducted in each fixed plot to obtain the following attributes: DBH, tree height, and other tree attributes.

1.2 Selection and cutting of standard wood

Standard wood was selected following a full tree survey. The following selection criteria were used: (1) wood was located within the plot; stems were representative of the plot with no disturbances (e.g., pests, fire, or anthropogenic activities); and the wood was healthy; and (2) based on the full tree survey data, a tree-sampling method was used to calculate the average basal area, and three trees closest to the average values were selected (i.e., standard trees).

1.3 Harvested tree measurements

This method is similar to the method used by Kankare et al. (2013). For each tree that was cut down, its height, DBH, and AGB were measured. The DBH was measured before the tree was cut down, and the tree height and AGB were measured after the tree was cut down. AGB was measured as follows. (1) All branches (with foliage on them) were removed from the stem and weighed (FW_{b+f}). (2) The stem was sectioned into meter-long pieces using a chainsaw and weighed separately (FW_s). (3) The weight of all branches (with foliage on them) plus the weight of all stems equals the weight of fresh AGB: $AGB = FW_{b+f} + FW_s$. (4) A total of four to six branch samples were systematically sampled from each tree at regular intervals over the entire crown length, and foliage samples were collected from each of the sampled branches. These fresh samples were weighed and used to calculate the proportion of foliage weight to branch plus foliage weight (p) and to calculate the water content of foliage (r_1) and water content of branches (r_2) after drying. (5) A total of six stem samples were systematically sampled from each tree at regular intervals over the entire stem length. These samples were used to calculate the water content of stems (r_3). (6) The fresh weight of each of the tissue samples (stems, branches, and foliage) was weighed in the field, and each tissue sample was placed in plastic bags separately to bring to the laboratory. (7) Fresh samples were oven dried at 85 °C to determine the constant dry weight. (8) According to the fresh weight and constant dry weight of each tissue sample, we calculated their rate of water content (r_1, r_2 and r_3). (9) Tree dry AGB was calculated:

$$AGB = FW_{b+f} \times p \times (1 - r_1) + FW_{b+f} \times (1 - p) \times (1 - r_2) + FW_s \times (1 - r_3)$$

(A.1).

2. Introduction to P-BSHADE

P-BSHADE is an extension of the BSHADE method, which stands for the best linear unbiased estimation (BLUE) model for biased spatial location data (Hu et al. 2013). With the BSHADE model, the spatial correlation and heterogeneity of the target data are added into the model using prior knowledge (such as forest AGB). In addition, through rectification of sample points, the BLUE model can estimate the target subjects. The strategy of the algorithm is to transform the problem into one of solving for the extremum of a multivariate function with constraint conditions, followed by using the Lagrange multiplier method and the overall estimate to acquire the corresponding parameters (Wang et al. 2011) (i.e., each sample in this method is given a certain weight so that the variance between each sample and the true value is minimized to achieve rectification).

Based on the BSHADE method, P-BSHADE is a BLUE-based interpolation method that considers both temporal and spatial heterogeneity. It can use biased samples to deduce the corresponding attributes of regions with missing samples. Therefore, the P-BSHADE model includes the following characteristics and assumptions: (1) the spatial distribution of the target data (such as forest AGB) is heterogeneous and (2) the correlations and differences among the target data in different forests (or sites) are included in the operation of the model (Xu et al. 2013). The performance of the P-BSHADE method has been tested using average annual temperature data in China from 1950 to 2000 (Xu et al. 2017).

The specific mathematical formula for the P-BSHADE model has been described (Hu et al. 2013; Xu et al. 2013).

2.1 Objective

The objective is to interpolate the AGB data of the target sample plot by using data acquired from other sample plots. A theoretical description is

$$\hat{y}_j = \Sigma_{i=1}^{n} w_{ij} y_i \tag{A.3}$$

where $\hat{y}_j$ is the AGB of the jth sample plot estimated by the P-BSHADE model, $(j = 1, \dots n, n = 30)$; y_i is the true AGB of the ith sample plot, $(i = 1, \dots n, n = 30)$; w_{ij} is the weight (contribution) of the true AGB of the ith sample plot to the AGB to be interpolated of the jth sample plot (when $j = 1,\ i = 2, 3, \dots, 30$; when $j = 2,\ i = 1,\ 3,\ 5, \dots$,30), and w_{ij} is calculated by the true AGB of the i-th sample plot and the allometric model estimation of the AGB in the j-th sample plot.

