Derivation of forest leaf area index from multi- and hyperspectral remote sensing data

M. Schlerf, C. Atzberger, M. Vohland, H. Buddenbaum, S. Seeling & J. Hill
University of Trier, Remote Sensing Department, Trier, Germany

Keywords: hyperspectral, vegetation indices, forest, leaf area index

ABSTRACT: This study evaluated systematically linear predictive models between vegetation indices (VIs) derived from radiometrically corrected airborne imaging spectrometer (HyMap) data and field measurements of leaf area index (LAI) (n=40). Ratio-based and soil-line related broadband VIs were calculated after HyMap reflectance had been spectrally resampled to Landsat TM channels. Hyperspectral VIs involved all possible types of 2-band combinations of RVI and PVI. Cross-validation procedure was used to assess the prediction power of the regression models. Analyses were performed on the entire data set or on subsets stratified according to stand age. A perpendicular vegetation index (PVI) based on wavebands at 1088 nm and 1148 nm was linearly related to leaf area index (LAI) (R2=0.67, RMSE=0.69m2m-2 (21% of the mean); after removal of one forest stand subjected to clearing measures: R2=0.77, RMSE=0.54m2m-2 (17% of the mean)). The study demonstrates that for hyperspectral image data, linear regression models can be applied to quantify LAI with good accuracy. Best hyperspectral VIs in relation with LAI are typically based on wavebands related to prominent water absorption features. Such VIs measure the total amount of canopy water; as the leaf water content is considered to be relatively constant in the study area, variations of LAI are retrieved.

1 INTRODUCTION

The majority of studies for extracting biophysical variables from remotely sensed data have used empirical techniques to relate the spectral measurements to biophysical parameters (Treitz & Howarth, 1999). While much work has been done with agricultural crops, relatively little research has been done on investigating the relationships between forest leaf area index (LAI) and satellite data (Fassnacht et al., 1997). Most of the studies on forests have used broadband VIs (e.g. NDVI, RVI) to derive LAI of coniferous forest stands; however, with varying success (Running et al., 1986; Badhwar et al., 1986; Spanner et al., 1990; Chen & Cihlar; 1996; Franklin et al., 1997; Brown et al., 2000; Chen et al., 2002;). Few studies looked at the suitability of high spectral resolution remote sensing data to derive the LAI of coniferous forests. Gong et al. (1995) tested CASI data using univariate and multivariate regression and a VI based algorithm and found strong relationships with reasonable low errors. Danson & Plummer (1995) tested if the red edge inflection point (REIP) is primarily controlled by forest canopy LAI using helicopter-borne spectroradiometer data and found a strong non-linear correlation between plot LAI and the REIP for Sitka spruce *(Picea sitchensis)*. For the same tree species, forest LAI was recently related to the canopy REIP computed from imaging spectrometer data (CASI) with success (Lucas et al. 2000).

Despite the research undertaken it is often unclear whether the high spectral resolution data offer advantages over broadband data (Danson & Rollin, 1997).

The overall aim of the work was to evaluate the information content of hyperspectral remote sensing data for the estimation of forest LAI in comparison with broadband data. More specific objectives were:

(i) to determine spectral vegetation indices (VI) that are best suited for characterising LAI

(ii) to compare and contrast traditional broad-band and hyperspectral VIs in terms of basic statistical characteristics of the predicted relative to the observed LAI.

The research was restricted to Norway spruce as only forest stands of this single species occur at a sufficient number within the selected test site. Besides, coniferous forests do not show a saturation of VI with increasing LAI, whereas the saturation effect limits the usefulness of optical remote sensing for biophysical parameter mapping for mature deciduous forests (Chen et al., 2002).

2 MATERIAL AND METHODS

The Idarwald forest (49°45'N, 7°10'E) is located on the north-western slopes of the Hunsrück mountain ridge, Germany. It covers an area of about 7,500 ha. In 1999, 42 relatively homogenous Norway spruce *(Picea abies)* stands were identified at the Idarwald study area. Within these stands, 42 square 0.09 ha plots were established. The central location of each ground plot was determined with an accuracy of about ±5 m using a differential global positioning system (GPS). Signal distortion within the forest stands did not give as high accuracies as specified by the manufacturer. From a total of 42 stands, one was not covered by the HyMap imagery, and another one was thinned in the period between ground data collection and the HyMap overpass. These two stands were excluded from the analysis reducing the dataset to 40 stands.

