Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

STUDYING THE RELATIONSHIP BETWEEN SOME ATTRIBUTES OF CONIFEROUS FORESTS AND SPECTRAL DATA FROM THE ASTER SATELLITE SENSOR

Petar Dimitrov, Eugenia Roumenina

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: petarkirilov@mail.bg

Abstract

The paper presents the results from a study aiming to assess the relationship between spectral data from the satellite sensor Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and five forest attributes measured in 29 plots in coniferous stands in the Rila Mountain. Biomass, volume and canopy cover showed moderate correlation with the radiance in the near-infrared band. Stand density was weakly correlated with spectral data, and basal area did not show statistically significant correlation. The relationship between ground-based and satellite data was modelled by ordinary least squares (OLS) and reduced major axis (RMA) regression. The models for predicting the different forest attributes had relative standard error of estimate between 14.4% (for canopy cover) and 51.8% (for volume).

1. Introduction

The information demand for different forest attribute data used to characterize productivity, structure and environmental functions of forests have increased in the recent years, along with the need for sustainable forest management. Remote sensing offers increasing variety of data types and the potential of these data for assessing forest characteristics is now extensively studied (Lefsky et al., 2001). Optical remote sensors are the most commonly used for forest research. Sensors like NOAA AVHRR are very useful to assess forest biomass at global scale (Dong et al., 2003). For studies at larger scale, the use of high-resolution data, such as Landsat, ASTER or

SPOT is needed. These sensors are suitable to map the highly fragmented forests in Bulgaria.

Forest attributes cannot be measured directly by remote sensing methods. However, they are related with stands' physiognomy and architecture, which in turn influence the reflectance characteristics captured by optical sensors (Franklin 2001). The reflectance in different spectral bands is correlated with ground-measured forest attributes, allowing their assessment and mapping by multispectral images. This empirical approach is widely used by the remote sensing community (Muukkonen and Heiskanen, 2005; Zheng et al., 2004; Hall et al., 2006).

To determine correlations, ground data at both plot level (Anaya et al., 2009; Healey et al., 2006) and stand level are used. In this work, data from a set of ground plots are used, because this allows gathering data for more forest attributes than these available in the stand datasets provided by forest authorities.

Coniferous forests in Bulgaria cover over 1.2 million hectares, which represents 31 % of the forest area (NSI, 2008). These forests are very important not only as a considerable timber resource, but also for their soil protecting, water regulating and other ecological functions. To characterize and manage them, information for many parameters is required. Dendrometry parameters describing the overstory layer, for example canopy cover, density, basal area, volume and biomass, are most commonly measured and used in practice. These parameters are related with many processes and forest functions; for example, canopy cover affects surface runoff and natural regeneration of stands (Raev, 1980; Stoyanova, 2006). Recently, increasing attention has been paid to the assessment of forest biomass for ecosystem productivity and carbon cycle studies (Houghton, 2005).

Information about forest attributes is obtained from the forest inventory, which provides accurate and relatively exhaustive data on a regular basis. In Bulgaria, forest inventories are carried out usually at ten-year intervals and data are gathered on stand level. Because of its flexibility (possibility for more frequent updates over arbitrary territory), remote sensing will be used in future more often as an additional source of forest information.

The objective of the study is to assess the relationship between some commonly used forest attributes and the spectral data from the Advanced

Spaceborne Thermal Emission and Reflection Radiometer (ASTER) instrument onboard the Terra satellite for coniferous forests.

The study area comprises part of the north-west section of the Rila Mountain and falls within the boundaries of the *Govedartsi* forestry section withn the *Borovets* State Forestry. The region features several types of coniferous forests (Petkov et al., 1966), comprising mainly Scots Pine (*Pinus sylvestris* L.) and Norway Spruce (*Picea abies* (L.) Karst.) and more rarely, Silver Fir (*Abies alba* Mill.) and Macedonian Pine (*Pinus peuce* Griseb.).

