



Multitemporal and multiresolution leaf area index retrieval for operational local rice crop monitoring



Manuel Campos-Taberner^{a,*}, Francisco Javier García-Haro^a, Gustau Camps-Valls^b, Gonçal Grau-Muedra^a, Francesco Nutini^c, Alberto Crema^c, Mirco Boschetti^c

^aDepartament de Física de la Terra i Termodinàmica, Facultat de Física, Universitat de València, Dr. Moliner, 50, Burjassot 46100, València, Spain

^bImage Processing Laboratory (IPL), Universitat de València, Catedrático A. Escardino, 9, Paterna 46980, València, Spain

^cInstitute for Electromagnetic Sensing of the Environment, Italian National Research Council, Via Bassini 15, Milan 20133, Italy

ARTICLE INFO

Article history:

Received 2 May 2016

Received in revised form 16 September 2016

Accepted 1 October 2016

Available online 14 October 2016

Keywords:

Crop monitoring

Rice

Leaf area index (LAI) retrieval

PROSAIL

Smartphone

Gaussian process regression (GPR)

Landsat

SPOT5 Take5

ABSTRACT

This paper presents an operational chain for high-resolution leaf area index (LAI) retrieval from multiresolution satellite data specifically developed for Mediterranean rice areas. The proposed methodology is based on the inversion of the PROSAIL radiative transfer model through the state-of-the-art nonlinear Gaussian process regression (GPR) method. Landsat and SPOT5 data were used for multitemporal LAI retrievals at high-resolution. LAI estimates were validated using time series of *in situ* LAI measurements collected during the rice season in Spain and Italy. Ground LAI data were collected with smartphones using PocketLAI, a specific phone application for LAI estimation. Temporal evolution of the LAI estimates using Landsat and SPOT5 data followed consistently the temporal evolution of the *in situ* LAI measurements acquired on several Mediterranean rice varieties. The estimates had a root-mean-square-error (RMSE) of 0.39 and 0.51 m²/m² in Spain and 0.38 and 0.47 m²/m² in Italy for Landsat and SPOT5 respectively, with a strong correlation ($R^2 > 0.92$) for both cases. Spatial-temporal assessment of the estimated LAI from Landsat and SPOT5 data confirmed the robustness and consistency of the retrieval chain. This paper demonstrates the importance of an adequate characterization of the underlying rice background in order to address changes in background condition related to water management. Results highlight the potential of the proposed chain for deriving multitemporal near real-time decametric LAI maps fundamental for operational rice crop monitoring, and demonstrate the readiness of the proposed method for the processing of data such as the recently launched Sentinel-2.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Green leaf area index (LAI) is a key biophysical parameter which represents half the total intercepting leaf area per unit ground surface area (Chen and Black, 1992). LAI plays an important role in vegetation processes such as photosynthesis and transpiration, and is connected to meteorological/climate and ecological land processes. LAI has been widely used in many agricultural and remote sensing studies (Carlson and Ripley, 1997; Clevers, 1988; Gower et al., 1999; Myneni et al., 1997). Concerning biomass and crop yield estimation, LAI estimates can be assimilated in crop models (Confalonieri et al., 2009) by means of forcing and/or recalibration techniques (Dorigo et al., 2007; Quaife et al., 2008). LAI retrieval from satellite data is among the main goals of the remote sensing community

(Chen et al., 2002; Colombo et al., 2003; Fang and Liang, 2005) as evidenced by the variety and usefulness of operational medium resolution products for vegetation monitoring from satellite sensors, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) (Myneni et al., 2002) and the *Système Pour l'Observation de la Terre* (SPOT) VEGETATION (Baret et al., 2007, 2013). Nevertheless, higher spatial resolutions (10–30 m) are needed to support crop management activities at a parcel level. In this context, the Landsat Data Continuity Mission (LDCM) (Roy et al., 2014) and the recently European Sentinel-2 Mission (Drusch et al., 2012; Malenovsky et al., 2012) provide valuable high-resolution (HR) information for a wide variety of land applications (Malenovsky et al., 2012) including crop monitoring.

From a methodological point of view, LAI retrieval can be faced following either statistical, physical, or hybrid methods (Camps-Valls et al., 2011; Verrelst et al., 2015a; Campos-Taberner et al., 2015b). Parametric statistical methods have been developed through empirical relations between *in situ* LAI acquisitions and vegetation indices

* Corresponding author.

E-mail address: Manuel.Campos@uv.es (M. Campos-Taberner).

derived from airborne and satellite spectra (Haboudane et al., 2004). Alternatively, non-parametric methods do not assume an explicit (parametric) relation between LAI and the spectral reflectance. Non-parametric models estimate a variable of interest using a training database of biophysical parameter and spectral data pairs. In the last decade, non-parametric techniques excelled in biophysical parameter retrieval, either following pure *statistical* or *hybrid* approaches. Many methods have been used in a wide range of applications (Atzberger and Richter, 2012; Verrelst et al., 2015a). Specifically, current operational vegetation products such as LAI are typically produced with neural networks (NN) (Baret et al., 2007, 2013). Nevertheless, in the recent years, Gaussian process regression (GPR) (Rasmussen and Williams, 2006) provided encouraging results in the framework of biophysical parameter estimation outperforming the rest (Camps-Valls et al., 2016; Lazaro-Gredilla et al., 2014; Verrelst et al., 2012a, 2015b).

In statistical approaches, concomitant *in situ* measurements of the biophysical parameter of interest and the associated spectral data from remote sensing platforms are used as a training database, whereas the hybrid approaches rely on a database generated by a radiative transfer model (RTM). The advantage of hybrid approaches is that a broad range of land cover situations can be simulated (e.g. up to hundred thousands), leading to a data set much bigger than what can be collected during a field campaign. RTMs are based on the physical knowledge describing the interactions between radiation, canopy elements, and the soil surface. RTMs have been used for modeling different types of vegetation, making them suitable for general-purpose retrieval applications (Jacquemoud et al., 2000). In particular, the PROSAIL RTM (Jacquemoud et al., 2009), which results from the PROSPECT leaf optical model (Jacquemoud and Baret, 1990) coupled with the SAIL canopy reflectance model (Verhoef, 1984), has been used in several remote sensing studies. PROSAIL has been successfully applied to a variety of crops (Duan et al., 2014). However, it is worth noting that RTM inversion poses several methodological problems: it may lead to high computational cost, and being an ill-posed problem, it may give rise to unstable results. Prior information related to the distribution of the canopy variables and representative background spectra can be implemented in the RTM to better address the inversion process (Combal et al., 2003; Meroni et al., 2004).

Direct validation comparing LAI predictions with *in situ* LAI measurements is needed to report the accuracy of LAI retrievals. LAI ground measurement methods can be divided into two categories: direct and indirect methods (Breda, 2003; Jonckheere et al., 2004). Direct methods involve destructive techniques and require a field effort in collecting leaf samples. Due to the difficulty of continuous applications of direct methods on a large scale, the use of indirect methods based on measurements of the transmission of radiation through the canopy have been widely used (Weiss et al., 2004). In this context, to exploit new technologies in crop monitoring, smartphones have been used for indirect rice LAI measurements reporting good consistency and performance compared with classical instruments such as LAI-2000 or LAI-2200 Plant Canopy Analyzers (LI-COR, Inc., Nebraska, USA) and digital hemispherical photography (DHP) (Campos-Taberner et al., 2016).

In this paper, we present an operational chain for LAI retrieval from Landsat and SPOT5 satellite data specifically calibrated for rice crops. One specific characteristic of rice cultivation is the land preparation followed by the pre-season flooding, this agronomic practice characterizes worldwide the majority of rice cropping systems (namely, irrigated rice, lowland rain-fed rice and deep-water rice) which account for over 90% of the 154 million ha (Maclean, 2002) cultivated with rice each year. In the European temperate rice cultivation areas, the background of the fields is dry at the beginning of the season, and it remains in this condition until early-May when the fields are starting to be flooded. From this date on, the soil

background is flooded most of the time except in some dates when the farmers pump out the water for agronomic purposes. From a remote sensing point of view, this characteristic determines a strong change in the soil background conditions from dry soil to standing water (Boschetti et al., 2014). Thus, the intermittent flooding in a paddy rice field generates an uncontrolled reflectance signature, which may confound the retrieval of rice LAI. It is therefore relevant to test methods for this specific crop due also to its worldwide importance.

In this study, LAI is estimated using hybrid methods through the generation of a reflectance and associated LAI database from the PROSAIL model and powerful nonlinear inversion methods, such as neural networks, kernel ridge regression (KRR) and Gaussian process regression. Based on experimental performance in our datasets, we focus on GPR. This work uses for the first time GPR as the regression tool for multitemporal LAI production chain during the rice cycle. This study shows results produced in the framework of the ERMES project (<http://www.ermes-fp7space.eu/>) where rice monitoring is performed exploiting seasonal remote sensing data and crop modeling as a demonstration of potential operation system. Actually, Landsat data were processed in near real time and the corresponding LAI maps were provided through the web-based geoportal in the project framework to the crop modelers. Contrarily, SPOT5 Take5 data were analyzed in backcasting to assess the potential use of Sentinel-2 data. Spatial and temporal consistency of the LAI estimates were validated with the available multitemporal *in situ* measurements collected in two ERMES study areas (Italy and Spain) from sowing up to rice LAI peak.

The remainder of the paper is organized as follows. Section 2 describes the study areas for which LAI maps are derived, the *in situ* LAI measurements and the remote sensing surface reflectance data. The parameterization of the PROSAIL model, as well as the theoretical basis of the regression methods, are presented in Sections 3.1 and 3.2 respectively. Section 4 evaluates the spatial-temporal consistency of the obtained LAI estimates and provides an analysis of LAI trends for different varieties and management practices as well as discusses the impact of rice background in multitemporal LAI retrievals. Finally, Section 5 outlines the main conclusions and discusses the utility of the obtained results in the framework of operational rice monitoring systems.

2. Data collection

2.1. Study areas

ERMES aims to develop a prototype of COPERNICUS down-stream services based on the assimilation of Earth observation (EO) and *in situ* data within crop modeling solutions dedicated to the rice sector. In this framework, the ERMES study areas have been selected in Spain (Valencian area), Italy (Piedmont and Lombardy rice district) and Greece (Thessaloniki area), which are the three countries responsible of 85% of total European rice production. In this study, we focus on the Italian and Spanish ERMES local study areas. Within each study area, rice is a common crop with a long tradition and economical value. Rice varieties in the Spanish area are mainly *Bomba* and *Senia* belonging to the Japonica group. Sowing activities are concentrated around May 10–15th and fields are managed by keeping them flooded for most of the time during the rice growing period. In Italy, the rice cropping systems are much more variable. About 180 varieties are cultivated covering both Japonica and Indica groups characterized by short and long cycles with a duration from 120 to > 150 days, respectively (Boschetti et al., 2009). For this reason, the sowing date can vary from the beginning of April to mid of May. Moreover, in Italy, the so-called dry seeding technique is increasing year after year. This technique consists of seeding rice in rows with a common seeder in dry soil condition and held without water until

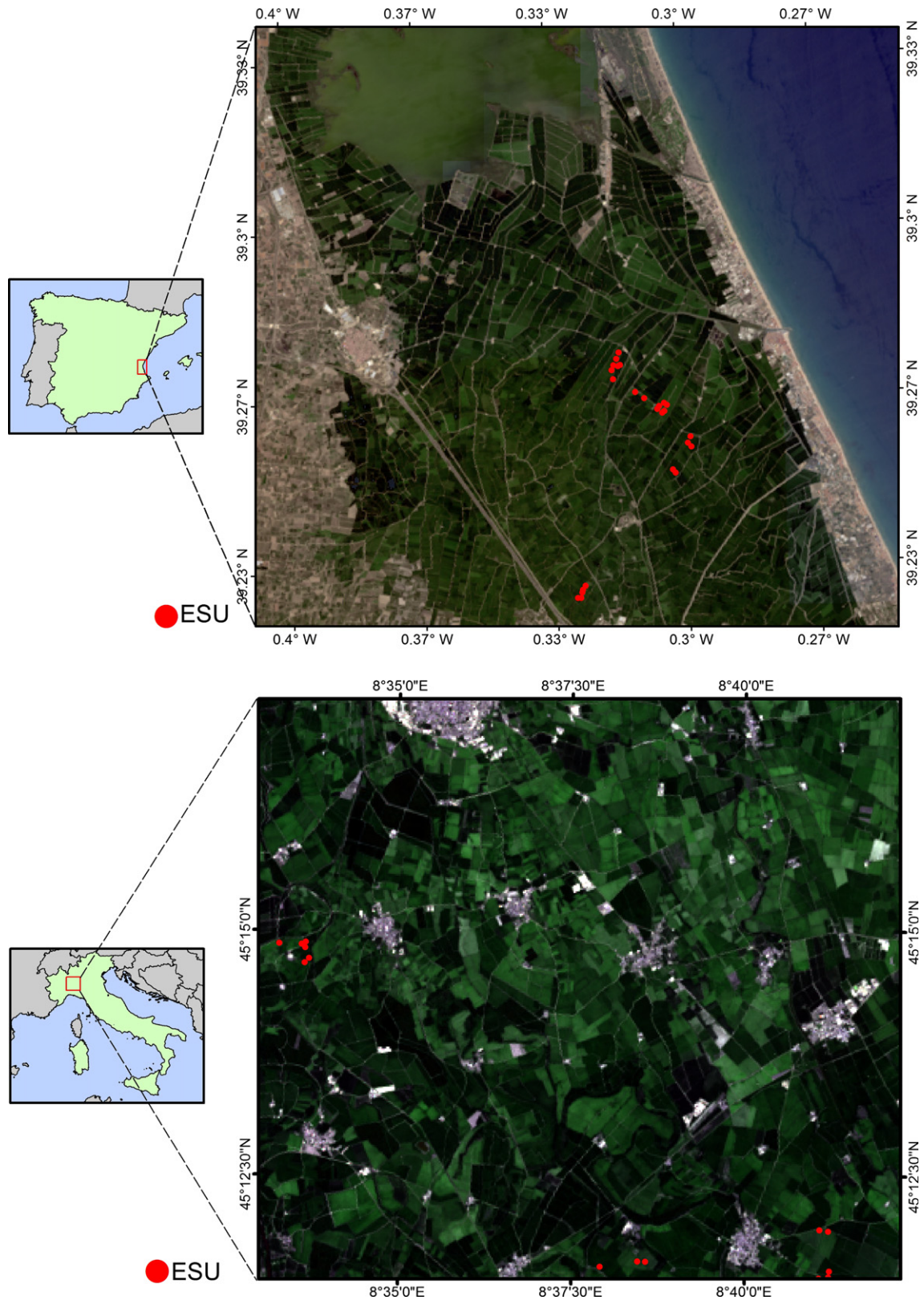


Fig. 1. Study areas: Landsat 8 OLI surface reflectance RGB composite of the Spanish study area on 10 August 2015 (top) and Landsat 8 OLI surface reflectance RGB composite of the Italian study area on 7 August 2015 (bottom). Locations of the ground measurements are displayed as red dots.

the unfolding of 2–4 leaves, and then crops are flooded. In 2014, it was estimated that more than 30% of the rice cultivated area adopted this technique.