At each plot, measurements of forest leaf area index (LAI) were carried out. LAI was estimated using a Li-Cor LAI-2000 instrument. The instrument was only operated under overcast sky conditions. Within each plot, Li-Cor measurements were taken at 10 regularly spaced points from which the average was calculated. Above the canopy measurements were taken in a nearby open field before entering the plots. Although Li-CoR measurements represent rather an effective plant area index instead of the real leaf area index due to the non-random distribution of leaves, no corrections were applied.

To improve LAI mapping it has been suggested to derive land cover specific VI-LAI relationships using land cover maps (Franklin et al., 1997; Chen et al., 2002). This stratification process is usually based on the species type. As in the present study just one single species is considered, it was decided to stratify the dataset according to stand age (Table 1).

Table 1. Resulting subsets after stratification according to forest stand age

Subset	Designation	Abbreviation	Age	n
1	total (pooled dataset)	t	10-148	40
2	medium to old	mo	30-148	35
3	old	o	80-148	17
4	medium	m	30-79	18
5	young	y	10-29	5

The hyperspectral HyMap sensor (Integrated Spectronics, Australia) was flown over the Idarwald test site at 17th July 1999. The data were recorded at 13:00 hrs CET at an average flying height of 1980 m above ground level and free of cloud cover. The resulting ground resolution was about 5 m with a full scene covering approximately 3 km x 10 km. From the original data cube with 128 bands between 400 and 2500 nm, 14 HyMap spectral channels with high noise were identified as bad bands and removed from the dataset. Consequently, subsequent analyses were restricted to the

M. Schlerf, C. Atzberger, M. Vohland, H. Buddenbaum, S. Seeling & J. Hill

remaining 114 bands. To correct the view angle effect, an across track illumination correction was applied to each spectral band independently (ENVI 3.4, Research Systems). For this purpose, a second-order polynomial was fitted to the data. Based on the fitted polynomials, a normalization procedure was applied. The radiometric correction steps applied to the HyMap data consisted of the atmospheric correction, the correction for illumination effects caused by topography, and the in-flight calibration of the HyMap sensor system (Hill & Mehl, 2003). The image data were geometrically corrected using the parametric geocoding software PARGE (Schläpfer et al., 1998). For this purpose, the required DEM with an original pixel size of 20 m was resampled to 5 m. Mean reflectance spectra were extracted from the image data at a fixed radius around the central position of each test plot (GPS measurement).

Table 2. Broadband and hyperspectral vegetation indices investigated in this study.
ρ=reflectance, TM=Thematic Mapper

Name	Abbrev-iation	Equation	Reference
Ratio vegetation index	RVI	$$RVI = \frac{\rho_{TM4}}{\rho_{TM3}}$$	Pearson & Miller, 1972
Normalised difference vegetation index	NDVI	$$NDVI = \frac{\rho_{TM4} - \rho_{TM3}}{\rho_{TM4} + \rho_{TM3}}$$	Rouse et al., 1974
Perpendicular vegetation index	PVI	$$PVI = \frac{\rho_{TM4} - a\rho_{TM3} - b}{\sqrt{1 + a^2}}$$ $a = 0.9, b = 0.1$	Richardson & Wiegand, 1977
Transformed soil-adjusted vegetation index	TSAVI	$$TSAVI = \frac{a(\rho_{TM4} - a\rho_{TM3} - b)}{a\rho_{TM4} + \rho_{TM3} - ab}$$ $a = 0.9, b = 0.1$	Baret et al., 1989
Mid-infrared vegetation index	MVI	$$MVI = \frac{\rho_{TM4}}{\rho_{TM5}}$$	Fassnacht et al., 1997
Greeness vegetation index	GVI	$$GVI = -0.2848\rho_{TM1} - 0.2435\rho_{TM2}$$ $$-0.5436\rho_{TM3} + 0.7243\rho_{TM4}$$ $$+0.0840\rho_{TM5} - 0.1800\rho_{TM7}$$	Christ and Cicone, 1984
Narrowband RVI	RVI$_{Bd1,Bd2}$	$$RVI_{Bd1,Bd2} = \frac{\rho_{Bd1}}{\rho_{Bd2}}$$	Thenkabail et al., 2000
Narrowband PVI	PVI$_{Bd1,Bd2}$	$$PVI_{Bd1,Bd2} = \frac{\rho_{Bd1} - a\rho_{Bd2} - b}{\sqrt{1 + a^2}}$$	

