Integrating Sparse LiDAR and Multisensor Time-Series Imagery From Spaceborne Platforms for Deriving Localized Canopy Height Model

Arnav Goel, Hunsoo Song, and Jinha Jung, *Senior Member, IEEE*

***Abstract*— Canopy height is a fundamental metric for extracting valuable information about forested areas. Over the past decade, the light detection and ranging (LiDAR) technology has provided a straightforward method for measuring canopy height using various platforms, including terrestrial, uncrewed aerial vehicles (UAVs), airborne, and satellite sensors. However, despite its global reach, spaceborne LiDAR data suffers from a sparse sampling pattern that fails to provide continuous global coverage. In contrast, satellites like LANDSAT deliver seamless and extensive coverage of the Earth's surface through spectral data. This study aims to develop a deep learning model to infer canopy heights from sparsely observed LiDAR data, utilizing the multisensor spectral data from spaceborne platforms. Specifically tailored for localized sites, the model focuses on regional-level canopy height estimation by leveraging the relationship between canopy height and multisensor time-series data from Landsat, Sentinel-2, and Sentinel-1. We first demonstrate the importance of integrating multisensor data by training three separate models: one using only Landsat data, one using only Sentinel-2 data, and a multimodal model that incorporates Landsat, Sentinel 1, and Sentinel 2 data to estimate LiDAR-derived canopy height. These models were tested on two sites in Indiana—Tippecanoe and Monroe counties—where the multimodal approach produced the best results, achieving RMSEs of 3.895 and 4.993 m, respectively. We then tested our multimodal model in two additional counties—Baker County, FL, USA and Piute County, UT, USA—where the model achieved an RMSE of 5.397 and 3.742 m, respectively.**

***Index Terms*— Canopy height estimation, localized model, multisensor satellite data.**

I. Introduction

FORESTS constitute one of the largest terrestrial ecosystems on our planet. They serve as one of the main pillars supporting the lives of humans and other animals [1]. As a carbon sink, forests act to sequester a significant amount of greenhouse gases from the Earth's atmosphere. They also play a crucial role in providing oxygen that sustains life while removing carbon dioxide from the atmosphere—a primary contributor to global warming, which endangers the lives of all living beings.

With the advent of human civilization and urban areas, forests have consistently fallen victim to depletion for timber use and land area. Consequently, this has resulted in a significant increase in greenhouse gases over the past 100 years, contributing to global warming [2]. Naturally, the maintenance and preservation of forest areas have become key sustainable development goals as declared by the United Nations [3].

Therefore, studying forest structures remains an essential research topic for monitoring changes in forest areas and estimating forest biomass. One key metric in understanding forest structure is canopy height, which denotes the height of trees in forested regions [4]. Temporally studying canopy height enables quantification of deforestation, tree degradation, and the effectiveness of any restoration policies implemented in the region.

Various methods are used to measure canopy height, each with its own advantages and disadvantages. The main tradeoff is between the accuracy and precision of the measurements and the scale of data collection. These methods include in situ field measurement, uncrewed aerial vehicle (UAV)-assisted light detection and ranging (LiDAR), airborne LiDAR, and satellite-based LiDAR data [5]. In situ field measurements are the most accurate but also the most labor intensive and time-consuming. Conversely, LiDAR measurements offer a less laborious remote sensing alternative, utilizing laser beams to estimate the heights of the targets they strike. Recently, UAV-based LiDAR systems have become increasingly popular for small-scale, forest-level data collection due to their high accuracy and reduced labor intensity compared to field measurements [6]. In addition, airborne LiDAR is effective for medium-scale data collection across counties or states, offering broader coverage than UAV LiDAR, albeit with a sparser point density.

Lately, satellite-based LiDAR systems have emerged in the past decade, enabling the measurement of canopy heights on a global scale at minimal cost. Notable examples include NASA's GLAS (ICESat), ATLAS (ICESat-2), and more recently, NASA's Global Ecosystem Dynamics Investigation (GEDI)—a full-waveform LiDAR instrument installed on the International Space Station (ISS). The main drawback of this approach is the limited accuracy and sparse sampling

Received 11 June 2024; revised 9 October 2024 and 16 December 2024; accepted 11 January 2025. Date of publication 17 February 2025; date of current version 7 March 2025. This work was supported by USDA NIFA grant #2023-68012-38992 and NRCS grant #NR233A750004G044. The findings and conclusions should not be construed to represent any agency determination or policy. *(Corresponding author: Jinha Jung.)*

Arnav Goel and Jinha Jung are with the Geospatial Data Science Laboratory and the Lyles School of Civil and Construction Engineering, Purdue University, West Lafayette, IN 47906 USA (e-mail: goel65@purdue.edu; jinha@purdue.edu).

Hunsoo Song is with the Geospatial Data Science Laboratory and the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47906 USA, and also with the Yale School of the Environment, Yale University, New Haven, CT 06511 USA (e-mail: hunsoo@purdue.edu).

Digital Object Identifier 10.1109/TGRS.2025.3542685

of LiDAR measurements. GEDI's waveform footprint has a reported horizontal accuracy of 25 m and has covered only 4% of the Earth's surface. Its eight observation tracks are 600 m apart, taking samples every 60 m [7]. Another satellite-based LiDAR mission, ICESAT-2, utilizes a photon-counting laser ranging system, unlike GEDI's full waveform LiDAR. The advantage of the photon-counting laser system is its 70-cm separation between each laser pulse along the track [8]. The primary goal of ICESAT-2 is to measure changes in the cryosphere, which informs its orbital design to maximize coverage over the poles and boreal forests at higher latitudes. Consequently, GEDI, with its specific design to estimate forest aboveground biomass and an orbital path optimized for these regions, is a more suitable source for LiDAR observations at the test sites that we shall discuss later in this article [7].

To address the limitations of LiDAR data, such as its scarcity and low sampling frequency, researchers have explored alternative satellite observations to infer canopy heights. An early study by Pascual et al. [9] effectively demonstrated this by linking canopy height from airborne LiDAR with spectral indices from Landsat ETM+ tiles. Similarly, LaRue et al. [11] found that NDVI strongly correlated with structural complexity, with adjusted R^2 values between 0.52 and 0.62. Moving beyond purely spectral analysis, García et al. [10] and Torres de Almeida et al. [22] investigated the integration of SAR data from platforms such as ALOS-PALSAR and Sentinel 1, respectively. Their studies suggested that while SAR data could enhance canopy measurements, the marginal improvements were often outweighed by the additional processing demands. These studies collectively suggest a promising yet complex relationship between canopy structure and other satellite observations, with a major focus on engineering useful features from spectral data to accurately infer canopy height.

Building on the exploration of connections between the structural information and satellite observations, Potapov et al. [12] advanced the field by introducing a global canopy height model that incorporates the GEDI sensor's canopy height product, integrated with the Landsat analysis ready dataset (ARD). One of the level 2 products of GEDI, known as relative canopy height, is derived from the full waveform collected by the sensor. Due to the sparse and limited nature of GEDI's data, Potapov et al. [12] trained a per-pixel bagged regression tree model to extrapolate the relative height 95 percentile (RH95) canopy height using Landsat ARD tiles. Aimed at developing a global model, their accuracy assessment resulted in an RMSE of 9.07 m and MAE of 6.36 m when compared with available airborne LiDAR data. To incorporate the varying times at which Landsat imagery was captured throughout the year, they created 16-day composites for each location [12]. In addition, they applied cloud masks to select only clear sky observations and derived annual metrics based on reflectance values corresponding to specific phenological stages [13].

Following the accelerated advancement of deep neural networks, Potapov's paper has succeeded in a couple of attempts at establishing global foundation models for canopy height extraction from RGB satellite imagery using LiDAR data from airborne and satellite platforms. Notably, Lang et al. [14] employed a convolutional neural network (CNN) architecture to extract canopy heights from 10-m resolution optical imagery from Sentinel 2. Their model reported an RMSE of 7.9 m and a bias of 1.7 m when compared with an independent airborne LiDAR dataset. Finally, Tolan et al. [15] achieved promising results by utilizing self-supervised vision transformers on high-resolution satellite imagery and training them on both airborne LiDAR and GEDI's spaceborne LiDAR data. Creating a canopy height map for the state of California and São Paulo, they achieved an MAE of 2.8 m at submeter spatial resolution.

Overall, progress in this field is marked by two distinct approaches: global foundation models built on large architectures and millions of images, yet still relying on single-temporal optical imagery [14], [15] and smaller scale models that integrate more diverse sensor information such as multispectral and SAR reflectance but rely on location-specific data sources and extensive feature engineering [9], [10], [11], [22].

In this article, we present a methodology that employs a locally tailored model while using diverse data sources that are ubiquitously available, ensuring that it can be reliably replicated at any global location. Inspired by the success of multitemporal features derived from Landsat's surface reflectance in the study by Potapov et al. [12], we adopt a deep learning approach capable of processing time-series data of surface reflectance from spectral satellites like Landsat to establish a relationship between changes in raw reflectance values in a region and the regional canopy height. We then develop a multimodal architecture that learns from time-series data from multisource satellites—Landsat and Sentinel 2 for their multispectral reflectance data, and Sentinel 1 for its active sensor radar data. By focusing on the county level, we present a framework capable of constructing localized canopy height models using limited LiDAR data from GEDI and multisensor data from Landsat, Sentinel 1, and Sentinel 2. As per our knowledge, this is the first paper that considers reflectance values as time series data and uses a multimodal recurrent neural network architecture to extract canopy height from multiple nonsynchronous satellite sources.

