Advanced Statistical Frameworks for Enhancing Stand-Level Forest Inventory: A Synthesis of Methods for Estimation, Updating, and Spatial Integration

Gemini, Deep Research AI, September 22, 2025

Section 1: Foundational Approaches to Stand-Level Volume Estimation

The primary objective of a forest inventory is to describe the quantity and quality of forest resources, providing the foundational data for sound forest management, timber sales, and operational planning.[1] At the stand level—a group of trees sufficiently uniform in composition, age, or condition to be a manageable unit—the central task is to obtain reliable estimates of merchantable wood volume, often disaggregated by tree species and size class.[2] This information is critical for determining the financial value of timber resources and making informed silvicultural decisions.[4] Historically, this process was dominated by direct field measurements, but it has evolved into a sophisticated discipline that integrates statistical sampling theory, advanced modeling, and remote sensing technologies to balance cost, accuracy, and reliability.[1]

1.1 Principles of Stand Characterization and Sampling

The immense scale of forest landscapes makes a 100-percent inventory of every tree both economically prohibitive and, for most management purposes, unnecessary. Complete inventories are not only expensive and tedious but also prone to increased non-sampling errors, such as incorrect data recording, as fatigue sets in.[1] Consequently, forest inventory relies on the principles of statistical sampling, whereby measurements from a representative subset of the forest are used to infer the characteristics of the entire population. The design of this sampling effort is a critical decision influenced by the inventory's objectives, the variability of the forest, and budgetary constraints.[1]

Ground-Based Estimation Methods

The traditional approach to estimating stand volume begins with direct measurements of individual trees within sample plots. The fundamental theory is to establish robust relationships between merchantable volume—a variable that cannot be measured directly on a standing tree without destructive harvesting—and dimensions that are easily and cost-effectively measured, such as diameter at breast height (DBH) and total tree height.[4]

One of the simplest conceptual methods is the **mean tree method**. In this approach, the tree of mean volume within a stand is identified and its volume is carefully measured (or estimated from a detailed function). This single volume is then multiplied by the total number of trees in the stand to derive an estimate of total stand volume. A more refined variant involves calculating the volume per unit of basal area for the mean tree and multiplying this ratio by the total stand basal area (G), a quantity that can be efficiently measured through enumeration or angle-count sampling.[6] To improve precision, this method can be applied independently to different strata within the stand, such as diameter or basal area classes. By summing the volumes of these individual strata, a more accurate total stand volume is obtained.[6]

However, the most common and robust ground-based methods rely on developing explicit mathematical relationships. This is achieved by creating **volume tables or volume equations** through regression analysis. The process involves felling a set of sample trees, carefully measuring their volume, and then using the method of least squares to fit a regression model that predicts volume from easily measured variables.[6] These models can vary in complexity:

- **1-Way Volume Tables:** Predict volume using only DBH.
- **2-Way Volume Tables:** Predict volume using both DBH and height, which is generally more accurate as it accounts for variation in tree form.
- **3-Way or 4-Way Volume Tables:** May incorporate additional variables like a form factor or taper measurement to further refine the estimate.[6]

A powerful stand-specific application of this principle is the **volume line**. This is a linear regression of individual tree volume (v) against its basal area (g), typically of the form $v=a+bg$. Once the regression constants 'a' and 'b' are determined for a particular stand from a subsample of trees, the volume of any tree in that stand can be estimated from its measured basal area. The total stand volume (V) can then be calculated efficiently by summing the individual tree volumes or, more simply, by using the formula $V=N\cdot a+b\cdot G$, where N is the total number of trees and G is the total stand basal area.[6] This approach effectively creates a custom, 1-way volume table for that specific stand at that point in time.

Large-scale inventory programs, such as the USDA's Forest Inventory and Analysis (FIA), operationalize this concept by developing regionally specific volume models for different species and species groups. These models are a core component of comprehensive estimation systems like the Component Ratio Method (CRM), which is used to estimate not just merchantable volume but also other components like biomass in the tree top, limbs, and stump.[7]

Statistical Sampling Designs

The choice of sampling design is crucial for ensuring that the collected data is representative and that the resulting estimates are statistically defensible. While numerous designs exist, a few are foundational to forest inventory.[1]

Systematic sampling is the most frequently used technique in forest inventory. In this design, sampling units (plots) are spaced at fixed intervals throughout the population, often on a grid.[1] Its popularity stems from several advantages: it is easy to implement in the field, its logic is straightforward to explain to field crews and stakeholders, and by spreading the sample evenly across the entire population, it often yields more precise results for the population mean than simple random sampling with the same number of plots.[1] This mechanical and uniform placement of plots generally makes inventories based on systematic sampling faster and less expensive to execute.[1]

Simple random sampling, where every possible combination of sampling units has an equal chance of being selected, is the theoretical basis for much of statistical inference but is less commonly used in practice due to logistical challenges and potential for inefficiency if sample plots cluster by chance.[1]

Stratified random sampling offers a significant improvement in efficiency over simple random sampling. In this method, the forest population is first divided into homogeneous subgroups, or strata, based on some characteristic like forest type, age class, or stand density. Each stratum is then sampled independently, typically with a random or systematic design.[1] This approach has two main advantages. First, it provides separate estimates and measures of precision for each stratum, which can be valuable for management. Second, because the variation within each homogeneous stratum is typically less than the variation across the population as a whole, stratification often yields a more precise estimate for the overall population mean for a given sampling intensity.[1] The delineation of these strata is increasingly accomplished using remotely sensed data, such as aerial photographs or satellite images, which provide a synoptic view of the landscape and allow for efficient classification of forest conditions prior to fieldwork.[1]

1.2 Modeling Stand Structure through Diameter Distributions

While estimates of total stand volume are useful, they provide an incomplete picture for management and economic planning. A stand composed of many small-diameter trees may have the same total volume as a stand with fewer large-diameter, high-value sawtimber trees. To address this, modern forest inventory focuses on characterizing the entire **diameter distribution**—the number of trees per unit area in each diameter class.[8] The diameter of a tree is strongly correlated with other critical variables like height, volume, and biomass, making the diameter distribution a powerful descriptor of stand structure and a direct indicator of the types of wood products that can be harvested.[8]

Probability Density Functions (PDFs)

Estimating the diameter distribution empirically by measuring every tree in a stand is impractical. Instead, statistical models known as probability density functions (PDFs) are used to describe and predict the shape of the distribution from a few key stand variables (e.g., mean diameter, basal area). This modeling approach significantly reduces the cost of forest inventories and improves the predictive power of growth models.[8] Several PDFs have proven effective in forestry due to their flexibility in fitting the wide variety of distribution shapes found in nature, from the bell-shaped curves of even-aged plantations to the reverse-J shapes of uneven-aged, old-growth forests.[10]

The **Weibull function** is arguably the most popular and widely used PDF for modeling diameter distributions. Its introduction to forestry by Bailey and Dell (1973) was a significant advancement, as its mathematical properties make it both highly flexible and straightforward to integrate for calculating the proportion of trees within any given diameter class.[10] It can be specified with two or three parameters: a location parameter (

a) defining the minimum diameter, a scale parameter (b), and a shape parameter (c). The two-parameter version is often sufficient and preferred for its parsimony.[8] Studies consistently show that the Weibull function provides excellent fits across diverse forest types and conditions.[8]

