IMPLEMENTING MULTI-SCALE AGRICULTURAL INDICATORS EXPLOITING SENTINELS

PRODUCT USER MANUAL
CROP MAPS

IMAGINES_RP6.3_PUM-CROP MAPS

ISSUE I2.00

EC Proposal Reference N° FP7-311766

Due date of deliverable: June 2015

Actual submission date: April 2016

Duration : 40 months

Start date of project: 01.11.2012

Name of lead partner for this deliverable: UCL

Book Captain: François Waldner (UCL)

Contributing Authors: Pierre Defourny (UCL)

Project co-funded by the European Commission within the Seventh Framework Program (2007-2013)		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

DOCUMENT RELEASE SHEET

Book Captain:	François Waldner	Date: 29.04.2016	Sign.
Approval:	R. Lacaze	Date: 24.05.2016	Sign.
Endorsement:	M. Koleva	Date:	Sign.
Distribution:	Public		

CHANGE RECORD

Issue/Revision	Date	Page(s)	Description of Change	Release
	27.11.2015	All	Initial issue	I1.00
	29.04.2016	All	Updated issue with results for South Africa	I2.00

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ATBD	Algorithm Theoretical Basis Document
CGL	Crop Group Layer
CTL	Crop Type Layer
EU	European Union
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
HRO	High Resolution Optical
GEO	Group on Earth Observation
GEOGLAM	GEO Global Agriculture Monitoring
LAI	Leaf Area Index
LDAS	Land Data Assimilation System
MRO	Medium Resolution Optical
MSAVI2	Modified Soil Adjusted Vegetation Index
NDVI	Normalized Difference Vegetation Index
PCL	Pre-seasonal Cropland Layer
PUM	Product User Manual
SSD	Service Specifications Document
SVP	Service Validation Plan
UCL	Université catholique de Louvain
UTM	Universal Transverse Mercator
LDAS	Land Data Assimilation System

1. BACKGROUND OF THE DOCUMENT

1.1. EXECUTIVE SUMMARY

The Copernicus program is the EU response to the increasing demand for reliable environmental data. The objective of the Copernicus Land Service is to continuously monitor and forecast the status of land territories and to supply reliable geo-information to decision makers, businesses and citizens to define environmental policies and take right actions. ImagineS intends to continue the innovation and development activities to support the operations of the Copernicus Global Land service, preparing the use of the new Earth Observation data, including Sentinels missions data, in an operational context. Moreover, ImagineS aims to favor the emergence of downstream activities dedicated to the monitoring of crop and fodder production, that are key for the implementation of the EU Common Agricultural Policy, of the food security policy, and could contribute to the Global Agricultural Geo-Monitoring Initiative (GEOGLAM) coordinated by the intergovernmental Group on Earth Observations (GEO).

The main objectives of IMAGINES are to (i) improve the retrieval of basic biophysical variables, mainly LAI, FAPAR and the surface albedo, identified as Terrestrial Essential Climate Variables, by merging the information coming from different sensors (PROBA-V and Landsat-8) in view to prepare the use of Sentinel missions data; (ii) develop qualified software able to process multi-sensor data at the global scale on a fully automatic basis; (iii) complement and contribute to the existing or future agricultural services by providing new data streams relying upon an original method to assess the above-ground biomass, based on the assimilation of satellite products in a Land Data Assimilation System (LDAS) in order to monitor the crop/fodder biomass production together with the carbon and water fluxes; (iv) demonstrate the added value of this contribution for a community of users acting at global, European, national, and regional scales.

Thus, ImagineS proposed a new method for mapping crop types along the season by combining information coming from Sentinel-1, -2 and -3 proxy imagery. The objective assigned is to derive reliable (accuracy target set at 85%) crop maps as soon in the season as possible capitalizing on the advantage of each sensor: the temporal coverage of Sentinel-3, the high resolution of Sentinel-2 and the weather independent acquisitions of Sentinel-1.

The Product User Manual (PUM) aims at presenting the characteristics of the crop maps and its format. The product consists in fact in three products: 1) a pre-seasonal cropland map derived at the beginning of the growing season, 2) a winter crop map derived at the end of the winter and 3) a calculated at every acquisition of a high resolution image along the growing season. The products are projected on a regular latitude/longitude grid with spatial

resolution ranging from 250-m to 10-m according to the spatial resolution of the images used. They are provided in single band GEOTIFF format containing the classes to be mapped.

1.2. SCOPE AND OBJECTIVES

The PUM is a self-contained document which gathers all necessary information to use the product on an efficient and reliable way. It gives an overview of the products properties, in terms of algorithm, technical characteristics and main validation results.

1.3. CONTENT OF THE DOCUMENT

The document is structured as follows:

- Chapter 2 summarizes the algorithm.
- Chapter 3 describes the technical characteristics of the products.
- Chapter 4 presents the quality assessment of the products.

1.4. RELATED DOCUMENTS

1.4.1. Inputs

Overview of deliverables acting as inputs to this document.

Document ID	Descriptor
IMAGINES_RP1.2_SSD	Service Specifications Document
IMAGINES_RP4.1_ATBD-CropMaps	Algorithmic Theoretical Basis Document of the Crop Maps

2. ALGORITHM

The algorithm, defined by UCL (Université catholique de Louvain) generates crop maps from proxy imagery for Sentinel-1, Sentinel-2 and Sentinel-3 imagery. These products are known as crop maps. It has to be highlighted that due to data availability constraints during the project lifetime, those proxy images were not evenly distributed in time along the growing season.

The products defined in this user manual cover two study sites. The first one is Tula Oblast (54.2048° N, 37.6185° E), Russia and the second is the Free State Province (28.4541° S, 26.7968° E), South Africa.

2.1. DEFINITION OF THE CLASSES

2.1.1. Classes for the study site in Russia

Table 1, Table 2, and Table 3 present the classes of the products.

