geospatial technologies

A Review of Methods for Mapping and Prediction of Inventory Attributes for Operational Forest Management

Kimberley D. Brosofske, Robert E. Froese, Michael J. Falkowski, and Asim Banskota

Forest inventory attributes are an important source of information for a variety of strategic and tactical forest management purposes. However, it is not possible or feasible for field inventories to be conducted contiguously across large areas, especially at a resolution fine enough to be useful for operational management. Therefore, a large number of quantitative modeling and prediction methods have been and are being developed and applied to predict and map forest attributes, with the goal of providing an accurate, spatially continuous, and detailed information base for practitioners of forestry and ecosystem management. This article reviews the most commonly used prediction techniques in the context of a comprehensive modeling framework that includes a discussion of methods, data sources, variable selection, and model validation. The methods discussed include regression, nearest neighbor, artificial neural networks, decision trees, and ensembles such as random forest. No single technique is revealed as universally superior for predicting forest inventory attributes; the ideal approach depends on goals, available training and ancillary data, and the modeler's interest in tradeoffs between realism and statistical considerations. Useful ancillary data included in the models tend to include climate and topographic variables as well as vegetation indices derived from optical remote sensing systems such as Landsat. However, the use of airborne LiDAR in modeling of forest inventory attributes is increasing rapidly and shows promise for operational forest management applications. Different considerations are encapsulated within a generalized model development framework that provides a structure against which tradeoffs can be evaluated.

Keywords: forest inventory, geospatial, imputation, mapping models, remote sensing

Detailed, accurate information about forest composition and structure is needed for effective management of forest ecosystems. Whereas the goal of forest inventory is to quantify the amount and type of forest resources and related attributes in a given area, the goal of forest mapping is to depict the spatial distribution of those resources and related attributes (Corona 2010). Detailed information on forest composition and structure is necessary for silviculture (e.g., Pond et al. 2014), ecological restoration and risk assessment (e.g., Pierce et al. 2009), assessment of biodiversity and habitat (e.g., Turner et al. 2003, Martinuzzi et al. 2009, Helmer et al. 2010), carbon management and reporting (e.g., Hall et al. 2006, Gibbs et al. 2007, Froese et al. 2010, Hoover and Rebain 2011), and forest health assessment and management (e.g., Solberg et al. 2004, Wolter et al. 2009), among other purposes. Forest inventories can provide this information for sampled plots, which can be summarized to the stand or higher levels via standard statistical procedures. However, forest and ecosystem managers benefit from spatially continuous maps of inventory data that can be tailored to specific objectives or projects.

Information needs differ across the forest planning spectrum. Although strategic planning efforts typically leverage coarse-resolution information, fine spatial resolutions such as stand-level summaries and tree-level attributes may be needed for tactical planning and operational forest management (Falkowski et al. 2009b). For example, tree-level data are needed for initialization of forest growth and dynamics models for use in operational management (Falkowski et al. 2010), whereas county-, state-, or regional-level inventory summaries are used for forest planning at the regional or national levels (McRoberts et al. 2010b). Efforts to conserve biodiversity and restore ecosystems for the benefit of wildlife communities often also require information at fine spatial scales. For example, Lesak et al. (2011) found forest structure attributes derived from airborne light

detection and ranging (LiDAR) to be useful in predicting songbird diversity in temperate forests. Because data can be easily summarized and scaled up, spatially continuous (i.e., "wall-to-wall") maps of inventory attributes at detailed to moderate spatial resolutions (e.g., 1−30 m) provide the greatest flexibility for meeting the widest variety of forest information needs by allowing researchers and forestry practitioners to tailor analyses to meet particular objectives.

In the context of operational management, forest inventory attributes may refer to a wide range of variables including individual tree measurements (e.g., species, diameter, height, volume, and biomass) as well as stand- or plot-level averages or totals of tree-level measurements (e.g., basal area, tree density, quadratic mean diameter, biomass, and volume). Collecting detailed inventory data across large, contiguous regions is not practical via traditional forest inventory methods because of time, expense, and other conditions; therefore, it is useful to have an effective way to predict attributes of interest at unsampled locations. A variety of statistical methods have been used toward these goals. Although regression is common, imputation, interpolation, and machine-learning algorithms are also frequently used, typically by taking advantage of remotely sensed and other ancillary data sets that are available across an area of interest. The ancillary data function as explanatory variables in models predicting necessary forest attributes, using ground samples from the forest inventory or other source (i.e., reference data) for model training and validation. Most comparisons of modeling and prediction methods that exist in the literature apply two or three methods to specific data sets (e.g., Berterretche et al. 2005, Hudak et al. 2008, Eskelson et al. 2009, Powell et al. 2010, Goerndt et al. 2011). Because the number of individual studies and methods used is large, it can be challenging to gain even a broad sense of strengths and weaknesses of the many different approaches.

In development of a prediction model, in addition to the difficulties associated with selection of an appropriate prediction method, the selection of ancillary data and variables can be daunting. Ancillary data can originate from many different sources, including satellite or airborne sensors (e.g., optical, radar, and LiDAR), digital models, aerial photo interpretation, field-based mapping efforts, or a combination of these. For example, prediction of fine-scale (e.g., tree-level) attributes might require high spatial or three-dimensional resolution in the source data (e.g., LiDAR instead of Landsat). It can also be difficult to decide which variables are most appropriate within a particular context. Although certain variables may be useful for one study area, they may be irrelevant to another with different primary environmental gradients. Conversely, there may be some predictors that are universally useful (e.g., vegetation indices such as the normalized difference vegetation index [NDVI]). Variable selection can itself be the essence of a modeling exercise when the purpose is to understand the relationship between inventory and predictor variables. However, even if generating accurate predictions is most important, selecting useful variables and discarding others can have a significant impact on model performance (Hudak et al. 2008, Packalén et al. 2012). Indeed, variable selection can be more important than method selection in modeling of forest inventory.

Both regional forest planning and operational forest management could benefit from detailed, geographically extensive forest inventories derived via statistical modeling techniques leveraging remotely sensed data. Although many studies have examined the use of various types of ancillary data as well as the utility of various approaches for mapping particular data sets (e.g., Cairns 2001, Lef-

sky et al. 2001, Labrecque et al. 2006, Maselli and Chiesi 2006, Wallerman and Holmgren 2007, Hudak et al. 2009, Goerndt et al. 2011, Tonolli et al. 2011), a cohesive synopsis of analytical methods for developing geospatial forest inventory mapping models in the context of the overall modeling process is particularly valuable. The primary objective of this article is to provide an overview of common approaches for generating maps of operational forest inventory attributes across large spatial extents. This overview includes a review and discussion of (1) common statistical prediction methods (including strengths and limitations), (2) several variable selection approaches suitable for different data sets and prediction techniques, (3) sources of modeling data and capabilities of different remote sensing data for predicting various forest inventory attributes, and (4) relevant aspects and approaches for validating and assessing the accuracy of the prediction models and map products.

Modeling and Prediction Methods

The methods that have been used for the prediction and mapping of detailed forest attributes from geospatial data are many and varied. Most models have been built to predict univariate responses, but the desire to reduce analysis time, maintain the covariance structure of the data, and ensure realistic combinations of attribute values has resulted in increased emphasis on multivariate techniques that predict a suite of attributes simultaneously (Moeur and Stage 1995, Lister 2009, Nothdurft et al. 2009, Falkowski et al. 2010). Many parametric approaches are common in the forestry literature, including maximum likelihood estimation (Hagner and Reese 2007, Baatuuwie and Van Leeuwen 2011), discriminant analysis (Thenkabail et al. 2004, van Aardt et al. 2008), and regression (le Maire et al. 2011, Tonolli et al. 2011). However, the nature of many forest inventory variables complicates the use of these methods. The popularity of nonparametric approaches means that these receive significant attention in our review. However, we include an overview of regression as well, given the importance and unique advantages of that method.

Regression

As a parametric technique, regression has traditionally been the most widely used method for predicting forest attributes of operational interest from ancillary data and remains a common analytical tool (e.g., Fiorella and Ripple 1993, Means et al. 1999, Wolter et al. 2009). For example, le Maire et al. (2011) used stepwise multiple linear regression to predict stand dominant height and stand merchantable wood volume from age, NDVI, and a number of bioclimatic variables in Brazil. Hudak et al. (2006) used multiple linear regression models to predict and map basal area and tree density in temperate coniferous forests. In addition, many of the studies detailing the operational use of airborne LiDAR data in forest inventories have used regression to predict forest inventory variables such as tree height, basal area, volume, biomass, and density (e.g., Næsset 2007, Woods et al. 2011). For example, Stephens et al. (2012) used a regression developed from airborne LiDAR height and canopy cover metrics to model total carbon stocks in New Zealand.

The most common form of regression for forest attribute prediction uses a linear or intrinsically linear multiple regression equation solved using ordinary least squares (OLS) (e.g., Næsset 2007, le Maire et al. 2011, Tonolli et al. 2011). Intrinsically linear equations

are those that may be linearized by transformation, such as multiplicative models that are linear on the log scale (e.g., Næsset 2002). When the coefficients are transformed, they may be estimated with OLS; however, if the coefficients are back-transformed to the original scale, adjustments may be necessary to correct for bias (e.g., Fehrmann et al. 2008). For example, Næsset et al. (2005) used a multiplicative regression model, linearized with a log transformation, to predict forest biophysical attributes (mean height, volume, and others) using derived attributes from LiDAR in Norway. In some cases, truly nonlinear models may better represent many of the ecological relationships of interest at the operational level. For example, Muukkonen and Heiskanen (2005) used nonlinear regression to model boreal forest attributes including volume, age, and overstory biomass as a function of ASTER spectral data but found that a linear model was sufficient for understory biomass.

Frequently, the usual OLS assumptions about errors (independent, homoskedastic, and Gaussian) (Sokal and Rohlf 2012) are not met, and remedial measures are used. These include transformation of the response variable (e.g., Frazer et al. 2011), weighted least squares (e.g., Brown et al. 1989), or a shift to alternative methods for model fitting, such as partial least squares (e.g., Coops et al. 2003, Townsend et al. 2003), generalized method of moments (e.g., Nord-Larsen and Schumacher 2012), or mixed-effects models using restricted maximum likelihood (e.g., Fehrmann et al. 2008). The optimum approach depends on the context of the problem. The use of generalized linear models is common when the goal is modeling distributions (e.g., diameter and height) and the error covariance matrix is explicitly non-Gaussian. See, for example, Breidenbach et al. (2008), who modeled diameter distribution using a Weibull function. Parameter estimates from OLS can be unstable in the presence of multicollinearity, which can lead to models that extrapolate poorly outside of the training data or are difficult to interpret. Careful screening during model construction is important, using diagnostics such as condition number (e.g., Popescu et al. 2003, Næsset et al. 2005) or variance inflation factor (Sokal and Rohlf 2012). Alternative methods like partial least squares or canonical correlation analysis (CCA) (e.g., Cohen et al. 2003, Schroeder et al. 2007) can reduce multicollinearity because they extract linear combinations of predictors that optimize the relationship with the response variable. However, such models are much more difficult to interpret and may not provide any measurable improvement over OLS in development or application (Næsset et al. 2005).

OLS regression assumes that there are no measurement errors in the explanatory variables, which is often not realistic, especially with regard to remote sensing variables (Curran and Hay 1986, Berterretche et al. 2005). Standard major axis (SMA) regression has been proposed and used increasingly as a way to account for the errors in both the response and explanatory variables (Cohen et al. 2003, Berterretche et al. 2005). The method is also known as reduced major axis or geometric mean regression, but the term SMA regression is preferred (Sokal and Rohlf 2012). Cohen et al. (2003) compared SMA and OLS regression for predicting leaf area index (LAI) and canopy cover and found that SMA regression maintained the variance structure of the observations in the predictions, whereas OLS regression did not. Schroeder et al. (2007) preferred SMA regression for this reason, using it to predict tree cover (in an absolute percentage of ground surface) from Landsat reflectance bands applied to normalized Landsat images in a date-invariant approach.

Recursive Partitioning: Classification and Regression Tree (CART) and Multivariate Adaptive Regression Splines (MARS)

CARTs, sometimes called decision trees, work by recursively subdividing training (i.e., reference) data into more and more homogeneous subgroups. Classification trees are those that predict a categorical response, such as forest type, whereas regression trees predict a continuous response, such as basal area. At each level, a splitting rule is developed to partition the observations into two subgroups: for classification, the rule maximizes intergroup variability, whereas for regression the rule minimizes the sum of squared residuals (Hastie et al. 2009). This results in a hierarchical "tree" that places observations into end nodes with other similar observations.

Decision trees have been used extensively for classification of forest and other vegetation types (e.g., Ruefenacht et al. 2008, Mora et al. 2010) as well as for prediction of forest structural attributes (e.g., Falkowski et al. 2005, Blackard et al. 2008, Helmer et al. 2010). Decision trees have a large number of characteristics that make them an attractive option for prediction in forestry. They are nonparametric and can readily accommodate nonlinear responses, variable interactions, both continuous and categorical explanatory variables, and missing values and are relatively unaffected by outliers or multicollinearity (Urban 2002, Piramuthu 2008). Decision trees can also handle large data sets and large numbers of variables and are useful for variable reduction because a measure of the relative importance of the different variables can be calculated. Here, relative importance is a sum of internal metrics used to construct individual trees and is different from "variable importance," commonly provided for some ensemble methods (see Hastie et al. (2009) for a detailed explanation). Trees can be represented graphically with ease, and therefore even complex data structures such as thresholds and nested responses can be interpreted readily (Young et al. 2009). Although typically used to predict a univariate response, De'ath (2002) demonstrated their utility in predicting multivariate ecological responses. Limitations of decision trees have been found to include potential instability in the models (i.e., different tree structures) that can result from slight changes to the input data (Prasad et al. 2006). Although use of closely correlated variables as surrogates for missing data allows flexibility in the use of incomplete data sets, this can result in very different end models if a variable used heavily as a surrogate is removed from the analysis. In addition, some decision tree algorithms (e.g., CART) (Hastie et al. 2009) also show selection bias, in which a predictor is more likely to be chosen if it has a greater number of distinct values or a larger sample size (Young et al. 2009). Some algorithms have been developed to counter this problem (e.g., GUIDE) (Loh 2011). A large imbalance of cases in the response classes can also result in a greater number of predictions falling within the more abundant class (He and Garcia 2009). The use of stratified sampling designs may help balance sample size across classes.

MARS (Friedman 1991) produces a model based on local relationships between response and predictors using a set of adaptive piecewise basis functions (i.e., linear regressions) to fit separate splines to intervals of the explanatory variables. The "knots" (endpoints of the intervals) and the identification of the predictors are determined by the basis functions in a forward and backward stepwise procedure that involves overfitting the model and then removing the knots that contribute least to model fit. For continuous variables, the discontinuous splits obtained with decision trees may be arbitrary. By combining spline fitting and recursive partitioning, MARS overcomes this problem and models continuous smooth

functions that reflect the local pattern. However, predictions can be unrealistic because the local nature of the data may affect the basis functions excessively; further, selection of appropriate parameters can be tedious and time-consuming, often involving trial and error (Prasad et al. 2006). MARS has been used successfully to predict forest stand biomass, age, diameter, and crown cover in the western United States (Moisen and Frescino 2002), and Muñoz and Felicísimo (2004) demonstrated its utility for mapping potential plant distributions in Latin America and Spain.

Artificial Neural Networks (ANNs)

ANN models are nonlinear, "black box" approaches to prediction and mapping that were originally based on the biological problem-solving process of the brain. The idea is that the data describing the unknown situation are presented to the trained model, which then provides the predictions. The training of the model involves "learning" the patterns from the observations, which takes place in different ways, depending on the ANN chosen. There are many types of ANN algorithms that can be used based on the specific objectives and characteristics of the data set. The most popular algorithm is the multilayer perceptron (MLP) (also known as multilayer feed-forward networks trained by back-propagation or simply back-propagation networks), although other algorithms such as Kohonen self-organizing maps and Hopfield networks have also been used with success (Lek and Guégan 1999). The MLP is a supervised model in that data with known outputs are used in the model training procedure. Data flow from the input layer containing the explanatory data through one or more hidden layers, with predictions provided in the output layer. The hidden layers are where the "learning" takes place and contain a number of nodes that represent a set of models or weights that are applied to the predictors to obtain the output values. These weights are determined in each hidden layer node by linearly combining all the input variables into a derived variable and then performing a nonlinear transformation on the derived value (Campbell et al. 2007). The learning process in the MLP algorithm consists of checking predictions based on a certain set of weights against the known values and then going back and correcting the weights to reduce the error if needed, until the network gives acceptable predictions. The user specifies the number of hidden layers and hidden layer nodes to use in the model. Multiple hidden layer nodes may increase the predictive ability of the model, but too many may result in model overparameterization and decreased predictive ability (Blackard and Dean 1999). Use of too many nodes will also increase the computing resources needed to train the model. Selection of appropriate model architecture and parameters can be challenging, typically requiring a trial-and-error process (Blackard and Dean 1999, Peng and Wen 1999, Niska et al. 2010).

Although not as widely used as other approaches in ecology, ANNs have been applied to forestry data, especially classification problems (Fitzgerald and Lees 1992, Peng and Wen 1999, Rogan et al. 2008). Fewer studies have used ANNs to predict continuous variables (but see Jensen et al. 1999, Moisen and Frescino 2002, Niska et al. 2010). Ingram et al. (2005) used ANNs to model plot- and stand-level basal areas from Landsat ETM+ spectral data in a tropical forest, but they were not able to obtain a significant ANN model for predicting stem density from the same predictors. A study by Niska et al. (2010) in the boreal forest achieved high accuracy using three ANN models to predict species-specific plot and stand volumes from a number of spectral and textural variables derived from aerial photographs and airborne LiDAR data.