Other commonly used PDFs include:

- **Johnson's SB function:** A highly flexible four-parameter distribution that can model a very wide range of shapes, including bimodal distributions, which can occur in complex, multi-cohort stands.[8] Its parameters relate to location (
  ϵ), scale (λ), and shape (γ, δ), which express skewness and kurtosis.[8]
- **Beta function:** Another flexible distribution defined on a finite interval, making it suitable for diameter data which has a natural minimum and maximum.[8]
- **Gamma, Normal, and Logit-Logistic functions** have also been applied successfully in specific contexts.[8]

Parameter Estimation Methods

The utility of a PDF depends on the ability to accurately estimate its parameters for a given stand at a specific point in time. Two primary methods have been developed for this purpose: parameter prediction and parameter recovery.[8]

The **parameter prediction** approach involves developing regression models that directly predict the PDF parameters (e.g., the Weibull 'b' and 'c' parameters) from stand-level variables like age, site index, or dominant height.[10] While conceptually simple, this method has been found to be less robust than the alternative.

The **parameter recovery** approach, introduced by Hyink and Moser (1983), is now the preferred method.[10] Instead of predicting the abstract PDF parameters directly, this technique first predicts tangible, biologically meaningful stand attributes and then "recovers" the PDF parameters from them. This ensures that the resulting distribution is consistent with the predicted aggregate properties of the stand. The stand attributes used for recovery are typically moments of the diameter distribution (e.g., arithmetic mean diameter, quadratic mean diameter (

dg), variance) or percentiles of the distribution (e.g., the 25th, 50th, or 95th diameter percentiles).[10]

The process for predicting a future diameter distribution using this method involves several steps:

1. **Project Stand Attributes:** Using current inventory data for a stand (e.g., number of trees N0, dominant height Hd0, quadratic mean diameter dq0 at age A0), a system of regression equations is used to predict these attributes at a future age A. For example, future survival (N) might be predicted from initial stocking and the time interval, while future dq might be predicted from initial stand conditions.[10]
2. **Recover PDF Parameters:** The predicted future stand attributes (e.g., predicted future dq and diameter variance) are then used as inputs to a system of equations that is solved to find the PDF parameters that would produce those exact attributes. For the two-parameter Weibull function, for instance, the shape parameter 'c' and scale parameter 'b' can be recovered from the predicted mean diameter and variance.[10]

This approach has been shown to be more accurate and stable than direct parameter prediction. Research evaluating different combinations of moments and percentiles for recovery has found that methods involving the predicted quadratic mean diameter tend to perform best.[10] This focus on accurately modeling the full diameter distribution provides a critical link between ecological inventory and economic valuation, as it allows managers to directly estimate the volume available by specific size classes, which is essential for timber appraisal and harvest planning.[8]

1.3 The Role of Auxiliary Data in Modern Inventories

The evolution of forest inventory has been profoundly shaped by the increasing availability of spatially comprehensive auxiliary data, primarily from remote sensing platforms. This has catalyzed a shift from purely design-based inventory frameworks, which rely solely on field plots for estimates, to model-assisted and model-based frameworks. These modern approaches use the statistical relationship between accurate, but sparse, field measurements and dense, but less direct, remote sensing measurements to produce more precise estimates and spatially explicit maps of forest attributes.[12]

Airborne Laser Scanning (LiDAR)

Airborne Laser Scanning, or LiDAR, has revolutionized forest inventory by providing highly accurate, three-dimensional information about the physical structure of the forest canopy.[14] A LiDAR sensor mounted on an aircraft emits laser pulses and measures the time it takes for them to reflect off objects and return. By differencing the returns from the canopy and the ground, a precise measurement of vegetation height can be obtained for vast areas.[16]

The dominant operational method for integrating LiDAR with field data is the **area-based approach (ABA)**.[16] In this two-stage process:

1. **Model Fitting:** Statistical models are developed that link forest attributes measured on ground plots (e.g., merchantable volume, basal area, biomass) to metrics derived from the LiDAR point cloud within the plot footprint. LiDAR metrics typically include statistics describing the height distribution of laser returns, such as height percentiles (e.g., 95th percentile), mean height, and measures of canopy density.[16]
2. **Prediction and Mapping:** The fitted models are then applied to a wall-to-wall grid of LiDAR metrics covering the entire inventory area. This produces a continuous map of the predicted forest attribute for every pixel (e.g., a 30 m x 30 m grid cell) in the landscape.[16]

When used within a design-based sampling framework, LiDAR data serves as a powerful auxiliary variable. For instance, a **ratio estimator** can be used where the LiDAR-derived canopy height is measured for the whole study area. This wall-to-wall information is then coupled with the standing volume measured on a much smaller sample of ground plots. The resulting estimate for total volume is substantially more precise than an estimate derived from the ground plots alone. In one case study, this LiDAR-assisted method produced a 95% confidence interval for total volume that was approximately two-thirds smaller than that from the ground-only survey, representing a dramatic gain in precision.[15]

Digital Photogrammetry (DP)

While LiDAR is highly accurate, its acquisition can be expensive. An increasingly popular and cost-effective alternative is **Digital Photogrammetry (DP)**, which generates 3D point clouds from overlapping stereo aerial photographs.[18] Using Structure from Motion (SfM) algorithms, modern photogrammetric software can match features across multiple images to reconstruct the 3D structure of the forest canopy, producing a point cloud similar in nature to that from LiDAR.[19]

This technology, especially when deployed on Unmanned Aerial Systems (UAS), offers a flexible and affordable means to acquire high-resolution 3D data.[20] The resulting point clouds can be used in an area-based approach, just like LiDAR data, to estimate key stand parameters such as timber volume, biomass, basal area, and height.[18] While DP point clouds primarily capture the outer canopy envelope and lack the ground-penetrating capability of LiDAR, studies have shown their accuracy to be comparable to LiDAR for many inventory applications, particularly when a high-quality Digital Elevation Model (DEM) is available from a previous LiDAR flight.[18] The higher frequency of aerial photo acquisition compared to LiDAR scans in many regions makes DP a compelling tool for updating inventories more regularly.[18]

Satellite Imagery

Satellite imagery has long been used in forest inventory, primarily as a tool for stratification. By classifying a landscape into broad categories (e.g., coniferous forest, deciduous forest, non-forest) using Landsat or similar satellite data, inventory plots can be allocated more efficiently, reducing the overall variance of estimates for attributes like forest area or total volume.[22]

With advancements in sensor technology and analytical methods, satellite data is now being used in more direct prediction roles. Multi-spectral data from sensors like Landsat, Sentinel-2, and MODIS can be used as predictor variables in both parametric regression models and non-parametric methods to create maps of forest attributes.[23] The unique spectral bands of modern sensors, such as the red-edge bands on Sentinel-2, are particularly sensitive to vegetation characteristics and can improve model performance.[24] Furthermore, very high-resolution satellite imagery (e.g., 30-50 cm from Maxar's WorldView satellites) combined with machine learning algorithms can now be used for tasks previously thought to require aerial data, such as individual tree crown delineation, tree counting for stocking assessments, and even stand health analysis.[25] While optical satellite data can suffer from signal saturation in dense forests and is limited by cloud cover, its broad-area coverage and frequent revisit times make it an indispensable tool for large-area monitoring and estimation.[24]

Section 2: Temporal Updating of Stand Inventories: Modeling Forest Dynamics

Forests are dynamic ecosystems, constantly changing due to growth, mortality, and disturbances. A static forest inventory, representing a single snapshot in time, quickly becomes outdated, diminishing its value for management planning.[18] Consequently, a central challenge in forest science is to accurately update stand-level estimates over time. This involves moving beyond simple remeasurement to the application of mathematical models that describe and project the processes of stand development. The state-of-the-art in this domain has evolved from deterministic growth projections to sophisticated data assimilation frameworks that continuously integrate new observations with model forecasts to maintain a current and precise understanding of the forest's state.