Table 1: Pre-seasonal Cropland Layer legend

Class	Numeric value	Hexadecimal color
cropland	14	#eead0e
broadleaved forest	50	#488957
needleleaved forest	70	#556b2f
grassland	100	#dfdfeb
urban	190	#ff0000
water	210	#3399ff

Table 2: Crop Group Layer legend

Class	Numeric value	Hexadecimal color
other crops	14	#3c3636
winter crops	15	#ffa000
broadleaved forest	50	#488957
needleleaved forest	70	#556b2f
grassland	100	#dfdfeb
urban	190	#ff0000
water	210	#3399ff

Table 3: Crop Type Layer legend

Class	Numeric value	Hexadecimal color
buckwheat	1	#f5f5dc
build up	2	#ff0000
forest	3	#008000
grassland	4	#dfdfeb
maize	5	#8a2be2
oats	6	#ffb6c1
potato	7	#333333
rapeseed	8	#794044
spring barley	9	#ffa500
spring wheat	10	#800000
water	11	#3399ff
winter wheat	12	#ffff00

2.1.2. Classes for the study site in South Africa

Table 4,

Table 5 and Table 6 present the classes of the products.

Table 4: Pre-seasonal Cropland Layer legend for the Free State Province (South Africa)

Class	Numeric value	Hexadecimal color
forest	0	#488957
cropland	1	#eead0e
bushland	2	#ACAC6E
shrubland	3	#662200
grassland	5	#dfdfeb
Water body	6	#3399ff
Bare soil	7	#B7AB92
urban	8	#ff0000
mines	9	#DDDDDD
wetland	10	#73A0A0

Table 5: Crop Group Layer legend for the Free State Province (South Africa)

Class	Numeric value	Hexadecimal color
non cropland	0	#B4B4B4
other crops	1	#3C3636
winter crops	2	#FFA000

Table 6: Crop Type Layer legend for the Free State Province (South Africa)

Class	Numeric value	Hexadecimal color
maize	1	#E3BE00
other crops	2	#009E45
planted pastures	3	#95DD7F
natural field	4	#D2E6AB
soybeans	5	#216000
stubble	6	#B8AC93
sunflower	7	#E6E600
weed	8	#E65C5C

2.2. PRE-SEASONAL CROPLAND LAYER ALGORITHM

2.2.1. Outline of the algorithm

At the beginning of the cropping season, a pre-seasonal cropland extent map is produced using a dedicated land cover change algorithm based on metrics of the previous year at moderate resolution. The best local land cover map available helps this early diagnosis. If none is available, a global land cover map (such as GlobCover) is used. Since those land cover map do not target agricultural land and since changes may have occurred, an image-to-map discrepancy has been developed to adjust it to the current regional conditions.

2.2.2. Classification methodology

2.2.2.1. Input data

- Seven or Ten-day mean composite time-series of moderate resolution vegetation index from the previous year [start of season year t-1 → start of season year t]
- Re-sampled and co-registered land cover map

2.2.2.2. Methodology

The core of the method relies on a change detection approach. It seeks to extract class probability distributions of a reference land cover map after removal of potential outliers (Figure 1). The classification considers four time-series features (mean, maximum, minimum, slope between the minimum and the maximum) that were extracted from the time-series themselves.

First, all the mixed classes were considered as outliers and set aside. Second, for each of the remaining land cover class, the selection of outliers relied on a probability threshold α

which specifies the limit at which an observation is considered an outlier. The distribution is iteratively trimmed until no more outliers are identified. For the normal case, the outlier detection criterion gives:
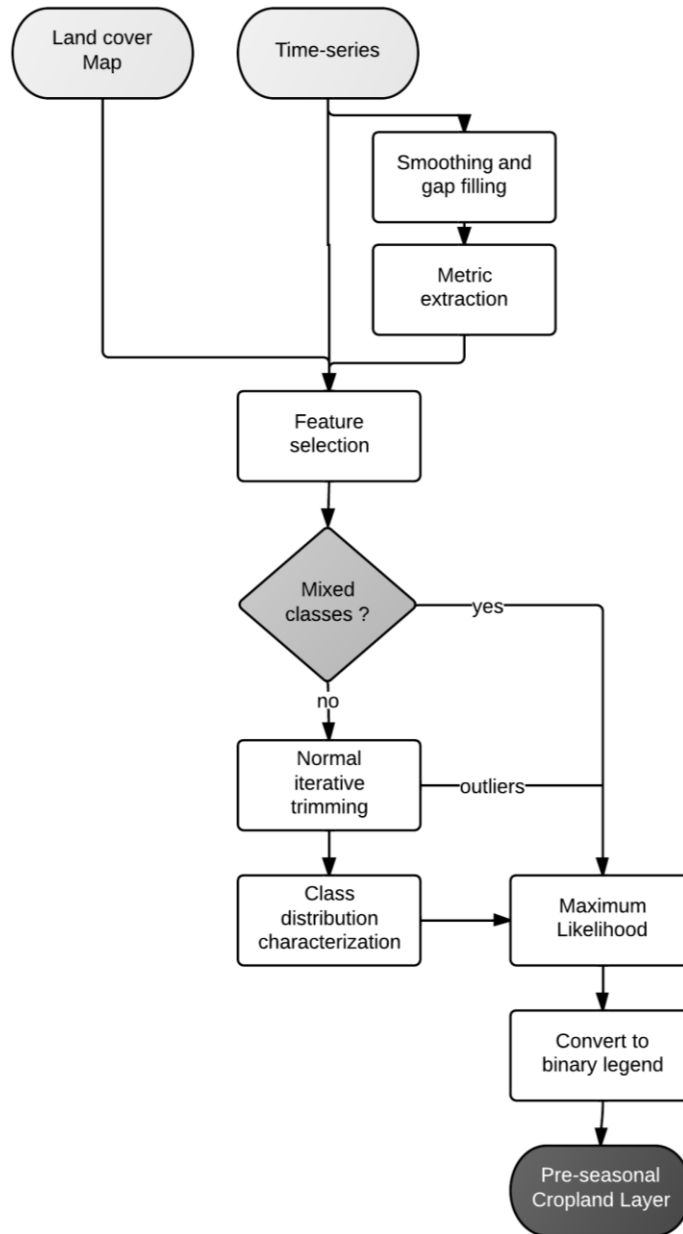

