Forest Ecology and Management 389 (2017) 199–210

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Estimating aboveground biomass of broadleaf, needleleaf, and mixed forests in Northeastern China through analysis of 25-m ALOS/PALSAR mosaic data

Jun Ma[a], Xiangming Xiao[a,b,*], Yuanwei Qin[b], Bangqian Chen[a,c], Yuanman Hu[d], Xiangping Li[a], Bin Zhao[a]

[a] Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, Fudan University, Shanghai 200433, China
[b] Department of Microbiology and Plant Biology, Center for Spatial Analysis, University of Oklahoma, Norman, OK 73019, USA
[c] Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Danzhou Investigation & Experiment Station of Tropical Crops, Ministry of Agriculture, Danzhou 571737, China
[d] Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China

A R T I C L E I N F O

Article history:
Received 2 September 2016
Received in revised form 9 December 2016
Accepted 17 December 2016

Keywords:
ALOS/PALSAR
Aboveground live biomass
Northeastern China
Nonlinear regression models
Boosted regression tree
Topographical and stand structure factors

A B S T R A C T

Aboveground biomass (AGB) of temperate forest plays an important role in global carbon cycles and needs to be estimated accurately. ALOS/PALSAR (Advanced Land Observing Satellite/Phased Array L-band Synthetic Aperture Radar) data has recently been used to estimate forest AGB. However, the relationships between AGB and PALSAR backscatter coefficients of different forest types in Northeastern China remain unknown. In this study, we analyzed PALSAR data in 2010 and observed AGB data from 104 forest plots in 2011 of needleleaf forest, mixed forest, and broadleaf forest in Heilongjiang province of Northeastern China. "*Poisson*" regression in generalized linear models (GLMs) and *BRT* (boosted regression tree) analysis in generalized boosted models (GBMs) were used to test whether the constructed PALSAR/AGB models based on individual forest types have better performance. We also investigated whether adding topographical and stand structure factors into the regression models can enhance the model performance. Results showed that GBM model had a better performance in fitting the relationships between AGB and PALSAR backscatter coefficients than did GLM model for needleleaf forest ($RMSE$ = 3.81 Mg ha^{-1}, R^2 = 0.98), mixed forest ($RMSE$ = 17.72 Mg ha^{-1}, R^2 = 0.96), and broadleaf forest ($RMSE$ = 19.94 Mg ha^{-1}, R^2 = 0.96), and performance of nonlinear regression models constructed on individual forest types were higher than that on all forest plots. Moreover, fitting results of GLM and GBM models were both enhanced when topographical and stand structure factors were incorporated into the predictor variables. Regression models constructed based on individual forest types outperform than that based on all forest plots, and the model performance will be enhanced when incorporating topographical and stand structure factors. With information of forest types, topography, and stand features, PALSAR data can express its full ability in accurate estimation of forest AGB.

1. Introduction

Temperate forests cover more than 6.4 billion hectares on the Earth, and approximately 41 Pg carbon is stored in its vegetation carbon pools, most of which is held in aboveground live biomass (AGB) (Dixon et al., 1994). In Northeastern (NE) China, the area of temperate forest is more than 38.3 million hectares and accounts for more than one third of the total forest area in China, and the carbon storage of forests in this area is about 1.4 Pg C and also accounts for about 30% of the total carbon storage in forests of China (Wang, 2006). Many factors have both positive and negative influence on forest aboveground biomass. On the one hand, human and natural disturbances, such as harvesting, fire, and pest disease, in history decreased the carbon density in NE China, which is lower than that in temperate forests of other regions over the world (Fang et al., 2001). Forests in NE China tended to be carbon source due to overharvesting and degradation during 1980s and 1990s (Piao et al., 2009). On the other hand, NE China locates in high latitude region where the climate has changed intensely since last century, and forest biomass in this region is boosted by the climate warming (Yang and Wang, 2005). More-

* Corresponding author at: Department of Botany and Microbiology, College of Arts and Sciences, Center for Spatial Analysis, College of Atmospheric & Geographic Sciences, University of Oklahoma, USA.

E-mail address: xiangming.xiao@ou.edu (X. Xiao).

http://dx.doi.org/10.1016/j.foreco.2016.12.020
0378-1127/© 2016 Published by Elsevier B.V.

over, although forests in NE China have experienced severe harvesting in history (Jiang et al., 2002; Yu et al., 2011), they had been one of the key objectives for conservation and reforestation in Natural Forest Resource Conservation Project of China since 2000 (Wei et al., 2014), and forest biomass in this region increased rapidly (Ma et al., 2016). Forest biomass in NE China has changed greatly during the past several decades. Therefore, accurate estimation of forest aboveground biomass has important significance in estimating the role of temperate forests in regional and global carbon cycle (Laurin et al., 2016) and developing science-based forest management practices.

There are a number of ways to estimate and monitor forest AGB (Brandeis et al., 2006; Soenen et al., 2010; FAO, 2015). Directly weighing individual components of trees is the most accurate way to estimate the biomass of trees (Parresol, 1999), but the method is hardly adopted because of its high cost of labor, money, and time. Conducting forest inventory and calculating forest biomass using allometric biomass equations based on DBH (diameter at breast height) and height of each tree is an efficient way (Gower et al., 1999; Wang, 2006). Although rich data of forest composition and structure can be obtained in forest inventory, it still has some deficiency in evaluating spatial distribution of forest biomass (Brown et al., 1999; Houghton et al., 2001). Moreover, it is also difficult to calculate the biomass of some tree species, as their allometric equations haven't been established yet. Remote sensing has offered a viable mean for estimating forest AGB at large spatial scales (Hansen et al., 2000; Myneni et al., 2001; Brown, 2002).

Estimation of forest AGB from remote sensing data starts with analysis of the relationship between remote sensing signals and AGB of training samples, and then applies this relationship (statistical model) to calculate AGB over the entire study area (Bastin et al., 2014). Data from optical sensors were used to estimate forest biomass, based on the relationship between forest AGB and vegetation indices, such as NDVI (normalized difference vegetation index) and EVI (enhanced vegetation index) (Huete et al., 2002; Nakaji et al., 2008). However, the applications with optical data are often limited by the lack of high quality images due to frequent clouds and saturation at low biomass level by the spectral bands and spectral indices (Nichol and Sarker, 2011). Data from LiDAR (Light detection and ranging) provide accurate three-dimension information like tree height and canopy vertical structure (Naesset, 2002; Goetz et al., 2009), and AGB is calculated using empirical equation of tree height and biomass (Lefsky et al., 1999; Zhao et al., 2009). Because of sophisticated technical equipment and high cost, airborne LiDAR images are not widely available and are less often used in biomass estimation at large spatial scales, including temperate forest of NE China (Tang et al., 2012; Zhang and Ni-meister, 2014).

Synthetic Aperture Radar (SAR) data such as L-band ALOS/PALSAR (Advanced Land Observing Satellite/Phased Array L-band Synthetic Aperture Radar) and X-band TerraSAR-X are widely available and have been increasingly used in estimation of forest AGB (Karjalainen et al., 2012; Vastaranta et al., 2014). PALSAR data were used to estimate AGB of forest plots from tropic and temperate forests to boreal forests in Africa, North America, Australia, and Russia (Lucas et al., 2007; Thiel et al., 2009; Lucas et al., 2010; Cartus et al., 2012; Sarker et al., 2012). Nonlinear regression models were developed to estimate forest AGB based on PALSAR backscatter coefficients; but the model structure and parameters vary substantially among these studies (Lucas et al., 2010; Englhart et al., 2011; Carreiras et al., 2012; Peregon and Yamagata, 2013). In addition, other forest stand properties (stand structure and complexity of understory layer) and topographical features vary among different forest types and affect forest AGB (Conard and Ivanova, 1997; Jobidon, 2000; Ma et al., 2015b). These factors also have influence on PALSAR backscatter coefficients (Lucas et al., 2010; Whittle et al., 2012; Atwood et al., 2014). Therefore, it may be useful to incorporate forest stand and topographical factors in the nonlinear regression models and to construct various regression models of different forest types for the purpose of accurate estimation of AGB.

In this study, we constructed the nonlinear relationship between PALSAR backscatter coefficients and forest AGB of different forest types in NE China, based on forest inventory data of 104 plots and PALSAR data. Forest types in NE China were divided into broadleaf forest, needleleaf forest, and mixed forest in our study. The objectives of this study were twofold: (1) determine the relationships between AGB and PALSAR backscatter coefficients by different forest types; (2) test the hypothesis that adding forest stand and topographical factors in the predictor variables of regression models can improve estimation of forest AGB.

2. Materials and methods

2.1. Study area

Our study area is the forest zone in Heilongjiang province of NE China, and it extends across 43.42°N–52.58°N, 118.06°E–135.16°E (Fig. 1). The topography is characterized by low mountains with elevation of 120–1000 m. The climate types are mid-temperate continental monsoon climate and cold- temperate continental monsoon climate. The annual mean temperature ranges from −2.8 °C in southern part to −3.2 °C in northern part. The average annual rainfall ranges from 530 mm to 800 mm, falling most in summer. Three main forest types are located in our study area, needleleaf forest in the northern part, mixed forest in the central part, and broadleaf forest in the southern part (Fig. 1). Based on our inventory data and previous studies (Ma et al., 2016; Ma et al., 2015b), species compositions of the three forest types are listed in Table 1.

2.2. Field inventory data and AGB calculation

In 2011, field inventory was carried out in various types of forests in Heilongjiang province. A total of 104 forest plots (Fig. 1) with the size of 20 m × 50 m were surveyed. These plots belong to three forest types: needleleaf forest, mixed forest, and broadleaf forest (Table S2). For each plot, location (latitude and longitude) of the central point, species name, diameter at breast height (DBH), and height of individual trees in the overstory layer were recorded. Because the lower limit of the applicable range of most biomass allometric equations in this study is about 5 cm, we only measured the trees that with a minimum DBH of 5 cm. trees with DBH less than 5 cm will be regarded as shrubs, and their biomass was calculated by direct measurement. Each plot was regarded as an individual sample in our analysis. The number of dead trees was quite few, therefore they were not included in the AGB of our survey. Within each tree plot, three 2 m × 2 m shrub plots and three 1 m × 1 m herb plots were selected randomly. Species name and abundance of each shrub and herb were recorded, and then the aboveground part of shrub and herb was harvested. These shrub and herb samples were taken into laboratory for further processing, and they were dried to a constant weight at 105 °C and then weighed. Considering the low growth rate of forests in this high latitude region, the increment of forest AGB for one year is negligible. Therefore, forest inventory results in 2011 were matching with PALSAR data in 2010.

