RETRIEVAL OF CONIFEROUS CANOPY CHLOROPHYLL CONTENT FROM HIGH SPATIAL RESOLUTION HYPERSPECTRAL DATA

Z. Malenovský [a, b, *], R. Zurita-Milla [b], L. Homolová [a, c], E. Martin [d], M. E. Schaepman [b], J.P. Gastellu-Etchegorry [d], R. Pokorný [a], J.G.P.W. Clevers [b]

[a] Inst. of Systems Biology and Ecology, Academy of Sciences of the Czech Republic, Poříčí 3b, 60300 Brno, Czech Republic - zbynek.malenovsky@gmail.com, (lucie.homolova, eradek)@brno.cas.cz

[b] Centre for Geo-Information, Wageningen University, Droevendaalsesteeg 3 / PO Box 47, 6700 AA Wageningen, The Netherlands - (raul.zurita-milla, michael.schaepman, jan.clevers)@wur.nl

[c] Dep. of Applied Geoinformatics and Cartography, Faculty of Science, Charles University in Prague, Albertov 6, 12843 Prague 2, Czech Republic

[d] Centre d'Etudes Spatiales de la Biosphère, UPS-CNES-CNRS-IRD, 18 Avenue Edouard Belin, BPI 2801, 31401 Toulouse, Cedex 9, France - (emmanuel.martin, jean-philippe.gastellu)@cesbio.cnes.fr

Commission VI, WG VI/4

KEY WORDS: Remote Sensing, Hyperspectral, High Resolution, Forestry, Vegetation, Estimation, Retrieval, Neural

ABSTRACT:

The discrete Anisotropic Radiative Transfer (DART) model, coupled with an adjusted version of the PROSPECT model, was used to retrieve total chlorophyll content (C_{ab}) of a complex Norway spruce (*Picea abies* (L.) Karst.) canopy from airborne hyperspectral data acquired at very high spatial resolution. The radiative transfer models were parameterized by using field measurements and observations collected from a young spruce stand growing at the permanent experimental site Bílý Kříž (the Moravian-Silesian Beskydy Mts., the Czech Republic, 18.53863°E, 49.50256°N, 936 m a.s.l.). A set of the hyperspectral images with a pixel-size of 0.4 m was acquired for the test site by an airborne AISA Eagle VNIR system in September 18th, 2004. An operational canopy C_{ab} estimation was carried out by means of a PROSPECT-DART inversion employing an artificial neural network (ANN) and a vegetation index $ANCB_{650-720}$. Both retrieval approaches used continuum removed reflectance values of six AISA Eagle spectral bands located between 650 and 720 nm. The C_{ab} inversion was only performed for direct sun exposed (sunlit) crown pixels in order to ensure a high quality (noiseless) reflectance signal. Results of both inversion approaches were similar, when validated against the ground measured C_{ab} of nine Norway spruce crowns. Coefficients of determination (R^2) between ground truth and remote sensing C_{ab} estimates were 0.78 and 0.76, respectively, with root mean square errors (RMSE) of 2.95 $\mu g\ cm^{-2}$ for the ANN and 3.36 $\mu g\ cm^{-2}$ for the $ANCB_{650-720}$ retrieval. The spatial patterns of C_{ab} values estimated by both inversion methods were consistent with each other. About 80% of the C_{ab} estimated values had an absolute difference smaller than 2 $\mu g\ cm^{-2}$.

1. INTRODUCTION

1.1 General Introduction

The total content of chlorophyll *a* and *b* (C_{ab}) is an important biomarker of the actual plant status. C_{ab} can also be used as an indicator of vegetation gross primary productivity (Gitelson et al., 2006). A number of studies has been carried out to design an appropriate algorithm to estimate the total chlorophyll content of structurally homogeneous agriculture crops (Daughtry et al., 2000; Haboudane et al., 2002), but also for structurally complex coniferous forest canopies (Zarco-Tejada et al., 2004), using optical remote sensing data. Canopy C_{ab} content can be spatially retrieved from hyperspectral images by using chlorophyll vegetation indices (le Maire et al., 2004) or by inversion of physical radiative transfer (RT) models (Gascon et al., 2004). However, if very high spatial resolution data (pixel-size of about 1 m) is used to analyze structurally heterogeneous canopies, the complexity of the inversion increases as the scene needs to be modeled with more details. Therefore, physically based C_{ab} content retrievals for such canopies need appropriate RT models capable to address the complexity of the scene at such a high spatial resolution.