Neural network approaches may be well suited to problems in which accurate predictions are the primary goal but, because of their black box nature, are less suited to situations in which the relationships underlying the predictions are of primary interest (Cartwright 2008). Their uncomplicated application to data sets containing nonlinearity, nonnormality, heterogeneity, and multicollinearity (Campbell et al. 2007) makes ANNs potentially useful tools for ecologists, particularly when the underlying processes and structure are not known (Peng and Wen 1999). In addition, they are able to use data originating from multiple sources in the same model (Atkinson and Tatnall 1997). However, Niska et al. (2010) note that careful variable selection is important when an ANN is used because variables that have no relationship with the response can have a large influence on the predictive power of the model. The computing time required for training and testing the model can also be prohibitive, and the method can be difficult to understand and implement effectively, especially with regard to optimal selection of the hidden layers and nodes (Blackard and Dean 1999, Peng and Wen 1999, Moisen and Frescino 2002).

Ensembles: Random Forest and Other Methods

A collection of predictive models may be combined into an ensemble to leverage their collective strengths (Hastie et al. 2009). Perhaps the most popular ensemble method at present is the random forest (RF) decision tree algorithm, which combines predictions from a myriad of individual decision trees (Breiman 2001, Cutler et al. 2007). RF uses bootstrapping to select samples that are then subjected to model fitting, but only a small number of randomly selected predictor variables are used to find the best split at each node, decreasing the correlation between trees and reducing bias. Each tree is grown fully and then used to predict the out-of-bag (OOB) observations (i.e., observations not included in the bootstrap sample). A simple majority or average of the predictions determines the class placement or prediction of an observation. RF maintains the advantages of simpler CART algorithms but is not subject to overfitting and tends to be less biased (Breiman 2001, Prasad et al. 2006). A measure of relative variable importance similar to that used for CARTs may be obtained, but RF can also use OOB samples to construct a different metric. Here, variable importance is calculated as the difference in prediction accuracy when OOB values for each variable are permuted, compared with the actual values (Hastie et al. 2009). Randomization nullifies the value of a variable; thus, a large decrease in accuracy when a variable is permuted implies that the variable is important.

Although the application of RF in forestry is relatively recent, many studies have had success using it to classify vegetation (Grossman et al. 2010) and predict various forest attributes such as live biomass (Baccini et al. 2004, Houghton et al. 2007), tree cover or species presence (Evans and Cushman 2009, Liknes et al. 2010), forest structural stage (Falkowski et al. 2009b), and disturbance (Stumpf and Kerle 2011), as well as individual tree attributes (Yu et al. 2011). For land cover classification, Gislason et al. (2006) achieved higher overall accuracy using RF than using a single decision tree, although the authors did not specify whether this difference was statistically significant. Although Prasad et al. (2006) found that RF could be computationally time-consuming, Gislason

et al. (2006) found it to be very fast. RF is less interpretable than simpler tree methods because it is cumbersome to investigate the large collection of individual trees. However, metrics such as variable importance aid in interpretation, and other techniques such as conditional density estimation can be used post hoc to reveal the underlying biophysical relationships between attributes of interest and important predictor variables (Falkowski et al. 2009b, Henareh Khalyani et al. 2012).

Although RF has some significant advantages over single-tree methods, some recent studies have shown that variable selection can have an important influence on the results. For example, Falkowski et al. (2009a) and Murphy et al. (2010) have suggested that multicollinearity among the predictors might decrease RF model accuracy. Because tree-based methods are robust to overfitting, they have been commonly applied in contexts in which there are a large number of predictors relative to the number of observations in the training data set. However, because RF uses only a subset of predictor variables to find the best split at each node, the result can be sensitive to the inclusion of a large number of predictors with no causal relation, which reduces the probability of selecting useful variables (Hastie et al. 2009). Thus, careful variable selection strategies during model construction may lead to significant improvements.

Other ensemble methods have been used less commonly for operational forest management. Bootstrap aggregation, or "bagging," is a method in which, as in RF, many decision trees are grown and averaged using bootstrapping of the training data set. It differs from RF in that all variables are used to find the best split at each node, and thus bagging lacks the advantages of RF when construction of less correlated trees is desired. Prasad et al. (2006) found that bagging trees performed better than simple regression trees and MARS, and about as well as RF, in mapping the distributions of four common forest tree species in the eastern United States under different climate scenarios. In their evaluation of land cover classification techniques, Gislason et al. (2006) also found that their best bagging classifiers performed approximately on par with RF but better than a single decision tree. Stochastic gradient boosting (SGB) (Friedman 2002) is another refinement of the single decision tree approach. In SGB, many small trees are built sequentially from "pseudo" residuals (the gradient of the loss function of the previous tree). At each iteration a tree is built from a random subsample of the data set, producing an incremental improvement in the model. Using only a fraction of the training data increases both computational speed and predictive accuracy, while also reducing overfitting (Moisen et al. 2006). SGB has been found to perform better than standard classification and regression trees and other modeling techniques for species modeling (Guisan et al. 2007), species mapping and basal area prediction (Moisen et al. 2006), prediction of aboveground biomass (Carreiras et al. 2012), and assessment of forest damage (Kupfer et al. 2008).

Nearest Neighbors

Nearest neighbor techniques (also called reference sample plot methods) (Kilkki and Päivinen 1987) predict the response variable(s) in an unsampled target unit (e.g., pixel) by computing a statistical distance metric between the target and reference samples, or "neighbors," and then assigning the value(s) of the closest neighbor(s) to the target unit. The distance metric is computed from ancillary (sometimes termed "auxiliary") variables drawn from the feature space (sensu McRoberts 2012) that are common to both the reference and target samples. In other words, the reference and target samples are linked via the ancillary variables, which are usually obtained from remote sensing. Much of the early work on nearest neighbor techniques originated and was conducted in the Nordic countries, especially Finland (Tomppo 1991, Tokola et al. 1996). This approach, often referred to generically as k-nearest neighbor (k-NN), can use any number (k) of nearest neighbors to impute the target value, usually by averaging or by taking the mode or median if the response is categorical. Besides the size of k, other choices to be made when this method is used include the distance measure (Chirici et al. 2008, Hudak et al. 2008) and neighbor weighting scheme (if any). Although nearest neighbor techniques can be used to predict a single response variable, they are especially useful for the prediction of multiple response variables simultaneously. For example, Vauhkonen et al. (2010) used nearest neighbors to simultaneously predict tree species, diameter, height, and stem volume from airborne LiDAR data in Finland.

Several distance metrics have been used with k-NN analysis to predict forest attributes (Chirici et al. 2008, Hudak et al. 2008). In addition to simple metrics that are independent of the response variable, such as Euclidean or Mahalanobis distance, many metrics that are based on the observations of the response have been developed. The most similar neighbor (MSN) approach (Moeur and Stage 1995) calculates distance (or similarity) using the results from CCA and then predicts values using $k = 1$. Both the canonical coefficients and canonical correlations are incorporated into a squared difference function to weight the ancillary variables according to their multivariate predictive power and allow preservation of the covariance structure of the response data set. MSN or a variant in which $k >1$ has been applied in a wide variety of studies to predict and map forest structural attributes (LeMay et al. 2008, Nothdurft et al. 2009), and others have used this approach to obtain tree lists for unsampled areas (Eskelson et al. 2008). The gradient nearest neighbor (GNN) technique is similar to MSN but differs in that it uses CCA to establish relationships between the ancillary data and observed response variables as an initial step (Ohmann and Gregory 2002, Ohmann et al. 2011). Distances are calculated in multivariate ordination space using axis scores to weight the ancillary variables and eigenvalues to weight the axes and attribute values of the single closest plot ($k = 1$) are assigned to the target pixel. Wilson et al. (2012) used a modified GNN approach to map tree species distributions at a moderate resolution in the eastern United States, noting that the approach has potential for operational application. Crookston and Finley (2008) incorporated the RF algorithm into k-NN as a distance measure using an RF proximity matrix, which is based on the number of times pairs of observations end up in the same terminal node, and modified the technique to allow for multiple response variables. RF-k-NN has been found to be a useful approach for imputation of complex forest attributes such as forest structural stage (Falkowski et al. 2009b) and snag characteristics (Eskelson et al. 2012), tree-level forest inventory data (Falkowski et al. 2010, Vauhkonen et al. 2010), and also forest inventory variables at the plot level (Hudak et al. 2008) and pixel level (Chirici et al. 2008, McInerney et al. 2010). Other distance metrics have also been developed. Chirici et al. (2008) used three distance measures to predict growing stock volume in Italy, including a modification of the Mahalanobis distance metric in which the variance-covariance matrix was computed in a fuzzy way, giving greater weight to observations of the response that are closer to the mean (fuzzy distance), distance modified by multiple regression, and distance modified by nonparametric weights. The Finnish multisource national forest

inventory uses a four-step genetic algorithm to optimize weights based on large-scale variation in forest attributes as well as spectral information (Tomppo and Halme 2004, Tomppo et al. 2009). Despite the popularity of nearest neighbor techniques in the literature, there are surprisingly few critical comparisons of distance metrics. Hudak et al. (2008) compared several distance metrics, concluding that the RF approach was superior to Euclidean and Mahalanobis distances as well as to MSN and GNN for predicting basal area and tree density in a temperate conifer forest. Chirici et al. (2008) found that, for their data, fuzzy distance outperformed the other distance metrics, especially with regard to maintaining the stability of the predictions. Vauhkonen et al. (2010) compared RF and k-MSN approaches to predicting a multivariate set of tree structural attributes in Finland and found only marginal differences.

In k-NN, the optimal number of neighbors to use in calculating the attributes of the target is not well defined. Although accuracy of the predictions tends to increase with increasing k, variation in the predictions decreases as a result of averaging of a large number of values, and the differences between observed and predicted values increase for extreme values (Tuominen et al. 2003, McRoberts 2012). In addition, Franco-Lopez et al. (2001) found a tradeoff between reductions in overall error and decreasing producer's accuracy with increasing k. Several authors have also noted that, because predicted values represent an average of observed values when $k > 1$, the values predicted for the target may be unrealistic (LeMay and Temesgen 2005); this can be especially true with multivariate response variables, which, when averaged, may produce a combination of attributes that is not seen in the observed population. Using a single neighbor may be preferred in some mapping applications to retain the full range of values in the data but using more than one neighbor might be more appropriate for accurate large-area prediction (Franco-Lopez et al. 2001). In the case of $k = 1$, it is extremely important to have a representative sample of the population because predictions are limited to the range of observed data. In many applications, it is desirable to optimize k using the reference data. McRoberts (2012) gives several options for optimization schemes, including choosing the value of k that minimizes a specific assessment criterion (e.g., root mean square error [RMSE]) or that minimizes over- or underprediction; he also suggests a multivariate optimization criterion.

Neighbor weighting in k-NN is most commonly accomplished using weights that are inversely proportional to the statistical distance between the target unit and the reference unit. Studies evaluating alternative weighting methods or optimization of the weighting parameter are scarce. Using a case study of univariate volume prediction from Landsat data in the boreal forest, McRoberts (2012) found that complex optimization schemes did not appreciably improve predictions over equal neighbor weighting. However, McRoberts (2012) recommended further study in the case of multiple response variables and larger sets of predictor variables. Weighting has been shown to be useful for accounting for spatial gradients (e.g., altitudinal or latitudinal) in multivariable models for volume by species group in Finland (Tomppo and Halme 2004) and for classification of forest into species and productivity groups in Finland and northern Italy (Tomppo et al. 2009).

Nearest neighbor approaches have been applied widely to the problem of predicting forest inventory information from remotely sensed data. Their popularity stems from several advantages. First, nearest neighbor approaches are intuitively simple and can therefore be readily understood and applied. They are also nonparametric in

that they make no assumptions regarding the underlying distribution of the data. However, they can be computationally demanding, especially with very large data sets as are likely to be encountered when forest attributes are mapped at fine resolutions across large geographic areas. A review of the advantages and disadvantages of alternative nearest neighbor approaches used in forestry applications was recently published by Eskelson et al. (2009).

Model Development

Clearly, a large number of choices are available to the modeler in both the range of available methods and the number of readily available predictor variables in the form of remotely sensed and geospatial data. Careful decisions about method and variables are essential in model development and are intricately linked to the objectives of each modeling problem. Fundamentally, decisions may be driven by the degree to which the modeler seeks to explain or understand the system versus generating accurate predictions (Caswell 1976). If the goal of the model is to gain understanding, then ease of interpretation and ability to discern between alternative and potentially causal explanations is key. When the goal is prediction, then decisions about method and variable selection may be driven principally by the degree to which they improve accuracy, with little concern for or attention to interpreting model dynamics. In both cases, discarding irrelevant and redundant predictor variables can be important.

Method Selection

Several authors have compared the effectiveness of various techniques for predicting forest composition and structure. Pierce et al. (2009) found that GNN generally performed better than linear regression models or classification trees in temperate marine forests but that regression performed better in the more complex forests of adjacent temperate steppe and Mediterranean forests, suggesting that the best technique may be reliant on the characteristics of the particular forest or landscape and that nonparametric methods are not always better than parametric methods. In the boreal forest, Maltamo et al. (2009) found that MSN more accurately predicted a number of tree-level size and quality attributes than did a regression approach. Eskelson et al. (2012) found that RF-NN provided more accurate predictions of snag density by decay class than a two-step parametric model but less accurate classifications of the presence or absence of large snags than logistic regression. Berterretche et al. (2005) compared SMA regression and two geostatistical methods for predicting LAI and concluded that SMA was the most practical option because it did not need a high density of reference measurements for application. ANNs have been found to outperform traditional statistical approaches such as regression and discriminant analysis in a variety of studies (Blackard and Dean 1999, Jensen et al. 1999, Vasconcelos et al. 2001, Foody et al. 2003). Using Landsat data to classify complex tropical rainforest types, Sesnie et al. (2010) found classification accuracies to be comparable for ANNs and RF. However, other studies have found performance of ANNs to be worse than (Serpico et al. 1996) or similar to (Niska et al. 2010) those of more conventional models.

As the variety of outcomes of these studies suggests, there is no consensus on which method is "best" in a general sense. The determination of best depends on several factors that may be subjective, including the purpose of the analysis, available resources, data characteristics, underlying biophysical conditions of the forest, and the
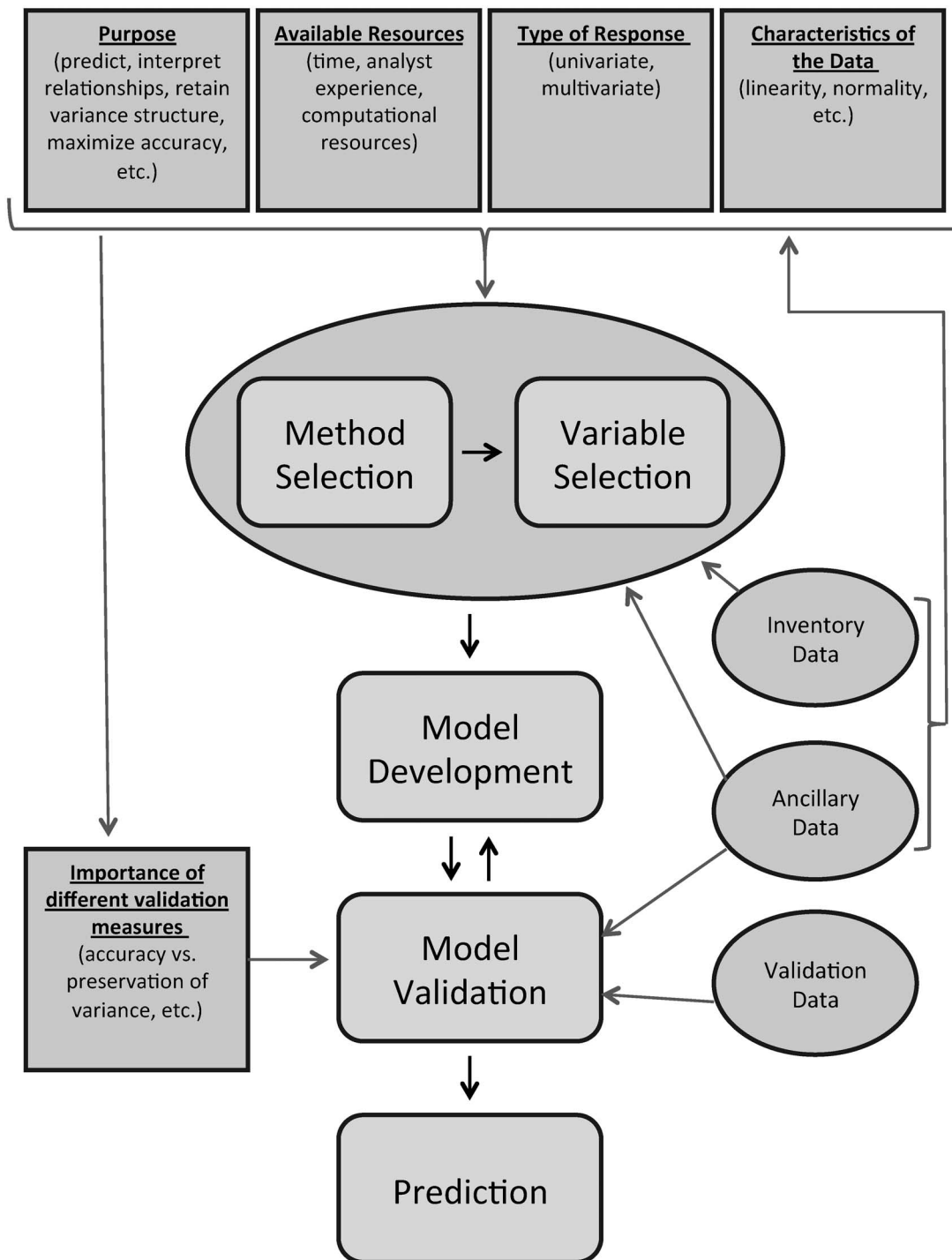
Figure 1. A diagrammatic representation of the model development and validation/evaluation process. Inputs include the problem context, which affects method and variable selection decisions. The final product is model predictions.

type of response variables (Figure 1; Powell et al. 2010) found that RF provided more accurate predictions (i.e., lower RMSE) of aboveground biomass than SMA or GNN, but the variance ratio indicated that SMA and GNN were better choices for preservation of the observed variance in the predictions. Eskelson et al. (2012) found that RF-NN produced greater accuracy than a parametric method, where accuracy was defined as the fraction of predictions within 30% of the true value. However, RF-NN also had a greater RMSE, revealing a tendency for k-NN to produce large over- and underpredictions across the range of the responses. Thus, there are

tradeoffs that need to be considered when an analytical approach is selected (Table 1).