Both hyperspectral and broadband VIs were computed from the radiometrically corrected HyMap imagery (Table 2). Ratio-based and orthogonal broadband VIs were calculated after the HyMap data had been resampled to Landsat TM spectral bands involving the appropriate Landsat TM 5 filter functions; the ratio-based VIs were RVI and NDVI; the orthogonal VIs were PVI and TSAVI;

additional VIs were MVI, and GVI. Hyperspectral (narrowband) indices were computed for RVI and PVI involving all possible two-band combinations of 114 channels (=12,996 band combinations). PVI requires site-specific soil line slopes (a) and intercepts (b). As no soil spectral data was available, standard values (a=0.9; b=0.1) were used. It was assumed that the soil line concept, originally defined for the red-nIR feature space, can be transferred into other spectral domains (Thenkabail et al., 2000). Hence, it was supposed that soil lines exist between all wavebands.

Linear regression was employed to evaluate the relationships between biophysical stand variables and VI. For conifers it has been shown, that linear regression models seem to be appropriate as saturation occurs only at relatively high densities (Chen & Cihalar, 1996; Chen et al., 2002). The cross-validation procedure was used to validate the regression models. The cross-validated RMSE is a good indicator of the accuracy of the model in predicting unknown samples.

Table 3. Summary statistics for forest stand variables (n=40)

	Mean	Standard deviation	Maximum	Minimum	Range
LAI [m^2 m^{-2}]	3.24	0.97	5.47	1.66	3.81

Table 4. Linear correlation between forest LAI and HyMap reflectance resampled to TM spectral bands (n=40); ** correlation coefficient significant at P<0.01, * correlation coefficient significant at P<0.05

	TM3	TM4	TM5
LAI	0.36*	0.76**	0.33*

3   RESULTS AND DISCUSSION

Although forest stands from very young to very old stands were probed, the variation of LAI values is rather moderate (Table 3). LAI was strongest correlated with TM4, less correlated with TM3, and least correlated with TM5 (Table 4). Correlations in all bands were positive. Correlation spectra were computed for reflectance and first derivative of reflectance (not shown). The strength of the relations with LAI generally decreases from nIR to mIR wavebands and is greater for first derivative reflectance spectra in opposition to reflectance spectra. The nIR region of reflectance spectra reveals the highest correlations followed by the green peak, the mIR region, and the chlorophyll absorption features in the red and blue wavebands.

An inverse relationship can be observed in Figure 1, where LAI is plotted against stand age for the probed Norway spruce stands of Idarwald test site. The inverse relation between LAI and stand age seems to be related to the fact that during stand development, crowns of individual trees expand and increase utilization of available growing space. The point at which crowns of different trees begin to interact is considered as being the peak LAI after which a rapid decrease takes place due to competition between the individual trees (Vose et al. 1994). It has been shown that stands of slow growing species (*pinus contora*) reached its maximum LAI at age 40 and that it lasted for about 30 years (Long & Smith, 1992; cited from Vose et al., 1994). In Figure 1, a peak LAI can be assumed at an age of 20. From an age of 20 onwards, a gradual decline in LAI from up to age 150 is evident. Maximum variation of LAI occurs at age 60-70. Obviously, some stands have been thinned lately whereas others have been thinned long time ago. After an increase of LAI up to age 100 a decrease of LAI is caused by gaps related to logging measures. It can be concluded that the observed age course of Norway spruce based on the probed sample is a result of both natural circumstances and actual management practices. It is also clear from Figure 1, that LAI can vary considerably within a single age class and thus, information on LAI can not simply be derived from an age class map.