1.2 Objectives

The main objective of the study was to develop a physically based approach to retrieve total chlorophyll content (C_{ab}) of a complex Norway spruce (*Picea abies* (L.) Karst.) canopy from very high spatial resolution (0.4 m) hyperspectral data. The motivation for this work was to produce high spatial resolution maps of canopy C_{ab}, which in turn could serve as a reference to validate satellite-based chlorophyll products. Additionally, we were using this exercise to better understand the up-scaling of chlorophyll retrieval algorithms from leave to crown level.

2. MATERIAL AND METHODS

2.1 General methodology

The general methodology used in this study is schematically depicted in the flowchart of figure 1.

* Corresponding author.

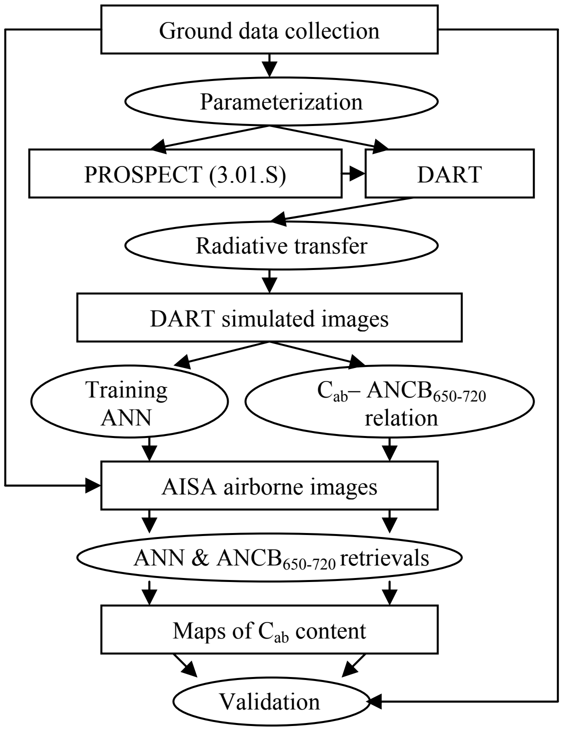

Figure 1. The methodological flowchart.

2.2 Study area and chlorophyll ground truth data

A montane forest stand of 28 years old Norway spruce trees, located at the permanent experimental research site Bily Kriz (the Moravian-Silesian Beskydy Mountains; 18.54°E, 49.50°N; altitude 936 m above sea level), was selected to conduct this study (figure 2). The average annual air temperature at this stand is about 5.5ºC, the average annual precipitation amounts around 1000-1400 mm. This regularly spaced plantation was established with three years old spruce seedlings in 1981. In 2004, the spruces were about 10 m tall with an average diameter at breast height (DBH) of 12.8 cm.

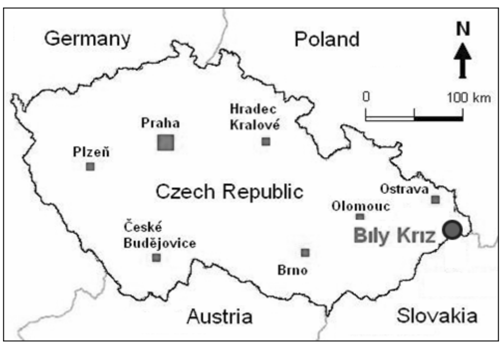

Figure 2. Geographic location of the experimental research plot Bílý Kříž (Europe, East border of the Czech Republic).

Needle samples collected from ten selected spruce trees were analyzed in a laboratory to obtain the canopy C_{ab} content that was later used to validate the results. Samples of last three age-classes were collected from a sunlit, transitional, and shaded branch cut off from each sampled tree. The representative C_{ab} content of each crown was computed as an average weighted by irradiation availability within the horizontal crown levels and abundance of the age-classes within the whole crown.