2. Materials and methods

2.1. Data

This study uses two types of data – ground-based and satellite. Ground-based data comprise measurements carried out in 29 test plots. The measurements were carried out in 2009 and 2010, whereas the test plots were distributed throughout the territory, so that stands of various age and biomass were comprised. Accounting for the inevitable errors during satellite images' geometric correction, the aspiration in specifying the test plots was that they fell among homogeneous stands, far from roads, cuttings or other features. Depending on the trees' age and density, the plots were sized between 5x5 m and 30x30 m and were commensurate with the ASTER pixel size. In each plot, the species and the diameter at breast height (dbh) of all trees higher than 2 meters were marked. The trees' height was measured by a heightmeter, whereas only in some plots, the heights were estimated by a stand height curve. The collected data were used to calculate the basal area (m^2/ha), the density (pcs/ha), the stem volume (m^3/ha) and the biomass (t/ha). The volume and the biomass were calculated by methods described by Beruchashvili and Zhuchkova (1997). In it, the stem volume is found as a function of the height, the breast diameter, and the stem form quotient. The biomass is calculated by multiplying the stem's volume by the density of the relevant tree species, adding to it the biomass of the branches, leaves and roots, determined as a percent of the stam's mass for the relevant tree species. Moreover, for each plot, the canopy cover was determined based on photos taken by a digital camera assembled on a tripod and levelled so that the optical axis was vertical (Cohen et al., 2003). The photos were taken by a standard zoom lens, adjusted so as to ensure a view angle of 30° along the frame's short side. On the smaller test plots, one photo was taken in the centre of the plot, and on the larger test plots – four nonoverlapping photos were taken. The canopy cover percent was determined by a round pallette (of diameter equal to the length of the frame's short side) imposed on the frame's centre. The statistical description of the data collected from the test plots is presented on *Table 1.*

The ASTER image used for the analysis dates from 1 October 2008 and features processing level 1A. Regretfully, in the images obtained after April 2007, the bands from the short-wave infrared range are unusable on account of a fault in the SWIR sensor. Therefore, the analysis uses only the three bands from the visible and near-infrared range (VNIR) (the thermal bands are not subject to this study). The three VNIR bands (band1-green (0.52-0.60 μm), band2-red (0.63-0.69 μm) and band3- near-infrared (0.76-0.86 μm)) were orthorectified in the ENVI software product as one HDF file. For the purpose, the SRTM digital elevation model (DEM) was used, with cell size of 90 m (USGS, 2006) and 9 GPS ground control points, identified on the red band. The orthorectification error was 2.2 pixels. ENVI applied automatically the calibration coefficients provided in the HDF file and converted the data into 32-bit radiance values (W/m^2/sr/μm). The image was resampled after the nearest neighbour method and had pixel size of 15 m.

Table 1. Statistical description of the data collected from the test plots

	Min	Max	Mean	St.dev.
Volume (m^3 ha^{-1})	81	983	480,0	287,1
Biomass (t ha^{-1})	41	518	252,3	150,9
Density (# ha^{-1})	175	19200	3078	5030,9
Basal area (m^2 ha^{-1})	23,2	84,4	49,4	16,3
Canopy cover (%)	51,1	96,4	76,9	13,9

2.2. Data processing and analysis

To assess the relationships between ground-measured parameters and satellite data, the radiance values from the image's three spectral bands for the relevant pixel in which each test plot falls were derived. Apart from the three spectral bands, to assess the forest attributes, two spectral indices were used as well: Normalized Difference Vegetation Index (NDVI = band3-band2/ band3+band2) and Simple Ratio (SR = band3/band2) (Tucker 1979).

The relationship between ground-based data and satellite variables was assessed through the correlation coefficient (*r*) (*Table 2*) and through graphs (*Figure 1*), after which for part of the forest attributes, regression models were developed. Models were developed only for the canopy cover, the biomass and the volume, since it was established that the relation of the basal area and stand density and the ASTER data is insignificant (*Figure 1*, *Table 2*).

