Effects of Forest Canopy Structure on Forest Aboveground Biomass Estimation Using Landsat Imagery

CHAO LI1, MINGYANG LI1, KOTARO IIZUKA2, JIE LIU3, KEYI CHEN4, AND YINGCHANG LI1

1Co-Innovation Center for Sustainable Forestry in Southern China, College of Forestry, Nanjing Forestry University, Nanjing 210037, China
2Center for Spatial Information Science, The University of Tokyo, Kashiwa 277-8568, Japan
3College of Landscape Architecture, Nanjing Forestry University, Nanjing 210037, China
4Research Institute of Forestry Policy and Information, Chinese Academy of Forestry, Beijing 100091, China

Corresponding author: Mingyang Li (lmy196727@njfu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 31770679, in part by the Priority Academic Program Development of Jiangsu Higher Education Institution (PAPD), and in part by the Top-notch Academic Programs Project (TAPP) of Jiangsu Higher Education Institutions, China, under Grant PPZY2015A062.

ABSTRACT Accurate remote sensing-based forest aboveground biomass (AGB) estimation is important for accurate understanding of carbon accounting and climate change at a large scale. However, over- and underestimation are common in the process, resulting in inaccurate AGB estimations. Here, the AGB was estimated and mapped by combining Landsat 8 images and forest inventory data in western Hunan Province, China. We used forest canopy density (FCD) mapper to quantify the forest canopy structure. The linear model (LR) and piecewise model with FCD gradients (classified by k-means clustering; sparse, medium, and dense) were developed to estimate AGB for each forest type (coniferous, broadleaf, mixed, and total forests). The piecewise model considered the following scenarios: piecewise model using the variables of LR model (PM), and piecewise model using the variables selected for different FCD gradients (PMV). The PM (R²:0.45–0.56) and PMV (R²:0.63–0.75) models showed better agreement between observed and predicted AGB than the LR (R²:0.18–0.27) models, and the PMV model was the most accurate for each forest type. The PM and PMV models performed better than LR models at different FCD gradients. The PM and PMV models can better alleviate the over- and underestimations of the LR models. At different FCD gradients, the PMV models had different variables, indicating that the correlation between the AGB and spectral variables was different. Overall, FCD is an important forest parameter that influences AGB estimation, and the piecewise model has potential to improve remote sensing-based AGB estimation.

INDEX TERMS Forest inventory data, forest aboveground biomass, Landsat 8 image, canopy structure, piecewise model, subtropical forest.

I. INTRODUCTION Forest ecosystems play a critical role in global climate change and global carbon cycle [1]. Forest biomass is an important forest ecosystem parameter used in environmental and climate modelling [2] to estimate forest productivity and the carbon sequestration rate [3], [4]. Accurate forest biomass estimation, especially aboveground biomass (AGB), is the basis for research on forest carbon sinks and forest structure and functions.

Remote sensing images are suitable for analyzing the spatiotemporal patterns and change trends of forest biomass. In the past 30 years, remote sensing has become an effective tool for estimating forest biomass at various scales [5]–[9]. Remote sensing-based methods mostly use the reflectance, spectral indexes, and image texture of the image to estimate AGB. Combining forest biomass data from forest plot inventory with various remote sensing variables has reportedly showed great potential for accurate biomass estimation in many countries [10]–[14]. However, the assessment of AGB is still affected by uncertainties when the remote sensing data are used in AGB assessment from plot estimates to
wall-to-wall extrapolation [15], thereby weakening the results at a large scale. Using the Landsat images, Zhao et al. [16] proved that overestimations and underestimations occurred for different forest types when the plot AGB was <40 and >120 Mg/ha, respectively. Rodríguez-Veiga et al. [17] used MODIS and ALOS PALSAR data for AGB estimation in Mexico and found that there were large uncertainties in the estimates and that the AGB was over- or underestimated in many cases. Jung et al. [18] studied the uncertainty caused by the plot location error when forest inventory data in combining with Landsat TM imagery were used for forest carbon estimation. They demonstrated that the relative RMSE values significantly decreased after the location errors were removed from the forest inventory plots. Using the optical and Spaceborne LiDAR data, Sun et al. [19] reported that sample size and prediction method have significant effects on the accuracy of AGB estimation; the accuracy of AGB estimation increased with the sample sizes. Furthermore, the allometric models used in field plot AGB, preprocessing of the images, and structure of forests can also result in inaccurate of remote sensing-based AGB estimation [20]–[22].

