Crop Biophysical Properties Estimation Based on LiDAR Full-Waveform Inversion Using the DART RTM

Sahar Ben Hmida, Abdelaziz Kallel, Jean-Philippe Gastellu-Etchegorry, Jean-Louis C. H. Roujean

Crop biophysical properties estimation based on LiDAR full waveform inversion using the DART RTM

Sahar Ben Hmida, Abdelaziz Kallel, Jean-Philippe Gastellu-Etchegorry and Jean-Louis Roujean

***Abstract*—This paper presents an inversion model that explores the potential of small footprint LiDAR waveforms for estimating crop biophysical properties. Here, we consider the height, LAI (Leaf Area Index) and ground reflectance of two maize and wheat fields. Within field, spatial heterogeneity is a source of difficulty for inversion approaches. For example, in the maize field, standard deviation is 0.16m for height and 0.6 for LAI. To mitigate this problem, we first classify all waveforms into maize and wheat clusters, and then assess the biophysical properties of each cluster with a procedure that uses a look up table (LUT) of waveforms. The latter ones are simulated by the Discrete Anisotropic Radiative Transfer (DART) model with 3D mockups of the crop fields for different crop properties and according to the LiDAR configuration. The approach was successfully tested with *in situ* measurements. Height is well estimated. Its root mean square error (RMSE) is 0.07 m and 0.04 m for maize and wheat, respectively. LAI estimate is also accurate (RMSE = 0.17) for maize except for wheat last growth stage (RMSE = 0.5), when LAI is very low because most leaves are dry. Field spatial heterogeneity explains that many clusters were needed to get accurate results.**

***Index Terms*—LiDAR waveform, Inversion, LUT, Crops biophysical properties, DART.**

I. Introduction

WHEAT and maize are the second and third major cereal crops in the world after rice. They are the most important food crops that provide the energy and proteins for human and animals. Estimating their yield and monitoring their health are then, very crucial. Vegetation height and leaf area index (LAI) are among the major biophysical parameters to estimate for studying canopy behavior and vegetation growth [1], [2], [3], [4]. LAI contributes particularly to different physical and biophysical processes like plant respiration and transpiration, photosynthesis, rainfall interception and vegetation carbon storage [5], [6], [7]. It explains that these biophysical parameters are helpful for managing irrigation, fertilization and insect control [8]. They can be measured in the field [9] with direct methods, but this approach is not suitable for large areas because it requires huge time, budget and destructive sampling [10].

The above remarks explain the widespread use of remote sensing approaches that rely on empirical relationships [11] and regression models [12], [13] between field measurements and vegetation indices derived from remote sensing data. Multi-spectral satellite images with high spatial resolution (e.g., Sentinel-2, Landsat, etc.) are commonly used to derive biophysical parameters such as vegetation LAI [14] chlorophyll amount [15], [16] and biomass [17]. However, present approaches with these passive sensors do not allow access to the crop 3D structure (e.g., canopy height) and suffer from limitations in vegetation parameters estimation especially in dense vegetation cover with high LAI values [18], [19], [20].

Light detection and ranging (LiDAR) is an active sensor that can characterize vegetation structure [21], [22], [23], [24] because it provides the time delay between the emitted and backscattered laser pulse. Recent studies were performed to estimate vegetation biophysical parameters such as LAI [1], [24], [25], [26], height [3], [27], [28], [29] and forest biomass [30], [31], [32]. Generally speaking, the accuracy of the retrieved parameters depends heavily on LiDAR point density [4].

Up to now, most of LiDAR research has focused on forests, in particular through the determination of parameters such as LAI and canopy height. Only few studies have been carried out to estimate crop biophysical parameters using terrestrial LiDAR system (TLS) [33], [34], [35] and airborne discrete return data [36], [37]. This small number of studies is mostly explained by the fact that the low height of crops compared to forests reduces the time of flight between the LiDAR first and last returns, which limits their accurate detection. In addition, the density of crop vegetation can also be a limiting factor because it decreases the penetration of the laser pulse through foliage.

Compared to the LiDAR discrete returns, LiDAR full waveforms record all backscattered energy from all illumined targets as a function of the range (height), which leads to more complete descriptions of the scattering events in vegetation canopy [38]. Nevertheless, we did not find in literature crop studies that are carried out with LiDAR waveforms. This explains why we investigated the possibility to estimate crop height and LAI with LiDAR waveforms.

The inversion of waveforms requires a good understanding of the physical processes that drive LiDAR laser beam interaction in vegetation canopies. It explains the development

S. Ben Hmida and A. Kallel are with the Advanced Technologies for Medecine and Signals Laboratory and the Digital Research Center of Sfax CRNS, Technopark of Sfax, PO Box 275, 3021 Sfax, Tunisia (e-mail: sahar.ben.hmida@gmail.com).

J.P. Gastellu-Etchegorry and S. Ben Hmida are with the Center for the study of the Biosphere from Space CESBIO, University of Toulouse, 31401 Toulouse cedex 9, France .

J.L. Roujean was with Météo France CNRM-GAME, F-31057 Toulouse, France.

Manuscript received Month day, year; revised Month day, year.

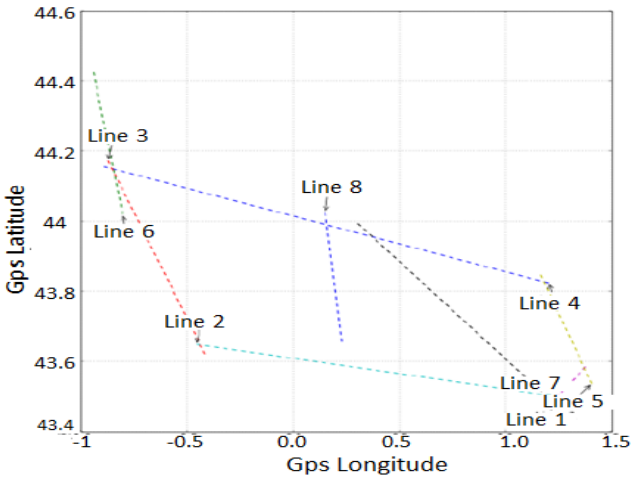


Fig. 1. Location of the 8 LiDAR flight lines in southwest France.

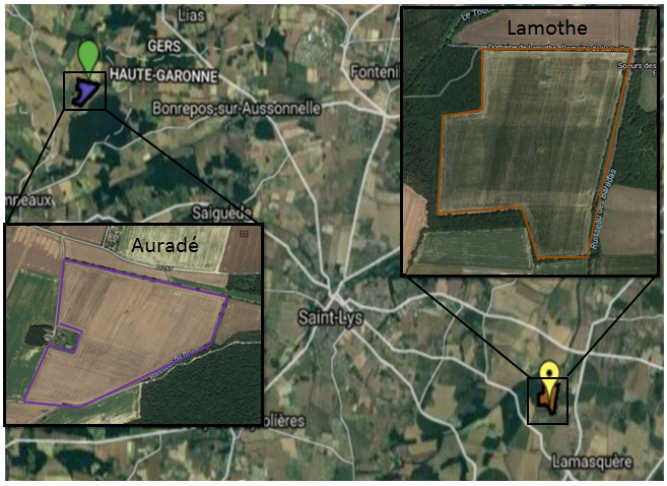


Fig. 2. Study areas from google maps. Purple lines delineate the experimental wheat field in Auradé (bottom left) and brown lines delineate the maize field in Lamothe (top right).

of Radiative Transfer Models (RTM) such as DART [39], [40] and Flight [41] which simulate this radiation interaction and consequently can link LiDAR observations and vegetation properties [24], [25]. The inversion of such models is considered to be a fully parametrized problem. Look Up Table (LUT), Quasi-Newton (QNT) and Neural Network (NNT) are the most popular methods to solve this inverse problem. They have been used in many studies to retrieve canopy properties [2], [3]. The LUT-based method is among the simplest ones to implement and less demanding in terms of computation. For one side, because it seeks from a table of simulated observations the closest one to the real measurement without *a priori* information requirement like QNT. For the other side, it does not need a large training database like the NNT. It has already been used to invert LiDAR waveform data in terms of forest biophysical parameters [3], [24], [42].

Here, we investigate the possibility to invert small footprint LiDAR waveforms in terms of crop properties. We consider the height, LAI and ground reflectance of two maize and wheat fields in southwest France. We present a LUT based inversion approach that relies on waveforms that are simulated with DART model [40].

II. Materials

The study is carried out in the frame of the research project Agriculture-Health-SPECTrometry (AHSPECT) funded by the European Facility for Airborne Research (EUFAR) and led by the National Center for Meteorological Research (CNRM). The aim is to study and monitor the health of the dominant crops of the southwest of France (wheat, maize and sunflower), near Toulouse. For this objective, on June 23, 2015, a LiDAR aboard the NERC DORNIER aircraft was flown along eight flights lines (Fig. 1) above selected and documented agricultural sites.

A. Study Area

Fig. 2 shows the two agricultural test sites that were considered: Lamothe ($43°29'47.7"N$, $1°14'15.9"E$) and Auradé ($43°32'55.5"N$, $1°06'21.4"E$). These two sites are characterized by hot and dry summers [17]. The usual crop rotation is "maize-wheat-maize-sunflower" for the Lamothe site and "wheat-sunflower-wheat-rape" for the Auradé site. At the time of the LiDAR campaign, most of surfaces (approximately 10 ha for Auradé and 25 ha for Lamothe) were planted with wheat and maize, respectively.

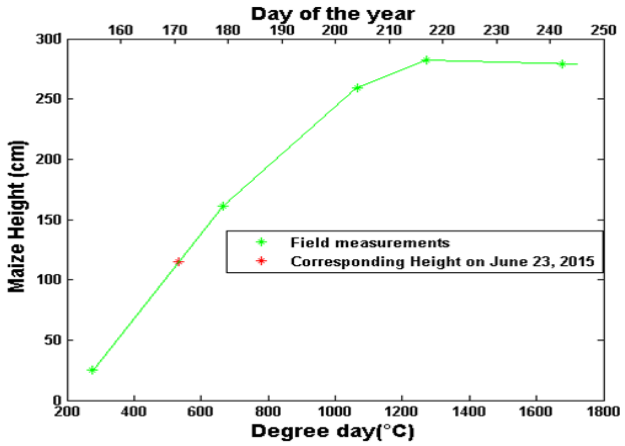
B. Field data

Field measurements were collected at the day of the LiDAR acquisition. This simultaneity is essential to validate our methodology. Ten sample plots were selected randomly in the Auradé wheat field, each one with a $25cm \times 25cm$ size. In each sample, the height values of all plants were measured which allowed us to compute the wheat mean height and its standard deviation (std). At the time of the LiDAR survey, the wheat was mature with zero LAI value, 0.49 m mean height and 0.07 m std.

