PFG 2015 / 1, 0055–0067
Stuttgart, February 2015

Article

Comparison of two Statistical Methods for the Derivation of the Fraction of Absorbed Photosynthetic Active Radiation for Cotton

Sylvia Lex, Sarah Asam, Fabian Löw & Christopher Conrad, Würzburg

Keywords: FAPAR, agriculture, RapidEye, NDVI, empirical regression, percentile approach

Summary: The fraction of absorbed photosynthetic active radiation (FAPAR) is an important input for modelling biomass increase and agricultural yield and can be calculated based on optical remote sensing data. In this study two remote sensing based approaches to derive the FAPAR for irrigated cotton in Fergana valley, Uzbekistan, are tested and compared: (i) FAPAR rescale from the normalized difference vegetation index (NDVI) ("percentile approach"), and (ii) an empirical regression approach based on NDVI. In the rescaling approach FAPAR was derived by relating upper and lower percentiles derived from the NDVI distribution of cotton fields from the entire study area to fixed FAPAR minima (bare soil) and maxima. NDVI was derived from multi-temporal 6.5 m RapidEye data acquired throughout 2011. For the regression approach FAPAR data was collected *in situ* from cotton fields during the vegetation season. The percentile approach delivered an RMSE of 0.10 whilst regression was only slightly better with an RMSE of 0.07. Hence, the percentile approach could be concluded as being a fast and easy alternative to field data demanding empirical regressions for the derivation of FAPAR on cotton fields.

Zusammenfassung: *Vergleich zweier statistischer Methoden zur Ableitung des Anteils absorbierter Photosynthese wirksamer Strahlung (FAPAR) für Baumwolle.* Der Anteil absorbierter photosynthesewirksamer Strahlung (FAPAR), welcher ein wichtiger Inputparameter für die Modellierung von Biomassezuwachs und Ernteerträgen ist, kann aus optischen Fernerkundungsdaten abgeleitet werden. In dieser Studie werden zwei verschiedene, fernerkundungsbasierte Verfahren zur Ableitung von FAPAR für Baumwollfelder miteinander verglichen: (i) die direkte Ableitung von FAPAR durch Umskalierung des Normalized Difference Vegetation Index (NDVI) („Perzentil-Ansatz") und (ii) eine empirische Regression von FAPAR-Feldmessungen mit dem NDVI. Bei der Umskalierung werden obere und untere Grenzwerte der NDVI-Verteilung (Perzentile) in Beziehung zu FAPAR Minimum und Maximum-Werten gesetzt. Die erforderlichen Geländedaten wurden während der Vegetationsperiode 2011 im Ferghanatal in Usbekistan aufgenommen. Der NDVI wurde aus multi-temporalen RapidEye-Daten desselben Jahres berechnet. Die direkte Ableitung von FAPAR resultiert in einem RMSE von 0.10, während mit der empirischen Regression nur ein leicht besseres Ergebnis (RMSE = 0.07) erzielt werden kann. Die Ergebnisse lassen auf eine gute Eignung des vergleichsweise einfachen Perzentil-Ansatzes als Alternative zu messdatenintensiven Methoden zur Bestimmung von FAPAR auf Baumwollfeldern schließen.

1 Introduction

Biophysical parameters like the fraction of absorbed photosynthetic active radiation (FAPAR) are important vegetation parameters for environmental monitoring. The FAPAR characterizes the potential of vegetation to absorb energy (Myneni et al. 1997) and is defined as the fraction of radiation that is absorbed by the vegetation canopy in the visible

www.schweizerbart.de
DOI: 10.1127/pfg/2015/0250 1432-8364/15/0250 $ 3.50

light (400 nm – 700 nm) for photosynthesis (Monteith 1972). Due to these properties, multi-temporal FAPAR can be used for modelling biomass accumulation and yield in agriculture (Delécolle et al. 1992).

Radiative transfer models (RTM) can achieve very high accuracies when modelling FAPAR but require numerous sensor- and site-specific input data such as illumination and viewing geometries, leaf structure, chlorophyll content or dry matter content (Jacquemoud et al. 2009). On the basis of RTM, a near-linear relationship between FAPAR and remotely sensed vegetation indices (VIs) was found (Choudhury 1987, Goward & Huemmrich 1992) which was the fundament for statistical FAPAR modelling.

The relation between VIs and FAPAR has been investigated for different biomes from in situ measurements to the scale of moderate resolution sensors. Linear regressions of in situ FAPAR and normalized difference VI (NDVI) measurements were found for grassland and agriculture in West Africa by Fensholt et al. (2004). Cristiano et al. (2010) systematically compared numerous VIs with FAPAR measurements of two different grass types in Argentina within the vegetation season and found green NDVI, NDVI and the optimized soil adjusted VI to be the best statistical estimators for FAPAR. Logarithmic correlations modestly outperformed linear correlations. Differences in plant architecture of the grass types or plant stress situations influenced the slope of the statistical relation between the VIs and FAPAR only negligibly. Increasing stress in turn negatively affected the strength of the correlation results. Investigations of the statistical relation between FAPAR and VIs derived from satellite data revealed a significant influence of vegetation cover fractions on FAPAR (Asrar et al. 1992). Others pointed at the necessity to carefully select the appropriate VI in order to avoid impact of the soil background colour, or to minimize errors for different vegetation classes (e.g. Choudhury 1987, Sellers et al. 1994, 1996).

In the light of the almost linear relationship between FAPAR and VIs, Sellers et al. (1994, 1996) proposed a straight forward method for global FAPAR retrieval independent from in situ data, based on Advanced Very High Resolution Radiometer (AVHRR) data. Percentiles from occurring NDVI values indicate the range of possible FAPAR values, e.g. in case of numerous vegetation classes, 98% and 5% NDVI refer to FAPAR values of 0.95 and 0.001, respectively (Sellers et al. 1996). NDVI-percentiles were calculated for each class separately. In this global approach, the NDVI percentiles were extracted from all available AVHRR pixel values of one vegetation class after correcting for varying illumination angles caused by different geographical latitudes (Sellers et al. 1994, 1996). As Sellers et al. (1992) empirically found a near-linearity between the FAPAR and simple ratio (SR: ratio between near infrared and red reflectance), Sellers et al. (1994) used SR for calculating FAPAR between the extreme values (percentiles) in the original formulation of their approach. However, the authors were aware that other VIs may better suit for some vegetation classes and soil types, as for instance shown by Choudhury (1987) who selected the NDVI for deriving FAPAR of crops on bright soils. Examples for the implementation of NDVI instead of SR in a percentile approach are given by Olofsson & Eklundh (2007) and Olofsson et al. (2007). The authors utilized Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI for FAPAR modelling on forested sites in Sweden and Denmark.