This article is organized into the following sections. Section II provides an overview of the datasets we used, the dependent and independent variables employed in our model, as well as details about the four test sites, where the experiments were conducted. Section III outlines the preprocessing procedure and the deep learning model employed as our framework. Later, in Sections IV and V, we present our experimental results and discuss our model's performance compared with independent airborne LiDAR data and state-of-the-art global canopy height models, and in Section VI, we provide the conclusion of our study.

II. Data and Test Sites

A. Spaceborne LiDAR Data

The GEDI is a full waveform LiDAR sensor installed on the Japanese Experiment Module-Exposed Facility (JEM-EF) of the ISS in December 2018. The LiDAR system consists

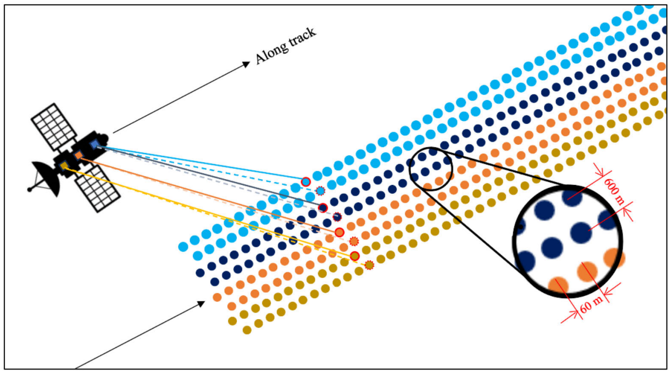


Fig. 1. GEDI sampling pattern.

of three lasers, two of which are referred to as "power" beams, while the third is divided into two beams known as the "coverage" beams. Each beam has a footprint diameter of approximately 25 m. These four beams generate eight tracks of data, with alternating shots dithered across the track. Each track is separated by around 600 m in the across-flight direction, and the center of each footprint is spaced 60 m apart in the along-track direction, as shown in Fig. 1. The GEDI team provides four levels of data products, depending on the extent of processing applied to the original observations. Level 1A represents the raw waveform data, which is geolocated in Level 1B. After the initial processing stage, Level 2A provides canopy top height and relative height metrics, and Level 2B offers canopy cover fraction and leaf area index. Finally, Level 3 provides a gridded product of Levels 2 and 4 gives information about the above-ground biomass (agb) [16].

For our research, we utilized the Level 2A data product of GEDI, which comprises the canopy height metrics. These metrics are computed by subtracting the elevation of the highest detected return from the lowest mode (corresponding to the ground) of the waveform [17]. GEDI also reports the height above ground of each energy quantile in the received waveform, which is the relative height metric. In the case of one of our test sites, Tippecanoe County, we downloaded 67 GEDI L2A files from the Earth Explorer website, corresponding to data collected from 2020 to 2022.

We examined the latitude and longitude corresponding to the highest return from each GEDI sample. If the latitude and longitude fell within the boundaries of the test site, we checked the sample's quality flag. The Level 2A product provides a quality flag for each observation. The flag is determined by a number of factors such as the energy, sensitivity, and amplitude of the returned signal. Only samples with a quality flag equal to 1—corresponding to the highest quality observation—were selected for further analysis [17].

B. Land Cover Classification Data

The National Land Cover Database (NLCD) is a dataset produced and distributed by the U.S. Geological Survey (USGS) [18], providing comprehensive descriptive data of land surfaces in the form of thematic classes.

With NLCD's definition, our test sites contain a total of ten classes for nonwater and nonurban landmass (out of the total 16 classes). These classes are listed in Table I—the ten nonwater and nonurban classes start from class number 31. Provided at a spatial resolution of 30 m, we use this data to mask out the water bodies and urban areas from our test sites, as they are beyond the scope of this article. Later, we also input these classes as supplementary information for our deep learning model.

TABLE I

NLCD 2019 Land Cover Classification Legend (Source: [29])

Class Number	NLCD Class
11	Open Water
12	Perennial Ice/Snow
21	Developed, Open Space
22	Developed, Low Intensity
23	Developed, Medium Intensity
24	Developed, High Intensity
31	Barren Land (Rock/Sand/Clay)
41	Deciduous Forest
42	Evergreen Forest
43	Mixed Forest
52	Scrub/Scrub
71	Grassland/Herbaceous
81	Pasture/Hay
82	Cultivated Crops
90	Woody Wetlands
95	Emergent Herbaceous Wetlands

C. Airborne LiDAR Data

Airborne LiDAR data from the USGS's 3-D elevation program (3DEP) was utilized for validating our test results. For Tippecanoe and Monroe Counties, Indiana the data was collected in Spring 2018. For Baker County, FL, USA, the data were collected between late 2018 and early 2020. Finally, for Piute County, UT, USA, there was not a single data collection campaign covering the entire county. The mountainous forest region on the left were surveyed as part of the Central Southern campaign and Statewide Kane campaign in 2020, the valley region in the middle was surveyed as part of the Southern campaign in 2018, while the Parker Mountain region on the right was surveyed as part of the Statewide South campaign in 2020. Even still, these only partially covered the entire Piute County, which we used for our accuracy assessment. According to the 3DEP quality assessment for vegetated areas, both the LiDAR point cloud and the digital terrain model (DTM) have submeter accuracy [19]. These datasets were employed to create a normalized digital height model (NDHM) that served as our validation dataset for assessing the accuracy of our results. The procedure for generating this validation dataset is discussed in Section III.

D. Multispectral Satellite Imagery

The Harmonized Landsat and Sentinel-2 (HLS) dataset is a collaborative effort between NASA and USGS. They provide multispectral surface reflectance from sensors onboard these satellites, namely the operational land imager (Landsat 8) and the multispectral instrument (Sentinel-2) [26], [27].

Sentinel's various spectral bands have native spatial resolutions ranging from 10 to 60 m, which are resampled to

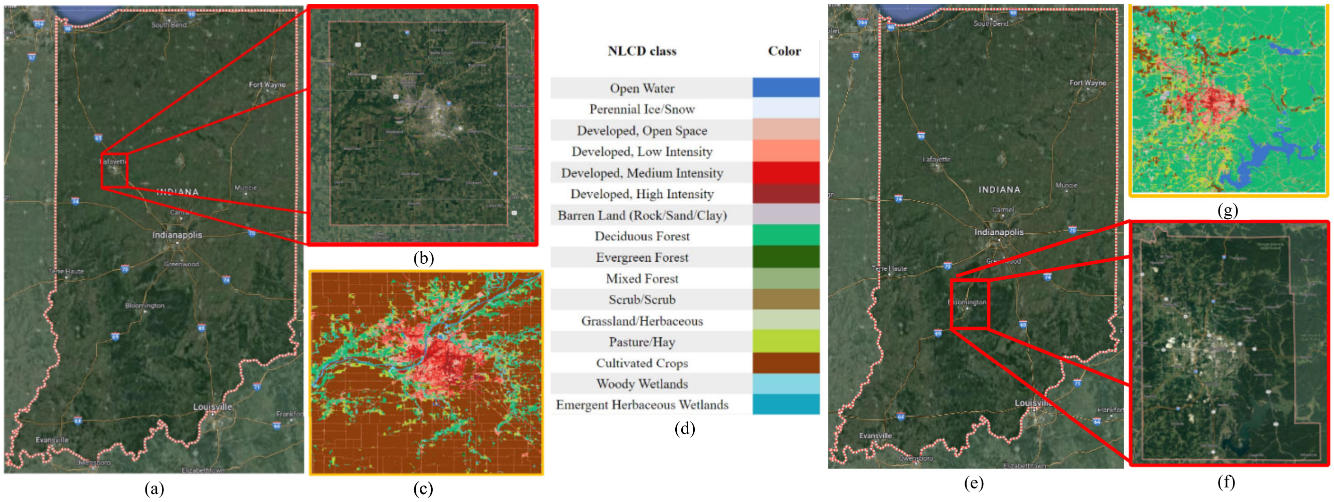


Fig. 2. (a) Tippecanoe County in Google satellite view of Indiana. (b) Google satellite view of Tippecanoe County. (c) NLCD map of Tippecanoe County. (d) NLCD map legend. (e) Monroe County in Google satellite view of Indiana. (f) Google satellite view of Monroe County. (g) NLCD map of Monroe County.

30 m to match Landsat's specifications. Multispectral data from these sensors undergo atmospheric correction, spatial co-registration, and common gridding. The dataset offers two primary products: L30 and S30. L30 consists of 30-m resolution data from Landsat, comprising ten bands covering the visible, near-infrared, shortwave infrared, and thermal infrared spectral ranges. The S30 product, corresponding to Sentinel-2, includes 13 bands covering an additional range of the near-infrared spectrum, including three red-edge bands and a broad near-infrared band. Unlike Landsat, the sensor on Sentinel-2 does not cover the thermal range.

The quality assessment for both HLS products is provided by the Fmask band, generated using the Fmask 4.2 software. Fmask assigns an 8-bit packed integer for each pixel, containing information about the presence of clouds, water, and snow. We unpack and modify the Fmask value for use as one of the features, with the exact modifications discussed in Section III.

E. Synthetic Aperture Radar Satellite Imagery

Sentinel 1, part of the European Space Agency's Copernicus Sentinel series, consists of two satellites equipped with C-band synthetic aperture radar (SAR) [21]. Working as an active sensor in the radar range of the electromagnetic spectrum, Sentinel 1 provides continuous data without interruption from clouds or aerosols.

