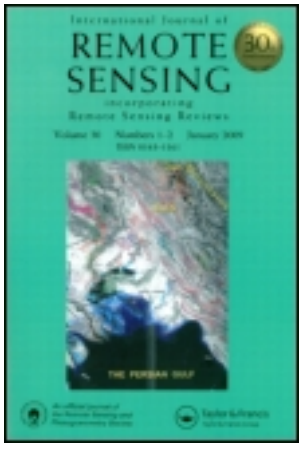

International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

Retrieval of forest growing stock volume by two different methods using Landsat TM images

Sheng Zheng[ab], Chunxiang Cao[a], Yongfeng Dang[c], Haibing Xiang[ab], Jian Zhao[ab], Yuxing Zhang[c], Xuejun Wang[c] & Hongwen Guo[d]

[a] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences and Beijing Normal University, Beijing 100101, PR China

[b] University of Chinese Academy of Sciences, Beijing 100049, PR China

[c] Academy of Forest Inventory and Planning, State Forestry Administration, Beijing 100714, PR China

[d] Anshan Wood & Wood Products Testing & Inspection Management Centre, Anshan 114001, PR China

Published online: 24 Dec 2013.

To cite this article: Sheng Zheng, Chunxiang Cao, Yongfeng Dang, Haibing Xiang, Jian Zhao, Yuxing Zhang, Xuejun Wang & Hongwen Guo (2014) Retrieval of forest growing stock volume by two different methods using Landsat TM images, International Journal of Remote Sensing, 35:1, 29-43, DOI: 10.1080/01431161.2013.860567

To link to this article: http://dx.doi.org/10.1080/01431161.2013.860567

International Journal of Remote Sensing, 2014
Vol. 35, No. 1, 29–43, http://dx.doi.org/10.1080/01431161.2013.860567

Taylor & Francis
Taylor & Francis Group

Retrieval of forest growing stock volume by two different methods using Landsat TM images

Sheng Zheng[a,b], Chunxiang Cao[a]*, Yongfeng Dang[c], Haibing Xiang[a,b], Jian Zhao[a,b], Yuxing Zhang[c], Xuejun Wang[c], and Hongwen Guo[d]

[a]*State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences and Beijing Normal University, Beijing 100101, PR China;* [b]*University of Chinese Academy of Sciences, Beijing 100049, PR China;* [c]*Academy of Forest Inventory and Planning, State Forestry Administration, Beijing 100714, PR China;* [d]*Anshan Wood & Wood Products Testing & Inspection Management Centre, Anshan 114001, PR China*

(*Received 22 March 2013; accepted 4 June 2013*)

Forest growing stock volume (GSV) is one of the most important indicators in the field of forest resources investigation and monitoring. This article describes the application of two different methods, the multiple stepwise regression (MSR) model and the back-propagation neural network (BPNN), to retrieve forest GSV using Landsat Thematic Mapper (TM) images and field data. The article describes the data used, the retrieval methods adopted, and the results achieved. The results show that the surface reflectance of six bands significantly correlated with forest GSV, as did six vegetation indices, factors from principal component analysis and tasselled cap transformation, and three terrain factors. Moreover, texture features including Band 1_{mean}, Band 2_{mean}, and Band 3_{mean} were highly correlated with forest GSV. An optimal MSR model that included three factors was established for retrieving forest GSV using 53 remote-sensing factors. Three factors were included in the model. Leave-one-out cross-validation demonstrated that the model worked well. Finally, BPNN was constructed and the predicted result was highly consistent with measured forest GSV. In a comparison of the retrieved results with the MSR model and BPNN, the MSR model was better at quantitatively finding the correlation between each remote-sensing factor and forest GSV, and a linear equation could be acquired. However, BPNN was better at predicting forest GSV based on the field data. Additionally, the retrieved map of forest GSV for the whole study area by BPNN was much more consistent with the Landsat TM false-colour composite than that retrieved by the MSR model.

1. Introduction

Sustainable forest management requires a great deal of data in regard to describing and quantifying forest variables. The retrieval of forest information has become increasingly important in recent decades because of issues related to global climate change. The forest growing stock volume (GSV) is one of the most important forest variables in the context of forest management. GSV is a major predictor for assessing above-ground biomass (Shvidenko et al. 2007) and is crucial for estimating the compartment (Jenkins et al. 2003) or total above-ground biomass (Somogyi et al. 2008), which is a fundamental variable for estimating net carbon dioxide exchange between the land surface and the atmosphere. Therefore, GSV is an important parameter not only for acquiring a more detailed

*Corresponding author. Email: cao413@irsa.ac.cn

evaluation of terrestrial biosphere models (Beer et al. 2006) but also for inversely estimating the parameters of biosphere models (Carvalhais et al. 2010).

To understand the status of national forest resources, China carries out a national forest continuous inventory every 5 years and the survey and design of a forest resource plan every 10 years (Yang et al. 2003). The investigation data used to survey and evaluate forest area are traditionally acquired manually, and this process requires a great deal of human and material resources. Remote-sensing technology provides a low-cost and effective way to obtain large-scale and high-precision data, which play a critical role in the investigation of GSV. Since the 1970s the use of satellite images, such as Landsat Thematic Mapper (TM) and Earth-observing Satellites (SPOT), for estimating continuous forest parameters has been widely studied (Kilkki and Paivinen 1987; Katila and Tomppo 2001; Cao et al. 2009).

There are three main types of remote-sensing data used for the retrieval of GSV: airborne lasers, Landsat TM, and lidar. Synthetic aperture radar data also play a major role in estimating GSV. In 1984, Nelson et al. (1984,1988) began estimating forest canopy density using airborne laser data, and estimated forest biomass and volume using laser data in 1988. Naesset et al. (2002) used three methods to estimate average tree height, the height of dominant tree species, average diameter, and other variables in 233 sample areas based on laser data. Lidar can extract vertical structure information of vegetation, and this technology can be applied to estimating forest GSV. Estimated forest GSV can be accurate for coniferous forests of average age or for single species using lidar data. Donoghue et al. (2007) estimated GSV using two methods in three study areas of Scotland using lidar data. He et al. (2012) studied forest stand biomass estimations using Advance Land Observing Satellite Phased Array type L-band Synthetic Aperture Radar (ALOS PALSAR) data based on lidar-derived prior knowledge in the Qilian Mountains, western China, and the results showed that at the stand level and in different biomass grades, backscatter coefficient increased with increase in forest biomass.

