The modeling of above ground biomass in ranges of Corbett Tiger Reserve using dual-polarization ALOS PALSAR data

Y. Kumar *, S. Singh and R.S. Chatterjee

Indian Institute of Remote Sensing, Dehradun, India

Date Received: 05-10-2017    Date Accepted: 07-12-2017

Abstract

The study has been carried out in the Pauri Garhwal district of Uttarakh and keeping the focus on Corbett Tiger Reserve (CTR). The total area of CTR covered in the scene is 889 sq. km. The main aim of the paper is to develop a model by establishing a relationship between backscatter coefficients generated from dual polarization L-band ALOS PALSAR data acquired in July 2008 and the field inventory data collected by Forest Survey of India team in 2010. A total of 120 sample plots data were collected in the area out of which 60 plots were used for the training of the model and the remaining 60 plots were left for the validation of the most significant model. The Simple regression analysis was computed between HH & HV backscatter as independent variable and per plot biomass as dependent variable. The Linear, Logarithmic and Polynomial best fit regression models were analyzed. It was found that the coefficient of determination is more with HV backscatter ($R^2=0.75$) using logarithmic model as compared among HV in linear and polynomial on one hand and HH in linear, logarithmic and polynomial on the other hand. To improve the accuracy and to know the combined effects of both the polarizations, multiple linear regression analysis (MLR) was applied. There was a significant improvement in correlation coefficients ($R^2=0.86$). The in-situ field inventory data shows that the biomass in the CTR ranges from 9.6 t/ha to 322.6 t/ha. The simple regression modelled biomass ranges from 26.2 t/ha to 401.43 t/ha, whereas the MLR modelled biomass ranges from 10.96 t/ha to 312.64 t/ha. The majority of the area was found to be in the range of 100 t/ha to 150 t/ha biomass. The coefficient of determination ($R^2$) between observed and predicted biomass was found to be 0.734 with simple regression, whereas it was found to be 0.83 with MLR.

Key words: Biomass, modeling and remote sensing

1. Introduction

In the regulation of global climate change “Forests” play a key role. It retains large amount of carbon over a long period and thus acts as both sink and source of carbon dioxide (CO₂). The estimation and monitoring of CO₂ source and sink are required for the greenhouse gas inventories, terrestrial carbon accounting and modelling climate change (Dobson et al., 1992; Falkowski et al., 2000; Schimel et al., 2001; Canadell et al., 2004; Houghton, 2005; Schulze, 2006; Heimann and Reichstein, 2008; Le Quere et al., 2009; Loarie et al., 2009). Approximately 50% of the carbon is stored in biomass, thus continuous and effective monitoring is required to estimate the vegetation biomass especially in forest ecosystem.

* Correspondence: yogesh.iirs@gmail.com

Tel: 

ISSN 2235-9370 Print/ISSN 2235-9362 Online © University of Sri Jayewardenepura

108
The forest ecosystem is changing; the rate of change, growth and addition of biomass should be carefully understood to develop a more accurate methodology for the estimation of factors responsible for change. The destruction in the forest ecosystem leads to the emission of CO$_2$ and other greenhouse gases which results in climate change. The climate change will result in large shifts in the distribution of forest biomes which in turn will significantly change the amounts and patterns of carbon storage in these ecosystems. The mapping and to understand the pattern of change in the forest above ground biomass (AGB) and carbon is an important and challenging task to perform.

In earlier days destructive method was adopted which was also known as the harvest method (Goetz et al., 2009; Malhi et al., 2004). According to Gibbs et al (2007) this method was the most direct method for the estimation of above ground biomass and the carbon stocks stored in the forest ecosystem. This method is used for development of location and species-specific allometric equations used for accessing biomass on large scale. The Allometric method (Chave et al., 2005; FAO, 1997; IPCC, 2006) is based on the principle that every components of trees shows relationship with each other. It is a non-destructive method for the estimation of biomass without felling and thus widely used. In the forest inventories different biometric parameters of trees like diameter at breast height, circumference at breast height (CBH), height, wood density, crown diameter etc. are measured and used to establish an allometric equation by establishing the relationship between these parameters with above ground biomass. In India biomass, carbon stock and carbon budget estimation is done by various workers (Ravindranath et al., 1997; Lal and Singh, 2000; Chhabra et al., 2002) on the basis of growing stock (GS) volume data of forest inventories and appropriate conversion factor related to both biomass and carbon. Remote sensing data are playing an important role in biomass assessment. An approach for the assessment of forest biomass and carbon is boosting day by day using remote sensing technology. As the biomass or carbon cannot be measured directly from remote sensing sensors, it needs the in-situ ground inventory data for establishing a relationship between the biomass and sensor signals (Rosenqvist et al., 2003). The optical remote sensing data was widely used for the mapping and modelling of AGB by establishing a relationship between spectral responses or vegetation indices derived from multispectral image and plot level biomass. The optical remote sensing has limited capacity to predict the accurate biomass because of low saturation level of the spectral bands and the derived spectral indices which results in poor correlation between spectral indices and biomass. The frequent cloud cover in the tropical region hindered the acquisition of high quality data in all weather conditions.

