

# Validation of the DART Model for Airborne Laser Scanner Simulations on Complex Forest Environments

Florian de Boissieu<sup>ID</sup>, Florence Heuschmidt<sup>ID</sup>, Nicolas Lauret<sup>ID</sup>, Dav M. Ebengo<sup>ID</sup>, Grégoire Vincent<sup>ID</sup>, Jean-Baptiste Féret<sup>ID</sup>, Tiangang Yin<sup>ID</sup>, Jean-Philippe Gastellu-Etchegorry<sup>ID</sup>, *Member, IEEE*, Josiane Costeraste, Marie-José Lefèvre-Fonollosa<sup>ID</sup>, and Sylvie Durrieu<sup>ID</sup>

**Abstract**—With the recent progresses in lidar technology for Earth remote sensing, the development of a reliable lidar simulator is becoming central in order to define specifications for new sensors, perform intercomparisons, train machine learning algorithms, and help transferring information from one scale to another. The discrete anisotropic radiative transfer (DART) model includes such a lidar simulator. Although already tested on several virtual scenes, the DART outputs still need to be rigorously evaluated against actual sensor acquisitions, especially on real complex scenes of various forest types, such as dense tropical forests. That is the purpose of the present study. A real airborne laser scanner (ALS) with full-waveform capacity was first radiometrically calibrated on targets of measured reflectance. The properties of the ALS system were then introduced in the DART model, along with a 3-D virtual scene built from terrestrial laser scans and spectroscopic measurements acquired on a forest plot near the calibration site. Finally, an ALS acquisition was simulated and the shape and magnitude of the waveforms were compared with real acquisitions. The comparison

between measured and simulated data was performed at different scales by aggregating waveform samples into a 3-D grid with a vertical resolution of 1 m and a horizontal resolution ranging from 2 to 80 m. Results showed a high similarity between simulated and measured waveforms at all scales with  $R^2 > 0.9$  and  $\text{NRMSE} < 10\%$ . These promising results open up numerous perspectives for improved spaceborne and airborne lidar data processing and for the development of new systems.

**Index Terms**—3-D, discrete anisotropic radiative transfer (DART), forest, lidar, point cloud, radiative transfer model, simulation, waveform.

## I. INTRODUCTION

LIDAR data has proved its suitability for the study of forest ecosystems as it provides a unique insight into the 3-D distribution of the vegetation structure from the top of the canopy to the ground. This information on vegetation structure is highly relevant to retrieve key forest structural and biophysical information, e.g., basal area, wood volume, biomass, plant and leaf area density profiles, plant and leaf area index, which are essential for sustainable forest management and for the modeling of ecosystem functioning [1], [2], [3], [4], [5]. However, processing and analyzing lidar signals (waveforms or point clouds) is complex owing to their 3-D characteristics. Moreover, ground truth measurements in a 3-D context to calibrate and validate models is a challenging task, since the instruments and protocols used for this purpose are either destructive, time consuming or both [6], [7]. If such a task is possible at small scale, it becomes unrealistic at large scale.

In this context, radiative transfer modeling plays an important role in improving lidar signal interpretation and forest parameter assessment, especially in structurally complex forests. 3-D radiative transfer models can take advantage of optically and structurally realistic 3-D mockups, and ray-tracing algorithms to simulate laser propagation (reflection, transmission, and absorption) through the elements of the mockup (leaves, branches, trunks), and to produce lidar waveforms and point clouds [8], [9], [10]. Resulting simulations have a large potential for multiple applications, including the adjustment of specifications for spaceborne lidar sensors [8], [11], [12], and the conception, intercomparison and training of algorithms dedicated to lidar data processing [13]. However, most of these applications imply

Manuscript received 20 November 2022; revised 3 March 2023 and 17 May 2023; accepted 10 July 2023. Date of publication 6 September 2023; date of current version 18 September 2023. This work was carried out in the frameworks of the LEAF-EXPEVAL and the HyperTropik/HyperBIO projects, supported by TOSCA Continental Surface Program of the Centre National d'Études Spatiales (CNES). This work was also supported by the Agence Nationale de la Recherche (BioCop project) under Grant ANR-17-CE32-0001. (*Corresponding author: Florian de Boissieu.*)

Florian de Boissieu, Jean-Baptiste Féret, and Sylvie Durrieu are with the INRAE, AgroParisTech, CIRAD, CNRS, University of Montpellier, TETIS, 34000 Montpellier, France (e-mail: florian.deboissieu@inrae.fr; jean-baptiste.feret@teledetection.fr; sylvie.durrieu@inrae.fr).

Florence Heuschmidt was with the CIRAD, CNRS, INRAE, IRD, AMAP, University of Montpellier, 34000 Montpellier, France. She is now an independent. (e-mail: florence.hj10@gmail.com).

Dav M. Ebengo was with the INRAE, AgroParisTech, CIRAD, CNRS, University of Montpellier, TETIS 34000, Montpellier, France. He is now with Université Clermont Auvergne, CNRS, MSH, F-63000 Clermont-Ferrand, France (e-mail: dav.ebengo\_mwampongo@uca.fr).

Grégoire Vincent is with the CIRAD, CNRS, INRAE, IRD, AMAP, University of Montpellier, 34000 Montpellier, France (e-mail: gregoire.vincent@ird.fr).

Tiangang Yin is with the Department of Land Surveying and Geo-Infomatics, The Hong Kong Polytechnic University, Hong Kong (e-mail: tiangang.yin.85@gmail.com).

Nicolas Lauret and Jean-Philippe Gastellu-Etchegorry are with the CNES, CNRS, UPS, IRD, INRAE, CESBIO, University of Toulouse, 31000 Toulouse, France (e-mail: nicolas.lauret@univ-tlse3.fr; jean-philippe.gastellu@iut-tlse3.fr).

Josiane Costeraste and Marie-José Lefèvre-Fonollosa are with the Earth Observation Satellite Systems, CNES Toulouse, 31000 Toulouse, France (e-mail: josiane.costeraste@cnes.fr; marie.lefevre-fonollosa@outlook.com).

Digital Object Identifier 10.1109/JSTARS.2023.3302030

the capacity to produce representations and signals that are consistent with real acquisitions on complex forest scenes, which requires accurate simulation of the radiative transfer within such complex scenes.

The discrete anisotropic radiative transfer model (DART) is a radiative transfer model originally developed to simulate satellite images, with particular interest to support space agencies in the definition of specifications for future satellite sensors. DART simulates the travel of light from a source (e.g., sun) to a sensor when passing through a virtual 3-D mockup of the scene (meshes or turbid voxels) characterized with optical and thermal properties. The initial DART model has been extended to single pulse [14] and later to multipulse lidar waveform simulations based on a quasi-Monte Carlo ray tracing approach [15]. More recently, the DART lidar model has been further accelerated using the Lux-core ray tracing library [9].

Among the evaluations of the DART model [16], [17], [18] and its intercomparison with other radiative transfer models [19], two studies drew our attention for this work. In [8], the DART lidar model was evaluated against the waveforms of a real acquisition with the laser vegetation imaging sensor at a height above ground level (AGL) of 10 km over a mixed forest (i.e., a footprint diameter of  $\sim 20$  m). The deciduous and coniferous crowns of the scene were simplified to ellipsoidal and conical turbid volumes with dimensions derived from allometric models, and uniform leaf area density and optical properties across individuals of the same kind. The DART lidar model was also evaluated against an ALS acquisition on a Maritime pine stand with different representations of the virtual scene (lollipop trees, turbid voxels, full mesh virtual trees) [20]. However, for that study the ALS waveforms were aggregated at the scale of the plot (diameter of 30 m) and the intensity was normalized before the comparison was made. Both studies [8], [20] were geared toward large footprint lidar waveforms, i.e., satellite lidar, and they compared simulated to real waveforms after normalizing the magnitude of the waveforms, limiting the evaluation to a structural comparison, i.e., the shape of the waveforms.

This article focuses on the capacities of the DART lidar model to reproduce both structural and radiometric information embedded in full-waveform ALS acquisitions over structurally complex forests.

In order to achieve this goal, a two-step experiment was set up. The first step aimed at evaluating the calibration factor of a full-waveform ALS sensor, which links the power received at the sensor aperture to the digital values recorded by the sensor. The calibration factor was estimated using ALS data and imaging spectroscopy acquired simultaneously over flat targets with contrasted spectral signatures. The second step aimed at comparing ALS waveforms acquired with the same sensor over a tropical forest plot to simulated waveforms obtained from DART simulations on the 3-D mockup of the forest plot. Section II describes the experimental sites and setup as well as the ALS sensor and data acquired for the study. Section III details the methods implemented for the calibration of the ALS sensor, its integration into DART, and the evaluation of the resulting simulations. The results are presented in Section IV and discussed

TABLE I  
SPECIFICATIONS OF LiDAR SENSOR RIEGL LMS-Q780

Symbol	Quantity	Value
$\lambda$	laser wavelength	1064 nm
$d\theta$	laser beam divergence at the $1/e^2$ points.	$\leq 0.25$ mrad
$P_{average}$	maximum average output power of laser	0.4 W
$PRR$	pulse repetition rate	400 kHz
$FWHM$	emitted pulse duration at half maximum, measured from data.	4.8 ns
$f_{ND}$	sampling frequency	1 GHz
$D$	size of the ALS beam aperture window used as the aperture diameter of the sensor in simulation.	75 mm

in Section V drawing perspectives for future work. Finally, Section VI concludes this article.

## II. MATERIAL

The experiment was conducted at the Paracou experimental station, French Guiana, a well-studied site in which several permanent rainforest plots have been regularly monitored since 1984 [21]. In September 2016, an experimental campaign was set up to simultaneously acquire airborne data, including full-waveform ALS data, imaging spectroscopy and very high resolution RGB orthophotos (ground sampling distance of 10 cm), as well as various field measurements with a terrestrial laser scanner (TLS) and a spectrometer. Two areas were specifically targeted for the current study. The first one, hereinafter referred to as “calibration plot,” corresponds to a drop zone, i.e., a flat clearing, in which reference targets were laid on the ground and used for the radiometric calibration of the ALS sensor. The second study area, hereinafter referred to as “forest plot” is included in one of the monitored forest plots of Paracou and was used for the comparison between measured and simulated ALS waveforms. TLS acquired over the forest plot was used to build a 3-D mockup for radiative transfer simulations. The following sections detail the specifications of the ALS sensor under investigation, as well as the data acquired and used for the sensor calibration and for the simulations with DART on both study areas.

### A. ALS Sensor Specifications

Full-waveform ALS data were acquired with a RIEGL LMS-Q780: the high sampling frequency of the backscattered signal provides the complete temporal distribution of the returned light energy. Sensor specifications are listed in Table I. These specifications were extracted from the datasheet supplied by the manufacturer [22], except for the emitted pulse duration at half maximum (FWHM) and the frequency of the digitization which were retrieved from the lidar data, i.e., RIEGL SDF files. The FWHM was estimated assuming a Gaussian shaped pulse after verification of this assumption on the pulse records. The assessment of a Gaussian shaped pulse can be found in Appendix C. As the diameter of aperture of the sensor (or its field of view) was not available in the specifications, it was assumed

to be equal to the size the beam aperture window given in the datasheet.

Additional properties can be derived from these specifications. The transmitted energy at full power  $E_T$  can be derived from the average power using (1), which gives a theoretical value of 1  $\mu\text{J}$  per pulse for that sensor

$$E_T = \frac{P_{\text{average}}}{\text{PRR}}. \quad (1)$$

A 1 GHz sampling frequency of the lidar signal corresponds to a sampling period of 1 ns, which represents a distance of 30 cm along the beam direction at light speed. Considering the path forward and backward of the signal, the expected ranging resolution for the LMS-Q780 is 15 cm.

Considering the upper bound of the laser beam divergence of this sensor, the laser footprint diameter at the ground level would be 22.5 cm for a flying altitude of 900 m AGL.

