MACHINE LEARNING
Science and Technology

PAPER • OPEN ACCESS

Deep-learning-based canopy height model generation from sub-meter resolution panchromatic satellite imagery

To cite this article: Charles J Abolt *et al* 2025 *Mach. Learn.: Sci. Technol.* **6** 015013

IOP Publishing Mach. Learn.: Sci. Technol. 6 (2025) 015013 https://doi.org/10.1088/2632-2153/ada47e

MACHINE LEARNING Science and Technology

OPEN ACCESS

RECEIVED
12 April 2024

REVISED
16 September 2024

ACCEPTED FOR PUBLICATION
30 December 2024

PUBLISHED
22 January 2025

PAPER

Deep-learning-based canopy height model generation from sub-meter resolution panchromatic satellite imagery

Charles J Abolt[1,*], Javier E Santos[1], Adam L Atchley[1], Lucas Wells[2], Daithi Martin[2], Russell A Parsons[3] and Rodman R Linn[1]

[1] Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, United States of America
[2] Silvx Labs, Missoula, MT, United States of America
[3] USFS RMRS Fire Sciences Laboratory, Missoula, MT, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: chuck.abolt@utexas.edu

Keywords: forestry, optical, lidar, fire, CNN

Supplementary material for this article is available online

Abstract

Canopy height models (CHMs) with sufficient resolution to distinguish individual trees are useful for a variety of applications. However, standard techniques to acquire such data, such as airborne lidar surveying, are often prohibitively expensive. Deep learning techniques for generating CHMs from high-resolution imagery are an attractive option to reduce costs. To date, success with these methods has been demonstrated using multichannel aerial photography and specialized satellite data products derived from multiple sensors, neither of which is commonly available at temporal resolutions finer than one year. Here we demonstrate a method to generate sub-meter resolution CHMs in three forests in California using a more abundant data source: sub-meter resolution, panchromatic satellite imagery from a single sensor. We show that phenology and species composition play important roles in model transferability; when trained using imagery from a single conifer forest in autumn, the model performs well on autumn imagery from a second conifer forest several hundred kilometers distant with no re-training. With modest additions to the training dataset, the same model generates minimally biased estimates of canopy height in both conifer and deciduous forests during multiple seasons. Because the model operates on satellite data with global coverage and a relatively short return interval, we propose its suitability to extrapolate tree-level canopy height data to remote regions and conduct high-temporal resolution monitoring of forest structure. We furthermore demonstrate the workflow's applicability to fire modeling by conducting simulations in forests populated by trees measured using both this approach and airborne lidar surveying. We find minimal differences in fire behavior relative to a baseline case in which only statistical distributions of tree height and crown area are known. This result underscores the value of forest structural information derived from our workflow for improving the fidelity of wildland fire simulations, among other ecological applications.

1. Introduction

Canopy height data are fundamental to many forestry applications, such as forest inventory and growth monitoring (Vastaranta *et al* 2013). In other disciplines, uses for canopy height data range from carbon stock assessments (Behera *et al* 2017), to habitat mapping (Coops *et al* 2016) and hydrologic modeling (Varhola and Coops 2013). Historically, the standard workflow for generating products with spatial coverage at the forest scale or greater has been to measure canopy height directly in field plots, then extrapolate these measurements to a standard area unit (acres or hectares), possibly with aid from remote sensing analysis (Tomppo *et al* 2010, Fischer and Traub 2019). In recent years, this general approach has remained common, though direct field measurements of tree height have often been supplemented or replaced by remotely sensed estimates, derived either from photogrammetry or airborne or satellite lidar (Potapov *et al* 2021, Lang *et al* 2022, California Forest Observatory 2023). Commonly, these methods yield medium to coarse resolution gridded products (i.e. cells $\geqslant$ 10 m) which, despite lacking sufficient detail to distinguish individual trees, are useful for many purposes, including estimation of crucial ecological properties such as above ground biomass or net primary production (Frolking *et al* 2009). When such data products have a temporal component, they have often been applied to quantify forest degradation and recovery following disturbance and to monitor timber harvesting activity (Frolking *et al* 2009, Hansen *et al* 2013, Ceccherini *et al* 2020). At continental to global scales, data products of medium to coarse spatial resolution are also commonly used to parameterize climate, hydrologic, and biogeochemical processes within the land surface components of Earth system models (Levis 2010, Fisher *et al* 2018).

Recently, improvements in the availability of fine-resolution (grid cells $\leqslant$ several meters) canopy height models (CHMs) have accelerated the extraction or modeling of tree-level information for forest inventories (Zhang *et al* 2015, Budei *et al* 2018, Dalagnol *et al* 2022). Fine-resolution CHMs are usually derived from airborne lidar surveying or photogrammetric analysis of aerial or satellite imagery. In contrast to coarser resolution estimates, these data products permit direct measurement of spatial variability in the distributions of tree-level parameters such as height or crown radius, which can inform estimates of forest species composition and improve understanding of ecosystem services (Kauranne *et al* 2017). Additionally, as climate change has recently driven shifts in fire occurrence and impacts across the globe, there has been growing interest in using high-resolution CHMs to produce detailed fuel maps. These data inform physics-rich fire models such as FIRETEC or the fire dynamics simulator, which simulate complex fire phenomena at spatial resolutions as fine as the decimeter-scale (Mell *et al* 2009, Parsons *et al* 2011, 2018, Ziegler *et al* 2017, Moinuddin and Sutherland 2020, Marcozzi *et al* 2023). Even in relatively low-intensity burns limited to surface fuels, forest structure plays an important role in model simulations due to its impact on wind fields and the distribution of litter, both of which influence spread rate, residence time, the soil heat flux, and other fire behavior metrics.

Despite their benefits, applications of CHMs with sufficient resolution to detect individual trees have been limited by the cost of conducting airborne lidar and photogrammetric surveys (Potapov *et al* 2021, Lang *et al* 2022). In some regions of the planet this problem is improving. Within the United States, for example, the 3D elevation program of the U.S. Geological Survey has produced an enormous volume of publicly available airborne lidar point cloud data over the past two decades which can be processed into CHMs. However, the resolution of the data varies, with most datasets having below eight points per square meter. While such data can be used to produce CHMs (Oh *et al* 2022), the accuracy with which individual trees can be detected has been shown to be significantly lower than in surveys with higher point density (Sparks *et al* 2022). In general, higher density lidar data tends to be from more recent acquisitions and is only available in selected areas. Because airborne lidar data is not available in many parts of the globe, some researchers are exploring the potential of developing CHMs from airborne imagery. For example, Li *et al* (2023) recently pioneered the application of computer vision and deep learning to extract tree-level canopy height data from aerial imagery acquired over Denmark and Finland. Shortly afterward, Wagner *et al* (2024) mapped canopy height across California using 0.6 m resolution multispectral (RGB + near infrared) aerial photography from the USDA-NAIP program, and Tolan *et al* (2024) mapped canopy height throughout the continental United States using Maxar Vivid2 mosaic imagery, which is derived from several commercial satellite sensors. These approaches are attractive and cost-effective relative to airborne lidar or photogrammetry. However, each approach presents limitations. For example, the high-resolution aerial photography used by Li *et al* (2023) and Wagner *et al* (2024) is not available globally and is typically collected at low temporal resolution. By contrast, the satellite imagery on which Tolan *et al*'s (2024) workflow relies is available anywhere on the planet's surface, but it is only updated at intervals of one to several years at a given area of interest. Moreover, although the CHMs produced by Tolan *et al* cover vast spatial domains, they typically lack sufficient detail to distinguish crowns, and therefore no attempt has been made to segment the results into trees or tree-approximate-objects (TAOs).

Here, we seek to apply a similar approach, but leverage much more broadly available data captured from a single spaceborne sensor. We advance the extraction of tree-level canopy height information from high-resolution imagery by demonstrating that sub-meter resolution CHMs can be generated through deep-learning-based processing of off-nadir (by up to 10°), single band (i.e. panchromatic) images from individual sensors among the WorldView constellation of satellites, which provide global imagery at 0.5 m or higher resolution. We designed our deep learning model to process panchromatic data as input due to its higher data availability relative to RGB and multispectral images and the consistently fine (50 cm or higher) resolution of the images. By leveraging data that is available from any of the three WorldView satellites, we maximize the temporal frequency with which an area of interest may be observed, which is advantageous for applications like growth monitoring or assessment of damage following natural disasters. Using this approach, we demonstrate that a model trained in a coniferous forest in central California robustly estimates

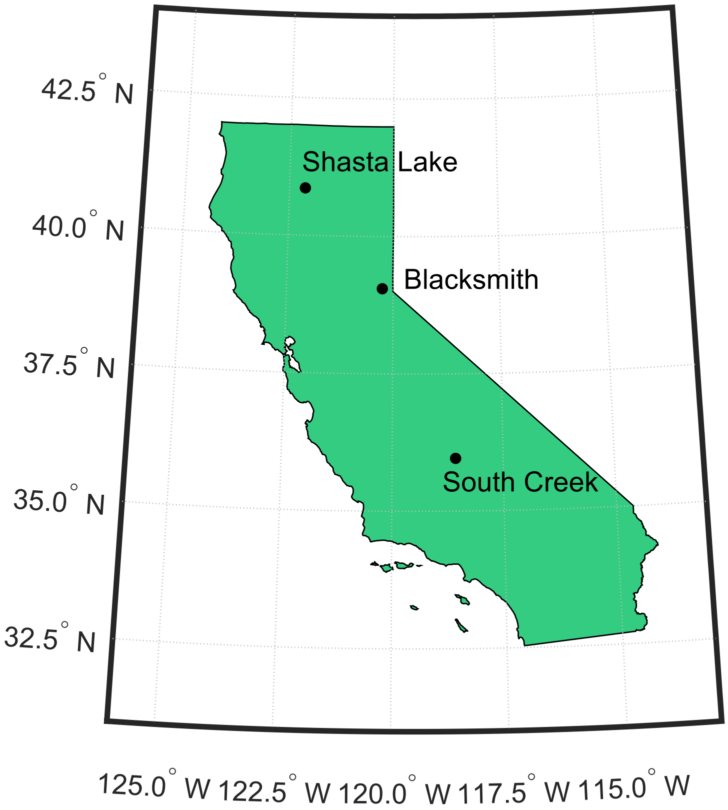

Figure 1. Locations of the three study areas within the state of California.

