EVALUATION OF LEAF AREA INDEX AND DRY MATTER PREDICTIONS FOR CROP GROWTH MODELLING AND YIELD ESTIMATION BASED ON FIELD REFLECTANCE MEASUREMENTS

Heike Gerighausen[1], Holger Lilienthal[1], Thomas Jarmer[2], and Bastian Siegmann[2]

1. Julius Kühn-Institut, Institute of Crop and Soil Science, 38116 Braunschweig, Germany; {heike.gerighausen / holger.lilienthal}(at)julius-kuehn.de
2. University of Osnabrueck, Institute for Geoinformatics and Remote Sensing, 49076 Osnabrueck, Germany; {tjarmer / bsiegmann}(at)igf.uni-osnabrueck.de

ABSTRACT

Leaf area index (*LAI*) and above ground biomass dry matter (*DM*) are key variables for crop growth monitoring and yield estimation. High prediction accuracies of these parameters are a vital prerequisite for sophisticated yield projections. The aim of the study was to examine the predictive ability of partial least squares regression (PLSR) for *LAI* and *DM* retrieval from hyperspectral (EnMAP), superspectral (Sentinel-2), and multispectral (Landsat 8, RapidEye) remote sensing data based on field reflectance measurements. Data was acquired from several crop types (wheat, rye, barley, rape, potato, sugar beet) during field campaigns in three different regions of Germany between the years 2011 and 2014. The field reflectance measurements were resampled to match the different spectral resolutions. Continuous reflectance and resampled data were transformed using five spectral pre-processing techniques. Continuous data were used for comparison and served as best case scenario. The predictive ability of the PLSR models for *LAI* and *DM* was examined with respect to the spectral resolution and the pre-processing techniques. To verify whether the composition of the data set had an effect on prediction quality, the entire data set (global) was divided in sub data sets (local) with respect to the region of acquisition, the year of acquisition and the crop type. Statistical models of the local data sets were compared with those based on the global data set. Generally, models were assessed with two validation strategies.

R^2 of the global PLSR models based on continuous field reflectance measurements and independent validation varied from 0.74 to 0.79 (*LAI*), and from 0.76 to 0.87 (*DM*). Root mean square error ranged between 0.70 and 0.74 $m^2\ m^{-2}$, and between 1.64 and 2.56 $t\ ha^{-1}$, respectively. There was no pre-processing method which consistently improved model performance. However, results pointed out that the technique should be chosen with respect to the sensor and the parameter of interest. Models based on hyperspectral information performed generally best. Prediction error increased with the superspectral sensor configuration by only 3% for *LAI*, and 16% for *DM*. Multispectral sensor configurations caused the prediction error to rise by up to 22% and 54%, respectively. A stratification into local data sets according to date of acquisition, sampling region and crop type partially increased the prediction performance. Cross-validation yielded higher prediction errors than independent validation in most cases.

KEYWORDS

Leaf area index, above ground biomass dry matter, field spectroscopy, PLSR, pre-processing.

INTRODUCTION

Crop yield estimations are an important source of information for food security and the assessment of the agricultural market situation on a national and a global scale (1,2). Further, yield maps are an important component of site-specific farming to evaluate the effect of management strategies (e.g., 3,4,5,6), although multi-temporal analysis of yield maps alone did not prove successful for the delineation of management zones (7,8). Remote sensing systems can provide spatially and

DOI: 10.12760/02-2015-2-06

spectrally differentiated information over large areas at regular intervals. Therefore, they may offer essential supplementary information to conventional yield statistics and to precision agriculture (9).

The most simple and straightforward approach to estimate yield from remote sensing images is the direct correlation between vegetation indices and yield (e.g., 10), but numerous studies pointed out the shortcomings of this technique (e.g., 11,12). A more advanced methodology is the assimilation of biophysical key variables such as Leaf Area Index (*LAI*) and above ground biomass into crop growth models (e.g., 13,14,15). Both variables can be retrieved from remote sensing data (e.g., 16,17,18) and it has been demonstrated that their assimilation into crop growth models can improve yield estimations (e.g., 19). In any case, a high quality of the input variables is a vital prerequisite for a sound yield monitoring and forecasting.

In the majority of studies, *LAI* and above ground biomass have been derived from remote sensing data using parametric empirical relationships between *in situ* measurements of the mentioned parameters and vegetation indices (e.g., 17,20,21). The use of spectral vegetation indices, however, has limitations for high biomass and high *LAI* values, since vegetation indices reach a saturation level for *LAI* greater than 2 (22,23). This problem may be mitigated to a certain extent by using indices with more than three bands or by applying other curve fitting types (24). For *LAI* retrieval, Rivera et al. (24) stated that including bands in the shortwave infrared region eliminated the saturation effect almost completely.

Beyond that, the exploitation of the full-resolution spectra using non-parametric regression algorithms has proven successful for the retrieval of biophysical variables, especially for hyperspectral systems. Non-parametric algorithms can be divided into linear (e.g., partial least squares regression (PLSR), principal component regression (PCR)) and nonlinear regression methods (e.g., support vector regression (SVR), random forest regression (RFR), neural networks (NNs), kernel ridge regression (KRR) and Gaussian processes regression (GPR), the latter also referred to as machine learning regression algorithms (MLRAs)) (25). PLSR became one of the most popular non-parametric methods for biophysical parameter retrieval, not least due to its availability in various software packages, its ease of use and fast processing times (26,27). PLSR was successfully adopted in a variety of vegetation studies and outperformed not only index based parametric approaches. Hansen and Schjoerring (28), for example, showed that PLSR using the information from all wavelengths significantly improved predictions for biomass and leaf nitrogen compared to a simple regression with traditional broad-band and narrow-band indices. For *LAI* and other crop variables, PLSR was at least as good as deploying parametric approaches and narrow-band indices. Pimstein et al. (29) achieved lower errors for dry biomass, water and nitrogen content of wheat using PLSR instead of a parametric approach with vegetation indices. Darvishzadeh et al. (30) tested several statistical techniques for *LAI* estimation in heterogeneous grassland. They also stated that estimations further improved when PLSR was applied. Siegmann and Jarmer (31) compared the capability of PLSR with two nonlinear regression methods, SVR and RF, to estimate winter wheat *LAI* from hyperspectral data. They showed that PLSR obtained clearly better results than SVR or RF if independent validation was applied, and it performed less sensitive to the applied validation techniques. The predictive power of PLSR models, however, may vary depending on the pre-processing technique (32), whereas pre-processing is supposed to remove spectral noise, normalize systematic changes or perform scatter correction. Although designated to efficiently handle highly correlated multiple spectral bands, PLSR may also gain comparable prediction accuracies if confined to a few specific wavelengths regions (33).

In recent years, and with the forthcoming of the European Sentinel-2 and Sentinel-3 mission, increased attention has been paid to the second group of non-parametric regression algorithms that are referred to as machine learning regression algorithms (25,27,34). Verrelst et al. (25) demonstrated that GPR is a powerful tool for vegetation parameter retrieval including *LAI*, outperforming other MLRAs (NN, SVR, KRR, GPR) in the majority of the test cases. According to Rivera Caicedo et al. (34), GPR and KRR proved most robust and were the best performing regression algorithms for *LAI* and leaf chlorophyll content. In their study, nonlinear regression algorithms (NN, KRR, GPR) generally excelled linear regression algorithms (PLSR, PCR) in terms of accuracy, bias, and

robustness. An urgent open question is how well these methods perform with large data sets (34). Verrelst et al. (27) also confirmed that MLRAs from the kernel family achieved a higher accuracy than parametric and linear non-parametric regression methods such as PLSR and PCR, whilst the latter were invincible in terms of processing speed. The authors further tested a physical approach relying on a lookup-table (LUT)-based inversion of a canopy radiative transfer model (RTM). Prediction accuracy of RTMs was similar to two-band indices and is in agreement with findings by Darvishzadeh et al. (30). Owing to their physical nature, RTMs are more generally applicable, and constitute an alternative strategy for biophysical parameter retrieval (e.g., 35,36). But the inversion of a canopy RTM is a complex task, because it is mostly under-determined (37) and ill-posed (38). Numerous strategies have been proposed to address the ill-posed problem including iterative numerical optimization, look-up tables (LUT), or hybrid approaches. LUT-based inversion routines are widely applied using various regularization strategies to improve the robustness of the LUT-based inversion, for example the use of a priori information, modified cost functions, or spatial or temporal constraints. A comprehensive overview of existing procedures is given in Verrelst et al. (27). The selection of the regularization method but also the choice of the RTM model and the applied inversion procedure strongly influence the performance and the partially expensive computational effort of this strategy of information retrieval (26,38).

In our study, we aimed at evaluating the performance of PLSR based on field reflectance measurements and *in situ* observations to predict *LAI* and above ground biomass from remote sensing data. Unlike a multiplicity of studies (e.g., 20,21,27,34,39), our research is based on field measurements from three different geographic regions and six crop cultivars acquired at various dates during the growing seasons in four years. Data from different regions, acquisition dates and crop types were used to ensure the applicability of the empirical approach on a regional scale. The analysis was executed in preparation of image based retrieval of *LAI* and biomass for yield estimations by means of linking remote sensing data and crop growth modelling.

Therefore, we firstly (i) tested the influence of pre-processing methods appropriate for data normalization and removal of systematic errors. Further, we aimed at identifying the best suited earth observation mission for our approach, including hyperspectral (EnMAP), superspectral (Sentinel-2), and multispectral (Landsat 8, RapidEye) systems. Hence, we secondly (ii) assessed the predictive power of the PLSR models when wavelength information was adjusted to five spectral sensor configurations. Thirdly (iii), we addressed potential effects on model prediction accuracy due to the heterogeneous composition of the data set. For this reason, the data set was divided in sub-groups and model performance was analysed with respect to the crop type, the year of acquisition and the region of acquisition. Analyses for items (i) and (ii) were performed on continuous and on resampled data whereby results based on continuous data served as best case scenarios. Effects due to database composition were explored with the continuous reflectance measurements only.