### B. ALS Data

The ALS data was available in RIEGL SDF files, containing raw full waveforms, and in LAS files containing the point clouds extracted from the SDF files with RIEGL's software suite. The LAS files were delivered with point data record format 1 including Extra Bytes [23], which gives for each echo, among other variables, the ALS shot time-stamp as well as the  $x$ ,  $y$ , and  $z$  coordinates, the intensity and the width of the echo [24].

The ALS shot pattern (sensor position and direction of each emitted pulse) was obtained from the sensor trajectory combined with the ALS point cloud. The sensor location was linearly interpolated from the trajectory data at the ALS shot time-stamps. The shot direction was computed from the line connecting the sensor location to the echoes of the shot. The point clouds were classified and a 0.5 m resolution digital terrain model (DTM) was computed from the points classified as ground.

For our experiments, the waveform samples needed to be georeferenced. The usual pipeline to convert the raw waveforms to a georeferenced point cloud is first to extract the echoes, second to compute the echoes coordinates in the sensor coordinate system, which are then successively converted to the geocentric coordinate system and projected in the local coordinate system taking into account the geoid. The same could be done directly with the waveform samples (instead of the echoes). However, the registration of the returned waveform blocks to the correct emitted pulse can be complex for multiple pulses in the air [25], which was the case for our dataset. Moreover, these steps are very sensitive to an imprecision on the location and orientation of the different instruments involved (GPS, IMU, and sensor). And as we cannot know in advance which waveform is crossing the scene, the whole SDF file would have to be processed before extracting the waveforms belonging to the area of interest. Therefore, we chose another strategy considering that the point cloud had been produced with the same pipeline already implemented in RIEGL's software suite.

The chosen alternative, hereinafter referred to as "SDF Based WaveForms" (SB-WF), consisted in extracting from the RIEGL SDF files the waveform sample blocks based on the time stamps of the echoes belonging to the area of interest. As the echo

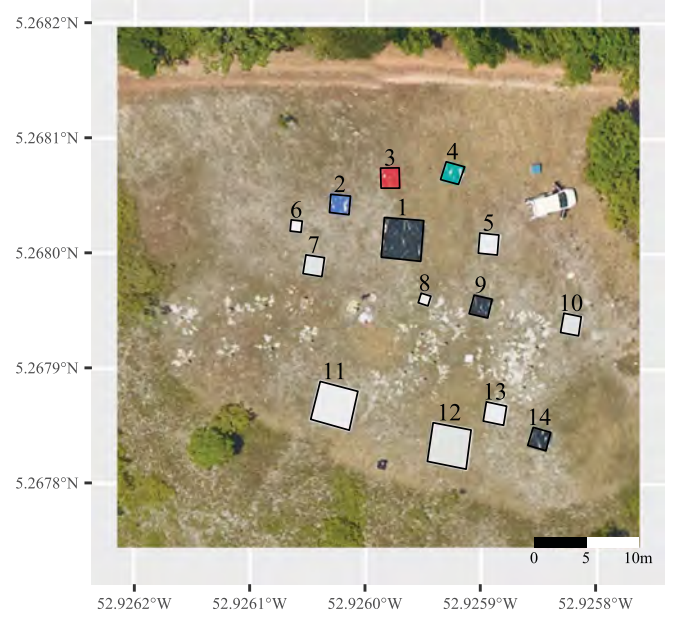


Fig. 1. Spatial configuration of the targets on the calibration plot (obtained from orthophotography).

time stamp was corresponding to the time of emission, we were able to discriminate the waveform blocks belonging to multiple pulses in the air according to the light travel time between the pulse emission and the echo. The waveform samples were positioned in space using the line between the sensor location at the time of shot and the echoes location. A background noise was noticed on the raw waveforms, probably resulting from solar and atmosphere radiations and the detector noise [26], [27], [28]. As the relevant information to model the background noise was not available, the raw waveforms were thresholded in order to reduce its impact on the analysis. Details on SDF-based waveform (SB-WF) reconstruction as well as waveform examples are given in Appendix D.

### C. Experimental Data on the Calibration Plot

In order to determine the calibration factor of the ALS sensor, 14 square tarps of various size (1 to 4 m width) with different reflective properties at the laser wavelength were laid out on the ground of the calibration plot. The targets were manually delineated by photointerpretation of the RGB orthophoto in order to georeference them. The reflectance at the laser wavelength of each target was extracted from imaging spectroscopy data acquired simultaneously to ALS data. Imaging spectroscopy was operated at a flying altitude of 450 m AGL leading to a spatial resolution of 0.8 m. Targets 6 and 8 (1 m width) were excluded from the study as the spatial resolution of imaging spectroscopy could not guarantee pure pixels. The central pixel was selected from each remaining target in order to avoid mixed pixels at the border of the targets. The hyperspectral band used for that operation had a central wavelength and a bandwidth of 1063.74 and 6.25 nm, respectively. Fig. 1 shows the display of the targets



TABLE II  
DIFFERENT ALS CONFIGURATIONS USED FOR CALIBRATION

Flying AGL (m)	Laser power (%)	Number of flightlines
450	6	2
450	12	2
650	12	2
900	12	17

on the RGB orthophoto. The size, color and reflectance values of the targets are available in Table III of Appendix B.

Automatic gain control (AGC) is implemented in some lidar sensors [29] in order to magnify the signal and get returns even from low-reflectance targets. However, these specifications were not provided by the manufacturer for the LMS-Q780. If such a feature was implemented in this lidar sensor, it would be expected to observe variations of the calibration factor value according to the sensor-to-target distance, i.e., the flying altitude and to the laser power. In order to investigate AGC and to account for it if needed, the airborne acquisitions were repeated over the calibration plot with different flying altitudes (450, 600, and 900 m) and different laser powers (6% and 12% of full power defined in Table I). Table II summarizes the properties of the 23 datasets acquired with different sensor settings and flight conditions that were used for the calibration, more details can be found in Appendix B.

#### D. Experimental Data on the Forest Plot

The forest plot corresponds to an  $80 \times 80$  m square subset of Plot P9 of the INRAE plot network monitored in Paracou. This undisturbed control plot was used to compare ALS data simulated with DART to real ALS data for a complex natural forest. The acquisitions covering the forest plot were performed with a flying altitude of 900 m and a laser power set to 12% of the full power. We selected one ALS flightline with minimum scan angle (lidar shots close to nadir) over the forest plot in order to perform the comparison. Both the structure and the optical properties of the forest elements on the site were measured to build a 3-D mockup of the forest. TLS data of the forest plot were acquired with a RIEGL VZ400 system to derive finely resolved forest structure. Scan positions were set in a regular 10 m grid. Capturing a complete sample of the scene at each location required two scans (upright and tilted by  $90^\circ$ ), owing to the  $100^\circ$  vertical field of view of the scanner. Therefore, 162 scans were collected from 81 locations. The angular resolution between sequentially fired pulses was set to  $0.04^\circ$ , resulting in approximately 22.4 million emitted pulses per scan (i.e., 3.6 billion for the entire plot). Up to five targets could be resolved per pulse, with a nominal ranging accuracy of 5 mm. The laser is characterized by a beam divergence of less than 0.3 mrad, and the diameter of the beam at emission is 7 mm (corresponding to a beam diameter of 22 mm at a range of 50 m). The pulse repetition rate was set to 300 kHz. All scans were co-registered using RiScanPro software. The position of the corners of the plots was registered using a total station and these corners were used

to transform the scans of the plot to the UTM 22N coordinate system used for the ALS data analysis.

Leaf optical properties (reflectance and transmittance), as well as bark reflectance were measured with an ASD FieldSpec-3 spectroradiometer equipped with a LiCor 1800-12 integrating sphere. Leaves were collected by climbers at several heights in the crown and for several species.

### III. METHOD

The waveform signal of a lidar system over a lambertian surface larger than the laser footprint can be described by the following [30], [31]:

$$E_{DN} = K \cdot \frac{\rho \cdot \cos(\alpha)}{4R^2} \cdot E_T \quad (2)$$

with  $E_T$  the transmitted energy,  $E_{DN}$  the received energy in Digital Number unit (DN),  $R$  the sensor-to-target distance,  $\rho$  the reflectance,  $\alpha$  the angle of incidence of the laser, and  $K$  the calibration factor. The calibration factor is dependent on the atmospheric transmission factor  $\eta_{atm}$ , which is considered to be constant during the ALS campaign (approx. 2 h), and on the detection and sampling system factor  $\eta_{sys}$  [31].

The validation of the DART lidar simulations was performed in two steps. The first step aimed at determining  $K$  on the calibration plot to link the intensity of the ALS signal simulated with DART to the intensity of the measured ALS signal, thus providing a means to convert the simulated signal into digital counts. The second step focused on the comparison between ALS acquisition and ALS simulation on the forest plot to validate the capacity to simulate realistic signals with DART in complex forests.

#### A. Calibration

Equation (2) was converted into (3) in order to determine  $K$  as the solution of a linear regression problem between the normalized received energy  $E_N$  and the reflectance  $\rho$ :

$$E_N = \frac{E_{DN} \cdot 4R^2}{E_T \cdot \cos(\alpha)} = K \cdot \rho. \quad (3)$$

For an opaque target of reflectance  $\rho$  and for a footprint fully contained within the target limits,  $E_{DN}$  is obtained from the following:

$$\sigma = \frac{\text{FWHM}}{2\sqrt{2\ln(2)}} \quad (4)$$

$$E_{DN} = I_{DN} \cdot \sigma \cdot \sqrt{2\pi} \quad (5)$$

with  $I_{DN}$  being the peak value of a Gaussian shaped return characterized by a full width at half maximum FWHM.  $I_{DN}$  and FWHM, correspond to the intensity and the pulsewidth values delivered by RIEGL software during point cloud extraction [24].

The  $E_{DN}$  values used in the linear regression were computed for each target and each flight line as the average values of all single lidar returns for which the footprint was fully included within the target limits. A linear regression was computed for each flightline in order to study the impact of the flying altitude and of the transmitted power on  $K$ , if any, as well as to evaluate

the variability of  $K$  between flightlines. Indeed, if any AGC was operating during the ALS acquisitions, it would be expected to compensate for changes in the returned intensity. In that case,  $K$  should be influenced by the laser power and the flying altitude according to (2).

Once estimated, the average value of  $K$  was further used to convert photons provided by the DART simulations into digital counts to make the direct comparison possible between simulated and measured signals.

## B. Simulations

The accuracy of the DART simulations was evaluated for simulations of ALS acquisitions over the flat targets (calibration plot) and the complex 3-D systems (forest plot). The simulations were computed with `pytools4dart` [32].

1) *ALS Sensor*: The sensor parametrization in the DART simulations was performed according to the sensor characteristics given in Section II. The magnitude of the lidar waveforms simulated by DART is expressed as a number of photons, which depends on the backscattered signal received by the sensor and on the digitization rate. It was converted into LMS-Q780 intensity in digital numbers (DN), applying the gain  $G$  as expressed in (6) and flooring the result to its integer part in order to reproduce quantization:

$$G = \frac{E_p \cdot f_{A/D} \cdot K}{D^2} \quad (6)$$

where  $E_p$  is the photon energy,  $f_{A/D}$  the frequency of digitization, and  $D$  the diameter of the receiver aperture (see Table I). For easy visualization of the whole simulation, the point cloud was extracted from the waveforms (in DN) with the Gaussian decomposition algorithm implemented in [33], [34].

2) *Calibration Plot*: The 3-D mockup of the calibration plot was built by draping the georeferenced target polygons over the DTM and transforming them into a 3-D format compatible with DART (i.e., Wavefront obj format [35]). The optical properties of the mockup elements were defined using the target reflectance values previously extracted from imaging spectroscopy.

The simulation of the ALS point cloud on the targets was computed for each of the 23 flightlines shot patterns. The measured point clouds could be directly compared to the corresponding simulated point clouds as the targets were opaque. Thus, the linear relationship between observation and simulation intensities was tested for each flightline.

### 3) Forest Plot:

a) *Mockup construction*: The conversion of TLS point clouds into a 3-D vegetation mockup of the forest plot was performed as follows [36].

- 1) Filter mixed points in original point cloud based on deviation value [24].
- 2) Classify leaf and wood points from the filtered TLS point cloud using a random forest model based on local geometric and reflectance indices (the balanced accuracy, evaluated for a manually segmented small scene of  $\sim 10$  million points, was found to be 0.94).

