HyperHeight LiDAR Compressive Sampling and Machine Learning Reconstruction of Forested Landscapes

Andres Ramirez-Jaime, *Graduate Student Member, IEEE*, Karelia Pena-Pena, Gonzalo R. Arce, *Life Fellow, IEEE*, David Harding, Mark Stephen, and James MacKinnon

***Abstract*— Light detection and ranging (LiDAR) remote sensing systems are deployed in various platforms including satellites, airplanes, and drones—which, in essence, determines the sampling characteristics of the underlying imaging system. Low-altitude LiDARs provide high photon count and high spatial resolution but only in very localized patches. Satellite LiDARs, on the other hand, provide measurements at a global scale but are limited by low photon count and their samples are sparsely apart along swath line trajectories that are far in between. This article describes a new class of satellite remote sensing LiDARs, aimed at overcoming the limitations of current satellite imaging systems. It exploits the principles of compressive sensing and machine learning (ML) to compressively sense Earth from hundreds of kilometers above Earth to then reconstruct the 3-D imagery with resolution and coverage, as if the data were collected from airborne platforms at just hundreds of meters in height. We introduce a novel representation of waveform altimetry profiles, coined hyperheight data cubes (HHDCs), which encompass rich information about the 3-D structure of a scene. Canopy height models (CHMs), digital terrain models (DTMs), and many other features of a scene that are embedded in HHDC are easily extracted with simple statistical quantiles. We introduce ML methods to reconstruct the compressive LiDAR measurements so as to attain high-resolution, dense coverage, and broad field-of-view per swath pass. ML training data are attained from NASA's G-LiHT imaging missions. Simulations with various types of forests across the US illustrate the power of the new LiDAR imaging systems.**

***Index Terms*— Canopy height model (CHM), compressive sampling, digital terrain model (DTM), light detection and ranging (LiDAR), machine learning (ML).**

I. Introduction

LIGHT detection and ranging (LiDAR) systems provide precise laser measurements of the distance between the sensor and objects of interest [1]. In remote sensing, they provide unique capabilities in collecting high-resolution elevation data that are unmatched by any other remote sensing technique [2]. Depending on the method for detecting the return signals, LiDARs provide three types of data. Digitized waveforms of the received laser energy, point clouds of discrete returns where the signal exceeds a threshold, or photon-counting point clouds composed of individually detected photons [3]. The resulting data products are used for scientific discovery and applied mapping purposes. In forested landscapes, for instance, they simultaneously measure ground topography beneath vegetation cover and characterize forest canopy structure including the spatial organization within the canopy and its height [4], [5]. This vegetation structural information is used for a variety of purposes, including estimation of the above-ground stored carbon, establishing how sunlight is intercepted by foliage, modeling the exchange of carbon with the atmosphere and characterizing forest habitats and their relationship to biodiversity, and the topography data are also used for many purposes, including analysis of landform evolution, hydrologic routing and flood modeling, landslide forecasting, and characterization of surface-rupturing earthquakes [3]. In glaciers, they measure sea ice thickness distribution [6], and in coastal and beach areas they enable bathymetry [7], [8], [9]. LiDARs are also used in computer vision to create 3-D images of scenes in a wide variety of applications such as autonomous driving [10], [11], [12] and robotics [13], [14].

While the underlying principles of LiDARs used in all these applications are similar, the return signals acquired differ significantly depending on the range of the target data. Short-range measurements collected by drones or unmanned aerial systems (UAS) provide very high-resolution (cm-scale) 3-D point cloud mapping [10], [11], [12]. Moderate resolution (m-scale) is feasible for larger areas up to entire countries using fixed-wing aircraft. NASA's Goddard LiDAR, Hyperspectral, Thermal (G-LiHT) mission is one such example where data are collected at 300 m, providing submeter resolution point clouds across a dozen regions of the US and Mexico spanning over 6500 km^2 [15]. For long-range remote sensing, NASA has conducted global LiDAR Earth observations from space in a series of missions with three waveform-recording systems, the shuttle laser altimeter (SLA) [16], the Ice, Cloud, and Land Elevation Satellite (ICESat) [17], and the Global Ecosystem Dynamics Investigation (GEDI) hosted on the International Space Station [18], and a photon-counting system, the second ICESat mission (ICESat-2) [19].

Manuscript received 23 January 2023; revised 26 October 2023; accepted 3 January 2024. Date of publication 29 January 2024; date of current version 31 January 2024. This work was supported in part by the National Aeronautics and Space Administration (NASA) under Grant 80NSSC19M0087 and Grant 80NSSC22K1099. *(Corresponding author: Andres Ramirez-Jaime.)*

Andres Ramirez-Jaime, Karelia Pena-Pena, and Gonzalo R. Arce are with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: aramjai@udel.edu).

David Harding, Mark Stephen, and James MacKinnon are with NASA's Goddard Space Flight Center (GSFC), Greenbelt, MD 20771 USA.

Digital Object Identifier 10.1109/TGRS.2024.3356389

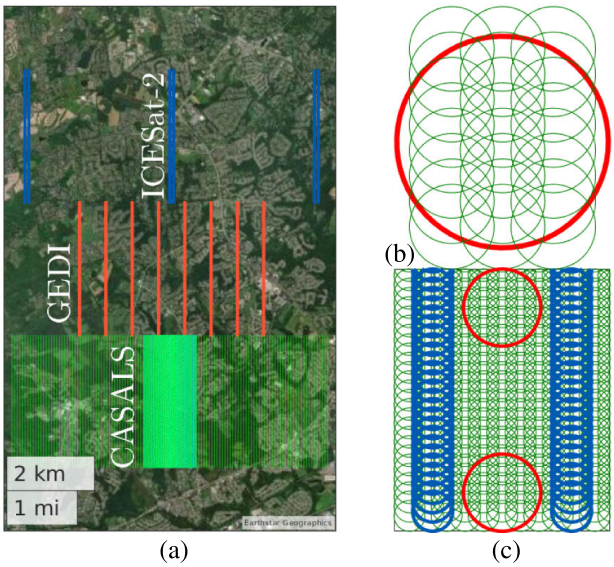

Fig. 1. Comparison of ICESat-2 (blue), GEDI (red), and CASALS (green) laser footprints showing (a) ICESat-2 and GEDI 2-D profiling along an orbit track compared with CASALS 3-D swath mapping, (b) overlapping CASALS footprints configured to emulated GEDI, and (c) highest density CASALS footprint configuration which achieves increased density along and across track sampling compared with the GEDI and ICESat-2 profiles.

Although these spaceflight missions acquire measurements at a global scale, thus providing a complete representation of the planet Earth's surface, the area covered is sparse. The data are only acquired along 2-D profiles and the 3-D spatial resolution is low for a variety of reasons. The number of profiles is limited, with one for SLA and ICESat, multiple laser beams forming six profiles for ICESat-2, and eight for GEDI, and the cross-track spacing between profiles is large, as illustrated in Fig. 1(a). The configuration of the ICESat-2 profiles is designed to determine ice sheet elevation change, and the GEDI configuration is designed to uniformly sample vegetation structure. The laser footprints are large (varying between 100 m for SLA and 10 m for ICESat-2), and the sampling along-track is discontinuous for the waveform systems (varying between 700 m for SLA and 30 m for GEDI) because high laser pulse energy is required for waveforms imposing a low pulse rate due to limited spacecraft power. ICESat-2 using low-power micropulses has overlapping footprints with 0.7-m spacing but only several photons per pulse are detected for vegetated landscapes requiring the combination of photons over tens of meters to have data comparable to waveform systems. Furthermore, unlike UAS and aircraft which can repeatedly observe locations many times even within a single day, spaceflight LiDAR temporal sampling is low limited by the repeat cycle of the spacecraft which is typically on the order of weeks or months.

Fig. 2 illustrates the vastly different 3-D imaging capabilities provided by the ICESat-2 and G-LiHT missions. While G-LiHT provides dense high-resolution and high-photon count 3-D imagery, ICESat-2 only offers low-resolution line profiles with very low photon counts. The figure depicts the ICESat-2 line profiles acquired over multiple swath paths of the satellite in the period of months. In between the ICESat-2 line profiles, vast amounts of landscape remain without sampling illumination. It is clear that Earth science in general would benefit greatly if satellite laser altimetry systems could provide 3-D imagery resolution and coverage comparable to that offered by G-LiHT type systems. To this end, the National Academies in the Earth Science Decadal Survey for future laser altimetry recommended to drastically improve LiDAR efficiency, enabling more coverage, higher resolution, with the use of smaller, cost-effective satellite platforms [20].

To address this need, new LiDAR imaging systems are being proposed. One such sensor is the adaptive wavelength scanning LiDAR (AWSL) [21] which is a part of the concurrent artificially intelligent spectrometry and adaptive LiDAR system (CASALS) payload being developed at Goddard Space Flight Center for a spaceflight mission [22]. CASALS will provide measurements supporting scientific studies and societal applications related to the carbon cycle and ecosystems, cryosphere response to climate change, natural hazards, and atmospheric clouds and aerosols. The AWSL uses a novel method to scan a beam to specified locations across 7 km using a photonic integrated circuit seed laser, wavelength tuning circuitry, a high-power fiber amplifier, and a wavelength-to-angle dispersive grating. The method achieves beam scanning with no mechanical parts, removing a risk for spaceflight use. By rapidly changing the laser wavelength and pulse rate, an essentially limitless pattern of laser footprint locations and spacings can be produced. For example, the ICESat-2 and GEDI profile patterns can be created for data continuity purposes or profiles can be closely spaced to map a swath. For swath mapping, the AWSL footprints planned to be 10 m could be spaced by 3 m along and across track to map a 1-km wide swath [bright green swath in Fig. 1(a)]. They can even be further concentrated in overlapping footprints for detailed mapping in a swath as wide as 200 m, or they can be distributed in a sparse pattern covering a much broader area [light green swath in Fig. 1(a)]. Fig. 1(b) and (c) shows a dense sampling of overlapping footprints planned for the CASALS LiDAR in comparison to GEDI and ICESat-2 footprints, respectively. The AWSL receiver uses a detector array to image the swath which has single-photon sensitivity and analog output which will be digitized at high speed to provide waveforms. Approximately 20 detected photons are expected per laser pulse for vegetation, $\sim 10\times$ that of ICESat-2. CASALS will include high-resolution multispectral imaging to characterize the composition, physical state, and biophysical properties of the locations being targeted for fusion with the LiDAR to enhance the science and applications utility of the observations.

The adaptive capability of the CASALS LiDAR makes it very well suited to use sparse sampling patterns optimized for the topographic and vegetation cover attributes of a landscape and for the objective for which the data are being acquired. A fundamental constraint will be the power provided by the spacecraft to CASALS which imposes a limit on the maximum achievable laser pulse rate. Therefore, a key goal is to establish how sparsely a landscape can be sampled while still providing the information necessary to achieve the mission objectives. By doing so, wider areas can be mapped within the power provided. The LiDAR sensing mechanism

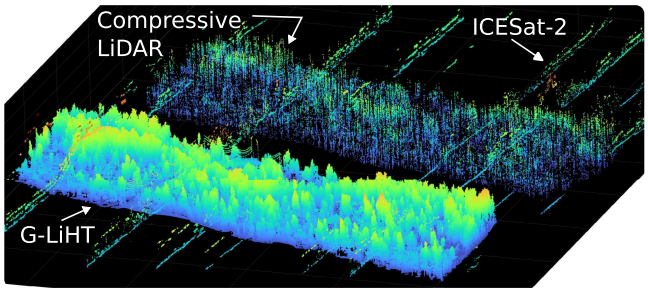

Fig. 2. LiDAR sampling strategies: dense coverage and high-resolution low-altitude G-LiHT, line profiles of ICESat-2, and wide field-of-view compressive sampling.

can thus be posed as a compressive sensing problem, and in this study, we explore sampling pattern designs and develop an inverse imaging approach based on machine learning (ML) to assess the accuracy with which forest structures can be reconstructed at resolutions higher than the sparse footprint sampling. The focus is on forest structure because it is one of the most challenging landscapes in which to achieve 3-D swath mapping at resolutions necessary for ecologic studies. Fig. 2 illustrates the underlying objective of this work where the Earth's landscape is compressively sampled from high satellite altitudes using wide field-of-view beam-steering. The compressive samples are then used to reconstruct a 3-D image representation of the scene as if we had imaged it with LiDAR sensors used from only a few hundred meters above Earth (G-LiHT in Fig. 2). The sensing profiles of ICESat-2 are also shown in Fig. 2 as a reference. The sensing and inverse problem mechanisms needed to accomplish these objectives are formulated in Sections IV and V.

