ISPRS Journal of Photogrammetry and Remote Sensing 235 (2026) 363–382

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

From canopy to ground via ForestGen3D: Learning cross-domain generation of 3D forest structure from aerial-to-terrestrial LiDAR

Juan Castorena [a,*], E. Louise Loudermilk [b], Scott Pokswinski [c], Rodman Linn [a]

[a] Los Alamos National Laboratories, Los Alamos, NM, 48124, USA
[b] Southern Research Station, Disturbance and Prescribed Fire Laboratory, 320 E Green St., Athens, GA 30606, USA
[c] New Mexico Consortium, 4200 W Jemez Rd 200, Los Alamos, NM 87544, USA

ARTICLE INFO

Keywords:
Forests
3D generation
Artificial intelligence
Terrestrial
Aerial
LiDAR
Denoising diffusion
Remote sensing
Wildfires

ABSTRACT

The 3D structure of living and non-living components in ecosystems plays a critical role in determining ecological processes and feedbacks from both natural and human-driven disturbances. Anticipating the effects of wildfire, drought, disease, or atmospheric deposition depends on accurate characterization of 3D vegetation structure, yet widespread measurement remains prohibitively expensive and often infeasible. We present ForestGen3D, a cross-domain generative framework that preserves ALS observed 3D forest structure while inferring missing sub-canopy detail. ForestGen3D is based on conditional denoising diffusion probabilistic models trained on co-registered ALS and terrestrial LiDAR (TLS) data, exploiting the empirical containment of TLS structure within the ALS derived spatial envelope. The model generates realistic TLS-like point clouds that remain spatially consistent with ALS geometry while smoothly reconstructing canopy interior, mid-story, and near-ground structure, enabling landscape-scalable reconstruction of full vertical forest structure. To ensure ecological plausibility, we introduce a geometric containment prior based on the convex hull of ALS observations and provide theoretical and empirical guarantees that generated structures remain spatially consistent. We evaluate ForestGen3D at tree, plot, and landscape scales using real-world data from mixed conifer ecosystems, and show through qualitative and quantitative geometric and distributional analyses that it produces high-fidelity reconstructions closely matching TLS reference data in terms of 3D structural similarity and downstream biophysical metrics, including tree height, DBH, crown diameter, and crown volume. We further demonstrate that the resulting expected point containment (EPC) serves as a practical proxy for generation quality in settings where TLS ground truth is unavailable. Our results demonstrate that ForestGen3D enhances the utility of ALS only environments by inferring ecologically plausible sub-canopy structure while faithfully preserving the landscape heterogeneity encoded in ALS observations, thereby providing a richer 3D representation for ecological analysis, structural fuel characterization and related remote sensing applications.

1. Introduction

Wildfires, drought, pathogens, and other natural disturbances are increasingly transforming forested landscapes, driving the need for improved ecological monitoring and fuel structure characterization. These disturbances are shaped not only by vegetation type but also by the three-dimensional (3D) arrangement of biomass across vertical and horizontal layers. The 3D structure of both living and non-living components of vegetation plays a critical role in regulating ecological processes, including energy and moisture fluxes, carbon storage, and fire behavior (Vierling et al., 2008; Kane et al., 2010; Loudermilk et al., 2017; Marcozzi et al., 2025). However, large-scale collection of detailed 3D vegetation structure remains a major bottleneck. Most landscape-scale ecological assessments still rely on field plots that are statistically upscaled using summary metrics (Bechtold and Patterson, 2005). To spatially represent forests, synthetic layouts are often generated by populating plots with idealized geometric models (e.g., cylinders, ovals), randomly placed according to field statistics (Köhler and Huth, 1998; Dufour-Kowalski et al., 2012; de Wergifosse et al., 2022; Maréchaux et al., 2021). While informative, these methods often overlook spatial relationships between vegetation components and fail to capture the full 3D spatial heterogeneity, both of which are crucial for modeling dynamic ecological processes. A related class of methods common in visualization, and gaming applications uses procedural forest generation, where tree forms and vegetation distributions are instantiated from

* Corresponding author.
E-mail address: jcastorena@lanl.gov (J. Castorena).

https://doi.org/10.1016/j.isprsjprs.2026.03.009
Received 19 September 2025; Received in revised form 21 January 2026; Accepted 9 March 2026
Available online 18 March 2026
0924-2716/© 2026 Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS).

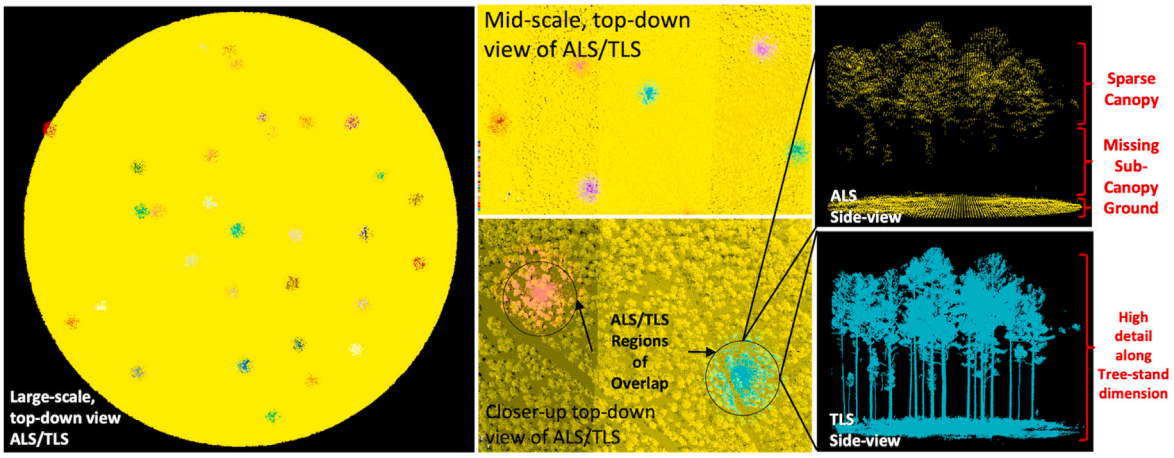


Fig. 1. Characteristics of ALS (in yellow) and TLS data at a surveying forest site, covering 2.5 km in diameter with 31 TLS scans randomly distributed (a color-coded blob per scan). Overlapping ALS in top-right shows a side-view point cloud with sparse/missing information below the canopy. Co-registered TLS in the bottom-right, on the other hand shows high detail along the tree-stand direction but is limited in the lat-long direction due to physical range limitations of TLS (as shown in the top-down images). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

parametric rules or stylized templates. Tools such as SpeedTree (Interactive Data Visualization, Inc., 2024) and Unity's Terrain Engine (Technologies, 2024) exemplify this approach, offering efficient generation of stylized vegetation for rendering purposes. While these models can quickly populate large virtual landscapes, they typically prioritize visual coherence and rendering performance over fidelity with real-world forest conditions, and are therefore less suitable for scientific analyses that depend on detailed and ecologically grounded 3D representations of forest fuel structure.

Over the past two decades, 3D remote sensing technologies particularly terrestrial laser scanning (TLS) (Pokswinski et al., 2021) and aerial laser scanning (ALS) (Hyyppä et al., 2008) have become increasingly valuable for vegetation mapping. TLS offers high-resolution 3D point clouds of stems, understory, and ground structure but is constrained to plot-scale deployment due to cost and labor demands. ALS, by contrast, provides broad landscape coverage and effectively captures canopy structure but suffers from occlusion and fails to resolve sub-canopy features in dense forests, even in full-waveform systems (Nelson et al., 1984; Mallet and Bretar, 2009; Castorena and Creusere, 2015). These limitations are illustrated in Fig. 1: while ALS provides landscape-scale coverage of the upper canopy, TLS offers detail below the canopy but lacks coverage in the lateral/longitudinal directions. Surprisingly, most methods in forest mapping and monitoring do not exploit the complementary nature of these two sensing modalities. Instead, they typically rely on either ALS or TLS independently, developing analysis pipelines tailored to a single data source. This leads to structural blind spots for instance, missing understory in ALS-based workflows or extrapolated spatial sparsity in TLS-based ones that fail to respect landscape heterogeneity and that persist in operational products and ecological models (Calders et al., 2020; Hyyppä et al., 2012).

In this work, we propose to bridge these sensing gaps by learning generative models of 3D fuel structure that leverage the strengths of both data sources. To achieve this, we develop ForestGen3D, a cross-domain generative modeling framework based on denoising diffusion probabilistic models (DDPMs). DDPMs are a class of deep generative models that learn the underlying data distribution by gradually adding noise to training examples and then learning to reverse this process through iterative denoising (Ho et al., 2020; Song et al., 2020). Once trained, the model can synthesize new samples by starting from pure Gaussian noise and progressively refining structure in a manner consistent with the learned distribution. In our case, the DDPM is conditioned on ALS inputs so that the denoising process respects the geometric information encoded in airborne LiDAR while generating plausible sub-canopy structure. These models are trained here on a new dataset of co-registered ALS/TLS tree 3D point clouds referred to as the *CoLiDAR Forest3D (ALS/TLS)* and collected across mixed conifer vegetation ecosystems. Once trained, the model can generate novel 3D vegetation samples by starting from random noise and iteratively refining structure through a learned denoising process, guided by ALS input as shown in Fig. 2. To the best of our knowledge, this is the first approach to leverage conditional diffusion models for cross-domain generation of 3D forest structure, bridging the gap between sparse, landscape-scale ALS and detailed, plot-scale, TLS representations. This enables scalable ecological modeling of forests in full vertical profile beyond existing upsampling or procedural simulation methods.

Existing 3D generative and completion models, such as PCN (Yuan et al., 2018), 3D-GAN (Wu et al., 2016), latent-GAN (Achlioptas et al., 2018), and PointFlow (Yang et al., 2019), have so far been developed and tested primarily on man-made objects from synthetic datasets such as ShapeNet. While effective in those controlled domains, their suitability for modeling real-world forest vegetation remains largely unknown. Forest structure presents fundamentally different challenges: complex topologies with intertwined crowns, occlusion-driven sparsity due to canopy cover, and vertical heterogeneity across understory, stems, and crowns. These characteristics may degrade the performance of models designed for compact, uniformly sampled shapes. Moreover, existing methods are typically trained without reference to real-world sensing modalities such as LiDAR, and therefore do not account for the anisotropic density, partial observability, and sensor-specific artifacts present in ecological data. In contrast, the objective of our approach is to generate ecologically realistic 3D vegetation structure directly from ALS/TLS data collected in mixed conifer forests, capturing both the natural variability of vegetation and the sensor dependent effects inherent in terrestrial and airborne LiDAR measurements. In the sections that follow, we present our dataset and training approach, describe the DDPM-based model, and evaluate ForestGen3D across tree, plot, and landscape scale experiments. Quantitative and qualitative results demonstrate that the proposed approach approximates TLS-level structural detail using ALS data alone.

2. ForestGen3D

The core problem we address is generating synthetic forests by improving the fidelity of individual tree structures beyond conventional coarse-shape models. To achieve this, we explicitly preserve ecosystem fidelity by conditioning the generation process on ALS inputs, which

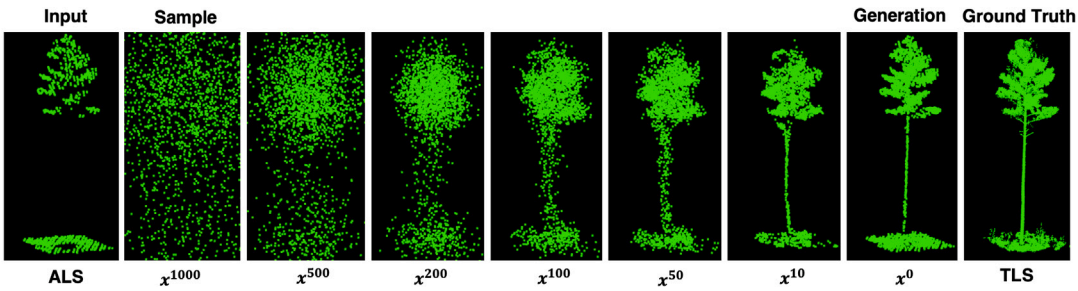


Fig. 2. ForestGen3D based on denoising diffusion probabilistic model (DDPM) formulation. Model iteratively denoises a sample guided by ALS measurements and generates a sample of forest tree 3D structure from the learned training distribution. Note that the 3D structural information provided by ALS is effective for constraining 3D structural generation in the vertical tree stand direction.

provide canopy level structure tied to real forest layouts. Rather than modeling entire plots or landscapes directly which can compromise fine-grained detail and impose significant computational demands, we focus on learning 3D structure at the tree scale, where the physical representation of vegetation is most accurate. This localized modeling approach allows for more efficient training and data preparation, reducing the need to capture broad ecosystem variation in a single model. Moreover, it enables scalable deployment by applying the generative model independently to each detected tree at the plot or landscape scale, under the assumption that only regions beneath vegetation are susceptible to occlusion in ALS data. To support this effort, we constructed a 3D forest dataset of co-registered ALS/TLS tree samples, which we used to train cross-domain denoising diffusion probabilistic models (DDPMs). These models learn the statistical distribution of tree 3D structure in a target domain (e.g., TLS) conditioned on observations in a reference domain (e.g., ALS). DDPMs start from random Gaussian noise and iteratively refine structure through a learned denoising process, enabling flexible and efficient generation of high-resolution 3D tree forms. Fundamentally, the framework learns to generate samples from the TLS distribution conditioned on a measured sample from the ALS distribution. Fig. 2 illustrates this process: an ALS input guides the denoising steps, which progressively reconstruct the full 3D structure of a tree, including stem and understory geometry missing from the input.

2.1. Training dataset

We construct a new 3D forest dataset *CoLiDAR-Forest3D (ALS/TLS) dataset*, containing 2900 co-registered TLS/ALS tree scans for training, including both individual trees and clusters of closely packed trees. The dataset was constructed using the co-registration approach from Castorena et al. (2025) to align TLS scans with overlapping ALS data and the deep-learning based tree detection method from Windrim and Bryson (2019) to extract individual trees. The raw data were collected from a mixed-conifer forest at Fort Stewart in the United States (Snider et al., 2025), primarily featuring species such as longleaf pine, slash pine, loblolly pine, and turkey oak. The understory consists of graminoids, forbs, and shrubs, including wiregrass, gallberry, and saw palmetto. TLS data was acquired using a Leica BLK360 LiDAR scanner with an 830 nm wavelength laser and a field of view (FOV) of 300° (vertical) and 360° (horizontal). Scans were collected using a single-scan acquisition strategy, in which the instrument is deployed once at a plot centered position rather than through multiple co-registered scans. Although single viewpoint TLS does not fully eliminate occlusion, especially in dense forest understories, this strategy has been shown to be both effective and efficient for characterizing tree and plot level structural attributes across a wide range of forest ecosystems. Prior work demonstrates that randomized single-scan deployments provide reliable summary estimates of forest characterizations, while enabling practical data collection at larger spatial and ecological extents (Pokswinski et al., 2021; Hawley et al., 2018; Loudermilk et al., 2009). Consequently, our dataset prioritizes broader ecological coverage and diversity of vegetation conditions rather than full multi-view TLS completeness, aligning with the objective of training a generative model that generalizes across landscape heterogeneity. Scans were taken from static tripod positions, with 240 scans randomly distributed across a 2.5 km diameter area. Each TLS scan contained approximately 8 million points, with around 4.6 million from the ground. Since most returns were concentrated within a 25 meter range, points beyond this threshold were removed to mitigate the non-uniform point density. A representative sampling pattern of TLS scans across a region is shown in Fig. 1, where each TLS scan is represented by a color-coded blob. ALS data was collected using a Galaxy T2000 LiDAR sensor mounted on a fixed-wing aircraft. The sensor operates at a 1064 nm wavelength and captures up to eight returns per pulse, with a 28° FOV, flown at an altitude of approximately 1675 m. The system achieves a point density of approximately 15 points per square meter and maintains an accuracy of 0.03 to 0.25 m RMSE within a 150–6500 meter range. The surveyed area was covered using multiple overlapping flight lines, including seven cross strips to reduce point sparsity and enhance coverage. Unlike TLS, ALS points are not affected by satellite signal occlusion and were globally aligned using a Trimble CenterPoint RTX RTK system for centimeter-level accuracy. A regional scale ALS point cloud is shown in Fig. 1 in yellow, with orthographic (top-down) views on the left and middle and a side-view on the top right.