Only few studies transferred the approach of Sellers et al. (1996) to high resolution data. Lobell et al. (2003) compared several combinations of VIs for wheat in agricultural areas of Mexico based on multi-temporal 30 m Landsat data (TM and ETM+). They tested (i) SR and (ii) NDVI using the percentile approach. With the resulting FAPAR from (i) and (ii) they calculated the mean FAPAR of both approaches (iii). Yield was calculated and the best results, i.e. mostly matching regional yield statistics, were obtained by averaging the estimates of the FAPAR from SR and NDVI.

This study targets at the application of the approach of Sellers et al. (1994, 1996) to multi-temporal high spatial resolution RapidEye data of 2011 for deriving the FAPAR in cotton ecosystems of Fergana Valley in Uzbekistan. In this region, area-wide approaches for mapping vegetation parameters without or using minimum field data appear to be useful for land and water management, because information on crop growth and crop yields is rare. The percentile approach was employed in three different experiments using literature values and satellite data derived soil background values for calculating the FAPAR range for the VI percentiles (NDVI and SR). The results were compared with simple linear regression results between field measurements of the FAPAR and NDVI. The NDVI was selected because it was identified as the best performing VI for the FAPAR derivation via linear regression in a pre-study conducted in the same study region (Lex et al. 2013). The test-plots were located on two farms with different cotton cultivation practises.

2 Study Area

The study area Fergana valley is located in the eastern part of Uzbekistan in the upstream part of the Syr Darja River. The climate is dry with a mean precipitation ranging from 100 mm to 500 mm and a mean temperature of 13.1 °C (Abdullaev et al. 2009). It is one of the largest and most intensively used agricultural areas in Uzbekistan where one third of the country's population lives (Reddy et al. 2012). Main crops in this region are winter-wheat, cotton, rice, and orchards. Vegetables like water-melons and tomatoes are also planted (Conrad et al. 2013). As the evaporation rate is about 1,200 mm per year, which exceeds precipitation by far, irrigation of the agricultural fields is necessary (Reddy et al. 2013).

Two locations in the Fergana valley were selected as study sites: The Water User Association (WUA) "Akbarabad" and the WUA "Azizbek" (see Fig. 1). In Akbarabad, cotton is sown at a row distance of 60 cm while in Azizbek farmers planted in 90 cm distance. Sowing date for cotton in both study sites was in the first half of April in 2011. Harvesting took place between the end of September and middle of October. All fields are characterized by similar soils which are loamy with an average content of sand between 40 and 50%, a silt content of 50 to 60% and a small clay fraction between 0 and 20% (Kenjabaev et al. 2013).

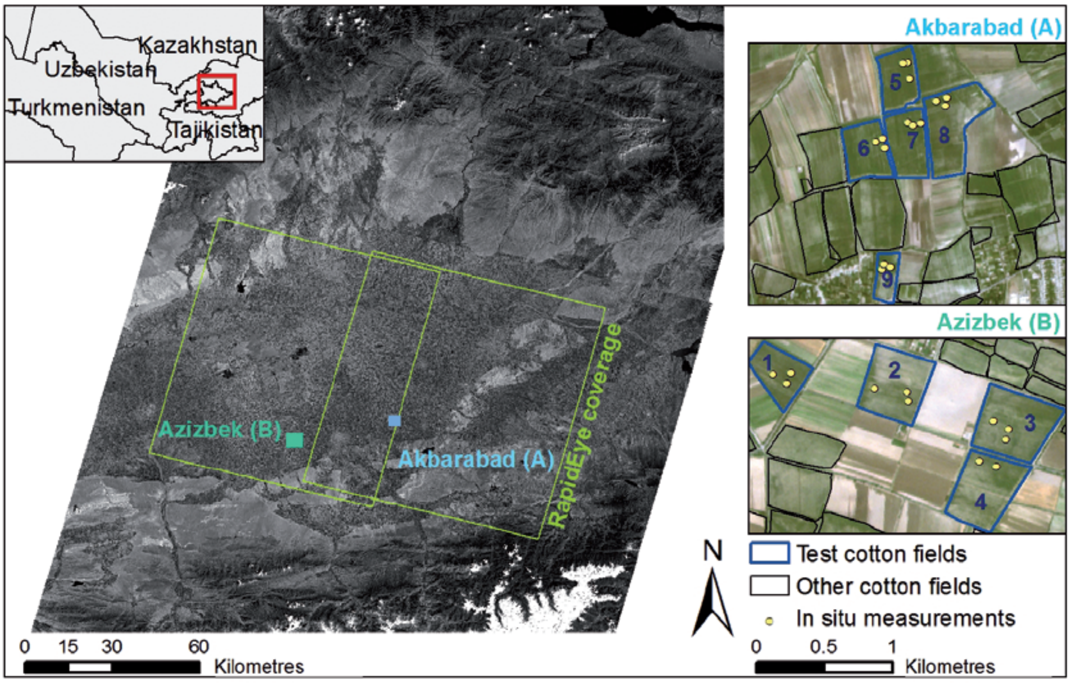

Fig. 1: The two study sites are located in the Fergana valley, Uzbekistan: Akbarabad (A) and Azizbek (B). RapidEye data, shown in the maps on the right, is from 29.7.2011 and displayed as true colour composite. Blue lines indicate the boundaries of test cotton fields and the *in situ* measurements are presented with yellow dots. The background map consists of two Landsat – TM images (only red band) from 3.6.2011 (southern part) and 22.8.2011 (northern part).