Currently, a variety of strategies are available for tackling the inverse problem that arises in compressed sensing scenarios. While there are optimization-based methods that can address this problem without requiring data, this study focuses on solutions that leverage deep learning techniques. This choice is primarily motivated by the remarkable performance deep learning has exhibited across multiple disciplines, making it a logical step to extend these approaches to new applications. Artificial intelligence is not a new concept in the field of remote sensing however. Neural networks, for instance, have been widely used for tasks such as hyperspectral image segmentation and classification [23], [24], detecting anomalous events in satellite imagery [25], recognizing targets using synthetic aperture radar (SAR) sensors [26], classifying LiDAR point clouds [27], and even fusing data from multiple sensors in multimodal systems [28]. While artificial intelligence has been successfully used for reconstructing RGB images and LiDAR point clouds, its application in the context of satellite LiDAR, especially with sparse measurements, is relatively unexplored. Although there have been notable studies such as those referenced in [29] and [30] focusing on reconstructing images in forested areas using airborne systems, to the best of our knowledge, there is no prior work addressing the significant challenge of enhancing the resolution of satellite LiDAR data to improve operational efficiency.

To characterize LiDAR systems capable of 3-D swath mapping in preparation for sending them into orbit, we first synthesize and emulate the underlying sensing mechanisms. The source data for the synthesis are the G-LiHT airborne sensor suite which acquires very high-density discrete-return point clouds [15], along with hyperspectral and thermal imaging. The G-LiHT LiDAR typical acquires long, 100-m wide swaths, along many locations across North America. The points are processed to synthesize an array of large circular footprints to represent 3-D waveform scenes. This way of presenting the output of a 3-D LiDAR as a volumetric waveform product is coined here as a *hyperheight data cube* (HHDC) due to its resemblance with a hyperspectral data cube.

The main contributions of this article are as follows.

1) We introduced a novel tensor-based representation coined HHDC for LIDAR systems, encompassing comprehensive 3-D information pertinent to ecological studies, including canopy profiles, digital terrain models (DTMs), etc. This representation exhibits the requisite flexibility to facilitate signal processing and ML applications.
2) We demonstrated the compressibility of this innovative representation through a wavelet study, thereby paving the way for the potential implementation of compressed sensing systems using artificial intelligence methodologies.
3) We presented a comprehensive methodology, enabling the derivation of all the traditional ecological representations via statistical analyses of the proposed tensor. This endeavor serves to bridge the gap between prior studies and the novel products to be delivered through satellite LiDAR systems.
4) We illustrated the solution of the inverse problem associated with HHDC, aiming to bridge the disparity between satellite measurements and airborne representations through the utilization of deep learning techniques.
5) We illustrated the influence of distinct sampling patterns on the quality of neural network reconstructions, both when the sampling percentage remains constant and when it changes.

The rest of this article is organized as follows: Section II presents the methodology for synthesizing large footprint LiDAR measurements acquired by waveform recording detectors. Section III introduces HHDCs and presents different ways of interpreting them. Section IV describes the physical attenuation characteristics of the photons as they travel from low altitudes to heights of satellite orbits. Section V casts the satellite LiDAR imaging system as a compressive sensing problem. Section VI presents the use of ML techniques to reconstruct high-resolution LiDAR HHDC from sparsely sampled data. Section VII presents the results of the ML system and their implication for satellite LiDAR systems. Finally, Section VIII presents the concluding remarks.

II. Waveform LiDARs

NASA has conducted several LiDAR missions, including GEDI and ICESat-2, using waveform and photon-counting LiDARs in which the sensors do not record 3-D atomized point clouds but rather they record large illuminated areas

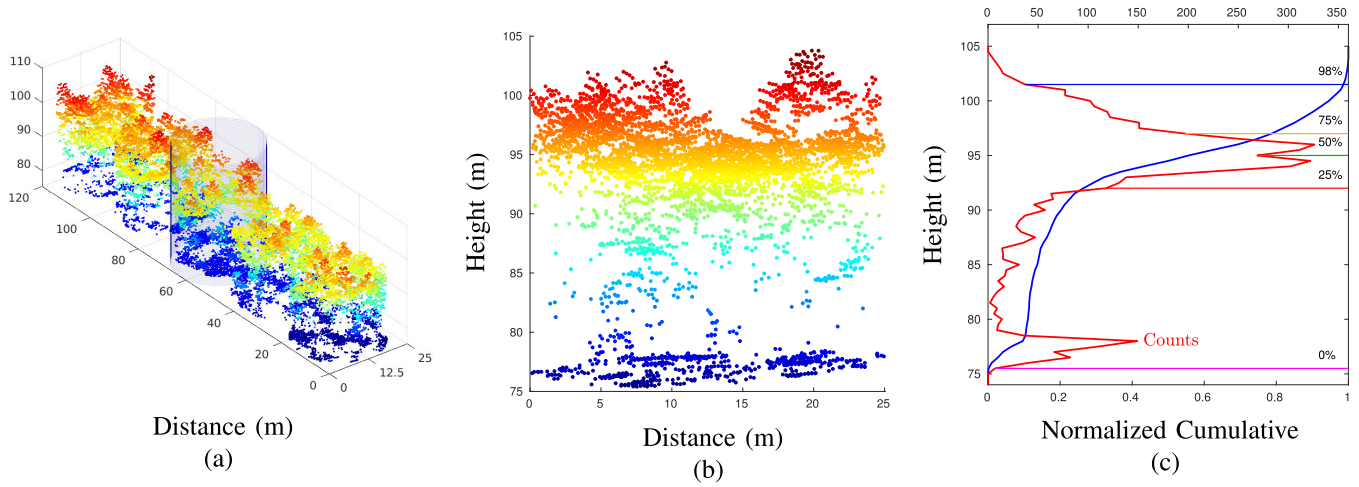

Fig. 3. (a) Collection of point clouds in individual 25-m footprints arranged as a swath. (b) Side view of a single footprint so the individual photons can be identified. (c) Histogram height distribution of the points of the footprint (red) and the cumulative distribution (blue) where five key percentiles can be seen. Note that the z-axis represents the height above sea level.

and generate a continuous or discrete waveform that describes the structure underneath. This contrasts with the traditional way of interpreting LiDAR images as creating a 3-D representation of the environment as seen in the scanning LiDARs of automobiles. To better understand the sensing phenomena in waveform LiDARs, consider the point cloud presented in Fig. 3(a). This point cloud was acquired in a forested area in Maine using NASA's airborne G-LiHT system using a scanning LiDAR and it shows the point clouds inside multiple 25-m-diameter cylinder footprints of a forest in Maine, USA. From this figure, the structure of trees in the scene can be seen as the resolution inside the volume of a 25-m footprint offers thousands of individual points (photons) in the point cloud. To understand the output of a discrete waveform LiDAR, it is needful to describe the height distribution of the points found within a single illuminated area. Fig. 3(b) shows a side view of the point cloud inside of a single 25-m footprint so that every photon is clearly visible within the tree structure, and in Fig. 3(c), the equivalent height distribution histogram is created using the heights of the individual points seen in the side view. Even though the individual photon location is lost while creating the histogram, all the key features used for the canopy study are still retained. Alongside the height histogram, some key percentiles are shown; these percentiles carry important information about the topographical structure, the health of the forest, ecological variables, among others. This height distribution is equivalent to the output of a photon-counting LiDAR illuminating the same 25-m region.

The phenomenon that is not modeled directly by simply constructing the histogram of heights is the Gaussian nature of the pulses sent by the laser in both time and space. To account for this, we leverage the way full-waveform LiDARs are emulated to then perform a discretization of the resulting signal. Following [18], the waveform returned to the detector can be modeled by summing the convolutions between the intensity of each point in the footprint and the Gaussian signal produced by the laser. Let the pulse p emitted by the laser be given by

$$p(z - z_i) = \frac{1}{\sigma_p \sqrt{2\pi}} e^{-\frac{(z-z_i)^2}{2\sigma_p^2}} \tag{1}$$

where σ_p is given by the Gaussian shape of the laser signal. The collected waveform is then described by

$$I(z) = \sum_{i}^{N} I_{w,i} * p(z - z_i) \tag{2}$$

where N represents the total number of points in the footprint, I_w represents the intensity of the points, and p represents the Gaussian signal of the laser. It is important to note that the intensity of the points decays in a Gaussian manner with respect to the center of the footprint according to

$$I_{w,i} = \frac{1}{\sigma_f \sqrt{2\pi}} e^{-\frac{(x_i - x_0)^2 + (y_i - y_0)^2}{2\sigma_f^2}} \tag{3}$$

where σ_f is determined by the size of the footprint, x_0 and y_0 are the coordinates of the center of the footprint, and x_i and y_i are the coordinates of point i. Equation (2) describes the full-waveform LiDAR which can be discretized to provide a photon-counting LiDAR as

$$I_d(z) = I(z) \sum_{k \in \mathbb{Z}} \delta(z - k\Delta z) \tag{4}$$

where Δz is the vertical resolution of the emulated photon-counting LiDAR. The parameter Δz is partially dependent on the physics of the selected laser, as can be determined by the duration of the pulse (1), and is also dependent on the hardware specifications of the used detector [31]. This is due to the fact that it is not feasible to attain arbitrary resolution given that the detector needs a certain amount of time to process each sample, thus rendering it impossible to differentiate photons that arrive at arbitrarily small time intervals. Fig. 4 shows a comparison between the raw height histogram and the waveform obtained by modeling the Gaussian phenomenon of the laser pulse.

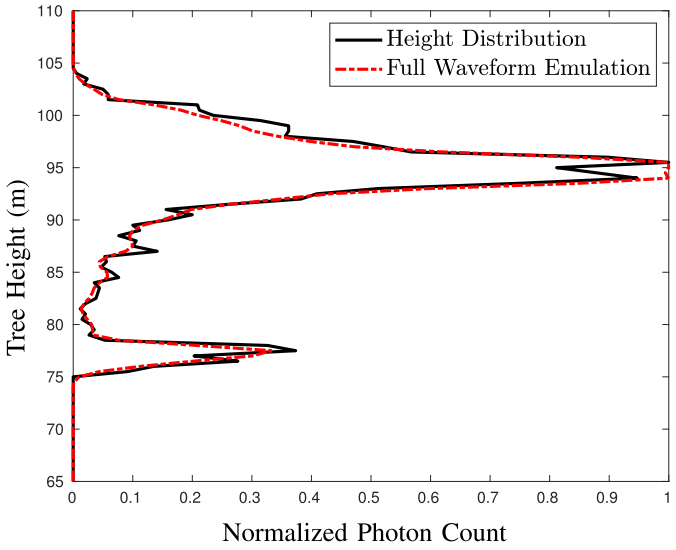

Fig. 4. Comparison between the height histogram and the continuous waveform created by the convolution of the points and the signal of the laser.

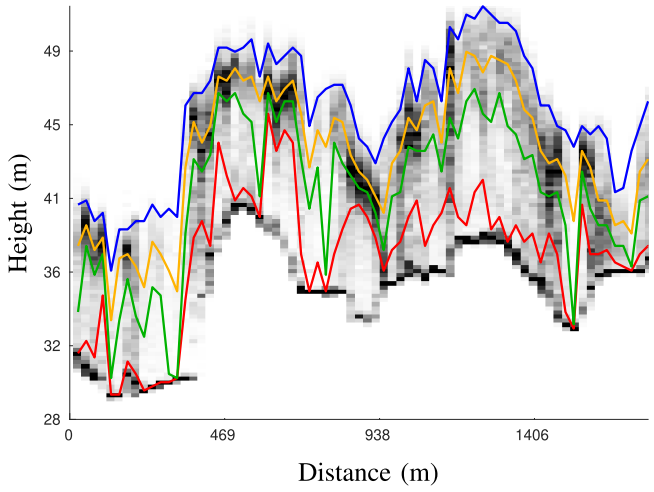

Fig. 5. CHP constructed by concatenating multiple height histograms over a swath path alongside key percentiles that allow studying ecological variables. The 98% (blue), 50% (green), and 25% (red) percentiles are shown along the CHP.