The DBH-based allometric equations from previous studies (Chen and Zhu, 1989; Wang, 2006) were adopted to calculate tree AGB (Table S1). The dry weight of shrub and herb samples of the three subplots within a tree plot represented the AGB of under-

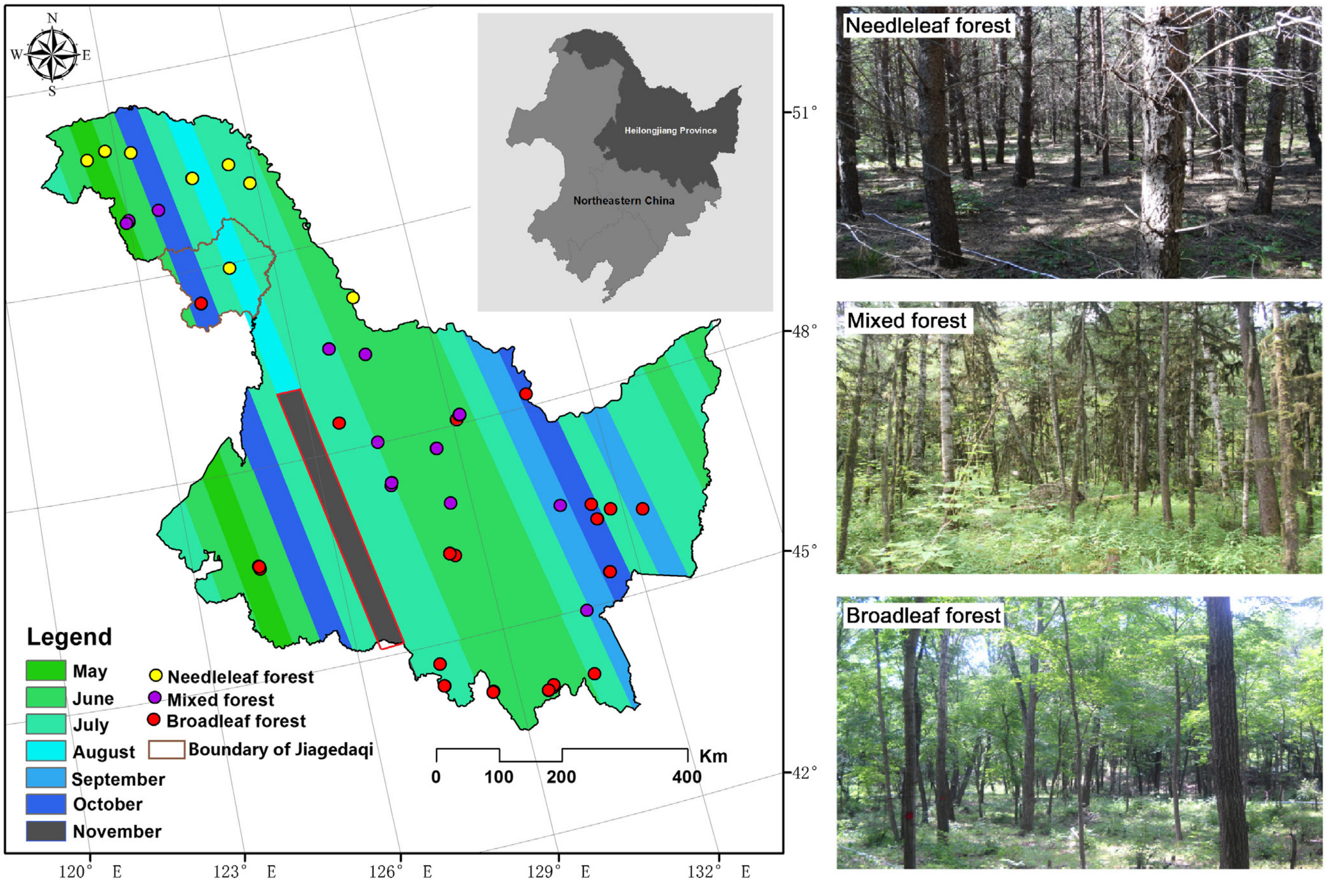

Fig. 1. Locations of forest sampling plots of different forest types and the acquisition dates of PALSAR mosaic dataset and field photos of needleleaf forest, mixed forest, and broadleaf forest in Heilongjiang Province of Northeastern China in 2010. The area with the brown boundary do not belong to Heilongjiang Province, however, we also conducted field investigation there. Note: the strip in the red box was obtained in November 2010, which is beyond the plant growing season and could be affected by snow, and thus it was replaced by PALSAR data in September 2009.

Table 1
Main species, plots number, and aboveground live biomass (AGB) of the three forest types that surveyed in Heilongjiang of Northeastern China. Min, Max, Mean, Median, and Std represent the minimum, maximum, mean, median and standard deviation values of AGB of different forest types.

Forest types	Main species	Plots number	AGB (Mg ha^{-1})				
			Min	Max	Mean	Median	Std
Needleleaf forest	*Larix gmelinii*, *Pinus sylvestris Linn.*, *Betula platyphylla*, *etc.*	20	11.2	169.6	84.0	79.3	47.9
Mixed forest	*Pinus. koraiensis*, *Picea koraiensis and Picea jezoensis*, *Abies nephrolepis*, *Fraxinus mandshurica*, *Ulmus japonica*, *Acer mono Maxim*, *etc.*	50	14.3	350.3	127.2	116.7	79.9
Broadleaf forest	*Populus davidiana*, *Betula costata*, *Quercus mongolica*, *Betula costata*, *etc.*	34	31.9	388.0	129.2	108.3	72.5
All forest plots		104	11.2	388.0	119.5	106.2	74.4

story layer. The AGB (Mg ha^{-1}) of a tree plot was calculated as the sum of all trees' AGB and understory layers' AGB in the tree plot. However, understory AGB only accounted for a small percentage of the total AGB, and the proportions of the understory AGB to the total AGB for needleleaf forest, mixed forest, and broadleaf forest were 0.71%, 0.99%, and 1.56%, respectively (Table S2).

At plot level, stand structure features were mainly reflected based on the inventory data. Tree density and median value of tree height in a plot were calculated. The two stand structure factors as well as their ranges and explanations or calculation formulas were listed in Table 2. Considering the high correlation between these two stand structure factors to forest AGB (Brown et al., 1989; Fang et al., 1996; Lefsky et al., 2002; Naesset, 2002) and the availability of the global forest canopy height data (Lefsky, 2010), they were also incorporated as the predictor variables of forest AGB estimation.

2.3. ALOS/PALSAR data

The 25-m PALSAR L-band orthorectified mosaic data with the Fine Beam Dual (FBD) model in 2009 and 2010 was downloaded from the ALOS Research and Application Project of EORC, Japan Aerospace Exploration Agency (http://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/data/). This dataset is aggregated from original observation with minimum response to surface moisture (Shimada et al., 2014). In high latitude regions, the growing season of forests is relative short (May to October) and there is large variation in snow cover, which may contribute to seasonal variability

Table 2
Ranges and calculation of topographical and forest stand structure variables used to be combined in generalized linear models (GLMs) and generalized boosted models (GBMs) analysis. In calculation of topographical wetness index (TWI), α represents flow area per cell, and β represents the slope (by radian).

Variables	Code	Range	Explanation or calculation formula
Elevation (m)	M1	121–1016	Elevation
Slope (°)	M2	0.8–22.5	Slope
Aspect (°)	M3	7.8–322.3	Aspect
Irasp	M4	0–1	$Irasp = \frac{1-\cos\left[\left(\frac{\pi}{180}\right)\cdot(\text{aspect}-30)\right]}{2}$
Relief	M5	3–59	Relief amplitude
STD	M6	0.9–17.8	Standard deviation of elevation of 3*3 pixels
TWI	M7	4.7–13.1	$TWI = \ln \alpha/\beta$
Tree Height (m)	N1	3.3–21.6	Median value of tree height of a plot
TD (stem ha^{-1})	N2	201–3670	Tree density

in SAR data (Santoro et al., 2011). We selected PALSAR data mainly in summer and early autumn when biomass data is considered to be most representative and PALSAR data receives least influence of moisture and snow. However, one PALSAR image strip was obtained in November 2010, and it was replaced by another strip of PALSAR data obtained in September 2009 (Fig. 1).

Gamma-naught HH and HV are included in the dataset, and the preprocessing of PALSAR data is completed by JAXA. Geometrical calibration of the PALSAR image was conducted using 90 m resolution SRTM (Shuttle Radar Topography Mission) Digital Elevation Model (DEM) (Shimada and Ohtaki, 2010). The digital numbers (DN) of PALSAR signal amplitude have been extracted and converted to gamma naught backscattering coefficients (dB) in decimal units using following equation (Shimada et al., 2009; Qin et al., 2015):

$$\gamma^0 = 10 \times \log_{10} DN^2 - CF$$

where γ^0 is the backscattering coefficient, DN is the PALSAR signal amplitude in HH or HV, CF is the calibration factor, depended on incidence angle (Shimada et al., 2009), and equals to −83.

In addition to HH and HV backscatter coefficients, we also calculated the sum, the difference, the ratio, the normalized difference, and square values of above coefficients (Table 3), and they were all used to explore the relationships between PALSAR data and forest AGB.