2.1 Projecting Stand Development with Growth and Yield Models

Growth and yield modeling is the discipline of using mathematical equations to predict the future condition of a forest based on its current state and site characteristics.[4] These models are the workhorses of long-term forest management, allowing planners to simulate the outcomes of different silvicultural strategies (e.g., thinning, fertilization, final harvest) and determine the optimal path to achieve desired future conditions.[4]

Modeling Philosophies and Resolution

Growth and yield models can be broadly categorized based on their underlying philosophy and level of detail.[28]

Empirical models are the most common type. They are built by fitting statistical relationships to long-term data from permanent sample plots. These models, such as traditional stand table projection systems, capture the observed patterns of growth in a specific region for specific species but may be unreliable when extrapolating to conditions outside of those represented in the fitting data (e.g., predicting the response to a novel thinning regime).[28] In contrast,

mechanistic models (or process-based models) attempt to simulate the fundamental biological processes of tree growth, such as photosynthesis, respiration, and nutrient cycling. Models like JABOWA, TREGRO, and SORTIE use ecological principles to predict how an individual tree responds to environmental factors like light, water, and temperature.[28] While theoretically more robust for predicting responses to novel conditions like climate change, they are often more complex and data-intensive than empirical models.[29]

Models also differ in their resolution, or the fundamental unit of simulation [26]:

- **Whole-Stand Models:** These are the simplest, predicting aggregate stand-level variables like total basal area, number of trees per acre, or total volume as a function of stand age, site index, and density.[26] They provide a broad overview of stand development but offer no detail on the size distribution of trees.

- **Diameter-Class Models:** These models disaggregate the stand into diameter classes and project the number of trees in each class over time. This provides more detailed information about stand structure and potential products than whole-stand models.[26]
- **Individual-Tree Models:** This is the most detailed and flexible approach. These models simulate the growth and survival of each individual tree within a sample plot, with stand-level dynamics emerging as the aggregate result of these individual processes.[30]

Individual-Tree Models in Practice

Individual-tree models represent the state-of-the-art for detailed stand projection. They typically consist of submodels that predict diameter growth, height growth, and the probability of mortality for each tree, often as a function of the tree's size, species, and the level of competition it faces from its neighbors.[30]

A key distinction is between **distance-dependent** and **distance-independent** models. Distance-dependent models require the spatial coordinates of every tree and calculate a competition index for each tree based on the size and proximity of its neighbors. While more biologically realistic, they are also more data-intensive. Distance-independent models, which are more common in operational settings, do not require tree locations and use a plot-level measure of competition, such as total stand basal area, to modify individual tree growth rates.[30]

The **Forest Vegetation Simulator (FVS)** is a premier example of a distance-independent, individual-tree growth modeling system widely used by the USDA Forest Service and other land managers.[31] FVS is not a single model but a family of geographically specific "variants," each calibrated for the tree species, forest types, and growth patterns of a particular region (e.g., the Inland Empire variant, the Southern variant). This structure allows it to simulate a wide range of silvicultural treatments—from commercial thinning to prescribed burning—for most major forest types across the United States.[31]

Another influential model is the **Prognosis Model for Stand Development**. A key feature of this model is its "self-calibration" process. It uses past growth data from increment cores taken from a specific stand to adjust its internal growth functions. This allows the model to adapt its predictions to the unique local site quality, genetic characteristics, and vigor of the trees in that particular stand, rather than relying solely on regional average growth rates. This makes it a powerful tool for creating a customized "prognosis" of how a specific stand will respond to alternative management prescriptions.[32]

The application of these models represents a significant leap from static inventory. They transform a snapshot of the forest into a trajectory, allowing managers to play out "what-if" scenarios over decades. However, these projections are inherently uncertain; they are forecasts, not facts, and their accuracy degrades over time as small errors in the growth models accumulate and unforeseen events occur.[11] This limitation sets the stage for a more dynamic approach to inventory updating.

2.2 State-of-the-Art: Data Assimilation Frameworks

The advent of frequent, low-cost remote sensing has created an opportunity to move beyond simple projection and toward a system of continuous inventory updating. **Data Assimilation** is the formal statistical framework for achieving this. Originating in fields like meteorology for weather forecasting, data assimilation provides a procedure for dynamically merging information from a forecasting model with new observations to produce an improved estimate of a system's current state.[14] Rather than discarding old inventory data when new data becomes available, data assimilation uses the old data, projected forward in time, as a valuable piece of prior information.[33]

This represents a fundamental paradigm shift. The forest stand is no longer viewed as a static object to be measured periodically, but as a dynamic system whose state trajectory is continuously tracked and corrected. The inventory is not a report, but a live, evolving estimate.

Core Components of a Data Assimilation System

A generic data assimilation system for a forest stand consists of a repeating cycle with four key components [14]:

1. **Initial State Estimate:** The process begins with an estimate of a stand parameter (e.g., growing stock volume) at time t0. Crucially, this is not just a point estimate but is characterized by a probability distribution, including a mean and a measure of its uncertainty (variance).
2. **Forecasting Model:** A growth and yield model is used to project the state estimate forward from time t0 to a future time t1. The model must also propagate the uncertainty, predicting not only the expected volume at t1 but also the variance of that prediction, which will typically increase over the forecast interval.
3. **New Observation:** At time t1, a new measurement of the stand parameter is obtained, typically from a remote sensing source like LiDAR or DP. This new observation also has its own estimate of uncertainty.
4. **Assimilation (or Update) Step:** A statistical algorithm combines the forecasted state (the "prior") with the new observation. This produces a new, updated estimate of the state at

time t1 (the "posterior"). This posterior estimate is more precise (i.e., has a smaller variance) than either the forecast or the new observation alone, as it optimally incorporates the information from both sources.[14] This posterior then becomes the initial state for the next cycle, which repeats when another observation becomes available at time t2.

