

Mapping Leaf Area Index in subtropical upland ecosystems using RapidEye imagery and the randomForest algorithm

Philip Beckschäfer⁽¹⁾, Lutz Fehrmann⁽¹⁾, Rhett D Harrison⁽²⁾, Jianchu Xu⁽³⁾, Christoph Kleinn⁽¹⁾

Canopy leaf area, frequently quantified by the Leaf Area Index (LAI), serves as the dominant control over primary production, energy exchange, transpiration, and other physiological attributes related to ecosystem processes. Maps depicting the spatial distribution of LAI across the landscape are of particularly high value for a better understanding of ecosystem dynamics and processes, especially over large and remote areas. Moreover, LAI maps have the potential to be used by process models describing energy and mass exchanges in the biosphere/atmosphere system. In this article we assess the applicability of the RapidEye satellite system, whose sensor is optimized towards vegetation analyses, for mapping LAI along a disturbance gradient, ranging from heavily disturbed shrub land to mature mountain rainforest. By incorporating image texture features into the analysis, we aim at assessing the potential quality improvement of LAI maps and the reduction of uncertainties associated with LAI maps compared to maps based on Vegetation Indexes (VI) solely. We identified 22 out of the 59 image features as being relevant for predicting LAI. Among these, especially VIs were ranked high. In particular, the two VIs using RapidEye's RED-EDGE band stand out as the top two predictor variables. Nevertheless, map accuracy as quantified by the mean absolute error obtained from a 10-fold cross validation (MAE_{CV}) increased significantly if VIs and texture features are combined (MAE_{CV} = 0.56), compared to maps based on VIs only (MAE_{CV} = 0.62). We placed special emphasis on the uncertainties associated with the resulting map addressing that map users often treat uncertainty statements only in a pro-forma manner. Therefore, the LAI map was complemented with a map depicting the spatial distribution of the goodness-of-fit of the model, quantified by the mean absolute error (MAE), used for predictive mapping. From this an area weighted MAE (= 0.35) was calculated and compared to the unweighted MAE of 0.29. Mapping was done using randomForest, a widely used statistical modeling technique for predictive biological mapping.

Keywords: Ecosystem Monitoring, Forest and Vegetation Parameters, Leaf Area Index (LAI), Hemispherical Photography, Map Uncertainty, Vegetation Indexes, Image Texture, Xishuangbanna

Introduction

Canopy leaf area is the main factor for primary production, energy exchange, transpiration, and other physiological attributes

related to ecosystem processes (Pierce & Running 1988, Gower & Norman 1991, Bonan 1993, Turner et al. 1999, Asner et al. 2003). The Leaf Area Index (LAI), defined

□ (1) Chair of Forest Inventory and Remote Sensing, Georg-August-Universität Göttingen, Büsgenweg 5, D-37077 Göttingen (Germany); (2) Key Laboratory for Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, Mengla, 666303 Yunnan (China); (3) World Agroforestry Centre, ICRAF-Kunming Office, Kunming, 650201 Yunnan (China)

@ Philip Beckschäfer (philip.beckschaefer@gmail.com)

Received: Feb 06, 2013 - Accepted: May 13, 2013

Citation: Beckschäfer P, Fehrmann L, Harrison RD, Xu J, Kleinn C, 2014. Mapping Leaf Area Index in subtropical upland ecosystems using RapidEye imagery and the randomForest algorithm. *iForest* 7: 1-11 [online 2013-10-07] URL: <http://www.sisef.it/iforest/contents/?id=ifor0968-006>

Communicated by: Giorgio Matteucci

as the projected leaf area per unit ground surface area, is frequently used to quantify the canopy leaf area. Thus, LAI is one of the key biophysical variables required by many process models describing the soil/plant/atmosphere system (Baret & Buis 2008, Latifi & Galos 2010). Because canopies and some of their characteristics can be directly observed from above, LAI is among the major variables of interest in remote sensing analyses (Lee et al. 2004), and its importance has led to considerable efforts to map its distribution over a variety of spatial and temporal scales (Cohen et al. 2003, Morisette et al. 2006, Zhao & Popescu 2009, Latifi & Galos 2010, Tang et al. 2012). The majority of studies used (semi-) empirical relationships between LAI and combinations of spectral bands, namely vegetation indexes (VI), for LAI mapping (Baret & Buis 2008, Vuolo et al. 2010). In a review paper on the relationships between remotely-sensed VIs and canopy attributes, Glenn et al. (2008) point out that VIs are often strongly related to light dependent physiological processes occurring in the upper canopy, but often exhibit only moderate relationships to detailed features of canopy architecture, such as LAI. VIs are generally regarded as important but have their limitations since they only utilize a fraction of the spectral information available in remote sensing data (Gonsamo & Pellikka 2012). Similarly, it has been pointed out that there is not enough evidence that spectral reflectance in the visible and near-infrared is sufficient to estimate LAI in forests, especially under close canopy situations (Lee et al. 2004). From their review, Glenn et al. (2008) concluded that remote sensing models exclusively based on VIs to estimate LAI are, in particular, subject to error and uncertainty. Another fact to be considered is that LAI is a variable that cannot be directly measured in the field. Therefore, all efforts to correlate VIs derived from remote sensing interpretation to field observations are characterized by the lack of ground-level measurable parameters, as both information are results of modeling efforts.

Additional information derived from remote sensing images with the potential to improve LAI predictions are texture features. Texture features quantify the spatial variability of pixel values within a neighborhood defined by a moving window, and thus, complement the spectral information with a spatial component (Colombo et al. 2003). As the variation in texture is related to changes in the spatial distribution of vegetation (Wulder et al. 1998), image texture can be linked to the spatial distribution of vegetation (Colombo et al. 2003).

The scope of this paper is to investigate the applicability of RapidEye imagery, which is optimized towards vegetation analyses, for

LAI mapping along a disturbance gradient, ranging from heavily disturbed shrub-land to mature mountain rainforest. By incorporating image texture features into the analysis we aim at assessing the potential quality improvement of LAI maps and the reduction of uncertainties associated with LAI maps compared to those based on VIs solely (Glenn et al. 2008). To predict LAI we use the Random Forest (RF) algorithm (Breiman 2001), an increasingly and widely used statistical modeling technique for predictive biological mapping (Prasad et al. 2006).

Field data come from forest inventory plots located in the uplands of Xishuangbanna, China. Xishuangbanna is located in the transition zone between tropical southeastern Asia and subtropical and temperate China,

resulting in the region with the highest biodiversity in China (Zhang & Cao 1995, Li et al. 2007). This high biological diversity is threatened by the expansion of rubber (*Hevea brasiliensis*) plantations (Xu et al. 2005, Li et al. 2007, Xu et al. 2013). In the time from 1976 to 2003 Xishuangbanna's estimated forest cover declined from 69% to less than 50% in conjunction with a decline of mean size of forest patches from 217 to 115 ha (Li et al. 2009). The forest type most affected by the expansion of rubber plantations was tropical seasonal rain forest (Li et al. 2007). For a better understanding of ecosystem dynamics and processes over large and remote areas, as faced in this region, LAI maps are of particularly high value (Chapin III et al. 2011).

Methods

Field data

The study site is located in Mengsong Administrative Village, Jinghong County, Xishuangbanna, Yunnan, China at an elevation of 800-2000 m a.s.l. (UTM/WGS84: 47N 656355 E, 2377646 N - Fig. 1). The subtropical climate of this region is influenced by the Indian monsoon; it has an annual mean temperature of 18 °C and an average rainfall of 1600-1800 mm, of which 80% is concentrated from May to October. Vegetation varies with altitude and the mosaic distribution of primary to secondary forest according to micro-environments.

LAI data were assessed on 28 inventory plots in May 2011. Each plot consisted of nine subplots arranged on a square grid with 50 m spacing (Fig. 1, lower panel). Plots covered a gradient from heavily disturbed shrub land, through secondary regrowth to mature mountain rainforest. A probability sampling design was implemented: (1) to allow for a statistically sound assessment of LAI throughout the study site; and (2) because it contributes to scientifically defensible accuracy assessment (Stehman & Czaplewski 1998). Plot locations were selected applying double sampling for stratification. A 500x500 m point grid was placed over the RapidEye image of the study site and each grid point classified into shrub land, regrowing forest, mature forest, and other land cover (e.g., settlements, water bodies, and mining claims). To ensure that sample plots were distributed over the whole area, the study site was divided into 16 equally sized primary units. From these primary units, 12 were randomly selected and within each one mature forest plot and one regrowing forest grid-point were selected at random. One shrub land grid-point was randomly drawn from every second of the selected primary units. The 28 selected grid-points became the SW corner of the sample plots.