canopy height in two forests several hundred kilometers distant, either with no additional training data all or with modest additions of training data from the target forests. To highlight the rich spatial detail of our results, we then show that the incorporation of forest structural information derived our workflow into a set of physics-rich fire simulations results in fire behavior that closely matches simulations informed by more expensive airborne lidar data. Because our model processes imagery from tasking satellites with a repeat intervals of several days, we propose its suitability as a tool to (1) generate CHMs with sufficient detail to distinguish tree crowns from regions lacking high-resolution airborne lidar surveys and (2) conduct high temporal resolution forest inventories in areas where fine-resolution CHMs are already available.

2. Study area and methods

2.1. Study areas

Our study focused on three forest sites in California (figure 1) from which airborne-lidar derived CHMs, acquired by the United States Forest Service (USFS) between 2012–2018, were available at 25 cm resolution (Xu *et al* 2018). At each site, panchromatic imagery from the WorldView constellation of satellites was downloaded from two or more dates through the EnhancedView Web Hosting Service (https://evwhs-digitalglobe.com; supplementary table S1). Prior to model training, the imagery at each site was visually inspected and compared against the CHM. Any images in which signs of disturbance (such as wildland fire, or anthropogenic forest clearing) were observed between the dates of the imagery and lidar data acquisition were discarded. Care was also taken to control for potential spatial misalignment between the satellite imagery and lidar data. Depending on the sensor, the upper bounds of geolocation errors in uncorrected WorldView imagery is estimated at ∼3–8 m. In practice, the accuracy is often far greater. For this study, we employed only data in which no misalignment was visually discernible between the satellite images and the lidar-derived CHM. Therefore, no image registration processes were necessary prior to training. For processing by our neural network, both the satellite imagery and the CHM were re-projected into the local UTM coordinate system and resampled at a uniform spatial resolution of 50 cm.

2.1.1. Blacksmith

Blacksmith, the largest study area in our analysis, is situated in the Sierra Nevada mountains near 120.52° W 39.05° N, approximately 20 km due west of Lake Tahoe. The heavily forested study area is dominated by

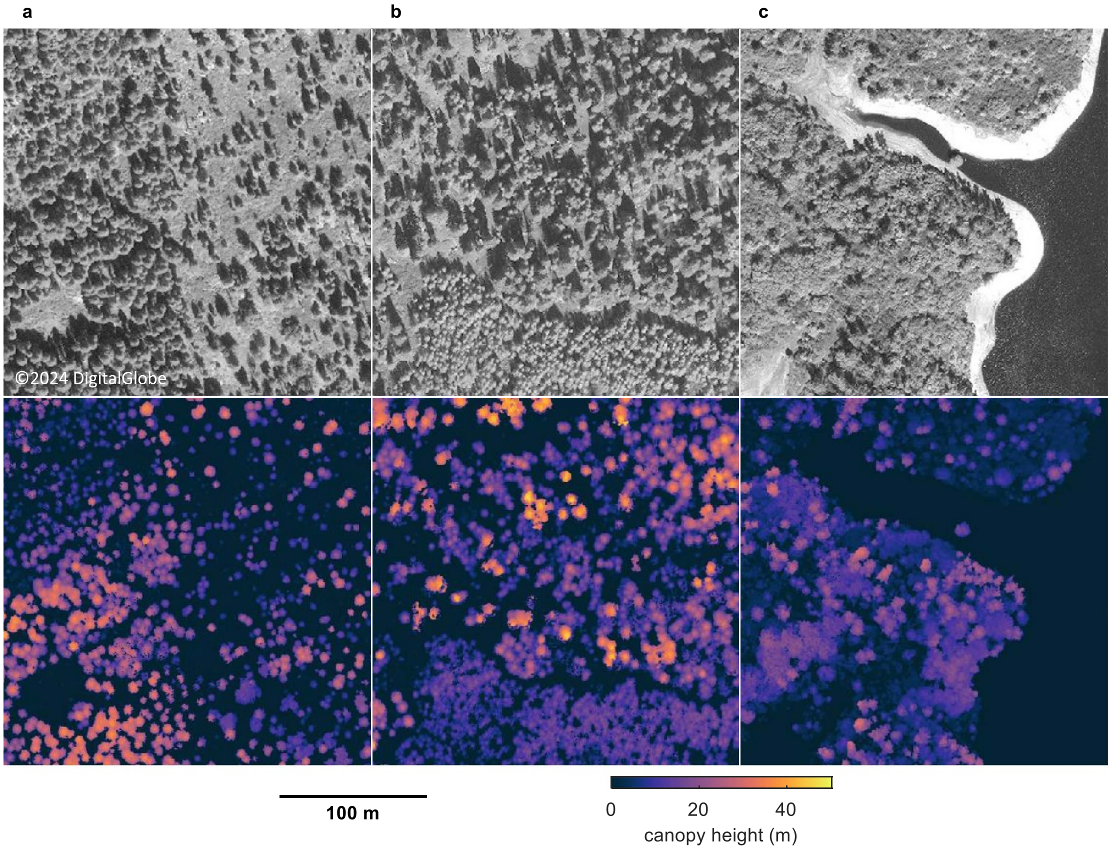

Figure 2. Samples of panchromatic imagery (top row) and lidar-derived CHMs (bottom row) from the (a) Blacksmith, (b), South Creek, and (c) Shasta Lake study areas.

conifers; USFS forest inventory analysis (FIA) forest types mapped within its boundaries include California mixed conifer (52% coverage), fir/spruce/mountain hemlock (10%), Douglas fir (8%), western oak (7%) and ponderosa pine (4%) (Gray 2012). A USFS CHM at the site was acquired in 2012 and covers 396 km^2 of terrain, ranging in elevation from 700–2200 m above sea level (masl), with a median elevation of 1570 masl. Samples of the satellite imagery and lidar-derived CHM available from Blacksmith are show in figure 2(a). Four cloud free satellite images overlapping the study area were downloaded: from 16 November 2011 covering 61.75 km^2 of the study area; 2 September 2012 (190.5 km^2); 14 October 2013 (313.25 km^2), and 31 October 2013 (54.75 km^2).

2.1.2. South creek

The South Creek study area is also located in the Sierra Nevada, approximately 400 km southeast of Blacksmith, near 118.57° W 35.96° N. A USFS CHM of the study area was acquired in 2012 and covers 64.2 km^2, extending in elevation from 1150–2420 masl, with a median of 1840 masl. FIA forest types represented include California mixed conifer (35% coverage), western oak (30%), ponderosa pine (7%) and fir/spruce/mountain hemlock (4%) (figure 2(b)). Two cloud free panchromatic satellite images overlapping the study area were available: from 5 August 2013 (overlapping 44.25 km^2 of the CHM) and 13 November 2013 (50 km^2).

2.1.3. Shasta lake

Shasta Lake, the smallest study area, is located in northern California's Klamath Mountains near 122.12° W 40.86° N, approximately 240 km northwest of Blacksmith. The USFS CHM at the site was acquired in 2013 and encompasses 48.8 km^2, centered on the basin of a stream flowing into one of the northeast arms of the reservoir. Elevation ranges from 58–2600 masl (median = 450 masl) and FIA forest types include western oak (50% coverage), California mixed conifer (35%), and ponderosa pine (7%) (figure 2(c)). Two cloud-free satellite images overlapping the study area were downloaded for the analysis, from 10 August 2011 (32.25 km^2) and 25 April 2018 (19.25 km^2).

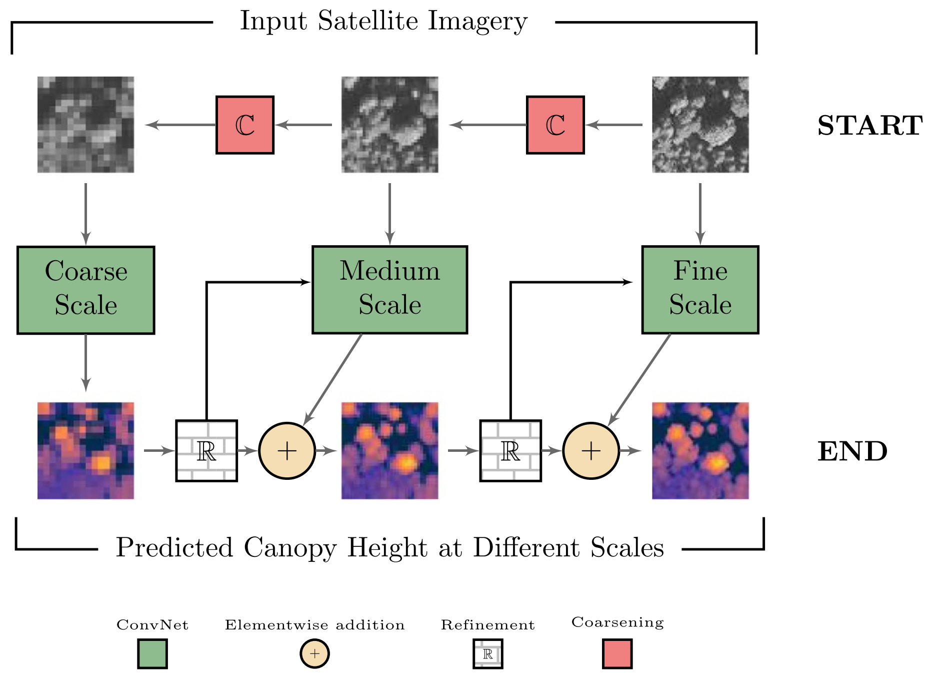

Figure 3. Flow of information through MS-Net, modified from Santos *et al* (2021). A set of fully convolutional neural networks, operating on copies of the input data which have been coarsened to various degrees, work together to generate a predicted CHM with identical size and resolution to the satellite image.