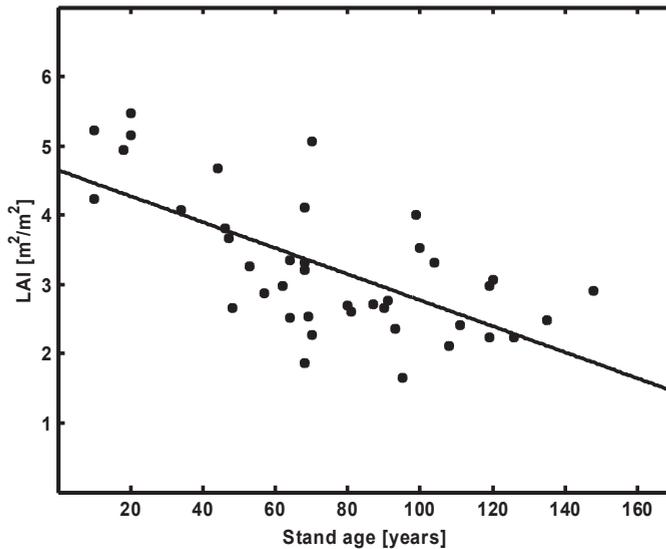M. Schlerf, C. Atzberger, M. Vohland, H. Buddenbaum, S. Seeling & J. Hill

Figure 1. LAI as a function of stand age for Norway spruce stands at Idarwald test site (n=40).

To determine optimal narrowband VIs, coefficients of determination (R^2) between all possible two-band VIs and LAI were computed. The results are illustrated in 2D-correlation plots (Figure 2). Based on the R^2 values in the 2D-correlation plots, band combinations that form the best indices were determined. These were considered as optimal indices and were named RVI_opt and PVI_opt. Up to three best performing indices were considered when they occurred in different wavelength regions. For example, in Figure 2, the indices that have highest correlation with LAI were extracted from the 2D-correlation plot range above 0.80; the three most dominant indices occur in the regions around 1346nm/1207nm, 1134nm/920nm, and 1802nm/1163nm. The band positions were then tabulated in Table 5. For the best performing narrowband index, cross-validated R^2 and RMSE were computed (Table 6).

The first step in the analysis of broadband VIs was to compare ratio to soil-adjusted VIs and to compare VIs based on Vis and nIR reflectance to those based on mIR and nIR reflectance. Broadband ratio VIs (RVI and NDVI) generally show relatively low values of R^2 and relatively high values of RMSE for all subsets. The soil-adjusted broadband models (PVI and TSAVI) perform significantly better for the pooled data set; obviously, background effects related to soil and litter were reduced. Whereas younger stands reveal a denser canopy with little background contribution to the signal, in older stands with gaps and a more open canopy the background may have a larger influence on reflectance. Hence, when the total age dynamic is considered, TSAVI and PVI perform better than RVI or NDVI but when age classes are considered separately no performance increase is observed. The findings of Brown et al. (2000), that soil-adjusted VIs compared to the RVI have a decreased sensitivity to forest LAI, can not be confirmed.

Broadband MVI is closer related to LAI than broadband RVI and NDVI (subset m and mo). The mIR band in combination with the nIR band seems to contain more information relevant to the characterisation of forest canopies than the combination of red and nIR bands. Recently, a closer relation of forest LAI to radiation in the mIR than to radiation in the Vis was found by Boyd et al. (2000) for tropical vegetation; the authors put forward that this could be the case also with boreal forests. The results found in the present study also support the suggestion by Fastnacht et al. (1997) that mIR bands may improve LAI estimation, particularly in more open forest stands.

Table 5. Best narrowband RVI and PVI derived from 2D-correlation plots for different subsets. Subset y was excluded from the analysis due to the small number of samples

	t (10-148 years)		mo (30-148 years)		o (80-148 years)		m (30-79 years)	
	λ [nm]	r^2	λ [nm]	r^2	λ [nm]	r^2	λ [nm]	r^2
RVI_opt	918/965	.65	896/965	.68			1346/1207	.85
	850/680	.55	1802/1043	.65	747/807	.60	1134/920	.85
	1165/750	.60	1346/1119	.65			1802/1163	.80
PVI_opt	1088/1148	.70	895/1134	.64	1445/2060	.56	918/1134	.83
	1148/807	.70					1320/1220	.82
							2220/457	.72