In addition to weighing the relative importance of the different validation measures, statistical considerations such as model structure and characteristics of the data are important. Linear regression techniques have the most restrictive data assumptions, several of which may be violated for many important forest attributes. However, regression has the advantage that the relationship between predictors and response is fixed in the coefficients; the resulting models can be recycled or embedded in scalable systems that are not

Table 1. Advantages and disadvantages associated with methods for modeling and predicting forest inventory attributes.

Method	Advantages	Disadvantages
Regression	Most techniques are familiar to analysts and users Robust predictions when assumptions are valid Efficient variable selection Extendable across sites with similar conditions Easy to implement/included in most statistical packages	Assumptions are often violated Linear variable relationships are usually assumed Collinearity among predictors complicates model building Cannot easily handle missing values Does not explicitly consider the spatial structure of the data Variable interactions must be prespecified
CART	Distribution free Nonlinear variable relationships allowed Relatively unaffected by multicollinearity and outliers Easily handles both categorical and continuous data Relatively simple to implement and interpret Can handle variable interactions without prespecification	Model structures of a single decision tree can be unstable given small changes to input data Some algorithms suffer from selection bias Does not explicitly consider the spatial structure of the data
MARS	Better at detecting global and linear data structure Models nonlinearity and interactions among variables Computationally efficient Both continuous and categorical predictors	Model tuning is cumbersome Extensively guided by local nature of data, predictions with new data are unstable
ANNs	Distribution free Underlying processes do not need to be defined/known Easily handles nonlinear data Can effectively handle very complex problems Collinearity among predictors is not a problem Effective with multisource data	"Black box" approach Steep learning curve Risk of overfitting Determination of optimal parameters can be difficult/time consuming (trial and error) Greater number of predictors may lead to poorer predictions Can produce values outside the data range Does not consider the spatial structure of the data
RF	Retains the benefits of CART and addresses many limitations Reduces tendency to overfitting and bias compared to CART Can estimate variable importance from cross-validation Can be combined with nearest neighbors (i.e., RF imputation)	Less interpretable than CART because many trees are averaged Potentially sensitive to the inclusion of too many predictors Careful variable selection may be important Does not explicitly consider the spatial structure of the data
NN	Distribution free Concept of "nearest neighbor" is intuitive Easy to implement Variants allow explicit incorporation of the relationships between predictors and response into the weights and distance measures Easily handles multivariate responses	Can be time consuming (search process) with large data sets Selection of appropriate parameters (k) is not always straightforward; may require trial and error Does not explicitly consider the spatial structure of the data without using spatial distance-based weight

dependent on context-specific samples, as in k-NN. This property is essential in common forest inventory applications, such as individual tree volume and allometric biomass equations (e.g., Jenkins et al. 2003), which, once developed, can be used in many subsequent contexts. ANNs are free of distributional assumptions and do not require knowledge or definition of any underlying processes, but as a black box approach, the relationships underlying the predictions are not known, so interpretation of the relationships is not possible. Decision trees and nearest neighbor methods are also distribution free and have several other advantages, but disadvantages may include selection bias or difficulty in selecting appropriate model parameters.

Consideration also should be given to whether the analysis will predict univariate or multivariate responses. Many studies have taken a univariate approach to predicting forest composition and structure at both large and small scales. For example, Iverson and Prasad (1998) used a number of climatic, soil, land cover, and topographic variables to individually predict the distributions of 80 tree species at the county level in the eastern United States with reasonable success for most species. However, although the univariate prediction approach was appropriate for the coarse resolution examined in their study, it could lead to biologically unrealistic predictions (i.e., unrealistic species combinations) at the finer resolutions typically required for operational use, suggesting that a multivariate approach might be the better option when composition for direct applications is mapped (Moeur and Stage 1995). Likewise, forest

structural variables that are modeled independently could lead to predicted conditions that are biologically impossible, such as positive basal area predictions in areas of no forest cover. Although all of the methods discussed in this article can be adapted to handle multivariate problems, nearest neighbor methods probably handle these most easily and may also help preserve the covariance structure in the data when $k = 1$, if that is a desirable outcome.

Practical considerations related to computational time, resources, and analyst experience also play a role in method selection. For example, Moisen and Frescino (2002) compared five approaches with forest attribute imputation, including ANNs, geostatistical methods, CART, general additive models, and MARS; the authors found comparable accuracies among all five techniques when applied to real data sets, but computational time for the ANN was slowest and was described as having the "potential to be cripplingly slow" for some parameter optimization procedures. In the past decade, advances in computer technology and algorithm improvements (e.g., Finley and McRoberts 2008) have decreased the computational time required for many procedures such that it may not be the constraint it once was, but it can still be a relevant consideration. In addition, some techniques such as ANNs can be difficult to implement and understand, requiring a relatively steep learning curve, whereas others such as regression, decision trees, and nearest neighbor methods may be much more intuitive and easily implemented through widely available software.

Variable Selection

Useful models of forest inventory variables include predictor variables that are functionally or statistically related to the response(s) and discard those that are irrelevant or redundant. Clearly, distinguishing between these classes is essential when the purpose of modeling is to understand the underlying system (Rykiel 1996, Froese and Robinson 2007). However, the inclusion of irrelevant or redundant (correlated) variables can also have negative effects when models are developed for prediction. Under classic OLS, models remain unbiased even if irrelevant variables are included, but correlated variables can render models sensitive to the particular sample of calibration data. Many other estimation models have been shown to be sensitive to irrelevant or redundant predictor variables as well (Packalén and Maltamo 2007, Niska et al. 2010, Banskota et al. 2013a). For example, in a k-NN imputation approach, McRoberts et al. (2002) found that the inclusion of covariates that were unrelated to the response variables decreased the overall prediction accuracy, and Banskota et al. (2013b) and Pal and Foody (2010) demonstrated the necessity of conducting an analysis of variable importance before OLS regression and SVM classification, respectively. Even approaches such as RF, which can deal with high-dimension data sets, are negatively affected by the inclusion of redundant variables in the same prediction model (Evans and Cushman 2009, Falkowski et al. 2010, Murphy et al. 2010). Variable selection can be particularly challenging when modeling forest inventory attributes, because there are typically a large number of readily available covariates in the form of remotely sensed and geospatial data.

Variable selection can be made on the basis of previous studies, through the implementation of a variable or model selection algorithm (Packalén and Maltamo 2007, Hudak et al. 2008, Lister 2009) or by conducting feature extraction techniques such as principal components analysis (PCA) or CCA (Lefsky et al. 2005a). Feature extraction techniques transform the original, high-dimension variable space into a subspace with a lower dimensionality, while preserving a high degree of the original variation (Sotoca et al. 2007). Several feature extraction techniques such as PCA (Pu and Gong 2004), singular value decomposition (Philips et al. 2009), and maximum noise fraction (Green et al. 1988) have been used in forest inventory and assessment applications. However, the use of linear combinations of predictors in approaches such as PCA may limit model interpretability because the individual predictors are combined in composite variables (Cohen et al. 2003). If the goal of the modeling effort is attaining accurate predictions rather than clear interpretability of the relationships, then techniques such as PCA might be useful. Another major limitation of these techniques is that transformations are carried out independent of response variables and are based on the global covariance matrix, which can blur distinguishable features related to the response variables (Miao et al. 2007). Conversely, CCA transforms independent variables in a manner that retains their covariance structure, which is useful in the prediction of the response variables (Lefsky et al. 2005a). The wavelet transform, which provides a multiscale representation of the original data, has been used for feature extraction in forestry applications, including forest canopy structure identification (Bradshaw and Spies 1992), crown closure mapping (Pu and Gong 2004), tropical species identification (Zhang et al. 2006), and pine species classification (Banskota et al. 2011). However, as with other feature extraction techniques, wavelet analysis lacks an automated method to select features and still needs a suitable technique for selecting a useful subset of the transformed variables.

Most studies use some form of variable or model selection algorithm to choose predictor variables. Stepwise variable selection is commonly used in OLS regression (e.g., Lefsky et al. 2001, Hudak et al. 2006, Næsset 2007). Several stepwise procedures (e.g., forward, backward, and mixed) and criteria (e.g., F-test, Akaike information criterion, Bayesian information criterion, and Mallows' Cp) can be used to select variables (Sokal and Rohlf 2012). Alternatively, the best subset techniques exhaustively compare all possible subsets and return the best subset of variables according to some user-defined criterion. For example, Hudak et al. (2006) used stepwise multiple regression and best subsets regression to predict basal area and tree density from multispectral satellite and airborne LiDAR data. Decision tree or ensemble algorithms such as RF have a built-in ability to identify relevant predictor variables based on variable importance measures such as "permutation importance" or "Gini importance" (Strobl et al. 2008). Although RF is considered to be insensitive to the number of predictor variables, removal of irrelevant variables can improve the performance of the algorithm on retraining (Svetnik et al. 2003). For example, both Hudak et al. (2008) and Falkowski et al. (2010) used similar iterative stepwise looping procedures that are analogous to backwards stepwise multiple regression to identify candidate predictor variables in RFs. Packalén et al. (2012) term this "variable selection by random forests." The RF method can also potentially act as a variable selection tool for other classifiers or regression techniques (Hudak et al. 2008, McInerney et al. 2010).

On performing variable selection, some studies have indicated the relative performances of each predictor or explored the reasons for particular variables being identified as important. This information is potentially very valuable for future efforts to model forest inventory data. Although relationships and important predictors may vary, depending on geographic region or scale, it is useful to have a starting point that could reduce time and effort in future model development and that could potentially (if consistencies are found) form the framework of a theory. Furthermore, some studies have suggested that variable selection may be more important than model selection when the goal is predictive accuracy. For example, Hudak et al. (2008) found that variable selection using RF improved the model fit for every nearest neighbor distance metric they compared. Packalén et al. (2012) found that optimal variable selection always improved model accuracy and in most cases was more important than their choice of nearest neighbor method.

Selecting Responses and Data Sets

Spatial forest inventory modeling can be complex and resource consuming. Informed selection of appropriate predictor variables a priori could increase the efficiency of the modeling process and the accuracy, utility, and effectiveness of the resulting model and map (Niska et al. 2010). We include here a summary of variables deemed important enough to have been included in many models and try to indicate when certain variables have been specifically determined to be more or less influential for the prediction of specific forest inventory attributes (also see Table 2).

Reference and Feature Data

National forest inventories (NFIs) have been used as sources for reference data in forest attribute mapping efforts worldwide (e.g., Tomppo et al. 1999, 2008, McRoberts et al. 2002, Gjertsen 2007, Hirata et al. 2009, McRoberts 2009b, Teissier du Cros and Vidal

Table 2. Common predictors of forest compositional and structural inventory attributes.

Response	Frequent predictors	References
Composition		
Forest/vegetation type	Ancillary variables[1] Multispectral imageries and/or derivatives[2] Hyperspectral imageries and/or derivatives[3]	Franklin (1986),[2] Wolter et al. (1995),[2] Blackard and Dean (1999)[1], Moisen and Frescino (2002),[1,2] Ruefenacht et al. (2008),[1,2] Grossman et al. (2010)[1,2]
Species composition	Ancillary variables[1] Multispectral imageries and/or derivatives[2] Hyperspectral imageries and/or derivatives[3] LiDAR metrics[4] High-resolution imageries[5]	Iverson and Prasad (1998),[1] Leckie et al. (2003),[4] Zhang et al. (2006),[2,4] Asner and Martin (2009),[3] Dong (2009),[4] Mora et al. (2010),[5] Banskota et al. (2011b),[3] Wilson et al. (2012)[1,2]
Structure		
Age/successional stage	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] High-resolution imageries[5]	Fiorella and Ripple (1993),[2] Jensen et al. (1999),[4] Moisen and Frescino (2002),[1,2] Lefsky et al. (2005b),[2,4] Chubey et al. (2006),[5] Falkowski et al. (2009a)[4]
Basal area	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] Inventory variables[6] Aerial photos[7]	Franklin (1986),[2] Means et al. (1999),[4] Franco-Lopez et al. (2001),[6] Ingram et al. (2005),[2] LeMay and Temesgen (2005),[1,6,7] Hudak et al. (2006),[2,4] Pierce et al. (2009),[1,2] Magnussen et al. (2010),[4] Hudak et al. (2008),[4] Falkowski et al. (2010)[4]
Canopy cover/crown closure	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] High-resolution imageries[5]	Means et al. (1999),[4] Falkowski et al. (2005),[2] Chubey et al. (2006),[5] Hall et al. (2006),[2] Falkowski et al. (2009a),[4] Pierce et al. (2009)[1,2]
Diameter	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] Inventory variables[6]	Leduc et al. (2001),[6] Moisen and Frescino (2002),[1,2] Næsset (2007),[4] Breidenbach et al. (2008),[6] Pierce et al. (2009),[1,2] Falkowski et al. (2010),[4] Vauhkonen et al. (2010)[4]
Density	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] Inventory variables[6] Aerial photos[7]	LeMay and Temesgen (2005),[1,6,7] Hudak et al. (2006),[2,4] Hudak et al. (2008),[4] Pierce et al. (2009),[1,2] Magnussen et al. (2010)[4]
Height	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] High-resolution imageries[5]	Means et al. (1999),[4] Watt et al. (2004),[4] Chubey et al. (2006),[5] Hall et al. (2006),[2] Hirata et al. (2009),[4] Maltamo et al. (2009),[4] Pierce et al. (2009),[1,2] Hill et al. (2010),[2,4] Magnussen et al. (2010),[4] Pascual et al. (2010),[2,4] Hummel et al. (2011),[4] le Maire et al. (2011),[1,2] Richardson and Moskal (2011)[4]
LAI	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] High-resolution imageries[5]	Turner et al. (1999),[2] Pocewicz et al. (2004),[2] Berterretche et al. (2005),[2] Banskota et al. (2013a),[3] Banskota et al. (2013b)[3]
Volume	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] Inventory variables[6]	Franco-Lopez et al. (2001),[6] Mäkelä and Pekkarinen (2004),[2] Temppo and Halme (2004),[2] Næsset et al. (2005),[2] Hall et al. (2006),[2] Chirici et al. (2008),[2] Næsset (2007),[4] Magnussen et al. (2010),[4] Niska et al. (2010),[4] Tonolli (2011)[4]
Biomass/AGB	Ancillary variables[1] Multispectral imageries and/or derivatives[2] LiDAR metrics[4] Synthetic aperture radar (SAR)[8]	Franklin (1986),[2] Means et al. (1999),[4] Foody et al. (2001),[2] Moisen and Frescino (2002),[1,2] Lefsky et al. (2005b),[2,4] Hall et al. (2006),[2] Blackard et al. (2008),[1,2] Næsset and Gobakken (2008),[4] Zheng et al. (2004),[2] Zhao et al. (2009),[4] Banskota et al. (2011b),[4,8] Goetz and Dubayah (2011),[4,8] Næsset et al. (2011)[4,8]

Ancillary variables refer to the variables related to climate (e.g., precipitation and temperature), topography (e.g., location and elevation and landform), soil, and land cover, individually or in combination. Image derivatives refer to the bands or features derived from different transformations (e.g., band ratios and textures). The superscripts in the cited references denote the corresponding predictor variables used in the studies.

2009). Reference data may also come from ground samples obtained from other government or industry-initiated forest inventories (e.g., Pierce et al. 2009, le Maire et al. 2011) or independent project-based samples (e.g., Zheng et al. 2004, Falkowski et al. 2010, Niska et al. 2010). Pierce et al. (2009) combined field inventory data from several federal sources to map wildland fuels and forest structure in temperate forests in an effort to provide needed information for fire risk assessment and fuel treatment planning. Usually, reference data come from field measurements of actual conditions, but in some instances they may be derived from other sources, especially when field measurements would be difficult or expensive to obtain. For example, indices (e.g., successional stage and defoliation) derived from very fine-resolution imagery or airborne LiDAR data have been used as sources for reference data (Schroeder et al. 2007, Eklundh et al. 2009, McRoberts et al. 2010a).

Using existing data from NFIs can reduce costs and help ensure rigor in sampling design and consistency in data collection. In many instances, the sampling schemes have been updated to include variables relevant to current international issues such as biodiversity, sustainability, acid deposition, and carbon accounting (McRoberts et al. 2010b). However, not all countries have NFIs, or if they do, they may contain insufficient data for some purposes. Further, forest inventories may be outdated and thus of less operational utility, especially in developing countries with limited financial resources. Depending on the purpose and desired map properties, there may also be problematic spatial or temporal mismatches, including global positioning system (GPS) position errors (McRoberts 2010). Confidentiality constraints that fuzz the precise spatial locations of the reference data can also prove a hindrance to fine-resolution analytical problems. For example, Prisley et al. (2009) found the

perturbed locations associated with plots of the USDA Forest Inventory and Analysis program (Bechtold and Patterson 2005) unsuitable for analyses involving topographic data at resolutions of 10–30 m.

Ancillary or "feature" data (sensu McRoberts 2012) are used as predictor variables and can originate from many sources, although the most common are from remote sensing. Other sources include field-based mapping efforts, such as soil surveys, and climate layers derived from interpolation of weather station data (e.g., Daly et al. 2008). Remotely sensed data are popular because they are spatially continuous, typically cheaper to obtain than field data and often have high temporal resolution. Passive optical remote sensors include Landsat, moderate resolution imaging spectroradiometer (MODIS), advanced very high resolution radiometer (AVHRR), and airborne visible/infrared imaging spectrometer (AVIRIS); because these have long histories and are widely available, they are especially useful for monitoring and offer unique possibilities for homogeneous global assessment systems. However, because passive sensors are not able to represent three-dimensional spatial patterns, they have historically been more useful for characterizing horizontal forest structure than vertical structure, especially in heterogeneous forests or those with high biomass (Drake et al. 2002, Lefsky et al. 2002, Powell et al. 2007). More recent methodological advances such as the use of multitemporal data in combination with canopy shadowing and light attenuation relationships have made optical data more practical for characterization of vertical forest structure (Wolter et al. 1995, Lefsky et al. 2001, Xiao et al. 2002, Powell et al. 2010). Some studies have demonstrated that predictions of forest attributes can be improved in heterogeneous forests by using multitemporal, multisensor, or ancillary data (e.g., Lefsky et al. 2001, Xiao et al. 2002, Sesnie et al. 2008, 2010, Helmer et al. 2012).

