



**IMPLEMENTING MULTI-SCALE AGRICULTURAL INDICATORS EXPLOITING SENTINELS**

**VEGETATION FIELD DATA AND PRODUCTION OF  
GROUND-BASED MAPS:**

**“GUANGDONG SITE, XUWEN, CHINA”  
APRIL TO JUNE, 2013**

**ISSUE I1.00**

EC Proposal Reference N° FP7-311766

Actual submission date : November 2014

Start date of project: 01.11.2012

Duration : 40 months

**Name of lead partner for this deliverable:** EOLAB

**Book Captain:** Consuelo Latorre (EOLAB)



Contributing Authors: Vicent Sendra, Fernando Camacho (EOLAB)

Pei Zhiyuan (CAAE)

Don Ball, Greg Gibbons, Ian Jarvis (JECAM)

<b>Project co-funded by the European Commission within the Seventh Framework Program (2007-2013)</b>		
<b>Dissemination Level</b>		
PU	Public	<b>X</b>
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

<b>Book Captain:</b>	C. Latorre	Date: 09.01.2015	Sign. 
<b>Approval:</b>	R. Lacaze	Date: 17.02.2015	Sign. 
<b>Endorsement:</b>	I. Marin-Moreno	Date:	Sign.
<b>Distribution:</b>			

## CHANGE RECORD

Issue/Revision	Date	Page(s)	Description of Change	Release
	09.01.2015	All	First Issue	I1.00

## TABLE OF CONTENTS

---

<b>1. Background of the Document</b> .....	<b>9</b>
1.1. Executive Summary .....	9
1.2. Portfolio .....	9
1.3. Scope and Objectives.....	10
1.4. Content of the Document .....	10
<b>2. Introduction</b> .....	<b>11</b>
<b>3. Study area</b> .....	<b>13</b>
3.1. Location .....	13
3.2. Description of The Test Site .....	13
<b>4. Ground measurements</b> .....	<b>15</b>
4.1. Material and Methods .....	15
4.2. Spatial Sampling Scheme .....	16
4.3. ground data.....	17
4.3.1. Data processing .....	17
4.3.2. Content of the Ground Dataset.....	17
<b>5. Evaluation of the Sampling</b> .....	<b>20</b>
5.1. Principles.....	20
5.2. Evaluation Based On NDVI Values.....	20
5.3. Evaluation Based On Convex Hull: Product Quality Flag. ....	21
<b>6. production of ground-based maps</b> .....	<b>23</b>
6.1. Imagery .....	23
6.2. The Transfer Function .....	23
6.2.1. The regression method.....	23
6.2.2. Band combination .....	24
6.2.3. The selected Transfer Function .....	25
6.3. The High Resolution Ground Based Maps .....	25
6.3.1. Mean Values .....	27
<b>7. Conclusions</b> .....	<b>28</b>
<b>8. Acknowledgements</b> .....	<b>29</b>
<b>9. References</b> .....	<b>30</b>

## LIST OF FIGURES

Figure 1: Location of Guangdong (Xuwen) site, China. ....	13
Figure 2: False color composition of TOA Reflectance Landsat 8 image over the study area. 20x20 km <sup>2</sup> area (Image used for the third campaign. Date of the image 19 <sup>th</sup> , May 2013.....	14
Figure 3: LAI-2000 optical sensor with 5 zenith angles. ....	15
Figure 4: Distribution of the sampling units (ESU) over the study area. ....	16
Figure 5: LA <sub>eff</sub> measurements acquired in Guangdong site (China) during the field campaign of 9 <sup>th</sup> April, 2013 .....	18
Figure 6: As in Figure 5 for 6 <sup>th</sup> May, 2013. ....	18
Figure 7: As in Figure 5 for 26 <sup>th</sup> May, 2013.....	18
Figure 8: As in Figure 5 for 19 <sup>th</sup> June, 2013 .....	18
Figure 9: Evolution of five selected ESUS of LA <sub>eff</sub> measurements among the four campaigns, Guangdong site – China, 2013. ....	19
Figure 10: Distribution of the measured biophysical variable LA <sub>eff</sub> over the ESUs. Top Left: First campaign (9 <sup>th</sup> April, 2013). Top Right: Second campaign (6 <sup>th</sup> May, 2013). Bottom Left: Third campaign (26 <sup>th</sup> May, 2013). Top Right: Fourth campaign (19 <sup>th</sup> June, 2013). Guangdong site (China), 2013. ....	19
Figure 11: Comparison of NDVI (TOA) distribution between ESUs (green dots) and over the whole image (Blue line), Guangdong (Xuwen) site. China (26 <sup>th</sup> May 2013). ....	21
Figure 12: Convex Hull test over 20x20 km <sup>2</sup> area centered at the test site and the selected area for validation (5x5 km <sup>2</sup> ): clear and dark blue correspond to the pixels belonging to the ‘strict’ and ‘large’ convex hulls. Red corresponds to the pixels for which the transfer function is extrapolating, and Grey for water areas. Guangdong (China). (26 <sup>th</sup> May 2013). ....	22
Figure 13: Test of multiple regression (TF) applied on different band combinations. Band combinations are given in abscissa (1=G, 2=RED, 3=NIR and 4=SWIR). The weighted root mean square error (RMSE) is presented in red along with the cross-validation RMSE in green. The numbers indicate the number of data used for the robust regression with a weight lower than 0.7 that could be considered as outliers. Third Campaign (26 <sup>th</sup> May, 2013). Guangdong, China.....	24
Figure 14: Scatter-plots between ground observations and their corresponding transfer function (TF) estimates using 4 bands combination. Full dots: Weight>0.7. Empty dots: 0<Weight<0.7. Third Campaign (26 <sup>th</sup> May, 2013). Guangdong, China.....	25
Figure 15: HR LA <sub>eff</sub> maps (20x20 km <sup>2</sup> ) retrieved on the Guangdong site. Third Campaign (26 <sup>th</sup> May, 2013). ....	26
Figure 16: HR LA <sub>eff</sub> maps (5x5 km <sup>2</sup> ) retrieved on the Guangdong site. Third Campaign (26 <sup>th</sup> May, 2013). ....	26

## LIST OF TABLES

---

<i>Table 1: Ground Campaigns and Landsat 8 Imagery available.....</i>	<i>12</i>
<i>Table 2: Coordinates and altitude of the test site (centre of 20x20 km<sup>2</sup> study area).....</i>	<i>13</i>
<i>Table 3: The Header used to describe ESUs with the ground measurements. ....</i>	<i>17</i>
<i>Table 4: Percentages over the selected area over the test site of Guangdong – China. (0: extrapolation of TF, 1: strict convex hull, 2: large convex hull, 3: water mask).....</i>	<i>22</i>
<i>Table 5: Acquisition properties of Landsat 8 data used for retrieving high resolution maps.....</i>	<i>23</i>
<i>Table 6: Transfer function applied to the whole site for LAI<sub>eff</sub>. RW for weighted RMSE, and RC for cross-validation RMSE. ....</i>	<i>25</i>
<i>Table 7: Mean value and standard deviation (STD) of the HR biophysical maps for the selected 3x3 km<sup>2</sup> area at Guangdong site. ....</i>	<i>27</i>
<i>Table 8: Content of the dataset.....</i>	<i>27</i>

