First Results From the Phenology-Based Synthesis Classifier Using Landsat 8 Imagery

D. Simonetti, E. Simonetti, Z. Szantoi, A. Lupi, and H. D. Eva

Abstract—A fully automatic phenology-based synthesis (PBS) classification algorithm was developed to map land cover based on medium spatial resolution satellite data using the Google Earth Engine cloud computing platform. Vegetation seasonality, particularly in the tropical dry regions, can lead conventional algorithms based on a single date image classification to "misclassify" land cover types, as the selected date might reflect only a particular stage of the natural phenological cycle. The PBS classifier operates with occurrence rules applied to a selection of single date image classifications of the study area to assign the most appropriate land cover class. Since the launch of Landsat 8 in 2013, it has been possible to acquire imagery at any point on the Earth every 16 days with exceptional radiometric quality. The relatively high global acquisition frequency and the open data policy allow near-real-time land cover mapping and monitoring with automated tools such as the PBS classifier. We mapped four protected areas and their 20-km buffer zones from different ecoregions in Sub-Saharan Africa using the PBS classifier to present its first results. Accuracy assessment was carried out through a visual interpretation of very high resolution images using a Web geographic information system interface. The combined overall accuracy was over 90%, which demonstrates the potential of the classifier and the power of cloud computing in geospatial sciences.

Index Terms—Google Earth Engine (GEE), image classification, land cover mapping, Landsat 8, phenology.

I. INTRODUCTION

SINGLE date classification (SDC) accuracy is often related to the acquisition date of the available imagery, but this involves the uncertainty of climate conditions and their effects on vegetation seasonality. Thus, high revisit frequency such as NASA's Moderate Resolution Imaging Spectroradiometer payload can help to generate time series to monitor abrupt land cover changes and dynamics [1]. However, its coarse ground spatial resolution (250 m) limits the applicability to discriminate different land covers [2], particularly in fragmented landscapes. As a result, smaller vegetation patches, illegal logging, and small-scale agricultural expansion may remain undetected [3].

Landsat imagery [Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+)] have also been used for time-series generation, focusing mainly on forest cover trends [2], [4] and, more recently, on Landsat 8 Operational Land Imager (OLI) for water quality changes [5]. As the Long-Term Acquisition Plan

(LTAP) has improved the number of exploitable acquisitions available, imagery from the latest OLI sensor can be collected and analyzed as time series for land cover/change mapping. This is an important step toward regional and global automated medium spatial resolution mapping and monitoring.

Vegetation seasonality affects both intra- and interannual scenes but is mainly evident in the growing (leaves on) and senescence (leaves off) periods. These processes alter the spectral values of the recorded imagery, particularly for herbaceous and deciduous vegetation [6]. Thus, land cover mapping or change studies based on single date imageries are affected and highly sensitive to the seasonality problem. However, when multiple observations are available within a season cycle, Landsat data can play a significant role, as its spatial resolution allows for the detection of smaller scale patches, changes, and fragmentation, which are prevalent in Africa and Asia [7]. However, to estimate the appropriate timescale to collect and analyze such imagery, the vegetation pattern and dynamics of a particular area have to be either well known or derived from historical Earth observation data.

Maps generated using phenological information (as the proposed phenology-based synthesis (PBS) classifier) are better at avoiding frequent misclassification of deciduous tree cover or herbaceous vegetation as "bare soil" during dry season and could help in monitoring various ecosystem processes such as habitat fragmentation, ecotonal habitats, carbon sequestration, succession, or patterns of biodiversity [6].

This letter demonstrates the PBS algorithm and its applicability in wall-to-wall land cover mapping using a geospatial cloud computing service (Google Earth Engine (GEE)— https://earthengine.google.org). The proposed methodology has been tested in different ecoregions in Sub-Saharan Africa using Landsat 8 data.