2.2. Deep learning workflow

We approached our task of extracting tree height information from satellite imagery as an image-to-image regression problem, in which a satellite image passed as input to the network yields a CHM of equal size and spatial resolution. To perform this task, we employed a modified version of the MultiScale Network for hierarchical regression (MS-Net), a deep learning framework which was originally developed to predict steady-state fluid velocity fields within porous media (Santos *et al* 2021). MS-Net comprises not a single neural network, but rather a collection of fully convolutional networks, which operate on copies of the input image which have been coarsened to various degrees and work cooperatively to generate a final prediction of the output (i.e. the CHM). A high-level schematic of the flow of information through MS-Net is shown in figure 3. When an input image is passed to MS-Net, both it and the target image are first coarsened a user-specified number of times, reducing the number of rows and columns in the images by a factor of two at each scale. At the coarsest scale, the image is processed by a fully convolutional (non-contracting) network, which is optimized to reproduce the coarsened target image as closely as possible. The coarse-level prediction is then 'refined' (or upscaled) to the next coarsest scale and concatenated with the corresponding input image. This concatenated image is passed through another fully convolutional network, the output of which is combined through element-wise addition with the prediction from the previous scale to generate a higher resolution prediction. The upscaling process repeats until a predicted CHM is generated with spatial resolution equal to the input.

The unique architecture of MS-Net was designed to maximize the workflow's field of vision (or the portion of an input image to which a single pixel in the prediction is sensitive) relative to computational cost. A key feature of MS-Net is that the fully convolutional network at the coarsest scale (which incurs the least cost, as it operates on the smallest arrays) contains the largest number of convolutional filters. This effectively tasks the network at the coarsest scale with extracting nearly all the information regarding the broad contours of the target, while the networks at the remaining scales (which employ fewer filters) add successively finer detail. In our implementation, information was processed through a workflow comprising three scales. Relative to the first version of MS-Net described by Santos *et al* (2021), the primary modifications were to replace the 3D convolutional kernels with 2D kernels and to use a ReLU activation function for the final layer of the networks, ensuring that estimated canopy height was never negative. The fully convolutional network at the finest scale had eight filters, and the number of filters increased by a factor of four at each successively coarser scale. At each scale, the convolutional network contained fifteen layers. The workflow in whole comprised 2.2×10^6 trainable parameters and was trained and executed on a single NVIDIA RTX A6000 GPU.

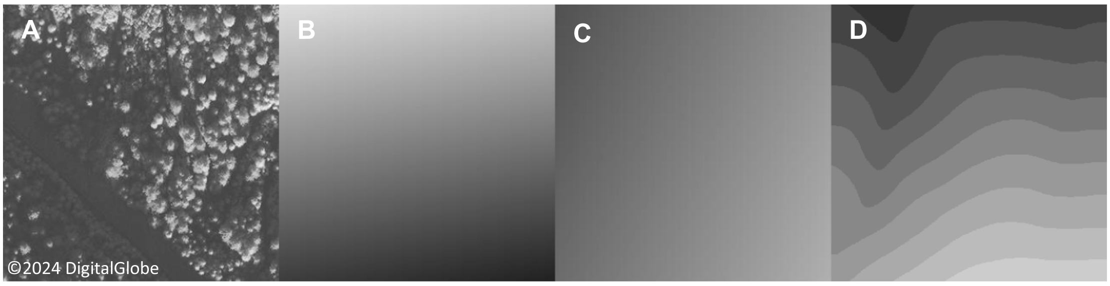

Figure 4. Input data processed by MS-Net included four features, the first of which was the panchromatic satellite image (a). Planes oriented normal to the solar angle (b) and sensor look angle (c) at the time of image acquisition provided additional contextual information. The underlying topography (d) was extracted from the National Elevation Dataset and converted to a grayscale image.

2.3. Data pre-processing and model training

Our workflow was trained on images at 50 cm resolution spatial resolution, measuring 512 × 512 pixels (or 256 m along each side). To prepare these images, the satellite images and corresponding CHMs were first divided into tracts of one square kilometer, which were each further subdivided into sixteen tiles of 250 × 250 m, with overlapping 3 m buffers. The satellite imagery, which was downloaded in 8-bit tiff format, was rescaled to assign every pixel as a double-precision number between 0 and 1. The pixel values of the target CHMs, which were downloaded as double-precision arrays with units of meters above the ground surface, were divided by 80 (or the approximate height of the tallest tree in the data set), to assign all pixels a value in the same interval.

In early iterations of the model, we found that the use of panchromatic satellite imagery presented two unique challenges. First, because the satellite always views the ground surface from an off-nadir perspective, the tops of trees tend to appear displaced from the ground directly underneath by a distance that scales with canopy height. Second, tree crowns are frequently found in the shadows of adjacent trees, which obscure details to a greater degree than would be expected in multi-band images. To mitigate errors caused by these issues, we decided to incorporate into the MS-Net workflow additional contextual information regarding the sensor look angle and the solar angle, extracted from the metadata of each satellite image. Additionally, we provided information about the underlying topography, which influenced shadow projection. Each of these three features was represented as an additional band in the input to the MS-Net workflow (figure 4). The sensor look angle and solar angle were represented as planes normal to the position in the sky of the satellite and the sun, respectively, at the time of image acquisition. Topographic data were extracted at 2 m resolution from the National Elevation Dataset, then resampled at 50 cm resolution and rescaled globally to the interval between 0 and 1, enforcing a consistent relationship between image intensity and topographic relief. Our final workflow therefore processed four bands of input imagery to perform canopy height inference.

2.4. Training procedure and evaluation of accuracy and transferability

To test the robustness and transferability of our model, we trained two iterations of MS-Net. In the first iteration, we trained and validated the model using all available data from Blacksmith (the largest study area, comprising 620 km^2 of imagery acquired over four dates), then immediately tested performance at the two smaller sites, South Creek and Shasta Lake. In the second iteration, we trained and validated the model using all available information from Blacksmith, supplemented with ∼25% of the available data from the two remaining sites (comprising an additional 36 km^2 of imagery acquired over four dates), then evaluated improvements to performance in the remaining data from the two smaller sites. In both iterations, 75% of the processed (i.e. non-test) data were used for training, reserving 25% for validation. The loss function was computed as the pixel-wise root mean square error (RMSE) between the target and prediction, and the training algorithm employed an Adam optimizer. Training proceeded until loss computed on the validation dataset reached a minimum, which required approximately two to three days of computing time on a single GPU.

Beyond the training loss function, model accuracy in the test datasets was evaluated using several metrics which were selected for their relevance to forest inventories and modeling of land surface and fire processes. These included the total fraction of the test region occupied by a forest canopy higher than 3 m and the number and median area of gaps in the canopy. Additionally, the predicted and target CHMs from each test forest were processed through a conventional, automated tree segmentation algorithm to compare the approximate number of trees extracted from each. The segmentation algorithm worked in two stages. In the first, point features representing individual tree-tops were demarcated by identifying local maxima in a version of the CHM which had been smoothed with a Gaussian filter. Maxima which did not appear to be surrounded by tree crowns of adequate area were then eliminated, with the minimum acceptable tree crown area described as a function of height at the tree-top (Pitkanen *et al* 2004). In the second stage, the estimated tree-tops were used as input in a marker-controlled watershed operation, which delineated the area of the canopy corresponding to each tree (Chen *et al* 2006). After comparing the number of trees segmented from the lidar-derived and satellite-derived CHMs, the tree-level accuracy of the canopy height predictions was evaluated by using the boundaries of individual trees segmented from the predicted CHM to compare maximum height in the predicted and target data. Note that, while we have referred to this procedure as tree delineation, this segmentation approach can be said to produce (TAOs, Jeronimo *et al* 2018), which are not necessarily individual trees. This is an important distinction as tree top maps of TAOs do not characterize forests in the same way traditional inventories do, but can rather be viewed as a discretized, object-based derivation of the continuous CHM data.

