



High-resolution multi-temporal forest canopy height mapping in California using GEDI LiDAR and multi-sensor remote sensing

Md Didarul Islam^{a,*}, Andrew S. Johnson^b, Jessica Miesel^a, Matthew Dickinson^c, Jens Stevens^d

^a Department of Forest, Rangeland and Fire Sciences, University of Idaho, USA

^b Department of Environmental Science, Policy, and Management, University of California, Berkeley, USA

^c USDA Forest Service Northern Research Station, Delaware, OH, USA

^d USDA Forest Service Pacific Northwest Research Station, Seattle, WA, USA

ARTICLE INFO

Keywords:

forest canopy height
GEDI LiDAR
multi-sensor remote sensing
XGBoost
machinelearning
uncertainty quantification
forest structure
multi-temporal mapping
wildfire disturbance

ABSTRACT

Accurate, high-resolution forest canopy height information is critical for monitoring forest structure, estimating aboveground biomass, and informing land management and climate resilience strategies. In particular, remotely-sensed canopy height estimates can provide an estimate of fuel volume for wildland fire effects analyses, including in the Wildland Urban Interface where fuels in the home ignition zone are influential. This study presents a scalable machine learning framework to estimate canopy height across California by integrating Global Ecosystem Dynamics Investigation (GEDI) LiDAR Relative Height measurements with a suite of multi-sensor remote sensing data. The model was trained on more than 5.6 million rigorously filtered GEDI footprints from California and Arizona to minimize noise and ensure robust performance across ecologically diverse and topographically complex landscapes. Predictor variables included optical (Sentinel-2), radar (Sentinel-1, PALSAR-2), topographic, land cover, and a range of derived spectral indices and spatial-textural features. Model validation using an independent GEDI hold-out set demonstrated strong performance, with an R^2 of 0.72, RMSE of 6.05 m, MAE of 4.31 m, and a slight positive bias of 0.80 m, closely matching the training metrics, suggesting a well-calibrated model. Compared to existing global canopy height products, the proposed regional approach shows improved agreement with GEDI observations in complex terrain and dense vegetation areas within California, while reflecting differences in training scope and objectives between regional and global models. A key outcome of this study is a 10-m resolution biennial canopy height time series for California (2016, 2018, 2020, 2022, and 2024) suitable for stand-scale and multi-epoch structural assessment. The resulting 10-m, biennial California canopy-height series for 2016–2024 is used to assess regrowth trajectories of canopy height and biomass in four major wildfire areas.

1. Introduction

Forest canopy height is a key structural attribute that reflects the vertical complexity of vegetation and plays an important role in characterizing ecosystem productivity, biomass distribution, and habitat structure (Lefsky et al., 2005; Shu et al., 2023; Zhang et al., 2014). It is widely used in applications such as aboveground biomass (AGBD) estimation, carbon accounting, biodiversity modeling, and forest management planning. The ability to monitor canopy height over time is increasingly relevant because of long-term forest changes in a changing climate and shorter-term disturbance impacts such as wildfire, drought,

or logging, and for tracking forest recovery and success on (Lang et al., 2023; Zang et al., 2024). While significant progress has been made in remote-sensing based forest monitoring, the availability of canopy height products with both high spatial resolution and consistent temporal coverage across large landscapes remains limited. Existing global datasets derived from spaceborne LiDAR and optical imagery offer valuable insights but often lack the resolution, regional specificity, or temporal granularity needed for local to sub-national decision-making (Lang et al., 2023; Potapov et al., 2021; Tolan et al., 2024). These gaps highlight the need for regionally optimized temporally flexible modeling approaches that can generate reliable canopy height estimates across

* Corresponding author.

E-mail address: didarul.ku.14@gmail.com (M.D. Islam).

<https://doi.org/10.1016/j.srs.2026.100488>

Received 25 February 2026; Received in revised form 3 August 2026; Accepted 3 August 2026

Available online 6 August 2026

2666-0172/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

diverse forested landscapes.

The western United States, and California in particular, exemplify a region where timely forest structure data are essential. California's forests encompass diverse ecological zones and complex terrain, but they are increasingly affected by overlapping stressors including prolonged droughts, pest outbreaks, intensified wildfires, and rapid land-use change (Westerling et al., 2011; Westerling and Bryant, 2008; Wong et al., 2020). Decades of fire suppression, combined with rising temperatures and increased ignition sources, have altered natural fire regimes, leading to excessive fuel buildup and shifts in canopy structure towards higher continuous cover and less spatial heterogeneity (North et al., 2015; Parks et al., 2016; Stephens et al., 2020a). Remote sensing decision support for wildland fuel management in this context primarily relies on spectral image differencing, which requires extensive ecosystem-specific calibration to accurately inform forest restoration efforts, fuel treatment prioritization, and post-disturbance recovery assessments (Miller and Thode, 2007; Miller and Quayle, 2015; Islam et al., 2022). Incorporating accurate, spatially explicit forest structure data such as canopy heights may decrease analytical workloads and significantly enhance ecosystem management and policy decisions in regions undergoing rapid ecological change.

Traditional approaches to canopy height mapping, such as field-based inventories and airborne LiDAR, are highly accurate but limited in spatial and temporal coverage due to their cost and operational constraints (Dubayah et al., 2020; Wulder et al., 2012). Recent advances in satellite remote sensing have improved the scalability of vegetation monitoring, yet most global canopy height products remain coarse in resolution (≥ 30 m) and limited in their ability to resolve structural heterogeneity within complex landscapes (Hayashi et al., 2013; Lang et al., 2023; Lefsky et al., 2005; Potapov et al., 2021; Saatchi and Favrichon, 2024; Wang et al., 2014; Xiao et al., 2025; Islam et al., 2026; Islam et al., 2024). The advent of NASA's GEDI LiDAR mission provides a valuable opportunity to address these challenges (Dubayah et al., 2020; Hancock et al., 2019). GEDI offers accurate measurements at the footprint level, that is, ~ 25 -m-diameter circular laser spots that return a vertical waveform of the canopy structure; however, coverage is limited to narrow orbital swaths, so predictive models are needed to generate continuous canopy height surfaces (Dubayah et al., 2020).

Machine Learning (ML) models have emerged as powerful tools for spatially generalizing GEDI observations by integrating a range of ancillary datasets, including optical imagery, synthetic aperture radar (SAR), and topographic variables. These models are particularly well-suited for capturing complex, nonlinear relationships between forest structure and environmental predictors, enabling the production of scalable and spatially continuous canopy height maps over large regions. Among the most widely used global canopy height products developed through such approaches are those by Potapov et al. (2021) and Lang et al. (2023), hereafter referred to as the UMD 2019; ETH 2019 products, respectively. The UMD 2019 product leverages Landsat-based spectral bands and vegetation indices to generate a 30-m resolution global canopy height dataset. In contrast, the ETH 2019 product utilizes Sentinel-2 imagery and ML to produce canopy height estimates at a finer 10-m resolution, and is widely regarded as one of the most reliable open-access datasets available for global forest monitoring (Lang et al., 2023; Potapov et al., 2021).

Despite substantial progress made by existing global Canopy Height Models (CHM), several limitations continue to constrain their accuracy and generalizability, particularly when applied across structurally diverse landscapes such as those found in California, or when extended across multiple time periods. First, most global models rely on GEDI data collected from a single year, which restricts their temporal transferability. As a result, these models often struggle to generalize across interannual variability in vegetation phenology, post-disturbance regrowth, and changing climate conditions (Lang et al., 2023; Potapov et al., 2021). Second, many of these models and data products lack the

spatial granularity and context sensitivity required for stand-level forest management. In regions with complex terrain, heterogeneous vegetation types, and diverse successional stages, such models may overlook key spatial patterns such as topographic shading or localized forest structure, ultimately limiting their ability to accurately capture canopy height variability (Long and Hatten, 2023; Potapov et al., 2021; Saatchi and Favrichon, 2024; Xiao et al., 2025).

A particularly persistent challenge lies in estimating canopy height in ecosystems with exceptionally tall forest structures and topographic complexity, such as California's Coast Range, Sierra Nevada, and Southern Cascades mountains. Coast Range forests feature coast redwood (*Sequoia sempervirens*), which regularly exceeds 80 m in canopy height. Significant difficulties for remote sensing exist here due to dense vertical layering, persistent fog, and shadowing effects, which diminish the reliability of both optical and radar signals. Regions featuring steep terrain further complicate sensor viewing geometries and increase the potential for geolocation or elevation-related errors in satellite data. Previous studies, including those by Potapov et al. (2021) and Lang et al. (2023), have consistently demonstrated that model performance declines in such high-biomass environments, particularly when training datasets underrepresent extreme canopy heights. In these cases, ML models tend to underfit the upper end of the canopy height distribution, with predictive saturation commonly occurring around 40 to 50 m. As a result, canopy heights in redwood forests or dense mixed-conifer stands are often systematically underestimated by 10 m or more. This issue is exacerbated in regions with mixed vegetation mosaics, where tall forest patches are adjacent to lower-stature vegetation types such as shrublands, regenerating forests, or agricultural clearings. The resulting spectral and structural heterogeneity poses challenges for models that lack spatial awareness or cannot account for contextual transitions between vegetation types.

These limitations underscore the urgent need for modeling frameworks that are better equipped to handle structural extremes and spatial complexity. Incorporating diverse training samples particularly from high-canopy and mixed-structure ecosystems along with multitemporal data that captures forest dynamics over time is essential for enhancing model robustness and generalizability. A more nuanced and adaptive approach to canopy height modeling is critical for accurately characterizing forest structure in ecologically variable biomes, with direct implications for ecosystem management and monitoring.

To address these gaps, this study develops a scalable ML based framework to estimate canopy height across California at 10-m resolution over biennial intervals. By integrating ~ 5.6 million quality control GEDI footprints with a multi-sensor remote sensing dataset including Sentinel-2 spectral bands, topographic metrics, radar-derived features from Sentinel-1 and ALOS PALSAR-2, land cover, and derived spectral indices related to vegetation, we generated a multitemporal canopy height product spanning from 2016 to 2024 in two-year increments. The model was trained on ecologically and topographically representative samples from both California and Arizona to ensure robust spatial generalization across diverse forest types and environmental gradients. The objectives of this work are threefold: (1) to develop a regionally calibrated modeling and mapping workflow that extends GEDI-trained canopy height estimation to years prior to the GEDI mission using year-specific predictor composites; (2) to generate a 10-m resolution canopy height dataset suitable for stand-level analysis across a large, topographically and ecologically diverse region; and (3) to provide an open-access, transferable modeling workflow for scalable, interpretable, and policy-relevant forest mapping and monitoring in other regions. Collectively, this study aims to fill a critical data gap in spatiotemporally resolved canopy height estimation, supporting more effective strategies for forest management, post-disturbance assessment, and climate resilience planning.

This study makes four practical contributions. First, it delivers a California-focused, 10 m canopy height product optimized for complex terrain using GEDI RH98 as reference and multi-sensor predictors.

Second, it provides a scalable tile-based implementation in Google Earth Engine for statewide inference. Third, it introduces a consistent biennial canopy-height series (2016–2024) designed for stand-scale mapping and disturbance-response monitoring. Finally, it provides clear appropriate-use guidance by documenting key limitations, including temporal extrapolation uncertainty, spatial imbalance in GEDI sampling, and reduced accuracy in very tall canopies.

2. Materials & methods

The canopy height mapping framework developed in this study follows an integrated, multi-sensor workflow designed to maximize structural information extraction and modeling efficiency (Fig. 1). The process begins with the acquisition of diverse satellite and ancillary datasets, which are transformed into informative predictor variables through spectral index computation, radar polarization combinations, textural measures such as GLCM, and spatial statistics (mean and standard deviation) calculated over 3×3 pixel neighborhoods. Model training is performed using optimized ML algorithms, followed by validation against independent GEDI reference data and airborne LiDAR datasets.

2.1. Study area

This study focused on California, a geographically and ecologically diverse state in the western United States. The landscape spans a wide range of topographic and climatic zones, from coastal ranges and inland

valleys to alpine mountains and arid deserts. This physical diversity supports a rich variety of forest ecosystems, including towering old-growth redwoods, dense mixed-conifer stands, oak woodlands, montane shrublands, and desert scrub. Forest structure and composition vary significantly with elevation, precipitation, soil type, and disturbance history, resulting in highly heterogeneous vegetation patterns across the state (Olson et al., 2001).

California includes thirteen distinct Level III ecoregions as defined by the U.S. Environmental Protection Agency (EPA), each with characteristic vegetation dynamics (Fig. 2) (Griffith et al., 2016). The Sierra Nevada and Klamath Mountains support some of the tallest and densest conifer forests in the state, including species like Douglas-fir, sugar pine, and giant sequoia, with canopy heights often exceeding 50 m. The Coast Range features temperate forests with coast redwoods - the tallest trees in the world along with tanoak and Douglas-fir. The Cascades and Eastern Cascades contain mixed fir, pine, and hemlock forests, transitioning into drier open-canopy woodlands in the east. In contrast, the Central Valley, Central Basin, and Northern Basin have limited forest cover, consisting mostly of riparian strips, savannas, or desert shrublands, where canopy height is generally low. The California Chaparral and Pine-Oak Mountains are characterized by fire-adapted shrublands and mixed hardwood-conifer forests, while the Mojave and Sonoran Deserts exhibit minimal vegetation, with scattered desert trees and shrubs under arid conditions. This ecological gradient from coastal fog-fed forests to interior deserts creates complex spatial patterns in canopy height and biomass distribution. The distribution of GEDI samples across these ecoregions is quantified in Fig. 9c, enabling assessment

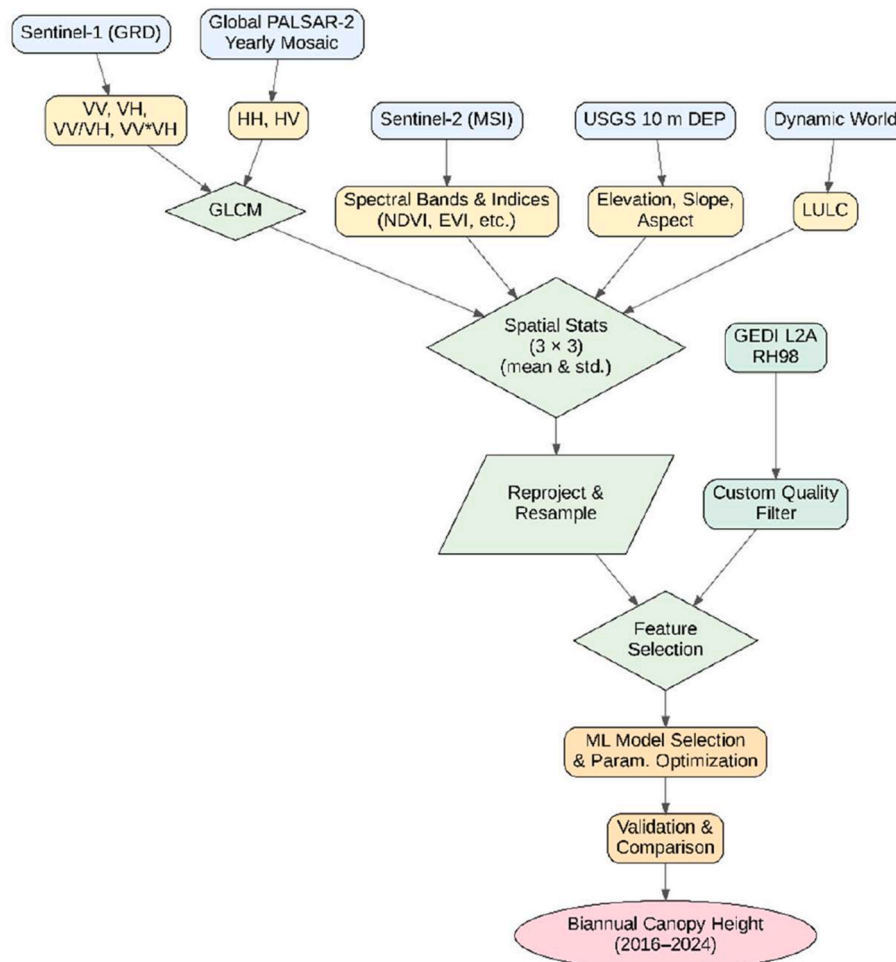


Fig. 1. Simplified methodological workflow outlining data acquisition, preprocessing, feature extraction, model training, validation, and biennial canopy height map generation for California.

