

A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping

Patrick Griffiths, Sebastian van der Linden, Tobias Kuemmerle, and Patrick Hostert

What is the motivation for this article?

Which key parameters were used for the parametric scoring?

What is the difference between annual and seasonal consistency?

Abstract—Information on the changing land surface is required at high spatial resolutions as many processes cannot be resolved using coarse resolution data. Deriving such information over large areas for Landsat data, however, still faces numerous challenges. Image compositing offers great potential to circumvent such shortcomings. We here present a compositing algorithm that facilitates creating cloud free, seasonally and radiometrically consistent datasets from the Landsat archive. A parametric weighting scheme allows for flexibly utilizing different pixel characteristics for optimized compositing. We describe in detail the development of three parameter decision functions: acquisition year, day of year and distance to clouds. Our test site covers 42 Landsat footprints in Eastern Europe and we produced three annual composites. We evaluated seasonal and annual consistency and compared our composites to BRDF normalized MODIS reflectance products. Finally, we also evaluated how well the composites work for land cover mapping. Results prove that our algorithm allows for creating seasonally consistent large area composites. Radiometric correspondence to MODIS was high (up to $R^2 > 0.8$), but varied with land cover configuration and selected image acquisition dates. Land cover mapping yielded promising results (overall accuracy 72%). Class delineations were regionally consistent with minimal effort for training data. Class specific accuracies increased considerably ($\sim 10\%$) when spectral metrics were incorporated. Our study highlights the value of compositing in general and for Landsat data in particular, allowing for regional to global LULCC mapping at high spatial resolutions.

Index Terms—Compositing, Landsat, large area mapping, random forests, remote sensing.

I. INTRODUCTION

REMOTE sensing based information on land use and land cover change (LULCC) is required at spatial resolutions higher than those currently available from existing global land cover products [15], [40]. This is because many LULCC processes, such as logging, deforestation, land abandonment or urban sprawl, represent critical drivers of global environmental change, but occur at spatial scales that cannot be resolved with coarse resolution data in many areas of the world. This is specifically true for many regions in Africa or South-East Asia,

Manuscript received August 11, 2012; revised October 12, 2012; accepted October 23, 2012. Date of publication January 18, 2013; date of current version September 20, 2013. This work was funded by the Belgian Science Policy, Research Program for Earth Observation Stereo II, contract SR/00/133, as part of the FOMO project (Remote sensing of the forest transition and its ecosystem impacts in mountain environments).

The authors are with the Geography Department, Humboldt-Universität zu Berlin, 10099 Berlin, Germany (corresponding author e-mail: patrick.griffiths@geo.hu-berlin.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2012.2228167

where fine scale processes and patterns prevail. The global Landsat archive allows for reconstructing LULCC back to the 1970s and therefore has become an integral component of global LULCC research [10], [34].

Despite large amounts of Earth Observation (EO) data available at spatial resolutions of 20 to 50 m (in the following referred to as high spatial resolution), land cover and LULCC products with such spatial detail are commonly not available across large areas [15]. The reason for this is that mapping and monitoring of land cover across large regions at high spatial resolutions still poses unique challenges [39]. These challenges partly relate to the spatial detail of Landsat, which comes at the cost of a relatively small swath. A great number of Landsat footprints are therefore often needed for a full regional coverage. Additionally, phenologically and radiometrically consistent datasets are required when analyzing LULCC and complex changes in vegetation cover in particular [36]. This is not easily feasible, considering that Landsat has a repeat frequency of 16-days and thus might only provide few or not even a single unclouded scene per growing season in many areas [28]. Data availability is further aggravated by discontinuities in image archives, as well as data or sensor related errors (e.g. the scan line correction (SLC) failure of Landsat 7 after May 2003 Arvidson *et al.* [2]). Thus, conceptually more advanced approaches have to be developed to allow for robust mapping and monitoring at high spatial resolution over large areas.

In many studies where land cover has been mapped at high resolution over large areas, unsupervised and/or supervised classification methods are used, requiring considerable user interaction and thus limiting the potential for automation [21], [38]. Land cover products at Landsat resolution also are available for Europe (CORINE land cover, [20]). However, CORINE data are derived through interpretation and digitization of Landsat data, have very high production costs and are limited to a minimum mapping unit of 25 ha [13]. A number of operational land cover monitoring approaches using Landsat data over regional to continental scales have been developed in Australia prior to the opening of the USGS Landsat archive [6], [16], [17]. Some studies have investigated the potential of transferring classification models across space or generating training data for an unclassified image from the overlap with an existing classification [30], [37], [53]. More recently, different groups have investigated the potential of exploiting annual time series of Landsat data for enhanced process understanding in forest ecosystems [22], [29]. However, existing approaches for large area mapping with Landsat data have only limited potential for automation, work only for simple class separation

problems or are thematically restricted to certain environments. Therefore, further improvements in large area mapping methods are required [9].

Image compositing offers opportunities to overcome restrictions in data availability and to improve large area mapping and monitoring at the same time. Compositing was originally developed for wide-swath sensors that frequently provide global coverage. New image datasets are created by selecting one specific observation from numerous acquisitions or averaging spectral values. For coarse resolution sensors, such as AVHRR, the main purpose was to reduce the influence of clouds on the signal. Most commonly, simple decision rules were applied, such as the maximum/minimum band value or NDVI [8], [9]. More advanced compositing approaches additionally take a pixel's view angle into account [47], [52]. However, compositing was not often considered when working with high resolution data, mostly due to high data costs and processing constraints. Recent developments therefore encourage image compositing for Landsat data. These include a changed data policy (i.e. free data access—Loveland and Dwyer [34]), enhanced preprocessing algorithms leading to improved standard image product quality [7], [54] and advances in storage and computational resources [41], [44]. These developments will also affect future satellite missions beyond Landsat-8, such as Sentinel-2, which will most likely adapt to these standards [12].

Pixel-based compositing (PBC) of high-resolution optical imagery, as opposed to scene-based compositing or mosaicking, offers advantages for large area LULCC analysis. Global analysis models can be based on a single, homogeneous and cloud free dataset that ideally provides a consistent radiometric response across large areas [19]. As all unclouded observations per pixel are extracted from the Landsat archive, different image metrics can be generated as valuable byproducts of the compositing process. Such metrics can for example be produced to capture relevant phenologic states in the seasonal cycle of vegetation or can correspond to descriptive statistics that provide a measure of average spectral response or spectral variability [18]. Above all, the change from a scene based-to a pixel-based perspective yields several improvements:

- analyses are no longer restricted to a few “best” images, where low cloud cover often had to be favored over seasonally suited acquisitions
- the artificial partitioning into footprints can be overcome
- the entire image archive can be exploited and partially useful image areas (e.g. in SCL-off data) can easily be included
- observation frequency is increased as pixels unaffected by clouds can be incorporated from cloud contaminated images
- observation frequency is additionally increased by across track overlap exploitation

Taking these considerations into account, PBC emerges as a valuable tool for large area applications using high resolution optical data. First examples of Landsat large area image composites have recently been produced to map boreal forest change in Russia or to monitor deforestation in tropical Africa [42], [43]. For the area of the United States, composited ETM+

top-of-atmosphere imagery has recently been made available via a web interface [46]. Existing approaches, however, either rely on evaluating a single, often spectral criterion, do not utilize surface reflectances or do not retain all spectral bands at the original 30 m resolution of Landsat data. In this paper we present a compositing algorithm that takes advantage of state-of-the-art Landsat pre-processing and that can be flexibly parameterized to user-specific needs. Most importantly, we may optimize a compositing product for specific process and change regimes, varying applications and different geographic regions. Accordingly, our objectives were to:

- 1) Develop an algorithm that generates cloud free, radiometrically and phenologically consistent Landsat composites, based on flexible rule sets for pixel selection
- 2) Evaluate the radiometric consistency of the output composites with respect to MODIS BRDF normalized reflectance products
- 3) Assess the suitability of the composites and the surplus value of spectral variability byproducts from compositing for regional land cover mapping

II. METHODOLOGY

A. Data Used and Test Region

We developed our algorithm and evaluated the outputs over a region in Eastern Europe covered by 42 Landsat World-Reference-System-2 footprints (Fig. 1). The test region ($\sim 850,000 \text{ km}^2$) features mountainous terrain with elevations of up to 2,600 m in the Carpathian Mountains, as well as the lowlands of the Pannonian Basin, extending towards the South and the Adriatic Sea. For the detailed assessment of radiometric consistency, 10 sub regions of $60 \times 60 \text{ km}$ were selected. These sub regions were purposefully located in the central part of the study region and within scene overlaps. They were chosen to represent a broad variety of land cover configurations (Fig. 1). The region features a temperate-continental climate with increasing maritime influences towards the southwest. The forests in the region comprise coniferous mixed and deciduous forests. Coniferous forests are mostly dominated by Norway spruce (*Picea abies*) and silver fir (*Abies alba*) and deciduous forests feature European beech (*Fagus sylvatica*), pedunculate oak (*Quercus robur*), hornbeam (*Carpinus betulus*) and lime (*Tilia cordata*). The regional growing season begins in April and ends in October, but varies over elevation ranges and in response to annual rainfall variability [45].

