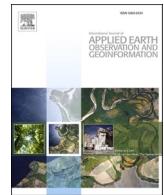




Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Improved estimation of forest stand volume by the integration of GEDI LiDAR data and multi-sensor imagery in the Changbai Mountains Mixed forests Ecoregion (CMMFE), northeast China



Lin Chen^{a,c}, Chunying Ren^{b,*}, Bai Zhang^b, Zongming Wang^{b,g}, Mingyue Liu^{d,e,f}, Weidong Man^{d,e,f}, Jiafu Liu^h

^a Institute of Remote Sensing and Earth Sciences, College of Science, Hangzhou Normal University, Hangzhou 311121, China

^b Northeast Institute of Geography and Agroecology, Key Laboratory of Wetland Ecology and Environment, Chinese Academy of Sciences, Changchun 130102, China

^c Zhejiang Provincial Key Laboratory of Urban Wetlands and Regional Change, Hangzhou 311121, China

^d College of Mining Engineering, North China University of Science and Technology, Tangshan 063210, China

^e Hebei Key Laboratory of Mining Development and Security Technology, Tangshan 063210, China

^f Hebei Industrial Technology Institute of Mine Ecological Remediation, Tangshan 063210, China

^g National Earth System Science Data Center, Beijing 100101, China

^h College of Tourism and Geographic Sciences, Jilin Normal University, Siping 136000, China

ARTICLE INFO

Keywords:

GEDI LiDAR
Sentinel imagery
Stand volume
Geographically weighted regression
Random forests
Changbai Mountains Mixed Forests Ecoregion

ABSTRACT

Forest stand volume is a vital indicator of productivity and carbon storage. Conventionally, stand volume is estimated from field samples, Synthetic Aperture Radar (SAR), and optical imagery, which suffer saturation problems. Although Light Detection and Ranging (LiDAR) technique degrades the signal saturation, its large-scale application is hindered by spatial continuity. To address this issue, recently released Global Ecosystem Dynamics Investigation (GEDI) LiDAR data, Sentinel-1 SAR, Sentinel-2 Multispectral Instrument (MSI), and Advanced Land Observing Satellite (ALOS) digital surface model (DSM) imagery were integrated for volume modeling and estimation under a point-line-polygon framework. The footprint-level LiDAR variables, as a linear bridge, were adopted to link field plots to the full-cover multi-sensor imagery. Results showed that volume of the Changbai Mountains Mixed Forests Ecoregion (CMMFE) displayed variations along the elevation gradient, ranging from 47.56 to 277.30 m³/ha with a mean value of 151.39 m³/ha. Additionally, the accuracy comparison based on independent validation samples indicated that integrating GEDI LiDAR data under a point-line-polygon framework performed better than the traditional point-polygon approach, which directly linked field samples to multi-sensor imagery. The corresponding estimated error declined from 22.08% to 15.21%. The canopy cover and tree height from LiDAR, elevation from L band InSAR, and spectral indices of MSI red-edge bands were key for stand volume mapping in heterogeneous temperate forests. This comparison also showed that the integration of LiDAR by a point-line-polygon framework adopted 2/3 of the modeling points but acquired more accurate estimation than a traditional approach only based on multi-sensor imagery, which implied less field sampling work was needed for similar research. Consequently, as a pioneering exploration of GEDI LiDAR data combined with multi-sensor imagery under the point-line-polygon framework, this study provides an efficient methodology for the volume estimation of heterogeneous forests.

1. Introduction

The Changbai Mountains Mixed Forests Ecoregion (CMMFE) includes the Changbai Mountain areas in China and North Korea. This ecoregion features the most naturally protected deciduous trees and

extensive late-successional conifer forests in northeast China (Olson et al., 2001). The region also supports numerous endemic plant species due to its temperate climate and relative isolation, but is central to various floral communities (Olson et al., 2001; Wang et al., 2011; Chen et al., 2019a). Forest stand volume, a common parameter of National

* Corresponding author.

E-mail address: renchy@iga.ac.cn (C. Ren).

Forest Inventory worldwide, forms the basis for ecosystem health and decision-making at various levels (Fazakas et al., 1999; Condés and McRoberts, 2017; Nilsson et al., 2017). The forest stand volume in CMMFE in northeastern China provides more than one-third of Chinese timber yield, and is also the baseline of aboveground biomass estimation (Wang, 2006; Zhang et al., 2015; National Forestry and Grassland Administration, 2019). Therefore, a spatially continuous estimation of stand volume in this ecoregion is essential for quantifying and monitoring forest-related carbon emissions, forest degradation, and ecological security (Chirici et al., 2016; Chen et al., 2020).

Traditionally, forest stand volume is estimated by measured values of canopy height and diameter from harvested trees based on a species-sensitive volume table, which is costly and spatially limited (Hyppä et al., 2000; Boisvenue et al., 2016). Combining various remote sensing data and field-measured samples has become an efficient methodology to generate full-cover estimations of forest stand volume (Santoro et al., 2011; Saarela et al., 2015; Mauya et al., 2019). Optical sensors were first applied to assess stand volume through field sampling due to the characteristic reflectance characterized by horizontal forest structures (Kilpeläinen and Tokola, 1999; Lausch et al., 2017; Chrysafis et al., 2019). However, owing to poor penetration capacity through clouds and forest canopies, optical sensor-based volume estimation is subjected to severe saturation problems (Hyppä et al., 2000; Long et al., 2019). Active microwave sensors with better penetrability and detailed vertical information derived from different frequency bands, polarizations, and imaging geometries, such as Synthetic Aperture Radar (SAR), are renowned for stand volume mapping (Lausch et al., 2017; Chen et al., 2020). Although, SAR still faces signal saturation with volume values above 200 m³/ha in closed heterogeneous forests (Santi et al., 2015; Zhang et al., 2019). Light Detection and Ranging (LiDAR) showed great potential on forest three-dimensional modeling and mapping over the past few years, reducing the saturation problems and improving the result precision (Cartus et al., 2012; Puliti et al., 2018). Because of a lack of space-borne sensors and space continuity, LiDAR data are still auxiliary in stand volume mapping at large scales.

The remote sensing data mentioned above were used for stand volume estimation and each has strong and weak points (Ranson et al., 2007; Thiel and Schmullius, 2016; de Souza et al., 2019; Chirici et al., 2020). Progress has been made by integrating multi-sensor remote sensing modeling for spatiotemporally uniform volume estimation (Lehmann et al., 2015; de Souza et al., 2019; Mauya et al., 2019; Xie et al., 2020). Hyde et al. (2006) mapped forest height and biomass based on variables from LiDAR, radar, and Landsat ETM⁺ sensors and concluded that the combination of all sensors was more accurate than LiDAR alone. Hawrylo and Węzyk (2018) used Sentinel-2 satellite imagery and airborne point cloud to estimate stand volume, and achieving more accurate predictions than using single-sensor data. LiDAR data are generally adopted to link sampling points of field measurements with full-cover imagery using two steps: field-measured volume is connected to LiDAR variables; then, modeling of LiDAR-estimated volume by satellite predictors for the final mapping is conducted (Matasci et al., 2018; Li et al., 2020). This approach is the point-line-polygon framework and has yielded accurate spatial mapping of forest parameters (Wang et al., 2020). However, due to restriction to free-access satellite sensors, large-scale estimation of forest stand volume by integration of LiDAR data and multi-sensor imagery under the point-line-polygon framework has rarely been reported.

Open-access GEDI LiDAR observations have been collected and processed to create useful results for stand volume estimation since April 2019 (Magruder et al., 2019; Dubayah et al., 2020; Potapov et al., 2021). Among them, canopy cover and height are linear with diameter at breast height (DBH) and tree height, respectively, and have great potential for forest volume estimation (Narine et al., 2019; Puliti et al., 2020). Limited by the footprint format and spatiotemporal resolution, the usage of GEDI data to acquire wall-to-wall forest stand volume still encounters difficulties without assistance from multi-sensor imagery. Famed for free

accessibility to C band SAR and multispectral instrument (MSI) imagery with fine spatiotemporal resolution and global coverage, Sentinel-1 and Sentinel-2 have capabilities of large-scale mapping of stand volume (Mauya et al., 2019; Chen et al., 2020). The C band SAR signal back-scattered from forests depends on the tree number, canopy cover, and height, as well as backgrounds, which makes it useful for forest volume estimation (Dos Reis et al., 2019; Liu et al., 2019a). Based on the strong absorption of red and blue-violet light during photosynthesis, the MSI signal reflected from forests contained information on biophysical variables and vegetation growth status for stand volume estimation (Dube et al., 2015; Ahmadi et al., 2020). The Advanced Land Observing Satellite (ALOS) digital surface model (DSM) provided accurate topographic indicators derived from L band interferometric SAR (InSAR) and was helpful in stand volume mapping as a representative of the local hydrothermal condition (Rahlf et al., 2014; Tadono et al., 2015). Hence, further exploration on how large-scale volume mapping would be achieved with the integration of GEDI LiDAR data and multi-sensor imagery from Sentinel-1, Sentinel-2 and ALOS DSM under the point-line-polygon method is necessary.

To address the gap mentioned previously, this study combined field samples, GEDI LiDAR data, and multi-sensor imagery under the point-line-polygon modeling framework to acquire an up-to-date map of forest stand volume in CMMFE during 2019. In details, the footprint-level GEDI LiDAR data were served to relate ground-measured points with full-cover imagery from Sentinel-1, Sentinel-2 and ALOS DSM for the large-area mapping. The objectives were to: (1) quantify the relationships between stand volume with variables from LiDAR, C band SAR, MSI and L band InSAR; (2) compare additional integrating GEDI LiDAR data under a point-line-polygon framework with a traditional point-polygon approach based only on multi-sensor imagery; and (3) map forest stand volume by integrating GEDI LiDAR data and multi-sensor imagery to provide managerial suggestions.

2. Material and methods

2.1. The study area

The study area is within CMMFE defined by the World Wide Fund for Nature (WWF) (WWF ID: PA0414). It is located at the border between the mountainous regions of Jilin and Heilongjiang Provinces in northeast China, and covers 204,157 ha (Fig. 1). Due to its four distinct seasons and monsoon-influenced humid continental climate, this region has an annual average temperature and precipitation of 5.14 °C and 636.14 mm, respectively (Mudanjiang Statistical Bureau, 2018; Yanbian Statistical Bureau, 2018). The forest types are mainly natural deciduous broadleaf and mixed broadleaf-conifer woodlands (Olson et al., 2001). Dominant tree species include *Betula costata* (Trautv.), *Tilia amurensis* (Rupr.), *Fraxinus mandschurica* (Rupr.), *Mongolian oak* (*Quercus* spp.), and *Juglans mandshurica* (Maxim.). Dark-brown earth and bog soil are the typical soils.

2.2. Data

2.2.1. Field data

Limited by labor, accessibility, and weather, the field campaign was only carried out within forest farms belonging to the Dunhua Forestry Bureau from May to July of 2019, using a stratified sampling design. Two teams collected measurements following national guidelines for forest resource surveys (MOF, 1982). In total, 1116 samples 25 m in size were measured (Fig. 1a). Stand volume was estimated by the measured tree height and DBH, i.e., 1.3 m from the ground, according to the tree volume tables from the National Standard of China (LY/T 1353–1999) (Forestry Administration of China, 1999).

