

Supplementary file S2

Tree matching procedure

Candidate matches were classified into three match-quality classes:

1. XY-distance < 2 m and Z-distance < 2 m;
2. XY-distance $\geq$ 2 m and Z-distance < 2 m; and
3. XY-distance $\geq$ 2 m and Z-distance $\geq$ 2 m.

Candidate matches were subsequently filtered by removing all matches with XY-distance $\geq$ 3 m. In addition, matches were removed when the Z-distance was $\geq$ 4 m for trees shorter than 25 m or $\geq$ 5 m for trees at least 25 m tall. These thresholds were selected following graphical inspection of the data and comparison of ALS-derived height maxima with field-measured height maxima at the plot level. The remaining candidate matches were ranked according to match-quality class and XY-distance, with match-quality class taking precedence.

SMOTE parameter selection

Under- and oversampling values ranging from 2 to 24 in steps of one were evaluated, resulting in all possible combinations of these values. The oversampling value specifies how many synthetic samples of the minority class, in this case European aspen, are created, with the total number of the minority class including both the original and synthetic samples. The undersampling value in turn specifies how many samples of the majority class, in this case non-aspen, are included by multiplying the undersampling value with the number of synthetic samples of the minority class.

We used five-fold tree-level cross-validated random forest for all these different degrees of under- and oversampling to select the combination with the best classification performance. To create the five folds, we ordered the European aspens and non-aspens, separately, by volume and repeatedly assigned numbers to the trees in these orders from one to five, corresponding to the folds. We next examined the effects of the different degrees of under- and oversampling on various metrics, namely European aspen's F_1 -score, user's accuracy, and producer's accuracy, as well as Cohen's kappa and overall accuracy. We used Euclidean distance calculations for these five metrics to select the degree of under- and oversampling with the lowest value: the optimum degree of under- and oversampling in our case.

Random forest hyperparameter optimisation

The hyperparameters we tuned were *mtry* ranging from 9 to 29 ($0.1\text{--}0.3 \times$ the number of features, rounded down to the nearest integer values), *minimum node size* ranging from 1 to 10, *sample fraction* ranging from 0.2 to 0.8, and the *number of trees* ranging from 490 to 1470 ($5\text{--}15 \times$ the number of features).