2.5. Application of our workflow to fire modeling

To demonstrate the utility of incorporating forest structural information derived from our workflow in physics-based fire behavior models, we simulated fires at two experimental plots using FIRETEC. The sites each measured 500 m by 500 m and were selected from the Blacksmith validation dataset and South Creek test dataset from the first iteration of training (supplementary figures S1 and S2). FIRETEC is a fire-atmosphere model that solves Navier–Stokes equations to represent the tight coupling of fire dynamics with atmospheric conditions and forest structure characteristics (Linn *et al* 2002). Finely resolved forest structures are represented in a discrete domain structure, characterized by two-meter horizontal discretization and variable vertical discretization (Linn *et al* 2005). The tight coupling between combustion and wind has been validated against prior observations of fire behavior (Linn and Cunningham 2005, Linn *et al* 2012, Hoffman *et al* 2015). Previous model experiments have demonstrated that forest structure conditions can be linked to characteristic fire behavior (Parsons *et al* 2017, Atchley *et al* 2021). Therefore, we conducted a set of experiments comparing the results of simulations that incorporate forest structural information derived from our workflow (hereafter called the machine learning, or ML simulations) with simulations of the same forests mapped with airborne lidar (ALS simulations).

In addition to comparing ALS and ML-derived forest structure characteristics, we generated a third case, referred to as an unstructured (US) domain, representing a condition of unknown tree placement. The data extracted from the ALS and ML CHMs included tree location, tree height, and crown radius for each tree (or TAO) in the domain. In the US case, we randomized tree location but retained tree height and crown radius variables from the ALS CHM, mimicking a situation in which tree population statistics are available, but spatially explicit tree-level positions are not. For all domains, the height-normalized crown base height and height to maximum crown radius were set to 0.4 and 0.5, respectively. Fine fuel density was set at 0.446 km m^{-3}, fuel moisture (wet mass over dry mass) was 112.5%, and size scale, in inverse meters, analogous to surface area to volume ratio, was 0.0005 for all trees in the domain. To focus comparisons on differences in canopy fuels, we employed the same homogenous surface fuel for all cases, characterized by a density of 0.23 kg m^{-3}, a fuel height of 0.35 m, and a fuel moisture of 6%.

CFD fire simulations model complex wind flows around trees that influence fire behavior. To ensure the wind field conditions properly coupled with the forest, we initialized the wind field by establishing a wind speed of 6 m per second at an elevation of 25 m. This condition was maintained for 14 min to stabilize within an US domain, which then provided the wind speed boundary conditions necessary for our fire simulations. Following this stabilization period, each simulation domain sustained the wind condition for an additional 25 s prior to the ignition event to ensure accurate wind flow dynamics. The ignition line was placed 20 m into the domain along the upwind boundary. The ignition line extended 400 m through the domain, stopping 50 m inside the lateral edges of the domain to reduce boundary condition effects. The fires themselves were allowed to develop over a 15 min duration. To compare the fire behavior results we calculated time series of maximum spread (quantified as the furthest downwind point to experience combustion, in m), mass of surface combustion (in kg), and fire intensity (quantified as the amount of energy released per area burned, in W m^{-2}). We also compared domain windspeed profiles as a diagnostic measure of how each domain's forest structure influenced air movement.

3. Results

In general, our methods generated realistic and minimally biased CHMs, which captured less fine-scale detail than the lidar-derived targets but bore a close visual resemblance in which individual trees could be identified by the human eye (figure 5). In the first iteration of training (using data from Blacksmith only,

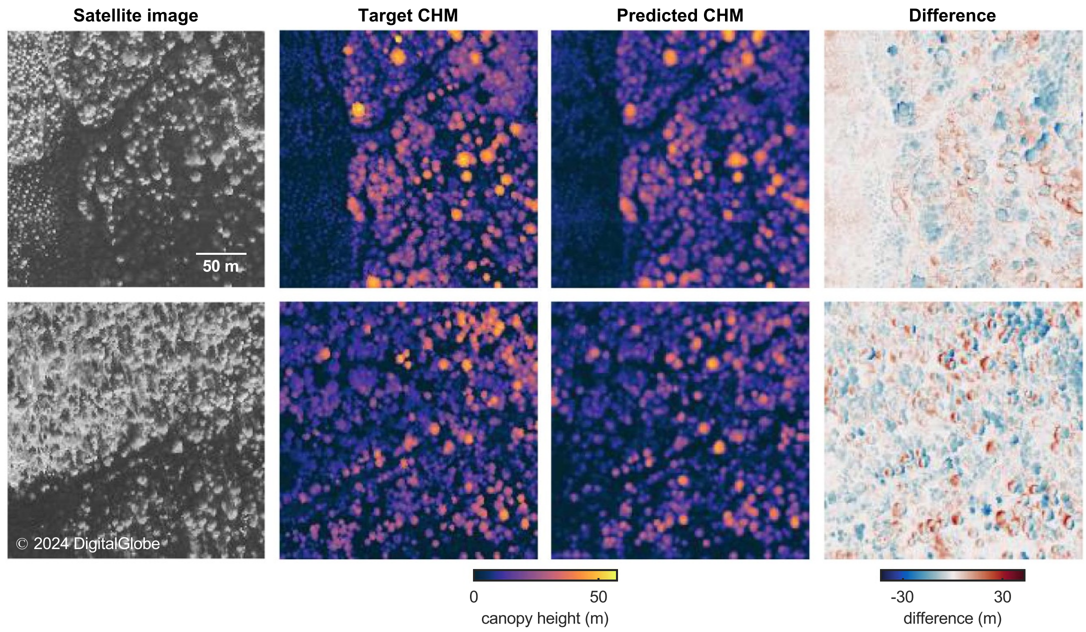

Figure 5. Panchromatic satellite imagery alongside target and predicted CHMs from subsets of the Blacksmith validation dataset (top row) and South Creek test dataset (bottom), after the first iteration of training (using Blacksmith data only).

which were acquired from the months of September through November), the model generated realistic CHMs both in the Blacksmith validation dataset and in test data acquired from South Creek in November 2013. The predicted CHMs were highly inaccurate in the remaining test datasets (South Creek, August 2013, and Shasta Lake, August 2011 and April 2018). After the second iteration of training (in which Blacksmith data were supplemented with smaller volumes of data from South Creek and Shasta Lake), accuracy improved markedly in the test datasets. Moreover, the incorporation into FIRETEC of forest structural information generated by our workflow in plots from Blacksmith and South Creek resulted in fire behavior which closely matched analogous runs informed by airborne lidar data.

3.1. Iteration 1

The first iteration of model training, based on data from Blacksmith only, required 60 h on a single GPU, after which pixel-wise RMSE in the validation dataset reached a minimum of 6.0 m. Tree-level canopy height estimates in the Blacksmith validation dataset correlated strongly with observations from the target dataset, with a mean absolute error (MAE) of 3.5–4.3 m (depending on the date of image acquisition) and a weak negative bias (figure 6). Tree-level height estimates were far less accurate in the four test datasets—with the notable exception of the South Creek dataset acquired in November 2013, where MAE (4.2 m) was comparable to Blacksmith, despite the model not having been exposed to any data from this site during training. Among the five highest performing datasets, 95% of errors in identifying tree height ranged from an underestimation of −13.8 m to an overestimation of 8.2 m. We note that the satellite imagery from this high-performing test dataset and the training data from Blacksmith were all acquired in autumn, and that both Blacksmith and South Creek contain predominantly conifer forests. In South Creek data from August 2013, and in data acquired at Shasta Lake (a site which contains greater concentrations of oak) during April 2011 and August 2018, the model consistently underestimated canopy height.

In general, the delineation of individual trees (or TAOs) within the target CHMs was less accurate than the estimates of canopy height (table 1). Within the Blacksmith validation datasets and the South Creek test dataset from November 2013, the model overestimated the total percent canopy cover within the imagery by 7%–11%, while underestimating the total number of trees by 21%–36%. This implies that individual trees in the predicted CHM were often clumped together into a single object during the automated crown segmentation process, and that the model was somewhat imprecise in delineating gaps in the forest canopy. Accordingly, the predicted CHMs contained significantly fewer gaps than the target CHMs, although median gap area was similar. As the total canopy area was larger and the number of delineated trees was smaller in the predicted CHMs, median predicted tree crown area was significantly higher than in the target dataset.

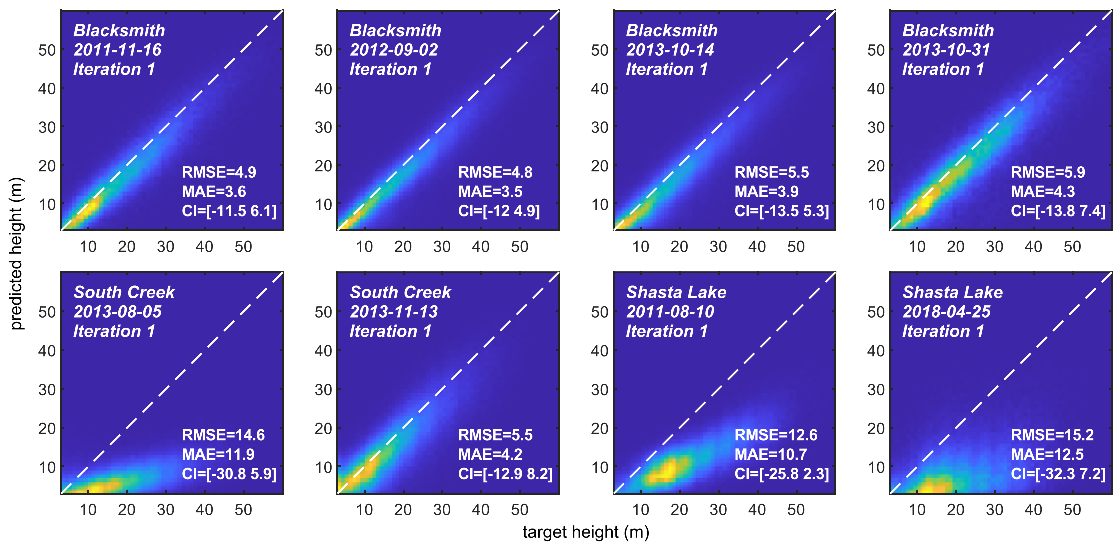

Figure 6. Tree-level comparison of predicted and target canopy height in the Blacksmith validation datasets (top row) and South Creek and Shasta Lake test datasets (bottom row) after the first iteration of training (using data from Blacksmith only). Confidence intervals (CI) refer to the 2.5 and 97.5 percentile errors between target and predicted tree height.

