Modeling

T M Barrett and J S Fried, USDA Forest Service PNW Research Station, Forest Inventory and Analysis, Portland, OR, USA

Published by Elsevier Ltd., 2004

Description and Terminology

Modeling is used extensively in forest inventory applications. Models, which are described as deliberate 'abstractions of a system,' are used to calculate forest attributes that cannot be easily measured, to understand how forest ecosystems function, to extrapolate forest attributes over space, and to project how forests change over time.

Until the early twentieth century, models used in forest inventory consisted of hand-drawn graphs and tables. Over the course of the past century, models were increasingly expressed in mathematical terms and modelers used statistical methods in model development and evaluation. The rapid development of computers in the past few decades has facilitated increasing sophistication of modeling techniques and broadened the scope of model application.

Modeling terminology in forestry follows that of other fields where modeling is applied. Models can be deterministic, where a single outcome is predicted, or stochastic, where random elements are incorporated. Statistical models are created with a development data set, but evaluated with an independent validation data set. Modelers sometimes distinguish between validation, a formal test against the independent data set, and verification, or the internal consistency of the model. Models are often tested with sensitivity analysis, the evaluation of model performance when input parameters are varied. Models are adjusted to local conditions through the use of calibration. Model complexity is influenced by model scope, the area of interest, and model resolution, the amount of detail in the model.

Modelers in forestry have liberally borrowed ideas and terminology from other fields, with the result of occasionally inconsistent use of terminology. For example, scale in geography represents the ratio of map units to real-world units, with the unfortunate result that small-scale is used to refer to spatial models (maps) of large land areas. In other fields, such as ecology, it is common to use small-scale to refer to things that are small in scope or area. In forestry literature, examples of both uses of the term may be found, and the meaning must be inferred from the context.

Some modeling terminology in forestry has developed coincident with particular classes of models that are specific to the field. Growth and yield models are one such example. Growth and yield models, developed within the field of forestry, are sometimes contrasted with process models, developed within the field of ecology. Both may model tree or stand growth, but are characterized by some general differences. Growth and yield models tend to be empirically based, use statistical techniques to relate variables that may not have a direct causal relationship, and are intended to aid prediction. Process models tend to be conceptually based, often use a large number of variables that represent the ecological components of a system, and are intended to increase understanding. In practice, most models have a mix of these characteristics, and the classes are imperfectly distinguished.

Growth models are often combined with forest planning models to project how inventory will change over time. Such models can have very substantial impacts on forest policy, even though they have at times proved inaccurate (**Figure 1**). Inventory projections such as those shown in **Figure 1** have been used in debates about how much forestland should be maintained in reserves, to influence policies of taxation and trade, and to support various types of forest regulation. It is likely that projections of forest inventory will play a role in international debate about carbon storage and global climate change.

Models, in short, play a variety of roles in forest inventory, from the relatively simple function of replacing or augmenting field measurements, to sophisticated policy models with far-reaching implications. The following sections illustrate the great variety of methods and applications.

Modeling Tree and Forest Attributes

For Efficiency in Forest Inventory

In some cases, field measurements of a tree or forest attribute are replaced by modeled values to increase efficiency of the inventory. Individual tree height, the length from the base of the tree trunk to the tip of the apical meristem (tree top), is an example of this type of model use. Tree heights can be measured in the field with a variety of techniques, including height poles, clinometers, and laser instruments. Tree height measurements are important values for calculating timber volume, biomass, or forest structural characteristics. However, tree height measurements take more time than measurements of many other individual tree attributes, such as tree diameter or