In the second step of the analysis, the broadband VIs were compared to the hyperspectral VIs to see if hyperspectral indices improve the prediction accuracy. All best narrowband RVI (RVI_opt) and PVI (PVI_opt) perform better than the corresponding broadband VIs. RMSE values of about 1 m^2 m^{-2} or larger for regression models between broadband RVI and LAI (subsets am, a, and m) decrease to values as low as 0.5 m^2 m^{-2} for the optimal narrowband RVI. The improvement of the narrowband models compared to the broadband models is not that distinct in the case of the PVI. Over all age classes PVI_opt performs better than RVI_opt; for subsets mo, m, and o, however, RVI_opt shows lower RMSE values than PVI_opt.

To summarise, it can be concluded that the hyperspectral dataset contains more information relevant to the estimation of forest LAI than multispectral data. However, this does not always seem to be valid: For instance, Broge & Mortensen (2002) came to the conclusion, that hypespectral VIs derived from field spectral measurements were not better at estimating green crop area index (a variable related to LAI) than traditional broadband VIs. For old stands of age 80 or more (subset o), relatively poor relationships were found in particularly for the broadband VIs. Also other studies reported problems with old stands that have been ascribed to shadow effects and a relatively dark background in the nIR (Spanner et al., 1990).

In the scatter plot between PVI_opt and LAI (Figure 3), even at high values of LAI no saturation is evident. A closer look reveals an outlier at position 4.1/-0.104. The corresponding forest stand of age 34 was identified in the image data. Whereas the 1999 image showed no abnormality in reflectance, in recently acquired HyMap data of 2003, a striped pattern was detected. From the spectral reflectance properties, the image pixels representing stripes could be identified as a mixture between tree crowns and forest litter. The striping pattern was caused by aisles that had been cut into the forest between the image acquisition and field measurements. Forest aisles allowed for the employment of harvesters in order to remove trees that had been exposed to game bite (personal communication of the responsible forest official, G. Womelsdorf, 25.02.2004).
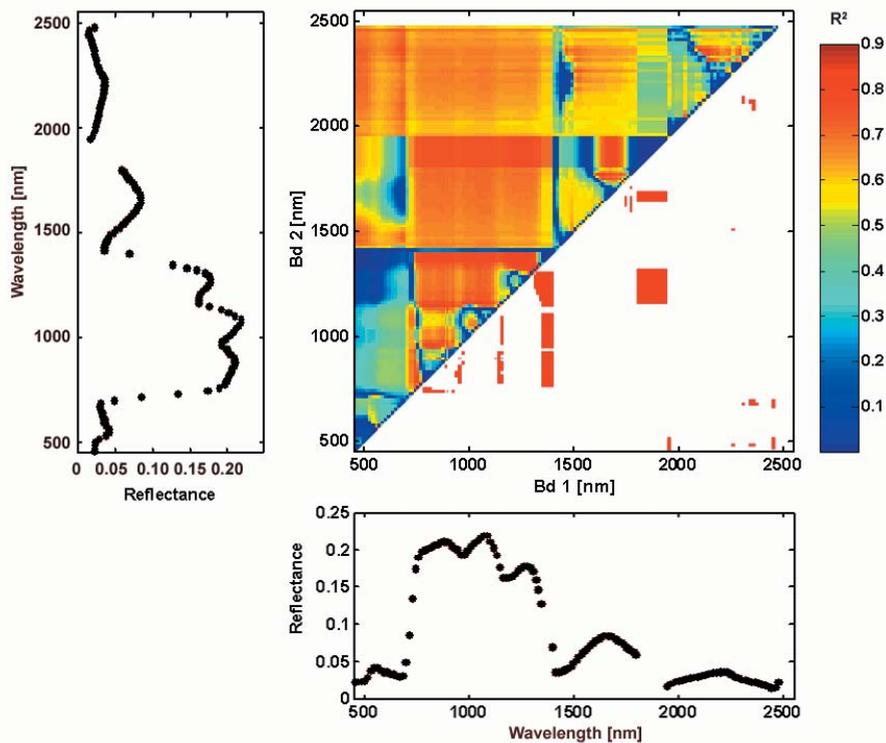
Figure 2. 2D-correlation plot that shows the correlation (R^2) between LAI and narrowband RVI values (subset m). The matrix is symmetrical, therefore just values above the diagonal are displayed. Below the diagonal, band combinations are marked in red where $R^2 > 0.75$. The displayed average reflectance spectrum of all measured forest plots eases the interpretation of the 2D-correlation plot.

