International Journal of Applied Earth Observation and Geoinformation 122 (2023) 103435

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Comparing harmonic regression and GLAD Phenology metrics for estimation of forest community types and aboveground live biomass within forest inventory and analysis plots

Aaron E. Maxwell [a,*], Barry T. Wilson [b], Justin J. Holgerson [b], Michelle S. Bester [a]

[a] West Virginia University Department of Geology and Geography, Morgantown, WV 26505, United States
[b] Northern Research Station, USDA Forest Service, St. Paul, MN 55108, United States

ARTICLE INFO

Keywords:
Landsat
Harmonic regression
Time series analysis
Phenology
Forest types
Aboveground live biomass

ABSTRACT

The Landsat multispectral time series is a valuable source of moderate spatial resolution data to support forest mapping and monitoring tasks. Using United States Department of Agriculture (USDA) Forest Service Forest Inventory and Analysis (FIA) plots within the states of Michigan, Oregon, and West Virginia, two methods to summarize time series observations, harmonic regression coefficients and Global Land Analysis & Discovery (GLAD) Phenology metrics, are compared for predicting forest community type, total aboveground live biomass (AGLBM), and species-specific AGLBM. Harmonic regression coefficients, which provided mean overall accuracies (OAs) between 62.8% and 73.1% and map image classification efficacies (MICEs) varying between 0.455 and 0.566 for the three studied states and calculated using multiple model replicates, generally provided better predictive performance for differentiating forest community types in comparison to GLAD phenology metrics. However, differences were not always statistically significant. DTM-derived terrain variables improved classification performance in some landscapes when using machine learning, random forest classification (for example, the highest obtained mean OA of 78.5% (MICE = 0.557) was obtained for West Virginia when using the combined harmonic regression and terrain variables). For the regression-based estimation of total and species-specific AGLBM, the spectral data used and the incorporation of terrain variables had less of an impact on model performance. Further, the regression models in Oregon provided larger mean R-squared values (0.452 to 0.490) in comparison to those in Michigan and West Virginia, where all R-squared values were lower than 0.200, suggesting that AGLBM prediction may be more or less challenging in different landscapes depending on forest characteristics, terrain, management practices, and disturbance histories.

1. Introduction

The multispectral data provided by the Landsat program, which began in the 1970s and continues to the present day, have many applications for forest mapping, monitoring, and assessment at a moderate spatial resolution (Landsat Legacy Project Team, 2017). Applications include monitoring and mapping forest fires and associated effects (e.g., (French et al., 2008; Parker et al., 2015; Soverel et al., 2010)), deforestation and fragmentation patterns (e.g., (Hansen et al., 2013; Margono et al., 2012; Schultz et al., 2016; Shimabukuro et al., 1998)), forest health and the impact of blights or disease (e.g., (Pasquarella et al., 2017; Royle and Lathrop, 1997; Vogelmann et al., 2009)), forest community types (e.g., (Beaubien et al., 1979; Bryant et al., 1980; Pasquarella et al., 2018)), and other forest attributes (e.g., species composition, biomass, and leaf area index) (Chen and Cihlar, 1996; Fassnacht et al., 1997; Huang et al., 2009; López-Serrano et al., 2020; Potapov et al., 2021). Forests are essential ecosystems in regards to the biodiversity they support; their roles in local-, regional-, and global-scale energy and chemical cycles, including the carbon and nitrogen cycles; their modulation of global climate change; and the ecosystem services they provide (Binkley, 2021; Bonan, 2008; Franklin et al., 2018; Jackson et al., 2008; Perry et al., 2008; Smail, 2010). As a result, continued research on the use of remotely sensed data to map, monitor, and manage forests is of great importance.

In 2008, the Landsat data archive was made open and freely available, which spurred a rapid increase in the download volume and use of

* Corresponding author.
E-mail address: Aaron.Maxwell@mail.wvu.edu (A.E. Maxwell).

https://doi.org/10.1016/j.jag.2023.103435
Received 12 April 2023; Received in revised form 25 June 2023; Accepted 24 July 2023

these data. For example, between 1972 and 2008, less than 3,000 scenes were sold per month. In contrast, and following the opening of the archive, nearly one million scenes were downloaded in 2009 with a subsequent 20-fold increase in downloads by 2017 (Zhu et al., 2019). The availability of the full data archive has also allowed easier analysis of multiple scenes collected on different dates, as opposed to individual scenes. This has fueled research on change detection, anomaly detection, time series analysis, and vegetation phenology (Adams et al., 2020; Hansen et al., 2016; Pasquarella et al., 2018, 2017; Potapov et al., 2021, 2020, 2008; Wilson et al., 2018). Such analyses are also supported by the availability of derived products and analysis ready data (ARD) (Dwyer et al., 2018; Wulder et al., 2019). Further, cloud computing platforms, such as Google Earth Engine (GEE) (Gorelick et al., 2017; Hansen et al., 2013) and the Microsoft Planetary Computer (Simoes et al., 2021; Xu et al., 2022), allow for the analysis of large volumes of time series data without downloading and locally processing individual scenes.

Multiple methods are available to generate derived metrics from a single image, such as using band indices (Lillesand and Kiefer, 1979), principal component analysis (Estornell et al., 2013; Lillesand and Kiefer, 1979), or tasseled cap transformations (Crist and Cicone, 1984; Lastovicka et al., 2020; Lillesand and Kiefer, 1979). Similarly, multiple methods can be used to derive a set of metrics from a time series of observations. Methods include generating monthly or seasonal composites (Adams et al., 2020; Flood, 2013), fitting models to detect anomalies or abrupt changes in surface reflectance (Zhu and Woodcock, 2014), summarizing the phenology curve (Nijland et al., 2016; Pasquarella et al., 2018; Rose and Nagle, 2021), or fitting a time series function that can capture patterns or oscillations at different temporal scales (e.g., monthly and seasonally) (Adams et al., 2020; Pasquarella et al., 2018; Wilson et al., 2018). Thus, it is important to consider what methods should be used to characterize or generate a set of derived metrics from the time series to support specific mapping and/or modeling needs.

This study compares Landsat-derived harmonic regression (Sellers et al., 1996, 1994; Wilson et al., 2018) and Global Land Analysis & Discovery (GLAD) Phenology Metrics Type A datasets for (1) differentiating broad forest community types, (2) estimating total aboveground live biomass (AGLBM), and (3) estimating species-specific AGLBM. Harmonic regression, which is described in more detail below, allows for estimating a periodic signal using a Fourier series of varying order or number of harmonics. In the context of multispectral data, this periodic signal is generally associated with a spectral reflectance band or derivative, such as a vegetation index (Sellers et al., 1996, 1994; Wilson et al., 2018). The Global Land Analysis & Discovery (GLAD) Phenology Type A ARD product, which is also described in more detail below offers a globally consistent characterization of land surface phenology as captured by the Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI) sensors (Potapov et al., 2020). Training and validation data in this study were derived from the United States Department of Agriculture (USDA) Forest Service Forest Inventory and Analysis (FIA) plots within the states of Michigan, Oregon, and West Virginia in the United States (USA).

2. Methods

2.1. Study areas and field data

In order to generalize the results of the study, three separate study areas were used: the states of Michigan, Oregon, and West Virginia, USA (Fig. 1). These states were chosen based on variability in landscape and forest characteristics and also due to the ability to establish data sharing