At each subplot center, hemispherical photographs were taken with a Nikon D70s digital single lens reflex camera equipped with a Sigma Circular Fisheye 4.5 mm 1:2.8 lens with a field of view of 180°. The camera was mounted on a tripod at 1.2 m height to characterize the canopy without the interfering presence of understory vegetation (Tagle et al. 2011). Vegetation within 0.5 m of the lens was removed, as this can lead to an inflation of the LAI estimate. The camera was leveled to face exactly the vertical using a bubble-level slotted into the flash socket. The camera was systematically orientated to magnetic north using a compass (Beaudet & Messier 2002). Photographs were taken without direct sunlight entering the lens (Rich 1989) in the early morning, late afternoon or on overcast days (Weiss et al. 2004). The basic camera settings were mode "P"

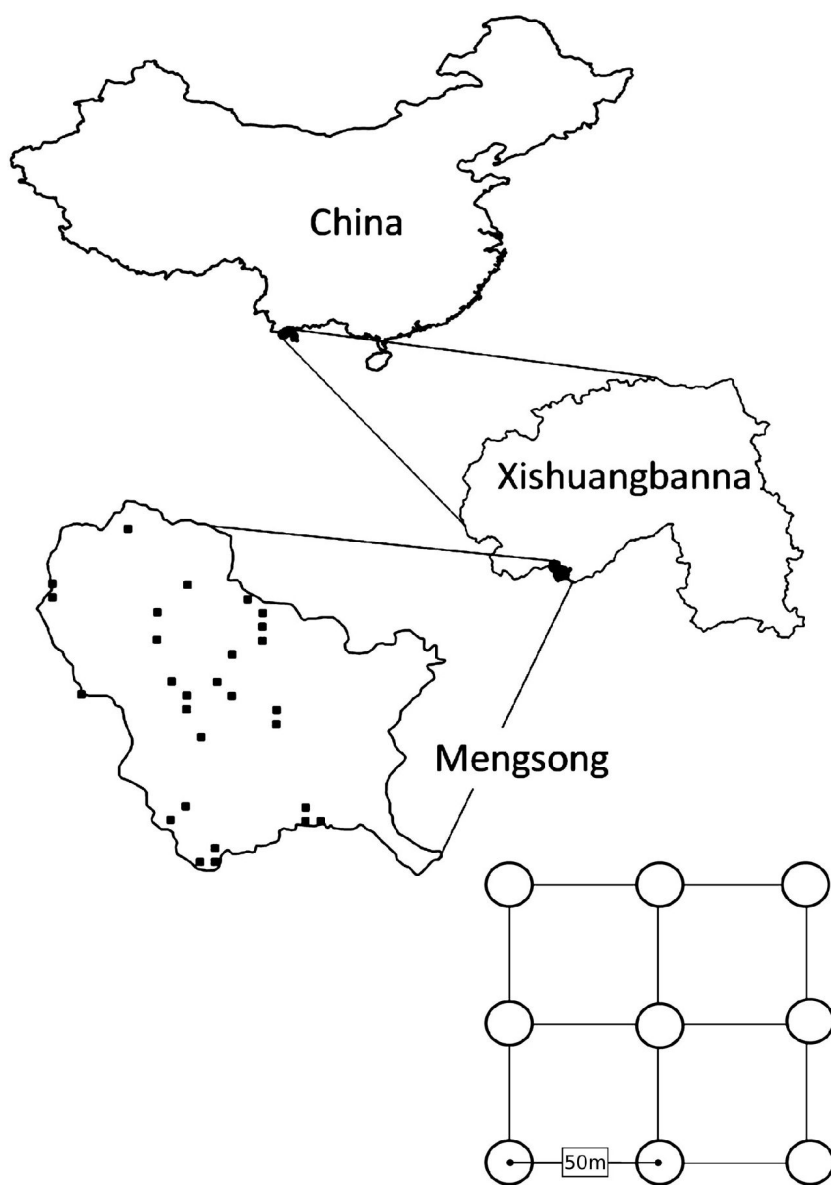


Fig. 1 - Location of the study site Mengsong in Xishuangbanna, China. Black squares in Mengsong map depict the locations of the 28 inventory plots. Plots consist of 9 subplots arranged on a square grid with 50 meter spacing.

Tab. 1 - Image features used as predictor variables.

RapidEye bands (wave length in nm)	BLUE (440 - 510), GREEN (520 - 590), RED (630 - 685), RED-EDGE (690 - 730), NIR (760 - 850)
Vegetation indexes	NDVI (Gamon et al. 1995), NDVI-RED-EDGE (Gitelson & Merzlyak 1997), NDVI-GREEN (Buschmann & Nagel 1993), RATIO, CHLOROPHYLL-GREEN-MODEL (CGM), CHLOROPHYLL-RED-EDGE-MODEL (CRM - Gitelson et al. 2005)
Texture indexes calculated on the NIR band (for moving window sizes of 15, 25, and 35 m each)	<p>Occurrence (Anys et al. 1994): ARITHMETIC MEAN (MEAN), STANDARD DEVIATION (SD), COEFFICIENT OF VARIATION (CV)</p> <p>Co-occurrence (Haralick et al. 1973): ANGULAR SECOND MOMENT (ASM), CONTRAST (CON), ENTROPY (ENT), INVERSE DIFFERENCE MOMENT (IDM), CORRELATION (COR), DISSIMILARITY (DIS), MAXIMUM PROBABILITY (MAXP), MEAN (MEANCO), VARIANCE (VARC), CLUSTER SHADE (CS), CLUSTER PROMINENCE (CP)</p> <p>Roughness: VECTOR DISPERSION K (ROUGH1 - Fisher 1953, Grohmann et al. 2009), AREA RATIO (ROUGH2 - Grohmann et al. 2009)</p>

(Programmed Auto), ISO 400, and matrix metering. Photographs were underexposed by -2 stops of exposure (Kato & Komiyama 2000) and stored in JPEG format (3008 x 2000 pixels resolution - Frazer et al. 2001). To the blue color planes of the 8-bit photographs, an automated global thresholding was applied to avoid variations in threshold setting by manual interpretation of photographs and to speed up the processing time (Jonckheere et al. 2004, Duveiller & Defourny 2010). The “Minimum” thresholding algorithm (Prewitt & Mendelsohn 1966) implemented in ImageJ (Schneider et al. 2012) was used. From the binarized photographs, LAI was derived with Gap Light Analyzer 2.0 (Frazer et al. 1999).

Remote sensing data

The RapidEye satellite system is optimized towards vegetation analyses and monitoring of agricultural and natural resources at relatively large cartographic scale (Tapsall et al. 2010). RapidEye’s 5 spectral bands (Tab. 1) have a native resolution of 6.5 m (resampled to 5 m). A unique feature of the sensor is the RED-EDGE band which possibly allows for better estimates of *e.g.* the chlorophyll content of vegetation (Viña & Gitelson 2005, Tapsall et al. 2010). Making use of this spectral band, specific VIs, such as the NDVI-RED-EDGE (Gitelson & Merzlyak 1997) and the CHLOROPHYLL-RED-EDGE-MODEL (Gitelson et al. 2005) have been developed. RapidEye imagery and its associated VIs have been shown useful in deriving biophysical variables (among them LAI) in the agricultural sector (Vuolo et al. 2010) and have been proven suitable for feature detection and land cover mapping in agricultural landscapes (Tapsall et al. 2010). Nevertheless, Vuolo et al. (2010) highlight that further validation work is required to test the applicability to different vegetation types and different geographical regions.

The Mengsong study site was covered by 2 cloud free RapidEye image tiles (ortho product level 3A), both acquired within a few seconds on the January 11th 2011. Images

were mosaicked and then pre-processed using the software developed by Magdon et al. (2011). Pre-processing involved an atmospheric correction based on MODIS atmosphere data by means of the Second Simulation of a Satellite Signal in the Solar Spectrum-Vector (6SV) model (Vermote et al. 1997) and a topographic correction based on a SRTM elevation model resampled to 30 m resolution.

The reflectance values of the 5 pre-processed RapidEye bands were used for the calculation of 6 VIs. For the near infrared band (NIR), texture features were calculated at 3 different spatial scales using moving windows of 15, 25, and 35 m side length. Roughness texture features were calculated in GRASS (GRASS Development Team 2012), occurrence and co-occurrence texture features and VIs were calculated using the software by Magdon et al. (2011). In total 59 image features were obtained (Tab. 1). All image features were aggregated to mean values on a spatial resolution of 20 m in order to decrease the effect of co-registration errors resulting from imperfect matching of imagery to the field sample locations (Muukkonen & Heiskanen 2005, Fuchs et al. 2009).