The model in (1) and (4) provides a relationship that can be used to emulate the output of a waveform LiDAR. If a small footprint LiDAR is available over a specific region, it is possible to approximate the output of a waveform LiDAR by studying the distribution of the heights of individual points within a small patch of a specific region. In this manner, footprint waveforms can be synthesized from a discrete size 25 m to match the image system under study, but this is a parameter that can be modified and helps represent in a general way the size of the footprint of the waveform LiDAR.

A. Canopy Height Profile (CHP) and HHDCs

It is common to represent a series of waveforms (histograms) concatenated side by side to study the characteristics of an entire region. For example, Fig. 3(a) depicts many 25-m-diameter areas that have been arranged together as a swath so that, as shown in Fig. 5, histograms of individual illuminated areas can be analyzed holistically. In Fig. 5, the x-axis represents the movement along the swath created in Fig. 3(a), and the grayscale represents the concentration of photons at a given height of the histogram; hence, it is easier to interpret attributes such as the structure of the topography (described by the lower part of the waveforms), or the points where the forest is the densest (described by areas where the graph has darker shades of gray). The representation shown in Fig. 5 is often referred to as a CHP [32] and is the traditional way of displaying data from NASA missions that implement waveform LiDARs. As mentioned before, some key percentiles are shown over the CHP. Particularly, the 25%, 50%, 75%, and 98% percentiles are shown in Fig. 5 since that they can be used for understanding how biomass behaves in the forest [33], [34].

Even though this way of analyzing the terrain using multiple histograms is quite useful, it lacks the ability to study not only along the length of the terrain, i.e., along the swath, but also across the width of the terrain, i.e., across the swath. To present a complete representation of the geographical area of interest, it would be necessary to show illuminated areas that are parallel to those shown in Fig. 3, so that more area is covered. Fig. 6(a) shows a G-LiHT point cloud of a forest along with the footprint array overlaid on top of it. The set of illuminated regions that will show waveforms along and across the swath. Of course, if one has this configuration of illuminated areas, the joint histogram representation is no longer a single 2-D image but rather multiple images placed next to each other that can be represented as a 3-D cube with complete terrain information. Fig. 6(b) depicts such representation by showing the terrain along and across the swath, as well as the height of the histograms and their intensity. Notably, this 3-D representation resembles a hyperspectral data cube in which multiple bands of the electromagnetic spectrum are plotted as a cube, and thus, we coin the height distributions of the canopy as an HHDC. Although the HHDC is a natural evolution of the CHP, it is more descriptive due to the fact that it has information on both the length and width of the swath.

The HHDC characterizes a 3-D space, and thus, it can be naturally represented as a three-order tensor $\mathcal{X} \in \mathbb{R}^{n\times m\times c}$ where n and m are the number of footprints across and along the swath, respectively, and c is the number of histogram bins associated with each footprint (see Fig. 7). Each waveform associated with a footprint is referred to as a tube scalar $\mathbf{x} \in \mathbb{R}^c$ [Fig. 7(b)]. The (i, j, k) entry of the tensor represents the number of photons found in the ith footprint location across the swath, the jth footprint location along the swath, and at the Δz k height, where Δz is the chosen vertical resolution. Vertical slices of HHDC, along the swath, produce canopy height planes as depicted in Fig. 5. Vertical slices along the perpendicular direction of the HHDC lead to similar across the swath canopy height planes. These multiple cuts will aid in the interpretation of the information embedded in the HHDC.

As Fig. 6(b) illustrates, HHDCs in general have well-defined structure across the spatial and height dimensions. In fact, the 3-D wavelet representation of HHDCs of natural forests and vegetation is highly sparse. Fig. 6(c) depicts the sorted magnitude and log-magnitude coefficients of a 3-D Symlets

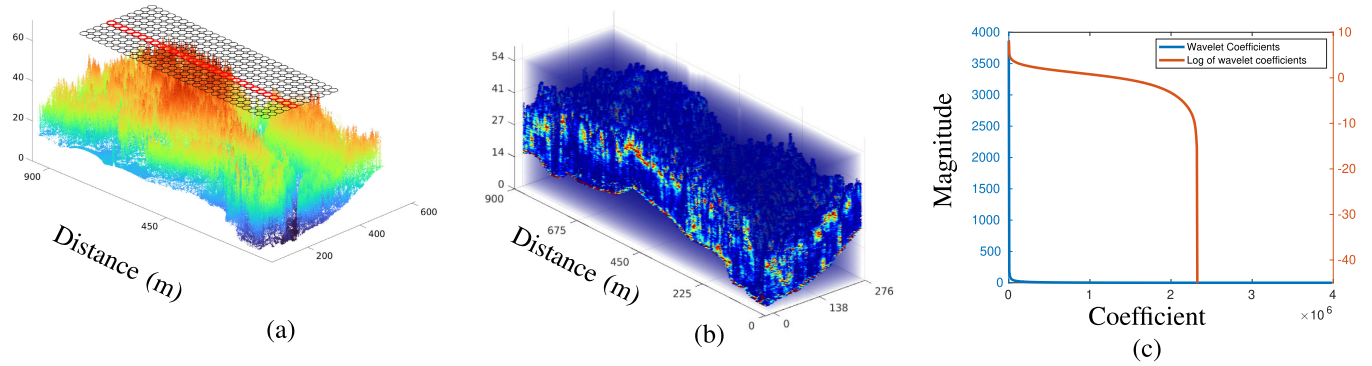

Fig. 6. (a) G-LiHT point cloud of a forest and the footprint array overlaid on top where a single swath line is highlighted. (b) Volumetric HHDC created from the array of waveform footprints. (c) Magnitude and log-magnitude of the HHDC that illustrate its sparsity in a 3-D Symlets 4 wavelet representation.

4 wavelet representation of the HHDC in Fig. 6(b). It can be seen that less than 5% of the coefficients are significant. Forests in other regions of the US such as Florida exhibit even higher sparsity. Sparsity indicates that HHDCs found in natural forestry can be concisely represented in some basis with just a very small number of coefficients since natural HHDCs exhibit correlation among adjacent pixels and also across height bins. This property is important as it opens the opportunity to design compressive sensing protocols capable of capturing the essential information content in HHDCs with just a small number of compressive measurements that sketch the 3-D information embedded in the HHDCs. The HHDCs of interest could then be accurately reconstructed from the small number of compressive measurements as it will be described shortly.

III. HHDC Interpretation

To describe the virtues of an HHDC and what it represents, consider again the high-resolution G-LiHT point cloud of a forest region shown in Fig. 6(a). From this point cloud, it is possible to construct an HHDC with 10-m footprints for which the laser shots have a separation across the swath of 6 m and along the swath of 3 m, and a vertical resolution of the histograms of 0.5 m. The HHDC tensor for this forest is shown in Fig. 6(b). Even though the information embedded in the HHDC is rich, it can be complicated to interpret. In this section, some subviews of the cube are presented to facilitate its interpretation, namely, view along the swath, view across the swath, DTM, and canopy height model (CHM).

A. View Along the Swath

A direct way to visualize an HHDC is by looking at it as a set of 2-D measurements over different lines of the terrain. For this case, what is being seen is a slice on the yz-axis of the HHDC generating a CHP, as shown in Fig. 7(c). Fig. 11(a) shows this view extracted directly from the HHDC cube.

B. Across the Swath

Given an HHDC, it is also possible to analyze what happens in the direction perpendicular to the direction of the moving detector platform. This type of visualization seeks to understand what happens in an xz slice of the cube, as shown in Fig. 11(c). It is important to note that this representation is similar to the one shown in Fig. 11(a) since both come from the same geographical region.

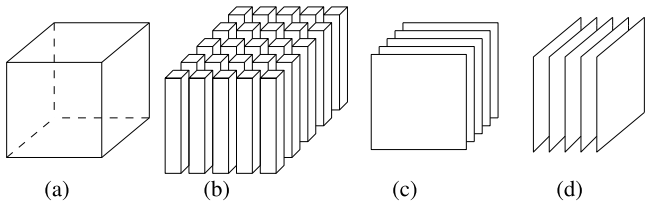

Fig. 7. (a) HHDC tensor representation divided into (b) tubes representing the information of each footprint, (c) along the swath slices, and (d) across the swath slices.

C. Digital Terrain Model

One of the traditional ways to study the output of a point cloud is to extract the well-known DTM. The DTM provides relevant information about the topography of the geographical area being studied. Remembering that the HHDC is constructed from the height distributions, it is possible to generate percentiles associated with the height distributions, which seek to analyze how the horizontal slices of the cube change. For example, in the case of DTM, the height at which all the distributions exceed a value above the noise floor (usually a value close to 2%–5%) is studied, and this information is condensed on a 2-D image. Fig. 8 (bottom) shows the construction of the DTM for the HHDC. It is important to highlight how this representation provides full information on how the topographic information of this forested region of Maine is changing simply using satellite emulations. Mathematically, given the histogram height distribution along the z-axis of the (xth, yth) footprint of the HHDC, $\mathcal{X} \in \mathbb{R}^{m\times n\times c}$, a sample probability density function on z can be obtained as

$$\mathcal{P}_{x,y,z}(z) = \frac{\mathcal{X}_{x,y,z}}{\sum_{k=1}^{c} \mathcal{X}_{x,y,k}} \tag{5}$$

for $\forall z \in \mathcal{Z}$. Note that the denominator of this equation exists to ensure a valid probability density function. Then, the cumulative distribution function along the z axis of the (xth, yth) footprint of the HHDC, $\mathcal{X}$, is obtained as

$$\mathcal{F}_{x,y,z} = \sum_{k=1}^{z} \mathcal{P}_{x,y,k}. \tag{6}$$

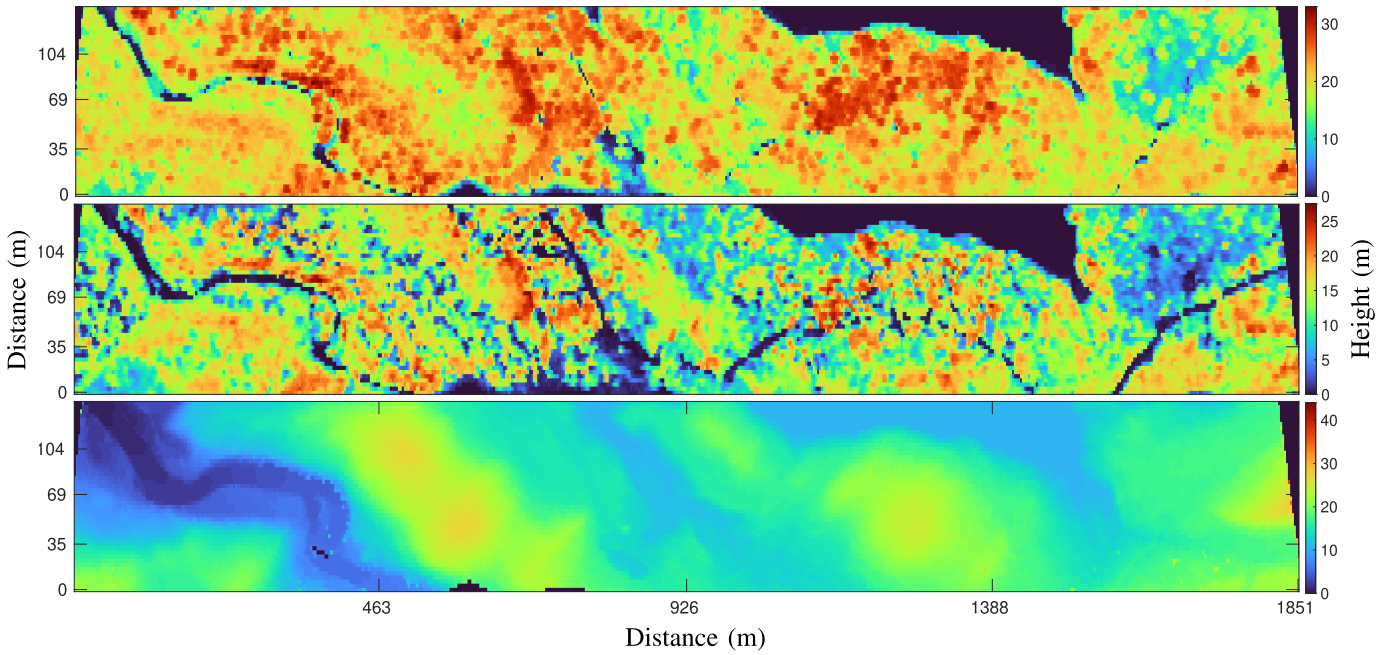

Fig. 8. (Top) CHM of the HHDC. (Middle) 50% percentile as it is typically used for biomass studies. (Bottom) DTM of the HHDC.