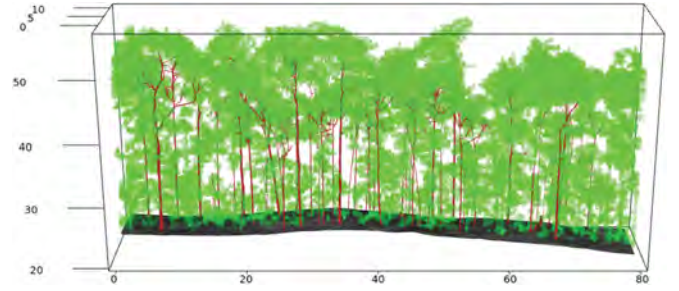


Fig. 2. Slice of  $80 \times 10$  m of the 3-D mockup structure, with the meshes of wood in brown, the voxels of foliage ( $LAD > 0.5 \text{ m}^2 \cdot \text{m}^{-3}$ ) in green and the DTM in dark grey.

- 3) Reintegrate mixed points and assign leaf or wood class label to mixed points based on the label of the nearest neighbor in the filtered point cloud.
- 4) Reconstruct the 3-D meshes of the wood elements (trunks and large branches) from the wood points with the Simpletree method implemented in the Computree software [37].
- 5) Voxelize the leaf point cloud with AMAPVox 1.0 [4], [38] to estimate the leaf area density at a resolution of  $0.5 \times 0.5 \times 0.5$  m.

The ground was represented by a regular triangle mesh obtained from the 0.5 m resolution digital terrain model raster.

DART can either consider each voxel as a turbid element or convert this turbid medium into a set of triangles of uniform size and randomly distributed in space and orientation within the voxel. In this study, the both representations were tested, and the triangle size was set to  $0.01 \text{ m}^2$  when the turbid-to-triangles conversion was done. The two representations will be further referred to as the “turbid” and “triangle” scenarios. Fig. 2 shows a slice of the mockup structure.

The optical properties of the different elements of the mockup were specified as follows.

- 1) For leaves, dominant tree crowns were first delineated based on the canopy height model (computed from the ALS point cloud) and relying also on RGB orthophotos and hyperspectral images [39]. Leaf optical properties measured in the field were assigned to leaf voxels for a subset of 22 crowns belonging to 19 species which were unambiguously identified in the RGB imagery. Default leaf optical properties were used for other leaf voxels. These default properties were simulated using PROSPECT-D [40] parametrized with the average of the values measured in the field when available (leaf chlorophyll content =  $40.8 \mu\text{g}/\text{cm}^2$ ; leaf carotenoid content =  $7.5 \mu\text{g}/\text{cm}^2$ ; leaf anthocyanin content =  $0.4 \mu\text{g}/\text{cm}^2$ ; leaf mass per area =  $12.3 \text{ mg}/\text{cm}^2$ ; leaf water content =  $0.012 \text{ cm}$ ; leaf structure parameter = 1.8).
- 2) Trunk meshes were defined with a deciduous bark reflectance measured in the field which was 0.42 at 1064 nm.
- 3) The ground optical property was chosen after evaluating the reflectance of single ALS ground returns on the forest plot, which showed an average reflectance of 0.42 (and 0.07 of standard deviation) at a wavelength of 1064 nm.

*b) Simulation/reality comparison:* For complex environments, comparing real and simulated waveforms instead of echoes appears to be most relevant. First, discrepancies between real and simulated echoes could arise as the exact approach used by the data provider for echo extraction is undocumented. Second, the 3-D representation (1) and the optical properties (2) of the canopy of the forest plot cannot be considered to be exactly the same as those of the real forest

- 1) The nonwoody vegetation is represented by voxels of  $0.5 \times 0.5 \times 0.5$  m, either turbid or filled with flat triangles of a fixed size and following a spherical leaf angle distribution, while in reality leaves are not triangular nor completely flat objects, they have variable sizes and orientations, and they show some degree of clumping. In addition, small branches and twigs were pooled with leaves in the wood-leaf segmentation step. The resulting differences in the 3-D distribution of targets between the real forest plot and its mockup representation are likely to induce a difference in the temporal distribution of the backscattered energy and thus in the shapes of the waveforms.
- 2) Another major difference between real and simulated forest stems from the optical properties of the elements of the scene, which have consequences on the magnitude of the waveforms: leaf optical properties show natural variability within individual tree crowns, while they are uniform within a crown in the mockup.

Thus, it is expected to have local discrepancies between measured and simulated waveforms with changes in the number, the location, the width and the amplitude of the echoes.

Comparing the simulated point clouds to the measured point clouds after applying a waveform decomposition algorithm would make the comparison highly dependent on the choice of the algorithm, especially in the forest environment where multiecho returns are prone to happen. Therefore, the comparison between simulations and measurements was made at the waveform level, aggregating the waveforms at various spatial resolutions to decrease the effect of the much undersampled variability of the 3-D mock-up properties as compared to the real forest plot.

To that end, the waveforms were spatialized, converting each waveform bin to a 3-D point. The resulting dense 3-D point cloud was further normalized to the DTM by removing the ground altitude from the  $z$  coordinate. This way, mixing of vegetation and ground signals was avoided during aggregation, enhancing the comparability across the different aggregation scales. The dense point cloud was then aggregated by summing the point intensities at various 3-D grid resolutions  $[r_x, r_y, r_z]$  with  $r_x = r_y \in \{2, 5, 10, 20, 40, 80\}$  m and  $r_z = 1$  m. It resulted in  $\{1600, 256, 64, 16, 4, 1\}$  vertical intensity profiles according to the horizontal resolution used for aggregation. The simulated intensity profiles were compared to the measured intensity profiles using the following set of metrics.

- 1) The correlation between each measured and simulated vertical intensity profile.

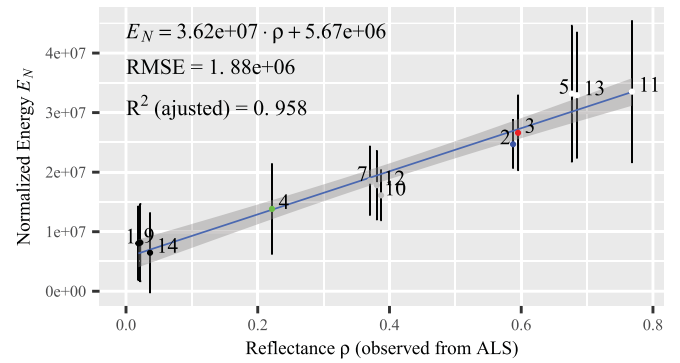


Fig. 3. Linear fit between the normalized energy measured from ALS and the reflectance measured from airborne imaging spectroscopy on the flightline acquired at 18:37 UTC (power = 6%, altitude = 467 m). The vertical bars correspond to the standard deviation of the normalized energy within target footprints.

- 2) The root mean square error normalized to the range of the measured intensity profile (Profile NRMSE).
- 3) The relative difference of the peak intensity between measured and simulated profiles, using measured as reference.
- 4) The root mean square error of the heights AGL of the intensity quantiles every 5% between the 5th and the 95th percentiles (Height RMSE).

The first two metrics measure similarity of the whole intensity profile and thus simultaneously provide information on the quality of the radiometric and geometric information of the simulated data. The third metric focuses on the maximum intensity, which is a key information for sensor calibration, while the last metric provides information on the shape of the profiles and thus measure the quality of the spatial distribution of the signal, i.e., the quality of the geometric information of simulated signals. Maximum intensity values from observation and simulation profiles were compared using their relative difference (with observation as the reference) so that they are comparable across aggregation resolutions. The overall assessment of simulation quality was based on the analysis of the distributions of these metrics for the set of profiles, the size of which depended on the aggregation resolution.

## IV. RESULTS

### A. Calibration

The relationship between the target reflectance measured from airborne imaging spectroscopy and the average value of  $E_N$  measured from ALS was analyzed independently for each flightline, in order to evaluate the variation between the different acquisitions. Fig. 3 illustrates the relationship for one flightline. It highlights the linear relationship between the measurements acquired simultaneously with the two airborne sensors, as well as the variability among ALS measurements belonging to the same target. This variability is explained by the directional effects induced by the surface roughness of the non-lambertian tarps, despite efforts to lay them out flat. However, averaging  $E_N$  over



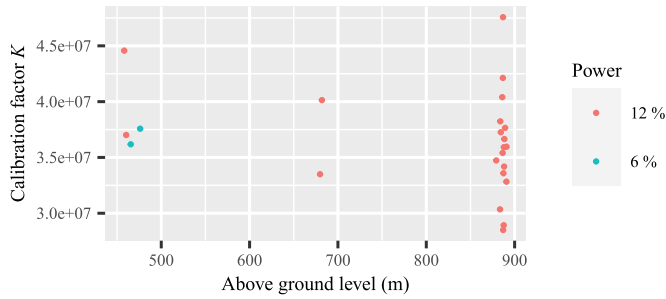


Fig. 4. Calibration factor of each acquisition in function of the laser power and the flying altitude.

each target resulted in a linear relationship between  $E_N$  and reflectance, characterized by a strong correlation (see Fig. 3).

The calibration factor  $K$  of the sensor was computed based on the linear relationship between airborne imaging spectroscopy data and ALS data acquired for each of the 23 flightlines over the calibration site. The values of  $K$  range from  $2.87 \cdot 10^7$  to  $4.76 \cdot 10^7$ , with an average value of  $3.65 \cdot 10^7$ , and its standard deviation was  $4.5 \cdot 10^6$ , which corresponds to a coefficient of variation of 12.5% across acquisitions. The results for each flightline are available in Appendix A.

The values of  $K$  were compared between different acquisition configurations, including changes in sensor power and altitude of acquisition, in order to identify a possible AGC in response to the average magnitude of the returned waveforms. Our results presented in Fig. 4 showed no evidence of a relationship between the value of  $K$  and the laser power or the flying altitude, suggesting that the RIEGL LMS-Q780 does not utilize an AGC, or, at least, not for the range of measured intensities (the maximum intensity measured over the targets was 2895 DN) at these flying altitude and laser power combinations.

The average value of  $K$  was further used as the calibration factor to transform the DART outputs before the comparison between the simulation and the real data.

## B. Simulations

1) *Calibration Plot:* The returned energy at the receiver aperture was compared between measured and simulated lidar acquisitions. The median of the returned energy per calibration target and per flightline was used as input for the model in order to minimize both the influence of the target size on the regression model (larger targets are sampled more than smaller targets), and the influence of surface roughness, which was not accurately reproduced in the simulation. Fig. 5 shows the slope of the regression model of each flightline as a function of the calibration factor of each acquisition. The average and the coefficient of variation of the slopes of the models were 0.992% and 11.6%, respectively. It confirms the capacity of DART to accurately reproduce the magnitude of the ALS intensities on flat and opaque targets larger than the lidar footprint such as the tarps used in this experiment.

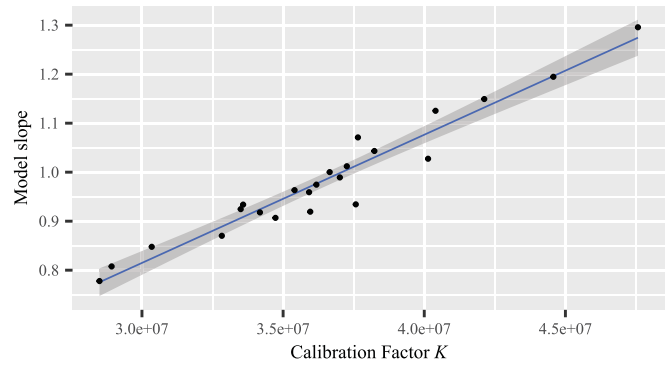


Fig. 5. Influence of  $K$  value (assessed by flightline) on the slope of the models between observations and simulations returned energy. Each point represents a flightline. The blue line is the linear regression between  $K$  and the model slopes.