Selection of predictor variables

Removing predictor variables with no predictive power may improve the performance of an algorithm and the interpretability of a model as well (Svetnik et al. 2004). To conduct a selection of predictor variables, we compiled a table containing the LAI value for each subplot location and the corresponding pixel values of all 59 image features as predictor variables. For data analysis the RF algorithm (R package RANDOMFOREST - Liaw & Wiener 2002) was used. RF makes no assumptions about the distribution of input data and is able to capture non-linear relationships involving complex high order interaction effects (Strobl et al. 2007). RF is an ensemble model which uses the results of many different models, in our case regression trees, to compute a prediction. To make

regression trees uncorrelated, at each node of a tree a different subset of predictor variables is randomly selected as potential split criteria (Horning 2010). Further, every regression tree is constructed using a different bootstrap sample of about 2/3 of the observations. The remaining 1/3 of observations, the so-called out-of-bag (OOB) data, is used for an internal cross validation quantifying the accuracy of the model (Horning 2010) and to rank the predictor variables by importance. The importance of a predictor variable is expressed as the relative increase in mean square error of the prediction of OOB data caused by a random permutation of values of that variable (Cutler et al. 2007). This ranking can be used to detect meaningful variables within a large set of variables (Díaz-Uriarte & de Andrés 2005, Horning 2010). RF shows high predictive accuracy and is applicable even to highly correlated variables (Strobl et al. 2008). Since our predictor variables are exclusively derived from RapidEye’s 5 spectral bands we expected them to be correlated to some extent.

We conducted a 2 step variable selection procedure to remove variables having no predictive power and those being redundant. We used the Boruta algorithm (R package BORUTA - Kursa & Rudnicki 2010) to eliminate variables without predictive power. Boruta assesses the relevance of variables for a decision by testing whether the importance of each individual predictor variable is significantly higher than the importance of a random variable (Leutner et al. 2012). To account for the stochasticity inherent to RF, the algorithm fits RF models iteratively until all predictor variables are classified as “accepted” or “rejected” at the 0.05 alpha level. Predictor variables which are not significantly better or worse than random variables are labeled “tentative” (Leutner et al. 2012). We computed the Boruta algorithm with maxRuns=1000 and ntree=500. The final set of all relevant predictor variables may contain highly correlated, redundant variables (Kursa & Rudnicki 2010).

To remove redundant variables and identi-

fy a parsimonious model, we applied backward elimination of variables (Svetnik et al. 2004, Diaz-Uriarte & de Andrés 2005). From the set of “accepted” predictor variables as ranked by the Boruta algorithm subsequently the least important variable was removed; following a RF model was built using the remaining predictor variables. This non-recursive removal of the least important variable was repeated until only 2 predictor variables were left. For each RF model its generalization performance was evaluated by calculating the mean absolute error obtained from a 10-fold cross validation (MAE_CV). In each fold, a random selection of 10% of data points was excluded as test data, then a RF model was fit on the remaining data and applied to predict the test data. Absolute differences between predicted and observed data values were averaged per fold and then averaged over all folds. This cross validation procedure was repeated 20 times (each time using different randomly chosen test data) to acquire stable MAE_CV values. Finally, MAE_CV values resulting from these repetitions were averaged and complemented with its standard deviation. Compared to the cross validation using OOB data, which is conducted internally by the RF algorithm, such an external cross validation is regarded to result in a more objective quality assessment of the model performance (Reunanen 2003, Svetnik et al. 2004). Further, by using a non-recursive approach and embedding the variable selection into an external cross validation, bias in performance evaluation due to

over-fitting is prevented (Cawley & Talbot 2010).

After fitting all RF models we selected a model with best efficiency in terms of the number of variables and the resulting MAE_CV. Following the principle of parsimony we selected the model with the fewest number of variables showing no significant increase of MAE_CV compared to the lowest MAE_CV. Wilcoxon’s rank sum test was used to test whether differences in MAE_CV were significant at $\alpha = 0.05$ level. Finally, a RF model with 8 predictor variables was selected and applied on the respective variables for LAI mapping.

Assessing map uncertainties

Given that map users often treat uncertainty statements only in a pro-forma manner (Fassnacht et al. 2006), we place special emphasis on uncertainties associated with the resulting map. Therefore, generalization performance of the RF model used for predictive mapping was evaluated by the MAE_CV, calculated as described above. The goodness-of-fit of the RF model was quantified by the mean absolute error (MAE). Since it was found that the goodness-of-fit was not uniformly distributed over the range of predicted LAI values, MAE was also calculated for distinct sections of predicted LAI values. Finally, the resulting LAI map was complemented with a map depicting the spatial distribution of MAE values per LAI class. Furthermore, areas covered by pixel values beyond the range of the available

training data were mapped to highlight that extrapolation beyond the range of available training data is problematic and that these predictions need to be interpreted cautiously (Leutner et al. 2012).

Eventually, we derived the following error estimates to assess the uncertainties associated with the resulting map:

1. MAE_CV obtained by 10-fold CV to provide an estimate of the generalization performance of a RF model trained on the entire sample.
2. MAE as a quantification of the goodness-of-fit of the RF model to the data.
3. Exploratory analysis revealed that the model fit is better for low (<1) and high (>3) LAI values than for intermediate ones (Fig. 2). Therefore, the range of predicted LAI values was subdivided into 3 classes and for each class MAE and confidence intervals were calculated.
4. An area-weighted MAE was calculated by weighting the MAE of each class with its areal extent.
5. To highlight that model fit was not even over the whole range of predicted LAI values a spatial distribution of per class MAE values was presented as a supplementary map (Fig. 7).
6. The share of the total image area covered by reflectance values that were beyond the range of training data was stated and the corresponding area mapped.

Assessing the influence of texture features on LAI prediction

To evaluate the influence of texture features on LAI predictions, RF models were built either using only VIs, only texture features, or both jointly. Generalization performance of these models was evaluated by the MAE_CV, calculated as described above. Wilcoxon’s rank sum test was applied to test whether differences in MAE_CV were significant at $\alpha = 0.05$. In this analysis only those VIs and texture features which were classified as relevant by the Boruta analysis were considered.

Results

Response variable

Observed LAI values ranged from 0 to 6.67 with a mean LAI of 2.75. Predicted LAI values covered a smaller range from 0.1 to 4.32 with a mean LAI of 2.74. In the scatterplot depicting observed vs. predicted LAI values (Fig. 2), it is obvious that low LAI values tended to be over-predicted, while high LAI values tended to be under-predicted. Two data points having exceptional high observed LAI values were grossly under-predicted by the RF model. A tendency towards a better model fit for high and low LAI values was visible. For the intermediate section of the LAI range, for which only few observations

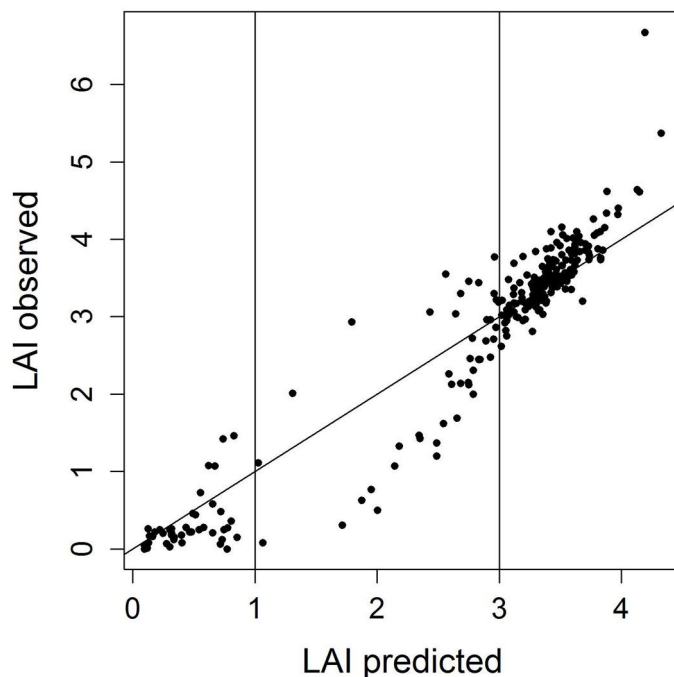


Fig. 2 - Scatterplot of predicted vs. observed LAI values. The subdivision of predicted LAI values into three classes (<1, 1-3, and >3) according to the fit of the RF model used for predicting LAI is indicated. Prediction accuracy was lowest for the intermediate LAI class (LAI 1-3).