We use the Level 1 ground range detected (GRD) data product from Google Earth Engine, which undergoes thermal noise removal, radiometric calibration, and terrain correction using the Sentinel 1 Toolbox. Specifically, we utilize two bands, "VV" and "VH," corresponding to the two polarizations collected by the Interferometric Wide Swath mode over land. As SAR data can penetrate through cloud, no cloud coverage information is provided with the data product.

F. Test Sites

To conduct our experiments, we selected four counties across the United States as test sites. Among these, Tippecanoe and Monroe Counties are located in the state of Indiana, with their specific locations illustrated in Fig. 2. According to the NLCD map, 77.46% of nonurban land in Tippecanoe County is covered by cultivated crops, followed by 11.85% covered by deciduous forests. In contrast, Monroe County's landscape is predominantly deciduous forests, accounting for 74.3% of the land cover, with 12.04% categorized as pasture/hay.

Our third test site was Baker County, where evergreen forests and woody wetlands dominate, collectively accounting for 81.37% of the land cover. Finally, we selected Piute County, a mountainous region characterized by shrubland (49.44%) in the valleys and evergreen forests (36.48%) covering the mountains (Fig. 3)

III. Methodology

A. Overview

The objective of this study is to develop a model that establishes a relationship between GEDI's height metric and multisensor reflectance values within a localized region, enabling the extrapolation of the height metric for the entire region. To explore the significance of each sensor's data, three different models are presented in this article. Model 1 is trained on the HLS L30 dataset, utilizing spectral reflectance bands collected by Landsat. Model 2 is trained on the HLS S30 dataset, comprising data collected by Sentinel-2's multispectral sensor. Finally, Model 3 showcases a multimodal approach that incorporates asynchronous time series input from Landsat, Sentinel 1, and Sentinel 2. In addition, each model is also given the pixel's NLCD class as one of the features. While NLCD is created using the Landsat dataset and is only available over the United States, an alternative for places outside the USA is Google's Dynamic World, which utilizes Sentinel data to create a global 10-m land cover dataset [24].

We conduct two sets of experiments in this article. First, we train and test Models 1–3 in two of our test sites—Tippecanoe County and Monroe County. This helps us to analyze how the input from different satellite sensors affect

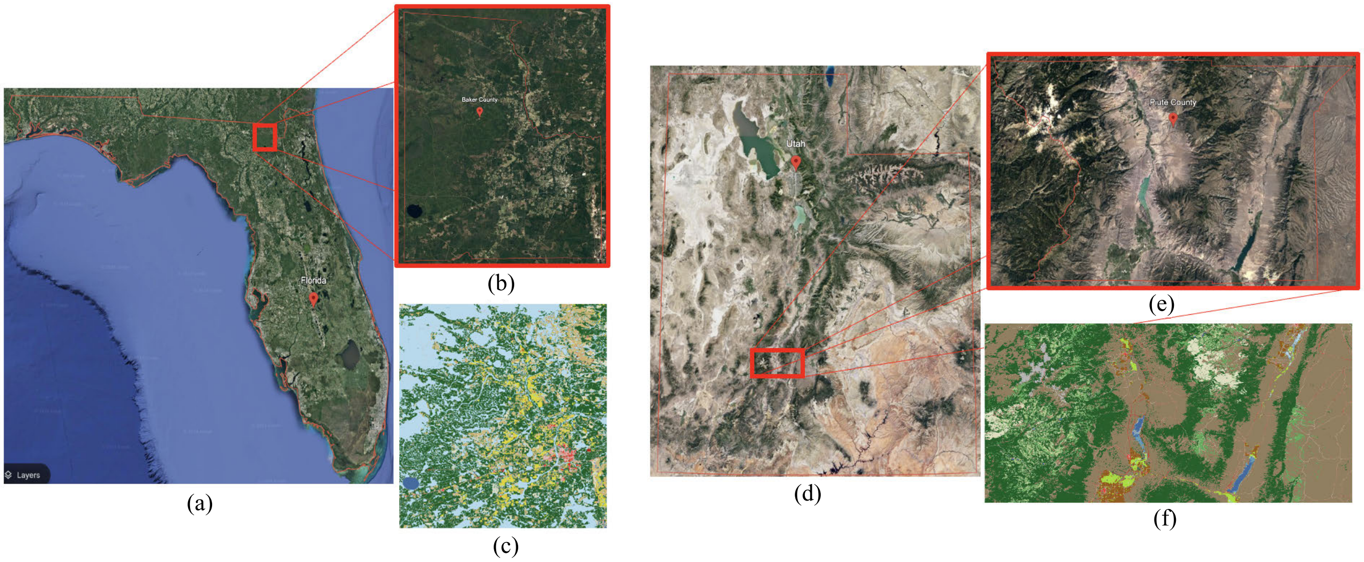


Fig. 3. (a) Baker County in Google satellite view of Florida. (b) Google satellite view of Baker County. (c) NLCD map of Baker County. (d) Piute County in Google satellite view of Utah. (e) Google satellite view of Piute County. (f) NLCD map of Piute County.

the overall accuracy of estimating the canopy height. Second, we train and test our Model 3 over Baker and Piute Counties, and compare the performance of our method against the state-of-the-art global canopy height models. Our models' output is the RH95 height for a particular pixel, with a spatial resolution of 30 m. The training process utilizes the entire training dataset of each county, and the final raster is tested against a validation dataset. The validation RH95 raster is generated using the LiDAR point cloud and DTM from the 3DEP dataset.

B. Preprocessing

The preprocessing stage of our model involves creating a raster for each of the aforementioned datasets, organizing the data in a structured manner for easy retrieval when constructing the feature vector. To illustrate the methodology, we will use Tippecanoe County as an example, with the same procedure applied to other counties. All calculations were performed using WGS 84 geodetic coordinates, i.e., latitude and longitude. We specifically used the geodetic coordinates to showcase a pipeline that could be applied to any location within the bounds of the GEDI footprint.

For Tippecanoe County, the base raster was established as follows: The top-left corner's coordinates were 40.56312 N and 87.09554 W, while the bottom-right coordinates were 40.21435 N and 86.69534 W. Utilizing the spatial resolution of HLS tiles over Tippecanoe County (0.00031443° decimal), we determined the number of rows and columns, resulting in 1272 rows and 1109 columns.

1) GEDI RH95: GEDI L2A files falling within the boundaries of our test site for the years 2020–2022 were downloaded. We discard the poor-quality observations using the quality flag provided with the data. For each sample, the latitude, longitude, and GEDI height metrics (RH90, RH95, and RH100) were stored. Subsequently, we compared these height metrics with our 3DEP validation dataset, finding that the RH95 height metric from GEDI exhibits the highest correlation and least mean bias error with the validation dataset. Hence, the RH95 height metric from GEDI was selected as our target variable for the model.

Using the base raster specifications, data from these files were saved in a sparse raster for each county. In cases where multiple RH95 height observations fell within the same cell, the values were averaged out.

2) Landsat Features: For our first feature set, we downloaded all Harmonized Landsat Sentinel-2 (HLS) L30 tiles covering our test sites in the years 2020–2022. A total of 69 rasters for Tippecanoe County were obtained, with each raster having 11 layers—10 surface reflectance bands and one quality assessment band.

3) Sentinel 2 Features: For our second feature set, we downloaded all HLS S30 tiles covering our test sites in the years 2020–2022. Due to the higher temporal resolution of Sentinel 2, we filtered out all tiles with more than 30% cloud coverage. This resulted in 103 Sentinel rasters for Tippecanoe County, each comprising 14 layers—13 surface reflectance bands and one quality assessment band.

4) Sentinel 1 Features: The last feature set was made using Sentinel 1 data. We had 176 tiles collected between 2020 and 2022 over Tippecanoe County. The spatial resolution of Sentinel 1 GRD product is approximately 10 m. To match the spatial resolution of HLS tiles, we applied a 5 × 5 mean filter and used nearest neighbors to create Sentinel 1 rasters for our test sites. This mean filter also helped clean the dataset by reducing random noise that is commonly found in radar data.

5) NLCD Raster: The NLCD 2019 (CONUS) raster was obtained from the USGS website and cropped according to the specified boundaries of our test sites. As the spatial resolution of NLCD was identical to Landsat, we utilized nearest neighbors to create an NLCD raster for all the counties.

This raster was employed to mask urban areas and water bodies from all our data rasters, as they were beyond the scope of our problem. The excluded NLCD classes were

11, 12, 21, 22, 23, and 24. Urban areas and water bodies were intentionally excluded due to their different relationship between height and spectral reflectance compared to natural landscapes. In addition, as mentioned before, these classes were used as additional input to our models, a discussion of which will follow in this section.

6) 3DEP RH95: Under the 3DEP initiative, each CONUS state conducted airborne LiDAR data collection which provides point cloud and DTM tiles.

The DTM tiles in Tippecanoe County have a spatial resolution of 2.5 feet and reports the ground height relative to the GRS 80 reference ellipsoid. To process the data, the point cloud files for Tippecanoe County were normalized using the DTM tiles. This normalization involved replacing the Z coordinate of each point in the cloud with the difference between the Z value and the DTM height of the corresponding pixel. The new Z value represents the height above ground for each point. Subsequently, the height-normalized point cloud was reprojected into WGS 84 geodetic coordinates, replacing the X and Y coordinates with longitude and latitude, respectively.

