IMPLEMENTING MULTI-SCALE AGRICULTURAL INDICATORS EXPLOITING SENTINELS

VEGETATION FIELD DATA AND PRODUCTION OF GROUND-BASED MAPS

"PSHENICHNE SITE, UKRAINE"

MULTI-TEMPORAL CAMPAIGN: APRIL - JULY, 2015

ISSUE I1.00

EC Proposal Reference N° FP7-311766

Actual submission date : April 2016

Start date of project: 01.11.2012　　　　　　　　　　　Duration : 40 months

Name of lead partner for this deliverable: EOLAB

Book Captain: María del Carmen Piñó (EOLAB)

Contributing Authors: Consuelo Latorre (EOLAB), Fernando Camacho (EOLAB),

Nataliia Kussul, Serhiy Skakun, Andrii Kolotii, Andrii Shelestov (NAS-SSAU)

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:	María del Carmen Piñó	Date: 01.04.2016	Sign.
Approval:	R. Lacaze	Date: 07.06.2016	Sign.
Endorsement:	M. Koleva	Date:	Sign.
Distribution:	Public		

CHANGE RECORD

Issue/Revision	Date	Page(s)	Description of Change	Release
	01.04.2016	All	First Issue	I1.00

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

CEOS	Committee on Earth Observation Satellite
CEOS LPV	Land Product Validation Subgroup
DG AGRI	Directorate General for Agriculture and Rural Development
DG RELEX	Directorate General for External Relations (European Commission)
DHP	Digital Hemispheric Photographs
ECV	Essential Climate Variables
EUROSTATS	Directorate General of the European Commission
ESU	Elementary Sample Unit
FAPAR	Fraction of Absorbed Photo-synthetically Active Radiation
FAO	Food and Agriculture Organization
FCOVER	Fraction of Vegetation Cover
GCOS	Global Climate Observing System
GEO-GLAM	Global Agricultural Geo- Monitoring Initiative
GIO-GL	GMES Initial Operations - Global Land (GMES)
GCOS	Global Climate Observing System
GMES	Global Monitoring for Environment and Security
GPS	Global Positioning System
IMAGINES	Implementing Multi-scale Agricultural Indicators Exploiting Sentinels
JECAM	Joint Experiment for Crop Assessment and Monitoring
LAI	Leaf Area Index
LDAS	Land Data Assimilation System
LUT	Look-up-table techniques
NAS	Space Research Institute National Academy of Sciences
SSA	State Space Agency
NSPI	Neighborhood Similar Pixel Interpolator
PAI	Plant Area Index
PROBA-V	Project for On-Board Autonomy satellite, the V standing for vegetation.
RMSE	Root Mean Square Error
SPOT /VGT	Satellite Pour l'Observation de la Terre / VEGETATION
SLT	Solar Local Time
TOC	Top of Canopy Reflectance
USGS	U.S. Geological Survey Science organization.
UNFCCC	United Nations Framework Convention on Climate Change
UTM	Universal Transverse Mercator coordinates system
VALERI	Validation of Land European Remote sensing Instruments
WGCV	Working Group on Calibration and Validation (CEOS)

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 intend 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 (http://land.copernicus.eu/global/).

The demonstration of high resolution (30m) products (Landsat-8 + PROBA-V) was done over demonstration sites of cropland and grassland in contrasting climatic and environmental conditions. Demonstration products are available on the ImagineS website (http://www.fp7-imagines.eu/pages/services-and-products/landsat-8-biophysical-products.php)

1.3. SCOPE AND OBJECTIVES

The main objective of this document is to describe the ground data collected by Integration-Plus for FP7 ImagineS project (Sub-contract FP7-311766-15/1.1), and the processing carried out by EOLAB to derive high resolution maps of the following biophysical variables:

- 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 two different LAI quantities (for green elements):
  - ➢ The effective LAI (LAIeff) derived from the description of the gap fraction as a function of the view zenith angle. In addition, effective LAI measures derived at 57.5^o are also provided in the ground database.
  - ➢ The actual LAI (LAI)corrected from the clumping index.
- Fraction of green Vegetation Cover (FCover), defined as the proportion of soil covered by vegetation, derived from the gap fraction between 0 and 10° of view zenith angle.
- Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), which is the fraction of the photosynthetically active radiation (PAR) absorbed by a vegetation canopy. We are also focused on green elements. PAR is the solar radiation reaching the canopy in the 0.4–0.7 µm wavelength region. We focused on the instantaneous 'black-sky' FAPAR at 10:00h Solar Local Time (SLT), which is the FAPAR under direct illumination conditions at a given solar position. In addition, two other quantities are provided: daily integrated FAPAR computed as the black-sky FAPAR integrated over the day and the 'white-sky' FAPAR, which is the FAPAR under diffuse illumination conditions.

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.

1.5. RELATED DOCUMENT

- ImagineS_RP7.5_FieldCampaign_Pshenichne2013: Field campaign and Data Processing report of the measurements collected in 2013 over Pshenichne site.

- ImagineS_RP7.5_FieldCampaign_Pshenichne2014: Field campaign and Data Processing report of the measurements collected in 2014 over Pshenichne site.

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 LAI 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 (Morisette 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. This will be 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 will be compared with the medium resolution satellite product at the spatial support of the product.

One of the Imagines demonstration sites selected to support the validation of Copernicus Global Land is a JECAM site located in Pshenichne, Ukraine. In the framework of ImagineS project, Integration-Plus in collaboration with the Space Research Institute NAS and the State Space Agency SSA Ukraine has carried out a multi-temporal field campaign (Table1) to characterize the vegetation biophysical parameters at the test site of Pshenichne. The field campaigns and processing of the collected ground data was founded by ImagineS (Sub-contract FP7-311766-15/1.1).

Multi-temporal Field Campaign: From 22th April to 23rd July, 2015.

Table 1: Ground Campaign dates and cloud-free Landsat imagery available.

CAMPAIGN	DATES	IMAGERY
First campaign	22th April 2015	Not available
Second campaign	3rd June 2015	Not available
Third campaign	23rd June 2015	LANDSAT8 TOA (25.06.2015)
Fourth campaign	7th July 2015	LANDSAT8 TOA (25.06.2015)
Fifth campaign	23rd July 2015	LANDSAT7 TOA (19.07.2015)

Teams involved in field collection: Natalia Kussul, Skakun Serhiy, Andrii Kolotii, Andrii Shelestov.

Contact: Natalia Kussul (kussul@mail.ru)

3. STUDY AREA

3.1. LOCATION

The experimental site is located around Pshenichne farm, in the region of Kiev, 50 km away from the capital (Figure 1). Ground measurements were conducted over selected fields located on the side of Pshenichne. The coordinates of the test site are shown in the Table 2.
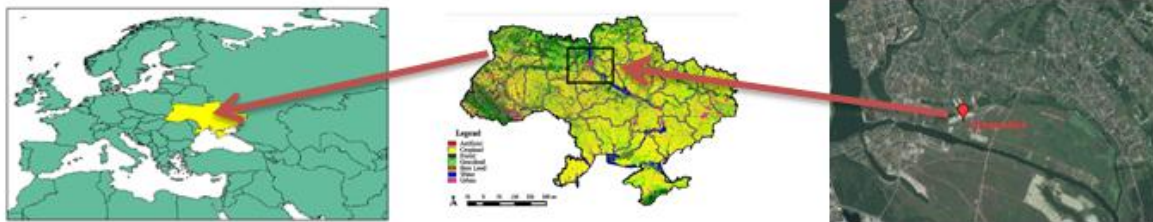
Figure 1: Location of Pshenichne site, Ukraine.

Table 2: Coordinates and altitude of the test site (centre).

Site Center	
Geographic Lat/lon, WGS-84 (degrees)	Latitude = 50.0765º E Longitude = 30.2322º N
Altitude	200 m

3.2. DESCRIPTION OF THE TEST SITE

The types of crops that can be found in the region of Pshenichne are winter wheat, maize, soybean and winter rapeseed. There is not a typical simple crop rotation in this region. Most producers use different crop rotations depending on specialization.

Figure 2 shows the test site of Pshenichne, where the different measured plots are shown for the fifth campaign over an NDVI-TOA image. The locations of ESUs for each campaign are shown in the Figure 3.
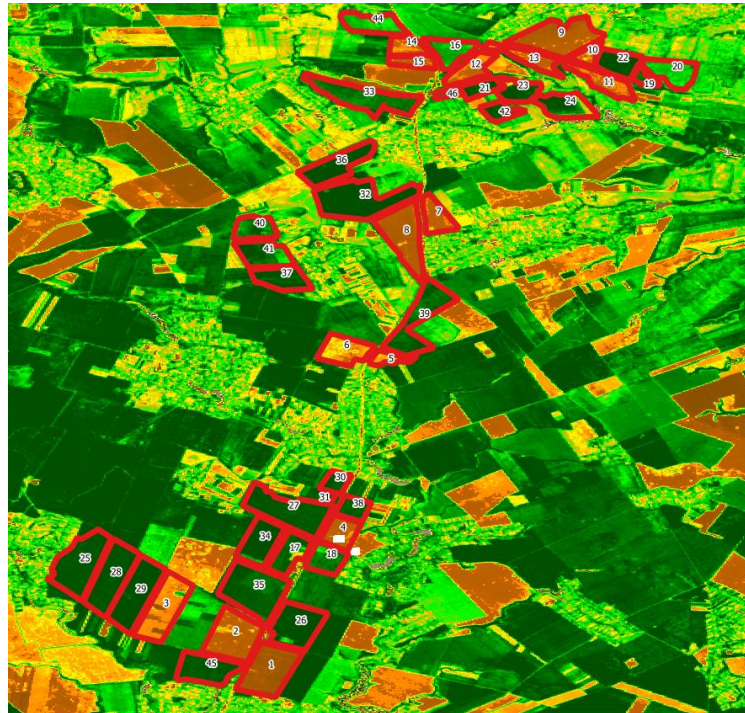
Figure 2: The Pshenichne test site, with the location of the measured plots for the fifth campaign (23rd July, 2015). Background: NDVI TOA Landsat-8 image.

4. GROUND MEASUREMENTS

The ground measurements were acquired and processed by Integration-Plus for the FP7 ImagineS project (under sub-contract FP7-311766-15/1.1).

4.1. MATERIAL AND METHODS

Digital Hemispherical Photographs (DHP) were acquired with a NIKON D70 and a CANON 550D digital cameras. Hemispherical photos allow the calculation of LAI, FAPAR and FCOVER measuring gap fraction trough an extreme wide-angle camera lens (i.e. 180º) (Weiss et al; 2004). It produces circular images that record the size, shape, and location of gaps, either looking upward from within a canopy or looking downward from above the canopy.

The hemispherical photos acquired during the field campaign were processed with the CAN-EYE software version 6.4 (http://www6.paca.inra.fr/can-eye) to derive LAI, FAPAR and FCOVER. It is based on a RGB color classification of the image to discriminate vegetation elements from background (i.e., gaps). This approach allows exploiting downward-looking photographs for short canopies (background = soil) as well as upward-looking photographs for tall canopies (background = sky). CAN-EYE software processes simultaneously up to 20 images acquired over the same ESU. Note that the N images were acquired with similar illumination conditions to limit the variation of color dynamics between images.

The processing is achieved in 3 main steps (Weiss et al., 2004). First, image pre-processing is performed, which includes removing undesired objects (e.g. operator, sun glint) and image contrast adjustments to ensure a better visual discrimination between vegetation elements and background. Second, an automatic classification (k-means clustering) is applied to reduce the total number of distinctive colours of the image to 324 which is sufficient to ensure accurate discrimination capacities while keeping a small enough number of colours to be easily manipulated. Finally, a default classification based on predefined colour segmentation is first proposed and then iteratively refined by the user. The allocation of the colours to each class (vegetation elements versus background) is the most critical phase that needs to be interactive because colours depend both on illumination conditions and on canopy elements. At the end of this process a binary image, background versus vegetation elements (including both green and non-green elements) is obtained.