From the last two decades the focus has changed from optical data to SAR data for the assessment. The main advantage of SAR data is its all weather and night availability with longer wavelength and deeper penetration depth, greater sensitivity to biomass and availability of data (Santoro et al., 2006; Santoro et al., 2009; Morel et al., 2011). The SAR data have been used numerously by scientists to estimate the retrieval of biomass using radar data and variations in the forest ecosystem biomass (Sader, 1987; Wu, 1987; Hussain et al., 1991; Dobson et al., 1992; Kasischke, 1992; Le Toan et al., 1992; kasischke et al., 1994a). The longer wavelength of SAR data (L and P bands) proves to be more useful than shorter wavelengths (X and C bands) because of increasing backscatter range. The strength of the relationship depends on the size of the sample plots (Mitchard et al., 2009; Saatchi et al., 2011) and hence should be carefully chosen and laid. The simplest approach of biomass modelling is used in the upper stretches of CTR i.e. the backscattering coefficient derived from the data is correlated with the field inventory data. This approach has been tested throughout different forest types in the world with high degree of correlation between observed and predicted forest biomass (Hussein et al., 1991; Le Toan et al., 1992; Dobson et al., 1992). The
potential of L-band radar backscatter to estimate aboveground biomass (AGB) has been studied for most forest types (Harrell et al., 1995; Imhoff, 1995; Kasischke et al., 1995; Le Toan et al., 1992; Lucas et al., 2010; Pulliainen et al., 1996; Santoro et al., 2006). The saturation of the SAR data is one of the challenge in the biomass modelling. The reported saturation level for the L-band data ranges between 40 t/ha (Luckman et al., 1997; Imhoff, 1995) to 150 t/ha (Kuplich et al., 2005; Lucas et al., 2007; Mitchard et al., 2009). The saturation level for biomass in X and C band is very low (30 t/ha to 50 t/ha). Remote sensing, being an advanced technology is quite useful for reliable estimation of vegetation biomass and carbon over large areas. Furthermore, remote sensing is also useful for stratification of forests and in selection of proper sample plots for enumeration which is otherwise not possible through convention methods. Most of the studies with optical sensors have estimated biomass indirectly because of the several inherent limitations of optical data such as: inability to penetrate the vegetation canopy, insufficient sensitivity to forest structure and above ground biomass, inadequate temporal frequency because of persistent cloud cover etc. It is proposed to evolve methods to improve the assessment of phytomass/Carbon using optical and Microwave remote sensing data and suggest method for improvements in estimates of biomass. Taking the advantage of the deeper penetration of longer wavelength in the forest canopy, an attempt was made to develop empirical relationships between microwave backscatter from satellite and the biomass levels so as to estimate the forest biomass of the study area. The significant empirical relationship was used for the spectral modelling of biomass in the whole study area.

2. Materials and Methods

2.1. The study area

The Corbett Tiger Reserve lies between the latitudes 29° 25' N to 29° 40'N & longitudes 78° 5' E to 79° 5' E. It spreads through 3 districts of Uttarakhand namely Pauri, Nainital, Almora and a small part falls in Amangarh, Bijnore district of eastern Uttar Pradesh. The Ramganga, Palain and Sonnandri River flow through these valleys. The vegetation in CTR is of forests, grasslands and riparian types. Flora diversity of CTR is very rich as the major portion of the reserve is confined to Bhabar tract of Shiwalik formation. There are 617 species of the flora under 410 genera 111 families of Angiosperms (Monocot-132, Dicots-462), 1 Gymnosperm and 22 Fern and fern allies. There are more than 110 tree species in the forest. Notably 73% is constituted by Sal (Shorea robusta) forests. A frequent associate of Sal is Adina cordifolia. The predominant species in the higher ridges is Bakli (Anogiesus latifolia) and the other associates are Bauhinia rausinosa, Lagerstromia parviflora, Cassia fistula, Semecarpus anacardium. Chir (Pinus roxburghii) the only conifer is confined to some of the highest ridges around Sultan. The river valley, high banks and islands are dominated by Delbergia sissoo. Lantana camara is profusely invading in the reserve, inhibiting the growth of other species. Cannabis sativa is also found extensively in the grasslands.

