ISPRS Journal of Photogrammetry and Remote Sensing 226 (2025) 187–203

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Deep learning and machine learning enable broad-scale woodland height, cover, and biomass mapping

Michael J. Campbell [a,*], Jessie F. Eastburn [b], Simon C. Brewer [a], Philip E. Dennison [a]

[a] School of Environment, Society & Sustainability, University of Utah, 260 South Central Campus Drive, Room 4625, Salt Lake City, UT 84112, USA
[b] Natural Resource Ecology Laboratory, Colorado State University, 1231 Libbie Coy Way, Fort Collins, CO 80523, USA

ARTICLE INFO

Keywords:
Convolutional neural networks
Random forests
National Agricultural Inventory Program
Airborne lidar
Aboveground biomass
Piñon-juniper woodlands

ABSTRACT

Accurate, spatially explicit quantification of vegetation structure in drylands can improve our understanding of the important role that these critical ecosystems play in the Earth system. In semiarid woodland settings, remote sensing of vegetation structure is challenging due to low tree height, cover, and greenness as well as limited spatial and temporal availability of airborne lidar data. These limitations have hindered the development of remote sensing applications in globally widespread and ecologically important dryland systems. In this study, we implement a U-Net convolutional neural network capable of predicting per-pixel, lidar-derived vegetation height in piñon-juniper woodlands using widely available, high-resolution aerial imagery. We used this imagery and modeled canopy height data to construct random forest models for predicting tree density, canopy cover, and live aboveground biomass. Trained and validated on a field dataset that spanned diverse portions of the vast range of piñon-juniper woodlands in the southwestern US, our models demonstrated high performance according to both variance explained ($R^2_{density} = 0.45$; $R^2_{cover} = 0.80$; $R^2_{biomass} = 0.61$) and predictive error (%$RMSE_{density} = 57$; %$RMSE_{cover} = 19$; %$RMSE_{biomass} = 42$). A comparative analysis revealed that, while performance was somewhat lower than models driven solely by airborne lidar, they vastly exceeded that of models driven by aerial imagery alone or a combination of Landsat, topography, and climate data. Although the structural predictive maps featured some artifacts from illumination and perspective differences inherent to aerial imagery, this workflow represents a viable pathway for spatially exhaustive and temporally consistent vegetation structure mapping in piñon-juniper and other dry woodland ecosystems.

1. Introduction

Mapping vegetation structure over space and time is critical to understanding the important roles that terrestrial ecosystems play in the Earth system (Harris et al., 2021; Pettorelli et al., 2018; Senf, 2022). Remote sensing plays a key role in this process, providing a robust, objective, and diverse array of datasets from which vegetation structure can be derived through various analytical means (Mitchell et al., 2017). However, tradeoffs in remote sensing platforms and sensors have prevented the development of a singular, one-size-fits-all solution for the spatially explicit estimation of vegetation structure (Dalla Mura et al., 2015; De Sy et al., 2012; Lefsky and Cohen, 2003). For example, data collected from airborne platforms often provide higher spatial resolution than data collected from satellites. Conversely, satellite data typically supply higher temporal resolution and product consistency. To further complicate matters, the types and importance of tradeoffs vary by ecosystem. For example, in a tropical forest, satellite optical imagery may be unable to accurately map vegetation structure due to persistent cloud cover and saturation of the relationship between vegetation indices and biomass (Hilker et al., 2012; Mutanga et al., 2023). In these ecosystems, airborne or satellite lidar may provide the best solution for mapping vegetation structure (Asner et al., 2012; Dubayah et al., 2022). Conversely, in a temperate grassland, satellite-based proxies of vegetation greenness may exceed lidar's ability to map cover or biomass, given the challenges associated with distinguishing grasses from the ground surface in a point cloud or waveform (Reinermann et al., 2020; Roberts et al., 2019; Shendryk, 2022). Thus, there remains a need to examine the relative strengths and weaknesses of different remote sensing-based approaches to vegetation structure mapping in an ecosystem-specific manner.

One class of ecosystems that warrants a dedicated and thoughtful investigation of remote sensing tradeoffs is that of dry woodland

* Corresponding author.
E-mail address: mickey.campbell@ess.utah.edu (M.J. Campbell).

https://doi.org/10.1016/j.isprsjprs.2025.05.016
Received 7 February 2025; Received in revised form 21 April 2025; Accepted 15 May 2025
Available online 23 May 2025

environments (Campos et al., 2018; David et al., 2022; Yang et al., 2012). Dryland ecosystems have received increasing attention in recent years for the critical role that they play in the global carbon cycle (Hanan et al., 2021; Lal, 2019; Reed et al., 2025). Though lower in biomass than their more mesic counterparts, dryland ecosystems are highly responsive to changes in climate and, as a result, play a substantive role in driving interannual variability in global carbon uptake (Ahlström et al., 2015; Gherardi and Sala, 2019; Poulter et al., 2014). Furthermore, with projected expansion to their extent, drylands are poised to comprise a growing portion of the global landmass, enhancing their ecological importance in a changing climate (Huang et al., 2016). The largest dry woodland ecosystem in the United States is piñon-juniper (PJ) woodlands (Miller et al., 2019). Spanning 10 states and featuring a geographic range larger than one quarter of the contiguous US, PJ woodlands have a complex ecological and management history (Romme et al., 2009). Mapping vegetation structure in PJ and other similar woodlands comes with a novel set of remote sensing challenges due to the vegetation's relatively short stature, low cover, and low photosynthetic activity (Smith et al., 2019). Thus, remote sensing approaches suitable for other tree-dominated ecosystems may not be appropriate in these semiarid woodlands.

Ideally, global multispectral satellite imaging (e.g., Landsat, Sentinel-2) would provide the foundation of vegetation structure mapping, given their global reach, forest stand-scale spatial resolution, and high temporal resolution. These sensors have shown some promise in dry woodland settings, with varying degrees of success mapping tree structural variables such as tree density (e.g., Humagain et al., 2017), canopy cover (e.g., Reinhardt et al., 2020), and biomass (e.g., Gizachew et al., 2016). However, the accuracy of models derived from satellite imagery alone tends to be considerably lower than woodland structural models driven by airborne lidar data, given lidar's capacity to characterize structure in three dimensions (Campbell et al., 2023; Krofcheck et al., 2016; Sankey et al., 2013; Wu et al., 2016). The primary limitation to operational vegetation mapping using airborne lidar is its lack of spatiotemporal availability. This could potentially be overcome with satellite lidar programs such as the Global Ecosystem Dynamics Investigation (GEDI; Dubayah et al., 2020); however, recent evidence has pointed towards limited success of utilizing GEDI data in short-stature, low-cover woodlands (Campbell et al., 2024; Li et al., 2023).

Several studies have demonstrated the potential utility of high-resolution aerial or satellite imagery for mapping woodland structure (e.g., Davies et al., 2010; Hulet et al., 2014). Textural information contained within the imagery and the ability to distinguish tree canopies from background materials have facilitated the accurate prediction of tree density, canopy cover, and biomass across forest types (Basu et al., 2015; Falkowski et al., 2017; Hogland et al., 2018). Although aerial imagery suffers some of the same spatiotemporal limitations as opportunistically collected airborne lidar, as well as challenges associated with spectral and radiometric resolutions (Maxwell et al., 2017), operational programs such as the US National Agricultural Inventory Program (NAIP) may offer a contiguous US-wide basis for vegetation structure mapping. Increasing availability of high-resolution satellite imagery from commercial vendors also offers an additional opportunity (Ahmad et al., 2021). However, in the absence of three-dimensional structural data, high-resolution imagery may still fall short of airborne lidar's accuracy in woodland ecosystems.

The relatively recent proliferation of convolutional neural networks (CNNs) in the realm of ecological remote sensing may enable the development of a vegetation structure mapping workflow that simultaneously leverages the benefits of high-resolution imagery and airborne lidar (Kattenborn et al., 2021). Recent efforts have demonstrated the successful, per-pixel estimation of canopy height using CNNs and other deep learning models trained with height data from lidar and satellite or aerial imagery (Chang et al., 2025; Malambo and Popescu, 2023; Tolan et al., 2024; Wagner et al., 2024). Applying this modeling framework to NAIP imagery in a dry woodland setting could potentially yield accurate and timely predictions of vegetation structure in these critical ecosystems.

The objectives of this research were as follows:

1. To introduce and describe a robust analytical framework for mapping vegetation structure in an ecologically diverse and geographically expansive dry woodland ecosystem.
2. To test the extent to which a U-Net CNN, trained on lidar-derived canopy height (response variable) and high-resolution aerial imagery (predictor data) can accurately predict vegetation height in this woodland environment.
3. To evaluate the degree to which tree density, canopy cover, and live aboveground biomass can be accurately predicted using a suite of imagery-derived predictor variables (including predicted canopy height) using a random forest model.
4. To compare our new modeling framework to vegetation structure random forest models driven by alternate sets of predictor variables, including: (1) high-resolution aerial imagery without predicted canopy height data; (2) moderate resolution satellite imagery, topography, and climate data; and (3) airborne lidar data.

2. Methods

2.1. Methods Summary

We strategically selected 100 sites, each 3x3km, to capture ecological variability throughout the range of PJ woodlands in the Western US. Within each site, we acquired NAIP imagery and airborne lidar data from the United States Geological Survey (USGS). We split the sites into training (n = 60), validation (n = 20), and testing (n = 20) groups. We trained and validated a U-Net CNN model to predict lidar-derived canopy height using NAIP image data at 0.6 m spatial resolution, evaluating the performance using the test data. We used the CNN and NAIP imagery to generate canopy height models (CHMs) for 18 sites containing a total of 180 field plots, each 30x30m, within which we collected tree structure data. We used random forests to model tree density, canopy cover, and live aboveground biomass using a suite of NAIP- and CHM-derived predictor variables, with performance evaluated using a leave one site out cross-validation procedure. We tested the extent to which a quantile-based bias correction procedure could improve the ability to predict extreme values. To assess the comparative strengths and weaknesses of these tree structural predictive models, we generated three other sets of models: (1) using NAIP data alone (i.e., without the CNN-derived CHM); (2) using Landsat, topography, and climate data; and (3) using airborne lidar data. Finally, we used our NAIP + CHM random forest model to map biomass throughout our 18 field sites and evaluated the spatial dimensions of the resulting maps.

2.2. Study area

Our study spanned the extent of PJ woodlands in the Western US (Fig. 1). Our analysis focused on two different sets of sites. The larger set (n = 100), labeled "CNN […] Sites" in Fig. 1, were used as the basis of training, validating, and testing a U-Net CNN model designed to predict canopy height using high-resolution aerial imagery. These sites were selected using a conditioned Latin hypercube sampling algorithm, which ensures that variability among a defined set of input variables is captured in a resultant set of sample points (Minasny and McBratney, 2006). For this study, those input variables included elevation and climate data from PRISM (to capture environmental variability; Daly et al. (1997); PRISM Climate Group, Oregon State University (2019)), Landsat 8 surface reflectance data (to capture spectral variability), and latitude and longitude (to capture spatial variability). The sampling area was restricted to the intersection of areas mapped as PJ by 2023 LANDFIRE Existing Vegetation Type (LANDFIRE, 2023) and areas that had USGS 3D Elevation Program quality level 2 or better airborne lidar

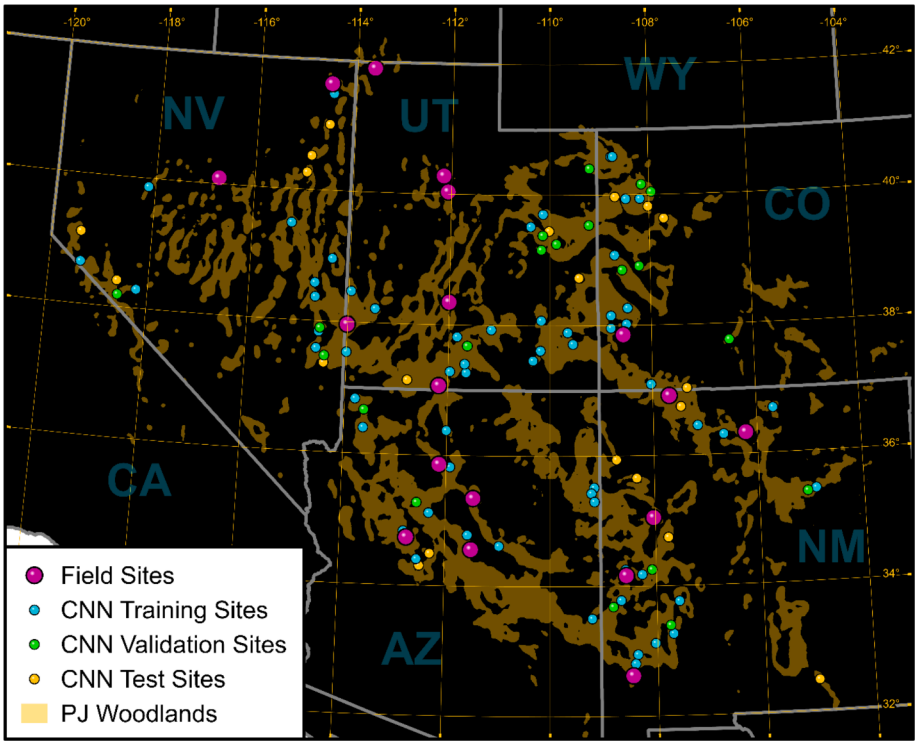

Fig. 1. Study area map highlighting the field sites and sample sites within the context of PJ woodlands in the Western United States. The field sites represent 18 sites within which we collected vegetation structure field data and the other sites represent 100 sites which were used to train, validate, and test a convolutional neural network (CNN) built to map vegetation height. The PJ woodland extent was derived from LANDFIRE Existing Vegetation Type. For display purposes, we applied a circular focal filter with a 5 km radius, thresholding the resulting data to only pixels containing at least 25 % PJ woodlands within their focal area.

data collected since 2013. The CNN sample sites were 3x3km each in size, and did not overlap spatially. The average distance between each site and its nearest neighbor was approximately 35 km.