The Spanish study area is located in selected farms within the rice district of Valencia, East of Spain (see Fig. 1 (top)) belonging to the *Albufera Natural Park*, which is included as a special protection area

in the Natura 2000 network by the European Commission allowing only rice crop practices (European Commission, 2011). The area has a typical Mediterranean climate, with an average annual temperature and humidity of 17 °C and 65%, respectively. The mean annual precipitation is approximately 430 mm, mainly recorded in autumn. The Spanish site is a homogeneous rice planting area of approximately 10 km × 20 km extension. The Italian study area is the Lomellina rice district, which is located in the south-western Lombardy region, between the Ticino, Sesia and Po rivers where rice is the dominant crop (>90%). It is the rural part of the Pavia province, which includes 58 municipalities, and is particularly renowned for its rice cultivation (see Fig. 1 (bottom)).

2.2. Ground measurements

In the framework of the 2015 ERMES activities, LAI ground measurements were conducted over the previous study areas. Based on land cover distribution in the areas and information provided by farmers at the very beginning of the rice season, a reliable sampling was achieved selecting ESUs (elementary sampling units) with different rice varieties and sowing dates in order to cover as much as possible the variability of the study areas. In Spain, the total number of ESUs was 40. For some of the ESUs, field acquisitions were made in all field campaigns, whereas in other ESUs the temporal frequency was approximately one every two field campaigns. In the Italian study area, 10 ESUs were fixed and considered during all the field campaigns, increasing the number of ESUs up to 19 for some dates (see Table 1). The same sampling scheme was used over each ESU, following the guidelines and recommendations of the Validation of Land European Remote sensing Instruments (VALERI) protocol (<http://w3.avignon.inra.fr/valeri/>). For the case of row crops, this protocol suggests to take measurements along small transects between rows incorporating some random acquisitions to prevent possible biases in the characterization of the row effect. We adopted the same schema also for fields with broadcast seeding. The size of the ESUs was approximately 20 m × 20 m, and the locations were far from the field borders. In order to characterize the spatial variability within each ESU, a range of 18 to 24 measurements was taken. This number of replicates allows to obtain a statistically significant mean LAI estimate per ESU. The center of the ESU was geo-located using a GPS for later matching and associate the mean LAI estimate with the corresponding satellite spectra.

LAI estimates were acquired in the two countries with smartphones using an app called PocketLAI (Confalonieri et al., 2013). PocketLAI computes indirect LAI measurements through the segmentation of images acquired at 57.5° below the canopy and showed good performance in canopies with different structures (Francone et al., 2014). PocketLAI estimates can reproduce destructive LAI measurements with acceptable results in terms of both reliability and accuracy (Confalonieri et al., 2013). A schematic of a single LAI measurement using PocketLAI and the theoretical background are shown in the supplementary material (see Appendix A). Since the

gap fraction–LAI relationship is not linear, it is not equivalent to first average the gap fraction and then estimate the LAI than the contrary (Weiss et al., 2004). However, specifically over rice crops, we have recently shown that LAI measurements taken with PocketLAI align well with other traditional acquisition techniques, such as plant canopy analyzers and digital cameras for hemispherical photography (Campos-Taberner et al., 2016). Together with PocketLAI data, ancillary information such as rice phenology and flood condition was taken on each ESU to better evaluate the computed LAI maps.

2.3. Landsat imagery

The United States Geological Survey (USGS) facilitates free access to Landsat archive data (Woodcock et al., 2008). In this work, Landsat 8 OLI data (30 m pixel resolution) were downloaded through the USGS Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) (<http://espa.cr.usgs.gov/>). The provisional Landsat 8 Surface Reflectance product was used (LaSRC) (Vermote et al., 2016). In order to focus on the rice areas only, the Landsat 8 OLI images were cropped to 1500 × 800 and 480 × 431 pixel size in the case of Spain and Italy, respectively. Images were available every 16 days in Italy. The Spanish study area lies in two Landsat paths within the same row (198/33 and 199/33), increasing the temporal resolution of the images to seven and nine days. Landsat 8 OLI surface reflectance spectral bands were filtered to relate only the blue (B), green (G), red (R), near infrared (NIR), and the two short wave infrared (SWIR1, SWIR2) channels with the ground LAI measurements in the retrieval process.

Cloud contamination is a common problem which limits the utility of passive optical multispectral images. To deal with this problem, Landsat 7 ETM+ data were used for increasing the temporal resolution of cloud free images. Nevertheless, Landsat 7 data can be affected by data gaps causing lost information of about 22% of the pixels in ETM+ images (referred to as SLC-off images) (Arvidson et al., 2006; Ju and Roy, 2008). In this study, we used the Neighborhood Similar Pixel Interpolator (NSPI) algorithm (Chen et al., 2011), which makes use of the neighboring pixels with similar spectral characteristics to predict the value of missing pixels using spatial and temporal information of gap free images (Campos-Taberner et al., 2015a). To compute the spectral similarity between the target pixel and neighboring pixels, only cloud-free observations are used, thus exploiting information about the cloud mask and the temporal trajectory. Fig. 2 shows the available cloud-free Landsat imagery acquired during the 2015 rice season over the study areas.

It is worth noting that Landsat 7 ETM+ and Landsat 8 OLI have different spectral response functions. This leads to slight differences in surface reflectance for each sensor. Nevertheless, differences are low when surface reflectance is simulated (see Fig. S2 of the supplementary material in Appendix A). On the other hand, another source of discrepancies in LAI retrievals is the atmospheric correction used for the generation of Landsat 7 ETM+ and Landsat 8

Table 1
Dates and number of ESUs sampled on the study areas during 2015.

Spain			Italy		
Date	DOY	ESUs	Date	DOY	ESUs
02/06	153	24	26/05	146	10
09/06	160	24	04/06	155	16
17/06	168	22	15/06	166	18
24/06	175	20	24/06	175	19
02/07	183	22	06/07	187	19
09/07	190	22	21/07	202	19
17/07	198	19	07/08	219	19
23/07	204	20			
03/08	215	27			

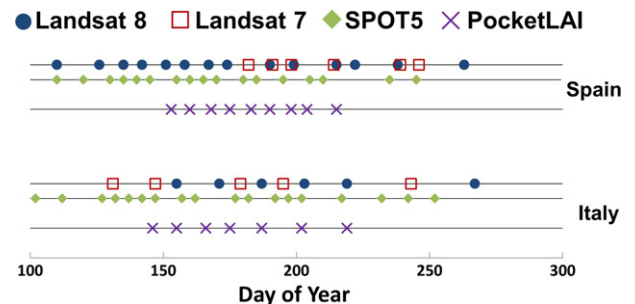


Fig. 2. Remote sensing data acquired during the 2015 rice season over the Spanish and Italian study areas and the *in situ* PocketLAI acquisitions.

OLI surface reflectance data. Although the Landsat 7 ETM+ atmospheric correction (LEDAPS) is less accurate than the Landsat 8 OLI atmospheric correction (Vermote et al., 2016) differences in surface reflectance and LAI estimates of two consecutive images are small (Roy et al., 2016).

2.4. SPOT5 imagery

With the aim of simulating multitemporal data from Sentinel-2 mission, the SPOT5 Take5 experiment provided 10 m pixel resolution data over selected sites every 5 days under constant angles from end-April to early-September 2015, thus covering the majority of the vegetation phase for summer crops in Europe. In this framework, ESA launched a call for proposals of sites selection, and the ERMES study areas were included in the 2015 SPOT5 Take5 acquisition plan. A total of 17 and 18 SPOT5 cloud-free images were acquired over the Spanish and the Italian study areas (see Fig. 2). The imagery were downloaded through the Theia land data center, which provides a top of canopy surface reflectance product (green, red, near infrared and short wave infrared channels) obtained applying an ortho-rectified process (Baillarin et al., 2008) and then computing an atmospheric correction with MACCS Software (Hagolle et al., 2008). Images were spatially cropped to 1440×1293 (Italy) and 4500×2400 (Spain) pixel size in order to match with the extent of the corresponding Landsat imagery.

Note that pixels corresponding to urban areas, water bodies and areas of 'no interest' such as non-rice crops were masked out during the retrieval process in order to avoid meaningless LAI estimates over those surfaces. The masking process was realized using official parcel boundaries and farmers crop declaration for CAP (European Common

Agricultural Policy) obtained from the Valencian government (<http://terrasit.gva.es/>) and the Lombardy regional authorities (<https://www.dati.lombardia.it/browse?category=Agricoltura>).

3. Retrieval methodology

A general outline of the proposed retrieval chain is shown in Fig. 3. The basic ingredients include the PROSAIL model and statistical regression algorithms for model inversion. Once the regression models are developed, we apply them to retrieve high-resolution LAI maps on the study areas from corresponding Landsat and SPOT5 surface reflectance data. In the following subsections we describe these components in detail.

3.1. PROSAIL model

The PROSAIL radiative transfer model (Jacquemoud et al., 2009) was used to build the database for training the retrieval model. It assumes the canopy as a turbid medium for which leaves are randomly distributed. PROSAIL simulates the top of canopy bi-directional reflectance in the range of 400 nm to 2500 nm as a function of input variables related to the structure of the canopy, the leaf optical properties, the background soil reflectance and the sun-view geometry. Leaf optical properties are expressed taking into account the mesophyll structural parameter (N), leaf chlorophyll (C_{ab}), dry matter (C_m), and water (C_w) contents. C_w was tied to the dry matter content ($C_w = C_m \times C_{wREL} / (1 - C_{wREL})$) assuming that green leaves have a relative water content (C_{wREL}) varying within a relatively small range (Baret et al., 2007) (see Table 2). On the other hand, canopy structure is characterized by the average leaf angle

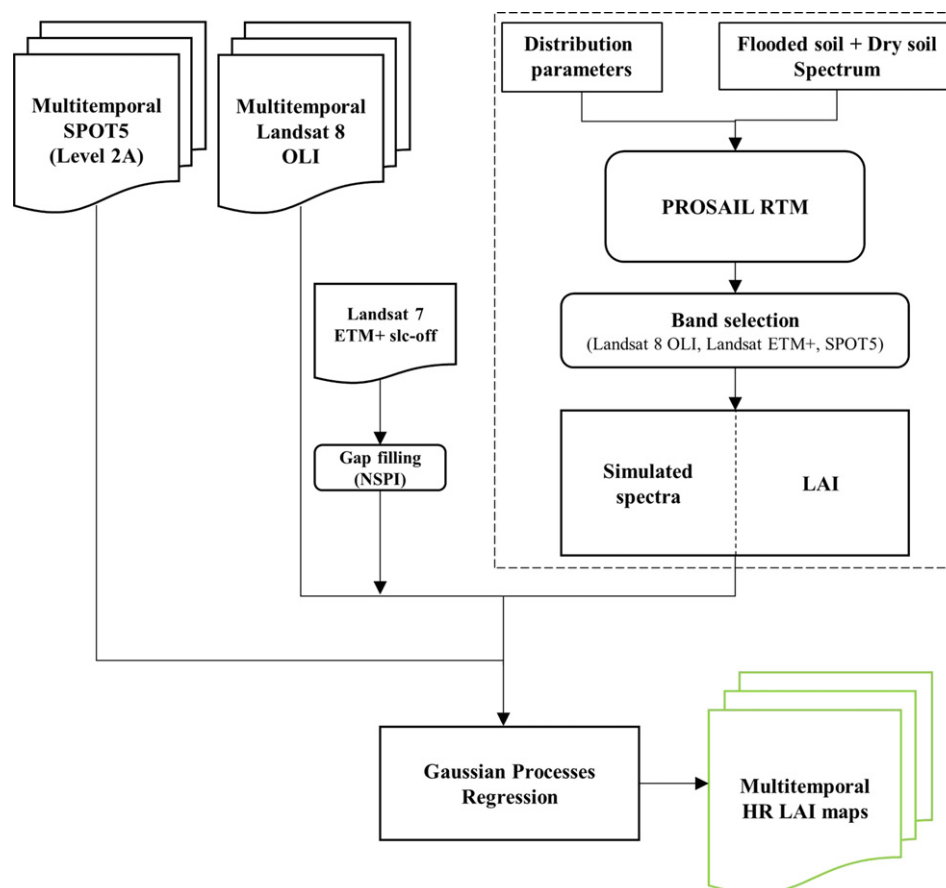


Fig. 3. Operational chain followed in the multitemporal LAI retrieval.

