ESTIMATING THE FOREST STAND VOLUME AND BASAL AREA USING PLEIADES SPECTRAL AND AUXILIARY DATA

M. Zahriban Heasari, [1,*] A. Fallah, [2] Sh. Shataee, [3] S. Kalbi [4], H. Persson [5]

[1] Dept. of Forestry, Gorgan University of Agricultural Sciences and Natural Resources,
Iran - mozhgan.zahriban@gmail.com
[2] Dept. of Forestry, Sari University of Agricultural Sciences and Natural Resources,
Iran - fallaha2007@yahoo.com
[3] Dept. of Forestry, Gorgan University of Agricultural Sciences and Natural Resources, Iran
Shataee@yahoo.com
[4] Dept. of Forestry, Sari University of Agricultural Sciences and Natural Resource, Iran
s.kalbi63@gmail.com
[5] Dept. of Forest Resource Management- Forest remote sensing, Swedish University of Agricultural Sciences,
Henrik.Persson@slu.se

Commission VI, WG VI/4

KEY WORDS: Pleiades; Auxiliary data; Spectral data; nonparametric methods; Volume; Basal area

ABSTRACT:

The aim of this study was to evaluate the improvements of volume and basal area estimations, when spectral data from the Pleiades were complemented with auxiliary data. The study area was located in the Darabkola's forest of Sari, Iran. In-situ data were collected for 144 circular sample plots, with 17.84 m radius, which were distributed using a simple random sampling design. Tree information included diameter at breast height (DBH) of all trees within the sample plots, and the height of some trees. By using DBH and tree height, the volume and basal area per hectare was also computed for each plot. Geometric and radiometric corrections of spectral data were applied to the images. In addition, the auxiliary maps of slope, aspect, elevation, soil pH and texture (through ground sampling and interpolation), precipitation and temperature (through interpolation of climate stations) were prepared. Digital values corresponding to ground plots were extracted from spectral bands and auxiliary data and considered as independent variables while volume and basal area were selected as dependent variables. The forest modeling was carried out using a non-parametric method of random forest (RF), using 70% of the sample plots as training data. The results were validated using the remaining 30% sample plots. The results indicated that by using both spectral and auxiliary data, the RMSE was reduced by 5% compared to using only spectral data for volume modeling. The corresponding advantage of using both spectral and auxiliary data was 1% to 3% when basal area was modeled.

1. Introduction

Forests cover approximately one third of the earth's surface and are the most important component of global carbon storage (Dixon et al. 1994). In addition, forests have important effects on human life in various aspects such as welfare, prosperity and happiness. Moreover, forests with their economic, social, and environmental values can ensure human life and sustainable development (FAO 2010). Therefore, in managing and planning for these valuable sources, useful information about their quantitative and qualitative characteristics is required (Lu et al. 2004). Quantitative features of forest stand structures such as volume and basal area can be considered as important factors in sustainable forest management (Zahriban et al. 2015). These data are usually obtained through field work which are not cost-effective (Lu et al. 2004). Therefore, remote sensing data as a valuable tool for mapping and monitoring these resources have been proposed (Stone 2010). The spectral data of remote sensing have been used in various fields of forestry like estimating the structural characteristics of forests (Franklin 1986; Leboeuf et al. 2012; Shataee et al. 2012; Mohamadi et al, 2017). However, many studies have shown that spectral data alone cannot exactly estimate structural characteristics of forests and to achieve this purpose, auxiliary data are needed (Holmgren et al. 2000). Indicated that standard error for

estimating stem volume using spectral data of TM was 0.36, while this value was 0.017 using spectral data with auxiliary data (stand age and height). (Mohammadi and Shataee 2007) Revealed that integrating the spectral data of Landsat with auxiliary data (slope, aspect, and elevation) reduced errors and increased the coefficients of determination in structural characteristics (volume and basal area). In addition, a study by (Gebreslasie et al. 2009) has shown that using ASTER data together with auxiliary data (age and site index) increased the coefficient of determination in estimating volume, basal area, stem numbers per hectare and stand height. On the other hand, besides parametric methods, multivariable non-parametric methods like decision trees have been developed and used in recent times in estimating quantitative features of forest. These methods have some advantages like easy interpretation, flexibility in using other auxiliary data, do not need to comply with any distribution, do not remove any datum and being multivariable (McRoberts et al. 2007). Random forest (here after RF) is a tree based and non-parametric method consisting of a large number of classification and regression trees which is expanded from the regression tree model (Breiman 2001). This method, compared with other methods, has high flexibility and ability to analyze different statistical data such as regression, classification, survival and non-supervised learning (Cutler et al. 2007). Therefore, the present study was conducted to