II. DATA AND METHODOLOGY

A. Landsat Data Set

The high temporal frequency (∼16 days), combined with the improved LTAP acquisition plan [8] and the distribution policy of the Landsat data, makes the data ideal for land cover mapping and change monitoring applications. GEE was utilized to process Landsat 8 data over several countries, with particular emphasis on Sub-Saharan Africa's semi-dry regions, as illustrated in Fig. 1. Although any area and size of interest could be easily mapped, assessing the map accuracy is still a demanding task due to the lack of up-to-date and globally distributed reference ground control points. Given such a limitation, we have selected four protected areas (PAs) and 20-km buffer zones to test and evaluate the performance of the PBS algorithm. As the selected PAs were located in different
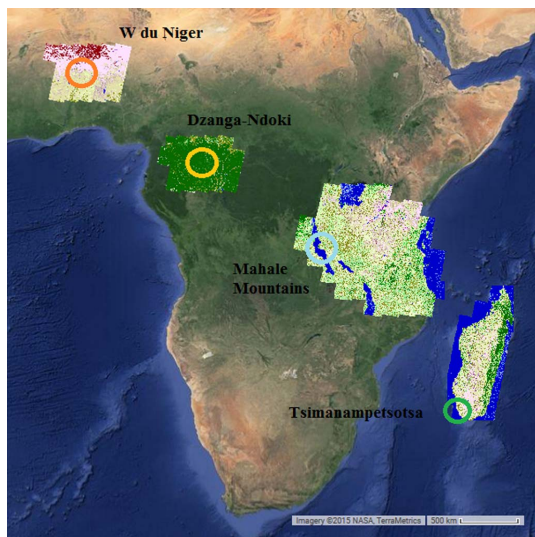
Fig. 1. PBS land cover maps of selected African countries (or part of) underlying an ideal transect crossing diverse biomes from northwest to southeast. Circles represent the location of the four selected PAs.
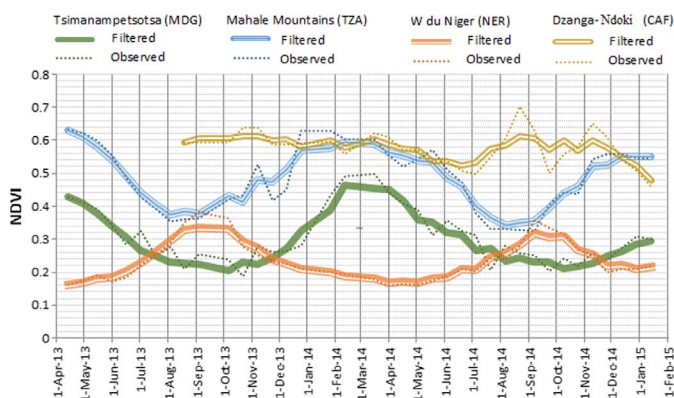

Fig. 2. Average NDVI trends based on Landsat 8 imagery for the investigated PAs. Dzanga-Ndoki National Park has no valid observations before August 2013 (100% clouds or missing on GEE archive).

ecoregions (see Fig. 1), their seasonal and vegetation patterns differed considerably (see Fig. 2). The peak and the end of the growing seasons are automatically extracted from the available Landsat 8 imagery (from 2013 onward) based on the averaged Normalized Difference Vegetation Index's (NDVI) profiles of the study areas to identify the most feasible set of imagery for the actual PBS algorithm. This fully automatic and self-adapting image selection procedure is a very important step in dry and semi-dry regions where the length of the growing season can fluctuate depending on the present weather patterns.

The four PAs were selected (see Table I) on the basis of location, very high spatial resolution satellite, ground data availability, and the need for up-to-date land cover mapping through the BIOPAMA project (http://biopama.jrc.ec.europa. eu/). Additionally, 1) the large land area of Park W du Niger will test the robustness of the algorithm using four adjacent scenes; and 2) Dzanga-Ndoki, characterized by persistent cloud coverage (only three cloud-free images over nine months), will force PBS to work with few cloud-free observations testing, at the same time, the effectiveness of the built-in cloud detection algorithm.