## LIST OF ACRONYMS

---

<b>CAAE</b>	Chinese Academy of Agricultural Engineering
<b>CEOS</b>	Committee on Earth Observation Satellite
<b>CEOS LPV</b>	Land Product Validation Subgroup
<b>DG AGRI</b>	Directorate General for Agriculture and Rural Development
<b>DG RELEX</b>	Directorate General for External Relations (European Commission)
<b>ECV</b>	Essential Climate Variables
<b>EUROSTATS</b>	Directorate General of the European Commission
<b>ESU</b>	Elementary Sample Unit
<b>FAPAR</b>	Fraction of Absorbed Photo-synthetically Active Radiation
<b>FAO</b>	Food and Agriculture Organization
<b>FCOVER</b>	Fraction of Vegetation Cover
<b>GCOS</b>	Global Climate Observing System
<b>GEO-GLAM</b>	Global Agricultural Geo- Monitoring Initiative
<b>GIO-GL</b>	GMES Initial Operations - Global Land (GMES)
<b>GCOS</b>	Global Climate Observing System
<b>GMES</b>	Global Monitoring for Environment and Security
<b>IMAGINES</b>	Implementing Multi-scale Agricultural Indicators Exploiting Sentinels
<b>JECAM</b>	Joint Experiment for Crop Assessment and Monitoring
<b>LAI</b>	Leaf Area Index
<b>LDAS</b>	Land Data Assimilation System
<b>LUT</b>	Look-up-table techniques
<b>PAI</b>	Plant Area Index
<b>PROBA-V</b>	Project for On-Board Autonomy satellite, the V standing for vegetation.
<b>RMSE</b>	Root Mean Square Error
<b>SPOT /VGT</b>	Satellite Pour l'Observation de la Terre / VEGETATION
<b>SCI</b>	GMES Services Coordinated Interface
<b>TOA</b>	Top of Atmosphere Reflectance
<b>UNFCCC</b>	United Nations Framework Convention on Climate Change
<b>UTM</b>	Universal Transverse Mercator coordinate system
<b>VALERI</b>	Validation of Land European Remote sensing Instruments
<b>WGCV</b>	Working Group on Calibration and Validation (CEOS)
<b>WGS-84</b>	World Geodetic System



# 1. BACKGROUND OF THE DOCUMENT

## 1.1. EXECUTIVE SUMMARY

The Copernicus Land Service has been built in the framework of the FP7 geoland2 project, which has set up pre-operational infrastructures. ImagineS intends to ensure the continuity of the innovation and development activities of geoland2 to support the operations of the global land component of the GMES Initial Operation (GIO) phase. In particular, the use of the future Sentinel data in an operational context will be prepared. Moreover, IMAGINES will favor the emergence of new downstream activities dedicated to the monitoring of crop and fodder production.

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.

Further, ImagineS will serve the growing needs of international (e.g. FAO and NGOs), European (e.g. DG AGRI, EUROSTATS, DG RELEX), and national users (e.g. national services in agro-meteorology, ministries, group of producers, traders) on accurate and reliable information for the implementation of the EU Common Agricultural Policy, of the food security policy, for early warning systems, and trading issues. ImagineS will also contribute to the Global Agricultural Geo-Monitoring Initiative (GEO-GLAM) by its original agriculture service which can monitor crop and fodder production together with the carbon and water fluxes and can provide drought indicators, and through links with JECAM (Joint Experiment for Crop Assessment and Monitoring).

## 1.2. PORTFOLIO

The ImagineS portfolio contains global and regional biophysical variables derived from multi-sensor satellite data, at different spatial resolutions, together with agricultural indicators, including the above-ground biomass, the carbon and water fluxes, and drought indices resulting from the assimilation of the biophysical variables in the Land Data Assimilation System (LDAS).

The production in Near Real Time of the 333m resolution products, at a frequency of 10 days, using PROBA-V data is carried out in the Copernicus Global Land Service. It starts by covering Europe only, and be gradually extended to the whole globe.

Meanwhile, ImagineS will perform in parallel off-line production over demonstration sites outside Europe. The demonstration of high resolution (30m) products (Landsat-8 + PROBA-

V) will be done over demonstration sites of cropland and grassland in contrasting climatic and environmental conditions.

### 1.3. SCOPE AND OBJECTIVES

The main objective of this document is to describe the field campaign and ground data collected at Guangdong site in China and the up-scaling of the ground data to produce a ground-based high resolution map of the following biophysical variable:

- Leaf Area Index (LAI), defined as half of the total developed area of leaves per unit ground surface area ( $m^2/m^2$ ). We focused on an effective LAI ( $LAI_{eff}$ ) derived from the description of the gap fraction as a function of the view zenith angle, for green elements.

### 1.4. CONTENT OF THE DOCUMENT

This document is structured as follows:

- Chapter 2 provides an introduction to the field experiment.
- Chapter 3 provides the location and description of the site.
- Chapter 4 describes the ground measurements, including material and methods, sampling and data processing.
- Chapter 5 provides an evaluation of the sampling.
- Chapter 6 describes the production of high resolution ground-based maps, and the selected “mean” values for validation.

## 2. INTRODUCTION

Validation of remote sensing products is mandatory to guaranty that the satellite products meets the user's requirements. Protocols for validation of global LAIeff products are already developed in the context of Land Product Validation (LPV) group of the Committee on Earth Observation Satellite (CEOS) for the validation of satellite-derived land products (Fernandes et al., 2014), and recently applied to Copernicus global land products based on SPOT/VGT observation (Camacho et al., 2013). This generic approach is made of 2 major components:

- The indirect validation: including inter-comparison between products as well as evaluation of their temporal and spatial consistency
- The direct validation: comparing satellite products to ground measurements of the corresponding biophysical variables. In the case of low and medium resolution sensors, the main difficulty relies on scaling local ground measurements to the extent corresponding to pixels size. However, the direct validation is limited by the small number of sites, for that reason a main objective of ImagineS is the collection of ground truth data in demonstration sites.

The content of this document is compliant with existing validation guidelines (for direct validation) as proposed by the CEOS LPV group (Morissette et al., 2006); the VALERI project (<http://w3.avignon.inra.fr/valeri/>) and ESA campaigns (Baret and Fernandes, 2012). It therefore follows the general strategy based on a bottom up approach: it starts from the scale of the individual measurements that are aggregated over an elementary sampling unit (ESU) corresponding to a support area consistent with that of the high resolution imagery used for the up-scaling of ground data. Several ESUs are sampled over the site. Radiometric values over a decametric image are also extracted over the ESUs. These are later used to develop empirical transfer functions for up-scaling the ESU ground measurements (e.g. Martínez et al., 2009). Finally, the high resolution ground based map is compared with the medium resolution satellite product at the spatial support of the product.

A multi-temporal field campaign from April to June to characterize the LAI over rice crops at the Guangdong - Xuwen (China) test site were carried out by the Chinese Academy of Agricultural Engineering, involved in JECAM initiative.

### Field Campaigns:

Table 1 shows the date of the four field campaigns performed over the rice at the Guangdong site. A cloud-free Landsat-8 acquisition was available close to the third campaign (26<sup>th</sup> May) for the up-scaling of the ground measurements.