We here aimed at producing image composites for three years, i.e. 2000, 2005 and 2010. Such even spaced, nominal dates are commonly used for quantifying LULCC [32]. Available imagery in the Landsat archive was not sufficient to achieve a full coverage from yearly data for each of these years. We therefore considered imagery acquired within \pm two years of the respective year for compositing. To differentiate between the five year interval composites and the actual acquisition date of the respective data set used in the compositing, we hereafter refer to 2000, 2005 and 2010 as “target years”. We used all available precision terrain corrected (L1T) Landsat TM and ETM+ imagery with less than 70% cloud cover from the USGS Landsat archive (Table I). Images acquired between

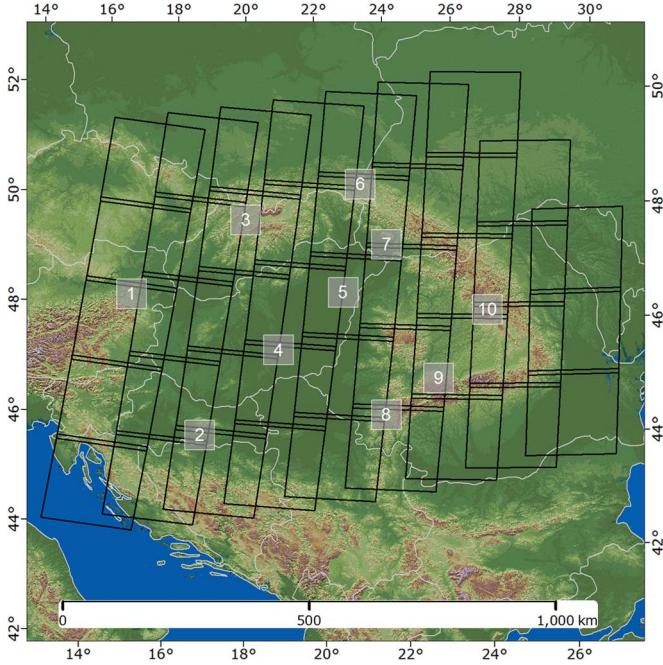


Fig. 1. Landsat footprints from path 190 / row 25 in the northwest to path 182 / row 29 in the southeast (overlaid on Shuttle Radar Topography Mission elevation data). Sub regions for evaluation purposes are indicated by grey rectangles.

TABLE I
LANDSAT DATA USED FOR THE COMPOSITING IN 2000, 2005 AND 2010

	2000	2005	2010
Total number of images	890	1478	1590
Total data volume [TByte]	0.80	1.82	1.51
Target year images [%]	29.1	15.3	34.7
Median day-of-year	181	203	198
Average cloud cover [%]	38.1	38.6	40.4

mid-November and mid-February were excluded to avoid the prevalent snow coverage, low sun elevation angles and shadowing, and consequently poor signal-to-noise-ratio.

We additionally acquired MODIS reflectance data to assess the consistency of spectral responses in Landsat composites. We considered MODIS surface reflectance to be a state of the art, large area image product that provides consistent spectral reference across space. We used the BRDF-adjusted, 16-day surface reflectance product (MCD43A4) with 500 m spatial resolution for 6 spectral bands to assess the radiometric response of our results. For this product, spectral responses are normalized to nadir viewing geometry, thus view-angle related effects are minimized [35]. Despite varying spatial resolutions, MODIS and Landsat data are well suited to be compared spectrally, as their equatorial overpass times are similar [25], both sensors have comparable spectral bands with corresponding center

wavelengths [7]. Furthermore, the atmospheric correction algorithm applied to MODIS data is similar to the correction method applied to the Landsat data in this study [36].

B. Pre-Processing

We atmospherically corrected all input images, converting digital numbers into surface reflectance to assure comparability of results across different Landsat sensors, footprints and acquisition dates. We used the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm to perform the correction [36]. The algorithm uses the Second Simulation of the Satellite Signal in the Solar Spectrum (6S, [49]) radiative transfer model in a similar manner as used to correct MODIS imagery. Surface reflectance bands 1–5 and 7 were used for further analysis at the original 30 m spatial resolution. Cloud masks were subsequently developed for all images using the FMASK algorithm [55]. The FMASK approach is an extended and improved rule-based method of the automated cloud cover assessment system (ACCA, [26]). Parameters such as cloud probability thresholds were set to conservative values to capture as many clouds and cloud shadows as possible. Additionally, we developed distance images providing the distance of each pixel to the next cloud or cloud shadow. All data was subsequently re-projected to Lambert Azimuthal Equal Area (LAEA) with a European Terrestrial Reference System 89 (ETRS89) datum so that imagery from different UTM zones can be composited into one single output dataset.

C. Image Compositing

As image frames of individual footprints are always shifted against each other, the greatest possible spatial extent of all inputs was determined and tile-wise parallel processing subsequently used. All data that spatially overlapped was extracted based on a tile-wise bounding-box test and the potentially unclouded observations according to the cloud masks were extracted.

During compositing, all available observations of a given pixel are assessed regarding their suitability for the composite. This assessment can be based on different parameters, for which a score is calculated for each available observation. The weighting of these parameter scores is obtained regarding the respective objectives pursued, e.g. with a high scoring weight on the day-of-year in the leaf-on season in case of temperate forests analysis. Subsequently, scores are summed up, and the spectral values of the observation with the highest score are written into the composite. All other available observations are used to produce spectral variability metrics. In the following section we describe the development of the parameter score functions.

We implemented a parametric weighting scheme to assess the suitability of observations, based on score functions for three parameters: (1) acquisition year (i.e. annual suitability), (2) acquisition day-of-year (DOY—seasonal suitability), as well as (3) the distance of a given pixel to the next cloud (risk of remnants of atmospheric disturbances). Compositing images from different dates and years requires considering the trade-offs between inter- and intra-annually matching acquisition dates, i.e.

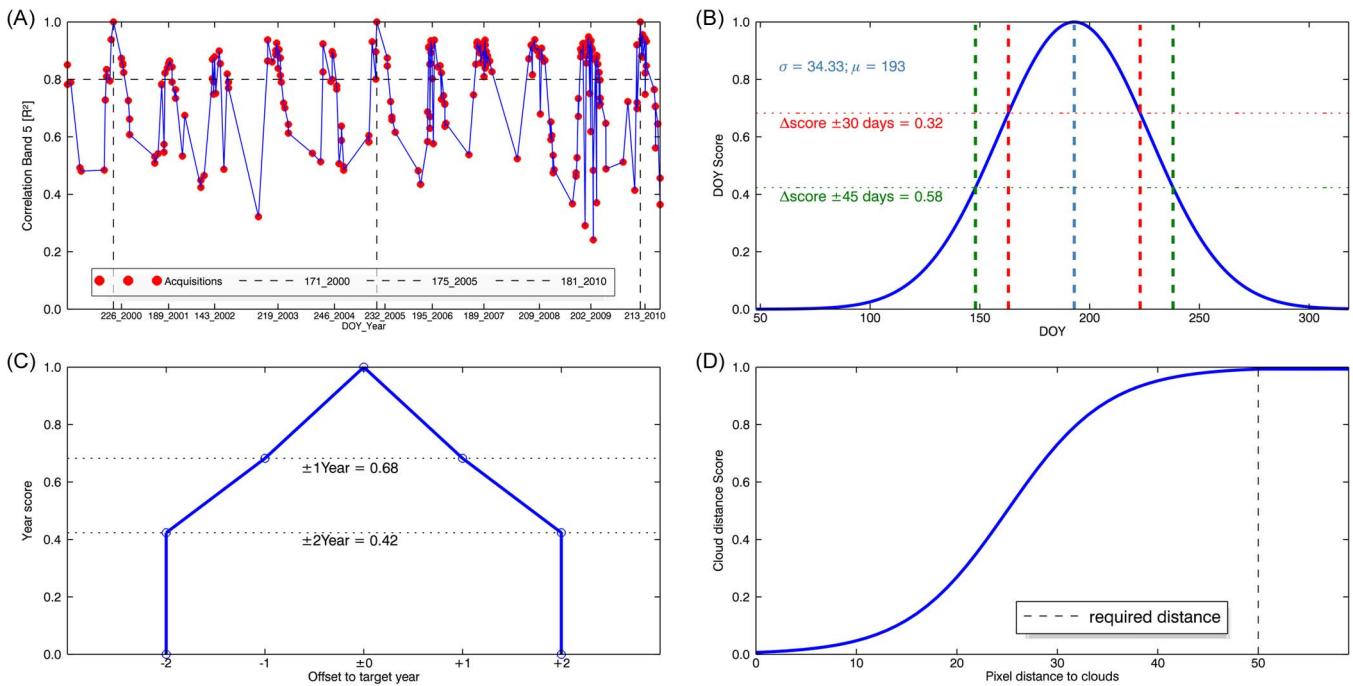


Fig. 2. Score functions: (a) Seasonal correspondence of available images with three selected reference images (indicated by vertical dashed lines) in Landsat band 5, horizontal dashed line indicates the $0.8 R^2$ cut-off value; (b) DOY score function with all year median DOY of 193 and all year standard deviation of 34.33, green and red dashed lines indicate the ± 30 and ± 45 day offset scores, respectively; (c) year score function; (d) cloud distance score function.

potential land cover changes and phenologic effects. As we focused on generic land use and land cover mapping, we aimed at producing composite datasets representing the seasonal period of main photosynthetic activity (i.e. a leaf-on state) and therefore first needed to assess the favored seasonal window.

We did this by first extracting all imagery from 1999 to 2011 available for selected areas with high data availability. From these, we selected three largely cloud free reference images for the target years (dashed vertical lines in Fig. 2(a)), and as close as possible to the middle of the year (DOY 183), which is an approximation of the peak photosynthetic activity in our study region. We then assessed how well all other images spectrally corresponded to these reference images. This assessment was based on stable broadleaf or mixed forest areas (i.e. areas with climate-dependent phenology). We then correlated shortwave infrared (SWIR—band five) reflectances of these areas for unclouded pixels in each image with reflectances in the respective reference image (target year \pm two years). We based this assessment on SWIR reflectance, which has been shown to be temporally more stable than near infrared (NIR), e.g. against understory or other background signal [10]. Fig. 2(a) shows the temporal variation of the correlation coefficient over time with a seasonal pattern becoming apparent for individual years.

We finally used an R^2 of 0.8 as a cut-off criterion to provide seasonally comparable observations and derived the median and standard deviation of these images. The resulting all year median DOY (193 or July 12th) was employed to anchor a Gaussian scoring function (Fig. 2(b)) with the resulting standard deviation for all images with an R^2 higher than 0.8 (34 days). We used this function to evaluate the seasonal suitability (i.e. DOY) of all images for all years according to (1):

$$Score_{DOY} = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_i - \mu}{\sigma})^2} \quad (1)$$

where σ denoted the all year DOY standard deviation, μ is the target DOY and x_i is the DOY for a given acquisition.