TABLE I

SELECTED PAs WITH WORLD DATABASE ON PAs' IDENTIFICATION NUMBER; THE ECOREGION/MAJOR HABITAT TYPE AS CLASSIFIED BY THE WORLD WILD FUND; LANDSAT PATH AND ROW COVERING THE AREAS, IMAGE ACQUISITION TIME RANGE, AND CLOUD COVERAGE DISTRIBUTION

Protected Area (WDPA ID) Surface	Ecoregions/Major Habitat Type	Landsat Path Row	Image Date From / To	% Cloud Coverage
Tanzania: **Mahale Mountain National Park** (7521) 1613 Km^2	Albertine Rift Montane Forests/Tropical and Subtropical Moist Broadleaf Forests	172 64	01-04-2013 08-09-2013 8 Images	
Central African Republic: **Dzanga-Ndoki National Park** (31458) 495 Km^2	Northwestern Congolian Lowland Forests / Tropical and Subtropical Moist Broadleaf Forests	182 58	01-11-2013 01-08-2014 12 Images (3 cloud free)	
Madagascar: **Tsimanampetsotsa National Park** (2307) 458 Km^2	Madagascar Spiny Thickets / Deserts and Xeric Shrublands	160 77	01-04-2013 01-10-2013 11 Images	
Niger : **W du Niger National Park** (818) 2200 Km^2	West Sudanian Savanna/Tropical and Subtropical Grasslands, Savannas and Shrublands	192-193 051-052	01-08-2013 01-05-2014 61 Images	

% Cloud Coverage legend: 0-10 %, 10-20 %, 20-30 %, 30-40 %, 40-50 %, 50-60 %, > 60 %

B. Image Preprocessing

Selected Landsat 8 images were preprocessed within the GEE environment converting digital number (DN) values to top-of-atmosphere (TOA) reflectance, a physical unit that allows putting data from different sensors/platforms onto a common radiometric scale minimizing spectral differences caused by acquisition time, Sun elevation, and Sun–Earth distance. Bands 1 and 9, thermals, and panchromatic were not processed. The TOA correction formula was as follows:

$$\rho_\lambda = \pi * L_\lambda * d^2 / \text{ESUN}_\lambda * \cos\theta_{\text{SZ}} \qquad (1)$$

where

ρ_λ TOA reflectance for band λ;
L_λ radiance for band $\lambda = M_\lambda * Q_{\text{cal}} + A_\lambda$;
M_λ band-specific multiplicative rescaling factor;
A_λ band-specific additive rescaling factor;
Q_{cal} quantized and calibrated standard product pixel values (DN in 16 bits);
d $(1 - 0.01672 * \cos(0.01745 * (0.9856 * (\text{Julian Day Image} - 4)))$;
θ_{SZ} local solar zenith angle;
ESUN   [2067.000, 1893.000, 1603.000, 972.6, 245.0, 79.72].

Reprojection to common reference system, scene mosaicking, and band stacking are automatically performed by GEE.

C. SDC

The proposed pixel-based SDC algorithm (see Table II), driven by predefined knowledge-based "fuzzy" rules built upon spectral signature collected at global scale, processes the input multispectral image into a discrete thematic layer of 13 classes

TABLE II
PSEUDOCODE OF THE SDC ALGORITHM AS IMPLEMENTED IN GEE; VALUES ARE IN TOA REFLECTANCE, MODIFIED FROM [9] AND [10]

1st step categorization based on NDVI	2nd step

[-1,0] CAT1	if Snowshape = 1 then **SI** else
	if Watershape and B >= 0.078 and 0.04 <=G <=0.12 and
	        max(SWIR1,SWIR2) <= 0.04 then **DWAT** else
	if R >= max(NIR,SWIR1,SWIR2) and 0.04 <=R <= 0.19 and
	        B > 0.078 and max(SWIR1,SWIR2) < 0.04 then **SWAT** else
	if B > 0.94 and G > 0.94 and R > 0.94 and NIR > 0.94 then **CL** else
	if Wetness > 5 then **DWAT** else **SS**

[0,0.45] CAT2	if Snowshape = 1 then **SI** else
	if B > 0.94 and G > 0.94 and R > 0.94 and NIR > 0.94 then **CL** else
	if Other_watershape and B > 0.078 and max(SWIR1 , SWIR2) < 0.058 then **SWAT** else
	if Cloudshape or Cloudshape1 or Cloudshape2 then **CL** else
	if B > G > R and NIR > 0.254 and B > 0.165 and NDVI < 0.40) then **CL** else
	if (B > G) and B > 0.27 and G > 0.21 and ABS(R-G) <= 0.1 and NIR > 0.35 then **CL** else
	if 0.13 > B > G > R < 0.05 and (B-NIR) < -0.04 then **SV** else
	if Wetness > 5 then **DWAT** else
	if 0.13 > B > G > R < 0.05 and B-NIR < 0.04 then **SV** else
	if (ABS(NIR-G) < 0.01 or (B-NIR) < 0.01) and B>G >R and NIR>0.06 then **SS** else
	if NDVI <= 0.09 and NIR < 0.4 and G <= R <= NIR then **OLD2** else
	if NDVI <= 0.20 and NIR > 0.3 and B <= G <= R <= NIR then **OLD1** else
	if NDVI >= 0.35and B >= G and ABS(R-G) < 0.04 then **SPV** else
	if NDVI >= 0.20 and ABS(R-G) < 0.05 then **OLD** else **OLL**