Based on the specifications of the established base raster, all points were grouped according to the raster pixel they fell under. For each raster pixel, the points belonging to that pixel were sorted by their height values. The 95th percentile height was chosen as the pixel value for the raster, referred to as 3DEP RH95. Selecting the 95th percentile height instead of the maximum height helps account for any outliers in the data collection procedure of 3DEP. These outliers could result from the laser beam hitting birds or drones flying above the canopies during the data collection process, which we aimed to avoid.

7) Building Feature Vector: We adopt a per-pixel training regime, where each of our samples corresponds to a particular pixel of the county's raster. Our features include time series reflectance values from one or more of the satellite sensors discussed above, and the one-hot vector of the pixel's NLCD class. Each of our models is trained over samples for which we have GEDI's RH95 height metric.

For a test site, two feature vectors were constructed—the training dataset and the test dataset. The training features comprised all the samples with GEDI satellite observations, excluding those classified as water bodies or urban land. For the test dataset, we compared the results of our model with an independent airborne LiDAR dataset. We included all samples in the county for which we had a 3DEP RH95 value, except those classified as water bodies or urban land.

For Tippecanoe County, the Landsat training dataset had a shape of (96 113, 69, 11), representing 96 113 training samples, 69 timestamps, and 11 bands. Meanwhile, the Sentinel 2 training dataset had a shape of (95 878, 103, 14), indicating 95 878 training samples, 103 timestamps, and 14 bands. Finally, the shape of Sentinel 1 training dataset was (95 912, 176, 2). The varying number of training samples in these datasets is due to the removal of samples with corrupted sensor values. The test dataset for Landsat and Sentinel 2 had 1.187 million samples, while Sentinel 1 had 1.184 million samples, each with corresponding numbers of timestamps and bands.

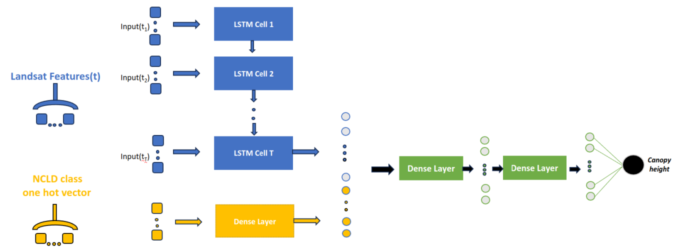


Fig. 4. Unimodal architecture for Model 1–estimates RH95 canopy height using HLS L30 time series dataset. HLS L30 was replaced with HLS S30 for Model 2.

8) Quality Assessment Band: The quality assessment band, named Fmask, is included in the HLS dataset and provides information on the presence of clouds and aerosols in the pixel. They are represented by an 8-bit packed binary integer [20]. With a packed binary integer, the numerical difference between two values does not correspond to the amount of contamination the pixel might have. To prepare it for our model, we converted all pixels without clouds (bit 1 equals 0) to 1. All pixels with clouds but low/climatology aerosol level (bit 7 equals 0) are given a value of 0.5. Finally, all pixels with clouds and moderate/high aerosol levels (bit 7 equals 1) are given a value of 0. The following is the code we developed to carry out the modification.

```
def unpack_Fmask(fmask):
    qa = 0

    if fmask/2 % 2 == 0:
        qa = 1 # Clear sky

    elif fmask/2 % 2 == 1:
        q = fmask // {2^7}

        if q % 2 == 0:
            qa = 0.5 # Partially clouded
        else:
            qa = 0 # Completely clouded

    return qa
```

C. Model Training and Evalution

1) Model Architecture: In this article, we employ two distinct deep learning architectures, both rooted in the recurrent neural network long short-term memory (LSTM) [25]. For Models 1 and 2, we construct a unimodal architecture utilizing a single layer of LSTM, followed by a fully connected dense layer, as shown in Fig. 4. In addition, we incorporate NLCD classes as input by separately passing the one-hot encoded vector of the class through a fully connected layer. The outputs from these two streams are concatenated and fed through two additional dense layers, culminating in the RH95 GEDI height metric output. Even though the NLCD classes arguably count as another modality, we consider these models unimodal as they learn from a single source of surface reflectance time series data.

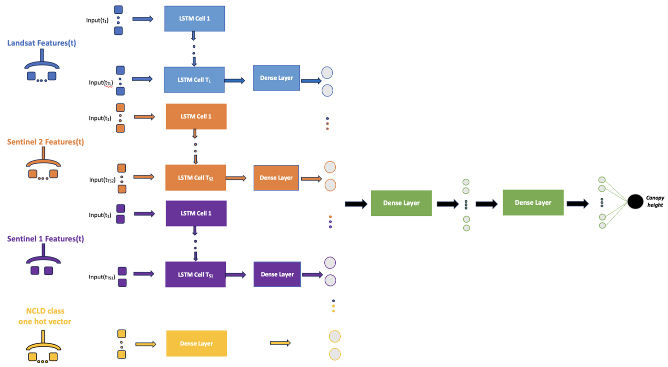


Fig. 5. Multimodal architecture for Model 3—estimates RH95 canopy height using HLS L30, HLS S30, and Sentinel-1 time-series datasets.

Model 3 constitutes a multimodal architecture capable of processing asynchronous time-series information from Landsat, Sentinel 1, and Sentinel 2. Here, the time-series data from each sensor undergoes a layer of LSTM and a fully connected layer separately, following which they are concatenated and sent through two additional fully connected layers, as shown in Fig. 5. Similar to Models 1 and 2, the encoded NLCD class is concatenated here as well. We employ the Adam optimizer as our activation function and utilize $L1$ loss for all our models. To set the number of nodes in the aforementioned layers, we utilize a hyperparameter tuner. This tuner executes the training regime ten times using different hyperparameters and selects the best ones based on the $L1$ loss of a set-aside validation dataset—the training data was split 80:20 into training and validation sets based on stratified sampling of canopy height.

As the final filtration step, we utilize the modified Fmask value to ensure that no Landsat or Sentinel date has more than 50% of training samples with moderate/high aerosol levels.

2) Validation: After completing the model training and predicting the heights for the test dataset, we proceeded to validate our results by comparing them with the 3DEP RH95 height. We calculated several performance metrics to assess our model's accuracy. These metrics included mean absolute error (MAE), root mean squared error (RMSE), R-square (R^2), and slope. The equations for each of these metrics are as follows.

1) *MAE:*

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i|. \tag{1}$$

2) *RMSE:*

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2}. \tag{2}$$

3) R^2*:*

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}. \tag{3}$$

Here, y_i represents the pixel values observed from the validation model, $\bar{y}$ represents the mean of those pixel values, x_i represents the pixel values predicted from our trained model, and n represents the total number of pixels. These metrics provide insights into our model's accuracy and performance. Our validation results in this article compare the overall performances of our models against the independently collected 3DEP dataset.

3) Global Canopy Products: The three global canopy map benchmarks that we compared our results against, are described as follows.

a) Potapov's model: Potapov et. al.'s model [13], accessible on their website [23], employs a bagged regression trees model on features derived from Landsat ARD Collection 1. Their model undergoes an extensive preprocessing procedure involving the generation of 16-day composites to ensure consistent features for a global-scale model, along with removing cloud-covered pixels. They extract 546 metrics from the Landsat files to train their model. To evaluate Potapov's model's performance in our study area, we conducted a comparative analysis of their model in all four counties against the 3DEP RH95 model.

b) ETH model: Lang et al.'s model [14], accessible through the Google Earth engine, employs a fully CNN over Sentinel 2 optical images, and produces a global canopy height model at a spatial resolution of 10 m. They use a sparse raster from GEDI RH98 to train their CNN model. To perform a fair comparison, we down-sample their model at our test sites to 30 m resolution by taking a mean of 9 pixels in each 30-m grid.

c) META model: Tolan et al.'s model [15], accessible through Amazon web services, uses a self-supervised learning regime over Maxar optical imagery for the states of California and São Paulo, followed by a CNN using GEDI RH95 measurements to produce a global canopy height model. Producing a canopy height model at 1-m resolution, we downsample their model to 30 m by taking 95 percentile height from the 900 pixels in each 30-m grid.

IV. Experimental Results

In this section, we present the outcomes from our three models, evaluating their performance and comparing them against the 3DEP RH95. Subsequently, we compare our multimodal model against the three state-of-the-art global canopy height model—Potapov et al. [12] (Potapov), Lang et al. [14] (ETH), and Tolan et al. [15] (META).

A. Comparison Among HLS L30 Model, HLS S30 Model, and Multimodal Model

For two of our test sites, Tippecanoe County, IN, USA, and Monroe County, IN, USA, we constructed RH95 CHMs using Model 1 (HLS L30), Model 2 (HLS S30), and Model 3 (a multimodal architecture). The output results and density plots are presented in Figs. 6 and 7, while the accuracy of these models compared to 3DEP RH95 is summarized in Table II.