It is important to understand that this cumulative distribution is only performed along the z-axis which is where the heights in the HHDC are confined. Given $\mathcal{F}$, the DTM of an HHDC, denoted as $\mathbf{F}_{\text{DTM}}$, can be described as

$$\mathbf{F}_{\text{DTM}} = \inf\{z | \mathcal{F}_{x,y,z} \geq \theta : \forall z \in \mathcal{Z}\} \tag{7}$$

where $\mathcal{Z}$ represents all the possible values for heights and θ represents the threshold that exceeds a minimum noise level before being considered as terrain. In general, the set $\mathcal{Z}$ is all real numbers, but clearly, when considering heights, it is possible to reduce this range significantly for the forests being studied, e.g., a forest might only have heights in the range 0–30 m.

D. Canopy Height Model

Finally, the fourth proposed visualization of the HHDC is the well-known CHM. This type of visualization is very similar to the DTM, with the difference that what is intended is to study the height of the trees with respect to the floor on which they are located. To do this, it is necessary to know the height of the trees at the waveform level, and, as with the DTM, it is sufficient to determine at what height the waveform reaches a maximum value (usually, the 98th percentile is assumed and not the 100th percentile to avoid noisy signals), and subtract it with the DTM mentioned above. An example of the CHM for the study forest is shown in Fig. 8 (top). It is important to note that this representation helps establish important properties of the forest, such as its maximum and average heights. Having mathematically defined the DTM, it is easy to describe the CHM (denoted here as $\mathbf{F}_{\text{CHM}}$) as follows:

$$\mathbf{F}_{\text{CHM}} = \mathbf{F}_{\text{DEM}} - \mathbf{F}_{\text{DTM}} \tag{8}$$

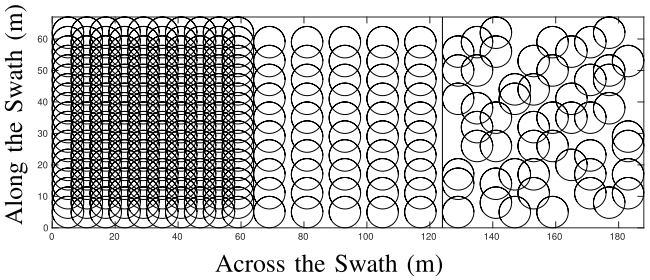

Fig. 9. CASALS LiDAR sampling patterns over the same region using overlapping footprints. (Left) Full sampling (for this case a total of 400 shots). (Middle) Bayer subsampling covering the same area with only 100 footprints. (Right) Blue noise subsampling with 100 footprints. Subsampling in space and photon count results in a compression ratio of 800:1.

where $\mathbf{F}_{\text{DEM}}$ (known as the digital elevation model) is defined as

$$\mathbf{F}_{\text{DEM}} = \inf\{z \,|\, \mathcal{F}_{x,y,z} \geq 0.98 : \forall z \in \mathcal{Z}\}. \tag{9}$$

A critical point to highlight is the flexibility of studying satellite measurements when represented as a hyperheight cube. All the visualizations mentioned have been used to study canopy for several decades. The cubic representation only retains these notions in such a way that all the tools used so far are still valid. However, it is essential to note that the HHDC can display all the visualizations simultaneously and consistently.

IV. Satellite Hyperheight Data Cubes

The aforementioned HHDCs (Fig. 6) were derived from high-density point clouds acquired by low-altitude LiDARs such as those provided by G-LiHT. If one were to sense the same area depicted in Fig. 6, but instead of using G-LiHT one would use a satellite LiDAR system such as CASALS,

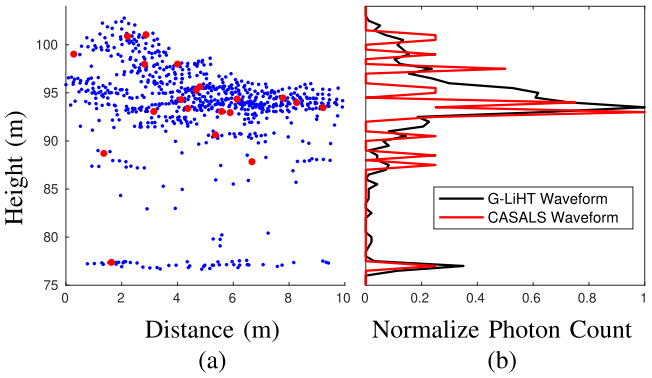

Fig. 10. (a) Photons returned from a 10-m footprint to G-LiHT (blue) and the compressive measurement of photons returned to CASALS (red). (b) Corresponding histogram waveforms. The histograms are normalized to their maximum value for sake of comparison.

what would the HHDC look like? To answer this question, one must consider the scale difference that exists between the traditional airborne measurements, such as G-LiHT, and the measurements acquired in a satellite system such as ICESat-2 or AWSL in CASALS. G-LiHT measurements are taken from an altitude of about 335 m above ground so that each footprint will have a very high photon density per footprint (∼1000 on a 10-m-diameter footprint), and hence they can have high-photon count resolution. However, ICESat-2's orbit is about 500 km above ground so that each illuminated footprint will return significantly fewer photons (∼100 photons per footprint). Satellite LiDAR HHDC's thus represents a noisy sketch of airborne LiDAR HHDC's, where the satellite-derived waveforms are poorly defined and quite noisy as illustrated in Fig. 10. Notably, two different effects cause substantial loss of photons: on one hand, there is attenuation caused by the distance of the laser, which, as is well-known, quickly drops with a factor of z^2, where z is the distance to the target [35]. This is amplified by the fact that the LiDAR instrument mainly operates on nonperfect reflective surfaces, e.g., tree canopies, so few photons reach the target reflect in unwanted directions or will not even reflect back. On the other hand, there is the physical phenomenon of atmospheric loss of photons as they travel to the satellite detector, which acts as a compressive projection of the underlying signals. The laser beam attenuation as it propagates through the atmosphere is mainly due to molecular absorption and particle scattering. Both the effects are dependent on the operating wavelength. The Beer–Lambert law is the fundamental model that determines the overall transmittance τ and is given by $\tau = \exp^{-\gamma z}$ where z is the propagation distance, and γ is the attenuation coefficient [36], [37]. Resolved over a single path length, the transmittance becomes $\tau = \exp^{-\int_0^z \gamma(z,f)dz}$ where the attenuation coefficient $\gamma(z, f)$ is defined by the absorption and scattering properties of the medium and the wavelength of the laser beam. For cases where $\gamma(z, f)$ is not constant along the path length, a more general expression of transmittance can be modeled as described in [37]. Fig. 10(a) depicts a side view rendition of a point cloud inside the 10-m footprint obtained with G-LiHT (blue points) and the attenuated compressive measurements obtained from a satellite LiDAR (red points). Note how the point density in the satellite system is substantially reduced to such an extent that there are only 20 photons in the whole footprint. A comparison of the shape of the waveforms of both the systems shows the challenge of working with footprints measured from orbit, as shown in Fig. 10(b) (for ease of comparison, the waveforms are normalized to their maximum value).

A. Overlapping Footprints

Several alternatives can be used to overcome the limited number of photons returning to the sensor onboard the satellite. On one hand, the GEDI strategy could be used in which large footprints (diameter of 25 m) are used to capture a larger number of photons. This, however, significantly reduces the spatial resolution, as shown in Fig. 1. Note how each footprint associated with GEDI covers large portions of the terrain, but there is a significant separation between them. For this reason, systems such as ICESat-2 choose to use smaller footprints (13-m diameter), which, while collecting fewer photons, can be captured successively so that there is overlap between contiguous footprints. This way, the lack of photons per footprint can be overcome by capturing many consecutive footprints, as shown in Fig. 1. The state-of-the-art systems, such as CASALS, aim to overcome the limitation on spatial sampling resolution using even smaller footprints (10-m diameter) and guaranteeing an overlap both along the direction of the satellite motion and across the trajectory, as shown at the left of Fig. 9. The waveform sampling pattern in this case thus exhibits footprint overlap along and across the swath.

To better understand the difference between HHDCs obtained from low-altitude high-density photon measurements such as G-LiHT and those attained with CASALs. Fig. 11(a) depicts the high-density photon count CHP for a swath of approximately 1.8 km long. In contrast, Fig. 11(b) shows the same swath but considering footprints with 20 photons (CASALs) including the mentioned footprint overlap. Note how, regardless of the lack of photons in the CASAL's footprints, the main features of the CHP are still there. For example, features such as terrain structure or tree density are still present. As will be shown later, it is possible to use measurements such as those presented in Fig. 11(b) and reconstruct the CHP presented in Fig. 11(a) using ML techniques.

V. Compressive Footprint Sensing

CASALS uses novel nonmechanically scanning LiDARs that uses wavelength tuning and a high-resolution grating to achieve high-speed and accurate cross-track beam-steering. In a swath path, CASALS can thus acquire the densely sampled waveform patterns shown in Fig. 9(a) where the footprints exhibit some overlap in both, along the swath and across the swath directions. Sensing of such a large amount of spatial information across hundreds of height bins, however, requires significant laser power and data storage resources. Notably, wavelength tuning in tunable lasers is not just capable of raster scans but can hop arbitrarily from any wavelength in the wavelength range to any other wavelength in the tunable

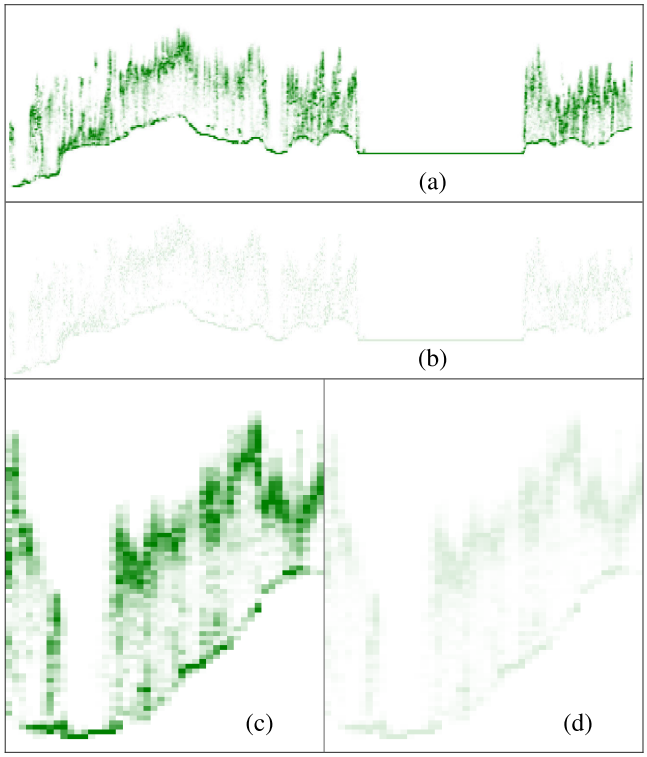

Fig. 11. (a) CHP using all the available photons. (b) CHP created using a limited amount of photons per footprint. Note how the CHP is not well-defined due to reduced photon resolution. CHP across the swath view using (c) full photon counts and (d) limited photon counts.

range, enabling arbitrary LiDAR scan patterns. This capability naturally enables the application of compressive sensing protocols to sense the HHDC data cube with fewer footprint measurements per area. Note that footprint subsampling imposes subsampling not only in the spatial axes of the scene but also on the photon count as these travel from the surface of Earth to the satellite's orbit. Data compression in CASALS will thus be attained directly, at the sensor, via footprint compressive sampling. The need for digital compression of the acquired footprints can be reduced or discarded altogether. The compressive measurements in CASALS are then realized first by beam-steering footprint sampling in the spatial dimension, followed by atmosphere extinction of the photons in each footprint that are attenuated as a function of distance.