evaluate volume and basal area of the forest using spectral data from the Pleiades satellite with site auxiliary data including soil, climate, and physiographic data using random forest algorithm. It is assumed that using the spectral data of the Pleiades satellite with the auxiliary data will improve structural characteristics and increase accuracy of estimates.

2.   Material and methods

2.1. Site description

The study was carried out in the Darabkola forest located in the Mazandaran province, Northern Iran (longitude 53˚ 27´ 9˚ E and latitude 36˚ 46´ 74˚ N) (Fig. 1). The study area is located in district I of the Darabkola forest, 2612 hectares in area (Anonymous 2004).
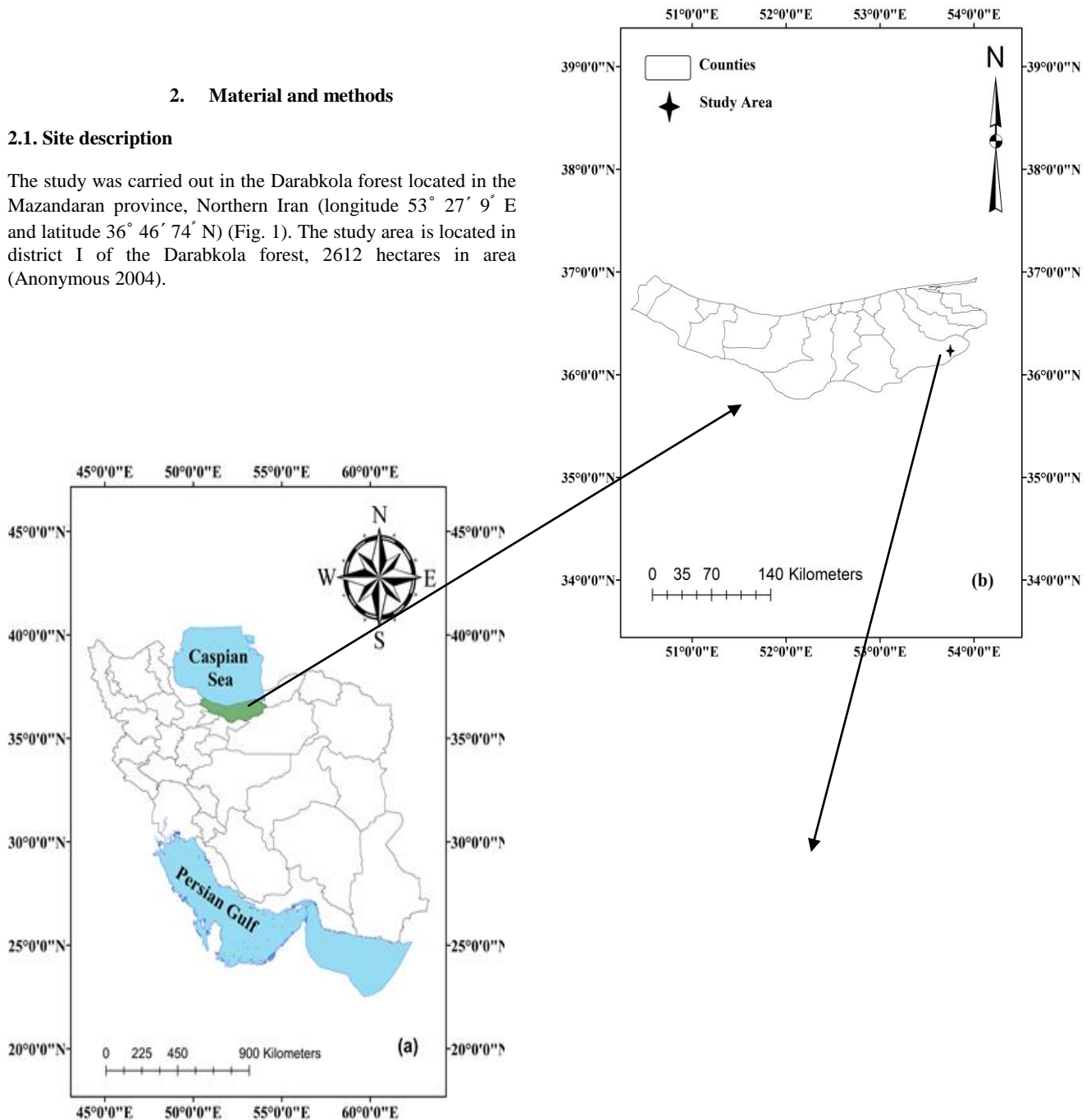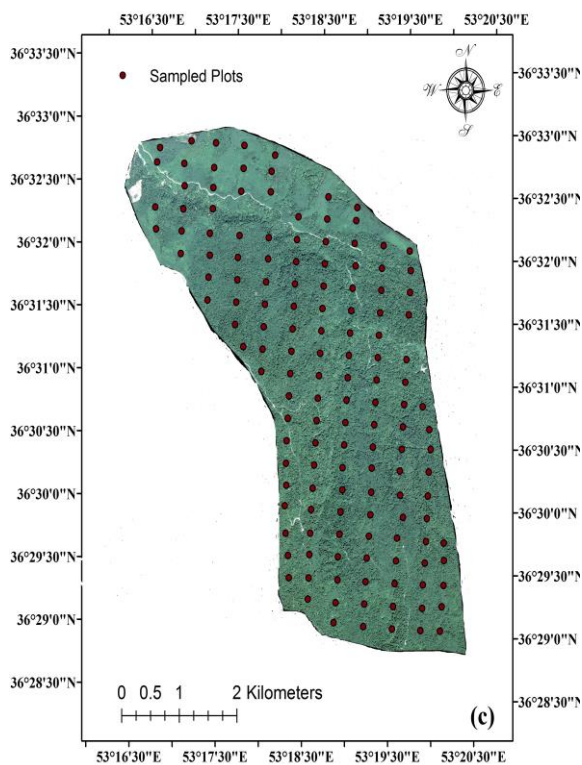