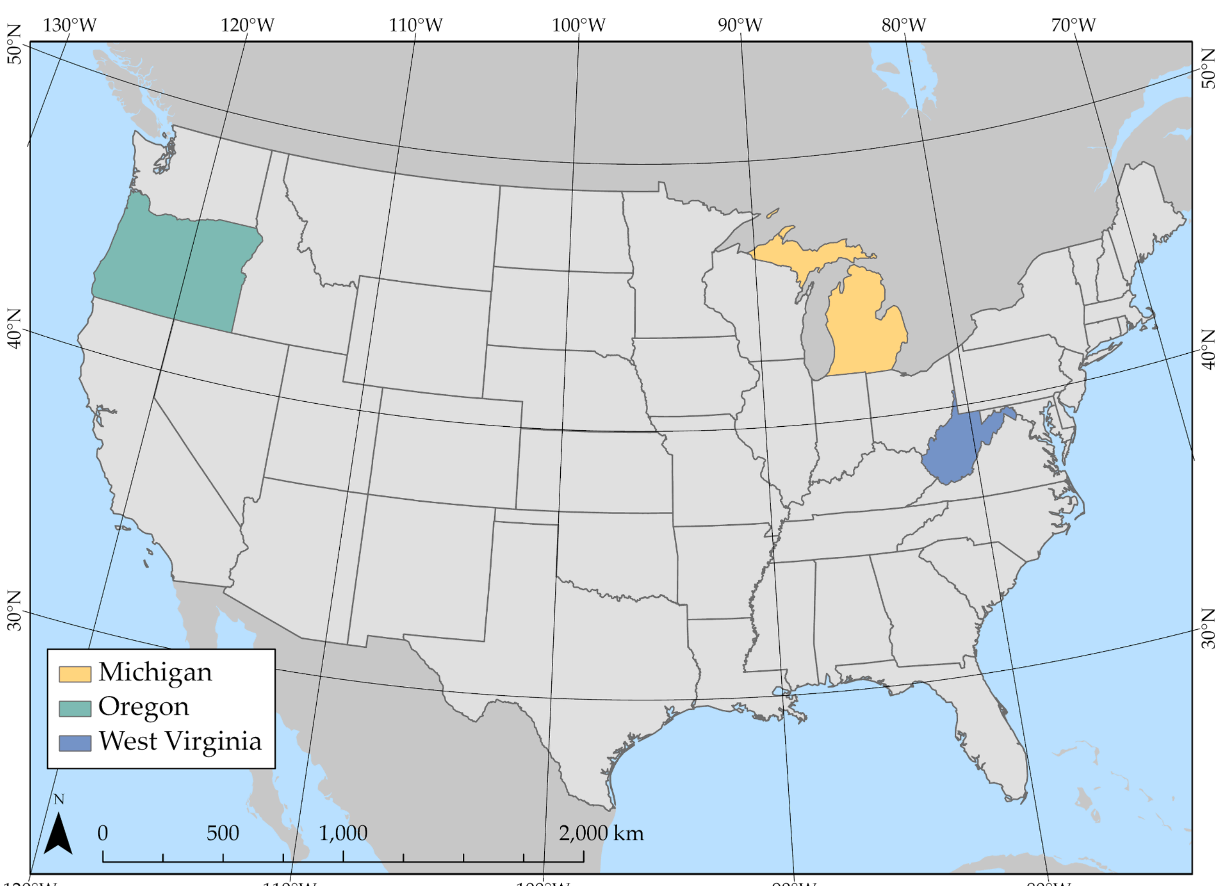

Fig. 1. Study areas as the states of Michigan, Oregon, and West Virginia in the United States. Map projection is USA Contiguous Albers Equal Area Conic.

agreements with the research stations in which they occur. The state of Michigan occurs at an average latitude of around 45°N and has elevations ranging from 174 to 603 m with an average elevation of 270 m. It experiences a continental climate with hot, humid summers, cold winters, and four distinct seasons. Temperatures range from a winter average of around −10 °C to a summer average of around 26 °C, and mean annual precipitation is between 760 and 1,020 mm. Landforms and surficial conditions in the state have been heavily influenced by Pleistocene glaciation. Oregon occurs at an average latitude of 44°N, has an elevation range of 0 to 3,329 m, and has an average elevation around 1,000 m. Climates are variable across the state with marked differences in both temperature and rainfall; the following climate types occur within the state: oceanic, Mediterranean, steppe, subarctic, and high desert. The state is mountainous with active tectonic activity and volcanism. West Virginia occurs within the Appalachian region of North America, has a mean latitude of around 39°N, and elevations ranging from 73 to 1,482 m with a mean elevation of 461 m. It experiences a humid subtropical climate with warm to hot and humid summers and cold winters. Temperatures average between −4 °C in the winter and 24 °C in the summer. Average annual precipitation ranges from 810 to 1,400 mm. In contrast to Michigan, Pleistocene glacial deposits are absent. Forests across all three states have been impacted by European settlement, agriculture, and timber harvesting (Hardwick et al., 2013). Forest communities occurring in each state are discussed below.

The FIA program collects, analyzes, and reports information about the nations' forests on an annual basis in order to document both forest status and change (Forest Inventory and Analysis National Program, n. d.). This requires the establishment and repeated field surveying of a large set of plots occurring on both public and private lands. To select plot locations, the country is divided into 6,000 acre (2,428 ha) hexagons, and a random coordinate is selected within each hexagon. Each individual plot consists of four subplots, one at the center and the other three positioned such that its center is 120 feet (36.6 m) from the center of the central plot and at an azimuth of 0°, 120°, or 240° (Bechtold and Patterson, 2005). Since the correct coordinates of FIA plots are not made publicly available and because we needed to extract pixel values at the plot locations, a data request was made to the USDA Forest Service Pacific Northwest Research Station to obtain data for Oregon and the Northern Research Station to obtain data for Michigan and West Virginia. Extraction of data at plot locations was conducted by the Spatial Data Services group of the FIA program, and tables were provided to the researchers without associated plot coordinate information in order to maintain data integrity and privacy.

Once data tables were provided, we filtered the data so that only plots that had a stocked forested condition status, had a forest type group defined, were single-condition, and had a non-zero reported total AGLBM were used in the analysis. This resulted in 1,642 plots for Michigan, 2,605 for Oregon, and 891 for West Virginia. A wide variety of forest type groups are differentiated by the FIA program. To reduce the number of categories for the classification assessment and to remove categories that were rare or had few samples, forest type groups were merged into a smaller number of categories, termed forest community types in this study. Forest type groups that were not abundant and could not be combined with other classes were assigned to an "Other" class. Table 1 provides the forest community types that were defined and the count of samples in each grouping per state. For the estimation of a continuous variable, we explored total AGLBM and a species-specific AGLBM. For Oregon, Douglas Fir (*Pseudotsuga menziesii*) was selected due to its abundance. Also due to abundance, White Oak (*Quercus alba*) was selected in Michigan and West Virginia. Fig. 2 below shows the distribution of total AGLBM in each state within the used plots in units of tons per acre. The mean total AGLBM for Michigan, Oregon, and West Virginia were 11.2, 20.2, and 19.0 tons per acre, respectively, while the standard deviations were 8.3, 23.8, and 12.3, respectively. The mean species AGLBM for Michigan, Oregon, and West Virginia were 0.2, 16.3, and 1.6 tons per acre, respectively, while the standard deviations were 1.4, 24.3, and 4.8, respectively.

Table 1
Forest classification/groupings of forest type groups and associated count of plots for each state.

State	Forest Community Type	Count
Michigan	Aspen/Birch	229
	Elm/Ash	110
	Spruce/Fir	216
	Maple/Mixed	562
	Oak Dominant	279
	Pine Dominant	234
	Other	12
Oregon	Alder/Maple	29
	Douglas Fir	857
	Spruce/Fir	420
	Hemlock/Sitka	75
	Lodgepole Pine	236
	Oak Dominant	44
	Tanoak/Laurel	54
	Larch	32
	Other	848
West Virginia	Elm/Ash	7
	Loblolly	9
	Maple/Mixed	242
	Oak Dominant	591
	Oak/Pine	18
	Spruce/Fir	8
	Pine Dominant	3
	Other	13

2.2. Harmonic regression metrics

Equation 1 represents a harmonic regression implementation where β_0 is the *y*-intercept, β_1 is a slope term that allows for including a long term increasing or decreasing linear trend (i.e., non-stationarity), and β_{2i} through β_{3i} represent coefficients associated with the cosine and sine terms that describe the oscillating signal. Two coefficients are estimated for each order or harmonic. Incorporating multiple harmonics allows for modeling multiple periodic signals or more complex patterns occurring over different time intervals (e.g., yearly or seasonally); however, adding many harmonics can result in overfitting. Similar to traditional regression methods, harmonic regression coefficients can be estimated using the ordinary least squares (OLS) method. The variable *t* represents the time for which a prediction is being made while *T* is the number of time divisions for one year or the largest temporal pattern being modeled. The number of harmonics included in the model is *n* while the current harmonic is denoted as *i*. Lastly, $\hat{y}$ is the estimated value of the measure of interest, such as the predicted normalized difference vegetation index (NDVI), at a given time (*t*) (Sellers et al., 1996, 1994; Shumway et al., 2000; Wilson et al., 2018). Fig. 3 conceptualizes two harmonic regression models with one and three harmonics. Note how adding additional harmonics allows for modeling more complex periodic patterns in the data.

$$\hat{y} = \beta_0 + \beta_1 2\pi t + \sum_{i=1}^{n}\left(\beta_{2i}\cos\left(\frac{i2\pi t}{T}\right) + \beta_{3i}\sin\left(\frac{i2\pi t}{T}\right)\right) \text{(Equation 1).}$$

In this study, harmonic regression coefficients were calculated using Landsat 8 OLI data at all plot locations. Specifically, all cloud free pixel observations at the plot location were extracted for all data collected between 2014 and 2018; the tasseled cap transformation (Baig et al., 2014; Crist and Cicone, 1984) was performed from the surface reflectance data to derive brightness, greenness, and wetness bands; and a 3rd order harmonic series was fit to the three bands separately using the available clear observations and the OLS method. The three harmonics represent oscillations occurring at the twelve-, six-, and four-month periods. This resulted in a total of 21 metrics: the sine and cosine coefficients for each harmonic and the *y*-intercept term calculated using the brightness, greenness, and wetness bands separately. The signal was assumed to be stationary, so no long-term trend was included in the model.