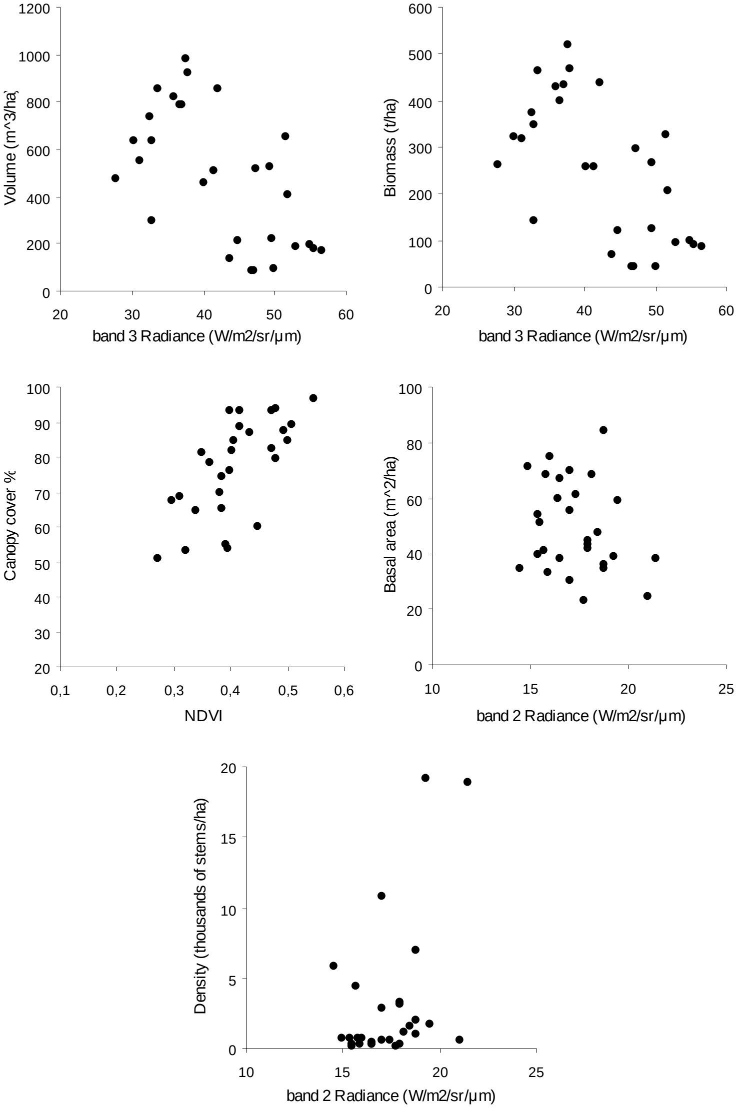

Fig. 1. Scatterplots of ground-measured forest attributes against the corresponding satellite variables, most strongly correlated with them.

To model statistically the relationships between forest attributes and spectral data, linear regression analysis was used since the established relationships were linear or close to the linear. Two types of regression were tested – the traditional ordinary least squares (OLS) method and the reduced major axis (RMA) method. The reduced major axis regression is used, when the measurement errors in the independent variable should be accounted for. Regression analysis by this method was made using a free-of-charge web-based application (Bohonak and Kim van der Linde, 2004).

To predict each forest attribute, the three spectral bands and spectral indices were tested consecutively as independent variables. In the final version, the band or index for which the coefficient of determination (r^2) was highest was chosen. With all forest attributes, the best predictor was the spectral band or index which is most strongly correlated with them. Addition of a second independent variable did not improve the models.

As predictors in the models, the third band (x_3) and NDVI (x_{NDVI}) were used. The volume and the biomass models have the following form:

(1) $$y = a + b * x_3 ,$$

and this for the canopy cover is:

(2) $$y = a + b * x_{NDVI} ,$$

where *a* and *b* are regression parameters.

2.3. Validation

To assess the accuracy of regression models, cross validation procedure was used. In it, the value of each test plot is predicted based on the observations from the other plots. This is necessitated by the small volume of the sample which does not allow part of it to be allocated for validation. The accuracy of the models was assessed by the standard error of estimate (*s*):

(4) $$s = \sqrt{\frac{\sum (y - y')^2}{(n-2)}} ,$$

where:

y is the actual value, *y′* is the predicted value, and *n* is the number of observation pairs. To allow for comparing the errors between the forest parameters, the standard error of estimate was also presented as a percentage of the parameter's mean actual value (s_r).

3. Results

The radiance in the near infrared band (band3) is most strongly correlated with the volume and the biomass measured on the plots, the correlation coefficients being accordingly -0.60 and -0.62 (*Table 2*). The strongest correlation with canopy cover is demonstrated by NDVI (r=0.67). The density and especially, the basal area, demonstrate very poor correlation with all satellite variables. In all cases, the correlation coefficients are not high which is also prompted by the scattered distribution of the graph points in *Figure 1*. A relatively clearer linear relationship is observed only between canopy cover and NDVI.