METHODS

Field reflectance measurements and sampling

The data set consists of 364 reflectance measurements of winter wheat, winter rye, winter barley, winter rape, sugar beet and potato which were acquired during 16 field operations between March and July in the years 2011 to 2014. Details on sampling dates and number of measurements are given in Table 1. The data was sampled in three different regions in Germany (Figure 1). Sampling was carried out on agricultural fields near the city of Köthen (R1) in the federal state of Saxony-Anhalt during three field campaigns. In Braunschweig (R2), Lower-Saxony, samples were acquired on test sites of the Julius Kühn-Institut. Near Baruth (R3), a city in the federal state of Brandenburg, sampling took place on a breeding nursery. On the experimental plots in Braunschweig and Baruth, different fertilisation treatments and irrigation schemes were applied. Partially, different genotypes were cultivated.

Canopy reflectance was recorded from 0.25 m² plots with field spectrometers (ASD Fieldspec Pro, SVC HR1024), covering a spectral range from 350 nm to 2500 nm. All reflectance curves were

resampled to 1 nm and smoothed using a Savitzky Golay filter (40) with a filter size of 32 and a smoothing polynomial of 4. Settings are based on experiences of a previous study (Gerighausen, unpublished). Afterwards, pre-processing transformations were applied if applicable (cp. statistical analysis). Due to atmospheric absorption from 1355 nm to 1425 nm and from 1785 nm to 1995 nm, noisy bands were removed.

Exactly at the location of the canopy reflectance measurements of each plot, *LAI* was measured either with the *LAI*-2000 plant canopy analyser (LI-COR, Inc.) or the Sunscan Canopy Analysis System SS1-Com-R4 (Delta-T Devices, Ltd). The amount of dry matter (*DM*) was determined for each sample plot by destructive techniques. Above ground biomass was cut and weighed after oven-drying at 60 °C for at least 24 h.

Table 1: Dates of sampling operations and number of measurements in the years 2011 to 2014.

Crop type	Sampling date with number of measurements in brackets (*LAI/DM*)	Sampling year	Sampling locations
wheat	8th May (*57/57*)	2011	R1
	24th(*67/67*), 25th May (*8/8*)	2012	R1
	25th April (*10/10*), 9th July (*14/14*)	2013	R2
	4th June (*3/3*)	2014	R2
rye	14th May (*5/5*),	2012	R2
	24th April (*15/15*), 28th May (*16/-*), 5th (*15/15*), 6th (*32/-*), 19th June (*15/15*), 9th July (*15/15*)	2013	R2
	15th May (*38/-*)	2013	R3
	4th June (*3/3*)	2014	R2
barley	4th June (*3/3*)	2014	R2
rape	13th March (*20/20*), 6th July (*6/6*)	2014	R2
potato, sugar beet	6th June (*10/10*), 4th July (*12/12*)	2014	R2

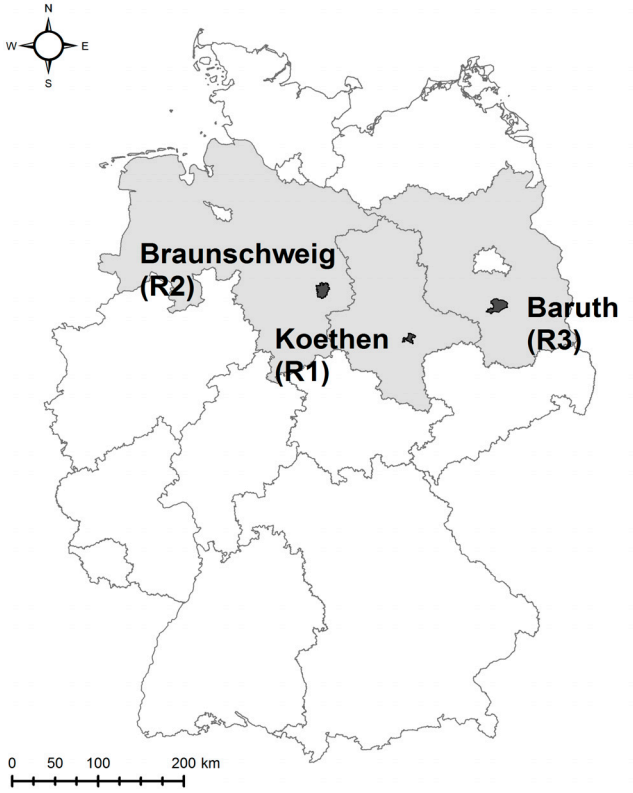

Figure 1: Location of the study areas in Germany.

Statistical analyses

The quantitative relationships between canopy reflectance and crop parameters were established by means of PLSR. Statistical models were set up for *LAI* and *DM*, respectively, for global and local data sets as described below. The analyses were performed with The Unscrambler® (CAMO Software AS) applying the PLSR according to Martens and Naes (41). PLSR is a bilinear multivariate modelling approach that projects information of the predictor (X) and the response (Y) variables onto a few so-called latent variables or factors. This way, the problem of variable selection is reduced. A detailed description of the PLSR algorithm is given by Wold et al. (42).

Model validation was performed by applying two different techniques. First, an independent validation (ival) was conducted by splitting the data into a calibration and a validation data set by systematic sampling. The data set was sorted according to *LAI* and *DM*, respectively, and every 2nd sample was chosen for calibration purposes. The rest was retained for validation only. This procedure was chosen to ensure that the validation data set is within the data range of the generated model. Independent validation can be considered the most reliable method to assess model accuracy, as samples are not involved in model calibration. However, model accuracy may be overestimated if samples are not entirely independent due to spatial autocorrelation. This problem is largely reduced in the data set due to a multiplicity of individual trial plots (232 out of 364) with different genotypes and treatments. Beyond that, model accuracy may be affected by the way the data sets are divided. In order to discover such undesirable effects, a random procedure using leave-one-out cross-validation (cv) was additionally performed on all data sets for comparison. If less than 30 samples were available due to thematic sub-grouping, a cross-validation was performed only. The maximum number of PLSR factors was set to 10. The optimal number of latent variables was determined by leave-one-out cross-validation and the model providing the lowest root mean square error (*RMSE*) was chosen. In order to find out if the validation strategy affects model performance, a two-tailed t-test of the paired sample residuals was carried out with a significance level of 0.05.

Statistical models were set up for several calibration/validation combinations to analyse the effects of the pre-processing transformation procedures (i), the spectral configuration (ii), and the composition of the data set (iii) on model performance. For items (i) and (ii), a “global” calibration/validation set containing all samples was created. Analyses for both items were performed on continuous and on resampled data whereby the continuous data served as best case scenario. For item (i), the following pre-processing techniques were applied to the global data set one by one: absorbance (log(1/*R*); ABS), unit vector normalisation (UVN), standard normal variate transformation (SNV), mean normalisation (MN) (43), and first derivative (Deriv1) according to Savitzky & Golay (40). For item (ii), the data was resampled to match the configuration of five sensor systems including EnMAP, Sentinel-2, RapidEye and Landsat 8 OLI (later referred to as Landsat 8), and additionally a ground based hyperspectral system akin the PentaSpek system (44).These sensors were selected for three main reasons: first, to assess the effect of narrowband hyperspectral, superspectral and broad-band multispectral information on the prediction accuracy; second, to examine potential implications of the spectral range, and finally, to verify the potential of the approach for existing but also for missions that have been launched recently or will be launched in the foreseeable future.

Table 2 gives an overview of the characteristics of the sensors. Details on the spectral bands for EnMAP and Sentinel-2 are based on simulated data sets which have been used for spectral resampling in this study. Band specifications may therefore slightly differ from the literature. Spectral resampling was accomplished using a Gaussian model with a FWHM equal to the band spacings (45). To verify whether the composition of the data base has an impact on the model accuracy (iii), a global calibration/validation set using the entire spectral database with continuous reflectance measurements was set up again. Subsequently, the entire data set was grouped into sub data sets with respect to the region of acquisition, the year of acquisition and the crop type. These sub data sets served as the basis for “local” calibration/validation sets. Descriptive statistics of the global and local data sets are summarized in Table 3.

The performance of the PLSR models was evaluated by the *RMSE* and the coefficient of determination (R^2). Further, prediction accuracy between models with a different data base was compared using the ratio of performance to deviation (*RPD*) after Chang et al. (46). They suggested three model categories: models with *RPD* > 2 (category A) which accurately predict parameters, models with *RPD* < 1.4 (category C) which have no prediction ability, and models with *RPD* between 1.4 and 2.0 which represent an intermediate class (category B). In addition, the ratio of performance of inter-quartile distance (*RPIQ*) after Bellon-Maurel et al. (47) was used. *RPIQ* is supposed to better account for the spread of populations than the widely used *RPD*, but there is no classification scheme for assessing the model quality using the *RPIQ* yet.

RESULTS AND DISCUSSION

Descriptive statistics

LAI varied from 0.06 to 6.22 $m^2\ m^{-2}$ for the entire data set (global) comprising 364 samples. The mean value was 2.68 $m^2\ m^{-2}$. Data range and mean of the local data sets differed in some cases markedly depending on the region of acquisition, the year of acquisition and the crop type (Table 3). The total number of rape and root crops (sugar beet, potato) samples was relatively low. The data showed a slight positive skew except for rye. The amount of above-ground dry matter (*DM*) was determined for 76 % of the sampled plots. It ranged between 0.23 and 20.68 $t\ ha^{-1}$ with a mean value of 6.60 $t\ ha^{-1}$. *DM* showed a bimodal distribution, because data could not yet be acquired sufficiently during all growing stages. In particular, little data is available in the period of booting to the beginning of florescence emergence (cereals and rape) and during tuber and beet root formation (potato, sugar beet). Similar to the *LAI*, local data sets partly exhibited distinct differences in data range and mean values. Some of the local data sets showed a positive skew (R2, Year 2014, rape). By contrast, the local data set based on samples taken in 2012 had a negative skew. It should be noted at this point that PLSR has the advantage of not being subject to assumptions about the data distribution, samples size or scale (48).