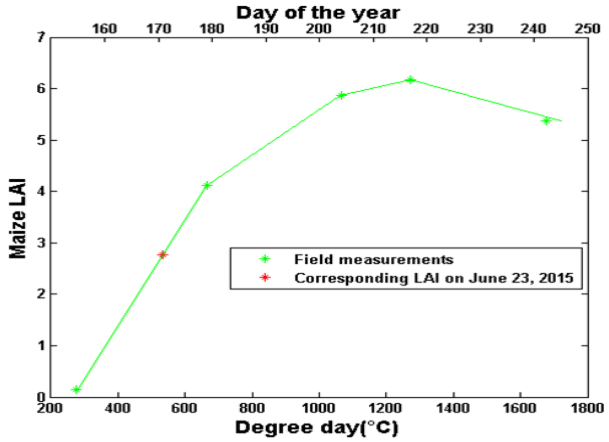
In Lamothe, field measurements were not collected at the time of LiDAR acquisition. However, maize field data were available at five dates before and after the LiDAR acquisition (Table I), which allowed us to assess the maize height and LAI by an interpolation procedure. For each measurement, five experiment plots were selected randomly in the field. In each plot, four maize plants were taken as samples at different places. The height and green leaves area of the twenty corns were measured. LAI was computed using the green leaves area and the inter-row and intra-row plant spacing. The interpolation procedure relies on the degree days (DD) since the date of sowing. Indeed, DD drives plant growth. It is computed with:

$$DD = \sum_{day} \left(\frac{T_{max} + T_{min}}{2} - T_{base} \right) \quad (1)$$

where T_{base} refers to the base temperature which is $10\,°C$ for maize corn [43], and T_{max} and T_{min} are the day maximum and minimum temperatures, respectively. They were measured



(a)



(b)

Fig. 3. Maize evolution of height (a) and LAI (b) according the degree day.

with a HMP155 thermal sensor. Fig. 3 shows the LAI and height evolution according to DD. On June 23, 2015, the cumulative temperature was $\approx 534\,^{\circ}$C, which allowed one to assess the height (≈ 1.15 m) and LAI (≈ 2.77). The height and LAI standard deviations were computed with the assumption that ratios height/std and LAI/std at the time of the LiDAR acquisition were equal to those ratios at the closest day of field measurement (i.e., July 1st, 2015). Hence, std was ≈ 0.16 m for height and ≈ 0.6 for LAI.

TABLE I
FIELD-MEASURED BIOPHYSICAL PARAMETERS FOR MAIZE CROPS AT DIFFERENT MEASUREMENT-DATE

Measurement date in 2015	Height: mean (m)	Height: standard deviation (std)	LAI: mean	LAI: standard deviation (std)
Jun, 2^{nd}	0.25	0.06	0.14	0.08
Jul, 1^{st}	1.61	0.22	4.11	0.9
Jul, 23	2.59	0.24	5.86	1.12
Aug, 6	2.82	0.1	6.17	0.8
Sep, 2^{nd}	2.78	0.15	5.37	0.97

C. LiDAR data

The LiDAR data were acquired on June 23, 2015 with a Leica Airborne laser scanner ALS50-II. This small footprint airborne sensor emits 4 to 9 ns laser pulses at 1064 nm wavelength and has a $\approx$ 22 cm diameter footprint at 1000 m altitude. It records either discrete points or full waveform data with 2 ns sampling rate (i.e., range resolution $\approx$ 30 cm). The echo intensity of the full waveform is recorded as a digital count at the predefined sample rate (i.e., bin). Here, the predefined number of bins per waveform is 256, which corresponds to 76.8 m. Among the 8 LiDAR flights lines, the lines 1 and 7 (Fig. 1) are over the studied fields in Lamothe and Auradé areas. LiDAR data were stored with the LiDAR data exchange format (LAS 1.3). Fig. 4 shows the location of the LiDAR data over the two studied maize and wheat fields: 29797 processed waveforms (i.e., $200\times15\ m^2$) in Auradé, and 18430 processed waveforms (i.e., $200\times9\ m^2$) in Lamothe. The density of the acquisition points is large ($\approx 10/m^2$), which is well adapted for accurate height estimation.

III. METHOD

The newly designed full waveform inversion method combines two complementary approaches: physical modeling and signal processing. Fig. 5 summarizes its three steps. In step 1, raw waveforms are pre-processed. In particular they are denoised, spatially averaged, and then classified into homogeneous clusters, which leads to a mean LiDAR pulse per cluster. In step 2, the three-dimensional (3-D) RTM DART model [40] simulates a Look up table of waveforms for specific biophysical characteristics (height, LAI and soil reflectance). In step 3, for each cluster, the inverse method selects the closest simulation of the LUT, which gives the desired biophysical properties. These three steps are detailed below.

A. LiDAR data Processing (step 1)

The spatial heterogeneity of environmental conditions (e.g., soil heterogeneity and moisture) in agricultural fields induces a spatial variability of crop development (vegetation height, LAI, etc.) [44], [45]. This statement was verified when analyzing the raw waveforms. Indeed, they were highly variable within each crop field. In addition, they were noisy. Hence, several pre-processing were applied before the inversion procedure itself, as described below.

1) Windowing: The so-called windowing procedure is a spatial averaging procedure that aims to remove the spatial variability of waveforms due to local environmental heterogeneity such as intra-rang spacing. It computes a mean waveform per spatial window. We considered three window sizes: $1\times1\ m^2$, $2\times2\ m^2$ and $3\times3\ m^2$ windows. One must note that this reduces considerably the noise level of waveforms.

2) Clustering (GMM distribution): The so-colled clustering procedure aims to remove outliers and to reduce the effect of noise. For that, it classifies waveforms into clusters, each one being characterized by a mean waveform. This, clustering is achieved with the Gaussian mixture model (GMM) one of the Expectation-Maximization algorithm application [46]. This algorithm seeks to find iteratively the maximum likelihood

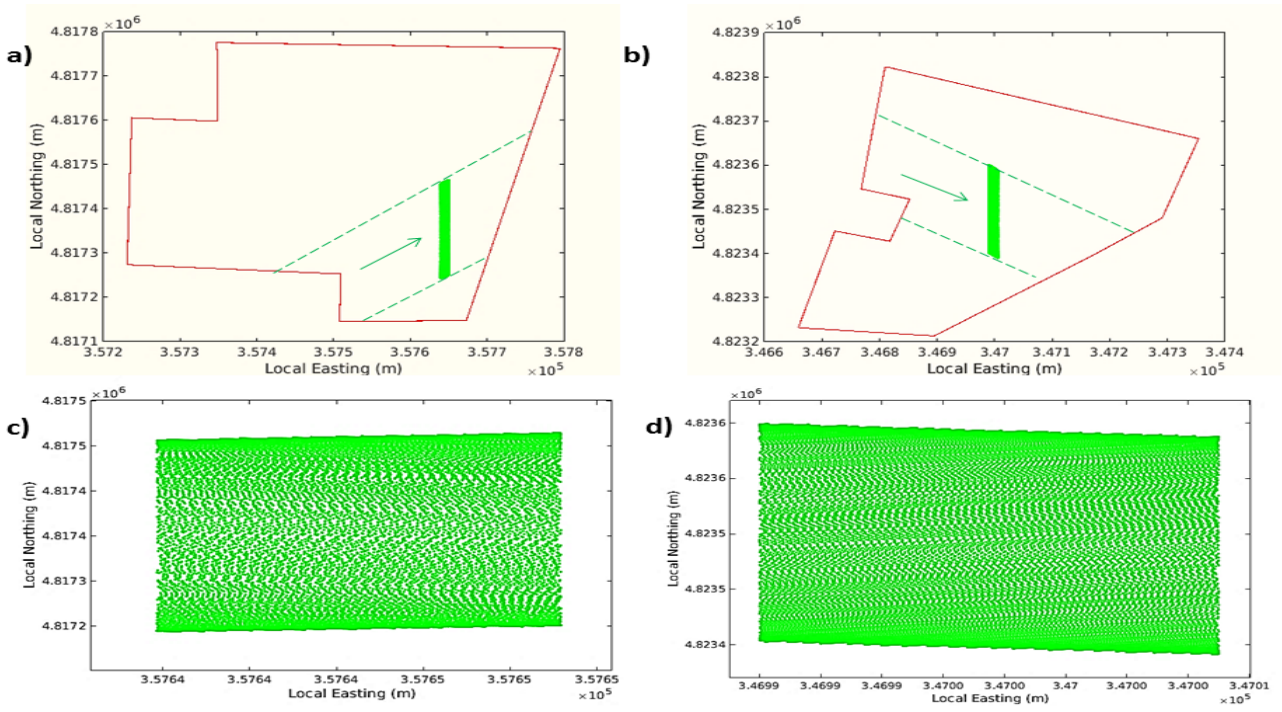


Fig. 4. LiDAR acquisitions in Lamothe (a) and Auradé (b). Red lines delineate the studied field, the green arrow indicates the LiDAR flightlines and green surfaces show the studied waveforms. c) and d) Zoom of the studied waveforms in a) and b), respectively.

estimator that an observation belongs to a given cluster. It starts from an initial guess probability $\gamma_{ij}^{(0)}$ to assess if the LiDAR observation $y_{i\in[1,n]}$ belongs to cluster $z_{j\in[1,k]}$, where n is the LiDAR observation number and k is the predefined number of clusters. Then, the method searches iteratively for the most likely combination $\theta^{(m)}$ between clusters and their related observations and its maximum probability estimator $\widehat{\theta}$:

$$\widehat{\theta}^{(m+1)} = argmax_{\theta} \sum_{j=1}^{k} (\sum_{i=1}^{n} \gamma_{ij}^{(m)}) \log \theta_j \tag{2}$$

Here $\gamma_{ij}^{(m)}$ is the membership relation between observation y_i and cluster z_j at m^{th} iteration:

$$\gamma_{ij}^{(m)} = P(z_j | y_i, \theta^{(m)}) \tag{3}$$

In order to get the best waveform clustering, this step is repeated several times before selecting the most likelihood result. Fig. 6 plots an example of waveforms that are classified into three clusters. The resulting clusters are homogeneous and different between each other.

3) Denoising method: LiDAR signals are always noisy due to background, detector and photon noise [23], [47], [48]. Hence, some denoising processing must be applied. For instruments with stable behavior, [49] suggest to use the noise tracking method with a specified threshold. This method considers that the waveform starting point is the bin in which the cumulative energy above the threshold reaches 1% of the total signal. Here, due to the small height of crops and the 2 ns sampling rate of the LiDAR sensor, the number of bins of the useful signal is too small to apply the noise tracking method. Hence, we adopted the simple threshold approach [50] with a noise threshold equal to the mean waveform amplitude during the propagation of laser pulse in the atmosphere plus six standard deviations. Fig. 7 shows an example of noise removal: after the waveform peak is detected, the first and last echoes are located by selecting waveform bins whose energies exceed the specified threshold.

