Received: 29 June 2022 | Accepted: 12 February 2023
DOI: 10.1111/2041-210X.14084

Methods in Ecology and Evolution

REVIEW

Active Remote Sensing for Ecology and Ecosystem Conservation

Effects of outliers on remote sensing-assisted forest biomass estimation: A case study from the United States national forest inventory

Jonathan A. Knott[1] | Greg C. Liknes[1] | Courtney L. Giebink[1,2] | Sungchan Oh[3] | Grant M. Domke[1] | Ronald E. McRoberts[4] | Valquiria F. Quirino[4,5] | Brian F. Walters[1]

[1]USDA Forest Service, St Paul, Minnesota, USA
[2]Oak Ridge Associated Universities, Oak Ridge, Tennessee, USA
[3]Institute for Plant Sciences, Purdue University, West Lafayette, Indiana, USA
[4]University of Minnesota, St Paul, Minnesota, USA
[5]North Dakota State University, Fargo, North Dakota, USA

Correspondence
Jonathan A. Knott
Email: jonathan.knott@usda.gov

Funding information
U.S. Forest Service

Handling Editor: Hooman Latifi

Abstract

1. Large-scale ecological sampling networks, such as national forest inventories (NFIs), collect in situ data to support biodiversity monitoring, forest management and planning, and greenhouse gas reporting. Data harmonization aims to link auxiliary remotely sensed data to field-collected data to expand beyond field sampling plots, but outliers that arise in data harmonization—questionable observations because their values differ substantially from the rest—are rarely addressed.
2. In this paper, we review the sources of commonly occurring outliers, including random chance (statistical outliers), definitions and protocols set by sampling networks, and temporal and spatial mismatch between field-collected and remotely sensed data. We illustrate different types of outliers and the effects they have on estimates of above-ground biomass population parameters using a case study of 292 NFI plots paired with airborne laser scanning (ALS) and Sentinel-2 data from Sawyer County, Wisconsin, United States.
3. Depending on the criteria used to identify outliers (sampling year, plot location error, nonresponse, presence of zeros and model residuals), as many as 53 of the 292 Forest Inventory and Analysis plot observations (18%) were identified as potential outliers using a single criterion and 111 plot observations (38%) if all criteria were used. Inclusion or removal of potential outliers led to substantial differences in estimates of mean and standard error of the estimate of biomass per unit area. The simple expansion estimator, which does not rely on ALS or other auxiliary data, was more sensitive to outliers than model-assisted approaches that incorporated ALS and Sentinel-2 data. Including Sentinel-2 predictors showed minimal increases to the precision of our estimates relative to models with ALS predictors alone.

4. Outliers arise from many causes and can be pervasive in data harmonization workflows. Our review and case study serve as a note of caution to researchers and practitioners that the inclusion or removal of potential outliers can have unintended consequences on population parameter estimates. When used to inform large-scale biomass mapping, carbon markets, greenhouse gas reporting and environmental policy, it is necessary to ensure the proper use of NFI and remotely sensed data in geospatial data harmonization.

KEYWORDS
airborne laser scanning, data harmonization, Forest Inventory and Analysis (FIA), lidar, Sentinel-2

1 | INTRODUCTION

Forests provide countless ecological and economic functions, so monitoring and estimating forest resources have become global priorities. National forest inventories (NFIs) have been established globally to support forest resource estimation for greenhouse gas reporting and economic valuations, but NFIs also support other ecological purposes (Chirici et al., 2012; EPA, 2022; Romijn et al., 2015; Tinkham et al., 2018; Tomppo et al., 2010). Applications of NFI data include ecological indicators of climate change (Horn et al., 2018; Knott et al., 2020; Will-Wolf et al., 2017; Zhu et al., 2012); below-ground mycorrhizal associations with tree species (Averill et al., 2022; Carteron et al., 2022; Jo et al., 2019); invasion biology of plants, pests and pathogens (Baer & Gray, 2022; Fei et al., 2019; Iannone et al., 2016); and biodiversity patterns and processes (Chirici et al., 2012; Fei et al., 2018; Winter et al., 2012; Zeller et al., 2018), among others. In addition to characterizing observed ecological patterns and processes, NFI data have also supported the parameterization of ecological models including spatially explicit landscape models (Purves et al., 2007; Suárez-Muñoz et al., 2021; Wang et al., 2013), tree growth models (Lichstein et al., 2010; Giebink, DeRose, et al., 2022; Heilman et al., 2022) and species distribution models (Crookston et al., 2010; Iverson et al., 2019; Prasad et al., 2020), which aim to predict forest responses to changing climate and disturbances.

NFIs collect data on forest attributes at regional to continental scales, and the sub-populations (land area) of interest included in NFIs range from forests impacted by specific activities (e.g. management and harvest), forests of specific ownership classes (e.g. only public land), to only forestland or all land (Tomppo et al., 2010). To capture this range, many NFIs use probability-based sampling designs: inventory plots are established across the landscape to collect a representative sample of the population of interest (Bechtold & Patterson, 2005; Tomppo et al., 2010). These strategic inventories (many sites collecting a relatively small quantity of data) provide a contrast to traditional, place-based ecological research networks (few sampling sites collecting a relatively large quantity of data). NFIs require careful planning to ensure estimation of parameters are consistent and generalizable across the population of interest (i.e. forestlands). Ongoing efforts to coordinate among NFIs internationally (Kovac et al., 2014; McRoberts, Tomppo, et al., 2012; Prasad et al., 2020; Tomppo & Schadauer, 2012) aim to ensure consistency in global efforts to provide estimates of how forests mitigate and contribute to climate change through the capture, storage and emissions of greenhouse gases (such as the United Nations Framework Convention on Climate Change, UNFCCC, Reducing Emissions from Deforestation and Forest Degradation, REDD+, and the Intergovernmental Panel on Climate Change, IPCC; IPCC, 2019).

The recent proliferation of remote sensing technologies has facilitated the estimation of forest resources by integrating NFI data with remotely sensed data. This process, hereafter 'data harmonization', combines geospatial data and/or data types from multiple sources (Global Forest Observations Initiative, 2020; Lister et al., 2020; White et al., 2013). Remote sensing provides the opportunity for NFIs to estimate forest resources beyond the small proportion of land inventoried by on-the-ground sampling, filling the gaps between NFI plots (Wilson et al., 2013; Yu et al., 2022). Remote sensing may also improve cost efficiency of NFIs, for example, stratifying by forest/non-forest to reduce sample size needed to meet precision standards, providing low-cost information about forest attributes if NFIs are given a small budget, or reducing travel expenses to remote areas where access may be difficult (Lister et al., 2020).

However, NFI definitions and protocols were often formulated and established decades before the emergence and proliferation of remote sensing (Lister et al., 2020; Tomppo et al., 2010), so geospatial data harmonization using NFI data can lead to outliers—defined broadly as potential mismatches among data from different sources that lead to values that are biologically untenable or statistically different from the rest (Kendall et al., 1982). For example, NFIs might define forest/non-forest based on canopy cover and/or land use, but remote sensing might detect forest-like land cover in areas that are considered non-forest by NFIs (e.g. trees outside forests in orchards, parks and residential areas; Thomas et al., 2021). NFIs must consider whether the definitions they adhere to and the attributes these definitions describe can be characterized through the harmonization of remotely sensed and field-collected data.

Here, we provide an overview of potential outliers that arise from geospatial data harmonization of remotely sensed and field-collected NFI data. Our study had two objectives: (1) to provide a literature review on common sources of outliers between remotely sensed and NFI data and (2) to illustrate the effects of retaining or removing outliers on estimates of above-ground live and standing dead biomass per unit area (hereafter 'biomass per unit area') using a case study in Sawyer County, Wisconsin, United States (US). We conclude with discussion on how outliers may affect other sampling efforts beyond NFIs and how the uncertainty that arises from geospatial data harmonization can have substantial effects on resource estimation that informs management, carbon markets and greenhouse gas reporting.

2 | A PRIMER ON REMOTE SENSING-ASSISTED ESTIMATION

A general data harmonization workflow for estimating forest parameters (e.g. biomass per unit area) for a population (study area) from field-collected and remotely sensed data includes the following steps. First, forest inventory data are collected at plots within the population of interest to estimate tree- and plot-level biomass. Individual tree measurements, usually diameter and height, are converted to tree-level biomass predictions using allometric models. The United States Department of Agriculture Forest Service, Forest Inventory and Analysis (FIA) program uses a component ratio method which first predicts the volume of different components of the tree and converts volume to biomass and carbon predictions using scaling factors based on wood densities and carbon fractions (Domke et al., 2012; Doraisami et al., 2022; Miles & Smith, 2009; Woodall et al., 2011). Other allometric models can be used with NFI data (including Jenkins et al., 2003; Chave et al., 2005; Chave et al., 2014; Colmanetti et al., 2018; Gonzalez-Akre et al., 2022). These predictions can be standardized to per unit area (e.g. megagrams of biomass per hectare) using expansion factors based on NFI sampling design that relate plot area and sample size to the population. Some NFIs make publicly available their inventory data and associated volume/biomass/carbon predictions (such as the FIA DataMart, https://apps.fs.usda.gov/fia/datamart/datamart.html) while others only release estimates of population parameters such as country- or region-level estimates of total forest area (Tomppo et al., 2010).

Next, remotely sensed data are compiled for the same population (Figure 1). Passive sensors rely on the reflectance of the sun from surfaces, whereas active sensors use their own light or other energy source (Lu et al., 2016). Generally, passive sensors provide information about the reflectance of the forest canopy, whereas active sensors have the potential to penetrate the canopy and provide additional information (Gonzalez et al., 2010; Vaglio Laurin et al., 2014). Remotely sensed data may be delivered as raster data with a fixed resolution (hereafter 'pixel based') or, in the case of lidar, as point clouds for which users are not constrained by a pre-determined grid and can 'clip' data to features (such as an NFI plot footprint) or calculate metrics and create summaries at user-defined resolution.

Concurrent forest inventory and remote sensing is possible when sample size is small and equipment is available (Stovall & Shugart, 2018). This is common practice for traditional ecological research networks that can commission their own remote sensing. However, concurrent field sampling and remote sensing for NFIs (generally a large sample size with many field crews) usually become economically or operationally unfeasible (Oh et al., 2022, but see Babcock et al., 2018). Instead, publicly available, large-scale aerial (such as uncrewed aerial systems or airplane-based platforms) or satellite (such as Landsat or Sentinel) remote sensing are typically used. Temporal resolution (revisit time) and data quality (due to inconsistent lighting condition or cloud cover) often necessitate the use of a time window around the field inventory date to get complete coverage of the population of interest. Once an appropriate sensor is chosen, either measurements from the sensor or metrics are calculated for NFI plots and the rest of the population units for the study area.

After compiling NFI plot data and remotely sensed data, one of several estimators used to estimate forest biophysical population parameters (e.g. biomass per unit area) is selected. Detailed descriptions of these estimators can be found in McRoberts et al. (2013), Ståhl et al. (2016), Lister et al. (2020) and McConville et al. (2020). Two estimators were the focus of this study. The simple expansion estimator (hereafter 'expansion estimator') relies solely on the probability-based NFI sample to generate population estimates. The model-assisted estimator uses remotely sensed data to reduce the variance (or increase the precision) of the estimate. To do this, the model-assisted estimator requires a model to be constructed relating NFI plot biomass to remotely sensed metrics. Then, this model is used to predict biomass for each population unit. Population estimates are based on the predicted values for the population units but are adjusted based on the NFI plot observations. More details on these estimators can be found in the case study below.