2) *Forest Plot:* A selection of the intensity profiles of both the turbid and triangle scenarios at different aggregation resolutions is presented in Fig. 6. Although it cannot be considered as fully representative of the waveform results, it illustrates the consistency between real and simulated intensity profiles for both scenarios. Fig. 6 also highlights that the peak at ground level is overestimated in simulations. The fact that the ground peak does not appear in Fig. 6 on the profiles at small grid size is a coincidence. An analysis of the distribution of this peak through the different profiles at 2 m grid size showed that the overestimation is not distributed evenly across the scene. Several factors can explain the ground peak overestimation and are discussed in Section V.

The comparison between the aggregated profiles of observations and simulations is summarized in Fig. 7 by intensity and height metrics. Overall, aggregating waveforms over larger areas improved the consistency between measured and simulated values. Such a regularizing effect was expected since the aggregation smooths out much of the local variability (both in terms of geometric and optical properties of the scatterers) that is not perfectly reproduced in the mockup derived from the TLS point cloud. At the coarsest spatial resolution (80 m), the intensity profiles showed a high correlation ( $R > 0.99$ ) and low error (NRMSE = 5% for the turbid scenario), attesting to a general fit of the structure and optical properties used in the simulated forest, as well as of the good performance of the DART model. The median correlation between the observed and the simulated intensities increased with the decrease in horizontal resolution and ranged from 0.92 at 2 m resolution up to 0.98 at 20 m resolution [see Fig. 7(a)], confirming the strong agreement between the shapes of the aggregated profiles at 20 m and beyond. The median RMSE computed over heights of intensity quantiles [see Fig. 7(d)] ranged from 1.5 at 2 m resolution to 0.6 m at 80 m resolution.

For the turbid scenario, the median of the profile NRMSE ranged from 12% at 2 m resolution to 6% at 20 m resolution and 5% at 80 m resolution [see Fig. 7(b)]. Fig. 7(c) shows the relative difference of the peak magnitudes of intensity profiles. It confirms that the maximum of intensity, independently of its

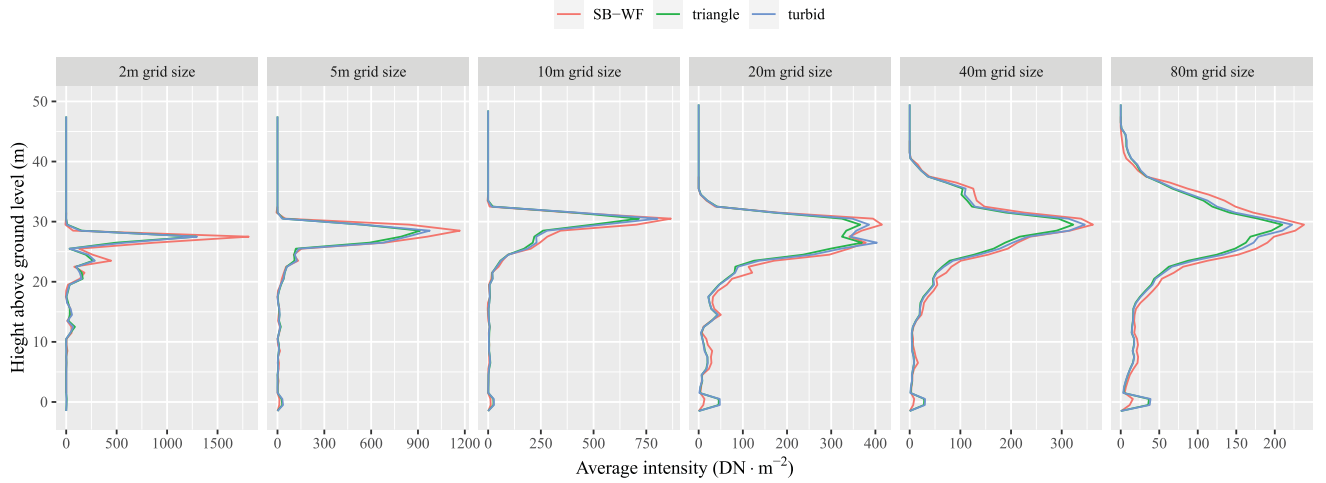


Fig. 6. Example of intensity profiles for the different cell sizes of analysis ranging from 2 to 80 m. SB-WF are derived from real data (with different waveform reconstruction methods). Triangle and turbid are derived from simulations with the different leaf voxel representations.

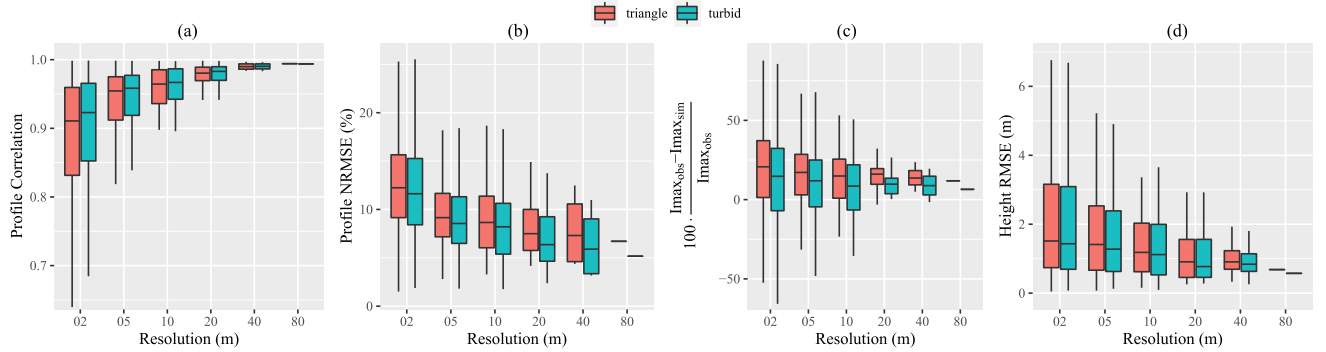


Fig. 7. Comparison between SB-WF and simulated intensity profiles for different aggregation resolutions. (a) Correlation. (b) NRMSE. (c) Relative difference of peak intensity. (d) RMSE of the heights AGL of intensity quantiles.

vertical location on the profile, was generally underestimated. The median error is 14% to 20% at 2 m resolution depending on the mockup type, and 10% to 16% at 20 m resolution. At the lowest resolution (80 m), the peak intensity of the profiles was underestimated by 6.5% for the turbid scenario and 12% for the triangle scenario. The scene employing the turbid vegetation model outperformed the one with triangles in terms of intensity, especially at the peak of intensity. But despite such discrepancies, the results highlight the global consistency between turbid and triangle scenarios.

## V. DISCUSSION

The radiometric validation of ALS waveforms simulated with the DART model first required to calibrate an ALS sensor on simple targets introducing the sensor specifications in the simulator. Once the radiometric calibration was achieved, simulated and real ALS data were compared for a complex forest plot. In this section, the results are discussed before considering the prospects for applications of this study.

### A. Calibration

The calibration factor  $K$  required in (2) was estimated using a set of flat targets laid on the ground in a clearing, the reflectance of which was measured by imaging spectroscopy. Fitting  $K$  in this way allowed relating the reflectance of the targets to the average intensity of the ALS returns of each flightline used in this study. Our experiment also showed that  $K$  could be assumed to be constant as  $K$  was found to be independent of laser power, flying altitude and reflectance, thus suggesting that the sensor did not include AGC. However, a variability in  $K$  values of 12.5% was observed between flightlines.

Investigating the variability of  $K$ , we found that the peak intensity of the ALS echoes was influenced by the signed value of the scan angle, i.e., the rotational angle at which the laser pulse was emitted taking into account the roll angle of the aircraft ( $0^\circ$  being nadir,  $-90^\circ$  the left side of the aircraft in the direction of flight). Indeed, Fig. 8 shows a linear regression between  $K$  and scan angle. Although the relationship between echoes peak intensity and scan angle was evidenced for the whole ALS



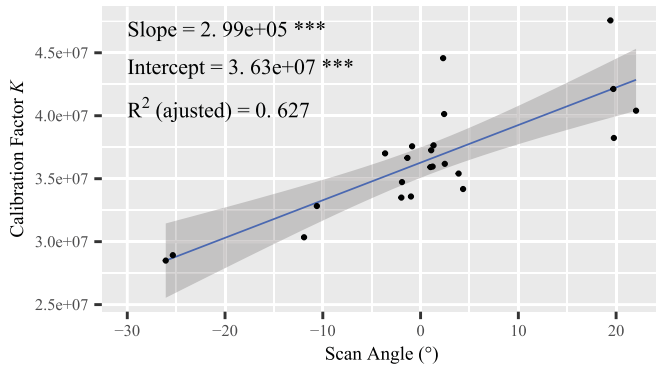


Fig. 8. Influence of the scan angle on the calibration factor  $K$ , each point corresponds to a flightline. Both the values for slope and intercept are significant at  $p \leq 0.001$  (\*\*\*)

dataset (not shown here), such a relationship was unexpected and we did not manage to formulate any solid physical justification for this influence except a possible issue with the sensor or the sensor set-up on the airplane (which could not be checked). In order to reduce the possible impact of the scan angle on the rest of the experiments, the scan used for the simulations over the forest plot was selected based on a criterion of minimum scan angle in that area.

When accounting for the effect of scan angle, the coefficient of variation of  $K$  between flightlines was reduced to 7.5%, i.e., 40% of the original variation, although 20% deviation still remained for the farthest values from the regression line as shown in Fig. 8. The remaining variability of  $K$  may be due to other factors such as the variability of target optical properties and the surface configuration. The targets were standard tarps with contrasted reflectance at the laser wavelength. However, these targets cannot be considered as lambertian surfaces and may have presented a specular component, as evidenced on the orthophotography (white reflections on colored tarps in Fig. 1). In addition, the flatness of the tarps could not be guaranteed during the field experiment. This may have resulted in specular reflection in various directions. A precise characterization of the bidirectional reflectance distribution function of the tarps complemented by a fine-scale surface model would have been needed to test the influence of this specular component on the variability of the intensity observed among the returns of the same tarp for a given flightline (see vertical bars in Fig. 3). It would also have helped to understand to what extent this variability contributed to the variation of  $K$ . The absorption of the signal in the atmosphere may be another factor explaining the variability of  $K$  across flightlines. Although the probability of changes in atmospheric conditions was minimized by performing the acquisition in clear sky conditions and within a short time period (2.5 hours between 13:10 and 15:40 local time), the tropical environment is prone to the presence of water vapor in the troposphere in concentrations that can vary significantly during the day [41]. This may lead to changes in absorption of the returned energy.

The simulation on the calibration plot showed accurate reproduction of the returned intensity over flat targets, confirming the

capacity of the DART model to reproduce the lidar signal on flat targets. These results over the calibration plot demonstrated the feasibility to calibrate an ALS sensor implementing a quite simple field experiment in order to obtain realistic simulations. This is an important point, as information on radiometric calibration of ALS sensors are rarely available and as independent calibrations in laboratory can be laborious and expensive. In the current study, the calibration experiment benefited from the simultaneous acquisition of ALS and imaging spectroscopy. The calibration protocol presented could be alleviated by the spectroscopic characterization of the targets in laboratory, an option that still has to be tested.

### B. Simulation in Dense Forest

The analysis of the simulations on the forest plot was less straightforward, as the environment is highly complex from both the structural and the radiometric point of view. Despite this complexity, consistency between simulated and real signals was found to be high. However, a detailed examination of the simulations revealed discrepancies with the observed signal. The most noticeable are the underestimation of the waveform intensity, the overestimation of the ground peak and the influence of the aggregation scale. In this section we discuss the possible origin of these discrepancies and draw the limits and areas of improvements of the current study.

1) *Origin of Discrepancies:* The ground peak overestimation in the simulations is an interesting case as it may involve both the quality of the 3-D scene and the sensor sensitivity.