Table 2: Sensor characteristics of the five systems tested using spectral resampling.

	EnMAP (49)	PentaSpek (34)	Sentinel-2 (50)	RapidEye (51)	Landsat 8 (52)
Spectral range (nm)	420-2450	200-1100 (effectively 400-925)	443-2190	440-850	430-2290
No. bands	242	551	13	5	7
Spectral bands[1,2] (nm)	420-1000 (6.5) 900-2450 (10)	400-925 (1)	442 (20) 492 (65) 559 (35) 662.5 (30) 702 (15) 744.5 (15) 782 (20) 842 (115) 867 (20) 942 (20) 1369 (30) 1607 (90) 2189.5 (180)	440-510 520-590 630-685 690-730 760-850	430-450 450-510 530-590 640-670 850-880 1570-1650 2110-2290
GSD[3] (nadir) (m)	30	0.15×0.82**	10 to 60	6.5	30

[1] for EnMAP, numbers in brackets indicate spectral sampling distance, [2] for PentaSpek and Sentinel-2, numbers in brackets indicate band width. [3] Ground Sampling Distance
** @ 0.5 km/h and a 45° viewing angle and one metre distance between sensor and object

Effect of pre-processing transformations

The results on the effect of different pre-processing transformations as applied to the field spectrometer data with 1 nm spectral resolution are shown in Table 4 and Table 5. Compared to no transformation, the *LAI* models based on independent validation showed a marginal increase of the coefficient of determination (R^2_{ival}) when SNV and first derivative were applied. $RMSE_{ival}$ decreased only in the case of first derivative transformation. All other pre-processing techniques had either a negligible effect on model performance or even decreased it. By contrast, cross-validated *LAI* models improved slightly when ABS, UVN and MN were applied. Except for two cases, cross-validation resulted in lower R^2 and higher *RMSE* than independent validation (Table 4). Results indicate that none of the tested pre-processing methods generally performs best. *DM* models were positively influenced by SNV no matter which validation procedures was adopted. $RMSE_{ival}$ was reduced by almost 10% compared to the raw reflectance spectra. Irrespective of the validation procedure, model performance also benefits if UVN or first derivative were applied.

By contrast, a conversion of reflectance to absorbance clearly decreased the predictive power of the models (Table 5). As reported for *LAI*, R^2 of the cross-validated *DM* models was lower and *RMSE* was higher in comparison to the independently validated models. The positive influence of the first derivative transformation as observed for the *LAI* and the *DM* models was also found by Pimstein et al. (29) when deriving dry biomass, water content, and nitrogen content. By evaluating the prediction accuracy for feed quality constituents in pasture, Thulin et al. (32) also stated that their best models were dominated by the first derivative. The *t*-test *p* values of the *LAI* models were all above 0.05, indicating that there is no significant difference between the models based on independent or cross-validation. However, for *DM*, *t*-test *p* values point out significant differences of coupled model residuals in four cases.

Table 3: Descriptive statistics of the model reference data base for the entire data set (global) and sub data sets (local) according to region of acquisition (R1: Köthen, R2: Braunschweig, R3: Baruth), year of acquisition (e.g., Y11: 2011) and crop type. The sub data set grain comprises wheat, rye and barley. The sub data set "root" combines sugar beet and potato. LAI in m^2 m^{-2}, DM in t ha^{-1}.

LAI													
	All	Region			Year				Crop				
	global	R1	R2	R3	Y11	Y12	Y13	Y14	wheat	rye	grain	rape	root
No.	364	132	194	38	57	80	170	57	159	154	316	26	22
Mean	2.68	2.88	2.35	3.67	1.65	3.75	2.71	2.09	2.65	2.92	2.78	1.88	2.08
Min	0.06	0.5	0.06	2.63	0.50	1.11	0.06	0.30	0.13	0.06	0.06	0.30	0.46
Max	6.22	5.95	6.22	5.22	3.32	5.95	6.22	5.01	5.95	6.22	6.22	4.52	5.01
SD	1.46	1.33	1.54	0.66	0.63	0.90	1.59	1.23	1.40	1.52	1.46	1.06	1.42
Dry matter													
	All	Region			Year				Crop				
	global	R1	R2	R3	Y11	Y12	Y13	Y14	wheat	rye	grain	rape	root
No.	278	132	146	-	57	80	84	57	159	68	230	26	22
Mean	6.60	7.59	5.71	-	3.53	10.38	5.97	5.28	7.56	5.80	7.11	5.65	2.44
Min	0.23	1.79	0.22	-	1.79	2.87	0.22	0.39	0.36	0.22	0.22	1.08	0.39
Max	20.68	13.77	20.68	-	6.17	13.77	18.02	20.68	18.02	15.31	18.02	20.68	5.52
SD	4.58	3.91	4.95	-	0.94	2.27	4.86	5.31	4.41	3.85	4.35	6.24	1.37

Working with spectra adjusted to the spectral configuration of hyper-, super- and multispectral sensors, as in the following section, revealed that there is obviously no "all-purpose" transformation. The performance and efficiency of the pre-processing transformation technique varied with the spectral configuration of the sensor system (results not shown). Several transformation techniques achieved similar results. This was the case for the *LAI* models based on PentaSpek-like reflectance values when SNV, MN and ABS were applied. Further, prediction accuracy was similar for the *DM* models

developed with Sentinel-2-like reflectance spectra and SNV, Deriv1 or MN. Based on EnMAP-like spectra, best results were achieved using UNV or SNV for *LAI* and *DM*, respectively. For Landsat 8-like spectra, UVN was the best choice for *LAI*, but for *DM* no pre-processing yielded the lowest error values. If spectra were adjusted to RapidEye spectral configuration, transformation to absorbance or UVN yielded best results for *LAI* and *DM*, respectively. Since first derivative transformation is inappropriate for multispectral sensors, it was not used for the latter two. As result of the investigations, the analysis in the next section was done using the best pre-processing transformation (cp. Table 6). Beyond that, findings gained for spectra with 1 nm resolution based on global data sets could not be transferred to the local data sets (results not shown). For the purpose of comparison, these models have been set up with a single pre-processing method that is UVN for *LAI*, and SNV for *DM* (cp. Table 7, Table 8). These techniques were selected as they performed best in many cases.

Table 4: Prediction accuracy of the LAI models with respect to different pre-processing techniques. none: no transformation, ABS: absorbance, UVN: unit vector normalisation, SNV: standard normal variate, MN: mean normalisation, Deriv1: first derivation.

Technique	No. LV_{ival}	R^2_{ival}	$RMSE_{ival}$ (m^2 m^{-2})	No. LV_{cv}	R^2_{cv}	$RMSE_{cv}$ (m^2 m^{-2})	*t*-test *p* value
none	7	0.76	0.71	5	0.72	0.77	0.78
ABS	6	0.76	0.71	6	0.75	0.73	0.24
UVN	3	0.75	0.73	4	0.75	0.73	0.98
SNV	4	0.79	0.72	3	0.73	0.75	0.39
MN	2	0.74	0.74	4	0.75	0.73	0.49
Deriv1	6	0.77	0.70	3	0.71	0.79	0.38

Table 5: Prediction accuracy of the DM models with respect to different pre-processing techniques. none: no transformation, ABS: absorbance, UVN: unit vector normalisation, SNV: standard normal variate, MN: mean normalisation, Deriv1: first derivation.

Technique	No. LV_{ival}	R^2_{ival}	$RMSE_{ival}$ (t ha^{-1})	No. LV_{cv}	R^2_{cv}	$RMSE_{cv}$ (m^2 m^{-2})	*t*-test *p* value
none	6	0.85	1.81	6	0.83	1.88	0.02
ABS	7	0.76	2.56	7	0.77	2.21	0.36
UVN	6	0.86	1.73	5	0.84	1.86	0.01
SNV	7	0.87	1.64	6	0.85	1.80	0.14
MN	5	0.85	1.80	5	0.83	1.88	0.00
Deriv1	6	0.85	1.76	6	0.84	1.86	0.04

Effect of spectral configuration

Results on the effect of spectral configuration on model performance based on independent validation are displayed in Figure 2 and Figure 3. Results of the cross-validated *LAI* and *DM* models are depicted in Figure 4 and Figure 5. Results indicate that resampling reflectance spectra to the spectral configuration of the EnMAP hyperspectral instrument or the Sentinel-2 superspectral instrument did not negatively affect the prediction accuracy of the *LAI* models. In fact, data reduction partly increased model prediction performance. Accordingly, scatter plots of measured versus predicted values of the EnMAP, PentaSpek and Sentinel-2 model show a little less variation along the 1:1 line than the SVC model disregarding the few outliers in all models (Figure 2, Figure 4). This result is in agreement with Herrmann et al. (55) who reported that continuous data did not provide any significant advantage over VENμS and Sentinel-2 data in *LAI* assessment. Results for the PentaSpek system are inconsistent for the two applied validation strategies. While cross-validation leads to an increase in prediction performance just as observed for EnMAP and Sentinel-2 like spec-

tra, independent validation causes a decrease. If spectra were adjusted to multispectral instruments, i.e. RapidEye and Landsat 8, prediction accuracy decreased as expected due to the strongly reduced spectral information. R^2ival dropped to 0.66 and 0.65, *RMSE*ival increased to 0.85 and 0.86, respectively (Figure 2). Cross-validated models reflected the same trend, but prediction accuracy was even lower than with independent validation (Figure 4). For Landsat 8, scatter plots show that samples with *LAI* less than 3 are increasingly overestimated. At the same time, the underestimation of *LAI* values greater than 5, which is present in all scatter plots, is augmented. For RapidEye, variation along the 1:1 line is generally amplified. All *t*-test *p* values were higher 0.05 indicating that coupled model residuals are the same except for the *LAI* model based on RapidEye (Table 6).