The sensing mechanism model is illustrated in Fig. 12, where the high-resolution G-LiHT data cube $\mathcal{X}_G \in \mathbb{R}^{n\times m\times c}$ is projected onto the compressed CASALS data cube $\mathcal{X}_C \in \mathbb{R}^{n_2\times m_2\times c_2}$ having much fewer footprint measurements and height bins, i.e., $n_2 < n$, $m_2 < m$, and $c_2 < c$. CS dictates that one can recover the high-resolution G-LiHT HHDC $\mathcal{X}_G$ of interest from the fewer measurements available in the CASALS HHDC $\mathcal{X}_C$. To make this possible, CS relies on two principles: sparsity, which characterizes the hyperheight scenes of interest, and incoherence, which shapes the footprint sensing structure. Sparsity requires that HHDCs found in nature can be concisely represented in some basis with just a small number of coefficients. As shown in Fig. 6(c), this is indeed the case with HHDCs where natural scenes exhibit correlation among adjacent tubes and also across height bins.

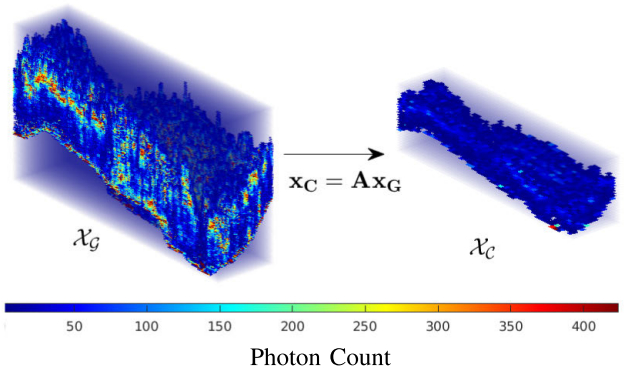

Fig. 12. Compressive sensing of the low-altitude G-LiHT HHDC $\mathcal{X}_G \in \mathbb{R}^{nmc}$ by the sensing matrix $\mathbf{A}$ leading to CASALS' compressed HHDC $\mathcal{X}_C \in \mathbb{R}^{n_2m_2c_2}$ with photon extinction, where $n_2 < n$, $m_2 < m$, and $c_2 \ll c$.

Incoherence refers to the structure of the sampling waveforms used in CS that unlike the signals of interest have a dense representation in the basis of signal representation. In CASALS, the sampling waveforms are limited to sparse ON-OFF laser-coded illumination and also by the random loss of photons as they transverse from the illuminated footprint areas to the detector at high altitude on board of the satellite. Both these operations are incoherent with the basis representation of HHDCs.

Let $\mathbf{x}_g \in \mathbb{R}^{nmc}$ and $\mathbf{x}_c \in \mathbb{R}^{n_2m_2c}$ be the reindexed 3-D vector representations of $\mathcal{X}_G$ and $\mathcal{X}_C$, respectively. The signal acquisition can then be written as an underdetermined system of equations $\mathbf{x}_c = \mathcal{A}_\Psi(\mathbf{x}_g) + \mathbf{n}$ where $\mathcal{A}(\cdot)$ is a nonlinear function that characterizes the sensing phenomena and Ψ is a matrix determining the sampling illumination pattern, which together characterize the forward imaging problem. $\mathbf{n}$ represents the observation noise. The overarching objective in compressive LiDAR is thus to compressively sense Earth from hundreds of kilometers above Earth to then reconstruct the 3-D imagery with resolution and coverage, as if the data were collected from airborne platforms at just hundreds of meters in height. The CS inverse problem reduces to solving an underdetermined nonlinear system of equations to recover θ such that the cost function

$$\arg\min_{\mathbf{x}_g} \left\|\mathbf{x}_c - \mathcal{A}_\Psi\left(\mathbf{x}_g\right)\right\|^2 + \lambda R\left(\mathbf{x}_g\right) \tag{10}$$

is minimized where $R(\cdot)$ is a regularizer function and λ is a regularization constant. CS specifically deals with characterizing the "goodness" of the scanning pattern [38], [39]. Random on–off patterns are commonly used is CS but random patterns that maximize the mutual distance among "on" samples, i.e., blue noise (BN), have been shown to have optimal properties in CS in a number of applications [39], [40], [41]. Fig. 9(b) and (c) depicts Bayer and blue noise sampling patterns, respectively [42], [43], [44]. Since the tunable laser is limited in the number of wavelength hops per unit of time, the advantage of CS in CASALS is thus enabling a much wider field-of-view swath sparse scanning, allowing the capacity of measuring Earth much faster, with fewer swath paths, and at high-photon count resolution. The sampling patterns in Fig. 9(b) and (c), for instance, sample the HHDC with 25% density thus making

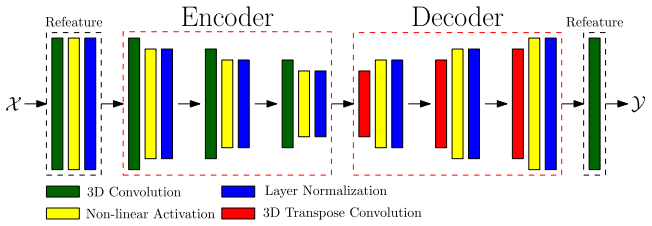

Fig. 13. Proposed 3-D Autoencoder architecture. This neural network allows to take HHDCs with incomplete and poorly defined features and create high-resolution HHDCs that could be used in environmental studies. This neural network uses 3-D convolutions and Gaussian error linear unit (GELU) nonlinear activation functions to perform the task.

it possible to cover the Earth's surface four times faster. For a 25% footprint sampling, the HHDC compressive sampling gives a compression ratio of 800:1.

VI. Reconstruction Algorithm

Given the compressive sampled footprints along the spatial and height (photon count) dimensions, the goal is then to reconstruct the high-resolution HHDC in all its dimensions including the photon counts. Several numerical and iterative optimization-based algorithms are available to solve the inverse CS problem. Typical computations performed by these algorithms include matrix pseudoinverses, sparse basis transformations, and vector–matrix multiplications. Given that the underlying signals in HHDCs are high dimensional, these calculations require a large number of floating point operations. Reconstructions of HHDC of large scenes are indeed overwhelming since they can take hours in desktop architectures [38], [39], [45], [46]. To avoid the computational bottleneck in the inverse problem, this work uses a learning-based algorithm trained with a large dataset of known HHDCs to reconstruct new unknown HHDCs. Currently, there are a great variety of algorithms that allow to perform similar tasks to that of reconstructing HHDCs, but it is important to highlight the two most important particularities of the problem under study: it is required, on one hand, to fill the inputs that were not taken by the laser, and on the other hand, it is required to increase the photon-level resolution of the HHDC. Considering these constraints, it is proposed to use a 3-D Convolutional Autoencoder as shown in Fig. 13. This architecture has been shown to be adequate for matrix completion and super-resolution tasks, because the latent spaces it generates within it are able to extract the important features of cubes and images, largely ignoring information that may be lost during the selected sampling.

To this end, it is necessary to consider some particularities from the proposed architecture shown in Fig. 13. First, it is imperative to consider the dimensions of the inputs and outputs of the system. Understanding that an HHDC is described by the offsets along and across the swath, as well as by the vertical resolution, it suffices only to assume a subset of cubes such that their dimensions are confined in dimensions that a neural network can understand. For this purpose, it is assumed that the HHDC is stored as a tensor $\mathcal{X} \in \mathbb{R}^{n\times m\times c}$ to be reconstructed by the Autoencoder. Therefore, it is clear that the dimensions for the input and output of the network are $n \times m \times c$. In the actual application, the dimensions n and m represent the number of samples along and across the swath, respectively, and the dimension c represents the number of bins chosen to represent the histograms of the cube. In principle, n could be very large since it would be given by the continuous line of samples taken as the satellite moves during its successive orbits; however, it is necessary to restrict the 3-D so that the input tensor is meaningful for the network. Therefore, it is proposed to restrict the HHDC reconstruction to tensors with dimensions $\mathcal{X} \in \mathbb{R}^{32\times32\times100}$. It may seem somewhat small dimensions for the terrestrial scale; however, if we consider offset along and across the swath of 3 and 6 m, using a vertical resolution of 0.5 m, it can be stated that during each reconstruction, the algorithm would be evaluating a ground surface of approximately 100 × 200 m for which the trees can be up to 50 m high. This is possible since the Autoencoder is making nonlinear transformation $g(\mathcal{X})$:$\mathbb{R}^{32\times32\times100} \rightarrow \mathbb{R}^{32\times32\times100}$, which seeks to take an incomplete and ill-defined input tensor $\mathcal{X} \in \mathbb{R}^{32\times32\times100}$ that has many missing inputs and has low photon counts and produces a high-resolution output tensor $\mathcal{Y} \in \mathbb{R}^{32\times32\times100}$ with complete information and full photon count.

Describing the architecture shown in Fig. 13, four fundamental blocks can be identified: 3-D convolutions, nonlinear activations, layer normalizations, and 3-D transpose convolutions. These blocks are specifically designed for the CLS application and were chosen based on the expected performance. It is important to highlight that the Autoencoder does not have explicit max pooling layers, and therefore, the bottleneck scheme is achieved solely and exclusively with the convolutions.

A. Input Refeature Block

One of the most important characteristics of working with convolutional layers, apart from the fact that they extract spatial information, is that they allow to increase the hyperdimensions of the input signals and create deep feature maps on which much more can be learned. The proposed Autoencoder uses convolutional layers but must start its structure with a block called Input Refeature used for transforming the input data into something more amenable for a convolutional layer. This block seeks to create pseudo-HHDCs from the original incomplete cube so that subsequent layers can more easily complete it. After passing through the Input Refeature Block, 32 cubes will be created from the original cube, but the spatial dimensions will not be reduced. That is, this block can be viewed as a transformation $g_1(\mathcal{X})$:$\mathbb{R}^{32\times32\times100\times1} \rightarrow \mathbb{R}^{32\times32\times100\times32}$ where the feature maps are obtained from convolutions performed with cubic kernels $\mathcal{K}_i^1$ of dimension $\mathbb{R}^{3\times3\times3}$ with nonlinear activation functions of Gaussian error linear unit (GeLU) type defined as $\mathcal{X}_{1i} = \sigma(\mathcal{X} * \mathcal{K}_i^1)$, where σ is the GeLU function described by $\sigma(x) = xP(X \leq x) = x\Phi(x)$ where $\Phi(x)$ is the Gaussian cumulative distribution function.

To improve the training and performance of the neural network, it is always recommended to consider normalization layers. For the design of the Autoencoder, it is proposed to use layer normalization because it is well-known that these types of normalizations are independent of the chosen batch

size and still guarantee stability in the hidden layers of the network. The normalizations must be performed considering the tensor nature of the inputs, and therefore, the output of the layer normalization is given by

$$\mathrm{LN}(\mathcal{X}_N) = \gamma\left(\frac{\mathcal{X}_N - \mu(\mathcal{X}_N)}{\left(\sqrt{\sigma(\mathcal{X}_N)^2 + \epsilon}\right)}\right) + \beta \tag{11}$$

where γ and β are the parameters to be learned during training, ϵ is a hyperparameter for algorithmic stability, and μ and σ^2 are, respectively, the mean and variance per channel N described as

$$\mu(\mathcal{X}_N) = \frac{1}{nmc}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{c}\mathcal{X}_{Nijk}, \quad \text{and}$$
$$\sigma^2(\mathcal{X}_N) = \frac{1}{nmc}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{c}\left(\mathcal{X}_{Nijk} - \mu(\mathcal{X}_N)\right)^2.$$

B. Encoder

The encoder follows a datapath similar to what was mentioned in the previous block, with the particularity that now it seeks to reduce the spatial dimensions of the cube to reach a latent representation that manages to capture as much information of the incomplete cube as possible. Evaluating the three steps shown in the encoder of Fig. 13, it can be understood that this is a set of transformations g_2:$\mathbb{R}^{32\times32\times100\times32} \rightarrow \mathbb{R}^{16\times16\times100\times64} \rightarrow \mathbb{R}^{14\times14\times100\times128} \rightarrow \mathbb{R}^{12\times12\times100\times128}$, where the reduction in the dimensions of the HHDCs, within the Autoencoder latent space, is associated with the selected strides and paddings. Considering that the layers are the same as in the previous block (3-D convolutions, nonlinear activations, and layer normalizations), no details are necessary on how they are mathematically described; it is only mentioned that the encoder receives a tensor $\mathcal{X}_1$ and generates a latent tensor $\mathcal{X}_l$.