In order to test whether the PALSAR data of the 104 forest plots can represent the forests in the entire study area, we calculated and compared the frequency and cumulative frequency curves of PALSAR HH and HV backscatter coefficients for (1) all forest pixels in Heilongjiang province, and (2) those pixels associated with the 104 forest sampling plots, respectively (Fig. 2), based on the forest and non-forest map from analysis of PALSAR and MODIS data (Qin et al., 2015). The frequency distribution of HH and HV backscatter coefficients for the 104 plots is similar to those associated with all pixels in Heilongjiang province indicating that our sampling sites are representative.

2.4. Topographical data

We downloaded the Digital Elevation Model (DEM) data at 30-m spatial resolution from the USGS (United States Geological Survey) website (http://www.usgs.gov/). We calculated elevation, slope, aspect, irradiation aspect (Irasp), relief amplitude (Relif), standard deviation of elevation of 3×3 pixels (STD), and topographic wetness index (TWI) of each forest plot. Irasp represents the amount of irradiation of a certain aspect, Relif and STD both represent the relief intensity of microtopography, and TWI reflects the moisture of circumstance induced by topographical conditions (Sorensen et al., 2006). Selection of these topographical factors was based on the consideration of their possible impacts on forest AGB or PALSAR data, and all these topographical factors can be calculated from DEM dataset, which can also be applied in forests of other regions. The calculation and ranges of these topographical factors were also listed in Table 2.

2.5. Regression models

A few studies showed strong linear relationships between logarithmic transform of AGB and predictor variables (Carreiras et al., 2012; Peregon and Yamagata, 2013). In this study, we first developed the linear-log regression models between forest AGB (natural logarithmic transformation) and PALSAR (HH, HV) backscatter coefficients. However, several other studies (Lucas et al., 2010; Cartus et al., 2012) reported that nonlinear regression models were considered as the best fit of the relationship between PALSAR data and forest AGB. Therefore, we also used both "*Poisson*" regression in generalized linear models (GLMs) and boosted regression tree (*BRT*) analysis in generalized boosted models (GBMs) to fit the relationship between forest AGB and PALSAR backscatter coefficients in this study. Considering topographical and forest stand structure factors have significant influence on both forest AGB and PALSAR backscatter coefficients, they were also incorporated in regression models to explore whether they can improve the performance of models.

"*Poisson*" regression normally has an advantage in fitting logarithmic model of variables, therefore, it can be used to build direct nonlinear relationship between AGB and predictor variables. *BRT* analysis is a machine learning approach used in nonlinear relationship analysis (Moisen et al., 2006; Elith et al., 2008), which couples

Table 3
Parameters estimates and fitting statistics of variables of full model using "*Poisson*" regression in generalized linear models (GLMs) from samples of different forest types. An asterisk means the significant ($P < 0.05$) effect of regression of the variable. γ^0 is the ALOS/PALSAR backscatter intensity (dB). The significance values of these coefficients were obtained for all coefficients run at the same time.

Parameters	Variables	Code	Needleleaf forest	Mixed forest	Broadleaf forest	All forest plots
α0	*Intercept*		−488.5*	3.7	−109.9*	−138.8*
α1	γ^0_{HH}	X1	−2579.6*	−924.3*	−266.1*	−513.5*
α2	γ^0_{HV}	X2	598. 9*	3.5	421.5*	195.3*
α3	γ^0_{HH+HV}	X3	993.8*	456.7*	−79.8	157.4*
α4	γ^0_{HV-HH}	X4	−1609.6*	−454.0*	−342.0*	−351.5*
α5	$\gamma^0_{HH/HV}$	X5	534.9*	−114.2*	138.9*	173.2*
α6	$\gamma^0_{(HV-HH)/(HH+HV)}$	X6	733.0*	151.5*	134.2*	214.5*
α7	${\gamma^0_{HH}}^2$	X7	−636.7*	−52.34	525.1*	58.5*
α8	${\gamma^0_{HV}}^2$	X8	−637.5*	−52.0	525.2*	58.5*
α9	${\gamma^0_{(HH+HV)}}^2$	X9	318.6*	26.0	−262.6*	−29.3*
α10	${\gamma^0_{(HV-HH)}}^2$	X10	318.0*	26.2	−262.6*	−29.2*
α11	${\gamma^0_{(HH/HV)}}^2$	X11	−3.1	81.6*	−47.9*	−50.3*
α12	${\gamma^0_{[(HV-HH)/(HH+HV)]}}^2$	X12	−414.8*	−138.8*	−7.3	−65.8*

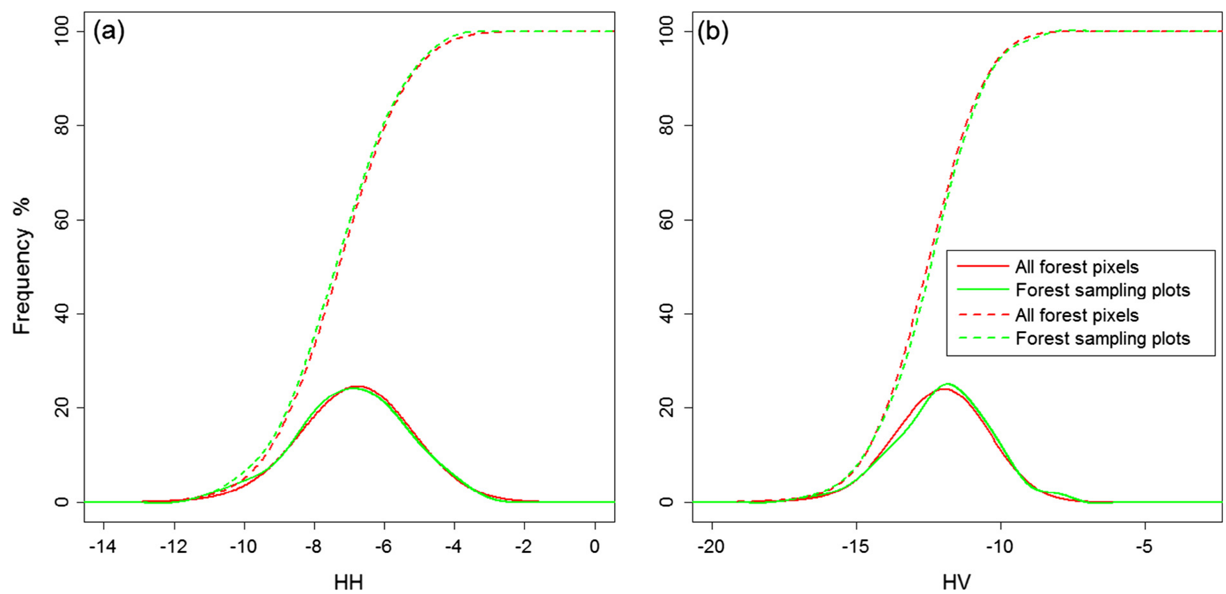

Fig. 2. Frequency and cumulative frequency curves of PALSAR HH and HV backscatter coefficients for (1) all forest pixels in Heilongjiang province, and (2) forest sampling plots in this study. (a) HH, (b) HV.

the strengths of two algorithms: regression trees and boosting. Regression trees are originated from the theories of classification and decision tree. Boosting is mainly based on a forward procedure which construct and combine a collection of models with the purpose of improving model performance. No transformation is needed in *BRT* analysis due to the ability in accommodating any data distribution.

In regression models, the response variable is forest AGB, and the predictor variables include PALSAR backscatter coefficients, forest stand structure and topographical factors. We used *R* software to conduct GLM and GBM ("*gbm*" package) fitting (R Development Core Team, 2011). In addition, the ranking of relative importance of individual predictor variables was also output by *BRT* analysis. Parameters including "*gaussian*" error distribution, a learning rate of 0.005, a bag fraction of 0.5, and 10-fold cross validation were set in *BRT* analysis. Root mean square error (*RMSE*) and R square (R^2) were used to evaluate the performance of fitting models. In the results, fitted AGB in "*Poisson*" regression of GLM and cross validation predicted AGB in BRT analysis of GBM were also generated.

2.6. Variables selection

In order to avoid over fitting, all PALSAR backscatter coefficients were firstly used in GLM regression models, and parameter estimates (full model) and the significance of each PALSAR backscatter coefficients were output (Table 3). Results showed that not all PALSAR backscatter coefficients were significantly correlated with forest AGB, and collinearity may exist among these coefficients. Therefore, based on the significance of each parameter of the full model, variables selection was conducted using all-subsets regressions ("*leaps*" packages in R) method to get the best fit model and to avoid collinearity. Variable selections were developed in both circumstances that the predictor variables include or not include topographical and stand structure factors. The highest adjust R^2 was used as the filter criteria to choose predictor variables that would be selected to construct nonlinear regression models. The chosen predictor variables of different forest types and all forest plots by all-subsets regressions were shown in Fig. 3. However, based on several previous studies (Saatchi et al., 2007; Carreiras et al., 2012; Peregon and Yamagata, 2013), HH and HV backscatter coefficients will be included if they are not chosen by the method of all-subsets regression. All the nonlinear regression models in GLM and GBM were constructed after variable selection.

3. Results

3.1. Single-variable linear-log regression models between AGB and PALSAR HH and HV data

Fig. 4a showed the relationships between PALSAR HH backscatter coefficients and AGB by individual forest types and all forest plots. As AGB increases, PALSAR HH also increased and reached saturation points at ~150 Mg ha^{-1} for needleleaf forest and ~100 Mg ha^{-1} for mixed forest. According to the linear-log regression models, the relationship between forest AGB and PALSAR HH backscatter coefficients were significant ($P < 0.05$) for needleleaf forest ($R^2 = 0.63$), mixed forest ($R^2 = 0.20$), and all forest plots ($R^2 = 0.13$) (Fig. 4a). Broadleaf forest had no significant logarithmic correlation between PALSAR HH and AGB.

Fig. 4b showed the relationships between PALSAR HV and AGB by individual forest types and all forest plots. The larger dynamic range of HV backscatter coefficients, in comparison to HH, clearly represented the sensitivity of HV to the variation in AGB. The scatterplots showed that saturation points of PALSAR HV vary from ~160 Mg ha^{-1} for needleleaf forest, ~130 Mg ha^{-1} for mixed forest, to ~100 Mg ha^{-1} for broadleaf forest. According to the linear-log regression models, the relationships between forest AGB and PALSAR HV backscatter coefficients were significant ($P < 0.05$) for needleleaf forest ($R^2 = 0.63$), mixed forest ($R^2 = 0.47$), broadleaf forest ($R^2 = 0.28$), and all forest plots ($R^2 = 0.41$), respectively.

