MAPPING OF CONIFEROUS FORESTS' STRUCTURAL ATTRIBUTES IN RILA MOUNTAIN, BULGARIA BY SATELLITE DATA

Petar Dimitrov[1]

[1] Space Research and Technology Institute, Bulgarian Academy of Sciences, **Bulgaria**

ABSTRACT – *Information about forest structure is needed in many ecology and wildlife habitat studies. Remote sensing is currently being evaluated by the science community as a source of such type of information. In the paper spectral data from SPOT5 and image texture data from QuickBird satellite sensors are utilized for predicting three forest attributes: dominant height, dominant diameter, and canopy cover in coniferous forests. The developed multiple regression models for the three attributes have RMSE of 3.1 m (14.2%), 7.4 cm (19.4%), and 10.0% (12.8%) respectively. These results are comparable with previous studies and show significant potential for remote sensing of at least some parameters of the structure of coniferous forests in this region. The regression models and satellite images were used for generation of raster surfaces, which can be used in GIS analyses and for mapping.*

Keywords: *forest structure, multispectral images, image texture, regression analysis, Rila Mountain*

INTRODUCTION

Most of the forest territories in the European countries are intensively managed and exhibit changes in comparison with their primary natural state. They are, however, among the most important ecosystems sustaining wildlife [1], which is increasingly acknowledged and accentuated in their management goals. The structure of the forest is a key characteristic which may affect the diversity and abundance of wildlife species. As it is connected with the properties of the habitat forest structure is considered in many ecological and habitat studies. These studies show that the habitat selection and use by particular species or group of species can be related with the quantity of different structure variables like tree height, density, canopy cover, amount of under storey plants, etc. [2, 3, 4, 5, 6]. It has been shown that forest structure affects not only the fauna but also the floristic diversity in the herbaceous layer [7]. Findings of the relationships between species and forest structure help to develop managemet regimes that protect wildlife [8, 4].

Spatially explicit data describing different aspects of forest structure are required for studying and protecting wildlife in forest areas including for such tools like habitat models [9] and monitoring. Geographic information systems (GIS) and remote sensing (RS) play significant role in accomplishing such tasks providing researchers and managers with environmental map layers [10]. Remote sensing is considered promising data source for wildlife habitat analysis [11, 10]. Data from remote sensing instruments can be used in combination with ground-based data for deriving forest structure information over a territory.

The study objective is to investigate the possibility of estimation and mapping some structural attributes in the coniferous forests of Rila Mountain by the combined use of spectral and textural information from satellite images.

MATERIALS AND METHODS

For this study a test area was chosen (Fig. 1) on the northern side of Rila Mountain (SW Bulgaria) characterized by several types of coniferous forest [12] composed mostly by Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.) and less Silver fir (*Abies alba* Mill.) and Macedonian pine (*Pinus peuce* Griseb.). The field data were collected from 32 plots, which location was chosen based on existing forest inventory maps and satellite images so as to cover the diversity of forest structural types in the region. The plots were with varying size (from 5×5 to 30×30m) depending on the age and density of the stand. At each plot the species and diameter at breast height (DBH) were recorded for each tree, and in most plots the height of trees was also measured. The missing heights were restored from the diameter of the stem using height curves calculated from the existing measurements from the plots. One to four vertical photographs of forest canopy were taken per plot using digital camera and vario-lens set at 30° field-of-view on the short side of the frame. The photographs were used to assess the canopy cover. Coordinates in the centre of each plot were measured by GPS. Based on the gathered field data three forest structure attributes were calculated: 1) dominant diameter (D_d) – the mean diameter of the five thickest stems in the plot; 2) dominant height (H_d) – the mean height of these same thickest stems in the plot; 3) canopy cover (CC) – the proportion of the forest floor covered by the vertical projection of the tree crowns.