C. Decoder

Once the latent tensor $\mathcal{X}_l$ is obtained, the spatial dimensions of which were decreased from the input tensor $\mathcal{X}$, but widely expanded a number in feature maps, it can be attempted to recover the original signal by means of a decoder block as shown in Fig. 13. Note that the decoder performs the inverse operations to the encoder (as expected) and therefore can be viewed as a set of transformations g_3:$\mathbb{R}^{12\times12\times100\times128} \rightarrow \mathbb{R}^{14\times14\times100\times128} \rightarrow \mathbb{R}^{16\times16\times100\times64} \rightarrow \mathbb{R}^{32\times32\times100\times32}$, where what is done is to increase the number of the spatial dimensions and decrease the hyperdimensions of the feature maps. However, it is key to consider that for the encoder 3-D convolutions are used while in the decoder 3-D transposed convolutions are implemented so that each pseudo-HHDC of the decoder can be expressed as

$$\mathcal{X}_i = \sigma\left(\mathcal{X}_l \circledast \mathcal{K}_i^j\right) \tag{12}$$

where $\circledast$ represents a transposed convolution and $\mathcal{K}_i^j$ represents the kernel of the jth transposed convolution.

D. Output Refeature Block

After the latent tensor $\mathcal{X}_l$ is processed by the decoder, a decoded tensor $\mathcal{X}_d \in \mathbb{R}^{32\times32\times100\times32}$ is obtained. Nonetheless, the output tensor is not a tensor composed of 32 HHDCs but only one. For this reason, a final layer is needed to assemble the 32 cubes and generate a representation containing the high-resolution HHDC. In brief, the Output Refeature simply reverses the operation performed by the Input Refeature so it is a transformation $g_4(\mathcal{X}_d)$:$\mathbb{R}^{32\times32\times100\times32} \rightarrow \mathbb{R}^{32\times32\times100\times1}$.

E. Error Metrics

To measure the accuracy of the reconstructions made by the neural network, it is necessary to propose metrics that allow measuring the differences between two HHDCs. For this purpose, six metrics are presented to help perform this task. Of course, only one of them is required when training the network, but the other five will help have a quantitative notion of how well the reconstruction task is being done. As it is widely accepted, to perform the training of the network, it is proposed to use an rms-type measure between the reconstructed HHDC $\hat{\mathcal{X}}$ and the desired HHDC $\mathcal{Y}$ as $\mathrm{rms} = 1/(nmc)(\sum_{n,m,c}(\hat{\mathcal{X}}_{i,j,k} - \mathcal{Y}_{i,j,k})^2)^{1/2}$. This metric gives a per photon difference between two HHDC, and thus, it gives the ability to tell how similar the footprint information is going to be. Having an rms like error could be misleading in some scenarios, and for that reason, an MAE-type error is also proposed as $\mathrm{MAE} = 1/(nmc)\sum_{n,m,c}|\hat{\mathcal{X}}_{i,j,k} - \mathcal{Y}_{i,j,k}|$ for comparing two HHDC. This metric gives a broad overview of the raw differences between two histograms by looking at the whole cube at once. Even though having footprint difference information is enough for training the model, it is also beneficial to have metrics that allow to compare the most important subviews of the cube, i.e., the CHM and the DTM. For that regard, rms and MAE are calculated for both the CHM and the DTM. It is critical to highlight that since the CHM and the DTM provide height information, these errors are no longer in units of photon counts but rather in meters.

F. Summary of the Neural Network

The 3-D convolutions in the Input Refeature block have a kernel size of $3 \times 3 \times 3$, a stride of 1 on every dimension, and a padding such that the input and output HHDCs have the same size. The 3-D convolutions of the encoder block also have a kernel size of $3 \times 3 \times 3$. The first 3-D convolutions on the encoder block have a stride of 2 on both x and y such that the dimensions of the HHDC are reduced by half on these axes, so no reduction is performed with padding. On the second and third 3-D convolutions of the encoder block, the bottleneck of the Autoencoder is achieved by doing convolutions with no padding and slowly reducing the dimensions of the HHDCs. All 3-D transpose convolutions on the decoder block have a kernel size of $2 \times 2 \times 2$; this is so that more high-frequency components can be recovered. The stride of the first two 3-D transpose convolutions is one on each dimension, but the output HHDC is padded with zeros so that the dimensions are slowly increased. The third 3-D transposed convolution of the decoder block has a stride of

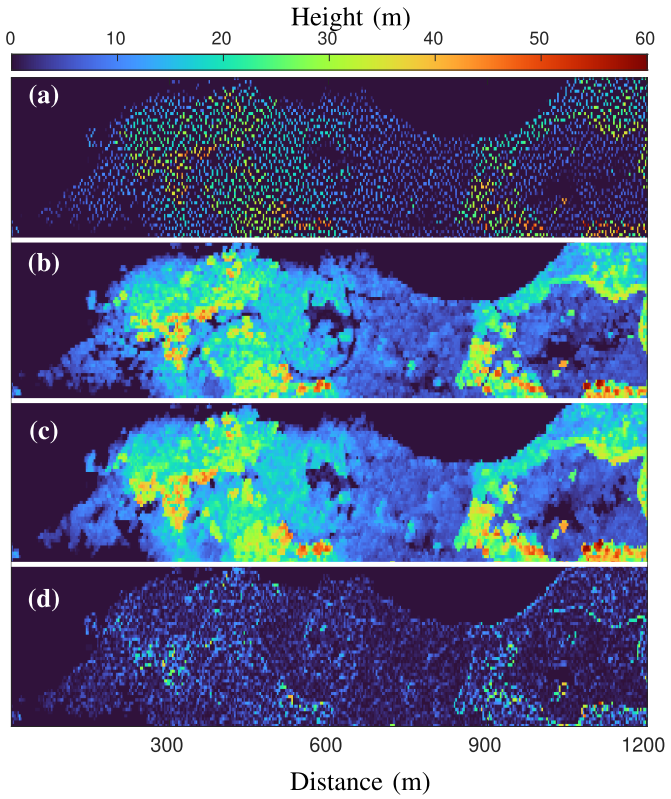

Fig. 14. CHM for the Oregon Willamette National Forest. (a) CHM of the CASALS input HHDC. (b) CHM for the ground truth. (c) CHM for the NN reconstructed HHDC. (d) Difference between the reconstruction and the ground truth.

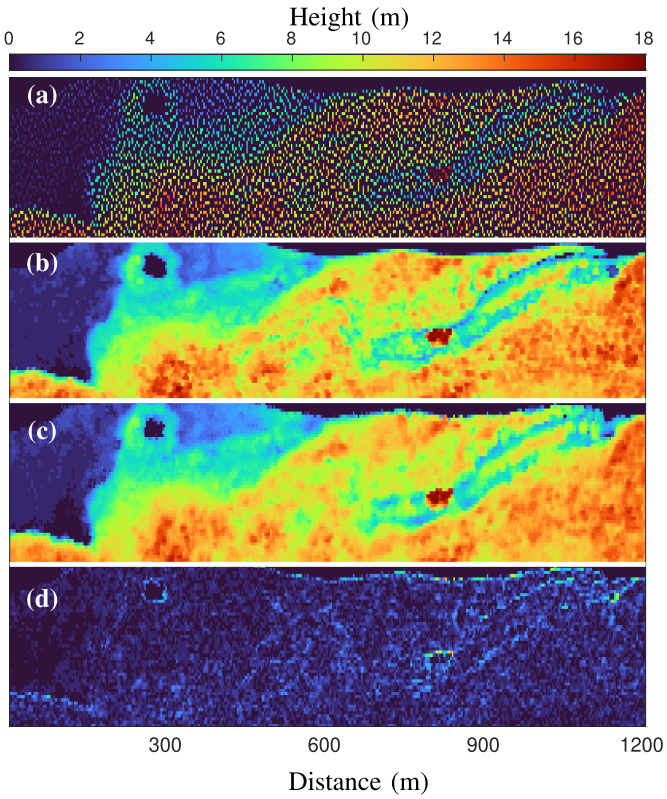

Fig. 15. CHM for the Florida Everglades National Park. (a) CHM of the CASALS input HHDC. (b) CHM for the ground truth (G-LiHT). (c) CHM for the HHDC reconstructed using the proposed neural network. (d) Difference between the reconstruction and the ground truth.

2 on both x and y such that the dimensions of the HHDC are doubled on these axes. Finally, to fully recover the HHDC, the Output Refeature block combines all the representations within the decoder as a single cube by doing 3-D convolutions with a kernel of $3 \times 3 \times 3$. All the nonlinear activations on the neural network are GELU functions and the numerical stability parameter of all the layer normalizations is 1×10^{-5}. The complete implementation of this neural network has around 30 million parameters and was trained with dual NVIDIA Quadro RTX5000 GPUs. As mentioned in Section VII, the training was performed on two specific geographical regions, and thus, it was performed twice. It should be pointed out that the neural network architecture is not changed regardless of the chosen region, i.e., only the weights are changed during training. During training, the ADAM algorithm with a learning rate of 0.001 was used, and performing cross-validation and optimizing an rms type cost function over the $32 \times 32 \times n$ cubes was implemented. The batch size was 128 HHDCs. This hyperparamenters were selected using a binary search over the first epochs of the training until the required performance was achieved. Understanding time-related complexity is crucial, and it can be approached from two angles: training and inference. It is important to differentiate between these, as they lead to different strategies. One involves performing inference on the ground, while the other does so directly in orbit. When it comes to training, the neural network was trained for 100 epochs, taking around 100 epochs to reach its best performance. To prevent overfitting, a stopping system with a validation set was used. On the inference side, it is worth noting that these experiments used powerful GPUs, resulting in reconstruction times of less than a second. However, when considering a full satellite implementation, factors such as hardware and onboard power constraints in orbit need to be taken into account.

VII. Results and Discussion

To test the power of the proposed methodology, two forested areas in the United States that are more than 4000 km apart and are part of the G-LiHT collections were used. The two areas chosen were the Everglades Natural Park in the state of Florida and the Willamette National Forest in the state of Oregon. All the tests consider footprints of 10 m in diameter with 3 m along the swath and 6 m across the swath. In addition, it is considered that the number of photons returning to orbit is around 20 photons per footprint and that the downsampling pattern used by the satellite to complete the CLS is a blue noise pattern with a sampling rate of 25%, as shown at the bottom of Fig. 9. This implies that the compression level, considering there may be an initial number of photons of more than 1000 per footprint, is close to 800:1. To compare the efficiency of the reconstruction algorithm, it is suggested to analyze the four views of the cube proposed above, i.e., CHM, DTM, along the swath, and across the swath.