The smaller set (n = 18) of sites, labeled "Field Sites" in Fig. 1, represent a series of variably sized areas (33 – 838 km^2), each of which contains between 7 and 12 field plots (180 in total) collected as part of a dedicated campaign to capture PJ woodland tree structure in 2021–2022 (Campbell et al., 2024). The site extents are defined by a bounding box around a 1 km buffer around each site's plots. As with the CNN sample sites, these sites were selected to capture ecological variability within the vast PJ range, though this was largely done on an opportunistic basis, factoring in additional considerations such as public land and road access. The field plots were used to train and validate random forest-driven models, which were applied to the predictive mapping of vegetation structure throughout each of the 18 sites.

2.3. Data Acquisition and Preprocessing

2.3.1. Field data

In the summers of 2021 and 2022, we collected 180 plots' worth of PJ tree structural data, distributed throughout 18 sites in the Western US (Fig. 1). Each plot was approximately 30x30m in size, and the stem diameters of all trees greater than 0.5 m in height were measured. Stem diameters were used to estimate per-tree aboveground biomass through allometry (Chojnacky et al., 2014). Tree-level biomass was aggregated to plot-level estimates of biomass density (in megagrams per hectare, or Mg/ha), which served as one of the three focal tree structural measures in this study. A second structural measure of interest we derived from our field plot data was tree density, measured in trees per hectare (trees/ha). The third structural measure of interest – canopy cover – was computed within field plots, but was not measured in the field. For this, we used airborne lidar, as it enables a more direct and spatially relevant measure of cover than field-based methods can provide. Canopy cover was calculated as the number of first return points equal to or greater than 1 m in aboveground height divided by the total number of first returns. More details on lidar data processing are provided in Section 2.3.3.

2.3.2. NAIP data

We acquired NAIP high spatial resolution aerial image data from the USGS EarthExplorer data portal within both the 100 CNN sample sites and the 18 field sites. These orthorectified and mosaicked images are delivered as digital ortho quarter quad tiles, that each span 3.75′ x 3.75′ plus a 300 m buffer. Spatial resolutions and collection dates vary by state, but NAIP imagery are collected typically every two years at a spatial resolution of either 0.3 m or 0.6 m. NAIP imagery has four spectral bands: red, green, blue, and near-infrared. Bands are provided as digital number (DN) values and are not corrected to surface reflectance.

For the 100 CNN sample sites, we used the most recent data available, as we considered it to be of the highest quality. The resulting imagery was collected between 2021 and 2023. For the 18 field sites, we wanted to ensure that the imagery captured pre-field measurement conditions, so we used imagery acquired from 2018 to 2020. For consistency, 0.3 m images were mean-aggregated to 0.6 m. The images were mosaicked and clipped to site extents. The mosaics were then subset into a series of 256x256 pixel image chips that overlapped by 16 pixels in both x and y dimensions, as these image chips would serve as the basis of training, validating, and testing a CNN model, which requires consistently-sized and relatively small input images. This yielded a total of 400 image chips for each of the 3x3km CNN sample sites, and a variable number of image chips for each of the field sites ranging from 1,598 to 42,406.

2.3.3. Lidar data

We acquired airborne lidar data from the USGS 3D Elevation

Program within the extent of each CNN sample site and field site (Snyder, 2012). Unlike NAIP, which is acquired programmatically on a statewide basis, lidar data from the 3D Elevation Program are compiled from individual projects through partnerships with various organizations. Within our 100 CNN sample sites, we used data from 25 different projects, whose collection dates ranged from 2013 to 08-24 to 2023–05-02, with a mean date of 2020–02-25. The pulse densities in these data ranged from 2.62 to 31.11 pulses/m^2 (mean = 7.18 pulses/m^2). Within our 18 field sites, we used data from 18 different projects, whose collection dates ranged from 2013 to 08-24 to 2021–11-17, with a mean date of 2019–04-23. The pulse densities ranged from 3.33 to 19.78 pulses/m^2 (mean = 8.16 pulses/m^2).

All lidar data processing was done in *R* using the *lidR* library (R Core Team, 2021; Roussel et al., 2020). Across all sites, points that were classified as noise or flagged as withheld were removed. Among the remaining points in each cloud, raw elevations were normalized to aboveground heights through subtraction of a terrain model interpolated using a triangulated irregular network. For the 100 CNN sample sites, 0.6 m-resolution CHMs were interpolated using the pit-free algorithm of Khosravipour et al. (2014), matching the resolution of the NAIP data. These CHMs were subset into the same overlapping 256x256 pixel extents used to create NAIP chips, as they would represent the per-pixel response variable for CNN training, validation, and testing purposes.

2.3.4. Landsat, Topography, and climate data

We acquired and generated a suite of 30 m spatial resolution spectral, topographic, and climatic datasets that we used to build a comparison vegetation structural predictive model (described later in Section 2.4.2.3) within our 18 field sites. We used Google Earth Engine to generate a series of 2020 seasonal cloud-free, cloud shadow-free, and snow-free surface reflectance composites of Landsat 8 OLI imagery (Gorelick et al., 2017). We used Collection 2 Tier 1 Level 2 surface reflectance to create "spring" (March-May), "summer" (June-August), and "autumn" (September-November) median reflectance composites. For each of these composites, we generated a variety of spectral indices (Table A1). We also acquired 30 m-resolution digital elevation models from the USGS within our field sites, which would serve as the basis for generating a suite of topographic derivative predictor variables (Table A2). The elevation models also served as inputs for downscaling climate data using ClimateNA (Wang et al., 2016). We generated 30-year (1991 – 2020) annual and seasonal averages for a diverse array of the climatic variables provided by ClimateNA (Table A3).

2.4. Modeling

2.4.1. Convolutional neural network

To build a model capable of predicting per-pixel canopy height at a high spatial resolution, we constructed a U-Net CNN (Ronneberger et al., 2015), as implemented in Python using Keras and TensorFlow version 2.7 (Fig. 2). We selected the U-Net framework for its widely regarded ability to generate per-pixel predictions at the original input resolution of the image data, while incorporating higher-level spatial patterns derived from convolutional filters applied at various scales in the learning process. We randomly split our 100 3x3km sample sites into training (n = 60), validation (n = 20), and test (n = 20). Each site had 400 256x256 chips of 4-band NAIP image data, which served as the input to the U-Net, and CHM data, which served as the response variable. We trained the network from scratch without the use of any pre-trained weights. In total, the U-Net was trained on 24,000 images, validated with 8,000 images, and tested on 8,000 images. The model architecture can be seen in Fig. 2. Our U-Net features an encoder (down-sampling) path with four convolutional blocks, each of which features two 3x3 convolutional layers, padded to ensure retention of the input spatial dimensions, and activated using a rectified linear unit (ReLu) function. The output layer from each block is down-sampled using 2x2 max pooling to capture higher-level shape, texture, and contextual information. After each pooling process, dropout is applied to reduce overfitting, the remainder of which is fed into a new convolution block with increasing numbers of convolution filters. The U-Net then proceeds through a decoder (up-sampling) pathway with four 2x2 deconvolution (or transposed convolution) blocks, with dropout between each block. At each level, the features from the corresponding level in the encoder path are merged with the up-sampled features, which preserves the relative spatial location of information from the encoding process. These merged layers are then fed through two 3x3 convolution layers with decreasing numbers of filters. Lastly, a 1x1 convolution layer with a

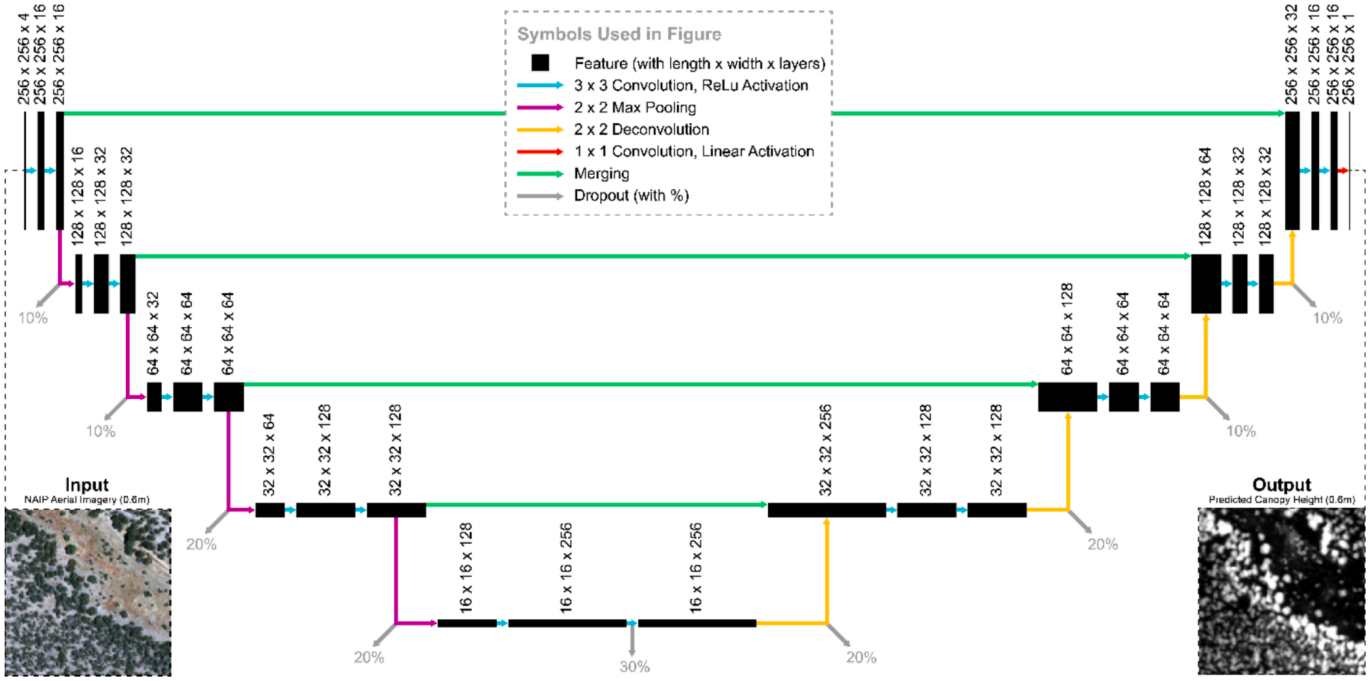

Fig. 2. U-Net convolutional neural network architecture used in this study.

linear activation function yields a predictive map of canopy height with the original extent and resolution of the input NAIP image.

The model was trained with an Adam optimizer (learning rate = 0.0001) and validated with a batch size of 64 images for 100 epochs. Training took approximately 2000 s per epoch on a PowerEdge R740 Server with 384 GB of RAM and an NVIDIA Tesla T4 GPU. Training and validation loss, as measured by mean squared error (MSE) in predicted canopy height, were monitored for each epoch, and the epoch with the lowest validation MSE was saved as the final model. This model was used to predict canopy height within the 8,000 test images. The output CHM chips were mosaicked to create site-wide CHMs using ArcGIS's Mosaic To New Raster tool, with the overlap function set to "blend", which determines output pixel values with a weighted mean of inputs according to their overlap position. This approach is designed to minimize edge effects. A random sample of 50,000 points was generated within each of the 20 test sample sites, at which CNN-modeled (i.e., predicted) CHM and lidar-derived (i.e., observed) CHM pixel values were extracted. Given the high spatial resolution of these maps (0.6 m), and the potential for small positional offsets between NAIP and lidar data to negatively impact a direct performance assessment, we also sought to understand how CNN performance varied over larger spatial extents than the individual pixel (Duncanson et al., 2025). Accordingly, for each of the 1 M random sample points, we also extracted mean predicted and observed CHM values within circular buffer areas around each point. We tested buffers of 1 m, 2 m, 4 m, 8 m, and 16 m, the last of which is of comparable scale to the 30 m pixel resolution of the stand-scale vegetation structural modeling conducted in this study (Section 2.4.2). Predicted versus observed canopy height was evaluated through linear regression and quantified by R^2 (Eq. (1), %RMSE (Eq. (2), and % bias (Eq. (3), defined below:

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

$$\%RMSE = 100 \times \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \Big/ \bar{y} \tag{2}$$

$$\%bias = 100 \times \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \Big/ \bar{y} \tag{3}$$

In Eqs. 1–3 above, y_i is the observed value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of all observed values for each data point in $[i \cdots n]$. We opted to use the normalized %RMSE and %bias in lieu of their raw counterparts to facilitate response variable- and ecosystem-independent comparison.

2.4.2. Vegetation structure modeling

2.4.2.1. Primary Models: NAIP + CHM. To build models capable of predicting tree density, canopy cover, and live aboveground biomass, we needed to generate a suite of predictor variables that might correlate to these biometrics. Our primary models relied on NAIP imagery and our CNN-generated CHM to derive these predictors. Because the data are collected at different times of day, month, and year, and because those time frames vary between neighboring states, NAIP imagery can feature spectral artifacts that manifest as tile boundaries between imagery collected at different times (Hogland et al., 2018). It is impossible to remove these artifacts entirely, but to minimize their impact, we opted to use normalized difference indices, rather than raw band values, in the generation of NAIP-based predictor variables (Maxwell et al., 2017). We computed every non-redundant set of normalized differences that the four-band data offered (Table 1).

To correspond with the spatial dimensions of our field plots (30x30m), we needed a set of predictor layers that had a 30 m spatial resolution. To that end, we needed to aggregate the 0.6 m NAIP imagery to 30 m. For each of the normalized difference indices in Table 1, we computed the following descriptive statistics among the 250 0.6x0.6 m pixels that comprised each 30x30m super-pixel in the output predictor layer: mean, standard deviation, coefficient of variation, quantiles (5th, 10th, 25th, 50th, 75th, 90th, 95th), skewness, and kurtosis. In all, 12 predictor variables were generated for each of the six normalized difference image datasets, yielding a set of 72 NAIP-based predictor variables, designed to capture the spectral characteristics of vegetation structure.

Table 1
Normalized difference spectral indices derived from National Agricultural Inventory Program (NAIP) imagery that served as the basis of generating vegetation structural predictor variables. Note that for NAIP data, band 1 represents reflectance in the red portion of the spectrum, band 2 is green, band 3 is blue, and band 4 is near infrared (NIR).