3 Data and Methods

3.1 *Field Campaign*

In total, nine cotton fields were selected within the two study sites in the south-eastern part of the valley (see Fig. 1). Four fields were located in the WUA "Azizbek" (study site B: Fields 1, 2, 3, 4) and five in the WUA "Akbarabad" (Study site A: Fields 5, 6, 7, 8, 9). The FAPAR ground truth data was collected between the beginning of June and the end of September 2011. On each field three environmental sampling units (ESUs) were established with each ESU consisting of twelve field points. Only on field 4 two ESUs were measured. An ESU describes a comparatively homogeneous area with a size of 20 m × 20 m. The twelve field points should integrate the within ESU variability of the field measurements. It has to be noted that in the later analysis, segments of RapidEye data were analysed, which completely covered the ESU. The total number of field points was 312 (26 ESUs with twelve field points each). Each field point was visited six times during the season (Tab. 1).

At each field point the FAPAR was measured using the Ceptometer AccuPAR LP-80 (Decagon Devices 2013). It is equipped with an 80 cm long bar with 80 equally distributed sensors enabling measurements of the photosynthetic active radiation (PAR) above and beneath the canopy. The average plant height was measured with a folding metre stick.

3.2 *Pre-Processing of Satellite Data*

RapidEye data (Tyc et al. 2005) was acquired throughout the measurement campaign (Tab. 1). In order to cover both study sites, two different RapidEye paths were analysed (Fig. 1). Each scene was atmospherically corrected using ATCOR2 (Richter & Schlaepfer 2012) within the processing chain CATENA that was developed at the German Aerospace Center (DLR) (DLR 2014). Geometric correction was done with the software ERDAS Autosync-Module (ERDAS 2010) with a resulting RMSE of below 2.2 m as described in Lex et al. (2013).

In situ data was not compared to pixel values but to a homogeneous surrounding of the field points. This step should account for both, uncertainties between the geolocation of in situ and satellite data, which may lead for incorrect sample pairs of in situ FAPAR and RapidEye NDVI, and the scale difference between 6.5 m RapidEye pixels and the ESUs. To cope with the prescribed uncertainty the data was segmented using the software eCognition (Trimble Germany GmbH 2010) as proposed by Fritsch et al. (2012). Therefore, a relatively small scale parameter (10) and the parameters shape and compactness (set to 0.9 each) were used. By doing so, a sub-segmentation of the cotton fields was achieved. For every available RapidEye acquisition, NDVI was calculated and spatially averaged within these sub-segments of the cotton fields.

Tab. 1: Acquisition timing of RapidEye data and corresponding field sampling campaigns (adapted from Lex et al. 2013)

Site A (Akbarabad)			**Site B (Azizbek)**		
RapidEye Acquisition	Beginning of fieldwork	End of fieldwork	RapidEye acquisition	Beginning of fieldwork	End of fieldwork
07.06.2011	08.06.2011	15.06.2011	15.06.2011	14.06.2011	16.06.2011
24.06.2011	24.06.2011	28.06.2011	23.06.2011	25.06.2011	29.06.2011
23.07.2011	15.07.2011	23.07.2011	07.07.2011	11.07.2011	13.07.2011
29.07.2011	30.07.2011	01.08.2011	29.07.2011	21.07.2011	29.07.2011
07.08.2011	10.08.2011	16.08.2011	07.08.2011	12.08.2011	15.08.2011
31.08.2011	26.08.2011	05.09.2011	23.08.2011	24.08.2011	25.08.2011

3.3 *Percentile Approach*

For the derivation of the FAPAR the formula proposed by Sellers et al. (1996) was utilized in this study (1):

$$FAPAR = \frac{(VI_i - VI_{98}) * (FAPAR_{\max} - FAPAR_{\min})}{VI_{98} - VI_5} + FAPAR_{\min} \quad (1)$$

where $FAPAR_{max}$ ($FAPAR_{min}$) refers to the maximum (minimum) possible FAPAR value of one class within the entire study area, VI_{98} (VI_5) corresponds with VI-values of the 98% (5%) percentile and VI_i is the actual VI-value at pixel i. In their global approach Sellers et al. (1996) suggested for vegetation classes that the FAPAR values of 0.950 ($FAPAR_{max}$) and 0.001 ($FAPAR_{min}$) denote the VI_{98} and VI_5, respectively.

For the RapidEye data in the Fergana Valley four variants of the original percentile approach were implemented using two different VIs and one alternative for the VI_5 percentile. First, the VIs were varied and (1) was implemented for cotton with the aforementioned propositions for the FAPAR extremes and VI percentiles. Here, simple ratio (SR, (2)) and the NDVI (3) were selected for implementing the FAPAR retrieval (SR: SEL-1, NDVI: SEL-2).

$$SR = \frac{NIR}{RED} \quad (2)$$

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3)$$

Second, VI_5 was substituted with the VI value of bare soil. This step was based on the assumption that taking into account only satellite measurements during the vegetation period within a comparatively small geographic region will underestimate FAPAR in early crop stages. In such cases, when all satellite data show crop cover, the 5% percentile of VI data could already reflect vegetation activity and in consequence FAPAR higher than 0.001, which was set in SEL-1 and SEL-2. Therefore, one additional RapidEye acquisition from 8.4.2011 was analysed. Accordingly, SEL-3 (SEL-4) is based on SR (NDVI) from pre-season bare soil situation.

VI analysis was conducted on a segment level. The percentiles (VI_{98} and VI_5) were calculated over all scenes of the RapidEye time series between the beginning of June and the end of September. This VI-value for bare soil in SEL-3 and SEL-4 is the mean value of all pixels within all sub-segments on cotton fields (section 3.2).

3.4 *Empirical Regression Approach*

Based on the in situ FAPAR data and the NDVI mean values of the sample segments, paired for corresponding acquisition dates (see Tab. 1), a simple linear regression equation including all observations aggregated at ESU level was established (REG). NDVI was selected as VI because according to Lex et al. (2013) best empirical regression results were obtained with SAVI and NDVI. As the NDVI is used in the percentile approach of Sellers et al. (1996), in this study NDVI was chosen. To account for the fact that study sites A and B varied in their management practices (section 2) and for a better assessment of the stability of the regression parameters (slope and offset) over the area, the regression approach was also separately applied to the two study sites (REG-A, REG-B).

A regression equation can only be established if the residuals of the model fit are normally distributed (Bahrenberg et al. 2010). The Lilliefors-Test is also valid for few observations (Lilliefors 1967) and was selected in this study to test the normal distribution of the residuals.

3.5 *Validation and Comparison*

Three quantitative assessment measures were applied for validating and comparing the quality of all experiments. This comparison between the results in turn enabled the identification of the most accurate approach and the evaluation of the quality difference when using the simplified percentile approach instead of a field work intensive regression.