To achieve more detailed predictions of inventory attributes, higher resolution data sources may be needed. Very high spatial resolution data from optical airborne or satellite sensors or airborne LiDAR show promise for prediction of both stand- and tree-level attributes, and research for operational purposes is increasing (Hudak et al. 2008, Næsset and Gobakken 2008, Falkowski et al. 2009a, Gleason and Im 2011, Næsset et al. 2011, Gobakken et al. 2012). For example, Mora et al. (2010) used panchromatic QuickBird satellite data to identify individual tree crowns and identify the major tree species present in a boreal forest. Breidenbach and Astrup (2012) used a canopy height model derived from optical airborne imagery for small area estimation of forest volume, height, and biomass. Gobakken et al. (2012) concluded that the use of airborne LiDAR was feasible for regional prediction of aboveground biomass in Norway. LiDAR has some advantages; unlike passive sensors, LiDAR has the ability to penetrate vegetation and clouds. Airborne LiDAR has been implemented or studied for use in an operational inventory in diverse forest types across the globe (e.g., Parker and Evans 2004, Jensen et al. 2006, Tickle et al. 2006, Næsset 2007, Hilker et al. 2008, Stephens et al. 2012). Despite the potential of airborne LiDAR, cost and availability continue to limit its operational feasibility (Wulder et al. 2008, Wolter et al. 2009), although it has been found to be cost-effective in some studies (e.g., Eid et al. 2004, Hummel et al. 2011). LiDAR sampling, in contrast to wall-to-wall LiDAR surveying, could prove to be an effective compromise for some operational inventory needs (Wulder et al. 2012). In the future, further development of spaceborne LiDAR systems could ultimately provide systematically collected forest canopy measurements in support of operational forest inventory and management.

Because of the limitations associated with acquisition of airborne LiDAR data, it is of interest whether Landsat or other optical remote sensing data are useful for predicting forest attributes at fine resolutions, particularly in heterogeneous forests. Unfortunately, there have been few studies that address this topic. Generally, optical remote sensing data tend to become poor predictors of forest biomass and vertical structure after the canopy closes or when vegetation is dense (Foody et al. 2001). For example, Watt et al. (2004) concluded that airborne LiDAR metrics were able to provide accurate predictions of forest height even after canopy closure in densely stocked conifer stands, while metrics from Landsat ETM+ and high-resolution IKONOS imagery were not. Nevertheless, Foody et al. (2001) obtained biomass predictions that were strongly correlated with measured values using Landsat data in tropical forests; they emphasized the importance of using all the available spectral data instead of vegetation indices such as NDVI that only incorporate some of the spectral wavebands. le Maire et al. (2011) were able to develop useful nonlinear regression models for predicting stand height and volume in eucalyptus stands, by using a time series of MODIS reflectance data to predict age from NDVI, and then predicting height and volume from models incorporating age, NDVI, and bioclimatic variables as predictors. These studies suggest that optical data may be useful if care is taken to incorporate a time series of images and to use appropriate indices. Few studies directly assessed the ability of Landsat versus airborne LiDAR to predict forest attributes at relatively fine resolutions. Hill et al. (2010) found that certain Landsat spectral indices were strongly correlated with canopy height as determined from airborne LiDAR measurements in a tropical rainforest, suggesting that the Landsat data could be used to predict height over a large area. In a similar study, Pascual et al. (2010) also found relationships between Landsat indices and airborne LiDAR canopy height. Hummel et al. (2011) examined a wider variety of forest structural attributes in a heterogeneous temperate forest landscape and found that imputations from Landsat data were not able to provide stand-level predictions of height, basal area, and volume (i.e., merchantable and total cubic meters) that were similar to estimates obtained from field measurements, although biomass and tree density estimates did not differ from stand means calculated from field data (Hummel et al. 2011). Airborne LiDAR data, however, provided estimates that were more consistent with field measurements. This finding seems to suggest that it may be possible to use Landsat data for prediction of forest attributes at operational (stand) scales, at least for some attributes, but that airborne LiDAR data may be more consistently useful. However, if tree-level or very fine-resolution information is desired, Landsat data are generally not appropriate (Wulder et al. 2004).

Spatial Mismatch

Although registration errors are often not addressed explicitly, several studies have demonstrated that errors caused by poor georeferencing of remotely sensed imagery and positional errors associated with ground reference points may contribute substantially to prediction error (Næsset et al. 2011, Valbuena et al. 2011, Bright et al. 2012). McRoberts (2010) found that a mean registration error of 30 m (i.e., the width of one Landsat pixel) resulted in more than half of the subplots in his study being associated with the wrong pixel. Errors associated with GPS units that are used to locate ground reference points can have similar effects. The GPS units used by the

USDA Forest Inventory and Analysis program have been reported to have errors ranging from 5 to 20 m (Cooke 2000). The combined uncertainties in GPS locations, image coordinates, and geometric errors can result in greater discrepancies among satellite images, airborne LiDAR, and field plots, especially in steep, mountainous areas and high-density forests (Valbuena et al. 2011, Bright et al. 2012, Jung et al. 2013). In a simulation study motivated by prior work with Landsat data, Patterson and Williams (2003) found that small registration errors (a 1 or 2 pixel shift) induced minimal bias but dramatically inflated the variance of estimates of forest area. These types of errors could result in forest attribute maps that are of questionable utility in the originally intended contexts. However, few recommendations exist in the literature on how to compensate for or correct the errors caused by image-to-plot mismatch. One such remedy is reported by Halme and Tomppo (2001), who showed that reassigning the locations of field plots to Landsat pixels within a 7 × 7 pixel window, using a weighted function of correlations among field and spectral values, can produce substantial reductions in prediction RMSE of volume per unit area using the k-NN method.

Relatively small plot location errors probably do not lead to large prediction errors when integrated with coarser-resolution raster data sets, such as MODIS. However, the many sources of spatial mismatch do raise questions about the utility of combining field plots with very fine resolution remotely sensed data such as LiDAR in operational mapping efforts. Gobakken and Næsset (2009) examined the impact of interaction between plot position error and plot size on the accuracy of least-squares predictions of stand height, basal area, and volume and found that larger field plots provided greater accuracy for a given position error. They also observed that LiDAR metrics were less sensitive to variations in plot size and coregistration error in young, dense, and homogeneous canopies than in open and heterogeneous stands. Gobakken and Næsset (2009) recommended that for poor sites where there are normally few stems it is important to have large plots and accurate plot positions to obtain precise estimates; this is not likely to be difficult to attain as relatively precise GPS positions would be expected in such forests compared with dense forests. Similar results were obtained by Mauro et al. (2011) while evaluating the effect of positioning error on tree height distribution measurements carried out in circular plots. They found that GPS positional errors had a greater impact on LiDAR metrics when extracted from small plots (6 m radius) than from larger sized plots (10 m radius). Using a Monte Carlo approach, Frazer et al. (2011) further demonstrated that large plots were substantially more resilient to coregistration error than were small plots.

Response Variables

Forest inventories collect information on a multitude of variables, the number of which is constantly increasing as new uses are found. Whereas traditional inventory variables have been those relevant to timber management, such as composition, diameter, and height, an increasing focus on nontimber objectives has broadened the scope to include data useful for applications such as forest health monitoring, biodiversity conservation, and carbon accounting, simultaneously increasing the costs and time involved in collecting the data (McRoberts and Tomppo 2007). Here, we examine forest inventory variables frequently predicted via these mapping models and discuss the ancillary variables commonly found to be useful as predictors (Table 2).

Composition

Information on forest composition is basic to almost all operational forest management and planning applications. One of the main uses of remote sensing data in forestry has been to classify forest cover types (e.g., Rogan et al. 2008, Ruefenacht et al. 2008, Knorn et al. 2009, Mora et al. 2010) (Table 2). Traditionally, classification of forest composition has taken the form of relatively broad type classes. For example, Ruefenacht et al. (2008) classified 142 forest types across the United States using a large number of spatial predictors, including MODIS satellite data and derived indices, soils, climate, topographic, and ecoregion variables. Such broad forest or cover type classes may, however, obscure the heterogeneity in actual ground conditions. Some remotely sensed data collected by satellite and aerial platforms have proven to be of use in predicting forest composition at detailed resolutions or for distinguishing individual tree species distributions. For example, Mora et al. (2010) used tree crown metrics from panchromatic Quickbird-2 images to delineate the four major species in forest stands in a boreal forest. Wilson et al. (2012) mapped tree species distributions in continental and subtropical forests using MODIS, climate, topographic, and ecoregion variables.

Accurate prediction of species composition for complex or diverse stands tends to be more difficult than for more homogeneous stands (e.g., Temesgen et al. 2003, Pierce et al. 2009, Helmer et al. 2012). Coarser resolution sensors may not provide the spatial detail needed to capture compositional diversity; while the broad spectral differences between coniferous and deciduous species are generally clear, distinguishing species within these classes from each other is much more difficult given the spectral resolution of many sensor systems. Therefore, distinguishing attributes of forests with high species diversity may require high spatial and spectral resolution data sources. Indeed, recent work by Asner and Martin (2009) has highlighted the potential for use of airborne, high spatial resolution, hyperspectral imagery to map taxonomic diversity in a highly diverse tropical forest. Although such approaches have merit, the relatively high costs and complexities associated with data acquisition, processing, and analysis have limited the application of hyperspectral remote sensing in the field of forestry.

Structure

Descriptions of forest structure are necessary for strategic and tactical silvicultural, forest health, restoration, conservation, and reporting objectives, as well for assessment of economic and noneconomic forest values. Remotely sensed and ancillary data have been used to predict a large number of forest structural attributes (Table 2) at both the plot and stand levels, including diameter (Pierce et al. 2009, Falkowski et al. 2010), height (Hall et al. 2006, Pierce et al. 2009), age (Jensen et al. 1999, Moisen and Frescino 2002), tree density (Franco-Lopez et al. 2001, Ingram et al. 2005, Pierce et al. 2009), basal area (Ingram et al. 2005, Pierce et al. 2009, Falkowski et al. 2010), volume (Franco-Lopez et al. 2001, Reese et al. 2002, Mäkelä and Pekkarinen 2004, Chirici et al. 2008, Niska et al. 2010), biomass (Moisen and Frescino 2002, Labrecque et al. 2006, Powell et al. 2010, Tuominen et al. 2010), LAI (Turner et al. 1999, Berterretche et al. 2005), and crown cover or canopy closure (Moisen and Frescino 2002, Hall et al. 2006). In addition to attributes of interest for operational forest management, ecologically meaningful variables are often predicted, including snag density

(Martinuzzi et al. 2009, Eskelson et al. 2012), coarse woody debris (Pesonen et al. 2008, Hudak et al. 2012), tree size inequality (Valbuena et al. 2013), and tree competition (Pedersen et al. 2012), among others.

Many forest structure variables are highly correlated, e.g., age, height, diameter, and volume. Therefore, it makes sense, when possible, to use inventory variables to predict other inventory variables. For example, le Maire et al. (2011) found that age-related variables explained substantial variation in volume and height of eucalyptus plantations in southern Brazil, though addition of NDVI and climatic variables improved these values significantly. Leduc et al. (2001) used stand age, site index, height, and density to predict diameter distributions in subtropical conifer plantations. Although useful, the lack of continuous spatial coverage of the predictors themselves is a problem when one is attempting to use imputation or interpolation to create continuous predictive output layers of inventory attributes.

Fortunately, remotely sensed data can be used to predict many forest structural attributes (Table 2). Spectral data from passive sensors (e.g., Landsat) have often been used. For example, Pocewicz et al. (2004) found NDVIc (a mid-infrared corrected version of NDVI that has improved sensitivity to plant area and canopy closure) to be a strong predictor of LAI in complex temperate mountainous forests. Fiorella and Ripple (1993) found strong negative correlations between Landsat TM spectral bands and stand age in maritime coniferous forests, but also found that the structural index (TM 4/5 ratio) had the highest correlation with stand age. Similarly, Hall et al. (2006) found negative correlations between two structural variables (stand height and crown closure) and Landsat ETM+ spectral bands, with bands 4 (near infrared) and 5 (mid-infrared) being most highly correlated with height and crown closure. Ingram et al. (2005) found reflectance of Landsat bands 3, 4, 5, and 7 to be useful predictors of basal area in tropical forests, whereas NDVI was not helpful.

Airborne LiDAR metrics have become increasingly popular, particularly with the rapid development of the technology. In the boreal forest, Næsset (2007) found that LiDAR-derived height and canopy density metrics were effective predictors of the variability in ground-measured values of six structural variables at the stand level, including mean and dominant heights, diameter, tree density, basal area, and volume; they also found that the predictions obtained from the models using airborne LiDAR data did not differ significantly from ground-measured values, indicating the potential for operational use in forest inventories. Airborne LiDAR metrics are especially useful in predicting height of dominant trees (Hirata et al. 2009, Maltamo et al. 2009). For example, Richardson and Moskal (2011) found LiDAR metrics useful for predicting density of taller trees, but not shorter trees, which they suggested might be a limitation of the technology. Falkowski et al. (2009a) found mean vegetation height and canopy cover metrics derived from airborne LiDAR measurements to be especially useful in classifying forest structural stages, and Hudak et al. (2008) found topographic and canopy structure variables derived from LiDAR measurements to be helpful when modeling basal area and tree density. Lefsky et al. (2005a) demonstrated that the vast number of metrics derived from large footprint LiDAR data could be grouped into three nonredundant factors (mean height, cover or leaf index area, and height variability) related to unique components of forest structure.

Biomass

Forest biomass is of particular current interest because of its relationship with climate (i.e., carbon sequestration and emissions) and as a potential alternative fuel source (i.e., biofuels), although biomass inventory also aids understanding of forest productivity, nutrient allocation, and fuel accumulation (Brown et al. 1999, Zhang and Kondragunta 2006, Saatchi et al. 2007). Biomass estimates are particularly relevant for carbon accounting programs such as measuring, reporting, and verification systems under the auspices of the United Nations Framework Convention on Climate Change and in quantifying impacts of implementing activities to reduce emissions from deforestation and degradation in developing countries (Global Observations of Forest and Land Cover Dynamics 2012). These programs require field measurements, which can be difficult to obtain, especially for developing countries with limited resources. However, remote sensing can facilitate carbon monitoring and measurement efforts by helping to stratify land use types for better representation of field samples, by direct prediction of structural parameters or by modeling approaches that can be used to calculate biomass through allometric relationships (Goetz and Dubayah 2011). Although total carbon pools include several components (e.g., aboveground biomass, belowground biomass, dead wood, litter, and soil organic matter), aboveground biomass (AGB) is the most commonly and easily measured (Zheng et al. 2004, Hall et al. 2006, Blackard et al. 2008) and is often used to predict carbon stored in other pools.

The most common approach used for biomass involves deriving biomass values for reference data, usually from conventional field-based forest inventory programs and then developing prediction models that use remote sensing data as feature variables (e.g., Baccini et al. 2004, Houghton et al. 2007, Powell et al. 2007, Nelson et al. 2012). Studies have shown that forest complexity may affect biomass predictions from optical data. For example, using Landsat spectral data, Powell et al. (2007) were able to build stronger biomass models in younger, coniferous forests than in older, deciduous forests. Houghton et al. (2007) predicted and mapped biomass in boreal forests by developing a relationship between ground-measured wood volume and MODIS data; they found that spectral saturation was less influential than different forest definitions on variation in biomass predictions and suggested that use of finer-resolution data, multitemporal data, and ancillary data on climate, soils, and topography could help to improve biomass models in areas where forest structure or composition is less well-defined. Indeed, Baccini et al. (2004) found nonlinear relationships between aboveground biomass and MODIS spectral data, precipitation, and elevation; incorporating these variables into a biomass model gave good results. Hall et al. (2006) found that modeling biomass and volume using height and crown closure (predicted from Landsat ETM+ data) was better than direct prediction of biomass and volume from Landsat data, compared with field observations. Most biomass prediction studies examined here tended to truncate the distribution by underpredicting in areas of high biomass and overpredicting in areas of low biomass.

LiDAR metrics show great promise for modeling biomass components. In an early study, Drake et al. (2002) successfully used four LiDAR height and ground return metrics to predict AGB in dense tropical forests. Næsset and Gobakken (2008) modeled AGB and belowground biomass, respectively, in models using metrics derived from airborne LiDAR as predictors in a boreal forest. Nelson et al. (2012) augmented forest inventory ground measurements with

LiDAR-derived height and density to predict AGB and found that profiling LiDAR surveys could provide estimates comparable to ground measurements. Zhao et al. (2009) demonstrated an approach to scale-invariant biomass prediction using LiDAR-derived canopy height distributions and canopy height quantile functions.

Biomass is rarely, if ever, directly observed in the reference data; instead, biomass at the scale used in modeling is commonly derived from tree-level measurements using a variety of allometric approaches. Usually, whole tree or tree component biomass is estimated as a function of diameter, and plot-level unit area estimates are calculated by summing (e.g., Zheng et al. 2004, Hall et al. 2006, Houghton et al. 2007, Powell et al. 2007). Biomass components may be available in NFI databases with no special processing required before model development (e.g., Blackard et al. 2008). Alternatively, biomass can be derived directly at the stand level using volume-to-biomass multipliers (Baccini et al. 2004) or in a two-step process in which foliage is predicted from feature variables and biomass is derived using allometric relationships with foliage (Zhang and Kondragunta 2006). Zhao et al. (2009) note that the biomass prediction method is often sensitive to the scale of the original reference data or the feature data used in the model; they proposed a scale-invariant approach in which biomass is modeled as a function of canopy height distribution metrics from fine-scale LiDAR data. Of note, Zhao et al. (2009) derived the reference data for their model by aggregating tree-level biomass predictions from published allometric functions of diameter, and diameter was derived from fine-scale LiDAR data.