Name	Abbreviation	Formula
Normalized Difference of Green and Red	NDGR	$(b2-b1)/(b2+b1)$
Normalized Difference of Blue and Red	NDBR	$(b3-b1)/(b3+b1)$
Normalized Difference of NIR and Red	NDNR	$(b4-b1)/(b4+b1)$
Normalized Difference of Blue and Green	NDBG	$(b3-b2)/(b3+b2)$
Normalized Difference of NIR and Green	NDNG	$(b4-b2)/(b4+b2)$
Normalized Difference of NIR and Blue	NDNB	$(b4-b3)/(b4+b3)$

To create an additional set of predictors that captured the three-dimensional characteristics of vegetation structure, we applied all of the same descriptive statistics to the CNN-derived CHMs. Given the rich structural information contained within CHMs, we opted to derive three additional sets of predictors. The first was aimed at capturing tree density. To do this, we used the local maximum filter function *lmf()* within the *lidR* library in *R* (Jean-Romain, 2023; R Core Team, 2021). It uses an adaptive focal filter to efficiently locate high points in a canopy height model. We thresholded the resulting points at four different heights (0.5, 1.0, 2.0, and 4.0 m) and calculated their densities, normalized by 30x30m pixel area. The second was aimed at capturing canopy cover. To do this, we used the same four height thresholds and calculated the proportion of 0.6 m pixels equal to or greater than each height, normalized by the total number of 0.6 m pixels in a 30x30m area. The third was aimed at capturing canopy volume. To do this, we once again used the four height thresholds; however, in this case, we multiplied the canopy heights of all pixels equal to or greater than those thresholds by their 0.6x0.6 m pixel area, summed them and normalized them by 30x30m pixel area. In all, this produced 12 descriptive statistical predictors, four density predictors, four cover predictors, and four volume predictors from the CHM.

The 72 NAIP and 24 CHM predictor variables were used to model tree density, canopy cover, and biomass within the 180 field plots using the following random forest-driven model framework. For each of the three response variables, the 96 candidate predictors were fed through the variable selection using random forests (VSURF) algorithm to yield a concise, non-redundant, and meaningful set of predictor variables (Genuer et al., 2015). The distilled data were then used to tune random forest's hyperparameters (number of variables tested at each decision tree split, minimum decision tree node size, and sample fraction used for bagging). This was done with *tuneRanger*, which uses model-based optimization to efficiently find a set of hyperparameters that minimize predictive error (Probst et al., 2019). Models were built and validated using a leave-one-site-out cross-validation procedure, which minimizes the effects of spatial autocorrelation on apparent model performance. To do so, plot data from each of the 18 field sites were iteratively set aside for validation, while the data from the other 17 sites were used to train the model. Predictions were made on the plot data that were left out. After all 18 iterations, the combined predictions were compared to their observed values, and performance was evaluated using a linear regression, R^2, %RMSE, and %bias between predictions and observations. Variable importance was assessed on each of the 18 iterations according

to two different measures: (1) permutation importance, as measure by the increase in %RMSE that a model yields when a predictor variable's values are randomly permuted; and (2) mean absolute Shapley additive explanations (SHAP) values, which quantify individual predictor variables' contribution to a model's prediction. For the former, we used the *ranger* library's permutation importance calculation. For the latter, we used the *R* package *treeshap* (Komisarczyk et al., 2024). For each we computed the mean and standard deviation among the 18 iterations, which yielded useful information on which predictor variables were found to be most valuable.

Given the propensity of random forests, and other ensemble machine learning algorithms, to yield relatively moderate predictions that underestimate high values and overestimate low values, we sought to determine whether a bias correction procedure could improve the prediction of extreme values (Zhang and Lu, 2012). To that end, we employed a quantile-driven approach that creates a correction factor for predictions driven by the differences in empirical distributions between out-of-bag predictions and training observations (Belitz and Stackelberg, 2021). The method is illustrated in Fig. 3.

2.4.2.2. Comparison models #1: NAIP only. In comparison to more traditional machine learning approaches, building a U-Net CNN model is not a trivial task; it requires algorithmic and programmatic familiarity, a lot of training and validation data, and a capable computer ideally equipped with a powerful graphics processing unit. In recognition of this potential limitation, we sought to understand the value of the CNN-derived CHM in the prediction of vegetation structure. In other words, if NAIP imagery alone can yield models with comparable accuracy, then perhaps it would not be worthwhile to build and apply a canopy height model U-Net. To that end, we replicated the exact modeling and assessment workflow described in Section 2.4.2.1, but simply omitted the CHM-derived predictor variables. For each of the three vegetation structural parameters of interest (density, cover, and biomass), we compared R^2 and %RMSE values between these models and the primary models to quantify performance differences.

2.4.2.3. Comparison models #2: Landsat + Topography + Climate. NAIP data are collected nationally in the contiguous US but, as mentioned previously, they are only available approximately every two years and can feature spatial and spectral artifacts due to imaging geometry and illumination conditions. Furthermore, their high spatial resolution creates data processing and storage challenges. We created a second comparison model that uses solely 30 m resolution Landsat 8 OLI imagery, topography data, and downscaled climate data (Section 2.3.4). In all, there were 155 Landsat, topographic, and climatic predictor variables (Tables A1-A3), all of which were fed through the same random forest modeling and assessment framework applied to the primary models. For each of the three vegetation structural parameters of interest (density, cover, and biomass), we compared R^2 and %RMSE values between these models and the primary models to quantify performance differences.

2.4.2.4. Comparison models #3: Airborne lidar. Although the intent of our CNN model was to offer the improved ability to predict vegetation structure using the resulting CHMs' height information, that improvement may still be limited in comparison to lidar point cloud data for two reasons. The first is that the CHM is a modeled product, and thus contains inherent error. The second is that a CHM merely captures the height of the tallest vegetation in a given pixel, whereas a point cloud contains discrete height measurements throughout the vertical canopy profile. Thus, it is likely that airborne lidar's ability to map vegetation structure remains superior to our NAIP + CHM model. To evaluate how close to lidar's benchmark performance our NAIP + CHM model gets, we generated a third set of comparison models driven solely by airborne lidar data. We generated two different sets of predictor variables from airborne lidar data: (1) CHM-derived variables; and (2) point cloud-derived variables. The former set mirrored the predictor variables generated from the CNN-modeled CHMs, for direct comparative purposes. The latter was derived using the *metrics_set3()* function within the *lidRmetrics* library in *R* (R Core Team, 2021; Tompalski, 2023). In all, there were 121 lidar predictor variables, all of which were fed through the same random forest modeling and assessment framework applied to the primary models. For each of the three vegetation structural parameters of interest (density, cover, and biomass), we compared R^2 and % RMSE values between these models and the primary models to quantify performance differences.

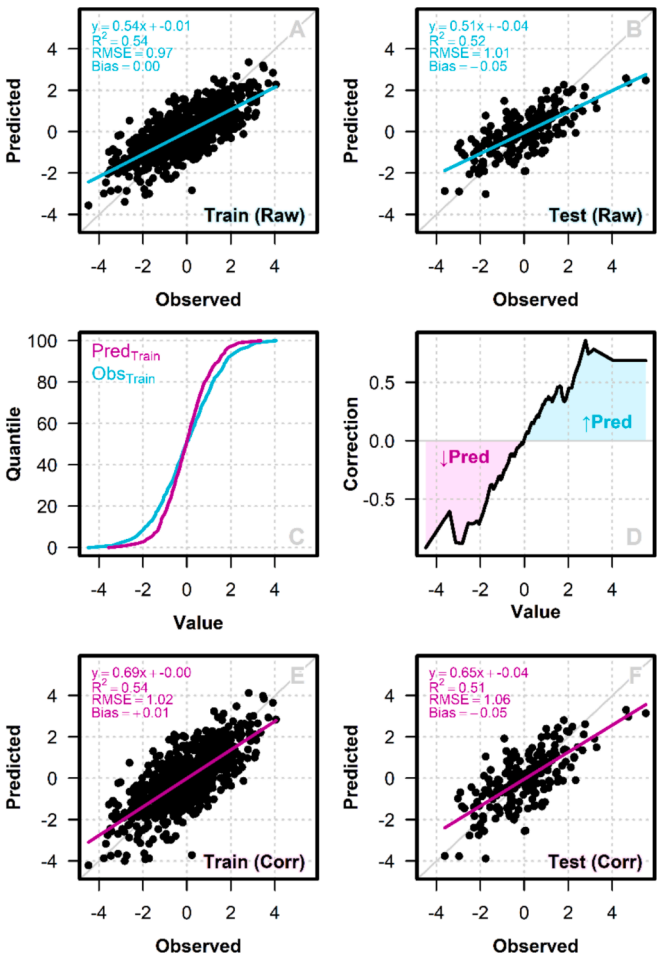

Fig. 3. Example demonstration of the bias correction approach evaluated in this study. A set of 1000 points were simulated to have a linearly correlated relationship, where the predictor variable was randomly sampled from a standard normal distribution and the response variable was equal to the predictor variable plus noise, also sampled from a standard normal distribution. The data were split into training (80%) and test (20%) data. A linear model was constructed using the training data. Applying this model to both the training (A) and test (B) data yielded predicted versus observed relationships that tended to underestimate high values and overestimate low values. By comparing the difference in quantile distributions between the training predictions and observations (C), we derive a correction factor (D) that can be applied to model predictions (E & F). The corrected models feature similar performance metrics with regression slopes closer to 1, suggesting an improved ability to predict extreme values.

2.4.3. Vegetation structure mapping

We applied our primary models (NAIP + CHM) to the prediction of the three vegetation structural metrics of interest (tree density, canopy cover, and biomass) across all 18 of our field sites. To ensure that we were avoiding ecological extrapolation, we masked the resulting maps to areas mapped as PJ woodlands according to the 2020 LANDFIRE

Existing Vegetation Type data. We used LANDFIRE 2020 data as they represented pre-field measurement conditions.

3. Results

3.1. Convolutional neural network results

The code used to produce our final canopy height predictive model, and the model itself, can be found here: https://github.com/mickeycampbell/pj-lidar-naip. The first few epochs of the U-NET CNN model saw substantial decreases in loss in both the training and validation data (Fig. 4). After that point, both training and validation loss continued to decrease, though training loss continued to decrease at a higher rate and more linearly than validation loss. The 97th epoch featured the lowest validation loss, and thus was retained as the final model used for predictive purposes. It is possible that further training may have yielded a marginally improved model, though it is likewise possible that beyond 100 epochs, overfitting may have begun to increase the validation error.

When applied to the prediction of canopy height within the test data, the model performed fairly well (Fig. 5). Given the relatively short height of PJ trees, the vast majority of predictions fell within the range of 0–8 m. Although the R^2 between predictions and observations at the 0.6 m pixel level was only 0.42, there is a high density of pixels within the 0–8 m range that fall fairly close to the 1:1 line, suggesting good prediction-observation agreement within the most relevant range of canopy heights (Fig. 5A). As we consider spatial units larger than that of the individual pixel as the basis of model performance assessment, we see that increasing area yields higher R^2 and lower %RMSE values (Fig. 5C). At the largest buffer tested (16 m), the model captured 57 % of variance in observed canopy height with a predictive error of 56 % (Fig. 5B). In other words, when aggregated to the stand scale, canopy height predictions are more accurate than the comparably noisy predictions at the sub-tree, 0.6 m pixel scale. Performance varied significantly by test site (Fig. A1), with sites 13, 17, and 18 yielding particularly low R^2 values, indicating a failure to capture prevailing trends in observed canopy height. Test site 12 had the highest predictive error, with an average %RMSE of 335, attributable to the very low observed height throughout the site, resulting in a very high mean %bias (+295).

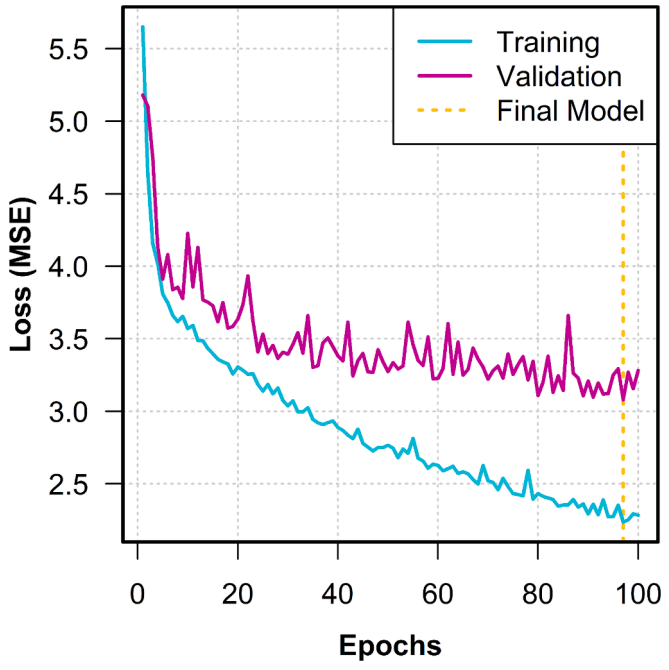

Fig. 4. Training and validation loss, as measured by mean squared error (MSE) in model-predicted canopy height in comparison to lidar-derived estimates thereof.

Fig. 6 depicts side by side comparisons of predicted and observed canopy heights, as well as the NAIP imagery used to make the predictions for geographic context for three of the 20 test sample sites (Figs. A2-A18 show the other 17 test sites). Indeed, a qualitative visual assessment of these maps revealed an impressive level of agreement between lidar-derived and CNN-predicted CHMs, with some notable caveats. There were a few extreme observed heights upwards of 40 m (Fig. 5A), which may have arisen from a number of sources, including the presence of taller trees (PJ transitions into fir and pine forests at higher elevations), powerlines, or cliffs. Cliff tops can get misclassified as non-ground points, given the extreme elevation changes that may occur over short distances. Test Area #13 in Fig. 6 illustrates one example of cliffs being misclassified as tall vegetation in the lidar CHM.

There is also a high concentration of points in Fig. 5A where the observed height is at or near zero, but the CNN predicted relatively taller vegetation. One clear example of this in the maps is Test Area #13 (Fig. 6), where there are large areas of the NAIP imagery that are completely shaded by steep relief and comparably low sun angles. In the complete absence of spectral information, the CNN was left to predict a moderate vegetation height, much of which is significantly overestimated (and some of which is also underestimated). Furthermore, whereas lidar point clouds provide discrete, three-dimensional measurements of vegetation structure, NAIP and other aerial imagery sometimes suffer from relief displacement, whereby features can appear to be laying on their side due to the perspective of the camera relative to features on the ground. In effect, this obscures ground area in the imagery, and elongates trees, both of which could explain the CNN predicting relatively tall trees where there were none.