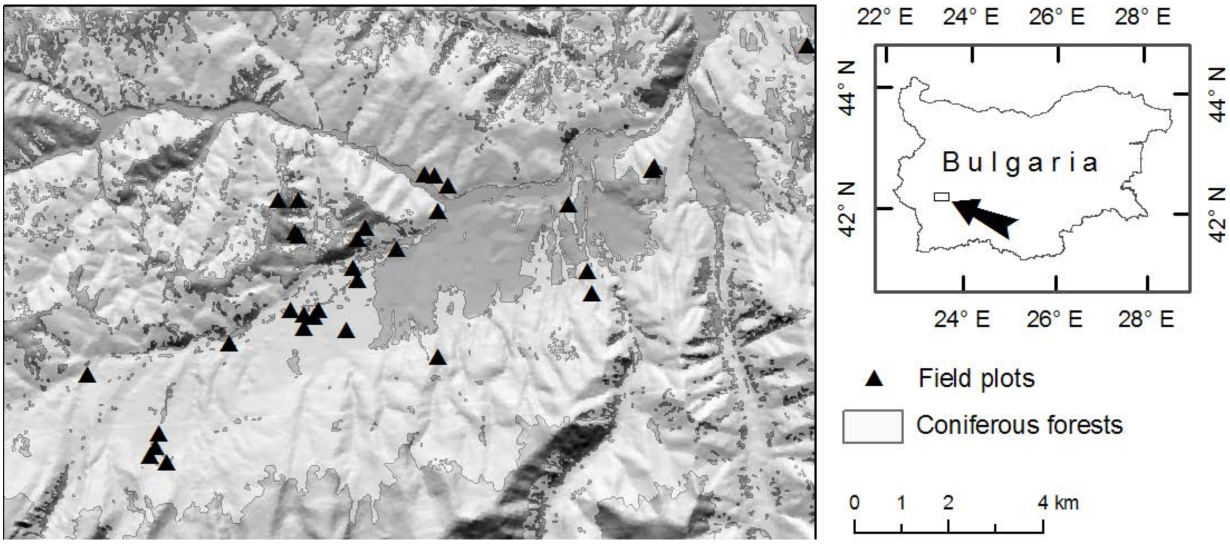


Fig.1 Location of the study area and distribution of the coniferous forests and field plots within its borders.

Two satellite images and their derivatives were used as a complementary remote sensing data – medium spatial resolution SPOT5 *HRG* image and high resolution QuickBird image. The SPOT 5 multispectral image (14.07.2008) has pixel size of 10m and four spectral bands: Green (0.49 – 0.60 μm), Red (0.61 – 0.68 μm), Near Infrared (NIR) (0.78 – 0.89 μm) and Shortwave Infrared (SWIR) (1.54 – 1.75 μm). The image was orthorectified and corrected for differences in terrain illumination using the SCS+C method [13]. Four spectral vegetation indices were calculated from the SPOT5 bands: Normalized Difference Vegetation Index (NDVI = NIR – Red / NIR + Red), Simple Ratio (SR = NIR / Red), Normalized Difference Infrared Index (NDII = NIR – SWIR / NIR + SWIR), and Structural Index (SI = NIR / SWIR). The QuickBird multispectral image (16.08.2007) has pixel size of 2.4 m and four spectral bands: B1 (0.43 – 0.54 μm), B2 (0.46 – 0.62 μm), B3 (0.59 – 0.71), and B4 (0.71 – 0.91 μm). Four types of gray-level co-occurrence matrix (GLCM) texture filters were applied to each of the QuickBird spectral bands, namely: *variance*, *homogeneity*, *contrast*, and *dissimilarity*. Only the texture images not the original spectral bands from QuickBird were used in the further analysis.

The values of the satellite variables (SPOT5 bands radiance, SPOT5 vegetation indexes, and QuickBird texture measures) were extracted for each field plot using GIS and statistically analyzed together with the ground data. One field plot was outside the QuickBird image footprint. The Pearson's correlation coefficient (r) for each pair of forest structure attribute and satellite variable was used as guide for selecting predictors for the linear regression analysis. Two regression models were developed for each forest structure attribute: the first using the spectral variable (band or vegetation index) most strongly correlated with the attribute, and the second using a texture variable in addition to the spectral one. A base-ten-logarithm transformation of the data was applied when needed to account for nonlinearity. In the back transformation of y the error term ε was taken into account (ε=mean square of the error from the regression/2; [14]). The accuracy statistics:

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

were calculated by the leave-one-out cross validation method, where y_i is the ground-measured value at plot i, $\hat{y}_i$ is the predicted value for plot i using data from the rest of the plots, and n is the number of plots used. The relative counterparts of these statistics - $RMSE_r$ and $Bias_r$ - were calculated as percent from the mean value of the forest structure attribute measured on the ground.