Next, we defined a preferred seasonal window to ensure favoring a certain range of DOYs within the target year over a more central DOY of previous or subsequent years. On the one hand, favoring an optimum DOY over choosing a pixel from the central target year is reasonable, as yearly phenological cycles largely drive image statistics in a central European setting. This implies that the DOY criterion is more important than the year itself. On the other hand, extending the number of years considered during compositing increases the likelihood of including land cover changes in the output datasets. Intra-annual NDVI profiles from MODIS data and field experience proved that images acquired within 30 days from the median DOY always related to phenological states of photosynthetic peak activity for forest areas. The peak of intra-annual NDVI profiles examined for deciduous and mixed forests always occurred within this seasonal window. During the selection of acquisition dates, a ± 30 day offset to the median DOY was therefore always favored over images acquired later or earlier than the target year. About ± 15 days around this central window of peak phenology, some of the regions' forest stands appeared to be in a slightly pre-peak phenologic state or already showed first signs of senescence. Therefore, we estimated a DOY offset of ± 45 days to be a suitable threshold to delineate the secondary period of main photosynthetic activity for temperate forests encountered in our test region. To balance the risk of including non-peak phenology (larger DOY offset) against the risk of including potential land cover changes (greater with increasing temporal offset from the target year), acquisitions from within 45 days of the target DOY (i.e. May 28th and August 26th) should always be favored over images acquired two years from the target year. We therefore assessed the difference in the DOY score for 30 and 45 days

TABLE II
OVERVIEW OF DIFFERENT COMPOSITING OUTPUTS PRODUCED PER TARGET YEAR RESULT

Output type	Description
Best observation composite	Pixels with highest score for evaluated parameters
Spectral variance metrics	Mean, standard deviation and range images for spectral bands 1-5 and 7.
	Sum of reflectance in nIR and swIR bands normalized by number of observations
Flag metrics	Path/row, Acquisition year/month/day, Acquisition DOY, Number of clear observations, determined decision score value, sun zenith during acquisition

(Fig. 2(b)) and constructed a piecewise linear year scoring function (Fig. 2(c)).

We scored a pixel's distance to clouds or cloud shadows based on a sigmoidal function in conjunction with a user defined value defining the minimum required distance to clouds/shadows. Examination of Landsat images with different atmospheric conditions indicated that many pixels directly adjacent to clouds are not always identified as such during cloud masking, but may still be strongly affected by neighboring clouds. While this influence can be assumed to decrease exponentially with increasing distance from clouds, the minimum required distance is not easily derivable from the input imagery alone. Therefore, this value needs to be defined depending on the abundance of imagery, providing more conservative (i.e. higher) values if lots of imagery is available and lower values in cases where imagery is relatively scarce. For this study we defined a minimum required distance of 50 pixels (dashed line in Fig. 2(d)), as a compromise between falsely excluded clear pixels and including undesired observations (e.g. affected by water vapor or haze). All pixels that were located closer to clouds than 50 pixels obtained a score according to (2):

$$Score_{CloudDist.} = \frac{1}{1 + e^{(-0.2 \cdot \min\{D_i, D_{req}\} - (\frac{D_{req} - D_{min}}{2}))}} \quad (2)$$

Where D_i indicates a given pixels distance to clouds, D_{req} is the defined minimum required distance, D_{min} is the minimum distance of the given pixel observations. As the data still contained some cloud remnants, a threshold based cloud-remnant removal was performed. Thresholds were based on the visible wavelengths' reflectance and the relation between visible, NIR and SWIR bands, yielding the final unclouded observations (hereafter: clear observations).

The provided scores are subsequently summed up with the cloud distance score weighted half of the scores for DOY and year, respectively. Finally all 6 spectral band values of the observation with the highest final score are written into the best observation composite. Should there be several observations with identical final scores, the one closest to the target year is selected. The observation featuring the lower reflectance in the blue band is selected if no winner could be determined during the previous steps, to additionally minimize atmospheric effects.

After the clear observations have been determined, additional datasets and metrics were assembled. Besides the best observation composites, we produced image variability measures and flag images. Spectral variability measures take advantage of all extracted clear observations. Thus they contain implicit information on land cover phenology and spectral-temporal stability. Here, we produced mean, standard deviation and range images for each spectral band. Additionally, we created a band providing the sum of SWIR/NIR bands normalized over the number of clear observations. The flag bands provide information on acquisition date (year and DOY), the number of clear observations, the selected image footprint, the score determined for the best observation as well as winner acquisition sun zenith angle. Table II summarizes the individual outputs produced during compositing and Fig. 3 provides examples of compositing outputs for the target year 2005.

D. Evaluating Annual, Seasonal and Radiometric Consistency

We evaluated the best observation composites for the three target years in terms of annual and seasonal consistency. This evaluation was based on the respective flag images for different cases relating to the decision for a certain acquisition date: The ideal case-1 has a pixel selected from the target year and a DOY within 30 days from the target DOY. The least desirable case would represent a two year offset from the target year and a DOY offset that is greater than 45 days (case-9). Table III provides an overview of the investigated temporal compositing cases.

The evaluation of radiometric consistency was based on the comparison with MODIS reflectance data. We acquired the temporally closest MODIS 16-day composite to the target DOY and year. We then adjusted our best observation composites to 500 m resolution using simple pixel aggregation. To assess how well our PBC outputs corresponded to the MODIS products, we derived band-wise Pearson correlation coefficients (R^2) and RMSE for the 10 sub regions and the three PBC outputs. We additionally assessed the temporal composition of selected pixels within the sub regions according to the 9 defined cases. We also provided the average number of available clear observations for each subset and PBC output.

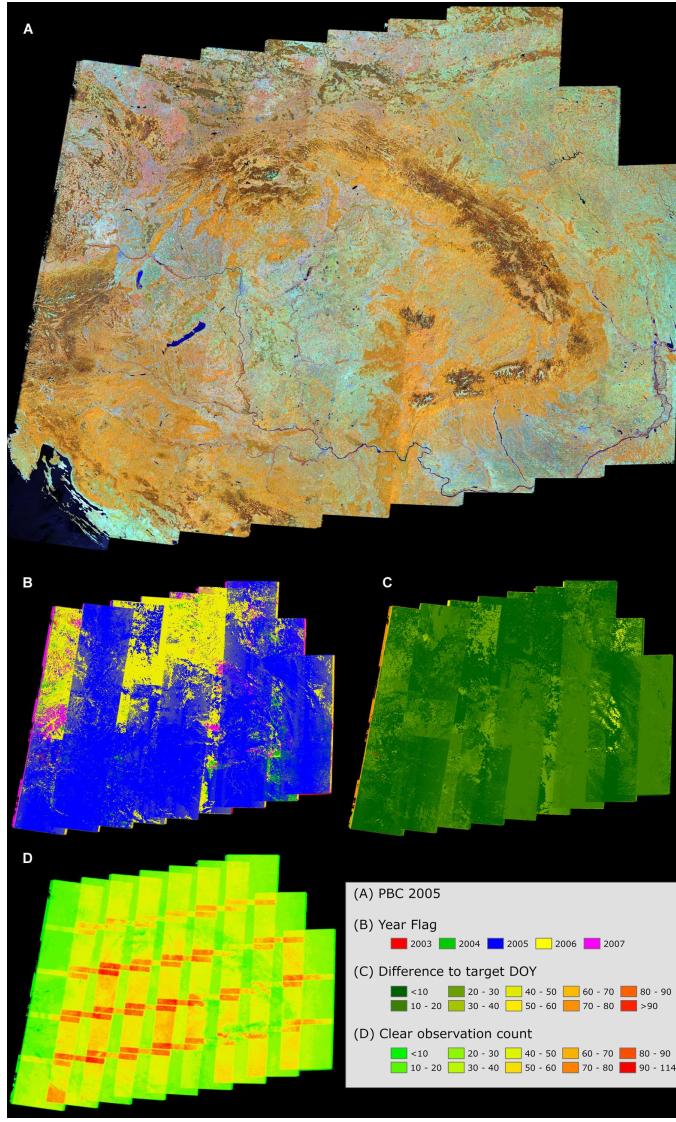


Fig. 3. Output examples of the 2005 PBC: (A) best observation composite ($RGB = 4,5,3$); (B) year flag image; (C) relative difference to target DOY; (D) number of clear observations per pixel.

E. PBC-Based Land Cover Mapping

Finally, we evaluated the performance of our composited image products for land cover classification. Land cover mapping performance was assessed using a random forest (RF) classifier [51] and a sample of interpreted randomly sampled points. We sampled a total of 2,000 points randomly with a minimum distance of 1 km to minimize spatial autocorrelation. The samples were then interpreted using VHR imagery along with the original Landsat imagery and ground truth data. We only retained samples if their interpreted land use/land cover did not change between 2003 and 2007, yielding a total sample size of 1,623 points. We aimed at achieving a thematic depth comparable to that of common global land cover products [1], [3], [14]. We consequently targeted at separating coniferous forest, mixed forest, deciduous forest, agriculture, grassland, built-up and water. We used a single RF model over the entire test region to assess if the composited input data allows for homogeneous classification results based on one global

TABLE III
EVALUATED CASES FOR TEMPORAL COMPOSING OF ACQUISITION DATES.
THE COLUMN “PRIORITY” INDICATES THE RANKED PREFERENCE OF CASES
FOR COMPOSING

Case #	Offset to target year	Offset to target DOY	Priority
Case 1	0 years	± 30 days	1
Case 2	0 years	> 30 days - ≤ 45 days	4
Case 3	0 years	> 45 days	7
Case 4	± 1 year	± 30 days	2
Case 5	± 1 year	> 30 days - ≤ 45 days	5
Case 6	± 1 year	> 45 days	8
Case 7	± 2 years	± 30 days	3
Case 8	± 2 years	> 30 days - ≤ 45 days	6
Case 9	± 2 years	> 45 days	9

TABLE IV
ACQUISITION YEARS OF PIXELS USED FOR COMPOSING (NOTE: NO INPUT
IMAGERY EXISTED FOR 1998 AND 2012)

Target Year	-2	-1	0	+1	+2
2000 [%]	---	9.52	65.91	21.30	3.27
2005 [%]	0.43	2.40	67.91	26.26	3.01
2010 [%]	0.13	8.06	91.24	0.57	---

classification model. We then trained different RF classification models for the entire test region for the 2005 PBC using different input feature stacks: (a) only the best observation composite, and (b) the best observation composite together with the spectral variability metrics. The RF models were trained using 500 individual trees and the number of features at each split was set to the square root of the number of input features [50]. Validation was based on 10-fold cross validation of the interpreted samples, with 10% of the samples left out for the validation of each resulting model that was trained with the remaining 90% of samples [31]. We derived overall accuracies, producer’s and user’s accuracy as well as kappa statistics.