2.3 Hyperspectral airborne images

An airborne/field campaign was carried out at the research site in September 18th, 2004. Several multi-directional flight lines were acquired in a star pattern with an aerial VNIR hyperspectral system AISA Eagle (SPECIM Ltd., Finland). The resulting AISA Eagle images had a pixel size of 0.4m and 64 spectral bands with a Full-Width-Half-Maximum (FWHM) of about 10 nm. Data for the radiometric and atmospheric corrections and for geo-orthorectification of the AISA images were collected simultaneously with the sensor over flight. The digital numbers were transformed into radiance values using sensor calibration files with the CaliGeo software. The atmospheric correction was performed using the ATCOR-4 model (Richter and Schlapfer, 2002). The ATCOR-4 model was also used to correct the brightness reflectance gradient within the airborne images in the across-track direction (BRDF correction). The midpoint of all the flight lines was situated over the testing forest stand. Therefore, we had the opportunity to investigate the influence of the different flight directions (different geometrical set-ups of the sun-object-sensor system) on the C_{ab} estimation from AISA Eagle images. For this purpose we selected an area of 116 by 110 m that was covered by three flight lines: flight line 1 (flown from East to West), flight line 2 (flown from North to South), and flight line 5 (flown from West to East). These three AISA Eagle testing subsets were classified using a maximum likelihood rule in three classes: background, sunlit and shaded crowns (figure 3). The classification results were used to extract the signal of the sunlit crown pixels employed in Cab retrieval and to compute the canopy cover (CC) of the observed Norway spruce forest stand (table 1).

2.4 Radiative transfer modelling

The leaf radiative transfer model PROSPECT (Jacquemoud and Baret, 1990) adjusted for Norway spruce needles (Malenovsky et al., 2006) was coupled with the 3D Discrete Anisotropic Radiative Transfer (DART) model (Gastellu-Etchegorry et al., 2004) to simulate hyperspectral images of the Norway spruce scene under investigation..

Parameter	Units	Values
Slope	Deg.	13.5°
Sun angles	Deg.	$\theta_s = 47.8°$, $\varphi_s = 183.4°$
Canopy cover	%	75, 85, 95
Leaf area index	m^2 m^{-2}	4, 5, 6, 7, 8, 9
Chlorophyll content	μg cm^{-2}	10, 25, 40, 55, 70,85
Visible simulated bands	nm	652.1, 661.4, 670.7, 680.1, 689.4
NIR simulated bands	nm	698.7, 708.1, 717.4

Table 1. Basic inputs used for PROSPECT-DART modelling parameterization.

PROSPECT input parameters C_m (dry matter content; C_m = 0.0118-0.0233 g cm^{-2}) and C_w (water content; C_w = 0.0365-0.0486 cm) were measured on the needle samples collected for the chlorophyll analysis. The structural parameter N was retrieved from reflectance and transmittance measurements of similar needle samples (N = 2.02-2.08). Chlorophyll content was varying between 10-85 μg cm^{-2} according to prior measurements (see table 1).