3 | OUTLIERS AND NONRESPONSE

The term 'outlier' describes observations that are questionable because their values are biologically untenable or statistically different from the rest, which brings into question whether they are mistakenly sampled from a different population or that the sampling procedures or model used to predict the attribute have failed to correctly characterize them (Kendall et al., 1982). This definition allows some observations to be characterized as outliers based on random effects—by chance they are outside of the range of the rest—while others may have known causes due to features of the sampling procedure or measurement protocol. Outliers due to random effects are determined by a threshold for a variable, for example, 1.5 times the interquartile range below the first quartile or above the third quartile. In regression, a threshold can be set for either observed values or their residuals (difference between observed and predicted values), for example, observations with the largest absolute values of residuals and are many standard deviations away from the mean (Draper & Smith, 1981). There are many other types of outliers where values may not reach a statistical threshold, but rather the values are abnormal given the inventory procedure or sensor specification. These

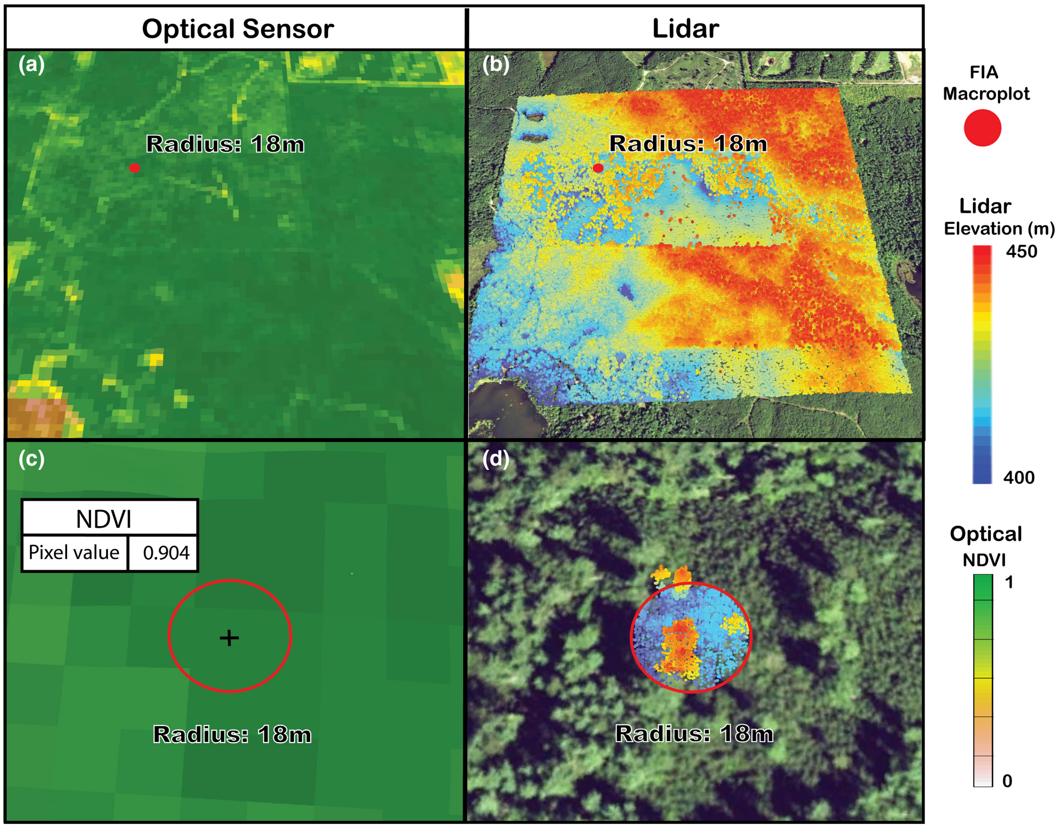

FIGURE 1 Passive and active remote sensing of forest inventory plots. (a, c) Passive optical sensors such as Sentinel-2 report metrics in rasterized (pixel) form, and plot values are typically based on either the average of overlaid pixels or the pixel that coincides with the plot centroid. (b, d) Active sensors such as lidar (right) generate data in point clouds that can be clipped to the inventory plot footprint and allows for the detection of features at a finer resolution. Red circles represent an example Forest Inventory and Analysis (FIA) central macroplot with radius of 18 m. Normalized Difference Vegetation Index (NDVI) (a, c) can range from −1 to 1, but the colour gradient was scaled to NDVI >0 for visualization purposes. Background imagery (b, d) was from the 2010 collection of the National Agriculture Imagery Program (NAIP).

procedural outliers can arise from a variety of sources: human error; accuracy and precision of equipment (such as Global Positioning System [GPS], receivers that are influenced by connectivity, weather, topography or canopy-related interference); sampling design and protocols (such as definitions of forest/non-forest and minimum size classes for measurement); and data availability (such as return window for satellite-based sensors); among others.

Nonresponse—the inability for NFIs to collect data at a given location or time—can lead to missing data. Nonresponse occurs for a variety of reasons for NFIs, including denied access on private lands or hazardous conditions (Domke et al., 2014; Patterson et al., 2012; Westfall et al., 2022). Nonresponse can also arise from plots that cannot be located during remeasurements by field crews, plots located in tidal zones and plots outside of geopolitical boundaries (Burrill et al., 2021). These types of missing data can be missing at random, missing completely at random or missing not at random (Little & Rubin, 2002; Rubin, 1987). Missing at random and missing completely at random occur when missingness is not linked to predictor or response variables. If the probability of nonresponse is equally likely and is not related to a particular cause in an NFI—that is nonresponse is missing completely at random—there are ways to compensate for missing data through the probability design of the NFI by adjusting expansion factors (the amount of land each plot represents; Domke et al., 2014). If nonresponse is missing at random (missing and observed data have systematic differences but these differences can be fully explained by other characteristics of the NFI), strata can be constructed to account for these differences (Domke et al., 2014). However, if data are missing not at random—that is they have a specific cause or are related to a specific stratum or category within the NFI—the probability sample can be compromised (Domke et al., 2014). For example, if nonresponse arises primarily from forests with characteristics unlike the rest of the population (e.g. ownership and management), the conclusions drawn from this subset of data are no longer applicable to all forestland.

4 | SOURCES OF OUTLIERS ARISING FROM GEOSPATIAL DATA HARMONIZATION

4.1 | Procedural outliers due to NFI protocols and definitions

Entities that conduct NFIs identify key populations of interest when designing and conducting field data collection. Setting thresholds and protocols are important for ensuring the population of interest is measured consistently (IPCC, 2019; Tomppo et al., 2010). Thresholds and protocols serve as assumptions in models, provide context to the interpretation of results and guide precision standards for NFIs (Bechtold & Patterson, 2005). Some examples include forest/non-forest definitions; minimum size thresholds for tree measurement; and a nested macro/sub/microplot design to reduce the effort and cost to measure and/or tally smaller trees and seedlings, among others (Figure 2). Thresholds and protocols are informed by entity objectives including cost considerations and precision standards, among many other factors (Tomppo et al., 2010).

NFIs pay close attention to where and when data are collected. In the FIA program, base-intensity field data are collected for approximately one plot per 2400 hectares of forestland. The FIA program collected data intermittently only on productive forestland for the first ~80 years of the program's existence ('periodic' inventory), but beginning in the late 1990s/early 2000s, the FIA program switched from periodic to annual inventories, for which a portion of the plots (e.g. 1/10 to 1/5) on all forestland within a state are measured each year, and each plot is remeasured on a 5- to 10-year interval (generally, 5–7 years in the eastern US and 10 years in the western US; Bechtold & Patterson, 2005). The annual inventory spreads the workload of periodic inventories over multiple years, while gaining better temporal resolution to detect change by revisiting plots every 5–10 years. Some areas intensify sampling above baseline sampling intensity by increasing the number of plots and/or reducing remeasurement period.

NFI thresholds limit what is measured on field plots, but remote sensors are often agnostic to these thresholds and have specifications of their own. The presence of understorey vegetation not measured in the NFI protocol may be predicted by remote sensing, but presence of other non-vegetation objects such as human-made objects or wildlife may also be predicted (Figure 2). NFIs may also set limits on the type of vegetation data collected, for example, whether the NFI includes protocols for down dead wood or herbaceous understorey vegetation (Bechtold & Patterson, 2005). Trees that are near the edge, but just outside of, the NFI plot boundary are not measured by field crews, but remote sensing could detect canopy features of those trees that extend into the plot; vice versa, trees that are just inside the NFI plot boundary will be measured by the field crew even though a substantial portion of the tree may fall outside of the plot (and is therefore clipped out of the lidar point cloud). For pixel-based products, unmeasured vegetation in and around the plot contributes to the value of the pixel.

4.2 | Spatial mismatch of NFI plots between field measurement and remote sensing

Co-registration error of NFI and remotely sensed data can occur due to inaccuracy of plot coordinate measurements made in the field, often obtained through satellite navigation (Dorigo et al., 2010; McRoberts et al., 2018; Pascual et al., 2021). Here, we refer to GPS because it is used by the FIA program but note that other NFIs may use different satellite navigation technologies (i.e. Global Navigation Satellite Systems) that have similar location accuracy concerns. Horizontal root mean square errors (RMSE) of consumer-grade GPS

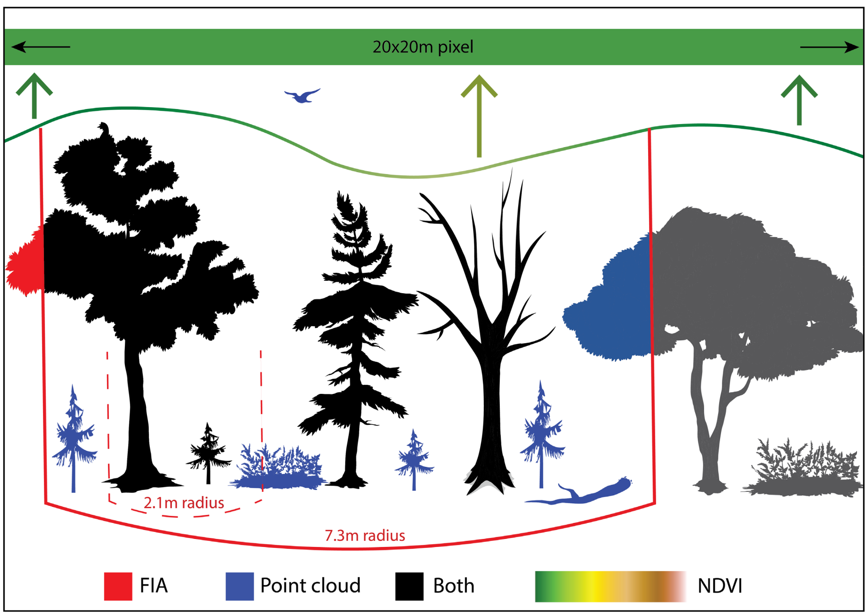

FIGURE 2 Detection of forest attributes by Forest Inventory and Analysis (FIA) and remote sensing. Red cylinders indicate FIA subplot (solid) and microplot (dashed) footprints that are clipped out of a point cloud. Features in black are measured by FIA and included in the point cloud; features in blue may be included in the point cloud but are not measured by FIA; features in red are measured by FIA but not included in the clipped point cloud; features in grey are not measured by FIA and are clipped out of the lidar point cloud but may contribute to pixel-based metrics of lower resolution satellite imagery such as Normalized Difference Vegetation Index (NDVI) computed from Sentinel-2 red and near-infrared bands.