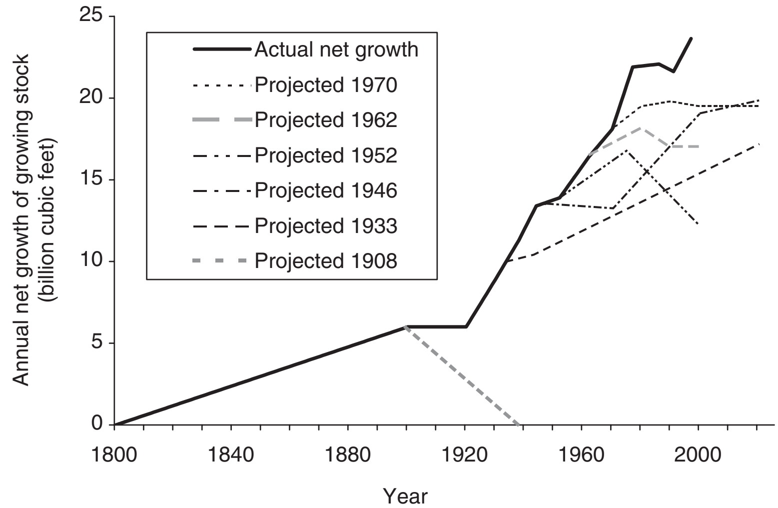


Figure 1 Annual net growth of timber in the USA, 1800–2000, and projections of future growth made in 1908, 1933, 1946, 1952, 1962, and 1970. Reproduced with permission from Clawson M (1979) Forests in the long sweep of American history. *Science* 204: 1168–1174, fig. 4. Copyright 1979 American Association for the Advancement of Science.

species. For this reason, some types of forest inventories measure heights for a subsample of trees and model heights for the rest.

Early height models for a geographic region used equations that expressed individual tree height as a simple polynomial function of tree trunk diameter; model forms have changed over time, but it is still common to use regression techniques to relate height to diameter (**Figure 2**). Although additional variables may be used, such as density or site quality attributes, the purpose of the model is prediction rather than an understanding of why different trees have different heights.

Site quality measurements are also sometimes based on models of tree heights. 'Site' is the local environment where a tree grows; 'site quality' refers to the potential productivity of this area. Site quality is most frequently measured with site indices, which are models of dominant tree height in relation to tree age. Other methods of measuring site quality include use of indicator plants, recent height growth, or physical attributes such as slope, aspect, and soil properties.

Crown cover is another inventory attribute that is commonly modeled. Crown cover is the percentage of the ground that is covered by the vertical projection of tree crowns or leaves. In recent decades, measurements of crown cover have been used in assessments of wildlife habitat and fire hazard. Similarly to height measurements, a variety of specialized methods have been developed to measure crown cover, including instruments such as densiometers or canopy cameras, and techniques such as interpretation of aerial photographs. A strong allometric relationship exists between tree crowns and tree trunk diameters; in the early days of aerial photography, this relationship was used to model tree diameters from crown widths, so that estimates of timber volume could be made. As interest in crown cover itself became important; the same allometric relationship was used to model crown widths from tree diameters. For continuous coverage of large land areas, crown cover information is often developed from models using a combination of remote sensing data and field inventories.

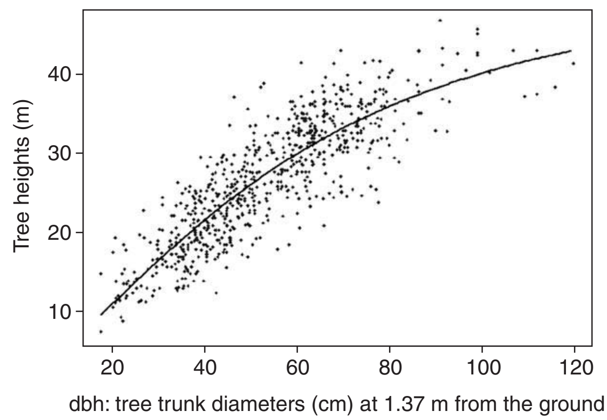


Figure 2 Nonlinear model (black line) of tree height as a function of diameter for *Pinus ponderosa*. Model is Richard's form: $\text{height} = 1.37 + 47.11\ (1 - \exp(-0.021\ \text{diameter}))\text{^}1.50$. Data are from Forest Inventory Analysis program, west coast of USA.