2.2.2. Multi-sensor data and pre-processing

The multi-sensor satellite data used in this study are listed in Table 1.

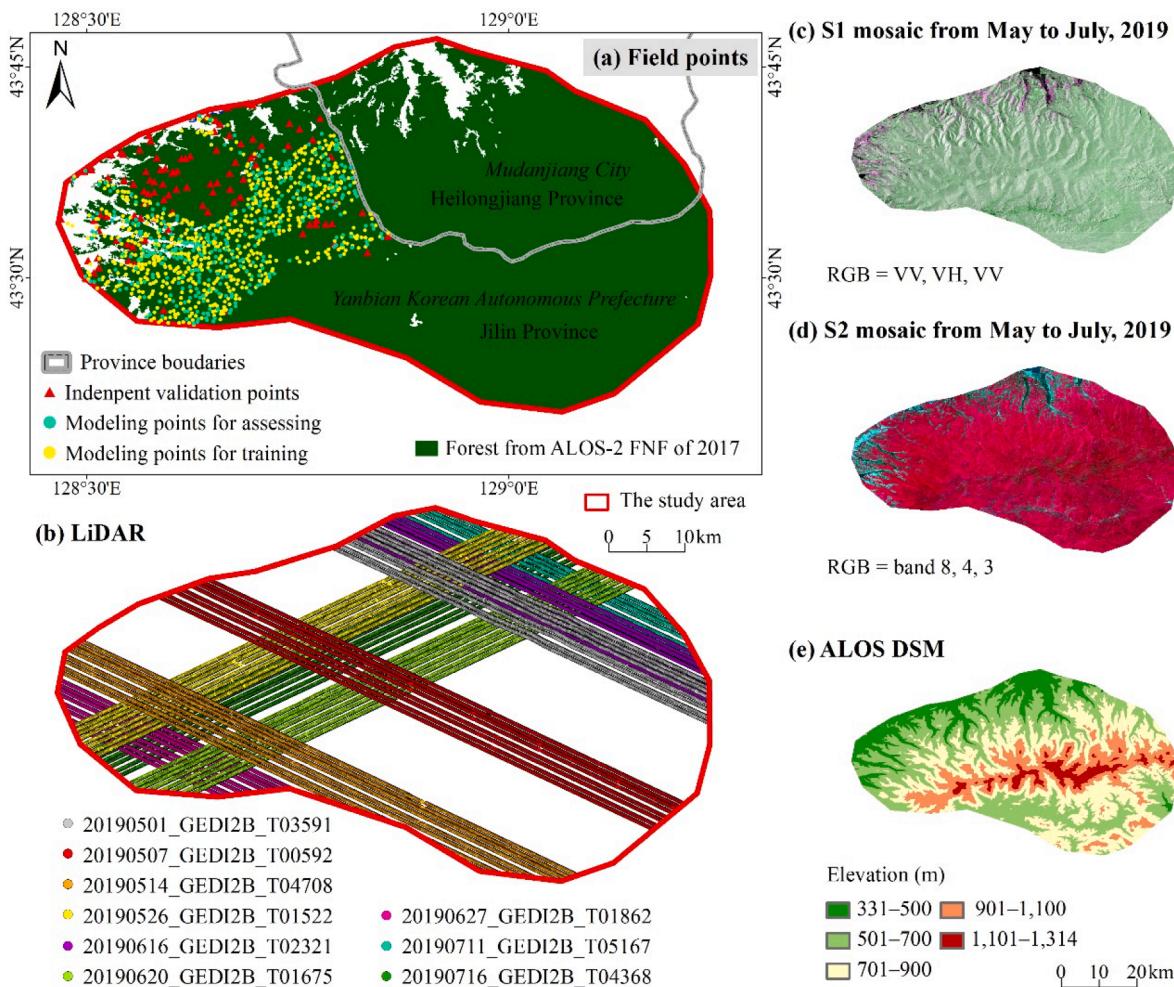


Fig. 1. The study area and ground sampled sites (a), as well as processed remote sensing dataset, including Global Ecosystem Dynamics Investigation (GEDI) Light Detection and Ranging (LiDAR) (b), Sentinel-1 (S1) Synthetic Aperture Radar (SAR) (c), Sentinel-2 (S2) Multispectral Instrument (MSI) L2A (d), Advanced Land Observing Satellite (ALOS) digital surface model (DSM) (e), and ALOS-2 Forest/Non-Forest (FNF) data.

In order to make the time phase of field samples and remote sensing data uniform, the GEDI L2B data from May to July of 2019 were downloaded from Land Processes Distributed Active Archive Center (LP DAAC) to acquire canopy cover and tree height (Figs. 1b and 2), which linearly explained stand volume to a large extent as demonstrated in previous research (Naesset, 1997; Giannico et al., 2016; Xu et al., 2019). GEDI data record the amount of energy returned by various tree components at different heights above the ground and contain structure information, such as surface topography and metrics of canopy height, cover, and vertical profiles (Dubayah et al., 2020; Potapov et al., 2021). The canopy cover and tree height from GEDI L2B (GEDI2B) are based on the directional gap probability profile derived from the L1B waveform (Ni-Meister et al., 2001). A total of 54,829 groups of the canopy cover, elevation of highest detected return (elev_highestreturn), and center of lowest mode (elev_lowestmode) were extracted from GEDI L2B data by rGEDI Package using the R software (Silva et al., 2020). The difference between elev_highestreturn and elev_lowestmode was calculated as a metric of the tree height. After deleting invalid values, 22,350 pairs of canopy cover and tree height were acquired.

The 1014 sampling sites within the valid GEDI lines were selected for stand volume modeling. The remaining 102 points were used for independent validation and accuracy comparison (Fig. 2). Modeling points were randomly divided into 2/3 for training and 1/3 to assess sets for building and evaluating the GEDI-based models.

The mosaic imagery of C band SAR from May to July of 2019 (Fig. 1c) were generated in the Google Earth Engine (GEE) platform by

22 Sentinel-1 images at ground-range-detected level, and included selection, masking, converting to normalized backscatter coefficient, and mosaic (Hird et al., 2017). Based on GEE, a series of 90 Sentinel-2A L2A satellite images were processed to composite median values of the multispectral bands at 10 m spatial resolution by removing various types of clouds and noise, and conducting compositing function (Fig. 1d) (Hird et al., 2017). The DSM data from ALOS and ALOS-2 Forest/Non-Forest (FNF) maps were downloaded from the Japan Aerospace Exploration Agency (Fig. 1a and e). The DSM data were re-sampled into 10 m spatial resolution for topographic indicator extraction (Fig. 2). All pre-processed satellite data were re-projected into the same projection.

2.3. Methods

For improved estimation of stand volume by integration of GEDI data, the workflow of this study, as shown in Fig. 2, contains three major sections: (1) the extraction of forest areas from ALOS-2 FNF maps and Sentinel-2 mosaic images using a hybrid approach, as reported in Section 2.3.1; (2) additionally integrating GEDI LiDAR data by establishing a point-line-polygon framework; and (3) traditional modeling by samples and multi-sensor imagery under the point-polygon approach for accuracy comparison. In the point-line-polygon framework, modeling points for training were linked to GEDI lines by geographically weighted regression (GWR) as presented in Section 2.3.2. Then GEDI lines were used to map wall-to-wall stand volume by random forests (RF) as presented in Section 2.3.3. Additionally, a traditional approach was also

Table 1

The adopted GEDI, S1, S2, ALOS DSM and FNF data.

achieved by RF based on modeling points and multi-sensor imagery. In other words, the RF models used in two frameworks differed in inputs with GEDI lines for the point-line-polygon framework and modeling points for traditional approach but were the same in outputs as stand

volume polygons. The main difference between point-line-polygon framework and traditional approach was that the former inserted the GEDI data as a linear bridge. Both frameworks adopted RF modeling and multi-sensor imagery, i.e., Sentinel-1, Sentinel-2, and ALOS DSM.

2.3.1. A hybrid approach on acquiring forest areas of 2019

A hybrid approach proposed by Wang and Chen (2020) was used to identify forested area, and included object-based image analysis (OBIA), stratified image classification, visual interpretation with manual modification, and dissolve (Fig. 3). This approach effectively accomplishes the change analysis with a consistent classification scheme and evaluates the existing forest cover mapping efforts.

The approach took classes from the ALOS-2 FNF maps, i.e., forest, non-forest and water, as the Phase I data (Fig. 3a). Cloud-free Sentinel-2 mosaic images (Fig. 3b) were used as the Phase II data. The workflow of this hybrid approach involved the segmentation process by OBIA on the eCognition Developer 9.0 software (Blaschke, 2010; Jia et al., 2020; Mao et al., 2020). Scale, shape/color, and compactness/smoothness are defined in the eCognition software to constrain the pixel-growing algorithm (Duro et al., 2012). After a “trial and error” process for testing the parameters, a satisfactory match between the object and forest was achieved, with 150, 0.1, and 0.5 for values of scale, shape, and compactness, respectively (Fig. 3c). A stratified image classification was used to assign classes of objects from Phase II by those from Phase I on eCognition (Fig. 3d). Then, visual interpretation and manual modification were conducted through segmented objects from Phase II data with the help of finer resolution images from Google Earth (Fig. 3e). Afterwards, dissolve was conducted to obtain the final forest map of 2019 at 10 m spatial resolution (Fig. 3f).

2.3.2. Prediction of stand volume lines from GEDI data by GWR

Many studies have demonstrated that stand volume has a generalized linear relationship with canopy cover and tree height using LiDAR data (Means et al., 2000; Tompalski et al., 2014; Bont et al., 2020). However, this relationship of volume with canopy cover and tree height varies by location. Thus, geographically weighted regression (GWR) was used to achieve forest stand volume extraction from GEDI data in this study (Fig. 4).