Three elements might have caused the overestimation of the ground peak in the simulations. First, the energy returned from the ground depends on the energy intercepted by the vegetation above. Thus, ground peak intensity overestimation could be the consequence of an underestimation of the vegetation density in the canopy or in the understory. Second, the ground was represented by a regular triangle mesh with uniform optical properties. However, the ground is usually covered by low vegetation, and both optical properties and 3-D structure (roughness and slope orientation) vary over space at different scales. Third, the detection limit of the real ALS sensor, triggering the recording of the waveform sample blocks (not simulated), may have censored small ground returns, leaving them unrecorded if they occurred far from the detected returns of the same pulse, i.e., outside the recorded waveform sample blocks.

Several considerations point towards an underestimation of the LAD to be the main cause of the ground peak overestimation. Indeed, Figs. 6 and 7 show a global underestimation of the energy of the profiles at all aggregation scale. This underestimation is even more prominent at the peak energy [see Figs. 6 and 7(c)]. In Vincent and Heuschmidt [36], the simulation with AMAPVox of diffuse light through the 3-D scene derived from TLS data was compared to the measurements obtained two years before (2014) from LAI2200 canopy analyser. The average predicted transmittance was 2.8% whereas the measured transmittance was 2.1% [36], a difference that could not be explained by a possible change in the canopy cover over a two year period.

An explanation of the underestimation of the LAD could come from the assumptions made when computing it from the TLS data. When scanning from below the canopy any pulse not generating a return is considered as traversing the canopy. In dense forest a small rate of “false no-return” pulses, i.e., pulses intercepted but with not enough energy returning to the sensor to trigger the corresponding echoes, might significantly bias the final transmittance estimate and lead to an underestimation of the LAD in the forest mockup.

According to Fig. 6 at 80 m resolution, the difference of the intercepted energy would be distributed all along the profile (with an inversion at the bottom). However, the top of the canopy constitutes the farthest target from the TLS view, thus the location where the LAD estimation error is expected to be the largest, while it is also the first target from the ALS view, thus the location where the ALS sampling is the densest. In order to have a clearer picture at the top of the canopy, we conducted the same analysis as in Section IV-B2, but aligning the  $z$ -coordinate of the waveform dense point cloud to the canopy height model (CHM) instead of the DTM. The details of this analysis are presented in Appendix E. Fig. 12 in Appendix E confirms that most of the energy loss occurred in the top meters of the canopy and that the underestimation of the intercepted energy at that point was ranging between 20% and 30%. These values match with the transmittance overestimation assessed in [36]. It suggests that LAD was underestimated for the top of canopy voxels of the forest mockup. The vertical distribution of the error observed in Fig. 6 (80 m resolution) could be the consequence of the aggregation of waveforms from lidar shots reaching the canopy at various heights.

In addition to a possible LAD underestimation, other factors could contribute to the underestimation of the intercepted energy in the simulations at the canopy level. As it has been mentioned in Section II, the real waveforms suffered from a background noise and potential ringing effect. Tests conducted without removing the background noise on the real waveforms (not presented here) doubled the NRMSE between actual and simulated profiles, as well as the relative difference of the peak of intensity of the profiles. Although it was reduced by shifting the SB-WF waveforms, the background noise was observed not to be constant and the residual noise could still contribute to an increase of the profile magnitude. Similarly, the high intensity returns can generate a so called ringing effect on the sensor signal, an effect that could also be part of the discrepancy observed between the real and the simulated waveforms.

2) *Limits and Future Improvements:* From a more global point of view, Fig. 7 underlines an influence of the aggregating scale, with better adequation of simulations with real signal at large than at small aggregation scales. Setting aside the possible LAD underestimation, it shows a performance variability (boxes hinges and whiskers in Fig. 7) increasing exponentially as the resolution increases. This suggests that at high resolutions, realistic simulations would need more effort in the reconstruction of a detailed mockup, by improving both the representation of the structure and the characterization of optical properties of the scene elements. Although much effort was devoted to building

a detailed 3-D scene, it remains a simplification of the reality. In addition to the likely aforementioned LAD underestimation, this is also true for the optical properties for which only a few leaf measurements were available from the experimental campaign leaving the majority of the leaf voxels of the mockup with a single default optical property. High tree-to-tree variability has been observed in apparent reflectance at the site [42]. Thus, simulations would certainly benefit from a greater effort in the leaf traits determination and in their distribution in the scene. This could be obtained from intensive field data collection, or from estimation of leaf traits from imaging spectroscopy data [43]. The conversion of voxels from turbid to a random distribution of triangles did not seem to enhance the structure representation. Thus, more investigation is needed to pinpoint the main factors driving the discrepancies between observations and simulations at small scales.

To this end, several areas of improvements can be expected from recent research. Progress in 3-D neural network segmentation [13], [44], [45], [46], [47] may help in enhancing the identification of the nature and the spatial distribution of the elements of the mockup, in particular leaf-wood segmentation that is a standard preliminary step to tree reconstruction from TLS data. Promising improvements have been made recently which could help increase the accuracy of the structure parameters. New estimators of LAD, likely to be less prone to bias than the one used in the present study, have also been proposed for single return scanners [17], [48], [49] and extended to multiple return scanners [50]. However, these estimators still need to be validated in a tropical forest context. Combining TLS data with UAV laser scanner [51] might also improve LAD estimates at the top of the canopy where first and main interceptions occur for ALS or spaceborne lidar systems. Leaf angle distribution [52] could be extracted from TLS data in principle, at least at close range, if not available in literature, and contribute to a better representation with triangles. Efforts to estimate leaf traits distribution from imaging spectroscopy [43] should also contribute to add realism in that part of the 3-D scene.

At the sensor level, Appendices D and E show that work remains to be done in order to understand, disentangle, and model the different sources of signal imperfections in order to increase simulation realism at small scale. We did not compare measured and simulated point clouds in this study because it raised additional questions on the processing chain of the ALS data, such as the methods implemented in RIEGL software for waveforms decomposition. Still, Fig. 9 highlights what can be achieved after a standard Gaussian decomposition of the forest plot simulation. It illustrates the level of consistency that could be expected with such a simulation as well as some of the differences noticed on the waveform comparison (e.g., the higher density of echoes in the simulated point cloud, particularly visible at the ground level). Accounting for sensor noise and noise filtering in the Gaussian decomposition process should lead to more realistic point cloud density. To that aim, increased transparency in sensor specifications and processing algorithms would be helpful.

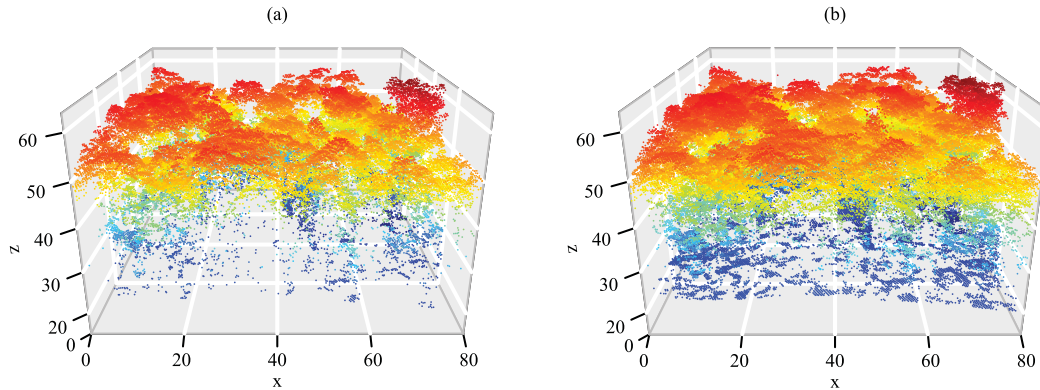


Fig. 9. Point clouds of the forest plot for (a) real data and (b) simulation, the color being function of the height.

### C. Prospects of DART Lidar Modeling Applications

Aggregated at the scale of the  $80 \times 80$  m plot, the simulated full-waveform showed a very close agreement with real data in terms of both structure (intensity profile correlation of 0.99, RMSE of the heights of intensity quantiles less than 1 m) and intensity (profile NRMSE is 5%). Aggregated at the scale of 20 m, the simulations still resulted in consistent intensity profiles (correlation of 0.98, median RMSE<sub>I<sub>q</sub></sub> around 1 m and median NRMSE of 6%).

Such results could be of high interest for the development of improved area-based models [53], that are used to predict forest stand attributes with variables derived from the analysis of ALS point clouds at the level of traditional field inventory, i.e., for forest plots with diameters ranging from about 20 to 30 m. Indeed, simulations could be used to better analyze, understand and model impacts of acquisition conditions, e.g., pulse density [54], [55], [56] or scan angles [57]. They could also provide large datasets and facilitate the set-up of data augmentation strategies to implement deep learning approaches [13] in a field of application for which acquiring large data sets with synchronous field measurements and ALS acquisitions is challenging.

The accuracy of the DART lidar model at 20 m scale is also interesting for spaceborne lidar as it corresponds to the ground footprint diameter of actual (and under development) sensors. Simulating large footprint waveforms for spatial missions (e.g., GEDI-NASA [58], MOLI-JAXA [59], LEAF-CNES [60], [61]) can be envisioned in several ways with the DART model. First, from TLS datasets: the quality of the waveforms obtained by summing up simulations of ALS data allows to be optimistic about the capacity to accurately simulate large footprint waveforms with DART in forest environments. Second, simultaneously simulating ALS data and large footprint waveforms could be used to validate and optimize simulators based on the conversion of discrete-return ALS data to a large-footprint full-waveform signal, such as the one developed for the calibration of algorithms and assessment of the GEDI mission accuracy [62]. For example, it would be of high interest to better model the contribution of multiple scattering when simulating large footprint waveforms from small footprint signals, for which the contribution of multiple scattering is negligible. In addition, it

could help to get a better handle on a possible change in wavelength between both airborne and spaceborne systems. Third, generating forest scenes from ALS data using voxelization could also be considered to simulate large footprint waveforms with DART. Independently of the way they are produced, the accurate simulation of such waveforms can already serve numerous applications such as the refinement of specifications for new spaceborne sensors, as well as training and testing algorithms and their adaptation between sensors with different specifications to keep continuity in product quality. These applications could be extended to smaller scale and resolution with the enhancement of the 3-D scene accuracy and the improvement of the ALS sensor model. Combined with the speed-up of the models [9], [10], [32], a simulator such as DART could become a key element to transfer information from one scale to another (local to global, terrain to spatial), and to enhance the simulation and thus the analysis of other sources (multi/hyperspectral, radar).

## VI. CONCLUSION

In this article, we investigated the capacities of the DART model to simulate ALS full-waveform acquisitions. The simultaneous acquisition of ALS and imaging spectroscopy over a calibration plot with ground targets allowed to evaluate the calibration factor linking the energy at the entrance of sensor aperture to the intensity values delivered by the sensor. The calibration factor was determined with a limited variability (CV 12% to 7.5% when taking into account effects stemming from the scan angle) and then used with the DART model to simulate full-waveform ALS acquisitions. Compared with the real ALS data, the simulations confirmed the accuracy of the DART model over a simple scene as well as a highly complex 3-D scene derived from TLS data acquired in a dense tropical forest. The multiscale analysis (2 to 80 m) of the forest plot simulations evidenced a strong correlation with the real data (92% to 99%) and small error in intensity value (NRMSE 12% to 5%) and in vertical distribution of the returned energy (RMSE 1.5 to 0.6 m). With such a level of confidence in the model, the analysis of the simulations allowed us to point out a possible underestimation of the leaf area density assessed from TLS data at the top of the canopy, and to identify several areas of improvement for both the 3-D scene and the sensor model. Finally, this work opens



TABLE III  
CALIBRATION TARGETS WITH THEIR SIZE, COLOR, AND REFLECTANCE AT LIDAR WAVELENGTH EXTRACTED FROM IMAGING SPECTROSCOPY

ID	Color	Size (m)	$\rho_{1064\text{nm}}$ ( $10^{-2}$ )
1	black	4	1.9
2	blue	2	58.7
3	red	2	59.5
4	green	2	22.2
5	white	2	67.7
6	white	1	53.6
7	grey	2	37.0
8	white	1	51.3
9	black	2	2.2
10	grey	2	38.7
11	white	4	76.8
12	grey	4	38.1
13	white	2	68.5
14	black	2	3.7

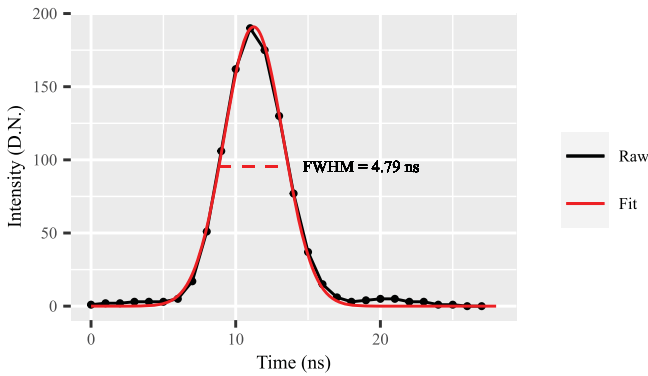


Fig. 10. Example waveform of an emitted pulse extracted from the RIEGL SDF file of the forest plot scan. The dashed line shows the FWHM which is 4.79 ns in that case.

up multiple prospects of applications, at short and long term, to improve the monitoring of forest environments from lidar based observations.