Figure 1 a Geographical locality of the research area in Iran; b Study area in Mazandaran Province; and c location of Sampled Plots on RGB color extracted from Pleiades satellite

2.2. Data used in the study
2.2.1 Ground data

The field information were sampled from 144 circular plots (10 acre each) on a systematic grid with a random starting location. The gridding was 330 m in the north direction and 500 m in the east direction (Fig. 1). In each plot, DBH of trees with more than 7.5 cm and height of two trees including the nearest tree to the center of the plot and the thick tree were recorded. The plot-based volume was calculated with the regional Tariff table, using diameter of tree species and basal area as measured DBH.

2.2.2. Topographic data

Digital topographic map of 1:25000 scale was used for analysing topographic data. These maps are in dgn format, in two-dimensional (2D) and (3D) modes. Their image system and coordinate grid used are UTM based on the 1984 European ellipse. In this research two map sheets NW 66621 and SW66632 were digitally used (2D and 3D). Digital elevation model (DEM) was prepared in GIS and elevation. Slope and aspect maps were extracted from the obtained DEM.

2.2.3. Soil data

According to topographic data and forest stand distribution, 50 soil samples (0-10 cm depth) were taken to determine soil pH and texture (Clay, silt, sand and silt and clay). The soil texture and pH maps were prepared using Kriging method in ArcGIS.