GWR simulates local relationships based on the locally smooth idea to individually calculate parameters that obeys a distance decay for each location with an unknown volume (Bunson et al., 1996, 1998). In detail, the closer to the location of a measurement, the greater the weight that is assigned (Chen et al., 2018). Locally weighted least square method is used to embed the spatial location of volume measurements into a regression (Fotheringham et al., 2002). The GWR model, shown as Eq. (1), was established for GEDI-based estimation of stand volume using GWR4 software based on training samples by determining necessary parameters, i.e., model and kernel type, and bandwidth selection method and criteria (Nakaya et al., 2014; Ahmed et al., 2017). The modeling accuracy of volume by GWR was evaluated by 338 assessing samples based on the root mean squared error (RMSE), mean error (ME) and coefficient of determination (R^2) (Chen et al., 2019b).

$$\begin{cases} \widehat{SV}_i = \beta_0(u_i, v_i) + \sum_{k=1}^2 \beta_k(u_i, v_i)x_{ik}^* + \varepsilon_i \\ \widehat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) S V \\ W(u_i, v_i) = f(d_i(u_i, v_i), b) \end{cases} \quad (1)$$

where (u_i, v_i) are the location of estimated point i ; $\beta_0(u_i, v_i)$ is the intercept; $\beta_k(u_i, v_i)$ is the coefficient of canopy cover or tree height; x^{*ik} is the value of explanatory variable k at point i , i.e., canopy cover or tree height; ε_i is the error and assumed to be normally distributed with zero mean and constant variance; X is the matrix of x^{*ik} ; SV is the vector of measured volume values; $W(u_i, v_i)$ is a weight matrix determined by $d_i(u_i, v_i)$ and b , including Gaussian and bi-square kernel types with the fixed or

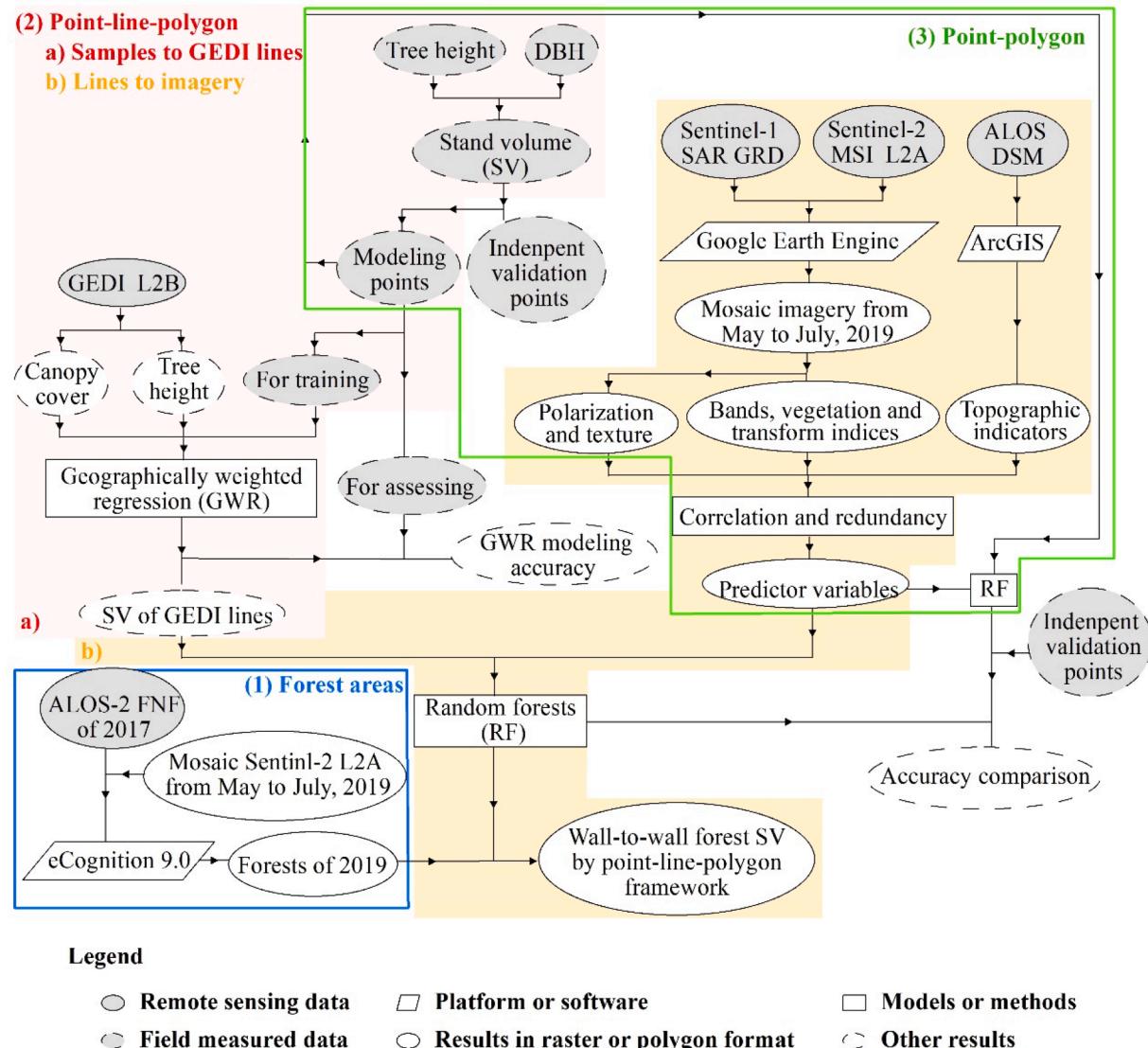


Fig. 2. Overall workflow of the point-line-polygon framework of forest SV estimation, compared with the traditional point-polygon way.

adaptive bandwidth, to ensure that those measurements closer to point i have more influence on the results; $d_i(u_i, v_i)$ is the distance between the measurements i and the location (u_i, v_i) ; and b is the optimal bandwidth for GWR, selected by a method and criterion, which controlled the effect of the distance on the weight value.

2.3.3. Estimation of stand volume polygons by RF

Before RF modeling, remote sensing indices were extracted and filtered to acquire predictor variables. In this study, 53 remote sensing indices were selected and extracted for volume estimation, with 22, 26, and 5 indicators from SAR, MSI, and DSM, respectively (Table 2). Sensitive to tree structure and insensitive to topography, normalized backscatter coefficients and texture characteristics were chosen and extracted from mosaic imagery by Sentinel-1 Toolbox in SNAP software (Dos Reis et al., 2019; Mauya et al., 2019; Chen et al., 2020). Results were promising, demonstrating that band reflectance and spectral indices with red-edge bands from MSI improved forest stand volume estimation (Puliti et al., 2018; Astola et al., 2019). Sentinel-2 Toolbox in SNAP software was used to calculate MSI indices from Sentinel-2 mosaic imagery (Table 2). Based on previous studies, related topographic indicators were extracted from ALOS DSM by Spatial Analyst of ArcGIS software (Wijaya et al., 2010; Chen et al., 2020).

Predictor variables for volume estimation were determined by

pairwise Pearson's correlation analysis after degrading the variable correlations, affecting output accuracy of RF models (Millard and Richardson, 2015). This consisted of two steps performed in SPSS software: the candidate selection of significantly related variables ($p < 0.05$); the candidate disposal by the collinearity ($r \geq 0.8$), with the remaining one most correlated to measured stand volume (Xu et al., 2018a; Chen et al., 2020).

RF has been extensively applied in forest parameter mapping using remote sensing data and has achieved improved accuracy by comparative assessments due to being less sensitive to noise in training samples (Shataee et al., 2012; Fassnacht et al., 2014; Zhao et al., 2019). It was built in WEKA software with two parameters, i.e., number of features for splitting the nodes and number of trees to be optimized (Chen et al., 2019b). It is also capable of identifying the importance of variables based on the mean variance decrease (Wittke et al., 2019). For comparison between the traditional and point-line-polygon frameworks for volume mapping, two groups of variables were input for RF modeling (Fig. 5). Models were compared by 102 independent validation points based on the RMSE, ME, R^2 , and the relative improvement (RI) (Chen et al., 2019b). The wall-to-wall stand volume estimation was the average of the membership. The final map of forest stand volume was obtained by the forest area mask of 2019 (Fig. 5).

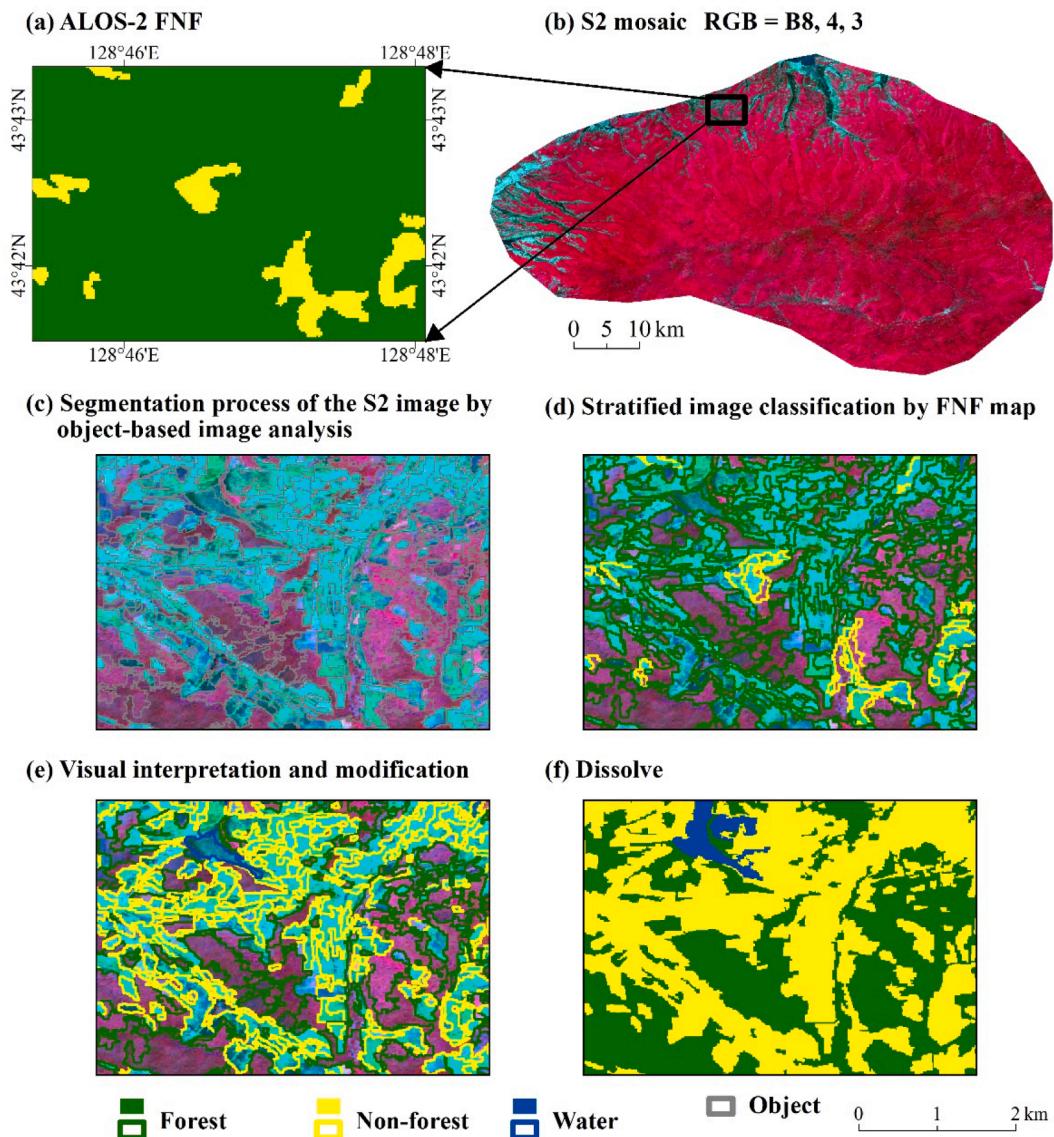


Fig. 3. A hybrid approach based on the ALOS-2 FNF map (a) and S2 image (b). It includes object-based image analysis with 150, 0.1, and 0.5 for values of scale, shape, and compactness, respectively (c), stratified image classification (d), visual interpretation with manual modification (e), and dissolve (f).