RESULTS AND DISCUSSION

The dominant height and dominant diameter show similar pattern in its correlations with the spectral variables from SPOT5. They are significantly (p<0.05) correlated with all bands and indices with the strongest correlation observed with the NIR band (-0.88 and -0.85 respectively) (Fig. 2). The correlation between NIR band and dominant height differ significantly from this of the other bands and indices except NDVI and SR (Table 1). The correlation between NIR band and dominant diameter differ significantly from this of red band, SWIR band and NDII (p<0.05) and from this of green band and SI if slightly higher significance level is accepted (Table 1). Canopy cover shows strongest correlation with the SI, but the correlation coefficients with the other indices and NIR band are very similar (Fig.2, Table 1). However the maximal correlation coefficient is only moderate (r=0.65, p<0.05), which indicates that this structural attribute is weakly related with the spectral data. In fact the correlations with the green, red and SWIR bands are not significant. The vegetation indices outperform the individual spectral bands and it appears that they contain more information about canopy cover. Moreover, the two indices (SI and NDII) incorporating the SWIR band have highest correlation coefficients despite the fact that the SWIR band alone is not correlated with the canopy cover.

The results for the dominant height are in agreement with Gerylo et al. [15], who found that the height of boreal coniferous stands is most strongly correlated with the near infrared band (band 4, 0.76-0.90 μm) of Landsat *TM*. Furthermore, the observed relationship between height and Landsat spectral bands had been negative because taller trees cast larger shadows that can create a decrease in the overall spectral response value detected by the satellite sensor [15]. The same mechanism seems to control the relationship between SPOT5 bands and dominant height in this study. Similarly, higher dominant diameter of a stand is connected with lower radiance values. This is expected since the dominant height and dominant diameter are strongly correlated to each other (r=0.96, p<0.001). In general as stand matures the size of the trees (diameter of bole, height, etc.) increase and more shadows appear.

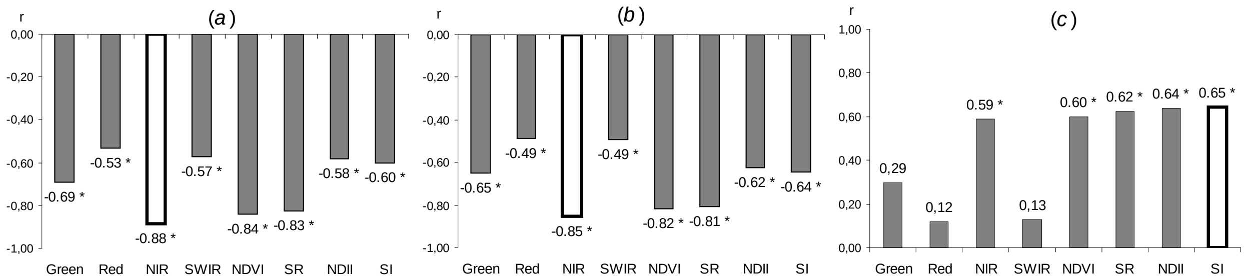


Fig. 2 Correlation coefficients of the spectral variables from SPOT5 and (a) the dominant height, (b) the dominant diameter and (c) the canopy cover. Values marked with asterisk (*) are statistically significant at the 0.05 level (n=32).