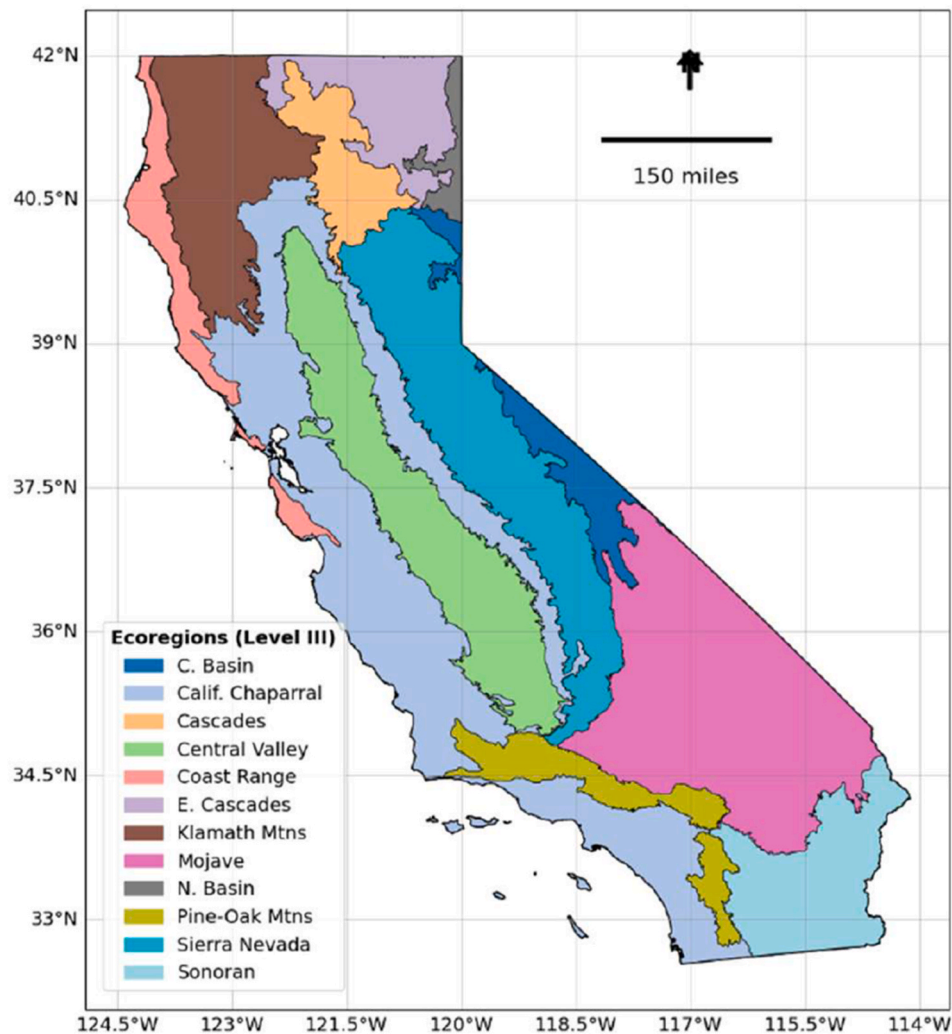


Fig. 2. Level III ecoregions of California based on the United States Environmental Protection Agency (EPA) classification (Griffith et al., 2016).

of whether regional performance differences relate to sampling density.

In addition to natural variability, California's forests have been increasingly shaped by disturbance, particularly wildfire. Wildfires burned approximately 6.9 million acres between 2020 and 2021, including the August Complex (2020) and Dixie Fire (2021), which were the two largest incidents in state history (CAL FIRE, 2024). Fire frequency and severity have escalated in recent decades, driven by the confluence of prolonged drought, rising temperatures, and decades of fuel accumulation due to historical fire suppression (Stephens et al., 2020b). These interacting climatic and anthropogenic drivers have led to major shifts in canopy structure, particularly in high-biomass regions such as the Sierra Nevada, Klamath Mountains, and California chaparral, making California a uniquely complex and dynamic landscape for studying forest structural change over time.

2.2. Data source

We compiled a comprehensive multi-source dataset to support canopy height modeling across California using remote sensing and environmental predictors. The target variable, canopy height, was derived from GEDI Level 2A Relative Height (RH) metrics, specifically the 98th percentile height (RH98). To ensure seasonal consistency and minimize the effects of snow cover and phenological variability, only GEDI acquisitions from April to September were retained. Predictor datasets were similarly restricted to this time frame to ensure temporal alignment.

The input features included a combination of optical, radar, topographic, and land cover data. Specifically, we incorporated Sentinel-1 backscatter (VV and VH), Sentinel-2 spectral bands and derived vegetation indices (e.g., NDVI, NIR), ALOS PALSAR-2 radar texture metrics, the USGS 10-m Digital Elevation Model (DEM) for elevation and slope, and Dynamic World 10-m land use/land cover classifications. Annual composites of optical and SAR predictors were created by calculating median reflectance or backscatter values from April to September, reducing atmospheric noise and aligning temporally with GEDI sampling. Each predictor was composited per year using the median of all valid observations within the April to September growing-season window, that is, a per-year median composite rather than a multi-year average or a maximum-value composite. Static topographic layers (elevation, slope, aspect) were used directly, and land cover was taken as the annual modal class. All datasets were processed separately by year and subsequently merged into a harmonized multitemporal database to support ML modeling. These annual predictor composites were later sampled at biennial output years (2016, 2018, 2020, 2022, 2024) to generate the final canopy-height series.

GEDI RH98 observations (2019–2022) are used to learn a statistical relationship between canopy height and multi-sensor predictors. We then apply the trained model to year-specific predictor composites to produce a consistent biennial series for 2016–2024. This “back-casting” does not assume constant canopy height through time. Instead, it assumes that the learned relationship between predictors and height is sufficiently stable to support broad-scale inference when predictors are

constructed consistently (same seasonal window, compositing strategy, and feature definitions). The resulting multi-temporal maps are intended primarily to capture major structural gradients and large changes (for example, disturbance and early recovery), rather than subtle growth increments.

2.2.1. GEDI level 2A

The GEDI mission, launched in 2018 aboard the International Space Station, provides high-resolution waveform LiDAR data for monitoring forest vertical structure. The Level 2A product contains geolocated RH metrics derived from waveform returns over ~25-m footprints. The geolocation of each GEDI shot is provided as the latitude and longitude of the highest return within the waveform (lat_highestreturn, lon_highestreturn), which represents the location of the tallest reflector within the footprint rather than the geometric footprint center. Predictor stacks were sampled at these coordinates and matched to the GEDI RH98 value at the same point. This approach was also used by [Lang et al. \(2023\)](#) in their global 10 m GEDI based canopy height product and is supported by [Tang et al. \(2023\)](#), who report GEDI geolocation accuracy of approximately 10 m post correction, comparable to the predictor pixel size. The 25 m footprint nominal extent therefore overlaps the 10 m predictor pixel within the typical geolocation tolerance rather than introducing a fundamental scale mismatch, and the residual scale uncertainty is partially absorbed by the multi pixel spatial context features described in Section 3 (3×3 neighborhood mean and standard deviation statistics computed for each spectral and index predictor), which encode the local variability that any pixel scale supervision against a 25 m footprint would imply. Among these, RH98, the relative height at which 98% of the waveform energy has returned, is widely recognized as a robust proxy for top-of-canopy height due to its reduced sensitivity to sub-canopy noise and clutter ([Dubayah et al., 2020](#); [Hancock et al., 2019](#); [Li et al., 2023](#)). RH98 values were extracted from GEDI Level 2A data acquired between April and September, spanning the years 2019 to 2022. To ensure data reliability across California's ecologically and topographically diverse terrain, a series of quality filters were applied within Google Earth Engine (GEE), with slight modifications to the criteria proposed by [Moudry et al. \(2024\)](#). The filtering criteria used to ensure the reliability of GEDI RH98 observations are summarized in [Table 1](#). After filtering, a total of 5,681,743 high-quality RH98 footprints were retained for model training and validation. The footprints span the four GEDI-era years 2019–2022, with the dataset distributed approximately proportionally across years; the relatively lower 2022 contribution reflects the GEDI mission's storage of the ISS instrument in March 2023. Additionally, Aboveground Biomass Density (AGBD) estimates

from GEDI Level 4A were used to filter noisy canopy height observations by exploiting the allometric relationship between tree height and biomass, whereby taller canopies generally correspond to higher biomass (see section 3.3.1).

2.2.2. Sentinel 2

Sentinel-2 provides high-resolution multispectral imagery that is widely used in vegetation monitoring and forest structure analysis. With its 10–60 m spatial resolution and 5-day revisit frequency, A suite of Sentinel-2 based spectral bands and vegetation indices were used to capture spatial variation in canopy density, moisture content, and photosynthetic activity providing valuable indirect information related to forest height, particularly when combined with topographic and structural predictors. All spectral bands with a native resolution greater than 20 m were resampled to 10 m using nearest neighbor interpolation to maintain consistent spatial alignment.

2.2.3. Sentinel 1

Sentinel-1 is a C-band Synthetic Aperture Radar (SAR) satellite mission that provides all-weather, day-and-night imaging capabilities, making it especially valuable for forest monitoring in regions frequently affected by cloud cover or complex topographic shading. Its dual-polarization backscatter data (VV and VH), available at 10-m spatial resolution, offer key forest structural information. In forested landscapes, Sentinel-1 backscatter is sensitive to vertical structure and biomass density through scattering processes such as surface, volume, and double-bounce interactions ([Fang et al., 2023](#); [Lang et al., 2023](#); [Nandy et al., 2021](#)). Sentinel-1 derived backscatter coefficients and polarization ratios were used as predictor variables to capture spatial variation in canopy structure. These radar-based metrics complement optical indices by improving sensitivity to structural complexity and surface roughness, thereby contributing to more accurate and spatially consistent canopy height estimation across California's diverse forest ecosystems.

2.2.4. PalsAR-2

PALSAR-2 is an L-band Synthetic Aperture Radar (SAR) sensor on-board the ALOS-2 satellite, well-suited for capturing forest structure due to its deeper canopy penetration compared to shorter-wavelength sensors ([Cartus et al., 2012](#); [Shimada et al., 2014](#)). We used the Level 2.2 ScanSAR backscatter products, which provide terrain-flattened gamma-naught coefficients for HH and HV polarizations over a wide 350 km swath at 25-m spatial resolution. For integration with other datasets, all PALSAR-2 layers were resampled to 10-m resolution using nearest-neighbor interpolation. The L-band backscatter from PALSAR-2 enhances sensitivity to canopy height and sub-canopy structure, particularly in dense and mountainous forest regions, complementing optical and C-band radar inputs in capturing vertical forest characteristics.

2.2.5. USGS 3DEP 10m DEM

Topographic information was incorporated using the 10-m resolution DEM from the U.S. Geological Survey's 3D Elevation Program (3DEP). This seamless elevation dataset provides full coverage across the conterminous United States (CONUS), with consistent vertical accuracy and national standardization. The DEM was used to characterize terrain variability, which is known to influence forest structure and sensor observation geometry. From the DEM, slope and aspect were derived to capture the influence of topographic gradients and solar orientation on canopy height patterns. These topographic variables provide valuable context for interpreting spatial variation in vegetation structure, especially in mountainous regions where elevation and exposure play a critical role.

2.2.6. NLCD & Dynamic World LULC

To improve the accuracy and contextual filtering of canopy height

Table 1

Summary of GEDI Level 2A filtering criteria used to retain high-quality canopy height measurements for model training.

Filter Condition	Description	Purpose
quality_flag = 1	Only high-quality waveforms	Ensures valid GEDI shots
degrade_flag = 0	No signal degradation	Removes compromised returns
num_detectedmodes > 0	Valid waveform mode detected	Confirms canopy-ground separation
sensitivity > 0.90	High sensitivity threshold	Filters weak or noisy returns
beam < 8	Exclude power beams	Retains coverage and full-power beams
elev_lowestmode > 0	Valid ground elevation detection	Removes ground-detection errors
rh98 > 0 and < 80	Physiologically valid canopy height	Excludes extreme outliers
surface_flag = 1	Surface identified	Confirms valid surface return
elevation_bias_flag = 0	No elevation bias present	Ensures vertical accuracy
LULC	Exclude water, built-up, bare, snow/ice	Retains only vegetated canopy

predictions, we incorporated two complementary land use/land cover (LULC) datasets: Dynamic World and the National Land Cover Database (NLCD). Dynamic World is a near-real-time, globally available LULC dataset developed by Google, assigning one of nine land cover classes (e. g., trees, grass, crops, built-up, barren, snow/ice) at each 10-m pixel (Brown et al., 2022). We used LULC both as a predictor and to mask non-vegetated pixels during model development. To reduce short-term noise and ensure seasonal consistency, we generated an April–September composite for each year by assigning the per-pixel modal class. This yielded a stable seasonal LULC layer aligned with the GEDI sampling window and the predictor variables.

NLCD (National Land Cover Database) is a U.S.-specific, 30-m resolution LULC dataset maintained by the Multi-Resolution Land Characteristics (MRLC) consortium (Homer et al., 2020). It provides thematically detailed land cover classifications across the conterminous United States for multiple years with consistent methodology that supports temporal analysis. We used NLCD at the post-processing stage to filter out non-vegetated areas from the final canopy height products. By aligning canopy height predictions with NLCD's forest-related classes, we ensured that modeled heights were retained only in ecologically valid vegetated regions. Together, these two LULC products enabled high-resolution, temporally aligned masking and filtering during both the model training and final product refinement stages.

2.3. Data processing & feature extraction

2.3.1. Filtering & improving GEDI RH98 data quality

To further improve the reliability of model training, we applied a targeted data cleaning step to remove implausible observations from the joint RH98 and AGBD distribution. Specifically, we sought to eliminate cases where RH98 values were disproportionately high or low relative to observed AGBD, which most commonly reflect known failure modes that affect both quantities, namely GEDI waveform misinterpretation, geolocation error, or allometric model breakdown in a typical canopy structure (Roy et al., 2021; Duncanson et al., 2022). A visual inspection of the raw data revealed a well defined allometric envelope, with most observations falling within a plausible biomass and height domain. To

formalize this boundary, we fitted a third degree polynomial regression to the mean RH98 and AGBD relationship ($R^2 = 0.882$) and defined upper and lower quantile regression limits ($q = 0.995$ and $q = 0.010$). Observations falling outside these bounds were considered implausible and removed. The cleaned dataset retained 5,594,978 out of 5,681,743 observations ($\sim 98.5\%$), preserving the realistic variability in canopy height and biomass while excluding extreme outliers (Fig. 3). This step functions as a conservative quality control screen applied prior to model fitting and was implemented uniformly across the training data; it is not tuned using map outputs or evaluation results. We note that GEDI L4A AGBD is derived from GEDI RH metrics via species and plant functional type specific allometric models (Duncanson et al., 2022), so the joint RH98 and AGBD distribution is not a comparison of two fully independent measurements. The filter is therefore not a verification of RH98 against an external standard but a consistency check between two GEDI derived quantities that share an underlying biophysical signal, and points outside the envelope mark disagreement between the two estimates, which in practice arises from the failure modes listed above. We acknowledge that some valid ecological outliers, for example tall thin trees or short dense trees, may also fall outside the envelope, but the conservative threshold ($q = 0.01, 0.995$) and high retention rate ($\sim 98.5\%$) limit the magnitude of this risk. We considered using only physical plausibility bounds on RH98 alone (for example, $0 < \text{RH98} < 90$ m) without the joint envelope, but this admits known waveform derived retrieval failures that degrade the model more substantially than the loss of a small fraction of ecological outliers.