Figure 1: Flowchart of the pre-seasonal cropland layer

$$(x - \mu)'\Sigma^{-1}(x - \mu) \leq \chi_p^2(\alpha)$$

where χ^2 is the upper $(100\alpha)^{th}$ percentile of a χ_p^2 distribution with *p* degrees of freedom. The threshold *α* was set at 0.05. The multivariate normal iterative trimming was applied on each pure thematic of the reference land cover map except the mosaic classes. The outliers were merged to the mosaic class pixels and are subsequently reclassified according to a maximum likelihood decision rule. The output classes were defined by the parameters of the trimmed classes. A priori probabilities were computed from the initial land cover map with mixed classes probabilities redistributed to the closest thematic class. The four metrics were scaled and centered to avoid confusion between the different metric units. The updated land cover map details as many classes as pure classes in the initial legend.

2.3. CROP GROUP LAYER ALGORITHM FOR TULA

2.3.1. Outline of the algorithm

At the end of the winter, a knowledge-based crop group layer distinguishing winter crops from summer crops is delivered thanks to an automated phenological object-based time-series classification of medium resolution data. A first step relies on a harmonic components analysis to spatially group pixels with the same temporal trajectories. As multi-date segmentation is known to perform better than single-date but requires the prior identification of key-dates, segmenting on the harmonic components is an alternative to overcome this constraint. Second, the automated adaptive recognition decision rule relies on the presence or absence of an observable winter growth peak.

2.3.2. Classification methodology

2.3.2.1. Input data

- Ten-day mean composite moderate resolution time-series of vegetation index from the previous year [01-07-year t-1 → 01-03-year t]
- Pre-seasonal cropland layer

2.3.2.2. Methodology

A two-step object-based classification addresses the binary classification of winter crops versus summer crops (Figure 2). In the first step, objects are derived by means of a segmentation on the harmonic components of the time-series. In the second, an automated time-series analysis based on phenological metrics classifies the objects in the two classes. The previously derived cropland masks out the non-agricultural areas.

Figure 2: Flowchart of the crop group layer algorithm.
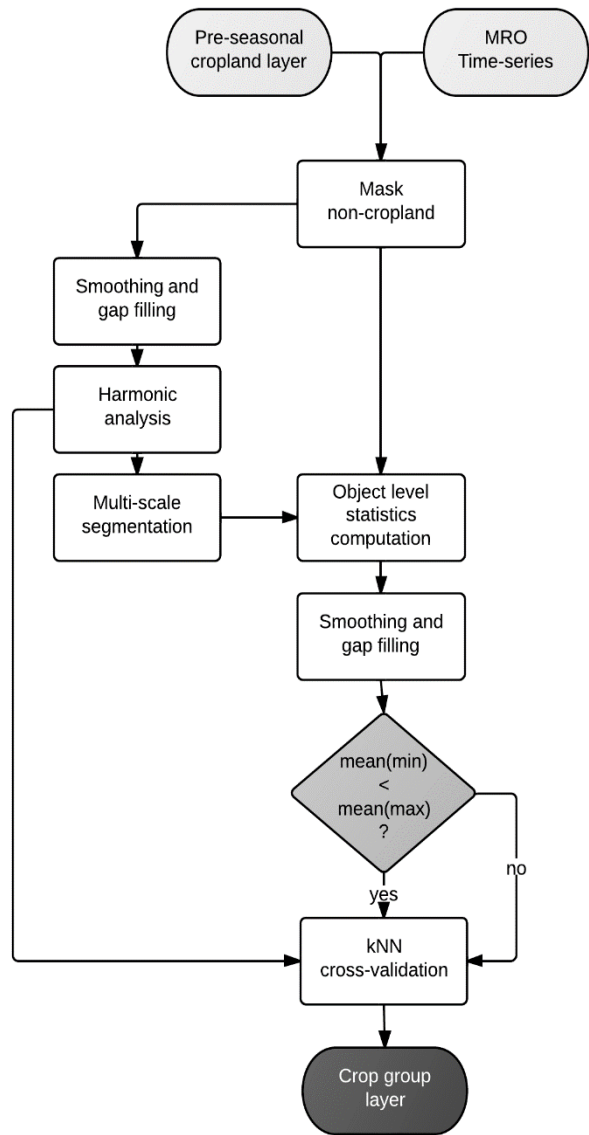
The harmonic analysis was applied pixel-wise. A Fourier transform converts the input time-series into a complex array with a real and an imaginary part. The real (a) and an imaginary (b) part can be converted to polar form. Each order i is defined by its phase φ_i and amplitude a_i:

