Forestry An International Journal of Forest Research

Institute of Chartered Foresters

Forestry 2015; **88**, 96–107, doi:10.1093/forestry/cpu032
Advance Access publication 11 September 2014

Estimation of forest structural information using RapidEye satellite data

Adelheid Wallner*, Alata Elatawneh, Thomas Schneider and Thomas Knoke

Center of Life & Food Sciences Weihenstephan, Institute of Forest Management, Technische Universität München, Hans-Carl-von-Carlowitz-Platz 2, Freising-85354, Germany

*Corresponding author. E-mail: adelheid.wallner@tum.de

Received 3 December 2013

Forest management plans in Bavaria are generally updated only once every 10 years. However, the increasingly dynamic forest structure due to climatic changes requires more frequent data collection in order to maintain up-to-date information. This study explored the use of RapidEye satellite data to provide more frequent updates to the information database. Forest structural information such as quadratic mean diameter (dq), basal area (BA), stem number (SN) and volume (V) were estimated using multi-seasonal analysis of three RapidEye datasets from 2009. Spectral indices and textural metrics provided additional image feature layers. Forest inventory plots were stratified based on the forest type. A correlation analysis was conducted between terrestrial inventory data and that derived from RapidEye data. A cross-validated stepwise forward regression analysis was performed for each forest type. The coefficient of determination and relative root mean square error (rRMSE) showed that stratification improved the regression models, which obtained determination measures ranging from 0.37 to 0.63 and rRMSE ranging from 25 to 131 per cent. Biases of the regression estimates were small, hence the results obtained from applying the models were of an acceptable level of accuracy. The analysis confirmed the potential of RapidEye data to support forest management.

Introduction

Bavaria contains 2.56 million hectares of forest area, which amounts to approximately one-third of its total land surface (Schnell and Bauer, 2005). As a result of legal decisions and due to European Union regulations, forest owners are obliged to fulfil several different national and international requirements. Examples of these requirements include evaluating and monitoring of NATURA 2000 areas and compliance with the Flora-Fauna-Habitat Directive. Furthermore, due to climate change, foresters are faced with the challenge of establishing forests with high level of structural and biological diversity. In view of these challenges, Felbermeier *et al*. (2010) conducted a study in Bavaria to analyse the requirements of foresters that can potentially be met using remote sensing techniques. The results showed a strong demand for information about forest structure at the forest enterprise level such as timber volume (*V*); stem number (SN) and basal area (BA). This information is usually derived from periodic national and regional forest inventories, which are carried out in time steps of 10 years in Bavaria. Up-to-date information is thus difficult to obtain.

Several studies have looked at this phenomenon at the local or global scales, by investigating the option of updating forest databases with remote sensing data (Wolter *et al*., 2009; Ozdemir and Karnieli, 2011; Rahlf, 2011; Stepper and Schneider, 2012). For forest management in Bavaria, remote sensing systems which deliver information at the 1:10 000 mapping scale are generally considered most appropriate. This scale represents spatial resolutions from 2 to a maximum of 10 m. Forest structural attributes are mainly derived by evaluating spectral information. The approach described in Wolter *et al*. (2009) used 'Systeme Pour l'Observation de la Terre' (SPOT) 5 satellite data to estimate forest structural attributes such as diameter at breast height (DBH), tree height and BA, via partial least squares regression. WorldView-2 data have also been tested as a basis for deriving information on stand development, forest change analysis and estimation of structural attributes via regression analyses, as described in Ozdemir and Karnieli (2011).

To estimate timber *V*, SN and BA, spectral information has usually been combined with textural information. Coburn and Roberts (2004), Kayitakire *et al*. (2006), Wunderle *et al*. (2007) and Ozdemir and Karnieli (2011) showed that textural features based on image grey-level co-occurrence matrices (GLCMs; Haralick *et al*., 1973; Haralick, 1979) are helpful tools for remote sensing image analysis in the forest. Textural analysis can be used to identify patterns in images based on colour, contrast, shape, tone, size and shadow which may represent actual differences among features on the ground. Coburn and Roberts (2004) found that the *Angular Second Moment (ASM)*, *Entropy* and *Contrast* features were the most relevant textural features when analysing 4 m spatial resolution optical images for three different types of coniferous stand classification. Kayitakire *et al*. (2006) showed that the GLCMs *ASM*, *Contrast*, *Variance*, *Homogeneity*, *Correlation* and *Entropy* were effective for estimating forest structure attributes such as BA, height, age, DBH and stand density using IKONOS-2 data of 1 m resolution. Rahlf (2011) suggested similar features (*Mean*, *Variance*, *Homogeneity*, *Contrast* and *Dissimilarity*)

for height estimation using RapidEye data. No publications were available that specifically addressed the estimation of attributes such as *V*, quadratic mean diameter (dq), BA and SN using RapidEye data.

Indices such as Normalized Differenced Vegetation Index (NDVI), Ratio Index (RE), Simple Ratio Index (SR), Near-Infrared-Green Ratio (GR) and Green-Red-Vegetation Index (VI) have often been used to study forest structural information (Eckert, 2006; Wolter *et al*., 2009). The RE index was identified in a study by Schneider *et al*. (2012) as an effective spectral index for the detection of forest structure change. The benefits of using multi-seasonal images were confirmed by Maselli *et al*. (2005), who estimated BA using a multi-seasonal set of Landsat 7 ETM + scenes. Subsequently, Elatawneh *et al*. (2013) conducted a classification using up to eight datasets from two different years to conduct multi-temporal RapidEye image analysis for tree species discrimination. Given the high temporal frequency of available RapidEye data, these scenes seem very suitable for this approach.

The work presented here focused on forest structural attributes estimation at the stand level using data obtained from the RapidEye high-resolution remote sensing system. The high repetition rate of the RapidEye sensor (each nadir scene is theoretically recaptured every 5.5 days), combined with the large swath width of 77 km and the affordable price of 0.95€ km^{-2} (RapidEye AG, 2011) offers a wide range of new opportunities. The aim of the research described in this paper was to examine the potential for using RapidEye data to estimate dq, SN, BA and *V* via regression analysis to support forest inventories. In particular, we evaluated the possibilities for improving forest structural attribute estimation by stratifying inventory plots, as well as the use of multi-seasonal datasets to increase model accuracy.

Materials and methods

Study area

The study area is situated in the south-eastern part of Bavaria within the municipality of Traunstein (47°51′42″N; 12°39′20″E, 600–700 m above sea level) and consists of ~234 ha of forest which is owned by the city of Traunstein and managed by the Chair for Forest Growth and Yield at the Technische Universität München. Spruce (*Picea abies* (L.) H. Karsten) (43 per cent), beech (*Fagus sylvatica* L.) (22 per cent) and fir (*Abies alba* Mill.) (13 per cent) are the dominant tree species in the area. Other tree species at the test site include ash (*Fraxinus excelsior* L.) (6 per cent), maple (*Acer pseudoplatanus* L.) (5 per cent) and scattered individuals of various other species (11 per cent). Depending on the management treatment used in the stands and the soil conditions, the forest structure varies from even-aged pure stands to uneven-aged mixed stands. The geology at the research site is known as the Swabian-Bavarian young moraine and molasse foothills (Schnell and Bauer, 2005), a landscape dominated by a mixture of plateaus and valleys.

Dependent variables (forest variables)

The following variables obtained from the terrestrial inventory data from 2008 were used as dependent variables in the regression analysis: dq, BA, SN and *V*. The inventory was carried out following the methods used in Bavarian State-owned forests (FER, 2011). In the test site forest, the inventory plots were systematically arranged in a 100 × 100 m sample grid. Each plot consisted of three concentric circles within which various forest information was noted – an inner circle of 31 m^2 (3.15 m radius) in which all trees with a DBH smaller than 10 cm were measured, an intermediate circle of 125 m^2 (6.31 m radius) wherein all trees with a DBH ranging between 10 and 30 cm were recorded and an outer circle of 500 m^2 (12.62 m radius) in which all trees with a DBH > 30 cm were measured.