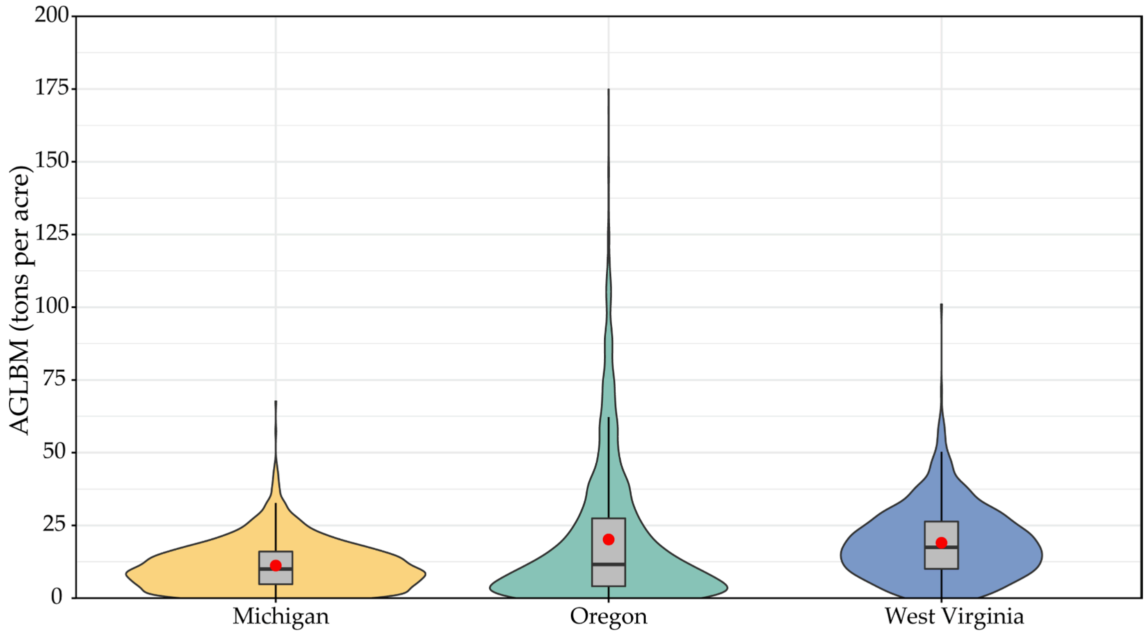

Fig. 2. Distribution of total aboveground live biomass (AGLBM) within each state for all used plot locations. Units are tons per acre. Red dot represents the mean AGLBM. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

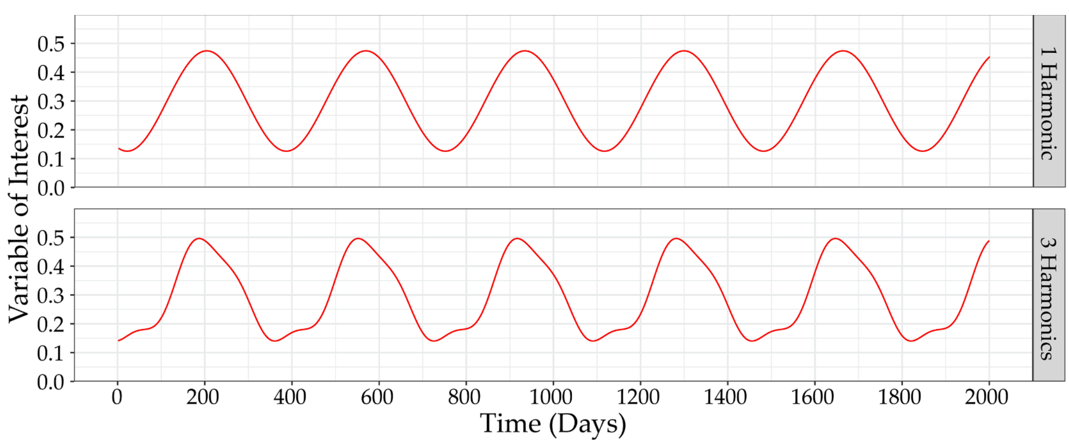

Fig. 3. Conceptualization of harmonic regression with varying number of harmonics.

Several studies have documented the value of harmonic regression for deriving predictor variables from a time series of multispectral data for estimating forest attributes. Wilson et al. (2018) proposed the use of harmonic regression coefficients derived from Landsat tasseled cap transformation bands for classification of forest types and estimation of forest metrics within FIA plots, including number of trees, basal area, and AGLBM. They documented a two- to three-fold improvement in explained variance for regression models and a 10% to 20% increase in classification accuracy in comparison to using monthly Landsat image composites. Adams et al. (2020) documented improved performance for differentiating forest community types and gradient compositions using Landsat-derived harmonic regression coefficients calculated from tasseled cap bands and the enhanced vegetation index (EVI) compared to Landsat-derived seasonal composites. Similarly, Pasquarella et al. (2018) noted better performance for differentiating forest community types using harmonic regression compared to seasonal composites and single-date images. They attributed this to improved separation between deciduous and conifer forest community types, which had distinguishable amplitudes in the yearly oscillations.

2.3. *GLAD Phenology metrics*

The methods used to generate GLAD Phenology metrics products are described by Potapov et al. (2020) while full documentation is available on the GLAD website (https://glad.umd.edu/ard/glad-landsat-ard-tools). First, all Landsat 5 TM, Landsat 7 ETM+, and Landsat 8 OLI images from 1997 to the present are categorized into quality tiers, and only those meeting the highest geometric and radiometric quality are used (i.e., Tier 1 scenes). Scenes including seasonal snow cover are also excluded. Remaining scenes are processed to estimate surface reflectance then combined to generate 16-day composites using the best available data. Prior to generating metrics relative to a selected processing year, gaps, such as those resulting from cloud cover, are filled using observations from prior years. If no gap-free data are available at a pixel location from prior years, linear regression is used to fill gaps using near pixel values. The goal is to generate a globally consistent and gap-filled ARD product with a regular 16-day time interval using the best available data (Potapov et al., 2020).

Once the 16-day composites are generated and gaps are filled, a large set of metrics are generated to summarize phenology conditions relative to a specified year. Metric sets include Type A, B, and C. We used the Type A metrics in this study since more measures are generated in comparison to the Type B set and because the Type C metrics were not available at the time that a data request was made to the FIA program. Further, metrics were derived relative to the year 2018. All 16-day composites from 2015 to 2018 were downloaded, and data from 2015 through 2017 were used to fill gaps in the 2018 data. A Perl (The Perl Programming Language - https://www.perl.org," n.d.) script provided by the GLAD program (Potapov et al., 2020) was used to generate the 341 metrics from the 16-day composites.

We made use of a total of 341 metrics, which are summarized in Table 2. In the GLAD ARD Type A product, metrics are calculated from the six spectral reflectance bands (blue, green, red, near infrared (NIR), shortwave infrared 1 (SWIR1), and shortwave infrared 2 (SWIR2)) and a set of band ratios derived from these bands. The spectral variability vegetation index (SVVI) after Coulter et al. (2016) is also included, which is calculated by subtracting the sum of the standard deviations of the NIR, SWIR1, and SWIR2 bands for the available image dates from the sum of the standard deviations of all included bands (Coulter et al., 2016). Summary statistics are then derived from the rankings of the data either defined using the current spectral band or index or a corresponding variable. Corresponding variables used are the NDVI, SVVI, and brightness temperature, which are derived from the Landsat thermal data (Potapov et al., 2020).

Although these data are globally available, they have seen limited application for estimating forest conditions or attributes. Potapov et al. (2021) combined the Type C metric set with canopy height estimates derived from the Global Ecosystem Dynamics Investigation (GEDI) light detection and ranging (LiDAR) instrument to generate a prediction of canopy heights for all forested Landsat pixels globally (Potapov et al., 2021). In a prior study, some of the authors of the current study compared the GLAD Phenology Type C metrics to harmonic regression coefficients and seasonal band medians, all of which were derived from Landsat data, to differentiate forest community types for the entire state of West Virginia, USA. We observed the best performance when combining harmonic regression coefficients with terrain variables derived from National Elevation Dataset (NED) digital terrain models (DTMs). The GLAD metrics generally outperformed seasonal composites but did not provide an accuracy comparable to the harmonic regression results (Hartley et al., 2022). This study expands upon our prior study by incorporating a large set of ground measurements from the FIA program; exploring both the classification of forest community types and the prediction of total and species-specific AGLBM; and investigating multiple states with varying ecological, environmental, and landscape alteration conditions.

Table 2
Summary of GLAD Phenology metrics Type A. Summary metrics are calculated for spectral bands or band indices based on rank relative to the current variable or a corresponding variable.