Table 6: Pre-processing techniques and t-test p values of the LAI and DM models for different spectral configuration.

Parameter	*LAI*		*DM*	
System	PP	*t*-test *p* value	PP	*t*-test *p* value
SVC	Deriv1	0.85	SNV	0.14
EnMAP	UVN	0.73	SNV	0.89
PentaSpek	SNV	0.61	SNV	0.05
RapidEye	ABS	0.00	UVN	0.00
Landsat 8	UVN	0.95	none	0.00
Sentinel-2	ABS	0.23	SNV	0.00

Comparable to *LAI*, the *DM* prediction models developed with spectra resampled to EnMAP band configuration showed an accuracy similar to that of models based on field reflectance. If spectra were transformed to match the PentaSpek and Sentinel-2 band configurations, model prediction accuracy slightly decreased but results were still promising. Adjusting spectra to RapidEye and Landsat 8 band configuration led to considerably lower prediction accuracies. R^2 was particularly low for RapidEye, and *RMSE* was correspondingly high regardless of the validation scheme (Figure 3, Figure 5). Generally, cross-validated models showed slightly higher prediction errors as observed before. Scatter plots of measured versus predicted values exhibited a large variation along the 1:1 line for the Landsat 8 and RapidEye models. Smaller *DM* values are generally overestimated, while *DM* is underestimated for higher values (Figure 3, Figure 5). *T*-test *p* values disclose significant differences between ival- and cv-models based on super- and multispectral data (Table 6).

The predictive ability of the *LAI* model based on data resampled to the PentaSpek system featuring a limited spectral range (cp. Table 2) was only slightly lower than the model based on field spectrometer data (SVC). For *DM*, the *RMSE* increased by about 20% but was in the range of the prediction error of the Sentinel-2 model. This may be regarded as a hint that it was not the data range (extended to the SWIR), but the spectral resolution, that was crucial for successful parameter predictions. The poor performance of the RapidEye and Landsat 8 model strengthen that assumption. However, Verrelst et al. (25) reported high prediction accuracies also for a set-up with the four Sentinel-2 MSI bands of 10 m spatial resolution and spectral bands between 490-665 nm and 842 nm which very much resemble four of five Rapid Eye bands. Thus, further research is needed to shed light on that issue. Improvements of the prediction accuracy using hyper- and superspectral sensors compared to multispectral systems are in conformity with previous studies. Twele et al. (17) reported higher prediction accuracy for EO-1 Hyperion than for Landsat Enhanced Thematic Mapper (ETM+). Pu et al. (53) achieved the best results with EO-1 Hyperion followed by EO-1 ALI and ETM+. The high potential of the Sentinel-2 mission for the retrieval of biophysical variables was also pointed out by Richter et al. (36) and Frampton et al. (54). Our results show that the spectral configuration of Sentinel-2 seems as appropriate as that of hyperspectral systems using PLSR especially for *LAI* retrieval. Taking the frequent revisit time of Sentinel-2 and its higher spatial configuration compared to EnMAP or Landsat 8 into account, this sensor will be ideal for use in vegetation and yield monitoring as stated previously (e.g., 36,54).

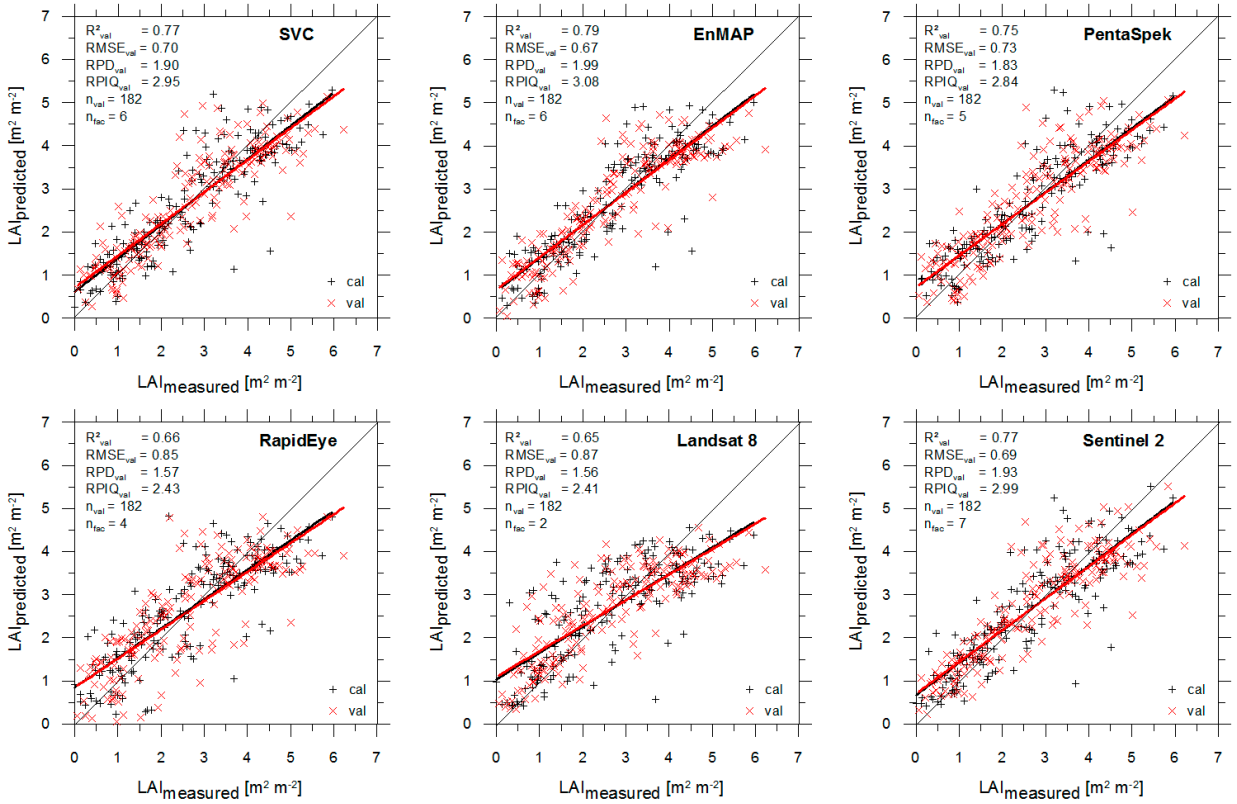

Figure 2: Plots of measured versus predicted values for LAI (m^2 m^{-2}) based on independent validated models with different spectral configurations. For spectral pre-processing see Table 6.

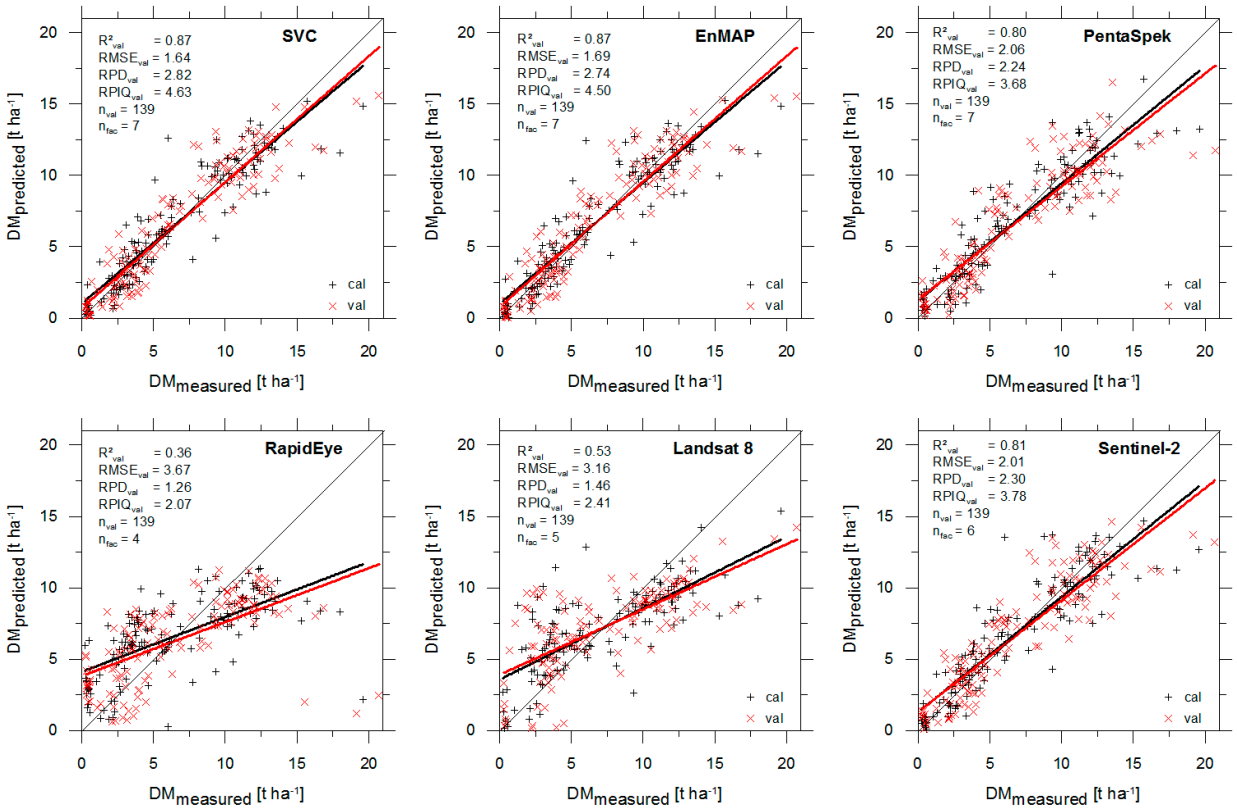

Figure 3: Plots of measured versus predicted values for DM (t ha^{-1}) based on independently validated models with different spectral configurations. For spectral pre-processing see Table 6.