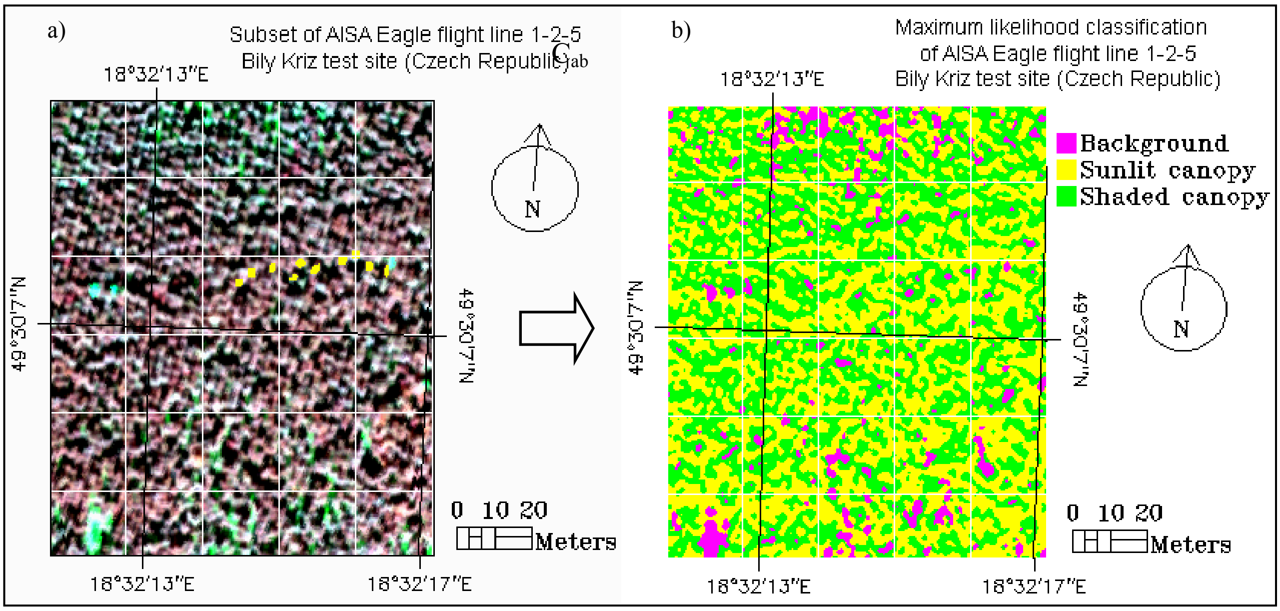

Figure 3. Testing subset of the AISA Eagle flight lines 1, 2, and 5: a) false colour RGB composition (NIR, Red, Green bands) of the Norway spruce test site with ten selected chlorophyll sampling crowns (yellow polygons), and b) thematic map of Maximum Likelihood automatic classification of the testing image subset.

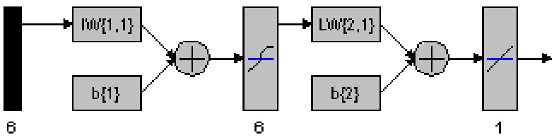

Figure 4. Design of the artificial neuron network (ANN) applied for chlorophyll content estimation (C_{ab}).

The DART model was parameterized by means of the ground data obtained at the test site during the ground/flight campaign. Other required allometric and eco-physiological measurements of the tree crowns were obtained from previous field measurements taking place that the same research site earlier in 1997 (Pokorný and Marek, 2000). Three-dimensional representations of a Norway spruce site of an area of 6 by 6 m were constructed using four (CC = 75%), five (CC = 85%), and six (CC = 95%) trees in case of a regular tree distribution and five (CC = 75%), six (CC = 85%), and seven (CC = 95%) trees in case of an irregular (clumped) stem distribution. Each Norway spruce tree, having a total height ranging between 9 – 11 m, was created using 11 crown horizontal levels. For each level were defined specific average leaf angles (from 25° to 40°) and specific leaf optical properties (reflectance, transmittance, and absorption). The leaf optical properties were simulated using the Norway spruce adjusted PROSPECT model and according to the proportion of the different needle age-classes that were found to be present in each of the crown levels. Destructive field measurements were used to parameterise the vertical and horizontal leaf distributions, spatially specific crown defoliation, the distribution of woody elements (DART uses superimposed parallelepipeds to represent trunks and pyramids to represent branches of first order, i.e. branches that grow directly from the trunk), and the distribution of woody tiny twigs (DART uses turbid media to represent branched smaller than 1 cm in diameter). The forest stand background, covering the continuous slope of 13.5°, was modelled as a mixture of bare soil and needle litter. The optical properties of the other scene surfaces (bark of trunks and branches, and forest background elements) were measured with an ASD FieldSpec Pro spectroradiometer connected to the Li-Cor integrating sphere Li-1800-12. They were defined in DART as to be of Lambertian nature. The radiative transfer through the atmosphere above the forest stand was not included. Therefore, the spectral bands simulated by DART corresponded to the top of canopy (TOC) bidirectional reflectance function (BRF) recorded by the AISA Eagle bands (mentioned in table 1).

2.5 Chlorophyll estimating methods

The total chlorophyll content of the Norway spruce crowns was estimated from the AISA images using a database of DART simulated images by means of two methods: i) an artificial neural network (ANN) and ii) an optical vegetation index, namely Area under curve Normalized to maximal Chlorophyll absorption Band depth between 650-720 nm ($ANCB_{650-720}$).

2.5.1 Artificial neural network:

The performance of several ANN architectures was tested using the neural network toolbox available in MATLAB ®. A two-layer feedforward backpropagation neural network was finally selected for the analysis. A tan-sigmoidal transfer function was used in the first layer and a linear transfer function in the output layer (figure 4). The network was trained using the sunlit DART simulated pixels in the wavelength range from 650 to 720 nm, transformed by means of continuum removal (Kokaly and Clark, 1999; Curran et al., 2001). Before the training of the network, the continuum removed BRF and the C_{ab} data were pre-processed so that they have a zero mean and a standard deviation of 1. The ANN had six neurons of the first layer, because each input requires a neuron. The output layer consisted of one neuron – C_{ab} content. The Levenberg-Marquardt optimization algorithm was selected for the training of the network (it is very fast, but it requires a large amount of memory). The 216 DART simulated scenes were split into 3 groups: i) training (50%), ii) validation (25%), and iii) testing (25%) sets. To avoid the so-called overfitting of the ANN, an early stopping technique was used during the training of the network. This means that the validation dataset was presented to the network simultaneously with the training dataset and when the error of the validation dataset was above a certain threshold then the training of the network was stopped (even though the error of the training dataset might still be declining).