Table 2
Distribution of the canopy, leaf and soil parameters within the PROSAIL RTM.

Parameter	Min	Max	Mode	Std	Type	
Canopy	LAI (m ² /m ²)	0	10	3.5	4.5	Gaussian
	ALA (°)	30	80	60	20	Gaussian
	Hotspot	0.1	0.5	0.2	0.2	Gaussian
	vCover	0.5	1	1	0.2	Truncated Gaussian
Leaf	N	1.2	2.2	1.5	0.3	Gaussian
	C _{ab} (μg · cm ⁻²)	20	90	45	30	Gaussian
	C _{dm} (g · cm ⁻²)	0.003	0.011	0.005	0.005	Gaussian
	C _{wREL}	0.6	0.8	–	–	Uniform
Soil	β _s	0.3	1.2	0.9	0.25	Gaussian

inclination (ALA), the LAI and the hot-spot parameter (Hotspot). A multiplicative brightness parameter (β_s) was introduced and applied to spectral flooded and dry soil signatures to represent different background reflectance types (Baret et al., 2007; Claverie et al., 2013). The system geometry was described by the solar zenith angle (θ_s), view zenith angle (θ_v), and the relative azimuth angle between both angles (Δθ).

Sub-pixel non-vegetated areas were found in the borders of rice fields but patches of bare/flooded soil, small water stripes and channels were found in paddies as well. These conditions can be due to water drainage, very poor soil fertility, flattening mechanical process and other causes related to agro-practices leading to high yield reduction. The interested reader is referred to Fig. S4 of the supplementary material attached in Appendix A for further details. Therefore, in order to account for these mixed conditions, a pixel can be represented by a linear mixture of vegetation (vCover) and bare/flooded soil (1-vCover) spectra. A linear spectral mixing model was assumed for the sake of simplicity. Taking this heterogeneity into account, the pixel reflectance can be expressed as $R = R_{veg} \times vCover + R_{soil} \times (1 - vCover)$, where R_{veg} and R_{soil} account for pure vegetation reflectance and background, respectively. This simple approach was introduced by Baret et al. (2007) to account for clumping at the landscape level. Note that when validating retrieved LAI of a mixed pixel, it is compared with LAI × vCover (see distributions of LAI and vCover in Table 2). In this study, vCover was assumed to be independent of LAI, following a truncated Gaussian distribution. In addition, a 5% of pure background spectra (vCover=0) were added to represent situations at the beginning of the season (no vegetation) and large patches of bare/flooded soil present during the rice season.

The leaf and canopy variables as well as the soil brightness parameter were randomly generated following specific distributions (see Table 2). The parameterizations were similar to other studies using high-resolution sensors (Bsaibes et al., 2009; Duveiller et al., 2011). Nevertheless, in this study, a site-specific parameterization of the PROSAIL model based on the available 2014 ERMES field measurements was selected in order to constrain the behavior of the model to Mediterranean rice areas reducing the equifinality of the ill-posed PROSAIL inversion process (Combal et al., 2003). During the ERMES 2014 rice season, leaf chlorophyll content was measured with a SPAD-502 chlorophyll meter (Campos-Taberner et al., 2016) thus allowing to constrain the C_{ab} range to Mediterranean rice values which typically vary from 35 to 70 μg · cm⁻² depending on the rice variety. This range was slightly extended in order to represent rice with high leaf chlorophyll concentration due to fertilization practices, as well as low leaf chlorophyll content caused either by possible diseases, blasts or nitrogen deficits. ALA distribution was selected for accounting specific leaf inclination during rice phenological stages (Zhang et al., 2013). In order to better constrain the retrieval to rice crops, a spectral library of underlying soil background was generated by considering signatures of homogeneous flooded and dry areas identified within rice fields in the study areas with Landsat and SPOT5 imagery, which was spectrally

matched with typical rice background spectra collected by Boschetti et al. (2014). Each background signature was selected randomly from the spectral library and multiplied by the soil brightness factor (β_s), which was assumed to follow a Gaussian distribution (see Table 2). β_s is assumed to be sensitive to soil moisture, roughness and geometrical configuration (Baret et al., 2007). A random Latin hypercube sampling design allowed to populate more evenly the canopy realization space (Mckay et al., 2000). For each sensor, a single dataset of PROSAIL simulations was performed, which included all geometrical configurations. The distributions for the system geometry were randomly generated based on information in imagery metadata.

3.2. Regression methods

In this paper we propose inverting PROSAIL using machine learning statistical algorithms. For this purpose, we used three representative nonlinear regression methods: the familiar artificial neural network, and two related kernel-based regression algorithms: the kernel ridge regression and the Gaussian process regression. This section reviews the three regression methods used and discusses about the implementation issues for the interested reader.

3.2.1. Neural networks (NN)

Artificial neural networks are based on the combination of simple nonlinear processing units, called neurons, into a fully connected hierarchical architecture. The network can model complex, nonlinear input-output relations, and has been the preferred regression and function approximation tool for decades for retrieving biophysical parameters. Actually, the vast majority of hybrid inversion methods consider the use of neural networks (Baret et al., 1995; Baret and Fourty, 1997; Smith, 1993) for retrieval of canopy parameters. Essentially, each neuron in a network performs a linear regression followed by a non-linear activation (sigmoid-like) function. Neurons of different layers are interconnected by weights that are adjusted during the training phase (Haykin, 1999). In order to train the network (i.e. fit the weights), one has to select a cost function (in our case the least squares loss) and an algorithm to do this (in our case the standard back-propagation algorithm). Several hyperparameters are involved as well and impact the solution: essentially, the number of hidden layers and neurons/nodes and the learning rate of the back-propagation algorithm.

3.2.2. Kernel methods

Kernel methods (Shawe-Taylor and Cristianini, 2004) owe their name to the use of kernel functions, which measure similarities between input data examples. We used two related and powerful kernel methods for regression: kernel ridge regression (Shawe-Taylor and Cristianini, 2004), and Gaussian process regression (Rasmussen and Williams, 2006). The KRR is considered as the nonlinear (kernel) version of the canonical least squares linear regression, while the GPR is a probabilistic approximation to nonparametric kernel-based regression, where both a predictive mean (point-wise estimates of LAI) and predictive variance (error bars for the LAI predictions) can be

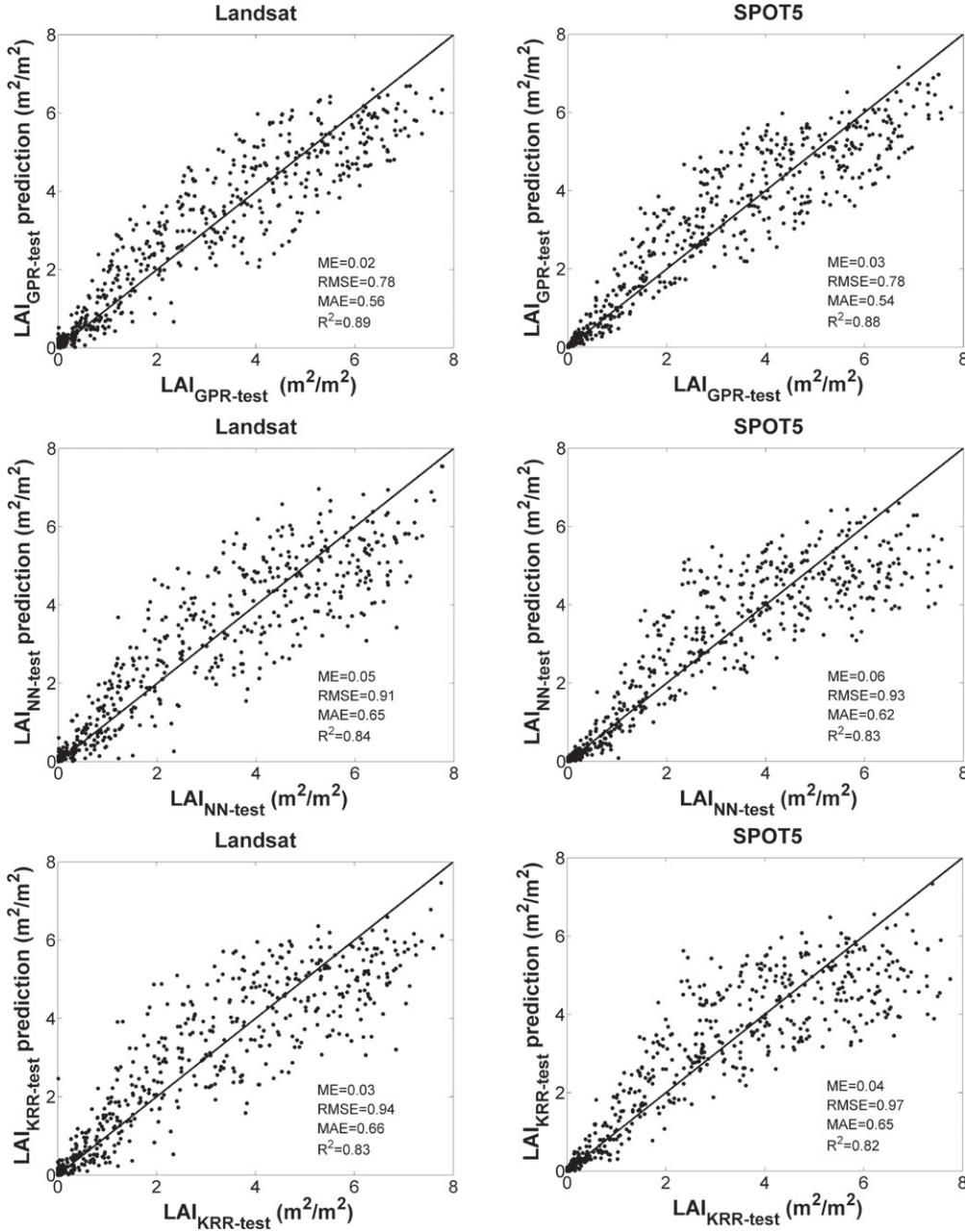


Fig. 4. GPR (top), NN (middle) and KRR (bottom) theoretical performances using Landsat 8 OLI (left) and SPOT5 (right) simulated data. Plots are obtained by applying the trained models on test data. Each method was trained following either a standard fold cross-validation within the training set (KRR and NN) or by maximizing the marginal log-likelihood of the observations, which is analytical (for the GPR).

derived. Notationally, both methods offer the same explicit form of the predictive model, which establishes a relation between the input (e.g., spectral data) $\mathbf{x} = [x^1, \dots, x^B] \in \mathbb{R}^B$ and the output variable (i.e., LAI) $y \in \mathbb{R}$ of the form:

$$\hat{y} = f(\mathbf{x}) = \sum_{i=1}^N \alpha_i k_{\theta}(\mathbf{x}_i, \mathbf{x}) + \alpha_0, \quad (1)$$

where $\{\mathbf{x}_i\}_{i=1}^N$ are the spectra used in the training phase, α_i is the weight assigned to each one of them, α_0 is the bias in the regression function, and k_{θ} is a kernel or covariance function (parametrized by a set of hyperparameters θ) that evaluates the similarity between the test spectrum and all N training spectra.

In order to generate a kernel regression model, one needs to specify a covariance/kernel function k_{θ} , to infer its hyperparameters θ and model weights α . For the KRR prediction model, we used the squared exponential (SE) kernel: $k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / (2\sigma^2))$, which is simply parameterized by $\theta = \sigma$ (also known as the kernel length-scale) that needs to be tuned typically by cross-validation. For the GPR prediction model, we used the so-called automatic relevance determination (ARD) kernel, as an alternative generalization of the isotropic SE prior:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \nu \exp\left(-\sum_{b=1}^B \frac{(x_i^{(b)} - x_j^{(b)})^2}{2\sigma_b^2}\right) + \sigma_n^2 \delta_{ij}, \quad (2)$$

where ν is a scaling factor, B is the number of bands, and σ_b is a dedicated parameter controlling the spread of the relations for each particular spectral band b . Model hyperparameters are collectively grouped in $\theta = [\nu, \sigma_n, \sigma_1, \dots, \sigma_B]$, and model weights α_i can be automatically optimized by maximizing the marginal likelihood in the training set (Rasmussen and Williams, 2006; Verrelst et al., 2012b). The obtained weights α_i after optimization give the relevance of each spectrum \mathbf{x}_i , while the inverse of σ_b represents the relevance of each band b . Hence, low values of σ_b indicate a higher informative content of this certain band b to the training function k . In this paper we study both the KRR and the GPR models paying attention to the band relevance conveyed by the inferred σ_b values, and the prediction uncertainty provided by the GPR model.

3.2.3. Model development and source code

Inference of the hyper-parameters for all methods and the weights for doing predictions was done as follows. We first generated 2000 data pairs (reflectances-LAI values) with PROSAIL, and used 70% for model selection (training set), and then evaluated and showed the results in the remaining 30% test set, which was never used or seen in model fitting. Even though 1400 samples could seem to be insufficient for training purposes, increasing this number of samples did not have a significant impact in the accuracy of the retrievals for all methods in the test set, indicating that they did not incur in any overfitting issue and highlighting the good representativity of the simulated data. For neural networks, the hyperparameters to be chosen were the number of hidden layers and neurons (for simplicity we evaluated one hidden layer and 2–30 hidden neurons) and the learning rate between 0.001 and 0.1 in log-scale. The bias input was set to -1 (not adjusted during training) and different initializations of the weights were tested. For the KRR model, we varied the length-scale σ (between 0.1 and 10 times the average distance between all training points in log-scale) and the regularization parameter σ_n (between 10^{-5} and 10^{-2} in log-scale). For the case of the GPR, we inferred the hyperparameters $\theta = [\nu, \sigma_n, \sigma_1, \dots, \sigma_B]$ and model weights using an optimization of the evidence. All experiments were conducted with our SimpleR MATLAB toolbox, freely available at <http://isp.uv.es/>. The toolbox is intended for practitioners with little expertise in machine learning, and that may want to assess advanced methods in their problems easily. The toolbox compares numerically and statistically the algorithms by simply entering the input-output (e.g. reflectances and LAI values) data matrices.

