Annual forest aboveground biomass changes mapped using ICESat/GLAS measurements, historical inventory data, and time-series optical and radar imagery for Guangdong province, China

Wenjuan Shen[a,b], Mingshi Li[a,b,*], Chengquan Huang[c], Xin Tao[d], Anshi Wei[e]

[a] College of Forestry, Nanjing Forestry University, Nanjing, 210037, China
[b] Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, 210037, China
[c] Department of Geographical Sciences, University of Maryland, College Park, MD, 20742, USA
[d] Department of Geography, University at Buffalo, Buffalo, NY 14261, USA
[e] Guangdong Provincial Center for Forest Resources Monitoring, Guangzhou, 510173, China

ARTICLE INFO

ABSTRACT

Forest aboveground biomass (AGB) assessments are essential for accurate understanding of carbon accounting under forest disturbance effects and climate change. We mapped AGB data (from 1986 to 2016) by combining the forest inventories and multisource remotely sensed data, including the Ice, Cloud, and Land Elevation Satellite data and Landsat dense time series imagery, and L-band Synthetic Aperture Radar (PALSAR) mosaic data in Guangdong, China. We used random forest (RF) and stochastic gradient boosting (SGB) algorithms to determine the optimal variables of statistical models for mapping and validation of the AGB purpose. Our results showed that the Geoscience Laser Altimeter System (GLAS)-based AGB correlated well ($R^2_{adj} = 0.89$, n = 277, p < 0.001, RMSE = 21.24t/ha) with those obtained using the field-based method that used an RF-based approach, although inevitably, there is a saturation problem. The combined remotely sensed optical and radar imagery and ancillary data sets for mapping AGB using the RF algorithm yielded a stronger ($R^2_{adj} = 0.86$, n = 558, p < 0.001, RMSE = 11.35t/ha) linear correlation with those produced using the GLAS waveform data than that produced using the SGB algorithm. The overall accuracy and Kappa coefficient of mapping forests based on the PALSAR-forest/non-forest Landsat-based phenology for AGB masking were approximately 92.1% and 0.83, respectively. Additionally, the total amount of AGB had increased from 1986 to 2016 by 55.9%. The same increasing trend was observed for total AGB in both mid-subtropical (from 42% to 62%) and south-subtropical (from 38% to 57%) evergreen broadleaved forests, whereas a decreasing trend was witnessed in the tropical forest, particularly after 2010. There was an upward trend of total AGB among the four economic zones of Guangdong; the mountainous area had the highest AGB value distribution, accounting for 58%–70%, followed by the Pearl River Delta region (20%–30%), the western coast of Guangdong (3%–9%), and the eastern coast of Guangdong (2%–7%). The resulting provincial continuous forest AGB maps will provide a better evaluation of carbon dynamic in southern China.

1. Introduction

Accurate forest biomass assessments are important for evaluating forest carbon stocks and terrestrial carbon dynamics (Fang et al., 2001; Houghton et al., 2009; Le Toan et al., 2011; Tian et al., 2017). Traditional forest inventories (Lu, 2006) or field-based allometric equations (Lucas et al., 2006) have been conducted for the calculation of forest aboveground biomass (AGB). A major barrier to quantifying AGB has

---

been identified as being the ground-based inventory that lacks time-liness and spatial coverage (Lu, 2006), especially for large-scale AGB measurements. For this reason, remotely sensed data have been widely applied in detecting forest biomass (Dong et al., 2003; Foody et al., 2003; Lefsky et al., 2002; Lu, 2006; Lu et al., 2016).

Integration of both optical data (e.g., HJ-1, MODIS, and Landsat) and radar data (Synthetic Aperture Radar (SAR); e.g., Phased Array L-band Synthetic Aperture Radar (PALSAR) on the Advanced Land Observing Satellite (ALOS) satellite) captures more useful information from canopy cover and canopy structure (Guo et al., 2010; Shen et al., 2016; Xing et al., 2016). Additionally, full-waveform, large-footprint, satellite-based Geoscience Laser Altimeter System (GLAS) data (Schutz et al., 2005) overcomes both the cost intensiveness of airborne LIDAR (light detection and ranging) data or high-resolution images, and the spatial limitations of field-based techniques. The GLAS-based waveform can be used to estimate large-scale forest structures, such as height and AGB (Chi et al., 2015; Duncanson et al., 2010; Huang et al., 2017; Lefsky et al., 2005; Mitchard et al., 2012; Saatchi et al., 2011; Simard et al., 2011; Sun et al., 2008; Yu et al., 2015).

Many studies have found the value of gaining high-quality field measurements when monitoring and validating AGB (Chen, 2013; Su et al., 2016; Wulder et al., 2008). A direct geolocation linkage between the field inventories and GLAS waveforms is the most straightforward approach to solve the spatially discontinuous GLAS footprints (its footprints with a nominal diameter of 70 m are spaced at 170 m along tracks and tens of kilometers across tracks) (Schutz et al., 2005) and ultimately estimate AGB (Baccini et al., 2012; Chi et al., 2015, 2017; Saatchi et al., 2011; Zhang et al., 2014b). Two major field measurement sources called National Forest Inventory(NFI) (Shen et al., 2016; Zhao et al., 2012) and Forest Management Planning Inventory (FMPI) (Lei et al., 2009; Xie et al., 2011) in China can be applied to estimate GLAS waveform-based AGB, excluding the different resolution sizes of air-borne LIDAR data (Nelson et al., 2017).

Extrapolating GLAS-based biomass estimates to MODIS has been successfully utilized over the past few years (Chi et al., 2015; Hu et al., 2016; Margolis et al., 2015; Su et al., 2016; Zhang et al., 2014b). The disadvantage of this approach is the mismatch in spatial resolutions between MODIS and field measurements, and inability to capture multi-year biomass dynamic changes evaluation. Landsat data of 30 m spatial resolution could match a majority of plots to predict spectrally forest biomass (Lu et al., 2016; Main-Knorn et al., 2013; Powell et al., 2010); for example, a study with two cloudless Landsat data from two different years for GLAS-based AGB extrapolation (Chi et al., 2017). The PALSAR image pixels are similar to the size of Landsat pixels and plot (Yu and Saatchi, 2016). The trend is to use multi-sensor remote-sensing systems integration, including optical, radar, and space-borne LIDAR data (i.e., GLAS), in view of the weakness of a single sensor, which indicates the great potential of GLAS-based forest biomass estimation.

Various approaches could be used to derive GLAS-based AGB from the field plot observations, including stepwise regression, SR; partial least-squares regression, PLSR; ordinary least squares, OLS; support vector regression, SVR; k nearest neighbors, kNN (Duncanson et al., 2010; Guo et al., 2010; Margolis et al., 2015; Zhang et al., 2014b). SR provided the widest range of biomass estimates, but with the greatest uncertainty in overfitting (Chi et al., 2017; Zhang et al., 2014a, b); poor validation performance of SR and PLSR was found (Zhang et al., 2014b). Few studies, however, have explored the random forest (RF) approach in GLAS-derived AGB. RF and stochastic gradient boosting (SGB) are both nonparametric modeling techniques. RF outperformed other regression methods (Coulston et al., 2012), although previous studies looking at individual year continuous models for AGB estimation pointed out that the SGB approach outperformed the RF, or both approaches showed extremely similar performance (Dube et al., 2014; Freeman et al., 2015). There has been little direct comparison of RF and

SGB for continuous GLAS waveform-extrapolated footprint AGB.

The specific objectives of the current study were as follows: (i) constructing RF-based GLAS waveform-derived AGB model from field inventory data; (ii) capturing GLAS waveform- extrapolated footprint AGB to the scene level based on Landsat time-series data (1986–2016) and PALSAR data (2007–2016); (iii) exploring and validating the ability of RF and SGB for quantifying AGB; and (iv) identify and map of annual forests for AGB masking from 1986 to 2016 based on the integration of the PALSAR-based forest/non-forest (FNF) and Landsat-based phenology variables.

2. Materials and data

2.1. The study area

Guangdong Province ($17.97 \times 10^4 \, \text{km}^2$) is located at 20°13′N to 25°31′N and from 109°39′E to 117°19′E (Fig. 1). Most of the areas have a mid-subtropical or south-subtropical monsoon climate–southern coastal region being a tropical monsoon climate. Its mean annual pre-cipitation and temperature ranges from 1300 to 2500 mm and 19 to 24 °C, respectively. According to the Chinese climatic zones and geo-graphic characteristics of forests (Ren et al., 2013), from north to south, Guangdong Province is divided into three forest zones (Fig. 1(b)): mid-subtropical typical evergreen broadleaved forest region, south-sub-tropical monsoon evergreen broadleaved forest region, and tropical monsoon forest or rainforest region (Zhou et al., 2017). Guangdong has a wet season from April to September and a dry season from November to January. February, March, and October are transition months (Wang et al., 2009). During these months, the wetness conditions can differ substantially across the province. For example, in March and April, the northern region is often wet, but the south is dry. In September, the pattern is reversed. According to the status of economic development, Guangdong Province has four economic regions, including the Pearl River Delta economic zone (PRD), the east coast economic zone (EasternGD), the west coast economic zone (WesternGD), and the northern Guangdong mountainous zone (Mountainous). Northern Guangdong is located in the valley basin and is surrounded by mountains, the PRD mainly consists of hilly lands and plains, the WesternGD is a large platform area, and the EasternGD is mostly plains.

2.2. Data acquisition and preprocessing

2.2.1. Forest inventory plots

China has a three-tiered inventory system: the first level is the NFI, the second level is the FMPI, and the third level is the Forest Operation Design Inventory (FODI) (Xie et al., 2011). A subcompartment or forest stand in FMPI is a contiguous trees region that is quite homogeneous or contains a bunch of forest features, such as species, age, site quality, average stand height, average stand diameter at breast height, and stem volume. Two years (2002 and 2007) of NFI data for the entire province and five years (2005–2009) of subcompartment data ("xiaoban" (XB)) from five cities (Guangzhou, Heyuan, Huizhou, Qingyuan, and Shao-guan) located within the study area were available for use in this study. NFI AGB (25.82 m × 25.82 m in size) was calculated from the dominant trees AGB by using the allometric equations (Shen et al., 2016). Forest stand (XB) (a range between 0.006 ha and 68 ha in size) AGB was de-rived from stand volume based on a biomass expansion factor devel-oped by the Guangdong Provincial Center for Forest Resources Mon-itoring. GLAS footprints were overlaid onto the subcompartments and NFI data to derive the field inventory-based AGB (Fig. 2). The overall data processing and analysis workflow was summarized in Fig. 3.

2.2.2. ICESat GLAS data

The GLAS's laser footprint on the surface is elliptical, with an

Fig. 1. Overview of the study area (a), including climatic zone-based forest zones and Landsat footprints intersecting with PALSAR forest/non-forest map at the spatial resolution of 70 m for 2015 in Guangdong (b), the location and numbers of all available GLAS footprints divided by four economic zones (c), and the location, numbers, and sources of relating the field plots AGB to GLAS footprints (d).

equivalent circular area of 70 m in diameter, and the space between footprints is 172 m (Schutz et al., 2005). A detailed description of GLAS can be found in Schutz et al. (2005) and Harding and Carabajal (2005). Table 1 lists GLAS waveform parameters that are sensitive to forest biomass resulting from GLA14 data. The relative heights (rh10 to rh100) are frequently used as AGB estimation indexes (Hayashi et al., 2015).