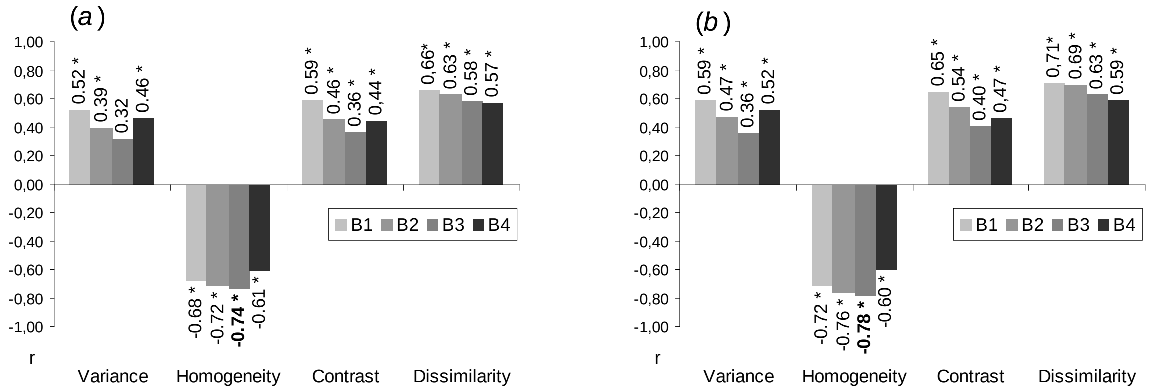
The relationship between satellite spectral data and canopy cover is not so straightforward in coniferous forests. In literature, examples of both positive and negative relationship of this forest attribute with reflectance/radiance can be found. In relatively open (prevailing canopy cover 30-60%) boreal forests in Canada Hall et al. [16] reported negative relationship with Landsat ETM+ spectral bands. The same is observed by Berterretche et al. [17] in forest with mean canopy cover of 39%. Even in the same area the sign of the relationship can be different depending on the species and its ecological specificity [15]. In the present study positive relationship between canopy cover and SPOT5 spectral bands and indices is observed with coniferous forests in the region having high closure (mainly 70-80%). The young naturally regenerated or artificial stands have maximal canopy cover. Their radiance is high because the canopy is denser and larger part of the crown/foliage is illuminated by the Sun. With the increase of the age forests appear darker as disused earlier and have generally lower canopy cover. However, the canopy cover is not so strongly related with the increasing shadowing in mature forests. One reason may be that some forests in the region are characterized with multilayer canopy structure, where shorter trees occupy gaps between taller trees. Thus, the canopy cover may retain high values even when great shadowing caused by the large tree crowns is present. It should be mentioned also that the influence of the understorey on the reflectance in the forests with low canopy cover is weak. In the study area these forests have tall trees that shade the bright understorey especially at higher Sun and satellite-sensor zenith angels.

Table 1 Results from the test of equality of correlation coefficients (H_0: $\rho_1 = \rho_2$). The band or index with the maximal correlation coefficient is compared with all other bands and indices for each forest structure parameter. The *p*-values of the test statistic Z are presented [18]. The bolded values indicate significant difference at the 0.05 level.

		NIR band			NIR band			SI
Dom. Height	Green band	**0.040**	Dom. Diameter	Green band	0.065	Can. cover	Green band	0.077
	Red band	**0.002**		Red band	**0.006**		Red band	**0.013**
	SWIR band	**0.005**		SWIR band	**0.006**		NIR band	0.727
	NDVI	0.522		NDVI	0.674		SWIR band	**0.015**
	SR	0.426		SR	0.598		NDVI	0.776
	NDII	**0.006**		NDII	**0.044**		SR	0.888
	SI	**0.008**		SI	0.061		NDII	0.978

Regardless of the spectral band used the *homogeneity* has consistently higher correlation coefficients with the forest structure attributes than the other three texture measures of the QuickBird image. The maximal correlations are observed with *homogeneity* of band B3 (the red band), with r=-0.74, -0.78, and 0.68 for dominant height, dominant diameter, and canopy cover respectively (Fig.3). Since the *homogeneity* measure is indicator of the smoothes of the texture it is negatively correlated with the structure attributes describing tree size and positively correlated with the canopy cover.

The pairs of regression models for the three forest structure attributes are presented in table 2. As predictors in the single regression models the NIR band and SI are used. As a second predictor in the multiple regression models the *variance* of band B4 (NIR) is chosen since its inclusion into the models results in maximal increase of the adjusted R^2, compared with the other texture measures. All models are significant (p<0.001); however the intercept of the single regression model for canopy cover is not statistically significant (p>0.05). It can be seen that the combination of spectral and texture variable provide for more accurate regression models than using only spectral variable. The degree of improvement is higher for the dominant diameter and dominant height, where decrease by 26%-30% in RMSE is observed and lower for the canopy cover, where RMSE decreases by 11%. The biases of all models are not significant and represent less than one percent of the mean value of the forest attributes (Table 2).