In previous studies, several spectral features of Landsat TM were tested and compared in relation to forest GSV estimates, including the normalized difference vegetation index, the environmental vegetation index, and a combination of spectral features. Forest GSV was found to be correlated with the spectral features to some extent (Gemmell 1995; Tomppo et al. 2002; Chirici et al. 2008). Gemmell (1995) studied the impact of each band, canopy density, forest stand size, and topographic features on GSV retrieval based on Landsat TM data. The results showed that TM4 and TM5 were highly correlated with GSV. The research results provided a theoretical basis for Landsat TM in the quantitative estimation of GSV. Fazakas et al. (1999) applied TM data and national forest inventory data to estimate GSV in central Sweden by the *k*-nearest-neighbour (*k*-NN) method. Tomppo et al. (2002) established a nonlinear regression model to estimate GSV in a large area by the non-parametric nearest-neighbour method based on Landsat TM, Indian Remote Sensing Satellites-1C Wide Field Sensor (IRS-1C WiFS), and sample data. The results showed that GSV could be retrieved more accurately when Landsat TM and IRS-1C WiFS were combined. Mäkelä and Pekkarinen (2004) applied the average of TM data pixels as a basis from which to estimate the GSV of different tree species using the *k*-NN algorithm. Diamantopoulou (2005) compared the artificial neural network (ANN) model and several other nonlinear models to estimate the GSV of pine trees and found that the root mean square error (RMSE) of the three-layer forward feedback ANN model was lower than that of the optimal nonlinear regression model. Gu et al. (2006) applied the *k*-NN method to Landsat TM images to improve the estimation of forest volume. Chirici et al. (2008) described the application of non-parametric and parametric

methods for estimating forest GSV using Landsat images on the basis of data measured in the field and integrated with ancillary information. The results cast a promising light on the use of non-parametric techniques for forest attribute estimation and mapping, with accuracy sufficiently high to support forest planning activities in complex landscapes.

A comprehensive analysis of the correlation between forest GSV and image information extracted from Landsat TM is currently lacking, and few studies have compared the retrieved results of forest GSV based on Landsat TM images using linear and nonlinear models (Mäkelä and Pekkarinen 2004; Gu et al. 2006; Chirici et al. 2008). In this article, we describe the application of a multiple stepwise regression (MSR) model and back-propagation neural network (BPNN) for retrieving forest GSV in Xiuyan using Landsat TM images from 2006 and field data from the national forest continuous inventory.

2. Study area and data

2.1. Study area

Our study area is located in Xiuyan in southeastern Liaoning, China (Figure 1). Xiuyan's latitude ranges from 40° 00′ to 40° 50′ N and its longitude from 122° 52′ to 123° 46′ E, constituting a total area of 4507 km^2. Xiuyan is a Manchu autonomous county and is under the administration of Anshan. Xiuyan has a population of 500,000, with 20 towns and 3 townships are under its administration. Xiuyan's forest cover is 73%, the main type being broadleaf forest consisting mainly of oak. In addition, there is a smaller population of poplar trees.

Xiuyan has a monsoon-influenced humid continental climate, characterized by hot, humid summers because of the monsoon and long, cold, and very dry winters because of

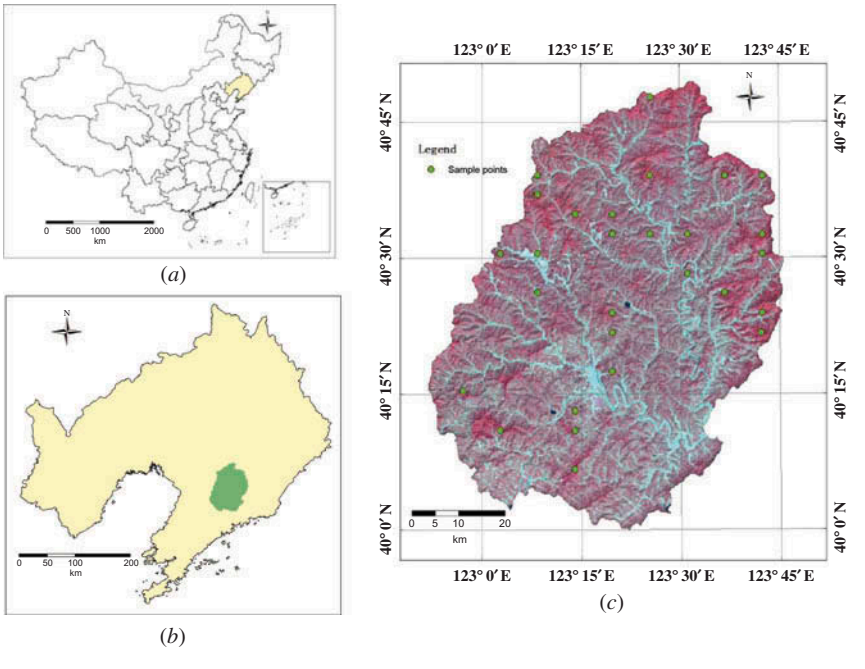

Figure 1. Study area. (*a*) Location of Liaoning in China; (*b*) location of Xiuyan in Liaoning; (*c*) false-colour composite of TM bands 4, 3, 2. The green points represent the selected sample points in the study area.

Table 1. Technical parameters of Landsat-5.

Band	Wavelength (μm)	Spectral range	Ground pixel resolution (m)
1	0.45–0.52	Blue	30
2	0.52–0.60	Green	30
3	0.63–0.69	Red	30
4	0.76–0.90	Near-infrared	30
5	1.55–1.75	Shortwave infrared	30
6	10.4–12.5	Thermal infrared	120
7	2.08–2.35	Shortwave infrared	30

the Siberian anticyclone. A majority of the annual rainfall occurs in July and August. The annual average temperature is 7.39°C, and the average annual rainfall is 775.8–933.8 mm (https://en.wikipedia.org/wiki/Xiuyan_Manchu_Autonomous_County).