[0.45,1] CAT3	if NDVI < 0.5 and NIR >= 0.15 then **SPV** else
	if NDVI < 0.5 and NIR < 0.15 then **SV** else
	   if NDVI < 0.55 and B <= NIR <= 0.15 then **SV** else
	   if NDVI < 0.55 and B <= NIR then **GRS** else
	   if NDVI < 0.55 and B > NIR then **SV** else
	     if NDVI < 0.65 and NIR >= 0.22 then **TCL** else
	     if NDVI < 0.65 and NIR >= 0.165 then **TCD2** else
	     if NDVI < 0.65 and NIR < 0.165 then **TCD1** else
	     if NDVI < 0.78 and NIR < 0.30 then **TCD** else
	     if NDVI >= 0.78 **SHR** else **TCL**

Where:
B: blue, G: green, R: red, NIR: band 4, SWIR1: band 5; SWIR2: band 7;
Watershape: ((B-G) > -0.2) >= G >= R >= NIR >= SWIR1)
Other_watershape: B >= G >= R <= NIR < R*1.3 < 0.12 > SWIR1 > SWIR2 and 0.039 < NIR < G
NDSI: (G - SWIR1) / (G + SWIR1)
Wetness: 66.96*B + 53.55*G + 23.61*R + 16.72*NIR – 194.53*SWIR1 – 137.19*SWIR2
Brightsoil: (B < 0.27 and Growing15) or (B < 0.27 and Growing14 and SWIR1-SWIR2 > 0.038)
Saturation: max(G,R,NIR)-min(G,R,NIR)/max(G,R,NIR)
Snowshape: min(B,G,R,NIR) > 0.30 and NDSI > 0.65
Cloudshape: min(B,G,R)> 0.17 and max(G,B,R,NIR) > 0.30 and (NIR/R) >= 1.3) and (NIR/G) >= 1.3 and (NIR/SWIR1) >= 0.95 and SWIR1 > min(B,G,R) and NDSI < 0.65
Cloudshape1: max(B,G,R,NIR,SWIR1,SWIR2) > 0.47 and min(B,G,R,NIR) > 0.37
Cloudshape2: min(B,G,R) > 0.21 and SWIR1 > min(B,G,R) and 0.2 <= Saturation <= 0.4 and max(G,R,NIR) >= 0.35 and NDSI > -0.
Growing14: B<G<R<NIR
Growing15: B<G<R<NIR<SWIR1

TABLE III
THEMATIC CLASSES ID, DESCRIPTION, COLOR, AND THE ASSOCIATED LAND COVER AS OBSERVED AT THE DATE OF THE IMAGE ACQUISITION

Class ID	Thematic Classes	Associated Land Cover
ND	No Data	-
WAT (DWAT/SWAT)	Water	Deep Water bodies/Rivers
CL	Clouds	Clouds
SI	Snow	Snow / Ice
TCD	Tree Cover Dark	Dense Forest/Dense Shrub
TCL	Tree Cover Light	Open Forest/Shrub
SHR	Shrub	Dense Shrub
GRS	Grassland	Dense Grassland/ Open Shrub
SPV	Sparse vegetation	Sparse Grassland/Sparse Shrub
OLL	Other Land Light	Light soil/rocks/sand
OLD	Other Land Dark	Dark soil/rocks/sand
SV	Shadowed Vegetation	Shadowed / Low Illuminated Vegetation
SS	Shadowed Soil	Shadowed Soil / Burnt Areas

(see Table III). Relying on TOA reflectance (float [0–1]), these rules, originally developed to classify SPOT4/5, AVNIR-2, DMC, and Landsat TM/ETM+ imagery, are fully capable to process Landsat OLI data sets.