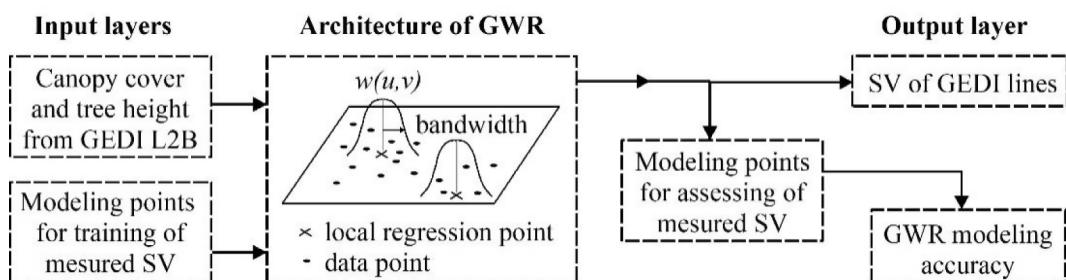


Fig. 4. Illustration of SV extraction from GEDI data by geographically weighted regression (GWR) modeling. $w(u, v)$ is a weight matrix to ensure that those observations more close to the unknown point affect more on the results.

3. Results

3.1. Volume extraction from GEDI data and GWR model

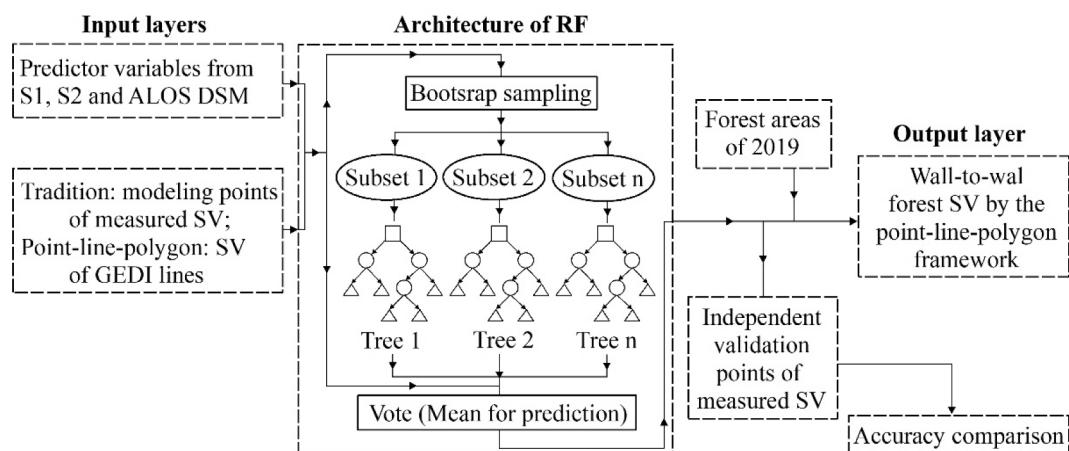
The measured values of forest volume ranged from 5 to 288.63 m³/ha and were mainly below 200 m³/ha (Fig. 6a). The measured values

were divided by the same frequency into six levels for the further comparison of the mapped volume (Fig. 6a). The median and standard deviation (SD) values of measured volume were 151.29 and 47.07 m³/ha, respectively (Fig. 6b). With the increase of elevation, the mean values of measured volume grew to the peak of 162.61 m³/ha, and then decreased to the low of 78.49 m³/ha.

Table 2

Remote sensing indices from the multi-sensor imagery for volume mapping.

Images	Variables	Description
S1 mosaic	Polarization	VV VH Normalized backscatter coefficient of vertical transmit-vertical channel in dB Normalized backscatter coefficient of vertical transmit-horizontal channel in dB
	Texture	VV/VH_CON Contrast VV/VH_DIS Dissimilarity VV/VH_HOM Homogeneity VV/VH_ASM Angular second moment VV/VH_ENE Energy VV/VH_MAX Maximum probability VV/VH_ENT Entropy VV/VH_MEA Gray-level co-occurrence matrix (GLCM) mean VV/VH_VAR GLCM variance VV/VH_COR GLCM correlation
		B2 Blue, 490 nm B3 Green, 560 nm B4 Red, 665 nm B5 Red edge, 705 nm B6 Red edge, 749 nm B7 Red edge, 783 nm B8 Near infrared, 842 nm B8a Near infrared, 865 nm B11 Short-wave infrared, 1610 nm B12 Short-wave infrared, 2190 nm
	Vegetation indices	RVI Ratio vegetation index, B8/B4 DVI Difference vegetation index, B8–B4 PVI Perpendicular vegetation index, $\sin(45^\circ) \times B8 - \cos(45^\circ) \times B4$ NDVI Normalized difference vegetation index, $(B8 - B4)/(B8 + B4)$ SAVI Soil adjusted vegetation index, $1.5 \times (B8 - B4)/(B8 + B4 + 0.5)$ NDVI5 Normalized difference vegetation index with bands 4 and 5, $(B5 - B4)/(B5 + B4)$ NLI5 Non-linear vegetation index with bands 4 and 5, $(B52 - B4)/(B52 + B4)$ NDVI6 Normalized difference vegetation index with bands 4 and 6, $(B6 - B4)/(B6 + B4)$ NDVI7 Normalized difference vegetation index with bands 4 and 7, $(B7 - B4)/(B7 + B4)$ NDVI8a Normalized difference vegetation index with bands 4 and 8a, $(B8a - B4)/(B8a + B4)$ MSI Moisture stress index, B8/B11 EVI5 Enhanced vegetation index with bands 4, 5 and 2, $2.5 * (B5 - B4)/(B5 + 6 * B4 - 7.5 * B2 + 1)$ S2REP Sentinel-2 red-edge position index, $705 + 35 \times [(B4 + B7)/2 - B5] \times (B6 - B5)$
	Transform indices	TCW Tasseled cap wetness, $0.1509 * B2 + 0.1973 * B3 + 0.3279 * B4 + 0.3406 * B8 + 0.7112 * B11 + 0.4572 * B12$ TCB Tasseled cap brightness, $0.3037 * B2 + 0.2793 * B3 + 0.4743 * B4 + 0.5585 * B8 + 0.5082 * B11 + 0.1863 * B12$ TCG Tasseled cap greenness, $-0.2848 * B2 - 0.2435 * B3 - 0.5436 * B4 + 0.7243 * B8 + 0.0840 * B11 - 0.1800 * B12$
DSM	Topographic indicators	H Elevation S Slope A Aspect M Surface roughness SPI Stream power index, $\ln[Ac \times \tan\beta \times 100]$

**Fig. 5.** Delineation of SV estimation from ALOS and Sentinel series imagery by random forests (RF) modeling.

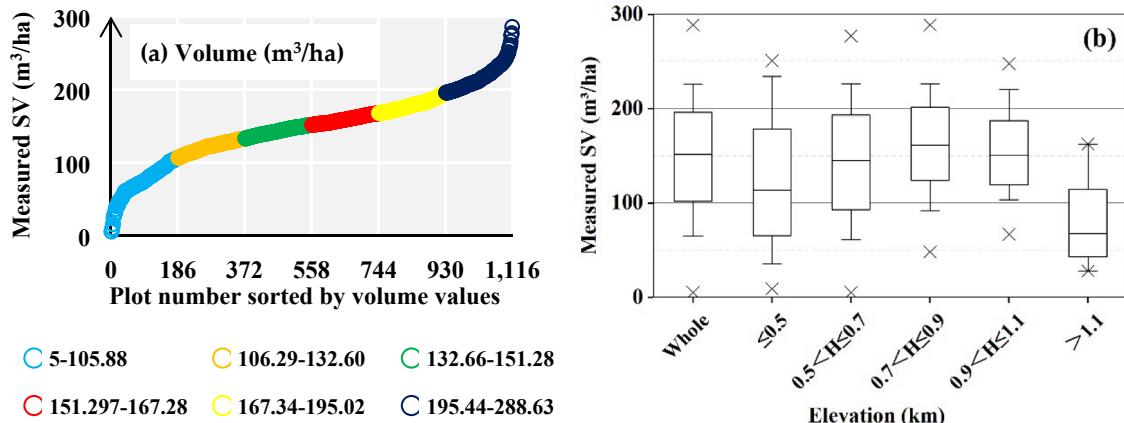


Fig. 6. Measured profiles of stand volume (SV) samples in the study site from Plot 1 to 111,6 (a) and distribution of SV values with different elevation levels (b). The boxes depict values within the range of the mean \pm SD, lines in boxes are median, squares denote means, the dash is the whisker of 5–95%, and crosses show the minimum and maximum values.

The GWR model built by the 676 training samples used the Gaussian approach, where the weight function was a fixed Gaussian kernel. This model found the optimal bandwidth was 0.016 by using a golden selection method and corrected Akaike information criterion (AICc, small sample bias corrected AIC). The accuracy, evaluated by the 338 assessing samples, achieved an RMSE, ME, and R^2 of 27.27 m³/ha, -1.92 m³/ha, and 0.81, respectively. The absolute mean values of coefficients for canopy cover of 21.9 generally showed a stronger ability to explain the relationship with volume than tree height at 0.21. The values of coefficient of variation (CV) indicated that the effects of tree height (CV = 14.3) had a stronger variability than that of canopy cover (CV = 3.3). The volume values extracted from 22,350 pairs of canopy cover and tree height by GWR ranged from 45.26 to 280.57 m³/ha. The median and SD values of measured volume were 150.78 and 22.36 m³/ha, respectively (Table 3). Overall, stand volume derived from GEDI lines has similar distribution to the measured value (Fig. 6).

3.2. Predictor variables

A total of 45 variables had a significant relationship with stand volume (Table 4), including 16 from SAR, 25 from MSI, and four from DSM. The VV backscatters had positive correlations with the measured volume. The GLCM mean of VV and contrast of VH were most related to volume among 10 kinds of texture features from SAR. Namely, the growth of texture regularity and smoothness of VV and VH backscatters indicated an increasing stand volume. It was also demonstrated that texture features from VV were more relevant to volume than those from VH.

As for MSI variables, B2, B3, B4, B5, B11, B12, TCW, and TWB were negatively related to volume, while the remaining 17 variables showed a positive correlation. The reflectance and spectral indices that were involved in featured red-edge bands of Sentinel-2 signified a close connection to the measured volume. MSI contributed more sensitive

variables for volume estimation than C band SAR.

All DSM-based topographic indicators displayed a strong positive influence on the growth of measured volume ($p < 0.01$), except SPI ($p < 0.05$). It was revealed that variables from DSM had similar performances with those from MSI. The predictor variables input in volume modeling were the 17 predictors marked “Yes” in Table 4, which were significantly relevant to volume ($p < 0.05$) without redundancy ($r < 0.8$). Overall, the VV texture characteristics of SAR, reflectance, and spectral indices from MSI, and DSM-derived elevation were comparatively important for volume mapping.