The concentration of points along the x axis where $y \approx 0$ in Fig. 5A points towards the inverse scenario – where observed height suggested the presence of trees but predicted height was at or near zero. One explanation for these cases can be seen in Test Area #2 (Fig. 6), which featured a high concentration of tree mortality. It would appear the CNN learned that image features needed to be green for them to be mapped as having some height. Thus, trees in various stages of mortality may have been perceived as being non-tree, and thus being predicted to have little or no height. Though unintended, in effect, this acts as a boon to vegetation structure mapping, as measures of tree density, cover, and biomass are focused on live vegetation. Indeed, it serves even as somewhat of an advantage over lidar data, which "sees" structure irrespective of condition, given the lack of spectral information contained in a single laser pulse.

3.2. Vegetation structure modeling results

Our primary models, driven by predictor variables from NAIP and the CNN-derived CHM, performed well at predicting vegetation structure (Fig. 7). Canopy cover was the highest-performing structural variable, with 80 % of variance in observed cover captured by the model and a predictive error (%RMSE) of 19 %. Live aboveground biomass was the second best, with an R^2 of 0.61 and %RMSE of 42 %. The tree density model performed the worst of the three, but still demonstrated a useful degree of predictiveness, with an R^2 of 0.45 and %RMSE of 57 %. All three models were fairly unbiased when computed in aggregate (mean of predictions minus observations). However, the predictive bias varies substantially according to the magnitude of observations, with high values being underestimated and low values being overestimated. This is evident in the slopes of the linear equations for each prediction versus observation trendline. Density featured the lowest regression line slope, suggesting the greatest propensity for making overly moderate predictions (or, conversely, the poorest ability to predict extreme values accurately). The bias corrected models were intended to, and to varying degrees succeeded at, reducing this predictive tendency. The regression slopes for the bias corrected models all increased, with that of canopy

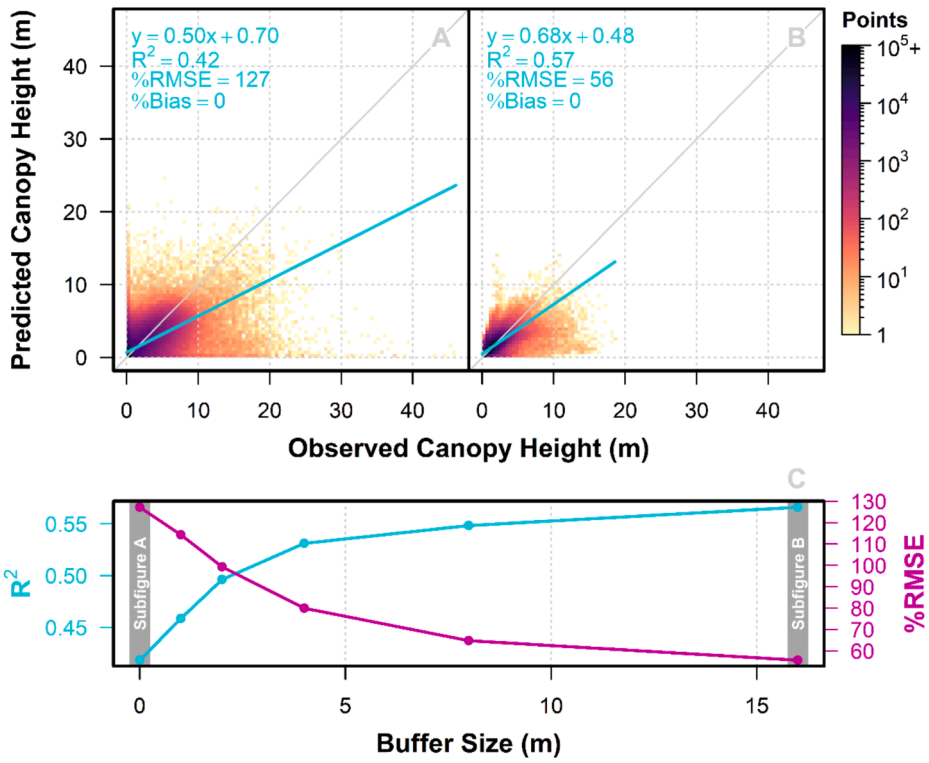

Fig. 5. Model performance of the U-Net convolutional neural network (CNN) for predicting vegetation height, compiled from a random sample of 1 M points equally distributed among the 20 test areas. (A) Predicted (y-axis) versus observed (x-axis) canopy height at the individual 0.6 m pixel level. (B) Predicted (y-axis) versus observed (x-axis) canopy height averaged within a 16 m buffer area around each sample point. (C) R^2 and %RMSE values across the different scales of model performance tested according to the buffer size around each sample point. Given the volume of points, the results in (A) and (B) are shown as a gradient of point densities on a log scale to facilitate visual interpretation in the heavily skewed data. The cyan lines in (A) and (B) represent an ordinary least-squares regression between predictions and observations, with associated statistics shown in the upper left. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cover almost reaching 1. However, the bias corrected models universally featured somewhat higher %RMSE. This highlights an important tradeoff to consider – whether it is more important to enable extreme predictions or if moderate but slightly less erroneous predictions are more desirable for a given application.

Despite being supplied with 96 candidate predictor variables, the variable selection algorithm yielded very parsimonious final sets of predictors, with the density model being based on 8 selected variables, cover on 12, and biomass on 10 (Fig. 8). The variable importance plots clearly reveal the key role that the CNN CHM played in the modeling process, with all three models featuring CHM-based variables atop the descending lists of importance according to both permutation importance and mean absolute SHAP values. Predictors representing canopy cover and volume at different height thresholds (CHM Cov 0.5 m, CHM Cov 1.0 m, CHM Vol 0.5 m, CHM Vol 1.0 m), along with the 50th percentile of height (CHM 50th), comprised the most important predictors across all three models. Two density-based variables (CHM Dens 1.0 m CHM Dens 2.0 m), also appeared in the tree density model, found to be most important according to SHAP values, which aligned with expectations. Among the spectral predictor variables, skewness and kurtosis of various normalized differences both emerged as important predictors, showing up eight times in Fig. 8. Five of the six normalized difference indices were featured in the final predictor set, with NDNR (also known as the normalized difference vegetation index) and NDNG (also known as the green normalized difference vegetation index) being most abundant.

Our primary models (NAIP + CHM) outperformed the NAIP-only and Landsat + topography + climate models in both R^2 and %RMSE across all three vegetation structural response variables (Table 2). The improved performance over the NAIP-only models, in particular, speaks to the value that the CNN CHM modeling procedure added to the ability to accurately predict vegetation structure. On average across all three response variables, NAIP-only explained approximately 19 % less variance and featured 15 % higher predictive error than the NAIP + CHM models. The relatively poor performance of Landsat + topography + climate is particularly noteworthy, given the ubiquity of these data sources in the remote sensing of vegetation literature. On average, these models explained 32 % less variance in observed vegetation structure and yielded 29 % higher prediction error than our primary models. Landsat's poor performance in modeling biomass is worth noting (R^2 = 0.29; %RMSE = 56 %), given the importance of live aboveground biomass as a proxy for carbon stocks. As anticipated, the models driven by airborne lidar data featured the highest performance of all. This is particularly true of tree density, where airborne lidar had a 50 % higher R^2 and a 23 % lower error than the primary NAIP + CHM model. The differences in the biomass model were less significant (12 % better R^2 and 10 % better RMSE than NAIP + CHM). Given that airborne lidar is often considered the "gold standard" for biomass mapping, the fact that our primary model achieved comparable biomass mapping performance is a valuable finding.

3.3. *Vegetation structure mapping results*

One example of the vegetation structure mapping results can be seen in Fig. 9. Given the correlation between them, the three structural variables convey the same spatial trends, such that higher density areas also tend to have higher cover and biomass. A close visual inspection, however, reveals some key differences. For example, cover tends to vary less over local scales, whereas density and biomass feature more localized hotspots. Furthermore, there are examples of areas that feature

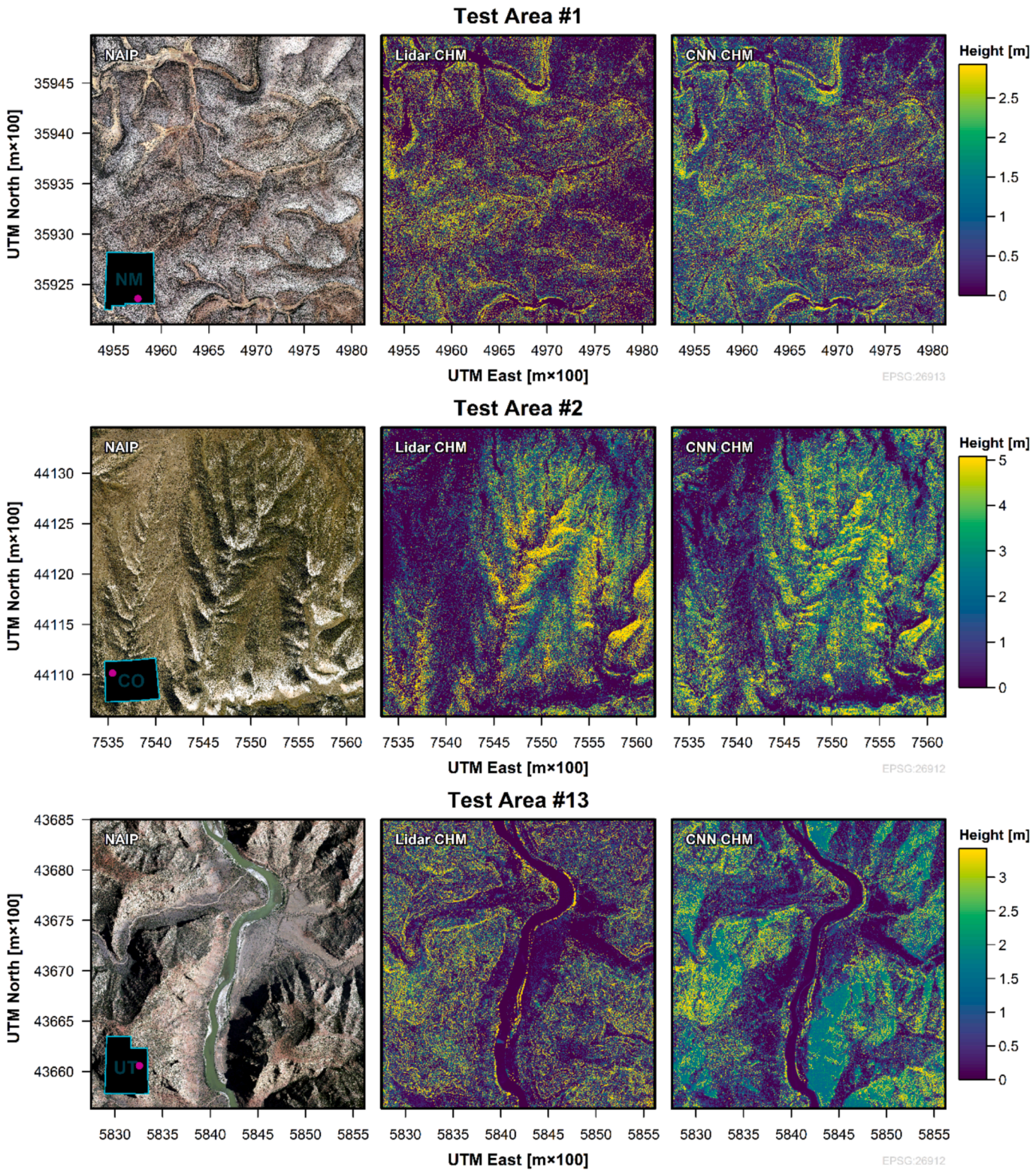

Fig. 6. Three examples from the 20 test sample sites where per-pixel vegetation height was modeled from aerial imagery using a convolutional neural network. Test Area #1 (top) represented a case where the model captured patterns of vegetation height with minimal error. Test Area #2 (middle) represented a case where tree mortality yielded patches of tree height underprediction. Test Area #13 (bottom) represented a case where topographically driven shadows yielded large swaths of uncertainty in vegetation height predictions and where steep cliffs yielded overestimation of vegetation height predictions. In all three examples, the locator map on the bottom left features a magenta point representing the test area's location within its US state. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

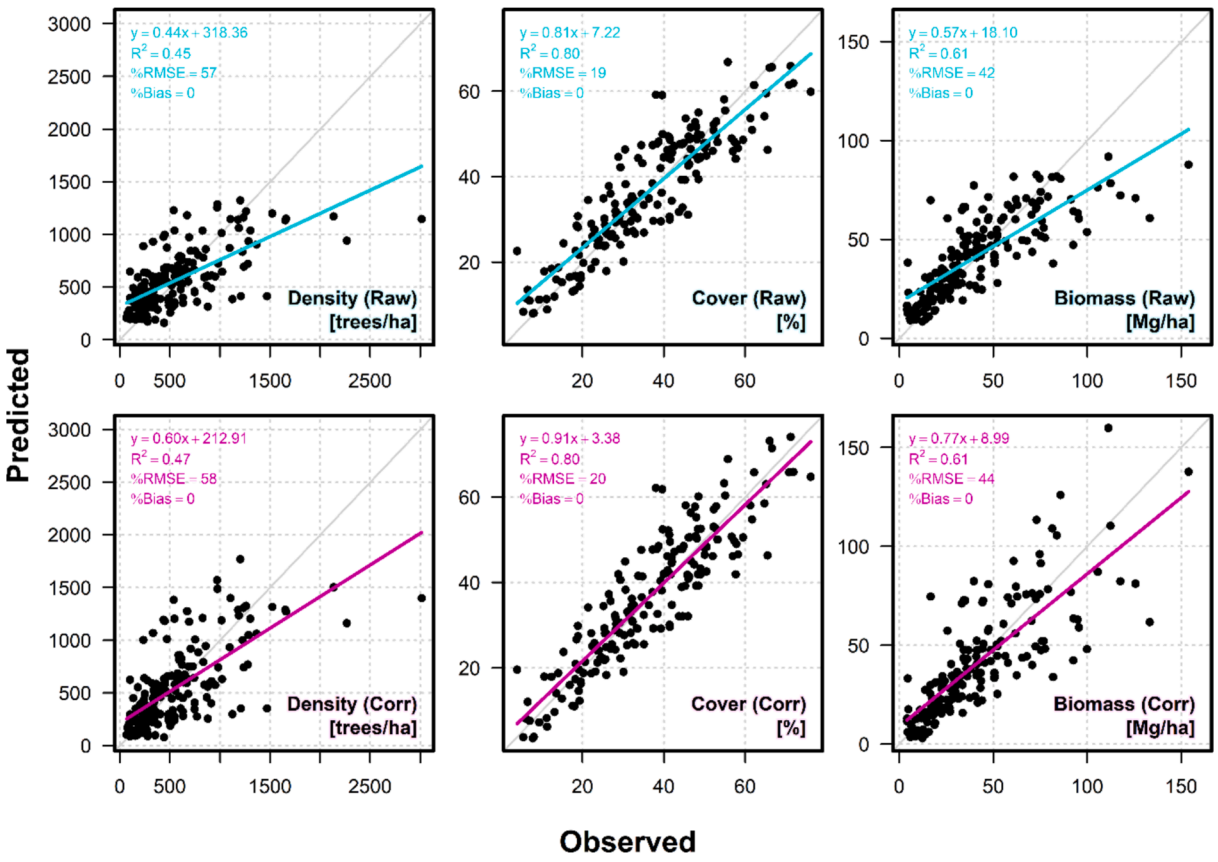

Fig. 7. Predictions (y-axis) versus observations (x-axis) for the NAIP + CHM random forest models within the 180 field plots. The top row represents the raw predictions, whereas the bottom row represents the bias-corrected predictions. All results are based on a leave one site out cross-validation procedure.

relatively higher cover, but lower density and biomass, such as in the southeastern portion of the site displayed in Fig. 9. Some NAIP artifacts are clear in the maps – especially Fig. 9B – where areas of higher and lower predicted canopy height are separated by a vertical flight line boundary. Note that this effect is also present in Test Area #14 (Fig. A12).