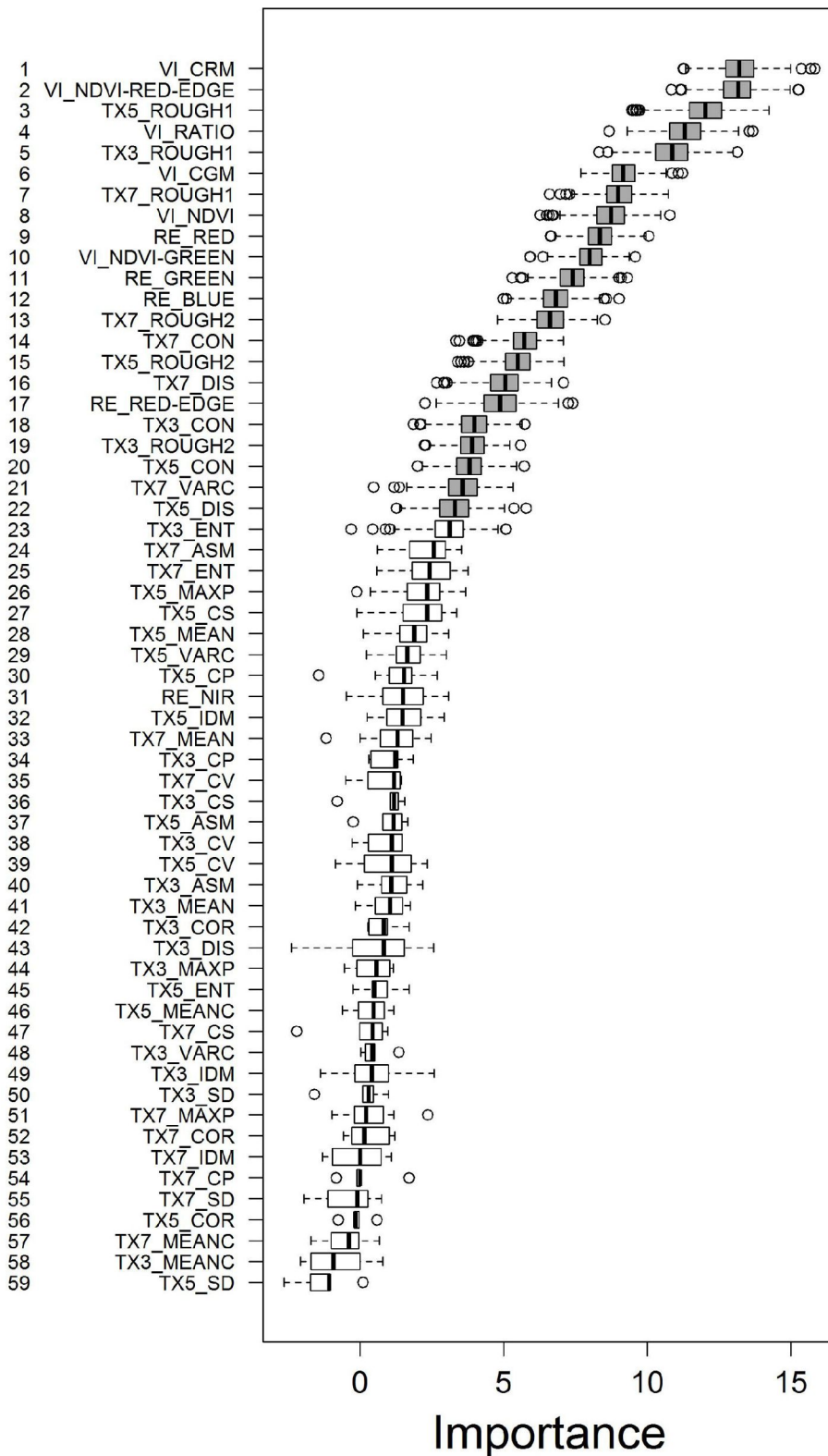


Fig. 3 - Ranking of predictor variables according to their importance assessed by the Boruta algorithm. Boruta generates random variables and tests whether correlations of these random variables with decisions are higher than correlations of real variables with decisions. Colouring: (grey): relevant variables; (white): irrelevant variables. Prefixes: (RE): RapidEye band; (VI): vegetation index; (TX): texture feature. The numbers (3, 5, and 7) following the TX-prefix refer to the moving window size of 15, 25, and 35 m side length, respectively.

were available, the accuracy of predictions was lower.

Predictor variables

Based on the Boruta analysis we identified 22 out of the 59 predictor variables as being relevant for predicting LAI. Among these, all

6 VIs were ranked at high positions (Fig. 3), with NDVI-GREEN ranked lowest at 10th position. The VIs using the RED-EDGE bands information, CHLOROPHYLL -RED-EDGE-MODEL (CRM) and NDVI- RED-EDGE, stand out as the top two predictor variables. Interestingly, the RED-EDGE and

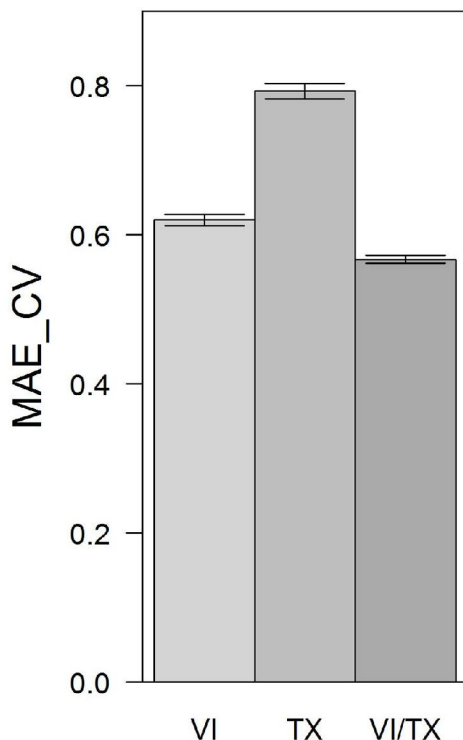
the NIR bands themselves were ranked distinctly lower at rank 17 and 31, respectively. The RapidEye bands RED, GREEN, and BLUE were listed at positions 9, 11, and 12 of the ranking.

The texture feature ROUGH1 was ranked 3rd, 5th, and 7th for moving window sizes 15,

25, and 35 m side length, respectively. ROUGH2, calculated for all 3 moving window sizes, was ranked among the relevant predictor variables at positions 13, 15, and

19. From the group of co-occurrence texture features only CON, DIS, and VARC were classified relevant. All occurrence texture features were classified irrelevant.

Fig. 4 - MAE_CV based on 20 repetitions of 10-fold cross validations for RF models using VIs only, texture features only, and VIs and texture features combined. Error bars depict standard deviations of MAE_CV values calculated over 20 repetitions.



- 1 VI_CRM
- 2 VI_NDVI_RED_EDGE
- 3 TX5_ROUGH1
- 4 VI_RATIO
- 5 TX3_ROUGH1
- 6 VI_CGM
- 7 TX7_ROUGH1
- 8 VI_NDVI
- 9 RE_RED
- 10 VI_NDVI_GREEN
- 11 RE_GREEN
- 12 RE_BLUE
- 13 TX7_ROUGH2
- 14 TX7_CON
- 15 TX5_ROUGH2
- 16 TX7_DIS
- 17 RE_RED_EDGE
- 18 TX3_CON
- 19 TX3_ROUGH2
- 20 TX5_CON
- 21 TX7_VARC
- 22 TX5_DIS

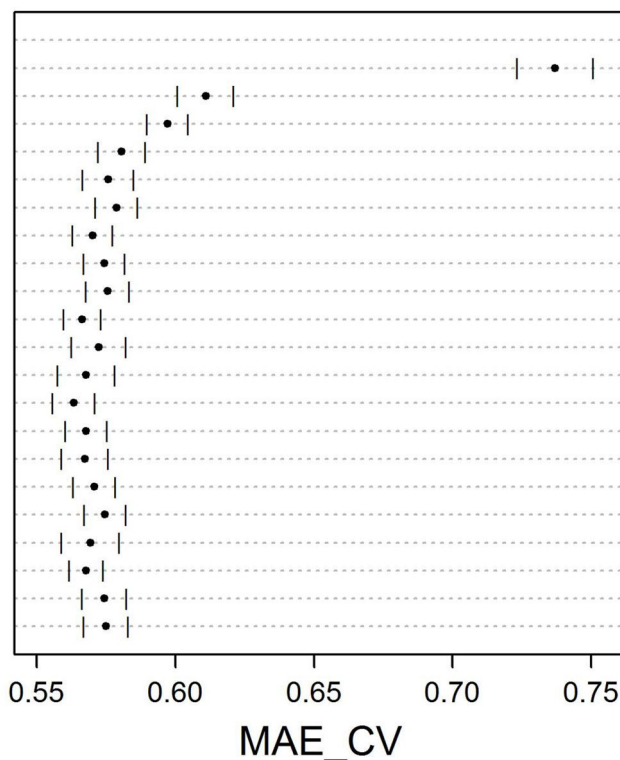


Fig. 5 - Backward elimination of predictor variables. Points represent mean MAE_CV (error bars: standard deviation) for RF models based on the respective predictor variable given on the y-axis combined with those listed above that predictor variable.

Comparing the MAE_CV values resulting for RF models based on VIs only, texture features only, and VIs and texture features combined, lowest MAE_CV was observed for the RF model based on the combination of VIs and texture features (MAE_CV = 0.57). The RF model based on VIs only was ranked second (MAE_CV = 0.62). The RF model exclusively using texture features performed worst (MAE_CV = 0.79 - Fig. 4). All observed differences were significant.

Backward selection of variables resulted in MAE_CV values for RF models as depicted in Fig. 5. The lowest MAE_CV = 0.56 was achieved by the RF model using the top 14 predictor variables. Since the MAE_CV of this model was not significantly different from the RF model using only the top 8 predictor variables (MAE_CV = 0.57, p = 0.1) we selected this more parsimonious RF model for predicting/mapping LAI. All MAE_CV values of RF models using less than 8 predictor variables were significantly different from the MAE_CV of the RF model having the lowest MAE_CV.