2.2.3. Landsat-based data

We used more than 1000 radiometrically and geometrically corrected Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper plus (ETM+), and Landsat 8 Operational Land Imager (OLI)

images (30 m) collected from the USGS/EROS Landsat archive, spanning from 1986 to 2016 (Fig. 4). Guangdong province is covered by 12 Landsat path/row tiles (Fig. 1(b)). Our goal was to use the highest quality cloud-free images; however, about 60 selected images had substantial (≤50% cloud cover) cloud contamination because of the lack of better images. Multiple images were composited to remove cloud and phenology effects (Shen et al., 2016). Landsat 5/7 and Landsat 8 images were radiometrically and atmospherically calibrated to surface reflectance based on the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm (Masek et al., 2006) and the Landsat 8 surface reflectance code (LaSRC) (Vogelmann et al., 2017), respectively. ETM+ scan-line-off (SLC) gap filling was conducted based on the gdal_fillnodata tool, which uses an inverse distance weighting (IDW) interpolation by interpolating the missing values at gaps using existing values (Alexandridis et al., 2013).

2.2.4. ALOS PALSAR-based data

This study used 25 m resolution ALOS PASAR (HH, HV polarizations) mosaic image data sets (off-Nadir 34.3°) from 2007–2010 and 2015–2016; they have been corrected and normalized geometrically and radiometrically according to topography (Shimada et al., 2014). We processed the PALSAR data by converting digital number to the backscatter coefficient in decibels for the HH and HV polarization components (Shimada et al., 2014). The enhanced Lee filter with a window size of 5×5 pixels was implemented to reduce speckles. The variables derived incorporated HH, HV polarizations, HH/HV (ratio), and HH-HV (difference). The PALSAR data were re-projected with the GLAS data for matching 70 m spatial resolution.

Fig. 2. Boxplot of AGB derived by relating field plots to GLAS footprints showing the mean (star), median (thick black line), and the lower and upper 25 percentiles (thin lines), and maximum and minimum values (black point).
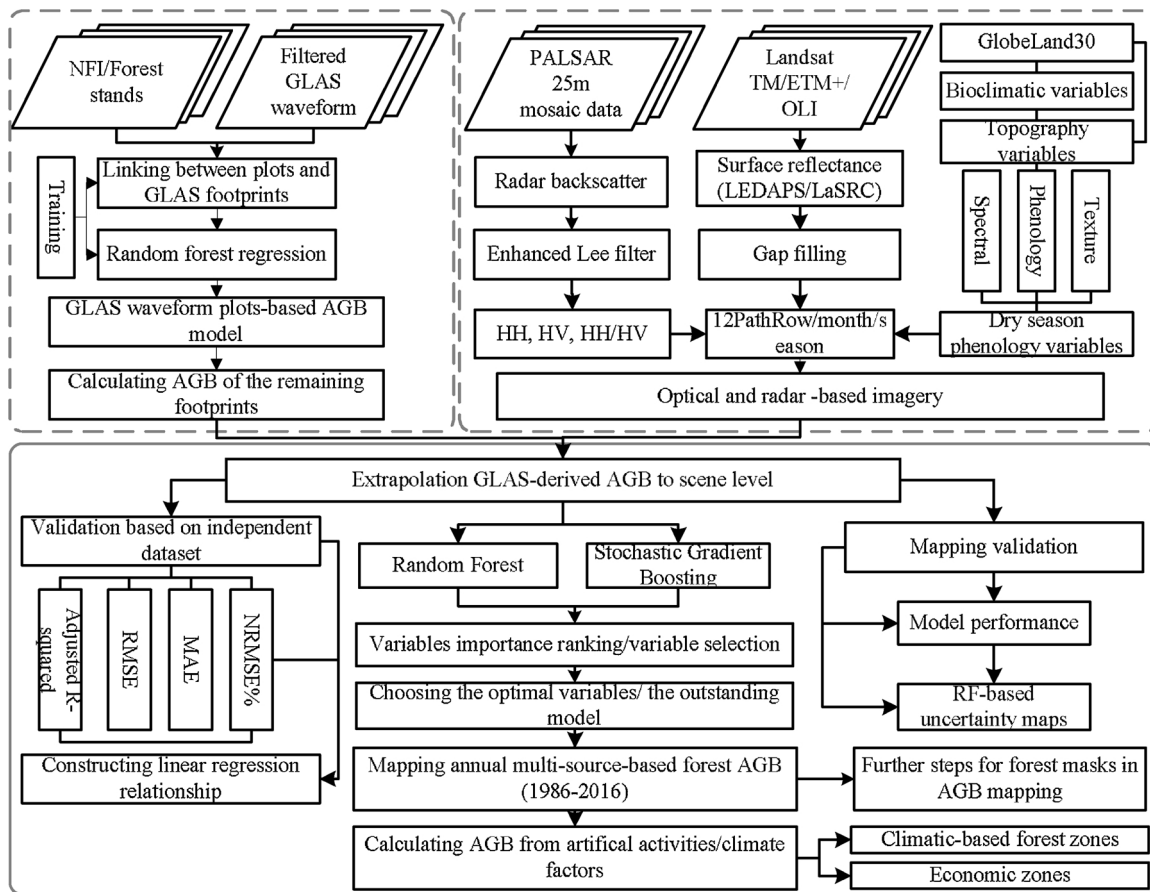
Fig. 3. Workflow diagram.

2.2.5. Auxiliary data

Topography variables (elevation and slope) from the NASA Shuttle Radar Topographic Mission (SRTM) digital elevation model (DEM) data were re-projected from 30 m resolution to 70 m spatial resolution with nearest neighbor resampling. WorldClim-based bioclimatic variables (annual mean temperature, annual temperature seasonality, annual precipitation, and annual precipitation seasonality, 1 km resolution) were prepared (Hijmans et al., 2005). This study used global land cover (China's 30 m GlobeLand30, GLC) product with 10 classes of mapping for the years 2000 and 2010 (Chen et al., 2015). These data were re-sampled from different resolutions to 70 m spatial resolution with nearest neighbor interpolation to make them consistent with 70 m in diameter of the GLAS footprint.

3. Methodology

3.1. Linking field-based AGB to GLAS waveform parameters

GLAS data acquired during several periods from 2003 to 2009 over

Guangdong was used in the research. To identify and eliminate spurious GLAS observations due to cloud contamination, terrain effects, and other factors, we removed GLAS footprints whose slopes were greater than 10° or were located over a non-forest area based on the MODIS land cover product (Los et al., 2012). After removal, this study used a total of 2788 GLAS footprints (Fig. 1(c)), of which 553 could be linked with field plots (Figs. 1(d) and 2). We used GLAS waveform-based parameters (Table 1) and plot-based AGB data as the input data in a random forest regression model to build a GLAS waveform parameters-based AGB prediction model (50% for training, 276). We then used GLAS waveform-based AGB models to calculate the remaining footprints.

3.2. Stochastic models

RF and SGB are two ensemble modeling techniques where the forest consists of a large number of regression trees. An observation passes down the tree through a series of splits, or nodes, at which a decision is made as to which direction to proceed based on values of the

Table 1
The commonly used GLAS waveform parameters for predicting forest AGB.

| Parameters	Definition
rh25, rh50, rh75, rh98	Quartiles heights for waveform energy to reach 25%, 50%,75%, and 98% of total energy starting from the signal end
gamp1, gamp2,…, gamp6	Amplitude of Gaussian peaks in G14 data (amplitude of up to six Gaussians)
gcnt1, gcnt2,…, gcnt6	Centroid range increment for up to six peaks (centroid of the Gaussian fit to the last peak)
gsig1, gsig2,…, gsig6	Sigma of Gaussian fits in G14 product (sigma of up to six Gaussians) standard deviation of the fit
pct_cover	The ratio of canopy return energy to the total energy of the waveform
ht	Top tree height (the distance between the signal beginning and the last Gaussian peak) from waveform
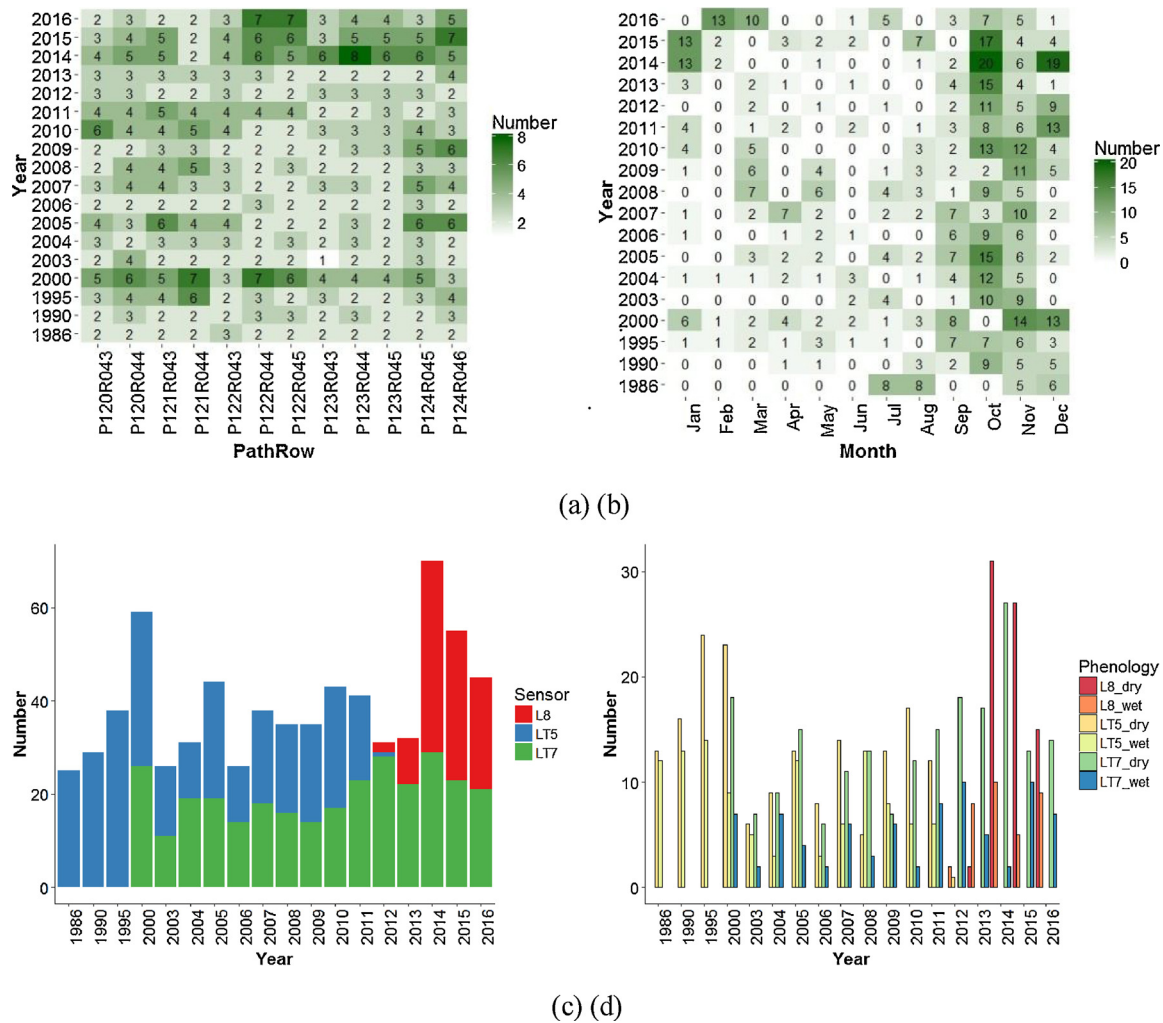
Fig. 4. Statistics of the number of Landsat images used by (a)12 paths/rows, (b) 12 months, (c) 3 sensors, and (d) dry season and wet season.

explanatory variables. Ultimately, a terminal node, or leaf, is reached and a predicted response is given, which is typically the mean of observations in the terminal node for a continuous response (Freeman et al., 2015). We used an RF and SGB regression tree implemented via the ModelMap package (Freeman, 2009) in R (R Development Core Team, 2008), respectively, to (1) identify important predictor variables in the optimal approach, (2) model the relationship between the GLAS waveform parameters and AGB, and (3) apply the model to the remaining footprints for deriving AGB as well as generating the GLAS waveform-extrapolated footprint AGB.