$$a_i = \sqrt{a_i^2 + b_i^2}$$

$$\varphi_i = \arctan \frac{b_i}{a_i}$$

As the high orders contain mainly noise, only the additive term and the two first harmonic are kept. Similar pixels were then spatially grouped based on the similarity of the two first harmonic and the additive term using the multi-resolution segmentation algorithm. In the second step corresponds to the classification of the objects. A knowledge-based adaptive decision rule has been defined based on a specific temporal pattern of the winter crop: the presence of a winter growth peak. For each object-level time-series, the algorithm extracts three metrics: (i) the snow date corresponds to the dekade during which the MSAVI2 (Modified Soil Adjusted Vegetation Index) signal falls below a threshold, (ii) the maximum of vegetation is the local maximum MSAVI2 value at the winter growth peak and (iii) the local minimum preceding the winter peak. The simplest decision rule would classify as winter crop every feature with a winter peak, but this hypothesis does not hold as other landscape elements also exhibit a maximum. Usually, those elements however are characterized by a rapid drop and then increase of the signal whereas winter crop shows a plateau corresponding to field preparation and sowing. To tighten the rule, the minimum and the maximum are converted into local average of the metric values: an object is classified as winter crop if $mean(max) > mean(min)$. This filters out objects without a significant increase in their respective trajectory. To strengthen the accuracy of the classification, a cross-validation based on the harmonic components is added. The 25-nearest (in Euclidean distance) vectors are identified and the final label is decided by majority vote. The harmonic component values are first standardized and then averaged per object.

2.4. CROP GROUP LAYER ALGORITHM FOR THE FREE STATE

2.4.1. Outline of the algorithm

At the end of the winter season, a knowledge-based crop group layer distinguishing winter crops from un-cropped areas is delivered thanks to an automated phenological object-based time-series classification of medium and high resolution resolution data. In a first step, the amplitude of the NDVI during the winter growing period relies on a harmonic components analysis to spatially group pixels with the same temporal trajectories. As multi-date segmentation is known to perform better than single-date but requires the prior identification of key-dates, segmenting on the harmonic components is an alternative to overcome this constraint. Second, the automated adaptive recognition decision rule relies on the presence or absence of an observable winter growth peak.

2.4.2. Classification methodology

2.4.2.1. Input data

- Seven-day mean composite moderate resolution time-series of vegetation index from the previous year [01-06-year t→ 01-11-year t]
- RapidEye images until 01-11-year t

- Radarsat-2 images until 01-11-year t
- National Land cover 2013-2014

2.4.2.2.   Methodology

Holding to the same approach as for vernalizing areas, a three-step winter crop detection was designed combining an automated detection approach at moderate resolution with an object-oriented approach at high resolution (Figure 3).
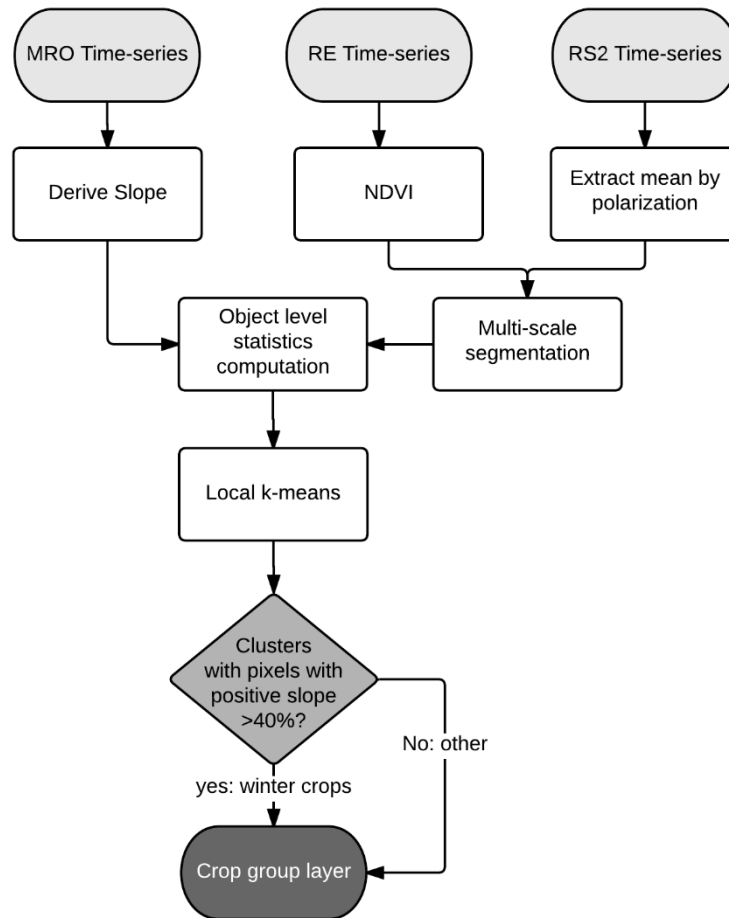
Figure 3: Flowchart for the winter crop classification with the Free State data set

First, the slope between the minimum and the maximumn was derived from the MSAVI2 MODIS time-series. Areas with positive values correspond to winter growing areas whereas slow decreasing values characterize, other land cover types such as winter grazing areas. Second, if using the slope appears as an efficient automated recognition method, the spatial resolution remained to low. To increase the spatial details, a segmentation was carried out on the NDVI high resolution optical images. Average object-level statistics were then

computed – mean σ_{VH}, σ_{VV}, NDVI and slope. Third, an unsupervised clustering was then applied locally (moving windows of 50x50-km) on sigma and NDVI. As setting the number of cluster is always a difficult task but of key importance, the local training allows to reduce the variability within the area to be classified. Clusters having 40% or more of their sample associated with a positive slope were labeled as winter crops. Refine the delineation of winter crop field but also include smaller fields for which the winter growth cycle was not picked up because of MODIS spatial resolution.

2.5. CROP TYPE LAYER ALGORITHM

2.5.1. Outline of the algorithm

Along the growing season, a multi-sensor crop specific classification is finally achieved. The crop type layer is updated as high resolution data acquisition progresses. Crops type maps are produced through an iterative segmentation, classification and fusion scheme of the moderate and high resolution optical and radar images.