**Table 1: Ground Campaigns and Landsat 8 Imagery available.**

<b>CAMPAIGN</b>	<b>DATES</b>	<b>IMAGERY</b>
<b>First campaign</b>	9 <sup>th</sup> April, 2013	Cloud contaminated
<b>Second campaign</b>	6 <sup>th</sup> May, 2013	Cloud contaminated
<b>Third campaign</b>	26 <sup>th</sup> May, 2013	LANDSAT8 TOA (19.05.2013)
<b>Fourth campaign</b>	19 <sup>th</sup> June, 2013	Cloud contaminated

**Team involved in field collection:**

CAAE: Zhiyuan, P.

**Contact:**

EOLAB: Fernando Camacho (fernando.camacho@eolab.es)

CAAE: Pei Zhiyuan (peizhiyuan@agri.gov.cn)

JECAM: Ian Jarvis (Ian.Jarvis@AGR.GC.CA)

### 3. STUDY AREA

#### 3.1. LOCATION

The Guangdong (Xuwen) site is located in the Guangdong province. Guangdong (traditional Chinese: 廣東) is a province on the South China Sea coast of the People's Republic of China. Formerly known as Canton or Kwangtung in English, Guangdong surpassed Henan and Sichuan to become the most populous province in China in January 2005. The provincial capital Guangzhou and economic hub Shenzhen are among the most populous and important cities in China. (See Figure 1).

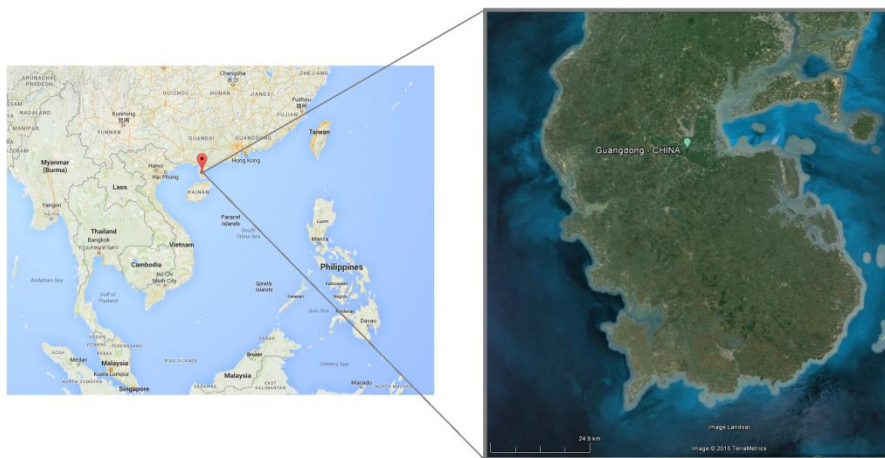


Figure 1: Location of Guangdong (Xuwen) site, China.

Table 2: Coordinates and altitude of the test site (centre of 20x20 km<sup>2</sup> study area).

Site Center	
Geographic Lat/lon, WGS-84 (degrees)	Latitude = 20°52'12" N Longitude = 110°04'48" E
Altitude	2 m

#### 3.2. DESCRIPTION OF THE TEST SITE

Figure 2 shows a Landsat-8 color composition of the Guangdong site. The paddy rice fields can be easily identified in the basin of the Nandu river, the longest river in Hainan Province, China.

Typical field size ranges from 0.1 to 10.0 ha. Crop types are mainly paddy rice, but also sugar cane, peanut, and other vegetable are cultivated. Typical crop rotation is paddy rice with two harvests per year. Sowing of main rice takes place in late March and June, to be

harvested in July. The second rice crop is sown in August and is harvested from late October to November. The climate is characterized by relatively high temperatures. The average temperature is 22° Celsius and the warmest month is July with an average temperature of 28° Celsius. The coolest month is January, with an average temperature of 14° Celsius. The rainfall is evenly distributed over the entire year. In summer, the rainfall comes from the tropical cyclones. The average amount of rainfall is 2402.8 mm. The soil texture is loam. The landscape topology is flatlands and hills.

(<http://www.jecam.org/?/site-description/china-guangdong>).



**Figure 2: False color composition of TOA Reflectance Landsat 8 image over the study area. 20x20 km<sup>2</sup> area (Image used for the third campaign. Date of the image 19<sup>th</sup>, May 2013.**

## 4. GROUND MEASUREMENTS

### 4.1. MATERIAL AND METHODS

#### 4.1.3 LI-COR LAI-2000 plant canopy analyzer

The LI-COR (Inc., Lincoln, Nebraska, 2009) LAI-2000 is a plant canopy analyzer used in the field campaign.

The instrument calculates Leaf Area Index (LAI) and other canopy attributes from light measurements made with a “fish-eye” optical sensor (148° field-of-view). Measurements made above and below the canopy are used to calculate canopy light interception at five zenith angles. The average probability of light penetration into the canopy is computed by

$$\overline{P(\theta_i)} = \frac{1}{N_{obs}} \sum_{j=1}^{N_{obs}} \frac{B_{ij}}{A_{ij}} \quad \text{Eq. (1)}$$

where the subscript  $i$  ( $i = 1 \dots 5$ ) refers to the optical sensor rings centered at  $\theta_i$  and  $j$  refers to the number of observational pairs ( $j = 1 \dots N_{obs}$ ).  $B_{ij}$  and  $A_{ij}$  are the  $j^{th}$  below and above canopy readings, respectively, for the  $i^{th}$  ring. The gap fraction for the  $i^{th}$  ring is computed from

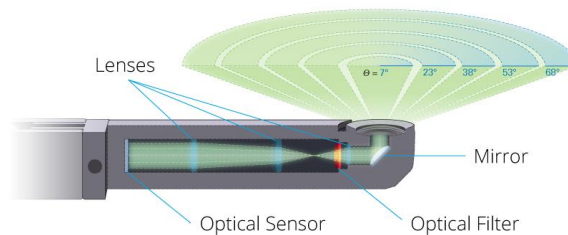
$$G_i = e^{\left(\overline{\ln P(\theta_i)}\right)} = e^{\left(\frac{1}{N_{obs}} \sum_{j=1}^{N_{obs}} \ln \frac{B_{ij}}{A_{ij}}\right)} \quad \text{Eq. (2)}$$

Assuming the foliage elements are randomly distributed in space, the effective PAI (PAI<sub>eff</sub>) can be estimated by the transmittance in the different view angles based on Miller’s formula (Miller, 1967).

$$PAI_{eff} = 2 \int_0^{\pi/2} -\ln P(\theta) \cos \theta \sin \theta d\theta \quad \text{Eq. (3)}$$

The amount of foliage in a vegetative canopy can be deduced from measurements of how quickly radiation is attenuated as it passes through the canopy. By measuring this attenuation at several angles from the zenith, foliage orientation information can also be obtained. The LAI-2200 measures the attenuation of diffuse sky radiation at five zenith angles simultaneously, arranged in concentric rings.