receivers are typically 5–10 m under forest canopy and is influenced by weather conditions, timing of satellites, canopy cover and topography that interrupt satellite connections (Kaartinen et al., 2015). Inaccuracy of coordinate measurements may cause a substantial number of actual plot centres to be located near the measured plot boundary or outside the plot area altogether. The spatial mismatch and resultant inaccuracy of biophysical variables can be partially alleviated by using survey-grade differential GPS receivers (dGPS) (Andersen et al., 2022; Pascual et al., 2021). RMSE of dGPS receivers are less than consumer-grade GPS receivers, with RMSE typically <1 m under closed forest canopy (Hasegawa & Yoshimura, 2003). However, managing dGPS receivers merely for plot centre measurement can be expensive and laborious for large-scale NFIs. Coordinate averaging can also alleviate inaccuracies (McRoberts et al., 2018) but is only possible for plots with multiple remeasurements. Gobakken and Næsset (2009) suggest that increasing plot size (increasing the ratio between plot size and expected spatial error) can help reduce the effects of spatial mismatch by increasing the area of overlap even with location inaccuracy; that is, estimates of forest attributes for larger plots are less sensitive to spatial errors.

Similarly, differences in spatial resolution of remotely sensed data and plot size can lead to spatial mismatch. When plot centroids are used to extract a single pixel from rasterized data, the actual measured area may contain additional pixels or vegetation outside of the measured area that may contribute to the pixel value, especially when plot size differs from pixel size (Figure 2; Hyyppa & Hyyppa, 2001; Muukkonen & Heiskanen, 2007). Plots may also be assigned to an incorrect pixel even when the pixel size is equal to plot size if spatial errors are present (McRoberts, 2010). In addition, mismatches between plot (often circular) and pixel (square) shape may occur (Figure 1c), although the effects of these mismatches are minimal (Packalen et al., 2022).

4.3 | Temporal mismatch

Remotely sensed data are rarely temporally coincident with field-collected data (Oh et al., 2022). Instead, most remote sensing occurs sporadically for aerial-based systems (such as county- or state-acquired lidar) or on daily to weekly time-scales for satellite-based systems (such as a 16-day return period for Landsat 8). In addition, data from satellite-based sensors may need to be compiled across a wider period to capture days without cloud cover. Because of temporal mismatches, forests can change between field measurement and remote sensing. Changes vary in magnitude and cause (biotic, abiotic or anthropogenic), from small, natural changes associated with stand development and aging, to moderate disturbances due to pests, wildlife or silviculture, to large disturbances that remove the forest completely such as stand-replacing disturbance (e.g. wildfire, flooding and windthrow), clear-cut harvesting and land conversion (Fitts et al., 2022). Depending on the sequence of events (remote sensing followed by disturbance followed by field sampling, or vice versa), forest parameters may be under- or over-estimated when harmonizing remotely sensed and NFI data. The selection of an appropriate time window or revisit time should consider the magnitude of change in the population parameter of interest. For example, biomass is accumulated slowly so a multiple-year window may be appropriate, but remote sensing to map wildfires may need a restricted window and/or use a sensor with a higher frequency revisit time.

4.4 | Statistical outliers

Some outliers create obvious data anomalies, where algorithms based on remotely sensed data predict the presence of trees and/or other vegetation even though field crews did not collect data on these plots (e.g. non-forest plots with trees in the FIA program). However, others are more subtle and may only be identified when considering a statistical model relating remotely sensed metrics to field-collected data. In such cases, plots have abnormally large model residuals which contribute to loss of accuracy and precision for resource estimates (Draper & Smith, 1981). Statistical outliers can be identified by evaluating the distribution of residuals and determining if any are larger than a certain threshold (McRoberts et al., 2017). Data can also have high leverage (different from the rest of the predictor values, detected by large Cook's distance), and may be influential (shift parameter estimates when retained/removed).

5 | HARMONIZING NFI DATA WITH LIDAR AND SENTINEL-2: A CASE STUDY FROM SAWYER COUNTY, WISCONSIN, USA

5.1 | Data and methods

5.1.1 | NFI data from the FIA program

We focused our case study on Sawyer County, WI, USA (ca. 45.89 N, 91.14 W, Figure S1) because (1) airborne laser scanning (ALS) lidar data had become recently available for the county (collected in 2017) and (2) the county has been well-sampled by the FIA program due to the large amounts of forest cover and double intensified sampling over the study period (292 plots, approximately one plot per 1200 ha forestland). Because the publicly available FIA database anonymizes some plot locations (perturbing coordinates and swapping plot attributes between a small proportion of similar plots within counties to maintain plot integrity and landowner privacy; Bechtold & Patterson, 2005), we used actual locations of FIA plots obtained from the National Information Management System internal to the FIA program. Plots measured between 2012 and 2019 were included to obtain full coverage of the study area. Each plot consisted of a central 168.1 m^2 (7.31 m radius) circular subplot with three surrounding subplots with centroids at 0-, 120- and 240-degree azimuths (Bechtold & Patterson, 2005). Because observations on subplots are spatially autocorrelated with each other, we focused solely on the central subplot.

We compiled biomass estimates for our study from the FIA database (Burrill et al., 2021). The FIA program uses a component ratio method to predict biomass for individual trees using diameter at breast height, tree height and other stand characteristics (Woodall et al., 2011). Predictions were calculated at the individual tree level for live and standing dead trees and expanded to per-unit-area (megagrams [metric tonnes] per hectare, $\mathrm{Mg\,ha^{-1}}$) using the FIA expansion factors for the proportion of forest represented on each FIA plot.

5.1.2 | Remotely sensed data collection and processing

ALS data for Sawyer County were acquired from the Wisconsin State Cartographer's Office. ALS data were collected in April of 2017 using a Leica ALS 70 lidar sensor at an altitude of 1800 m with a 40-degree field of view. The average pulse density was approximately three returns per m^2. Applying the area-based approach to lidar inventory (Næsset, 2002; White et al., 2013), we summarized vegetation heights within each central subplot using mean, maximum and quantile return heights in 5% intervals, and vertical standard deviation, skewness, kurtosis and entropy of the distribution of return heights as estimates of vertical heterogeneity. The same metrics were then generated for a rectangular grid (pixels) tessellating the county at a resolution that closely matched the area of the central subplots. We excluded return points less than 2 m above ground to filter out understorey vegetation and greater than 45.7 m (the tallest recorded tree for Wisconsin in the FIA database) to filter out false returns that are considered noise (Figure 2). We also set height = 0 when there were fewer than 10 vegetation returns (<2% vegetation on average) per pixel or FIA plot because these were suspected to be incorrectly identified as vegetation by the instrumentation's point classification system.

Surface reflectance data from the European Space Agency's Copernicus Sentinel-2 (Drusch et al., 2012) were acquired from 2017 to 2020. We removed clouds, split data into seasons and composited across years using the median value for each cloud-free pixel. We calculated Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index and generated tasselled cap bands following Shi and Xu (2019). The resulting collection of Sentinel-2 data included 33 bands at 20×20 m pixel resolution. To facilitate the use of both the lidar and Sentinel-2 data, we used two versions of ALS data: (1) ALS metrics were calculated from vegetation returns in the circular FIA central subplot and tessellated to a subplot-sized grid (described above) for models harmonizing FIA data with ALS only, and (2) gridded ALS pixel data were reprojected, resampled and aligned to the 20×20 m resolution Sentinel-2 images for models harmonizing FIA data with both ALS and Sentinel-2. Plots were assigned remotely sensed metrics from this 20×20 m ALS raster dataset based on the plot centre coordinates and coincident pixels.

5.1.3 | Identifying outliers

We used multiple criteria to identify potential outliers and constructed subsets of the data with and without these potential outliers (Table 1). First, scatter graphs of biomass versus ALS metrics (Figure 3) revealed that many plots had positive ALS return heights but no biomass reported by FIA. Although these could be the result of some of the following causes, we simply flagged each of these plots as potential outliers. Second, following McRoberts et al. (2018), we identified plots with potential GPS errors. We consulted all records of each plot remeasurement since FIA began the annual survey in 1999. For each plot in each year of sampling, we calculated the difference between the recorded plot coordinates and the average of the coordinates across all remeasurements of each plot (generally 3–5 measurements per plot). Then, we identified plots whose coordinates deviated by more than 9.29 m from the average plot coordinates; these would share less than 25% of the same area (McRoberts et al., 2018). Third, temporal mismatch is a likely issue when comparing NFI data to lidar data. Because NFI plots were visited across many years to complete the full panel, plot measurement dates for only 51 of 292 plots (17%) coincided with the year of ALS data collection. We flagged plots as potential outliers when there was a clear mismatch between biomass and ALS return height (zero biomass reported by FIA but positive ALS return height) and the NFI plot was inventoried before or after 2017 (i.e. ALS collection year). Fourth, although we planned to flag all causes of nonresponse (10 possible nonresponse codes in the FIA database; Burrill et al., 2021), only denied access plots occurred in our dataset. Fifth, plots with zero biomass recorded by FIA and ALS indicating no positive vegetation returns are theoretically correctly identified; however, inclusion of excessive zeros (zero-inflation of the predictor and response distributions) may question the validity of model assumptions as described below (Ancelet et al., 2010). When a model with a non-zero intercept is used to estimate biomass as a function of remotely sensed metrics, plot observations at the origin increase the variance (have a non-zero residual) and may be influential (outside of the distribution of the rest of data) which can have non-negligible effects on the shape of the relationship. When a model is forced through the origin, these plots have zero residuals and only affect the relationship via increase/decrease in sample size (the value of n in Equations 5 and 6 below). Sixth, our data showed likely heteroscedasticity (Figure 3), so we used an approach that accounted for unequal variances to identify plots with statistically significant residuals following McRoberts et al. (2018). To do so, we constructed a nonlinear model of biomass as a function of ALS mean return height (Z_{mean}) (see Equation 3 below), grouped observations into 12 equal size groups ordered on Z_{mean}, calculated within-group variance of residuals and mean of Z_{mean}, estimated within-group variance as a function of within-group Z_{mean}, and estimated within-group variance (Var_{pred}) for each observation.

TABLE 1 Description, criteria and quantity of different outliers identified in Sawyer County.