III. RESULTS

The resulting composites were free of clouds and homogeneous in appearance even though data from different years and different seasons was broadly mixed (Fig. 3(a), (b), (c)). Between 91% (2010) and 66% (2000) of all pixels consisted of observations acquired within the respective target year (Table IV). No more than 3.5% of all pixels selected during compositing were acquired two years earlier or later than any of the target

TABLE V
TEMPORAL COMPOSITION OF CASES IN THE THREE BEST OBSERVATION COMPOSITES

	<i>Case-1</i>	<i>Case-2</i>	<i>Case-3</i>	<i>Case-4</i>	<i>Case-5</i>	<i>Case-6</i>	<i>Case-7</i>	<i>Case-8</i>	<i>Case-9</i>
2000 [%]	51.65	13.60	0.66	30.21	0.49	0.12	3.02	0.04	0.21
2005 [%]	66.51	1.21	0.19	28.50	0.06	0.09	2.74	0.15	0.55
2010 [%]	88.18	2.83	0.23	8.51	0.06	0.06	0.05	0.04	0.04

years. Even though the seasonal range spanned spring to autumn in all three composites, the majority of selected observations captured peak phenology of green vegetation.

Accordingly, between 52% (2000) and 88% (2010) of selected pixels were case-1 pixels, i.e. acquired within the target year and within 30 days of the target DOY (Table V). For target year 2000, about 14% of pixels were acquired within 45 days of the target DOY (case-2), while the share of case-2 pixels was considerably lower for the other years (<3%). The distribution of case-4 pixels was relatively even for the 2000 and 2005 result (30% and 29%, respectively) and considerable lower for the 2010 result (8.5%). Thus, seasonal consistency was highest within the 2010 outputs, with more than 96% of pixels acquired within the preferred seasonal window (cases 1, 3, 7).

The share of pixels with an unfavorable seasonal state (i.e. DOY offset to target DOY >45 days) was below 1% for all PBC outputs. The number of extracted cloud free observations per pixel ranged from one to 114 for the 2005 composite. On average, between 21 (2000) and 39 (2005) unclouded observations were available for each pixel. The number of clear observations was especially increased within the across track overlap areas (Fig. 3(d)). For mountainous areas, the number of available unclouded observations was considerably lower (Fig. 3(d)). These areas also exhibited higher year and DOY offsets (Fig. 3(b), (d)).

The PBCs resembled spectral and radiometric characteristics of the MODIS 16-day reflectance composites well (Fig. 4). However, the land cover composition in the sub regions influenced the level of agreement. Forests generally lead to high correlations while spectrally dynamic land cover types, such as agriculture or grasslands, suffered from spectral variation or land cover changes. Second, the visible bands generally exhibited lower correspondence between MODIS and PBC reflectances, presumably due to atmospheric effects. In many cases, this effect seemed to be related to considerable spectral scatter in the MODIS visible bands (e.g. Fig. 4(d), (j)).

Sub region 1 was strongly dominated by broadleaved and coniferous forests within mountainous terrain (Fig. 4(a), (b)). The 2005 and 2010 results achieved R^2 values higher than 0.8 in five bands (Table VI). The 2010 PBC resulted in correlations of 0.9 in the green and red spectral bands. The temporal composition of pixels was very different between the 2005 and 2010 results (Table VII). More than 80% of the 2005 result consisted of pixels acquired with at least one year offset. In the case of the 2010 result, 83% of pixels related to case-1. The temporal composition in the 2000 case was more variable with 3% of case-2 and case-3 pixels. The average number of extracted clear obser-

vations was relatively low (15) and the corresponding MODIS image lacked more than 50% of observations (Fig. 4(a)).

Forest dominated sub region 3 is located in mountainous terrain at higher elevations where cloud coverage is frequent (Fig. 4(c), (d)). Here the year 2000 result had the lowest average number of clear observation available (13), about 48% of pixels were case-2 observations. The corresponding MODIS data for this period lacked more than 90% of observations. The composition of acquisition dates for the 2000 result covered a wider range of cases and correlations to MODIS bands were comparatively low. The 2005 result (~74% case-4) achieved higher R^2 values in five out of six bands (Fig. 4(d)), even though the 2010 output consisted of over 98% of case-1 pixels.

Sub region 5 featured mostly agricultural land and grassland and was located at lower elevations (Fig. 4(e)). Here the 2000 result performed worst in four out of 6 bands, relating to 41% of case-4 pixels and 12% of case-7 pixels. The corresponding MODIS product featured approximately 10% of no data pixels and the number of clear observations was as low as 19. With more than 99% of case-1 acquisitions, the 2010 result achieved the highest correlations to the MODIS bands.

Sub region 6 exhibited a mosaic of grassland and agriculture dominated valleys and forested slopes and ridges over medium elevation terrain (Fig. 4(f)). For the 2000 result, relatively few acquisitions were available (21) and consequently only 24% of pixels corresponded to case-1, while the majority (68%) were case-4 pixels. About 7% of pixels were acquired within the extended seasonal window. The achieved R^2 values were the highest in all bands. The 2010 result contained 21% of case-2 pixels but the 2005 output, with 53% case-4 acquisitions, mostly achieved lower R^2 values.

Sub region 8 featured a variety of land cover types over a range of elevations. The highest correlations were achieved in the 2005 result (in 5 out of 6 bands) even though one third of pixels were acquired with one year offset to the target year. Most of these pixels were located over forested slopes in the eastern part of the sub region (Fig. 4(h)). The 2010 result did not achieve high correlations despite of featuring more than 90% of case-1 pixels. Sub region 9 again was strongly dominated by coniferous forests (Fig. 4(i), (j)). Here the highest overall correspondence to MODIS reflectances was achieved in the NIR bands for the 2005 result (0.92). The 2000 and 2005 results both achieved similarly high correlations and both were composed of more than 90% of case-1 pixels. The NIR and SWIR bands for the 2010 result achieved comparably high correlations and 1/3 of pixels were case-4 observations.

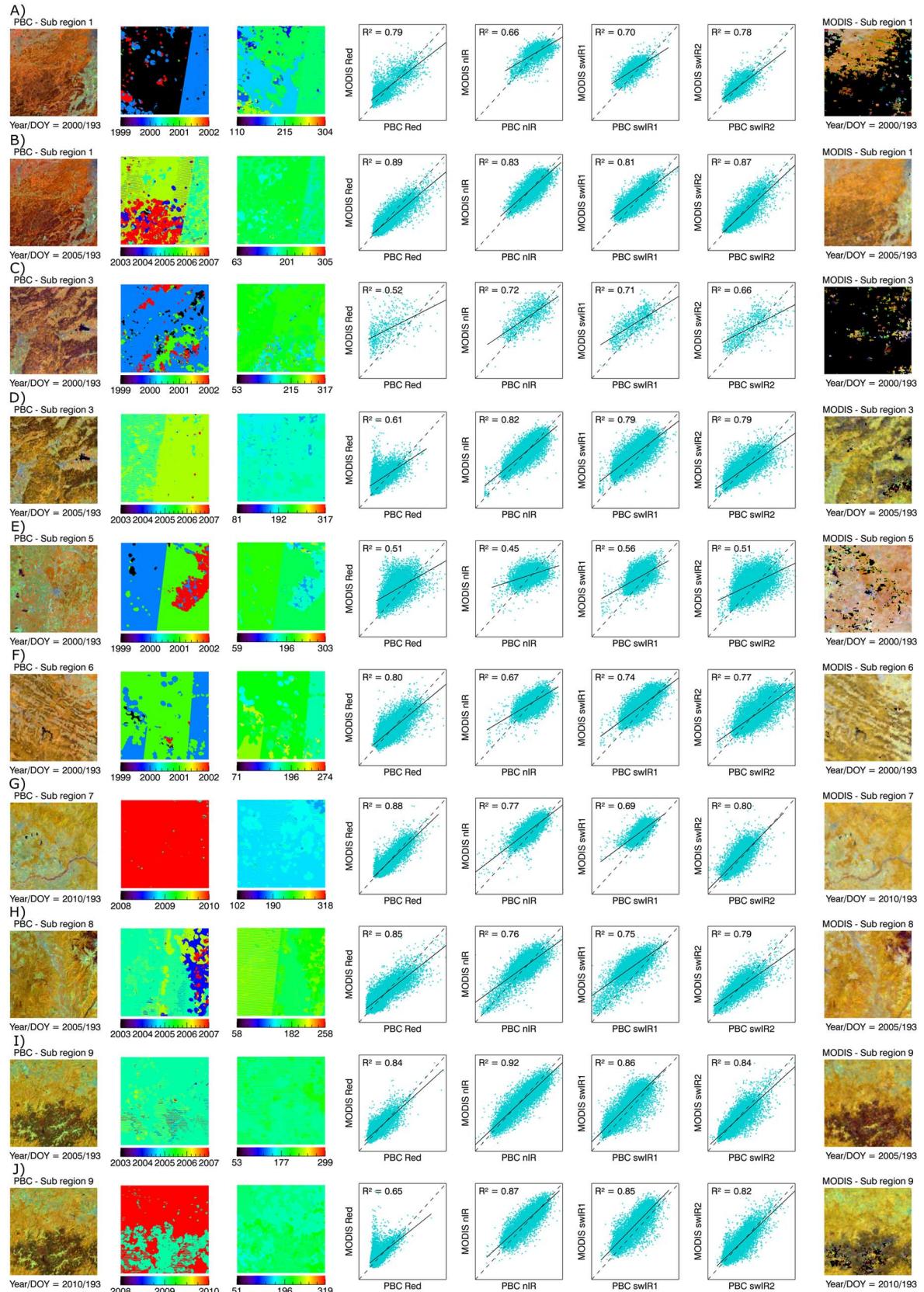


Fig. 4. Radiometric consistency evaluation (from left to right): (column 1) aggregated PBC imagery (RGB = 4, 5, 3); (column 2) corresponding year flag images; (column 3) corresponding DOY flag images; (columns 4–7) band-wise scatterplots for red, NIR, SWIR1 and SWIR2 bands of MODIS (Y axis) and PBC (X axis) data; (column 8) corresponding MODIS nadir BRDF-adjusted 16-day reflectance product (MCD43A4, RGB = 2, 6, 1).