2.2. Satellite data

The Phased Array type L-band Synthetic Aperture Radar (PALSAR) is an active microwave sensor using L-band frequency to achieve cloud-free and day-and-night land observation. ALOS PALSAR Fine Beam dual Polarization (FBD) scene was obtained from the Alaska satellite Facility. The data of July 2008 was downloaded which consists of two bands in HH and HV polarization having spatial resolution of 15.85 m.
2.3. Field data collection

A two stage sampling design is formulated for national forest inventory. In the first stage the country is divided into homogeneous strata based on physiography, climate and vegetation. Samples of 10 percent districts proportion to their sizes are selected randomly for the detailed inventory. For each selected districts, Survey of India (SOI) topo sheets of 1:50,000 scales (size 15’×15’) are divided into 36 grids of 21/2’×21/2’ which is further divided into sub-grids of 11/4’×11/4’ forming the basic sampling frame. Two of these sub-grids are then randomly selected to lay out the sample plots. The sample plot of size 31.62×31.62m was laid and the diameter at breast height (DBH) of all tress having DBH 10cm and above using caliper, double bark thickness using 62 steel scale, height of the tress using Hypsometer and crown width were measured and recorded. The sample plot is divided into sub-sample plots of 5×5m for herbs and 1×1m for shrubs.

2.4. Plot level biomass

The allometric equations developed by Forest Survey of India (1996) were used. The data base of the field data was created in the MS Excel sheet and analyzed in SPSS software. The circumferences at breast height were converted into diameter and basal area was calculated. The volume of each tree within the plot was estimated using aforementioned allometric relationship. The selection of volume equation for a species depends upon the ‘n’ (total number of sample tree on which regression equation are based) and ‘R²’ (Coefficient of determination). The value obtained from the equation was multiplied with wood specific gravity (Forest Research Institute, 1996) to estimate the biomass.

\[ = \times \]

The total biomass of all the trees within the plot were obtained by multiplying the obtained biomass with Biomass Expansion Factor (BEF) and the oven dry weight of shrubs, herbs and litters were all added to get the plot level biomass which was further taken on to pixel level biomass.
2.5. Image processing

The dual polarization ALOS PALSAR data acquired is imbedded with inherent speckle noise which reduces the appearance of data. Multilook operator was used to reduce the inherent speckle and to get a nominal image pixel size. The terrain correction was carried out using 30m resolution SRTM DEM obtained USGS earth explorer in ASF tool of PolSAR-Pro followed by speckle filtering. The data was provided with the digital numbers (DN) which was converted in to backscattering coefficient using the formula:

\[ \sigma = 10 \cdot 10^{\frac{CF}{10}} \]

Where:

\[ CF = -83 \text{ [dB]} \]

The formula was applied on both the HH and HV polarizations to get backscattered images. The backscattering coefficients values were different for both the polarizations. The range boundaries of the Corbett Tiger Reserve were used in Arc-GIS to extract the area falls under the reserve. The sampling was done for the whole region to reduce the uncertainty in the modelling.

2.5.2. Biomass modelling and mapping

The Global Positioning System (GPS) locations of the sample plots were converted into point shape using Arc GIS software. Half of the sample plot information was used for the training of models for the assessment of biomass and the remaining half was used for the validation of the model. The plot information used for the training was overlaid on back scattered image of HH and HV polarizations. The sigma naught (\(\sigma^0\)) backscattering coefficient values were extracted from the plots. The backscattered values for both the polarizations were correlated with the plot level biomass. The best fit model was selected for the modelling of biomass in the Corbett Tiger Reserve.

3. Results and Discussion

3.1. Plot level biomass

The destructive methodology has been preceded by non-destructive methodology for biomass estimation. The plot level basal area was taken as variable for the estimation of biomass. A significant coefficient of correlation (\(R^2=0.94\)) was found between the basal area and biomass in 120 sample plots. The biomass in the region ranges between 10.12 t/ha and 322.61 t/ha. The majority of the area was found to be in the range of 100 t/ha to 150 t/ha biomass. The possible explanation for significant degree of determination might be related to the fact that this region is a protected area in which periodic silvicultural practices has been applied for its management.