After filtering out implausible RH98–AGBD outliers, we examined the distribution of the remaining RH98 values to better characterize the structural composition of the training dataset and evaluate implications for model performance. The raw RH98 distribution is strongly right-skewed, with most canopy height values concentrated between 5 and 25 m and a pronounced peak near 10–15 m (Figure A1, Appendix). The distribution gradually tapers off toward taller canopy heights, reflecting the natural prevalence of mid-story vegetation and the relative rarity of exceptionally tall forest structures in California. The overlaid red curve, representing a fitted normal distribution, demonstrates a poor alignment with the histogram particularly in the tail highlighting the non-Gaussian

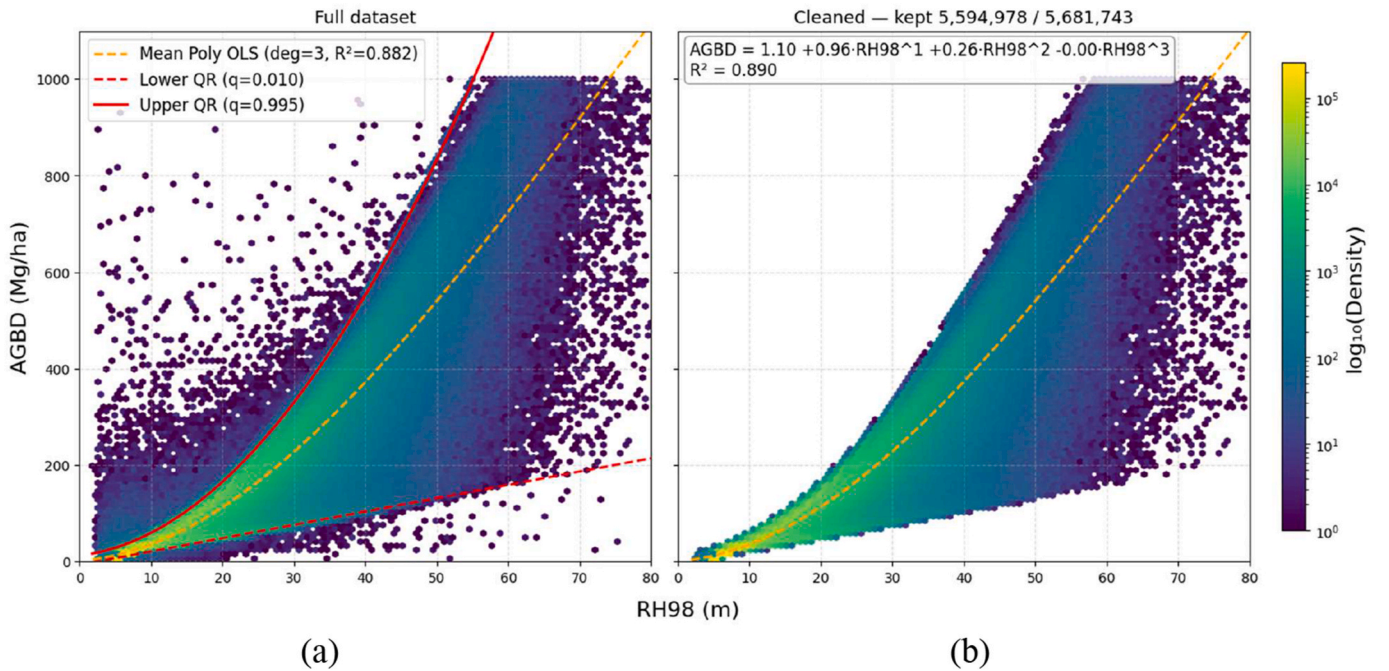


Fig. 3. Relationship between GEDI relative height at 98th percentile (RH98, in m) and Aboveground Biomass Density (AGBD, in Mg ha^{-1}). (a) Full dataset with mean third-degree polynomial regression (yellow dashed line) and upper and lower quantile regression bounds at $q = 0.995$ and $q = 0.010$ (red and orange lines). (b) Cleaned dataset after applying quantile-based filtering, retaining $\sim 98.5\%$ of the original observations.

nature of the RH98 data. Such skewness can lead to variance instability and biased residuals, potentially degrading the performance of regression-based ML models. To address this issue, we applied a natural logarithmic transformation to the RH98 values prior to model training. The resulting log-transformed distribution aligns more closely with the overlaid normal distribution curve. The transformation significantly reduced skewness, stabilized variance, and compressed the influence of high-end outliers while preserving the ordinal relationships among observations. Model predictions were back transformed to the original scale via the inverse $\expm1$ operation. We acknowledge that this back transform is non linear and that, under standard assumptions, the predicted conditional median in log space recovers the conditional median rather than the conditional mean on the original scale (Jensen's inequality). To check the magnitude of this potential bias, we computed the Duan (1983) smearing estimator from the training residuals: $S = \text{mean}(\exp(\epsilon)) = 0.95$, which would correspond to a uniform downward shift of approximately 5%. Applying smearing as a multiplicative correction shifted the global test bias from +0.77 m to -0.10 m but worsened the underestimation in tall canopies (>40 m), where bias became more negative by approximately 2 m in each of the 40–50, 50–60, 60–70, and 70–90 m bins. This empirical result indicates that the height dependent underestimation observed in our model is not driven by the log back transform but by predictor saturation in spectral and radar features at the upper canopy range, consistent with the saturation patterns reported in other GEDI based canopy height products (Potapov et al., 2021; Lang et al., 2023). We therefore did not apply a smearing correction in the released product.

2.3.2. Spectral, textural, and statistical feature extraction

To enhance the spatial representation of canopy structure and minimize noise from pixel-level variability, we applied a 3×3 window-based neighborhood operation across all relevant spectral bands and derived indices. For each input variable, we computed two key spatial statistics: the mean and the standard deviation within the 3×3 kernel, capturing local central tendency and variability in spectral response. This approach allowed the model to incorporate contextual information that is often critical for identifying gradual changes in canopy density, vegetation transitions, or edge effects. In addition to basic statistical summaries, we extracted texture metrics from Synthetic Aperture Radar (SAR) datasets - Sentinel-1 (VV and VH) and PALSAR-2 (HH and HV) - using the Gray-Level Co-occurrence Matrix (GLCM) method with a 3×3 (size = 1) window. The texture features included contrast, dissimilarity, and sum average for each polarization channel. These metrics quantify spatial heterogeneity in backscatter values, offering insights into canopy roughness, vertical layering, and forest structure complexity that are not captured by raw intensity alone. All input layers, including original bands, vegetation indices, derived statistics, and texture measures, were resampled or processed to a uniform 10-m spatial resolution using nearest-neighbor interpolation. This ensured consistent pixel alignment across datasets sourced from varying sensors and spatial scales, such as Sentinel-2 (10–60 m), Sentinel-1 (10 m), PALSAR-2 (25 m), and USGS DEM (10 m). After harmonizing spatial resolution and stacking all bands, the final dataset comprised 64 variables (see Appendix A), representing a diverse set of spectral, structural, topographic, and texture-based predictors used in the canopy height mapping framework.

2.3.3. Feature selection using tree-based importance scores

To streamline the input predictor set and improve generalization, we employed a feature selection approach based on tree-based variable importance scores. Specifically, we used the SelectFromModel method from the scikit-learn library, which evaluates feature importance using a fitted XGBoost model (see Section 3.4.2 for modeling details). Importance scores in XGBoost are based on the gain metric, which quantifies the improvement in model performance contributed by each feature across decision tree splits. Features with importance scores above the median threshold were retained for subsequent modeling. From an

initial set of 64 (see Tabel A1, Appendix) variables including spectral bands, vegetation and burn indices, neighborhood statistics, texture features, topographic metrics, and land cover class, we selected a reduced set of 25 informative predictors (Table 2). These features capture diverse structural, spectral, and terrain-related signals important for explaining spatial variation in canopy height.

2.4. Model development & hyperparameter tuning

2.4.1. Model training and hyperparameter optimization

To model canopy height across California, we implemented a supervised ML framework using three regression-tree based ensemble candidate algorithms: eXtreme Gradient Boosting (XGBoost), Random Forest (RF), and Light Gradient Boosting Machine (LightGBM). To ensure optimal algorithm selection and mitigate manual bias in model configuration, we adopted an Automated Machine Learning (AutoML) strategy. This approach automatically evaluated and compared the performance of each candidate model using cross-validated metrics on the training set. The primary selection criteria were predictive accuracy and generalizability across ecologically diverse and topographically complex forest regions. The AutoML process identified XGBoost as the best-performing algorithm, offering a favorable balance between accuracy, speed, and interpretability. Its ability to handle nonlinear relationships and heterogeneous feature importance made it especially well-suited for integrating multi-source remote sensing and topographic data.

Table 2

List of selected predictor variables retained after feature selection using XGBoost gain-based importance.

Category	Variable	Description	Source/Method
Optical	B5	Red Edge 1 band	Sentinel-2 (raw band)
	B6	Red Edge 2 band	Sentinel-2 (raw band)
	B8	Near Infrared (NIR) band	Sentinel-2 (raw band)
	B11	Shortwave Infrared 1 (SWIR1) band	Sentinel-2 (raw band)
	B12	Shortwave Infrared 2 (SWIR2) band	Sentinel-2 (raw band)
SAR	VV	VV polarization backscatter	Sentinel-1
	VH	VH polarization backscatter	Sentinel-1
	HV	HV polarization backscatter	PALSAR-2
Spectral Indices	mean_MBI	Mean Modified Bare Soil Index	Sentinel-2, 3×3 mean
	NDVI	Normalized Difference Vegetation Index	Sentinel-2, $\text{NDVI} = (B8 - B4)/(B8 + B4)$
	mean_NDVI	Mean NDVI in 3×3 neighborhood	Sentinel-2, 3×3 mean
	mean_B8	Mean NIR reflectance	Sentinel-2 B8, 3×3 mean
	mean_B12	Mean SWIR reflectance	Sentinel-2 B12, 3×3 mean
	std_SEI	Standard deviation of Shadow Enhanced Index	Sentinel-2, 3×3 std dev
	TCB	Tasseled Cap Brightness	Sentinel-2
	std_TCB	Standard deviation of TCB	Sentinel-2, 3×3 std dev
	TCG	Tasseled Cap Greenness	Sentinel-2
	std_PVI	Standard deviation of PVI	Sentinel-2, 3×3 std dev
SAR Derived	VV_VH	Interaction term of VV and VH	Sentinel-1, $\text{VV} \times \text{VH}$
	HH_savg	GLCM Sum Average texture of HH	PALSAR-2, GLCM texture
	HV_savg	GLCM Sum Average texture of HV	PALSAR-2, GLCM texture
Topo-graphical	label	Dominant land cover class	Dynamic World
	DEM	Elevation	USGS 3DEP 10m
	slope	Terrain slope	Derived from DEM
	aspect	Terrain aspect	Derived from DEM

Following model selection, we used Optuna, a Bayesian optimization framework, to perform systematic hyperparameter tuning (Akiba et al., 2019). Optuna efficiently explored the hyperparameter space using a probabilistic search strategy to identify the configuration that maximized cross-validated performance while minimizing overfitting. Key parameters such as learning rate, tree depth, number of estimators, and regularization weights were tuned iteratively. Model training and evaluation were conducted on Google Colab Pro, leveraging a Tesla T4 GPU to accelerate processing and manage the computational demands of over 5.6 million GEDI footprints and their associated feature layers.

2.4.2. Model complexity and generalization

To assess model generalization and minimize the risk of overfitting, we systematically evaluated predictive performance using standard accuracy metrics (R^2 , RMSE, and MAE; defined in the Model Evaluation Metrics section) across a range of regression tree depths. The selected configuration and its performance across tree depths are reported in the Results (Table 3).

2.4.3. Model validation with ALS LiDAR

After selecting the final model, we evaluated it using grid-based spatial-block cross-validation to reduce bias from spatial autocorrelation among GEDI footprints. Validation was performed using the same statewide tiling scheme described below, withholding geographically distinct grid cells for testing while training on the remaining cells to ensure spatial separation between training and validation samples. To further assess the accuracy of the modeled canopy height product, we validated predictions against high-resolution airborne LiDAR-derived CHMs from the U.S. Geological Survey's 3D Elevation Program (3DEP). We obtained publicly available LAS files from four airborne LiDAR acquisitions across Northern and Central California conducted between 2020 and 2022: Santa Cruz County (March–April 2020), Southern Sierra (June–July 2020), Sierra Nevada (June–August 2022), and Northern California (September–October 2022). The compressed LAZ files were downloaded and processed using the lidR package in R. For each region, CHMs were generated by first normalizing the point cloud and then applying a maximum height function within 0.5-m resolution grid cells, capturing the tallest vegetation returns per pixel. These CHMs were subsequently resampled to 10-m resolution using a maximum-value aggregation method to ensure compatibility with the model-derived estimates. Airborne LiDAR validation provides an independent structural check of the modeled canopy height estimates. However, the available ALS acquisitions cover only a subset of California and do not represent the full range of forest types, structural conditions, and environmental gradients statewide. Consequently, ALS-based evaluation supports local structural realism within sampled regions but should not be interpreted as comprehensive statewide validation.

2.4.4. Model Evaluation Metrics

To evaluate the predictive performance of the ML models and assess the quality of the generated data product, three commonly used statistical metrics were employed: Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE).

R^2 : measures the proportion of variance in the observed canopy height that is explained by the model. A higher R^2 value indicates better

explanatory power.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad 1$$

where y_i is the observed canopy height value, $\hat{y}_i$ is the estimated height, $\bar{y}$ is the mean of observed height, and n is the number of observations.

RMSE: quantifies the square root of the average squared differences between observed and predicted values. It penalizes larger errors more heavily and is sensitive to outliers.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad 2$$

MAE: represents the average of the absolute differences between observed and predicted values, offering an intuitive measure of overall prediction error.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad 3$$

Together, these metrics provide a comprehensive evaluation of model performance, capturing both the goodness of fit and the magnitude of prediction errors.

2.5. Scalable statewide canopy height workflow

To generate canopy height predictions across the entire state of California, we developed a scalable and automated workflow that integrated GEE and the trained XGBoost model. Given GEE's processing limitations for large spatial domains, we divided the state into 60 km × 60 km tiles, resulting in 144 non-overlapping spatial units that collectively covered the full extent of California (Gorelick et al., 2017). For each tile, the workflow automatically ingested and preprocessed the required remote sensing and environmental variables, exported them to cloud storage, and applied the trained ML model to generate tile-level canopy height predictions at 10-m resolution.

This modular workflow was designed to run efficiently in a distributed fashion, enabling consistent model inference across tiles without manual intervention. Once predictions for all tiles were completed, the outputs were mosaicked to create seamless, wall-to-wall canopy height maps for the state. This approach supported the scalable generation of high-resolution canopy height maps across multiple years, maintaining spatial consistency, computational efficiency, and predictive accuracy throughout the process.

2.6. Per-Pixel Uncertainty Quantification

To provide spatially explicit uncertainty for each biennial canopy-height layer, we implemented a bootstrap ensemble (Efron and Tibshirani, 1994). We drew 30 bootstrap resamples of the training data and refit the XGBoost model on each, holding the tuned hyperparameters fixed. For every predicted pixel, the standard deviation of the 30 back-transformed predictions (σ_{boot}) quantifies model and sampling (epistemic) uncertainty. Because bootstrap variance captures reducible error but not systematic bias, we additionally computed a total uncertainty, $\sigma_{total} = \sqrt{(\sigma_{boot}^2 + \sigma_{resid}^2)}$, where σ_{resid} is the residual spread within the corresponding 5 m predicted-height bin, estimated from the held-out test set. A calibrated 90% prediction interval was defined as $\hat{y} \pm 1.645 \cdot c \cdot \sigma_{total}$, with a single global scalar $c = 0.945$ tuned so that empirical coverage on held-out GEDI test data matched the nominal level (achieved coverage = 0.90). The resulting per-pixel σ and prediction-interval raster's accompany each biennial layer; the product is demonstrated over the Island Fire perimeter in Fig. 13, with per-year uncertainty statistics reported in Table S1, with per-year uncertainty

Table 3

Model performance metrics across varying regression tree depths for XGBoost, showing training and test RMSE, MAE, and R^2 values.