A normal measurement with the LAI-2200 consists of a minimum of ten numbers: five of the numbers are the signals from the five detectors when the optical sensor was above the vegetation, and the remaining five are the readings made with the sensor below the vegetation. For both readings, the sensor is looking up at the sky. Five values of canopy transmittance are calculated from these readings by dividing corresponding pairs. (Figure 3)

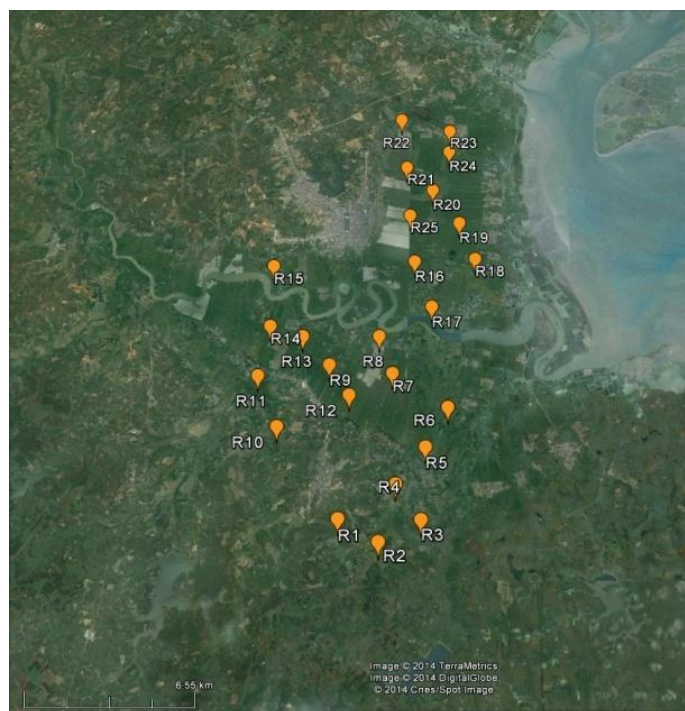


**Figure 3: LAI-2000 optical sensor with 5 zenith angles.**

## 4.2. SPATIAL SAMPLING SCHEME

A total of 25 ESUs (Elementary Sample Unit) of one land cover types (RICE) were characterized during the campaigns. The sampling scheme within each ESU was not provided by the team in charge of the measurements.

The centre of the ESU was geo-located. Figure 4 shows the distribution of the sampling units over the experimental site. The ground measurements are spread across the paddy rice fields.



**Figure 4: Distribution of the sampling units (ESU) over the study area.**



## 4.3. GROUND DATA

### 4.3.1. Data processing

LAI<sub>eff</sub> data was estimated with LICOR-LAI2000 (Equ.3), no additional processing was done.

### 4.3.2. Content of the Ground Dataset

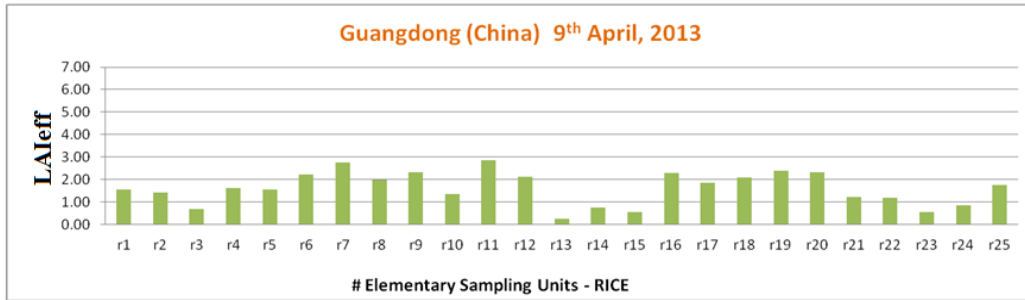
Each ESU is described according to a standard format. The header of the database is shown in Table 3.

**Table 3: The Header used to describe ESUs with the ground measurements.**

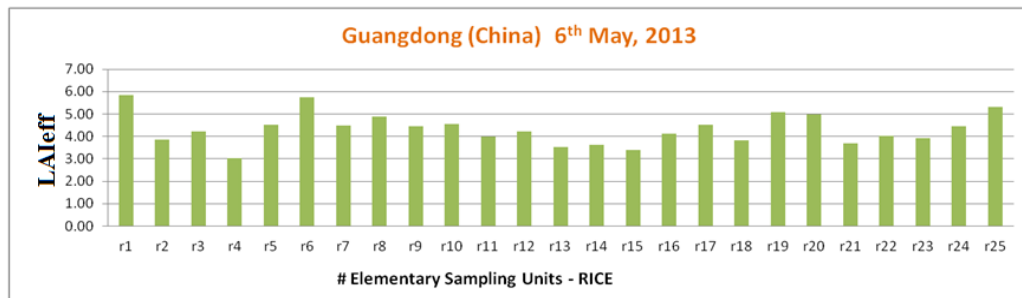
Column	Var.Name	Comment	
1	Plot #	Number of the field plot in the site	
2	Plot Label	Label of the plot in the site	
3	ESU #	Number of the Elementary Sampling Unit (ESU)	
4	ESU Label	Label of the ESU in the campaign	
5	Northing Coord.	Geographical coordinate: Latitude (°), WGS-84	
6	Easting Coord.	Geographical coordinate: Longitude (°), WGS-84	
7	Extent (m) of ESU (diameter)	Size of the ESU <sup>(1)</sup>	
8	Land Cover	Detailed land cover	
9	Start Date (dd/mm/yyyy)	Starting date of measurements	
10	End Date (dd/mm/yyyy)	Ending date of measurements	
11	LAI <sub>eff</sub>	Method	Instrument
12		Nb. Replications	Number of Replications
13		PRODUCT	Methodology
14		Uncertainty	Standard deviation

Four field campaigns have been performed from April to June, 2013.

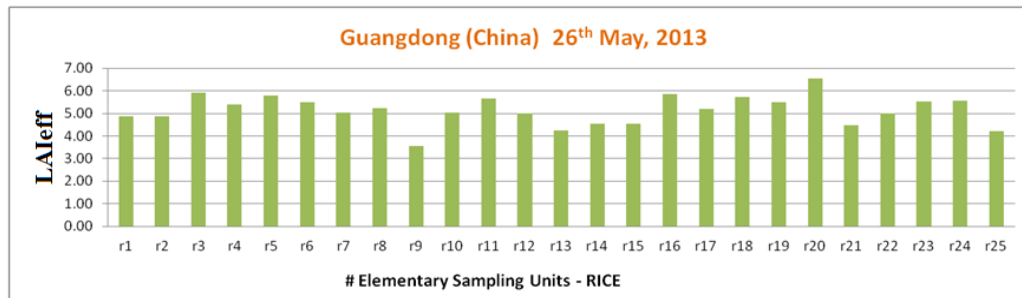
Figure 5 to Figure 8 show the measurements obtained during the field experiments. Figure 5 shows the measurements for all the field campaigns. For the first campaign, LAI<sub>eff</sub> values vary from 0.3 to 2.9. For the second campaign, LAI<sub>eff</sub> values are typically between 3 (ESU r4) and 6 (ESU r1) (Figure 6). Slightly higher values were reported for the third campaign (Figure 7). Finally, Figure 8 shows the LAI<sub>eff</sub> values covering the last stage of the rice crop, with some fields harvested.