The heterogeneity of forest stand structures (both horizontal and vertical structures) may lead to the over- and underestimation of AGB. Sanga-Ngoie et al. [23] demonstrated that the use of Landsat imagery can improve forest biomass and carbon estimation in Japan by considering forest types and stand age structures. Pargal et al. [24] used the Fourier transform textural ordination (FOTO) method to obtain canopy information from the very high-resolution (VHR) imagery; with this approach, non-saturated forest structures can be obtained and AGB estimations can be improved. Li et al. [25] demonstrated that the use of Landsat imagery can improve forest carbon estimation by considering forest canopy structure (forest crown density). In remote sensing-based AGB estimation using Landsat imagery, Main-Knorn et al. [26] reported that overestimations usually occurred in sparse forests with low AGB values (density below 0.4), and underestimations usually occurred in relatively dense forests (density above 0.9) with high AGB values.

However, only a few studies have evaluated the effects of FCD on remote sensing-based AGB estimation although the optical sensor mainly provides the information of forest canopy. The relationships between forest parameters and spectral variables may be different under different forest structures. Therefore, the AGB was modelled by spectral variables under different forest types and FCD gradients to investigate the effects of forest structure on the accuracy of AGB estimation. Here, we derived forest canopy information from Landsat 8 images, covering thousands of hectares of coniferous forests (CFF), broadleaf forests (BLF), and mixed BLF and CFF forests (MXF) in subtropical region in China. We combined the spectral variables for AGB estimations with the FCD gradients to: (i) assess the potential use of FCD in AGB estimation; and (ii) infer regional-scale AGB distribution in subtropical forests.

### II. MATERIALS AND METHODS

#### A. STUDY AREA

Located in the western Hunan Province, the study area is surrounded by the Wuling, Xuefeng Mountains, and Yunnan - Guizhou Plateau (Figure 1). The topography of the region is complex and the environment is fragile. The climate in the study area is a typical subtropical monsoon humid climate with continental characteristics, influenced by the mountains. This is a key forestry area in Hunan Province with rich natural resources. The regional forests are dominated by evergreen coniferous and broadleaf forest, deciduous and mixed forests [27].

#### B. CALCULATION OF PLOT AGB

We used 681 forest plots that were systematically allocated based on a grid of 4 km × 8 km on the map of Hunan province [28], [29] including CFF, BLF, and MXF. All plots were 0.067 ha in size and were surveyed by the National Forest Continuous Inventory (NFCI) in 2014. The conversion equations between the stand volume and biomass developed by Li et al. [30] was used to calculate the plot AGB. The spatial distribution of the forest inventory sample plots is shown in Figure 1. Table 1 presents the plot AGB for different forest types; the average AGB of the plots was 50.37 Mg/ha and the standard deviation was 34.55 Mg/ha. The MXF had the largest mean AGB followed by the BLF, and CFF.

#### C. LANDSAT IMAGES AND INFORMATION EXTRACTION

Two surface reflectance productions of Landsat 8 Operational Land Imager (OLI) [31] acquired in 2014 were downloaded
### TABLE 2. Summary of predictor variables for AGB estimation.

| Type                        | Definition                                                                 |
|-----------------------------|-----------------------------------------------------------------------------|
| Single band (6)             | $B_i = 1/b_i$, $i=2,\ldots,7$                                             |
| Inversions of bands (6)     | $SR_d = b_i/b_j$, $i, j=2,\ldots,7, i \neq j$                             |
| Two-band ratios (30)        | $SR_d = b_i/b_j$, $i, j=2,\ldots,7$                                       |
| Three-band ratios (60)      | Normalized Difference Vegetation Index (NDVI), Atmospherically Resistant Vegetation Index (ARVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI) |
| Vegetation indices (4)      | Mean ($b_{MOD}$), Homogeneity ($b_{MOD}$), Contrast ($b_{CON}$), Dissimilarity ($b_{DIS}$), Entropy ($b_{ENT}$), Angular Second Moment ($b_{SAM}$), Correlation ($b_{COR}$), Variance ($b_{VAR}$) |
| Textural images (144)       | Using moving window size with $j \times j$, $j=3,5$, and 7 respectively |

(https://earthexplorer.usgs.gov/) and used in this study. The path of the images was 125 and the rows of the images were 40 and 41, respectively. Radiometric and atmospheric correction of the images were already performed by USGS. Therefore, to eliminate the influence of topography, topographic correction was performed using the C-correction approach [32]. The 2014 forest classification (CFF, BLF, and MXF) map of Hunan province was downloaded from the ESA website (http://maps.elie.ucl.ac.be/CCI/viewer/index.php).