B. LUT Generation (step 2)

1) DART Radiative transfer model: The DART (Discrete Anisotropic Radiative Transfer) model was used to simulate LiDAR waveforms. It is one of the most accurate and complete models that simulate radiative transfer from the visible to thermal infrared, in the Earth landscapes and the atmosphere. It has been developed since 1993 at CESBIO. It operates on three-dimensional heterogeneous natural (forest, grassland, etc.) and urban landscapes (buildings, house, etc.) in three modes: flux tracking [39], Monte Carlo [51] and LiDAR [40]. This model simulates Earth landscapes as a juxtaposition of voxels that are filled with fluids, facets (i.e., triangles and parallelograms) and the so-called turbid vegetation. The latter is defined by the optical properties, volume density and angular distribution of its foliar elements. Facets are used to simulate solid elements such as woody elements, urban elements, foliar elements, etc. They are defined by their orientation, area and optical properties. Furthermore, DART can simulate complex 3D scenes with imported 3D matrices of LAI and 3D objects (wheat, corn, rice, cherry tree, etc.) made of facets. 3D radiation interaction is simulated with two approaches: volume interaction for turbid medium [52] and surface interaction for facets [53]. In the LiDAR mode, Monte Carlo and flux tracking methods [40] are used to track each photon that is launched

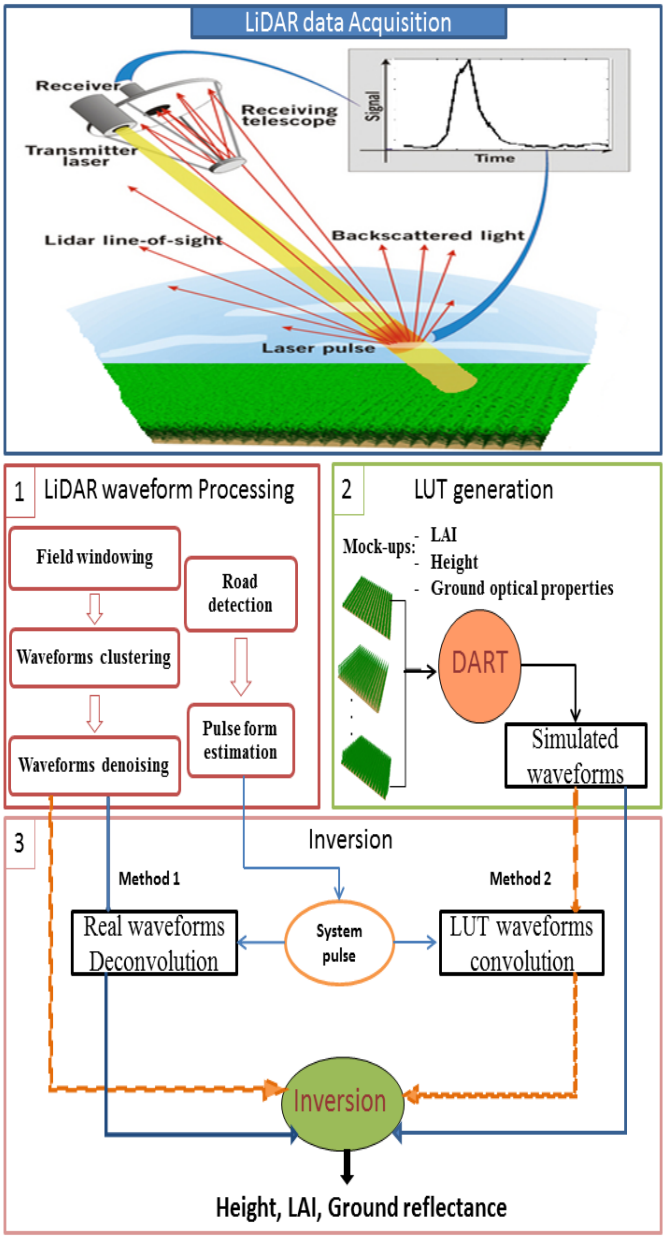


Fig. 5. Schematic diagram of the inversion method.

from the sensor until it is absorbed, it exits the scene or it is captured by the sensor.

2) Look up Table Generation: Here, the term LUT refers to a set of LiDAR waveforms that are simulated for the considered LiDAR instrument configuration (cf. II-C) and for a number of wheat and maize crop realizations. A realistic field crop mock-up was simulated per realisation, with constant and variable input parameters. Constant parameters were derived from field measurements (e.g., inter-row spacing and number of crops feet per m^2) or bibliography. Foliar optical properties were derived from the Prospect leaf model [54], [55]. The variable parameters are for different plant growth stages: stem length, LAI, presence of tassel, seeds, corn, etc. Table II gives the values of two constant parameters (crop density and inter-row spacing) and the ranges of three variable

TABLE II
INPUT PARAMETERS DESCRIPTION USED FOR CREATION OF THE LUT.

Crop	Crop density (plants/m^2)	Inter-row spacing (m)	Height (m)		LAI		Soil reflectance	
			Min	Max	Min	Max	Min	Max
Maize	9	0.755	0.3	2.6	0.24	5.9	0.3	0.6
Wheat	250	0.17	0.08	0.55	0	2.7	0.3	0.6

input parameters (soil reflectance and crop height and LAI). In addition to field measurements, the BBCH (Biologische Bundesanstalt fr Land und Forstwirtschaft) scale [56] was used to set the plant characteristics per considered growth stage: leaf shapes, number of leaves per plant, distance between leaves and presence of plant organs (e.g., tassel, corn, seeds, etc.) according to the stem length. To obtain accurate 3D plant mock-ups per growth stage, the 3D maize and wheat plants of the DART database were adapted by adding or removing leaves and resizing plant organs with the "Blender"[1] software. Then, DART used these 3D plants to create 3D mock-ups of maize and wheat fields. The area of these mock-ups is 9 m^2. Fig. 8 and 9 show mock-ups at several growth stages. For each mock-up, different LiDAR footprint positions were simulated, and their waveforms were averaged to obtain a statistically representative LiDAR waveform which was stored in the LUT.

C. Inversion (step 3)

1) LiDAR pulse delay processing: A LiDAR waveform can be considered as the convolution of the LiDAR system pulse and the cross section profile of the observed target [49], [57]. Individual target locations and distributions can be identified by application of fitting function [57] or by deconvolution [58]. However, in our study, the system pulse is unknown and consequently must be estimated. For that, roads were treated as reference targets. Indeed, with the assumption that roads are horizontal, which implies that their time response is very short, their recorded signal is nearly the LiDAR pulse. Road surfaces were therefore located and their waveforms were extracted. We considered that the pulse form is the average waveform of roads. Fig. 10 shows the estimated system pulse as well as a typical LiDAR waveform of a maize field. It illustrates that roads have a time response that is shorter than maize.

2) Inverse problem solution: Our inversion model compares the measured and simulated waveforms (Fig. 5). For that, it can use one of the two following comparison methods. In one method, the LiDAR waveform of each cluster is deconvolved with the estimated system pulse and then compared with the waveforms of the LUT. The deconvolution of the LiDAR signal uses the Wiener filter [59]. In the other method, the LUT waveforms are convolved with the estimated system pulse, and then compared with the waveform of each cluster. Here, the results of the two methods are compared. Results are found by minimizing a cost function which is the Root Mean Square Error RMSE between the measured waveform (W_R) and the simulated waveforms of the LUT (W_{LUT}), for each cluster. Before computing the RMSE per cluster, the simulated and

[1]Blender : https://www.blender.org/

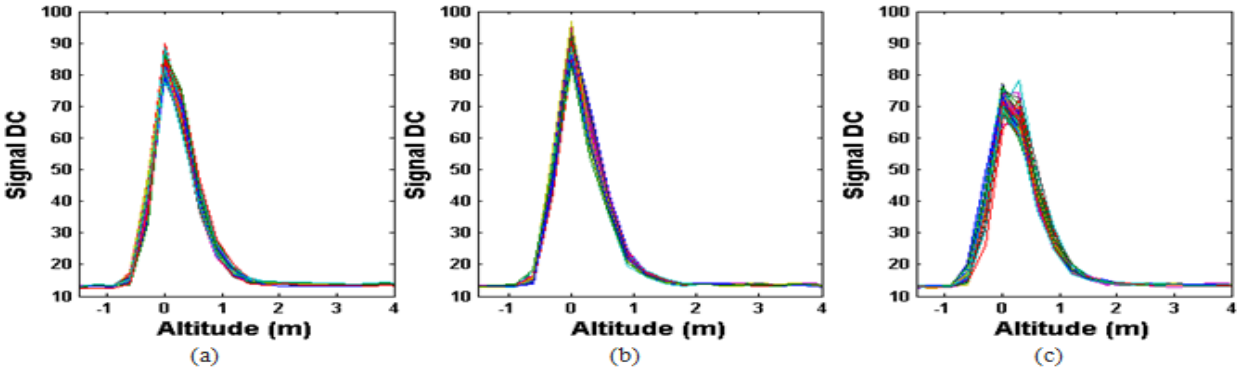


Fig. 6. Example of waveforms classification into 3 clusters: a) cluster 1, b) cluster 2 and c) cluster 3. DC stands for digital count.

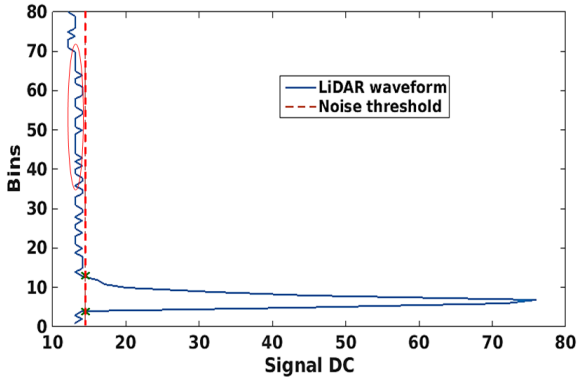


Fig. 7. Noise threshold (dashed red curve) for typical LiDAR waveform (blue curve) for the maize field. The two green crosses show the first and last echo locations. DC stands for digital count.The red ellipse indicates the bins that are used to calculate the mean LiDAR signal in the atmosphere

real waveforms were normalized relative to their maximum peak energies in order to be independent of the pulse energy, which ensured a maximal waveform compatibility.

$$RMSE = \sqrt{\frac{1}{n_{bin}} \sum_{i=1}^{n_{bin}} (W_R^i - W_{LUT}^i)^2} \qquad (4)$$

where n_{bin} is the number of bins of the denoised LiDAR signal (cf. III-A3).

The searched crop parameters (i.e., soil reflectance and canopy height and LAI) are assumed to be the parameters of the LUT simulation that has the minimal RMSE value.

IV. Results

Our inversion model was applied to the LiDAR data acquired over the maize and wheat sites. Its two comparison methods were used, and results were compared. Its performance was evaluated with a sensitivity analysis to the field window size (i.e., local spatial heterogeneity) and to the number of clusters used for classifying the LiDAR waveforms. For that, the histogram of each estimated vegetation property (LAI, height, soil reflectance) was computed and its statistical results were compared to field measurements.

A. Wheat crop

Wheat results assessed with the convolution of simulated LUT waveforms and the deconvolution of measured LiDAR waveforms are detailed below.

1) Convolution of the LUT: Fig. 11 displays the histograms of the wheat height, LAI and ground reflectance values that are assessed by convolving the LUT, for three window sizes and two cluster numbers (i.e., 3 and 5). Fig. 12 and Table III show their mean values, standard deviation (std) and standard deviation of their difference with *in situ* measurements.

The height and LAI estimations do not depend on the window size, whatever the number of clusters. For instance, we note two height values (0.5 m and 0.55 m: Fig. 11a, b and c) and two LAI values (0.38 and 0.67 in Fig. 11d, e and f). However, the spatial distribution of these estimated values changes with the window size. For the $1m \times 1m$ window size, height is 0.5 m for $\approx$ 90% of the field surface and 0.55 m for $\approx$ 10% of the wheat field. For the $2m \times 2m$ window size, the 0.5 m and 0.55 m proportions are $\approx$ 20% and 80%, respectively. For the $3m \times 3m$ window size, the 0.5 m and 0.55 m heights have similar proportions. The soil reflectance estimation is from 0.3 to 0.6 (Fig. 11g, h and i) depending on the number of clusters and the window size. However, more than 85% of the estimated values are around 0.3 and 0.4, which is realistic.