Each TLS scan was initially positioned using GPS-based estimates in global coordinates. These positions were used to extract a corresponding 50 meter radius ALS pointcloud subset around each TLS scan center. The automatic and targetless co-registration approach from Castorena et al. (2025) was applied to refine the position provided by GPS while also aligning the TLS scan to the ALS point-cloud relative to rotations. Each co-registered pair was manually verified by an expert to ensure qualitative alignment between spatial features. Tree detection was performed using the deep learning (DL) based method from Windrim and Bryson (2019), which detects trees or tightly clustered groups of trees/vegetation by fitting 2D bounding boxes from a top-down bird's-eye view and then expanding each bounding box to 3D informed by the lowest-to-highest point elevation difference. This approach ensures that the resulting 3D bounding boxes capture the full vertical extent of vegetation, from the top of the canopy to the lowest ground point. These bounding boxes were used to extract tree-level point clouds by truncating the points in the ALS and TLS scans inside each 3D box. To mitigate the non-uniform point density in TLS, we truncated each point cloud to a 25 m radius and downsampled each TLS scan to a 0.10 cm voxel size. Each tree scan was then centered and normalized within the $[0,1]^3$ range. The dataset was manually curated to remove incomplete trees (e.g., those significantly cut by bounding boxes), and incorrect detections. The final dataset consists of 2900 TLS/ALS tree or tree-cluster pairs. Fig. 3 illustrates nine ALS/TLS example pairs, with ALS trees in yellow (top row) and corresponding TLS scans in aqua (bottom row). For validation, we construct a set that consists of 300 trees sampled from plots that are geographically distinct from those used for training and testing. These trees span the same range

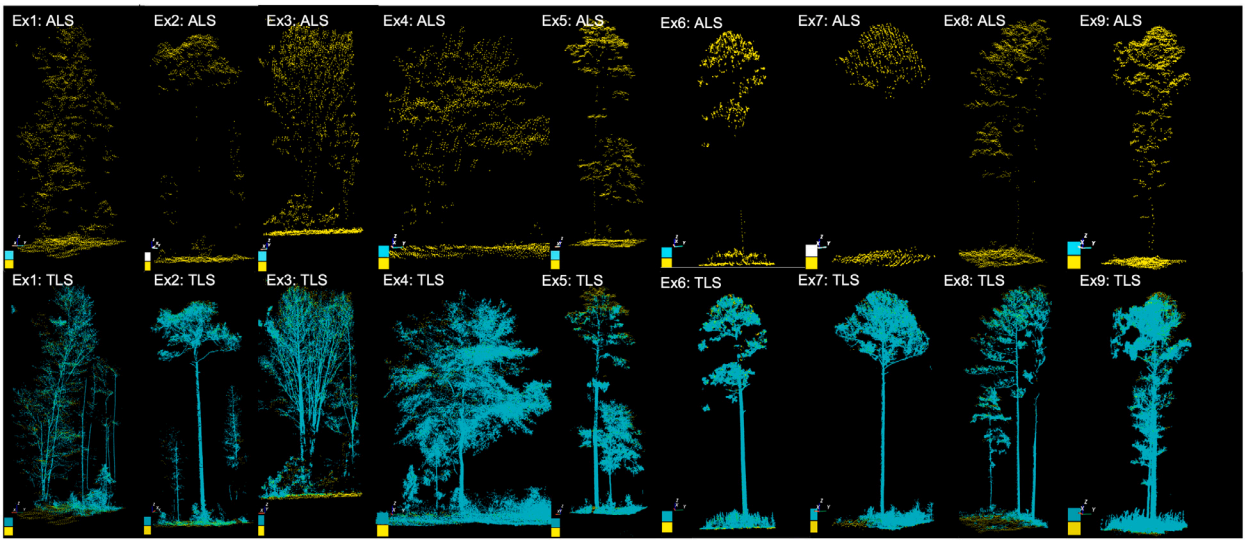


Fig. 3. 9 examples of the 3D training dataset consisting of 2900 examples of tree-segmented and co-registered aerial lidar scanning (ALS)/terrestrial lidar scanning (TLS) trees. ALS trees in yellow in the top row, corresponding co-registered TLS trees in aqua color in the bottom row.

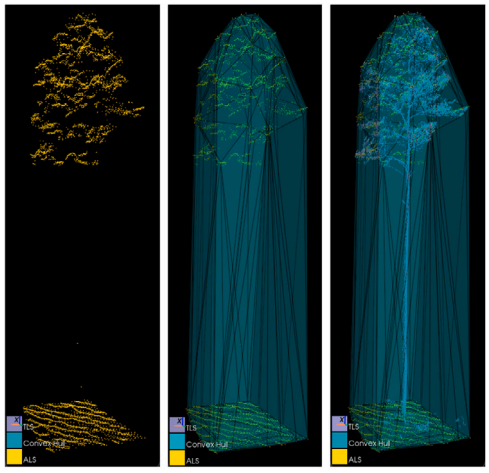


Fig. 4. Empirical geometric relationship between ALS and TLS point clouds. ALS (in yellow) captures upper-canopy geometry, forming a sparse convex envelope (in green), while TLS points (in purple) representing stems, branches, and understory are largely contained within this envelope. This empirical property motivates the containment assumption exploited in this work when learning cross-domain mappings between ALS and TLS data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of mixed-conifer vegetation types but originate from separate ALS/TLS co-registered plots, ensuring that the model is evaluated on structurally similar but spatially independent examples. The validation set is used exclusively for hyperparameter tuning, monitoring reconstruction error during training, and selecting the best checkpoint under an early stopping criterion. Note that the test set is held out entirely until final evaluation, and no samples in the validation set appear in the training or test splits.

Due to the physical operation principles of LiDAR, TLS and ALS yield structurally complementary but distributionally distinct 3D point cloud characteristics. ALS captures forest structure from an aerial, top-down perspective and provides relatively complete sampling of the upper canopy features, while TLS operates from the ground and more effectively resolves stems, understory and internal structure. This leads to a broader spatial support for ALS compared to TLS. Let $\mathbf{X} \in \mathbb{R}^{N\times 3}$ and $\mathbf{Y} \in \mathbb{R}^{M\times 3}$ denote the TLS and ALS point clouds for an individual tree (or small group of trees), and let $\mathcal{B}_{\text{ALS}} = \text{Conv}(\mathbf{Y})$ be the convex hull (i.e., convex envelope) of the ALS data. *Assumption:* We adopt a localized containment assumption at the individual tree scale, where TLS points are expected to lie within the convex envelope of their corresponding co-registered ALS points with high probability (see Fig. 4). This is expressed as:

$$\mathbb{E}_{(\mathbf{X},\mathbf{Y})\sim D}\left[\mathbb{E}_{\mathbf{x}\sim\mathbf{X}}[1(\mathbf{x}\in\mathcal{B}_{\text{ALS}})]\right] \geq 1-\epsilon \tag{1}$$

for a small constant $\epsilon \ll 1$, with $(\mathbf{X},\mathbf{Y}) \sim D$ drawn from the dataset of co-registered tree level TLS/ALS pairs. Here, $\mathbf{x} \sim \mathbf{X}$ are 3D points uniformly sampled from the TLS point cloud, and $1(\cdot)$ is the indicator function. In our empirical analysis of the training *CoLIDAR-Forest3D (ALS/TLS)* dataset, approximately 3.9% of TLS points fall outside the ALS convex hull, i.e., $\epsilon = 0.04$, validating the assumption at the tree scale. We emphasize that this prior is not assumed to hold globally at the plot or landscape scale and that a variety of factors may violate containment. However, the small deviations from containment that may be locally present are expected and well justified, as they can result from sensor misalignment, resolution differences, vegetation visible from the ground but occluded from the aerial view, and inherent measurement noise in both ALS and TLS systems. Despite these localized discrepancies, the overall containment expectation introduces a geometric prior that is implicitly learned by our conditional generative model. This prior not only helps constrain synthesis within plausible ALS-guided structure, but also provides a useful means of measuring the spatial fidelity of the generated points relative to the ALS envelope.

2.2. 3D cross-domain denoising diffusion generative models

Denoising diffusion probabilistic models are generative deep neural networks that operate within a variational inference framework (Ho et al., 2020). Their primary objective is to learn an approximation of the underlying training data distribution while enabling flexible and efficient sample generation from this learned distribution. In our context, we adopt DDPMs in a *cross-domain* setting, where the goal is to generate TLS-like 3D vegetation structure conditioned on co-registered ALS measurements. This enables the synthesis of dense sub-canopy and near-ground details in regions where ALS data is sparse or occluded. To this end, we train the model to approximate the conditional distribution $p(\mathbf{X} \mid \mathbf{Y})$, where ALS point clouds represent sparse, canopy dominated

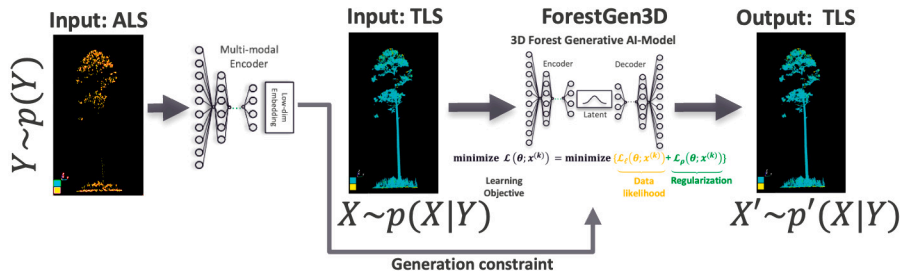


Fig. 5. Schematic of ALS-to-TLS cross-domain ForestGen3D. The forward process encodes a TLS sample into Gaussian noise, while the reverse process learns to denoise a sample conditioned on ALS input following a denoising diffusion probabilistic model (DDPM) formulation, reconstructing high-resolution 3D vegetation structure.

measurements, and TLS point clouds provide detailed ground-to-canopy interior structural information. During inference, the model begins from isotropic Gaussian noise and iteratively refines the sample using structural cues provided by ALS input, producing 3D point clouds that resemble high-resolution TLS samples respecting ALS fidelity.

The DDPM architecture consists of two components: a forward diffusion process that progressively corrupts the input point cloud $\mathbf{X}^0 \in \mathbb{R}^{N\times 3}$ over T steps by multiplying it with a decaying factor while adding Gaussian noise, and a learned reverse process that denoises the corrupted sample back to a clean reconstruction. This yields a training trajectory: $\mathbf{X}^0 \to \mathbf{X}^1 \to \cdots \to \mathbf{X}^T \sim \mathcal{N}(0, \mathbf{I})$, as the forward (diffusion) process, $\mathbf{X}^T \to \mathbf{X}^{T-1} \to \cdots \to \hat{\mathbf{X}}^0$ as the reverse (denoising) process. The model conditions each denoising step on ALS observations $\mathbf{Y}$, enabling it to learn structure consistent with the canopy geometry while generating plausible TLS content beneath. The generative backbone used in ForestGen3D corresponds to the Conditional Generation Network (CGNet) architecture in Lyu et al. (2021), which enables us to condition denoising based on ALS point cloud measurements. CGNet consists of three principal components: a PointNet++ based denoising network (Qi et al., 2017), a ConditionNet that encodes ALS structure, and a FeatureMapper that fuses both representations. The denoising network follows a hierarchical, U-Net architecture that extracts features from noisy point clouds using multi-scale grouping and radius based neighborhood search. Four hierarchical levels are used, employing radii of 0.1, 0.2, 0.4, and 0.8 m with 32 neighbors sampled at each scale. Feature dimensions progress through $\{32, 64, 128, 256, 512\}$, and residual multi layer perceptron (MLP) blocks with skip connections are used to improve stability across diffusion steps. The diffusion timestep is embedded using a 128 dimensional positional encoding and concatenated to point level features. ALS-based conditioning is provided through the ConditionNet, which mirrors the hierarchical structure of the denoising network and extracts both local and global structural features from the ALS point cloud. Its feature dimensions follow $\{32, 32, 64, 64, 128\}$, enabling the network to capture the vertical distribution patterns, and scan sparsity. The resulting ALS features are aligned with the denoising hierarchy and supplied to the FeatureMapper, which integrates noisy point cloud features and ALS information using FiLM-style modulation (Perez et al., 2018). This module applies shared MLP feature transforms, radius-based grouping, and an attention mechanism with grouped feature transformations and batch normalized activations. Together, these components ensure that the reverse diffusion process is guided both by ALS derived structure and by the learned statistical distribution of TLS geometry. A high-level schematic of the model architecture is shown in Fig. 5.

The model is trained on point clouds $\mathbf{X}, \mathbf{Y} \sim p_{data}(\mathbf{X}, \mathbf{Y}) \in \mathbb{R}^{N\times 3} \times \mathbb{R}^{M\times 3}$, sampled from aligned TLS/ALS pairs. The forward diffusion process follows a first-order Markov chain by adding Gaussian noise over T steps (e.g., $T = 1000$). The reverse process learns to approximate $p(\mathbf{X}^{t-1} \mid \mathbf{X}^t, \mathbf{Y})$, thereby denoising the latent variable back to a plausible TLS point cloud. Fig. 2 illustrates the generative process in a leave-one-out setting, where a withheld TLS scan is reconstructed using only its associated ALS input. The figure shows the denoising trajectory starting from an initial sample $\mathbf{X}^T \sim \mathcal{N}(0, \mathbf{I})$ and progressively refining toward a tree structure that closely approximates the held-out TLS reference. Notably, the model is able to synthesize sub-canopy features that are missing from the ALS input, highlighting the DDPM's capacity to infer occluded structural details.

Remark (*Containment Consistency Under ELBO*)**.** Because the conditional DDPM is trained via ELBO minimization (Ho et al., 2020) to approximate the true conditional distribution $p_{\text{data}}(\mathbf{X} \mid \mathbf{Y})$, and the training data satisfies an approximate containment condition at the tree scale, we expect the learned generative model $p_\theta(\mathbf{X} \mid \mathbf{Y})$ to inherit this spatial bias.

Formally, we can bound the expected point containment (EPC) for tree-level generation as:

$$\mathbb{E}_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{X}\sim p_\theta(\cdot|\mathbf{Y})}\left[\mathbb{E}_{\mathbf{x}\sim\mathbf{X}}\left[1(\mathbf{x} \in \mathcal{B}_{\text{ALS}}(\mathbf{Y}))\right]\right]\right] \geq 1 - \epsilon - \delta \tag{2}$$

where ϵ reflects the empirical containment gap (e.g., 3.9% in our CoLiDAR Forest3D dataset), and δ accounts for model approximation error due to ELBO. Since the convex hull is computed per tree (or tight tree cluster), the spatial bias is preserved during training and inference, and we observe in practice that containment violations remain rare and localized. A full proof of Eq. (2) is provided in the Appendix.

3. Experimentation

The experiments in this section are designed to address two key questions: (1) Can conditional DDPM models effectively learn the detailed 3D structural distribution of forest vegetation from TLS data with ALS guidance? and (2) Is the approach scalable to plot-scale and landscape-scale deployment, enabling accurate synthesis of tree-level structure and improvement of downstream ecological biometrics commonly used in forest management?. The first set of experiments, presented in Section 3.2, evaluates the generative capacity of DDPMs at the tree scale. We assess whether the model's inductive biases and architecture are well-suited for representing the hierarchical and self-similar structure of vegetation. This includes analysis of the model's approximation quality via reconstruction error, as well as ablation studies that explore the sensitivity of generation quality to key DDPM hyperparameters such as diffusion steps. To quantitatively assess reconstruction error, we compute Chamfer Distance (CD) and Earth Mover's Distance (EMD) (Fan et al., 2017) between generated and reference TLS point clouds, using held-out tree samples. In addition, we examine the relationship between CD, EMD, and our expected point containment (EPC) metric. The goal of this analysis is to evaluate whether containment can serve as a practical proxy for CD and/or EMD in scenarios where ground-truth TLS data is unavailable. The second set of experiments, detailed in Sections 3.4–3.6, demonstrates the model's generalization and scalability from tree to landscape scale scenarios in mixed conifer ecosystems. At the tree and plot scales, we evaluate the consistency of ALS-guided generation in matching spatial extents and density patterns observed in ground truth TLS scans.