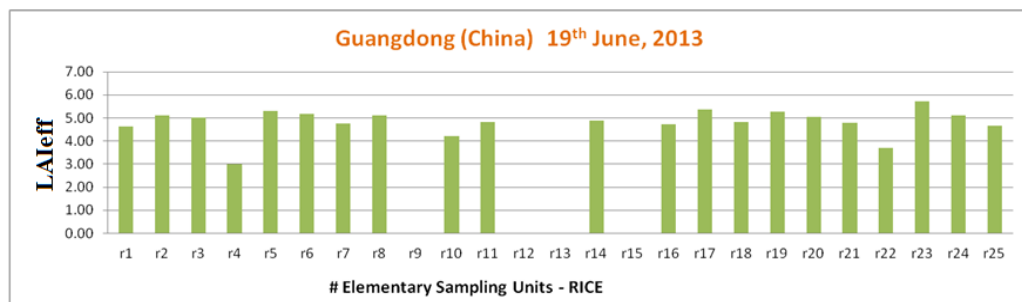
**Figure 5: LAIeff measurements acquired in Guangdong site (China) during the field campaign of 9<sup>th</sup> April, 2013**



**Figure 6: As in Figure 5 for 6<sup>th</sup> May, 2013.**

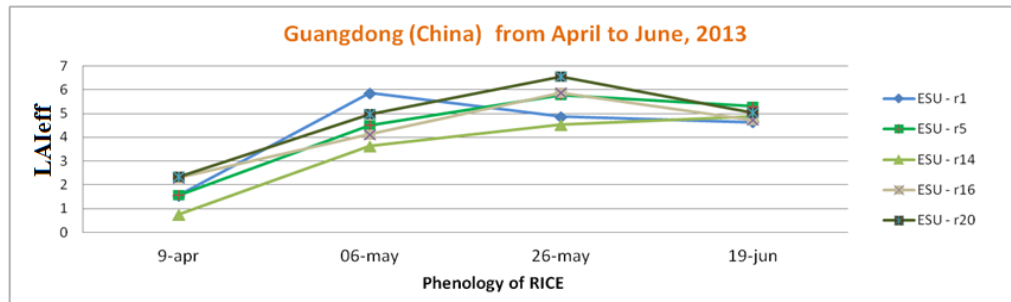


**Figure 7: As in Figure 5 for 26<sup>th</sup> May, 2013.**



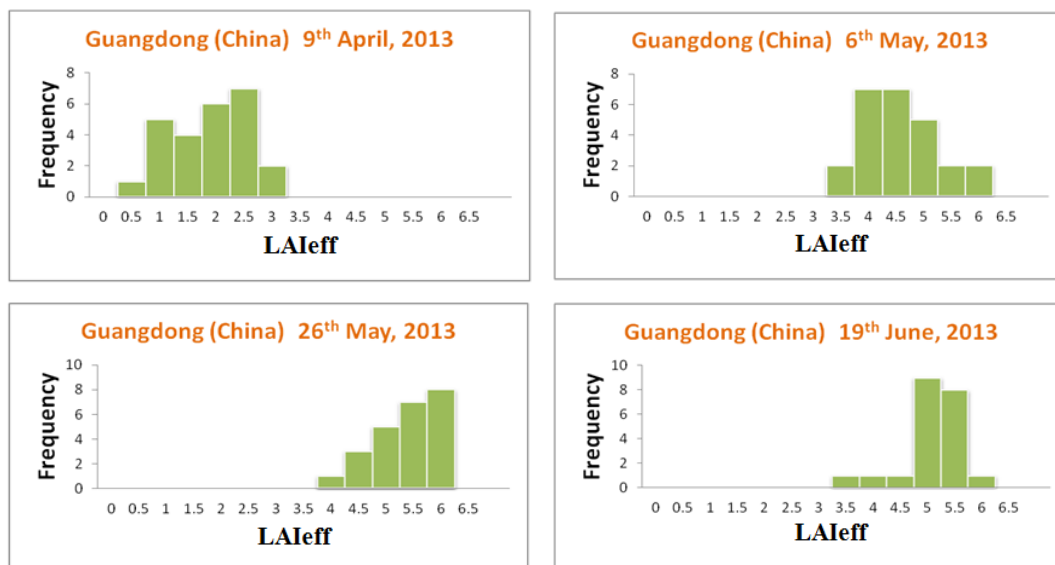
**Figure 8: As in Figure 5 for 19<sup>th</sup> June, 2013**

Figure 9 describes the phenology trends of several ESUs selected among the field experiment. Note that ESU 1 (r1) shows a different phenology than in the other fields probably due to a different type of rice.



**Figure 9: Evolution of five selected ESUS of LAIeff measurements among the four campaigns, Guangdong site – China, 2013.**

Figure 10 describes the distribution of the measurements for all field campaigns. It shows the evolution of the rice fields from growing stage in April to maximum in the third campaign at the end of May, as Figure 9 presents for several representative ESUs. Note that in June the values are slightly lower than at the end of May, which could mean the plant enter into the maturation stage so the plant activity is mainly in developing the grain rather than the foliar elements.



**Figure 10: Distribution of the measured biophysical variable LAIeff over the ESUs. Top Left: First campaign (9<sup>th</sup> April, 2013). Top Right: Second campaign (6<sup>th</sup> May, 2013). Bottom Left: Third campaign (26<sup>th</sup> May, 2013). Top Right: Fourth campaign (19<sup>th</sup> June, 2013). Guangdong site (China), 2013.**

## 5. EVALUATION OF THE SAMPLING

### 5.1. PRINCIPLES

Based on previous field activities, the data set sampling was concentrated in the most representative areas. 25 ESUs were used for the up-scaling with the high resolution imagery.

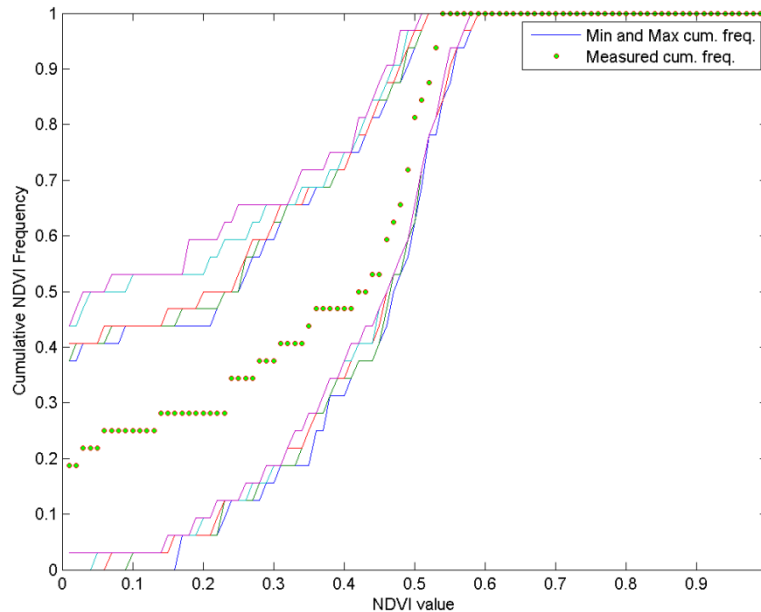
### 5.2. EVALUATION BASED ON NDVI VALUES

The sampling strategy is evaluated using the Landsat-8 image by comparing the NDVI distribution over the site (the extended area  $20 \times 20 \text{ km}^2$ ) with the NDVI distribution over the ESUs (Figure 11). As the number of pixels is drastically different for the ESU and whole site (WS) it is not statistically consistent to directly compare the two NDVI histograms. Therefore, the proposed technique consists in comparing the NDVI cumulative frequency of the two distributions by a Monte-Carlo procedure which aims at comparing the actual frequency to randomly shifted sampling patterns. It consists in:

1. computing the cumulative frequency of the  $N$  pixel NDVI that correspond to the exact ESU locations; then, applying a unique random translation to the sampling design (modulo the size of the image)
2. computing the cumulative frequency of NDVI on the randomly shifted sampling design
3. repeating steps 2 and 3, 199 times with 199 different random translation vectors.