The mean crop height (Fig. 12) is well estimated for any cluster number: RMSE values are lower than 0.05 m (Table III). The best result is obtained with the $1m \times 1m$ window size (RMSE $\approx$ 0.01 m). Its estimated std is smaller than the std measured in the field. This is explained by the variability reduction due to the local (window) and global (cluster) averaging processes. LAI is less accurately retrieved than height. Its RMSE is around 0.5. As already mentioned, at the time of LiDAR acquisition, the actual LAI is null since wheat was in the last growth stage "Ripening" in which the wheat leaves were dried.

2) Deconvolution of real LiDAR data: Fig. 13 illustrates the case of the deconvolution of the real LiDAR waveforms. It shows the estimated height, LAI and soil reflectance according to the field window size for two cluster numbers. The height estimated values are 0.5 m (90%) and 0.55 m (10%) and the LAI estimated values are 0.38 and 0.67. The results are slightly

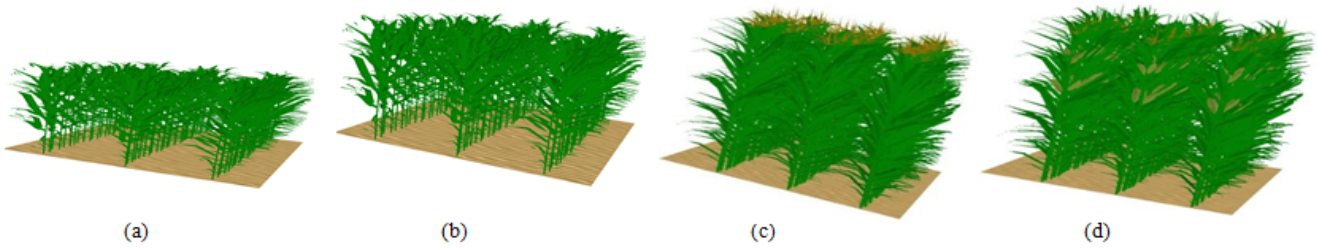


Fig. 8. Four DART maize mock-ups. a) BBCH 1: leaf development (height h = 60 cm). b) BBCH 3: stem stretching (h = 1.2 m). c) BBCH 5: maize panicle (h = 2 m). d) BBCH 8: seed maturation (h = 2.6 m).

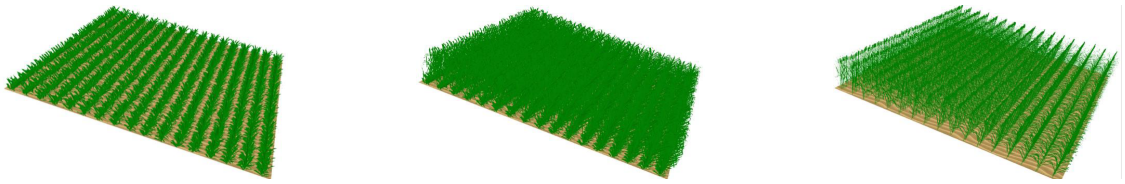


Fig. 9. Three DART wheat mock-ups. a) BBCH 1: Leaf development (h = 16 cm). b) BBCH 3: stem stretching (h = 45 cm). c) BBCH 8: wheat kernels maturation (h = 50 cm).

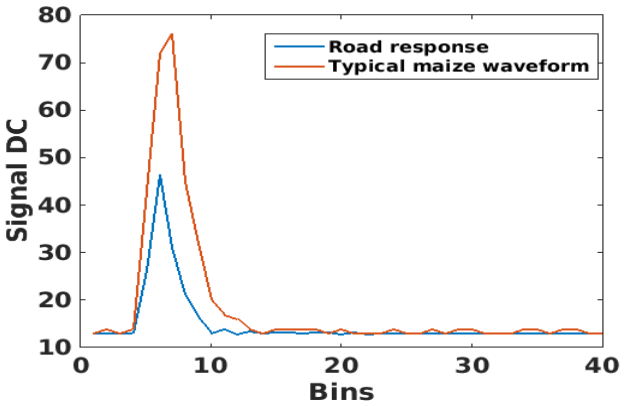


Fig. 10. Road LiDAR waveform (blue line) and typical maize LiDAR waveform.

TABLE III
MEAN ESTIMATED WHEAT FIELD HEIGHT, LAI AND SOIL REFLECTANCE, FOR 3 FIELD WINDOW SIZES AND 3 (BLUE COLOR) AND 5 (RED) CLUSTERS USED IN THE "LUT CONVOLUTION" APPROACH.

		Window size					
		$1m \times 1m$		$2m \times 2m$		$3m \times 3m$	
		Cluster Number		Cluster Number		Cluster Number	
Estimated Parameter		3	5	3	5	3	5
Soil reflectance	Mean	0.4	0.42	0.31	0.32	0.34	0.34
	std	0.02	0.07	0.05	0.06	0.05	0.05
LAI	Mean	0.38	0.4	0.63	0.61	0.54	0.53
	std	0	0.07	0.09	0.11	0.14	0.14
	RMSE	0.38	0.40	0.63	0.61	0.54	0.53
Height	Mean	0.50	0.50	0.55	0.54	0.53	0.52
	std	0.01	0.01	0.01	0.02	0.02	0.02
	RMSE	0.01	0.01	0.05	0.05	0.04	0.03

dependent on the cluster number and the window size. The distribution of the estimated parameter values is comparable

TABLE IV
MEAN ESTIMATED WHEAT FIELD HEIGHT, LAI AND SOIL REFLECTANCE, FOR 3 FIELD WINDOW SIZES AND 3 (BLUE COLOR) AND 5 (RED COLOR) CLUSTERS, USING THE "DECONVOLUTION OF REAL WAVEFORMS" APPROACH.

		Window size					
		$1m \times 1m$		$2m \times 2m$		$3m \times 3m$	
		Cluster Number		Cluster Number		Cluster Number	
Estimated Parameter		3	5	3	5	3	5
Soil reflectance	Mean	0.49	0.49	0.49	0.48	0.49	0.5
	std	0.04	0.04	0.02	0.03	0.11	0.09
LAI	Mean	0.38	0.42	0.38	0.38	0.42	0.41
	std	0	0.10	0.05	0.05	0.09	0.08
	RMSE	0.38	0.42	0.38	0.39	0.42	0.41
Height	Mean	0.5	0.50	0.50	0.50	0.50	0.50
	std	0	0.017	0.008	0.008	0.016	0.015
	RMSE	0.01	0.01	0.01	0.01	0.01	0.01

to that obtained with the LUT convolution approach with small differences in window percentages. Soil reflectance values vary between 0.3 and 0.6 depending on the window size.

Fig. 14 and Table IV plot the mean estimated parameters against the *in situ* measurements. The mean height RMSE is ≈ 0.01 m and the mean LAI RMSE is ≈ 0.4. It appears that the deconvolution of real LiDAR waveforms leads to slightly better results than the convolution of the LUT.

B. Maize crop

Maize results assessed with the convolution of simulated LUT waveforms and the deconvolution of measured LiDAR waveforms are detailed below. Compared to the wheat field, the number of clusters is larger in order to take into account the maize larger spatial variability.

1) Convolution of the LUT: The maize height, LAI, and soil reflectance histograms, obtained with the LUT waveform convolution, are shown in Fig. 15 for 3 cluster numbers (5,

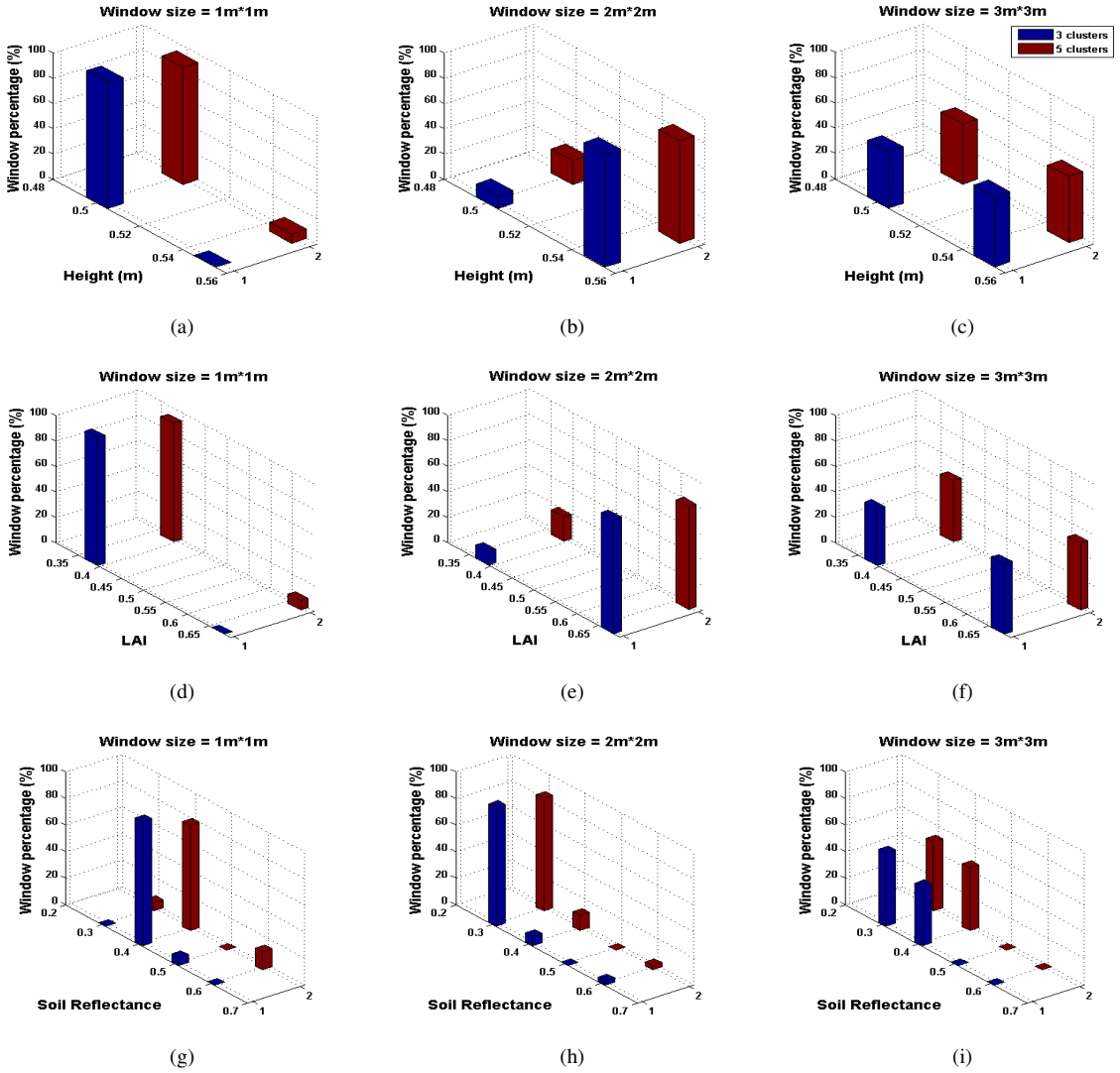


Fig. 11. Wheat field histograms, using the LUT convolution method: height (top), LAI (middle) and soil reflectance (bottom), for 3 field window sizes and 3 (blue color) and 5 (red) clusters.