3.2. Multi-variable nonlinear regression models between AGB and PALSAR data

Improvement of performance was found in the GLM regression models for needleleaf forest (*RMSE* = 21.47 Mg ha^{-1}, $R^2 = 0.82$), mixed forest (*RMSE* = 56.46 Mg ha^{-1}, $R^2 = 0.45$), and broadleaf forest (*RMSE* = 52.02 Mg ha^{-1}, $R^2 = 0.44$) (Table 4), in comparison to the linear-log relationships between AGB and single HH and HV

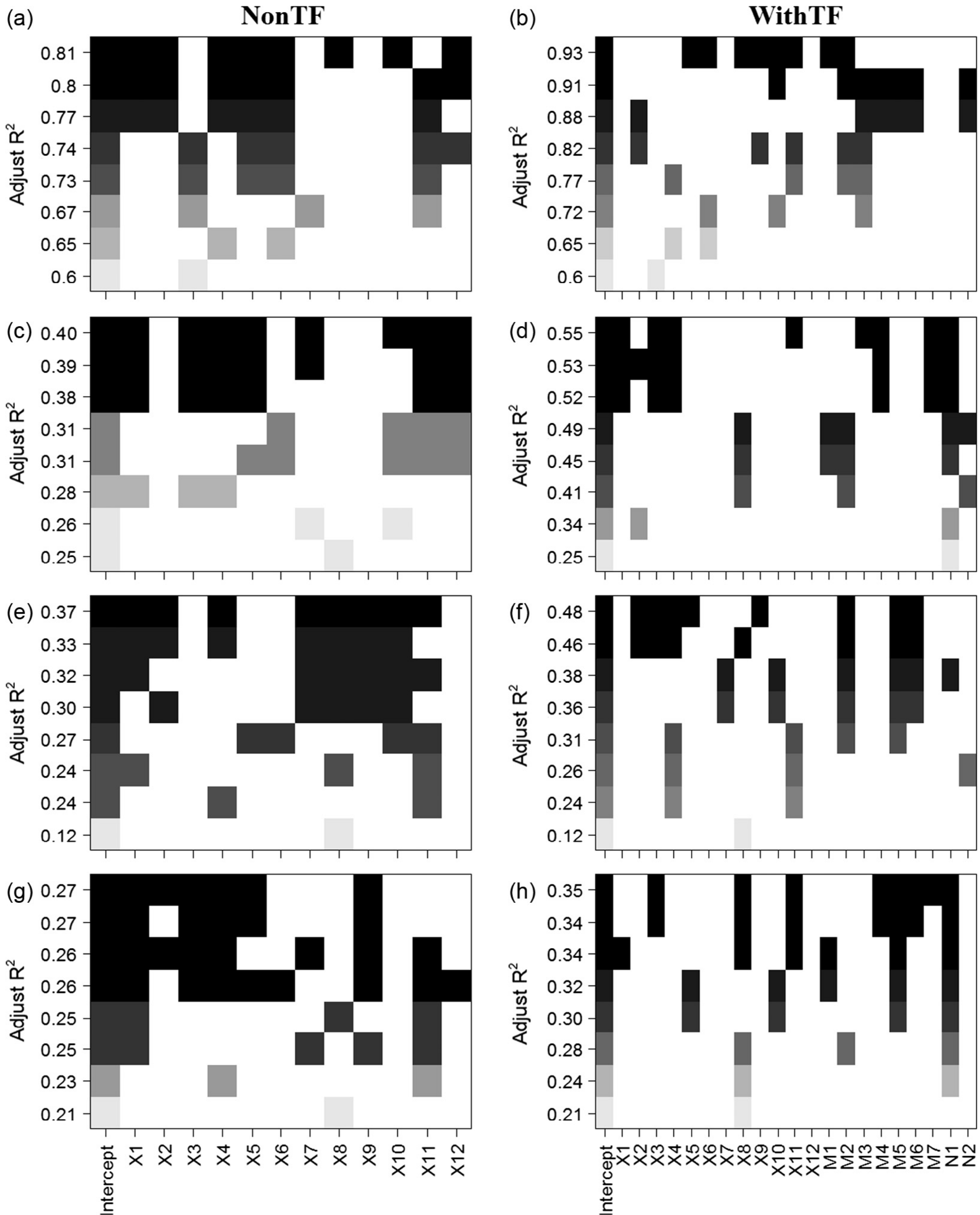

Fig. 3. Selected predictor variables of different forest types and all forest sampling plots in nonlinear regression (GLM and GBM) models using all-subsets regression method. (a) and (b): Needleleaf forest, (c) and (d): Mixed forest, (e) and (f): Broadleaf forest, and (g) and (h): All forest plots. NonTF represents PALSAR data only, and WithTF represents PALSAR data and topographical and forest stand structure factors.

backscatter coefficients (see Fig. 4). Moreover, higher correlations were found in the GBM regression models for needleleaf forest (*RMSE* = 3.81 Mg ha^{-1}, R^2 = 0.98), mixed forest (*RMSE* = 17.72 Mg ha^{-1}, R^2 = 0.96), broadleaf forest (*RMSE* = 19.94 Mg ha^{-1}, R^2 = 0.96), and all forest plots (*RMSE* = 25.99 Mg ha^{-1}, R^2 = 0.90).

The estimated AGB from the GLM and GBM models was significantly correlated with observed AGB, respectively. In "*Poisson*" regression of GLM (Fig. 5), the correlations between fitted AGB and observed AGB of needleleaf forest (R^2 = 0.80), mixed forest (R^2 = 0.50), and broadleaf forest (R^2 = 0.49) outperformed that of all forest plots (R^2 = 0.36). Similar pattern was also found in *BRT* analysis of GBM (Fig. 6). Correlations between cross validation predicted AGB and observed AGB in needleleaf forest (R^2 = 0.98), mixed forest (R^2 = 0.93), and broadleaf forest (R^2 = 0.91) were all stronger than that in all forest plots (R^2 = 0.36).

The most important factors in nonlinear regression models based on multi-variables of PALSAR coefficients that influence the estimation of AGB for needleleaf forest, mixed forest, broadleaf forest were γ^0_{HH} (X1), γ^0_{HH+HV} (X3), $\gamma^0_{(HH+HV)^2}$ (X9), and γ^0_{HV} (X2), respectively. Their relative importance were 34.8%, 31.0%, 23.0%, and 29.6%, respectively (Table 5).

3.3. Effect of topographical and stand structure factors on regression models

The GLM models had better performance for needleleaf forest (*RMSE* = 10.75 Mg ha^{-1}, R^2 = 0.94), mixed forest (*RMSE* = 46.70 Mg ha^{-1}, R^2 = 0.58), broadleaf forest (*RMSE* = 45.52 Mg ha^{-1}, R^2 = 0.54), and all forest plots (*RMSE* = 56.10 Mg ha^{-1}, R^2 = 0.40)

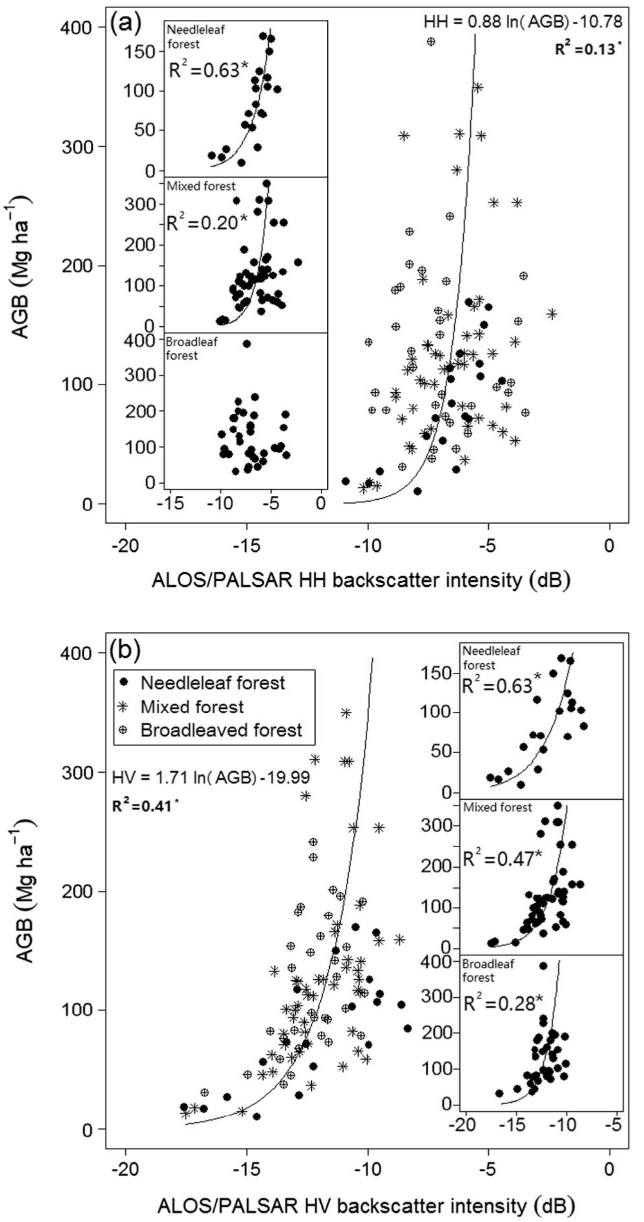

Fig. 4. The relationship between forest aboveground biomass (AGB) and PALSAR backscatter coefficients by individual forest types and all forest sampling plots. (a) HH, (b) HV. Logarithmic regressions are used to fit the relationship, and asterisks after the R square values indicate significant correlations ($P < 0.05$).

when topographical and stand structure factors were incorporated (Table 4). The GBM models also show improved performance for various forest types but not for all forest plots (Table 4).