Table 2. Correlation coefficients of forest attributes and spectral data

	band1	band2	band3	NDVI	SR
Volume (m^3ha^{-1})	-0.467*	-0.451*	**-0.603*****	-0.453*	-0.480**
Biomass ($t\ ha^{-1}$)	-0.497**	-0.490**	**-0.618*****	-0.449*	-0.475**
Basal area (m^2ha^{-1})	-0.136	**-0.234**	-0.211	-0.072	-0.123
Density (#/ha)	0.389*	**0.432***	0.373*	0.245	0.224
Canopy cover (%)	0.247	0.109	0.553**	**0.667*****	0.655***

*p < .05 **p < .01 ***p < .001

Table 3 presents the parameters and the accuracy assessments of traditional regression models (ordinary least squares method) for the three forest attributes most strongly correlated with satellite data. Satellite spectral data explain only between 36% and 45% of the variations of the three forest attributes, the canopy cover model featuring the highest r^2, while timber volume coefficient of determination is lowest. The relative standard error of estimate varies between 13.7% and 48.6% for the different parameters. The cross validation results provide somewhat higher values for the relative standard error of estimate.

Table 4 presents the regression results after the RMA method. Again, the canopy cover model features the smallest error, and the stem volume model features the greatest error. During the validation, as well as during the modelling itself, s_r are higher with this type of regression, than with the ordinary least squares regression.

Table 3. Results from the ordinary least squares (OLS) regression

	Model				Validation	
	Parameters*	r^2	s	s_r	s	s_r
Volume (m^3ha^{-1})	a =1342.1 b =-20.3	0.36	233.3	48.6	248.4	51.8
Biomass (t ha^{-1})	a =717.4 b =-10.9	0.38	120.8	47.9	128.6	51.0
Canopy cover (%)	a =21.5 (*p*=0.09) b =135.2	0.45	10.5	13.7	11.1	14.4

* The *p* values lower than 0,001 are not shown; *s*-standard error of estimate; s_r-relative standard error of estimate.

Table 4. Results from the reduced major axis (RMA) regression.

	Model				Validation	
	Parameters	r^2	s	s_r	s	s_r
Volume (m^3ha^{-1})	a =1911.0 b =-33.7	0.36	260.7	54.3	276.9	57.7
Biomass (t ha^{-1})	a =1004.0 b =-17.7	0.38	134.3	53.2	142.5	56.5
Canopy cover (%)	a =-6.1 b =202.5	0.45	11.5	15.0	12.1	15.8

4. Discussion and conclusion

4.1. Correlations

All considered characteristics of the stand except for the basal area are statistically significantly correlated with at least one of the three ASTER bands. Nevertheless, the values of *r* are low, since no strict linear relationship is observed (*Figure 1*). The near-infrared band appears to be the best volume and biomass predictor. On the overall, the spectral indices NDVI and SR manifest significantly lower correlations with forest attributes compared to the spectral bands themselves. Only canopy cover is strongly correlated with NDVI.

The volume and the biomass of the studied coniferous stands are negatively related to spectral bands and indices. With growth of the trees and increase of the forest’s age, the reflectance in the visible spectrum range decreases. The same holds for the reflectance in the near-infrared range, although green vegetation is a good reflector in this part of the spectrum. The reason for this tendency is that the reflectance in the near-infrared range is strongly affected by the shadows between the individual tree crowns, which take shape and become thicker with increase of the forest's age. Canopy cover is positively correlated with spectral bands. This contradicts previous studies (Hall et al., 2006), where negative correlation has been established. The discrepancy may be explained by the different structure and peculiarities of the forests in the examined regions. In this study, forests with great (70-80%) canopy cover prevail, while the stands examined by Hall et al. (2006) feature mostly canopy cover between 30 and 60%. Actually, the correlation established in this study represents only the interval for canopy cover between 50 and 100%. The physical explanation of the observed positive correlation is related with the fact that the greatest canopy cover is typical for young planted stands which look bright on the image. The forests with smaller canopy cover look darker, since these are mostly mature stands where the shadowing of the adjacent crowns and soil is strong.

4.2. Assessment of the regression models

To model the relationships between forest attributes and spectral characteristics, liner regression was used. The obtained standard errors of estimate are high, but close to those from previous studies (Hyyppä et al., 2000). Thus, for instance, the non-linear model used by Muukkonen and Heiskanen (2005) to predict volume features relative RMSE error of 44.8% (calculated using independent data), while in this study, the relative standard error of estimate is 51.8 % (*Table 3*). In the first case, however, the authors use data at stand level.

The cross validation of the biomass assessment model showed that the relative error is 51%. For comparison, the error with a similar study using data from Landsat TM is 47% (Lefsky et al., 2001). Canopy cover assessed using ASTER data has relative error of 14.4%. For comparison, Hall et al. (2006) model the canopy cover with relative RMSE error of 12%, using bands 3 (0.63-0.69 μm), 4 (0.75-0.90 μm) and 7 (2.09-2.35 μm) of Landsat ETM+.

