Front. For. China 2008, 3(1): 10–15
DOI 10.1007/s11461-008-0013-z

RESEARCH ARTICLE

Mingshi LI, Ying TAN, Jie PAN, Shikui PENG

Modeling forest aboveground biomass by combining spectrum, textures and topographic features

Abstract Many textural measures have been developed and used for improving land cover classification accuracy, but they rarely examined the role of textures in improving the performance of forest aboveground biomass estimations. The relationship between texture and biomass is poorly understood. In this paper, SPOT5 HRG datasets were ortho-rectified and atmospherically calibrated. Then the transform of spectral features is introduced, and the extraction of textural measures based on the Gray Level Co-occurrence Matrix is also implemented in accordance with four different directions (0°, 45°, 90° and 135°) and various moving window sizes, ranging from 3×3 to 51×51. Thus, a variety of textures were generated. Combined with derived topographic features, the forest aboveground biomass estimation models for five predominant forest types in the scenic spot of the Mausoleum of Sun Yat-Sen, Nanjing, are identified and constructed, and the estimation accuracies exhibited by these models are also validated and evaluated respectively. The results indicate that: 1) Most textures are weakly correlated with forest biomass, but minority textural measures such as ME, CR and VA play a significantly effective and critical role in estimating forest biomass; 2) The textures of coniferous forest appear preferable to those of broad-leaved forest and mixed forest in representing the spatial configurations of forests; and 3) Among the topographic features including slope, aspect and elevation, aspect has the lowest correlation with the biomass of a forest in this study.

Keywords SPOT5 HRG, textural measures, topographic features, biomass, modeling

1 Introduction

The Above Ground Biomass (AGB) of forest controls the latent carbon emissions from logging or forest fires. Accurate extraction and change monitoring of forest biomass play an extremely vital role in deeply understanding the significance of local, regional even global climate changes, and environmental and ecological degradation. Conventional methods of acquiring these information on forest AGB have been the empirical estimation by means of models based on extensive and intensive measurements of forest structural parameters, including height, diameter at breast height (DBH) and crown closure etc (Fang et al., 1996). Obviously, this way of mapping forest biomass over time and space at the regional and global scales will be time-consuming, costly and labor-intensive. With regard to forest AGB extraction, modern remote sensing techniques have a positive advantage, such as improved estimation accuracies over conventional methods mentioned above. Nevertheless, due to the heterogeneities of forest stand structures and spectral saturation of remote sensing detection, the establishment of estimation models of forest biomass relying solely on spectral information has the deficiencies of low accuracies and weak adaptabilities (Zhao and Li, 2001). This paper explores the establishment and validation of biomass estimation models for broadleaved forests and coniferous forests by combining textural features derived from remote sensing imagery, spectral features and topographical features. Topographical features include elevation, slope and aspect, which control the spatial distribution of forest biomass. Ultimately, the impacts of different textural measures on the spatial variability of different forest types are intensively analyzed as well.

2 Study site

The Purple Mountains Scenic Spot (32°02′09″–32°05′48″N, 118°48′24″–118°52′59″E) is taken as our study. It is the so-called "green lung" of Nanjing City

Translated from *Remote Sensing Information*, 2006, 6: 6–9 [译自: 遥感信息]

Mingshi LI (✉), Ying TAN, Jie PAN, Shikui PENG
College of Forest Resources and Environment, Nanjing Forestry University, Nanjing 210037, China
E-mail: nfulms@yahoo.com.cn

and it is also the typical prototype of a modern urban forest, with an extremely high ecological status. The Purple Mountains are located at the eastern suburb of Nanjing City. Its total area is about 3010 hm^2. Its peak is about 448.8 m. The annual mean precipitation is about 900–1000 mm and the annual mean temperature is around 15.7°C. In our study site, Masson pine, black pine, exotic pine, Chinese sweet gum, oak, bamboo etc. are distributed. Forest coverage reaches 68%.