This provides a total population of  $N = 199 + 1$  (actual) cumulative frequency on which a statistical test at acceptance probability  $1 - \alpha = 95\%$  is applied: for a given NDVI level, if the actual ESU density function is between two limits defined by the  $N\alpha / 2 = 5$  highest and lowest values of the 200 cumulative frequencies, the hypothesis assuming that WS and ESU NDVI distributions are equivalent is accepted, otherwise it is rejected.

Figure 11 shows the NDVI (Top Of Atmosphere) distribution over a  $20 \times 20 \text{ km}^2$  area centred in Guangdong site for the 26<sup>th</sup> May, 2013 campaign. The sampling is good over the whole site as it was comprised between the 5 highest and 5 lowest cumulative frequencies. The sampling presents a small bias towards higher NDVI values as the sampling was concentrated over densest rice fields.



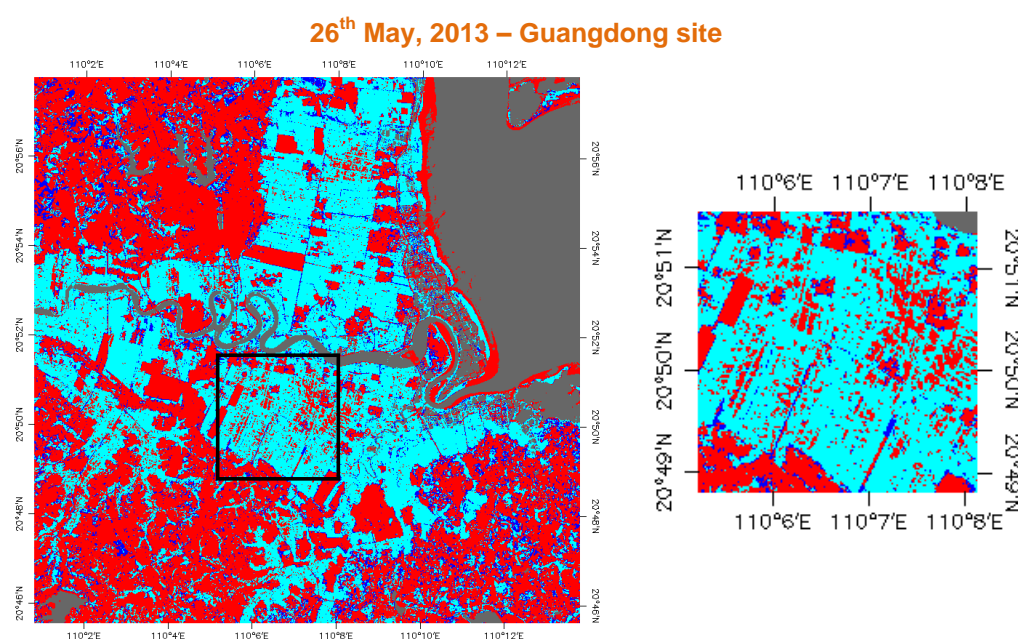
**Figure 11: Comparison of NDVI (TOA) distribution between ESUs (green dots) and over the whole image (Blue line), Guangdong (Xuwen) site, China (26<sup>th</sup> May 2013).**

### 5.3. EVALUATION BASED ON CONVEX HULL: PRODUCT QUALITY FLAG.

The interpolation capabilities of the empirical transfer function used for up-scaling the ground data using decametric images is dependent of the sampling (Martinez et al., 2009). A test based on the convex hulls was also carried out to characterize the representativeness of ESUs and the reliability of the empirical transfer function using the different combinations of the selected bands (green, red, NIR and SWIR) of the Landsat 8 image. A flag image is computed over the reflectances. The result on convex-hulls can be interpreted as:

- pixels inside the 'strict convex-hull': a convex-hull is computed using all the Landsat 8 reflectances corresponding to the ESUs belonging to the class. These pixels are well represented by the ground sampling and therefore, when applying a transfer function the degree of confidence in the results will be quite high, since the transfer function will be used as an interpolator;
- pixels inside the 'large convex-hull': a convex-hull is computed using all the reflectance combinations ( $\pm 5\%$  in relative value) corresponding to the ESUs. For these pixels, the degree of confidence in the obtained results will be quite good, since the transfer function is used as an extrapolator (but not far from interpolator);
- pixels outside the two convex-hulls: this means that for these pixels, the transfer function will behave as an extrapolator which makes the results less reliable. However, having a priori information on the site may help to evaluate the extrapolation capacities of the transfer function.

Figure 12 shows the results of the Convex-Hull test (i.e., Quality Flag image) for the Guangdong site over a 20x20 km<sup>2</sup> and 5x5 km<sup>2</sup> areas. The strict and large convex-hulls are high around the test site (41 % over the 20x20 km<sup>2</sup> area and 70% over the 5x5 km<sup>2</sup> validation area). The QF map shows also that there is a quite large area where the transfer function behaves as extrapolator corresponding to different land use. Nevertheless, the results obtained in the empirical maps seem to be reliable, but should be considered with caution. Note that the Convex-Hull test provides information on the representativeness of the sampling, informing on the confidence of the interpolation performed by the empirical function. Furthermore, water areas (river and sea) in the image were masked if NDVI-TOA values were lower than 0.0001.



**Figure 12: Convex Hull test over 20x20 km<sup>2</sup> area centered at the test site and the selected area for validation (5x5 km<sup>2</sup>): clear and dark blue correspond to the pixels belonging to the ‘strict’ and ‘large’ convex hulls. Red corresponds to the pixels for which the transfer function is extrapolating, and Grey for water areas. Guangdong (China). (26<sup>th</sup> May 2013).**

Table 4 shows the percentages of the different results over the study area 20x20km<sup>2</sup> and the selected area 5x5km<sup>2</sup> over the test site of Guangdong-Xuwen, China.

**Table 4: Percentages over the selected area over the test site of Guangdong – China. (0: extrapolation of TF, 1: strict convex hull, 2: large convex hull, 3: water mask).**

Field Campaigns, 2013		Quality Flags (%)							
DATE		20x20 km <sup>2</sup>				5x5 km <sup>2</sup>			
Convex hull values		0	1	2	3	0	1	2	3
<b>26<sup>th</sup> May</b>		41%	32%	9%	17%	29%	64%	6%	1%

## 6. PRODUCTION OF GROUND-BASED MAPS

### 6.1. IMAGERY

The Landsat 8 image was acquired the 19<sup>th</sup> May 2013 (see Table 5 for acquisition properties). Four spectral bands were selected from 530 nm to 1650 nm with a nadir ground sampling distance of 30 m. For the transfer function analysis, the input satellite data used is Top of Atmosphere (TOA) reflectance. The original projection is UTM 49 North, WGS-84.