3.2. Biomass modelling and estimation

The satellite image was extracted for the ranges lies within Corbett Tiger Reserve. The aim was to estimate the biomass in the tiger reserve, thus the remaining regions were not used in the modelling. The Figure 3 & 4 shown below are the backscatter image of both the polarizations falls within the tiger reserve. The backscattering coefficients for the field plots ranged from -31.52 to 12.75 dB in HV polarization (Figure 3) whereas it ranges from -26.76 to 16.64 dB in HH polarization (Figure 4). The HV backscatter was observed to be less as compared with HH polarization.
The less value of HV polarization is due to its multiple interactions with the forest canopy as compared with HH polarization. The less negative values and higher positive values were discarded before setting up the relationship.

![Figure 3: HV Backscatter image of ALOS PALSAR](image1)
![Figure 4: HH Backscatter image of ALOS PALSAR](image2)

The biomass values were to be predicted (dependent variable) on the basis of backscatter coefficients (independent variable). Thus, simple regression analysis was performed for the analysis. The plot level biomass were plotted on the Y-axis and the backscatter values were plotted on the X-axis to obtain scatter plot.

The best fit regression models were tried on the data sets. The linear best fit regression line between plot level biomass and HV and HH backscatter are shown in Figure 5a & 5b. The coefficient of determination is more with HV backscatter ($R^2=0.75$). It shows that 75% variability in the biomass can be addressed by HV backscatter. It was found low with HH backscatter ($R^2=0.45$). Similarly, the logarithmic best fit regression line was applied with both the polarizations. It was found that the logarithmic regression provides the best regression model with HV polarization having a highest coefficient of determination ($R^2=0.754$) with respect to HH ($R^2=0.48$) polarization (Figure 5c & 5d).

Polynomial equation was considered to be least significant among other equations for the modelling. The figure 5e and 5f represents its regression with HV and HH polarization. The coefficient of determination ($R^2$) was observed to be 0.73 with HV backscatter and 0.48 with HH polarization. It was observed from the graphs that the regression lines were saturating when the range of biomass crosses 150 t/ha. The regression graphs are shown below in Figure 5.
5a. Linear relationship between HV and plot biomass

\[ y = 0.5564x - 19.758 \]
\[ R^2 = 0.7508 \]

5b. Linear relationship between HH and plot biomass

\[ y = 0.5003x - 13.672 \]
\[ R^2 = 0.4482 \]

5c. Logarithmic relationship between HV and plot biomass

\[ y = 4.8561\ln(x) - 24.796 \]
\[ R^2 = 0.7543 \]

5d. Logarithmic relationship between HH and plot Biomass

\[ y = 4.1071\ln(x) - 17.746 \]
\[ R^2 = 0.4814 \]

5e. Polynomial relationship between HV and plot Biomass

\[ y = -0.0153x^2 + 0.8901x - 21.212 \]
\[ R^2 = 0.7386 \]

5f. Polynomial relationship between HH and plot Biomass

\[ y = -0.0319x^2 + 1.1146x - 16.107 \]
\[ R^2 = 0.4862 \]

Figure 5: Regression graphs
Among different curve progression (linear, logarithmic, polynomial) of correlation the optimal equation with high coefficient of determination ($R^2=0.75$) was derived from HV backscatter in logarithmic model when used independently and was found to be:

$$= 4.8561 \ln - 24.796$$

(1)

Where:

- $Y$ = Backscatter coefficient ($\sigma^0 [\text{dB}]$)
- $X$ = Biomass

The results were compared with the studies carried out on the similar principle across the world. It was found that the backscatter from the forests depends on the structural properties (Imhoff, 1995a). It has been demonstrated that there is a strong relationship between backscatter coefficients and above ground biomass within a particular forest types (Le Toan et al., 1992; Dobson et al., 1992; Imhoff, 1995b). A model based approach had been investigated for stem wise forest stem volume retrieval using JERS-1 L-band SAR data in Sweden, Finland and Siberia. In dense forest the backscatter shows a difference of ~4dB, whereas in sparse forests, the backscatter depends on the dielectric properties of the forest floor showing smaller difference throughout the year (Santoro et al., 2006). A similar study had been conducted by Luckman et al., (1997) in the central Amazon basin using ERS-1 & JERS-1 satellite. It was concluded that the longer wavelength (L-Band) is more suitable to discriminate between different levels of forest biomass up to a certain threshold because of its deeper penetration into the vegetation canopy. The cross polarized backscatter is more sensitive to changes in biomass density because of its crown scattering mechanism. In the case with shorter wavelength (C-Band), it has been difficult to differentiate between vegetation and bare soil when it is dry. In the simple regression analysis only one independent variable was used for the prediction of biomass in the region. To improve the accuracy and to know the combined effects of both the polarizations, multiple regression analysis was applied. The multi-linear regression analysis has been done using plot level biomass as dependent variable and HH & HV polarizations as independent variables. There was a significant improvement in correlation coefficients ($R^2=0.86$). The equation developed using multi-linear regression analysis was found to be:

$$= 1.364 - 0.098 + 28.20$$

(2)