3.2.1. The implementation of RF modeling

Two tuning parameters, such as ntrees and mtry, were required to implement RF (Freeman et al., 2015). "ntrees" controls the total number of independent trees, or the number of trees required to stabilize variable importance and variable interaction (Breiman, 2001; Liaw and Matthew, 2002). "mtry" controls the number of predictor variables randomly sampled to determine each split (Freeman et al., 2015). Here, the RF model had 500 ntrees and a mtry default calculation (the square root of the total number of predictor variables) used for modeling, in which each tree was built from one half of the training data's random sampling (Breiman, 2001). The higher the percent increase in mean square error (MSE) (PercentIncMSE) and increase in NodePurity (IncNodePurity), the stronger the importance of these predictor variables (Freeman et al., 2015).

3.2.2. The implementation of SGB

Several tuning parameters (shrinkage, bagging fraction, interaction depth, and number of trees) were required to implement SGB (Freeman et al., 2015). Like RF, SGB also provides variable importance measures (Friedman, 2001). An advantage of SGB is that it is not necessary to preselect or transform predictor variables. Unlike RF, in SGB, the use of too many trees may lead to overfitting and affect model performance. To help improve model accuracy, model-fitting parameters were selected (shrinkage: 0.01; n.trees: 2500; v.fold: 1; interaction.depth: 10; bag.fraction: 0.5; n.minobsinnode: 10) in the current analysis.

Variable selection is based on "variable selection using random forests" (VSURF) (Genuer et al., 2015). The most important objects are varselect.thres (VI mean, VI standard deviation), varselect.interp (OOB error), and varselect.pred (OOB error), which contain the set of variables selected after the thresholding, interpretation, and prediction steps, respectively (Genuer et al., 2015). The less important predictor variables were then removed, and the optimal predictor variables were selected to produce the final model and predict AGB based on seven years' field observations-based RF and SGB model. The gbm package (Ridgeway, 2007) calculates the relative influence of the reduction of squared error attributable to each variable (decrease in squared error) (Friedman, 2001, 2002). In this study, we compared the importance ranking between the RF model and the SGB model based on the optimal predictor variables and derived the best model for AGB mapping (Freeman et al., 2015).

Table 2
Validation data set of the imagery extrapolation.

Year		2003	2004	2005	2006	2007	2008	2009
No. of pixels		85	119	75	74	94	91	20
Field plots	Obs. AGB(t/ha)	10.9–175.3	4.1–344.0	3.0–337.9	3.2–289.7	3.6–256.6	3.9–371.9	26.1–247.1
	Obs. Mean (t/ha)	47.0	48.5	50.1	47.6	46.3	53.9	55.6
RF (reduced)	Pred. AGB (t/ha)	28.9–130.3	25.5–171.6	19.6–186.4	19.8–182.9	20.7–161.6	20.4–209.2	33.9–164.5
	Pred. Mean (t/ha)	50.2	51.8	51.8	49.1	47.8	54.1	54.6
SGB + RF (reduced)	Pred. AGB (t/ha)	25.1–139.2	26.5–151.3	28.0–142.1	27.0–137.6	28.0–149.8	25.9–142.5	34.8–106.7
	Pred. Mean (t/ha)	55.2	52.6	65.0	54.4	65.4	65.3	52.2
SGB	Pred. AGB (t/ha)	52.3–114.3	78.3–119.8	53.4–113.8	59.3–118.1	55.2–115.0	54.8–111.2	56.8–86.1
(reduced)	Pred. Mean (t/ha)	86.9	92.6	83.6	85.2	85.2	83.3	83.3

3.3. Extrapolating GLAS-derived AGB estimates to Landsat and PALSAR data

Similar to the GLAS waveform parameters modeling for AGB, after AGB was obtained for footprints, we explored RF and SGB regression models to extrapolate the AGB from the GLAS footprints to the scene level based on Landsat, PALSAR and ancillary data sets-based predictor variables (Appendix Table A1). Since GLAS was not a mapping data, we produced an annual AGB map with 70 m resolution covering the entire study area (4/5 samples were for training).

3.3.1. Explanatory variables based on image data

Numerous spectral, phenology, climate, topography, land cover type, texture measurements, and PALSAR polarizations were used for AGB assessment (Appendix Table A1). Spectral indices and tasseled cap transformations (R package RStoolbox, http://bleutner.github.io/RStoolbox) were acquired from seven years of the best quality wet season/dry season Landsat imagery that matched with field observations acquisition years. Dry or wet season normalized difference vegetation index (NDVI) time series are defined as consisting of near monthly NDVI images from September or October of the previous year, to March or April of the current year, and March or April to August or September of the current year, respectively (Gessner et al., 2013), based on the feature of local dry and wet climate (Fig. 4(d), Appendix Table A1). The phenology variables, including maximum, mean, median, minimum, range, and standard deviation from time-series NDVI data in the dry season, were then derived to distinguish forest from crops and grasses based on photosynthetic activity

(Karlson et al., 2015). Only images during the dry season were used because most of them had little to no cloud cover. We used the gray level co-occurrence matrix approach (3 × 3 pixels, offset ([1,1]), and a 32 gray level quantization) implemented in the R package glcm (Zvoleff, 2015) to produce eight texture variables from the six Landsat bands. HH, HV, and HH/HV were also included in the textural analysis.

3.3.2. Model assessment and validation

We conducted the reduced predictor variables data sets for the GLAS waveform AGB estimation model by using RF modeling. Both RF and SGB modeling were for a provincial scale AGB extrapolation. Regarding the GLAS footprints, we assessed the predictive accuracy of the model based on validation data sets (1/2, 277/553) by OOB predictions on the training data. Regarding extrapolation, the remaining 20% of the samples (558) were reserved for imagery extrapolation validation, and the independent validation data sets approach was based on the data set in Table 2. We calculated the adjusted R-squared (R^2_{adj}), the root mean square error (RMSE), the normalized root-mean-square error (NRMSE %), and the mean absolute error (MAE) to assess the model performance. In this study, we validated the linear regression relationships between the predicted data based on AGB extrapolation results from RF and SGB and the observed data from GLAS waveform predictions. RF-based mapping quality was also determined by uncertainty maps derived from calculating the mean, standard deviation, and the coefficient of variation (CV) for each pixel from the predictions of each of the independent randomly generated trees from the RF model (Freeman et al., 2015). The uncertainty was highest when the modeled trees were
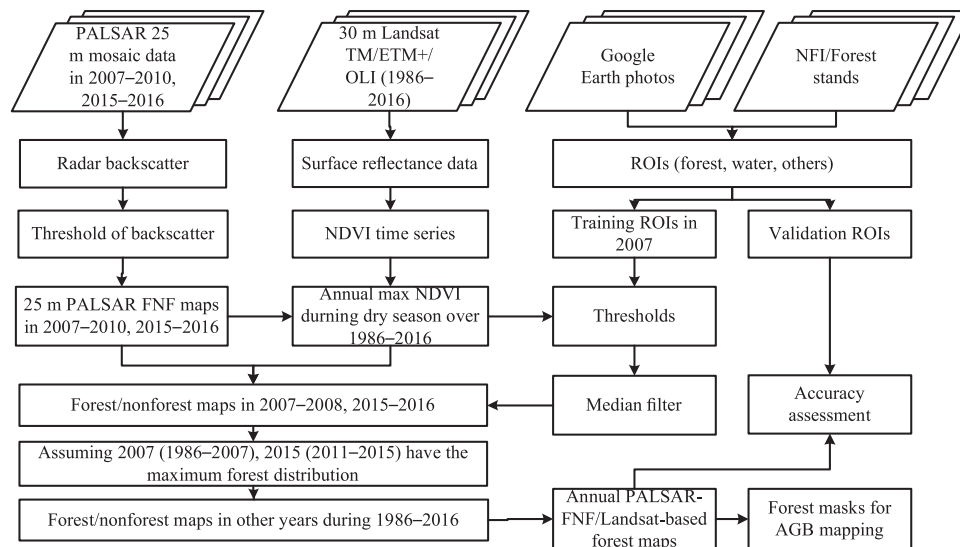

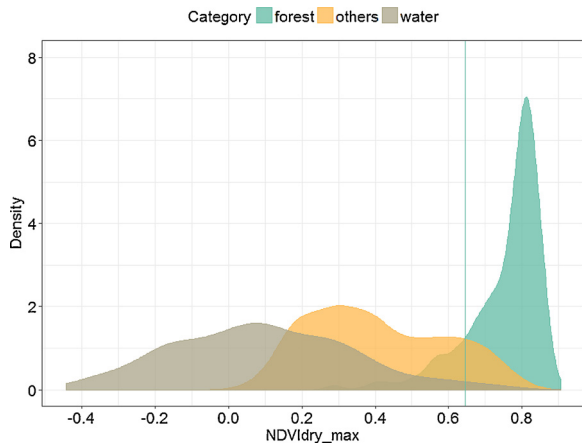**Fig. 5.** Workflow for PALSAR-FNF/Landsat-based forest mapping to mask AGB.

Fig. 6. Density plots of forest/non-forest (water/others) in 2007 PALSAR in the maximum NDVI during the dry season.

not in agreement; for example, some trees predicting low AGB when others predicted high (Freeman et al., 2015). In SGB, the trees were not independent, thus, the SGB uncertainty map was not included.

3.4. Forest distribution mapping based on Landsat and PALSAR

3.4.1. PALSAR-forest/non-forest(FNF)/Landsat-based forest mapping

The global FNF map was generated by using the region-dependent threshold of backscatter with an accuracy of more than 84% (http://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/fnf_index.htm). Here, forest was defined as a unit of land (> 0.5 ha) with forest cover over 10% (FAO, 2012). The FNF image data sets produced from global 25 m PALSAR mosaic products from 2007 to 2010, 2015 (Fig. 1(b)), and 2016 (Fig. 5) were used. Then, all of the FNF maps were resampled to 70 m to match AGB spatial resolution.

We developed a pixel-based approach to produce annual forest maps for AGB masking from 1986 to 2016, using the integration of the PALSAR-based FNF and Landsat images. For example, the highest NDVI value was used to further differentiate PALSAR-based forests from other land cover types (Fig. 6). Here, cloud-free wet season Landsat images were limited, and most of the tree species were considered evergreen. The maximum NDVI value of dry season Landsat images was $NDVI_{max}$.

The $NDVI_{max}$ Gaussian kernel density (Silverman,1986) for training ROIs (2030) of the forest (750 polygons, 7,801,180 pixels), other types (short for others) (912 polygons, 24,777 pixels), and water (368 polygons, 54,615 pixels) was plotted (Fig. 6). Finally, the threshold (0.645) of $NDVI_{max}$ (solid line) was used to separate 80% of the forest pixels from 99% of the others' pixels, and 95% of the forest pixels from 99% of the water pixels. To reduce speckle noise effects (e.g., salt-and-pepper noise), we subsequently used a median filter (window size, 5×5). The 2007 and 2015 PALSAR data were used to detect forest maps before 2007, and from 2011 to 2015, respectively, because there were no long time series PALSAR data. The hypothesis that the 2007 and 2015 PALSAR data produced the maximum forest area based on the national forestry yearbook of China (8.27×10^6 ha, 9.06×10^6 ha), respectively, during 1986–2007 and 2011–2015, has been adopted. We used the PALSAR-FNF/Landsat-based forest mapping algorithm to generate annual forest maps in 1986–2016 (Fig. 5).