3 Data and methods

The data captured and used in this study include the topographical map covering this area, with a scale of 1:10000, SPOT HRG imagery, acquired on November 9, 2002, containing four multi-spectral bands (10 m spatial resolution) and a panchromatic band (2.5 m spatial resolution), and the fruits of forest resources inventories, which were conducted in 2002 to generate a forest distribution map and relevant databases.

The used methodology in this paper can be outlined in Fig. 1.

3.1 Vectorization, digital elevation model (DEM) generation and extraction of topographic features

The contours, discrete elevation control points and forest distribution information are manually vectorized respectively after digital topographic map and forest distribution map are obtained. All spatial data are georeferenced to Lon/Lat WGS 84. The forest distribution map is vectorized as a polygon layer and the central coordinates of each stand subcompartment are automatically exported by means of a geographic information system (GIS). Linear interpolation technique is employed to generate DEM on the basis of vectorized contours and discrete elevation control points. Based on generated DEM, a spatial modeler is implemented to produce slope and aspect images. All products produced from the above procedures are prerequisite to the subsequent remotely sensed imagery processing and biomass modeling.

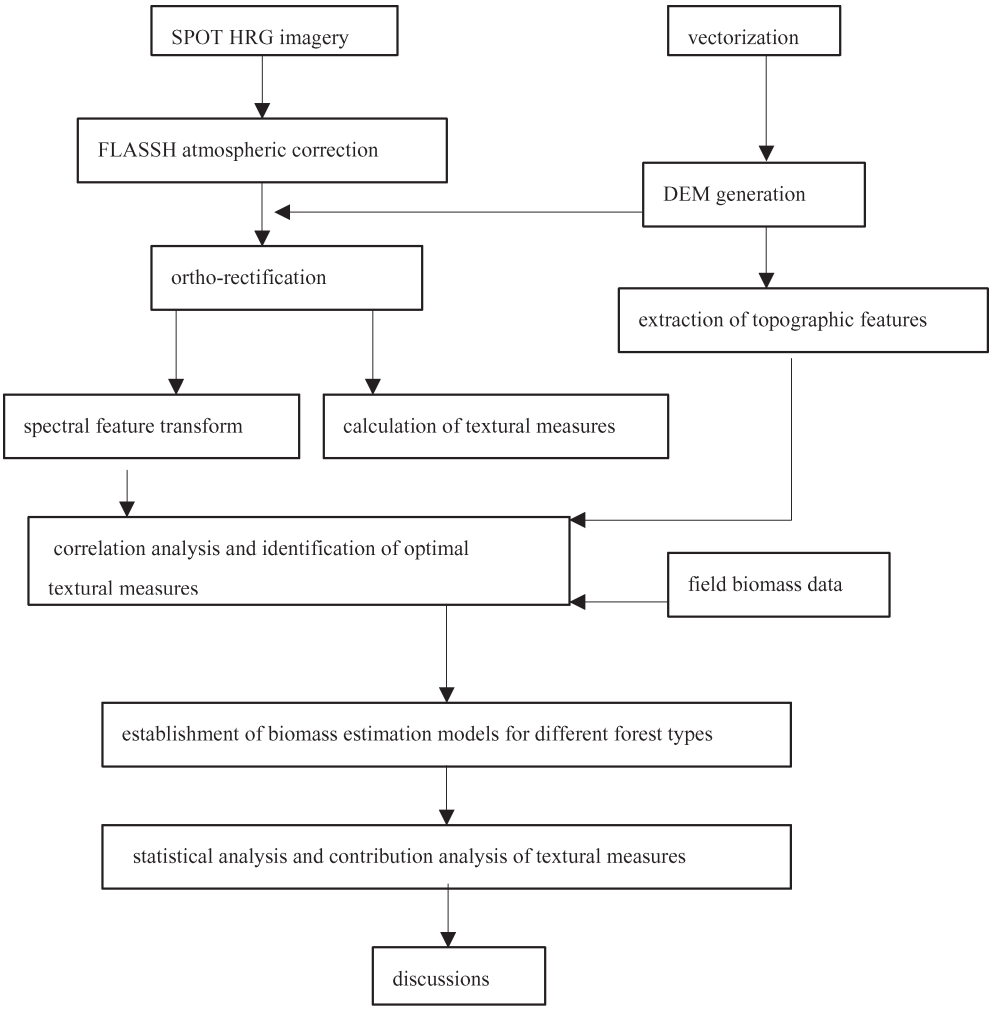

Fig. 1 Study framework of forest aboveground biomass modeling

3.2 Remote sensing (RS) data processing

3.2.1 SPOT HRG imagery pre-processing

Atmospheric correction of SPOT HRG imagery is carried out by using fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) module, which has been embedded into the ENVI4.2 package. Prior to doing this correction, the conversion from the digital number (DN) value of each pixel to radiance on the top of the atmosphere needs to be completed by means of gains and offsets for different spectral bands, then the converted radiance images need to be organized from BSQ structure to BIP or BIL structure. Conversion from radiance to surface reflectance is conducted by specifying the following parameters: the longitude and latitude of the image center (gained from the header file), sensor type, elevation of sensor, average altitude of study area, pixel size, atmospheric model (mid-latitude, summer), aerosol model (urban), visibility (20 km) etc. By means of generated DEM, ortho-rectification of SPOT HRG imagery is followed and all 4 multi-spectral bands are georeferenced and resampled to match 2.5 m spatial resolution panchromatic band. In this georeferencing process, a geometric error of 0.46 pixel is identified.