At the regional scale, we deploy the model across 200 meter radius ALS-only regions and assess the realism of synthesized vegetation. For both plot and landscape cases, we compute downstream ecological biometrics including tree height, diameter at breast height (DBH), crown diameter (CrD), and crown volume (CrV) based on ALS, TLS, ALS+TLS, and ALS+ForestGen3D samples. These comparisons highlight the model's ability to augment sparse ALS with realistic sub-canopy structure, leading to improved and more complete biometric estimates. Furthermore, we confirm that the containment prior is preserved in inference time landscape scale deployments, as evidenced by low out of envelope deviation statistics across diverse mixed-conifer landscapes.

3.1. Implementation details

ForestGen3D is implemented in Python/PyTorch with CUDA acceleration. All experiments were conducted on an NVIDIA Tesla V100S-PCIE-32 GB GPU, with a total training time of approximately 20 h for 100 epochs. The diffusion process follows a standard DDPM formulation, using a linear noise schedule with $\beta_0 = 1 \times 10^{-4}$, $\beta_T = 2 \times 10^{-2}$, and $T = 1000$ diffusion steps, although we also test other T settings. The model is trained using a mean squared error (MSE) loss, while the Earth Mover's Distance (EMD) and Chamfer distance metric is optionally monitored during validation to ensure that the learned denoising process maintains geometric stability. Absolute coordinates, local neighborhood features, global shape descriptors, and ALS conditioning are included throughout the denoising process. Grouped feature attention, batch normalization, and residual connections are enabled across all modules to stabilize training. Radius neighborhood search is used throughout both encoding and decoding, while a k-nearest neighbors (kNN) strategy with $k = 8$ is employed for feature propagation during decoding. At inference, this configuration produces a complete 3D tree structure in approximately 0.4 s. To support reproducibility and reuse, the ForestGen3D model implementation, trained weights, and associated configuration files are expected to be released publicly upon publication of the manuscript.

3.2. Training and validation

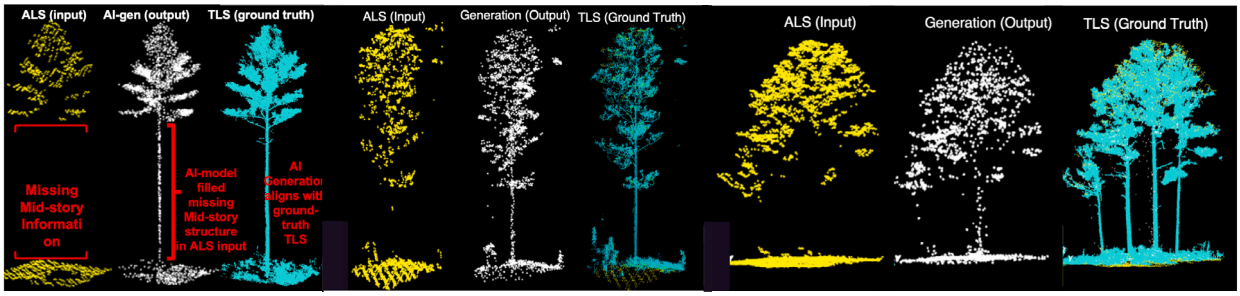
To evaluate the inductive biases and approximation fidelity of our DDPM-based generative framework, we conduct a series of ablation experiments. During training, each batch consists of 16 randomly selected ALS/TLS example pairs. Each point cloud is uniformly subsampled to a fixed size of $N = 2048$. This random subsampling enhances dataset diversity by introducing spatial variability across training epochs, enabling the model to encounter a broader range of structural configurations over time. Although the $N = 2048$ resolution may appear limiting, it applies only at the level of individual trees. When aggregated over larger spatial extents such as plots or landscapes, this per tree resolution naturally scales to yield detailed reconstructions across entire forested regions. Moreover, higher-resolution representations can be achieved through ensemble sampling, in which multiple stochastic generations are produced for the same conditional ALS input and merged coherently to create denser 3D tree structures. As an illustration of the capabilities of the generative model to approximate 3D tree structures given the existing settings, we include Fig. 6 showcasing three examples in the training set with ALS input (first column), AI-generation (second column) and ground truth TLS (third column) in each example. Here, we see that our model is able to generate samples at good qualitative approximations of the 3D structural features of TLS distributions conditional on the ALS inputs across the entire vertical tree-stand direction. Figs. 6(a) and 6(b) depict instances where the ALS input contains a single tree, which the model accurately approximates, demonstrating its ability to infer complex internal structures from sparse input. Fig. 6(c) further illustrates its performance in closely packed tree settings. Here, ground truth TLS shows four trees with tightly packed canopies, while the model generates fewer distinct trunks than the ground truth. Despite this, the output remains visually realistic and spatially consistent with what a human observer might infer from the incomplete ALS scan.

Ablation studies to analyze model reconstruction performance where conducted by varying the number of diffusion steps $T \in \{200, 400, 600, 1000, 2000\}$, diffusion parameters $\beta_0 = 10^{-4}$, $\beta_T = 0.02$, and batch size 16 constant. These experiments aim to analyze the error of the model at training time and also the spatial containment consistency in Eq. (2). We track reconstruction error across training iterations using both Chamfer Distance (CD) and Earth Mover's Distance (EMD), computed between generated and ground truth TLS point clouds on a held-out validation set of 300 tree examples. As shown in Figs. 7(a)–7(b), both CD and EMD consistently decrease with training iterations across all diffusion schedules, indicating convergence and improving fidelity. Notably, models with higher diffusion steps (e.g., $T = 1000$) exhibit slower convergence but ultimately achieve slightly lower reconstruction error, while smaller T values converge quickly with slightly higher error floors.

In parallel, we assess the containment behavior of generated point clouds using the Expected Point Containment metric, which quantifies the proportion of generated points lying within the ALS convex envelope. Fig. 8(a) shows that EPC over training iterations mirror reconstruction improvement, suggesting that containment behavior improves with model learning. To investigate the link between spatial consistency and reconstruction fidelity, we plot CD and EMD directly against EPC in Figs. 8(b) and 8(c), respectively. Across all diffusion schedules, we observe a strong inverse correlation: as EPC increases, both CD and EMD rapidly decrease, ultimately plateauing at low error values for high containment ratios. This relationship reveals an important insight: when generated trees remain well contained within the convex hull of ALS inputs, as measured by EPC, this behavior serves as a meaningful indicator of generative quality. In particular, high EPC values consistently appear alongside low reconstruction errors (e.g., CD, EMD), suggesting that EPC can serve as a practical proxy for generation quality in scenarios where ground-truth is unavailable. Moreover, the observed alignment between containment behavior and reconstruction fidelity empirically supports the theoretical guarantee presented in the remark of Section 2.2, demonstrating that the learned conditional distribution $p_\theta(\mathbf{X}|\mathbf{Y})$ inherits the spatial containment property embedded in the training data through ELBO minimization. These findings not only validate our containment assumption, but also highlight its practical utility for model evaluation in operational forest structure generation pipelines lacking TLS ground truth.

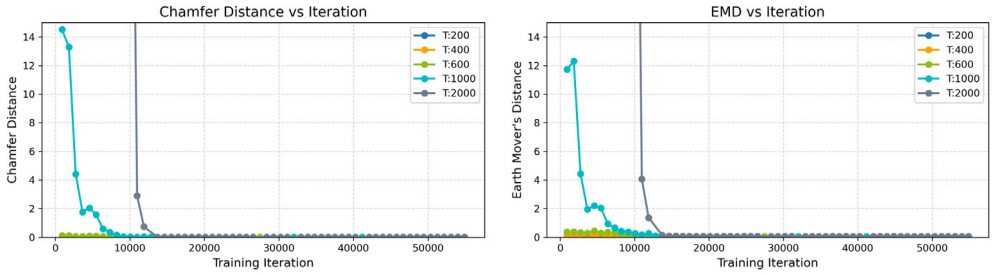
3.3. Testing dataset

For testing, we used 50 TLS scans collected from geographic regions that do not overlap with those of the training and validation areas, though sharing similar ecosystem characteristics and collected with the same sensors used for the training data. We evaluated the model's performance across three spatial scales. At the tree scale, we followed the same co-registration and tree detection procedures used during training, resulting in 1457 individual test trees. This setup enabled direct, one-to-one comparisons between generated samples and corresponding TLS ground truth reference data at high structural fidelity. At the plot scale, we constructed test plots by aggregating ALS data over spatial extents corresponding to each TLS scan, truncating ALS coverage to match the TLS field of view (e.g., 25 m in radius). This allowed for evaluation of the model's ability to reconstruct localized forest structure. Finally, at the regional scale, we assessed scalability and extrapolation by deploying the model over multiple 200 meter radius areas containing ALS data. These larger, fully independent areas tested the model's ability to synthesize realistic vegetation structure across broader, ALS only landscapes. This evaluation strategy enables comprehensive assessment of ForestGen3D's robustness and flexibility from fine-grained, tree-level detail to broad-scale generation in mixed conifer ecosystems.



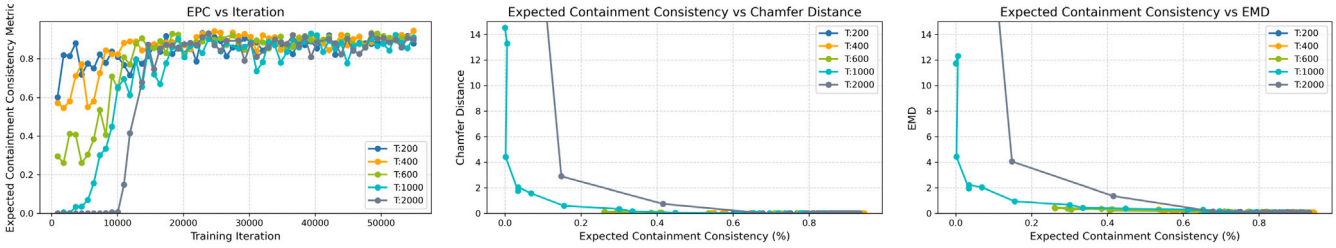
(a) Example 1. (b) Example 2. (c) Example 3.

Fig. 6. Three examples of ForestGen3D generation at the tree-scale conditioned by ALS inputs in the training set. This qualitative examples showcase the architecture's capacity to encode and represent the 3D structural details of trees.



(a) Chamfer Distance (CD) (b) Earth Movers Distance (EMD)

Fig. 7. Chamfer Distance (CD) and Earth's movers distance (EMD) errors in the validation set as a function of training iterations. Both metrics consistently decrease with training iterations across all diffusion schedules.



(a) Expected Point Containment (EPC) (b) Chamfer Distance (CD) versus EPC (c) Earth Movers Distance (EMD) versus EPC

Fig. 8. Relationship between containment behavior measured by expected point containment (EPC) and reconstruction error metrics in validation set. There is a strong inverse correlation between EPC and error metrics across all denoising time-step schedules T: as EPC increases, both Chamfer Distance (CD) and Earth's movers distance (EMD) rapidly decrease.

3.4. Tree-scale ALS-guided generation

In Fig. 9 we include three representative qualitative comparisons from the test set to illustrate differences between completion/generation approaches across a range of tree structures. First two columns across all three cases represent ALS input and TLS ground truth, while the remaining columns correspond to the completion/generation resulting from the application of different trained models. In all three examples, PCN (Yuan et al., 2018) tends to oversmooth crowns, recovering coarse silhouettes while missing high-frequency branching detail. The other baseline generative models (3D-GAN (Wu et al., 2016), latent-GAN (Achlioptas et al., 2018), and PointFlow (Yang et al., 2019)) produce shapes that appear closer to the TLS shape, yet are less consistent across examples in general, in comparison to those from our ForestGen3D model. For instance, in the second example, latent-GAN (l) yields a form whose stem does not resemble the straight trunk structure of the TLS reference. Similarly, PointFlow in (f) generates a discontinuous tree stem of larger a diameter compared to ground truth TLS. 3D-GAN and ForestGen3D produce consistently closer results to TLS with canopy structure close to that of ALS and coherent stem-to-crown transitions, recognizable branching patterns, and overall shapes that better match human expectations of realistic tree architecture. These qualitative outcomes mirror the quantitative CD and EMD error, and EPC consistency results below each figure, highlighting the capabilities of generative models to produce realistic structural properties of forests. Note also, that the generated points are generally tightly packed and bounded by the ALS canopy points. This behavior stems from the inherent nature of the ALS training data, where the distribution of the ALS convex-hull is typically broader and encompasses the finer details captured by TLS. Consequently, the generated points do not generally escape the span of the ALS canopy diameter. From a statistical standpoint, as the generative approach aims to produce samples from a distribution that is largely encompassed by the ALS input, it is inherently constrained, making it less likely for the generated points to statistically escape the overall ALS convex envelope.

Table 1 shows a quantitative evaluation summarizing the CD/EMD and EPC metrics over the test dataset of 1457 test examples. This table summarizes the resulting metrics for the completion PCN (Yuan et al., 2018) method and the generative models including 3D-GAN (Wu

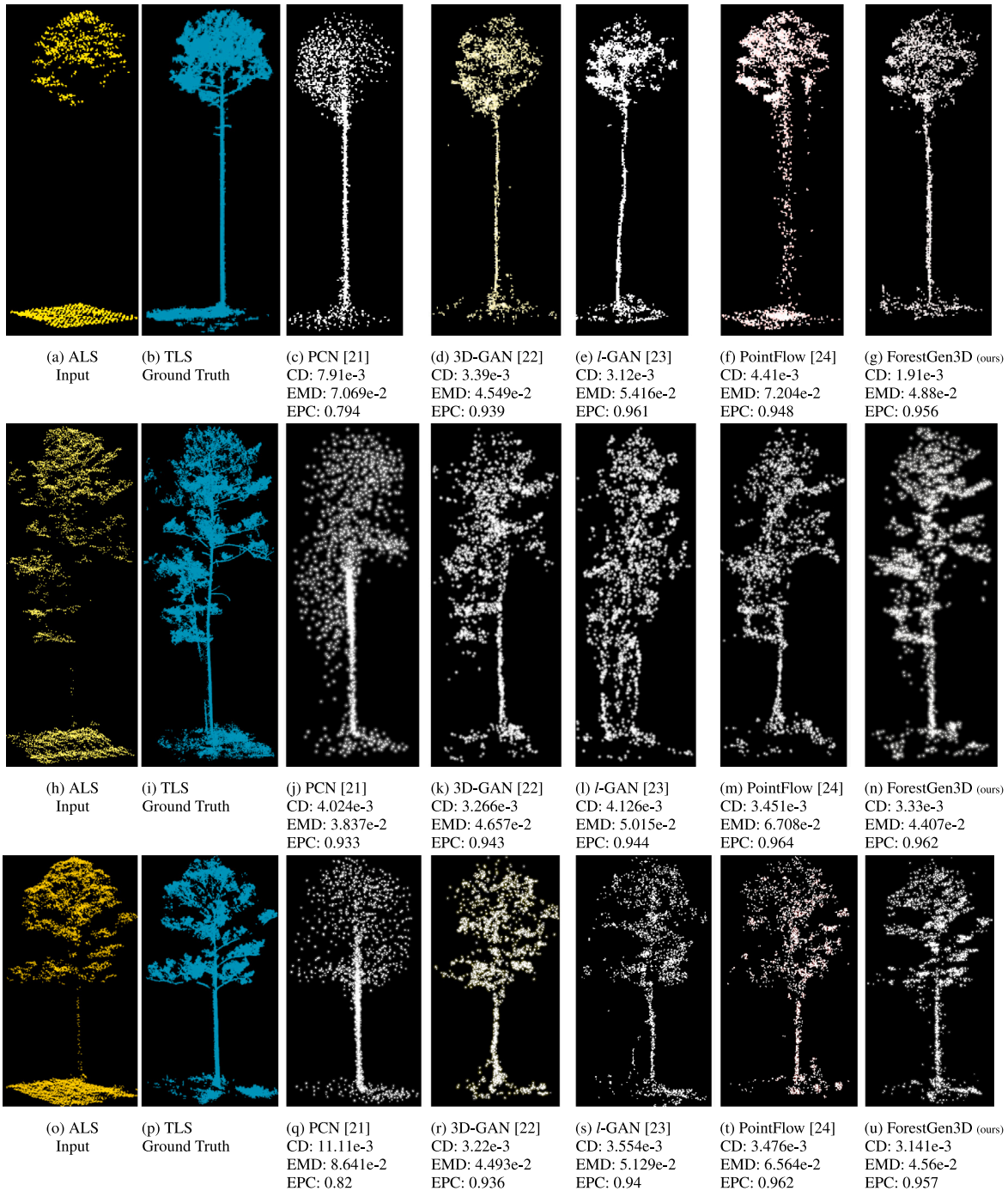


Fig. 9. Qualitative and Quantitative comparison of ALS-conditioned tree structure generation across three representative examples. (a,h,o) ALS input and (b,i,p) TLS reference; (c,j,q) PCN, (d,k,r) 3D-GAN, (e,l,s) latent-GAN, (f,m,t) PointFlow, and (g,n,u) ForestGen3D (ours). PCN oversmooths crowns, producing coarse silhouettes with little fine-scale detail. The other baselines generate shapes that more closely resemble TLS references but with inconsistencies such as distorted trunks (latent-GAN in (l)) or more discontinuous stems (PointFlow in (f)). 3D-GAN and ForestGen3D yield more coherent reconstructions, with ForestGen3D showing smaller metric errors with respect to ground truth. Pairwise point cloud distance metrics included are Chamfer Distance (CD), Earth's movers distance (EMD) and Expected Point containment (EPC).