2.5.2. Classification methodology

2.5.2.1. Input data

- Seven or Ten day moderate resolution vegetation index time-series [01-07-year t-1 → harvest] (see Figure 4 and Figure 5 for an overview of the temporal distribution)
- Preprocessed, re-sampled and co-registered radar data [start of spring - harvest]
- Preprocessed, RapidEye data [start of spring - harvest]
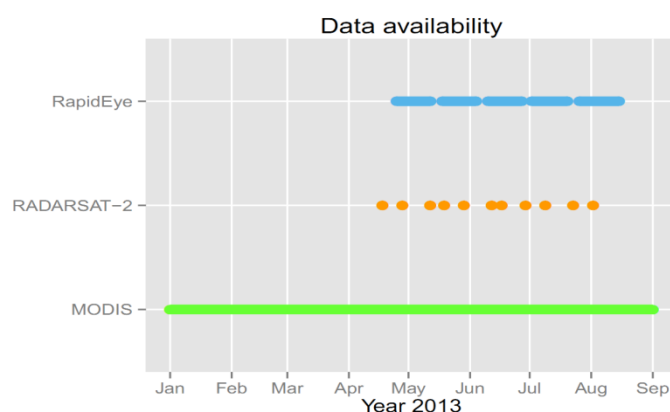- Crop type field observations

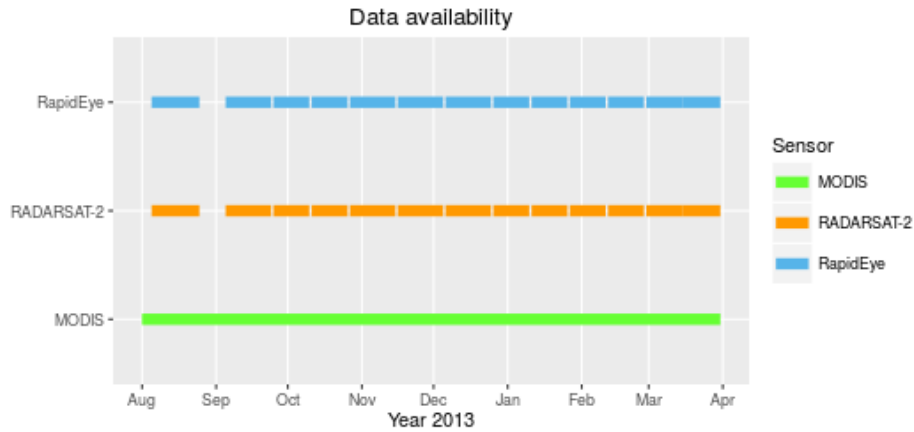**Figure 4: Data acquisition for the Russian site.**

Figure 5: Data acquisition for the South African site.
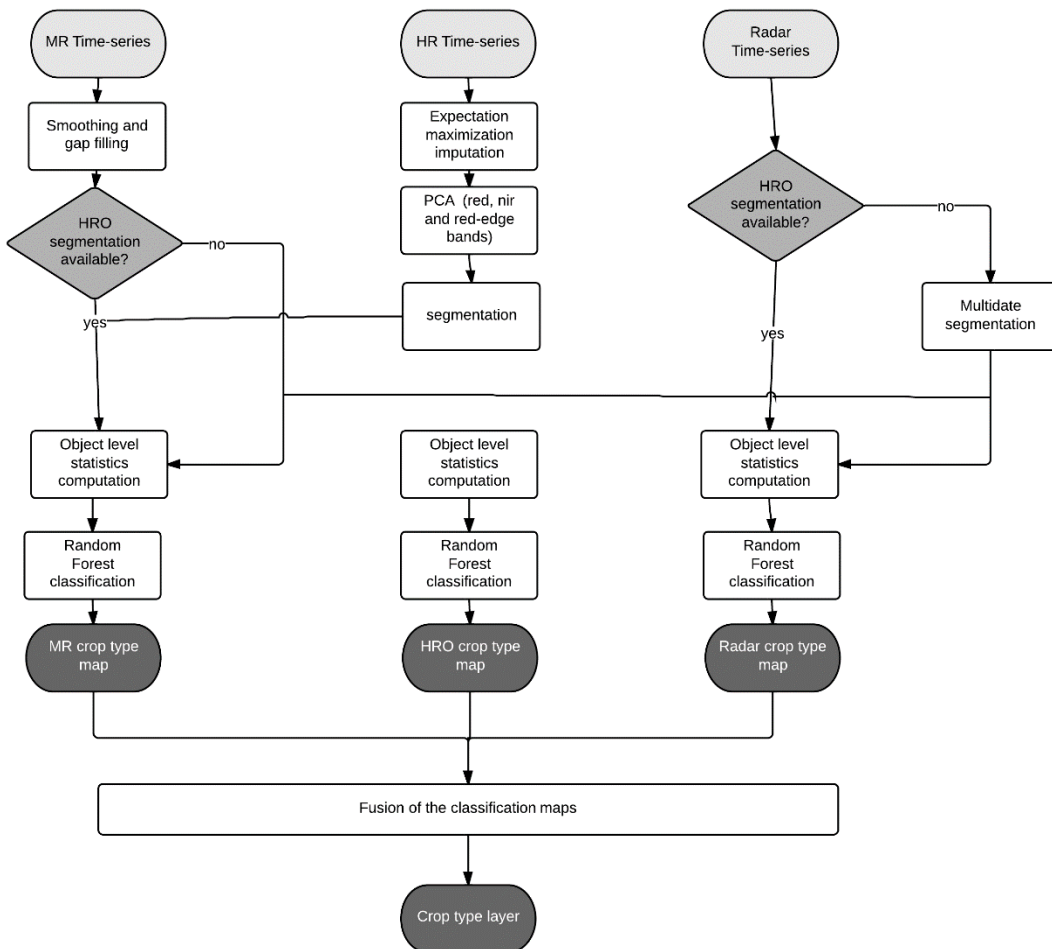
2.5.2.2. Methodology

Figure 6: Flowchart of the crop type layer algorithm.