The inventory data from each of the 228 field plots were then allocated to one of four strata. The method used – stratification into pure and mixed plots of deciduous and coniferous trees – was based on the approach employed by Ďurský (2000), Heurich (2006), Latifi *et al*. (2012) and Straub *et al*. (2013). This approach, where forest plot data for the different forest types (strata) are analysed separately, has been shown to reduce the estimation error of dq, BA, SN and *V*. Only measurements from trees with a DBH of >7 cm were used to calculate dq, BA, SN and *V* (Kramer and Akça, 2008). Given that restriction, 12 field plots were excluded as they contained only young trees (DBH < 7 cm). Therefore, data from only 216 field plots out of 228 were used to calculate the aforementioned attributes. The dominant tree species were identified based on a calculation of the total basal area of each species of tree within each plot. All plots that contained a BA of <20 per cent of species other than the dominant species in that stand were considered 'pure plots'. Thus, 80 per cent was chosen as the threshold for assignment of a plot to the 'pure' stratum (Ďurský, 2000; Heurich, 2006). Accordingly, 55 plots were allocated to the 'pure coniferous' (pc) stratum, j_1 and 20 plots to the 'pure deciduous' (pd) stratum, j_2. The remaining plots were assigned either to stratum j_3 – 'deciduous dominant' (dd) (57 plots) or stratum j_4 – 'coniferous dominant' (cd) (84 plots). The stratified field plots were located in stands in different developmental stages. The summary statistics for the structural information for both the stratified and unstratified sample plots are listed in Table 1.

Table 1 Summary statistics for forest structural information of all forest inventory sample plots ($n = 216$) and for four strata of the same sample plots based on the species composition in the stand

Attributes	Stratum (j)	Number of sample plots (n)	Standard deviation	Min.	Mean	Max.
dq (cm)	Unstratified	216	13	8	33	63
	cd (j_4)	84	13	10	34	62
	pc (j_1)	55	13	8	36	61
	dd (j_3)	57	13	9	33	63
	pd (j_2)	20	15	8	26	54
SN (ha^{-1})	Unstratified	216	181	20	334	3364
	cd (j_4)	84	350	20	306	1943
	pc (j_1)	55	521	20	370	3364
	dd (j_3)	57	438	20	368	2505
	pd (j_2)	20	173	20	232	560
BA (m^2 ha^{-1})	Unstratified	216	14	1	23	67
	cd (j_4)	84	11	5	23	58
	pc (j_1)	55	17	1	31	67
	dd (j_3)	57	8	4	19	47
	pd (j_2)	20	9	1	9	34
V (m^3 ha^{-1})	Unstratified	216	181	2	259	931
	cd (j_4)	84	145	21	256	734
	pc (j_1)	55	236	2	363	931
	dd (j_3)	57	123	10	217	527
	pd (j_2)	20	138	2	106	496

pc (j_1) = pure coniferous stratum j_1; pd (j_2) = pure deciduous stratum j_2; dd (j_3) = deciduous dominant stratum j_3; cd (j_4) = coniferous dominant stratum j_4; dq = quadratic mean diameter (cm); SN = stem number (ha^{-1}); BA = basal area (m^2 ha^{-1}); V = volume (m^3 ha^{-1}).

Independent variables (remote sensing-based variables)

The RapidEye satellite system is a constellation of five satellites carrying identical sensors, all of which were launched at the end of 2008 (RapidEye AG, 2011). Each sensor collects data in five bands – blue (440–510 nm), green (520–590 nm), red (630–685 nm), Red Edge (690–30 nm) and near-infrared (NIR) (760–850 nm) – at a spatial resolution of 6.5 m which is resampled to 5 m during pre-processing. The data used in this study consisted of a set of 3A correction level (after radiometric, sensor and geometric correction, with a resampled pixel size of 5 m) RapidEye scenes acquired during the growing season in 2009. The first dataset (17 May) was considered to be representative of spring with the trees just reaching full leaf, while the second (1 August) was acquired during the peak of the growing season when the maximum photosynthetic activity takes place. The third scene (7 September) coincided approximately with the beginning of autumn when the leaves began to change colour. Thus, the three images were chosen in order to capture each of the most important seasonal phenological states of the trees located in the study area. Quality control of the geometric correction was conducted using ENVI® 4.3 (ENVI, 2005) and atmospheric correction was applied to the RapidEye data using the ATCOR 3 model, which was implemented in PCI-Geomatica (Richter, 2011).

The RapidEye image pixels were segmented into image objects using the eCognition® Developer 8.9 software package. A multi-resolution algorithm was chosen for this segmentation (Baatz and Schäpe, 2000), using the following settings: scale – 10, shape – 0.5 and compactness – 0.8. Based on their position and size (500 m^2 circular polygons) the inventory plots were used to create a segmentation boundary (Figure 1). The image objects contained within this segmentation boundary were used as training objects for extracting the textural and spectral features. Subsequently, the segmentation was used to classify the entire test site into broadleaf and coniferous forest. The test site was then subdivided into cells, each with a size of 15 × 15 m. These cells and the classification were used to stratify the area into pc, cd, pd and dd as the basis for the estimation of the forest structure attributes shown in Figure 1. This approach was used to regionalize the models to the test site.

In addition to the mean spectral values of the five RapidEye bands, the spectral values calculated from neighbourhood statistics (Skewness and Standard Deviation (Std.Dev.)) and various vegetation indices were also used for the study. In order to derive appropriate vegetation indices to analyse the effects of spectral reflectance and absorption in forest stands (Richardson and Everitt, 1992), the following vegetation indices were computed from the RapidEye data:

NDVI	(NIR − RED)/(NIR + RED)	Rouse *et al*. (1974)	(1)
RE	GREEN × RED EDGE/RED	Schneider *et al*. (2012)	(2)
SR	NIR/RED	Jordan (1969)	(3)
GR	NIR/GREEN	Lyon *et al*. (1998)	(4)
VI	GREEN/RED	Kanemasu (1974)	(5)
Brightness	(BLUE + GREEN + RED + RED EDGE + NIR)/5	Trimble Reference Book (2013)	(6)

Eight GLCMs – *Mean*, *Std.Dev.*, *Correlation*, *Dissimilarity*, *Entropy*, *ASM*, *Contrast and Homogeneity* – were selected for this study, as our review of the literature led us to determine that they were the most relevant for forest structure attribute estimation. The values for the different texture features were calculated within the image objects derived from the segmentation process as described above using the eCognition® Developer 8.9 software package. The formulas used for the GLCM texture features are described in the Trimble Reference Book (2013).

Statistical methods

The final dataset of independent variables used for the statistical analysis was made up of 120 texture features (derived from three datasets of five bands each), 18 spectral ratio features and 45 spectral features. These features were used to estimate the forest structure information dq, SN, BA and *V* (dependent variables) for comparison with the same attributes calculated using actual data that had been previously collected at 216 terrestrial field plots stratified into the four strata as described above. For each dependent variable, either a simple or a multiple regression model was established. The prediction ability of the model with and without stratification as well as the effect of seasonality was tested.

To analyse the relationships between the dependent and the independent variables, the Pearson's correlation coefficient (*r*) was calculated for each pair (*P*-value of <0.05 and 0.01; Table 2). Collinearity of the predictors was checked manually through examination of the bivariate correlations. The variance inflation factors (VIFs), as measures of multicollinearity, were analysed for each of the predictors.

The forest structural information was modelled as a function of the spectral and textural features derived from the RapidEye satellite images. A stepwise forward selection algorithm implemented with the SPSS command 'reg.' was chosen to determine the remote sensing-based independent variables most suitable for inclusion in the model. This selection process was used to determine the independent variables with the largest positive or negative correlations with the outcome (dependent) variables. A variable was retained in the model if it significantly improved the ability of the model to predict the outcome. Then, the next variable with a large semi-partial correlation with the outcome was considered. A variable was only retained if its influence on the dependent variable ($\hat{y}$) was found to be significant. The process was ended when there were no further statistically significant variables remaining. To avoid an artificial increase in the coefficient of determination (R^2), a restriction of four for the amount of independent variables and a cut-off value for VIF of less than four was required in order for an independent variable to remain in the model, according to Castillo-Santiago *et al*. (2010) and Ozdemir and Karnieli (2011). The basic model function of a multiple regression can be written as follows:

$$\hat{y}_i = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_i x_i \quad (1)$$

where $\hat{y}_i$ is the predicted value for the dependent variable (in this case, the forest variable to be estimated), x_i are the independent variables (here, derived from the remote sensing data) $i = 1,\ldots,m$ and b_i represents the coefficients to be estimated $i = 0,\ldots, m$.

As a final step, we checked the assumptions of the model for linearity and normality. The standardized residuals of the different models were first visually assessed for normality using a histogram with density plot and normal probability plot of the residuals. We also analysed the dispersion and shape of the standardized residuals for homogenous variances when plotted against the standardized predicted values. The quality of the model as a whole was checked by evaluating the *F*-ratio from the output of the analysis of the variance (ANOVA). The influence of each of the independent variables in the models was checked via *t* test (significant $P < 0.05$).