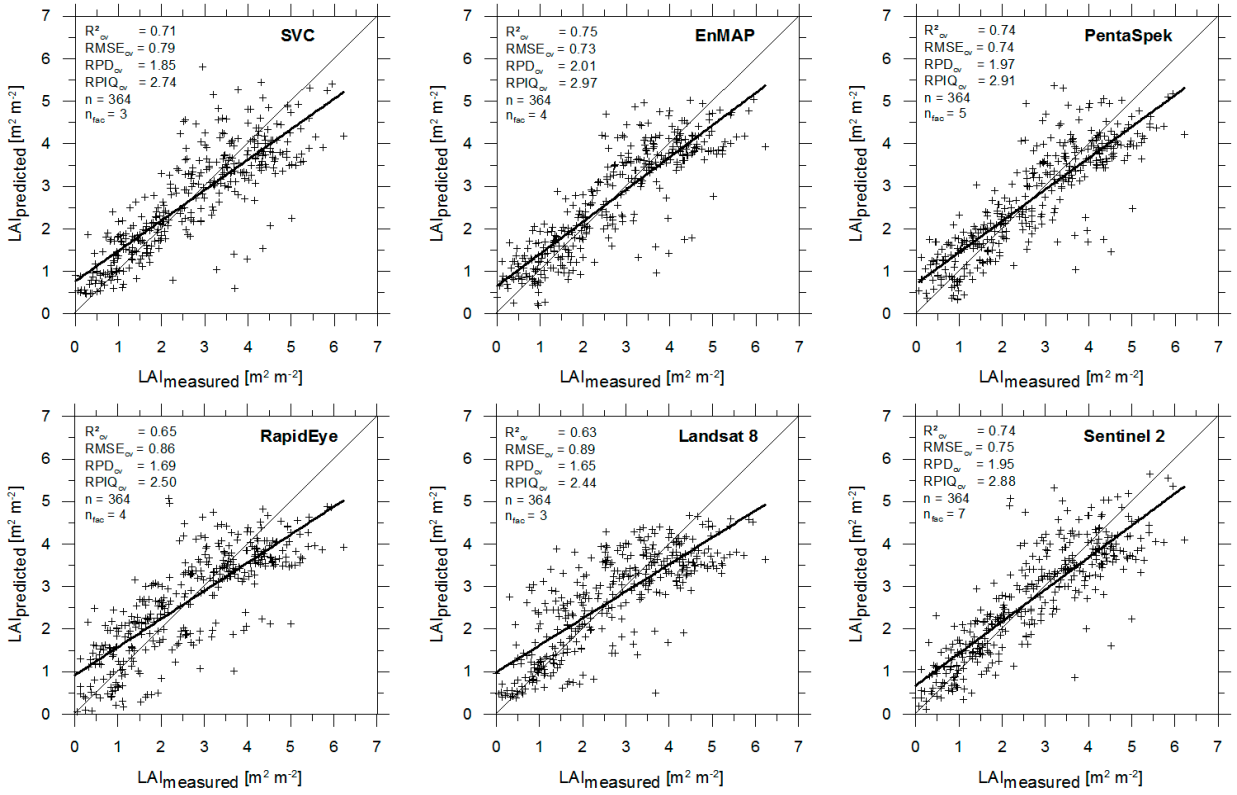

Figure 4: Plots of measured versus predicted values for LAI (m^2 m^{-2}) based on cross-validated models with different spectral configurations. For spectral pre-processing see Table 6.

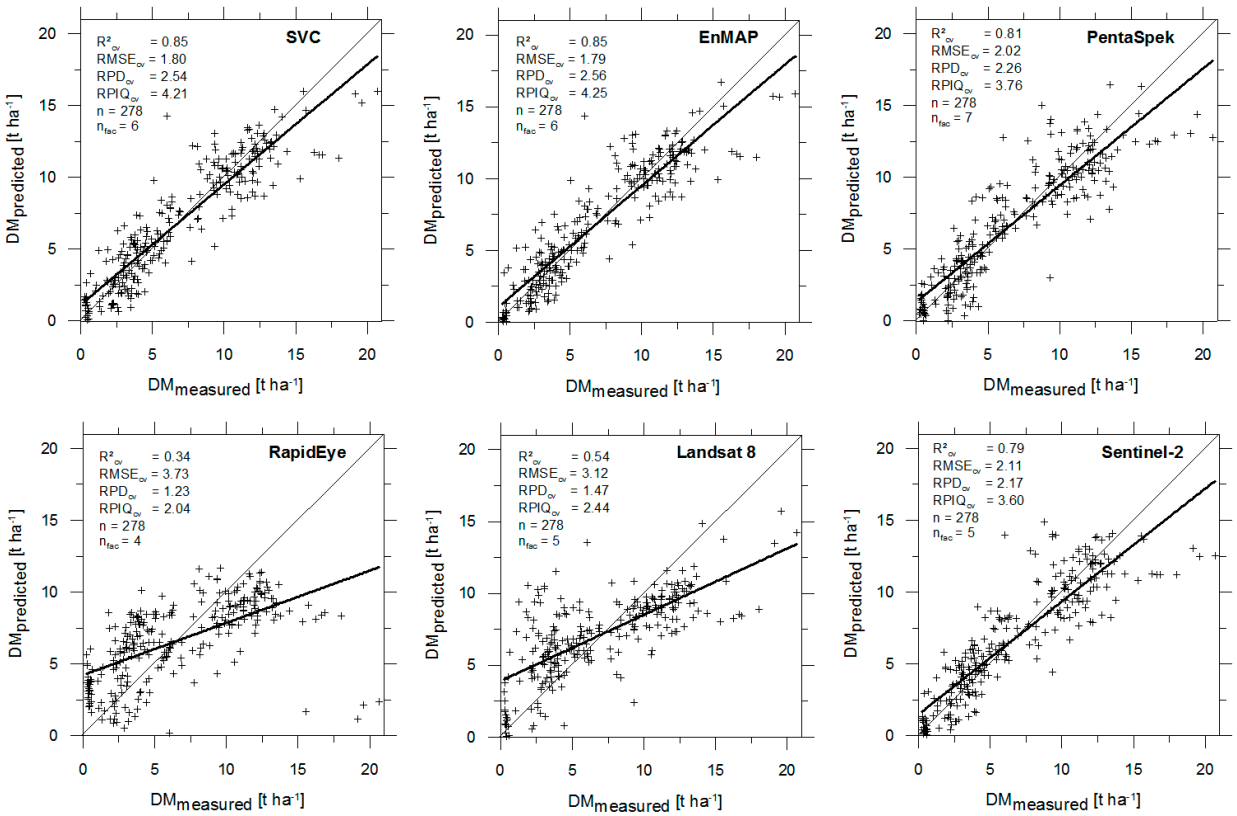

Figure 5: Plots of measured versus predicted values for DM (t ha^{-1}) based on cross-validated models with different spectral configurations. For spectral pre-processing see Table 6.

Effect of data base composition

Subsetting the data into local data sets did not necessarily result in an improvement of model accuracy. Further, the behaviour of the local models was inconsistent for *LAI* (Table 7) and *DM* (Table 8). For *LAI*, local models according to the region of acquisition improved compared to the global model for two regions (R1, R2). The plots of measured versus predicted *LAI* values visualise these results by noticeably less scattering, especially in the R1 model (Figure 6a,b). Prediction accuracy decreased for the Baruth test site (R3). Grouping the data set according to the year of acquisition was of little benefit except for the year 2011, and most notably for the year 2013. $RPD_{ival/cv}$ values above 2 indicated category A models with a high predictive accuracy. $RPIQ_{ival/cv}$ of the 2013 model was almost twice as high as for the global model. The $RPIQ_{ival/cv}$ for the 2011 model dropped, however, compared to the global model as a consequence of the much smaller data range and a large number of samples with *LAI* values below 2 (Table 3). This becomes clearly obvious in the scatter plot (Figure 6c,d). If the samples were divided with respect to the crop type, prediction ability increased considerably for wheat and rye (Table 7, Figure 6e,f). A positive effect was achieved even with a combined model for all grain crops. These models belong to category A, which was confirmed by elevated $RPIQ_{ival/cv}$ values. The local models for rape and root crops have only poor predictive ability with an RPD_{cv} less than 1.35 (category C) (Table 7). Despite the improved prediction accuracy of selected local models, the trend to underestimate *LAI* values greater than 5 persisted. Further, the cross-validated model for rye cannot appropriately account for *LAI* values near 3 (Figure 6). This was not observed when independent validation was performed with samples withheld from calibration. Correspondingly, the *t*-test *p* value of the models for rye is 0.01, pointing out that there is a difference between them. *T*-test *p* values of the coupled residuals of the R2, R3 and the 2013 model were also lower than 0.05. In that case, differences between cross-validation and independent validation were not apparent in the scatter plots (not shown).