3.4.2. Validation and accuracy assessment of forest maps

We assessed the PALSAR-FNF/Landsat-based forest maps using a confusion matrix based on the validation plots. We acquired 507 plots of forests and non-forests based on these materials (Google Earth, NFI, and subcompartment data in 2007) (Fig. 5).

3.5. Mapping forest biomass

The PALSAR-FNF/Landsat-based annual forest maps provided baseline forest masks to overlay onto the annual AGB maps to extract the AGB patterns just for forest pixels (Fig. 5). All of the AGB prediction maps in different paths/rows were composited after they were re-projected from the corresponding Universal Transverse Mercator projection to the latitude/longitude coordinates. The final annual AGB maps were clipped by the Guangdong Province administrative boundary.

4. Results

4.1. AGB estimates from GLAS waveform parameters

Variable selection based on random forests gives a plot of the results in four graphs showing numbers of selected variables (Fig. 7). First, the four variables with negative importance (variables gcnt1, gamp6, gsig5 and gsig6) based on VI mean and VI standard deviation were eliminated. Then, the model with five variables based on OOB error, which
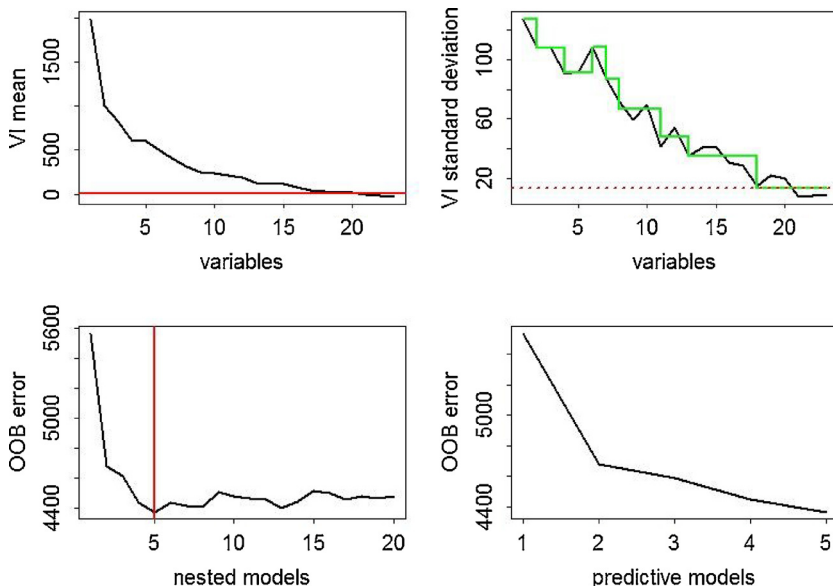

Fig. 7. Variable selection based on random forests for the GLAS waveform AGB data. The top graphs illustrate the thresholding for removing negative importance variables based on VI mean and VI standard deviation; bottom left and bottom right graphs are associated with interpretation and prediction showing the number of selected variables based on OOB error, respectively.
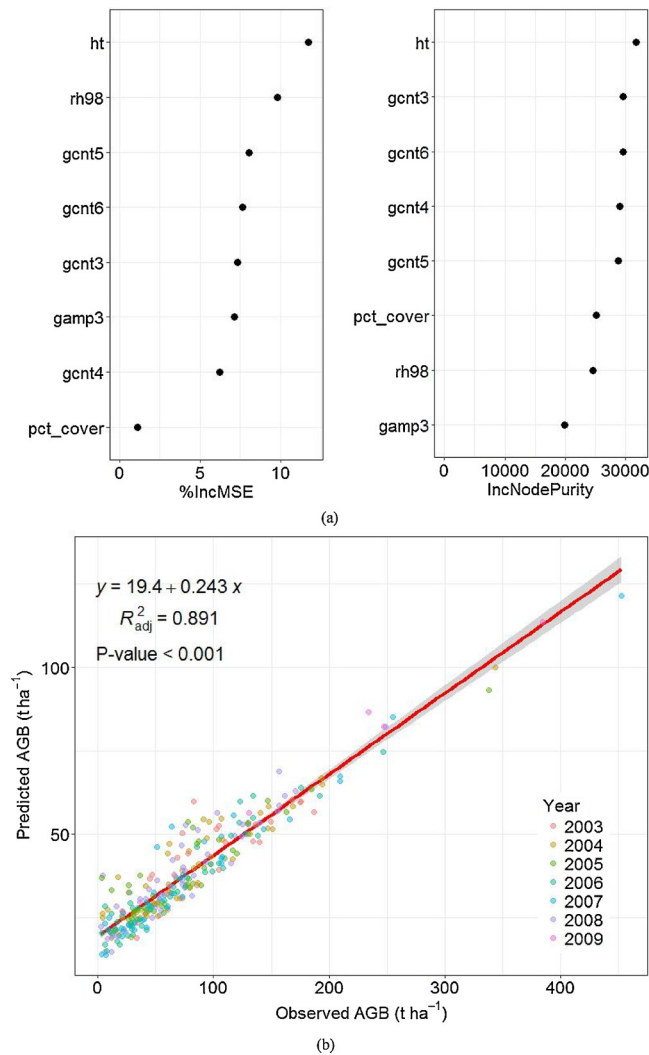
Fig. 8. Variable importance ranking based on OOB error. The left and right graphs illustrate the percent increase in mean square error(MSE) (PercentIncMSE) and increase in NodePurity (IncNodePurity) with the importance of theses predictor variables, respectively (a). The linear relationship of the RF-based GLAS waveform parameters predicted AGB and the field measurements-based observed AGB, the red solid line is the prediction curve, the gray solid line indicates the 95% confidence intervals (CI lines) and the shaded confidence area for the predictions (b).

contained all of the most important variables, was selected. Also, three variables (gcnt4, gamp3, and pct_cover) were used for the GLAS waveform AGB mapping. Ultimately, the reduced variables for the prediction of GLAS waveform AGB included the following: ht, rh98, gcnt5, gcnt3, gcnt6, pct-cover, gcnt4, and gamp3 (Fig. 8(a)). The GLAS waveform-derived AGB correlated well with ground-based AGB

($R^2_{adj} = 0.89$, RMSE = 21.24 t/ha, NRMSE = 4.83%, MAE = 42.77 t/ha, p-value < 0.001; Fig. 8(b)).

4.2. PALSAR-FNF/Landsat forest map in Guangdong

Table 3 presents the accuracy assessment of the PALSAR-FNF/Landsat forest maps with ground truth based regions of interest (ROI). We used a total of 209 (1,560,236 pixels) ground truth forest polygon ROIs and 298 (53,065 pixels) non-forest ROIs in 2007 for validation. The overall accuracy, user accuracy, producer accuracy and Kappa coefficient for the forest classification was 92.1%, 99.4%, 81.3%, and 0.83 (95% confidence interval (CI): 0.8941, 0.943), respectively. The assessment results demonstrated that the FNF/Landsat forest map had reasonably good accuracy.

4.3. Assessment of AGB retrieval performance

Table 4 shows the prediction statistics results from the selected variables according to RF and SGB. We ranked the predictor variables on the basis of the PercentIncMSE and IncNodePurity. Through removing the least important predictor variables based on VSURF function, OOB error was reduced, and 15 variables still remained. The derived reduced variables from a test set with OOB predictions (Fig. 9(a)) included the following: DEM, annual mean temperature (AMT), annual precipitation (AP), temperature seasonality (TS), precipitation seasonality (PS), PALSAR HV, slope, EVI, NDMI, the mean and median of dry season NDVI, TCW, TCG, Landsat SWIR, and SWIR2. Combining the imagery and other geospatial data estimated AGB yielded a strong correlation with the GLAS waveform-extrapolated footprint AGB ($R^2_{adj} = 0.858$, RMSE = 11.35 t/ha, NRMSE = 3.22%, MAE = 9.02 t/ha, p-value < 0.001) (Fig. 9 (c)). To make comparisons with RF, we selected 15 predictor variables in SGB1. Fig. 9(b) indicates that TS, AP, ndvi_dryrange, slope, elevation, HH, AMT, SWIR1, NIR, HV, SAVI, NDMI, blue, TCW, and TCB from the entire data sets are ranked high in predicting AGB, according to the variable relative importance determined by the SGB algorithm. The reduced variables yielded an R^2_{adj} of −0.0001 and RMSE of 30.1 t/ha, NRMSE of 9.42%, MAE of 43.92 t/ha, and p-value > 0.001 (Fig. 9(d)). Meanwhile, the comparison of relative variable importance in RF and SGB2 based on the same reduced variables was presented in Fig. 10. Based on these variables used in RF, the AGB estimation of SGB2 was slightly improved ($R^2_{adj} = 0.03$, RMSE = 29.63 t/ha, NRMSE = 8.54%, MAE = 25.48 t/ha, p-value < 0.001) compared with that of the SGB1 model using the optimal variables (Fig. 9(e)).

One pixel per plot for 2003–2009 was prepared for validation, producing 20–120 reference pixels for predicting AGB (Table 2). The distribution of the observed and the predicted AGB were shown in Table 2. The SGB model produced the least accurate results among all of the methods (Fig. 9). Overall, the RF model was superior to the other models according to the validation R^2_{adj} and RMSE. Without regard to some prediction errors, RF-based reduced variables could bring the most accurate AGB predictions and maps.

The annual AGB maps with 70 m resolution were created by using

Table 3
The accuracy assessment of forest classification in 2007.

	Class	Ground truth (pixels)		Total Classified Pixels	User Accuracy (%)	Commission Error (%)
		F	NF			
Classification	F	1,269,092	178	1,269,270	99.4%	0.58%
	NF	291,144	52,887	344,031	88.4%	11.6%
Total ground-truth pixels		1,560,236	53,065	1,613,301		
Producer accuracy (%)		81.3%	99.7%		Overall accuracy = 92.1%	
Omission error (%)		18.66%	0.34%		Kappa coefficient = 0.83	

Table 4
Comparisons of AGB assessment results using RF and SGB.

Statistical methods	Input data set	No. of variables (Reduced)	R^2_{adj}	RMSE (t/ha)
RF	Elevation, ndvi_drymedian, ndvi_drymean, AMT, AP, TS, PS, HV, slope, EVI, SWIR1, SWIR2, NDMI, TCW, TCG	15	0.858	11.35
SGB1	TS, AP, ndvi_dryrange, slope, elevation, HH, AMT, SWIR1, NIR, HV, SAVI, NDMI, Blue, TCW, TCB	15	− 0.0001	30.10
SGB2	Elevation, ndvi_drymedian, ndvi_drymean, AMT, AP, TS, PS, HV, slope, EVI, SWIR1, SWIR2, NDMI, TCW, TCG	15	0.0307	29.63

the optimal RF model; for example, the 2010 AGB is shown in Fig. 11. These figures illustrate the RF-predicted AGB pattern and its corresponding uncertainty. We assume that global land cover(GLC) covering Guangdong in 2010 represents the ground truth, as seen in Google Earth photos (Fig. 11). When the CV value was high, the mean of AGB was low, and the map of RF uncertainty (Fig. 11) highlighted small anomalous areas of high uncertainty (Freeman et al., 2015; Shen et al., 2016). We must be very cautious of those locations with high CV values when using the predicted AGB maps because some trees predict low AGB and others predict higher AGB.