Accuracy assessment

The prediction error of the selected models was estimated by the leave-one-out cross-validation method (LOOCV), a special case of a k-time cross-validation (Latifi *et al*., 2010; Straub *et al*., 2010). LOOCV was conducted for each regression model in the four strata (j_1,j_2,j_3,j_4) as well as for the models without stratification. Owing to the relatively low number of observations in each of the strata ($j_1 = 55$, $j_2 = 20$, $j_3 = 57$, $j_4 = 84$), the LOOCV approach was preferred over *k*-fold cross-validation. The predicted value for the k_jth observation (inventory plot *k* from the stratum *j*) was calculated using the regression equation obtained by fitting the model without the k_jth observation. The method starts with all observations and proceeds by eliminating one observation and using the model to predict the dependent variable for the observation that was eliminated. This process was repeated until

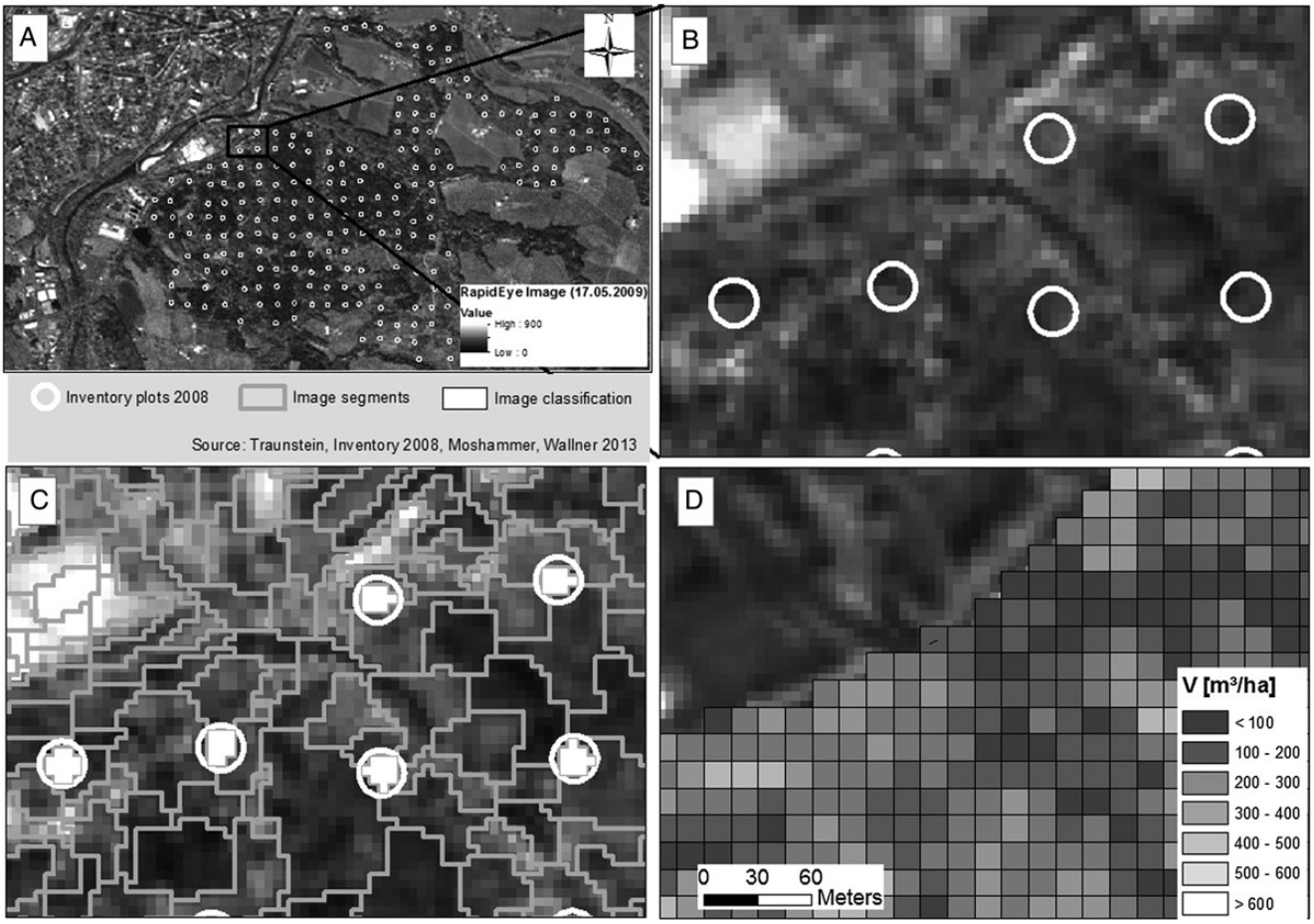

Figure 1 Processing steps used for spectral value extraction. (A) Test site with inventory grid. (B) Inventory plot size and position. (C) Segmentation into objects. (D) Mapping of *V* predictions on a grid of 15 × 15 m spatial resolution.

each observation had been used once for validation ($k_j - 1$) (Stone, 1974). The performance was then evaluated by means of the root mean square error (RMSE) and the relative RMSE (rRMSE) (Witten *et al*., 2011). The bias was calculated using the following formula:

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

where y_i is the observed variable for plot j, $\hat{y}_i$ is the predicted variable for plot j, $\bar{y}_i$ is the mean of the observed variable for plot j, and n is the number of sample plots used for validation. The relative bias as a diagnostic measure was calculated as follows:

$$\text{Bias} = [\%] = \frac{\text{Bias}}{\bar{y}} \times 100 \quad (3)$$

The accuracy of the regionalization of each of the models was checked using a test of means. A mean value from the modelled data for a particular forest attribute (e.g. *V*) was calculated by selecting 216 cells at random from the regionalized data and calculating the mean of those values. This process was repeated 1000 times per model in order to create a distribution of means from the regionalized data from each model for each attribute. These means were then compared with the mean values for each of these forest attributes that had been derived from the terrestrial inventory data.

Results

Estimation accuracy (Regression models)

A summary of the basic information obtained from the analyses described above is given in Table 1. Analysis of the correlation coefficients revealed that only a small number of the extracted spectral features were significantly correlated with the forest variables as shown in Table 2. The vegetation indices NDVI and VI showed no significant correlation with the forest structure variables. Moreover, a number of variables were not chosen for inclusion in the models because of multicollinearity.

The image features showing high correlation to the dependent variables were used as independent variables in the regression analyses. The results of the different models and the variables selected are given in Table 3. The values of the *t* test (Table 3) showed that all independent variables had a significant impact on the dependent variables. In Table 4, the calculated values are summarized and the strength of the relationships between the models and the dependent variables (multiple correlation coefficients, coefficient of determination, $\text{RMSE}_{\text{Model}}$ and the *F*-ratio of the ANOVA) is clearly demonstrated. The *F*-ratios for all of the models tested were significant and thus showed the principally significant relationship between dependent and independent

Table 2 Correlation coefficients, *r*, for the linear relationships between forest structural information unstratified and stratified, compared with features (spectral reflectance, texture features and indices) derived from RapidEye images

	dq	SN	BA	*V*
(a) Unstratified ($n = 216$)				
Mean NIR 05	−0.445**	0.188**		
GLCM Entropy Green 09	0.171*	0.060*		
GLCM Std.Dev. Red Edge 08	−0.203**			
Skewness Red 09	−0.154*			
Skewness NIR 08		0.188**		
SR 09		0.240**		
Brightness 08			−0.657**	−0.636**
GLCM Mean Red 05			−0.139*	
GLCM ASM NIR 08			0.173*	
Std.Dev. Red 05			−0.264**	
GR 09				0.280**
(b) cd (j_4) ($n = 84$)				
SR 09	−0.503**			
Mean Green 05	−0.381**			
GLCM Contrast Red Edge 08	−0.233*			
GLCM Contrast NIR 05	0.273*			
GLCM ASM NIR 08		0.388**		
GLCM Homogeneity NIR 05		0.274*		
Mean NIR 05		0.227*		
Skewness Red 08		−0.263*		
Brightness 08			−0.576**	−0.613**
GLCM Std.Dev. Green 09			0.293**	0.301**
GLCM ASM Red 08				0.376**
(c) pc (j_1) ($n = 55$)				
RE 05	−0.644**			
Std.Dev. Green 05	0.350**			
SR 05		−0.644**		
GLCM ASM Red Edge 05		0.310*	0.434**	
Brightness 08			−0.721**	
Brightness 05				−0.714**
GLCM Entropy Red 09				−0.330*
Mean Red 09				−0.497**
GLCM Correlation NIR 05				0.296*
(d) dd (j_3) ($n = 57$)				
GLCM Std.Dev. Red Edge 08	−0.374**			−0.339**
GLCM Mean Green 09	−0.364**			
GLCM Correlation Red 09	0.270*			
GLCM Entropy Green 09	0.331*			
GLCM ASM Green 09		0.428**		
Mean NIR 05		0.322*		
GLCM Std.Dev. NIR 05		0.350**		
GLCM Mean Blue 09			−0.379**	−0.358**
Mean Red Edge 08			−0.310*	
GR 09			0.334*	0.399**
GLCM Entropy Blue 05			−0.311*	
Brightness 08				−0.290*
(e) pd (j_2) ($n = 20$)				
GLCM Std.Dev. NIR 08	0.607**			
Std.Dev. Red Edge 08		−0.550*		
GLCM Std.Dev. Blue 08			−0.604**	−0.545*
GLCM ASM RedEdge 09			0.487*	0.533*