Table 7 Prediction accuracy of the LAI models with respect to data base composition (thematic groups). The pre-processing technique was UVN. R1: Köthen, R2: Braunschweig, R3: Baruth. ival: independent validation, cv: cross-validation

	No. LV_{ival}	R^2_{ival}	$RMSE_{ival}$ (t ha^{-1})	RPD_{ival}	$RPIQ_{ival}$	No. LV_{cv}	R^2_{cv}	$RMSE_{cv}$ (m^2 m^{-2})	RPD_{cv}	$RPIQ_{cv}$	*t*-test *p* value
	Global										
All	3	0.75	0.73	1.82	2.82	4	0.75	0.73	2.00	2.96	0.95
	Local: Region										
R1	6	0.79	0.61	2.20	3.72	3	0.82	0.58	2.32	3.96	0.94
R2	7	0.77	0.73	2.12	3.58	6	0.76	0.76	2.03	3.46	0.00
R3	3	0.40	0.65	1.06	1.75	4	0.60	0.42	1.56	2.72	0.00
	Local: Year										
2011	5	0.75	0.30	2.02	2.15	3	0.68	0.30	2.08	2.22	0.48
2012	3	0.36	0.69	1.27	1.51	2	0.41	0.70	1.30	1.55	0.17
2013	3	0.87	0.57	2.80	5.28	3	0.87	0.57	2.78	5.33	0.00
2014	4	0.52	0.81	1.47	1.97	3	0.37	0.98	1.25	1.70	0.46
	Local: Crop type										
wheat	4	0.85	0.54	2.58	3.95	4	0.85	0.55	2.56	3.91	0.37
rye	6	0.84	0.60	2.55	4.57	2	0.81	0.67	2.26	4.12	0.01
grain	3	0.76	0.71	2.06	3.54	3	0.80	0.66	2.21	3.81	0.91
rape	-	-	-	-	-	1	0.20	0.96	1.10	1.69	-
root	-	-	-	-	-	7	0.47	1.05	1.35	2.08	-

For *DM*, no clear trend was observable for the local models with respect to the region of acquisition. On the one hand, $R^2_{ival/cv}$ and $RPD_{ival/cv}$ decreased except for the R2 model based on inde-

pendent validation. On the other hand, $RPIQ_{ival/cv}$ increased for the R1 models but decreased for the R2 models due to the differences in the underlying data distribution. Splitting up the samples according to the year of acquisition leads to a substantial increase in prediction accuracy, if only data from 2013 were included. This is in agreement with the findings reported for *LAI*. However, this time, poor results were obtained for the data from 2011. Instead, $RPD_{ival/cv}$ of the 2014 model was about as high as that of the global model. The limited $RPIQ_{ival/cv}$ values of the model can be explained by the high frequency of samples with little dry matter. Yet, the scatter along the 1:1 line is relatively homogenous with a minor trend to underestimate higher values (Figure 7a,b). Crop type-specific *DM* models (except for root crops) belong to category A (RPD > 2) just as the global model. The partitioning of the samples by crop type, however, provided contradicting results for the two applied validation strategies. In contrast to the independently validated models, performance increased if cross-validation was applied. These differences correspond with *t*-test *p* values of 0.0007 and 0.02 for wheat and grain model pairs. Setting up a separate model for rape improved the prediction performance, but only a few samples of high biomass fit the regression line at this stage and $RPIQ_{cv}$ is low (Figure 7c). More samples need to be acquired for evidence. The root model showed poor results just as observed for *LAI*.

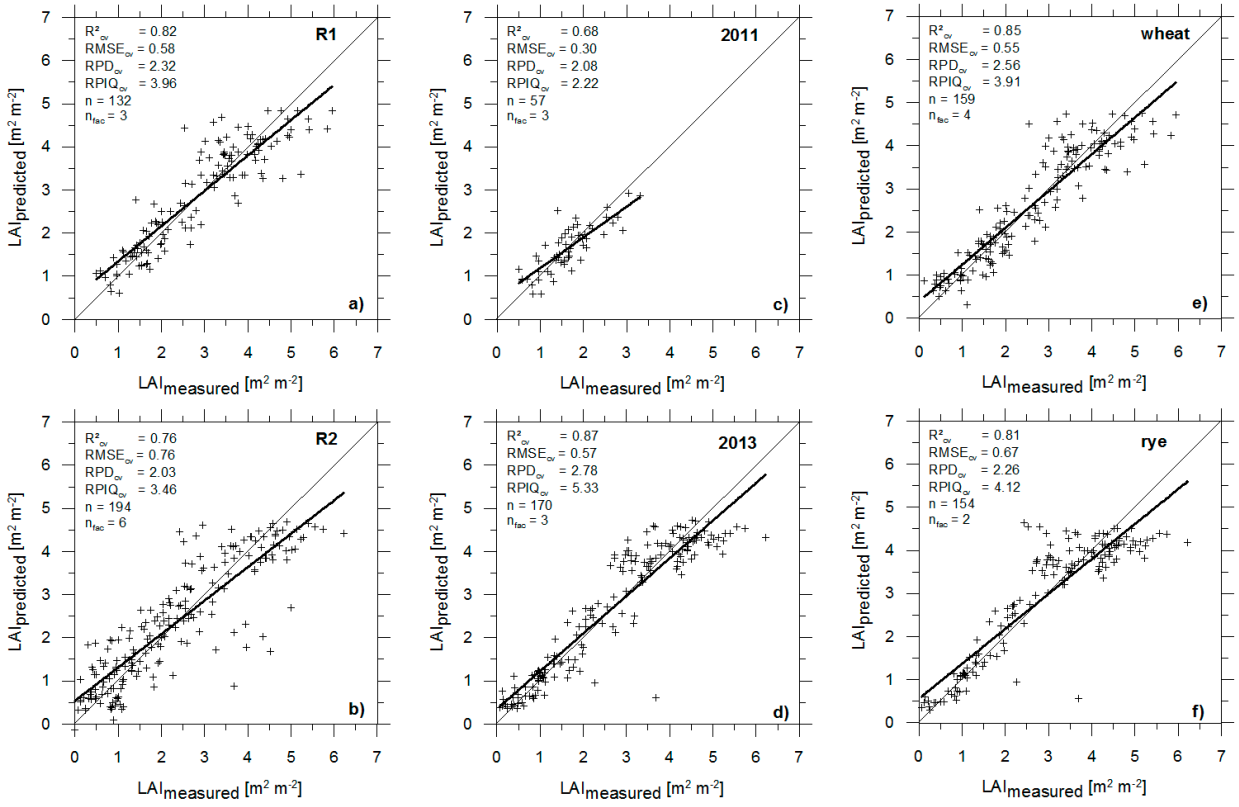

Figure 6: Selected plots of measured versus predicted values for local LAI models based on cross-validation.

The precision obtained for *LAI* and *DM* with global and local PLSR models in our study is within the precision obtained by previous authors (e.g., 29,55). Results confirmed that obviously neither the region of acquisition nor the year of acquisition or the crop type alone determine model performance, as the separation of the global data set into local data sets had both positive and negative effects in each of the three thematic groups for *LAI* and for *DM*. Rather, it seems, that the data range and the variation related to the dates and the frequency of sampling, have an additional impact on model performance. In most cases, prediction accuracy of the local models improved, if the local data sets possessed a similar data range and variation as the global model. For *LAI*, this applies to the R1, R2, 2013, wheat, rye and grain model. For *DM*, this is applicable to the 2013, 2014, wheat, and grain models as well as to the model for winter rape (cf. Table 3, Table 7, Table 8).

However, the prediction accuracy of the *LAI* model based on data from 2011 increased despite a much smaller data range and little variation. In fact, the 2011 data set originates from a one-day field campaign in one crop type just like the data set of the R3 model. However, the latter showed a considerable decline in model performance due to a general overestimation of *LAI* values. Further, the *DM* model for region 2 (R2) was not better than the global model, although data range and variation were equivalent to the global data set. An explanation for that behaviour has not been found so far. The generally poor results for root crops may be attributed to the following aspects. First, the cultivation of sugar beet and potatoes in rows make correct *LAI* measurements of 0.25 m² plots more difficult than in relatively homogeneous canopies of grain crops or rape. Second, diurnal variations of the canopy structure, especially on very hot days, may influence measurements. Beyond that, data on sugar beet and potato is hitherto only available from two dates during the growing season (cp. Table 1). More samples covering the entire growing season are needed to gain a further insight into the performance of the models.

Table 8: Prediction accuracy of the DM models with respect to data base composition (thematic groups). The pre-processing technique was SNV. R1: Köthen, R2: Braunschweig, R3: Baruth. cv: cross-validation

	No. LV_{ival}	R^2_{ival}	$RMSE_{ival}$ (t ha^{-1})	RPD_{ival}	$RPIQ_{ival}$	No. LV_{cv}	R^2_{cv}	$RMSE_{cv}$ (t ha^{-1})	RPD_{cv}	$RPIQ_{cv}$	*t*-test *p* value
	Global										
All	7	0.87	1.64	2.82	4.63	6	0.85	1.80	2.54	4.21	0.13
	Local: Region										
R1	2	0.85	1.50	2.62	5.00	4	0.84	1.55	2.52	4.87	0.25
R2	7	0.88	1.74	2.89	3.95	5	0.83	2.02	2.45	3.35	0.91
R3	-	-	-	-	-	-	-	-	-	-	-
	Local: Year										
2011	1	0.52	0.61	1.46	1.82	2	0.44	0.71	1.33	1.66	0.35
2012	2	0.41	1.64	1.31	1.72	3	0.36	1.83	1.24	1.54	0.85
2013	4	0.92	1.37	3.61	6.55	4	0.90	1.57	3.08	5.66	0.00
2014	6	0.88	1.74	2.94	2.29	7	0.84	2.12	2.50	1.88	0.02
	Local: Crop type										
wheat	7	0.86	1.62	2.70	4.82	6	0.85	1.71	2.58	4.59	0.00
rye	3	0.80	1.75	2.26	3.15	5	0.86	1.43	2.69	3.81	0.25
grain	6	0.86	1.64	2.67	4.50	6	0.86	1.65	2.63	4.47	0.02
rape	-	-	-	-	-	4	0.88	2.16	2.89	0.99	-
root	-	-	-	-	-	2	0.41	1.08	1.27	1.66	-