The central insight of data assimilation is that the combination of the forecast and the new observation is weighted by their respective uncertainties. A very precise forecast will be adjusted only slightly by a noisy new measurement, whereas a very uncertain forecast will be heavily corrected by a precise new measurement.[35] This elevates the role of uncertainty from a simple metric for reporting final precision to a critical, operational input that governs the entire estimation system. This rigorous accounting for uncertainty is what enables more advanced, risk-aware forest planning, such as stochastic optimization, which can explicitly consider the probability of different outcomes.[11]

Statistical Engines for Data Assimilation

Two main classes of statistical methods are used to perform the assimilation step [35]:

The **Kalman Filter** is a standard and computationally efficient algorithm for data assimilation, widely used in many engineering and scientific fields.[37] It assumes that the errors in both the forecasting model and the new observations are normally distributed. In its basic form, it applies to linear systems. Since forest growth is non-linear, a variant called the

Extended Kalman Filter (EKF) is typically used. The EKF linearizes the non-linear growth model at each time step using a Taylor series approximation.[35] The "Kalman Gain" is the weighting factor calculated at each step that determines how much the forecast is adjusted based on the new data; it is a function of the relative uncertainties of the two information sources.[36] While powerful, studies have shown that the EKF can sometimes underestimate the true variance of the estimates.[34]

Bayesian Methods provide a more general and flexible approach to data assimilation. In the Bayesian framework, the true state of the system is conceived of as a probability distribution. The forecast from the growth model provides the "prior distribution" for the state at time t1. The new observation is used to define a "likelihood function." Bayes' Theorem is then used to combine the prior and the likelihood to produce the "posterior distribution," which represents the updated state of knowledge.[14] This posterior distribution then serves as the prior for the next time step. Bayesian methods are advantageous because they are not restricted to assumptions of normality and can handle any form of error distribution.[35] They provide a unified and philosophically coherent manner for updating beliefs in the light of new evidence.[29]

Challenges in Implementation

While data assimilation holds immense promise, its operational implementation in forestry faces several significant challenges [14]:

- **Data Frequency and Quality:** data assimilation is most advantageous when new, low-precision (and thus low-cost) data becomes available at short intervals to correct an accurate growth model.[34] The increasing availability of data from sources like digital photogrammetry and satellites is now making this feasible.[18]
- **Temporally Correlated Errors:** A critical assumption of simple data assimilation algorithms is that the errors of the observations at different time points are independent. However, in practice, errors from remote sensing-based predictions are often strongly and positively correlated over time. For example, a specific sensor is likely to respond similarly to particular forest conditions (e.g., underestimating volume in very dense stands) at every acquisition. This correlation violates the model assumptions and can severely limit the gain in precision from assimilation, as the new data provides less "new" information than assumed. Accounting for these correlated error structures is a major area of ongoing research.[14]
- **Model and Uncertainty Characterization:** Effective data assimilation requires not only a growth model that can make predictions but one that can also accurately quantify the uncertainty of those predictions. Developing such models is a non-trivial task.[35] Similarly, the uncertainty of each remote sensing-based estimate must be well-characterized.
- **Handling Disturbances:** The continuous, gradual change described by a growth model is abruptly broken by events like thinning, clear-cutting, fire, or windthrow. A practical data assimilation system must include a mechanism to detect these major disturbances, at which point the assimilation algorithm must be broken and re-initialized with a new starting state.[14]

Section 3: Improving Local Precision by Leveraging Regional Data

A persistent challenge in forest inventory is the mismatch between the scale of data collection and the scale of management decision-making. Large-scale programs, such as national forest inventories (NFIs), are designed to provide statistically robust estimates of forest resources over broad geographic areas like states or regions. However, operational forest management occurs at the local level—the stand or a small group of stands. At this fine scale, NFI data is often too sparse to yield estimates with acceptable precision, creating a critical information gap.[13]

Small Area Estimation (SAE) is a suite of statistical techniques designed specifically to bridge this scale gap by combining information from the broad-scale survey with auxiliary data to produce reliable estimates for small domains where sample data is limited or even non-existent.[13]

3.1 The Small Area Problem in Forest Inventory

The USDA Forest Service's Forest Inventory and Analysis (FIA) program is a prime example of a large-scale inventory system. It is built on a systematic sample of permanent plots distributed across the entire United States, providing essential information on the nation's forest resources.[40] Due to the sampling intensity, estimates derived from FIA data are most accurate for large geographic units, such as multi-county regions or entire states.[40]

When a forest manager or planner needs an estimate for a much smaller "domain of interest"—such as a single county, a specific ownership type within that county, or an individual forest stand—the number of FIA plots falling within that domain is typically very small.[41] Attempting to calculate an estimate using only the one or two plots that happen to be in the area (a "direct" estimate) will result in an estimate with an extremely large standard error, rendering it too imprecise to be useful for making high-stakes management decisions.[13] In many small areas, there may be no sample plots at all, making direct estimation impossible. This is the "small area problem": the need for precise local information is high, but the availability of direct local data is low.[41] SAE methods were developed to solve this problem by "borrowing strength" from other data sources to augment the limited local information.[13]

This process makes large, publicly funded regional datasets directly relevant and useful for local, operational problems. It takes a strategic-level resource and transforms it into a tactical-level tool. The development and application of SAE in forestry is a direct response to this fundamental mismatch in scales, providing a statistically rigorous bridge between regional data collection and stand-level management.

3.2 Small Area Estimation (SAE) Methodologies

The core principle of SAE is to use a statistical model to link the variable of interest (e.g., stand volume), which is measured on a sparse sample of plots, to auxiliary variables that are available for every unit in the entire population, including the small areas of interest.[13] This model, fit using data from a larger region, allows information from data-rich areas to inform the estimates in data-sparse areas. SAE methods can be categorized into three main types of estimators.[13]

Direct Estimators

A direct estimator uses only the sample observations (y) that fall within the small area domain of interest (d). For example, the Horvitz-Thompson estimator, which weights each observation by the inverse of its inclusion probability, is a common design-based direct estimator.[13] The primary advantage of direct estimators is that they are design-unbiased, meaning they do not rely on model assumptions. Their major disadvantage, which defines the small area problem, is their high variance and instability when the sample size within the domain is small.[13]

Indirect (Synthetic) Estimators

An indirect, or synthetic, estimator relies entirely on the linking model and the auxiliary data (x). It does not use the direct sample observations from within the small area, even if they are available. The process involves fitting a model, such as y=f(x), using data from a large area, and then applying that model to the auxiliary data for the small area to produce a "synthetic" prediction.[13] The stand-level predictions commonly generated in an area-based LiDAR inventory—where a model is fit using plots from across a landscape and then applied to predict the mean value for every pixel within a stand—are functionally equivalent to synthetic estimates.[16] Synthetic estimators have low variance because they are based on a stable model fit with a large amount of data, but they can be biased if the model does not fit the specific conditions of the small area well.[13]

Composite Estimators

Composite estimators represent the most common and powerful SAE approach. They seek to strike an optimal balance between the low bias of a direct estimator and the low variance of an indirect estimator by taking a weighted average of the two.[13] The weight assigned to each component is typically inversely proportional to its variance, so that more weight is given to the more precise component.[13]

The **Fay-Herriot (FH) model** is a foundational and widely used area-level composite estimator.[42] It is called an "area-level" model because it works with aggregated data: a direct estimate for each small area and the mean of the auxiliary variables for each small area. The model links these two sets of data and includes a random area effect that accounts for variability between the small areas that is not explained by the auxiliary variables.[42] The final FH estimate for a given area is a weighted combination of its direct estimate and its model-based synthetic prediction. As the sample size in an area increases, the direct estimate becomes more precise, and the FH model automatically gives it more weight. Conversely, for an area with very few samples, the model gives more weight to the more stable synthetic prediction.[13]