It is worth mentioning that the highest MAE_CV was observed for the RF model using only the top two predictor variables, the VIs CRM and NDVI-RED-EDGE. However, if the texture feature ROUGH1 is included into the RF model, MAE_CV drops considerably (p < 0.0001).

LAI map - spatial prediction of LAI

We produced a wall-to-wall LAI map of the study area by applying the selected RF model to the corresponding combination of image features (Fig. 6). MAE_CV of the LAI map was 0.57 and the MAE was 0.29. The area-weighted MAE (0.35) was slightly higher than the unweighted MAE (Tab. 2). This difference occurred because: (1) the MAE of the intermediate LAI class was noticeably higher than the MAE of the other classes; and (2) the intermediate class covered roughly one third (29% - Fig. 7) of the mapping area, thus, its MAE was weighted by a factor of the same magnitude as those of the other two classes (Tab. 2). The larger differences between observed and predicted LAI values occurring in the intermediate LAI class did not influence the unweighted MAE too much since the total number of observations in this class was lower compared to those of the other two classes (Tab. 2).

Exploratory analysis of image areas associated with high MAE values revealed that high MAE values cannot be directly related to specific vegetation types. Furthermore, the variability of predictor variables for the corresponding LAI class was not higher than for the other classes (Tab. 3).

Overall, 7.46% of the study site were covered by pixel values beyond the range of our training data (Fig. 7). Visual inspection revealed that primarily water bodies, settle-

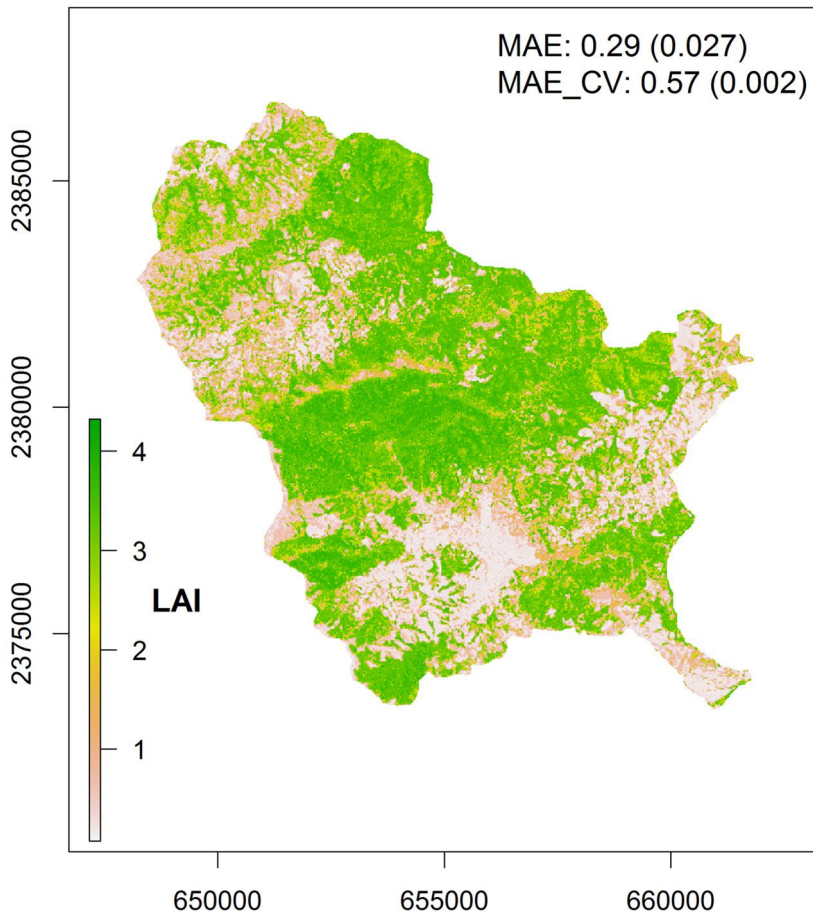


Fig. 6 - LAI map for the Mensgong study site (January 11th 2011). Map uncertainty was approximated by the MAE and the MAE_CV (complemented with their estimated standard errors - in parenthesis).

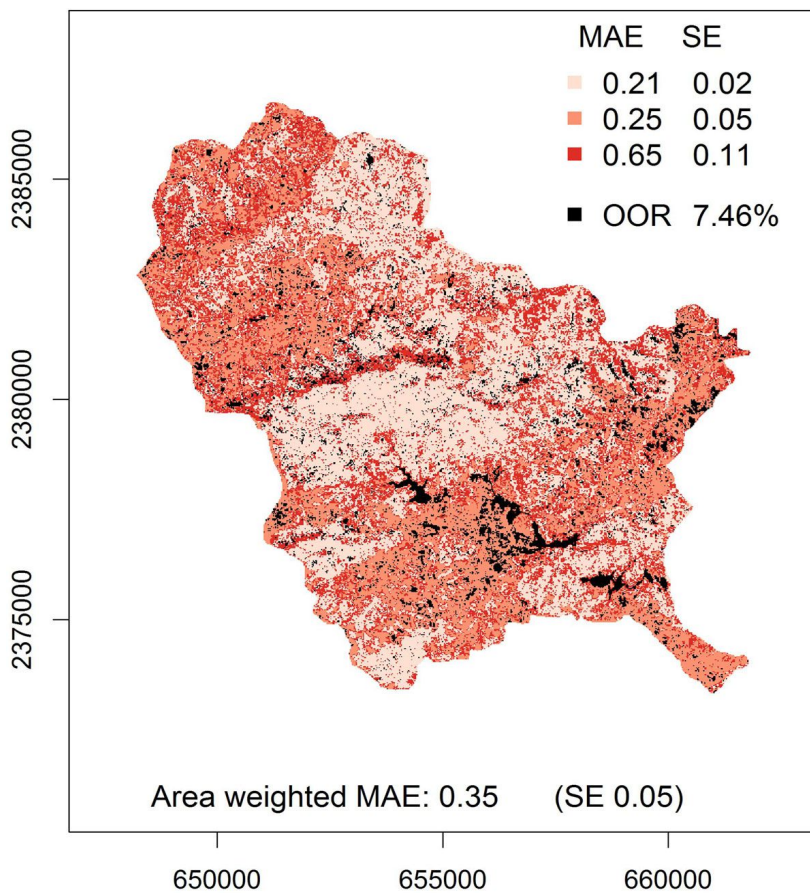


Fig. 7 - Map of model fit, quantified by the MAE, as a proxy for prediction uncertainty associated with LAI map (Fig. 6). Overall, 7.46 % of the mapping area were covered by pixel values that were out of range (OOR) of training data. (SE): standard error.

Tab. 2 - Mean Absolute Error (MAE) calculated for 3 sections of predicted LAI values.

LAI section	No. of observations	MAE				Area (pixel)	Weight	Area-weighted MAE	
		Value	%	SD	SE			Value	SE
<1	45	0.25	57.81	0.3	0.05	63913	0.27	0.07	0.01
1-3	43	0.65	25.93	0.71	0.11	68113	0.29	0.19	0.03
>3	164	0.21	6.07	0.3	0.02	102546	0.44	0.09	0.01

Tab. 3 - Coefficient of variation of predictor variables' pixel values per LAI class standardized by the respective number of observations per LAI class.

LAI (observed)	VI_CRM	VI_NDVI-RED-EDGE	TX5_ROUGHNES1	VI_RATIO	TX3_ROUGHNES1	VI_CGM	TX7_ROUGHNES1	VI_NDVI
<1	0.77	0.52	1.82	0.34	1.73	0.7	1.86	0.47
1-3	0.53	0.33	2.29	0.26	2.21	0.55	2.13	0.22
>3	0.09	0.05	0.34	0.03	0.37	0.13	0.33	0.03

ments/roads, bare agricultural land, and mining claims were represented by these pixel values. Nevertheless, out-of-range pixel values were also found within vegetated areas, yet no vegetation type could be identified as being particularly affected.

Discussion

We produced a high resolution map depicting the spatial variability of LAI by combining RapidEye data and LAI estimates obtained from field sampling. Such map has the potential to be used in spatially distributed modeling of vegetation productivity, evapotranspiration, and surface energy balance (Turner et al. 1999, Tang et al. 2012), and it can support a better understanding of ecosystem processes over large and remote areas (Chapin III et al. 2011). This is particularly important for regions like Xishuangbanna that harbor a high biodiversity (Zhang & Cao 1995, Li et al. 2007) which is threatened by the dramatic expansion of rubber plantations (Xu et al. 2013, Li et al. 2007). The results from this study provide information to make inferences about ecosystem dynamics of our study area. The presented approach to map LAI using the RF algorithm has the potential to be transferred to other geographical regions harboring different vegetation types. Transferred to another region, the resulting LAI maps presumably carry an error which differs from that reported by our study. Therefore, the user should make sure that map uncertainties are described adequately.