In Tippecanoe County, the RMSE values for Models 1–3 are 4.145, 4.087, and 3.875 m, respectively. Meanwhile, in Monroe County, the RMSE values for Models 1–3 are 5.230, 5.228, and 4.993 m, respectively. The accuracy metrics

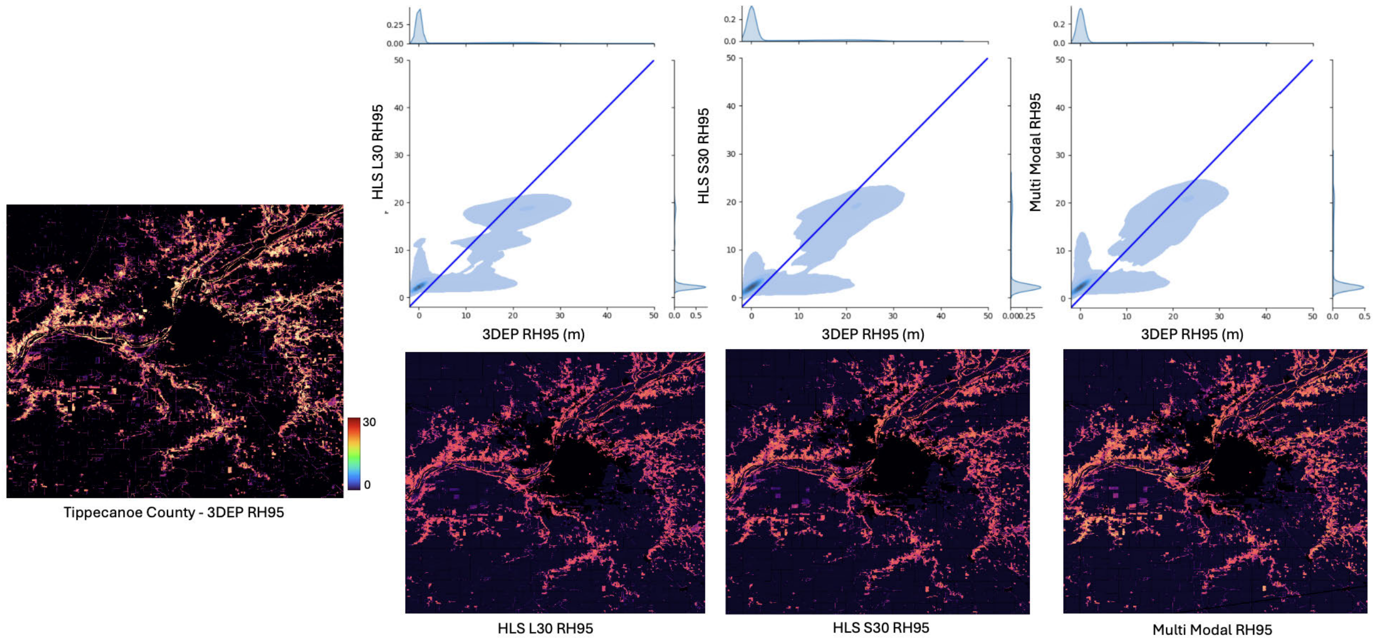


Fig. 6. Tippecanoe County: model results. 3DEP RH95 is the validation height map, while HLS L30 RH95, HLS S30 RH95, and multimodal RH95 are the estimated height map from Models 1–3, respectively. The three density plots show the performance of each model with respect to 3DEP RH95.

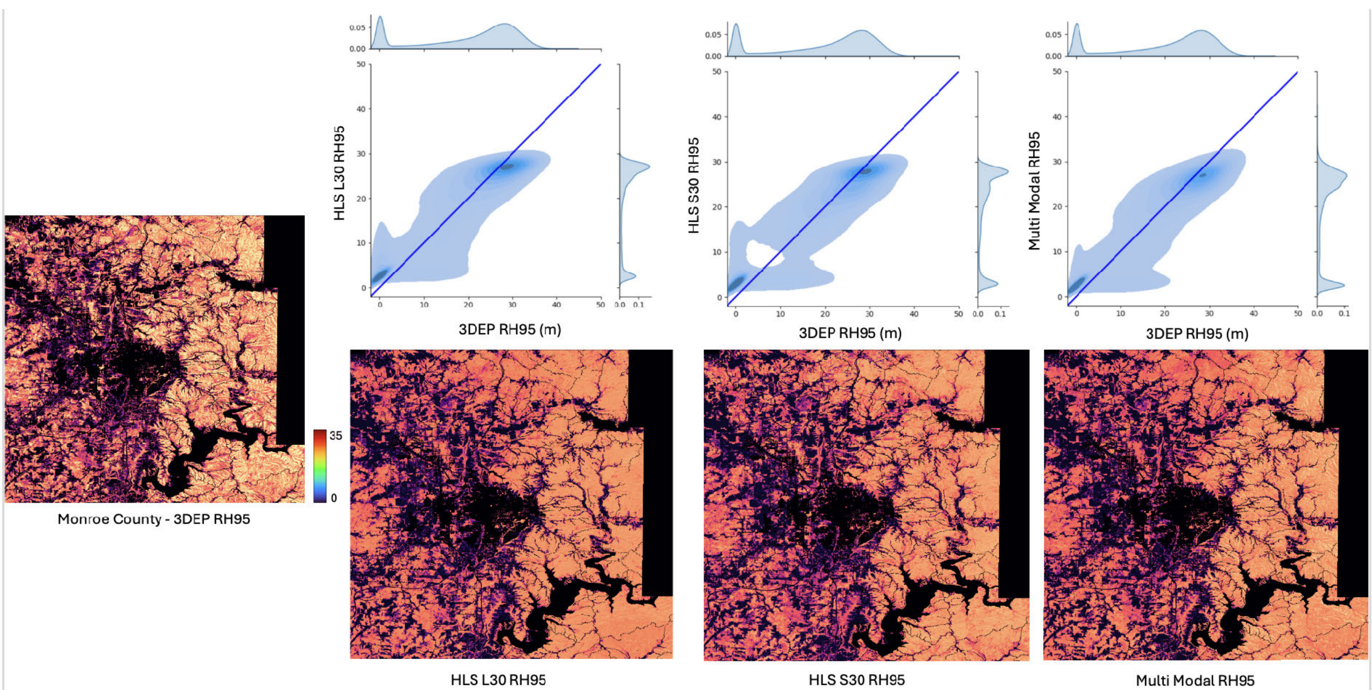


Fig. 7. Monroe County: Model results. 3DEP RH95 is the validation height map, while HLS L30 RH95, HLS S30 RH95, and multimodal RH95 are the estimated height map from Models 1–3, respectively. The three density plots show the performance of each model with respect to 3DEP RH95.

indicate a general trend where the HLS S30 model outperforms the HLS L30 model. This improvement is likely due to the higher temporal resolution of the Sentinel-2 satellite compared to Landsat. In addition, Sentinel-2 offers three extra red-edge bands in the near-infrared spectrum that are unavailable in the Landsat dataset, further enhancing performance.

The KDE plot for both test sites reveals that the HLS L30 model underestimates forest heights greater than 25 m, with a plateau around 20 m in Tippecanoe County and 30 m in Monroe County. Although the HLS S30 model mitigates this issue to some extent, the multimodal model surpasses both, accurately estimating heights beyond 30 m. Furthermore, this model demonstrates superior ability in capturing the in-class variability of forest land cover, as shown in both the RH95 map and the KDE plot.

The enhanced performance of the multimodal model underscores the advantage of integrating the spectral reflectance from both the Landsat and Sentinel-2 datasets, as they

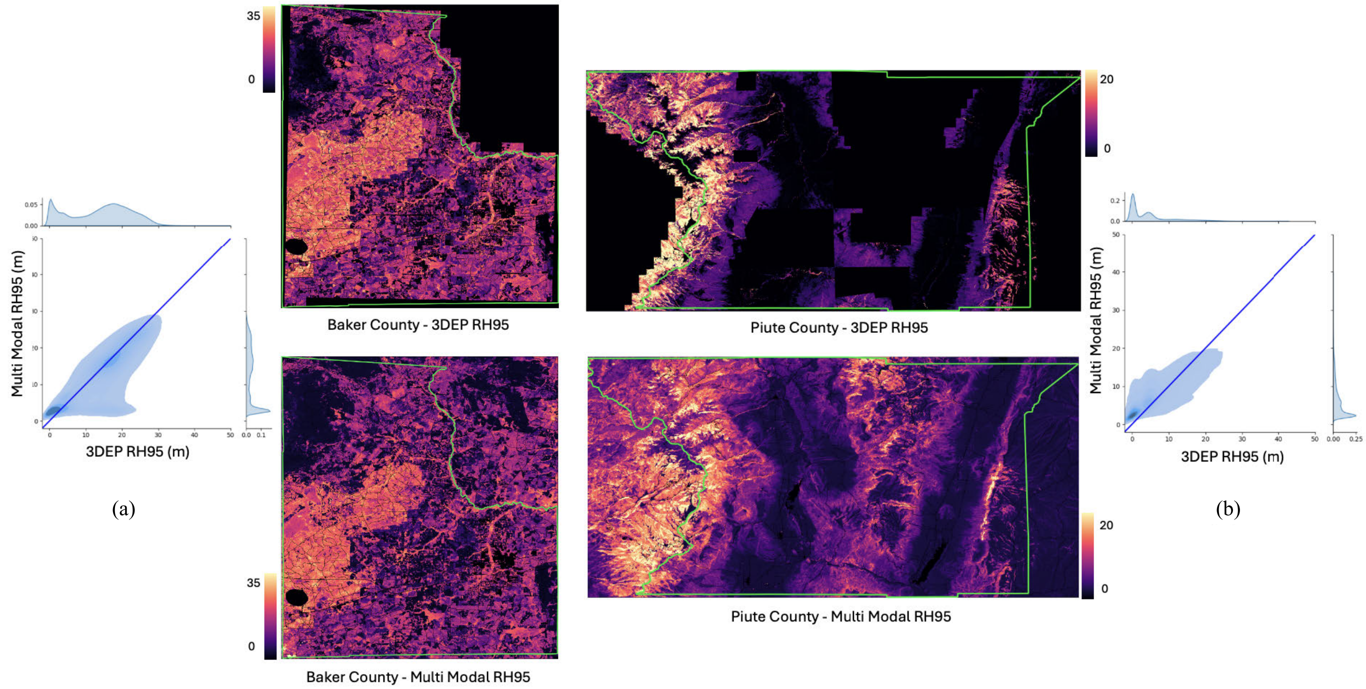


Fig. 8. (a) Baker County: model results. (b) Piute County: model results. 3DEP RH95 is the validation height map, while multimodal RH95 is the estimated height map from Model 3. The density plots show the performance of the model with respect to 3DEP RH95.