Other advanced composite estimators, such as the **Empirical Best Linear Unbiased Predictor**

(EBLUP), are "unit-level" models that work with individual plot data rather than area aggregates. These models, often formulated as linear mixed-effects models, can provide more efficient estimates if a strong relationship exists at the plot level.[13]

3.3 The Synergy of SAE and Remote Sensing

The practical implementation and success of modern SAE in forestry are inextricably linked to the availability of high-quality, wall-to-wall remote sensing data. The statistical models at the heart of SAE require auxiliary variables that are not only strongly correlated with the forest attributes of interest but are also available for the entire landscape, covering both sampled and unsampled locations.[13] Remote sensing technologies are uniquely capable of providing this critical data layer.[13]

Data from LiDAR, digital photogrammetry, and multispectral or hyperspectral satellite sensors serve as the ideal auxiliary information for SAE models.[13] For example, LiDAR-derived metrics like canopy height and cover can be used in an FH or EBLUP model to predict stand volume or biomass. The remote sensing data provides the spatial fabric that connects the sparsely located field plots, allowing the model to "interpolate" or predict forest attributes in the spaces between plots with statistical rigor.[46]

This synergy fundamentally increases the economic value and utility of both NFI field data and remote sensing acquisitions. By themselves, NFI plots are often too sparse for local management, and uncalibrated remote sensing data lacks ground-truthed accuracy. When fused through an SAE framework, they create a high-resolution, spatially explicit, and statistically defensible information product that is far more valuable than the sum of its parts.[11] This powerful combination allows forest managers to leverage the investment in national-scale inventory programs to support local, operational decisions, effectively democratizing access to high-quality forest information.

Section 4: Integrated Statistical Paradigms for Modern Forest Inventory

The increasing complexity and richness of forest inventory data—characterized by hierarchical structures, spatial and temporal correlations, and the fusion of multiple data sources—demand statistical paradigms that are more sophisticated than standard regression techniques. Modern forest biometrics relies on integrated frameworks that can properly account for these data structures to produce accurate and defensible estimates. Mixed-effects models provide the tools to handle hierarchical data, the Hierarchical Bayesian paradigm offers a unifying framework for integrating information across scales, and non-parametric methods provide flexible, data-driven alternatives for mapping and imputation.

4.1 Accounting for Data Structure with Mixed-Effects Models

orest inventory data are rarely simple, independent observations. Instead, they possess a nested or hierarchical structure: multiple trees are measured within a single plot, multiple plots may be located within a stand, and the same plot is often remeasured over time (longitudinal data).[38] Observations within the same group (e.g., trees on the same plot) tend to be more similar to each other than to observations from different groups, violating the critical assumption of independence that underlies standard regression models.[38]

Mixed-effects models are the standard statistical approach for analyzing such structured data.[38] They address this challenge by partitioning the variation in the data into two components [47]:

- **Fixed Effects:** These are the conventional regression parameters that represent population-level average relationships. For example, in a height-diameter model, the fixed effects would describe the average relationship between a tree's height and its diameter for all trees in the population.
- **Random Effects:** These parameters capture the idiosyncratic deviations of individual subjects or groups (e.g., plots) from the population average. For instance, a random effect for a plot could account for the fact that all trees on that particular plot are, on average, taller for their diameter than the population mean, perhaps due to superior site quality.

By explicitly modeling this group-level variation, mixed-effects models produce more accurate and efficient estimates of the fixed-effect parameters and provide a mechanism for making subject-specific predictions.[43] In forestry, they are widely used to develop generalized height-diameter models that can be calibrated to local stand conditions, to model longitudinal data in growth and yield studies, and to build the unit-level models used in Small Area Estimation.[38] The inclusion of random effects correctly accounts for the sources of variation in the data, preventing the underestimation of parameter variance that can occur when these correlations are ignored.[43]

4.2 The Power of Hierarchical Bayesian (HB) Modeling

The Hierarchical Bayesian (HB) paradigm offers a comprehensive and philosophically coherent framework that can naturally integrate the solutions to both the temporal updating problem (Data Assimilation) and the spatial scaling problem (Small Area Estimation). It extends the principles of Bayesian inference to models with the kind of nested data structures common in forestry.[46]

The Bayesian Philosophy

In contrast to the classical (frequentist) approach where model parameters are considered fixed, unknown constants, Bayesian inference treats parameters as random variables that have probability distributions.[29] The process begins by specifying a

prior probability distribution for each parameter, which quantifies existing knowledge or belief about its value before observing the current data. This prior information is then formally combined with the information contained in the data, which is expressed through a **likelihood function**, using Bayes' Theorem. The result is a **posterior probability distribution**, which represents the updated state of knowledge about the parameter after accounting for the new data.[29] This posterior distribution provides a complete characterization of uncertainty, from which not only a point estimate (like the mean or median) but also credible intervals (the Bayesian equivalent of confidence intervals) can be derived.[29]

Integrated Applications in Forest Inventory

The hierarchical structure of HB models makes them perfectly suited for forestry applications. A model can be built with multiple levels that mirror the data's structure—for example, a level for individual trees, nested within a level for plots, which are in turn nested within a level for stands or ecological regions.[46] This allows the model to "share strength" or "borrow information" across the levels in a natural, probabilistic way.

This framework provides a powerful and unified approach to the key challenges in modern inventory:

- **Hierarchical Bayesian SAE:** The Bayesian framework is a natural fit for Small Area Estimation. Area-level models like the Fay-Herriot and unit-level mixed models have direct Bayesian analogues.[46] The HB approach offers several advantages, including the ability to incorporate prior information about the degree of similarity between areas, more flexible modeling of the random effects, and the production of a full posterior distribution for the estimate in each small area, providing a richer quantification of uncertainty than a simple point estimate and standard error.[45]
- **Bayesian Data Assimilation:** As noted previously, Bayesian methods are one of the two primary engines for Data Assimilation. The sequential nature of Bayesian updating—where the posterior from one time step becomes the prior for the next—is the very definition of a data assimilation system. This provides a unified way to handle estimation and prediction over time.[14]
- **Bayesian Model Calibration:** For complex, mechanistic growth models with many

parameters, Bayesian calibration is an extremely powerful tool. It can estimate the full joint posterior distribution of all model parameters simultaneously, revealing correlations and uncertainties that are difficult to assess with classical methods. This allows for a robust propagation of parameter uncertainty through the model to quantify the predictive uncertainty of its outputs.[29]

The HB paradigm is not merely another tool in the toolbox; it is emerging as a comprehensive statistical "language" that can express and solve the interconnected spatial and temporal problems at the heart of modern forest inventory within a single, coherent probabilistic framework.