The root-mean-square error (RMSE, (4)) returns an average absolute deviation:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \in_i^2} \quad (4),$$

where n is the number of FAPAR measurements and $\in$ is the deviation between model result and the respective *in situ* data. The mean direction of deviation is given by the BIAS (5):

$$BIAS = \frac{1}{n}\sum_{i=1}^{n} \in_i \quad (5).$$

The coefficient of determination (R^2) was employed for estimating the strength of the statistical correlation between the model output and the field measurements.

4 Results and Discussion

4.1 *Analyses of in situ Data*

The temporal development of the FAPAR varied between all cotton fields during the entire vegetation period. However, the most remarkable discrimination was found between the temporal development of the fields in the study sites A and B (Akbarabad, Azizbek). Already in time step 2 (Site A: 23.6.2011, Site B: 24.6.2011) FAPAR increases more slowly in study site B than in study site A (Akbarabad) (Fig. 2a). Later, in time steps 3 to 6 (7.7.2011/23.7.2011, 29.7.2011, 7.8.2011, and 23.8.2011/31.8.2011), the boxplots of the fields in Azizbek (B) remained on a significant lower FAPAR level than in Akbarabad (A).

Plant heights developed similar to the FAPAR within the season (Fig. 2b). Plants on the study site A were permanently higher (average of 76 cm) than those on the study site B (average of 63 cm). Both, FAPAR and plant height indicate a comparatively high vegetation cover in site A. The latter can be substantiated by taking into account the differences in cotton cultivation practises. Farmers in the study site B grow cotton with a row distance of 90 cm, whilst fields in the study site A were organized in rows with distances of 60 cm (section 2).

Varying vegetation heights within neighboured observations indicate a more heterogeneous structure of the vegetation canopy. The latter is influencing the scattering of light as described by Myneni & Williams (1994). Accordingly, the distribution of FAPAR corresponded with the deviation of plant height measurements on both observation sites. For most field observation periods, FAPAR and plant height boxes show narrow distribution between the upper and lower box boundaries in study site B. In contrast, variability of both variables was closer to the observed mean values in the study site A.

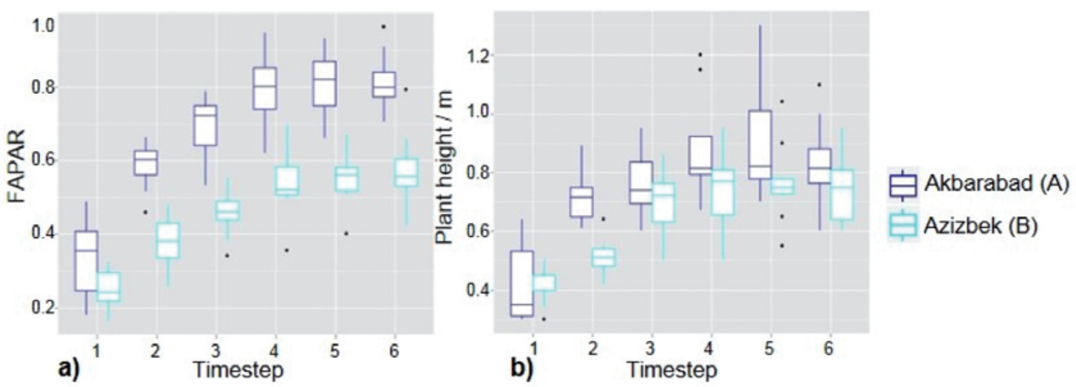

Fig. 2: a) Seasonal development of FAPAR b) development of plant heights for study sites A (Akbarabad, in blue colour) and B (Azizbek, in cyan colour); the box plots show the distribution of FAPAR and plant height values for each time step for both study sites A and B. The bar within the box is the median, the edges of the box show the interquartile range and the edges of the whiskers represent the minimum and maximum values. All values that are higher than 1.5 interquartile range are plotted as dots (FAPAR = fraction of absorbed photosynthetic active radiation).

4.2 *Percentile Approach*

Fig. 3a compares the in situ measured FAPAR with the results from the percentile approach experiments SEL-1 and SEL-2. For VI_{98} a NDVI value of 0.84 was received and VI_5 was found to be 0.14.

The FAPAR values, calculated based on SR (SEL-1), were lowest and fit the in situ data with a RMSE of 0.281 (R^2 = 0.72, BIAS = -39%, Tab. 2). SR was less sensitive to low FAPAR ranges, which led to an underestimation of in situ measurements (see blue dots in Fig. 4). The result queries the presumption of a linear relation between the SR and the FAPAR for cotton ecosystems in the study region.

The scatterplot between the FAPAR measured in situ and the FAPAR derived from the NDVI (SEL-2) shows a clear linear relation. The validation measures of SEL-2 (RMSE: 0.1, R^2: 0.87) exceeded that of SEL-1. However, the BIAS of 4.29% indicates a small overestimation of the model SEL-2 which can be assigned to high FAPAR levels (Fig. 3, red dots). These overestimations could be corrected by adjustment of the NDVI-percentiles, e.g. utilizing maximal field measurements of the FAPAR could be envisaged. The latter would go beyond the scope of this study, in which the percentile approach was tested as a variant for the FAPAR derivation without field measurements. The results of SEL-2 comply with observations of Olofsson & Eklundh (2007) who observed RMSE values between 0.03 and 0.67 using the percentile approach with NDVI for the FAPAR estimations for Scandinavian forest classes.