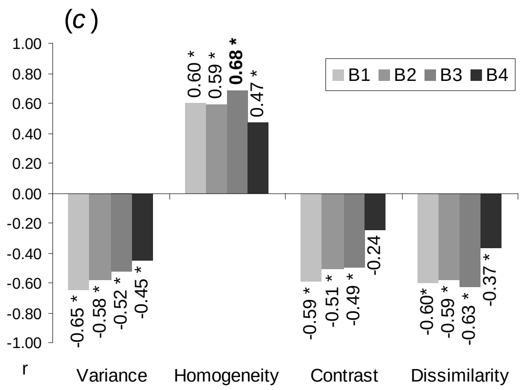


Fig. 3 Correlation coefficients of the texture measures from the four QuickBird bands and (a) the dominant height, (b) the dominant diameter and (c) the canopy cover. Values marked with asterisk (*) are statistically significant at the 0.05 level (n=31).

The NIR band of SPOT5 alone or in combination with the *variance* of QuickBird band B4 explains 79%-89% of the variance of the dominant height. Predictions using the multiple regression model are generally very close to the ground-measured values (Fig.4) and have RMSE of 3.1 m ($RMSE_r$=14.2%). These results are close to that reported by other authors. For example Kayitakire et al. [19] modeled tree height with RMSE of 2.2 m (10%) using IKONOS texture data, while Hall et al. [16] with RMSE of 2.8 m using Landsat ETM+ bands. Also Lefsky et al. [20] predict the maximal height by multiple regression with standard error (SE) of 38% using Landsat TM and with SE of 23% using an airborne laser altimetry system called SLICER. Therefore, the accuracy of the proposed models is comparable even with those obtained with more sophisticated Lidar sensors.

Table 2 Regression models

№		Adj. R^2	MSE	F	b	RMSE	$RMSE_r$	Bias	$Bias_r$	Bias sig.
1	$H_d=b_0+b_1*\log(NIR)$	0.791	-	118.1 ***	b_0 = 182.99 *** b_1 = -95.09 ***	4.4	20.4	0.02	0.1	0.980
2	$H_d=b_0+b_1*\log(NIR)+b_2*\log(V_{B4})$	0.886	-	117.9 ***	b_0 = 152.47 *** b_1 = -85.78 *** b_2 = 8.34 ***	3.1	14.2	0.02	0.1	0.977
3	$\log(D_d)=b_0+b_1*NIR$	0.750	0.022	94.1 ***	b_0 = 2.71 *** b_1 = -0.02 ***	10.0	27.0	0.4	1.0	0.841
4	$\log(D_d)=b_0+b_1*NIR+b_2*\log(V_{B4})$	0.905	0.007	143.3 ***	b_0 = 1.90 *** b_1 = -0.02 *** b_2 = 0.34 ***	7.4	19.4	0.01	0.0	0.992
5	Canopy cover=b_0+b_1*SI	0.397	-	21.4 ***	b_0 = 17.99 b_1 = 6.11 ***	11.2	14.3	0.04	-0.1	0.984
6	Canopy cover=$b_0+b_1*SI+b_2*V_{B4}$	0.544	-	18.9 ***	b_0 = 31.16 * b_1 = 5.58 *** b_2 = -0.11 **	10.0	12.8	-0.1	-0.1	0.964

***, **, and * indicate significance at the 0.001, 0.01, and 0.05 level respectively. V_{B4} is the *variance* texture measure derived from band B4 of QuickBird. The mean square of the error from the regression (MSE) is provided for the equations with transformed dependent variable.

Similarly, large part of the variance in the dominant diameter (75%-90%) is explained by the NIR band of SPOT5 and the *variance* of QuickBird band B4. Predictions using the multiple regression model correspond well to the ground-measured values especially for the smaller diameters (Fig.4). The RMSE is 7.4 cm (19.4%). In a previous study [20] multiple regression model for predicting mean DBH of dominant trees had SE=32% when Landsat TM data ware used and SE=19% when data from SLICER ware used.

