PLOT-LEVEL FOREST VOLUME ESTIMATION USING QUICK BIRD SATELLITE DATA IN THE HYRCANIAN FOREST

Salam Yazdani*[1], Shaban Shataee[2], Manouchehr Babanejad[3]
*[1]*MSc. Student of forestry, Forestry department, Gorgan University of Agricultural sciences and natural resources, Gorgan, Iran, Tel:+98171-2245882, yazdani2709k@yahoo.com*
[2]Forestry department, Gorgan University of Agricultural sciences and natural resources, Gorgan, Iran,
[3]Department of Statistic, Basic Sciences faculty, Golestan University, Gorgan, Iran

KEYWORDS: Forest quantitative, Quickbird, Hyrcanian forest, Texture analysis, linear and non-linear regression

Abstract
Awareness of the stand volume on forest is essential for management. The objective of study was estimation of forest stand volume using Quickbird data in the Hyrcanian forests of Iran. The images were orthorectified using DEM and ground control points. The proper processing analyses including principal component analysis, pansharp merging, band rationing and vegetation indices were done using main bands. The different texture analyses by different kernel sizes were also applied on the multispectral and panchromatic bands. By random cluster sampling method, 112 plots with size of 10*10 meters, the diameter at breast height (DBH) and height of some trees were measured in natural stand and hand transplant. The volume per hectare for each sample plot was calculated using DBH and height of trees. The positions of center of plots were registered by DGPS in a processing kinematics method. The plot based spectral values of main and processed bands were extracted. After normalizing the data, the genetic algorithm was used to select the more correlated bands among 227 main and processed bands. The results showed that the homogeneity, correlation, GLDV mean (equivalent to dissimilarity), variance and dissimilarity texture analyzed bands with sizes of 4*4 and 8*8 kernel sizes and homogeneity with kernel size of 20*20 on panchromatic band were more correlated with volume variable. Correlation analysis between volume and the best selected spectral bands were investigated by linear and non-linear regression methods. The performances of estimations of best models with highest R^2adj were examined using 20 percent of unused plots by relative RMSe and Bias measures. The results showed that linear regression model could estimated stand volume by R^2adj=0.26, RMSE=68.04% and Bias=25.01%, but the best non-linear logarithm model could estimate the volume with R^2adj=0.31, RMSE=65.91% and Bias=20.77%. For improving the estimations, using of the non-parametric algorithms may be produce better result in a future work.

1. INTRODUCTION

Estimation of forest structure variables from remotely sensed data has been one of main interesting for scientists from the early days of acquiring the digital remote sensing. Among remote sensing data sources, the visual aerial photographs have extensively used in the mapping and inventory of forest stands in past decades. However, utility of these data is time-consuming, subjective and highly dependent on interpreter experience. Automated image analysis techniques provide a faster, alternative method of retrieving forest parameters from space-borne and air-borne digital imagery (Kayitakire, et al., 2006). Added to these problems, the fast changing nature of forest areas is another problem that require to repeatedly assessments at short time intervals (Mohammadi et al. 2010). Up-to-date information on forest resources and monitoring ongoing spatial processes of

forest landscape is great important for successful and sustainable management of forest resources (Mohammadi et al. 2010). Due to these reasons, other methods of estimating forest characteristics for larger areas such as digital remote sensing are often used in early decades. Estimation and prediction of forest attributes are often done trough correlation analysis between forest field-based attributes and spectral information of air born or space born remote sensing data trough parametric and nonparametric regression analyses. Parametric liner and non-linear regression methods are the most used statistical techniques for modeling forest characteristics because of its easy use and straightforward interpretability (Aertsen et al. 2010; Curt et al. 2001). Several studies have focused on using medium resolution data sources such as Landsat TM/ETM+ and SPOT-HRV data to estimate forest inventory variables, i.e. stand volume, basal area, mean height, density and cover type (Woodcock et al. 1997; Franco-Lopez et al. 2001; Pax et.al., 2001). Mohammadi et al., (2010) modeled forest stand volume and tree density using Landsat ETM$^+$ data in the Iranian's Hyrcanian forests, with adjusted R^2 (43% and 73.4%) and RMSE of 97.49 m3/ha and 170.13 n/ha), respectively. Kayitakire et al. (2006) retrieved forest structure variables based on image texture analysis and IKONOS-2 imagery in eastern Belgium. Their estimations were based on texture features that were derived from the grey-level co-occurrence matrix (GLCM). The coefficients of determination, R^2, of the best model ranged from 0.76 to 0.82 for top height, circumference, stand density and age variables. Basel area was found to be weakly correlated to texture variable (R^2=0.35). Sivanpillai et al., (2006) analyzed the relationship between Landsat ETM+ reflectance value and stand characteristics of commercially managed loblolly pine *(pinus teade L.*) in east Texas and could predict the stand age and tree density with R^2=0.78 and R^2=0.60, respectively. Huiyan et al., (2006) investigated estimation of forest volumes by integrating Landsat TM imagery and forest inventory data in East China by *K*nn. The estimation error (RMSE) of total trees was 44.2%, but in species level, the RMSE for Larix stands was 51.7%, and for the Korean pine and broad leave stands were over 71.7% and 88.19%, respectively.
This paper describes a study of the relationship between Quickbird data and forest stand volume in the Shastkolate forest, Golestan Province, north of Iran. These paper the capability of liner and non-liner analyses for estimation of the volume in the Hyrcanian forest, northern of Iran. The aim of is to compare and evaluate one statistical liner and non-liner regression for modeling stand volume in the Hyrcanian forest, northern of Iran.