Faux Inventory or "Tree Lists"

Increasingly detailed forest inventory information is needed for many objectives. For example, technological advances have improved our ability to predict forest growth and dynamics through simulation models that require tree lists or stand tables for initialization (e.g., the Forest Vegetation Simulator) (Dixon 2002). Ultimately, detailed maps of forest structure and composition that provide tree lists allow the greatest flexibility for varied uses because they can be manipulated or aggregated as needed to suit the objectives of a particular analysis. Another advantage of predictive modeling of tree lists is that they are equivalent to predicting compatible multivariate inventory attributes; i.e., if tree lists are imputed as if they were plot-based inventory data ("faux inventory"), then derived attributes such as basal area, volume, and density are automatically compatible. Separate predictions of composition, structure, or biomass may not be required.

Tree lists can be approximated through diameter distributions. For example, Maltamo et al. (2007) modeled diameter distribution using a Weibull function and LiDAR data as predictor variables. However, distributions are approximations and are incomplete compared with inventory data that routinely include multiple attributes (e.g., diameter, height, species, and quality). An alternative is to use methods that impute entire tree lists. Temesgen et al. (2003) used MSN to predict tree lists for complex temperate forest stands with aerial data on species composition, crown closure, elevation, biogeoclimatic ecosystem classification zones, height, age, and site class as predictors. When subsequently used as inputs to a growth-and-yield simulator, the modeled tree lists produced average stand volume predictions that closely approximated the volume obtained using the actual tree list, although results were sometimes less accurate for individual stands. Wallerman and Holmgren (2007) used LiDAR and optical image data as predictors with a similar approach,

using MSN, and achieved sufficiently accurate estimates of volume and density in a boreal forest. Building on these results, Duvemo et al. (2007) evaluated the cost/benefit tradeoff for a strategy using field plots against imputation and concluded that, for some interest rate scenarios, imputation was the superior approach both in terms of cost and risk. More recently, Falkowski et al. (2010) predicted tree-level forest inventory data from LiDAR height and topographic metrics, achieving high accuracy overall although the error was much higher for inventory metrics that incorporated small trees. Advances in LiDAR technology may enable rapid advances in tree list modeling in the future; identification of individual tree crowns from high pulse density LiDAR data may be operationally feasible, and tree level variables may be determined directly from the LiDAR point cloud or based on various other LiDAR features that are extracted for the tree segments (Lindberg et al. 2010, 2013, Hyyppä et al. 2012, Vastaranta et al. 2012). Overstory trees will still occlude understory trees from view of the sensor, impeding the prediction of a full tree list, though some approaches have been proposed to address this issue (e.g., Maltamo et al. 2007, Breidenbach et al. 2010).

Model Validation and Assessment

Model development and validation are closely linked. Once a method and appropriate variables have been identified, the model can be run and its performance assessed using a number of techniques. Results from model testing can inform strategies to refine the initial model or replace it entirely; therefore, model building is often considered an iterative process (Rykiel 1996, Froese and Robinson 2007). Despite much agreement on the utility of model validation, the literature is replete with alternative approaches and arguments about their relative appropriateness. There is some consensus that validation can at least be divided into questions about realism and questions about predictive accuracy (Caswell 1976, Rykiel 1996, Robinson and Ek 2000, Froese and Robinson 2007). Especially with the latter, comparison of predictions with measurements of the target population is the usual basis for model validation.

Ideally, data-based model validation is done with independent data drawn from the population to which the model will be applied. Here we define independent as being entirely separate from the data used to train or calibrate the model. This is in contrast to cross-validation, or "data splitting," in which a single data set is split into two subsets, one used for training and the other for validation. Use of independent data has the advantage of generating estimates of accuracy that are free of both the model and the training data. The model is tested in the context of interest; i.e., on the population on which it is to be used, not the population from which it was constructed. If independent data are not available, cross-validation may be used instead. This approach has been argued to provide no information beyond fit and lack of fit statistics derived from the model fit, at least for regression models, and may merely demonstrate that the subsets come from the same population (Kozak and Kozak 2003). However, with cross-validation accuracy can be determined without making any assumptions, parametric or otherwise, about the distribution of model coefficients or the training data. This allows for model-free estimates of prediction accuracy and permits statistical inferences about accuracy to be made for nonstatistical models (Robinson and Ek 2000), which is particularly useful for ANN black boxes.

Sometimes splitting the observations into training and testing

groups is not feasible (i.e., very small data sets), and independent data are lacking. In that case, double cross-validation (also called *k*-fold cross-validation) may be used to estimate the error. Double cross-validation is an iterative procedure that involves dividing a data set into *k* subsets, training the model using all subsets but one, and then testing the model on the subset that was left out (Kozak and Kozak 2003). This process repeats until all subsets have been used once as the test set and then averages the errors obtained from the repeated tests. In this way, all of the observations are used to train the model and to test the model, maximizing the information gleaned from the data set. Double cross-validation has been shown to be nearly unbiased with respect to estimating error, although its error estimates can be highly variable (Efron and Tibshirani 1997). Bootstrapping, which is similar to cross-validation except that error is estimated from subsamples drawn with replacement from the original sample, has also been used widely to estimate error and can sometimes provide better error estimates than cross-validation, generally resulting in smaller variability although bias tends to be larger (Efron 1983, Efron and Tibshirani 1997, Lendasse et al. 2003). However, bootstrapping does not work equally well for all analytical methods; although it is a popular method of error estimation for ANNs, it works less well with decision trees (Breiman et al. 1984, Kohavi 1995).

Categorical Response

Among the scientific community, no broadly accepted standardized methods exist for accuracy assessment or reporting of categorical data, although agencies may develop standards for accuracy in certain contexts (Foody 2002). However, most studies report an overall accuracy measure such as the percentage of correctly classified cases that quantifies how the predictions as a whole compare with the observed data. Because overall accuracy gives no information on how error is distributed among the classes, measures of producer's and user's accuracies are also needed to provide a better assessment of the utility of the model/map for different purposes (Foody 2002, Fassnacht et al. 2006, McRoberts et al. 2010a). Producer's accuracy refers to errors of omission, revealing how many of the test observations of a particular class are labeled correctly, whereas user's accuracy refers to errors of commission, which reveal how many of the classified pixels in a given class are actually what they say they are. Reporting an error (or "confusion") matrix is generally advised to fully disclose relevant information. Although normalization of the error matrix has been a common practice for almost 30 years (Congalton et al. 1983), this practice has been criticized increasingly in recent years for interpretation problems and because bias in parameters estimated from the normalized matrix has been demonstrated (Stehman 2004, McRoberts et al. 2010a). Therefore, it is generally advised that the raw error matrix be reported (Foody 2002). There is no universal standard for deciding whether or not a classification can be considered "accurate enough," but some authors have provided guidance with respect to acceptable levels of accuracy (e.g., Thomlinson et al. 1999). Even so, whether or not a classification is accurate enough often depends on subjective decisions related to the prospective use of the map (Fassnacht et al. 2006).

The error matrix, as an accuracy measure, is not without criticisms (Foody 2002). It typically assumes that classes are discrete and that all cases (pixels, stands, and others) belong entirely in a single class. However, heterogeneity often exists at subpixel resolutions (i.e., mixed pixels), and sometimes a pixel could belong to more than one class. Modification of accuracy measures to include "fuzzy" class boundaries may help with this (Townsend 2000, Fassnacht et al. 2006). In addition, the severity of misclassification may differ among classes (e.g., a high misclassification rate may be less important among similar classes but extremely important among other classes). Thus, interpretation of the error matrix and derived accuracy measures is not always straightforward. Another perceived issue with the error matrix is that it does not take into account cases that have been allocated into the correct class purely by chance (Foody 2002). The κ coefficient (Cohen 1960) has become a popular statistic for estimating the difference between actual and chance agreement. Because variances can be calculated, differences between two κ coefficients can be tested statistically, although often the statistical assumptions (e.g., independence of the samples) may not be met (Foody 2002, 2004). Recent criticism of the κ coefficient suggests that the adjustment for chance agreement is conceptually flawed, and, furthermore, whether a classification is correct by chance or design is irrelevant and misleading to map users (Stehman and Foody 2009). In their evaluation of a wide variety of accuracy measures, Liu et al. (2007) found a high correlation between the κ value and overall accuracy, suggesting that κ was simply a downscaled version of overall accuracy; they recommended using overall accuracy as a primary measure and advised against using the κ coefficient despite its popularity. The error matrix is increasingly considered to be inadequate for full accuracy reporting, and current thinking emphasizes the need for techniques for constructing inferences in the form of confidence intervals (McRoberts 2010). Such techniques are of particular value for comparison and evaluation purposes as estimation protocols shift from plot-based methods to geospatial methods to accommodate operational needs and limitations.

Continuous Response

For continuous data, measures of precision (unsystematic error) and bias (systematic error) are considered essential for assessing the accuracy of modeled data. Most studies commonly report fit statistics, either the RMSE or root mean squared difference (RMSD) (Stage and Crookston 2007), sometimes augmented by other measures (e.g., R^2 or relative RMSE). RMSD is an analog to RMSE commonly reported for nearest neighbor models; the notational distinction is important because the RMSD includes different components of error than the usual RMSE reported in regression (for details, see Stage and Crookston 2007). None is an ideal measure, however.

R^2 is a measure of the variation explained by the least squares regression model but does not actually measure the difference between two data sets (Ji and Gallo 2006). RMSE is not standardized and therefore has units identical to the quantity being predicted, so it is often not comparable among different models (McRoberts 2009a, Riemann et al. 2010). Furthermore, RMSE is a metric of precision only for an unbiased estimator; otherwise, it incorporates both. It is important to keep in mind that both R^2 and RMSE have been used as fit statistics for trained models, but they can also be calculated for comparisons of predicted to measured values as a better measure of how well the model predicts new cases. Graphical plots of predicted versus actual values of a variable give a visual representation of bias, often clearly showing where over- or under-prediction occurred. Prediction bias is calculated by taking the average of the differences between the observed and predicted values. The bias measure can be tested statistically to determine whether it

is significantly different from zero, and, if so, under- or overprediction can be determined. For example, Niska et al. (2010) calculated a test statistic, tbias, as $bias/(s/n - 2)$, where s is the SD of the differences between the actual and predicted values. This statistic was used to determine whether significant bias was present in models developed to predict stem volumes at plot and stand levels.

Taken alone, these precision and bias measures are not sufficient to assess a modeled spatial data set. In addition to error associated strictly with the modeling process and variability in the data sets, uncertainty and error may result from several other factors, including spatial mismatches in data, registration, and other locational errors, spatial variation in error due to sensors, and uncertainty and error in the reference data (Foody 2002, Ji and Gallo 2006, Riemann et al. 2010). All of these error sources make accuracy assessments difficult to interpret, especially when only one or two indices are used. It is important to have several assessment measures because model or map choice may depend on different criteria for different uses (Foody 2002, Riemann et al. 2010). A map that is acceptable for one purpose may not be for another. For example, Powell et al. (2010) found that an RF model generally performed better than two other methods in terms of minimizing prediction error of biomass, but SMA regression and GNN did a better job at variance preservation. Thus, an effective assessment provides enough information to evaluate whether a modeled data set is appropriate for varied uses.

Riemann et al. (2010) proposed a potentially useful set of guidelines for assessing modeled output for continuous variables, summarizing what the authors considered to be the five characteristics of effective assessments: (1) use of multiple types of assessment, (2) a description of the characteristics relevant to how the data set will be used, (3) a description of the location of error, (4) assessment across a range of scales, and (5) timeliness and consistent application of assessments. Operationally, the Riemann et al. (2010) protocol proposes (1) comparing the empirical cumulative distribution functions of the observed and predicted data sets for several scales and testing them using the Kolmogorov-Smirnov test, (2) examining overall agreement across several scales using (i) 1:1 scatterplots and SMA regression lines, (ii) an agreement coefficient that is then recalculated for both systematic and unsystematic error based on the proximity of the SMA line to the 1:1 line and the degree of scatter around the SMA line, (iii) differences between means and RMSE, and (iv) maps of the plot-based and modeled estimates at different scales, (3) examining spatial and distributional patterns of local differences by comparing predictions versus plot-based confidence intervals and producing choropleth maps of predictions with respect to plot confidence intervals, and (4) examining local variability via choropleth maps of SD of modeled estimates.

In addition to the detailed maps and accuracy measures described above, equivalence tests have been proposed as a way to statistically test whether two data sets differ from each other (Robinson and Froese 2004, Robinson et al. 2005). Unlike traditional significance testing, which tests the null hypothesis of no differences between the two data sets, equivalence tests place the burden of proof on the model, testing the null hypothesis that the two data sets are different and thereby requiring evidence for model validation through rejection of the null hypothesis. Leites et al. (2009) used equivalence tests, among other evaluation criteria, to evaluate crown ratio models and their influence on projected stand variables in temperate conifer forests. Likewise, Falkowski et al. (2010) used equivalence tests to evaluate imputation models predicting tree-level inventory data against NFI data in the same forest type. They also assessed the

model predictions for operational accuracy by conducting equivalence tests on the projections from a forest growth model obtained using both the predicted and observed data for parameterization.

Conclusions

Predicting forest inventory attributes from ancillary remotely sensed or other data involves a number of important considerations at each step of the process (Figure 1). Each of the methods we reviewed has particular strengths and weaknesses, and no analytical technique examined emerged as superior for all cases. The choice of method should be determined carefully by identifying the specific goals of an analysis, the type of response required, the characteristics of the ancillary data, and the resources available. Many studies reach the conclusion that the approach chosen and the amount of attention dedicated to variable selection is probably more important than method selection to generate precise and useful results.

Identifying the purpose of the analysis is probably the single most important first step and will help rule out inappropriate analytical methods. For example, "black box" techniques such as ANN provide no information on relationships among variables, so if interpretation of such relationships is important, then ANN is probably not an appropriate analysis technique. Further, specific variants of nearest neighbor analysis may be better for preserving the covariance structure of the data, whereas other techniques might provide greater predictive accuracy. It is up to the user to decide which is most important. Moreover, analyst and computing resources are not a trivial consideration. Whereas regression, nearest neighbor, and decision trees are conceptually easy to understand and implement, ANN in particular is not straightforward and typically requires large amounts of computing time for implementation.

Careful consideration of the types and characteristics of the response and ancillary variables can also help determine the best analytical method for a particular problem. Even when multivariate responses are of interest, most previous studies have built individual models for each forest attribute response. However, more recent work is trending toward development of multiresponse models for several reasons, including maintenance of covariance structure and realistic outcomes as well as simply reducing the time and expense associated with the analysis phase. For example, multivariate imputation of forest inventory plot identification numbers that are linked to the inventory data through lookup tables permits prediction of any of the structural or compositional variables measured on the inventory plots by imputing a single plot identification. Because only a single model needs to be developed, this type of analysis reduces the work and resources involved in the prediction process. Thus, multivariate responses may be best suited to nearest neighbors analysis, whereas any of the methods might perform well with a univariate response.

Other characteristics of the data such as non-normality and nonlinear relationships might suggest use of approaches that do not rely on those assumptions, such as decision trees, nearest neighbors, or ANNs. If there are missing data, regression is not recommended (Table 1). Some techniques such as decision trees and nearest neighbors can readily incorporate several ancillary variables, as well as both categorical and continuous predictors, better than other techniques such as regression. Although regression can handle both categorical and continuous data, inclusion of categorical predictors can often be a tedious and inefficient task, especially if there are a large number of categories or classes.

A wide variety of remotely sensed and other ancillary variables

have been used to predict forest composition and structural attributes. Traditionally, species composition has been classified from the spectral bands of remote sensing images into broad cover type classes, but other data such as NDVI, land cover, topography, climate, soils, and tree crown metrics have also been commonly used (Table 2). Many studies have found measured forest inventory variables such as height or age to be useful in predicting other structural inventory attributes such as biomass or volume; however, for mapping purposes, the lack of spatially continuous coverage of the predictors is problematic, and remotely sensed data have proven to be valuable because of their spatial continuity. When prediction of individual tree variables is desired, metrics from active sensors such as LiDAR are being used increasingly and effectively for the purpose of providing information on vertical structure (e.g., canopy height) and horizontal structure (e.g., canopy cover and tree density) that can be further used in models predicting attributes such as volume or biomass. The use of LiDAR metrics, in particular, for predicting forest inventory attributes holds great promise for forest management and has been implemented in an operational context in some countries. Both compositional and structural attributes tend to be more difficult to predict accurately in complex stands than in simple stands, and various authors have suggested the use of fine-resolution, multitemporal, and ancillary data on climate, soils, and topography to help overcome this problem.

Method and variable selection are intertwined in the sense that characteristics of both the response and predictor variables help determine which method is most appropriate to use (Figure 1). A problematic trend apparent from our review is the lack of consistency in the literature with regard to testing and reporting the relative contributions of ancillary variables to models predicting forest attributes. If reported consistently, such information could be very useful for refining future modeling efforts and improving predictions as well as permitting generalizations and theoretical advances to be made. In many cases, variables were included with no apparent rationale other than availability, and no variable reduction procedure or other test of variable contributions was used, which impeded our efforts to compare and generalize. Therefore, we strongly recommend that future efforts report the relative contributions of the various predictor variables to the models that are developed. Formal study leveraging data fusion schemes may aid in standardization in the future.

