# ESTIMATION OF FOREST BIOMASS BASED ON MULITI-SOURCE REMOTE SENSING DATA SET
# -A CASE STUDY OF SHANGRI-LA COUNTY

Feng Wanwan[1, 2], Wang Leiguang[2], Xie Junfeng[3], Yue Cairong[1], Zheng Yalan[1], Yu Longhua[1]

[1] Southwest Forestry University, Faculty of Forestry, Kunming 650224, China
[2] Southwest Forestry University, Big data and Artificial Intelligence Research Institute, Kunming 650224, China
[3] Satellite Surveying and Mapping Application Centre, Beijing 100048, China

**Commission VI, WG VI/4**

**KEY WORDS:** Texture, Regression model, Shangri-La county, Multi-source data, Forest biomass estimation

**ABSTRACT:**

Forest biomass is an important indicator for the structure and function of forest ecosystems, and an accurate assessment of forest biomass is crucial for understanding ecosystem changes. Remote sensing has been widely used for inversion of biomass. However, in mature or over-mature forest areas, spectral saturation is prone to occur. Based on existing research, this paper synthesizes domestic high resolution satellites, ZY3-01 satellites, and GLAS14-level data from space-borne Lidar system, and other data set. Extracting texture and elevation features respectively, for the inversion of forest biomass. This experiment takes Shangri-La as the research area. Firstly, the biomass in the laser spot was calculated based on GLAS data and other auxiliary data, DEM, the second type inventory of forest resources data and the Shangri-La vector boundary data. Then, the regression model was established, that is, the relationship between the texture factors of ZY3-01 and biomass in the laser spot. Finally, by using this model and the forest distribution map in Shangri-La, the biomass of the whole area is obtained, which is $1.3972 \times 10^8$t.

## 1. INTRODUCTION

Forest biomass is the result of long-term production and metabolism of forest ecosystems. Advanced technologies, such as remote sensing, can measure forest biomass from stand to area scales. The Geoscience Laser Altimeter System (GLAS) waveform data have been used widely in the estimation of forest ecological parameters in recent years (Lefsky et al.2005; Li et al.2017; Yu et al.2010; Soenen et al.2010).Since the GLAS laser footprint is 70 meters in diameter and 150 meters apart each other, limiting biomass estimates at the regional scale. In order to estimate the biomass of forest land from a regional scale, other optical remote sensing data sets need to be combined. Wang et al. (2012) combine GLAS and MODIS data to estimate the forest vegetation height and establish a regression model with the measured biomass, and then estimate the forest biomass of Yunnan Province by the model. Pang et al. (2011) use airborne LiDAR data to estimate the aboveground biomass of the Greater Mekong sub-region (GMS). Xing (2016) takes the forest district of Wangqing County, Jilin Province, as the research area, GLAS data and MODIS remote sensing image data as the data source, and extracts the vertical forest structure parameters and the forest horizontal structure parameters respectively. By combining these parameters, establishes the forest above ground biomass inversion model, which improves the estimation accuracy and realized the continuous mapping of the study area.
Spectral features extracted from optical remote sensing sensors are widely used in the process of modeling, but suffer from the cloud cover and signal saturation, especially in the forest areas. Texture features naturally become dominant features, that reveal the horizontal structure information of the image content and reflects the spatial variation and the spatial correlation characteristics of the image gray values. Meng et al. (2017) build the relationship with forest aboveground biomass (AGB) using vegetation indices and textural indices based on high spatial resolution Worldview-2 multispectral images, confirm that textural information can improve certain accuracy in forest AGB estimation.

The optical data characterizing the horizontal structure of the canopy are fused with the GLAS data characterizing the vertical structure of the stand, meanwhile, with other auxiliary data set, the improvement of the accuracy of the biomass estimation can be expected. So based on previous study, we combine ZY-3 satellite, Aster DEM, and the second type inventory of forest resources to establish a regression model, thus achieving the seamless retrieval of biomass in Shangri-La Forest Region.