Spectral Band or Index	Ranking Variable	Summary Metrics
Blue	Self	Minimum (min)*
Green		Maximum (max)*
Red	NDVI	Second lowest value (smin)*
NIR		Second highest value (smax)*
SWIR1	SVVI	Median
SWIR2		Mean smin to Q2*
NDVI	Brightness	Mean Q2 to smax*
(NIR-SWIR1)/(NIR + SWIR1)	Temperature	Mean min to Q1*
		Mean Q3 to max*
(Blue-Green)/(Blue + Green)		Mean Q1 to Q3
		Mean of all
(Blue-Red)/(Blue + Red)		Mean of all values except min and max
(Blue-NIR)/(Blue + NIR)		Amplitude (max – min)*
		Amplitude (smax – smin)*
(Green-Red)/(Green + Red)		Amplitude (Mean Q2 to max – Mean Q2 to min)*
(Green-NIR)/(Green + NIR)		Amplitude (Mean Q3 to max – Mean Q2 to min)*
(SWIR1-SWIR2)/(SWIR1 + SWIR2)		
SVVI		

* indicates metrics that are calculated when defining ranks using a corresponding variable; all metrics are calculated when determining ranks using the current variable. NIR = near infrared; SWIR1 = shortwave infrared 1; SWIR2 = shortwave infrared 2; NDVI = normalized difference vegetation index; SVVI = spectral variability vegetation index; Q1 = first quartile; Q2 = second quartile; Q3 = third quartile.

2.4. Digital terrain variables

Digital terrain variables derived from DTMs have generally been shown to be useful for forest type differentiation in some landscapes (see, for example, (Adams et al., 2020; Hartley et al., 2022; Pasquarella et al., 2018)) since they provide landscape information associated with relative topographic position, steepness, slope orientation, surface roughness, and incoming solar radiation, which can impact site characteristics and forest compositions and characteristics (Franklin, 2020; Maxwell and Shobe, 2022). Terrain metrics can be especially useful when forest classes may have limited spectral differences but favor different landscape conditions (Adams et al., 2020; Hartley et al., 2022). As a result and using ArcGIS Pro (Download ArcGIS Pro—ArcGIS Pro | Documentation, n.d.) and custom Python/ArcPy scripts (ArcGIS Pro Python reference—ArcGIS Pro | Documentation, n.d.; Welcome to Python.org, n.d.), we generated a set of digital terrain variables from the 1/3rd arc-second (10 m) spatial resolution National Elevation Dataset (NED) DTM data (Gesch et al., 2002). Table 3 below summarizes the twelve calculated metrics, which were selected to capture a variety of terrain characteristics. In order to match the spatial resolution of the Landsat-derived variables, all metrics were calculated at a 10 m spatial resolution then resampled to a 30 m spatial resolution using pixel aggregation. For variables that rely on moving window-based calculations, a circular window with a 3 cell, or 30 m, radius was used.

2.5. Modeling, Validation, and comparison

Random forest (Breiman, 2001), a machine learning algorithm, was used to perform the classification task of differentiating forest community types and also the regression tasks of estimating total and species-

Table 3
Terrain derivatives used as predictor variables in the study with associated descriptions or equations. Z = elevation; Slp = topographic slope; Asp = topographic aspect; σ^2 = variance.

Variable	Description/Equation
Linear Aspect	$270 - \frac{360}{2\pi} \times \arctan^2(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y})$
Elevation	Bare-ground surface height
Topographic Dissection Index	$\frac{z - z_{min}}{z_{max} - z_{min}}$
Heat Load Index	Index associated with annual direct incoming solar radiation taking into account latitude, slope, and aspect
Topographic Roughness Index	$\sigma^2(z)$
Surface Area Ratio	$\frac{\text{Cell Size}^2}{\text{Cos(Slope in Degrees)}}$
Site Exposure Index	$Slp \times \cos(\pi\frac{Asp - 180}{180})$
Slope (Degrees)	$\arctan\left(\sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\right)\left(\frac{180}{\pi}\right)$
Mean Slope	Calculates slope within a moving window
Slope Position	$z\text{-}z_{mean}$
Surface Relief Ratio	$\frac{z_{mean} - z_{min}}{z_{max} - z_{min}}$
Topographic Radiation Aspect Index	$\frac{1 - \cos\left(\left(\frac{\pi}{180}\right) \times (Asp - 30)\right)}{2}$

specific AGLBM. It is a nonparametric, ensemble decision tree method in which (1) each tree in the ensemble uses a random subset of the training data, which are selected using bootstrapping or random sampling with replacement, and (2) each decision rule or binary split is based on a single variable selected from a random subset of all available predictor variables. The goal of these augmentations is to reduce the correlation between the decision trees in the ensemble by providing each tree with a different subset of the training samples and limiting what variables are available to partition the data. The goal is to train a set of weak models that are collectively strong due to reduced correlation (Breiman, 2001).

Instead of randomly dividing the data into separate training and testing sets and in order to be able to train multiple models and assess the variability in performance, a total of 20 models were generated for each feature space combination. A total of four feature spaces were tested: harmonic regression coefficients (H), harmonic regression coefficients plus terrain variables (HT), GLAD metrics (G), and GLAD metrics plus terrain variables (GT). Each replicate used a different random split of the available samples into training and testing partitions with 70% of the samples used to train the model and 30% reserved for assessment. When predicting the forest community types, we performed stratified random sampling to maintain the relative proportions of each class in the training and testing sets. For each separate model we also tuned the number of random variables available for splitting at each node (*mtry*) hyperparameter using 10-fold cross-validation and the random training samples selected for that specific model. For the classification problem, the *mtry* parameter that provided the best Kappa statistic was selected. For the regression problems, the *mtry* parameter that provided the best root mean square error (RMSE) was selected. All models were trained using the ranger (Wright and Ziegler, 2017) and caret (Kuhn, 2021) packages in the R language and data science environment (R Core Team, 2020), and 500 trees were used in each ensemble. For the classification problem and in order to combat the negative impacts of class imbalance, classes were weighted based on the inverse of their abundance in the dataset while training the models.

Due to the large number of variables that constitute the GLAD Phenology Type A metrics, we also conducted principal component analysis (PCA) (Estornell et al., 2013) separately for each model replicate and using only the training data for that model run to avoid a data leak. A total of 15 principal component bands were calculated, which was found to be adequate to capture more than 99% of the variability in the original 341 variables. Since PCAs were calculated separately for each model run, loadings differed for each replicate.

For the forest community type classification problem, model assessment was performed using the withheld testing data. Each replicate was evaluated separately using different withheld testing sets in order to capture the variability in predicted performance. The classifications were assessed using overall accuracy (OA), which is calculated as the proportion of the testing samples that were correctly classified (Congalton and Green, 2019), and map image classification efficacy (MICE) (Shao et al., 2021). MICE, in contrast to the more traditional Kappa metric which has been suggested to have logical flaws and for which its use is being discouraged, uses only the reference class margin totals to adjust OA for chance agreement. Kappa uses both the reference and classification margin totals. MICE is robust to accuracy inflation due to chance agreement, and it is particularly informative when the number of samples per class are imbalanced (Shao et al., 2021). Lastly, we report the percent of the validation samples in which the correct forest community type was within the top three predicted probabilities (top3). Metrics were calculated using the caret (Kuhn, 2021) and rfUtilities (Evans and Murphy, 2018) packages in R (R Core Team, 2020).

Also for the forest community type classification problem, we assessed the predicted class probabilities using area under the receiver operating characteristic curve (AUC ROC) and area under the precision-recall curve (AUC PRC). The ROC curve considers class producer's accuracy (1 – omission error) at multiple binary thresholds. It can be overly optimistic when classes are imbalanced. In contrast, the precision-recall curve takes into account the producer's and user's accuracy (1 – commission error), defined relative to the positive case, and can be especially useful when classes are imbalanced because user's accuracy is impacted by class proportions (Fan et al., 2006; Saito and Rehmsmeier, 2015; Tharwat, 2020). Since more than two classes are differentiated in this study, multi-class AUC ROC and AUC PR were calculated using the multiROC R package (Wei and Wang, 2018). Micro-average AUCs were calculated by aggregating true positive (TP), true negative (TN), false positive (FP), and false negative (FN) samples across all classes, which is more sensitive to class imbalance than the alternative macro-averaging method, which averages the metric for each class and gives each class equal weight in the calculation (Wei and Wang, 2018).

For the regression assessment, we used the RMSE and prediction R-squared metrics. RMSE is calculated in the units of the variable being predicted, in this case tons per acre, and represents the average absolute residual error. R-squared represents the proportion of variance in the response variable that is explained by the model and is unitless (James et al., 2013). We made use of the Metrics package (Hamner and Frasco, 2018) in R.