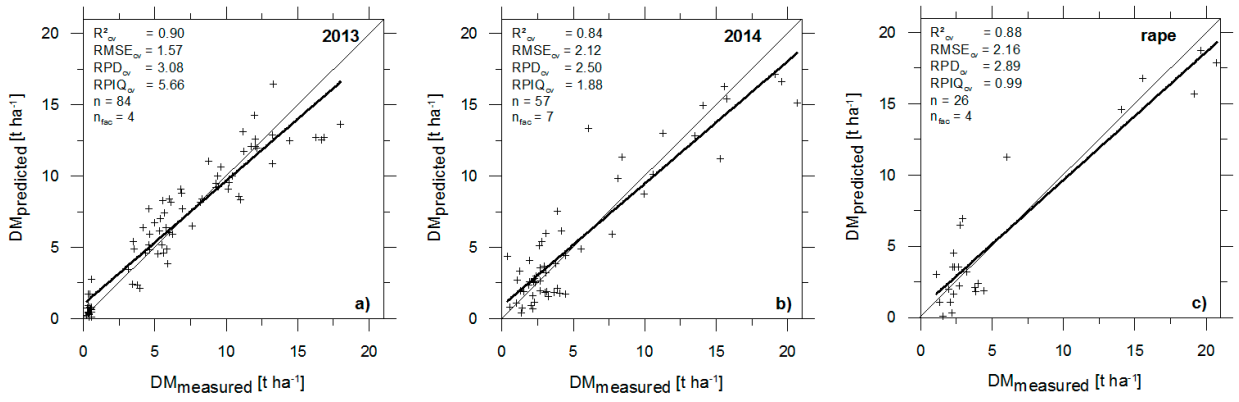

Figure 7: Selected plots of measured versus predicted values for local DM models based on cross-validation.

Plotting the X-scores of the PLSR model can be used to identify object similarities and dissimilarities (42). The X1-X2 and X2-X4 score plots of the global *LAI* model (Figure 8) displayed a distinct separation of the rape and root crop samples, in particular potato. A separation of samples is also evident in the lower half of the scatter plots of the global *LAI* models (Figure 2, Figure 4) and corresponds to the potato samples. This may be seen as an indication that the crop type affects the model performance, and consequently, individual crop models should be set up to estimate the *LAI*. A similar conclusion based on a study of wheat and potato was drawn by Herrmann et al. (55) who analysed the suitability of PLSR and vegetation indices for *LAI* estimation. By contrast, Kross et al. (20) reported the opposite. They showed that indices were insensitive to crop types with different canopy and leaf structure (corn and soybean) which would be a large benefit with regard to operability. For the *DM* model, X-score plots did not show distinct crop groups. This is also reflected by lower deviations between R^2 and RPD of the global and the local crop models as compared to *LAI*.

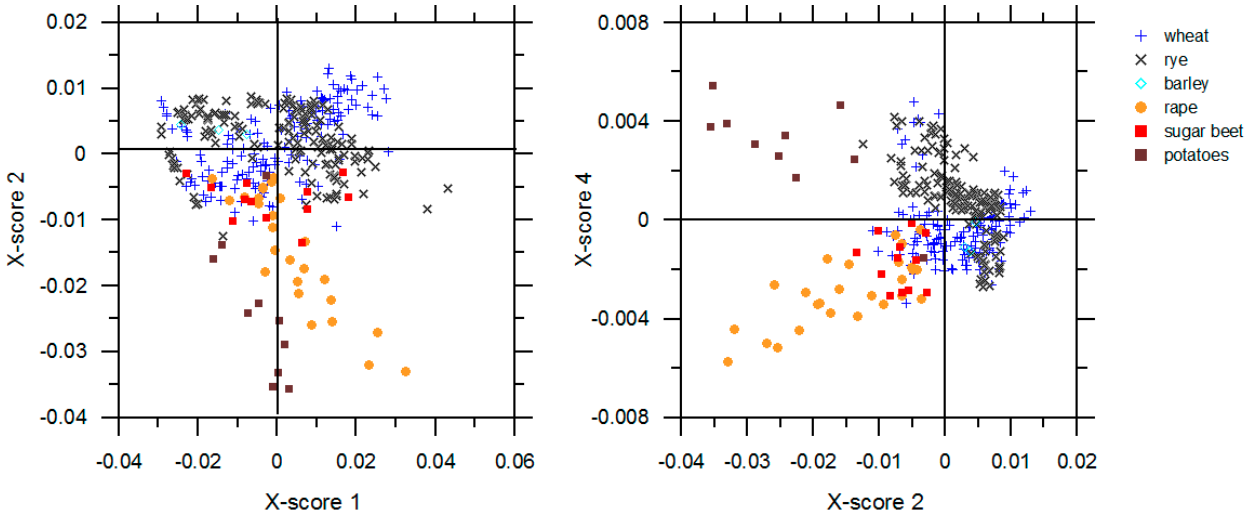

Figure 8: Plot of XX-scores of the cross-validated global LAI model. Samples are marked according to the crop type.

CONCLUSIONS

In this study, we presented an approach to estimating *LAI* and *DM* based on field reflectance measurements from different geographic regions, acquisition times and crop types using PLSR. There was no all-purpose preprocessing method but results pointed out that the technique should be chosen with respect to the sensor and the parameter of interest. In general, pre-processing had less effect on the prediction accuracy of the *LAI* models than on the prediction accuracy of the *DM* models. Hyperspectral sensors (EnMAP, PentaSpek) and the superspectral sensor (Sentinel-2) performed as well as field spectrometer data for *LAI* assessment. For *DM*, prediction accuracy decreased for all tested sensor configurations, but not for EnMAP. The prediction accuracy of the *DM* models for multispectral sensors (Landsat 8, RapidEye) was generally low. The approach seems therefore best suited for the hyperspectral EnMAP mission, but results for the superspectral instrument on board the Sentinel-2 satellites were promising, too. However, application of the models to image data will require additional properties not considered in this study to be examined including the sensitivity of the predictions with respect to the spatial resolution and the noise of the sensor, the viewing angle during image acquisition, and the quality of atmospheric correction with cloud detection. Grouping the data set according to the year of acquisition, the region of acquisition and the crop type did partly improve the model accuracy, but the mechanisms behind have not yet been fully understood. Additional data for crops like rape, sugar beet, and potato at various crop development stages is needed to gain a further insight. At that point, results suggest that crop-specific models are best for predicting *LAI*.

Generally, model prediction accuracy varied depending on the deployed validation strategy. Comparing the performance of model pairs for a tested configuration, i.e. the $wheat_{cv}$ vs. $wheat_{ival}$ model, prediction accuracy was mostly somewhat lower using cross-validation. Moreover, model performance did not always show the same trend for both validation techniques, i.e., when comparing $wheat_{cv}$ vs. $global_{cv}$ or $wheat_{ival}$ vs. $global_{ival}$. These results clearly demonstrated the need for additional sampling to extend the validation data base and to verify the accuracy and the robustness of the regression models. In future work, selected models shall be applied to image data with spectral characteristics tested herein. A validation of these image-based results with additional *in situ* reference measurements of the target variables will provide further information on the validity of model errors. In view of the low prediction ability of the PLSR models for multispectral sensors, such as Landsat 8 and RapidEye, non-parametric methods able to perform nonlinear data fitting (MLRAs) may provide a more powerful alternative.

ACKNOWLEDGEMENTS

The research in this paper was funded by the German Federal Ministry of Economic Affairs and Energy (Support code 50EE1315) within the framework of the RiflE project on *Regional detection of the current agricultural yield potential by remote sensing*. A major part of the field work was carried out within the HyLand project. This project was funded with financial resources of the Federal Ministry of Economics and Technology on the basis of a decision of the German Parliament, grant number 50 EE 1014. The support is gratefully acknowledged. The authors also wish to thank Karl Segl from the Remote Sensing section at GFZ German Research Centre for Geosciences for providing the simulated EnMAP and Sentinel-2 data used for spectral resampling.

REFERENCES

1 Soares J, C Justice, P Kosuth, O Leo, J S Parihar, D Williams, B Wu & I Jarvis, 2011. The Launch of the G20 Global Agricultural Geo-Monitoring Initiative (GEO-GLAM). Document 14, GEO-VIII Meeting 16-17 November 2011, 6 pp.