3.2.2 Spectral feature transform of SPOT HRG imagery

After atmospheric correction and ortho-rectification of the original five bands, principal components analysis (PCA) is followed to generate three principal components, *PC*1, *PC*2 and *PC*3. The contribution of three principal components reached 97.3% of the original spectral information. Among them, *PC*1 accounts for 75.8% variance contribution. In addition, normalized difference vegetation (*NDVI*) and ratio vegetation index (*RVI*) are derived from different arithmetic combinations of near-infrared band and red band. So far, ten features lay the spectral foundations of following biomass modeling.

3.2.3 Generation of textures

Generally, textures of remotely sensed imagery are understood as having spatial variability and periodic occurrence, or as a repeated local pattern (textural element) and its alignment rules (Haralick, 1979; Chica and Abarca, 2000; Sakari and Anssi, 2005). Texture is an important feature of RS imagery. It primarily reflects the roughness of surfaces, reveals the configuration information of different objects in an image and their relationships with surroundings, or refracts the spatial fundamentals of land cover or land use types (Wang, 1994; Xue and Wang, 1997; Huang et al., 2003). Statistically, texture is a unity of local variabilities and spatial correlations. At present, methods of extracting texture are mainly categorized into three classes: statistical methods, configurative methods and spectrum decomposition methods. Statistical methods are employed in this study to primarily depict the stochastic distribution and spatially statistical characteristics of local patterns or textural elements. For statistical methods, the configuration of textural features is measured by spatial frequencies or by densities of local patterns. Gray Level Co-occurrence Matrix (GLCM) is a typical representative of statistical methods for textural extraction (He and Li, 1990; Marceau and Howarth, 1990). Extracting textures by using GLCM method needs to specify two parameters, one is the lag distance between pixels and detecting directions, and the other is the moving window size. Depending upon nine textural formulas listed in Table 1 and PC1 image, lag distance of one pixel (doing so complies with the rationale of spatial correlation) and four detecting directions, 0°, 45°, 90° and 135° are specified to calculate textural features. During this calculation, moving window size ranges from 3 × 3 to 51 × 51, to probe the effects of various window sizes on the extraction of textural measures.