2.5.2 Chlorophyll vegetation index: The $ANCB_{650-720}$ is a variant of the optical index $ANMB_{650-725}$ (Area under curve Normalized to Maximal Band depth between 650-725 nm), described in detail in Malenovsky et al. (2006). The only difference is that the $ANCB_{650-720}$ uses a shorter wavelength range (650-720 nm) and that the normalization of the continuum removed area is done using the maximum chlorophyll absorption wavelength (in this case 670.7 nm). The $ANCB_{650-720}$ computed from the sunlit crown pixels of the DART simulated images was statistically related to the C_{ab} used during the DART simulations. Finally, the obtained equation was applied to the sunlit pixels of the AISA Eagle images.

3. RESULTS AND DISCUSSION

The performance of the trained ANN was assessed using the testing dataset. Results were quite satisfactory: R^2 equal 0.9988 and root-mean-square-error (RMSE) equal to 0.40 μg.cm^{-2}. The $ANCB_{650-720}$ statistical exponential relationship established using the DART simulated images was:

$$\ln(C_{ab}) = 7.3903 - 7984.0135/(ANCB_{650-720})^2 \quad (1)$$

with an R^2 of 0.9989.

3.1 Crown chlorophyll estimation

Figure 5 illustrates the validation of chlorophyll content that was estimated for each of the sampled crown used to create the validation dataset. The graphs show that both methods, applied to the three AISA flight lines, gave very similar results. The RMSE between measured and retrieved C_{ab} values was found to be about 3 μg.cm^{-2} for both methods. An outlier was identified with both methods (see graphs A, B). This point corresponds to a crown that grows next to a highly reflective metal meteorological tower. Apparently the reflectance of this tower contaminated the spectral signature of this crown and caused the C_{ab} overestimation. In general, the $ANCB_{650-720}$ index yielded lower C_{ab} estimates for low chlorophyll concentrations, while the estimates for high C_{ab} concentrations were fitting well with the results found for the ANN approach (see graph C). Comparison with the ground truth measurements suggests that the $ANCB_{650-720}$ approach is less accurate and probably less sensitive to retrieve lower chlorophyll content values. Finally,