The CAN-EYE software computes biophysical variables from gap fraction as follows:

Effective LAI (LAIeff): Among the several methods described in Weiss et al (2004), the effective LAI estimation in the CAN-EYE software is performed by model inversion. The effective LAI is estimated from the Plant Area Index (PAI) which is the variable estimated by CAN-EYE, as no distinction between leaves or other plant elements are made from the gap fraction estimates. PAI is very close to the effective LAI for croplands when pictures are taken downward looking, whereas larger discrepancies are expected for forest when pictures are taken upward looking. Effective LAI is directly retrieved by inverting Eq. (1) (Poisson

model) and assuming an ellipsoidal distribution of the leaf inclination using look-up-table (LUT) techniques.

$$P_0(\theta_v, \varphi_v) = e^{-N \cdot (\theta_v, \varphi_v)} = e^{-G \cdot (\theta_v, \varphi_v) \cdot \frac{LAI_{eff}}{\cos(\theta_v)}}$$ Eq. (1)

A large range of random combinations of LAI (between 0 and 10, step of 0.01) and ALA (Average Leaf Angle) (10º and 80º, step of 2º) values is used to build a database made of the corresponding gap fraction values (Eq.1) in the zenithal directions defined by the CAN-EYE user (60º for the DHP collection in this field campaign). The process consists then in selecting the LUT element in the database that is the closest to the measured P_0. The distance (cost function C_k) of the kth element of the LUT to the measured gap fraction is computed as the sum of two terms. The first term computes a weighted relative root mean square error between the measured gap fraction and the LUT one. The second term is the regularization term that imposes constraints to improve the PAI estimates. Two equations are proposed for the second "regularization" term:

(1) constraint used in CAN-EYE V5.1 on the retrieved ALA values that assume an average leaf angle close to 60º ± 03º, and

(2) constraint used in CAN-EYE V6.1 on the retrieved PAI value that must be close from the one retrieved from the zenithal ring at 57º. This constraint is more efficient, but it can be computed only when the 57º ring is available (i.e., COI≥60º)

The software also proposed other ways of computing PAI and ALA effective using Miller's formula (Miller, 1967) which assumed that gap fraction only depends from view zenith angle. Furthermore, the CAN-EYE makes an estimation using the Welles and Norman (1991) method used in LAI-2000 for 5 rings. These LAI2000-like estimates were not used here as are based on the same Miller's formula but using limited angular sampling.

LAI: The actual LAI that can be measured only with a planimeter with however possible allometric relationships to reduce the sampling, is related to the effective leaf area index through:

$$LAI_{eff} = \lambda_0 \cdot LAI$$ Eq. (2)

Where λ_0 is the clumping index. In CAN-EYE, the clumping index is computed using the Lang and Xiang (1986) logarithm gap fraction averaging method, although some uncertainties are associated to this method (Demarez et al., 2008). The principle is based on the assumption that vegetation elements are locally assumed randomly distributed. Values of clumping index given by CAN-EYE are in certain cases correlated with the size of the cells used to divide photographs.

*As the CAN-EYE software provides different results (CEV6.1, CEV5.1 and Miller's) for LAIeff and LAI variables; an average LAI value was provided as ground estimate, and the standard deviation of the different method LAI estimates was reported as the uncertainty of the estimate (see associated **2015_VGM_Pshenichne.xls** file). Note that for LAI, only CEV6.1 and CEV5.1 were used.*

FCOVER is retrieved from gap fraction between 0 to 10°.

$$FCOVER = 1 - P_0 \cdot (0 - 10°)$$ Eq. (3)

The standard deviation calculated over different pictures is provided by the CAN-EYE software. The FCover along with the standard deviation (uncertainty) is provided in the field data file (2015_VGM_Pshenichne.xls).

FAPAR: As there is little scattering by leaves in that particular spectral domain due to the strong absorbing features of the photosynthetic pigments, FAPAR is often assumed to be equal to FIPAR (Fraction of Intercepted Photosynthetically Active Radiation), and therefore directly related to the gap fraction. The actual FAPAR is the sum of two terms, weighted by the diffuse fraction in the PAR domain: the 'black sky' FAPAR that corresponds to the direct component and the 'white sky' or the diffuse component.

The instantaneous "Black-sky FAPAR" (FPAR[BS]) is given at a solar position (date, hour and latitude). Depending on latitude, the CAN-EYE software computes the solar zenith angle every solar hour during half the day (there is symmetry at 12:00). The instantaneous FAPAR is then approximated at each solar hour as 1 minus the gap fraction in the corresponding solar zenith angle:

$$FAPAR^{BS}(\theta_S) = 1 - P_0 \cdot (\theta_S)$$ Eq. (4)

The daily integrated "black sky" (or direct) FAPAR is computed as the following:

$$FAPAR_{Day}^{BS} = \frac{\int_{sunset}^{sunrise} cos(\theta_S) \cdot [1 - P_0 \cdot (\theta_S)] \cdot d\theta}{\int_{sunset}^{sunrise} cos(\theta_S) \cdot d\theta}$$ Eq. (5)

The "white-sky" (or diffuse) FAPAR is computed as the following:

$$FAPAR^{WS} = \frac{1}{\pi} \int_0^{2\cdot\pi} \int_0^{\frac{\pi}{2}} P_0 cos(\theta_S) sin(\theta_S) d\theta d\varphi = 2 \cdot \int_0^{\frac{\pi}{2}} P_0 cos(\theta_S) sin(\theta_S) d\theta$$ Eq. (6)

The CAN-EYE software provides the three FAPAR variables. Instantaneous black-sky FAPAR values at 10:00h SLT were up-scaled. No uncertainty is provided on the FAPAR estimates.

4.2. SPATIAL SAMPLING SCHEME

A pseudo-regular sampling was used within each ESU of approximately 30x30 m^2. The center of the ESU was geo-located using a Global Positioning System (GPS). A total of 31 ESUs for the first campaign, 27 ESUs for the second and the third campaigns, and finally 28 ESUs for the fourth and fifth campaigns were characterized. At the first campaign, winter crops (rapeseed and wheat) were the predominant harvest. In the last campaigns (end of June-July), they were replaced by Maize and Soybean. The distribution of sampling units per land cover type for the five different campaigns is shown in the able 3: . The spatial sampling over the test site for the different campaigns is shown in Figure 3.
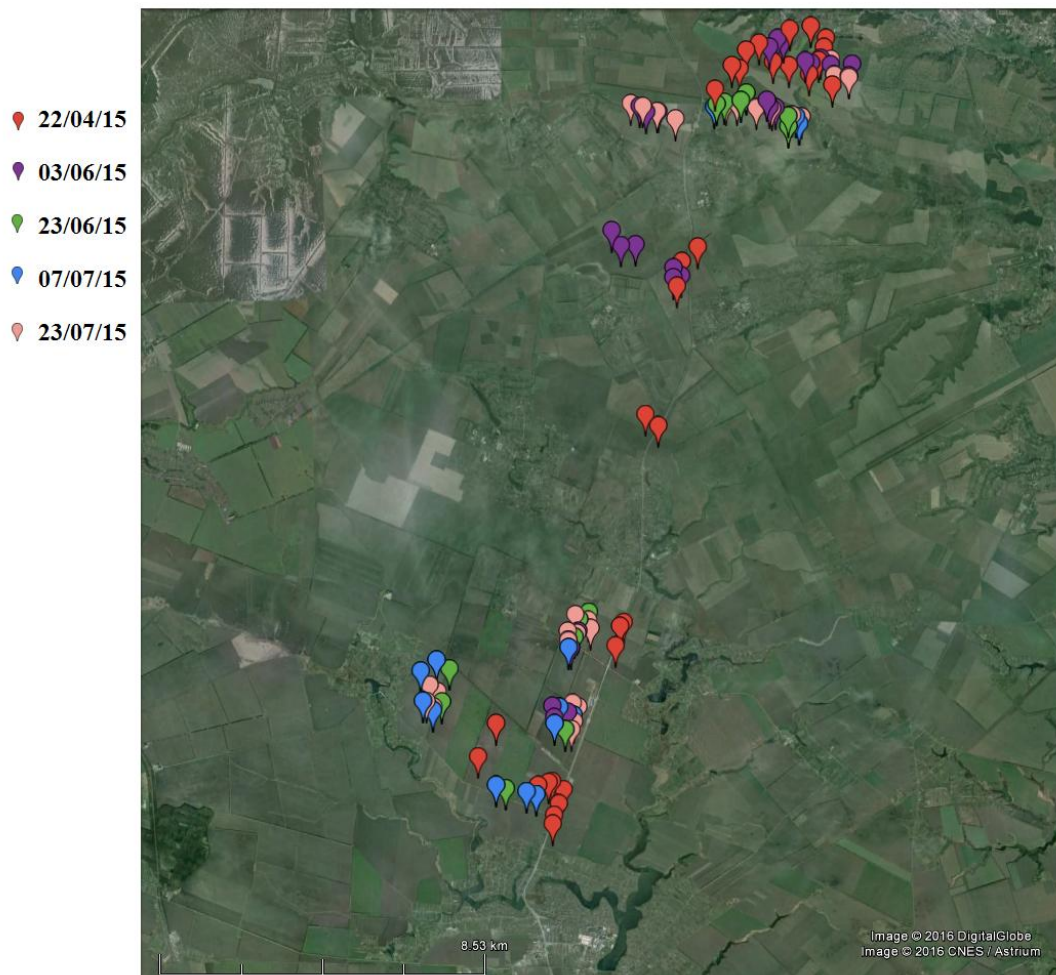
Figure 3: Location of the ESUs over the Pshenichne site in Google Earth for the five campaigns of 2015.

able 3: Distribution of ESUs per land cover type in Pshenichne, 2015.

Land Use	Number of ESUs				
	First Campaign	Second Campaign	Third campaign	Fourth campaign	Fifth campaign
Winter Rapeseed	6	0	0	0	0
Winter Wheat	25	9	0	0	0
Soybean	0	9	9	13	13
Maize	0	9	18	15	14

Erreur ! Source du renvoi introuvable. summarizes the distribution of the ESUs in the study area per crop type acquired during the five field campaigns, where the percentage of each vegetation type sampled is shown.
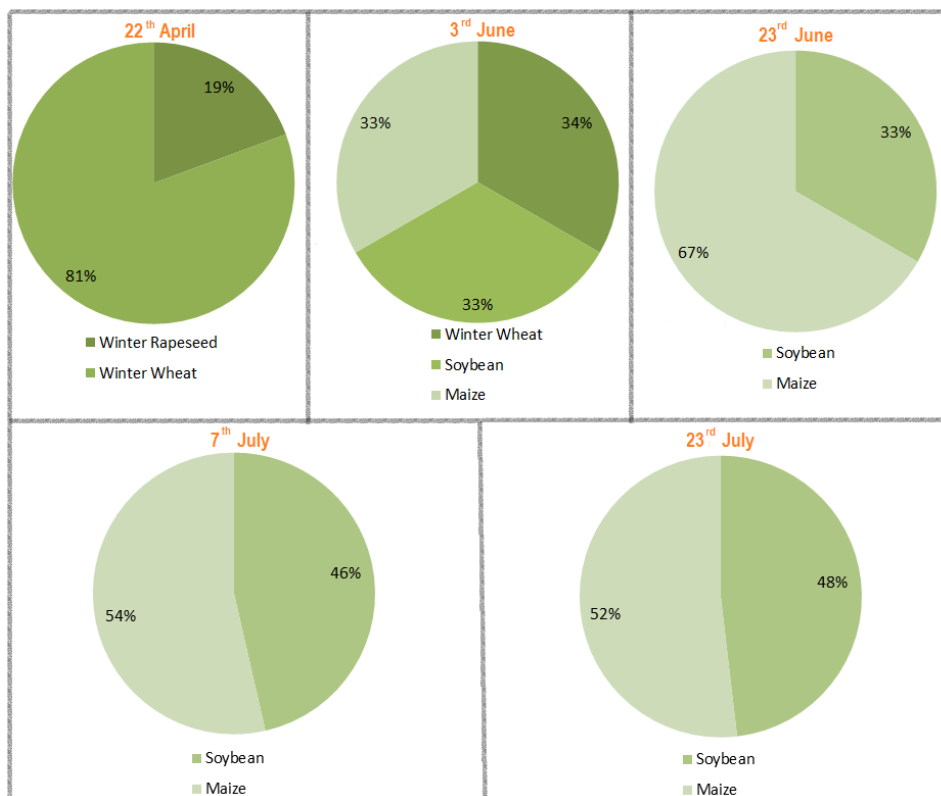
Figure 4: Percentage of land cover type sampled in Pshenichne, 2015.