## 2. STUDY AREA

Shangri-La county, under the jurisdiction of Diqing Tibetan autonomous prefecture, is located in longitude 99 °20 '~ 100 ° 19', latitude 26 °52 '~ 28 °52', surrounded by the Jinsha River, The terrain is tilted from northwest to southeast. This area belongs to Mountain cold temperate monsoon climate.
Shangri-La's forest resources are extremely rich. The common forest types are *Quercus semecarpifolia Smith, Pinus yunnanensis, Pinus densata, Abies georgei Orr and Spruce.* The city has a forest area of 757 105. 2 hm$^2$, forest coverage rate reached 74. 99 % (Song et al.2007).

## 3. METHODS AND RESULTS

Based on the existing research, this paper uses the spatial information of remote sensing data set, including elevation and texture features, and other supporting data to achieve the estimation of forest biomass in Shangri-La. Finally, compares with other research results. The flow chart is shown in Figure 3.

### 3.1 Data processing

The overall process is shown in Figure 3. The first step is the pre-processing of various remote sensing data, followed by feature extraction and selection, establishing a regression relationship between independent variables, that is feature set, and biomass, and finally combining the forest distribution map to calculate the biomass of the entire region.
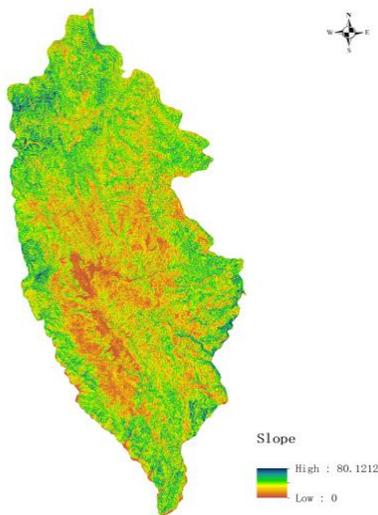


Figure 1. GLAS laser footprint position



Figure 2. Shangri-La slope map

**Step1: Selecting the 2008 GLAS 14 laser data to calculate the average height of all spots in the Shangri-La district.**
The GLAS 14 data is global land surface altimetry data, which can be derived from GLAS 05 and GLAS 06 and records surface height information. This data contains latitude and longitude, elevation and other information. In this study, the average height of the light spot in Shangri-La was extracted by using the GLAS 14 data. The average height within the spot is shown in Figure 4.

**Step2: Removing all laser spots with slope greater than 50 degrees.**
When the laser radar spot projects on the ground, the GLAS14 extracted from the ground is flat and the confidence of the tree height information is high. When the ground has a certain gradient, it will broaden the waveform and the tree height information is inaccurate.
Therefore, according to the latitude and longitude position of the spot, combined with the slope information, amended the tree height information, and the light spot with a slope of less than 50 degrees is retained. Slope map shown in Figure 2.

**Step3: Calculating the volume in the laser spot.**
Wang et al.(2015) selected Shangri-La as the research area, combined with measured data, and established a linear relationship between the average tree height within the laser spot of the Geoscience Laser Altimeter System (GLAS) data and the forest volume ($r^2$=0.7156).
The linear regression relationship between GLAS average canopy height and field measured stock volume is:

$$Y = 0.009 \times X + 18.58 \tag{1}$$

where   X: GLAS average height
        Y: Stock volume

The volume in the entire effective laser spot was calculated by this model.

**Step4: Choosing different dominant tree species of pure laser spot.**
Using the second type inventory of forest resources data to classify the light spots according to the dominant tree species, the light spot of the pure forest area was extracted. Part of the result is shown in Figure 4.