Spectral variables (Table 2) including the original bands (bands 2–7), vegetation indices, and various image transformations (inversions, band ratios, and texture measures) [33] were calculated as independent variables to estimate AGB.

In addition, the FCD Mapper, which was developed by the International Tropical Timber Organization (ITTO) and the Japan Overseas Forestry Consultants Association [34], [35], were used to map the FCD of the study area. Forest canopy density (FCD) represents the proportion of the total area covered by the crown projection [36] and is an important factor related to forest structure. In forest inventory, FCD is a key factor that can reflect the degree of canopy closure, tree space utilization, stand density, and forest growth [37]. In addition, FCD is an important indicator used to determine the strength of tending cutting, and estimate the AGB and forest volume. There is no unified criterion to define the classes of FCD. It was difficult to know a priori, how many FCD gradients would be necessary for piecewise model. K-means clustering is commonly used unsupervised non-hierarchical clustering algorithm for data grouping, it is intuitive, efficient, and rapid [38], [39]. K-means clustering divides the data into k clusters ensuring that the data in one cluster have high similarity, whereas have low similarity with the data in other clusters [40]. In this study, k-means clustering algorithm was used to classify the FCD.

**D. AGB ESTIMATION MODEL**

Stepwise regression was used to select the optimal variables for remote sensing-based AGB models by the correlation analyzing results between plot AGB and the spectral variables. Linear regression models (LR) were developed using the selected variables as predictors and the AGB from plot inventory data as response variable for different forest types.

The accuracy of the remote sensing-based AGB model was very low in dense and sparse forests. Therefore, the piecewise linear model [41] for different forest types, which divided the whole data into segmented consecutive linear models with different coefficients, was developed according to the FCD gradients. The PM model of each forest type was developed using the variables of LR models. To ensure each FCD class has a sufficient number of forest plots, k-means clustering was used to test the best FCD classes for the PM models by taking different k values (2, 3, 4, 5, and 6). Besides, under the optimal cluster of the FCD, the relationships between AGB of different segments and spectral variables may be different. Therefore, in this study, we also selected spectral variables for each segment. The piecewise model with different variables in each segment (PMV) for different forest types was also developed.

Different scenarios of this study were designed based on the three models and forest types: (1) the LR and PM models were developed using the LR model variables as independent variables, and the PMV models were developed using the variables selected for each FCD gradient; (2) 12 AGB models were developed based on three modelling algorithms and four forest types. The following workflow was used in Figure 2.

**E. MODEL COMPARISON AND EVALUATION**

Ten-fold cross validation was performed, and the coefficient of determination ($R^2$), root mean square error (RMSE) and relative RMSE (rRMSE) [42] were calculated for each model to evaluate the performance of the models. The RMSE and rRMSE values for each segment were calculated and compared. Besides, the model residuals were also calculated and analyzed. Using the best model, 10 AGB maps were generated by 10-fold cross validation. The mean and standard deviation (Std) of 10 AGB maps were calculated as final AGB map and the uncertainty map, respectively.

In addition, we tested whether the improving of PM model resulted from the decrease in the sample size. To this end, we randomly selected $n$ subsamples ($n$ ranged from 10 to
the largest subsample number of a forest type (e.g., 107 for CFF) of each forest type, and then the RMSE distribution of predicted AGB for the subsamples was analyzed. The selection of random subsamples for each forest type is a cyclic process, which is shown in Figure 3:

- Subset generation: a subsample was randomly selected according to the determination condition of \( n \). The start point of the subsample selection was 10 (\( n = 10 \)). The dataset was input into the LR model.
- Subset modelling: the AGB of the subsamples was modelled to calculate the RMSE. For each \( n \), the subset generation and modelling steps were repeated 1000 times; then, 1000 RMSE values were generated for each \( n \).
- Stopping criterion: used to determine when the subsample generation step should stop. Here, we set two stopping criteria: (1) for each \( n \), we counted the number of RMSE values, and when the number reached 1000, the subset generation and modelling steps were stopped; and (2) if the \( n \) was larger than the largest subsample number at an FCD gradient.