Max Depth	RMSE		MAE		R2	
	Train	Test	Train	Test	Train	Test
6	6.30	6.40	4.55	4.59	0.69	0.68
10	5.61	6.05	4.05	4.31	0.76	0.72
15	3.11	5.23	2.18	3.55	0.93	0.79
20	0.82	4.84	0.48	2.78	0.99	0.82

maps shown in Fig. S2. This ensemble-based approach follows uncertainty estimation used in comparable large-area canopy-height products (Lang et al., 2023).

3. Results

We first examine how the selected predictors relate to GEDI canopy height. These relationships are largely nonlinear, which is consistent with our use of a nonlinear machine-learning model. Scatterplots showing the relationships between GEDI-derived canopy height (RH98) and selected remote sensing predictors, with each point color-coded by AGBD, are presented in Fig. 4. While the linear regressions capture overall directional trends, the scatterplots reveal curvilinear and saturating patterns, particularly at the lower and upper ends of the RH98 range. For example, in bands like NDVI and HV_sav, the relationship with RH98 plateaus at higher canopy heights, while variables such as B11 and B5 exhibit steep declines at low RH98 followed by leveling off. These nonlinearities, especially at structural extremes, underscore the limitations of linear models and the importance of using nonlinear ML algorithms to better capture the ecological complexity of canopy-sensor interactions.

Infrared reflectance (Bands 11 and mean_B12) and red-edge reflectance (Band 5) exhibited strong negative correlations with RH98 ($r = -0.64$, -0.63 , and -0.61 , respectively). These spectral bands are sensitive to canopy moisture, shadowing, and the visibility of lower canopy layers. In taller and more structurally complex forests, the optical signal becomes increasingly dominated by scattering and absorption in the upper canopy, limiting the sensor's ability to detect reflected light from the subcanopy or forest floor. As a result, overall reflectance in the SWIR and red-edge wavelengths decreases with increasing canopy height. This inverse relationship highlights how closed, vertically layered canopies suppress subcanopy contributions to the satellite signal, particularly in moisture-sensitive spectral regions.

Radar texture metrics, particularly HV_sav ($r = 0.57$), exhibited strong positive associations with RH98. This metric, derived from L-band PALSAR-2 HV backscatter using GLCM, captures local spatial heterogeneity in radar return intensity. Taller, more structurally complex forests tend to generate stronger volumetric scattering, especially

from woody elements such as branches and trunks. Unlike optical sensors, L-band SAR interacts with these canopy components through scattering processes, providing complementary structural information that enhances canopy height modeling.

Moderate positive correlations were observed for the Shadow-Enhanced Index (SEI) ($r = 0.44$) and the Perpendicular Vegetation Index (PVI) ($r = 0.33$), indicating that these spectral indices respond to increased canopy height. In taller forests, SEI values are elevated due to more continuous and homogeneous shadowing patterns associated with dense canopy cover. Similarly, PVI increases as closed canopies reduce soil exposure and enhance the vegetation signal in red-NIR spectral space. These patterns suggest that mature, high-biomass forests exhibit greater spectral homogeneity and vegetation dominance. In mountainous areas, terrain shading may also create cooler, moister microclimates that support taller vegetation, though these effects may vary with aspect and elevation.

Lastly, mean NDVI exhibits a saturating positive relationship with RH98 ($r = 0.61$), as expected. While NDVI generally increases with canopy greenness and vegetation density, its sensitivity declines beyond moderate canopy closure typically around 20 to 30 m of RH98, resulting in a plateau at higher canopy heights. This saturation effect reflects NDVI's limited ability to distinguish between structurally complex, high-biomass forests once a certain level of vegetation cover is reached. The AGBD gradient further illustrates this limitation: even within narrow NDVI ranges, biomass can vary substantially, especially in forests with complex vertical layering and multistory canopies.

Together, the selected predictors including spectral reflectance, vegetation indices, SAR backscatter and texture, topographic attributes, and land cover capture diverse and complementary aspects of forest structure. These variables effectively represent key drivers of canopy height, such as biomass density, vertical complexity, moisture content, and terrain context. Their strong and interpretable relationships with canopy height confirm their suitability as covariates in the ML model, providing a robust foundation for accurately estimating canopy height across ecologically and topographically heterogeneous landscapes.

Building on these predictors, we next assess the predictive performance of the trained model. The performance of the XGBoost model, trained using an optimized set of hyperparameters, was evaluated using

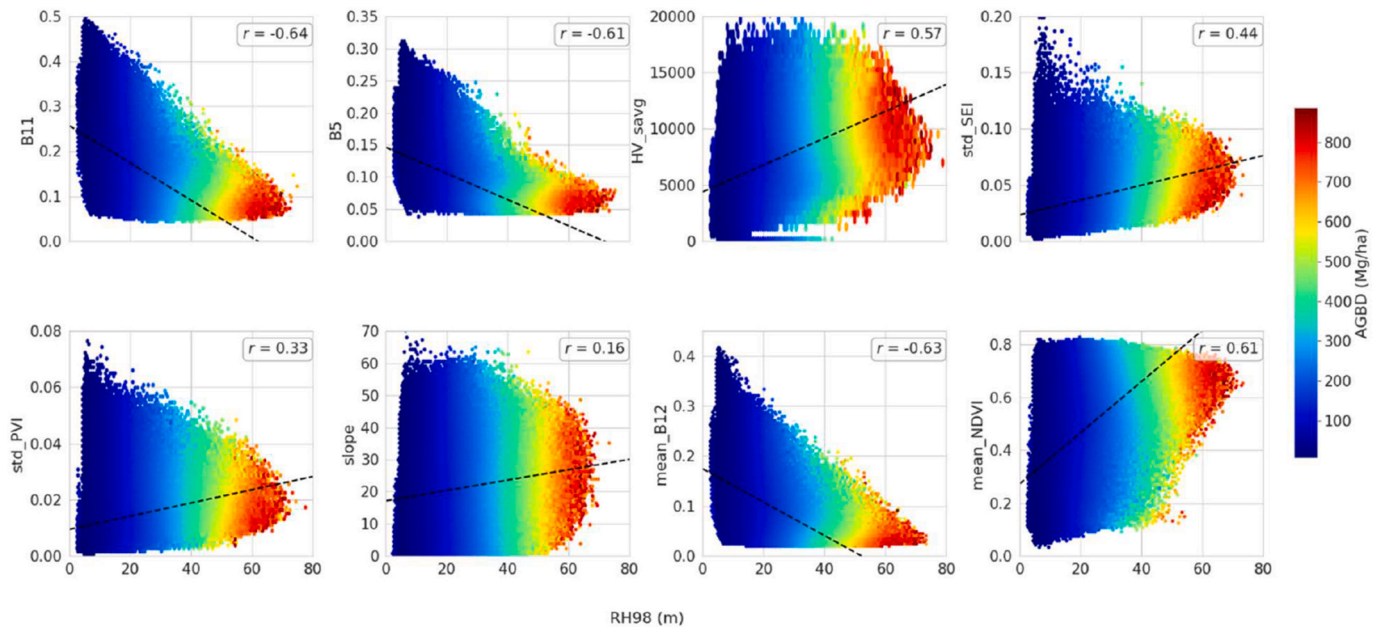


Fig. 4. Relationships between GEDI-derived canopy height (RH98, m) and selected remote sensing predictors, with points colored by mean Aboveground Biomass Density (AGBD, Mg ha⁻¹). Predictors include Sentinel-2 spectral bands (unitless reflectance, 0–1 scale), Sentinel-1 backscatter (dB), texture metrics, and topographic derivatives.

learning curves derived from a 90% training and 10% hold-out test split. The model exhibited stable convergence across 500 boosting rounds, with both training and test Log RMSE values decreasing smoothly and plateauing near epoch 400 (see Figure A2, Appendix). Importantly, the gap between training and test RMSE remained small throughout, indicating strong generalization capacity with no signs of overfitting.

As expected, deeper regression trees (e.g., max_depth = 15 or 20) achieved near-perfect fits on the training data ($R^2 = 0.925$ and 0.994 , respectively); however, their performance degraded on the test set, indicating clear signs of overfitting. In contrast, shallower models (e.g., max_depth = 6 and 10) demonstrated more stable generalization. The model with max_depth = 10 provided the most balanced performance, yielding a test RMSE of 6.05 m, MAE of 4.31 m, and R^2 of 0.72, closely aligned with the training metrics (Table 3). As a result, the max_depth = 10 model was selected for final statewide canopy height prediction.

The final XGBoost model was configured with a maximum regression tree depth of 10, 500 boosting rounds, and a learning rate of 0.1. To reduce the risk of overfitting and enhance model generalization, a subsampling rate of 0.6 was applied to the training data, and the full set of features was used per regression tree (colsample_bytree = 1.0). The model was trained using GPU acceleration, allowing for efficient and scalable processing across large datasets.

To understand which predictors drove the model's performance, we examined the gain-based feature importance of the final model. Sentinel-2 spectral bands, particularly B11 (SWIR2) and B5 (red-edge), contributed the highest predictive gain, followed by the land-cover class (label). Several spatial and statistical predictors, such as HV_savg, std_SEI, and std_PVI, further improved the model's ability to capture local textural variability. Topographic attributes (for example, slope, aspect, and DEM) and other spectral and radar variables showed moderate influence, confirming that integrating optical, SAR, and terrain metrics provides a comprehensive basis for estimating canopy height across ecologically complex regions (Fig. 5).

Annual and aggregate model performance is shown through scatterplots (Fig. 6a) comparing estimated and observed canopy height values for 2019–2022 and the combined test dataset. Across all years, the XGBoost model demonstrated strong predictive agreement, with R^2 values ranging from 0.67 to 0.75 and RMSE between 5.66 m and 6.69 m.

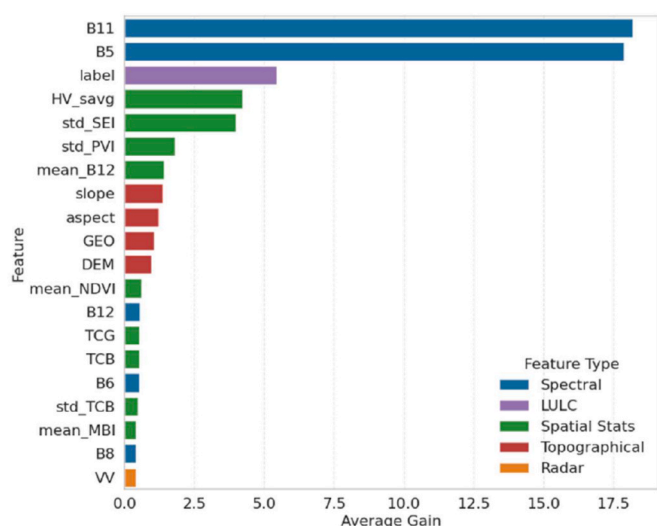


Fig. 5. Gain-based feature importance for the top predictors in the final XGBoost model. Gain (dimensionless, summing to 1 across all features) quantifies the relative improvement in model performance contributed by each feature across decision tree splits. Predictors are categorized by variable type: spectral bands (Sentinel-2), radar (Sentinel-1 and PALSAR-2), topographic, and texture/statistical.

The results reflect stable performance despite interannual variation in sample size and forest conditions, including differences in vegetation structure and phenology. In all cases, prediction densities cluster along the 1:1 line, and bias remains minimal ranging from -0.14 m to -1.37 m suggesting the model's suitability for year-to-year prediction. The combined analysis across all years yielded an R^2 of 0.73 and RMSE of 5.93 m, further validating the model's robustness and temporal transferability for statewide canopy height mapping.

The year-stratified evaluation in Fig. 6a serves as direct evidence of temporal stability across the GEDI training period: model performance varied only narrowly across 2019, 2020, 2021, and 2022 ($R^2 = 0.75, 0.74, 0.72, 0.67$; RMSE = 5.67–6.11 m; median bias between +0.8 and +1.3 m). This narrow performance band, taken across 539,611 held-out footprints, indicates that the biennial maps are produced under consistent calibration. Direct Δ Height validation at statewide scale is not feasible because publicly available repeat airborne LiDAR in California within the study window is sparse and confined to small areas (e.g., NEON sites at ~ 5 – 15 km² each). As an illustrative single-site check at one such area, we compared predicted and airborne-LiDAR Δ Height at NEON's Teakettle (TEAK) site over 2018–2024 (Supplementary Fig. S1); the agreement was reasonable (RMSE = 3.0 m, MAE = 2.2 m, bias = -0.15 m), but this ~ 1 – 2 km² single-forest-type stand is far too limited to serve as a statewide temporal validation or to support any product-level conclusion, and is reported solely as localized supporting evidence. The year-stratified evaluation therefore provides the strongest defensible evidence that the model produces consistent predictions across the years used to generate the biennial product. Natural canopy growth in undisturbed mature forest (typically 0.2–0.5 m over a two-year interval) is below the per-pixel noise floor of any of these validation methods, so the product is not designed to resolve subtle pixel-level growth but to capture structural changes whose magnitude exceeds the noise floor, most notably disturbance-driven canopy loss and post-disturbance recovery, as demonstrated in the wildfire application in Section 5 (Figs. 11 and 12).

Consistent with these findings, Fig. 6b illustrates model residuals across 5-m canopy height intervals using box plots. The model performs best for shorter to mid-range canopy structures (0–25 m), where prediction errors are symmetrically distributed around zero, and variance is minimal. However, a clear pattern of underestimation emerges in taller height bins, particularly beyond 50 m, where the mean prediction error becomes increasingly negative. The widest boxes and most negative mean errors appear in the 75–80 m bin, highlighting challenges in modeling extreme canopy heights. These limitations are likely due to sample sparsity in high-canopy areas, increased structural heterogeneity, or saturation in remote sensing signals. Together, these results confirm that while the model achieves high accuracy across most forest types, caution is warranted when interpreting predictions in old-growth or exceptionally tall forests.

A spatially explicit per-pixel uncertainty layer accompanies every biennial canopy-height map (Per-Pixel Uncertainty Quantification, Methods). For each pixel we release a bootstrap ensemble uncertainty (σ_{boot}), a height-stratified total uncertainty (σ_{total}), and a calibrated 90% prediction interval, enabling end-users to attach a confidence estimate to any pixel and to distinguish real structural change from model error. Uncertainty is lowest in short-to mid-canopy stands and increases in tall, structurally complex forest (Fig. 13), consistent with the residual behavior in Fig. 6b. The binned residual distributions by canopy height (Fig. 6b), against independent airborne LiDAR (Fig. 7b), and across EPA Level III ecoregions (Fig. 9) provide complementary references for interpreting this uncertainty.