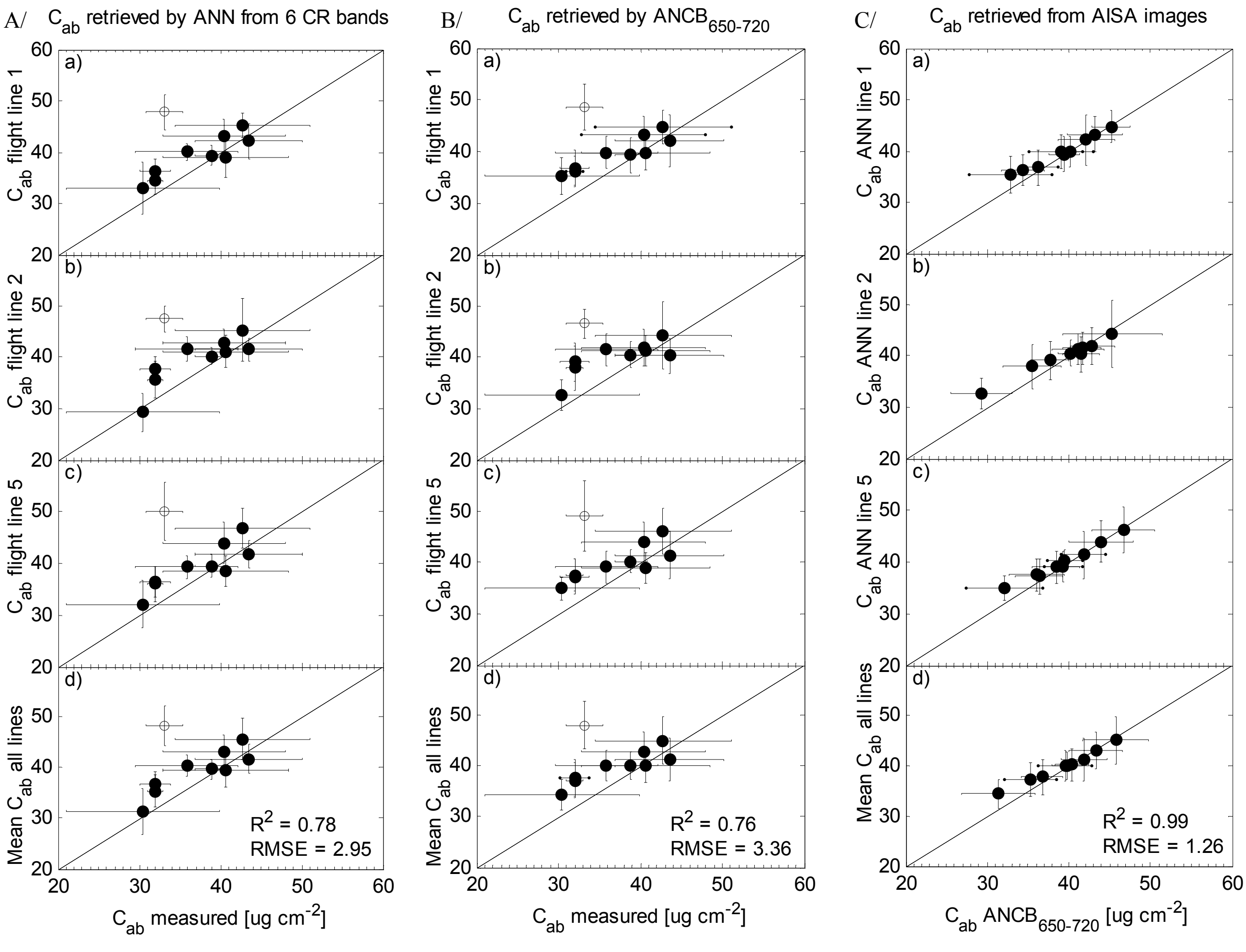

Figure 5. Validation of chlorophyll content (C_{ab}) retrieved for ten sampled spruce crowns from the AISA Eagle flight line 1 (a), flight line 2 (b), flight line 5 (c), and mean values of all three flight lines (d) using: A/ an artificial neural network (ANN), B/ optical index $ANCB_{650-720}$, and C/ reciprocal comparison of both methods. (Each dot symbol represents one tree crown, empty dot means an outlier, horizontal bars represent two standard deviations of measured C_{ab} values (A,B) or retrieved by $ANCB_{650-720}$ (C), and vertical bars represent two standard deviations of C_{ab} values estimated by ANN (A,C) and $ANCB_{650-720}$ (B). Solid line represents a one-to-one relationship, R^2 = coefficient of determination, RMSE = root mean square error).