4. Results and validation

This section is devoted to show the theoretical performance of the regression methods and experimental evidence of the performance of the proposed processing scheme. We pay attention to the

derived HR LAI maps, and illustrate the usefulness for multitemporal rice crop monitoring through a temporal evolution analysis of LAI estimates compared to *in situ* measurements in the different test sites.

4.1. Accuracy assessment over the simulated dataset

It was necessary to build a dedicated model for each of the three remote sensing datasets (Landsat 7 ETM+, Landsat 8 OLI and SPOT5) depending on their spectral bands and angular configurations. We first evaluate the GPR performance compared to NN and KRR (see Fig. 4) over the test dataset (600 samples) corresponding to Landsat 7 ETM+, Landsat 8 OLI and SPOT5 datasets (Landsat 7 ETM+ results were similar to Landsat 8 OLI and are not shown for brevity). Hereafter, we will refer simply to ‘Landsat’ estimates, irrespectively of being for Landsat 7 ETM+ or Landsat 8 OLI. Likewise, we refer to GPR-Landsat and GPR-SPOT5 to the GPR models built using Landsat and SPOT5 data, respectively. Results revealed good accuracy and low bias in both Landsat and SPOT5 simulated reflectances. GPR outperformed NN and KRR in all statistical quality measures (see Fig. 4). GPR-Landsat and GPR-SPOT5 performances were robust and very similar, revealing biases of 0.02 and 0.03 m^2/m^2 respectively, while a root mean squared error (RMSE) of 0.78 m^2/m^2 and high determination coefficients ($R^2 > 0.87$) were obtained in both cases. Fig. 5 exhibits the evolution of the average RMSE as a function of the number of predictions, and shows a consistent better performance (i.e. improved robustness to reduced-sized datasets) of GPR over NN and KRR. The curves are the result of averaging a number of realizations, and for each realization, we computed the RMSE with a fixed number of predictions chosen at random. In the limit of 100%, one obtains the results in Fig. 4. This particular plot tries to analyze models’s robustness to local consistency and reliability of the estimated error (Montavon et al., 2013). The shape of the obtained curves are similar for all methods (high variance when few predictions were evaluated, and rapid convergence to stable RMSE), yet better for GPR. These results revealed GPR as the most accurate and robust regression method in both Landsat and SPOT5 test sets.

As discussed earlier (cf. Section 3.2), GPR provides a σ_b parameter whose reciprocal represents the relevance of each band in the regression. Hence, it is possible to identify the most relevant bands used by the GPR model, which it is not affordable when using NN and KRR. Specifically, the most relevant bands in both cases were the green and near infrared bands (see Fig. 6), which showed the theoretical consistency between GPR model behaviors. The identification of the most valuable bands for LAI retrieval was expected since the green and near infrared bands are more related to the greenness of the vegetation (useful for assessing plant vigor) and leaf area index features,

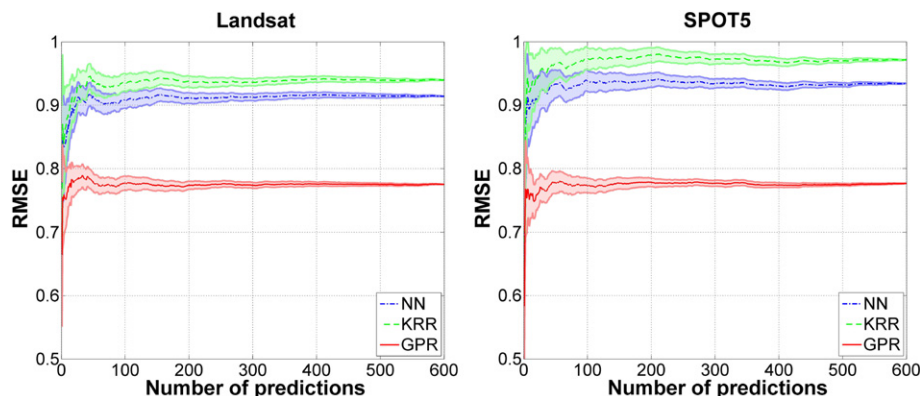


Fig. 5. Performance in the test set averaged over 100 random realizations of the Landsat and SPOT5 training-test data splitting as a function of the number of used predictions.

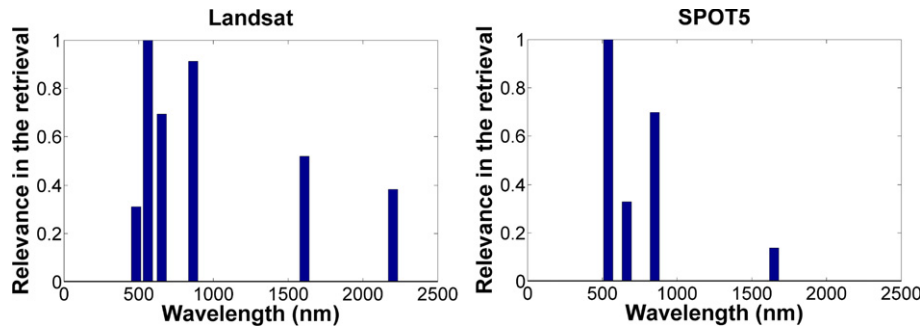


Fig. 6. Relevance band histograms in GPR-Landsat (left) and GPR-SPOT5 (right).

respectively. Similar band ranking was obtained by Verrelst et al. (2015b) although the results are not directly comparable because of the different spatial and spectral resolutions.

In this work, all six Landsat optical bands were used for the GPR-Landsat LAI retrievals although in many studies the blue band is not used because of atmospheric effects. Nevertheless, blue band was also considered for LAI retrieval in other studies (Atzberger and Richter, 2012; Borel, 2010; Verrelst et al., 2015b) since the blue spectrum holds information valuable for LAI and phenology (Huete et al., 2002). Actually, GPR is very robust to moderate-to-high dimensional spaces (i.e. few more input variables do not impact results negatively). A comparison of LAI estimates obtained with or without using the Landsat blue band revealed a slight improvement when including the blue band in the GPR retrieval (see Fig. S3 and Table S1 of the supplementary material in Appendix A).

4.2. Accuracy assessment over ground LAI

With the goal of assessing the accuracy of the retrievals, RMSE between the estimates and the *in situ* measurements was computed. RMSE values of 0.39, and 0.38 m^2/m^2 were found in Spain and Italy respectively, showing good accuracy between GPR-Landsat map values and the *in situ* LAI measurements (see Fig. 7). The GPR-SPOT5 retrievals revealed good accuracies as well, showing RMSE values of 0.51 and 0.47 m^2/m^2 in Spain and Italy respectively. Other different statistics, such as the mean error (ME), mean absolute error (MAE) and the coefficient of determination were also computed to evaluate the bias, accuracy, and the goodness-of-fit between GPR-Landsat and GPR-SPOT5 predictions and measurements. A remarkably good correspondence between satellite retrievals and *in situ* measurements was found in the Spanish site, with very low bias for both

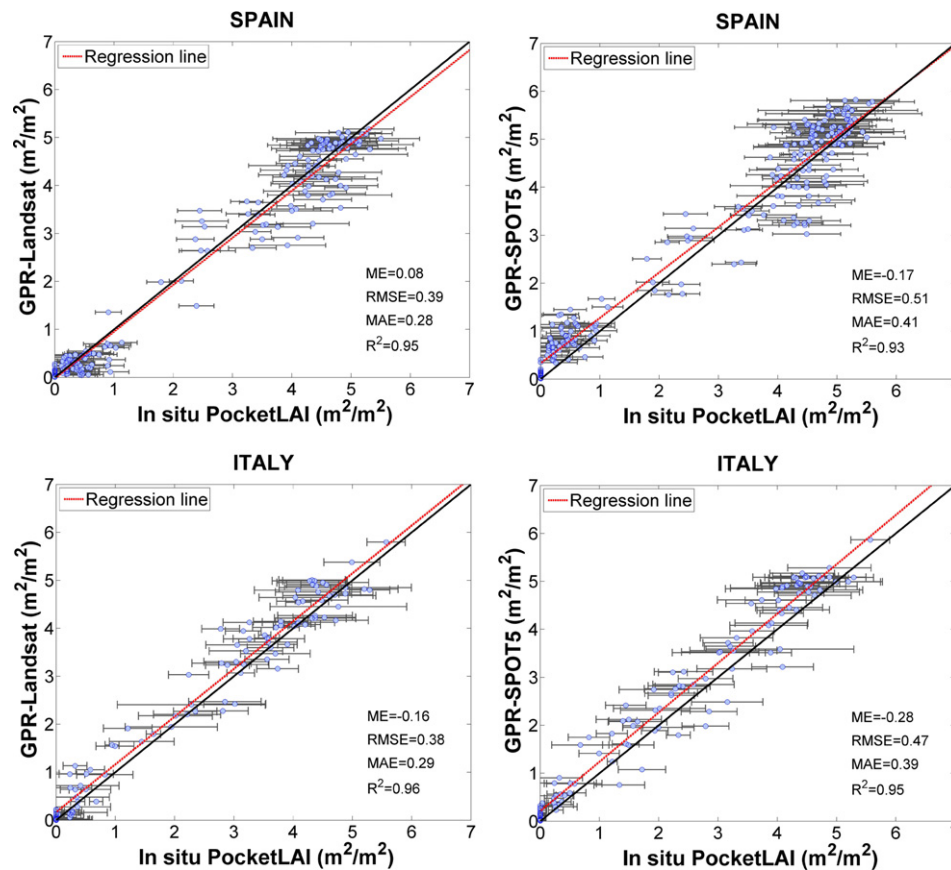


Fig. 7. Scatter plots of estimated LAI values using Landsat (left) and SPOT5 (right) data versus *in situ* LAI measurements acquired with PocketLAI over Spain (up) and Italy (bottom). Standard deviation of measurements is drawn as horizontal error bars as well as bisector line (black line). For the sake of visualization, GPR prediction uncertainty ($\approx \pm 1$) are not shown.

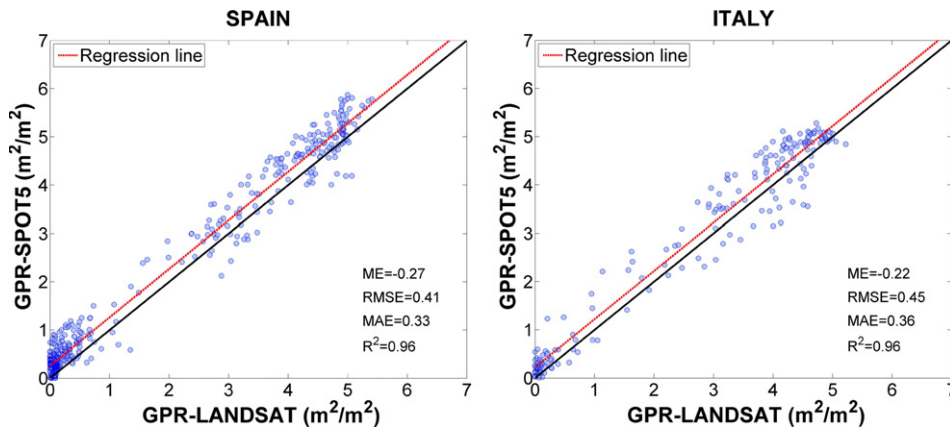


Fig. 8. Scatter plots of the estimated LAI values of the GPR-Landsat and GPR-SPOT5 retrievals in (left) Spain and (right) Italy.

GPR-Landsat and GPR-SPOT5 models. A good agreement and low biases were also observed for the Italian site. In all cases, very high correlations were found with $R^2 > 0.92$. Error bars in Fig. 7 refer to the standard deviation of the field measurements, and are thus related with the heterogeneity of the ESU. In addition, retrievals were also computed using KRR and NN showing slightly less accurate LAI estimates. The interested reader is referred to Fig. S5 of the supplementary material attached in Appendix A.