Similar SDC algorithm robustness and scalability among the aforementioned sensor have been confirmed by [10] and [11]; however, estimating the individual SDC's accuracy values at larger scale would be extremely time consuming and probably not a meaningful exercise since the algorithm delivers broad thematic categories derived from spectral properties observed during a precise time in the vegetation cycle. Fig. 3 shows SDC's thematic maps for Mahale Mountains; each of them gradually differs from the preceding as the vegetation condition is changing over time, accordingly to the weather patterns/
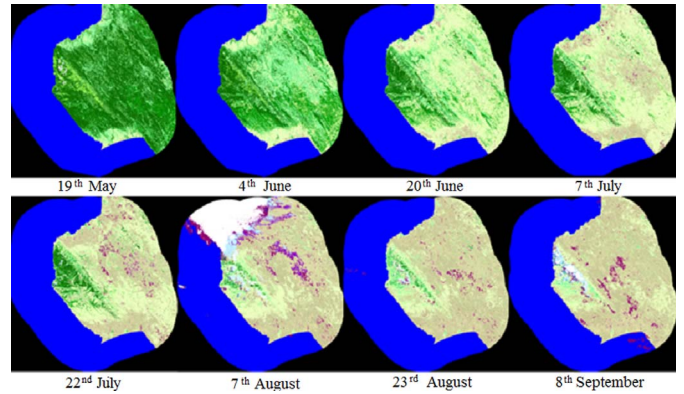

Fig. 3. Sequence of SDC thematic maps for Mahale Mountain National Park. The upper left corner image corresponds to the peak of the growing season, whereas the lower right corner image corresponds to the end of the dry season (2013). See Table III for legend and color.

TABLE IV
FREQUENCY RULES TABLE WITH THE PERCENTAGE OF OBSERVATIONS PER CLASS AS MODIFIED FROM [9]

Classes	ND	CL	SI	WAT	SV	TCD	TCL	SHR	GRS	SPV	OLL	OLD	SS
No data	100	0											
Cloud + No data		100											
Snow/Ice			100										
Permanent Water				100									
Seasonal Water		≥ 0		>60	≥ 0	≥ 0	≥ 0	≥ 0	≥ 0	≥ 0	≥ 0	≥ 0	≥ 0
Evergreen Closed Forest					≥ 0	>80	≥ 0	≥ 0					
Evergreen Open Forest					≥ 0	>40	>40	≥ 0	≥ 0				
Evergreen Shrub					≥ 0	≥ 0	≥ 0	>60	<10				
Close Deciduous Forest					≥ 0		>60		≥ 0	≥ 0	≥ 0	≥ 0	<20
Open Deciduous Forest					≥ 0		>20	>35		≥ 0		<35	
Grassland					≥ 0		<10	≥ 0	>50	≥ 0		<30	
Dark Rocks / Urban										≥ 0	≥ 0	>80	
Dark Soil											≥ 0	>80	≥ 0
Light Soil/ Sand										<20	>80	≥ 0	≥ 0

seasons. While mapping deciduous vegetation, any SDCs might represent the "optimal" observation since the peak of the vegetation cycle is offering information about the greenness and density (shrub or tree cover), and the leafless ones are confirming the presence of deciduous vegetation.

D. PBS

The PBS algorithm combines SDCs into a reliable land cover map based on predefined frequency rules determined on the basis of 1) basic assumptions for "stable" classes such as permanent water/ice or evergreen vegetation and 2) empirical thresholds based on observations collected at the global scale. These rules have been coded (see Table IV) in the algorithm, avoiding any user interaction/training collection while executing the PBS.

Thematic map products (SDC) reduce the data dimensionality and thus help to develop a simple and robust classification model that can perform relatively fast. The "frequency rule" analyzes the thematic class of each pixel in the same geolocation from all available SDC maps looking for a defined pattern during the observed period (see Table IV). Observations such as "No data" (ND) and "Clouds/Shadow" (CL/SH), as detected by the SDC, are excluded from the analysis, except when no other valid observations are available; in this latter case, PBS output will inherit the "ND" or "CL/SH" label accordingly.

Here are two examples of the frequency rule definition as derived from Table IV.