2.2.4. Climate data

The data of climate parameters including precipitation and temperature from 1991 to 2011 were used. Four and seven nearest stations were selected for temperature and precipitation, respectively. In order to produce temperature and precipitation maps, the correlation between elevation and precipitation in the study area was obtained (Eq 1-4). Then, Temperature and precipitation maps were prepared by correlation between elevation and climatic parameters in ArcGIS.

$$Y=0.302x+691.2 \qquad R^2 = 0.571 \qquad (1)$$

$$Y=0.069x+21.8 \qquad R^2 = 0.761 \qquad (2)$$

$$Y= -0.050x+13.07 \qquad R^2 = 0.480 \qquad (3)$$

$$Y= 0.053x+16.46 \qquad R^2 = 0.615 \qquad (4)$$

Where   x = digital elevation model
            Y = climate map

2.2.5. Satellite data

In this study, multispectral images of Pleiades satellite from the spring of 2013 were used including four spectral bands with a spatial resolution of 2 m and a panchromatic band providing a resolution of 0.5 m. The images have16-bit radiometric resolution.

2.3. Pre-processing and processing of satellite images

Geometric correction and orthorectification were applied on images before they were used in this study. Geometry correction was optimal by comparing with vector layers of roads. In addition, to reduce the illumination effects due to the slopes or uneven topography on the radiometric image, illumination was corrected by DEM in ERDAS 2013. In addition to the main bands, other processed bands with proper processing were created for the modeling process. Texture analysis is one of the most appropriate processing methods to estimate the characteristics of forest structure. Texture analysis is a function of spatial variations of pixels which expresses softness, roughness, smoothness and regularity of each level. In this study, the characteristics of texture analysis including homogeneity, c moment, GLDV of angular second moment, GLDV of entropy, GLDV of mean variance, GLDV of contrast, and inverse were used (Carr and Miranda 1998; Soh and Satsoulis. 1999; Solberg 1999) Regarding spatial resolution of the visible bands (2 m) and sample plot size (1000 m2), the 11 × 11 and 15 × 15 pixels for red and infrared bands (2 m) and the 41 × 41 and 61 × 61 pixels for panchromatic band (0.5 m) were used. In addition, to create plant indexes, the results of similar studies were applied (Hall et al. 2006; Khorami 2004; Mohamadi 2007; Gebreslasie et al. 2009; Yazdani 2011; Noorian 2013).

2.4. Extraction of spectral values and auxiliary data

Digital values corresponding to ground plots were extracted from the main and processed spectral bands and auxiliary data as independent variables for modeling.

2.5. Modeling
2.5.1. Random forest

The modeling of volume and basal area was conducted by spectral data and integration of spectral and auxiliary data. According to (Shataee et al. 2012) three parameters should be considered for creating a random forest. First, the proportions of sub-samples were considered as 50, 60, 70, 80, and 90 in this study. Secondly, in the modeling variable in each node of tree,

D%	D	RMSE%	RMSE	Estimates of K	Optimal tress
-0.46	-1.42	47.67	144.23	7	350
-1.26	-3.83	47.98	145/17	8	350
-0.093	-0.28	46.85	141.75	9	350
-0.50	-1.51	46.59	140.97	10	350
-0.74	-2.26	47.07	142.43	11	350

the number of variables in each node is the root square of total number of independent variables and subsequently the modeling was obtained to be ±2 of this value. In addition, optimal K was found by repetition of variables. Moreover, the third parameter was the number of trees which was 400 in this study.

2.6. The validation of models

In order to evaluate the performance of the models, a leave-one-out method was used (REF). The root means squared error RMSE (Eq. 3), relative RMSE (Eq. 4), mean difference (Eq. 5) and relative mean difference (Eq. 6).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