et al., 2016), latent *l*-GAN (Achlioptas et al., 2018), PointFlow (Yang et al., 2019) and our ForestGen3D denoising diffusion based model. For reference, example CD/EMD/EPC values are also reported in Fig. 9 for their corresponding examples. Beyond the numerical values, the metrics highlight important differences between the approaches. PCN, while competitive in CD, performs worst in EMD, reflecting its bias toward capturing coarse, low frequency geometry at the expense of fine structural detail. Latent-GAN shows the largest CD, due to plausible over-compressing structural features in the latent layers, while 3D-GAN improves CD, it still remains above our model. PointFlow achieves stronger balance across CD and EMD, though it is also higher than the performance of our model. In contrast, ForestGen3D consistently achieves the lowest CD and EMD, while also attaining the highest EPC, indicating superior alignment with TLS reference structure and better recovery of fine-scale details. These results quantitatively demonstrate that ForestGen3D more effectively learns the underlying distribution of forest structure from ALS/TLS data, achieving overall improved performance compared to both completion/generation-based baselines.

Table 1
Quantitative generation and completion errors evaluated on the test set. Metrics include Chamfer Distance (CD), Earth Mover's Distance (EMD), and Expected Point Containment (EPC). CD and EMD quantify geometric similarity between generated point clouds and TLS references, while EPC measures spatial containment with the ALS convex envelope when TLS ground truth is unavailable. Lower CD and EMD values and higher EPC indicate better reconstruction quality and structural fidelity.

Method	Distance metrics		
	CD	EMD	EPC (%)
PCN (Yuan et al., 2018)	4.63e−3	9.58e−2	0.921
3D-GAN (Wu et al., 2016)	4.37e−3	5.761e−2	0.926
l-GAN (Achlioptas et al., 2018)	6.10e−3	5.59e−2	0.862
PointFlow (Yang et al., 2019)	4.52e−3	5.429e−2	0.939
ForestGen3D (Ours)	**4.30e−3**	**5.06e−2**	**0.945**

3.5. Plot-scale ALS-guided generation

We also conduct both qualitative and quantitative analyses at the plot scale, centered at TLS scans and covering a 25 meter radius. For the qualitative evaluation, we visually compare structural reconstructions obtained from ALS, TLS, ALS+TLS and a hybrid approach combining ALS with ForestGen3D generation enhancement. For each plot, individual trees were first detected using 3D bounding boxes, and our ForestGen3D model was then applied to generate plausible understory structure beneath each detected tree. Regions lacking tree or shrub like vegetation are left unaltered under the assumption that they do not suffer from occlusion in ALS. To support the quantitative assessment, we perform an evaluation in the downstream task of vegetation biometric estimates derived from each data source: ALS only, TLS only, ALS+TLS and ALS enhanced with ForestGen3D. The combined ALS+TLS data serves as ground truth, given its ability to capture detailed, high-resolution vegetation structure across the entire vertical extent of the tree, including the ground, overstory and understory components. Fig. 10 illustrates four representative testset examples (one per row) at plot-scale. The first column illustrates the ALS point cloud, the second column shows the corresponding co-registered ground truth TLS scan, and the third column presents the results of aggregating the ALS scan with the generated under-story from our ForestGen3D model. Note that in all examples, the ALS plots suffer from either missing or highly sparse understory information between the ground and tree foliage, as can be observed in the images in the first column where points are missing. TLS, on the other hand, provides high resolution in the elevation/altitude direction, as shown in the second column, where the main tree stem is finely resolved along the vertical tree-stand direction. The third column, showing the point cloud from ALS+ForestGen3D, illustrates that our approach is capable of enhancing or generating realistic understory structure at the plot scale, where the generated stems and ground features fill in gaps and transition smoothly from the available ALS data.

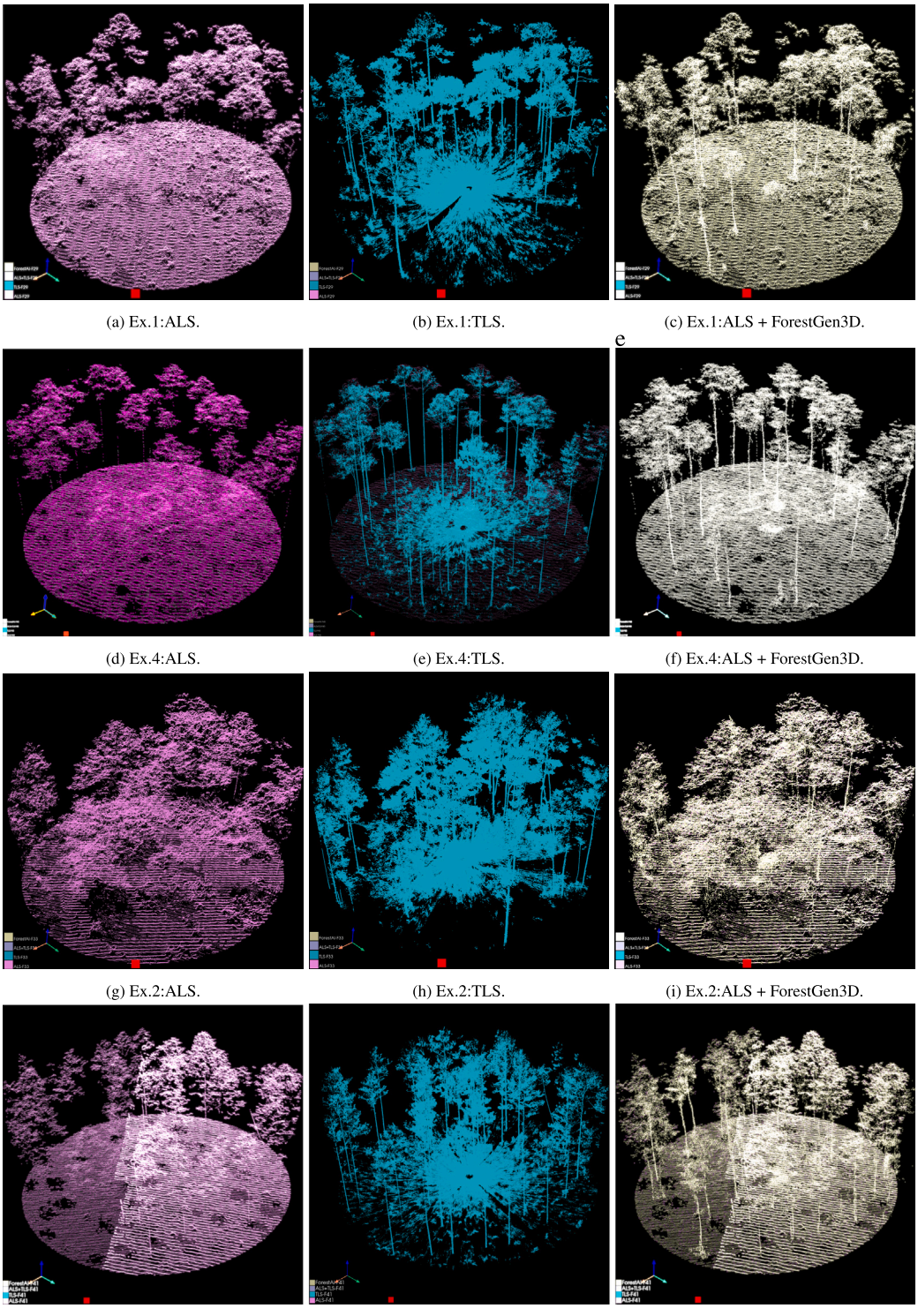
3.5.1. Do ForestGen3D-derived biometrics approximate TLS-derived distributions?

Biometric estimation from 3D point clouds is treated here as a downstream evaluation task rather than the primary objective of ForestGen3D. The goal of this analysis is to assess whether the additional sub-canopy structure generated by the model, while respecting ALS-provided geometry, produces biophysical metric distributions that are consistent with those derived from ALS+TLS. Accordingly, we first evaluate similarity at the level of biometric distributions, which offers a robust and sensor-agnostic assessment of structural realism, and then complement this analysis with per-tree RMSE-based experiments where reliable cross-source tree correspondences are available. This combined evaluation mitigates biases introduced by source-specific tree detection and segmentation differences while still enabling direct quantitative comparison when appropriate.

In the first set of experiments for this task, we compute biophysical tree metric distributions derived from ALS, TLS, ALS+TLS and ALS+ForestGen3D data sources. For this, we use a histogram per metric (one per column) depicting the distribution of key biophysical tree attributes, offering insight into the variability of tree structures within the sampled plot region. Such histograms were computed over co-registered sources at individual plots of 25 meter radius. To produce these histograms, we evaluate tree height, diameter at breast height (DBH), crown diameter (CrD), and crown volume (CrV) using established deep learning–based tree segmentation methods (Windrim and Bryson, 2019). These models segment each point cloud within detected vegetation bounding boxes into structural components: ground, low vegetation, stems, branches, and foliage, which are then used as inputs to geometric fitting procedures that yield the final biophysical metrics. Tree height is estimated using the highest elevation point within each detected tree, subtracting a projection of that point onto the ground surface. DBH is computed following the method of Rijal et al. (2012) applied to segmented stems, with robustness improved through a RANSAC based linear interpolation of fitted circles across multiple heights around breast height to mitigate the effect of outliers. Crown diameter is estimated by projecting segmented foliage points onto a horizontal plane and measuring the maximum pairwise distance. Crown volume is computed using a k-means–based convex hull approach (Zhu et al., 2021), which partitions foliage points into spatial clusters, computes the convex hull of each cluster, and sums their volumes. This approach mitigates overestimation caused by empty space inclusion when applying a single convex hull to the full crown.

In Fig. 11, integrated ALS+TLS data (third row across all examples) provides the most comprehensive representation of forest structure by combining ALS's canopy coverage from the top with TLS's detailed near-ground sampling. As such, biometrics derived from ALS+TLS are treated as a gold-standard reference for comparison. ALS-only data (first row) efficiently captures canopy-top structure and yields reliable estimates of tree height and crown diameter, but lacks the resolution necessary to characterize lower-canopy and understory features due to occlusion. This limitation leads to sparse and biased DBH estimates across most examples, while height remains closer to the ALS+TLS reference because it depend less directly on stem-level information. TLS-only data (second row) provides dense sampling of stems and lower-canopy elements, resulting in well-defined DBH distributions. TLS-derived crown diameter and height distributions are generally contained within ALS-derived ranges, consistent with the assumption that TLS crown points are spatially embedded within the ALS envelope. Smaller height values observed in TLS reflect understory trees occluded in ALS data or missed upper canopy. Differences in crown volume distributions reflect the complementary directional sampling strengths of ALS (horizontal canopy extent) and TLS (vertical crown structure). The ALS+ForestGen3D results (last row) indicate that the proposed model compensates for ALS structural blind spots while preserving ALS derived landscape geometry. The resulting biometric distributions consistently approximate those of the ALS+TLS reference across all evaluated metrics. Improvements are most pronounced for DBH and crown volume, indicating recovery of sub-canopy and stem structure absent in ALS-only data. Importantly, tree height and crown diameter distributions remain largely contained within the corresponding ALS-derived ranges, supporting the model's geometric containment guarantees under the ELBO objective. Together, these results indicate that ForestGen3D generates structurally plausible 3D vegetation consistent with TLS-derived distributions while maintaining fidelity to ALS-provided canopy geometry. Other representative example illustrations can be found in Figs. 15–17 in the Appendix.

Beyond qualitative assessment, we quantitatively evaluate model performance in this downstream task, by comparing the distributions of tree height, diameter at breast height, crown diameter, and crown volume derived from the different 3D sources. Because plot-scale biometric distributions are typically sparse (on the order of 10–30 trees



(a) Ex.1:ALS. (b) Ex.1:TLS. (c) Ex.1:ALS + ForestGen3D.

(d) Ex.4:ALS. (e) Ex.4:TLS. (f) Ex.4:ALS + ForestGen3D.

(g) Ex.2:ALS. (h) Ex.2:TLS. (i) Ex.2:ALS + ForestGen3D.

(j) Ex.3:ALS. (k) Ex.3:TLS. (l) Ex.3:ALS + ForestGen3D.

Fig. 10. Representative 3D structural comparisons across sensing sources. Each row corresponds to a distinct 25 m radius plot with overlapping ALS and TLS coverage. Columns show ALS (left), TLS (middle), and ALS+ForestGen3D (right). ForestGen3D reconstructs missing vertical stem and mid-story structure absent in ALS, smoothly transitioning from ALS-observed canopy and ground returns into the generated canopy interior and sub-canopy regions.