2.2. *Data*

The measured field data are from the national forest continuous inventory from 2005. This inventory is carried out by the State Forestry Administration, PR China, which aims to understand the status of forest resource and trends on a large scale. The main survey factors are average tree diameter and height and GSV. The survey is carried out periodically at fixed sample points, and the area of each sample point is 0.0667 ha. The density of sample points is set according to the required estimate precision of forest land area, growing stock, amount of growth and consumption. Distance between points can vary from 2 × 2 to 8 × 8 km (FAO 2007); in Xiuyan this is 4 km from north to south and 8 km from west to east.

With regard to remote-sensing data, we chose one high-quality image (30 m ground pixel resolution) acquired on 18 September 2006, from Landsat TM data. The technical parameters of Landsat-5 are shown in Table 1. Other auxiliary data used include the Advanced Spaceborne Thermal Emission and Reflection Radiometer Digital Elevation Model (ASTER DEM) with a resolution of 30 m, and we generated factors including slope and aspect for the study area according to ASTER DEM.

3. Methods

3.1. *Pre-processing*

The pre-processing of Landsat TM images mainly includes geometric correction, orthorectification, atmospheric correction, image cropping, and classification. In this article, atmospheric correction of the Landsat TM image was performed using the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) software package in Environment for Visualizing Images (ENVI). The FLAASH module was developed by Spectral Sciences under the sponsorship of the US Air Force Research Laboratory (Adler-Golden et al. 1999). The module incorporates Moderate Resolution Atmospheric Transmission (MODTRAN) 4 radiation transfer code with all MODTRAN atmosphere and aerosol types to calculate a unique solution for each image. The input parameters of the MODTRAN model can be divided into five categories of parameter: satellite operation, sensor, atmospheric, observation geometry, and surface (Berk et al. 1999; Zheng et al. 2011). Image cropping was utilized to extract Xiuyan from the Landsat TM image,

and classification based on the maximum likelihood methods was applied to extract the forest area in Xiuyan. According to sample points with the attribute of land-cover types from the national forest continuous inventory in Xiuyan, we identified land-cover types from the Landsat TM image: forest and no-forest areas including farmland, settlement, and water. According to the classification map, forest cover is 70%, which is close to the actual 73% forest cover. In addition, we calculated the confusion matrix and kappa coefficient (0.9) for the resulting classification.

The pre-processing for the measured field forest GSV data mainly included choosing sample points of forest GSV greater than 0 from the national forest continuous inventory in Xiuyan and then deleting abnormal points. Twenty-eight field data points were saved for future study (Figure 1(*c*)). Because the ground pixel resolution of the Landsat TM image is 30 m, the area of each pixel is 900 m^2. However, the area of each sample point from the national forest continuous inventory is 0.0667 hm^2 or 667 m^2. Therefore, before establishing the retrieval model for forest GSV, we calculated forest GSV of the pixel corresponding to the sample point using Equation (1):

$$V_p = \frac{900}{667} \times V_s, \tag{1}$$

where V_p refers to the GSV of one sample point and V_s represents the GSV of the pixel corresponding to the sample point.

3.2. *Remote-sensing factors*

Based on the Landsat TM data after pre-processing, we acquired the surface reflectance for each band of the Landsat TM (except band 6) and then acquired the vegetation index through band calculation of the Landsat TM image, obtaining factors from principal component analysis (PCA) and tasselled cap transformation. In addition, we obtained a further 30 variables through texture analysis using a window size of 3 × 3 pixels for the Landsat TM image. All the remote-sensing factors used are listed in Table 2.

In Table 2, Band 1 refers to the surface reflectance of TM band 1, as do Band 2, Band 3, etc. Band 5/Band 4 and Band 2/Band 4 represent Band 5 divided by Band 4 and Band 2 divided by Band 4, respectively. Brightness, Greenness, and Third denote the brightness, greenness, and third component after tasselled cap transformation, respectively. PCA 1, PCA 2, and PCA 3 represent the first three components after PCA. Equations and explanations of the vegetation index including DVI, SAVI, ARVI, PVI, NDVI, etc., are as follows.

Table 2. Remote-sensing factors calculated from the Landsat TM image.

Factor type	Remote-sensing factors
Band reflectance	Band 1, Band 2, Band 3, Band 4, Band 5, Band 7
Band math	DVI, SAVI, ARVI, PVI, NDVI, RVI, (Band 5)/(Band 4), (Band 2)/(Band 4)
Tasselled cap transformation	Brightness, Greenness, Third
Principal component analysis	PCA 1, PCA 2, PCA 3
Terrain	DEM, Aspect, Slope
Texture analysis	Range, mean, variance, entropy, skewness

DVI (difference vegetation index) is very sensitive to changes in soil background and can identify vegetation and water (Jordan 1969):

$$\mathrm{DVI} = (\mathrm{Band4}) - (\mathrm{Band3}). \tag{2}$$

NDVI (normalized difference vegetation index) increases the vegetation response capacity and is currently the most widely used vegetation index:

$$\mathrm{NDVI} = \frac{(\mathrm{Band\ 4}) - (\mathrm{Band\ 3})}{(\mathrm{Band\ 4}) + (\mathrm{Band\ 3})}. \tag{3}$$

Compared with NDVI, SAVI (soil-adjusted vegetation index) adds the soil adjustment coefficient L, which is determined by its actual situation and ranges from 0 to 1;

$$\mathrm{SAVI} = \frac{((\mathrm{Band\ 4}) - (\mathrm{Band\ 3})) \times (1 + L)}{((\mathrm{Band\ 4}) + (\mathrm{Band\ 3})) \times (3 + L)}. \tag{4}$$

To eliminate the influence of the atmosphere, Kaufman and Tanre (1992) introduced ARVI (atmospherically resistant vegetation index). This index corrects for radiance in the red channel using the radiance difference between the blue and red channels when defining NDVI;

$$\mathrm{ARVI} = \frac{(\mathrm{Band\ 4}) - (2 \times (\mathrm{Band\ 3}) - (\mathrm{Band\ 1}))}{(\mathrm{Band\ 4}) + (2 \times (\mathrm{Band\ 3}) - (\mathrm{Band\ 1}))}. \tag{5}$$

PVI (perpendicular vegetation index) eliminates the influence of soil background and is less sensitive to atmospheric influences than the other vegetation indices:

$$\mathrm{PVI} = \frac{(\mathrm{Band\ 4}) - a \times (\mathrm{Band\ 3}) - b}{\sqrt{(1 + a^2)}}. \tag{6}$$

In Equation (6), both a and b represent coefficients. Owing to the markedly different responses to green vegetation between the red and near-infrared bands, the ratio vegetation index (RVI) can fully express the difference between these:

$$\mathrm{RVI} = (\mathrm{Band\ 4})/(\mathrm{Band\ 3}). \tag{7}$$

Terrain factors include ASTER DEM, with a resolution of 30 m, and slope and aspect, which were generated using ASTER DEM.