Correlations between fitted AGB and observed AGB in GLM models for needleleaf forest ($R^2 = 0.95$), mixed forest ($R^2 = 0.66$), broadleaf forest ($R^2 = 0.60$), and all forest plots ($R^2 = 0.43$) were all increased when topographical and stand structure factors were incorporated into predictor variables of the GLM models (Fig. 5). In the GBM models, except for all forest plots ($R^2 = 0.34$), higher correlations between cross validation predicted AGB and observed AGB for needleleaf forest ($R^2 = 0.99$), mixed forest ($R^2 = 0.99$), and broadleaf forest ($R^2 = 0.92$), were also detected when topographical and stand structure factors were included (Fig. 6).

Some topographical factors (M1, M2, M5, and M7) and tree height (N1) emerged to be the top five most important factors that influence forest AGB of various forest types and all forest plots when topographical and stand structure factor were incorporated into nonlinear regression models (Table 5). Especially, tree height (N1), with relative importance of 29.5%, emerged to be the most important variables influence forest AGB for mixed forest.

4. Discussion

4.1. The relationship between forest AGB and individual PALSAR HH and HV data

Significant logarithmic correlation was found between forest AGB and PALSAR HH backscatter coefficients in all forest plots, but fitted equation only explained a relatively small ($R^2 = 0.13$) proportion of variance (Fig. 4). At the meantime, a better ($R^2 = 0.41$) performance in logarithmic correlation between HV backscatter and forest AGB was identified. Logarithmic correlations between forest AGB and HV backscatter coefficients of needleleaf forest, mixed forest and broadleaf forest were higher than those between AGB and HH backscatter coefficients. These findings prove that HV backscatter coefficients generally have higher sensitivity in quantifying forest AGB, which has been also reported by some previous studies (Mitchard et al., 2009; Lucas et al., 2010).

The degree of logarithmic correlations between forest AGB and PALSAR HH and HV backscatter coefficients declines from needleleaf forest to mixed forest and broadleaf forest. This may be attributed to the following reasons. First, the vertical and spatial complexity varies among different forest types (Kane et al., 2013). The structure of leaves and branches in needleleaf forest is tighter than that in mixed and broadleaf forests, and it allows more scattering information to be obtained by PALSAR sensor (Carreiras et al., 2012). Therefore, AGB of needleleaf forest can be estimated more accurately. Second, the complexity of species composition (Zenner and Hibbs, 2000; McElhinny et al., 2005) and AGB of understory layers increases from needleleaf forest to mixed forest and broadleaf forest (Table S2) which may result in relatively larger error in estimating AGB from forest inventory data in broadleaf forest and mixed forest. This may cause lower correlation between forest AGB and PALSAR backscatter coefficients in mixed forest and broadleaf forest than in needleleaf forest.

Table 4
Root mean square error (RMSE) and R^2 (without/with topographical and forest stand structure factors) of fitted model in "*Poisson*" regression of generalized linear models (GLMs) and 10-fold cross-validation in boosted regression tree (BRT) analysis of generalized boosted models (GBMs) from samples of different forest types. NonTF and WithTF represent analysis without and with topographical and forest stand structure factors, respectively. Topographical and forest stand structure factors in this study are all listed in Table 2.

Statistics	Needleleaf forest		Mixed forest		Broadleaf forest		All forest plots	
	NonTF	WithTF	NonTF	WithTF	NonTF	WithTF	NonTF	WithTF
GLM								
RMSE (Mg ha^{-1})	21.47	10.75	56.46	46.70	52.02	45.52	59.14	56.10
R^2	0.82	0.94	0.45	0.58	0.44	0.54	0.32	0.40
GBM								
RMSE (Mg ha^{-1})	3.81	2.18	17.72	11.30	19.94	19.30	25.99	29.66
R^2	0.98	0.99	0.96	0.98	0.96	0.97	0.90	0.88

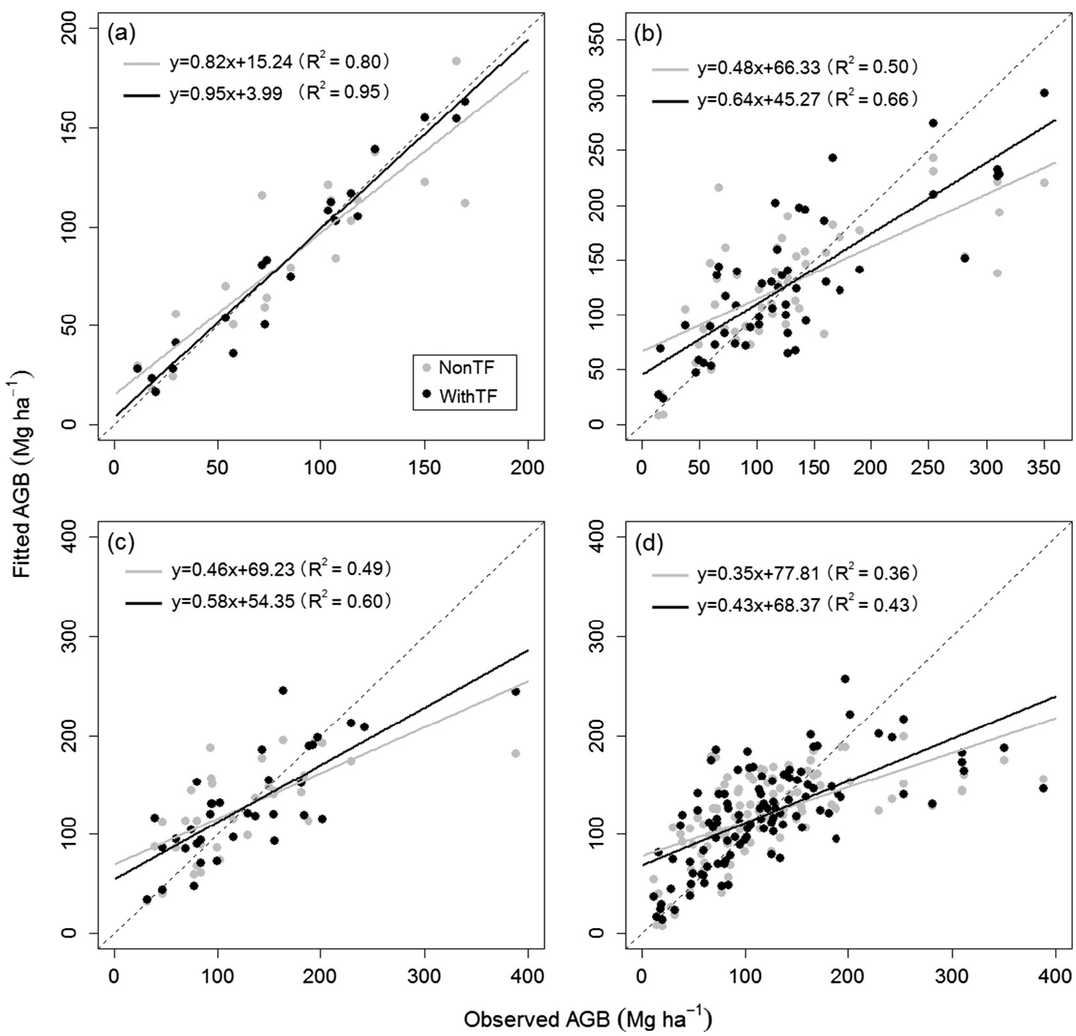

Fig. 5. The relationship between fitted AGB generated using "*Poisson*" regression in generalized linear models (GLMs) and observed AGB of various forest types based on predictor variables from (1) PALSAR data only (**NonTF**), and (2) PALSAR data and topographical and forest stand structure factors (**WithTF**). (a) Needleleaf forest, (b) Mixed forest, (c) Broadleaf forest, and (d) All forest sampling plots.

Both HH and HV backscatter coefficients could reach saturation as forest AGB increases. This phenomenon was reported by many publications (Luckman et al., 1997; Austin et al., 2003; Saatchi et al., 2007; Englhart et al., 2011; Carreiras et al., 2012), and some of them have pointed out that the saturation level was approximate at 100 Mg ha^{-1} in tropic forests. Our results showed that saturation levels varied among the three forest types in the temperate areas, for example, the saturation level of needleleaf forest, approximately up to 150 Mg ha^{-1}, was higher than the other two types. Generally, the complexity of vertical structure and understory composition in needleleaf forest is simpler when compared to mixed and broadleaf forests (Fig. 1). It is likely that the PALSAR sensor is more sensitive to AGB of needleleaf forest, and the saturation level therefore increases.

Highly consistent agreements of PALSAR HH/HV frequency and cumulative frequency curves between our field inventory plots (N = 104) and all forest pixels of Heilongjiang province (Fig. 2) indicate that the selected forest plots can well represent the forests in this area. Therefore, the relationships between PALSAR backscatter coefficients and forest AGB, especially for individual forest types, are reliable in estimating forest AGB at large spatial scale. Moreover, when AGB of some plots that reaches a high level (over 150 Mg ha^{-1}), their HH and HV backscatter coefficients maintain at about −6 dB and −11 dB (Fig. 4), respectively, which are higher than about 80% of the plots in current forests (Fig. 2). This shows that most of the forests in NE China are in the low level of AGB and suggests that a great potential of increasing AGB exists in the forests of NE China.