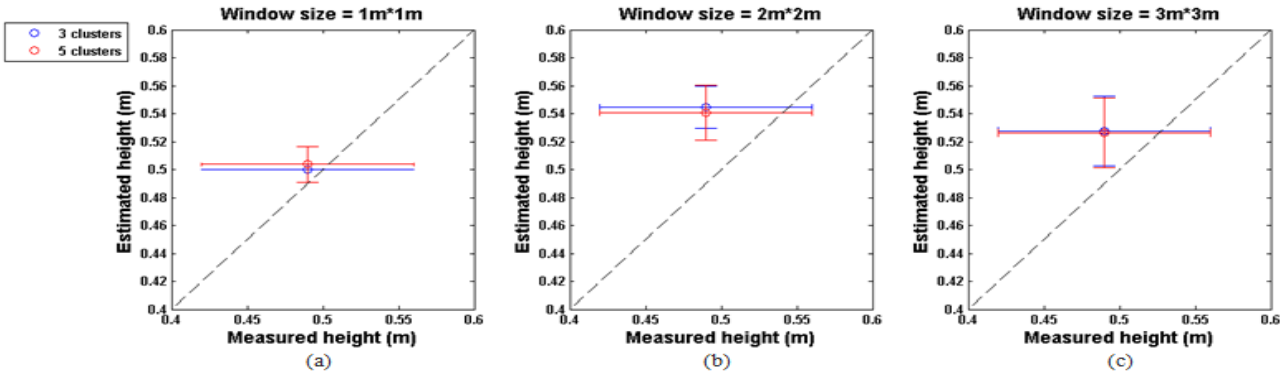


Fig. 12. Estimated wheat height against *in situ* measurement collected on 23 June 2015, for 3 (bleu) and 5 (red) clusters and 3 field window sizes: $1m^2$ (a), $4m^2$ (b), and $9m^2$ (c), using the LUT convolution. Circles represent the mean of estimated height. Vertical and horizontal error bars represent the standard deviation of the estimated height and field measurements, respectively.

8 and 10) and 3 window sizes. Fig. 16 indicates their mean values, std and difference compared to *in situ* measurements.

Height values estimated using $1m^2$, $4m^2$ and $9m^2$ window sizes have almost the same distribution: two major groups with low and high values. The group with low heights (i.e., 0.7 m, 0.8 m and 0.9 m) has an occurrence percentage smaller than

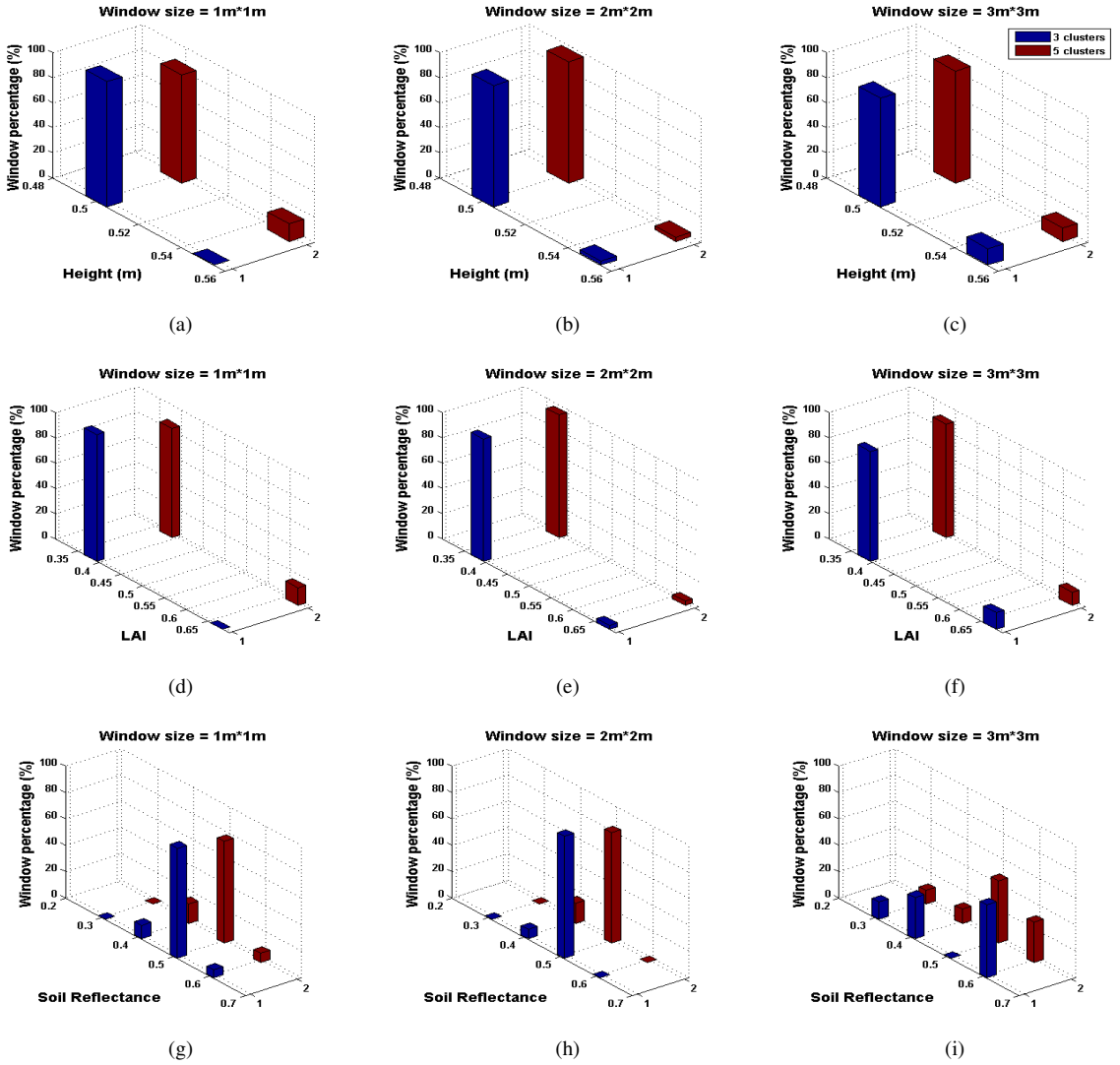


Fig. 13. wheat field histograms, using the waveform deconvolution: height (top), LAI (middle) and soil reflectance (bottom), for 3 field window sizes and 3 (blue color) and 5 (red) clusters.

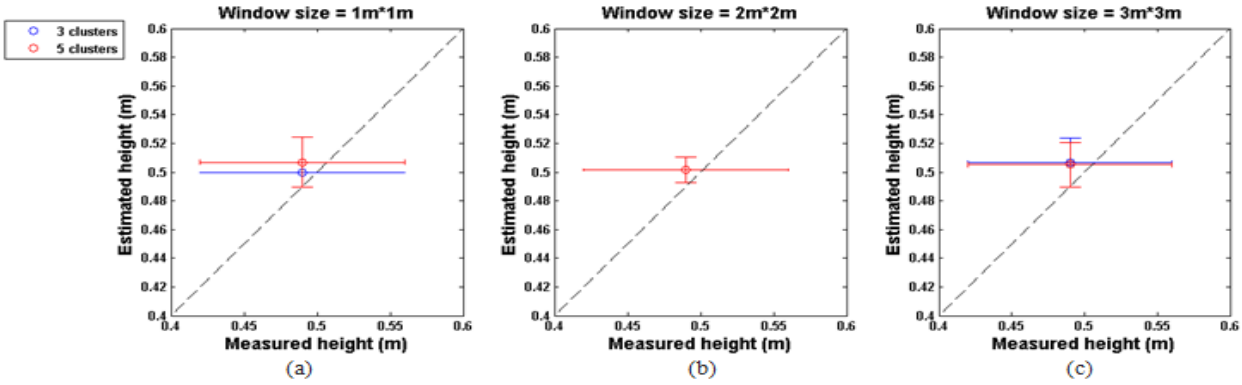


Fig. 14. Estimated wheat height against *in situ* measurement collected on 23 June 2015, for 3 (bleu) and 5 (red) clusters and 3 field window sizes: $1m^2$ (a), $4m^2$ (b), and $9m^2$ (c), using the deconvolution of LiDAR waveforms. Circles represent the mean of estimated height. Vertical and horizontal error bars represent the standard deviation of the estimated height and field measurements, respectively.

25%, whereas the group with high values (i.e., 1.2 m and 1.3 m) has a percentage larger than 75%. The retrieved LAI and height distributions are well correlated. Two LAI groups are distinguished: a most present group with LAI between 2.5 and 3, and a less present group with LAI from 0.98 up to 1.5. The estimated soil reflectance value ranges from 0.3 up to 0.6,

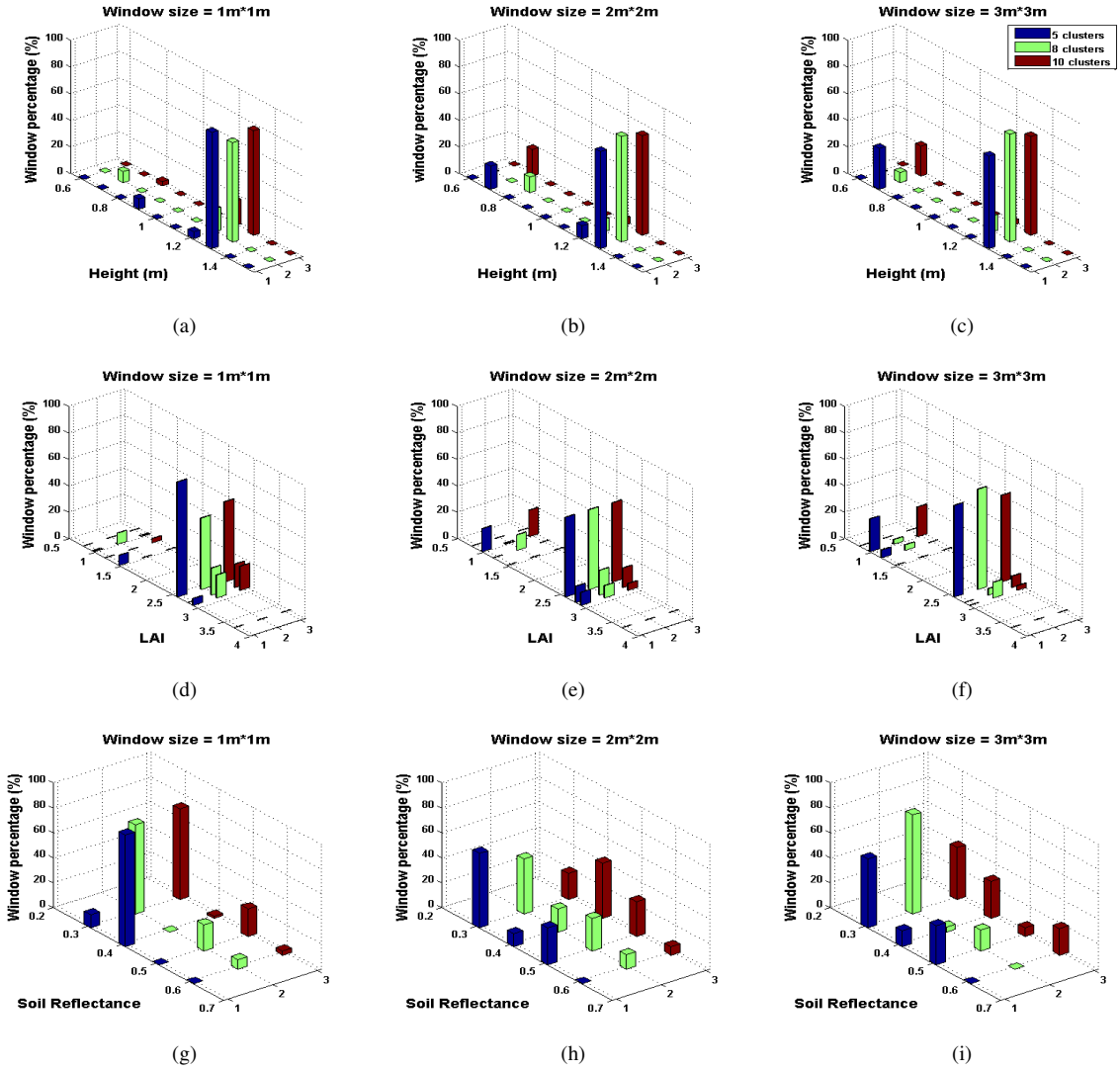


Fig. 15. Maize field histograms, using the LUT convolution method: height (top), LAI (middle) and soil reflectance (bottom), for 3 field window sizes and 5 (blue color), 8 (green) and 10 (red) clusters.