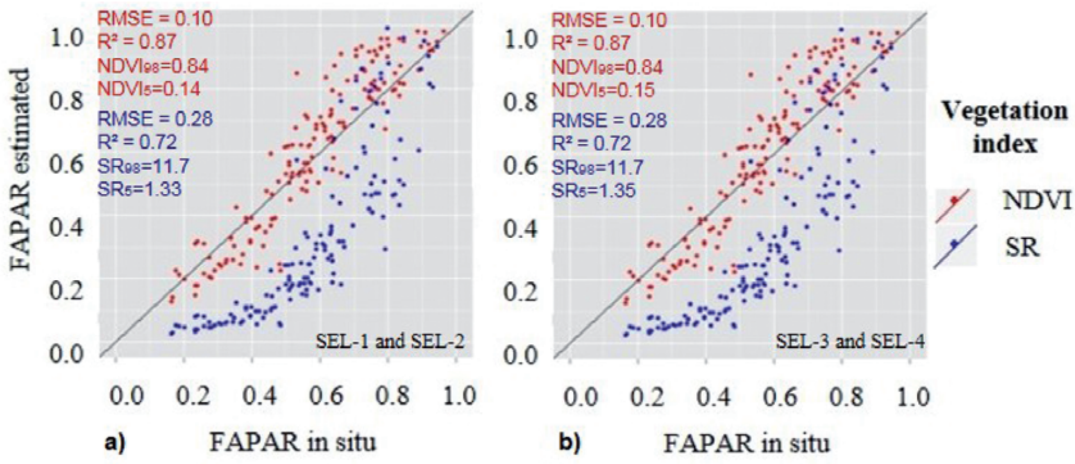

Fig. 3: a) Results of the percentile approach in contrast to the field measurements: with the SR (SEL-1, blue dots) and the NDVI (SEL-2, red dots); b) Results of the percentile approach in contrast to the field measurements: with the SR (SEL-3, blue dots) and NDVI (SEL-4, red dots).

Tab. 2: Assessment of the percentile approach (SEL-1 to SEL-4) based on BIAS, RMSE, and R^2.

Method	BIAS (in %)	RMSE	R²	Number of observations
SEL-1	-39.646	0.281	0.7227219	148
SEL-2	4.293	0.098	0.8741476	148
SEL-3	-39.865	0.283	0.7227219	148
SEL-4	3.612	0.099	0.8741476	148
SEL-2-A*	7.954	0.116	0.8001162	86
SEL-2-B*	-3.125	0.064	0.8903632	62

*A and B refer to the respective study sites.

The experiments SEL-3 and SEL-4 returned validation measures which are almost the same as those of SEL-1 and SEL-2, respectively (Fig. 3b, Tab. 2). The lower percentile boundary (VI_5) used in SEL-1 and SEL-2 was with a NDVI value of 0.14 (SR value of 1.33) very similar to the soil VI approach (SEL-3 and SEL-4), which revealed a NDVI value of 0.15 (SR-value of 1.35) for VI_5. The VI values composing VI_5 in SEL-1 and SEL-2 originated all from the RapidEye observations in June (7.6.2011). Distinguishing between NDVI acquired in June and April (8.4.2011) led to negligible differences, which could be confirmed by visual inspection. The results show that the percentile approach can be easily applied to multi-temporal remote sensing data if sufficient NDVI values represent soil conditions. In this study, despite including only data after sowing, soil representing NDVI values were suffered due to the presence of one time step in initial growing phases. However, SEL-4 presented a feasible variant for the FAPAR derivation in case of absent early season data characterized by a very low vegetation cover.

The performance of the percentile approach was further analyzed separately for the study sites A and B. Only SEL-2 is presented as this analysis revealed the lowest absolute numerical deviation between model results and in situ measurements expressed by the RMSE (Tab. 2). SEL-4, which delivered nearly the same results as SEL-2, due to the similar value of VI_5 for SEL-2 and SEL-4, is not presented here as the resulting FAPAR was very close to that from SEL-2. The scatterplots between the FAPAR, derived by the NDVI-percentiles, and field measurements (Fig. 4) show a higher deviation from the theoretically perfect result (diagonal line) for the study site A than for the study site B. Accordingly, the RMSEs of 0.06 and 0.12 approve a slightly higher performance of the SEL-2 model in Azizbek (study site B) than in Akbarabad (study site A). These differences might be explained by the growing conditions differing between the study sites as described in section 4.1. In the study site A the cotton plants are higher than those in the study site B. A taller plant drops more shadow than a smaller plant, especially during hours that are deviating from noon, when lower sun elevation occurs. Additionally, the plant heights observed in study site A within one time step show higher variability than in study site B. As a result shadow lengths are not constant. In case of small row distances as recorded in the study site A, the shadow dropping of one cotton row might fall into the next row. Accordingly, uncertainties of the FAPAR measurements in situ, as well as the satellite derived NDVI are higher than in the study site B, which in total confirms findings of Tewolde et al. (2005). However, the high precision level of both study sites indicate the stability of the percentile approach to derive the FAPAR of cotton under different cultivation practices.

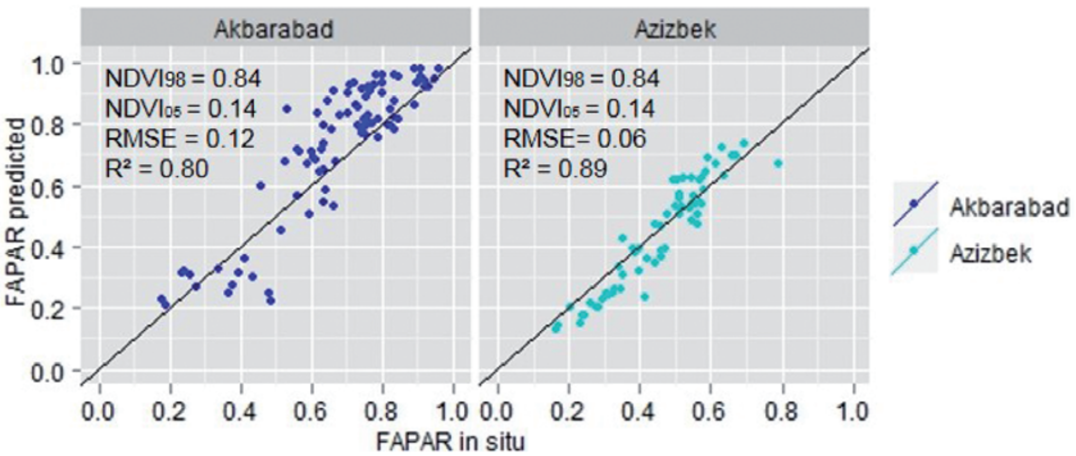

Fig. 4: Percentile approach with NDVI (SEL-2) in the study sites Akbarabad (left) and Azizbek (right).

4.3 *Simple Linear Regression*

The Lilliefors test for normal distribution of the residuals, which is the precondition for empirical regressions, delivered a p-value of 0.8753 for the regression approach applied to the entire set of field samples and NDVI values (REG). Accordingly, the hypothesis of normal distribution cannot be rejected as the p-value exceeded 0.05. Also for the separate analysis of the two study sites, the Lilliefors test showed a normal distribution of the residuals ($p = 0.5796$ for REG-A and $p = 0.2006$ for REG-B).