PCA is a way of identifying patterns in data and expressing the data to highlight their similarities and differences. The other main advantage of PCA is that once these patterns in the data have been identified and data have been compressed (i.e. by reducing the number of dimensions), little information is lost (Jiang et al. 2006; Francioso et al. 2008). We selected the first three principal components, which were marked as PCA 1, PCA 2, and PCA 3.

Tasselled cap transformation is a useful tool for compressing spectral data into a few bands associated with physical scene characteristics (Crist and Cicone 1984). This transformation can maximize the separation of soil and vegetation. We chose the first

three bands generated from tasselled cap transformation, which were marked as Brightness, Greenness, and Third.

Texture is one of the most important features of remote-sensing images. Texture analysis refers to the process of extracting texture features through image processing techniques and analysing the texture quantitatively or qualitatively. In this article, five texture features – range, mean, variance, entropy, and skewness – were derived from the Landsat TM image.

3.3. Regression model

To retrieve forest GSV for a large-scale area, we established the simple regression model based on all remote-sensing factors. We then chose the stepwise regression method to establish the MSR model. The regression model was performed using Statistical Product and Service Solutions (SPSS) software, version 16.0 (SPSS Inc., Chicago, IL, USA). A two-tailed analysis was used for all statistical tests, and p-values < 0.05 were considered statistically significant.

With regard to the validation of the results retrieved by a multiple regression model, we applied the leave-one-out cross-validation. An explanation of leave-one-out cross-validation follows here. First, among all sample points, we excluded one every time and established the model using the remainder of the sample points. We then predicted the excluded sample point using the established model. We repeated this process until every sample point was predicted once. Then, we validated the retrieved results based on measured field forest GSV and predicted forest GSV.

3.4. Back-propagation neural network

BPNN is considered one of the most effective methods for solving ANNs. Because ANNs have proved to be powerful tools that are particularly suited for various tasks in deducing uncertainties and because the models have the ability to learn and extract x–y relationships from the training samples, their flexibility is a decisive asset compared with parametric techniques that require the assumption of a specifically hard model form (Bai and Jin 2005). BPNN is an error back-propagation algorithm for the multi-layer network and is based on the main principle that several parameters will have an impact on forecasting factors as network inputs and will form a network with foresting factors as the output. The network is used to carry out self-organizing learning, achieving targets by ceaselessly projecting the impact factors of the nonlinear mapping relationship between the expectations and observations with a smaller mean square error (MSE). BPNN contains an input layer, a hidden layer, and an output layer. The weights of the neural network are a feed-forward or feedback process through a number of interconnected neurons, which are primarily assigned randomly then adjusted by decreasing MSE (Cao et al. 2010).

4. Results

4.1. Simple regression model

As Table 2 shows, we divided the remote-sensing factors into six types. To quantify the impact of each factor on forest GSV, we extracted each remote-sensing factor's value in the sample point using the nearest-neighbour resampling technique and then analysed the correlation between forest GSV and each remote-sensing factor using Pearson correlation analysis. The results are shown in Table 3. Subsequently, we established the simple

Table 3. Pearson correlation coefficients (r) between remote-sensing factors and measured forest GSV.

Remote-sensing factors	r	Sig.	Remote-sensing factors	r	Sig.	Remote-sensing factors	r	Sig.
Band 1	–0.694	0.000	Greenness	–0.322	0.094	Band $3_{entropy}$	–0.247	0.205
Band 2	–0.660	0.000	Third	0.438	0.020	Band $3_{skewness}$	0.343	0.074
Band 3	–0.584	0.001	DEM	0.446	0.017	Band 4_{range}	0.564	0.002
Band 4	–0.461	0.014	Aspect	0.159	0.419	Band 4_{mean}	–0.446	0.017
Band 5	–0.518	0.005	Slope	0.482	0.009	Band $4_{variance}$	0.511	0.005
Band 7	–0.506	0.006	Band 1_{range}	–0.198	0.313	Band $4_{entropy}$	0.356	0.063
DVI	–0.401	0.034	Band 1_{mean}	–0.801	0.000	Band $4_{skewness}$	–0.207	0.291
SAVI	–0.373	0.05	Band $1_{variance}$	–0.181	0.356	Band 5_{range}	0.109	0.581
PVI	–0.403	0.033	Band $1_{entropy}$	0.072	0.718	Band 5_{mean}	–0.548	0.003
NDVI	0.316	0.102	Band $1_{skewness}$	0.029	0.884	Band $5_{variance}$	0.038	0.848
ARVI	0.145	0.463	Band 2_{range}	–0.126	0.522	Band $5_{entropy}$	–0.018	0.926
RVI	0.318	0.099	Band 2_{mean}	–0.736	0.000	Band $5_{skewness}$	–0.325	0.092
(Band 5)/(Band 4)	–0.149	0.450	Band $2_{variance}$	–0.196	0.316	Band 7_{range}	0.002	0.992
(Band 2)/(Band 4)	–0.402	0.034	Band $2_{entropy}$	0.076	0.702	Band 7_{mean}	–0.541	0.003
PCA 1	0.492	0.008	Band $2_{skewness}$	0.047	0.813	Band $7_{variance}$	–0.037	0.851
PCA 2	0.206	0.294	Band 3_{range}	–0.393	0.039	Band $7_{entropy}$	0.171	0.384
PCA 3	0.056	0.779	Band 3_{mean}	–0.674	0.000	Band$7_{skewness}$	0.006	0.977
Brightness	–0.540	0.003	Band $3_{variance}$	–0.361	0.059			

Note: Sig., Significance.

Table 4. Model summary of simple regression model.