When zoomed in to the level of individual trees near one of the image mosaic seamlines (Fig. 10), a few notable artifacts emerge. First, the imagery to the right of the seamline (predicted as having higher density/cover/biomass) featured higher relief displacement due to having been captured from a higher off-nadir viewing angle. This effectively elongates the trees, artificially increasing the projected two-dimensional area of each of each tree. This would act to inflate many of the predictor variables (e.g., cover, volume), yielding downstream vegetation structural overprediction. Secondly, the shadow length and orientation differences point to different illumination conditions, where the imagery on the right of the seamline was collected with a lower solar zenith angle. These shadow areas were predicted to have some vegetation height, likely since shaded portions of trees (which have height) may be indistinguishable from tree shadows on the ground (which have no height). Furthermore, the CNN may have learned that longer shadows correspond to taller trees, increasing the predicted heights of individual trees. Taken together, variability in viewing perspective and illumination conditions appeared that manifested in the NAIP imagery acted to yield artifacts in the vegetation structure prediction maps, though cover and density appear more robust to these artifacts than biomass.

4. Discussion

Our primary objective in this study was to introduce and describe a robust analytical framework for mapping vegetation structure in an ecologically diverse and geographically expansive dry woodland ecosystem. To that end, we have presented a robust, multistage data fusion workflow, which leverages the individual strengths and overcomes the weaknesses of NAIP imagery and airborne lidar, to map tree density, canopy cover, and biomass across the wide range of PJ woodlands in the Western US. By leveraging both deep and machine learning techniques, the methods that we have described herein could be applied in a repeatable manner to quantify important woodland biometrics on broad spatial scales.

NAIP's strengths lie primarily in its high two-dimensional spatial resolution and its systematic, statewide, collection frequency. To understand not just how vegetation structure varies over space but also over time requires remote sensing data collected with some level of temporal consistency. Given that NAIP is both spatially and temporally exhaustive in the contiguous US, it could potentially serve as a useful basis for vegetation structure mapping. Indeed, our model comparison results show that NAIP alone can capture prevailing trends in tree density, canopy cover, and biomass. This finding adds to a fairly lengthy body of existing scientific literature where NAIP or other comparable airborne image datasets have shown promise towards accurate vegetation mapping (e.g., Basu et al., 2015; Erker et al., 2019; Hogland et al., 2018; Sunde et al., 2020). Indeed, even some have demonstrated success in PJ and other dry woodland ecosystems (Coates et al., 2017; Davies et al., 2010; Gustafson et al., 2018; Hulet et al., 2014; Ku and Popescu, 2019). However, to the authors' knowledge, none of these efforts have designed their study specifically to question NAIP's ability to map PJ tree structure on an ecosystem-wide basis. Here we have shown that not only is NAIP capable of mapping vegetation structure in ecologically narrow portions of PJ's vast range, but our leave-one-site-out cross-validation is a clear demonstration that structure can be mapped across ecologically diverse portions of the range with relatively high accuracy

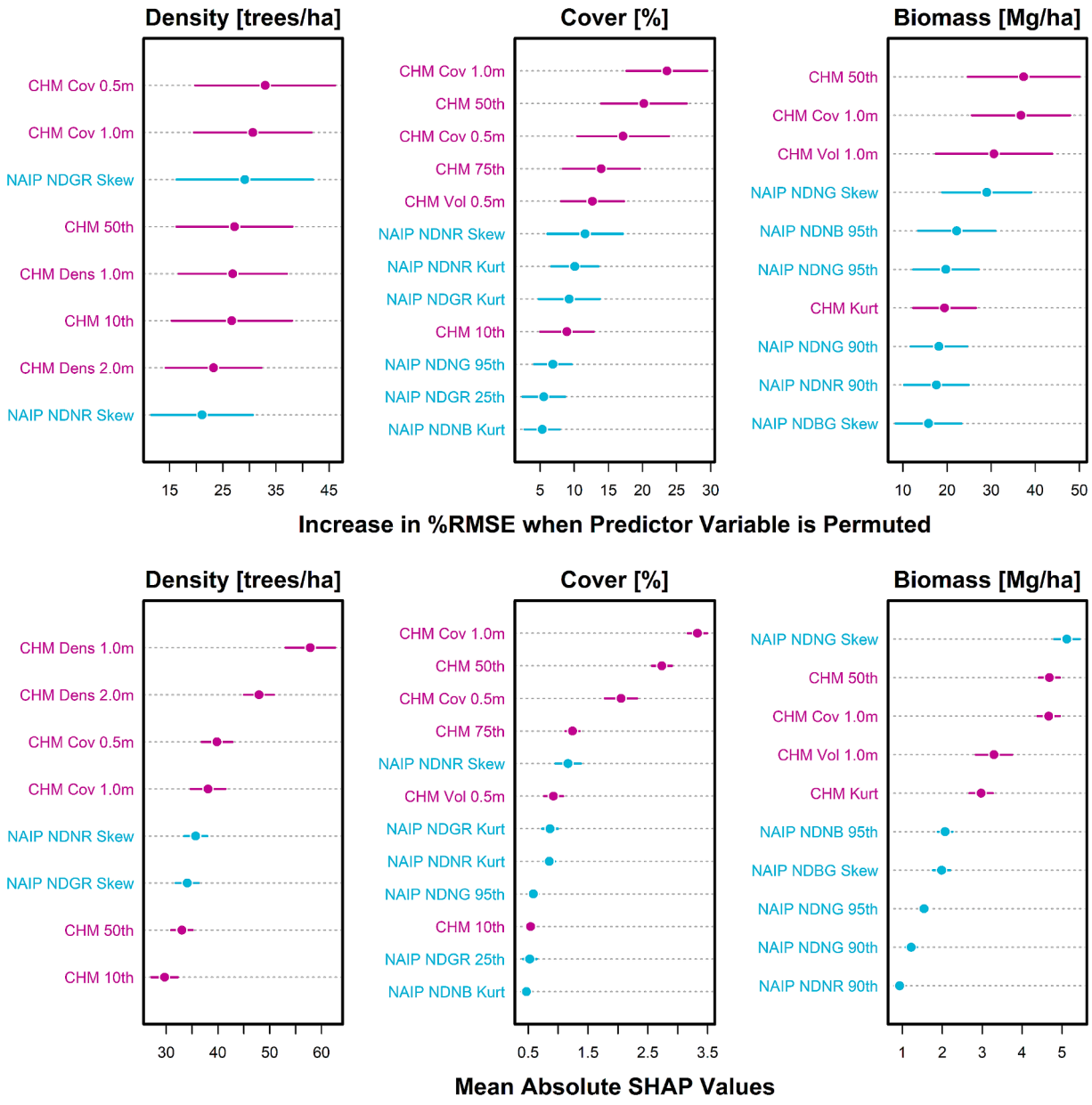

Fig. 8. Variable importance for the three tree structural random forest models. The points represent the mean and the lines represent the standard deviation of importance computed across the 18 folds in the leave one site out cross-validation procedure. The CHM-based predictor variables are shown in magenta, whereas the NAIP imagery-based predictor variables are shown in cyan. The top row provides permutation importance, measured by the increase in predictive percent root mean squared error (%RMSE) that occurs when a predictor variable's values are randomly permuted. The bottom row provides mean absolute Shapley additive explanations (SHAP) values, which quantify individual predictor variables' marginal contribution to a model's prediction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Eastburn et al., 2024).

Airborne lidar adds a third dimension (height), which is critical for vegetation structure mapping, but absent in passive optical aerial image data. Vegetation height, and the various derivative statistics one can extract from lidar-based height measurements, have been widely shown to be an essential component of vegetation structure mapping (Zolkos et al., 2013). Our lidar-based predictive model outperformed all others, highlighting just how impressive its vegetation structure mapping capabilities are, despite the inherent remote sensing challenges posed by our focal dryland ecosystem. However, airborne lidar, at least to date, are not collected with the same programmatic frequency in the US as NAIP imagery, and far less so than satellite imaging programs. Although the USGS 3D Elevation Program has made great strides towards nationwide lidar data collection, these data can quickly become outdated as vegetation structure changes in response to afforestation and deforestation, the latter of which can occur rapidly and on broad spatial scales (Snyder, 2012). Thus, airborne lidar alone, though certainly considered a gold standard for providing vegetation structural snapshots in time, is currently poorly suited for temporally consistent vegetation mapping.

Through the use of a U-Net CNN, we built a model capable of predicting per-pixel canopy height using only NAIP imagery. As a result, we can fuse the critical vertical dimension to NAIP, at once overcoming NAIP's two-dimensional nature and also lidar's infrequent, opportunistic temporal availability. Although the CNN-derived CHMs are not as accurate as those derived from lidar, our study showed that adding CHM-based predictors greatly enhanced the ability to accurately map vegetation structure in comparison to NAIP alone. To be sure, we are not

Table 2
Performance metrics (R^2 and RMSE) for the vegetation structural predictive models based on four different sets of predictor variables. All four predictor sets contain the performance metrics, but the three comparison models (NAIP, Landsat + Topo + Clim, and Airborne Lidar) also feature parenthetical proportional comparisons to the primary NAIP + CHM model to convey how much better or worse they are performing relative to the central model described in this study. All metrics are based on the raw (not bias-corrected) models. The not applicable (NA) values for the Airborne Lidar-driven Cover model statistics is meant to highlight that, since lidar was used to derive these response variables, the model performance should be ignored.

Response Variable	Performance Metric	NAIP + CHM	NAIP Only	Landsat + Topo + Clim	Airborne Lidar
Density	R^2	0.45	0.35 (−23 %)	0.37 (−17 %)	0.68 (+50 %)
	%RMSE	57	63 (+9%)	62 (+7%)	44 (−23 %)
Cover	R^2	0.80	0.72 (−9%)	0.58 (−27 %)	NA
	%RMSE	19	22 (+17 %)	28 (+45 %)	NA
Biomass	R^2	0.61	0.46 (−25 %)	0.29 (−52 %)	0.69 (+12 %)
	%RMSE	42	50 (+20 %)	56 (+35 %)	38 (−10 %)

the first to predict lidar-based canopy height using high-resolution aerial or satellite imagery (Chang et al., 2025; Malambo and Popescu, 2023; Tolan et al., 2024; Wagner et al., 2024). However, to our knowledge, we are the first to demonstrate this analysis technique with an ecosystem-specific focus on PJ woodlands. In some ways, PJ may be ideally suited to this task, given the trees' relatively wide spacing and the resultant spatial variability in image texture and pattern, both of which CNNs rely on for learning purposes. In other ways, however, PJ woodlands pose novel challenges, such as their relatively short stature, and low variability therein. In many forests, the difference between a sapling and mature tree can be tens of meters. In PJ, that same difference may only be one or two meters. So, while CNN's may excel at capturing the patterns of individual trees, their ability to accurately attribute heights to those trees may be limited by the small amount of training variability in the response variable supplied to the model. This manifested in our results, where even small shifts in tree shadow length and orientation led to substantial differences in height predictions, which yielded downstream vegetation structural prediction uncertainties. Future work should focus on replicating this analytical workflow in other tree-dominated ecosystems to evaluate its broad applicability. Given Allred et al. (2025)'s recent release of calibrated NAIP-lidar CHM imagery pairs, this process could greatly simplified.

Although U-Nets have been widely shown to be effective at a range of image analysis tasks in the ecological remote sensing discipline, they are but one category of deep learning architecture that we could have employed to map canopy heights. Indeed, even within the general U-Net framework there are many different ways to design the network, including the use of various backbones (e.g., ResNet), with and without pre-trained weights, to improve model performance (Cao and Zhang, 2020; He et al., 2016). It is possible that an entirely different deep learning algorithm, such as a transformer-based approach, may have yielded superior canopy height predictive performance (Chang et al., 2025; Wagner et al., 2024). An exhaustive comparison of different deep learning algorithms, architectures, and hyperparameters was beyond the scope of this study. Instead, our focus was on demonstrating one relatively simple, widely used CNN technique within a broader analytical framework aimed at using modeled canopy heights as important predictors of ecologically valuable biometrics (density, cover, and biomass). Future research should explore alternative deep learning approaches in the hopes of improving model performance. Furthermore, given the fact that many foundational CNN models are trained on 3-band (RGB) imagery, there may be value in establishing foundational models trained on 4-band NAIP imagery in the future.