TABLE II
COMPARISON OF MODEL PERFORMANCE IN TIPPECANOE AND MONROE COUNTIES

Metric	HLS L30 Model	HLS S30 Model	Multi Modal Model (MMM)
Tippecanoe County			
MAE	3.117m	3.037m	3.015m
RMSE	4.145m	4.087m	3.895m
R^2	0.744	0.751	0.774
Slope	0.612	0.632	0.687
Monroe County			
MAE	3.982m	3.933m	3.763m
RMSE	5.230m	5.228m	4.993m
R^2	0.779	0.778	0.798
Slope	0.749	0.752	0.783

complement each other's temporal gaps. In addition, the inclusion of Sentinel-1 data, which provides active radar bands, further improves the model's robustness.

B. Comparison With Global Canopy Height Models

We applied the multimodal architecture to two additional counties—Baker County and Piute County. The resulting RH95 maps and density plots are presented in Fig. 8.

In this section, we compare our model's results across the four test sites against global canopy height models. Fig. 9 shows the canopy height maps from Potapov, ETH, and META for Tippecanoe and Monroe Counties, alongside their density plots against our 3DEP validation dataset. Similarly, Fig. 10 presents their results for Baker and Piute Counties.

In Baker County and Piute County, we achieved RMSE values of 5.397 and 3.742 m, respectively (Table III). The KDE plots from each of the four test sites show consistent results along the one-to-one line, particularly in forest-dense regions of Monroe and Baker Counties. In both counties, our model achieved the best R^2 and RMSE values compared to the global canopy height products.

One limitation of our model is the slight overestimation in cropland and shrub-dominated areas, which are predominant in Tippecanoe and Piute Counties, respectively. This overestimation negatively impacts our MAE in the regions, where the global canopy models—defaulting to zero height—tend to perform better.

C. Boundary Pixel Error

Upon inspecting the KDE plots across all our models, a recurring issue contributing to prediction errors became apparent: the presence of flat concentrations of points parallel to the X- and Y-axes. These points represent either nonzero true heights being predicted as close to zero, or near-zero heights being predicted as nonzero. From KDE plots in Figs. 9 and 10, we see that similar phenomenon can be seen in the benchmark datasets as well. Further investigating the cause of this error, we conducted a spatial analysis of the locations of these pixels in our model. Predominantly found at the boundaries of land cover classes, such as the edges of forested areas, these pixels were spread across the county at all our test sites.

To quantify the effect of this phenomenon, we leveraged the NLCD and eliminated the boundary points of all land cover types. We then recalculated our accuracy metrics for the remaining points. As shown in Table IV, this adjustment resulted in noticeable improvement across all our test sites. The RMSE in Tippecanoe County decreased by nearly a meter, now at 3.037 m, while in Monroe County, the RMSE dropped to 4.305 m. Similarly, the RMSE in Baker County and Piute

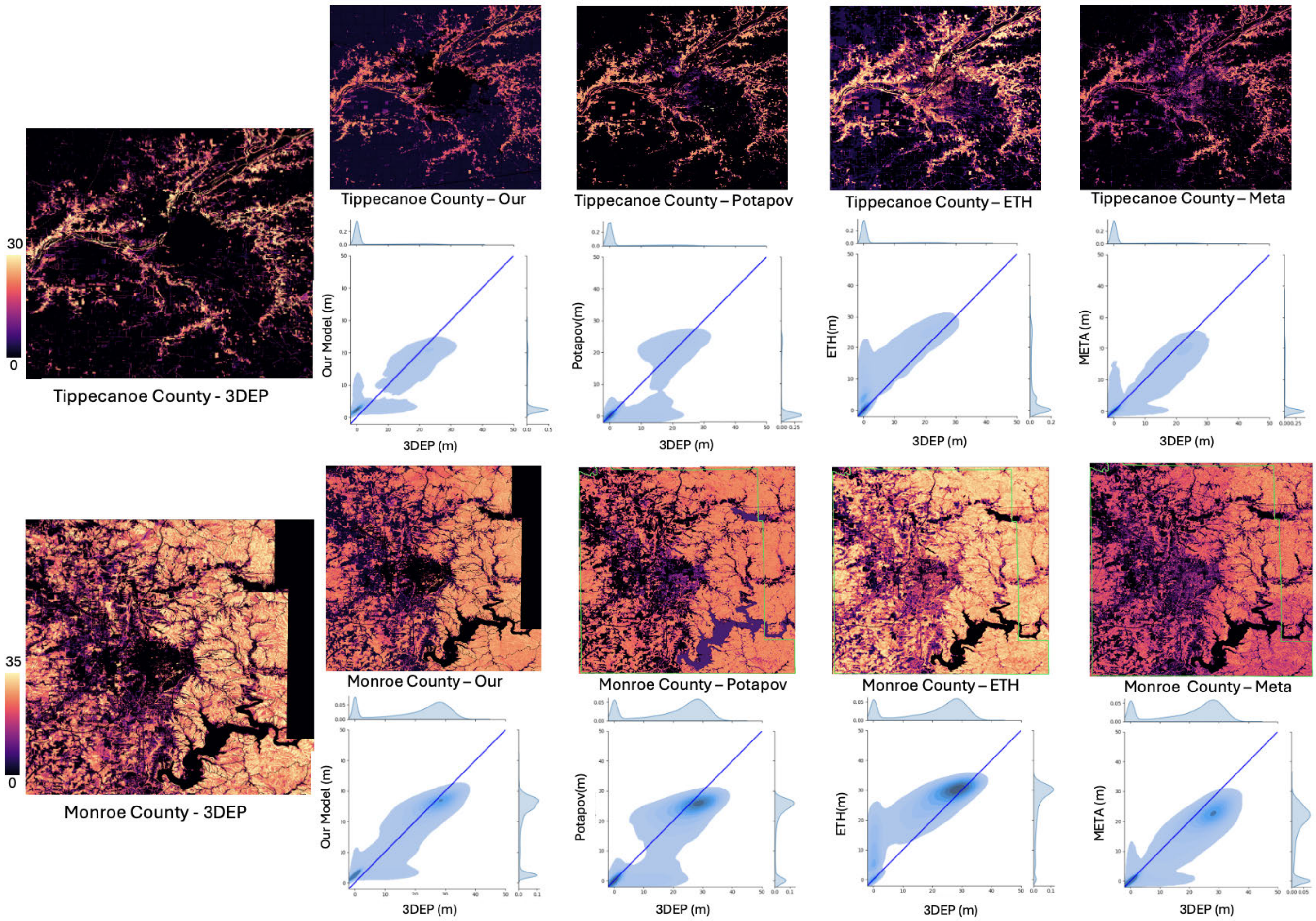


Fig. 9. Our, Potapov, ETH, and META canopy height models for Tippecanoe and Monroe Counties, and their density plots against 3DEP RH95.

County reduced to 5.095 and 3.515 m, respectively. This improvement, which can be found across various accuracy metrics, illustrates that the discrepancy between the model's output and the validation data is higher at the boundaries of land cover classes than otherwise and is one of the main contributors to the overall error.

The errors associated with these boundary pixels may arise from various factors. Given that we are working at a spatial resolution of 30 m, it is highly likely that these pixels are mixed, with reflectance originating from multiple land cover types. In addition, intrinsic spatial referencing errors across different platforms could be a contributing factor, particularly at boundaries where there is a sudden change in height values. Addressing this issue is a broader challenge that requires further investigation in the future.

V. Discussion

Accurate and timely canopy height modeling is essential for tracking forest health, carbon storage, and ecosystem changes. Precise models enable better monitoring of deforestation and biomass, while timely updates help capture rapid environmental shifts. These insights are critical for improving land management and informing climate action, aligning with the goals of this study to enhance localized canopy height estimation using multisensor satellite data.

This study aimed to improve localized canopy height estimation by integrating sparse LiDAR with multisensor satellite data. From the literature, we see that previous attempts at creating localized models depended heavily on hand-crafted features and traditional ML algorithms to estimate canopy heights. On the other hand, in recent years, studies have employed deep learning to create larger scale canopy height models. Though some studies were able to generate global canopy height models, they offer less flexibility in terms of the height metric they use, or for which time frame the model was trained for. This ambiguity can hinder timely and accurate applications where precise data vintage is critical. Consequently, smaller scale forest studies require a pipeline that is easily trainable with limited resources and time.