2. METHODS

2.1. Study area

The study area is located in the east north of Iran, eastern part of the Golestan province (Figure 1), comprising about 1714 hectares, extending from 36° 43´to 36° 48´ N latitudes and 54° 21´ to 54° 24´ E longitudes. The elevation ranges between 220 to 1100 m above mean sea level. The main tree species are *Parrotia Persica*, *Carpinus betulus, Acer cappadocicum, Cerasus avium, fagus orientalis, diospyros lotus,* and *Quercus castaneafolia.*

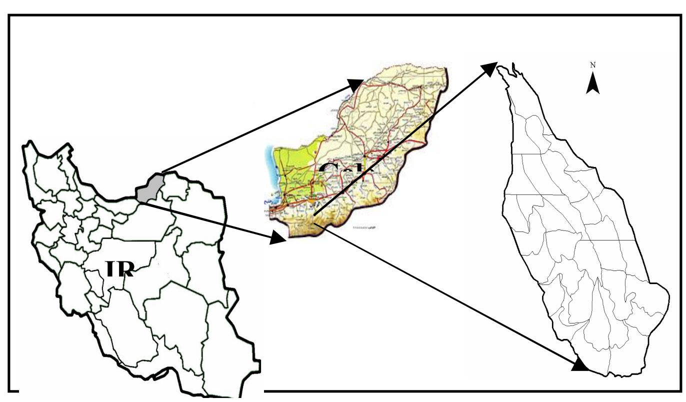

Figure 1: Location of the study area in the Golestan Province of Iran

2.2. Field data

In summer 2010, the filed information was gathered trough random cluster sampling method, and in 23 clusters, 112 plots with 10×10 meters size area. The center coordinates of plots were recorded using Differential Global Positioning System (DGPS) devices by post processing kinematics method. In each plot, kind of tree species, height of some trees, diameter at breast height (DBH) were measured for all trees with DBH greater than 7.5 cm. The plot level volume was computed using a local volume table, containing diameter at breast height ($d_{1.3}$) and height, to estimation volume of different species in plots.

2.3. Image processing

The Quickbird multispectral and panchromatic images acquired on 7 October 2007 were used in this study. The images were georeferenced and orthorectified using 24 ground control points (GCPs) collected by (DGPS) receiver and a digital elevation model (DEM). The root mean squared error (RMSE) for the panchromatic image were obtained about 0.76 and 0.67 m for the X and Y coordinates, whereas, the RMSE for the multispectral image were obtained about 0.97 and 1 pixels for the X and Y coordinates. The geometric precision of image was also verified using road vector layer and some unused field collected DGPS control points. After geometric rectification, relevant vegetation indices were generated by arithmetic computations. The tasseled cap transformation was also applied to generate the brightness, greenness and wetness components by applying coefficients on the spectral bands (Table 1). The method is widely used in vegetation mapping and monitoring applications (Mohammadi et al., 2010). In addition, pan sharpening fusion method has employed for merging the spectral bands with panchromatic band. We also performed principal component analysis on main bands to produce principal components. In addition, the texture analysis with different kernel sizes of 4×4, 8×8 and 12×12 pixel on main bands and sizes 20×20, 40×40 and 60×60 pixel on Panchromatic band were applied to produce texture bands. Finally, 227 main and pseudo bands were created. Averaged reflectance values of all main and processed bands were extracted on the plot area.

Table 1: The TCT coefficients for Quickbird multispectral bands (Yarbrough & Easson, 2005).