**Step5: The volume is converted to biomass.**
Based on the conversion relationship between the stock volume and the biomass proposed by Huang (2008) (see Table 1), the biomass of different light spot groups was calculated. Spotted biomass was used as the dependent variable for the second model of this research.
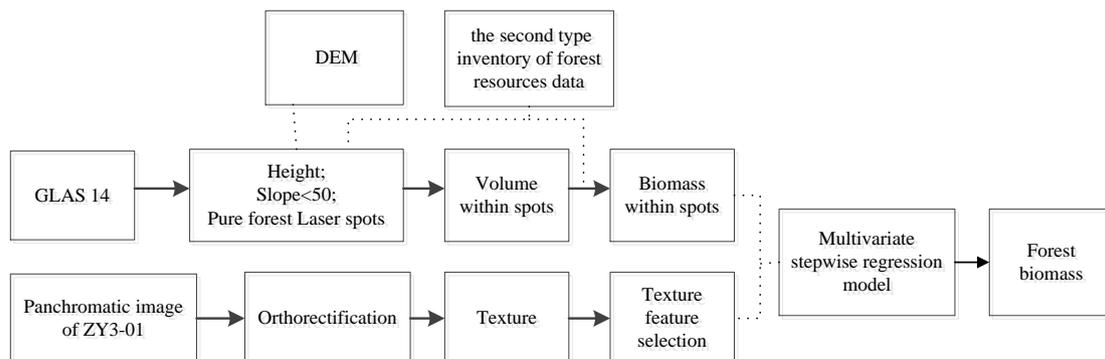


Figure 3. Flow chart

| Forest structure | Biomass (w) - Volume (v) Regression equation |
|---|---|
| *Picea asperata Mast.*<br>*Abies fabri (Mast.)Craib* | $w = 2.137 \times v^{0.7532}$ |
| *Pinus yunnanensis* | $w = 0.8569 \times v^{0.8564}$ |
| *Pinus armandii Franch and other pine species* | $w = 0.5272 \times v^{1.0793}$ |
| *Quercus acutissima and other Hardwood* | $w = 0.6573 \times v^{1.0502}$ |
| *Softwood* | $w = 2.1529 \times v^{0.6085}$ |
| *Platycladus orientalis(L.)Francoptmxjjkmsc* | $w = 1.2237 \times v^{0.8592}$ |
| *Eucalyptus robusta Smith* | $w = 2.5878 \times v^{0.6261}$ |
| *Phoebe zhennan S. Lee et F. N. Wei* | $w = 1.0726 \times v^{1.0128}$ |

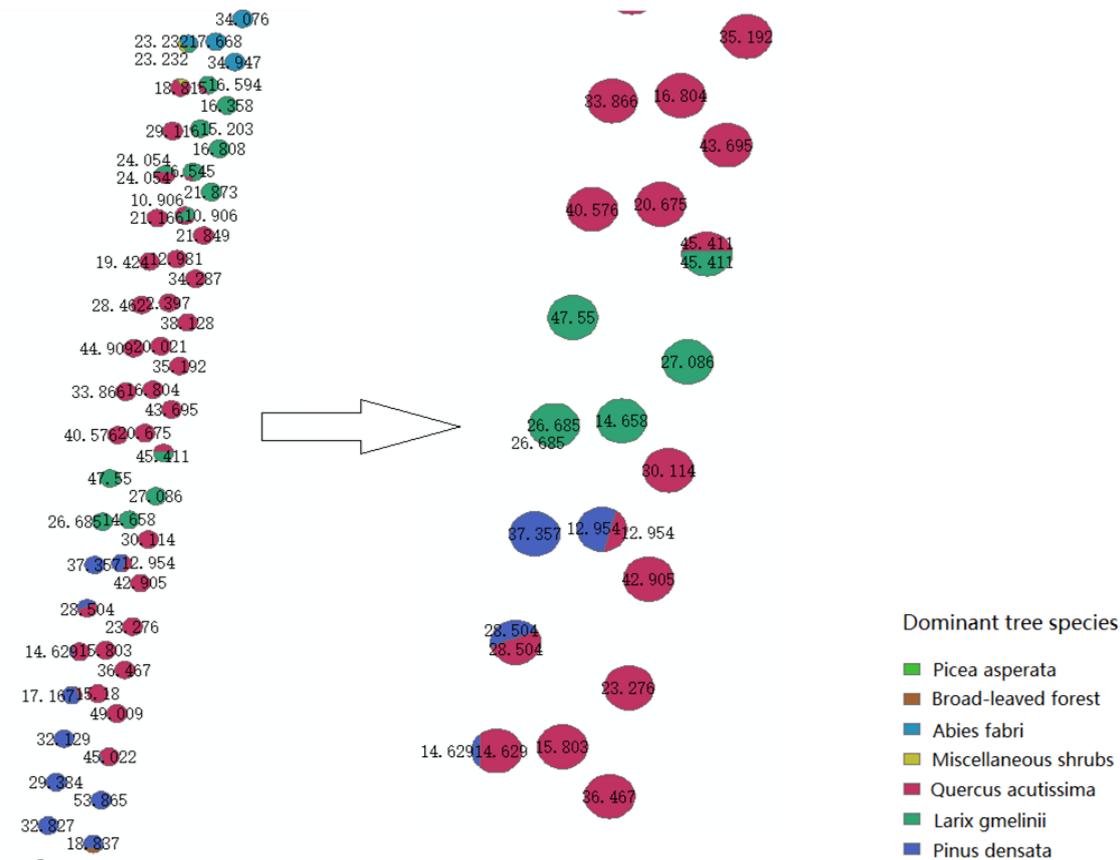Table 1. Relationship between biomass and stock volume