## APPENDIX

### A. Overview

This Appendix provides additional information on the material used for the experiments, and complementary analysis that helped understanding the results.

Section B gives specifications on the calibration target and on the flightlines used for the calibration experiment. Section C and Section D provide more details on the emitted and returned pulses extracted from RIEGL SDF files. Section E gives supplementary elements for CHM instead of DTM aligned waveform aggregation.

### B. Specifications of Calibration Experiment

Table III details the specifications of the calibration targets in size, color and apparent reflectance at 1064 nm extracted from imaging spectroscopy.

TABLE IV  
FLIGHTLINES USED IN THE CALIBRATION EXPERIMENT, WITH THE LOCAL TIME AT WHICH THEY STARTED, THE FRACTION OF POWER USED BY THE LASER, THE HEIGHT ABOVE GROUND LEVEL OF THE SENSOR, THE AVERAGE SCAN ANGLE OVER THE CALIBRATION PLOT, AND THE REGRESSION RESULTS FOR THE ESTIMATION OF THE CALIBRATION FACTOR ( $K$ , RMSE AND  $R^2$ )

UTC Time	Power (%)	AGL (m)	Scan Angle ( $^\circ$ )	$K$ ( $10^7$ )	RMSE ( $10^6$ )	$R^2$ (%)
16:12	12	891	1.2	3.6	3.41	87
16:37	12	888	-25.3	2.89	1.1	97
16:41	12	884	19.8	3.82	1.9	96
16:44	12	891	-10.6	3.28	1.89	95
16:48	12	887	3.9	3.54	2.63	92
16:52	12	888	4.3	3.42	2.45	92
16:56	12	883	-11.9	3.03	1.53	96
16:59	12	887	19.4	4.76	1.65	96
17:05	12	887	19.7	4.21	0.94	98
17:09	12	887	-26.1	2.85	1.41	95
17:51	12	886	22	4.04	1.76	96
17:55	12	879	-1.9	3.47	1.75	96
17:57	12	888	1	3.59	2.23	94
18:01	12	888	-1.3	3.67	1.3	98
18:10	12	889	1.3	3.77	2.21	95
18:13	12	887	-1	3.36	1.68	96
18:16	12	884	1.1	3.73	2.6	93
18:20	12	680	-2	3.35	1.39	97
18:24	12	682	2.4	4.01	2.64	93
18:27	12	460	-3.6	3.7	2.63	92
18:30	12	458	2.3	4.46	2.91	94
18:33	6	476	-0.9	3.76	1.88	96
18:37	6	465	2.5	3.62	1.88	96

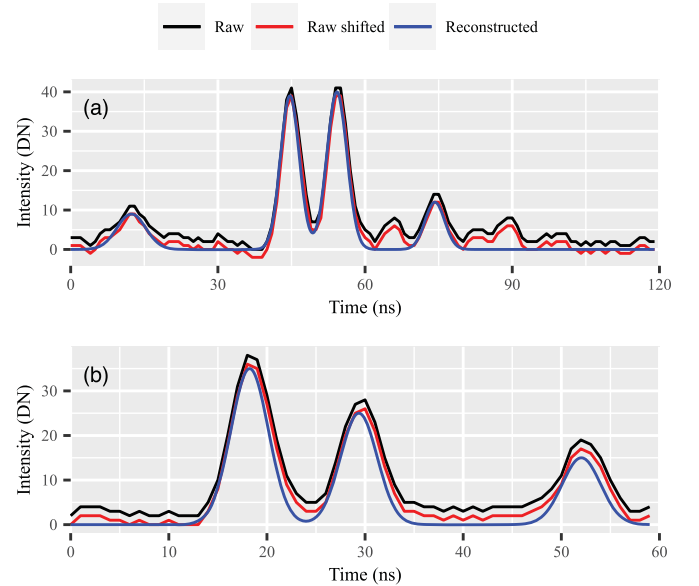


Fig. 11. Examples waveforms of two pulses returns. In black, the raw waveform was extracted from the RIEGL SDF file of the forest plot scan. In red, the raw waveform was shifted by  $-2$  DN. In blue, the reconstructed waveform was reconstructed from the point cloud information.

Table IV lists the specifications of the flightlines used in the calibration experiment, as well as the regression results obtained for each of them.

### C. Emitted Waveforms

The outputs of the lidar sensor LMS-Q780 includes SDF file that contains records (i.e., waveform sample blocks) of the

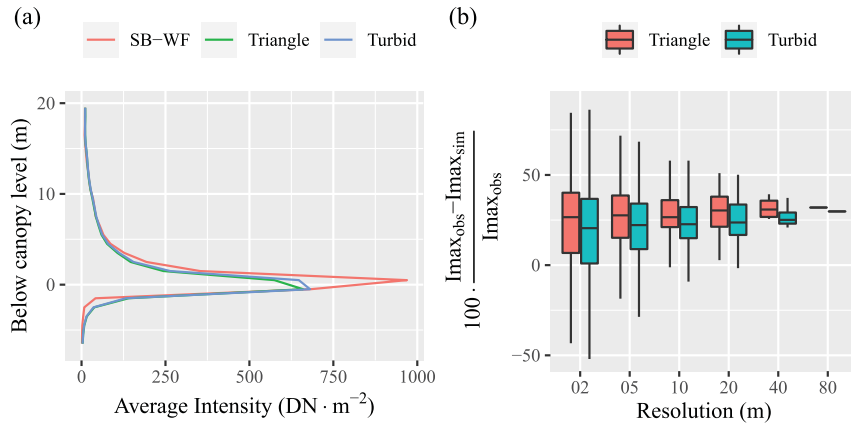


Fig. 12. Comparison of the intensity profiles derived from real and simulated data after alignment of waveforms to the CHM (only the simulation with turbid voxel considered here). (a) Intensity profile aggregation at 80 m grid size. (b) Relative difference of peak intensity for the different aggregation resolutions.

emitted waveforms and the returned waveforms. During our study, these files were used in order to check the assumption of a Gaussian shaped emitted pulse, to evaluate the FWHM value of the emitted pulse and to validate our waveform reconstruction strategy.

The nonlinear least square regression of a Gaussian fit on the emitted pulses over the forest plot shows an average NRMSE of 1.2% and a FWHM average value of 4.815 ns with a standard deviation of 0.018 ns (i.e., 0.4%). These results validate our hypothesis of a Gaussian shaped emitted pulse.

Fig. 10 shows an example of emitted pulse extracted from the RIEGL SDF files of the campaign, and the corresponding Gaussian fit. Although the fitting results are good, one can also notice a small local maximum after the main pulse that may be due to ringing.

#### D. Reconstructed Waveforms

Two alternative methods were used to prepare the ALS waveforms for the comparison to the simulations. The SB-WF reconstruction was performed concatenating the waveform sample blocks extracted from the ALS SDF files, and removing the background noise as described hereinafter. The blocks of the same shot are aligned using the fit of the emitted pulse in order to determine the starting time of the shot, the interpolated trajectory to determine the location of the sensor and the echo location (from the point cloud) to determine the echo to sensor distance (and thus light travel time).

The Gaussian-based waveform (GB-WF) reconstruction was performed as a sum of Gaussians parameterized by the intensity and the width of the echoes, and positioned at the time of flight between the sensor and the echoes (deduced from the distance between the sensor and the echo). In order to limit the size of the dataset, the reconstruction was made only in the interval  $\pm 5\sigma$  around each echo. The GB-WF waveform reconstruction method is fast to compute and benefits from RIEGL signal filtering (background noise, ringing effect, misleading signal due to sensor sensitivity, ...). However, as small echoes can be difficult to separate from the noise or the ringing effect, they are usually censored by the sensor detection limit either during the

acquisition (i.e., not recorded) or during the data postprocessing stage, as data providers usually prefer to avoid having noise in the final dataset at the cost of a small loss of information.

In the current study, the difference between the SDF raw waveforms and GB-WF was used to evaluate the magnitude of the background noise, which showed a median value of 2 DN. Considering the background noise as additive to the signal returned from the target, the raw waveform magnitude was shifted down by 2 DN in order to reduce the impact of the noise on the comparisons with the simulations. The shifted raw waveforms are the ones used in the analysis and referred to as SB-WF in the body of this manuscript.

A couple of examples of returned waveforms extracted from SDF files (raw waveform) and the corresponding reconstructed waveforms from Gaussian sums are shown in Fig. 11. In Fig. 11(a), one can notice the noise level as well as the small echoes discarded during post-processing (the detection limit was evaluated at 7 DN). When the shift was larger than the noise, generating negative values, the samples were set to zero intensity. In Fig. 11(b), it can be noticed that the background noise signal is not constant. Thus, some residual noise signal is still included even after shifting the raw waveforms. More investigation is needed in order to better characterize the background noise and eventually include it in the simulations.

#### E. CHM Normalized Waveform Analysis

The same comparison as in Section IV-B2 was conducted after aligning the z-coordinate of the waveform dense point clouds to the CHM instead of the DTM, the CHM being computed at 0.5-m and with a pit-free algorithm [63]. The aggregated profiles at 80 m grid size are presented in Fig. 12(a) and the relative error at peak magnitude in Fig. 12(b). Both figures evidenced a concentration of energy loss at top of canopy going, from 20% to 30% depending on the grid size and the type waveform considered as real data.

#### ACKNOWLEDGMENT

The authors have no conflicts of interest to declare.