depending on the window size, with a [0.3 0.4] range for 70% of windows, which is realistic.

We analyzed the effect of the cluster number on the results accuracy. Mean maize height and LAI estimates do not depend on the cluster number for the $1m^2$ window size, with low RMSE around 0.07 m and 0.21, respectively, and corresponding std equal to 0.16 m and 0.4, respectively. For the other window sizes, the height and LAI are best estimated with a large cluster number. For example, height is estimated with low RMSE (0.01 m) for ten clusters (Fig. 16b and c) and LAI is estimated with relatively low RMSE (0.24) for eight clusters (Fig. 16e and f). As expected, classifications, and consequentltly inversion products, are improved if the number of clusters increases. For example, the retrieval of LAI with ten clusters, leads to std equal to 0.64, which is similar to *in situ* measurements (std equal to 0.6).

2) Deconvolution of real LiDAR waveforms: Retrieved maize height, LAI and field soil reflectance are shown in Fig. 17 for 3 cluster numbers and 3 window sizes using the waveform deconvolution. Their mean estimated values and their std are plotted in Fig. 18 against the field measurements.

Fig. 17 shows that two main classes are observed for different window sizes. The most present class (occurrence percentage $\approx 85\%$) is characterized by a height of 1.2 m and LAI of 2.9. The height and LAI values of the second estimated class are equal to 1.3 m and 2.6, respectively for 11% of windows. The rest of windows (percentage $\approx 4\%$) is distributed differently from a window size to another. The estimated soil reflectance is concentrated on two close values (0.3 and 0.4) independently of the cluster number and the window size.

The comparison between inversion results and field measurements shows that mean height and LAI values are accurately retrieved for different window sizes and cluster numbers with RMSE smaller than 0.06 m and 0.17, respectively. In particular, the height and LAI are best estimated with ten clusters, with RMSE equal to 0.03 m and 0.03, respectively. Their std is 0.12 m and 0.4, respectively, whereas the std

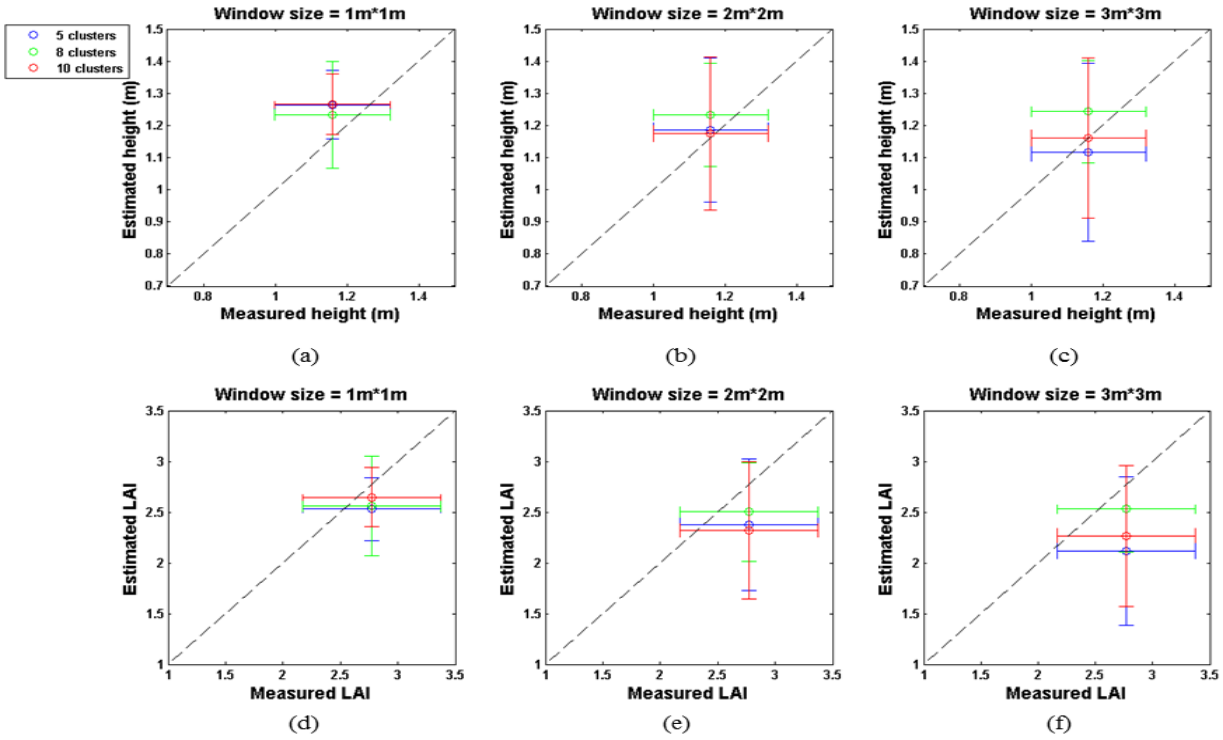


Fig. 16. Estimated maize height against *in situ* measurement collected on 23 June 2015, for 5 (blue color), 8 (green) and 10 (red) clusters, and 3 field window sizes: $1m^2$, $4m^2$, and $9m^2$, using the LUT convolution. Circles represent the mean estimated height. Vertical and horizontal error bars represent the standard deviation of the estimated height and field measurements, respectively.

of field measurements is 0.16 m and 0.6, respectively. The similarity between estimated and measured std (same order) proves that the designed approach can represent heterogeneity in the maize field.

The accuracy of the estimated results by the two compared methods: convolution of LUT and the deconvolution of the actual waveforms are comparable for maize height. However, the real waveform deconvolution method displayed more accurate estimation for LAI.

V. Discussion and Conclusion

This study presents a new inversion method of small footprint LiDAR full waveform for crop parameter retrieval. This method is based on a look up table of waveforms that are simulated with the 3-D DART model for specific crop biophysical properties and LiDAR configuration. The designed inversion model proved to have the potential to estimate the height and LAI of maize and wheat fields in the Auradé and Lamothe regions, in southwest France. It uses two approaches for comparing measured LiDAR waveforms and DART simulated LUT waveforms: (i) comparison of real waveforms with LUT waveforms that are convolved, and (ii) comparison of LUT waveforms with LiDAR measured waveforms that are deconvolved. The two methods lead to accurate assessments for the height and LAI in wheat field independently of the number of clusters used in the waveform classification procedure and the size of the analysis window representative of the local crop heterogeneity. Nevertheless, best results were obtained using the method that deconvolves the acquired LiDAR waveforms. In that case, the height is estimated with a RMSE as low as 0.01 m. Wheat LAI estimates were difficult to validate, since dry leaves were not considered when measuring the area of leaves in field and laboratory measurements. Wheat was in the "Ripening" growth stage, in which the kernels were mature, the stem development has been accomplished and leaves were totally dry. Then, wheat crops were almost ready for harvest. The small spatial heterogeneity in the wheat field explains the small standard deviation values that are obtained: 0.02 m and 0.1 for height and LAI, respectively.

For the maize field, the ranges of height (0.7 m to 1.3 m) and LAI (0.98 to 2.9) values are better detected using the method that convolves the LUT waveforms independently of the window size and the cluster number. Around 75% of the field area is associated to the higher height values and 25% is associated to low values. The estimated soil reflectance is concentrated on 0.3 and 0.4 in 75% of the maize field. Furthermore, the best estimated mean height and LAI are obtained with large cluster number with RMSE equal to 0.01 m and 0.24, respectively. The assessed height and LAI standard deviations are similar to those that were measured in the field (0.16 m and 0.6, respectively). This similarity supports the potential of LiDAR for detecting the maize plants heterogeneity in the field. Compared to the LUT convolution method, the deconvolution of the LiDAR waveforms lead to estimated parameter values with smaller ranges: 1.2 m, 1.3 m and 1.5 m for height, and 2.6, 2.9 and 3.7 for LAI.