Figs. 14 and 15 show the CHMs for both Florida and Oregon HHDCs. In both the figures, the top image corresponds to the input, the middle corresponds to the ground truth (Gt),

TABLE I
METRICS

	Florida				Oregon			
Metric	Bayer 25%	BN 5%	BN 10%	BN 25%	Bayer 25%	BN 5%	BN 10%	BN 25%
RMS (Photons)	27.71	33.57	31.17	**25.79**	1.33	1.71	1.64	**0.68**
RMS DTM (m)	**0.36**	0.39	0.37	0.42	8.49	112.25	10.23	**6.66**
RMS CHM (m)	**1.53**	2.54	1.90	1.57	10.46	16.97	13.46	**8.00**
MAE (Photons)	10.64	13.69	12.44	**10.45**	0.14	0.19	0.18	**0.06**
MAE DTM (m)	0.06	0.08	0.06	**0.06**	4.19	99.72	4.24	**2.09**
MAE CHM (m)	0.94	1.52	1.13	**0.84**	5.41	8.77	6.21	**2.87**
RMS 50% (m)	**1.87**	2.35	2.40	1.90	9.21	13.32	11.08	**6.99**
MAE 50% (m)	**1.38**	1.83	1.86	1.42	4.89	6.66	4.83	**2.12**

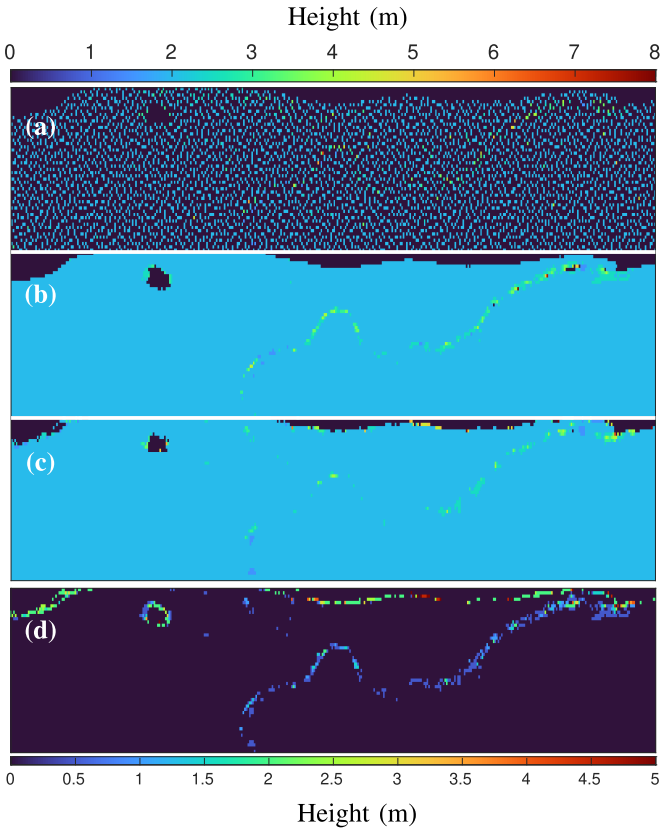

Fig. 16. DTM for the Florida Everglades National Park. (a) DTM of the input CASALS HHDC. (b) DTM for the ground truth (G-LiHT). (c) DTM for the HHDC reconstructed using the proposed neural network. (d) Difference between the reconstruction and the ground truth.

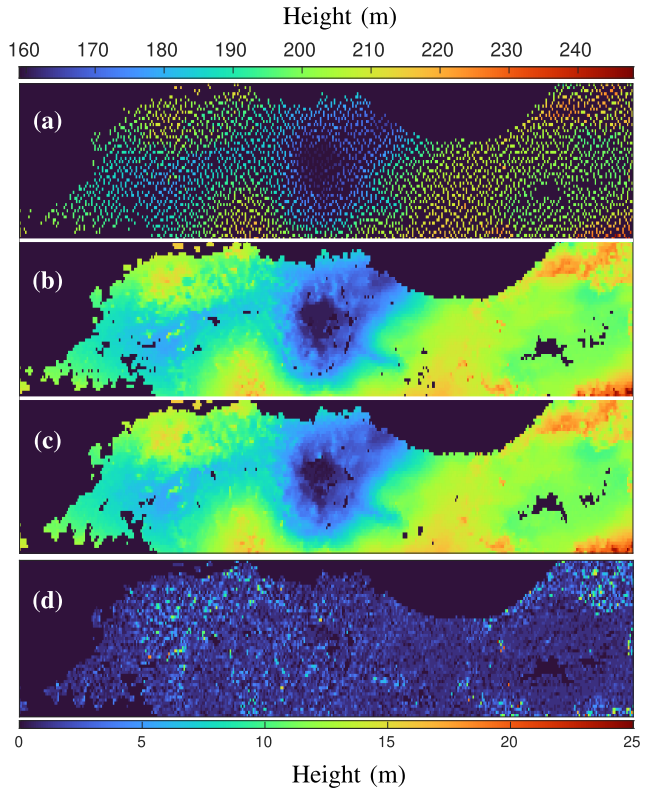

Fig. 17. DTM for the Oregon Willamette National Forest. (a) DTM of the CASALS input HHDC. (b) DTM for the ground truth (G-LiHT). (c) DTM for the NN reconstructed HHDC. (d) Difference between the reconstruction and the ground truth.

and the bottom corresponds to the reconstruction. Looking at the CHMs, it is clear that the algorithm is able to recover the most important features of the HHDC such as the tree crowns and the rivers. Figs. 16 and 17 show the DTMs for both Florida and Oregon HHDCs. As in the previous figure, the order of the images is input, Gt, and reconstruction. As with the CHM, the algorithm is able to recover all relevant features of the DTM. Figs. 18–21 show the profiles along and across the swath for Florida and Oregon. It is interesting to note that both the representations are very similar given that they show the same geographical location; however, it is important to show both the representations as the footprint separation is different between the two and this could result in potential loss of quality in the across the swath reconstruction. Finally, it is essential to note that although the neural network is the same for both the cases, i.e., Florida and Oregon, it was trained with the data according to the reconstructed location.

From these results, it can be understood how powerful learning algorithms are when implementing CLS problems. From several of the input figures, it is clear that many cube features disappear when photons or footprints are removed. This is particularly notable in the along and across the swath figures. However, after performing the corresponding reconstructions, features such as the tree canopies in the CHM or the roughness profile in the along the swath are recovered well enough to perform any ecological or environmental study. To conclude the study, it is crucial to show quantitative samples of the system performance. In Table I, the results of the six proposed

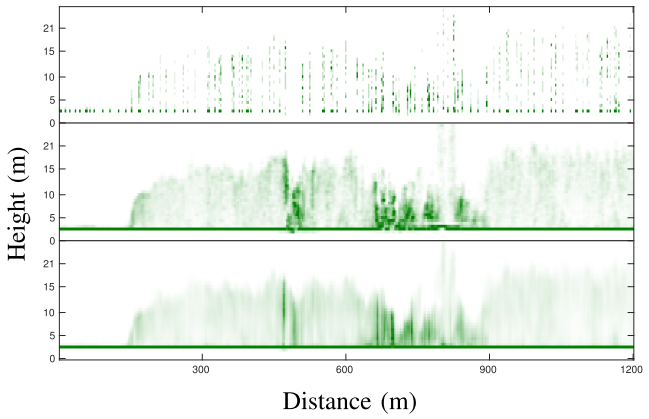

Fig. 18. Along the Swath view for the Florida Everglades National Park. (Top) Along the swath view of the CASALS input HHDC. (Middle) Along the swath view for the ground truth (G-LiHT). (Bottom) Along the swath view for the NN reconstructed HHDC.

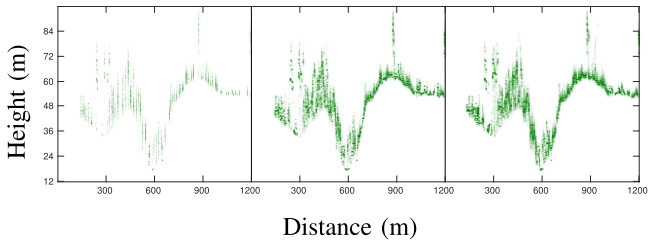

Fig. 19. Along the swath view of the Oregon Willamette National Forest using (Left) CASALS HHDC and (middle) ground truth (G-LiHT) HHDC. (Right) Reconstructed HHDC using the proposed neural network.

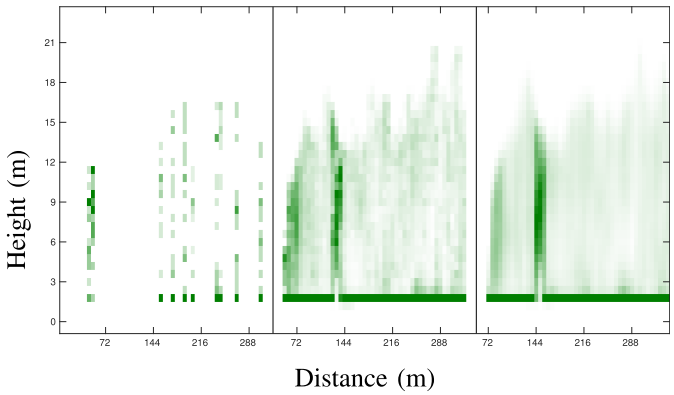

Fig. 20. Across the swath views for the Florida Everglades National Park. (Left) Across the swath view of CASALS HHDC. (Middle) Across the swath view for the ground truth (G-LiHT). (Right) Across the swath view for the NN reconstructed HHDC. Note how the neural network is able to reconstruct the profile using adjacent profiles sampled from the blue noise.

metrics are summarized. It is necessary to understand what the differences between the metrics represent and why they are not the same between the two locations. To begin with, if the DTM and CHM metrics are examined for both the locations, it can be seen that in Oregon, they are always higher. This is because the trees are taller there, and consequently, the metric will always yield a higher value. However, if reconstructions are graphically evaluated, it can be seen that they are of the same quality. To complement the study, different sampling patterns and ratios are also provided, e.g., blue noise with 10%, but

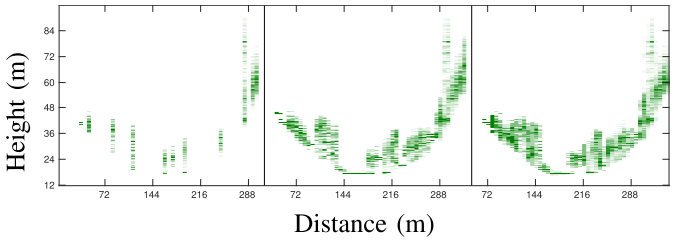

Fig. 21. Across the swath views for the Oregon Willamette National Forest. (Left) Across the swath view of CASALS HHDC. (Middle) Across the swath view for the ground truth (G-LiHT). (Right) Across the swath view for the NN reconstructed HHDC.

the proposed downsampling method outperforms every other scheme in most metrics.

VIII. Conclusion

This work has presented a complete methodology to perform the emulation of new satellite LiDAR technologies, as well as a methodology to expand the capabilities of these new systems through compressed sensing techniques. A new way to model the outputs of LiDAR systems has also been proposed. All the information measured by a satellite is condensed into a 3-D representation known as an HHDC. Through this new cubic representation, it is possible to obtain all the visualizations typically used in the literature to study the canopy while knowing that all of them can be used simultaneously. In addition, this work has addressed a machine learning study that allows taking incomplete HHDCs given the characteristics of satellite LiDAR and generating cubes that have airborne LiDAR systems' features. The main result is the ability of the proposed neural network to not only recover the footprints lost due to compressed sensing strategies but also how it is able to generate representations that are very dense at the photon level.

Several paths are being followed to continue this research. On one hand, a detailed study of how changing the footprint loss system affects the learning algorithm can be carried out. Furthermore, techniques can be explored to determine the best footprint configuration according to the region being sensed, for example, how to change the sampling pattern if a very dense forest or a desert is being measured. Another possible way to continue this work is to generate algorithms to increase the resolution of each footprint so that somehow the proposed overlaps are eliminated, and thus, a super-resolution system can be obtained. Finally, it is possible to further improve the performance of the reconstructions by implementing different neural networks such as generative adversarial networks and vision transformer networks.

References

[1] R. O. Dubayah and J. B. Drake, "LiDAR remote sensing for forestry," *J. Forestry*, vol. 98, no. 6, pp. 44–46, 2000.

[2] M. Dassot, T. Constant, and M. Fournier, "The use of terrestrial LiDAR technology in forest science: Application fields, benefits and challenges," *Ann. Forest Sci.*, vol. 68, no. 5, pp. 959–974, Aug. 2011.

[3] D. Harding, "Pulsed laser altimeter ranging techniques and implications for terrain mapping," in *Topographic Laser Ranging and Scanning*. 2nd ed. Boca Raton, MA, USA: CRC Press, Feb. 2018, pp. 201–220.