It should be noted that deep learning is not the only approach for getting height information from NAIP imagery. The NAIP program has begun in recent years to make digital aerial photogrammetric point cloud data available for purchase. Recent studies have shown varying degrees of success at using these data products to map tree height (e.g., Prior et al., 2022; Ritz et al., 2022; Schroeder et al., 2022), though lidar-derived terrain elevations are typically needed to normalize surface elevations to aboveground heights. To our knowledge, there have been no quantitative comparisons of deep learning-derived and NAIP photogrammetric canopy height models. A future comparison like this would enable a valuable assessment of tradeoffs between height estimation approaches.

Although we did not explicitly test it, our CNN demonstrated a useful degree of temporal transferability. The training, validation, and test data represented aerial imagery collected over the course of three years, collected with different specifications, at different times of year. Furthermore, we applied the model to the prediction of vegetation heights within our field sites, whose image data were collected over three different years. Although we did not test the accuracy of height predictions in our field sites, the strength of the vegetation structural models, which were heavily influenced by CNN CHM-derived predictors, suggests that heights were modeled with a useful degree of accuracy. Future research should aim to test the temporal consistency of predictions, by applying the CNN model to the same location in successive image captures. Given that our CNN showed evidence of the ability to distinguish between live and dead trees, a future study could explore the CNN's ability to specifically quantify disturbance severity.

The poor performance of Landsat, topography, and climate data at explaining vegetation structural variability in PJ woodlands is particularly noteworthy. Several studies have used Landsat as the primary or secondary basis of vegetation structure mapping in PJ woodlands, though most of these were undertaken at comparably local scales (Brewer et al., 2017; Campbell et al., 2021; Huang et al., 2010; Sankey and Glenn, 2011; Yang et al., 2012). One exception to this is Filippelli et al. (2020), who mapped PJ biomass across the entire Great Basin using Landsat, topography, and climate data; however, even their results showed poor agreement with field-based biomass estimates at the pixel-level. Another is Reinhardt et al. (2020), who mapped PJ canopy cover at a similar scale to Filippelli et al. (2020), yielding predictive results quite similar in performance to our study, as canopy cover yielded the best performance of the structural metrics we modeled. At best, topography and climate data can capture the ecological niche of a vegetation type, and perhaps even quantify biomass trends within that niche as a function of environmental factors such as moisture availability (Xu et al., 2015). However, given the complex management-driven and naturally occurring disturbance history in PJ woodlands, optical imagery is needed to augment environmentally driven biomass predictions. The problem is that Landsat, with 30 m spatial resolution, may be poorly suited to mapping structure in these sparsely canopied, low-productivity woodlands. The high exposure of understory rock, soil, grasses, forbs, and shrubs, coupled with tree canopies with relatively low greenness means that the materials driving surface reflectance and spectral indices are not dominantly the tree canopies themselves (Smith et al., 2019). So, while Landsat may be able to map PJ biomass locally, where understory conditions are held relatively constant, when understory conditions vary between ecologically distinct sites, Landsat falters.

A close examination of our NAIP + CHM variable importance may explain some of Landsat's poor performance. As Fig. 8 highlights, nearly all of the NAIP-based predictor variables found to be most important were either skewness/kurtosis measures or high (90th or 95th)

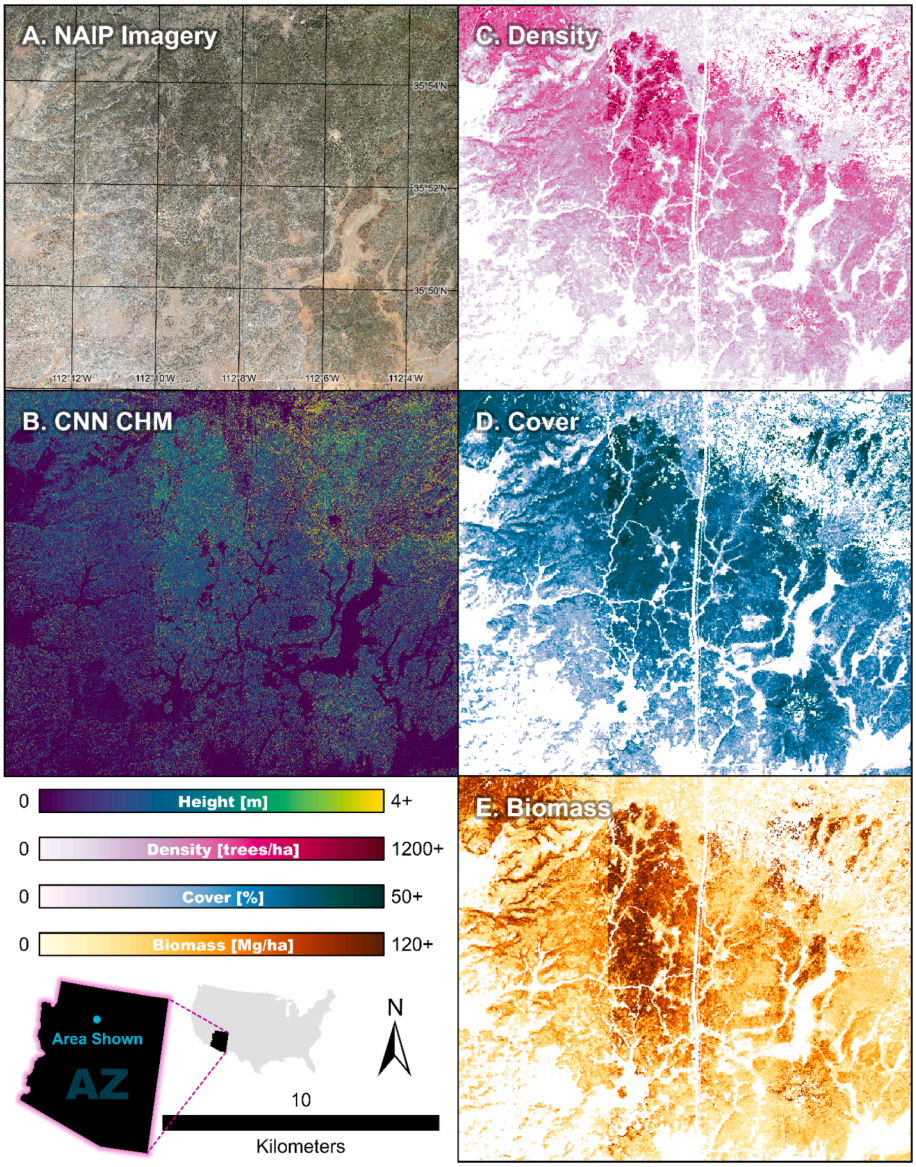

Fig. 9. Results of the vegetation structure mapping in one of the 18 field sites in Northern Arizona, USA, including (A) National Agricultural Inventory Program (NAIP) imagery; (B) the convolutional neural network (CNN)-modeled canopy height model (CHM); (C) random forest-modeled tree density; (D) random forest-modeled canopy cover; and (E) random forest-modeled live aboveground biomass. (A-B) are at 0.6 m resolution whereas (C-E) are at 30 m resolution, masked to areas mapped as piñon-juniper woodlands in 2020 LANDFIRE Existing Vegetation Type data. (A-E) are shown in EPSG:26912.

percentiles of spectral indices. This highlights the importance of non-normal distributions of spectral constituents within an aggregated 30 m pixel, and how high-end spectral index values (e.g., tree canopies with a dominantly barren soil background) are more predictive than measures of central tendency. The most Landsat-like predictor one could derive from aggregation of NAIP data would be a mean, but no mean values were selected in any of our models, suggesting that spectral averaging over 900 m^2 areas does not adequately capture vegetation structure in PJ and perhaps other dry woodlands that feature spectrally and/or structurally similar vegetation.

A limitation of our primary modeling workflow comes from the artifacts inherent to aerial image data. Acknowledged by other studies that have relied on NAIP and comparable image datasets (Hogland et al., 2018; Maxwell et al., 2017; Richardson and Moskal, 2014), differences in solar illumination and viewing perspective can drive vegetation predictive uncertainty. There are a few ways to potentially mitigate these effects. First, ensuring that imagery captured with wide-ranging illumination and viewing perspective conditions are included in the training and validation data could enhance a CNN's ability to account for these effects in the prediction of vegetation heights. Secondly, including solar zenith angle metadata in the modeling process – in the CNN and/or the vegetation structure random forests – could offset the effects of shadowing. This would be difficult to do with NAIP data, as time of day and day of year metadata are not readily available on a per-pixel basis. Third, the use of multitemporal aerial imagery could minimize scene-specific artifacts. By applying a CNN to the prediction on image data collected consecutively, averaging the height predictions could potentially improve results. Fourth, using high-resolution satellite imagery (e.g., Tolan et al., 2024), rather than aerial imagery, would both minimize relief displacement and enable more temporally specific image capture, minimizing illumination effects. Lastly, incorporating other predictor variables into the CNN modeling workflow, such as topography (e.g., Sha et al., 2020), could help improve the model learn environmental patterns of tree height in addition to optical patterns.

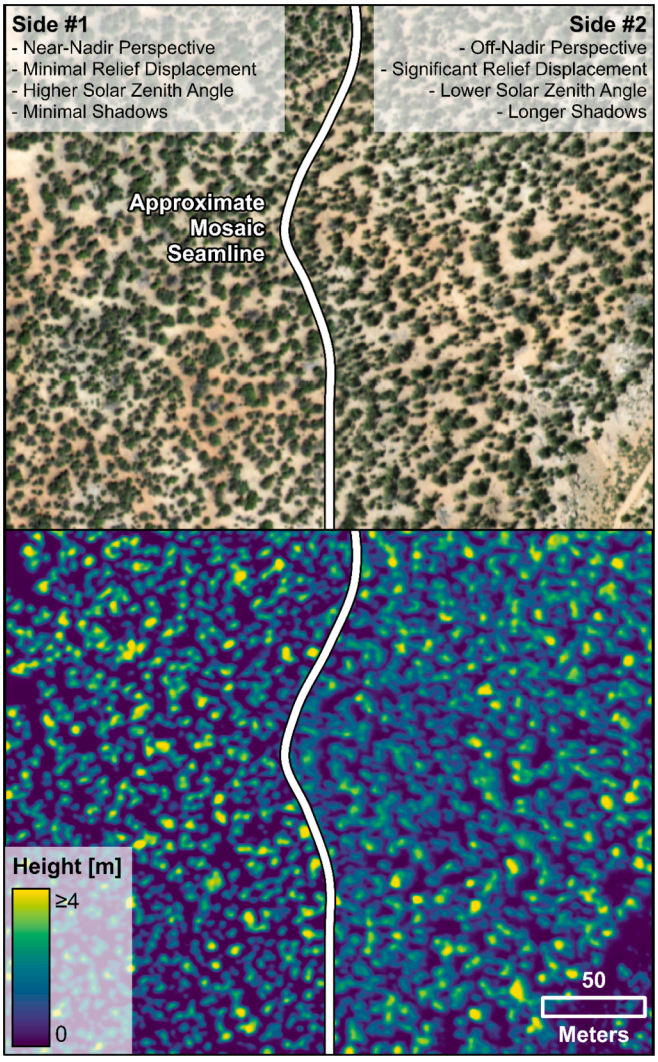

Fig. 10. A close-up view of the apparent seamline between two adjacent areas of National Agricultural Inventory Program (NAIP) imagery collected under different viewing perspective and solar illumination conditions (top), and the associated predicted canopy height data (bottom) conveying the artifacts that yielded spatial discontinuities in the vegetation structure prediction maps.

5. Conclusions

Given the importance of dryland ecosystems in global land-climate interactions, we need to develop remote sensing-driven workflows for consistently and accurately mapping their vegetation structure over space and time. In this study, we have introduced one such workflow that overcomes some of the most critical tradeoffs faced by individual remotely sensed datasets, potentially enabling broad-scale and temporally consistent vegetation structure mapping. Although our models performed somewhat poorer than those driven by predictor variables derived from airborne lidar data, our reliance on NAIP imagery, which is collected every two years in the US, provides the ability to evaluate change over time, which opportunistically collected lidar data do not. Furthermore, our model greatly outperformed one driven by Landsat, topography, and climate data, whose comparably poor performance emphasizes the complexity of mapping vegetation in structurally complex dry woodlands. Although our focus in this study was on dry woodlands, the workflow we have introduced should be tested in other tree-dominated ecosystems. Future research should aim to evaluate whether our approach could yield similar predictive performance in more mesic, forested settings. Furthermore, more research is needed to test the temporal consistency of this analytical approach. Given the sometimes-inconsistent spectral characteristics of aerial imagery, understanding the limitations of our approach to quantifying vegetation structural change over time will inform the degree to which our approach can be operationalized for broad-scale, systematic mapping of vegetation structure.

CRediT authorship contribution statement

Michael J. Campbell: Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jessie F. Eastburn:** Writing – review & editing, Investigation. **Simon C. Brewer:** Writing – review & editing, Methodology. **Philip E. Dennison:** Writing – review & editing, Supervision, Resources, Conceptualization.

Funding

This work was supported by the National Aeronautics and Space Administration [grant number 80NSSC21K0948].

Declaration of competing interest

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

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2025.05.016.

References

Ahlström, A., Raupach, M.R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J.G., Friedlingstein, P., Jain, A.K., Kato, E., Poulter, B., Sitch, S., Stocker, B.D., Viovy, N., Wang, Y.P., Wiltshire, A., Zaehle, S., Zeng, N., 2015. The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink. Science 348, 895–899. https://doi.org/10.1126/science.aaa1668.

Ahmad, A., Gilani, H., Ahmad, S.R., 2021. Forest Aboveground Biomass Estimation and Mapping through High-Resolution Optical Satellite Imagery—A Literature Review. Forests 12, 914. https://doi.org/10.3390/f12070914.

Allred, B.W., McCord, S.E., Morford, S.L., 2025. Canopy height model and NAIP imagery pairs across CONUS. Sci. Data 12, 322. https://doi.org/10.1038/s41597-025-04655-z.

Asner, G.P., Mascaro, J., Muller-Landau, H.C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J.S., van Breugel, M., 2012. A universal airborne LiDAR approach for tropical forest carbon mapping. Oecologia 168, 1147–1160. https://doi.org/10.1007/s00442-011-2165-z.