The RMSE of the multiple regression model for canopy cover is 10.0 (percents canopy cover - pcc). For comparison Hall et al. [16] modeled the canopy cover with RMSE of 12.0 pcc using bands 3 (0.63-0.69 µm), 4 (0.75-0.90 µm) and 7 (2.09-2.35 µm) of Landsat ETM+, and Cohen et al. [21] with RMSE of 10.4 pcc using the same sensor. In this study the error is relatively low; however, the model has low R^2, which shows that only small part of canopy cover variations can be explained with the used satellite variables. On Fig.4 it can be seen also that field plots with canopy cover under 60% are overestimated and in general the correspondence with the ground-measured data is not very good. As the regression models do not permit extrapolations outside the range of the empirical data used for their development the proposed equations cannot provide reliable estimates for forests with canopy cover lower than 50%. Such condition is however not very common for Bulgarian coniferous forests except where certain types of felling are present.

This study may not reveal the full potential of image texture information since only four GLCM texture measures are tested using single set of parameters for the analysis: window size 3×3 pixels and one pixel offset on x and y direction. Thoroughly investigation of the contribution of the image spatial information to the forest structure prediction needs different input parameters to be tested in the GLCM analysis [19]. Their optimization may result in further improvement of the regression models.

When ground data from field plots are used, geolocation errors may cause that ground data were inaccurately linked with the satellite image data. However, this type of errors probably has minor influence on the proposed in this study models because of the good orthorectification of the images (RMSE=0.5 pixels for SPOT5 and RMSE=1.4 pixels for QuickBird assessed by 9 independent check points) and the homogeneous stands in which the plots are located.

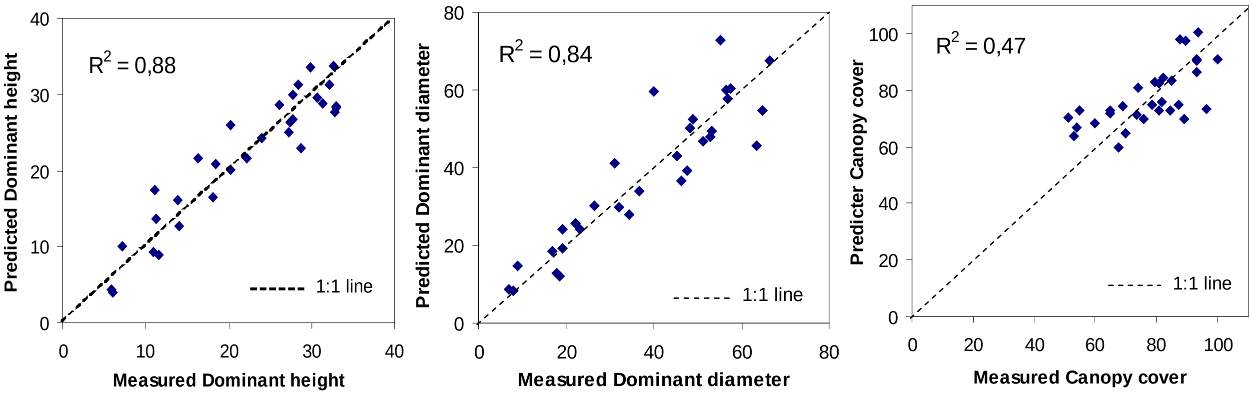


Fig. 4 Predicted values obtained by the leave-one-out cross validation against the ground measured values of the three forest structure attributes. Results are for the multiple regression models (№ 2,4, and 6 in Table2).

The three multiple regression models are used to generate raster surfaces characterizing the spatial distribution of dominant height, dominant diameter, and canopy cover. For this aim the *variance* image from QuickBird is first resampled to 2.5 m and then degraded by factor of 2 so as to gain the same resolution as the SPOT5 image (10 m). Mask obtained by unsupervised classification of the SPOT5 image is applied to assign zero value for all cover types different from coniferous forest. Areas covered by Mountain Pine (*Pinus mugo* Turra) are manually excluded. The generated raster surfaces can be used together with other data for different analysis in geographic information system (GIS). By reclassification of the raster values forest structure attributes can be presented in interval scale which meets particular needs and maps like these in Fig.5 can be prepared.