Using the 20 model replicates for each feature space, we assessed the differences in performance using a pairwise Kruskal-Wallis test. This nonparametric method compares differences between groups using rank order and adjusts *p*-values for multiple comparisons (Dalgaard, 2002). For each pairwise comparison, the null hypothesis is that the median values of the two groups are equal while the alternative is that the median values are not equal. For the forest type classification problem, we compared the MICE metrics while we compared RMSE for the total and species-specific AGLBM predictions. A 95% confidence interval and a 0.05 alpha were used.

3. Results and discussion

3.1. Forest type classification

Table 4 below provides the mean performance metrics for each state and for each feature space, which was calculated from the 20 model replicates, while Fig. 4 shows the distribution of OA and MICE for each feature space combination within each state. In regards to models that did not incorporate terrain metrics, the harmonic regression coefficients (H) generally outperformed the predictions made using the 15 principal components calculated from the GLAD metrics (G). OA differences were 0.031, 0.016, and 0.037 for Michigan, Oregon, and West Virginia, respectively. Similar trends were noted for the other metrics. When adding the terrain metrics, improvements were seen in comparison to

Table 4
Mean assessment results for 20 replicates of forest type classification using four different feature sets. OA = overall accuracy; MICE = map image classification efficacy; top3 = proportion of samples in which the correct forest type was in the top three predicted class probabilities; ROC AUC = area under the receiver operating characteristic curve; ROC PRC = area under the precision-recall curve; G = GLAD; GT = GLAD plus terrain derivatives; H = harmonic regression; HT = harmonic regression + terrain derivatives.

State	Feature Set	OA	MICE	top3	AUC ROC	AUC PRC
Michigan	G	0.597	0.492	0.894	0.900	0.671
	GT	0.620	0.520	0.906	0.908	0.684
	H	0.628	0.530	0.908	0.910	0.692
	HT	0.631	0.535	0.911	0.914	0.695
Oregon	G	0.658	0.544	0.901	0.935	0.709
	GT	0.691	0.588	0.937	0.949	0.737
	H	0.674	0.566	0.926	0.945	0.736
	HT	0.698	0.598	0.945	0.955	0.759
West Virginia	G	0.694	0.368	0.958	0.946	0.728
	GT	0.772	0.529	0.961	0.958	0.789
	H	0.731	0.445	0.977	0.962	0.786
	HT	0.785	0.557	0.973	0.969	0.830

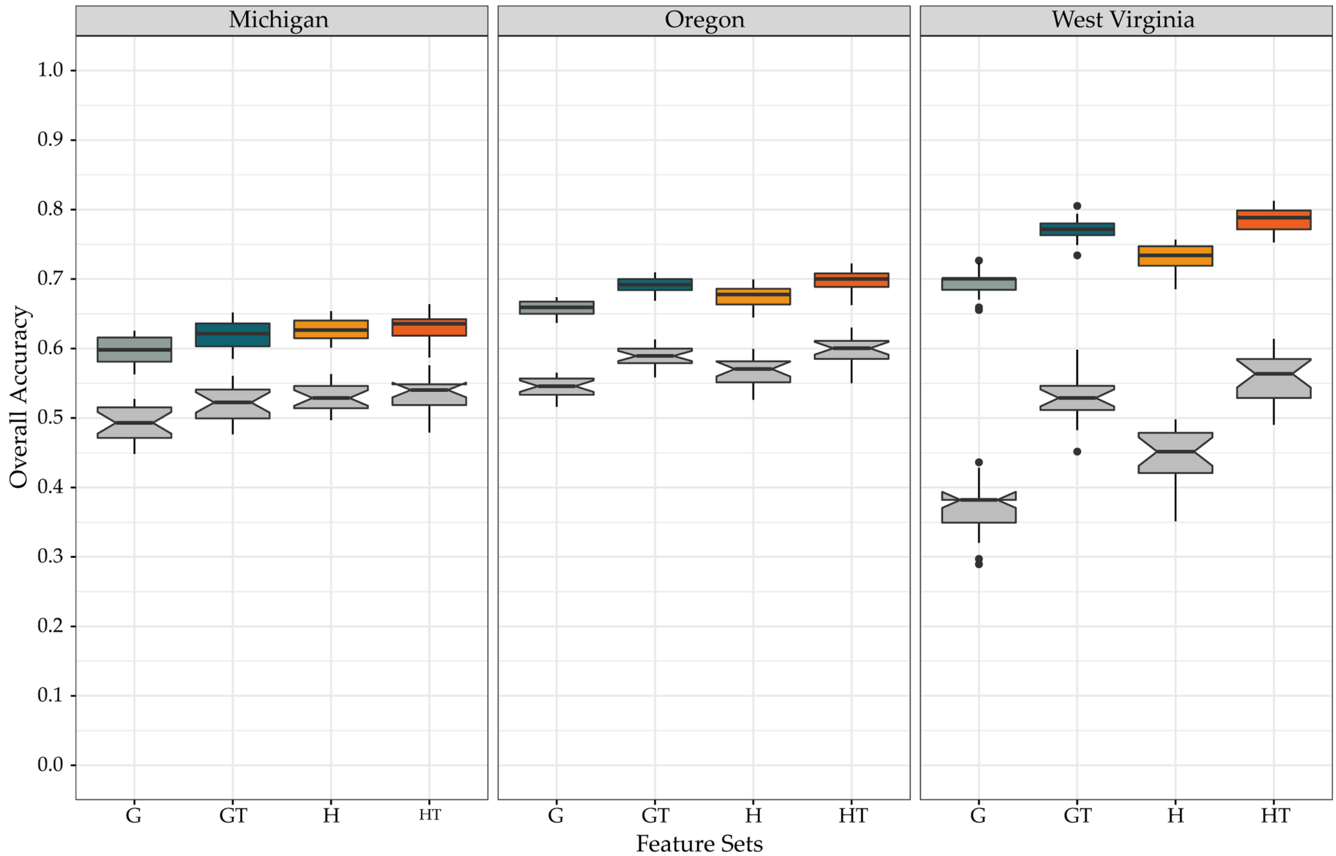

Fig. 4. Distribution of overall accuracy (OA) and map image classification efficacy (MICE) for 20 model replicates of each of the four feature spaces. Colored boxplots represent OA and gray boxplots represent MICE. G = GLAD; GT = GLAD plus terrain derivatives; H = harmonic regression; HT = harmonic regression + terrain derivatives.

only using the harmonic regression or GLAD Phenology metrics, highlighting the value of incorporating digital terrain data when attempting to differentiate certain forest community types. This supports findings from Adams et al. (2020), Pasquarella et al. (2018), and Hartley et al. (2022). These findings generally align with the findings from our prior study in West Virginia, Hartley et al. (2022), in which the GLAD Type C metrics were used, and training and validation data were derived from field plot data provided by the West Virginia Natural Heritage Program (Hartley et al., 2022). Harmonic regression generally outperformed the GLAD metrics, and digital terrain variables were important for improving predictive performance.

It should be noted that the top3 metric was above 0.89 for all states and all metric sets. In other words, the correct forest community type had a predicted probability in the top three classes more than 89% of the time. This suggests that a large proportion of the error is generally associated with confusion between specific classes. The AUC ROC metric was generally lower than the AUC PRC metric, which we attribute to the high degree of class imbalance or number of samples in each forest community type (Saito and Rehmsmeier, 2015).

Table 5 shows the Kruskal-Wallis pairwise comparison results for the forest community type differentiation using the different feature spaces within each state, which were calculated using the 20 replicates and compares the median of the MICE metrics based on rank. Using a 95% confidence interval and 0.05 alpha, the G and GT results were statistically different for all states. Further, the H and HT results were statistically significant for Oregon and West Virginia but not Michigan. This highlights the value of incorporating digital terrain variables for forest community type differentiation in some landscapes. We attribute the reduced value of adding terrain variables in Michigan in comparison to Oregon and West Virginia to lower topographic variability, which could result in topographic factors being less important for differentiating forest community types in this landscape. However, other confounding factors include different disturbance and management practices and histories, different community types, and the impact of glaciation.

Table 5
Kruskal-Wallis pairwise comparison results for predicting forest community types using different feature spaces. Statistically significant comparisons are shaded gray. G = GLAD; GT = GLAD plus terrain derivatives; H = harmonic regression; HT = harmonic regression + terrain derivatives; UCI = upper 95% confidence interval; LCI = lower 95% confidence interval.