3.3. Volume estimation by RF models

With 300 trees and the number of features set to four, the RF models based on point-line-polygon and traditional frameworks were built separately. The results showed that the attribute importance ranking of the RF model acquired by point-line-polygon and traditional frameworks was similar (Fig. 7). Generally, H, S2REP, B12, PVI, and S were the most important variables for volume modeling. Topographic indicators, reflectance, and spectral indices displayed more influences than texture features from VV channel. The impacts of the VV backscatter and texture features from VH channel were marginal.

Table 5 showed the accuracy of the estimated volume from point-line-polygon and traditional models through 102 validation samples. To better evaluate and compare the accuracy, the mean measured value of volume (148.78 m³/ha in Fig. 6) served to divide the ME and RMSE. ME values revealed that both two models overestimated forest volume. The comparison demonstrated that integration of GEDI LiDAR data as a linear bridge between the sample points and remote sensing imagery was more accurate than directly modeling relationship between measured volume and images (Fig. 8). Combining GEDI lines remarkably improved accuracy of volume modeling by 31% (6.87 m³/ha) according to RMSE values.

3.4. Forest stand volume of CMMFE

After implementing the hybrid approach, forest area of 2019 was acquired. The distribution of forest volume based on the point-line-polygon framework was mapped by the RF model mentioned above, and then non-forest areas, i.e., water and non-forest, were masked out (Fig. 9a). To compare with field-measured values, estimated values of forest stand volume were displayed at six levels. Each level had an equal number of measured sample sites (Fig. 6a). Estimated stand volume was close to field measured values, which was illustrated by the same pattern at each level (Fig. 9a). Under the point-line-polygon framework, the

Table 3

Description statistics of the GEDI-based canopy cover, tree height and stand volume.

Variables	Minimum	Maximum	Mean	Median	Standard Deviation (SD)
Canopy cover	0.01	0.98	0.35	0.27	0.33
Tree height (m)	1.46	45.78	12.17	10.15	7.39
Stand volume	45.26	280.57	146.81	150.78	22.36

Table 4

The selection result of predictor variables derived from multi-sensor imagery with a *p*-value of the *t*-test being below 0.05 marked '*' and *p*-value below 0.01 marked '**'.

Images	Related variables	<i>r</i>	Collinear with	Predictors
S1 mosaic	VV	0.07*	/	Yes
	VV_Dis	-0.10**	VV_Hom, VV_Ent	Yes
	VV_Hom	0.14**	VV_Dis, VV_ASM, VV_Ene, VV_MAX, VV_Ent	No
	VV_ASM	-0.13**	VV_Hom, VV_Ene, VV_MAX, VV_Ent	No
	VV_Ene	0.14**	VV_Hom, VV_ASM, VV_MAX, VV_Ent	Yes
	VV_MAX	-0.14**	VV_Hom, VV_ASM, VV_Ene, VV_Ent	No
	VV_Ent	-0.13**	VV_Dis, VV_Hom, VV_ASM, VV_Ene, VV_MAX	No
	VV_Mea	0.22**	VV_Var	Yes
	VV_Var	0.15**	VV_Mea	No
	VV_Cor	0.19**	/	Yes
	VH_Con	-0.13**	VH_Dis	Yes
	VH_Dis	-0.10**	VH_Con, VH_Hom, VH_Ent	No
	VH_Hom	0.08*	VH_Dis, VH_Ent	No
	VH_Ent	-0.09**	VH_Dis, VH_Hom	Yes
	VH_Mea	0.07*	/	Yes
	VH_Cor	0.10**	/	Yes
S2 mosaic	B2	-0.11**	B3, B4, B5, B11, B12, RVI, NDVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a, MSI	No
	B3	-0.10**	B2, B4, B5, B11, RVI, NDVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a, MSI	Yes
	B4	-0.22**	B2, B3, B5, B11, B12	No
	B5	-0.10**	B2, B3, B4, B11, B12, RVI, NDVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a, MSI, TCW	No
	B6	0.23**	B7, B8, B8a, DVI, S2REP, TCG	No
	B7	0.25**	B6, B8, B8a, DVI, S2REP, TCG	No
	B8	0.24**	B6, B7, B8a, DVI, S2REP, TCG	No
	B8a	0.25**	B6, B7, B8, DVI, S2REP, TCG	No
	B11	-0.15**	B2, B3, B4, B5, B12, TCW, TCB	No
	B12	-0.19**	B2, B4, B5, B11	Yes
	RVI	0.71*	B2, B3, B5, NDVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a, MSI	No
	DVI	0.28**	B6, B7, B8, B8a, S2REP, TCG	No
	PVI	0.17**	/	Yes
	NDVI	0.07*	B2, B3, B5, RVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a, MSI	No
	SAVI	0.07*	B2, B3, B5, RVI, NDVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a, MSI	No
	NDVI5	0.07*	B2, B3, B5, RVI, SAVI, NDVI, NLIS, NDVI6, NDVI7, NDVI8a, MSI	No
	NLIS	0.07*	B2, B3, B5, RVI, SAVI, NDVI, NDVI5, NDVI6, NDVI7, NDVI8a, MSI	No
	NDVI6	0.07*	B2, B3, B5, RVI, NDVI, SAVI, NDVI5, NLIS, NDVI7, NDVI8a, MSI	No
	NDVI7	0.07*	B2, B3, B5, RVI, NDVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a, MSI	No
	NDVI8a	0.07*	B2, B3, B5, RVI, NDVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, MSI	No
	MSI	0.07*	B2, B3, B5, RVI, NDVI, SAVI, NDVI5, NLIS, NDVI6, NDVI7, NDVI8a	No
	S2REP	0.29**	B6, B7, B8, B8a, DVI, TCG	Yes
	TCW	-0.11**	B5, B11, TCB	No
	TCB	-0.19**	B11, TCW	Yes

Table 4 (continued)

Images	Related variables	<i>r</i>	Collinear with	Predictors
DSM	TCG	0.28**	B6, B7, B8, B8a, DVI, S2REP	No
	H	0.27**		Yes
	S	0.14**	M	Yes
	M	0.10**	S	No
	SPI	0.06*		Yes

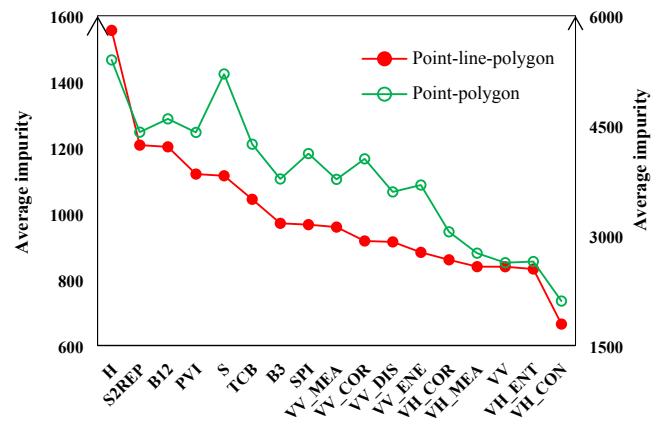


Fig. 7. Attribute importance of RF models for the forest stand volume estimation.

Table 5

Accuracy assessment of RF modeling on stand volume according to independent validation data.

Frameworks	ME		RMSE		<i>R</i> ²	RI
	m ³ /ha	%	m ³ /ha	%		
Point-line-polygon	-7.03	-4.72	22.63	15.21	0.88	0.31
Point-polygon	-9.76	-6.56	32.85	22.08	0.80	/

forest stand volume in the study area was estimated from 47.56 to 277.30 m³/ha, with mean and SD values of 151.39 and 33.12 m³/ha, respectively.

The middle region of the study area with the highest elevation (>1100 m) had the smallest mean value of 99.89 m³/ha, with volume ranging from 65.50 to 213.58 m³/ha (Fig. 9). In the north area with medium altitude (700 < *H* ≤ 900 m), the largest mean value was 155.16 m³/ha, with volume ranging from 47.56 to 277.30 m³/ha. Overall, the distribution of estimated forest stand volume was near the measured volume (Figs. 6b and 9b). The major errors were overestimation of small values in the low altitude (≤ 700 m) area and underestimation of large volume in the medium elevation region.

4. Discussion

4.1. Role of multi-sensor variables on volume estimation

The absolute values of coefficients of canopy cover and tree height demonstrated the role of LiDAR variables on stand volume estimation. However, the coefficients varied by location. Among the 22,350 pairs of GEDI data, the absolute values of the majority of coefficients of canopy cover were larger than those of tree height, while only 913 coefficients of tree height, which located in the higher altitude (>700 m), have larger absolute values. This indicated that the spatial variation of stand volume was greatly influenced by canopy cover from LiDAR in the study area, whereas in the higher altitude area tree height was much more influence on stand volume. This may be caused by changes of vegetation

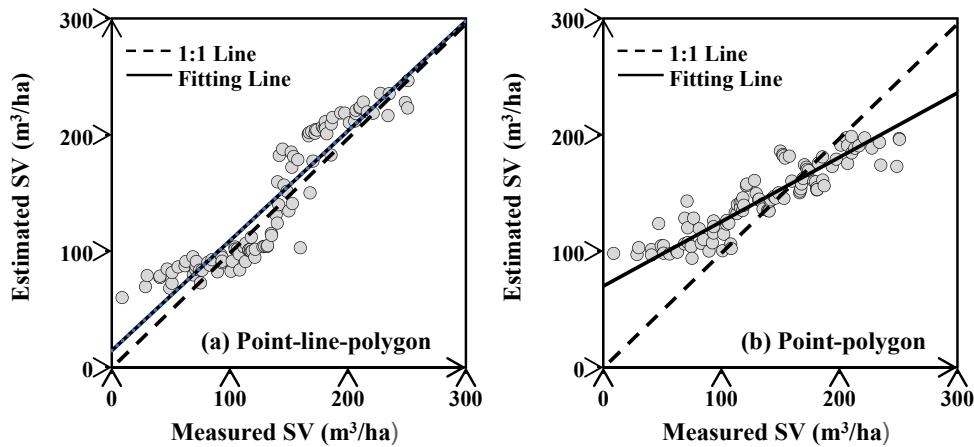


Fig. 8. Scatter plots of estimated versus measured SV from validation samples.

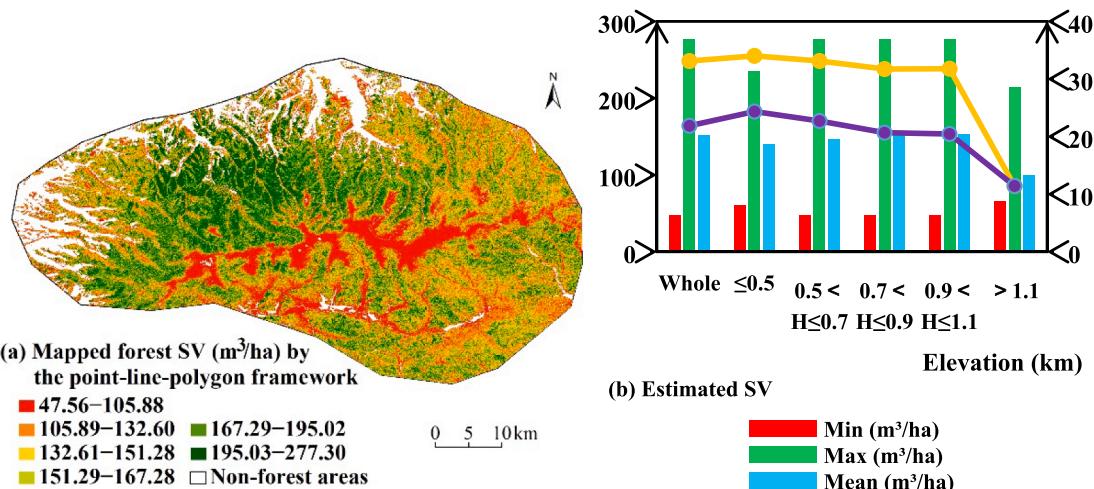


Fig. 9. The map (a) and values alone different levels of elevation (b) of estimated forest SV by the point-line-polygon framework derived from GEDI LiDAR data, ALOS and Sentinel series imagery. SD and CV denote standard deviation and coefficient of variation, respectively.

in vertical zones along elevation of the Changbai Mountains Mixed Forests. In fact, with elevation increasing, stand density decreased and coniferous forests became dominant (Wang et al., 2011), leading to greater influence of tree height on volume accumulation than canopy cover.