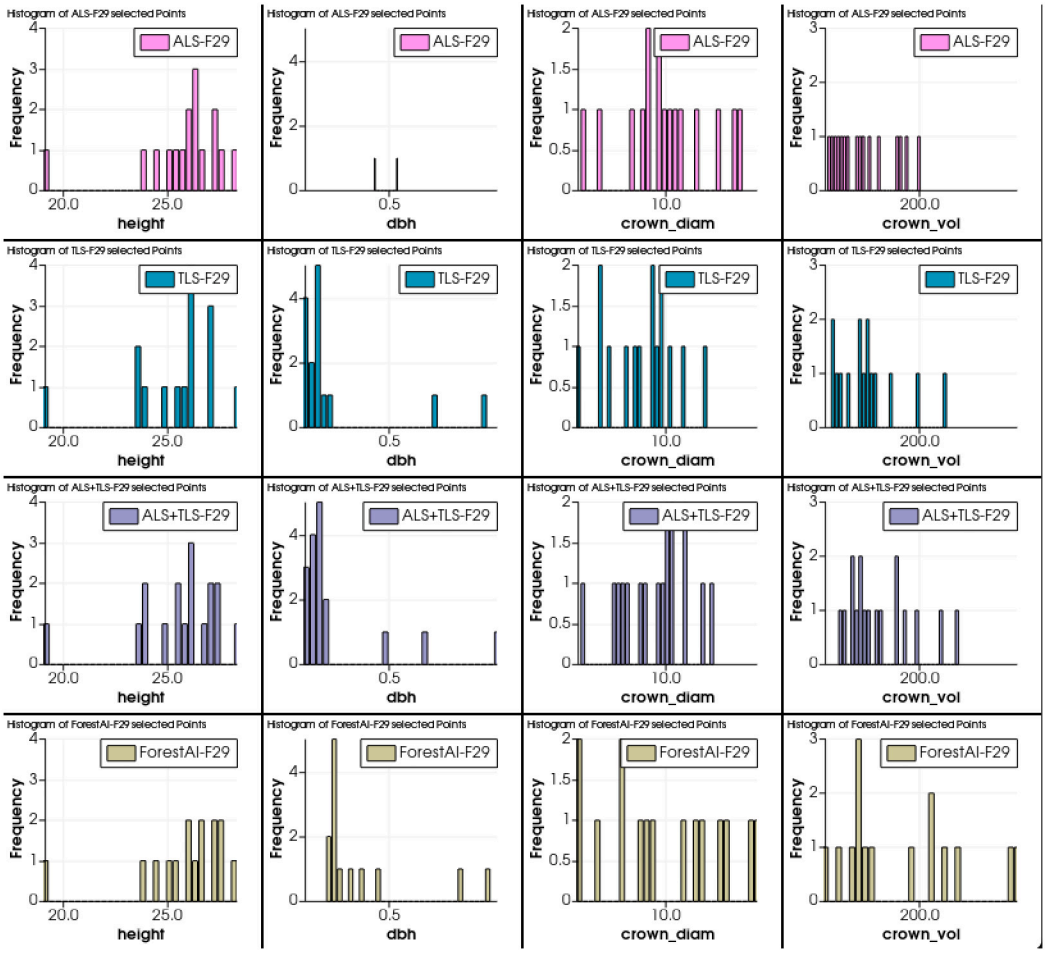


Fig. 11. Plot-scale distributions of biophysical tree metrics derived from different sensing sources. Rows correspond to ALS (row 1), TLS (row 2), ALS+TLS (row 3), and ALS+ForestGen3D (row 4). Columns show distributions for tree height (col. 1), diameter at breast height (DBH; col. 2), crown diameter (CrD; col. 3), and crown volume (CrV; col. 4). Each histogram represents metrics computed over a 25 m radius plot. ALS+TLS is treated as the structural reference, while ALS+ForestGen3D closely approximates its metric distributions, indicating recovery of sub-canopy structure while preserving ALS-derived spatial constraints.

per plot), we adopt the Wasserstein distance (WD) as our primary distributional comparison metric. WD is particularly well suited for this domain, as it remains stable under limited sample sizes and partial support overlap, unlike divergence-based measures such as the Kullback–Leibler divergence. The Wasserstein distance (see Panaretos and Zemel (2019) for an overview) provides an intuitive measure of dissimilarity by quantifying the minimum cost required to transform one distribution into another and is therefore also referred to as the Earth Mover's Distance. We explicitly distinguish this use of WD for comparing biometric distributions from the EMD metric which operates directly on 3D point sets.

Table 2 summarizes average Wasserstein distances computed across all 51 plots in the test set. Across all examples and metrics, ALS+ForestGen3D consistently exhibits the smallest Wasserstein distances relative to ALS+TLS. Note in general, that in terms of distances the smallests' are observed relative to the sensing source that provides greater information content for the structural features underlying that metric. For example, Wasserstein distances for tree height are generally smaller between ALS+TLS and ALS than between ALS+TLS and TLS, reflecting the fact that ALS offers more uniform and less occlusion affected sampling of canopy tops and ground surfaces. Conversely, metrics such as DBH, which depend strongly on stem level sampling, show closer agreement between ALS+TLS and TLS. Overall, the results here indicates that the biometric distributions derived from ALS+ForestGen3D generation most closely approximate those obtained from the combined ALS+TLS data, which provides the most complete characterization of forest structure. These findings support that ForestGen3D effectively recovers below-canopy structural information missing from ALS-only observations while remaining consistent with ALS-provided canopy geometry. Note that WD for crown volume is shown in cubic meters, hence its high value difference compared to the other metrics.

Table 2
Average Wasserstein distances between plot-scale biophysical metric distributions derived from ALS, TLS, and ALS+ForestGen3D relative to ALS+TLS, computed over all 51 plots in the test set. Results confirm that ALS-ForestGen3D yields biometric distributions that most closely approximate those derived from ALS+TLS across all evaluated metrics, demonstrating consistent performance across diverse plots within the mixed-conifer ecosystem.

Sensing source	Plot-scale biophysical metrics			
	Height	DBH	CrD	CrV
ALS	0.94	0.29	0.81	4.87^3
TLS	1.85	0.22	1.11	4.33^3
ALS + ForestGen3D	0.82	0.205	0.88	4.24^3

The second portion of the biometric estimation task directly assesses the added value of ForestGen3D relative to conventional ALS-based allometric approaches. This consists in estimating DBH from ALS-derived crown diameter and tree height using established allometric relationships following Jucker et al. (2017). Model parameters were

Table 3
Root mean square error (RMSE) of biophysical metrics derived from different sensing sources relative to the ALS+TLS reference. Metrics include tree height, diameter at breast height (DBH), crown diameter (CrD), and crown volume (CrV). DBH estimates for ALS-only inputs are obtained using an allometric model based on crown diameter and height following Jucker et al. (2017). Results are reported for a test of 265 matched trees across all sources within the Ft. Stewart 2024 dataset. Lower RMSE values indicate closer agreement with the ALS+TLS benchmark, with ALS+ForestGen3D achieving the lowest error across all metrics.

Sensing source	Plot-scale biophysical metrics			
	Height	DBH	CrD	CrV
ALS	0.531	0.208 (allometric (Jucker et al., 2017))	0.566	5.112^3
TLS	1.508	0.143	0.6127	4.029^3
ALS + ForestGen3D	0.480	0.112	0.561	2.109^3

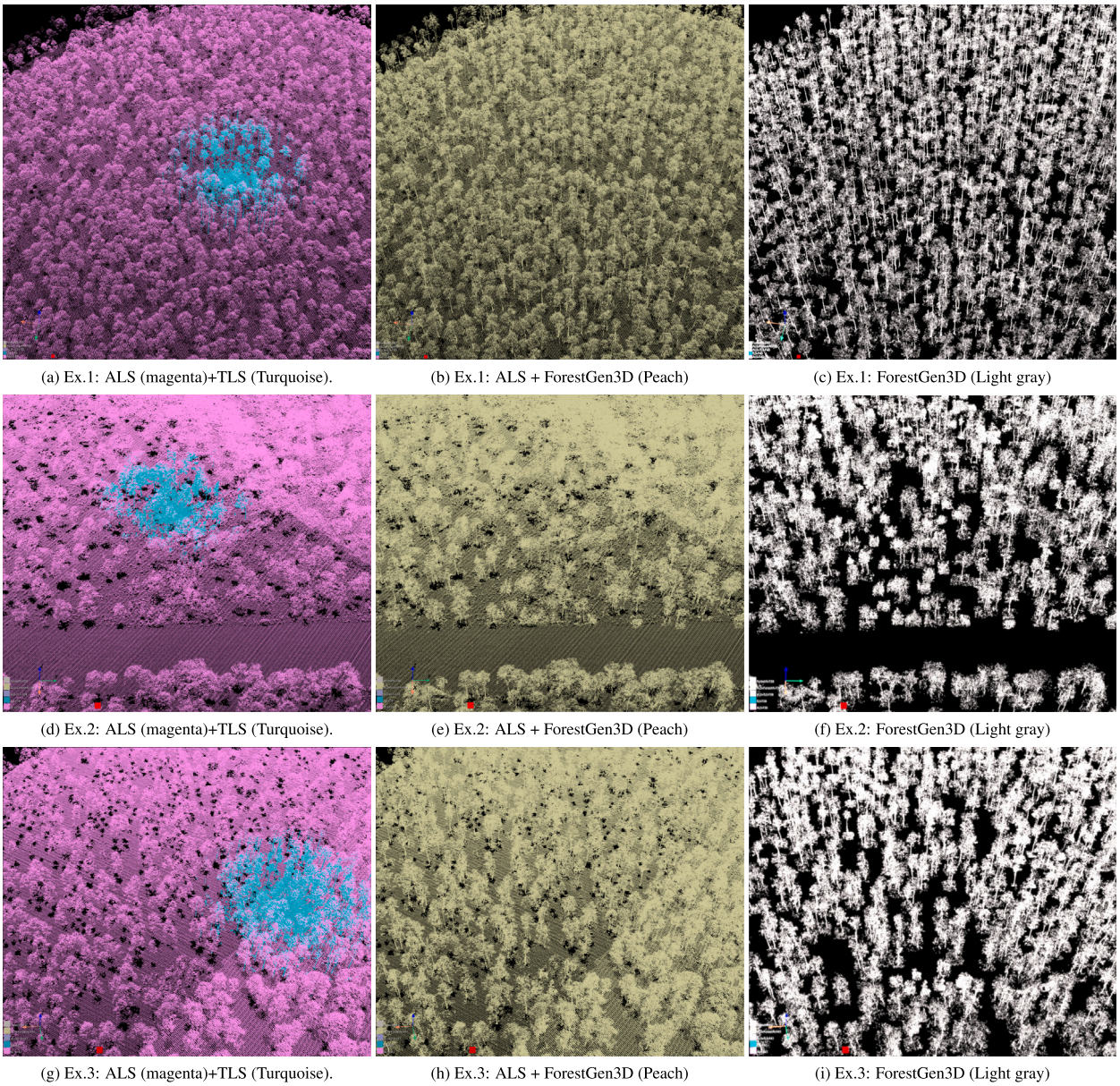
trained using 453 matched trees from the Ft. Stewart 2022 dataset, drawn from geographically distinct plots relative to the test set but within the same landscape and vegetation regime. Evaluation was performed on 246 test matched trees from the Ft. Stewart 2024 dataset, with ALS+TLS point clouds used as reference. To mitigate nonuniform sampling effects associated with fixed single scan TLS acquisitions, each point cloud was truncated to a 25 m radius around the TLS scan center. Trees were matched across co-registered ALS, TLS, ALS+TLS, and ALS+ForestGen3D sources using a 0.5 m spatial neighborhood. Table 3 summarizing RMSE's across all four evaluated biometrics: tree height, DBH, crown diameter, and crown volume show that ALS+ForestGen3D achieved the lowest RMSE. In particular, DBH error was reduced from 0.208 m (ALS allometry (Jucker et al., 2017)) to 0.112 m (ALS+ForestGen3D), while crown volume error was reduced by more than 50%. These results demonstrate that ForestGen3D not only reconstructs realistic 3D structure but also improves downstream biometric estimation relative to conventional ALS based allometric models.

3.6. Regional-scale ALS-guided generation

To evaluate the scalability of the ForestGen3D model, we applied it to generate 3D forest structure over a larger spatial extent, a 200-meter radius region centered at TLS scan geo-referenced locations. While the TLS scan itself covers a range of only 25 m, this broader area enables us to test the model's ability to extrapolate beyond small-scale. Notably, the ALS data used in this region was not part of the training set. The generation procedure at this larger scale mirrors the approach used for plot-level synthesis. First, we apply a tree detection algorithm across the entire ALS region of interest to identify areas with occlusion when viewed from above, which are typically regions where airborne LiDAR fails to resolve understory structure. Second, we extract 3D bounding boxes around these detected trees and apply the ForestGen3D model independently within each bounding box using the local ALS point cloud as input. Finally, the resulting generated point clouds are aggregated over the full 200-meter radius area keeping track of their localization to produce a coherent forest 3D reconstruction. Fig. 12 illustrate representative snapshot examples of the resulting generation/reconstruction. In these visualizations, point clouds are color-coded as follows: magenta for ALS, turquoise for TLS, peach for ALS + ForestGen3D generation, and white for the generated ForestGen3D samples alone. In this test case where TLS data is not available, the model generates understory structure consistent with the characteristic statistical distribution observed in the training data. Importantly, these large-scale results corroborate the findings from the plot-scale experiments. ForestGen3D generates foliage and branching structures that blend smoothly into those observed in the ALS data, maintaining spatial consistency and fidelity. In understory regions lacking ALS points, the model synthesizes realistic and plausible tree trunks in locations and with diameters that appear structurally consistent with the associated crown geometry provided by ALS. When sparse ALS points are present in the understory, the generated output enhances and complements this limited data by reinforcing plausible stem/branch geometry. At the ground level, the model enriches the representation of shrubs and low vegetation while also improving the resolution of ground returns and the near-ground structure.

To explicitly evaluate whether the generated 3D structure remains spatially constrained by the ALS observations, particularly in the absence of TLS ground truth at landscape scales, we next introduce a complementary evaluation based on the expected point containment (EPC) in Eq. (1). This analysis quantifies the extent to which generated points respect the implicit spatial bounds defined by the ALS convex hull, providing a geometry aware validation of large scale structural plausibility. For this, we analyzed the EPC metric and the distribution of distances between generated points falling outside the ALS convex envelope and their nearest point on the envelope surface. Fig. 13 summarizes this analysis across 10 evaluation plots, each representing a 200 meter radius ALS-only region labeled by plot name on the x-axis. Each violin plot shows the distribution of out-of-envelope distances, computed only for generated points that lie outside the ALS convex hull. The line inside each violin represents the median, while the shape reflects the underlying distance histogram. Text annotations above each violin indicate: (i) N, the number of trees generated per 200 m radius region; (ii) *Out*, the ratio of generated points outside the ALS convex hull (relative to the total number of points); and (iii) the mean $\pm$ standard deviation of the out-of-envelope distances. Across all plots, the fraction of points outside the ALS envelope remains consistently low, ranging from 1.4% to 1.6%, and mean distances remain below 0.2 m, demonstrating good quality generations via the EPC. The results in Fig. 13 provide an important consistency check linking landscape scale ALS geometry with the fine scale structure generated by ForestGen3D. Because ALS often undersamples subcanopy regions, a generative model could in principle hallucinate structure that violates spatial constraints or extends far beyond the areas supported by ALS evidence. The violin plot analysis quantitatively shows that this does not occur in practice: the model generates new subcanopy detail while remaining tightly confined to the ALS spatial envelope. This demonstrates that ForestGen3D enriches incomplete ALS point clouds without distorting broader landscape geometry and, consistent with earlier findings linking high EPC to low Chamfer Distance (CD) and Earth Mover's Distance (EMD) under TLS supervision, indicates that the model achieves low 3D reconstruction error and high-quality generation at the landscape scale.

Summarizing, across all evaluations, ForestGen3D demonstrates strong and consistent quantitative performance in reconstructing subcanopy forest structure from ALS input alone. At the individual tree scale, the model achieves low reconstruction error relative to TLS references, as shown by Chamfer distance, Earth's Moving Distance, and expected point containment (Fig. 7, Fig. 8, Table 1). Biophysical attributes derived from generated point clouds tree height, DBH, crown diameter, and crown volume closely track TLS measurements and outperform or complement ALS-based or ALS-allometry based estimates (Tables 2, 3), indicating that the model recovers realistic structural detail not observable from ALS alone. At landscape scale, ForestGen3D preserves global spatial consistency with ALS geometry, with the EPC metric and out-of-envelope distance distributions (Fig. 13) confirming that more than 98% of generated points remain contained within the ALS convex hull and deviations remain minimal (<0.2 m). Together, these quantitative analyses demonstrate that ForestGen3D provides a reliable and ecologically coherent enhancement of ALS-only forest structure, producing high-fidelity 3D reconstructions across individual-tree, plot-scale and landscape-level settings.



(a) Ex.1: ALS (magenta)+TLS (Turquoise). (b) Ex.1: ALS + ForestGen3D (Peach) (c) Ex.1: ForestGen3D (Light gray)

(d) Ex.2: ALS (magenta)+TLS (Turquoise). (e) Ex.2: ALS + ForestGen3D (Peach) (f) Ex.2: ForestGen3D (Light gray)

(g) Ex.3: ALS (magenta)+TLS (Turquoise). (h) Ex.3: ALS + ForestGen3D (Peach) (i) Ex.3: ForestGen3D (Light gray)

Fig. 12. Landscape-scale generation examples from three different areas (one per row). Columns show (a,d,g) broad ALS (magenta) with *in-situ* TLS reference (turquoise), (b,e,h) ALS combined with ForestGen3D output (peach), and (c,f,i) ForestGen3D alone (light gray). Each snapshot corresponds to a circular area of 200 m radius. ForestGen3D was applied selectively to regions with occlusion in ALS, filling in subcanopy structure while leaving unoccluded regions unchanged (black areas indicate no generation). Note that these examples are outside the training set.