4.3. GROUND DATA

4.3.1. Data processing

As said above the data processing was carried out by Integration-Plus. EOLAB preformed the quality control of the ground data set.

Firstly, a visual inspection of the CAN-EYE report, including the hemispherical photos, the classification, and the result was undertaken to look for suspicious ESUs (due to problems in the image or poor classification of the soil/vegetation). This was marked in the database as "suspicious" ESU. Two examples are shown below:

Figure 4 shows an example of the DHP collection over an ESU of maize. The outcome of the processing is quite influenced by the saturated image (marked in red in Figure 4). This image was removed to make the DHP sampling more consistent, and then the ESU was re-processed.
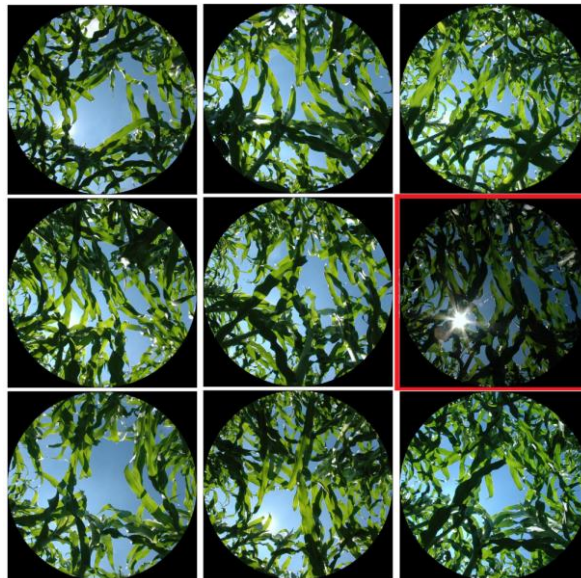
Figure 4: ESU 12 (MA35B) over Pshenichne, Ukraine, 2015. Third field campaign.

Figure 5 shows an example of suspicious value due to the misclassification. As can be readily observed in the image, the brightest parts of the leaves are miss-classified as soils, which lead to underestimation of retrieved values.

Underestimated

Figure 5: ESU 6 (MA28A) over Pshenichne, Ukraine, 2015. Third field campaign. Left: Original picture. Right: Vegetation (green) /soil (brown) classification.

Secondly, the consistency among the different variables was checked according to the expected exponential trend (LAI vs FAPAR) and linear trend (FAPAR vs FCOVER) (Figure 6). Those ESUs were the differences between FAPAR and FCOVER were larger than 0.25 (FAPAR > FCOVER) were investigated as suspicious of FCOVER underestimation (e.g. due to the small FOV of the DHP image used for FCOVER estimation). This was marked as "possible FCOVER underestimation".
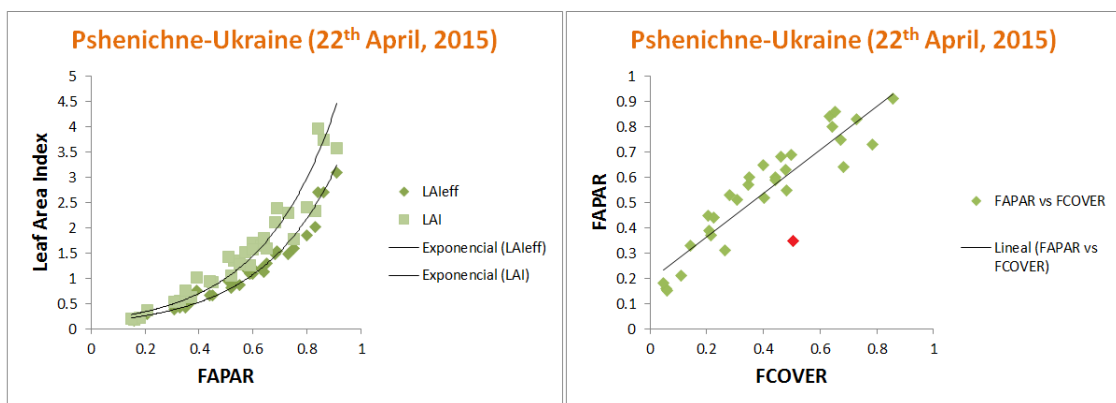
Figure 6: Inter-comparison of the measured biophysical variables. LAI versus FAPAR (Left) and FAPAR versus FCOVER (Right). Pshenichne, Ukraine, 2015.

Finally, clumping index values were also assessed over very homogeneous and dense crops canopies were clumping index should be very close to 1. Lower values of the clumping

index could be indicative of uncertainties in the classification of the DHP (due to shaded areas). These ESUs were marked as "very homogeneous, clumping should be close to 1". The reliability of the LAI values in these ESUs is lower. Figure 7 shows on example of very homogeneous canopies (the DHP picture and its classification made by CAN-EYE). For this type of homogeneous canopies the clumping index should be close to unity, and however the CAN-EYE outcome is 0.77, according to the soil/vegetation image. As can be observed, areas classified as soils corresponds shaded vegetation, and then the retrieved values (LAIeff, FAPAR, FCOVER) are underestimated. The lower values of the clumping index (0.77) tends to correct the underestimation in the LAI (LAIeff/clumping), but the uncertainty of the retrieved LAI is also large due to misclassification of vegetation as soil. 7 ESUs (of 28) during the fourth campaign and 6 (of 28) ESUs during the fifth campaign showed this issue which could introduce underestimations (mainly for LAIeff, FAPAR and FCOVER) in the ground dataset used for the transfer function.
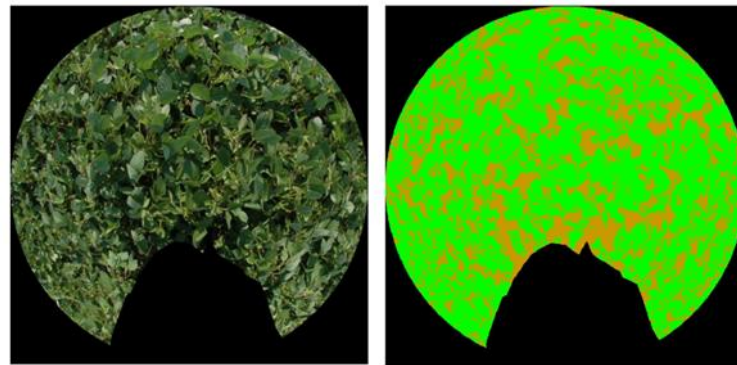
Figure 7: ESU 28 (SO33C) over Pshenichne, Ukraine, 2015. Fourth field campaign. Left: Original picture. Right: Vegetation and soil classification.

As a result of the quality control performed by EOLAB several "suspicious" or "possible FCOVER underestimation" ESUs were re-processed by EOLAB. Both values (Integrated-Plus and EOLAB) are keep in the database. The suspicious ESUs and values re-processed by EOLAB are indicated in the database. Re-processed values by EOLAB were included in the database only when different than the original values (provided by Integrated-Plus).

Finally, few additional ground control points (GCP in the database) were added to the ground dataset to better constrain the empirical transfer function.

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. All ground values can be found in the associated file "**2015_VGM_Pshenichne.xlsx**"

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 [1]
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+4*j (*)	LAI	Method	Instrument
12+4*j (*)		Nb. Replications	Number of Replications
13+4*j (*)		PRODUCT	Methodology
14+4*j (*)		Uncertainty	Standard deviation

*LAIeff, LAI, FAPAR and FCOVER

Erreur ! Source du renvoi introuvable. shows the distribution of LAIeff, LAI, FAPAR and FCOVER values for the multi-temporal campaign (the ground measurements per ESU can be found in the Annex I). During the first campaign (22th of April), LAI values are typically lower than 2.5 except for few ESUs of winter wheat (up to 4), whereas FCOVER and FAPAR show distribution of values ranging between 0 and 1. For the second campaign (3rd of June), values are concentrated in low or high values (LAI typically lower than 1 or higher than 4.5, and FAPAR typically lower than 0.4 or higher than 0.9). The exception is the FCOVER with distribution of values ranging typically between 0 and 0.7. This is explained due to the under-estimation of the FCOVER for dense winter wheat (ESUs 1 to 6 in the excel file), where FCOVER << FAPAR. For the third campaign, LAI values are typically around 3, FAPAR around 0.7 and FCOVER around 0.6. For the fourth and fifth campaigns, larger values are obtained, with LAI up to 5, and FAPAR and FCOVER values typically on 0.9.
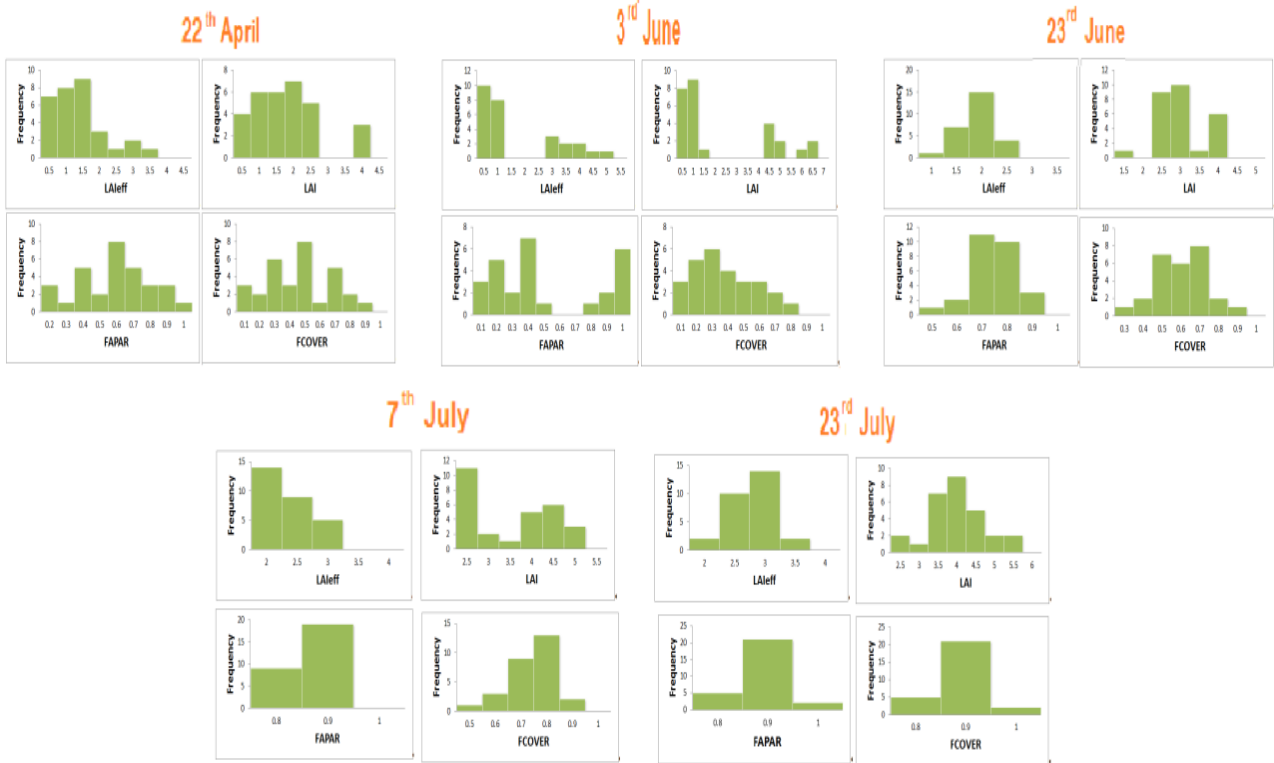
Figure 8: Distribution of the measured biophysical variables for the multi-temporal field campaigns. Pshenichne, 2015.