4.4. Forest AGB estimation in Guangdong Province

We used the GLAS-based biomass estimates to produce annual provincial AGB maps based on multitemporal Landsat and PALSAR imagery using the RF algorithm. Thus, accurate estimates of AGB changes in Guangdong Province from 1986 to 2016 were obtained. The Gaussian kernel density plot of spatial AGB from 1986 to 2016 are shown in Fig. 12, an AGB value range of 50–80 t/ha accounts for more than 50% of forest regions. The estimated total AGB and other statistics in Guangdong forests from 1986 to 2016 are summarized in Fig. 13(a); the regional mean AGB had a slight decreasing from 66.084 (t/ha) to 60.312 (t/ha) by − 8.7%, and the maximum AGB density ranged from 184.5 to 224.84 (t/ha), the total AGB across Guangdong markedly increased, from 1.54×10^8 t (1986) to 3.50×10^8 t (2016), an increase of 55.9%.

In terms of the climatic zone-based forest zones (Fig. 13(b)), a similar increasing trend was observed for total AGB in both the mid-subtropical and south-subtropical, that is, an increase from 0.89×10^8 t and 0.65×10^8 t in 1986 to 1.82×10^8 t and 1.68×10^8 t in 2016, respectively, by 51.2% and 61.1%, accounting for 42–62% and 38–57% of the total AGB. In the tropical area, AGB decreased sharply from 110,000 t in 1986 to 1526 t in 2016, particularly after 2010. In terms of the economic zones (Fig. 13(c)), there was an upward tendency among these four economic zones. Mountainous regions had the highest AGB range of values–the lowest value being of 1.08×10^8 t (1986) and the highest value being 3.05×10^8 t (2014). Mountainous accounted for 58–70% of the total AGB, followed by the PRD region (20–30%), western coast of Guangdong (3–9%), and eastern coast of Guangdong (2–7%). The estimated total AGB were from 0.32×10^8 t to 1.1×10^8 t in PRD, 0.42×10^7 t to 2.8×10^7 t on the eastern coast of Guangdong, and 0.5×10^7 t to 4.0×10^7 t on the western coast of Guangdong. Moreover, the total AGB trend on the western coast of Guangdong and PRD closely corresponded with the trend of the mountainous area, which had a tendency of increasing relatively slowly to the maximum, then going down, exhibiting differences from the eastern coast of Guangdong. The derived AGB values in the aforementioned three regions increased rapidly from 1986–1990, and then decreased with slight fluctuations until 2008, especially in the mountainous area, and followed by a gradual increase up to 2014, unlike the eastern coast of Guangdong, which dropped until 1990. An increase then occurred until 2009, with another decline in 2008, and finally, a sharp decline after 2014.

5. Discussion

5.1. GLAS-based waveform predictions validation

We prepared GLAS waveform parameters as input parameters in the GLAS-based AGB estimation. After the variable importance estimation, the optimal variables were determined. We found that ht (top tree height from waveform) and rh98 (quartiles heights for waveform energy to reach 98% of total energy starting from signal end) have a relatively strong relationship with GLAS-based AGB for evergreen broadleaved forest and rainforest, which agreed with previous research that suggested that relative heights (h75 or h100) were positively correlated with AGB estimates (Chi et al., 2017; Guo et al., 2010). Rh98 (similar to h100) was useful for estimating AGB in Guangdong.

The model-based approach can be applied in regions of enough numbers of field observations to GLAS footprints, providing an opportunity to estimate AGB in homogeneous areas where field inventories are sparse (Healey et al., 2012). Chi et al. (2015) reported that the logarithm function models were explained well by the adjusted R^2 range of 0.62 to 0.76 in the needleleaf, broadleaf, and mixed forests in the subtropical and tropical zones, including Guangdong. Zhang et al. (2014a,b) used SR, PLSR, and SVR for GLAS-based AGB estimation at field-measured points and reported an R^2 of 0.82 and an RMSE of 32.69 Mg/ha for the final chosen bootstrapping SVR in northeastern China. The adjusted R^2 of our study reached 0.89 (RMSE = 21.24 t/ha, NRMSE = 4.83%). Obviously, the results from the RF model are superior to the above-mentioned final performances. However, there is a saturation problem, because the RF model overestimated low values and underestimated high values (Fig. 8(b)). Generally, the RF method can reduce errors while making more accurate GLAS-based AGB predictions.

5.2. Making comparisons between forest AGB estimates and existing results

5.2.1. Chosen predicted variables and validation

A variety of spectral variables sensitive to AGB were included. The relative importance of reduced variables in the RF model explained more than that of SGB (Fig.9) because the SGB model is quite sophisticated since changing any settings can affect the optimal values of other settings (Freeman et al., 2014). Vegetation phenology is highly sensitive to climate change. Phenology also controls vegetation feedback to the climate system by influencing the seasonality of albedo, surface roughness length, canopy conductance, and fluxes of water, energy, CO_2 and biogenic volatile organic compounds (Richardson et al., 2013). Inclusion of phenology variables generally improved AGB predictions, especially two variables: the median and mean of dry season NDVI, which were chosen in the optimal RF model (Fig. 11). The relative importance of the AGB predictions (Fig. 9(a)) shows that the time series dry season NDVI contains phenology and soil moisture seasonal variations to distinguish between tree cover and background (Karlson et al., 2015). We only used NDVI to characterize AGB during the dry season, but other spectral indices or data from all of the seasons could be used; for example, tasseled cap greenness and wetness can be used because of their stronger relationship to AGB (Fig. 9(a)).
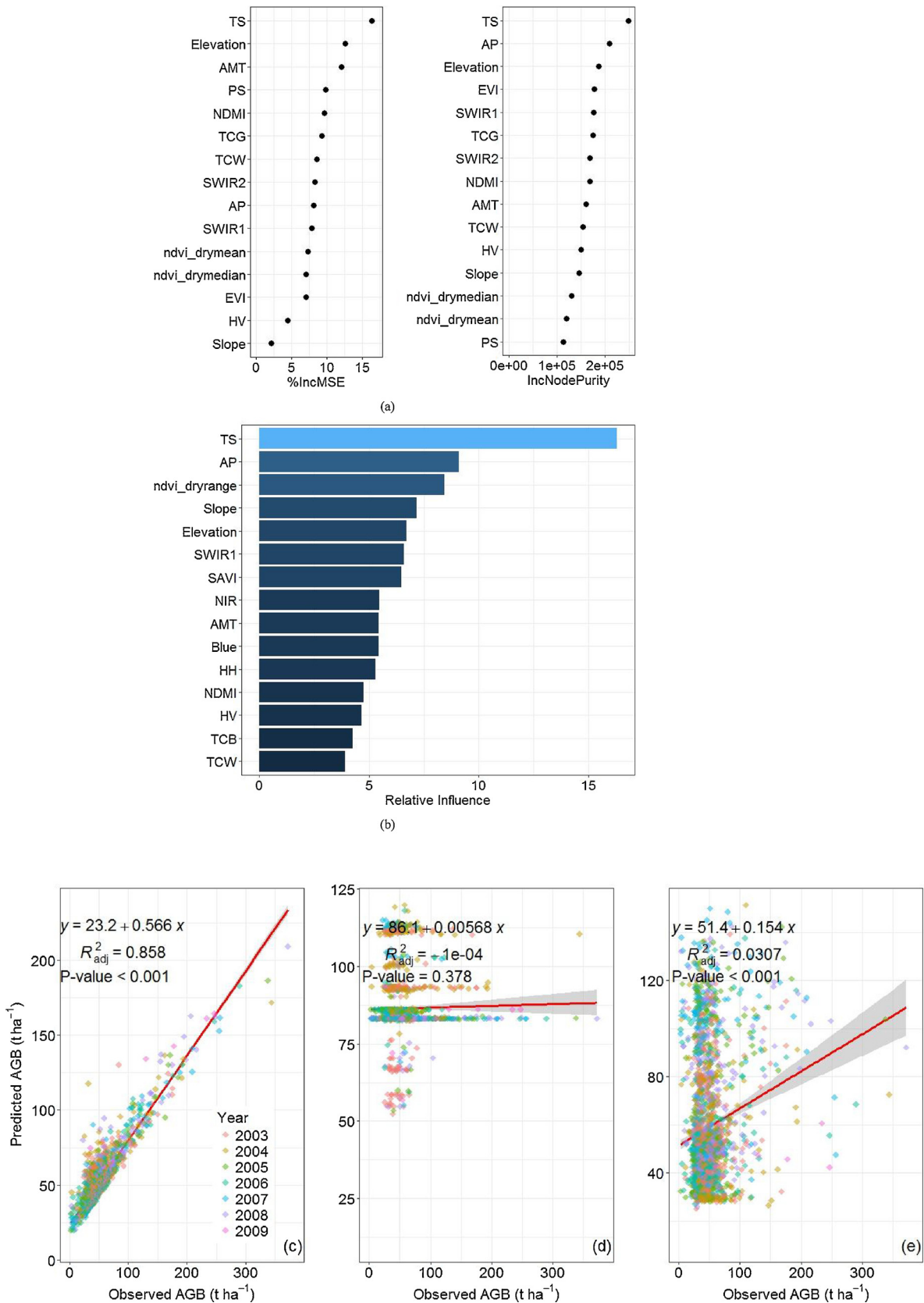
Fig. 9. Variable importance ranking based on OOB error in RF(a), and the relative influence of the reduction of squared error attributable to each variable (decrease in squared error) in SGB1 (b), respectively; relationship between GLAS waveform- extrapolated footprint aboveground biomass (AGB) (predicted) and RF-based GLAS waveform parameters predicted AGB (observed) based on RF (c), SGB1 (d) and SGB2 (e) models, red solid line is the prediction curve, grey solid line indicates the 95% confidence intervals (CI-lines) and the shaded confidence area for the predictions (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Variable Importance Comparison
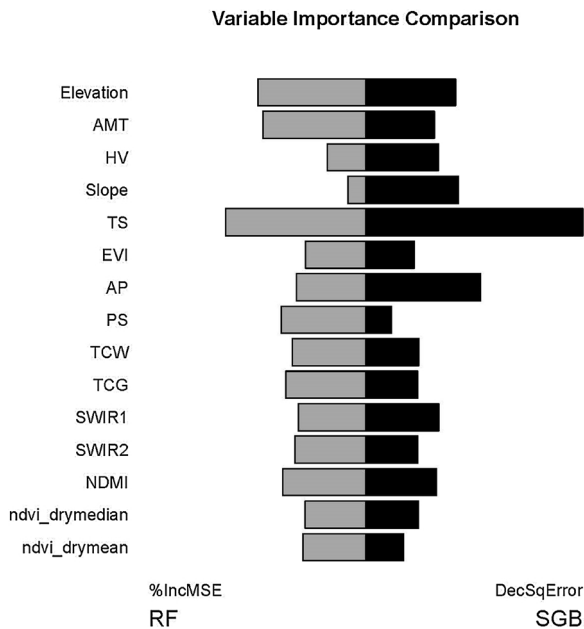
Fig. 10. A comparison of variable importance between the RF and the SGB.

Fig. 12. Kernel density distribution pattern of AGB estimated from 1986 to 2016; points in different colors in the legend represent lines in different years.

Understanding biomass-climate relationships is important for biomass and carbon estimations from the past to future since climate is evolving. Although climatic variables have been used in biomass predictions (Poudel et al., 2011; Zhang et al., 2014b), we only know a little about how climate affects biomass variations(Stegen et al., 2011). Climate was a key influencer in the Guangdong AGB estimation, especially for temperature seasonality, the annual mean temperature, and the annual precipitation, among all of the climatic variables (Fig.9(a)).