State	Set	Difference	UCI	LCI	*p*-value
Michigan	G to GT	−20.70	−40.08	−1.32	2.89E-02
	G to H	−28.75	−48.13	−9.37	0.000543
	H to GT	−8.05	−27.43	11.32	1
	G to HT	−33.55	−52.93	−14.17	3.00E-05
	GT to HT	−12.85	−32.23	6.53	0.481042
	H to HT	−4.80	−24.18	14.58	1
Oregon	G to GT	−37.75	−57.13	−18.37	2.00E-06
	G to H	−18.45	−37.83	0.93	0.072145
	H to GT	19.30	−0.08	38.68	0.051664
	G to HT	−45.2	−64.58	−25.82	0
	GT to HT	−7.45	−26.83	11.93	1
	H to HT	−26.75	−46.13	−7.37	0.001628
West Virginia	G to GT	−43.125	−62.49	−23.76	0
	G to H	−17.225	−36.59	2.14	0.113754
	H to HT	25.9	6.53	45.27	0.002512
	G to HT	−52.25	−71.62	−32.88	0
	GT to HT	−9.125	−28.49	10.24	1
	H to HT	−35.025	−54.39	−15.66	1.10E-05

The G and H datasets were only statistically different for Michigan while the GT and HT datasets were not statistically significant for any of the three states. This suggests that while harmonic regression coefficients generally provided higher OA, MICE, AUC ROC, and AUC PRC, the improvement was generally not statistically significant. Practically, this suggests that GLAD metrics may be a suitable choice when attempting to differentiate forest community types at a moderate spatial resolution using spectral or spectral and terrain predictor variables.

3.2. AGLBM prediction

The mean RMSE and R-squared for the 20 model replicates for each feature space in each state and for both the total and species-specific AGLBM are reported in Table 6 while Fig. 5 shows the distribution of predicted R-squared and RMSE. Generally, a larger proportion of the total variance in total and Douglas Fir AGLBM were explained in the Oregon models than the total and White Oak AGLBM in Michigan and West Virginia. However, RMSE was lower for Michigan and West Virginia. We attribute these differences partially to the variability in AGLBM between FIA plots within the different states. As noted above, the standard deviation of total AGLBM for Michigan, Oregon, and West Virginia were 1.4, 24.3, and 4.8, respectively. Thus, the Oregon FIA plot data were much more variable than the data from the other two states. Thus, direct comparisons between the three states may not be appropriate. Also, the species chosen to explore species-specific AGLBM was different for Oregon vs. Michigan and West Virginia. However, even given these differences, the results generally suggest stronger performance in Oregon in comparison to the other two states, perhaps due to the stronger topographic influence. For example, all R-squared values, regardless of the feature space or whether total or species-specific AGLBM was being predicted, were above 0.40 in Oregon and lower than 0.20 in the other two states. This could partially be a result of differences in inherent variability, as noted above. However, it could also be related to differences in disturbance histories, management practices, and forest characteristics, such as the vertical distribution of the biomass, the amount of subcanopy vegetation, and the spacing and sizes of trees within the FIA plots.

In contrast to the forest community type classification results, the performance of the different feature spaces was less pronounced for predicting total and species-specific AGLBM. Adding terrain variables generally had less of an impact and there were less differences between the harmonic regression and GLAD Phenology metrics models. Based on the pairwise Kruskal-Wallis analysis, for estimating total AGLBM only the Oregon G and HT models were statistically different (p-value = 0.022). All other pairs were suggested to not have statistically different median RMSEs. For the species-specific AGLBM, no feature spaces in any states had statistically significantly different median RMSEs at an alpha of 0.05. In summary, the feature space, including representing the spectral data using harmonic regression coefficients vs. GLAD Phenology metrics and including both spectral and terrain variables, had less of an impact on predicting total or species-specific AGLBM than differentiating forest community types. Further, forest community type differentiation was generally found to be more accurate than biomass estimation using the Landsat-derived spectral measures with or without the terrain variables.

Table 6
Mean assessment results for 20 replicates of total AGLBM and species-specific AGLBM predictions using four different feature spaces. RMSE = root mean square error; G = GLAD; GT = GLAD plus terrain derivatives; H = harmonic regression; HT = harmonic regression + terrain derivatives.

		Total		Species	
State	Feature Set	RMSE	R-Squared	RMSE	R-Squared
Michigan	G	7.833	0.134	1.376	0.009
	GT	7.801	0.142	1.374	0.012
	H	7.733	0.158	1.332	0.087
	HT	7.750	0.157	1.347	0.059
Oregon	G	17.906	0.452	17.442	0.468
	GT	17.667	0.467	17.194	0.484
	H	17.432	0.484	17.592	0.461
	HT	17.369	0.490	17.326	0.478
West Virginia	G	12.305	0.011	4.498	0.035
	GT	12.192	0.028	4.472	0.048
	H	12.188	0.027	4.484	0.039
	HT	12.138	0.034	4.446	0.054

Several prior studies have noted the benefit of using time series aggregation or summarization methods, such as harmonic regression, that characterize spectral variability, seasonality, and phenology, to predict AGLBM (for example, (H. Nguyen et al., 2019; Wilson et al., 2018; Zhu and Liu, 2015)). Prior studies have also noted the value of using machine learning-based regression methods, such as RF and support vector machines (SVM), for predicting AGLBM from spectral variables (for example, (López-Serrano et al., 2020; Wu et al., 2016)). Although adding terrain derivatives offered only minimal improvements in this study, as documented by the lack of statistical difference between models, other studies have noted benefits (e.g., (López-Serrano et al., 2020)). A survey of prior studies also suggests variability in reported model performance, as estimated using R-squared and/or the RMSE metric. We argue that the variabilities documented both within this study and between prior studies may be at least partially attributed to differences in the forest types and species compositions, vertical stratification, disturbance histories, landscape characteristics, and spatial extents of study areas used in the research, which highlights the value of using multiple, disparate study areas to generalize findings. This also highlights the value of large datasets, such as those provided by the FIA program, for assessing modeling methods.

4. Conclusions and recommendations

This study offered a comparison of two methods to summarize a time series of Landsat observations, harmonic regression coefficients and GLAD Phenology metrics, for estimating forest characteristics, specifically forest community type, total AGLBM, and species-specific AGLBM. We also incorporated terrain variables derived from NED DTM data. FIA plots served as the training and validation data in the study, and three separate states were examined: Michigan, Oregon, and West Virginia.

Our findings support the use of harmonic regression coefficients as a means to summarize spectral time series data. Harmonic regression coefficients generally provided better predictive performance for differentiating forest community types in comparison to GLAD phenology metrics. However, differences were not always statistically significant, suggesting that GLAD ARD products may be adequate if harmonic regression coefficients are not available. Regardless of the spectral data used, DTM-derived terrain variables improved classification performance in some landscapes when using machine learning, random forest classification. For the regression-based estimation of total and species-specific AGLBM, the spectral data used and the incorporation of terrain variables had less of an impact on model performance. Further, the regression models in Oregon were more accurate than those in Michigan and West Virginia, suggesting that AGLBM prediction may be more or less challenging in different landscapes depending on forest characteristics (e.g., species composition and understory abundance), terrain, management practices, and disturbance histories.

The development of ARD products relying on harmonic regression methods would be valuable to aid users in adopting these metrics. Further, this could add consistency to harmonic regression and coefficient estimation methods between studies and applied mapping and monitoring tasks. Further, the development of standard terrain metric datasets from existing DTM data, such as those provided by the United States Geological Survey (USGS) 3DEP program, could be valuable for forest mapping and monitoring tasks that rely on empirical, machine learning methods.