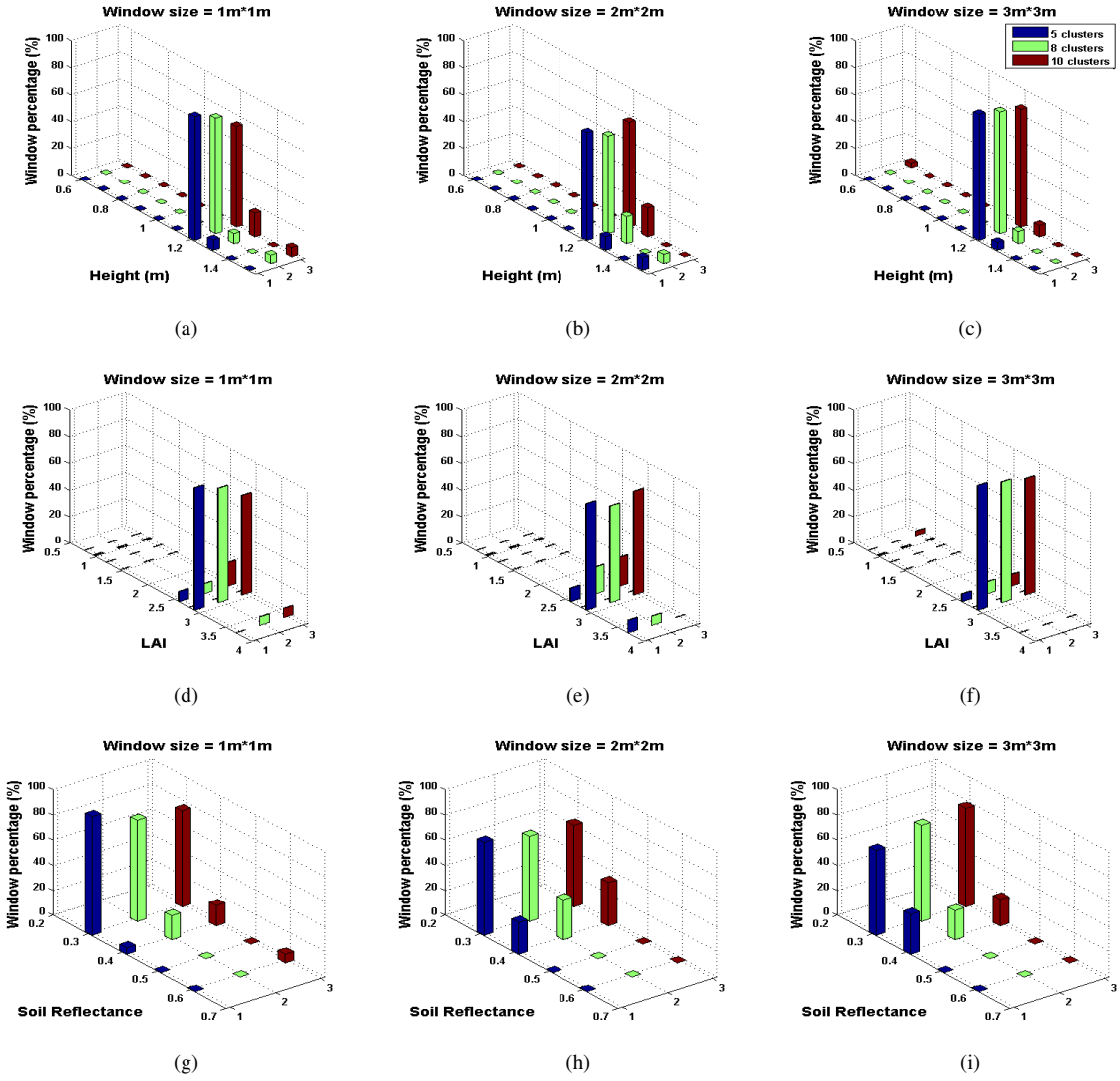


Fig. 17. Maize field histograms, using the waveform deconvolution: height (top), LAI (middle) and soil reflectance (bottom), for 3 field window sizes and 5 (blue), 8 (green) and 10 (red) clusters.

Indeed, more than 85% of windows are concentrated on one class characterized by 1.2 m and 2.9 for height and LAI, respectively. The rest of the window percentage is scattered between the other classes independently of the cluster numbers and the window size. Nonetheless, the mean maize height was derived with small RMSE of 0.03 m using ten clusters in the classification procedure. LAI was accurately estimated with RMSE around 0.04 using a large cluster number compared to 0.24 using the LUT convolution method.
Most estimates of soil reflectance values were around 0.3 and 0.4 values which is realistic. However, it is important to stress that varying the soil reflectance values in addition to the crop plant properties when generating the LUT would probably improve the estimate of the vegetation parameters. In addition, the good performance of such inversion model was likely due to the realistic modeling of the crops mock ups for different growth stages.

Furthermore, the accuracy of the obtained results, independently of the choice of the window size ($1m^2$, $4m^2$ and $9m^2$) representative of the local heterogeneity within fields illustrates the efficiency and the robustness of the waveform inversion model. Moreover, the fact to obtain good results tends to validate our assumption that the LiDAR sent pulse could be estimated as the response of the roads.

Better validation results could be obtained if more detailed and precise field measurements were available. Ideally, the availability of geolocalized field data would allow spatial comparisons of crop parameters that are estimated and *in situ* measurements. Furthermore, results would be also improved by combining LiDAR acquisition with hyperspectral images, for example for assessing crop biochemical properties such as the vegetation chlorophyll content and water content.

ACKNOWLEDGMENT

Data acquisition at FR-Lam is mainly funded by ICOS-France. The experimental site also benefited from the support and facilities of the Regional Spatial Observatory (OSR), CNRS (Centre National de la Recherche Scientifique), CNES

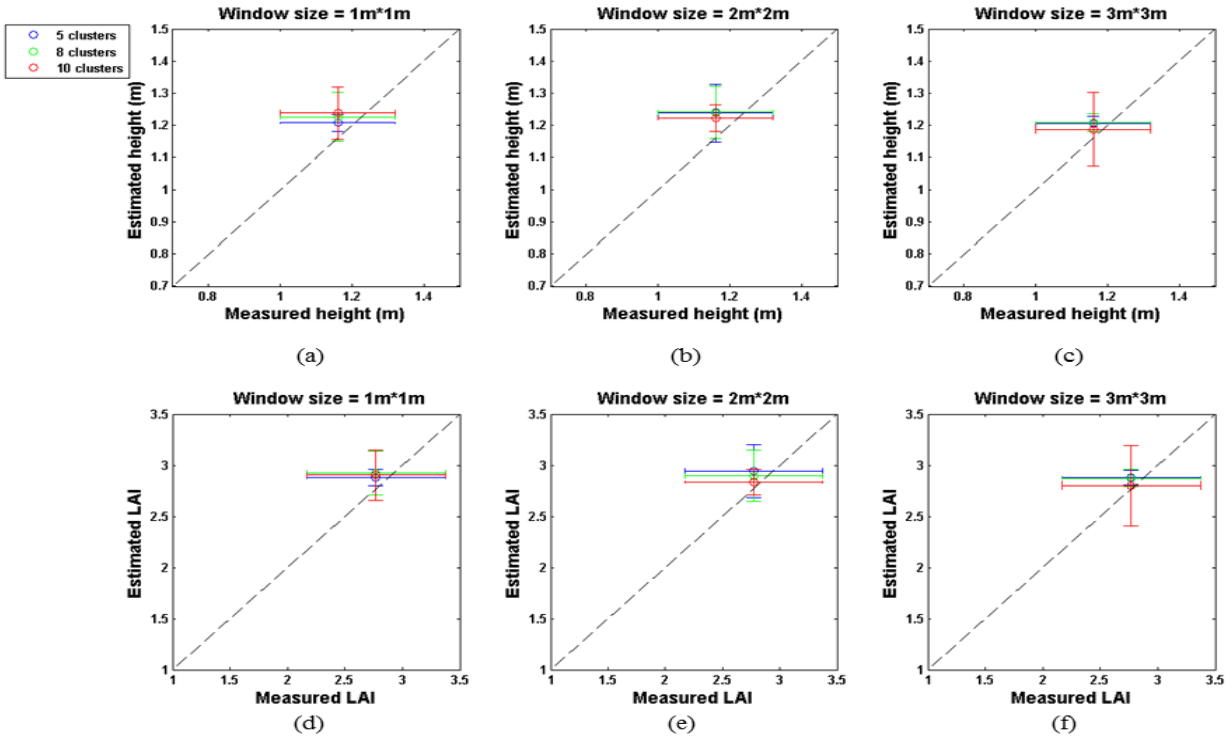


Fig. 18. Estimated maize height against *in situ* measurement, for 5 (blue color), 8 (green) and 10 (red) clusters, and 3 field window sizes: $1m^2$, $4m^2$, and $9m^2$, using the deconvolution of real waveforms. Circles represent the mean estimated height. Vertical and horizontal error bars represent the standard deviation of the estimated height and field measurements, respectively.

(Centre National d'Etudes Spatiales) and the University of Toulouse. We are grateful to Nicole Ferroni and Tiphaine Tallec for their involvement in vegetation measurements. We thank C. Marais-Sicre for providing the map of the vegetation classification on the southwest of France. This work was supported by the CNES in the frame of the TOSCA project Stem-Leaf.

REFERENCES

[1] H. Tang, M. Brolly, F. Zhao, A. H. Strahler, C. L. Schaaf, S. Ganguly, G. Zhang, and R. Dubayah, "Deriving and validating leaf area index (lai) at multiple spatial scales through lidar remote sensing: A case study in sierra national forest, ca," *Remote Sensing of Environment*, vol. 143, pp. 131–141, 2014.

[2] B. Combal, F. Baret, M. Weiss, A. Trubuil, D. Mace, A. Pragnere, R. Myneni, Y. Knyazikhin, and L. Wang, "Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem," *Remote sensing of environment*, vol. 84, no. 1, pp. 1–15, 2003.

[3] B. Koetz, F. Morsdorf, G. Sun, K. Ranson, K. Itten, and B. Allgower, "Inversion of a lidar waveform model for forest biophysical parameter estimation," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 1, pp. 49–53, 2006.

[4] S. Luo, J. M. Chen, C. Wang, X. Xi, H. Zeng, D. Peng, and D. Li, "Effects of lidar point density, sampling size and height threshold on estimation accuracy of crop biophysical parameters," *Optics Express*, vol. 24, no. 11, pp. 11 578–11 593, 2016.

[5] J. M. Chen, A. Govind, O. Sonnentag, Y. Zhang, A. Barr, and B. Amiro, "Leaf area index measurements at fluxnet-canada forest sites," *Agricultural and Forest Meteorology*, vol. 140, no. 1, pp. 257–268, 2006.

[6] S. T. Gower and J. M. Norman, "Rapid estimation of leaf area index in conifer and broad-leaf plantations," *Ecology*, vol. 72, no. 5, pp. 1896–1900, 1991.

[7] D. Ellsworth and P. Reich, "Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest," *Oecologia*, vol. 96, no. 2, pp. 169–178, 1993.

[8] N. Zhang, M. Wang, and N. Wang, "Precision agriculturea worldwide overview," *Computers and electronics in agriculture*, vol. 36, no. 2, pp. 113–132, 2002.

[9] I. Jonckheere, S. Fleck, K. Nackaerts, B. Muys, P. Coppin, M. Weiss, and F. Baret, "Review of methods for in situ leaf area index determination: Part i. theories, sensors and hemispherical photography," *Agricultural and forest meteorology*, vol. 121, no. 1, pp. 19–35, 2004.

[10] C. Atzberger, "Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs," *Remote Sensing*, vol. 5, no. 2, pp. 949–981, 2013.

[11] J. T. Morisette, F. Baret, J. L. Privette, R. B. Myneni, J. E. Nickeson, S. Garrigues, N. V. Shabanov, M. Weiss, R. A. Fernandes, S. G. Leblanc *et al.*, "Validation of global moderate-resolution lai products: A framework proposed within the ceos land product validation subgroup," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, p. 1804, 2006.

[12] W. B. Cohen, T. K. Maiersperger, S. T. Gower, and D. P. Turner, "An improved strategy for regression of biophysical variables and landsat etm+ data," *Remote Sensing of Environment*, vol. 84, no. 4, pp. 561–571, 2003.

[13] R. Colombo, D. Bellingeri, D. Fasolini, and C. M. Marino, "Retrieval of leaf area index in different vegetation types using high resolution satellite data," *Remote Sensing of Environment*, vol. 86, no. 1, pp. 120–131, 2003.

[14] C. Atzberger and K. Richter, "Spatially constrained inversion of radiative transfer models for improved lai mapping from future sentinel-2 imagery," *Remote Sensing of Environment*, vol. 120, pp. 208–218, 2012.

[15] V. C. Laurent, W. Verhoef, A. Damm, M. E. Schaepman, and J. G. Clevers, "A bayesian object-based approach for estimating vegetation biophysical and biochemical variables from apex at-sensor radiance data," *Remote Sensing of Environment*, vol. 139, pp. 6–17, 2013.