The LAI values observed in this research were slightly lower but generally in accordance with other studies assessing LAI in tropical/ subtropical ecosystems. In a review paper Asner et al. (2003) reported a mean LAI of 4.9 for tropical evergreen broadleaf forests. Roberts et al. (2004) obtained LAI values ranging from 4.1 to 8.0 for tropical lowland rainforests, with a tendency for higher values in Asia. Along a gradient co-

vering open pasture, secondary forests, re-generation forests after selective logging, and old-growth forests Tang et al. (2012) mapped LAI using waveform LIDAR at La Selva, Costa Rica and observed mean values of 1.74, 5.20, 5.41, and 5.62 LAI, respectively. Nevertheless, there are few studies for direct comparisons since different definitions of LAI are frequently used (Barclay 1998) and different methods for LAI determination are applied in the field (Bréda 2003, Asner et al. 2003). Furthermore, a significant fraction of literature on LAI does not describe the methodology used in sufficient detail, thus, comparability is hampered (Asner et al. 2003, Beckschäfer et al. 2013). Differences between studies may also arise from seasonal changes in LAI due to changes in rainfall volume and other climatic parameters (Bréda 2003). In our study, field data and remote sensing data were acquired within the dry season, which might have caused slightly lower LAI values compared to those found in the literature.

Except for two outliers showing observed LAI values distinctly higher than predicted, a good agreement between observed and predicted LAI values was found. MAE was highest for the intermediate LAI class containing lower numbers of observations. Higher MAE values might be explained by the greater spatial variability of LAI values observed in this class resulting from heterogeneous shrub vegetation and trees scattered within grasslands. In such a heterogeneous landscape slight location errors in field and satellite data might result in perceptible prediction errors. Using a pixel size of 20x20 m we tried to account for these errors but in some cases they may still occur. Moreover, it needs to be taken into account that LAI as derived from hemispherical photographs is a modeled value in itself, which might already carry an unknown error. Further, a tendency towards an under-prediction of high LAI values and an over-prediction of low LAI va-

lues has been observed. This is because the response from a RF model, in the case of regression, is a value resulting from averaging the predictions made by all trees within a RF (Horning 2010).

VIs derived from RapidEye data, especially those making use of the RED-EDGE band, appeared to be important for predicting LAI. The importance of the RED-EDGE information results from the sensitivity of the respective electromagnetic spectrum (680-740 nm) to vegetation chlorophyll content that shows an abrupt rise in reflectance caused by vegetation. This is related to strong chlorophyll absorption and high internal leaf scattering of plant tissue (Schuster et al. 2012). By choosing the RED-EDGE band, instead of the RED band for the NDVI calculation, a lower saturation over highly vegetated area is achieved (Tapsall et al. 2010). Surprisingly, the RED-EDGE band itself was only ranked 17th among the relevant predictor variables. The higher ranking of VIs might be explained by their general ability to reduce the impacts of confounding factors such as soil reflectance and atmospheric effects on reflectance values (Baret & Guyot 1991, Lu 2005).

Besides VIs, the texture features ROUGH1 and ROUGH2 were ranked high among the relevant predictor variables. MAE_CV of RF models including VIs and texture features jointly was significantly lower than MAE_CV of RF models exclusively based on VIs. This shows the potential of texture features derived from RapidEye data to improve the quality of LAI maps and to reduce the associated uncertainties. Similar effects of the inclusion of texture features have been reported for IKONOS satellite data (Colombo et al. 2003) and airborne CASI imagery (Wulder et al. 1998). Nevertheless, most texture features were classified irrelevant by the Boruta analysis in our study, pointing to the difficulty to identify which specific textural characteristic is represented by each of the

texture features (Wulder et al. 1998). Texture features vary with the characteristics of the landscape under investigation and image types used (Lu 2005), and besides the identification of appropriate texture features, suitable moving window sizes and image bands need to be determined (Chen et al. 2004). Clear guidelines on how to select appropriate texture features are still lacking; hence, the generation of suitable texture features is a challenging task (Lu 2005). In our study, texture features were calculated for the NIR band using three moving window sizes of 15, 25, and 35 m side length. Whether a different set of texture features would be among the relevant predictor variables if calculated for larger moving window sizes or different spectral bands, was not investigated here but should be considered in future mapping efforts. For LAI mapping it might be of particular interest whether texture features calculated for the two VIs identified as being most relevant for predictions would further increase map accuracy.

Using only a reduced set of predictor variables did not substantially enhance the predictive performance of the RF models. This confirms previous studies stating that RF is generally able to deal with large amounts of non-informative or redundant variables (Diaz-Uriarte & de Andrés 2005, Leutner et al. 2012). Nevertheless, reducing the set of predictor variables might reduce computational cost and increase the interpretability of the predictions made by the model (Svetnik et al. 2004). In our opinion, reducing model complexity and providing map users with understandable descriptions of the methods used to create a map should be an integral part of predictive mapping. Through this, informed assessments of appropriate and inappropriate uses of maps can be made by the user (Fassnacht et al. 2006).

Map interpretation and inference are directly affected by map accuracy. Therefore, users should not rely on maps without associated estimates of error (Card 1982). Addressing this issue, we provided estimates of data fit and generalization performance of the RF model. Following the recommendation of Mitchard et al. (2011), we additionally complemented the produced LAI map with an estimated spatial distribution of accuracy. From this map an area weighted MAE was calculated. By providing an area weighted MAE and the corresponding MAE map, the user is able to evaluate the usefulness of the map at hand. Mapping of areas having pixel values which were beyond the range of reflectance values available in the training data appeared to be valuable information. Exploratory analysis revealed that such areas were mainly declared for landscape elements such as water bodies, settlements, or mining areas for which LAI predictions would be questionable.

Our results demonstrate the suitability of RapidEye data to retrieve LAI across a range of landscape classes including forests. Thus, the applicability of RapidEye imagery to deriving LAI for agricultural areas (Tapsall et al. 2010, Vuolo et al. 2010) can be broadened to include forest ecosystems. This is especially valuable, as forest LAI is regarded as one of the most important structural variables for understanding ecosystem processes (Bonan 1993).

Acknowledgments

Our thanks are due to the Advisory Group on International Agricultural Research (BEAF) at the German Agency for International Cooperation (GIZ) and the German Ministry for Economic Cooperation (BMZ) for funding the research project MMC (Making the Mekong Connected, Project No. 08.7860.3-001.00) within which this study had been carried out. We are also grateful to all members of the MMC-project for their support. We acknowledge the German National Space Agency DLR (Deutsches Zentrum für Luft- und Raumfahrt e.V.) for the delivery of RapidEye images as part of the RapidEye Science Archive (Proposal No. 390). The provision of RapidEye images was funded by the German Federal Ministry of Economics and Technology. The responsibility for any use of the results lies with the user. We also thank Mr. Paul Magdon for his support in pre-processing RapidEye images.