Badgley, G., Field, C.B., Berry, J.A., 2017. Canopy near-infrared reflectance and terrestrial photosynthesis. Sci. Adv. 3, e1602244. https://doi.org/10.1126/sciadv.1602244.

Baig, M.H.A., Zhang, L., Shuai, T., Tong, Q., 2014. Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. Remote Sens. Lett. 5, 423–431. https://doi.org/10.1080/2150704X.2014.915434.

Basu, S., Ganguly, S., Nemani, R.R., Mukhopadhyay, S., Zhang, G., Milesi, C., Michaelis, A., Votava, P., Dubayah, R., Duncanson, L., Cook, B., Yu, Y., Saatchi, S., DiBiano, R., Karki, M., Boyda, E., Kumar, U., Li, S., 2015. A Semiautomated Probabilistic Framework for Tree-Cover Delineation From 1-m NAIP Imagery Using a High-Performance Computing Architecture. IEEE Trans. Geosci. Remote Sens. 53, 5690–5708. https://doi.org/10.1109/TGRS.2015.2428197.

Belitz, K., Stackelberg, P.E., 2021. Evaluation of six methods for correcting bias in estimates from ensemble tree machine learning regression models. Environ. Model Softw. 139, 105006. https://doi.org/10.1016/j.envsoft.2021.105006.

Böhner, J., Antonić, O., 2009. Chapter 8 Land-Surface Parameters Specific to Topo-Climatology, in: Hengl, T., Reuter, H.I. (Eds.), Developments in Soil Science, Geomorphometry. Elsevier, pp. 195–226. Doi: 10.1016/S0166-2481(08)00008-1.

Brewer, W.L., Lippitt, C.L., Lippitt, C.D., Litvak, M.E., 2017. Assessing drought-induced change in a piñon-juniper woodland with Landsat: a multiple endmember spectral mixture analysis approach. Int. J. Remote Sens. 38, 4156–4176. https://doi.org/10.1080/01431161.2017.1317940.

Campbell, M.J., Dennison, P.E., Kerr, K.L., Brewer, S.C., Anderegg, W.R.L., 2021. Scaled biomass estimation in woodland ecosystems: Testing the individual and combined

capacities of satellite multispectral and lidar data. Remote Sens. Environ. 262, 112511. https://doi.org/10.1016/j.rse.2021.112511.
Campbell, M.J., Eastburn, J.F., Dennison, P.E., Vogeler, J.C., Stovall, A.E.L., 2024. Evaluating the performance of airborne and spaceborne lidar for mapping biomass in the United States' largest dry woodland ecosystem. Remote Sens. Environ. 308, 114196. https://doi.org/10.1016/j.rse.2024.114196.
Campbell, M.J., Eastburn, J.F., Mistick, K.A., Smith, A.M., Stovall, A.E.L., 2023. Mapping individual tree and plot-level biomass using airborne and mobile lidar in piñon-juniper woodlands. Int. J. Appl. Earth Obs. Geoinf. 118, 103232. https://doi.org/10.1016/j.jag.2023.103232.
Campos, V.E., Gatica, G.M., Cappa, F.M., Giannoni, S.M., Campos, C.M., 2018. Remote sensing data to assess compositional and structural indicators in dry woodland. Ecol. Ind. 88, 63–70. https://doi.org/10.1016/j.ecolind.2018.01.032.
Cao, K., Zhang, X., 2020. An Improved Res-UNet Model for Tree Species Classification Using Airborne High-Resolution Images. Remote Sens. (Basel) 12, 1128. https://doi.org/10.3390/rs12071128.
Chang, T., Ndegwa, K., Gros, A., Landau, V.A., Zachmann, L.J., State, B., Gritts, M.A., Miller, C.W., Rutenbeck, N.E., Conway, S., Bayes, G., 2025. VibrantVS: A High-Resolution Vision Transformer for Forest Canopy Height Estimation. Remote Sens. (Basel) 17, 1017. https://doi.org/10.3390/rs17061017.
Chojnacky, D.C., Heath, L.S., Jenkins, J.C., 2014. Updated generalized biomass equations for North American tree species. Forestry (lond) 87, 129–151. https://doi.org/10.1093/forestry/cpt053.
Coates, P.S., Prochazka, B.G., Ricca, M.A., Gustafson, K.B., Ziegler, P., Casazza, M.L., 2017. Pinyon and Juniper Encroachment into Sagebrush Ecosystems Impacts Distribution and Survival of Greater Sage-Grouse. Rangeland Ecology & Management, Woody invasion of western rangelands: Using grouse as focal species for ecosystem restoration 70, 25–38. Doi: 10.1016/j.rama.2016.09.001.
Dalla Mura, M., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., Benediktsson, J.A., 2015. Challenges and Opportunities of Multimodality and Data Fusion in Remote Sensing. Proc. IEEE 103, 1585–1601. https://doi.org/10.1109/JPROC.2015.2462751.
Daly, C., Taylor, G.H., Gibson, W.P., 1997. The PRISM approach to mapping precipitation and temperature, in: Proc., 10th AMS Conf. on Applied Climatology. pp. 20–23.
David, R.M., Rosser, N.J., Donoghue, D.N.M., 2022. Remote sensing for monitoring tropical dryland forests: a review of current research, knowledge gaps and future directions for Southern Africa. Environ. Res. Commun. 4, 042001. https://doi.org/10.1088/2515-7620/ac5b84.
Davies, K.W., Petersen, S.L., Johnson, D.D., Bracken Davis, D., Madsen, M.D., Zvirzdin, D.L., Bates, J.D., 2010. Estimating Juniper Cover From National Agriculture Imagery Program (NAIP) Imagery and Evaluating Relationships Between Potential Cover and Environmental Variables. Rangel. Ecol. Manage. 63, 630–637. https://doi.org/10.2111/REM-D-09-00129.1.
De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., Crombé, P., 2013. Application of the topographic position index to heterogeneous landscapes. Geomorphology 186, 39–49. https://doi.org/10.1016/j.geomorph.2012.12.015.
De Sy, V., Herold, M., Achard, F., Asner, G.P., Held, A., Kellndorfer, J., Verbesselt, J., 2012. Synergies of multiple remote sensing data sources for REDD+ monitoring. Current Opin. Environm. Sustain. 4/6 Clim. Syst. 4, 696–706. https://doi.org/10.1016/j.cosust.2012.09.013.
Dubayah, R., Armston, J., Healey, S.P., Bruening, J.M., Patterson, P.L., Kellner, J.R., Duncanson, L., Saarela, S., Ståhl, G., Yang, Z., Tang, H., Blair, J.B., Fatoyinbo, L., Goetz, S., Hancock, S., Hansen, M., Hofton, M., Hurtt, G., Luthcke, S., 2022. GEDI launches a new era of biomass inference from space. Environ. Res. Lett. 17, 095001. https://doi.org/10.1088/1748-9326/ac8694.
Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. Sci. Remote Sens. 1, 100002. https://doi.org/10.1016/j.srs.2020.100002.
Duncanson, L., Hunka, N., Jucker, T., Armston, J., Harris, N., Fatoyinbo, L., Williams, C. A., Atkins, J.W., Raczka, B., Serbin, S., Keller, M., Dubayah, R., Babcock, C., Cochrane, M.A., Hudak, A., Hurtt, G.C., Montesano, P.M., Moskal, L.M., Park, T., Saatchi, S., Silva, C.A., Tang, H., Vargas, R., Weiskittel, A., Wessels, K., Goetz, S.J., 2025. Spatial resolution for forest carbon maps. Science 387, 370–371. https://doi.org/10.1126/science.adt6811.
Eastburn, J.F., Campbell, M.J., Dennison, P.E., Anderegg, W.R., Barrett, K.J., Fekety, P. A., Flake, S.W., Huffman, D.W., Kannenberg, S.A., Kerr, K.L., Sánchez Meador, A.J., Vogeler, J.C., 2024. Ecological and climatic transferability of airborne lidar-driven aboveground biomass models in Piñon-Juniper woodlands. Giscience & Remote Sensing 61, 2363577. https://doi.org/10.1080/15481603.2024.2363577.
Erker, T., Wang, L., Lorentz, L., Stoltman, A., Townsend, P.A., 2019. A statewide urban tree canopy mapping method. Remote Sens. Environ. 229, 148–158. https://doi.org/10.1016/j.rse.2019.03.037.
Evans, J.S., Murphy, M.A., Ram, K., 2023. spatialEco: Spatial Analysis and Modelling Utilities.
Falkowski, M.J., Evans, J.S., Naugle, D.E., Hagen, C.A., Carleton, S.A., Maestas, J.D., Khalyani, A.H., Poznanovic, A.J., Lawrence, A.J., 2017. Mapping tree canopy cover in support of proactive prairie grouse conservation in Western North America. rangeland ecology & management, woody invasion of western rangelands: using grouse as focal species for ecosystem restoration 70, 15–24. Doi: 10.1016/j.rama.2016.08.002.
Filippelli, S.K., Falkowski, M.J., Hudak, A.T., Fekety, P., Vogeler, J.C., Khalyani, A.H., Rau, B.M., Strand, E.K., 2020. Monitoring pinyon-juniper cover and aboveground biomass across the Great Basin. Environ. Res. Lett. https://doi.org/10.1088/1748-9326/ab6785.
Genuer, R., Poggi, J.-M., Tuleau-Malot, C., 2015. VSURF: An R Package for Variable Selection Using Random Forests. The R Journal 7, 19–33.
Gherardi, L.A., Sala, O.E., 2019. Effect of interannual precipitation variability on dryland productivity: A global synthesis. Glob. Chang. Biol. 25, 269–276. https://doi.org/10.1111/gcb.14480.
Gizachew, B., Solberg, S., Næsset, E., Gobakken, T., Bollandsås, O.M., Breidenbach, J., Zahabu, E., Mauya, E.W., 2016. Mapping and estimating the total living biomass and carbon in low-biomass woodlands using Landsat 8 CDR data. Carbon Balance Manag. 11, 13. https://doi.org/10.1186/s13021-016-0055-8.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, Big Remotely Sensed Data: Tools, Applications and Experiences 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.
Gustafson, K.B., Coates, P.S., Roth, C.L., Chenaille, M.P., Ricca, M.A., Sanchez-Chopitea, E., Casazza, M.L., 2018. Using object-based image analysis to conduct high- resolution conifer extraction at regional spatial scales. Int. J. Appl. Earth Obs. Geoinf. 73, 148–155. https://doi.org/10.1016/j.jag.2018.06.002.
Hanan, N.P., Milne, E., Aynekulu, E., Yu, Q., Anchang, J., 2021. A Role for Drylands in a Carbon Neutral World? Front. Environ. Sci. 9. https://doi.org/10.3389/fenvs.2021.786087.
Hardisky, M., Klemas, V., Smart, R., 1983. Close The influence of soil salinity, growth form and leaf moisture on the spectral radiance of Spartina alterniflora canopies. Photogramm. Eng. Remote Sens. 49, 77–83.
Harris, N.L., Gibbs, D.A., Baccini, A., Birdsey, R.A., de Bruin, S., Farina, M., Fatoyinbo, L., Hansen, M.C., Herold, M., Houghton, R.A., Potapov, P.V., Suarez, D.R., Roman-Cuesta, R.M., Saatchi, S.S., Slay, C.M., Turubanova, S.A., Tyukavina, A., 2021. Global maps of twenty-first century forest carbon fluxes. Nat. Clim. Chang. 11, 234–240. https://doi.org/10.1038/s41558-020-00976-6.
Hawbaker, T.J., Vanderhoof, M.K., Schmidt, G.L., Beal, Y.-J., Picotte, J.J., Takacs, J.D., Falgout, J.T., Dwyer, J.L., 2020. The Landsat Burned Area algorithm and products for the conterminous United States. Remote Sens. Environ. 244, 111801. https://doi.org/10.1016/j.rse.2020.111801.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition. In: Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.
Hijmans, R.J., Bivand, R., Pebesma, E., Sumner, M.D., 2023. terra: Spatial Data Analysis.
Hilker, T., Lyapustin, A.I., Tucker, C.J., Sellers, P.J., Hall, F.G., Wang, Y., 2012. Remote sensing of tropical ecosystems: Atmospheric correction and cloud masking matter. Remote Sens. Environ. 127, 370–384. https://doi.org/10.1016/j.rse.2012.08.035.
Hogland, J., Anderson, N., St. Peter, J., Drake, J., Medley, P., 2018. Mapping forest characteristics at fine resolution across large landscapes of the Southeastern United States using NAIP imagery and FIA field plot data. ISPRS Int. J. Geo Inf. 7, 140. https://doi.org/10.3390/ijgi7040140.
Horn, B.K.P., 1981. Hill shading and the reflectance map. Proc. IEEE 69, 14–47. https://doi.org/10.1109/PROC.1981.11918.
Huang, C., Asner, G.P., Barger, N.N., Neff, J.C., Floyd, M.L., 2010. Regional aboveground live carbon losses due to drought-induced tree dieback in piñon–juniper ecosystems. Remote Sens. Environ. 114, 1471–1479. https://doi.org/10.1016/j.rse.2010.02.003.
Huang, J., Yu, H., Guan, X., Wang, G., Guo, R., 2016. Accelerated dryland expansion under climate change. Nat. Clim. Chang. 6, 166–171. https://doi.org/10.1038/nclimate2837.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.
Hulet, A., Roundy, B.A., Petersen, S.L., Bunting, S.C., Jensen, R.R., Roundy, D.B., 2014. Utilizing National Agriculture Imagery Program Data to Estimate Tree Cover and Biomass of Piñon and Juniper Woodlands. Rangel. Ecol. Manage. 67, 563–572. https://doi.org/10.2111/REM-D-13-00044.1.
Humagain, K., Portillo-Quintero, C., Cox, R.D., Cain, J.W., 2017. Mapping Tree Density in Forests of the Southwestern USA Using Landsat 8 Data. Forests 8, 287. https://doi.org/10.3390/f8080287.
Jean-Romain, 2023. Jean-Romain/lidRplugins.
Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS J. Photogramm. Remote Sens. 173, 24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010.
Key, C.H., Benson, N.C., 1999. The Normalized Burn Ratio (NBR): A Landsat TM radiometric measure of burn severity. United States Geological Survey, Northern Rocky Mountain Science Center.(Bozeman, MT).
Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., Hussin, Y.A., 2014. Generating Pit-free Canopy Height Models from Airborne Lidar [WWW Document]. Doi: info: doi/10.14358/PERS.80.9.863.
Komisarczyk, K., Kozminski, P., Maksymiuk, S., Kapsner, L.A., Spytek, M., Krzyzinski, M., Biecek, P., 2024. treeshap: Compute SHAP Values for Your Tree-Based Models Using the "TreeSHAP" Algorithm.
Krofcheck, D.J., Litvak, M.E., Lippitt, C.D., Neuenschwander, A., 2016. Woody Biomass Estimation in a Southwestern U.S. Juniper Savanna Using LiDAR-Derived Clumped Tree Segmentation and Existing Allometries. Remote Sens. (Basel) 8, 453. https://doi.org/10.3390/rs8060453.
Ku, N.-W., Popescu, S.C., 2019. A comparison of multiple methods for mapping local-scale mesquite tree aboveground biomass with remotely sensed data. Biomass Bioenergy 122, 270–279. https://doi.org/10.1016/j.biombioe.2019.01.045.
Lal, R., 2019. Carbon cycling in global drylands. Curr. Clim. Change Rep. 5, 221–232. https://doi.org/10.1007/s40641-019-00132-z.