4.2. Nonlinear relationships between AGB and multi-variable PALSAR backscatter data of various forest types

Improvement of fitting results between AGB and multi-variable PALSAR backscatter data were detected in GLM and GBM models for different forest types and all forest plots (Table 4) when compared to the linear-log regression between AGB and single backscatter coefficients (Fig. 4). This is in line with many previous studies (Lucas et al., 2007; Lucas et al., 2010; Englhart et al., 2011; Cartus et al., 2012) which demonstrates that simple logarithmic regression models are not suitable for estimating AGB of forest with complex compositions. Regression models constructed based on multi-variable of PALSAR data have higher ability in detecting canopy structure and retrieving forest AGB (Saatchi et al., 2007; Carreiras et al., 2012). These studies constructed models based on single PALSAR HH/HV and their square value as well as the mean, minimum, maximum, and standard deviation of PALSAR HH and HV backscatter coefficients in North America and West Africa, respectively. However, predictor variables in this study

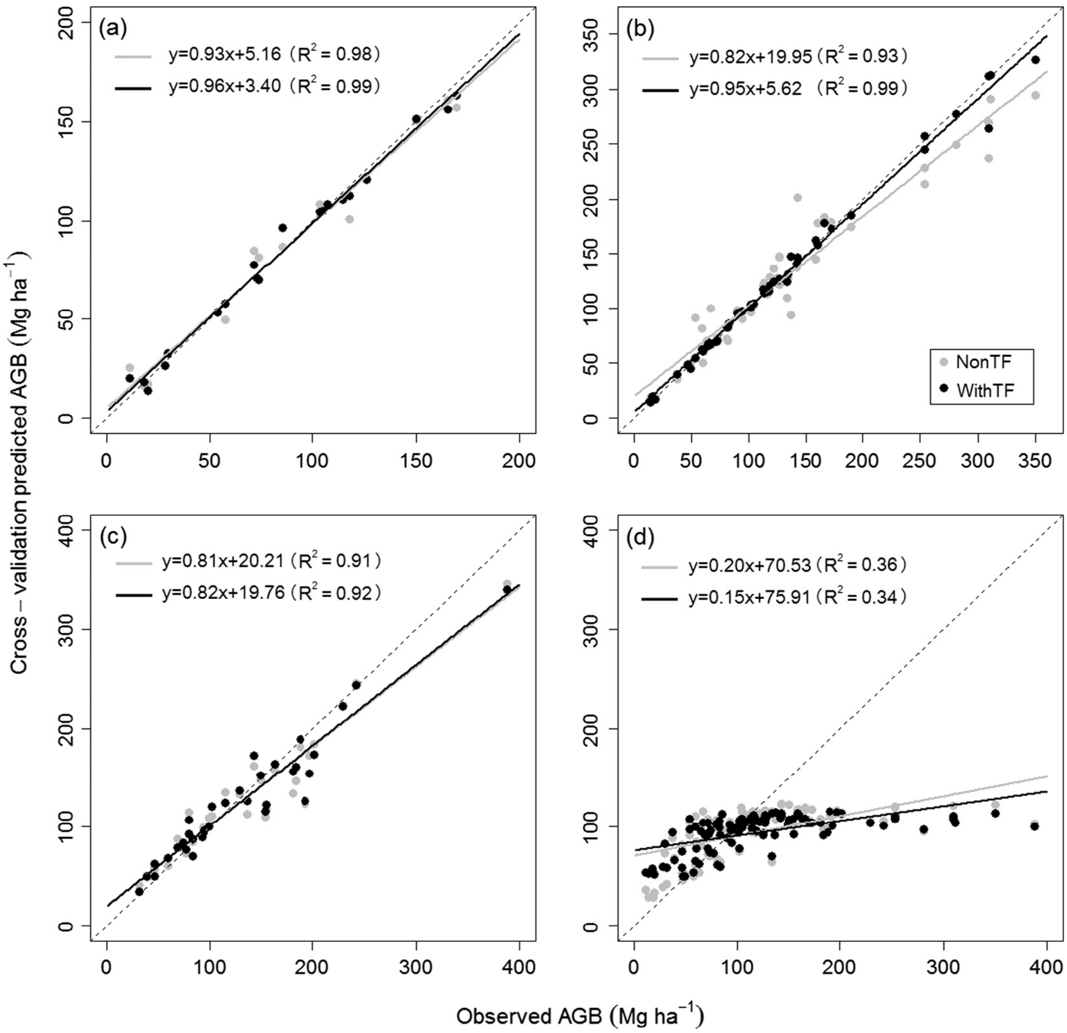

Fig. 6. The relationship between predicted AGB generated using 10-fold cross validation in boosted regression tree (*BRT*) analysis of generalized boosted models (GBMs) and observed AGB of various forest types based on predictor variables from (1) PALSAR data only (**NonTF**), and (2) PALSAR data and topographical and forest stand structure factors (**WithTF**). (a) Needleleaf forest, (b) Mixed forest, (c) Broadleaf forest, and (d) All forest sampling plots.

Table 5
Relative importance (%) of the top five most important predictor variables on forest AGB fitting of different forest types in GBM modelling. **NonTF** represents PALSAR data only, while **WithTF** represents PALSAR data and topographical and forest stand structure factors.

Forest types	NonTF		WithTF	
	Variables	Relative importance (%)	Variables	Relative importance (%)
Needleleaf forest	γ^0_{HH}	34.77	$\gamma^0_{(HH+HV)}{}^2$	42
	γ^0_{HV}	21.33	$\gamma^0_{HV}{}^2$	14.22
	$\gamma^0_{HV}{}^2$	15.19	$\gamma^0_{HH/HV}$	11.21
	$\gamma^0_{HV\text{-}HH}$	11.92	Elevation	10.74
	$\gamma^0_{(HV\text{-}HH)/(HH+HV)}$	5.74	Slope	9.62
Mixed forest	γ^0_{HH+HV}	30.96	Tree height	29.48
	$\gamma^0_{HV\text{-}HH}$	19.66	$\gamma^0_{HV\text{-}HH}$	13.66
	$\gamma^0_{HH/HV}$	14.06	γ^0_{HH+HV}	11.70
	$\gamma^0_{(HV\text{-}HH)}{}^2$	13.25	$\gamma^0_{(HH/HV)}{}^2$	11.00
	γ^0_{HH}	9.95	TWI	10.80
Broadleaf forest	$\gamma^0_{(HH+HV)}{}^2$	23.04	γ^0_{HV}	24.72
	γ^0_{HV}	21.72	$\gamma^0_{HV\text{-}HH}$	16.91
	$\gamma^0_{(HH/HV)}{}^2$	14.18	Relief	14.49
	$\gamma^0_{HV\text{-}HH}$	14.09	Slope	13.36
	$\gamma^0_{HV}{}^2$	9.95	γ^0_{HH+HV}	13.17
All forest plots	γ^0_{HV}	29.61	$\gamma^0_{HV}{}^2$	22.98
	$\gamma^0_{HV\text{-}HH}$	20.84	Tree height	12.33
	γ^0_{HH}	17.52	TWI	12.14
	$\gamma^0_{HH/HV}$	15.57	$\gamma^0_{(HH/HV)}{}^2$	11.56
	γ^0_{HH+HV}	9.97	Relief	11.06

are partly in line with previous studies (Saatchi et al., 2007; Carreiras et al., 2012) and suggest that γ^0_{HH} (X1), γ^0_{HV} (X2), γ^0_{HH+HV} (X3), and $\gamma^0_{(HH+HV)^2}$ (X9) are the most important factors in nonlinear regression models used to estimate forest AGB (Table 5). The possible reason is that the stand compositions of the forests and topographical conditions in our study is quite different, and it makes differences in the sensitivity of various PALSAR backscatter coefficients.

Nonlinear regression models (both GLM and GBM) that constructed on individual forest types all had higher correlations ($R^2 > 0.95$) but lower *RMSE* (<20 Mg ha^{-1}) than those of all forest plots (Table 4). Similarly, fitting results of nonlinear regression models, improvement of the correlations between observed AGB and fitted or cross validation predicted AGB of various forest types were also identified. This is in line with the fitting results between AGB and single PALSAR backscatter coefficients (Fig. 4), and it demonstrates that constructing nonlinear regression models between AGB and PALSAR data based on individual forest types can enhance the predictability. Originally, the complexity of stand structure and species compositions are different among needleleaf forest, mixed forest, and broadleaf forest (Zenner and Hibbs, 2000; McElhinny et al., 2005), just as the variation of the understory AGB among different forest types in this study. Although the proportion of the understory AGB to the total AGB is generally less than 2% and (Table S2), the understory layer still attenuates the PALSAR backscatter signal. Better performance of nonlinear regression models in needleleaf forest is attributed to its simple and clear understory composition (Ma et al., 2016). The predictability nonlinear regression models will certainly decline when various forest types are mixed together, which induces higher complexity of change patterns of PALSAR data with forest AGB.

In all forest plots, both fitted AGB in GLM models and cross validation predicted AGB in GBM models were underestimated when real observed AGB higher than 120 Mg ha^{-1} (Fig. 5–7). This phenomenon is also emerged in several previous studies (Englhart et al., 2011; Carreiras et al., 2012; Peregon and Yamagata, 2013), and the beginning point of underestimation ranges from 100 Mg ha^{-1} to 200 Mg ha^{-1}. The explanation is mainly that predicted AGB in 10-fold cross validation statistic approach may be skewed in nonlinear regression models (Friedman, 2001). Besides, most of the underestimated samples belong to mixed and broadleaf forests which have higher AGB than needleleaf forest. It is possible that the complex stand structure and understory composition of mixed and broadleaf forests decreased the predictability of regression model of all forest plots and then AGB was underrated. However, this problem is solved when constructing regression models using data of each forest type, which was shown in high correlation between observed AGB and either fitted or predicted AGB of individual needleleaf, mixed, and broadleaf forest. Especially in needleleaf forest, GBM regression model well fitted the relationship between AGB and PALSAR data (*RMSE* = 3.81 Mg ha^{-1}, $R^2 = 0.98$), and this suggests that PALSAR data may widely be used to estimate needleleaf forest AGB in high latitude area. Our results also suggest that accurate information on different forest types is essentially for the estimation of forest AGB using nonlinear regression models. Several global and regional maps of forests are already available (Friedl et al., 2010; Gong et al., 2013; Grekousis et al., 2015) and need to be carefully investigated for their likely effects on AGB estimation in the AGB.