3.2. Iteration 2

The second iteration of model training, in which data from Blacksmith were supplemented with data from South Creek and Shasta Lake, required 72 h, after which pixel-wise RMSE in the validation dataset reached a minimum of 6.1 m. Relative to the first iteration, MAE in the Blacksmith validation dataset increased marginally, while accuracy in the test datasets showed pronounced improvements (figure 7). Within the test datasets, MAE was somewhat greater at the oak-dominated Shasta Lake site (4.4–5 m) than at South Creek (4 m). In both the Blacksmith validation datasets and the South Creek/Shasta Lake test datasets, predicted canopy height showed a slight negative bias relative to the target data. Among all eight datasets, 95% of errors in tree height estimation ranged from an underestimation by 15.5 m to an overestimation by 10 m.

As in the first iteration of training, the model tended to over-estimate the total canopy coverage and the area of individual tree crowns, while underestimating the total number of trees and the median area of gaps in the forest canopy (table 2). The overestimation of canopy coverage remained consistent with the first iteration, at 7%–11%, although the underestimation of tree count increased somewhat, averaging ∼40%. Tree delineation errors were greatest in the complex, oak-dominated forests at Shasta Lake, where the number of trees delineated in the predicted CHMs was roughly half the number extracted from the target data.

3.3. Fire modeling results

At both Blacksmith and South Creek, a visual comparison of FIRETEC results reveals similar behavior in runs incorporating forest structural information derived from ALS and ML, relative to the US runs which assumed a random distribution of tree location (figure 8). Metrics including spread rate, fuel consumption, and fire intensity showed close agreement among the ALS and ML domains, whereas fires in the US domains showed divergent behavior (figure 9). At both sites, fires in the ALS and ML domains spread faster than in the US domains and showed very similar dynamics in fire intensity. The simulated fire behavior was driven by convective wind conditions, which determined spread rates and reflected effective forest structure characteristics which were captured in both the ALS-derived and ML-derived domains. Domain-averaged eastward wind profiles just prior to ignition are shown in figure 10. In all simulations, the forest canopy created drag, which resulted in a zone of relatively slow windspeeds, extending from approximately 8–20 m above the ground surface, overlying a zone of slightly faster windspeeds below the canopy. The canopy and subcanopy windspeeds are the factors most likely to determine fire behavior differences among the simulations. Here we see that the wind profiles in the ALS and ML domains were consistently closer to one another than to the wind profile in the US domain, especially at South Creek. At both test sites, the US domains showed a slower canopy and subcanopy windspeed, which resulted in slower spread rates and differing fire intensity dynamics.

Table 1. Canopy cover, tree approximate object, and gap count statistics from Blacksmith validation datasets and South Creek and Shasta River test datasets after first iteration of training (using data from Blacksmith only).

Site	Date	Canopy Cover (%)		TAO Count		Median Crown Area (m^2)		Gap Count		Median Gap Area (m^2)	
		Target	Pred	Target	Pred	Target	Pred	Target	Pred	Target	Pred
Blacksmith	16 November 2011	68	75	183 595	120 252	19.5	38.75	43 157	12 131	3.5	4
Blacksmith	2 September 2012	49	56	579 425	371 247	18	37.75	108 051	68 972	3.75	6
Blacksmith	14 October 2013	54	63	921 220	583 382	17.5	38	178 260	84 693	3.5	4
Blacksmith	31 October 2013	71	82	180 167	127 172	24	45	45 981	5877	3.5	4
South Creek	5 August 2013	54	40	541 226	316 780	18.75	24.75	88 760	64 130	4.25	6
South Creek	13 November 2013	50	61	583 000	462 153	19	31.75	96 368	60 734	4.25	4
Shasta Lake	10 August 2011	81	81	355 346	253 923	26.75	43.25	75 897	6647	3.5	6
Shasta Lake	25 April 2018	77	53	204 431	133 944	26.25	29.75	41 164	14 824	3.5	4

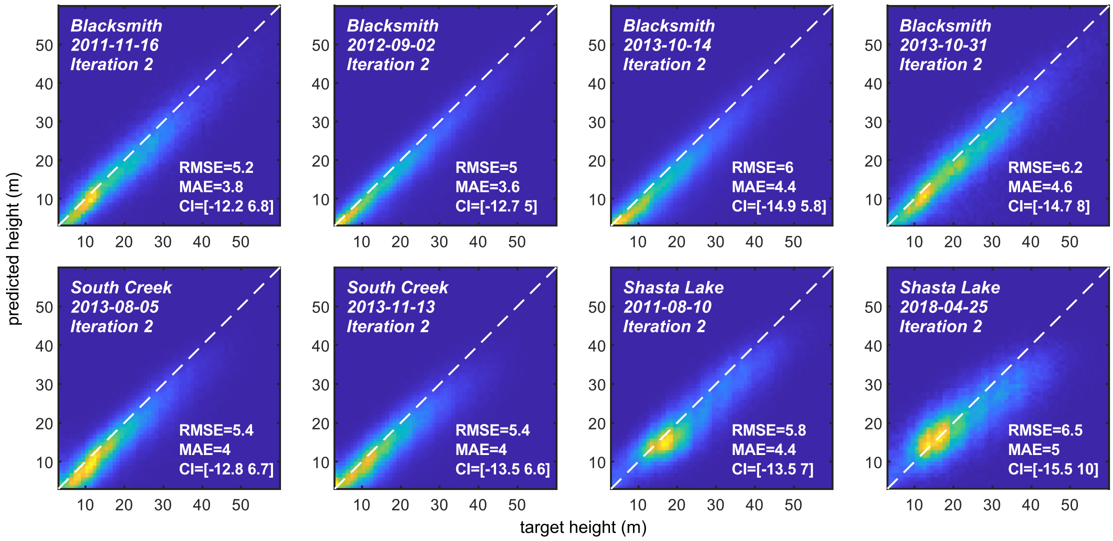

Figure 7. Tree-level comparison of predicted and target canopy height in the Blacksmith validation datasets (top row) and South Creek and Shasta Lake test datasets (bottom row) after the second iteration of training. Confidence intervals (CI) refer to the 2.5 and 97.5 percentile errors between target and predicted tree height.

4. Discussion

4.1. Seasonality and species composition as controls on model accuracy

The future value of our workflow is tied directly to its transferability to forests and images outside the training dataset. The results we present in this manuscript are promising in this regard. In Iteration 1 of training, the model performed poorly in three of the four test datasets, which comprised imagery of forests several hundred kilometers distant from Blacksmith (the site of training and validation). However, in one exception, the South Creek test data from November 2013, performance was comparable to Blacksmith, demonstrating minimal bias in canopy height estimates and a low MAE in estimates of individual tree height (figures 5 and 6). This outcome is encouraging and indicates that phenology and species composition likely play important roles in model transferability. The high performing test dataset from South Creek comprised the only test data to be acquired from a mixed conifer forest in Autumn, matching the conditions that characterize the training and validation data. In contrast, performance was poor in South Creek imagery acquired during summer, and in summer and springtime imagery from Shasta Lake, a site characterized by extensive oak growth. The finding that seasonality plays a role in model performance, even in a predominantly evergreen forest, mirrors conclusions from a large body of prior work on extracting forest attributes from coarser resolution sensors such as Landsat (e.g. Maeda *et al* 2014, Adams *et al* 2020, Chen *et al* 2020). In our present work, the dependence on seasonality may be especially strongly linked to the solar angle, which varied markedly in spring and summertime test images relative to the autumn conditions in the training and validation datasets.

Both the flexibility and the limitations of our workflow were demonstrated in the second iteration of training, where performance in all test datasets improved markedly following the incorporation of modest amounts of test data from Iteration 1 into the training and validation data, but performance at the Blacksmith survey area decreased marginally. In other words, the model became much more generalizable as the diversity of forest types on which it was trained increased, but performance in the mixed conifer forests of Blacksmith declined slightly relative to a model that was trained on that forest type alone. This outcome highlights that curation of training data from forests of diverse species composition and all times of year will be crucial to the improvement of model performance over ever larger spatial and temporal domains. As next steps in improving model robustness, we propose to investigate how clustering techniques can be employed to organize the growing volume of training data into meaningful subsets used to train specialized models. We often think of forests changing gradually over some Euclidean distance, following the first law of geography such that a given forest is most similar to those adjacent. However, considering the appearance of a forest from the perspective of a satellite, different forest types existing in disparate parts of the world can look more similar than the same forest at a different time of day or season. In other words, clustering data geo-spatially or by forest type might be inferior, in terms of model stability and performance, to organizing data by features and patterns in image space. We plan to explore this in more detail in future work.

Table 2. Canopy cover, tree approximate object, and gap count statistics from Blacksmith validation datasets and South Creek and Shasta River test datasets after second iteration of training (using data from all forests).