Outlier class	Description	Criteria	Number of outliers (% out of 292 total)
Zero biomass, positive ALS returns	ALS indicated the presence of trees, but FIA reported no biomass	Z_{mean}>0, biomass = 0	49 (16.8%)
GPS errors	Spatial mismatch between coordinates of the sample and the mean coordinates across all remeasurements used to clip the ALS point cloud and extract Sentinel-2 variables	Recorded GPS coordinates >9.29 m from average coordinates	16 (5.5%)
Temporal mismatches	FIA collected data before or after ALS data collection, but there was a mismatch (FIA reported no biomass but ALS indicated the presence of trees)	FIA measurement year ≠ 2017, FIA reported 0 biomass but ALS vegetation return heights >0 m	FIA before lidar: 31 (10.6%) FIA after lidar: 9 (3.1%)
Nonresponse	FIA plot was not measured due to nonresponse. In Sawyer County, the only cause of nonresponse was denied access	Non-sampled reason code field populated in FIA database	14 (4.8%)
Double zero	No positive vegetation returns from ALS, and FIA reported no biomass	Z_{mean} = 0, biomass = 0	50 (17.1%)
Statistical outliers	Large residuals (observed − predicted biomass)	Residuals outside of confidence interval while accounting for heteroscedasticity (see *Identifying outliers* section)	12 (4.1%)
All	Plots meet any of the outlier criteria	See above criteria	111 (38.0%)

Abbreviations: ALS, airborne laser scanning; FIA, Forest Inventory and Analysis; GPS, Global Positioning System.

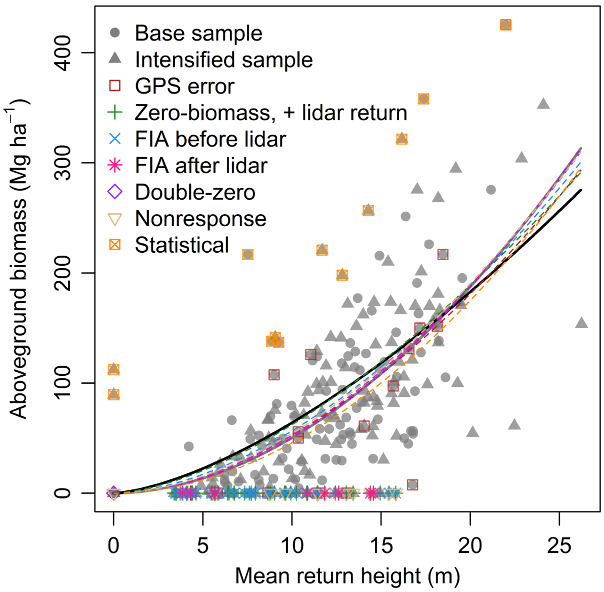

FIGURE 3 Relationship between biomass and mean airborne laser scanning (ALS) return height and the presence of outliers. Grey symbols indicate Forest Inventory and Analysis (FIA) plots (circles = base intensity, triangles = intensified sample) and are marked with an additional symbol if they were part of an outlier subset. Solid black and grey lines indicate the relationship when all or none of the outliers are removed, respectively. Dashed lines indicate the relationship when different classes of outliers are removed. Note that temporal mismatch classes (FIA before and after lidar) are subsets of the zero-biomass, positive lidar return class, and that 50 plots (double zero) are at the origin.

We then evaluated which observed residuals fell outside of the 95% confidence interval: $\mathrm{CI} = 0 \pm 1.96\sqrt{\mathrm{Var}_{\mathrm{pred}}}$.

5.1.4 | Estimating biomass per unit area using simple expansion and model-assisted estimators

We used two approaches for estimating county-wide mean above-ground biomass per unit area, the simple expansion estimator and the model-assisted estimator (Cochran, 1991; McRoberts, Gobakken, et al., 2012a; McRoberts et al., 2013; Särndal et al., 2003). The expansion estimator relies on the assumption of an equal probability NFI sample and does not require auxiliary data (except for the identification of outliers). Using the expansion estimator, mean biomass per unit area ($\hat{\mu}_{\mathrm{EXP}}$) can be expressed as:

$$\hat{\mu}_{\mathrm{EXP}} = \frac{1}{n-1}\sum_{i=1}^{n} y_i, \quad (1)$$

and variance of the estimate of the mean ($\hat{\mathrm{Var}}(\hat{\mu}_{\mathrm{EXP}})$) can be estimated as:

$$\hat{\mathrm{Var}}(\hat{\mu}_{\mathrm{EXP}}) = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{\mu}_{\mathrm{EXP}})^2, \quad (2)$$

where y_i is the total biomass at each of the n FIA plots (McRoberts, Gobakken, et al., 2012). The expansion estimator serves as a baseline to compare the effectiveness of incorporating remotely sensed data in the model-assisted estimator.

For the model-assisted estimator, we compared two models that related biomass to remotely sensed data. First, we used a nonlinear regression model of the form:

$$y_i = \beta_1 x_i^{\beta_2} + \varepsilon_i, \tag{3}$$

where β_1 and β_2 are parameters to be estimated and x_i is the remotely sensed metric of interest, and ε_i is the residual error. This model was fit in R (version 3.6.0, R Core Team, 2019) using the nonlinear least squares ('nls') function.

Second, we used linear regression models of the form:

$$y_i = \beta_0 + \sum_{j=1}^{J} \beta_j x_{i,j} + \varepsilon_i, \tag{4}$$

and

$$y_i = \sum_{j=1}^{J} \beta_j x_{i,j} + \varepsilon_i, \tag{5}$$

where β_j is the jth parameter to be estimated for each of J remotely sensed metrics $x_{i,j}$. Equation 4 differs from Equation 5 in that it includes a non-zero intercept. Fitting linear models and calculation of the model-assisted estimators were performed in R using the linear model ('lm') function.

Using the different model forms with the model-assisted estimator, the estimate of mean biomass per unit area can be expressed as:

$$\hat{\mu}_{\mathrm{MA}} = \frac{1}{N} \sum_{i=1}^{N} \hat{y}_i - \frac{1}{n} \sum_{i=1}^{n} \left(\hat{y}_i - y_i\right), \tag{6}$$

where N is the population size (pixels in the county), and $\hat{y}_i$ is the predicted value of biomass from Equations 3, 4, or 5. The variance of the biomass estimate can be expressed as:

$$\hat{\mathrm{Var}}\left(\hat{\mu}_{\mathrm{MA}}\right) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \left(\varepsilon_i - \bar{\varepsilon}\right)^2, \tag{7}$$

where $\varepsilon_i = \hat{y}_i - y_i$.

To show how the inclusion or removal of outliers affects variance estimates relative to a null baseline, we calculated a variance ratio (VR):

$$\mathrm{VR}_{\mathrm{NULL/EST}} = \frac{\hat{\mathrm{Var}}\left(\hat{\mu}_{\mathrm{NULL}}\right)}{\hat{\mathrm{Var}}\left(\hat{\mu}_{\mathrm{EST}}\right)}, \tag{8}$$

where $\hat{\mathrm{Var}}\left(\hat{\mu}_{\mathrm{NULL}}\right)$ is the variance of the expansion or model-assisted estimator of the full sample and $\hat{\mathrm{Var}}\left(\hat{\mu}_{\mathrm{EST}}\right)$ is the variance obtained on an outlier-removed subset using Equations 2 and 7 for expansion and model-assisted estimators, respectively. The VR was also used to compare model-assisted estimators to the null expansion estimator using all data (i.e. no outliers removed); this special case is relative efficiency and can be interpreted as the factor by which the sample size would have to be increased for the expansion estimator to produce the same variance as the model-assisted estimator.

We used two data sources for predictors in model-assisted estimators, ALS and Sentinel-2. We calculated a correlation coefficient between biomass and a suite of ALS and Sentinel-2 metrics and chose the predictors with the largest correlations to include in our models for the model-assisted estimators (Figure S2). Our models included a nonlinear model with mean vegetation return height (Z_{mean}) from ALS and a linear model with autumn tasselled cap band 2 (greenness, TC2) from Sentinel-2. We also included the linear model with the lowest corrected Akaike information criterion that combined ALS with Sentinel-2, which used ALS Z_{mean} and Sentinel-2 spring NDVI.

5.2 | Case study results and discussion

In total, we flagged 111 of 292 FIA plots (38.0%) as potential outliers. Most of these outliers (99, 33.9% of the total FIA plots) had zero biomass recorded by FIA, and 50 plots (17.1%) also had zero positive vegetation returns (Table 1, Figure 3). However, the remaining 61 plots (20.8%) had potential issues. For example, plots with no biomass but positive ALS returns were abundant (49 plots, 16.8%). Of these, 31 plots (10.6%) were measured before, and 9 plots (3.1%) were measured after 2017 (the year ALS data were collected). In addition, we found 16 plots (5.5%) with potential GPS error. Of these plots, 4 plots (1.4%) had no recorded biomass, but 12 plots (4.1%) had some reported biomass. Nonresponse from denied access accounted for 14 plots (4.8%). Finally, our analysis of model residuals indicated that 12 plots (4.1%) were statistical outliers. Although the proportions of these different types of outliers might be context dependent, extrapolating to the full NFI of approximately 130,000 plots, the removal of the most abundant outlier class (zero biomass but positive ALS returns) could lead to upwards of 22,000 plots removed from national-scale biomass estimation.

We produced estimates of county-wide biomass per unit area when retaining and removing outliers from the analysis using expansion and model-assisted estimators. We compared mean and variance of biomass estimates, and whether model-assisted versus expansion estimators were more sensitive to the inclusion of outliers (Figure 4, Tables 2 and 3). We found that mean biomass estimates were smaller using the nonlinear model-assisted estimator (Equation 3) versus the expansion estimator (all points below the 1:1 line in Figure 4) but were similar when comparing the nonlinear model-assisted estimator using Z_{mean} versus the linear model-assisted (with no intercept) estimator using Z_{mean} + NDVI as predictors (Equation 5, Figure S3). The model using only the Sentinel-2 TC2 as a predictor was less accurate than models that included ALS Z_{mean} and was removed from the rest of the analysis. Relative efficiency was >1.88 (approximately two- to threefold reduction in variance, Table 3) when using the linear and nonlinear model-assisted versus expansion estimator, regardless of the removal of any class of outlier, indicating a successful harmonization of ALS, Sentinel-2 and FIA plot data. VR comparing the model-assisted estimators on the full sample versus outlier-removed subsets showed that removing statistically significant

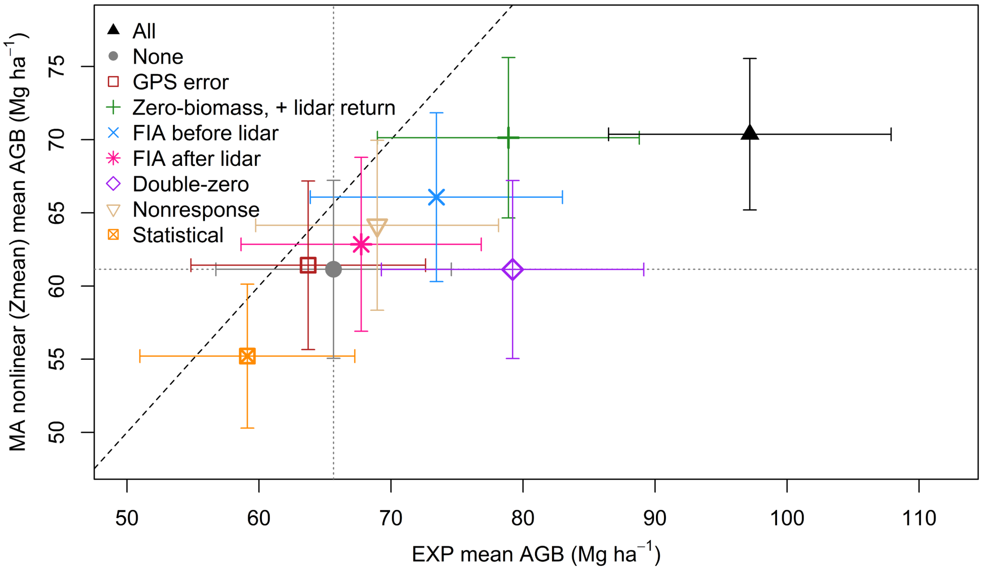

FIGURE 4 Mean biomass per unit area (AGB, Mg ha^{-1}) estimated using expansion (EXP) and model-assisted (MA) estimators. Error bars represent 95% confidence intervals. Dashed line indicates 1:1 relationship and dotted lines indicate null model (no outliers removed) mean biomass.