Terrain topography in Guangdong is quite complex, and results in the GLAS waveform being stretched (Pang et al., 2011). It was concluded that the regression approach-based SRTM range is one of the outstanding variables for AGB estimation (Boudreau et al., 2008). Chi et al. (2015) pointed out that the bias between biomass estimates under a low slope ($< 20°$) and field-based biomass is smaller than that of those with a high slope ($> 20°$) or all of the slopes. Here, chosen GLAS footprints falling on slopes less than 10° to estimate AGB were considered appropriate. Wu et al. (2016) suggested that the RF algorithm with R^2 (0.63) and RMSE (26.44 t/ha) exhibited the best performance and was better than stochastic gradient boosting, when using Landsat-based AGB estimation in northwestern Zhejiang Province. Chi et al. (2017) revealed that Landsat, GLAS, and DEM-based AGB estimations based on the RF approach had an R^2 of 0.72 and an RMSE of 25.24 Mg/ha. Our results within an R^2_{adj} of 0.858 and an RMSE of 11.35 t/ha induced by
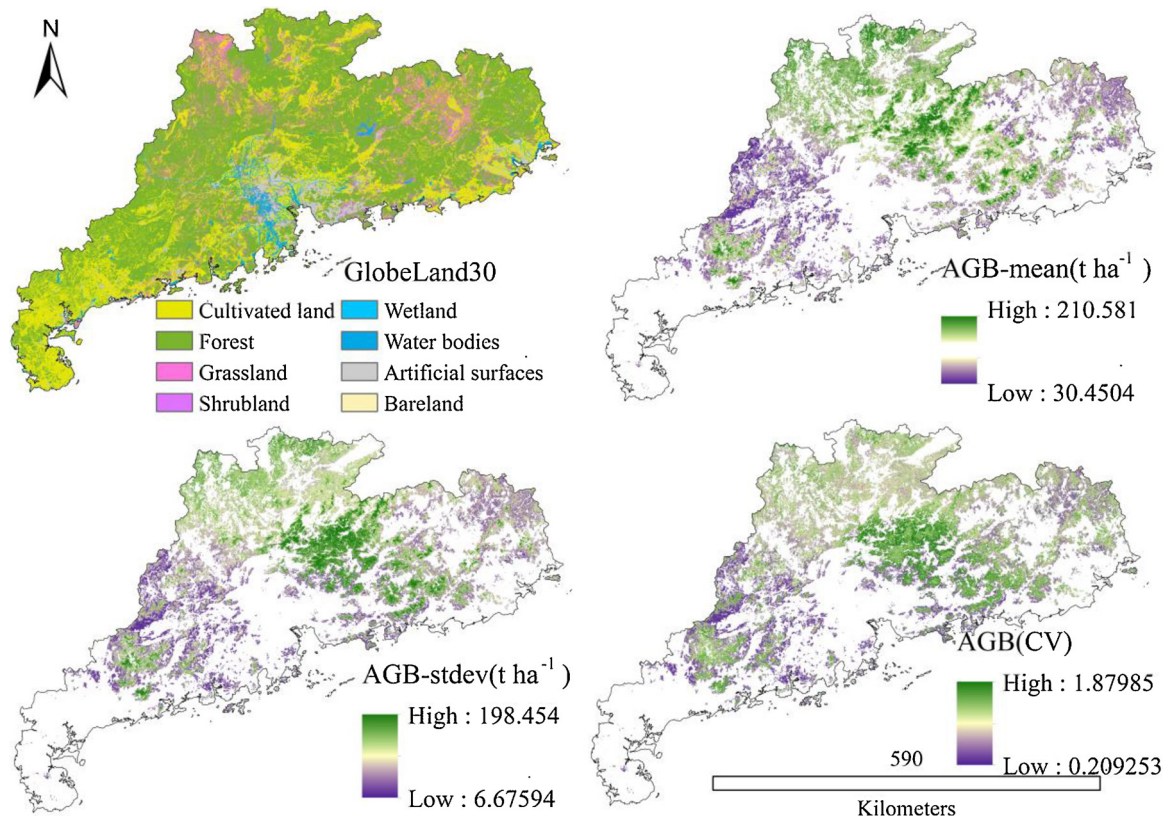

Fig. 11. The predicted AGB pattern map in 2010 and its corresponding uncertainties depicted by the statistics of mean, coefficients of variation and standard deviations combined with GLC in 2010 in Guangdong.
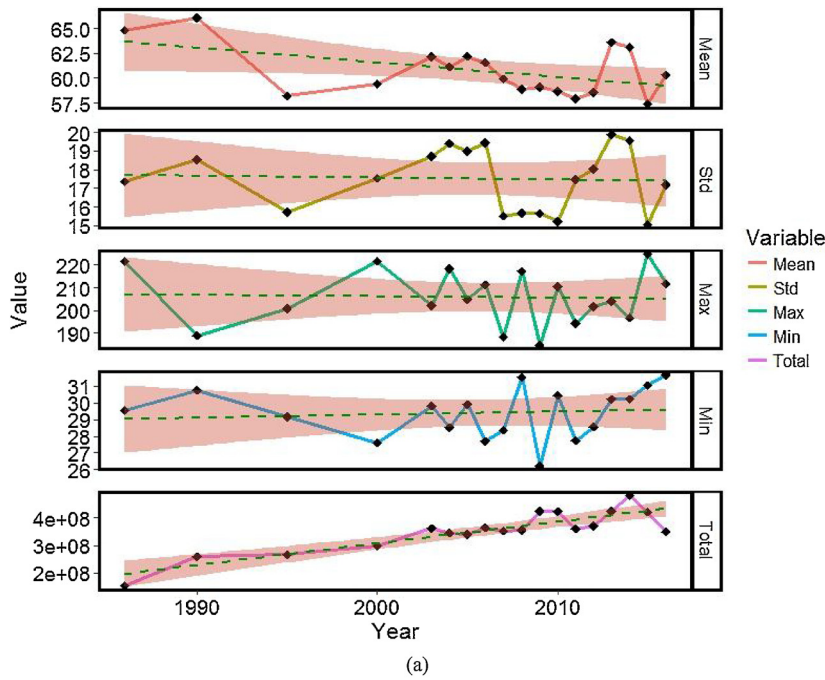
(a)

(b)

(c)

Fig. 13. Total AGB changes in Guangdong from 1986 to 2016, and the statistical parameters (mean, standard deviation, maximum and minimum values) of the predicted AGB maps, the dashed line is the prediction curve against the time, the orange pink solid line indicates the 95% confidence intervals (CI lines) and the shaded confidence area for the predictions (a); Changes in total AGB over three climate zone-based forest zones (b); Changes in total AGB over the four economic zones, including the Mountainous area, Pearl River Delta region, eastern coast of Guangdong and western coast of Guangdong, and the dotted line represents a decrease in the Mountainous region in 2008 (c).
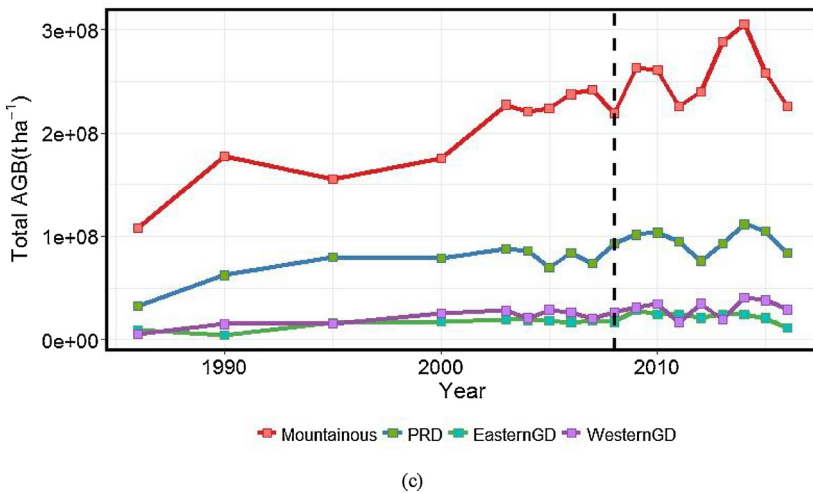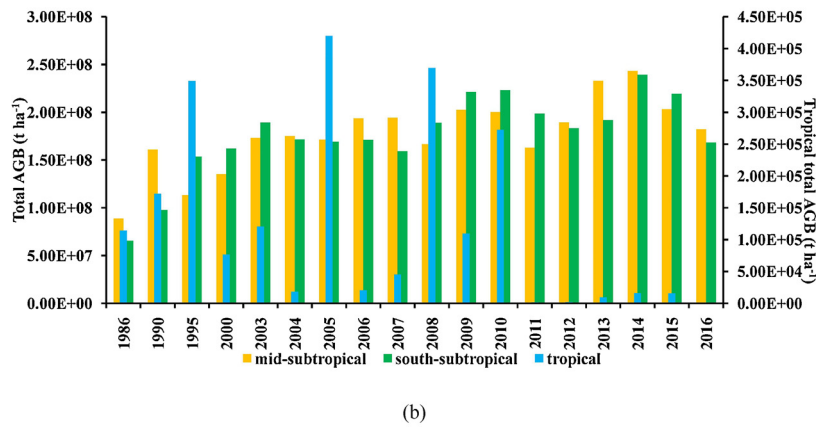
the RF model are superior to the above-mentioned studies, and that of using the SGB model.

5.2.2. Forest classification assessment

Many researchers used the MOD12Q1 global land cover (Chi et al., 2015; Hu et al., 2016; Su et al., 2016; Zhang et al., 2014b) and Landsat classification product (Chi et al., 2017; Margolis et al., 2015; Nelson et al., 2017) in place of forest. However, in view of different definitions and mapping means of land cover (Healey et al., 2012), there are striking differences among land cover maps (Fritz and See, 2008). The overall accuracy and Kappa coefficient of our findings at 92.1% and 0.83, respectively, were acceptable when long-time forest classification maps were generated based on a pixel-based algorithm.

5.2.3. Forest AGB estimation and comparisons

Referring to the RF-based modeling results, it appears that the mean AGB approximated 69.73 and 74.53 t/ha in 2007 and 2011, respectively, in northern Guangdong. However, the RF model-based AGB assessment within the VCT-based forest type in northern Guangdong derived from Landsat, PALSAR, and NFI showed a mean AGB of 59.61 and 60.48 t/ha, respectively (Shen et al., 2016). The results of this study are much higher than those of the former estimation results and a mean AGB of 59.92 and 64.56 t/ha from China NFI, respectively. The predicted total AGB of 3.5×10^8 t in 2016 in the entirety of Guangdong was lower than that of the AGB data (5.944×10^8 t) from the Guangdong Forestry and Ecological Situation Bulletin. Such discrepancies are attributable to different forest definition used in the two cases. Du et al. (2014) used MODIS and NFI to estimate the entire Guangdong AGB, and the underestimated biomass distribution was found between 0 and 40.4 Mg/ha.