When modeled values are used to replace inventory measurements, as may occur with tree heights or crown cover, both bias and imprecision can contribute to errors. Although modeled values may be less accurate than measured values, the decreased costs of data collection make the technique useful for some types of inventory where efficiency is particularly important. Knowledge of the accuracy of the models is critical to correct use. For example, the imprecision of the height model shown in **Figure 2** would make it unsuitable for many applications.

Computed Inventory Attributes

Even when an inventory contains an extended list of field-measured attributes (e.g., the US Forest Inventory and Analysis Program includes up to 50 tree attributes in its core measurement protocol), the attributes of greatest interest to inventory users almost always require modeling. Timber volume has long been the most commonly modeled tree attribute. Now, a recent focus centers on estimating biomass and carbon, not only for standing live trees but also for standing and down dead wood, such as snags and logs. At the plot level, all manner of management-relevant attributes are routinely estimated, such as biodiversity, wildlife habitat potential, fire hazard, and susceptibility to insect attack.

Whole-tree volume equations are widely available for most tree species utilized as timber and can be found in articles, published compendiums, and documentation of inventories. For the most common species, it is not uncommon to find several equations to choose from, for example, equations designed primarily for a particular region, for various merchantability and processing standards, in various units of measure, or for old versus second growth stands. Most such equations take field measured diameter at breast height (dbh), height, and sometimes site index as inputs, and produce estimates of either total bole volume or bole volume within locally prevailing merchantability limits (e.g., above a specified stump height up to a minimum top diameter). For species not commonly utilized, the selection of equations is often quite limited, and inventory compilers must choose between using equations developed for other species believed to be similar or resort to using basic geometric formulas, an approach that fails to account for variation of tree form or taper. Such formulas are also useful for estimating the volumes of down wood pieces (dead woody material on the ground, such as logs or branches), with the choice of formulae constrained by which diameters are measured (e.g., large end, small end, or at the point of transect intersection).

Some inventories involve field estimation of the numbers of logs in each tree, and, sometimes, the diameter at the breakpoints between logs (e.g., via reloskop). This opens the possibility for modeling volume for each log using volume formulas or tables designed around log size, or modeling tree volume based on dbh and the number of logs in the tree. This extra effort is sometimes justified when more precise merchantable wood volume estimates are needed.

Tree biomass is frequently estimated to assess biological productivity and ecological dynamics, to characterize fuel loadings for fire hazard assessment, and to serve as a basis for modeling carbon stores and dynamics. Bole biomass can be estimated as a scalar multiple of cubic volume, where scale factors reflect wood density and are estimated separately for each species or species group. Branch and leaf/needle biomass attributes are usually estimated via allometric relationships with dbh, and sometimes site index and/or height, which are developed via destructive sampling – harvesting, clipping, drying, and weighing of these plant parts from a representative sample of trees for each species of interest. Biomass of shrubs and herbs is also of great interest, and can be estimated from cover and height measurements taken in the field; however, there can be considerable variation among species, and cover-based biomass equations have been devised for very few shrub and herb species, making the use of equations from 'comparable' species a common fallback.

Amounts of elemental carbon can be derived as a scalar multiple of biomass, with different scale factors used for woody and herbaceous plant parts. Carbon and biomass amounts in down wood must be adjusted for the degree of decay/decomposition that has already occurred in the down-wood piece. Adjustment factors are generally specified when a down-wood decay class system is established for field classification of decay degree.

Wildlife habitat is another increasingly important forest attribute that can be modeled from forest inventory data. Habitat for individual species may be described by forest attributes such as average tree size or canopy cover. For example, a 2002 analysis of habitat for northern spotted owls (*Strix occidentalis*) in Washington, USA, found that averaged roosting habitat variables were tree dbh of 37 cm, canopy cover of 84%, 51 snags ha^{-1}, and 120 tonnes ha^{-1} of down wood. Such information can be paired with forest inventory data to build models estimating how much habitat might exist for this species. Such a model would be an example of calculating habitat attributes specific to an individual species.