4.3. Comparison of both regression methods

The accuracy of the models, discussed in the previous section, refers to the results from the ordinary least squares regression. The use of the reduced major axis (RMA) method results in greater standard errors which might be expected. Nevertheless, RMA regression is often used with satellite data to determine different forest characteristics (Healey et al., 2006). The advantage of this method is that, with it, the measured quantity variance is preserved during its modelling (Cohen et al., 2003). As may be seen from *Figure 2,* with the ordinary least squares method, the modelled biomass values feature smaller variance compared to the real ones. On the

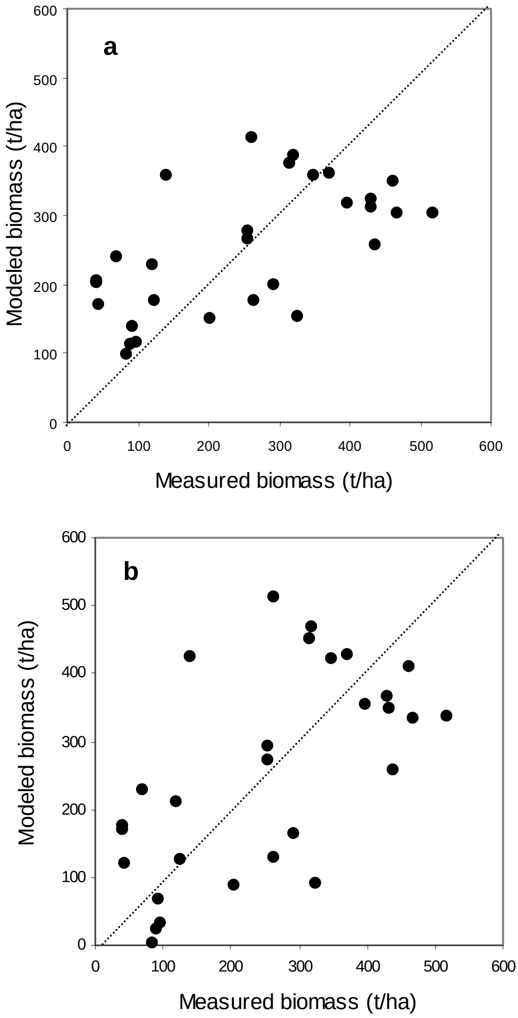

Fig. 2. Measured versus modelled biomass; (a) after the ordinary least squares method, and (b) after the reduced major axis method

other hand, with the reduced major axis method, the variation of the real values is preserved during modelling. Cohen et al. (2003) also point out that when there are errors in the measurement of both the independent and the dependent variable, it is not reasonable to use the ordinary least squares method. This is exactly the case when applying remote sensing to determine forest vegetation characteristics, since errors are observed both in ground-based data, as well as in the reflectance measured by the satellite sensors. Both methods have their advantages and disadvantages; therefore, they are used simultaneously in remote sensing.

***

The established correlations between the attributes of coniferous forests in the Rila Mountain and spectral data from the ASTER satellite sensor comply with the results under similar conditions in the boreal forests of Europe and North America. Most forest attributes are most strongly correlated with the radiance values in the near-infrared band. The absolute value of the maximal correlation coefficients for the various parameters varies between 0.23 (for the basal area) and 0.67 (for the canopy cover), while the relative standard errors of estimate lie in the interval between 14.4% (for the canopy cover) and 51.8% (for the volume). The results show that ASTER data may be used with the greatest fidelity to assess the canopy cover of coniferous forests in the region. The achievement of more accurate assessments, with admissible error levels for various applications, requires improvement of data geometric precision, and the use of greater number of spectral bands. Studies show that predictions accuracy increases when, besides spectral data, additional information is used, such as stand age map (Zheng et al., 2004) or canopy cover map. These possibilities will be examined in a future publication.