**Table 5: Acquisition properties of Landsat 8 data used for retrieving high resolution maps.**

LANDSAT 8 METADATA	
Platform / Instrument	LANDSAT_8 / OLI_TIRS
Path	124
Row	46
Spectral Range (selected bands)	B1(green) : 0.53-0.59 $\mu\text{m}$ B2(red) : 0.64-0.67 $\mu\text{m}$ B3(NIR) : 0.85-0.88 $\mu\text{m}$ B4(SWIR) : 1.58-1.65 $\mu\text{m}$
	26 <sup>th</sup> May campaign, 2013
Acquisition date	2013-05-19
Illumination Azimuth angle	86.98119962 °
Illumination Elevation angle	68.85006975 °
Ground control points	80
Geometric RMSE	5.577

### 6.2. THE TRANSFER FUNCTION

#### 6.2.1. The regression method

If the number of ESUs is enough, multiple robust regression 'REG' between ESUs reflectance and the considered biophysical variable can be applied (Martínez et al., 2009): we used the 'robustfit' function from the Matlab statistics toolbox. It uses an iteratively re-weighted least squares algorithm, with the weights at each iteration computed by applying the bi-square function to the residuals from the previous iteration. This algorithm provides lower weight to ESUs that do not fit well.

The results are less sensitive to outliers in the data as compared with ordinary least squares regression. At the end of the processing, two errors are computed: weighted RMSE

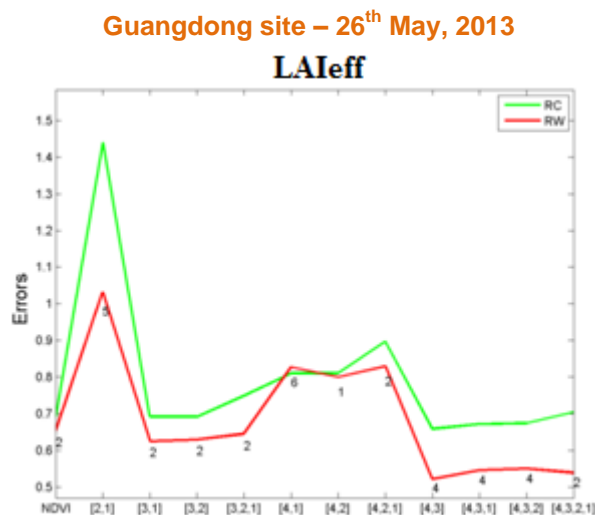
(RW) (using the weights attributed to each ESU) and cross-validation RMSE (RC) (leave-one-out method).

As the method has limited extrapolation capacities, a flag image (Figure 12), based on the convex hulls, is included in the final ground based map in order to inform the users on the reliability of the estimates.

### 6.2.2. Band combination

Figure 13 shows the results obtained for all the possible band combinations using TOA reflectances. Attending specifications of minimal RMSE, it has been chosen for the intensive campaign (26<sup>th</sup> May, 2013): band 1 (green), band 2 (red), band 3 (Near Infrared) and band 4 (Short Wave Infrared) combination of (1, 2, 3, 4) = (G, R, NIR, SWIR).

This combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE, the weighted RMSE (lowest value) and the number of rejected points.



**Figure 13: Test of multiple regression (TF) applied on different band combinations. Band combinations are given in abscissa (1=G, 2=RED, 3=NIR and 4=SWIR). The weighted root mean square error (RMSE) is presented in red along with the cross-validation RMSE in green. The numbers indicate the number of data used for the robust regression with a weight lower than 0.7 that could be considered as outliers. Third Campaign (26<sup>th</sup> May, 2013). Guangdong, China.**

Figure 14 shows scatter-plot between ground observations and their corresponding transfer function (TF) estimates for the selected band combination. A good correlation is observed for the LAI<sub>eff</sub> with points distributed along the 1:1 line, and no systematic errors. The empirical transfer function slightly overestimates the few control points selected over bare and water areas, however these areas cover only a very small fraction of the selected validation area, and the impact is expected to be very low.



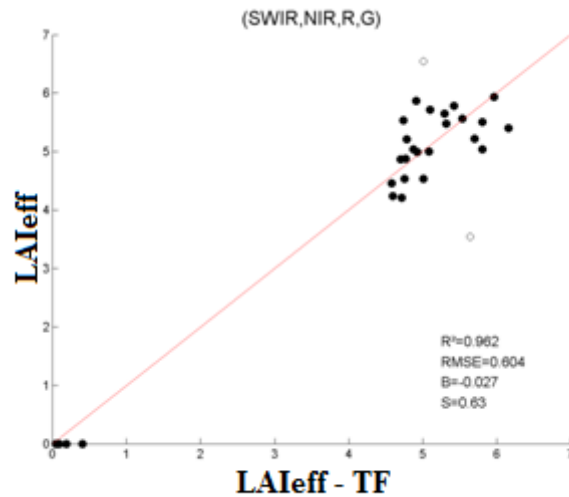


Figure 14: Scatter-plots between ground observations and their corresponding transfer function (TF) estimates using 4 bands combination. Full dots: Weight>0.7. Empty dots: 0<Weight<0.7. Third Campaign (26<sup>th</sup> May, 2013). Guangdong, China.

### 6.2.3. The selected Transfer Function

The applied transfer function is detailed in Table 6, along with its weighted RMS (RW) and cross validated RMS (RC) errors.

Table 6: Transfer function applied to the whole site for LAIeff. RW for weighted RMSE, and RC for cross-validation RMSE.

Variable	Band Combination	RW	RC
26 <sup>th</sup> May, 2013 – Guangdong site			
LAIeff	$3.04464 - 0.00069 \cdot (\text{SWIR}) + 0.00059 \cdot (\text{NIR})$ $+0.00158 \cdot (\text{R}) - 0.00169 \cdot (\text{G})$	0.54	0.65

## 6.3. THE HIGH RESOLUTION GROUND BASED MAPS

The high resolution maps are obtained applying the selected transfer function (Table 6) to the Landsat-8 TOA reflectance image. Figure 15 displays the LAIeff over a 20x20 km<sup>2</sup> area, and Figure 16 shows the LAIeff over the 5x5 km<sup>2</sup> area

Guangdong site – 26<sup>th</sup> May, 2013

LAI<sub>eff</sub>

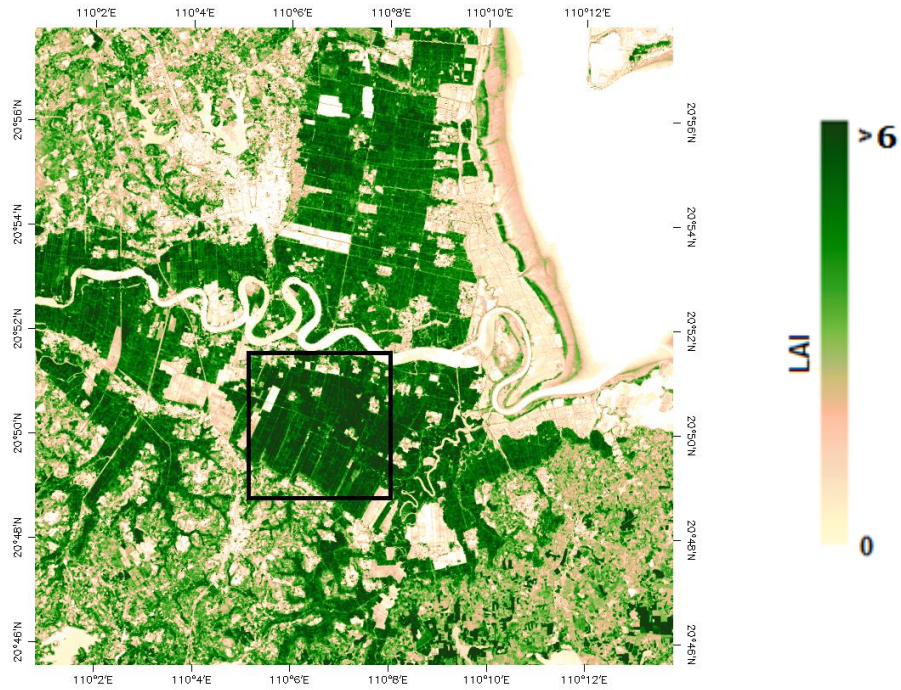


Figure 15: HR LAI<sub>eff</sub> maps (20x20 km<sup>2</sup>) retrieved on the Guangdong site. Third Campaign (26<sup>th</sup> May, 2013).