Once a method has been selected and appropriate response and predictor variables have been identified, the formal model can be developed, tested, refined if necessary based on the results of the validation, and used to predict and map forest attributes across a spatially continuous area (Figure 1). Part of the difficulty in comparing different predictive techniques or even simply different data sets is the lack of a standardized validation protocol for prediction models. Most studies report RMSE, but this measure is not standardized, rendering it less useful for purposes of comparison outside a particular study. Other measures are reported sporadically but not consistently enough to be useful for comparison purposes. As discussed earlier, Riemann et al. (2010) proposed a comprehensive assessment protocol for prediction models using geospatial data that would, if adopted widely, provide valuable information to allow for broader comparisons and advances in model development, ultimately benefiting forest planning and operations as well as the wider forestry community.

Literature Cited

ASNER, G.P., AND R.E. MARTIN. 2009. Airborne spectranomics: Mapping canopy chemical and taxonomic diversity in tropical forests. *Front. Ecol. Environ.* 7:269–276.

ATKINSON, P.M., AND A.R.L. TATNALL. 1997. Introduction: Neural networks in remote sensing. *Int. J. Remote Sens.* 18:699–709.

BAATUUWIE, N.B., AND I.L. VAN LEEUWEN. 2011. Evaluation of three classifiers in mapping forest stand types using medium resolution imagery: A case study in the Offinso Forest District, Ghana. *Afr. J. Environ. Sci. Technol.* 5:25–36.

BACCINI, A., M.A. FRIEDL, C.E. WOODCOCK, AND R. WARBINGTON. 2004. Forest biomass estimation over regional scales using multisource data. *Geophys. Res Lett.* 31:L10501.

BANSKOTA, A., R.H. WYNNE, P. JOHNSON, AND B. EMESSIENE. 2011. Synergistic use of very high-frequency radar and discrete-return lidar for estimating biomass in temperate hardwood and mixed forests. *Ann. For. Sci.* 68:347–356.

BANSKOTA, A., R.H WYNNE, S.P. SERBIN, N. KAYASTHA, V.A. THOMAS, AND P.A. TOWNSEND. 2013a. Utility of the wavelet transform for LAI Estimation using hyperspectral data. *Photogramm. Eng. Remote Sens.* 79:653–662.

BANSKOTA, A., R.H. WYNNE, V.A. THOMAS, S.P. SERBIN, N. KAYASTHA, J.P. GASTELLU-ETCHEGORRY, AND P.A. TOWNSEND. 2013b. Investigating the utility of wavelet transforms for inverting a 3-D radiative transfer model using hyperspectral data to retrieve forest LAI. *Rem. Sens.* 5:2639–2659.

BECHTOLD, W.A., AND P.L. PATTERSON. 2005. *The enhanced Forest Inventory and Analysis program: National sampling design and estimation procedures.* USDA For. Serv., Gen. Tech. Rep. SRS-80, Southern Research Station, Asheville, NC. 85 p.

BERTERRETCHE, M., A.T. HUDAK, W.B. COHEN, T.K. MAIERSPERGER, S.T. GOWER, AND J. DUNGAN. 2005. Comparison of regression and geostatistical methods for mapping leaf area index (LAI) with Landsat ETM+ data over a boreal forest. *Rem. Sens. Environ.* 96:49–61.

BLACKARD, J.A., AND D.J. DEAN. 1999. Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. *Comput. Electron. Agr.* 24:131–151.

BLACKARD, J.A., M.V. FINCO, E.H. HELMER, G.R. HOLDEN, M.L. HOPPUS, D.M. JACOBS, A.J. LISTER, ET AL. 2008. Mapping US forest biomass using nationwide forest inventory data and moderate resolution information. *Rem. Sens. Environ.* 112:1658–1677.

BRADSHAW, G.A., AND T.A. SPIES. 1992. Characterizing canopy gap structure in forests using wavelet analysis. *J. Ecol.* 80:205–215.

BREIDENBACH, J., AND R. ASTRUP. 2012. Small area estimation of forest attributes in the Norwegian national forest inventory. *Eur. J. For. Res.* 131:1255–1267.

BREIDENBACH, J., C. GLÄSER, AND M. SCHMIDT. 2008. Estimation of diameter distributions by means of airborne laser scanner data. *Can. J. For. Res.* 38:1611–1620.

BREIDENBACH, J., E. NÆSSET, V. LIEN, T. GOBAKKEN, AND S. SOLBERG. 2010. Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Rem. Sens. Environ.* 114:911–924.

BREIMAN, L. 2001. Random forests. *Machine Learn.* 45:5–32.

BREIMAN, L., J.H. FRIEDMAN, R.A. OLSHEN, AND C.J. STONE. 1984. *Classification and regression trees.* Chapman & Hall, New York. 368 p.

BRIGHT, B.C., J.A. HICKE, AND A.T. HUDAK. 2012. Estimating aboveground carbon stocks of a forest affected by mountain pine beetle in Idaho using lidar and multispectral imagery. *Rem. Sens. Environ.* 124:270–281.

BROWN, S., A.J.R. GILLESPIE, AND A.E. LUGO. 1989. Biomass estimation

methods for tropical forests with applications to forest inventory data. *For. Sci.* 35:881–902.

BROWN, S.L., P. SCHROEDER, AND J.S. KERN. 1999. Spatial distribution of biomass in forests of the eastern USA. *For. Ecol. Manage.* 123:81–90.

CAIRNS, D.M. 2001. A comparison of methods for predicting vegetation type. *Plant Ecol.* 156:3–18.

CAMPBELL, N., K. MACKENZIE, A.F. ZUUR, E.N. IENO, AND G.M. SMITH. 2007. Fish stock identification through neural network analysis of parasite fauna. P. 449–462 in *Analysing ecological data*, Zuur, A.F., E.N. Ieno, and G.M. Smith (eds.). Springer, New York.

CARREIRAS, J.M.B., M.J. VASCONCELOS, AND R.M. LUCAS. 2012. Understanding the relationship between aboveground biomass and ALOS PALSAR data in the forests of Guinea-Bissau (West Africa). *Rem. Sens. Environ.* 121:426–442.

CARTWRIGHT, H.M. 2008. Artificial neural networks in biology and chemistry–The evolution of a new analytical tool. *Methods Mol. Biol.* 458:1–13.

CASWELL, H. 1976. The validation problem. P. 313–325 in *Systems analysis and simulation in ecology*, vol. IV, Patten, B. (ed.). Academic Press, New York.

CHIRICI, G., A. BARBATI, P. CORONA, M. MARCHETTI, D. TRAVAGLINI, F. MASELLI, AND R. BERTINI. 2008. Non-parametric and parametric methods using satellite images for estimating growing stock volume in alpine and Mediterranean forest ecosystems. *Rem. Sens. Environ.* 112:2686–2700.

CHUBEY, M.S., S.E. FRANKLIN, AND M.A. WULDER. 2006. Object-based analysis of Ikonos-2 imagery for extraction of forest inventory parameters. *Photogramm. Eng. Rem. Sens.* 72:383–394.

COHEN, J. 1960. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20:37–46.

COHEN, W.B., T.K. MAIERSPERGER, S.T. GOWER, AND D.P. TURNER. 2003. An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Rem. Sens. Environ.* 84:561–571.

CONGALTON, R.G., R.G. ODERWALD, AND R.A. MEAD. 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogramm. Eng. Rem. Sens.* 49:1671–1678.

COOKE, W.H. 2000. Forest/non-forest stratification in Georgia with Landsat Thematic Mapper data. P. 28–30 in *Proc. of the First annual Forest Inventory and Analysis symposium*, McRoberts, R.E., G.A. Reams, and P.C. Van Deusen (eds.). USDA For. Serv., Gen. Tech. Rep. NC-213, North Central Research Station, St. Paul, MN.

COOPS, N.C., M.L. SMITH, M.E. MARTIN, AND S.V. OLLINGER. 2003. Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data. *IEEE Trans. Geosci. Rem. Sens.* 41:1338–1346.

CORONA, P. 2010. Integration of forest mapping and inventory to support forest management. *iForest* 3:59–64.

CROOKSTON, N.L., AND A.O. FINLEY. 2008. yaImpute: An R package for *k*NN imputation. *J. Stat. Softw.* 23:1–16.

CURRAN, P.J., AND A.M. HAY. 1986. The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogramm. Eng. Rem. Sens.* 52:229–241.

CUTLER, D.R., T.C. EDWARDS JR., K.H. BEARD, A. CUTLER, K.T. HESS, J. GIBSON, AND J.J. LAWLER. 2007. Random forests for classification in ecology. *Ecology* 88:2783–2792.

DALY, C., M. HALBLEIB, J.I. SMITH, W.P. GIBSON, M.K. DOGGETT, G.H. TAYLOR, J. CURTIS, AND P.P. PASTERIS. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.* 28:2031–2064.

DE'ATH, G. 2002. Multivariate regression trees: A new technique for modeling species- environment relationships. *Ecology* 83:1105–1117.

DIXON, G.E. (COMP.). 2002. *Essential FVS: A user's guide to the Forest Vegetation Simulator.* USDA For. Serv., Internal Report, Fort Collins, CO. 226 p.

DONG, P. 2009. Characterization of individual tree crowns using three-di-

mensional shape signatures derived from LiDAR data. *Int. J. Rem. Sens.* 30:6621–6628.

DRAKE, J.B., R.O. DUBAYAH, D.B. CLARK, R.G. KNOX, J.B. BLAIR, M.A. HOFTON, R.L. CHAZDON, J.F. WEISHAMPEL, AND S. PRINCE. 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. *Rem. Sens. Environ.* 79:305–319.

DUVEMO, K., A. BARTH, AND J. WALLERMAN. 2007. Evaluating sample plot imputation techniques as input in forest management planning. *Can. J. For. Res.* 37:2069–2079.

EFRON, B. 1983. Estimating the error rate of a prediction rule: Improvements on cross-validation. *J. Am. Stat. Assoc.* 78:316–331.

EFRON, B., AND R. TIBSHIRANI. 1997. Improvements on cross-validation: The .632+ bootstrap method. *J. Am. Stat. Assoc.* 92:548–560.

EID, T., T. GOBAKKEN, AND E. NÆSSET. 2004. Comparing stand inventories based on photo interpretation and laser scanning by means of cost-plus-loss analyses. *Scand. J. For. Res.* 19:512–523.

EKLUNDH, L., T. JOHANSSON, AND S. SOLBERG. 2009. Mapping insect defoliation in Scots pine with MODIS time-series data. *Rem. Sens. Environ.* 113:1566–1573.

ESKELSON, B.N.I., H. TEMESGEN, AND T.M. BARRETT. 2008. Comparison of stratified and non- stratified most similar neighbour imputation for estimating stand tables. *Forestry* 81:125–134.

ESKELSON, B.N.I., H. TEMESGEN, AND J.C. HAGAR. 2012. A comparison of selected parametric and imputation methods for estimating snag density and snag quality attributes. *For. Ecol. Manage.* 272:26–34.

ESKELSON, B.N.I., H. TEMESGEN, V. LEMAY, T.M. BARRETT, N.L. CROOKSTON, AND A.T. HUDAK. 2009. The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scand. J. For. Res.* 24:235–246.

EVANS, J.S., AND S.A. CUSHMAN. 2009. Gradient modeling of conifer species using random forests. *Landsc. Ecol.* 24:673–683.

FALKOWSKI, M.J., J.S. EVANS, S. MARTINUZZI, P.E. GESSLER, AND A.T. HUDAK. 2009a. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Rem. Sens. Environ.* 113:946–956.

FALKOWSKI, M.J., P.E. GESSLER, P. MORGAN, A.T. HUDAK, AND A.M.S. SMITH. 2005. Characterizing and mapping forest fire fuels using ASTER imagery and gradient modeling. *For. Ecol. Manage.* 217:129–146.

FALKOWSKI, M.J., A.T. HUDAK, N.L. CROOKSTON, P.E. GESSLER, E.H. UEBLER, AND A.M.S. SMITH. 2010. Landscape-scale parameterization of a tree-level forest growth model: A *k*-nearest neighbor imputation approach incorporating LiDAR data. *Can. J. For. Res.* 40:184–199.

FALKOWSKI, M.J., M.A. WULDER, J.C. WHITE, AND M.D. GILLIS. 2009b. Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery. *Prog. Phys. Geog.* 33:403–423.

FASSNACHT, K.S., W.B. COHEN, AND T.A. SPIES. 2006. Key issues in making and using satellite-based maps in ecology: A primer. *For. Ecol. Manage.* 222:167–181.

FEHRMANN, L., A. LEHTONEN, C. KLEINN, AND E. TOMPPO. 2008. Comparison of linear and mixed-effect regression models and a *k*-nearest neighbour approach for estimation of single-tree biomass. *Can. J. For. Res.* 38:1–9.

FINLEY, A.O., AND R.E. MCROBERTS. 2008. Efficient *k*-nearest neighbor searches for multi-source forest attribute mapping. *Rem. Sens. Environ.* 112:2203–2211.

FIORELLA, M., AND W.J. RIPPLE. 1993. Analysis of conifer forest regeneration using Landsat Thematic Mapper data. *Photogramm. Eng. Rem. Sensing.* 59:1383–1388.

FITZGERALD, R.W., AND B.G. LEES. 1992. The application of neural networks to the floristic classification of remote sensing and GIS data in complex terrain. P. 570–574 in *Proc. of the XVII Congress ISPRS.* American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

FOODY, G.M. 2002. Status of land cover classification accuracy assessment. *Rem. Sens. Environ.* 80:185–201.

FOODY, G.M. 2004. Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogramm. Eng. Rem. Sens.* 70:627–633.

FOODY, G.M., D.S. BOYD, AND M.E.J. CUTLER. 2003. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Rem. Sens. Environ.* 85:463–474.

FOODY, G.M., M.E. CUTLER, J. MCMORROW, D. PELZ, H. TANGKI, D.S. BOYD, AND I. DOUGLAS. 2001. Mapping the biomass of Bornean tropical rain forest from remotely sensed data. *Global Ecol. Biogeogr.* 10:379–387.

FRANCO-LOPEZ, H., A.R. EK, AND M.E. BAUER. 2001. Estimation and mapping of forest stand density, volume, and cover type using the *k*-nearest neighbors method. *Rem. Sens. Environ.* 77:251–274.

FRANKLIN, J. 1986. Thematic mapper analysis of coniferous forest structure and composition. *Int. J. Rem. Sens.* 7:1287–1302.

FRAZER, G.W., S. MAGNUSSEN, M.A. WULDER, AND K.O. NIEMANN. 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Rem. Sens. Environ.* 115:636–649.

FRIEDMAN, J.H. 1991. Multivariate adaptive regression splines. *Ann. Stat.* 19:1–67.

FRIEDMAN, J.H. 2002. Stochastic gradient boosting. *Comp. Stat. Data Anal.* 38:367–378.

FROESE, R.E., AND A.P. ROBINSON. 2007. A validation and evaluation of the Prognosis individual-tree basal area increment model. *Can. J. For. Res.* 37:1438–1449.

FROESE, R.E., D.R. SHONNARD, C.A. MILLER, K.P. KOERS, AND D.M. JOHNSON. 2010. An evaluation of greenhouse gas mitigation options for coal-fired power plants in the US Great Lakes States. *Biomass Bioenerg.* 34:251–262.

GIBBS, H.K., S. BROWN, J.O. NILES, AND J.A. FOLEY. 2007. Monitoring and estimating tropical forest carbon stocks: Making REDD a reality. *Environ. Res. Lett.* 2:045023.

GISLASON, P.O., J.A. BENEDIKTSSON, AND J.R. SVEINSSON. 2006. Random forests for land cover classification. *Pattern Recogn. Lett.* 27:294–300.

GJERTSEN, A.K. 2007. Accuracy of forest mapping based on Landsat TM data and a kNN-based method. *Rem. Sens. Environ.* 110:420-430.

GLEASON, C.J., AND J. IM. 2011. A review of remote sensing of forest biomass and biofuel: Options for small-area applications. *GISci. Rem. Sens.* 48:141–170.

GOBAKKEN, T., AND E. NÆSSET. 2009. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* 39:1036–1052.

GOBAKKEN, T., E. NÆSSET, R. NELSON, O.M. BOLLANDSÅS, T.G. GREGOIRE, G. STÅHL, S. HOLM, H.O. ØRKA, AND R. ASTRUP. 2012. Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. *Rem. Sens. Environ.* 123:443–456.

GOERNDT, M.E., V.J. MONLEON, AND H. TEMESGEN. 2011. A comparison of small-area estimation techniques to estimate selected stand attributes using LiDAR-derived auxiliary variables. *Can. J. For. Res.* 41:1189–1201.

GOETZ, S., AND R. DUBAYAH. 2011. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Manage.* 2:231–244.

GLOBAL OBSERVATIONS OF FOREST AND LAND COVER DYNAMICS. 2012. *A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals associated with deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation.* GOFC-GOLD Rep. Version COP18-1, GOFC-GOLD

Land Cover Project Office, Wageningen Univ., Wageningen, The Netherlands. 219 p.

GREEN, A.A., M. BERMAN, P. SWITZER, AND M.D. CRAIG. 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Rem. Sens.* 26:65–74.

GROSSMAN, E., J. OHMANN, J. KAGAN, H. MAY, AND M. GREGORY. 2010. Mapping ecological systems with a random forest model: Tradeoffs between errors and bias. *Gap Anal. Bull.* 17:16–22.

GUISAN, A., N.E. ZIMMERMANN, J. ELITH, C.H. GRAHAM, S. PHILLIPS, AND A.T. PETERSON. 2007. What matters for predicting the occurrences of trees: Techniques, data, or species' characteristics? *Ecol. Monogr.* 77:615–630.

HAGNER, O., AND H. REESE. 2007. A method for calibrated maximum likelihood classification of forest types. *Rem. Sens. Environ.* 11:438–444.

HALL, R.J., R.S. SKAKUN, E.J. ARSENAULT, AND B.S. CASE. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *For. Ecol. Manage.* 225:378–390.

HALME, M., AND E. TOMPPO. 2001. Improving the accuracy of multisource forest inventory estimates to reducing plot location error—A multicriteria approach. *Rem. Sens. Environ.* 78:321–327.

HASTIE, T., R. TIBSHIRANI, AND J. FRIEDMAN. 2009. *The elements of statistical learning: Prediction, inference and data mining*, 2nd ed. Springer Verlag, New York. 763 p.