TABLE 2 Estimates of mean ($\hat{\mu}$) and variance ($\widehat{\text{Var}}(\hat{\mu})$) of above-ground biomass (Mg ha^{-1}) using model-assisted linear (MA_lin) and nonlinear (MA_nl) and simple expansion (EXP) estimators.

Outliers removed	n	$\hat{\mu}_{EXP}$	$\widehat{\text{Var}}(\hat{\mu}_{EXP})$	$\hat{\mu}_{MA\_nl}$	$\widehat{\text{Var}}(\hat{\mu}_{MA\_nl})$	$\hat{\mu}_{MA\_lin}$	$\widehat{\text{Var}}(\hat{\mu}_{MA\_lin})$
None	292	65.65	20.70	61.14	9.63	63.84	11.02
FIA after lidar	283	67.74	21.54	62.85	9.19	65.52	10.59
Statistically significant	280	59.12	17.24	55.21	6.30	58.15	7.71
Nonresponse	278	68.95	22.02	64.15	8.77	66.68	10.12
GPS error	276	63.73	20.55	61.42	8.63	63.88	9.68
FIA before lidar	261	73.44	23.72	66.07	8.65	69.48	9.69
Zero biomass, positive lidar returns	243	78.89	25.60	70.13	7.83	73.68	8.77
Double zero	242	79.21	25.70	61.13	9.63	54.38	10.53
All	181	97.18	29.81	70.37	6.96	66.48	7.51

Abbreviations: FIA, Forest Inventory and Analysis; GPS, Global Positioning System.

TABLE 3 Relative efficiency (RE) comparing the variance of the null expansion (EXP) estimator (all data included) to the model-assisted (MA) estimators and variance ratio (VR) comparing the MA estimator from outlier-removed subsets to the null model-assisted estimator (all data included).

Outliers removed	MA nonlinear RE	MA nonlinear VR	MA linear RE	MA linear VR
None	2.15	—	1.88	—
FIA after lidar	2.25	1.05	1.95	1.04
Statistically significant	3.29	1.53	2.68	1.43
Nonresponse	2.36	1.10	2.05	1.09
GPS error	2.40	1.16	2.14	1.14
FIA before lidar	2.39	1.11	2.14	1.14
Zero biomass, positive lidar returns	2.64	1.23	2.36	1.26
Double zero	2.15	1.00	1.97	1.05
All	2.97	1.38	2.76	1.47

Abbreviations: FIA, Forest Inventory and Analysis; GPS, Global Positioning System.

outliers had the largest reduction of variance ($VR_{NONE/STAT}$ = 1.53 and 1.43 for the nonlinear and linear model-assisted estimators, respectively, Table 3).

Removal of outliers had larger effects on the expansion estimator than on the model-assisted estimator. Our model-assisted estimator used a nonlinear model that is forced through the intercept which led to plots with biomass = 0 having zero residuals. Removal of any type of zero biomass outliers had larger effects on estimates of mean biomass per unit area using the expansion estimator, where removing zero-biomass outliers decreased denominator in Equation 1 (Figure 4). Because GPS errors happen sporadically (i.e. approximately equal number of positive and negative residuals, Figure 3), we found that the removal of data for plots with potential erroneous GPS coordinates had the smallest differences for estimates of mean biomass per unit area (Table 2). In total, when all outliers were removed, estimates of mean and variance of biomass per unit area increased by 48% and 44%, respectively, in the expansion estimator, versus 15% increase in estimated mean and 28% reduction in estimated variance of biomass in the nonlinear model-assisted estimator.

6 | WHAT TO DO WITH OUTLIERS?

There are many proposed methods for dealing with outliers. Often, outliers that arise from protocol are considered in the QA/QC procedure of data analysis and are removed from the sample (McRoberts et al., 2017). These data are often removed from the analysis because issues can be easily traced to their source within the NFI protocol/definitions or data quality issues. However, the choice to retain or remove these data can have consequences for the estimation of forest population parameters (e.g. mean biomass per unit area), especially when using an estimator that relies solely on the field-collected data (i.e. the expansion estimator). In our case study, 4.1%–16.8% of FIA plots may be outliers if defined by a single criterion or 38% if all criteria are used. The removal from the analysis greatly reduces sample size and brings into question whether the sample can still be characterized as a probability sample of the NFI. The contribution of each sample unit to population estimates using design-based estimators decreases as sample size increases (Equations 1 and 6), so the removal of outliers in small area estimation (when sample size is already limited) can lead to unstable population estimates (Breidenbach & Astrup, 2012; Goerndt et al., 2011).

Another option is to use methods that are more robust to outliers. Using zero-inflated models that account for the presence of structural zeros (i.e. those that arise from protocol definitions) in addition to random zeros (i.e. 'true' zeros that arise from variability in the attribute of interest) can help reduce the effects of zeros on the relationship between predictor and response, but methods to do so are rarely applied to forest biomass estimation (Ancelet et al., 2010). Rubin (1987) suggests that instead of assuming these values are zeros or removing them from the dataset, multiple imputation (i.e. imputing the value multiple times based on a distribution rather than a single time based on the mean) can help to retain the sample size while accounting for uncertainty in the 'true' value of the missing data. Models such as k-nearest neighbours, random forests and other machine learning algorithms are potentially robust to outliers, but interpretability of relationships and sensitivity to other characteristics of the data (such as sample size) may hinder their applicability (McRoberts et al., 2017; Yang et al., 2022).

There is the option of doing nothing with the outliers, and the effects may vary based on the estimator. For the expansion estimator, when outliers are removed, the sample size decreases, which can cause an increase in uncertainty (Equation 2). Our case study from Wisconsin shows this trade-off: estimated variance increased when all types of outliers except statistically significant residuals or GPS errors were removed. By using a model-assisted estimator, estimated variance was reduced ($VR > 1$) for all subsets, indicating that model-assisted estimators were less sensitive to reduced sample size.

There are programmatic changes that can be made to alleviate future effects of outliers. Lister et al. (2020) highlight the use of remote sensing in pre-field analysis by FIA (i.e. to reduce the likelihood and associated cost of sending a field crew to measure a non-forested area) and how the evolution of the technology and imagery used in this process have increased precision of the NFI. Continued improvements to pre-field remote sensing can help reduce the number of outliers by improved detection of forest changes that spatially and temporally align with field inventories. Uncertainty in plot locations when using commercial grade GPS receivers (errors as great as 10 m) can be partially alleviated if NFIs invest in survey grade dGPS units. When the use of dGPS is economically or operationally unfeasible, some GPS errors can be reduced through coordinate averaging (McRoberts et al., 2018), although in our case study, GPS errors did not strongly influence the accuracy or precision of mean biomass estimates (Figure 3). Other protocol definitions could be adjusted to better ensure that remote sensing-based models are predicting the same attributes that are being measured in the field. For example, sampling of smaller trees and other understorey vegetation could be expanded to the full subplot area (Figure 2). In our case study, we filtered out ALS vegetation returns <2 m to mimic the lack of understorey sampling in the FIA protocol (Bechtold & Patterson, 2005). Methods for accounting for nonresponse, such as imputing estimates based on previous measurements and/or other NFI plots in the population and improving the process to request access, could reduce the likelihood of nonresponse-related outliers and their impacts on estimation (Domke et al., 2014; Westfall et al., 2022; Westfall & Wilson, 2022). Mismatches in observed and predicted biomass resulting from temporal mismatches between field sampling and remote sensing may be due to harvest or other disturbance, so incorporating more information about drivers of change may help better diagnose temporal mismatches (Edgar & Westfall, 2022; Fitts et al., 2022).

7 | EXTENDING BEYOND THE NFI

Our review focused on NFIs, but discussion on outliers from geospatial data harmonization could be applied to other survey data. The quantity of ecological research networks has increased over recent years and serve a variety of purposes, including measuring gas fluxes using eddy covariance (regional flux tower sampling networks embedded within the international FLUXNET network), sampling aquatic and terrestrial biodiversity (United States National Ecological Observatory Network, NEON, Forest Global Earth Observatory Network, ForestGEO, among others), and coordinating experiments across sites (Long Term Ecological Research, LTER, Nutrient Network, NutNet, among others). These networks are conducting their own remote sensing (such as the airborne observation platform from NEON or eddy covariance data from FLUXNET) and/or compiling other sources of remotely sensed data in a similar manner to the workflow described here (Davies et al., 2021; Friend et al., 2007; Kamoske et al., 2022; Musinsky et al., 2022). For example, NEON data have been used in data harmonization workflows for studying patterns of biodiversity (Kamoske et al., 2022), but the relationship between remotely sensed and on-the-ground measurements may be less reliable than expected (Pau et al., 2022). ForestGEO sites collect field inventory data over the course of many years, but remote sensing (commissioned at ForestGEO sites or compilation of other remote sensing) may not be temporally coincident (Davies et al., 2021). As these ecological research networks continue to grow and expand their sampling and analyses, they should continue to address questions about the validity of their data harmonization workflows.

Methods for harmonizing Global Ecosystems Dynamics Investigation (GEDI) satellite-based lidar with other sensors and field inventory data are of recent interest for global biomass mapping (Kellner et al., 2022; Silva et al., 2021). The GEDI L4B gridded biomass product was built using a suite of biomass models (Duncanson et al., 2022), and Dubayah et al. (2022) showed high correlation between GEDI biomass and an aggregated FIA biomass product (64,000 ha hexagons containing an average of 20 FIA plots; Menlove & Healey, 2020). Due to the location errors of FIA plots and GEDI footprints (~10 m each on average), it is uncertain whether spatial co-registration of these two data sources can be achieved at the plot/footprint level (Roy et al., 2021). Additionally, GEDI is near the end of its life span, and it is uncertain if GEDI or a similar programme will exist in the future. Practitioners incorporating GEDI or any other remotely sensed data in geospatial data harmonization workflows should consider whether the longevity of a particular remote sensing mission will preclude its use in future applications.