The statistical analysis of the entire dataset (REG) achieved an RMSE of 0.07 and R^2 of 0.87 (Tab. 3). According to the statistical nature of high correlation coefficients BIAS was negligible (<< 0.01%). Previous studies achieved similar RMSE values for FAPAR derivation using linear regression, e.g. 0.086 in an agricultural landscape in Spain based on Landsat NDVI (Ridao et al. 1998), or R^2 of 0.61 in a steppe landscape in South America modelled with MODIS NDVI (Cristiano et al. 2010) and 0.86 – 0.96 in the savannah in Senegal, also based on MODIS NDVI (Fensholt et al. 2004).

Similar statistical relations were found for the two study sites, however, the results varied slightly. The coefficient of determination for study site B (REG-B) was with a value of 0.89 higher than for study site A (REG-A), where R^2 was 0.80. Also RMSE (0.05) of REG-B, i.e. the total deviation between statistically derived FAPAR and in situ measurements was lower than of REG-A (0.08).

The same reasons as stated for the percentile approach, i.e. high variability of plant heights and a resulting inhomogeneity in site A, can explain these differences between the results among the study sites. The scattering of the value pairs from the FAPAR in situ data and the NDVI around the regression line for each study site is given in Fig. 5. The same patterns as for the percentile approach can be discovered.

In Azizbek (study site B) the FAPAR in situ data ranged from 0.17 to 0.79 while in Akbarabad (study site A) the data range was between 0.18 and 0.96. During the entire season the FAPAR-values of study site B- remained

Tab. 3: Assessment of the simple linear regression (REG, REG-A, REG-B) based on BIAS, RMSE, and R^2.

Method	BIAS (in %)	RMSE	R^2	Number of observations
REG	0.0001	0.07	0.87	148
REG-A[1)]	-0.0006	0.08	0.80	86
REG-B[1)]	-0.0001	0.05	0.89	62

1) A and B refer to the respective study sites.

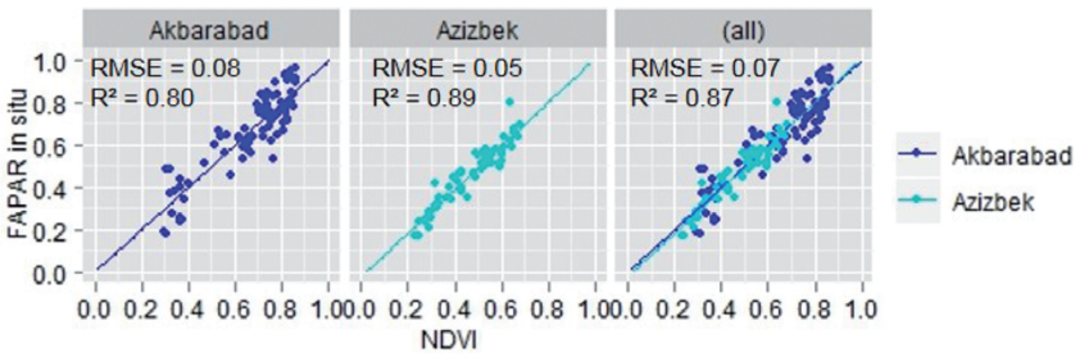

Fig. 5: Scatter plots of the in situ FAPAR and the NDVI, as well as regression lines for the study sites (Akbarabad, REG-A and Azizbek, REG-B) and both sites (REG).

Tab. 4: Regression equations (simple linear regression).

Method	Regression equations
REG	FAPAR = -0.01300 + 1.00965 * NDVI[1)]
REG-A[2)]	FAPAR = -0.00026 + 0.99297 * NDVI
REG-B[2)]	FAPAR = -0.02145 + 1.02324 * NDVI

1) Adopted from Lex et al. (2013)
2) A and B refer to the respective study sites.

below the FAPAR-level of study site A (see Fig. 2a). Even though value ranges differed between the two study sites, similar regression equations were derived for all datasets (Tab. 4 and Fig. 5). As stated above, the management can have an influence on the crop growth. However, the similarity of the slopes received from the regression demonstrate a negligible influence of the two cultivation practices (cotton row distances) observed in this study on the quality of the linear relationship between NDVI and FAPAR.

4.4 *Comparison of the Approaches*

A comparison of RMSE and BIAS indicates only minor quantitative deviations between the optimal percentile approach SEL-2 (and SEL-4) and the simple empirical regression (REG). Of course, REG matches the dataset exactly to field data, but it is noteworthy that no field data has been used for the derivation of FAPAR in the percentile approaches. Even in SEL-4 (inclusion of a period when fields are completely free of vegetation) only satellite datasets were utilized.

Multi-temporal FAPAR maps received from SEL-2 and REG within study site B are shown in Fig. 6. In the first periods (15.6.2011 and 6.7.2011) no differences in the colour levels occurred. Moderately higher values for FAPAR derived from the percentile approach are visible in the late season situation (right part of Fig. 6, 23.8.2011), which can be attributed to the aforementioned overestimations of FAPAR in case of high vegetation cover (section 4.1). However, the comparison of the maps shows the same spatial patterns of