I) Evergreen Closed Forest is characterized by two variables: 1) each SDC has to be in the range [SV, TCD, and TCL]; and 2) at least 80% of the time is TCD.

II) Closed Deciduous Forest's peak of the growing season is represented by the dark forest class (TCD) followed by the brighter (TCL and SHR) classes before the leaf-off period (OLL) and dark other land (OLD) classes.

Additional classes such as seasonal flooded vegetation and snow (not incurred in the study areas) have been omitted from Table IV, although implemented in the PBS considering the dominant land cover class together with at least 15% of either water or snow. The classification schema follows the top-down class order and stops as soon as a condition has been satisfied, avoiding that a pixel fits in two or more classes.

Ideally, PBS requires one image every 16 days (Landsat 8 frequency) from the peak of the growing season until leaf senescence. During particular conditions (short growing season and frequent cloud coverage), some PBS land cover pixels might be based on a single observation; however, this is unlikely as such conditions are quite in contradiction: regions with a short flowering season are generally characterized by low precipitations and, therefore, low cloud coverage. On the other hand, evergreen humid areas might have few observations because of clouds; however, such land cover is not much affected by seasonal variability, and thus, even few observations would deliver correct results. To overcome this issue, a cloud coverage frequency map comes along the PBS as a per-pixel quality control flag. The SDC algorithm might misclassify land cover in individual thematic maps and, thus, lower accuracy values; however, it delivers only 13 broad thematic classes, keeping the inaccuracy values at a minimal level, and the PBS model is robust enough to overpass such misclassifications as demonstrated by the observed final accuracy. Additionally, reducing error propagation [10] introduced by a two-stage model merging the SDC and PBS into a single-stage rule-based algorithm might not be the optimal solution since 1) the need of accurate cloud and shadows masks is crucial and requires SDC; 2) the implementation of the PBS model using 13 thematic layers (variables) is easier to tune and control rather than adopting a model working with series of TOA reflectance and/or nondiscrete indices; and 3) PBS rules are simple and easy to adapt to other specific needs.

E. Land Cover/Use Legend

The land cover legend (see Table V) includes some representative land use classes to allow the correct labeling of the validation data set and to further develop land use discrimination functions within the PBS.

F. Accuracy Assessment

The validation data collection for the evaluation of the PBS algorithm's accuracy has been done using an in-house Web geographic information system interface based on GeoExt and Google Earth API in order to simultaneously display the following: 1) up-to-date very high resolution (VHR) imagery from Google resources (Digital Globe, CNES/Spot and Astrium Image, Rapid-Eye); 2) the set of ground reference points

TABLE V
LAND COVER CLASSES AND COLORS OF THE PBS MAPS

PBS ID	Land Cover Class	PBS ID	Land Cover Class
10	Evergreen Closed Tree Cover (>40%)	15	Tree or Shrub crops
11	Evergreen Open Tree Cover (40% to 15 %)	16	Arable Agriculture
13	Evergreen Shrub Cover	27	Flooded Agriculture
14	Evergreen Herbaceous Cover	17	Flooded Shrub and Herbaceous
23	Deciduous Closed Tree Cover (>40%)	18	Open Water
24	Deciduous Open Tree Cover (40% to 15 %)	19	Urban
25	Deciduous Shrub Cover	20	Bare Areas
26	Deciduous Herbaceous Cover	21	Clouds & Shadows
		22	Snow & Ice

TABLE VI
NUMBER OF GROUND CONTROL POINTS AND OVERALL
ACCURACY VALUES OF THE SELECTED PA

Protected Area	# of GRDS	Overall Accuracy (%)
Mahale Mountain National Park	4404	92.26
Dzanga-Ndoki National Park	3089	98.25
Tsimanampetsotsa National Park	1096	83.75
W du Niger National Park	2098	93.61
Cumulative (4 PAs)	10687	93.37

(GRDS) to be interpreted following a regular grid (2 km × 2 km for Park W and 1 km × 1 km for the others); and 3) available Landsat 8 data. At each confluence point, a circle (60-m radius) was created, and the corresponding dominant land cover/land use class was visually assigned. Too heterogeneous samples (no absolute majority) or samples not covered by VHR images were excluded from the data sets. For ease of interpretation, the experts have the possibility to adjust bands and stretch combination of the Landsat 8 imagery, as well as select the best VHR image using the "Google Time Slider" API's feature. Producer's, user's, and overall accuracy values [12] were calculated based on the collected GRDS.