III. RESULTS

A. FCD GRADIENTS AND MODEL VARIABLES

We tested the different \( k \) values of the PM models in this study (Figure 4). For BLF and total forests, the PM model had the largest \( R^2 \) when the \( k = 3 \). For the MXF and CFF forests, the increases of \( R^2 \) values were not significant from \( k = 3 \) to \( k = 6 \). In addition, when the \( k \leq 3 \), all the segmented linear models were significant; when the \( k > 3 \), at least one non-significant segmented linear model was obtained for each forest type. Therefore, the \( k = 3 \) was selected to divide the FCD gradients. The optimal \( k \) value was also used in the PMV models to compare the PM and PMV models. According to the \( k \) value, the FCD map was divided into three gradients, namely, sparse, medium, and dense. The statistics of the AGB in each FCD gradient are shown in Table 5.

The variables of the LR, PM and PMV models for all forest types were selected based on the correlations of plot AGB and spectral variables (Table 3). Three spectral variables were selected in the LR and PM models of each forest type. For the PMV models, the selected variables of each segment were different for each forest type.

B. AGB MODEL PERFORMANCES

Using the selected variables, 12 AGB models were generated. By cross validation, the fitting results of all the models are...
shown in Table 4. The LR models ($R^2 < 0.28$) could explain less than 30% variations of the plot AGB for each forest type. The $R^2$ value of the total forest LR model was only 0.18 with an RMSE value of 32.59 Mg/ha. The MXF LR model had the highest $R^2$ value of 0.27 with an RMSE value of 27.7 Mg/ha. For the CFF and BLF LR models, the $R^2$ values were 0.23 and 0.24, the RMSE values were 24.43 Mg/ha and 35.27 Mg/ha, respectively. The PM and PMV models for each forest type were more accurate than the LR models. The $R^2$ values of the PM models ranged from 0.45 to 0.56, while the RMSE values ranged from 18.67 to 28.6 Mg/ha. The $R^2$ value of the CFF PM model increased the most (0.33) compared with that of the LR model, and the RMSE value of the total PM model decreased the most (6.9 Mg/ha) compared with that of the LR model. The PMV models explained the most plot AGB variations for each forest type. The $R^2$ values of the PMV models ranged from 0.63 to 0.75, while the RMSE values ranged from 14.95 to 23.48 Mg/ha. The $R^2$ value of the CFF PMV model increased the most (0.49) compared with that of the LR model, and the RMSE value of the BLF PMV...
model decreased the most (11.79 Mg/ha) compared with that of the BLF LR model. In addition, for these forest types, the $R^2$ values of the PMV models increased by more than 0.15 compared with that of the PM models, and the RMSE values decreased by more than 3.7 Mg/ha compared with that of the PM models. The rRMSE values of PMV models were all lowest. The fitting results of the PMV models were the best for each forest type compared with those of the LR and PM models.

The scatterplots of these models are shown in Figure 5. For the scatterplots of LR models, overestimations were detected when the plot AGB value was $<30$ Mg/ha, and underestimations were appeared when the plot AGB was $>90$ Mg/ha, in each forest type. For the scatterplots of PM models, over-estimations and underestimations were observed when the AGB values was $<20$ Mg/ha and $>140$ Mg/ha, respectively, in each forest type. The scatterplots of the PMV models were the most closet to the 1:1 line than those of the LR and PM models, over- and underestimations were significantly alleviated.

C. COMPARISONS OF AGB ESTIMATION AT DIFFERENT FCD GRADIENTS

The RMSE and rRMSE values at three FCD gradients were calculated (Figure 6). The RMSE values of the PMV models were significantly lower than those of the PM and LR models in the sparse, medium, and dense plots, respectively (Figure 6a). Furthermore, when compared with the LR models, the RMSE values of the PMV and PM models decreased more in the sparse and dense plots than in the medium and total plots. The rRMSE values of the LR models were all larger than 50%, and were significantly higher than those of the PM and PMV models, especially in the sparse plots (Figure 6b). In addition, the residuals of the LR, PM, and PMV models in the sparse and dense plots were calculated (Figure 7). When the LR models were used, more than 50% of the residuals in the sparse plots were $<-15$ Mg/ha and approximately 50% of the residuals in the dense plots were $>15$ Mg/ha. This showed that the models had serious over- and underestimations in these two FCD gradients, respectively. Most of the residuals of the PMV and PM models in the sparse and dense plots were between $-15$ to $15$ Mg/ha, and these two models were better than LR model and can alleviate the over- and underestimation. The RMSE and residuals in the sparse and dense plots of the PMV models were the best compared with those of the PM and LR models.