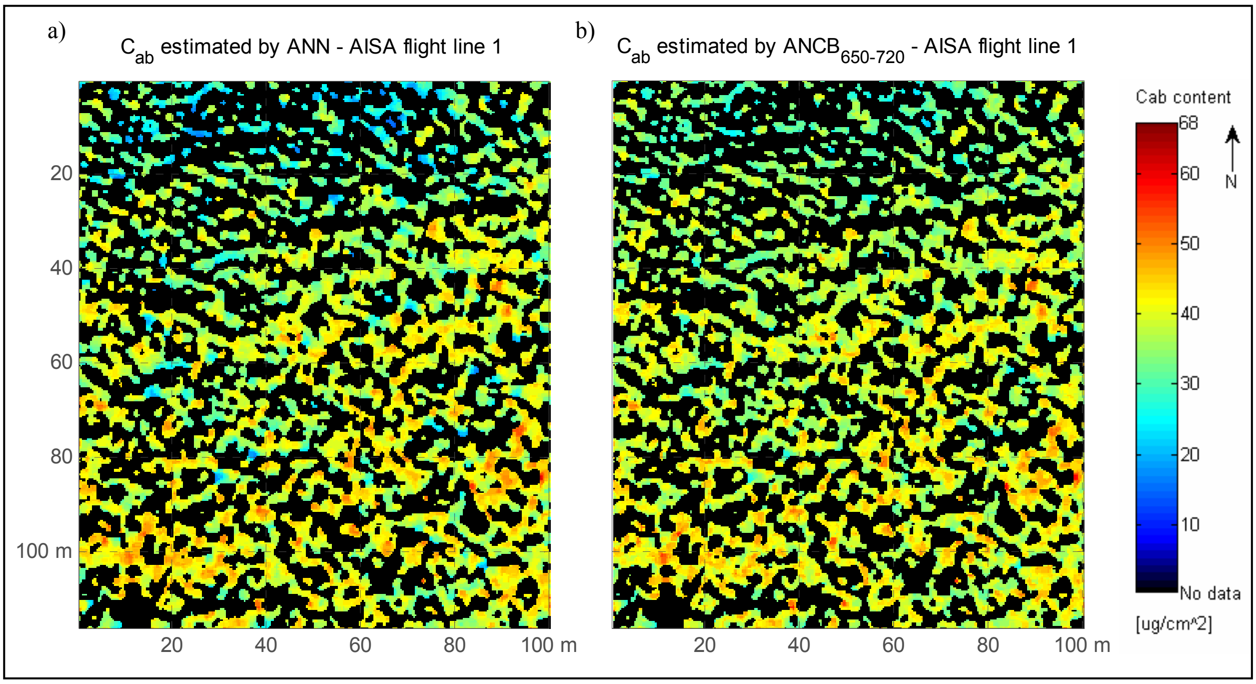

Figure 6. Maps of chlorophyll content (C_{ab}) retrieved from sunlit Norway spruce crown pixels of the AISA Eagle testing subset (flight line 1) by means of: a) an artificial neural network (ANN), b) the optical index ANCB$_{650-720}$. (Black colour indicates non-analysed shaded canopy and background).

we did not see significant differences in crown C_{ab} estimated from the different AISA flight lines. Therefore, one can conclude that both methods are most likely independent of the sun-object-sensor system set-up. Nevertheless, these methods also need to be tested for more geometrically heterogeneous ecosystems, e.g., mature Norway spruce stand with lower fractional canopy cover and significant portion of the understory, to support this conclusion.

3.2 Canopy chlorophyll estimation

The results of the chlorophyll estimation for the flight line 1 are shown in figure 6. A spatial pattern of these maps, similar for both retrieval methods, reveals lower C_{ab} values at the northern part and higher C_{ab} estimates at the southern part of the tested AISA subset. A comparable spatial pattern was found also for the flight lines 2 and 5 (maps not presented). Subtraction of the ANCB$_{65-720}$ chlorophyll map from the one produced using ANN resulted in an almost symmetrical Gaussian distribution (figure 7a). However, the high frequencies found for the positive differences mean that the ANCB$_{65-720}$ estimates are in general higher than the estimated obtained with ANN. Those positive differences mainly correspond with the lower C_{ab} content values, as already shown in the previous section of the crown level results. Looking at the absolute differences between the chlorophyll content retrieved by the ANCB$_{65-720}$ and the ANN, one can see that about 80% of all investigated pixels contained chlorophyll of the same concentration or plus/minus 2 $\mu g.cm^{-2}$ in maximum. Next 10% of the pixels were different by means of about plus/minus 4 $\mu g.cm^{-2}$ and the remaining 10% represented differences of plus/minus 6 to 10 $\mu g.cm^{-2}$ (see figure 7b). The mean difference was found to be around 1.8 $\mu g.cm^{-2}$. Similar results were also obtained when analyzing the testing subsets of the AISA flight lines 2 and 5. These findings support a hypothesis that the sun-object-sensor geometrical set-up does not play an important role when retrieving green foliar pigments from hyperspectral data of spatial resolution below 1 m.

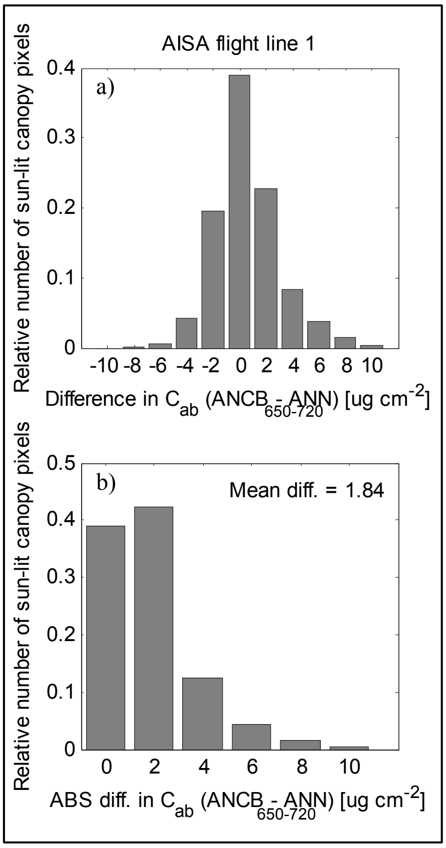

Figure 7. Histograms counting: a) differences, and b) absolute differences of the ANCB$_{650-720}$ and ANN chlorophyll maps produced for the testing subset of AISA flight line 1.