5. EVALUATION OF THE SAMPLING

5.1. EVALUATION BASED ON NDVI VALUES

The sampling strategy is evaluated using the Landsat-8 and Landsat-7 (see Table 5) TOA images by comparing the NDVI distribution over the site with the NDVI distribution over the ESUs. 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 (module the size of the image)

2. Computing the cumulative frequency of NDVI on the randomly shifted sampling design

3. Repeating steps 1 and 2, 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 - α = 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 9 shows that the ESUs NDVI TOA distribution is good over the whole site (i.e., comprised between the highest and lowest cumulative frequencies) for the third (23rd June) and fourth (7th July) campaigns. However, for the fifth campaign (23rd July) the distribution of the sampling is biased towards highest NDVI values for low and high vegetation cover.
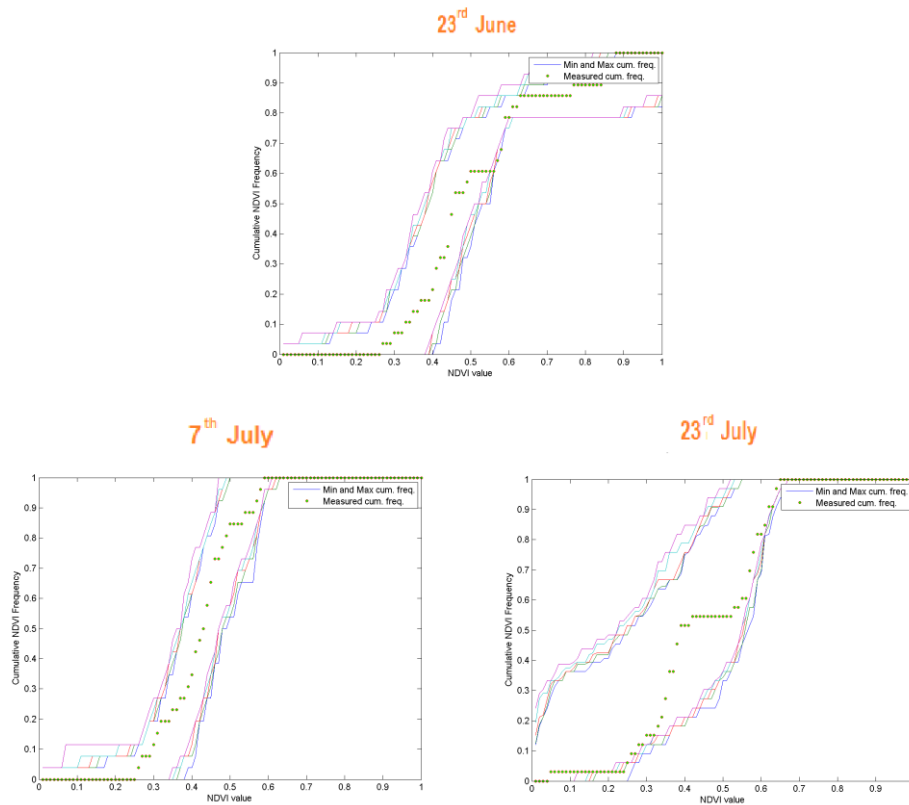
Figure 9: Comparison of NDVI TOA distribution between ESUs and over the whole image, Pshenichne-Ukraine, 2015 (no cloud-free images were available for the first and second campaigns).

5.2. 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 of the Landsat-7 and Landsat-8 TOA images. 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 TOA 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 (± 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 10 shows the results of the Convex-Hull test (i.e., Quality Flag images) for the Pshenichne site over the 5x5 km^2 study area and the extended 20x20 km^2 area. For the 5x5 km^2 study area, the percentage of good interpolation confidence of the transfer function goes up to 49% for the third campaign, 48% for the fourth and 51% for the last campaign (Table 4). For the extended area (20x20 km^2), the percentage of good interpolation confidence of the transfer function goes up to 37% for the third, 36% for the fourth and 54% for the last campaign (Table 4).

23rd June	7th July	23rd July
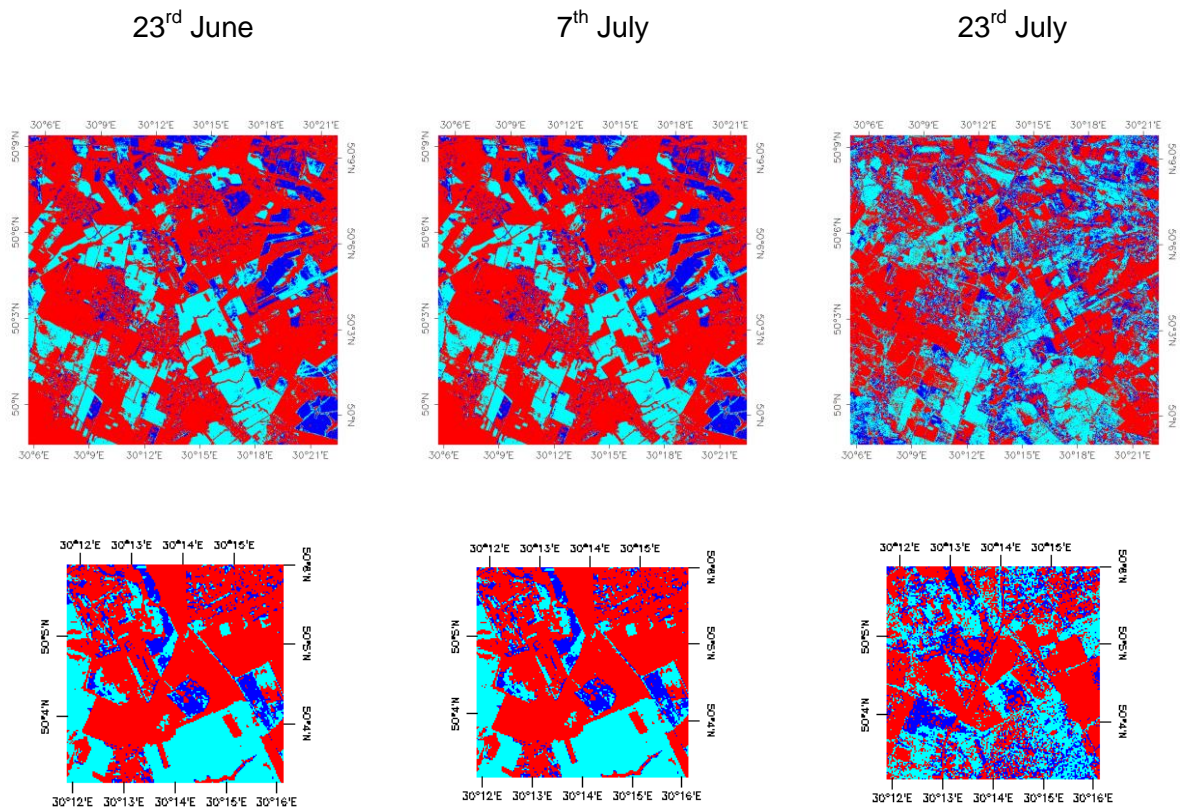
Figure 10: Convex Hull test over 20x20 km^2 (Top) and 5x5 km^2 (Bottom) area over Pshenichne site, Ukraine. (2015). Clear and dark blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls. Red corresponds to the pixels for which the transfer function behaves as extrapolator.

Table 4: Percentages over the 20x20 km^2 and 5x5 km^2 area over the test site of Pshenichne (Ukraine) Convex hull values: 0=extrapolation of TF, 1=strict convex hull and 2=large convex hull.

Field Campaigns	Quality Flags (%)							
Size	20x20 km^2				5x5 km^2			
DATE	0	1	2	1&2	0	1	2	1&2
23rd June, 2015	64	24	13	37	51	40	9	49
7th July, 2015	64	23	13	36	52	39	9	48
23rd July, 2015	49	34	17	51	46	34	20	54

6. PRODUCTION OF GROUND-BASED MAPS

6.1. IMAGERY

The Landsat-8 (OLI) image acquired the 25th of June was used for both the third and fourth campaigns. As no cloud-free Landsat-8 image was available for the other campaigns, a Landsat-7 image was used instead for the last campaign. No cloud freed Landsat imagery was available for the other campaigns. The Landsat-7 (ETM+) image was acquired the 19th July 2015 (see Table 5 for acquisition properties). Four spectral bands were selected from 500 nm to 1750 nm with a nadir ground sampling distance of 30 m. The original projection is UTM 36 North, WGS-84.

As all Landsat-7 scenes collected since May of 2003 have data gaps, it is needed to perform a correction. A number of methods have been used to fill the gaps of Landsat-7 data. Based on the assumption that the same-class neighboring pixels exhibit similar patterns of spectral differences between dates, we used a simple and effective method to interpolate the values of the pixels within the gaps. This method is the Neighborhood Similar Pixel Interpolator (NSPI). Results indicate that NSPI can restore the value of un-scanned pixels very accurately, and that it works especially well in heterogeneous regions (Chen et al., 2011; Latorre et al., 2014). Figure 11 show a vertical profile of the Landsat-7 image, before and after the gap filling correction. Note that this gap filling correction is only for Landsat-7 images.
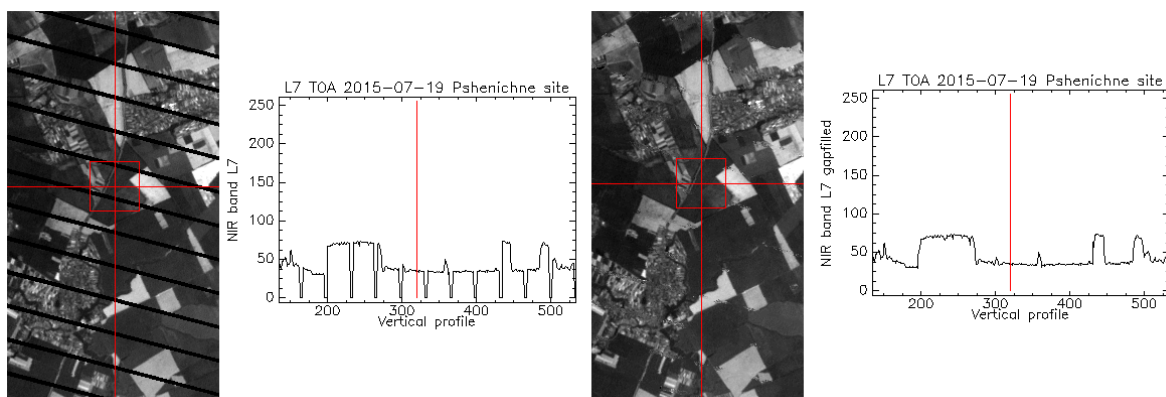

Figure 11: Vertical profile over the Pshenichne site, for NIR band of original Landsat-7 image (Left) and gap filled Landsat-7 image (Right). 23rd July, 2015.

Table 5: Acquisition properties of Landsat-8 and Landsat-7 data used for retrieving high resolution maps.

Pshenichne	Landsat-8 METADATA		Landsat-7 METADATA
Platform / Instrument	Landsat-8 / OLI_TIRS		Landsat-7/ETM+
Path	181		181
Row	25		25
Selected Bands	B3(green) : 0.53-0.59 μm		B2(green) : 0.52-0.60 μm
	B4(red) : 0.64-0.67 μm		B3(red) : 0.63-0.69 μm
	B5(NIR) : 0.85-0.88 μm		B4(NIR) : 0.77-0.90 μm
	B6(SWIR1) : 1.58-1.65 μm		B5(SWIR1) : 1.55-1.75 μm
Field Campaigns	Third campaign	Fourth Campaign	Fifth Campaign
	23rd June, 2015	7th July, 2015	23rd July, 2015
Acquisition date	25/06/2015	25/06/2015	19/07/205
	16:04:02	16:04:02	18:46:23
Illumination Azimuth angle	146.261°	146.261°	146.693°
Illumination Elevation angle	59.763°	59.763°	57.142°
Ground control points verify	167	167	151
Geometric RMSE Verify	5.981	5.981	4.925

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

As the method has limited extrapolation capacities, a flag image for each transfer function (Figure 10) is included in the ground based maps, in order to inform the users on the confidence of the transfer function estimates.