3.7. Discussion

Our findings confirm that DDPM models can generate realistic 3D representations of forest tree structures. The model's ability to preserve fidelity to the true 3D structure depends on the informativeness of additional conditioning data sources. Furthermore, the computational complexity of the generative process proposed here scales linearly with area coverage, enabling efficient deployment across large landscapes without increasing model complexity. With current implementations taking $\approx$ 0.4 s per tree-scale generation in a Tesla V100S-PCIE-32GB GPU and with possibilities to further parallelize generation. However, the approach has several limitations. First, the generative model is limited by the distribution it has learned from the training data, which in this case primarily consists of mixed coniferous forest types. When applied to generate trees with structural characteristics not well represented in the training set such as species with drastically different crown architectures, stem densities, branching patterns or ground vegetation (e.g., small shrubs, debri), the model may produce structures that deviate from true vegetation distributions. As a result, generalization to out-of-distribution ecosystems, such as boreal forests, tropical canopies, or hardwood-dominated stands, remains open. Addressing this limitation will require either broadening the training dataset to cover more ecological diversity or incorporating domain adaptation techniques to enable transfer across forest types with different structural priors. While the ALS convex hull provides a useful geometric prior for bounding plausible TLS generation, it is important to note that this containment assumption is approximate. The convex envelope may under/overestimate volume in sparsely sampled regions, and containment does not guarantee structural realism on its own. However, we empirically observe that TLS scans lie within the ALS envelope with over 96% consistency in the training set, and that generation fidelity to the true structure and containment are strongly correlated. These findings support its use as a practical diagnostic metric in large-scale,

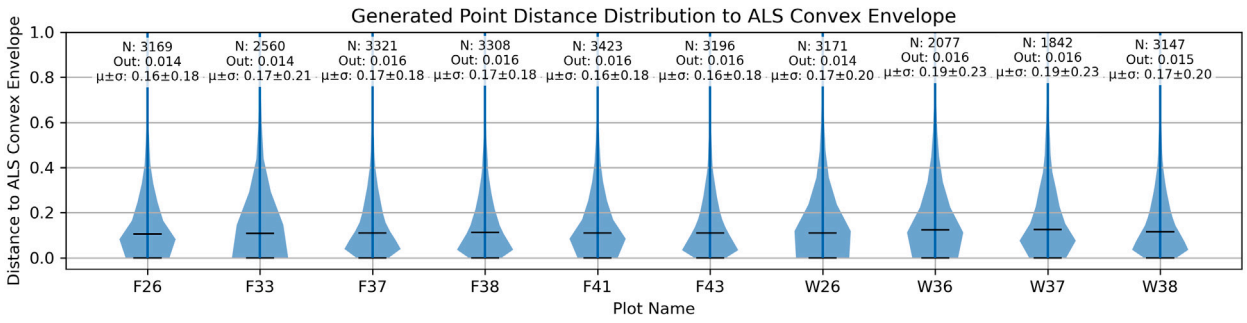


Fig. 13. Distribution of point-to-envelope distances between ForestGen3D generated points and the ALS derived convex hull across ten landscape-scale (200 m radius) regions. The x-axis represents different geo-located areas, the y-axis indicates the minimum Euclidean distance from generated points lying outside the ALS convex hull to their nearest point on the hull surface. The violin shape represents the distribution of these distances; the internal black bar denotes the median. Annotations above each plot report: N, the number of trees generated in the region; *Out*, the fraction of generated points falling outside the ALS envelope; and the mean $\pm$ standard deviation ($\mu \pm \sigma$) of out-of-envelope distances. Across all regions, the proportion of outside points remains below 1.6%, and deviations remain small (typically <0.2 m). These results show that ForestGen3D generates structure that remains spatially consistent with the ALS envelope while enriching sub-canopy detail, confirming that the model respects the geometry encoded in ALS measurements.

ground-truth limited deployments. In addition, while single-scan TLS provides a practical and widely used means of capturing tree-level structure, we recognize that multi-scan, multi-view TLS acquisitions would offer a more complete 3D characterization of vegetation by substantially reducing occlusion, especially in lower and mid-story foliage. Access to such high-quality, fully co-registered TLS datasets would benefit generative modeling in two ways. First, it would provide a less biased training distribution, allowing the DDPM to learn understory geometry from point clouds that better approximate the true forest structure. Second, improved completeness in the TLS training dataset would likely enhance the model's ability to reproduce fine-scale geometric features and further improve downstream LiDAR-based biometric estimate tasks. Exploring multi-scan TLS datasets therefore represents a promising direction for future work, particularly for advancing the model's capacity to learn and generate detailed understory and ground-fuel structure.

In dense forest stands, tree detection is often challenged by dense canopies that occlude crown boundaries in ALS data. As a result, the detection pipeline may assign a single bounding box to multiple adjacent trees, which in turn affects the input to the generative model. Depending on the structural cues available in the ALS input, such as crown asymmetry, partial separation, or variations in canopy height the generative model may synthesize either a single fused structure or multiple distinct stems. Because the training dataset includes several examples of closely packed multi-tree configurations, the capacity to distinguish between adjacent trees when sufficient information is present in the point cloud. Nevertheless, in cases where ALS input lacks clear separation cues, the resulting generations may overestimate stem diameter or merge tree forms. To improve robustness in these scenarios, future work will explore both during training and deployment, forest context-aware adaptive tree detection methods, including the use of segmentation masks for individual crown delineation as in Yang et al. (2020), as well as stem count aware conditioning strategies to better guide the generative process when multiple trees are present within a single detection. Finally, the model does not explicitly account for temporal variations in forest structure caused by seasonal changes, prescribed burns, wildfires, thinning, or disease. Ignoring these dynamic factors can create a mismatch between generated and real world forest conditions, limiting the model's applicability for long-term ecological monitoring and fuel characterization. The aim in future studies is to develop models with additional training/conditional data informative of the potential temporal dynamics either occurring naturally or from human intervention.

Despite these limitations, the ability to reconstruct fine-scale 3D vegetation structure from ALS at tree, plot, and regional scales offers immediate value for next-generation wildfire and ecosystem modeling frameworks. ForestGen3D generates spatially explicit 3D point clouds that preserve the heterogeneous landscape geometry encoded in ALS while reconstructing missing sub-canopy and stem level structure. This capability is particularly important for advancing next-generation fire behavior and fuel modeling systems based on FIRETEC (Linn et al., 2002), QUIC-Fire (Linn et al., 2020), and FASTFuels (Marcozzi et al., 2025), which are beginning to use 3D representations of canopy gaps, ladder fuels, and vertical fuel continuity. By producing high-resolution 3D forest structure that remains spatially consistent with ALS observations, ForestGen3D enables high-resolution 3D models to be applied in ALS-only environments while supplying the sub-canopy detail required for physically realistic simulation. While ALS-based allometric models have proven effective for estimating scalar forest attributes such as DBH, biomass, and crown dimensions from canopy-level measurements, they fundamentally operate on summary metrics and do not reconstruct explicit three-dimensional geometry. In this work, biometric estimation is therefore treated as a downstream evaluation task rather than the primary objective. Our results show that ForestGen3D not only reconstructs realistic 3D structure but also yields biophysical metric distributions that closely match those derived from TLS. Quantitative evaluations using distributional similarity metrics and RMSE demonstrate that ALS+ForestGen3D improves DBH estimation relative to conventional ALS-based allometry, indicating that the model captures ecologically meaningful sub-canopy and stem-level variability.

4. Conclusion

We present ForestGen3D, a scalable generative modeling framework that produces high-resolution 3D forest structure from aerial LiDAR alone. By leveraging denoising diffusion probabilistic models trained on co-registered aerial and terrestrial LiDAR data, ForestGen3D reconstructs sub-canopy detail that is typically occluded in ALS, enabling realistic and ecologically grounded 3D generations across landscapes. Our work introduces a geometric containment prior, referred to as the expected point consistency, preserved over model generations ensuring spatial plausibility, and we show that this prior can serve as a useful proxy for reconstruction quality in the absence of ground truth TLS. Empirical results across multiple spatial scales confirm that ForestGen3D approximates TLS structure with high fidelity in both geometric and ecological terms. Biophysical metric distributions derived from generated data such as tree height, DBH, crown diameter, and volume closely match those from full ALS+TLS data, demonstrating the model's utility in ecological monitoring and risk analysis. Despite these strengths, the model's generalizability is constrained by the structural diversity present in the training dataset. Ecosystems with markedly different vegetation forms or disturbance regimes may require adaptation via domain transfer or broader training sets. In addition, the

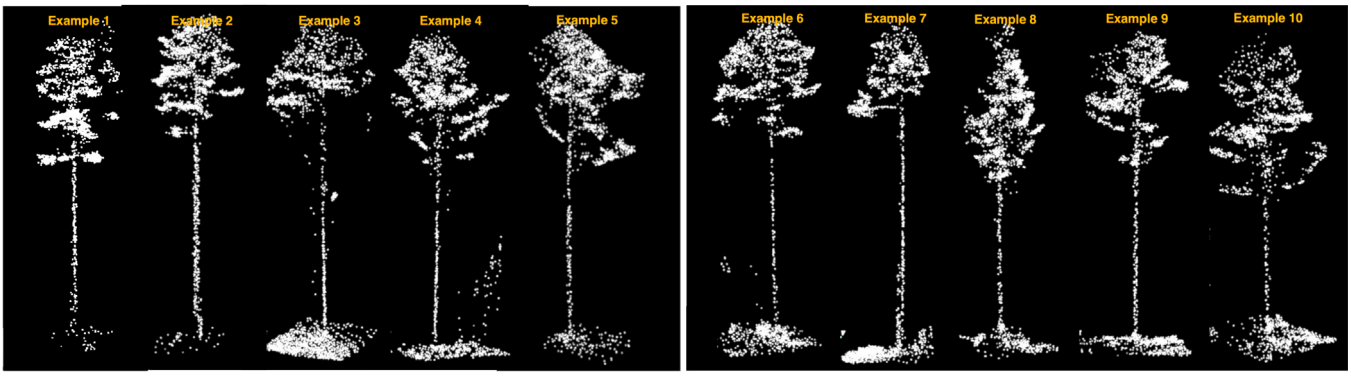


Fig. 14. ForestGen3D model generation under the unconditional setting: Sample from Gaussian distribution and generate tree structures according to a training TLS scan dataset.

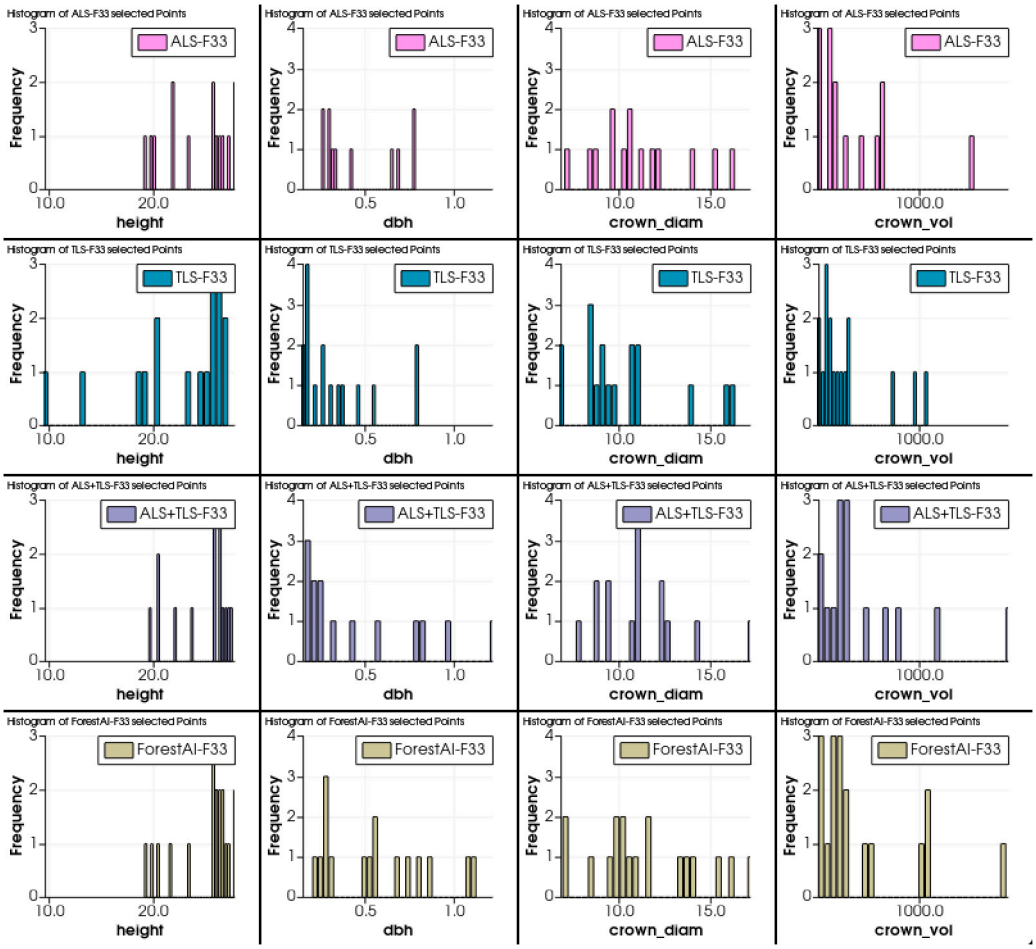


Fig. 15. Plot-scale distributions of biophysical tree metrics derived from different sensing sources. Rows correspond to ALS (row 1), TLS (row 2), ALS+TLS (row 3), and ALS+ForestGen3D (row 4). Columns show distributions for tree height (col. 1), diameter at breast height (DBH; col. 2), crown diameter (CrD; col. 3), and crown volume (CrV; col. 4). Each histogram represents metrics computed over a 25 m radius plot. ALS+TLS is treated as the structural reference, while ALS+ForestGen3D closely approximates its metric distributions, indicating successful recovery of sub-canopy structure while preserving ALS-derived spatial constraints.

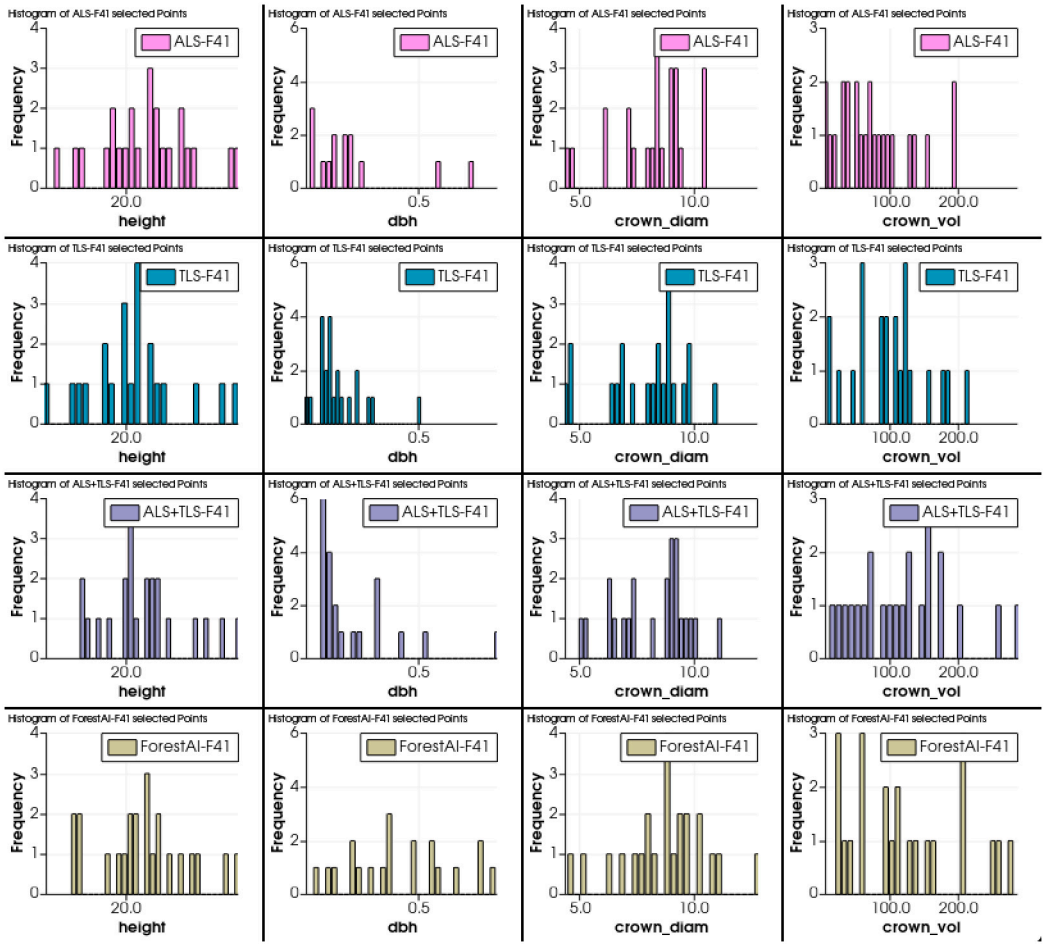


Fig. 16. Plot-scale distributions of biophysical tree metrics derived from different sensing sources. Rows correspond to ALS (row 1), TLS (row 2), ALS+TLS (row 3), and ALS+ForestGen3D (row 4). Columns show distributions for tree height (col. 1), diameter at breast height (DBH; col. 2), crown diameter (CrD; col. 3), and crown volume (CrV; col. 4). Each histogram represents metrics computed over a 25 m radius plot. ALS+TLS is treated as the structural reference, while ALS+ForestGen3D closely approximates its metric distributions, indicating successful recovery of sub-canopy structure while preserving ALS-derived spatial constraints.

model currently does not explicitly account for temporal dynamics such as seasonal change, management treatments, or natural disturbance recovery. Future work will extend the framework to incorporate more ecosystem diversity, temporal dynamics, and time-aware conditioning to further improve fidelity and robustness. Overall, ForestGen3D provides a practical and extensible solution to the challenge of generating detailed 3D forest structure in ALS only environments, with promising applications in ecological forecasting, remote sensing, and wildfire modeling.