Three different processing chains have been developed, one per sensor (Figure 6). The only information shared is the spatial delimitation of the objects. Each sensor has its own classifiers and the resulting classifications are fused into the final classification.

Missing values and smoothing

As classifiers cannot handle with missing observations, reconstruction of missing data is a mandatory step. MRO time-series are smoothed and gap-filled by means of the Whittaker filter (Eilers, 2003). For HRO, due to the short length of the HRO time-series (only 5 coverages in Russia), traditional filters for filling the missing values were discarded. Instead, a temporal linear interpolation was run the incomplete HRO to impute them. At the first acquisition of HRO data imputation with linear interpolation is infeasible; missing data are thus imputed thanks to a weighted mean of the k-nearest neighbors identified with the radar data.

Segmentation

The Sentinel-2 like images is the primary source for segmentation because of their higher spatial resolution. The NDVI of each images were stacked and put into a principal component analysis that will reduce the dimension for the segmentation while maximizing the variance of the output components. A multi-resolution segmentation algorithm was then applied on the time stack of NDVI. Besides, if no HRO image is available at the acquisition of the first radar images, the segmentation is applied on the radar data only with its respective parameters.

Objects statistics computation

Object-level statistics are computed using the latest segmentation available. Only average values of the input data are computed, that is: the mean MSAVI2, σ_{VH}, σ_{VV}, and blue, green, red, infra-red and red-edge.

Training of the classifiers

At each acquisition of a high resolution data (either radar or optical), new classifiers are trained to include the new observations. One random forest with 800 trees is trained by data source: one for the MODIS data, one for the Radarsat-2 data and one for the RapidEye data. It should be that several classifiers have been tested and that the random forest was outperforming the others for every sensor. Random forests are insensitive to noise and overtraining. Their internal selection of variables is expected to prevent from the side effects of high dimensionality.

Classification

At each acquisition of a HR image, the classifiers are applied per data type on all available images to produce three sensor-specific crop type maps along with the membership probability of a pixel to each class.

Map fusion

As the availability of the training data might be a constraint and because the availability of calibration data might change from year to year, the approach chosen to fuse the results of the classifiers is based on their soft outputs, i.e., their class membership. The class membership are of the following form:

$$pr(x) = \{pr_1(x), pr_2(x), \ldots, pr_i(x), \ldots, pr_n(x)\}$$

where $pr_i(x)$ is the estimated membership degree of a object x to class i, and n the number of classes. The vectors pr_{MODIS} *(x,)* $pr_{RADARSAT}$ *(x) and* $pr_{RAPIDEYE}$ *(x)* for each object resulting from the three independent classification are then multiplied class-wise and the class that maximizes the class membership is taken as the final output of the classification.

2.6. LIMITATIONS OF THE PRODUCT

The method combines three different kinds of sensor and might be extended to more which reduces the sensor dependence and enhances the probability of acquiring during a good discrimination window. This combination makes the classification method more robust to address the large spectrum of agro-systems of the world.

However, due to constraints in the data acquisition, the multi-sensor time-series of does not cover the entire season, e.g. RapidEye and Radarsat2 images are only available at the beginning of the spring. More importantly, they do not cover the previous season. Therefore, the pre-seasonal cropland layer is extracted at the resolution of MODIS. The jump at 10m later on in the season might thus lead to some inconsistencies. Regarding the pre-seasonal cropland layer, the reference has to be of reasonable accuracy in order to produce accurate results. Indeed, if a class is dominated by mislabeled pixels, the trimming would tend to remove the correctly labeled data because of the spectral distribution shift introduced by the mislabeled pixels. Besides, other features might appear more relevant in other agro-systems and could easily replace the four proposed here. In the case of high cloud coverage during the winter, the identification of the winter growth peak might be less accurate leading to more confusion between winter crops and summer crops. In the full-season case, defining the parameters for the segmentations and the number of clusters for the multi-source crop group recognition would need some fine tuning to be adapted on a new site. Finally for the machine learning step, texture, vegetation indices and temporal features might be considered to increase the accuracy of the recognition. The main constraint for operational and large scale application remains the availability of training data.

3. PRODUCT DESCRIPTION

3.1. FILE FORMAT AND CONTENT

The Pre-seasonal Crop Layer (PCL), Crop Group Layer (CGL) and Crop Type Layer (CTL) products follows the following naming standard:

ImagineS_CropMaps_<Acronym>_<YYYYMMDD>_<AREA>_V<version>

where:

- <Acronym> is the short name of each product i.e. PCL, CGL or CTL

- <YYYYMMDD> gives the temporal location of the file. YYYY, MM, DD denote the year, the month, the day, respectively.

- <AREA> gives the spatial coverage of the file. In our case, <AREA> is either Tula or FreeStateProvince.

- <version > gives the version number of the product.

The Crop Maps products are distributed in zip archives. The archive contains following files:

- A single band geotiff containing the PCL;
- A single band geotiff containing the CGL;
- Multiple single band geotiffs containing the CTL at different dates;
- Three csv files containing the legend of the products (PCL, CGL and CTL) used. The csv have all three columns: attT which provides the class, numT which provides the associated numeric value of a given class and colT which gives the color associated to a class in hexadecimal notation as described in Table 1, Table 2, and Table 3, respectively.

3.2. PRODUCT CHARACTERISTICS

3.2.1. Spatial information.

The products are available for two study sites: the Tula Oblast (Russia) for the year 2013 and the Free State Province (South Africa) for the year 2013. The Tula oblast covers 25 700km² and is centered at 53°55′N 37°35′E. The Free State province covers 129,825 km² and is centered at 28.4541° S, 26.7968° E.