To assess the structural accuracy of our predicted CHM, we validated its outputs against high-resolution airborne LiDAR-derived canopy heights from four regions in California. As shown in Fig. 7a, ALS-derived canopy heights demonstrated stronger agreement with our model-predicted heights ($R^2 = 0.40$, RMSE = 6.90 m, Bias = -1.74 m) compared to GEDI reference height ($R^2 = 0.32$, RMSE = 7.37 m,

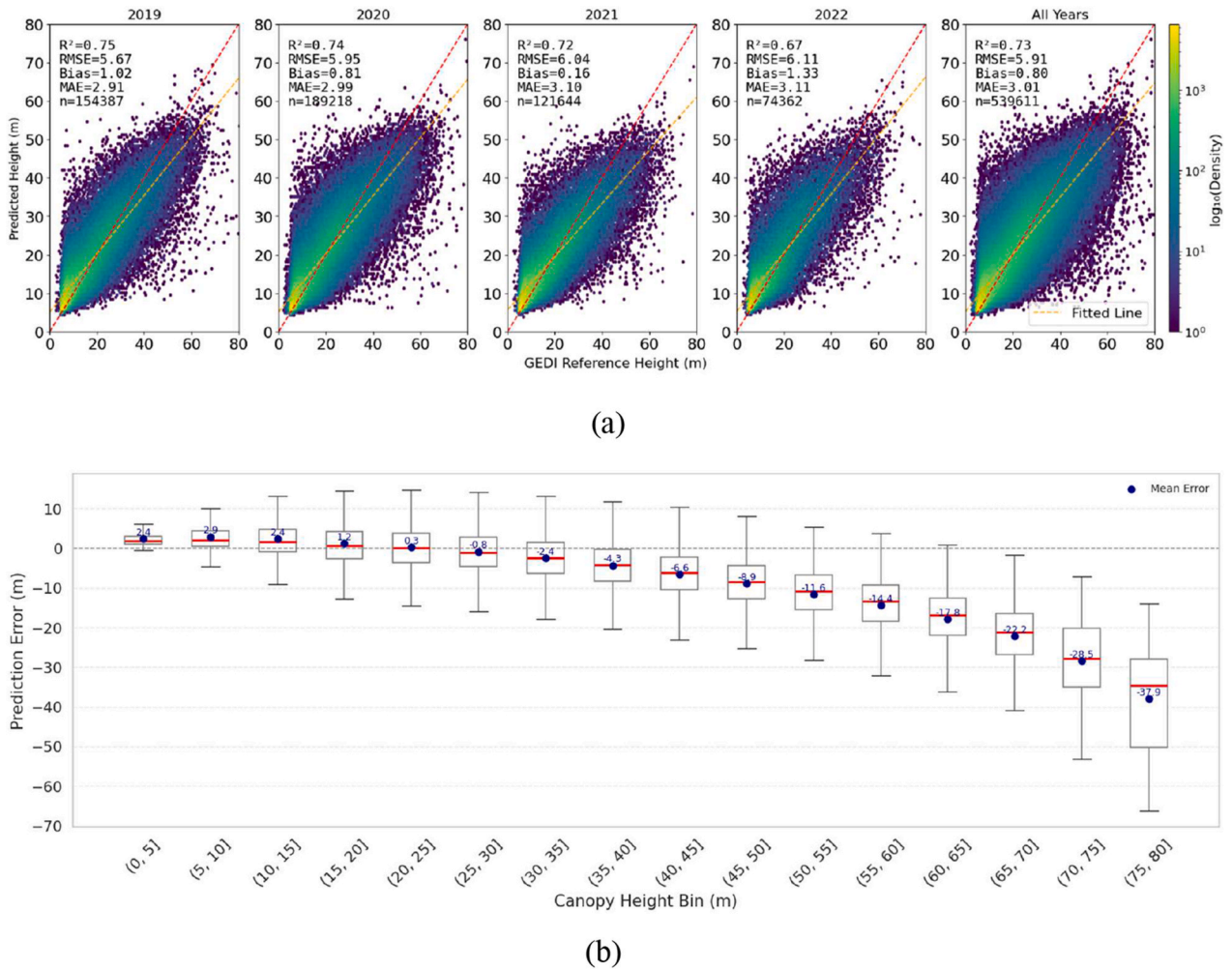


Fig. 6. Model prediction accuracy and residual behavior across years and canopy height ranges. (a) Estimated versus GEDI reference height (m) for 2019–2022 and all years combined, computed on 539,611 held-out test footprints. The red dashed line is the 1:1 line, the orange line is the least-squares fit, and the colormap shows $\log_{10}$ point density. (b) Residuals (predicted – observed, m) by 5 m RH98 bins. Boxes show the interquartile range with the median in red; whiskers span the 5th–95th percentiles; blue dots mark the mean.

Bias = -2.52 m). Although the model exhibited a slightly higher mean absolute error (MAE = 4.14 m) than GEDI (MAE = 3.81 m), the reduction in bias and improved R^2 indicate closer structural correspondence with the airborne LiDAR reference. Binned error analysis (Fig. 7b) further revealed that the model performed well for shorter canopies (0 – 25 m) and underestimation bias increased for taller forests (>30 m). These results confirm that the post-processed estimated canopy height closely agrees with airborne LiDAR-derived canopy height and improves upon GEDI reference height in capturing spatial variation. These findings support the utility of the modeled product for forest monitoring and structural analysis.

3.1. Cross-comparison with global canopy height products and ecoregion analysis

To benchmark model performance, this study compared the generated canopy height estimates with two widely used global products - Lang et al. (ETH, 2019) and Potapov et al. (UMD, 2019) - using held-out GEDI RH98 data as reference height. This evaluation ensures a consistent and independent validation framework for all three models across structurally and ecologically diverse forest conditions in California. As shown in Fig. 8a, the proposed method achieves the strongest alignment with GEDI observations ($R^2 = 0.68$, RMSE = 6.53 m, MAE = 3.84 m), substantially outperforming ETH 2019 ($R^2 = 0.48$, RMSE = 8.34 m) and

UMD 2019 ($R^2 = 0.21$, RMSE = 10.27 m). Predictions from the proposed model are more tightly clustered along the 1:1 line, particularly in mid-to high-canopy ranges, whereas both global products exhibit greater dispersion and error, especially in tall forest stands. These differences are further clarified in Fig. 8b, which presents the mean prediction error across canopy height bins using held-out GEDI reference height data. While all models perform similarly at lower heights (<25 m), errors increase significantly for taller canopies. The UMD 2019 product exhibits the most pronounced underestimation, with average errors exceeding -50 m in the 75 – 80 m bin. ETH 2019 shows a similar but less severe trend, with errors increasing notably beyond 60 m. In contrast, our proposed method maintains a smaller and more consistent bias across height ranges, indicating stronger representation of vertical canopy structure and reduced saturation effects in tall forests.

Differences in performance likely reflect differences in objectives, training data scope, and predictor availability between global and regional products, as well as California's structural extremes and terrain complexity. The UMD model is based exclusively on Landsat data, while the ETH model relies solely on Sentinel-2 imagery - both of which, while globally available, offer limited structural sensitivity and may saturate in dense or multistory canopies. In contrast, our proposed method integrates a diverse suite of predictors, including optical features from Sentinel-2, SAR metrics from Sentinel-1 and ALOS PALSAR-2, terrain derivatives (elevation, slope, aspect), land cover classifications, and a

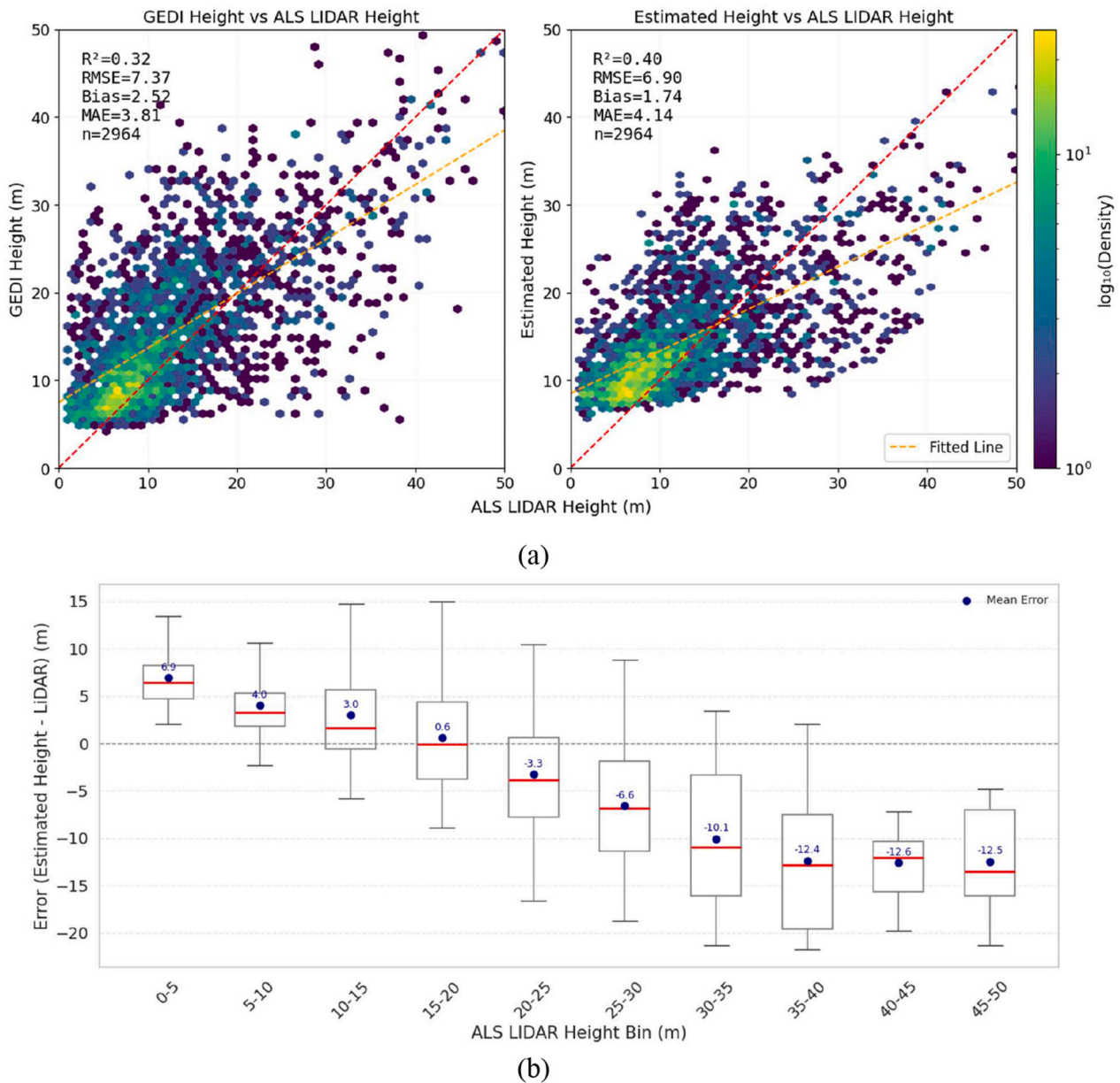


Fig. 7. Validation of canopy height estimates using Airborne Laser Scanning (ALS) LiDAR-derived canopy height models (CHM, m) at four California reference regions. (a) Scatterplots comparing ALS canopy height (m) with GEDI reference RH98 (left panel) and with our model predicted height (right panel). (b) Binned error distribution (predicted – ALS reference, m) across canopy height classes.

range of spectral indices and spatial-statistical variables. This multi-sensor, multi-feature approach enhances the model's ability to capture complex biophysical gradients and reduces prediction bias in ecologically heterogeneous landscapes. Moreover, California's rugged topography and ecological variability ranging from coastal redwoods and montane conifers to mixed hardwoods and chaparral pose particular challenges for global models trained on broad-scale, globally averaged relationships. The moderate performance of the ETH and UMD products in this region likely reflects their limited ability to generalize across such fine-scale structural and environmental variability. By contrast, our regionally tailored approach captures localized canopy dynamics more effectively, yielding improved spatial accuracy and structural realism for forest inventory and monitoring applications.

To evaluate spatial variation in model performance, we analyzed predicted canopy height residuals across Level III ecoregions in California using held-out GEDI reference height. The distribution of GEDI-observed canopy heights varied substantially across ecoregions, with

the Sierra Nevada, Klamath Mountains, and Coast Range exhibiting the tallest canopy structures, while the Central Valley, Sonoran, and Mojave regions had the shortest (Fig. 9a). The residuals (predicted - GEDI height) reveal that the model maintained relatively unbiased predictions across most ecoregions, with median residuals clustered near zero (Fig. 9b). However, slight underestimations were observed in high-canopy regions like the Sierra Nevada and Coast Range, consistent with earlier findings of increased bias in tall forests. Fig. 9c confirms the robustness of these evaluations by displaying the number of validation samples per ecoregion on a logarithmic scale, indicating strong sample representation across diverse ecological contexts. These results suggest stable behavior across many California ecoregions under the current validation design. However, GEDI sampling density is spatially imbalanced, and areas with fewer footprints likely have higher uncertainty and potentially reduced generalization.

The final output of this study is a series of high-resolution (10-m) statewide canopy height maps for California, spanning the years 2016,

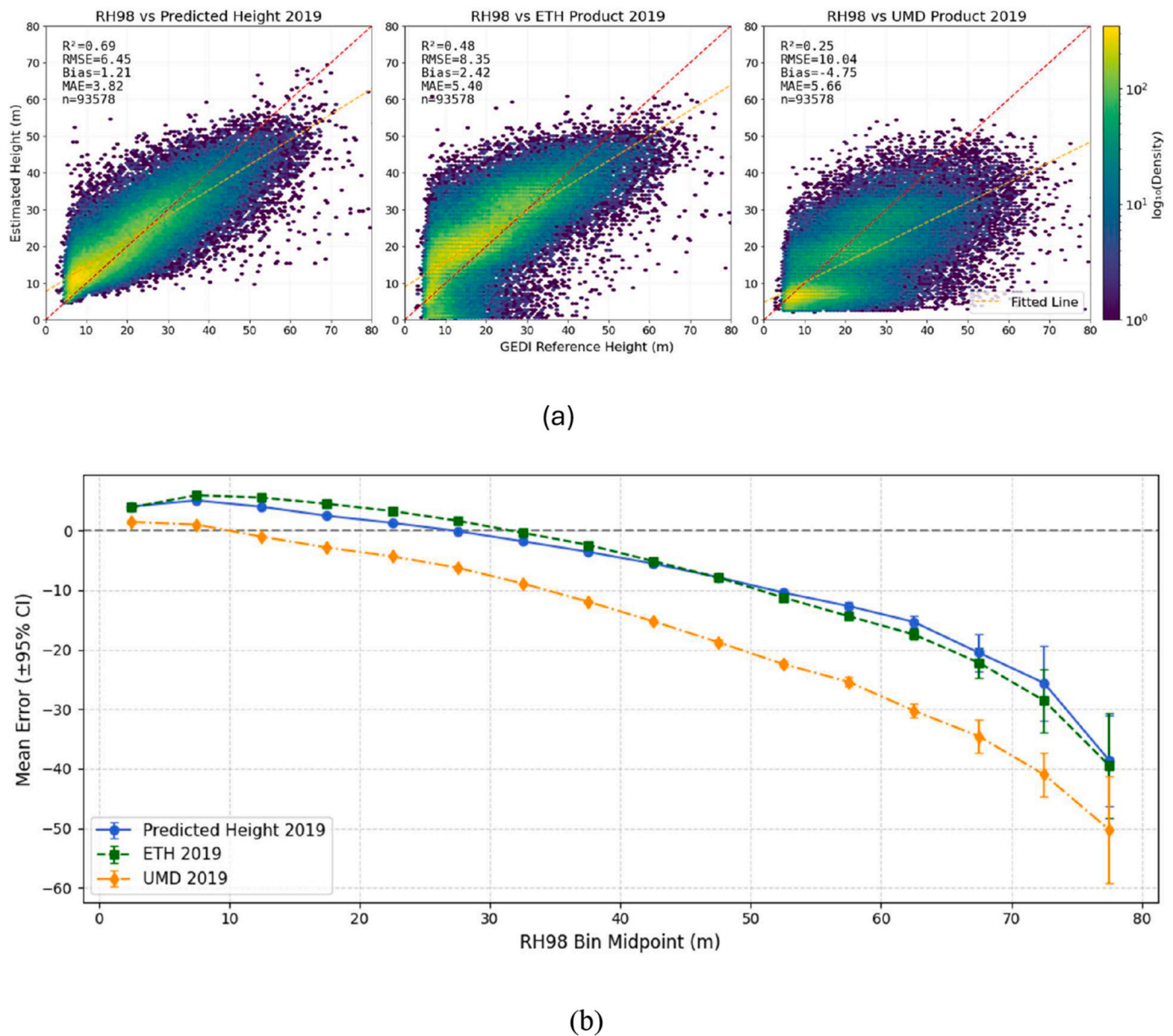


Fig. 8. Comparison of our model with two global canopy height products: [Lang et al. \(2023\)](#); ETH Zurich, denoted ETH) and [Potapov et al. \(2021\)](#); University of Maryland, denoted UMD). (a) Scatterplots of estimated versus GEDI RH98 height (m) for the three products, illustrating accuracy and bias across the canopy height range. (b) Mean prediction error (predicted – GEDI reference, m) across RH98 canopy height bins for the year 2019, computed on the same held-out test set for all three products.