NFI data are being used to assess carbon baselines and additionality—the relative increases in carbon stored on the landscape due to forest management practices such as delayed harvest, reforestation and prevention of land conversion—as potential natural climate solutions (Badgley et al., 2022; Giebink, Domke, et al., 2022). These baselines rely on the assumption that field-collected data and remote sensing are properly linking remotely sensed metrics to carbon stocks or biomass. But, as shown in our case study, the precision and accuracy of these baseline estimates may be sensitive to potential outliers. Furthermore, if local calibration and validation of model estimates and/or data products are conducted and sample plots are removed to compensate for outliers identified as part of this process, that may jeopardize the underlying probability sample used to calculate design-based estimates of the population parameter of interest (Anderson-Teixeira & Belair, 2022; Badgley et al., 2022).

In summary, our review and case study serve as a note of caution to researchers and practitioners aiming to harmonize remotely sensed data with field-collected data to ensure that what is being measured in the field is being predicted by models based on their remote sensing technologies of choice. The inclusion or removal of outliers may have large effects on the estimation of forest resources, and therefore careful attention should be paid to the workflow used and whether the treatment of outliers will impact the conclusions drawn from geospatial data harmonization.

AUTHOR CONTRIBUTIONS

Jonathan A. Knott, Greg C. Liknes, Ronald E. McRoberts, Grant M. Domke and Valquiria F. Quirino developed the idea; Greg C. Liknes and Brian F. Walters compiled and processed data; Jonathan A. Knott, Greg C. Liknes and Ronald E. McRoberts analysed the data; Jonathan A. Knott drafted the initial manuscript; Courtney L. Giebink created conceptual figures (Figures 1 and 2); and all authors contributed to writing and editing.

ACKNOWLEDGEMENTS

Funding for this project was through the USDA Forest Service, Northern Research Station, Forest Inventory and Analysis Program. The findings and conclusions in this publication are solely of the authors and should not be construed to represent any official USDA or US Government determination or policy.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to disclose.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.14084.

DATA AVAILABILITY STATEMENT

NFI data were obtained from the USDA Forest Service, FIA Program. Publicly available FIA data are available through the FIA DataMart (https://apps.fs.usda.gov/fia/datamart/datamart.html); however, as noted in the methods, we used a federally protected version of the FIA database to obtain actual (unperturbed) plot locations and biomass data for our analyses. Lidar data are available from the Wisconsin State Cartographer's Office (https://www.sco.wisc.edu/data/elevationlidar/). Sentinel-2 data are available through Google Earth Engine (https://earthengine.google.com/).

ORCID

Jonathan A. Knott https://orcid.org/0000-0003-3856-8454
Courtney L. Giebink https://orcid.org/0000-0003-0635-0020
Sungchan Oh https://orcid.org/0000-0003-2337-9693
Grant M. Domke https://orcid.org/0000-0003-0485-0355
Valquiria F. Quirino https://orcid.org/0000-0002-5150-2735
Brian F. Walters https://orcid.org/0000-0001-7408-9704

REFERENCES

Ancelet, S., Etienne, M. P., Benoît, H., & Parent, E. (2010). Modelling spatial zero-inflated continuous data with an exponentially compound Poisson process. *Environmental and Ecological Statistics*, *17*(3), 347–376. https://doi.org/10.1007/s10651-009-0111-6

Andersen, H.-E., Strunk, J., & McGaughey, R. J. (2022). Using high-performance global navigation satellite system technology to improve Forest Inventory and Analysis plot coordinates in the Pacific region. In *Gen. Tech. Rep. PNW-GTR-1000* (Vol. *1000*, pp. 1–38). U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 38 p. (online only).

Anderson-Teixeira, K. J., & Belair, E. P. (2022). Effective forest-based climate change mitigation requires our best science. *Global Change Biology*, *28*(4), 1200–1203. https://doi.org/10.1111/gcb.16008

Averill, C., Fortunel, C., Maynard, D. S., van den Hoogen, J., Dietze, M. C., Bhatnagar, J. M., & Crowther, T. W. (2022). Alternative stable states of the forest mycobiome are maintained through positive feedbacks. *Nature Ecology and Evolution*, *6*, 375–382. https://doi.org/10.1038/s41559-022-01663-9

Babcock, C., Finley, A. O., Andersen, H. E., Pattison, R., Cook, B. D., Morton, D. C., Alonzo, M., Nelson, R., Gregoire, T., Ene, L., Gobakken, T., & Næsset, E. (2018). Geostatistical estimation of forest biomass in interior Alaska combining Landsat-derived tree cover, sampled airborne lidar and field observations. *Remote Sensing of Environment*, *212*, 212–230. https://doi.org/10.1016/j.rse.2018.04.044

Badgley, G., Freeman, J., Hamman, J. J., Haya, B., Trugman, A. T., Anderegg, W. R. L., & Cullenward, D. (2022). Systematic over-crediting in California's forest carbon offsets program. *Global Change Biology*, *28*(4), 1433–1445. https://doi.org/10.1111/gcb.15943

Baer, K. C., & Gray, A. N. (2022). Biotic predictors improve species distribution models for invasive plants in Western U.S. forests at high but not low spatial resolutions. *Forest Ecology and Management*, *518*, 120249. https://doi.org/10.1016/j.foreco.2022.120249

Bechtold, W. A., & Patterson, P. L. (2005). *The enhanced forest inventory and analysis program—National sampling design and estimation procedures* (Vol. *80*). USDA Forest Service, Southern Research Station.

Breidenbach, J., & Astrup, R. (2012). Small area estimation of forest attributes in the Norwegian National Forest Inventory. *European Journal of Forest Research*, *131*(4), 1255–1267. https://doi.org/10.1007/s10342-012-0596-7

Burrill, E. A., DiTommasa, A. M., Turner, J. A., Pugh, S. A., Menlove, J., Christiansen, G., Perry, C. J., & Conkling, B. L. (2021). *The Forest inventory and analysis database: Database description and user guide version 9.0 for phase 2* (p. 1024). U.S. Department of Agriculture, Forest Service. http://www.fia.fs.fed.us/library/database-documentation

Carteron, A., Vellend, M., & Laliberté, E. (2022). Mycorrhizal dominance reduces local tree species diversity across US forests. *Nature Ecology and Evolution*, *6*, 370–374. https://doi.org/10.1038/s41559-021-01634-6

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B. W., Ogawa, H., Puig, H., Riéra, B., & Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, *145*(1), 87–99. https://doi.org/10.1007/s00442-005-0100-x

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B. C., Duque, A., Eid, T., Fearnside, P. M., Goodman, R. C., Henry, M., Martínez-Yrízar, A., Mugasha, W. A., Muller-Landau, H. C., Mencuccini, M., Nelson, B. W., Ngomanda, A., Nogueira, E. M., Ortiz-Malavassi, E., … Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*, *20*(10), 3177–3190. https://doi.org/10.1111/gcb.12629

Chirici, G., McRoberts, R. E., Winter, S., Bertini, R., Bröändli, U. B., Asensio, I. A., Bastrup-Birk, A., Rondeux, J., Barsoum, N., & Marchetti, M. (2012). National forest inventory contributions to forest biodiversity monitoring. *Forest Science*, *58*(3), 257–268. https://doi.org/10.5849/forsci.12-003

Cochran, W. G. (1991). *Sampling Techniques* (3rd ed.). John Wiley & Sons.

Colmanetti, M. A. A., Weiskittel, A., Barbosa, L. M., Shirasuna, R. T., de Lima, F. C., Ortiz, P. R. T., Catharino, E. L. M., Barbosa, T. C., & do Couto, H. T. Z. (2018). Aboveground biomass and carbon of the highly diverse Atlantic Forest in Brazil: Comparison of alternative individual tree modeling and prediction strategies. *Carbon Management*, *9*(4), 383–397. https://doi.org/10.1080/17583004.2018.1503040

Crookston, N. L., Rehfeldt, G. E., Dixon, G. E., & Weiskittel, A. R. (2010). Addressing climate change in the forest vegetation simulator to assess impacts on landscape forest dynamics. *Forest Ecology and Management*, *260*(7), 1198–1211. https://doi.org/10.1016/j.foreco.2010.07.013

Davies, S. J., Abiem, I., Abu Salim, K., Aguilar, S., Allen, D., Alonso, A., Anderson-Teixeira, K., Andrade, A., Arellano, G., Ashton, P. S., Baker, P. J., Baker, M. E., Baltzer, J. L., Basset, Y., Bissiengou, P., Bohlman, S., Bourg, N. A., Brockelman, W. Y., Bunyavejchewin, S., … Zuleta, D. (2021). ForestGEO: Understanding forest diversity and dynamics through a global observatory network. *Biological Conservation*, *253*, 108907. https://doi.org/10.1016/j.biocon.2020.108907

Domke, G. M., Woodall, C. W., Smith, J. E., Westfall, J. A., & McRoberts, R. E. (2012). Consequences of alternative tree-level biomass estimation procedures on U.S. forest carbon stock estimates. *Forest Ecology and Management*, *270*, 108–116. https://doi.org/10.1016/j.foreco.2012.01.022

Domke, G. M., Woodall, C. W., Walters, B. F., McRoberts, R. E., & Hatfield, M. A. (2014). Strategies to compensate for the effects of nonresponse on forest carbon baseline estimates from the national forest inventory of the United States. *Forest Ecology and Management*, *315*, 112–120. https://doi.org/10.1016/j.foreco.2013.12.031

Doraisami, M., Kish, R., Paroshy, N. J., Domke, G. M., Thomas, S. C., & Martin, A. R. (2022). A global database of woody tissue carbon concentrations. *Scientific Data*, *9*(284), 1–12. https://doi.org/10.1038/s41597-022-01396-1

Dorigo, W., Hollaus, M., Wagner, W., & Schadauer, K. (2010). An application-oriented automated approach for co-registration of forest inventory and airborne laser scanning data. *International Journal of Remote Sensing*, *31*(5), 1133–1153. https://doi.org/10.1080/01431160903380581

Draper, N. R., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). John Wiley and Sons.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

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. *Environmental Research Letters*, *17*(9), 095001. https://doi.org/10.1088/1748-9326/ac8694

Duncanson, L., Kellner, J. R., Armston, J., Dubayah, R., Minor, D. M., Hancock, S., Healey, S. P., Patterson, P. L., Saarela, S., Marselis, S., Silva, C. E., Bruening, J., Goetz, S. J., Tang, H., Hofton, M., Blair, B., Luthcke, S., Fatoyinbo, L., Abernethy, K., … Zgraggen, C. (2022). Aboveground biomass density models for NASA's global ecosystem dynamics investigation (GEDI) lidar mission. *Remote Sensing of Environment*, *270*, 112845. https://doi.org/10.1016/j.rse.2021.112845

Edgar, C. B., & Westfall, J. A. (2022). Timing and extent of forest disturbance in the Laurentian mixed Forest. *Frontiers in Forests and Global Change*, *5*, 1–15. https://www.frontiersin.org/articles/10.3389/ffgc.2022.963796

EPA. (2022). *Inventory of U.S. greenhouse gas emissions and sinks: 1990–2020* (p. 205). U.S. Environmental Protection Agency, EPA 430-R-22-003. https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2020