5.3. Major drivers for forest AGB dynamics

Although several data sets support an overall forest AGB increase in Guangdong, the direction and magnitude of the forest AGB change is still unknown because of the large uncertainties in remote sensing-derived estimates, which can be explained by human activities and climate effects. The local government proposed an afforestation plan of "Green Guangdong" in 1985 and achieved it in 1993; in 1994 and 2005, the government encouraged the development of sustainable and modern forestry (Ren et al., 2013), which might be one of the causes of the significant increase in forest biomass density in Southeast China (including Guangdong) from 1985 to 2005 (Piao, 2005). The AGB decrease in 2008 (Fig. 13(c), dotted line) in the mountainous area and eastern coast of Guangdong can be explained by freezing rain and snow disasters in 2008, resulting in a loss of one-tenth of China's forests (Stone, 2008). The major land-use changes, such as an increase in tropical economic forests (e.g., eucalyptus plantations and rubber trees) (Zhu, 2017) and the frequent typhoons and droughts (Wang et al., 2016; Zhang et al., 2017), directly result in the tropical forest having large variations before 2010, especially in the Leizhou Peninsula (the largest eucalyptus distribution region lies at the southern part of mainland China) (Fig. 1). The Leizhou Peninsula has experienced gradual forest loss, water shortages, and ecosystem degradation biodiversity loss (Zhang et al., 2017)—for instance, a 60 year drought disaster occurred in 2015, which may explain the sharply decreasing AGB up to 2016 (Fig. 13(b)).

However, most mountainous areas in northern Guangdong have had slow economic development and have retained natural forests and plantations to provide important ecological protection in Guangdong, suggesting why northern Guangdong has the highest forest AGB (Fig. 13(c)) (Ren et al., 2013). Since the early 1980s, the Pearl River Delta has witnessed a rapid development of population and economy, large plantations have quickly replaced natural forests, and the mature plantations have led to a higher forest AGB stock (Ren et al., 2013), lower than that of northern Guangdong (Fig. 13(c)). Similarly, western

Guangdong conducted massive reclamation and establishment of commercial plantations, bringing many eucalyptus forests; the coastal shelter belt of eastern Guangdong planted water conservation forest (Ren et al., 2013). Young plantations occupying the large areas in the western and eastern regions may make forest AGB storage low (Ren et al., 2013). Additionally, the establishment of using native broad-leaved plantations excluding eucalyptus and pine plantations contributed to high forest AGB storage (Ren et al., 2013). Thus, forest management for high forest AGB and carbon stock should be examined.

5.4. Uncertainty in detection of AGB

The main reason for uncertainty for total forest AGB detection was different sources and spectral discrepancy in various tree species, which ultimately affected the production of land-cover maps (Zheng et al., 2008). Using some of the GLAS waveform parameters to relate AGB was more accurate than using an allometric equation to estimate AGB by airborne LIDAR-based canopy height estimated from the GLAS waveform (Hayashi et al., 2013). The GLAS footprints were from 2003 to 2009; however, the field plots were surveyed from 2005–2009. The time discrepancy between the field plots (2005) and the observations (2003 and 2004), and the inconsistent plot sizes, can result in errors in AGB estimation. The next-generation space-borne LIDAR systems will have an important role in future forest resource monitoring (Hayashi et al., 2015; Healey et al., 2012; Margolis et al., 2015).

6. Conclusions

In the study, the estimation method of forest AGB distribution in Guangdong Province was developed through the combination of multisource data sets, namely the RF-based GLAS waveform-derived AGB model and the RF-based GLAS waveform- extrapolated footprint AGB model. The final two models were used to effectively construct GLAS waveform-based AGB prediction framework. As validated by field inventory data and uncertainty maps, the estimated models generated reliable forest AGB maps with some basic prediction error. By integrating phenology variables and climate variables, we obtained much interannual and seasonal, or spatio-temporal dynamics of forest AGB. Although combining multiple sources (field measurements, remote sensing information and the optimal machine learning algorithms) to develop models has some uncertainties, findings from this study will provide a basis for exploring the performance of forest AGB change detection in subtropical and tropical regions, which can help policymakers and the remote-sensing community understand forest biomass carbon stocks associated with afforestation and deforestation in plantations and rainfall forests under the impacts of forest disturbance and climate change.

Appendix A

Table A1
Summary of the six types of predictor variables including speCctrum, phenology, climate, topography, texture, and PALSAR components for AGB estimation models.

Type	Variable	Formula	Ref.	Description
Spectrum	R,G,B, NIR,SWIR1,SWIR2			Landsat 5,7,8 bands
	NDVI	$(NIR - R)/(NIR + R)$	(Rouse et al., 1973)	Normalized Difference Vegetation Index
	NDMI	$(NIR - SWIR)/(NIR + SWIR)$	(Gao, 1996)	Normalized Difference Moisture Index
	EVI	$2.5 \times (NIR - R)/(NIR + 6 \times R - 7.5 \times B + 1)$	(Huete et al., 2002)	Enhanced vegetation index
	SAVI	$(1 + 0.5) \times (NIR\text{-}R)/(NIR + R + 0.5)$	(Huete, 1988)	Soil-adjusted Vegetation Index
Phenology (dry season NDVI)	Maximum, mean, median, minimum, range, standard deviation		(Gessner et al., 2013)	Cumulative NDVI
Climate surfaces (Worldclim)	AMP		(Hijmans et al., 2005)	Annual Mean Precipitation (mm)
	PS			Precipitation Seasonality (Coefficient of Variation)
	AMT			Annual Mean Temperature (℃)
	TS			Temperature Seasonality (standard deviation × 100)
Tasseled cap transformations	TCB, TCG, TCW		(Baig et al., 2014; Crist and Cicone, 1984; Huang et al., 2002)	Brightness, Greenness, Wetness (L5, L7, L8)
Topography	Elevation, slope			
Land cover type	LC		(Chen et al., 2015)	GlobeLand30
Texture (window sizes 3 × 3pixels)	mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, correlation		(Haralick et al., 1973)	GLCM texture measures
PALSAR	HH, HV, HH/HV			PALSAR components

Appendix B. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.agrformet.2018.04.005.

The supplementary materials provide one data sheet for each initially collected plot measurment and GLAS waveform parameter values ("mydata1_AGB_GLASwaveforms.csv"), another data sheet for GLAS-derived AGB and various values ("mydata2_AGB_variables.csv"), also the RF and SGB models' running R code ("LUT_1200432016_reduce.csv" is a "raster look up table" to map RF-based AGB map).

References

Alexandridis, T.K., Cherif, I., Kalogeropoulos, C., Monachou, S., Eskridge, K., Silleos, N., 2013. Rapid error assessment for quantitative estimations from Landsat 7 gap-filled images. Remote Sens. Lett. 4, 920–928.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S., Houghton, R.A., 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nat. Clim. Change 2, 182–185.

Baig, M.H.A., Zhang, L., Shuai, T., Tong, Q., 2014. Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. Remote Sens. Lett. 5, 423–431.

Boudreau, J., Nelson, R.F., Margolis, H.A., Beaudoin, A., Guindon, L., Kimes, D.S., 2008. Regional aboveground forest biomass using airborne and spaceborne LiDAR in Quebec. Remote Sens. Environ. 112, 3876–3890.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30m resolution: a POK-based operational approach. ISPRS J. Photogramm. 103, 7–27.

Chen, Q., 2013. Lidar remote sensing of vegetation biomass. Remote Sens. Nat. Resour. 399, 399–420.

Chi, H., Sun, G., Huang, J., Guo, Z., Ni, W., Fu, A., 2015. National Forest aboveground biomass mapping from ICESat/GLAS data and MODIS imagery in China. Remote Sens. 7, 5534–5564.

Chi, H., Sun, G., Huang, J., Li, R., Ren, X., Ni, W., Fu, A., 2017. Estimation of Forest aboveground biomass in Changbai Mountain region using ICESat/GLAS and Landsat/TM data. Remote Sens. 9, 707.

Coulston, J.W., Moisen, G.G., Wilson, B.T., Finco, M.V., Cohen, W.B., Brewer, C.K., 2012. Modeling percent tree canopy cover a pilot study. Photogramm. Eng. Rem. S 78, 715–727.

Crist, E.P., Cicone, R.C., 1984. A physically-based transformation of thematic mapper data - the Tm tasseled cap. IEEE Trans. Geosci. Remote Sens. 22, 256–263.

Dong, J.R., Kaufmann, R.K., Myneni, R.B., Tucker, C.J., Kauppi, P.E., Liski, J., Buermann, W., Alexeyev, V., Hughes, M.K., 2003. Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. Remote Sens. Environ. 84, 393–410.

Du, L., Zhou, T., Zou, Z., Zhao, X., Huang, K., Wu, H., 2014. Mapping forest biomass using remote sensing and national forest inventory in China. Forests 5, 1267–1283.

Dube, T., Mutanga, O., Elhadi, A., Ismail, R., 2014. Intra-and-inter species biomass prediction in a plantation forest: testing the utility of high spatial resolution spaceborne multispectral RapidEye sensor and advanced machine learning algorithms. Sensors (Basel) 14, 15348–15370.

Duncanson, L.I., Niemann, K.O., Wulder, M.A., 2010. Integration of GLAS and Landsat TM data for aboveground biomass estimation. Can. J. Remote Sens. 36, 129–141.

Fang, J.Y., Chen, A.P., Peng, C.H., Zhao, S.Q., Ci, L., 2001. Changes in forest biomass carbon storage in China between 1949 and 1998. Science 292, 2320–2322.

FAO, 2012. In: Nations., F.a.A.O.o.t.U (Ed.), Global Forest Resource Assessment (FRA), Rome, Italy.

Foody, G.M., Boyd, D.S., Cutler, M.E.J., 2003. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. Remote Sens. Environ. 85, 463–474.

Freeman, E., 2009. ModelMap: An R Package for Modeling and Map Production Using Random Forest and Stochastic Gradient Boosting. . USDA Forest Service, Rocky Mountain Research Station, 507 25th street, Ogden, UT, USA. http://CRAN.R-project.org/.eafreeman@fs.fed.us.

Freeman, E.A., Moisen, G.G., Coulston, J.W., Wilson, B.T., 2015. Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance1. Can. J. For. Res. 46, 323–339.

Freeman, E.D., Larsen, R.T., Peterson, M.E., Anderson, C.R., Hersey, K.R., Mcmillan, B.R., 2014. Effects of male-biased harvest on mule deer: implications for rates of pregnancy, synchrony, and timing of parturition. Wildl. Soc. B 38, 806–811.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Annal. Stat. 29, 1189–1232.

Friedman, J.H., 2002. Stochastic gradient boosting. Comput. Stat. Data Anal. 38, 367–378.

Fritz, S., See, L., 2008. Identifying and quantifying uncertainty and spatial disagreement

in the comparison of global land cover for different applications. Global Change Biol. 14, 1057–1075.

Gao, B.C., 1996. NDWI - a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266.

Genuer, R., Poggi, J.M., Tuleau-Malot, C., 2015. VSURF: An R Package for Variable Selection Using Random Forests. R J, vol. 7. pp. 19–33.

Gessner, U., Machwitz, M., Conrad, C., Dech, S., 2013. Estimating the fractional cover of growth forms and bare surface in savannas. A multi-resolution approach based on regression tree ensembles. Remote Sens. Environ. 129, 90–102.

Guo, Z., Chi, H., Sun, G., 2010. Estimating forest aboveground biomass using HJ-1 satellite CCD and ICESat GLAS waveform data. Sci. China Earth Sci. 53, 16–25.

Haralick, R.M., Shanmuga, K., Dinstein, I., 1973. Textural Features for Image Classification. Ieee T Syst Man Cyb Smc3. pp. 610–621.

Harding, D.J., Carabajal, C.C., 2005. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. Geophys. Res. Lett. 32.

Hayashi, M., Saigusa, N., Oguma, H., Yamao, Y., Yamagata, Y., Takao, G., 2013. Applying ICESat/GLAS Data to Estimate Forest Aboveground Biomass on Hokkaido, Japan. In AGU Fall Meeting Abstracts.