Landfire, 2023. LANDFIRE Existing Vegetation Type, v.2.3.0.
Lefsky, M.A., Cohen, W.B., 2003. Selection of Remotely Sensed Data. In: Wulder, M.A., Franklin, S.E. (Eds.), Remote Sensing of Forest Environments: Concepts and Case Studies. Springer, US, Boston, MA, pp. 13–46. https://doi.org/10.1007/978-1-4615-0306-4_2.
Li, X., Wessels, K., Armston, J., Hancock, S., Mathieu, R., Main, R., Naidoo, L., Erasmus, B., Scholes, R., 2023. First validation of GEDI canopy heights in African savannas. Remote Sens. Environ. 285, 113402. https://doi.org/10.1016/j.rse.2022.113402.
Liu, H.Q., Huete, A., 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Trans. Geosci. Remote Sens. 33, 457–465. https://doi.org/10.1109/TGRS.1995.8746027.
Malambo, L., Popescu, S., 2023. Image to Image Deep Learning for Enhanced Vegetation Height Modeling in Texas. Remote Sens. (Basel) 15, 5391. https://doi.org/10.3390/rs15225391.
Maxwell, A.E., Warner, T.A., Vanderbilt, B.C., Ramezan, C.A., 2017. Land Cover Classification and Feature Extraction from National Agriculture Imagery Program (NAIP) Orthoimagery: A Review. Photogramm. Eng. Remote Sens. 83, 737–747. https://doi.org/10.14358/PERS.83.10.737.
McCune, B., Keon, D., 2002. Equations for potential annual direct incident radiation and heat load. J. Veg. Sci. 13, 603–606. https://doi.org/10.1111/j.1654-1103.2002.tb02087.x.
Miller, R.F., Chambers, J.C., Evers, L., Williams, C.J., Snyder, K.A., Roundy, B.A., Pierson, F.B., 2019. The ecology, history, ecohydrology, and management of pinyon and juniper woodlands in the Great Basin and Northern Colorado Plateau of the western United States. Gen. Tech. Rep. RMRS-GTR-403. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 284 p. 403.
Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Comput. Geosci. 32, 1378–1388. https://doi.org/10.1016/j.cageo.2005.12.009.
Mitchell, A.L., Rosenqvist, A., Mora, B., 2017. Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. Carbon Balance Manag. 12, 1–22. https://doi.org/10.1186/s13021-017-0078-9.
Mutanga, O., Masenyama, A., Sibanda, M., 2023. Spectral saturation in the remote sensing of high-density vegetation traits: A systematic review of progress, challenges, and prospects. ISPRS J. Photogramm. Remote Sens. 198, 297–309. https://doi.org/10.1016/j.isprsjprs.2023.03.010.
Pettorelli, N., Schulte to Bühne, H., Tulloch, A., Dubois, G., Macinnis-Ng, C., Queirós, A. M., Keith, D.A., Wegmann, M., Schrodt, F., Stellmes, M., Sonnenschein, R., Geller, G. N., Roy, S., Somers, B., Murray, N., Bland, L., Geijzendorffer, I., Kerr, J.T., Broszeit, S., Leitão, P.J., Duncan, C., El Serafy, G., He, K.S., Blanchard, J.L., Lucas, R., Mairota, P., Webb, T.J., Nicholson, E., 2018. Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward. Remote Sens. Ecol. Conserv. 4, 71–93. 10.1002/rse2.59.
Poulter, B., Frank, D., Ciais, P., Myneni, R.B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y.Y., Running, S.W., Sitch, S., van der Werf, G.R., 2014. Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. Nature 509, 600–603. https://doi.org/10.1038/nature13376.
Prior, E.M., Thomas, V.A., Wynne, R.H., 2022. Estimation of mean dominant height using NAIP digital aerial photogrammetry and lidar over mixed deciduous forest in the southeastern USA. Int. J. Appl. Earth Obs. Geoinf. 110, 102813. https://doi.org/10.1016/j.jag.2022.102813.
PRISM Climate Group, Oregon State University, 2019. 30-Year Normals.
Probst, P., Wright, M.N., Boulesteix, A.-L., 2019. Hyperparameters and tuning strategies for random forest. WIREs Data Min. Knowl. Discovery 9, e1301.
Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sens. Environ. 48, 119–126. https://doi.org/10.1016/0034-4257(94)90134-1.
R Core Team, 2021. R: A language and environment for statistical computing.
Reed, S.C., Feldman, A.F., Hanan, N.P., Moore, D.J.P., Ojima, D.S., Smith, W.K., Wessels, K., Amaral, C., Babst, F., Biederman, J., Litvak, M.E., Macbean, N., Poulter, B., Scott, R.L., Babst-Kostecka, A., Fu, Z., Green, J.K., Kokaly, R.F., Swap, R. J., Serbin, S.P., Tucker, C.J., Wang, L., Watts, J.D., Wolfe, G.M., Flores, A., Leaf, J.R., Washington-Allen, R., Prentice, K., Kachergis, E., Reyes, J., Ryan, J., Sanclements, M., Loescher, H., Leidner, A., Swetnam, T., Cook, B., 2025. The ARID Scoping Study Final Report. ORNLDAAC. https://doi.org/10.3334/ORNLDAAC/2397.
Reinermann, S., Asam, S., Kuenzer, C., 2020. Remote Sensing of Grassland Production and Management—A Review. Remote Sens. (Basel) 12, 1949. https://doi.org/10.3390/rs12121949.
Reinhardt, J.R., Filippelli, S., Falkowski, M., Allred, B., Maestas, J.D., Carlson, J.C., Naugle, D.E., 2020. Quantifying Pinyon-Juniper Reduction within North America's Sagebrush Ecosystem. Rangel. Ecol. Manage. 73, 420–432. https://doi.org/10.1016/j.rama.2020.01.002.
Richardson, J.J., Moskal, L.M., 2014. Uncertainty in urban forest canopy assessment: Lessons from Seattle, WA, USA. Urban For. Urban Green. 13, 152–157. https://doi.org/10.1016/j.ufug.2013.07.003.
Riley, S.J., DeGloria, S.D., Elliot, R., 1999. Index that quantifies topographic heterogeneity. Intermountain J. Sci. 5, 23–27.
Ritz, A.L., Thomas, V.A., Wynne, R.H., Corey Green, P., Schroeder, T.A., Albaugh, T.J., Burkhart, H.E., Carter, D.R., Cook, R.L., Campoe, O.C., Rubilar, R.A., Rakestraw, J., 2022. Assessing the utility of NAIP digital aerial photogrammetric point clouds for estimating canopy height of managed loblolly pine plantations in the southeastern United States. Int. J. Appl. Earth Obs. Geoinf. 113, 103012. https://doi.org/10.1016/j.jag.2022.103012.
Roberts, D.W., Cooper, S.V., 1989. Concepts and techniques of vegetation mapping. Land classifications based on vegetation: applications for resource management 90–96.
Roberts, K.C., Lindsay, J.B., Berg, A.A., 2019. An Analysis of ground-point classifiers for terrestrial LiDAR. Remote Sens. (Basel) 11, 1915. https://doi.org/10.3390/rs11161915.
Romme, W.H., Allen, C.D., Bailey, J.D., Baker, W.L., Bestelmeyer, B.T., Brown, P.M., Eisenhart, K.S., Floyd, M.L., Huffman, D.W., Jacobs, B.F., Miller, R.F., Muldavin, E. H., Swetnam, T.W., Tausch, R.J., Weisberg, P.J., 2009. Historical and modern disturbance regimes, stand structures, and landscape dynamics in piñon–juniper vegetation of the Western United States. Rangel. Ecol. Manage. 62, 203–222. https://doi.org/10.2111/08-188R1.1.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
Rouse, J.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS.
Roussel, J.-R., documentation), D.A. (Reviews the, features), F.D.B. (Fixed bugs and improved catalog, segment_snags), A.S.M. (Implemented wing2015 for, track_sensor), B.J.-F. (Contributed to R. for, track_sensor), G.D. (Implemented G. for, 2020. lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications.
Sankey, T., Glenn, N., 2011. Landsat-5 TM and lidar fusion for sub-pixel juniper tree cover estimates in a western rangeland. Photogramm. Eng. Remote Sens. 77, 1241–1248. https://doi.org/10.14358/PERS.77.12.1241.
Sankey, T., Shrestha, R., Sankey, J.B., Hardegree, S., Strand, E., 2013. Lidar-derived estimate and uncertainty of carbon sink in successional phases of woody encroachment. J. Geophys. Res. Biogeo. 118, 1144–1155. https://doi.org/10.1002/jgrg.20088.
Schroeder, T.A., Obata, S., Papeş, M., Branoff, B., 2022. Evaluating statewide NAIP photogrammetric point clouds for operational improvement of national forest inventory estimates in mixed hardwood forests of the Southeastern U.S. Remote Sens. (Basel) 14, 4386. https://doi.org/10.3390/rs14174386.
Senf, C., 2022. Seeing the system from above: the use and potential of remote sensing for studying ecosystem dynamics. Ecosystems 25, 1719–1737. https://doi.org/10.1007/s10021-022-00777-2.
Sha, Y., Ii, D.J.G., West, G., Stull, R., 2020. Deep-Learning-Based Gridded Downscaling of Surface Meteorological Variables in Complex Terrain. Part I: Daily Maximum and Minimum 2-m Temperature. Doi: 10.1175/JAMC-D-20-0057.1.
Shendryk, Y., 2022. Fusing GEDI with earth observation data for large area aboveground biomass mapping. Int. J. Appl. Earth Obs. Geoinf. 115, 103108. https://doi.org/10.1016/j.jag.2022.103108.
Smith, W.K., Dannenberg, M.P., Yan, D., Herrmann, S., Barnes, M.L., Barron-Gafford, G. A., Biederman, J.A., Ferrenberg, S., Fox, A.M., Hudson, A., Knowles, J.F., MacBean, N., Moore, D.J.P., Nagler, P.L., Reed, S.C., Rutherford, W.A., Scott, R.L., Wang, X., Yang, J., 2019. Remote sensing of dryland ecosystem structure and function: Progress, challenges, and opportunities. Remote Sens. Environ. 233, 111401. https://doi.org/10.1016/j.rse.2019.111401.
Snyder, G.I., 2012. The 3D Elevation Program: summary of program direction (USGS Numbered Series No. 2012–3089), Fact Sheet. U.S. Geological Survey, Reston, VA.
Sunde, M.G., Diamond, D.D., Elliott, L.F., Hanberry, P., True, D., 2020. Mapping high-resolution percentage canopy cover using a multi-sensor approach. Remote Sens. Environ. 242, 111748. https://doi.org/10.1016/j.rse.2020.111748.
Tolan, J., Yang, H.-I., Nosarzewski, B., Couairon, G., Vo, H.V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., Vamaraju, J., Moutakanni, T., Bojanowski, P., Johns, T., White, B., Tiecke, T., Couprie, C., 2024. Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. Remote Sens. Environ. 300, 113888. https://doi.org/10.1016/j.rse.2023.113888.
Tompalski, P., 2023. lidRmetrics.
Wagner, F.H., Roberts, S., Ritz, A.L., Carter, G., Dalagnol, R., Favrichon, S., Hirye, M.C. M., Brandt, M., Ciais, P., Saatchi, S., 2024. Sub-meter tree height mapping of California using aerial images and LiDAR-informed U-Net model. Remote Sens. Environ. 305, 114099. https://doi.org/10.1016/j.rse.2024.114099.
Wang, T., Hamann, A., Spittlehouse, D., Carroll, C., 2016. Locally downscaled and spatially customizable climate data for historical and future periods for North America. PLoS One 11, e0156720. https://doi.org/10.1371/journal.pone.0156720.
Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., Grehan, A.J., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Mar. Geod. 30, 3–35. https://doi.org/10.1080/01490410701295962.
Wu, Z., Dye, D., Vogel, J., Middleton, B., 2016. Estimating forest and woodland aboveground biomass using active and passive remote sensing. Photogramm. Eng. Remote Sens. 82, 271–281. https://doi.org/10.14358/PERS.82.4.271.
Xu, Y., Franklin, S.B., Wang, Q., Shi, Z., Luo, Y., Lu, Z., Zhang, J., Qiao, X., Jiang, M., 2015. Topographic and biotic factors determine forest biomass spatial distribution in a subtropical mountain moist forest. For. Ecol. Manage. 357, 95–103. https://doi.org/10.1016/j.foreco.2015.08.010.
Yang, J., Weisberg, P.J., Bristow, N.A., 2012. Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. Remote Sens. Environ. 119, 62–71. https://doi.org/10.1016/j.rse.2011.12.004.

Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth Surf. Proc. Land. 12, 47–56. https://doi.org/10.1002/esp.3290120107.
Zhang, G., Lu, Y., 2012. Bias-corrected random forests in regression. J. Appl. Stat. 39, 151–160. https://doi.org/10.1080/02664763.2011.578621.
Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sens. Environ. 128, 289–298. https://doi.org/10.1016/j.rse.2012.10.017.