G. Results

The overall accuracy values for the individual parks (all over 80%), as well as the cumulative (93.37%) presented in Table VI, confirm the robustness of the proposed PBS model in discriminating a variety of land cover types using a limited number of observations and handling possible SDC inaccuracy values or disturbance factors such as clouds or shadows.

However, the cumulative confusion matrix for the four PAs (see Table VII) reveals two main areas of misclassification among 1) *Deciduous Tree Cover* and *Deciduous Shrub Cover*, mainly derived by the uncertainty in estimating the height of the trees (> 5 m) or shrub (< 5 m) when using optical data, and 2) between *Deciduous Closed* and *Open Tree Cover* as observed in the less accurate Tsimanampetsotsa site, where those two classes account for 49.7% of the samples.

Land use classes (1.38% of the samples) not implemented in the PBS lower the general accuracy, with major emphasis on Park W du Niger, where they account for more than 3% of the collected samples, mainly distributed in the northern part as confirmed by the Globcover 2009 (http://due.esrin.esa.int/globcover) map in Fig. 4.

TABLE VII
CUMULATIVE CONFUSION MATRIX FOR THE FOUR PAs. AUTOMATIC CLASSIFICATION LAND COVER CLASSES IN COLUMNS, GROUND TRUTH OBSERVATIONS IN ROWS. NO OCCURRENCES OF CLASS CLOUD & SHADOWS (21) AND SNOW & ICE (22) HAVE BEEN OBSERVED WITHIN THE STUDY AREAS AND ARE THEREFORE OMITTED FROM THE TABLE

	10	11	13	14	15	16	17	18	19	20	23	24	25	26	27	Total	PA
10	3175		8								2	3				3188	99.59
11	20	42									3	2				67	62.68
13	12	1	37										1			51	72.54
14	8										1		1			10	0
15	1		2													3	0
16			1				1				6	35	32	48		123	0
17							11	2				2	2			17	64.70
18								2394								2394	100
19									1				2	4		7	0
20										3			1			4	75
23	1										543	98	4			646	84.05
24							1				95	1037	110	3		1246	83.22
25										1	20	51	2359	63		2494	94.58
26	1	1									5	9	19	378		413	91.52
27	1										3	9	1	1		15	
Total	3219	44	48	0	0	0	13	2396	0	5	678	1247	2531	497	0	10687	
UA	98.63	95.45	77.08	0	0	0	84.61	99.91	0	60	80.08	83.16	93.20	76.05	0		93.37

Fig. 4. PBS, GLC30, and Globcover 2009 land cover maps of the four PAs.

III. CONCLUSION

A PBS has been developed for the Landsat 8 sensor and works in a fully automatic manner without additional input requirements such as supplementary imagery or training data. The PBS implementation in GEE offers a fast and efficient mapping tool applicable at either local or country level. However, assessing map accuracy is still a demanding task due to the lack of up-to-date and globally distributed reference ground control points. Given such a limitation, four study areas have been identified across different ecoregions in order to perform a robust accuracy assessment. The achieved overall accuracy values ($> 90\%$) demonstrate the potential of the PBS in large-scale land cover mapping based on freely available data.

Cross comparison with existing land cover maps (see Fig. 4) highlights the following: 1) the inaccuracy of 30-m products such as GLC30 (http://www.globallandcover.com/GLC30Download/index.aspx) when looking at semi-dry areas as W du Niger or Tsimansmpetsotsa, where the vast majority of deciduous forest and shrub have been classified as grassland; and 2) the recognizable alignment of the general PBS land cover distribution pattern with Globcover 2009 (http://due.esrin.esa.int/globcover), although the 300 mt resolution and the years the latter product refers to do not make it suitable for local and up-to-date mapping studies.

Land cover changes are definitely playing a disturbance role in time-series analysis, but we would like to remark that the proposed PBS model makes use of a limited number of observations mainly concentrated between the peak and the end of the growing season, limiting misclassifications to the generated seasonal map. Indeed, using PBS and GEE is possible to map large geographical areas quickly every year and with high accuracy values, particularly in dynamic environments. PBS rules for land use discrimination such as agriculture and built-up areas are subject of further developments.