$$RMSE\% = \frac{RMSE}{\bar{y}} \times 100 \qquad (6)$$

$$\bar{D} = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \qquad (7)$$

$$\bar{D}\% = \frac{\bar{D}}{\bar{y}} \times 100 \qquad (8)$$

Where y_i = the observed and

$\hat{y}_i$ = the model- predicted value for plot

n = the number of plots

$\bar{y}$ = the mean of the observations

3. Results
3.1. Estimating the volume using random forest algorithm

The random forest algorithm is important to determine the optimal number of trees. Applying the initial 400 trees and their graph shows the changes in their RMSE in comparison with the number of evaluating and training samples to be optimal at 250 trees (Fig 2).
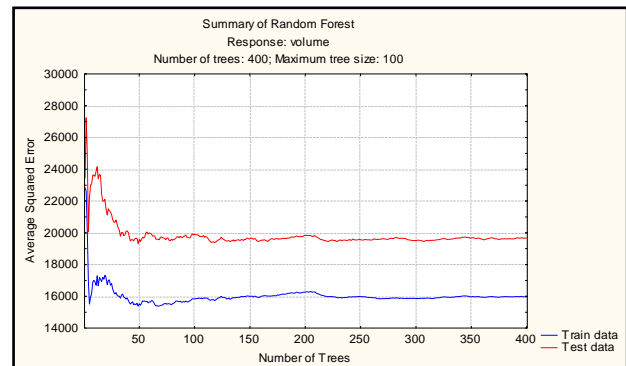

Figure. 2: average squared error changes in the number of trees for volume

Table. 1: The evaluation of the random forest to determine the most appropriate parameter for volume using the spectral data

Table 1 shows the results of the evaluation criteria of the random forest algorithm with the different number of Kin estimating volume using only spectral data. The results revealed that modeling with 10 variables in each node (k=10) had a minimum of RMSE and was selected as optimal K in estimating the volume using the Pleiades data.

Table.2: The evaluation of the random forest to determine the most appropriate parameter for volume using spectral and auxiliary data

D%	D	RMSE%	RMSE	Estimates of K	Optimal trees
-2.26	-6.85	42.73	129.28	8	250
-2.80	-8.47	41.87	126.70	9	250
-2.35	-7.11	41.58	125.81	10	250
-2.11	-6.39	41.02	124.11	11	250
-3.62	-10.97	41.81	126.52	12	250

Table 2 represents the evaluation criteria resulted from the random forest with different number of k in estimating volume using spectral and auxiliary data.

3.2 Estimating the basal area using random forest algorithm

The number of optimal trees is eminent to determine random forest algorithm. The results showed 250 trees were optimal (figure 3).
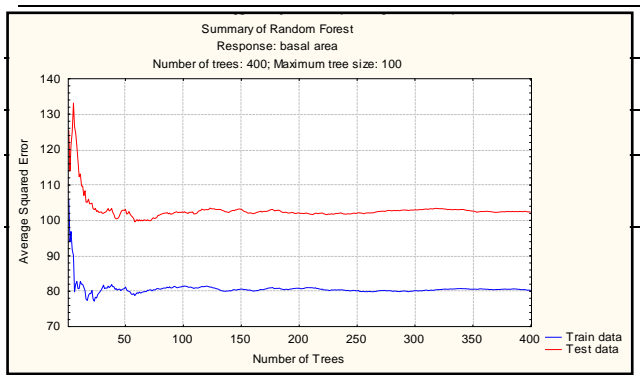
Figure. 3: average squared error changes in the number of trees for basal area

Table. 3: The evaluation of the random forest to determine the most appropriate parameter for basal area using spectral data

Table 3 shows the results of the evaluation criteria of the random forest algorithm with the different number of Kin

D%	D	RMSE%	RMSE	Estimate of K	Optimal trees
-1.35	-0.32	38.81	9.22	8	250
-1.50	-0.35	38.72	9.20	9	250
-1.80	-0.42	38.15	9.06	10	250
-0.79	-0.18	39.43	9.37	11	250
-1.12	-0.26	38.37	9.11	12	250

estimating basal area using only spectral data. The results revealed that modeling with 11 variables in each node (k=11) had a minimum of RMSE and was selected as optimal K in estimating the basal area using the Pleiades data.

Table.4: The evaluation of the random forest to determine the most appropriate parameter for basal area using spectral and auxiliary data.

Table 4 shows the evaluation criteria resulted from the random forest with different number of k in estimating basal area using

Volume (per hectare)				
Data	D%	D	RMSE%	RMSE
Spectral data	-0.50	-1.51	46.59	140.97
Spectral and auxiliary data	-2.11	-6.39	41.02	124.11

spectral and auxiliary data.