Another common approach uses a standardized forest classification system that is then given species-specific rating values. For instance, one class in the California Wildlife Habitat Relationships system is tree species of red fir with average tree diameter greater than 61 cm and canopy cover $>60\%$; for feeding habitat, this class is rated as 'high' for the spotted owl, 'medium' for the big brown bat (*Eptesicus fuscus*) and 'low' for the mountain lion (*Puma concolor)*. Additional habitat elements, such as presence of water or decaying logs, can contribute to these ratings of habitat suitability. Both species-specific and generalized classification systems are most often nonspatial, but an active area of research is spatially based models that factor in attributes such as patch size, proximity to water, distance between patches, and extent and type of edges.

The most common accuracy assessment method is to correlate model predictions with the presence or absence of wildlife species. Habitat alone is not enough to understand wildlife population distributions, because of impacts such as hunting, predation, intra- and interspecies competition, and migration. None the less, habitat classification is a useful method for understanding how impacts that change forest species, age, or structure may affect wildlife species.

The increasing frequency of large, wildland fires and broad agreement that, in many areas, fire exclusion has led to changes in forest structure that make today's forests more vulnerable to stand-replacement fire, has generated substantial interest in predicting fire and fuel attributes from inventory data. Such models can be ad hoc, for example, indices of stand density or standing biomass thought to be related to potential fire impacts. Others, which involve processing inventory data through formal simulation models, predict attributes related to fuel laddering and crown fire, such as torching and crowning indices, which depend on height to crown base and crown bulk density (itself a modeled attribute). When such models are coupled with stand-projection models, attributes like crown fire potential can be evaluated at many points in the trajectory of a stand, and under a variety of silvicultural management and fuel treatment regimes.

Other fire-modeling approaches utilize whatever field-measured inventory data attributes are available. They impute additional attributes by relating measured attributes to those in a database of reference plots, and ultimately predict levels of a multitude of surface fuel, surface fire behavior, crown fire potential, and fuel consumption attributes. Enthusiasm for all of these fire-modeling approaches is tempered by the scarce validation evidence associated with these models, and the substantial challenges to obtaining such evidence in adequate quantities.

Inventory data are also used to model surface fuel characteristics via classification (e.g., into stylized, surface fuel 'models') based on forest type. When a multipurpose inventory includes down-wood measurements (most commonly collected via line-intersect sampling transects), along with litter depth and mass and understory height and cover, surface fuel loadings by fuel size class can be directly estimated, potentially enhancing the specificity and accuracy of fire behavior and outcome predictions.

Spatial Modeling of Forest Inventory

Paper maps were the earliest form of spatial models used in forestry and they continue to enjoy wide use. Primitive, computer-based mapping systems were developed in the late 1960s, and these ultimately evolved into the modern geographic information systems (GIS) now used routinely in forest inventory applications.

Field measurements of forest inventory rely on sampling, or the selection of units from a larger population. The desirability of spatially continuous models as a basis for many applications has motivated the development of numerous techniques for extrapolating plot measurements to larger landscapes. Aerial photographs have long been used to map a forest through delineating stands (contiguous groups of trees that are similar in age, species, or structure). Forest attributes may be assigned directly by the air photo interpreter. For attributes that are difficult to estimate, an alternative is to delineate stand boundaries through photo interpretation, but assign attributes calculated from inventory plots or transects that were measured within the stand boundaries.

Since the 1970s, remote sensing imagery has been used to provide continuous spatial models of forest attributes (*see* **Resource Assessment**: GIS and Remote Sensing). Remote sensors do not directly measure the attributes of interest. With the most commonly used sensors, the digital data represent electromagnetic radiation intensity in different spectral bands or wavelengths. A number of standardized ratios of intensity in one band relative to another, and texture measures (typically based on variance in reflectance within a moving window) have also proved helpful. With the aid of additional sources of information such as geo-referenced plot information or stand maps, mathematical models are then used to relate this digital data to the forest attributes of interest. Although Thematic Mapper data collected by LANDSAT satellites have been most widely used, other remote sensing information such as AVHRR, SPOT, SLAR, and aerial infrared video have proved useful for specific applications. As part of the earth-observing system (EOS), a variety of new sensors began to be launched in 1999, and may prove useful for large-area ecological modeling of forests. There is also great interest in using the hyperspatial (~ 1 m resolution) and hyperspectral data recently available from space-based platforms to provide species-level and even tree-level representations of the forest, and many believe that light detection and ranging (LIDAR) imagery shows great promise for remote sensing of canopy structure and subcanopy vegetation layers.