Figure 4. Laser spot classification and average height

**Step6: Extraction of model factors.**

Texture information extracted according to different windows and different directions setting based on high spatial resolution ZY-3 Panchromatic image, shown in Figure 5, was utilized to build the relationship with forest aboveground biomass as well as assess the accuracy of estimation model. We applied Gray Level Co-occurrence Matrix (GLCM) algorithm for extracting eight texture features from ZY-3 panchromatic image, including, mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation. The window sizes are set as $3 \times 3$ to $31 \times 31$, and the horizontal and 45 degrees directions are selected, In this way, a total of 240 texture features were obtained.

In order to reduce redundancy and validity of information, Pearson's correlation coefficient was used to measure the correlation with the biomass in the light spot respectively. Finally, 83 significant correlations were obtained.

**3.2 The establishment of regression models**

We select multivariate stepwise regression model for biomass inversion. All factors are normalized before building model. The result of multivariate stepwise regression model is:

$$Y = -0.028 \times X_1 - 0.021 \times X_2 + 0.141 \quad (2)$$

where

Y: Laser spot biomass
$X_1$: $25 \times 25$ mean of horizontal direction
$X_2$: $13 \times 13$ variance of horizontal direction.

It confirmed that Texture and biomass were negatively correlated.

Although 83 texture factors have been initially filtered, the overfitting problem might still occur due to the large number of input variables and results in the decrement of the robustness of regression model.

Feature weighting algorithm, proposed by Kira (1992), is a computational feature weighting algorithm. But it can process only two categories of data .In 1994, Kononeneill improve the algorithm, called RelieF, which can achieve multiple classification and regression problems. The main idea of the algorithm is to assign adaptive weight to every feature according to the relationship between the feature and classes, and then select the feature subset according to the weight value. RelieF algorithm is capable of avoiding overfitting in the training process, and hence it was used to perform feature selection.

In this experiment, for the texture feature set, its weight was calculated using RelieF algorithm. The feature weights are shown in Figure 6. Based on weight values, the top 10 features were selected for modeling.
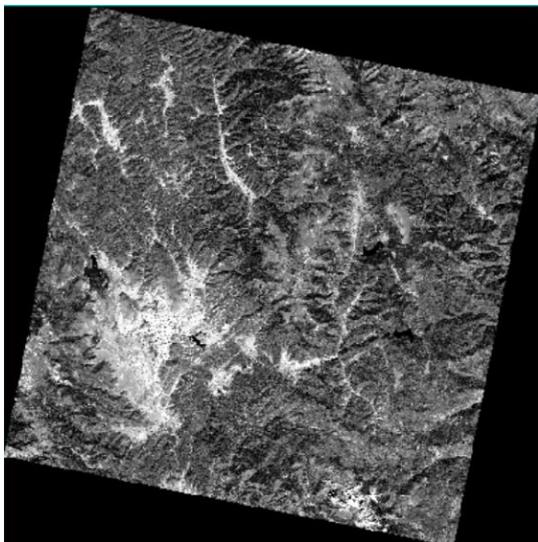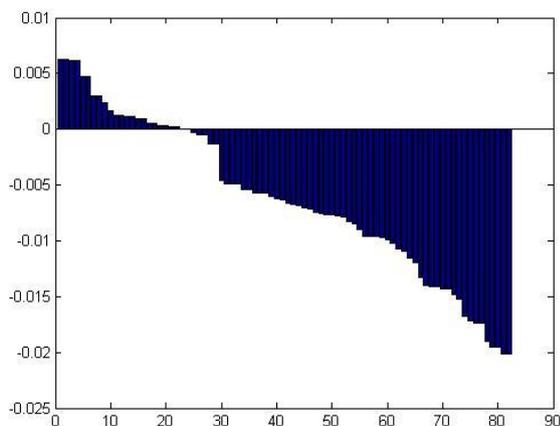