4.3. Comparison of GPR-Landsat and GPR-SPOT5

Besides the aforementioned theoretical performance of the GPR-Landsat and GPR-SPOT5, the experimental consistency between estimates was assessed with the inter-comparison of multitemporal LAI estimates. The closest acquisition dates between Landsat and SPOT5 were taken into account for the comparison. After the SPOT5 resampling to Landsat resolution, the comparison was achieved averaging the LAI value of the valid estimates computed over the 3×3 pixels if more than 5 out of the 9 pixels were valid (Morissette et al., 2006)

in order to reduce coregistration errors between images and inconsistencies associated to differences in the point spread functions. In general, the estimates provided by both GPRs were highly correlated. However, a slight overestimation of the SPOT5 based retrievals was observed in both study areas (see Fig. 8). In addition, spatial consistency between retrievals was performed for the closest temporal Landsat and SPOT5 images during the rice growing season. The closest useful available images were found in Italy in 2015 July 22nd for Landsat and in 2015 July 21st in the case of SPOT5. Difference ($\Delta = GPR_{SPOT5} - GPR_{Landsat}$) LAI map was computed after the GPR-SPOT5 LAI map resampling into Landsat resolution (30 m). Statistical differences between GPR-SPOT5 and GPR-Landsat from ANOVA (one-way analysis of variance) were computed revealing F -statistic and p -value of 2.860 and 0.001 respectively, highlighting that the two distributions are not statistically different. The obtained results show 77% of the pixels fall within $\pm 0.5 \text{ m}^2/\text{m}^2$ interval and only 1% of the pixels reveal differences higher than $\pm 2 \text{ m}^2/\text{m}^2$ (see Fig. 9 (right)). The prevailing light green color in Fig. 9 (left) demonstrate that the retrievals from the two sensors are coherent and most of the rice areas have

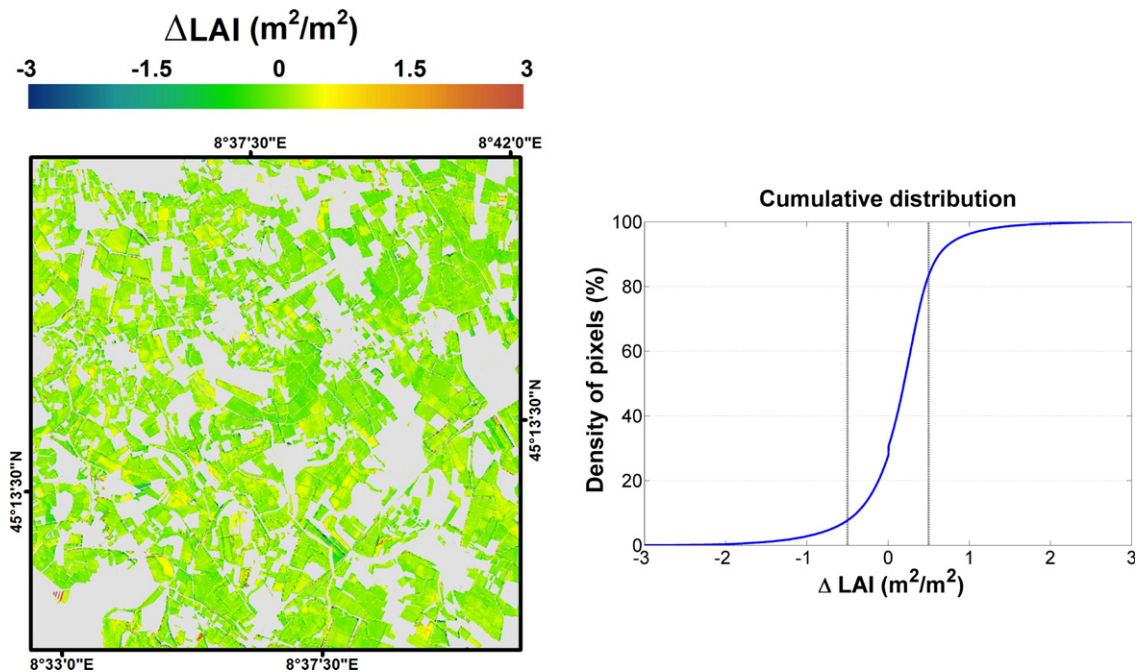


Fig. 9. (Left) LAI differences map between GPR-SPOT5 (2015 July 21st) and GPR-Landsat (2015 July 22nd), and (right) the corresponding cumulative distribution values. Grey mask covers non-rice areas.

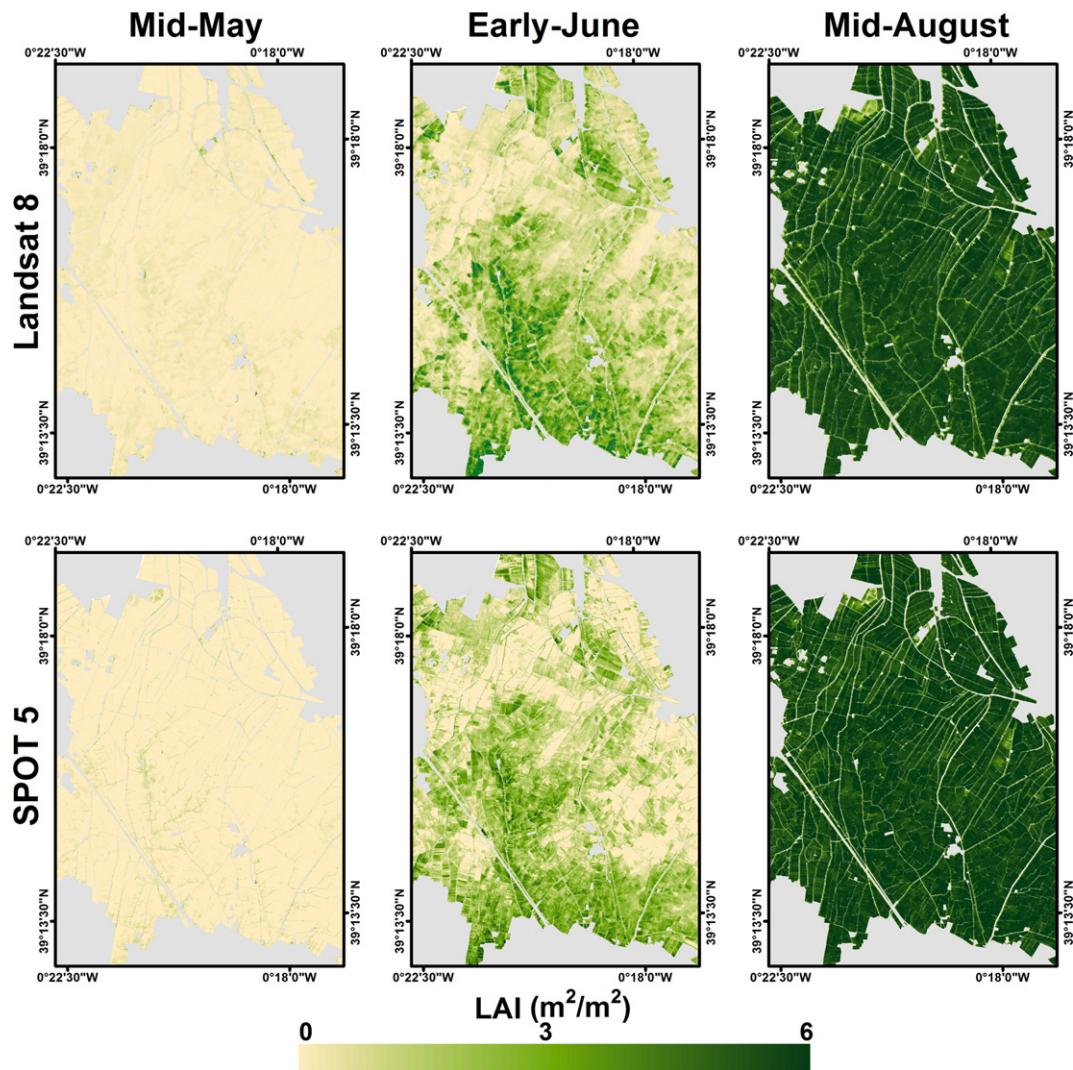


Fig. 10. Estimated LAI maps derived using Landsat (top) and SPOT5 (bottom) images over the Spanish study area in mid-May (left), early-June (middle) and mid-August (right). Grey mask covers non-rice areas.

LAI differences around zero. Bigger differences can be partially due to the different spatial resolution between Landsat and SPOT5.

4.4. Spatio-temporal analysis

Multitemporal LAI maps were derived for the study areas during the 2015 rice season. Fig. 10 shows the GPR-Landsat and GPR-SPOT5 LAI maps derived for the very beginning of season (mid-May), early growing season (early-June), and maximum leaf rice development (mid-August) in the Spanish site. The corresponding HR estimated LAI maps for the Italian study area are shown in Fig. 11. First inspection of the maps indicates the occurrence of very low LAI values corresponding to mid-May since the sowing dates were around May, 10–15th in Spain. In Italy, for the same period, some rice fields have higher LAI values (>2) because of the early sowing of some rice varieties. In the Spanish site, it can be seen the expected rice emergence in the early-June LAI maps (see Fig. 10 (middle)), while the early-July Italian maps (Fig. 11 (middle)) show already higher LAI estimates due to the advanced phenological growing state corresponding to the rice stem elongation phase. Eventually, the mid-August maps (Figs. 10 and 11 right panels) show the highest LAI estimates because rice plants reached the heading phase and the LAI seasonal peak. All 2015 derived high-resolution maps can be investigated by registering and

joining as local user into the ERMES web-based geo-portal (<http://ermes.dlsi.uji.es/prototype/geoportal/>) or can be found in the ERMES catalogs (<http://get-it.ermes-fp7space.eu/>).

The *in situ* LAI data points allowed us to compare the temporal evolution of field measurements over the study areas (see Fig. 12). In general, LAI estimates derived from the proposed algorithm using GPR-Landsat and GPR-SPOT5 data agree with regard to the seasonal rice phenological cycle in the two countries and followed the temporal dynamics of the ground measurements. The different LAI evolutions that can be observed in Fig. 12, show coherent temporal behaviors as a consequence of either different rice varieties or sowing dates that determine a shift in the development curve. The interested reader is referred to Fig. S6 of the supplementary material attached in Appendix A for further details. It is interesting to note that the difference between Landsat 7 ETM+ and Landsat 8 OLI did not induce large difference in LAI retrieval. For example, LAI estimates of Landsat-8 DOY=190 and Landsat-7 DOY=191 for the Spanish site showed differences lower than $0.25 \text{ m}^2/\text{m}^2$ (see blue profiles in Fig. 12).

On the other hand, both GPR-Landsat and GPR-SPOT5 models also provided a prediction uncertainty for the LAI estimates (see Fig. 12). Although these uncertainties cannot be used as a validation *per se*, they can be useful to draw conclusions about the quality of the

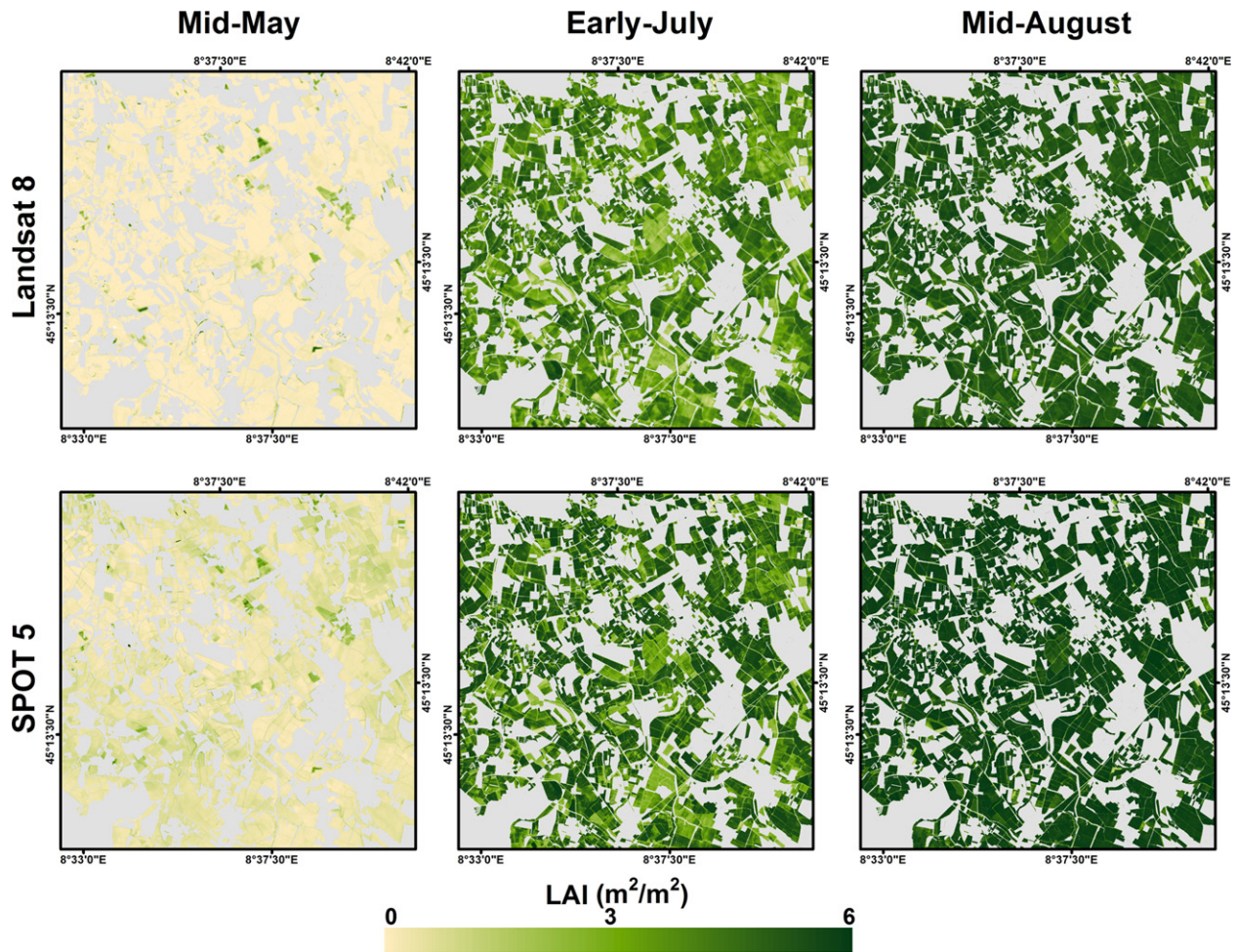


Fig. 11. Estimated LAI maps derived using Landsat (top) and SPOT5 (bottom) images over the Italian study area in mid-May (left), early-July (middle) and mid-August (right). Grey mask covers non-rice areas.

retrievals. Apparently, the prediction uncertainties look constant for the multitemporal LAI estimates, nevertheless a deeper look reveals a change in the behavior: at the beginning of the rice season (no vegetation) values are high (≈ 1.25) and when rice starts to emerge (about day of year 140–150) the uncertainty decreases significantly (≈ 0.95), while during the rice development period it remains virtually constant. This reflects the fact that the simulations include a large amount of cases with intermediate and high LAI values. It is worth mentioning that the GPR prediction uncertainty only depends on the training reflectance. Thus, if the test and training reflectance are deemed similar, the uncertainty for the test set decreases because we are dealing with similar input features. Therefore, prediction uncertainty must be interpreted as a qualitative variable associated to the estimates: the higher the uncertainty, the lower the confidence on the associated estimate. As a matter of fact, this interpretation comes from the PROSAIL training data set used in the GPRs since unrepresented spectra, such as non-vegetated bodies, lead to higher uncertainties for its associated retrievals. In Section 4.5, we analyze the impact of using different background spectra in PROSAIL on LAI prediction uncertainties.