The mathematical models used to create spatial models of forest attributes have evolved over time. Automated or semiautomated classification has been the standard method for mapping forests from satellite imagery for over two decades, including supervised and unsupervised approaches. Classification relies on statistical procedures, usually based on maximum likelihood estimation, to assign each pixel in the landscape to a forest-type class. Density and size classes are also sometimes assigned from image data. Classification accuracy is typically assessed via a confusion matrix – essentially a cross-tabulation of predicted and observed values, where the latter are derived from a set of ground control points that are either visited in the field or manually interpreted from aerial photographs. It is generally the case that the larger the number of classes (e.g., forest type by size by density) attempted in the classification, the lower the overall accuracy of the classification. Despite experimentation with various fuzzy classification approaches that assign partial credit for near-misses, overall accuracy rarely exceeds 85%.

Disillusionment with classification accuracy and the comparatively greater flexibility (e.g., in modeling derived attributes) that comes with continuous modeling of inventory attributes has led to the increasing popularity of imputation approaches to spatial interpolation of inventory plot data, including most similar neighbor (MSN), gradient nearest neighbor (GNN), and k nearest neighbor (kNN) variants. The first two impute all of a single inventory plot's attributes to all unsampled pixels in the landscape judged most similar on the basis of a similarity- or gradient-based weight matrix (**Figure 3**), and the third combines the k plots nearest in attribute space to an unsampled pixel to assign attributes to that pixel. Neural nets have also been used to merge inventory data and remote sensing data to provide spatial models of vegetation. Developers report that the technique can be used to develop maps of acceptable accuracy rapidly, but the models function as 'black boxes' that provide little insight into how a map was produced. The accuracy of any of these imputed maps can be assessed via cross-validation with samples not used in the development of an imputation model. As with other types of models, the appropriate use of spatial, GIS, and remote sensing models requires understanding the accuracy of the model for the specific application being tested.