Hayashi, M., Saigusa, N., Yamagata, Y., Hirano, T., 2015. Regional forest biomass estimation using ICESat GLAS spaceborne LiDAR over Borneo. Carbon Manage. 6, 19–33.

Healey, S.P., Patterson, P.L., Saatchi, S., Lefsky, M.A., Lister, A.J., Freeman, E.A., 2012. A sample design for globally consistent biomass estimation using lidar data from the Geoscience Laser Altimeter System (GLAS). Carbon Balance Manage. 7, 1–10.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.

Houghton, R.A., Hall, F., Goetz, S.J., 2009. Importance of biomass in the global carbon cycle. J. Geophys. Res. Biogeosci. 114, 1–13.

Hu, T., Su, Y., Xue, B., Liu, J., Zhao, X., Fang, J., Guo, Q., 2016. Mapping global Forest aboveground biomass with spaceborne LiDAR, optical imagery, and Forest inventory data. Remote Sens. 8, 565.

Huang, C., Wylie, B., Yang, L., Homer, C., Zylstra, G., 2002. Derivation of a tasselled cap transformation based on Landsat 7 at-satellite reflectance. Int. J. Remote Sens. 23, 1741–1748.

Huang, H., Liu, C., Wang, X., Biging, G.S., Chen, Y., Yang, J., Gong, P., 2017. Mapping vegetation heights in China using slope correction ICESat data, SRTM, MODIS-derived and climate data. ISPRS J. Photogramm. 129, 189–199.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.

Karlson, M., Ostwald, M., Reese, H., Sanou, J., Tankoano, B., Mattsson, E., 2015. Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian Woodlands using Landsat 8 and random Forest. Remote Sens. 7, 10017–10041.

Le Toan, T., Quegan, S., Davidson, M.W.J., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H., Ulander, L., 2011. The BIOMASS mission: mapping global forest biomass to better understand the terrestrial carbon cycle. Remote Sens. Environ. 115, 2850–2860.

Lefsky, M.A., Cohen, W.B., Harding, D.J., Parker, G.G., Acker, S.A., Gower, S.T., 2002. Lidar remote sensing of above-ground biomass in three biomes. Global Ecol. Biogeogr. 11, 393–399.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Espirito-Santo, F.D., Hunter, M.O., de Oliveira, R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. Geophys. Res. Lett. 32.

Lei, X.D., Tang, M.P., Lu, Y.C., Hong, L.X., Tian, D.L., 2009. Forest inventory in China: status and challenges. Int. For. Rev. 11, 52–63.

Liaw, A., Matthew, W., 2002. Classification and regression by randomForest. R. News 2, 18–22.

Los, S.O., Rosette, J.A.B., Kljun, N., North, P.R.J., Chasmer, L., Suarez, J.C., Hopkinson, C., Hill, R.A., van Gorsel, E., Mahoney, C., Berni, J.A.J., 2012. Vegetation height and cover fraction between 60A degrees S and 60A degrees N from ICESat GLAS data. Geosci. Modell. Dev. 5, 413–432.

Lu, D.S., 2006. The potential and challenge of remote sensing-based biomass estimation. Int. J. Remote Sens. 27, 1297–1328.

Lu, D.S., Chen, Q., Wang, G.X., Liu, L.J., Li, G.Y., Moran, E., 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. Int. J. Digit. Earth 9, 63–105.

Lucas, R.M., Cronin, N., Lee, A., Moghaddam, M., Witte, C., Tickle, P., 2006. Empirical relationships between AIRSAR backscatter and LiDAR-derived forest biomass, Queensland, Australia. Remote Sens. Environ. 100, 407–425.

Main-Knorn, M., Cohen, W.B., Kennedy, R.E., Pflugmacher, D., Griffiths, P., Hostert, P., 2013. Monitoring coniferous forest biomass change using a Landsat trajectory-based approach. Remote Sens. Environ. 139, 277–290.

Margolis, H.A., Nelson, R.F., Montesano, P.M., Beaudoin, A., Sun, G., Andersen, H.-E., Wulder, M.A., 2015. Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. Can. J. For. Res. 45, 838–855.

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., Lim, T.K., 2006. A Landsat surface reflectance dataset for North America, 1990-2000. IEEE Geosci. Remote Sens. Lett. 3, 68–72.

Mitchard, E.T.A., Saatchi, S.S., White, L.J.T., Abernethy, K.A., Jeffery, K.J., Lewis, S.L., Collins, M., Lefsky, M.A., Leal, M.E., Woodhouse, I.H., Meir, P., 2012. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: overcoming problems of high biomass and persistent cloud. Biogeosciences 9, 179–191.

Nelson, R., Margolis, H., Montesano, P., Sun, G., Cook, B., Corp, L., Andersen, H.-E., deJong, B., Pellat, F.P., Fickel, T., Kauffman, J., Prisley, S., 2017. Lidar-based estimates of aboveground biomass in the continental US and Mexico using ground, airborne, and satellite observations. Remote Sens. Environ. 188, 127–140.

Pang, Y., Lefsky, M., Sun, G., Ranson, J., 2011. Impact of footprint diameter and off-nadir pointing on the precision of canopy height estimates from spaceborne lidar. Remote Sens. Environ. 115, 2798–2809.

Piao, S., 2005. Forest biomass carbon stocks in China over the past 2 decades: estimation based on integrated inventory and satellite data. J. Geophys. Res. 110.

Poudel, B.C., Sathre, R., Gustavsson, L., Bergh, J., Lundström, A., Hyvönen, R., 2011. Effects of climate change on biomass production and substitution in north-central Sweden. Biomass Bioenergy 35, 4340–4355.

Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., Ohmann, J.L., 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: a comparison of empirical modeling approaches. Remote Sens. Environ. 114, 1053–1068.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0. http://www.R-project.org.

Ren, H., Chen, H., Li, L., Li, P., Hou, C., Wan, H., Zhang, Q., Zhang, P., 2013. Spatial and temporal patterns of carbon storage from 1992 to 2002 in forest ecosystems in Guangdong, Southern China. Plant Soil 363, 123–138.

Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173.

Ridgeway, G., 2007. Generalized Boosted Models: A Guide to the Gbm Package. Update 1.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the Great plains with ERST. NASA Spec. Publ. 351, 309–317.

Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. PNAS 108, 9899–9904.

Schutz, B.E., Zwally, H.J., Shuman, C.A., Hancock, D., DiMarzio, J.P., 2005. Overview of the ICESat mission. Geophys. Res. Lett. 32.

Shen, W.J., Li, M., Huang, C.Q., Wei, A.S., 2016. Quantifying live aboveground biomass and forest disturbance of mountainous natural and plantation forests in Northern Guangdong, China, based on multi-temporal Landsat, PALSAR and field plot data. Remote Sens. 8, 595.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). Remote Sens. Environ. 155, 13–31.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. CRC press.

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. 116.

Stegen, J.C., Swenson, N.G., Enquist, B.J., White, E.P., Phillips, O.L., Jorgensen, P.M., Weiser, M.D., Mendoza, A.M., Vargas, P.N., 2011. Variation in above-ground forest biomass across broad climatic gradients. Global Ecol. Biogeogr. 20, 744–754.

Stone, R., 2008. Natural disasters - ecologists report huge storm losses in China's forests. Science 319, 1318–1319.

Su, Y., Guo, Q., Xue, B., Hu, T., Alvarez, O., Tao, S., Fang, J., 2016. Spatial distribution of forest aboveground biomass in China: estimation through combination of spaceborne lidar, optical imagery, and forest inventory data. Remote Sens. Environ. 173, 187–199.

Sun, G., Ranson, K., Kimes, D., Blair, J., Kovacs, K., 2008. Forest vertical structure from GLAS: an evaluation using LVIS and SRTM data. Remote Sens. Environ. 112, 107–117.

Tian, X., Yan, M., van der Tol, C., Li, Z., Su, Z., Chen, E., Li, X., Li, L., Wang, X., Pan, X., Gao, L., Han, Z., 2017. Modeling forest above-ground biomass dynamics using multi-source data and incorporated models: a case study over the qilian mountains. Agric. For. Meteorol. 246, 1–14.

Vogelmann, J., Khoa, P., Lan, D., Shermeyer, J., Shi, H., Wimberly, M., Duong, H., Huong, L., 2017. Assessment of forest degradation in Vietnam using Landsat time series data. Forests 8, 238.

Wang, D., Wang, X., Liu, L., Wang, D., Huang, H., Pan, C., 2016. Evaluation of CMPA precipitation estimate in the evolution of typhoon-related storm rainfall in Guangdong, China. J. Hydroinform. 18, 1055–1068.

Wang, H., Mo, J., Lu, X., Xue, J., Li, J., Fang, Y., 2009. Effects of elevated nitrogen deposition on soil microbial biomass carbon in major subtropical forests of southern China. Front. For. China 4, 21–27.

Wu, C., Shen, H., Shen, A., Deng, J., Gan, M., Zhu, J., Xu, H., Wang, K., 2016. Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. J. Appl. Remote Sens. 10, 035010.

Wulder, M.A., White, J.C., Fournier, R.A., Luther, J.E., Magnussen, S., 2008. Spatially explicit large area biomass estimation: Three approaches using forest inventory and remotely sensed imagery in a GIS. Sensors-Basel 8, 529–560.

Xie, X., Wang, Q., Dai, L., Su, D., Wang, X., Qi, G., Ye, Y., 2011. Application of China's national Forest continuous inventory database. Environ. Manage. 48, 1095–1106.

Xing, Y., Qiu, S., Ding, J., Tian, J., 2016. Estimation Of Regional Forest Aboveground Biomass Combining ICESat-GLAS Waveforms And HJ-1A/HSI Hyperspectral Imageries. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLI-B7. pp. 731–737.

Yu, Y., Saatchi, S., 2016. Sensitivity of L-Band SAR backscatter to aboveground biomass of global forests. Remote Sens. 8, 522.

Yu, Y., Yang, X., Fan, W., 2015. Estimates of forest structure parameters from GLAS data and multi-angle imaging spectrometer data. Int. J. Appl. Earth Obs. Geoinf. 38, 65–71.

Zhang, H., Cheng, W., Qiu, X., Feng, X., Gong, W., 2017. Tide-surge interaction along the east coast of the Leizhou Peninsula, South China Sea. Cont. Shelf Res. 142, 32–49.

Zhang, J., Huang, S., Hogg, E.H., Lieffers, V., Qin, Y., He, F., 2014a. Estimating spatial variation in Alberta forest biomass from a combination of forest inventory and remote

sensing data. Biogeosciences 11, 2793–2808.

Zhang, Y.Z., Liang, S.L., Sun, G.Q., 2014b. Forest biomass mapping of northern eastern china using GLAS and MODIS. Ieee J-Stars 7, 140–152.

Zhao, F., Guo, Q., Kelly, M., 2012. Allometric equation choice impacts lidar-based forest biomass estimates: a case study from the Sierra national Forest. CA. Agric. For. Meteorol. 165, 64–72.

Zheng, D.L., Heath, L.S., Ducey, M.J., 2008. Satellite detection of land-use change and effects on regional forest aboveground biomass estimates. Environ. Monit. Assess.

144, 67–79.

Zhou, L., Wu, J., Mo, X., Zhou, H., Diao, C., Wang, Q., Chen, Y., Zhang, F., 2017. Quantitative and detailed spatiotemporal patterns of drought in China during 2001-2013. Sci. Total Environ. 589, 136–145.

Zhu, H., 2017. The tropical forests of Southern China and conservation of biodiversity. Bot. Rev. 83, 87–105.

Zvoleff, A., 2015. Glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R. http://CRAN.R-project.org/package=glcm.R.