4.3. Comparing performance between GLM and GBM models

In this study, we evaluated two popular nonlinear regression models ("*Poisson*" regression of GLM and *BRT* analysis of GBM), and both of them showed good fittings of the relationship between forest AGB and PALSAR backscatter coefficients (Table 4). This suggests that nonlinear regression is an appropriate method to fit the relationship between forest AGB and PALSAR data, which is widely used in AGB estimation from remote sensing data (Garestier and Le Toan, 2010; Morel et al., 2011). The improvement of fitting statistics of the GBM models over the GLM models in individual forest types demonstrates that *BRT* analysis outperforms the "*Poisson*" regression in constructing the nonlinear relationship. *BRT* analysis couples the advantages of decision tree and boosting simultaneously and have been tested in quite a lot researches in prediction or classification (Carreiras et al., 2006; Ma et al., 2016). Better performance of the GBM models was also shown in the high correlations (all $R^2 > 0.9$) between cross validation predicted AGB and observed AGB for various forest types.

Some other regression algorithms have also been used in fitting relationship between forest AGB and PALSAR data. For example, bagging stochastic gradient boosting algorithm was adopted to fit the regression model between AGB and PALSAR backscatter coefficients in tropic forest (Carreiras et al., 2012), and the correlation (R^2) between predicted AGB and observed AGB was 0.144 that is far lower than the correlation in our study. This indicates *BRT* analysis has a considerable ability in fitting regressions between AGB and PALSAR backscatter coefficients.

4.4. The effect of topographical and stand structure factors on nonlinear regression models

Fitting results in nonlinear regression models for different forest types and all forest plots were enhanced (Table 4) when topographical and stand structure factors were added into predictor variables. Meanwhile, correlations between observed AGB and either fitted AGB in the GLM models or cross validation predicted AGB in the GBM models for various forest types also increased (Figs. 5–7). This shows that topographical factors and stand structure factors have certain impacts on forest AGB estimation in regression models, which has also been reported by previous studies (Takyu et al., 2003; Tateno et al., 2004). Moreover, PALSAR backscatter coefficients are also influenced by topographical and stand structure factors. Slope, aspect, wetness, and spatial and vertical structure in canopy all have some impacts on PALSAR backscatter coefficients (Attarchi and Gloaguen, 2014). Both response and predictor variables in nonlinear regression models are affected by topographical and stand structure factors, and the fitting results will be certainly influenced.

Stand structure factors emerged in the top five important variables that influence AGB in needleleaf forest, mixed forest, and broadleaf forest when these factors incorporated into *BRT* analysis in the GBM models (Table 5). In this study, we found that tree height (N1) is an important variable in the GLM and GBM regression models for mixed forest (29.5%) and all forest plots (12.3%). This reflects the important role of stand structure factors in estimating forest AGB. Canopy height is highly related to AGB and widely used in predicting forest AGB from allometric equations or LiDAR (Light Detection And Ranging) inversion (Chave et al., 2005; Saatchi et al., 2011). A global forest canopy height map, generated from satellite observation data, is now available (Lefsky, 2010), and thus it is feasible to the index of forest height to estimate forest AGB.

PALSAR data are affected by topography (Rosenqvist et al., 2007; Shimada et al., 2009). Although topographic correction of PALSAR data was conducted using digital elevation model (DEM), our results showed that relief amplitude (M5) have moderate impact on forest AGB in nonlinear regression models for broadleaf forest (14.5%) and all forest plots (11.1%). This indicates that topographic correction using DEM cannot eliminate the negative influence of microtopography on PALSAR. Moreover, topographic wetness index (M7) also has some impact on estimating forest AGB using nonlinear regression models. Topographic wetness index is calculated based on flow accumulation of each cell of land and represents the moisture of the circumstance at some extent. The main reason may attribute to the high sensitivity of PALSAR data to moisture and snow/ice as reported by a previous study (Thiel et al., 2009). Therefore, topographical conditions should be considered carefully especially in future mapping of forest AGB.

5. Conclusions

In this study, "*Poisson*" regression in GLM models and *BRT* analysis in GBM models were both used to construct the nonlinear relationship between forest AGB and PALSAR backscatter coefficients of three forest types in NE China. Topographical and stand structure factors were also evaluated and incorporated in the regression models. Although HV backscatter coefficient has a higher ability in estimating AGB of forest in NE China, both the GLM and GBM nonlinear regression models fit the relationship between forest AGB and PALSAR data better, and the GBM model generally outperformed the GLM models in estimating forest AGB. Regression models constructed based on individual forest types are better than that based on all forest plots. Incorporating topographical and stand structure factors into nonlinear regression models can enhance the fitting for forest AGB, especially topographic wetness index and tree height are two important factors for the estimation of forest AGB.

Acknowledgements

This research was funded by China Postdoctoral Science Foundation (2015M581519) and Natural Science Foundation of China (41601181 and 41571408). We thank Rencang Bu, Yu Chang, Zaiping Xiong, and Miao Liu for their field work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foreco.2016.12.020.

References

Attarchi, S., Gloaguen, R., 2014. Classifying complex mountainous forests with l-band sar and landsat data integration: a comparison among different machine learning methods in the hyrcanian forest. Remote Sens. 6, 3624–3647.
Atwood, D.K., Andersen, H.E., Matthiss, B., Holecz, F., 2014. Impact of Topographic Correction on Estimation of Aboveground Boreal Biomass Using Multi-temporal, L-Band Backscatter. Ieee J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 3262–3273.
Austin, J.M., Mackey, B.G., Van Niel, K.P., 2003. Estimating forest biomass using satellite radar: an exploratory study in a temperate Australian Eucalyptus forest. For. Ecol. Manage. 176, 575–583.
Bastin, J.F., Barbier, N., Couteron, P., Adams, B., Shapiro, A., Bogaert, J., De Canniere, C., 2014. Aboveground biomass mapping of African forest mosaics using canopy texture analysis: toward a regional approach. Ecol. Appl. 24, 1984–2001.
Brandeis, T.J., Delaney, M., Parresol, B.R., Royer, L., 2006. Development of equations for predicting Puerto Rican subtropical dry forest biomass and volume. For. Ecol. Manage. 233, 133–142.
Brown, S., 2002. Measuring carbon in forests: current status and future challenges. Environ. Pollut. 116, 363–372.
Brown, S., Gillespie, A.J.R., Lugo, A.E., 1989. Biomass estimation methods for tropical forests with applications to forest inventory data. Forest Sci. 35, 881–902.
Brown, S.L., Schroeder, P., Kern, J.S., 1999. Spatial distribution of biomass in forests of the eastern USA. For. Ecol. Manage. 123, 81–90.
Carreiras, J.M.B., Pereira, J.M.C., Campagnolo, M.L., Shimabukuro, Y.E., 2006. Assessing the extent of agriculture/pasture and secondary succession forest in the Brazilian Legal Amazon using SPOT VEGETATION data. Remote Sens. Environ. 101, 283–298.
Carreiras, J.M.B., Vasconcelos, M.J., Lucas, R.M., 2012. Understanding the relationship between aboveground biomass and ALOS PALSAR data in the forests of Guinea-Bissau (West Africa). Remote Sens. Environ. 121, 426–442.
Cartus, O., Santoro, M., Kellndorfer, J., 2012. Mapping forest aboveground biomass in the Northeastern United States with ALOS PALSAR dual-polarization L-band. Remote Sens. Environ. 124, 466–478.
Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Folster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.P., Nelson, B.W., Ogawa, H., Puig, H., Riera, B., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 87–99.
Chen, C., Zhu, J., 1989. The Biomass Manual of Main Tree Species in Northeastern China. China Forestry Press, Beijing.
Conard, S.G., Ivanova, G.A., 1997. Wildfire in Russian boreal forests – potential impacts of fire regime characteristics on emissions and global carbon balance estimates. Environ. Pollut. 98, 305–313.
Dixon, R.K., Brown, S., Houghton, R.A., Solomon, A.M., Trexler, M.C., Wisniewski, J., 1994. Carbon pools and flux of global forest ecosystems. Science 263, 185–190.
Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813.
Englhart, S., Keuck, V., Siegert, F., 2011. Aboveground biomass retrieval in tropical forests – the potential of combined X- and L-band SAR data use. Remote Sens. Environ. 115, 1260–1271.
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.
Fang, J.Y., Liu, G.H., Xu, S.L., 1996. Biomass and net production of forest vegetation in China. Acta Ecol. Sinica 16, 497–508 (in Chinese).
FAO, 2015. Global Forest Resources Assessment 2015: How Have the World's Forets Changed? FAO, Rome.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.
Garestier, F., Le Toan, T., 2010. Forest modeling for height inversion using single-baseline InSAR/Pol-InSAR Data. IEEE Trans. Geosci. Remote Sens. 48, 1528–1539.
Goetz, S.J., Baccini, A., Laporte, N.T., Johns, T., Walker, W., Kellndorfer, J., Houghton, R.A., Sun, M., 2009. Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. Carbon Balance Manage. 4. 2-2.
Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM + data. Int. J. Remote Sens. 34, 2607–2654.
Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimation of leaf area index, f(APAR), and net primary production of terrestrial ecosystems. Remote Sens. Environ. 70, 29–51.
Grekousis, G., Mountrakis, G., Kavouras, M., 2015. An overview of 21 global and 43 regional land-cover mapping products. Int. J. Remote Sens. 36, 5309–5335.
Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. Int. J. Remote Sens. 21, 1331–1364.
Houghton, R.A., Lawrence, K.T., Hackler, J.L., Brown, S., 2001. The spatial distribution of forest biomass in the Brazilian Amazon: a comparison of estimates. Glob. Change Biol. 7, 731–746.
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.
Jiang, H., Apps, M.J., Peng, C.H., Zhang, Y.L., Liu, J.X., 2002. Modelling the influence of harvesting on Chinese boreal forest carbon dynamics. For. Ecol. Manage. 169, 65–82.
Jobidon, R., 2000. Density-dependent effects of northern hardwood competition on selected environmental resources and young white spruce (Picea glauca) plantation growth, mineral nutrition, and stand structural development - a 5-year study. For. Ecol. Manage. 130, 77–97.
Kane, V.R., Lutz, J.A., Roberts, S.L., Smith, D.F., McGaughey, R.J., Povak, N.A., Brooks, M.L., 2013. Landscape-scale effects of fire severity on mixed-conifer and red fir forest structure in Yosemite National Park. For. Ecol. Manage. 287, 17–31.
Karjalainen, M., Kankare, V., Vastaranta, M., Holopainen, M., Hyyppa, J., 2012. Prediction of plot-level forest variables using TerraSAR-X stereo SAR data. Remote Sens. Environ. 117, 338–347.
Laurin, G.V., Puletti, N., Chen, Q., Corona, P., Papale, D., Valentini, R., 2016. Above ground biomass and tree species richness estimation with airborne lidar in tropical Ghana forests. Int. J. Appl. Earth Obs. Geoinf. 52, 371–379.
Lefsky, M.A., 2010. A global forest canopy height map from the moderate resolution imaging spectroradiometer and the geoscience laser altimeter system. Geophys. Res. Lett. 37.
Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2002. Lidar remote sensing for ecosystem studies. Bioscience 52, 19–30.
Lefsky, M.A., Harding, D., Cohen, W.B., Parker, G., Shugart, H.H., 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. Remote Sens. Environ. 67, 83–98.
Lucas, R., Armston, J., Fairfax, R., Fensham, R., Accad, A., Carreiras, J., Kelley, J., Bunting, P., Clewley, D., Bray, S., Metcalfe, D., Dwyer, J., Bowen, M., Eyre, T., Laidlaw, M., Shimada, M., 2010. An Evaluation of the ALOS PALSAR L-Band Backscatter-Above Ground Biomass Relationship Queensland, Australia: Impacts of Surface Moisture Condition and Vegetation Structure. Ieee J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 576–593.