3.2.2. Projection and Grid Information

3.2.2.1. Data set for the Tula Oblast

The product was processed in a local UTM 37N projection system in a WGS84 coordinate system. The spatial resolution of the products depends of the input data of the products: 250m for the PCL, 250m for the CGL and 10m for the CTL.

3.2.2.2. Data set for the Free State Province

The product was processed in WGS84 coordinate system without projection. The spatial resolution of the products depends of the input data of the products: 250m for the PCL, 10m for the CGL and 10m for the CTL.

3.2.3. Temporal Information

The processing chain was designed to work in near real time. Therefore, the temporal information "YYYYMMD" in the filename is the "actual" date corresponding to the product value, e.g., the last date of acquisition. For the PCL, it corresponds to the last date of the MODIS time-series, i.e., the first date of the season. Similarly for the CGL, it corresponds to the last date of the MODIS time-series. For the CTL, it refers to the last date of the RapidEye/Radarsat2 coverage.

3.2.4. Data Policies

Mention these products in any publication or communication with the following citation: "***The research leading to these products has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) ImagineS project under the Grant Agreement N°311766. These products are the property of UCL under copyright ImagineS. They are generated from the RapidEye and Radarsat date distributed by CSCDA of ESA, and from MODIS data distributed by NASA.***"

3.2.5. Contact

The Crop Maps products are available through the IMAGINES website (http://fp7-imagines.eu/pages/services-and-products/crop-maps.php). Connection information to access the products are given after registration.

Scientific contact: Université catholique de Louvain (UCL), Belgium

Email address: francois.waldner@uclouvain.be, pierre.defourny@uclouvain.be

4. QUALITY ASSESSMENT

Different accuracy metrics, derived from the confusion matrix, were selected for the quality assessment. The overall accuracy evaluated the overall effectiveness of the algorithm, while the F-score, the users' accuracy and the producers' accuracy measure the accuracy of a class (Foody, 2002).

4.1. ACCURACY ASSESSMENT FOR THE TULA OBLAST

4.1.1. Pre-seasonal Cropland Layer

The algorithm described above was applied to produce the crop group layer of the growing season of 2013. Its accuracy was assessed and showed a global accuracy of 79% (Table 7).

Table 7: Confusion matrix of the Pre-seasonal Cropland Layer for the Tula site

Map \ Reference	Non-cropland	Cropland	User accuracy
Non-cropland	6142	1559	80%
Cropland	515	1762	77%
Producer Accuracy	92%	53%	Overall Accuracy: 79%

4.1.2. Crop Group Layer

The winter crop detection algorithm was applied for the 2013 winter over the Tula oblast (Table 8). For the winter crop class, one can observe an omission rate of 44% (100-Producer accuracy). This omission error was investigated thanks to the Pareto boundary (Boschetti, Flasse, & Brivio, 2004). At their early development stages, winter crops exhibit a large within field heterogeneity. The error matrix does not consider contextual influence of mixed pixels on the product accuracy. The Pareto boundary method is an alternative to deal with these shortcomings.

Table 8: Confusion matrix of the Crop Group Layer for the Tula site

Map \ Reference	Other crops	Winter Wheat	User accuracy
Other crops	348	74	82%
Winter crop	35	144	80%
Producer Accuracy	91%	66%	Overall Accuracy: 82%

4.1.3. Crop Type Layer

The Crop Type Layer was derived for the 2013 summer season in Tula (Figure 7: Evolution of the overall accuracy of the Crop Type Layer product along the 2013 season for the Tula site). The overall accuracy saturates then around 90% until the end of the season. The forest belt pattern is well captured and the resolution allows to map grassland accuracely while it was not achievable with MODIS only.
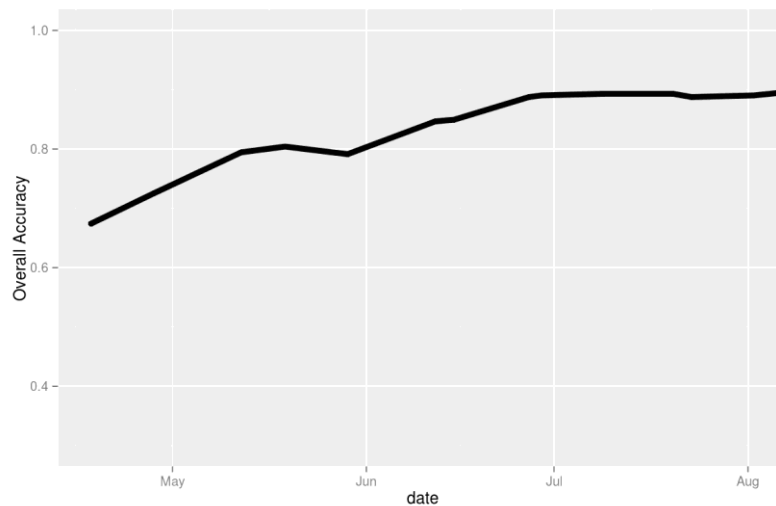

Figure 7: Evolution of the overall accuracy of the Crop Type Layer product along the 2013 season for the Tula site

4.2. ACCURACY ASSESSMENT FOR THE FREE STATE PROVINCE

4.2.1. Pre-seasonal Cropland Layer

The PCL for the Free State was validated with a sample of a thousand points randomly distributed over the area of interest. The overall accuracy reached 76% (Table 9). One can observed the relatively high omission error rate for the cropland class which might be attributed omitted cropland in areas with high cropland fragmentation where MODIS spatial resolution is insufficient to resolve the fields.

Table 9: Confusion matrix of the Pre-seasonal Cropland Layer for the Free State site

	Non Cropland	Cropland	Users' Accuracy
Non cropland	626	125	83.4%
Cropland	80	169	67.9%
Producer's accuracy	88.7%	57.5%	Overall Accuracy: 75.9%

4.2.2. Crop Group Layer

The winter crop detection method was applied over the whole Free State Province for the 2013 winter season. The overall accuracy obtained is 97% (Table 10).