Model	Variable	R^2	Std. error of the estimate	Sig.
1	Band 1	0.482	3.0975	0.000
2	Band 2	0.436	3.2326	0.000
3	Band 3	0.342	3.4925	0.001
4	Band 4	0.212	3.8195	0.014
5	Band 5	0.268	3.6817	0.005
6	Band 7	0.256	3.7129	0.006
7	DVI	0.161	3.9422	0.034
8	PVI	0.163	3.9383	0.033
9	(Band 2)/(Band 4)	0.162	3.9408	0.034
10	PCA 1	0.242	3.7461	0.008
11	Brightness	0.292	3.6221	0.003
12	Third	0.192	3.8692	0.020
13	DEM	0.199	3.8514	0.017
14	Slope	0.232	3.7711	0.009
15	Band 1_{mean}	0.642	2.5754	0.000
16	Band 2_{mean}	0.542	2.9142	0.000
17	Band 3_{range}	0.154	3.9580	0.039
18	Band 3_{mean}	0.455	3.1783	0.000
19	Band 4_{range}	0.318	3.5549	0.002
20	Band 4_{mean}	0.199	3.8513	0.017
21	Band $4_{variance}$	0.261	3.7004	0.005
22	Band 5_{mean}	0.300	3.6012	0.003
23	Band 7_{mean}	0.293	3.6196	0.003

Note: Std., Standard.

regression model using the remote-sensing factor, which was significantly correlated with forest GSV. A summary of the model is shown in Table 4.

First, we analysed the correlation between the surface reflectance of the six bands and measured forest GSV, and found significant correlation in all cases. The model summary shows that among these six bands, Band 1 had the highest correlation (R^2 0.482) and the lowest standard error of the estimate (3.0975). After analysing the correlation between variables from band math and measured forest GSV, DVI, PVI, and (Band 2)/(Band 4) were found to be significantly correlated with forest GSV. In addition to these factors frequently used in previous studies, we also found that among the terrain factors and those from tasselled cap transformation and PCA, DEM, Slope, PCA 1, Brightness, and Third were significantly correlated with forest GSV. Finally, we analysed the correlation between 30 factors from texture analysis for each band and measured forest GSV. The results showed that Band 1_{mean}, Band 2_{mean}, and Band 3_{mean} were highly correlated with forest GSV ($R^2 > 0.45$).

4.2. *Multiple regression model*

Using simple regression, we analysed the correlation between each remote-sensing factor and forest GSV (Table 2) and established a simple regression model for retrieving forest GSV. The surface reflectance of six bands, vegetation index, and all other available ancillary information were correlated with forest GSV to varying degrees, so these were considered as potential independent variables for forest GSV estimation. We established an MSR model using these remote-sensing factors. Even when a variable was not

Table 5. Model summary of the MSR model.

Model	Variable	R^2	Std. error of the estimate	Sig.
1	Band 1_{mean}, Band $3_{skewness}$, PCA 3	0.80	2.02434	0.000

Table 6. Coefficients of the MSR model.

Model	Independent variable	Dependent variable	Nonstandardized coefficients	Std. error of the nonstandardized coefficients	Sig.
1	Constant,	Measured	12.289	1.028	0.000
	Band 1_{mean},	forest	–743.653	83.289	0.000
	Band $3_{skewness}$,	GSV	2.179×10^{-8}	0.000	0.001
	PCA 3		–93.446	33.801	0.011

significantly correlated with forest GSV, it still could be introduced to the MSR model if it, combined with other variables, could influence forest GSV significantly. Therefore, the MSR model in this article was the optimal MSR model based on the 53 remote-sensing factors listed in Table 2. The model summary and model coefficients are shown in Tables 5 and 6, respectively. The resulting regression model is shown in Equation (8), with R^2 of 0.80. In the optimal MSR model, Band 1_{mean}, Band $3_{skewness}$, and PCA 3 were included, and the predicted forest GSV using this model was basically consistent with the measured forest GSV (Figure 2):

$$\begin{aligned} GSV = {} & (12.289 - 743.653) \times (\text{Band } 1_{mean}) + (2.179 \times 10^{-8}) \\ & \times (\text{Band } 3_{skewness}) - 93.446 \times (\text{PCA } 3). \end{aligned} \tag{8}$$

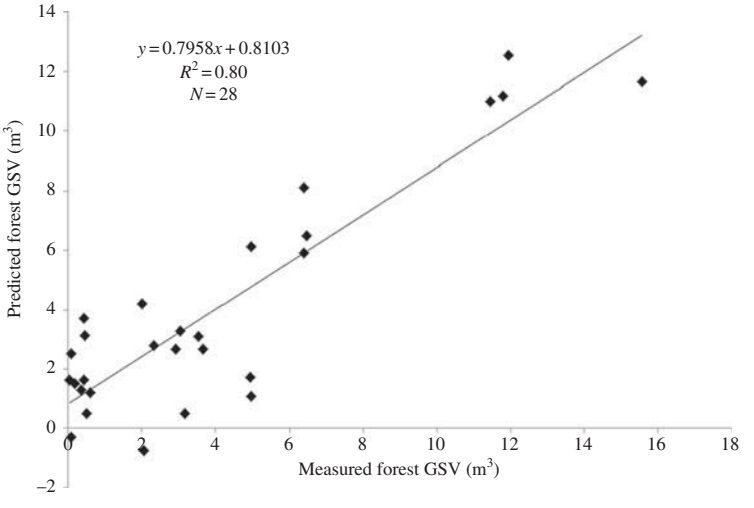

Figure 2. Relation between measured and predicted forest GSV.

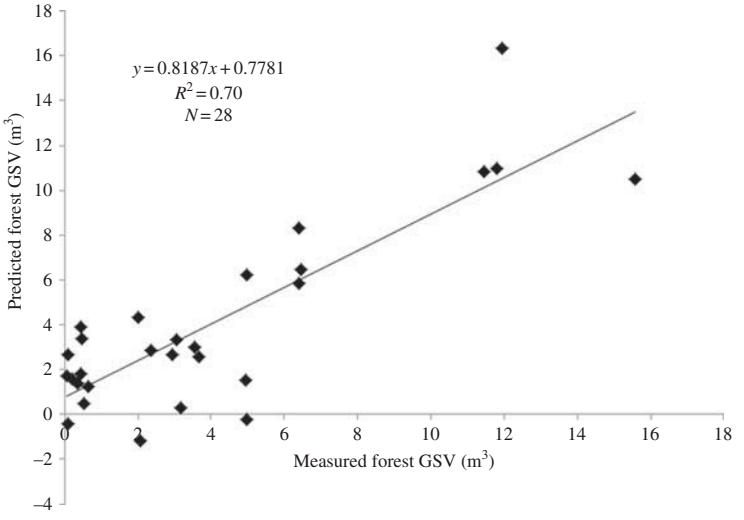

Figure 3. Leave-one-out cross-validation result of forest GSV from the MSR model.