Table 1 List of textural measures

number	textural measure	formula
1	Mean (*ME*)	$ME = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}$
2	Variance (*VA*)	$VA = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}(i-ME)^2$
3	Homogeneity (*HO*)	$HO = \sum_{i,j=0}^{N-1} i \frac{P_{i,j}}{1+(i-j)^2}$
4	Contrast (*CO*)	$CO = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}(i-j)^2$
5	Dissimilarity (*DI*)	$DI = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}\lvert i-j\rvert$
6	Entropy (*EN*)	$EN = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}(-\ln P_{i,j})$
7	Second Moment (*SM*)	$SM = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}^2$
8	Correlation (*CR*)	$CR = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}\left[\frac{(i-ME)(j-ME)}{\sqrt{VA_i \cdot VA_j}}\right]$
9	Skewness (*SK*)	$SK = \frac{\lvert\sum (P_{i,j}-ME)^3\rvert}{(N^2-1)(VA)^{3/2}}$

Notes: $P_{i,j} = V_{i,j} \Big/ \sum_{i,j=0}^{N-1} V_{i,j}$, V_{ij} refers to the DN value of pixel located at i row and j column; N is the size of moving window when calculating textural features.

3.3 Biomass modeling of forest types

Referring to the empirical models for conversion from stock volume to aboveground biomass (Fang et al., 1996), the volume of each subcompartment is converted into aboveground biomass according to the predominant species or species group. In total, there are 367 forest stand subcompartment conversions being accomplished. In accordance with the central coordinates of each subcompartment, a data set consisting of spectral, textural and topographic features and field biomass data for each subcompartment is automatically extracted and organized for biomass modeling.

3.3.1 Selection of textural features for modeling

The features, available for modeling, include 225 textural measures (nine textural measures by 15 moving window sizes), ten spectral features (transformed features plus original bands) and three topographic features, including elevation, slope and aspect. As for 225 textural features, we need to identify those textural measures and corresponding moving window sizes, which closely correlate with biomasses of different forest types. Correlation analysis is used to identify them and the analytical results are summarized in Table 2.

3.3.2 Construction and validation of models

The data of five predominant species or species groups are standardized by standard deviation for subsequent modeling. For data of each species or species group, 80% of the total subcompartments is randomly picked to simulate the biomass estimation model and the remaining 20% is used to validate the established model. Eliminating some features based on collinearity diagnoses (Spss tutorial), those regression equations, which have the highest R^2_a (the adjusted correlation coefficient), are regarded as the optimal regression equations for biomass estimation (Fan and Mei, 2002). R^2_a is defined as follows:

$$R_a^2 = 1 - \left(\frac{n-1}{n-p}\right)\frac{SSE_P}{SST} = 1 - \frac{MSE_P}{\left(\frac{SST}{n-1}\right)} \quad (1)$$

where P stands for the number of dependents in the equation; n denotes the amount of samples when fitting models; SST denotes the total variance of independents; and SSE_P stands for the variance of errors.

4 Results and discussion

The results of correlation analysis in Table 2 show the linear linkages between biomass and various features. For Masson pine, its biomass has relatively strong interactions with XS3, XS4, PC1, PC2, elevation and texture mean (51). However, using these features as dependents does not imply that the biomass estimation model will have a high estimation accuracy, because probably existing interactions among these features will reduce the capability of these dependents to explain