Figure 5. Orthophoto map of ZY3-01



Figure 6. Texture feature weight graph

### 3.3 Seamless estimation of biomass

GLAS is carried on the Ice, Cloud and land Elevation Satellite (ICE-Sat), whose laser pulses not only reach the top of the crown, but also reach the ground and provide complete waveform data, widely used in forest structure parameters and forest biomass estimation. However, Laser footprint is discretely distributed, so limiting biomass estimates at the regional scale. In order to estimate the biomass of forest land on a regional scale, other optical remote sensing data needs to be combined.

The panchromatic image of ZY3-01, with a resolution of 2.1 meters and a width of 51 kilometers, requires a total of 15 scenes throughout Shangri-La County. We selected panchromatic images with a cloudiness of less than 2 for orthorectification. After the mosaic procedure, calculated the mean and variance. Using formula 2 in combination with the forest type map, achieved the inversion of forest biomass in the entire region, which is $1.3972 \times 10^8$t.

### 3.4 Results comparison

Yue (2011) uses the support vector machine regression model to estimate the biomass of 4 dominant species in Shangri-La County. The results showed that the biomass of 4 dominant species totaled $1.19 \times 10^8$t. Wang (2013) establish models of regional forest biomass estimations for 4 dominant species in Shangri-La forest. Result reveals a total biomass of $1.14 \times 10^8$t.

### 4. CONCLUSIONS

This research explored the capability of the biomass within the light spot and textural information for AGB assessment in Shangri-La City, Yunnan Province. Then the optical image is used to estimate the whole area's forest biomass. We got the following conclusions:

(1) The total forest biomass of Shangri-La county is $1.3972 \times 108$t. The results are in the same order of magnitude as the other two studies, which mirrors that our method is effective.

(2) The horizontal direction texture features, mean with $25 \times 25$ moving window sizes and variance with $13 \times 13$ moving window sizes are negatively correlated with the forest biomass in Shangri-La County.

Tree height is an important parameter for quantitative estimation of aboveground biomass and is also necessary for obtaining primary forest-related and biodiversity information. The accurate forest tree height value can help to establish a variety of ecological models such as carbon cycle. In this study, by extracting the tree height information of GLAS14 data and using some of the existing studies, the biomass in the light spot was estimated. Then the correlation between ZY3-01 texture factor and biomass was analysed .The regression model was established. At last, based on the forest classification map contributed by the second type inventory of forest resources data, the biomass of Shangri-La was calculated.

### 5. DISCUSSIONS

The remote sensing techniques provide an efficient way to estimate ecological index. Many domestic and foreign studies often use medium or low resolution images to estimate biomass, such as Landsat series satellites (Lu, 2016) and MODIS. However, in mature forest area, the problem of spectral saturation easily appears. Wang etc. (2012) conduct an analysis of the uncertainty of forest biomass. The results show that the Landsat TM data are not suitable for mature forest biomass estimation. The main sources of uncertainty are the change of input sample data and the data saturation problem related to optical sensors. Lidar overcomes the shortcomings of TM and

provides better biomass estimation performance. The texture information of high resolution remote sensing image can make up for this problem to some extent. For the purpose of this study, for the ZY3-01 data, extract the panchromatic image texture factor for biomass estimation, but there are also the following problems:

The area of the ZY3-01 panchromatic image is 2500 km$^2$ for a single scene, the revisit period is 5 days, and the whole Shangri-La contains 15 panoramic images. In quantitative remote sensing, the quality of the image, such as the cloudiness, is crucial to the result. In this study, the cloudiness of the selected images is less than 5%, so that the screening time cannot achieve the same image acquisition time. Although the selected images are of the same timing, different years have a certain impact on the estimation of actual biomass.