6.2.2. Band combination

Figure 12 shows the errors (RW, RC) obtained for the several band combinations using TOA reflectance. The selected combination for all the variables and images is: NDVI TOA combination.

In spite of the band combination of [NDVI] doesn't give the lowest values of RC and RW in all the transfer functions, this combination was selected since it provides similar errors than the best combination, but the NDVI assures good consistency of the maps over the whole area, including those regions where measurements were not taken. For instance over bare areas or senescent regions, the NDVI assures low values, as well as consistency among variables (LAI, FAPAR, FCOVER). This is not the case for other band combinations, where inconsistencies among variables are obtained for regions outside the convex hull (low confidence of the transfer function).

Figure 12: Test of multiple regressions (TF) applied on different band combinations. 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.

6.2.3. The selected Transfer Function

The applied transfer function is detailed in **Erreur ! Source du renvoi introuvable.**, along with its weighted and cross validated errors.

For the FAPAR and FCOVER, a linear relationship with NDVI was selected:

$$FAPAR = a + b \cdot NDVI$$ Eq. (7)

$$FCOVER = a + b \cdot NDVI$$ Eq. (8)

For the LAIeff and LAI, an exponential relationship with NDVI was selected according to Baret et al., (1989):

$$LAIeff = a + b \cdot ln\left(\frac{NDVI_\infty - NDVI}{NDVI_\infty - NDVI_S}\right)$$ Eq. (9)

$$LAI = a + b \cdot ln\left(\frac{NDVI_\infty - NDVI}{NDVI_\infty - NDVI_S}\right)$$ Eq. (10)

Where *b* represents the extinction coefficient which depends on the average leaf angle inclination, solar zenith angle and diffuse reflectance and transmittance of the leaves. "b" was set empirically with the ground data for each transfer function, as well as the residuals "a". NDVIs represents the typical NDVI of bare soil areas and NDVI∞ represents the NDVI of fully developed canopies, both assumed to be constant over the image. NDVIs were set to 0.22 and NDVI∞ to 0.62 for the third and fourth campaigns and NDVIs was set to 0.14 and NDVI∞ to 0.68 for the fifth campaign.

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

Variable	Band Combination	RW	RC
23rd June, 2015		**Third Campaign**	
LAIeff	$0.305 - 1.247 \cdot ln\left(\frac{0.62-NDVI}{0.62-0.22}\right)$	0.22	0.22
LAI	$0.856 - 1.246 \cdot ln\left(\frac{0.62-NDVI}{0.62-0.22}\right)$	0.68	0.65
FAPAR	$- 0.311 + 2.08 \cdot NDVI$	0.05	0.04
FCOVER	$- 0.303 + 1.777 \cdot NDVI$	0.11	0.10
7th July, 2015		**Fourth Campaign**	
LAIeff	$0.597 - 1.202 \cdot ln\left(\frac{0.62-NDVI}{0.62-0.22}\right)$	0.36	0.34
LAI	$0.625 - 2.332 \cdot ln\left(\frac{0.62-NDVI}{0.62-0.22}\right)$	0.95	0.87
FAPAR	$- 0.284 + 2.270 \cdot NDVI$	0.06	0.05
FCOVER	$- 0.240 + 2.00 \cdot NDVI$	0.09	0.08
23rd July, 2015		**Fifth Campaign**	
LAIeff	$0.385 - 1.534 \cdot ln\left(\frac{0.68-NDVI}{0.68-0.14}\right)$	0.33	0.36
LAI	$0.462 - 2.353 \cdot ln\left(\frac{0.68-NDVI}{0.68-0.14}\right)$	0.50	0.50
FAPAR	$- 0.024 + 1.608 \cdot NDVI$	0.06	0.06
FCOVER	$- 0.013 + 1.489 \cdot NDVI$	0.12	0.12

Figure 13 shows scatter-plots between ground observations and their corresponding transfer function (TF) estimates for the selected bands combinations.
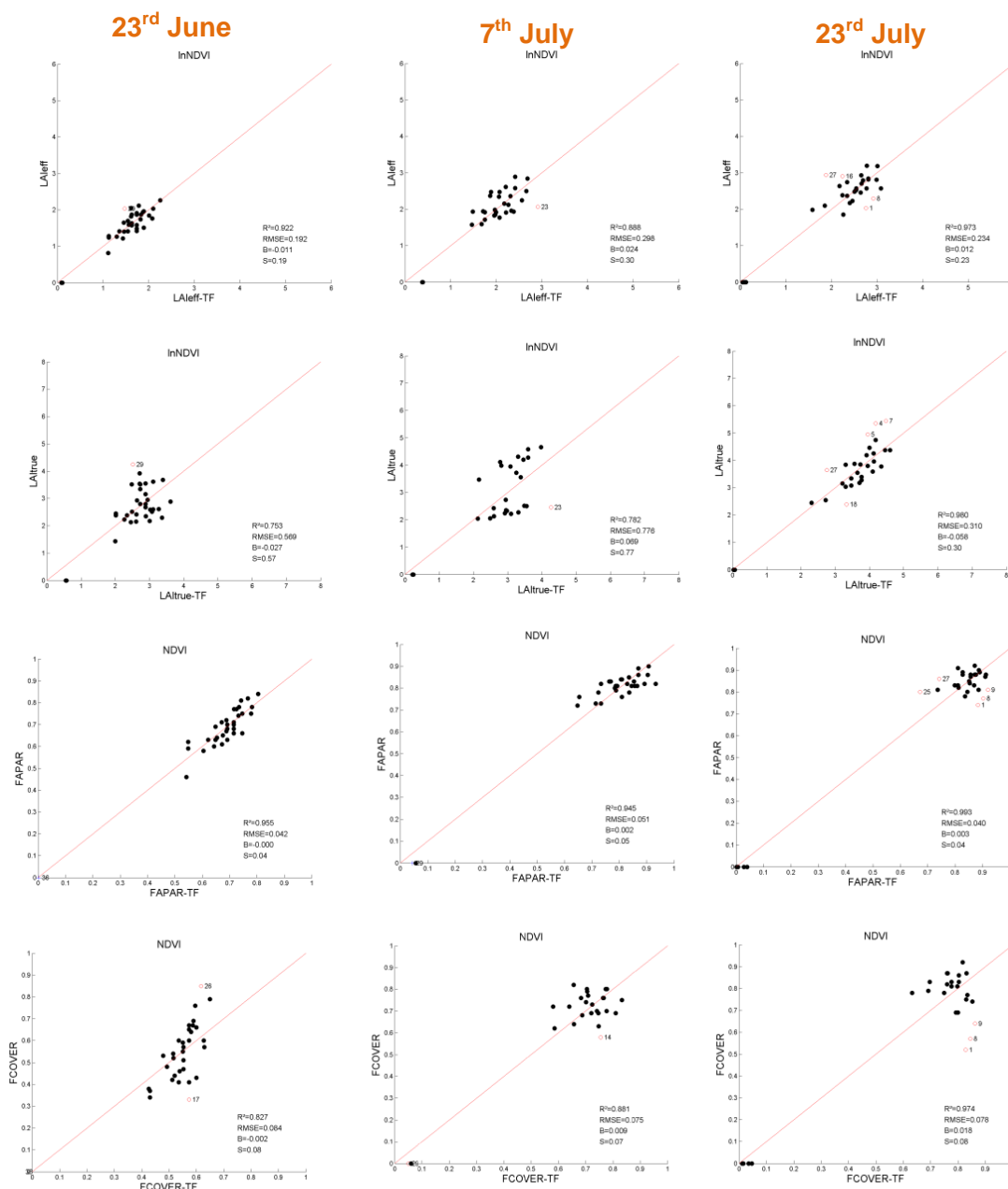
Figure 13: LAIeff, LAI, FAPAR and FCOVER results for regression on reflectance using NDVI combination. Full dots: Weight>0.7. Empty dots: 0<Weight<0.7.

A good correlation is observed for all the variables with points distributed along the 1:1 line and no mean bias, and small scattering. Better correlation and lower errors are obtained for FAPAR and LAIeff, whereas some scattering is observed for LAI and FCOVER.

The RMSE values of all biophysical variables are shown in the Table 7. Note the very good RMSE values for FAPAR and LAIeff within the GCOS requirements on accuracy for LAI, FAPAR satellite products (i.e., LAI max(0.5, 20%), FAPAR max(0.05, 10%)).

Table 7: RMSE values obtained from the Scatter-Plots for LAIeff, LAI, FAPAR and FCOVER.

Pshenichne	RMSE			
	LAIeff	LAI	FAPAR	FCOVER
23rd June, 2015	0.19	0.57	0.04	0.08
7th July, 2015	0.30	0.78	0.05	0.08
23rd July, 2015	0.23	0.31	0.04	0.08

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 (and Landsat-7) TOA reflectance. Figure 14, Figure 15, Figure 16 Figure 17 present the empirical TF biophysical maps over the extended 20x20 km^2 area and the 5x5 km^2 area for LAIeff, LAI, FAPAR (10:00 SLT) and FCOVER, respectively.

Figure 10 shows the Quality Flags included in the final product. The temporal and spatial variation of the different plots can be readily observed.
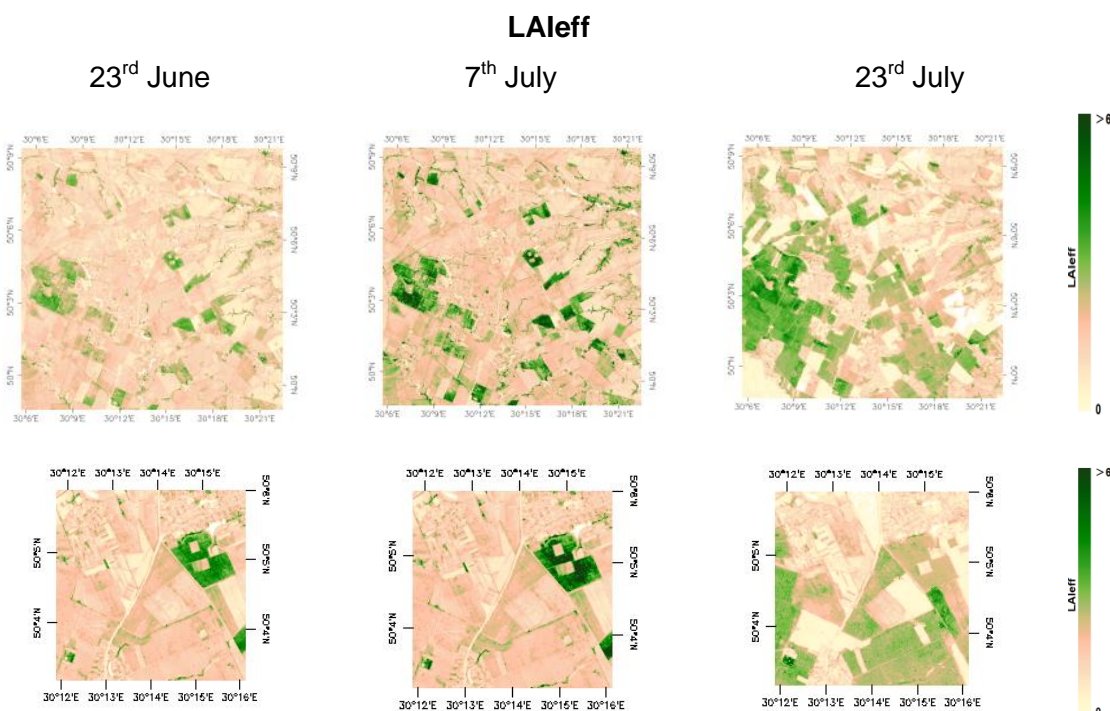