In our proposed pipeline, we utilized the HLS dataset, which synchronizes the tiling structure of Landsat 8/9 and Sentinel-2 A/B, providing surface reflectance values from the corresponding sensors. In our first set of experiments, we built models using data solely from the Landsat counterpart of the HLS dataset (HLS L30), solely from the Sentinel-2 counterpart (HLS S30), and then proposed a multimodal architecture that integrates data from both satellites, along with SAR data from Sentinel-1. As the results demonstrate progressive improvements in accuracy metrics from Models

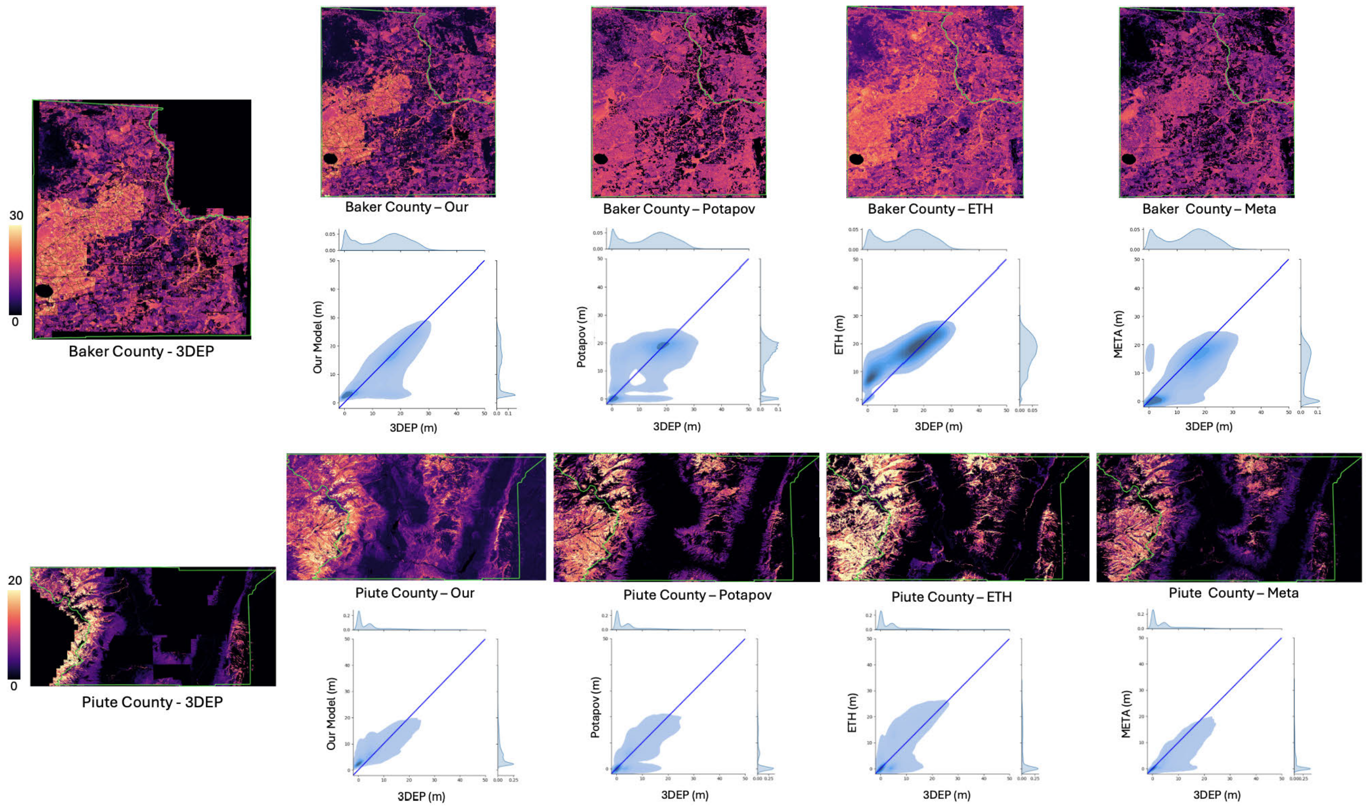


Fig. 10. Our, Potapov, ETH, and META canopy height models for Baker and Piute Counties, and their density plots against 3DEP RH95.

1 to 3, our experiments underscore the advantage of the three additional red-edge bands in Sentinel-2 for predicting various canopy heights [28], as well as the benefits of incorporating data from different satellites, each complementing the other's temporal gaps. Additionally, while radar data (e.g., Sentinel-1) has commonly been used to derive 3-D information about Earth's surfaces, such as building structures, this study reveals that radar can also be highly beneficial for analyzing forest structure.

In our second set of experiments, we compared the results of our multimodal model with various global canopy height products across the four test sites. These sites were chosen to evaluate our model's performance in different land cover conditions. Notably, our model outperformed the global models most significantly in Baker and Monroe Counties, both of which are heavily forested regions. As a general trend across all our test sites, we achieved better R^2 and RMSE compared with the benchmark models. However, our performance in terms of MAE was more limited, particularly in Tippecanoe and Piute Counties. Both RMSE and R^2 penalize large deviations from the true values more heavily, indicating that our model excels at distinguishing between different land cover types and differentiating taller canopies, like trees, from shorter canopies, such as crops and shrubs. On the other hand, in Tippecanoe and Piute Counties, our model tends to slightly overestimate canopy heights for crops and shrubs, contributing to the higher MAE. Since MAE linearly penalizes deviation from the true value, a greater number of samples with small differences can have a more significant impact on MAE than a smaller number of samples with larger differences.

It is worth noting that the ETH and META models are originally provided at 10- and 1-m resolutions, respectively, but we downsampled them to 30-m resolution for comparison with our 30-m 3DEP RH95. We only present their result to show our model's comparable performance against them. In addition, our results demonstrate that beyond extracting spatial features from RGB imagery as done by the ETH and META models, temporal changes in reflectance values can also be leveraged to infer canopy heights—particularly when multispectral satellites such as Landsat and Sentinel-2 provide a wide array of spectral bands. The future research should explore incorporating both spatial and temporal features into network architectures to enhance canopy height estimation at local and global scales.

Finally, using a localized approach against a global foundation model involves certain tradeoffs. While the primary motivation behind our model is to create a transparent and customizable pipeline for users conducting smaller scale studies, it inherently limits its applicability outside the geographical extent of its training data. Scaling the multimodal LSTM network to a larger geographical scope requires additional preprocessing steps, primarily to synchronize time-series inputs across different locations. As our model leverages the time series of reflectance values from satellites, the epochs from each satellite are consistent only within the scope of their respective tiles. To expand this approach to larger scales, it is essential to harmonize the time series from different locations.

TABLE III
COMPARISON OF MODEL PERFORMANCE ACROSS DIFFERENT COUNTIES

Metric	Our MMM*	Potapov	ETH	META
Tippecanoe County				
MAE	3.015m	1.815m	3.073m	1.623m
RMSE	3.895m	4.870m	5.160m	3.836m
R^2	0.774	0.646	0.600	0.780
Slope	0.687	0.798	0.960	0.741
Monroe County				
MAE	3.763m	4.499m	5.393m	5.083m
RMSE	4.993m	6.890m	7.338m	6.741m
R^2	0.798	0.615	0.564	0.632
Slope	0.783	0.797	0.667	0.732
Baker County				
MAE	3.668m	5.527m	4.387m	4.653m
RMSE	5.397m	7.361m	5.559m	6.297m
R^2	0.581	0.221	0.559	0.429
Slope	0.749	0.502	0.565	0.684
Piute County				
MAE	2.956m	2.536m	3.087m	2.179m
RMSE	3.742m	4.999m	4.623m	3.775m
R^2	0.631	0.342	0.437	0.625
Slope	0.614	0.732	0.939	0.667

*MMM: Multi Modal Model

TABLE IV
ANALYZING BOUNDARY PIXEL ERROR IN OUR MULTIMODAL MODEL

Metric	Before removing Boundary Pixels	After removing Boundary Pixels
Tippecanoe County		
MAE	3.015m	2.561m
RMSE	3.895m	3.037m
R^2	0.774	0.744
Slope	0.687	0.701
Monroe County		
MAE	3.763m	3.356m
RMSE	4.993m	4.305m
R^2	0.798	0.815
Slope	0.783	0.754
Baker County		
MAE	3.668m	3.395m
RMSE	5.397m	5.095m
R^2	0.581	0.635
Slope	0.749	0.768
Piute County		
MAE	2.956m	2.811m
RMSE	3.742m	3.515m
R^2	0.631	0.629
Slope	0.614	0.652

This could involve creating 15 day or monthly aggregates of images from each satellite, akin to the methodology adopted by Potapov et al. [13].

VI. CONCLUSION

This article introduces a novel framework for deriving canopy height observations from multimodal satellite reflectance values. Through experimentation in our four test sites, we have demonstrated the feasibility of localized site extrapolation using raw surface reflectance values obtained from satellites, such as Landsat, Sentinel-1, and Sentinel-2. Leveraging deep learning architectures, our framework achieves performance comparable to state-of-the-art results, particularly by treating temporal reflectance data as time series and employing recurrent neural networks like LSTM.

In recent years, significant research efforts have been directed toward using deep learning architectures that estimate canopy heights from satellite imagery and airborne/spaceborne LiDAR data. However, these efforts have predominantly relied on RGB imagery, inferring canopy height based solely on spectral information from the three bands and the spatial context within the images. Given the availability of a continuous influx of diverse information, including multispectral bandwidths and radar data, our objective was to explore the potential of incorporating such data to enhance canopy height inference. Our results support this hypothesis, demonstrating progressively enhanced performance from Landsat to Sentinel-2, which offers broader coverage of the electromagnetic spectrum, to a multimodal model that also integrates SAR data from Sentinel-1. Since these datasets are readily accessible through application programming interfaces (APIs), future work could focus on automating the preprocessing steps and further streamlining the creation of training datasets. Additional improvements might involve exploring super-resolution techniques to mitigate boundary-pixel errors and evaluating the robustness of our model across other biomes, such as tropical and boreal forests. Finally, our findings advocate for leveraging temporal variations in reflectance data, as these variations offer deeper insights into how canopy height influences the spectral signatures observed from space.