*P-value ≤0.05 (two-tailed) considered significant.
**P-value ≤0.01 (two-tailed) considered significant.
05 = May; 08 = August; 09 = September.

Table 3 Regression models for the forest structural information (dependent variables) and the RapidEye image features (independent variables) for unstratified and stratified forest sample plots

Dependent variables	Predictive model based on image feature			Std. error	Beta	VIF	t	Sig.	R^2
Unstratified									
dq (cm)	b_0	38.30		14.43			2.66	0.009	30
	b_1	−0.10	Mean NIR 05	0.01	−0.46	1.0	−7.93	<0.001	
	b_2	10.10	GLCM Entropy Green 09	2.65	0.22	1.0	3.81	<0.001	
	b_3	−0.56	GLCM Std.Dev. Red Edge 08	0.19	−0.18	1.0	−3.03	0.003	
	b_4	−3.68	Skewness Red 09	1.39	−0.15	1.0	−2.65	0.009	
Stratified									
dq (cm) (cd, j_4)	b_0	65.12		7.74			8.41	<0.001	38
	b_1	−1.48	SR 09	0.36	−0.39	1.1	−4.10	<0.001	
	b_2	−0.34	Mean Green 05	0.15	−0.21	1.2	−2.18	0.032	
	b_3	−0.003	GLCM Contrast Red Edge 08	0.001	−0.22	1.0	−2.41	0.018	
	b_4	0.002	GLCM Contrast NIR 05	0.001	0.21	1.1	2.28	0.025	
dq (cm) (pc, j_1)	b_0	49.01		3.85			12.73	<0.001	55
	b_1	−0.11	RE 05	0.02	−0.65	1.0	−6.95	<0.001	
	b_2	1.52	Std.Dev. Green 05	0.39	−0.36	1.0	3.87	<0.001	
dq (cm) (dd, j_3)	b_0	156.00		44.45			3.51	0.001	43
	b_1	−1.07	GLCM Std.Dev. Red Edge 08	0.28	−0.40	1.0	−3.80	<0.001	
	b_2	−0.99	GLCM Mean Green 09	0.28	−0.38	1.0	−3.62	0.001	
	b_3	18.82	GLCM Correlation Red 09	8.20	0.25	1.1	2.30	0.026	
	b_4	10.56	GLCM Entropy Green 09	4.79	0.24	1.1	2.20	0.032	
dq (cm) (pd, j_2)	b_0	−94.44		37.24			−2.54	0.021	37
	b_1	2.54	GLCM Std.Dev. NIR 08	0.78	0.61	1.0	3.24	0.005	
Unstratified									
SN (ha^{-1})	b_0	1503.36		393.11			3.82	<0.001	19
	b_1	−405.03	GLCM Entropy Green 09	90.87	−0.28	1.0	−4.46	<0.001	
	b_2	16.67	SR 09	10.33	0.12	1.4	1.61	0.108	
	b_3	214.56	Skewness NIR 08	51.37	0.27	1.0	4.18	<0.001	
	b_4	0.93	Mean NIR 05	0.32	0.22	1.5	2.92	0.004	
Stratified									
SN (ha^{-1}) (cd, j_4)	b_0	−1081.64		258.38			−4.19	<0.001	36
	b_1	81384.42	GLCM ASM NIR 08	19394.81	0.38	1.0	4.20	<0.001	
	b_2	10705.03	GLCM Homogeneity NIR 05	2839.64	0.34	1.0	3.77	<0.001	
	b_3	1.16	Mean NIR 05	0.38	0.27	1.0	3.04	0.003	
	b_4	−101.54	Skewness Red 08	41.27	−0.23	1.0	−2.46	0.016	
SN (ha^{-1}) (pc, j_1)	b_0	−1404.78		412.42			−3.41	0.001	31
	b_1	50.09	SR 05	12.47	0.46	1.0	4.02	<0.001	
	b_2	85672.99	GLCM ASM Red Edge 05	29071.60	0.34	1.0	2.95	0.005	
SN (ha^{-1}) (dd, j_3)	b_0	−2765.63		651.45			−4.25	<0.001	40
	b_1	21669.98	GLCM ASM Green 09	6635.68	0.35	1.0	3.27	0.002	
	b_2	2.02	Mean NIR 05	0.58	0.38	1.1	3.46	0.001	
	b_3	41.34	GLCM Std.Dev. NIR 05	12.10	0.38	1.1	3.42	0.001	
SN (ha^{-1}) (pd, j_2)	b_0	495.41		99.89			4.96	<0.001	30
	b_1	−17.87	Std.Dev. Red Edge 08	6.40	−0.55	1.0	−2.79	0.012	
Unstratified									
BA (m^2 ha^{-1})	b_0	90.89		17.67			5.15	<0.001	48
	b_1	−0.41	Brightness 08	0.03	−0.63	1.1	−12.22	<0.001	
	b_2	−0.37	GLCM Mean Red 05	0.13	−0.14	1.0	−2.77	0.006	
	b_3	909.06	GLCM ASM NIR 08	385.11	0.12	1.0	2.36	0.019	
	b_4	−0.56	Std.Dev. Red 05	0.28	−0.10	1.1	−2.01	0.046	

Continued

Table 3 *Continued*

Dependent variables	Predictive model based on image feature			Std. error	Beta	VIF	*t*	Sig.	R^2
Stratified									
BA (m^2 ha^{-1}) (cd, j_4)	b_0	8.84		12.96			0.68	0.497	40
	b_1	−0.37	Brightness 08	0.06	−0.56	1.0	−6.54	<0.001	
	b_2	0.81	GLCM Std.Dev. Green 09	0.26	0.27	1.0	3.08	0.003	
BA (m^2 ha^{-1}) (pc, j_1)	b_0	39.55		11.77			3.36	0.001	58
	b_1	−0.56	Brightness 08	0.08	−0.65	0.9	−7.00	<0.001	
	b_2	2126.20	GLCM ASM Red Edge 05	769.55	0.26	0.9	2.76	0.008	
BA (m^2 ha^{-1}) (dd, j_3)	b_0	35.95		8.22			4.37	<0.001	40
	b_1	− 0.09	GLCM Mean Blue 09	0.03	−0.33	1.1	−2.96	0.005	
	b_2	−0.14	Mean RedEdge 08	0.06	−0.27	1.1	−2.45	0.018	
	b_3	16.71	GR 09	6.38	0.30	1.1	2.62	0.012	
	b_4	−3.22	GLCM Entropy Blue 05	1.29	−0.28	1.1	−2.50	0.015	
BA (m^2 ha^{-1}) (pd, j_2)	b_0	−5.14		8.20			−0.63	0.539	53
	b_1	−0.19	GLCM Std.Dev. Blue 08	0.06	−0.54	1.0	−3.23	0.005	
	b_2	1560.58	GLCM ASM Red Edge 09	644.67	0.41	1.0	2.42	0.027	
Unstratified									
V (m^3 ha^{-1})	b_0	479.74		75.50			6.35	<0.001	42
	b_1	−5.30	Brightness 08	0.47	−0.60	1.1	−11.27	<0.001	
	b_2	192.30	GR 09	74.74	0.14	1.1	2.57	0.011	
Stratified									
V (m^3 ha^{-1}) (cd, j_4)	b_0	5.95		166.61			0.04	0.972	49
	b_1	−4.672	Brightness 08	0.75	−0.53	1.1	−6.24	<0.001	
	b_2	11.57	GLCM Std.Dev. Green 09	3.35	0.28	1.0	3.45	0.001	
	b_3	158.28	GLCM ASM Red 08	65.85	0.20	1.1	2.40	0.019	
V (m^3 ha^{-1}) (pc, j_1)	b_0	1481.75		336.31			4.41	<0.001	63
	b_1	−5.12	Brightness 05	0.91	−0.55	1.3	−5.62	<0.001	
	b_2	−164.42	GLCM Entropy Red 09	74.91	−0.19	1.0	−2.20	0.033	
	b_3	−9.66	Mean Red 09	3.99	−0.23	1.2	−2.42	0.019	
	b_4	358.87	GLCM Correlation NIR 05	169.40	0.19	1.1	2.12	0.039	
V (m^3 ha^{-1}) (dd, j_3)	b_0	666.78		170.87			3.90	<0.001	42
	b_1	279.92	GR 09	90.13	0.33	1.0	3.11	0.003	
	b_2	−7.41	GLCM Std.Dev. Red Edge 08	2.77	−0.29	1.0	−2.68	0.010	
	b_3	−1.32	GLCM Mean Blue 09	0.42	−0.34	1.0	−3.13	0.003	
	b_4	−2.04	Brightness 08	0.93	−0.24	0.9	−2.21	0.032	
V (m^3 ha^{-1}) (pd, j_2)	b_0	−162.57		127.56			−1.27	0.220	51
	b_1	−2.60	GLCM Std.Dev. Blue 08	0.94	−0.48	1.0	−2.77	0.013	
	b_2	26997.605	GLCM ASM Red Edge 09	10031.29	0.46	1.0	2.69	0.015	