We tested whether the improvement in the PM model resulted from the decrease in the sample size (Figure 8). The mean RMSE values of random samples were 24.54 Mg/ha (CFF), 35.97 Mg/ha (BLF), 26.18 Mg/ha (MXF), and 30.63 Mg/ha (Total), which were consistent with the RMSE values of the LR models, but were higher than those of the PM and PMV models. The results indicated that the improvement in the PM model was not caused by the decrease in the
sample size. In addition, the RMSE values of the sparse and medium plots of the PM and PMV models and the RMSE values of dense plots of the PMV models were lower than the mean RMSE values of random samples, indicating that using the FCD gradients can increase the model quality in AGB estimations.

**D. AGB DISTRIBUTION PATTERN AND UNCERTAINTY**

The AGB and uncertainty distributions of the study area were mapped and analyzed using the PMV models of forest types (Figures 9 and 10). The southeastern and northwestern parts of the study area presented high AGB values (>120 Mg/ha); the northeastern part presented low AGB values (<60 Mg/ha) (Figure 9a). A large proportion of the region presented an AGB value of <60 Mg/ha, followed by areas with an AGB value of 60–120 Mg/ha. Only less than 3.5% of the region had an AGB of >150 Mg/ha (Figure 10a). The uncertainty (Std) for a large proportion of the study area was <4 Mg/ha (Figure 10b). The distribution of uncertainty consistent with the distribution of AGB. The area with higher AGB presented higher uncertainty, whereas the area with lower AGB presented relatively lower uncertainty (Figure 9).

---

**FIGURE 7.** Frequency distribution of the residuals using three models in sparse and dense plots by forest type.

**FIGURE 8.** Distribution of the RMSE for the predicted AGB of random samples.
IV. DISCUSSION

A. IMPROVING AGB ESTIMATION

In this study, significant over- and underestimation were observed when the remote sensing-based LR AGB models were used. The relationships between spectral variables and AGB may not be simply linear, and the LR model cannot handle complex non-linear relationships. This is one of the reasons that the LR model could only reflect approximately 25% of the AGB variations in this study. Obviously, the non-linear models or the nonparametric models may improve the AGB estimation as reported previously [33], [43]. However, there were still over- and underestimations in remote sensing-based AGB estimation using the nonlinear or nonparametric models [44], [45]. We assumed that the FCD gradient significantly affects the accuracy of the remote sensing-based AGB estimates, and we proposed to use the piecewise model to alleviate uncertainties and improve AGB estimation accuracy. Finally, the piecewise models (both the PM and PMV models) performed better than the LR models in all forest types, and the over- and underestimation of the LR models were significantly alleviated by piecewise models. In addition, the PMV models in this study performed better than the nonparametric remote sensing-based AGB models used in other studies in Hunan province [44], [46], especially in the relatively low and high AGB ranges. Therefore, our results demonstrated that the piecewise models by FCD gradients were useful in improving AGB estimation.

In the estimation of AGB using the LR, PM, and PMV models, the selected variables for each forest type were different. Furthermore, the stratified models by forest type obviously improved the estimations compared with the total forest models, and this consistent with the findings of the previous study [47]. In addition, we selected the variables for each FCD gradient and developed the PMV models. The results showed that different relationships existed between the AGB and spectral variables under different forest structures (forest type and FCD gradient). Incorporating forest structures into remote sensing-based AGB models might result in improved performance.
B. EFFECTS OF FCD GRADIENTS ON AGB ESTIMATION

The FCD value is significantly correlated with forest growth, volume, and biomass. Dense forests (high FCD values) are fully closed and show strong crown competition, which could influence tree growth. High FCD value results in the saturation problem in the optical sensors. Sparse forests (low FCD values) are not fully closed and show weak crown competition. Low FCD value results in the spectral mixture problem in the optical sensors.

The residuals of the LR, PM, and PMV models were calculated, and when the residuals were \(-15\) Mg/ha and \(>15\) Mg/ha, the estimations were proposed to be significantly uncertain. In this study, significant overestimations were observed for the sparse plots in the LR models and significant underestimations were observed for the dense plots in the LR models. These results indicated that the uncertainties that the LR models had mainly existed in the sparse and dense plots; this is consistent with the findings of previous study [26]. The optical sensors are insensitive to the forest biomass change when the forest reaches a certain density or has a complex structure; this results in the underestimation of biomass values [8], [48]–[50]. This is the saturation problem of the optical sensors. Gemmell [51] identified the problem of data saturation when the forest biomass was estimated using Landsat in the Rocky Mountains. Lu [52] found that high AGB values were easily underestimated in the Brazilian Amazon. In the sparse or young forests with low biomass values, the AGB is always overestimated due to the mixed effects of the undergrowth vegetation and soil. Eriksson et al. [53] pointed out that understory vegetation affected the radiation flux in remote sensing images of forest canopies. Pisek et al. [54] stated that the influence of understory vegetation was more apparent in areas where the stand density is low and caused uncertainty in the leaf area index inversion. The spectral mixture and saturation problems reduced the differences in spectral variables for different forest structures. The relations between plot AGB and spectral variables cannot be fully reflected especially in the sparse and dense forests. The piecewise models in this study divided the whole plots into different segmented linear regressions, and it can fully reflect the relationships between AGB and spectral variables in different segments. Using the piecewise models, which consider the FCD gradient, the overestimations in sparse plots and underestimations in very dense plots were alleviated, indicating that considering the forest canopy structure can solve the saturation and spectral mixture problems in AGB estimation to some extent. Therefore, the forest canopy structure is a key factor that affects the accuracy of AGB estimation.