Figure 14: High resolution ground-based LAIeff maps obtained for the Pshenichne site (2015). Top: 20x20 km^2 area. Bottom: 5x5 km^2 area.
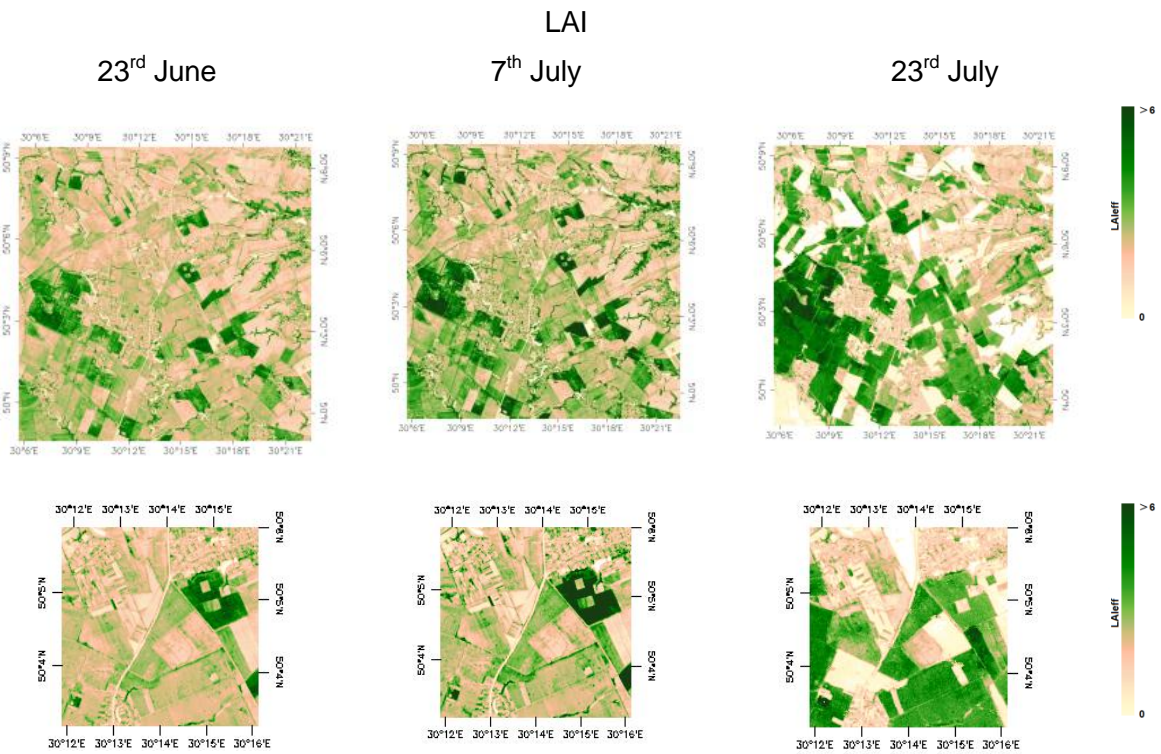
LAI

23rd June   7th July   23rd July

Figure 15: High resolution ground-based LAI maps obtained for the Pshenichne site (2015). Top: 20x20 km^2 area. Bottom: 5x5 km^2 area.
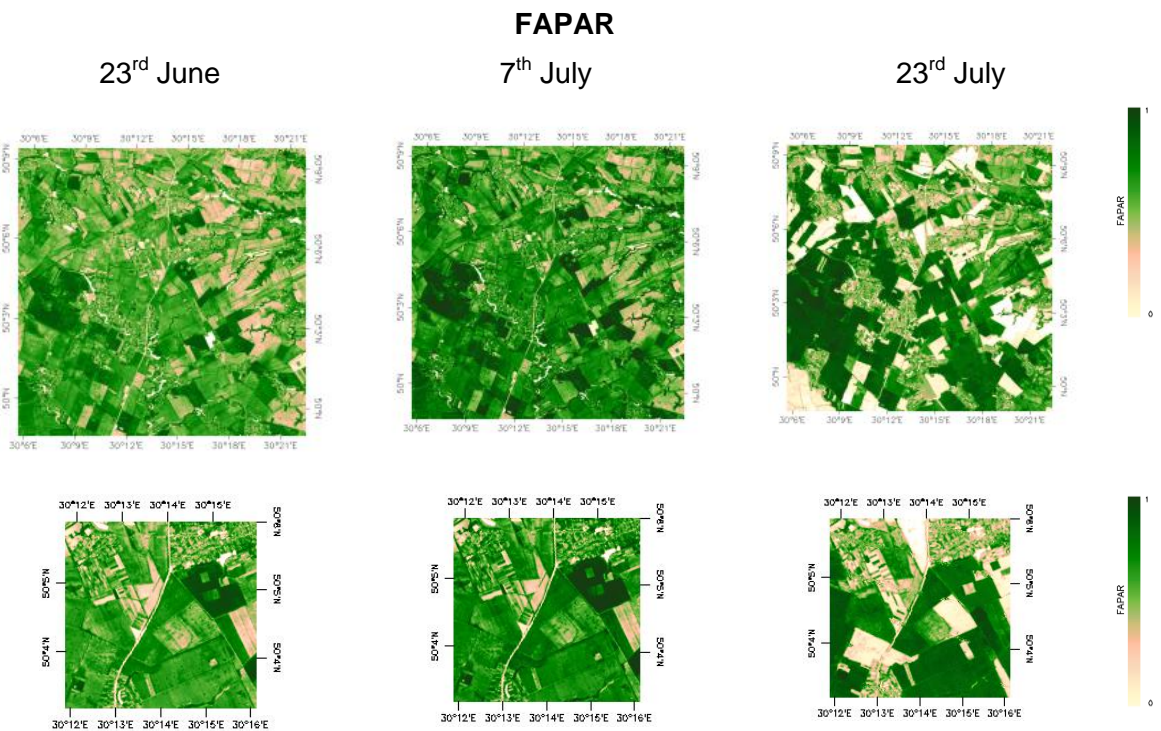
FAPAR

23rd June   7th July   23rd July

Figure 16: High resolution ground-based FAPAR (10:00h SLT) maps obtained for the Pshenichne site (2015). Top: 20x20 km^2 area. Bottom: 5x5 km^2 area.
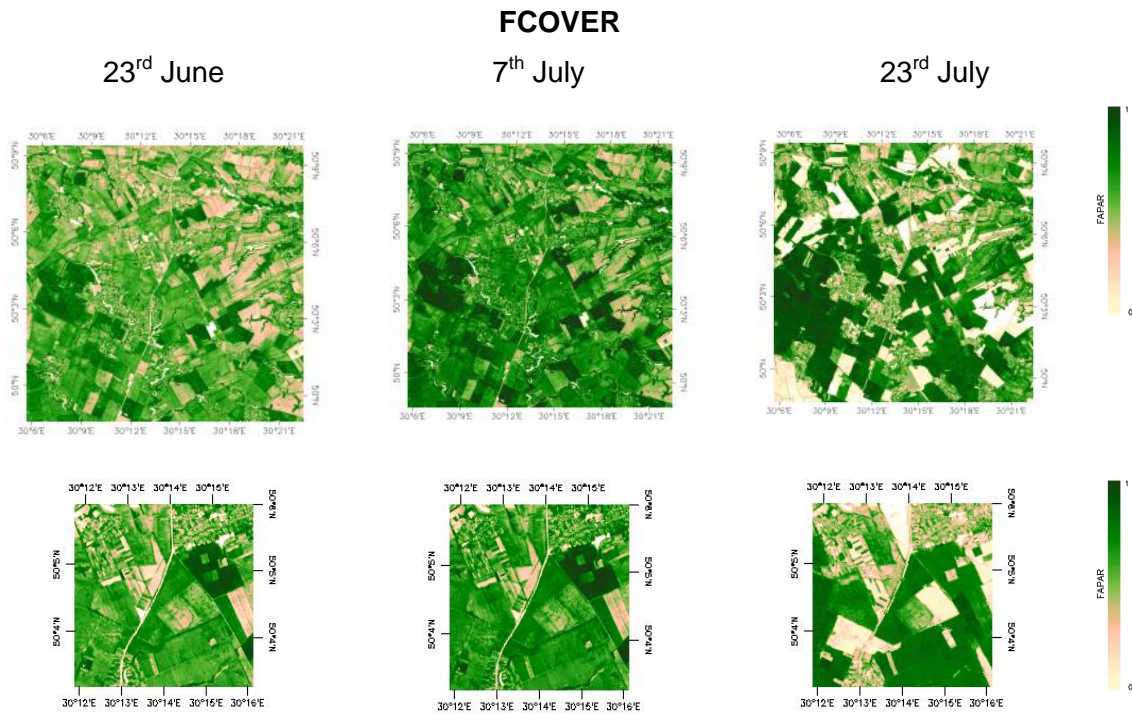
FCOVER

23rd June   7th July   23rd July

Figure 17: High resolution ground-based FCOVER maps obtained for the Pshenichne site (2015). Top: 20x20 km^2 area. Bottom: 5x5 km^2 area.

Figure 18 shows scatters plots between the retrieved maps that show the good consistency of the ground-based maps, showing the exponential (LAI vs FAPAR) and linear (FAPAR vs FCOVER) trend observed with the ground data.
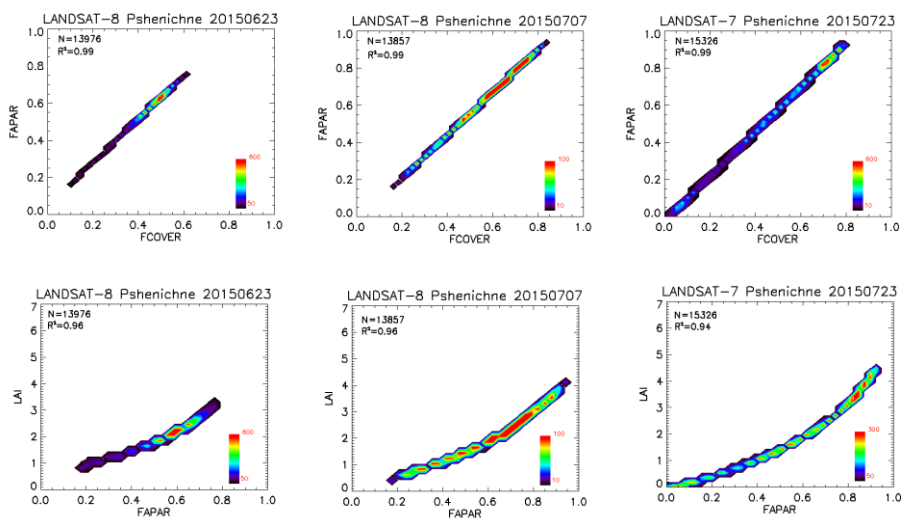

Figure 18: Scatter plots to LAI vs FAPAR and FAPAR vs FCOVER for all the 3 campaigns over Pshenichne-Ukraine.

6.3.1. Mean Values

Mean values of a 3x3 km^2 area centred in the test site are provided for validation of 1 km satellite products to reduce co-registration and PSF errors, and in agreement with the CEOS OLIVE direct dataset (Table 8). 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^2, and 20x20 km^2.

Table 8: Mean values and standard deviation (STD) of the HR biophysical maps for the selected 3 x 3 km^2 area at Pshenichne site (Ukraine)

Pshenichne 3x3 km^2	LATITUDE				LONGITUDE			
	50.0765º N				30.2322º E			
	Mean Values				STDV Values			
	LAIeff	LAI	FAPAR	FCOVER	LAIeff	LAI	FAPAR	FCOVER
23rd June, 2015	1.37	2.36	0.58	0.46	0.64	0.91	0.17	0.15
7th July, 2015	1.86	2.61	0.69	0.62	1.02	1.37	0.19	0.17
23rd July, 2015	1.47	2.12	0.56	0.53	0.92	1.41	0.29	0.27

Table 9 describes the content of the geo-biophysical maps in the nomenclature: "BIO_YYYYMMDD_SENSOR_Site _ETF_Area"

Where:

BIO stands for Biophysical (LAIeff, LAI, FAPAR and FCOVER)

SENSOR = LANDSAT8 or LANDSAT7

YYYYMMDD = Acquisition date

Site = Pshenichne

ETF stands for Empirical Transfer Function

Area = 20x20 and 5x5

Table 9: Content of the dataset.