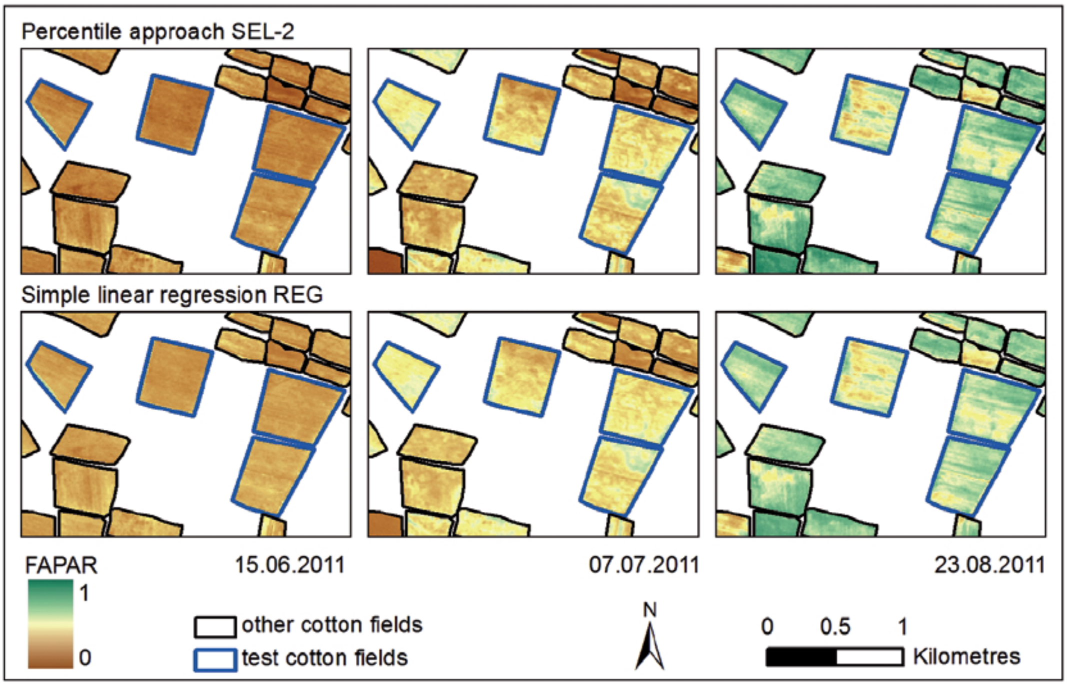

Fig. 6: FAPAR of cotton fields received from the experiments SEL-2 (upper part) and REG (lower part) in study site B, Azizbek, for the time steps 15.06.2011, 7.7.2011 and 23.8.2011, representing the early, mid, and late season situations of the vegetation period. On the fields with a blue outline *in situ* data were collected.

FAPAR variations throughout the vegetation period and underlines the similarity of the approaches.

5 Conclusions

In this study two methods for deriving the FAPAR from multi-temporal RapidEye data for cotton were compared. The study was implemented in two irrigated agricultural study sites in the Fergana Valley, Uzbekistan. As base line, an empirical regression that relates in situ FAPAR measurements to NDVI, computed from RapidEye data at six satellite image acquisitions over the growing season, was established. Furthermore, the FAPAR was calculated by linearly scaling the NDVI percentiles to maxima and minima of class-specific FAPAR values, following the approach of Sellers et al. (1996). The performances of the two methods were compared by opposing RMSE, BIAS, and the coefficient of determination revealed from correlation between the in situ FAPAR and modelled data. The maxima and minima of FAPAR, which were in the originally formulation assigned to the 98% and 5% percentile of the NDVI distribution, i.e. 0.95 and 0.001, were transferred to the cotton class at the local scale. The application of the percentile approach, which is without field measurements, resulted in accuracies comparable to that of the linear regression, also on a similar accuracy level received from empirical experiments in other agro-ecosystems: The percentile approach delivered an RMSE of 0.10 whilst regression was only slightly better with an RMSE of 0.07. This demonstrated that the FAPAR of cotton fields could be derived independently from field measurements by applying the percentile approach. Differences in canopy diversity introduced by variable plant heights within the field or distances between the rows did not show a significant impact, neither on the empirical regression nor on the percentile approach.

However, generalized statements about the usefulness of the percentile approach for FAPAR estimations of cotton ecosystems are not easy to conclude. A wide range of NDVI values was available, because the entire vegetation period was covered by satellite data. RapidEye scenes were acquired in both the initial and the main vegetative phase of the vegetation period, when quasi non-vegetated soil cover and most dense vegetation cover was measured by the NDVI, respectively. Another issue to be considered is that this study was conducted in the Fergana Valley, hence conclusions from this study might not necessarily apply to other cotton ecosystems e.g. with other climatic conditions or environmental or management settings (soils, field sizes, cultivation practises). Thus, the presented study has a more explorative character in terms of transferability but it can be concluded as an encouraging approach for accurate FAPAR modelling without in situ data, which can be subsequently used for crop yield estimations of cotton.

Acknowledgements

The authors thank the foundation “fiat panis” for financing the field campaign and the German Aerospace Center (DLR) for providing data from the RapidEye Science Archive in RESA project 461. This study was partly conducted within the project CAWa (“Water for Central Asia”, FKZ AA7090002), financed by the German Foreign Office, and the project Techs4TimeS (FKZ 50EE1353), funded by the German Federal Ministry for Economic Affairs and Energy. The authors are grateful for the support of the Scientific Information Center of the Interstate Coordination Water Commission of the Central Asia (SIC ICWC) in Tashkent, Uzbekistan during the field campaign.

References

Abdullaev, I., Kazbekov, J., Manthritilake, H. & Jumaboev, K., 2009: Participatory water management at the main canal: A case from South Ferghana canal in Uzbekistan. – Agricultural Water Management **96** (2): 317–329.

Asrar, G., Myneni, R.B. & Choudhury, B.J., 1992: Spatial Heterogeneity in Vegetation Canopies and Remote Sensing of Absorbed Photosynthetically Active Radiation: A Modeling Study. – Remote Sensing of Environment **41** (2–3): 85–103.

Bahrenberg, G., Giese, E., Mevenkamp, N. & Nipper, J. (eds.), 2010: Statistische Methoden in der Geographie. – Univariate und bivariate Statistik. – 282 pp., Borntrager, Stuttgart.

Choudhury, B.J., 1987: Relationships Between Vegetation Indices, Radiation Absorption, and Net Photosynthesis Evaluated by a Sensitivity Analysis. – Remote Sensing of Environment **22** (2): 209–233.

Conrad, C., Rahmann, M., Machwitz, M., Stulina, G., Paeth, H. & Dech, S., 2013: Satellite based calculation of spatially distributed crop water requirements for cotton and wheat cultivation in Fergana Valley, Uzbekistan. – Global and Planetary Change **110:** 88–98.

Cristiano, P.M., Posse, G., Di Bella, C.M. & Jaimes, F.R., 2010: Uncertainties in FPAR estimation of grass canopies under different stress situations and differences in architecture. – International Journal of Remote Sensing **31** (15): 4095–4109.

Decagon Devices, 2013: AccuPAR PAR/LAI Ceptometer. – Model LP-80, 82, Decagon Devices Inc., Pullman, Washington DC, USA.