The importance of SAR and MSI variables on volume estimation was revealed by correlation coefficients (Table 4) and attribute importance of RF models. The elevation, as the proxy of L band InSAR height, was of prime importance on volume estimation in this study. This could be due to InSAR height being directly proportional to volume (Gama et al., 2010; Solberg et al., 2013). Stand volume of CMMFE showed variations along the elevation gradient (Fig. 6 and Fig. 9), which is consistent with previous studies on Changbai flora regions (Li et al., 2011; Chen et al., 2020). The mature *Mongolian oak* (*Quercus* spp.) dominant natural forest site in the lower altitude region of the study area had the largest measured volume of 288.63 m³/ha. This may be explained by the negative effect of temperature on the productivity of *Mongolian oak* forests (Wu et al., 2019). The spatial variation of stand volume was smaller than reports on nearby areas (Wang et al., 2011; Chen et al., 2019a). The coarse resolution in this highly heterogeneous landscape and smaller changes in altitude were the reasons for smaller variation in estimated volume. The stand volume of forests in the southeastern part of the study area, where surface elevation changed dramatically, had greater spatial variations (Figs. 1e and 9a). The close relation between

elevation and forest stand volume in the study area was mainly because moisture, temperature, and plant diversity changed remarkably with altitude in the Changbai Mountain areas (Guo et al., 2014; Shen et al., 2014; Cong et al., 2019).

MSI variables have strong ability to retrieve vegetation types and horizontal structures such as canopy cover and DBH, particularly the reflectance from red-edge bands as a proxy of the chlorophyll content, leading to the great influence on volume mapping in this study and other research (Peña et al., 2012; Lausch et al., 2017). The collinearity and redundancy of MSI variables were noticeable, which showed that the selection of predictors was essential to volume mapping based on optical sensors. Backscatters from SAR sensors were related to roughness and water content of vegetation as well as stand volume owing to penetrability (Lu et al., 2016; Chen et al., 2020). However, the influence of backscatters from Sentinel-1 on stand volume mapping was marginal. This resulted from the saturation problem of C band SAR. The measured values of stand volume were partially above 200 m³/ha in the study area (Fig. 6), which was greater than the conventional saturation value of C band backscatters (Santi et al., 2015). The results showed that texture characteristics of SAR imagery were much more beneficial to volume estimation than original backscatters especially that of VV channel, owing to the reduction of impacts to the heterogeneity (Dos Reis et al., 2019; Morin et al., 2019).

The comparison of volume estimation based on variables from four

sensors denoted that LiDAR data performed better than other imagery ($\text{RMSE}_{\text{LiDAR-lines}} = 27.27 \text{ m}^3/\text{ha}$, $\text{RMSE}_{\text{traditional}} = 32.85 \text{ m}^3/\text{ha}$). This is because LiDAR directly measures the distribution of vegetation along the vertical axis and provides three-dimension structure features (Lefsky et al., 1999; Lin, 2019). These characteristics result in a precise assessment of the total volume and spatial organization of vegetation material as well as a larger saturation value (Patenaude et al., 2004; Xu et al., 2018b). However, the coarse temporal resolution of GEDI LiDAR data leads to a lack of coverage of the study area in the same phenological phase, which may be improved by the combination with airborne LiDAR data in future works. In short, canopy cover and tree height from LiDAR, topographic indices from L band InSAR, and spectral indices of red-edge band from MSI were recommended for stand volume estimation based on the satellite data at a large-scale in heterogeneous forests.

4.2. Point-line-polygon versus traditional frameworks

The RMSE of estimation by the traditional approach was 22.08%, which was at the average level of published regional maps of volume using the same point-polygon method (Kilpeläinen and Tokola, 1999; Chirici et al., 2008; Immitzer et al., 2016; Xu et al., 2018a; Dos Reis et al., 2019; Liu et al., 2019b; Hawrylo et al., 2020). In this study, the traditional model built by imagery from Sentinel-1 C band SAR, Sentinel-2 MSI, and ALOS DSM acquired more accurate volume than by data from single optical sensors such as Landsat and RapidEye (Mäkelä and Pekkarinen, 2004; Xu et al., 2018a). This can be explained by the higher resolution of Sentinel series and advantages of using multi-sensor imagery. However, previous traditional models which added LiDAR-derived predictors obtained a higher accuracy than the point-polygon model of this study (Xu et al., 2018a; Hawrylo et al., 2020). It was also revealed that integration of LiDAR-derived predictors can reduce the saturation problem from optical and C band SAR sensors. However, limited by costs, integration of air-borne LiDAR data with imagery in traditional models has rarely been reported for spatially continues volume mapping at a regional scale, more commonly adopting a point-level comparison (Xu et al., 2018a; Hawrylo et al., 2020).

The point-line-polygon model integrated space-borne LiDAR data and multi-sensor imagery in two steps. Moreover, because of the coarser spatiotemporal resolution of GEDI LiDAR data than Sentinel and ALOS imagery, it is difficult to integrate all sensors using a traditional approach for volume mapping. The assessment result of the point-line-polygon model was more accurate than that of the traditional model (Table 5 and Fig. 8), implying that the integration of GEDI data and multi-sensor imagery is suitable for large-scale forest stand volume mapping. Fig. 8 depicted that the error of the traditional model of underestimating large values and overestimating small volume was reduced. Possibly because stand volume lines extracted from GEDI data as the bridge reduce the saturation problem and heterogeneity, better matching the predictors from satellite remote sensing imagery. Thus, the accuracy improvement of volume extraction from GEDI lines is crucial in future studies. There were 676 actual samples used in the modeling process under the point-line-polygon framework. However, modeling samples of the traditional model were 1014. This inferred that the point-line-polygon framework could lessen workloads of field sampling in comparison to the traditional framework. Limited by field sampling area and temporal resolution of GEDI data, mapped results still had issues with estimation of the whole range of stand volume, especially the minimum and maximum levels of volume values. Furthermore, the RF algorithm only predicts values within the range covered by the measured samples. Thus, the representativeness and accuracy of stand volume extracted from GEDI lines were in high demand for RF modeling. In the future work, other mapping algorithms, e.g., Cubist, and support vector machine, can be used to attempt mapping of stand volume to upgrade precision.

This study concluded that integration of GEDI data and multi-sensor imagery under the point-line-polygon framework was a promising

methodology for mapping of forest stand volume. It should be noted that volume extraction from GEDI lines and mapping algorithm were integral for determining accuracy.

4.3. Uncertainty and forest management

Uncertainty is considered a major issue associated with the remotely sensed volume. It is derived from in-situ measurements, calculation of measured volume, remote sensing predictors, and modeling. Tree height measurements and tree volume tables contribute to reference value uncertainty of forest stand volume in this study. This uncertainty can be reduced by the repeated measurements and updated volume tables in future work. Additionally, the uncertainty of multi-sensor derived variables was considered in this study. Their uncertainty was reduced by similar acquisition time among variables and also with field data. Abundant volume-related variables articulated in published papers (Table 2) were also extracted and filtered to find suitable predictors in this study. The sampling size matched the spatial resolution of GEDI data, but was inconsistent with the multi-sensor imagery, increasing the uncertainty. The suitable sampling size should be explored further. The uncertainty of modeling was shown as the RMSE value in Part 4.2. Accurate estimation of stand volume from GEDI lines was essential in reducing the uncertainty of modeling. Various methods of spatial prediction and GEDI-based variables will be adopted for future mapping of forest stand volume.

For sustainable forest management, certain measures should be taken based on spatial variations of stand volume. Medium-altitude forests with high stand density should be thinned to increase space and resources. Low-altitude forests should be enclosed for cultivation. Increment felling is urgent in smaller volume regions. The advanced reproduction as well as the man-made promotion of natural regeneration need be conducted at a high-altitude ($\geq 700 \text{ m}$).

5. Conclusions

The principal merit of this study is the first mapping of forest stand volume in the CMMFE of northeast China from field samples, GEDI LiDAR data, and multi-sensor imagery of Sentinel-1, Sentinel-2, and ALOS DSM, under a point-line-polygon framework. Specifically, partial-coverage GEDI LiDAR lines acted as a bridge to relate field plots and full-cover multi-sensor imagery.

The results indicated that additional integration of GEDI LiDAR data under a point-line-polygon framework ($R^2 = 0.88$ and $\text{RMSE} = 15.21\%$) improved accuracy by 31% compared with the traditional point-polygon method, which linked field samples to multi-sensor imagery directly ($R^2 = 0.80$ and $\text{RMSE} = 22.08\%$). Canopy cover and tree height from GEDI LiDAR, elevation from L band InSAR, and spectral indices of MSI red-edge bands were crucial for stand volume estimation in heterogeneous temperate forests. Results also demonstrated that the integration of LiDAR by a point-line-polygon framework adopted 2/3 modeling points but acquired the more accurate estimation than a traditional approach based only on multi-sensor imagery, implying less field sampling work was needed for similar research. Overall, integration of GEDI LiDAR data and multi-sensor imagery produced a more accurate mapping of forest stand volume with less field measurements.

The results illustrated that stand volume of CMMFE showed variations along the elevation gradient, which ranged from 47.56 to 277.30 m^3/ha . The forests in the middle of the study area with the highest altitude had the smallest mean value of $99.89 \text{ m}^3/\text{ha}$. Whereas the northern forests with medium altitude had the largest mean value of $155.16 \text{ m}^3/\text{ha}$. As the pioneering exploration of stand volume mapping by the integration of GEDI LiDAR data and multi-sensor imagery under the point-line-polygon framework, this study provides other researchers an informative methodology on carbon changes and supports decision makers in responding to global warming in the CMMFE.