Fei, S., Jo, I., Guo, Q., Wardle, D. A., Fang, J., Chen, A., Oswalt, C. M., & Brockerhoff, E. G. (2018). Impacts of climate on the biodiversity-productivity relationship in natural forests. *Nature Communications*, *9*(1), 5436. https://doi.org/10.1038/s41467-018-07880-w

Fei, S., Morin, R. S., Oswalt, C. M., & Liebhold, A. M. (2019). Biomass losses resulting from insect and disease invasions in US forests. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(35), 17371–17376. https://doi.org/10.1073/pnas.1820601116

Fitts, L. A., Domke, G. M., & Russell, M. B. (2022). Comparing methods that quantify forest disturbances in the United States' national forest inventory. *Environmental Monitoring and Assessment*, *194*(304), 1–17. https://doi.org/10.1007/s10661-022-09948-z

Friend, A. D., Arneth, A., Kiang, N. Y., Lomas, M., Ogée, J., Rödenbeck, C., Running, S. W., Santaren, J.-D., Sitch, S., Viovy, N., Ian Woodward, F., & Zaehle, S. (2007). FLUXNET and modelling the global carbon cycle. *Global Change Biology*, *13*(3), 610–633. https://doi.org/10.1111/j.1365-2486.2006.01223.x

Giebink, C. L., DeRose, R. J., Castle, M., Shaw, J. D., & Evans, M. E. K. (2022). Climatic sensitivities derived from tree rings improve predictions of the Forest vegetation simulator growth and yield model. *Forest Ecology and Management*, *517*, 120256. https://doi.org/10.1016/j.foreco.2022.120256

Giebink, C. L., Domke, G. M., Fisher, R. A., Heilman, K. A., Moore, D. J. P., DeRose, R. J., & Evans, M. E. K. (2022). The policy and ecology of forest-based climate mitigation: Challenges, needs, and opportunities. *Plant and Soil*, *479*, 25–52. https://doi.org/10.1007/s11104-022-05315-6

Global Forest Observations Initiative. (2020). *Integrating remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests* (3rd ed., pp. 333). USDA Forest Service.

Gobakken, T., & Næsset, E. (2009). Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, *39*(5), 1036–1052. https://doi.org/10.1139/X09-025

Goerndt, M. E., Monleon, V. J., & Temesgen, H. (2011). A comparison of small-area estimation techniques to estimate selected stand attributes using LiDAR-derived auxiliary variables. *Canadian Journal of Forest Research*, *41*(6), 1189–1201. https://doi.org/10.1139/x11-033

Gonzalez, P., Asner, G. P., Battles, J. J., Lefsky, M. A., Waring, K. M., & Palace, M. (2010). Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. *Remote Sensing of Environment*, *114*(7), 1561–1575. https://doi.org/10.1016/j.rse.2010.02.011

Gonzalez-Akre, E., Piponiot, C., Lepore, M., Herrmann, V., Lutz, J. A., Baltzer, J. L., Dick, C. W., Gilbert, G. S., He, F., Heym, M., Huerta, A. I., Jansen, P. A., Johnson, D. J., Knapp, N., Král, K., Lin, D., Malhi, Y., McMahon, S. M., Myers, J. A., … Anderson-Teixeira, K. J. (2022). Allodb: An R package for biomass estimation at globally distributed extratropical forest plots. *Methods in Ecology and Evolution*, *13*(2), 330–338. https://doi.org/10.1111/2041-210X.13756

Hasegawa, H., & Yoshimura, T. (2003). Application of dual-frequency GPS receivers for static surveying under tree canopies. *Journal of Forest Research*, *8*(2), 103–110. https://doi.org/10.1007/s103100300012

Heilman, K. A., Dietze, M. C., Arizpe, A. A., Aragon, J., Gray, A., Shaw, J. D., Finley, A. O., Klesse, S., DeRose, R. J., & Evans, M. E. K. (2022). Ecological forecasting of tree growth: Regional fusion of tree-ring and forest inventory data to quantify drivers and characterize uncertainty. *Global Change Biology*, *28*(7), 2442–2460. https://doi.org/10.1111/gcb.16038

Horn, K. J., Thomas, R. Q., Clark, C. M., Pardo, L. H., Fenn, M. E., Lawrence, G. B., Perakis, S. S., Smithwick, E. A. H., Baldwin, D., Braun, S., Nordin, A., Perry, C. H., Phelan, J. N., Schaberg, P. G., Clair, S. B. S., Warby, R., & Watmough, S. (2018). Growth and survival relationships of 71 tree species with nitrogen and sulfur deposition across the conterminous U.S. *PLoS ONE*, *13*(10), e0205296. https://doi.org/10.1371/journal.pone.0205296

Hyyppa, H. J., & Hyyppa, J. M. (2001). Effects of stand size on the accuracy of remote sensing-based forest inventory. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(12), 2613–2621. https://doi.org/10.1109/36.974996

Iannone, B. V., Potter, K. M., Guo, Q., Liebhold, A. M., Pijanowski, B. C., Oswalt, C. M., & Fei, S. (2016). Biological invasion hotspots: A trait-based perspective reveals new sub-continental patterns. *Ecography*, *39*(10), 961–969. https://doi.org/10.1111/ecog.01973

IPCC. (2019). *2019 refinement to the 2006 IPCC guidelines for national greenhouse gas inventories* (Vol. 4). Agriculture, Forestry and Other Land Use.

Iverson, L. R., Peters, M. P., Prasad, A. M., & Matthews, S. N. (2019). Analysis of climate change impacts on tree species of the eastern US: Results of DISTRIB-II modeling. *Forests*, *10*(4), 302.

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, *49*(1), 12–35. https://doi.org/10.1093/forestscience/49.1.12

Jo, I., Fei, S., Oswalt, C. M., Domke, G. M., & Phillips, R. P. (2019). Shifts in dominant tree mycorrhizal associations in response to anthropogenic impacts. *Science Advances*, *5*(4), eaav6358.

Kaartinen, H., Hyyppä, J., Vastaranta, M., Kukko, A., Jaakkola, A., Yu, X., Pyörälä, J., Liang, X., Liu, J., Wang, Y., Kaijaluoto, R., Melkas, T., Holopainen, M., & Hyyppä, H. (2015). Accuracy of kinematic positioning using global satellite navigation systems under forest canopies. *Forests*, *6*(9), 3218–3236. https://doi.org/10.3390/f6093218

Kamoske, A. G., Dahlin, K. M., Read, Q. D., Record, S., Stark, S. C., Serbin, S. P., & Zarnetske, P. L. (2022). Towards mapping biodiversity from above: Can fusing lidar and hyperspectral remote sensing predict taxonomic, functional, and phylogenetic tree diversity in temperate forests? *Global Ecology and Biogeography*, *31*, 1440–1460. https://doi.org/10.1111/geb.13516

Kellner, J. R., Armston, J., & Duncanson, L. (2022). Algorithm theoretical basis document for GEDI footprint aboveground biomass density. *Earth and Space Science*, *9*, e2022EA002516. https://doi.org/10.1029/2022EA002516

Kendall, M. G., Buckland, W. R., & Marriott, F. H. (1982). *A dictionary of statistical terms* (Vol. 10). Longman Publishing Group.

Knott, J. A., Jenkins, M. A., Oswalt, C. M., & Fei, S. (2020). Community-level responses to climate change in forests of the eastern United States. *Global Ecology and Biogeography*, *29*(8), 1299–1314. https://doi.org/10.1111/geb.13102

Kovac, M., Bauer, A., & Ståhl, G. (2014). Merging national forest and national forest health inventories to obtain an integrated forest resource inventory—Experiences from Bavaria, Slovenia and Sweden. *PLoS ONE*, *9*(6), e100157. https://doi.org/10.1371/journal.pone.0100157

Lichstein, J. W., Dushoff, J., Ogle, K., Chen, A., Purves, D. W., Caspersen, J. P., & Pacala, S. W. (2010). Unlocking the forest inventory data: Relating individual tree performance to unmeasured environmental factors. *Ecological Applications*, *20*(3), 684–699. https://doi.org/10.1890/08-2334.1

Lister, A. J., Andersen, H., Frescino, T., Gatziolis, D., Healey, S., Heath, L. S., Liknes, G. C., McRoberts, R., Moisen, G. G., Nelson, M., Riemann, R., Schleeweis, K., Schroeder, T. A., Westfall, J., & Tyler Wilson, B. (2020). Use of remote sensing data to improve the efficiency of national forest inventories: A case study from the United States national forest inventory. *Forests*, *11*(12), 1–41. https://doi.org/10.3390/f11121364

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. John Wiley & Sons.

Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., & Moran, E. (2016). A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*, *9*(1), 63–105. https://doi.org/10.1080/17538947.2014.990526

McConville, K. S., Moisen, G. G., & Frescino, T. S. (2020). A tutorial on model-assisted estimation with application to forest inventory. *Forests*, *11*(2), 1–25. https://doi.org/10.3390/f11020244

McRoberts, R. E. (2010). The effects of rectification and global positioning system errors on satellite image-based estimates of forest area. *Remote Sensing of Environment*, *114*(8), 1710–1717. https://doi.org/10.1016/j.rse.2010.03.001

McRoberts, R. E., Chen, Q., & Walters, B. F. (2017). Multivariate inference for forest inventories using auxiliary airborne laser scanning data. *Forest Ecology and Management*, *401*, 295–303. https://doi.org/10.1016/j.foreco.2017.07.017

McRoberts, R. E., Chen, Q., Walters, B. F., & Kaisershot, D. J. (2018). The effects of global positioning system receiver accuracy on airborne laser scanning-assisted estimates of aboveground biomass. *Remote Sensing of Environment*, *207*(September 2017), 42–49. https://doi.org/10.1016/j.rse.2017.09.036

McRoberts, R. E., Gobakken, T., & Næsset, E. (2012). Post-stratified estimation of forest area and growing stock volume using lidar-based stratifications. *Remote Sensing of Environment*, *125*, 157–166. https://doi.org/10.1016/j.rse.2012.07.002

McRoberts, R. E., Næsset, E., & Gobakken, T. (2013). Inference for lidar-assisted estimation of forest growing stock volume. *Remote Sensing of Environment*, *128*, 268–275. https://doi.org/10.1016/j.rse.2012.10.007

McRoberts, R. E., Tomppo, E. O., Schadauer, K., & Ståhl, G. (2012). Harmonizing national forest inventories. *Forest Science*, *58*(3), 189–190. https://doi.org/10.5849/forsci.12-042

Menlove, J., & Healey, S. P. (2020). A comprehensive forest biomass dataset for the USA allows customized validation of remotely sensed biomass estimates. *Remote Sensing*, *12*(24), Article 24. https://doi.org/10.3390/rs12244141

Miles, P. D., & Smith, W. B. (2009). *Specific gravity and other properties of wood and bark for 156 tree species found in North America*. USDA Forest Service.