4.3 The Application of Non-Parametric and Imputation Techniques

Alongside the sophisticated parametric models described above, a suite of flexible, data-driven non-parametric methods plays a crucial role in operational forest inventory, particularly for creating spatially explicit maps and filling in missing data. These methods make fewer assumptions about the underlying statistical distributions and functional forms of relationships in the data.

k-Nearest Neighbors (k-NN) Imputation

The **k-Nearest Neighbors (k-NN)** method is one of the most widely used non-parametric techniques in forestry.[52] It is an imputation method, meaning its primary purpose is to fill in missing values. In a typical forest inventory application, the goal is to predict a suite of forest attributes (Y-variables, e.g., volume, biomass, basal area), which are measured on field plots, for every pixel in a landscape where they are unobserved.[54]

The k-NN process works as follows [52]:

1. **Feature Space:** A "feature space" is defined by a set of auxiliary variables (X-variables) that are available for both the field plots (the "reference" set) and all the pixels in the landscape (the "target" set). These X-variables are typically derived from remote sensing data, such as the spectral bands of a satellite image or topographic variables from a DEM.
2. **Distance Calculation:** For a given target pixel, the algorithm calculates the "distance" in the multi-dimensional feature space between that pixel and every reference plot. This distance measures the similarity of the target and reference units based on their auxiliary variables. Various distance metrics can be used, from simple Euclidean distance to more complex metrics derived from machine learning algorithms like Random Forest.[52]
3. **Imputation:** The algorithm identifies the k reference plots that are closest (most similar) to

the target pixel. The predicted values for the Y-variables at the target pixel are then calculated as the (possibly weighted) average of the values from those *k* nearest neighbors.

The primary application of k-NN is to create wall-to-wall maps of multiple forest attributes simultaneously.[52] A key strength of the method is that, because it imputes the full vector of attributes from the donor plots, it tends to preserve the natural variance and covariance structure of the original data better than methods that predict each variable with a separate regression model.[54]

The choice between a parametric, model-based approach like SAE and a non-parametric approach like k-NN involves a critical trade-off. Model-based methods provide a rigorous inferential framework, allowing for the calculation of statistically defensible estimates of uncertainty (e.g., the variance of the estimate for a specific county).[42] Non-parametric methods like k-NN are highly flexible and often easier to implement for generating predictive maps, but they struggle to produce reliable, statistically sound variance estimates for their predictions.[54] The choice of method therefore depends heavily on the end goal: if the objective is a visually appealing map for general planning, k-NN may be sufficient; if the objective is a legally or financially defensible estimate with a known level of precision for a specific area, a model-based SAE approach is superior.

Section 5: Synthesis and Future Directions

The science of forest inventory has undergone a profound transformation, evolving from a discipline reliant on manual field measurements and simple stand averages into a highly quantitative field that integrates data across multiple spatial and temporal scales using sophisticated statistical frameworks. The modern stand-level inventory system is not a single method but a cohesive, multi-faceted process designed to provide accurate, up-to-date, and spatially explicit information to support sustainable forest management in an increasingly complex world. This synthesis has detailed the foundational principles, the methods for temporal updating, the techniques for improving local precision, and the integrated statistical paradigms that underpin the state-of-the-art.

5.1 An Integrated View of the Modern Inventory System

A comprehensive, modern system for stand-level forest inventory can be conceptualized as a workflow that seamlessly integrates the key methodologies discussed throughout this report. This integrated system addresses the core challenges of cost, precision, temporal relevance, and spatial scale.

The process begins with **Data Fusion and Initial State Characterization**. The foundation is still a network of high-quality field plots, but these are no longer the sole source of information. They serve as the "ground truth" for calibrating and validating models that leverage wall-to-wall auxiliary data from remote sensing platforms like LiDAR and Digital Photogrammetry. An area-based approach is used to build robust predictive models linking field-measured attributes (e.g., volume, biomass) to 3D structural metrics from the remote sensing data. To provide detailed information for product assortment and economic valuation, the initial stand state is often characterized not just by total volume but by its full diameter distribution, modeled using flexible probability density functions like the Weibull.

Next, the system addresses the **Spatial Scaling and Localization** problem. The models developed in the first stage are often built using data from a large regional survey, such as the FIA. To produce precise estimates for specific, small management units (stands or counties), Small Area Estimation techniques are employed. Using the remote sensing data as the critical linking auxiliary information, SAE models "borrow strength" from the regional dataset to refine the local estimates, effectively bridging the gap between the scale of data collection and the scale of management. This step transforms a strategic, regional inventory into an operational, local one.

Simultaneously, the system must maintain **Temporal Relevance** through dynamic updating. The static, localized inventory is projected forward in time using individual-tree or stand-level growth and yield models like FVS. This provides a forecast of future conditions. However, to avoid the accumulation of model error, this forecast is continuously corrected within a Data Assimilation framework. As new remote sensing data becomes available at regular intervals (e.g., every 1-3 years), it is assimilated with the model forecast to produce a new, more precise estimate of the current stand state.

The entire integrated system is held together by an **Overarching Statistical Paradigm**. Increasingly, Hierarchical Bayesian modeling is recognized as the ideal framework for this task. Its structure naturally accommodates the nested nature of forest data (trees in plots in stands), and its probabilistic approach provides a coherent way to implement both SAE (borrowing strength across space) and data assimilation (updating beliefs over time), all while providing a comprehensive quantification of uncertainty at every stage of the process. The table below summarizes the distinct roles and characteristics of the three most advanced frameworks that form the core of this modern system.

Table 1: Comparison of Major Stand Updating and Estimation Frameworks.

Feature	Data Assimilation	Small Area Estimation (SAE)	Hierarchical Bayesian (HB)

			Modeling
Primary Goal	To provide the most precise estimate of a system's **current state** by sequentially integrating new observations with model forecasts over **time**.	To provide precise estimates for small **spatial domains** (areas) where direct sampling is insufficient, by "borrowing strength" from other areas and auxiliary data.	To provide a flexible and coherent probabilistic framework for inference that can unify complex spatial, temporal, and hierarchical data structures while robustly quantifying uncertainty.
Core Problem Addressed	Temporal data fusion; keeping estimates up-to-date.	Spatial scaling; improving local precision from regional data.	Integrating complex data structures; comprehensive uncertainty accounting.
Key Data Requirements	A time-series of observations (e.g., remote sensing), a dynamic forecasting model (e.g., growth model), and estimates of uncertainty for both.	A sample survey covering a large area (e.g., FIA), wall-to-wall auxiliary data correlated with the variable of interest (e.g., LiDAR), and a linking model.	Data with hierarchical structure (spatial or temporal), prior information on parameters, and a fully specified probabilistic model (likelihood and priors).
Handling of Uncertainty	Uncertainty is an operational input; the relative variance of the forecast and	Primarily focused on reducing the Mean Squared Error (MSE) of the final	Provides a full posterior probability distribution for every parameter and

	new observation determines the update (Kalman Gain).	estimate by optimally balancing the bias of an indirect estimator with the variance of a direct one.	prediction, offering the most complete characterization of uncertainty.
Key Strengths	Dynamically updates inventories; improves precision over time; formally combines models and data.	Makes large-scale inventories relevant at the local/management scale; significantly improves precision in data-sparse areas.	Unifying framework for SAE and data assimilation; flexible model specification; formal incorporation of prior knowledge; robust uncertainty propagation.
Major Limitations	Sensitive to model misspecification and correlated errors in observation time-series; requires good estimates of uncertainty for all components.	Relies on the strength of the linking model; can be biased if the relationship between auxiliary and response variables is weak or inconsistent across domains.	Can be computationally intensive (MCMC); specification of prior distributions can be challenging and influential.