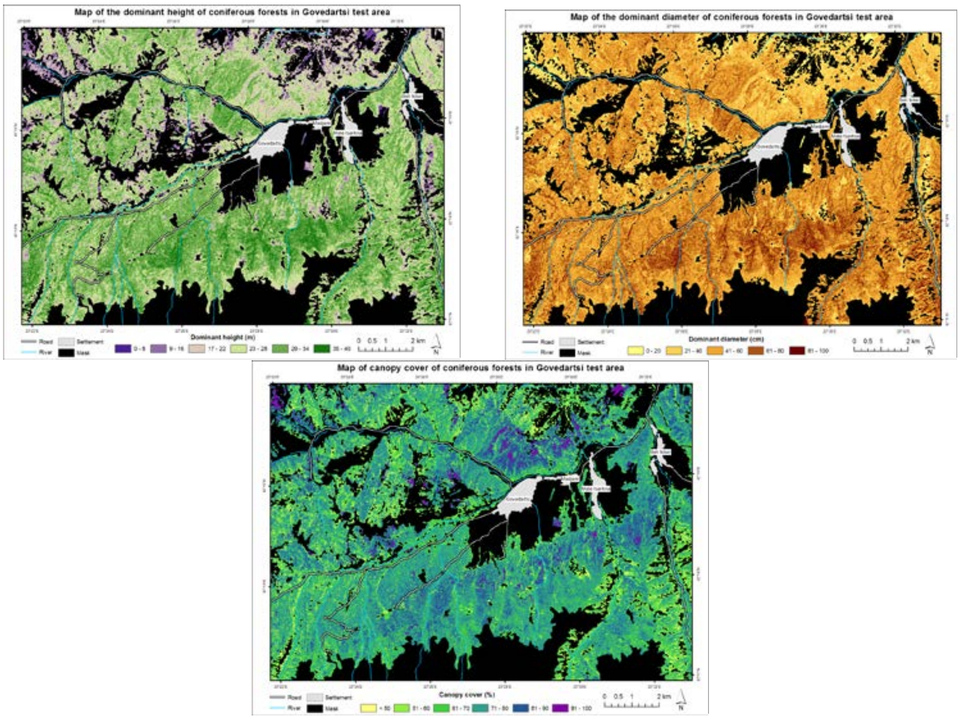


Fig. 5 Maps of dominant height, dominant diameter, and canopy cover.

CONCLUSION

This study shows that the dominant height, dominant diameter, and canopy cover of Rila Mountain's coniferous forests can be estimated with reasonable accuracy using satellite images. For the single regression models where only spectral information from SPOT5 is used $RMSE_r$ is between 14% and 27%, whereas for the multiple regression where additional texture information from QuickBird is utilized $RMSE_r$ is between 13% and 19%. Thus, it can be concluded that adding QuickBird texture variable is of considerable benefit to the development of more accurate models. The combined use of data from the two examined here satellite sensors is proved productive, however it is not always possible to have the two types of data. Therefore, applying spectral and textural information from single satellite sensor may be more practical alternative. For this aim for example SPOT5 can be used and from its panchromatic band to be derived textural information. This possibility will be analyzed in further studies.

Maps generated using this approach can be readily incorporated in GIS databases and used for natural resources management and wildlife conservation projects. Overlaying the satellite-generated maps of forest structure with spatial data from studies of avian or other forest species in the region will reveal if relationship between remotely sensed structural attributes and species distribution patterns exist. This is needed to assess the appropriateness of the maps for habitat studies.

ACKNOWLEDGMENTS

The satellite images used in the study were granted by SPOT Image and Digital Globe under the Planet Action Initiative. Part of the field work was funded by ESF with Contract No. BG051PO001/07/3.3-02/63/170608 between the SRTI-BAS and the Ministry of Education, Youth and Science of Republic of Bulgaria.

REFERENCES

[1] BirdLife European Forest Task Force (EFTF). Bulgarian-Romanian Forest Mapping Project. Final report. 69p. 2009.

[2] Bocca M., Carisio L., & Rolando A. Habitat use, home ranges and census techniques in the Black Woodpecker Dryocopus martius in the Alps. Ardea, vol. 95/1, pp 17–29, 2007.