CRediT authorship contribution statement

Juan Castorena: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **E. Louise Loudermilk:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Data curation. **Scott Pokswinski:** Resources, Data curation. **Rodman Linn:** Validation, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Research presented in this article was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20220024DR and 20260010DR. This research was also funded by the Department of Defense, Environmental Security Technology Certification Program, grant number RC24-8162. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

Appendix A. Expected point consistency guarantee via ELBO minimization

Assumptions. Let $(\mathbf{X}_0, \mathbf{Y})$ denote a paired TLS/ALS point cloud sample drawn from the empirical training distribution D. Let $\mathcal{B}_{\text{ALS}}(\mathbf{Y}) = \text{Conv}(\mathbf{Y})$ denote the convex hull of the ALS point cloud. Define the point-level containment indicator $C(\mathbf{x}, \mathbf{Y}) = 1(\mathbf{x} \in \mathcal{B}_{\text{ALS}}(\mathbf{Y}))$.

We assume:

1. **Empirical containment:** TLS points are spatially bounded by the ALS convex hull in expectation:

$$\mathbb{E}_{(\mathbf{X}_0,\mathbf{Y})\sim D}\left[\mathbb{E}_{\mathbf{x}\sim\mathbf{X}_0}\left[C(\mathbf{x},\mathbf{Y})\right]\right] \geq 1-\epsilon$$

for a small constant $\epsilon \in [0, 1]$.

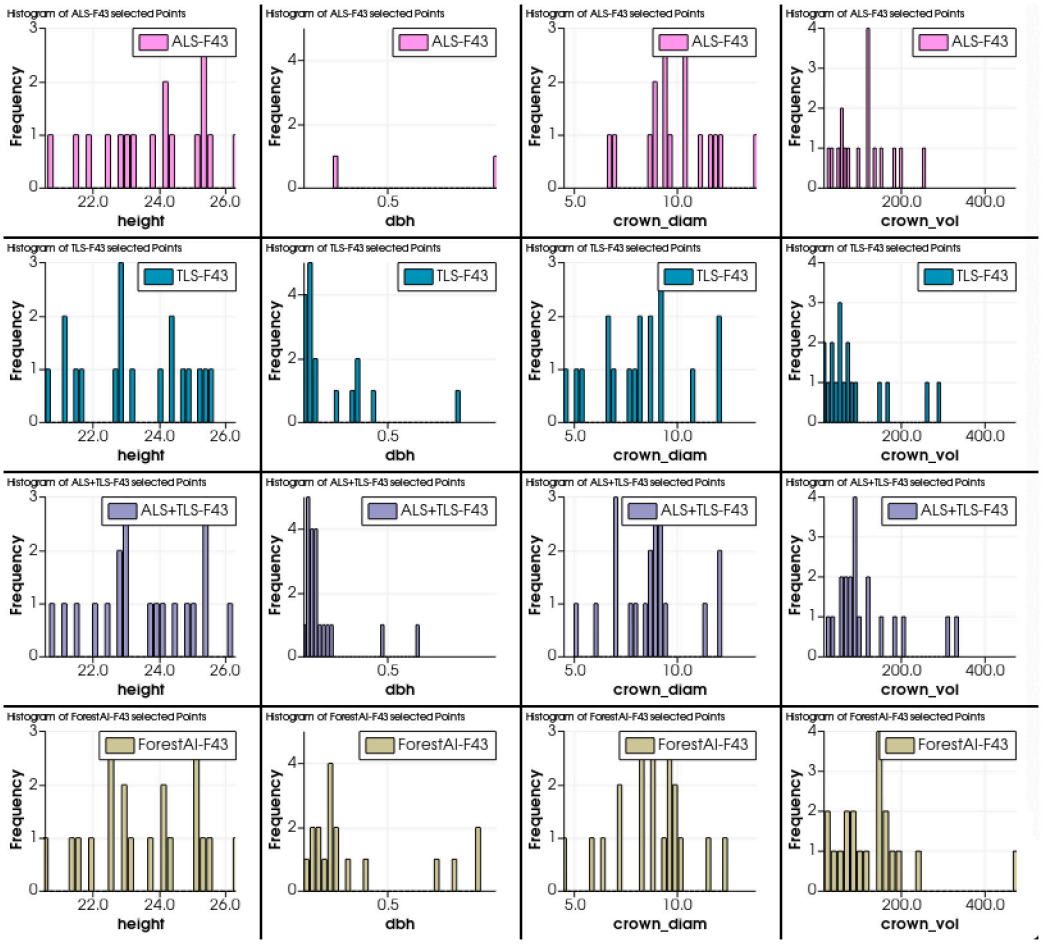


Fig. 17. Plot-scale distributions of biophysical tree metrics derived from different sensing sources. Rows correspond to ALS (row 1), TLS (row 2), ALS+TLS (row 3), and ALS+ForestGen3D (row 4). Columns show distributions for tree height (col. 1), diameter at breast height (DBH; col. 2), crown diameter (CrD; col. 3), and crown volume (CrV; col. 4). Each histogram represents metrics computed over a 25 m radius plot. ALS+TLS is treated as the structural reference, while ALS+ForestGen3D closely approximates its metric distributions, indicating successful recovery of sub-canopy structure while preserving ALS-derived spatial constraints.

2. **Model approximation:** The learned model $p_\theta(\mathbf{X}_0 \mid \mathbf{Y})$ satisfies a KL divergence bound:

$$\mathrm{KL}(p_{\text{data}}(\mathbf{X}_0 \mid \mathbf{Y}) \parallel p_\theta(\mathbf{X}_0 \mid \mathbf{Y})) \le \delta_{\mathrm{KL}}$$

for all $\mathbf{Y}$.

Scope Note. The containment assumption stated above is specific to the tree scale, where ($\mathbf{X}_0$ and $\mathbf{Y}$) are co-registered ALS and TLS point clouds corresponding to individual trees or small tree clusters. We do not assume global containment over entire plots or landscape-scale point clouds, where occlusion and ALS sparsity can invalidate this condition. In our dataset, this per-tree containment assumption is empirically validated, with over 96% of TLS points contained within the ALS convex hull on average.

Goal. Show that the expected containment of generated points satisfies:

$$\mathbb{E}_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{X}_0\sim p_\theta(\cdot|\mathbf{Y})}\left[\mathbb{E}_{\mathbf{x}\sim\mathbf{X}_0}\left[C(\mathbf{x},\mathbf{Y})\right]\right]\right] \ge 1-\epsilon-\delta$$

for some $\delta = \sqrt{2\delta_{\mathrm{KL}}}$, using Pinsker's inequality.

Proof. Let

$$f(\mathbf{X}_0,\mathbf{Y}) = \mathbb{E}_{x\sim\mathbf{X}_0}\left[C(\mathbf{x},\mathbf{Y})\right]$$

be the average containment score for a TLS sample $\mathbf{X}_0$ with respect to the ALS envelope. By assumption (1), we have:

$$\mathbb{E}_{(\mathbf{X}_0,\mathbf{Y})\sim D}\left[f(\mathbf{X}_0,\mathbf{Y})\right] \ge 1-\epsilon.$$

Now fix a conditioning ALS point cloud $\mathbf{Y}$. Then:

$$\left|\mathbb{E}_{\mathbf{X}_0\sim p_{\text{data}}(\cdot|\mathbf{Y})} f(\mathbf{X}_0,\mathbf{Y}) - \mathbb{E}_{\mathbf{X}_0\sim p_\theta(\cdot|\mathbf{Y})} f(\mathbf{X}_0,\mathbf{Y})\right| \le \|p_{\text{data}}(\cdot \mid \mathbf{Y}) - p_\theta(\cdot \mid \mathbf{Y})\|_{\mathrm{TV}}.$$

Here, $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation norm, a standard metric between probability distributions. It bounds the discrepancy in expectations of bounded functions and is upper-bounded by KL divergence via Pinsker's inequality (Cover and Thomas, 2006). Applying such inequality:

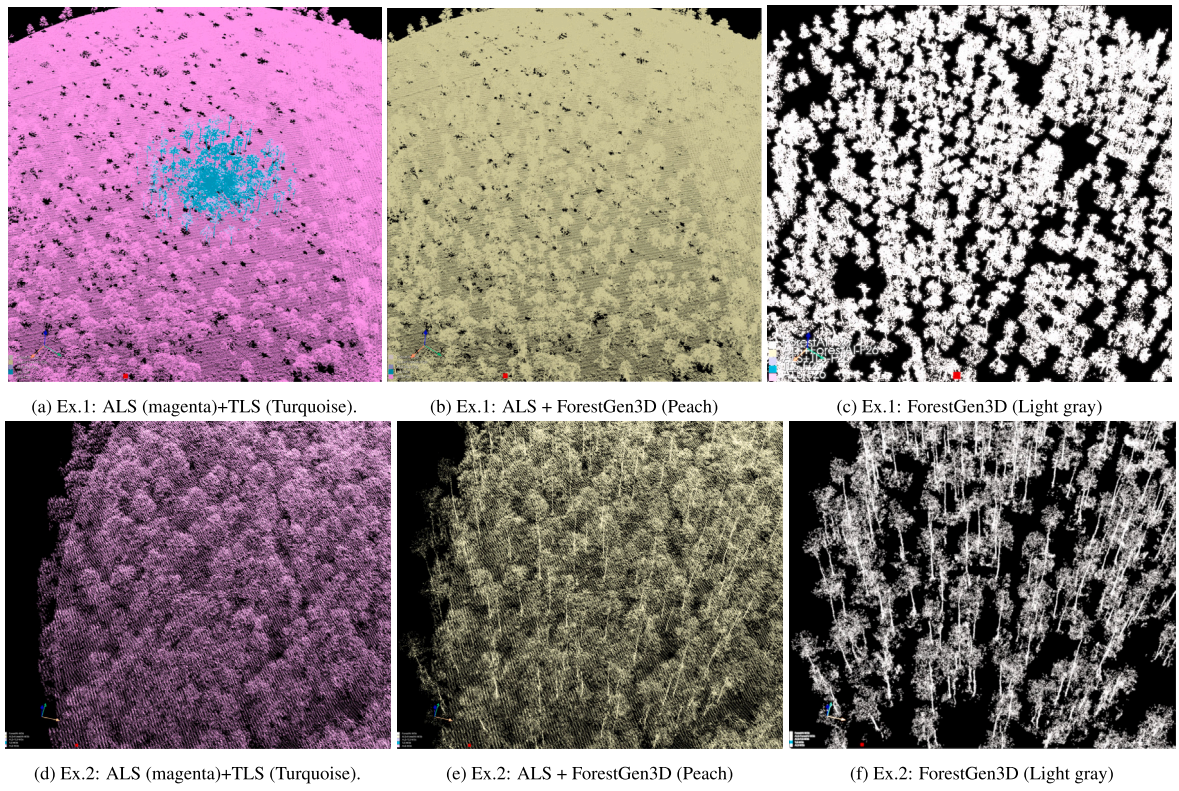
$$\|p_{\text{data}}(\cdot \mid \mathbf{Y}) - p_\theta(\cdot \mid \mathbf{Y})\|_{\mathrm{TV}} \le \sqrt{\frac{1}{2}\mathrm{KL}(p_{\text{data}} \parallel p_\theta)} \le \sqrt{\frac{1}{2}\delta_{\mathrm{KL}}}.$$

Since f is bounded in $[0,1]$, we obtain:

$$\mathbb{E}_{\mathbf{X}_0\sim p_\theta(\cdot|\mathbf{Y})} f(\mathbf{X}_0,\mathbf{Y}) \ge \mathbb{E}_{\mathbf{X}_0\sim p_{\text{data}}(\cdot|\mathbf{Y})} f(\mathbf{X}_0,\mathbf{Y}) - \sqrt{\frac{1}{2}\delta_{\mathrm{KL}}}.$$

Taking the expectation over $\mathbf{Y}$ and combining with the empirical containment assumption:

$$\mathbb{E}_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{X}_0\sim p_\theta(\cdot|\mathbf{Y})} f(\mathbf{X}_0,\mathbf{Y})\right] \ge \mathbb{E}_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{X}_0\sim p_{\text{data}}(\cdot|\mathbf{Y})} f(\mathbf{X}_0,\mathbf{Y})\right] - \sqrt{2\delta_{\mathrm{KL}}} \ge 1-\epsilon-\sqrt{2\delta_{\mathrm{KL}}}.$$



(a) Ex.1: ALS (magenta)+TLS (Turquoise). (b) Ex.1: ALS + ForestGen3D (Peach) (c) Ex.1: ForestGen3D (Light gray)

(d) Ex.2: ALS (magenta)+TLS (Turquoise). (e) Ex.2: ALS + ForestGen3D (Peach) (f) Ex.2: ForestGen3D (Light gray)

Fig. 18. Additional examples of landscape-scale ForestGen3D generation. Each row corresponds to a distinct 200 m radius area centered on a TLS scan location. Columns show (left) ALS combined with TLS reference data, (middle) ALS combined with ForestGen3D output, and (right) ForestGen3D-generated points alone. Point clouds are color-coded as follows: magenta for ALS, turquoise for TLS, peach for ALS + ForestGen3D, and white for ForestGen3D-generated points. ForestGen3D is applied selectively to regions exhibiting occlusion in the ALS data, generating sub-canopy structure while preserving ALS-observed geometry.

Therefore:

$$\mathbb{E}_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{X}_0\sim p_\theta(\cdot|\mathbf{Y})}\left[\mathbb{E}_{\mathbf{x}\sim\mathbf{X}_0}\left[1(x\in B_{\mathrm{ALS}}(\mathbf{Y}))\right]\right]\right] \geq 1-\epsilon-\sqrt{2\delta_{\mathrm{KL}}}. \quad \square$$

Appendix B. Unconditional 3D tree generation with DDPM

In some regions, TLS or co-registered ALS–TLS data may be unavailable due to logistical constraints, cost, or historical limitations. In such cases, it becomes necessary to resort to statistical approximations such as species specific structural priors or to borrow training data from ecologically similar surrogate areas. One advantage of the proposed DDPM framework is that it supports unconditional generation, allowing plausible 3D vegetation structure to be synthesized even without ALS guidance. This is particularly useful in applications where only species composition, vegetation type, or biophysical statistical information is known to populate a geo-referenced region. We include here several illustrative examples at the tree-scale that demonstrate that the model, even in the absence of ALS conditioning, can produce tree structures that are structurally realistic and ecologically consistent with those found in the training distribution.