t test significant *P*-value <0.05.

variables. Also, the analysis of the VIF showed that there was no multicollinearity among the independent variables used in the models (Table 3). Furthermore, the test for normality using the histogram with density plot and normal probability plot of residuals revealed nearly normally distributed residuals in general. It can, therefore, be concluded that the selected regression models are appropriate. Figure 2 shows two example plots of the standardized residuals against the standardized predicted values for *V* and dq for the stratum pc. Similar plots were analysed for all of the predicted structural attributes showing no particular trend in the residuals. The points were well distributed on both sides of the 1:1 line; no major shift was found in any of the models. Only the *V* model for stratum pd showed a small shift of 2.67 per cent as bias.

Table 3 shows the coefficient of determination (R^2) obtained for each model of the unstratified and stratified forest sample plots. The major findings in terms of the best predictors for each forest variable can be summarized as follows: the variable dq was best predicted using the GLCMs *Contrast Red Edge* and *Contrast NIR*, *Std.Dev. Red Edge* and *Std.Dev. NIR*, *Mean Green*, *Correlation Red* and *Entropy Green*. The indices SR and RE, in combination with any of the spectral features Mean Green, Mean NIR, Std.Dev. Green and Skewness Red, also explained dq quite well. For stratum pc, we obtained the highest R^2 (0.55) when the spectral features, Std.Dev. Green and the index RE were used. The lowest R^2 (0.37) within the stratification was obtained in stratum pd using the GLCMs *Std.Dev. NIR*. The accuracy check of the regionalization of dq compared with the terrestrial data showed a difference in the mean value of 0.47 cm.

The best estimate of SN was found using the GLCMs *ASM Green*, *ASM Red Edge* and *ASM NIR*, *Homogeneity NIR*, *Std.Dev. NIR* and

Table 4 Strength of the relationship between each of the simple and multiple regression models of the stratified test site

Dependent variable	Multiple correlation coefficient (R)	Coefficient of determination (R^2)	$RMSE_{Model}$	F	Significance
dq (cm) (cd)	0.62	0.38	10.1	12.2	<0.001
dq (cm) (pc)	0.73	0.55	8.8	31.1	<0.001
dq (cm) (dd)	0.66	0.43	10.0	9.9	<0.001
dq (cm) (pd)	0.61	0.37	12.4	10.5	0.005
SN (ha^{-1}) (cd)	0.60	0.36	286.6	11.3	<0.001
SN (ha^{-1}) (pc)	0.56	0.31	441.2	11.7	<0.001
SN (ha^{-1}) (dd)	0.63	0.40	347.7	11.9	<0.001
SN (ha^{-1}) (pd)	0.55	0.30	148.2	7.8	0.012
BA ($m^2\ ha^{-1}$) (cd)	0.63	0.40	8.3	27.2	<0.001
BA ($m^2\ ha^{-1}$) (pc)	0.76	0.58	11.3	36.2	<0.001
BA ($m^2\ ha^{-1}$) (dd)	0.63	0.40	6.7	8.5	<0.001
BA ($m^2\ ha^{-1}$) (pd)	0.73	0.53	6.6	9.5	0.002
V ($m^3\ ha^{-1}$) (cd)	0.70	0.49	105.6	25.2	<0.001
V ($m^3\ ha^{-1}$) (pc)	0.80	0.63	149.0	21.2	<0.001
V ($m^3\ ha^{-1}$) (dd)	0.65	0.42	97.4	9.4	<0.001
V ($m^3\ ha^{-1}$) (pd)	0.71	0.51	102.7	8.7	0.002

ANOVA F-value probability P-value ≤ 0.05.

Entropy Green. Furthermore, the index SR and the spectral features Mean NIR, Std.Dev. Red Edge, Skewness Red and Skewness NIR also fit the models well. The coefficient of determination for stratified SN varied between 0.30 (pd) and 0.40 (dd) (Table 4). The R^2 in stratum pd was best explained with the spectral feature Std.Dev. Red Edge; and in stratum dd, with texture feature GLCMs *ASM Green* and *Std.Dev. NIR* and the spectral feature Mean NIR. The accuracy test of the regionalized models for the mean value of the predicted SN (393 ha^{-1}) compared with the mean value derived from terrestrial measurements (334 ha^{-1}) showed a difference of $-59\ ha^{-1}$.

The GLCMs texture features *Std.Dev. Green* and *Std.Dev. Blue*, *ASM Red Edge* and *ASM NIR*, *Mean Blue* and *Mean Red*, as well as *Entropy Blue* best explained BA. The indices Brightness and GR and the spectral features Mean Red Edge and Std.Dev. Red were also included in these models. As shown in Table 4, the variation of R^2 among the different strata for the models of BA ranged from 0.40 (cd and dd) to 0.58 (pc). Regionalization accuracy showed a difference of $-6.12\ m^2\ ha^{-1}$ between the predicted mean value and the measured mean value.

V was best explained by the texture feature GLCMs *Std.Dev. Green, Std.Dev. Blue* and *Std.Dev. Red Edge*, *ASM Red* and *ASM Red Edge*, *Entropy Red*, *Correlation NIR* and *Mean Blue*. The indices Brightness and GR and the spectral feature Mean Red contributed to these models as well. The R^2 for the stratified models of V ranged between 0.42 (dd) and 0.63 (pc) (Table 4). The mean value for V predicted by the models across the entire test site was only 18.83 cubic meters less than that found using on the ground measures.

All regression models derived were statistically significant ($P \leq 0.05$). Relatively low R^2 values ranging from 0.31 to 0.43 were calculated for the models predicting dq (in strata cd, dd and pd), SN (all strata), BA (strata cd and dd) and V (stratum dd). In contrast, dq (stratum pc), BA (strata pc and pd) and V (strata cd, pc and pd) were relatively well predicted, as the models explained a considerable amount of the variability in these forest structural attributes. The calculated coefficients of determination for these models and the RMSEs derived from the LOOCV were 0.49 (109.0 $m^3 ha^{-1}$), 0.51 (118.9 $m^3\ ha^{-1}$), 0.53 (7.2 $m^2\ ha^{-1}$), 0.55 (9.0 cm), 0.58 (11.6 $m^2\ ha^{-1}$) and 0.63 (157.1 $m^3\ ha^{-1}$), respectively (Tables 4 and 5).

Stratification approach

The stratification of the forest sample plots based on forest type yielded a significant improvement in the results of the regression analyses and provided a clearer picture of the strata for which the model prediction worked particularly well. This was especially true for the variation in bias. For nearly all attributes, the highest values for the RMSE, rRMSE and bias were those for the pd stratum. The results for both the stratified and unstratified approaches are given in Table 5. For all strata, this approach resulted in a lower RMSE than that found for the unstratified estimation of all attributes: for dq, the estimated rRMSE was reduced from 33.6 to 31.2 per cent. The RMSE for the stratified estimation of SN (111.6 per cent) was lower than the 114.4 per cent found for the unstratified estimation and the error in the estimation of BA was reduced from 44.0 to 39.8 per cent. Finally, the stratified estimation of standing volume had a lower RMSE about 47.7 per cent compared with that found for the unstratified estimation about 54.1 per cent.