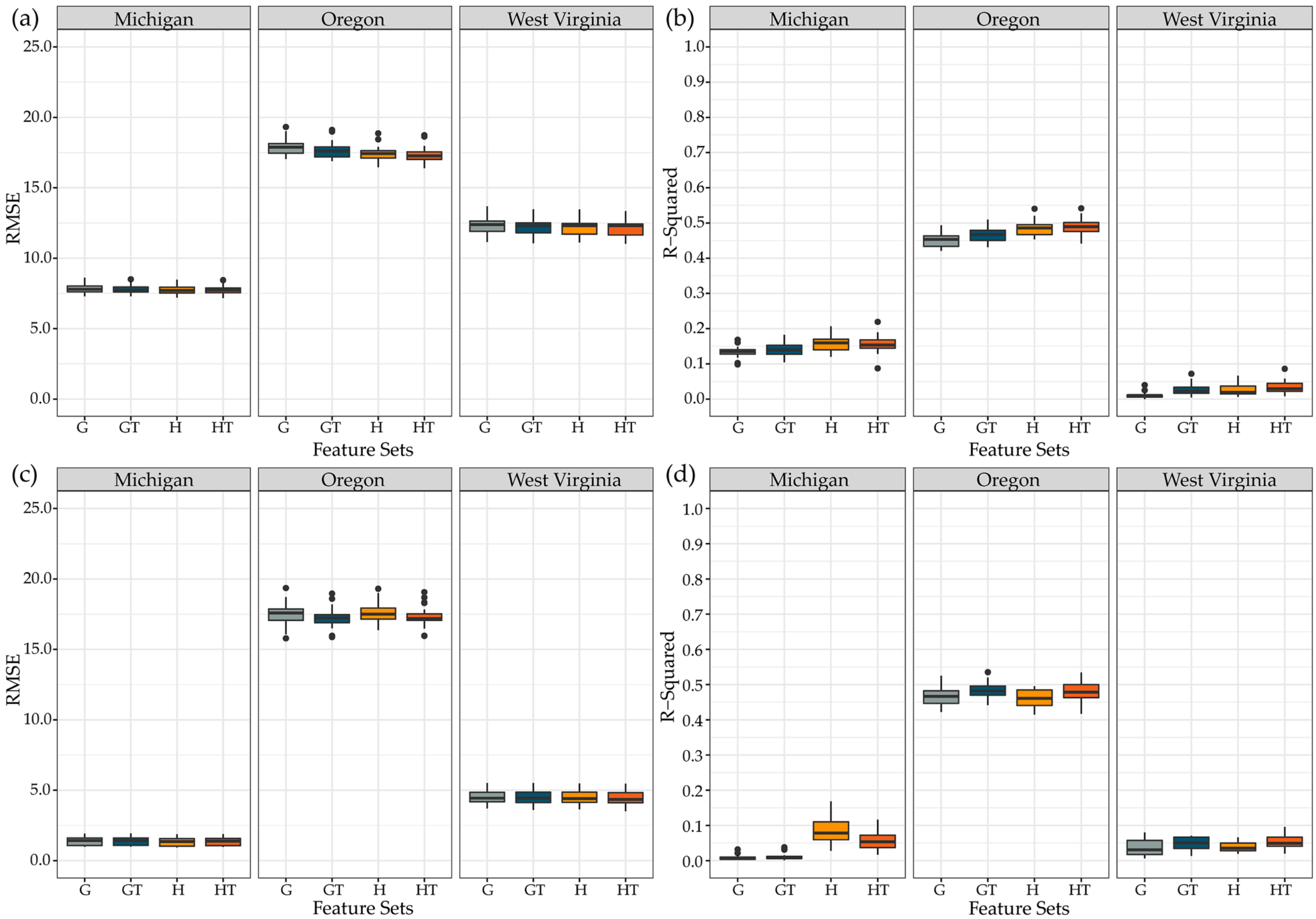

Fig. 5. Distribution of RMSE (a and c) and predicted R-squared (b and d) for 20 model replicates of each of the four feature spaces for estimating total (a and b) and species-specific AGLBM (c and d). RMSE = root mean square error; G = GLAD; GT = GLAD plus terrain derivatives; H = harmonic regression; HT = harmonic regression + terrain derivatives.

CRediT authorship contribution statement

Aaron E. Maxwell: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Barry T. Wilson:** Data curation, Writing – review & editing. **Justin J. Holgerson:** Data curation, Writing – review & editing. **Michelle S. Bester:** Formal analysis, Investigation, Writing – review & editing.

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.

Data availability

The authors do not have permission to share data.

Acknowledgments

We would like to thank two anonymous reviewers whose comments strengthened the work. Funding was provided by AmericaView, which is supported by the U.S. Geological Survey under Grant/Cooperative Agreement No. G18AP00077. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U. S. Government.

References

Adams, B., Iverson, L., Matthews, S., Peters, M., Prasad, A., Hix, D.M., 2020. Mapping Forest Composition with Landsat Time Series: An Evaluation of Seasonal Composites and Harmonic Regression. Remote Sens. 12, 610. https://doi.org/10.3390/rs12040610.

ArcGIS Pro Python reference—ArcGIS Pro | Documentation [WWW Document], n.d. URL https://pro.arcgis.com/en/pro-app/latest/arcpy/main/arcgis-pro-arcpy-reference.htm (accessed 6.11.22).

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 (5), 423–431.

Beaubien, J., et al., 1979. Forest type mapping from Landsat digital data. Photogramm. Eng. Remote Sens. 45 (8), 1135–1144.

Bechtold, W.A., Patterson, P.L., 2005. The enhanced forest inventory and analysis program–national sampling design and estimation procedures. USDA Forest Service, Southern Research Station.

Binkley, D., 2021. Forest Ecology: An Evidence-based Approach. John Wiley & Sons.

Bonan, G.B., 2008. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. Science 320, 1444–1449. https://doi.org/10.1126/science.1155121.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Bryant, E., Dodge, A.G., Warren, S.D., 1980. Landsat for practical forest type mapping- A test case. Photogramm. Eng. Remote Sens. 46, 1575–1584.

Chen, J.M., Cihlar, J., 1996. Retrieving leaf area index of boreal conifer forests using Landsat TM images. Remote Sens. Environ. 55 (2), 153–162.

Congalton, R.G., Green, K., 2019. Assessing the accuracy of remotely sensed data: principles and practices. CRC Press.

Coulter, L.L., Stow, D.A., Tsai, Y.-H., Ibanez, N., Shih, H., Kerr, A., Benza, M., Weeks, J. R., Mensah, F., 2016. Classification and assessment of land cover and land use change in southern Ghana using dense stacks of Landsat 7 ETM+ imagery. Remote Sens. Environ. 184, 396–409.

Crist, E.P., Cicone, R.C., 1984. A physically-based transformation of Thematic Mapper data—The TM Tasseled Cap. IEEE Trans. Geosci. Remote Sens. GE-22 (3), 256–263.

Dalgaard, P., 2002. Analysis of variance and the Kruskal-Wallis test. Introd. Stat. R 111–127.