C. SEGMENTS (K VALUES) OF THE PIECWISE MODEL

The PM and PMV models were all piecewise linear regression models, and the basis of segments was FCD gradients in the whole area. In this study, k-means clustering algorithm was used to classify the FCD and when the k was 3, the PM models performed the best. In addition, in the PM models of different forest types, no significant segmented linear models existed when the k was >3. This result indicated that the variables selected by the LR models cannot fully explain the AGB at different FCD gradients when the gradients were >3. The PMV models were developed by selecting variables for different FCD gradients, therefore, all segmented linear models in the PMV models were significant. However, the three gradients of the FCD were also used in the PMV models in this study for comparison with the PM models. In the future, more k values should be tested in the PMV models to identify the best segment number of piecewise model to obtain the most accurate AGB results. In addition, by comparing the RMSE values of random samples and each segment (FCD gradient), we concluded that the increase in the PM model quality was not caused by the decrease in sample size due to the FCD gradients.

D. AGB IN THE STUDY AREA

The study region is a key forestry but a rural mountain area with a low level of economic development in Hunan Province. According to the AGB spatial map, the southeastern and northwestern parts with the relative high altitude in the study area (mountain) had the relatively high AGB values; the northeastern part where the altitude is relatively low (plain) had the low AGB values. This distribution pattern was similar to the pattern reported by previous study [44]. Due to the low forest management level, attach importance to afforestation and despise to manage, the forest structure is unreasonable. Therefore, more than 60% of the forests had lower AGB values (<60 Mg/ha) and only 3.5% of the forests had relatively high AGB values (>150 Mg/ha) (Figure 9a). In addition, although the distribution of the FCD is consistent with the AGB, the distributions of some parts still differed. In the remote mountain regions, the young growth and half-mature forests are difficult to thin in time resulting in high stand densities and a low biomass. In the mature and over-mature forests, the timber cannot be harvested in time because of the limitation of the Natural Forest Protection Policy, and this may affect the afforestation and regeneration of forests and result in low stand densities and a high biomass. Therefore, slight over- and underestimations were still observed in the PM and PMV models.

V. CONCLUSION

Estimation models of forest AGB in western Hunan Province was developed by combining the Landsat 8 OLI spectral variables and FCD gradients, namely LR (linear model), PM model (piecewise model), and PMV model (piecewise model using variables for different FCD gradients). The three models were used to develop an AGB prediction framework by considering the stratification and non-stratification of forest types. As evaluated using plot inventory AGB and spatial distribution map, the AGB estimations of the PMV models were the most accurate. Using the FCD gradients, the PM and PMV models performed better than the LR models and alleviated the uncertainties associated with the LR models.
Besides, the relationship between AGB and spectral variables in different FCD gradients were different. The approach developed in this study will provide more accurate AGB estimates at a regional scale. This will promote better understanding of the forest carbon stocks and provide basic information for forest management at a regional scale. In other areas, our approach is easily replicable for AGB/carbon estimation using Landsat images, but the variables in the models may differ because of different regional conditions.