Table 5 represents applying spectral and auxiliary data in comparison with applying only spectral data in estimating both volume

Table 5 represents applying spectral and auxiliary data in comparison with applying only spectral data in estimating basal area

4. Conclusion

The present study was conducted to estimate volume and basal area using spectral data of Pleiades and auxiliary data including physiographic, soil and climate using random forest algorithm in the North of Iran.

The results have shown that RMSE and mean difference in estimating volume were 41.02 and 2.11%, and for basal area were 38.15 and -1.80, respectively. Integrating the spectral data with auxiliary data improved the modeling results of volume by

D%	D	RMSE%	RMSE	Estimates of K	Optimal trees
-0.37	-0.087	41.36	9.83	7	225
0.18	0.042	41.07	9.76	8	225
0.46	0.11	41.38	9.83	9	225
-0.60	-0.14	41.33	9.82	10	225
-0.66	-0.15	40.94	9.73	11	225

a 5% decrease in RMSE, and this value was 2.5% for basal area. The results of this study are consistent with those obtained by (Hasanzadeh-Naverdi et al. 2000; Kazemi et al. 2012; Valipour et al. 2013; Noshadi et al. 2014). They have shown that auxiliary data including precipitation, temperature, slope, and elevation affect volume and basal area. In addition (Gebreslasie et al. 2009; Mohammadi and Shataee 2007; Holmgren et al. 2000). Used auxiliary data (physiographic, stand age, site index, soil, and climate) with spectral data in order to improve their results. Their results showed this integration increases the determination coefficient and reduces RMSE which are consistent with the results of this study. The present study suggests that using auxiliary data with spectral data can improve the modeling results.

some possible future research would be (i) Using other auxiliary data such as soil age, habitat index, other soil parameters such as (calcium, phosphorus, potassium, carbon and nitrogen) ii) modeling volume and basal area using different sources of remote sensing data, e.g. unmanned aerial vehicle in combination, ALS and Radar data.

References

Anonymous. 2004. "The Revised proposal of Darabkola Forestry plan, the administration of natural resources," Mazandaran province. Iran

Breiman, L. 2001. Random forests. Machine Learning, 5–32.

Carr J. R., and Miranda F. P. 1998. The semivariogram in Comparison to the co-occurrence matrix for classification of Image texture. IEEE Transactions on Geosciences and Remote Sensing, 36 (6):1945-1952.

Cutler D. R., Edwards T C. Karen J. Beard H. Cutler A. Hess K.T. Gibson J. and Lawler J J. 2007. Random Forests for Classification in Ecology, Ecology, 88(11): 2783-2792.

Dixon R K., Brown S., Houghton R A., Solomon A M. Trexler M C. and Wisniewski J. 1994. Carbon Pools and flux global Forest ecosystems Science, 263: 185- 190.

FAO. 2010. Global Forest Resources Assessment 2010 – main Report. FAO Forestry Paper No. 163. Rome.

www.fao.org/docrep/013/i1757e/i1757e00.htm.

Franklin J. 1986. Thematic mapper analysis of coniferous forest structure and composition. *International Journal of Remote Sensing*, 7(10): 1287-1301.

Gebreslasie M T., Ahmed F B., Jan AN., and Adrdt V 2009. Predicting Forest Structural Using Ancillary Data and ASTER Satellite Data. *International Journal of Applied Earth Observation and Geoinformation*, 12S: 23S–26S.

Hall R J., Skakun R S., Arsenault E J., and Case B S. 2006. Modeling forest stand structure attributes using Landsat ETM+ Data: application to mapping of aboveground Biomass and stand volume. Forest Ecology and Management, 225, pp. 378–390.

Hasanzadeh-Naverdi A., namiranian M., marvi mohajer M. And azizi p. 2000. The effect of slope and elevation on volume Of natural beech forest in north of Iran. *Iranian Journal of Natural Resources*, 53(3): 201-215.

Holmgren J., Joyce S., Nilsson M., And Olsson H. 2000. Estimating Stem Volume and Basal Area in Forest Compartments by Combining Satellite Image Data with Field Data. *Scandinavian Journal of Forest Research*, 15(1): 103-111.