The rate of LAI development depends on the occurrence of specific phenological phases. Green-up corresponds to the end of the tillering phase, the rapid LAI increase occurs during stem elongation phase and the plateau of maximum LAI is reached in correspondence of heading for all the flowering period. From the analysis of LAI profiles for different ESUs, especially in Italy, it is possible to appreciate

a variability in the occurrence of the mentioned phenological stages due to different sowing dates and cultivated varieties. A deeper inter-comparison of LAI time series derived from GPR-SPOT5 in the Italian case study was conducted over selected monitored fields to better exhibit these behaviors. Fig. 13 reports the average time series of LAI for each monitored field in 2015. A buffer of 5 m was considered in order to get rid of the field border in the average computation. In the figure, the different monitored parcels are indicated in numerical order and different varieties in color (see legend), dotted vertical lines represent the date of sowing and the sowing practices as provided by the farmers (i.e. direct sowing in dry soil or seed broadcasting in flood condition in red and blue color respectively). It is possible to appreciate how rice is usually cultivated in monoculture and sowed in dry bare soil or in water (low/zero LAI value at the beginning of the series). It is noteworthy that anomalous high early LAI values in some fields (fields # 68 or # 384) are indicative of a cover crop preceding the rice season (observation confirmed by farmers communication). An early crop establishment period, about DOY 92 (fields # 50 and # 55), determined crop presence (e.g. LAI value) around DOY 130 to 150 when other fields (fields # 61, # 282 and # 384) were just sown. As a consequence of different sowing dates and varieties, the periods of maximum LAI plateau and senescence occur in different moment. Those fields with an early sowing (DOY 90 to 111) reached the LAI peak around DOY 180 whereas the variety sowed around DOY 140 (Sirio CL, Selenio) reached the LAI peak after DOY 200. The interested reader is referred to Fig. S7 in the

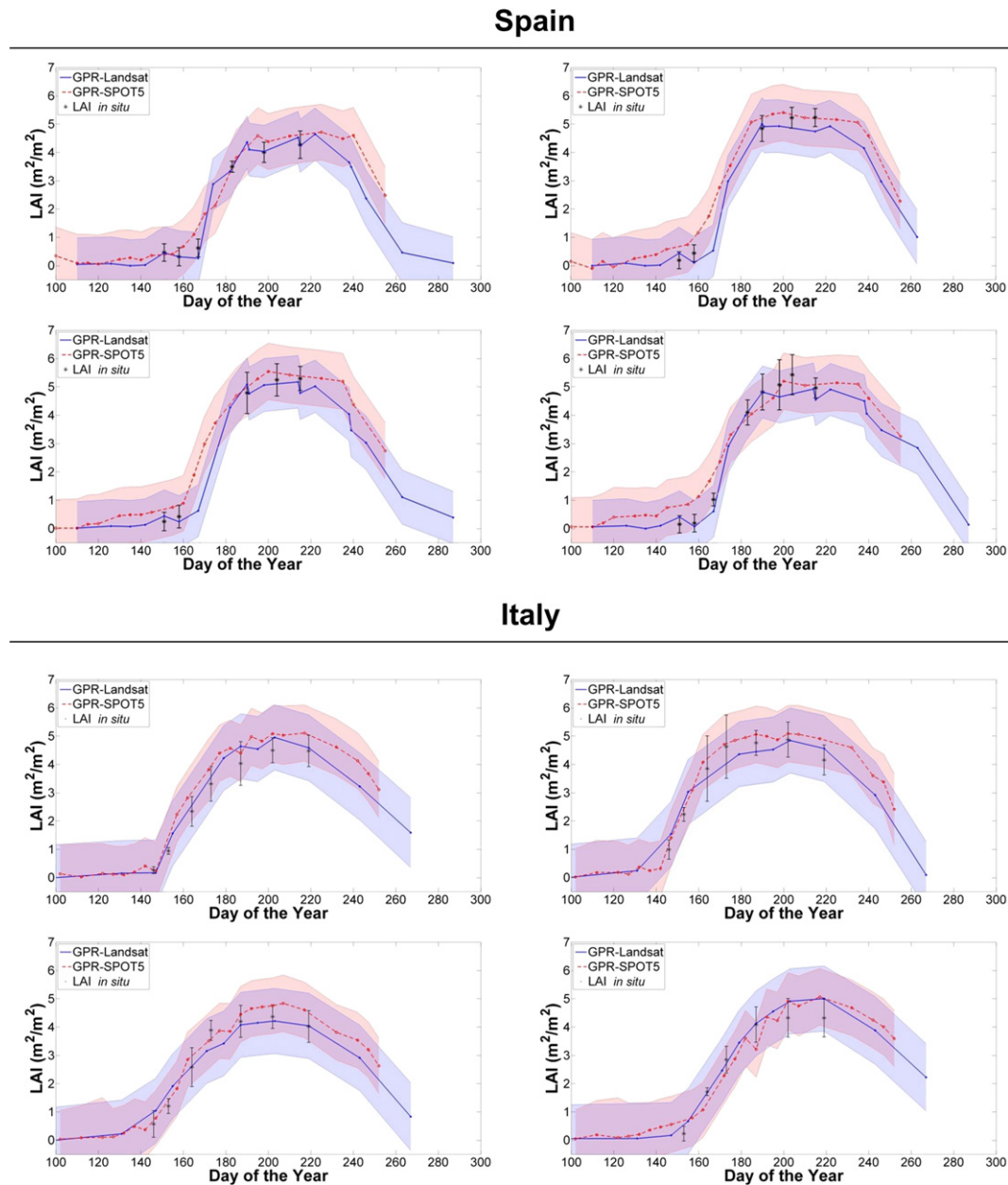


Fig. 12. Temporal dynamics of the Landsat and SPOT5 LAI estimates and the field LAI measurements within four representative ESUs in (top) the Spanish and (bottom) the Italian study areas. The prediction uncertainty provided by the GPRs is shown in shaded blue (GPR-Landsat) and orange (GPR-SPOT5) around the mean prediction; in violet the overlap. The standard deviations of the *in situ* measurements are displayed as error bars.

supplementary material attached in Appendix A for further analysis on LAI temporal behavior according to sowing date and cultivated variety.

The analysis of profiles reveals that LAI estimates are in agreement with the qualitative behavior of rice growth in the different conditions. LAI trends do not show artifacts due to changes in spectral background conditions related to water management. When soil changes from dry and flood condition (or vice versa) in the early growing stages (e.g. for dry sowing field # 106 and # 118; for flood sowing field # 62 and # 102) anomalies in LAI retrievals can be appreciated. The small fluctuations within each temporal profile are mainly due to the residual of atmospheric contamination; for example the SPOT5 image acquired at DOY 142 shows a haze pattern from east to west that can justify the small peak in several plots both in flood (# 102) and dry conditions (# 120). The LAI values of

these peaks/drops are anyway within the expected uncertainty of the retrieval.

These results highlight the usefulness of the retrieved LAI time series at high resolution to perform field level crop monitoring and site specific assimilation in crop models.

4.5. Assessing the influence of rice background in LAI retrievals

In this study, we used a representative set of background spectra taking into account all possible expected conditions of the field. The influence of background characteristics in LAI retrievals was assessed introducing different types of background spectra in the training database. Firstly, we trained and inverted the PROSAIL model using typical spectra of dry bare soils which were present at the very beginning of the rice season. Secondly, we used a set of flooded soil

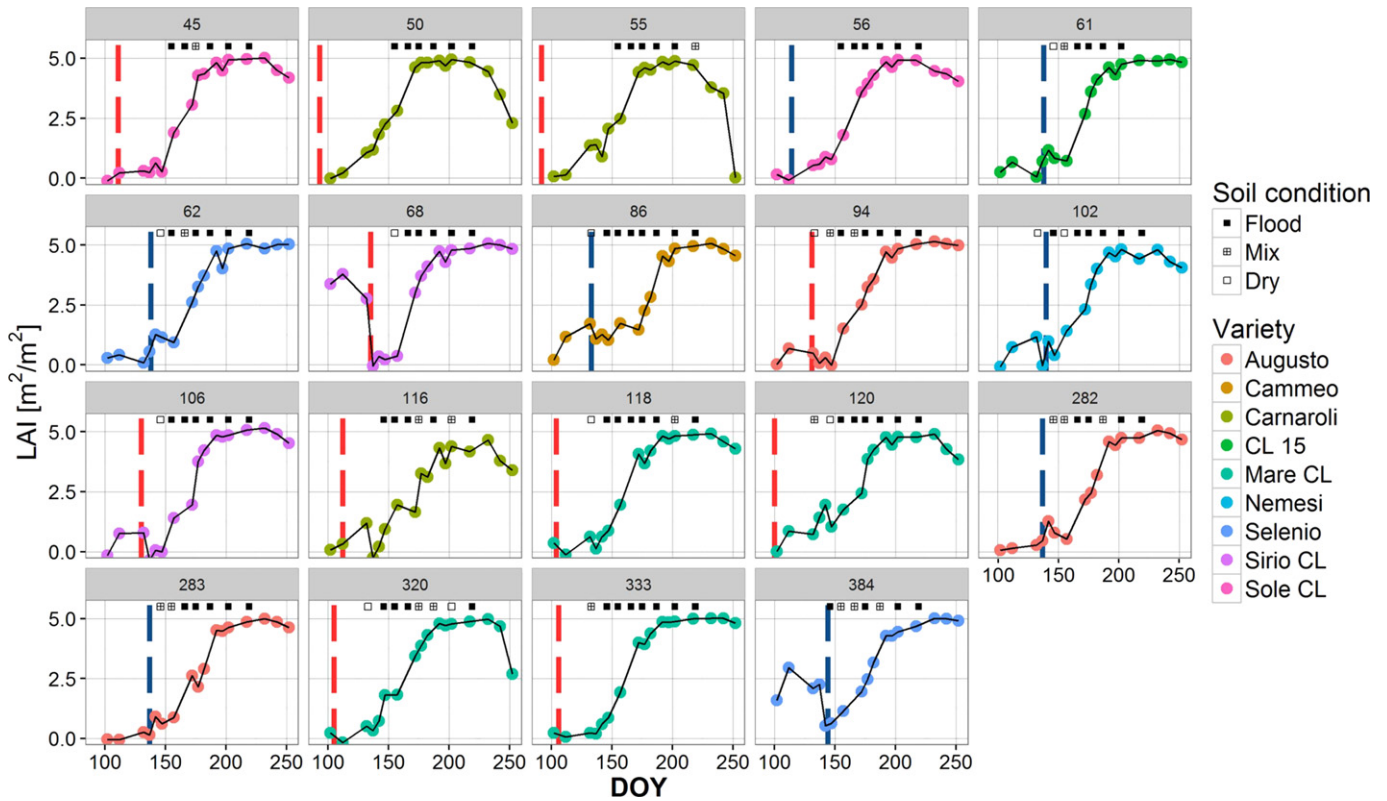


Fig. 13. Average behavior of estimated LAI over the monitored fields. Dotted vertical bars indicate the sowing date in dry soil (red) or in food condition (blue). Black squares indicates the soil condition recorded during the periodical field campaign.

signatures representing typical background conditions during the rice growing (from sowing to harvest). Eventually, we used both dry and flooded spectra. Note that during these experiments the distributions of PROSAIL parameters for the crop were kept constant, and only the set of background spectra was changed (i.e. dry, flooded and both dry + flooded conditions). Table 3 exhibits the statistics when estimates from Landsat over Spain are compared with *in situ* LAI data. Similar results were obtained in the case of Italy and SPOT5 data (results not shown for brevity). The best results were obtained when a background spectral library composed by flooded and dry soils was used. Conversely, when considering only flooded or dry backgrounds high errors were found during the initial development stages of the plants (during tillering and before panicle formation), as the reflectance of the background constitutes a significant component of the overall spectral signature recorded by the sensor. Thus, a correct characterization of the soil conditions in the training database is mandatory in order to obtain time series of realistic LAI estimates and avoid meaningless values in the first rice growing stages.

Fig. 14 shows a typical example of the SPOT5 LAI evolution retrieved for a rice pixel in a representative Italian and Spanish ESUs. They reflect a common situation in which the soil was dry until DOY

132 (Italian area) and DOY 130 (Spanish area), and flooded immediately before sowing and during plant development. The figure compares the retrievals taking into account both dry and flooded conditions (red line), and the influence of using only dry soils (yellow line) or flooded soils (blue line). It can be observed that the use of a complete and representative spectral database (flooded + dry background) in the PROSAIL simulation provides realistic LAI evolutions (red line) which agree well with field measurements. It should be noted that in this situation LAI remains close to 0 before the rice emergence (from DOY 100 to 150 approx.), irrespectively of the flooded/non-flooded condition of the soil background. In a scenario where the training database does not characterize the flooded conditions (yellow line), the retrieved LAI is unrealistically high after the flooding during the first crop development stages, but as LAI increases canopy closure minimizes the importance of spectral background in the retrieval. The prediction uncertainty reflects this unreliability of the estimates, producing very high values (about 3) in the first crop stages (DOY 125–160). Similarly, when the training database does not characterize the dry conditions (blue line), the retrieved LAI is also unrealistically high before the real flooding of the fields.