		Canopy Cover (%)		TAO Count		Median Crown Area (m^2)		Gap Count		Median Gap Area (m^2)	
Site	Date	Target	Pred	Target	Pred	Target	Pred	Target	Pred	Target	Pred
Blacksmith	16 November 2011	68	76	183 595	110 458	19.5	45.25	43 157	11 020	3.5	1.75
Blacksmith	2 September 2012	49	57	579 425	339 287	18	43.75	108 051	140 997	3.75	1.25
Blacksmith	14 October 2013	54	64	921 220	524 478	17.5	45.75	178 260	94 662	3.5	1.25
Blacksmith	31 October 2013	71	82	180 167	118 963	24	50.25	45 981	6921	3.5	2
South Creek	5 August 2013	54	64	541 226	290 893	18.75	49.5	88 760	80 217	4.25	1.75
South Creek	13 November 2013	50	61	583 000	380 173	19	42.25	96 368	37 376	4.25	2
Shasta Lake	10 August 2011	81	88	355 346	174 141	26.75	77	75 897	5453	3.5	1.75
Shasta Lake	25 April 2018	77	86	204 431	95 884	26.25	80.5	41 164	7116	3.5	1

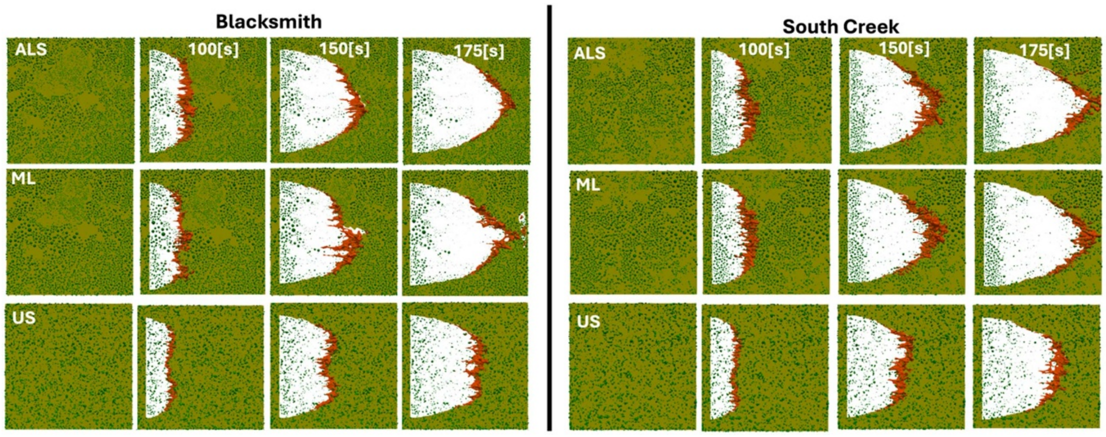

Figure 8. Planar view fire progression visualizations before ignition, 100 s after ignition, 150 s after ignition, and 175 s after ignition.

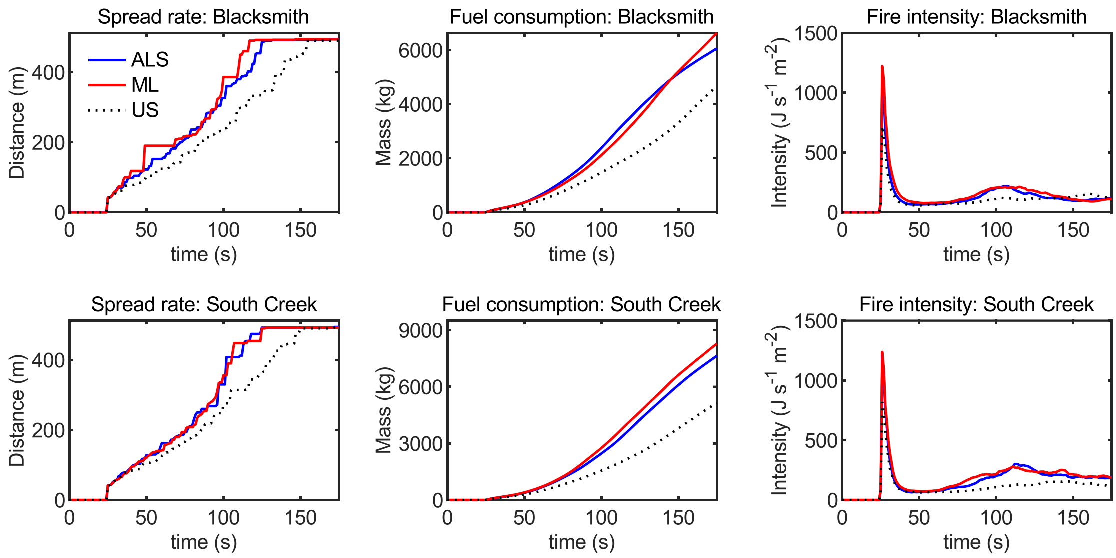

Figure 9. Fire behavior metrics, including spread rate, fuel consumption and fire intensity are plotted for the Blacksmith (top) and South Creek (bottom) domains.

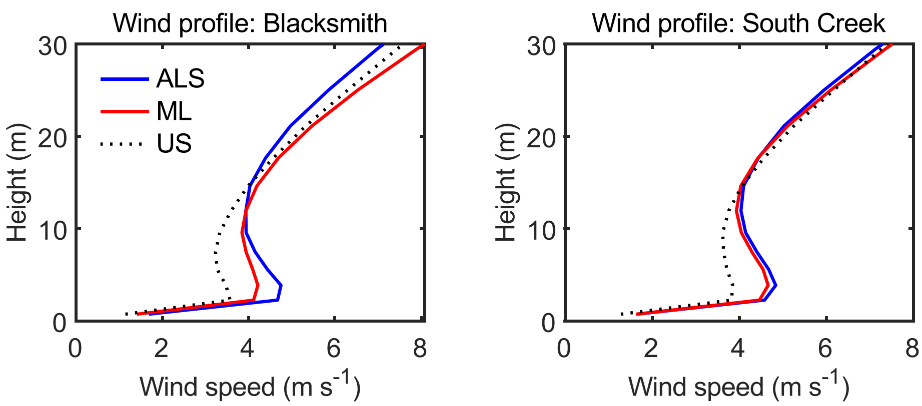

Figure 10. Domain average wind profiles before ignition for Blacksmith (left) and South Creek (right).

Compared with workflow's performance at estimating canopy height, tree delineation was more challenging and less successful, as our segmentation algorithm tended to identify fewer TAOs of larger area in the machine-learning derived CHMs relative to the airborne lidar-derived target data. This problem was especially pronounced at Shasta Lake, where the canopies of neighboring oak trees tended to grow together, increasingly resembling a single object as textural information from the canopy was lost—a problem which has also been demonstrated in other remote sensing modalities (e.g. Jeronimo *et al* 2018). Given that the challenge was most pronounced in deciduous forests, we propose that it might be mitigated in future iterations of our workflow by including a pre-processing step in which a mask of deciduous tree growth is included as input feature, providing the neural network greater context for processing these areas. Such a mask could be produced, for example, by analyzing a leaf-off image from the area of interest, in which conifers and deciduous trees will be easily distinguished (e.g. Wolter *et al* 1995, Khorrami *et al* 2022).

4.2. Comparison of our model with prior work

Among recent efforts to derive CHMs from optical imagery, one of the most ambitious and successful has been the work of Li *et al* (2023), who processed airborne RGB imagery to derive tree-level inventories of canopy height across the entire nation of Denmark. The authors subsequently demonstrated that with relatively modest amounts of training data, they could leverage transfer learning to adapt their workflow to forests in Finland. To our knowledge, while earlier efforts had been made to extract CHMs from optical data (e.g., Lang *et al* 2019, Potapov *et al* 2021, Fayad *et al* 2024), Li *et al*'s study was the first to achieve sufficient spatial resolution to extract individual trees, or objects approaching the granularity of individual trees. A key distinction between Li *et al*'s workflow and ours is the data source. While the off-nadir, panchromatic satellite imagery we operate on presents unique challenges relative to the nadir-view, multiband airborne imagery they employed, it is more abundant in both space and time, which may increase future opportunities to map remote regions or to quickly evaluate changes to forest structure following disturbances such as wildland fires, wind storms, disease and insect outbreaks, and landslides.

In terms of model performance, Li *et al*'s workflow generated somewhat more accurate estimates of tree height, with a global MAE of 2.9 m. In comparison, the results we report here varied in MAE from 3.5–5 m, depending on the forest to which our workflow was applied. We consider this discrepancy in model performance both negligible and expected, given the coarser pixel size and lack of multispectral information in our input imagery. Evaluating the relative skill of our workflows at tree counting is less straightforward than canopy height, as Li *et al* compared the number of trees extracted from their workflow with ground-based measurements, while we compared tree counts from our workflow with counts derived from processing of airborne lidar data. In general, given that the airborne lidar tree counts we used likely underestimated the true number of trees on the ground, it is likely that our workflow resulted in a greater undercount. Nonetheless, one major similarity between our studies was the finding that species composition highly impacted tree counts, as Li *et al* likewise encountered greater challenges in deciduous forests relative to conifers.