Kazemi M., Balapour sh., and rezaee f. 2012. Studying the Climatic parameters (temperature and rainfall) on annual Growth of poplar. *Iranian Journal of Wood and Forest Science And Technology*, 19(2): 143-152.

Khorami K R. 2004. Investigation of the potential of Landsat7 ETM+ data in Volume estimating of beech forest stands (case Study: Sangedeh area in north of Iran). M.Sc. Thesis, Faculty of Natural Resources, University of Tehran, Karaj, pp. 80.

Leboeuf A., Fournier R A., Luther J E., Beaudoin A., and Guindon L 2012. Forest Attribute Estimation of Northeastern Canadian Forests using QuickBird Imagery and a Shadow Fraction Method. Forest Ecology and Management, 266: 66-74.

Lu D., Mausel P., Brondizio E., and Moran E. 2004. Relationships between forests stand parameters and Landsat TM spectral response in the Brazilian Amazon Basin. Forest Ecology Management, 198:149-167.

Mc Robert R., Tomppo E. Finley A. and Heikkinen J. 2007. Estimating aerial means and variances of forest attributes using The k-nearest neighbors' technique and satellite imagery. Remote Sensing of Environment, 111, pp. 466–480.

Mohamadi J. 2007. Estimation of some characteristics of a Forest to explore the possibility of creating spatial prediction Models using satellite spectral data (Case study, oak Loveh). M.Sc. Thesis, Faculty of Natural Resources, The agriculture And natural resources university of Gorgon, Iran, pp.71.

Mohammadi J., and Shaban Sh. 2007. Forest stand density Mapping using Landsat-ETM+ data. Loveh forests, north of Iran, 28th Asian conferences of remote sensing, Malaysia, 12-16 November 2007, 10-27.

Mohammadi, J., Shataee, SH., Namiranian, M., & Næsset, E. 2017. Modeling biophysical properties of broad-leaved stands in the Hyrcanian forests of Iran using fused airborne laser scanner data and UltraCam-D images. *International Journal of Applied Earth Observation and Geoinformation*, 61, 32–45.

Noorian N. 2013. Comparing the images using Mediumresolution (Aster and TM) and high resolution (Quickbird) in estimating quantitative characteristics of forest, M.Sc. thesis, the agriculture and natural resources university of Gorgon, Iran, pp. 120.

Noshadi H., namiranian M., atarod p. and hosseinzade j. 2014. The effect of physiographic parameters on quantitative Characteristics of Quercus brantii in west of Iran. Iranian Journal of forest and wood products, 67(1): 73-84.

Shataee Sh. Kalbi S. and Fallah A. 2012. Forest attributes Imputation using machine-learning methods and ASTER data: Comparison of k-NN, SVR and random forest Regression Algorithms. International Journal of Remote Sensing, 33 (19): 6254-6280.

Soh, L. K., and C. Satsoulis. 1999. "Texture analysis of SAR Sea ice imagery using gray level co-occurrence matrices," IEEE Transaction on Geosciences and remote sensing, 37(2):780-795.

Solberg, A. H. S. 1999. "Contextual data fusion applied to Forest map revision," IEEE Transaction on Geosciences and Remote sensing, 37(3): 1234-1243.

Spies, T. A. 2010. "Forest structure: a key to the ecosystem," Northwest Sci. 72, 34–36.

Stone, R., 2010. "Earth-observation summit endorses global Data sharing," Science, 330, 902.

Valipour A., namiranian M., ghazanfari H., heshmat vaezin M., Jozef lekser M., and pliniger T. 2013. The relationship between Structural characteristic of oak forests and physiographic Parameters in Zagros forests, Iran. Iranian Journal of Forest and Poplar Research. 1: 30-47.

Yazdani S. 2011. Estimating some quantitative characteristics Of forest using Quickbird data. M.Sc. thesis, the agriculture and Natural resources University of Gorgon, Iran, pp. 129.

Zahriban M., Fallah A., Shataee Sh and Kalbi S. 2015 Estimating Quantitative forest attributes using Pleiades Satellite data and non-parametric algorithms in Darabkola Forests, Mazandaran. Iranian Journal of Forest and Poplar Research Vol. 23 No. 3.