CRediT authorship contribution statement

Lin Chen: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Writing - original draft. **Chunying Ren:** Conceptualization, Funding acquisition, Supervision, Project administration, Resources, Writing - review & editing. **Bai Zhang:** Conceptualization, Supervision, Writing - review & editing. **Zongming Wang:** Supervision, Funding acquisition, Writing - review & editing. **Mingyue Liu:** Investigation, Software, Validation. **Weidong Man:** Investigation, Software, Validation, Funding acquisition. **Jiafu Liu:** Funding acquisition, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the colleagues who participated in the field surveys and data collection. We thank the National Earth System Science Data Center (<http://www.geodata.cn>) for providing geographic information data. This study is supported by National Key Research and Development Project of China, China (No. 2016YFC0500300), Start-up Foundation of Hangzhou Normal University, China (No. 4085C50220204092), Jilin Scientific and Technological Development Program, China (No. 20170301001NY), National Natural Science Foundation of China, China (No. 41977411), the funding from Youth Innovation Promotion Association of Chinese Academy of Sciences, China (No. 2017277, 2012178), and Nature Science Foundation of Hebei Province, China (No. D2019209317), as well as National Earth System Science Data Center of China.

References

- Ahmadi, K., Kalantar, B., Saeidi, V., Harandi, E.K.G., Janizadeh, S., Ueda, N., 2020. Comparison of machine learning methods for mapping the stand characteristics of temperate forests using multi-spectral Sentinel-2 data. *Remote Sens.* 12, 3019.
- Ahmed, M.A.A., Abd-Elrahman, A., Escobedo, F.J., Cropper, W.P., Martin, T.A., Timilsina, N., 2017. Spatially-explicit modeling of multi-scale drivers of aboveground forest biomass and water yield in watersheds of the Southeastern United States. *J. Environ. Manag.* 199, 158–171.
- Astola, H., Häme, T., Sirro, L., Molinier, M., Kilpi, J., 2019. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. *Remote Sens. Environ.* 223, 257–273.
- Blaschke, T., 2010. Object-based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* 65, 2–16.
- Bont, L.G., Hill, A., Waser, L.T., Bürgi, A., Ginzler, C., Blattert, C., 2020. Airborne-laser-scanning-derived auxiliary information discriminating between broadleaf and conifer trees improves the accuracy of models for predicting timber volume in mixed and heterogeneously structured forests. *Forest Ecol. Manag.* 459, 117856.
- Boisvenue, C., Smiley, B.P., White, J.C., Kurz, W.A., Wulder, M.A., 2016. Integration of Landsat time series and field plots for forest productivity estimates in decision support models. *For. Ecol. Manag.* 376, 284–297.
- Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* 28, 281–298.
- Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1998. Geographically Weighted Regression—Modelling Spatial Non-stationarity. In: Workshop on Local Indicators of Spatial Association; University of Leicester: Leicester, UK, pp. 431–443.
- Cartus, O., Kellndorfer, J., Rombach, M., Walker, W., 2012. Mapping canopy height and growing stock volume using airborne Lidar, ALOS PALSAR and Landsat ETM+. *Remote Sens.* 4, 3320–3345.
- Chen, L., Ren, C., Zhang, B., Wang, Z., 2020. Multi-sensor prediction of stand volume by a hybrid model of support vector machine for regression kriging. *Forests* 11, 296.
- Chen, L., Ren, C., Zhang, B., Wang, Z., Xi, Y., 2018. Estimation of forest above-ground biomass by geographically weighted regression and machine learning with Sentinel imagery. *Fores* 9, 582.
- Chen, L., Ren, C., Zhang, B., Wang, Z., Wang, Y., 2019a. Mapping spatial variations of structure and function parameters for forest condition assessment of the Changbai Mountain National Nature Reserve. *Remote Sens.* 11, 3004.
- Chen, L., Wang, Y., Ren, C., Zhang, B., Wang, Z., 2019b. Assessment of multi-wavelength SAR and multispectral instrument data for forest aboveground biomass mapping using random forest kriging. *Forest Ecol. Manag.* 447, 12–25.
- Chirici, G., Barbati, A., Corona, P., Marchetti, M., Travaglini, D., Maselli, F., Bertini, R., 2008. Non-parametric and parametric methods using satellite images for estimating growing stock volume in alpine and Mediterranean forest ecosystems. *Remote Sens. Environ.* 112, 2686–2700.
- Chirici, G., Giannetti, F., McRoberts, R.E., Travaglini, D., Pecchi, M., Maselli, F., Chiesi, M., Corona, P., 2020. Wall-to-wall spatial prediction of growing stock volume based on Italian National Forest Inventory plots and remotely sensed data. *Int. J. Appl. Earth Obs. Geoinformation* 84, 101959.
- Chirici, G., Mura, M., McInerney, D., Py, N., Tomppo, E.O., Waser, L.T., Travaglini, D., McRoberts, R.E., 2016. A meta-analysis and review of the literature on the k-nearest neighbors technique for forestry applications that use remotely sensed data. *Remote Sens. Environ.* 176, 282–294.
- Chrysafis, I., Mallinis, G., Tsakiri, M., Patias, P., 2019. Evaluation of single-date and multi-seasonal spatial and spectral information of Sentinel-2 imagery to assess growing stock volume of a Mediterranean forest. *Int. J. Appl. Earth Obs. Geoinformation* 77, 1–14.
- Condés, S., McRoberts, R.E., 2017. Updating national forest inventory estimates of growing stock volume using hybrid inference. *Forest Ecol. Manag.* 400, 48–57.
- Cong, Y., Li, M., Liu, K., Dang, Y., Han, H., He, H., 2019. Decreased temperature with increasing elevation decreases the end-season leaf-to-wood reallocation of resources in deciduous *Betula ermanii* Cham. *Trees For.* 10, 166.
- de Souza, G.S.A., Soares, V.P., Leite, H.G., Gleriani, J.M., do Amaral, C.H., Ferraz, A.S., de Freitas Silveira, M.V., dos Santos, J.F.C. Velloso, S.G.S., Domingues, G.F., Silva, S., 2019. Multi-sensor prediction of Eucalyptus stand volume: a support vector approach. *ISPRS J. Photogramm. Remote Sens.* 156, 135–146.
- Dos Reis, A.A., Franklin, S.E., de Mello, J.M., Junior, F.W.A., 2019. Volume estimation in a Eucalyptus plantation using multi-source remote sensing and digital terrain data: a case study in Minas Gerais State, Brazil. *Int. J. Remote Sens.* 4, 2683–2702.
- Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurt, G., Kellner, J., Luthcke, S., Armstrong, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marsalis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The Global Ecosystem Dynamics Investigation: high-resolution laser ranging of the Earth's forests and topography. *Sci. Remote Sens.* 1, 100002.
- Dube, T., Mutanga, O., Abdel-Rahman, E.M., Ismail, R., Slotow, R., 2015. Predicting *Eucalyptus* spp. stand volume in Zululand, South Africa: an analysis using a stochastic gradient boosting regression ensemble with multi-source data sets. *Int. J. Remote Sens.* 36, 3751–3772.
- Duro, D.C., Franklin, S.E., Dube, M.G., 2012. Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. *Int. J. Remote Sens.* 33, 4502–4526.
- Fassnacht, F.E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P., Koch, B., 2014. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens. Environ.* 154, 102–114.
- Fazakas, Z., Nilsson, M., Olsson, H., 1999. Regional forest biomass and wood volume estimation using satellite data and ancillary data. *Agr. For. Meteorol.* 98–99, 41–7425.
- Forestry Administration of China, 1999. Tree Volume Tables (National standard # LY/T 1353-1999); Forestry Administration of China: Beijing, China.
- Fotheringham, A.S., Brunsdon, C., Charlton, M.E., 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley, Chichester, UK.
- Gama, F.P., dos Santos, J.R., Mura, J.C., 2010. Eucalyptus biomass and volume estimation using interferometric and polarimetric SAR data. *Remote Sens.* 2, 939–956.
- Giannico, V., Laforteza, R., John, R., Sanesi, G., Pesola, L., Chen, J., 2016. Estimating stand volume and above-ground biomass of urban forests using LiDAR. *Remote Sens.* 8, 339.
- Guo, D., Zhang, H., Hou, G., Zhao, J., Liu, D., Guo, X., 2014. Topographic controls on alpine treeline patterns on Changbai Mountain, China. *J. Mt. Sci.* 11, 429–441.
- Hawrylo, P., Francini, S., Chirici, G., Giannetti, F., Parkitna, K., Krok, G., Metelsztedt, K., Lisiarczuk, M., Sterenczak, K., Ciesielski, M., Wezyk, P., Socha, J., 2020. The use of remotely sensed data and polish NFI plots for prediction of growing stock volume using different predictive methods. *Remote Sens.* 12, 3331.
- Hawrylo, P., Wezyk, P., 2018. Predicting growing stock volume of Scots Pine stands using Sentinel-2 satellite imagery and airborne image-derived point clouds. *Forests* 9, 274.
- Hird, J.N., DeLancey, E.R., McDermid, G.J., Kariyeva, J., 2017. Google Earth Engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping. *Remote Sens.* 9, 1315.
- Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., Hunsaker, C., 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sens. Environ.* 102, 63–73.
- Hyppä, J., Hyppä, H., Inkkinen, M., Engdahl, M., Linko, S., Zhu, Y., 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecol. Manag.* 128, 109–120.
- Immitzer, M., Stepper, C., Böck, S., Straub, C., Atzberger, C., 2016. *Forest Ecol. Manag.* 359, 232–246.
- Jia, M., Mao, D., Wang, Z., Ren, C., Zhu, Q., Li, X., Zhang, Y., 2020. Tracking long-term floodplain wetland changes: a case study in the China side of the Amur River Basin. *Int. J. Appl. Earth Obs. Geoinformation* 92, 102185.
- Kilpeläinen, P., Tokola, T., 1999. Gain to be achieved from stand delineation in LANDSAT TM image-based estimates of stand volume. *Forest Ecol. Manag.* 124, 105–112.