5.2 Emerging Research and Operational Challenges

While the statistical frameworks described represent a mature and powerful suite of tools, the field of forest inventory continues to advance, driven by new technologies, evolving management needs, and persistent scientific challenges. Several key areas represent the frontiers of research and the primary hurdles to full operational implementation.

- **Advanced Data Fusion:** The proliferation of new remote sensing platforms offers unprecedented opportunities but also new challenges. Future inventory systems will need to fuse data from an even more diverse array of sensors, including spaceborne LiDAR (e.g., NASA's GEDI), Synthetic Aperture Radar (SAR), which can penetrate clouds, and hyperspectral sensors that provide detailed information on species and forest health. Developing statistical models that can optimally integrate these multi-modal data sources is a key research priority.

- **Dynamic Models for a Changing World:** Forest growth and yield models are the engines of temporal projection. There is a critical need to improve these models to better account for the effects of climate change, which may alter growth rates and species suitability, and to more accurately simulate the impacts of both natural disturbances (fire, insects, disease) and novel silvicultural practices.[11] The value of information provided by an inventory system is directly tied to the quality of its predictive models; inaccurate forecasts can lead to sub-optimal decisions and significant economic losses.[11]
- **Computational and Big Data Hurdles:** The methods described, particularly Hierarchical Bayesian models fit with Markov Chain Monte Carlo (MCMC) simulation and the processing of massive, wall-to-wall remote sensing datasets, are computationally intensive.[29] Developing more efficient algorithms and leveraging high-performance computing infrastructure are essential for making these advanced techniques operationally feasible for routine, large-scale inventory updates.
- **Comprehensive Uncertainty Propagation:** A recurring theme is the central importance of uncertainty. While individual methods provide estimates of uncertainty for their specific outputs, a major ongoing challenge is to fully propagate uncertainty through the entire, multi-stage modeling chain. This includes accounting for measurement error in field plots, positioning errors, sampling error, error in remote sensing data, and error from every statistical model used in the process. Providing decision-makers with a final estimate that reflects this total, cascaded uncertainty is the ultimate goal for enabling truly risk-informed forest management.[11]
- **Operationalization and Knowledge Transfer:** Finally, there remains a gap between the cutting-edge methods developed in the research community and the tools used by many on-the-ground forest managers. Bridging this gap requires the development of user-friendly software, accessible training materials, and clear demonstrations of the value proposition—in terms of improved decision-making and economic outcomes—of adopting these more complex but powerful statistical frameworks.[11]

Works cited

1. Sampling methods - Assessing the status of logged-over production forests : The development of a rapid appraisal technique, accessed September 22, 2025, https://www.fao.org/4/ac838e/AC838E03.htm
2. 2.3. Stand Dynamics: Stand Structure - SFA Silviculture, accessed September 22, 2025, https://www.sfasilviculture.com/index.php/textbook/2-3-stand-dynamics-stand-structure
3. Dynamic spatial regression models for space-varying forest stand tables, accessed

September 22, 2025, https://experts.umn.edu/en/publications/dynamic-spatial-regression-models-for-space-varying-forest-stand-
4. Forest Growth and Yield | Mississippi State University Extension Service, accessed September 22, 2025, https://extension.msstate.edu/publications/forest-growth-and-yield
5. Forest Inventory & Management_Manual.pdf, accessed September 22, 2025, https://www.forestcarbonpartnership.org/sites/fcp/files/fcp-docs/2015/October/Forest%20Inventory%20%26%20Management_Manual.pdf
6. Stand Volume - The Australian National University, accessed September 22, 2025, https://fennerschool-associated.anu.edu.au/mensuration/BrackandWood1998/S_VOLUME.HTM
7. Estimation of Merchantable Bole Volume and Biomass above Saw log Top in the National Forest Inventory of the United States - ResearchGate, accessed September 22, 2025, https://www.researchgate.net/publication/259819530_Estimation_of_Merchantable_Bole_Volume_and_Biomass_above_Saw_log_Top_in_the_National_Forest_Inventory_of_the_United_States
8. Modelling diameter distribution of natural forests in Pueblo Nuevo ..., accessed September 22, 2025, https://www.scielo.org.mx/scielo.php?pid=S2007-11322022000500075&script=sci_arttext&tlng=en
9. Modeling the Diameter Distribution of Mixed Uneven-Aged Stands in the South Western Carpathians in Romania - MDPI, accessed September 22, 2025, https://www.mdpi.com/1999-4907/12/7/958
10. USE OF THE WEIBULL FUNCTION TO PREDICT FUTURE ..., accessed September 22, 2025, https://www.srs.fs.usda.gov/pubs/gtr/gtr_srs156/gtr_srs156_053.pdf
11. Assessing the importance of detailed forest inventory information using stochastic programming - Canadian Science Publishing, accessed September 22, 2025, https://cdnsciencepub.com/doi/10.1139/cjfr-2023-0218
12. Dynamic spatial regression models for space-varying forest stand ..., accessed September 22, 2025, https://www.researchgate.net/publication/267760021_Dynamic_spatial_regression_models_for_space-varying_forest_stand_tables
13. Review and Synthesis of Estimation Strategies to Meet ... - Frontiers, accessed September 22, 2025, https://www.frontiersin.org/journals/forests-and-global-change/articles/10.3389/ffgc.2022.813569/full
14. Data assimilation in stand-level forest inventories, accessed September 22, 2025, https://cdnsciencepub.com/doi/10.1139/cjfr-2013-0250
15. Area-based lidar-assisted estimation of forest standing volume, accessed September 22, 2025, https://cdnsciencepub.com/doi/abs/10.1139/X08-122
16. Stand validation of lidar forest inventory modeling for a managed southern pine forest, accessed September 22, 2025, https://www.fs.usda.gov/pnw/pubs/journals/pnw_2023_strunk001.pdf
17. Stand validation of lidar forest inventory modeling for a managed southern pine forest, accessed September 22, 2025, https://cdnsciencepub.com/doi/10.1139/cjfr-2022-0032
18. Full article: Updating of forest stand data by using recent digital photogrammetry in combination with older airborne laser scanning data - Taylor & Francis Online, accessed September 22, 2025, https://www.tandfonline.com/doi/full/10.1080/02827581.2021.1936153