In addition, GLAS data acquirement is ended in 2009, laser energy was depleted in 2009, and data quality was low. The GLAS 14 data in this experiment was in 2008 and ZY3-01 data were after 2012. This, in turn, is a major factor in the estimation of images.

In summary, for a large range of biomass estimation, whether it is homologous data, or between different sources data, the problem of time mismatch cannot be avoided. In the later stage, only by shortening the observation period of the earth observation on the remote sensing, increasing the observation range to the earth, expanding the earth observation means and improving the research technique can we improve the accuracy of large area estimation and better serve the research of global change.

## ACKNOWLEDGEMENTS

## REFERENCES

Cao Chunxiang, Ni Xiliang, Chen Wei, He Qisheng, 2015. Remote sensing diagnosis of Forest Aboveground Biomass. Science Press, first edition.

Huang Congde, 2008. Carbon Storage in Forest Ecosystems and Its Spatial differentiation characteristics. Sichuan Agricultural University.

Huang Kebiao, Pang Yong, 2011. Aboveground biomass analysis of forest in the Greater Mekong sub region based on remote sensing. Resource Science. 33(10):1863-1869.

Li Fu, Zhao Dan, 2017. Variations in forest aboveground biomass in Miyun Reservoir of Beijing over the past two decades.Journal of Soils and Sediments, (6):1-11.

Li Zengyuan, Pang Yong, Liu Qingwang, 2015. Inversion technique and method of Lidar forest parameters. Science Press, November .First edition.

Lu Chi，Zhang Jialong，Wang Aiyun,et al. 2016. Rebuilding the model on the biomass estimation of Pinus densata in Shangri-la City based on Landsat TM in Shangri-la City Based on Landsat TM. Forest Inventory and Planning. 41(6).

Lu Dengsheng, Chen Qi, Wang Guangxing, et al., 2012. Aboveground forest biomass estimation with Landsat and LiDAR Data and uncertainty analysis of the Estimates.

Hindawi Publishing Corporation.International Journal of Forestry Research. Article ID 436537, 16 pages.

Kira K, Rendell LA, 1992. The feature selection problem: Traditional methods and a new algorithm. AAAI; (2): 129-134.

Michael A. Lefsky, David J. Harding, 2005. Estimates of forest canopy height and aboveground biomass using ICESat. GEOPHYSICAL RESEARCH LETTERS, VOL. 32, L22S02.

Meng Shili, Pang Yong etc., 2017. Estimation of aboveground biomass in a temperate forest using texture information from high spatial resolution Worldview-2 data. Journal of Remote Sensing, (5):812-824.

S A Soenen , D R Peddle, 2010. Estimating aboveground forest biomass from canopy reflectance model inversion in mountainous terrain. Remote Sensing of Environment. 114(7):1325-1337.

Song Farong, Zhang Zhanzhong, 2007. Investigation Report on Planning and Design of Forest Resources in Shangri - La County, Yunnan Province. Yunnan Provincial Forestry Investigation and Planning Institute of Dali Branch.

Wang Cheng, 2015. Data Processing and Application of Spaceborne Lidar. Science Press.

Wang Jinliang,Cheng Feng,Wang Cheng,et al., 2012. Primary Discussion on the Potential of Forest Volume Estimating Using ICESat-GLAS Data in Complex Terrain Area-A Case Study of Shangri-la,Yunnan Province.Remote Sensing Technology and Application. 27(1):45-50.

Wang Jinliang, Cheng Pengfei, Xu Shen et al., 2013. Forest biomass estimation in Shangri-La based on the remote sensing, Journal of Zhejiang A&F University. 30(3):325-329.

Xing Yanqiu et al., 2016. Methods and applications of satellite lidar inversion of forest biomass . Science Press. First edition.

Yifan Yu, Sassan Saatchi, 2010. Regional distribution of forest height and biomass from multi-sensor data fusion. Journal of geophysical reaearch,vol.115, G00E12.

Yue Cairong, 2011. Forest biomass estimation in Shangri-La County based on remote sensing.