## REFERENCE

- [1] M. Bouvier, S. Durrieu, R. A. Fournier, and J.-P. Renaud, "Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data," *Remote Sens. Environ.*, vol. 156, pp. 322–334, Jan. 2015, doi: [10.1016/j.rse.2014.10.004](https://doi.org/10.1016/j.rse.2014.10.004).
- [2] A. Munoz, J. Bock, J. M. Monnet, J. P. Renaud, A. Jolly, and C. Riond, "Independent validation of prediction of forest parameters from airborne laser scanning," *Fr. Rev. Photogrammetry Remote Sens.*, vol. 211/212, pp. 81–92, 2015.
- [3] M. Maltamo, K. Eerikainen, J. Pitkanen, J. Hyypä, and M. Vehmas, "Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions," *Remote Sens. Environ.*, vol. 90, no. 3, pp. 319–330, Apr. 2004, doi: [10.1016/j.rse.2004.01.006](https://doi.org/10.1016/j.rse.2004.01.006).
- [4] G. Vincent et al., "Mapping plant area index of tropical evergreen forest by airborne laser scanning: A cross-validation study using LAI2200 optical sensor," *Remote Sens. Environ.*, vol. 198, pp. 254–266, Sep. 2017, doi: [10.1016/j.rse.2017.05.034](https://doi.org/10.1016/j.rse.2017.05.034).
- [5] G. Vincent et al., "Accuracy of small footprint airborne LiDAR in its predictions of tropical moist forest stand structure," *Remote Sens. Environ.*, vol. 125, pp. 23–33, Oct. 2012, doi: [10.1016/j.rse.2012.06.019](https://doi.org/10.1016/j.rse.2012.06.019).
- [6] M. Soma, F. Pimont, S. Durrieu, and J.-L. Dupuy, "Enhanced measurements of leaf area density with T-LiDAR: Evaluating and calibrating the effects of vegetation heterogeneity and scanner properties," *Remote Sens.*, vol. 10, no. 10, Oct. 2018, Art. no. 1580, doi: [10.3390/rs10101580](https://doi.org/10.3390/rs10101580).
- [7] S. Wei et al., "An assessment study of three indirect methods for estimating leaf area density and leaf area index of individual trees," *Agricultural Forest Meteorol.*, vol. 292/293, Oct. 2020, Art. no. 108101, doi: [10.1016/j.agrformet.2020.108101](https://doi.org/10.1016/j.agrformet.2020.108101).
- [8] J.-P. Gastellu-Etchegorry et al., "Simulation of satellite, airborne and terrestrial LiDAR with DART (I): Waveform simulation with quasi-Monte Carlo ray tracing," *Remote Sens. Environ.*, vol. 184, pp. 418–435, Oct. 2016, doi: [10.1016/j.rse.2016.07.010](https://doi.org/10.1016/j.rse.2016.07.010).
- [9] X. Yang et al., "Comprehensive LiDAR simulation with efficient physically-based DART-Lux model (I): Theory, novelty, and consistency validation," *Remote Sens. Environ.*, vol. 272, Apr. 2022, Art. no. 112952, doi: [10.1016/j.rse.2022.112952](https://doi.org/10.1016/j.rse.2022.112952).
- [10] L. Winiwarter et al., "Virtual laser scanning with HELIOS++: A novel take on ray tracing-based simulation of topographic full-waveform 3D laser scanning," *Remote Sens. Environ.*, vol. 269, Feb. 2022, Art. no. 112772, doi: [10.1016/j.rse.2021.112772](https://doi.org/10.1016/j.rse.2021.112772).
- [11] P. R. J. North, J. A. B. Rosette, J. C. Suárez, and S. O. Los, "A Monte Carlo radiative transfer model of satellite waveform LiDAR," *Int. J. Remote Sens.*, vol. 31, no. 5, pp. 1343–1358, Mar. 2010, doi: [10.1080/01431160903380664](https://doi.org/10.1080/01431160903380664).
- [12] S. Durrieu and R. F. Nelson, "Earth observation from space – The issue of environmental sustainability," *Space Policy*, vol. 29, no. 4, pp. 238–250, Nov. 2013, doi: [10.1016/j.spacepol.2013.07.003](https://doi.org/10.1016/j.spacepol.2013.07.003).
- [13] J. Morel, A. Bac, and T. Kanai, "Segmentation of unbalanced and in-homogeneous point clouds and its application to 3D scanned trees," *Vis. Comput.*, vol. 36, no. 10–12, pp. 2419–2431, Oct. 2020, doi: [10.1007/s00371-020-01966-7](https://doi.org/10.1007/s00371-020-01966-7).
- [14] T. Yin, J. P. Gastellu-Etchegorry, E. Grau, N. Lauret, and J. Rubio, "Simulating satellite waveform LiDAR with DART model," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2013, pp. 3029–3032, doi: [10.1109/IGARSS.2013.6723464](https://doi.org/10.1109/IGARSS.2013.6723464).
- [15] T. Yin, N. Lauret, and J.-P. Gastellu-Etchegorry, "Simulation of satellite, airborne and terrestrial LiDAR with DART (II): ALS and TLS multi-pulse acquisitions, photon counting, and solar noise," *Remote Sens. Environ.*, vol. 184, pp. 454–468, Oct. 2016, doi: [10.1016/j.rse.2016.07.009](https://doi.org/10.1016/j.rse.2016.07.009).
- [16] L. Landier et al., "Calibration of urban canopies albedo and 3D shortwave radiative budget using remote-sensing data and the DART model," *Eur. J. Remote Sens.*, vol. 51, no. 1, pp. 739–753, Jan. 2018, doi: [10.1080/22797254.2018.1462102](https://doi.org/10.1080/22797254.2018.1462102).
- [17] T. Yin, B. D. Cook, and D. C. Morton, "Three-dimensional estimation of deciduous forest canopy structure and leaf area using multi-directional, leaf-on and leaf-off airborne LiDAR data," *Agricultural Forest Meteorol.*, vol. 314, Mar. 2022, Art. no. 108781, doi: [10.1016/j.agrformet.2021.108781](https://doi.org/10.1016/j.agrformet.2021.108781).
- [18] Z. Zhen et al., "Quantitative analysis of DART calibration accuracy for retrieving spectral signatures over urban area," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10057–10068, 2021.
- [19] J.-L. Widlowski et al., "The fourth phase of the radiative transfer model intercomparison (RAMI) exercise: Actual canopy scenarios and conformity testing," *Remote Sens. Environ.*, vol. 169, pp. 418–437, Nov. 2015, doi: [10.1016/j.rse.2015.08.016](https://doi.org/10.1016/j.rse.2015.08.016).
- [20] E. Grau et al., "Modelling full waveform LiDAR data on forest structures at plot level: A sensitivity analysis of forest and sensor main characteristics on full - waveform simulated data," in *Proc. SilviLaser*, 2015, pp. 146–148.
- [21] S. Gourlet-Fleury, J.-M. Guehl, and O. Laroussinie, *Ecology and Management of a Neotropical Rainforest: Lessons Drawn From Paracou, a Long-Term Experimental Research Site in French Guiana*. Amsterdam, The Netherlands: Elsevier, 2004.
- [22] RIEGL Laser Measurement Systems, "Datasheet LMS-Q780 Riegl," Horn, Austria, Mar. 24, 2015.
- [23] "LAS specifications 1.4 - R15," ASPRS, Baton Rouge, LA, USA, Jul. 9, 2019.
- [24] RIEGL Laser Measurement Systems GmbH, "Whitepaper: LAS extrabytes implementation in RIEGL software," Apr. 15, 2019. [Online]. Available: [http://www.riegl.com/uploads/tx\\_pxpriegl/downloads/Whitepaper\\_LAS\\_extrabytes\\_implementation\\_in\\_RIEGL\\_Software\\_2019-04-15.pdf](http://www.riegl.com/uploads/tx_pxpriegl/downloads/Whitepaper_LAS_extrabytes_implementation_in_RIEGL_Software_2019-04-15.pdf)
- [25] P. Rieger, "Range ambiguity resolution technique applying pulse-position modulation in time-of-flight scanning LiDAR applications," *Opt. Eng.*, vol. 53, no. 6, 2014, Art. no. 061614.
- [26] S. Hancock et al., "Waveform LiDAR over vegetation: An evaluation of inversion methods for estimating return energy," *Remote Sens. Environ.*, vol. 164, pp. 208–224, Jul. 2015, doi: [10.1016/j.rse.2015.04.013](https://doi.org/10.1016/j.rse.2015.04.013).
- [27] A. Chauve, C. Mallet, F. Bretar, S. Durrieu, and M. Descelligny, "Processing full-waveform LiDAR data: Modelling raw signals," in *Proc. Int. Soc. Photogrammetry Remote Sens. Workshop Laser Scanning SilviLaser*, 2007, pp. 102–107.
- [28] T.-A. Teo and W.-Y. Yeh, "The benefit of the geospatial-related waveforms analysis to extract weak laser pulses," *Remote Sens.*, vol. 10, no. 7, Jul. 2018, Art. no. 1141, doi: [10.3390/rs10071141](https://doi.org/10.3390/rs10071141).
- [29] A. Vain, X. Yu, S. Kaasalainen, and J. Hyypä, "Correcting airborne laser scanning intensity data for automatic gain control effect," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 3, pp. 511–514, Jul. 2010.
- [30] X. Li, L. Ma, and L. Xu, "Empirical modeling for non-Lambertian reflectance based on full-waveform laser detection," *Opt. Eng.*, vol. 52, no. 11, Nov. 2013, Art. no. 116110, doi: [10.1117/1.OE.52.11.116110](https://doi.org/10.1117/1.OE.52.11.116110).
- [31] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *Int. Soc. Photogrammetry Remote Sens. J. Photogramm. Remote Sens.*, vol. 60, no. 2, pp. 100–112, Apr. 2006, doi: [10.1016/j.isprs.2005.12.001](https://doi.org/10.1016/j.isprs.2005.12.001).
- [32] F. de Boissieu, E. Chraïbi, C. Lavalley, and J.-B. Féret, "pytools4dart: Python API to DART radiative transfer simulator," 2019. Accessed: Feb. 12, 2019. [Online]. Available: <https://gitlab.com/pytools4dart/pytools4dart>
- [33] F. de Boissieu, "gdecomp: Gaussian decomposition of a LiDAR waveform," 2020. Accessed: Mar. 28, 2022. [Online]. Available: <https://gitlab.com/pytools4dart/gdecomp>
- [34] T. Yin, J. Qi, J.-P. Gastellu-Etchegorry, S. Wei, B. D. Cook, and D. C. Morton, "Gaussian decomposition of LiDAR waveform data simulated by dart," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2018, pp. 4300–4303, doi: [10.1109/IGARSS.2018.8518945](https://doi.org/10.1109/IGARSS.2018.8518945).
- [35] "Wavefront OBJ file format," Jan. 23, 2020. Accessed: Feb. 12, 2023. [Online]. Available: <https://www.loc.gov/preservation/digital/formats/fdd/fdd000507.shtml>
- [36] G. Vincent and F. Heuschmidt, "Rapport CNES Leaf-ExpVal IRD AMAP," IRD, Montpellier, France, Tech. Rep., Apr. 2019.
- [37] J. Hackenberg, H. Spiecker, K. Calders, M. Disney, and P. Raumanen, "SimpleTree — An efficient open source tool to build tree models from TLS clouds," *Forests*, vol. 6, no. 11, pp. 4245–4294, Nov. 2015, doi: [10.3390/f6114245](https://doi.org/10.3390/f6114245).
- [38] P. Verley, J. Heurtebize, and G. Vincent, "AMAPVox: LiDAR data voxelisation software," Version 1.1.6, 2019. Accessed: Mar. 28, 2022. [Online]. Available: <https://amap-dev.cirad.fr/projects/amapvox/wiki>
- [39] M. Aubry-Kientz et al., "Multisensor data fusion for improved segmentation of individual tree crowns in dense tropical forests," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3927–3936, 2021.
- [40] J.-B. Féret, A. A. Gitelson, S. D. Noble, and S. Jacquemoud, "PROSPECT-D: Towards modeling leaf optical properties through a complete life-cycle," *Remote Sens. Environ.*, vol. 193, pp. 204–215, May 2017, doi: [10.1016/j.rse.2017.03.004](https://doi.org/10.1016/j.rse.2017.03.004).