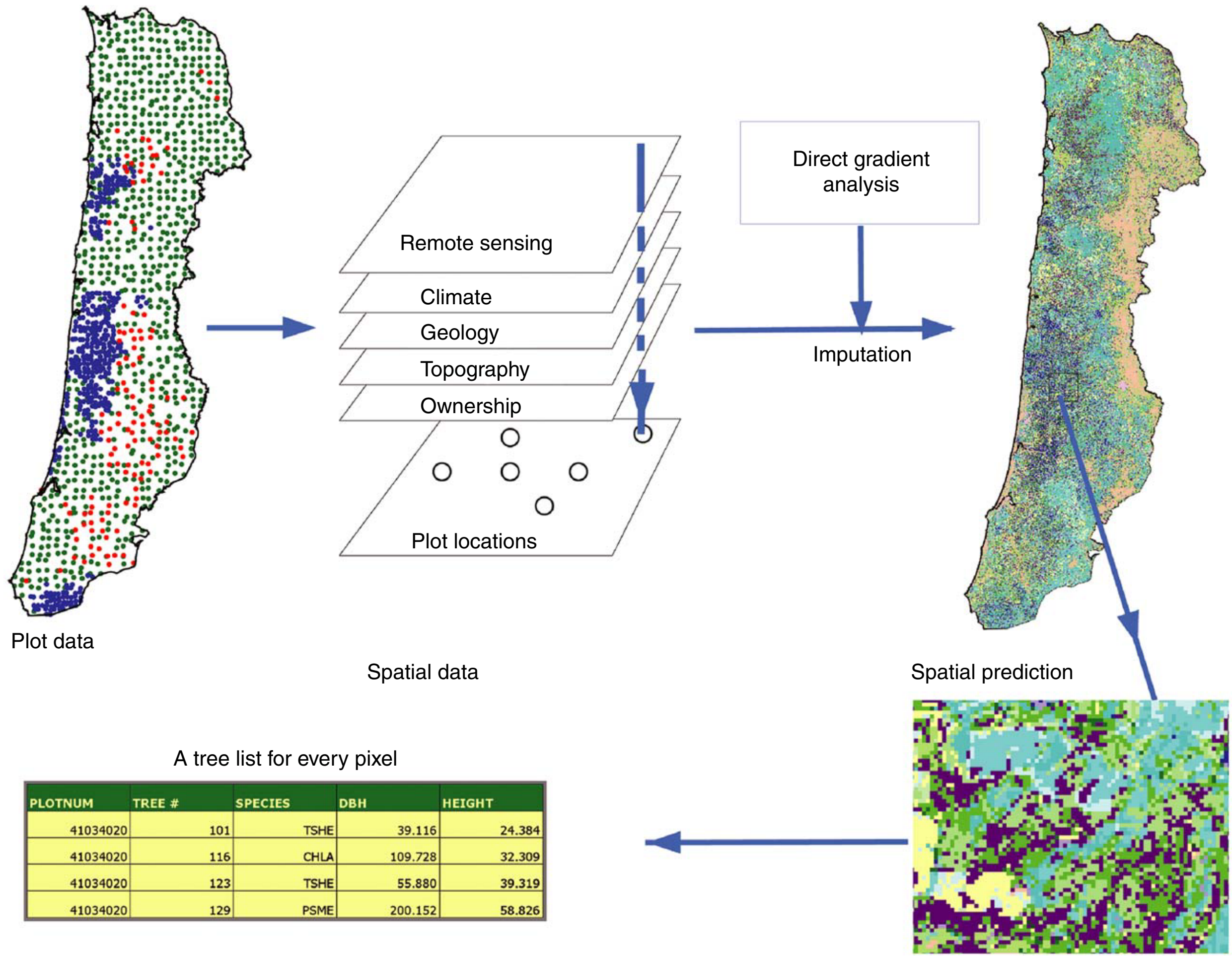
One application that combines spatial and temporal modeling is inventory updating. Continuous forest inventory (CFI) systems often measure a portion of plots each year. This approach results in having a portion of inventory data always current, but it also means that most data are from previous years. The national inventory of the USA uses such a system, and a variety of methods have been proposed to update inventory to the current year. Some methods involve imputation, or the substitution of information from similar trees or plots. Other methods use modeling procedures to update the plots, with or without the use of ancillary remote sensing data.

Temporal Modeling of Forest Inventory

Temporal modeling helps us understand how forests have changed in the past and predict how they will change in the future. Temporal modeling that accounts for management alternatives often serves as a basis for planning, and can help ensure the sustainability of forest resources.

Yield tables, which display timber volumes for forests of different ages, were one of the earliest and simplest temporal models. Simulation models that dynamically project forest attributes over time now serve as the basis for most temporal modeling (*see* **Mensuration**: Yield Tables, Forecasting, Modeling and Simulation). Growth models of various types are typically embedded within forest planning models, and use initial conditions and a set of assumptions to project forest characteristics over time. The planning model is used to vary the assumptions, establish a basis for choosing among management alternatives, and to synthesize and display forest-wide outputs. For example, a nonspatial individual tree growth model that predicts height and diameter growth for every tree may be used to project forest inventory for a set of silvicultural alternatives. The planning model might combine the output from the growth model with cost and revenue information, include an optimization algorithm to allow decision-makers to test different management strategies, and link to a GIS or visualization program to display what the forest would look like in the future.