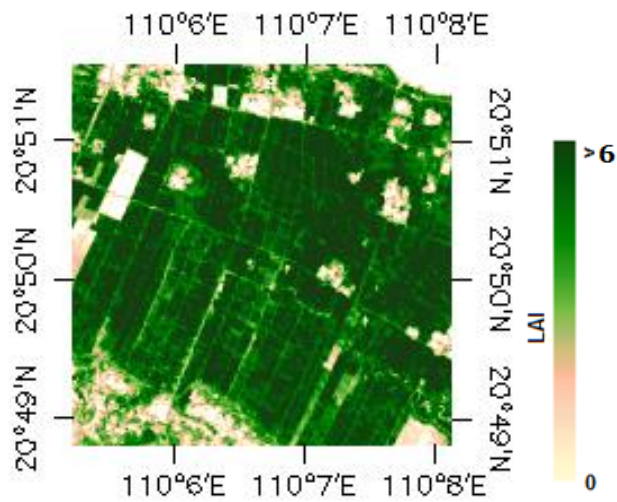


Figure 16: HR LAI<sub>eff</sub> maps (5x5 km<sup>2</sup>) retrieved on the Guangdong site. Third Campaign (26<sup>th</sup> May, 2013).

### 6.3.1. Mean Values

Mean value of a 3x3 km<sup>2</sup> area centred in the test site is provided for validation of 1 km satellite products in order to reduce co-registration and PSF errors, and in agreement with the CEOS OLIVE direct dataset (Table 7). For the validation of coarser resolutions product (e.g. MSG products) a larger area should be considered. For this reason, empirical maps are provided at 5x5 km<sup>2</sup> and 20x20 km<sup>2</sup>.

**Table 7: Mean value and standard deviation (STD) of the HR biophysical maps for the selected 3x3 km<sup>2</sup> area at Guangdong site.**

NAME	COORDINATES		LAI <sub>eff</sub>	
	LATITUDE	LONGITUDE	MEAN	STD
RICE	20.837	110.111	5.33	1.00

Table 8 describes the content of the geo-biophysical maps in the nomenclature: "BIO\_YYYYMMDD\_SENSOR\_Site ETF\_Area" files, where:

- BIO stands for Biophysical (LAI<sub>eff</sub>)
- SENSOR = Landsat-8
- YYYYMMDD = Campaign date
- Site = Guangdong
- ETF stands for Empirical Transfer Function
- Area = window size 20x20 km<sup>2</sup> and 5x5 km<sup>2</sup>

**Table 8: Content of the dataset.**

Parameter	Dataset name	Range	Variable Type	Scale Factor	No Value
LAI effective	LAI <sub>eff</sub>	[0, 7]	Integer	1000	-1
Quality Flag	QFlag	0,1,2,3 (*)	Integer	N/A	-1

(\*) 0 means extrapolated value (low confidence), 1 strict interpolator (best confidence), 2 large interpolator (medium confidence), 3 water mask.

## 7. CONCLUSIONS

The FP7 ImagineS project continues the innovation and development activities to support the operations of the Copernicus Global Land service. Guangdong, in Xuwen (China), is one of the JECAM sites where LAI measurements over paddy rice fields were collected in 2013. The ground data was shared with ImagineS for the validation of Copernicus Global Land products.

This report first presents the ground data collected during a multi-temporal field experiment from April to June, 2013. The dataset includes 25 elementary sampling units that provide LA<sub>eff</sub> values estimated with LAI-2000 to characterize the rice fields in the study area.

Secondly, high resolution ground-based maps of the biophysical variables have been produced over the site. Ground-based maps have been derived using high resolution imagery (Landsat 8) according with the CEOS LPV recommendations for validation of low resolution satellite sensors. Transfer function has been derived by multiple robust regressions between ESUs reflectance and the effective LAI. The spectral bands combination to minimize RMSE and errors (weighted RMSE and cross-validation RMSE) were band 1 (green), band 2 (red) band 3 (Near Infrared) and band 4 (Short Wave Infrared) combination. The RMSE value estimated is 0.6 for LA<sub>eff</sub>, with no bias.

The quality flag map based on the convex-hull analysis shows very good quality for the validation area (70% at 5x5 km<sup>2</sup>). There is a large area in the contours of the image (20x20km<sup>2</sup>) corresponding to different land use far away from the sampled rice area, where the transfer function behaves as extrapolator. However, the results obtained in the map seem to be reliable.

The biophysical variable maps are available in geographic (UTM 19 South projection WGS-84) coordinates at 30 m resolution. Mean values and standard deviation for LA<sub>eff</sub> was computed over an area of 3x3 km<sup>2</sup>, centered at the selected validation area.

## 8. ACKNOWLEDGEMENTS

This work is supported by the FP7 IMAGINES project under Grant Agreement N°311766. Landsat-8 HR imagery is provided through the USGS Global Visualization service. This work is done in collaboration with the consortium implementing the Global Component of the Copernicus Land Service.

Thanks to the Chinese Academy of Agricultural Engineering, people from Guangdong site and JECAM for providing ground dataset.

## 9. REFERENCES

- Baret, F. and Fernandes, R. (2012). Validation Concept. VALSE2-PR-014-INRA, 42 pp.
- Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., and Weiss, M. (2013). GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment*, 137: 310-329.
- Demarez, V., Duthoit, S., Baret, F., Weiss, M. and Dedieu, G. (2008). Estimation of leaf area and clumping indexes of crops with hemispherical photographs. *Agricultural and Forest Meteorology*. 148, 644-655.
- Fernandes, R., Plummer, S., Nightingale, J., et al. (2014). Global Leaf Area Index Product Validation Good Practices. CEOS Working Group on Calibration and Validation - Land Product Validation Sub-Group. *Version 2.0: Public version made available on LPV website*.
- Martínez, B., García-Haro, F. J., & Camacho, F. (2009). Derivation of high-resolution leaf area index maps in support of validation activities: Application to the cropland Barrax site. *Agricultural and Forest Meteorology*, 149, 130–145.
- Miller, J.B. (1967). A formula for average foliage density. *Aust. J. Bot.*, 15:141-144
- Morisette, J. T., Baret, F., Privette, J. L., Myneni, R. B., Nickeson, J. E., Garrigues, S., et al. (2006). Validation of global moderate-resolution LAI products: A framework proposed within the CEOS land product validation subgroup. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1804–1817.
- Weiss, M., Baret, F., Smith, G.J., Jonckheere, I. and Coppin, P., (2004). Review of methods for in situ leaf area index (LAI) determination. Part II. Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology*. 121, 37–53.
- Weiss M. and Baret F. (2010). CAN-EYE V6.1 User Manual
- Welles, J.M. and Norman, J.M., 1991. Instrument for indirect measurement of canopy architecture. *Agronomy J.*, 83(5): 818-825.