Lucas, R.M., Mitchell, A.L., Rosenqvist, A., Proisy, C., Melius, A., Ticehurst, C., 2007. The potential of L-band SAR for quantifying mangrove characteristics and change: case studies from the tropics. Aquatic Conservation-Marine and Freshwater Ecosystems 17, 245–264.
Luckman, A., Baker, J., Kuplich, T.M., Yanasse, C.D.F., Frery, A.C., 1997. A study of the relationship between radar backscatter and regenerating tropical forest biomass for spaceborne SAR instruments. Remote Sens. Environ. 60, 1–13.
Ma, J., Bu, R.C., Liu, M., Chang, Y., Han, F.L., Qin, Q., Hu, Y.M., 2016. Recovery of understory vegetation biomass and biodiversity in burned larch boreal forests in Northeastern China. Scand. J. For. Res. 31, 382–393.
Ma, J., Bu, R.C., Liu, M., Chang, Y., Qin, Q., Hu, Y.M., 2015. Ecosystem carbon storage distribution between plant and soil in different forest types in Northeastern China. Ecol. Eng. 81, 353–362.
McElhinny, C., Gibbons, P., Brack, C., Bauhus, J., 2005. Forest and woodland stand structural complexity: Its definition and measurement. For. Ecol. Manage. 218, 1–24.
Mitchard, E.T.A., Saatchi, S.S., Woodhouse, I.H., Nangendo, G., Ribeiro, N.S., Williams, M., Ryan, C.M., Lewis, S.L., Feldpausch, T.R., Meir, P., 2009. Using satellite radar backscatter to predict above-ground woody biomass: a consistent relationship across four different African landscapes. Geophys. Res. Lett. 36.
Moisen, G.G., Freeman, E.A., Blackard, J.A., Frescino, T.S., Zimmermann, N.E., Edwards Jr., T.C., 2006. Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. Ecol. Model. 199, 176–187.
Morel, A.C., Saatchi, S.S., Malhi, Y., Berry, N.J., Banin, L., Burslem, D., Nilus, R., Ong, R. C., 2011. Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data. For. Ecol. Manage. 262, 1786–1798.
Myneni, R.B., Dong, J., Tucker, C.J., Kaufmann, R.K., Kauppi, P.E., Liski, J., Zhou, L., Alexeyev, V., Hughes, M.K., 2001. A large carbon sink in the woody biomass of Northern forests. Proc. Natl. Acad. Sci. U.S.A. 98, 14784–14789.
Naesset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80, 88–99.
Nakaji, T., Ide, R., Takagi, K., Kosugi, Y., Ohkubo, S., Nasahara, K.N., Saigusa, N., Oguma, H., 2008. Utility of spectral vegetation indices for estimation of light conversion efficiency in coniferous forests in Japan. Agric. For. Meteorol. 148, 776–787.
Nichol, J.E., Sarker, M.L.R., 2011. Improved biomass estimation using the texture parameters of two high-resolution optical sensors. IEEE Trans. Geosci. Remote Sens. 49, 930–948.
Parresol, B.R., 1999. Assessing tree and stand biomass: a review with examples and critical comparisons. Forest Sci. 45, 573–593.
Peregon, A., Yamagata, Y., 2013. The use of ALOS/PALSAR backscatter to estimate above-ground forest biomass: a case study in Western Siberia. Remote Sens. Environ. 137, 139–146.
Piao, S., Fang, J., Ciais, P., Peylin, P., Huang, Y., Sitch, S., Wang, T., 2009. The carbon balance of terrestrial ecosystems in China. Nature 458, 1009–U1082.
Qin, Y.W., Xiao, X.M., Dong, J.W., Zhang, G.L., Shimada, M., Liu, J.Y., Li, C.G., Kou, W.L., Moore, B., 2015. Forest cover maps of China in 2010 from multiple approaches and data sources: PALSAR, Landsat, MODIS, FRA, and NFI. ISPRS J. Photogramm. Remote Sens. 109, 1–16.
R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
Rosenqvist, A., Shimada, M., Ito, N., Watanabe, M., 2007. ALOS PALSAR: a Pathfinder mission for global-scale monitoring of the environment. IEEE Trans. Geosci. Remote Sens. 45, 3307–3316.
Saatchi, S., Halligan, K., Despain, D.G., Crabtree, R.L., 2007. Estimation of forest fuel load from radar remote sensing. IEEE Trans. Geosci. Remote Sens. 45, 1726–1740.
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. Proc. Natl. Acad. Sci. U.S.A. 108, 9899–9904.
Santoro, M., Beer, C., Cartus, O., Schmullius, C., Shvidenko, A., McCallum, I., Wegmueller, 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.
Sarker, M.L.R., Nichol, J., Ahmad, B., Busu, I., Rahman, A.A., 2012. Potential of texture measurements of two-date dual polarization PALSAR data for the improvement of forest biomass estimation. ISPRS J. Photogramm. Remote Sens. 69, 146–166.
Shimada, M., Isoguchi, O., Tadono, T., Isono, K., 2009. PALSAR radiometric and geometric calibration. IEEE Trans. Geosci. Remote Sens. 47, 3915–3932.
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.
Shimada, M., Ohtaki, T., 2010. Generating large-scale high-quality SAR mosaic datasets: application to PALSAR data for global monitoring. Ieee J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 637–656.
Soenen, S.A., Peddle, D.R., Hall, R.J., Coburn, C.A., Hall, F.G., 2010. Estimating aboveground forest biomass from canopy reflectance model inversion in mountainous terrain. Remote Sens. Environ. 114, 1325–1337.
Sorensen, R., Zinko, U., Seibert, J., 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrol. Earth Syst. Sci. 10, 101–112.
Takyu, M., Aiba, S.I., Kitayama, K., 2003. Changes in biomass, productivity and decomposition along topographical gradients under different geological conditions in tropical lower montane forests on Mount Kinabalu, Borneo. Oecologia 134, 397–404.
Tang, X., Liu, D., Wang, Z., Jia, M., Dong, Z., Liu, J., Xu, W., 2012. Estimation of forest aboveground biomass based on remote sensing data: A review. Chin. J. Ecol. 31, 1311–1318.
Tateno, R., Hishi, T., Takeda, H., 2004. Above- and belowground biomass and net primary production in a cool-temperate deciduous forest in relation to topographical changes in soil nitrogen. For. Ecol. Manage. 193, 297–306.
Thiel, C.J., Thiel, C., Schmullius, C.C., 2009. Operational large-area forest monitoring in Siberia using ALOS PALSAR summer intensities and winter coherence. IEEE Trans. Geosci. Remote Sens. 47, 3993–4000.
Vastaranta, M., Holopainen, M., Karjalainen, M., Kankare, V., Hyyppa, J., Kaasalainen, S., 2014. TerraSAR-X stereo radargrammetry and airborne scanning LiDAR height metrics in imputation of forest aboveground biomass and stem volume. IEEE Trans. Geosci. Remote Sens. 52, 1197–1204.
Wang, C.K., 2006. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests. For. Ecol. Manage. 222, 9–16.
Wei, Y.W., Zhou, W.M., Yu, D.P., Zhou, L., Fang, X.M., Zhao, W., Bao, Y., Meng, Y.Y., Dai, L.M., 2014. Carbon storage of forest vegetation under the Natural Forest Protection Program in Northeast China. Acta Ecol. Sin. 34, 5696–5705.
Whittle, M., Quegan, S., Uryu, Y., Stueewe, M., Yulianto, K., 2012. Detection of tropical deforestation using ALOS-PALSAR: a sumatran case study. Remote Sens. Environ. 124, 83–98.
Yang, J., Wang, C., 2005. Soil carbon storage and flux of temperate forest ecosystems in northeastern China. Acta Ecol. Sin., 2875–2882
Yu, D., Zhou, L., Zhou, W., Ding, H., Wang, Q., Wang, Y., Wu, X., Dai, L., 2011. Forest management in northeast China: history, problems, and challenges. Environ. Manage. 48, 1122–1135.
Zenner, E.K., Hibbs, D.E., 2000. A new method for modeling the heterogeneity of forest structure. For. Ecol. Manage. 129, 75–87.
Zhang, X., Ni-meister, W., 2014. Remote sensing of forest biomass. In: Hanes, J.M. (Ed.), Biophysical Applications of Satellite Remote Sensing, pp. 63–98.
Zhao, K., Popescu, S., Nelson, R., 2009. Lidar remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers. Remote Sens. Environ. 113, 182–196.