2 Atzberger C, 2013. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. Remote Sensing, 5(2): 949-981

3 Blackmore S, 2000. The interpretation of trends from multiple yield maps. Computers and Electronics in Agriculture, 26: 37-51

4 Schepers A R, J F Shanahan, M A Liebig, J S Schepers, S H Johnson & A Luchiari, 2004. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. Agrononomy Journal, 96: 195-203

5 Maine N, J Lowenberg-DeBoer, W T Nell & Z G Alemu, 2010. Impact of variable-rate application of nitrogen on yield and profit: a case study from South Africa. Precision Agriculture, 11(5): 448-463

6 Scharf P C, D K Shannon, H L Palm, K A Sudduth, S T Drummond, N R Kitchen, L J Mueller, V C Hubbard & L F Oliveira, 2011. Sensor-based nitrogen applications out-performed producer-chosen rates for corn in on-farm demonstrations. Agronomy Journal, 103: 1683-1691

7 Blackmore S, R J Godwin & S Fountas, 2003. The analysis of spatial and temporal trends in yield map data over six years. Biosystems Engineering, 84(4): 455-466

8 Milne AE, R Webster, D Ginsburg & D Kindred, 2012. Spatial multivariate classification of an arable field into compact management zones based on past crop yields. Computers and Electronics in Agriculture, 80:17-30

9 Mulla D J, 2013. Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. Biosystems Engineering, 114: 358-371

10 Ren J, Z Chen, Q Zhou & H Tang, 2008. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China, International Journal of Applied Earth Observation and Geoinformation, 10: 403-413

11 Uno Y, S Prasher, R Lacroix, P Goel, Y Karimi, A Viau & R Patel, 2005. Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. Computer and Electronics in Agriculture, 47: 149-161

12 Zarco-Tejada P, S Ustin & M Whiting M, 2005. Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. Agronomy Journal, 97(3): 641-653

13 Bach H & W Mauser, 2003. Methods and Examples for Remote Sensing Data Assimilation in Land Surface Process Modeling. IEEE Transactions on Geoscience and Remote Sensing, 41(7): 1629-1637

14 Fang H, S Liang, G Hoogenboom, J Teasdale & M Cavigelli, 2008. Corn‐yield estimation through assimilation of remotely sensed data into the CSM‐CERES‐Maize model. International Journal of Remote Sensing, 29(10): 3011-3032

15 Machwitz M, L Giustarini, C Bossung, D Frantz, M Schlerf, H Lilienthal, L Wandera, P Matgen, L Hoffmann & T Udelhoven, 2014. Enhanced biomass prediction by assimilating satellite data into a crop growth model. Environmental Modelling & Software, 62: 437-453

16 Zheng G & L M Moskal, 2009. Retrieving Leaf Area Index (LAI) using remote sensing: Theories, methods and sensors. Sensors, 9: 2719-2745

17 Twele A, S Erasmi & M Kappas, 2008. Spatially explicit estimation of Leaf Area Index using EO-1 Hyperion and Landsat ETM+ Data: Implications of spectral bandwidth and shortwave infrared data on prediction accuracy in a tropical montane environment. GIScience & Remote Sensing, 45(2): 229-248

18 Marshall M & P Thenkabail, 2015. Advantage of hyperspectral EO-1 Hyperion over multispectral IKONOS, GeoEye-1, WorldView-2, Landsat ETM+, and MODIS vegetation indices in crop biomass estimation. ISPRS Journal of Photogrammetry and Remote Sensing, 108: 205-218

19 Dente L, G Satalino, F Mattia & M Rinaldi, 2008. Assimilation of leaf area index derived from ASAR and MERIS data into CERES-Wheat model to map wheat yield. Remote Sensing of Environment, 112: 1395-1407

20 Kross A, H McNairn, D Lapen, M Sunohara & C Champagne, 2015. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. International Journal of Applied Earth Observation and Geoinformation, 34: 235-248

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

22 Baret F & G Guyot, 1991. Potentials and limits of vegetation indices for *LAI* and APAR assessment. Remote Sensing of Environment, 35: 161-173

23 Haboudane D, J R Miller, E Pattey, P J Zarco-Tejada & I B Strachan, 2004. Hyperspectral vegetation indices and novel algorithms for predicting green *LAI* of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sensing of Environment, 90: 337-352

24 Rivera J P, J Verrelst, J Delegido, F Veroustraete & J Moreno, 2014. On the semi-automatic retrieval of biophysical parameters based on spectral index optimization. Remote Sensing, 6: 4927–4951

25 Verrelst J, J Muñoz, L Alonso, J Delegido, J P Rivera, G Camps-Valls & J Moreno, 2012. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. Remote Sensing of Environment, 118: 127-139

26 Verrelst J, J P Rivera, F Veroustraete, J Muñoz-Marí, J G P W Clevers, G Camps-Valls & J Moreno, 2015. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – A comparison. ISPRS Journal of Photogrammetry and Remote Sensing, 108: 260-272

27 Verrelst J, G Camps-Valls, J Muñoz-Marí, J P Rivera, F Veroustraete, J G P W Clevers & J Moreno, 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. ISPRS Journal of Photogrammetry and Remote Sensing, 108: 273-290

28 Hansen P M & J K Schjoerring, 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sensing of Environment, 86(4): 542-553

29 Pimstein A, A Karnieli & D J Bonfil, 2007. Wheat and maize monitoring based on ground spectral measurements and multivariate data analysis. Journal of Applied Remote Sensing, 1: 013530

30 Darvishzadeh R, C Atzberger, A Skidmore & M Schlerf, 2011. Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. ISPRS Journal of Photogrammetry and Remote Sensing, 66: 894-906

31 Siegmann B & T Jarmer, 2015. Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data. International Journal of Remote Sensing, 36(18): 4519-4534

32 Thulin S, M J Hill, A Held, S Jones & P Woodgate, 2012. Hyperspectral determination of feed quality constituents in temperate pastures: Effect of processing methods on predictive relationships from partial least squares regression. International Journal of Applied Earth Observation and Geoinformation, 19: 322-334

33 An N, A L Goldsby, K P Price & D J Bremer, 2015. Using hyperspectral radiometry to predict the green leaf area index of turfgrass. International Journal of Remote Sensing, 36(5): 1470-1483

34 Rivera Caicedo J P, Verrelst J, Muñoz-Marí J, Moreno J, 2014. Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(4): 1249-1259

35 Dorigo W, R Richter, F Baret, R Bamler & W Wagner, 2009. Enhanced automated canopy characterization from hyperspectral data by a novel two step radiative transfer model inversion approach. Remote Sensing, 1: 1139-1170

36 Richter K, T B Hank, F Vuolo, W Mauser & G D'Urso, 2012. Optimal exploitation of the Sentinel-2 spectral capabilities for crop leaf area index mapping. Remote Sensing, 4: 561-582

37 Jacquemoud S, W Verhoef, F Baret, C Bacour, P J Zarco-Tejada, G P Asner, C François & S L Ustin, 2009. PROSPECT+SAIL models: A review of use for vegetation characterization. Remote Sensing of Environment, 113: S56-S66

38 Combal B, F Baret, M Weiss, A Trubuil, D Macé, A Pragnère, R Myneni, Y Knyazikhin & L Wang, 2002. Retrieval of canopy biophysical variables from bidirectional reflectance Using prior information to solve the ill-posed inverse problem. Remote Sensing of Environment, 84: 1-15

39 Jarmer T, 2013. Spectroscopy and hyperspectral imagery for monitoring summer barley. International Journal of Remote Sensing, 34: 6067-6078

40 Savitzky A & M J E Golay, 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry, 36(8): 1627-1639

41 Martens H & T Næs, 1989. Multivariate Calibration (John Wiley & Sons) 504 pp.

42 Wold S, M Sjöström & L Eriksson, 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems, 58: 109-130

43 CAMO Software AS, 2015. The Unscrambler® X version 10.2 Help Manual

44 Lilienthal H & E Schnug, 2010. Bodengestützte Erfassung räumlich hochaufgelöster Hyperspektraldaten. Das Penta-Spek System. Bornimer Agrartechnische Berichte, 73, edited by M Zude & M Kraft (16. Workshop Computer-Bildanalyse in der Landwirtschaft - Computerized Image Analysis in Agriculture, Braunschweig) 86-93

45 Exelis Visual Information Solutions, Inc., 2014. ENVI 5.2 Help Manual

46 Chang C-W, D A Laird, M J Mausbach & C R Hurburgh Jr., 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Science Society of America Journal, 65: 480-490

47 Bellon-Maurel V, E Fernandez-Ahumada, B Palagos, J-M Roger & A McBratney, 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. Trends in Analytical Chemistry, 29(9): 1073-1081

48 Haenlein M & A M Kaplan, 2004. A beginner's guide to partial least squares analysis. Understanding statistics, 3(4): 283-297

49 Kaufmann H, L Guanter, K Segl, S Chabrillat, S Hofer, K-P Foerster, B Sang, T Stuffler, A Mueller, R Mueller, R Richter, C Chlebek & G Rossner, 2010. EnMAP – An advanced optical payload for Earth observation – Abstracts. Art, Science and Applications of Reflectance Spectroscopy Symposium (Boulder, Colorado 2010).

50 Segl K, R Richter, T Küster & H Kaufmann, 2012. End-to-end sensor simulation for spectral band selection and optimization with application to the Sentinel-2 mission. Applied Optics, 51(4): 439-449

51 BlackBridge, 2015. Satellite Imagery Product Specifications. Version 6.0, April 2015, URL:

52 Roy D P, M A Wulder, T R Loveland, C E Woodcock, R G Allen, M C Anderson, D Helder, J R Irons, D M Johnson, R Kennedy, T A Scambos, C B Schaaf, J R Schott, Y Sheng , E F Vermote, A S Belward, R Bindschadler, W B Cohen, F Gao, J D Hipple, P Hostert, J Huntington, C O Justice, A Kilic, V Kovalskyy, Z P Lee, L Lymburner, J G Masek, J McCorkel, Y Shuai, R Trezza, J Vogelmann, R H Wynne & Z Zhu, 2014. Landsat-8: Science and product vision for terrestrial global change research, Remote Sensing of Environment, 145: 154-172

53 Pu R, P Gong & Q Yu, 2008. Comparative analysis of EO-1 ALI and Hyperion, and Landsat ETM+ data for mapping forest crown closure and Leaf Area Index. Sensors, 8: 3744-3766

54 Frampton W J, J Dash, G Watmough & E J Milton, 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. Journal of Photogrammetry and Remote Sensing, 82: 83-92

55 Herrmann I, A Pimstein, A Karnieli, Y Cohen, V Alchantis D J Bonfil, 2011. LAI assessment of wheat and potato crops by VENμs and Sentinel-2 bands. Remote Sensing of Environment, 114: 2141-2151