Musinsky, J., Goulden, T., Wirth, G., Leisso, N., Krause, K., Haynes, M., & Chapman, C. (2022). Spanning scales: The airborne spatial and temporal sampling design of the National Ecological Observatory Network. *Methods in Ecology and Evolution*, *13*(9), 1866–1884. https://doi.org/10.1111/2041-210X.13942

Muukkonen, P., & Heiskanen, J. (2007). Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: A possibility to verify carbon inventories. *Remote Sensing of Environment*, *107*(4), 617–624. https://doi.org/10.1016/j.rse.2006.10.011

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, *80*(1), 88–99. https://doi.org/10.1016/S0034-4257(01)00290-5

Oh, S., Jung, J., Shao, G., Shao, G., Gallion, J., & Fei, S. (2022). High-resolution canopy height model generation and validation using USGS 3DEP LiDAR data in Indiana, USA. *Remote Sensing*, *14*(935), 1–18. https://doi.org/10.3390/rs14040935

Packalen, P., Strunk, J., Maltamo, M., & Myllymäki, M. (2022). Circular or square plots in ALS-based forest inventories—Does it matter? *Forestry: An International Journal of Forest Research*, *96*(1), 49–61. https://doi.org/10.1093/forestry/cpac032

Pascual, A., Guerra-Hernández, J., Cosenza, D. N., & Sandoval-Altalerrea, V. (2021). Using enhanced data co-registration to update Spanish National Forest Inventories (NFI) and to reduce training data under LiDAR-assisted inference. *International Journal of Remote Sensing*, *42*(1), 106–127. https://doi.org/10.1080/01431161.2020.1813346

Patterson, P. L., Coulston, J. W., Roesch, F. A., Westfall, J. A., & Hill, A. D. (2012). A primer for nonresponse in the US forest inventory and analysis program. *Environmental Monitoring and Assessment*, *184*(3), 1423–1433. https://doi.org/10.1007/s10661-011-2051-5

Pau, S., Nippert, J. B., Slapikas, R., Griffith, D., Bachle, S., Helliker, B. R., O'Connor, R. C., Riley, W. J., Still, C. J., & Zaricor, M. (2022). Poor relationships between NEON airborne observation platform data and field-based vegetation traits at a mesic grassland. *Ecology*, *103*(2), 1–10. https://doi.org/10.1002/ecy.3590

Prasad, A., Pedlar, J., Peters, M., McKenney, D., Iverson, L., Matthews, S., & Adams, B. (2020). Combining US and Canadian forest inventories to assess habitat suitability and migration potential of 25 tree species under climate change. *Diversity and Distributions*, *26*(9), 1142–1159. https://doi.org/10.1111/ddi.13078

Purves, D. W., Lichstein, J. W., & Pacala, S. W. (2007). Crown plasticity and competition for canopy space: A new spatially implicit model parameterized for 250 north American tree species. *PLoS ONE*, *2*(9), e870. https://doi.org/10.1371/journal.pone.0000870

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Romijn, E., Lantican, C. B., Herold, M., Lindquist, E., Ochieng, R., Wijaya, A., Murdiyarso, D., & Verchot, L. (2015). Assessing change in national forest monitoring capacities of 99 tropical countries. *Forest Ecology and Management*, *352*, 109–123. https://doi.org/10.1016/j.foreco.2015.06.003

Roy, D. P., Kashongwe, H. B., & Armston, J. (2021). The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Science of Remote Sensing*, *4*, 100024. https://doi.org/10.1016/j.srs.2021.100024

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.

Särndal, C.-E., Swensson, B., & Wretman, J. (2003). *Model assisted survey sampling*. Springer Science & Business Media.

Shi, T., & Xu, H. (2019). Derivation of tasseled cap transformation coefficients for Sentinel-2 MSI at-sensor reflectance data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(10), 4038–4048. https://doi.org/10.1109/JSTARS.2019.2938388

Silva, C. A., Duncanson, L., Hancock, S., Neuenschwander, A., Thomas, N., Hofton, M., Fatoyinbo, L., Simard, M., Marshak, C. Z., Armston, J., Lutchke, S., & Dubayah, R. (2021). Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. *Remote Sensing of Environment*, *253*(September 2020), 112234. https://doi.org/10.1016/j.rse.2020.112234

Ståhl, G., Saarela, S., Schnell, S., Holm, S., Breidenbach, J., Healey, S. P., Patterson, P. L., Magnussen, S., Næsset, E., McRoberts, R. E., & Gregoire, T. G. (2016). Use of models in large-area forest surveys: Comparing model-assisted, model-based and hybrid estimation. *Forest Ecosystems*, *3*(1), 5. https://doi.org/10.1186/s40663-016-0064-9

Stovall, A. E. L., & Shugart, H. H. (2018). Improved biomass calibration and validation with terrestrial lidar: Implications for future LiDAR and SAR missions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(10), 3527–3537. https://doi.org/10.1109/JSTARS.2018.2803110

Suárez-Muñoz, M., Mina, M., Salazar, P. C., Navarro-Cerrillo, R. M., Quero, J. L., & Bonet-García, F. J. (2021). A step-by-step guide to initialize and calibrate landscape models: A case study in the Mediterranean mountains. *Frontiers in Ecology and Evolution*, *9*, 1–19. https://www.frontiersin.org/articles/10.3389/fevo.2021.653393

Thomas, N., Baltezar, P., Lagomasino, D., Stovall, A., Iqbal, Z., & Fatoyinbo, L. (2021). Trees outside forests are an underestimated resource in a country with low forest cover. *Scientific Reports*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-86944-2

Tinkham, W. T., Mahoney, P. R., Hudak, A. T., Domke, G. M., Falkowski, M. J., Woodall, C. W., & Smith, A. M. S. (2018). Applications of the United States forest inventory and analysis dataset: A review and future directions. *Canadian Journal of Forest Research*, *48*(11), 1251–1268. https://doi.org/10.1139/cjfr-2018-0196

Tomppo, E., Gschwantner, T., Lawrence, M., McRoberts, R. E., Gabler, K., Schadauer, K., Vidal, C., Lanz, A., Ståhl, G., & Cienciala, E. (2010). *National forest inventories: Pathways for common reporting* (Vol. *1*). Springer.

Tomppo, E. O., & Schadauer, K. (2012). Harmonization of national forest inventories in Europe: Advances under cost action E43. *Forest Science*, *58*(3), 191–200. https://doi.org/10.5849/forsci.10-091

Vaglio Laurin, G., Chen, Q., Lindsell, J. A., Coomes, D. A., Del Frate, F., Guerriero, L., Pirotti, F., & Valentini, R. (2014). Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *89*, 49–58. https://doi.org/10.1016/j.isprsjprs.2014.01.001

Wang, W. J., He, H. S., Spetich, M. A., Shifley, S. R., Thompson, F. R., III, Larsen, D. R., Fraser, J. S., & Yang, J. (2013). A large-scale forest landscape model incorporating multi-scale processes and utilizing forest inventory data. *Ecosphere*, *4*(9), art106. https://doi.org/10.1890/ES13-00040.1

Westfall, J. A., Schroeder, T. A., McCollum, J. M., & Patterson, P. L. (2022). A spatial and temporal assessment of nonresponse in the national forest inventory of the U.S. *Environmental Monitoring and Assessment*, *194*(8), 530. https://doi.org/10.1007/s10661-022-10219-0

Westfall, J. A., & Wilson, B. T. (2022). Nonresponse bias in change estimation: A national forest inventory example. *Forestry: An International Journal of Forest Research*, *95*(3), 301–311. https://doi.org/10.1093/forestry/cpab056

White, J. C., Wulder, M. A., Varhola, A., Vastaranta, M., Coops, N. C., Cook, B. D., Pitt, D., & Woods, M. (2013). A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *Forestry Chronicle*, *89*(6), 722–723. https://doi.org/10.5558/tfc2013-132

Will-Wolf, S., Jovan, S., & Amacher, M. C. (2017). Lichen elemental content bioindicators for air quality in upper Midwest, USA: A model for large-scale monitoring. *Ecological Indicators*, *78*, 253–263. https://doi.org/10.1016/j.ecolind.2017.03.017

Wilson, B. T., Woodall, C. W., & Griffith, D. M. (2013). Imputing forest carbon stock estimates from inventory plots to a nationally continuous coverage. *Carbon Balance and Management*, *8*(1), 1. https://doi.org/10.1186/1750-0680-8-1

Winter, S., Böck, A., & McRoberts, R. E. (2012). Estimating tree species diversity across geographic scales. *European Journal of Forest Research*, *131*(2), 441–451. https://doi.org/10.1007/s10342-011-0518-0

Woodall, C. W., Heath, L. S., Domke, G. M., & Nichols, M. C. (2011). *Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the U.S. forest inventory, 2010* (p. 3). USDA Forest Service.

Yang, Q., Su, Y., Hu, T., Jin, S., Liu, X., Niu, C., Liu, Z., Kelly, M., Wei, J., & Guo, Q. (2022). Allometry-based estimation of forest aboveground biomass combining LiDAR canopy height attributes and optical spectral indexes. *Forest Ecosystems*, *9*, 100059. https://doi.org/10.1016/j.fecs.2022.100059

Yu, Y., Saatchi, S., Domke, G. M., Walters, B., Woodall, C., Ganguly, S., Li, S., Kalia, S., Park, T., Nemani, R., Hagen, S. C., & Melendy, L. (2022). Making the US national forest inventory spatially contiguous and temporally consistent. *Environmental Research Letters*, *17*(6), 065002. https://doi.org/10.1088/1748-9326/ac6b47

Zeller, L., Liang, J., & Pretzsch, H. (2018). Tree species richness enhances stand productivity while stand structure can have opposite effects, based on forest inventory data from Germany and The United States of America. *Forest Ecosystems*, *5*(1), 1–17. https://doi.org/10.1186/s40663-017-0127-6

Zhu, K., Woodall, C. W., & Clark, J. S. (2012). Failure to migrate: Lack of tree range expansion in response to climate change. *Global Change Biology*, *18*(3), 1042–1052. https://doi.org/10.1111/j.1365-2486.2011.02571.x

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Map of study area and examples of remotely sensed data. Dots indicate publicly available (i.e., perturbed) FIA plot coordinates.

Figure S2. Correlation between Sentinel-2 metrics and aboveground biomass (TOTAL_AGB). Variables with the highest correlation with TOTAL_AGB (TC2_fa and ndvi_sp) were used to build a linear model for the model-assisted estimator.

Figure S3. Relationship between linear and nonlinear model-assisted (MA) estimators. Nonlinear estimator used the model form $y_i = \beta_1 x_i^{\beta_2} + \epsilon_i$ (Equation 3) where $x_i = Z_{mean}$ for the ith plot, and the linear estimator used the model form $y_i = \sum_{j=1}^{J} \beta_j x_{i,j} + \epsilon_i$ (Equation 5) where $x_{i,1} = Z_{mean}$ and $x_{i,2}$ = Sentinel-2 spring NDVI for the ith plot. Dashed line indicates 1:1 relationship and horizontal and vertical dotted lines indicate null model (no outliers removed) mean aboveground biomass per unit area (AGB).

How to cite this article: Knott, J. A., Liknes, G. C., Giebink, C. L., Oh, S., Domke, G. M., McRoberts, R. E., Quirino, V. F., & Walters, B. F. (2023). Effects of outliers on remote sensing-assisted forest biomass estimation: A case study from the United States national forest inventory. *Methods in Ecology and Evolution*, *14*, 1587–1602. https://doi.org/10.1111/2041-210X.14084