Table 6. Cross-validated R^2 (first line) and cross-validated RMSE (second line) for linear regression between broadband and hyperspectral vegetation indices and forest stand LAI. The best VIs are typed in bold. Subset y was excluded from the analysis due to the small number of samples.

	Subset	t	mo	m	o
	n	*40*	*35*	*17*	*18*
Broadband VI	RVI	.40	.42	.40	.17
		1.23	.92	1.05	1.67
	NDVI	.43	.44	.44	.17
		1.13	.86	.95	1.72
	PVI	.57	.29	.23	.22
		.87	1.29	1.7	1.7
	TSAVI	.61	.36	.30	.24
		.80	1.06	1.35	1.55
	MVI	.38	.54	.58	.35
		1.3	.71	.71	1.83
	GVI	.58	.31	.24	.24
		.85	1.2	1.62	1.55
Hyper-spectral VI	RVI	.62	.64	.77	.57
	_opt	.78	.58	.45	.57
	PVI	.67	.45	.41	.52
	_opt	.69	.86	1.02	.58

Derivation of forest leaf area index from multi– and hyperspectral remote sensing data
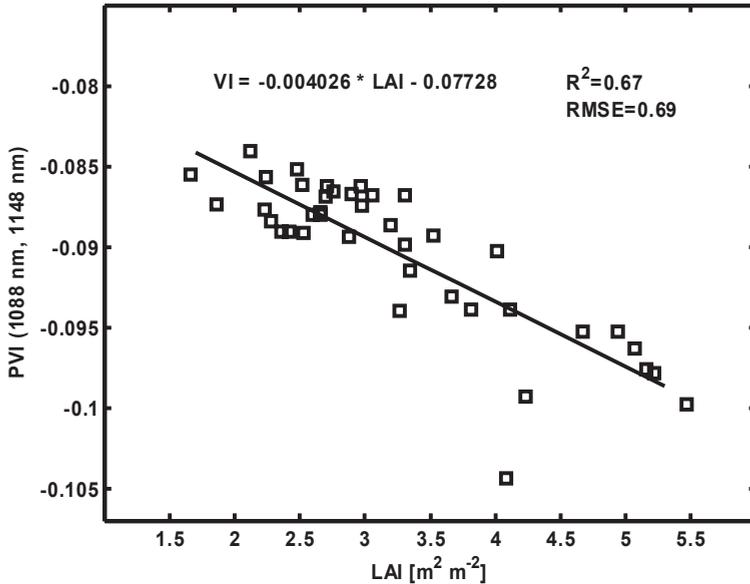
Figure 3. Linear regression between best narrowband PVI and LAI. Values of R^2 and RMSE are cross-validated. When the outlier at position 4.1/-0.104 is removed, cross-validated R^2 increases to 0.77 and RMSE decreases to 0.54 m^2 m^{-2}.

4   CONCLUSIONS

The following conclusion can be drawn from the research:
- Forest leaf area index (LAI) and crown volume (VOL) can be estimated with good accuracy from hyperspectral remote sensing data,
- Orthogonal compared to ratio VIs are better suited to characterise forest LAI and VOL,
- Hyperspectral data contains more information relevant to the estimation of the forest stand variables LAI and VOL than multispectral data.

M. Schlerf, C. Atzberger, M. Vohland, H. Buddenbaum, S. Seeling & J. Hill

REFERENCES

Badhwar, G.D.; Macdonald, R.B.; Hall, F.G. 1986. Spectral characterization of biophysical characteristics in a boreal forest: Relationship between Thematic Mapper band reflectance and leaf area index for aspen. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24: 322-326.

Baret, F.; Guyot, G.; Major, D.J. 1989. Crop biomass evaluation using radiometric measurements. *Photogrammetria (PRS)*, 43: 241-256.

Boyd, D.S.; Wicks, T.E.; Curran, P.J. 2000. Use of middle infrared radiation to estimate the leaf area index of a boreal forest. *Tree Physiology*, 20: 755-760.