Multi-seasonal approach

In addition to the investigation of the impacts of forest stand stratification on forest attribute estimation, the value of using multi-seasonal image components was analysed. The influence of multiple datasets is summarized in Table 6. The model with dq as the dependent variable showed an R^2 of 29.8 per cent in stratum cd using mono-temporal image analysis, while the

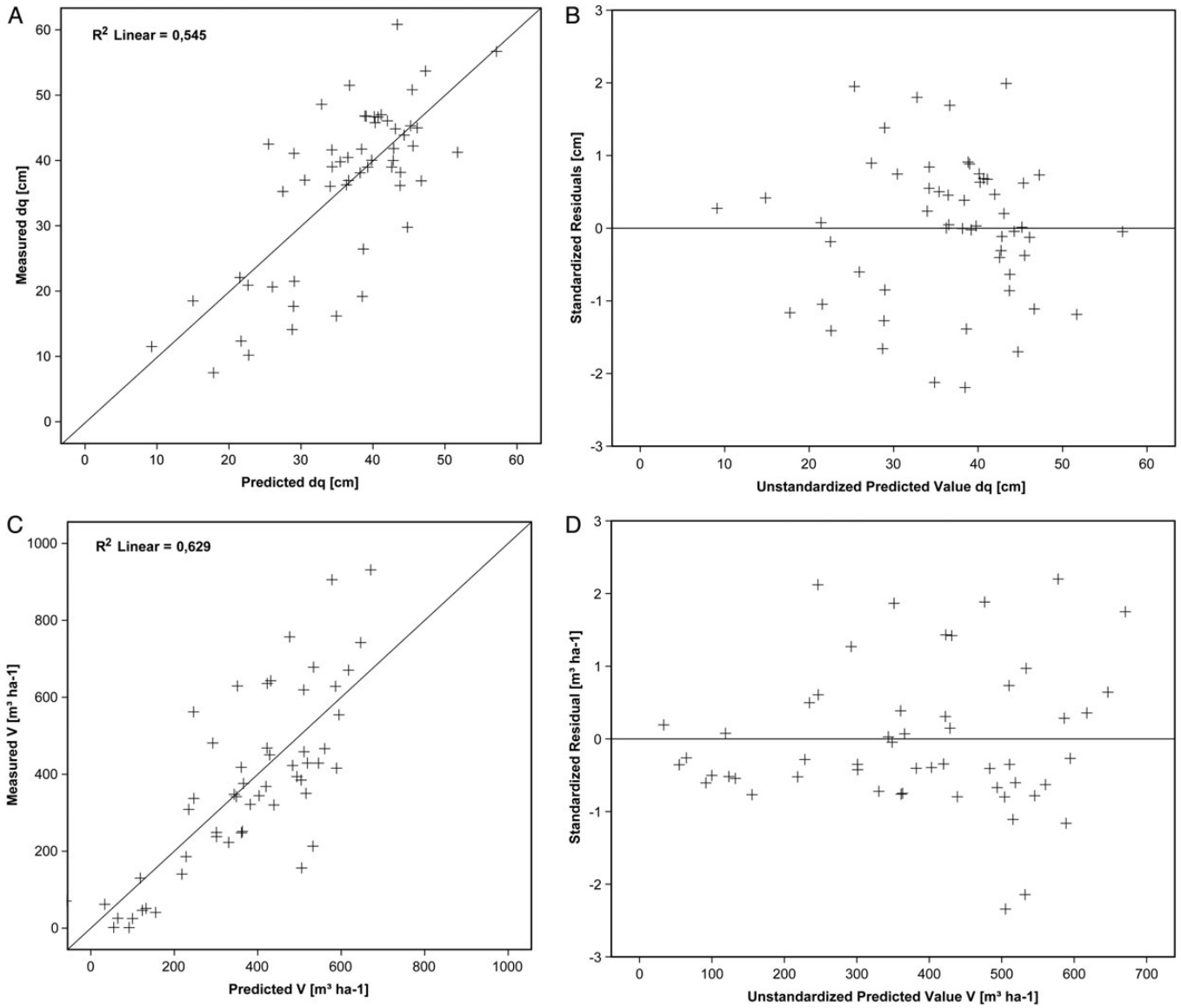

Figure 2 Field-measured forest structural information *vs* satellite-derived forest structural information and their residuals. Upper plot (A) scatter plot of dq and (B) residual plot of dq in stratum pure coniferous and the lower plot (C) scatterplot of *V* and (D) residual plot of *V* in stratum pure coniferous.

analysis of this stratum using multi-seasonal images increased the coefficient of determination to 38.0 per cent. Similar results were achieved with the same attribute in stratum dd, where the analysis using the mono-temporal image showed an R^2 of 24.5 per cent compared with an R^2 of 43.0 per cent obtained using multi-seasonal image analysis. Smaller improvements in R^2 for dq were also found for strata pc and pd, although the differences were not as large (Table 6). An increase in R^2 was also seen for SN and BA when using the multi-seasonal approach. The estimation showed better results for attribute SN in stratum dd and attribute BA for strata pc and cd using the mono-temporal approach. If acquisition of only one image were possible, the best seasonal image for the estimation of attribute dq would be the image from May for the strata cd and pc; while the best results for single image analysis for the deciduous strata were achieved with the data from August and September. Similar results were seen for attribute V.

Discussion and conclusions

Among the four forest variables investigated, the variables dq, Ba and *V* were most accurately estimated by the models. We found nothing in the literature with which we compare our results for dq. A comparison of the mean values predicted using the remote sensing data with the terrestrially measured mean values revealed overestimations in the model results for SN ($-59\ \mathrm{ha}^{-1}$), BA ($-6.12\ \mathrm{m}^2\ \mathrm{ha}^{-1}$) and *V* ($-18.83\ \mathrm{m}^3\ \mathrm{ha}^{-1}$). The models of dq showed an underestimation of the mean value of 0.47 cm. Our study also supports the findings about the forest variables SN and BA conducted by Ozdemir and Karnieli (2011). In that study, the estimated forest variable SN showed a R^2 of 38 per cent (RMSE $110\ \mathrm{ha}^{-1}$) and BA a R^2 of 54 per cent (RMSE $1.79\ \mathrm{m}^2\ \mathrm{ha}^{-1}$) using 2 m resolution (WorldView-2) data. Kayitakire *et al*. (2006) estimated the forest variable BA with a resulting R^2 of

Table 5 Plot level LOOCV RMSEs, relative RMSEs (%) and bias using stratified and unstratified stepwise-selected variables

	dq		SN		BA		*V*	
	RMSE (cm) rRMSE (%)	Bias (cm) rBias (%)	RMSE (ha^{-1}) rRMSE (%)	Bias (ha^{-1}) rBias (%)	RMSE ($m^2\ ha^{-1}$) rRMSE (%)	Bias ($m^2\ ha^{-1}$) rBias (%)	RMSE ($m^3\ ha^{-1}$) rRMSE (%)	Bias ($m^3\ ha^{-1}$) rBias (%)
cd	10.5	0.08	308.6	4.70	8.5	−0.01	109.0	−0.09
	31.1	0.24	100.8	1.50	37.1	−0.04	42.7	−0.04
pc	9.0	−0.02	483.0	3.37	11.6	0.03	157.1	2.52
	24.9	−0.06	130.5	0.91	37.0	0.10	43.3	0.69
dd	10.6	0.05	390.7	3.89	7.4	−0.06	106.4	−0.51
	32.5	0.15	103.8	1.03	39.8	−0.32	49.1	−0.24
pd	13.0	0.31	158.6	3.30	7.2	0.25	118.9	1.99
	49.9	1.19	68.3	1.42	76.3	2.67	112.2	1.88
Stratified	10.4	0.11	372.9	4.03	9.0	0.01	123.3	0.65
	31.2	0.33	111.6	1.21	39.8	0.05	47.7	0.25
Unstratified	11.2	0.00	382.1	2.54	10.0	0.02	140.0	−13.31
	33.6	0.00	114.4	0.76	44.0	0.11	54.1	−5.14

35 per cent (RMSE 6.85 $m^2\ ha^{-1}$) using 1 m resolution data (IKONOS-2). A comparison with the results published by Straub *et al*. (2013) for *V* and BA at the stand level at the same test site showed similar results to ours. The RMSE they achieved for *V* estimation using the stratification approach was 34.3 per cent for aerial image data and 37.7 per cent for LiDAR data, while our study showed a RMSE of 47.7 per cent using RapidEye data. The RMSE we found for the attribute BA was 39.8 per cent, in comparison with the results found by Straub *et al*. (2013) using the stratification approach for both aerial images (35.6 per cent) and LiDAR data (31.6 per cent). Our study showed slightly higher RMSE values than those obtained by Straub *et al*. (2013), which could be interpreted as an effect of the lower spatial resolution of the image data we used. These results confirm that while a pixel size of 5 m spatial resolution can be problematic for forest attribute estimation, these data can still compete with higher resolution data in terms of the results they bring. However, the pixel sizes of RapidEye data do not allow single tree detection; accordingly, the number of trees within one inventory plot cannot be determined with high accuracy using these data (Fehlert, 1984; Kenneweg *et al*., 1991).