The proposed algorithm will be able to handle the upcoming Sentinel 2 [13] data as confirmed by the pretesting of the SPOT4 (Take5) imagery and given its free distribution policy [14].

Further development, example, and code (GEE Javascript/Python and the stand-alone offline Python implementation) will be available at http://bioval.jrc.ec.europa.eu.
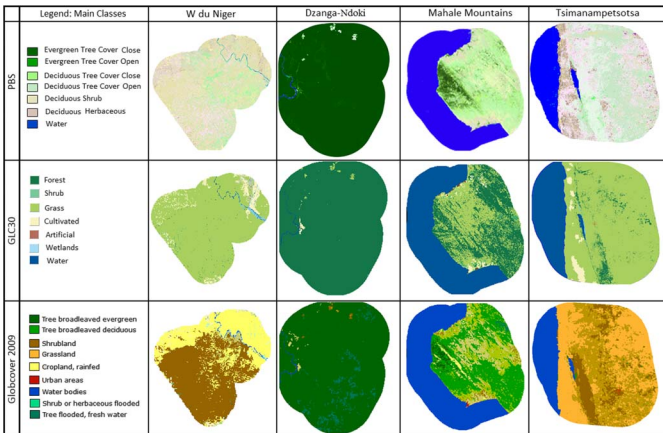
REFERENCES

[1] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sens. Environ.*, vol. 105, no. 2, pp. 142–154, Nov. 2006.

[2] D. J. Hayes and W. B. Cohen, "Spatial, spectral and temporal patterns of tropical forest cover change as observed with multiple scales of optical satellite data," *Remote Sens. Environ.*, vol. 106, no. 1, pp. 1–16, Jan. 2007.

[3] Y. E. Shimabukuro *et al.*, "Deforestation detection in Brazilian Amazon region in a near real time using Terra MODIS daily data," in *Proc. IEEE IGARSS*, 2004, vol. 5, pp. 3405–3408.

[4] E. A. Lehmann, J. F. Wallace, P. A. Caccetta, S. L. Furby, and K. Zdunic, "Forest cover trends from time series Landsat data for the Australian continent," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 21, pp. 453–462, Apr. 2013.

[5] F. L. Lobo, M. P. F. Costa, and E. M. L. M. Novo, "Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities," *Remote Sens. Environ.*, vol. 157, pp. 170–184, Feb. 2015.

[6] J. E. Vogelmann, G. Xian, C. Homer, and B. Tolk, "Monitoring gradual ecosystem change using Landsat time series analyses: Case studies in selected forest and rangeland ecosystems," *Remote Sens. Environ.*, vol. 122, pp. 92–105, Jul. 2012.

[7] E. F. Lambin, H. J. Geist, and E. Lepers, "Dynamics of land-use and land-cover change in tropical regions," *Annu. Rev. Environ. Resour.*, vol. 28, no. 1, pp. 205–241, Nov. 2003.

[8] J. C. White and M. A. Wulder, "The Landsat observation record of Canada: 1972–2012," *Can. J. Remote Sens.*, vol. 39, no. 6, pp. 455–467, Feb. 2014.

[9] E. Simonetti, D. Simonetti, and D. Preatoni, "Phenology-based land cover classification using Landsat 8 time series," Joint Res. Centre, Ispra, Italy, EUR 26841 EN, 2014.

[10] Z. Szantoi and D. Simonetti, "Fast and robust topographic correction method for medium resolution satellite imagery using a stratified approach," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 6, no. 1, pp. 1921–1933, Aug. 2013.

[11] A. Baraldi *et al.*, "Automatic spectral-rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye, and DMC/SPOT-1/-2 imagery—Part I: System design and implementation," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1299–1325, Mar. 2010.

[12] M. Story and R. G. Congalton, "Accuracy assessment: A user's perspective," *Photogramm. Eng. Remote Sens.*, vol. 52, no. 3, pp. 397–399, 1986.

[13] M. Drusch *et al.*, "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, May 2012.

[14] W. Turner *et al.*, "Free and open-access satellite data are key to biodiversity conservation," *Biol. Conserv.*, vol. 182, pp. 173–176, Feb. 2015.