TABLE VI

RESULTS OF THE EVALUATION OF THE RADIOMETRIC CONSISTENCY ASSESSED FOR THE 10 SUB REGIONS (FIRST COLUMN), COEFFICIENTS OF DETERMINATION (R^2) AND ROOT MEAN SQUARED ERRORS (RMSE) FOR THE 6 CORRESPONDING MODIS AND AGGREGATED PBC BANDS

Radiometric consistency: PBC vs. MODIS (R^2 / RMSE)						
	TM 1 vs. MOD 3	TM 2 vs. MOD 4	TM 3 vs. MOD 1	TM 4 vs. MOD 2	TM 5 vs. MOD 6	TM 7 vs. MOD 7
Sub region 1	2000 0.35 / 195.59	0.7 / 168.1	0.79 / 173.4	0.66 / 396.24	0.7 / 334.17	0.78 / 207.68
	2005 0.77 / 133.78	0.88 / 96.93	0.89 / 125.73	0.83 / 331.04	0.81 / 275.93	0.87 / 175.09
	2010 0.81 / 124.25	0.9 / 100.67	0.91 / 121.81	0.82 / 308.56	0.85 / 337.87	0.88 / 180.75
Sub region 2	2000 0.85 / 86.71	0.86 / 131.1	0.87 / 130.46	0.76 / 491.67	0.76 / 436.25	0.83 / 195.08
	2005 0.75 / 87.2	0.85 / 107.66	0.85 / 102.1	0.74 / 371.07	0.68 / 424.3	0.79 / 157.82
	2010 0.77 / 95.59	0.88 / 98.77	0.88 / 108.85	0.79 / 289.42	0.79 / 402.81	0.85 / 169.67
Sub region 3	2000 0.11 / 542.31	0.5 / 436.72	0.52 / 374.51	0.72 / 593.87	0.71 / 488.62	0.66 / 276.88
	2005 0.1 / 232.82	0.62 / 223.65	0.61 / 185.9	0.82 / 472.36	0.79 / 356.3	0.79 / 168.38
	2010 -0.04 / 273.7	0.51 / 217.42	0.57 / 195.98	0.8 / 436.19	0.8 / 358.94	0.79 / 190.44
Sub region 4	2000 0.48 / 168.23	0.56 / 156.07	0.58 / 218.5	0.48 / 397.8	0.6 / 477.01	0.62 / 380.35
	2005 0.79 / 162.31	0.81 / 79.74	0.81 / 119.01	0.73 / 288.33	0.73 / 333.02	0.77 / 261.31
	2010 0.42 / 114.96	0.66 / 121.81	0.68 / 142.27	0.73 / 370.34	0.66 / 443.23	0.67 / 303.31
Sub region 5	2000 0.31 / 134.78	0.51 / 166.53	0.51 / 209.91	0.45 / 483.42	0.56 / 490.78	0.51 / 374.38
	2005 0.42 / 130.15	0.64 / 99.31	0.56 / 166.74	0.48 / 557.04	0.5 / 318.36	0.48 / 388.15
	2010 0.63 / 140.23	0.75 / 83.04	0.77 / 126.9	0.75 / 354.93	0.67 / 337.17	0.69 / 250.56
Sub region 6	2000 0.52 / 143.5	0.82 / 92.04	0.8 / 86.86	0.67 / 390.69	0.74 / 293.05	0.77 / 144.44
	2005 0.08 / 243.88	0.53 / 246.88	0.48 / 207.4	0.67 / 477.81	0.63 / 361.75	0.64 / 167.69
	2010 0.1 / 146.07	0.61 / 159.65	0.55 / 138.57	0.67 / 462.88	0.65 / 391.88	0.64 / 157.7
Sub region 7	2000 0.69 / 152.95	0.77 / 149.84	0.72 / 235.84	0.7 / 428.37	0.61 / 383.6	0.66 / 352.63
	2005 0.31 / 144.94	0.75 / 139.49	0.8 / 146.53	0.74 / 343.01	0.59 / 299.43	0.75 / 184.4
	2010 0.67 / 114.32	0.85 / 94.51	0.88 / 91.11	0.77 / 335.81	0.69 / 386.71	0.8 / 143.04
Sub region 8	2000 0.76 / 85.19	0.81 / 157.34	0.76 / 198.61	0.43 / 582.38	0.7 / 562.07	0.68 / 311.28
	2005 0.57 / 113.38	0.84 / 89.71	0.85 / 81.64	0.76 / 374.26	0.75 / 365.26	0.79 / 124.11
	2010 -0.1 / 235.41	0.4 / 192.15	0.42 / 165.45	0.74 / 355.06	0.73 / 324.72	0.73 / 120.13
Sub region 9	2000 0.86 / 84.24	0.86 / 121.46	0.86 / 158.29	0.84 / 442.9	0.87 / 377.5	0.87 / 227.75
	2005 0.53 / 107.88	0.84 / 99.92	0.84 / 102.28	0.92 / 343.05	0.86 / 310.78	0.84 / 151.11
	2010 0.02 / 206.79	0.55 / 179.75	0.65 / 161.17	0.87 / 399.34	0.85 / 370.79	0.82 / 166.82
Sub region 10	2000 0.69 / 106.84	0.83 / 107.15	0.81 / 118.25	0.84 / 430.16	0.86 / 356.22	0.82 / 194.65
	2005 0.27 / 116.39	0.74 / 140.13	0.7 / 125.58	0.89 / 383.13	0.83 / 356.38	0.82 / 130.09
	2010 -0.09 / 188.15	0.55 / 161.04	0.47 / 140.84	0.83 / 441.1	0.81 / 338.59	0.77 / 131.69

Results from the land cover classification underlined the radiometric consistency of the results. The composited imagery is radiometrically stable enough to train and apply a single classification model over the entire region. The achieved overall accuracies (Table VIII) were 65% for the classifications using only the best observation composite (run (a)), and 72% for the run using the variability measures as additional input features (run (b)). There were considerable differences among both classification runs. Overall, the grassland class (training sample size 102) achieved the lowest accuracies with user's accuracies of 14% and 9% for run (a) and run (b) respectively. Forest classes achieved better results. For example the coniferous forest class was validated with a user's accuracy of 54% and 60%, and even higher user's accuracies of 84% and 92% for classification runs (a) and (b), respectively. The mixed forest class showed the weakest performance among the three forest classes with relatively high omission and commission errors. We found considerable improvements in classification accuracies when the spectral variability measures were included as input features. The agriculture, built up and water classes achieved high accuracies in both classification runs and were validated with even

higher accuracies when the variability metrics were additionally used. User's accuracies improved by about 10% during classification run (b) for the deciduous and mixed forest as well as the built up classes. Similarly, producer's accuracies improved considerably for mixed forest, grassland and built up classes. It is interesting to note that compositing artifacts related to SLC-off scan line errors were visible in some areas of the land cover map when using only the PBC. However, these artifacts were not visible in the map resulting from the PBC plus the variability measures as inputs. Moreover, the result obtained from only using the PBC as input features appeared to be stronger affected by salt-and-pepper like speckle in the classified image. Also the delineation of built-up areas was spatially more consistent using the PBC and variability measures for the classification (Fig. 5 top).

IV. DISCUSSION AND CONCLUSIONS

In this paper we describe a new algorithm for pixel based compositing of Landsat data. Our approach takes advantage of automated pre-processing approaches and facilitates generating regional image data sets from the Landsat archive. For selecting

TABLE VII

RESULTS OF THE EVALUATION OF THE TEMPORAL COMPOSITION OF PIXELS ACCORDING TO THE 9 DEFINED CASES ASSESSED FOR THE 10 SUB REGIONS (FIRST COLUMN), PROVIDED FOR EACH TARGET YEAR COMPOSITE. ADDITIONALLY, THE AVERAGE NUMBER OF AVAILABLE UNCLOUDED OBSERVATIONS IS PROVIDED