[4] X. Liu, "Airborne LiDAR for DEM generation: Some critical issues," *Prog. Phys. Geography, Earth Environ.*, vol. 32, no. 1, pp. 31–49, Feb. 2008.
[5] M. Jaboyedoff et al., "Use of LiDAR in landslide investigations: A review," *Natural Hazards*, vol. 61, no. 1, pp. 5–28, Mar. 2012.
[6] A. Neuenschwander and L. Magruder, "The potential impact of vertical sampling uncertainty on ICESat-2/ATLAS terrain and canopy height retrievals for multiple ecosystems," *Remote Sens.*, vol. 8, no. 12, p. 1039, Dec. 2016.
[7] J. L. Irish and T. E. White, "Coastal engineering applications of high-resolution LiDAR bathymetry," *Coastal Eng.*, vol. 35, nos. 1–2, pp. 47–71, Oct. 1998.
[8] G. C. Guenther, "Airborne LIDAR bathymetry," in *Digital Elevation Model Technologies and Applications: The Dem User's Manual*, D. F. Maune, Ed., 2nd ed. Bethesda, MD, USA: American Society for Photogrammetry and Remote Sensing, 2007, pp. 253–320.
[9] V. Klemas, "Beach profiling and LiDAR bathymetry: An overview with case studies," *J. Coastal Res.*, vol. 277, pp. 1019–1028, Nov. 2011.
[10] S. Royo and M. Ballesta-Garcia, "An overview of LiDAR imaging systems for autonomous vehicles," *Appl. Sci.*, vol. 9, no. 19, p. 4093, Sep. 2019.
[11] H. Wang, B. Wang, B. Liu, X. Meng, and G. Yang, "Pedestrian recognition and tracking using 3D LiDAR for autonomous vehicle," *Robot. Auto. Syst.*, vol. 88, pp. 71–78, Feb. 2017.
[12] D. Zermas, I. Izzat, and N. Papanikolopoulos, "Fast segmentation of 3D point clouds: A paradigm on LiDAR data for autonomous vehicle applications," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 5067–5073.
[13] U. Weiss and P. Biber, "Plant detection and mapping for agricultural robots using a 3D LiDAR sensor," *Robot. Auto. Syst.*, vol. 59, no. 5, pp. 265–273, May 2011.
[14] C. Premebida, O. Ludwig, and U. Nunes, "LiDAR and vision-based pedestrian detection system," *J. Field Robot.*, vol. 26, no. 9, pp. 696–711, 2009.
[15] B. Cook et al., "NASA Goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager," *Remote Sens.*, vol. 5, no. 8, pp. 4045–4066, Aug. 2013.
[16] J. L. Bufton, D. J. Harding, and J. B. Garvin, "Shuttle laser altimeter," in *Proc. Shuttle Small Payloads Symp.*, 1999, pp. 1–10.
[17] B. E. Schutz, H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio, "Overview of the ICESat mission," *Geophys. Res. Lett.*, vol. 32, no. 21, Nov. 2005.
[18] 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, Feb. 2019.
[19] H. Tang, J. Armston, S. Hancock, S. Marselis, S. Goetz, and R. Dubayah, "Characterizing global forest canopy cover distribution using spaceborne LiDAR," *Remote Sens. Environ.*, vol. 231, Sep. 2019, Art. no. 111262.
[20] *National Academies of Sciences and Medicine, National Academies of Sciences, Engineering, and Medicine, Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation From Space*, The National Academies Press, Washington, DC, USA, 2018, doi: 10.17226/24938.
[21] G. Yang et al., "Adaptive wavelength scanning LiDAR (AWSL) for 3D mapping from space," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2022, pp. 4268–4271.
[22] D. J. Harding et al., "CASALS: A LiDAR and spectrometry SmallSat for a future polar altimeter mission," in *Proc. AGU Fall Meeting Abstracts*, 2020, pp. 20–28.
[23] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.
[24] B. Pan, X. Xu, Z. Shi, N. Zhang, H. Luo, and X. Lan, "DSSNet: A simple dilated semantic segmentation network for hyperspectral imagery classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 11, pp. 1968–1972, Nov. 2020.
[25] Y. K. Takehisa Yairi, "Telemetry-mining: A machine learning approach to anomaly detection and fault diagnosis for space systems," in *Proc. 2nd IEEE Int. Conf. Space Mission Challenges Inf. Technol. (SMC-IT)*, Jun. 2006, p. 8.
[26] S. Chen and H. Wang, "SAR target recognition based on deep learning," in *Proc. Int. Conf. Data Sci. Adv. Analytics (DSAA)*, Oct. 2014, pp. 541–547.
[27] A. Diab, R. Kashef, and A. Shaker, "Deep learning for LiDAR point cloud classification in remote sensing," *Sensors*, vol. 22, no. 20, p. 7868, Oct. 2022.
[28] D. Hong et al., "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021.
[29] R. C. Shinde, S. S. Durbha, and A. V. Potnis, "LiDARCSNet: A deep convolutional compressive sensing reconstruction framework for 3D airborne LiDAR point cloud," *ISPRS J. Photogramm. Remote Sens.*, vol. 180, pp. 313–334, Oct. 2021.
[30] R. C. Shinde and S. S. Durbha, "Deep convolutional compressed sensing-based adaptive 3D reconstruction of sparse LiDAR data: A case study for forests," *Remote Sens.*, vol. 15, no. 5, p. 1394, Mar. 2023.
[31] A. E. Siegman, *Lasers*. Sausalito, CA, USA: Univ. Science Books, 1986.
[32] D. J. Harding, M. A. Lefsky, G. G. Parker, and J. B. Blair, "Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests," *Remote Sens. Environ.*, vol. 76, no. 3, pp. 283–297, Jun. 2001.
[33] L. Duncanson et al., "Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma county, California," *Remote Sens. Environ.*, vol. 242, Jun. 2020, Art. no. 111779.
[34] W. Qi, S. Saarela, J. Armston, G. Ståhl, and R. Dubayah, "Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI LiDAR data," *Remote Sens. Environ.*, vol. 232, Oct. 2019, Art. no. 111283.
[35] J. Carter et al., "An introduction to LiDAR technology, data, and applications," *NOAA Coastal Services Center*, vol. 2, p.17, Nov. 2012.
[36] S. A. Sulaiman, A. K. A. A. Ismail, and M. N. F. N. Azman, "Scattering effects in laser attenuation system for measurement of droplet number density," *Energy Proc.*, vol. 50, pp. 79–86, Jan. 2014.
[37] T. Fahey, M. Islam, A. Gardi, and R. Sabatini, "Laser beam atmospheric propagation modelling for aerospace LiDAR applications," *Atmosphere*, vol. 12, no. 7, p. 918, Jul. 2021.
[38] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[39] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 105–115, Jan. 2014.
[40] Y. Wu, P. Ye, I. O. Mirza, G. R. Arce, and D. W. Prather, "Experimental demonstration of an optical-sectioning compressive sensing microscope (CSM)," *Opt. Exp.*, vol. 18, no. 24, pp. 24565–24578, 2010.
[41] C. V. Correa, H. Arguello, and G. R. Arce, "Spatiotemporal blue noise coded aperture design for multi-shot compressive spectral imaging," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 33, no. 12, pp. 2312–2322, 2016.
[42] T. S. Rao, G. R. Arce, and J. P. Allebach, "Analysis of ordered dither for arbitrary sampling lattices and screen periodicities," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 11, pp. 1981–2000, 1990.
[43] D. L. Lau, R. Ulichney, and G. R. Arce, "Blue- and green-noise halftoning models," *IEEE Signal Process. Mag.*, vol. 20, no. 4, pp. 28–38, Jul. 2003.
[44] D. L. Lau, G. R. Arce, and N. C. Gallagher, "Digital halftoning by means of green-noise masks," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 16, no. 7, pp. 1575–1586, 1999.
[45] J. Tan, Y. Ma, H. Rueda, D. Baron, and G. R. Arce, "Compressive hyperspectral imaging via approximate message passing," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 389–401, Mar. 2016.
[46] H. Arguello, C. V. Correa, and G. R. Arce, "Fast lapped block reconstructions in compressive spectral imaging," *Appl. Opt.*, vol. 52, no. 10, pp. 35–45, 2013.

Andres Ramirez-Jaime (Graduate Student Member, IEEE) received the B.S. and M.Sc. degrees in electronic engineering from Universidad de Los Andes, Bogotá, Colombia, in 2013 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA.

He was a Visiting Scientist at the Institute of Robotics and Industrial Computing, Technical University of Catalonia, Barcelona, Spain. Before joining the University of Delaware, he was a Professor with the Faculty of Engineering, University of La Sabana, Cundinamarca, Colombia, where he was the Leader of the UNISABANA Herons SPL Robotics Team and part of the UNISABANA Herons Ventilator Team. His research interest include generative models, computer vision, and deep learning.

Karelia Pena-Pena received the B.Sc. degree in electrical engineering from Universidad de Los Andes, Mérida, Venezuela, in 2017, and the M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Delaware, Newark, DE, USA, in 2020 and 2023, respectively.

Her research interests include graph signal processing, natural language processing, computer vision, machine learning, and optimization.

Gonzalo R. Arce (Life Fellow, IEEE) is currently the Charles Black Evans Distinguished Professor of electrical and computer engineering and a J. P. Morgan-Chase Senior Faculty Fellow with the Institute of Financial Services Analytics, University of Delaware, Newark, DE, USA. He holds 25 U.S. patents and is the coauthor of four books. His research interests include computational imaging, data science, and machine learning.

Dr. Arce held the 2010 and 2017 Fulbright-Nokia Distinguished Chair of Information and Communications Technologies with Aalto University, Espoo, Finland. He was a recipient of an NSF Research Initiation Award. He was a elected Fellow of the OPTICA, SPIE, AAIA, and the National Academy of Inventors.

David Harding received the B.S. and Ph.D. degrees in geological sciences from Cornell University, Ithaca, NY, USA, in 1980 and 1988, respectively.

Since 1991, he has been a Research Scientist with the NASA Goddard Space Flight Center, Greenbelt, MD, USA, where he was a member of the Planetary Geodynamics Laboratory in the Sciences and Exploration Directorate, and is currently a member of the Biospheric Sciences Laboratory. He is the Principal Investigator for the airborne Slope Imaging Multi-polarization Photo-Counting LiDAR. He was a member of the Ice, Cloud, and Land Elevation Satellite Science Team and ICESat-2 Science Definition Team. His research interests include topographic expression of land surface processes, structure of vegetation, and physical properties of snow, ice, and water, in particular by developing and using advanced airborne and spaceflight laser altimeter systems.

Mark Stephen received the B.S. degree in physics from the University of Delaware, Newark, DE, USA, in 1992, and the M.S. and Ph.D. degrees in applied physics from the University of Maryland, Baltimore, MD, USA, in 2003 and 2008, respectively.

He has worked at NASA's Goddard Space Flight Center, Greenbelt, MD, USA, since 1991, developing laser and electro-optics technologies for space-based applications. His research activities have included lasers, optical components, and laser instruments. He developed diode-pumped solid-state laser systems with an emphasis on laser diode array pumps and space qualification of these components, use of fiber amplifier technology for gas detection, and remote sensing using laser spectroscopy and new laser architectures. He worked on several satellite programs including the Geoscience Laser Altimeter System (GLAS), the Mercury Laser Altimeter (MLA), and the Lunar Orbiter Laser Altimeter (LOLA). He was the Product Development Lead for the Advanced Technology Laser Altimeter System (ATLAS) Laser, which is currently flying aboard ICESat-2. He is currently working on photonic integrated circuits and mapping LiDARs for space applications.

James MacKinnon received the B.S. and M.S. degrees in computer engineering from the University of Florida, Gainesville, FL, USA, in 2014 and 2016, respectively.

He is a Computer Engineer at the NASA Goddard Space Flight Center, Greenbelt, MD, USA, in the Science Data Processing Branch. His most recent work include developing the payload data processing FPGA design for the NASA-developed CeREs CubeSat and being a Principal Investigator on an Internal Research and Development Project with the goal of designing a neural network for detecting wildfires from multispectral imagery. His expertise includes FPGA design for high-performance, space-based data processing systems, machine learning for science data, and reliable software design.