Other recent work extending the generation of CHMs from aerial and spaceborne imagery include the efforts of Wagner *et al* (2024), who generated a CHM for the state of California by processing four-channel (RGB + near infrared) NAIP aerial imagery at 0.6 m resolution, and Tolan *et al* (2024), who processed Maxar Vivid2 mosaic imagery to map sub-meter resolution canopy height across the continental United States and several jurisdictions in other nations. While neither of these efforts involved post-processing to extract tree-level statistics, the predicted CHMs of Wagner *et al* were comparable to our own, characterized by a slight negative bias in estimated canopy height and a MAE that varied from 0.05 m in a desert with very low-growing vegetation to 12.02 m in a redwood forest with very tall trees. Visually, the CHMs produced by Wagner *et al* contain sufficient detail to discern tree approximate objects. In contrast, the CHMs produced by Tolan *et al*, which like ours are derived from satellite imagery, lack enough information regarding canopy structure to discern individual trees or TAOs (see figure 7 in Wagner *et al* (2024), for a visual comparison). However, the effort by Tolan *et al* covers a much broader spatial domain, and it also resulted in a comparable mean annual error of 2.6–6.8 m, depending on the validation site.

A central distinction between these contemporary efforts and our own is the deep learning framework. Our workflow employed MS-Net, while Wagner *et al* used a traditional U-Net and Tolan *et al* used transformer networks. Another, perhaps more consequential difference, is the imagery source. Wagner *et al* relied on aerial imagery from a program limited to the United States, which collects new images every two to three years. Tolan *et al* relied on composite product derived from several satellite images, which is available globally, but typically no more frequently than once a year. In contrast, our study is the first to demonstrate proof-of-concept for a workflow that generates sub-meter-resolution CHMs from individual spaceborne sensors, which can capture an area of interest at time intervals as short as several days, increasing the potential to monitor forest structural changes in response to disturbances such as wildland fire.

To summarize these distinctions, several recent studies have advanced machine-learning-based generation of CHMs from high resolution imagery, but ours is the first to achieve sufficient detail to discern individual trees or tree approximate objects while using satellite imagery as a data source. Li *et al* (2023) and

Wagner *et al* (2024) generated CHMs in which TAOs were discernible but used aerial photographs as a data source. Tolan *et al* (2024) processed satellite imagery across a broad spatial domain, but generated CHMs with insufficient structural detail to distinguish tree crowns. Because our workflow has been demonstrated on images captured from a single spaceborne sensor with a repeat interval of days, as opposed to a mosaic product produced using imagery spanning one to three years, it unlocks new potential for high temporal resolution monitoring of changes to forest structure.

We note that one innovation from earlier work which might benefit future iterations of our own is the reliance of Li *et al*'s (2023) workflow on two separate neural networks: the first for canopy height estimation, and a second for crown delineation and tree counting. An advantage to deploying this strategy is that the accuracy of the tree counts was not dependent on the capacity of the first neural network to reproduce the texture of the forest canopy. In some cases, for example, the gridded canopy height product generated by Li *et al*'s first neural network correctly estimated canopy height in forest stands measuring several hundred meters across or larger but yielded insufficient textural detail to discern the boundaries of individual tree crowns. By contrast, our workflow relied on a single neural network to produce a CHM with sufficient detail to be segmented into individual tree crowns using a conventional, watershed-based method. As we consider our workflow's negative bias in tree counts (relative to its mostly negligible bias in tree height estimates) its main area of underperformance, we propose that future iterations may benefit by likewise incorporating a second neural network for crown delineation. This step would have the potential to reduce errors in tree counts and estimates of crown area while taking advantage of a recent burst of research activity into crown boundary mapping from high resolution images.

4.3. Usefulness of our approach for physics-rich fire modeling

The value of forest structural information inferred from our workflow was demonstrated in FIRETEC simulations at Blacksmith and South Creek sites. Despite the differences in tree counts and canopy areas extracted from ML and ALS, we found that the two methods yielded very similar fire behaviors relative to a third scenario in which forest statistics (including tree height and canopy area) were identical to ALS but the placement of individual trees was randomized. The differences in fire behavior between ALS and ML on the one hand, and US runs on the other, reflected the heterogeneous area distribution of canopy gaps and the tendency for similarly sized trees to cluster together in the former. These forest characteristics have been shown in prior work to influence fire behavior (Finney 2001, Knapp and Keeley 2006, Pimont *et al* 2017), especially in numerical investigations using computational fluid dynamics simulations (e.g. Hoffman *et al* 2015, Parsons *et al* 2017, Ziegler *et al* 2017). Relative to the US scenarios, the rapid fire spread across the ML and ALS domains was because the presence of several large gaps in the forest increased wind entrainment below and within the canopy (Pimont *et al* 2011, Atchley *et al* 2021). Here we note that the decreased wind speed within 10 m of the surface in the ML relative to the ALS runs (figure 9) may reflect the tendency of our workflow to slightly over-estimate total canopy coverage (tables 1 and 2). We also acknowledge that, while the incorporation of forest structural information at Blacksmith and South Creek resulted in higher spread rates, this result should not be universally expected, as increased forest detail representation in some cases may slow or even stop fire progression in computational fluid dynamics simulators (Atchley *et al* 2021). Nonetheless, the close matches in fire spread rates, fuel consumption rates, and fire intensity dynamics among the ML and ALS simulations at our test forests demonstrate that our workflow can capture the most effectively important forest structural characteristics that influence wildland fire behavior.

4.4. Limitations of our study

We consider segmentation of the canopy into individual tree crowns or TAOs as our workflow's central area of under-performance, as tree counts from its predicted CHMs were 20%–50% lower than tree-counts from co-located ALS-derived CHMs. The importance of this error varies by application. On one hand, we demonstrated that at present, the incorporation of tree-level forest structural detail derived from our workflow and from ALS results in comparable simulated fire behavior in a physics-rich CFD model. On the other hand, forestry applications such as growth monitoring and species inventories would be more significantly limited by this shortcoming. To improve this issue, one method could be the incorporation of a second neural network tasked with defining tree crown boundaries directly from the input imagery, which was critical to the workflow of Li *et al* (2023). Another limitation is that, at present, our proof-of-concept is relatively limited in geographic scope, focusing on three forests in California. The present results indicate that, as we extend future iterations of our workflow across broader geographies, matching both the species composition and the seasonality of images in the training dataset with those used for inference will be critical to success. Additionally, it is possible that the incorporation of additional features—for example, RGB or infrared bands from the WorldView 2 and 3 satellites—may also improve both model transferability to new forests and the recovery of canopy structure detail for crown segmentation. The primary tradeoffs involved in this method are that, by eliminating the WorldView 1 satellite, reliance on RGB or multispectral imagery would reduce the temporal frequency with which data can be acquired and the historical domain over which it is available. However, future investigation into the extent to which the information in these extra bands improves model performance relative to single-band panchromatic images is warranted.

5. Conclusions

We present a workflow that employs convolutional neural networks to generate CHMs with sufficient resolution to distinguish individual trees using off-nadir, panchromatic satellite imagery as input. Although the resulting CHMs have similar or slightly higher MAE than analogous approaches built to process airborne, multispectral imagery and multi-band satellite imagery, our methods present several distinct advantages as the input data are more abundant. We demonstrate that the species composition and phenology in the training and validation datasets strongly impact model transferability to new forests. Like prior approaches, our workflow tends to underestimate total tree counts, as neighboring trees with complex canopy structures are frequently clumped into tree-approximate objects. Performance in this regard might be improved in the future by: (1) including a mask that distinguishes between conifers and deciduous trees as an input feature, and (2) delegating the task of tree segmentation to a separate neural network. Preliminary investigations using a physics-rich fire simulator demonstrate a high degree of similarity in results from runs incorporating forest structural information derived either from our workflow or directly from airborne lidar surveying. Overall, fire simulations with realistic forest structures derived from either method resulted in faster spread rates than in runs with homogeneously structured forests, due to a more realistic canopy gap size distribution, which influenced wind dynamics. These results underscore the potential of our workflow to improve the realism of fire simulations in forests where high-resolution CHMs are otherwise unavailable. The combination of a mapping approach leveraging broadly available data and cutting-edge physics-based fire modeling offers promise for significantly expanding applications of such advanced fire modeling over a much broader geographic range than has been possible before. Future refinements in both the mapping approaches and ongoing development of advanced fire models will further expand capabilities for assessing potential fire behavior, ecosystem impacts, and risks to communities and firefighters going forward.

Data availability statement

Commercial satellite imagery used in this work was acquired through the EnhancedView Web Hosting Service (https://evhws.digitalglobe.com). Image legacy IDs and metadata are presented in the supplemental information. Samples of the imagery appear in the figures of this manuscript with permission of the private vendor. Code to implement MS-Net is publicly available at https://github.com/je-santos/ms_net.

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

Support for this work included funding through the Laboratory Directed Research and Development program of Los Alamos National Laboratory under Project Numbers 20220024DR, 20230667ER, and 20240448ER.