In conclusion, the unrealistic effects produced by wrong spectra background can be very critical when time series analysis of rice crop dynamics is conducted. The increase of LAI at the beginning of the season would determine an unrealistic simulation of crop growth and identification of a wrong sowing date/emergence period respectively. This assessment demonstrates the importance of training correctly the RTM to produce a reliable data set for the retrieval process. In addition, retrieval of prediction uncertainty together with the LAI map can be useful to automatically identify non-agronomical areas and targets for which the model was not calibrated and/or performed well.

Table 3
Statistical scores between the Landsat estimated LAI taking into account different background conditions and ground measurements over Spain. In all field conditions, the number of samples used for the generation of the statistics was 171.

	ME	RMSE	MAE	R ²
Dry	-0.50	1.09	0.85	0.81
Flooded	0.10	0.62	0.40	0.90
Dry+flooded	0.08	0.39	0.28	0.95

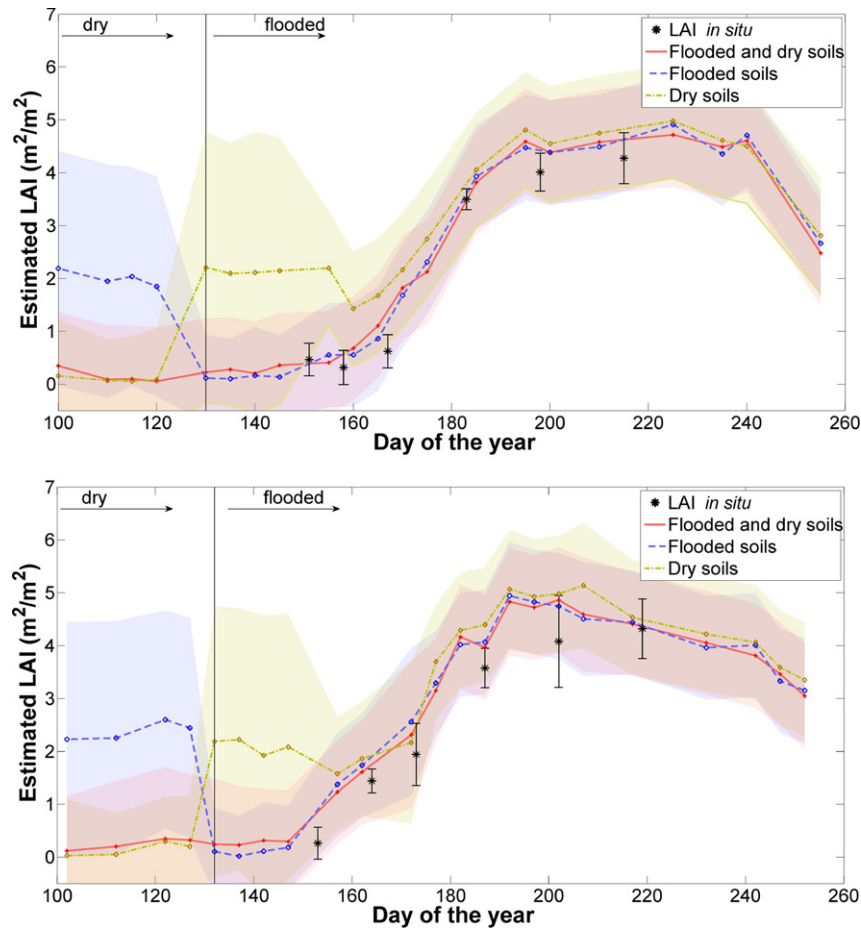


Fig. 14. Temporal evolution of GPR-SPOT5 LAI estimates within an ESU using different soil background in (top) Spain and (bottom) Italy.

5. Conclusions

In this study, a fully operational chain for deriving high-resolution LAI maps over Mediterranean rice crops is presented. The main novelties of this work include the following: 1) the use of sound machine learning algorithms trained on simulated RTM data specifically generated to characterize rice cropping features, 2) LAI estimates were originally validated with measurements acquired by a smartphone app (at affordable cost both in time, post-processing and human resources), 3) spatially and time-resolved estimations of LAI (not just static estimates) were produced, 4) different areas with particularly different rice crop specificities and spectral background conditions related to water management were analyzed, 5) seasonal remote sensing data from two multispectral sensors were exploited allowing to simulate the future potential of the operational chain with the Sentinel-2 constellation, 6) band ranking for rice monitoring where vegetation and water features make the problem more challenging was studied, and 7) the estimation of prediction uncertainties in a temporal manner were derived from the Gaussian process model and their validity and consistency were assessed. The proposed approach was illustrated in two rice study areas characterized by different rice varieties and agro-practices.

The production of the multitemporal local LAI maps was based on the PROSAIL RTM inversion with Gaussian process regression and Landsat/SPOT5 surface reflectance data. GPR proved to be a highly efficient and accurate method to invert the PROSAIL model, outperforming NN and KRR. However, the flexibility of the processing chain allows its application to any other regression method of choice. Theoretical performances of the regressions were satisfactory, and

a deeper analysis of the GPR-Landsat and GPR-SPOT5 training process showed that both models identified the green and near infrared channels as the most relevant bands in the retrieval process.

The multitemporal HR LAI maps captured the range and the temporal evolution of rice growth and resulted in agreement with corresponding *in situ* LAI measurements obtained from PocketLAI in two different countries. In addition, a comparison of the multitemporal estimates provided by both GPR-Landsat and GPR-SPOT5 retrievals showed good temporal consistency between them. However, a slight overestimation of the GPR-SPOT5 was observed during the period of the rice development.

In the development of the operational chain, the only user-demanding requirement refers to the provision of physical prior knowledge: site-specific characteristics need to be introduced in the model parameterization, such as rice plant and soil characteristics of the interested study area. Results highlight how a correct characterization of the underlying rice background is needed to obtain realistic LAI estimates during the initial development stages of the plants. The use of a training database comprising flooded and dry soil signatures showed to be robust against changes in background condition related to water management.

GPR retrievals also provided associated uncertainty for the predictions as a very valuable side information. This uncertainty is helpful a) for the scientist as a product to understand where potential errors in the retrieval exist, and b) for expert users, such as crop modelers, in order to weight LAI estimates in the assimilation process according to their error/goodness. The prediction uncertainty can also be used for diagnosing the presence of surfaces not addressed in the simulated database. Therefore, inspection of the uncertainty

provided with insights to (1) refine the selection of PROSAIL inputs, such as certain backgrounds not initially included in the preliminary retrieval, and (2) identify non-rice pixels thus identifying possible errors in the rice mask.

The near-real time production of HR LAI maps is useful in planning the management practices (i.e. fertilization) to minimize the yield pattern variability within each parcel. This information is particularly important for precision farming activities, since farmers are expecting to be supported in prescription map production for top-dress fertilization. A dense temporal data set of LAI maps is also fundamental to perform expert crop monitoring and/or improve crop model estimations exploiting assimilation techniques. It is important to mention that, whenever no external information able to represent the real variability in the field is provided for crop modeling, model simulation will provide the same results. This is the case when the aim is to apply crop models in an operational way at a parcel level. It is in fact not possible to obtain micro-meteorological information able to provide information changing from field to field. Moreover, it is not realistic to have detailed soil maps that usually exist at regional level with a scale ranging from 1:25.000 to 1:100.000. In this case, if the sowing date and variety are the same, or slightly different, the only way to capture the real spatio-temporal changes in crop development and production is to assimilate exogenous observation of crop status such as the information provided by EO LAI maps.

These results demonstrate the consistency and robustness of the presented processing chain for rice monitoring, which can be considered satisfactory for the production of HR LAI maps, and suitable to be assimilated by crop models. The proposal is aimed to improve crop monitoring and is specially suited for precision agriculture applications. Future work will include the use of the Sentinel-2 constellation whose spectral and temporal characteristics will make possible to increase the temporal resolution of the LAI estimates due to the combination with Landsat imagery, leading to further high-resolution multi-sensor studies in the framework of crop monitor and management specially in near real time. The presented processing chain has been operational during the 2016 ERMES rice season.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement 606983, the European Research Council (ERC) funding under the ERC-CoG-2014 SEDAL under grant agreement 647423, and the Land-SAF (the EUMETSAT Network of Satellite Application Facilities) project. The authors would like to deeply thank the Spanish and Italian rice farmers for allowing collecting *in situ* data within their fields and Cassandra Lab. for providing the PocketLAI application.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.rse.2016.10.009>.

References

- Arvidson, T., Goward, S., Gasch, J., Williams, D., 2006. Landsat-7 long-term acquisition plan. *Photogramm. Eng. Remote. Sens.* 72, 1137–1146. <http://dx.doi.org/10.14358/PERS.72.10.1137>.
- Atzberger, C., Richter, K., 2012. Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. *Remote Sens. Environ.* 120, 208–218. <http://dx.doi.org/10.1016/j.rse.2011.10.035>.
- Baillarin, S., Gigord, P., Hagolle, O., 2008. Automatic registration of optical images, a stake for future missions: application to ortho-rectification, time series and mosaic products. *Geoscience and Remote Sensing Symposium, 2008. IGARSS* 2008. IEEE International. vol. 2. II-1112–II-1115. <http://dx.doi.org/10.1109/IGARSS.2008.4779194>.
- Baret, F., Clevers, J., Steven, M., 1995. The robustness of canopy gap fraction estimates from red and near-infrared reflectances: a comparison of approaches. *Remote Sens. Environ.* 54, 141–151. [http://dx.doi.org/10.1016/0034-4257\(95\)00136-0](http://dx.doi.org/10.1016/0034-4257(95)00136-0).
- Baret, F., Fourty, T., 1997. Estimation of leaf water content and specific leaf weight from reflectance and transmittance measurements. *Agronomie* 17, 455–464. <http://dx.doi.org/10.1051/agro:19970903>.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J.L., Leroy, M., 2007. LAI, FAPAR and FCOVER CYCLOPES global products derived from VEGETATION: part 1: principles of the algorithm. *Remote Sens. Environ.* 110, 275–286. <http://dx.doi.org/10.1016/j.rse.2007.02.018>.
- Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholczyk, P., Smets, B., 2013. Geov1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: principles of development and production. *Remote Sens. Environ.* 137, 299–309. <http://dx.doi.org/10.1016/j.rse.2012.12.027>.
- Borel, C.C., 2010. *Vegetative canopy parameter retrieval using 8-band data*. Digital-Globe 8-Band Research Challenge 2010.
- Boschetti, M., Nutini, F., Manfron, G., Brivio, P.A., Nelson, A., 2014. Comparative analysis of normalised difference spectral indices derived from modis for detecting surface water in flooded rice cropping systems. *PLoS one* 9, e88741. <http://dx.doi.org/10.1371/journal.pone.0088741>.
- Boschetti, M., Stroppiana, D., Brivio, P., Bocchi, S., 2009. Multi-year monitoring of rice crop phenology through time series analysis of MODIS images. *Int. J. Remote Sens.* 30, 4643–4662. <http://dx.doi.org/10.1080/01431160802632249>.
- Breda, N.J.J., 2003. Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. *J. Exp. Bot.* 54, 2403–2417. <http://dx.doi.org/10.1093/jxb/erg263>.
- Bsaibes, A., Courault, D., Baret, F., Weiss, M., Olioso, A., Jacob, F., Hagolle, O., Marloie, O., Bertrand, N., Desfond, V., Kzemipour, F., 2009. Albedo and LAI estimates from FORMOSAT-2 data for crop monitoring. *Remote Sens. Environ.* 113, 716–729. <http://dx.doi.org/10.1016/j.rse.2008.11.014>.
- Campos-Taberner, M., García-Haro, F.J., Moreno, A., Gilabert, M.A., Martínez, B., Sánchez-Ruiz, S., Camps-Valls, G., 2015a. Development of an earth observation processing chain for crop bio-physical parameters at local scale. *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. pp. 29–32. <http://dx.doi.org/10.1109/IGARSS.2015.7325689>.
- Campos-Taberner, M., García-Haro, F., Moreno, A., Gilabert, M., Sanchez-Ruiz, S., Martínez, B., Camps-Valls, G., 2015b. Mapping leaf area index with a smartphone and Gaussian processes. *Geosci. Remote Sens. Lett. IEEE* 12, 2501–2505. <http://dx.doi.org/10.1109/LGRS.2015.2488682>.
- Campos-Taberner, M., García-Haro, F.J., Confalonieri, R., Martínez, B., Moreno, A., Sánchez-Ruiz, S., Gilabert, M.A., Camacho, F., Boschetti, M., Busetto, L., 2016. Multitemporal monitoring of plant area index in the valencia rice district with pocketlai. *Remote Sens.* 8, 202. <http://dx.doi.org/10.3390/rs8030202>.
- Camps-Valls, G., Muñoz-Marí, J., Verrelst, J., Mateo, F., Gomez-Dans, J., 2016. A survey on Gaussian processes for Earth observation data analysis. *IEEE Geosci. Remote Sens. Mag.* 4, 58–78. <http://dx.doi.org/10.1109/MGRS.2015.2510084>.
- Camps-Valls, G., Tuia, D., Gómez-Chova, L., Jiménez, S., Malo, J., 2011. *Remote Sensing Image Processing*, vol. 5. Morgan & Claypool Publishers. <http://dx.doi.org/10.2200/S00392ED1V01Y201107VM012>.
- Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.* 62, 241–252. [http://dx.doi.org/10.1016/S0034-4257\(97\)00104-1](http://dx.doi.org/10.1016/S0034-4257(97)00104-1).
- Chen, J., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S., White, H., Hall, R., Peddle, D., King, D., Trofymow, J., Swift, E., der Sanden, J.V., Pellikka, P., 2002. Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sens. Environ.* 80, 165–184. [http://dx.doi.org/10.1016/S0034-4257\(01\)00300-5](http://dx.doi.org/10.1016/S0034-4257(01)00300-5).
- 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 Sens. Environ.* 115, 1053–1064. <http://dx.doi.org/10.1016/j.rse.2010.12.010>.
- Chen, J.M., Black, T.A., 1992. Defining leaf area index for non-flat leaves. *Plant, Cell Environ.* 15, 421–429. <http://dx.doi.org/10.1111/j.1365-3040.1992.tb00992.x>.
- Claverie, M., Vermote, E.F., Weiss, M., Baret, F., Hagolle, O., Demarez, V., 2013. Validation of coarse spatial resolution LAI and FAPAR time series over cropland in southwest France. *Remote Sens. Environ.* 139, 216–230. <http://dx.doi.org/10.1016/j.rse.2013.07.027>.
- Clevers, J., 1988. The derivation of a simplified reflectance model for the estimation of leaf area index. *Remote Sens. Environ.* 25, 53–69. [http://dx.doi.org/10.1016/0034-4257\(88\)90041-7](http://dx.doi.org/10.1016/0034-4257(88)90041-7).
- Colombo, R., Bellingeri, D., Fasolini, D., Marino, C.M., 2003. Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sens. Environ.* 86, 120–131. [http://dx.doi.org/10.1016/S0034-4257\(03\)00094-4](http://dx.doi.org/10.1016/S0034-4257(03)00094-4).
- Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, D., Pragnère, A., Myneni, R., Knyazikhin, Y., Wang, L., 2003. Retrieval of canopy biophysical variables from bidirectional reflectance: using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* 84, 1–15. [http://dx.doi.org/10.1016/S0034-4257\(02\)00035-4](http://dx.doi.org/10.1016/S0034-4257(02)00035-4).
- Confalonieri, R., Foi, M., Casa, R., Aquaro, S., Tona, E., Peterle, M., Boldini, A., Carli, G.D., Ferrari, A., Finotto, G., Guarneri, T., Manzoni, V., Movedi, E., Nisoli, A., Paleari, L., Radici, I., Suardi, M., Veronesi, D., Bregaglio, S., Cappelli, G., Chiodini, M., Dominoni, P., Francone, C., Frasso, N., Stella, T., Acutis, M., 2013. Development of an app for estimating leaf area index using a smartphone. *Truthness and precision*