Different purposes require approaches that vary in scope (spatial and temporal scale) and resolution. Planning models may cover very short time frames for small land areas, or they may cover long time frames and large land areas. Foresters divide



PLOTNUM	TREE #	SPECIES	DBH	HEIGHT
41034020	101	TSHE	39.116	24.384
41034020	116	CHLA	109.728	32.309
41034020	123	TSHE	55.880	39.319
41034020	129	PSME	200.152	58.826

Figure 3 Inventory plot data can be imputed via the method of gradient nearest neighbors using multivariate techniques like canonical correspondence analysis. This approach creates a full suite of internally consistent plot attributes for every pixel in a forested landscape based on the similarity of remote sensing and ancillary geographic information systems (GIS) layers at plot and unsampled locations.

planning models into hierarchical levels to reflect this range. Strategic models may be used to project forest inventory attributes over a 50–200-year time horizon, typically use time steps of 10 or 20 years, are used for forests of 20 000 to 200 000 ha, and generalize vegetation into fairly broad classes. Tactical models may be used to project forest inventory over a 10–100-year time frame, are used for forests of 1000–20 000 ha, may use smaller time steps, and may allow spatially explicit projections of vegetation over time. Operational models may be used to project forest attributes over short time horizons of 10–20 years, often use single-year time steps, are applied to forests of less than 1000 ha, recognize stands as unique entities, and are focused on short-term decisions such as what harvesting system to use, which roads to build, and which roads to close. As computing power has grown exponentially over the past two decades, these hierarchical distinctions have become blurred and it has become common to include greater detail and spatial resolution for large land areas and long time horizons.

Strategic, tactical, and operational planning models use a variety of mathematical techniques to aid in selecting silvicultural and management alternatives. Until the 1990s, linear and goal programming were the most common solution methods. Spatial restrictions between neighboring stands, whether imposed by regulation or necessitated by management goals, required different techniques. Integer programming and a number of heuristic methods, including simulated annealing, genetic algorithms, and tabu search, have been applied to spatially restricted planning problems.

Extremely large land areas are addressed with forest sector models, which are used to project forest inventory on a national or international scale. Forest sector models differ from planning models by relying