2018, 2020, 2022, and 2024 ([Fig. 10](#)). These maps were generated using the trained model with optimized parameters, applied across 144 spatial tiles covering the full extent of the state. The spatial consistency and temporal resolution of these data products enable the monitoring of forest structure change over time at unprecedented detail. This statewide canopy height time series provides a valuable resource for ecological monitoring, wildfire recovery assessment, and long-term forest management planning in California's diverse and disturbance-prone landscapes.

4. Discussion

The proposed framework demonstrates strong performance and broad applicability across California's diverse forest landscapes. A key strength lies in the integration of multitemporal and multi-sensor predictor variables including Sentinel-2 optical bands, Sentinel-1 and

PALSAR-2 radar features, topographic derivatives, land cover classes, and a suite of spectral and spatial indices. This combination enables the model to capture the nonlinear, multidimensional relationships that shape forest vertical structure. The use of over 5.6 million quality-filtered GEDI footprints from multiple years further strengthens model robustness, helping to mitigate phenological variability, atmospheric interference, and seasonal inconsistencies. Model performance remains consistent across time and space, as evidenced by stable metrics (test $R^2 = 0.72$, RMSE = 6.05 m), minimal overfitting ([Figure A2](#), Appendix), and spatially unbiased residuals across Level III ecoregions ([Fig. 9](#)). Together, these features support the generalizability of the model for statewide mapping at 10-m resolution; the variable importance ranking we report provides further insight into why the multi-sensor design contributes to this performance.

A useful way to understand why our model performs differently from existing global canopy height products is to compare the variable

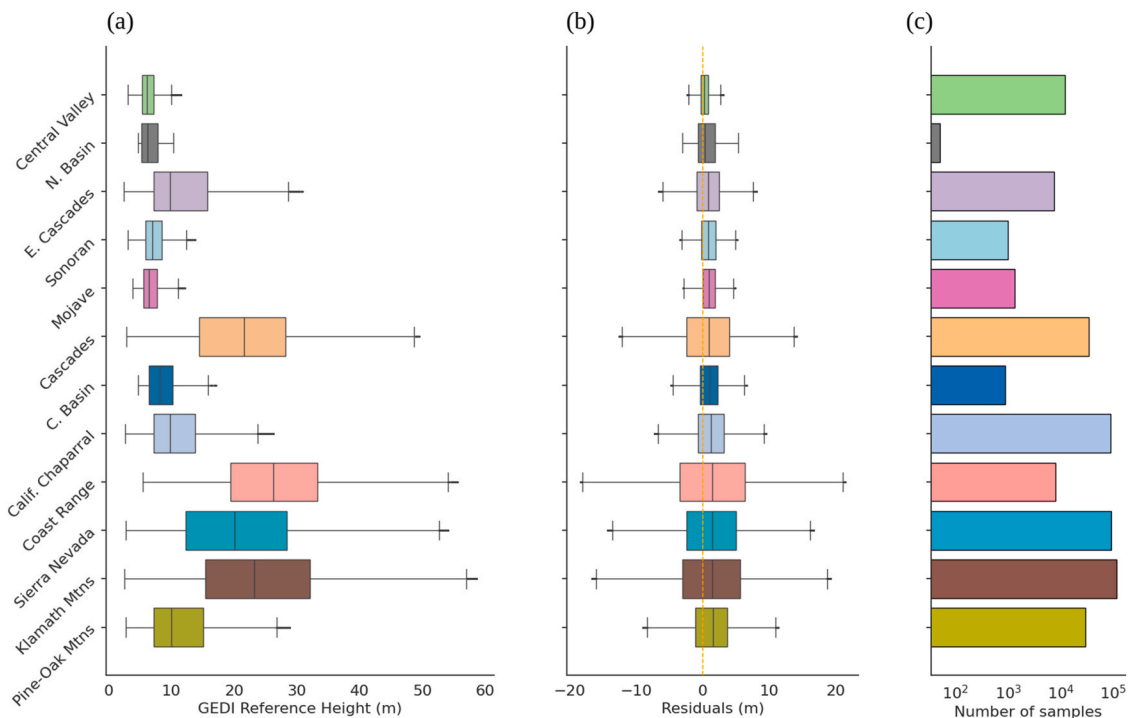


Fig. 9. Evaluation of canopy height model (CHM) performance across the 13 EPA Level III ecoregions of California. (a) Distribution of GEDI reference canopy height (m) by ecoregion. (b) Residuals between predicted and observed GEDI height (predicted - observed, m) by ecoregion. (c) Log₁₀-scaled GEDI sample count (test partition, proportional to training density) per ecoregion, showing validation density.

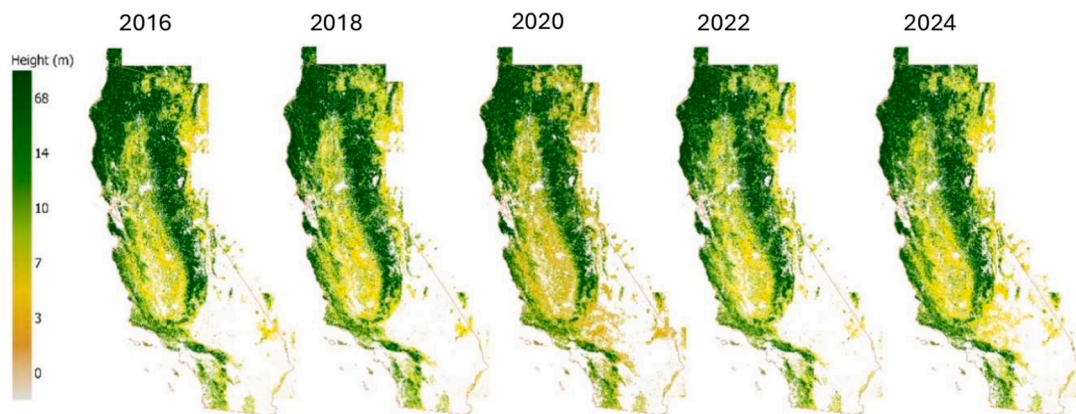


Fig. 10. Time series of 10-m resolution canopy height maps (m) for California for 2016, 2018, 2020, 2022, and 2024, generated using the proposed workflow.

importance rankings. Both Potapov et al. (2021), who used Landsat multitemporal surface reflectance with phenological indicators, and Lang et al. (2023), who used Sentinel-2 alone at 10 m, derived their global products from a single optical sensor family. Our model, by contrast, integrates four independent sensor systems: Sentinel-2 optical reflectance and indices, Sentinel-1 C-band radar backscatter and texture, ALOS PALSAR-2 L-band radar backscatter and texture, and USGS 3DEP elevation. The variable importance ranking we obtain reflects this multi-sensor design in three important ways.

First, the dominance of Sentinel-2 shortwave infrared (B11) and red-edge (B5) is consistent with the same band-family ranking reported by Lang et al. (2023) and with the SWIR-led ranking reported by Potapov et al. (2021) for Landsat. This convergence across studies confirms that shortwave and red-edge reflectance carry a transferable biophysical signal of canopy structure that is robust across spatial scales and forest types.

Second, PALSAR-2 L-band HV backscatter with spatial averaging

(HV_{savg}) ranks among the top predictors in our model. This ranking cannot emerge from single-optical-sensor pipelines by construction: optical reflectance saturates over dense or tall canopies, while L-band radar penetrates further into the canopy than the C-band Sentinel-1 used in some recent global pipelines and provides volumetric structural information that is particularly informative in California's tall-conifer forests. Including PALSAR allows the model to access a structural signal that global optical-only products cannot.

Third, statistical texture predictors computed across both optical and radar features (for example std_SEI and std_PVI) rank in our top six. Texture information at this scale captures fine-grained stand heterogeneity that helps disambiguate similar-height stands of different structural composition and is largely unresolvable in coarser-resolution global products. Land cover label also ranks higher in our model (third) than is typical in global studies, because a single-state model can exploit categorical vegetation contrasts (oak woodland, mixed conifer, chaparral, redwood) that global products must remain agnostic to in order to

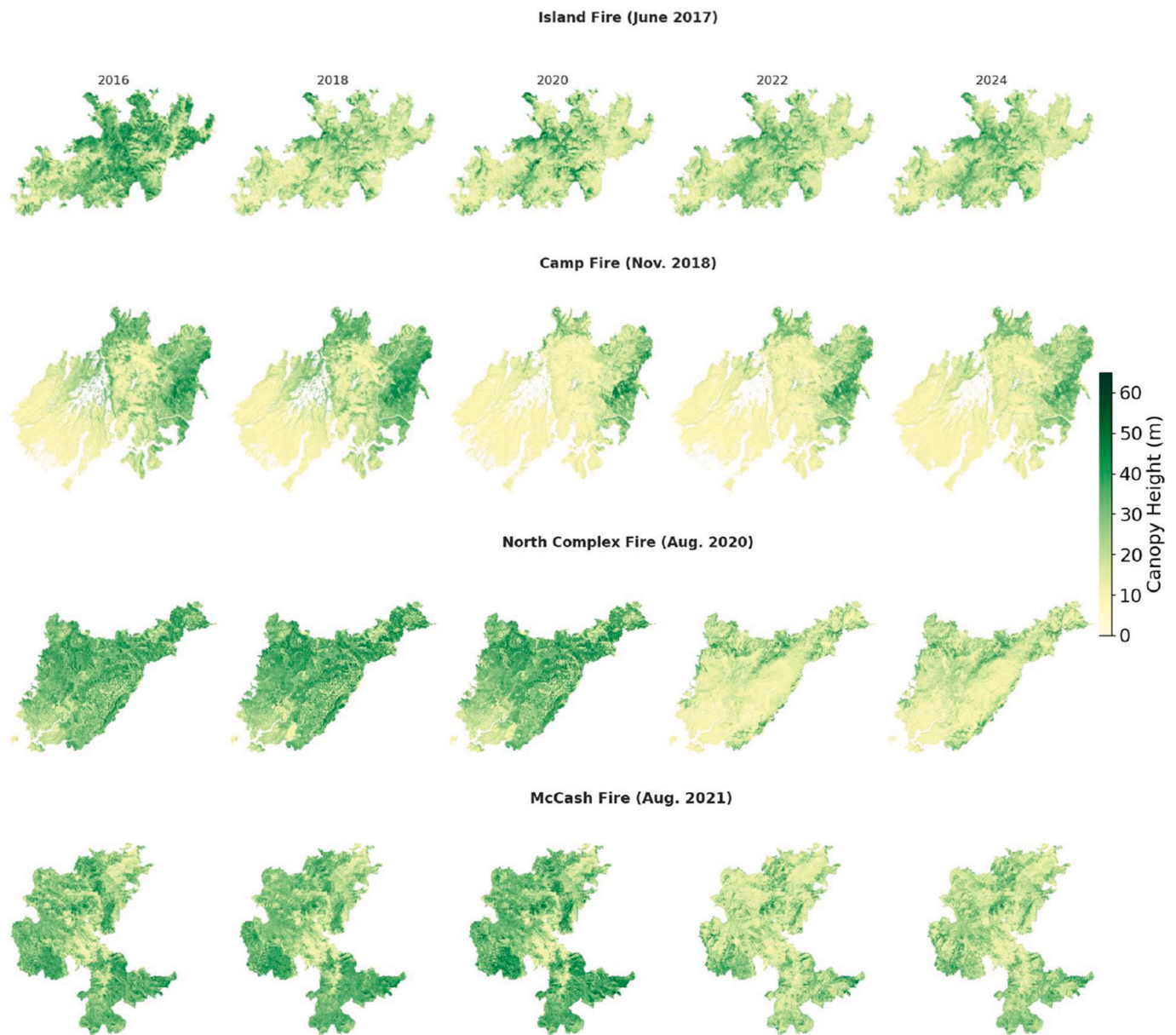


Fig. 11. Time series maps of canopy height (m, color-scaled) from 2016 to 2024 across four major wildfire footprints in California: Island Fire (2017), Camp Fire (2018), North Complex Fire (2020), and McCash Fire (2021). Each panel shows the canopy height map for one biennial epoch.

maintain transferability across biomes.

Together, these patterns illustrate the comparative advantage of a regional, multi-sensor, high-resolution approach over single-sensor global products. Multi-sensor fusion specifically enables the model to draw structural information from L-band radar that is unavailable to optical-only pipelines, and high resolution combined with regional scope enables the model to exploit fine-scale texture and land cover stratification that global products cannot.

Despite these strengths, the model - like the global canopy height products developed by Potapov et al. (2021) and Lang et al. (2023) - systematically underpredicts canopy height in the tallest forest classes, particularly above 50 m RH98. However, it still outperforms these products in terms of accuracy and bias reduction, especially in the structurally complex ecosystems of California. This persistent bias across models underscores the ongoing challenges of estimating extreme canopy structures using satellite remote sensing, particularly in dense, multistory forests such as coastal redwoods and old-growth conifer systems. This limitation is most evident in the residual plots (Figs. 6b

and 7b) and the binned prediction error curves (Fig. 8b), where model performance degrades in regions like the Coast Range and Klamath Mountains. Several factors likely contribute to this underestimation: (1) limited representation of very tall canopies in the training dataset, (2) signal saturation in optical indices such as NDVI and SWIR under dense foliage conditions, and (3) insufficient vertical sensitivity in most remotely sensed predictors. While certain variables such as L-band radar texture metrics (e.g., HV_savg) and shadow-enhanced indices (e.g., SEI) help mitigate this issue by capturing aspects of structural complexity, they are not sufficient alone to fully resolve height saturation in forests with multistory vertical layering. Quantitatively, the binned residual analysis (Fig. 8b) shows mean underestimation increasing from approximately -3 m in the 30–40 m bin to approximately -12 m in the 50–60 m bin and exceeding -19 m in the 60–70 m bin, with the largest deviations observed in the 70–80 m bin where coastal redwoods and Sierra Nevada giant sequoias reach the upper limits of canopy height. Two design choices in the modeling pipeline partially address this saturation. First, the training data were filtered using a joint

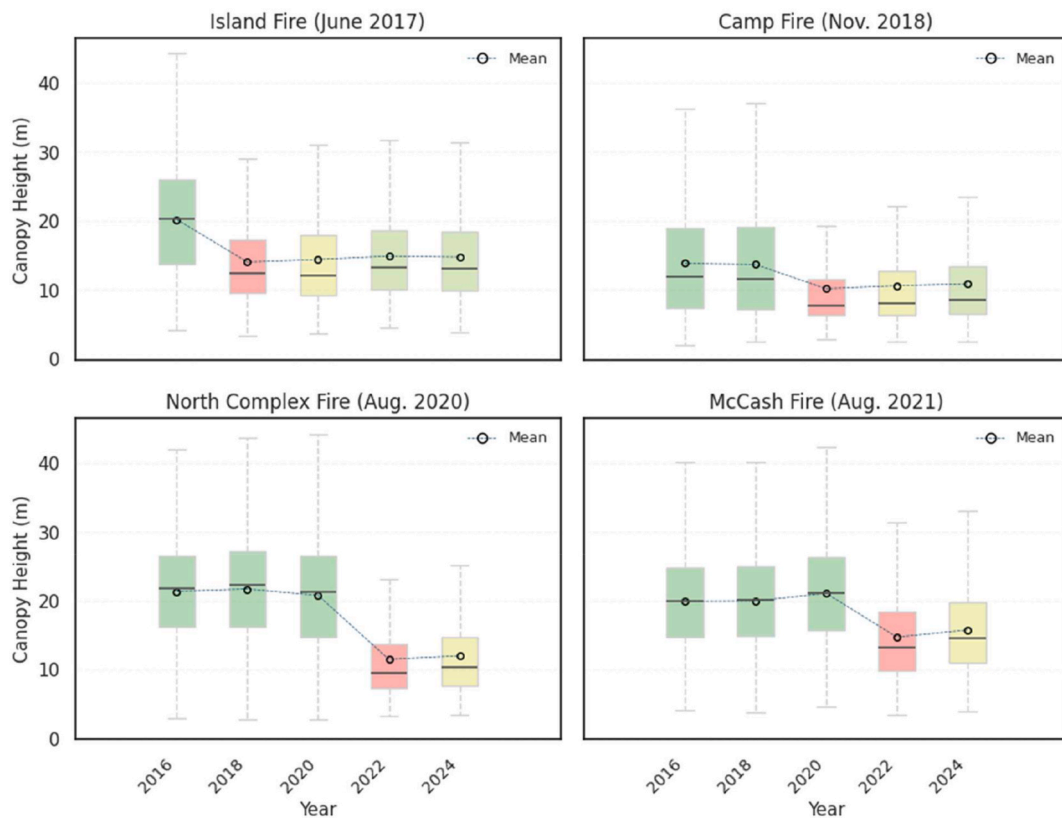


Fig. 12. Boxplots of predicted canopy height (m) by year within each burn perimeter, for the four wildfire footprints shown in Fig. 11. Green shows pre-fire years, red shows the fire year, and yellow shows the first and second post-fire years. Boxes indicate the interquartile range with medians, whiskers show the 5th–95th percentiles, and the dotted line with open circles marks the mean. The patterns highlight fire-year canopy collapse and early recovery.