	A1	A2	A3	A4
Brightness	0.319	0.542	0.490	0.604
Greenness	-0.121	-0.331	-0.517	0.780
Wetness	0.652	0.375	-0.639	0.163

2.4. Statistical analyses

The Kolmogronov-Smirnov test was used to determine data normality. After verifying normality, in order to find the best-correlated independent bands with volume variable, the genetic algorithm was applied between all independent variables (original bands, vegetation indices, fusion bands, tasseled cap and PCA components and texture bands). Among 227 independent variables, the genetic algorithm was selected the 20 best independent variables to model stand volume. The parametric liner and non-liner regressions analyses by best subset regression method were used to describe relationship between volume as dependent and Quickbird main and processed bands as independent variables on 85 percent of plots in the modeling processes. The best subset regression analysis identifies the best fitting regression model that can be constructed with the predictor variables. Liner and non-liner regressions analysis selected a subset of independent variables that explains most of the variability in the dependent variable. Independent variables of the final model were selected based on a combination of both their individual contribution to the model, adjusted coefficient of determination ($R^2_{adj.}$) and residual mean square error (MSe) (Rawling et al., 1998).

2.5. Model validation

The validity of performances were examined using regression diagnostics metrics i.e. root mean square error (RMSe), relative RMSe, bias and relative bias, and using the independent and unused 15 % plots (18 plots). In addition, some common graphical diagnostic tools (McRobert, 2009) were used to illustrate the quality of performances.

$$RMSe = \sqrt{\sum_{i=1}^{m}(est_i - obs_i)} \Big/ m \quad (1)$$

$$\text{Bias} = \sum_{i=1}^{m}(est_i - obs_i) \Big/ m \quad (3)$$

$$RMSe\% = \frac{\sqrt{\sum_{i=1}^{m}(est_i - obs_i)} \Big/ m}{\sum_{i=1}^{m}(obs_i)/m} *100 \quad (2)$$

$$\text{Bias}\% = \frac{\sum_{i=1}^{m}(est_i - obs_i) \Big/ m}{\sum_{i=1}^{m}(obs_i)/m} *100 \quad (4)$$

Where *est* is estimation values from implementation of algorithms in m validation samples, *obs* is observation values and m is number of validation samples

3. RESULTS

The results of normality test showed that all variables had a normal. The results of genetic algorithm showed that texture variables of homogeneity, correlation, GLDV mean (equivalent to dissimilarity), variance and dissimilarity with kernel sizes of 4×4, 8×8 pixel on main bands and variable homogeneity in size 20×20 pixel on panchromatic band were more correlated with volume. Linear and non-liner combinations of variables produced by texture analysis could better predict stand volume (Table 2). The results showed that estimation of stand volume in linear regression (R^2adj=0.26, PRMSE=68.04% and Bias=25.01%) and non-linear logarithm (R^2adj=0.31, PRMSE=65.91% and Bias=20.77%) hade best results.

Table 2: Overview of the predictor variables selected by the stand volume models developed with tow techniques.

Dependent variable	*regression*	*best Independent variables*	*kernel sizes*
Stand volume (m3/ha)	linear	Homogeneity, Variance, Dissimilarity	4*4
		Homogeneity	20*20
		Correlation, GLDV Mean	8*8
		GLDV Mean, Homogeneity	8*8
		Correlation, Dissimilarity	8*8
Stand volume (m3/ha)	non-linear logarithm	Homogeneity, Variance	4*4
		Dissimilarity	4*4
		Homogeneity	20*20
		Correlation, GLDV Mean	8*8
		GLDV Mean, Homogeneity	8*8
		Correlation, Dissimilarity	8*8
		Variance	12*12

A linear and non-liner combination of variables produced by texture analysis could better predicted stand volume (Table 3). The results showed that estimation of stand volume in linear regression (R^2adj=0.26, PRMSE=68.04% and Bias=25.01%) and non-linear logarithm (R^2adj=0.31, PRMSE=65.91% and Bias=20.77%) hade best results (Table 3).

Table 3: Overview of the predictor variables selected by the stand volume models developed with two techniques (liner and non-liner regression).

Dependent variable	*Independent variables*	*Coefficient*	*Constant*	*R^2_{adj}*	*PRMSE*	*PBias*
Stand volume (m3/ha)	Homogeneity	4256.561	325.691	0.26	68.04%	25.01%
	Dissimilarity	-61.01				
	Variance	16.75				
	Homogeneity	4839.85				
	Correlation	659.21				
	GLDV Mean	-16.75				
	GLDV Mean	-8.44				
	Homogeneity	1178.87				
	Correlation	329.85				
	Dissimilarity	-80.26				
Stand volume (m3/ha	Homogeneity	8.11	1.831	0.31	65.99%	20.77%
	Variance	0.013				
	Dissimilarity	- 0.031				
	Homogeneity	11.76				
	Correlation	1.437				
	GLDV Mean	-0.042				
	GLDV Mean	-0.021				
	Homogeneity	7.04				
	Correlation	1.021				
	Dissimilarity	0.30				
	Variance	0.22				