Textural features such as GLCM *Dissimilarity* and indices like NDVI and VI showed little importance as predictors in the regression analysis. However, the use of textural features improved our results, thus confirming the findings of Kayitakire *et al*. (2006) using IKONOS-2 data, Wunderle *et al*. (2007) using SPOT 5 data and Ozdemir and Karnieli (2011) using WorldView 2-data. The GLCMs *Contrast and Entropy* best explained the forest variables BA and *V*, as described in the studies by Kayitakire *et al*. (2006) and Ozdemir and Karnieli (2011). Also the window size of 15 × 15 m used here confirmed its effectiveness, as published in the work by Kayitakire *et al*. (2006). The best results obtained in their study comparing the effectiveness of different window sizes (5 × 5 m, 15 × 15 m and 25 × 25 m) were for the models with a window size of 15 × 15 m.

Wolter *et al*. (2009) and Eckert (2006) used various vegetation indices to estimate structural attributes in their studies. In our study, the best results were achieved using the indices SR, GR, RE and Brightness. The effectiveness of these indices for estimating BA and *V* can be explained after Eckert (2006) as follows: a low value for this attribute implies the presence of stands of coniferous forest with shady areas and relatively low stand density where one would expect high values for *V* and BA, while higher values for these indices imply broadleaved forest with a closed canopy. Furthermore, the spectral feature Mean NIR was chosen three times in different models as a predictor. Eckert (2006) reported that low NIR reflectance is related to old-growth, low stand density and stands with high timber volume, whereas high reflectance in the NIR implies small timber and mixed timber stages with high stand densities. In our study, Mean NIR proved to be a good predictor for the attribute SN.

Our results also show that the stratification approach generally improved the estimation of forest structural information. Heurich (2006) and Latifi *et al*. (2012) used similar approaches involving stratification of forest plots into coniferous, deciduous and mixed strata, while Straub *et al*. (2013) used only coniferous and deciduous strata. Another finding of the study presented here was that forest information could be predicted with higher accuracy in coniferous stands than in deciduous stands. Wolter *et al*. (2009) attributed a similar finding to the greater contrast between sunlight and shadow in coniferous stands. Also, the stronger correlation among the dependent structural variables dq, BA and *V* for conifers improved the results obtained for the pc and cd strata. One concern is the relatively high bias in the pd strata (2.67 per cent). The bias could be a result of small trees being excluded from the analysis. Alternatively, an overestimation in the model may have been caused by either the low number of inventory plots available (for pd, $n = 20$) or the high level of variation within the forest stands in Traunstein, as described by Heurich (2006) when using LiDAR data to estimate forest information in the Bavarian Forest National Park. The estimated forest variable SN showed improvements in the models with species-based stratification approach. However, the RMSE and bias showed high SN values as compared with the other three forest variables. A similar result was also observed by Ozdemir and Karnieli (2011). This could be improved by stratifying based on the age of stands as suggested by the same authors.

The potential advantages of the multi-seasonal approach were also analysed in a study by Maselli *et al*. (2005), who estimated

Table 6 Comparison of the R^2 and the absolute and relative RMSEs found using multi-temporal and mono-temporal analysis for the estimation of forest structural information

	R^2	RMSE	rRMSE
dq	(%)	(cm)	(%)
Multi-seasonal cd	38.0	10.5	31.1
Mono-temporal 05 cd	29.8	10.9	32.5
Multi-seasonal pc	55.0	9.0	24.9
Mono-temporal 05 pc	54.5	9.0	24.9
Multi-seasonal dd	43.0	10.6	32.5
Mono-temporal 08 dd	24.5	11.8	36.4
Multi-seasonal pd	37.0	13.0	49.9
Mono-temporal 09 pd	36.9	13.0	49.9
SN	(%)	(ha^{-1})	(%)
Multi-seasonal cd	36.0	308.6	100.8
Mono-temporal 09 cd	24.5	323.4	105.7
Multi-seasonal pc	31.0	483.0	130.5
Mono-temporal 05 pc	31.0	483.1	130.6
Multi-seasonal dd	40.0	390.7	103.8
Mono-temporal 05 dd	45.4	383.8	101.8
Multi-seasonal pd	30.0	158.6	68.3
Mono-temporal 09 pd	30.2	158.7	68.3
BA	(%)	(m^2 ha^{-1})	(%)
Multi-seasonal cd	40.0	8.5	37.2
Mono-temporal 09 cd	41.5	8.8	38.4
Multi-seasonal pc	58.0	11.6	37.0
Mono-temporal 05 pc	62.2	11.1	35.7
Multi-seasonal dd	40.0	7.4	39.8
Mono-temporal 08 dd	23.9	8.0	42.6
Multi-seasonal pd	53.0	7.2	76.3
Mono-temporal 09 pd	36.0	8.0	86.5
V	(%)	(m^3 ha^{-1})	(%)
Multi-seasonal cd	49.0	109.0	42.7
Mono-temporal 08 cd	44.9	110.8	43.3
Multi-seasonal pc	63.0	157.1	43.3
Mono-temporal 08 pc	54.2	168.5	46.4
Multi-seasonal dd	42.0	106.4	49.1
Mono-temporal 09 dd	27.3	110.2	50.8
Multi-seasonal pd	51.0	118.9	112.2
Mono-temporal 09 pd	29.7	125.6	118.4

BA using a seasonal time series of Landsat 7 ETM + datasets. However, Maselli *et al*. (2005) reported only a marginal and inconsistent improvement in estimation through the use of multi-seasonal datasets. Our study showed similar findings for SN and BA, depending on the stratum. In all strata, significant improvements were obtained for the estimation of the attributes dq and *V* using the multi-seasonal approach. Furthermore, the study presented here found that the relationship between terrestrial measurements of forest attributes and spectral information is more obvious when images from the beginning or the end of the growing season are used, confirming the findings of Maselli *et al*. (2005) as well as those of Elatawneh *et al*. (2013) for a forest in Freising, Germany.

The study presented here examined the potential for the estimation of forest structural attributes using forest stand-type stratification and high-resolution, multi-seasonal RapidEye data. Object-based extraction of image features, spectral bands, indices and textural features from RapidEye sensor data provided an efficient basis for estimating forest attributes using simple or multiple linear regression analysis. The regionalization of the data to the test site showed good results. Furthermore, RapidEye data have the advantages of high areal coverage and the affordable price of only 0.95€ per km^2 (RapidEye AG, 2011). Thus, the costs of forest inventories could be reduced by using RapidEye archive data for forest structural estimation. Another major advantage of this data is the high temporal resolution. The RapidEye satellite system offers a repetition rate of 5.5 days (nadir view), whereas a forest inventory will not likely be repeated more often than every 10 years. Therefore, RapidEye data can provide timely information about changes in the forest structure and thus can be used to simulate interim forest inventories. Furthermore, it can provide the input data necessary to initialize forest growth simulators.

Both the methodology used and the RapidEye data evaluated, however, bring to light some shortcomings of our study. The relatively high RMSE obtained can most likely be attributed to the mixed pixel effect of the 5 m resolution data used. However, our aim was to estimate structural information on a regional scale, which means at the enterprise level. Thus, the RMSE found in our study calculated at a stand level is sufficient to fulfil the expected requirements, particularly when estimating attributes for the entire forest area. The application of the chosen predictors to different forest areas and different RapidEye datasets could lead to different results under different circumstances; thus, the approach needs to be tested in further studies. The study did, however, show the potential applicability of high-resolution data as a first step in supporting forest management plans by estimating forest structural information. The map produced can be used as preliminary advice to forest enterprises for forest management plan preparation. In future, we also intend to assess how well the estimated values work to initialize a growth simulator.