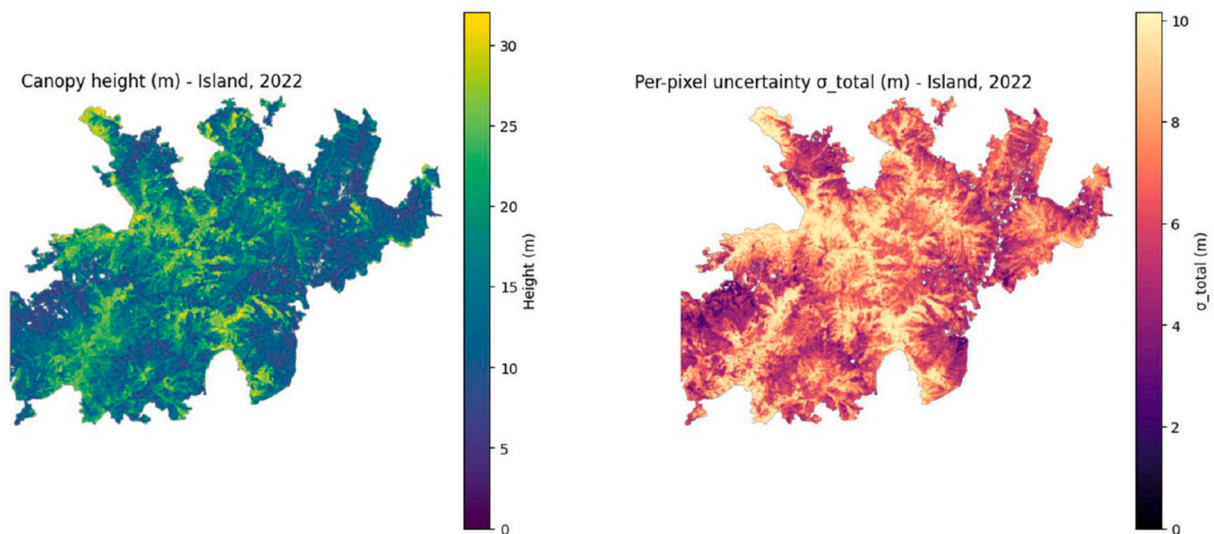


Fig. 13. Per-pixel uncertainty of the canopy-height product, demonstrated for the Island Fire perimeter (2022). (a) Predicted canopy height (m). (b) Per-pixel bootstrap uncertainty σ (m), computed as the standard deviation of a 30-member XGBoost bootstrap ensemble. Uncertainty increases in tall and structurally complex stands. Both panels share the same 10 m grid and MTBS burn-perimeter mask.

RH98–AGBD quantile envelope rather than a univariate height filter, which retains ecologically valid tall-canopy observations rather than truncating the upper tail of the training distribution (Section 3). Second, we applied a height-proportional sample weighting scheme during model fitting (weights scaled by raw RH98 normalized to maximum) so that tall-canopy footprints exert proportionally more influence during training than would be the case under uniform weighting. We tested a

more aggressive quadratic weighting (weights = $(h/h_{\max})^2$) but observed that it shifted global bias upward by ~ 1.6 m without meaningfully reducing the tall-canopy underestimation, indicating that the upper-tail bias is dominated by predictor saturation in the optical and C-band radar features rather than by under-representation in the loss function. Longer-wavelength L-band SAR (PALSAR-2) and the texture features derived from it are included in our predictor stack specifically

to extend the upper sensitivity range, and as shown in the variable importance ranking (Fig. 5) they contribute substantially to model performance; however, their effective resolution and penetration limits prevent them from fully resolving the >60 m canopy regime where most California giant sequoia and coastal redwood stands sit. Stratified modeling approaches and the integration of forthcoming NISAR L-band time-series data are promising paths to extend the upper sensitivity limit in subsequent iterations of the framework.

To characterize five-year forest structural dynamics, we analyze canopy height change across the 2016 to 2024 series within four major wildfire footprints, where the magnitude of change exceeds the per-pixel uncertainty of the product. The model's utility extends beyond static mapping, offering high-resolution, multitemporal insights into forest disturbance and recovery. Figs. 11 and 12 illustrate this application in the context of wildfire impacts across four major fire footprints: Island Fire (2017), Camp Fire (2018), North Complex Fire (2020), and McCash Fire (2021). Fig. 11 presents canopy height time series maps (2016–2024), which show abrupt structural collapse post-fire and spatial heterogeneity in recovery patterns. For example, the Camp Fire region experienced widespread canopy loss between 2018 and 2020, with limited regrowth visible by 2024. In contrast, the McCash Fire zone exhibited faster canopy rebound, particularly along its periphery highlighting variability shaped by burn severity, vegetation type, and post-fire climate.

Biennial canopy height distributions in Fig. 12 quantify the spatial patterns seen in Fig. 11 and track structural change over time. All four fire sites show substantial declines in median height and structural heterogeneity immediately after fire. The North Complex Fire site, for instance, shows a drop in median canopy height from ~20 m in 2020 to ~12 m in 2022. While modest structural recovery is observed in the McCash and North Complex fire perimeters by 2024, the Camp and Island Fire sites show little rebound, potentially due to more severe fire intensity, delayed regeneration, or persistent stressors such as drought or erosion. These results highlight the model's value in quantifying both the immediate and longer-term ecological impacts of wildfire.

In combination, the results of this study emphasize the importance of regional, data-rich, ML approaches for capturing forest structure across ecologically complex landscapes. The canopy height product provides critical insights into disturbance dynamics, recovery trajectories, and long-term vegetation change. While limitations remain in representing extremely tall canopies, the model represents a substantial improvement over existing global products and provides a scalable, interpretable framework for future expansion across CONUS.

5. Conclusion

This study presents a scalable, high-resolution canopy height mapping framework for California, leveraging GEDI LiDAR-derived canopy height data and multi-source satellite observations within an optimized XGBoost regression pipeline. Using GEE for preprocessing and GPU-accelerated training, we generated biennial 10-m resolution canopy height maps for 2016, 2018, 2020, 2022, and 2024 across California's ecologically diverse forested landscapes. The model demonstrated strong predictive performance ($R^2 = 0.72$; RMSE = 6.05 m on the test set), with limited overfitting under the current evaluation design. Shortwave infrared and red-edge spectral bands, spatial texture metrics, and topographic variables emerged as key predictors of canopy structure, highlighting the value of integrating spectral, structural, and contextual features for high-resolution canopy height estimation at statewide extent. External validation using airborne LiDAR-derived CHMs provided an independent structural check of the modeled

outputs, with strongest agreement in mid-range canopy heights.

A key contribution of this study is the temporal depth of the canopy height product and its application for disturbance-response monitoring. By examining canopy height patterns across four major wildfire footprints, we show that the biennial series captures expected directional changes associated with disturbance and early recovery. This wildfire analysis is presented as an application vignette rather than a causal attribution assessment, and inter-period differences may also reflect sensor effects and modeling uncertainty, particularly in areas with complex terrain or sparse GEDI sampling.

Several limitations should be considered. First, although our evaluation accounts for spatial dependence through spatially separated validation, performance can vary with the blocking design and the spatial distribution of training samples, and airborne-LiDAR validation is geographically limited and does not represent all statewide forest types. Second, the 2016 and 2018 epochs involve temporal extrapolation beyond GEDI's operational window (back-cast from 2019 to 2022 training), so those layers carry additional uncertainty, and the product is designed to capture disturbance-scale change rather than subtle inter-annual growth, which lies below the per-pixel noise floor. Third, the released per-pixel uncertainty layer characterizes model and sampling variance and the height-dependent residual spread, but does not fully capture systematic bias in the tallest canopies (>50 m), where predictions should be interpreted with corresponding caution (Fig. 6b).

Future work could improve predictions in highly variable or tall-canopy conditions, such as old-growth redwood forests, by incorporating additional structural information, including expanded GEDI waveform metrics and upcoming spaceborne radar observations (for example, NISAR). Expanding this framework to other regions and biomes, together with distribution-free conformal prediction to complement the bootstrap uncertainty layer introduced here, will support broader efforts to develop temporally consistent and ecologically meaningful forest height monitoring systems in an era of accelerating environmental change.

CRediT authorship contribution statement

Md Didarul Islam: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Andrew S. Johnson:** Conceptualization, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Jessica Miesel:** Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing. **Matthew Dickinson:** Funding acquisition, Investigation, Project administration, Resources, Software, Supervision, Writing – review & editing. **Jens Stevens:** Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Software, Supervision, Writing – review & editing.

Declaration of competing interest

We have read and have abided by the statement of ethical standards for manuscripts submitted to the Science of Remote Sensing. The authors have no conflicts of interest to declare.

Acknowledgments

This work was supported by the USDA Forest Service Washington Office and Northern Research Station under agreement 23-JV-11242306-069, and by McIntire-Stennis Project IDAZ-MS-0132.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.srs.2026.100488>.

Appendix:

Table A1

Initial set of 64 variables used in the analysis.

Variable	Description	Source
B2	Blue band	Sentinel-2
B3	Green band	Sentinel-2
B4	Red band	Sentinel-2
B5	Red Edge 1	Sentinel-2
B6	Red Edge 2	Sentinel-2
B7	Red Edge 3	Sentinel-2
B8	NIR	Sentinel-2
B11	SWIR1	Sentinel-2
B12	SWIR2	Sentinel-2
NDVI	Normalized Difference Vegetation Index	Sentinel-2
EVI	Enhanced Vegetation Index	Sentinel-2
NDMI	Normalized Difference Moisture Index	Sentinel-2
NDBI	Normalized Difference Built-up Index	Sentinel-2
SAVI	Soil-Adjusted Vegetation Index	Sentinel-2
TCB	Tasseled Cap Brightness	Sentinel-2
TCG	Tasseled Cap Greenness	Sentinel-2
PVI	Perpendicular Vegetation Index	Sentinel-2
ASI	Artificial Surface Index	Sentinel-2
BAEI	Built-Up Area Extraction Index	Sentinel-2
ENDISI	Enhanced Normalized Difference Impervious Surface Index	Sentinel-2
MBI	Modified Bare Soil Index	Sentinel-2
MNDWI	Modified NDWI	Sentinel-2
NBI	New Built-up Index	Sentinel-2
SEI	Shadow Enhanced Index	Sentinel-2
EBBI	Enhanced Built-up and Bareness Index	Sentinel-2
IOR	Iron Oxide Ratio	Sentinel-2
FMR	Ferrous Mineral Ratio	Sentinel-2
SFMR	Simple Ferrous Mineral Ratio	Sentinel-2
NBR	Normalized Burn Ratio	Sentinel-2
mean_NBR	Mean NBR in 3×3	Sentinel-2
std_NBR	Standard deviation of NBR	Sentinel-2
mean_NDVI	Mean NDVI in 3×3	Sentinel-2
std_NDVI	Standard deviation of NDVI	Sentinel-2
mean_TCB	Mean TCB in 3×3	Sentinel-2
std_TCB	Standard deviation of TCB	Sentinel-2
mean_TCG	Mean TCG in 3×3	Sentinel-2
std_TCG	Standard deviation of TCG	Sentinel-2
mean_PVI	Mean PVI in 3×3	Sentinel-2
std_PVI	Standard deviation of PVI	Sentinel-2
mean_SEI	Mean SEI	Sentinel-2
std_SEI	Standard deviation of SEI	Sentinel-2
mean_MBI	Mean MBI	Sentinel-2
mean_EBBI	Mean EBBI	Sentinel-2
mean_IOR	Mean IOR	Sentinel-2
mean_NIR	Mean NIR	Sentinel-2
mean_SWIR	Mean SWIR	Sentinel-2
std_NIR	Standard deviation of NIR	Sentinel-2
std_SWIR	Standard deviation of SWIR	Sentinel-2
VV	Vertical transmit and receive (VV)	Sentinel-1
VH	Vertical transmit, horizontal receive (VH)	Sentinel-1
VV_VH	VV * VH interaction term	Sentinel-1
VV_contrast	GLCM contrast (VV)	Sentinel-1
VH_contrast	GLCM contrast (VH)	Sentinel-1
HH	HH polarization backscatter	PALSAR-2
HV	HV polarization backscatter	PALSAR-2
HV_HH	Sum of HV and HH	PALSAR-2
HH_contrast	GLCM contrast (HH)	PALSAR-2
HH_diss	GLCM dissimilarity (HH)	PALSAR-2
HH_savg	GLCM sum average (HH)	PALSAR-2
HV_contrast	GLCM contrast (HV)	PALSAR-2
HV_diss	GLCM dissimilarity (HV)	PALSAR-2
HV_savg	GLCM sum average (HV)	PALSAR-2
DEM	Digital Elevation Model (10 m)	USGS 3DEP
label	Dynamic World Land Cover (mode)	Dynamic World

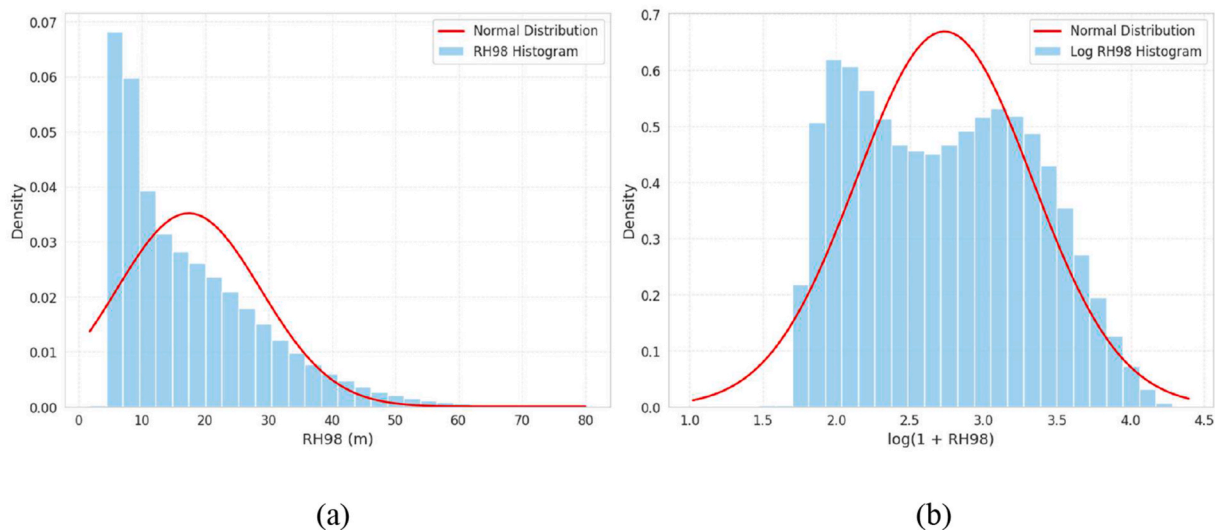


Fig. A1. Distributions of GEDI relative height at 98th percentile (RH98) used for model training. (a) Raw RH98 values (m) exhibit strong right skew, with most observations below 25 m. (b) Natural-log-transformed RH98 values show reduced skewness and a more symmetric distribution suitable for regression modeling.