References

1. Anaya, J., E. Chuvieco, A. Palacios-Orueta. Aboveground biomass assessment in Colombia: A remote sensing approach. *Forest Ecology and Management*, 257, 2009, 1237–1246.
2. Beruchashvili, N., V. Zhuchkova. Methods for complex physical-geographical investigations: The manual. Moscow University Press, Moscow, 1997, 320 p. (In Russian).
3. Bohonak, A. J. & Kim van der Linde. 2004. RMA: Software for Reduced Major Axis regression, Java version. Website: http://www.kimvdlinde.com/professional/rma.html

4. Cohen, W. B., T. K. Maiersperger, S. T. Gower, D. P. Turner. An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment*, 84, 2003, 561–571.
5. Dong, J., R. Kaufmann, R. Myneni, C. Tucker, P. Kauppi, J. Liski, W. Buermann, V. Alexeyev, M. Hughes. Remote sensing estimates of boreal and temperate forest woody biomass: Carbon pools, sources, and sinks. *Remote Sensing of Environment*, 84, 2003, 393– 410
6. Franklin, S. E. Remote sensing for sustainable forest management. Lewis Publishers. 2001, 407 p.
7. Hall, R. J., R. S. Skakun, E. J. Arsenault, B. S. Case. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *Forest Ecology and Management*, 225, 2006, 378–390.
8. Healey, S. P., Z. Yang, W. B. Cohen, D. J. Pierce. Application of two regression-based methods to estimate the effects of partial harvest on forest structure using Landsat data. *Remote Sensing of Environment*, 101, 2006, 115–126
9. Houghton, R. A. Aboveground Forest Biomass and the Global Carbon Balance. *Global Change Biology*, 11, 2005, 945–958, doi: 10.1111/j.1365-2486.2005.00955.x.
10. Hyyppä, J., H. Hyyppä, M. Inkinen, M. Engdahl, S. Linko, Yi-Hong Zhu. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, 128, 2000, 109-120
11. Lefsky, M. A., W. B. Cohen, T. A. Spies. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Can. J. For. Res.*, 31, 2001, 78-87
12. Muukkonen, P., J. Heiskanen. Estimating biomass for boreal forests using ASTER satellite data combined with standwise forest inventory data. *Remote Sensing of Environment*, 99, 2005, 434 – 447.
13. NSI. 2008. Statistical yearbook 2008. National Statistical Institute, Republic of Bulgaria, Sofia.
14. Petkov, P, N. Penev, M. Marinov, S. Nedjalkov, Z. Naoumov, D. Garelkova, G. Antonov. Forest types and organization of forestry in technical section Govedartsi of the Samokov Forestry. *Gorskostopanska nauka (Forest Science)*, Vol. III, № 2, 1966, 87-109 (in Bulgarian)
15. Raev, Iv. Surface runoff in spruce stands undergoing active economic intervention. *Newsletter of the Bulgarian Geographical Society*, Vol. XVIII (XXVIII), 1980. (In Bulgarian).
16. Stoyanova, N. Natural regeneration of the forest ecosystems of *Picea abies* (L.) Karsten in the Rila Mountain. In: Ivan Raev (Ed.) Environment and structure of the Norway spruce forests in the Rila Mountain, Pensoft. Sofia-Moscow, 2006, 131-145 (in Bulgarian).
17. Tucker, C. J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment*, 8, 1979, 127-150.

18. USGS. 2006. Shuttle Radar Topography Mission, 3 Arc Second scene SRTM_ff03_p183r031, Filled Finished-A 2.0, Global Land Cover Facility, University of Maryland, College Park, Maryland, February 2000.

19. Zheng, D., J. Rademacher, J. Chen, T. Crow, M. Bresee, J. Le Moine, Soung-Ryoul Ryu. Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. *Remote Sensing of Environment*, 93, 2004, 402– 411.

ИЗСЛЕДВАНЕ НА ВРЪЗКАТА МЕЖДУ НЯКОИ ТАКСАЦИОННИ ПОКАЗАТЕЛИ НА ИГЛОЛИСТНИ НАСАЖДЕНИЯ И СПЕКТРАЛНИТЕ ДАННИ ОТ САТЕЛИТНИЯ СЕНЗОР ASTER

П. Димитров, Е. Руменина

Резюме

В статията са представени резултатите от проведено изследване за оценка на връзката между спектралните данни от сателитния сензор Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) и пет таксационни показателя, измерени в 29 пробни площадки в иглолистни насаждения в Рила планина. Биомасата, обемът и склопеността показват умерена корелация с яркостта в близкия инфрачервен канал. Гъстотата на дървостоя е слабо корелирана със спектралните данни, а кръговата площ не показва статистически значима корелация. Връзката между наземните и сателитните данни е моделирана с помощта на регресионен анализ по метода на най-малките квадрати и на редуцираната главна ос. Съставените регресионни модели за оценка на отделните таксационни параметри имат относителни стандартни грешки на оценките в интервала от 14,4 % (за склопеността) до 51,8 % (за обема).