19. Photogrammetry and LiDAR for Forest Industry - GeoAI, accessed September 22, 2025, https://geoai.au/photogrammetry-and-lidar-for-forest-industry/
20. Forest Stand Inventory Based on Combined Aerial and Terrestrial Close-Range Photogrammetry - MDPI, accessed September 22, 2025, https://www.mdpi.com/1999-4907/7/8/165
21. Effective UAV Photogrammetry for Forest Management: New Insights on Side Overlap and Flight Parameters - MDPI, accessed September 22, 2025, https://www.mdpi.com/1999-4907/15/12/2135
22. Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique - Northern Research Station, accessed September 22, 2025, https://www.nrs.fs.usda.gov/pubs/jrnl/2002/nc_2002_mcroberts_007.pdf
23. Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data - ResearchGate, accessed September 22, 2025, https://www.researchgate.net/publication/222669813_Estimation_of_forest_stand_volumes_by_Landsat_TM_imagery_and_stand-level_field-inventory_data
24. Regional Forest Volume Estimation by Expanding LiDAR Samples Using Multi-Sensor Satellite Data - MDPI, accessed September 22, 2025, https://www.mdpi.com/2072-4292/12/3/360
25. Creating Forest Inventory from High-Resolution Satellite Images, accessed September 22, 2025, https://blog.maxar.com/earth-intelligence/2018/creating-forest-inventory-from-high-resolution-satellite-images
26. 9.1: Growth and Yield Models - Statistics LibreTexts, accessed September 22, 2025, https://stats.libretexts.org/Bookshelves/Applied_Statistics/Natural_Resources_Biometrics_(Kiernan)/09%3A_Modeling_Growth_Yield_and_Site_Index/9.01%3A_Growth_and_Yield_Models
27. Modelling the stand dynamics after a thinning induced partial mortality: A compensatory growth perspective - Frontiers, accessed September 22, 2025, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2022.1044637/full
28. Modelling in Forest Management, accessed September 22, 2025, https://www.fs.usda.gov/ne/burlington/research/ne4454/pdf_pubs/forest_mgmt_modelling.pdf
29. (PDF) Bayesian calibration of process-based forest models: Bridging ..., accessed September 22, 2025, https://www.researchgate.net/publication/7868457_Bayesian_calibration_of_process-based_forest_models_Bridging_the_gap_between_models_and_data
30. A Review of Methods for Updating Forest Monitoring System Estimates, accessed September 22, 2025, https://www.nrs.fs.usda.gov/pubs/gtr/gtr_nc212/gtr_nc212_494.pdf
31. USDAForestService/ForestVegetationSimulator: Forest ... - GitHub, accessed September 22, 2025, https://github.com/USDAForestService/ForestVegetationSimulator
32. Prognosis Model for Stand Development - USDA Forest Service, accessed September 22, 2025, https://www.fs.usda.gov/rm/pubs_int/int_rp137.pdf
33. Potential of using data assimilation to support forest planning, accessed September 22, 2025, https://cdnsciencepub.com/doi/10.1139/cjfr-2016-0439
34. Data assimilation in stand-level forest inventories | Request PDF - ResearchGate, accessed September 22, 2025, https://www.researchgate.net/publication/259225487_Data_assimilation_in_stand-

level_forest_inventories
35. Data Assimilation in Forest Inventory: First Empirical Results - MDPI, accessed September 22, 2025, https://www.mdpi.com/1999-4907/6/12/4384
36. (PDF) Data Assimilation in Forest Inventory: First Empirical Results - ResearchGate, accessed September 22, 2025, https://www.researchgate.net/publication/290441108_Data_Assimilation_in_Forest_Inventory_First_Empirical_Results
37. Importance of Calibration for Improving the Efficiency of Data Assimilation for Predicting Forest Characteristics - MDPI, accessed September 22, 2025, https://www.mdpi.com/2072-4292/14/18/4627
38. Bayesian approach for modelling non-linear longitudinal/hierarchical data with random effects in forestry - Oxford Academic, accessed September 22, 2025, https://academic.oup.com/forestry/article/85/1/17/642941
39. Data assimilation in forest inventories at stand level - SLU, accessed September 22, 2025, https://pub.epsilon.slu.se/id/document/14015916
40. Small Area Estimation Techniques | NCASI, accessed September 22, 2025, https://www.ncasi.org/technical-studies/forestry/sustainability-and-fiber-supply/small-area-estimation-techniques/
41. Small Area Estimation in Forest Inventories: New Needs, Methods, and Tools - Frontiers, accessed September 22, 2025, https://www.frontiersin.org/research-topics/16045/small-area-estimation-in-forest-inventories-new-needs-methods-and-tools/magazine
42. new approach to small area estimation: improving forest management unit estimates with advanced preprocessing in a multivariate Fay–Herriot model | Forestry - Oxford Academic, accessed September 22, 2025, https://academic.oup.com/forestry/article/98/4/605/7919161
43. Hybrid estimation based on mixed-effects models in forest inventories, accessed September 22, 2025, https://cdnsciencepub.com/doi/10.1139/cjfr-2016-0298
44. Design-based properties of some small-area estimators in forest inventory with two-phase sampling - Canadian Science Publishing, accessed September 22, 2025, https://cdnsciencepub.com/doi/abs/10.1139/cjfr-2012-0381
45. Hierarchical Bayesian models for small area estimation of county-level private forest landowner population - Canadian Science Publishing, accessed September 22, 2025, https://cdnsciencepub.com/doi/abs/10.1139/cjfr-2017-0154
46. Hierarchical Bayesian Small Area Estimation Using ... - Frontiers, accessed September 22, 2025, https://www.frontiersin.org/journals/forests-and-global-change/articles/10.3389/ffgc.2021.752911/full
47. Hybrid estimation based on mixed-effects models in forest inventories, accessed September 22, 2025, https://cdnsciencepub.com/doi/abs/10.1139/cjfr-2016-0298
48. Incorporating stand parameters in nonlinear height-diameter mixed-effects model for uneven-aged Larix gmelinii forests - Frontiers, accessed September 22, 2025, https://www.frontiersin.org/journals/forests-and-global-change/articles/10.3389/ffgc.2024.1491648/full
49. Bayesian Hierarchical Models can Infer Interpretable Predictions of Leaf Area Index From Heterogeneous Datasets - Frontiers, accessed September 22, 2025, https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2021.780814/pdf

50. Estimating Tree Height-Diameter Models with the Bayesian Method - PMC, accessed September 22, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC3953559/
51. Small area estimation of forest parameters | Statistical Ecology and Forest Science Lab, accessed September 22, 2025, https://doserlab.com/project/sae/
52. A Comparison of Imputation Approaches for Estimating Forest Biomass Using Landsat Time-Series and Inventory Data - MDPI, accessed September 22, 2025, https://www.mdpi.com/2072-4292/10/11/1825
53. Applying an Efficient k -Nearest Neighbor Search to Forest Attribute Imputation, accessed September 22, 2025, https://www.researchgate.net/publication/43287975_Applying_an_Efficient_k-Nearest_Neighbor_Search_to_Forest_Attribute_Imputation
54. The roles of nearest neighbor methods in imputing missing data in ..., accessed September 22, 2025, https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1216&context=usdafsfacpub
55. k-nearest neighbor imputation of forest inventory variables in new hampshire - Northern Research Station, accessed September 22, 2025, https://www.nrs.fs.usda.gov/pubs/jrnl/2005/ne_2005_lister_001.pdf
56. Evaluating k-Nearest Neighbor (kNN) Imputation Models for Species-Level Aboveground Forest Biomass Mapping in Northeast China - MDPI, accessed September 22, 2025, https://www.mdpi.com/2072-4292/11/17/2005
57. Evaluating k-nearest neighbor (kNN) imputation models for species-level aboveground forest biomass mapping in northeast China - USGS.gov, accessed September 22, 2025, https://www.usgs.gov/publications/evaluating-k-nearest-neighbor-knn-imputation-models-species-level-aboveground-forest