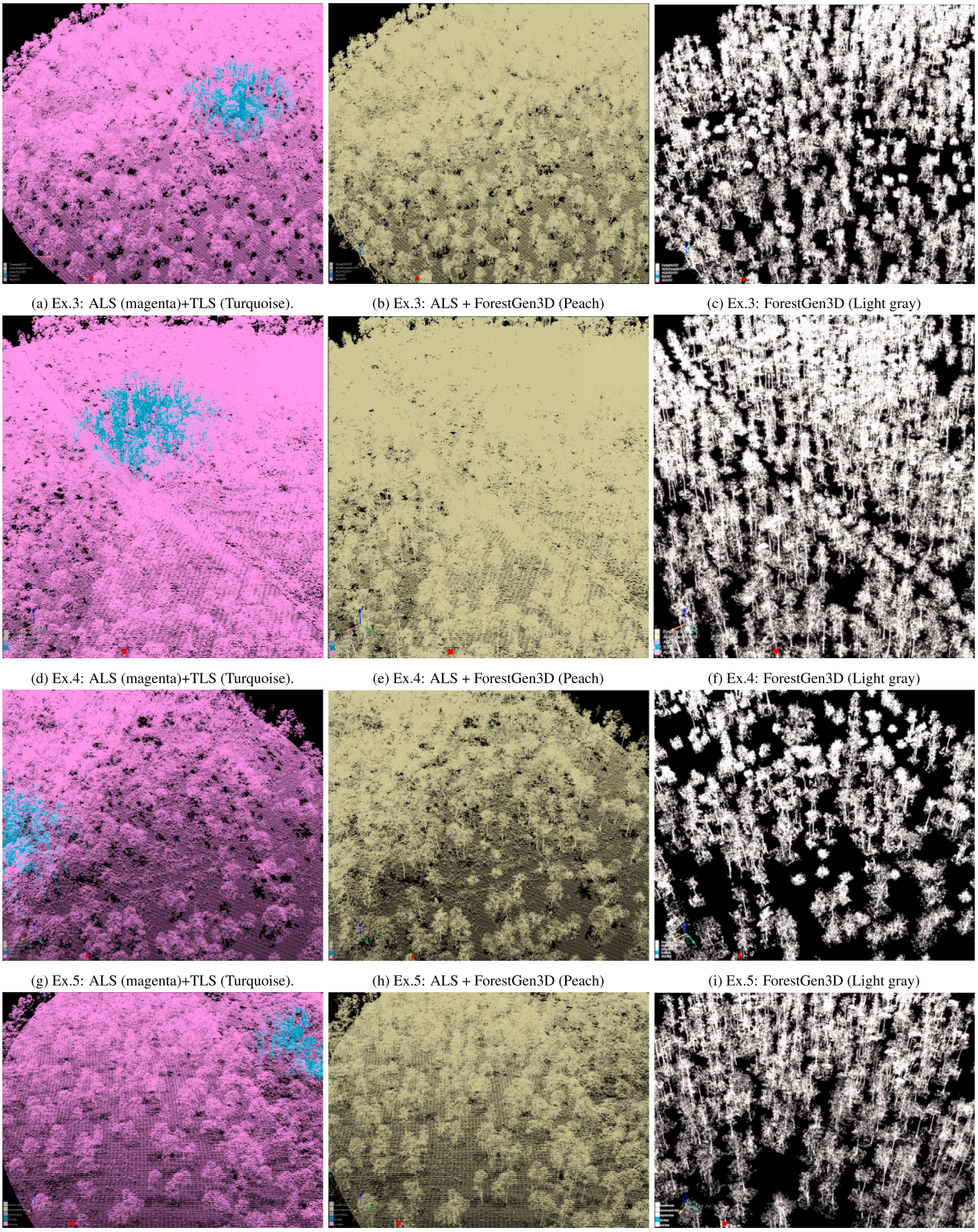
Fig. 14 presents ten illustrative samples from the model, trained on the 2900 extracted trees from TLS scans without using ALS conditioning. The generated tree structures exhibit fine details along the vertical tree-stand direction, mirroring the characteristics of the TLS training data. The model successfully learns the spatial distribution of key tree components: trunk, branches, leaves, and ground, along with their spatial relationships. Additionally, the model captures variations in canopy structure despite being trained on a relatively small dataset. The generated trees predominantly resemble mixed conifer species, including longleaf, slash, and loblolly pines, reflecting their prevalence in the training dataset. To the best of our knowledge, this model achieves an unprecedented level of detail in tree generation directly from real LiDAR forestry data.

Appendix C. Additional comparisons of biophysical metric distributions

Figs. 15, 16, and 17 present additional plot-scale examples (25 m radius) of biophysical metric distributions for tree height, DBH, crown diameter, and crown volume, derived from ALS, TLS, ALS+TLS, and ALS+ForestGen3D point clouds. Across all examples, ALS+ForestGen3D consistently compensates for ALS structural blind spots while preserving ALS-derived spatial geometry. Tree height and crown diameter distributions remain largely contained within ALS-derived ranges, whereas the most pronounced improvements are observed for DBH and crown volume, reflecting recovery of sub-canopy and lower-canopy structure absent in ALS-only data. Collectively, these results further confirm that ForestGen3D produces structurally plausible 3D vegetation whose biometric distributions closely align with TLS-derived references while maintaining fidelity to ALS-provided canopy geometry.

Appendix D. Additional regional-scale generations

This section provides additional visual examples of ForestGen3D outputs at the regional scale. Figs. 18 and 19 show spatial snapshots of 200 m radius regions, with point clouds color-coded as follows: magenta for ALS, turquoise for TLS, peach for ALS combined with ForestGen3D output, and white for ForestGen3D-generated points alone. Note that ForestGen3D is applied only within regions identified as occluded in the ALS data. Thus, figures displaying only the generated points contain empty areas corresponding to locations where no synthesis was performed. When combined with the original ALS observations, these regions are filled by existing ALS ground and canopy returns, resulting in a complete point cloud representation. The generated points primarily populate the lower canopy and sub-canopy layers and extend toward the ground surface, following the spatial constraints imposed



(a) Ex.3: ALS (magenta)+TLS (Turquoise). (b) Ex.3: ALS + ForestGen3D (Peach) (c) Ex.3: ForestGen3D (Light gray)

(d) Ex.4: ALS (magenta)+TLS (Turquoise). (e) Ex.4: ALS + ForestGen3D (Peach) (f) Ex.4: ForestGen3D (Light gray)

(g) Ex.5: ALS (magenta)+TLS (Turquoise). (h) Ex.5: ALS + ForestGen3D (Peach) (i) Ex.5: ForestGen3D (Light gray)

(j) Ex.6: ALS (magenta)+TLS (Turquoise). (k) Ex.6: ALS + ForestGen3D (Peach) (l) Ex.6: ForestGen3D (Light gray)

Fig. 19. Additional landscape-scale ForestGen3D examples. Rows show different 200 m radius areas; columns show ALS+TLS (left), ALS+ForestGen3D (middle), and ForestGen3D alone (right). Colors denote ALS (magenta), TLS (turquoise), ALS+ForestGen3D (peach), and generated points (white).

by the ALS geometry. The synthesized structures exhibit continuous transitions between ALS-observed and generated regions, yielding coherent tree-level and stand-level configurations. Each 200 m radius region contains approximately 3000 trees. With an average inference time of about 0.4 s per tree, generation of a full region requires on the order of 20 minutes, illustrating the computational feasibility of applying ForestGen3D at regional scales.

References

Achlioptas, P., Diamanti, O., Mitliagkas, I., Guibas, L., 2018. Learning representations and generative models for 3d point clouds. In: International Conference on Machine Learning. PMLR, pp. 40–49.

Bechtold, W.A., Patterson, P.L., 2005. The Enhanced Forest Inventory and Analysis Program National Sampling Design and Estimation Procedures. no. 80, USDA Forest Service, Southern Research Station.

Calders, K., Adams, J., Armston, J., Bartholomeus, H., Bauwens, S., Bentley, L.P., Chave, J., Danson, F.M., Demol, M., Disney, M., et al., 2020. Terrestrial laser scanning in forest ecology: Expanding the horizon. Remote Sens. Environ. 251, 112102.

Castorena, J., Creusere, C.D., 2015. Sampling of time-resolved full-waveform LIDAR signals at sub-Nyquist rates. IEEE Trans. Geosci. Remote Sens. 53 (7), 3791–3802.

Castorena, J., Dickman, L.T., Killebrew, A.J., Gattiker, J.R., Linn, R., Loudermilk, E.L., 2025. ForestAlign: Automatic forest structure-based alignment for multi-view TLS and ALS point clouds. Sci. Remote. Sens. 11, 100194.

Cover, T.M., Thomas, J.A., 2006. Elements of Information Theory, second ed. Wiley-Interscience.

de Wergifosse, L., André, F., Goosse, H., Boczon, A., Cecchini, S., Ciceu, A., Collalti, A., Cools, N., D'Andrea, E., De Vos, B., Hamdi, R., Ingerslev, M., Knudsen, M.A., Kowalska, A., Leca, S., Matteucci, G., Nord-Larsen, T., Sanders, T.G., Schmitz, A., Termonia, P., Vanguelova, E., Van Schaeybroeck, B., Verstraeten, A., Vesterdal, L., Jonard, M., 2022. Simulating tree growth response to climate change in structurally diverse oak and beech forests. Sci. Total Environ. 806, 150422. http://dx.doi.org/10.1016/j.scitotenv.2021.150422, URL https://www.sciencedirect.com/science/article/pii/S0048969721054991.

Dufour-Kowalski, S., Courbaud, B., Dreyfus, P., Meredieu, C., De Coligny, F., 2012. Capsis: an open software framework and community for forest growth modelling. Ann. For. Sci. 69, 221–233.

Fan, H., Su, H., Guibas, L.J., 2017. A point set generation network for 3d object reconstruction from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 605–613.

Hawley, C.M., Loudermilk, E.L., Rowell, E.M., Pokswinski, S., 2018. A novel approach to fuel biomass sampling for 3D fuel characterization. MethodsX 5, 1597–1604.

Ho, J., Jain, A., Abbeel, P., 2020. Denoising diffusion probabilistic models. Adv. Neural Inf. Process. Syst. 33, 6840–6851.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., Maltamo, M., 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. Int. J. Remote Sens. 29 (5), 1339–1366.

Hyyppä, J., Yu, X., Hyyppä, H., Vastaranta, M., Holopainen, M., Kukko, A., Kaartinen, H., Jaakkola, A., Vaaja, M., Koskinen, J., et al., 2012. Advances in forest inventory using airborne laser scanning. Remote. Sens. 4 (5), 1190–1207.

Interactive Data Visualization, Inc., 2024. SpeedTree. URL https://store.speedtree.com/. Version 10.x.x. Interactive Data Visualization, Inc..

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van Ewijk, K.Y., Forrester, D.I., Haeni, M., et al., 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. Global Change Biol. 23 (1), 177–190.

Kane, V.R., McGaughey, R.J., Bakker, J.D., Gersonde, R.F., Lutz, J.A., Franklin, J.F., 2010. Comparisons between field-and LiDAR-based measures of stand structural complexity. Can. J. Forest Res. 40 (4), 761–773.

Köhler, P., Huth, A., 1998. The effects of tree species grouping in tropical rainforest modelling: simulations with the individual-based model FORMIND. Ecol. Model. 109 (3), 301–321.

Linn, R.R., Goodrick, S.L., Brambilla, S., Brown, M.J., Middleton, R.S., O'Brien, J.J., Hiers, J.K., 2020. QUIC-fire: A fast-running simulation tool for prescribed fire planning. Environ. Model. Softw. 125, 104616.

Linn, R., Reisner, J., Colman, J.J., Winterkamp, J., 2002. Studying wildfire behavior using FIRETEC. Int. J. Wildland Fire 11 (4), 233–246.

Loudermilk, E.L., Hiers, J.K., O'Brien, J.J., 2017. The role of fuels for understanding fire behavior and fire effects. In: Ecological Restoration and Management of Longleaf Pine Forests. CRC Press, pp. 107–122.

Loudermilk, E.L., Hiers, J.K., O'Brien, J.J., Mitchell, R.J., Singhania, A., Fernandez, J.C., Cropper, W.P., Slatton, K.C., 2009. Ground-based LIDAR: a novel approach to quantify fine-scale fuelbed characteristics. Int. J. Wildland Fire 18 (6), 676–685.

Lyu, Z., Kong, Z., Xu, X., Pan, L., Lin, D., 2021. A conditional point diffusion-refinement paradigm for 3d point cloud completion. arXiv preprint arXiv:2112.03530.

Mallet, C., Bretar, F., 2009. Full-waveform topographic lidar: State-of-the-art. ISPRS J. Photogramm. Remote Sens. 64 (1), 1–16.

Marcozzi, A., Wells, L., Parsons, R., Mueller, E., Linn, R., Hiers, J.K., 2025. FastFuels: Advancing wildland fire modeling with high-resolution 3D fuel data and data assimilation. Environ. Model. Softw. 183, 106214. http://dx.doi.org/10.1016/j.envsoft.2024.106214.

Maréchaux, I., Langerwisch, F., Huth, A., Bugmann, H., Morin, X., Reyer, C., Seidl, R., Collalti, A., Dantas de Paula, M., Fischer, R., et al., 2021. Tackling unresolved questions in forest ecology: The past and future role of simulation models. Ecol. Evol. 11, 3746–3770.

Nelson, R., Krabill, W., MacLean, G., 1984. Determining forest canopy characteristics using airborne laser data. Remote Sens. Environ. 15 (3), 201–212.

Panaretos, V.M., Zemel, Y., 2019. Statistical aspects of Wasserstein distances. Annu. Rev. Stat. Appl. 6 (1), 405–431.

Perez, E., Strub, F., De Vries, H., Dumoulin, V., Courville, A., 2018. Film: Visual reasoning with a general conditioning layer. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 32, 1.

Pokswinski, S., Gallagher, M.R., Skowronski, N.S., Loudermilk, E.L., Hawley, C., Wallace, D., Everland, A., Wallace, J., Hiers, J.K., 2021. A simplified and affordable approach to forest monitoring using single terrestrial laser scans and transect sampling. MethodsX 8, 101484.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Adv. Neural Inf. Process. Syst. 30.

Rijal, B., Weiskittel, A.R., Kershaw, Jr., J.A., 2012. Development of height to crown base models for thirteen tree species of the North American Acadian region. For. Chron. 88 (1), 60–73.

Snider, M., Pokswinski, S., Wilder, S., 2025. Fort Stewart 2025: Contextual data for wildland fire science initiative (WFSI) research burn campaign. http://dx.doi.org/10.60594/W4S88P, WFSI Data Portal.

Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B., 2020. Score-based generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456.

Technologies, U., 2024. Unity terrain engine. URL https://unity.com/. Part of the Unity engine, for creating terrain. Unity Technologies.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., Clawges, R.M., 2008. Lidar: shedding new light on habitat characterization and modeling. Front. Ecol. Environ. 6 (2), 90–98.

Windrim, L., Bryson, M., 2019. Forest tree detection and segmentation using high resolution airborne lidar. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, IEEE, pp. 3898–3904.

Wu, J., Zhang, C., Xue, T., Freeman, B., Tenenbaum, J., 2016. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. Adv. Neural Inf. Process. Syst. 29.

Yang, G., Huang, X., Hao, Z., Liu, M.-Y., Belongie, S., Hariharan, B., 2019. Pointflow: 3d point cloud generation with continuous normalizing flows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4541–4550.

Yang, J., Kang, Z., Cheng, S., Yang, Z., Akwensi, P.H., 2020. An individual tree segmentation method based on watershed algorithm and three-dimensional spatial distribution analysis from airborne LiDAR point clouds. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 13, 1055–1067.

Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M., 2018. Pcn: Point completion network. In: 2018 International Conference on 3D Vision. 3DV, IEEE, pp. 728–737.

Zhu, Z., Kleinn, C., Nölke, N., 2021. Assessing tree crown volume—A review. For.: Int. J. For. Res. 94 (1), 18–35.