ACKNOWLEDGMENT

This work is based upon efforts supported by the PERSEUS grant, #2023-68012-38992 under USDA NIFA and the EFFICACI grant, #NR233A750004G044 under USDA NRCS. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NIFA/NRCS or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

REFERENCES

[1] Y. Pan, R. A. Birdsey, O. L. Phillips, and R. B. Jackson, "The structure, distribution, and biomass of the world's forests," *Annu. Rev. Ecol., Evol., Systematics*, vol. 44, no. 1, pp. 593–622, Nov. 2013, doi: 10.1146/annurev-ecolsys-110512-135914.

[2] R. A. Houghton et al., "Carbon emissions from land use and land-cover change," *Biogeosciences*, vol. 9, no. 12, pp. 5125–5142, Dec. 2012, doi: 10.5194/bg-9-5125-2012.

[3] *Work of the Statistical Commission Pertaining to the 2030 Agenda for Sustainable Development*, United Nations, New York, NY, USA, Jul. 2017.

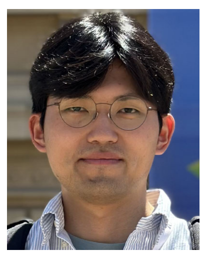
[4] S. Roy et al., "Estimating above ground biomass (AGB) and tree density using Sentinel-1 data," *Spatial Modeling in Forest Resources Management* (Environmental Science and Engineering). Cham, Switzerland: Springer, pp. 259–280, Oct. 2021, doi: 10.1007/978-3-030-56542-8_11.

[5] S. Illarionova, D. Shadrin, V. Ignatiev, S. Shayakhmetov, A. Trekin, and I. Oseledets, "Estimation of the canopy height model from multispectral satellite imagery with convolutional neural networks," *IEEE Access*, vol. 10, pp. 34116–34132, 2022, doi: 10.1109/ACCESS.2022.3161568.

[6] L. Wallace, A. Lucieer, C. Watson, and D. Turner, "Development of a UAV-LiDAR system with application to forest inventory," *Remote Sens.*, vol. 4, no. 6, pp. 1519–1543, May 2012, doi: 10.3390/rs4061519.
[7] R. Dubayah et al., "The global ecosystem dynamics investigation: High-resolution laser ranging of the Earth's forests and topography," *Sci. Remote Sens.*, vol. 1, Jun. 2020, Art. no. 100002, doi: 10.1016/j.srs.2020.100002.
[8] T. Markus et al., "The ice, cloud, and land elevation satellite-2 (ICESat-2): Science requirements, concept, and implementation," *Remote Sens. Environ.*, vol. 190, pp. 260–273, Mar. 2017, doi: 10.1016/j.rse.2016.12.029.
[9] C. Pascual, A. García-Abril, W. B. Cohen, and S. Martín-Fernández, "Relationship between LiDAR-derived forest canopy height and Landsat images," *Int. J. Remote Sens.*, vol. 31, no. 5, pp. 1261–1280, Mar. 2010, doi: 10.1080/01431160903380656.
[10] M. García, S. Saatchi, S. Ustin, and H. Balzter, "Modelling forest canopy height by integrating airborne LiDAR samples with satellite radar and multispectral imagery," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 66, pp. 159–173, Apr. 2018, doi: 10.1016/j.jag.2017.11.017.
[11] E. A. LaRue et al., "Linking Landsat to terrestrial LiDAR: Vegetation metrics of forest greenness are correlated with canopy structural complexity," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 73, pp. 420–427, Dec. 2018, doi: 10.1016/j.jag.2018.07.001.
[12] P. Potapov et al., "Mapping global forest canopy height through integration of GEDI and Landsat data," *Remote Sens. Environ.*, vol. 253, Feb. 2021, Art. no. 112165, doi: 10.1016/j.rse.2020.112165.
[13] P. Potapov et al., "Landsat analysis ready data for global land cover and land cover change mapping," *Remote Sens.*, vol. 12, no. 3, p. 426, Jan. 2020, doi: 10.3390/rs12030426.
[14] N. Lang, W. Jetz, K. Schindler, and J. D. Wegner, "A high-resolution canopy height model of the Earth," *Nature Ecol. Evol.*, vol. 7, no. 11, pp. 1778–1789, Sep. 2023, doi: 10.1038/s41559-023-02206-6.
[15] J. Tolan et al., "Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial LiDAR," *Remote Sens. Environ.*, vol. 300, Jan. 2024, Art. no. 113888, doi: 10.1016/j.rse.2023.113888.
[16] *GEDI: Products*. Accessed: Aug. 2, 2023. [Online]. Available: https://gedi.umd.edu/data/products/
[17] M. Hofton and J. B. Blair, "Algorithm theoretical basis document (ATBD) for GEDI transmit and receive waveform processing for L1 and L2 products," Goddard Space Flight Center, Greenbelt, MD, USA, Tech. Rep., 2019. [Online]. Available: https://lpdaac.usgs.gov/documents/581/GEDI_WF_ATBD_v1.0.pdf
[18] J. Wickham, S. V. Stehman, D. G. Sorenson, L. Gass, and J. A. Dewitz, "Thematic accuracy assessment of the NLCD 2016 land cover for the conterminous United States," *Remote Sens. Environ.*, vol. 257, May 2021, Art. no. 112357, doi: 10.1016/j.rse.2021.112357.
[19] *IN Indiana Statewide LiDAR 2017 B17-West*, United States Geological Survey (USGS), Reston, VA, USA, Airborne LiDAR Rep., Dec. 2020.
[20] M. Claverie, J. G. Masek, J. Ju, and J. L. Dungan, "Harmonized landsat-8 Sentinel-2 (HLS) product users guide," Nat. Aeronaut. Space Admin. (NASA), Product Version 1.2, Washington, DC, USA, Tech. Rep., 2017.
[21] *Sentinel-1: ESA's Radar Observatory Mission for GMES Operational Services*, 1st ed., K. Fletcher, Ed., Noordwijk, The Netherlands: ESA Communications, Mar. 2012.
[22] C. Torres de Almeida et al., "Canopy height mapping by sentinel 1 and 2 satellite images, airborne LiDAR data, and machine learning," *Remote Sens.*, vol. 14, no. 16, p. 4112, Aug. 2022, doi: 10.3390/rs14164112.
[23] P. Potapov et al. (2019). *Global Forest Canopy Height*. Accessed: Aug. 2, 2023. [Online]. Available: https://glad.umd.edu/dataset/gedi
[24] C. F. Brown et al., "Dynamic world, near real-time global 10 m land use land cover mapping," *Sci. Data*, vol. 9, no. 1, p. 251, Jun. 2022, doi: 10.1038/s41597-022-01307-4.
[25] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
[26] J. Masek et al., "HLS operational land imager surface reflectance and TOA brightness daily global 30m v2.0," NASA EOSDIS Land Processes DAAC, Tech. Rep. HLSL30.002, 2021, doi: 10.5067/HLS/HLSL30.002.
[27] J. Masek et al., "HLS Sentinel-2 MSI surface reflectance daily global 30m v2.0," NASA EOSDIS Land Processes DAAC, Tech. Rep. HLSS30.002, 2021, doi: 10.5067/HLS/HLSS30.002.
[28] W. Li, Z. Niu, R. Shang, Y. Qin, L. Wang, and H. Chen, "High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with sentinel-1, Sentinel-2 and landsat-8 data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 92, Oct. 2020, Art. no. 102163, doi: 10.1016/j.jag.2020.102163.
[29] J. Wickham, S. V. Stehman, D. G. Sorenson, L. Gass, and J. A. Dewitz, "Thematic accuracy assessment of the NLCD 2016 land cover for the conterminous United States," *GIScience Remote Sensing*, vol. 60, no. 1, 2023, doi: 10.1080/15481603.2023.2181143.

Arnav Goel received the B.Tech. degree in civil engineering from the Indian Institute of Technology Kanpur, Kanpur, India, in 2021, the M.S. degree in civil engineering from Purdue University, West Lafayette, IN, USA, and the Ph.D. degree in geomatics, specialization in computational engineering from the Lyles School of Civil Engineering, Purdue University, in 2023.

His research focuses on using satellite remote sensing data to study natural ecosystems, such as forests and developing state-of-the-art tools for Earth observation.



Hunsoo Song received the B.S. and M.S. degrees in civil and environmental engineering from Seoul National University, Seoul, South Korea, in 2018 and 2020, respectively, and the Ph.D. degree in civil engineering, specialized in geomatics and computer engineering from Purdue University, West Lafayette, IN, USA, in May 2024.

He is currently a Geospatial Data Scientist and a Civil and Environmental Engineer, focused on the intersections of urban inequity, accessibility, and climate challenges. He is currently a Post-Doctoral Associate with the Hixon Center for Urban Sustainability and the Seto Laboratory, Yale University, New Haven, CT, USA. His research interests include advancing urban sustainability through innovative 3-D urban mapping and analysis, leveraging deep learning methods, and geospatial digital twin simulations.

Jinha Jung (Senior Member, IEEE) received the B.S. and M.S. degrees in civil, urban, and geosystem engineering from Seoul National University, Seoul, South Korea, in 2003 and 2005, respectively, and the Ph.D. degree from the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN. USA, in 2011.

He is currently an Associate Professor with the Lyles School of Civil Engineering, Purdue University. His research interests include advanced light detection and ranging (LiDAR) data analysis for interdisciplinary research, leveraging his specialties in remote sensing, geospatial data science, and high-performance computing.