	Radiometric consistency: Temporal composition of pixels									Clear observations
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	
Sub region 1	2000	25.19	3.06	3.14	64.83	0.21	0.01	3.56	0.00	0.00
	2005	17.90	0.08	0.31	59.61	0.00	0.00	22.09	0.00	0.01
	2010	82.71	0.15	0.18	16.93	0.00	0.00	0.01	0.00	0.02
Sub region 2	2000	77.03	21.66	0.00	1.29	0.00	0.00	0.02	0.00	0.00
	2005	96.30	0.00	0.00	3.66	0.00	0.00	0.03	0.00	0.00
	2010	99.66	0.01	0.00	0.33	0.00	0.00	0.00	0.00	0.00
Sub region 3	2000	26.54	47.66	0.92	17.07	0.00	0.51	7.08	0.20	0.01
	2005	26.13	0.01	0.03	73.61	0.00	0.00	0.21	0.00	0.00
	2010	98.34	0.66	0.00	0.97	0.00	0.02	0.01	0.00	0.00
Sub region 4	2000	78.11	7.95	0.10	11.88	0.00	0.00	1.96	0.00	0.00
	2005	94.09	0.00	0.00	5.89	0.00	0.00	0.01	0.00	0.00
	2010	99.67	0.00	0.00	0.32	0.00	0.00	0.00	0.00	0.00
Sub region 5	2000	44.20	0.19	2.78	40.64	0.00	0.00	12.19	0.00	0.00
	2005	77.31	0.00	0.00	22.67	0.00	0.00	0.02	0.00	0.00
	2010	99.12	0.01	0.00	0.87	0.00	0.00	0.00	0.00	0.00
Sub region 6	2000	23.58	7.01	0.27	68.45	0.01	0.15	0.52	0.00	0.00
	2005	46.17	0.15	0.00	53.42	0.00	0.00	0.26	0.00	0.00
	2010	70.68	20.45	0.01	8.61	0.18	0.00	0.07	0.00	0.00
Sub region 7	2000	62.98	2.07	0.00	34.94	0.00	0.00	0.00	0.00	0.00
	2005	46.17	0.15	0.00	53.42	0.00	0.00	0.26	0.00	0.00
	2010	99.54	0.00	0.00	0.46	0.00	0.00	0.00	0.00	0.00
Sub region 8	2000	20.90	36.98	0.00	42.11	0.00	0.00	0.00	0.00	0.00
	2005	65.83	0.00	0.00	31.41	0.11	0.01	2.64	0.00	0.00
	2010	91.65	0.67	0.00	7.67	0.00	0.00	0.01	0.00	0.00
Sub region 9	2000	99.14	0.42	0.00	0.43	0.00	0.00	0.00	0.00	0.00
	2005	92.05	0.00	0.00	6.30	0.00	0.00	1.64	0.00	0.00
	2010	68.63	0.44	0.00	30.91	0.00	0.00	0.02	0.00	0.00
Sub region 10	2000	73.61	1.91	0.00	24.46	0.00	0.00	0.02	0.00	0.00
	2005	80.71	7.75	0.00	8.85	0.00	0.00	2.68	0.00	0.00
	2010	65.99	0.33	0.00	33.68	0.00	0.00	0.00	0.00	0.00

the most suitable from all available observations, we implemented a parametric weighting scheme. We described the development of score functions that facilitate homogeneous surface reflectance composites at 30 m resolution. Seasonal and annual consistency was satisfying and complete coverage was achieved even over mountainous terrain with frequent cloud cover.

We further achieved high radiometric correspondence to MODIS reflectance products. The level of correlation varied according to land cover composition. Our results underpin that optimizing phenologic consistency is more relevant to achieve radiometrically consistent composites than using data from the target year itself—at least in a temperate setting. We successfully utilized PBC to map land cover using a relatively small training set. Overall and class specific accuracies improved considerably when spectral variability metrics were additionally provided to the classifier.

With regard to the opening of the Landsat archive and considerable advancements of pre-processing methods, compositing is a logical consequence of these recent developments. Our

approach takes advantage of automated procedures for cloud masking [55] and atmospheric correction [36]. These methods allow for full automatization and processing of large volumes of data. Using atmospherically corrected surface reflectance data is a great asset, as relative radiometric normalization is not easily feasible with thousands of input scenes from different years and seasons. Our implementation of the here presented compositing methodology builds on parallel processing capacities and was run on a relatively modest multi-core server. However, the composites were produced within a time window comparable to time invested 10 years ago into manual cloud masking for single scene classifications. The potential to increase the performance using more advanced high performance computing environments (e.g. graphics processing units, cloud computing) thus is considerable [33].

We base the decision for image selection during compositing on a flexible parametric weighting scheme that evaluates available observations for their suitability. While this decision system can be easily extended to include additional param-

TABLE VIII

RESULT OF THE CROSS VALIDATION OF THE CLASSIFICATION RESULTS FOR (A) ONLY BEST OBSERVATION BANDS AS INPUT, AND (B) THE BEST OBSERVATION BANDS AND THE SPECTRAL VARIABILITY MEASURES (PA = Producer's accuracy, UA = user's accuracy, OAC = overall accuracy, CF = coniferous forest, MF = mixed forest, DF = deciduous forest, AG = agriculture, GL = grassland, BU = built up, WT = water). ADDITIONALLY THE NUMBER OF SAMPLES PER CLASS FOR TRAINING IS PROVIDED

Class	<i>RF classification model</i>				
	(a) only PBC		(b) PBC + Stats		
	PA [%]	UA [%]	PA [%]	UA [%]	# samples
CF	53.75%	83.98%	59.60%	92.19%	256
MF	36.19%	16.03%	50.41%	26.16%	237
DF	62.11%	69.82%	69.70%	74.85%	338
AG	76.71%	85.62%	81.50%	93.93%	577
GL	50.00%	13.73%	75.00%	8.82%	102
BU	72.22%	32.50%	87.50%	43.75%	80
WT	90.00%	81.82%	95.83%	69.70%	33
OAC	64.70%		71.47%		
Kappa	0.54		0.62		

ters (e.g. local solar incidence angle, off-nadir pixel position, etc.), we here focused on three important parameters that are essential for achieving homogeneous composites. Designing the scoring functions for the selection of annual and seasonal acquisition dates appeared most important. We developed these scoring functions empirically based on the available data and knowledge of the regional seasonality. We used intra-annual vegetation profiles from MODIS for validation purposes only. This approach allows to parameterize the temporal score functions independent of the availability of MODIS data and thus also retrospectively for data since the beginning of the Landsat archive. However, the potential of integrating MODIS-scale data for seasonal fine-tuning during compositing or to account for annual variation in phenology is considerable and should be subject of future research. While the performance of the scoring functions in other regions still needs to be evaluated the parameterization can be derived in a similar manner elsewhere. If annual consistency is focal e.g., when analyzing deforestation in quickly regrowing tropical forest regions, the annual weight has to be increased relative to the seasonal weight [5]. Moreover, the weighting scheme can be parameterized to produce datasets tailored for specific mapping applications. For example, a leaf-on vegetation state could potentially be supplemented by a senescent seasonal state to improve the differentiation of forests types [4]. However, as Landsat data is predominantly acquired during the growing season in the temperate zone, the annual window for compositing would have to be increased considerably in order to achieve full coverage when compositing spring or autumn observations.

The results were annually and seasonally overall consistent for all three output composites. The annual consistency was slightly higher for the 2005 and 2010 result. The reason for this is that more images were available during the corresponding time periods, as the long term data acquisition plan was adjusted to compensate for the SLC failure after 2003 [28]. The ability to produce PBCs from the Landsat archive for large areas with surface reflectance at 30 m spatial resolution is a great asset for

many applications [48]. Within our results we achieved full regional coverage also for areas with persistent cloud cover. In such cases, the seasonal and annual offset of selected pixels was relatively high to enable a complete coverage (i.e. up to two years offset to the target and possibly DOY offsets greater than 45 days). The corresponding MODIS 16-day products often also lacked many observations within these areas (e.g. Fig. 4(a), (c)).

Our results showed that PBCs can achieve high radiometric consistency across space and time. In many cases, our results strongly correlated with the MODIS reflectance product even when imagery from different years was used for compositing. Larger shares of pixels with a greater seasonal offset (i.e. greater than 30 days) increased variance in the correspondences between MODIS and aggregated PBC bands. Multi annual imagery was increasingly incorporated into the PBC for mountainous areas with frequent cloud cover where MODIS data did not provide sufficient observations for a given 16-day period (e.g. Fig. 4(a), (c)). Land cover composition had a marked effect on the correlation between Landsat PBC and MODIS data. We showed that for areas exhibiting a natural phenology (especially forests), seasonal consistency over multiple years is more important than annual consistency (i.e. cases 1, 4 and 7 are more crucial than cases 2, 5 and 8). Areas with spectrally variable and spatially heterogeneous land cover configurations, such as agricultural areas, only achieved high correlations with MODIS data when annual and seasonal acquisition dates were very similar. For these areas, land cover changes (i.e. cropped vs. barren fields) that occurred during the multi-year compositing period considerably decreased spectral correspondence to the MODIS data. Specifically in the case of fine-textured land cover configurations, introducing MODIS point spread functions during aggregation might have influenced our analysis. However, the averaging effect related to introducing the PSF would not have altered the overall results [24].

As an example for an application using PBCs, we performed a set of land cover classifications. Results showed that the PBCs proved suitable for consistent and easy to implement land cover mapping over large areas. We used a relatively small set of randomly selected and visually labeled samples for training and a generic training strategy of the classification models. Regardless of this simple and straightforward approach, class delineations were consistent across the entire region. Several classes achieved consistently high accuracies over a heterogenous region covered by 42 Landsat footprints. Our training sample was too small to fully characterize all spectral varieties of some heterogeneous classes (e.g. grassland, compare Table VIII). Nevertheless, results demonstrated what can be achieved with only one iteration of training, based solely on temporally stable and randomly selected training samples. In the future, training data automation will certainly further improve the classification accuracies [23], [48], but was beyond the scope of this paper. As class boundaries are not always easily separable using a single "best" observation, we incorporated spectral variability measures that can be effortlessly derived from the best pixel assessment. Incorporating spectral variability measures led to considerable improvements in mapping accuracies for all classes. Moreover, compositing artifacts resulting from incorporated SLC-off imagery and