[3] Nikolov S.C. Effect of stand age on bird communities in late-successional Macedonian pine forests in Bulgaria. Forest Ecology and Management, vol. 257, pp 580–587, 2009.

[4] Payer D.C. & Harrison D.J. Influence of forest structure on habitat use by American marten in an industrial forest. Forest Ecology and Management, vol. 179, pp 145–156, 2003.

[5] Heinze E., Boch S., Fischer M., Hessenmöller D., Klenk B., Müller J., Prati D., Ernst-Detlef Schulze, Seele C., Socher S., & Halle S. Habitat use of large ungulates in northeastern Germany in relation to forest management. Forest Ecology and Management, vol. 261, pp. 288–296, 2011.

[6] Åberg J., Swenson J.E., & Angelstam P. The habitat requirements of hazel grouse (*Bonasa bonasia*) in managed boreal forest and applicability of forest stand descriptions as a tool to identify suitable patches. Forest Ecology and Management, vol. 175, pp. 437-444, 2003.

[7] Battles J.J., Shlisky A.J., Barrett R.H., Heald R.C. & Allen-Diaz B.H. The effects of forest management on plant species diversity in s Sierran conifer forest. Forest Ecology and Management, vol. 146, pp. 211-222, 2001.

[8] Sturtevant B.R., Bissonette J.A. & Long J.N. Temporal and spatial dynamics of boreal forest structure in western Newfoundland: silvicultural implications for marten habitat management. *Forest Ecology and Management*, vol. 87, pp. 13-25, 1996.

[9] Kroll A.J. & Haufler J.B. Development and evaluation of habitat models at multiple spatial scales: A case study with the dusky flycatcher. Forest Ecology and Management, vol. 229, pp. 161–169, 2006.

[10] McDermid G.J., Hall R.J., Sanchez-Azofeifa G.A., Franklin S.E., Stenhouse G.B., Kobliuk T. & LeDrew E.F. Remote sensing and forest inventory for wildlife habitat assessment. Forest Ecology and Management, vol. 257, pp. 2262–2269, 2009.

[11] Hyde P., Dubayah R., Walker W., Blair J.B., Hofton M. & Hunsaker C. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. Remote Sensing of Environment, vol. 102, pp. 63–73, 2006.

[12] Petkov P., Penev N., Marinov M., Nedjalkov S., Naoumov Z., Garelkova D. & Antonov G. Forest types and organization of forestry in technical section Govedartsi of the Samokov Forestry. Gorskostopanska nauka, vol. 3/2, pp. 87-109, 1966. (in Bulgarian).

[13] Soenen S.A., Peddle D.R. & Coburn C.A. SCS+C: a modified Sun-canopy-sensor topographic correction in forested terrain. IEEE Transactions in geoscience and remote sensing, vol. 43/9, pp. 2148-2159, 2005.

[14] Newman M.C. Regression analysis of log-transformed data: statistical bias and its correction. Environmental Toxicology and Chemistry, vol. 12, pp. 1129-1133, 1993.

[15] Gerylo G.R., Hall R.J., Franklin S.E., Gooderham S. & Gallagher L. Modeling forest stand parameters from Landsat Thematic Mapper (TM) data. In: Proceedings of the 22nd Canadian Remote Sensing Symposium, August 21-25, 2000, Victoria, British Columbia. pp. 405-413, 2000.

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

[17] Berterretche M., Hudak A.T., Cohen W.B., Maiersperger T.K., Gower S.T. & Dungan J. Comparison of regression and geostatistical methods for mapping Leaf Area Index (LAI) with Landsat ETM+ data over a boreal forest. Remote Sensing of Environment, vol. 96, pp. 49–61, 2005.

[18] SYSTAT Software, Inc. MYSTAT: A Student Version of SYSTAT, v.12.02. Users Manual, 2007

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

[20] Lefsky M.A., Cohen W.B. & Spies T.A. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. Can. J. For. Res., vol. 31, pp. 78-87, 2001.

[21] Cohen W.B., Maiersperger T.K., Gower S.T. & Turner D.P. An improved strategy for regression of biophysical variables and Landsat ETM+ data. Remote Sensing of Environment, vol. 84, pp. 561–571, 2003.