Broge, N.H.; Mortensen, J.V. 2002. Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. *Remote Sensing of Environment*, 81: 45-57.

Brown, L.; Chen, J.M.; Leblanc, S.G. 2000. A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: An image and model analysis. *Remote Sensing of Environment*, 71: 16-25.

Chen, J.M.; Cihlar, J. 1996. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment*, 55: 153-162.

Chen, J.M.; Pavlic, G.; Brown, L. 2002. Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurments. *Remote Sensing of Environment*, 80: 165-184.

Christ, E.P. & Cicone, R.C. 1984. A physically-based transformation of Thematic Mapper data - the TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22: 256-263.

Danson, F. M.; Plummer, S. E. 1995: Red-edge response to forest leaf area index. *International Journal of Remote Sensing*, 16: 183-188.

Danson, F.M.; Rollin, E.M. 1997. Information content of high spectral resolution reflectance data for forest canopies. Aus: Guyot, G.; Phulpin (Eds.): *Physical measurements and signatures in remote sensing.* Rotterdam (Balkema): 457-462.

Fassnacht, K.S.; Gower, S.T.; MacKenzie, M.D. 1996. Estimating the leaf area index of north central Wisconsin forests using the Landsat Thematic Mapper. *Remote Sensing of Environment*, 61: 229-245.

Franklin, S.E.; Lavigne, M.B.; Deuling, M.S. 1997. Estimation of forest leaf area index using remote sensing and GIS data for modelling net primary production. *International Journal of Remote Sensing*, 18: 3459-3471.

Gong, P.; Pu, R.; Miller, J.R. 1995. Coniferous forest leaf area index estimation along the Oregon transect using Compact Airborne Spectrographic Imager data. *Photogrammetric Engineering & Remote Sensing*, 61: 1107-1117.

Hill, J.; Mehl, W. 2003. Georadiometrische Aufbereitung multi- und hyperspektraler Daten zur Erzeugung langjähriger kalibrierter Zeitreihen. *Photogrammetrie-Fernerkundung-Geoinformation*, 1/2003: 7-14.

Lucas, N. S.; Curran, P. J.; Plummer, S. E.; Danson, F. M. 2000: Estimating the stem carbon production of a coniferous forest using an ecosystem simulaton model driven by the remotely sensed red edge. *International Journal of Remote Sensing*, 21: 619-631.

Pearson, R.L. & Miller, L.D. 1972. Remote mapping of standing crop biomass for estimation of the productivity of the short-grass prairie, Pawnee National Grasslands, Colorado. *Proceedings of the 8th International Symposium on Remote Sensing of Environment, ERIM International*: 1357-1381.

Richardson, A.J. & Wiegand, C.L. 1977. Distinguishing vegetation from soil background information. *Photogrammetric Engineering & Remote Sensing*, 43: 1541-1552.

Rouse, J.W.; Haas, R.H.; Schell, J.A. 1974. Monitoring the vernal advancement of retrogradiation of natural vegetation. In: *NASA/GSFC, Type III, Final Report*. Greenbelt, MD, USA: 1-371.

Running, S.W.; et al. 1986. Remote sensing of coniferous forest leaf area. *Ecology*, 67: 273-276.

Schläpfer D.; Schaepman M.E.; Itten K.I. 1998. PARGE: Parametric Geocoding Based on GCP- Calibrated Auxiliary Data. In: *SPIE Int. Symp. on Opt. Sc., Eng. and Instr.* San Diego, CA: 334-344.

Spanner, M.A.; Pierce, L.A.; Peterson, D.A. 1990. Remote sensing of temperate coniferous forest leaf area index - The influence of canopy closure, understory vegetation and background reflectance. *International Journal of Remote Sensing*, 11: 95-111.

Thenkabail, P.S.; Smith, R.B.; De Paw, E. 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71: 158-182.

Treitz, P.M.; Howarth, P.J. 1999. Hyperspectral remote sensing for estimating biophysical parameters of forest ecosystems. *Progress in Physical Geography*, 23: 359-390.

Vose, J.M.; Dougherty, P.M.; Long, J.N. 1994. Factors influencing the amount and distribution of leaf area of pine stands. *Ecological Bulletins*, 43: 102-114.