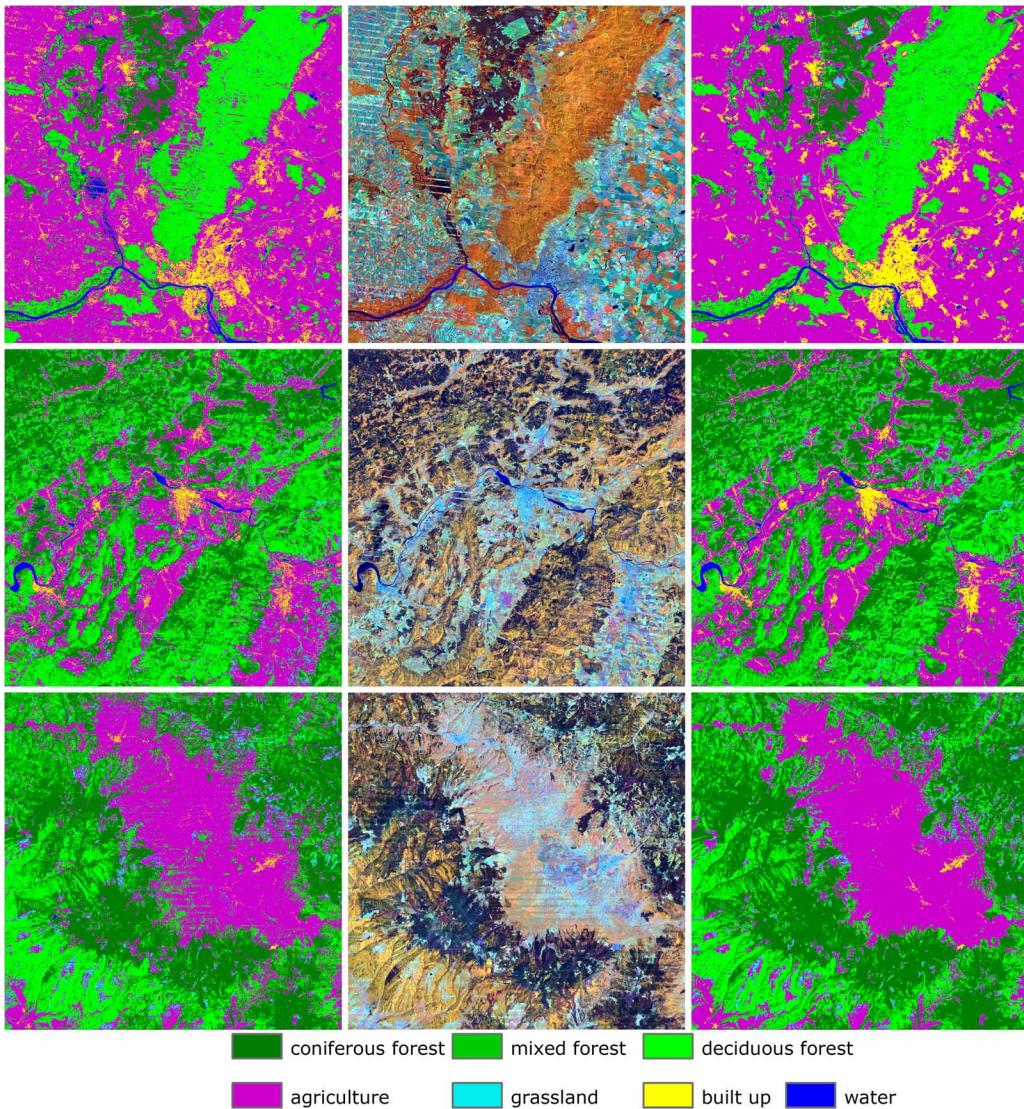


Fig. 5. Classification results (60×60 km) for the best observation composite only (left column) and for the best observation composite plus spectral variability measures (right column). Corresponding regions of the 2005 PBC in $\text{RGB} = 4, 5, 3$ (center column).

salt-and-pepper effects are effectively eliminated when including variability metrics in the classification process, as these artifacts do not translate into averaged spectral measurements (Fig. 5 top). Many classes were delineated with much greater consistency and omission errors were minimized when we provided these additional input features (e.g. built up or mixed forest classes).

Compositing Landsat surface reflectance data provides stable radiometrically consistent data sets across large areas. Such products allow for integrated change classification approaches based on multiple composites, targeted at capturing different points in time. Such a methodology is commonly assumed to outperform other LULCC analysis methods, such as post classification comparison [11]. So far, our methodology exploits the temporal information only partially using the spectral variability measures. More explicit utilization of the temporal context (e.g. seasonally weighted variability measures) from all available unclouded observations per pixel will further enhance information extraction. Future versions of our PBC algorithm will therefore address the potential of different phenology

measures and draw on exploiting the feature space available from multiple composites.

Our approach will profit from data continuity through forthcoming satellite missions such as the Landsat Data Continuity Mission and the Sentinel-2 constellation [12], [27]. Increased image acquisition capabilities and temporal repeat frequencies will provide unique opportunities for cross-sensor compositing and improved process understanding from such integrated data analyses.

ACKNOWLEDGMENT

The authors acknowledge support through the USGS-NASA Landsat Science Team. T. Kuemmerle gratefully acknowledges support through the Einstein Foundation (Berlin, Germany) and the European Union (VOLANTE, FP7-ENV- 2010-265104). The authors would like to thank M. Baumann, B. Jakimow, A. Rabe, D. Pflugmacher and V. C. Radeloff for insightful discussions, and also G. Aronson for support with the data pre-processing.

REFERENCES

- [1] O. Arino, P. Bicheron, F. Achard, J. Latham, R. Witt, and J.-L. Weber, "GLOBCOVER—The most detailed portrait of Earth," *European Space Agency ESA Bulletin*, pp. 24–31, 2008.
- [2] T. Arvidson, S. Goward, J. Gasch, and D. Williams, "Landsat-7 long-term acquisition plan: Development and validation," *Photogramm. Eng. Remote Sens.*, vol. 72, pp. 1137–1146, 2006.
- [3] E. Bartholome and A. S. Belward, "GLC2000: A new approach to global land cover mapping from Earth observation data," *Int. J. Remote Sens.*, vol. 26, pp. 1959–1977, 2005.
- [4] M. Baumann, M. Ozdogan, T. Kuemmerle, K. J. Wendland, E. Esipova, and V. C. Radeloff, "Using the Landsat record to detect forest-cover changes during and after the collapse of the Soviet Union in the temperate zone of European Russia," *Remote Sens. Environ.*, vol. 124, pp. 174–184, 2012.
- [5] M. Broich, M. C. Hansen, P. Potapov, B. Adusei, E. Lindquist, and S. V. Stehman, "Time-series analysis of multi-resolution optical imagery for quantifying forest cover loss in Sumatra and Kalimantan, Indonesia," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 13, pp. 277–291, 2011.
- [6] P. Caccetta, S. Furby, J. O'Connell, J. Wallace, and X. Wu, "Continental monitoring: 34 years of land cover change using Landsat imagery," in *Proc. 32nd Int. Symp. Remote Sens. Environ.*, Costa Rica, 2007.
- [7] G. Chander, X. Xiong, T. Choi, and A. Angal, "Monitoring on-orbit calibration stability of the Terra MODIS and Landsat 7 ETM+ sensors using pseudo-invariant test sites," *Remote Sens. Environ.*, vol. 114, pp. 925–939, 2010.
- [8] B. J. Choudhury, N. E. Digirolamo, and T. J. Dorman, "A comparison of reflectances and vegetation indices from three methods of compositing the AVHRR-GAC data over Northern Africa," *Remote Sens. Rev.*, vol. 10, pp. 245–263, 1994.
- [9] J. Cihlar, "Land cover mapping of large areas from satellites: Status and research priorities," *Int. J. Remote Sens.*, vol. 21, pp. 1093–1114, 2000.
- [10] W. B. Cohen and S. N. Goward, "Landsat's role in ecological applications of remote sensing," *Bioscience*, vol. 54, pp. 535–545, 2004.
- [11] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: A review," *Int. J. Remote Sens.*, vol. 25, pp. 1565–1596, 2004.
- [12] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese, and P. Bargellini, "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.*, 2012.
- [13] J. Feranec, G. Hazeu, S. Christensen, and G. Jaffrain, "Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia)," *Land Use Policy*, vol. 24, pp. 234–247, 2007.
- [14] M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang, "MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote Sens. Environ.*, vol. 114, pp. 168–182, 2010.
- [15] S. Fritz, L. See, I. McCallum, C. Schill, M. Obersteiner, M. van der Velde, H. Boettcher, P. Havlik, and F. Achard, "Highlighting continued uncertainty in global land cover maps for the user community," *Environ. Res. Lett.*, vol. 6, 2011.
- [16] S. Furby, P. A. Caccetta, J. Wallace, X. Wu, J. O'Connell, J. Chia, S. Collings, A. Traylen, and D. Devereux, "Continental scale land cover change monitoring in Australia using Landsat imagery," in *Studying, Modelling and Sense Making of Planet Earth*, Mytilene, Lesvos, Greece, 2008a.
- [17] S. L. Furby, P. Caccetta, J. F. Wallace, X. Wu, J. O'Connell, S. Collings, A. Traylen, and D. Devereux, "Recent Developments in Landsat-Based Continental Scale Land Cover Change Monitoring in Australia," in *XXI Congr. Int. Soc. Photogrammetry and Remote Sensing*, Beijing, China, 2008b.
- [18] M. C. Hansen, R. S. DeFries, J. R. G. Townshend, R. Sohlberg, C. Dimiceli, and M. Carroll, "Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data," *Remote Sens. Environ.*, vol. 83, pp. 303–319, 2002.
- [19] M. C. Hansen and T. R. Loveland, "A review of large area monitoring of land cover change using Landsat data," *Remote Sens. Environ.*, 2012.
- [20] Y. Heymann, C. Steenmans, G. Croisille, and M. Bossard, "CORINE land cover," in *Technical Guide*. Luxembourg: Office for Official Publications of the European Union, 1994, p. 137.
- [21] C. Homer, J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. VanDriel, and J. Wickham, "Completion of the 2001 National Land Cover Database for the conterminous United States," *Photogramm. Eng. Remote Sens.*, vol. 73, pp. 337–341, 2007.
- [22] C. Huang, S. N. Coward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann, "An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks," *Remote Sens. Environ.*, vol. 114, pp. 183–198, 2010.
- [23] C. Huang, K. Song, S. Kim, J. R. G. Townshend, P. Davis, J. G. Masek, and S. N. Goward, "Use of a dark object concept and support vector machines to automate forest cover change analysis," *Remote Sens. Environ.*, vol. 112, pp. 970–985, 2008.
- [24] C. Q. Huang, J. R. G. Townshend, S. L. Liang, S. N. V. Kalluri, and R. S. DeFries, "Impact of sensor's point spread function on land cover characterization: Assessment and deconvolution," *Remote Sens. Environ.*, vol. 80, pp. 203–212, 2002.
- [25] T. Hwang, C. Song, P. V. Bolstad, and L. E. Band, "Downscaling real-time vegetation dynamics by fusing multi-temporal MODIS and Landsat NDVI in topographically complex terrain," *Remote Sens. Environ.*, vol. 115, pp. 2499–2512, 2011.
- [26] R. R. Irish, J. L. Barker, S. N. Goward, and T. Arvidson, "Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm," *Photogramm. Eng. Remote Sens.*, vol. 72, pp. 1179–1188, 2006.
- [27] J. R. Irons, J. L. Dwyer, and J. A. Barsi, "The next Landsat satellite: The Landsat Data Continuity Mission," *Remote Sens. Environ.*, 2012.
- [28] J. Ju and D. P. Roy, "The availability of cloud-free Landsat ETM plus data over the conterminous United States and globally," *Remote Sens. Environ.*, vol. 112, pp. 1196–1211, 2008.
- [29] R. E. Kennedy, Z. Yang, and W. B. Cohen, "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms," *Remote Sens. Environ.*, vol. 114, pp. 2897–2910, 2010.
- [30] J. Knorn, A. Rabe, V. C. Radeloff, T. Kuemmerle, J. Kozak, and P. Hostert, "Land cover mapping of large areas using chain classification of neighboring Landsat satellite images," *Remote Sens. Environ.*, vol. 113, pp. 957–964, 2009.
- [31] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. 14th Int. Joint Conf. Artificial Intelligence*, Montreal, Quebec, Canada, 1995, vol. 2, pp. 1137–1143, Morgan Kaufmann Publishers Inc..
- [32] *Land Use and Land Cover Change. Local Processes and Global Impacts*, E. F. Lambin and H. J. Geist, Eds. Berlin, Heidelberg, New York: Springer Verlag, 2006.
- [33] C. A. Lee, S. D. Gasster, A. Plaza, C.-I. Chang, and B. Huang, "Recent developments in high performance computing for remote sensing: A review," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 4, pp. 508–527, 2011.
- [34] T. R. Loveland and J. L. Dwyer, "Landsat: Building a strong future," *Remote Sens. Environ.*, 2012.
- [35] W. Lucht, C. B. Schaaf, and A. H. Strahler, "An algorithm for the retrieval of albedo from space using semiempirical BRDF models," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, pp. 977–998, 2000.
- [36] J. G. Masek, E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, J. Kutler, and T. K. Lim, "A Landsat surface reflectance dataset for North America, 1990–2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, pp. 68–72, 2006.
- [37] I. Olthof, C. Butson, and R. Fraser, "Signature extension through space for northern landcover classification: A comparison of radiometric correction methods," *Remote Sens. Environ.*, vol. 95, pp. 290–302, 2005.
- [38] I. Olthof, R. Latifovic, and D. Pouliot, "Development of a circa 2000 land cover map of northern Canada at 30 m resolution from Landsat," *Can. J. Remote Sens.*, vol. 35, pp. 152–165, 2009.
- [39] M. Pax-Lenney, C. E. Woodecock, S. A. Macomber, S. Gopal, and C. Song, "Forest mapping with a generalized classifier and Landsat TM data," *Remote Sens. Environ.*, vol. 77, pp. 241–250, 2001.
- [40] A. Pekkarinen, L. Reithmaier, and P. Strobl, "Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, pp. 171–183, 2009.
- [41] A. Plaza, Q. Du, Y.-L. Chang, and R. L. King, "High performance computing for hyperspectral remote sensing," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 4, pp. 528–544, 2011.
- [42] P. Potapov, S. Turubanova, and M. C. Hansen, "Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia," *Remote Sens. Environ.*, vol. 115, pp. 548–561, 2011.