References

- Anys H, Bannari A, He D, Morin D (1994). Texture analysis for the mapping of urban areas using airborne MEIS-II images. In: Proceedings of the "First International Airborne Remote Sensing Conference and Exhibition". Strasbourg (France), vol. 3, pp. 231-245.
- Asner GP, Scurlock JMO, Hicke AJ (2003). Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecology and Biogeography* 12 (3): 191-205. - doi: [10.1046/j.1466-822X.2003.00026.x](https://doi.org/10.1046/j.1466-822X.2003.00026.x)
- Barclay HJ (1998). Conversion of total leaf area to projected leaf area in lodgepole pine and douglas-fir. *Tree Physiology* 18 (3): 185-193. - doi: [10.1093/treephys/18.3.185](https://doi.org/10.1093/treephys/18.3.185)
- Baret F, Buis S (2008). Estimating canopy characteristics from remote sensing observations: review of methods and associated problems. In: "Advances in Land Remote Sensing: System, Modelling, Inversion and Application" (Liang S ed). Springer, New York, USA, pp. 173-201.
- Baret F, Guyot G (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment* 35: 161-173. - doi: [10.1016/0034-4257\(91\)90009-U](https://doi.org/10.1016/0034-4257(91)90009-U)
- Baudet M, Messier C (2002). Variation in canopy openness and light transmission following selection cutting in northern hardwood stands: an assessment based on hemispherical photographs. *Agricultural and Forest Meteorology* 110 (3): 217-228. - doi: [10.1016/S0168-1923\(01\)00289-1](https://doi.org/10.1016/S0168-1923(01)00289-1)
- Beckschäfer P, Seidel D, Kleinn C, Xu J (2013). On the exposure of hemispherical photographs in forests. *iForest* 6: 228-237. - doi: [10.3832/ifor0957-006](https://doi.org/10.3832/ifor0957-006)
- Bonan G (1993). Importance of leaf area index and forest type when estimating photosynthesis in boreal forests. *Remote Sensing of Environment* 43 (3): 303-314. - doi: [10.1016/0034-4257\(93\)90072-6](https://doi.org/10.1016/0034-4257(93)90072-6)
- Bréda NJJ (2003). Ground based measurements of leaf area index: a review of methods, instruments and current controversies. *Journal of Experimental Botany* 54 (392): 2403-2417. - doi: [10.1093/jxb/erg263](https://doi.org/10.1093/jxb/erg263)
- Breiman L (2001). Random forests. *Machine Learning* 45: 5-32. - doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
- Buschmann C, Nagel E (1993). *In vivo* spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *International Journal of Remote Sensing* 14 (4): 711-722. - doi: [10.1080/01431169308904370](https://doi.org/10.1080/01431169308904370)
- Card D (1982). Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing* 48: 431-439.
- Cawley GC, Talbot NLC (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research* 11: 2079-2107.
- Chapin III FS, Matson PA, Vitousek P (2011). Principles of terrestrial ecosystem ecology (2nd edn). Springer, New York USA, pp. 529. - doi: [10.1007/978-1-4419-9504-9](https://doi.org/10.1007/978-1-4419-9504-9)
- Chen D, Stow DA, Gong P (2004). Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case. *International Journal of Remote Sensing* 25 (11): 2177-2192. - doi: [10.1080/01431160310001618464](https://doi.org/10.1080/01431160310001618464)
- Cohen WB, Maiersperger TK, Gower ST, Turner DP (2003). An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment* 84(4): 561-571. - doi: [10.1016/S0034-4257\(02\)00173-6](https://doi.org/10.1016/S0034-4257(02)00173-6)
- Colombo R, Bellingeri D, Fasolini D, Marino CM (2003). Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sensing of Environment* 86 (1): 120-131. - doi: [10.1016/S0034-4257\(03\)00094-4](https://doi.org/10.1016/S0034-4257(03)00094-4)
- Cutler D, Edwards Jr T, Beard K, Cutler A, Hess K, Gibson J, Lawler J (2007). Random forests for classification in ecology. *Ecology* 88 (11): 2783-2792. - doi: [10.1890/07-0539.1](https://doi.org/10.1890/07-0539.1)
- Diaz-Uriarte R, de Andrés S (2005). Variable selection from random forests: application to gene expression data. Technical report. [online] URL: <http://ligarto.org/rdiaz/Papers/rfVS/random-ForestVarSel.html>
- Duveiller G, Defourny P (2010). Batch processing of hemispherical photography using object-based

- image analysis to derive canopy biophysical variables. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII-4/C7. [online] URL: <http://www.isprs.org/proceedings/XXXVIII/4-C7/pdf/Duveiller.pdf>
- Fassnacht KS, Cohen WB, Spies TA (2006). Key issues in making and using satellite-based maps in ecology: a primer. *Forest Ecology and Management* 222: 167-181. - doi: [10.1016/j.foreco.2005.09.026](https://doi.org/10.1016/j.foreco.2005.09.026)
- Fisher R (1953). Dispersion on a sphere. *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* 217 (1130): 295-305. - doi: [10.1098/rspa.1953.0064](https://doi.org/10.1098/rspa.1953.0064)
- Frazier G, Canham C, Lertzman K (1999). Gap Light Analyzer (GLA), version 2.0: imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs, users manual and program documentation. Simon Fraser University, Burnaby, British Columbia, and the Institute of Ecosystem Studies, Millbrook, New York, USA.
- Frazier G, Fournier R, Trofymow J, Hall R (2001). A comparison of digital and film fisheye photography for analysis of forest canopy structure and gap light transmission. *Agricultural and Forest Meteorology* 109 (4): 249-263. - doi: [10.1016/S0168-1923\(01\)00274-X](https://doi.org/10.1016/S0168-1923(01)00274-X)
- Fuchs H, Magdon P, Kleinn C, Flessa H (2009). Estimating aboveground carbon in a catchment of the siberian forest tundra: combining satellite imagery and field inventory. *Remote Sensing of Environment* 113 (3): 518-531. - doi: [10.1016/j.rse.2008.07.017](https://doi.org/10.1016/j.rse.2008.07.017)
- Gamon J, Field C, Goulden M, Griffin K, Hartley A, Joel G, Peñuelas J, Valentini R (1995). Relationships between ndvi, canopy structure, and photosynthesis in three californian vegetation types. *Ecological Applications* 5 (1): 28-41. - doi: [10.2307/1942049](https://doi.org/10.2307/1942049)
- Gitelson A, Viña A, Ciganda V, Rundquist D, Arkebauer T (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters* 32 (8): L08403.
- Gitelson AA, Merzlyak MN (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing* 18 (12): 2691-2697. - doi: [10.1080/014311697217558](https://doi.org/10.1080/014311697217558)
- Glenn EP, Huete AR, Nagler PL, Nelson SG (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: what vegetation indices can and cannot tell us about the landscape. *Sensors* 8 (4): 2136-2160. - doi: [10.3390/s8042136](https://doi.org/10.3390/s8042136)
- Gonsamo A, Pellikka P (2012). The sensitivity based estimation of leaf area index from spectral vegetation indices. *ISPRS Journal of Photogrammetry and Remote Sensing* 70: 15-25. - doi: [10.1016/j.isprsjprs.2012.03.009](https://doi.org/10.1016/j.isprsjprs.2012.03.009)
- Gower S, Norman J (1991). Rapid estimation of leaf area index in conifer and broad-leaf plantations. *Ecology* 72 (5): 1896-1900. - doi: [10.2307/1940988](https://doi.org/10.2307/1940988)
- GRASS Development Team (2012). Geographic Resources Analysis Support System (GRASS GIS) Software. Open Source Geospatial Foundation, USA. [online] URL: <http://grass.osgeo.org>
- Grohmann C, Smith M, Riccomini C (2009). Surface roughness of topography: a multi-scale analysis of landform elements in midland valley, Scotland. In: *Proceedings of the conference "Geomorphometry 2009"* (Purves R, Gruber S, Straumann R, Hengl T eds). University of Zurich (CH), pp. 140-148.
- Haralick RM, Shanmugam K, Dinstein I (1973). Textural features for image classification. *Systems, Man and Cybernetics. IEEE Transactions on SMC* 3 (6): 610-621.
- Horning N (2010). Random forests: an algorithm for image classification and generation of continuous fields data sets. In: *Proceeding of the "International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences"*. Hanoi (Vietnam) 9-11 Dec 2010. JVGC - Japan-Vietnam Geoinformatics Consortium, Institute for Environment and Resources, Vietnam.
- Jonckheere I, Fleck S, Nackaerts K, Muys B, Coppin P, Weiss M, Baret F (2004). Review of methods for *in situ* leaf area index determination: Part I. Theories, sensors and hemispherical photography. *Agricultural and Forest Meteorology* 121 (1): 19-35. - doi: [10.1016/j.agrformet.2003.08.027](https://doi.org/10.1016/j.agrformet.2003.08.027)
- Kato S, Komiyama A (2000). A calibration method for adjusting hemispherical photographs to appropriate black-and-white images. *Journal of Forest Research* 5 (2): 109-111. - doi: [10.1007/BF02762529](https://doi.org/10.1007/BF02762529)
- Kursa M, Rudnicki W (2010). Feature selection with the boruta package. *Journal of Statistical Software* 36 (11): 1-13.
- Latifi H, Galos B (2010). Remote sensing-supported vegetation parameters for regional climate models: a brief review. *iForest* 3: 98-101 - doi: [10.3832/ifor0543-003](https://doi.org/10.3832/ifor0543-003)
- Lee K, Park Y, Kim S, Park J, Woo C, Jang K (2004). Remote sensing estimation of forest lai in close canopy situation. In: *Proceedings of the 20th ISPRS Congress (Altan O ed)*. Istanbul (Turkey) 12-23 July 2004. ISPRS Archives, vol. 35, Part B7. [online] URL: <http://www.isprs.org/proceedings/XXXV/congress/comm7/papers/77.pdf>
- Leutner BF, Reineking B, Müller, J, Bachmann, M, Beierkuhnlein, C, Dech, S, Wegmann, M (2012). Modelling forest alpha-diversity and floristic composition on the added value of lidar plus hyperspectral remote sensing. *Remote Sensing* 4 (9): 2818-2845. - doi: [10.3390/rs4092818](https://doi.org/10.3390/rs4092818)
- Li H, Aide T, Ma Y, Liu W, Cao M (2007). Demand for rubber is causing the loss of high diversity rain forest in sw China. *Biodiversity and Conservation* 16: 1731-1745. - doi: [10.1007/s10531-006-9052-7](https://doi.org/10.1007/s10531-006-9052-7)
- Li H, Ma Y, Liu W, Liu W (2009). Clearance and fragmentation of tropical rain forest in Xishuangbanna, SW, China. *Biodiversity and conservation* 18 (13): 3421-3440. - doi: [10.1007/s10531-009-9651-1](https://doi.org/10.1007/s10531-009-9651-1)
- Liaw A, Wiener M (2002). The randomforest package. *R News* 2 (3): 18-22.
- Lu D (2005). Aboveground biomass estimation using landsat tm data in the Brazilian amazon. *International Journal of Remote Sensing* 26 (12): 2509-2525. - doi: [10.1080/01431160500142145](https://doi.org/10.1080/01431160500142145)
- Magdon P, Fuchs H, Fischer C, Kleinn C (2011). Forest cover monitoring using RapidEye: a case study in costa rica. In: "RapidEye Science Archive", (RESA)-Erste Ergebnisse, GITO mbH Verlag, pp. 47-56.
- Mitchard E, Saatchi S, Lewis S, Feldpausch T, Gerard F, Woodhouse I, Meir P (2011). Comment on "A first map of tropical africa's above-ground biomass derived from satellite imagery". *Environmental Research Letters* 6 (4): 049001. - doi: [10.1088/1748-9326/6/4/049001](https://doi.org/10.1088/1748-9326/6/4/049001)
- Morisette J, Baret F, Privette J, Myneni R, Nickerson J, Garrigues S, Shabanov N, Weiss M, Fernandes R, Leblanc S, Kalacska M, Sanchez-Azofeifa G, Chubey M, Rivard B, Stenberg P, Rautiainen M, Voipio P, Manninen T, Pilant A, Lewis T, Iames J, Colombo R, Meroni M, Busetto L, Cohen W, Turner D, Warner E, Petersen G, Seufert G, Cook R (2006). Validation of global moderate-resolution lai products: a framework proposed within the ceos land product validation subgroup. *Geoscience and Remote Sensing, IEEE Transactions* 44 (7): 1804-1817. - doi: [10.1109/TGRS.2006.872529](https://doi.org/10.1109/TGRS.2006.872529)
- Muukkonen P, Heiskanen J (2005). Estimating biomass for boreal forests using aster satellite data combined with standwise forest inventory data. *Remote Sensing of Environment* 99 (4): 434-447. - doi: [10.1016/j.rse.2005.09.011](https://doi.org/10.1016/j.rse.2005.09.011)
- Pierce L, Running S (1988). Rapid estimation of coniferous forest leaf area index using a portable integrating radiometer. *Ecology* 69 (6): 1762-1767. - doi: [10.2307/1941154](https://doi.org/10.2307/1941154)
- Prasad A, Iverson L, Liaw A (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9: 181-199. - doi: [10.1007/s10021-005-0054-1](https://doi.org/10.1007/s10021-005-0054-1)
- Prewitt JMS, Mendelsohn ML (1966). The analysis of cell images. *Annals of the New York Academy of Sciences* 128 (3): 1035-1053. - doi: [10.1111/j.1749-6632.1965.tb11715.x](https://doi.org/10.1111/j.1749-6632.1965.tb11715.x)
- Reunanen J (2003). Overfitting in making comparisons between variable selection methods. *Journal of Machine Learning Research* 3: 1371-1382.
- Rich P (1989). A manual for analysis of hemispherical canopy photography. Technical report, Los Alamos National Lab, New Mexico, USA.
- Roberts J, Gash J, Tani M (2004). Controls on evaporation in lowland tropical rainforest. In: "Forests, Water and People in the Humid Tropics: Past, Present and Future Hydrological Research for Integrated Land and Water Management" (Bonell M, Bruijnzeel LA eds), Cambridge University Press, Cambridge, UK, pp. 287.
- Schneider C, Rasband W, Eliceiri K (2012). NIH image to ImageJ: 25 years of image analysis.