As expected, the estimates of the two properties in Eq. (4) are unbiased:

$$E(y_j) = E(\hat{y}_j) \tag{A.4}$$

Minimum estimation variance is expressed as

$$\min_w \left[\sigma_{\hat{y}_j}^2 = E(\hat{y}_j - y_i)^2 \right] \tag{A.5}$$

where E is the statistical expectation.

2.2 Ratio of data from the target sample plot to those from other sample plots

The ratio between data from the target sample plot to those from other sample plots is one of the most important inputs for estimating the ABG of the target sample plot and is an index of heterogeneity in the AGB spatial distribution. The relationship between data from the target sample plot and from the other sample plots is expressed as

$$b_{ij}Ey_j = Ey_i \tag{A.6}$$

In most cases, the AGB of any two plots are not equal, and the relationship between them can be further expressed as the relative bias b_{ij} between the mathematical expectation of y_j and y_i. Considering Eq. (A.3), Eq. (A.6) can be written as

$$\sum_{i=1}^{n} w_{ij} b_{ij} = 1 \tag{A.7}$$

This equation is generally valid for nonhomogeneous conditions. Clearly, the determination of b_{ij} requires calculating the coefficients w_{ij} $(i = 1, \dots, n, j = 1, \dots, n)$, which is addressed in the following section.

2.3 Weight estimation

The main challenge in estimation is finding the weights w_{ij} that satisfy the unbiased condition and that minimize estimation variance:

$$\sigma_{\hat{y}_j}^2 = E\left(\hat{y}_j - y_i\right)^2 = C\left(\hat{y}_j\hat{y}_j\right) + C(y_i y_i) - 2C\left(\hat{y}_j y_i\right) \tag{A.8}$$

These weights can be calculated by minimizing the estimation variance and taking unbiasedness into account:

$$\begin{bmatrix} C(y_1y_1) & \cdots & C(y_1y_n) & b_{1j} \\ \vdots & \ddots & \vdots & \vdots \\ C(y_ny_1) & \cdots & C(y_ny_n) & b_{nj} \\ b_{1j} & \dots & b_{nj} & 0 \end{bmatrix} \begin{bmatrix} w_{1j} \\ \vdots \\ w_{nj} \\ \mu \end{bmatrix} = \begin{bmatrix} C(y_1y_j) \\ \vdots \\ C(y_ny_j) \\ 1 \end{bmatrix} \quad \text{(A.9)}$$

where μ is a Lagrange multiplier. The minimized variance in the estimation error can then be written as

$$\sigma_y^2 = \sigma_{y_i}^2 + \Sigma_{i=1}^{n}\Sigma_{k=1}^{n}C(y_iy_k) - 2\Sigma_{i=1}^{n}w_{ij}C(y_iy_j) + 2\mu(\Sigma_{i=1}^{n}w_{ij}b_{ij} - 1) \quad \text{(A.10)}$$

References

Hu, M.G., Wang, J.F., Zhao, Y., Jia, L. 2013. A B-SHADE based best linear unbiased estimation tool for biased samples. Environmental Modelling & Software 48(48) 93-97.

Kankare, V., Räty, M., Yu, X., Holopainen, M., Vastaranta, M., Kantola, T., Hyyppä, J., Hyyppä, H., Alho, P. and Viitala, R. 2013. Single tree biomass modelling using airborne laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing, 85, 66-73.

Wang, J.F., Reis, B.Y., Hu, M.G., George, C., Yang, W.Z., Qiao, S., Li, Z.J., Li, X.Z., Lai, S.J., Chen, H.Y., 2011. Area Disease Estimation Based on Sentinel Hospital Records. Plos One 6(8) e23428.

Xu C.D., Wang J.F., Li Q.X. 2017. A New Method for Temperature Spatial Interpolation Based on Sparse Historical Stations. Journal of Climate 31 1757-1770.

Xu, C.D., Wang, J.F., Hu, M.G., Li, Q.X., 2013. Interpolation of missing temperature data at meteorological stations using P-BSHADE*. Journal of Climate 26(19) 7452-7463.