HE, H., AND E.A. GARCIA. 2009. Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* 21:1263–1284.

HELMER, E.H., T.S. RUZYCKI, J. BENNER, S.M. VOGGESSER, B.P. SCOBIE, C. PARK, D.W. FANNING, AND S. RAMNARINE. 2012. Detailed maps of tropical forest types are within reach: Forest tree communities for Trinidad and Tobago mapped with multiseason Landsat and multiseason fine-resolution imagery. *For. Ecol. Manage.* 279:147–166.

HELMER, E.H., T.S. RUZYCKI, J.M. WUNDERLE JR., S. VOGESSER, B. RUEFENACHT, C. KWIT, T.J. BRANDEIS, AND D.N. EWERT. 2010. Mapping tropical dry forest height, foliage height profiles and disturbance type and age with a time series of cloud-cleared Landsat and ALI image mosaics to characterize avian habitat. *Rem. Sens. Environ.* 114:2457–2473.

HENAREH KHALYANI, A., M.J. FALKOWSKI, AND A.L. MAYER. 2012. Classification of Landsat images based on spectral and topographic variables for land cover change detection in Zagros forests. *Int. J. Rem. Sens.* 33:6956–6974.

HILKER, T., M.A. WULDER, AND N.C. COOPS. 2008. Update of forest inventory data with lidar and high spatial resolution satellite imagery. *Can. J. Rem. Sens.* 34:5–12.

HILL, R.A., D.S. BOYD, AND C. HOPKINSON. 2010. Integrating airborne LiDAR and Landsat ETM+ data for large area assessment of forest canopy height in Amazonia. In *Silvilaser 2010, the 10th International conference on LiDAR applications for assessing forest ecosystems*, 14–17 September 2010, Freiburg, Germany.

HIRATA, Y., N. FURUYA, M. SUZUKI, AND H. YAMAMOTO. 2009. Airborne laser scanning in forest management: Individual tree identification and laser pulse penetration in a stand with different levels of thinning. *For. Ecol. Manage.* 258:752–760.

HOOVER, C.M., AND S.A. REBAIN. 2011. *Forest carbon estimation using the Forest Vegetation Simulator: Seven things you need to know.* USDA For. Serv., Gen. Tech. Rep. NRS-77, Northern Research Station, Newtown Square, PA. 20 p.

HOUGHTON, R.A., D. BUTMAN, A.G. BUNN, O.N. KRANKINA, P. SCHLESINGER, AND T.A. STONE. 2007. Mapping Russian forest biomass with data from satellites and forest inventories. *Environ. Res. Lett.* 2:045032.

HUDAK, A.T., N.L. CROOKSTON, J.S. EVANS, M.J. FALKOWSKI, A.M.S.

SMITH, P.E. GESSLER, AND P. MORGAN. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Can. J. Rem. Sens.* 32:126–138.

HUDAK, A.T., N.L. CROOKSTON, J.S. EVANS, D.E. HALL, AND M.J. FALKOWSKI. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Rem. Sens. Environ.* 112:2232–2245.

HUDAK, A.T., J.S. EVANS, AND A.M.S. SMITH. 2009. LiDAR utility for natural resource managers. *Rem. Sens.* 1:934–951.

HUDAK, A.T., E.K. STRAND, L.A. VIERLING, J.C. BYRNE, J. EITEL, S. MARTINUZZI, AND M.J. FALKOWSKI. 2012. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Rem. Sens Environ.* 123:25–40.

HUMMEL, S., A.T. HUDAK, E.H. UEBLER, M.J. FALKOWSKI, AND K.A. MEGOWN. 2011. A comparison of accuracy and cost of LiDAR versus stand exam data for landscape management on the Malheur National Forest. *J. For.* 109:267–273.

HYYPPÄ, J., X. YU, H. HYYPPÄ, M. VASTARANTA, M. HOLOPAINEN, A. KUKKO, H. KAARTINEN, ET AL. 2012. Advances in forest inventory using airborne laser scanning. *Rem. Sens.* 4:1190–1207.

INGRAM, J.C., T.P. DAWSON, AND R.J. WHITTAKER. 2005. Mapping tropical forest structure in southeastern Madagascar using remote sensing and artificial neural networks. *Rem. Sens. Environ.* 94:491–507.

IVERSON, L.R., AND A.M. PRASAD. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecol. Monogr.* 68:465–485.

JENKINS, J.C., D.C. CHOJNACKY, L.S. HEATH, AND R.A. BIRDSEY. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49:12–35.

JENSEN, J.L.R., K.S. HUMES, T. CONNER, C.J. WILLIAMS, AND J. DE-GROOT. 2006. Estimation of biophysical characteristics for highly variable mixed-conifer stands using small-footprint lidar. *Can. J. For. Res.* 36:1129–1138.

JENSEN, J.R., F. QIU, AND M. JI. 1999. Predictive modelling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. *Int. J. Rem. Sens.* 20:2805–2822.

JI, L., AND K. GALLO. 2006. An agreement coefficient for image comparison. *Photogramm. Eng. Rem. Sens.* 72:823–833.

JUNG, J., S. KIM, S. HONG, K. KIM, E. KIM, J. IM, AND J. HEO. 2013. Effects of national forest inventory plot location error on forest carbon stock estimation using *k*-nearest neighbor algorithm. *ISPRS J. Photogramm.* 81:82–92.

KILKKI, P., AND R. PÄIVINEN. 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. *Univ. Helsinki Dept. For. Mensuration Manage. Res. Notes* 19:209–215.

KNORN, J., A. RABE, V.C. RADELOFF, T. KUEMMERLE, J. KOZAK, AND P. HOSTERT. 2009. Landcover mapping of large areas using chain classification of neighboring Landsat satellite images. *Rem. Sens. Environ.* 113:957–964.

KOHAVI, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. P. 1137–1145 in *Proc. of the 14th international joint conference on artificial intelligence.* Morgan Kaufman Publishers, San Mateo, CA.

KOZAK, A., AND R. KOZAK. 2003. Does cross validation provide additional information in the evaluation of regression models? *Can. J. For. Res.* 33:976–987.

KUPFER, J.A., A.T. MYERS, S.E. MCLANE, AND G.N. MELTON. 2008. Patterns of forest damage in a southern Mississippi landscape caused by Hurricane Katrina. *Ecosystems* 11:45–60.

LABRECQUE, S., R.A. FOURNIER, J.E. LUTHER, AND D. PIERCEY. 2006. A comparison of four methods to map biomass from Landsat-TM and inventory data in western Newfoundland. *For. Ecol. Manage.* 226:129–144.

LECKIE, D.G., F.A. GOUGEON, N. WALSWORTH, AND D. PARADINE.

2003. Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Rem. Sens. Environ.* 85:355–369.

LE MAIRE, G., C. MARSDEN, Y. NOUVELLON, C. GRINAND, R. HAKAMADA, J.L. STAPE, AND J.P. LACLAU. 2011. MODIS NDVI time-series allow the monitoring of *Eucalyptus* plantation biomass. *Rem. Sens. Environ.* 115:2613–2625.

LEDUC, D.J., T.G. MATNEY, K.L. BELLI, AND V.C. BALDWIN JR. 2001. *Predicting diameter distributions of longleaf pine plantations: A comparison between artificial neural networks and other accepted methodologies.* USDA For. Serv., Res. Paper SRS-25, Southern Research Station, Asheville, NC. 18 p.

LEFSKY, M.A., W.B. COHEN, G.G. PARKER, AND D.J. HARDING. 2002. Lidar remote sensing for ecosystem studies. *BioScience* 52:19–30.

LEFSKY, M.A., W.B. COHEN, AND T.A. SPIES. 2001. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Can. J. For. Res.* 31:78–87.

LEFSKY, M.A., A.T. HUDAK, W B. COHEN, AND S.A. ACKER. 2005a. Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. *Rem. Sens. Environ.* 95:517–531.

LEFSKY, M.A., D.P. TURNER, M. GUZY, AND W.B. COHEN. 2005b. Combining lidar estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity. *Rem. Sens. Environ.* 95:549–558.

LEITES, L.P., A.P. ROBINSON, AND N.L. CROOKSTON. 2009. Accuracy and equivalence testing of crown ratio models and assessment of their impact on diameter growth and basal area increment predictions of two variants of the Forest Vegetation Simulator. *Can. J. For. Res.* 39:655–665.

LEK, S., AND J.F. GUÉGAN. 1999. Artificial neural networks as a tool in ecological modeling, an introduction. *Ecol. Model.* 120:65–73.

LEMAY, V., J. MAEDEL, AND N.C. COOPS. 2008. Estimating stand structural details using nearest neighbor analyses to link ground data, forest cover maps, and Landsat imagery. *Rem. Sens. Environ.* 112:2578–2591.

LEMAY, V., AND H. TEMESGEN. 2005. Comparison of nearest neighbor methods for estimating basal area and stems per hectare using aerial auxiliary variables. *For. Sci.* 51:109–119.

LENDASSE, A., V. WERTZ, AND M. VERLEYSEN. 2003. Model selection with cross-validations and bootstraps—Application to time series prediction with RBFN models. P. 573–580 in *Artificial neural networks and neural information processing—ICANN/ICONIP* 2003, Kaynak, O., E. Alpaydin, E. Oja, and L. Xu (eds.). Springer-Verlag Lecture Notes in Computer Science 2714, Springer-Verlag, Berlin, Germany.

LESAK, A.A., V.C. RADELOFF, T.J. HAWBAKER, A.M. PIDGEON, T. GOBAKKEN, AND K. CONTRUCCI. 2011. Modeling forest songbird species richness using LiDAR-derived measures of forest structure. *Rem. Sens. Environ.* 115:2823–2835.

LIKNES, G.C., C.H. PERRY, AND D.M. MENEGUZZO. 2010. Assessing tree cover in agricultural landscapes using high-resolution aerial imagery. *J. Terres. Observ.* 2:38–55.

LINDBERG, E., J. HOLMGREN, K. OLOFSSON, J. WALLERMAN, AND H. OLSSON. 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. *Int. J. Rem. Sens.* 31:1175–1192.

LINDBERG, E., J. HOLMGREN, K. OLOFSSON, J. WALLERMAN, AND H. OLSSON. 2013. Estimation of tree lists from airborne laser scanning using tree model clustering and k-MSN imputation. *Rem. Sens.* 5:1932–1955.

LISTER, A. 2009. Landscape scale mapping of forest inventory data by nearest neighbor classification. P. 265–272 in *Proc. of the Eighth annual forest inventory and analysis symposium*, McRoberts, R.E., G.A. Reams, P.C. Van Deusen, and W.H. McWilliams (eds.). USDA For. Serv., Gen. Tech. Rep. WO-79, Washington, DC.

LIU, C., P. FRAZIER, AND L. KUMAR. 2007. Comparative assessment of the measures of thematic classification accuracy. *Rem. Sens. Environ.* 107:606–616.

LOH, W.Y. 2011. *User manual for GUIDE 11.1.* Department of Statistics, Univ. of Wisconsin, Madison, WI. 104 p.

MAGNUSSON, S., E. NÆSSET, AND T. GOBAKKEN. 2010. Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. *Rem. Sens. Environ.* 114:700–712.

MÄKELÄ, H., AND A. PEKKARINEN. 2004. Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data. *For. Ecol. Manage.* 196:245–255.

MALTAMO, M., J. PEUHKURINEN, J. MALINEN, J. VAUHKONEN, P. PACKALÉN, AND T. TOKOLA. 2009. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. *Silva Fenn.* 43:507–521.

MALTAMO, M., A. SUVANTO, AND P. PACKALÉN. 2007. Comparison of basal area and stem frequency diameter distribution modelling using airborne laser scanner data and calibration estimation. *For. Ecol. Manage.* 247:26–34.

MARTINUZZI, S., L.A. VIERLING, W.A. GOULD, M.J. FALKOWSKI, J.S. EVANS, A.T. HUDAK, AND K.T. VIERLING. 2009. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Rem. Sens. Environ.* 113:2533–2546.

MASELLI, F., AND M. CHIESI. 2006. Evaluation of statistical methods to estimate forest volume in a Mediterranean region. *IEEE Trans. Geosci. Rem. Sens.* 44:2239–2250.

MAURO, F., R. VALBUENA, J. MANZANERA, AND A. GARCIA-ABRIL. 2011. Influence of Global Navigation Satellite System errors in positioning inventory plots for tree-height distribution studies. *Can. J. For. Res.* 41:11–23.

MCINERNEY, D.O., J. SUAREZ-MINGUEZ, R. VALBUENA, AND M. NIEUWENHUIS. 2010. Forest canopy height retrieval using LiDAR data, medium-resolution satellite imagery and kNN estimation in Aberfoyle, Scotland. *Forestry* 83:195–206.

MCROBERTS, R.E. 2009a. Diagnostic tools for nearest neighbors techniques when used with satellite imagery. *Rem. Sens. Environ.* 113:489–499.

MCROBERTS, R.E. 2009b. A two-step nearest neighbors algorithm using satellite imagery for predicting forest structure within species composition classes. *Rem. Sens. Environ.* 113:532–545.

MCROBERTS, R.E. 2010. The effects of rectification and global positioning system errors on satellite image-based estimates of forest area. *Rem. Sens. Environ.* 114:1710–1717.

MCROBERTS, R.E. 2012. Estimating forest attribute parameters for small areas using nearest neighbors techniques. *For. Ecol. Manage.* 272:3–12.

MCROBERTS, R.E., W.B. COHEN, E. NÆSSET, S.V. STEHMAN, AND E.O. TOMPPO. 2010a. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scand. J. For. Res.* 25:340–367.

MCROBERTS, R.E., M.D. NELSON, AND D.G. WENDT. 2002. Stratified estimation of forest area using satellite imagery, inventory data, and the *k*-nearest neighbors technique. *Rem. Sens. Environ.* 82:457–468.

MCROBERTS, R.E., AND E.O. TOMPPO. 2007. Remote sensing support for national forest inventories. *Rem. Sens. Environ.* 110:412–419.

MCROBERTS, R.E., E.O. TOMPPO, AND E. NÆSSET. 2010b. Advances and emerging issues in national forest inventories. *Scand. J. For. Res.* 25:368–381.

MEANS, J.E., S.A. ACKER, D.J. HARDING, J.B. BLAIR, M.A. LEFSKY, W.B. COHEN, M.E. HARMON, AND W.A. MCKEE. 1999. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Rem. Sens. Environ.* 67:298–308.

MIAO, X., P. GONG, R. PU, R.I .CARRUTHERS, AND J.S. HEATON. 2007. Applying class-based feature extraction approaches for supervised classification of hyperspectral imagery. *Can. J. Rem. Sens.* 33:162–175.

MOEUR, M., AND A.R. STAGE. 1995. Most similar neighbor: An improved sampling inference procedure for natural resource planning. *For. Sci.* 41:337–359.

MOISEN, G.G., E.A. FREEMAN, J.A. BLACKARD, T.S. FRESCINO, N.E. ZIMMERMANN, AND T.C. EDWARDS JR. 2006. Predicting tree species presence and basal area in Utah: A comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecol. Model.* 199:176–187.

MOISEN, G.G., AND T.S. FRESCINO. 2002. Comparing five modelling techniques for predicting forest characteristics. *Ecol. Model.* 157:209–225.

MORA, B., M.A. WULDER, AND J.C. WHITE. 2010. Identifying leading species using tree crown metrics derived from very high spatial resolution imagery in a boreal forest environment. *Can. J. Rem. Sens.* 36:332–344.

MUÑOZ, J., AND Á.M. FELICÍSIMO. 2004. Comparison of statistical methods commonly used in predictive modelling. *J. Veg. Sci.* 15:285–292.

MURPHY, M.A., J.S. EVANS, AND A. STORFER. 2010. Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology* 91:252–261.

MUUKKONEN, P., AND J. HEISKANEN. 2005. Estimating biomass for boreal forests using ASTER satellite data combined with standwise forest inventory data. *Rem. Sens. Environ.* 99:434–447.

NÆSSET, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Rem. Sens. Environ.* 80:88–99.

NÆSSET, E. 2007. Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. *Scand. J. For. Res.* 22:433–442.

NÆSSET, E., O.M. BOLLANDSÅS, AND T. GOBAKKEN. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Rem. Sens. Environ.* 94:541–553.

NÆSSET, E., AND T. GOBAKKEN. 2008. Estimation of above- and belowground biomass across regions of the boreal forest zone using airborne laser. *Rem. Sens. Environ.* 112:3079–3090.

NÆSSET, E., T. GOBAKKEN, S. SOLBERG, T.G. GREGOIRE, R. NELSON, G. STÅHL, AND D. WEYDAHL. 2011. Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area. *Rem. Sens. Environ.* 115:3599–3614.

NELSON, R., T. GOBAKKEN, E. NÆSSET, T.G. GREGOIRE, G. STÅHL, S. HOLM, AND J. FLEWELLING. 2012. Lidar sampling—Using an airborne profiler to estimate forest biomass in Hedmark County, Norway. *Rem. Sens. Environ.* 123:563–578.

NISKA, H., J.-P. SKÖN, P. PACKALÉN, T. TOKOLA, M. MALTAMO, AND M. KOLEHMAINEN. 2010. Neural networks for the prediction of species-specific plot volumes using airborne laser scanning and aerial photographs. *IEEE Trans. Geosci. Rem. Sens.* 48:1076–1085.

NORD-LARSEN, T., AND J. SCHUMACHER. 2012. Estimation of forest resources from a country wide laser scanning survey and national forest inventory data. *Remote Sens. Environ.* 119:148–157.

NOTHDURFT, A., J. SABOROWSKI, AND J. BREIDENBACH. 2009. Spatial prediction of forest stand variables. *Eur. J. For. Res.* 128:241–251.

OHMANN, J.L., AND M.J. GREGORY. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Can. J. For. Res.* 32:725–741.

OHMANN, J.L., M.J. GREGORY, E.B. HENDERSON, AND H.M. ROBERTS. 2011. Mapping gradients of community composition with nearest-neighbor imputation: Extending plot data for landscape analysis. *J. Veg. Sci.* 22:660–676.