We validated the results retrieved by the MSR model through leave-one-out cross-validation for the 28 field data points. The cross-validation results are shown in Figure 3, with an R^2 of 0.70 and RMSE of 2.3 m^3, demonstrating that the model was highly accurate. Based on the MSR model shown in Equation (8), we retrieved forest GSV in the study area shown in Figures 4(*a*) and (*c*). In this figure, white represents a non-forested area. Forest GSV of most pixels in the study area ranged from 0 to 2 m^3.

4.3. Back-propagation neural network

There were three main steps to achieving prediction of forest GSV by BPNN: network creation, learning, and network simulation. In the process of learning, the input layer included seven basic factors – the reflectance of six bands and DEM – and the output layer was forest GSV. There were two main reasons for only including these seven basic factors and excluding the other remote-sensing factors. First, the other remote-sensing factors could be calculated from these seven basic factors. Second, the accuracy of the predicted results met our requirement using these seven basic factors. The predicted result was highly consistent with measured forest GSV, with R^2 of 1 and RMSE of 0.0014 m^3 for the 28 field data points. We then used this trained network to predict forest GSV in the study area, and the results are shown in Figures 4(*b*) and (*d*).

5. Discussion

This article focused on retrieving forest GSV using Landsat TM images and measured data using both the MSR model and BPNN. The measured data in the study area were from the national forest continuous inventory, and 28 field data points were saved after deleting abnormal points. By analysing the correlation between forest GSV and the 53 remote-sensing factors, we found that the surface reflectance of all six bands was significantly correlated with forest GSV, and Band 1 had the highest correlation among the six bands. We then analysed the correlation between forest GSV and other

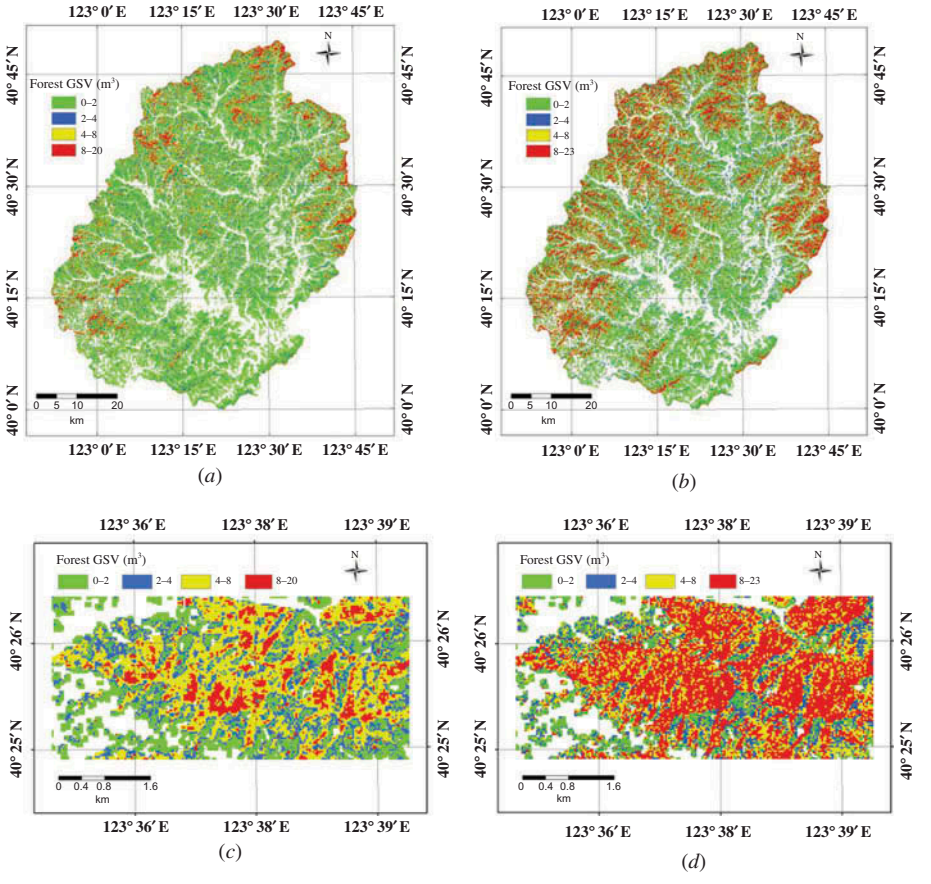

Figure 4. Forest GSV maps of the study area. (*a*) Map retrieved from the MSR model; (*b*) map retrieved from BPNN; (*c*) and (*d*) forest GSV of the local region retrieved from the MSR model and BPNN, respectively.

remote-sensing factors including vegetation index, PCA, tasselled cap transformation, and terrain. Each type of remote-sensing factor was correlated with forest GSV to some extent, but the coefficient of the Pearson correlation was not high.

In previous studies on retrieval of forest GSV, most researchers studied the correlation between this and band surface reflectance or vegetation index, whereas few researchers took texture features into account. In this article, considering that the texture information might be related to the retrieval of forest GSV, we calculated the coefficients of the Pearson correlation between forest GSV and texture features, including each band's data range, mean, variance, entropy, and skewness. The results showed that forest GSV was significantly correlated with many texture features and was particularly highly correlated with Band 1_{mean}, Band 2_{mean}, and Band 3_{mean}, because the texture feature considered the spatial relationship between the pixels, which is different from pixel-based spectral features (Atkinson et al. 2000). The texture feature improves the accuracy of image analysis and feature extraction (Berberoglu et al. 2000). Levesque et al. (2003), Moskal et al. (2004), and Kayitakire et al. (2006) studied forest health, pest status, and tree age, respectively, based on image texture analysis, and the results showed that the accuracy of the extracted information was higher using texture features than spectral features. In the

present article, from the false-colour composite of TM bands 4 (red), 3 (green), and 2 (blue), the distribution of forest GSV was consistent with texture feature to some extent. For example, the texture feature of an area with dense vegetation was more uniform, whereas that of an area with sparse vegetation exhibited greater differences. The texture features of different species were also somewhat different. Hence, the accuracy of retrieving the forest GSV could be improved by the use of texture information.