Table 2 Correlation analysis of biomass with spectrum, topography and texture features

feature	correlation coefficients masson pine	exotic pine	Chinese sweet gum	oak	mixed forest
*XS*1	–0.117	–0.206	–0.064	0.165	0.158
*XS*2	–0.233	–0.256	–0.216	0.001	0.193
*XS*3	–0.499	0.178	–0.108	0.150	0.079
*XS*4	–0.447	–0.069	–0.220	–0.001	0.186
PAN	–0.148	–0.178	0.041	0.017	0.164
*PC*1	–0.369	–0.294	0.033	0.014	0.182
*PC*2	–0.327	–0.084	0.093	–0.008	0.279
*PC*3	0.281	0.348	0.258	0.174	0.162
NDVI	–0.126	0.267	0.181	0.035	0.256
RVI	–0.116	0.252	0.171	0.045	0.276
Elevation	–0.286	0.018	–0.118	–0.268	–0.217
Slope	–0.212	–0.164	–0.233	–0.173	–0.242
Aspect	–0.033	–0.002	–0.094	0.076	–0.030
ME	–0.444(51)	–0.334(3)	0.042(27)	0.144(29)	0.267(35)
VA	0.156(31)	–0.124(9)	0.157(9)	0.146(7)	–0.053(3)
HO	–0.040(23)	0.248(9)	–0.247(7)	0.122(29)	0.094(5)
CO	0.166(11)	–0.185(9)	0.131(9)	0.135(5)	0.057(31)
DI	0.139(11)	–0.230(9)	0.211(7)	0.120(5)	0.112(31)
EN	0.093(13)	–0.319(9)	0.257(11)	–0.035(29)	0.048(3)
SM	0.052(5)	0.415(11)	–0.213(5)	0.228(29)	–0.067(11)
CR	0.198(7)	0.427(9)	–0.143(9)	–0.146(5)	–0.134(3)
SK	0.112(7)	–0.156(9)	–0.142(11)	0.098(5)	–0.056(7)

Notes: the number in the brackets refers to the optimal moving window size when extracting this texture

Table 3 Biomass estimation models and corresponding statistics

forest type	optimal biomass estimation model	correlation coefficient R^2 (R^2_a)	model validation (RMS of predictive error)	contribution of textures (R^2 change value)
Masson pine	$Y = 0.001 + 0.261XS1 - 0.375XS3 + 0.488PAN + 0.289PC2 - 0.228PC3\text{-}0.382\text{Elevation} - 0.807ME(51) + 0.211VA(31) + 0.410HO(23) + 0.408CO(11) + 0.227SM(5) + 0.159CR(7)$	0.582(0.397)	0.801	0.205
Exotic pine	$Y = 0.003 - 0.754XS3 + 2.628PC1 - 2.806PC2 - 0.669PC3 + 7.976NDVI - 5.050RVI - 1.489ME(3) + 1.039VA(9) - 0.876EN(9)$	0.642(0.422)	0.743	0.258
Chinese sweet gum	$Y = -0.000036 - 0.342XS2 - 0.581XS3 + 0.335PC3 + 3.679NDVI - 3.199RVI + 0.446ME(27) + 1.175VA(9) - 4.294HO(7) - 4.964DI(7) + 0.446EN(11) + 0.226SM(5) + 0.471CR(9)$	0.518(0.388)	0.854	0.120
Oak	$Y = -0.117XS2 + 0.185XS3 + 0.173\text{Slope} - 0.403\text{Elevation} + 0.369SM(29) - 0.291\text{CR}(5)$	0.263(0.208)	0.996	0.122
Mixed forest	$Y = -0.744XS3 + 0.705XS4 - 0.958PC1 + 1.933PC2 + 0.524PC3 - 1.175NDVI\text{-}0.320\text{Slope} - 175VA(3) + 0.129EN(3) + 0.162CR(3)$	0.298(0.237)	0.982	0.044

Notes: contribution of textures could be understood as the reduction of explanation capabilities of non-textural dependents when eliminating all textures from these multivariant regression models.