PACKALÉN, P., AND M. MALTAMO. 2007. The k-MSN method for the

prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. *Rem. Sens. Environ.* 109:328–341.

PACKALÉN, P., H. TEMESGEN, AND M. MALTAMO. 2012. Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory. *Can. J. Rem. Sens.* 38:557–569.

PAL, M., AND G.M. FOODY. 2010. Feature selection for classification of hyperspectral data by SVM. *IEEE Trans. Geosci. Rem. Sens.* 48:2297–2307.

PARKER, R.C., AND D.L. EVANS. 2004. An application of LiDAR in a double-sample forest inventory. *West. J. Appl. For.* 19:95–101.

PASCUAL, C., A. GARCÍA-ABRIL, W.B. COHEN, AND S. MARTÍN-FERNÁNDEZ. 2010. Relationship between LiDAR-derived forest canopy height and Landsat images. *Int. J. Rem. Sens.* 31:1261–1280.

PATTERSON, P.L., AND M.S. WILLIAMS. 2003. Effects of registration errors between remotely sensed and ground data on estimators of forest area. *For. Sci.* 49:110–118.

PEDERSEN, R.Ø., O.M. BOLLANDSÅS, T. GOBAKKEN, AND N. NÆSSET. 2012. Deriving individual tree competition indices from airborne laser scanning. *For. Ecol. Manage.* 280:150–165.

PENG, C., AND X. WEN. 1999. Recent applications of artificial neural networks in forest resource management: An overview. P. 15–22 in *Environmental decision support systems and artificial intelligence*, Corté, U., and M. Sànchez-Marrè (eds.). Tech. Rep. WS-99-07, AAAI Press, Menlo Park, CA.

PESONEN, A., M. MALTAMO, K. EERIKÄINEN, AND P. PACKALÉN. 2008. Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *For. Ecol. Manage.* 255:3288–3296.

PHILIPS, R.D., C.E. BLINN, L.T. WATSON, AND R.H. WYNNE. 2009. An adaptive noise-filtering algorithm for AVIRIS data with implications for classification accuracy. *IEEE Trans. Geosci. Rem. Sens.* 47:3168–3179.

PIERCE, K.B. JR., J.L. OHMANN, M.C. WIMBERLY, M.J. GREGORY, AND J.S. FRIED. 2009. Mapping wildland fuels and forest structure for land management: A comparison of nearest neighbor imputation and other methods. *Can. J. For. Res.* 39:1901–1916.

PIRAMUTHU, S. 2008. Input data for decision trees. *Expert Syst. Appl.* 34:1220–1226.

POCEWICZ, A.L., P. GESSLER, AND A.P. ROBINSON. 2004. The relationship between effective plant area index and Landsat spectral response across elevation, solar insolation, and spatial scales in a northern Idaho forest. *Can. J. For. Res.* 34:465–480.

POND, N.C., R.E. FROESE, AND L.M. NAGEL. 2014. Sustainability of the selection system in northern hardwood forests. *For. Sci.* doi:10.5849/forsci.12-113.

POPESCU, S.C., R.H. WYNNE, AND R.F. NELSON. 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Can. J. For. Res.* 29:564–577.

POWELL, S.L., W.B. COHEN, S.P. HEALEY, R.E. KENNEDY, G.G. MOISEN, K.B. PIERCE, AND J.L. OHMANN. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Rem. Sens. Environ.* 114:1053–1068.

POWELL, S.L., W.B. COHEN, AND R.E. KENNEDY. 2007. Analysis of biomass dynamics derived from spectral trajectories. In *Proc. of the Eleventh Forest Service remote sensing applications conference*, American Society for Photogrammetry and Remote Sensing, Bethesda, MD.

PRASAD, A.M., L.R. IVERSON, AND A. LIAW. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* 9:181–199.

PRISLEY, S.P., H.-J. WANG, P.J. RADTKE, AND J. COULSTON. 2009. Combining FIA plot data with topographic variables: Are precise locations needed? In *Proc. of the Forest Inventory and Analysis symposium 2008*, McWilliams, W., G. Moisen, and R. Czaplewski (eds.). USDA For.

Serv., Proc. RMRS-P-56CD, Rocky Mountain Research Station, Park City, UT. 839 p.

PU, R., AND P. GONG. 2004. Wavelet transform applied to EO-1 hyperspectral data for forest LAI and crown closure mapping. *Rem. Sens. Environ.* 91:212–224.

REESE, H., M. NILSSON, P. SANDSTRÖM, AND H. OLSSON. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Comput. Electron. Agr.* 37:37–55.

RICHARDSON, J.J., AND L.M. MOSKAL. 2011. Strengths and limitations of assessing forest density and spatial configuration with aerial LiDAR. *Rem. Sens. Environ.* 115:2640–2651.

RIEMANN, R., B.T. WILSON, A. LISTER, AND S. PARKS. 2010. An effective assessment protocol for continuous geospatial datasets of forest characteristics using USFS Forest Inventory and Analysis (FIA) data. *Rem. Sens. Environ.* 114:2337–2352.

ROBINSON, A.P., R.A. DUURSMA, AND J.D. MARSHALL. 2005. A regression-based equivalence test for model validation: Shifting the burden of proof. *Tree Physiol.* 25:903–913.

ROBINSON, A.P., AND A.R. EK. 2000. The consequences of hierarchy for modeling in forest ecosystems. *Can. J. For. Res.* 30:1837–1846.

ROBINSON, A.P., AND R.E. FROESE. 2004. Model validation using equivalence tests. *Ecol. Model.* 176:349–358.

ROGAN, J., J. FRANKLIN, D. STOW, J. MILLER, C. WOODCOCK, AND D. ROBERTS. 2008. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Rem. Sens. Environ.* 112:2272–2283.

RUEFENACHT, B., M.V. FINCO, M.D. NELSON, R. CZAPLEWSKI, E.H. HELMER, J.A. BLACKARD, G.R. HOLDEN, ET AL. 2008. Conterminous US and Alaska forest type mapping using Forest Inventory and Analysis data. *Photogramm. Eng. Rem. Sens.* 74:1379–1388.

RYKIEL, E.J. JR. 1996. Testing ecological models: The meaning of validation. *Ecol. Model.* 90:229–244.

SAATCHI, S., K. HALLIGAN, D.G. DESPAIN, AND R.L. CRABTREE. 2007. Estimation of forest fuel load from radar remote sensing. *IEEE Trans. Geosci. Rem. Sens.* 45:1726–1740.

SCHROEDER, T.A., W.B. COHEN, AND Z. YANG. 2007. Patterns of forest regrowth following clearcutting in western Oregon as determined from a Landsat time-series. *For. Ecol. Manage.* 243:259–273.

SERPICO, S.B., L. BRUZZONE, AND F. ROLI. 1996. An experimental comparison of neural and statistical non-parametric algorithms for supervised classification of remote-sensing images. *Pattern Recogn. Lett.* 17:1331–1341.

SESNIE, S.E., B. FINEGAN, P.E. GESSLER, S. THESSLER, Z.R. BENDANA, AND A.M.S. SMITH. 2010. The multispectral separability of Costa Rican rainforest types with support vector machines and random forest decision trees. *Int. J. Rem. Sens.* 31:2885–2909.

SESNIE, S.E., P.E. GESSLER, B. FINEGAN, AND S. THESSLER. 2008. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Rem. Sens. Environ.* 112:2145–2159.

SOKAL, R.R., AND F.J. ROHLF. 2012. *Biometry*, 4th ed. W.H. Freeman, New York. 937 p.

SOLBERG, S., E. NÆSSET, H. LANGE, AND O.M. BOLLANDSÅS. 2004. Remote sensing of forest health. *ISPRS Arch.* XXXVI-8/W2:161–166.

SOTOCA, J.M., F. PLA, AND J.S. SÁNCHEZ. 2007. Band selection in multispectral images by minimization of dependent information. *IEEE Trans. Syst. Man Cybernet. C* 37:258–267.

STAGE, A.R., AND N.L. CROOKSTON. 2007. Partitioning error components for accuracy-assessment of near-neighbor methods of imputation. *For. Sci.* 53:62–72.

STEHMAN, S.V. 2004. A critical evaluation of the normalized error matrix in map accuracy assessment. *Photogramm. Eng. Rem. Sens.* 70:743–751.

STEHMAN, S.V., AND G.M. FOODY. 2009. Accuracy assessment. P. 297–309 in *The SAGE handbook of remote sensing*, Warner, T.A.,

M.D. Nellis, and G.M. Foody (eds.). SAGE Publications Ltd, London, UK.

Stephens, P.R., M.O. Kimberley, P.N. Beets, T.S.H. Paul, N. Searles, A. Bell, C. Brack, and J. Broadley. 2012. Airborne scanning LiDAR in a double sampling forest carbon inventory. *Rem. Sens. Environ.* 117:348–357.

Strobl, C., A.L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinform.* 9:307.

Stumpf, A., and N. Kerle. 2011. Object-oriented mapping of landslides using Random Forests. *Rem. Sens. Environ.* 115:2564–2577.

Svetnik, V., A. Liaw, C. Tong, J.C. Culberson, R.P. Sheridan, and B.P. Feuston. 2003. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comp. Sci.* 43:1947–1958.

Teissier du Cros, R., and C. Vidal. 2009. Mapping Martinique's forests and other natural lands for land planning and development. P. 31–38 in *Proc. of the Eighth annual Forest Inventory and Analysis symposium*, McRoberts, R.E., G.A. Reams, P.C. Van Deusen, and W.H. McWilliams (eds.). USDA For. Serv., Gen. Tech. Rep. WO-79, Washington, DC.

Temesgen, H., V.M. LeMay, K.L. Froese, and P.L. Marshall. 2003. Imputing tree-lists from aerial attributes for complex stands of southeastern British Columbia. *For. Ecol. Manage.* 177:277–285.

Thenkabail, P.S., E.A. Enclona, M.S. Ashton, C. Legg, and M.J. De Dieu. 2004. Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. *Rem. Sens. Environ.* 90:23–43.

Thomlinson, J.R., P.V. Bolstad, and W.B. Cohen. 1999. Coordinating methodologies for scaling landcover classifications from site-specific to global: Steps toward validating global map products. *Rem. Sens. Environ.* 70:16–28.

Tickle, P.K., A. Lee, R.M. Lucas, J. Austin, and C. Witte. 2006. Quantifying Australian forest floristics and structure using small footprint LiDAR and large scale aerial photography. *For. Ecol. Manage.* 223:379–394.

Tokola, T., J. Pitkänen, S. Partinen, and E. Muinonen. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *Int. J. Rem. Sens.* 17:2333–2351.

Tomppo, E. 1991. Satellite imagery-based national inventory of Finland. *Int. Arch, Photogramm. Rem. Sens.* 28:419–424.

Tomppo, E.O., C. Gagliano, F. De Natale, M. Katila, and R.E. McRoberts. 2009. Predicting categorical forest variables using an improved k-nearest neighbour estimator and Landsat imagery. *Rem. Sens. Environ.* 113:500–517.

Tomppo, E., C. Goulding, and M. Katila. 1999. Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. *Scand. J. For. Res.* 14:182–192.

Tomppo, E., and M. Halme. 2004. Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: A genetic algorithm approach. *Rem. Sens. Environ.* 92:1–20.

Tomppo, E., H. Olsson, G. Ståhl, M. Nilsson, O. Hagner, and M. Katila. 2008. Combining national forest inventory field plots and remote sensing data for forest databases. *Rem. Sens. Environ.* 112:1982–1999.

Tonolli, S., M. Dalponte, M. Neteler, M. Rodeghiero, L. Vescovo, and D. Gianelle. 2011. Fusion of airborne LiDAR and satellite multispectral data for the estimation of timber volume in the Southern Alps. *Rem. Sens. Environ.* 115:2486–2498.

Townsend, P.A. 2000. A quantitative fuzzy approach to assess mapped vegetation classifications for ecological applications. *Rem. Sens. Environ.* 72:253–267.

Townsend, P.A., J.R. Foster, R.A. Chastain Jr., and W.S. Currie. 2003. Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS. *IEEE Trans. Geosci. Rem. Sens.* 41:1347–1354.

Tuominen, S., K. Eerikäinen, A. Schibalski, M. Haakana, and A. Lehtonen. 2010. Mapping biomass variables with a multi-source forest inventory technique. *Silva Fenn.* 44:109–119.

Tuominen, S., S. Fish, and S. Poso. 2003. Combining remote sensing, data from earlier inventories, and geostatistical interpolation in multi-source forest inventory. *Can. J. For. Res.* 33:624–634.

Turner, D.P., W.B. Cohen, R.E. Kennedy, K.S. Fassnacht, and J.M. Briggs. 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Rem. Sens. Environ.* 70:52–68.

Turner, W., S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger. 2003. Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.* 18:306–314.

Urban, D.L. 2002. Classification and regression trees. P. 222–232 in *Analysis of ecological communities*, McCune, B., and J.B. Grace (eds.). MjM Sofware Design, Gleneden Beach, OR.

Valbuena, R., M. Maltamo, S. Martín Fernández, P. Packalén, C. Pascual, and G.J. Nabuurs. 2013. Patterns of covariance between airborne laser scanning metrics and Lorenz curve descriptors of tree size inequality. *Can. J. Rem. Sens.* doi:10.5589/m13-012.

Valbuena, R., F. Mauro, F.J. Arjonilla, and J.A. Manzanera. 2011. Comparing airborne laser scanning-imagery fusion methods based on geometric accuracy in forested areas. *Rem. Sens. Environ.* 1158:1942–1954.

van Aardt, J.A.N., R.H. Wynne, and J.A. Scrivani. 2008. Lidar-based mapping of forest volume and biomass by taxonomic group using structurally homogenous segments. *Photogramm. Eng. Rem. Sens.* 74:1033–1044.

Vasconcelos, M.J.P., S. Silva, M. Tomé, M. Alvim, and J.M.C. Pereira. 2001. Spatial prediction of fire ignition probabilities: Comparing logistic regression and neural networks. *Photogramm. Eng. Rem. Sens.* 67:73–81.

Vastaranta, M., V. Kankare, M. Holopainen, X. Yu, J. Hyyppä, and H. Hyyppä. 2012. Combination of individual tree detection and area-based approach in imputation of forest variables using airborne laser data. *ISPRS J. Photogramm.* 67:73–79.

Vauhkonen, J., I. Korpela, M. Maltamo, and T. Tokola. 2010. Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sens. Environ.* 114:1263–1276.

Wallerman, J., and J. Holmgren. 2007. Estimating field-plot data of forest stands using airborne laser scanning and SPOT HRG data. *Rem. Sens. Environ.* 110:501–508.

Watt, P.J., D.N.M. Donoghue, K.B. McManus, and R.W. Dunford. 2004. Predicting forest height from Ikonos, Landsat and LiDAR imagery. *ISPRS Arch.* XXXVI-8/W2:228–231.

Wilson, B.T., A.J. Lister, and R.I. Riemann. 2012. A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *For. Ecol. Manage.* 271:182–198.

Wolter, P.T., D.J. Mladenoff, G.E. Host, and T.R. Crow. 1995. Improved forest classification in the northern Lake States using multi-temporal Landsat imagery. *Photogramm. Eng. Rem. Sens.* 61:1129–1144.

Wolter, P.T., P.A. Townsend, and B.R. Sturtevant. 2009. Estimation of forest structural parameters using 5 and 10 meter SPOT-5 satellite data. *Rem. Sens. Environ.* 113:2019–2036.

Woods, M., D. Pitt, M. Penner, K. Lim, D. Nesbitt, D. Etheridge, and P. Treitz. 2011. Operational implementation of a LiDAR inventory in boreal Ontario. *For. Chron.* 87:512–528.

Wulder, M.A., C.W. Bater, N.C. Coops, T. Hilker, and J.C. White. 2008. The role of LiDAR in sustainable forest management. *For. Chron.* 84:807–826.

WULDER, M.A., R.J. HALL, N.C. COOPS, AND S.E. FRANKLIN. 2004. High spatial resolution remotely sensed data for ecosystem characterization. *BioScience* 54:511–521.

WULDER, M.A., J.C. WHITE, R.F. NELSON, E. NÆSSET, H.O. ØRKA, N.C. COOPS, T. HILKER, C.W. BATER, AND T. GOBAKKEN. 2012. Lidar sampling for large-area forest characterization: A review. *Rem. Sens. Environ.* 121:196–209.

XIAO, X., S. BOLES, J. LIU, D. ZHUANG, AND M. LIU. 2002. Characterization of forest types in northeastern China, using multi-temporal SPOT-4 VEG-ETATION sensor data. *Rem. Sens. Environ.* 82:335–348.

YOUNG, T.M., Y. WANG, D.G. HODGES, AND F.M. GUESS. 2009. Decision tree applications for forestry and forest products manufacturers. P. 104–115 in *Proc. of the 2008 Southern Forest Economics Workers annual meeting*, Siry, J., P. Bettinger, T. Harris, T. Tye, S. Baldwin, and K. Merry (eds.). Center for Forest Business Publ. No. 30, Univ. of Georgia, Athens, GA.

YU, X., J. HYYPPÄ, M. VASTARANTA, M. HOLOPAINEN, AND R. VIITALA.
2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS J. Photogramm.* 66:28–37.

ZHANG, J., B. RIVARD, A. SÁNCHEZ-AZOFEIFA, AND K. CASTRO-ESAU. 2006. Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HY-DICE imagery. *Rem. Sens. Environ.* 105:129–141.

ZHANG, X., AND S. KONDRAGUNTA. 2006. Estimating forest biomass in the USA using generalized allometric models and MODIS land products. *Geophys. Res. Lett.* 33:L09402.

ZHAO, K., S. POPESCU, AND R. NELSON. 2009. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Rem. Sens. Environ.* 113:182–196.

ZHENG, D., J. RADEMACHER, J. CHEN, T. CROW, M. BRESEE, J. LE MOINE, AND S.R. RYU. 2004. Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. *Rem. Sens. Environ.* 93:402–411.