This equation 2 obtained from independent HV backscatter and biomass using logarithmic model and the equation 3 obtained from MLR analysis were used for above ground biomass mapping in CTR. Both the models were run independently to predict the biomass from plot level to the whole study area. The modelled biomass using equation 2 varies from 26.2 t/ha to 401.43 t/ha. The modelled biomass using equation 3 varies from 10.96 t/ha to 312.64 t/ha. The predicted biomass range using equation 3 was very close with the field data because of the combined potential of both the polarization. A simple approach to evaluate the model is to regress predicted verses observed values. Thus, the biomass maps obtained through the modelling using both the equations 2 & 3 were plotted against the remaining in-situ plot biomass (60 sample plots) left for the validation of the model. The observed plot biomass was represented on X-axis and the predicted biomass was represented on Y-axis. A significant coefficient of determination ($R^2=0.734$) is obtained between observed biomass and predicted biomass in case of values predicted by single regression analysis.
using HV backscatter (fig 6). Whereas, a strong coefficient of determination has been observed with Multiple Linear Regression based biomass map and *in-situ* data ($R^2=0.83$) (Figure 7). The total of 83% variability can be addressed by observed value to explain predicted values using MLR based equations whereas only 73.4% variability can be addressed using simple regression model. The result explains that both the parameters are essential for the estimation of biomass using microwave data. The figure 8 represents the biomass distribution map in the ranges of Corbett Tiger Reserve. The map clearly shows the majority of the area is dominated with the biomass range between 100 t/ha to 150 t/ha. It was observed during the field visit that few regions which were showing high biomass region i.e. more than 200 t/ha was actually not existing. The reason for this overestimation was due to the topographic distortion of layover and foreshortening in the SAR image. This topographic effect can be reduced by using different incidence angle SAR images in both the ascending and descending mode. The accuracy of the model can further be improved by utilizing the capabilities of fully polarimetric data. The semi-emperical models such as water cloud model (WCM), Extended water cloud model (EWCM) will have the capabilities to improve the biomass estimation in the region. The launch of NISAR mission in 2020 is going to be a great opportunity for mapping the biomass and carbon stock all across the country with a reliable and accurate means.

Figure 6: Observed vs Predicted

Figure 7: Observed vs Predicted Biomass using MLR.
4. Conclusions

The potential of L-band ALOS PALSAR dual polarization data has been investigated for the biomass retrieval in the CTR. The L-band ALOS PALSAR data is proved to be sensitive to the differences in the biomass and thus very useful in the AGB mapping in CTR. The simple regression analysis and multi-linear regression have been tried for biomass mapping in CTR. The best correlated model is derived from the relationship between backscatter coefficient of ALOS PALSAR and plot level biomass. Among the two polarizations HH & HV, HV found to be strongly correlated with plot biomass. The field inventory data shows that the biomass in the CTR ranges from 9.6 t/ha to 322.6 t/ha. The modelled biomass represents that the majority of the area is dominated with the biomass ranges from 100 t/ha to 150 t/ha. The total biomass in the upper stretches of CTR covering an area of 889 sq. km is found to be 8.9 million tonnes. A significant coefficient of determination is observed between the observed and predicted biomass on 95% confidence interval modelled using MLR. The MLR proves to be more effective for modelling as compared with SLR modelling. It was observed that the sample plot should be large enough to obtain a better relationship between backscattering coefficients and AGB as the area is rich in biodiversity and have complex structure of vegetation. The intensity of backscatter also gets saturated around 100 Mg/ha. Some uncertainties have been observed through the area having steep slopes. The layover and foreshortening were the major causes for the uncertainty. The overall accuracy of the result can further be enhanced using Quad polarization data and through the technique of Polarimetric decomposition.

Acknowledgements

The work in this paper was carried out with the support of ‘Forest Inventory Unit’ Forest Survey of India, Dehradun. The author wishes to thank Director General (DG) and Assistant Director General (ADG) Forest Survey of India for valuable guidance and support.

References