An optimal MSR model for retrieving forest GSV based on the 53 remote-sensing factors was established. In addition to Band 1_{mean}, which was most highly correlated with forest GSV, Band $3_{skewness}$ and PCA 3 were also included in the regression model. The R^2 of the regression model was 0.80, which was better than any simple regression model based on a single remote-sensing factor. Leave-one-out cross-validation was applied to validating the retrieved result, and the predicted forest GSV was basically consistent with measured forest GSV, with R^2 of 0.70 and RMSE of 2.3 m^3.

In addition to the MSR model, we built a BPNN to retrieve forest GSV. The input layer included seven variables – the reflectance of six bands and DEM. The predicted result from BPNN was highly consistent with measured forest GSV for the 28 field data points, a much better result than that of the MSR model. In addition, in comparing Figures 4(*c*) and (*d*), the forest GSV map retrieved from BPNN was much more consistent with the texture feature of the Landsat TM image than that retrieved from the MSR model.

During the process of building BPNN, we selected the appropriate learning algorithms and trained the network. We then predicted forest GSV using the trained network. Owing to the 'black box' operation of BPNN, we could not analyse the correlation between each factor and the forest GSV quantitatively. However, in the MSR model, we could clearly understand the impact of each factor on forest GSV and establish the linear regression equation.

Similar studies have focused on retrieval of forest GSV using optical remote-sensing data, the remote-sensing factors being mainly band surface reflectance or vegetation index. This article describes retrieval of forest GSV using two different methods using Landsat TM images. We found that the accuracy of retrieving forest GSV could be improved when taking into account texture information from Landsat TM images, so texture information in regard to optical remote-sensing data is important for retrieving forest GSV. Moreover, we compared the advantages and disadvantages of the MSR model and BPNN using Landsat TM images, and we should choose the appropriate method to meet our specific needs.

6. Conclusion

Three important conclusions can be drawn from this study. First, the comprehensive analysis of the correlation between forest GSV and image information extracted from Landsat TM images showed that in addition to the surface reflectance of six bands and other remote-sensing factors described in previous studies, forest GSV was highly correlated with certain texture factors. Second, based on the 53 remote-sensing factors, an optimal MSR model was established for retrieving forest GSV with R^2 of 0.80. Leave-one-out cross-validation demonstrated that the regression model worked well, with R^2 of 0.70. Moreover, the regression model was much better than any simple regression model using a single factor. Third, the predicted forest GSV using BPNN was highly consistent with measured forest GSV. Although BPNN failed to analyse the impact of each factor on forest GSV and the retrieval equation, it was better at both predicting forest GSV for the field data and retrieving the forest GSV map for the study area.

Acknowledgements

The authors thank the Academy of Forestry Inventory and Planning, State Forestry Administration, PR China, for providing the forest inventory data. We also wish to thank all those who have helped with the article.

Funding

This article was supported by the National Natural Science Foundation of China [grant number 41171330]; the National High Technology Research and Development Programme of China (863 Program) [grant number 2013AA12A302]; the Evaluation of Wetland Ecosystem.

References

Adler-Golden, S. M., M. W. Matthew, L. S. Bernstein, R. Y. Levine, A. Berk, S. C. Richtsmeier, P. K. Acharya, G. P. Anderson, J. W. Felde, and J. Gardner. 1999. "Atmospheric Correction for Shortwave Spectral Imagery Based on MODTRAN4." *International Society for Optics and Photonics* 3753: 61–69.

Atkinson, P., and P. Lewis. 2000. "Geostatistical Classification for Remote Sensing: An Introduction." *Computers & Geosciences* 26: 361–371.

Bai, Y. P., and Z. Jin. 2005. "Prediction of SARS Epidemic by BP Neural Networks with Online Prediction Strategy." *Chaos, Solitons & Fractals* 26: 559–569.

Beer, C., W. Lucht, C. Schmullius, and A. Shvidenko. 2006. "Small Net Carbon Dioxide Uptake by Russian Forests During 1981–1999." *Geophysical Research Letters* 33: L15403.

Berberoglu, S., C. Lloyd, P. Atkinson, and P. Curran. 2000. "The Integration of Spectral and Textural Information Using Neural Networks for Land-Cover Mapping in the Mediterranean." *Computers & Geosciences* 26: 385–396.

Berk, A., G. Anderson, P. Acharya, J. Chetwynd, L. Bernstein, E. Shettle, M. Matthew, and S. Adler-Golden. 1999. MODTRAN4 User's Manual. Air Force Research Laboratory, Space Vehicles Directorate.

Cao, C. X., C. Y. Chang, M. Xu, J. A. Zhao, M. X. Gao, H. Zhang, J. P. Guo, J. Guo, L. Dong, Q. He, L. Bai, Y. Bao, W. Chen, S. Zheng, Y. Tian, W. Li, and X. Li. 2010. "Epidemic Risk Analysis After the Wenchuan Earthquake Using Remote Sensing." *International Journal of Remote Sensing* 31: 3631–3642.

Cao, C. X., M. Xu, Q. S. He, and H. Zhang. 2009. "Trends Analysis on Forest Health Research Using Multi-Source Remote Sensing Data." *Journal of Remote Sensing* 13: 401–407.

Carvalhais, N., M. Reichstein, P. Ciais, G. J. Collatz, M. D. Mahecha, L. Montagnani, D. Papale, S. Rambal, and J. Seixas. 2010. "Identification of Vegetation and Soil Carbon Pools Out of Equilibrium in a Process Model via Eddy Covariance and Biometric Constraints." *Global Change Biology* 16: 2813–2829.

Chirici, G., A. Barbati, P. Corona, M. Marchetti, D. Travaglini, F. Maselli, and R. Bertini. 2008. "Non-Parametric and Parametric Methods Using Satellite Images for Estimating Growing Stock Volume in Alpine and Mediterranean Forest Ecosystems." *Remote Sensing of Environment* 112: 2686–2700.