APPENDIX

| TABLE 5. Statistics of plot AGB by FCD gradients. |
|-----------------------------------------------|
| Forest type | FCD gradients | No. of plots | Minimum (Mg/ha) | Mean (Mg/ha) | Maximum (Mg/ha) | Std (Mg/ha) |
|--------------|---------------|--------------|----------------|--------------|----------------|-------------|
| Sparse       | Medium        | 107          | 27.33          | 61.59        | 12.88          |
| CFF          | Medium        | 100          | 36.11          | 40.31        | 18.19          |
| Dense        |               | 68           | 72.96          | 154.48       | 34.14          |
| Sparse       | Medium        | 83           | 29.77          | 61.42        | 24.65          |
| BLF          | Medium        | 91           | 54.66          | 164.82       | 33.63          |
| Dense        |               | 86           | 77.66          | 223.35       | 46.12          |
| Sparse       | Medium        | 43           | 36.49          | 113.34       | 20.72          |
| MXF          | Medium        | 54           | 56.79          | 146.68       | 28.19          |
| Dense        |               | 49           | 72.13          | 187.60       | 34.96          |
| Sparse       | Medium        | 235          | 1.05           | 29.89        | 114.42         |
| Total        |               | 245          | 49.78          | 164.82       | 27.78          |
| Dense        |               | 203          | 74.58          | 223.35       | 38.72          |

ACKNOWLEDGMENT

The authors thank the ESA (http://maps.elie.ucl.ac.be/CCI/viewer/index.php) and USGS (http://glovis.usgs.gov/) for providing the Land Cover map and Landsat 8 OLI images. They also thank their editors, as well as the anonymous reviewers for their valuable comments.

REFERENCES

[1] R. K. Dixon, A. M. Soloman, S. Brown, R. A. Houghton, M. C. Trexier, and J. Wisniewski, “Carbon pools and flux of global forest ecosystems,” Science, vol. 263, no. 5144, pp. 185–190, Jan. 1994.
[2] L. Du, T. Zhou, X. Zhao, H. Wu, and D. Wu, “The spatial distribution of forest biomass in China using remote sensing and national forest inventory,” in Proc. IEEE Geosci. Remote Sens. Symp., Jul. 2014, pp. 737–740.
[3] J. Fang, “Changes in forest biomass carbon storage in China between 1949 and 1998,” Science, vol. 292, no. 5525, pp. 2320–2322, Jun. 2001.
[4] S. Luo, C. Wang, X. Xi, F. Pan, D. Peng, J. Zou, S. Nie, and H. Qin, “Fusion of airborne LiDAR data and hyperspectral imagery for above-ground and belowground forest biomass estimation,” Ecol. Indicators, vol. 73, pp. 378–387, Feb. 2017.
[5] D. Timothy, M. Ouisimo, S. Cletah, S. Adelabu, and B. Tisiti, “Remote sensing of aboveground forest biomass: A review,” Trop. Ecol., vol. 57, no. 2, pp. 125–132, 2016.
[6] J. M. Barbosa, E. N. Broadbent, and M. D. Bitencourt, “Remote sensing of aboveground biomass in tropical secondary forests: A review,” Int. J. Forest Res., vol. 2014, May 2014, Art. no. 715796.
[7] L. Kumar, P. Sinha, S. Taylor, and A. F. Alqurashi, “Review of the use of remote sensing for biomass estimation to support renewable energy generation,” J. Appl. Remote Sens., vol. 9, no. 1, Jun. 2015, Art. no. 097696.
[8] D. Lu, Q. Chen, G. Wang, L. Liu, G. Li, and E. Moran, “A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems,” Int. J. Digit. Earth, vol. 9, no. 1, pp. 63–105, Jan. 2016.
[9] D. C. Marvin, G. P. Asner, D. E. Knapp, C. B. Anderson, R. E. Martin, F. Sinca, and R. Tupayachi, “Amazonian landscapes and the biases in field studies of forest structure and biomass,” Proc. Nat. Acad. Sci. USA, vol. 111, no. 48, pp. E5224–E5232, Dec. 2014.
[10] A. Safari and H. Sohrabi, “Ability of Landsat-8 OLI derived texture metrics in estimating aboveground carbon stocks of coppice Oak Forests,” Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. XL-I-B8, pp. 751–754, Sep. 2016.
[32] P. M. Teillet, B. Guindon, and D. G. Goodenough, “On the slope-aspect correction of multispectral scanner data,” Can. J. Remote Sens., vol. 8, no. 2, pp. 84–106, Dec. 1982.

[33] H. Sun, G. Qie, G. Wang, Y. Tan, J. Li, Y. Peng, Z. Ma, and C. Luo, “Increasing the accuracy of mapping urban forest carbon density by combining spatial modeling and spectral unmixing analysis,” Remote Sens., vol. 7, no. 11, pp. 15114–15139, Nov. 2015.

[34] A. Rikimaru and S. Miyatake, Development of Forest Canopy Density Mapping and Monitoring Model using Indices of Vegetation, Bare soil and Shadow. Amsterdam: The Netherlands: Geospatial World, 1997.