- Nature Methods 9 (7): 671-675. - doi: [10.1038/nmeth.2089](https://doi.org/10.1038/nmeth.2089)
- Schuster C, Förster M, Kleinschmit B (2012). Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *International Journal of Remote Sensing* 33 (17): 5583-5599. - doi: [10.1080/01431161.2012.666812](https://doi.org/10.1080/01431161.2012.666812)
- Stehman SV, Czaplewski RL (1998). Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sensing of Environment* 64 (3): 331-344. - doi: [10.1016/S0034-4257\(98\)00010-8](https://doi.org/10.1016/S0034-4257(98)00010-8)
- Strobl C, Boulesteix A, Zeileis A, Hothorn T (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC bioinformatics* 8 (1): 25. - doi: [10.1186/1471-2105-8-25](https://doi.org/10.1186/1471-2105-8-25)
- Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A (2008). Conditional variable importance for random forests. *BMC Bioinformatics* 9 (1): 307. - doi: [10.1186/1471-2105-9-307](https://doi.org/10.1186/1471-2105-9-307)
- Svetnik V, Liaw A, Tong C, Wang T (2004). Application of Breiman's random forest to modeling structure-activity relationships of pharmaceutical molecules. In: "Multiple classifier systems" (Roli F, Kittler J, Windeatt T eds). *Lecture Notes in Computer Science*, vol. 3077, Springer, Berlin, Heidelberg, pp. 334-343.
- Tagle MAG, Jimenez Perez J, Himmelsbach W (2011). Impact of firewood extraction on leaf area index and canopy openness in mixed pine-oak forests in northeast Mexico. *Forstarchiv* 82 (1): 20-25.
- Tang H, Dubayah R, Swatantran A, Hofton M, Sheldon S, Clark DB, Blair B (2012). Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica. *Remote Sensing of Environment* 124: 242-250. - doi: [10.1016/j.rse.2012.05.005](https://doi.org/10.1016/j.rse.2012.05.005)
- Tapsall B, Milenov P, Tasdemir K (2010). Analysis of RapidEye imagery for annual landcover mapping as an aid to European Union (EU) common agricultural policy. In: "ISPRS TC VII Symposium - 100 Years ISPRS" (Wagner W, Székely B eds). *IAPRS*, Vienna, Austria, pp. 568-573.
- Turner D, Cohen W, Kennedy R, Fassnacht K, Briggs J (1999). Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment* 70 (1): 52-68. - doi: [10.1016/S0034-4257\(99\)00057-7](https://doi.org/10.1016/S0034-4257(99)00057-7)
- Vermote E, Tanré D, Deuze J, Herman M, Morcrette J (1997). Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *Geoscience and Remote Sensing, IEEE Transactions* 35 (3): 675-686. - doi: [10.1109/36.581987](https://doi.org/10.1109/36.581987)
- Viña A, Gitelson AA (2005). New developments in the remote estimation of the fraction of absorbed photosynthetically active radiation in crops. *Geophysical Research Letters* 32 (17): L17403.
- Vuolo F, Atzberger C, Richter K, D'Urso G, Dash J (2010). Retrieval of biophysical vegetation products from RapidEye imagery. In: "ISPRS TC VII Symposium - 100 Years ISPRS" (Wagner W, Székely B eds). *IAPRS*, Vienna, Austria, vol. 38, pp. 281-286.
- Weiss M, Baret F, Smith G, Jonckheere I, Coppin P (2004). Review of methods for *in situ* leaf area index (LAI) determination: Part II. Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology* 121 (1): 37-53. - doi: [10.1016/j.agrformet.2003.08.001](https://doi.org/10.1016/j.agrformet.2003.08.001)
- Wulder MA, LeDrew EF, Franklin SE, Lavigne MB (1998). Aerial image texture information in the estimation of northern deciduous and mixed wood forest leaf area index (LAI). *Remote Sensing of Environment* 64 (1): 64-76. - doi: [10.1016/S0034-4257\(97\)00169-7](https://doi.org/10.1016/S0034-4257(97)00169-7)
- Xu J, Fox J, Vogler J, Yongshou Z, Lixin Y, Jie Q, Leisz S (2005). Land-use and land-cover change and farmer vulnerability in Xishuangbanna prefecture in southwestern China. *Environmental Management* 36 (3): 404-413. - doi: [10.1007/s00267-003-0289-6](https://doi.org/10.1007/s00267-003-0289-6)
- Xu J, Grumbine RE, Beckschäfer P (2013). Landscape transformation through the use of ecological and socioeconomic indicators in Xishuangbanna, Southwest China, Mekong region. *Ecological Indicators*. - doi: [10.1016/j.ecolind.2012.08.023](https://doi.org/10.1016/j.ecolind.2012.08.023)
- Zhang J, Cao M (1995). Tropical forest vegetation of Xishuangbanna, SW China and its secondary changes, with special reference to some problems in local nature conservation. *Biological Conservation* 73 (3): 229-238. - doi: [10.1016/0006-3207\(94\)00118-A](https://doi.org/10.1016/0006-3207(94)00118-A)
- Zhao K, Popescu S (2009). Lidar-based mapping of leaf area index and its use for validating global carbon satellite LAI product in a temperate forest of the southern USA. *Remote Sensing of Environment* 113 (8): 1628-1645. - doi: [10.1016/j.rse.2009.03.006](https://doi.org/10.1016/j.rse.2009.03.006)