4. CONCLUSIONS

In this paper we have tested two physically based methods to retrieve the total chlorophyll content of a complex Norway spruce canopy from hyperspectral data acquired at very hight spatial resolution (pixel size of 0.4 m). The first method was based on an ANN inversion of a number of PROSPECT-DART model simulations. The second method exploited a physically inspired vegetation index trained with the radiative transfer simulations that were used for the ANN inversion. The following main conclusions can be drawn from results of this study:

- The C_{ab} content estimated by the ANN and the $ANCB_{650-720}$ approaches was consistent at crown and canopy level, and also comparable with the measured ground truth.
- The spatial pattern of C_{ab} estimates was similar for all tested flight directions.
- The continuum removal transformation was successfully used to minimize residual differences between the PROSPECT-DART simulated reflectance and reflectance of the real hyperspectral images.
- Pure vegetation reflectance is needed for an accurate C_{ab} estimation by means of both tested retrieving methods.
- The $ANCB_{650-720}$ does not seem to be sufficiently sensitive to retrieve low C_{ab} values.

5. REFERENCES

Curran, P.J., Dungan, J.L. and Peterson, D.L., 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the Kokaly and Clark methodologies. *Remote Sensing of Environment,* 76(3), pp. 349-359.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., de Colstoun, E. and Brown McMurtreyIII, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment,* 74(2), pp. 229-239.

Gascon, F., Gastellu-Etchegorry, J.P., Lefevre-Fonollosa, M.J. and Dufrene, E., 2004. Retrieval of forest biophysical variables by inverting a 3-D radiative transfer model and using high and very high resolution imagery. *International Journal of Remote Sensing,* 25(24), pp. 5601-5616.

Gastellu-Etchegorry, J.P., Martin, E. and Gascon, F., 2004. DART: a 3D model for simulating satellite images and studying surface radiation budget. *International Journal of Remote Sensing,* 25(1), pp. 73-96.

Gitelson, A.A., Vina, A., Verma, S.B., Rundquist, D.C., Arkebauer, T.J., Keydan, G., Leavitt, B., Ciganda, V., Burba, G.G. and Suyker, A.E., 2006. Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity. *Journal of Geophysical Research-Atmospheres,* 111(D8), pp. 1-13.

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J. and Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment,* 81(2-3), pp. 416-426.

Jacquemoud, S. and Baret, F., 1990. Prospect - a Model of Leaf Optical-Properties Spectra. *Remote Sensing of Environment,* 34(2), pp. 75-91.

Kokaly, R.F. and Clark, R.N., 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment,* 67(3), pp. 267-287.

le Maire, G., Francois, C. and Dufrene, E., 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing of Environment,* 89(1), pp. 1-28.

Malenovsky, Z., Albrechtova, J., Lhotakova, Z., Zurita-Milla, R., Clevers, J.G.P.W., Schaepman, M.E. and Cudlin, P., 2006. Applicability of the PROSPECT model for Norway spruce needles. *International Journal of Remote Sensing,* 27(23-24), pp. 5315-5340.

Malenovsky, Z., Ufer, C.M., Lhotakova, Z., Clevers, J.G.P.W., Schaepman, M.E., Albrechtova, J. and Cudlin, P., 2006. A new hyperspectral index for chlorophyll estimation of a forest canopy: Area under curve Normalized to Maximal Band depth between 650-725 nm. *EARSeL eProceedings,* 5, pp. 161-172.

Pokorný, R. and Marek, M.V., 2000. Test of accuracy of LAI estimation by LAI-2000 under artificially changed leaf to wood area proportions. *Biologia Plantarum,* 43(4), pp. 537-544.

Richter, R. and Schlapfer, D., 2002. Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: atmospheric/topographic correction. *International Journal of Remote Sensing,* 23(13), pp. 2631-2649.

Zarco-Tejada, P.J., Miller, J.R., Harron, J., Hu, B., Noland, T.L., Goel, N., Mohammed, G.H. and Sampson, P., 2004. Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies. *Remote Sensing of Environment,* 89(2), pp. 189-199.

6. ACKNOWLEDGEMENT

This work was supported by the ESA/PECS project No. 98029 and the Research Plan AV0Z60870520 of the Institute of Systems Biology and Ecology, Academy of Sciences of the Czech Republic, v.v.i. Authors would like to acknowledge the access to the CESNET METACentrum computing cluster facilities (Brno, Czech Republic), provided under the research intent MSM6383917201.