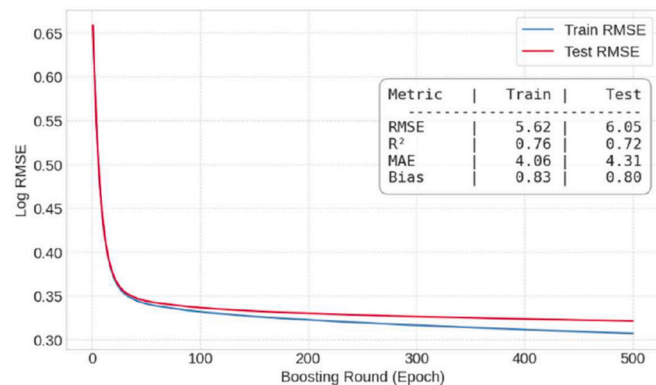


Fig. A2. Learning curves showing the natural logarithm of Root Mean Square Error (Log RMSE, dimensionless after log transform) for training and held-out test sets across 500 boosting rounds of the XGBoost model. The inset table summarizes final performance metrics (RMSE in m, R^2 dimensionless), indicating strong model generalization.

Data availability

Data will be made available on request.

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2623–2631. <https://doi.org/10.1145/3292500.3330701>.

Brown, C.F., Brumby, S.P., Guzder-Williams, B., Birch, T., Hyde, S.B., Mazzariello, J., Czerwinski, W., Pasquarella, V.J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., Tait, A.M., 2022. Dynamic world, near real-time global 10 m land use land cover mapping. *Sci. Data* 9 (1), 251. <https://doi.org/10.1038/s41597-022-01307-4>.

California Department of Forestry and Fire Protection (CAL FIRE), 2024. Top 20 Largest California Wildfires. <https://www.fire.ca.gov/-/media/calfire-website/our-impact/fire-statistics/top-20-largest-ca-wildfires.pdf>. (Accessed 6 August 2026) accessed.

Cartus, O., Santoro, M., Kelndorfer, J., 2012. Mapping forest aboveground biomass in the northeastern United States with ALOS PALSAR dual-polarization L-band. *Remote Sens. Environ.* 124, 466–478. <https://doi.org/10.1016/j.rse.2012.05.029>.

Duan, Naihua, 1983. Smearing estimate: a nonparametric retransformation method. *J. Am. Stat. Assoc.* 78 (383), 605–610. <https://doi.org/10.1080/01621459.1983.10478017>.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing* 1, 100002. <https://doi.org/10.1016/j.srs.2020.100002>.

Duncanson, L., Kellner, J., Armston, J., Dubayah, R., Minor, D.M., Hancock, S., Healey, S. P., Patterson, P.L., Saarela, S., Marselis, S., Silva, C.E., Bruening, J., Goetz, S.J., Tang, H., Hofton, M., Blair, B., Luthcke, S., Fatoyinbo, L., Abernethy, K., Alonso, A., Andersen, H.-E., Aplin, P., Baker, T.R., Barbier, N., Bastin, J.F., Biber, P., Boeckx, P., Bogaert, J., Boschetti, L., Boucher, P.B., Boyd, D.S., Burslem, D.F.R.P., Calvo-Rodriguez, S., Chave, J., Chazdon, R.L., Clark, D.B., Clark, D.A., Cohen, W.B., Coomes, D.A., Corona, P., Cushman, K.C., Cutler, M.E.J., Dalling, J.W., Dalponte, M., Dash, J., de-Miguel, S., Deng, S., Ellis, P.W., Erasmus, B., Fekety, P.A., Fernandez-Landa, A., Ferraz, A., Fischer, R., Fisher, A.G., García-Abril, A., Gobakken, T., Hacker, J.M., Heinrich, M., Hill, R.A., Hopkinson, C., Huang, H., Hubbell, S.P., Hudak, A.T., Huth, A., Imbach, B., Jeffery, K.J., Katoh, M., Kearsley, E., Kenfack, D., Kljun, N., Knapp, N., Král, K., Krůček, M., Labrière, N., Lewis, S.L., Longo, M., Lucas, R.M., Main, R., Manzanares, J.A., Martínez, R.V., Mathieu, R., Memiaghe, H., Meyer, V., Mendoza, A.M., Moneris, A., Montesano, P., Morsdorf, F., Næsset, E., Naidoo, L., Nilus, R., O'Brien, M., Orwig, D.A., Papathanassiou, K., Parker, G., Phillips, C., Phillips, O.L., Pisek, J., Poulsen, J.R., Pretzsch, H., Rüdiger, C., Saatchi, S., Sanchez-Azofeifa, A., Sanchez-Lopez, N., Scholes, R., Silva, C.A., Simard, M., Skidmore, A., Stereńczak, K., Tanase, M., Torresan, C., Valbuena, R., Verbeeck, H., Vrska, T., Wessels, K., White, J.C., White, L.J.T., Zahabu, E., Zraggen, C., 2022. Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission. *Remote Sens. Environ.* 270, 112845. <https://doi.org/10.1016/j.rse.2021.112845>, 112845. <https://linkinghub.elsevier.com/retrieve/pii/S0034425721005654>.

Efron, B., Tibshirani, R.J., 1994. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, ISBN 9780429246593. <https://doi.org/10.1201/9780429246593>. <https://www.taylorfrancis.com/books/9781000064988>. (Accessed 15 May 1994).

Fang, G., Xu, H., Yang, S.-I., Lou, X., Fang, L., 2023. Synergistic use of Sentinel-1, Sentinel-2, and landsat 8 in predicting forest variables. *Ecol. Indic.* 151, 110296. <https://doi.org/10.1016/j.ecolind.2023.110296>.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.

- Griffith, G.E., Omernik, J.M., Smith, D.W., Cook, T.D., Tallyn, E., Moseley, K., Johnson, C.B., 2016. Ecoregions of California. In: *Open-File Report* (Nos. 2016–1021). U.S. Geological Survey. <https://doi.org/10.3133/ofr20161021>.
- Hancock, S., Armston, J., Hofton, M., Sun, X., Tang, H., Duncanson, L.I., Kellner, J.R., Dubayah, R., 2019. The GEDI simulator: a large-footprint waveform lidar simulator for calibration and validation of spaceborne missions. *Earth Space Sci.* 6 (2), 294–310. <https://doi.org/10.1029/2018EA000506>.
- Hayashi, M., Saigusa, N., Oguma, H., Yamagata, Y., 2013. Forest canopy height estimation using ICESat/GLAS data and error factor analysis in Hokkaido, Japan. *ISPRS J. Photogrammetry Remote Sens.* 81, 12–18. <https://doi.org/10.1016/j.isprsjprs.2013.04.004>.
- Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., Gass, L., Funk, M., Wickham, J., Stehman, S., Auch, R., Riitters, K., 2020. Conterminous United States land cover change patterns 2001–2016 from the 2016 national land cover database. *ISPRS J. Photogrammetry Remote Sens.* 162, 184–199. <https://doi.org/10.1016/j.isprsjprs.2020.02.019>.
- Islam, Md. D., Di, L., Zhang, C., Yang, R., Qu, J.J., Tong, D., Guo, L., Lin, L., Pandey, A., 2024. A decision rule and machine learning-based hybrid approach for automated land-cover type local climate zones (LCZs) mapping using multi-source remote sensing data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 17, 8271–8290. <https://doi.org/10.1109/jstars.2024.3386389>. <https://ieeexplore.ieee.org/document/10494866/>.
- Islam, Md. D., Sultana, M., Aboulola, O.I., 2026. Blue-carbon monitoring in the Sundarbans: 10 m canopy height and biomass from GEDI LIDAR and multi-sensor fusion, 2016–2024. *Remote Sens. Appl.: Soc. Environ.* 42, 102099. <https://doi.org/10.1016/j.rsase.2026.102099>. <https://linkinghub.elsevier.com/retrieve/pii/S2352938526002326>.
- Islam, Md. Didarul, Di, L., Mia, Md. R., Sithi, M.S., 2022. Deforestation mapping of sundarbans using multi-temporal sentinel-2 data & transfer learning. 2022 10th International Conference on Agro-geoinformatics (Agro-GeoInformatics) 1–4. <https://doi.org/10.1109/agro-geoinformatics55649.2022.9858968>, IEEE. <https://ieeexplore.ieee.org/document/9858968/>.
- Lang, N., Jetz, W., Schindler, K., Wegner, J.D., 2023. A high-resolution canopy height model of the Earth. *Nat. Ecol. Evol.* 7 (11), 1778–1789. <https://doi.org/10.1038/s41559-023-02206-6>.
- Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Del Bom Espirito-Santo, F., Hunter, M.O., de Oliveira Jr., R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys. Res. Lett.* 32 (22). <https://doi.org/10.1029/2005GL023971>.
- Li, X., Wessels, K., Armston, J., Hancock, S., Mathieu, R., Main, R., Naidoo, L., Erasmus, B., Scholes, R., 2023. First validation of GEDI canopy heights in African savannas. *Remote Sens. Environ.* 285, 113402. <https://doi.org/10.1016/j.rse.2022.113402>.
- Long, J.L., Hatten, T.D., 2023. LANDFIRE. In: *Fact Sheet* (Nos. 2023–3044). U.S. Geological Survey. <https://doi.org/10.3133/fs20233044>.
- Miller, J.D., Quayle, B., 2015. Calibration and validation of immediate post-fire satellite-derived data to three severity metrics. *Fire Ecol.* 11 (2), 12–30. <https://doi.org/10.4996/fireecology.1102012>. <https://fireecology.springeropen.com/articles/10.4996/fireecology.1102012>. (Accessed 1 August 2015).
- Miller, J.D., Thode, A.E., 2007. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sens. Environ.* 109 (1), 66–80. <https://doi.org/10.1016/j.rse.2006.12.006>. <https://linkinghub.elsevier.com/retrieve/pii/S0034425706005128>.
- Moudry, V., Prošek, J., Marselis, S., Marešová, J., Šárová, E., Gdulová, K., Kozhoridze, G., Torresani, M., Rocchini, D., Eltner, A., Liu, X., Potůčková, M., Sedová, A., Crespo-Peremarch, P., Torralba, J., Ruiz, L.A., Perrone, M., Špatenková, O., Wild, J., 2024. How to Find Accurate Terrain and Canopy Height GEDI Footprints in Temperate Forests and Grasslands? *Earth Space Sci.* 11 (10), e2024EA003709. <https://doi.org/10.1029/2024ea003709>. <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024EA003709>. (Accessed 25 October 2024).
- Nandy, S., Srinet, R., Padalia, H., 2021. Mapping forest height and aboveground biomass by integrating ICESat-2, Sentinel-1 and Sentinel-2 data using random forest algorithm in northwest himalayan foothills of India. *Geophys. Res. Lett.* 48 (14). <https://doi.org/10.1029/2021GL093799>. e2021GL093799.
- North, M.P., Stephens, S.L., Collins, B.M., Agee, J.K., Aplet, G., Franklin, J.F., Fulé, P.Z., 2015. Reform forest fire management. *Science* 349 (6254), 1280–1281. <https://doi.org/10.1126/science.aab2356>.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on Earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *Bioscience* 51 (11), 933–938. [https://doi.org/10.1641/0006-3568\(2001\)051%255B0933:TEOTWA%25D2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051%255B0933:TEOTWA%25D2.0.CO;2).
- Parks, S.A., Miller, C., Abatzoglou, J.T., Holsinger, L.M., Parisien, M.-A., Dobrowski, S.Z., 2016. How will climate change affect wildland fire severity in the Western US? *Environ. Res. Lett.* 11 (3), 035002. <https://doi.org/10.1088/1748-9326/11/3/035002>.
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J. B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and landsat data. *Remote Sens. Environ.* 253, 112165. <https://doi.org/10.1016/j.rse.2020.112165>.
- Roy, D.P., Kashongwe, H.B., Armston, J., 2021. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Sci. Remote Sens.* 4, 100024. <https://doi.org/10.1016/j.srs.2021.100024>, 100024. <https://linkinghub.elsevier.com/retrieve/pii/S2666017221000110>.
- Saatchi, S., Favrichon, S., 2024. *Global Ecosystem Dynamics Investigation (GEDI) Global Vegetation Height Metrics from GEDI and ICESat2* (Version 1, P. 0 MB) [COG]. ORNL Distributed Active Archive Center. <https://doi.org/10.3334/ORNLDAA/2294>.
- Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sens. Environ.* 155, 13–31. <https://doi.org/10.1016/j.rse.2014.04.014>.
- Shu, M., Li, Q., Ghafoor, A., Zhu, J., Li, B., Ma, Y., 2023. Using the plant height and canopy coverage to estimation maize aboveground biomass with UAV digital images. *Eur. J. Agron.* 151, 126957. <https://doi.org/10.1016/j.eja.2023.126957>.
- Stephens, S.L., Westerling, A.L., Hurteau, M.D., Peery, M.Z., Schultz, C.A., Thompson, S., 2020a. Fire and climate change: conserving seasonally dry forests is still possible. *Front. Ecol. Environ.* 18 (6), 354–360. <https://doi.org/10.1002/fee.2218>.
- Stephens, S.L., Westerling, A.L., Hurteau, M.D., Peery, M.Z., Schultz, C.A., Thompson, S., 2020b. Fire and climate change: conserving seasonally dry forests is still possible. *Front. Ecol. Environ.* 18 (6), 354–360. <https://doi.org/10.1002/fee.2218>.
- Tang, H., Stoker, J., Luthcke, S., Armston, J., Lee, K., Blair, B., Hofton, M., 2023. Evaluating and mitigating the impact of systematic geolocation error on canopy height measurement performance of GEDI. *Remote Sens. Environ.* 291, 113571. <https://doi.org/10.1016/j.rse.2023.113571>, 113571. <https://linkinghub.elsevier.com/retrieve/pii/S0034425723001220>.
- Tolan, J., Yang, H.-I., Nosarzewski, B., Couairon, G., Vo, H.V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., Vamaraju, J., Moutakanni, T., Bojanowski, P., Johns, T., White, B., Tiecke, T., Couprie, C., 2024. Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. *Remote Sens. Environ.* 300, 113888. <https://doi.org/10.1016/j.rse.2023.113888>.
- Wang, X., Huang, H., Gong, P., Liu, C., Li, C., Li, W., 2014. Forest canopy height extraction in rugged areas with ICESat/GLAS data. *IEEE Trans. Geosci. Rem. Sens.* 52 (8), 4650–4657. <https://doi.org/10.1109/TGRS.2013.2283272>.
- Westerling, A.L., Bryant, B.P., 2008. Climate change and wildfire in California. *Clim. Change* 87 (1), 231–249. <https://doi.org/10.1007/s10584-007-9363-z>.
- Westerling, A.L., Bryant, B.P., Preisler, H.K., Holmes, T.P., Hidalgo, H.G., Das, T., Shrestha, S.R., 2011. Climate change and growth scenarios for California wildfire. *Clim. Change* 109 (1), 445–463. <https://doi.org/10.1007/s10584-011-0329-9>.
- Wong, S.D., Broader, J.C., Shaheen, S.A., 2020. Review of California wildfire evacuations from 2017 to 2019. <https://doi.org/10.7922/G29G5K2R>.
- Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bader, C.W., Gobakken, T., 2012. Lidar sampling for large-area forest characterization: a review. *Remote Sens. Environ.* 121, 196–209. <https://doi.org/10.1016/j.rse.2012.02.001>.
- Xiao, Y., Wang, Q., Xi, H., Tong, X., 2025. A 30 m canopy height map in China created by fusion of multiple relative height metrics. *IEEE Trans. Geosci. Rem. Sens.* 63, 1–16. <https://doi.org/10.1109/TGRS.2025.3572524>.
- Zang, J., Qiu, F., Zhang, Y., 2024. A global dataset of forest regrowth following wildfires. *Sci. Data* 11 (1), 1052. <https://doi.org/10.1038/s41597-024-03896-8>.
- Zhang, G., Ganguly, S., Nemani, R.R., White, M.A., Milesi, C., Hashimoto, H., Wang, W., Saatchi, S., Yu, Y., Myneni, R.B., 2014. Estimation of forest aboveground biomass in California using canopy height and leaf area index estimated from satellite data. *Remote Sens. Environ.* 151, 44–56. <https://doi.org/10.1016/j.rse.2014.01.025>.