Parameter	Dataset name	Range	Variable Type	Scale Factor	No Value
LAI effective	LAIeff	[0, 7]	Integer	1000	-1
LAI	LAI	[0, 7]	Integer	1000	-1
FAPAR (Daily)	FAPAR	[0, 1]	Integer	10000	-1
Fraction of Vegetation Cover	FCOVER	[0, 1]	Integer	10000	-1
Quality Flag	QFlag	0,1,2 (*)	Integer	N/A	-1

[*] 0 means extrapolated value (low confidence), 1 strict interpolator (best confidence), 2 large interpolator (medium confidence)

7. CONCLUSIONS

The FP7 ImagineS project continues the innovation and development activities to support the operations of the Copernicus Global Land service. One of the ImagineS demonstration sites corresponds to the Pshenichne JECAM site, in the Province of Kiev, in Ukraine.

This report presents the ground data collected and processed by Integration-Plus during five intensive field campaigns: 22th of April, 3rd of June, 23rd of June, 7th of July and 23rd of July 2015. The dataset includes 31, 27, 27, 28 and 28 elementary sampling units, respectively, where digital hemispherical photographs were taken and processed with the CAN-EYE software to provide LAI, LAIeff, FAPAR and FCOVER values to characterize the major cultivated vegetation of the area: winter rapeseed, winter wheat, soybean and maize.

High resolution ground-based maps of the biophysical variables were produced over the site. Ground-based maps were derived using high resolution imagery (Landsat-8 TOA in the third and fourth campaigns and Landsat-7 TOA in the fifth campaign) according with the CEOS LPV recommendations for validation of low resolution satellite sensors. Transfer functions were derived by multiple robust regressions between ESUs reflectance and the several biophysical variables. The band combination of [NDVI] was selected for the transfer function to assure good consistency of the retrieved values over the whole region and periods. The RMSE values for the transfer function estimates are ranging between 0.19 and 0.30 for LAIeff, 0.31 and 0.78 for LAI, 0.04 and 0.05 for FAPAR, and finally 0.08 for FCOVER, with no mean bias. Remarkably good is the performance for LAIeff and FAPAR.

The quality flag maps based on the convex-hull analysis show acceptable quality around the study area. The percentages corresponding to good interpolation capabilities for the 5x5 km^2 study area are 49%, 48% and 51% for the third, fourth and fifth campaigns, respectively.

The biophysical variable maps are available in geographic (UTM 32 North projection WGS-84) coordinates at 30 m resolution over the 20x20 km^2 and 5x5 km^2 over the site. Mean values and standard deviation over a validation area of 3x3 km^2 for LAIeff, LAI, FCOVER and FAPAR were computed centered at the validation test site.

8. ACKNOWLEDGEMENTS

9. REFERENCES

Baret, F., G. Guyot and Major, D. (1989). Crop biomass evaluation using radiometric measurements. Photogrammetria 43:241-256.

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.*

Chen, J., Zhu X., Vogelmann J. E., Gao F., Jin S. (2011). A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. Remote sensing of environment. 115: 1053-1064.

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. I48, 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.

Latorre, C., Camacho, F., Sánchez, J., Kussul, N., Serhiy, S., Kravchenko, O., (2013). "Vegetation Field Data and Production of Ground-Based Maps: 14th May, 15th June and 15th July. Pshenichne site, Ukraine" report. (Available at ImagineS website: http://fp7-imagines.eu/pages/documents.php).

Latorre, C., Camacho, F., Piñó, C., Kussul, N., Serhiy, S., Koloti, A., Shelestov, A. (2014). "Vegetation Field Data and Production of Ground-Based Maps: 12th June, 31st July. Pshenichne site, Ukraine" report. (Available at ImagineS website: http://fp7-imagines.eu/pages/documents.php).

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.