[35] A. Rikimaru, P. S. Roy, and S. Miyatake, “Tropical forest cover density mapping,” Tropical Ecol., vol. 43, no. 1, pp. 39–47, 2002.

[36] H. Prettsch, Forest Dynamics, Growth and Yield. Berlin, Germany: Springer, 2009.

[37] X. Meng, Forest Mensuration, 3rd ed. Beijing, China: Forestry Publishing House, 2006.

[38] R. Cordeiro de Amorim and B. Mirkin, “Minkowski metric, feature weighting and anomalous cluster initializing in K-means clustering,” Pattern Recognit., vol. 45, no. 3, pp. 1061–1075, Mar. 2012.

[39] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, “An efficient K-means clustering algorithms: Analysis and implementation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 881–892, Jul. 2002.

[40] A. S. Ahmar, D. Napitupulu, R. Rahim, R. Hidayat, Y. Sonatha, and M. Azmi, “Using K-means clustering to cluster provinces in indonesia,” J. Phys., Conf. Ser., vol. 1028, Jun. 2018, Art. no. 012006.

[41] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. New York, NY, USA: Springer, 2009.

[42] A. Safari, H. Sohrabi, S. Powell, and S. Shahtaei, “A comparative assessment of multi-temporal Landsat 8 and machine learning algorithms for estimating aboveground carbon stock in coppice oak forests,” Int. J. Remote Sens., vol. 38, no. 22, pp. 6407–6432, Nov. 2017.

[43] K. Liu, J. Wang, W. Zeng, and J. Song, “Comparison and evaluation of three methods for estimating forest above ground biomass using TM and GLAS data,” Remote Sens., vol. 9, no. 4, p. 341, Apr. 2017.

[44] V. Li, C. Li, M. Li, and Z. Liu, “Influence of variable selection and forest type on forest aboveground biomass estimation using machine learning algorithms,” forests, vol. 10, pp. 1–22, 2019.

[45] M. Karlson, M. Ostwald, H. Reese, J. Sanou, B. Tankoano, and E. Mattsson, “Mapping tree canopy cover and aboveground biomass in sudano-sahelian woodlands using Landsat 8 and random forest,” Remote Sens., vol. 7, no. 8, pp. 10017–10041, Aug. 2015.

[46] J. Zhi, Z. Huang, H. Sun, and G. Wang, “Mapping forest ecosystem biomass density for Xiangjiang River Basin by combining plot and remote sensing data and comparing spatial extrapolation methods,” Remote Sens., vol. 9, no. 3, pp. 1–23, 2017.

[47] Y. Gao, D. Lu, G. Li, G. Wang, Q. Chen, L. Liu, and D. Li, “Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region,” Remote Sens., vol. 10, no. 4, p. 627, Apr. 2018.

[48] D. Lu, Q. Chen, G. Wang, E. Moran, M. Batistella, M. Zhang, G. V. Laurin, and D. Saah, “Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates,” Int. J. Forest Res., vol. 2012, Jun. 2012, Art. no. 436537.

[49] G. M. Foody, D. S. Boyd, and M. E. J. Cutler, “Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions,” Remote Sens. Environ., vol. 85, no. 4, pp. 463–474, Jun. 2003.

[50] D. P. Turner, W. B. Cohen, R. E. Kennedy, K. S. Fassnacht, and J. M. Briggs, “Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites,” Remote Sens. Environ., vol. 70, no. 1, pp. 52–68, Oct. 1999.

[51] F. M. Gemmell, “Effects of forest cover, Terrain, and scale on timber volume estimation with thematic mapper data in a rocky mountain site,” Remote Sens. Environ., vol. 51, no. 2, pp. 291–305, Feb. 1995.

[52] D. Lu, “Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon,” Int. J. Remote Sens., vol. 26, no. 12, pp. 2509–2525, Jun. 2005.

[53] H. M. Eriksson, L. Ekbladh, A. Kuusk, and T. Nilsen, “Impact of understory vegetation on forest canopy reflectance and remotely sensed LAI estimates,” Remote Sens. Environ., vol. 103, no. 4, pp. 408–418, Aug. 2006.

[54] J. Pisk, J. M. Chen, J. R. Miller, J. R. Freemantle, J. I. Peltonenimo, and A. Salmi, “Mapping forest background reflectance in a boreal region using multitemporal compact airborne spectrographic imagery data,” IEEE Trans. Geosci. Remote Sens., vol. 48, no. 1, pp. 499–510, Jan. 2010.