4. DISCUSSION AND CONCLUSION

In this study, linear and non-linear regression models were evaluated for predicting and modeling the plot level forest volume in the Hyrcanian forest of Iran. The results showed that the best model for volume estimates comprised by the finer kernel size texture analyzed bands. It can concluded that the texture analysis is more useful in forest attribute estimation on fine and high resolution data compared to other used analyzed bands (original bands, vegetation indices, fusion bands and principal components). It should also be noted that the best estimates of the forest volume were achieved when the texture analysis by a window or kernel size of 4×4, 8×8 pixels on main bands and 20×20 pixels on panchromatic band. The linear and non-linear regression models incorporated by variables of produced using texture analysis were modeled the volume with R^2adj=0.26 and R^2adj=0.31, respectively. The R^2 values were lower than values those obtained in some studies (Hall et al., 2006; Huiyan et al., 2006 and Pax et al., 2001). In addition, RMSE obtained in this study were higher than obtained when direct estimation used to predict stand volume compared to these studies (Jensen, 2004; Mohammadi et al. 2010; Pax et al., 2001, Rawling et al., 1997). The results showed that texture analysis features derived from Quickbird spatial resolution image are promising for estimating coniferous forest stand volume. For access to better results in estimation of plot level forest volume, use of other modeling and predicting methods such as using non-parametric methods and algorithm may be useful. In this study, the plot sizes were 100 square meters, however, considering and gathering tree information in the bigger field plot sizes may be improve the results.

References

[1] Aertsen W., Kint V., Orshoven J. V., Ozkan K, and Muys B., 2010. Comparison and ranking of different modeling techniques for prediction of site Index in Mediterranean mountain forests. Ecological Modeling, 221, 1119-1130.

[2] Curt T., Bouchaud M., Agrech G., 2001. Predicting site index of Douglas-fir plantation from ecological variables in the Massif Central area of France. Forest Ecology Management. 149:61-74

[3] Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbor method. Remote Sensing of Environment, 77(3), 251–274.

[4] Hall R.J., Skakun, R. S., Arsenault E. J., Case B.S., 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest Ecology Management .225, 378–390.

[5] Huiyan, G., Limin, D., Gang.W., Dong. X., Shunzhong.W. and Hui.W, 2006. Estimation of forest volumes by integrating Landsat TM imagery and forest inventory data. Science in China Series E: Technological Sciences, Vol. 49, 54-62.

[6] Jensen, J. R., 2004. Introductory Digital Image Processing, A Remote Sensing Perspective. Prentice-Hall, New Jersey.

[7] Kayitakire F., Hamal C and Defourny, P., 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. Remote Sensing of Environment. 102: 390–401.

[8] Lillesand T, Keifer R. w., Chipman J. w. (2000). Remotely sensing and image interoperation. Johan Wiley and Sons, New York.

[9] Makela, H., Pekkarinen, A., 2004. Estimation of forest stands volumes by Landsat TM imagery and stand-level field-inventory data. Forest Ecology Management. 196, 245–255.

[10] Mohammadi J. and Shataee Sh., Yaghmaee F. and Mahiny A. S., 2010. Modeling Forest Stand volume and Tree Density Using Landsat ETM+ Data. International Journal of Remote Sensing, 31(11): 2959-2975.

[11] Pax, L. M., Woodcock, C. E., Macomber, S. A., Gopal, S., & Song, C. (2001). Forest mapping with a generalized classifier and Landsat TM data. Remote Sensing of Environment, 77(3), 241–250.

[12] Rawling J. O., Sastry G. P., and Dickey D. A. (1998). Applied analysis. Springer Press, New York, Berlin and Heidelberg, 650p.

[13] Sivanpillai, R., Smith, C.T., Srinivasan., R, Messina, M. G.,2006. Estimation of managed loblolly pine stands age and density with Landsat ETM+ data. Forest Ecology Management, 223, 247-254.

[14] Woodcock, C. E., Collins, J. B., Jakabhazy, V. D., Li, X. W., Macomber, S. A., & Wu, Y. C. (1997). Inversion of the Li–Strahler canopy reflectance model for mapping forest structure. IEEE Transactions on Geosciences and Remote Sensing, 35(2), 405–414.

[15] Yarbrough L.D., Easson, G., 2005. Quickbird 2 tasseled cap transformation coefficients: A comparison of derivation methods. Pecora 16 “Global Priorities in Land Remote Sensing, Sioux Falls, South Dakota.