Delécolle, R., Maas, S.J., Guérif, M. & Baret, F., 1992: Remote sensing and crop production models: present trends. – ISPRS Journal of Photogrammetry and Remote Sensing **47** (2–3): 145–161.

DLR, 2014: http://www.dlr.de/eoc/desktopdefault.aspx/tabid-5444/9113_read-17838/ (21.3.2014).

ERDAS, 2010: IMAGINE AutoSync. – ERDAS Inc., Atlanta, GA, USA.

Fensholt, R., Sandholt, I. & Rasmussen, M.S., 2004: Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using *in situ* measurements. – Remote Sensing of Environment **91** (3–4): 490–507.

Fritsch, S., Machwitz, M., Ehammer, A., Conrad, C. & Dech, S., 2012: Validation of the collection 5 MODIS FPAR product in a heterogeneous agricultural landscape in arid Uzbekistan using multitemporal RapidEye imagery. – International Journal of Remote Sensing **33** (21): 6818–6837.

Goward, S.N. & Huemmrich, K.F., 1992: Vegetation Canopy PAR Absorptance and the Normalized Difference Vegetation Index: An Assessment Using the SAIL Model. – Remote Sensing of Environment **39** (2): 119–140.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C. & Ustin, S.L., 2009: PROSPECT+SAIL models: A review of use for vegetation characterization. – Remote Sensing of Environment **113:** 56.

Kenjabaev, S., Forkutsa, I., Bach, M. & Frede, H.G., 2013: Performance evaluation of the BUDGET model in simulating cotton and wheat yield and soil moisture in Fergana valley. – Ahmed, M.N. & Pawlowski, I. (eds.): International Conference and Young Researchers' Forum: 'Natural resource use in Central Asia: institutional challenges and the contribution of capacity building', Giessen.

Lex, S., Conrad, C. & Schorcht, G., 2013: Analyzing the seasonal relations between in situ fPAR / LAI of cotton and spectral information of RapidEye. – Borg, E., Daedelow, H. & Johnson, R. (eds.): From the Basics to the Service. – GITO, Berlin.

Lilliefors, H.W., 1967: On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. – Journal of the American Statistical Association **62** (318): 399–402.

Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J. & Benning, T.L., 2003: Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. – Agriculture, Ecosystems & Environment **94** (2): 205–220.

Monteith, J.L., 1972: Solar Radiation and Productivity in Tropical Ecosystems. – Journal of Applied Ecology **9** (3): 747–766.

Myneni, R.B. & Williams, D.L., 1994: On the relationship between FAPAR and NDVI. – Remote Sensing of Environment **49** (3): 200–211.

Myneni, R.B., Nemani, R., Ramakrishna, R. & Running, S.W., 1997: Estimation of Global Leaf Area Index and Absorbed Par Using Radiative Transfer Models. – IEEE Transactions on Geoscience and Remote Sensing **35** (6): 1380–1393.

Olofsson, P. & Eklundh, L., 2007: Estimation of absorbed PAR across Scandinavia from satellite measurements. Part II: Modeling and evaluating the fractional absorption. – Remote Sensing of Environment **110** (2): 240–251.

Olofsson, P., Eklundh, L., Lagergren, F., Jönsson, P. & Lindroth, A., 2007: Estimating net primary production for Scandinavian forests using data from Terra/MODIS. – Advances in Space Research. **39** (1): 125–130.

Reddy, J.M., Muhammedjanov, S., Jumaboev, K. & Eshmuratov, D., 2012: Analysis of cotton water productivity in Fergana Valley of Central Asia. – Agricultural Sciences **3** (6): 822–834.

Reddy, J.M., Jumaboev, K., Matyakubov, B. & Eshmuratov, D., 2013: Evaluation of furrow irrigation practices in Fergana Valley of Uzbekistan. – Agricultural Water Management **117:** 133–144.

Richter, R. & Schlaepfer, D., 2012: Atmospheric / Topographic Correction for Satellite Imagery. – ATCOR-2/3 User Guide, Version 8.2 BETA, February 2012.

Ridao, E., Conde, J.R. & Mínguez, M., 1998: Estimating fAPAR from Nine Vegetation Indices for Irrigated and Nonirrigated Faba Bean and Semileafless Pea Canopies. – Remote Sensing of Environment **66** (1): 87–100.

Sellers, P.J., Berry, J.A., Collatz, G.J., Field, C.B. & Hall, F.G., 1992: Canopy Reflectance, Photosynthesis, and Transpiration. III. A Reanalysis Using Improved Leaf Models and a New Canopy Integration Scheme. – Remote Sensing of Environment **42** (3): 187–216.

Sellers, P.J., Tucker, C.J., Collatz, G.J., Los, S.O., Justice, C.O., Dazlich, D.A. & Randall, D., 1994: A global 1° by 1° NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. – International Journal of Remote Sensing **15** (17): 3519–3545.

Sellers, P.J., Los, S.O., Tucker, C.J., Justice, C.O., Dazlich, D.A., Collatz, G.J. & Randall, D., 1996: A Revised Land Surface Parameterization (SiB2) for Atmospheric GCMs. Part II: The Generation of Global Fields of Terrestrial Biophysical Parameters from Satellite Data. – Journal of Climate **9:** 706–737.

Tewolde, H., Sistani, K.R., Rowe, D.E., Adelli, A. & Tsegaye, T., 2005: Estimating cotton leaf area index nondestructively with a light sensor. – Agronomy Journal **97:** 1158–1163.

Trimble Germany GmbH, 2010: eCognition® Developer 8.64.0. – User Guide, 250, Trimble Germany GmbH, München.

Tyc, G., Tulip, J., Schulten, D., Krischke, M. & Oxfort, M., 2005: The RapidEye mission design. – Acta Astronautica **56** (1–2): 213–219.

Addresses of the Authors:

Sylvia Lex, Sarah Asam, Fabian Löw & Christopher Conrad, Department of Remote Sensing, Department of Geography and Geology, Julius-Maximilians-Universität Würzburg, Oswald-Külpe-Weg 86, D-97074 Würzburg, Germany, Tel. +49-931-31-84140, e-mail: {sylvia.lex}{sarah.asam}{fabian.loew}{christopher.conrad}@uni-wuerzburg.de, web: http://www.fernerkundung.geographie.uni-wuerzburg.de

Manuskript eingereicht: Juni 2014
Angenommen: September 2014