- determination and comparison with other indirect methods. *Comput. Electron. Agric.* 96, 67–74. <http://dx.doi.org/10.1016/j.compag.2013.04.019>.
- Confalonieri, R., Rosenmund, A.S., Baruth, B., 2009. An improved model to simulate rice yield. *Agron. Sustain. Dev.* 29, 463–474. <http://dx.doi.org/10.1051/agro/2009005>.
- Dorigo, W., Zurita-Milla, R., de Wit, A., Brazile, J., Singh, R., Schaepman, M., 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *Int. J. Appl. Earth Obs. Geoinf.* 9, 165–193. <http://dx.doi.org/10.1016/j.jag.2006.05.003>.
- Drusch, M., Bello, U.D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* 120, 25–36. <http://dx.doi.org/10.1016/j.rse.2011.11.026>.
- Duan, S.-B., Li, Z.-L., Wu, H., Tang, B.-H., Ma, L., Zhao, E., Li, C., 2014. Inversion of the PROSAIL model to estimate leaf area index of maize, potato, and sunflower fields from unmanned aerial vehicle hyperspectral data. *Int. J. Appl. Earth Obs. Geoinf.* 26, 12–20. <http://dx.doi.org/10.1016/j.jag.2013.05.007>.
- Duveiller, G., Weiss, M., Baret, F., Defourny, P., 2011. Retrieving wheat green area index during the growing season from optical time series measurements based on neural network radiative transfer inversion. *Remote Sens. Environ.* 115, 887–896. <http://dx.doi.org/10.1016/j.rse.2010.11.016>.
- European Commission, 2011. *Natura 2000*. (Accessed on May 2016) http://ec.europa.eu/environment/nature/natura2000/index_en.htm
- Fang, H., Liang, S., 2005. A hybrid inversion method for mapping leaf area index from MODIS data: experiments and application to broadleaf and needleleaf canopies. *Remote Sens. Environ.* 94, 405–424. <http://dx.doi.org/10.1016/j.rse.2004.11.001>.
- Francone, C., Pagani, V., Foi, M., Cappelli, G., Confalonieri, R., 2014. Comparison of leaf area index estimates by ceptomat and pocketlai smart app in canopies with different structures. *Field Crop Res.* 155, 38–41. <http://dx.doi.org/10.1016/j.fcr.2013.09.024>.
- Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimation of leaf area index, FAPAR, and net primary production of terrestrial ecosystems. *Remote Sens. Environ.* 70, 29–51. [http://dx.doi.org/10.1016/S0034-4257\(99\)00056-5](http://dx.doi.org/10.1016/S0034-4257(99)00056-5).
- Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* 90, 337–352. <http://dx.doi.org/10.1016/j.rse.2003.12.013>.
- Hagolle, O., Dedieu, G., Mougenot, B., Debaecker, V., Duchemin, B., Meygret, A., 2008. Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: application to Formosat-2 images. *Remote Sens. Environ.* 112, 1689–1701. <http://dx.doi.org/10.1016/j.rse.2007.08.016>.
- Haykin, S., 1999. *Neural Networks—A Comprehensive Foundation*. 2nd ed., Prentice Hall.
- Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., Ferreira, L., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. [http://dx.doi.org/10.1016/S0034-4257\(02\)00096-2](http://dx.doi.org/10.1016/S0034-4257(02)00096-2).
- Jacquemoud, S., Bacour, C., Poilvé, H., Frangi, J.-P., 2000. Comparison of four radiative transfer models to simulate plant canopies reflectance: direct and inverse mode. *Remote Sens. Environ.* 74, 471–481. [http://dx.doi.org/10.1016/S0034-4257\(00\)00139-5](http://dx.doi.org/10.1016/S0034-4257(00)00139-5).
- Jacquemoud, S., Baret, F., 1990. Prospect: A model of leaf optical properties spectra. *Remote Sens. Environ.* 34, 75–91. [http://dx.doi.org/10.1016/0034-4257\(90\)90100-Z](http://dx.doi.org/10.1016/0034-4257(90)90100-Z).
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. PROSPECT + SAIL models: a review of use for vegetation characterization. *Remote Sens. Environ.* 113, Supplement 1, S56–S66. <http://dx.doi.org/10.1016/j.rse.2008.01.026>.
- Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., Baret, F., 2004. Review of methods for in situ leaf area index determination: part I. Theories, sensors and hemispherical photography. *Agric. For. Meteorol.* 121, 19–35. <http://dx.doi.org/10.1016/j.agrformet.2003.08.027>.
- Ju, J., Roy, D.P., 2008. The availability of cloud-free landsat ETM+ data over the conterminous united states and globally. *Remote Sens. Environ.* 112, 1196–1211. <http://dx.doi.org/10.1016/j.rse.2007.08.011>.
- Lazaro-Gredilla, M., Tsisias, M., Verrelst, J., Camps-Valls, G., 2014. Retrieval of biophysical parameters with heteroscedastic Gaussian processes. *Geosci. Remote Sens. Lett. IEEE* 11, 838–842. <http://dx.doi.org/10.1109/LGRS.2013.2279695>.
- Maclean, J.L., 2002. *Rice almanac: source book for the most important economic activity on earth*. *Int. Rice Res. Inst.*
- Malenovsky, Z., Rott, H., Cihlar, J., Schaepman, M.E., García-Santos, G., Fernandes, R., Berger, M., 2012. Sentinels for science: potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sens. Environ.* 120, 91–101. <http://dx.doi.org/10.1016/j.rse.2011.09.026>.
- Mckay, M.D., Beckman, R.J., Conover, W.J., 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 42, 55–61. <http://dx.doi.org/10.1080/00401706.2000.10485979>.
- Meroni, M., Colombo, R., Panigada, C., 2004. Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. *Remote Sens. Environ.* 92, 195–206. <http://dx.doi.org/10.1016/j.rse.2004.06.005>.
- Montavon, G., Braun, M.L., Krueger, T., Muller, K.R., 2013. Analyzing local structure in kernel-based learning: explanation, complexity, and reliability assessment. *IEEE Signal Process. Mag.* 30, 62–74. <http://dx.doi.org/10.1109/MSP.2013.2249294>.
- Morisette, J., Baret, F., Privette, J., Myneni, R., Nickeson, J., Garrigues, S., Shabanov, N., Weiss, M., Fernandes, R., Leblanc, S., Kalacska, M., Sanchez-Azofeifa, G., Chubey, M., Rivard, B., Stenberg, P., Rautiainen, M., Voipio, P., Manninen, T., Pilaht, A., Lewis, T., Iliames, J., Colombo, R., Meroni, M., Busetto, L., Cohen, W., Turner, D., Warner, E., Petersen, G., Seufert, G., Cook, R., 2006. Validation of global moderate-resolution LAI products: a framework proposed within the CEOS land product validation subgroup. *IEEE Trans. Geosci. Remote Sens.* 44, 1804–1817. <http://dx.doi.org/10.1109/TGRS.2006.872529>.
- Myneni, R., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G., Lotsch, A., Friedl, M., Morisette, J., Votava, P., Nemani, R., Running, S., 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* 83, 214–231. [http://dx.doi.org/10.1016/S0034-4257\(02\)00074-3](http://dx.doi.org/10.1016/S0034-4257(02)00074-3).
- Myneni, R., Ramakrishna, R., Nemani, R., Running, S., 1997. Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Trans. Geosci. Remote Sens.* 35, 1380–1393. <http://dx.doi.org/10.1109/36.649788>.
- Quaife, T., Lewis, P., Kauwe, M.D., Williams, M., Law, B.E., Disney, M., Bowyer, P., 2008. Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter. *Remote Sens. Environ.* 112, 1347–1364. <http://dx.doi.org/10.1016/j.rse.2007.05.020>.
- Rasmussen, C.E., Williams, C.K.I., 2006. *Gaussian Processes for Machine Learning*. The MIT Press, New York.
- Roy, D., Kovalsky, V., Zhang, H., Vermote, E., Yan, L., Kumar, S., Egorov, A., 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* 185, 57–70. <http://dx.doi.org/10.1016/j.rse.2015.12.024>.
- Roy, D., Wulder, M., Loveland, T., C.E. W., Allen, R., Anderson, M., Helder, D., Irons, J., Johnson, D., Kennedy, R., Scambos, T., Schaaf, C., Schott, J., Sheng, Y., Vermote, E., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Hipple, J., Hostert, P., Huntington, J., Justice, C., Kilic, A., Kovalsky, V., Lee, Z., Lyburner, L., Masek, J., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R., 2014. Landsat-8: science and product vision for terrestrial global change research. *Remote Sens. Environ.* 145, 154–172. <http://dx.doi.org/10.1016/j.rse.2014.02.001>.
- Shawe-Taylor, J., Cristianini, N., 2004. *Kernel Methods for Pattern Analysis*. Cambridge University Press.
- Smith, J., 1993. LAI inversion using backpropagation neural network trained with multiple scattering model. *IEEE Trans. Geosci. Remote Sens.* 31, 1102–1106. <http://dx.doi.org/10.1109/36.263783>.
- Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. *Remote Sens. Environ.* 16, 125–141. [http://dx.doi.org/10.1016/0034-4257\(84\)90057-9](http://dx.doi.org/10.1016/0034-4257(84)90057-9).
- Vermote, E., Justice, C., Claverie, M., Franch, B., 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* 185, 46–56. <http://dx.doi.org/10.1016/j.rse.2016.04.008>.
- Verrelst, J., Camps-Valls, G., Muñoz-Mari, J., Rivera, J.P., Veroustraete, F., Clevers, J.G., Moreno, J., 2015a. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – a review. *ISPRS J. Photogramm. Remote Sens.* 108, 273–290. <http://dx.doi.org/10.1016/j.isprsjprs.2015.05.005>.
- Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J., Moreno, J., 2012a. Retrieval of vegetation biophysical parameters using Gaussian process techniques. *IEEE Trans. Geosci. Remote Sens.* 50, 1832–1843. <http://dx.doi.org/10.1109/TGRS.2011.2168962>.
- Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G., Moreno, J., 2012b. Machine learning regression algorithms for biophysical parameter retrieval: opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* 118, 127–139. <http://dx.doi.org/10.1016/j.rse.2011.11.002>.
- Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Mari, J., Clevers, J.G., Camps-Valls, G., Moreno, J., 2015b. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – a comparison. *ISPRS J. Photogramm. Remote Sens.* 108, 260–272. <http://dx.doi.org/10.1016/j.isprsjprs.2015.04.013>.
- Weiss, M., Baret, F., Smith, G., Jonckheere, I., Coppin, P., 2004. Review of methods for in situ leaf area index (LAI) determination: part II. Estimation of LAI, errors and sampling. *Agric. For. Meteorol.* 121, 37–53. <http://dx.doi.org/10.1016/j.agrformet.2003.08.001>.
- Woodcock, C., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thakabail, P., Vermote, E., Vogelmann, J., Wulder, M., Wynne, R., 2008. Free access to landsat imagery. *Science* 320, 1011. <http://dx.doi.org/10.1126/science.320.5879.1011a>.
- Zhang, X., Lu, C., Hu, N., Yao, K., Zhang, Q., Dai, Q., 2013. Simulation of canopy leaf inclination angle in rice. *Rice Science* 20, 434–441. [http://dx.doi.org/10.1016/S1672-6308\(13\)60161-4](http://dx.doi.org/10.1016/S1672-6308(13)60161-4).