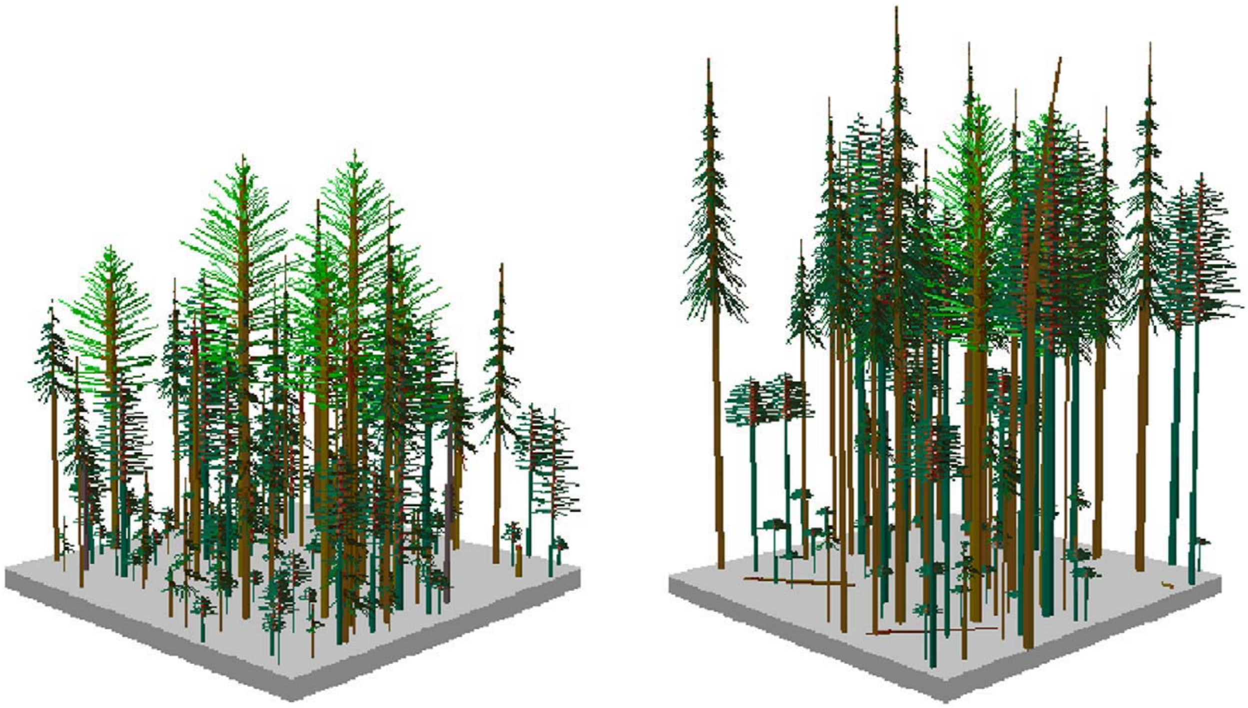


Figure 4 Visualization model of hemlock-fir inventory plot as measured in 2000 (left), and after 100 years using models of growth, mortality, and snag decay (right). Plot data are from Forest Inventory and Analysis program, growth model used is Forest Vegetation Simulator program, and visualization program is the Stand Visualization System.

on supply and demand functions rather than treating prices as fixed. Forest sector models also differ in that they are not explicitly linked to management decisions, although they may project future inventory levels for a variety of possible scenarios. Forest sector models provide useful insight into how forest inventory levels change in response to trade restrictions, technology development, or general demand trends.

Along with growth models, forest planning models, and forest sector models, many other kinds of models aid in projecting inventory attributes over time. Snag models may be used to understand how dead trees are recruited, decay, and fall down. Forest fire simulation models are used to predict how forest stands or landscapes would burn under different assumptions of weather, initial conditions, and suppression strategies. Ecological process models have been used to understand how net primary productivity and biomass would change under different climate and atmospheric carbon dioxide levels. Visualization models are used to generate graphic displays of how trees, stands, and landscapes would look in the future under alternative management scenarios. Models are frequently combined in any particular application (for example, see **Figure 4**). Relatively little research has been devoted to the effect on error and accuracy of predictions when models of different spatial and temporal scales are combined.

Modeling has become a regular component of day-to-day forest management. Improvements in validation may follow as the specific fields of application mature. The vast change in the past century, from simple hand-drawn graphical models of basic allometric relationship, to elaborate computer models intended to simulate global ecosystem processes, indicates that modeling in forestry will continue to be an active area of research.

See also: **Biodiversity**: Biodiversity in Forests. **Inventory**: Forest Inventory and Monitoring. **Landscape and Planning**: Forest Amenity Planning Approaches. **Mensuration**: Yield Tables, Forecasting, Modeling and Simulation. **Operations**: Forest Operations Management. **Plantation Silviculture**: Sustainability of Forest Plantations. **Resource Assessment**: GIS and Remote Sensing.

Further Reading

Avery TE and Burkhart HE (2002) *Forest Measurements*, 5th edn. New York: McGraw-Hill.

Davis LS, Johnson KN, Bettinger P, and Howard T (2001) *Forest Management*, 4th edn. New York: McGraw-Hill.

DeStefano S and Haight RG (eds) (2002) *Forest Wildlife–Habitat Relationships: Population and Community*

Responses to Forest Management. Bethesda, MD: Society of American Foresters.

Kohn KA and Franklin JF (eds) (1997) *Creating a Forestry for the 21st Century: The Science of Ecosystem Management*. Washington, DC: Island Press.

Ohmann JL and Gregory MJ (2002) Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research* 32(4): 725–741.

Moeur M and Stage AR (1995) Most similar neighbor: an improved sampling inference procedure of natural resource planning. *Forest Science* 41: 337–359.

Waring RH and Running SW (1998) *Forest Ecosystems: Analysis at Multiple Scales*, 2nd edn. San Diego, CA: Academic Press.