Table 10: Confusion matrix of the Crop Group Layer for the Free State site

Map \ Reference	Other crops	Winter crops	Users' Accuracy
Other Crops	2171	29	98.7%
Winter Crops	80	169	68.9%
Producer's accuracy	97.8%	78.2%	Overall Accuracy: 96.8%

4.2.3. Crop Type Layer

The CTL for the Free State was validated with an independent set of about 1200 field observations. The overall accuracy reached a maximum of 78% at the end of the season (Figure 8). The maize and the planted pastures classes (the two most dominant classes) were the two classes with the highest accuracy (Fscore >0.8) (Figure 9).
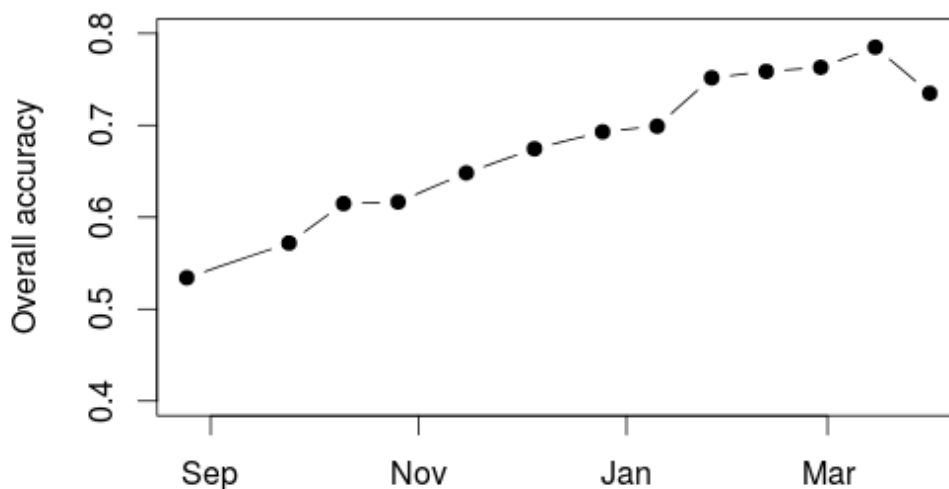

Figure 8: Evolution of the Overall Accuracy along the season for different classes included in the CTL for the Free State Province
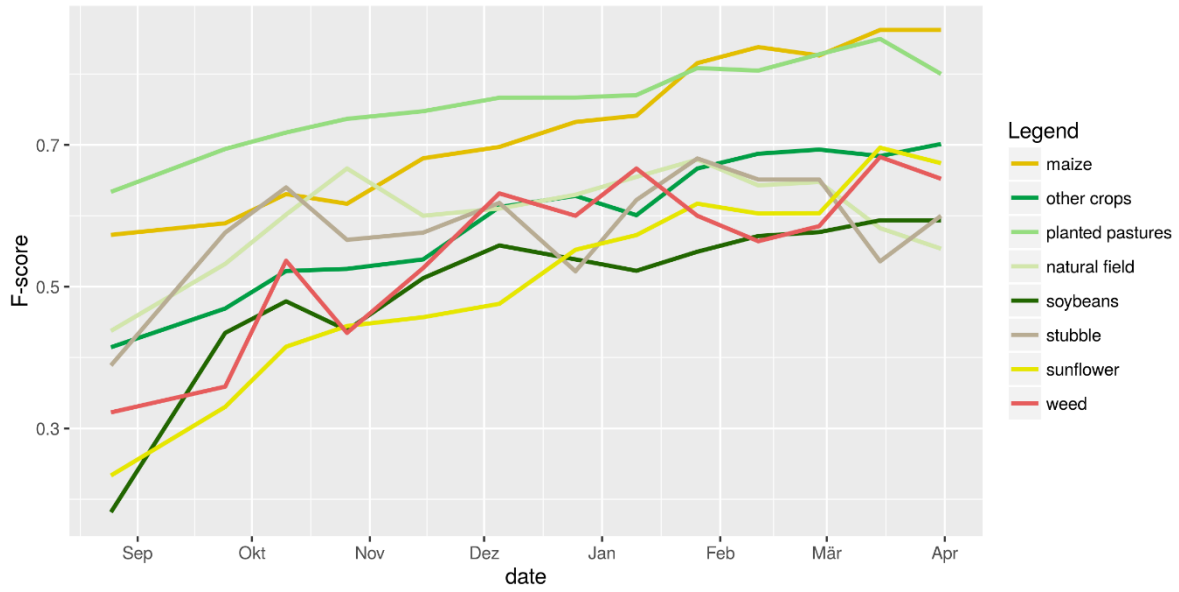
Figure 9: Evolution of the F-score along the season for different classes included in the CTL for the Free State Province

5. BIBLIOGRAPHY

Baatz, M., & Schäpe, A. (2000). Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung*.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

Desclée, B., Bogaert, P., & Defourny, P. (2006). Forest change detection by statistical object-based method. *Remote Sensing of Environment*, 102(1-2):1–11.

Eilers, P. H. (2003). A Perfect Smoother. *Analytical Chemistry* , 75(14):3631–3636.

Hégarat-Mascle, S. L., Richard, D., & Ottlé, C. (2003). Multi-scale data fusion using Dempster-Shafer evidence theory. *Integrated Computer-Aided Engineering*, 10(1), 9-22.

Jakubauskas, M., Legates, D., & Kastens, J. (2002). Crop identification using harmonic analysis of time-series AVHRR NDVI data. *Computers and Electronics in Agriculture*, 37(1-3):127–139.

Radoux, J. (2010). Updating land cover maps by GIS-driven analysis of very high resolution satellite images. *PhD Thesis, Université catholique de Louvain*.