- Lausch, A., Erasmi, S., King, D.J., Magdon, P., Heurich, M., 2017. Understanding forest health with remote sensing—part II—a review of approaches and data models. *Remote Sens.* 9, 129.
- Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D., 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-Fir Western Hemlock Forests. *Remote Sens. Environ.* 70, 339–361.
- Lehmann, E.A., Caccetta, P., Lowell, K., Mitchell, A., Zhou, Z.S., Held, A., Milne, T., Tapley, I., 2015. SAR and optical remote sensing: assessment of complementarity and interoperability in the context of a large-scale operational forest monitoring system. *Remote Sens. Environ.* 156, 335–348.
- Long, J., Lin, H., Wang, G., Sun, H., Yan, E., 2019. Mapping growing stem volume of Chinese Fir plantation using a saturation-based multivariate method and quad-polarimetric SAR images. *Remote Sens.* 11, 1872.
- Li, Q., Tang, L., Ren, Y., 2011. Temporal dynamics and spatial variations of forest vegetation carbon stock in Liaoning Province, China. *J. Forestry Res.* 22, 519–525.
- Li, W., Niu, Zheng, Shang, R., Qin, Y., Wang, L., Chen, H., 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. *Int J Appl Earth Obs Geoinformation* 92, 102163.
- Lin, C., 2019. Improved derivation of forest stand canopy height structure using harmonized metrics of full-waveform data. *Remote Sens. Environ.* 235, 111436.
- Liu, Y., Gong, W., Xing, Y., Hu, X., Gong, J., 2019a. Estimation of the forest stand mean height and aboveground biomass in Northeast China using SAR Sentinel-1B, multispectral Sentinel-2A, and DEM imagery. *ISPRS J. Photogramm. Remote Sens.* 151, 277–289.
- Liu, J., Wang, X., Wang, T., 2019b. Classification of tree species and stock volume estimation in ground forest images using Deep Learning. *Comput. Electron. Agr.* 166, 105012.
- Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E., 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *Int. J. Digit. Earth* 9, 63–105.
- Magruder, L., Neunschwander, A., Neumann, T., Kurtz, N., Duncanson, L., Dubayah, R., 2019. NASA's ICESat-2 and GEDI missions for land and vegetation applications. In: Proceedings of the 21st EGU General Assembly, Vienna, Austria.
- Mäkelä, H., Pekkarinen, A., 2004. Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data. *Forest Ecol. Manag.* 196, 245–255.
- Mao, D., Wang, Z., Du, B., Li, L., Tian, Y., Jia, M., Zeng, Y., Song, K., Jiang, M., Wang, Y., 2020. National wetland mapping in China: a new product resulting from object-based and hierarchical classification of Landsat 8 OLI images. *ISPRS J. Photogramm. Remote Sens.* 164, 11–25.
- Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Zald, H.S.J., 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. *Remote Sens. Environ.* 209, 90–106.
- Mayya, E.W., Koskinen, J., Tegel, K., Hämäläinen, J., Kauranne, T., Käyhkö, N., 2019. Modelling and predicting the growing stock volume in small-scale plantation forests of Tanzania using multi-sensor image synergy. *Forests* 10, 279.
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning lidar. *Photogramm. Eng. Rem. S.* 66, 1367–1371.
- Millard, K., Richardson, M.C., 2015. On the importance of training data sample selection in RF classification: a case study in peatland ecosystem mapping. *Remote Sens.* 7, 1–26.
- MOF (Ministry of Forestry), 1982. Standards for Forestry Resource Survey. China Forestry Publisher, Beijing, China.
- Morin, D., Planells, M., Guyon, D., Villard, L., Mermoz, S., Bouvet, A., Thevenon, H., Dejoux, J.-F., Toan, T.L., Dedieu, G., 2019. Estimation and mapping of forest structure parameters from open access satellite images: Development of a generic method with a study case on coniferous plantation. *Remote Sens.* 11, 1275.
- Mudanjiang Statistical Bureau, 2018. Mudanjiang Statistical Yearbook 2018.
- Naeset, E., 1997. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sens. Environ.* 61, 246–253.
- Nakaya, T., Charlton, M., Lewis, P., Brunsdon, C., Yao, J., Fotheringham, S., 2014. GWR4 User Manual, Windows Application for Geographically Weighted Regression Modelling. Ritsumeikan University, Kyoto, Japan.
- Narine, L.L., Popescu, S., Neunschwander, A., Zhou, T., Srinivasan, S., Harbeck, K., 2019. Estimating aboveground biomass and forest canopy cover with simulated ICESat-2 data. *Remote Sens. Environ.* 224, 1–11.
- National Forestry and Grassland Administration, 2019. China Forestry and Grassland Statistical Yearbook 2019.
- Nilsson, M., Nordkvist, K., Jonzén, J., Lindgren, N., Axensten, P., Wallerman, J., Egberth, M., Larsson, S., Nilsson, L., Eriksson, J., Olsson, H., 2017. A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the National Forest Inventory. *Remote Sens. Environ.* 194, 447–454.
- Ni-Meister, W., Jupp, D.L.B., Dubayah, R., 2001. Modeling lidar waveforms in heterogeneous and discrete canopies. *IEEE Trans. Geosci. Rem. Sens.* 39, 1943–1958.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettenberg, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* 51, 933–938.
- Patenaude, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J., Dawson, T.P., 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sens. Environ.* 93, 368–380.
- Peña, M.A., Brenning, A., Sagredo, A., 2012. Constructing satellite-derived hyperspectral indices sensitive to canopy structure variables of a Cordilleran Cypress (*Austrocedrus chilensis*) forest. *ISPRS J. Photogramm.* 74, 1–10.
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J. B., Hoffon, M., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* 253, 112165.
- Puliti, S., Breidenbach, J., Astrup, R., 2020. Estimation of forest growing stock volume with UAV laser scanning data: can it be done without field data? *Remote Sens.* 12, 1245.
- Puliti, S., Saarela, S., Gobakken, T., Ståhl, G., Naesset, E., 2018. Combining UAV and Sentinel-2 auxiliary data for forest growing stock volume estimation through hierarchical model-based inference. *Remote Sens. Environ.* 204, 485–497.
- Rahlf, J., Breidenbach, J., Solberg, S., Naesset, E., Astrup, R., 2014. Comparison of four types of 3D data for timber volume estimation. *Remote Sens. Environ.* 155, 325–333.
- Ranson, K.J., Kimes, D., Sun, G., Nelson, R., Kharuk, V., Montesano, P., 2007. Using MODIS and GLAS data to develop timber volume estimates in central Siberia. In: Proceedings, International Geoscience and Remote Sensing Symposium (IGARSS), 2007. Barcelona, Spain, pp. 2306–2309.
- Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., Hyppäät, J., 2015. Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information. *Remote Sens. Environ.* 158, 431–440.
- Santi, E., Paloscia, S., Pettinato, S., Chirici, G., Mura, M., Maselli, F., 2015. Application of neural networks for the retrieval of forest woody volume from SAR multifrequency data at L and C bands. *Eur. J. Remote Sens.* 48, 673–687.
- Santoro, M., Beer, C., Cartus, O., Schmullius, C., Shvidenko, A., McCallum, I., Wegmüller, U., Wiesmann, A., 2011. Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR ScanSAR backscatter measurements. *Remote Sens. Environ.* 115, 490–507.
- Shataee, S., Kalbi, S., Fallah, A., Pelz, D., 2012. Forest attribute imputation using machine-learning methods and ASTER data: comparison of K-NN, SVR and random forest regression algorithms. *Int. J. Remote Sens.* 33, 6254–6280.
- Shen, C., Liang, W., Shi, Y., Lin, X., Zhang, H., Wu, X., Xie, G., Chain, P., Grogan, P., Chu, H., 2014. Contrasting elevational diversity patterns between eukaryotic soil microbes and plants. *Ecology* 95, 3190–3202.
- Silva, C.A., Hamamura, C., Valbuena, R., Hancock, S., Cardil, A., Broadbent, E.N., Almeida, D.R.A., Silva Junior, C.H.L., Klauber, C., 2020. rGEDI: NASA's Global Ecosystem Dynamics Investigation (GEDI) Data Visualization and Processing, version 0.1.2, accessed on April 1 2020, available at: <https://CRAN.R-project.org/package=rGEDI>.
- Solberg, S., Astrup, R., Breidenbach, J., Nilsen, B., Weydahl, D., 2013. Monitoring spruce volume and biomass with InSAR data from TanDEM-X. *Remote Sens. Environ.* 139, 60–67.
- Tadono, T., Takaku, J., Tsutsui, K., Oda, F., Nagai, H., 2015. Status of "ALOS World 3D (AW3D)" global DSM generation. In: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, Milan Italy.
- Thiel, C., Schmullius, C., 2016. The potential of ALOS PALSAR backscatter and InSAR coherence for forest growing stock volume estimation in Central Siberia. *Remote Sens. Environ.* 173, 258–273.
- Tompalski, P., Coops, N.C., White, J.C., Wulder, M.A., 2014. Simulating the impacts of error in species and height upon tree volume derived from airborne laser scanning data. *Forest Ecol. Manag.* 327, 167–177.
- Wang, C.K., 2006. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests. *Forest Ecol. Manag.* 222, 9–16.
- Wang, D., Wan, B., Liu, J., Su, Y., Guo, Q., Qiu, P., Wu, X., 2020. Estimating aboveground biomass of the mangrove forests on northeast Hainan Island in China using an upscaling method from field plots, UAV-LiDAR data and Sentinel-2 imagery. *Int. J. Appl. Earth Obs. Geoinformation* 85, 101986.
- Wang, Y., Wu, Z., Yuan, X., Zhang, H., Zhang, J., Xu, J., Lu, Z., Zhou, Y., Feng, J., 2011. Resources and ecological security of the Changbai Mountain region in Northeast Asia. In: Wang, Y. (Ed.), *Remote Sensing of Protected Lands*. CRC Press, Boca Raton, FL, USA, pp. 203–232.
- Wang, Y., Chen, L., 2020. A hybrid approach for mapping salt marsh vegetation. In: Y. Wang (Ed.), *Coastal and Marine Environments*, second ed., CRC Press, Boca Raton, FL, USA, pp. 299–306.
- Wijaya, A., Kusnadi, S., Gloaguen, R., Heilmeier, H., 2010. Improved strategy for estimating stem volume and forest biomass using moderate resolution remote sensing data and GIS. *J. For. Res.* 21, 1–12.
- Wittke, S., Yu, X.W., Karjalainen, M., Hyppäät, J., Puttonen, E., 2019. Comparison of two-dimensional multitemporal Sentinel-2 data with three dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. *Int. J. Appl. Earth Obs. Geoinform.* 76, 167–178.
- Wu, Z., Zhang, Z., Wang, J., 2019. Estimating the productive potential of five natural forest types in northeastern China. *For. Ecosyst.* 6, 42.
- Xie, B., Cao, C., Xu, M., Bashir, B., Singh, R.P., Huang, Z., Lin, X., 2020. Regional forest volume estimation by expanding LiDAR samples using multi-sensor satellite data. *Remote Sens.* 12, 360.
- Xu, C., Manley, B., Morgenroth, J., 2018a. Evaluation of modelling approaches in predicting forest volume and stand age for small-scale plantation forests in New Zealand with RapidEye and LiDAR. *Int. J. Appl. Earth Obs. Geoinformation* 73, 386–396.
- Xu, L., Shi, Y., Fang, H., Zhou, G., Xu, X., Zhou, Y., Tao, J., Ji, B., Xu, J., Li, C., Chen, L., 2018b. Vegetation carbon stocks driven by canopy density and forest age in subtropical forest ecosystems. *Sci. Total Environ.* 631–632, 619–626.

- Xu, Y., Chao, L., Sun, Z., Jiang, L., Fang, J., 2019. Tree height explains stand volume of closed-canopy stands: evidence from forest inventory data of China. *For. Ecol. Manage.* 438, 51–56.
- Yanbian Statistical Bureau, 2018. Yanbian Statistical Yearbook 2018.
- Zhao, Q., Yu, S., Zhao, F., Tian, L., Zhao, Z., 2019. Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments. *Forest Ecol. Manag.* 434, 224–234.
- Zhang, H., Zhu, J., Wang, C., Lin, H., Long, J., Zhao, L., Fu, H., Liu, Z., 2019. Forest growing stock volume estimation in subtropical mountain areas using PALSAR-2 L-Band PolSAR data. *Forests* 10, 276.
- Zhang, L., Dong, T., Xu, W., Ouyang, Z., 2015. Assessment of habitat fragmentation caused by traffic networks and identifying key affected areas to facilitate rare wildlife conservation in China. *Wildlife Res.* 42, 266–279.