- [41] I. Moradi, P. Arkin, R. Ferraro, P. Eriksson, and E. Fetzer, "Diurnal variation of tropospheric relative humidity in tropical regions," *Atmospheric Chem. Phys.*, vol. 16, no. 11, pp. 6913–6929, Jun. 2016, doi: [10.5194/acp-16-6913-2016](#).
- [42] G. Vincent et al., "Multi-sensor airborne LiDAR requires intercalibration for consistent estimation of light attenuation and plant area density," *Remote Sens. Environ.*, vol. 286, Mar. 2023, Art. no. 113442, doi: [10.1016/j.rse.2022.113442](#).
- [43] D. M. Ebengo, F. de Boissieu, G. Vincent, C. Weber, and J.-B. Féret, "Simulating imaging spectroscopy in tropical forest with 3D radiative transfer modeling," *Remote Sens.*, vol. 13, no. 11, Jan. 2021, Art. no. 2120, doi: [10.3390/rs13112120](#).
- [44] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," in *Proc. 31st Int. Conf. Neural Inform. Process. Syst.*, Red Hook, NY, USA, Dec. 2017, pp. 5105–5114.
- [45] L. Landrieu and M. Simonovsky, "Large-scale point cloud semantic segmentation with superpoint graphs," in *IEEE/CVF Proc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4558–4567, doi: [10.1109/CVPR.2018.00479](#).
- [46] Q. Hu et al., "RandLA-Net: Efficient semantic segmentation of large-scale point clouds," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11105–11114, doi: [10.1109/CVPR42600.2020.01112](#).
- [47] S. Krisanski et al., "Forest structural complexity tool—An open source, fully-automated tool for measuring forest point clouds," *Remote Sens.*, vol. 13, no. 22, Jan. 2021, Art. no. 4677, doi: [10.3390/rs13224677](#).
- [48] F. Pimont, M. Soma, and J.-L. Dupuy, "Accounting for wood, foliage properties, and laser effective footprint in estimations of leaf area density from multiview-LiDAR data," *Remote Sens.*, vol. 11, no. 13, Jan. 2019, Art. no. 1580, doi: [10.3390/rs11131580](#).
- [49] M. Soma, F. Pimont, D. Allard, R. Fournier, and J.-L. Dupuy, "Mitigating occlusion effects in leaf area density estimates from terrestrial LiDAR through A SPECIFIC KRIGING METHOD," *Remote Sens. Environ.*, vol. 245, Aug. 2020, Art. no. 111836, doi: [10.1016/j.rse.2020.111836](#).
- [50] G. Vincent, F. Pimont, and P. Verley, "Various PAD/LAD estimators implemented in AMAPVox 1.8," *DataSuds*, Nov. 23, 2022, doi: [10.23708/1AJNMP](#).
- [51] L. Terryn et al., "Quantifying tropical forest structure through terrestrial and UAV laser scanning fusion in Australian rainforests," *Remote Sens. Environ.*, vol. 271, Mar. 2022, Art. no. 112912, doi: [10.1016/j.rse.2022.112912](#).
- [52] M. B. Vicari, J. Pisek, and M. Disney, "New estimates of leaf angle distribution from terrestrial LiDAR: Comparison with measured and modelled estimates from nine broadleaf tree species," *Agricultural Forest Meteorol.*, vol. 264, pp. 322–333, Jan. 2019, doi: [10.1016/j.agrformet.2018.10.021](#).
- [53] K. Parkitna et al., "Modelling growing stock volume of forest stands with various ALS area-based approaches," *Forestry, Int. J. Forest Res.*, vol. 94, no. 5, pp. 630–650, Dec. 2021, doi: [10.1093/forestry/cpab011](#).
- [54] P. Tompalski, J. C. White, N. C. Coops, and M. A. Wulder, "Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data," *Remote Sens. Environ.*, vol. 227, pp. 110–124, Jun. 2019, doi: [10.1016/j.rse.2019.04.006](#).
- [55] D. R. A. de Almeida et al., "Optimizing the remote detection of tropical rainforest structure with airborne LiDAR: Leaf area profile sensitivity to pulse density and spatial sampling," *Remote Sens.*, vol. 11, no. 1, Jan. 2019, Art. no. 92, doi: [10.3390/rs11010092](#).
- [56] P. Packalen, J. Strunk, T. Packalen, M. Maltamo, and L. Mehtätalo, "Resolution dependence in an area-based approach to forest inventory with airborne laser scanning," *Remote Sens. Environ.*, vol. 224, pp. 192–201, Apr. 2019, doi: [10.1016/j.rse.2019.01.022](#).
- [57] O. R. van Lier, J. E. Luther, J. C. White, R. A. Fournier, and J.-F. Côté, "Effect of scan angle on ALS metrics and area-based predictions of forest attributes for balsam fir dominated stands," *Forestry*, vol. 95, no. 1, pp. 49–72, Jan. 2022, doi: [10.1093/forestry/cpab029](#).
- [58] R. Dubayah et al., "The global ecosystem dynamics investigation: High-resolution laser ranging of the Earth's forests and topography," *Sci. Remote Sens.*, vol. 1, Jun. 2020, Art. no. 100002, doi: [10.1016/j.srs.2020.100002](#).
- [59] S. Daisuke, N. T. Trung, M. Rei, S. Yoshito, I. Tadashi, and K. Toshiyoshi, "Progress of the ISS based vegetation LiDAR mission, Moli - Japan's first space-based LiDAR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 3467–3470, doi: [10.1109/IGARSS39084.2020.9323332](#).
- [60] S. Durrieu et al., "Preliminary studies for a vegetation lidar/LiDAR space mission in France," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2013, pp. 4332–4335, doi: [10.1109/IGARSS.2013.6723793](#).
- [61] S. Durrieu et al., "Waveform simulations and analysis of LiDAR data for the development of a space-based vegetation LiDAR system," in *Proc. Int. Workshop Vegetation LiDAR Appl. Space*, Jan. 2016, p. 40.
- [62] S. Hancock et al., "The GEDI simulator: A large-footprint waveform LiDAR simulator for calibration and validation of spaceborne missions," *Earth Space Sci.*, vol. 6, no. 2, pp. 294–310, 2019, doi: [10.1029/2018EA000506](#).
- [63] J.-R. Roussel et al., "lidR: An R package for analysis of airborne laser scanning (ALS) data," *Remote Sens. Environ.*, vol. 251, Dec. 2020, Art. no. 112061, doi: [10.1016/j.rse.2020.112061](#).



**Florian de Boissieu** received the M.Sc. degree in physics and signal processing and the Ph.D. degree in signal processing from Grenoble Institut National Polytechnique, Grenoble, France, in 2006 and 2010, respectively.

From 2011 to 2021, he worked as a Postdoc on processing methods for optical remote sensing of environment, on vegetation, geology, and ocean color, for CNRS, IRD, INRAE, and CIRAD in New Caledonia and France. Since 2022, he has been affiliated with the National Institute of Agriculture, Food and Environment (INRAE), and works with TETIS research unit, Montpellier, France. His current research interests include the development of processing methods and softwares for lidar data and multi-/hyperspectral imaging of vegetation, as well as associated radiative transfer modeling.



**Florence Heuschmidt** received the graduate degree in agronomist engineer from VETAGRO SUP School, Lempdes, France, in 2018.

She evaluated several wood/leaf classification methods and tree reconstruction methods in tropical forests as part of her end-of-study internship in 2018. She did forestry specialization from BORDEAUX SCIENCES AGRO School. The internship project was carried out within UMR AMAP located in Montpellier. She worked for CNRS using this preliminary study to generate a 3-D mockup built from terrestrial lidar scan using tropical forest field data. Between 2019 and 2022 she was working at "Groupe Coopération Forestière" as a forest Engineer.



**Nicolas Lauret** received the master's degree in computer science and digital image engineering, and the Ph.D. degree, both from University Toulouse III Paul Sabatier, Toulouse, France, in 2008 and 2012, respectively.

Since 2008, he has been working with CESBIO Laboratory, Toulouse, as a developer of the three-dimensional radiative transfer model DART. Since 2022, he has affiliated with the French National Centre for Scientific Research (CNRS), Montpellier, France.



**Dav M. Ebengo** received the agronomist engineer diploma, in water and forest from the University of Kinshasa, Kinshasa, Democratic Republic of the Congo, in 2014, the M.Sc. degree in soil sciences from the University Paris-Saclay, Paris, France, in 2017, and the Ph.D. degree in geomatics from AgroParis-Tech, Montpellier, France, in 2021.

His research interests include radiative transfer modeling, remote sensing of plant ecophysiology, and biodiversity mapping.



**Grégoire Vincent** received the M.Sc. degree in agronomy from the Institut National Polytechnique de Lorraine, Vandoeuvre les Nancy, France, in 1989, and the Ph.D. degree in systems modeling from Université de Lyon, Lyon, France, in 1993.

He is a Forest Ecologist. His current research interests include leveraging new remote sensing technologies to improve our understanding of tropical forest response to ongoing climate change.



**Jean-Baptiste Féret** received the diploma degree in agricultural engineering from Montpellier SupAgro, Montpellier, France, in 2005, and the Ph.D. degree in environmental sciences from Université Pierre et Marie Curie, Paris, France, in 2009.

He worked with the Carnegie Institution for Science, at the Department of Global Ecology, Stanford, CA, USA, from 2010 to 2013 in Asner Lab. He worked for CNES at CESBIO lab in Toulouse, France, in 2014, in the context of the preparation of applications in biodiversity mapping using future

spaceborne imaging spectrometers using 3-D radiative transfer modeling. He is affiliated with the National Research Institute of Agriculture and Environment (INRAE) since 2014, and works at TETIS research unit, Montpellier, France. His research interests include optical remote sensing and radiative transfer modeling of vegetation, vegetation traits monitoring, and biodiversity mapping.



**Tiangang Yin** received the B.Sc. degree in physics and the M.Eng. degree in computer engineering from National University of Singapore, Singapore, in 2008 and 2011, respectively, and the Ph.D. degree in remote sensing and geoscience from the Centre d'Etudes Spatiales de la Biosphère (CESBIO), Paul Sabatier University, Toulouse, France, in 2015.

He is an Assistant Professor with the Department of Land Surveying and Geo-Informatics (LSGI), The Hong Kong Polytechnic University, Hong Kong. Prior to joining LSGI, was a research scientist jointly at

NASA Goddard Space Flight Center (GSFC) and the Earth System Science Interdisciplinary Center of the University of Maryland. Since 2012, he has been leading research and development of several modules (lidar, photogrammetry, urban atmosphere) for the discrete anisotropic radiative transfer (DART) model.



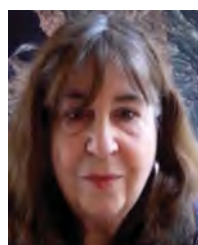
**Jean-Philippe Gastellu-Etchegorry** (Member, IEEE) received the B.Sc. degree in electricity from ENSEEIHT, Toulouse, France, in 1978, the Aggregation degree in physics from Paris 6 University, Paris, France, in 1981, and the Ph.D. degrees in solar physics and remote sensing from Toulouse III University (UT3), Toulouse, in 1983 and 1989, respectively.

He was with the PUSPICS-Remote Sensing Center, Gadjah Mada University, Yogyakarta, Indonesia, and Southeast Asian Ministers of

Education Organization, Bangkok, Thailand, from 1984 to 1990. He was with the Department of Remote Sensing, Centre d'Etude Spatiale des Rayonnements, Toulouse, and the Centre d'Etudes Spatiales de la Biosphère, Toulouse, in 1995. He is currently a Professor with UT3. Since 1993, he has been developing the discrete anisotropic radiative transfer (DART) model (<https://dart.omp.eu>). His research interests include 3-D radiative transfer modeling.

**Josiane Costeraste** received the Ph.D. degree in physics from the National Institute of Applied Sciences, Toulouse, France, in 1986.

She worked with CNES as an Instrument Engineer on various Earth observation systems and then as Head of the optical instruments department. Since 2011, she has been in charge of Earth Observation pre-phase A studies.



**Marie-José Lefèvre-Fonollosa** received the Doctor of Geophysics degree, in 1976, and the Ph.D. degree in space geosciences, in 1983.

She worked as a Researcher in the BRGM before to join the CNES in Toulouse, where she headed the Research & Technology Programme for Earth Observation and Universe Sciences. Senior Expert in Satellite Systems since 2006, she was in charge of preliminary phases of Earth Observation satellites which prepared the french space projects; she is retired from CNES since 2019.



**Sylvie Durrieu** received the M.Sc. degree in forest and plant biology from the University of Nancy, Nancy, France, in 1989, and the engineer degree and the Ph.D. degree in remote sensing from the French Institute of Forestry, Agricultural and Environmental Engineering (ENGREF-AgroParisTech), Paris, France, in 1989 and 1994, respectively.

She received an accreditation to supervise research in Earth and Water Sciences from the University of Montpellier (ED N° 584) in 2018. From 1994 to 1999 she worked with the French National Forest Inventory (IFN) and carried out R&D in the field of remote sensing applied to forest inventory and mapping, focusing on optical imagery technology. Since 1999 she has been working with the INRAE, Montpellier, France. Her research interests include lidar technology for forest ecosystem monitoring. She has developed methods for characterizing and monitoring forest resources and biodiversity from 3-D lidar data. She was PI of LEAF (Lidar for Earth And Forests), a project of space lidar mission submitted to ESA in 2008, and supported by CNES since 2013 as a potential mid-term space mission.