Acknowledgements

The authors are grateful to RESA of the DLR for the provision of RapidEye data and especially for their excellent support. Great thanks also go to the Chair for Forest Growth and Yield, Technische Universität München, the site manager of the test site forest, Mr Fischer, and the Bavarian State Institute of Forestry for providing the reference data. The authors would also like to thank the anonymous reviewers for their thorough review and Laura Carlson for language editing.

Funding

This work was supported by funding from the Federal Ministry of Economics and Technology (number 50EE0919) within the programme 'synergistic use of RapidEye and TerraSAR-X data for applications' of the Space Agency of the German Aerospace Centre (DLR). RapidEye data were provided via the RapidEye Science Archive (RESA) of the German Aerospace Centre (DLR) under project number 317.

Conflict of interest statement

None declared.

References

Baatz, M. and Schäpe, A. 2000 Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. *J. Photogram. Remote Sens.* **58**, 12–23.

Castillo-Santiago, M.A., Ricker, M. and de Jong, B.H. 2010 Estimation of tropical forest structure from SPOT-5 satellite images. *Int. J. Remote Sens.* **31**, 2767–2782.

Coburn, C.A. and Roberts, A.C.B. 2004 A multiscale texture analysis procedure for improved forest stand classification. *Int. J. Remote Sens.* **25**, 4287–4308.

Ďurský, J. 2000 *Einsatz von Waldwachstumssimulatoren für Bestand, Betrieb und Großregion*. Habilitationsschrift an der Forstwissenschaftlichen Fakultät der Technischen Universität München, 223 pp.

Eckert, S. 2006 *A Contribution to Sustainable Forest Management in Patagonia. Object-oriented Classification and Forest Parameter Extraction based on ASTER and Landsat ETM+ Data*. Dissertationsschrift der Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich, 180 pp.

Elatawneh, A., Rappl, A., Reshush, N., Schneider, T. and Knoke, T. 2013 Forest tree species identification using phenological stages and RapidEye data: a case study in the forest of Freising. In *5. RESA Workshop From the Basics to the Service*. Boldt (ed). GITO, pp. 23–38.

ENVI. 2005 User's Guide. 4.3. edn.

Fehlert, G.P. 1984 *Kalibrierung von MSS-Satellitenbilddaten zur Auswertung zeitlicher Reflexionsänderungen an Fichtenbeständen.* DFVLR Forschungsbericht, pp. 44–84.

Felbermeier, B., Hahn, A. and Schneider, T. 2010 Study on user requirements for remote sensing applications in forestry. *Proceeding of the ISPRS Comission VII Symposium*. Wien.

FER. 2011 *Richtlinien für die mittel- und langfristige Forstbetriebsplanung in den Bayerischen Staatsforsten. Forsteinrichtungsrichtlinie – FER 2011*. Bayerische Staatsforsten (Bavaria State Forest Enterprise).

Haralick, R.M. 1979 Statistical and structural approaches to texture. *Proc. IEEE* **67**, 786–804.

Haralick, R.M., Shanmugam, K. and Dinstein, I.H. 1973 Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **6**, 610–621.

Heurich, M. 2006 *Evaluierung und Entwicklung von Methoden zur automatisierten Erfassung von Waldstrukturen aus Daten flugzeuggetragener Fernerkundungssensoren*. Schriftenreihe des Wissenschaftszentrums Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München und der Bayerischen Landesanstalt für Wald und Forstwirtschaft.

Jordan, C.F. 1969 Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **50**, 663–666.

Kanemasu, E.T. 1974 Seasonal canopy reflectance patterns of wheat, sorghum, and soybean. *Remote Sens. Environ.* **3**, 43–47.

Kayitakire, F., Hamel, C. and Defourny, P. 2006 Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sens. Environ.* **102**, 390–401.

Kenneweg, H., Förster, B. and Runkel, M. 1991 Diagnose und Erfassung von Waldschäden auf der Basis von Spektralsignaturen. *In DLR Abschlußdokumentation – Untersuchung und Kartierung von Waldschäden mit Methoden der Fernerkundung*. Teil A.

Kramer, H. and Akça, A. 2008 *Leitfaden zur Waldmesslehre*. 5th edn. Sauerländer, 266 pp.

Latifi, H., Nothdurft, A. and Koch, B. 2010 Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. *Forestry* **83**, 395–407.

Latifi, H., Nothdurft, A., Straub, C. and Koch, B. 2012 Modeling stratified forest attributes using optical/LIDAR features in a central European landscape. *Int. J. Digit. Earth* **5**, 106–132.

Lyon, J.G., Yuan, D., Lunetta, R.S. and Elvidge, C.D. 1998 A change detection experiment using vegetation indices. *Photogram. Eng. Remote Sens.* **2**, 143–150.

Maselli, F., Chirici, G., Bottai, L., Corona, P. and Marchetti, M. 2005 Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images. *Int. J. Remote Sens.* **26**, 3781–3796.

Ozdemir, I. and Karnieli, A. 2011 Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *Int. J. Appl. Earth Observ. Geoinf.* **13**, 701–710.

Rahlf, J. 2011 *Möglichkeiten der Waldabgrenzung und Bestandeshöhenabschätzung mit Hilfe von LiDAR- und RapidEye-Daten*. Master's thesis, Institute of Forest Management, Technische Universität München, 83 pp.

RapidEye AG. 2011 *Satellite imagery product specifications*. 3.2 edn. RapidEye AG.

Richardson, A.J. and Everitt, J.H. 1992 Using spectral vegetation indices to estimate rangeland productivity. *Geocarto Int.* **7**, 63–69.

Richter, R. 2011 Atmospheric and Topographic Correction for Airborne Imagery. ATCOR-3 user guide 3.1 edn. Technical report, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Germany.

Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W. 1974 Monitoring vegetation systems in the great plains with ERTS. In *Proceedings of the ERTS-1 3rd Symposium*, vol 1. NASA SP-351. NASA, Washington, pp. 309–317.

Schneider, T., Tian, J., Elatawneh, A., Rappl, A. and Reinartz, P. 2012 Tracing structural changes of a complex forest by a multiple systems approach. In *Proceedings of 1st European Association of Remote Sensing Laboratories Workshop on Temporal Analysis of Satellite Images*, EARSeL, Mykonos, Greece, 159–165.

Schnell, A. and Bauer, A. 2005 Die zweite Bundeswaldinventur 2002: Ergebnisse für Bayern. Berichte der Bayerischen Landesanstalt für Wald und Forstwirtschaft. *LWF Wissen* **49**, 4–103.

Stepper, C. and Schneider, T. 2012 Altersschätzung von Waldbeständen mittels Texturmerkmalen. *Allg. Forst Z. Waldwirtschaft* **67**, 20–21.

Stone, M. 1974 Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc.* **36**, 111–147.

Straub, C., Weinacker, H. and Koch, B. 2010 A comparison of different methods for forest resource estimation using information from airborne laser scanning and CIR orthophotos. *Eur. J. Forest Res.* **129**, 1069–1080.

Straub, C., Stepper, C., Seitz, R. and Waser, L.T. 2013 Möglichkeiten einer stratifizierten Schätzung von forstlichen Kenngrößen unter Verwendung von amtlichen Stereo-Luftbildern, Orthophotos und Laserscannerdaten. In *33. Wissenschaftlich – Technische Jahrestagung der DGPF Band 22/2013*, Seyfert (ed.). DGPF pp. 186–193.

Trimble. 2013 *eCognition Developer 8.9 Reference Book*. 8.9.1 edn. Trimble Germany GmbH, 249 pp.

Witten, I., Frank, E. and Hall, M. 2011 *Data Mining. Practical Machine Learning Tools and Techniques*. 3rd edn. Morgan Kaufmann, 584 pp.

Wolter, P.T., Townsend, P.A. and Sturtevant, B.R. 2009 Estimation of forest structural parameters using 5 and 10 meter SPOT-5 satellite data. *Remote Sens. Environ.* **113**, 2019–2036.

Wunderle, A.L., Franklin, S.E. and Guo, X.G. 2007 Regenerating boreal forest structure estimation using SPOT 5 pan sharpened imagery. *Int. J. Remote Sens.* **28**, 4351–4364.