biomass variability. Meanwhile, statistical analysis shows that certain features, though they have weak correlations with biomass, as the inputs of estimation model, will improve estimation performance. Additionally, Table 2 tells us that aspect always has weak linkages with biomass of each forest type, but slope and elevation have relatively high and negative correlation with biomass. This phenomenon corresponds with the actual surroundings of forest growth. Generally, a forest, locating at abrupt and high elevation surroundings, will have a poor growth rate with a small biomass accumulation. As for the biomass of exotic pine, elevation has a quite weak and positive correlation with it (a correlation coefficient of 0.018). After examining the data of 30 subcompartments, we notice that their elevations of central points range varied little, from 50 m to 70 m. Statistically, this small coefficient of 0.018 results from uncontrollable errors. As for all textural measures, only mean and variance correlate with biomass closely. Table 3 shows that models for coniferous biomass estimation have higher accuracy than models for biomass estimation of broad-leaved forests. The fifth column in Table 3 indicates that it is very significant to introduce different textures into estimation models for different forest types, especially, for estimation of coniferous biomass. Introduction of textures leads to a big improvement of models in explaining the variance of biomass. Nevertheless, for the estimation of broad-leaved biomass, the role of textures is small, mainly due to the irregular crowns of broad-leaved forests and their poor repeatability. Textural measures emphasize the spatial repeatability and regular alignment of textural elements; hence, using it to represent the spatial configuration characteristics of broad-leaved forests will definitely have a poor performance. For example, for a mixed forest consisting of crowns with different sizes, heights and shapes, its spatial repeatability will be extremely chaotic, so there is no way for nine textural measures to distinguish them effectively.

5 Conclusions

The findings in this research could be outlined as follows: 1) Textural measures based on GLCM are effective and significant to characterize the spatial configurations of forests, especially for coniferous forests. However, for the spatial representation of broad-leaved forests, they play a poor role; 2) As for nine textural measures mentioned in this paper, ME and VA do a good job in characterizing forest spatial configuration, almost showing up in every estimation model and being somewhat capable of explaining the variance of biomass; and 3) Spectral features and topographical features are also important controls in representing forest configuration parameters.

Acknowledgements This study was funded by Natural Science Study Plans for Universities and Colleges in Jiangsu Province (05KJD220096) and Science and Technology Innovation Projects sponsored by Nanjing Forestry University (CX05-010-4). We also sincerely thank BIAN Lei, MIAO Qi, SHI Zheng, SUN Xin and LIN Chen, the 2006 undergraduates from Nanjing Forestry University, for their contributions to data extraction and standardization.

References

Chica M, Abarca F (2000). Computing geostatistical image texture for remotely sensed data classification, Comput Geosci, 26: 373–383

Fan J C, Mei C L (2002). Data Analysis. Beijing: Science Press, 94–117 (in Chinese)

Fang J Y, Liu G H, Xu S L (1996). Biomass and net primary productivity of forest vegetation in China. Acta Ecol Sin, 16(5): 497–508 (in Chinese)

Haralick R (1979). Statistical and structural approaches to texture. P IEEE, 67(5): 786–804

He D C, Wang L (1990). Texture unit, textural spectrum and texture analysis. IEEE T Geosci Remote Sens, 28: 509–512

Huang Y D, Li P J, Li Z X (2003). The application of geostatistical image texture to remote sensing lithological classification. Remote Sens Land Resour, (3): 45–49 (in Chinese)

Marceau D, Howarth P (1990). Evaluation of the gray-level co-occurrence matrix method for land cover classification using SPOT imagery. IEEE T Geosci Remote Sens, 28: 513–519

Sakari T, Anssi P (2005). Performance of different spectral and textural aerial photograph features in multi-source forest inventory. Remote Sens Environ, 94: 256–268

Wang R S (1994). Image Understanding. Changsha: Press of National University of Defense Technology, 145–183 (in Chinese)

Xue C S, Wang X (1997). Textural analysis of remotely sensed imagery based on fractal geometry and its application. Geoll Sci Technol Inform, 16(Supp.): 99–105 (in Chinese)

Zhao X W, Li C G (2001). Quantitative Estimation of Forest Resources Based on "3S" Technologies. Beijing: China Science and Technology Press, 18–41 (in Chinese)