- [43] P. V. Potapov, S. A. Turubanova, M. C. Hansen, B. Adusei, M. Broich, A. Altstatt, L. Mane, and C. O. Justice, "Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data," *Remote Sens. Environ.*, 2012.
- [44] J. A. Richards, "Analysis of remotely sensed data: The formative decades and the future," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, pp. 422–432, 2005.
- [45] T. Rotzer and F. M. Chmielewski, "Phenological maps of Europe," *Climate Research*, vol. 18, pp. 249–257, 2001.
- [46] D. P. Roy, J. Ju, K. Kline, P. L. Scaramuzza, V. Kovalskyy, M. Hansen, T. R. Loveland, E. Vermote, and C. Zhang, "Web-enabled Landsat Data (WELD): Landsat ETM plus composited mosaics of the conterminous United States," *Remote Sens. Environ.*, vol. 114, pp. 35–49, 2010.
- [47] B. Tan, C. E. Woodcock, J. Hu, P. Zhang, M. Ozdogan, D. Huang, W. Yang, Y. Knyazikhin, and R. B. Myneni, "The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions," *Remote Sens. Environ.*, vol. 105, pp. 98–114, 2006.
- [48] J. R. Townshend, J. G. Masek, C. Huang, E. F. Vermote, F. Gao, S. Chanman, J. O. Sexton, M. Feng, R. Narasimhan, D. Kim, K. Song, D. Song, X. Song, P. Noojipady, B. Tan, M. C. Hansen, M. Li, and R. Wolfe, "Global characterization and monitoring of forest cover using Landsat data: Opportunities and challenges," *Int. J. Digital Earth*, 2012.
- [49] E. Vermote, E. Tanre, J. L. Deuze, M. Herman, and J. J. Morcrette, "Second simulation of the satellite signal in the solar spectrum. An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, pp. 675–686, 1997.
- [50] B. Waske, S. van der Linden, J. A. Benediktsson, A. Rabe, and P. Hostert, "Sensitivity of support vector machines to random feature selection in classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, pp. 2880–2889, 2010.
- [51] B. Waske, S. van der Linden, C. Oldenburg, B. Jakimow, A. Rabe, and P. Hostert, "imageRF—A user-oriented implementation for remote sensing image analysis with Random Forests," *Environmental Modelling & Software*, vol. 35, pp. 192–193, 2012.
- [52] R. E. Wolfe, D. P. Roy, and E. Vermote, "MODIS land data storage, gridding, and compositing methodology: Level 2 grid," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, pp. 1324–1338, 1998.
- [53] C. E. Woodcock, S. A. Macomber, M. Pax-Lenney, and W. B. Cohen, "Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors," *Remote Sens. Environ.*, vol. 78, pp. 194–203, 2001.
- [54] M. A. Wulder, J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock, "Opening the archive: How free data has enabled the science and monitoring promise of Landsat," *Remote Sens. Environ.*, 2012.
- [55] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, 2012.



Patrick Griffiths received a Diploma degree in physical Geography from Humboldt University zu Berlin in 2008. He is a research assistant and Ph.D. candidate in remote sensing at the Geomatics Lab, Humboldt University.

His current research interests include automated image processing approaches for land change analysis, long-term remote sensing data archives, time-series analysis and multi-sensor analysis approaches.



Sebastian van der Linden received the M.Sc. degree in Applied Environmental Sciences from the University of Trier, Germany, in 2002 and the Ph.D. degree from Humboldt-Universität zu Berlin, Germany, in 2008.

He is a Senior Scientist at the Geomatics Laboratory of Humboldt-Universität zu Berlin, Germany, and is Managing Director of the University's Integrative Research Institute on Transformations of Human-Environment Systems-IRI THESys (www.iri-thesys.org). From 1999 to 2000, he studied at the University of Edinburgh, U.K. In 2006, he spent three months as a Visiting Scholar with the Lamont-Doherty Earth Observatory, Columbia University, New York, USA.

Dr. van der Linden's scientific interests include the development, implementation and application of machine learning based approaches for qualitative and quantitative analysis of imaging spectroscopy data. His current research is closely linked to the German hyperspectral satellite mission *EnMAP* and he coordinates the development of the *EnMAP-Box*. He is author and co-author of more than 18 journal publications and book chapters ranging from methodological to applied remote sensing.

Dr. van der Linden is a Reviewer for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS and IEEE Journal on Selected Topics of Applied Remote Sensing. He was selected one of IEEE JSTARS' Reviewers of the Year for 2011.



Tobias Kuemmerle received the Ph.D. degree in Geography from Humboldt-University Berlin.

He is Professor for Biogeography and Conservation Biology at the Geography Department of Humboldt-University Berlin. His research focuses on understanding dynamics in land systems, on exploring how land use and climate change affect biodiversity patterns, and on assessing how to balance human resource use and biodiversity conservation. Remote sensing imagery, especially for mapping land use and land cover change across broad geographic extents, plays a crucial role in many of his research projects.



Patrick Hostert received his first degree in physical geography from Trier University, Germany, in 1994, the M.Sc. degree in GIS from Edinburgh University, UK, in 1995, and the Ph.D. degree from Trier University, Germany, in 2001.

He was Assistant Professor at Humboldt University Berlin, Germany, from 2002 to 2006, and holds a full professorship in Geomatics with a focus on Remote Sensing at Humboldt University Berlin, Germany, since 2006. He is founder and head of the Geomatics Lab and since 2008 director/vice-director of the Geography Department at Humboldt University Berlin. Dr. Hostert is member of the scientific advisory group of the Environmental Monitoring and Analysis Program—Germany's hyperspectral EnMAP mission. Since 2013, he is a member of the Landsat Science Team. His research in remote sensing focuses on large area mapping and monitoring, time series analysis of remote sensing data, and hyperspectral image analysis.