10. ANNEX I. GROUND MEASUREMENTS.

The measurements for all field campaigns are shown from Figure 20 to Figure 24.
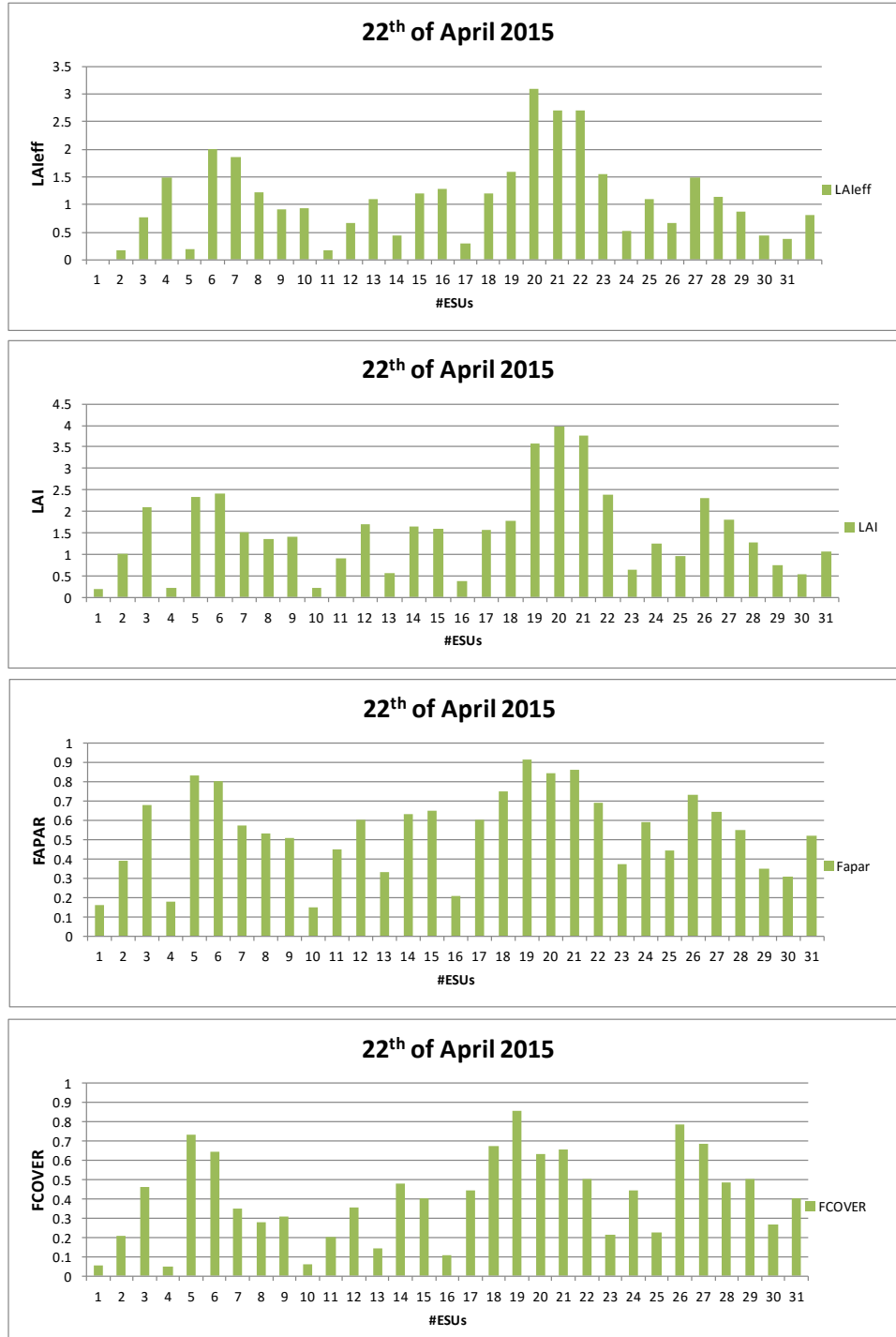
Figure 19: LAIeff, LAI, FAPAR and FCOVER measurements acquired in Pshenichne site, Ukraine. Field campaign 22th April 2015.
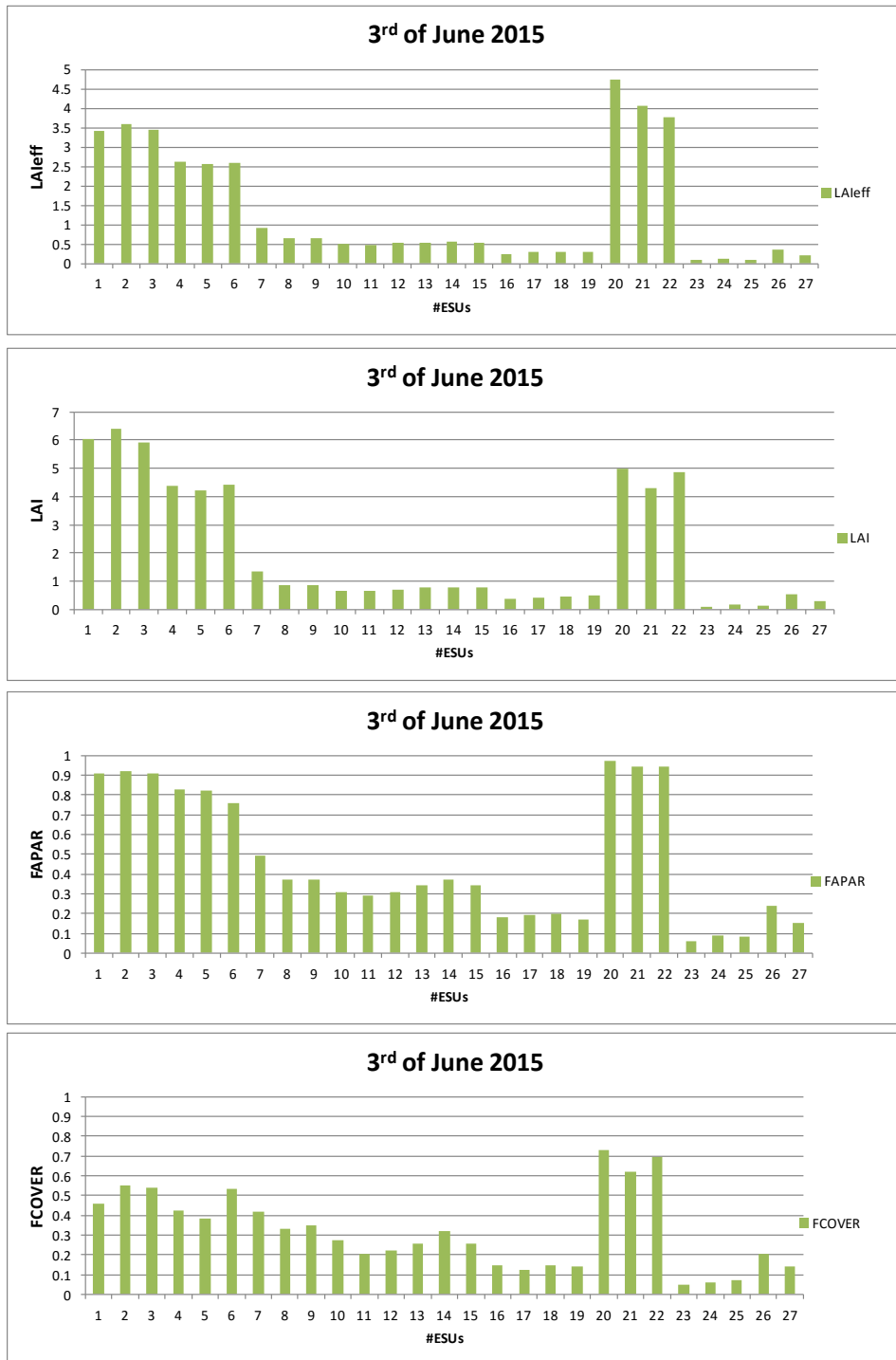
Figure 20: LAIeff, LAI, FAPAR and FCOVER measurements acquired in Pshenichne site, Ukraine. Field campaign 3rd June 2015.
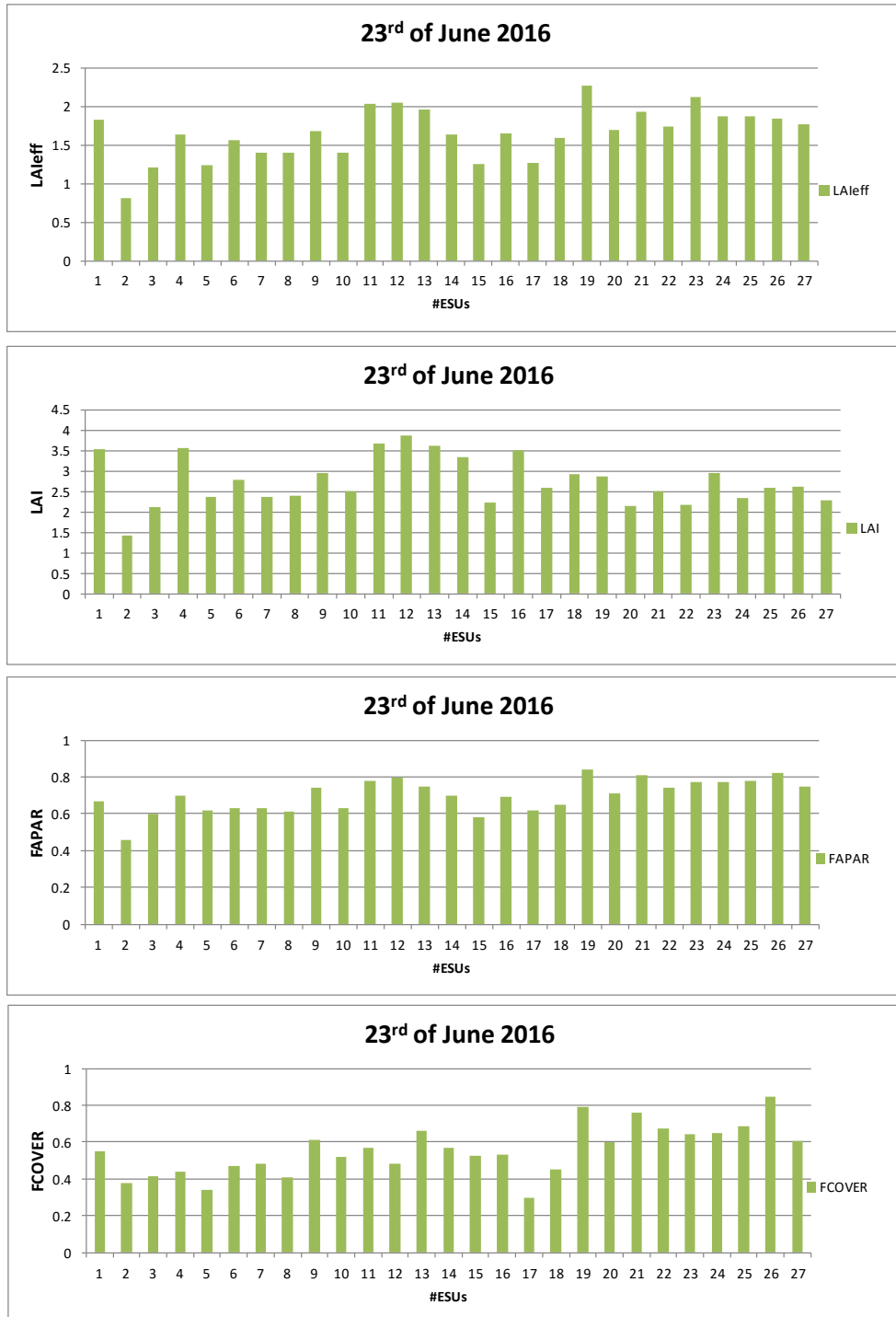
Figure 21: LAIeff, LAI, FAPAR and FCOVER measurements acquired in Pshenichne site, Ukraine. Field campaign 23rd June 2015.
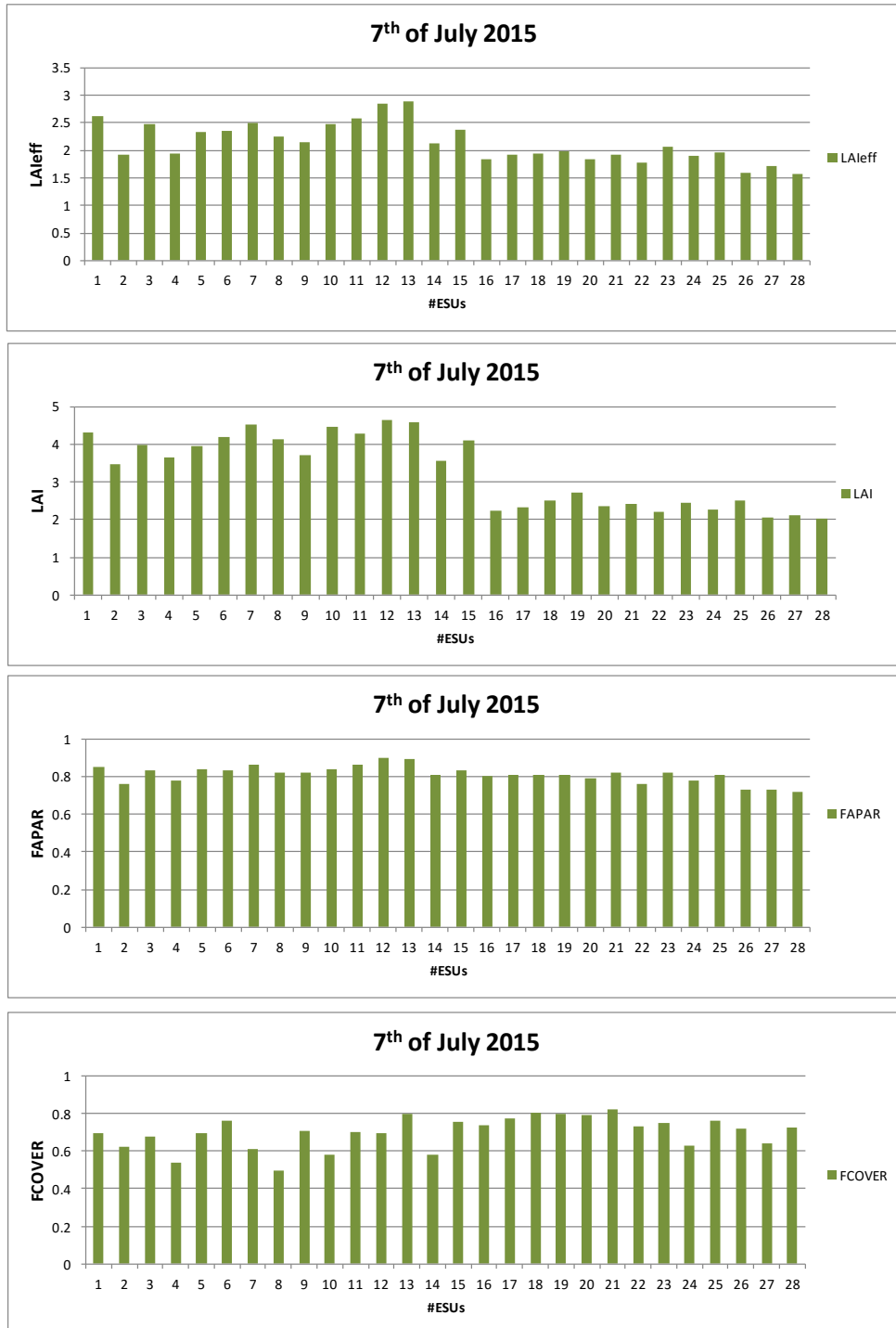
Figure 22: LAIeff, LAI, FAPAR and FCOVER measurements acquired in Pshenichne site, Ukraine. Field campaign 7th July 2015.
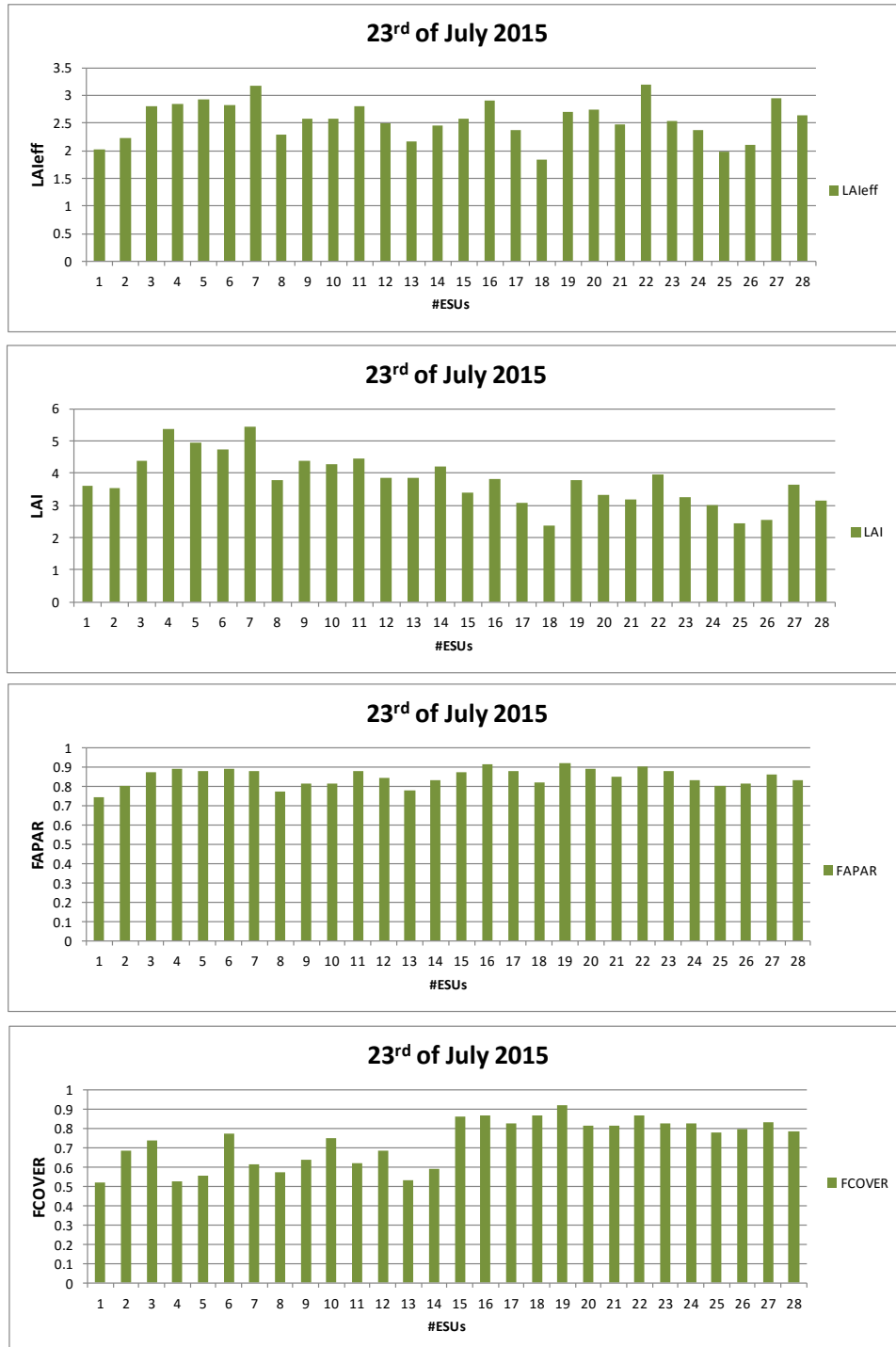
Figure 23: LAIeff, LAI, FAPAR and FCOVER measurements acquired in Pshenichne site, Ukraine. Field campaign 23rd July 2015.