Conflict of interest

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

ORCID iDs

Charles J Abolt https://orcid.org/0000-0001-6052-199X
Javier E Santos https://orcid.org/0000-0002-2404-3975
Adam L Atchley https://orcid.org/0000-0003-2203-1994
Lucas Wells https://orcid.org/0000-0001-9210-8112
Russell A Parsons https://orcid.org/0000-0001-5091-8993
Rodman R Linn https://orcid.org/0000-0002-5746-4880

References

Adams B, Iverson L, Matthews S, Peters M, Prasad A and Hix D M 2020 Mapping forest composition with landsat time series: an evaluation of seasonal composites and harmonic regression *Remote Sens.* **12** 610

Atchley A L, Linn R, Jonko A, Hoffman C, Hyman J D, Pimont F, Sieg C and Middleton R S 2021 Effects of fuel spatial distribution on wildland fire behavior *Int. J. Wildland Fire* **30** 179–89

Behera S K, Sahu N, Mishra A K, Bargali S S, Behera M D and Tuli R 2017 Aboveground biomass and carbon stock assessment in an Indian tropical deciduous forest and relationship with stand structural attributes *Ecol. Eng.* **99** 513–24

Budei B C, St-Onge B, Hopkinson C and Audet F-A 2018 Identifying the genus or species of individual trees using a three-wavelength airborne lidar system *Remote Sens. Environ.* **204** 632–47

California Forest Observatory 2023 A statewide tree-level forest monitoring system (Salo Sciences, Inc) (available at: https://forestobservatory.com)

Ceccherini G, Duveiller G, Grassi G, Lemoine G, Avitabile V, Pilli R and Cescatti A 2020 Abrupt increase in harvested forest area over Europe after 2015 *Nature* **538** 72–77

Chen D, Loboda T V and Hall J V 2020 A systematic evaluation of influence of selection process on remote sensing-based burn severity indices in North American boreal forest and tundra ecosystems *ISPRS J. Photogramm. Remote Sens.* **159** 63–77

Chen Q *et al* 2006 Isolating individual trees in a savanna woodland using small footprint lidar data *Photogramm. Eng. Remote Sens. Spatial Inf. Sci.* **36** 187–91

Coops N C, Tompaski P, Nijland W, Rickbeil G J M, Nielsen S E, Bater C W and Stadt J J 2016 A forest structure habitat index based on airborne lidar scanning data *Ecol. Indic.* **67** 346–57

Dalagnol R, Wagner F H, Emilio T, Streher A S, Galvão L S, Ometto J P H B and Aragão L E O C 2022 Canopy palm cover across the Brazilian Amazon forests mapped with airborne LiDAR data and deep learning *Remote Sens. Ecol. Conserv.* **8** 301–614

Fayad I *et al* 2024 Hy-TeC: a hybrid vision transformer model for high-resolution and large-scale mapping of canopy height *Remote Sens. Environ.* **302** 113945

Finney M A 2001 Design of regular landscape fuel treatment patterns for modifying fire growth and behavior *For. Sci.* **47** 219–28

Fischer C and Traub B 2019 *Swiss National Forest Inventory: Methods and Models of the Fourth Assessment* (Springer)

Fisher R A *et al* 2018 Vegetation demographics in Earth System Models: a review of progress and priorities *Glob. Change Biol.* **24** 35–54

Frolking S *et al* 2009 Forest disturbance and recovery: a general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure *J. Geophys. Res. Biogeosci.* **114** G00E02

Gray A N 2012 *Forest Inventory and Analysis Database of the United States of America (FIA)*

Hansen M C *et al* 2013 High-resolution global maps of 21st century forest cover change *Science* **342** 850–3

Hoffman C M, Canfield J, Linn R R, Mell W, Sieg C H, Pimont F and Ziegler J 2015 Evaluating crown fire rate of spread predictions from physics-based models *Fire Technol.* **52** 221–37

Jeronimo S M, Kane V R, Churchill D J, McGaughey R J and Franklin J F 2018 Applying LiDAR individual tree detection to management of structurally diverse forest landscapes *J. For.* **116** 336–46

Kauranne T, Pyankov S, Junttila V, Kedrov A, Tarasov A, Kuzmin A, Peuhkurinen J, Villikka M, Vartio V-M and Sirparanta S 2017 Airborne laser scanning based forest inventory: comparison of experimental results for the Perm Region, Russia and prior results from Finland *Forests* **8** 72

Khorrami R A, Naeimi Z, Wing M, Ahani H and Bateni S M 2022 A new multistep approach to identify leaf-off poplar plantations using airborne imagery *J. Geogr. Inf. Syst.* **14** 634–51

Knapp E E and Keeley J E 2006 Heterogeneity in fire severity within early season and late season prescribed burns in a mixed-conifer forest *Int. J. Wildland Fire* **15** 37–45

Lang N, Kalischek N, Armston J, Schindler K, Dubayah R and Wegner J D 2022 Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles *Remote Sens. Environ.* **268** 112760

Lang N, Schindler K and Wegner J D 2019 Country-wide high-resolution vegetation height mapping with Sentinel-2 *Remote Sens. Environ.* **233** 111347

Levis S 2010 Modeling vegetation and land use in models of the Earth System *WIREs Clim. Change* **1** 765–907

Li S *et al* 2023 Deep learning enables image-based tree counting, crown segmentation, and height prediction at national scale *PNAS Nexus* **2** pgad076

Linn R, Anderson K, Winterkamp J, Brooks A, Wotton M, Dupuy J-L, Pimont F and Edminster C 2012 Incorporating fire wind data into FIRETEC simulations of the International Crown Fire Modeling Experiment (ICFME): preliminary lessons learned *Can. J. For. Res.* **42** 879–98

Linn R and Cunningham P 2005 Numerical simulations of grass fires using a coupled atmosphere-fire model: basic fire behavior and dependence on windspeed *J. Geophys. Res. Atmos.* **110** D13

Linn R, Reisner J, Colman J J and Winterkamp J 2002 Studying wildfire behavior using FIRETEC *Int. J. Wildland Fire* **11** 233–46

Linn R, Winterkamp J, Colman J J, Edminster C and Bailey J D 2005 Modeling interactions between fire and atmosphere in discrete element fuel beds *Int. J. Wildland Fire* **14** 37–48

Maeda E, Heiskanen J, Thijs K W and Pellikka P K E 2014 Season-dependence of remote sensing indicators of tree species diversity *Remote Sens. Lett.* **5** 404–12

Marcozzi A A, Johnson J V, Parsons R A, Flanary S J, Seielstad C A and Downs J Z 2023 Application of LiDAR derived fuel cells to wildfire modeling at laboratory scale *Fire* **6** 394

Mell W, Maranghides A, McDermott R and Manzello S L 2009 Numerical simulation and experiments of burning douglas fir trees *Combust. Flame* **156** 2023–41

Moinuddin K A M and Sutherland D 2020 Modelling of tree fires and fires transitioning from the forest floor to the canopy with a physics-based model *Math. Comput. Simul.* **175** 81–95

Oh S, Jung J, Shao G, Shao G, Gallion J and Fei S 2022 High-resolution canopy height model generation and validation using USGS 3DEP LiDAR data in Indiana, USA *Remote Sens.* **14** 935

Parsons R A, Mell W E and McCauley P 2011 Linking 3D spatial models of fuels and fire: effects of spatial heterogeneity on fire behavior *Ecol. Modelling* **222** 679–91

Parsons R *et al* 2018 Modeling thinning effects of fire behavior with STANDFIRE *Ann. For. Sci.* **75** 1–10

Parsons R, Linn R, Pimont F, Hoffman C, Sauer J, Winterkamp J, Sieg C and Jolly W 2017 Numerical investigation of aggregated fuel spatial pattern impacts on fire behavior *Land* **6** 43

Pimont F, Dupuy J-L, Linn R R and Dupont S 2011 Impacts of tree canopy structure on wind flows and fire propagation simulated with FIRETEC *Ann. For. Sci.* **68** 523–30

Pimont F, Dupuy J-L, Linn R R, Parsons R and Martin-stpaul N 2017 Representativeness of wind measurements in fire experiments: lessons learned from large-eddy simulations in a homogeneous forest *Agric. For. Meteorol.* **232** 479–88

Pitkanen J M *et al* 2004 Adaptive methods for individual tree detection on airborne laser based canopy height models *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **36** 187–91

Potapov P *et al* 2021 Mapping global forest canopy height through integration of GEDI and Landsat data *Remote Sens. Environ.* **253** 112165

Santos J E, Yin Y, Jo H, Pan W, Kang Q, Viswanathan H S, Prodanović M, Pyrcz M J and Lubbers N 2021 Computationally efficient multiscale neural networks applied to fluid flow in complex 3D porous media *Transp. Porous Media* **140** 241–72

Sparks A M, Carrao M V and Smith A M 2022 Cross-comparison of individual tree detection methods using low and high pulse density airborne laser scanning data *Remote Sens.* **14** 3480

Tolan J *et al* 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

Tomppo E *et al* 2010 *National Forest Inventories* (Springer)

Varhola A and Coops N C 2013 Estimation of watershed-level distributed forest structure metrics relevant to hydrologic modeling using LiDAR and Landsat *J. Hydrol.* **487** 70–86

Vastaranta M, Wulder M A, White J C, Pekkarinen A, Tuominen S, Ginzler C, Kankare V, Holopainen M, Hyyppä J and Hyyppä H 2013 Airborne laser scanning and digital stereo imagery measures of forest structure: comparative results and implications to forest mapping and inventory update *Can. J. Remote Sens.* **39** 389–95

Wagner F, Roberts S, Ritz A L, Carter G, Dalagnol R, Favrichon S, Hirye M C M, Brandt M, Ciais P and Saatchi S 2024 Sub-meter tree height mapping of California using aerial images and LiDAR-informed U-net model *Remote Sens. Environ.* **305** 114099

Wolter P T *et al* 1995 Using multi-temporal Landsat imagery *Photogramm. Eng. Remote Sens.* **61** 1129–43

Xu Q *et al* 2018 *LiDAR-derived above Ground Biomass and Uncertainty for California Forests, 2005–2014* (ORNL DAAC) (https://doi.org/10.3334/ORNLDAAC/1357)

Zhang C, Zhou Y and Qui F 2015 Individual tree segmentation from LiDAR point clouds for urban forest inventory *Int. J. Remote Sens.* **29** 1339–66

Ziegler J P, Hoffman C, Battaglia M and Mell W 2017 Spatially explicit measurements of forest structure and fire behavior following restoration treatments in dry forests *For. Ecol. Manage.* **386** 1–12