Crist, E. P., and R. C. Cicone. 1984. "A Physically-Based Transformation of Thematic Mapper Data – The TM Tasseled Cap." *IEEE Transactions on Geoscience and Remote Sensing* GE-22: 256–263.

Diamantopoulou, M. J. 2005. "Artificial Neural Networks as an Alternative Tool in Pine Bark Volume Estimation." *Computers and Electronics in Agriculture* 48: 235–244.

Donoghue, D. N., P. J. Watt, N. J. Cox, and J. Wilson. 2007. "Remote Sensing of Species Mixtures in Conifer Plantations Using LiDAR Height and Intensity Data." *Remote Sensing of Environment* 110: 509–522.

FAO. 2007. "Brief on National Forest Inventory China." MAR-SFM Working paper 16/2007.

Fazakas, Z., M. Nilsson, and H. Olsson. 1999. "Regional Forest Biomass and Wood Volume Estimation Using Satellite Data and Ancillary Data." *Agricultural and Forest Meteorology* 98: 417–425.

Francioso, L., A. Forleo, A. Taurino, P. Siciliano, L. Lorenzelli, V. Guarnieri, A. Adami, and G. Agnusdei. 2008. "Linear Temperature Microhotplate Gas Sensor Array for Automotive Cabin Air Quality Monitoring." *Sensors and Actuators B: Chemical* 134: 660–665.

Gemmell, F. 1995. "Effects of Forest Cover, Terrain, and Scale on Timber Volume Estimation with Thematic Mapper Data in a Rocky Mountain Site." *Remote Sensing of Environment* 51: 291–305.
Gu, H. Y., L. M. Dai, G. Wu, D. Xu, S. Z. Wang, and H. Wang. 2006. "Estimation of Forest Volumes by Integrating Landsat TM Imagery and Forest Inventory Data." *Science in China Series E: Technological Sciences* 49: 54–62.
He, Q. S., C. X. Cao, E. X. Chen, G.-Q. Sun, F.-L. Ling, Y. Pang, H. Zhang, W.-J. Ni, M. Xu, Z.-Y. Li, and X.-W. Li. 2012. "Forest Stand Biomass Estimation Using ALOS PALSAR Data Based on LiDAR-Derived Prior Knowledge in the Qilian Mountain, Western China." *International Journal of Remote Sensing* 33: 710–729.
Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2003. "National-Scale Biomass Estimators for United States Tree Species." *Forest Science* 49: 12–35.
Jiang, W., J. W. Tao, and L. L. Wang. 2006. "A Novel Palmprint Recognition Algorithm Based on PCA&FLD." Proceeding of International Conference on Digital Telecommunications, Cote d'Azur, August 28–31.
Jordan, C. F. 1969. "Derivation of Leaf-Area Index from Quality of Light on the Forest Floor." *Ecology* 50: 663–666.
Katila, M., and E. Tomppo. 2001. "Selecting Estimation Parameters for the Finnish Multisource National Forest Inventory." *Remote Sensing of Environment* 76: 16–32.
Kaufman, Y. J., and D. Tanre. 1992. "Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS." *IEEE Transactions on Geoscience and Remote Sensing* 30: 261–270.
Kayitakire, F., C. Hamel, and P. Defourny. 2006. "Retrieving Forest Structure Variables Based on Image Texture Analysis and IKONOS-2 Imagery." *Remote Sensing of Environment* 102: 390–401.
Kilkki, P., and R. Päivinen. 1987. *Reference Sample Plots to Combine Field Measurements and Satellite Data in Forest Inventory* 210–215. Research Notes, 19. Helsinki: Department of Forest Mensuration and Management, University of Helsinki.
Lévesque, J., and D. J. King. 2003. "Spatial Analysis of Radiometric Fractions from High-Resolution Multispectral Imagery for Modelling Individual Tree Crown and Forest Canopy Structure and Health." *Remote Sensing of Environment* 84: 589–602.
Mäkelä, H., and A. Pekkarinen. 2004. "Estimation of Forest Stand Volumes by Landsat TM Imagery and Stand-Level Field-Inventory Data." *Forest Ecology and Management* 196: 245–255.
Moskal, L., and S. Franklin. 2004. "Relationship Between Airborne Multispectral Image Texture and Aspen Defoliation." *International Journal of Remote Sensing* 25: 2701–2711.
Næsset, E. 2002. "Predicting Forest Stand Characteristics with Airborne Scanning Laser Using a Practical Two-Stage Procedure and Field Data." *Remote Sensing of Environment* 80: 88–99.
Nelson, R., W. Krabill, and G. MacLean. 1984. "Determining Forest Canopy Characteristics Using Airborne Laser Data." *Remote Sensing of Environment* 15: 201–212.
Nelson, R., W. Krabill, and J. Tonelli. 1988. "Estimating Forest Biomass and Volume Using Airborne Laser Data." *Remote Sensing of Environment* 24: 247–267.
Shvidenko, A., D. Schepaschenko, S. Nilsson, and Y. Bouloui. 2007. "Semi-Empirical Models for Assessing Biological Productivity of Northern Eurasian Forests." *Ecological Modelling* 204: 163–179.
Somogyi, Z., M. Teobaldelli, S. Federici, G. Matteucci, V. Pagliari, G. Grassi, and G. Seufert. 2008. "Allometric Biomass and Carbon Factors Database." *IForest-Biogeosciences and Forestry* 1: 107–113.
Tomppo, E., M. Nilsson, M. Rosengren, P. Aalto, and P. Kennedy. 2002. "Simultaneous Use of Landsat-TM and IRS-1c WiFS Data in Estimating Large Area Tree Stem Volume and Aboveground Biomass." *Remote Sensing of Environment* 82: 156–171.
Yang, Y. Q., Y. M. Wu, P. Huang, and C. M. Hou. 2003. "Estimated Stand Based on RS and GIS." *Journal of South China Agricultural University* 24: 73–77.
Zheng, S., X. Zhao, H. Zhang, Q. S. He, C. X. Cao, and L. F. Chen. 2011. "Atmospheric Correction on CCD Data of HJ-1 Satellite and Analysis of Its Effect." *Journal of Remote Sensing* 15: 709–721.