Download ArcGIS Pro—ArcGIS Pro | Documentation [WWW Document], n.d. URL https://pro.arcgis.com/en/pro-app/latest/get-started/download-arcgis-pro.htm (accessed 6.11.22).
Dwyer, J.L., Roy, D.P., Sauer, B., Jenkerson, C.B., Zhang, H.K., Lymburner, L., 2018. Analysis Ready Data: Enabling Analysis of the Landsat Archive. Remote Sens. 10, 1363. https://doi.org/10.3390/rs10091363.
Estornell, J., Martí-Gavilá, J.M., Sebastiá, M.T., Mengual, J., 2013. Principal component analysis applied to remote sensing. Model. Sci. Educ. Learn. 6, 83–89.
Evans, J.S., Murphy, M.A., 2018. rfUtilities.
Fan, J., Upadhye, S., Worster, A., 2006. Understanding receiver operating characteristic (ROC) curves. CJEM 8, 19–20. https://doi.org/10.1017/S1481803500013336.
Fassnacht, K.S., Gower, S.T., MacKenzie, M.D., Nordheim, E.V., Lillesand, T.M., 1997. Estimating the leaf area index of north central Wisconsin forests using the Landsat Thematic Mapper. Remote Sens. Environ. 61 (2), 229–245.
Flood, N., 2013. Seasonal Composite Landsat TM/ETM+ Images Using the Medoid (a Multi-Dimensional Median). Remote Sens. 5, 6481–6500. https://doi.org/10.3390/rs5126481.
Forest Inventory and Analysis National Program [WWW Document], n.d. URL https://www.fia.fs.usda.gov/ (accessed 1.5.23).
Franklin, S.E., 2020. Interpretation and use of geomorphometry in remote sensing: a guide and review of integrated applications. Int. J. Remote Sens. 41 (19), 7700–7733.
Franklin, J.F., Johnson, K.N., Johnson, D.L., 2018. Ecological forest management. Waveland Press.
French, N.H., Kasischke, E.S., Hall, R.J., Murphy, K.A., Verbyla, D.L., Hoy, E.E., Allen, J. L., 2008. Using Landsat data to assess fire and burn severity in the North American boreal forest region: an overview and summary of results. Int. J. Wildland Fire 17, 443–462.
Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., Tyler, D., 2002. The national elevation dataset. Photogramm. Eng. Remote Sens. 68, 5–32.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. Big Remotely Sensed Data: tools, applications and experiences 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.
H. Nguyen, T., Jones, S., Soto-Berelov, M., Haywood, A., Hislop, S., 2019. Landsat time-series for estimating forest aboveground biomass and its dynamics across space and time: A review. Remote Sens. 12 (1), 98.
Hamner, B., Frasco, M., 2018. Metrics: Evaluation Metrics for Machine Learning.
Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science 342, 850–853. https://doi.org/10.1126/science.1244693.
Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., Moore, R., 2016. Humid tropical forest disturbance alerts using Landsat data. Environ. Res. Lett. 11 (3), 034008.
Hardwick, S.W., Shelley, F.M., Holtgrieve, D.G., 2013. The Geography of North America: Environment, Culture, Economy. Pearson Education.
Hartley, F.M., Maxwell, A.E., Landenberger, R.E., Bortolot, Z.J., 2022. Forest Type Differentiation Using GLAD Phenology Metrics, Land Surface Parameters, and Machine Learning. Geographies 2, 491–515.
Huang, C., Goward, S.N., Schleeweis, K., Thomas, N., Masek, J.G., Zhu, Z., 2009. Dynamics of national forests assessed using the Landsat record: Case studies in eastern United States. Remote Sens. Environ. 113 (7), 1430–1442.
Jackson, R.B., Randerson, J.T., Canadell, J.G., Anderson, R.G., Avissar, R., Baldocchi, D. D., Bonan, G.B., Caldeira, K., Diffenbaugh, N.S., Field, C.B., Hungate, B.A., Jobbágy, E.G., Kueppers, L.M., Nosetto, M.D., Pataki, D.E., 2008. Protecting climate with forests. Environ. Res. Lett. 3 (4), 044006.
James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning. Springer.
Kuhn, M., 2021. caret: Classification and Regression Training.
Landsat Legacy Project Team, 2017. Landsat's Enduring Legacy: Pioneering Global Land Observations from Space. American Society for Photogrammetry and Remote Sensing.
Lastovicka, J., Svec, P., Paluba, D., Kobliuk, N., Svoboda, J., Hladky, R., Stych, P., 2020. Sentinel-2 Data in an Evaluation of the Impact of the Disturbances on Forest Vegetation. Remote Sens. 12, 1914. https://doi.org/10.3390/rs12121914.
Lillesand, T.M., Kiefer, R.W., 1979. Remote sensing and image interpretation(Book). N. Y. John Wiley Sons Inc 1979 p. 624.
López-Serrano, P.M., Cárdenas Domínguez, J.L., Corral-Rivas, J.J., Jiménez, E., López-Sánchez, C.A., Vega-Nieva, D.J., 2020. Modeling of Aboveground Biomass with Landsat 8 OLI and Machine Learning in Temperate Forests. Forests 11, 11. https://doi.org/10.3390/f11010011.
Margono, B.A., Turubanova, S., Zhuravleva, I., Potapov, P., Tyukavina, A., Baccini, A., Goetz, S., Hansen, M.C., 2012. Mapping and monitoring deforestation and forest degradation in Sumatra (Indonesia) using Landsat time series data sets from 1990 to 2010. Environ. Res. Lett. 7 (3), 034010.
Maxwell, A.E., Shobe, C.M., 2022. Land-surface parameters for spatial predictive mapping and modeling. Earth-Sci. Rev. 226, 103944.
Nijland, W., Bolton, D.K., Coops, N.C., Stenhouse, G., 2016. Imaging phenology; scaling from camera plots to landscapes. Remote Sens. Environ. 177, 13–20.
Parker, B.M., Lewis, T., Srivastava, S.K., 2015. Estimation and evaluation of multi-decadal fire severity patterns using Landsat sensors. Remote Sens. Environ. 170, 340–349. https://doi.org/10.1016/j.rse.2015.09.014.
Pasquarella, V.J., Bradley, B.A., Woodcock, C.E., 2017. Near-real-time monitoring of insect defoliation using Landsat time series. Forests 8, 275.
Pasquarella, V.J., Holden, C.E., Woodcock, C.E., 2018. Improved mapping of forest type using spectral-temporal Landsat features. Remote Sens. Environ. 210, 193–207. https://doi.org/10.1016/j.rse.2018.02.064.
Perry, D.A., Oren, R., Hart, S.C., 2008. Forest ecosystems. JHU press.
Potapov, P., Yaroshenko, A., Turubanova, S., Dubinin, M., Laestadius, L., Thies, C., Aksenov, D., Egorov, A., Yesipova, Y., Glushkov, I., Karpachevskiy, M., Kostikova, A., Manisha, A., Tsybikova, E., Zhuravleva, I., 2008. Mapping the World's Intact Forest Landscapes by Remote Sensing. Ecol, Soc, p. 13.
Potapov, P., Hansen, M.C., Kommareddy, I., Kommareddy, A., Turubanova, S., Pickens, A., Adusei, B., Tyukavina, A., Ying, Q., 2020. Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping. Remote Sens. 12, 426. https://doi.org/10.3390/rs12030426.
Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J. B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. Remote Sens. Environ. 253, 112165 https://doi.org/10.1016/j.rse.2020.112165.
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rose, M.B., Nagle, N.N., 2021. Characterizing Forest Dynamics with Landsat-Derived Phenology Curves. Remote Sens. 13, 267. https://doi.org/10.3390/rs13020267.
Royle, D.D., Lathrop, R.G., 1997. Monitoring hemlock forest health in New Jersey using Landsat TM data and change detection techniques. For. Sci. 43, 327–335.
Saito, T., Rehmsmeier, M., 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PloS One 10(3), e0118432.
Schultz, M., Clevers, J.G., Carter, S., Verbesselt, J., Avitabile, V., Quang, H.V., Herold, M., 2016. Performance of vegetation indices from Landsat time series in deforestation monitoring. Int. J. Appl. Earth Obs. Geoinformation 52, 318–327.
Sellers, P.J., Tucker, C.J., Collatz, G.J., Los, S.O., Justice, C.O., Dazlich, D.A., Randall, D. A., 1994. A global 1 by 1 NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. Int. J. Remote Sens. 15 (17), 3519–3545.
Sellers, P.J., Tucker, C.J., Collatz, G.J., Los, S.O., Justice, C.O., Dazlich, D.A., Randall, D. A., 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data. J. Clim. 9 (4), 706–737.
Shao, G., Tang, L., Zhang, H., 2021. Introducing image classification efficacies. IEEE. Access 9, 134809–134816.
Shimabukuro, Y.E., Batista, G.T., Mello, E.M.K., Moreira, J.C., Duarte, V., 1998. Using shade fraction image segmentation to evaluate deforestation in Landsat Thematic Mapper images of the Amazon region. Int. J. Remote Sens. 19 (3), 535–541.
Shumway, R.H., Stoffer, D.S., Stoffer, D.S., 2000. Time series analysis and its applications. Springer.
Simoes, R., Camara, G., Queiroz, G., Souza, F., Andrade, P.R., Santos, L., Carvalho, A., Ferreira, K., 2021. Satellite image time series analysis for big earth observation data. Remote Sens. 13, 2428.
Smail, R.A., 2010. Forest land conversion, ecosystem services, and economic issues for policy: A review.
Soverel, N.O., Perrakis, D.D.B., Coops, N.C., 2010. Estimating burn severity from Landsat dNBR and RdNBR indices across western Canada. Remote Sens. Environ. 114, 1896–1909. https://doi.org/10.1016/j.rse.2010.03.013.
Tharwat, A., 2020. Classification assessment methods. Appl. Comput. Inform. 17 (1), 168–192.
The Perl Programming Language - www.perl.org [WWW Document], n.d. URL https://www.perl.org/ (accessed 6.25.22).
Vogelmann, J.E., Tolk, B., Zhu, Z., 2009. Monitoring forest changes in the southwestern United States using multitemporal Landsat data. Remote Sens. Environ. 113 (8), 1739–1748.
Wei, R., Wang, J., 2018. multiROC: Calculating and Visualizing ROC and PR Curves Across Multi-Class Classifications.
Welcome to Python.org [WWW Document], n.d. . Python.org. URL https://www.python.org/ (accessed 1.6.23).
Wilson, B.T., Knight, J.F., McRoberts, R.E., 2018. Harmonic regression of Landsat time series for modeling attributes from national forest inventory data. ISPRS J. Photogramm. Remote Sens. 137, 29–46.
Wright, M.N., Ziegler, A., 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J. Stat. Softw. 77, 1–17. https://doi.org/10.18637/jss.v077.i01.
Wu, C., Shen, H., Shen, A., Deng, J., Gan, M., Zhu, J., Xu, H., Wang, K., 2016. Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. J. Appl. Remote Sens. 10.
Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., Cohen, W.B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J.D., Hostert, P., Hughes, M.J., Huntington, J., Johnson, D.M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T.A., Schaaf, C., Schott, J.R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J.C., Wynne, R.H., Zhu, Z., 2019. Current status of Landsat program, science, and applications. Remote Sens. Environ. 225, 127–147. https://doi.org/10.1016/j.rse.2019.02.015.
Xu, C., Du, X., Jian, H., Dong, Y.i., Qin, W., Mu, H., Yan, Z., Zhu, J., Fan, X., 2022. Analyzing large-scale Data Cubes with user-defined algorithms: A cloud-native approach. Int. J. Appl. Earth Obs. Geoinformation 109, 102784.

Zhu, X., Liu, D., 2015. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. ISPRS J. Photogramm. Remote Sens. 102, 222–231.
Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171.
Zhu, Z., Wulder, M.A., Roy, D.P., Woodcock, C.E., Hansen, M.C., Radeloff, V.C., Healey, S.P., Schaaf, C., Hostert, P., Strobl, P., Pekel, J.-F., Lymburner, L., Pahlevan, N., Scambos, T.A., 2019. Benefits of the free and open Landsat data policy. Remote Sens. Environ. 224, 382–385. https://doi.org/10.1016/j.rse.2019.02.016.