[16] S. Chaabouni and A. Kallel, "Inversion of vegetation canopy reflectance based on variationnel multiscale approach," in *Advanced Technologies for Signal and Image Processing (ATSIP), 2016 2nd International Conference*. IEEE, 2016, pp. 494–498.

[17] M. Claverie, V. Demarez, B. Duchemin, O. Hagolle, D. Ducrot, C. Marais-Sicre, J.-F. Dejoux, M. Huc, P. Keravec, P. Béziat *et al.*, "Maize and sunflower biomass estimation in southwest france using high spatial and temporal resolution remote sensing data," *Remote Sensing of Environment*, vol. 124, pp. 844–857, 2012.

[18] A. A. Abuelgasim, R. A. Fernandes, and S. G. Leblanc, "Evaluation of national and global lai products derived from optical remote sensing instruments over canada," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1872–1884, 2006.

[19] J. J. Richardson, L. M. Moskal, and S.-H. Kim, "Modeling approaches to estimate effective leaf area index from aerial discrete-return lidar," *Agricultural and Forest Meteorology*, vol. 149, no. 6, pp. 1152–1160, 2009.

[20] L. She-Zhou, W. Cheng, X. Xiao-Huan, N. Sheng, X. Shao-Bo, and W. Yi-Ping, "Forest leaf area index estimation using combined ice-sat/glas and optical remote sensing image," *JOURNAL OF INFRARED AND MILLIMETER WAVES*, vol. 34, no. 2, pp. 243–249, 2015.

[21] J. B. Drake, R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. Prince, "Estimation of tropical forest structural characteristics using large-footprint lidar," *Remote Sensing of Environment*, vol. 79, no. 2, pp. 305–319, 2002.

[22] H. Tang, R. Dubayah, A. Swatantran, M. Hofton, S. Sheldon, D. B. Clark, and B. Blair, "Retrieval of vertical lai profiles over tropical rain forests using waveform lidar at la selva, costa rica," *Remote Sensing of Environment*, vol. 124, pp. 242–250, 2012.

[23] T. Allouis, S. Durrieu, P. Chazette, J.-S. Bailly, J. Cuesta, C. Véga, P. Flamant, and P. Couteron, "Potential of an ultraviolet, medium-footprint lidar prototype for retrieving forest structure," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, pp. S92–S102, 2011.

[24] I. Bye, P. North, S. Los, N. Kljun, J. Rosette, C. Hopkinson, L. Chasmer, and C. Mahoney, "Estimating forest canopy parameters from satellite waveform lidar by inversion of the flight three-dimensional radiative transfer model," *Remote Sensing of Environment*, vol. 188, pp. 177–189, 2017.

[25] S. Ben Hmida, A. Kallel, J.-P. Gastellu-Etchegorry, and A. B. Hamida, "Variational multiscale approach to lai profile inversion based on lidar full waveform measurements," in *Advanced Technologies for Signal and Image Processing (ATSIP), 2016 2nd International Conference on*. IEEE, 2016, pp. 511–516.

[26] H. Ma, J. Song, J. Wang, Z. Xiao, and Z. Fu, "Improvement of spatially continuous forest lai retrieval by integration of discrete airborne lidar and remote sensing multi-angle optical data," *Agricultural and Forest Meteorology*, vol. 189, pp. 60–70, 2014.

[27] N. Glenn, L. Spaete, T. Sankey, D. Derryberry, S. Hardegree, and J. Mitchell, "Errors in lidar-derived shrub height and crown area on sloped terrain," *Journal of Arid Environments*, vol. 75, no. 4, pp. 377–382, 2011.

[28] J. Cuesta, P. Chazette, T. Allouis, P. H. Flamant, S. Durrieu, J. Sanak, P. Genau, D. Guyon, D. Loustau, and C. Flamant, "Observing the forest canopy with a new ultra-violet compact airborne lidar," *Sensors*, vol. 10, no. 8, pp. 7386–7403, 2010.

[29] X. Shang and P. Chazette, "Interest of a full-waveform flown uv lidar to derive forest vertical structures and aboveground carbon," *Forests*, vol. 5, no. 6, pp. 1454–1480, 2014.

[30] G. Vincent, D. Sabatier, L. Blanc, J. Chave, E. Weissenbacher, R. Pélissier, E. Fonty, J.-F. Molino, and P. Couteron, "Accuracy of small footprint airborne lidar in its predictions of tropical moist forest stand structure," *Remote Sensing of Environment*, vol. 125, pp. 23–33, 2012.

[31] G. Frazer, S. Magnussen, M. Wulder, and K. Niemann, "Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of lidar-derived estimates of forest stand biomass," *Remote Sensing of Environment*, vol. 115, no. 2, pp. 636–649, 2011.

[32] C. S. Neigh, R. F. Nelson, K. J. Ranson, H. A. Margolis, P. M. Montesano, G. Sun, V. Kharuk, E. Næsset, M. A. Wulder, and H.-E. Andersen, "Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne lidar," *Remote Sensing of Environment*, vol. 137, pp. 274–287, 2013.

[33] L. Zhang and T. E. Grift, "A lidar-based crop height measurement system for miscanthus giganteus," *Computers and electronics in Agriculture*, vol. 85, pp. 70–76, 2012.

[34] N. Tilly, D. Hoffmeister, Q. Cao, S. Huang, V. Lenz-Wiedemann, Y. Miao, and G. Bareth, "Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice," *Journal of Applied Remote Sensing*, vol. 8, no. 1, pp. 083 671–083 671, 2014.

[35] J. U. Eitel, T. S. Magney, L. A. Vierling, T. T. Brown, and D. R. Huggins, "Lidar based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status," *Field Crops Research*, vol. 159, pp. 21–32, 2014.

[36] I. Davenport, R. Bradbury, G. Anderson, G. Hayman, J. Krebs, D. Mason, J. Wilson, and N. Veck, "Improving bird population models using airborne remote sensing," *International Journal of Remote Sensing*, vol. 21, no. 13-14, pp. 2705–2717, 2000.

[37] W. Li, Z. Niu, N. Huang, C. Wang, S. Gao, and C. Wu, "Airborne lidar technique for estimating biomass components of maize: A case study in zhangye city, northwest china," *Ecological Indicators*, vol. 57, pp. 486–496, 2015.

[38] S. Hancock, M. Disney, J.-P. Muller, P. Lewis, and M. Foster, "A threshold insensitive method for locating the forest canopy top with waveform lidar," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3286–3297, 2011.

[39] J.-P. Gastellu-Etchegorry, V. Demarez, V. Pinel, and F. Zagolski, "Modeling radiative transfer in heterogeneous 3-d vegetation canopies," *Remote sensing of environment*, vol. 58, no. 2, pp. 131–156, 1996.

[40] J.-P. Gastellu-Etchegorry, T. Yin, N. Lauret, T. Cajgfinger, T. Gregoire, E. Grau, J.-B. Feret, M. Lopes, J. Guilleux, G. Dedieu *et al.*, "Discrete anisotropic radiative transfer (dart 5) for modeling airborne and satellite spectroradiometer and lidar acquisitions of natural and urban landscapes," *Remote Sensing*, vol. 7, no. 2, p. 1667, 2015.

[41] P. North, J. Rosette, J. Suárez, and S. Los, "A monte carlo radiative transfer model of satellite waveform lidar," *International Journal of Remote Sensing*, vol. 31, no. 5, pp. 1343–1358, 2010.

[42] F. D. Schneider, R. Leiterer, F. Morsdorf, J.-P. Gastellu-Etchegorry, N. Lauret, N. Pfeifer, and M. E. Schaepman, "Simulating imaging spectrometer data: 3d forest modeling based on lidar and in situ data," *Remote Sensing of Environment*, vol. 152, pp. 235–250, 2014.

[43] G. Kenny and P. Harrison, "Thermal and moisture limits of grain maize in europe: model testing and sensitivity to climate change," *Climate research*, vol. 2, no. 2, pp. 113–129, 1992.

[44] R. Webster, "Is soil variation random?" *Geoderma*, vol. 97, no. 3, pp. 149–163, 2000.

[45] A. B. McBratney, M. M. Santos, and B. Minasny, "On digital soil mapping," *Geoderma*, vol. 117, no. 1, pp. 3–52, 2003.

[46] M. R. Gupta, Y. Chen *et al.*, "Theory and use of the em algorithm," *Foundations and Trends® in Signal Processing*, vol. 4, no. 3, pp. 223–296, 2011.

[47] E. P. Baltsavias, "Airborne laser scanning: basic relations and formulas," *ISPRS Journal of photogrammetry and remote sensing*, vol. 54, no. 2, pp. 199–214, 1999.

[48] S. Hancock, "Understanding the measurement of forests with waveform lidar," Ph.D. dissertation, UCL (University College London), 2010.

[49] S. Hancock, J. Armston, Z. Li, R. Gaulton, P. Lewis, M. Disney, F. M. Danson, A. Strahler, C. Schaaf, K. Anderson *et al.*, "Waveform lidar over vegetation: An evaluation of inversion methods for estimating return energy," *Remote Sensing of Environment*, vol. 164, pp. 208–224, 2015.

[50] Q. Chen, "Retrieving vegetation height of forests and woodlands over mountainous areas in the pacific coast region using satellite laser altimetry," *Remote Sensing of Environment*, vol. 114, no. 7, pp. 1610–1627, 2010.

[51] M. Disney, P. Lewis, and P. North, "Monte carlo ray tracing in optical canopy reflectance modelling," *Remote Sensing Reviews*, vol. 18, no. 2-4, pp. 163–196, 2000.

[52] J. Gastellu-Etchegorry, E. Martin, and F. Gascon, "Dart: a 3d model for simulating satellite images and studying surface radiation budget," *International journal of remote sensing*, vol. 25, no. 1, pp. 73–96, 2004.

[53] J.-P. Gastellu-Etchegorry, "3d modeling of satellite spectral images, radiation budget and energy budget of urban landscapes," *Meteorology and atmospheric physics*, vol. 102, no. 3-4, p. 187, 2008.

[54] S. Jacquemoud and F. Baret, "Prospect: A model of leaf optical properties spectra," *Remote sensing of environment*, vol. 34, no. 2, pp. 75–91, 1990.

[55] B. Hosgood, S. Jacquemoud, G. Andreoli, J. Verdebout, G. Pedrini, and G. Schmuck, "Leaf optical properties experiment 93," *Joint Res. Center, Eur. Comm., Inst. Remote Sensing Applications, Ispra, Italy, Tech. Rep. EUR*, vol. 16, p. 095, 1995.

[56] U. Meier, "Stades phénologiques des mono-et dicotylédones cultivées," *Centre Fédéral de Recherches Biologiques pour l Agriculture et les Forêts*, 2001.

[57] M. A. Hofton, J. B. Minster, and J. B. Blair, "Decomposition of laser altimeter waveforms," *IEEE Transactions on geoscience and remote sensing*, vol. 38, no. 4, pp. 1989–1996, 2000.

[58] A. Roncat, C. Briese, J. Jansa, and N. Pfeifer, "Radiometrically calibrated features of full-waveform lidar point clouds based on statistical moments," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 2, pp. 549–553, 2014.

[59] R. C. Gonzalez and R. E. Woods, "Image processing," *Digital image processing*, vol. 2, 2007.